MASARYK UNIVERSITY

FACULTY OF SCIENCE

DEPARTMENT OF EXPERIMENTAL BIOLOGY

ENGINEERING BACTERIA, THEIR ENZYMES, AND METABOLIC PATHWAYS FOR BIOTECHNOLOGICAL PROCESSING OF WASTE COMPOUNDS

HABILITATION THESIS

PAVEL DVOŘÁK

BRNO 2022

ABSTRACT

This habilitation thesis is a commented collection of my 15 peer-reviewed impacted publications that contributed to the fields of bioengineering and biotechnology. Biotechnology and particularly its environmental and industrial branches can take great advantage of recent advances in microbial bioengineering. Many environmental processes including biodegradation of polluting chemicals or mineralization of waste organic matter largely depend on the catalytic activities of microorganisms and especially of bacteria that are endowed with rapid growth and richness of metabolic pathways. This natural potential of bacterial catalysts can be further expanded behind the foreseen horizon using the new tools and approaches of protein engineering, metabolic engineering, and synthetic biology – the disciplines that together form three major pillars of modern microbial bioengineering. The thesis discusses the impact of these disciplines on the area of biotechnology with increasing significance – bioprocessing of waste compounds. The first part of the introductory text describes the adoption of engineering principles in the knowledge-aided designs of better molecular and whole-cell bacterial catalysts for the biodegradation and potential valorization of anthropogenic waste chemicals that are harmful to the environment. Special attention is dedicated to halogenated hydrocarbons and namely to the synthetic toxic chemical 1,2,3-trichloropropane. The second part highlights the role of microbial bioengineering in the preparation of enhanced bacterial cell factories that can valorize substrates from waste plant biomass in environment-friendly bioprocesses. Our recent attempts to upgrade soil bacterium *Pseudomonas putida* for this purpose are featured as an example.

ACKNOWLEDGEMENT

First of all, I want to thank my wife Monika and my daughters Zita and Zoja for their support, patience, and for giving me a home where I love to return to after work.

My big thanks go to my former supervisors and mentors Jiří Damborský, Zbyněk Prokop, Víctor de Lorenzo Prieto, and Pablo I. Nikel from whom I learnt the most and who showed me the beauties of science but explained me also its pitfalls.

I also want to thank my former and current colleagues, friends, and collaborators especially from Loschmidt Laboratories (Masaryk University), Molecular Environmental Microbiology Laboratory (CNB-CSIC Madrid), and Section of Microbiology of the Department of Experimental Biology (Masaryk University) for numerous priceless advice and enjoyable moments we experienced in and out of the laboratory walls.

Last but not least, I want to express my gratitude for the work with all those brilliant young people – graduate and undergraduate students - whose enthusiasm and interest is the fuel of my days.

ABBREVIATIONS

Acetyl-CoA, Acetyl coenzyme A

ATP, Adenosine triphosphate

CBP, Consolidated bioprocessing

CCR, Carbon catabolite repression

CLEAs, Cross-linked enzyme aggregates

DCP, 2,3-Dichloropropan-1-ol

DDT, Dichlorodiphenyltrichloroethane

DBTA cycle, Design-Build-Test-Analyze cycle

DNA, Deoxyribonucleic acid

ELISA, Enzyme-linked immunosorbent assays

E-value, Enantioselectivity expressed as enantiomeric ratio

GC-MS, Gas chromatography-mass spectrometry

GFP and CFP, Green and cyan fluorescent protein, respectively

His, Histidine

HLDs, Haloalkane dehalogenases

HPLC, High-performance liquid chromatography

IPTG, Isopropyl β-D-1-thiogalactopyranoside

kDa, Kilodaltons (molecular mass unit)

LC, Lignocellulose

mcl-PHA, Medium chain length polyhydroxyalkanoates

MHETase, A hydrolase that cleaves mono-terephthalic acid, a product of polyethylene terephthalate (PET) degradation

PETase, An esterase that catalyzes the hydrolysis of PET

PPP, Pentose phosphate pathway

PVA, Polyvinyl alcohol

RNA, Ribonucleic acid

ROS, Reactive oxygen species

SDS-PAGE, Sodium dodecyl sulphate–polyacrylamide gel electrophoresis

TCA cycle, Tricarboxylic acid cycle

TCP, 1,2,3-Trichloropropane

WT, Wild type

CONTENT

1. Structure of the thesis and my contribution to the presented articles……1

2. Introduction……5

3. Section I: Engineering bacterial enzymes and biochemical pathways for biodegradation of anthropogenic halogenated waste compounds……7

3.1 Halogenated hydrocarbons and dehalogenating enzymes……11

3.2 1,2,3-trichloropropane and the means for its removal……12

3.3 Synthetic biochemical pathway for biodegradation and potential valorization of TCP2……14

3.4 Improvement of the TCP pathway by protein engineering……15

3.5 Improvement of the TCP pathway by metabolic engineering and synthetic biology……18

3.6 *In vitro* and *in silico* reconstruction of the TCP pathway……19

3.7 Assembly and optimization of TCP pathway in heterologous bacterial host *Escherichia coli*……21

3.8 Metabolic burden and metabolite toxicity as key challenges for the engineering of biodegradation pathways *in vivo*……24

3.9 Immobilization of TCP pathway as an alternative strategy for biotechnological processing of toxic industrial waste……27

3.10 Perspectives of biotransformation of TCP……30

4. Section II: Engineering bacteria for utilization and valorization of substrates from lignocellulosic residues……33

4.1 *Pseudomonas putida* is an attractive bacterial host for LC biotechnology……35

4.2 Expanding the substrate scope of *P. putida* towards D-xylose and D-cellobiose……38

4.3 Valorization of lignocellulosic substrates by engineered *P. putida*……41

4.4 Empowering *P. putida* with surface-displayed designer protein scaffolds……43

4.5 Perspectives of using *P. putida* for the valorization of LC residues……45

5. Conclusion……48

6. References……50

7. Articles……58

1. STRUCTURE OF THE THESIS AND MY CONTRIBUTION TO THE PRESENTED ARTICLES

By the time of writing this thesis, I have published 16 articles (10 as the first or corresponding author) in impacted international journals and three patents. This thesis is a commented collection of 15 peer-reviewed articles published between years 2009 and 2021 in collaboration with my colleagues from Loschmidt Laboratories and Microbial Bioengineering Laboratory at Masaryk University in Brno, Czech Republic, and Molecular Environmental Microbiology Laboratory at Spanish National Centre for Biotechnology (CNB-CSIC) in Madrid, Spain, where I worked in this time interval.

The papers are sorted into two major sections. The first section (papers 1 – 10) is dedicated to the study and engineering of molecular and bacterial biocatalysts for biodegradation and potential valorization of waste halogenated hydrocarbons. Recalcitrant anthropogenic compound 1,2,3-trichloropropane is used as a model example. This collection also includes two review articles on biotechnological applications of haloalkane dehalogenases (Koudelakova et al., 2013) and on the application of modern bioengineering disciplines for the preparation of pollutant-removing bacteria (Dvořák et al., 2017) that were very well accepted by the research community and gathered numerous citations. The second section (papers 11 – 15) encompasses studies in which metabolic engineering and synthetic biology approaches were adopted to upgrade robust environmental bacterium *Pseudomonas putida* for biotechnological processing of diverse lignocellulosic substrates. The introductory text is divided correspondingly. My contribution to the 15 selected studies in terms of experimental work, supervision of students, manuscript preparation, and research direction is summarized in the tables below. Eight key first-author and corresponding-author research articles that significantly contributed to the studied field are underlined. The papers are listed below in the order that corresponds to their citation in the introductory text. In the section References, these 15 articles are highlighted bold for better orientation.

[1] Dvořák, P., Nikel, P.I., Damborský, J., and de Lorenzo, V. (2017) Bioremediation 3.0: Engineering pollutant-removing bacteria in the times of systemic biology. *Biotechnology Advances*. *35*, 845-866. (2017 IF = 12.451)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
-	-	85	-

[2] Koudelakova, T., Bidmanova, T., **Dvorak, P.**, Pavelka, A., Chaloupkova, R., Prokop, Z., and Damborsky, J. (2013) Haloalkane dehalogenases: Biotechnological applications. *Biotechnology Journal*. *8*, 32-45. (2013 IF = 3.237, most cited article in Biotechnology Journal in 2014).

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
-	-	15	-

[3] Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., **Dvorak, P.**, Prokop, Z., Stsiapanava, A., Kuty, M., Kuta-Smatanova, I., Dohnalek, J., Kulhanek, P., Wade, R.C., and Damborsky, J. (2009) Pathways and mechanisms for product release in the engineered haloalkane dehalogenases explored using classical and random acceleration molecular dynamics simulations. *Journal of Molecular Biology*. *392*, 1339-1356. (2009 IF = 4.031)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
10	-	10	-

[4] Dvorak, P., Kurumbang, N.P., Bendl, J., Brezovsky, J., Prokop, Z., and Damborsky, J. (2014) Maximizing the efficiency of multi-enzyme processes by stoichiometry optimization. *ChemBioChem*. *15*, 1891-1895. (2014 IF = 3.157)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
70	25	90	50

[5] Kurumbang, N.P.*, **Dvorak, P.***, Bendl, J., Brezovsky, J., Prokop, Z., and Damborsky, J. (2014) Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. *ACS Synthetic Biology*. *3*, 172-181. (*shared first author, 2014 IF = 4.433)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
50	50	50	50

[6] Dvorak, P., Chrast, L., Nikel. P.I., Fedr, R., Soucek, K., Chaloupkova, R., de Lorenzo, V., Prokop, Z., and Damborsky, J. (2015) Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Microbial Cell Factories*. 14, 201. (2015 IF = 4.151, among the most influential articles in MCF in 2015 and 2016 based on Altmetric.com)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
65	50	90	50

[7] Demko, M., Chrást, L., **Dvořák, P.**, Damborský, J., and Šafránek, D. (2019) Computational modelling of metabolic burden and substrate toxicity in *Escherichia coli* carrying a synthetic metabolic pathway. *Microorganisms*. *7*, 553. (2019 IF = 3.864)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
10	20	10	15

[8] Dvorak, P., Bidmanova, S., Prokop, Z., and Damborsky, J. (2014) Immobilized synthetic pathway for biodegradation of toxic recalcitrant pollutant 1,2,3-trichloropropane. *Environmental Science and Technology*. *48*, 6859–6866. (2014 IF = 5.478, ACS Editors' Choice May 24th 2014, the best Technology article out of 1,500 published papers in 2014 in Environmental Science and Technology)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
95	10	90	50

[9] Brezovsky, J., Babkova, P., Degtjarik, O., Fortova, A., Gora, A., Iermak, I., Rezacova, P., **Dvorak, P.**, Kuta Smatanova, I., Prokop, Z., Chaloupkova, R., and Damborsky, J. (2016) Engineering a de novo transport tunnel. *ACS Catalysis*. *6*, 7597-7610. (2016 IF = 10.720)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
10	-	5	-

[10] Vanacek, P., Sebestova, E., Babkova, P., Bidmanova, S., Daniel, L., **Dvorak, P.**, Stepankova, V., Chaloupkova, R., Brezovsky, J., Prokop, Z., and Damborsky, J. (2018) Exploration of enzyme diversity by integrating bioinformatics with expression analysis and biochemical characterization. *ACS Catalysis*. *8*, 2402–2412. (2018 IF = 12.025)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
10	5	5	5

[11] Dvořák, P. and de Lorenzo, V. (2018) Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. *Metabolic Engineering*. *48*, 94–108. (2020 IF = 8.115)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
100	-	90	90

[12] Dvořák, P.*, Kováč, J., and de Lorenzo, V. (2020) Biotransformation of D-xylose to D-xylonate coupled to medium-chain-length polyhydroxyalkanoate production in cellobiose-grown *Pseudomonas putida* EM42. *Microbial Biotechnology*. *13*, 1273-83. (*corresponding author, 2020 IF = 4.181)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
85	100	100	90

[13] Dvořák, P.*, Bayer, E.A., and de Lorenzo, V. (2020) Surface display of designer protein scaffolds on genome-reduced strains of *Pseudomonas putida*. *ACS Synthetic Biology*. *9*, 2749–64. (*co-corresponding author, 2020 IF = 5.229)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
100	-	90	90

[14] Espeso, D.R., **Dvořák, P.**, Aparicio, T., and de Lorenzo, V. (2020) An automated DIY framework for experimental evolution of *Pseudomonas putida*. *Microbial Biotechnology*. *14,* 2679-2685. (2020 IF = 4.181)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
30	10%	20	20

[15] Dvořák, P., Alvarez-Carreño, C., Ciordia, S., Paradela, A., and de Lorenzo, V. (2021) An updated structural model of the A domain of the *Pseudomonas putida* XylR regulator poses an atypical interplay with aromatic effectors. *Environmental Microbiology*. *23*, 4418-4433. (2021 IF = 5.491)

Experimental work (%)	Supervision (%)	Manuscript (%)	Research direction (%)
50	-	90	40

2. INTRODUCTION

With about 13 % of total carbonaceous biomass, bacteria represent the second most abundant form of life on Earth after plants.[1] They can be found in all ecosystems - marine, terrestrial, and aerial – where they contribute to numerous essential environmental processes including biomass decomposition and mineralization, biogeochemical cycling, soil fertilization, digestion and fermentation, or biodegradation of organic pollutants. Bacteria can catalyze all these processes thanks to the immense wealth of their catabolic and anabolic biochemical pathways. For instance, currently available metabolic models of a simple single *Escherichia coli* cell include over 2,700 biochemical reactions (http://bigg.ucsd.edu/). The great biocatalytic potential of bacteria and other microorganisms has been tamed and exploited by humans for thousands of years in traditional biotechnologies. But only recently, with the rise of modern analytical and bioengineering disciplines, humanity can make the most of the existing microbial biodiversity and even expand its biocatalytic capacity towards completely synthetic new-to-nature reactions.

Bacterial enzymes, biochemical routes, and whole-cell catalysts modified and optimized for desired tasks by the means of protein engineering, metabolic engineering, and synthetic biology find their use in diverse biotechnological processes. Modern bioengineering technologies are being recognized as essential weapons for coping with some of the global challenges of the 21st century not only by scientists but also by policy makers and the business community (www.weforum.org/agenda/2021/10/top-emerging-technologies-10-years/). One pronounced challenge is the accumulation of waste – gaseous, solid, and liquid - generated in unprecedented amounts by the growing population of men. The unregulated waste release results in well-known global consequences including climate change, plastic pollution, or contamination of soils and waters with toxic anthropogenic chemicals. Fast-growing and rapidly adapting bacteria are natural agents that help the ecosystems to cope with anthropogenic organic pollution. Degradation of the large portion of the massive oil spill from British Petrol Deepwater Horizon in the Gulf of Mexico in 2010 by marine bacterial communities is one of the examples that emphasize the remarkable self-healing potential of ecosystems endowed with broad microbial biodiversity.[2] However, a considerable part of man-made organic waste is recalcitrant (resistant to biological degradation) and cannot be processed easily

neither by natural means nor by biotechnologies. Examples of recalcitrant organic compounds include but are not restricted to some pesticides (e.g., atrazine, pentachlorophenol), industrial solvents and dielectric fluids (e.g., dichloroethane, polychlorinated biphenyls), warfare agents (yperite or defoliant Agent Orange made of 2,4,5-trichlorophenoxyacetic acid and 2,4-dichlorophenoxyacetic acid), or petroleum-based plastics. The majority of these chemicals have been introduced into the environment relatively recently and even bacteria have not had enough time to evolve efficient enzymatic complements for their complete mineralization. Lignocellulosic residues (a non-eatable portion of waste plant matter from agriculture, forestry, food industry, or from municipal wastes) form a special category of organic waste that is difficult to degrade and utilize biologically due to its structure and complexity.

With the increasing pressure on economic transformation towards circular economy and sustainable development (https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_en), lignocellulosic residues, dumped petroplastics, atmospheric CO_2, certain halogenated hydrocarbons, and other abundant types of organic waste attract attention not only as troublesome pollutants that need to be removed but also as potential new resources of carbon and energy for bioproduction of valuable chemicals – biofuels, biopolymers, and biopharmaceuticals.[3–5] Environmental bacteria and their enzymes used in such biotechnological conversions are sometimes limited by insufficient substrate scope, poor productivity, and low robustness (bacteria) or by suboptimal activity, selectivity, or stability (enzymes), The encountered bottlenecks can be mitigated or completely removed by modern bioengineering strategies that enable the preparation of more efficient biocatalysts for the degradation and valorization of diverse groups of organic waste chemicals.

This thesis summarizes my contribution to the study and engineering of molecular and whole-cell bacterial biocatalysts for biodegradation of and value-adding to two types of organic anthropogenic waste compounds – halogenated hydrocarbons and lignocellulosic residues.

3. SECTION I: ENGINEERING BACTERIAL ENZYMES AND BIOCHEMICAL PATHWAYS FOR BIODEGRADATION OF ANTHROPOGENIC HALOGENATED WASTE COMPOUNDS

The last decades have witnessed an unprecedented release of anthropogenic chemicals into the environment. The majority of xenobiotics of anthropogenic origin first appeared in natural ecosystems with the boom of the chemical industry in the mid of 20th century (some xenobiotics such as explosive 2,4,6-trinitrotoluene appeared on the scene even earlier with the military activities). New production technologies, organic chemicals, and materials with "magical" properties improved the life quality and comfort of millions and saved other millions from starvation. Oil-based plastics revolutionized transportation, dressing, or packaging. Synthetic drugs could fight serious diseases including cancer. Halogenated chemicals skyrocketed progress in organic chemistry and enabled the Green Revolution (or the Third Agricultural Revolution) in the 1950s and 1960s. However, what initially appeared as a pure blessing soon turned into a potential thread. Published in 1962, Rachel Carson's book Silent Spring first alerted to the devastating effects of synthetic pesticides in complex natural ecosystems. The publicity of the book and the related environmental movement resulted in the ban on dichlorodiphenyltrichloroethane known as DDT use for agricultural purposes in the United States and in the foundation of the U.S. Environmental Protection Agency. Many infamous cases of other troublesome synthetic chemicals followed. From the Thalidomide drug responsible for serious birth defects, polychlorinated biphenyls causing environmental toxicity, to the recently disclosed phenomenal problem of global microplastics pollution.

One of the major problems of many xenobiotics is their low biodegradability and accumulation in the environment. Under normal circumstances, microorganisms and particularly bacteria are the key players in natural biotransformations of organic compounds in aerobic and anaerobic conditions.[6] Their vast biodegradation potential has been exploited by humans since 19th century to which we date the first wastewater treatment plants.[7] However, some xenobiotics are hard nuts to crack even for bacteria. The reason is probably a lack of an efficient biodegradation pathway for a given chemical due to the insufficient time for its evolution, or unsuccessful evolution which ended in a deadlock, or high toxicity of a persistent compound and its degradation intermediates,[8]

Where natural means were failing, new technologies could bring a promise for the solution of the recalcitrance problem.

The rise of recombinant DNA technology and genetic engineering in the 1970s and 1980s, which was enabled by multiple new discoveries in the fields of molecular biology and biochemistry, seemed to provide such a solution. New genetic engineering tools were developed and first tested in model microorganisms, particularly in enterobacterium *Escherichia coli*. It was thus logical that bacteria and their faulty enzymes and biochemical routes for the catabolism of organic pollutants became the early targets of genetic engineers. Assembly of a complete biodegradation pathway in a suitable bacterial host by combining genes from several natural sources (so-called *patchwork assembly* approach) was believed to provide superbugs capable of mineralization of recalcitrant chemicals such as polychlorinated biphenyls or chlorotoluenes.[9,10] Despite certain success in the assembly of the synthetic biochemical traits in model and environmental bacteria, these initial attempts failed to generate vital degraders of xenobiotics. The major cause was the lack of insight into the properties and behavior of the constructed pathways implanted into the context of the surrogate host cell. This information deficit has been substantially narrowed in the last two decades of rapid development in systemic fields of metabolic engineering, protein engineering, synthetic and systems biology.

Metabolic engineering can be defined as the application of recombinant DNA methods to restructure metabolic networks for improved production or utilization of a certain substance.[11] Nowadays metabolic engineering often takes advantage of systems biology which aims to characterize and predict cell behavior by combining computational methods with high-throughput omics technologies – genomics, transcriptomics, proteomics, metabolomics, fluxomics, etc. This synergy gave rise to multi-disciplinary systems metabolic engineering which works with much more input information than traditional metabolic engineering.[12] Moreover, metabolic engineering also adopts tools and approaches of synthetic biology whose main focus is the understanding and taming of biological system complexity through the design of synthetic genetic circuits and devices using synthetic DNA technology.[13] It recruits work hierarchy, standards, and principles such as the Design-Build-Test-Analyze (DBTA) cycle known from civil engineering and electrical engineering for the designs of living systems. Last but not least, modern metabolic engineering often employs enzyme variants with activity, selectivity,

or stability improved by protein engineering. Proteins are modified either by rational design, that takes advantage of mutations targeted into the specific sites of protein structure defined by a computational design, or by directed evolution, that exploits random mutagenesis and massive screening or selection protocols.[14]

The overall impact of the aforementioned engineering disciplines on the area of biodegradation and bioremediation of environmental pollutants by molecular and whole-cell bacterial catalysts is thoroughly discussed and critically evaluated in our review article.[5] There, we also propose a workflow inspired by the DBTA cycle (nowadays routinely used during preparation of microbial cell factories for the production of valuable chemicals) for engineering natural or synthetic biodegradation pathways and bacterial degraders (Fig. 1). In this thesis, I focus on engineering biocatalysts for the degradation and potential valorization of halogenated waste chemicals that become persistent pollutants once released into the environment. I use our work on the synthetic catabolic pathway for 1,2,3-trichloropropane as an example of such an effort.

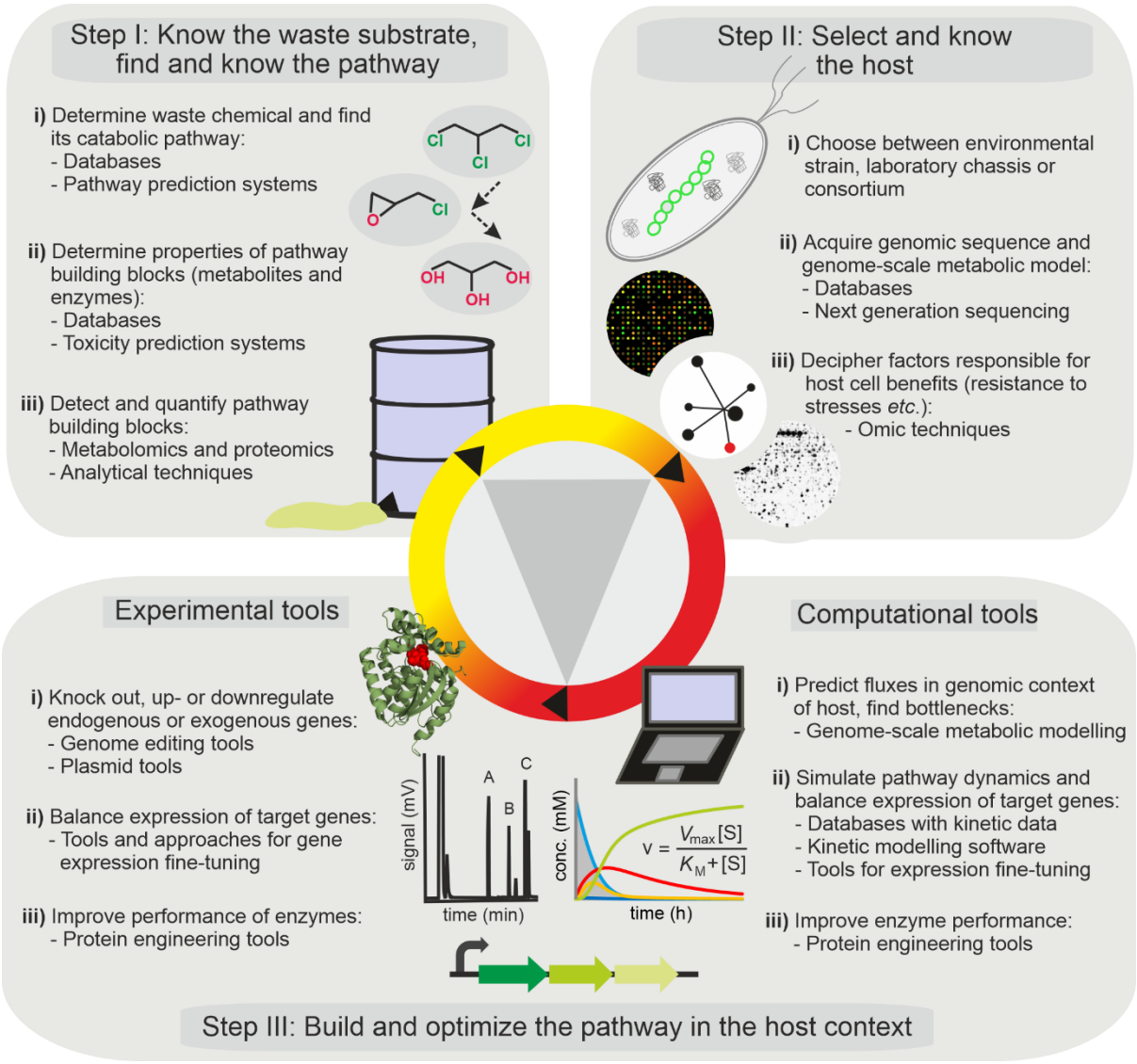


Figure 1 Workflow for the engineering of biodegradation pathways using tools and approaches of metabolic engineering, protein engineering, and synthetic biology. The Figure depicts the roadmap to superior bacterial catalysts for the biodegradation of waste organic chemicals that may act as environmental pollutants. Adopted from Dvořák et al. (2017)[5] and modified.

3.1 HALOGENATED HYDROCARBONS AND DEHALOGENATING ENZYMES

Halogenated hydrocarbons are organic chemicals in which at least one hydrogen atom is replaced by halide (F, Cl, Br, or I). They form a substantial portion of the problematic xenobiotics mentioned in the previous chapter. These compounds have been used on a large scale as pesticides (e.g., DDT, lindane, chlordane, atrazine) warfare agents (Agent Orange, yperite), soil fumigants (e.g., 1,2-dibromoethane), dielectric fluids (e.g., polychlorinated biphenyls), or industrial solvents and precursors in organic synthesis (e.g., 1,2-dichloroethane, 1,2,3-trichloropropane). Even though synthetic organohalides have thousands of natural structural analogs (for instance marine product 2,4-dibromophenol or actinomycete product dichloroaminobutyrate), many act as persistent organic pollutants toxic to the living organisms and regulated by the Stockholm convention.[15,16] Toxic and genotoxic effects have been well described on mammalian models especially for chlorinated anthropogenic chemicals.[17,18] The persistence and toxicity of halogenated hydrocarbons are often caused by the lack of abiotic and biotic transformation and local accumulation to the harmful concentrations. The major limitation for the complete mineralization of certain synthetic halogenated compounds is the occurrence of incomplete catabolic pathways in microorganisms and the corresponding accumulation of toxic intermediates that are generated after initial biotransformation reactions.[8]

Biotic transformation of halogenated hydrocarbons mainly (but not exclusively) by bacterial degradation predominates in the environment. Bacterial degraders of organohalide possess a broad palette of enzymes that can labilize or cleave carbon-halogen bonds via one of the known biochemical mechanisms: hydrolytic substitution, oxidation, or reductive dehalogenation in anaerobic conditions.[19] Hydrolytic enzymes of haloalkane dehalogenases (HLDs, EC 3.8.1.5) and halohydrin dehalogenases (class of lyases EC 4.5.1.-) play a prominent role in bacterial dehalogenation. These industrially relevant enzymes convert halogenated compounds to haloalcohols and haloalcohols to the corresponding epoxides, respectively.[20,21] Epoxide residues can be then hydrolyzed by available epoxide hydrolases to alcohols that can be further processed by microbial metabolism.

The rate and selectivity of the initial dehalogenation step are often pivotal for the successful degradation of a given chemical. HLDs are α/β-hydrolases that catalyze the cleavage of a carbon-halogen bond via nucleophilic substitution (S_N2 mechanism) followed by the addition of water (Fig. 2). No co-factor is needed. Due to the attractive substrate specificity (processing of over a hundred chlorinated, brominated, and iodinated aliphatic compounds by HLDs have been reported), catalytic mechanisms, frequent enantioselectivity, and high robustness (many HLDs can be produced as soluble proteins in heterologous *Escherichia coli* host) of members of this enzyme family, haloalkane dehalogenases have been considered for numerous applications described in our review article from 2013.[20] These also include biodegradation or recycling of halogenated by-products and waste compounds from the chemical industry that can act as toxic persistent pollutants if unintentionally or intentionally released into the environment. This is the case of 1,2,3-trichloropropane.

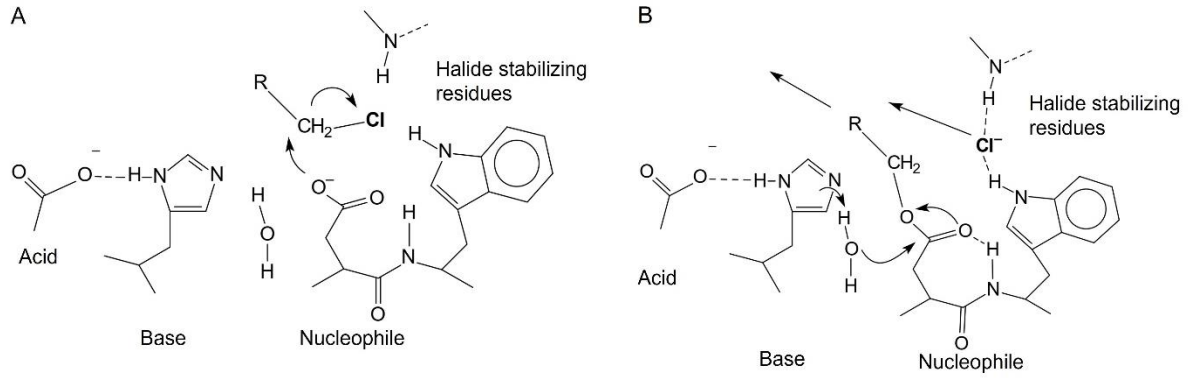


Figure 2 General catalytic mechanism of haloalkane dehalogenases. **A)** Formation of the covalent intermediate by S_N2 substitution. **B)** Hydrolysis of the intermediate. Adopted from Janssen (2004)[22] and modified.

3.2 1,2,3-TRICHLOROPROPANE AND THE MEANS FOR ITS REMOVAL

1,2,3-trichloropropane (TCP) is an anthropogenic compound recognized by the US Environmental Protection Agency as an emerging contaminant of ground waters.[23] TCP is a volatile colorless liquid produced worldwide by major chemical companies such as Shell or Dow Chemical Company in quantities reaching 50,000 metric tons annually. It has been used as a solvent and precursor in the chemical synthesis of epichlorohydrin,

hexafluoropropylene, dichloropropene, or polysulfone liquid polymers. TCP was spread into the environment and became a dreaded contaminant of soils and waters in states of EU, USA, and Asia mainly due to the improper waste management and its intentional incorporation in soil fumigant 1,3-dichloropropene.[24] High quantities of TCP can be found at industrial hazardous waste sites. Cases of accidental ingestion or inhalation of air contaminated with TCP revealed its acute toxicity in humans. Based on the studies with mice and rats, TCP is an anticipated human carcinogen.[25] Due to the health risks, in 2017 the State of California assigned an extraordinary low maximum contaminant level for TCP of 5 ppt (parts per trillion).[26]

Physicochemical properties of TCP such as low reactivity, low water solubility, and high density contribute to its persistence in the environment[24]. The half-life of its abiotic hydrolysis at pH 7.0 and 25 °C is estimated to be hundreds of years. Conventional remediation techniques including pump and treat technologies and the sorption to granular activated carbon, or vacuum extraction combined with chemical oxidation are relatively inefficient and expensive. Promising is the reductive conversion of TCP by zero-valent zinc. Biotic conversion can occur under aerobic and anaerobic conditions. Reductive dehalogenation of TCP in anoxic conditions by bacterial strains from the genus *Dehalogenimonas* was reported,[27] but the drawback of this conversion is that it gives rise to toxic products (allyl chloride or allyl alcohol). Moreover, the individual steps of this conversion are still unclear, though, several possible pathways for TCP transformation under reducing conditions have been recently proposed by computational chemistry methods.[28] Aerobic bioconversion of TCP is an attractive alternative as it can lead to the complete mineralization of the chemical and it supports bacterial growth. Theoretical calculations indicated that complete mineralization of TCP by bacterial metabolism is thermodynamically feasible.[29] However, the only reported natural organisms that could utilize TCP aerobically - methanotrophic bacterium *Methylosinus trichosporium* OB3b and four propane-oxidizing bacteria of genera *Rhodococcus*, *Mycobacterium*, and *Sphingopyxis* - could do so only in co-metabolic mode using their methane monooxygenase and propane monooxygenase, respectively.[30,31] Toxicity of TCP and oxidative stress caused by presumed reaction intermediates of oxidative catabolism of chlorinated aliphatics (e.g., epoxides) might pose a serious bottleneck for the natural evolution of an efficient mineralization pathway. Due to the urgent need for efficient degradation of TCP and the lack of natural means for such a task, this compound got into the sight of bioengineers.

3.3 SYNTHETIC BIOCHEMICAL PATHWAY FOR BIODEGRADATION AND POTENTIAL VALORIZATION OF TCP

The synthetic catabolic pathway for aerobic utilization of TCP was designed by the team of prof. Dick Janssen in 1999.[32] The design was based on two previous discoveries. The first was the isolation of Gram-negative bacterium *Agrobacterium radiobacter* AD1 which was capable of aerobic utilization of dihalogenated alcohols (1,3-dichloropropan-1-ol or 3-chloropropane-1,2-diol) and epoxides (epichlorohydrin, glycidol).[33] The product of the dehalogenation reactions – glycerol - was further phosphorylated and oxidized and served as a carbon and energy source for the growth of the host bacterium. The second was the characterization of new HLD DhaA from *Rhodococcus rhodochrous* NCIMB13064 by Kulakova et al.[34] DhaA was the first enzyme capable of TCP dehalogenation. It converted prochiral TCP into both (*R*)- and (*S*)- enantiomers of 2,3-dichloropropan-1-ol (DCP) with a slight preference for (*R*)-DCP. However, genes for haloalcohol metabolism were missing in *Rhodococcus* and the bacterium could not utilize TCP. The synthetic TCP pathway proposed by Prof. Janssen's team (Fig. 3) could lead to the complete mineralization of TCP. But it also offered the theoretical option of valorization of this industrial waste chemical via streaming of glycerol from TCP to a desirable bioproduct. Glycerol has been reported as a suitable substrate for microbial production of, e.g., citric acid, 1,3-propanediol, or bacterial bio-based plastics polyhydroxyalkanoates.[35] The initial effort was nonetheless fully concentrated on TCP biodegradation. Bosma and co-workers completed the artificial route by heterologous plasmid-based expression of *dhaA* gene in *A. radiobacter* AD1.[32] The resulting recombinant grew on brominated propanes but not on TCP. This failure was attributed to the poor catalytic efficiency of wild-type DhaA with the substrate of anthropogenic origin (k_{cat}/K_m=40 $s^{-1}.M^{-1}$). Therefore, the attempts to improve the performance of the synthetic TCP pathway logically started with the protein engineering of DhaA.

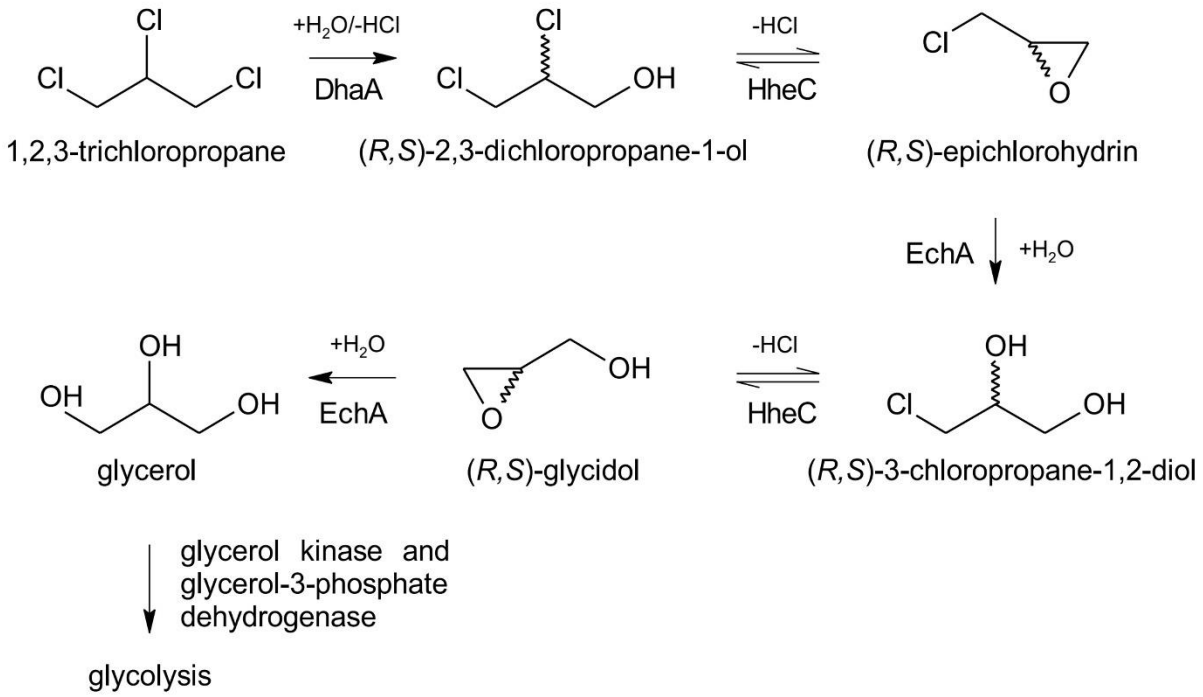


Figure 3 Five-step biotransformation of 1,2,3-trichloropropane into glycerol by the enzymes of the synthetic biodegradation pathway: haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCIMB 13064, haloalcohol dehalogenase HheC and epoxide hydrolase EchA from *Agrobacterium radiobacter* AD1.

3.4 IMPROVEMENT OF THE TCP PATHWAY BY PROTEIN ENGINEERING

In the first trial to improve the TCP pathway, Bosma and colleagues adopted DNA shuffling and error-prone PCR technologies to introduce random mutations in *dhaA* gene.[36] Subsequent screening of approximately 20,000 clones revealed mutant variant DhaA-M2 with two amino acid substitutions Cys176Tyr and Tyr273Phe and ~ 7-fold improved catalytic efficiency (k_{cat}/K_m=280 $s^{-1}.M^{-1}$) when compared with DhaA WT. Molecular modeling with available DhaA crystal structure indicated that the Cys176Tyr mutation allowed more productive binding of TCP in the active site of DhaA, which was probably further enhanced by the second substitution. The recombinant *A. radiobacter* strain bearing dhaA-M2 on a plasmid with a strong constitutive promoter was able to utilize about 0.72 mol (~106 mg, 3.6 mM) of TCP in a 200 mL culture within a 10-day interval. This was a significant improvement in comparison with the previous study. However, the technology still suffered from the reduced viability of the host bacterium in the TCP concentrations higher than 1 mM and from the low flux through the pathway

caused by the suboptimal activity of DhaA with TCP and unmatched enantioselectivity of DhaA and HheC. The latter issue was revealed by the detected accumulation of (*S*)-DCP in culture supernatants. The pathway bottlenecks disabled the continuous growth of engineered strain on TCP in the prepared packed-bed reactor.

In the following years, several studies focused on the characterization of HheC and EchA enzymes and on their enhancement via directed evolution or targeted mutagenesis.[37–40] Yet, none of these studies aimed at the TCP problem. The substantial progress in TCP pathway engineering was achieved in the late 2000s by the research team of prof. Jiří Damborský (Loschmidt Laboratories). First, Banáš and co-workers used molecular dynamics simulations and quantum mechanics calculations to elucidate the impact of Cys176Tyr mutation in DhaA mutant on its activity and enantioselectivity with TCP.[41] This work highlighted the importance of two major access routes that connect the occluded active site of DhaA with the bulk solvent – an upper tunnel and a tunnel slot – for TCP catalysis. This information was then used by Pavlová and colleagues for further improvement of DhaA activity with TCP.[42] The authors adopted a computer-assisted design for targeting new mutations in the access routes of DhaA. Three new substitutions (Ile135Phe, Val245Phe, and Leu246Ile), which appeared beneficial for TCP catalysis, were introduced into the protein tunnels of DhaA-M2 by site-directed and saturation mutagenesis. The resulting five-spot mutant DhaA 31 showed ~ 26-fold higher catalytic efficiency with TCP (k_{cat}/K_m=280 $s^{-1}.M^{-1}$) than DhaA WT (Fig. 4). (*R*)- and (*S*)-DCP were formed from TCP in almost equimolar ratio. The structural basis of improved activity with TCP was revealed by molecular dynamics simulations with the obtained crystal structure of DhaA 31 and kinetic analyses. Introduction of bulky Phe and Ile residues in the access tunnels reduced the massive inflow of water molecules in the buried active site of DhaA and increased the occurrence of the productive conformation of the substrate and subsequent nucleophilic substitution.

In our parallel study, four new point mutants of DhaA with substitutions in another three positions in the vicinity of the protein tunnels (W141, A145, and A172) were prepared and biochemically characterized together with four mutants obtained during previous mutagenesis enterprise.[43] Kinetics of all eight mutants with TCP as well as inhibition with DCP product and proper folding of proteins was determined. Subsequent molecular dynamics simulations confirmed two previously observed access routes (main tunnel p1

and slot tunnel p2a) and identified three new pathways (p2b and p2c as branches of p2a slot, and transient tunnel p3) that connected the bulk solvent with the buried active site of DhaA. The simulations revealed the relevance of individual routes for the exchange of water molecules and/or release of reaction products (chloride, DCP) from the active site of DhaA, and emphasized the importance of solvation of reaction products for their release from the protein. We also described how individual mutations modify the accessibility of the pathways and mechanism of ligand exchange.

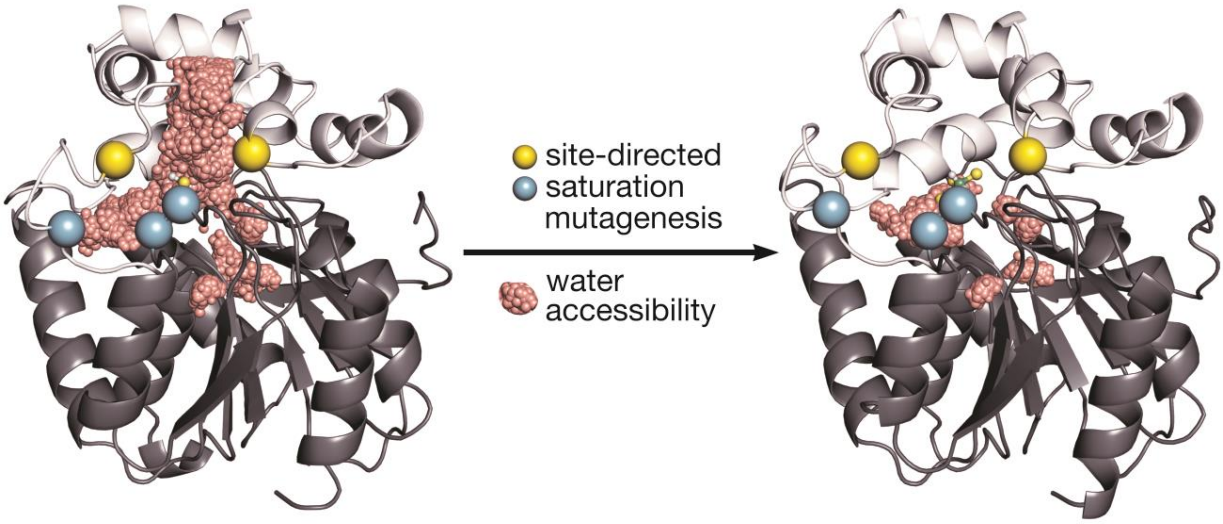


Figure 4 Structural basis of enhanced activity of DhaA31 mutant (on right) with 1,2,3-trichloropropane (TCP). Five amino acid substitutions (Ile135Phe, Cys176Tyr, Val245Phe, Leu246Ile, Tyr273Phe) introduced into the structure of wild-type haloalkane dehalogenase DhaA (on left) reduced the solvation of the active site. This promoted the formation of the activated complex of TCP and catalytic nucleophile Asp106 as water molecules sterically hinder access of TCP to this residue. Adopted from Pavlová et al. (2009).[42]

The three last research works described above shed the light on the conversion of the synthetic substrate by natural and engineered enzyme variants and paved the way towards the design of more efficient biocatalysts for anthropogenic chemicals. The advertised strategy of engineering enzyme tunnels for modification of properties of enzymes with the buried active sites has become widely accepted by the protein engineering community.[44]

Through the activity of DhaA with TCP was substantially improved, the enantioselectivity problem protruding in the preferential conversion of (*R*)-DCP by HheC and accumulation

of (*S*)-DCP in the pathway was not solved. No natural halohydrin dehalogenase capable of preferential conversion of (*S*)-DCP was available. Hence, van Leeuwen and colleagues from Prof. Janssen's group applied five rounds of saturation mutagenesis in multiple positions both close and distant to the active site of DhaA31 to obtain a variant converting TCP preferentially to (*R*)-DCP.[45] Indeed, new mutant DhaA r5-90R (abbreviated here as DhaA90R) with thirteen amino acid substitutions and substantially improved (*R*) selectivity (e.e. = 90%) compared to DhaA31 (e.e. = 13%) was obtained after screening of 5,500 clones. The drawback of this enantioselective variant was the low activity with TCP which dropped back to the level of wild-type DhaA or even below it (k_{cat}/K_m=25 $s^{-1}.M^{-1}$).

Several new DhaA protein variants with altered activity and enantioselectivity were thus available. The time has come to try other than protein engineering strategies that could probe the utility of these new mutants for TCP degradation and enhance the performance of the whole synthetic pathway.

3.5 IMPROVEMENT OF THE TCP PATHWAY BY METABOLIC ENGINEERING AND SYNTHETIC BIOLOGY

The TCP pathway represents a remarkable model system of a synthetic biodegradation route for a toxic anthropogenic waste compound. It is relatively simple, it requires no cofactors, and all its three enzymes are sufficiently robust and can be heterologously produced in *E. coli* and purified. We took the advantage of these characteristics and decided to dissect the pathway functioning, reveal the remaining bottlenecks, and use the obtained information for its knowledge-driven engineering and improvement. The previously proposed DBTA cycle (Fig. 1) was adopted in this endeavor. Our workflow included several major steps: **(a)** We first reconstructed the pathway in vitro using purified enzymes and we determined properties of individual pathway building blocks (enzymes and metabolites). **(b)** We then adopted existing or prepared new analytical methods for quantification of pathway building blocks and we built a kinetic model of the whole route which was used for the optimization of *in vitro* enzymatic cascade. **(c)** The pathway was then re-built *in vivo* in selected heterologous host *E. coli* BL21(DE3) and the mathematical model was employed for the evaluation of pathway limits in the living bacterial cells. (**d**) *In vivo* system enabled the study of specific parameters including

metabolite toxicity and metabolic burden that stemmed from the expression of exogenous genes. These parameters were included in the upgraded version of the mathematical model that can be used for new designs of bacterial TCP degraders. **(e)** Last but not least, an immobilized version of the synthetic TCP pathway was prepared and tested and its advantages over the in vivo system were discussed. The following chapters provide a more detailed description of these individual steps.

3.6 *IN VITRO* AND *IN SILICO* RECONSTRUCTION OF THE TCP PATHWAY

In vitro enzymatic cascades have certain advantages over their *in vivo* counterparts. Out of the context of a living cell, they do not suffer from metabolic cross-talks and prospective toxicity of pathway intermediates, they can be easily manipulated by simple mixing of cascade components in a test tube, and their study can provide essential information on pathway dynamics and constraints.[46] These attractive features of *in vitro* multi-enzyme systems have been examined and praised in numerous studies working with up to dozens of purified enzymes or with enzymes in cell lysates.[47] We hypothesized that the reconstruction of TCP route *in vitro* will shed a light on anticipated pathway bottlenecks and yield strategies for their mitigation or evasion. We purified available pathway enzymes including DhaA mutant variants using affinity chromatography (DhaA WT, DhaA31, DhaA90R, EchA) or anion exchange chromatography (HheC) and we determined their kinetic parameters with respective substrates from TCP pathway in the identical experimental conditions.[48] Altogether 16 steady-state kinetic parameters including two inhibition constants were used for the assembly of a kinetic model of TCP pathway. All pathway reactions in the model followed basic Michaelis-Menten kinetics. Kinetic models of metabolic pathways are invaluable tools for the assessment of dynamic responses of the metabolic network to the changing conditions. In such term, they are better than other types of mathematical models that describe cellular biochemistry.[49] However, accurate experimentally validated kinetic models of cell-free multi-enzyme systems are still rare.

One-pot multi-enzyme conversion of TCP to glycerol was first evaluated with DhaA WT (DhaA31 or DhaA90R), HheC, and EchA mixed in a 1:1:1 ratio. The newly developed GC-

MS method enabled parallel quantification of all pathway metabolites except for glycerol product in time. The time course of the reaction disclosed accumulation of (*S*)-DCP and glycidol – the latter caused by the inhibition of the last reaction step by TCP substrate and glycerol product. The kinetic model was validated against the experimental data and subsequently used for simulations that should determine loadings of individual pathway enzymes needed to achieve TCP conversion to glycerol in a defined time interval. These simulations allowed us to dissect the impact of protein engineering and of the optimization of enzyme loadings on TCP pathway efficiency. The obtained computational simulations were verified by one-pot experiments with purified enzymes and the experimental data matched the simulations very well. DhaA31 was confirmed as the best DhaA variant for the fastest conversion of TCP to glycerol. Its inclusion into the *in vitro* pathway could alone reduce the required total enzyme loading by 25 % (Fig. 5). If the ratio of enzymes in the pathway variant with DhaA31 was optimized by the kinetic modeling, up to 56 % of the total enzyme in the reaction mixture could be saved when compared with the basic unoptimized variant of the pathway including DhaA WT, HheC, and Echa mixed in 1:1:1 ratio.

To conclude, this study provided an in-depth description of TCP pathway bottlenecks and it gathered a quantity of new information that could be employed for planned knowledge-driven engineering of the route in a heterologous bacterial host. On top of this, we developed a workflow for optimizing in vitro multienzyme processes by kinetic modeling and biocatalyst stoichiometry tuning. Such strategy and freely available computer code, which was provided to the community, can be used for the refinement of other multi-enzyme cascades.

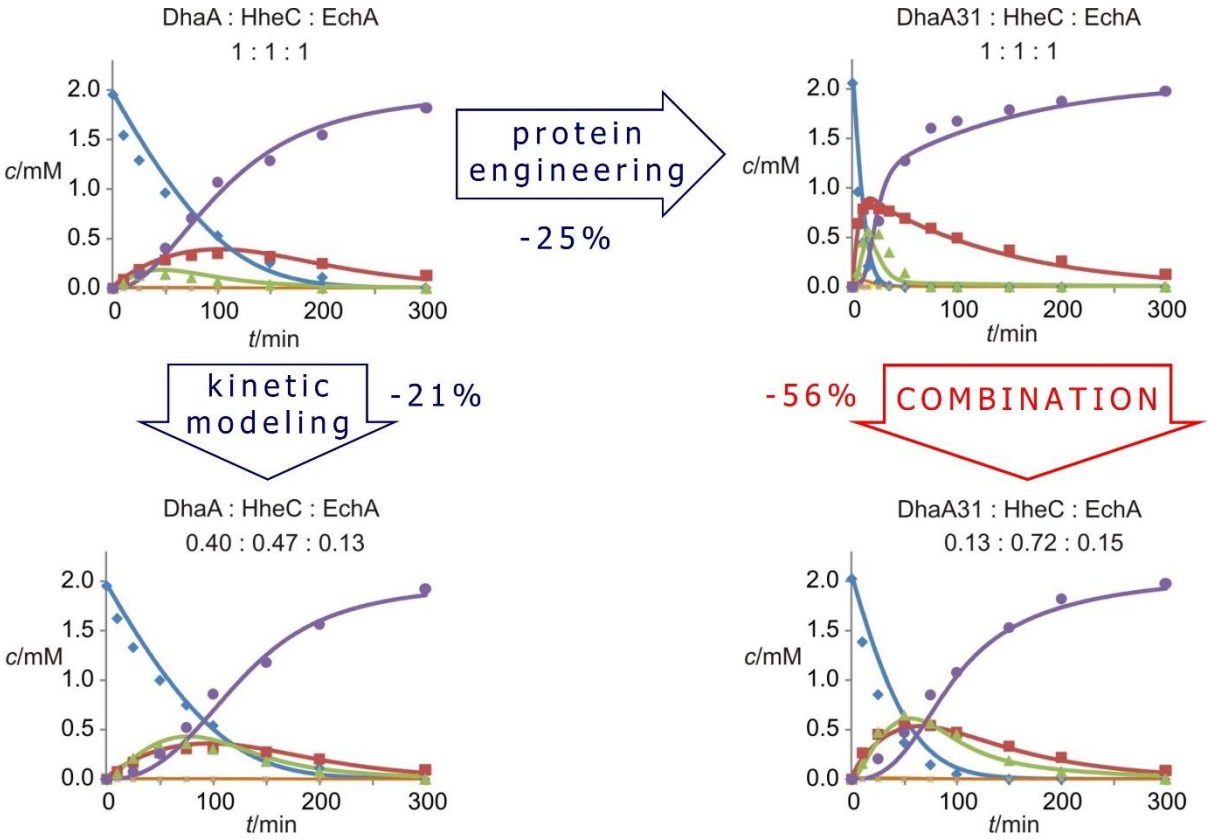


Figure 5 Separate and combined effects of kinetic modeling and protein engineering on in vitro three-enzyme conversion of 1,2,3-trichlororopane (TCP). Modeled and experimental results are indicated by solid lines and symbols, respectively. The goal was 95% conversion of TCP (2 mM) into glycerol within 300 min time interval by a non-optimized (1:1:1) or optimized ratio of pathway enzymes (DhaA/DhaA31, HheC, EchA). The portion of saved total enzyme in the reaction is shown in %. Blue line: TCP, red line: 2,3-dichloropropan-1-ol, orange line: 3-chloropropane-1,2-diol, green line: glycidol, violet line: glycerol.

3.7 ASSEMBLY AND OPTIMIZATION OF TCP PATHWAY IN HETEROLOGOUS BACTERIAL HOST *ESCHERICHIA COLI*

Bacteria are optimal host organisms for environmental and industrial biotechnologies. They have short generation time, usually low nutritional demand, and, depending on the species, can grow in diverse conditions including low or high pH, ionic strength, temperature, etc. As self-replicating biological micromachines, they can propagate the inserted genetic information and produce high quantities of protein biocatalysts. It is thus favorable to implant biodegradation or biosynthetic pathways of interest into a suitable bacterial cell factory that can perform the desired task. *E. coli* is the best-studied bacterium with a broad palette of metabolic engineering and SynBio tools available. Re-

factored strains of *E. coli* have been used for large-scale production of biofuels, biopolymers, and biopharmaceuticals, as well as for biodegradation of industrial wastes and environmental pollutants.[50] In the context of our previous *in vitro* work on the TCP pathway, *E. coli* offered one notable benefit. In contrast to *A. radiobacter* AD1, *E. coli*'s genome does not encode genes of enzymes from TCP pathway and the metabolites of this route, except for glycerol that can be easily metabolized, are alien to this bacterium. Hence, once implanted in *E. coli*, TCP pathway could act as orthogonal – parallel – to the native metabolism of this host and the metabolic crosstalk of the synthetic route and innate metabolic network should be minimal. This was the major prerequisite for the successful application of the previously constructed kinetic model for the description and optimization of the TCP pathway *in vivo*.

We introduced TCP pathway into popular *E. coli* strain BL21(DE3) often employed in metabolic engineering studies.[50] Three basic variants of the pathway including DhaA WT, DhaA31, or DhaA90R were assembled by cloning the respective genes in Duet plasmids.[51] This commercial (Merck) IPTG (Isopropyl β-D-1-thiogalactopyranoside)-inducible system allows for the tunable expression of multiple exogenous genes from combinations of compatible Duet plasmids in a single *E. coli* BL21(DE3) cell. The mathematical model of TCP cascade was enriched with two new experimentally determined parameters that were highly relevant for TCP catalysis *in vivo*: (i) toxicity of individual pathway metabolites to *E. coli* and (ii) copy numbers of Duet plasmids in *E. coli* cell. The upgraded model was then used to propose combinations of Duet plasmids whose use for the assembly of TCP pathway would lead to the fastest conversion of TCP to glycerol and the lowest possible accumulation of toxic intermediates. TCP and non-accumulating epichlorohydrin were identified in conducted growth tests as the most toxic chemicals in the synthetic pathway. Hence, fast removal of TCP became the priority and the model logically proposed pathway variants with DhaA31 as the most efficient. Several pathway variants proposed by the model were constructed and their performance was experimentally validated by using pre-induced resting cells of *E. coli* with overexpressed pathway enzymes. Production of pathway enzymes in individual *E. coli* degraders was quantified by sodium dodecyl sulfate–polyacrylamide gel electrophoresis and enzyme activity measurements. The excellent match between simulations and experimental data including the ratio of pathway enzymes in constructed strains and time courses of TCP conversion to glycerol recorded with pre-induced *E. coli* degraders confirmed

orthogonality of the synthetic route in the host cell. *E. coli* degraders with DhaA31 and the optimized ratio of pathway enzymes showed the highest viability of all prepared recombinants after the incubation with the toxic substrate.

However, none of the constructed strains was able to grow in a minimal salt medium with TCP used as a sole carbon source. We proposed several reasons of this unsatisfactory situation: (i) high toxicity of TCP towards *E. coli* (already 1 mM TCP is toxic for the host cell), (ii) metabolic burden caused by the heterologous expression of the synthetic pathway from Duet plasmids, and mainly (iii) low flux through the TCP pathway and insufficient production of glycerol that can be used by the cells as a carbon source for maintenance and growth. The latter issue could be solved by another round of protein engineering of DhaA. Using our model, we calculated the parameters of the "optimal" DhaA variant which would secure high flux through the TCP pathway and sufficient production of glycerol for sustainable growth of the host bacterium (Fig. 6). Such DhaA catalyst would have to possess either an order of magnitude higher catalytic efficiency (k_{cat}/K_m ≥2300 $s^{-1}.M^{-1}$), higher enantioselectivity (E-value ≥20) in favor of (*R*)-DCP production (E-value of DhaA31 is close to 1), or combination of both improved parameters (k_{cat}/K_m ≥700 $s^{-1}.M^{-1}$, E-value ≥10). Another option is the modification of enantioselectivity of HheC towards the equally fast conversion of (*R*)- and (*S*)-DCP or the obtention of a new (*S*)-selective halohydrin dehalogenase with reasonably high activity towards DCP. Unfortunately, neither the superb mutant of DhaA nor the (*S*)-selective halohydrin dehalogenase were available at that time and have not been acquired thus far. In our follow-up work we, therefore, focused on the two remaining problems of metabolic burden and substrate toxicity.

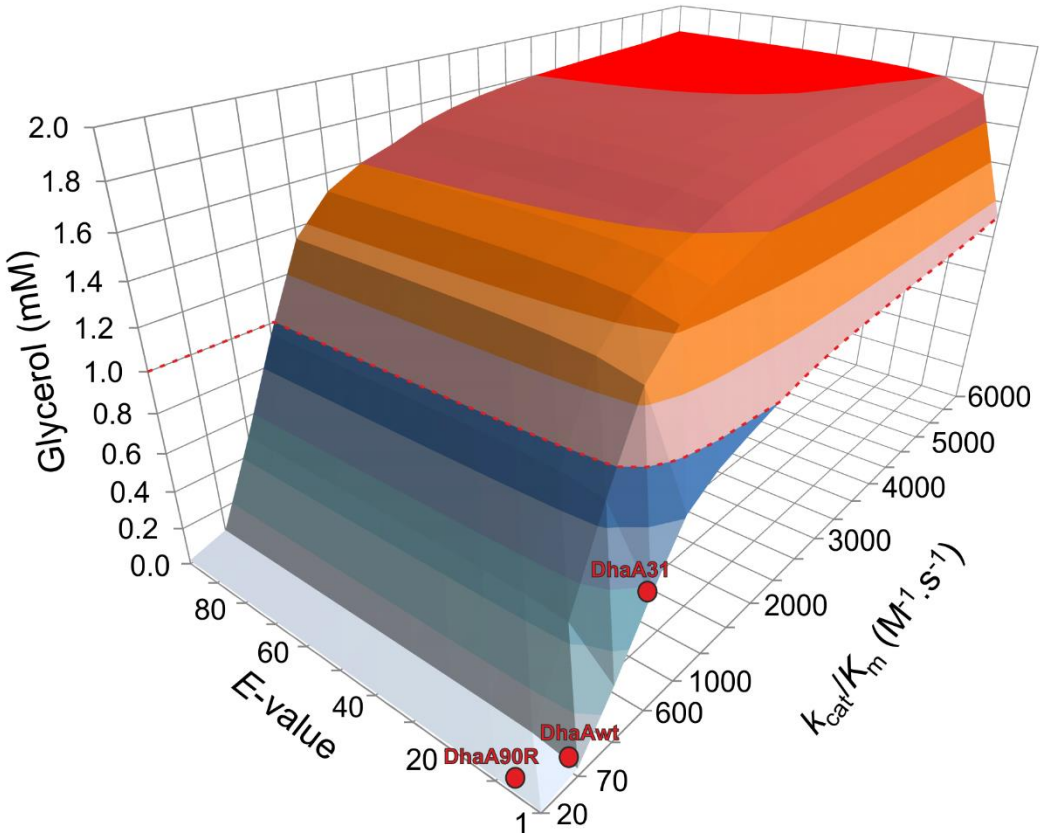


Figure 6 Hypersurface plot which shows the effect of catalytic efficiency (k_{cat}/K_m) and enantioselectivity (*E*-value) of DhaA on the production of glycerol from 2 mM 1,2,3-trichloropropane by the synthetic pathway. Red dots indicate the positions of the three mutant variants of DhaA discussed in the text. 1 mM glycerol (red dashed line) should be sufficient for the sustainable growth of engineered *Escherichia coli*. Adopted from Kurumbang et al. (2014).[51]

3.8 METABOLIC BURDEN AND METABOLITE TOXICITY AS KEY CHALLENGES FOR THE ENGINEERING OF BIODEGRADATION PATHWAYS *IN VIVO*

Besides suboptimal catalytic performance of enzymes from the target metabolic pathway, toxicity of a substrate, metabolic intermediates, or product(s) of the route and metabolic burden stemming from the expression of the pathway genes in a host cell are two key factors that may affect the fruitfulness of a given engineering enterprise.[52,53] These stress elements are often encountered during the engineering of biodegradation and biosynthetic pathways and the gathered knowledge on their mitigation is vast. Studies that describe and dissect the combined effects of these stress factors on the microbial host are nonetheless scarce.

In our work published in Microbial Cell Factories, we evaluated the consequences of simultaneous incidence of TCP toxicity and metabolic burden from exogenous metabolic pathway produced via IPTG-inducible LacI^{Q}/P_{lacUV5}-T7 expression system in *E. coli* BL21(DE3) cells.[54] BL21(DE3) cells and the inducible T7 RNA polymerase-dependent expression system based partially in the genome of the bacterium and partially on the compatible pET or Duet vectors are widely used for the controllable expression of recombinant proteins or whole biochemical pathways in metabolic engineering studies. But their utility for the production of a synthetic biodegradation pathway and conversion of toxic anthropogenic waste such as TCP had not been thoroughly investigated prior to our work.

To address this knowledge gap, we employed the *E. coli* BL21(DE3) degraders from our former study[51] and newly prepared control strains bearing empty Duet plasmids to dissect the contributions of TCP toxicity and metabolic burden from the pathway expression to the reduced viability of the strains performing the harsh biotechnological task. This was achieved by the combination of methods including conversions of TCP to glycerol by pre-induced *E. coli* resting cells and subsequent evaluation of changes in cellular morphology, physiology, and viability by plating experiments, flow cytometry, and transmission electron microscopy. Interestingly, the obtained data showed that the toxic effect of TCP on *E. coli* cells is substantially exacerbated by the addition of TCP. The addition of 0.2 mM IPTG reduced the viability of *E. coli* host with empty plasmids more than incubation with TCP. But the combination of both IPTG and TCP had the most devastating effect which could be partially alleviated by the presence of efficient variant of TCP pathway with DhaA31 and fast removal of the toxic substrate. We demonstrated that the cells pre-induced with IPTG and exposed to 2 mM TCP suffer from increased permeability of the plasmatic membrane, reduced membrane potential, and massive formation of reactive oxygen species (ROS). Though the mode of action of TCP in bacterial cells is not fully uncovered, it is known that chlorinated aliphatics and the derived epoxides can form DNA adducts and induce genetic damage.[29] Our investigation highlights another effect of lipid bilayer de-stabilization (caused probably by saturation of fatty acids via halogenation or peroxidation) and subsequent collapse of electron transport chain followed by the burst of ROS (e.g., hydrogen peroxide or superoxide radicals) that cause oxidative stress and potentially also cell death. The membrane

damage is exacerbated by the hydrophobic nature of TCP and yet unknown mechanism triggered by the addition of IPTG to the bacterial cells.

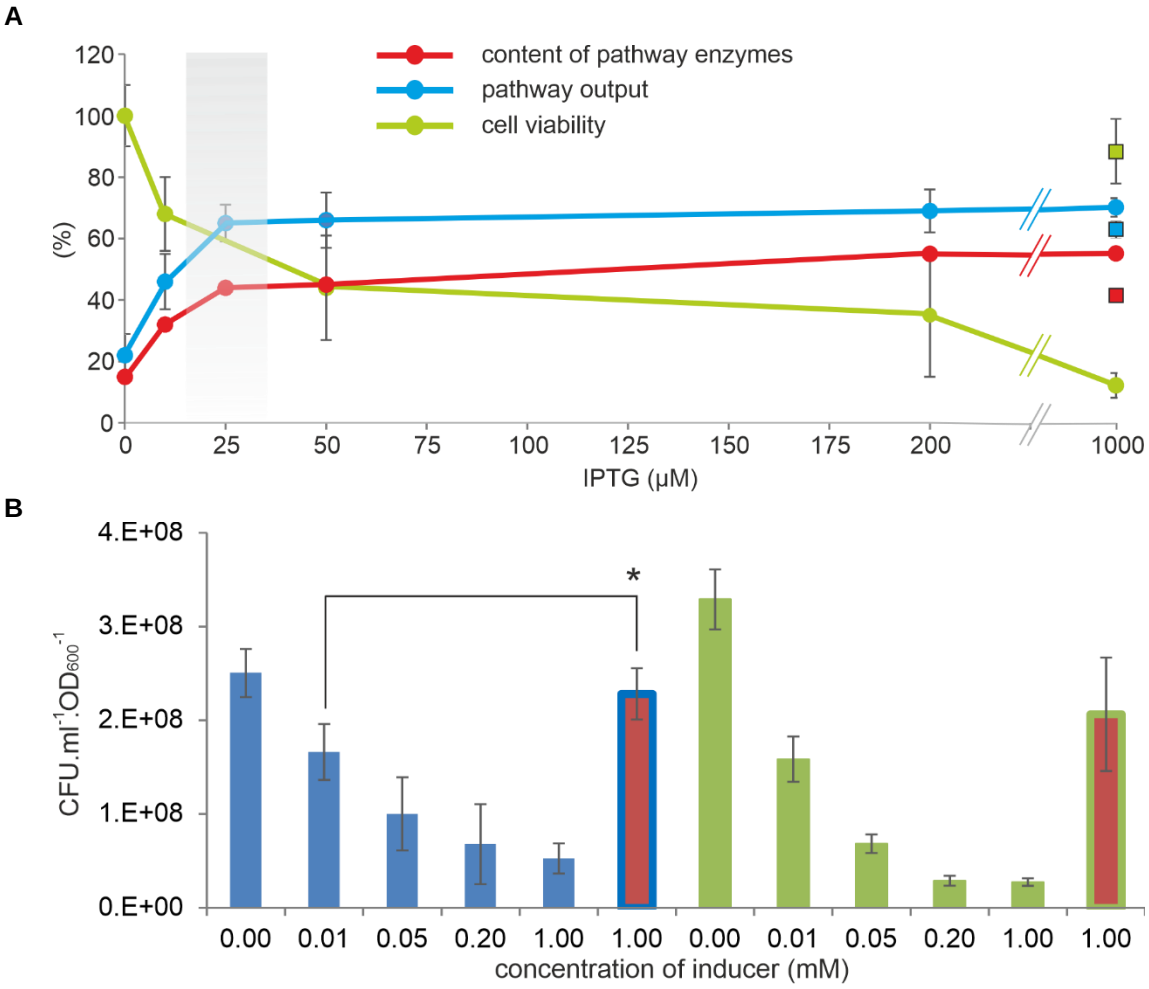


Figure 7 The effect of the concentration of IPTG inducer on the performance of the TCP pathway and the viability of the host bacterium *Escherichia coli* and comparison with the use of natural inducer lactose. **A)** Summarized effects of IPTG concentration on the expression of pathway genes (shown as portion of pathway enzymes in total cytoplasmic protein), pathway output (conversion of TCP to glycerol), and cell viability. Shown data points are means from three independent experiments ± standard deviation. Values determined for engineered *E. coli* degrader pre-induced with 1 mM lactose are indicated by squares. **B)** Viability of pre-induced *E. coli* resting cells with (blue columns) and without (green columns) TCP pathway after incubation with 2 mM TCP. Asterisks denote significantly higher (at $P < 0.05$) cell count of *E. coli* cells pre-induced with 1 mM lactose (red columns) when compared with the count of cells pre-induced with the lowest tested concentration of IPTG (0.01 mM). Shown data points are means from at least four independent experiments ± standard deviation. CFU, colony forming units. Adopted from Dvořák et al. (2015).[54]

Taken together, the study evidenced negligible metabolic burden from the TCP pathway enzymes but significant load from the used plasmids and synthetic inducer which magnified toxicity of the anthropogenic substrate. The outcomes of this work are of general interest for the biotech community because it showed that IPTG is far from being an innocuous molecule (the paper was among the most influential articles in Microbial Cell Factories in 2015 and 2016 based on Altmetric.com). We demonstrated that its negative effect can be effectively mitigated by the fine-tuning of its concentration (already 25 μM IPTG was enough for full induction of TCP pathway expression) or by using natural P_{lac} inducer lactose instead (Fig. 7). In addition, the collected experimental data enabled calibration of a next-generation mathematical model of TCP pathway.[55] For the first time, this model addressed the combined effects of metabolic burden and toxicity exacerbation in the context of the growing bacterial population. The upgraded model and the comprehensive knowledge gained during the above-described engineering endeavor will be essential for any further attempt to refine the TCP pathway performance in *E. coli* host and instrumental for similar projects coping with biotechnological processing of toxic chemicals. Still, one important lesson from our work is that the effects of numerous disadvantageous factors encountered during harsh biotechnological conversions can add up or even multiply to form a serious hurdle for efficient *in vivo* catalysis.

3.9 IMMOBILIZATION OF TCP PATHWAY AS AN ALTERNATIVE STRATEGY FOR BIOTECHNOLOGICAL PROCESSING OF TOXIC INDUSTRIAL WASTE

The experienced problems of high TCP toxicity, the suboptimal performance of the TCP pathway, and metabolic burden from pathway overexpression made us search for an alternative to bacterial degraders. As already described in one of the previous chapters, cell-free biocatalysis offers several considerable advantages over the use of microbial cell factories.[46] *In vitro* biological systems assembled from components obtained from living organisms (purified proteins or cell lysates) are usually not compromised by toxicity issues. Cell-free enzymatic cascades provide appreciable freedom in adjusting the ratio of individual proteins and total enzyme load in the reaction that cannot be exploited with a living organism. The metabolic burden is not a topic for discussion in *in vitro* systems. However, the use of free purified enzymes is not very practical neither economical and

the way of their stabilization and recycling must be found to set up viable cell-free biotechnology. A popular option is enzyme immobilization. We hypothesized that the immobilized TCP pathway could surpass the major drawbacks of bacterial degraders.

Dozens of techniques and support materials for the immobilization of proteins or whole cells are currently available.[56] But prior to our work, no immobilization protocols were established for DhaA31, HheC, and HheC, not talking about the lack of protocols for the immobilization of whole catabolic cascades. We encapsulated the three purified TCP pathway enzymes into the lens-shaped polyvinyl alcohol (PVA) particles known under the commercial name LentiKats at that time (Fig. 8A).[57] This low-cost technique had been previously found useful for the immobilization of haloalkane dehalogenase LinB[58] and the LentiKats technology had already been applied for large-scale applications such as wastewater treatment or synthesis of beta-lactam antibiotics.[59] PVA particles with encapsulated enzymes can be easily separated from the reaction mixture and recycled. However, the pores in the PVA gel are too big to arrest proteins smaller than 80 kDa. Consequently, tetrameric HheC (Mw = 112 kDa) was encapsulated directly while monomeric molecules of DhaA31 and EchA (Mw = 33 and 34 kDa, respectively) were first precipitated with ammonium sulfate and interconnected in cross-linked enzyme aggregates (CLEAs) using dextran polyaldehyde.[60] Immobilization reduced the activity of all three enzymes with their respective substrates from TCP pathway by 27–83 %. Interestingly, the overall efficiency of the cascade of immobilized enzymes mixed in a 1:1:1 ratio (measured as % of the conversion of 5 mM TCP to glycerol) was reduced only by 11 % when compared to the reaction with free enzymes. The long-term stability of immobilized enzymes did not differ significantly from the stability of their free counterparts. But in contrast to free enzymes, PVA particles were perfectly re-usable. After 10 successive cycles of TCP conversion to glycerol in batch operation, the immobilized route retained 77 % of its initial efficiency.

The major endurance test for the immobilized cascade was its adoption for the continuous removal of TCP from contaminated water in a bench-scale packed bed reactor (Fig. 8B). The two columns of the reactor were packed with PVA lentils containing DhaA31 (column 1) or with a mixture of LentiKats bearing HheC and EchA (column 2). The total mass ratio of the three enzymes was kept 1:1:1. During 10 weeks of operation at room temperature, the bioreactor processed 11 L of buffered water (pH 8.2)

contaminated with TCP concentrations as high as 10 mM (1.47 g/L, which is the water solubility limit of TCP). In total, 65.5 mmol of 67.7 mmol of toxic TCP that entered the system was converted to other chemicals (efficiency 97 %) out of which final product glycerol made 52.6 mmol and the rest was DCP and glycidol that slowly cumulated in the system during the reaction course. The immobilized cascade thus proved its utility for a long-term continuous conversion of a remarkable load of a highly toxic waste compound to attractive commodity chemical. One can imagine the use of an upgraded version of this biotechnology for a large-scale valorization of troublesome TCP residues in industry. Besides, kinetic parameters of DhaA31 bring a promise for the potential application of the immobilized cascade for biodegradation of much lower concentrations of TCP in contaminated groundwaters.[61] The developed cell-free biotechnology definitely needs further refinement and validation before any "real-world" application (all the potential drawbacks are discussed in our paper). But already this laboratory beta version showed the great potential of *in vitro* systems for the bioprocessing of toxic waste chemicals. The immobilized biocatalysts (encapsulated in form of purified enzymes or cell lysates) are robust, recyclable, compatible with concentrations of TCP that would kill a bacterial host, can operate under mild non-sterile conditions, and are not limited with restrictions for genetically modified organisms.

I am personally very glad that this pioneering study (to the best of our knowledge, this was the first report on an immobilized synthetic biodegradation pathway with engineered enzyme) drew the attention of a broader audience and its quality was repeatedly appreciated by the research community. It became a part of a monthly collection of papers selected by the editors of the American Chemical Society and received a prize for the best Technology article out of 1,500 papers in Environmental Science and Technology in 2014.

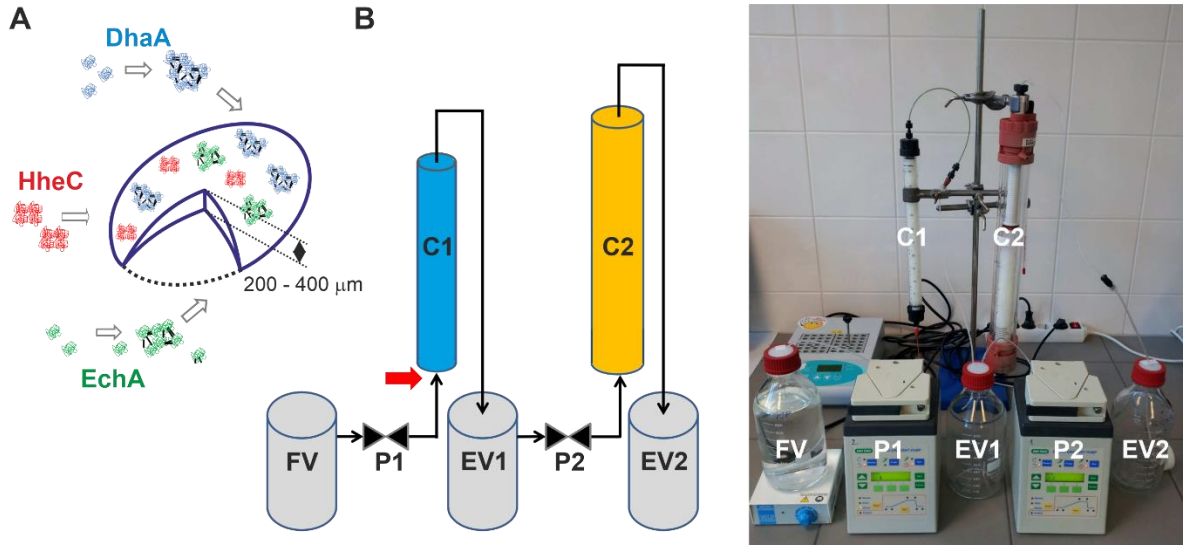


Figure 8 Schematic illustration of the immobilized TCP pathway and its use in the bench-top packed bed reactor for removing 1,2,3-trichloropropane from the water under continuous operation. **A)** Haloalkane dehalogenase DhaA and epoxide hydrolase EchA were first aggregated in cross-linked enzyme aggregates (CLEAs) and then encapsulated into lens-shaped polyvinyl alcohol particles. Molecules of tetrameric halohydrin dehalogenase HheC were encapsulated directly. **B)** The constructed reactor consisted of two glass columns (C1 and C2) packed with immobilized biocatalysts, two peristaltic pumps (P1 and P2), and three glass vessels connected with polytetrafluoroethylene tubing. A feed vessel (FV) contained 1 L of buffered (0.1 M Tris-SO_4) water with up to 10 mM TCP. Effluent vessel 1 (EV1) was used to collect samples downstream of column 1 and as a feeding bottle for column 2; effluent vessel 2 (EV2) was used to collect final samples downstream of column 2. The site for determining the TCP concentration in the input of column 1 is indicated with a red arrow.

3.10 PERSPECTIVES OF BIOTRANSFORMATION OF TCP

Even after more than two decades of intense research, the issue of TCP has not been convincingly solved. The synthetic biodegradation pathway has been already remarkably improved by the means of protein engineering, metabolic engineering, and synthetic biology. But the persisting kinetic bottlenecks backpedal the flux through the route and make the constructed bacterial TCP degraders inefficient. For the same reason, relatively high loads of pathway enzymes are needed for the discussed immobilized cell-free cascade which makes the biotechnology costly. Further protein engineering of key limiting enzymes – DhaA and HheC – is one possible way towards needed improvement. Unfortunately, although activity and enantioselectivity of DhaA31 and HheC have been the targets of protein engineers in recent years[62–64] no new variant with the relevancy for

TCP biotransformation was obtained. Already the previous sophisticated engineering exercises that led to mutants DhaA31 and DhaA90R indicated that further improvement of this enzyme would not be easy. Maybe it is time to start thinking out of the box and try a new unconventional engineering strategy.

One such strategy parallel to the standard optimization of active sites and modifications in existing tunnels can be the opening of completely new tunnel(s) in the structure of an enzyme with a buried ligand-binding pocket. In the study of Brezovsky and co-workers, we rehearsed the opening of a new tunnel on an example of haloalkane dehalogenase LinB which is a close relative of DhaA (they both belong to the phylogenetic subfamily of haloalkane dehalogenases HLD-II).[65] The original main tunnel p1 in LinB structure was first closed by the introduced disulfide bridge and the activity of the mutant with the set of halogenated substrates was substantially reduced. Then, a new parallel tunnel was opened nearby the original p1 by targeted saturation mutagenesis of three pre-selected bulky amino acids (Trp140, Phe143, Ile211). The best resulting mutant with a new tunnel confirmed by X-ray crystallography and molecular modeling not only restored the activity with most of the tested halogenated substrates to the level of wild-type LinB but it even outperformed the parental enzyme with some of the biotechnologically relevant compounds (e.g., pesticide and chemical precursor 1,2-dibromoethane) and became the most active HLD known to date. Modifying the access of substrate, reaction products, and water molecules into the active site of DhaA previously led to the best currently available variant DhaA31. Though the effect of the opening of a new tunnel is difficult to predict, it is possible that this strategy applied on DhaA WT or DhaA31 could be a way forward also in the TCP story.

Another promising, yet conventional, option is bioinformatic searches for sequences encoding new dehalogenating enzymes in the richness of available genome and metagenome data. The challenging angle of this approach is a need for high-throughput protocols for rapid biochemical characterization of dozens to hundreds of pre-selected proteins.[66] In the recent paper of Vanacek et al., we proposed and realized a workflow encompassing computational screening, expression analysis, and biochemical and biophysical characterization of multiple members from a selected enzyme family.[67] The workflow was validated against the HLD family. Initial in silico screening including sequence searches, multiple-sequence alignment, homology modeling, or calculations of

cavities and tunnels prioritized 530 sequences of putative HLDs. 20 genes were synthesized and expressed in heterologous host *E. coli* providing eight soluble proteins with detectable dehalogenase activity measured with a set of halogenated substrates by robotic screening at several temperatures. Further kinetic and structural characterization revealed several enzymes with unique activity (HLD with the highest known k_{cat}/K_m of 96.8 mM^{-1} s^{-1}), stability (an enzyme with a melting temperature of 71°C), and substrate specificity (HLD active with sulfur mustard). No enzyme active with TCP or DCP was identified but the study paved the way for other even more extensive searches that could possibly achieve this goal in near future.

And finally, the searches for new robust microbes in sites polluted with TCP or DCP could unveil completely new biodegradation pathways for these chemicals. In 2012, Arif and colleagues reported strain *Pseudomonas putida* MC4 isolated from a site polluted with chlorinated alkanes close to a chemical storage facility in Rotterdam.[68] The strain could utilize both enantiomers of DCP for growth via the action of its oxidative pathway starting with quinohemoprotein alcohol dehydrogenase DppA showing dehalogenating activity. The strain was subsequently empowered with *dhaA31* gene integrated into chromosome and tested for growth on TCP in 25-day fed-batch cultivation.[69] Though the bacterium could not sustain concentrations of TCP higher than 0.33 mM (probably due to the formation of toxic intermediates or by-products of oxidative TCP metabolism) slow growth of the recombinant organism on the anthropogenic chemical was indeed observed (μ = 0.001–0.008 h^{-1}). Engineered *P. putida* MC4 thus represents an example that inspires all researchers interested in the TCP problem in their further investigations that will once lead to the large-scale biodegradation of this recalcitrant anthropogenic chemical.

4. SECTION II: ENGINEERING BACTERIA FOR UTILIZATION AND VALORIZATION OF SUBSTRATES FROM LIGNOCELLULOSIC RESIDUES

Lignocellulose (LC) is a non-edible portion of plant biomass made of polysaccharides, phenols, and proteins. Over 200 million tonnes of lignocellulosic and cellulosic residues from agricultural, forestry, municipal, and household solid wastes out of the total amount of about 900 million tonnes are potentially available annually for sustainable biotechnological purposes in the EU.[70] Even higher amounts should be available in the US.[71] The efficient use of renewable LC residues for bioproduction of biopolymers, biofuels, or biopharmaceuticals can facilitate a truly green economy.[72] Historically, glucose and xylose, obtained from the most abundant cellulose (25-55%) and hemicellulose (11-50%) fractions of LC have been the substrates of interest.[73] In existing bioprocesses, focused mostly on bioethanol production, LC biomass is first pre-treated and fractionated using (physico)chemical methods such as milling, grinding, acid or alkaline hydrolysis, or steam explosion. Then, the cellulose and hemicellulose fractions are hydrolyzed by off-site produced cellulases and hemicellulases and the resulting monomeric glucose and xylose are fermented typically by an engineered *Escherichia coli*, *Zymomonas mobilis*, or a yeast strain. However, due to the hindrances in LC bioprocessing caused by the recalcitrance and complex structure of LC feedstocks (Fig. 9), the majority of biochemicals produced nowadays on large scale (aliphatic compounds, aromatics, biopolymers) are still obtained by fermentation of the 1st generation substrates such as edible starchy biomass or pure starch-derived glucose.[74] The 1st generation biotechnologies thus compete with food production and are not sustainable. The efficient economical bioconversion of LC remains to be one of the major technologic challenges of our time.

The 2nd generation biotechnologies based on lignocellulosic substrates must efficiently co-process glucose and xylose with other abundant sugar monomers (e.g., arabinose, mannose) and oligomers (cello- and xylooligosaccharides) and aromatic chemicals derived from the lignin portion of LC (10-40%) to become economically viable.[71,75,76] This is not happening because traditional microbial hosts that have been used for processing of LC-derived sugars, such as *E. coli* or *Saccharomyces cerevisiae*, suffer from certain drawbacks: (i) they are sensitive to toxic effects of inhibitory products, namely acids,

aromatic and heterocyclic compounds, resulting from lignocellulose pre-treatment, (ii) they can valorize neither lignin nor cello- and xylooligosaccharides, major by-products of lignocellulose pre-treatment and saccharification, respectively, (iii) they cannot assimilate hexoses and pentoses simultaneously due to the innate regulatory mechanisms causing carbon catabolite repression (CCR). The price and feasibility of LC pre-treatment and fractionation, and cellulose and hemicellulose hydrolysis by cellulases and hemicellulases produced off-site in fungi remain to be limiting factors too despite the recent advances in these technologies.[73] Consequently, appeals for the study of alternative strategies and microbial hosts that would lead to the wiser use of LC residues emerge.[75,77–79]

Parallel conversion of carbon sources from all three major polymeric components of LC to a single product or several products by microorganisms that can also make their own (hemi)cellulases and ligninases and thus reduce the need for added depolymerizing enzymes is considered a promising direction toward sustainable LC biotechnologies.[77–79] An extreme case of such a scenario in which both biomass degradation and bioproduct formation capacities are joint in a single microorganism is called consolidated bioprocessing (CBP). The question is how to achieve that? Nature's best solutions combined with cutting-edge metabolic engineering and synthetic biology adopted in the DBTA cycle may provide the answer. Functional expression and secretion of cellulases and hemicellulases have been already engineered in biofuel-producing bacteria such as *E. coli* or in yeast strains.[80,81] One of the first inspiring examples of such effort was published by Bokinsky and co-workers who designed a synthetic consortium of *E. coli* strains with implanted genes encoding cellulases and xylanases with secretion signals to degrade and valorize cellulose and hemicellulose fractions of LC.[80] The consortium was able to produce small quantities of advanced biofuels (pinene, butanol, fatty acid ethyl esters) from pre-treated switchgrass. *E. coli*, *Z. mobilis*, *S. cerevisiae*, and other industrially relevant bacteria and yeasts were also engineered for co-utilization of glucose with LC-derived pentoses (D-xylose, L-arabinose). In addition, potent cellulolytic bacteria (e.g., *Clostridia*) and fungi (e.g., *Trichoderma reesei*) and ligninolytic white rot fungi have been intensively studied, either individually or in a form of synthetic consortia, for their use in bioprocessing of LC feedstocks.[79,82] However, a single microorganism with the suite of properties for co-valorization of cellulose, hemicellulose, and lignin in harsh conditions

of a bioprocess is still not available. Robust environmental bacteria such as certain pseudomonads can be suitable candidates for this challenging task.

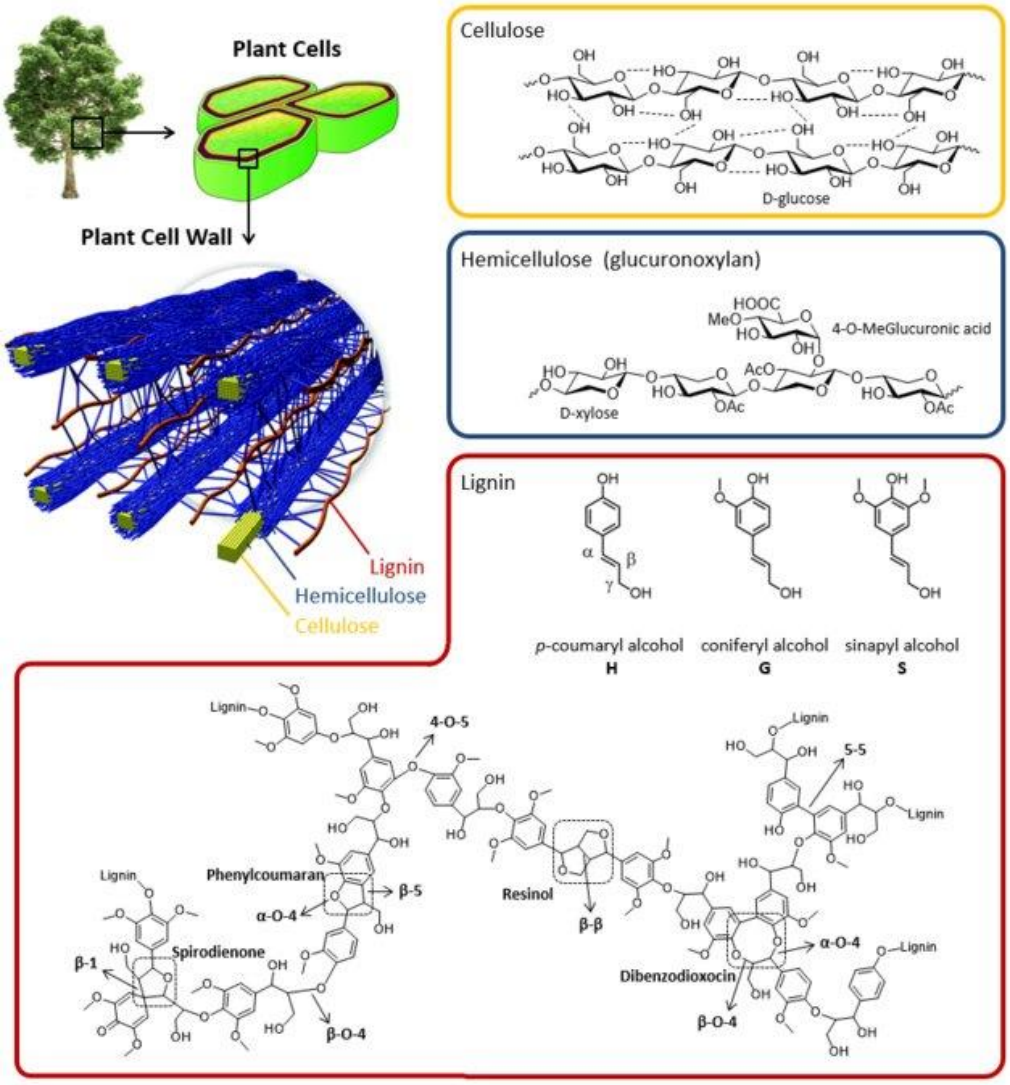


Figure 9 Schematic illustration of lignocellulose microfibril in plant cell wall and its free major polymeric components – cellulose, hemicellulose, and lignin. Adopted from Van den Bosch et al. (2018).[83]

4.1 *PSEUDOMONAS PUTIDA* IS AN ATTRACTIVE BACTERIAL HOST FOR LC BIOTECHNOLOGY

Pseudomonas putida KT2440 (class Gammaproteobacteria, family *Pseudomonadaceae*) is the best characterized and safe pseudomonad. This Gram-negative bacterial workhorse was derived from the saprophytic strain mt-2 which thrives in environments polluted with aromatic solvents. Being known for its robustness, *P. putida* KT2440 has been used

for decades as a model organism in the biodegradation and bioremediation field.[5] Quite recently, it has drawn the attention of the biotech community for its ability to degrade high molecular weight lignin with its ligninolytic enzymes and utilize lignin-derived aromatics (*p*-coumarate, ferulate, benzoate etc.) for the growth and production of bacterial bioplastics – polyhydroxyalkanoates.[75]

Strain KT2440 is a paradigmatic producer of cytoplasmic medium chain length polyhydroxyalkanoates (mcl-PHA). Out of all the studied bio-based plastics, fully biodegradable and biocompatible PHA represent the most promising alternative to their petrochemical analogs.[84] The mcl-PHA (6 - 14 C atoms) and their copolymers have better elastomeric properties and are more structurally diverse than short chain length PHA (scl-PHA; produced, e.g., by *Cupriavidus necator* or recombinant *E. coli*), which makes them promising materials for biomedicine, cosmetics, or the packaging industry. Their commercial production is still limited (was predicted to reach ~0.25 % of global plastic production in 2020) but could be boosted by the utilization of cheap non-edible carbon sources such as LC substrates. Pseudomonads are the best mcl-PHA producers among both natural and recombinant organisms. Engineered *P. putida* KT2440 was shown to convert glucose or lignin-born aromatics via acetyl-CoA and fatty acid de novo biosynthesis into high quantities (up to 67 and 54 % w/w_{DCW}, respectively) of mcl-PHA in nitrogen-limited medium.[85,86]

P. putida KT2440 has many other desirable characteristics for its use in LC biotechnology and in biotechnology enterprises in general.[87] It possesses HV1 safety certification, grows rapidly in defined minimal media (e.g., the growth rate in M9 medium with glucose is ~0.6 h^{-1}), and has overall low nutritional demand. KT2440 and its engineered derivatives show considerable resistance to oxidative stress and to LC-born inhibitory chemicals, many of which can even serve as growth substrates for this bacterium (besides already mentioned aromatics also acetate or furfural). Last but not least, *P. putida* has been proven suitable for large-scale aerobic fermentations and biotechnological production of diverse valuable chemicals from glucose or aromatics (besides PHA, e.g., *cis*,*cis*-muconic acid, rhamnolipids, terpenoids), it is amenable to genetic manipulations, and a broad palette of engineering tools has been already developed for this organism. These tools enabled the preparation of useful genome-reduced derivatives of KT2440 with altered physiology.

For instance, 11 non-adjacent genomic deletions in strain designated EM42 (300 genes including prophages, flagellum-encoding operons, endonucleases etc., ~4.3% of the whole genome) were shown to improve the expression of heterologous genes, increase intracellular ATP level, and enhance the host's genetic stability, viability, and thermal tolerance.[88] Another strain mentioned further in this section, EM371, lacks the majority of the non-essential outer membrane features.[89] Altogether 230 genes (~4.7% of the entire genome) that encode structures for bacterial motion (flagella), biofilm formation, or interaction with surroundings (e.g., pili, curli, fimbriae, exopolysaccharides, adhesins, lipopolysaccharides) were removed. *P. putida* EM371 with a "shaved" surface was proposed as a better strain for secretion of recombinant proteins.

Figure 10 Electron micrographs of *Pseudomonas putida* KT2440 and genome-reduced derivative strain EM42 which lacks polar flagella (in the cutout).

Nevertheless, *P. putida* KT2440 as a potential chassis for LC biotechnology has its drawbacks too. Out of all monomeric sugars present in LC hydrolysates, it utilizes only hexoses D-glucose and D-mannose, but not abundant pentoses D-xylose or L-arabinose. Though *P. putida*'s genome encodes 34 glycoside hydrolases including a couple of endoglucanases and a β-glucosidase (http://www.cazy.org/), it is not a typical cellulolytic bacterium and it does not grow on (hemi)cellulosic polymers or oligomers. During my postdoc stay in the laboratory of prof. Víctor de Lorenzo in CNB-CSIC Madrid I studied the carbohydrate metabolism of *P. putida* and tried to mitigate these drawbacks using therein developed genetic engineering tools. I now continue in this effort with my colleagues in the Microbial Bioengineering Laboratory recently established at Masaryk University.

4.2 EXPANDING THE SUBSTRATE SCOPE OF *P. PUTIDA* TOWARDS D-XYLOSE AND D-CELLOBIOSE

Out of LC sugars, cellulosic glucose is most frequently used for the biotechnological production of valuable chemicals.[73] It is obtained via the action of at least three enzymatic activities – endoglucanase, exoglucanase, and β-glucosidase - after hydrolysis of pretreated cellulose fraction of LC. Pentose sugars D-xylose and L-arabinose form a substantial portion of the hemicellulosic fraction of LC. D-xylose is together with D-glucose the most abundant sugar monomer in the majority of LC residues. For instance, it is the prime monomeric sugar in rice straw hydrolysates and forms ~1/3 of all sugar monomers in corn stover hydrolysates.[90,91] This makes xylose another very attractive LC-born carbon source for biotechnological purposes including the biomanufacturing of polyhydroxyalkanoates.[92] However, not all biotechnologically relevant bacteria can utilize xylose (e.g., *Z. mobilis*, *Corynebacterium glutamicum* do not possess catabolic pathways for xylose) and those that can (e.g., *E. coli*, *Bacillus subtilis*) often use glucose as a preferred carbon source due to the complex mechanisms of CCR. Also, *P. putida* KT2440 lacks biochemical traits for catabolism of pentoses and it was our goal during my postdoctoral stay in CNB-CSIC to establish xylose metabolism in this host. We presumed that *P. putida* with exogenous catabolic traits for xylose could co-utilize xylose with glucose because sugars do not play as central role in the nutrition of this bacterium as they do, e.g., in *E. coli*.

There are at least three major known catabolic pathways for xylose utilization in bacteria (Fig. 10).[93] Xylose can be assimilated via the isomerase pathway. It includes xylose isomerase (XylA) and xylulokinase (XylB) activities that funnel xylose to xylulose 5-phosphate and pentose phosphate pathway (PPP). Alternatively, xylose can be streamed to the tricarboxylic acid (TCA) cycle via α-ketoglutarate formed as a product of the oxidative Weimberg pathway. Another oxidative route – the Dahms pathway – converts xylose to pyruvate and glycolaldehyde via 2-keto-3-deoxy-D-xylonate. Each of these three routes was employed to engineer bacterial hosts with xylose negative metabolism including certain pseudomonads. In our work published in 2018, we described implanting of the xylose isomerase pathway from *E. coli* into *P. putida* KT2440 derivative, strain EM42.[94] We first identified that *P. putida* converts xylose non-productively into xylonic acid via the activity of its periplasmic glucose dehydrogenase Gcd. We deleted the

corresponding *gcd* gene using the method based on the I-*SceI* endonuclease mediated homologous recombination. Synthetic *xylABE* operon encoding XylA, XylB, and XylE xylose/H^+ symporter from *E. coli* was then introduced into *P. putida* Δ*gcd* mutant on low copy plasmid with strong constitutive promoter em7. The resulting recombinant grew in minimal medium with xylose (specific growth rate μ was 0.11 h^{-1}) used as a sole source of carbon and energy. A newly introduced transporter played a crucial role in the xylose utilization by engineered *P. putida*.

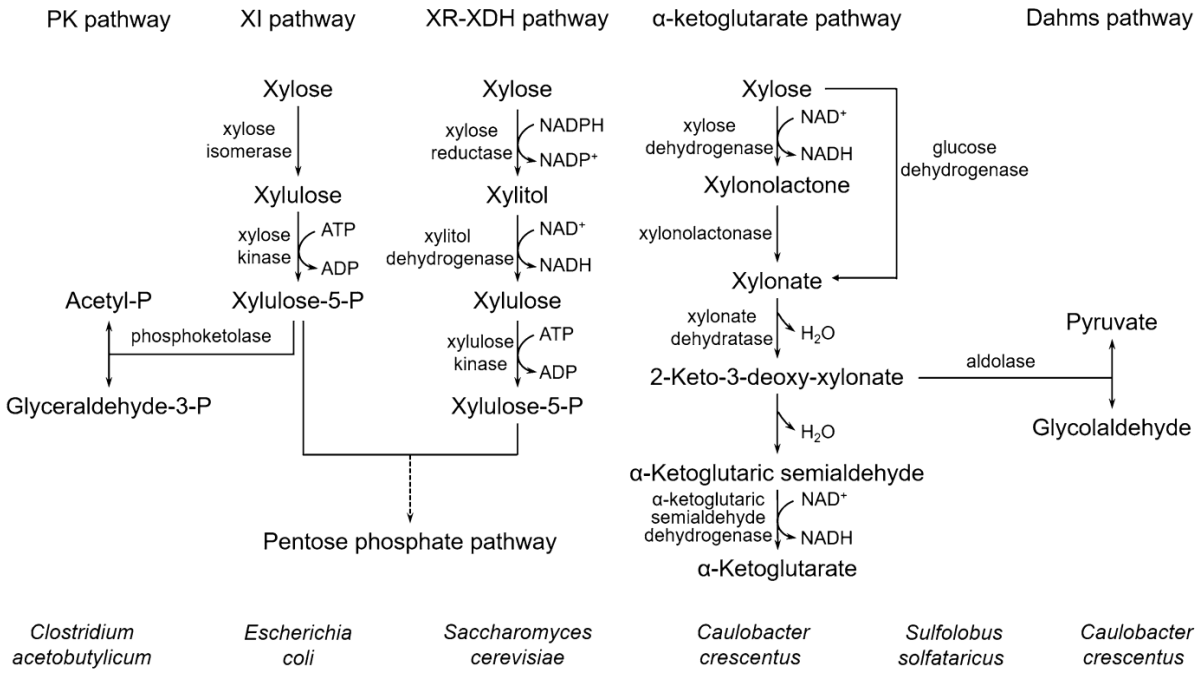


Figure 10 Microbial catabolic pathways for D-xylose. Note that α-ketoglutare pathway is also known as Weimberg pathway. Abbreviations: PK, phosphoketolase; XI, xylose isomerase; XR, xylose reductase; XDH, xylitol dehydrogenase. Adopted from Zhao et al. (2020).[93]

In the same study, we also enlarged the biochemical network of *P. putida* EM42 with exogenous β-glucosidase BglC from cellulolytic bacterium *Thermobifida fusca*. The gene of this cytoplasmic enzyme with em7 constitutive promoter was implanted into the chromosome of *P. putida* using the Tn5 minitransposon system. It enabled rapid growth (μ = 0.35 h^{-1}) of the resulting mutant in minimal medium with cellobiose – the shortest cellooligosaccharide and frequent side product of cellulose hydrolysis. For instance, in corn stover hydrolysate cellooligosaccharides form ~16 % of all saccharides.[91] Robust microbial hosts with established cellooligosaccharide metabolism are thus desirable. Moreover, as was proven also in our study, bacteria capable of active cellooligosaccharide

transport through the cell wall via their adenosine triphosphate (ATP)-dependent sugar transporters can serve energy because they can transfer two or more glucose at a cost of 1 mol ATP. Importantly, we demonstrated that engineered *P. putida* could indeed fully co-utilize xylose with cellobiose and glucose with no sign of CCR. Simultaneous utilization of multiple sugars and especially hexoses and pentoses is another highly desirable property of bacterial cell factories for LC biotechnology as it can reduce capital costs and operating expenses.[91] It allows parallel processing of glucose and xylose in a single pot also in fed-batch and continuous bioprocesses without accumulation of non-preferred sugar (usually xylose) which occurs in cultures with microorganisms tied by CCR.

Since our first attempt to engineer xylose metabolism in *P. putida*, more groups reported the implantation of pentose catabolic pathways in KT2440.[95,96] The most successful were Elmore and co-workers (2020) from National Renewable Energy Laboratory in USA.[91] They prepared a new KT2440 Δ*gcd* strain bearing *E. coli xylABE* genes in chromosome together with exogenous genes encoding transketolase (*tktA*) and transaldolase (*talB*) activities from the pentose phosphate pathway of *E. coli*. Growth of this strain on xylose was improved up to $\mu = 0.32\ h^{-1}$ by adaptive laboratory evolution in a defined minimal medium which enhanced the expression of *xylE* gene. In addition, another chromosomal insertion of recently discovered oxidative L-arabinose pathway from *Burkholderia ambifaria* AMMD (five genes) and arabinose/H+ symporter from E. coli promoted rapid growth ($\mu = 0.38\ h^{-1}$) of the strain on arabinose. Remarkably, KT2440 Δ*gcd* mutant with another deletion of *crc* gene (encodes a major regulator of CCR in *P. putida*) and two exogenous pentose catabolic routes was able to co-utilize glucose with xylose, arabinose, *p*-coumarate, and acetate in mock and real corn stover hydrolysates. This study for the first time presented *P. putida* as a bacterial platform for potential co-valorization of major carbon sources from all three fractions of LC. However, the practical utility of the obtained strain is limited by a long lag phase (≥14 h) observed in cultures with mixed carbon sources. And the conversion of utilized sugars and aromatics into a valuable product was not included in the study. There is therefore still plenty of space for further improvement of *P. putida*'s capacity to co-process multiple carbon sources and for the study of mechanisms that allow it. A lack of fundamental understanding of the co-metabolism of soluble LC-born substrates – pentoses, hexoses, and aromatic chemicals – forms a major knowledge gap that prevents the engineering of more efficient microbial cell factories for LC valorization. Understanding the co-function of *P. putida*'s innate and

implanted catabolic pathways will help to fill this gap and remove metabolic bottlenecks for carbon streaming into central metabolic intermediates such as pyruvate and acetyl-CoA and final bioproducts including PHA.

4.3 VALORIZATION OF LIGNOCELLULOSIC SUBSTRATES BY ENGINEERED *P. PUTIDA*

The main aim of LC biotechnology is to convert low-cost second-generation substrates from LC residues into higher value biochemicals. Establishing the efficient catabolism of LC sugars and aromatics is just the first step in this process. Wild-type or engineered *P. putida* KT2440 was applied in biotransformations or bioconversions that added value to pure glucose or aromatic chemicals (*p*-coumarate, ferulate, benzoate), as well as to LC hydrolysates or waste streams from LC pyrolysis.[85,91,97] However, reports on co-valorization of two or more LC substrates are rare. This situation is slowly changing. Simultaneous streaming of carbon and energy from several sources into one or more valuable compounds can improve the economics of bioprocesses that still cannot compete with the production of chemicals from fossil fuels.[78,98,99] The ideal scenario is if one of the products is intracellular and one extracellular, so they can be easily separated during downstream processing.

In the study that followed our achievement with the co-utilization of multiple sugars in engineered *P. putida* EM42, we employed the recombinant with inserted β-glucosidase BglC for the co-valorization of two new substrates – xylose and cellobiose.[**100**] Cellobiose was cleaved in the cytoplasm to glucose and funneled through glycolysis to acetyl-CoA and mcl-PHA while D-xylose was transformed by the activities of periplasmic Gcd and lactonase to D-xylonic acid which was released into the medium. In the previous work, xylonate was considered an undesired side product of xylose utilization via the isomerase pathway.[94] But because the growth of our EM42 recombinant on xylose was still too slow for bioproduction purposes, we decided to take advantage of the identified xylose oxidation. Xylonate is also recognized as a platform molecule of considerable biotechnological interest.[101] It can be used as a non-food derivative of gluconic acid, as a chelator, complexing agent, or a precursor of ethylene glycol and polyesters. Besides *P. putida* also some other biotechnologically relevant bacteria (e.g., *Gluconobacter oxydans*,

Klebsiella pneumoniae) including certain characterized pseudomonads such as *P. fragi* can naturally oxidize sugars with their glucose or xylose dehydrogenases.[102,103] These and several other microorganisms including engineered *Escherichia coli* with implanted xylose dehydrogenase were used for xylonate production from xylose.[101] But, to the best of our knowledge, xylonate was never co-produced with PHA.

Using shake flask cultures and a spectrum of analytical techniques (HPLC, chemical and enzymatic assays), we demonstrated that recombinant *P. putida* EM42 can transform xylose to xylonate with a high yield reaching 0.85 ± 0.06 g g^{-1} in optimized reaction conditions (buffered medium, high agitation). The biotransformation could be performed by both resting cells (these could be even recycled and used for xylonate production repeatedly) and cells growing on cellobiose. Cellobiose was shown to be a better substrate than monomeric glucose for growth-associated xylonate production because the disaccharide was not processed by glucose dehydrogenase and did not block the oxidation of xylose by this enzyme. PHA granules within the bacteria grown in minimal medium with 5 g l^{-1} cellobiose and 10 g l^{-1} xylose were visualized by confocal microscopy after staining the cells with Nile Red and then quantified in collected biomass using gas chromatography of the methanolysed polyester. The mcl-PHA content in cell dry weight (21 % w/w), yield (0.05 g g^{-1} cellobiose), titre (0.26 g l^{-1}), as well as monomer composition (3-hydroxydecanoate formed >75%) after 48 h culture were close to the values reported for *P. putida* KT2440 grown on glucose.[104] Xylonate yield in the same experiment reached 0.52 g g^{-1} xylose after 48 h and the maximum volumetric productivity was 156 mg l^{-1} h^{-1}.

The above-mentioned productivities must be further improved to approach the parameters required for the viable bioprocess. But already this pioneering study helped to map the potential of recombinant *P. putida* EM42 as a microbial platform for the parallel valorization of two second-generation LC substrates. It identified *P. putida* as a cell factory for high-yield biotransformation of xylose to xylonate and an alternative to established hosts for this reaction such as *Gluconobacter oxydans*. The production of mcl-PHA from cellooligosaccharide cellobiose was also reported for the first time in this work.

4.4 EMPOWERING *P. PUTIDA* WITH SURFACE-DISPLAYED DESIGNER PROTEIN SCAFFOLDS

Besides the rising interest in using *P. putida* in LC biotechnologies, it is now drawing much attention also as a potential microbial platform for the valorization of another type of organic polymeric waste – synthetic plastics.[105] Both these endeavors would make much use of an efficient molecular system that would deliver required de-polymerizing activities (cellulases, ligninolytic enzymes, PETase, MHETase) to the surface of the bacterium. *P. putida* lacks such apparatus and cannot degrade complex recalcitrant substrates without the externally added purified enzymes. A robust secretion system verified in *P. putida* could be adapted for many other applications including the display of antibodies, metal-binding proteins, or enzymes for bioremediation purposes, development of whole-cell electrochemical biosensor, or screening of libraries of antimicrobial peptides.

In the study published in 2020 in ACS Synthetic Biology, we tested four type V secretion pathway proteins known as autotransporters for surface display of small designer protein scaffolds that could bind multiple recombinant enzymes to the bacterial exterior (Fig. 11).**[106]** The binding apparatus was inspired by natural cellulosomes, complex enzymatic assemblies found on the surface of the most efficient cellulolytic bacteria such as *Clostridium thermocellum*. These organisms display so-called scaffoldin proteins with cohesin binding domains that define attachment of carbohydrate-active enzymes with corresponding dockerin domains to the cell surface.[107] The strong ($K_D \sim 10^{-9} – 10^{-10}$ M) highly specific non-covalent interactions between cohesins and dockerins enable orchestration of cellulases in natural cellulosomes which can degrade cellulose up to 50-fold more efficiently than free enzymes. Natural scaffoldins may consist of up to dozens of cohesins but such large structures (up to hundreds of kDa) are difficult to manipulate, not talking about their display on the surface of a heterologous microbial host. Small synthetic scaffoldins and whole designer cellulosomes (so-called minicellulosomes) were assembled on the surface of several domesticated Gram-positive bacteria such as *B. subtilis*, or *Lactococcus lactis*.[107,108] Display of designer scaffoldins on the cellular surface of a Gram-negative bacterium has never been achieved and there was a critical lack of knowledge on parameters that affect such an enterprise. We aimed to fill this knowledge gap by testing the display of several miniscaffoldins in *P. putida* strains EM42 and EM371.

These strains have substantial differences in the complexity of bacterial surface, which imposes structural constraints for recombinant protein secretion, and therefore were an outstanding study system for our work.

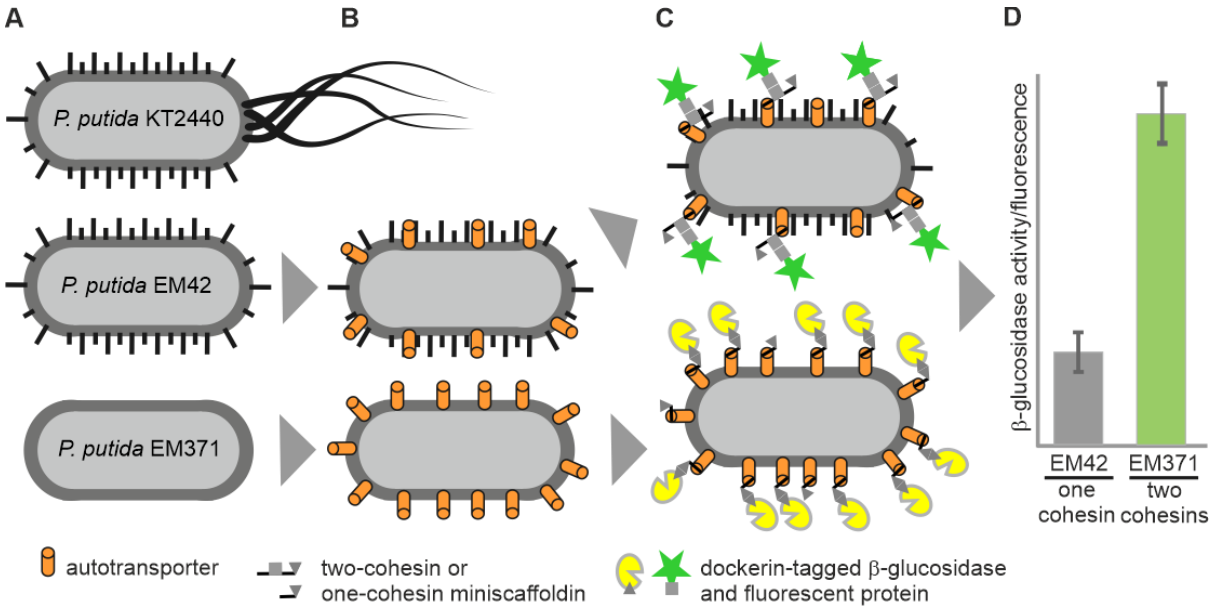


Figure 11 Schematic illustration of *Pseudomonas putida* engineering for the surface display of designer protein scaffoldins with cohesin binding domains. **A**) Strain EM42 and EM371, genome-reduced derivatives of *P. putida* KT2440, were used in this study (note that surface structures present in EM42 and absent in EM371 are shown as black lines). **B**) Ag43 autotransporter from *Escherichia coli* (orange columns) was selected out of four tested type V secretion systems for surface display of miniscaffoldins. **C**) Miniscaffoldins with single or two cohesins were displayed on bacterial surface via Ag43. **D**) The efficiency of the binding of dockerin-tagged recombinant proteins (β-glucosidase or fluorescent proteins) to the surface of EM42 and EM371 strain was evaluated and quantified. Adopted from Dvořák et al. (2020)[106] and modified.

We used overlap extension PCR to construct three miniscaffoldin variants (two of 16 and one of 35 kDa) with one or two cohesins from two different cellulolytic bacteria (*C. thermocellum* and *Acetivibrio cellulolyticus*) and chimeric β-glucosidase BglC and fluorescent proteins (green GFP and cyan CFP) tagged with corresponding dockerins. The latter three proteins were used as reporters for verification of scaffoldin display on *P. putida* surface. A battery of methods including enzyme-linked immunosorbent assays (ELISA), enzyme activity assays, SDS-PAGE, western blotting, and dot blots was adopted to verify and quantify expression of individual recombinant genes and to test the performance of resulting protein chimeras after their purification via 6xHis tag affinity chromatography. Scaffoldin variants were delivered to the surface of EM42 and EM371

by autotransporter Ag43 from *E. coli* which was selected as the most efficient among the four tested type V secretion systems.[109] Cells with exposed scaffoldins were incubated with purified dockerin-tagged BglC and the whole-cell β-glucosidase activity was determined to quantitatively evaluate the efficiency of the surface display. This in-house developed assay was supported by similar tests with purified dockerin-tagged fluorophores, measurements of whole-cell fluorescence in UV-VIS spectrophotometer, and observation of *P. putida* cells in a confocal microscope.

We could conclude that both single- and two-cohesin variants of designer scaffoldins were successfully displayed on the surface of EM42 and EM371. Shaved strain EM371 was, nonetheless, a better platform for the display of these synthetic structures. Single EM371 cell could attach up to 11,000 β-glucosidase molecules to its surface almost three times more than EM42. Our work thus highlighted the relevance of the cell surface engineering strategy for enhanced secretion of recombinant proteins in industrially relevant bacteria. It was also demonstrated that the size of displayed scaffolds is an important factor. Interestingly, bigger scaffoldins were better accessible on the bacterial surface for binding with dockerin-tagged molecules than the smaller ones.

P. putida is thus the first Gram-negative bacterium that can serve as a future carrier for cellulosome-like structures. We are currently working on the improvement of the available designer scaffoldins and our next goal is to establish *P. putida*-based whole-cell system for overproduction and release of dockerin-tagged depolymerizing enzymes.

4.5 PERSPECTIVES OF USING *P. PUTIDA* FOR THE VALORIZATION OF LC RESIDUES

The revealed ability of *P. putida* KT2440 and its derivatives to utilize and valorize multiple components of LC hydrolysates is a very appealing feature that attracts the attention of bioengineers and biotechnologists. Several studies from the last couple of years have aimed at the further broadening of *P. putida*'s substrate scope and at engineering and extending its biosynthetic pathways that could add value to the second-generation carbon sources. KT2440 was endowed with new biochemical traits for utilization of furfural and 5-(hydroxymethyl)furfural (major inhibitors in LC hydrolysates) or galactose.[110,111] It was evolved and rationally engineered towards

enhanced production of *cis*,*cis*-muconic acid from glucose and increased mcl-PHA synthesis from solubilized lignin, respectively.[85,112] Bator and colleagues calculated conversions of xylose to 14 valuable biochemicals (rhamnolipids, succinate, pyocyanin, glutamate etc.) using a genome-scale metabolic model of *P. putida* KT2440 with inserted reactions of Weimberg, Dahms, or isomerase pathway and evidenced that the last route provides the highest theoretical yields in case of 12 of these products.[95] Hemicellulosic pentoses are very attractive substrates also to produce polyhydroxyalkanoates which were, however, not included in the calculations.[92] The growth of our EM42 Δ*gcd* mutant with exogenous xylose isomerase pathway is still suboptimal for biotechnological purposes. We currently try to map catabolism of xylose in this strain by using theoretical and experimental analyses of carbon fluxes (flux balance analysis with genome-scale metabolic model of *P. putida* KT2440 and metabolic flux analysis with ^{13}C-labeled xylose, respectively) and we use the obtained information on possible metabolic constraints for its knowledge-driven improvement.

In the study of Espeso et al., an alternative evolutionary approach was adopted for the acceleration of growth of *P. putida* EM42 Δ*gcd* with *xylABE* operon delivered in the random spot in chromosome by Tn5 minitransposon-based system.[113] This strain with the isomerase pathway expressed from the chromosome grew substantially slower (μ=0.06 h^{-1}) than the strain with *xylABE* genes on the low-copy pSEVA2213 plasmid. The strain was evolved for 42 days in a dual-chamber semi-continuous log-phase bioreactor with an anti-biofilm layout. Despite the flawless performance of the developed do-it-yourself automated framework during the long time interval, the growth rate on xylose of *P. putida* EM42 Δ*gcd*::*xylABE* was improved only modestly (μ=0.09 h^{-1}). Bator and co-workers sub-cultured our strain with isomerase pathway on pSEVA2213 plasmid 30 times in minimal medium with xylose as a sole carbon source but did not observe any improvement of growth.[95] These experiments indicated that serious metabolic bottleneck(s) backpedaled the performance of EM42 recombinant on xylose. Our recent fluxomic analyses helped to shed a light on these bottlenecks and enabled the construction of new *P. putida* strains with an almost doubled growth rate on xylose (unpublished work).

All recombinant variants of *P. putida* KT2440 described in the aforementioned studies carried exogenous pathways expressed from strong constitutive promoters (e.g., P_{tac} or

P_{em7}). Such a setup is, however, not optimal because it forces the cells to constantly redirect their resources into the overproduction of foreign proteins. Dynamic gene expression response to the presence or absence of a 2G substrate is a tempting scenario that could reduce the metabolic burden from implanted modules and possibly further boost the performance of engineered *P. putida* variants. For instance, XylR- and AraC-dependent xylose- and arabinose-inducible promoters from *E. coli* could be adopted and trained for conditional expression of pentose catabolism modules in *P. putida* recombinants only in presence of hemicellulosic sugars in growth medium.[114] A similar system for conditional gene expression triggered by lignin-derived aromatics would be also highly attractive. We have recently updated a structural model of the A domain of *P. putida* XylR regulator which triggers transcriptional initiation in *xyl* operons for degradation of *m*-xylene or toluene upon binding of the aromatic effector.**[115]** The model was prepared by molecular threading using available structural data on related regulators. In/out distribution of residues was then validated by *in vivo* protein surface mapping with cross-linking reagent benzyl-bromide which reacted with accessible nucleophilic residues. The inducer-recognition pocket of the A domain, which is well defined in the new model, could be modified by targeted or random mutations to allow binding of lignin monomers such as *p*-coumarate, benzoate, guaiacol, or ferulate that are structurally related to natural XylR effectors.

The separation of engineered pathway modules among the members of a synthetic microbial community (single-species or multi-species) is another interesting strategy for the reduction of metabolic load which attracts the increasing attention of the bioengineering community.[116] This approach could lower the burden from the constitutive expression of 11 exogenous genes in *P. putida* KT2440 constructed for the co-utilization of glucose, xylose, arabinose, acetate, and lignin aromatics[91] and potentially shorten the long lag phase observed in the cultures with the single generalist strain. The designs of reliable robust artificial communities face many challenges, e.g., consortium instability in longer time intervals, the emergence of cheater strains, the imbalanced performance of individual consortium members.[116] But the adoption of this strategy specifically in LC biotechnologies that struggle with the feedstock complexity and recalcitrance makes lots of sense.[117]

Synthetic co-culture of several engineered strains could also help to meet the challenge of on-line supplementation of dockerin-tagged depolymerizing enzymes and *in vivo* assembly of cellulosome-like structures on *P. putida* surface.[118] In the envisioned cooperative single-species consortium, the autolytic strain overexpresses required enzymes and sacrifices itself to release and deliver biocatalysts to the scaffoldins on the surface of the production strain that can transform lignocellulosic polymers and oligomers directly to the target chemical. This biotechnology is being developed in our laboratory to complement our previous work and pave the way to new bacterial cell factories for the valorization of LC residues.

5. CONCLUSION

The introductory text of this thesis documents on the numerous referenced studies remarkable progress that has been recently achieved in the bioengineering of bacteria for biotechnological processing of organic waste compounds. There is no doubt that our understanding of microbial metabolism and physiology has reached new dimensions thanks to the available emerging technologies. The experience collected during the last two decades of protein and metabolic engineering and synthetic and systems biology paves the way to the new upcoming era of bioengineering-driven biodegradation and biosynthesis. However, the path to the "optimal" microbial cell factory that fulfils all the criteria for a practical application can be still crooked. Biodegradation and biotechnological valorization of organic waste compounds is a very attractive concept. But it still faces many hindrances on the levels of both strain and process engineering. The two model examples discussed herein – biodegradation and potential upcycling of toxic halogenated TCP via synthetic metabolic pathway assembled *in vitro* or in a living bacterium, and biotechnological valorization of lignocellulosic residues and derived substrates - represent extremely challenging enterprises. Our work shed more light on the studied systems and helped to remove some of the essential bottlenecks for the preparation of better enzymatic and bacterial biocatalysts for the given purposes.

The herein summarized studies on the designer TCP pathway have substantially contributed to the invasion of synthetic biology into the field of biodegradation and bioremediation of environmental pollutants. We have expanded the understanding of

biocatalysis of anthropogenic halogenated hydrocarbons and introduced several new concepts. We demonstrated that: (**i**) the combination of protein engineering and kinetic modeling of *in vitro* reconstructed biochemical cascade can greatly reduce the total enzyme load needed for the substrate conversion, (**ii**) mathematical model assembled from parameters collected *in vitro* (enzyme kinetics) and *in vivo* (substrate and metabolite toxicity, plasmid copy number) can be used for reliable computer-driven modular design of orthogonal biodegradation pathway in a suitable heterologous host, (**iii**) widely used synthetic inducer molecule IPTG can exacerbate substrate toxicity in popular bacterial host *E. coli* BL21(DE3), the use of natural inducer lactose can alleviate this effect, (**iv**) *in vitro* immobilized catabolic cascades can offer an alternative solution to microbial degraders for removal of toxic chemical waste.

By the same token, our research on *P. putida* enabled the expansion of its substrate scope towards abundant LC-derived pentoses and oligosaccharides and presented this organism as a suitable bacterial chassis for the parallel valorization of multiple LC substrates. The developed system for surface attachment of recombinant proteins will expand the utility of *P. putida* towards numerous new applications including facilitated degradation of organic polymeric waste. New validated structural model of XylR transcriptional regulator will be instrumental for the designs of molecular and whole-cell bacterial biosensors for aromatic chemicals.

I believe that these and other similar endeavors will lead to the next-generation biocatalysts that will overcome barriers for biotechnological upcycling of organic waste compounds and facilitate the transition to sustainable development and circular economy.

6. REFERENCES

(1) Bar-On, Y. M.; Phillips, R.; Milo, R. The Biomass Distribution on Earth. *Proc. Natl. Acad. Sci.* **2018**, *115* (25), 6506–6511. https://doi.org/10.1073/pnas.1711842115.

(2) Dombrowski, N.; Donaho, J. A.; Gutierrez, T.; Seitz, K. W.; Teske, A. P.; Baker, B. J. Reconstructing Metabolic Pathways of Hydrocarbon-Degrading Bacteria from the Deepwater Horizon Oil Spill. *Nat. Microbiol.* **2016**, *1* (7), 1–7. https://doi.org/10.1038/nmicrobiol.2016.57.

(3) Blank, L. M.; Narancic, T.; Mampel, J.; Tiso, T.; O'Connor, K. Biotechnological Upcycling of Plastic Waste and Other Non-Conventional Feedstocks in a Circular Economy. *Curr. Opin. Biotechnol.* **2020**, *62*, 212–219. https://doi.org/10.1016/j.copbio.2019.11.011.

(4) Arancon, R. A. D.; Lin, C. S. K.; Chan, K. M.; Kwan, T. H.; Luque, R. Advances on Waste Valorization: New Horizons for a More Sustainable Society. *Energy Sci. Eng.* **2013**, *1* (2), 53–71. https://doi.org/10.1002/ese3.9.

(5) Dvořák, P.; Nikel, P. I.; Damborský, J.; de Lorenzo, V. Bioremediation 3.0: Engineering Pollutant-Removing Bacteria in the Times of Systemic Biology. *Biotechnol. Adv.* 2017, *35* (7), 845–866. https://doi.org/10.1016/j.biotechadv.2017.08.001.

(6) Alexander, M. *Biodegradation and Bioremediation, Second Edition*, 2 edition.; Academic Press: San Diego, 1999.

(7) Litchfield, C. Thirty Years and Counting: Bioremediation in Its Prime? *BioScience* **2005**, *55* (3), 273–279. https://doi.org/10.1641/0006-3568(2005)055[0273:TYACBI]2.0.CO;2.

(8) Janssen, D. B.; Dinkla, I. J. T.; Poelarends, G. J.; Terpstra, P. Bacterial Degradation of Xenobiotic Compounds: Evolution and Distribution of Novel Enzyme Activities. *Environ. Microbiol.* **2005**, *7* (12), 1868–1882. https://doi.org/10.1111/j.1462-2920.2005.00966.x.

(9) Rojo, F.; Pieper, D. H.; Engesser, K. H.; Knackmuss, H. J.; Timmis, K. N. Assemblage of Ortho Cleavage Route for Simultaneous Degradation of Chloro- and Methylaromatics. *Science* **1987**, *238* (4832), 1395–1398.

(10) Haro, M. A.; de Lorenzo, V. Metabolic Engineering of Bacteria for Environmental Applications: Construction of *Pseudomonas* Strains for Biodegradation of 2-Chlorotoluene. *J. Biotechnol.* **2001**, *85* (2), 103–113.

(11) Bailey, J. E. Toward a Science of Metabolic Engineering. *Science* **1991**, *252* (5013), 1668–1675.

(12) Choi, K. R.; Jang, W. D.; Yang, D.; Cho, J. S.; Park, D.; Lee, S. Y. Systems Metabolic Engineering Strategies: Integrating Systems and Synthetic Biology with Metabolic Engineering. *Trends Biotechnol.* **2019**, *37* (8), 817–837. https://doi.org/10.1016/j.tibtech.2019.01.003.

(13) Stephanopoulos, G. Synthetic Biology and Metabolic Engineering. *ACS Synth. Biol.* **2012**, *1* (11), 514–525. https://doi.org/10.1021/sb300094q.

(14) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. Engineering the Third Wave of Biocatalysis. *Nature* **2012**, *485* (7397), 185–194. https://doi.org/10.1038/nature11117.

(15) Oberg, G. The Natural Chlorine Cycle--Fitting the Scattered Pieces. *Appl. Microbiol. Biotechnol.* **2002**, *58* (5), 565–581. https://doi.org/10.1007/s00253-001-0895-2.

(16) Kurihara, T.; Esaki, N. Bacterial Hydrolytic Dehalogenases and Related Enzymes: Occurrences, Reaction Mechanisms, and Applications. *Chem. Rec. N. Y. N* **2008**, *8* (2), 67–74. https://doi.org/10.1002/tcr.20141.

(17) Faroon, O.; Jones, D.; de Rosa, C. Effects of Polychlorinated Biphenyls on the Nervous System. *Toxicol. Ind. Health* **2000**, *16* (7–8), 305–333. https://doi.org/10.1177/074823370001600708.

(18) Crebelli, R.; Andreoli, C.; Carere, A.; Conti, L.; Crochi, B.; Cotta-Ramusino, M.; Benigni, R. Toxicology of Halogenated Aliphatic Hydrocarbons: Structural and Molecular

Determinants for the Disturbance of Chromosome Segregation and the Induction of Lipid Peroxidation. *Chem. Biol. Interact.* **1995**, *98* (2), 113–129.

(19) Agarwal, V.; Miles, Z. D.; Winter, J. M.; Eustáquio, A. S.; El Gamal, A. A.; Moore, B. S. Enzymatic Halogenation and Dehalogenation Reactions: Pervasive and Mechanistically Diverse. *Chem. Rev.* **2017**, *117* (8), 5619–5674. https://doi.org/10.1021/acs.chemrev.6b00571.

(20) Koudelakova, T.; Bidmanova, S.; Dvorak, P.; Pavelka, A.; Chaloupkova, R.; Prokop, Z.; Damborsky, J. Haloalkane Dehalogenases: Biotechnological Applications. *Biotechnol. J.* 2013, *8* (1), 32–45. https://doi.org/10.1002/biot.201100486.

(21) Schallmey, A.; Schallmey, M. Recent Advances on Halohydrin Dehalogenases—from Enzyme Identification to Novel Biocatalytic Applications. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 7827–7839. https://doi.org/10.1007/s00253-016-7750-y.

(22) Janssen, D. B. Evolving Haloalkane Dehalogenases. *Curr. Opin. Chem. Biol.* **2004**, *8* (2), 150–159. https://doi.org/10.1016/j.cbpa.2004.02.012.

(23) Emerging Contaminant − 1,2,3-Trichloropropane (TCP); Fact Sheet; U.S. Environmental Protection Agency: Washington, DC, 2009.

(24) Samin, G.; Janssen, D. B. Transformation and Biodegradation of 1,2,3-Trichloropropane (TCP). *Environ. Sci. Pollut. Res. Int.* **2012**, *19* (8), 3067–3078. https://doi.org/10.1007/s11356-012-0859-3.

(25) US EPA National Center for Environmental Assessment, I. O. IRIS Toxicological Review of 1,2,3-Trichloropropane (External Review Draft) https://cfpub.epa.gov/ncea/risk/recordisplay.cfm?deid=186355 (accessed 2021 -10 -04).

(26) *California Department of Public Health 1,2,3-Trichloropropane*; California Department of Public Health, 2009. https://www.waterboards.ca.gov/drinking_water/certlic/drinkingwater/123TCP.html (accessed 30 October 2021).

(27) Bowman, K. S.; Nobre, M. F.; da Costa, M. S.; Rainey, F. A.; Moe, W. M. Dehalogenimonas Alkenigignens Sp. Nov., a Chlorinated-Alkane-Dehalogenating Bacterium Isolated from Groundwater. *Int. J. Syst. Evol. Microbiol.* **2013**, *63* (Pt 4), 1492–1498. https://doi.org/10.1099/ijs.0.045054-0.

(28) Torralba-Sanchez, T. L.; Bylaska, E. J.; Salter-Blanc, A. J.; Meisenheimer, D. E.; Lyon, M. A.; Tratnyek, P. G. Reduction of 1,2,3-Trichloropropane (TCP): Pathways and Mechanisms from Computational Chemistry Calculations. *Environ. Sci. Process. Impacts* **2020**, *22* (3), 606–616. https://doi.org/10.1039/C9EM00557A.

(29) Dolfing, J.; Janssen, D. B. Estimates of Gibbs Free Energies of Formation of Chlorinated Aliphatic Compounds. *Biodegradation* **1994**, *5* (1), 21–28. https://doi.org/10.1007/BF00695210.

(30) Bosma, T.; Janssen, D. B. Conversion of Chlorinated Propanes by *Methylosinus trichosporium* OB3b Expressing Soluble Methane Monooxygenase. *Appl. Microbiol. Biotechnol.* **1998**, *50* (1), 105–112. https://doi.org/10.1007/s002530051263.

(31) Wang, B.; Chu, K.-H. Cometabolic Biodegradation of 1,2,3-Trichloropropane by Propane-Oxidizing Bacteria. *Chemosphere* **2017**, *168*, 1494–1497. https://doi.org/10.1016/j.chemosphere.2016.12.007.

(32) Bosma, T.; Kruizinga, E.; de Bruin, E. J.; Poelarends, G. J.; Janssen, D. B. Utilization of Trihalogenated Propanes by *Agrobacterium radiobacter* AD1 through Heterologous Expression of the Haloalkane Dehalogenase from *Rhodococcus sp.* Strain M15-3. *Appl. Environ. Microbiol.* **1999**, *65* (10), 4575–4581.

(33) Wijngaard, A. J. V. D.; Janssen, D. B.; Witholt, B. Degradation of Epichlorohydrin and Halohydrins by Bacterial Cultures Isolated from Freshwater Sediment. *J. Gen. Microbiol.* **1989**, *135* (8), 2199–2208. https://doi.org/10.1099/00221287-135-8-2199.

(34) Kulakova, A. N.; Larkin, M. J.; Kulakov, L. A. The Plasmid-Located Haloalkane Dehalogenase Gene from *Rhodococcus rhodochrous* NCIMB 13064. *Microbiol. Read. Engl.* **1997**, *143 (Pt 1)*, 109–115.

(35) Yang, F.; Hanna, M. A.; Sun, R. Value-Added Uses for Crude Glycerol--a Byproduct of Biodiesel Production. *Biotechnol. Biofuels* **2012**, *5*, 13. https://doi.org/10.1186/1754-6834-5-13.

(36) Bosma, T.; Damborský, J.; Stucki, G.; Janssen, D. B. Biodegradation of 1,2,3-Trichloropropane through Directed Evolution and Heterologous Expression of a Haloalkane Dehalogenase Gene. *Appl. Environ. Microbiol.* **2002**, *68* (7), 3582–3587.

(37) Nardini, M.; Ridder, I. S.; Rozeboom, H. J.; Kalk, K. H.; Rink, R.; Janssen, D. B.; Dijkstra, B. W. The X-Ray Structure of Epoxide Hydrolase from *Agrobacterium radiobacter* AD1. An Enzyme to Detoxify Harmful Epoxides. *J. Biol. Chem.* **1999**, *274* (21), 14579–14586.

(38) de Jong, R. M.; Tiesinga, J. J. W.; Rozeboom, H. J.; Kalk, K. H.; Tang, L.; Janssen, D. B.; Dijkstra, B. W. Structure and Mechanism of a Bacterial Haloalcohol Dehalogenase: A New Variation of the Short-Chain Dehydrogenase/Reductase Fold without an NAD(P)H Binding Site. *EMBO J.* **2003**, *22* (19), 4933–4944. https://doi.org/10.1093/emboj/cdg479.

(39) Tang, L.; Torres Pazmiño, D. E.; Fraaije, M. W.; de Jong, R. M.; Dijkstra, B. W.; Janssen, D. B. Improved Catalytic Properties of Halohydrin Dehalogenase by Modification of the Halide-Binding Site. *Biochemistry* **2005**, *44* (17), 6609–6618. https://doi.org/10.1021/bi047613z.

(40) Rui, L.; Cao, L.; Chen, W.; Reardon, K. F.; Wood, T. K. Protein Engineering of Epoxide Hydrolase from *Agrobacterium radiobacter* AD1 for Enhanced Activity and Enantioselective Production of (R)-1-Phenylethane-1,2-Diol. *Appl. Environ. Microbiol.* **2005**, *71* (7), 3995–4003. https://doi.org/10.1128/AEM.71.7.3995-4003.2005.

(41) Banáš, P.; Otyepka, M.; Jeřábek, P.; Petřek, M.; Damborský, J. Mechanism of Enhanced Conversion of 1,2,3-Trichloropropane by Mutant Haloalkane Dehalogenase Revealed by Molecular Modeling. *J. Comput. Aided Mol. Des.* **2006**, *20* (6), 375–383. https://doi.org/10.1007/s10822-006-9071-1.

(42) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. Redesigning Dehalogenase Access Tunnels as a Strategy for Degrading an Anthropogenic Substrate. *Nat. Chem. Biol.* **2009**, *5* (10), 727–733. https://doi.org/10.1038/nchembio.205.

(43) Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kuty, M.; Kuta-Smatanova, I.; Dohnalek, J.; Kulhanek, P.; Wade, R. C.; Damborsky, J. Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations. *J. Mol. Biol.* 2009, *392* (5), 1339–1356. https://doi.org/10.1016/j.jmb.2009.06.076.

(44) Kokkonen, P.; Bednar, D.; Pinto, G.; Prokop, Z.; Damborsky, J. Engineering Enzyme Access Tunnels. *Biotechnol. Adv.* **2019**, *37* (6), 107386. https://doi.org/10.1016/j.biotechadv.2019.04.008.

(45) van Leeuwen, J. G. E.; Wijma, H. J.; Floor, R. J.; van der Laan, J.-M.; Janssen, D. B. Directed Evolution Strategies for Enantiocomplementary Haloalkane Dehalogenases: From Chemical Waste to Enantiopure Building Blocks. *Chembiochem Eur. J. Chem. Biol.* **2012**, *13* (1), 137–148. https://doi.org/10.1002/cbic.201100579.

(46) Lim, H. J.; Kim, D.-M. Cell-Free Synthesis of Industrial Chemicals and Biofuels from Carbon Feedstocks. *Curr. Opin. Biotechnol.* **2022**, *73*, 158–163. https://doi.org/10.1016/j.copbio.2021.08.002.

(47) Bowie, J. U.; Sherkhanov, S.; Korman, T. P.; Valliere, M. A.; Opgenorth, P. H.; Liu, H. Synthetic Biochemistry: The Bio-Inspired Cell-Free Approach to Commodity Chemical Production. *Trends Biotechnol.* **2020**, *38* (7), 766–778. https://doi.org/10.1016/j.tibtech.2019.12.024.

(48) Dvorak, P.; Kurumbang, N. P.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Maximizing the Efficiency of Multienzyme Process by Stoichiometry Optimization. *Chembiochem Eur. J. Chem. Biol.* 2014, *15* (13), 1891–1895. https://doi.org/10.1002/cbic.201402265.

(49) Hameri, T.; Fengos, G.; Ataman, M.; Miskovic, L.; Hatzimanikatis, V. Kinetic Models of Metabolism That Consider Alternative Steady-State Solutions of Intracellular Fluxes and Concentrations. *Metab. Eng.* **2019**, *52*, 29–41. https://doi.org/10.1016/j.ymben.2018.10.005.

(50) Pontrelli, S.; Chiu, T.-Y.; Lan, E. I.; Chen, F. Y.-H.; Chang, P.; Liao, J. C. *Escherichia coli* as a Host for Metabolic Engineering. *Metab. Eng.* **2018**, *50*, 16–46. https://doi.org/10.1016/j.ymben.2018.04.008.

(51) Kurumbang, N. P.; Dvorak, P.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Computer-Assisted Engineering of the Synthetic Pathway for Biodegradation of a Toxic Persistent Pollutant. *ACS Synth. Biol.* 2013. https://doi.org/10.1021/sb400147n.

(52) Wu, G.; Yan, Q.; Jones, J. A.; Tang, Y. J.; Fong, S. S.; Koffas, M. A. G. Metabolic Burden: Cornerstones in Synthetic Biology and Metabolic Engineering Applications. *Trends Biotechnol.* **2016**, *34* (8), 652–664. https://doi.org/10.1016/j.tibtech.2016.02.010.

(53) Nicolaou, S. A.; Gaida, S. M.; Papoutsakis, E. T. A Comparative View of Metabolite and Substrate Stress and Tolerance in Microbial Bioprocessing: From Biofuels and Chemicals, to Biocatalysis and Bioremediation. *Metab. Eng.* **2010**, *12* (4), 307–331. https://doi.org/10.1016/j.ymben.2010.03.004.

(54) Dvorak, P.; Chrast, L.; Nikel, P. I.; Fedr, R.; Soucek, K.; Sedlackova, M.; Chaloupkova, R.; de Lorenzo, V.; Prokop, Z.; Damborsky, J. Exacerbation of Substrate Toxicity by IPTG in *Escherichia coli* BL21(DE3) Carrying a Synthetic Metabolic Pathway. *Microb. Cell Factories* 2015, *14*, 201. https://doi.org/10.1186/s12934-015-0393-3.

(55) Demko, M.; Chrást, L.; Dvořák, P.; Damborský, J.; Šafránek, D. Computational Modelling of Metabolic Burden and Substrate Toxicity in *Escherichia coli* Carrying a Synthetic Metabolic Pathway. *Microorganisms* 2019, *7* (11), 553. https://doi.org/10.3390/microorganisms7110553.

(56) Datta, S.; Christena, L. R.; Rajaram, Y. R. S. Enzyme Immobilization: An Overview on Techniques and Support Materials. *3 Biotech* **2013**, *3* (1), 1–9. https://doi.org/10.1007/s13205-012-0071-7.

(57) Dvorak, P.; Bidmanova, S.; Damborsky, J.; Prokop, Z. Immobilized Synthetic Pathway for Biodegradation of Toxic Recalcitrant Pollutant 1,2,3-Trichloropropane. *Environ. Sci. Technol.* 2014, *48* (12), 6859–6866. https://doi.org/10.1021/es500396r.

(58) Bidmanova, S.; Damborsky, J.; Prokop, Z. Immobilization of Haloalkane Dehalogenase LinB from *Sphingobium japonicum* UT26 for Biotechnological Applications. *J. Biocatal. Biotransformation* **2013**, *02* (01). https://doi.org/10.4172/2324-9099.1000106.

(59) Krasňan, V.; Stloukal, R.; Rosenberg, M.; Rebroš, M. Immobilization of Cells and Enzymes to LentiKats®. *Appl. Microbiol. Biotechnol.* **2016**, *100* (6), 2535–2553. https://doi.org/10.1007/s00253-016-7283-4.

(60) Sheldon, R. A.; van Pelt, S. Enzyme Immobilisation in Biocatalysis: Why, What and How. *Chem. Soc. Rev.* **2013**, *42* (15), 6223–6235. https://doi.org/10.1039/c3cs60075k.

(61) Stucki, Gerhard.; Thueer, Markus. Experiences of a Large-Scale Application of 1,2-Dichloroethane Degrading Microorganisms for Groundwater Treatment. *Environ. Sci. Technol.* **1995**, *29* (9), 2339–2345. https://doi.org/10.1021/es00009a028.

(62) Guo, C.; Chen, Y.; Zheng, Y.; Zhang, W.; Tao, Y.; Feng, J.; Tang, L. Exploring the Enantioselective Mechanism of Halohydrin Dehalogenase from *Agrobacterium radiobacter* AD1 by Iterative Saturation Mutagenesis. *Appl. Environ. Microbiol.* **2015**, *81* (8), 2919–2926. https://doi.org/10.1128/AEM.04153-14.

(63) Wang, X.; Han, S.; Yang, Z.; Tang, L. Improvement of the Thermostability and Activity of Halohydrin Dehalogenase from *Agrobacterium radiobacter* AD1 by Engineering C-Terminal Amino Acids. *J. Biotechnol.* **2015**, *212*, 92–98. https://doi.org/10.1016/j.jbiotec.2015.08.013.

(64) Zhang, X.-J.; Deng, H.-Z.; Liu, N.; Gong, Y.-C.; Liu, Z.-Q.; Zheng, Y.-G. Molecular Modification of a Halohydrin Dehalogenase for Kinetic Regulation to Synthesize Optically Pure (S)-

Epichlorohydrin. *Bioresour. Technol.* **2019**, *276*, 154–160. https://doi.org/10.1016/j.biortech.2018.12.103.

(65) Brezovsky, J.; Babkova, P.; Degtjarik, O.; Fortova, A.; Gora, A.; Iermak, I.; Rezacova, P.; Dvorak, P.; Smatanova, I. K.; Prokop, Z.; Chaloupkova, R.; Damborsky, J. Engineering a de Novo Transport Tunnel. *ACS Catal.* 2016, *6* (11), 7597–7610. https://doi.org/10.1021/acscatal.6b02081.

(66) Colin, P.-Y.; Kintses, B.; Gielen, F.; Miton, C. M.; Fischer, G.; Mohamed, M. F.; Hyvönen, M.; Morgavi, D. P.; Janssen, D. B.; Hollfelder, F. Ultrahigh-Throughput Discovery of Promiscuous Enzymes by Picodroplet Functional Metagenomics. *Nat. Commun.* **2015**, *6* (1), 10008. https://doi.org/10.1038/ncomms10008.

(67) Vanacek, P.; Sebestova, E.; Babkova, P.; Bidmanova, S.; Daniel, L.; Dvorak, P.; Stepankova, V.; Chaloupkova, R.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Exploration of Enzyme Diversity by Integrating Bioinformatics with Expression Analysis and Biochemical Characterization. *ACS Catal.* 2018, *8* (3), 2402–2412. https://doi.org/10.1021/acscatal.7b03523.

(68) Arif, M. I.; Samin, G.; van Leeuwen, J. G. E.; Oppentocht, J.; Janssen, D. B. Novel Dehalogenase Mechanism for 2,3-Dichloro-1-Propanol Utilization in *Pseudomonas putida* Strain MC4. *Appl. Environ. Microbiol.* **2012**, *78* (17), 6128–6136. https://doi.org/10.1128/AEM.00760-12.

(69) Samin, G.; Pavlova, M.; Arif, M. I.; Postema, C. P.; Damborsky, J.; Janssen, D. B. A *Pseudomonas putida* Strain Genetically Engineered for 1,2,3-Trichloropropane Bioremediation. *Appl. Environ. Microbiol.* **2014**, *80* (17), 5467–5476. https://doi.org/10.1128/AEM.01620-14.

(70) Searles, S.; Malins, C. (**2013**) ICCT White Paper. Availability of Cellulosic Residues and Wastes in the EU. [WWW document]. URL https://www.theicct.org/sites/default/files/publications/ICCT_EUcellulosic-waste-residues_20131022.pdf.

(71) Valdivia, M.; Galan, J. L.; Laffarga, J.; Ramos, J.-L. Biofuels 2020: Biorefineries Based on Lignocellulosic Materials. *Microb. Biotechnol.* **2016**, *9* (5), 585–594. https://doi.org/10.1111/1751-7915.12387.

(72) Isikgor, F. H.; Becer, C. R. Lignocellulosic Biomass: A Sustainable Platform for the Production of Bio-Based Chemicals and Polymers. *Polym. Chem.* **2015**, *6* (25), 4497–4559. https://doi.org/10.1039/C5PY00263J.

(73) Taha, M.; Foda, M.; Shahsavari, E.; Aburto-Medina, A.; Adetutu, E.; Ball, A. Commercial Feasibility of Lignocellulose Biodegradation: Possibilities and Challenges. *Curr. Opin. Biotechnol.* **2016**, *38*, 190–197. https://doi.org/10.1016/j.copbio.2016.02.012.

(74) Kawaguchi, H.; Hasunuma, T.; Ogino, C.; Kondo, A. Bioprocessing of Bio-Based Chemicals Produced from Lignocellulosic Feedstocks. *Curr. Opin. Biotechnol.* **2016**, *42*, 30–39. https://doi.org/10.1016/j.copbio.2016.02.031.

(75) Beckham, G. T.; Johnson, C. W.; Karp, E. M.; Salvachúa, D.; Vardon, D. R. Opportunities and Challenges in Biological Lignin Valorization. *Curr. Opin. Biotechnol.* **2016**, *42*, 40–53. https://doi.org/10.1016/j.copbio.2016.02.030.

(76) Parisutham, V.; Chandran, S.-P.; Mukhopadhyay, A.; Lee, S. K.; Keasling, J. D. Intracellular Cellobiose Metabolism and Its Applications in Lignocellulose-Based Biorefineries. *Bioresour. Technol.* **2017**, *239*, 496–506. https://doi.org/10.1016/j.biortech.2017.05.001.

(77) Lynd, L. R. The Grand Challenge of Cellulosic Biofuels. *Nat. Biotechnol.* **2017**, *35* (10), 912–915. https://doi.org/10.1038/nbt.3976.

(78) Baral, N. R.; Sundstrom, E. R.; Das, L.; Gladden, J.; Eudes, A.; Mortimer, J. C.; Singer, S. W.; Mukhopadhyay, A.; Scown, C. D. Approaches for More Efficient Biological Conversion of Lignocellulosic Feedstocks to Biofuels and Bioproducts. *ACS Sustain. Chem. Eng.* **2019**, *7* (10), 9062–9079. https://doi.org/10.1021/acssuschemeng.9b01229.

(79) Liu, Y.-J.; Li, B.; Feng, Y.; Cui, Q. Consolidated Bio-Saccharification: Leading Lignocellulose Bioconversion into the Real World. *Biotechnol. Adv.* **2020**, *40*, 107535. https://doi.org/10.1016/j.biotechadv.2020.107535.

(80) Bokinsky, G.; Peralta-Yahya, P. P.; George, A.; Holmes, B. M.; Steen, E. J.; Dietrich, J.; Lee, T. S.; Tullman-Ercek, D.; Voigt, C. A.; Simmons, B. A.; Keasling, J. D. Synthesis of Three Advanced Biofuels from Ionic Liquid-Pretreated Switchgrass Using Engineered *Escherichia coli. Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108* (50), 19949–19954. https://doi.org/10.1073/pnas.1106958108.

(81) Anandharaj, M.; Lin, Y.-J.; Rani, R. P.; Nadendla, E. K.; Ho, M.-C.; Huang, C.-C.; Cheng, J.-F.; Chang, J.-J.; Li, W.-H. Constructing a Yeast to Express the Largest Cellulosome Complex on the Cell Surface. *Proc. Natl. Acad. Sci.* **2020**, *117* (5), 2385–2394. https://doi.org/10.1073/pnas.1916529117.

(82) Camarero, S.; Martínez, M. J.; Martínez, A. T. Understanding Lignin Biodegradation for the Improved Utilization of Plant Biomass in Modern Biorefineries. *Biofuels Bioprod. Biorefining* **2014**, *8* (5), 615–625. https://doi.org/10.1002/bbb.1467.

(83) Van den Bosch, S.; Koelewijn, S.-F.; Renders, T.; Van den Bossche, G.; Vangeel, T.; Schutyser, W.; Sels, B. F. Catalytic Strategies Towards Lignin-Derived Chemicals. *Top. Curr. Chem.* **2018**, *376* (5), 36. https://doi.org/10.1007/s41061-018-0214-3.

(84) Choi, S. Y.; Cho, I. J.; Lee, Y.; Kim, Y.-J.; Kim, K.-J.; Lee, S. Y. Microbial Polyhydroxyalkanoates and Nonnatural Polyesters. *Adv. Mater.* **2020**, *32* (35), 1907138. https://doi.org/10.1002/adma.201907138.

(85) Salvachúa, D.; Rydzak, T.; Auwae, R.; De Capite, A.; Black, B. A.; Bouvier, J. T.; Cleveland, N. S.; Elmore, J. R.; Huenemann, J. D.; Katahira, R.; Michener, W. E.; Peterson, D. J.; Rohrer, H.; Vardon, D. R.; Beckham, G. T.; Guss, A. M. Metabolic Engineering of *Pseudomonas putida* for Increased Polyhydroxyalkanoate Production from Lignin. *Microb. Biotechnol.* **2020**, *13* (1), 290–298. https://doi.org/10.1111/1751-7915.13481.

(86) Poblete-Castro, I.; Rodriguez, A. L.; Lam, C. M. C.; Kessler, W. Improved Production of Medium-Chain-Length Polyhydroxyalkanoates in Glucose-Based Fed-Batch Cultivations of Metabolically Engineered *Pseudomonas putida* Strains. *J. Microbiol. Biotechnol.* **2014**, *24* (1), 59–69. https://doi.org/10.4014/jmb.1308.08052.

(87) Anna Weimer; Kohlstedt, M.; Volke, D. C.; Nikel, P. I.; Wittmann, C. Industrial Biotechnology of *Pseudomonas putida*: Advances and Prospects. *Appl. Microbiol. Biotechnol.* **2020**, *104* (18), 7745–7766. https://doi.org/10.1007/s00253-020-10811-9.

(88) Martínez-García, E.; Nikel, P. I.; Aparicio, T.; de Lorenzo, V. *Pseudomonas* 2.0: Genetic Upgrading of *P. putida* KT2440 as an Enhanced Host for Heterologous Gene Expression. *Microb. Cell Factories* **2014**, *13*, 159. https://doi.org/10.1186/s12934-014-0159-3.

(89) Martínez-García, E.; Fraile, S.; Rodríguez Espeso, D.; Vecchietti, D.; Bertoni, G.; de Lorenzo, V. Naked Bacterium: Emerging Properties of a Surfome-Streamlined *Pseudomonas putida* Strain. *ACS Synth. Biol.* **2020**, *9* (9), 2477–2492. https://doi.org/10.1021/acssynbio.0c00272.

(90) Obruca, S.; Benesová, P.; Maršálek, L.; Márova, I. Use of Lignocellulosic Materials for PHA Production. Chemical and Biochemical Engineering Quarterly [online]. Croatian Society Of Chemical Engineers, **2015**, *29*(2), 135-144. https://doi.org/10.15255/CABEQ.2014.2253.

(91) Elmore, J. R.; Dexter, G. N.; Salvachúa, D.; O'Brien, M.; Klingeman, D. M.; Gorday, K.; Michener, J. K.; Peterson, D. J.; Beckham, G. T.; Guss, A. M. Engineered *Pseudomonas putida* Simultaneously Catabolizes Five Major Components of Corn Stover Lignocellulose: Glucose, Xylose, Arabinose, p-Coumaric Acid, and Acetic Acid. *Metab. Eng.* **2020**, *62*, 62–71. https://doi.org/10.1016/j.ymben.2020.08.001.

(92) Dietrich, K.; Dumont, M.-J.; Del Rio, L. F.; Orsat, V. Sustainable PHA Production in Integrated Lignocellulose Biorefineries. *New Biotechnol.* **2019**, *49*, 161–168. https://doi.org/10.1016/j.nbt.2018.11.004.

(93) Zhao, Z.; Xian, M.; Liu, M.; Zhao, G. Biochemical Routes for Uptake and Conversion of Xylose by Microorganisms. *Biotechnol. Biofuels* **2020**, *13* (1), 21. https://doi.org/10.1186/s13068-020-1662-x.

(94) Dvořák, P.; de Lorenzo, V. Refactoring the Upper Sugar Metabolism of *Pseudomonas putida* for Co-Utilization of Cellobiose, Xylose, and Glucose. *Metab. Eng.* 2018, *48*, 94–108. https://doi.org/10.1016/j.ymben.2018.05.019.

(95) Bator, I.; Wittgens, A.; Rosenau, F.; Tiso, T.; Blank, L. M. Comparison of Three Xylose Pathways in *Pseudomonas putida* KT2440 for the Synthesis of Valuable Products. *Front. Bioeng. Biotechnol.* **2020**, *7*, 480. https://doi.org/10.3389/fbioe.2019.00480.

(96) Lim, H. G.; Eng, T.; Banerjee, D.; Alarcon, G.; Lau, A. K.; Park, M.-R.; Simmons, B. A.; Palsson, B. O.; Singer, S. W.; Mukhopadhyay, A.; Feist, A. M. Generation of *Pseudomonas putida* KT2440 Strains with Efficient Utilization of Xylose and Galactose via Adaptive Laboratory Evolution. *ACS Sustain. Chem. Eng.* **2021**, *9* (34), 11512–11523. https://doi.org/10.1021/acssuschemeng.1c03765.

(97) Jayakody, L. N.; Johnson, C. W.; Whitham, J. M.; Giannone, R. J.; Black, B. A.; Cleveland, N. S.; Klingeman, D. M.; Michener, W. E.; Olstad, J. L.; Vardon, D. R.; Brown, R. C.; Brown, S. D.; Hettich, R. L.; Guss, A. M.; Beckham, G. T. Thermochemical Wastewater Valorization via Enhanced Microbial Toxicity Tolerance. *Energy Environ. Sci.* **2018**, *11* (6), 1625–1638. https://doi.org/10.1039/C8EE00460A.

(98) Larroude, M.; Celinska, E.; Back, A.; Thomas, S.; Nicaud, J.-M.; Ledesma-Amaro, R. A Synthetic Biology Approach to Transform *Yarrowia lipolytica* into a Competitive Biotechnological Producer of β-Carotene. *Biotechnol. Bioeng.* **2018**, *115* (2), 464–472. https://doi.org/10.1002/bit.26473.

(99) Wang, Y.; Ling, C.; Chen, Y.; Jiang, X.; Chen, G.-Q. Microbial Engineering for Easy Downstream Processing. *Biotechnol. Adv.* **2019**, *37* (6), 107365. https://doi.org/10.1016/j.biotechadv.2019.03.004.

(100) Dvořák, P.; Kováč, J.; de Lorenzo, V. Biotransformation of D-Xylose to D-Xylonate Coupled to Medium-Chain-Length Polyhydroxyalkanoate Production in Cellobiose-Grown *Pseudomonas putida* EM42. *Microb. Biotechnol.* 2020, *13* (4), 1273-1283. https://doi.org/10.1111/1751-7915.13574.

(101) Toivari, M. H.; Nygård, Y.; Penttilä, M.; Ruohonen, L.; Wiebe, M. G. Microbial D-Xylonate Production. *Appl. Microbiol. Biotechnol.* **2012**, *96* (1), 1–8. https://doi.org/10.1007/s00253-012-4288-5.

(102) Buchert, J.; Viikari, L. The Role of Xylonolactone in Xylonic Acid Production by *Pseudomonas fragi*. *Appl. Microbiol. Biotechnol.* **1988**, *27* (4), 333–336. https://doi.org/10.1007/BF00251763.

(103) Meijnen, J.-P.; de Winde, J. H.; Ruijssenaars, H. J. Engineering *Pseudomonas putida* S12 for Efficient Utilization of D-Xylose and L-Arabinose. *Appl. Environ. Microbiol.* **2008**, *74* (16), 5031–5037. https://doi.org/10.1128/AEM.00924-08.

(104) Poblete-Castro, I.; Binger, D.; Rodrigues, A.; Becker, J.; Martins Dos Santos, V. A. P.; Wittmann, C. In-Silico-Driven Metabolic Engineering of *Pseudomonas putida* for Enhanced Production of Poly-Hydroxyalkanoates. *Metab. Eng.* **2013**, *15*, 113–123. https://doi.org/10.1016/j.ymben.2012.10.004.

(105) Werner, A. Z.; Clare, R.; Mand, T. D.; Pardo, I.; Ramirez, K. J.; Haugen, S. J.; Bratti, F.; Dexter, G. N.; Elmore, J. R.; Huenemann, J. D.; Peabody, G. L.; Johnson, C. W.; Rorrer, N. A.; Salvachúa, D.; Guss, A. M.; Beckham, G. T. Tandem Chemical Deconstruction and Biological Upcycling of Poly(Ethylene Terephthalate) to β-Ketoadipic Acid by *Pseudomonas putida* KT2440. *Metab. Eng.* **2021**, *67*, 250–261. https://doi.org/10.1016/j.ymben.2021.07.005.

(106) Dvořák, P.; Bayer, E. A.; de Lorenzo, V. Surface Display of Designer Protein Scaffolds on Genome-Reduced Strains of *Pseudomonas putida*. *ACS Synth. Biol.* 2020, *9* (10), 2749–2764. https://doi.org/10.1021/acssynbio.0c00276.

(107) Artzi, L.; Bayer, E. A.; Moraïs, S. Cellulosomes: Bacterial Nanomachines for Dismantling Plant Polysaccharides. *Nat. Rev. Microbiol.* **2017**, *15* (2), 83–95. https://doi.org/10.1038/nrmicro.2016.164.

(108) Wieczorek, A. S.; Martin, V. J. J. Effects of Synthetic Cohesin-Containing Scaffold Protein Architecture on Binding Dockerin-Enzyme Fusions on the Surface of *Lactococcus lactis*. *Microb. Cell Factories* **2012**, *11*, 160. https://doi.org/10.1186/1475-2859-11-160.

(109) van der Woude, M. W.; Henderson, I. R. Regulation and Function of Ag43 (Flu). *Annu. Rev. Microbiol.* **2008**, *62*, 153–169. https://doi.org/10.1146/annurev.micro.62.081307.162938.
(110) Peabody, G. L.; Elmore, J. R.; Martinez-Baird, J.; Guss, A. M. Engineered *Pseudomonas putida* KT2440 Co-Utilizes Galactose and Glucose. *Biotechnol. Biofuels* **2019**, *12* (1), 295. https://doi.org/10.1186/s13068-019-1627-0.
(111) Guarnieri, M. T.; Ann Franden, M.; Johnson, C. W.; Beckham, G. T. Conversion and Assimilation of Furfural and 5-(Hydroxymethyl)Furfural by *Pseudomonas putida* KT2440. *Metab. Eng. Commun.* **2017**, *4*, 22–28. https://doi.org/10.1016/j.meteno.2017.02.001.
(112) Bentley, G. J.; Narayanan, N.; Jha, R. K.; Salvachúa, D.; Elmore, J. R.; Peabody, G. L.; Black, B. A.; Ramirez, K.; De Capite, A.; Michener, W. E.; Werner, A. Z.; Klingeman, D. M.; Schindel, H. S.; Nelson, R.; Foust, L.; Guss, A. M.; Dale, T.; Johnson, C. W.; Beckham, G. T. Engineering Glucose Metabolism for Enhanced Muconic Acid Production in *Pseudomonas putida* KT2440. *Metab. Eng.* **2020**, *59*, 64–75. https://doi.org/10.1016/j.ymben.2020.01.001.
(113) Espeso, D. R.; Dvořák, P.; Aparicio, T.; de Lorenzo, V. An Automated DIY Framework for Experimental Evolution of *Pseudomonas putida*. *Microb. Biotechnol.* 2020. *14* (6), 2679-2685. https://doi.org/10.1111/1751-7915.13678.
(114) Tang, R.-Q.; Wagner, J. M.; Alper, H. S.; Zhao, X.-Q.; Bai, F.-W. Design, Evolution, and Characterization of a Xylose Biosensor in *Escherichia coli* Using the XylR/XylO System with an Expanded Operating Range. *ACS Synth. Biol.* **2020**, *9* (10), 2714–2722. https://doi.org/10.1021/acssynbio.0c00225.
(115) Dvořák, P.; Alvarez-Carreño, C.; Ciordia, S.; Paradela, A.; de Lorenzo, V. An Updated Structural Model of the A Domain of the *Pseudomonas putida* XylR Regulator Poses an Atypical Interplay with Aromatic Effectors. *Environ. Microbiol.* 2021, *23* (8), 4418–4433. https://doi.org/10.1111/1462-2920.15628.
(116) Wang, R.; Zhao, S.; Wang, Z.; Koffas, M. A. Recent Advances in Modular Co-Culture Engineering for Synthesis of Natural Products. *Curr. Opin. Biotechnol.* **2020**, *62*, 65–71. https://doi.org/10.1016/j.copbio.2019.09.004.
(117) Jiang, Y.; Dong, W.; Xin, F.; Jiang, M. Designing Synthetic Microbial Consortia for Biofuel Production. *Trends Biotechnol.* **2020**, *38* (8), 828–831. https://doi.org/10.1016/j.tibtech.2020.02.002.
(118) Honjo, H.; Iwasaki, K.; Soma, Y.; Tsuruno, K.; Hamada, H.; Hanai, T. Synthetic Microbial Consortium with Specific Roles Designated by Genetic Circuits for Cooperative Chemical Production. *Metab. Eng.* **2019**, *55*, 268–275. https://doi.org/10.1016/j.ymben.2019.08.007.

7. ARTICLES

Articles from Section I: Engineering bacterial enzymes and biochemical pathways for biodegradation of anthropogenic halogenated waste compounds

[1] Dvořák, P., Nikel, P.I., Damborský, J., and de Lorenzo, V. (2017) Bioremediation 3.0: Engineering pollutant-removing bacteria in the times of systemic biology. *Biotechnology Advances*. *35*, 845-866. (2017 IF = 12.451)

[2] Koudelakova, T., Bidmanova, T., **Dvorak, P.**, Pavelka, A., Chaloupkova, R., Prokop, Z., and Damborsky, J. (2013) Haloalkane dehalogenases: Biotechnological applications. *Biotechnology Journal*. *8*, 32-45. (2013 IF = 3.237).

[3] Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., **Dvorak, P.**, Prokop, Z., Stsiapanava, A., Kuty, M., Kuta-Smatanova, I., Dohnalek, J., Kulhanek, P., Wade, R.C., and Damborsky, J. (2009) Pathways and mechanisms for product release in the engineered haloalkane dehalogenases explored using classical and random acceleration molecular dynamics simulations. *Journal of Molecular Biology*. *392*, 1339-1356. (2009 IF = 4.031)

[4] Dvorak, P., Kurumbang, N.P., Bendl, J., Brezovsky, J., Prokop, Z., and Damborsky, J. (2014) Maximizing the efficiency of multi-enzyme processes by stoichiometry optimization. *ChemBioChem*. *15*, 1891-1895. (2014 IF = 3.157)

[5] Kurumbang, N.P.*, **Dvorak, P.***, Bendl, J., Brezovsky, J., Prokop, Z., and Damborsky, J. (2014) Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. *ACS Synthetic Biology*. *3*, 172-181. (*shared first author, 2014 IF = 4.433)

[6] Dvorak, P., Chrast, L., Nikel. P.I., Fedr, R., Soucek, K., Chaloupkova, R., de Lorenzo, V., Prokop, Z., and Damborsky, J. (2015) Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. *Microbial Cell Factories*. 14, 201. (2015 IF = 4.151)

[7] Demko, M., Chrást, L., **Dvořák, P.**, Damborský, J., and Šafránek, D. (2019) Computational modelling of metabolic burden and substrate toxicity in *Escherichia coli* carrying a synthetic metabolic pathway. *Microorganisms*. *7*, 553. (2019 IF = 3.864)

[8] Dvorak, P., Bidmanova, S., Prokop, Z., and Damborsky, J. (2014) Immobilized synthetic pathway for biodegradation of toxic recalcitrant pollutant 1,2,3-trichloropropane. *Environmental Science and Technology*. *48*, 6859–6866. (2014 IF = 5.478)

[9] Brezovsky, J., Babkova, P., Degtjarik, O., Fortova, A., Gora, A., Iermak, I., Rezacova, P., **Dvorak, P.**, Kuta Smatanova, I., Prokop, Z., Chaloupkova, R., and Damborsky, J. (2016) Engineering a de novo transport tunnel. *ACS Catalysis*. *6*, 7597-7610. (2016 IF = 10.720)

[10] Vanacek, P., Sebestova, E., Babkova, P., Bidmanova, S., Daniel, L., **Dvorak, P.**, Stepankova, V., Chaloupkova, R., Brezovsky, J., Prokop, Z., and Damborsky, J. (2018) Exploration of enzyme diversity by integrating bioinformatics with expression analysis and biochemical characterization. *ACS Catalysis*. *8*, 2402–2412. (2018 IF = 12.025)

Biotechnology Advances 35 (2017) 845–866

Contents lists available at ScienceDirect

Biotechnology Advances

journal homepage: www.elsevier.com/locate/biotechadv

Research review paper

Bioremediation 3.0: Engineering pollutant-removing bacteria in the times of systemic biology

Pavel Dvořák[a], Pablo I. Nikel[b], Jiří Damborský[c,d], Víctor de Lorenzo[a,*]

[a] *Systems and Synthetic Biology Program, Centro Nacional de Biotecnología (CNB-CSIC), 28049 Madrid, Spain*
[b] *The Novo Nordisk Foundation Center for Biosustainability, 2800 Lyngby, Denmark*
[c] *Loschmidt Laboratories, Centre for Toxic Compounds in the Environment RECETOX, Department of Experimental Biology, Faculty of Science, Masaryk University, 62500 Brno, Czech Republic*
[d] *International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, 656 91 Brno, Czech Republic*

ARTICLE INFO

Keywords:
Bioremediation
Biodegradation pathway engineering
Emerging pollutants
Environmental biotechnology
Systemic biology
Metabolic engineering
Systems biology
Synthetic biology

ABSTRACT

Elimination or mitigation of the toxic effects of chemical waste released to the environment by industrial and urban activities relies largely on the catalytic activities of microorganisms—specifically bacteria. Given their capacity to evolve rapidly, they have the biochemical power to tackle a large number of molecules mobilized from their geological repositories through human action (*e.g.*, hydrocarbons, heavy metals) or generated through chemical synthesis (*e.g.*, xenobiotic compounds). Whereas naturally occurring microbes already have considerable ability to remove many environmental pollutants with no external intervention, the onset of genetic engineering in the 1980s allowed the possibility of rational design of bacteria to catabolize specific compounds, which could eventually be released into the environment as bioremediation agents. The complexity of this endeavour and the lack of fundamental knowledge nonetheless led to the virtual abandonment of such a recombinant DNA-based bioremediation only a decade later. In a twist of events, the last few years have witnessed the emergence of new *systemic* fields (including systems and synthetic biology, and metabolic engineering) that allow revisiting the same environmental pollution challenges through fresh and far more powerful approaches. The focus on contaminated sites and chemicals has been broadened by the phenomenal problems of anthropogenic emissions of greenhouse gases and the accumulation of plastic waste on a global scale. In this article, we analyze how contemporary systemic biology is helping to take the design of bioremediation agents back to the core of environmental biotechnology. We inspect a number of recent strategies for catabolic pathway construction and optimization and we bring them together by proposing an engineering workflow.

1. Introduction

Increasing pollution of air, soils, ground and surface waters constitutes a major threat to public health both in developing countries as well as in industrialized countries including EU states, the USA, India and China. The majority of contaminants that affect soils and waters are heavy metals and organic compounds such as mineral oil hydrocarbons, polyaromatic hydrocarbons, benzene derivatives, and halogenated hydrocarbons. Many of organic polluting compounds for agricultural (the pesticides dichlorodiphenyltrichloroethane, atrazine, and pentachlorophenol), industrial (solvents such as dichloroethane or dielectric fluids such as polychlorinated biphenyls) or military use (explosives such as 2,4,6-trinitrotoluene) are xenobiotics of anthropogenic origin. There is also a spectrum of so-called *emerging contaminants* (Table 1), *i.e.*, substances long present in the environment whose presence and negative effects have only recently been recognized (Petrie et al., 2015). The list can be further broadened with petroleum-derived plastics and some chemicals originally considered to be *green*, including certain types of bioplastics or ionic liquids (Amde et al., 2015). Despite the recalcitrant nature of some of these polluting compounds, many are more or less susceptible to biodegradation (Alexander, 1999). In addition to these traditional causes of environmental deterioration, the recent decades have witnessed the onset of ramped-up levels of anthropogenic emissions of CO_2 and other greenhouse gases and their ensuing impact on climatic change. Whereas the chemicals themselves are simple (CO_2, CH_4, N_2O), the challenge here is less their biodegradation than their recapture in a non-gaseous form.

The major entity that causes large-scale transformations in the biosphere are microorganisms and their metabolic pathways. Microbes degrade toxic chemicals *via* complete mineralization or co-metabolism,

* Corresponding author.
E-mail address: vdlorenzo@cnb.csic.es (V. de Lorenzo).

http://dx.doi.org/10.1016/j.biotechadv.2017.08.001
Received 4 June 2017; Received in revised form 1 August 2017; Accepted 4 August 2017
Available online 05 August 2017
0734-9750/ © 2017 Elsevier Inc. All rights reserved.

Table 1
Emerging contaminants.
Sources: http://toxics.usgs.gov, http://www.eugris.info (Petrie et al., 2015).

Groups of products	Classes of chemicals	Examples
Human and veterinary pharmaceuticals	Antibiotics, anti-parasitic agents, ionophores	Amoxicillin, erythromycin, metronidazol, tetracycline, lincomycin, sulfathiazole
	Stimulants and drugs including anti-inflammatory, anti-diabetic, anti-epileptic, anti-hypertensive, or anti-cancer drugs, anticoagulants, hallucinogens, analgesics, β-blockers, anti-depressants, lipid regulators, or erectile dysfunction drugs	Amphetamine, cocaine, caffeine, nicotine, propranolol, ibuprofen, codeine, carbamazepine, bezafibrate, metformin, fluoxetine, warfarine, valsartan, tramadol, morphine, methandone, diazepam, ephedrine, tamoxifen
	Hormones including natural and synthetic estrogens, androgens	Estrone, estriol, testosterone, progesterone, mestranol (ovulation inhibitor), cholesterol
Industrial and household wastewater products	Insecticides, plasticizers, detergents, flame retardants, polycyclic aromatic hydrocarbons, antioxidants, solvents, disinfectants, fumigants, fragrances, preservatives	Carbaryl, chlorpyrifos, diethylphtalate, *p*-nonylphenol, tri(2-chloroethyl)phosphate, naphtalene, anthracene, 2,6-di-*tert*-butylphenol, 1,2,3-trichloropropane, phenol, 1,4-dichlorobenzene, acetophenone
Personal care products	Insect repellents, polycyclic musks, sunscreen agents, fragrances, antiseptics	Bisphenol A, 1-benzophenone, methylparaben, *N*,*N*-diethyltoluamide, triclosan
Nanomaterials	Miscelaneous	Nanosilver, alumina nanoparticles, titanium dioxide, fullerenes, carbon black

in aerobic or anaerobic conditions. Advantageous properties such as small genome size, relative simplicity of the cell, short replication times, rapid evolution and adaptation to the new environmental conditions made microbes, and particularly bacteria, favourable candidates for bioremediation technologies, that is *in situ* or *ex situ* removal of polluting chemicals from the environment using biological agents. The removal of environmental pollution caused by the extensive activities of industrial society is a serious topic that draws the attention of biotechnologists. This is because beyond the medical and environmental consequences, the situation signs considerable potential for growth of eco-industry focused on clean-up technologies and removal of environmental contaminants. In fact, valorization of waste chemicals accumulating in industry is one of the pillars of the circular economy and the 4th Industrial Revolution (Schmidt, 2012; Wigginton et al., 2012).

The earliest attempts at directed bioremediation, although not formalized as such at the time, dated back to the late 19th century with the origins of the first wastewater treatment plants (Litchfield, 2005). Bioremediation began in earnest some 45 years ago with the isolation of culturable bacteria from contaminated sites and studying their degradation pathways. The first report on enhanced *in situ* bioremediation of soil contaminated with peroleum-derived hydrocarbons was published in 1975 by Raymond et al. (1975). Natural microbial degraders were later applied with success in world-wide and local biotechnological processes including large-scale wastewater denitrification, uranium removal, and degradation of 1,2-dichloroethane from groundwater or the organophosphorus pesticide coumaphos from cattle-dip waste (Francis and Mankin, 1977; Lovley et al., 1991; Mulbry et al., 1998; Stucki and Thueer, 1995). The advent of technologies for pollutant removal using naturally emerging microorganisms could be called the era of *Bioremediation 1.0*. Even so, a number of specific chemicals, especially of anthropogenic origin, including persistent organic pollutants such as dichlorodiphenyltrichloroethane (DDT), trichloroethylene, 1,2,3-trichloropropane, some polychlorinated biphenyls (PCB) or dioxins continued to be resistant to natural biodegradation due to lack of efficient microbial catabolic traits whose evolution was not sufficiently rapid or ended in a deadlock (Janssen et al., 2005).

Initial discoveries in molecular biology and progress in biological engineering disciplines seemed to provide a partial solution for such challenges through rational interventions in the metabolic networks of selected microbial hosts. The rise of recombinant DNA technology allowed the transformation of bioremediation from empirical practice into an excercise in genetic engineering, giving rise to what we might term *Bioremediation 2.0*. The goal of the new field was to engineer whole microbes, their biodegradation pathways, and the corresponding enzymes towards *in situ* mineralization of target pollutants. Such *superbugs* were expected to provide an economically feasible, environmentally friendly alternative to the costly conventional technologies for pollutant removal available at the time (Ramos et al., 2011). The late 1980s and early 1990s represented the golden era of biodegradation research, with numerous engineering attempts following the pioneering work by Chakrabarty and co-workers (Kellogg et al., 1981). They described the preparation of recombinant *Pseudomonas putida* strains able to break down crude oil by the *plasmid-assisted molecular breeding*, that is, propagation of novel catabolic capabilities through directed bacterial conjugation and plasmid transfer. The persistence of many xenobiotics was attributed mainly to the absence of complete degradative pathways in a single organism (Brenner et al., 1994; Reineke and Knackmuss, 1979). Recruitment of complementary enzyme sequences by conjugative gene transfer and so called *patchwork assembly* of several existing natural pathways in a suitable host was believed to generate functional synthetic routes that would allow for the complete mineralization of persistent target compounds such as PCB (Lehrbach et al., 1984; Ramos et al., 1987; Rojo et al., 1987).

Despite some success with the *patchwork* strategy and engineering of *superbugs* with extended substrate scope in laboratory conditions, this initial and rather naïve approach led to many disappointments as well (Cases and de Lorenzo, 2005; de Lorenzo, 2009). A prominent example was the case of engineered *Pseudomonas* strains that did not grow on 2-chlorotoluene as the only carbon source, even though they possessed all the genetic components presumed necessary for substrate mineralization (Haro and de Lorenzo, 2001). From a contemporary perspective, such failures can be explained by lack of insight into important factors such as: (i) thermodynamic feasibility of assembled catabolic networks, (ii) kinetic characteristics of enzymes and physicochemical properties of metabolites, (iii) expression levels of pathway modules, (iv) cross-talk between exogenous and endogenous metabolic routes, and (v) stress responses and changes in overall host cell physiology after introduction of new metabolic modules and exposure to toxic substrates and metabolites (de Lorenzo, 2009; Ramos et al., 2011).

Fortunately, the last decade has witnessed the onset of what can be called *systemic biology*, which merges different approaches of systems biology, metabolic engineering, and synthetic biology, for the sake of understanding and reprograming biological systems. Systemic biology has the potential to remove the unknowns and bottlenecks encountered in past trials and paves the way towards the era of *Bioremediation 3.0*. The joint power of the systemic biology disciplines can ensure that biodegradation and bioremedation using genetically modified microorganisms will remain a vital concept deserving of the full attention of new generations of bioengineers.

In this article we review the applications of novel engineering strategies to the design and evolution of microbial biodegradation

pathways and whole-cell degraders from the last decade, and propose an optimal workflow for pathway design, construction and optimization. In particular, we discuss the potential of state-of-the-art *systemic* technologies not yet fully employed for this purpose including new ways to genetically engineer superior CO_2 scavengers. Lastly, the perspectives of microbial cell factories tailored for biodegradation and bioremediation are critically evaluated.

2. Approaches and tools of systems biology and metabolic engineering for tailoring biodegradation pathways

One key objective of systems biology is to gain comprehensive, quantitative understanding of living cells by combining high-throughput technologies and computational methods to characterize and predict cell behaviour (Dai and Nielsen, 2015). Metabolic engineering, first defined as a new scientific discipline by Bailey (1991), is now understood as the practice of optimizing genetic and regulatory processes within the cells to (i) improve the yield and productivity of native products synthesized by organisms, (ii) establish the synthesis of products new to the host cell, and, two points especially relevant for biodegradation, to (iii) extend the range of substrates or improve substrate uptake, and (iv) improve overall cell robustness (Nielsen et al., 2014). These goals can be achieved by engineering natural metabolic pathways in the host cell or synthetic routes assembled from enzymes originating from different organisms. The aims of systems biology coincide fully with the objectives of metabolic engineering. These two disciplines are now inseparable, complement each other, and have even merged into a field of systems metabolic engineering (Lee et al., 2012). Metabolic engineers use systems biology computational tools and 'omic' techniques to gain deeper insight into the genetic and physiological background of target organisms, to model enzymatic reactions and to determine the constraints for efficient biocatalysis. To overcome these constraints, computational tools are applied together with established experimental protocols, now frequently strengthened with synthetic biology standards and fine-tuning adaptive laboratory evolution. The *Design-Build-Test-Analyze* (DBTA) cycle is repeated until performance of the engineered cell factory is optimized and a cost-effective process can be established (Paddon and Keasling, 2014). It is tempting and logical to use the same intellectual workflow to engineer biodegradation pathways (Fig. 1). The portfolio of systems metabolic engineering tools applicable for such purposes will be discussed in more detail in the Section 2 of the review.

2.1. Step 1: get to know the contaminant and find a suitable catabolic pathway

The compound to be degraded or removed from the contaminated environment or industrial waste site is usually the very first component known in a project focused on biodegradation or bioremediation. The polluting chemicals show diverse physicochemical properties, can occupy heterogenous physical niches in the environment, and exist in concentrations ranging from ng/L to mg/L (Meckenstock et al., 2015). Industrial waste chemicals, prohibited pesticides, or warfare agents such as 1,2,3-trichloropropane, γ-hexachlorocyclohexane, or yperite, respectively, can be available for degradation as stock piles or mixed with environmental matrix. When bioremediation is a method of choice for removal of the target compound, another challenge arises: how to find the most suitable pathway for degradation and, ideally, complete mineralization of the chemical in the huge amount of genetic, biochemical and microbiological data available. As pointed out by Nobel laureate Sydney Brenner, we are currently "…drowning in a sea of data, thirsting for knowledge…". New computational tools and algorithms, combined with common sense, are the only way to cope with the complexity of life, find the needle in a haystack, and obtain valuable output – the natural or synthetic pathway or set of pathways most suitable for a specific biodegradation task in a selected microbial host. Computational tools can provide the user with relevant information on physicochemical properties of the target compound and the basic building blocks of the catabolic pathway – enzymes and metabolites. A wide spectrum of databases and prediction systems that provide useful information on a number of chemicals and biodegradative or biosynthetic routes has been developed over the years. Comprehensive reviews that focus on evaluation of computational tools mainly for biosynthetic pathway design have been published in the last few years (Long et al., 2015; Medema et al., 2012). Chemical and biodegradation databases as well as pathway and toxicity prediction systems that might be useful for evaluating existing biodegradation pathways or engineering new ones were reviewed by Arora and Bae (2014). To avoid duplication, here we aim to discuss mainly the updates and applications of representative systems for biodegradation pathway design.

2.1.1. Databases

The University of Minnesota Biocatalysis/Biodegradation Database and Pathway Prediction System (UM-BBD/PPS) is a remarkable tool that has garnered microbial pathways for xenobiotics for over 20 years and can be considered an advisable first-choice search engine (Gao et al., 2010). In 2014, UM-BBD/PPS rights were assigned to Eawag, the Swiss Federal Institute of Aquatic Science and Technology and the database got a new name EAWAG-BBD/PPS. EAWAG-BBD (http://eawag-bbd.ethz.ch) can not only search through 543 microorganism entries, 219 metabolic routes, 1503 reactions, 1396 chemicals, and 993 enzymes, but also allows prediction of novel biotransformations leading from target xenobiotic substrate to the molecule that can be metabolized by central catabolic pathways of a host cell. The enviPath tool (The Environmental Contaminant Biotransformation Pathway Resource; https://envipath.org/) was recently introduced as a rebuilt version of EAWAG-BBD/PPS (Wicker et al., 2016). Registered user can employ enviPath for design of particular biochemical routes and personal databases with biotransformation data as well as for prediction of new catabolic pathways. In 2017, the same team introduced a new database implanted in enviPath platform, Eawag-Soil (Latino et al., 2017). This public database is a unique repository of data collected during laboratory simulations on aerobic degradation of pesticides in soil, including biotransformation half-lifes.

Although they provide access to much more data in addition to biodegradation pathways and related content, MetaCyc and BioCyc databases by SRI International should be mentioned as well (Caspi et al., 2016). With its 2526 experimentally elucidated pathways from 2844 different organisms from all domains of life, MetaCyc (https://metacyc.org/) is one of the largest repositories of metabolism data and serves primarily as an on-line encyclopedia in which metabolic routes can be predicted from available sequenced genomes, browsed through, or mutually compared. The majority of listed compounds and reactions now also include the Gibbs free energy values (Caspi et al., 2016). BioCyc (https://biocyc.org/) is a collection of 9389 organism-specific Pathway/Genome Databases (PGDB; the number tripled in the last three years!). Each PGDB encompasses the sequenced genome and predicted metabolic network of single organism. Moreover, information on individual components such as predicted operons, metabolites, enzymes and their reactions, or transport systems is provided (Caspi et al., 2016). BioCyc also provides several useful tools for navigating, visualizing, and analyzing the PGDB, and for analyzing omics data. Both MetaCyc and BioCyc are linked to numerous other databases. UniProt (www.uniprot.org/), a collaboration between the European Bioinformatics Institute, the SIB Swiss Institute of Bioinformatics and the Protein Information Resource, supplies missing information on function, sequence, or taxonomy of proteins from metabolic or signalling pathways and is further interconnected with more specific protein databases such as RCSB Protein Data Bank or BRENDA, thus completing the picture with structural or kinetic data (Berman et al., 2000; Chang et al., 2015; UniProt Consortium, 2015).

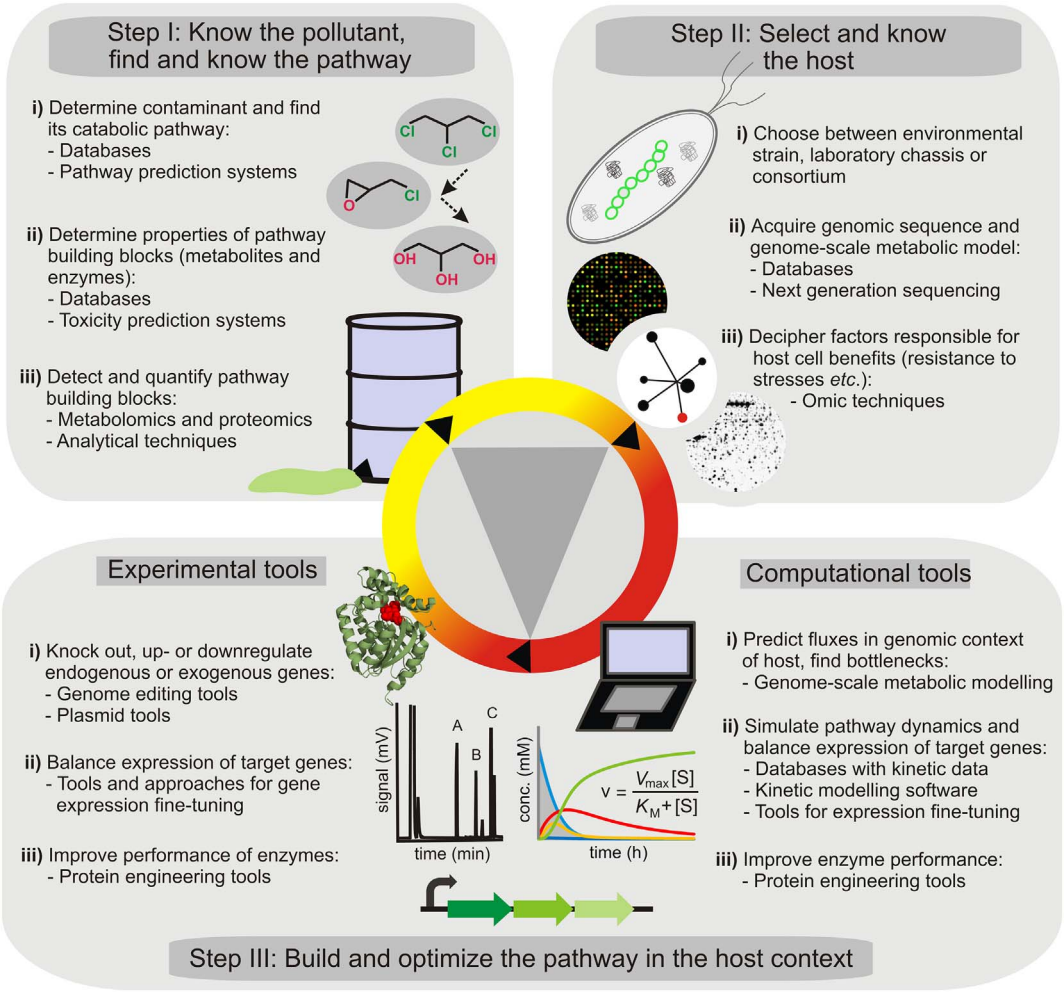


Fig. 1. Proposed workflow for engineering biodegradation pathways using systemic biology approaches and tools. The Figure sketches the roadmap to contructing superior bacterial catalysts for environmental bioremediation that capitalize on systems and synthetic biology, as explained in this review.

2.1.2. Pathway prediction systems and toxicity prediction algorithms

The development of reliable pathway prediction systems is of major importance for the field of biodegradation as for the numerous anthropogenic chemicals released into the environment the catabolic pathways are either completely unknown or poorly understood. These tools consider many factors including substrate specificities, binding sites or reaction mechanisms of enzymes, structural changes in substrate-product pairs and pathway distance from substrate to product (Medema et al., 2012). The pathway prediction system of EAWAG-BBD/PPS or enviPath, predicts biodegradation routes based on available biotransformation rules derived from reactions found in the EAWAG-BBD database or in the literature (Gao et al., 2010; Wicker et al., 2016). It can highlight initiating reaction types likely to occur in aerobic environments, which in some cases lead to the complete mineralization of the contaminant (Fig. 2), although, it does not define the thermodynamic feasibility of the proposed pathways or the specific enzymes that catalyse proposed reactions. This latter information can be supplemented by the complementary prediction algorithm of a new public database RAPID (http://rapid.umn.edu/rapid/), currently being developed by the group of Lawrence P. Wackett at University of Minnesota. The idea of future utility of combined EAWAG-BBD and RAPID algorithms was provided recently by Aukema and co-workers who used these tools to predict initial metabolism of a set of emerging contaminants that include recalcitrant pharmaceuticals, alkyl phthalates, or fragrance compounds previously shown to be degraded by *Paraburkholderia xenovorans* LB400 (Aukema et al., 2016).

In case of anthropogenic contaminants with unknown or uncomplete natural catabolic pathways, the Biochemical Network Integrated Computational Explorer (BNICE.ch) can be recommended as well (Hatzimanikatis et al., 2005). The authors from the Vassily Hatzimanikatis' group at the Swiss Federal Institute of Technology have been developing BNICE.ch, one of the first tools for uncovering and describing new enzymatic reactions based on known biochemistry, for > 15 years. In the latest version, the method was coupled with a manually-curated KEGG database, generally recognized as one of the most complete repositories of metabolic data (Kanehisa et al., 2014). The catalogued biochemistry of the KEGG database was converted into 361 bidirectional generalized rules which were applied to explore the possible space of all metabolic reactions that link the compounds reported in the KEGG and are potentially found in nature (Hadadi et al., 2016). As a result, 6528 KEGG reactions and 137,416 known or completely new enzymatic reactions between two or more KEGG compounds were organized in a reaction web-based database named ATLAS of Biochemistry (http://lcsb-databases.epfl.ch/atlas/). Importantly, the thermodynamic feasibility of *de novo* generated reactions was evaluated and the parameter of Gibbs free energy is a part of the majority of records in the ATLAS database. To the benefit of the reaction prediction system, according to authors, up to 80% of the newly added KEGG reactions in 2015 already existed in ATLAS as novel reactions.

In biodegradation pathway design, BNICE.ch was first used by Finley and co-workers to propose all possible catabolic routes for 4-chlorobiphenyl, phenanthrene, γ-hexachlorocyclohexane, and 1,2,4-

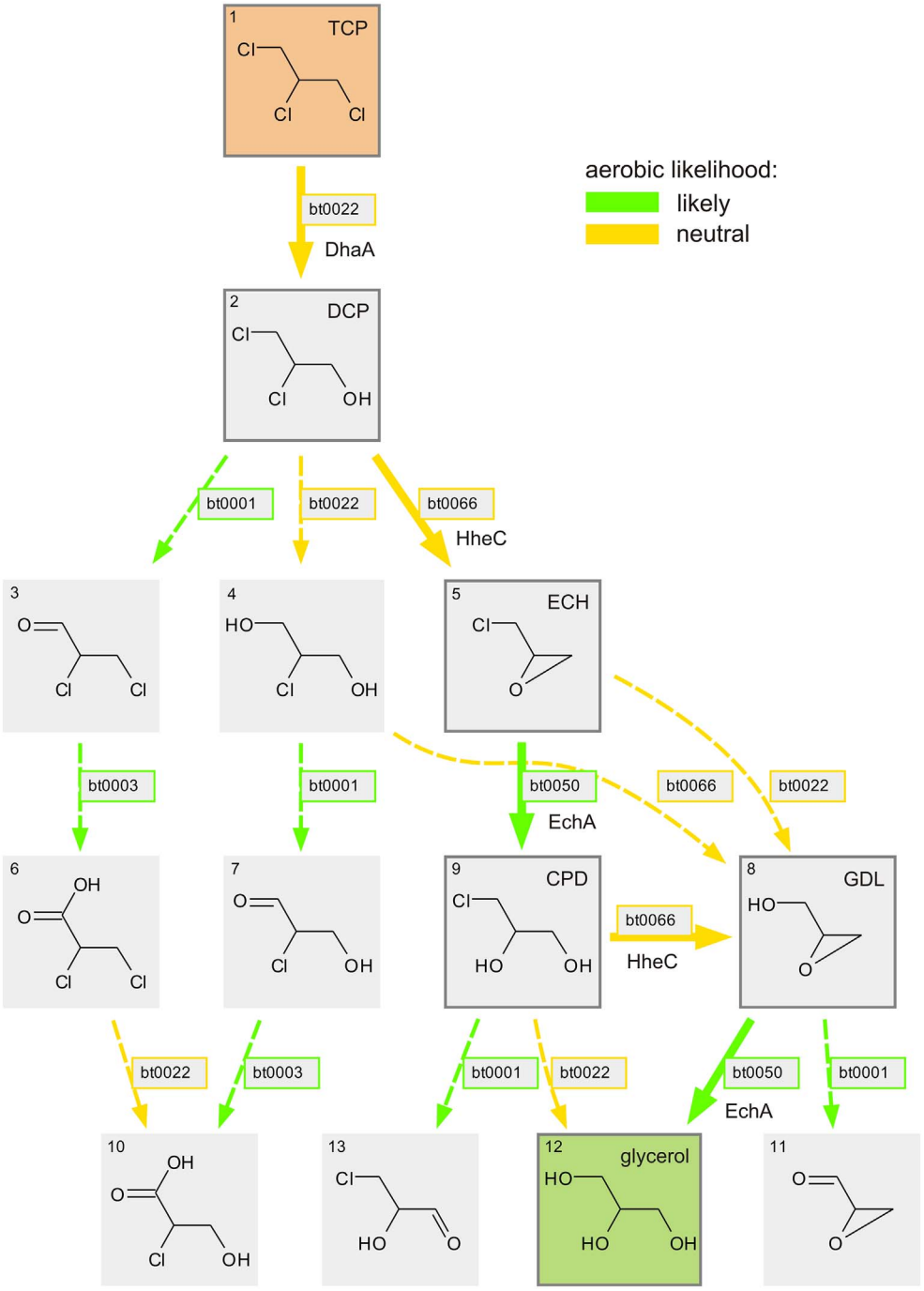


Fig. 2. Initial aerobic biotransformations of the emerging contaminant 1,2,3-trichloropropane (TCP) generated by the EAWAG-BBD Pathway Prediction System. Reactions are tagged with one of the 249 biotransformation rules used for the predictions. Reactions proven experimentally to be catalysed by haloalkane dehalogenase DhaA, epoxide hydrolase EchA and halohydrin dehalogenase HheC in aerobic conditions are highlighted with thick arrows (Bosma et al., 1999). The scheme was adopted and modified from http://eawag-bbd.ethz.ch/predict/. Aerobic likelihood specifies whether the reaction will occur in aerobic conditions, exposed to air, in soil (moderate moisture) or water, at neutral pH, 25 °C, with no competing compounds. Abbreviations: DCP, 2,3-dichloropropane-1-ol; ECH, epichlorohydrin; CPD, 3-chloropropane-1,2-diol; GDL, glycidol.

trichlorobenzene (1,2,4-TCB), compounds that represent various classes of xenobiotics (Finley et al., 2009). BNICE.ch not only reproduced the experimentally verified pathways, but also found completely novel reactions, taking into account the starting compound, the requested length of the pathway and the broader reaction rules of the Enzyme Commission classification system. The 15 novel pathways for 1,2,4-TCB were subsequently probed using thermodynamic analysis and found energetically feasible. In the follow-up study, the authors evaluated the new routes in the context of *Pseudomonas putida* KT2440 cellular metabolism (Finley et al., 2010). They expanded the available *P. putida* metabolic model by including the pathways obtained from BNICE.ch and metabolic flux analysis was used to predict the maximum biomass generated using 1,2,4-TCB as the sole carbon source. In this way, interesting alternative pathways were proposed that couple 1,2,4-TCB utilisation with biomass formation. Although this work was purely theoretical, it suggests the way for those who wish to use the increasing power of computational modelling to design biodegradation routes or explore the fate of xenobiotics in the environment.

Prediction tools for assisting the design of robust microbial biosensors are highly desirable as well. Such biosensors are often based on regulatory proteins or riboswitches able to translate the presence of a given metabolite or small molecule into a readily measurable readout (Gredell et al., 2012; Schallmey et al., 2014; Cardinale et al., 2017). But what to do when biosensors for a compound of interest do not exist? An interesting approach to tackle the issue was presented recently by Delépine et al. (2016). Their SensiPath algorithm (http://sensipath.micalis.fr/) stems from the concept of retrosynthesis introduced previously by the same laboratory (Carbonell et al., 2014). SensiPath

applies more than 9,000 unique reaction rules collected from BRENDA, Metacyc, and Rhea (Morgat et al., 2015) databases to find enzymes (currently limited to max. two-step reactions) able to convert a target undetectable organic chemical into a molecule that can be recognized by a known natural sensing system. In the theoretical part of the follow-up study the authors applied SensiPath to expand the range of thereby detectable compounds among chemical structures collected from several databases (Libis et al., 2016). The method almost tripled (from 169 to 477) the number of theoretically detectable compounds in the TOX21 database (Krewski et al., 2009). The predictive power of the method was validated in the experimental part of the study. Metabolic modules of up to two enzymes proposed by SensiPath were introduced into *E. coli* bearing a sensing module for the product of enzymatic conversion at stake. Functional whole-cell biosensors for cocaine, nitroglycerin, or parathion were constructed in such way. Despite this remarkable contribution, the number of undetectable organic chemicals remains at a high > 94% of all compounds in TOX21 database. Strengthening the computational power of tools such as SensiPath will hopefully improve the situation in the near future.

The predicted reaction networks can also be pruned using toxicity estimation algorithms, to prevent the formation of compounds highy toxic to the host (Benfenati, 2007). Besides quantitative structure activity relationship (QSAR) models, which calculate toxicity based on the physical characteristics of the chemical structures (molecular descriptors; Eriksson et al., 2003), other models have recently been developed to predict toxicity based, for example, on chemical-chemical interactions (Chen et al., 2013). QSAR models nonetheless remain the paradigm in the field. An interesting example of a QSAR model-based computational tool (https://absynth.issb.genopole.fr/Bioinformatics/) is the toxicity prediction web server EcoliTox (Planson et al., 2012). The dataset obtained from screening a diversified chemical library of 166 compounds for toxicity in *Escherichia coli* was used to develop a predictor of toxicity of metabolites throughout the metabolome of this popular bacterial host. Tools similar to the EcoliTox server could be integrated into a computational framework for improved design of native or heterologous biodegradation pathways and used to fine-tune their expression in selected microbial hosts. A powerful virtual screening approach, currently applied predominantly in drug design studies (Buryska et al., 2016), could be used similarly for synthetic pathway predictions to avoid selection of enzymes whose catalytic machinery might be inhibited by molecules present in a host cell.

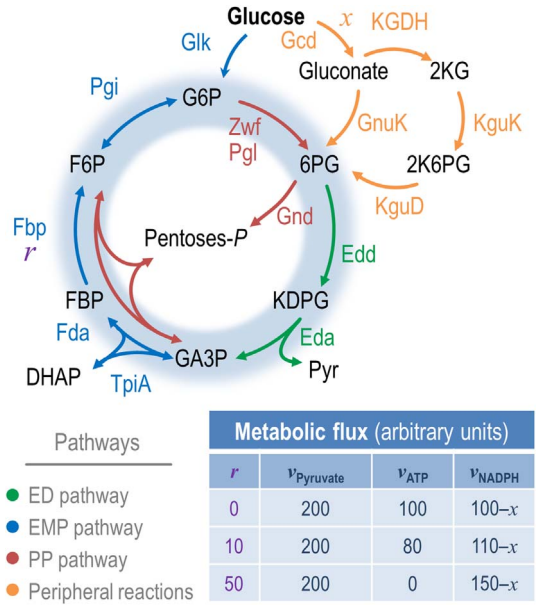
2.1.3. Detection and quantification of pathway building blocks

Once the suitable pathway is selected, another issue arises, that of how to detect and quantify the pathway building blocks - metabolites and enzymes. Detection and accurate quantification of metabolites in complex reaction mixtures and of functional enzymes produced within the cell provide valuable information on pathway performance and its bottlenecks, as well as a background for further optimization of the route. Here, the experimental tools and approaches come into play for the first time in our workflow scheme (Fig. 1). The best sensitivity and separation power is currently provided by chromatographic separation of metabolites followed by mass spectrometry analysis (Büscher et al., 2009). Information on physicochemical properties of compounds (boiling point, polarity, molecular weight) generated in the pathway is nonetheless crucial for correct choice of a detection method, be it gas chromatography, liquid chromatography, or capillary electrophoresis (Büscher et al., 2009). Of the numerous chemical databases, PubChem (https://pubchem.ncbi.nlm.nih.gov/), created in 2004 by the US National Institutes of Health, is the world's largest free chemistry repository, into which anyone can deposit data on structures and their properties (Kim et al., 2016). The database has grown to > 92 million structures. Bioactivity data of chemical substances, manually withdrawn from the scientific literature, are saved in PubChem Substance and Compound databases. Integration with other databases allows including pharmacology, toxicology, drug target information, safety or handling information to the annotation of chemical records. PubChem also hosts data from important regulatory agencies, including the US Environmental Protection Agency. In 2007, another free repository, ChemSpider (http://www.chemspider.com/), was created and later purchased by the UK Royal Society of Chemistry; it provides access to over 59 million entries from 487 diverse data sources and can be recommended as another rich source of knowledge on experimental or predicted physicochemical properties of compounds.

For pathway enzyme detection and absolute quantification, SDS-PAGE or better, Western blot analysis, followed by determination of enzyme activity have been standard protocols in metabolic engineering for many years (Kurumbang et al., 2014). When applied to multiple-enzyme pathways, however, these methods are tedious and time-consuming. Selected Reaction Monitoring Mass Spectrometry protocols developed in the past few years seem to provide a promising alternative for absolute protein quantification in the future, as they allow rapid simultaneous quantification of many proteins in the cell, with good selectivity and sensitivity regardless of organism of origin (Batth et al., 2012).

2.2. Step 2: select and get to know a suitable microbial host

In addition to catabolic pathway choice or design, the selection of a suitable host is a crucial initial step in any engineering project focused on the development of a whole-cell degrader. Over the years, several microbial hosts have been considered for application in biodegradation processes, but no single, naturally isolated bacterial strain possesses all the desired characteristics of the optimal degrader. Soil bacteria could satisfy many of these requirements because of the conditions they face naturally in the niches in which they thrive, including exposure to environmental contaminants and to competing and predatory species. These microorganisms have versatile metabolic lifestyles that allow them to adapt to changing conditions, sometimes adverse (oxidative stress, temperature challenges, osmotic perturbations). *P. putida* is a ubiquitous rhizosphere colonizer that belongs to the wide (if somewhat fuzzy) group of fluorescent *Pseudomonas*. *P. putida* KT2440 is the best-characterized saprophytic laboratory pseudomonad that has retained its ability to survive and thrive in the environment. Its entire chromosome sequence is available since 2002 (Belda et al., 2016; Nelson et al., 2002, Nikel et al., 2014). This strain derives from *P. putida* mt-2, a naturally occurring species able to degrade several aromatic compounds through the activities encoded in the catabolic TOL plasmid. The physiological and metabolic properties of pseudomonads argue for their selection as the starting point for engineering of biodegradation processes. In addition to rapid growth and low nutrient demand, their extremely versatile metabolism and high capacity to supply redox power, providing endurance to oxidative stress, are several advantages that render *P. putida* an interesting host for degradation applications (Nikel et al., 2016a). Most of these properties arise from a very robust central metabolism (Fig. 3), in which the combined activity of enzymes from different pathways (*i.e.*, the EDEMP cycle; Nikel et al., 2015) endows the cells with a large NADPH supply (Chavarría et al., 2013). The operativity of the EDEMP cycle enables the formation of ATP and NADPH at different rates depending on the amount of trioses phosphate recycled. Assuming that *P. putida* has (i) a NAD^+ dependent glyceraldehyde-3-*P* dehydrogenase and a $NADP^+$ dependent glucose-6-*P* dehydrogenase, (ii) pyruvate as the end product of glycolytic pathways, and (iii) a negligible flux through 2-ketogluconate, the yields of pyruvate, ATP, and NADPH are as indicated in the inset to Fig. 3. The ability to sustain high NADPH formation rates is a requisite needed in a bacterial host for the implantation of biodegradation pathways (Lorenzo and Loza-Tavera, 2011). As will be shown in part in following sections, these beneficial properties are reflected in the increasing number of metabolic engineering studies in which *P. putida* appears as one major player.



Metabolic flux (arbitrary units)			
r	$v_{Pyruvate}$	v_{ATP}	v_{NADPH}
0	200	100	100–x
10	200	80	110–x
50	200	0	150–x

Fig. 3. The action of the EDEMP cycle is illustrated with a theoretical example of flux distribution in glucose-grown *P. putida* KT2440. Total carbon uptake is considered to be 100 (arbitrary units), and glucose can be split into the phosphorylative (Glk) or oxidative (x) branches. The net formation rates (v) of pyruvate (Pyr), ATP, and NADPH are indicated in the inset table, with r representing the amount of triose phosphates recycled to hexose phosphates through the action of the EDEMP cycle. Note that part of the glucose can also be oxidized through the action of the so-called *peripheral reactions*, represented by the overall flux x in the scheme, and that two trioses are formed per glucose consumed, *i.e.*, a glucose uptake flux of 100 (arbitrary units) will result in a pyruvate flux of 200. Abbreviations: ED pathway, Entner–Doudoroff pathway; EMP pathway: Embden–Meyerhof–Parnas pathway; PP pathway, pentose phosphate pathway; G6P, glucose-6-P; F6P, fructose-6-P; FBP, fructose-1,6-P_2; DHAP, dihydroxyacetone-*P*; GA3P, glyceraldehyde-3-*P*; 6PG, 6-phosphogluconate; KDPG, 2-keto-3-deoxy-6-phosphogluconate; 2KG, 2-ketogluconate; and 2K6PG, 2-keto-6-phosphogluconate. Figure adapted and modified from (Nikel et al., 2016a).

2.2.1. *Omics techniques in studies of bacterial degraders*

Systems biology, and particularly 'omic' techniques, contributed markedly to rational selection of suitable hosts for metabolic engineering. The enormous amount of genetic information read with first- or next-generation sequencing techniques together with powerful new algorithms and computational tools to aid with data mining, allowed considerable progress in our understanding of dynamic interactions inside and in between cells and the environments they inhabit (Bouhajja et al., 2016; Vilchez-Vargas et al., 2010). Available genomic data also enabled new insights into the evolution of microbial metabolism towards biodegradation of well-recognized xenobiotics such as 1,3,5-triazine (Shapir et al., 2007). 16S rRNA-phylogenetic microarrays and the functional gene microarrays have been used to monitor microbial community dynamics and to track activities of microbes in the polluted environments, respectively (Chakraborty et al., 2012). The reconstructed genome-scale metabolic models, together with transcriptomic, proteomic, metabolomic, and flux analyses in various growth and stress conditions, completely changed our view of popular biodegradative bacterial hosts including the paradigmatic *P. putida* (Sohn et al., 2010).

For instance, fluxomic analyses applied to *P. putida* KT2440 helped to reveal its distinct strategies for dealing with carbon sources, *e.g.*, by favouring the routes that generate NADPH, which promotes greater resistance to oxidative stress (Chavarría et al., 2013; Nikel et al., 2013). Examination of transcriptomic data uncovered the significant portion of the *P. putida* genome (≥ 20%) that is differentially expressed when the cells are grown on diverse substrates and emphasized the role of a suite of global regulators (J. Kim et al., 2013). Recent complete structural re-annotation of the KT2440 strain genome allowed identification of 1485 CDSs associated to 1898 chemical reactions, prediction of catabolic pathways for 92 compounds (carbon, nitrogen, and phosphorus sources) and upgrading of the available genome-scale metabolic model (Belda et al., 2016; Puchałka et al., 2008). 'Omic' analyses greatly increased insight into the genetic and physiological background of some other pseudomonads besides KT2440, including *P. pseudoalcaligenes* CECT5344, which might be used for biodegradation of industrial cyanide-containing wastes, or the 2,4,6-trinitrotoluene biotransforming *P. putida* JLR11 (Pascual et al., 2015; Wibberg et al., 2016). At the time of writing this article, there are 3133 drafted and 215 completed *Pseudomonas* genomes currently available in the *Pseudomonas* Genome Database (http://www.pseudomonas.com/), perhaps the most comprehensive genomic database dedicated to a single bacterial genus (Winsor et al., 2016).

Complete genomic sequences are currently available for many other potentially useful degraders such as *Cupriavidus necator* JMP134, a bacterium with nearly 300 genes involved in the catabolism of aromatics, including chloroaromatics (Lykidis et al., 2010). Other examples include, but are not limited to, the PCBs-degrading *Acidovorax* sp. strain KKS102 (Ohtsubo et al., 2012), the cyclic hydrocarbons-degrading *Alicycliphilus denitrificans* strains BC and K601 (Oosterkamp et al., 2013), or the oil-degrading bacterium *Oleispira antarctica*, whose genome has provided useful information as to how bacteria mitigate oil spills in cold environments (Kube et al., 2013). Genomic sequences of these and many other microorganisms with biodegradation capabilities can be searched through in comprehensive databases such as NCBI genome database (https://www.ncbi.nlm.nih.gov/genome/) or MicroScope (http://www.genoscope.cns.fr/agc/microscope/home/), a microbial genome annotation and analysis platform (Vallenet et al., 2016), developed in the Laboratory of Bioinformatics Analyses for Genomics and Metabolism at the National Sequencing Centre (Evry, France); it provides a complete pipeline for annotations and comparative analyses of up to 6000 microbial genomes, of which hundreds were manually curated for accuracy.

Despite interest in study of individual bacterial species, the fascinating *in situ* systems biology exercise that followed one of the most alarming ecological disasters in modern human history, the 2010 BP Deepwater Horizon spill in the Gulf of Mexico, emphasized the irreplaceable role of bacterial communities in natural bioremediation. Bacteria that can convert oil-derived alkanes and aromatics into the biomass were identified based on (i) genome reconstructions *via* metagenome sequencing of the DNA from the stable-isotope-probing experiments (Dombrowski et al., 2016), and (ii) cultivation trials using oil samples from sea surface and deep sea collected during the outflow (Gutierrez et al., 2013). Only the concerted action of many microbial species including *Cycloclasticus*, *Alcanivorax*, members of *Rhodospirillales*, *Alteromonas* and others, augmented by the specific environment of the oil spill, allowed the degradation of the complex mixture that consisted of as many as 1000 compounds. Metagenomic strategies help to shed light also on the composition and dynamics of microbial consortia that participate in natural biodegradation of chorinated ethenes in contaminated groundwater (Adetutu et al., 2015) or polycyclic aromatic hydrocarbons in soils (Guazzaroni et al., 2013).

2.3. *Step 3: build the pathway and optimize its performance in the context of host metabolism*

Optimization of the proposed biochemical pathway in the context of host cell metabolism is often the most demanding part of the engineering project. This is especially true for *de novo* synthetic pathways composed of enzymes from different sources, implanted into a new bacterial host. Codon composition of introduced genes might not be optimal for expression in a heterologous host, enzyme activities might not be well balanced or might cross-talk with the native metabolic network, pathway intermediates can accumulate and inhibit the growth of cells, or cofactors, ATP or redox carriers might be lacking. The

portfolio of systemic tools that can help overcome these issues is expanding constantly. Here, we will center on the major concepts that have proven useful in tailoring biodegradation pathways and microbial degraders. We will discuss the application of several of these modern tools including the example of engineering the synthetic catabolic route for 1,2,3-trichloropropane (TCP, Fig. 2), an industrial byproduct of anthropogenic origin now being recognized as a groundwater contaminant (Samin and Janssen, 2012). TCP is representative of difficult-to-degrade halogenated aliphatic compounds, a tough case for remediation technologists due to its physicochemical properties, toxicity in living organisms, and absence of efficient natural catabolic pathway (Janssen et al., 2005; Samin and Janssen, 2012). The enzymes later used to compose the first synthetic pathway that allowed mineralization of TCP by an engineered bacterium had already been described in the late 1980s and early 90s (Bosma et al., 1999). With a history spanning > 25 years, the engineering of this biodegradation pathway thus represents one of the most systematic efforts of its kind.

2.3.1. Computational tools for pathway and strain optimization

The tremendous increase in the number of sequenced genomes (the number of complete genomes in the NCBI database has more than doubled in the last two years) of unicelluar and multicellular organisms, including the microbial degraders mentioned above, led to the reconstruction of genome-scale metabolic models representing all (anotated) biochemical reactions that take place in the living cell (King et al., 2015). Mathematical models of metabolic networks have a central role in metabolic engineering, and in pathway and strain optimization. Modelling can be used to analyze the selected pathway (*e.g.*, to determine the distribution of metabolic fluxes) and identify reactions that must be modified to improve its performance. Genome-scale metabolic models are built by compiling data on genes related to biochemical reactions from databases such as KEGG and BioCyc. The genes are compared to already finalized reconstructions of related organisms to find homologous reactions; the draft model is then refined and verified by simulations. Flux Balance Analysis (FBA) and Metabolic Flux Analysis (MFA) are two major approaches that use metabolic models to predict intracellular fluxes (Stephanopoulos, 1999). FBA is a theoretical concept that uses the stoichiometric coefficients for each reaction in the system as the set of constraints for optimization. MFA determines intracellular fluxes from measurable rates of metabolite uptake and secretion in growth medium. In a typical experiment, substrates labelled with the non-radioactive isotope ^{13}C are used in chemostat-grown cultures to trace fluxes through a cellular network. Computational tools such as Cobra 2.0 for Matlab, OptKnock, or k-OptForce can implement constraint-based flux calculations and suggest strategies based on knock-out, upregulation or downregulation of target genes to optimize metabolite production without compromising cell growth (Burgard et al., 2003; Chowdhury et al., 2014; Schellenberger et al., 2011). In a study by Izallalen and colleagues, OptKnock was used together with the constraint-based *in silico* model of *Geobacter sulfurreducens* to determine gene deletions which could lead to increased respiration rates (Izallalen et al., 2008). *Geobacter* species are well recognized for their ability to degrade radioactive and toxic metals and have been explored for their utility in microbial fuel cells (Shi et al., 2016). These processes are nonetheless limited by electron transfer rates in cells. OptKnock calculations followed by genetic constructions by the authors resulted in recombinant *Geobacter* cells with decreased ATP levels, slower growth rates, and predicted higher respiration rates (Izallalen et al., 2008). Subsequent genome-wide analysis of gene transcript levels also verified the *in silico* predictions.

Although FBA- and MFA-based approaches can handle genome-scale metabolic models and require no information on enzyme kinetics, their output only gives a steady-state approximation of the dynamic reality in the living cell. Kinetic modelling is an alternative to static flux calculations when reliable kinetic parameters of pathway enzymes and some other input data are available, *e.g.*, specific surface area of the cell, permeability coefficients for substrates, metabolite concentrations, or approximate enzyme amount in the cell (Chowdhury et al., 2015). Kinetic constants can be obtained from enzyme databases such as BRENDA or can be determined experimentally. One should keep in mind that most kinetic constants deposited in databases are not standardized and also measured in *in vitro*, which might be a limiting factor when such data are applied to kinetic modelling in living systems (Costa et al., 2011). Computational tools such as COPASI or *E*-Cell are used to assemble the pathway model in the form of kinetic and differential equations, and to simulate reaction time courses for various conditions (Mendes et al., 2009; Takahashi et al., 2003). Kinetic models of whole metabolic networks are the holy grail of metabolic engineering but their application in strain design is limited by factors such as unreliability of parameters, non-universality of rate laws, need to implement regulatory events, or extremely high demand on computational power (Chowdhury et al., 2015). In contrast, the use of kinetic modelling to engineer individual metabolic pathways or simpler *in vitro*-assembled networks is currently more feasible (Muschiol et al., 2015).

Recent studies by Dvorak, Kurumbang and colleagues are notable example of kinetic modelling applied in engineering of synthetic biodegradation pathways (Dvorak et al., 2014b; Kurumbang et al., 2014). As a model, these studies centered on detoxification of the industrial waste compound and water pollutant TCP to glycerol - reaction catalysed by the haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCIMB 13064 and the haloalcohol dehalogenase HheC and the epoxide hydrolase EchA from *Agrobacterium radiobacter* AD1 (Bosma et al., 1999). *In vitro* steady-state kinetic data obtained for HheC, EchA, and three variants of DhaA were utilized to develop a kinetic model of the reaction that proved useful for revealing major bottlenecks in the pathway, that is (i) low activity of wild-type DhaA on TCP and (ii) unbalanced enantioselectivity and enantiospecificity of the first and the second enzyme in the route, DhaA and HheC, respectively (Dvorak et al., 2014b). Modelling was further applied to predict optimal enzyme stoichiometry and fine-tuning overall pathway efficiency, both in *in vitro* conditions and in the heterologous microbial host *E. coli*. In *E. coli*, the pathway was established as a modular, tunable system (Kurumbang et al., 2014). Absolute amounts of pathway enzymes in the cell, copy numbers of used plasmid vectors, and toxicity levels of pathway substrate and intermediates were included in the mathematical model as additional parameters. Variants of the pathway, predicted to promote better host survival in a medium with toxic TCP, were verified experimentally. The observed precise matching between calculated and experimental data confirmed the performance of the pathway in *E. coli* and stressed the power of kinetic modelling when reliable parameters are available. Finally, the kinetic parameters of an optimal DhaA variant that would ensure sufficient TCP conversion to glycerol as well as host cell growth on the toxic substrate were both predicted by mathematical modelling.

This study shows that, for engineering and implanting heterologous pathways, complete knowledge of the host cell metabolic background is not a must, and a reliable strain design can be achieved with a relatively simple mathematical model (Dvorak et al., 2014b). Even so, similar studies that combine theoretical and experimental approaches in the design of pathways and strains for biodegradation are still rare. Most recent work in the biodegradation field is restricted to application of purely experimental techniques from a repertoire of metabolic engineering, discussed in the following section.

2.3.2. Experimental tools for pathway and strain optimization

Once the bottleneck reaction steps are identified and a solution is proposed, experimental techniques are applied to target the appropriate genes or regulatory mechanisms. Engineering is traditionally conducted at the level of gene expression, which affects the quantity of protein. Genes in native or synthetic pathways are usually overexpressed by introducing plasmid vectors with extra copies of desired coding sequences. One popular approach, also used for modular assembly of the

TCP biodegradation pathway described above, is to combine Novagen Duet plasmid vectors (Merck Millipore, Germany) or derivatives such as ePathBrick vectors, which carry compatible replication origins and antibiotic markers (Kurumbang et al., 2014; Tolia and Joshua-Tor, 2006; Xu et al., 2012). This system allows effective propagation and maintenance of up to four plasmids in a single *E. coli* BL21 cell with λDE3 lysogen and modular assembly of multi-gene pathways. SEVA (Standard European Vector Architecture) plasmids with standardized architecture and nomenclature are an alternative modular vector system tailored to allow cloning or expression of heterologous genes in *P. putida*, *E. coli* and other Gram-negative bacteria, with no specific need for strain pretreatment (Martínez-García et al., 2015). Up to four plasmids, each with one of the six available antibiotic markers (ampicillin, kanamycin, cloramphenicol, streptomycin, tetracycline, gentamicin), four independent replication origins (RK2, pBBR1, pRO1600/ColE1, RSF1010), and diverse cargoes, can be combined in a single host cell to fulfill the user's needs. A rich collection of constructs is already freely available (http://seva.cnb.csic.es/).

A set of compatible SEVA plasmids was used, for instance, in a study of Nikel and de Lorenzo (2013), who described extrication of *P. putida* KT2440 from its strictly aerobic nature, which prevents its use in industrial anaerobic fermentors and certain bioremediation applications (Nikel and de Lorenzo, 2013). Two scenarios were considered to explain the inability of *P. putida* to grow in anoxic conditions, (i) unbalanced energy charge due to limited activity of the respiratory chain, and (ii) lack of appropriate pathways for anoxic NADH re-oxidation. This two-fold problem was tackled by recruiting the pyruvate decarboxylase (*pdc*) and alcohol dehydrogenase II (*adhB*) genes from the anaerobe *Zymomonas mobilis* and the acetate kinase (*ackA*) gene from the facultative aerobe *E. coli*, to manipulate energy generation and redox balance in anoxic conditions. To evaluate the potential of the recombinant host for anoxic biotransformations, the authors used it as a host for the haloalkane dehalogenase genes from *P. pavonaceae* strain 170. The resulting recombinant, which bore two synthetic operons on compatible plasmids pSEVA234 and pSEVA428, not only survived in anoxic conditions, but also degraded the environmental pollutant 1,3-dichloropropene. These results highlight the possibility of harnessing the full potential of *P. putida* KT2440 as a robust biocatalyst by precise control of its energy and redox metabolism.

Despite the widespread use of recombinant plasmids in proof-of-concept studies, metabolic engineers must bear in mind that the introduction and massive expression of heterologous genes can affect host fitness, due either to additional metabolic load or to depletion of essential cofactors in redox reactions (Dvorak et al., 2015; Glick, 1995; Wu et al., 2016). This is even more important in the design of robust bacterial degraders that must cope with a spectrum of toxic substrates and/or pathway intermediates, and might need an additional energy charge and reducing cofators to cope with oxidative stress (Dvorak et al., 2015; Nikel et al., 2013). By-passing such difficulties is possible by tuning inducer concentration, applying lower-copy-number plasmids with weaker promoters, or by enzyme-mediated co-factor recycling through overexpression of NAD^+ kinase, transhydrogenases or dehydrogenases (Dvorak et al., 2015; Nikel et al., 2016b).

As an alternative, expression of heterologous genes directly from the host chromosome can be beneficial for greater stability of the desired genotype/phenotype and improved host viability (Martínez-García et al., 2014a; Santos and Yoshikuni, 2014; St-Pierre et al., 2013; Zobel et al., 2015). Homologous recombination and site-specific or random transposition-based techniques have been developed to enable chromosomal insertion of single genes, gene clusters or whole synthetic operons, or knockouts of genes that encode competing metabolic pathways (Loeschcke et al., 2013; Martínez-García et al., 2014a; Santos and Yoshikuni, 2014; St-Pierre et al., 2013; Zobel et al., 2015).

In a study by Samin and co-workers, a mutant variant of haloalkane dehalogenase DhaA was introduced into the chromosome of 2,3-dichloro-1-propanol-degrading *P. putida* MC4 under the control of a strong constitutive promoter using the Tn5 transposon-based delivery vector (Samin et al., 2014). Although the complete mineralization pathway in strain MC4 is not yet fully understood, the recombinant *Pseudomonas* was the first microorganism shown to grow aerobically on TCP, both in shaken flasks and in packed-bed reactor in continuous-flow conditions. Mini-transposon systems have also been used successfully for chromosomal integration of the whole catabolic cluster, as in the study by Wittich and Wolf who transplanted 11 genes from three distinct bacteria into *Cupriavidus necator* H850, giving rise to the first designer bacterium to show aerobic growth on a wide range of PCB, including the two commercial pesticides Aroclor 1221 and Aroclor 1232 (Wittich and Wolff, 2007).

Gene expression balancing is undoubtedly a powerful metabolic engineering concept. In addition to the approaches mentioned above for rather rough manipulation of gene expression, numerous techniques for fine-tuning expression have been developed in recent years. Surgical cuts, including computationally designed ribosome binding sites, refined architecture of the whole expression cassette, mutagenesis of promoter regions, tailored stability of mRNA molecules, or altered gene order in the operon were proven useful for optimizing bisynthetic pathways (Boyle and Silver, 2012). In several cases, expression adjustment was also considered in biodegradation pathway design. de la Peña Mattozzi and co-workers achieved more rapid paraoxon hydrolysis in a recombinant *P. putida* strain by altering the order of three genes in the operon for enzymes that convert one of the initial pathway intermediates (de la Peña Mattozzi et al., 2006). Expression of the *dszB* gene, which encodes 2-hydroxybiphenyl-2-sulfinate sulfinolyase, the last rate-limiting enzyme in the dibenzothiophene biodesulfurization pathway, was improved by rearranging gene order and removing the overlapping structure in the *dsz* operon (Li et al., 2007). Nonetheless, especially in cases of synthetic metabolic pathways composed of biocatalysts from diverse organisms, the activity, selectivity, stability or inhibition bottlenecks of individual enzymes can be too far-reaching to be solved by tuning expression of the corresponding genes. Moreover, overexpression of endogenous or exogenous genes often results in a metabolic burden and lower host viability due to overconsumption of metabolic precursors (amino acids, rRNA, ATP, reducing cofactors) to fuel the synthesis of non-essential proteins (Glick, 1995; Wu et al., 2016). Protein engineering of bottleneck enzymes then becomes a more reasonable approach.

2.3.3. Protein engineering to eliminate bottlenecks of biodegradation pathways

Over the last two decades, three distinct strategies were developed to allow the construction and identification of mutant enzymes with desirable properties (Bornscheuer et al., 2012). The earliest approach, *rational design*, exploits various computational techniques such as, molecular docking, homology modelling and molecular dynamics simulations, together with site-directed mutagenesis protocols to generate targeted mutations that result in single or several modified variants of an enzyme (Damborsky and Brezovsky, 2014). In contrast, *directed evolution* uses random mutagenesis methods and is beneficial especially in cases when neither the structure nor the catalytic mechanism of the enzyme is available (Brakmann, 2001). Random methods of directed evolution are combined with elements of rational enzyme modification to bypass certain limitations of both approaches. This focused directed evolution approach targets several specific residues or certain protein regions selected on the basis of prior structural and functional knowledge (Bornscheuer et al., 2012). Mutation hot spots are chosen experimentally and, on a much larger scale, computationally, using powerful new algorithms and statistical tools such as HotSpot Wizzard or 3DM tool (Bendl et al., 2016; Damborsky and Brezovsky, 2014; Kuipers et al., 2010). A clear trend in recent years is the introduction of protein engineering into the metabolic engineering workflow.

The activities of many enzymes that act on anthropogenic compounds are derived from promiscuous activities that are generally very inefficient (Khersonsky and Tawfik, 2010). In such cases, protein

engineering is the only possible solution. Numerous reports from the first decade of the 2000s describe application of common protein engineering methods such as error-prone PCR, DNA shuffling, site-directed or saturation mutagenesis for engineering activity, or selectivity of individual catabolic enzymes to halogenated hydrocarbons (Parales and Ditty, 2005; Wittich et al., 2010). Recently, the new challenge of accumulating plastic waste has attracted the attention of protein engineers, who try to enhance the activities of certain bacterial or fungal enzymes such as esterases, lipases or polyster hydrolases to synthetic polymers, including polyethylene terephthalate (PET) or polyurethanes (Wierckx et al., 2015). For instance, Wei and colleagues applied a comparison of crystal structures and molecular docking combined with side-directed mutagenesis to improve the activity of cutinase TfCut2 from *Thermobifida fusca* KW3 towards PET at elevated temperatures that promote degradation of this oil-derived plastic (Wei et al., 2016). Subsequent kinetic and *in silico* energetic analyses confirmed that the improvement in PET hydrolysis was the result of a relief of product inhibition caused by single point mutation in the enzyme's surface-exposed active site. New unique PETase and MHETase, shown to help bacterium *Indonella sakaiensis* to depolymerize and utilize PET, are promissing candidates for protein and metabolic engineering exercises that might lead to biotechnological recycling of the polymer (Bornscheuer, 2016; Yoshida et al., 2016a). Alas, such a promise is not devoid of controversy about the very enzymes that could do the job (Yang et al., 2016; Yoshida et al., 2016b)

The studies reporting the implantation of constructed mutants in the context of the whole biodegradative pathway are nonetheless rather rare. Iwakiri and co-workers successfully applied *Alcaligenes sp*. KF711 harbouring engineered monooxygenase $P450_{CAM}$ for the dehalogenation of pentachloroethane to trichloroethane in anoxic conditions (Iwakiri et al., 2004). In another study, promiscuous toluene *ortho*-monooxygenase and epoxide hydrolase were engineered using DNA-shuffling and saturation mutagenesis; the use of mutant genes allowed more rapid aerobic degradation of chlorinated ethenes, with less accumulation of stress-inducing intermediates in recombinant *E. coli* bearing the synthetic pathway (Lee et al., 2006; Rui et al., 2004).

Mutants of DhaA, the enzyme that initiates dehalogenation of TCP, were used in three of the previously mentioned studies that focused on TCP pathway engineering (Dvorak et al., 2014b; Kurumbang et al., 2014; Samin et al., 2014). The DhaA31 variant, with 29-fold improved catalytic efficiency towards the chlorinated substrate, was prepared by combining molecular dynamics simulations of product release from the buried active site with site-directed and saturation mutagenesis in selected hot spots (Pavlova et al., 2009). Another promising mutant with the potential to remove the second serious bottleneck in the synthetic TCP pathway—enantioselectivity of DhaA and high enantiospecificity of HheC—originated from the work of van Leeuwen et al. (2012). The authors targeted DhaA31 to obtain variants that convert TCP predominantly into (*R*)-DCP, which can be further converted by HheC at a much higher rate than the (*S*)-enantiomer. Five rounds of focused directed evolution using saturation mutagenesis with restricted codon sets provided mutant DhaA r5-90R with 13 new amino acid substitutions, and substantially improved (*R*)-selectivity. Mutagenesis nevertheless affected the activity with TCP, which dropped to wild-type DhaA levels; the mutant was later shown in *in vitro* and *in vivo* studies to be of no value for further improvement of TCP pathway efficiency (Dvorak et al., 2014b; Kurumbang et al., 2014). Additional engineering input is thus needed to obtain DhaA variants with improved activity and modified enantioselectivity that would promote smooth flux through the synthetic pathway (Fig. 4) and allow growth of bacterial recombinants in minimal medium with the toxic substrate (Kurumbang et al., 2014).

It will be necessary to accelerate the processes of laboratory evolution and screening of new enzymes with enhanced properties in mutant and metagenomic libraries to make protein engineering of greater value in optimizing natural and synthetic catabolic pathways (Bouhajja et al., 2016). Selection couples an improved enzyme property with host survival, which allows even $> 10^9$ clones/enzyme variants to be tested in a reasonable time. New selection assays based on toxic substrate conversions into a harmless utilizable metabolite, coupled with fluorescence-activated sorting of surviving cells could, for example, be applied for the enrichment of new dehalogenase variants from libraries prepared by directed evolution or a semi-rational approach (Fernández-Álvaro et al., 2011; Fibinger et al., 2015). Several recent studies presented workflows for high-throughput screening and characterization of improved enzyme variants or novel enzyme activities. These test schemes are based on: (i) bioinformatic pre-screening combined with high-throughput experimental characterization of candidate proteins (Bastard et al., 2014), (ii) completely automated robotic platforms manipulating clones growing in microtitre plates (Dörr et al., 2016), (iii) microcapillary arrays coupled with fluorescent assays and laser extraction for recovery of live clones (Chen et al., 2016), or (iv) single-cell sorting based on a fluorescent signal from clones with an implemented synthetic genetic circuit responding to a specific metabolite (Choi et al., 2014). *In silico* screening of molecules that fit the active site cavity of an available enzyme structure, or reverse screening of enzymes that can accommodate target substrate could supplement experimental methods and reduce the time needed for discovery of new biocatalysts for recalcitrant chemicals. Indeed, the potential of *in silico* screening methods was demonstrated when investigating cytochrome P450-mediated metabolism of xenobiotics (Raunio et al., 2015). In a recent study by Aukema and colleagues, docking and molecular dynamics simulations allowed selection of biphenyl dioxygenase from *Paraburkholderia xenovorans* LB400 as the best candidate enzyme for metabolizing carbamazepine, one of the most commonly identified recalcitrant pharmaceuticals in rivers (Aukema et al., 2016). The experimentally verified rate of carbamazepine degradation by *P. xenovorans* cells was 40-times greater than the best reported rates so far.

As shown in the preceding sections, the available toolbox of systems biology, metabolic engineering and protein engineering applicable for re-factoring microbes and their catabolic pathways towards more efficient biodegradation of waste chemicals and recalcitrant pollutants is becoming inordinately large. Even so, the defined workflows, standards and universally applicable principles for strain optimization were long missing in the field of metabolic engineering, despite which, the aim to engineer living systems on a rational basis remained prevalent in the community. This goal led metabolic engineers to adopt standards and strategies from another closely related field of synthetic biology, that was established primarily to program living systems with a high predictability.

3. Synthetic biology approaches and tools for biodegradation pathway engineering

The principal underlying thought in synthetic biology is that any living system can be considered a set of separate usable components that can be combined by the means of biological engineering in new arrangements to alter existing features or generate new ones (de Lorenzo and Danchin, 2008). Such biological engineering can be simplified by applying principles adopted from electronic engineering and computer science to produce predictable, robust systems (genetic control systems, metabolic pathways, chromosomes and whole cells) with non-natural functions (Paddon and Keasling, 2014). This would be achieved through the fabrication of thoroughly characterized, standardized, recyclable parts. From its very beginnings, synthetic biology has been confronted with the need to define its practice precisely and to develop clear, generally applicable strategies, as this was the only way to guarantee and implement engineering principles in a field as stochastic as biology. With its tools, standards and the DBTA cycle strategy, synthetic biology has contributed as has no other scientific discipline to rationalize metabolic engineering. Its bottom-up approaches and synthetic devices are being implemented to make more robust, better controllable microbial cell factories. Synthetic biology

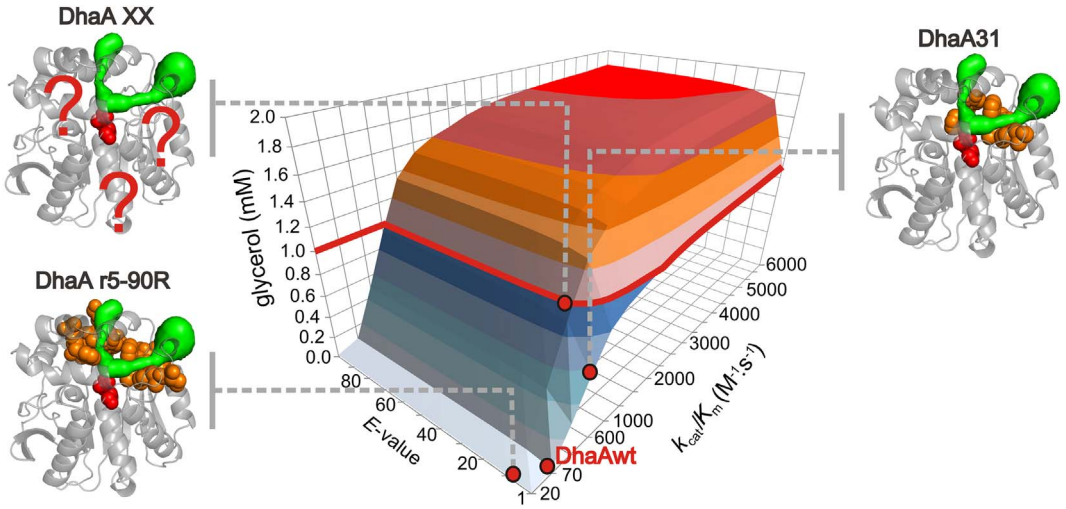


Fig. 4. Hypersurface plot describing the effect of catalytic efficiency (k_{cat}/K_m) and enantioselectivity (*E*-value) of haloalkane dehalogenase DhaA on the production of glycerol in the TCP pathway. The hypersurface was calculated using the mathematical model of the TCP pathway (Dvorak et al., 2014b) using the constraints of Kurumbang and co-workers (Kurumbang et al., 2014). Positions of four DhaA variants, DhaAwt (wild type, k_{cat}/K_m = 70 M^{-1} s^{-1}, *E*-value = 1), DhaA r5-90R (k_{cat}/K_m = 20 M^{-1} s^{-1}, *E*-value = 10), DhaA31 (k_{cat}/K_m = 600 M^{-1} s^{-1}, *E*-value = 1), and hypothetical DhaAXX (k_{cat}/K_m = 700 M^{-1} s^{-1}, *E*-value = 10), are indicated by red dots. Note that the majority of mutations (orange spheres) introduced into DhaA r5-90R and DhaA31 are adjacent to the active site (red spheres) and to the access tunnels (in green). Hypothetical enzyme DhaAXX with 10-fold improved catalytic efficiency and 10-fold improved enantioselectivity compared to DhaAwt, would enable production of 1 mM glycerol (red line) from 2 mM TCP in the culture of recombinant *E. coli* BL21(DE3) within 24 h time interval. Such glycerol concentration will be sufficient to support the cell growth. Figure was adopted and modified from (Kurumbang et al., 2014). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

can, and to some extent already contributes, the same service with its new concepts to the fields of biodegradation and bioremediation. In the the Section 3, we will discuss examples of beneficial inclusion of synthetic biology into strategies for engineering pathways and whole cells for biodegradation of waste and polluting chemicals, and perspectives for these initiatives.

3.1. Development of robust microbial chassis for biodegradation of toxic chemicals

Genetic instability and negatively affected fitness are frequently encountered drawbacks of the native microbial hosts tailored by metabolic engineering for specific biodegradation purposes. Synthetic biology offers possible solutions for these problems by implementing DNA synthesis and genome editing strategies. Both curiosity and prospective practical applications drive synthetic biologists to generate microbial cells with deleted parts of their genomes that encode redundant, cryptic, or even deleterious functions. Such cells, endowed with a reduced genome as foundation to house and support heterologous genetic parts, are collectively known as *chassis* (Adams, 2016). An extreme case of a *chassis* is the so-called *minimal cell* (Glass et al., 2006). Chassis can be prepared by *de novo* synthesis of a reduced target genome and its implantation into a suitable cell envelope, or through systematic deletions of non-essential genes in the genomes of an existing natural host. Microarray-based oligonucleotide synthesis is currently used to provide a substrate (usually 5–50 oligos) for construction of larger (usually 200–3000 bp) synthetic fragments (Kosuri and Church, 2014). Scarless methods including the popular Gibson assembly, uracil assembly, Golden Gate technology, ligase cycling reaction, or yeast recombination are used to combine sequence-verified gene-length fragments in even larger complexes (Casini et al., 2015). DNA synthesis also allows preparation of standardized genetic parts (promoters, ribosome binding sites, genes, terminators and so on) with verified codon-optimized sequences. Despite breakthroughs in gene synthesis and DNA assembly methods, this last approach for *chassis* construction seems far more feasible at the moment, considering the lesser demand of chromosome editing experiments, the constantly expanding portfolio of genetic tools, and the growing list of microorganisms with intentionally reduced genomes (Martínez-García and de Lorenzo, 2016; Si et al., 2015).

Discarding non-essential cell functions shows considerable potential not only for biosynthetic purposes (Hutchison et al., 2016), but also for designer biodegradation and bioremediation. A recent report on the systematic deletion of 11 non-adjacent genomic regions in *P. putida* KT2440 (Fig. 5) is a unique example of genome streamlining in a popular bacterial host with well-defined biodegradation capabilities (Martínez-García et al., 2014b). In all, 300 genes were eliminated, that is, 4.3% of the entire KT2440 strain genome, using the scar-less deletion procedure based on homologous recombination after *in vivo* DNA cleavage by the homing *Saccharomyces cerevisiae* nuclease I-*SceI* (Martínez-García and de Lorenzo, 2012). A suite of functions was targeted, including the complete flagellar machinery, whose assembly and function drains ATP from the cells and consumes NAD(P)H. Four prophages, two transposons, and three components of DNA restriction-modification systems were also eliminated to minimize genetic instability. The resulting strains designated *P. putida* EM42 and EM383 (the latter lacks the *recA* gene that encodes recombinase A) showed clearly superior growth properties and improved overall physiological vigour compared to wild-type KT2440. Moreover, due to the higher $NADPH/NADP^+$ ratio, the reduced-genome strains also better tolerated endogenous oxidative stress, a property that provides a crucial advantage for catalysing harsh biodegradation reactions such as aerobic dehalogenation of chlorinated pollutants (Nikel et al., 2013).

Despite their reliability and robustness, the procedures based on these genome editing tools are time- and labour-intensive, and restricted to a limited number of model microorganisms. This problem is now being challenged by new protocols that profit from the type II bacterial Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated protein (Cas), Multiple Automated Genome Editing (MAGE), and combinations thereof (Barrangou and van Pijkeren, 2016; Wang et al., 2009). Nonetheless, these technologies also suffer from some weaknesses such as generation of numerous off-target mutations. The off-target problem of the CRISPR-Cas system was recently mitigated by engineering high-fidelity Cas9 nucleases (Kleinstiver et al., 2016). In a similar manner, the number of off-target mutations during MAGE was reduced substantially when temperature-sensitive control of the endogenous methyl-directed mismatch repair system was introduced into the *E. coli* host together with a dominant

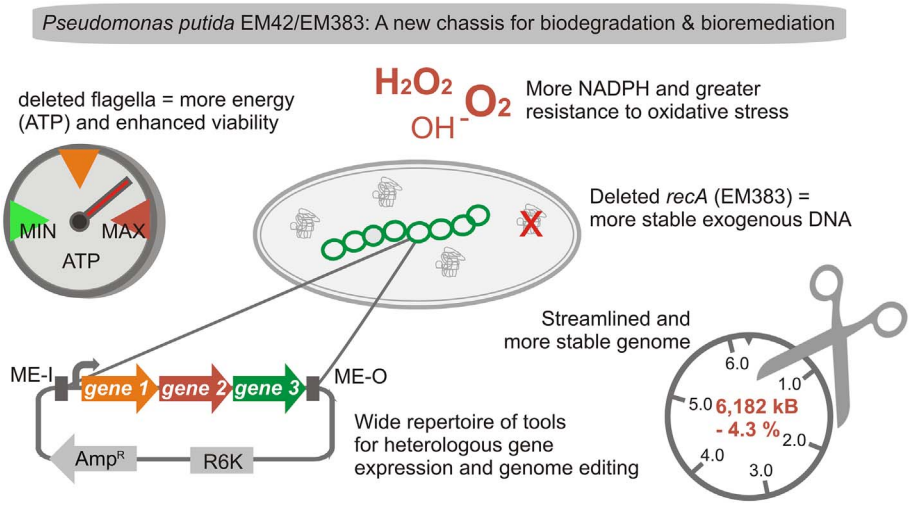


Fig. 5. Engineered *Pseudomonas putida* KT2440 as a new robust chassis for biodegradation and bioremediation. Martínez-García and colleagues used in-house genome editing tools to streamline the chromosome of *Pseudomonas putida* KT2440 (Martínez-García et al., 2014b). Deletion of 4.3% of the original genome gave rise to the new platform strains EM42 and EM383, whose advantageous properties make them useful chassis for engineering biodegradation pathways and other applications.

negative mutator allele of the *E. coli* mismatch repair protein MutL (Nyerges et al., 2016). By placing this highly conserved allele together with λ Red recombinase genes and a temperature-sensitive λ repressor on a single broad-host range plasmid, Nyerges and co-workers provided a simplified MAGE variant (pORTMAGE) that allowed genome editing and mutant library generation in several biotechnologically and clinically relevant bacterial species, with no need for prior modifications of parental strains. The recent discovery of the Ssr protein from *P. putida* DOT-T1E, a functional homologue of the β protein of the λ Red recombination system, paved the way for oligonucleotide-mediated multiplex genome engineering in pseudomonads, including *P. putida* KT2440 and its derived platform strains EM42 and EM383 (Aparicio et al., 2016). Continuing facilitated reduction of their genomes (deletions of multiple repetitive sequences or certain proteases) and genomes of other relevant bacterial hosts will provide more reliable, robust *chassis* for biodegradation pathway engineering.

3.2. Development of synthetic microbial consortia for enhanced biodegradation and bioremediation of pollutants

An alternative strategy that allows for the emergence of more robust microbe-based bioprocesses is just the opposite of the systematic development of bacterial *chassis* based on individual wild-type strains. Natural microbial consortia perform complicated biocatalytic tasks such as lignocellulose degradation, wastewater purification, or intestinal food digestion. By dividing the labour and metabolic burden, the consortium members become more flexible when facing complex environmental conditions. Distribution of catabolic capacity among cells of a single organism, among different bacterial strains, or even among species from different kingdoms is pivotal; this is especially true for biodegradation of complex toxic chemicals, which encompasses many steps and harmful intermediates with diverse physicochemical properties. The idea of exploiting consortia for directed biodegradation of polluting compounds and bioremediation of contaminated sites thus lies ready to hand.

Application of natural microbial consortia has already shown promise for bioremediation of sites polluted with anthropogenic pesticides such as diclofop methyl or atrazine (Baghapour et al., 2013; Wolfaardt et al., 1994). Together with the new field of synthetic biology, attempts are being made to modify the structure of original natural consortia or develop completely new artificial organizations of co-operating microorganisms. For example, co-culture of engineered *E. coli* SD2 and *P. putida* KT2440 was used successfully to mineralize the insecticide parathion in shaken flasks and in biofilm culture (Gilbert et al., 2003). In another study, complete mineralization of 2,4,6-tribromophenol, a flame-retardant intermediate and pesticide, was achieved using an artificial anaerobe consortium of reductive debrominator *Dehalobacter sp*. FTH1, a hydrogen supplier *Clostridium sp*. Ma13, and the 4-chorophenol-degrading strain *Desulfatiglans parachlorophenolica* DS (Li et al., 2015). Recently, Martínez and co-workers introduced an attractive alternative strategy for engineering the 4S pathway (Fig. 6A), a paradigmatic bioprocess for removal of the recalcitrant sulphur from aromatic heterocycles in fuels (Martínez et al., 2016). The pathway encoded by *dszABCD* genes converts the model compound dibenzothiophene (DBT) into sulphur-free 2-hydroxybiphenyl (2HBP). The authors initially synthesized and re-arranged the original genes and optimized transcriptional and translational signals for *P. putida* KT2440. By dividing the pathway into three separate modules expressed individually, they were able to define previously unreported bottlenecks in the pathway; that is, the inhibitory effect of intermediate 2HBP-sulfinate (HBPS) on DszA and DszC monooxygenases and HBPS leakage into the culture medium, from which it could not be transported back into the cells for further processing. To prevent HBPS accumulation, two *P. putida* strains bearing *dszC1-D1* or *dszB1A1-D1* modules were mixed as a suspension of resting cells. When the strains were combined at a 1:4 ratio and added with cell-free extract containing extra DszB, almost 100% of DBT was transformed into 2HBP in assay conditions.

Considerable attention is currently being dedicated to developing strategies for biodegradation of oil and oil-derived chemicals. Patowary and co-workers designed a potent microbial consortium for prospective decontamination of sites exposed to polycyclic aromatic hydrocarbons from crude oil (Patowary et al., 2016). The authors first inspected the biodegradative capacity for total petroleum hydrocarbons of 23 bacterial isolates from petroleum-contaminated soils, and subsequently designed 14 artificial consortia based on the five most efficient isolates. The best designer consortium, comprised of two biosurfactant-producing, hydrocarbon-degrading strains of *Bacillus pumilus* and *B. cereus*, showed up to 84% degradation of total petroleum hydrocarbons after five weeks, as verified by gravimetric, FTIR and GC/MS analyses. Application of natural or re-designed bacteria and their consortia also holds considerable promise for decomposition of oil-derived plastic waste (Skariyachan et al., 2016; Yoshida et al., 2016a, 2016b).

New tricks for more efficient biodegradation strategies could be adopted from inspiring current research on synthetic consortia designed to decompose and valorize lignocellulosic biomass (Minty et al., 2013). Tozakidis and co-workers reported application of an engineered consortium of three *P. putida* strains displaying thermophilic endoglucanase, exoglucanase, or β-glucosidase for concerted hydrolysis of cellulose at elevated temperatures (Tozakidis et al., 2016). Some studies describe surface display of whole designer cellulosomes – synthetic enzymatic nanomachines whose natural counterparts are produced by certain cellulolytic bacteria that exploit clustered surface-attached

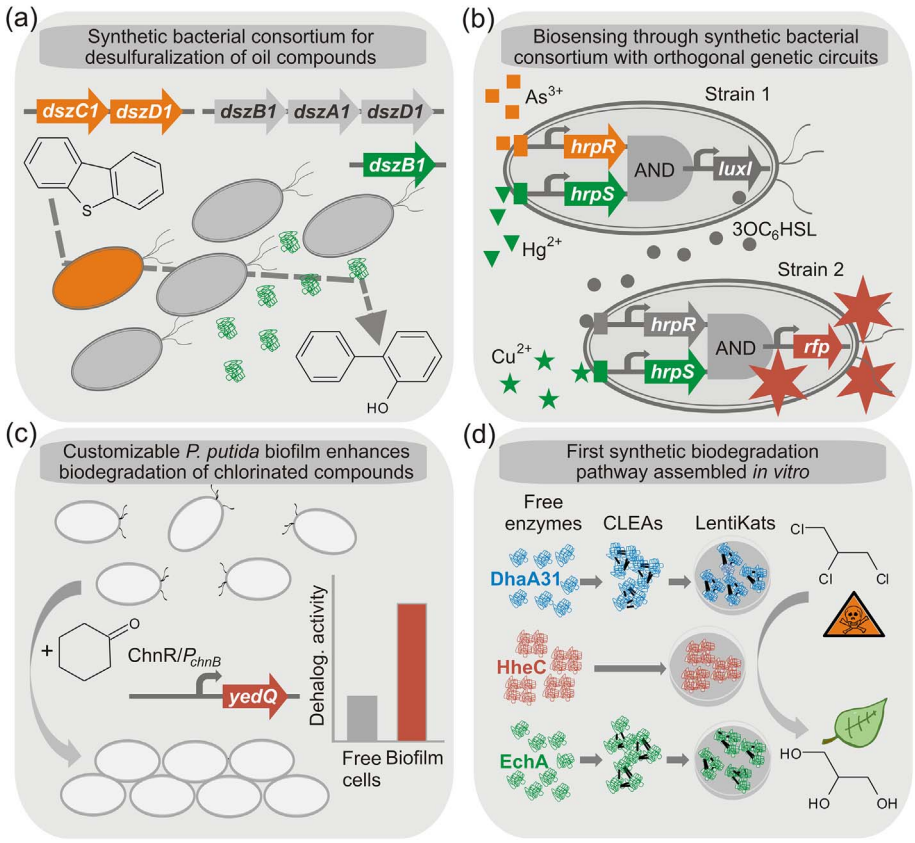


Fig. 6. Examples of synthetic biology approaches applied to engineering biodegradation pathways and whole-cell degraders. (a) Martínez and co-workers divided the 4S pathway for sulphur removal from aromatic heterocycles into two modules (*dszC1-D1* and *dszB1A1-D1*) that were expressed individually in two *P. putida* KT2440 strains to form a synthetic consortium (Martínez et al., 2016). Combining the strains at a 1:4 ratio and mixing them with cell-free extract from the third strain producing DszB resulted in almost complete conversion of dibenzothiophene into sulphur-free 2-hydroxybiphenyl in assay conditions. (b) Wang et al. constructed a synthetic bacterial consortium in which each of the two *E. coli* strains functioned as a double input sensor with synthetic AND gates for detecting arsenic, mercury, and copper ions and *quorum sensing* molecules ($3OC_6HSL$; Wang et al., 2013). Strain 2 generated red fluorescence only in the presence of all three metal ions in culture. (c) Benedetti and colleagues achieved directed transition between planktonic and biofilm lifestyles of *P. putida* KT2440 by controlling cyclic di-GMP levels with an orthogonal genetic device composed of the *yedQ* diguanylate cyclase gene from *E. coli* under the control of the cyclohexanone-responsive ChnR/P_{chnB} regulatory node from *Acinetobacter johnsonii* (Benedetti et al., 2016). Designed cells with a synthetic operon for 1-chlorobutane biodegradation formed inducible biofilms with higher dehalogenase activity than free, planktonic bacteria. (d) Dvorak and co-workers assembled the first *in vitro* synthetic biodegradation pathway for 1,2,3-trichloropropane (TCP) by combining engineered haloalkane dehalogenase DhaA31, halohydrin dehalogenase HheC, and epoxide hydrolase EchA, immobilized in cross-linked enzyme aggregates (CLEAS) and polyvinyl alcohol particles (LentiKats) (Dvorak et al., 2014a). The immobilized pathway converted high concentrations of toxic TCP into glycerol in contaminated water for more than two months of continuous operation in a bench-top packed bed reactor. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

cellulases for efficient depolymerization of cellulose and hemicellulose (S. Kim et al., 2013; Moraïs et al., 2014). The vast majority of the work that exploits surface display systems for whole-cell removal of environmental contaminants has been restricted to individual proteins such as laccases, methyl parathion hydrolase, or triphenylmethane reductase (Gao et al., 2014; Liu et al., 2016; Yang et al., 2012). It is tempting to expand available surface display technologies with cellulosome parts for the design of synthetic consortia that would catalyse more complex biodegradation reactions, including polymeric substrates or toxic metabolites, outside the cells.

Despite their true biotechnological potential, microbial consortia remain too complex for directed long-term application, either for biosynthesis or for biodegradation. It is still difficult to engineer homeostasis and evolutionary stability in an artificial multi-cellular system, whose behaviour becomes unpredictable with time (Escalante et al., 2015). Attempts to adopt principles of intercellular communication for synthetic consortia are just at their beginnings (Hennig et al., 2015; Scott and Hasty, 2016). Intensive co-operation of genetic engineers and evolutionary biologists with microbial ecologists and chemists will obviously be needed to derive a deeper understanding of the processes that rule the formation and survival of natural microbial consortia. Lessons learnt from nature can be implemented and combined with completely anthropogenic orthogonal genetic devices for more reliable design of stable multicellular systems.

3.3. Development of orthogonal systems in bacteria to enhance pollutant biodegradation and bioremediation

Orthogonal, or in other words parallel, independent systems[1] is a key concept in synthetic biology as long as such quality makes live systems more amenable to *bona fide* engineering. Orthogonalization involves *e.g.*, use of unnatural genetic codes (for example, quadruplet rather than triplet), alternative transcription-translation machineries, toggle switches, and genetic circuits for assembly of novel metabolic and signalling pathways from proteins containing non-natural amino acids (An and Chin, 2009; Wang et al., 2012). Such metabolic pathways could have new functions and would synthesize or catabolize a fresh spectrum of compounds. Ideally, orthogonal pathways/modules interact minimally with their natural counterparts in the host cell. This improves predictability of the component behaviour and prevents inhibitory cross-talk following introduction of the pathway or circuit into an existing metabolic network (Kim and Copley, 2012; Kurumbang et al., 2014). In an orthogonal module the input and output molecules of genetic circuits—or enzymes and metabolites of synthetic biochemical routes—are not degraded or inactivated by side-reactions or physico-chemical conditions of the host cell. While complete orthogonality is virtually impossible in biological systems, a practicable level of context-independence is indeed feasible. Some examples of claimed orthogonal systems include simple synthetic genetic circuits, switches, or designed metabolic pathways. These have the potential to improve function of individual engineered bacteria or synthetic consortia as whole-cell biosensors and degraders of recalcitrant chemicals.

Sensitivity, robustness and applicability of microbial biosensors for detecting heavy metals or halogenated hydrocarbons can be greatly improved by implementing orthogonal genetic devices (Bereza-Malcolm et al., 2015). Thus far, simple microbial biosensor designs based on one input and two-part regulatory systems with promoter and reporter providing a coloured, fluorescent, bioluminescent, or electric signal prevail in the literature (Ravikumar et al., 2017). Reports on more complex multi-input systems based on Boolean logic gates are still rare, but give us a clue to future developments in the field. As first outlined by the work of Wang et al. (2016), orthogonally acting multi-

[1] The mathematical and geometrical concept of *orthogonality* has been adopted first by computing science and then by synthetic biology to signify operative independence. System A is orthogonal to system B if A does not influence B—and *vice versa*.

input/multi-output logic gates will be especially valuable for detecting mixtures of polluting chemicals (organic, inorganic, or both) by providing a specific signal for each of the constituents (He et al., 2016; Wang et al., 2011). Developing such biosensors is of practical importance, as polluted environments are usually burdened with numerous contaminants that have diverse physicochemical properties. The authors engineered a set of two-input *E. coli*-based biosensors with synthetic AND gates using signalling sensory modules from native two-component signal transduction pathways, as well as *hrpR* and *hrpS* genes along with their HrpL promoter element from *P. syringae*. Using these sensors, they detected arsenic, mercury, copper, and zinc ions, as well as quorum sensing molecules in aqueous environment (Wang et al., 2011). The most advanced design resulted in a triple-input AND logic-gated biosensor that formed a synthetic bacterial consortium in which each of the two members acted as a double input sensor (Fig. 6B). In the presence of arsenic and mercury, the first strain formed a quorum sensing molecule ($3OC_6HSL$) that diffused freely into the medium and was sensed together with Cu^{2+} by the second strain, which provided a red fluorescent signal (Wang et al., 2013). The fluorescent response was seen only when all three metal ions were present in the culture.

Reliable orthogonal devices will be crucial for engineering artificial cell-cell communication, for instance, taking advantage of bacterial quorum sensing systems (Hennig et al., 2015; Scott and Hasty, 2016). Programming the dynamics of subpopulations of synthetic consortia that perform complex tasks will be only possible with parallel genetic circuits and signalling molecules unknown to the host cells (Chen et al., 2015). As stressed by Silva-Rocha and de Lorenzo (2014), catabolic pathways for recalcitrant and xenobiotic compounds could be a reliable source of wiring devices and molecules These pathways are frequently found in specific types of organisms, and metabolic crosstalk can thus be avoided by implementing components of such networks in suitable hosts. For instance, aromatic molecules such as benzoate or phenol can be used as signal inducers that are recognized by regulatory proteins that trigeger expression of target genes. The recent determination of the crystal structure of the sensory domain in the phenol-responsive transcription activator PoxR ilustrates how aromatics are sensed in bacteria (Patil et al., 2016). This study paves the way for wider application of rationally engineered transcription regulators responsive to aromatics in the design of synthetic genetic circuits.

The dynamic behaviour of the whole biodegradation pathway can also be better studied when the route is transferred into a distant host chassis where it acts orthogonally. The orthogonal nature of the synthetic pathway for TCP biodegradation in the surrogate host *E. coli* BL21 (DE3), described in Section 2.3.1., allowed a very good match between *in silico* predicted and *in vivo* metabolite concentrations and precise description of the bottlenecks (Kurumbang et al., 2014), as well as deciphering the contribution of metabolic burden and substrate/metabolite toxicity to the fitness cost of TCP biotransformation by whole-cell catalysts (Dvorak et al., 2015). Despite their benefits for such mechanistic and proof-of-concept studies, laboratory *E. coli* strains might not be optimally suited for the harsh conditions that accompany biodegradation and bioremediation processes (Adams, 2016; Nicolaou et al., 2010). Development of reliable *chassis* amenable to implantation of orthogonal genetic devices is therefore desirable. These new cell platforms should be based on robust environmental strains of *Pseudomonas*, *Rhodococcus*, *Deinococcus*, and other microorganisms with broad metabolic versatility and natural resistance to organic solvents or heavy metals (Adams, 2016).

One can also take advantage of other specific properties of environmental bacteria including formation of a robust biofilm. As shown recently by Benedetti and colleagues, the actual physical forms of whole-cell biocatalysts can be mastered using orthogonal genetic parts (Fig. 6C) through the so-called *synthetic morphology* approach (Benedetti et al., 2016). *Pseudomonads* generate biofilms with biophysical properties that depend on the species. Cyclic di-GMP (*c*-di-GMP) is the key signal molecule that rules the complex regulatory network mediating the transition between planktonic cells and biofilm formation. This feature was exploited to design an orthogonal genetic device for manipulating the native *c*-di-GMP biochemistry in *P. putida*, thereby controlling biofilm formation (Benedetti et al., 2016). The *E. coli yedQ* gene (encoding diguanylate cyclase, the enzyme that synthesizes *c*-di-GMP from GTP) was placed under control of a tightly regulated cyclohexanone-responsive expression system. A synthetic operon, encoding the enzymes needed for 1-chlorobutane biodegradation, was also introduced in the engineered, biofilm-forming *P. putida* strain. Upon activation of the corresponding genetic modules with appropriate inducers, the resulting *P. putida* biocatalyst displayed high dehalogenase activity in robust biofilms.

The modern approaches of biological engineering accelerate the development of whole-cell biocatalysts for *in situ* bioremediation or biosensing of emerging environmental pollutants. Field applications of genetically modified microorganisms are nonetheless restricted by current legislation and hindered by the unreliable behaviour of recombinant microbes in complex, fluctuating environments. Synthetic biology can help to tackle the major problems through the means described at the beginning of this chapter, *i.e.*, non-canonical genetic codes and xenobiochemistry, which can prevent diffusion of undesired sequences into the environmental gene pool, differentiate recombinants from their natural counterparts, and improve robustness and reliability of prepared genetic devices, whole-cell degraders and biosensors. The examples of such endeavours were recently reviewed extensively (Schmidt, 2010; Schmidt and de Lorenzo, 2016). These systems could be combined with bioluminiscence, fluorescence, or DNA watermarking technologies to track the environmental fate of designer degraders, or with inducible suicide systems for discarding degraders once their mission is completed (Liss et al., 2012; Liu et al., 2010; Paul et al., 2005). Yet the path to applied xenobiology is still crooked, and many ethical questions remain. Another concept from the portfolio of synthetic biology that will be discussed in the following section has the potential to fill the time gap until we obtain reliable whole-cell catalysts and alternative solutions also acceptable for GMO critics will be provided.

3.4. Engineering microbial biodegradation pathways in vitro

The major principle of cell-free synthetic biology is that purified biomolecules or components in crude cell extracts replace intact cells for constructing complex biomolecular systems (Hodgman and Jewett, 2012). To date, metabolic networks that encompass > 10 enzymes have been reconstructed *in vitro* (Rollin et al., 2015; Schwander et al., 2016). Cell-free metabolic pathways allow easy verification of biocatalyst function, determination of kinetic parameters, as well as evaluation of the network by kinetic modelling (Santacoloma et al., 2011). Note, however, that assemby of degradative pathways in cell-free systems (whether with purified enzymes or cell extracts) is not only to prototype and parametrize new routes, but also for direct use in the field (Karig, 2017). Despite the drawbacks that cell-free systems cannot self-propagate, that certain biomolecules can be sensitive to oxidizing environments, and that their encapsulation and large-scale production might be costly, these systems offer an appealing means to circumvent problems associated to GMO release to the environment. Apart from increasing safety, they offer other clear benefits. For instance, *in vitro* systems can operate in the presence of toxins that would inhibit or kill live cells, and metabolites, regulators and enzymes can be produced in optimized concentrations without interfering with cell components. Most important for release and predictability, the evolutionary opportunities of such non-live agents are zero, making the emergence of unexpected properties virtually impossible.

In vitro metabolic networks can suffer from suboptimal efficiency due to the lower enzyme concentration compared to the extremely dense cell cytoplasm (Hodgman and Jewett, 2012). This issue can be mitigated by enzyme immobilization and improved spatial organization

via synthetic protein or DNA scaffolds that reduce diffusion of pathway intermediates and promote substrate channeling (Siu et al., 2015). Alternatively, biocatalysts can be precipitated and covalently interconnected in cross-linked enzyme aggregate particles (Sheldon, 2011); this method was applied successfully to synthesize toxic nucleotide analogues by an immobilized five-enzyme synthetic pathway (Scism and Bachmann, 2010). The biochemical route described was completed with functional ATP regeneration, which demonstrates that *in vitro* technologies can cope with another possible drawback – limited cofactor recycling.

Thus far, mainly single immobilized enzymes or whole-cell biocatalysts have been exploited in biotechnological processes intended to remove polluting chemicals such as polyethylene terephthalate, 2,4-dinitrophenol, atrazine, or inorganic nitrates (Barth et al., 2016; Dehghanifard et al., 2013; Mutlu et al., 2015; Trögl et al., 2012). *In vitro* assays with cell-free extracts or purified enzymes, respectively, helped to identify an unknown anaerobic pathway for complete phthalate degradation to CO_2 in *Thauera chlorobenzoica* 3CB-1 or to decipher sequestration of the highly toxic intermediate tetrachlorobenzoquinone in pentachlorophenol degradation pathway in *Sphingobium chlorophenolicum* (Ebenau-Jehle et al., 2017; Yadid et al., 2013). Geueke et al. showed the benefits of an *in vitro* engineering approach in the model of γ-HCH biodegradation pathway (Geueke et al., 2013). They used a system of two purified enzymes, the dehydrochlorinase LinA and the haloalkane dehalogenase LinB, which initiate biotransformation of γ-HCH, a prohibited insecticide. They separately incubated five isomers that form technical HCH (which used to be applied frequently instead of pure γ-HCH) with various LinA:LinB ratios and determined metabolic profiles of sequential biotransformations. Analyses of these profiles helped determine the environmental fate of HCH isomers, and showed that the original HCH degradation pathway is optimized by evolution for γ-HCH, but not for other isomers that co-pollute contaminated sites.

The first report of *in vitro* assembly of a fully functional synthetic biodegradation pathway was published by Dvorak and co-workers (Dvorak et al., 2014a). The authors adopted the cell-free strategy using immobilized engineered haloalkane dehalogenase DhaA31, halohydrin dehalogenase HheC and epoxide hydrolase EchA for TCP biotransformation to harmless, valuable glycerol in contaminated water (Fig. 6D). The practical utility of bacterial recombinants that mineralize TCP (desccribed in Sections 2.3.1 and 2.3.2) is limited by substrate toxicity and by legislative barriers on the application of GMO (Kurumbang et al., 2014; Samin et al., 2014). Probing the function of the pathway in *in vitro* conditions therefore lied ready to hand. The authors successfully immobilized the pathway in the form of purified enzymes or cell-free extracts in cross-linked enzyme aggregates and lens-shaped polyvinyl alcohol particles. The immobilized pathway showed almost the same efficiency of TCP-to-glycerol conversion as mixture of free enzymes. Physicochemical properties of polyvinyl alcohol lentils allowed recovery from the reaction mixture and recycling of the pathway. In addition, the immobilized enzymes retained > 50% of their initial activities for over 2 months of continuous operation in a bench-top packed bed reactor. The study indicated that the immobilized route removes TCP from heavily contaminated water with a pollutant concentration (~1 g/L) that would be detrimental to the living degraders. One can anticipate that such conversions of toxic TCP to useful glycerol could even pave the way for prospective valorization of this waste chemical.

Although this biotechnology requires further validation and tuning, the work showed that *in vitro* assembly of natural or synthetic enzymatic pathways with engineered enzymes is a promising concept for the biodegradation of polluting compounds and toxic industrial waste products. Transport of such pathways *via* non-live carriers is related conceptually to efforts to produce and release complex therapeutic agents for human use. In both cases, their application demands the development of a sort of *environmental Galenic science* that enables supply of the engineered biological remedy when and where needed.

4. Towards planet-wide bioremediation interventions: CO_2 capture as a large-scale challenge

Environmental deterioration due to emissions of recalcitrant chemicals seems to pale when compared to the problem of global warming caused by the release of greenhouse gases that originate in human activities. The bulk of these gases comprise four natural molecules, carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O) and ozone (O_3), as well as one class of xenobiotics, the chlorofluorocarbons or CFC. Given the diluted, aereal and global nature of this problem, the strategies contemplated for tackling these chemical species differ considerably from those that centre on site-specific pollution issues. Most proposals thus far focus much more on reducing emissions than on capture and returning to innocuous chemical or mineral forms. All these compounds have their own natural biogeochemical cycle and, in theory, simply reducing their input into the biosphere should restore the original equilibrium (Rockström et al., 2017). Recent ecological thought nonetheless postulates that the impact of climate change in many ecosystems is already irreversible by natural means, and call for large-scale interventions that could help to resolve this impasse (de Lorenzo et al., 2016). In this case, what types of options could be considered from a systemic biology perspective?

Some incipient attempts have appeared recently, mostly in the field of engineering bacteria with a greater capacity for non-photosynthetic fixing of CO_2 (Erb and Zarzycki, 2016). One option is to use the six alternative pathways of autotrophic carbon fixation known in nature (Fuchs, 2011; Hicks et al., 2017) as a starting point to optimize such natural processes. Another possibilty is the invention of new routes by rewiring or modifying existing enzymes in a new configuration. A remarkable example of this solution is the work of Antonovsky et al. (2016), who showed that combining metabolic rewiring, recombinant protein expression, and laboratory evolution enables the production of sugars and other biomass constituents by a fully functional Calvin-Benson-Bassham (CBB) cycle in *E. coli*. In the bacteria designed in this way, carbon is fixed *via* a refactored CBB cycle, whereas reducing power and energy are derived from pyruvate oxidation (Fig. 7). It is interesting to note how evolutionary approaches adjust fluxes and generate connections between the designed CBB cycle and the host biochemical network, that cannot be predicted or designed.

A separate approach was recently developed by Schwander et al. (2016), who developed a synthetic route for continuous CO_2 fixation *in vitro* by composing a cycle formed of 17 enzymes that convert CO_2 into organic molecules. The cycle was draughted by metabolic retrosynthesis with enzymes from nine organisms from three domains of life, further optimized through rounds of protein engineering and metabolic proofreading. Although this completely designed pathway has not yet been implemented *in vivo* (*e.g.*, in a bacterial carrier) it is a question of time until it is done and propagated throughout an entire microbial community (de Lorenzo et al., 2016).

Finally, today's systemic approaches enable the invention of enzymes from scratch through *de novo* protein design followed by directed evolution, which opens perspectives for improving the natural state of affairs regarding fixation of CO_2 (and other greenhouse gases) as well as valorization of this waste as biotechnological feedstocks (Huang et al., 2016; Renata et al., 2015). A separate matter is the large-scale deployment of the microbial agents for CO_2 retrieval. Existing technology allows their implementation in biofilters and bioreactors that capture emissions at source. But how can we spread such desirable catalytic properties at a very large, even at planetary level to make a difference? Such technologies are not yet available—let alone that they raise a large number of safety questions. But they must necessarily be developed if we are not just to control greenhouse gas discharges, but also aim at their eventual reduction to pre-industrial levels (de Lorenzo et al., 2016).

Unfortunately, greenhouse gases are not the only globally widespread molecules to be concerned about. Much before global warming

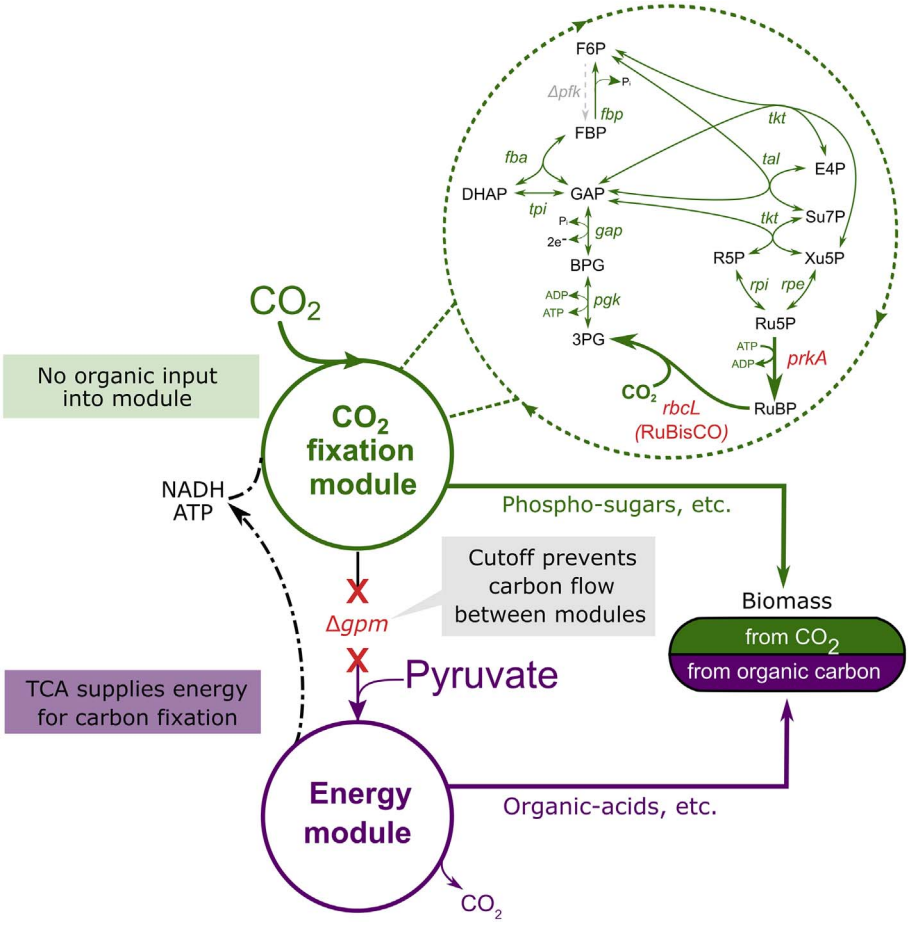


Fig. 7. Hemiautotrophic growth of engineered *E. coli* allowed by decoupling energy production and carbon fixation. The hemiautotrophic growth was observed after performing three distinct steps: (i) Calvin-Benson-Bassham (CBB) cycle for the biosynthesis of sugars from CO_2 was completed by introducing only two exogenous enzymatic activities, phosphoribulokinase (*prkA*) and RuBisCO, (ii) phosphoglycerate mutase genes *gpmA* and *gpmM* were deleted to separate central metabolism into two sub-systems, *i.e.*, a CO_2-fixing part that included PrkA, RuBisCO, upper glycolysis, and the pentose phosphate pathway, and an energy-supplying module that employed lower glycolysis and the TCA cycle, (iii) the engineered strain bearing further mutations (Δ*pfkA*, Δ*pfkB*, Δ*zwf*) was exposed to the strong selective pressure towards greater carbon fixation in xylose-limited chemostat with a surplus pyruvate and CO_2. Emerged mutations made designed CBB cycle functional and the hemiautotrophic strain that replaced utilisation of xylose with CO_2 fixation dominated the population. Figure adapted from Antonovsky et al. (2016).

associated to CO_2 emissions became so conspicuous, widespread pollution by endocrine disruptors (*e.g.*, alkylphenols, bisphenols, DDT, PCBs, polybrominated diphenyl ethers, phthalates and perfluorooctanoic acid) was well documented. Moreover, other *silent pollutants* originate in the pharmaceutical industry *e.g.*, antibiotics (Marathe et al., 2017) and hormones. They end up in a large variety of ecosystems in a diluted but still active form. Finally, plastics (*e.g.*, polycarbonate, polystyrene, polyethylene terephthalate, low/high-density polyethylene, polypropylene and polyvinyl chloride) are currently receiving a considerable focus as a major environmental problem. A number of microbial strains able to totally or partially act upon these compounds have been isolated (see above) but in general the biodegradation process is very slow. The global nature of these types of pollution thus requires the development of new, efficient microbial pathways for their elimination, quite a challenge for contemporary synthetic biologists and metabolic engineers (de Lorenzo, 2017). But—as is the case for CO2 capture mentioned above—pathways and hosts able to deploy them are not enough. Effective strategies for global dispersion of biodegradative activities of interest are badly needed—perhaps inspired in gene drives or some type of self-propagating, massive horizontal gene transfer (de Lorenzo et al., 2016).

5. Conclusions and perspectives

The practical examples described in the preceding sections show that technologies of systemic biology offer promise not only for the renaissance of bioremediation using anthropogenically enhanced microbial degraders, but also for entry to the new era of *Bioremediation 3.0*. The bottlenecks described at the beginning of this text are one by one being released or removed with the help of new computational and experimental tools. Databases and pathway prediction systems help the user to select suitable chemically and energetically feasible traits for desired tasks. Multi-omic analyses provide valuable information that facilitates choice of a suitable host. Genome-scale and kinetic models applied hand-in-hand with genetic engineering techniques enable tailoring of gene expression, reduction of metabolic burden, and pathway optimization in the context of host metabolism. The visionary concepts of synthetic biology and global scale CO_2 bioremediation open up new dimensions for engineering microbial degraders. General work schemes and standards for design, building, and fine-tuning of selected biochemical routes, networks and whole-cell biocatalysts are simultaneously becoming better defined and accepted by the community.

Despite the clear progress over the last decade of biochemical and biological engineering, the vast complexity of the living cell remains the major hurdle for any attempt for fully rational pathway design. In the case of biosensing and biodegradation pathway design and prospective applications of the genetically-modified microbes, this problem is exacerbated by the complexity of intercellular and interspecific interactions and by the still poorly understood interplay between the biotic and abiotic factors that govern contaminant biodegradation in polluted ecosystems (de Lorenzo, 2008; Meckenstock et al., 2015). We assume that, in principle, there are two ways to address this challenge, now and in future.

The first admits that complexity of biological systems will not be resolved anytime soon and thus bets on combining rational designs with natural evolution at the end of the DBTA cycle, thus expanding it to DBTAE (Design-Build-Test-Analyze-Evolve). Adaptive laboratory evolution (ALE) is a *high calibre* tool of this approach. ALE is now frequently used by academic and industrial laboratories to evolve desired properties in environmental bacteria and for fine-tuning recombinant microorganisms, but its principles have been well known since the time of the very first attempts to modify microbial phenotypes (Kellogg et al.,

1981). In biodegradation research, exposing microorganisms with desirable catabolic traits to the target chemical(s) for prolonged intervals in chemostat or shaken-flask cultures enabled isolation of bacteria able to use persistent compounds such as 2,4,5-trichlorophenoxyacetic acid (the defoliating Agent Orange), or helped improve atrazine degradation rates by *Pseudomonas* sp. ADP or 2,4-dichlorophenoxyacetic by *Ralstonia* sp. TFD41 (Devers et al., 2008; Kellogg et al., 1981; Nakatsu et al., 1998). Nonetheless, the potential of ALE in engineering microbial degraders has yet not been fully exploited. ALE will be an ideal tool for empowering rationally pre-designed recombinants with resistance to diverse environmental stresses encountered in directed biodegradation processes or in polluted ecosystems (Nicolaou et al., 2010; Oide et al., 2015; Wang et al., 2016). When combined with whole-genome sequencing and reverse engineering, ALE as an evolutionary "blind" approach can paradoxically contribute to further rationalizing the field by exposing otherwise unpredictable complex changes in multitrait phenotypes.

The second approach bets on technology without compromise, and will take advantage of continuing technological revolution in biology and related sciences. We recently witnessed the attempt of synthetic biologists to implement automation as we know it from electronic engineering in the design of orthogonal genetic circuits that would coordinate the actions of selected microbial chassis (Nielsen et al., 2016). The genetic "guts" and pathway enzymes of the chassis can be designed and synthesized *de novo* or be tailored rapidly, with surgical precision, using new genome editing and protein engineering tools (Huang et al., 2016; Hutchison et al., 2016; Nyerges et al., 2016; Ran et al., 2013; Renata et al., 2015). Although the circuit design tools still work predominantly with simple logic gates, the robustness and utility of the available minimal cells is limited, and *de novo*-designed enzymes have low activity, we can assume that the efficiency, reliability and time demand of such exercises will be further improved by progress in computer science and artificial intelligence development that could crack the complexity of life. A purely technological approach would omit any unknowns from the DBTA cycle with help of reliable computational simulations, thus expanding it to DSBTA (Design-*Simulate*-Build-Test-Analyze) and providing space for construction of whole-cell catalysts perfectly tailored for specific biodegradation tasks.

The challenges of both approaches applied to directed *in situ* bioremediation encompass the development of test schemes that will expose engineered organisms to mimicked environmental conditions, including limiting carbon sources, lack of oxygen or redox recycling, mass transfer perturbations, or interspecific predation (Meckenstock et al., 2015). A critical issue will be the choice and establishment of new chassis organisms better suited to field biotransformations than the currently available laboratory strains in terms of resistance to harsh conditions including extreme pH, temperature, osmotic pressure or fluctuating concentrations of toxic chemicals (Adams, 2016). These new strains must simultaneously be reliable, amenable to genetic manipulation, have mapped genomes and, ideally, metabolic models at hand. The *in situ* bioremediation quests will be complicated by the varying nature of polluting compounds, their mixtures and transformation products, all of which can be present at a single site (Sarigiannis and Hansen, 2012). Future microbial degraders must therefore be designed to cope with several problem compounds in parallel, and not only in terms of degradation or mineralization. The succesful conversion of dangerous contaminants into usable molecules that can be further valorized by whole-cell or cell-free biocatalysts is a desirable starting point for the development of new technologies that extend standard concepts of biodegradation and bioremediation (Dvorak et al., 2014a; Wigginton et al., 2012). Microbe-driven trash-to-treasure conversions are known from established biomass valorization processes, but this concept can also be implemented in wastewater treatment, metal and plastic recycling, or in direct valorization of industrial by-product streams or captured greenhouse gases (Bornscheuer, 2016; Koutinas et al., 2014; Nancharaiah et al., 2016; Wierckx et al., 2015).

Even so, the major challenge to the entire field of engineered degraders is not related to continued progress in the development of improved technologies, which is inevitable, but rather to the dissemination and popularization of achievements already attained. Public environmental concerns and regulatory constraints have not just restricted field tests of genetically modified microbes, but also affect the quality of fundamental research and, by the same token, overall progress in the field. The exercises described in this review will only have a chance to become real enterprises if the general population is inspired by biotechnology and accepts it as an inseparable part of our daily routine, as it did for information technology. The new systemic disciplines that make life less puzzling as well as emerging do-it-yourself movement in biotechnology can undoubtedly contribute to a general change of attitude (de Lorenzo and Schmidt, 2017). Such change is vital, because our ability to minimize production of waste and polluting chemicals, remove existing waste or use it in bioprocesses to form value-added compounds will shape the fate of our society on a global scale.

Acknowledgements

PD is the holder of the Marie Sklodowska-Curie grant No. 704410 (FUTURE). The work in Authors' Laboratory was funded by the CAMBIOS Project of the Spanish Ministry of Economy and Competitiveness RTC-2014-1777-3 (MINECO), HELIOS Project of the Spanish Ministry of Economy and Competitiveness BIO 2015-66960-C3-2-R (MINECO/FEDER) and the ARISYS (ERC-2012-ADG-322797), EmPowerPutida (EU-H2020-BIOTEC-2014-2015-6335536), and Raft4Biotech (720776) contracts of the European Union. JD is supported also by the Czech Ministry of Education (LQ1605, LO1214, LM2015055), and Czech Grant Agency (GA16-06096S). The financial support of The Novo Nordisk Foundation to PIN is also gratefully acknowledged.

References

Adams, B.L., 2016. The next generation of synthetic biology chassis: moving synthetic biology from the laboratory to the field. ACS Synth. Biol. 5, 1328–1330. http://dx.doi.org/10.1021/acssynbio.6b00256.

Adetutu, E.M., Gundry, T.D., Patil, S.S., Golneshin, A., Adigun, J., Bhaskarla, V., Aleer, S., Shahsavari, E., Ross, E., Ball, A.S., 2015. Exploiting the intrinsic microbial degradative potential for field-based in situ dechlorination of trichloroethene contaminated groundwater. J. Hazard. Mater. 300, 48–57. http://dx.doi.org/10.1016/j.jhazmat.2015.06.055.

Alexander, M., 1999. Biodegradation and Bioremediation. Academic Press.

Amde, M., Liu, J.-F., Pang, L., 2015. Environmental application, fate, effects, and concerns of ionic liquids: a review. Environ. Sci. Technol. 49, 12611–12627. http://dx.doi.org/10.1021/acs.est.5b03123.

An, W., Chin, J.W., 2009. Synthesis of orthogonal transcription-translation networks. Proc. Natl. Acad. Sci. U. S. A. 106, 8477–8482. http://dx.doi.org/10.1073/pnas.0900267106.

Antonovsky, N., Gleizer, S., Noor, E., Zohar, Y., Herz, E., Barenholz, U., Zelcbuch, L., Amram, S., Wides, A., Tepper, N., Davidi, D., Bar-On, Y., Bareia, T., Wernick, D.G., Shani, I., Malitsky, S., Jona, G., Bar-Even, A., Milo, R., 2016. Sugar synthesis from CO_2 in *Escherichia coli*. Cell 166, 115–125. http://dx.doi.org/10.1016/j.cell.2016.05.064.

Aparicio, T., Jensen, S.I., Nielsen, A.T., de Lorenzo, V., Martínez-García, E., 2016. The Ssr protein (T1E_1405) from *Pseudomonas putida* DOT-T1E enables oligonucleotide-based recombineering in platform strain P. putida EM42. Biotechnol. J. 11, 1309–1319. http://dx.doi.org/10.1002/biot.201600317.

Arora, P.K., Bae, H., 2014. Integration of bioinformatics to biodegradation. Biol. Proced. Online 16, 8. http://dx.doi.org/10.1186/1480-9222-16-8.

Aukema, K.G., Escalante, D.E., Maltby, M.M., Bera, A.K., Aksan, A., Wackett, L.P., 2016. In silico identification of bioremediation potential: carbamazepine and other recalcitrant personal care products. Environ. Sci. Technol. http://dx.doi.org/10.1021/acs.est.6b04345.

Baghapour, M.A., Nasseri, S., Derakhshan, Z., 2013. Atrazine removal from aqueous solutions using submerged biological aerated filter. J. Environ. Health Sci. Eng. 11, 6. http://dx.doi.org/10.1186/2052-336X-11-6.

Bailey, J.E., 1991. Toward a science of metabolic engineering. Science 252, 1668–1675.

Barrangou, R., van Pijkeren, J.-P., 2016. Exploiting CRISPR-Cas immune systems for genome editing in bacteria. Curr. Opin. Biotechnol. 37, 61–68. http://dx.doi.org/10.1016/j.copbio.2015.10.003.

Barth, M., Honak, A., Oeser, T., Wei, R., Belisário-Ferrari, M.R., Then, J., Schmidt, J., Zimmermann, W., 2016. A dual enzyme system composed of a polyester hydrolase

and a carboxylesterase enhances the biocatalytic degradation of polyethylene terephthalate films. Biotechnol. J. 11, 1082–1087. http://dx.doi.org/10.1002/biot.201600008.
Bastard, K., Smith, A.A.T., Vergne-Vaxelaire, C., Perret, A., Zaparucha, A., De Melo-Minardi, R., Mariage, A., Boutard, M., Debard, A., Lechaplais, C., Pelle, C., Pellouin, V., Perchat, N., Petit, J.-L., Kreimeyer, A., Medigue, C., Weissenbach, J., Artiguenave, F., De Berardinis, V., Vallenet, D., Salanoubat, M., 2014. Revealing the hidden functional diversity of an enzyme family. Nat. Chem. Biol. 10, 42–49. http://dx.doi.org/10.1038/nchembio.1387.
Batth, T.S., Keasling, J.D., Petzold, C.J., 2012. Targeted proteomics for metabolic pathway optimization. Methods Mol. Biol. 944, 237–249. (Clifton NJ). http://dx.doi.org/10.1007/978-1-62703-122-6_17.
Belda, E., van Heck, R.G.A., José Lopez-Sanchez, M., Cruveiller, S., Barbe, V., Fraser, C., Klenk, H.-P., Petersen, J., Morgat, A., Nikel, P.I., Vallenet, D., Rouy, Z., Sekowska, A., Martins Dos Santos, V.A.P., de Lorenzo, V., Danchin, A., Médigue, C., 2016. The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. Environ. Microbiol. 18, 3403–3424. http://dx.doi.org/10.1111/1462-2920.13230.
Bendl, J., Stourac, J., Sebestova, E., Vavra, O., Musil, M., Brezovsky, J., Damborsky, J., 2016. HotSpot Wizard 2.0: automated design of site-specific mutations and smart libraries in protein engineering. Nucleic Acids Res. 44, W479–487. http://dx.doi.org/10.1093/nar/gkw416.
Benedetti, I., de Lorenzo, V., Nikel, P.I., 2016. Genetic programming of catalytic *Pseudomonas putida* biofilms for boosting biodegradation of haloalkanes. Metab. Eng. 33, 109–118. http://dx.doi.org/10.1016/j.ymben.2015.11.004.
Benfenati, E., 2007. Predicting toxicity through computers: a changing world. Chem. Cent. J. 1, 32. http://dx.doi.org/10.1186/1752-153X-1-32.
Bereza-Malcolm, L.T., Mann, G., Franks, A.E., 2015. Environmental sensing of heavy metals through whole cell microbial biosensors: a synthetic biology approach. ACS Synth. Biol. 4, 535–546. http://dx.doi.org/10.1021/sb500286r.
Berman, H.M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T.N., Weissig, H., Shindyalov, I.N., Bourne, P.E., 2000. The Protein Data Bank. Nucleic Acids Res. 28, 235–242.
Bornscheuer, U.T., 2016. Microbiology. Feeding on plastic. Science 351, 1154–1155. http://dx.doi.org/10.1126/science.aaf2853.
Bornscheuer, U.T., Huisman, G.W., Kazlauskas, R.J., Lutz, S., Moore, J.C., Robins, K., 2012. Engineering the third wave of biocatalysis. Nature 485, 185–194. http://dx.doi.org/10.1038/nature11117.
Bosma, T., Kruizinga, E., de Bruin, E.J., Poelarends, G.J., Janssen, D.B., 1999. Utilization of trihalogenated propanes by agrobacterium radiobacter AD1 through heterologous expression of the haloalkane dehalogenase from Rhodococcus sp. strain M15-3. Appl. Environ. Microbiol. 65, 4575–4581.
Bouhajja, E., Agathos, S.N., George, I.F., 2016. Metagenomics: probing pollutant fate in natural and engineered ecosystems. Biotechnol. Adv. 34, 1413–1426. http://dx.doi.org/10.1016/j.biotechadv.2016.10.006.
Boyle, P.M., Silver, P.A., 2012. Parts plus pipes: synthetic biology approaches to metabolic engineering. Metab. Eng. 14, 223–232. http://dx.doi.org/10.1016/j.ymben.2011.10.003.
Brakmann, S., 2001. Discovery of superior enzymes by directed molecular evolution. Chembiochem Eur. J. Chem. Biol. 2, 865–871.
Brenner, V., Arensdorf, J.J., Focht, D.D., 1994. Genetic construction of PCB degraders. Biodegradation 5, 359–377.
Burgard, A.P., Pharkya, P., Maranas, C.D., 2003. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol. Bioeng. 84, 647–657. http://dx.doi.org/10.1002/bit.10803.
Buryska, T., Daniel, L., Kunka, A., Brezovsky, J., Damborsky, J., Prokop, Z., 2016. Discovery of novel Haloalkane dehalogenase inhibitors. Appl. Environ. Microbiol. 82, 1958–1965. http://dx.doi.org/10.1128/AEM.03916-15.
Büscher, J.M., Czernik, D., Ewald, J.C., Sauer, U., Zamboni, N., 2009. Cross-platform comparison of methods for quantitative metabolomics of primary metabolism. Anal. Chem. 81, 2135–2143. http://dx.doi.org/10.1021/ac8022857.
Carbonell, P., Parutto, P., Baudier, C., Junot, C., Faulon, J.-L., 2014. Retropath: automated pipeline for embedded metabolic circuits. ACS Synth. Biol. 3, 565–577. http://dx.doi.org/10.1021/sb4001273.
Cardinale, S., Tueros, F.G., Sommer, M.O.A, 2017. Genetic-metabolic coupling for targeted metabolic engineering. Cell Rep. 20, 1029–1037. http://dx.doi.org/10.1016/j.celrep.2017.07.015.
Cases, I., de Lorenzo, V., 2005. Genetically modified organisms for the environment: stories of success and failure and what we have learned from them. Int. Microbiol. 8, 213–222.
Casini, A., Storch, M., Baldwin, G.S., Ellis, T., 2015. Bricks and blueprints: methods and standards for DNA assembly. Nat. Rev. Mol. Cell Biol. 16, 568–576. http://dx.doi.org/10.1038/nrm4014.
Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L.A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D.S., Karp, P.D., 2016. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 44, D471–480. http://dx.doi.org/10.1093/nar/gkv1164.
Chakraborty, R., Wu, C.H., Hazen, T.C., 2012. Systems biology approach to bioremediation. Curr. Opin. Biotechnol. 23, 483–490. http://dx.doi.org/10.1016/j.copbio.2012.01.015.
Chang, A., Schomburg, I., Placzek, S., Jeske, L., Ulbrich, M., Xiao, M., Sensen, C.W., Schomburg, D., 2015. BRENDA in 2015: exciting developments in its 25th year of existence. Nucleic Acids Res. 43, D439–446. http://dx.doi.org/10.1093/nar/gku1068.
Chavarría, M., Nikel, P.I., Pérez-Pantoja, D., de Lorenzo, V., 2013. The Entner-Doudoroff pathway empowers *Pseudomonas putida* KT2440 with a high tolerance to oxidative stress. Environ. Microbiol. 15, 1772–1785. http://dx.doi.org/10.1111/1462-2920.12069.
Chen, L., Lu, J., Zhang, J., Feng, K.-R., Zheng, M.-Y., Cai, Y.-D., 2013. Predicting chemical toxicity effects based on chemical-chemical interactions. PLoS One 8, e56517. http://dx.doi.org/10.1371/journal.pone.0056517.
Chen, Y., Kim, J.K., Hirning, A.J., Josić, K., Bennett, M.R., 2015. Synthetic biology. Emergent genetic oscillations in a synthetic microbial consortium. Science 349, 986–989. http://dx.doi.org/10.1126/science.aaa3794.
Chen, B., Lim, S., Kannan, A., Alford, S.C., Sunden, F., Herschlag, D., Dimov, I.K., Baer, T.M., Cochran, J.R., 2016. High-throughput analysis and protein engineering using microcapillary arrays. Nat. Chem. Biol. 12, 76–81. http://dx.doi.org/10.1038/nchembio.1978.
Choi, S.-L., Rha, E., Lee, S.J., Kim, H., Kwon, K., Jeong, Y.-S., Rhee, Y.H., Song, J.J., Kim, H.-S., Lee, S.-G., 2014. Toward a generalized and high-throughput enzyme screening system based on artificial genetic circuits. ACS Synth. Biol. 3, 163–171. http://dx.doi.org/10.1021/sb400112u.
Chowdhury, A., Zomorrodi, A.R., Maranas, C.D., 2014. k-OptForce: integrating kinetics with flux balance analysis for strain design. PLoS Comput. Biol. 10, e1003487. http://dx.doi.org/10.1371/journal.pcbi.1003487.
Chowdhury, A., Khodayari, A., Maranas, C.D., 2015. Improving prediction fidelity of cellular metabolism with kinetic descriptions. Curr. Opin. Biotechnol. 36, 57–64. http://dx.doi.org/10.1016/j.copbio.2015.08.011.
Costa, R.S., Machado, D., Rocha, I., Ferreira, E.C., 2011. Critical perspective on the consequences of the limited availability of kinetic data in metabolic dynamic modelling. IET Syst. Biol. 5, 157–163. http://dx.doi.org/10.1049/iet-syb.2009.0058.
Dai, Z., Nielsen, J., 2015. Advancing metabolic engineering through systems biology of industrial microorganisms. Curr. Opin. Biotechnol. 36, 8–15. http://dx.doi.org/10.1016/j.copbio.2015.08.006.
Damborsky, J., Brezovsky, J., 2014. Computational tools for designing and engineering enzymes. Curr. Opin. Chem. Biol. 19, 8–16. http://dx.doi.org/10.1016/j.cbpa.2013.12.003.
de la Peña Mattozzi, M., Tehara, S.K., Hong, T., Keasling, J.D., 2006. Mineralization of paraoxon and its use as a sole C and P source by a rationally designed catabolic pathway in *Pseudomonas putida*. Appl. Environ. Microbiol. 72, 6699–6706. http://dx.doi.org/10.1128/AEM.00907-06.
de Lorenzo, V., 2008. Systems biology approaches to bioremediation. Curr. Opin. Biotechnol. 19, 579–589. http://dx.doi.org/10.1016/j.copbio.2008.10.004.
de Lorenzo, V., 2009. Recombinant bacteria for environmental release: what went wrong and what we have learnt from it. Clin. Microbiol. Infect. 15 (Suppl. 1), 63–65. http://dx.doi.org/10.1111/j.1469-0691.2008.02683.x.
de Lorenzo, V., 2017. Seven microbial bio-processes to help the planet. Environ. Microbiol. http://onlinelibrary.wiley.com/doi/10.1111/1751-7915.12816/full.
de Lorenzo, V., Danchin, A., 2008. Synthetic biology: discovering new worlds and new words. EMBO Rep. 9, 822–827. http://dx.doi.org/10.1038/embor.2008.159.
de Lorenzo, V., Schmidt, M., 2017. The do-it-yourself movement as a source of innovation in biotechnology - and much more. Microb. Biotechnol. 10, 517–519. http://dx.doi.org/10.1111/1751-7915.12715.
de Lorenzo, V., Marlière, P., Solé, R., 2016. Bioremediation at a global scale: from the test tube to planet Earth. Microb. Biotechnol. 9, 618–625. http://dx.doi.org/10.1111/1751-7915.12399.
Dehghanifard, E., Jonidi Jafari, A., Rezaei Kalantary, R., Mahvi, A.H., Faramarzi, M.A., Esrafili, A., 2013. Biodegradation of 2,4-dinitrophenol with laccase immobilized on nano-porous silica beads. Iranian J. Environ. Health Sci. Eng. 10. http://dx.doi.org/10.1186/1735-2746-10-25.
Delépine, B., Libis, V., Carbonell, P., Faulon, J.-L., 2016. SensiPath: computer-aided design of sensing-enabling metabolic pathways. Nucleic Acids Res. 44, W226–231. http://dx.doi.org/10.1093/nar/gkw305.
Devers, M., Rouard, N., Martin-Laurent, F., 2008. Fitness drift of an atrazine-degrading population under atrazine selection pressure. Environ. Microbiol. 10, 676–684. http://dx.doi.org/10.1111/j.1462-2920.2007.01490.x.
Dombrowski, N., Donaho, J.A., Gutierrez, T., Seitz, K.W., Teske, A.P., Baker, B.J., 2016. Reconstructing metabolic pathways of hydrocarbon-degrading bacteria from the Deepwater Horizon oil spill. Nat. Microbiol. 1, 16057. http://dx.doi.org/10.1038/nmicrobiol.2016.57.
Dörr, M., Fibinger, M.P.C., Last, D., Schmidt, S., Santos-Aberturas, J., Böttcher, D., Hummel, A., Vickers, C., Voss, M., Bornscheuer, U.T., 2016. Fully automatized high-throughput enzyme library screening using a robotic platform. Biotechnol. Bioeng. 113, 1421–1432. http://dx.doi.org/10.1002/bit.25925.
Dvorak, P., Bidmanova, S., Damborsky, J., Prokop, Z., 2014a. Immobilized synthetic pathway for biodegradation of toxic recalcitrant pollutant 1,2,3-trichloropropane. Environ. Sci. Technol. 48, 6859–6866. http://dx.doi.org/10.1021/es500396r.
Dvorak, P., Kurumbang, N.P., Bendl, J., Brezovsky, J., Prokop, Z., Damborsky, J., 2014b. Maximizing the efficiency of multienzyme process by stoichiometry optimization. Chembiochem Eur. J. Chem. Biol. 15, 1891–1895. http://dx.doi.org/10.1002/cbic.201402265.
Dvorak, P., Chrast, L., Nikel, P.I., Fedr, R., Soucek, K., Sedlackova, M., Chaloupkova, R., de Lorenzo, V., Prokop, Z., Damborsky, J., 2015. Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway. Microb. Cell Factories 14, 201. http://dx.doi.org/10.1186/s12934-015-0393-3.
Ebenau-Jehle, C., Mergelsberg, M., Fischer, S., Brüls, T., Jehmlich, N., von Bergen, M., Boll, M., 2017. An unusual strategy for the anoxic biodegradation of phthalate. ISME J. 11, 224–236. http://dx.doi.org/10.1038/ismej.2016.91.
Erb, T.J., Zarzycki, J., 2016. Biochemical and synthetic biology approaches to improve photosynthetic CO_2-fixation. Curr. Opin. Chem. Biol. 34, 72–79. http://dx.doi.org/10.1016/j.cbpa.2016.06.026.
Eriksson, L., Jaworska, J., Worth, A.P., Cronin, M.T.D., McDowell, R.M., Gramatica, P.,

2003. Methods for reliability and uncertainty assessment and for applicability evaluations of classification- and regression-based QSARs. Environ. Health Perspect. 111, 1361–1375.
Escalante, A.E., Rebolleda-Gómez, M., Benítez, M., Travisano, M., 2015. Ecological perspectives on synthetic biology: insights from microbial population biology. Front. Microbiol. 6, 143. http://dx.doi.org/10.3389/fmicb.2015.00143.
Fernández-Álvaro, E., Snajdrova, R., Jochens, H., Davids, T., Böttcher, D., Bornscheuer, U.T., 2011. A combination of in vivo selection and cell sorting for the identification of enantioselective biocatalysts. Angew. Chem. Int. Ed. Eng. 50, 8584–8587. http://dx.doi.org/10.1002/anie.201102360.
Fibinger, M.P.C., Davids, T., Böttcher, D., Bornscheuer, U.T., 2015. A selection assay for haloalkane dehalogenase activity based on toxic substrates. Appl. Microbiol. Biotechnol. 99, 8955–8962. http://dx.doi.org/10.1007/s00253-015-6686-y.
Finley, S.D., Broadbelt, L.J., Hatzimanikatis, V., 2009. Computational framework for predictive biodegradation. Biotechnol. Bioeng. 104, 1086–1097. http://dx.doi.org/10.1002/bit.22489.
Finley, S.D., Broadbelt, L.J., Hatzimanikatis, V., 2010. In silico feasibility of novel biodegradation pathways for 1,2,4-trichlorobenzene. BMC Syst. Biol. 4, 7. http://dx.doi.org/10.1186/1752-0509-4-7.
Francis, C.W., Mankin, J.B., 1977. High nitrate denitrification in continuous flow-stirred reactors. Water Res. 11. http://dx.doi.org/10.1016/0043-1354(77)90061-6.
Fuchs, G., 2011. Alternative pathways of carbon dioxide fixation: insights into the early evolution of life? Annu. Rev. Microbiol. 65, 631–658. http://dx.doi.org/10.1146/annurev-micro-090110-102801.
Gao, J., Ellis, L.B.M., Wackett, L.P., 2010. The University of Minnesota Biocatalysis/Biodegradation Database: improving public access. Nucleic Acids Res. 38, D488–491. http://dx.doi.org/10.1093/nar/gkp771.
Gao, F., Ding, H., Feng, Z., Liu, D., Zhao, Y., 2014. Functional display of triphenylmethane reductase for dye removal on the surface of *Escherichia coli* using N-terminal domain of ice nucleation protein. Bioresour. Technol. 169, 181–187. http://dx.doi.org/10.1016/j.biortech.2014.06.093.
Geueke, B., Garg, N., Ghosh, S., Fleischmann, T., Holliger, C., Lal, R., Kohler, H.-P.E., 2013. Metabolomics of hexachlorocyclohexane (HCH) transformation: ratio of LinA to LinB determines metabolic fate of HCH isomers. Environ. Microbiol. 15, 1040–1049. http://dx.doi.org/10.1111/1462-2920.12009.
Gilbert, E.S., Walker, A.W., Keasling, J.D., 2003. A constructed microbial consortium for biodegradation of the organophosphorus insecticide parathion. Appl. Microbiol. Biotechnol. 61, 77–81. http://dx.doi.org/10.1007/s00253-002-1203-5.
Glass, J.I., Assad-Garcia, N., Alperovich, N., Yooseph, S., Lewis, M.R., Maruf, M., Hutchison, C.A., Smith, H.O., Venter, J.C., 2006. Essential genes of a minimal bacterium. Proc. Natl. Acad. Sci. U. S. A. 103, 425–430. http://dx.doi.org/10.1073/pnas.0510013103.
Glick, B.R., 1995. Metabolic load and heterologous gene expression. Biotechnol. Adv. 13, 247–261.
Guazzaroni, M.-E., Herbst, F.-A., Lores, I., Tamames, J., Peláez, A.I., López-Cortés, N., Alcaide, M., Del Pozo, M.V., Vieites, J.M., von Bergen, M., Gallego, J.L.R., Bargiela, R., López-López, A., Pieper, D.H., Rosselló-Móra, R., Sánchez, J., Seifert, J., Ferrer, M., 2013. Metaproteogenomic insights beyond bacterial response to naphthalene exposure and bio-stimulation. ISME J. 7, 122–136. http://dx.doi.org/10.1038/ismej.2012.82.
Gredell, J.A., Frei, C.S., Cirino, P.C., 2012. Protein and RNA engineering to customize microbial molecular reporting. Biotechnol. J. 7, 477–499. http://onlinelibrary.wiley.com/doi/10.1002/biot.201100266/full.
Gutierrez, T., Singleton, D.R., Berry, D., Yang, T., Aitken, M.D., Teske, A., 2013. Hydrocarbon-degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation and DNA-SIP. ISME J. 7, 2091–2104. http://dx.doi.org/10.1038/ismej.2013.98.
Hadadi, N., Hafner, J., Shajkofci, A., Zisaki, A., Hatzimanikatis, V., 2016. ATLAS of biochemistry: a repository of all possible biochemical reactions for synthetic biology and metabolic engineering studies. ACS Synth. Biol. 5, 1155–1166. http://dx.doi.org/10.1021/acssynbio.6b00054.
Haro, M.A., de Lorenzo, V., 2001. Metabolic engineering of bacteria for environmental applications: construction of Pseudomonas strains for biodegradation of 2-chlorotoluene. J. Biotechnol. 85, 103–113.
Hatzimanikatis, V., Li, C., Ionita, J.A., Henry, C.S., Jankowski, M.D., Broadbelt, L.J., 2005. Exploring the diversity of complex metabolic networks. Bioinforma. Oxf. Engl. 21, 1603–1609. http://dx.doi.org/10.1093/bioinformatics/bti213.
He, W., Yuan, S., Zhong, W.-H., Siddikee, M.A., Dai, C.-C., 2016. Application of genetically engineered microbial whole-cell biosensors for combined chemosensing. Appl. Microbiol. Biotechnol. 100, 1109–1119. http://dx.doi.org/10.1007/s00253-015-7160-6.
Hennig, S., Rödel, G., Ostermann, K., 2015. Artificial cell-cell communication as an emerging tool in synthetic biology applications. J. Biol. Eng. 9, 13. http://dx.doi.org/10.1186/s13036-015-0011-2.
Hicks, N., Vik, U., Taylor, P., Ladoukakis, E., Park, J., Kolisis, F., Jakobsen, K.S., 2017. Using prokaryotes for carbon capture storage. Trends Biotechnol. 35, 22–32. http://dx.doi.org/10.1016/j.tibtech.2016.06.011.
Hodgman, C.E., Jewett, M.C., 2012. Cell-free synthetic biology: thinking outside the cell. Metab. Eng. 14, 261–269. http://dx.doi.org/10.1016/j.ymben.2011.09.002.
Huang, P.-S., Boyken, S.E., Baker, D., 2016. The coming of age of de novo protein design. Nature 537, 320–327. http://dx.doi.org/10.1038/nature19946.
Hutchison, C.A., Chuang, R.-Y., Noskov, V.N., Assad-Garcia, N., Deerinck, T.J., Ellisman, M.H., Gill, J., Kannan, K., Karas, B.J., Ma, L., Pelletier, J.F., Qi, Z.-Q., Richter, R.A., Strychalski, E.A., Sun, L., Suzuki, Y., Tsvetanova, B., Wise, K.S., Smith, H.O., Glass, J.I., Merryman, C., Gibson, D.G., Venter, J.C., 2016. Design and synthesis of a minimal bacterial genome. Science 351, aad6253. http://dx.doi.org/10.1126/science.aad6253.
Iwakiri, R., Yoshihira, K., Ngadiman, Null, Futagami, T., Goto, M., Furukawa, K., 2004. Total degradation of pentachloroethane by an engineered Alcaligenes strain expressing a modified camphor monooxygenase and a hybrid dioxygenase. Biosci. Biotechnol. Biochem. 68, 1353–1356. http://dx.doi.org/10.1271/bbb.68.1353.
Izallalen, M., Mahadevan, R., Burgard, A., Postier, B., Didonato, R., Sun, J., Schilling, C.H., Lovley, D.R., 2008. Geobacter sulfurreducens strain engineered for increased rates of respiration. Metab. Eng. 10, 267–275. http://dx.doi.org/10.1016/j.ymben.2008.06.005.
Janssen, D.B., Dinkla, I.J.T., Poelarends, G.J., Terpstra, P., 2005. Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities. Environ. Microbiol. 7, 1868–1882. http://dx.doi.org/10.1111/j.1462-2920.2005.00966.x.
Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., Tanabe, M., 2014. Data, information, knowledge and principle: back to metabolism in KEGG. Nucleic Acids Res. 42, D199–205. http://dx.doi.org/10.1093/nar/gkt1076.
Karig, D.K., 2017. Cell-free synthetic biology for environmental sensing and remediation. Curr. Opin. Biotechnol. 45, 69–75. http://dx.doi.org/10.1016/j.copbio.2017.01.010.
Kellogg, S.T., Chatterjee, D.K., Chakrabarty, A.M., 1981. Plasmid-assisted molecular breeding: new technique for enhanced biodegradation of persistent toxic chemicals. Science 214, 1133–1135.
Khersonsky, O., Tawfik, D.S., 2010. Enzyme promiscuity: a mechanistic and evolutionary perspective. Annu. Rev. Biochem. 79, 471–505. http://dx.doi.org/10.1146/annurev-biochem-030409-143718.
Kim, J., Copley, S.D., 2012. Inhibitory cross-talk upon introduction of a new metabolic pathway into an existing metabolic network. Proc. Natl. Acad. Sci. U. S. A. 109, E2856–2864. http://dx.doi.org/10.1073/pnas.1208509109.
Kim, S., Baek, S.-H., Lee, K., Hahn, J.-S., 2013a. Cellulosic ethanol production using a yeast consortium displaying a minicellulosome and β-glucosidase. Microb. Cell Factories 12, 14. http://dx.doi.org/10.1186/1475-2859-12-14.
Kim, J., Oliveros, J.C., Nikel, P.I., de Lorenzo, V., Silva-Rocha, R., 2013b. Transcriptomic fingerprinting of *Pseudomonas putida* under alternative physiological regimes. Environ. Microbiol. Rep. 5, 883–891. http://dx.doi.org/10.1111/1758-2229.12090.
Kim, S., Thiessen, P.A., Bolton, E.E., Chen, J., Fu, G., Gindulyte, A., Han, L., He, J., He, S., Shoemaker, B.A., Wang, J., Yu, B., Zhang, J., Bryant, S.H., 2016. PubChem substance and compound databases. Nucleic Acids Res. 44, D1202–1213. http://dx.doi.org/10.1093/nar/gkv951.
King, Z.A., Lloyd, C.J., Feist, A.M., Palsson, B.O., 2015. Next-generation genome-scale models for metabolic engineering. Curr. Opin. Biotechnol. 35, 23–29. http://dx.doi.org/10.1016/j.copbio.2014.12.016.
Kleinstiver, B.P., Pattanayak, V., Prew, M.S., Tsai, S.Q., Nguyen, N.T., Zheng, Z., Joung, J.K., 2016. High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. Nature 529, 490–495. http://dx.doi.org/10.1038/nature16526.
Kosuri, S., Church, G.M., 2014. Large-scale de novo DNA synthesis: technologies and applications. Nat. Methods 11, 499–507. http://dx.doi.org/10.1038/nmeth.2918.
Koutinas, A.A., Vlysidis, A., Pleissner, D., Kopsahelis, N., Lopez Garcia, I., Kookos, I.K., Papanikolaou, S., Kwan, T.H., Lin, C.S.K., 2014. Valorization of industrial waste and by-product streams via fermentation for the production of chemicals and biopolymers. Chem. Soc. Rev. 43, 2587–2627. http://dx.doi.org/10.1039/c3cs60293a.
Krewski, D., Andersen, M.E., Mantus, E., Zeise, L., 2009. Toxicity testing in the 21st century: implications for human health risk assessment. Risk Anal. 29, 474–479. http://dx.doi.org/10.1111/j.1539-6924.2008.01150.x.
Kube, M., Chernikova, T.N., Al-Ramahi, Y., Beloqui, A., Lopez-Cortez, N., Guazzaroni, M.-E., Heipieper, H.J., Klages, S., Kotsyurbenko, O.R., Langer, I., Nechitaylo, T.Y., Lünsdorf, H., Fernández, M., Juárez, S., Ciordia, S., Singer, A., Kagan, O., Egorova, O., Petit, P.A., Stogios, P., Kim, Y., Tchigvintsev, A., Flick, R., Denaro, R., Genovese, M., Albar, J.P., Reva, O.N., Martínez-Gomariz, M., Tran, H., Ferrer, M., Savchenko, A., Yakunin, A.F., Yakimov, M.M., Golyshina, O.V., Reinhardt, R., Golyshin, P.N., 2013. Genome sequence and functional genomic analysis of the oil-degrading bacterium Oleispira antarctica. Nat. Commun. 4, 2156. http://dx.doi.org/10.1038/ncomms3156.
Kuipers, R.K., Joosten, H.-J., van Berkel, W.J.H., Leferink, N.G.H., Rooijen, E., Ittmann, E., van Zimmeren, F., Jochens, H., Bornscheuer, U., Vriend, G., dos Santos, V.A.P.M., Schaap, P.J., 2010. 3DM: systematic analysis of heterogeneous superfamily data to discover protein functionalities. Proteins 78, 2101–2113. http://dx.doi.org/10.1002/prot.22725.
Kurumbang, N.P., Dvorak, P., Bendl, J., Brezovsky, J., Prokop, Z., Damborsky, J., 2014. Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. ACS Synth. Biol. 3, 172–181. http://dx.doi.org/10.1021/sb400147n.
Latino, D.A.R.S., Wicker, J., Gütlein, M., Schmid, E., Kramer, S., Fenner, K., 2017. Eawag-Soil in enviPath: a new resource for exploring regulatory pesticide soil biodegradation pathways and half-life data. Environ. Sci.: Processes Impacts 19, 449–464. http://dx.doi.org/10.1039/c6em00697c.
Lee, J., Cao, L., Ow, S.Y., Barrios-Llerena, M.E., Chen, W., Wood, T.K., Wright, P.C., 2006. Proteome changes after metabolic engineering to enhance aerobic mineralization of cis-1,2-dichloroethylene. J. Proteome Res. 5, 1388–1397. http://dx.doi.org/10.1021/pr060008t.
Lee, J.W., Na, D., Park, J.M., Lee, J., Choi, S., Lee, S.Y., 2012. Systems metabolic engineering of microorganisms for natural and non-natural chemicals. Nat. Chem. Biol. 8, 536–546. http://dx.doi.org/10.1038/nchembio.970.
Lehrbach, P.R., Zeyer, J., Reineke, W., Knackmuss, H.J., Timmis, K.N., 1984. Enzyme recruitment in vitro: use of cloned genes to extend the range of haloaromatics degraded by Pseudomonas sp. strain B13. J. Bacteriol. 158, 1025–1032.
Li, G.-Q., Ma, T., Li, S.-S., Li, H., Liang, F.-L., Liu, R.-L., 2007. Improvement of dibenzothiophene desulfurization activity by removing the gene overlap in the dsz

operon. Biosci. Biotechnol. Biochem. 71, 849–854. http://dx.doi.org/10.1271/bbb.60189.
Li, Z., Yoshida, N., Wang, A., Nan, J., Liang, B., Zhang, C., Zhang, D., Suzuki, D., Zhou, X., Xiao, Z., Katayama, A., 2015. Anaerobic mineralization of 2,4,6-tribromophenol to CO_2 by a synthetic microbial community comprising *Clostridium*, *Dehalobacter*, and *Desulfatiglans*. Bioresour. Technol. 176, 225–232. http://dx.doi.org/10.1016/j.biortech.2014.10.097.
Libis, V., Delépine, B., Faulon, J.-L., 2016. Expanding biosensing abilities through computer-aided design of metabolic pathways. ACS Synth. Biol. 5, 1076–1085. http://dx.doi.org/10.1021/acssynbio.5b00225.
Liss, M., Daubert, D., Brunner, K., Kliche, K., Hammes, U., Leiherer, A., Wagner, R., 2012. Embedding permanent watermarks in synthetic genes. PLoS One 7, e42465. http://dx.doi.org/10.1371/journal.pone.0042465.
Litchfield, C., 2005. Thirty years and counting: bioremediation in its prime? Bioscience 55, 273–279. http://dx.doi.org/10.1641/0006-3568(2005)055[0273:TYACBI]2.0.CO;2.
Liu, X., Germaine, K.J., Ryan, D., Dowling, D.N., 2010. Genetically modified Pseudomonas biosensing biodegraders to detect PCB and chlorobenzoate bioavailability and biodegradation in contaminated soils. Bioeng. Bugs 1, 198–206. http://dx.doi.org/10.4161/bbug.1.3.12443.
Liu, J., Tan, L., Wang, J., Wang, Z., Ni, H., Li, L., 2016. Complete biodegradation of chlorpyrifos by engineered *Pseudomonas putida* cells expressing surface-immobilized laccases. Chemosphere 157, 200–207. http://dx.doi.org/10.1016/j.chemosphere.2016.05.031.
Loeschcke, A., Markert, A., Wilhelm, S., Wirtz, A., Rosenau, F., Jaeger, K.-E., Drepper, T., 2013. TREX: a universal tool for the transfer and expression of biosynthetic pathways in bacteria. ACS Synth. Biol. 2, 22–33. http://dx.doi.org/10.1021/sb3000657.
Long, M.R., Ong, W.K., Reed, J.L., 2015. Computational methods in metabolic engineering for strain design. Curr. Opin. Biotechnol. 34, 135–141. http://dx.doi.org/10.1016/j.copbio.2014.12.019.
Lorenzo, V., Loza-Tavera, H., 2011. Microbial bioremediation of chemical pollutants: how bacteria cope with multi-stress environmental scenarios. In: Storz, G., Hengge, R. (Eds.), Bacterial Stress Responses, Second edition. ASM Press, Washington, DC, pp. 481–492. http://dx.doi.org/10.1128/9781555816841.ch30.
Lovley, D.R., Phillips, E.J.P., Gorby, Y.A., Landa, E.R., 1991. Microbial reduction of uranium. Nature 350, 413–416. http://dx.doi.org/10.1038/350413a0.
Lykidis, A., Pérez-Pantoja, D., Ledger, T., Mavromatis, K., Anderson, I.J., Ivanova, N.N., Hooper, S.D., Lapidus, A., Lucas, S., González, B., Kyrpides, N.C., 2010. The complete multipartite genome sequence of Cupriavidus necator JMP134, a versatile pollutant degrader. PLoS One 5, e9729. http://dx.doi.org/10.1371/journal.pone.0009729.
Marathe, N.P., Pal, C., Gaikwad, S.S., Jonsson, V., Kristiansson, E., Larsson, D.G.J., 2017. Untreated urban waste contaminates Indian river sediments with resistance genes to last resort antibiotics. Water Res. 124, 388–397. http://dx.doi.org/10.1016/j.watres.2017.07.060.
Martínez, I., Mohamed, M.E.-S., Rozas, D., García, J.L., Díaz, E., 2016. Engineering synthetic bacterial consortia for enhanced desulfurization and revalorization of oil sulfur compounds. Metab. Eng. 35, 46–54. http://dx.doi.org/10.1016/j.ymben.2016.01.005.
Martínez-García, E., de Lorenzo, V., 2012. Transposon-based and plasmid-based genetic tools for editing genomes of gram-negative bacteria. Methods Mol. Biol. 813, 267–283. (Clifton NJ). http://dx.doi.org/10.1007/978-1-61779-412-4_16.
Martínez-García, E., de Lorenzo, V., 2016. The quest for the minimal bacterial genome. Curr. Opin. Biotechnol. 42, 216–224. http://dx.doi.org/10.1016/j.copbio.2016.09.001.
Martínez-García, E., Aparicio, T., de Lorenzo, V., Nikel, P.I., 2014a. New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. Front. Bioeng. Biotechnol. 2, 46. http://dx.doi.org/10.3389/fbioe.2014.00046.
Martínez-García, E., Nikel, P.I., Aparicio, T., de Lorenzo, V., 2014b. Pseudomonas 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. Microb. Cell Factories 13, 159. http://dx.doi.org/10.1186/s12934-014-0159-3.
Martínez-García, E., Aparicio, T., Goñi-Moreno, A., Fraile, S., de Lorenzo, V., 2015. SEVA 2.0: an update of the Standard European Vector Architecture for de-/re-construction of bacterial functionalities. Nucleic Acids Res. 43, D1183–1189. http://dx.doi.org/10.1093/nar/gku1114.
Meckenstock, R.U., Elsner, M., Griebler, C., Lueders, T., Stumpp, C., Aamand, J., Agathos, S.N., Albrechtsen, H.-J., Bastiaens, L., Bjerg, P.L., Boon, N., Dejonghe, W., Huang, W.E., Schmidt, S.I., Smolders, E., Sørensen, S.R., Springael, D., van Breukelen, B.M., 2015. Biodegradation: updating the concepts of control for microbial cleanup in contaminated aquifers. Environ. Sci. Technol. 49, 7073–7081. http://dx.doi.org/10.1021/acs.est.5b00715.
Medema, M.H., van Raaphorst, R., Takano, E., Breitling, R., 2012. Computational tools for the synthetic design of biochemical pathways. Nat. Rev. Microbiol. 10, 191–202. http://dx.doi.org/10.1038/nrmicro2717.
Mendes, P., Hoops, S., Sahle, S., Gauges, R., Dada, J., Kummer, U., 2009. Computational modeling of biochemical networks using COPASI. Methods Mol. Biol. 500, 17–59. (Clifton NJ). http://dx.doi.org/10.1007/978-1-59745-525-1_2.
Minty, J.J., Singer, M.E., Scholz, S.A., Bae, C.-H., Ahn, J.-H., Foster, C.E., Liao, J.C., Lin, X.N., 2013. Design and characterization of synthetic fungal-bacterial consortia for direct production of isobutanol from cellulosic biomass. Proc. Natl. Acad. Sci. U. S. A. 110, 14592–14597. http://dx.doi.org/10.1073/pnas.1218447110.
Moraïs, S., Shterzer, N., Lamed, R., Bayer, E.A., Mizrahi, I., 2014. A combined cell-consortium approach for lignocellulose degradation by specialized *Lactobacillus plantarum* cells. Biotechnol. Biofuels 7, 112. http://dx.doi.org/10.1186/1754-6834-7-112.
Morgat, A., Axelsen, K.B., Lombardot, T., Alcántara, R., Aimo, L., Zerara, M., Niknejad, A., Belda, E., Hyka-Nouspikel, N., Coudert, E., Redaschi, N., Bougueleret, L., Steinbeck, C., Xenarios, I., Bridge, A., 2015. Updates in Rhea—a manually curated resource of biochemical reactions. Nucleic Acids Res. 43, D459–464. http://dx.doi.org/10.1093/nar/gku961.
Mulbry, W., Ahrens, E., Karns, J., 1998. Use of a field-scale biofilter for the degradation of the organophosphate insecticide coumaphos in cattle dip wastes. Pestic. Sci. 52, 268–274. http://dx.doi.org/10.1002/(SICI)1096-9063(199803)52:3<268::AID-PS719>3.0.CO;2-2.
Muschiol, J., Peters, C., Oberleitner, N., Mihovilovic, M.D., Bornscheuer, U.T., Rudroff, F., 2015. Cascade catalysis—strategies and challenges en route to preparative synthetic biology. Chem. Commun. (Camb.) 51, 5798–5811. http://dx.doi.org/10.1039/c4cc08752f.
Mutlu, B.R., Yeom, S., Wackett, L., Aksan, A., 2015. Modelling and optimization of a bioremediation system utilizing silica gel encapsulated whole-cell biocatalyst. Chem. Eng. J. 259, 574–580. http://dx.doi.org/10.1016/j.cej.2014.07.130.
Nakatsu, C.H., Korona, R., Lenski, R.E., de Bruijn, F.J., Marsh, T.L., Forney, L.J., 1998. Parallel and divergent genotypic evolution in experimental populations of Ralstonia sp. J. Bacteriol. 180, 4325–4331.
Nancharaiah, Y.V., Mohan, S.V., Lens, P.N.L., 2016. Biological and bioelectrochemical recovery of critical and scarce metals. Trends Biotechnol. 34, 137–155. http://dx.doi.org/10.1016/j.tibtech.2015.11.003.
Nelson, K.E., Weinel, C., Paulsen, I.T., Dodson, R.J., Hilbert, H., Martins dos Santos, V.A.P., Fouts, D.E., Gill, S.R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R.T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Chris Lee, P., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J.A., Timmis, K.N., Düsterhöft, A., Tümmler, B., Fraser, C.M., 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ. Microbiol. 4, 799–808.
Nicolaou, S.A., Gaida, S.M., Papoutsakis, E.T., 2010. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofuels and chemicals, to biocatalysis and bioremediation. Metab. Eng. 12, 307–331. http://dx.doi.org/10.1016/j.ymben.2010.03.004.
Nielsen, J., Fussenegger, M., Keasling, J., Lee, S.Y., Liao, J.C., Prather, K., Palsson, B., 2014. Engineering synergy in biotechnology. Nat. Chem. Biol. 10, 319–322. http://dx.doi.org/10.1038/nchembio.1519.
Nielsen, A.A.K., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D., Voigt, C.A., 2016. Genetic circuit design automation. Science 352, aac7341. http://dx.doi.org/10.1126/science.aac7341.
Nikel, P.I., de Lorenzo, V., 2013. Engineering an anaerobic metabolic regime in *Pseudomonas putida* KT2440 for the anoxic biodegradation of 1,3-dichloroprop-1-ene. Metab. Eng. 15, 98–112. http://dx.doi.org/10.1016/j.ymben.2012.09.006.
Nikel, P.I., Pérez-Pantoja, D., de Lorenzo, V., 2013. Why are chlorinated pollutants so difficult to degrade aerobically? Redox stress limits 1,3-dichloroprop-1-ene metabolism by Pseudomonas pavonaceae. Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci. 368, 20120377. http://dx.doi.org/10.1098/rstb.2012.0377.
Nikel, P.I., Martínez-García, E., de Lorenzo, V., 2014. Biotechnological domestication of pseudomonads using synthetic biology. Nat. Rev. Microbiol. 12, 368–379. http://dx.doi.org/10.1038/nrmicro3253.
Nikel, P.I., Chavarría, M., Fuhrer, T., Sauer, U., de Lorenzo, V., 2015. *Pseudomonas putida* KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. J. Biol. Chem. 290, 25920–25932. http://dx.doi.org/10.1074/jbc.M115.687749.
Nikel, P.I., Chavarría, M., Danchin, A., de Lorenzo, V., 2016a. From dirt to industrial applications: *Pseudomonas putida* as a synthetic biology chassis for hosting harsh biochemical reactions. Curr. Opin. Chem. Biol. 34, 20–29. http://dx.doi.org/10.1016/j.cbpa.2016.05.011.
Nikel, P.I., Pérez-Pantoja, D., de Lorenzo, V., 2016b. Pyridine nucleotide transhydrogenases enable redox balance of *Pseudomonas putida* during biodegradation of aromatic compounds. Environ. Microbiol. 18, 3565–3582. http://dx.doi.org/10.1111/1462-2920.13434.
Nyerges, Á., Csörgő, B., Nagy, I., Bálint, B., Bihari, P., Lázár, V., Apjok, G., Umenhoffer, K., Bogos, B., Pósfai, G., Pál, C., 2016. A highly precise and portable genome engineering method allows comparison of mutational effects across bacterial species. Proc. Natl. Acad. Sci. U. S. A. 113, 2502–2507. http://dx.doi.org/10.1073/pnas.1520040113.
Ohtsubo, Y., Maruyama, F., Mitsui, H., Nagata, Y., Tsuda, M., 2012. Complete genome sequence of Acidovorax sp. strain KKS102, a polychlorinated-biphenyl degrader. J. Bacteriol. 194, 6970–6971. http://dx.doi.org/10.1128/JB.01848-12.
Oide, S., Gunji, W., Moteki, Y., Yamamoto, S., Suda, M., Jojima, T., Yukawa, H., Inui, M., 2015. Thermal and solvent stress cross-tolerance conferred to *Corynebacterium glutamicum* by adaptive laboratory evolution. Appl. Environ. Microbiol. 81, 2284–2298. http://dx.doi.org/10.1128/AEM.03973-14.
Oosterkamp, M.J., Veuskens, T., Talarico Saia, F., Weelink, S.A.B., Goodwin, L.A., Daligault, H.E., Bruce, D.C., Detter, J.C., Tapia, R., Han, C.S., Land, M.L., Hauser, L.J., Langenhoff, A.A.M., Gerritse, J., van Berkel, W.J.H., Pieper, D.H., Junca, H., Smidt, H., Schraa, G., Davids, M., Schaap, P.J., Plugge, C.M., Stams, A.J.M., 2013. Genome analysis and physiological comparison of Alicycliphilus denitrificans strains BC and K601(T.). PLoS One 8, e66971. http://dx.doi.org/10.1371/journal.pone.0066971.
Paddon, C.J., Keasling, J.D., 2014. Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nat. Rev. Microbiol. 12, 355–367. http://dx.doi.org/10.1038/nrmicro3240.
Parales, R.E., Ditty, J.L., 2005. Laboratory evolution of catabolic enzymes and pathways. Curr. Opin. Biotechnol. 16, 315–325. http://dx.doi.org/10.1016/j.copbio.2005.03.008.

Pascual, J., Udaondo, Z., Molina, L., Segura, A., Esteve-Núñez, A., Caballero, A., Duque, E., Ramos, J.L., van Dillewijn, P., 2015. Draft genome sequence of *Pseudomonas putida* JLR11, a facultative anaerobic 2,4,6-trinitrotoluene biotransforming bacterium. Genome Announc. 3. http://dx.doi.org/10.1128/genomeA.00904-15.

Patil, V.V., Park, K.-H., Lee, S.-G., Woo, E., 2016. Structural analysis of the phenol-responsive sensory domain of the transcription activator PoxR. Structure 1993 (24), 624–630. (Lond. Engl.). http://dx.doi.org/10.1016/j.str.2016.03.006.

Patowary, K., Patowary, R., Kalita, M.C., Deka, S., 2016. Development of an efficient bacterial consortium for the potential remediation of hydrocarbons from contaminated sites. Front. Microbiol. 7, 1092. http://dx.doi.org/10.3389/fmicb.2016.01092.

Paul, D., Pandey, G., Jain, R.K., 2005. Suicidal genetically engineered microorganisms for bioremediation: need and perspectives. Bioessays 27, 563–573. http://dx.doi.org/10.1002/bies.20220.

Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R.C., Tsuda, M., Nagata, Y., Damborsky, J., 2009. Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. Nat. Chem. Biol. 5, 727–733. http://dx.doi.org/10.1038/nchembio.205.

Petrie, B., Barden, R., Kasprzyk-Hordern, B., 2015. A review on emerging contaminants in wastewaters and the environment: current knowledge, understudied areas and recommendations for future monitoring. In: Water Res., Occurrence, Fate, Removal and Assessment of Emerging Contaminants in Water in the Water Cycle (From Wastewater to Drinking Water). 72. pp. 3–27. http://dx.doi.org/10.1016/j.watres.2014.08.053.

Planson, A.-G., Carbonell, P., Paillard, E., Pollet, N., Faulon, J.-L., 2012. Compound toxicity screening and structure-activity relationship modeling in *Escherichia coli*. Biotechnol. Bioeng. 109, 846–850. http://dx.doi.org/10.1002/bit.24356.

Puchałka, J., Oberhardt, M.A., Godinho, M., Bielecka, A., Regenhardt, D., Timmis, K.N., Papin, J.A., Martins dos Santos, V.A.P., 2008. Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. PLoS Comput. Biol. 4, e1000210. http://dx.doi.org/10.1371/journal.pcbi.1000210.

Ramos, J.L., Wasserfallen, A., Rose, K., Timmis, K.N., 1987. Redesigning metabolic routes: manipulation of TOL plasmid pathway for catabolism of alkylbenzoates. Science 235, 593–596.

Ramos, J.-L., Marqués, S., van Dillewijn, P., Espinosa-Urgel, M., Segura, A., Duque, E., Krell, T., Ramos-González, M.-I., Bursakov, S., Roca, A., Solano, J., Fernádez, M., Niqui, J.L., Pizarro-Tobias, P., Wittich, R.-M., 2011. Laboratory research aimed at closing the gaps in microbial bioremediation. Trends Biotechnol. 29, 641–647. http://dx.doi.org/10.1016/j.tibtech.2011.06.007.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., Zhang, F., 2013. Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. 8, 2281–2308. http://dx.doi.org/10.1038/nprot.2013.143.

Raunio, H., Kuusisto, M., Juvonen, R.O., Pentikäinen, O.T., 2015. Modeling of interactions between xenobiotics and cytochrome P450 (CYP) enzymes. Front. Pharmacol. 6, 123. http://dx.doi.org/10.3389/fphar.2015.00123.

Ravikumar, S., Baylon, M.G., Park, S.J., Choi, J.-I., 2017. Engineered microbial biosensors based on bacterial two-component systems as synthetic biotechnology platforms in bioremediation and biorefinery. Microb. Cell Factories 16, 62. http://dx.doi.org/10.1186/s12934-017-0675-z.

Raymond, R.L., Jamison, V.W., Hudson, J.O., 1975. Final Report on Beneficial Stimulation of Bacterial Activity in Ground Waters Containing Petroleum Products. Sun Ventures, Incorporated.

Reineke, W., Knackmuss, H.J., 1979. Construction of haloaromatics utilising bacteria. Nature 277, 385–386.

Renata, H., Wang, Z.J., Arnold, F.H., 2015. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. Angew. Chem. Int. Ed. Eng. 54, 3351–3367. http://dx.doi.org/10.1002/anie.201409470.

Rockström, J., Gaffney, O., Rogelj, J., Meinshausen, M., Nakicenovic, N., Schellnhuber, H.J., 2017. A roadmap for rapid decarbonization. Science 355, 1269–1271. http://dx.doi.org/10.1126/science.aah3443.

Rojo, F., Pieper, D.H., Engesser, K.H., Knackmuss, H.J., Timmis, K.N., 1987. Assemblage of ortho cleavage route for simultaneous degradation of chloro- and methylaromatics. Science 238, 1395–1398.

Rollin, J.A., Martin del Campo, J., Myung, S., Sun, F., You, C., Bakovic, A., Castro, R., Chandrayan, S.K., Wu, C.-H., Adams, M.W.W., Senger, R.S., Zhang, Y.-H.P., 2015. High-yield hydrogen production from biomass by in vitro metabolic engineering: mixed sugars coutilization and kinetic modeling. Proc. Natl. Acad. Sci. U. S. A. 112, 4964–4969. http://dx.doi.org/10.1073/pnas.1417719112.

Rui, L., Kwon, Y.M., Reardon, K.F., Wood, T.K., 2004. Metabolic pathway engineering to enhance aerobic degradation of chlorinated ethenes and to reduce their toxicity by cloning a novel glutathione S-transferase, an evolved toluene o-monooxygenase, and gamma-glutamylcysteine synthetase. Environ. Microbiol. 6, 491–500. http://dx.doi.org/10.1111/j.1462-2920.2004.00586.x.

Samin, G., Janssen, D.B., 2012. Transformation and biodegradation of 1,2,3-trichloropropane (TCP). Environ. Sci. Pollut. Res. Int. 19, 3067–3078. http://dx.doi.org/10.1007/s11356-012-0859-3.

Samin, G., Pavlova, M., Arif, M.I., Postema, C.P., Damborsky, J., Janssen, D.B., 2014. A pseudomonas putida strain genetically engineered for 1,2,3-trichloropropane bioremediation. Appl. Environ. Microbiol. 80, 5467–5476. http://dx.doi.org/10.1128/AEM.01620-14.

Santacoloma, P.A., Sin, G., Gernaey, K.V., Woodley, J.M., 2011. Multienzyme-catalyzed processes: next-generation biocatalysis. Org. Process. Res. Dev. 15, 203–212. http://dx.doi.org/10.1021/op1002159.

Santos, C.N.S., Yoshikuni, Y., 2014. Engineering complex biological systems in bacteria through recombinase-assisted genome engineering. Nat. Protoc. 9, 1320–1336. http://dx.doi.org/10.1038/nprot.2014.084.

Sarigiannis, D.A., Hansen, U., 2012. Considering the cumulative risk of mixtures of chemicals - a challenge for policy makers. Environ. Health 11 (Suppl. 1), S18. http://dx.doi.org/10.1186/1476-069X-11-S1-S18.

Schallmey, M., Frunzke, J., Eggeling, L., Marienhagen, J., 2014. Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. Curr. Opin. Biotechnol. 26, 148–154. http://dx.doi.org/10.1016/j.copbio.2014.01.005.

Schellenberger, J., Que, R., Fleming, R.M.T., Thiele, I., Orth, J.D., Feist, A.M., Zielinski, D.C., Bordbar, A., Lewis, N.E., Rahmanian, S., Kang, J., Hyduke, D.R., Palsson, B.Ø., 2011. Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. Nat. Protoc. 6, 1290–1307. http://dx.doi.org/10.1038/nprot.2011.308.

Schmidt, M., 2010. Xenobiology: a new form of life as the ultimate biosafety tool. Bioessays 32, 322–331. http://dx.doi.org/10.1002/bies.200900147.

Schmidt, M. (Ed.), 2012. Synthetic Biology: Industrial and Environmental Applications, 1st edition. Wiley-Blackwell.

Schmidt, M., de Lorenzo, V., 2016. Synthetic bugs on the loose: containment options for deeply engineered (micro)organisms. Curr. Opin. Biotechnol. 38, 90–96. http://dx.doi.org/10.1016/j.copbio.2016.01.006.

Schwander, T., Schada von Borzyskowski, L., Burgener, S., Cortina, N.S., Erb, T.J., 2016. A synthetic pathway for the fixation of carbon dioxide in vitro. Science 354, 900–904. http://dx.doi.org/10.1126/science.aah5237.

Scism, R.A., Bachmann, B.O., 2010. Five-component cascade synthesis of nucleotide analogues in an engineered self-immobilized enzyme aggregate. Chembiochem Eur. J. Chem. Biol. 11, 67–70. http://dx.doi.org/10.1002/cbic.200900620.

Scott, S.R., Hasty, J., 2016. Quorum sensing communication modules for microbial consortia. ACS Synth. Biol. 5, 969–977. http://dx.doi.org/10.1021/acssynbio.5b00286.

Shapir, N., Mongodin, E.F., Sadowsky, M.J., Daugherty, S.C., Nelson, K.E., Wackett, L.P., 2007. Evolution of catabolic pathways: genomic insights into microbial s-triazine metabolism. J. Bacteriol. 189, 674–682. http://dx.doi.org/10.1128/JB.01257-06.

Sheldon, R.A., 2011. Characteristic features and biotechnological applications of cross-linked enzyme aggregates (CLEAs). Appl. Microbiol. Biotechnol. 92, 467–477. http://dx.doi.org/10.1007/s00253-011-3554-2.

Shi, L., Dong, H., Reguera, G., Beyenal, H., Lu, A., Liu, J., Yu, H.-Q., Fredrickson, J.K., 2016. Extracellular electron transfer mechanisms between microorganisms and minerals. Nat. Rev. Microbiol. 14, 651–662. http://dx.doi.org/10.1038/nrmicro.2016.93.

Si, T., Xiao, H., Zhao, H., 2015. Rapid prototyping of microbial cell factories via genome-scale engineering. Biotechnol. Adv. 33, 1420–1432. http://dx.doi.org/10.1016/j.biotechadv.2014.11.007.

Silva-Rocha, R., de Lorenzo, V., 2014. Engineering multicellular logic in bacteria with metabolic wires. ACS Synth. Biol. 3, 204–209. http://dx.doi.org/10.1021/sb400064y.

Siu, K.-H., Chen, R.P., Sun, Q., Chen, L., Tsai, S.-L., Chen, W., 2015. Synthetic scaffolds for pathway enhancement. Curr. Opin. Biotechnol. 36, 98–106. http://dx.doi.org/10.1016/j.copbio.2015.08.009.

Skariyachan, S., Manjunatha, V., Sultana, S., Jois, C., Bai, V., Vasist, K.S., 2016. Novel bacterial consortia isolated from plastic garbage processing areas demonstrated enhanced degradation for low density polyethylene. Environ. Sci. Pollut. Res. Int. 23, 18307–18319. http://dx.doi.org/10.1007/s11356-016-7000-y.

Sohn, S.B., Kim, T.Y., Park, J.M., Lee, S.Y., 2010. In silico genome-scale metabolic analysis of *Pseudomonas putida* KT2440 for polyhydroxyalkanoate synthesis, degradation of aromatics and anaerobic survival. Biotechnol. J. 5, 739–750. http://dx.doi.org/10.1002/biot.201000124.

Stephanopoulos, G., 1999. Metabolic fluxes and metabolic engineering. Metab. Eng. 1, 1–11. http://dx.doi.org/10.1006/mben.1998.0101.

St-Pierre, F., Cui, L., Priest, D.G., Endy, D., Dodd, I.B., Shearwin, K.E., 2013. One-step cloning and chromosomal integration of DNA. ACS Synth. Biol. 2, 537–541. http://dx.doi.org/10.1021/sb400021j.

Stucki, G., Thueer, M., 1995. Experiences of a large-scale application of 1,2-dichloroethane degrading microorganisms for groundwater treatment. Environ. Sci. Technol. 29, 2339–2345. http://dx.doi.org/10.1021/es00009a028.

Takahashi, K., Ishikawa, N., Sadamoto, Y., Sasamoto, H., Ohta, S., Shiozawa, A., Miyoshi, F., Naito, Y., Nakayama, Y., Tomita, M., 2003. E-cell 2: multi-platform E-cell simulation system. Bioinforma. Oxf. Engl. 19, 1727–1729.

Tolia, N.H., Joshua-Tor, L., 2006. Strategies for protein coexpression in *Escherichia coli*. Nat. Methods 3, 55–64. http://dx.doi.org/10.1038/nmeth0106-55.

Tozakidis, I.E.P., Brossette, T., Lenz, F., Maas, R.M., Jose, J., 2016. Proof of concept for the simplified breakdown of cellulose by combining *Pseudomonas putida* strains with surface displayed thermophilic endocellulase, exocellulase and β-glucosidase. Microb. Cell Factories 15, 103. http://dx.doi.org/10.1186/s12934-016-0505-8.

Trögl, J., Krhůtková, O., Pilařová, V., Dáňová, P., Holíček, R., Kohlová, M., Hejda, S., Smrčka, J., Bouška, A., Křiklavová, L., 2012. Removal of nitrates from high-salinity wastewaters from desulphurization process with denitrifying bacteria encapsulated in Lentikats Biocatalyst. Int. J. Environ. Sci. Technol. 9, 425–432. http://dx.doi.org/10.1007/s13762-012-0048-4.

UniProt Consortium, 2015. UniProt: a hub for protein information. Nucleic Acids Res. 43, D204–212. http://dx.doi.org/10.1093/nar/gku989.

Vallenet, D., Calteau, A., Cruveiller, S., Gachet, M., Lajus, A., Josso, A., Mercier, J., Renaux, A., Rollin, J., Rouy, Z., Roche, D., Scarpelli, C., Médigue, C., 2016. MicroScope in 2017: an expanding and evolving integrated resource for community expertise of microbial genomes. Nucleic Acids Res. http://dx.doi.org/10.1093/nar/gkw1101.

van Leeuwen, J.G.E., Wijma, H.J., Floor, R.J., van der Laan, J.-M., Janssen, D.B., 2012.

Directed evolution strategies for enantiocomplementary haloalkane dehalogenases: from chemical waste to enantiopure building blocks. Chembiochem Eur. J. Chem. Biol. 13, 137–148. http://dx.doi.org/10.1002/cbic.201100579.
Vilchez-Vargas, R., Junca, H., Pieper, D.H., 2010. Metabolic networks, microbial ecology and "omics" technologies: towards understanding in situ biodegradation processes. Environ. Microbiol. 12, 3089–3104. http://dx.doi.org/10.1111/j.1462-2920.2010.02340.x.
Wang, H.H., Isaacs, F.J., Carr, P.A., Sun, Z.Z., Xu, G., Forest, C.R., Church, G.M., 2009. Programming cells by multiplex genome engineering and accelerated evolution. Nature 460, 894–898. http://dx.doi.org/10.1038/nature08187.
Wang, B., Kitney, R.I., Joly, N., Buck, M., 2011. Engineering modular and orthogonal genetic logic gates for robust digital-like synthetic biology. Nat. Commun. 2, 508. http://dx.doi.org/10.1038/ncomms1516.
Wang, K., Schmied, W.H., Chin, J.W., 2012. Reprogramming the genetic code: from triplet to quadruplet codes. Angew. Chem. Int. Ed. Eng. 51, 2288–2297. http://dx.doi.org/10.1002/anie.201105016.
Wang, B., Barahona, M., Buck, M., 2013. A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals. Biosens. Bioelectron. 40, 368–376. http://dx.doi.org/10.1016/j.bios.2012.08.011.
Wang, L., Xue, C., Wang, L., Zhao, Q., Wei, W., Sun, Y., 2016. Strain improvement of *Chlorella* sp. for phenol biodegradation by adaptive laboratory evolution. Bioresour. Technol. 205, 264–268. http://dx.doi.org/10.1016/j.biortech.2016.01.022.
Wei, R., Oeser, T., Schmidt, J., Meier, R., Barth, M., Then, J., Zimmermann, W., 2016. Engineered bacterial polyester hydrolases efficiently degrade polyethylene terephthalate due to relieved product inhibition. Biotechnol. Bioeng. 113, 1658–1665. http://dx.doi.org/10.1002/bit.25941.
Wibberg, D., Bremges, A., Dammann-Kalinowski, T., Maus, I., Igeño, M.I., Vogelsang, R., König, C., Luque-Almagro, V.M., Roldán, M.D., Sczyrba, A., Moreno-Vivián, C., Blasco, R., Pühler, A., Schlüter, A., 2016. Finished genome sequence and methylome of the cyanide-degrading *Pseudomonas pseudoalcaligenes* strain CECT5344 as resolved by single-molecule real-time sequencing. J. Biotechnol. 232, 61–68. http://dx.doi.org/10.1016/j.jbiotec.2016.04.008.
Wicker, J., Lorsbach, T., Gütlein, M., Schmid, E., Latino, D., Kramer, S., Fenner, K., 2016. enviPath—the environmental contaminant biotransformation pathway resource. Nucleic Acids Res. 44, D502–508. http://dx.doi.org/10.1093/nar/gkv1229.
Wierckx, N., Prieto, M.A., Pomposiello, P., de Lorenzo, V., O'Connor, K., Blank, L.M., 2015. Plastic waste as a novel substrate for industrial biotechnology. Microb. Biotechnol. 8, 900–903. http://dx.doi.org/10.1111/1751-7915.12312.
Wigginton, N., Yeston, J., Malakoff, D., 2012. Working with waste. More treasure than trash. Introduction. Science 337, 662–663. http://dx.doi.org/10.1126/science.337.6095.662.
Winsor, G.L., Griffiths, E.J., Lo, R., Dhillon, B.K., Shay, J.A., Brinkman, F.S.L., 2016. Enhanced annotations and features for comparing thousands of Pseudomonas genomes in the Pseudomonas genome database. Nucleic Acids Res. 44, D646–653. http://dx.doi.org/10.1093/nar/gkv1227.
Wittich, R.-M., Wolff, P., 2007. Growth of the genetically engineered strain *Cupriavidus necator* RW112 with chlorobenzoates and technical chlorobiphenyls. Microbiology 153, 186–195. http://dx.doi.org/10.1099/mic.0.29096-0.
Wittich, R.-M., van Dillewijn, P., Ramos*, J.L., 2010. Rational construction of bacterial strains with new/improved catabolic capabilities for the efficient breakdown of environmental pollutants. In: Timmis, K.N. (Ed.), Handbook of Hydrocarbon and Lipid Microbiology. Springer, Berlin Heidelberg, pp. 1247–1254.
Wolfaardt, G.M., Lawrence, J.R., Robarts, R.D., Caldwell, S.J., Caldwell, D.E., 1994. Multicellular organization in a degradative biofilm community. Appl. Environ. Microbiol. 60, 434–446.
Wu, G., Yan, Q., Jones, J.A., Tang, Y.J., Fong, S.S., Koffas, M.A.G., 2016. Metabolic burden: cornerstones in synthetic biology and metabolic engineering applications. Trends Biotechnol. 34, 652–664. http://dx.doi.org/10.1016/j.tibtech.2016.02.010.
Xu, P., Vansiri, A., Bhan, N., Koffas, M.A.G., 2012. ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. ACS Synth. Biol. 1, 256–266. http://dx.doi.org/10.1021/sb300016b.
Yadid, I., Rudolph, J., Hlouchova, K., Copley, S.D., 2013. Sequestration of a highly reactive intermediate in an evolving pathway for degradation of pentachlorophenol. Proc. Natl. Acad. Sci. U. S. A. 110, E2182–2190. http://dx.doi.org/10.1073/pnas.1214052110.
Yang, J., Liu, R., Jiang, H., Yang, Y., Qiao, C., 2012. Selection of a whole-cell biocatalyst for methyl parathion biodegradation. Appl. Microbiol. Biotechnol. 95, 1625–1632. http://dx.doi.org/10.1007/s00253-011-3792-3.
Yang, Y., Yang, J., Jiang, L., 2016. Comment on A bacterium that degrades and assimilates poly(ethylene terephthalate). Science 353, 759. http://dx.doi.org/10.1126/science.aaf8305.
Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., Oda, K., 2016a. A bacterium that degrades and assimilates poly(ethylene terephthalate). Science 351, 1196–1199. http://dx.doi.org/10.1126/science.aad6359.
Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., Toyohara, K., Miyamoto, K., Kimura, Y., Oda, K., 2016b. Response to comment on A bacterium that degrades and assimilates poly(ethylene terephthalate). Science 353, 759. http://dx.doi.org/10.1126/science.aaf8625.
Zobel, S., Benedetti, I., Eisenbach, L., de Lorenzo, V., Wierckx, N., Blank, L.M., 2015. Tn7-based device for calibrated heterologous gene expression in *Pseudomonas putida*. ACS Synth. Biol. 4, 1341–1351. http://dx.doi.org/10.1021/acssynbio.5b00058.

Web references

http://toxics.usgs.gov, accessed 21.5.2017.

http://www.eugris.info, accessed 21.5.2017.

http://eawag-bbd.ethz.ch, accessed 29.7.2017.

https://envipath.org/, accessed 29.7.2017.

https://metacyc.org/, accessed 29.7.2017.

https://biocyc.org/, accessed 29.7.2017.

www.uniprot.org/, accessed 22.5.2017.

http://rapid.umn.edu/rapid/, accessed 22.5.2017.

http://lcsb-databases.epfl.ch/atlas/, accessed 22.5.2017.

http://sensipath.micalis.fr/, accessed 26.7.2017.

https://ntp.niehs.nih.gov/results/tox21/, accessed 26.7.17.

https://absynth.issb.genopole.fr/Bioinformatics/, accessed 22.5.2017.

https://pubchem.ncbi.nlm.nih.gov/, accessed 29.7.2017.

http://www.chemspider.com/, accessed 29.7.2017.

http://www.pseudomonas.com/, accessed 29.7.2017.

https://www.ncbi.nlm.nih.gov/genome/, accessed 22.5.2017.

www.genoscope.cns.fr/agc/microscope/home/, accessed 22.5.2017.

http://seva.cnb.csic.es/, accessed 22.5.2017.

Review

Haloalkane dehalogenases: Biotechnological applications*

Tana Koudelakova[1], Sarka Bidmanova[1], Pavel Dvorak[1], Antonin Pavelka[1], Radka Chaloupkova[1], Zbynek Prokop[1,2] and Jiri Damborsky[1,2]

[1] Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment, Masaryk University, Brno, Czech Republic
[2] Enantis Ltd., Brno, Czech Republic

Haloalkane dehalogenases (EC 3.8.1.5, HLDs) are α/β-hydrolases which act to cleave carbon-halogen bonds. Due to their unique catalytic mechanism, broad substrate specificity and high robustness, the members of this enzyme family have been employed in several practical applications: (i) biocatalytic preparation of optically pure building-blocks for organic synthesis; (ii) recycling of by-products from chemical processes; (iii) bioremediation of toxic environmental pollutants; (iv) decontamination of warfare agents; (v) biosensing of environmental pollutants; and (vi) protein tagging for cell imaging and protein analysis. This review discusses the application of HLDs in the context of the biochemical properties of individual enzymes. Further extension of HLD uses within the field of biotechnology will require currently limiting factors – such as low expression, product inhibition, insufficient enzyme selectivity, low affinity and catalytic efficiency towards selected substrates, and instability in the presence of organic co-solvents – to be overcome. We propose that strategies based on protein engineering and isolation of novel HLDs from extremophilic microorganisms may offer solutions.

Received	24 MAY 2012
Revised	30 JUN 2012
Accepted	20 JUL 2012

Supporting information available online

Keywords: Biocatalysis · Biodegradation · Biosensors · Cell imaging · Protein analysis

1 Introduction

Haloalkane dehalogenases (EC 3.8.1.5, HLDs) are α/β-hydrolases which convert halogenated compounds to corresponding alcohols, halides and protons [1]. Hydrolytic cleavage of a carbon-halogen bond proceeds by the S_N2 mechanism (Fig. 1) followed by the addition of water [2, 3]. Water is the only co-factor required for catalysis. Moreover, HLDs are relatively stable and easy to handle, which makes them attractive for both academic research and practical applications. The importance of dehalogenating enzymes is highlighted by the number of reviews that describe different biochemistry, genetics and applications, mainly in the field of biodegradation [4–19]. Specialized reviews focused on soil-based discovery of novel dehalogenases [20] and hexachlorocyclohexane degradation [21–24] have also been published. The aims of the present review are to summarize the properties of characterized HLDs with regards to potential application and to critically discuss their possible practical use.

2 Properties of HLDs

HLDs represent one of the best characterized enzyme families that act on halogenated hydrocarbons and their derivatives [25]. During the past twenty five years, 17 different bacterial HLDs have been identified and at least partly biochemically characterized [26–38] (Prudnikova et al., in preparation). Characteristics of HLDs are summarized in Figure 2. The properties important for various practical applications, such as expression, substrate

Correspondence: Prof. Jiri Damborsky, Loschmidt Laboratories, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic
E-mail: jiri@chemi.muni.cz

Abbreviations: HLD, haloalkane dehalogenase; **SSG**, substrate specificity group

*We dedicate this paper to the founder of this field, our mentor and friend, Professor Dick B. Janssen from the Groningen University.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

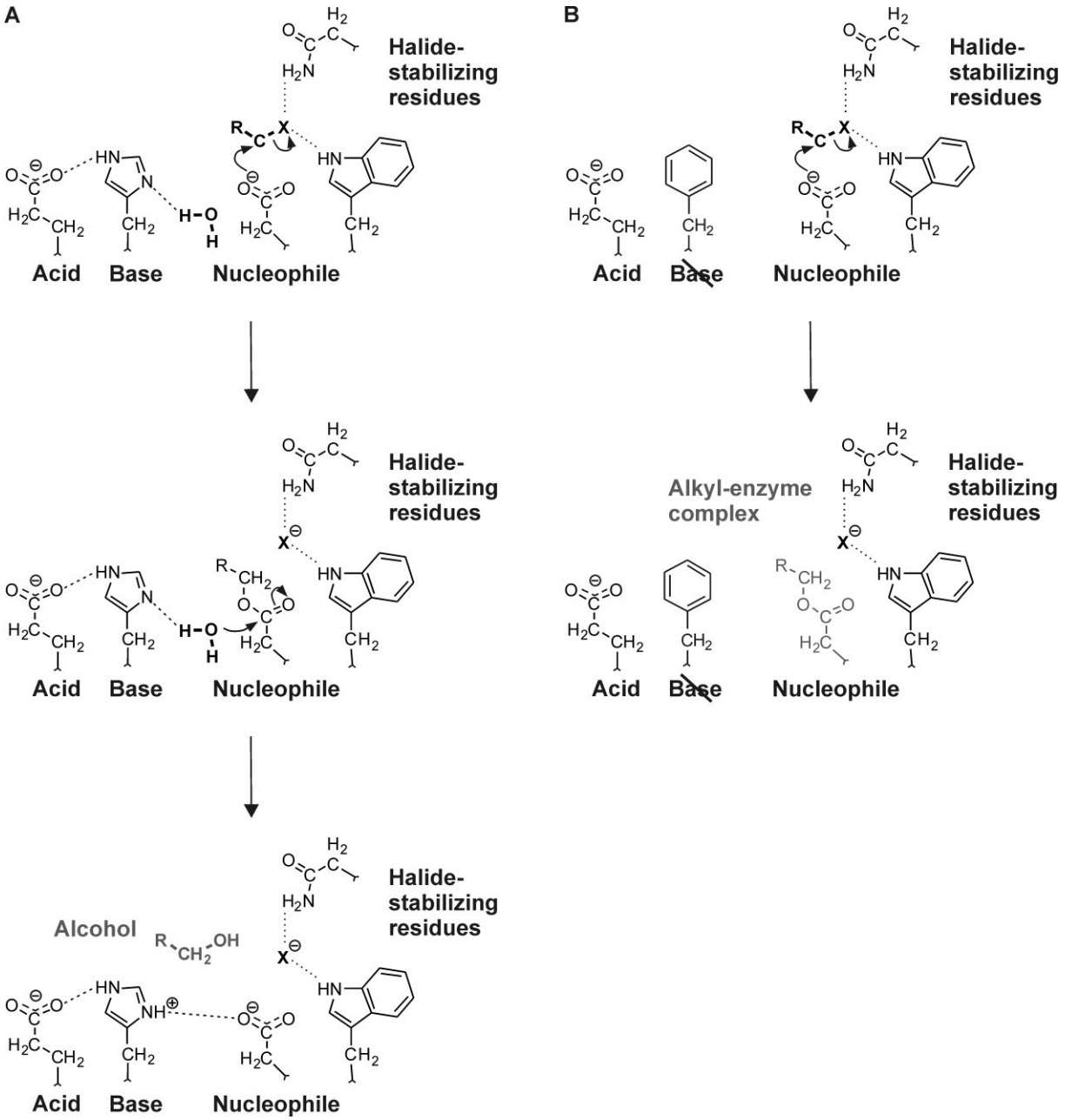


Figure 1. Simplified scheme of the reaction mechanism of HLDs. (**A**) Formation of an alcohol catalyzed by a wild-type enzyme. (**B**) Irreversible formation of an alkyl-enzyme complex in the HaloTag, where the catalytic histidine is mutated into phenylalanine. The different products of reactions are in gray. Depicted according to Verschueren et al. [2] and Los et al. [105].

specificity, catalytic efficiency, enantioselectivity and stability, are discussed in the following paragraphs.

2.1 Expression

The expression of HLDs in their natural hosts was studied with DhaA, DhlA and LinB, which were isolated from bacterial strains utilizing halogenated compounds as a sole carbon source [26–28]. Whereas DhlA and LinB are expressed constitutively in and *Xanthobacter autotrophicus* GJ10 and *Sphingobium japonicum* UT26, respectively [39, 40], the expression of DhaA in *Rhodococcus rhodochrous* NCIMB 13064 is regulated by a repressor responding to 1-chlorobutane and 1-haloalkane levels in the cellular environment [41–43]. However, strains with constitutive expression of DhaA have also been described [43]. Although the expression and function of other HLDs in their original strains has not been identified, most

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

A

	DatA	DbeA	DbjA	DhaA	DhlA	DhmA	DmbA	DmbB	DmbC	DmlA	DmmA	DpcA	DppA	DrbA	LinB
Expression															
Heterologous host	EC	EC	EC	EC	EC	RE	EC	EC	MS	EC	EC	EC	EC	EC	EC
Non-optimized yield (mg L^{-1})	30	40	50	120	70	0.8	10	1	3	50	150	30		4	50
Activity and specificity															
Overall activity	**	**	***	**	**		**		*			**		*	***
Chlorinated substrates	*	*	**	**	**	**	**		*	*	*	*	x	*	**
Brominated substrates	**	**	***	***	***	**	**	*	**	*	*	***	***	*	***
Iodinated substrates	**	**	***	***	***	**	**		**		*	***		*	***
Substrate specificity group	IV	IV	I	I	I		II		IV			IV		III	I
Catalytic efficiency															
K_m range (mM)	0.07-2		0.04-4	0.005-11	0.01-48		0.2-5		0.02		0.2	4-6	0.8-45	0.06	0.005-2
k_{cat} range (s^{-1})	0.04-2		0.02-4	0.08-14	0.09-4		0.08-9		0.07		2	1-3	0.4-10	0.1	0.1-41
Enantioselectivity															
α-Bromoesters	*	**	**	*	x							*			*
α-Bromoamides			**	*						*					**
β-Bromoalkanes	*	*	*	x								x			x
Stability															
Predicted *pI*	5.9	6.9	5.9	5.0	4.7	8.8	5.1	6.5	9.2	6.2	5.0	5.4	4.9	7.1	5.0
pH profile maximum	9.8	8.4	9.7	8.0-9.5	8.2	8.0	9.0	6.5	8.3			8.7	8.0-9.0	9.2	8.2, 8.8
Melting temperature (°C)	48	59	54	50	39		52	63[a]	52[a]			34		45[a]	48
Temp. profile maximum (°C)	40	45-55	50	30-37	37	50	45	45	40			25	30-37	50	40
Structural informations															
Predicted MW (kDa)	33.5	33.8	34.1	33.2	35.1	33.8	33.7	33.4	32.2	34.2	33.3	34.0	32.6	36.3	33.1
Units in quaternary structure	1-2[b]		1-2-4[b]	1	1				∞		2[c]	1	1	∞	1
Sample structure		UP[d]	3AFI	1CQW	2YXP		2QVB				3U1T		2XT0		1MJ5
Volume of active-site cavity ($Å^3$)		232	237	255	132		239				358		351		295
Tunnel bottleneck radius (Å)		1.7	2.0	1.4	0.8		2.0				2.1		1.9		1.4
Protein Engineering studies															
Constructed mutants			*	***	***	**									***

[a]Recalculated. [b]Dependent on conditions. [c]Predicted. [d]Unpublished data (Prudnikova et al.).

B

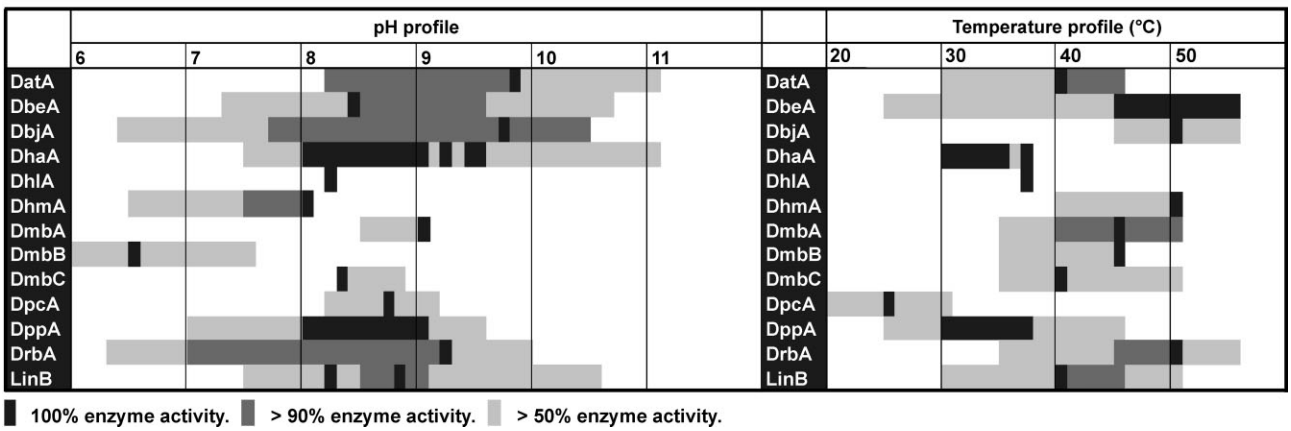


Figure 2. Biochemical and structural properties of HLDs. (**A**) Summary of properties of DatA [35], DbeA [46] (Prudnikova et al., in preparation), DbjA [31, 45–47, 58], DhaA [27, 45–47, 49, 50, 55, 119, 120], DhlA [26, 46, 48, 50, 56, 121], DhmA [29, 32], DmbA [30, 46, 122], DmbB [30], DmbC [33, 46], DmlA [31, 47], DmmA [37], DpcA [38], DppA [36], DrbA [33, 46], LinB [45–47, 51, 123]. The isoelectric point (*pI*) and molecular weight (MW) were predicted by Expasy server [124]. The detailed description of the thermal denaturation measurement and the calculation of cavity properties is given in the Supporting information. Abbreviations: EC, *Escherichia coli*; MS, *Mycobacterium smegmatis*; RE, *Rhodococcus erythropolis*; ×, *, ** and ***, qualitative measures; ∞, multimer. (**B**) pH and temperature profile.

HLDs can be heterologously expressed in *Escherichia coli* (Fig. 2) [26–39, 42] (Prudnikova et al., in preparation). Protein yields obtainable from shake flask cultivations under laboratory conditions range between 1–150 mg L^{-1}, depending on the particular heterologous enzyme expressed and expression system used. Improved expression may be achieved by optimized codon usage and the use of a suitable expression system.

2.2 Substrate specificity

HLDs possess broad substrate specificity. Cleavage of the carbon-halogen bond by HLDs has been reported using more than a hundred chlorinated, brominated and iodinated aliphatic compounds containing a monohalogenated sp^3 hybridized carbon as substrate (Fig. 3A, Supporting information, Tables S1 and S2) [1, 44–47]. Compounds

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

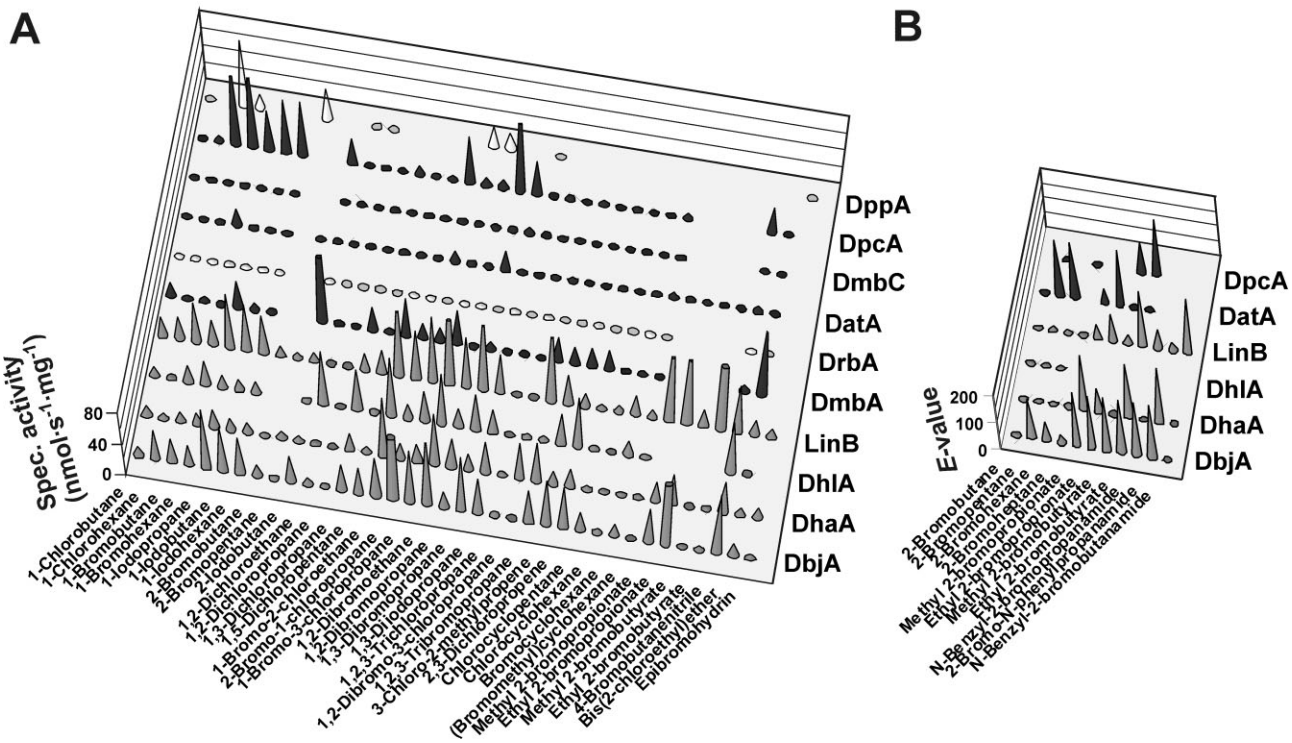


Figure 3. Substrate specificity and enantioselectivity of HLDs. (**A**) Substrate specificity of HLDs. The enzymes are colored according to membership to individual substrate specificity groups: SSG-I (DbjA, DhaA, DhlA and LinB), SSG-II (DmbA), SSG-III (DrbA) and SSG-IV (DatA, DmbC and DpcA). DppA is unclassified. The columns of specific activities higher than 80 nmol s^{-1} mg^{-1} are trimmed for clarity. (**B**) Enantioselectivity of HLDs.

halogenated on sp^2 hybridized carbon, multihalogenated on a single carbon, fluorinated, or aromatic are not converted [1]. A recent study of substrate specificity divided HLDs into four substrate specificity groups (SSGs) [46]. HLDs that convert a number of resistant chlorinated substrates were categorized as one group (SSG-I: DbjA from *Bradyrhizobium japonicum* USDA 110, DhaA, DhlA and LinB). Another group preferring brominated and iodinated compounds was identified (SSG-IV: DatA from *Agrobacterium tumefaciens* C58, DbeA from *Bradyrhizobium elkanii* USDA94 and DmbC from *Mycobacterium bovis* 5033/66) (Fig. 2). SSG-II and SSG-III were each populated with only one enzyme (DmbA from *Mycobacterium bovis* 5033/66 and DrbA from *Rhodopirellula baltica* SH1, respectively), reflecting the unique substrate preferences of these enzymes. The study also identified the "universal" substrates, i.e. 1-bromobutane, 1-iodopropane, 1-iodobutane, 1,2-dibromoethane and 4-bromobutanenitrile and the "poor" substrates, i.e. 1,2-dichloropropane, 1,2,3-trichloropropane, chlorocyclohexane and (bromomethyl)cyclohexane of analyzed HLDs.

2.3 Catalytic efficiency

HLDs exhibit catalytic efficiencies ranging from 10^4 to 10^5 M^{-1} s^{-1} with their best (brominated) substrates [48–53]. Even though these numbers are far from the theoretical maximal value of 10^8-10^9 M^{-1} s^{-1} for diffusion-limited enzymes, they are close to the recently estimated median catalytic efficiency of an "average" enzyme: 10^5 M^{-1} s^{-1} [54]. Catalytic efficiency achieved with anthropogenic chlorinated substrates, such as 1,2,3-trichloropropane, is significantly lower (40 M^{-1} s^{-1} for DhaA), making protein engineering of the enzyme to improve catalytic efficiencies essential to gain performance sufficient for industrial applications [52]. The Michaelis-Menten constants and the rate constants reported for HLDs vary across substrates by five and four orders of magnitude, respectively (K_m: 0.005 – 45 mM and k_{cat}: 0.04 – 40 mM) [30, 31, 33, 35–38, 48–51].

2.4 Enantioselectivity

Initial study of DhaA and DhlA revealed only a weak enantioselectivity of HLDs with selected chiral and prochiral halogenated substrates (Supporting information, Table S4) [55, 56]. While DhaA was assayed using 2-bromobutane, 2-bromopentane and 2-bromo-1-phenylpropane by Janssen et al. [55], both DhaA and DhlA were tested with six short-chain dihalogenated alkanes, six terminally monohalogenated esters and four halogenated prochiral propanes by Pieters et al. [56]. The subsequent

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

studies by Prokop et al. [45, 57] revealed that DhaA, LinB and DbjA strongly discriminates between enantiomers of several α-brominated esters (Fig. 3B, Supporting information, Table S3), and DbjA additionally exhibits high enantioselectivity towards two β-brominated alkanes. The same three enzymes were successfully used for kinetic resolution of α-bromoamides [47]. Most recently, three novel HLDs enantioselective towards α-brominated esters were discovered [35, 38] (Prudnikova et al., in preparation). The enzyme DatA also exhibits exceptional enantioselectivity towards two β-brominated alkanes [35]. All enantioselective HLDs tested so far have exhibited the preference for (*R*)-brominated substrates [35, 38, 45, 47].

2.5 Stability

Increasing reaction temperature to 35–50°C has a positive effect on the catalysis of HLDs. Above these temperatures, a rapid drop in activity is observed [58]. The only exception is cold-adapted DpcA from *Psychrobacter cryohalolentis* K5, which shows its highest activity at 25°C [38]. Although elevated temperature increases the catalytic rate, HLD structures start to denature between 40–50°C (Fig. 2, Supporting information, Table S3) [30, 33, 35, 58]. Similarly, extremely acidic or alkalic conditions are not suitable for HLDs, which generally perform catalysis best at pH 8-10 (Fig. 2, Supporting information, Table S3) [26–28, 30, 33, 35, 58]. The study of the behavior of DbjA, DhaA and LinB in the presence of 14 water-miscible organic solvents showed that the resistance of individual enzymes to organic co-solvents differs (Stepankova et al., submitted). While DbjA was stable in most of the co-solvents tested up to concentrations of 20% (v/v), the activity of LinB and DhaA gradually decreased at a co-solvent concentration of 5% and 10%, respectively.

3 Applications of HLDs

HLDs have attracted considerable attention due to their unique catalytic mechanism, broad substrate specificity and robustness. They have been used for biocatalytic preparation of optically pure building-blocks for organic synthesis; recycling of by-products from chemical processes; bioremediation of toxic environmental pollutants; decontamination of chemical warfare agents; biosensing environmental pollutants; and protein tagging for cell imaging and protein analysis (Fig. 4).

3.1 Biocatalysis

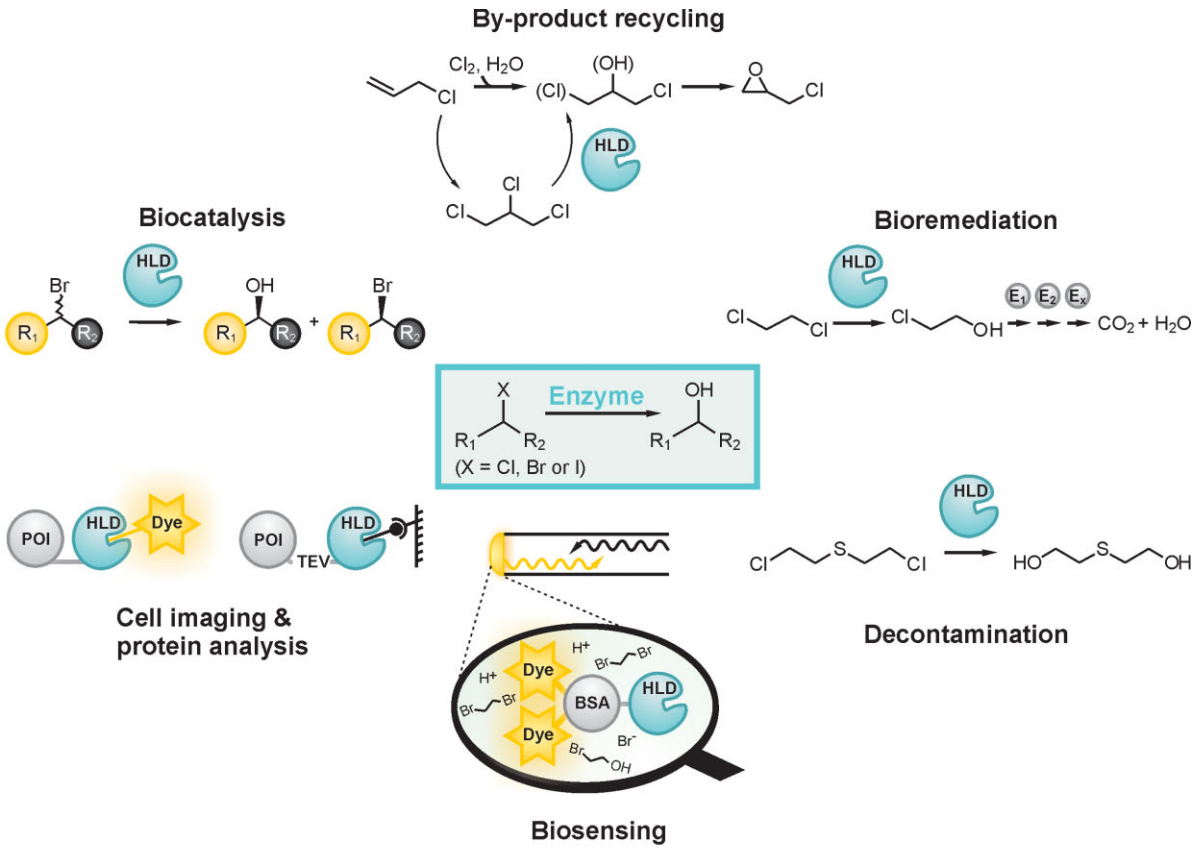
The discovery of high enantioselectivity of HLDs (Supporting information, Table S3) has opened up the possibility to utilize HLDs to prepare optically pure (*S*)-β-bromoalkanes, (*S*)-β-alcohols, (*S*)-α-brominated esters, (*S*)-α-hydroxyesters [45, 57], (*S*)-α-bromoamides and (*S*)-α-hydroxyamides [47], useful building blocks in the production of pharmaceuticals. For instance, (*S*)-2-pentanol is an intermediate in the synthesis of potential anti-Alzheimer's disease drugs [59], ethyl (*S*)-2-hydroxypropionate is used for preparation of lofexidine [60], and optically pure amides are precursors of non-natural peptides, vitamins and antibiotics [61] (Fig. 5).

Enantioselectivity of an enzyme can be described either by enantiomeric excess (*e.e.*, the excess of one enantiomer over the other) or the *E* value (*E*, the ratio of specificity constants for the prefered or non-prefered enantiomer). Besides high enantioselectivity (*e.e.* > 99%, *E* > 100), further requirements for the application of enzyme to a biocatalytic process have been proposed: (i) conversion of more than 98%; (ii) substrate loading higher than 100 g L^{-1}; (iii) substrate/biocatalyst ratio higher than 50 with maximal biocatalyst loading 5 g L^{-1}; and (iv) reaction time less than 24 h [62]. Due to the character of kinetic resolution of racemic substrates, the theoretical yield is limited to 50%, unless an additional chemical or enzymatic process [61] or racemization of the substrate [63] is used. Activation of the hydroxyl group of three different (*S*)-α-hydroxyamides by methanesulfonyl chloride has enabled the enantioconvergent chemoenzymatic synthesis of optically active *N*-alkyl (*R*)-α-azidoamides, *N*-alkyl (*R*)-α-benzylaminoamides, *N*-alkyl (*R*)-α-phenoxyamides and *N*-alkyl (*R*)-α-ethylthioamides [61] (Fig. 6). The compounds were prepared within two days to give a 57-96% yield and *e.e.* of 28–98%. The first reported dynamic kinetic resolution using a HLD employed polymer-based phosphonium bromide as a racemization agent for *N*-phenyl-2-bromopropionamide and incorporated an experimental setup with DbjA in a semipermeable membrane compartment [63]. A 63% yield and an *e.e.* of 95% were achieved using this process.

The use of semipermeable membrane reactors and organic solvents could solve the issue of low substrate loadings under aqueous reaction conditions, which are reported as approximately 0.3 g L^{-1} [45, 47]. In the case of (*S*)-2-pentanol preparation, such low substrate loading is the biggest disadvantage compared to the alternative kinetic resolution catalyzed by *Candida antarctica* lipase B which employs racemic 2-pentanol as both substrate and solvent [59]. Nevertheless, processes catalyzed by HLDs generally possess favorable substrate/biocatalyst ratios (>50). Reactions can be completed within 2 hour to 2 day time frame [45, 47].

A further factor contributing to the economical feasibility of the process is the price of biocatalyst production. Enzyme production yields have a dramatic effect on production costs as they can vary from units of mg L^{-1} in non-optimized systems up to 15 g L^{-1} and 25 g L^{-1} in optimized intracellular and extracellular systems, respectively [64]. The cost of enzyme should be less than 10% of total process cost [65]. Therefore, the current expression yields

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



	DatA	DbeA	DbjA	DhaA	DhlA	DmbA	DmlA	DpcA	LinB
Biocatalysis									
KR of α-bromoesters	●	●	●	●				●	●
KR of β-bromoalkanes	●	●	●						
KR of α-bromoamides			●	●			●		●
DKR of α-bromoamides			●						
Chemoenzymatic synthesis			●						●
By-product recycling									
Epichlorohydrin production				●					
Bioremediation									
DCA degradation					○				
TCP degradation				✦					
CB degradation				●					
Lindane degradation									●
Phytoremediation				●	●				
Decontamination									
Sulfur mustard decontamination				●		●			●
Biosensing									
Bioassays				●	●				
Biosensors - whole cells				○	○				
Biosensors - enzymes				✦					○
Cell imaging & protein analysis									
Binding of fluorescent probe			✦	✦					
Cell imaging				✦					
Protein labeling				✦					
Protein immobilization				✦					
Protein expression				✦					
Protein purification				✦					
Protein interaction studies				✦					

○ Commercial application, ● Lab-scale application, ✦ Engineered protein.

Figure 4. Biotechnological applications of HLDs. Summary of reported applications of DatA [35], DbeA (Prudnikova et al., in preparation), DbjA [45, 47, 57, 61, 63, 125], DhaA [14, 44, 45, 47, 57, 66, 78, 79, 83, 87, 89, 92, 94, 96, 105–115, 125], DhlA [75, 76, 93, 97], DmbA [44, 89], DmlA [47], DpcA [38] and LinB [28, 44, 45, 47, 57, 61, 89, 95]. Abbreviations: BSA, bovine serum albumin; CB, 1-chlorobutane, DCA; 1,2-dichloroethane; DKR, dynamic kinetic resolution; KR, kinetic resolution; POI, protein of interest; TCP, 1,2,3-trichloropropane; TEV, cleavage site for TEV protease.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

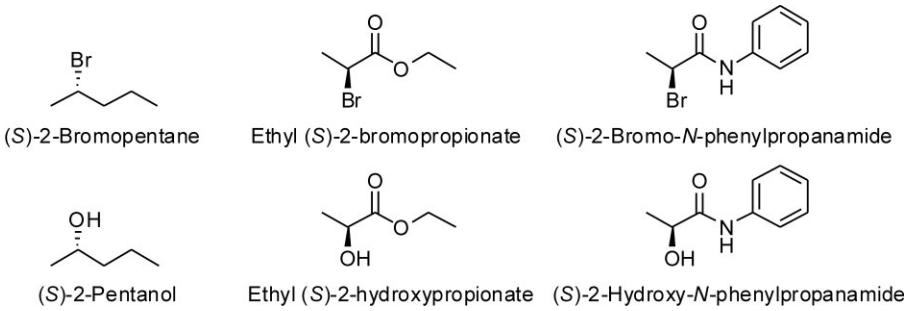


Figure 5. Examples of optically pure compounds obtained from kinetic resolutions employing HLDs.

of HLDs reaching 150 mg L^{-1} under laboratory conditions require optimization.

3.2 By-product recycling

1,2,3-Trichloropropane, 1,2-dichloropropane, and 1,2-dichlorobutane are by-products from the chemical synthesis of epichlorohydrin, propylene oxide, and butylene oxide, respectively [14, 20, 66] (Fig. 7). Recycling of the halogenated aliphatic compounds to halohydrins could save manufacturing costs; therefore the Dow Chemical Company attempted to develop an efficient catalyst to drive halohydrin formation (Supporting information, Table S5). The initial culture enrichment of various soil and water samples from Midland County, Michigan in medium with a mixture of 1-chlorobutane, 2-chlorobutane, and 1-chloro-2-methylpropane led to isolation of an *Rhodococcus* sp. TDTM003 possessing a variant of DhaA [20, 66].

Application of an enzyme biocatalyst for large-scale processing requires suitable specificity, but also favorable kinetic properties and good stability [14, 67]. Although wild type DhaA can convert 1,2,3-trichloropropane into required 2,3-dichloropropane-1-ol [66] and immobilized enzyme kept 10% of its activity in neat substrate [67], the technology was not competitive due to poor kinetic parameters of the biocatalyst (K_m > 1 mM, k_{cat} < 1 s^{-1}) and the presence of product inhibition (Supporting information, Table S5).

The limitations described above have motivated follow-up DNA enrichment studies and protein engineering experiments [20, 52, 68–70]. The enrichment of specific sequences from a soil sample using degenerated oligonucleotides led to discovery of 19 new genes and the characterization of four novel enzymes active with 1,2,3-trichloropropane [20]. Three of them lacked product inhibition, but their kinetic parameters were comparable with those reported for DhaA (Supporting information, Table S5).

The significant improvement of DhaA kinetic parameters was reported in three protein engineering studies [52, 68, 69]. A mutant possessing a four-fold higher catalytic constant with 1,2,3-trichloropropane than the wild-type was obtained after one round of error-prone PCR by Gray et al. [68]. Bosma et al. [69] constructed a variant M2 using DNA shuffling and error-prone PCR that was nearly eight-times more efficient in conversion of 1,2,3-trichloropropane than DhaA. Computer modeling provided a deeper understanding of the improved conversion of 1,2,3-trichloropropane by the variant M2 [71] and additional mutations were introduced to the access tunnels of DhaA by saturation mutagenesis [52, 72]. The resultant five-fold mutant DhaA31 exhibited a 26-times higher

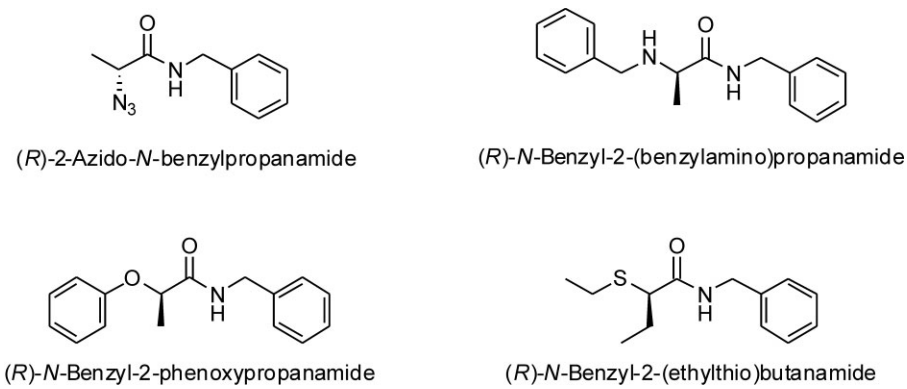


Figure 6. Examples of optically pure compounds obtained from chemoenzymatic syntheses employing HLDs.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

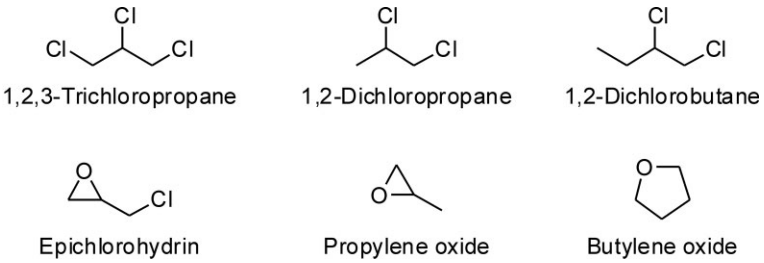


Figure 7. Examples of halogenated by-products from chemical processes (up) and the desired products (down) catalyzed by HLDs.

catalytic efficiency than the wild-type enzyme (with K_m close to 1 mM and $k_{cat} > 1$ s^{-1}; Supporting information, Table S5).

In addition to kinetic parameters, enantioselectivity and stability of DhaA also have been optimized. Van Leeuwen et al. [73] used DhaA31 as a template for additional five rounds of directed evolution and obtained enantiocomplementary mutants r5-90R and r5-97S, which can be used for the preparation of high-value chiral epichlorohydrins. Gray et al. [68, 74] obtained a remarkably thermostable DhaA variant (designated Dhla8), which exhibited a half-life of 138 min at 80°C by a combination of mutations from gene site saturation mutagenesis.

3.3 Bioremediation

Bioremediation is one of the most promising fields to apply HLDs because their natural substrates, chlorinated aliphatic and cyclic compounds, are prominent environmental pollutants. The compounds, such as 1,2-dichloroethane, 1-chlorobutane and hexachlorocyclohexane (Fig. 8), have been introduced into the environment through their use as organic solvents or pesticides [14, 18, 24, 52, 75].

The most thriving use of HLDs described to date is the application of the 1,2-dichloroethane utilizing strain *X. autotrophicus* GJ10 in a full-scale groundwater purification plant in Lübeck, Germany [13, 76]. The plant operates at 8–12°C and treats water contaminated by 1,2-dichloroethane in concentrations from 2–15 mg L^{-1} to the final concentration 0.01 mg L^{-1} with a speed of 5–20 m^3 h^{-1}. Previous attempts to utilize the same strain in air-lift and extractive membrane bioreactors for removal of 1,2-dichloroethane from the contaminated water or gas has remained at the level of laboratory studies [13, 77].

Several studies have been focused on microbial biodegradation of 1,2,3-trichloropropane. A continuous effort to enrich 1,2,3-trichloropropane-degrading organisms from environmental samples was not successful. Therefore, attempts to construct a synthetic biochemical pathway for 1,2,3-trichloropropane utilization in bacterium were made. Initially, the wild-type DhaA from *Rhodococcus* sp. m15-3 was heterogously expressed in *Agrobacterium radiobacter* AD1, the strain reported to possess a biochemical pathway for utilization of halogenated alcohols [78]. An engineered strain with an assembled pathway can use 1,2,3-tribromopropane and 1,2-dibromo-3-chloropropane as a sole carbon and energy source but its growth on 1,2,3-trichloropropane was reported to be poor [79]. In a follow-up study, the previously described DhaA mutant M2 (Supporting information, Table S5) was used instead of wild-type DhaA [69]. However, even an eight-fold improvement of DhaA catalytic efficiency was not sufficient to provide feasible biodegradation of 1,2,3-trichloropropane.

Continuous solid-gas biofilters with dehydrated bacterial cells containing DhaA or DhlA have been extensively studied for removal of 1-chlorobutane from waste gas [80–83]. Pre-treatment of the dehydrated bacterial cells of *E. coli* BL21(DE3) containing DhaA with ammonia vapor improved the stability of the catalyst and retained maximal reaction rates for 60 hours [83].

Biodegradation of four hexachlorocyclohexane isomers (α-, β-, γ- and δ-) was recently reviewed by Lal et al. [24]. Significant reductions in hexachlorocyclohexane levels (30–100%) were achieved after the growth stimulation of hexachlorocyclohexane-degrading strains at four different contaminated sites. Inoculation of hexachlorocyclohexane-degrading strains into diverse samples of contaminated soil led to successful biodegradation of hexachlorocyclohexane as well. The two most recent papers reported in vitro biotransformation of ε-hexachlorocyclohexane, heptachlorocyclohexane [84] and isomers of hexabromocyclododecane [85] (Fig. 8) into mono- and/or dihydroxylated products by LinB from *Sphingobium indicum* B90A.

Due to the restricted use of genetically modified bacteria for environmental clean-up applications, the possibility of phytoremediation has been investigated as well [75, 86–88]. Tobacco plants containing a haloalkane and a haloacid dehalogenase from *X. autotrophicus* GJ10 or an apoplast-targeted haloalkane dehalogenase from *Rhodococcus* sp. were recently engineered. These transgenic plants could degrade 1,2-dichloroethane [75] and 1-chlorobutane [87], respectively, and represent a potentially cost-effective and efficient remediation strategy [75].

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

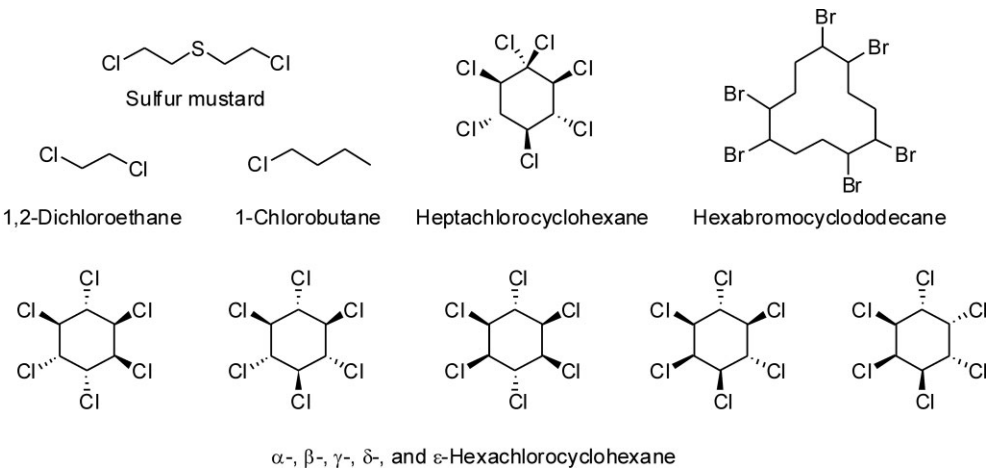


Figure 8. Examples of the warfare agent sulfur mustard and common halogenated environmental pollutants catalyzed by HLDs.

3.4 Decontamination

The ability of DhaA, DmbA and LinB to degrade the warfare agent sulfur mustard (yperite, Fig. 8) has been reported [44, 89]. While spontaneous hydrolysis of sulfur mustard leads to formation of a sulfonium ion, an unwanted reaction intermediate with adverse health effects, dehalogenase hydrolysis of sulfur mustard proceeds via a non-toxic intermediate [44]. The application of HLDs offers a safe and environmentally friendly alternative to chemical detoxification. The concept has been exploited in the enzyme-based decontamination mixture Yperzyme Kit [90] developed by the company Enantis. The major limiting factor of current technology is the poor water solubility of the target molecule. If sulfur mustard is predissolved in an organic solvent miscible with water, the agent readily dissolves in the mixture and rapid enzymatic hydrolysis can occur [91]. However, the unsatisfactory stability of the enzymes in the presence of organic co-solvents can limit such an application. Novel HLDs hydrolyzing sulfur mustard or the use of immobilized/engineered enzymes with improved stability could eliminate these limitations [44] (Prokop et al., patent application).

3.5 Biosensing

Detection of chemicals using biosensors provides a simple and rapid alternative to the traditional analytical techniques, such as gas and liquid chromatography, which are time-consuming, expensive and not suitable for continuous measurements [92–95]. Initial attempts to construct biosensors for halogenated substrates (Supporting information, Table S6) led to a system with dehalogenating cells immobilized in polymer beads as a biological component and ion selective electrodes as a transducer [96, 97].

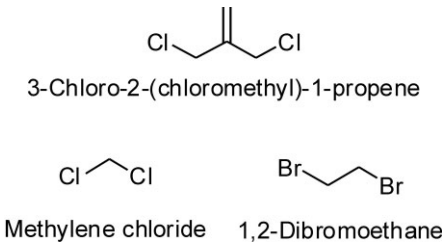


Figure 9. Examples of halogenated compounds detectable by biosensors employing HLDs.

The subsequently constructed biosensors utilized cells immobilized directly on the sensing membrane of ion selective electrode [92] or on the layer with a fluorescent pH indicator applied on optical fiber [93, 94]. This concept has been commercialized by the company OptiEnz Sensors, which offers biosensors for the detection of 1,2-dichloroethane, 1,2-dibromoethane, methylene chloride and other halogenated aliphatic hydrocarbons (Fig. 8 and 9). These biosensors possess low detection limits and enable simultaneous detection of analytes using muti-channel optoelectronics.

The first biosensor exploiting purified HLD was reported recently [95]. The biosensor tip consists of a layer of the enzyme and the fluorescent pH indicator conjugated to bovine serum albumin [98] (Bidmanova et al., unpublished data). The enzyme-based biosensor offers significantly shorter measurement time compared to conventional methods, exhibits broad temperature and pH working range and stability (Bidmanova et al., unpublished data). This sensing concept was exploited by the company Enantis in the biosensor EnviroPen, for detection of common halogenated environmental pollutants, e.g. 1,2-dibromoethane, 3-chloro-2-(chloromethyl)-1-propene and 1,2,3-trichloropropane (Figs. 7 and 9). There is no natural form of HLD possessing sufficient activity towards 1,2,3-trichloropropane, so the biosensor

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

suitable for the detection of this compound uses variant DhaA31 (Supporting information, Table S6).

The performance of the biosensor is critically influenced by the selection of an appropriate immobilization method for the biological component. The dehalogenating cells were encapsulated on the transducer [92–94, 96, 97], while the purified enzyme was covalently bound [95]. Two previous reports on covalent immobilization of HLDs described covalent binding of DhaA onto alumina support impregnated by polyethyleneimine [67] and combination of affinity and covalent binding of DhlA on iron oxide nanoparticles [99]. In this case, cross-linking of LinB with glutaraldehyde was favorable due to the highest activity retention and negligible leakage compared to other tested methods (Bidmanova et al., submitted).

Other critical characteristics of a biosensor are selectivity, sensitivity and functionality under operating conditions. Due to their wide substrate scope, individual HLDs are not specific to the particular substrate, but act within whole class of halogenated compounds [95]. Limited selectivity of HLD-based biosensor can be improved through protein engineering or using sensor arrays employing enzymes with different selectivity [94]. Such a strategy is applicable for the detection of 1,2-dichloroethane, which is converted only by DhlA [46].

The detection limits of current biosensors employing HLDs vary from 1.0 to 25 000 $\mu g\ L^{-1}$ [92–95]. Although such detection limits can be sufficient for the monitoring of contaminated sites, a lower limit of detection is required for analysis of drinking water [94]. The limits of 1,2-dibromoethane, 1,2-dichloroethane, 1,2-dichloropropane, 1,2-dibromo-3-chloropropane and lindane in drinking water set by US Environmental Protection Agency range from 0.05 $\mu g\ L^{-1}$ to 5 $\mu g\ L^{-1}$. To decrease the detection limit, protein engineering needs to be applied.

While monitoring contaminants in the environment, the temperature and pH under which the enzymatic reaction proceeds strongly depends on properties of the application site, e.g. a normal groundwater system has a constant temperature derived from the average surface temperature and geothermal gradient, and pH in the range of 6.5 to 8 [100]. However, acidification caused by industry can lead to pH values lower than 5.5 [101]. Novel enzymes functional at low temperatures and pH would be an advantage for the preparation of biosensors suitable for such environments.

3.6 Cell imaging and protein analysis

The wealth of knowledge on HLDs provided a theoretical basis for the design of a versatile fusion tag (HaloTag) system, which can be used for both in vivo and in vitro analysis of mammalian proteins [102–106]. The technology is based on irreversible covalent binding of miscellaneously labeled ligands by the mutant dehalogenase DhaA, which forms a genetically encoded fusion with a target protein [102–105]. The covalent binding of a chlorinated ligand is enabled by the mutation of the catalytic base His272Phe, which impairs the final hydrolysis step and leads to irreversible trapping of the alkyl-enzyme intermediate (Fig. 1B) [104]. The synthetic ligands consist of a chloroalkane linker attached to various fluorescent dyes, affinity handles, or solid surfaces (Supporting information, Table S7).

This innovative application of HLDs has been developed by the Promega Corporation, which has demonstrated its successful application to fluorescence imaging [102–105]; protein expression and purification from *E. coli* [107] and mammalian cells [108, 109]; protein expression and immobilization from a cell-free extract [110]; and protein-protein or protein-nucleic acid interaction studies (including arrays and chip technology) [111–115]. The HaloTag technology is currently protected by a number of patents, e.g. [106, 116–118] and has been already utilized in about 40 independent studies (see Supporting information).

To overcome a low affinity of the wild type DhaA to long chlorinated ligands, saturation mutagenesis with high-throughput fluorescence screening was employed [105, 106]. The beneficial mutations (Lys175Met, Cys176Gly, and Tyr273Leu) were combined in a single mutant, synergistically improving binding rate by four orders of magnitude, corresponding to the apparent second-order rate constant $10^6\ M^{-1}\ s^{-1}$ [105]. Good expression of HaloTag fusion proteins in mammalian or *E. coli* cells was supported by removal of harmful sequence motifs and optimization of codon usage according to human codon bias, while avoiding low usage codons of *E. coli* [118]. HaloTag was further engineered for maximal soluble expression of fusion proteins in *E. coli* using rational design and directed evolution [107]. This optimization yielded a fusion tag with increased stability called HaloTag7.

4 Conclusions

To date, 17 HLDs have been characterized and nine of them possess properties suitable for biotechnological applications, as described in section 2. HLDs have already been successfully used for laboratory-scale production of optically pure organocompounds; to neutralize sulphur mustard; for bioremediation of groundwater contaminated with 1,2-dichloroethane and biodegradation of hexachlorocyclohexane; in biosensors suitable for the detection of environmental pollutants; and for protein tagging. Several other potential applications still require additional research to solve limiting factors, such as low expression, product inhibition, insufficient enzyme selectivity, low affinity and catalytic efficiency towards selected substrates, or instability in the presence of organic co-solvents. Recently introduced high-throughput methods for

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

exploiting natural diversity and laboratory protein evolution have significantly contributed to the preparation of optimized HLDs. We propose this trend will continue in the foreseeable future and will further expand the practical usefulness of this fascinating enzyme family.

The authors would like to thank T. Prudnikova, I. Kuta Smatanova and P. Rezacova for providing the unpublished structure of DbeA for analysis of the enzyme's active-site parameters and to E. Chovancova and J. Brezovsky for their useful comments on the manuscript. The work was supported by the Grant Agency of the Czech Republic (203/08/0114, P503/12/0572 and P207/12/0775), the Grant Agency of the Czech Academy of Sciences (IAA401630901) and the European Regional Development Fund (CZ.1.05/2.1.00/01.0001). The work of AP was supported by Brno PhD Talent Scholarship provided by Brno City Murlicipality.

Zbynek Prokop and Jiri Damborsky are co-founders of Enantis.

5 References

[1] Damborsky, J., Rorije, E., Jesenska, A., Nagata, Y. et al., Structure-specificity relationships for haloalkane dehalogenases. *Environ. Toxicol. Chem.* 2001, *20*, 2681–2689.

[2] Verschueren, K. H., Seljée, F., Rozeboom, H. J., Kalk, K. H., Dijkstra, B. W., Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature* 1993, *363*, 693–698.

[3] Damborsky, J., Koca, J., Analysis of the reaction mechanism and substrate specificity of haloalkane dehalogenases by sequential and structural comparisons. *Protein Eng.* 1999, *12*, 989–998.

[4] Hardman, D. J., Biotransformation of halogenated compounds. *Crit. Rev. Biotechnol.* 1991, *11*, 1–40.

[5] Leisinger, T., Bader, R., Microbial dehalogenation of synthetic organohalogen compounds – hydrolytic dehalogenases. *Chimia* 1993, *47*, 116–121.

[6] Pries, F., van der Ploeg, J., Dolfing, J., Janssen, D., Degradation of halogenated aliphatic-compounds – the role of adaptation. *FEMS Microbiol. Rev.* 1994, *15*, 279–295.

[7] Fetzner, S., Lingens, F., Bacterial dehalogenases – biochemistry, genetics, and biotechnological applications. *Microbiol. Rev.* 1994, *58*, 641–685.

[8] Janssen, D. B., Pries, F., van der Ploeg, J. R., Genetics and biochemistry of dehalogenating enzymes. *Annu. Rev. Microbiol.* 1994, *48*, 163–191.

[9] Slater, J., Bull, A., Hardman, D., Microbial dehalogenation. *Biodegradation* 1995, *6*, 181–189.

[10] Leisinger, T., Biodegradation of chlorinated aliphatic compounds. *Curr. Opin. Biotechnol.* 1996, *7*, 295–300.

[11] Slater, J., Bull, A., Hardman, D., Microbial dehalogenation of halogenated alkanoic acids, alcohols and alkanes. *Adv. Microb. Physiol.* 1997, *38*, 133–176.

[12] Copley, S. D., Microbial dehalogenases: enzymes recruited to convert xenobiotic substrates. *Curr. Opin. Chem. Biol.* 1998, *2*, 613–617.

[13] Fetzner, S., Bacterial dehalogenation. *Appl. Microbiol. Biotechnol.* 1998, *50*, 633–657.

Jiri Damborsky is the Loschmidt Chair Professor at the Department of Experimental Biology and the Research Centre for Toxic Compounds in the Environment, Masaryk University, Brno, Czech Republic. He received his master's degree in 1993 and a PhD in Microbiology in 1997 at Masaryk University. His research interests lie in the field of protein engineering and synthetic biology. His research group develops new concepts and software tools for the rational design of enzymes and bacteria with improved properties for biocatalysis, biotransformation and biosensing. He has published more than 100 original articles and co-authored two patents for the use of enzymes in decontamination and biocatalysis. He is a co-founder of the biotechnology spin-off company Enantis and a holder of the EMBO/HHMI Scientist award of the European Molecular Biology Organisation and the Howard Hughes Medical Institute.

Tana Koudelakova works as a research specialist in the Loschmidt Laboratories at Masaryk University, Brno, Czech Republic. She received her master's degree in molecular biology and genetics in 2004 and a PhD in environmental chemistry in 2011 at Masaryk University. Her main research interests have been focused on engineering substrate specificity, enantioselectivity and stability of haloalkane dehalogenases. Thanks to scholarships from University of Greifswald (2007) and the Federation of European Microbiological Societies (2008), she has acquired skills in the field of directed evolution in the laboratory of Professor U. T. Bornscheuer, University of Greifswald, Germany. Tana Koudelakova is a co-author of eight original articles and one patent application in the fields of biochemistry and biotechnology.

[14] Swanson, P., Dehalogenases applied to industrial-scale biocatalysis. *Curr. Opin. Biotech.* 1999, *10*, 365–369.

[15] Janssen, D., Oppentocht, J., Poelarends, G., Microbial dehalogenation. *Curr. Opin. Biotechnol.* 2001, *12*, 254–258.

[16] Janssen, D. B., Evolving haloalkane dehalogenases. *Curr. Opin. Chem. Biol.* 2004, *8*, 150–159.

[17] Janssen, D. B., Dinkla, I. J. T., Poelarends, G. J., Terpstra, P., Bacterial degradation of xenobiotic compounds: evolution and distribution of novel enzyme activities. *Environ. Microbiol.* 2005, *7*, 1868–1882.

[18] Janssen, D., Biocatalysis by dehalogenating enzymes. *Adv. Appl. Microbiol.* 2007, *61*, 233–252.

[19] Scott, C., Pandey, G., Hartley, C. J., Jackson, C. J. et al., The enzymatic basis for pesticide bioremediation. *Indian J. Microbiol.* 2008, *48*, 65–79.

[20] Gray, K. A., Richardson, T. H., Robertson, D. E., Swanson, P. E., Subramanian, M. V., Soil-based gene discovery: a new technology to accelerate and broaden biocatalytic applications. *Adv. Appl. Microbiol.* 2003, *52*, 1–27.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[21] Johri, A. K., Dua, M., Tuteja, D., Saxena, R. et al., Genetic manipulations of microorganisms for the degradation of hexachlorocyclohexane. *FEMS Microbiol. Rev.* 1996, *19*, 69–84.
[22] Kumar, S., Mukerji, K., Lal, R., Molecular aspects of pesticide degradation by microorganisms. *Crit. Rev. Microbiol.* 1996, *22*, 1–26.
[23] Lal, R., Dadhwal, M., Kumari, K., Sharma, P. et al., *Pseudomonas* sp. to *Sphingobium indicum*: a journey of microbial degradation and bioremediation of hexachlorocyclohexane. *Indian J. Microbiol.* 2008, *48*, 3–18.
[24] Lal, R., Pandey, G., Sharma, P., Kumari, K. et al., Biochemistry of microbial degradation of hexachlorocyclohexane and prospects for bioremediation. *Microbiol. Mol. Biol. Rev.* 2010, *74*, 58–80.
[25] Damborsky, J., Chaloupkova, R., Pavlova, M., Chovancova, E., Brezovsky, J., Structure-function relationships and engineering of haloalkane dehalogenases, in: Timmis, K. N. (Ed.), *Handbook of Hydrocarbon and Lipid Microbiology*, Springer, Berlin, Heidelberg 2010, pp. 1081–1098.
[26] Keuning, S., Janssen, D. B., Witholt, B., Purification and characterization of hydrolytic haloalkane dehalogenase from *Xanthobacter autotrophicus* GJ10. *J. Bacteriol.* 1985, *163*, 635–639.
[27] Yokota, T., Omori, T., Kodama, T., Purification and properties of haloalkane dehalogenase from *Corynebacterium* sp. strain m15-3. *J. Bacteriol.* 1987, *169*, 4049–4054.
[28] Nagata, Y., Miyauchi, K., Damborsky, J., Manova, K. et al., Purification and characterization of a haloalkane dehalogenase of a new substrate class from a gamma-hexachlorocyclohexane-degrading bacterium, *Sphingomonas paucimobilis* UT26. *Appl. Environ. Microbiol.* 1997, *63*, 3707–3710.
[29] Jesenska, A., Bartos, M., Czernekova, V., Rychlik, I. et al., Cloning and expression of the haloalkane dehalogenase gene *dhmA* from *Mycobacterium avium* N85 and preliminary characterization of DhmA. *Appl. Environ. Microbiol.* 2002, *68*, 3724–3730.
[30] Jesenska, A., Pavlova, M., Strouhal, M., Chaloupkova, R. et al., Cloning, biochemical properties, and distribution of mycobacterial haloalkane dehalogenases. *Appl. Environ. Microbiol.* 2005, *71*, 6736–6745.
[31] Sato, Y., Monincova, M., Chaloupkova, R., Prokop, Z. et al., Two rhizobial strains, *Mesorhizobium loti* MAFF303099 and *Bradyrhizobium japonicum* USDA110, encode haloalkane dehalogenases with novel structures and substrate specificities. *Appl. Environ. Microbiol.* 2005, *71*, 4372–4379.
[32] Pavlova, M., Klvana, M., Jesenska, A., Prokop, Z. et al., The identification of catalytic pentad in the haloalkane dehalogenase DhmA from *Mycobacterium avium* N85: reaction mechanism and molecular evolution. *J. Struct. Biol.* 2007, *157*, 384–392.
[33] Jesenska, A., Monincova, M., Koudelakova, T., Hasan, K. et al., Biochemical characterization of haloalkane dehalogenases DrbA and DmbC, representatives of a novel subfamily. *Appl. Environ. Microbiol.* 2009, *75*, 5157–5160.
[34] Chan, W. Y., Wong, M., Guthrie, J., Savchenko, A. V. et al., Sequence- and activity-based screening of microbial genomes for novel dehalogenases. *Microb. Biotechnol.* 2010, *3*, 107–120.
[35] Hasan, K., Fortova, A., Koudelakova, T., Chaloupkova, R. et al., Biochemical characteristics of the novel haloalkane dehalogenase DatA, isolated from the plant pathogen *Agrobacterium tumefaciens* C58. *Appl. Environ. Microbiol.* 2011, *77*, 1881–1884.
[36] Hesseler, M., Bogdanovic, X., Hidalgo, A., Berenguer, J. et al., Cloning, functional expression, biochemical characterization, and structural analysis of a haloalkane dehalogenase from *Plesiocystis pacifica* SIR-1. *Appl. Microbiol. Biotechnol.* 2011, *91*,1049–1060.
[37] Gehret, J. J., Gu, L., Geders, T. W., Brown, W. C. et al., Structure and activity of DmmA, a marine haloalkane dehalogenase. *Protein Sci.* 2012, *21*, 239–248.
[38] Drienovska, I., Chovancova, E., Koudelakova, T., Damborsky, J., Chaloupkova, R., Biochemical characterization of a novel haloalkane dehalogenase from a cold-adapted bacterium. *Appl. Environ. Microbiol.* 2012, *78*, 4995–4998.
[39] Janssen, D. B., Pries, F., van der Ploeg, J., Kazemier, B. et al., Cloning of 1,2-dichloroethane degradation genes of *Xanthobacter autotrophicus* GJ10 and expression and sequencing of the *dhlA* gene. *J. Bacteriol.* 1989, *171*, 6791–6799.
[40] Nagata, Y., Natsui, S., Endo, R., Ohtsubo, Y. et al., Genomic organization and genomic structural rearrangements of *Sphingobium japonicum* UT26, an archetypal γ-hexachlorocyclohexane-degrading bacterium. *Enzyme Microb. Tech.* 2011, *49*, 499–508.
[41] Curragh, H., Flynn, O., Larkin, M. J., Stafford, T. M. et al., Haloalkane degradation and assimilation by *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology* 1994, *140*, 1433–1442.
[42] Kulakova, A. N., Larkin, M. J., Kulakov, L. A., The plasmid-located haloalkane dehalogenase gene from *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology* 1997, *143,* 109–115.
[43] Poelarends, G. J., Kulakov, L. A., Larkin, M. J., van Hylckama Vlieg, J. E., Janssen, D. B., Roles of horizontal gene transfer and gene integration in evolution of 1,3-dichloropropene- and 1,2-dibromoethane-degradative pathways. *J. Bacteriol.* 2000, *182*, 2191–2199.
[44] Prokop, Z., Oplustil, F., de Frank, J., Damborsky, J., Enzymes fight chemical weapons. *Biotechnol. J.* 2006, *1*, 1370–1380.
[45] Prokop, Z., Sato, Y., Brezovsky, J., Mozga, T. et al., Enantioselectivity of haloalkane dehalogenases and its modulation by surface loop engineering. *Angew. Chem. Int. Ed. Engl.* 2010, *49*, 6111–6115.
[46] Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M. et al., Substrate specificity of haloalkane dehalogenases. *Biochem. J.* 2011, *435*, 345–354.
[47] Westerbeek, A., Szymanski, W., Wijma, H. J., Marrink, S. J. et al., Kinetic resolution of alpha-bromoamides: experimental and theoretical investigation of highly enantioselective reactions catalyzed by haloalkane dehalogenases. *Adv. Synth. Catal.* 2011, *353*, 931–944.
[48] Schanstra, J. P., Kingma, J., Janssen, D. B., Specificity and kinetics of haloalkane dehalogenase. *J. Biol. Chem.* 1996, *271*, 14747–14753.
[49] Bosma, T., Pikkemaat, M. G., Kingma, J., Dijk, J., Janssen, D. B., Steady-state and pre-steady-state kinetic analysis of halopropane conversion by a *Rhodococcus* haloalkane dehalogenase. *Biochemistry* 2003, *42*, 8047–8053.
[50] Schindler, J. F., Naranjo, P. A., Honaberger, D. A., Chang, C. H. et al., Haloalkane dehalogenases: steady-state kinetics and halide inhibition. *Biochemistry* 1999, *38*, 5772–5778.
[51] Kmunicek, J., Hynkova, K., Jedlicka, T., Nagata, Y. et al., Quantitative analysis of substrate specificity of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *Biochemistry* 2005, *44*, 3390–3401.
[52] Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R. et al., Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. *Nat. Chem. Biol.* 2009, *5*, 727–733.
[53] Chaloupkova, R., Sykorova, J., Prokop, Z., Jesenska, A. et al., Modification of activity and specificity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26 by engineering of its entrance tunnel. *J. Biol. Chem.* 2003, *278*, 52622–52628.
[54] Bar-Even, A., Noor, E., Savir, Y., Liebermeister, W. et al., The moderately efficient enzyme: evolutionary and physicochemical trends shaping enzyme parameters. *Biochemistry* 2011, *50*, 4402–4410.
[55] Janssen, D. B., Gerritse, J., Brackman, J., Kalk, C. et al., Purification and characterization of a bacterial dehalogenase with activity toward halogenated alkanes, alcohols and ethers. *Eur. J. Biochem.* 1988, *171*, 67–72.
[56] Pieters, R. J., Lutje Spelberg, J. H., Kellogg, R. M., Janssen, D. B., The enantioselectivity of haloalkane dehalogenases. *Tetrahedron Lett.* 2001, *42*, 469–471.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[57] Prokop, Z., Damborsky, J., Janssen, D. B., Nagata, Y., *Method of production of optically active halohydrocarbons and alcohols using hydrolytic dehalogenation catalysed by haloalkane dehalogenases*. US Patent 7632666, 2009.

[58] Chaloupkova, R., Prokop, Z., Sato, Y., Nagata, Y., Damborsky, J., Stereoselectivity and conformational stability of haloalkane dehalogenase DbjA from *Bradyrhizobium japonicum* USDA110: the effect of pH and temperature. *FEBS J.* 2011, *278*, 2728–2738.

[59] Patel, R. N., Biocatalysis: synthesis of chiral intermediates for pharmaceuticals. *Curr. Org. Chem.* 2006, *10*, 1289–1321.

[60] Biedermann, J., Leon-Lomeli, A., Borbe, H. O., Prop, G., Two stereoisomeric imidazoline derivatives. Synthesis, optical, and alpha 2-adrenoceptor activities. *J. Med. Chem.* 1986, *29*, 1183–1188.

[61] Szymanski, W., Westerbeek, A., Janssen, D. B., Feringa, B. L., A simple enantioconvergent and chemoenzymatic synthesis of optically active alpha-substituted amides. *Angew. Chem. Int. Ed. Engl.* 2011, *50*, 10712–10715.

[62] Luetz, S., Giver, L., Lalonde, J., Engineered enzymes for chemical production. *Biotechnol. Bioeng.* 2008, *101*, 647–653.

[63] Westerbeek, A., Szymanski, W., Feringa, B. L., Janssen, D. B., Dynamic kinetic resolution process employing haloalkane dehalogenase. *ACS Catal.* 2011, *1*, 1654–1660.

[64] Tufvesson, P., Lima-Ramos, J., Nordblad, M., Woodley, J. M., Guidelines and cost analysis for catalyst production in biocatalytic processes. *Org. Process Res. Dev.* 2011, *15*, 266–274.

[65] Cheetham, P., What makes a good biocatalyst? *J. Biotechnol.* 1998, *66*, 3–10.

[66] Swanson, P. E., *Method for conversion of halogenated hydrocarbons to halohydrins*. US Patent 5372944, 1994.

[67] Dravis, B., Swanson, P., Russell, A., Haloalkane hydrolysis with an immobilized haloalkane dehalogenase. *Biotechnol. Bioeng.* 2001, *75*, 416–423.

[68] Gray, K., Richardson, T., Kretz, K., Short, J. et al., Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. *Adv. Synth. Catal.* 2001, *343*, 607–617.

[69] Bosma, T., Damborsky, J., Stucki, G., Janssen, D., Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. *Appl. Environ. Microbiol.* 2002, *68*, 3582–3587.

[70] Zhao, H., A pH-indicator-based screen for hydrolytic haloalkane dehalogenase, in: Arnold, F. H., Georgiou, G. (Eds.), *Directed Enzyme Evolution: Screening and Selection Methods*, Humana Press, New Jersey, 2003, pp. 213–222.

[71] Banas, P., Otyepka, M., Jerabek, P., Petrek, M., Damborsky, J., Mechanism of enhanced conversion of 1,2,3-trichloropropane by mutant haloalkane dehalogenase revealed by molecular modeling. *J. Comput. Aided Mol. Des.* 2006, *20*, 375–383.

[72] Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R. et al., Pathways and mechanisms for product release in the engineered haloalkane dehalogenases explored using classical and random acceleration molecular dynamics simulations. *J. Mol. Biol.* 2009, *392*, 1339–1356.

[73] van Leeuwen, J. G. E., Wijma, H. J., Floor, R. J., van der Laan, J.-M., Janssen, D. B., Directed evolution strategies for enantiocomplementary haloalkane dehalogenases: from chemical waste to enantiopure building blocks. *ChemBioChem* 2012, *13*, 137–148.

[74] Kretz, K. A., Richardson, T. H., Gray, K. A., Robertson, D. E. et al., Gene site saturation mutagenesis: a comprehensive mutagenesis approach. *Meth. Enzymol.* 2004, *388*, 3–11.

[75] Mena-Benitez, G. L., Gandia-Herrero, F., Graham, S., Larson, T. R. et al., Engineering a catabolic pathway in plants for the degradation of 1,2-dichloroethane. *Plant Physiol.* 2008, *147*, 1192–1198.

[76] Stucki, G., Thuer, M., Experiences of a large-scale application of 1,2-dichloroethane degrading microorganisms for groundwater treatment. *Environ. Sci. Technol.* 1995, *29*, 2339–2345.

[77] Freitas dos Santos, L. M., Livingston, A. G., Novel membrane bioreactor for detoxification of VOC wastewaters: biodegradation of 1,2-dichloroethane. *Water Res.* 1995, *29*, 179–194.

[78] van den Wijngaard, A. J., Janssen, D. B., Witholt, B., Degradation of epichlorohydrin and halohydrins by bacterial cultures isolated from freshwater sediment. *J. Gen. Microbiol.* 1989, *135*, 2199–2208.

[79] Bosma, T., Kruizinga, E., de Bruin, E. J., Poelarends, G. J., Janssen, D. B., Utilization of trihalogenated propanes by *Agrobacterium radiobacter* AD1 through heterologous expression of the haloalkane dehalogenase from *Rhodococcus* sp. strain M15-3. *Appl. Environ. Microbiol.* 1999, *65*, 4575–4581.

[80] Erable, B., Goubet, I., Lamare, S., Seltana, A. et al., Nonconventional hydrolytic dehalogenation of 1-chlorobutane by dehydrated bacteria in a continuous solid-gas biofilter. *Biotechnol. Bioeng.* 2005, *91*, 304–313.

[81] Erable, B., Maugard, T., Goubet, I., Lamare, S., Legoy, M. D., Biotransformation of halogenated compounds by lyophilized cells of *Rhodococcus erythropolis* in a continuous solid-gas biofilter. *Process Biochem.* 2005, *40*, 45–51.

[82] Erable, B., Goubet, I., Lamare, S., Legoy, M. D., Maugard, T., Haloalkane hydrolysis by *Rhodococcus erythropolis* cells: comparison of conventional aqueous phase dehalogenation and nonconventional gas phase dehalogenation. *Biotechnol. Bioeng.* 2004, *86*, 47–54.

[83] Erable, B., Goubet, I., Seltana, A., Maugard, T., Non-conventional gas phase remediation of volatile halogenated compounds by dehydrated bacteria. *J. Environ. Manage.* 2009, *90*, 2841–2844.

[84] Bala, K., Geueke, B., Miska, M. E., Rentsch, D. et al., Enzymatic conversion of ε-hexachlorocyclohexane and a heptachlorocyclohexane isomer, two neglected components of technical hexachlorocyclohexane. *Environ. Sci. Technol.* 2012, *46*, 4051–4058.

[85] Heeb, N. V., Zindel, D., Geueke, B., Kohler, H.-P. E., Lienemann, P., Biotransformation of hexabromocyclododecanes (HBCDs) with LinB-an HCH-converting bacterial enzyme. *Environ. Sci. Technol.* 2012, *46*, 6566–6574.

[86] Naested, H., Fennema, M., Hao, L., Andersen, M. et al., A bacterial haloalkane dehalogenase gene as a negative selectable marker in *Arabidopsis*. *Plant J.* 1999, *18*, 571–576.

[87] Uchida, E., Ouchi, T., Suzuki, Y., Yoshida, T. et al., Secretion of bacterial xenobiotic-degrading enzymes from transgenic plants by an apoplastic expressional system: an applicability for phytoremediation. *Environ. Sci. Technol.* 2005, *39*, 7671–7677.

[88] Jez, J. M., Toward protein engineering for phytoremediation: possibilities and challenges. *Int. J. Phytoremediation* 2011, *13*, 77–89.

[89] Prokop, Z., Damborsky, J., Oplustil, F., Jesenska, A., Nagata, Y., *Method of detoxication of yperite by using haloalkane dehalogenases*. US Patent 7888103, 2011.

[90] Alamo-Bethencourt, V., Aldridge, S., Coombs, A., DeFrancesco, L., Huggett, B., Osborne, R., Mustard gas enzyme, news in brief. *Nat. Biotech.* 2007, *25*, 1197–1198.

[91] Blum, M.-M., Richardt, A., in: Richardt A., Blum, M.-M. (Eds.), *Decontamination of Warfare Agents*, Wiley-VCH Verlag, Weinheim 2008, pp. 135–162.

[92] Peter, J., Hutter, W., Stöllnberger, W., Hampel, W., Detection of chlorinated and brominated hydrocarbons by an ion sensitive whole cell biosensor. *Biosens. Bioelectron.* 1996, *11*, 1215–1219.

[93] Campbell, D., Müller, C., Reardon, K., Development of a fiber optic enzymatic biosensor for 1,2-dichloroethane. *Biotechnol. Lett.* 2006, *28*, 883–887.

[94] Reardon, K. F., Campbell, D. W., Müller, C., Optical fiber enzymatic biosensor for reagentless measurement of ethylene dibromide. *Eng. Life Sci.* 2009, *9*, 291–297.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

[95] Bidmanova, S., Chaloupkova, R., Damborsky, J., Prokop, Z., Development of an enzymatic fiber-optic biosensor for detection of halogenated hydrocarbons. *Anal. Bioanal. Chem.* 2010, *398*, 1891–1898.
[96] Hutter, W., Peter, J., Swoboda, H., Hampel, W. et al., Development of a microbial bioassay for chlorinated and brominated hydrocarbons. *Anal. Chim. Acta* 1995, *306*, 237–241.
[97] Peter, J., Hutter, W., Stöllnberger, W., Karner, F., Hampel, W., Semicontinuous detection of 1,2-dichloroethane in water samples using *Xanthobacter autotrophicus* GJ10 encapsulated in chitosan beads. *Anal. Chem.* 1997, *69*, 2077–2079.
[98] Bidmanova, S., Hlavacek, A., Damborsky, J., Prokop, Z., Conjugation of 5(6)-carboxyfluorescein and 5(6)-carboxynaphthofluorescein with bovine serum albumin and their immobilization for optical pH sensing. *Sens. Actuators. B Chem.* 2012, *161*, 93–99.
[99] Johnson, A. K., Zawadzka, A. M., Deobald, L. A., Crawford, R. L., Paszczynski, A. J., Novel method for immobilization of enzymes to magnetic nanoparticles. *J. Nanopart. Res.* 2008, *10*, 1009–1025.
[100] Jordana, S., Piera, E. B., Natural groundwater quality and health. *Geol. Acta* 2004, *2*, 175–188.
[101] Scheidleder, A., Grath, J., Winkler, G., Stärk, U. et al., *Groundwater quality and quantity in Europe*, Office for Official Publications of the European Communities, Copenhagen, 1999, pp. 78–82.
[102] Los, G. V., Darzins, A., Karassina, N., Zimprich, C. et al., HaloTag interchangeable labeling technology for cell imaging and protein capture. *Cell Notes* 2005, *11*, 2–6.
[103] Los, G., Learish, R., Karassina, N., Zimprich, C. et al., HaloTag technology: cell imaging and protein analysis. *Cell Notes* 2006, *14*, 10–14.
[104] Los, G. V., Wood, K., The HaloTag: a novel technology for cell imaging and protein analysis. *Methods Mol. Biol.* 2007, *356*, 195–208.
[105] Los, G., Encell, L., McDougall, M., Hartzell, D. et al., HaloTag: A novel protein labeling technology for cell imaging and protein analysis. *ACS Chem. Biol.* 2008, *3*, 373–382.
[106] Darzins, A., Encell, L., Los, G. V., Wood, K. V. et al., *Covalent tethering of functional groups to proteins and substrates therefor*. US Patent 7425436, 2008.
[107] Ohana, R. F., Encell, L. P., Zhao, K., Simpson, D. et al., HaloTag7: a genetically engineered tag that enhances bacterial expression of soluble proteins and improves protein purification. *Protein Expr. Purif.* 2009, *68*, 110–120.
[108] Mendez, J. L., Ohana, R. F., Hurst, R., Murphy, N. et al., Highly efficient protein and complex purification from mammalian cells using the HaloTag technology. *Biotechniques* 2011, *51*, 276–277.
[109] Ohana, R. F., Hurst, R., Vidugiriene, J., Slater, M. R. et al., HaloTag-based purification of functional human kinases from mammalian cells. *Protein Expr. Purif.* 2011, *76*, 154–164.
[110] Leippe, D. M., Zhao, K. Q., Hsiao, K., Slater, M. R., Cell-free expression of protein kinase a for rapid activity assays. *Anal. Chem. Insights* 2010, *5*, 25–36.
[111] Nath, N., Hurst, R., Hook, B., Meisenheimer, P. et al., Improving protein array performance: focus on washing and storage conditions. *J. Proteome Res.* 2008, *7*, 4475–4482.
[112] Urh, M., Hartzell, D., Mendez, J., Klaubert, D. H., Wood, K., Methods for detection of protein-protein and protein-DNA interactions using HaloTag. *Methods Mol. Biol.* 2008, *421*, 191–209.
[113] Hartzell, D. D., Trinklein, N. D., Mendez, J., Murphy, N. et al., A functional analysis of the CREB signaling pathway using HaloCHIP-chip and high throughput reporter assays. *BMC Genomics* 2009, *10*.
[114] Hurst, R., Hook, B., Slater, M. R., Hartnett, J. et al., Protein-protein interaction studies on protein arrays: Effect of detection strategies on signal-to-background ratios. *Anal. Biochem.* 2009, *392*, 45–53.
[115] Daniels, D. L., Mendez, J., Mosley, A. L., Ramisetty, S. R. et al., Examining the complexity of human RNA polymerase complexes using HaloTag technology coupled to label free quantitative proteomics. *J. Proteome Res.* 2012, *11*, 564–575.
[116] Darzins, A., Encell, L., Johnson, T., Klaubert, D. et al., *Method of immobilizing a protein or molecule via a mutant dehalogenase that is bound to an immobilized dehalogenase substrate and linked directly or indirectly to the protein or molecule*. US Patent 7888086, 2011.
[117] Wood, K. V., Klaubert, D., Los, G. V., Bulleit, R. F. et al., *Covalent tethering of functional groups to proteins*. US Patent RE42931, 2011.
[118] Wood, K. V., Klaubert, D., Los, G. V., Bulleit, R. F. et al., *Compositions comprising a dehalogenase substrate and a fluorescent label and methods of use*. US Patent 7867726, 2011.
[119] Sallis, P. J., Armfield, S. J., Bull, A. T., Hardman, D. J., Isolation and characterization of a haloalkane halidohydrolase from *Rhodococcus erythropolis* Y2. *J. Gen. Microbiol.* 1990, *136*, 115–120.
[120] Newman, J., Peat, T. S., Richard, R., Kan, L. et al., Haloalkane dehalogenases: structure of a *Rhodococcus* enzyme. *Biochemistry* 1999, *38*, 16105–16114.
[121] Liu, X., Hanson, B. L., Langan, P., Viola, R. E., The effect of deuteration on protein structure: a high-resolution comparison of hydrogenous and perdeuterated haloalkane dehalogenase. *Acta Crystallogr. D Biol. Crystallogr.* 2007, *63*, 1000–1008.
[122] Mazumdar, P. A., Hulecki, J. C., Cherney, M. M., Garen, C. R., James, M. N. G., X-ray crystal structure of *Mycobacterium tuberculosis* haloalkane dehalogenase Rv2579. *Biochim. Biophys. Acta* 2008, *1784*, 351–362.
[123] Oakley, A. J., Klvana, M., Otyepka, M., Nagata, Y. et al., Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26 at 0.95 Å resolution: dynamics of catalytic residues. *Biochemistry* 2004, *43*, 870–878.
[124] Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I. et al., ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 2003, *31*, 3784–3788.
[125] Jesenska, A., Sykora, J., Olzynska, A., Brezovsky, J. et al., Nanosecond time-dependent Stokes shift at the tunnel mouth of haloalkane dehalogenases. *J. Am. Chem. Soc.* 2009, *131*, 494–501.

© 2012 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

doi:10.1016/j.jmb.2009.06.076 J. Mol. Biol. (2009) 392, 1339–1356

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Pathways and Mechanisms for Product Release in the Engineered Haloalkane Dehalogenases Explored Using Classical and Random Acceleration Molecular Dynamics Simulations

Martin Klvana[1], Martina Pavlova[1], Tana Koudelakova[1], Radka Chaloupkova[1], Pavel Dvorak[1], Zbynek Prokop[1], Alena Stsiapanava[2], Michal Kuty[2,3], Ivana Kuta-Smatanova[2,3], Jan Dohnalek[4], Petr Kulhanek[5], Rebecca C. Wade[6] and Jiri Damborsky[1]*

[1]*Loschmidt Laboratories, Institute of Experimental Biology and National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5/A4, 625 00 Brno, Czech Republic*

[2]*Institute of Physical Biology, University of South Bohemia, Zámek 136, 373 33 Nové Hrady, Czech Republic*

[3]*Institute of Systems Biology and Ecology, Academy of Sciences of the Czech Republic, Zámek 136, 373 33 Nové Hrady, Czech Republic*

[4]*Institute of Macromolecular Chemistry, Academy of Sciences of the Czech Republic, Heyrovského náměstí 2, 162 06 Praha 6, Czech Republic*

[5]*Laboratory of Computational Chemistry, National Centre for Biomolecular Research, Faculty of Science, Masaryk University, Kamenice 5/A4, 625 00 Brno, Czech Republic*

Eight mutants of the DhaA haloalkane dehalogenase carrying mutations at the residues lining two tunnels, previously observed by protein X-ray crystallography, were constructed and biochemically characterized. The mutants showed distinct catalytic efficiencies with the halogenated substrate 1,2,3-trichloropropane. Release pathways for the two dehalogenation products, 2,3-dichloropropane-1-ol and the chloride ion, and exchange pathways for water molecules, were studied using classical and random acceleration molecular dynamics simulations. Five different pathways, denoted p1, p2a, p2b, p2c, and p3, were identified. The individual pathways showed differing selectivity for the products: the chloride ion releases solely through p1, whereas the alcohol releases through all five pathways. Water molecules play a crucial role for release of both products by breakage of their hydrogen-bonding interactions with the active-site residues and shielding the charged chloride ion during its passage through a hydrophobic tunnel. Exchange of the chloride ions, the alcohol product, and the waters between the buried active site and the bulk solvent can be realized by three different mechanisms: (i) passage through a permanent tunnel, (ii) passage through a transient tunnel, and (iii) migration through a protein matrix. We demonstrate that the accessibility of the pathways and the mechanisms of ligand exchange were modified by mutations. Insertion of bulky aromatic residues in the tunnel corresponding to pathway p1 leads to reduced accessibility to the ligands and a change in mechanism of opening from permanent to transient. We propose that engineering the accessibility of tunnels and the mechanisms of ligand exchange is a powerful strategy for modification of the functional properties of enzymes with buried active sites.

© 2009 Elsevier Ltd. All rights reserved.

**Corresponding author.* E-mail address: jiri@chemi.muni.cz.

Abbreviations used: CC loop, C-terminal cap domain loop (α7/α8 loop); CL, chloride anion; DCL, 2,3-dichloropropane-1-ol; TCP, 1,2,3-trichloropropane; MD, molecular dynamics; NC loop, N-terminal cap domain loop (β6/α4 loop); RAMD, random acceleration molecular dynamics; PDB, Protein Data Bank.

0022-2836/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.

[6]*Molecular and Cellular Modeling Group, EML Research, Schloss-Wolfsbrunnenweg 33, D-69118 Heidelberg, Germany*

Received 11 March 2009; received in revised form 25 June 2009; accepted 29 June 2009
Available online 3 July 2009

Edited by D. Case

Keywords: haloalkane dehalogenase; product release; random acceleration molecular dynamics; tunnel; water exchange

Introduction

Many globular enzymes possess active sites buried in the protein core, and there is growing evidence that the access of substrates to the active site or release of products can be a determinant of their catalytic activity[1–13] and substrate specificity.[1–3,5–8,10,13–21] Mechanisms of ligand exchange between buried active sites and bulk solvent and the effects of mutations on the exchange process are often less well understood than the mechanisms of chemical reactions taking place in the active sites.

The wealth of knowledge that has been acquired about haloalkane dehalogenases (EC 3.8.1.5) in the past two decades makes these enzymes a good model system to study fundamental principles of enzymatic function. Haloalkane dehalogenases belong to the α/β-hydrolase superfamily of enzymes[22] and catalyze hydrolytic dehalogenation of various halogenated aliphatic hydrocarbons to a corresponding alcohol and a halide.[23–25] The reaction is accomplished by a catalytic pentad composed of a nucleophile, a base, a catalytic acid, and two halide-stabilizing residues.[26–29] Two kinetically observable chemical steps are as follows: (i) bimolecular nucleophilic substitution (S_N2) leading to the formation of a halide anion and alkyl-enzyme intermediate and (ii) nucleophilic addition (Ad_N) of a water molecule yielding a tetrahedral intermediate.[30–32] The reaction takes place in a hydrophobic active-site cavity located at the interface of the α/β-hydrolase domain and the helical cap domain.[27,33] The binding of hydrophobic substrates in the buried active site is favorable, shielding the reactive center from bulk water.[33] On the other hand, halide and alcohol products formed during the reaction must be released to allow another substrate molecule to enter the active site for the next catalytic cycle and this process can be rate-limiting.[30,31]

Two tunnels connecting the buried active site with the bulk solvent can be identified in the X-ray crystal structures of DhaA from *Rhodococcus* sp. available in the Protein Data Bank (PDB) structural database: 1CQW, 1BN6, and 1BN7.[27] These tunnels were previously named the main tunnel and the slot tunnel[34,35] and can serve as potential product release pathways. The main tunnel is open in all three structures, whereas the slot tunnel is open only in the structure 1BN6 with five ordered water sites in its mouth opening. A surface representation of structures 1CQW and 1BN7 shows that the slot tunnel is blocked and indicated only by a bulge protruding from the active site. Besides the two tunnels, a significantly deep surface depression near the slot mouth is nearly connected to the active site in all the three structures. The active site is hydrated by three water molecules in 1CQW with iodide bound between the two halide-stabilizing residues. On the other hand, the structure 1BN7 contains an acetate molecule in the active site and the active site of the structure 1BN6 contains an unidentified ligand.[27] Based on available structural information, several intriguing questions regarding the function of the tunnels and the role of solvent for release of products from the buried active site of DhaA can be formulated. Which of the two tunnels is used for release of a halide and an alcohol and for exchange of water molecules between the buried active site and bulk solvent? Are there any other product release pathways and water exchange pathways besides the two tunnels observed in the crystal structures? What is the mechanism for release of products from the buried active site? Could we change accessibility of the pathways by introducing mutations in the tunnels?

Molecular dynamics (MD) simulation is an established approach to study the dynamic behavior of proteins and can be applied for the investigation of release pathways for the two dehalogenation products of 1,2,3-trichloropropane (TCP) conversion in DhaA, that is, chloride anion (CL) and 2,3-dichloropropane-1-ol (DCL). The process of product release from the buried active site of DhaA to bulk solvent may occur on millisecond or longer time scales[31] and, therefore, can be too computationally demanding for classical MD simulations. This problem can be handled by the use of random acceleration molecular dynamics (RAMD) simulations.[14] RAMD is an enhanced sampling technique that

makes ligand release from a buried enzyme active site observable in computationally accessible simulation times.[15] RAMD resembles classical MD simulation except that an additional force is applied to the center of mass of the ligand in a randomly chosen direction. RAMD has been used to investigate substrate and product release pathways in cytochrome P450 enzymes,[14–17] unbinding of retinoic acid from retinoic acid receptor,[36] and release pathways for retinal in rhodopsin.[37]

Here, MD simulations are used to study release of products and exchange of water solvent in the wild-type haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCIMB13064 and eight DhaA mutants. The mutants carry substitutions in the residues lining two tunnels observed in the crystal structures of DhaA. All possible release pathways for two products and exchange pathways for the water molecules (ligands) are identified, and the mechanisms of ligand exchange between the buried active site and bulk solvent are distinguished. The effects of mutations on accessibility of the pathways and the mechanisms of ligand exchange are analyzed and compared with the crystal structures of the wild-type enzyme and three of the mutants.

Results

Construction and characterization of mutants with modified tunnels

Four variants of DhaA (mutants 04, 21, 27, and 31) carrying mutations in the residues lining the tunnels identified by protein X-ray crystallography were obtained by focused directed evolution of DhaA towards improved activity with TCP.[38,39] Here, we complemented this set by another four protein variants (mutants 14, 15, 51, and 52) constructed by site-directed mutagenesis with the aim of introducing additional variation in the main tunnel and the slot tunnel (Table 1 and Fig. 1). Mutant 14 was designed to contain a bulky residue in position 135 located in the slot tunnel. This variant represented a counterpart to mutant 04 carrying a bulky substitution in the main tunnel. Mutant 15 combined bulky residues in the slot tunnel (position 135) and the main tunnel (position 176). Mutants 51 and 52 carried bulky residues in four and five varied positions, respectively (Fig. 1). All constructed mutants were characterized for their activity with TCP using steady-state kinetics, for binding of DCL using inhibition kinetics, and for proper folding by circular dichroism (CD) spectroscopy.

All mutants showed an increase in the rate of TCP conversion except mutant 14, confirming the earlier proposal that a bulky residue at position 176 is essential for enhanced activity with TCP.[40] The highest increase in activity was determined for mutants carrying two aromatic substitutions in the main tunnel. The Michaelis constants determined for the mutants with TCP were similar to those for wild-type DhaA in all variants except mutant 51. An increase in the Michaelis constant for mutant 51 was compensated by the A145F mutation in mutant 52 (Table 1). The same level of K_m of the wild-type enzyme and the mutants corresponds with the fact that the residues targeted by mutagenesis are localized in the access tunnels rather than in the active site. The inhibition of wild-type DhaA and three mutants, 21, 27, and 31, by DCL was studied at a single concentration of TCP and several inhibitor concentrations. The dehalogenase activity of the DhaA variants decreased with increasing concentration of DCL. All tested mutants were inhibited at millimolar concentrations of DCL (K_i ranging from 2.08 to 4.42 mM) similarly to the wild-type enzyme (K_i = 2.50 mM).

Far-UV CD spectra of wild-type DhaA and its mutants exhibited one positive peak at 195 nm and two negative features at 208 and 222 nm, characteristic of α-helical content. Mutants 04, 14, 15, and 31 showed a similar intensity of their CD spectra to wild-type DhaA, confirming that the secondary structure of these enzymes was not significantly affected by the introduced mutations. Mutants 21, 27, 51, and 52 exhibited changes in the intensity of the measured CD spectra compared to wild-type DhaA, which indicated that the inserted mutations had an effect on the specific packing of the residues

Table 1. Mutants and their kinetic parameters for TCP conversion

	Variable residues									
	Slot tunnel			Main tunnel			Slot/main tunnel			
DhaA	135	141	246	145	172	176	245	K_m[a] (mM)	k_{cat}[a] (s^{-1})	k_{cat}/K_m (s^{-1} M^{-1})
WT	I	W	L	A	A	C	V	1.0 (±0.2)	0.04 (±0.01)	40
04	I	W	L	A	A	Y	V	1.7 (±0.1)	0.24 (±0.01)	141
14	F	W	L	A	A	C	V	1.5 (±0.3)	0.05 (±0.02)	33
15	F	W	L	A	A	Y	V	1.8 (±0.2)	0.23 (±0.01)	128
21[b]	L	F	I	A	A	Y	F	1.2 (±0.2)	0.55 (±0.04)	458
27[b]	V	W	I	A	A	Y	F	1.1 (±0.1)	1.02 (±0.06)	927
31[b]	F	W	I	A	A	Y	F	1.2 (±0.1)	1.26 (±0.07)	1050
51	F	W	I	A	F	Y	F	7.1 (±1.1)	0.21 (±0.02)	30
52	F	W	I	F	F	Y	F	1.7 (±0.2)	0.11 (±0.01)	65

[a] Standard deviation is given in parentheses.
[b] From Pavlova *et al.*[39]

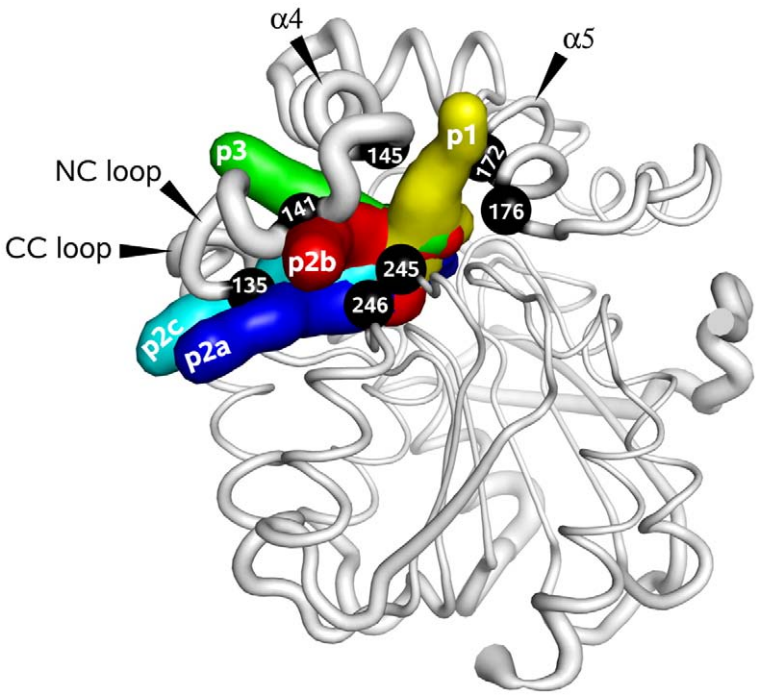


Fig. 1. Spatial location of pathways identified in the wild-type and eight mutants of DhaA from *R. rhodochrous* NCIMB13064. Pathways are represented by the surface and mapped on the crystal structure of DhaA from *Rhodococcus* sp. (PDB ID: 1CQW). The thickness of the ribbon corresponds to the crystallographic *B*-factors. Mutated residues are represented by black balls and labeled by the residue identifier. The color coding of the individual pathways is the same in all figures. p1, the main tunnel; p2a, the slot tunnel. All molecular graphics were created using PyMOL (DeLano Scientific, San Francisco, CA).

in the secondary structure of these enzymes, especially on the overall number of residues contributing to the α-helical content. Although there was an obvious difference in the intensity of the CD spectra among the DhaA variants, the absence of a significant change in the shape of the CD spectra of wild-type DhaA and the mutants suggested that the substitutions did not have a deleterious effect on the folding of the mutants.

Identification of pathways by MD simulations

The structures of the enzyme–product complexes for dehalogenation of TCP were modeled by molecular docking of *R*- and *S*-DCL, respectively, into wild-type DhaA complexed with CL. Both *R*- and *S*-DCL had similar conformations and positions in the active site with respective docked energies of -5.26 and -5.41 kcal mol^{-1} and both formed a hydrogen bond to the D106 side chain and a close contact to the imidazole ring of H272 in the active site. None of the mutated residues in the eight mutants studied overlapped with the binding site of DCL. Therefore, the same initial DCL coordinates were used in all mutants. The complexes were subjected to MD simulations to adjust the conformation of DCL in the active site and the conformations of the mutated residues. During these simulations, DCL preserved its docked orientation, characterized by the strong interactions with the catalytic residues D106 and H272. The overall stability of the complexes was supported by a flattened RMSD around 1.5 Å and a radius of gyration around 17.8 Å for all complexes. The equilibration MD trajectories were investigated for spontaneous release of the products (DCL and CL) and for exchange of water molecules between the buried active site and bulk solvent. In total, one release pathway for CL (p1), no release pathway for DCL, and five pathways for water molecules (p1, p2a, p2b, p2c, and p3) were observed for the DhaA variants. Another set of MD simulations was performed with CL replaced by a water molecule to model the system after release of halide from the active site. The overall stability of the complexes was supported by a flattened RMSD around 1.4 Å and a radius of gyration around 17.8 Å for all complexes. The MD trajectories were investigated for exchange of water molecules and spontaneous release of DCL. Altogether, three pathways were observed for water molecules (p1, p2a, and p2b) with the DhaA variants. No release of DCL was observed in any of the equilibration MD simulations, justifying the use of RAMD simulations to enhance DCL release. RAMD simulations of DCL release in wild-type DhaA and its mutants resulted in five pathways (p1, p2a, p2b, p2c, and p3). All existing pathways and mechanisms for exchange of CL, DCL, and water molecules between the active site and bulk solvent in wild-type DhaA and its mutants are summarized in Table 2 and in Figs. 2 and 3.

Pathway p1

Description of p1

p1 corresponds to the main tunnel identified in the crystal structures of DhaA. p1 is defined mainly by two helices of the cap domain α4 (F144, A/F145, and F149) and α5 (A/F172 and C/Y176) and, to a smaller extent, by the β8/α10 loop (H272) and the β8/α9 loop (V/F245). The main tunnel was permanently open in wild-type DhaA and mutant 14. The tunnel was widened during the passage of DCL, CL, and

Table 2. Occurrence of mechanisms of ligand exchange between the buried active site and bulk solvent in the wild-type (WT) and eight mutants (numbered) of DhaA during classical MD simulations and during RAMD simulations with a random force applied to DCL

Ligand	Pathway	Permanent tunnel	Transient tunnel	Protein matrix
CL	p1	WT	15	—
	p2a	—	—	—
	p2b	—	—	—
	p2c	—	—	—
	p3	—	—	—
DCL	p1	WT, 14	04, 15, 21, 27, 31, 51, 52	—
	p2a	—	WT, 04	—
	p2b	—	27	—
	p2c	—	21	—
	p3	—	WT, 27	—
Water	p1	WT, 14	04, 15, 21, 27, 31, 52	—
	p2a	—	04, 27	—
	p2b	—	WT, 04, 14, 15, 27, 31, 51, 52	—
	p2c	—	04	—
	p3	—	—	21

water molecules. The introduction of one or more aromatic substitutions in p1 caused a change in the opening mechanism from permanent tunnel to the transient tunnel (Fig. 4). Decreased accessibility of the tunnel for the studied ligands was observable already after a single substitution, C176Y in mutant 04, and was further pronounced by cumulatively introduced substitutions V245F, A172F, and A145F in mutants 31, 51, and 52 (Fig. 5d). The accessibility of p1 was controlled by the aromatic residues F144 and F149 in wild-type DhaA; by C176Y, V245F, A172F, and A145F in mutants 04, 15, 21, 27, 31, 51, and 52; by the varying length of the N-terminal part of α4 helix (143-EFA/F-145); by a flexible residue located in the middle of the α5 helix (G171); and by the varying size of the C-terminal part of the α5 helix (C/Y176).

Release of CL through p1

The release of CL was observed only through p1 and only in wild-type DhaA and mutant 15 (Fig. 5a). No release of the CL was observed in the classical MD and RAMD (data not shown) simulations of mutants 04, 14, 21, 27, 31, 51, and 52. The release process in wild-type DhaA started with rapid hydration of the active site by crystallographic water molecules located in p2a, followed by water molecules from bulk water entering the active site through p1 and p2b. The position of the two crystallographic water molecules in p2a was occupied by I135F in mutant 15. Therefore, only p1 and p2b were used by bulk water molecules to access the active site of mutant 15. The water molecules entered the active site and competed with CL for hydrogen-bonding interactions with the halide-stabilizing residues. Eventually, they enabled release of CL by making a hydration shell composed of three to five water molecules that accompanied the CL through the broadly widened p1. The entire release event took about 200 ps counting from the initial destabilization of the CL from the halide-stabilizing residues in both wild-type DhaA and mutant 15. The vacant binding site between the halide-stabilizing residues became occupied by a water molecule. Other water molecules in the active site competed with DCL for a hydrogen-bonding interaction with D106. DCL retained the hydrogen bond to D106 during the entire simulation in mutant 15, whereas in wild-type DhaA, this interaction was broken and DCL interacted with D106 indirectly through a water bridge.

Release of DCL through p1

p1 was the most frequently used pathway for the release of DCL in wild-type DhaA and all the mutants (Fig. 5b). The release process was strongly affected by the hydrogen bond between DCL and the nucleophile D106. This interaction had to be broken by water molecules to enable release of DCL from the active site and entry to the main tunnel. DCL was further attracted to the main tunnel by favorable electrostatic and van der Waals interactions with the catalytic base H272 and occasional hydrogen-bonding interactions with its $N^{\varepsilon 2}$. When the hydrogen bond to H272 was formed, DCL used exclusively pathway p1 for the release. DCL could

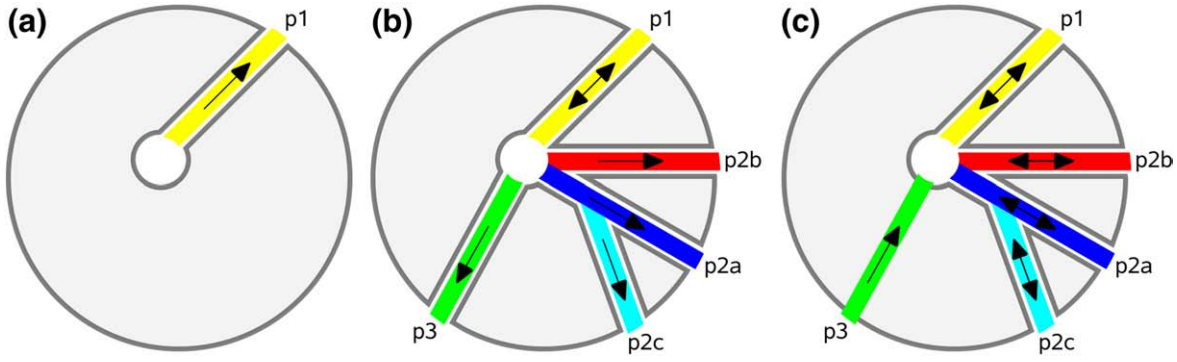


Fig. 2. Schematic representation of pathways for CL (a), DCL (b), and water molecules (c) in the wild-type and eight mutants of DhaA. Arrows indicate direction of passage of ligands through a tunnel or protein matrix (see Fig. 3). p1, the main tunnel; p2a, the slot tunnel.

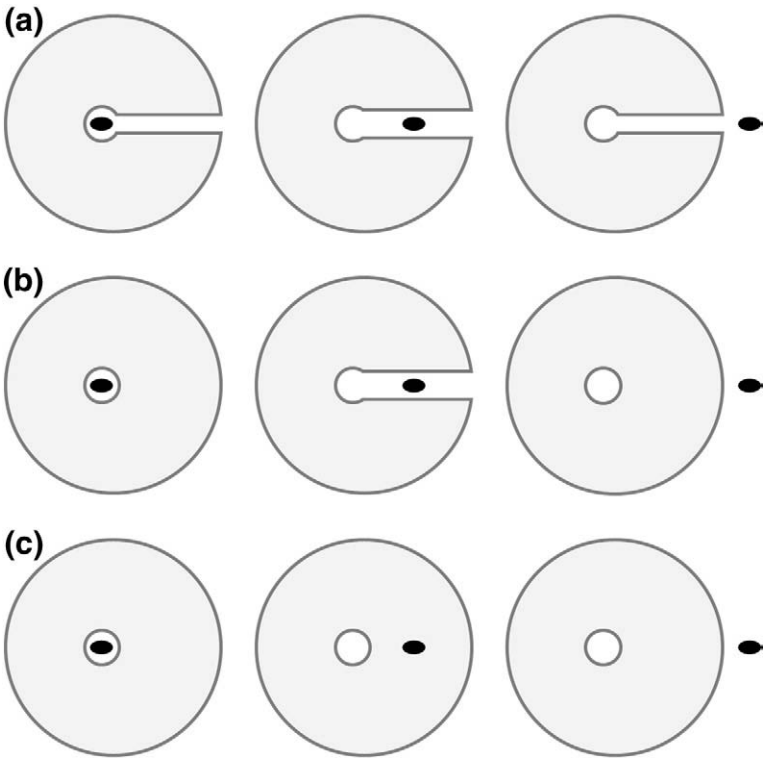


Fig. 3. Schematic representation for three mechanisms of ligand exchange between the buried active site of DhaA and bulk solvent: passage through a permanent tunnel (a), passage through a transient tunnel (b), passage through a protein matrix (c).

also form occasional hydrogen bonds to the thiol group of C176, the hydroxy group of C176Y, and the backbone carbonyl of A172. During the passage, DCL induced opening of p1 by promoting conformational changes in aromatic side chains lining p1. In wild-type DhaA, F144 was occasionally pushed to the bulk solvent to allow release of DCL out of the protein structure. The introduction of bulky aromatic substitutions C176Y and V245F significantly decreased the frequency of release of DCL through p1. The frequency was again increased by introducing additional aromatic substitutions A172F and A145F. The release of DCL from the variants carrying mutations C176Y and V245F required the simultaneous flip of both the aromatic residues followed by a conformational change of F144 to form an open

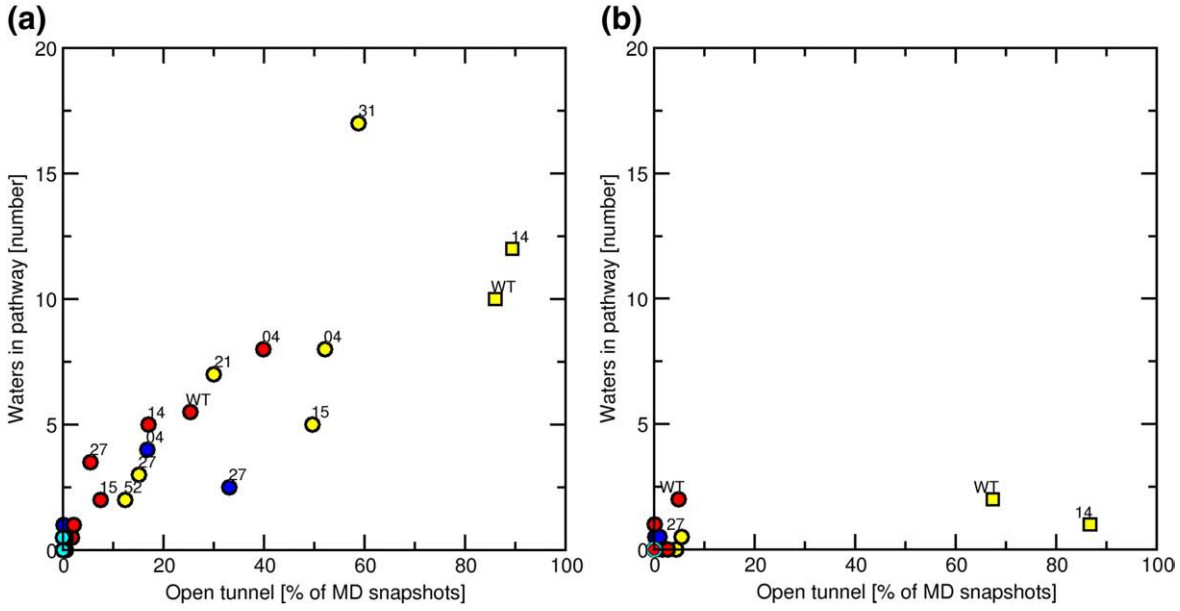


Fig. 4. The effect of mutations on the opening and solvent accessibility of permanent (squares) and transient tunnels (circles) in the wild-type and eight mutants of DhaA with CL (a) and without CL (b) in the active site (see Table 2). Each value represents the average from two independent simulations. Color coding of the pathways: p1, yellow; p2a, blue; p2b, red; p2c, cyan; p3, green.

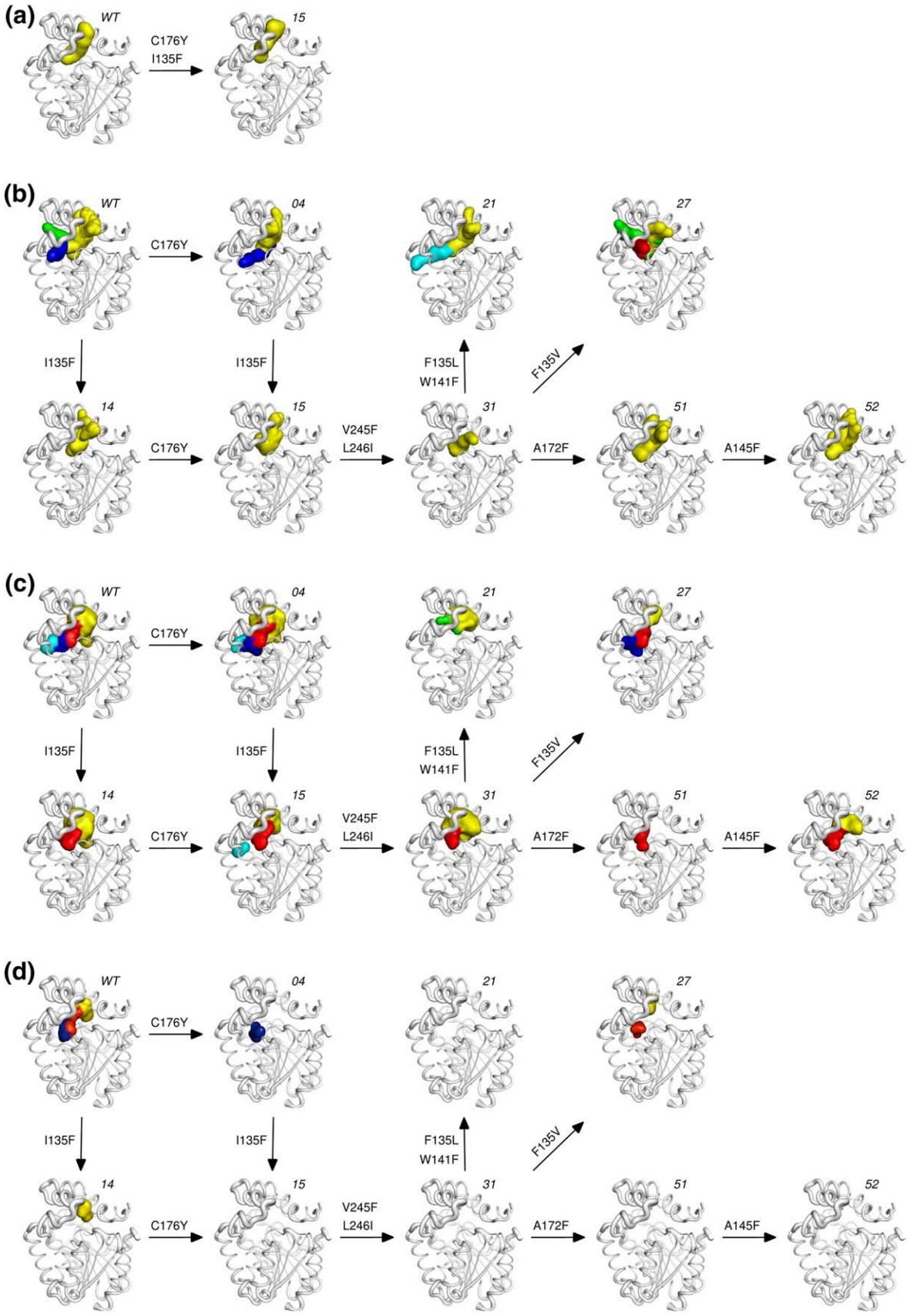


Fig. 5. Effect of mutations on the accessibility of the pathways for CL (a), DCL (b), water molecules with CL in the active site (c), and water molecules without CL in the active site (d). The pathways are depicted by a surface model of selected atoms of ligands in MD or RAMD simulations (see Methods) and mapped onto the crystal structure of DhaA from *Rhodococcus* sp. (PDB ID: 1CQW). DhaA is shown in ribbon; the pathways are color coded: p1, yellow; p2a, blue; p2b, red; p2c, cyan; p3, green.

gate. The release of DCL through p1 in mutant 51 required successive conformational changes in the order A172F, V245F, and C176Y. Additionally, A145F in mutant 52 caused hindrance of DCL release through p1. This steric hindrance was characterized by repeated entry of DCL into the main tunnel and its return back to the active site during the simulations.

Exchange of water molecules through p1

p1 was used by water molecules in all variants, except mutant 51, in the presence of CL in the active site (Fig. 5c). The accessibility of p1 for water molecules was much lower in the absence of CL in the active site (Fig. 5d); aromatic residues introduced into the main tunnel of mutants 04, 15, 21, 27, 31, 51, and 52 prevented the exchange of water molecules through p1 within the 2-ns time scale of the MD simulations.

Pathway p2a

Description of p2a

p2a is equivalent to the slot tunnel observed in the crystal structure 1BN6 of wild-type DhaA. p2a is defined by the β7/α9 loop (L246) and the β6/α4 loop [or N-terminal cap domain loop (NC loop); R133, I135, and W141]. p2a formed an open tunnel only during passage of DCL or water through p2a in the proteins with I135 (wild-type and mutant 04) or V135 (mutant 27). The accessibility of p2a was clearly controlled by residue 135 and by flexibility of the NC loop, the most mobile element of the DhaA structure according to the crystallographic and MD *B*-factors.

Release of DCL through p2a

p2a was accessible for DCL only in wild-type DhaA and mutant 04. The release of DCL through p2a was initiated by the breakage of a hydrogen bond between DCL and D106, immediately followed by formation of a DCL–water–D106 interaction with DCL located in the entry to the slot tunnel. The breakage of the water-bridge interaction resulted in fast passage of DCL towards R133 located at the outer opening of p2a. R133 made strong electrostatic and van der Waals interactions with DCL and additionally formed an unstable salt link with E140 and a stable hydrogen bond with the carbonyl of L246. DCL could form a hydrogen bond with the side chain of R133 when the salt link is naturally broken. Breakage of the salt link was, however, not essential for release of DCL to the bulk solvent.

Exchange of water molecules through p2a

p2a was rarely used by water molecules, and exchange of the solvent occurred only in wild-type DhaA and mutant 04 in both the presence and absence of CL in the active site and in mutant 27 only in the presence of CL in the active site (Fig. 5c and d). A single substitution, I135F, was sufficient to block exchange of water molecules through p2a.

Pathway p2b

Description of p2b

p2b corresponds to the deep surface depression near the slot tunnel opening in the crystal structures of wild-type DhaA. p2b is formed by the β7/α9 loop (V245F and L246I) and the NC loop (R133, I/L/V/F135, and W/F141). p2b formed a tunnel during the passage of DCL or water molecules through the pathway. p2b could be opened simultaneously with p1 and p2a. The accessibility of p2b was controlled by the aromatic residues W/F141 and V245F as well as by the flexibility of the NC loop.

Release of DCL through p2b

Release of DCL through p2b was observed only in mutant 27 carrying the mutation I135V (Fig. 5b). The release process started by breakage of the hydrogen bond between DCL and D106, followed by a DCL–water–D106 interaction and a transient hydrogen bond between DCL and the backbone carbonyl groups of E130 and I132. DCL entered p2b by inducing a simultaneous flip of the V245F and W141 side chains. DCL moved between the two aromatic side chains and established a hydrogen bond to the R133 side chain and the V245F backbone without perturbation of the salt link between R133 and E140 and the hydrogen bond between R133 and L246I. This hydrogen-bonding network hindered the smooth release of DCL to the bulk water.

Exchange of water molecules through p2b

p2b was used by water molecules in all DhaA variants, except mutant 21, in the presence of CL in the active site (Fig. 5c). p2b was the only pathway for exchange of water molecules in mutant 51 and the preferred pathway in mutant 27. p2b was always preferred over p2a, p2c, and p3. The accessibility of p2b for water molecules was significantly decreased in the absence of the CL in the active site. p2b was not used by water molecules in any variant except mutant 27 in MD simulations without CL (Fig. 5d).

Pathway p2c

Description of p2c

p2c was identified as a branch of p2a, which was not indicated by any of the crystal structures of wild-type DhaA. p2c was formed by the NC loop (R133, I135, and W/F141) and the α7/α8 loop [or C-terminal cap domain loop (CC loop); P210 and P212]. p2c formed an open tunnel only in wild-type DhaA and mutant 21 and only in the presence of CL in the active site. The bulkier DCL promoted opening of p2c for a longer time during its release in RAMD simulations. The accessibility of p2c was controlled

by the flexibility of the NC and CC loops and by a β-bridge interaction between P210 and A212 of the CC loop and I135 of the NC loop. The β-bridge must be disrupted to allow exchange of a ligand between bulk solvent and p2c.

Release of DCL through p2c

Release of DCL through p2c was observed only in mutant 21 carrying the unique substitution W141F (Fig. 5b). The release process started by obligatory breakage of the hydrogen bond between DCL and D106, followed by formation of the water bridge DCL–water–D106 positioning DCL near the entrance to p2a and p2c. After breakage of the water bridge, DCL moved along W141F, inducing a flip of its side chain, disrupted the A212–I135L β-bridge interaction, and established a hydrogen bond with the backbone carbonyl group of R133. The arginine showed strong electrostatic and van der Waals attraction for DCL. While maintaining the hydrogen bond to R133, DCL reoriented its carbon chain towards the β-bridge. After release of DCL from R133, DCL moved between the NC loop and the CC loop, causing breakage of the A212–I135L β-bridge interaction followed by release of DCL to bulk water and the immediate reconstruction of the β-bridge. The loss of the hydrogen-bonding interactions between the NC loop and the CC loop during DCL release through p2c was partially compensated for by a hydrogen bond between the residues of the β-bridge and DCL.

Exchange of water molecules through p2c

p2c was rarely used for exchange of water molecules and occurred only in wild-type DhaA and mutant 04 and only in the presence of CL in the active site (Fig. 5c and d). The accessibility of the active site for water molecules through p2c was blocked by I135F, but the β-bridge could be disrupted by water molecules from bulk solvent as observed for the mutant in the presence of CL in the active site.

Pathway p3

Description of p3

p3 was not observable in any crystal structure of wild-type DhaA. p3 was located between helix α4 (F149), the NC loop (W138 and W141), and the CC loop (L209 and I211). p3 was relatively short and straight compared to p2a, p2b, and p2c but was generally unfavorable for accommodation of water molecule or DCL due to steric clashes with the protein structure. p3 formed an open tunnel only upon passage of DCL through the pathway, whereas water migrated through the protein matrix. The open tunnel existed for the shortest period of time (on the order of picoseconds) among the pathways before its disappearance. Water in p3 moved solely through the protein matrix without formation of a tunnel (Fig. 3).

Release of DCL through p3

The release of DCL through p3 occurred in wild-type DhaA and in mutant 27 (Fig. 5b). After obligatory release of DCL from the nucleophile D106, a DCL–water–D106 interaction was established, and DCL was positioned between W141 and F149, whose side chains acted as a gate and flipped to open up the p3 pathway. The immediate release of DCL to the bulk solvent was blocked by the side chain of W138, and DCL was temporarily enclosed in an isolated cavity. DCL eventually pushed W138 into the bulk water and was released out of the protein. After DCL release, W138 made a fast flip to its original conformation to completely block p3.

Exchange of water molecules through p3

p3 served as the exchange pathway for one water molecule in mutant 21 carrying the mutation W141F and CL in the active site (Fig. 5c).

Structural characterization of three mutants with modified tunnels

Protein X-ray crystallography was used to determine the structures of mutants 04, 14, and 15 of DhaA from *R. rhodochrous* NCIMB13064 to atomic resolution (A.S. *et al.*, unpublished results). The mutant structures were compared with the structures of DhaA from *Rhodococcus* sp. available in the structural database.[27] This analysis deciphered the effect of the substitutions located in the main tunnel (mutant 04), the slot tunnel (mutant 14), and both the main and slot tunnels (mutant 15) on the accessibility of the active site and the mechanisms of ligand exchange (Fig. 6).

The main tunnel was open in all crystal structures of DhaA variants carrying the wild-type C176, while the single C176Y substitution blocked this tunnel. This observation was in agreement with MD simulations showing the main tunnel to be mostly closed in mutants 04 and 15, unless it was temporarily opened by a passing ligand. The side chain of Y176 was resolved in two different conformations in the crystal structure of mutant 04. The distance between oxygen atoms of 4-hydroxyphenyl groups of the two conformations is 4.1 Å, and there are most likely two distinct conformational states of this bulky residue at the mobile C-terminus of the α5 helix. Accommodation of the 4-hydroxyphenyl group of Y176 in the place formerly occupied by the thiol group of C176 results in one conformation pointing towards the active site and another placing the aromatic ring close to the carbonyl of A172. This observation is consistent with the gatekeeping function of Y176 and the observation of both conformations of Y176 in the MD simulations.

The opening of the slot tunnel in the structure 1BN6 was due to the presence of different rotamers of I135 and R133 side chains together with 0.8 Å displacement of the backbone of five residues of the NC loop (133-RPIPT-137), compared to structures 1CQW and 1BN7. This is in accordance with the

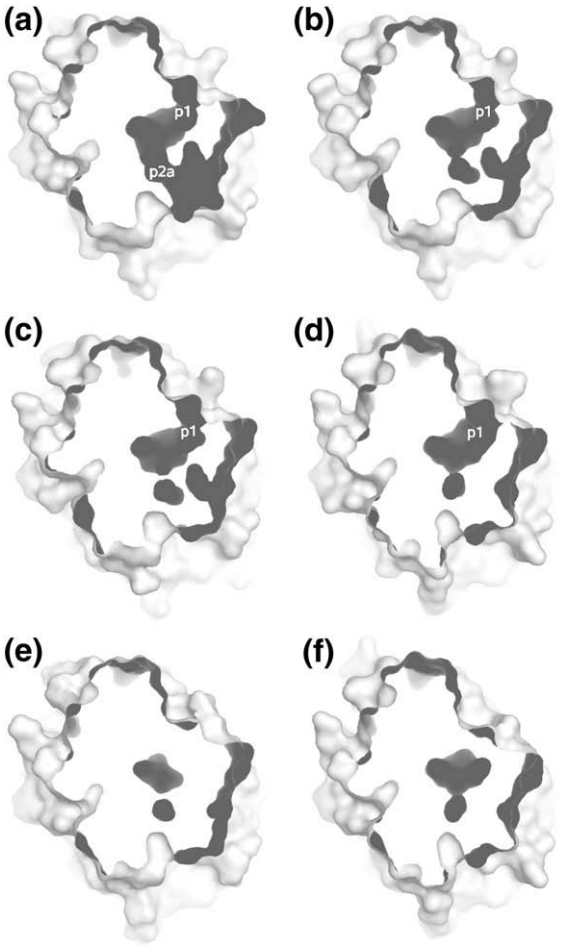


Fig. 6. The crystal structures of DhaA and its mutants ordered by the accessibility of their active sites *via* the main tunnel and the slot tunnel. DhaA from *Rhodococcus* sp. (PDB ID: 1BN6) with the main tunnel and the slot tunnel open (a), DhaA from *Rhodococcus* sp. (PDB ID: 1CQW and 1BN7) and mutant 14 of DhaA from *R. rhodochrous* NCIMB13064 (PDB ID: 3G9X) with the main tunnel open and the slot tunnel closed (b–d), mutants 04 and 15 of DhaA from *R. rhodochrous* NCIMB13064 (PDB ID: 3FBW and 3FWH) with both the main tunnel and the slot tunnel closed (e and f). The protein structures are visualized as a slice through the surface representation with tunnels and cavities colored in dark gray. p1, the main tunnel; p2a, the slot tunnel.

important role of the highly mobile NC loop in controlling the accessibility of the slot tunnel in RAMD simulations of DCL release through the p2a pathway. The displacement of the NC loop in the structure 1BN6 could be due to the presence of an unknown ligand represented by an extensive electron density in the active site and five water molecules located at the mouth of the slot tunnel.[27] On the contrary, the active site of 1CQW contains only water molecules and 1BN7 contains an acetate ion. The slot tunnel in structures 1CQW and 1BN7 was represented only by an isolated cavity containing two water molecules, while no crystallographic water molecules were identified in the region corresponding to the mouth of the slot tunnel.

The NC loop of mutant 04 was resolved in a conformation similar to that in structures 1CQW and 1BN7, resulting in a closed slot tunnel. The introduction of a bulky I135F substitution further reduced accessibility of the slot tunnel of the crystal structures of mutants 14 and 15. The NC loop of mutants 14 and 15 aligned with the corresponding region of 1BN6, suggesting slight displacement of the loop upon introduction of the bulky aromatic side chain. This was in agreement with MD simulations that showed I135 to switch between four different conformations, whereas F135 adopted a single conformation similar to those observed in the crystal structures of mutants 04 and 15. Based on available crystallographic data, we conclude that the slot tunnel belongs to a transient type of tunnel in all the studied proteins, while the main tunnel changes from being a permanent tunnel in the proteins with the wild-type cysteine in position 176 to being a transient tunnel in mutants 04 and 15 carrying tyrosine at position 176.

Discussion

MD simulations identified five pathways for product release and water exchange

Two distinct tunnels, named the main tunnel and the slot tunnel, could be identified in the crystal structures of the haloalkane dehalogenases.[34,35] The simulations conducted with wild-type DhaA confirm the relevance of the main tunnel (pathway p1) and the slot tunnel (pathway 2a) for release of products and exchange of water molecules between the buried active site and the bulk solvent. p1 is observed as the only release pathway for CL. The release of CL was observed only in wild-type DhaA and mutant 15, most probably due to limited time available in the classical MD simulations for the CL to become hydrated. Nevertheless, the proposed role of p1 for CL release is supported by the presence of iodide anions in the wild-type X-ray crystal structure 1CQW.[27] The structure reveals two iodide binding sites: (i) iodide anion positioned between the halide-stabilizing residues N41 and W107 and (ii) iodine covalently attached to the S^{γ} of C176.[27] An imaginary line connecting the two iodide binding sites goes through the p1 pathway. p1 is also the dominant release pathway for DCL and for exchange of water molecules between the active site and bulk solvent. p2a functions as an auxiliary pathway for DCL and water molecules. The release of DCL through the pathway corresponding to p2a was also observed in the recent MD simulations of enzyme–product complexes of the phylogenetically closely related haloalkane dehalogenase LinB.[41]

Besides p1 and p2a, the simulations revealed the additional p2b and p2c for water molecules and p3 for DCL in wild-type DhaA. While p2b can be related to a deep surface depression in the three available crystal structures of DhaA,[27] no indication for p2c or p3 can be found in these structures. DCL passing

through p3 forms a transient tunnel, while water molecules pass directly through the protein matrix of the region p3. MD simulations suggest that p2b and p2c function as auxiliary pathways for water molecules, especially in the presence of charged CL in the active site. Interestingly, p2b is the preferred route for the exchange of water molecules over p2a and is observed in all but one DhaA variant.

It is noteworthy that all the pathways are located along highly mobile secondary-structure elements of the DhaA cap domain comprising the NC loop, the variable N-terminal part of the α4 helix, the breakage point in the α5 helix at G171, and the variable C-terminal part of the α5 helix and the CC loop.[34] The larger backbone fluctuations during the passage of CL, DCL, or water molecules through a pathway compared to the free, unliganded pathway also suggest that the mobility can be further enhanced by interactions with small molecules. Apparently, some of the functionally relevant pathways are not observable in the crystal structures and MD simulations are necessary for their identification.[14–17,36,37] Previous studies on other systems have also shown that flexibility of loops and helices controls accessibility of the active site, for example, in cytochrome P450,[15,17,18,42–50] acetylcholinesterase,[4] gpH1 receptor,[49] and haloalkane dehalogenase LinB.[41] We further demonstrate that all relevant pathways in wild-type DhaA and its mutants can be identified by monitoring of the exchange of solvent between the buried active site and the protein surface.

Solvation of products is essential for their release from buried active site

Specific interactions between the two products, CL and DCL, the protein, and water solvent are essential for release of products from the buried active sites of haloalkane dehalogenases. CL formed during the dehalogenation reaction is strongly bound in between two halide-stabilizing residues, which are present in all currently known haloalkane dehalogenases.[26–29,33,50–55] The halide ion is positioned in between the two halide-stabilizing residues in all crystal structures of haloalkane dehalogenases, with the exception of enzyme–substrate complexes with the halide-binding site occupied by the substrate molecule.[27,33,54] The stabilization of halide ions is weaker in DhaA than in DhlA due to the different chemistry of the halide-stabilizing residues. DhaA possesses tryptophan and asparagine, whereas DhlA has two tryptophans.[53,55] The stronger stabilization of the halide product in DhlA, together with its occluded active site and different location of a catalytic acid, may explain the limitation of the reaction cycle by the halide release.[2,30] In DhaA, halide release is a fast process, showing no effect on overall kinetics.[38] The release of CL from DhaA is clearly triggered by water molecules. The CL bound in the active site induces a strong electrostatic field, which attracts water molecules from the bulk solvent to the active site not only through tunnels p1 and p2a but also through three auxiliary water pathways p2b, p2c, and p3. Water molecules compete with the halide-stabilizing residues for favorable interactions with CL and facilitate the release of CL from the active site. This proposal is in very good agreement with the classical MD simulations of product release from the haloalkane dehalogenase LinB.[41] In DhaA, CL leaves the active site through p1 surrounded by water molecules. The polar residue K175 located at the mouth opening of the main tunnel of DhaA guides solvated CL out of the main tunnel to the surrounding solvent.

Both molecular docking and MD simulations suggest that the DCL, formed during the dehalogenation reaction, makes a hydrogen bond with the nucleophile D106. This interaction could explain inhibition of TCP dehalogenation by DCL product. Disruption of this interaction is assisted by water molecules and represents one of the limiting events of DCL release. Several polar residues located along the release pathways make contacts with DCL during its release. Very important is van der Waals and electrostatic attraction of DCL by H272 in p1. Strong interaction of the alcohol product with the conserved catalytic histidine has been described for the enzyme–DCL complex of LinB[56] as well as for other enzyme–product complexes of this enzyme.[54,57–59] Another important interaction made by DCL, when moving away from the active site *via* p2a, p2b, and p2c, is with the polar R133. R133 controls p2a opening by transient formation of a salt link to E140. Favorable electrostatic and van der Waals interactions are provided by several aromatic side chains in p1, p2b, and p3. Bulky aromatic side chains function as gatekeepers and must undergo conformational change to increase the accessibility of the pathway for bulky DCL. If two or more aromatic side chains form the gate, they change conformation in a simultaneous or consecutive way, depending on their location along the pathway. DCL moves through a pathway taking advantage of temporarily increased local space due to natural protein breathing motions, which may be further enhanced by a DCL molecule. This is in agreement with studies describing the accessibility of pathways that are being controlled locally by (i) hydrogen-bonding and salt link interactions, described for cytochrome P450[14,16,43–45,59–62] and acetylcholinesterase,[63,64] and (ii) aromatic gating, described for cytochrome P450,[14,16,42,43] acetylcholinesterase,[4,64–67] NADH oxidase,[19] horseradish peroxidase,[68,69] and myoglobin.[70]

Mutations change the accessibility of the pathways

The eight mutants of the DhaA haloalkane dehalogenase studied here carry various substitutions in the residues lining the main tunnel and the slot tunnel. The mutations show diverse effects on the accessibility of individual pathways for individual ligands. Four out of five possible pathways (p1, p2a, p2b, and p2c) are accessible for water molecules in wild-type DhaA, whereas one pathway (p3) becomes accessible for water molecules only after

introducing the W141F substitution in mutant 21. Residue 141, together with F149, forms a gate that controls accessibility of p3. The release of DCL through p3 is observed in mutant 27, but not in mutant 21, suggesting that p3 is hardly accessible for bulky alcohol regardless of the residue in position 141. On the other hand, W141F and I135V are required to allow accessibility of p2b and p2c pathways for DCL. I135 is the key residue controlling the access to p2a, p2b, and p2c because its substitution to bulky phenylalanine effectively closes up p2a, p2b, and p2c for DCL and p2a and p2c for water molecules.

The only pathway that could not be completely blocked is p1. This is the main release pathway for DCL and is accessible even after the cumulative introduction of four aromatic substitutions. Introduction of aromatic substitutions in p1 (C176Y, V245F, A172F, and A145F) seems to decrease its accessibility for DCL in the presence of two substitutions but surprisingly seems to restore the accessibility for DCL by the third and fourth substitutions in mutants 51 and 52.

It is noteworthy that many aromatic residues are packed close to each other in the active site (H272, Y273, F152, and F168) and p1 (F144, A145F, F149, A172F, C176Y, and V245F) of mutant 52. We propose that an effective gating is established because product release is not impaired in mutant 52 and the water molecules can enter the occluded active site once CL is formed during the dehalogenation reaction. Such an aromatic gating is a common way by which enzymes with buried active sites control accessibility during the reaction cycle. An interesting example of aromatic gating has been described for acetylcholinesterase.[4,63,65,71] The gate of the main gorge of acetylcholinesterase is formed by four aromatic residues, and their pinching movement,[4,71] together with loop motion[63] and strong dipole moment,[65] is responsible for the enzyme operating near the diffusion-limited rate.

Mutations change the mechanism of ligand exchange

Three different mechanisms for exchange of products and water solvent were observed in the haloalkane dehalogenase DhaA and its mutants: (i) passage through a permanent tunnel, (ii) passage through a transient tunnel, and (iii) passage through a protein matrix. These mechanisms have analogies in other proteins. The exchange of ligands through permanent tunnels has been described for numerous proteins possessing a crystallographically observable tunnel in their structure.[3,4,8,10,17,18,20,34,41,46,72,73] The exchange through the transient tunnels corresponds to exchange through a so-called naturally fluctuating bottleneck,[74] which is a common mechanism to transiently enable access and egress of ligands in and out of the active site in the regions of lower density of protein atoms. This mechanism was previously reported for cytochrome P450,[42,44,47,59] acetylcholinesterase,[66,71] NADH oxidase,[19] T4 lysosyme,[75] and horseradish peroxidase.[69] Passage of the ligands through the protein matrix is a well-documented phenomenon for gas migration in heme proteins.[60,70,72,76,77]

To study the effect of mutations on the mechanism of ligand exchange, we attempted to assign one of three mechanisms to every ligand exchange observed in our molecular dynamic simulations. By comparing the mechanisms of the wild-type enzyme with its mutants, we demonstrated that substitutions introduced into the tunnels changed not only the accessibility of the individual pathways (described in the previous section) but also the mechanism of ligand exchange in the case of the p1 pathway. This pathway follows the permanent tunnel in the wild-type enzyme. The crystal structures of the mutants show that a single aromatic substitution in the tunnel results in its closure. However, MD reveals that ligands can pass through the p1 pathway even in the mutants with four aromatic substitutions. This is possible due to ligand-induced changes in the protein structure that cause the pathway to open up transiently to allow release of products. The opening of p1 can also be induced by water molecules entering the active site through p1 due to strong electrostatic attraction by CL. This solvation of the active site through the transiently opened tunnel is observed even in mutant 52 with the most occluded active site. The substitutions for aromatic residues in p1 clearly changed the mechanism for ligand exchange from the passage through a permanent tunnel to the passage through a transient tunnel.

Concluding remarks

We conclude from our study of the wild-type haloalkane dehalogenase DhaA from *R. rhodochrous* NCIMB13064 and its eight mutants that the ligands can be exchanged between the buried active site and the bulk solvent by five different pathways, denoted p1, p2a, p2b, p2c, and p3, and by three mechanisms, namely, passage through a permanent tunnel, passage through a transient tunnel, and passage through a protein matrix. Two out of the five pathways (p1 and p2a) are observable in the crystal structures, while all the three other pathways were identified by MD simulations.

The release of CL proceeds exclusively *via* pathway p1 and is accompanied by solvation of the negatively charged ion by water molecules, breakage of attractive interactions with the halide-stabilizing residues N41 and W107, and attraction by the polar K175, positioned at the tunnel opening. The release of DCL proceeds *via* all five pathways and requires the initial breakage of the hydrogen bonds between the product molecule, the nucleophile D106, and the catalytic histidine H272. Release of DCL *via* ligand-induced pathways is enabled by the high mobility of the protein backbone and by progressive rotations of the protein side chains leading to formation of transient tunnels. Point mutations systematically introduced into the p1 and p2a tunnels, identified

in the crystal structures, modulate the accessibility of the individual pathways and lead to a changed mechanism of ligand passage in p1. We propose that the accessibility and mechanisms of ligand passage in enzymes with buried active sites can be modulated by mutations introduced into the exchange pathways. These mutations may lead to pronounced effects on enzymatic activities and substrate specificities.

Materials and Methods

Mutagenesis and DNA sequencing

Established methods were employed for the preparation of plasmid DNA, digestion of plasmid and PCR-amplified DNA fragments with restriction endonucleases, ligation, agarose gel electrophoresis, and transformation of *Escherichia coli* cells.[78] The construction of the recombinant genes *dhaA04His*, *dhaA21His*, *dhaA27His*, and *dhaA31His* was described by Pavlova *et al.*[39] The mutant recombinant genes *dhaA14His* and *dhaA15His* were obtained using the QuikChange Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) according to the manufacturer's instructions. Plasmids pUC18::*dhaAHis* and pAQN::*dhaA04His* were used as templates.[79] The recombinant gene *dhaA14His* was afterwards recloned into the expression vector pAQN. The mutant genes *dhaA51His* and *dhaA52His* were constructed by site-directed mutagenesis using the principle of inverted PCR that was carried out according to the protocol provided with Phusion polymerase (Finnzymes, Espoo, Finland). pAQN::*dhaA31His* and pAQN::*dhaA51-His* were used as templates. The nucleotide sequences of all mutants were determined by the dideoxy chain termination method using an automated DNA sequencer, ABI PRISM 310 genetic analyzer (Applied Biosystems, Foster City, CA).

Protein expression and purification

The expression and purification of wild-type DhaA and mutants 04, 21, 27, and 31 were described by Pavlova *et al.*[39] The recombinant plasmids pAQN::*dhaA14His*, pAQN::*dhaA15His*, pAQN::*dhaA51His*, and pAQN::*dhaA52His* were transformed to *E. coli* BL21 cells. Fresh transformants were used to inoculate 2 L of Luria–Bertani medium (Sigma-Aldrich, St. Louis, MO) with ampicillin (100 μg/mL) and cultivated at 37 °C to an optical density of 0.5 at 600 nm. Protein expression was induced by addition of isopropyl-D-1-thiogalactopyranoside to a final concentration of 0.5 mM in Luria–Bertani medium. Cells were harvested by centrifugation at 8000*g* for 10 min after 4 h of cultivation at 30 °C. During harvesting, cells were washed and then resuspended in 20 mM KH_2PO_4 buffer (pH 7.5). Harvested cells were kept at −65 °C. Defrosted cell suspensions were disrupted by sonication with Soniprep 150 (Sanyo Gallenkamp, Loughborough, UK) or ultrasonic processor Hielscher UP200S (Hielscher Ultrasonics, Teltow, Germany), and the lysates were centrifuged at 21,000*g* for 1 h. The collected cell-free extracts were purified using FPLC Akta (Amersham Pharmacia Biotech, USA) and HiTrap Chelating column with affinity resin (Amersham Biosciences, Freiburg, Germany) charged with Ni^{2+} and equilibrated with purification buffer (pH 7.5) composed of 16.4 mM K_2HPO_4, 3.6 mM KH_2PO_4, and 0.5 M NaCl containing 10 mM imidazole. Unbound and weakly bound fractions were washed out with the purification buffer containing 50 mM imidazole. Histidine-tagged proteins were eluted with the purification buffer containing 300 mM imidazole. Purified proteins were dialyzed against 50 mM phosphate buffer (pH 7.5) composed of 41 mM K_2HPO_4 and 9 mM KH_2PO_4. Protein concentrations were determined by the method of Bradford (Sigma-Aldrich).

CD spectroscopy

CD spectra were recorded at room temperature (22 °C) using a spectrometer Jasco J-810 (Jasco, Tokyo, Japan). Data were collected from 190 to 260 nm at 100 nm/min, 1 s response time, and 2 nm bandwidth using a 0.1-cm quartz cuvette. Averages of 10 individual scans were corrected for absorbance caused by the buffer and expressed in terms of the mean residue ellipticity (Θ_{MRE}). Secondary-structure content was calculated from the CD spectra using K2D and Self-Consistent methods[80,81] implemented in the program DICROPROT.[82]

Enzyme kinetics

Steady-state kinetic constants K_m and k_{cat} for the conversion of TCP by wild-type DhaA and mutants 04, 14, 15, 21, 27, 31, 51, and 52 were assayed with TCP using the initial velocity measurements described previously.[7] The substrate concentration was assayed by a gas chromatography system equipped with a flame ionization detector Trace GC 2000 (Thermo Finnigan, San Jose, CA) and a DB-FFAP capillary column 30 m × 0.25 mm × 0.25 mm (J&W Scientific, Folsom, CA). The method described previously by Iwasaki *et al.* was used for determination of the product concentration.[83] The K_m and k_{cat} constants were calculated using the computer program Origin 6.1 (OriginLab, Northampton, MA).

Inhibition kinetics

Inhibition constants of wild-type DhaA and mutants 21, 27, and 31 for DCL were determined by monitoring the initial rates of TCP conversion at various DCL concentrations by the spectrophotometric method of Iwasaki. The substrate concentration was constant (1.4 mM), and the inhibitor concentrations varied between 1.4 and 35.0 mM. The velocity of the reaction without the inhibitor was measured as a negative control. Reactions were performed in duplicates. Halide concentrations were determined at several times (10, 20, 30, and 40 min) in order to obtain at least three data points in the initial phase of conversion. Initial substrate and inhibitor concentrations were determined before reaction initiation by gas chromatograph equipped with a flame ionization detector Trace GC 2000 (Thermo Finnigan) and a DB-FFAP capillary column 30 m × 0.25 mm × 0.25 mm (J&W Scientific). Steady-state inhibition constant K_i of DhaA variants for TCP conversion were determined by the initial rate of enzymatic activity using the program ORIGIN 6.1 (OriginLab).

Preparation of structures for molecular modeling

The crystal structure of 1CQW was truncated by five residues at the C-terminus and modified with three substitutions V172A, I209L, and A292G (numbered according to the *dhaA* gene, which differs from the numbering in the

structure by 11 amino acids) to mimic the structure of DhaA from *R. rhodochrous* NCIMB13064 (wild-type DhaA) used in experiments. Mutants 04, 14, 15, 21, 27, 31, 51, and 52 were prepared by the mutagenesis wizard of PyMOL 0.97,[84] selecting the most frequently occurring rotamer that had no steric overlap with the neighboring protein atoms. A three-dimensional model of DCL was built in PyMOL and geometry optimized by the AM1 method of MOPAC 2000 using the following keywords: SCFCRT = 1D-12, EF, GNORM = 0.0001, STEP = 15, POINTS = 12, LET, and PRECISE.[85] The optimized structure was refined by energy minimization using GAUSSIAN 94 employing the Hartree–Fock method and 6-31G* basis set.[86] Partial atomic charges were fitted to reproduce the electrostatic potential calculated with GAUSSIAN using the RESP module of AMBER 8.[87]

Molecular docking

Docking of DCL to wild-type DhaA, considering both *R*- and *S*-DCL, was performed using AUTODOCK 3.05.[88] Rotatable bonds were assigned to DCL using the Autotors module of AUTODOCK. All crystallographic water molecules were removed, the main tunnel iodide was removed, and the active-site iodide was replaced by a CL. Polar hydrogens were added to wild-type DhaA using WHATIF 5,[89] and solvation parameters were added using the Autogrid module of AUTODOCK. A grid box of $81 \times 81 \times 81$ points in x, y, and z dimensions was used with a grid spacing of 0.25 Å. The grid was centered on the C^{γ} of H272 to ensure that the entire active site, the open main tunnel, and the closed slot tunnel were encompassed by the box. The electrostatic map and atomic interaction maps for all atom types of DCL, that is, carbon, oxygen, chlorine, and hydrogen, were calculated using AUTOGRID. Fifty independent docking calculations were performed for *R*- and *S*-DCL using the AUTODOCK module of AUTODOCK using a Lamarckian genetic algorithm for global and a Solis&Wets algorithm for local search with an initial population size of 50 and default AUTODOCK 3.05 settings for elitism and crossover. A maximum of 27,000 generations or 1,500,000 energy evaluations were performed. The resulting conformations were clustered with a tolerance of 0.5 Å. The lowest-energy representations of the highest-populated clusters of *R*- and *S*-DCL in wild-type DhaA were selected as starting conformations for subsequent MD simulations of wild-type DhaA and its mutants.

Classical MD simulations

The AMBER94 force field[90] was used. CL was approximated by the IM atom type of the 1994 Cornell force field[90] and was assigned a charge of −1e. Alternatively, the CL was replaced by one water molecule to approximate the situation with the CL absent. All crystallographic water molecules not overlapping with the docked DCL were added to the complexes. Nonpolar hydrogens were added to the protein using the Leap module of AMBER 8. Seventeen sodium cations were added with Leap to ensure a neutral net charge of the system. Finally, the complexes were immersed in a rectangular box of TIP3P[91] water molecules with a minimum wall thickness of 10 Å using Leap and subjected to an equilibration protocol using the Sander module of AMBER. The equilibrations consisted of the following steps: (i) 300 steps of steepest descent energy minimization of all non-crystallographic atoms, that is, all hydrogens on protein, water, and DCL atoms; (ii) 20 ps periodic boundary condition MD of water, sodium cations, and DCL at constant temperature of 300 K (using the weak-coupling algorithm)[92] and constant pressure of 1 atm (with isotropic position scaling) with the rest of the system harmonically restrained with a 500 kcal mol^{-1} $Å^{-2}$ force constant; (iii) four consecutive steepest descent energy minimizations of 300 steps each with a decreasing restraint on the protein backbone with a force constant of 500, 125, 5, and 0 kcal mol^{-1} $Å^{-2}$ respectively; (iv) unrestrained MD using the same parameters as those for the 20 ps of MD but raising the temperature from 0 to 300 K during the initial 200 ps. The trajectories were propagated for 2.0 to 2.8 ns to ensure acquisition of well-equilibrated and stable systems. A time step of 2 fs was used with application of SHAKE algorithm[93] to bonds involving hydrogens and a particle mesh Ewald treatment of Coulombic interactions. The cutoff distance for the nonbonded interactions was 10 Å. Snapshots were gathered every 0.5 ps.

Random acceleration MD simulations

RAMD simulation[14] resembles classical MD simulation except that an additional force is applied to the center of mass of the ligand in a randomly chosen direction. After a user-defined number of time steps, the distance traveled by the ligand is compared to a threshold parameter. If the ligand does not reach the threshold distance, a new, randomly chosen direction is given to the force on its center of mass; otherwise, the force direction is maintained. The process is iterated until the ligand has been released into the bulk solvent. RAMD simulations were carried out using the AMBER 8 software package†.

RAMD was performed on DCL in the CL-free complexes for wild-type DhaA and its mutants. First, the proper setting of the RAMD parameters was tested on wild-type DhaA complexed with *R*- and *S*-DCL, respectively. A random acceleration of 0.25, 0.20, 0.15, 0.1, 0.09, 0.08, 0.07, 0.06, 0.05, 0.04, 0.03, 0.02, and 0.01 kcal $Å^{-1}$ g^{-1} applied to the center of mass of DCL; a number of time steps (10, 20, 40, and 80); and a threshold distance of 0.001, 0.002, 0.004, and 0.008 Å were tested. RAMD simulations were performed with various combinations of the values of the three parameters. Three different snapshots of the MD simulations from a well-equilibrated region were used as starting structures for RAMD simulations, resulting in six RAMD trajectories for each combination of the parameters considering *R*- and *S*-DCL together. The maximum duration of RAMD simulations was set to 1 ns. If DCL left the protein for the bulk water and the distance between center of mass of D106 and DCL exceeded 30 Å, the simulations were halted. RAMD simulations were selected for detailed analysis when the parameters used resulted in the release of *R*- or *S*-DCL within 1 ns in at least 4 out of 6 simulations and lasted for at least 20 ps. The final settings were 0.04 and 0.05 kcal $Å^{-1}$ g^{-1} for the random acceleration, 10 for the number of time steps, and 0.002 and 0.004 Å for the threshold distance, resulting in a total of 24 RAMD simulations of wild-type DhaA. The parameters derived from RAMD simulations of wild-type DhaA were adopted for the mutants. RAMD simulations of the mutants were also performed on *R*- and *S*-DCL, but only the final MD snapshot was used. Altogether, 62 RAMD trajectories of DhaA mutants were recorded. No difference in the preferential release through different pathways or in the mechanism of the release was obvious for the *R*- and *S*-

† The RAMD patch is available at http://projects.villa-bosch.de/mcm/software/amber

enantiomers of DCL. Therefore, *R*-DCL and *S*-DCL were further considered to provide variability in the MD trajectories only. RAMD simulations were also performed in the presence of CL (data not shown).

Analysis of MD simulations

The stability of the trajectories was assessed by RMSD and radius of gyration using the Ptraj and Carnal modules of AMBER. Stable parts of the MD simulations were decided visually by plotting RMSD *versus* time. The stability of the secondary elements was calculated in PyMOL for each snapshot of a trajectory using the DSSP method.[94] *B*-factors and geometrical parameters (distances, angles, and dihedrals) were measured using Ptraj. Hydrogen bonds were identified using <2.76 Å and >120° for the distance and angle thresholds, respectively. Distances between the center of mass of the two hydrogen atoms, $H^{\varepsilon 1}$ of W107 and $H^{\delta 22}$ of N41, and CL and between the center of mass of D106 and DCL were measured with Ptraj to identify regions of long residence time for CL or DCL during release from the active site to the bulk water. Residues within 7 Å of the center of mass of DCL along the trajectories were assumed to constitute the release pathways. Seven angstroms ensured that all first shell residues were included. Water molecules that resided in the protein interior or entered the protein during MD and RAMD simulations were identified as those located at a distance less than 8 Å from any atom of D106 in at least one snapshot using Carnal. The selected pathway residues and the internal water molecules aided analysis of important events using VMD 1.8.5[95] and PyMOL where especially behavior of the pathway residues and the internal water molecules were monitored. Release pathways for CL in MD and DCL in RAMD simulations were visualized by PyMOL as a surface representation of CL or all positions of a central carbon of DCL. The surface representations of all release pathways for CL and DCL from all RAMD simulations were superimposed onto the crystal structures of DhaA (PDB codes 1CQW and 1BN6) and clustered by visual inspection according to overlap between the surface representations of the pathways and the contribution of the same secondary-structure elements. Clusters were annotated by a number, and branching clusters were further distinguished by a letter. Water pathways were identified in MD simulations by superimposing all positions of oxygen atoms of all internal water molecules onto the crystal structures of DhaA in PyMOL. Opening and closing of tunnels connecting the active site to the bulk solvent and their inducibility by CL, DCL, and water were visualized by PyMOL as a slice through the solvent-accessible surface representation (with a probe radius of 1.4 Å) of the protein with 5- and 1-ps windows for MD and RAMD simulations, respectively. Tunnels were classified as closed or open according to whether the solvent-accessible surface shows the active site isolated from (closed) or connected to (open) the bulk solvent.

Interaction energies

All snapshots were extracted from the selected RAMD trajectories. All water molecules were removed and the structures were adjusted by 150 steps of steepest descent energy minimization with the implicit generalized Born solvation model II[96] using Sander. The van der Waals and electrostatic interaction energies between DCL and every protein residue were calculated for each of the energy-minimized structures using the Anal module of AMBER.

Accession numbers

Coordinates and structure factors of the X-ray crystal structures of the haloalkane dehalogenase DhaA mutants have been deposited in the PDB with accession numbers 3FBW (mutant 04; C176Y), 3G9X (mutant 14; I135F), and 3FWH (mutant 15; I135F+C176Y).

Acknowledgements

We acknowledge financial support from the Ministry of Education of the Czech Republic (LC06010 to M.P., MSM0021622412 to Martin Klvana, and MSM0021622413 to Z.P.), the Grant Agency of the Czech Republic (201/07/0927 to Jiri Damborsky and 203/08/0114 to R.C.), the Grant Agency of the Czech Academy of Sciences (IAA401630901 to Jiri Damborsky), the North Atlantic Treaty Organization (EST.CLG.980504—NATO Linkage Grant to R.W. and Jiri Damborsky), and the Klaus Tschira Foundation (R.W.). We acknowledge the Supercomputing Center Brno for computational resources.

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2009.06.076

References

1. Zamocky, M., Herzog, C., Nykyri, L. M. & Koller, F. (1995). Site-directed mutagenesis of the lower parts of the major substrate channel of yeast catalase A leads to highly increased peroxidatic activity. *FEBS Lett.* **367**, 241–245.
2. Pikkemaat, M. G., Ridder, I. S., Rozeboom, H. J., Kalk, K. H., Dijkstra, B. W. & Janssen, D. B. (1999). Crystallographic and kinetic evidence of a collision complex formed during halide import in haloalkane dehalogenase. *Biochemistry*, **38**, 12052–12061.
3. Sevinc, M. S., Mate, M. J., Switala, J., Fita, I. & Loewen, P. C. (1999). Role of the lateral channel in catalase HPII of *Escherichia coli*. *Protein Sci.* **8**, 490–498.
4. Tara, S., Helms, V., Straatsma, T. P. & McCammon, J. A. (1999). Molecular dynamics of mouse acetylcholinesterase complexed with huperzine A. *Biopolymers*, **50**, 347–359.
5. Schmitt, J., Brocca, S., Schmid, R. D. & Pleiss, J. (2002). Blocking the tunnel: engineering of *Candida rugosa* lipase mutants with short chain length specificity. *Protein Eng.* **15**, 595–601.
6. Scott, E. E., He, Y. Q. & Halpert, J. R. (2002). Substrate routes to the buried active site may vary among cytochromes P450: mutagenesis of the F–G region in P450 2b1. *Chem. Res. Toxicol.* **15**, 1407–1413.
7. Chaloupkova, R., Sykorova, J., Prokop, Z., Jesenska, A., Monincova, M., Pavlova, M. *et al.* (2003). Modification of activity and specificity of haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26 by engineering of its entrance tunnel. *J. Biol. Chem.* **278**, 52622–52628.

8. Chelikani, P., Carpena, X., Fita, I. & Loewen, P. C. (2003). An electrical potential in the access channel of catalases enhances catalysis. *J. Biol. Chem.* **278**, 31290–31296.
9. Fishman, A., Tao, Y., Bentley, W. E. & Wood, T. K. (2004). Protein engineering of toluene 4-monooxygenase of *Pseudomonas mendocina* KR1 for synthesizing 4-nitrocatechol from nitrobenzene. *Biotechnol. Bioeng.* **87**, 779–790.
10. Jakopitsch, C., Droghetti, E., Schmuckenschlager, F., Furtmüller, P. G., Smulevich, G. & Obinger, C. (2005). Role of the main access channel of catalase–peroxidase in catalysis. *J. Biol. Chem.* **280**, 42411–42422.
11. Zhou, Y., Zheng, Q., Li, Z., Zhang, Y., Sun, M., Sun, C. *et al.* (2006). On the human CYP2C9∗13 variant activity reduction: a molecular dynamics simulation and docking study. *Biochimie*, **88**, 1457–1465.
12. Kotik, M., Stepanek, V., Kyslik, P. & Maresova, H. (2007). Cloning of an epoxide hydrolase-encoding gene from *Aspergillus niger* M200, overexpression in *E. coli*, and modification of activity and enantioselectivity of the enzyme by protein engineering. *J. Biotechnol.* **132**, 8–15.
13. Silberstein, M., Damborsky, J. & Vajda, S. (2007). Exploring the binding sites of the haloalkane dehalogenase DhlA from *Xanthobacter autotrophicus* GJ10. *Biochemistry*, **46**, 9239–9249.
14. Lüdemann, S. K., Lounnas, V. & Wade, R. C. (2000). How do substrates enter and products exit the buried active site of cytochrome P450cam? 1. Random expulsion molecular dynamics investigation of ligand access channels and mechanisms. *J. Mol. Biol.* **303**, 797–811.
15. Schleinkofer, K., Sudarko, Winn, P. J., Lüdemann, S. K. & Wade, R. C. (2005). Do mammalian cytochrome P450s show multiple ligand access pathways and ligand channeling? *EMBO Rep.* **6**, 584–589.
16. Winn, P. J., Lüdemann, S. K., Gauges, R., Lounnas, V. & Wade, R. C. (2002). Comparison of the dynamics of substrate access channels in three cytochrome P450s reveals different opening mechanisms and a novel functional role for a buried arginine. *Proc. Natl Acad. Sci. USA*, **99**, 5361–5366.
17. Wade, R. C., Winn, P. J., Schlichting, I. & Sudarko (2004). A survey of active site access channels in cytochromes P450. *J. Inorg. Biochem.* **98**, 1175–1182.
18. Cojocaru, V., Winn, P. J. & Wade, R. C. (2007). The ins and outs of cytochrome P450s. *Biochim. Biophys. Acta*, **1770**, 390–401.
19. Hritz, J., Zoldak, G. & Sedlak, E. (2006). Cofactor assisted gating mechanism in the active site of NADH oxidase from *Thermus thermophilus*. *Proteins*, **64**, 465–476.
20. Wen, Z., Baudry, J., Berenbaum, M. R. & Schuler, M. A. (2005). Ile115Leu mutation in the SRS1 region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. *Protein Eng. Des. Sel.* **18**, 191–199.
21. Huang, X. & Raushel, F. M. (2000). An engineered blockage within the ammonia tunnel of carbamoyl phosphate synthetase prevents the use of glutamine as a substrate but not ammonia. *Biochemistry*, **39**, 3240–3247.
22. Ollis, D. L., Cheah, E., Cygler, M., Dijkstra, B., Frolow, F., Franken, S. M. *et al.* (1992). The α/β hydrolase fold. *Protein Eng.* **5**, 197–211.
23. Janssen, D. B., Pries, F., van der Ploeg, J., Kazemier, B., Terpstra, P. & Witholt, B. (1989). Cloning of 1,2-dichloroethane degradation genes of *Xanthobacter autotrophicus* GJ10 and expression and sequencing of the *dhlA* gene. *J. Bacteriol.* **171**, 6791–6799.
24. Nagata, Y., Imai, R., Sakai, A., Fukuda, M., Yano, K. & Takagi, M. (1993). Isolation and characterization of tn5-induced mutants of *Pseudomonas paucimobilis* UT26 defective in gamma-hexachlorocyclohexane dehydrochlorinase (LinA). *Biosci. Biotechnol. Biochem.* **57**, 703–709.
25. Curragh, H., Flynn, O., Larkin, M. J., Stafford, T. M., Hamilton, J. T. & Harper, D. B. (1994). Haloalkane degradation and assimilation by *Rhodococcus rhodochrous* NCIMB13064. *Microbiology*, **140**, 1433–1442.
26. Krooshof, G. H., Kwant, E. M., Damborsky, J., Koca, J. & Janssen, D. B. (1997). Repositioning the catalytic triad aspartic acid of haloalkane dehalogenase: effects on stability, kinetics, and structure. *Biochemistry*, **36**, 9571–9580.
27. Newman, J., Peat, T. S., Richard, R., Kan, L., Swanson, P. E., Affholter, J. A. *et al.* (1999). Haloalkane dehalogenases: structure of a *Rhodococcus* enzyme. *Biochemistry*, **38**, 16105–16114.
28. Janssen, D. B. (2004). Evolving haloalkane dehalogenases. *Curr. Opin. Chem. Biol.* **8**, 150–159.
29. Chovancova, E., Kosinski, J., Bujnicki, J. M. & Damborsky, J. (2007). Phylogenetic analysis of haloalkane dehalogenases. *Proteins*, **67**, 305–316.
30. Schanstra, J. P., Kingma, J. & Janssen, D. B. (1996). Specificity and kinetics of haloalkane dehalogenase. *J. Biol. Chem.* **271**, 14747–14753.
31. Bosma, T., Pikkemaat, M. G., Kingma, J., Dijk, J. & Janssen, D. B. (2003). Steady-state and pre-steady-state kinetic analysis of halopropane conversion by a *Rhodococcus* haloalkane dehalogenase. *Biochemistry*, **42**, 8047–8053.
32. Prokop, Z., Monincova, M., Chaloupkova, R., Klvana, M., Nagata, Y., Janssen, D. B. & Damborský, J. (2003). Catalytic mechanism of the haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26. *J. Biol. Chem.* **278**, 45094–45100.
33. Verschueren, K. H., Seljee, F., Rozeboom, H. J., Kalk, K. H. & Dijkstra, B. W. (1993). Crystallographic analysis of the catalytic mechanism of haloalkane dehalogenase. *Nature*, **363**, 693–698.
34. Otyepka, M. & Damborsky, J. (2002). Functionally relevant motions of haloalkane dehalogenases occur in the specificity-modulating cap domains. *Protein Sci.* **11**, 1206–1217.
35. Petrek, M., Otyepka, M., Banas, P., Kosinova, P., Koca, J. & Damborsky, J. (2006). CAVER: a new tool to explore routes from protein clefts, pockets and cavities. *BMC Bioinformatics*, **7**, 316.
36. Carlsson, P., Burendahl, S. & Nilsson, L. (2006). Unbinding of retinoic acid from the retinoic acid receptor by random expulsion molecular dynamics. *Biophys. J.* **91**, 3151–3161.
37. Wang, T. & Duan, Y. (2007). Chromophore channeling in the G-protein coupled receptor rhodopsin. *J. Am. Chem. Soc.* **129**, 6970–6971.
38. Bosma, T., Damborsky, J., Stucki, G. & Janssen, D. B. (2002). Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. *Appl. Environ. Microbiol.* **68**, 3582–3587.
39. Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M. *et al.* (2009). Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. *Nat. Chem. Biol.* (in press). doi:10.1038/nchembio.205.

40. Banas, P., Otyepka, M., Jerabek, P., Petrek, M. & Damborsky, J. (2006). Mechanism of enhanced conversion of 1,2,3-trichloropropane by mutant haloalkane dehalogenase revealed by molecular modeling. *J. Comput.-Aided Mol. Des.* **20**, 375–383.
41. Negri, A., Marco, E., Damborsky, J. & Gago, F. (2007). Stepwise dissection and visualization of the catalytic mechanism of haloalkane dehalogenase LinB using molecular dynamics simulations and computer graphics. *J. Mol. Graphics Modell.* **26**, 643–651.
42. Lüdemann, S. K., Lounnas, V. & Wade, R. C. (2000). How do substrates enter and products exit the buried active site of cytochrome P450cam? 2. Steered molecular dynamics and adiabatic mapping of substrate pathways. *J. Mol. Biol.* **303**, 813–830.
43. Li, W., Liu, H., Scott, E. E., Gräter, F., Halpert, J. R., Luo, X., Shen, J. & Jiang, H. (2005). Possible pathway (s) of testosterone egress from the active site of cytochrome P450 2B1: a steered molecular dynamics simulation. *Drug Metab. Dispos.* **33**, 910–919.
44. Dunn, A. R., Dmochowski, I. J., Bilwes, A. M., Gray, H.B. & Crane, B. R. (2001). Probing the open state of cytochrome P450cam with ruthenium-linker substrates. *Proc. Natl Acad. Sci. USA*, **98**, 12420–12425.
45. Yao, H., McCullough, C. R., Costache, A. D., Pullela, P. K. & Sem, D. S. (2007). Structural evidence for a functionally relevant second camphor binding site in P450cam: model for substrate entry into a P450 active site. *Proteins*, **69**, 125–138.
46. Podust, L. M., Poulos, T. L. & Waterman, M. R. (2001). Crystal structure of cytochrome P450 14α-sterol demethylase (CYP51) from *Mycobacterium tuberculosis* in complex with azole inhibitors. *Proc. Natl Acad. Sci. USA*, **98**, 3068–3073.
47. Seifert, A., Tatzel, S., Schmid, R. D. & Pleiss, J. (2006). Multiple molecular dynamics simulations of human P450 monooxygenase CYP2C9: the molecular basis of substrate binding and regioselectivity toward warfarin. *Proteins*, **64**, 147–155.
48. Prasad, S., Mazumdar, S. & Mitra, S. (2000). Binding of camphor to *Pseudomonas putida* cytochrome P450cam: steady-state and picosecond time-resolved fluorescence studies. *FEBS Lett.* **477**, 157–160.
49. Strasser, A. & Wittmann, H. (2007). LIGPATH: a module for predictive calculation of a ligand's pathway into a receptor-application to the gpH_1-receptor. *J. Mol. Model.* **13**, 209–218.
50. Kennes, C., Pries, F., Krooshof, G. H., Bokma, E., Kingma, J. & Janssen, D. B. (1995). Replacement of tryptophan residues in haloalkane dehalogenase reduces halide binding and catalytic activity. *Eur. J. Biochem.* **228**, 403–407.
51. Damborsky, J., Kuty, M., Nemec, M. & Koca, J. (1997). A molecular modelling study of the catalytic mechanism of haloalkane dehalogenase: 1. Quantum chemical study of the first reaction step. *J. Chem. Inf. Comput. Sci.* **37**, 562–568.
52. Damborsky, J. & Koca, J. (1999). Analysis of the reaction mechanism and substrate specificity of haloalkane dehalogenases by sequential and structural comparisons. *Protein Eng.* **12**, 989–998.
53. Schindler, J. F., Naranjo, P. A., Honaberger, D. A., Chang, C. H., Brainard, J. R., Vanderberg, L. A. & Unkefer, C. J. (1999). Haloalkane dehalogenases: steady-state kinetics and halide inhibition. *Biochemistry*, **38**, 5772–5778.
54. Marek, J., Vevodova, J., Smatanova, I. K., Nagata, Y., Svensson, L. A., Newman, J. *et al.* (2000). Crystal structure of the haloalkane dehalogenase from *Sphingomonas paucimobilis* UT26. *Biochemistry*, **39**, 14082–14086.
55. Bohac, M., Nagata, Y., Prokop, Z., Prokop, M., Monincova, M., Tsuda, M. *et al.* (2002). Halide-stabilizing residues of haloalkane dehalogenases studied by quantum mechanic calculations and site-directed mutagenesis. *Biochemistry*, **41**, 14272–14280.
56. Monincova, M., Prokop, Z., Vevodova, J., Nagata, Y. & Damborsky, J. (2007). Weak activity of haloalkane dehalogenase LinB with 1,2,3-trichloropropane revealed by X-ray crystallography and microcalorimetry. *Appl. Environ. Microbiol.* **73**, 2005–2008.
57. Streltsov, V. A., Prokop, Z., Damborsky, J., Nagata, Y., Oakley, A. & Wilce, M. C. J. (2003). Haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26: X-ray crystallographic studies of dehalogenation of brominated substrates. *Biochemistry*, **42**, 10104–10112.
58. Oakley, A. J., Klvana, M., Otyepka, M., Nagata, Y., Wilce, M. C. J. & Damborsky, J. (2004). Crystal structure of haloalkane dehalogenase LinB from *Sphingomonas paucimobilis* UT26 at 0.95 Å resolution: dynamics of catalytic residues. *Biochemistry*, **43**, 870–878.
59. Deprez, E., Gerber, N. C., Di Primo, C., Douzou, P., Sligar, S. G. & Hui Bon Hoa, G. (1994). Electrostatic control of the substrate access channel in cytochrome P450cam. *Biochemistry*, **33**, 14464–14468.
60. Di Primo, C., Deprez, E., Sligar, S. G. & Hui Bon Hoa, G. (1997). Origin of the photoacoustic signal in cytochrome P450cam: role of the Arg186-Asp251-Lys178 bifurcated salt bridge. *Biochemistry*, **36**, 112–118.
61. Lounnas, V. & Wade, R. C. (1997). Exceptionally stable salt bridges in cytochrome P450cam have functional roles. *Biochemistry*, **36**, 5402–5417.
62. Oprea, T. I., Hummer, G. & Garcia, A. E. (1997). Identification of a functional water channel in cytochrome P450 enzymes. *Proc. Natl Acad. Sci. USA*, **94**, 2133–2138.
63. Kovach, I. M., Qian, N. & Bencsura, A. (1994). Efficient product clearance through exit channels in substrate hydrolysis by acetylcholinesterase. *FEBS Lett.* **349**, 60–64.
64. Enyedy, I., Kovach, I. & Brooks, B. (1998). Alternate pathways for acetic acid and acetate ion release from acetylcholinesterase: a molecular dynamics study. *J. Am. Chem. Soc.* **120**, 8043–8050.
65. Ripoll, D. R., Faerman, C. H., Axelsen, P. H., Silman, I. & Sussman, J. L. (1993). An electrostatic mechanism for substrate guidance down the aromatic gorge of acetylcholinesterase. *Proc. Natl Acad. Sci. USA*, **90**, 5128–5132.
66. Van Belle, D., De Maria, L., Iurcu, G. & Wodak, S. J. (2000). Pathways of ligand clearance in acetylcholinesterase by multiple copy sampling. *J. Mol. Biol.* **298**, 705–726.
67. Xu, Y., Shen, J., Luo, X., Silman, I., Sussman, J. L., Chen, K. & Jiang, H. (2003). How does huperzine a enter and leave the binding gorge of acetylcholinesterase? Steered molecular dynamics simulations. *J. Am. Chem. Soc.* **125**, 11340–11349.
68. Khajehpour, M., Rietveld, I., Vinogradov, S., Prabhu, N. V., Sharp, K. A. & Vanderkooi, J. M. (2003). Accessibility of oxygen with respect to the heme pocket in horseradish peroxidase. *Proteins*, **53**, 656–666.
69. Zelent, B., Kaposi, A., Nucci, N., Sharp, K., Dalosto, S., Wright, W. & Vanderkooi, J. (2004). Water channel

of horseradish peroxidase studied by the charge-transfer absorption band of ferric heme. *J. Phys. Chem. B*, **108**, 10317–10324.
70. Teeter, M. M. (2004). Myoglobin cavities provide interior ligand pathway. *Protein Sci.* **13**, 313–318.
71. Bui, J. M., Henchman, R. H. & McCammon, J. A. (2003). The dynamics of ligand barrier crossing inside the acetylcholinesterase gorge. *Biophys. J.* **85**, 2267–2272.
72. Ye, X., Yu, A. & Champion, P. M. (2006). Dynamics of nitric oxide rebinding and escape in horseradish peroxidase. *J. Am. Chem. Soc.* **128**, 1444–1445.
73. Henchman, R. H., Tai, K., Shen, T. & McCammon, J. A. (2002). Properties of water molecules in the active site gorge of acetylcholinesterase from computer simulation. *Biophys. J.* **82**, 2671–2682.
74. Zwanzig, R. (1992). Dynamical disorder: passage through a fluctuating bottleneck. *J. Chem. Phys.* **97**, 3587–3589.
75. Carugo, O. & Argos, P. (1998). Accessibility to internal cavities and ligand binding sites monitored by protein crystallographic thermal factors. *Proteins*, **31**, 201–213.
76. Cohen, J., Arkhipov, A., Braun, R. I. & Schulten, K. (2006). Imaging the migration pathways for O_2, CO, NO, and Xe inside myoglobin. *Biophys. J.* **91**, 1844–1857.
77. Lavalette, D., Tétreau, C. & Mouawad, L. (2006). Ligand migration and escape pathways in haem proteins. *Biochem. Soc. Trans.* **34**, 975–978.
78. Maniatis, T., Fritsch, E. F. & Sambrook, J. (1982). *Molecular Cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
79. Stsiapanava, A., Koudelaková, T., Lapkouski, M., Pavlova, M., Damborsky, J. & Smatanova, I. K. (2008). Crystals of DhaA mutants from *Rhodococcus rhodochrous* NCIMB13064 diffracted to ultrahigh resolution: crystallization and preliminary diffraction analysis. *Acta Crystallogr., Sect. F*, **64**, 137–140.
80. Andrade, M. A., Chacon, P., Merelo, J. J. & Moran, F. (1993). Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng.* **6**, 383–390.
81. Sreerama, N. & Woody, R. W. (1993). A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal. Biochem.* **209**, 32–44.
82. Deleage, G. & Geourjon, C. (1993). An interactive graphic program for calculating the secondary structure content of proteins from circular dichroism spectrum. *Comput. Appl. Biosci.* **9**, 197–199.
83. Iwasaki, I., Utsumi, S. & Ozawa, T. (1952). New colorimetric determination of chloride using mercuric thiocyanate and ferric ion. *Bull. Chem. Soc. Jpn.* **25**, 226.
84. DeLano, W. (2002). *The PyMOL Molecular Graphics System.* DeLano Scientific, Palo Alto, CA.
85. Stewart, J. J. (1990). MOPAC: a semiempirical molecular orbital program. *J. Comput.-Aided Mol. Des.* **4**, 1–105.
86. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Gill, P. M. W., Johnson, B. J., Robb, M. A. *et al.* (1995). Gaussian 94, Revission D.4 Gaussian, Inc., Pittsburgh, PA.
87. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E. *et al.* (2004). AMBER 8 University of California, San Francisco, CA.
88. Morris, G. M., Goodsell, D. S., Halliday, R. S., Huey, R., Hart, W. E., Belew, R. K. & Olson, A. J. (1998). Automated docking using a Lamarckian genetic algorithm and an empirical binding free energy function. *J. Comput. Chem.* **19**, 1639–1662.
89. Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. *J. Mol. Graphics*, **8**, 52–56.
90. Cornell, W., Cieplak, P., Bayly, C., Gould, I., Merz, K., Ferguson, D. *et al.* (1996). A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. *J. Am. Chem. Soc.* **118**, 2309.
91. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W. & Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
92. Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A. & Haak, J. R. (1984). Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
93. Ryckaert, J., Ciccotti, G. & Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of *n*-alkanes. *J. Comput. Phys.* **23**, 327–341.
94. Kabsch, W. & Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers*, **22**, 2577–2637.
95. Humphrey, W., Dalke, A. & Schulten, K. (1996). VMD: visual molecular dynamics. *J. Mol. Graphics*, **14**, 33–38.
96. Onufriev, A., Bashford, D. & Case, D. A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized Born model. *Proteins*, **55**, 383–394.

ChemPubSoc Europe

DOI: 10.1002/cbic.201402265

Maximizing the Efficiency of Multienzyme Process by Stoichiometry Optimization

Pavel Dvorak,[a, b] Nagendra P. Kurumbang,[a] Jaroslav Bendl,[a, b, c] Jan Brezovsky,[a] Zbynek Prokop,[a, b] and Jiri Damborsky*[a, b]

Multienzyme processes represent an important area of biocatalysis. Their efficiency can be enhanced by optimization of the stoichiometry of the biocatalysts. Here we present a workflow for maximizing the efficiency of a three-enzyme system catalyzing a five-step chemical conversion. Kinetic models of pathways with wild-type or engineered enzymes were built, and the enzyme stoichiometry of each pathway was optimized. Mathematical modeling and one-pot multienzyme experiments provided detailed insights into pathway dynamics, enabled the selection of a suitable engineered enzyme, and afforded high efficiency while minimizing biocatalyst loadings. Optimizing the stoichiometry in a pathway with an engineered enzyme reduced the total biocatalyst load by an impressive 56%. Our new workflow represents a broadly applicable strategy for optimizing multienzyme processes.

In vitro multienzyme processes have great potential for the biosynthesis of fine and bulk chemicals, for bioremediation and biosensing.[1] Studies on two-enzyme systems dominate the literature, but systems of three,[2] four,[3] and even twelve[4] or thirteen[5] enzymes are known. Multienzyme systems are superior to single-enzyme biocatalysis in that they can catalyze more-complex synthetic schemes. In vitro multienzyme networks enable simpler process control than analogous in vivo systems and suffer less from reactant toxicity.[6] However, in both system types reaction bottlenecks often arise from imbalanced enzyme properties.

Protein engineering is often used to improve the catalytic properties and stability of enzymes.[7] Many engineered enzymes can be used in multienzyme processes, but methods for predicting their impact on productivity are lacking. Kinetic modeling is essential for analyzing enzymatic reactions and can enable their rational optimization.[1c,8] However, there are only few accurate kinetic models of in vitro multienzyme systems.[2,3,5,9] Available models rarely have experimental support, and their use in optimizing processes with engineered enzymes has not been adequately explored.

The aim of this study was to develop a workflow for optimizing multienzyme processes by using kinetic modeling to predict the effects of varying enzyme stoichiometry and employing available engineered enzymes. Our model system (Scheme 1) was a synthetic metabolic pathway for the five-step biotransformation of toxic industrial waste product 1,2,3-trichloropropane (TCP) into glycerol (GLY).[10] This pathway has been previously assembled inside living cells[11] and is based on haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCIMB 13064,[12] haloalcohol dehalogenase HheC from *Agrobacterium radiobacter* AD1,[13] and epoxide hydrolase EchA from *Agrobacterium radiobacter* AD1.[14] Herein, three DhaA variants were assessed: 1) wild-type DhaA, and the previously constructed mutants 2) DhaA31 (improved activity)[15] and 3) DhaA90R (increased enantioselectivity).[16] Kinetic models were built for pathways with each variant, and the enzyme stoichiometry was optimized under defined constraints.

We initially prepared soluble enzymes with purities of $\geq$95% for DhaA and EchA and $\geq$85% for HheC (Supporting Information). To validate the one-pot multienzyme biotransformation of TCP into GLY, 1 mg each of purified DhaA, HheC, and EchA were mixed in 10 mL of Tris-SO_4 buffer (pH 8.5) and incubated with 2 mM TCP at 37 °C for 300 min. The enzymes have similar molecular weights (34.1, 29.3, and 36.5 kDa, respectively), so mass ratio roughly equals molar ratio. A GC method for detecting and quantifying TCP and all pathway intermediates in a single analysis was developed and used to monitor the five-step process (see the Experimental Section and Figure S1 in the Supporting Information). The time course for the conversion confirmed pathway viability and revealed two major bottlenecks: 1) conversion of TCP into 2,3-dichloropropan-1-ol (DCP), and 2) mismatched selectivity of DhaA and HheC, thus causing accumulation of (*S*)-DCP (Scheme 1 and Figure S2). Similar bottlenecks have been identified in vivo.[11]

We then investigated two DhaA mutants with properties tuned to address these bottlenecks. DhaA31 was previously constructed in our laboratory by using computer-aided directed evolution of protein tunnels.[15] Its catalytic rate towards TCP is 32 times that of wild-type DhaA. Mutant DhaA90R was obtained by van Leeuwen and co-workers ("DhaAr5-90R") by

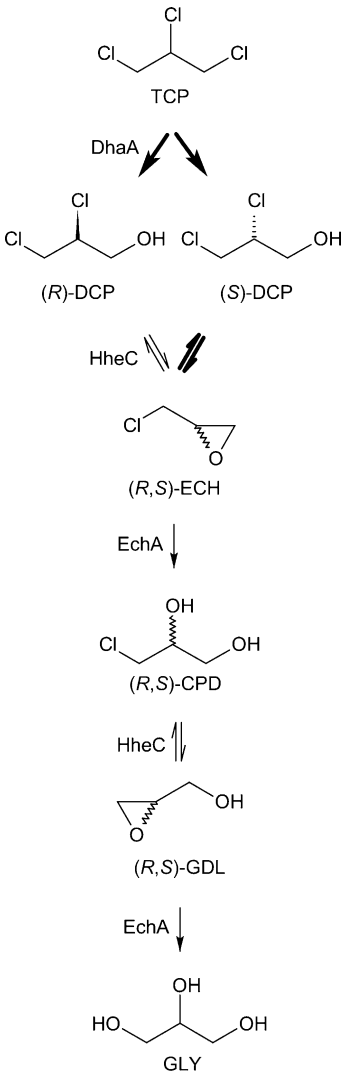
[a] *P. Dvorak, Dr. N. P. Kurumbang, J. Bendl, Dr. J. Brezovsky, Dr. Z. Prokop, Prof. J. Damborsky*
Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX
Faculty of Science, Masaryk University
Kamenice 5/A13, 625 00 Brno (Czech Republic)
E-mail: jiri@chemi.muni.cz

[b] *P. Dvorak, J. Bendl, Dr. Z. Prokop, Prof. J. Damborsky*
International Clinical Research Center
St. Anne's University Hospital Brno
Pekarska 53, 656 91 Brno (Czech Republic)

[c] *J. Bendl*
Department of Information Systems
Faculty of Information Technology, Brno University of Technology
Bozetechova 1, 612 00 Brno (Czech Republic)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201402265.

 © 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim



Scheme 1. Synthetic pathway for the three-enzyme biotransformation of 1,2,3-trichloropropane. Five consecutive steps are catalyzed by the haloalkane dehalogenase DhaA, from *Rhodococcus rhodochrous* NCIMB 13064, haloalcohol dehalogenase HheC from *Agrobacterium radiobacter* AD1, and epoxide hydrolase EchA from *Agrobacterium radiobacter* AD1. 1,2,3-trichloropropane (TCP) is converted via (*R*)- (*S*)-2,3-dichloropropane-1-ol (DCP), epichlorohydrin (ECH), 3-chloropropane-1,2-diol (CPD), and glycidol (GDL) to glycerol (GLY). Key bottlenecks are indicated by bold arrows.

focused directed evolution of DhaA31. It has the same activity as wild-type DhaA but seven times higher specificity for the (*R*)-DCP configuration.[16] Time courses for the three-enzyme conversion of TCP with purified DhaA31 or DhaA90R were recorded, as for the wild type. The differences between the resulting conversion profiles were consistent with the kinetic properties of DhaA, DhaA31, and DhaA90R (Figure S2 and Tables S2–S4).

Sixteen steady-state kinetic parameters were determined for the purified enzymes to establish a kinetic model of the pathway (Table 1 and the Supporting Information). All the studied reactions exhibited Michaelis–Menten kinetics. The conversion of glycidol (GDL) into GLY was described by a Michaelis–Menten equation, with two inhibition constants defining the inhibitory effects of GLY (product inhibition constant K_i = 1.00 mM) and TCP (shared equilibrium inhibition constant K_c = 0.21 mM; Supporting Information). These effects together with the substrate preference of EchA for epichlorohydrin (ECH) caused GDL accumulation (Figure S2).

Table 1. Experimental steady-state kinetic parameters in the kinetic model.

DhaA		HheC	
$K_{m,TCP}$ [mM]	1.01 ± 0.08	$K_{m,(R)\text{-DCP}}$ [mM]	2.49 ± 0.16
$k_{cat,TCP,(R)\text{-DCP}}$ [s^{-1}]	0.04[a]	$K_{m,(S)\text{-DCP}}$ [mM]	3.33 ± 0.51
$k_{cat,TCP,(S)\text{-DCP}}$ [s^{-1}]	0.03[a]	$K_{m,CPD}$ [mM]	0.86 ± 0.07
DhaA31		$k_{cat,(R)\text{-DCP}}$ [s^{-1}]	1.81 ± 0.05
$K_{m,TCP}$ [mM]	1.79 ± 0.09	$k_{cat,(S)\text{-DCP}}$ [s^{-1}]	0.08 ± 0.00
$k_{cat,TCP,(R)\text{-DCP}}$ [s^{-1}]	0.58[b]	$k_{cat,CPD}$ [s^{-1}]	2.38 ± 0.06
$k_{cat,TCP,(S)\text{-DCP}}$ [s^{-1}]	0.47[b]	EchA	
DhaA90R		$K_{m,ECH}$ [mM]	0.09 ± 0.08
$K_{m,TCP}$ [mM]	12.56 ± 2.99	$K_{m,GDL}$ [mM]	3.54 ± 0.09
$k_{cat,TCP,(R)\text{-DCP}}$ [s^{-1}]	0.19[c]	$k_{cat,ECH}$ [s^{-1}]	14.37 ± 0.52
$k_{cat,TCP,(S)\text{-DCP}}$ [s^{-1}]	0.02[c]	$k_{cat,GDL}$ [s^{-1}]	3.96 ± 0.08

[a] Determined from the ratio of (*R*)- and (*S*)-DCP production as 56 and 44% of 0.07 ± 0.00 s^{-1} for $k_{cat,TCP,(R)\text{-DCP}}$ and $k_{cat,TCP,(S)\text{-DCP}}$, respectively (Supporting Information, Experimental Section). [b] Determined as 55 and 45% of 1.05 ± 0.02 s^{-1} for $k_{cat,TCP,(R)\text{-DCP}}$ and $k_{cat,TCP,(S)\text{-DCP}}$, respectively. [c] Determined as 90 and 10% of 0.21 ± 0.03 s^{-1} for $k_{cat,TCP,(R)\text{-DCP}}$ and $k_{cat,TCP,(S)\text{-DCP}}$, respectively.

The kinetic model was validated against experimental data for the three-enzyme conversion of 2 mM TCP (Figure S2). The two datasets were in good agreement: with 1:1:1 mixtures of the three enzymes (total 3 mg), the predicted (and measured) productivities of the DhaA, DhaA31, and DhaA90R pathways were 72% (62%), 85% (85%) and 45% (42%), respectively. No enzyme inactivation occurred, but the model could be extended to include inactivation constants if necessary.

The model was then used to predict the DhaA, HheC, and EchA loadings needed to achieve 95% conversion of TCP into GLY under the chosen conditions by simulating stepwise increases in enzyme loading within the reaction system until the productivity goal was reached (see the Experimental Section and the Supporting Information). At a DhaA/HheC/EchA mass ratio of 1:1:1, the wild-type haloalkane dehalogenase pathway reached the productivity goal with 2.4 mg of each enzyme (total 7.2 mg; Figure 1). The DhaA31 and DhaA90R pathways required 1.8 mg and 4 mg of each enzyme, respectively, (i.e., total enzyme loads of 5.4 and 12 mg) to reach the goal. The differences in the modeled time courses and quantities of enzyme required to achieve 95% conversion for the three pathway variants demonstrate the profound effects of introducing engineered DhaA variants (Figure 1). Despite pronounced accumulation of DCP and GDL during the initial 25 min, the DhaA31 pathway was around 25% more efficient than the wild-type version. DhaA31 significantly accelerated the conversion of TCP and thus accelerated the consumption of accumulated GDL by suppressing TCP's inhibitory effect. In contrast, DhaA90R reduced system efficiency despite effectively minimizing DCP accumulation. This selective but catalytically inefficient mutant was thus not beneficial.

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

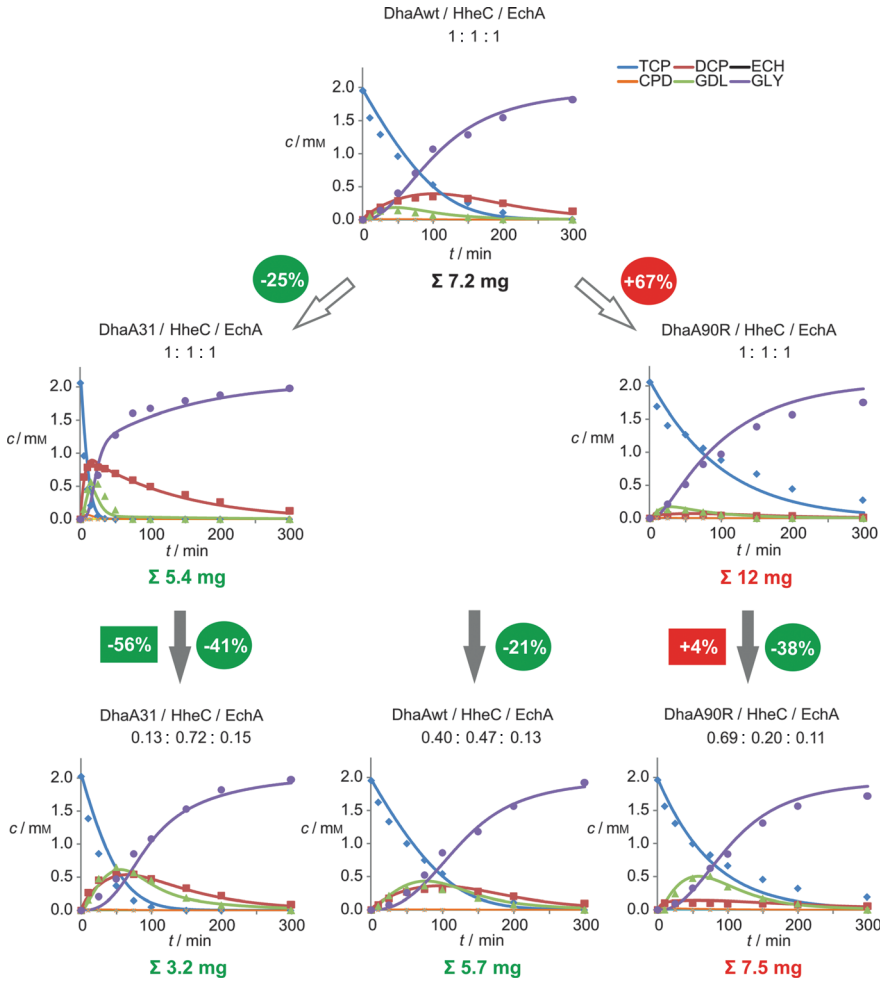


Figure 1. Optimization of three-enzyme conversion of 1,2,3-trichloropropane by kinetic modeling and employment of engineered enzyme variants. Calculated and measured results are indicated by solid lines and symbols, respectively. The following parameters were constrained: reaction volume (10 mL), initial TCP concentration (2 mM), and reaction duration (300 min). The initial optimization goal was 95% conversion of TCP into GLY within 300 min by a 1:1:1 ratio (wild-type enzyme). The calculated and experimentally verified total enzyme loading (Σ) required to achieve this was 7.2 mg. The wild-type DhaA enzyme was then replaced with mutants DhaA31[23] or DhaA90R[24] to study the effect of their kinetics on the pathway (white arrows). Further optimization was achieved by tuning the enzyme ratios (gray arrows). Reductions in total enzyme loading achieved by employment of mutants or stoichiometry optimization alone and in combination are shown in circles and squares, respectively. Experimental concentrations of TCP, DCP, ECH, CPD, and GDL were determined by GC. Concentrations of GLY were determined spectrophotometrically. Data points represent mean values from three independent experiments. (Error bars are omitted for clarity; standard deviations are provided in the Supporting Information; Tables S2–S4).

We then evaluated the effects of enzyme stoichiometry on efficiency. An algorithm was used to minimize the total enzyme load without sacrificing productivity (Experimental Section and Animation S1 in the Supporting Information). Optimal enzyme mass ratios were calculated for each pathway. The modeled time courses for individual reactions were similar in each case, but the optimized ratios and total enzyme loadings differed significantly between pathways (Figure 1). Enzyme stoichiometry optimization increased efficiency in the DhaA, DhaA31, and DhaA90R pathways and reduced the total enzyme loading required for >95% conversion, by 21, 41, and 38%, respectively.

All optimization simulations were validated by testing the enzyme mass ratios in vitro. The resulting data agreed very closely with the predictions (Figure 1). The optimized DhaA90R pathway was around 10% less productive than predicted, possibly because the K_m of DhaA90R for TCP was underestimated due to the limited water solubility of TCP (~10 mM). In all other cases, the optimized systems achieved productivities of 94–98% (Tables S2–S4). Because the experimental time courses only reflect optimal cases based on predefined constraints, we used the simulated data to create 3D isoproductivity charts to show the effects of varying the loadings of DhaA variant, HheC, and EchA (Figure 2). These show the limiting components for each pathway and can be used to identify solutions with similar productivities.

Our results demonstrate that both modifying enzyme kinetic parameters and optimizing enzyme stoichiometry improved the efficiency of the studied multienzyme system. However, the far simpler process of stoichiometry optimization had a greater impact than introducing engineered enzymes. The optimized pathway using wild-type DhaA required a similar total enzyme load to the non-optimized pathway using the engineered DhaA31 (5.7 vs. 5.4 mg), thus showing that kinetic modeling alone can provide excellent solutions in certain cases. Naturally, the best result was achieved with an optimized pathway using engineered DhaA31: additive effects in this case reduced the catalyst load required for 95% productivity by 56% relative to the unoptimized wild-type pathway (Figure 1). This would be very economically beneficial in a large-scale industrial process.

In summary, we present experimentally validated in silico optimization of a multienzyme process by biocatalyst stoichiometry tuning. Our workflow entails 1) experimental verification of process viability, 2) determination of enzyme kinetics, 3) identification of pathway bottlenecks and selection of suitable engineered enzymes, 4) development of a robust and accurate kinetic model, 5) experimental model validation, and 6) process optimization by in silico enzyme stoichiometry modeling. Recent progress in efficient enzyme stabilization[17] and ex vivo

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

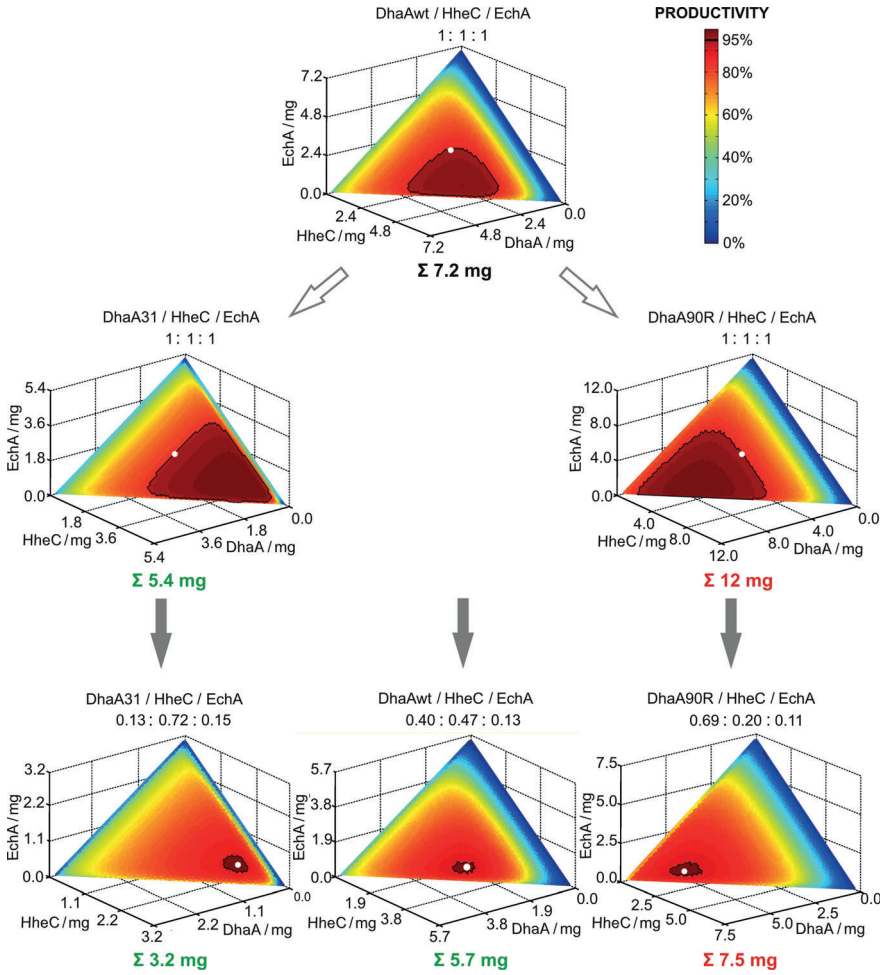


Figure 2. Isoproductivity color charts showing how biocatalyst concentration affects productivity in the three-enzyme conversion of 1,2,3-trichloropropane into glycerol. Applied constraints are as in Figure 1; the optimal solutions in each figure are indicated by white dots. Values on axes represent the loading of the corresponding enzyme in the reaction mixture. The black line indicates the threshold productivity (95%).

cofactor regeneration[18] together with the development of integrated analytical techniques,[19] increases in computational power, and the growing availability of kinetic data, will enable the refinement of many useful biotransformations. We believe that the workflow described herein and implemented in the provided computer code (Supporting Information) represents a widely applicable strategy for rapid optimization of multienzyme processes.

Experimental Section

Computational optimization of the multienzyme conversion of TCP: The initial constraints in modeling multienzyme conversion of TCP (to verify the developed kinetic model) were as follows: TCP starting concentration 2 mM (initial experimental concentration as close to 2 mM as possible; Tables S2–S4); the multienzyme reaction was allowed to proceed for 300 min; the concentration of each enzyme (DhaA variant, HheC, and EchA) was 0.1 mg mL^{-1} (total enzyme loading in 10 mL, 3 mg). Dynamic simulations of the multienzyme system based on a series of Michaelis–Menten equations (Supporting Information) were performed by using code written in Python 2.7 (Software S1 in the Supporting Information). The equations were expressed in differential form and integrated by using Euler's method (step size, 0.18 s). The effects of using different DhaA variants on the performance of the multienzyme process were evaluated with the following constraints: at least 95% conversion of TCP into GLY was to be achieved within 300 min when using DhaA variant, HheC, and EchA in an unoptimized mass ratio (1:1:1) at a total enzyme loading of 3.0 mg. The optimization algorithm increased the total enzyme loading in a stepwise fashion, by using increments of around 0.1 mg (i.e., increasing the loading of each individual enzyme by 0.066 mg in each step) until 95% conversion was surpassed. The enzyme stoichiometry in each of three pathways was then optimized to identify the DhaA variant/HheC/EchA ratio that would yield 95% conversion of 2 mM TCP into GLY within 300 min while minimizing the combined loading of the three enzymes. The algorithm searched for the most efficient ratio of three enzymes in the system starting with a total enzyme loading of 3 mg. The total enzyme loading within the system was increased in 0.1 mg increments until the most efficient ratio surpassed 95% conversion. For each total enzyme loading, 496 enzyme ratios were evaluated with step of 3% of a given total amount.

Multienzyme conversion of TCP in batch experiments: The multienzyme conversion of TCP (2 mM) was assayed in Tris-SO_4 buffer (10 mL, 50 mM, pH 8.5) in 25 mL micro-flasks sealed with Mininert valves (Alltech, USA) at 37 °C. The reaction was initiated by adding a defined amounts of purified DhaA variant, HheC, and EchA. Samples were periodically taken, mixed 1:1 with acetone containing hexanol (4 mM) as an internal standard, and analyzed by GC to determine the concentrations of TCP, DCP, ECH, CPD, and GDL (Supporting Information). Selected samples were analyzed by GC-MS to verify the identities of the metabolites. The concentration of GLY in the reaction mixture was determined spectrophotometrically by using the Free Glycerol Assay Kit (BioVision, Milpitas, CA). Details are provided in the Supporting Information.

Acknowledgements

We thank Lukas Maier for assistance with the enzymatic synthesis of (S)-2,3-dichloropropan-1-ol. This work was funded by the Grant Agency of the Czech Republic (P503/12/0572), the European Regional Development Fund (CZ.1.05/1.1.00/02.0123), the Min-

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

istry of Education of the Czech Republic (LO1214, CZ.1.07/2.3.00/20.0239), the SoMoPro Program (SYNTBIO No. 2SGA2873), the European Union (REGPOT 316345), the Research and Application of Advanced Methods in ICT (FIT-S-14–2299), the "Employment of the Best Young Scientists for International Cooperation Empowerment" Program (CZ1.07/2.3.00/30.0037). CERIT-SC is acknowledged for providing access to their computing facilities (CZ.1.05/3.2.00/08.0144).

Keywords: biocatalysis · biotransformations · kinetic modeling · multienzyme reaction · stoichiometry optimization

[1] a) I. Ardao, E. T. Hwang, A.-P. Zeng, *Adv. Biochem. Eng./Biotechnol.* **2013**, *137*, 153; b) Z. Findrik, D. Vasić-Rački, *Chem. Biochem. Eng. Q.* **2009**, *23*, 545; c) P. A. Santacoloma, G. Sin, K. V. Gernaey, J. M. Woodley, *Org. Process Res. Dev.* **2011**, *15*, 203; d) S. Schoffelen, J. C. M. van Hest, *Curr. Opin. Struct. Biol.* **2013**, *23*, 613.

[2] N. Ishii, Y. Suga, A. Hagiya, H. Watanabe, H. Mori, M. Yoshino, M. Tomita, *FEBS Lett.* **2007**, *581*, 413.

[3] Z. Findrik, D. Vasić-Rački, *Biotechnol. Bioeng.* **2007**, *98*, 956.

[4] A. Itoh, Y. Ohashi, T. Soga, H. Mori, T. Nishioka, M. Tomita, *Electrophoresis* **2004**, *25*, 1996.

[5] I. Ardao, A.-P. Zeng, *Chem. Eng. Sci.* **2013**, *87*, 183.

[6] a) F. Lopez-Gallego, C. Schmidt-Dannert, *Curr. Opin. Chem. Biol.* **2010**, *14*, 174; b) A. Meyer, R. Pellaux, S. Panke, *Curr. Opin. Microbiol.* **2007**, *10*, 246; c) S. Billerbeck, J. Härle, S. Panke, *Curr. Opin. Biotechnol.* **2013**, *24*, 1037.

[7] a) A. S. Bommarius, J. K. Blum, M. J. Abrahamson, *Curr. Opin. Chem. Biol.* **2011**, *15*, 194; b) U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robinson, *Nature* **2012**, *485*, 185.

[8] a) D. G. Blackmond, *Angew. Chem. Int. Ed.* **2005**, *44*, 4302; *Angew. Chem.* **2005**, *117*, 4374; b) R. Xue, J. M. Woodley, *Bioresour. Technol.* **2012**, *115*, 183.

[9] W. Van Hecke, A. Bhaqwat, R. Ludwiq, J. Dewulf, D. Haltrich, H. Van Lanqenhove, *Biotechnol. Bioeng.* **2009**, *102*, 1475.

[10] G. Samin, D. B. Janssen, *Environ. Sci. Pollut. Res. Int.* **2012**, *19*, 3067.

[11] a) T. Bosma, E. Kruizinga, E. J. de Bruin, G. J. Poelarends, D. B. Janssen, *Appl. Environ. Microbiol.* **1999**, *65*, 4575; b) T. Bosma, J. Damborský, G. Stucki, D. B. Janssen, *Appl. Environ. Microbiol.* **2002**, *68*, 3582; c) N. P. Kurumbang, P. Dvorak, J. Bendl, J. Brezovsky, Z. Prokop, J. Damborsky, *ACS Synth. Biol.* **2014**, *3*, 172.

[12] A. N. Kulakova, M. J. Larkin, L. A. Kulakov, *Microbiology* **1997**, *143*, 109.

[13] J. E. T. van Hylckama Vlieg, L. Tang, J. H. Lutje Spelberg, T. Smilda, G. J. Poelarends, T. Bosma, A. E. J. van Merode, M. W. Fraaije, D. B. Janssen, *J. Bacteriol.* **2001**, *183*, 5058.

[14] R. Rink, M. Fennema, M. Smids, U. Dehmel, D. B. Janssen, *J. Biol. Chem.* **1997**, *272*, 14650.

[15] M. Pavlova, M. Klvana, Z. Prokop, R. Chaloupkova, P. Banas, M. Otyepka, R. C. Wade, M. Tsuda, Y. Nagata, J. Damborsky *Nat. Chem. Biol.* **2009**, *5*, 727.

[16] J. G. van Leeuwen, H. J. Wijma, R. J. Floor, J.-M. van der Laan, D. B. Janssen, *ChemBioChem* **2012**, *13*, 137.

[17] a) T. Koudelakova, R. Chaloupkova, J. Brezovsky, Z. Prokop, E. Sebestova, M. Hesseler, M. Khabiri, M. Plevaka, D. Kulik, I. Kuta Smatanova, P. Rezacova, R. Ettrich, U. T. Bornscheuer, J. Damborsky, *Angew. Chem. Int. Ed.* **2013**, *52*, 1959; *Angew. Chem.* **2013**, *125*, 2013; b) M. T. Reetz, J. D. Carballeira, A. Vogel, *Angew. Chem. Int. Ed.* **2006**, *45*, 7745; *Angew. Chem.* **2006**, *118*, 7909.

[18] a) L. Lauterbach, O. Lenz, K. A. Vincent, *FEBS J.* **2013**, *280*, 3058; b) R. A. Scism, B. O. Bachmann, *ChemBioChem* **2010**, *11*, 67.

[19] M. Bujara, M. Schümperli, R. Pellaux, M. Heinemann, S. Panke, *Nat. Chem. Biol.* **2011**, *7*, 271.

Received: May 26, 2014
Published online on August 5, 2014

© 2014 Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim

ACS SyntheticBiology

Research Article

pubs.acs.org/synthbio

Computer-Assisted Engineering of the Synthetic Pathway for Biodegradation of a Toxic Persistent Pollutant

Nagendra Prasad Kurumbang,[†,⊥] Pavel Dvorak,[†,‡,⊥] Jaroslav Bendl,[†,§] Jan Brezovsky,[†] Zbynek Prokop,*[,†,‡] and Jiri Damborsky*[,†,‡]

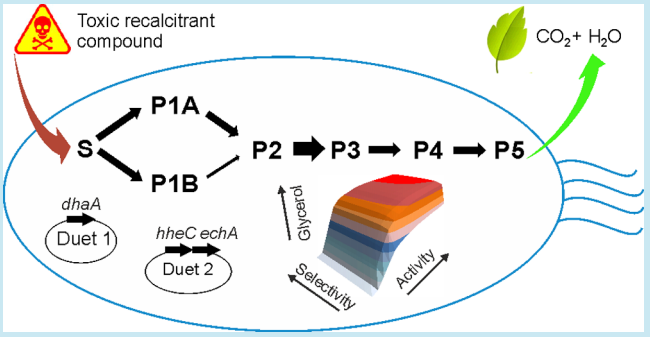
[†]Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic

[‡]International Clinical Research Center, St. Anne's University Hospital Brno, Pekarska 53, 656 91 Brno, Czech Republic

[§]Department of Information Systems, Faculty of Information Technology, Brno University of Technology, Bozetechova 1, 612 00 Brno, Czech Republic

Supporting Information

ABSTRACT: Anthropogenic halogenated compounds were unknown to nature until the industrial revolution, and microorganisms have not had sufficient time to evolve enzymes for their degradation. The lack of efficient enzymes and natural pathways can be addressed through a combination of protein and metabolic engineering. We have assembled a synthetic route for conversion of the highly toxic and recalcitrant 1,2,3-trichloropropane to glycerol in *Escherichia coli*, and used it for a systematic study of pathway bottlenecks. Optimal ratios of enzymes for the maximal production of glycerol, and minimal toxicity of metabolites were predicted using a mathematical model. The strains containing the expected optimal ratios of enzymes were constructed and characterized for their viability and degradation efficiency. Excellent agreement between predicted and experimental data was observed. The validated model was used to quantitatively describe the kinetic limitations of currently available enzyme variants and predict improvements required for further pathway optimization. This highlights the potential of forward engineering of microorganisms for the degradation of toxic anthropogenic compounds.



KEYWORDS: *activity, enantioselectivity, kinetic modeling, protein and metabolic engineering, synthetic pathway, toxicity*

Downloaded via MASARYK UNIV on September 2, 2021 at 20:20:24 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

Halogenated hydrocarbons, which are often anthropogenic compounds, are widely used for agricultural, industrial and military purposes.[1] Once introduced into the environment, these halogenated hydrocarbons often persist and cause a serious threat to natural ecosystems and human health. Natural catabolic pathways for their mineralization are inefficient or lacking, primarily due to the fact that most of the discussed chemicals were present in the environment until last century. Engineering bacteria toward enhanced biodegradation capacities has been intensively studied; however, only limited success has been achieved so far.[2,3] One of the most common failures is because of an imbalance in the expression or catalytic properties of enzymes employed in synthetic routes. This factor often leads to an accumulation of toxic intermediates, insufficient flux through the pathway, and limited fitness of the host organism.[4−8]

A number of approaches dealing with the unbalanced properties of enzymes in engineered pathways have been reported in recent years. Improvement of pathway performance can be achieved through the introduction of engineered enzyme variants,[9,10] kinetic modeling of the system[11,12] or modular optimization of the pathway.[13−15] Although these examples come predominantly from the field of biosynthesis of valuable chemicals, adoption of such synthetic biology tools holds considerable promise for rational tuning of pathways for biodegradation of industrial waste.[16]

1,2,3-Trichloropropane (TCP) is a man-made compound, produced as a byproduct during the commercial manufacture of epichlorohydrin, and used as an intermediate in chemical industries and as a solvent for oils, fats, waxes, and resins.[17] It is an emerging toxic groundwater pollutant and suspected carcinogen, which spreads to the environment mainly due to improper waste management.[17−19] No naturally occurring bacterial strain capable of aerobic utilization of TCP has been isolated from nature thus far. However, TCP can be converted to harmless glycerol (GLY) via a five-step catabolic pathway assembled with haloalkane dehalogenase (DhaA)[20] from *Rhodococcus rhodochrous* NCIMB 13064, haloalcohol dehalogenase (HheC),[21] and epoxide hydrolase (EchA)[22] from *Agrobacterium radiobacter* AD1 (Figure 1). The first recombinant strain partially degrading TCP was constructed using heterolo-

Special Issue: SB6.0

Received: September 26, 2013
Published: December 6, 2013

 © 2013 American Chemical Society

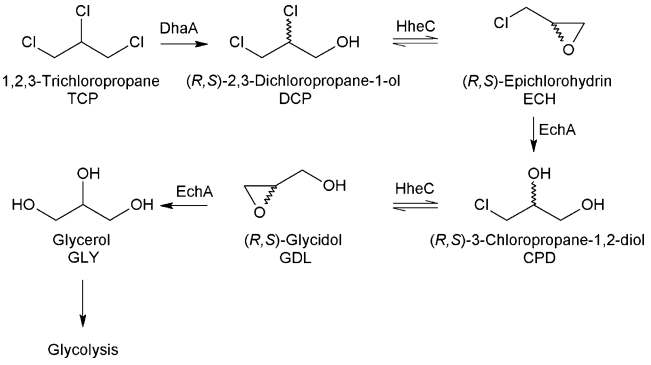


Figure 1. A synthetic pathway for the biodegradation of TCP assembled in *Escherichia coli* BL21 (DE3). Enzyme sources: DhaA[20] from *Rhodococcus rhodochrous* NCIMB 13064, HheC[21] from *Agrobacterium radiobacter* AD1, and EchA[22] from *Agrobacterium radiobacter* AD1.

gous expression of DhaA or an 8-times more active mutant form of DhaAM2 in its natural host *A. radiobacter* AD1, possessing the remainder of the pathway.[23,24] Although some increase in optical density of cultures were observed in the latter case, the efficiency of TCP mineralization was insufficient for supporting significant growth. It was proposed that the poor activity of haloalkane dehalogenase toward TCP, and formation of toxic intermediates represent primary bottlenecks of the pathway. Furthermore, equimolar production of the (*S*) and (*R*) enantiomers of 2,3-dichloropropane-1-ol (DCP) from prochiral TCP by non-selective DhaA and the resulting accumulation of (*S*)-DCP caused by high enantioselectivity of HheC toward (*R*)-DCP was observed and discussed as the factor limiting the flux of carbon through the pathway. The effects of the described bottlenecks in the context of the whole synthetic pathway have not been studied in detail.

Significant efforts have been invested in the past few years in engineering the first enzyme of the TCP pathway: haloalkane dehalogenase (DhaA).[25−28] Constructed mutants can possibly help overcome the previously mentioned limitations of the TCP pathway. Mutant DhaA31, constructed in our laboratory using computer-assisted directed evolution, showed 26-fold higher catalytic efficiency toward TCP than the wild type enzyme (DhaAwt: k_{cat} = 0.04 s^{-1}, k_{cat}/K_m = 40 s^{-1} M^{-1}; DhaA31: k_{cat} = 1.26 s^{-1}, k_{cat}/K_m = 1050 s^{-1} M^{-1}), while its enatioselectivity with the same substrate remained unchanged.[26] Mutant r5−90R (referred here as DhaA90R), obtained after five rounds of directed evolution of DhaA31 by van Leeuwen et al.,[27] converted prochiral TCP predominantly into (*R*)-DCP (ee 90%), which is the preferred substrate for the selective HheC. DhaA90R possessed similar catalytic efficiency toward TCP as the wild-type enzyme (k_{cat} = 0.16 s^{-1}; k_{cat}/K_m = 25 M^{-1} s^{-1}). Thus each of the mutants, DhaA31 and DhaA90R, show improvement in one of the limiting factors of the TCP pathway.

In this study, we carried out a systematic analysis and rational optimization of the synthetic metabolic pathway for the biodegradation of TCP. Both engineered variants of DhaA were integrated to evaluate the effects of increased activity and improved selectivity on the performance of the pathway. A recently developed kinetic model[29] with newly incorporated toxicity and plasmid parameters was employed for rational selection of suitable plasmid combinations and balancing the stoichiometry of the enzyme DhaA/DhaA31/DhaA90R:HheC:EchA. The pathway was optimized toward faster removal of toxic metabolites and higher production of GLY. Several *E. coli* constructs were prepared on the basis of the predictions and characterized for their viability and ability to degrade TCP. Advantages and limitations of the engineered DhaA variants in the context of the pathway were analyzed, and obtained information was used to propose further optimization steps.

RESULTS AND DISCUSSION

The haloalkane dehalogenase DhaA, the haloalcohol dehalogenase HheC, and the epoxide hydrolase EchA were transferred into a heterologous host, *Escherichia coli* BL21 (DE3), for systematic analysis and rational optimization of a synthetic metabolic pathway for the biodegradation of TCP. The three enzymes and five metabolites of the pathway do not naturally occur in *E. coli*, making this system suitable for the analysis of pathway bottlenecks.

Toxicity of TCP and Intermediate Metabolites for *E. coli*. Both toxicity of metabolites and the flux of carbon through the pathway represent the factors dictating the viability of the host organism. Toxicity of TCP and its metabolites toward *E. coli* has not been reported. Thus, we tested the effects of various concentrations of TCP, DCP, epichlorohydrin (ECH), 3-chloropropane-1,2-diol (CPD) and glycidol (GDL) on growth of the *E. coli* BL21 (DE3) culture in mineral medium supplied with glucose. The concentration at which 20% of the total cell population was inhibited (IC_{20}) was calculated for each compound (Figure S1, Supporting Information). Results revealed that TCP and epoxide ECH are the most toxic among all the tested compounds with IC_{20} values of 1.35 and 1.41 mM, respectively (Figure 2). The toxicity of both halogenated

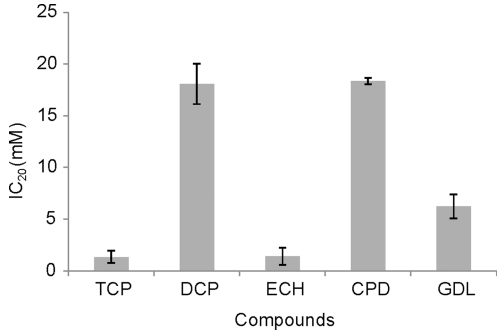


Figure 2. Toxicity of TCP and intermediate metabolites for *E. coli* BL21 (DE3). Toxicity is expressed as inhibition concentration IC_{20}.

alcohols, out of which DCP tends to accumulate in the pathway, was 1 order of magnitude lower. The second epoxide, GDL, had an intermediate toxic effect with an IC_{20} of 6.25 mM. The analysis of TCP, DCP, and GDL accumulation in *E. coli* showed that TCP is the only compound that has the potential for cellular accumulation due to its higher hydrophobicity (Table S1, Supporting Information). The hydrophobicity and accumulation of the other metabolites was significantly lower because of the introduced polar hydroxyl groups. These results suggest that concentrations of metabolites from the TCP pathway detected in the supernatant of the buffer or medium are a good representation of their intracellular concentrations.

Construction of TCP Degraders by the Assembly of an Engineered Pathway in *E. coli*. The TCP pathway was assembled in two modules, separating expression of the first enzyme (DhaA) from expression of the second and third enzymes (HheC and EchA). The two major bottlenecks of the pathway, (i) poor activity of the first enzyme with TCP and (ii)

dx.doi.org/10.1021/sb400147n | *ACS Synth. Biol.* 2014, 3, 172−181

Table 1. Theoretical and Experimental (in Bold) Ratios of Enzymes in the Constructed Degraders

degrader	plasmid combinations	theoretical and experimental ratio DhaA:HheC:EchA	sum of enzymes[a]
degWT	pCDF-dhaAwt + pETDuet-echA-hheC	0.20:0.40:0.40	
		0.28:0.38:0.34	1.00
degWT-opt	pETDuet-dhaAwt + pCDF-echA-hheC	0.50:0.25:0.25	
		0.56:0.25:0.19	0.94
deg90R	pCDF-dhaA90R + pETDuet-echA-hheC	0.20:0.40:0.40	
		0.16:0.40:0.45	1.03
deg90R-opt	pETDuet-dhaA90R + pACYC-echA-hheC	0.67:0.16:0.17	
		0.70:0.12:0.18	0.95
deg31	pCDF-dhaA31 + pETDuet-echA-hheC	0.20:0.40:0.40	
		0.12:0.42:0.46	1.06
deg31-opt	pCDF-dhaA31 + pACYC-echA-hheC	0.50:0.25:0.25	
		0.60:0.16:0.24	0.74

[a]The sum of the three enzymes in each degrader was calculated from the density of corresponding bands on SDS-PAGE gels and compared to the degWT value (set as 1.00).

formation of toxic intermediates, were thus dissected. Modular assembly of the pathway was carried out using combinations of two out of three Duet vectors: pACYCDuet-1, pCDFDuet-1, and pETDuet-1.[30] The Duet vector system has already proven its utility in modular engineering of biosynthetic multienzyme pathways.[15,31−33] Individual Duet vectors can be combined in a single host cell due to the different origins of the replication and antibiotic resistance marker (see Table 2).[34] The modular expression strength can be calculated using the promoter strength and plasmid copy number.[13,15,33,34] The expression levels of subcloned genes can be estimated from the copy numbers of combined plasmids, since all Duet vectors contain the same *T7* promoter and the differences in expression of individual genes from the TCP pathway under this promoter were found to be negligible (Figure S2, Supporting Information).

The starting variants of the TCP pathway were constructed by subcloning *dhaA*wt, *dhaA*90R or *dhaA*31 into the second multiple cloning site of pCDFDuet-1, and *echA* and *hheC* into the first and second multiple cloning sites of pETDuet-1, respectively. This combination should provide an almost equal ratio of subcloned enzymes based on reported copy numbers of used plasmids. The combination of plasmids, consisting of the first enzyme (DhaA) in pCDFDuet-1, and the second and third (HheC and EchA) enzymes in pETDuet-1, was the same in all three variants. Plasmid pairs were cotransformed into *E. coli*, resulting in three constructs denoted degWT, deg90R, and deg31 (Table 1). Degraders were cultivated in LB medium using a standardized cultivation protocol. An approximate estimation of expression profiles obtained from sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis of cell free extract (CFE) showed that the relative ratio of enzymes in the three degraders was close to 0.2:0.4:0.4 (Figure S3, Supporting Information).

Optimization of the ratio of the three enzymes in the TCP pathway was conducted using an extended version of the previously developed kinetic model.[29] The model is based on the Michaelis−Menten steady-state kinetic parameters of DhaA variants, HheC, and EchA, determined with their corresponding substrates *in vitro*. The concentration of each enzyme is kept constant at 0.1 mg/mL. New constraints defining the toxicity of individual metabolites and copy numbers of the used plasmids determining the expression level of enzymes were introduced for optimization of the pathway *in vivo*. The model was used to calculate all possible combinations of the two plasmids within the defined constraints (Tables S2−S4, Supporting Information). Calculated time courses of the modeled multienzyme conversion of 2 mM TCP at a time interval of 300 min were visualized using Gnuplot (Figures S4−S6, Supporting Information). Twelve resulting combinations for each of the pathway variants and the corresponding twelve relative expression ratios were ranked according to (i) the efficiency of GLY production and (ii) the level of the overall toxicity in the system, assuming additivity of the metabolites' toxic effects. Plasmid combinations used in the nonoptimized degraders degWT, deg90R, and deg31 (pCDF-dhaA variant+pETDuet-echA-hheC) were ranked lower compared to highly ranked degraders with optimized expression levels.

A single combination of plasmids (pETDuet-dhaA90R +pACYC-echA-hheC) showed the best rank, leading to the highest GLY production and the lowest toxicity for the degrader employing the DhaA90R variant (Table S3, Supporting Information). Variants with DhaAwt and DhaA31 showed three and four equally good combinations, respectively, from which one combination was selected for experimental construction (Tables S2 and S4, Supporting Information). The plasmid combination pETDuet-dhaAwt+pCDF-echA-hheC was selected for degWT-opt and pCDF-dhaA31+pACYC-echA-hheC for deg31-opt (Table 1). Comparison of the calculated values representing GLY production and overall toxicity showed that both degWT-opt and deg90R-opt produced 2-fold more GLY and less toxic metabolites, in comparison to degWT and deg90R (Tables S2 and S3, Supporting Information). Calculated values for plasmid combinations bearing genes of catalytically efficient DhaA31 plus HheC and EchA showed that this biochemical pathway cannot be effectively optimized for GLY production within the defined constraints (Table S4, Supporting Information), and deg31-opt was expected to provide improvement only in terms of a lower level of toxic metabolites.

Experimental Characterization of the TCP Degraders. The resting cells of six degraders were experimentally characterized in terms of their (i) expression profile, (ii) viability, and (iii) degradation capacity. The degraders were cultivated in LB medium using a standardized cultivation protocol with heterologous protein expression induced by isopropyl-β-D-thiogalactopyranoside (IPTG). The degraders with overexpressed enzymes of the TCP pathway were harvested and disintegrated using sonication. Equal amounts of CFEs were loaded on SDS-PAGE, and the expression profiles of six degraders were analyzed by densitometry (Figure 3). Relative ratios of DhaA variants, HheC and EchA were calculated from

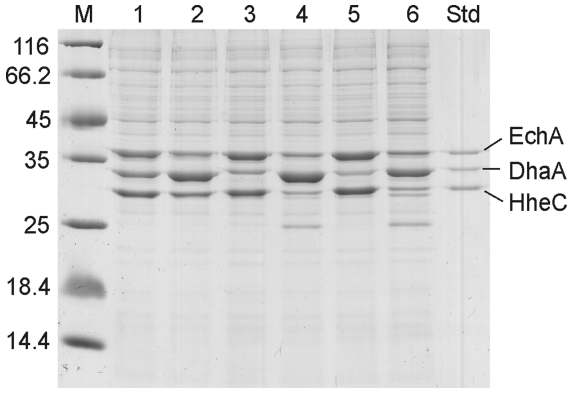


Figure 3. SDS-PAGE analysis of the expression profiles of the six constructed degraders. M: protein marker (units in kDa), 1: degWT, 2: degWT-opt, 3: deg90R, 4: deg90R-opt, 5: deg31, 6: deg31-opt, and Std: standard sample with purified DhaAwt, HheC, and EchA containing 0.25 μg of each enzyme. The theoretical molecular weights of the DhaA variants, HheC and Echa are 34, 29, and 35 kDa, respectively.

band densities and compared with the theoretical relative ratios predicted by modeling (Table 1). An excellent agreement between theoretical and experimental values was observed. The relative sum of the three enzymes from the TCP pathway was calculated as a sum of the band densities and revealed that degraders with an optimized ratio contained equal or lower total amounts of the enzymes than nonoptimized ones.

The viability and ability to degrade TCP was tested with resting cells of the degraders resuspended in synthetic mineral medium (SMM) with a final OD_{600} of 0.1. Cells were incubated at 30 °C with 2 mM TCP and no additional carbon source. Time course of TCP and DCP concentrations was monitored by GC analysis over 120 h (Figure 4A–C). No significant accumulation of other metabolites of the TCP pathway was observed. Viabilities of the six degraders and host with and without TCP as controls are presented in Figure 4D. No GLY was detected in cell cultures suggesting its utilization by the cells. The production of GLY via the TCP pathway was confirmed in parallel by incubation of induced cells of deg31 and deg31-opt in the presence of 2 mM TCP and 5 mM glucose (Figure S7,

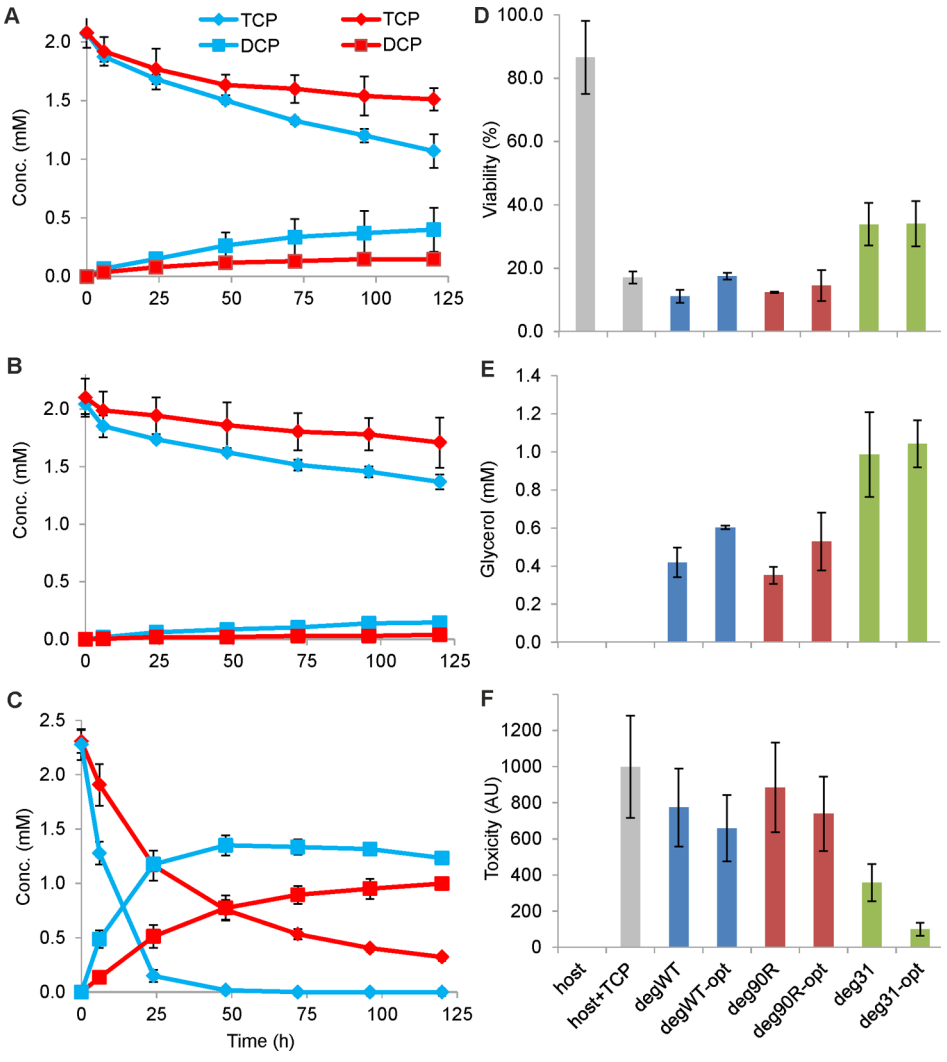


Figure 4. Degradation profiles and viability of the six degraders. (A–C) Degradation profiles of the degraders containing DhaAwt, DhaA90R, and DhaA31, respectively. The diamonds and squares refer to concentrations of TCP and DCP of nonoptimized (red) and optimized (blue) variants. (D) Viability of the host and degraders after 120 h incubation in SMM medium with TCP, relative to the viability at the beginning of the experiment. (E) GLY production calculated from concentrations of TCP and DCP detected at the end of incubation. (F) Overall toxicity of metabolites to degraders presented in arbitrary units (AU). Standard deviations were obtained from three independent experiments.

dx.doi.org/10.1021/sb400147n | *ACS Synth. Biol.* 2014, 3, 172–181

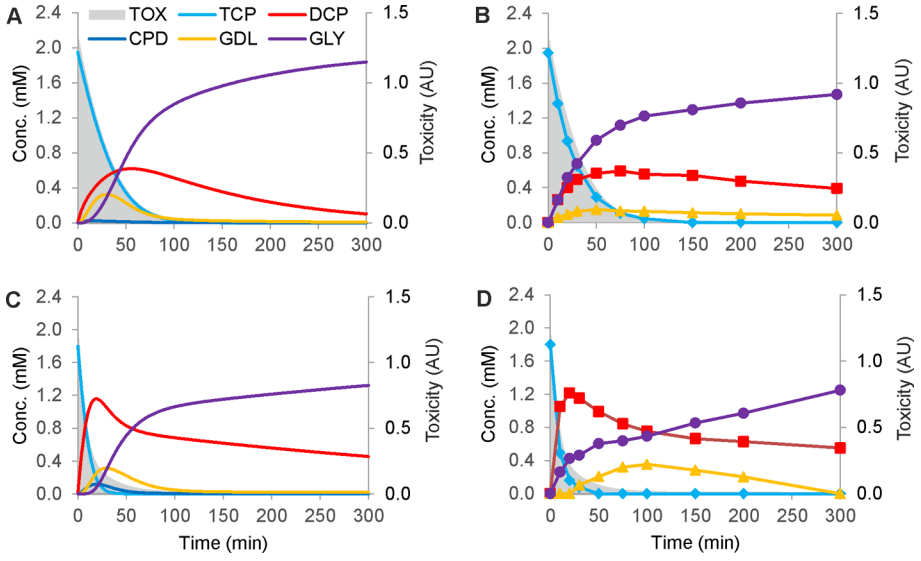


Figure 5. Predicted and experimentally determined degradation profiles of constructs containing the engineered haloalkane dehalogenase DhaA31. Theoretical (A) and experimental (B) profiles of deg31; theoretical (C) and experimental (D) profiles of deg31-opt. Toxicity is denoted as the gray area under the graph and given in arbitrary units (AU). Concentrations of GLY were calculated from concentrations of other detected metabolites. Experimental data points are the mean values of two replicates.

Supporting Information). Theoretical concentrations of produced GLY in cultures without glucose and relative toxicity for each degrader were calculated from recorded time course concentrations of TCP and DCP (Figure 4E,F).

The averaged data from the three independent experiments revealed that all three degraders with rationally optimized ratios of pathway enzymes degraded TCP faster than nonoptimized constructs. Experimental time courses of TCP and DCP concentrations correlated with the relative ratios of enzymes in the degraders. Differences in profiles corresponded to the catalytic parameters of the DhaA variant and the relative ratio of haloalkane dehalogenase with respect to the other two associated enzymes (Figure 4A–C and Table 1). Deg90R and deg90R-opt with selective DhaA90R showed minimal accumulation of DCP, but their degradation potential was limited by a low catalytic efficiency with TCP. Deg31 and deg31-opt, both containing DhaA31, the nonselective variant with improved catalytic efficiency toward TCP, showed the fastest conversion of TCP, but also the most significant accumulation of DCP. These two effects were even more pronounced in the deg31-opt with higher expression of DhaA31.

Viability reflects the physiological state of the cells heterologously expressing the synthetic pathway after their exposure to toxic TCP (Figure 4D). DegWT-opt showed a small but statistically significant difference in cell viability compared to degWT, suggesting that optimization of the expression levels provided benefits to the cells. The difference in viability of deg90R and deg90R-opt was statistically insignificant: the low production of GLY was not balanced by a higher selectivity of DhaA90R and reduced accumulation of DCP. The most striking result was the superior viability of both degraders containing the DhaA31 enzyme, connected to a higher production of GLY and/or reduced toxicity compared to the degraders expressing DhaAwt and DhaA90R. Deg31-opt removes the toxic metabolites faster than deg31, but the effect of the reduced toxicity on the viability of this degrader is not visible, suggesting that both degraders surpassed some threshold level beyond which toxicity does not play a role and only the amount of produced GLY is the key factor determining cell viability. Further improvement in viability should be achieved by a more efficient production of GLY. Approximately 1 mM concentration of GLY produced within 5 days by the best variants, deg31 and deg31-opt, is probably not sufficient to provide energy for growth and compensation for the oxidative stress induced during the aerobic mineralization of toxic chlorinated substrates.[35]

The ability of the degraders to recover from short-term (300 min) exposure to a higher concentration (3.5 mM) of TCP was tested (Figure S8, Supporting Information). Resulting growth curves corresponded well with the viabilities of degraders calculated from plating after 120 h incubation with a lower concentration of TCP. Only deg31 and deg31-opt recovered faster than the host BL21 (DE3) without the synthetic pathway (control) and with comparable growth. Degraders degWT and deg90R, with a nonoptimized ratio of pathway enzymes, and deg90R-opt, with an optimized ratio but poor conversion of TCP, recovered slower than the control. Only degWT-opt showed comparable growth with host cells without the introduced pathway, again confirming the benefits of pathway optimization. The data confirm that the constructs containing DhaA31 are not significantly affected by an increased metabolic burden.[36]

Identification of Pathway Bottlenecks. Two constructs carrying DhaA31 and showing the highest viability were selected for further characterization. Preinduced cells of elevated cell densities (OD_{600} 3.5) were used to secure complete degradation of 2 mM TCP during a shorter incubation time (300 min). The mass ratio of enzymes in both degraders was estimated by SDS-PAGE and verified with measurements of catalytic activities in CFE (Table S5, Supporting Information). Obtained data were used for determination of the enzyme concentrations and applied to calculations of the degradation profiles of deg31 and deg31-opt. In both cases the experimental profiles showed a good agreement with predictions (Figure 5). Accumulation of DCP and GDL was observed for both constructs. A more pronounced accumulation of DCP was observed for deg31-opt, with faster removal of TCP and a corresponding lower toxicity of the system. The accumulated DCP was analyzed by chiral GC and found to be mainly composed of the (*S*)-enantiomer (Figure S7

dx.doi.org/10.1021/sb400147n | *ACS Synth. Biol.* 2014, 3, 172–181

Supplementary Information). Theoretical conversions of TCP to GLY at the end of incubation calculated from experimental concentrations of TCP at time 0 min and concentrations of the accumulated metabolites at 300 min was similar for both deg31 and deg31-opt (75 and 69%). This is in agreement with the data previously collected over 120 h for conversion of 2 mM TCP and suggests that the degraders containing DhaA31 cannot be further optimized for better production of GLY without further engineering of DhaA properties (Table S4, Supporting Information).

Currently available DhaA variants with increased activity[26] and improved enantioselectivity[27] provided only partial improvement of the overall efficiency of TCP degradation, even after optimization in the ratio of enzymes of the synthetic pathway. The level of GLY production required for cellular maintenance and growth of the degrader continued to be limited by insufficient selectivity of the DhaA variant with enhanced activity, and *vice versa* by insufficient activity of the variant with increased enantioselectivity. Following these findings, we applied our mathematical model to estimate the effect of catalytic activity and enantioselectivity on the production of GLY (Figure 6). We

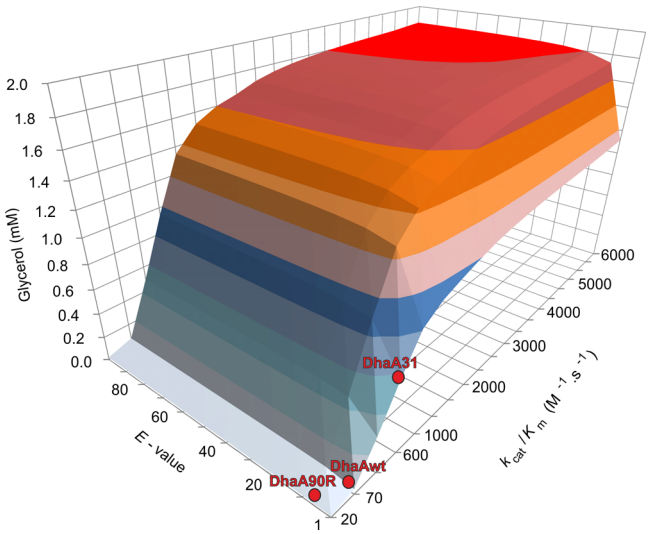


Figure 6. Hypersurface plot describing the effect of catalytic efficiency (k_{cat}/K_m) and enantioselectivity (*E*-value) of DhaA on the production of GLY. The positions of the three variants of DhaA are indicated by red dots. The substrate TCP is supplied at 2 mM. The kinetic constants of DhaAwt, DhaAR90, and DhaA31 were experimentally determined[29] and differ slightly from data reported in the literature.[26,27]

measured the minimum concentration of GLY required for cellular maintenance and observable growth of the host cells for OD_{600} 0.1 and determined that at least 1 mM of GLY should be supplied to the culture within the time interval of 24 h (Figure S9, Supporting Information). The amount of GLY required for maintenance and growth could be higher in a culture containing toxic substances like TCP and intermediate metabolites due to the extra burden caused by their toxicity. The theoretical production of GLY by the culture of the best degrader deg31-opt with OD_{600} of 0.1 was calculated to be 0.6 mM per 24 h, suggesting that DhaA31 does not possess the required catalytic properties. To increase GLY production to 1 mM per 24 h, the model suggested the need for further improvement of the catalytic efficiency of DhaA31 by 4-fold, or improvement of the enantioselectivity by 20-fold (Figure S10, Supporting Information). Alternatively, the employment of a DhaA variant with combined improvement of catalytic efficiency by 1.2-fold and enantioselectivity by 10-fold in favor of production of (*R*)-DCP would produce 1 mM GLY within 24 h. A DhaA variant possessing the properties enabling the highest possible pathway efficiency (1.6 mM GLY per 24 h) should have 4-fold improvement in catalytic efficiency and 20-fold higher enantioselectivity compared to DhaA31.

Conclusions. The toxicity of the substrate/metabolites and insufficient carbon/energy flow can limit viability of the host organism and represent possible bottlenecks in the TCP degradation pathway. The most toxic substances within the pathway are TCP (IC_{20} = 1.35 mM) and ECH (IC_{20} = 1.41 mM). Other intermediates of the pathway are significantly less toxic. Since accumulation of ECH inside the cells does not take place, the hydrophobic substrate TCP represents the single most important toxic substance of the pathway. A concentration of GLY ≥ 1 mM supports cellular maintenance and growth of *E. coli* BL21 (DE3) with an initial OD_{600} 0.1.

A kinetic model was employed to improve the TCP degradation pathway using engineered enzyme variants and balanced enzyme ratios. Predictions were evaluated in terms of minimized toxicity of intermediates and maximized production of GLY, and showed excellent correspondence with experimental data. Expression levels of the enzymes can also be reliably predicted from a combination of plasmids with different copy numbers. A statistically significant increase in the viability of degWT-opt compared to degWT demonstrates that optimization of the enzyme ratio is a useful approach for improvement of overall pathway performance, and that the kinetic model employing *in vitro* measured kinetic parameters of individual enzymes is applicable for rational pathway engineering. The model can be used for the design of constructs with optimized expression levels and prediction of minimum requirements for catalyst properties.

Toxicity of the substrate does not limit viability of the cells employing the highly active variant of the first enzyme, haloalkane dehalogenase DhaA31, in the pathway. This is due to rapid conversion of the highly toxic TCP to the less toxic DCP. However, the growth and viability of the cells are limited by the insufficient production of GLY. Production of GLY can be improved by further engineering the first enzyme of the pathway toward a higher catalytic efficiency (k_{cat}/K_m > 2300 s^{-1} M^{-1}) or higher enantioselectivy (*E*-value > 20), or by a combination of both properties (k_{cat}/K_m > 700 s^{-1} M^{-1} and *E*-value > 10). Removal of the enantioselectivity of the second enzyme in the pathway, haloalcohol dehalogenase HheC, represents another option for improvement of the carbon/energy flow. Construction of a new generation of catalysts for degradation of TCP is currently ongoing in our laboratory.

METHODS

Chemicals, Media, Strains and Plasmids. TCP, DCP, ECH, CPD, GDL and GLY standards were purchased from Sigma-Aldrich (USA). All chemicals used in this study were of analytical grade. All restriction enzymes and DNA ligase were purchased from New England Biolabs (USA). A free Glycerol Assay Kit was acquired from BioVision (USA). Luria Broth (LB) (Sigma Aldrich, USA) was used for routine cultures. A synthetic mineral medium (SMM)[37] containing 5.4 g of $Na_2HPO_4 \cdot 12H_2O$, 1.4 g of KH_2PO_4, 0.5 g of $(NH_4)_2SO_4$, 0.2 g of $MgSO_4 \cdot 7H_2O$, 2 mL of trace elements solution,[38] and 1 mL of vitamin B1 (10 g/mL) per 1 L was used for selection experiments. M9 minimal medium (Sigma Aldrich, USA) containing 0.2 g of

dx.doi.org/10.1021/sb400147n | *ACS Synth. Biol.* 2014, 3, 172−181

Table 2. Plasmids Used and Recombinants Constructed

duet vectors[a]	origin	copy no.	cloned gene	recombinant plasmids	restriction sites
pACYCDuet-1[b]	P15A	10−12	*echA, hheC*	pACYC-*echA-hheC*	*echA* (*Nco*I/*Hind*III), *hheC* (*Nde*I/*Kpn*I)
pCDFDuet-1[c]	CloDF13	20−40	*dhaAwt*	pCDF-*dhaA*wt	*Nde*I/*Kpn*I
			dhaA31	pCDF-*dhaA*31	*Nde*I/*Xho*I
			dhaA90R	pCDF-*dhaA*90R	*Nde*I/*Kpn*I
			echA, hheC	pCDF-*echA-hheC*	*echA* (*Nco*I/*Hind*III), *hheC* (*Nde*I/*Kpn*I)
pETDuet-1[d]	ColE1	40	*dhaAwt*	pETDuet-*dhaA*wt	(*Nde*I/*Kpn*I)
			echA, hheC	pETDuet-*echA-hheC*	*echA* (*Nco*I/*Hind*III), *hheC* (*Nde*I/*Kpn*I)
			dhaA90R	pETDuet-*dhaA*90R	(*Nde*I/*Kpn*I)

[a]Source: Novagen. [b]CmR: chloramphenicol resistance. [c]SmR: streptomycin/spectinomycin resistance. [d]AmpR: ampicillin resistance.

$MgSO_4 \cdot 7H_2O$ and 2 mL of trace elements per 1 L and 10 mM glucose were used for toxicity tests. *Escherichia coli* DH5α (Life Technologies, USA) was used in cloning and plasmid propagation. *E. coli* BL21 (DE3) (Life Technologies, USA) was used as a heterologous host for expression of the synthetic pathway for the biodegradation of TCP. Plasmids pET21b, pACYCDuet-1, pETDuet-1, pCDFDuet-1 (Novagen, Germany) were used for subcloning and modular construction of pathway variants.

Molecular Techniques and Culture Conditions. The genes of the haloalcohol dehalogenase encoding HheC[21] and the epoxide hydrolase encoding EchA[22] from *Agrobacterium radiobacter* AD1, together with genes of the wild-type haloalkane dehalogenase from *Rhodococcus rhodochrous* NCIMB 13064 encoding DhaAwt[20] and the mutants DhaA90R[27] and DhaA31[26] were commercially synthesized (GeneArt, Germany). A tag sequence of six histidine codons was attached downstream from the gene in all cases except for *hheC*. Sequences of all genes except for *hheC* were codon-optimized for expression in *E. coli* during gene synthesis. Synthetic genes were subcloned into the *Nde*I and *Bam*HI restriction sites of pET21b. An alternative *Nco*I restriction site was introduced at the beginning of the *echA* gene to enable cloning into the first multiple cloning sites of Duet vectors. Upon introduction of the *NcoI* site, the second codon of the *echA* gene, ACT encoding threonine, was substituted for GCA encoding alanine. The constructs pET21b-*dhaA*, pET21b-*dhaA*31, pET21b-*dhaA*90R, pET21b-*hheC*, and pET21b-*echA* were used for transformation of competent cells of *E. coli* DH5α using the heat-shock method for plasmid propagation. Isolated plasmids were used for transformation of competent cells of *E. coli* BL21 (DE3) for evaluation of individual gene expression under similar conditions. Subcloning of the genes coding for the synthetic pathway into Duet plasmids is summarized in Table 2. *E. coli* BL21 (DE3) cotransformants, prepared by modular combination of recombinant Duet vectors are summarized in Table 1. All plasmid constructs were verified by sequencing (GATC, Germany).

Precultures of *E. coli* DH5α, BL21 (DE3) host cells, and cotransformants were prepared by growth on LB medium with or without antibiotics at 37 °C overnight. Final concentrations of respective antibiotics (ampicillin 100 μg/mL, chloramphenicol 34 μg/mL, streptomycin 50 μg/mL) were used in the cultures with cells containing a single plasmid. Half concentrations of two relevant antibiotics were used in precultures and cultures of cotransformants. A standard protocol was developed for the cultivation of the degraders. Precultures were used to inoculate fresh LB medium, and cultures were grown at 37 °C until the cell density reached 1 at OD_{600}. Expression of recombinant enzymes was induced by 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and cultivation continued at 20 °C. Cells were harvested during the late exponential phase by centrifugation (5000*g*, 15 min, 4 °C), washed three times with sterile ice-cold SMM, M9 medium or 50 mM sodium phosphate buffer (pH 7.0), and used in further experiments.

Analytical Techniques. A gas chromatograph 6890N with a flame ionization detector (GC-FID) and mass spectrometer (GC-MS) 5975C MSD (Agilent Technologies, USA), with the capillary column ZB-FFAP 30 m × 0.25 mm × 0.25 μm (Phenomenex, USA) were used for routine analysis and quantification of TCP and its metabolites. The separation method for both GC-FID and GC-MS used an inlet temperature of 250 °C, split ratio 20:1, helium carrier gas with an initial flow of 0.6 mL/min for 1 min, followed by a flow gradient of 0.2 mL/min/min from 0.6 to 1.8 mL/min, and an oven temperature program set initially to 50 °C for 1 min, followed by a temperature gradient of 25 °C/min from 50 to 220 °C with holding for 2 min.

Determination of Toxicity of the Substrate and Intermediate Metabolites. A growth test in M9 medium containing 10 mM glucose was used to determine the toxicity of TCP and its intermediates. The growth of *E. coli* cultures was evaluated at 37 °C for 6 h using 0.5, 1, 2, 4, 5, 10, 15, and 20 mM of TCP, DCP, ECH, CPD, and GLD. To avoid the evaporation of volatile compounds, cultivation was done in 25 mL glass vials with a screw cap mininert valve (Sigma-Aldrich, USA). The concentration of TCP and intermediate compounds was monitored by GC. Samples (0.5 mL) were withdrawn and extracted with acetone (1:1), containing hexan-1-ol as an internal standard and centrifuged for 2 min at 18000*g*. The acetone extracts (2 μL) were injected directly into GC. *E. coli* cultures without the addition of the tested compounds were used as a negative control. For growth monitoring, 1 mL samples were withdrawn in 1 h intervals and measured at OD_{600}. Toxicity data were fit into polynomial equations from which the inhibitory concentration of individual compounds (IC_{20} value) was determined.

Determination of Expression Levels of the Enzymes of the TCP Pathway. The expression levels of DhaAwt, DhaA31, DhaA90R, HheC, and EchA were determined in *E. coli* BL21 (DE3) cells transformed with a pET21b vector subcloned with corresponding genes, as well as in the degraders prepared by standardized cultivation procedures. Washed cells were resuspended in 10 mL of 50 mM sodium phosphate buffer, and cell density was adjusted to 3.5 at OD_{600}. 1 U of DNaseI per 1 mL of cell suspension was added. Cells were disrupted with 5 cycles of sonication using a Hielscher UP200S (Teltow, Germany) ultrasonic processor with 0.3 s pulses and an amplitude 85%. Each cycle consisted of 5 min sonication followed by 5 min cooling at 4 °C. The cell lysate was centrifuged for 1 h at 18000*g* at 4 °C, and the resulting cell-free extract (CFE)

was decanted. The concentration of total protein in CFEs was determined using Bradford reagent (Sigma Aldrich, USA). Samples of CFE containing 5 μg of total protein were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). CFE prepared from *E. coli* BL21 (DE3) cells without plasmids was used as a control. For determination of the mass ratio, the method was calibrated with standard samples containing 0.25, 0.5, 1, 1.5, and 2 μg of each enzyme DhaA, EchA and HheC in purified form. Purification of DhaA, DhaA31, DhaA90R, HheC and EchA has been described elsewhere.[29] Gels were stained with Coomassie Brilliant Blue R-250 (Fluka, Switzerland) and analyzed using a GS-800 Calibrated Imaging Densitometer (Bio-Rad, USA). The amounts of DhaAwt, DhaA90R, DhaA31, HheC and EchA in samples of CFE applied to the gel and the corresponding relative and mass ratios of the enzymes in the individual degraders were estimated from trace densities of selected bands.

Verification of Expression Levels by Activity Measurements. The expression levels of DhaAwt, DhaA31, HheC, and EchA estimated from SDS-PAGE analysis of CFE samples using deg31 and deg31-opt were verified by the measurement of individual enzyme activity in CFE. The activity of DhaA, HheC, and EchA was estimated using 10 mM TCP, 20 mM DCP, and 10 mM ECH, respectively. Substrates were dissolved in 10 mL of 50 mM Tris-SO_4 buffer (pH 8.5) at 37 °C. The reaction was initiated with the addition of CFE to a volume corresponding to 1 mg of DhaA, 1 mg of HheC or 1 mg of EchA, as was estimated from the expression levels analyzed by SDS-PAGE. CFE from *E. coli* BL21 (DE3) without plasmids was used as a negative control. Samples (0.5 mL) of the reaction mixtures were withdrawn at certain time intervals and mixed with a 0.5 mL solution of acetone in hexan-1-ol and analyzed by GC. Specific activity was calculated as a decrease in substrate concentration in μmol per min per mg of the enzyme.

Mathematical Modeling and Optimization of the TCP Pathway. The mathematical model describing the TCP pathway was constructed using kinetic equations and parameters obtained from *in vitro* enzyme kinetics with the purified enzymes and corresponding substrates (TCP, DCP, ECH, CPD, and GDL). All kinetic experiments were performed in 50 mM Tris-SO_4 buffer at pH 8.5 and 37 °C. Determination of kinetic parameters and construction of the model and its experimental verification for *in vitro* conditions is described elsewhere.[29] Initial parameters for pathway optimization were as follows: starting concentration of TCP in 10 mL reaction, 2 mM; time interval of multienzyme reaction, 300 min; total mass of enzymes in the pathway, 3 mg. Allowed combinations of Duet vectors were prepared using the following criteria: only two out of three employed Duet vector derivatives (pACYC, pCDF and pETDuet) can be combined; either one or two genes coding for enzymes from the TCP pathway can be expressed from one vector; only the *echA* gene possesses the *Nco*I restriction site for subcloning into the first multiple cloning site of Duet vectors, thus it can be expressed only from a vector with another subcloned gene and not separately. The relative ratio of enzymes in a simulation was then determined by a particular combination of plasmids carrying individual enzymes and a copy number of individual plasmids (10 for pACYCDuet-1, 20 for pCDFDuet-1, and 40 for pETDuet-1). For all allowed combinations of plasmids, dynamic simulations of the multienzyme system based on a series of Michaelis−Menten equations were performed using the Python 2.7 programming language. The differential form of equations was integrated using Euler's method with a fixed step size of 0.05 min. The toxicity effect at individual times along degradation profiles was calculated using previously obtained polynomial equations fitted to experimental toxicity data and the actual concentration of each compound at a given time. The overall toxicity effect was calculated by integration of individual effects along the profile.

Multienzyme reactions including variants of the DhaA enzyme with modified catalytic properties were modeled using the following adjustments of the previously described process. The total mass of enzymes was set to 0.1 mg, the time-interval of the multienzyme reaction was set to 24 h. A modification of the catalytic performance of DhaA was as follows: K_M was set to 1 mM, while the overall k_{cat} was varied across the range of 1 to 1000 min^{-1}. The enantioselectivity of DhaA expressed as a ratio of catalytic rate constants leading toward production of individual enantiomers of DCP was assigned for values ranging from 1 to 200. To avoid the limitation of the relative ratio of enzymes in the simulation by the employed plasmids, the mass of individual enzymes were assigned a value from 0.0 to 0.1 mg with intervals of 0.0125 mg. The ratio resulting in the highest end-point production of GLY was considered as representative of the optimized pathway.

Characterization of the Degraders in Minimal Medium with TCP. Preinduced cells of the six constructed degraders (Table 1) were washed with sterile ice cold SMM medium and resuspended in the same medium. A final concentration of 2 mM TCP was added to 15 mL of sterile SMM in 25 mL glass vials with a screw cap mininert valve (Sigma-Aldrich, USA) and incubated at 30 °C with shaking for 2 h to allow for complete dissolution. Washed cells were adjusted to a final cell density 0.1 at OD_{600}. Vials containing BL21 (DE3) cells with and without TCP were used as controls. Flasks were continuously incubated for 5 days in a shaking incubator NB-205 (N-Biotek, South Korea) at 200 rpm and 30 °C. Samples for analysis of TCP and metabolites were withdrawn at 0, 6, 24, 48, 72, 96, and 120 h. Viability of cells at the beginning and end of their incubation with TCP was tested by plating a sample of cell suspension diluted 2×10^4-times with ice-cold sterile SMM onto LB agar plates. Plates were incubated overnight at 37 °C, and grown colonies were counted using the Colony Picker CP7200 (Norgren Systems, USA).

Recovery of Degraders in LB Medium. Preinduced cells of the six constructed degraders were incubated at 37 °C with shaking in 25 mL glass vials with a screw cap mininert valve containing 10 mL of 50 mM sodium phosphate buffer (pH 7.0) with 3.5 mM TCP. The OD_{600} of cell suspensions was 3.5. BL21 (DE3) cells without plasmids were used as controls. After 5 h of incubation, cell suspension samples (10 μL) were transferred into wells of a 96-well microtiter plate containing 100 μL of LB medium. The plate was incubated at 37 °C with shaking (900 rpm) in a shaking incubator (PST-100 HL Thermo Shaker, BIOSAN). Cell growth was monitored by measuring OD_{600} at 30 min intervals using a microtiter-plate reader (Tecan, Switzerland).

Degradation of TCP in Buffer by Preinduced Resting Cells. Cell suspensions of preinduced deg31, deg31-opt and *E. coli* BL21 (DE3) without plasmids (negative control) were diluted with sterile 50 mM sodium phosphate buffer to a final OD_{600} 7. Glass vials (25 mL) with a screw cap mininert valve containing 7.5 mL of the sterile 50 mM sodium phosphate buffer of pH 7.0 and 4 mM TCP were prepared separately and incubated for 1 h at 37 °C with shaking. The reaction was initiated by mixing 7.5 mL of the cell suspension with 7.5 mL of buffer and dissolved TCP. The final concentration of TCP in 15

dx.doi.org/10.1021/sb400147n | *ACS Synth. Biol.* 2014, 3, 172−181

mL of the cell suspension was 2 mM and the final theoretical OD_{600} was 3.5. Cell suspension samples (0.5 mL) were quenched in 0.5 mL of acetone with hexan-1-ol, vortexed, and centrifuged at 18000*g* for 2 min. The concentration of metabolites in the supernatant was analyzed using GC. The presence of GLY in the cell suspension was verified using a Free Glycerol Assay Kit (BioVision, USA).

■ ASSOCIATED CONTENT

Supporting Information

This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*E-mail: zbynek@chemi.muni.cz.

*E-mail: jiri@chemi.muni.cz.

Author Contributions

⊥N. P. Kurumbang and P. Dvorak contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This research was funded by the SoMoPro Program (Incoming Grant SYNTBIO No. 2SGA2873 to N.P.K.) and The Grant Agency of the Czech Republic (P503/12/0572). Research leading to these results has received financial contribution from the European Community within the Seventh Framework Program (FP/2007-2013) under Grant Agreement No. 229603. The research has also been cofinanced by the South Moravian Region. Additional funding was obtained from the European Regional Development Fund (CZ.1.05/2.1.00/01.0001, CZ.1.05/1.1.00/02.0123) and the Ministry of Education of the Czech Republic (CZ.1.07/2.3.00/30.0037, CZ.1.07/2.3.00/20.0239). The work of J.B. was supported by the "Employment of the Best Young Scientists for International Cooperation Empowerment" Program (CZ1.07/2.3.00/30.0037) and cofinanced through the European Social Fund and the state budget of the Czech Republic. CERIT-SC is acknowledged for providing access to their computing facilities under the program Center CERIT scientific Cloud (CZ.1.05/3.2.00/08.0144).

■ REFERENCES

(1) Bhatt, P., Kumar, S. M., Mudliar, S., and Chakrabarti, T. (2007) Biodegradation of chlorinated compounds—a review. *Crit. Rev. Environ. Sci. Technol. 37*, 165−198.

(2) de Lorenzo, V. (2009) Recombinant bacteria for environmental release: what went wrong and what we have learnt from it. *Clin. Microbiol. Infect. 1*, 63−65.

(3) Wittich, R. M., van Dillewijn, P., and Ramos, J. L. (2010) Rational construction of bacterial strains with new/improved catabolic capabilities for the efficient breakdown of environmental pollutants, in *Handbook of Hydrocarbon and Lipid Microbiology* (Timmis, K. N., Ed.), pp 1247−1254, Springer-Verlag, Berlin.

(4) Haro, M. A., and de Lorenzo, V. (2001) Metabolic engineering of bacteria for environmental applications: construction of *Pseudomonas* strains for biodegradation of 2-chlorotoluene. *J. Biotechnol. 85*, 103−113.

(5) Cámara, B., Herrera, C., González, M., Couve, E., Hofer, B., and Seeger, M. (2004) From PCBs to highly toxic metabolites by the biphenyl pathway. *Environ. Microbiol. 6*, 842−850.

(6) Pollmann, K., Wray, V., and Pieper, D. H. (2005) Chloromethylmuconolactones as critical metabolites in the degradation of chloromethylcatechols: recalcitrance of 2-chlorotoluene. *J. Bacteriol. 187*, 2332−2340.

(7) Martínez, P., Agulló, L., Hernández, M., and Seeger, M. (2007) Chlorobenzoate inhibits growth and induces stress proteins in the PCB-degrading bacterium *Burkholderia xenovorans* LB400. *Arch. Microbiol. 188*, 289−297.

(8) de la Peña Mattozzi, M., Tehara, S. K., Hong, T., and Keasling, J. D. (2006) Mineralization of paraoxon and its use as a sole C and P source by a rationally designed catabolic pathway in *Pseudomonas putida*. *Appl. Environ. Microbiol. 72*, 6699−6706.

(9) Leonard, E., Ajikumar, P. K., Thayer, K., Xiao, W. H., Mo, J. D., Tidor, B., Stephanopoulos, G., and Prather, K. L. (2010) Combining metabolic and protein engineering of a terpenoid biosynthetic pathway for overproduction and selectivity control. *Proc. Natl. Acad. Sci. U. S. A. 107*, 13654−13659.

(10) Zhang, K., Li, H., Cho, K. M., and Liao, J. C. (2010) Expanding metabolism for total biosynthesis of the nonnatural amino acid L-homoalanine. *Proc. Natl. Acad. Sci. U. S. A. 107*, 6234−6239.

(11) Bujara, M., Schümperli, M., Pellaux, R., Heinemann, M., and Panke, S. (2011) Optimization of a blueprint for *in vitro* glycolysis by metabolic real-time analysis. *Nat. Chem. Biol. 7*, 271−277.

(12) Li, R. D., Li, Y. Y., Lu, L. Y., Ren, C., Li, Y. X., and Liu, L. (2011) An improved kinetic model for the acetone-butanol-ethanol pathway of *Clostridium acetobutylicum* and model-based perturbation analysis. BMC. *Syst. Biol. 5*, S12.

(13) Ajikumar, P. K., Xiao, W. H., Tyo, K. E, Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B., and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in *Escherichia coli*. *Science 330*, 70−74.

(14) Xu, P., Gu, Q., Wang, W., Wong, L., Bower, A. G., Collins, C. H., and Koffas, M. A. (2013) Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nat. Commun. 4*, 1409.

(15) Wu, J., Du, G., Zhou, J., and Chen, J. (2013) Metabolic engineering of *Escherichia coli* for (2S)-pinocembrin production from glucose by a modular metabolic strategy. *Metab. Eng. 16*, 48−55.

(16) Copley, S. D. (2009) Evolution of efficient pathways for degradation of anthropogenic chemicals. *Nat. Chem. Biol. 5*, 559−566.

(17) Samin, G., and Janssen, D. B. (2012) Transformation and biodegradation of 1,2,3-trichloropropane (TCP). *Environ. Sci. Pollut. Res. Int. 19*, 3067−3078.

(18) Kielhorn, J., Könnecker, G., Pohlenz-Michel, C., Schmidt, S., and Mangelsdorf, I. (2003) *Concise International Chemical Assessment Document 56: 1,2,3-Trichloropropane*, WHO, Geneva.

(19) Gehlhaus, M., Foster, S., Hogan, K., and Holdsworth, G. (2009) *Toxicological Review of 1,2,3-Trichloropropane*, EPA/635/R-08/010F, US EPA, Washington, D.C.

(20) Kulakova, A. N., Larkin, M. J., and Kulakov, L. A. (1997) The plasmid-located haloalkane dehalogenase gene from *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology. 143*, 109−115.

(21) van Hylckama Vlieg, J. E., Tang, L., Lutje Spelberg, J. H., Smilda, T., Poelarends, G. J., Bosma, T., van Merode, A. E., Fraaije, M. W., and Janssen, D. B. (2001) Halohydrin dehalogenases are structurally and mechanistically related to short-chain dehydrogenases/reductases. *J. Bacteriol. 183*, 5058−5066.

(22) Rink, R., Fennema, M., Smids, M., Dehmel, U., and Janssen, D. B. (1997) Primary structure and catalytic mechanism of the epoxide hydrolase from *Agrobacterium radiobacter* AD1. *J. Biol. Chem. 272*, 14650−14657.

(23) Bosma, T., Kruizinga, E., de Bruin, E. J, Poelarends, G. J, and Janssen, D. B. (1999) Utilization of trihalogenated propanes by *Agrobacterium radiobacter* AD1 through heterologous expression of the haloalkane dehalogenase from *Rhodococcus* sp. strain M15-3. *Appl. Environ. Microbiol. 65*, 4575−4581.

(24) Bosma, T., Damborsky, J., Stucki, G., and Janssen, D. B. (2002) Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. *Appl. Environ. Microbiol. 68*, 3582−3587.

(25) Gray, K., Richardson, T., Kretz, K., Short, J., Bartnek, F., Knowles, R., Kan, L., Swanson, P., and Robertson, D. (2001) Rapid evolution of

reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. *Adv. Synth. Catal. 343*, 607−617.
(26) Pavlova, M., Klvana, M., Prokop, Z., Chaloupkova, R., Banas, P., Otyepka, M., Wade, R. C., Tsuda, M., Nagata, Y., and Damborsky, J. (2009) Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. *Nat. Chem. Biol. 5*, 727−733.
(27) van Leeuwen, J. G., Wijma, H. J., Floor, R. J., van der Laan, J. M., and Janssen, D. B. (2012) Directed evolution strategies for enantiocomplementary haloalkane dehalogenases: from chemical waste to enantiopure building blocks. *ChemBioChem. 13*, 137−148.
(28) Koudelakova, T., Chaloupkova, R., Brezovsky, J., Prokop, Z., Sebestova, E., Hesseler, M., Khabiri, M., Plevaka, M., Kulik, D., Kuta Smatanova, I., Rezacova, P., Ettrich, R., Bornscheuer, U. T., and Damborsky, J. (2013) Engineering enzyme stability and resistance to an organic cosolvent by modification of residues in the access tunnel. *Angew. Chem., Int. Ed. 52*, 1959−1963.
(29) Dvorak, P., Kurumbang, N. P., Bendl, J., Brezovsky, J., Prokop, Z., and Damborsky, J. (2014). Manuscript under preparation.
(30) Novagen (2004) *Duet Vectors—User Protocol*, EMD Biosciences, Darmstadt, Germany.
(31) Nawabi, P., Bauer, S., Kyrpides, N., and Lykidis, A. (2011) Engineering *Escherichia coli* for biodiesel production utilizing a bacterial fatty acid methyltransferase. *Appl. Environ. Microbiol. 77*, 8052−8061.
(32) Zhang, C., Liu, L., Teng, L., Chen, J., Liu, J., Li, J., Du, G., and Chen, J. (2012) Metabolic engineering of *Escherichia coli* BL21 for biosynthesis of heparosan, a bioengineered heparin precursor. *Metab. Eng. 14*, 521−527.
(33) Xu, P., Vansiri, A., Bhan, N., and Koffas, M. A. (2012) ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. *ACS Synth. Biol. 1*, 256−266.
(34) Tolia, N. H., and Joshua-Tor, L. (2006) Strategies for protein expression in *Escherichia coli*. *Nat. Methods 3*, 55−64.
(35) Nikel, P. I., Pérez-Pantoja, D., and de Lorenzo, V. (2013) Why are chlorinated pollutants so difficult to degrade aerobically? Redox stress limits 1,3-dichloprop-1-ene metabolism by *Pseudomonas pavonaceae*. *Philos. Trans. R. Soc., B 368*, 20120377.
(36) Diaz Ricci, J. C., and Hernández, M. E. (2000) Plasmid effects on *Escherichia coli* metabolism. *Crit. Rev. Biotechnol. 20*, 79−108.
(37) Janssen, D. B., Pries, F., Vanderploeg, J., Kazemier, B., Terpstra, P., and Witholt, B. (1989) Cloning of 1,2-dichloroethane degradation genes of *Xanthobacter autotrophicus* GJ10 and expression and sequencing of the *dhlA* gene. *J. Bacteriol. 171*, 6791−6799.
(38) Assis, H. M. S. (1993) *Biochemical characterization of a haloalcohol dehalogenase from Arthrobacter sp. H10a*. Ph.D. Thesis, University of Kent, Kent, U.K.

Dvorak *et al. Microb Cell Fact (2015) 14:201*
DOI 10.1186/s12934-015-0393-3

Microbial Cell Factories

RESEARCH

Open Access

Exacerbation of substrate toxicity by IPTG in *Escherichia coli* BL21(DE3) carrying a synthetic metabolic pathway

Pavel Dvorak[1,2], Lukas Chrast[1,2], Pablo I. Nikel[3], Radek Fedr[2,4], Karel Soucek[2,4,6], Miroslava Sedlackova[5], Radka Chaloupkova[1,2], Víctor de Lorenzo[3], Zbynek Prokop[1,2] and Jiri Damborsky[1,2*]

Abstract

Background: Heterologous expression systems based on promoters inducible with isopropyl-β-D-1-thiogalactopyranoside (IPTG), e.g., *Escherichia coli* BL21(DE3) and cognate $LacI^{Q}/P_{lacUV5}$-T7 vectors, are commonly used for production of recombinant proteins and metabolic pathways. The applicability of such cell factories is limited by the complex physiological burden imposed by overexpression of the exogenous genes during a bioprocess. This burden originates from a combination of stresses that may include competition for the expression machinery, side-reactions due to the activity of the recombinant proteins, or the toxicity of their substrates, products and intermediates. However, the physiological impact of IPTG-induced conditional expression on the recombinant host under such harsh conditions is often overlooked.

Results: The physiological responses to IPTG of the *E. coli* BL21(DE3) strain and three different recombinants carrying a synthetic metabolic pathway for biodegradation of the toxic anthropogenic pollutant 1,2,3-trichloropropane (TCP) were investigated using plating, flow cytometry, and electron microscopy. Collected data revealed unexpected negative synergistic effect of inducer of the expression system and toxic substrate resulting in pronounced physiological stress. Replacing IPTG with the natural sugar effector lactose greatly reduced such stress, demonstrating that the effect was due to the original inducer's chemical properties.

Conclusions: IPTG is not an innocuous inducer; instead, it exacerbates the toxicity of haloalkane substrate and causes appreciable damage to the *E. coli* BL21(DE3) host, which is already bearing a metabolic burden due to its content of plasmids carrying the genes of the synthetic metabolic pathway. The concentration of IPTG can be effectively tuned to mitigate this negative effect. Importantly, we show that induction with lactose, the natural inducer of P_{lac}, dramatically lightens the burden without reducing the efficiency of the synthetic TCP degradation pathway. This suggests that lactose may be a better inducer than IPTG for the expression of heterologous pathways in *E. coli* BL21(DE3).

Keywords: Metabolic burden, Substrate toxicity, *Escherichia coli*, Heterologous metabolic pathway, Isopropyl β-D-1-thiogalactopyranoside, Lactose, 1,2,3-trichloropropane, Metabolic engineering

Background

Escherichia coli is among the most widely used microbial hosts in both fundamental and applied research [1]. *E. coli* strain BL21(DE3) carries an inducible T7 RNA polymerase-dependent expression system that allows for the simple manipulation and tuning of protein production levels, and it has become a laboratory workhorse [1–4]. This strain carrying commercial pET vectors or their derivatives has been the host of choice in numerous studies on recombinant protein expression [5, 6]. More recently, it has found applications as a cell factory for heterologous expression of entire biochemical pathways

*Correspondence: jiri@chemi.muni.cz
[1] Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic
Full list of author information is available at the end of the article

BioMed Central

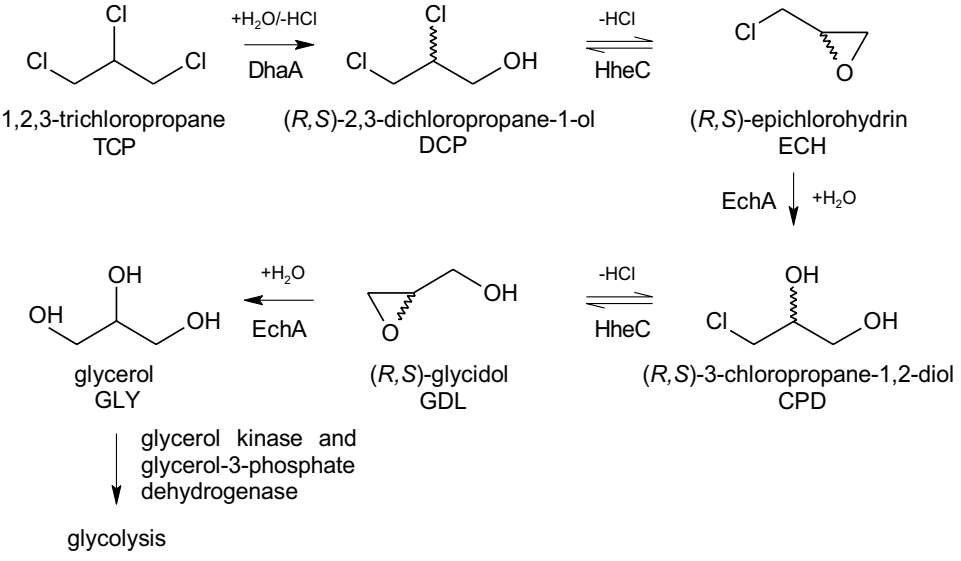
© 2015 Dvorak et al. This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

in the emerging fields of metabolic engineering and synthetic biology [6–10].

Despite its popularity, the *E. coli* BL21(DE3) and $LacI^Q/P_{lacUV5}$-T7 expression system suffers from certain drawbacks that mainly stem from the rapid and strong overexpression of recombinant proteins triggered by exposure to the synthetic inducer IPTG. The negative effects on host fitness associated with redirecting cells' metabolic capacities to achieve high levels of protein production are known as the *metabolic burden* or *metabolic load* [11]. The burden is often attributed to the overconsumption of metabolic precursors (e.g., amino acids, rRNAs, ATP, and reducing power) to fuel the synthesis of non-essential foreign proteins [12] or the energetically demanding maintenance and replication of plasmid vectors bearing heterologous genes and selection markers [13–15]. Fitness costs associated with the activities of the foreign proteins, which may cross-talk with the host's extant metabolic network [11] and burdens linked to the components of the expression system, such as the IPTG inducer and its import into the cell, are also frequently discussed [16, 17].

In addition to metabolic burden originating from the expression of foreign pathway components, the microbial cell factories used for biosynthesis of value-added chemicals or biodegradation of polluting compounds may be challenged by the toxicity of the processed substrate or its metabolic intermediates. These issues must be accounted for when considering the evolution of metabolic routes for biodegradation in natural and heterologous hosts [18–22]. Toxicity problems have also been encountered during the engineering of biosynthetic pathways for fatty acids, 1,3-propanediol, amorphadiene, taxadiene, and ethanol in *E. coli* [9, 23–27]. In addition, studies on the consolidated bioprocessing of lignocellulose have highlighted the potential adverse effects of inhibitory molecules in biomass-hydrolysate substrates [28]. Our current understanding of cellular responses to the exogenous and endogenous stressors that may be encountered during bioprocesses is extensive [29]. However, the combined effects of multiple simultaneous stresses on the hosts and their engineered induction systems have not been examined in depth.

To address this knowledge gap, we examined a recombinant strain of *E. coli* BL21(DE3) under conditions that provide an extreme combination of exogenous and endogenous stresses. The studied strain bears foreign genes encoding a five-step synthetic metabolic pathway for converting the industrial waste product and emerging environmental pollutant 1,2,3-trichloropropane (TCP) into the commodity chemical glycerol (Scheme 1). We have previously assembled this pathway in *E. coli* BL21(DE3) under the control of the $LacI^Q/P_{lacUV5}$-T7 system, and used the resulting construct to investigate reported pathway bottlenecks observed in vivo in the environmental bacterium *Agrobacterium radiobacter*



Scheme 1 Five-step biotransformation of 1,2,3-trichloropropane into glycerol by the enzymes of the synthetic biodegradation pathway. The pathway consists of haloalkane dehalogenase DhaA from *Rhodococcus rhodochrous* NCIMB 13064 [33] with the haloalcohol dehalogenase HheC and epoxide hydrolase EchA from *Agrobacterium radiobacter* AD1 [34, 35]. Computer-assisted protein engineering was used to improve activity of the haloalkane dehalogenase towards 1,2,3-trichloropropane, leading to the development of the 32-fold more active and 26-fold more efficient mutant DhaA31 [31]. Formed glycerol can be utilized in central catabolic pathways of the host cell

AD1 and in vitro in a version of the pathway constructed using immobilized enzymes [20, 22, 30]. By using protein engineering, metabolic engineering, and synthetic biology techniques, we were able to improve the pathway's performance [22, 31, 32] and identify two important factors limiting its output: an imbalance in the enantioselectivity of the pathway's enzymes, and the toxicity of the substrate and various pathway intermediates, which reduced the fitness of *E. coli* constructs incubated with TCP [20, 22].

Here, we assess the physiological consequences of using the IPTG-inducible system for the heterologous expression of this TCP degradation pathway. Our results demonstrate a negative synergistic effect of the inducer IPTG and the substrate TCP. Toxicity of TCP in resting *E. coli* BL21(DE3) cells is greatly exacerbated by pre-induction with IPTG. This effect was observed in the cells affected by metabolic burden from corresponding plasmids. Compared to this exacerbation effect, the physiological burden due to the production and presence of the pathway enzymes was minor. Moreover, we show that it is possible to moderate the pathway's fitness cost by tuning the IPTG concentration or reduce it dramatically by replacing IPTG with the natural inducer lactose. These findings highlight the need to take great care when selecting inducible systems for heterologous expression of metabolic pathways that catalyse harsh reactions, and to finely tune the fitness-productivity trade-off by optimizing the identity and concentration of the inducer.

Results and discussion

Biotransformation of TCP by resting *E. coli* BL21(DE3) cells carrying variants of a synthetic degradation pathway

Variants of the synthetic pathway featuring either the wild-type haloalkane dehalogenase or the 26-fold catalytically more efficient mutant DhaA31 were introduced into *E. coli* BL21(DE3) [22]. This host was selected because neither the enzymes nor the metabolites of the TCP pathway occur naturally in its metabolic network and because of the broad repertoire of commercially available Duet vectors for this strain, which enable the tunable co-expression of multiple genes in a single cell [5]. This resulted in the construction of a flexible system with limited risk of metabolic cross-talk, in which the expression of the three pathway components could be manipulated orthogonally [22, 32]. Three previously constructed *E. coli* BL21(DE3) degraders designated degWT, deg31, and deg31opt were tested (Table 1). *E. coli* degWT carries a variant of the TCP pathway based on the wild-type DhaA together with HheC and EchA, expressed in a relative ratio of 0.24:0.36:0.40, respectively, as determined after pre-induction with 0.2 mM IPTG (Additional file 1: Fig. S1). Strains deg31 and deg31opt both carry the TCP pathway featuring the engineered dehalogenase DhaA31,

Table 1 Bacterial strains and plasmids used in this study

Plasmid or strain	Relevant characteristics[a]	Source or reference
Plasmids		
pCDFDuet-1	Expression vector; *CDF* (pCloDF13 replicon), encodes two multiple cloning sites, each of which is preceded by a T7 promoter, a *lac* operator, and a ribosome binding site; Sm^R	Merck Millipore, Germany
pCDF-dhaAwt	pCDFDuet-1 carrying wild-type *dhaA* gene (encoding haloalkane dehalogenase) from *Rhodococcus rhodochrous* NCIMB 13064	[22]
pCDF-dhaA31	pCDFDuet-1 carrying mutant variant (*dhaA31*) of haloalkane dehalogenase gene	[22]
pETDuet-1	Expression vector; *ColE1* (pBR322 replicon), encodes two multiple cloning sites, each of which is preceded by a T7 promoter, a *lac* operator, and a ribosome binding site; Amp^R	Merck Millipore, Germany
pETDuet-echA-hheC	pETDuet-1 carrying wild-type *echA* and *hheC* genes (encoding epoxide hydrolase and haloalcohol dehalogenase) from *Agrobacterium radiobacter* AD1	[22]
pACYCDuet-1	Expression vector; *P15A* (pACYC184 replicon), encodes two multiple cloning sites, each of which is preceded by a T7 promoter, a *lac* operator, and a ribosome binding site; Cm^R	Merck Millipore, Germany
pACYC-echA-hheC	pACYC-1 carrying the wild-type *echA* and *hheC* genes (encoding epoxide hydrolase and haloalcohol dehalogenase) from *Agrobacterium radiobacter* AD1	[22]
Escherichia coli		
DH5α	Cloning host; Φ80*lacZ*ΔM15 *recA1 endA1 gyrA96 thi-1 hsdR17*($r_K^- m_K^+$) *supE44 relA1 deoR* Δ(*lacZYA-argF*)*U169*	Life Technologies, USA
BL21(DE3)	Expression host; F^- *ompT gal dcm lon hsdS*$_B$($r_B^- m_B^-$) λ(DE3 [*lacI*^{Q} *lacUV5-T7 gene 1 ind1 sam7 nin5*])	Life Technologies, USA
Host control	*E. coli* BL21(DE3) transformed with pCDF and pETDuet, control strain with empty plasmids	This study
degWT	*E. coli* BL21(DE3) transformed with pCDF-dhaAwt and pETDuet-echA-hheC	[22]
deg31	*E. coli* BL21(DE3) transformed with pCDF-dhaA31 and pETDuet-echA-hheC	[22]
deg31opt	*E. coli* BL21(DE3) transformed with pCDF-dhaA31 and pACYC-echA-hheC	[22]

[a] Antibiotic markers: *Sm* streptomycin, *Amp* ampicillin, *Cm* chloramphenicol

but the relative ratio of the three enzymes in deg31 is 0.14:0.41:0.45 while in deg31opt it is 0.63:0.16:0.21. Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) experiments indicated that the three pathway enzymes together accounted for a similar proportion of the total soluble protein fraction produced by all three degraders: 52 % for degWT, 54 % for deg31, and 44 % for deg31opt.

Pre-induced (0.2 mM IPTG) resting cells of each of the recombinant strains and a host control (Table 1) were incubated in phosphate buffer with 2 mM TCP, and the time-courses of TCP biotransformation over a 5 h interval were recorded (Fig. 1). The reaction profiles revealed fundamental differences between the strains with respect to the initial rates of TCP dehalogenation, accumulation of intermediates, and overall efficiency of glycerol formation. The theoretical concentrations of glycerol, otherwise rapidly utilized by *E. coli*, could be calculated from the experimental concentrations of TCP and detected intermediates by virtue of the pathway's orthogonal nature [22, 32]. While deg31opt benefited from the fastest first step (Fig. 1d), the best balanced pathway with the highest glycerol production was deg31 (Fig. 1c). On the other hand, degWT suffered from slow TCP conversion (Fig. 1b), prolonged exposure to the toxic substrate, and insufficient pathway output. As expected, the host control (carrying the corresponding empty plasmids, Fig. 1a) showed no activity toward TCP in the closed batch system.

The three *E. coli* recombinants and the control strain lacking the synthetic route represent a suitable model system for studying the contribution of metabolic burden and substrate/metabolite toxicity to the fitness cost of TCP biotransformation by whole-cell catalysts.

Assessment of metabolic burden and substrate toxicity effects by plating

Cell viability, estimated by plating, is a key physiological parameter that should reflect the individual strains' ability to cope with the stresses caused by the metabolic burden and TCP exposure [36]. *E. coli* degraders pre-induced with 0.2 mM IPTG and host controls were plated before and after 5 h of incubation in phosphate buffer with or without 2 mM TCP. The percentage of surviving cells was calculated after incubation.

The data obtained from plating before incubation (Fig. 2a) illustrate the separate effects of individual elements of the overall metabolic burden imposed on the cells. Several factors, including the presence of plasmid DNA and the associated selection markers, the addition of IPTG, and the burden due to heterologous pathway expression affected the degraders' viability in parallel even before the addition of the toxic substrate. The most pronounced impact at this stage was due to plasmid maintenance and the associated constitutive expression of selection marker genes from the Duet vectors, as well as the presence of IPTG. The presence of two medium-to-high copy plasmids pCDF (20–40 copies per cell) and pETDuet (~40 copies per cell) reduced viability by 50 % ($P < 0.01$; Fig. 2a). The "pre-induction" of the host control with empty plasmids reduced viability by almost 40 % relative to the non-induced control ($P < 0.01$). The expression of pathway enzymes in the *E. coli* degraders further reduced viability by about 20 % ($P < 0.05$). Additional reduction of viability of deg31opt compared to degWT and deg31 can be potentially attributed to the difference in antibiotic selection markers among three recombinants.

The data collected after 5 h incubation with or without TCP (Fig. 2b and Additional file 1: Fig. S2) included several unexpected observations. Surprisingly, TCP (initially added at a concentration of 2 mM) had only minor or negligible effects on the viability of non-induced control cells bearing empty plasmids; there was no significant difference in viability between these cells and the host controls that were exposed to neither TCP nor IPTG. This was unexpected because TCP has been reported to strongly inhibit growing cells of *E. coli* BL21(DE3) and natural hosts such as *A. radiobacter* AD1 or *Pseudomonas putida* MC4 even at concentrations 50 % lower than those used here [20, 22, 37]. On the other hand, the detrimental effect of IPTG was statistically significant ($P < 0.01$) (Fig. 2b and Additional file 1: Fig. S2). The most striking observation was that the relative viability of cells pre-induced with IPTG and then exposed to TCP was almost 90 % lower than that of host controls exposed to neither substance ($P < 0.01$) (Fig. 2b). This dramatic loss of viability does not correspond to a simple sum of the two compounds' individual effects; it is obvious that the IPTG exacerbated the toxicity of TCP. The fact that IPTG exacerbated the toxicity of TCP and not vice versa was confirmed by plating of three recombinants bearing the synthetic biodegradation pathway (Fig. 2b). Because these degraders had functional pathways for TCP degradation, they were better able to tolerate its presence. Importantly, the faster the conversion of TCP by the pathway enzymes, the greater the degraders' viability. Deg31opt, which achieved a rapid initial conversion of TCP but accumulated significant quantities of the intermediates DCP and GDL, survived the 5 h incubation almost as well as the host control that was not exposed to the substrate. These data are consistent with the previously reported results of growth arrest tests, which indicated that TCP is the most toxic compound in the pathway [22].

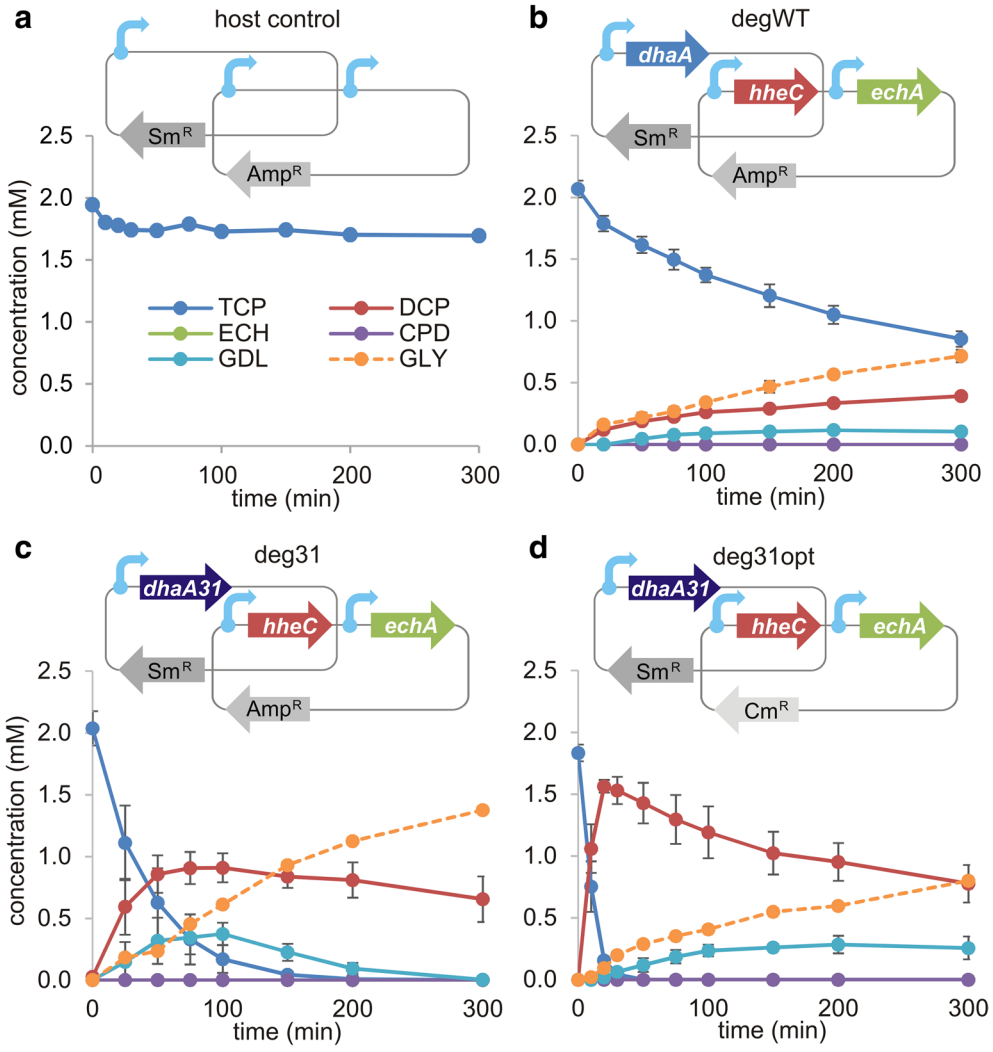


Fig. 1 Biotransformation of TCP catalysed by different *Escherichia coli* BL21(DE3) recombinants. **a** Control strain carrying the empty pCDF and pETDuet plasmids with streptomycin and ampicillin resistance marker genes, respectively. The *blue arrows* indicate individual T7 promoters. **b** The degrader degWT, which carries the haloalkane dehalogenase gene (*dhaA*) on pCDF and the haloalcohol dehalogenase (*hheC*) and epoxide hydrolase (*echA*) genes on pETDuet. **c** The degrader deg31, which carries the haloalkane dehalogenase mutant (*dhaA31*) gene on pCDF and two remaining genes of the degradation pathway on pETDuet. **d** The degrader deg31opt, which carries the *dhaA31* gene on pCDF and the two remaining genes of the degradation pathway on pACYC along with a chloramphenicol marker gene. The relative ratios of the TCP pathway genes produced by the degraders degWT, deg31, and deg31opt are 0.24:0.36:0.40, 0.14:0.41:0.45 and 0.63:0.16:0.21, respectively; the corresponding theoretical conversions of TCP into glycerol (GLY) are 35, 68, and 44 %, respectively. *Error bars* represent standard deviations calculated from three independent experiments. Theoretical concentrations of GLY were calculated from experimentally determined concentrations of TCP and intermediates. Sm^R streptomycin marker gene; Amp^R ampicillin marker gene; Cm^R chloramphenicol marker gene; *DCP* 2,3-dichloropropan-1-ol; *ECH* epichlorohydrin; *CPD* 3-chloropropane-1,2-diol; *GDL* glycidol. Note that the *green line* representing ECH is not visible because this intermediate does not accumulate at detectable levels during the experiment

Assessment of metabolic burden and substrate toxicity effects by multi-parameter flow cytometry

Multi-parameter flow cytometry allows simultaneous determination of several biochemical and physical variables immediately after sample preparation and hence should provide more accurate information on the cells' physiological status than can be obtained by plating [38, 39]. This technique also offers key information on the heterogeneity of bacterial populations. Unlike plating, it does not underestimate the number of viable cells in original cultures in cases when a fraction of the population has experienced sublethal injury and lost the capacity to grow [39, 40].

Incubation of the three degraders and host controls was conducted under the conditions mentioned in the previous section. The samples withdrawn after 5 h incubation

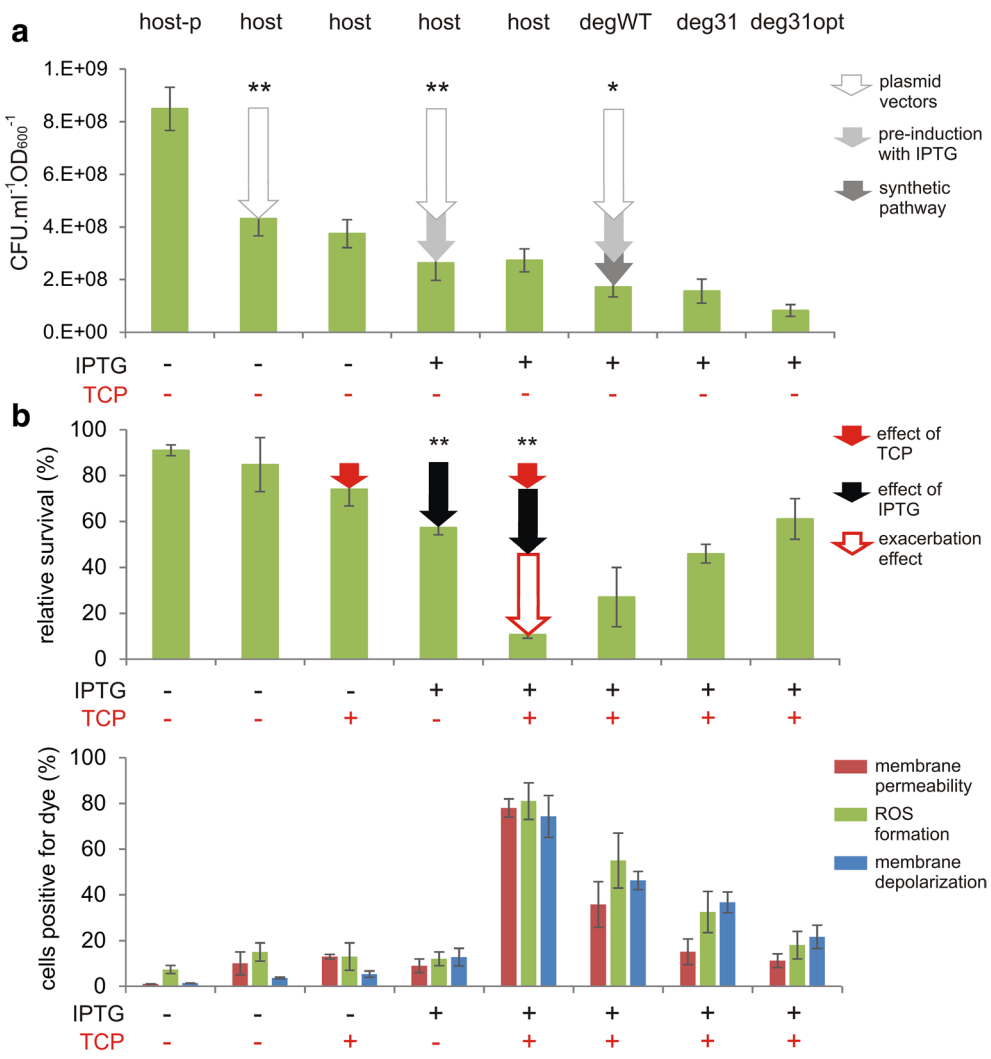


Fig. 2 Effects of metabolic burden and TCP toxicity on physiological parameters of *Escherichia coli* BL21(DE3) cells and three recombinants expressing the synthetic metabolic pathway. **a** Viability of cells non-induced or pre-induced with IPTG determined by plating before incubation in phosphate buffer. The effects of metabolic burden stemming from the presence of plasmids, pre-induction with 0.2 mM IPTG, and expression of the synthetic pathway are indicated by *coloured arrows*. *Asterisks* denote significance in decrease of cell count caused by each of three effects at either $P < 0.05$ (*) or $P < 0.01$ (**) when compared with preceding condition. **b** The percentage of surviving cells (*upper graph*) and the corresponding physiological parameters determined by flow cytometry (*lower graph*) after incubation in buffer with or without 2 mM TCP. The separate effects of TCP, IPTG, and the exacerbation of TCP toxicity in cells pre-induced with IPTG are indicated by *coloured arrows*. *Asterisks* denote significant difference in the decrease of cell count caused by each of three effects at $P < 0.01$ when compared with preceding condition. Physiological parameters including membrane permeability, formation of reactive oxygen species (ROS), and membrane depolarization were evaluated by staining the cells with appropriate dyes as explained in the Methods section. *Error bars* represent standard deviations calculated from at least five independent experiments. *CFU* colony forming units; *host-p E. coli* BL21(DE3) without plasmids; *host E. coli* BL21(DE3) with the empty pETDuet and pCDF plasmids

with or without TCP were stained with propidium iodide (PI), 6-carboxy-2′,7′-dichlorodihydrofluorescein diacetate (carboxy-H_2DCFDA), or *bis*-(1,3-dibutylbarbituric acid) trimethine oxonol [DiBAC$_4$(3)]. Dyes for labelling of nucleic acids, such as PI, which only enters cells with compromised membranes, are commonly used with membrane potential-sensitive dyes such as DiBAC$_4$(3), which binds the lipid-containing intracellular components, to study bacterial viability [38, 39, 41]. Carboxy-H_2DCFDA is a chemically reduced, acetylated and carboxylated form of fluorescein that has found many applications as a general indicator for the presence of reactive oxygen species (ROS) in eukaryotic and prokaryotic cells [42]. Recent studies on bacterial utilization of chlorinated aliphatic hydrocarbons suggest that this process is associated with strong oxidative stress and we

hypothesized that TCP would also evoke such a physiological response [42, 43]. The saturation of unsaturated fatty acids by halogenation or lipid peroxidation causes changes in membrane fluidity, resulting in the collapse of electron transport chains and premature electron transfer to oxygen, accompanied by ROS formation, membrane permeabilization, and cell death [44, 45].

The end-point flow cytometry protocol adopted in our study was less sensitive than plating in demonstrating the burden caused by plasmids (Fig. 2b), possibly because the presence of heterologous DNA and the constitutive expression of selection markers did not directly alter properties targeted by cytometry but instead imbalanced the cells' overall energy status. However, the flow cytometry approach using selected fluorochromes was useful in exposing the toxic effects of IPTG and TCP. There were no significant differences ($P > 0.05$) between the non-induced host controls with empty plasmids, regardless of their TCP exposure status, with respect to any of the three variables examined in these experiments (membrane permeability, ROS formation, and membrane potential; see Fig. 2b). However, the proportion of cells staining positive for $DiBAC_4(3)$ increased up to threefold in the control treated with IPTG alone ($P < 0.01$). The same effect was also observed in deg31, whose response to induction and incubation with TCP was studied in more detail (Fig. 3). The fraction of the bacterial population staining positively with all three dyes increased many-fold when pre-treatment with IPTG was combined with TCP exposure, confirming the previously observed exacerbating effect and indicating that the action of TCP in bacterial cells is accompanied by extensive ROS

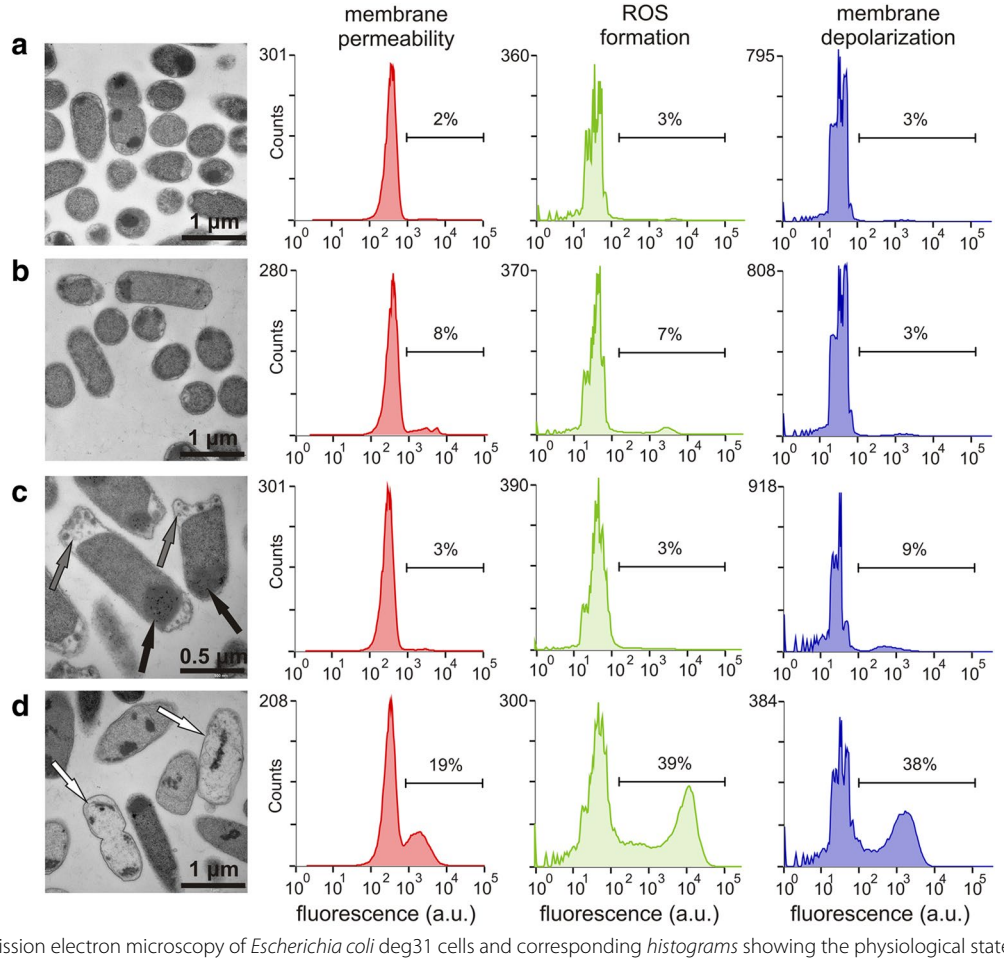


Fig. 3 Transmission electron microscopy of *Escherichia coli* deg31 cells and corresponding *histograms* showing the physiological state of populations stained with selected fluorescent dyes. **a** Non-induced cells incubated in phosphate buffer. **b** Non-induced cells incubated in phosphate buffer with 2 mM TCP. **c** Cells pre-induced with 0.2 mM IPTG incubated in phosphate buffer. **d** Cells pre-induced with 0.2 mM IPTG and incubated in phosphate buffer with 2 mM TCP. *Black arrows* indicate bodies that presumably consist of overexpressed heterologous proteins, *grey arrows* indicate separations of the inner and outer cell membranes, and *white arrows* indicate dead or dying cells

formation (Figs. 2b, 3). Membrane depolarization, ROS accumulation, and membrane permeability were reduced in *E. coli* recombinants expressing the synthetic biodegradation pathway, with the degree of reduction being proportional to the strains' initial rates of TCP conversion. Interestingly, the exacerbation of compound toxicity by pre-induction of the BL21(DE3) host control with IPTG was confirmed also in experiments using the model toxic compound *tert*-butyl hydroperoxide (TBHP), an organic peroxide and strong ROS formation promoting oxidative agent (Additional file 1: Fig. S3). This suggests that described exacerbation phenomenon should not be restricted only to our model toxic chemical TCP.

Membrane depolarization and ROS formation in vivo are dynamic processes. To follow their kinetics in *E. coli* degraders we also performed time-resolved measurements (Additional file 1: Fig. S4). The number of cells stained by $DiBAC_4(3)$ and carboxy-H_2DCFDA increased linearly over time in all of the stressed recombinants except for deg31. Interestingly, this degrader exhibited an initial burst of $DiBAC_4(3)$ and carboxy-H_2DCFDA fluorescence but these signals then plateaued or fell slightly. We assume that the characteristic profiles of $DiBAC_4(3)$ and carboxy-H_2DCFDA fluorescence for deg31 are linked to its unique variant of the synthetic biodegradation pathway and the corresponding time-course of TCP biotransformation. In contrast to the other strains, TCP, 2,3-dichloropropane-1-ol (DCP) and glycidol (GDL) were present at relatively high concentrations in the deg31 reaction mixture between minutes 50 and 100 of the measurement period. This may have caused synergistic toxicity, increasing the number of cells with depolarized membranes and enhanced ROS formation. Such joint effects are common; they have been observed for organophosphate pesticides, fluorosurfactants, and heavy metals, among others [46–48]. Subsequent freezing or moderate fall in the intensity of the signals was attributed to the parallel removal of TCP and GDL and the production of glycerol, which is known to be an efficient stress protectant in yeast and solvent-tolerant strains of *E. coli* [49, 50].

The variant of the synthetic pathway present in deg31 seemed to provide the best compromise in terms of coping with toxicity while efficiently converting TCP into harmless glycerol and was therefore selected for further investigation.

Assesment of metabolic burden and substrate toxicity effects by electron microscopy

Metabolic burden and toxicity can cause changes in the morphology of bacterial hosts [11, 51]. We therefore used transmission electron microscopy to study the changes in morphology of induced and non-induced deg31 cells after 5 h incubation with or without 2 mM TCP (Fig. 3). Pictures of induced and non-induced *E. coli* host control cells with empty plasmids are shown in (Additional file 1: Fig. S5). Microscopic observations were followed by multi-parameter flow cytometry of deg31 cells stained with PI, carboxy-H_2DCFDA, or $DiBAC_4(3)$.

Incubation of non-induced degraders with TCP produced only a small proportion of dead bacterial cells, and the morphology of cells treated in this way did not generally differ from that of cells unexposed to the toxic substrate (Fig. 3a, b). The proportion of cells stained by carboxy-H_2DCFDA and PI increased moderately, but no effect on the membrane potential was observed. Pre-induced bacteria producing recombinant proteins intensively formed visible inclusion bodies and showed frequent separation of the cytoplasmic membrane from the outer membrane at the poles of the cell (Fig. 3c). Since the majority of the recombinant proteins obtained from cell lysates were soluble (data not shown), we believe that the observed bodies consisted of active enzymes.

The combination of pre-induction and incubation with TCP produced the most pronounced morphological changes and was accompanied by substantial increases in the number of cells staining positively for PI, carboxy-H_2DCFDA, and $DiBAC_4(3)$ (Fig. 3d). Numerous dead or dying cells with damaged cytoplasmic membranes and leaked contents were clearly visible. Even so, a significant portion of the population resisted the combined effects of IPTG and TCP over the 5 h treatment period. This was mainly due to the well-balanced synthetic pathway of deg31 and its fast conversion of TCP to glycerol. Bistability is a common phenomenon and can be attributed to noise in the expression of the multigenic traits responsible for toxicity tolerance and the graduated stress response in the bacterial population [29, 36, 52]. In summary, our microscopic observations of deg31 populations treated under four different conditions were consistent with previous results and supported the conclusion that pre-induction with IPTG exacerbates TCP toxicity.

Toxicity exacerbation effect rises in cells experiencing metabolic burden from plasmids

Given that both the *E. coli* degraders and the host controls used in this work had to cope with the metabolic burden of the Duet plasmids and $LacI^Q/P_{lacUV5}$-T7 expression system, we decided to include *E. coli* controls without these components in the following experiment. For this purpose, we used *E. coli* BL21(DE3) without plasmids and the cloning strain *E. coli* DH5α, which lacks both the $LacI^Q/P_{lacUV5}$-T7 expression system and the *lac* operon. By employing the plating and flow cytometry protocols described above, we found that the exacerbation effect was modest or completely absent in both

strains (Additional file 1: Fig. S6A, B). This suggests that the double stress from IPTG and TCP is only manifested in strains carrying Duet plasmids and the corresponding expression system.

We presume that the studied recombinants could not efficiently suppress the double stress from IPTG and the toxic substrate due to the metabolic burden imposed by plasmids and the corresponding shortage of resources required for cell maintenance. It seems that the chemical structure of IPTG, its transport or presence inside the cell triggers physiological changes that facilitate the manifestation of TCP toxicity in *E. coli* BL21(DE3) cells with Duet plasmids.

Reducing metabolic burden and toxicity exacerbation by tuning IPTG concentration

Next, we attempted to reduce the burden imposed on *E. coli* recombinants by optimizing the concentration of IPTG. We looked for the lowest possible concentration of inducer that would minimize the fitness cost without significantly compromising the system's efficiency at degrading TCP. Deg31 cells were pre-induced with IPTG concentrations ranging from 0.01 to 1.00 mM and the resulting effects on cell viability and pathway efficiency were studied (Fig. 4; Additional file 1: Fig. S7).

E. coli BL21(DE3) with the empty pETDuet and pCDF plasmids was used as a control for plating to assess the burden imposed on the host by IPTG exposure in the absence of the heterologous pathway (Fig. 4). The viability of the pre-induced degrader and control was checked before and after 5 h of incubation in phosphate buffer with 2 mM TCP and compared to that of cells not exposed to IPTG (Fig. 4a, b). The percentage of surviving cells after incubation was calculated in each case (Fig. 4c). These experiments indicated that even in the absence of TCP there was an inverse correlation between the IPTG concentration and the viability of the recombinant *E. coli* (Fig. 4a). The deg31 strain suffered more from increasing IPTG concentrations than the control, probably due to the additional burden of expressing genes encoding the synthetic pathway. The opposite was true after 5 h of incubation because the TCP-catabolizing deg31 strain was more resistant to exacerbated toxicity than the host control (Fig. 4b, c). The synthetic pathway strain exhibited similar survival rates at all tested IPTG concentrations except for the highest—the relative viability of deg31 pre-induced with 1 mM IPTG was 100 %. We assumed that this outlying value was due to underestimation of the number of viable cells before incubation due to the intensive stress experienced by the degrader upon induction with such a high concentration of IPTG and the resulting presence of viable-but-not-culturable cells

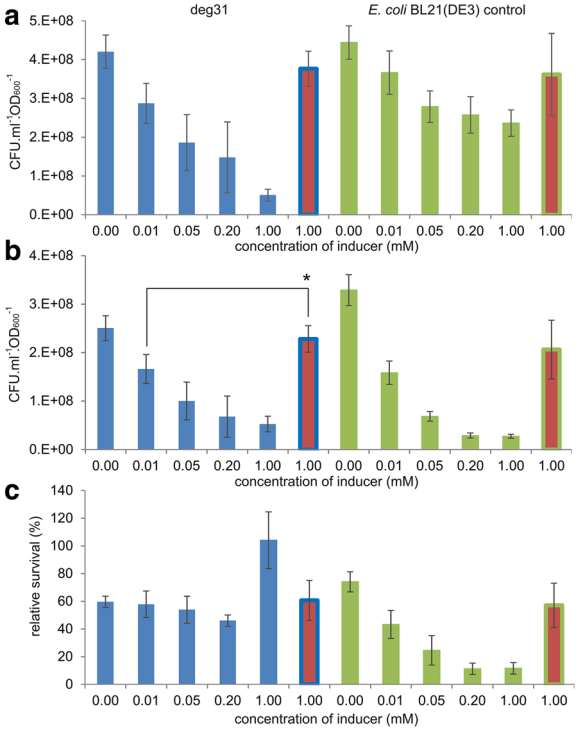


Fig. 4 Viability of *Escherichia coli* deg31 and host control strains after pre-induction with diverse concentrations of IPTG or 1 mM lactose, before and after incubation with TCP. **a** Viability of deg31 and *E. coli* BL21(DE3) with the empty pETDuet and pCDF plasmids as determined by plating of cells pre-induced with different concentrations of IPTG or 1 mM lactose (*red columns*) before incubation in phosphate buffer with TCP. **b** Viability of cells after incubation in buffer with 2 mM TCP. *Asterisks* denote significantly higher (at $P < 0.05$) cell count of deg31 pre-induced with 1 mM lactose when compared with the count of cells pre-induced with the lowest tested concentration of IPTG (0.01 mM). **c** Fraction of surviving cells calculated as the difference in the $CFUs.ml^{-1}.OD_{600}^{-1}$ before and after incubation with TCP. *Error bars* represent standard deviations calculated from at least four independent experiments. *CFU* colony forming units

in the suspension [40]. Plating experiments demonstrated that a fraction of the bacterial population regained its capacity to grow and reproduce some time after treatment with this high concentration of IPTG (Additional file 1: Fig. S8).

The time-courses of TCP biotransformation with resting cells of deg31 pre-induced with IPTG at 1.00, 0.20, 0.05, 0.025, and 0.01 mM (Additional file 1: Fig. S7) and densitometric analysis of SDS polyacrylamide gels with the corresponding samples of cell-free extracts (Additional file 1: Fig. S9, Table S1) showed that: (i) the relative ratio of pathway enzymes and the shape of the degradation profile did not change substantially with the inducer concentration, while (ii) the content of the three pathway enzymes in the total soluble protein of the recombinant

bacteria decreased from 55 % (1 mM IPTG) to 32 % (0.01 mM IPTG) and the pathway's output decreased from 70 to 46 %. Appreciable conversion of TCP was also achieved with non-induced cells due to the leakiness of the T7 promoter and basal expression of genes within the pathway (Additional file 1: Fig. S7).

Figure 5 summarizes the balance of the three parameters discussed in the preceding sections: (i) host viability, (ii) the cellular expression of pathway enzymes, and (iii) the output of the synthetic biodegradation route. We conclude that the minimal IPTG concentration that allows sufficient expression of genes within the pathway to achieve a reasonable output is 0.025 mM. Similar inducer concentrations that allow full gene expression have been reported for single recombinant proteins such as β-galactosidase, green fluorescent protein, and rhamnulose-1-phosphate aldolase [53–55]. Note that IPTG concentration of 0.025 mM is eight times lower than the originally tested inducer concentration and up to 40-times lower than the values reported in the scientific literature describing engineering of heterologous pathways in *E. coli* [9, 56]. Induction with lower amounts of IPTG improved the host's fitness. However, even the lowest concentration of 0.01 mM reduced the viability of the *E. coli* degrader by 30 % relative to non-induced deg31, and by up to 50 % relative to the non-induced host control ($P < 0.05$ in both cases; Fig. 4b). We therefore investigated the scope for replacing IPTG with an alternative inducer.

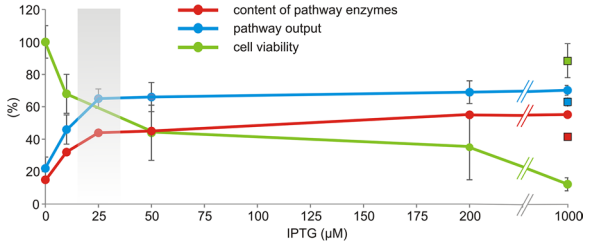


Fig. 5 Summarized effects of IPTG concentration on gene expression levels, pathway output, and cell viability in pre-induced *Escherichia coli* deg31 cells. Viability was determined by plating pre-induced deg31 cells resuspended in phosphate buffer before incubation with TCP. Pathway output was expressed as the theoretical conversion of TCP into glycerol at the end of 5 h degradation experiments with pre-induced, resting deg31 cells (see also Additional file 1: Fig. S5). The content of TCP pathway enzymes was estimated by analyzing cell-free extracts obtained from pre-induced cells by sodium dodecyl sulfate polyacrylamide gel electrophoresis (Additional file 1: Fig. S7 and Table S1). Two gels were analysed by densitometry and mean values are shown. *Error bars* represent standard deviations calculated from three independent experiments. Values determined for deg31 pre-induced with 1 mM lactose are indicated by *squares*

Reducing metabolic burden and toxicity exacerbation by inducing with lactose

Lactose is a natural inducer of the *lac* operon and can be employed as a cheaper alternative to synthetic IPTG. It has been proven to induce expression of recombinant proteins in *E. coli* to the same extent as IPTG on both laboratory and industrial scales [12, 57, 58]. In contrast to IPTG, lactose is a substrate of β-galactosidase (encoded by *lacZ*) and thus can be metabolized by cells with an intact *lac* operon, including *E. coli* BL21(DE3). Therefore, concentrations of lactose up to 30 mM are commonly used to induce the expression of cloned genes at levels that can be achieved with ≤1 mM IPTG [58].

We investigated the pre-induction of deg31 cells with 1 mM lactose. We assumed that this concentration would be sufficient to induce the expression of the TCP degradation pathway genes at levels that would confer adequate degradation efficiency. This expectation was confirmed by recorded time-courses of TCP biotransformation and the finding that pathway enzymes accounted for up to 41 % of the cells' total soluble protein under these conditions (Additional file 1: Figs. S7 and S9, and Table S1). The theoretical conversion of TCP into glycerol under these conditions was 63 %. These values are close to those observed for deg31 cells pre-induced with 0.025 or 0.05 mM IPTG. Most importantly, the deg31 cells pre-induced with lactose exhibited higher viability before and after 5 h incubation with TCP compared to bacteria treated with IPTG at any of the concentrations tested ($P < 0.05$; Fig. 4). The same relieving effect was observed for the host control. In term of survival, the cells pre-induced with lactose performed almost as well as their non-induced counterparts (Fig. 4c). In keeping with the plating results, flow cytometric analysis of host control and deg31 cells pre-induced with lactose revealed significantly lower levels of stressed cells with depolarized membranes than were observed for *E. coli* strains pre-induced with IPTG ($P < 0.01$; Additional file 1: Fig. S10). These results indicated again that the action of IPTG in the studied recombinants was consistently accompanied by changes in membrane properties.

A higher viability of cells induced with lactose instead of IPTG was previously reported for the expression of single heterologous proteins [12, 56, 58]. This effect was attributed to the delayed, milder induction achieved with the natural inducer. In contrast to synthetic IPTG, which can enter the cell rapidly both by passive diffusion and with the help of the LacY lactose permease, lactose can only enter the cytoplasm via the permease. Moreover, lactose must be converted into allolactose by β-galactosidase before binding to the *lac* repressor whereas IPTG binds to the repressor directly. Our results confirm that lactose imposes a lower metabolic

burden than IPTG, suggesting that it is also a more suitable inducer for the expression of entire heterologous pathways in *E. coli* BL21(DE3). This is particularly important for *E. coli* BL21(DE3) cells carrying synthetic pathways that degrade toxic compounds or produce toxic intermediates.

Conclusions

We studied the toxicity of TCP and its derived metabolites in *E. coli* BL21(DE3) carrying an inducible synthetic metabolic pathway comprising three catabolic genes [22]. The presence of plasmid DNA, exposure to IPTG, and the metabolic burden of the heterologous pathway were shown to affect the viability of *E. coli* degraders even before contact with the toxic substrate. The most pronounced impact was attributed to plasmid maintenance: the presence of two medium-to-high copy number plasmids (pCDF and pETDuet) reduced cell viability by 50 %. Pre-induction of host control cells carrying the empty plasmids caused a 40 % decrease in viability, while expression of the genes within the pathway further reduced viability by about 20 %. Treatment with exogenous TCP at a concentration of 2 mM had only a very minor or no effect on the viability of non-induced control cells carrying empty plasmids. The pronounced reduction in survival triggered by adding the substrate to pre-induced cells was unexpected, suggesting that TCP toxicity was exacerbated by pretreatment of the cells with IPTG. The recombinants clearly benefited from the presence of a functioning TCP degradation pathway—the faster the conversion of TCP, the greater the cell viability.

Flow cytometry with selected fluorochromes proved to be useful in dissecting the toxic effects of IPTG and TCP. Variations in the three monitored physiological parameters—membrane permeability, ROS formation, and membrane potential—was insignificant between the non-induced host controls irrespective of exposure to TCP. The fraction of cells staining positively for individual dyes rose significantly when IPTG was added together with TCP, verifying the previously observed exacerbation effect and showing that TCP exposure causes intensive ROS formation in bacterial cells. All three parameters were reduced in *E. coli* BL21(DE3) recombinants expressing the synthetic biodegradation pathway, and the degree of reduction increased in parallel with the initial rate of TCP degradation. Cell counts and flow cytometry data were in good agreement with the results of electron microscopy experiments.

The potential toxicity of IPTG and its negative effect on cell growth has been repeatedly reported [9, 16, 60, 61]. This effect was attributed to various factors including the induced changes in the host metabolism [16], the active uptake of the inducer into the cell by LacY [17, 62], and the rapid induction of heterologous gene expression [12, 56, 59]. However, that widely used inducer can exacerbate the toxicity of another compound, leading to damage of host cells experiencing metabolic burden, has not been previously reported and should be of wide interest. *E. coli* BL21(DE3) or similar hosts with the DE3 lysogen and well defined plasmid vectors are widely used because they make it possible to precisely balance the expression of heterologous pathway genes in metabolic engineering and synthetic biology [8, 10, 22, 25]. Still, insufficient attention has been paid to the possible exacerbation of multiple endogenous and exogenous stressors that impose a complex burden on such cell factories. IPTG is often applied at sub-millimolar concentrations [9, 63, 64] although the amount needed for full induction of heterologous genes can be an order of magnitude lower. The optimal concentration of IPTG in any given case may be system-specific [56], but the relatively simple experiments can be used to guide the optimization of its concentration. Importantly, the use of lactose as an alternative inducer should be considered: we observed that replacing IPTG with lactose dramatically reduced the burden experienced by the transformed bacterial cells, suggesting that it may be a better inducer than IPTG for the expression of heterologous pathways in *E. coli* BL21(DE3).

Methods

Chemicals and growth media

TCP, DCP, epichlorohydrin (ECH), 3-chloropropane-1,2-diol (CPD), GDL, and glycerol standards were purchased from Sigma-Aldrich Co. (St. Louis, MO, USA). All chemicals used in this study were of analytical grade. The fluorescent indicators carboxy-H_2DCFDA and DiBAC$_4$(3) were purchased from Life Technologies Inc. (Waltham, MA, USA) and PI from Sigma-Aldrich Co. IPTG was purchased from Duchefa Biochemie B.V. (Haarlem, The Netherlands). Lactose and TBHP solution (5–6 M in decane) were from Sigma-Aldrich Co. A Free Glycerol Assay Kit was acquired from BioVision Inc. (Milpitas, CA, USA). LB medium (Sigma-Aldrich Co.) was used for routine cultivation of *E. coli*.

Bacterial strains and growth conditions

Plasmids and *E. coli* strains used in this study are summarized in Table 1. *E. coli* DH5α was used as a control strain lacking *lacZYA* operon. *E. coli* BL21(DE3) was used as a heterologous host for inducible expression of the synthetic biodegradation pathway and also as a control strain bearing the *lacZYA* operon and LacI^Q/P_{lacUV5}-T7 expression system. Three variants of the pathway were constructed previously in a modular way by combining genes encoding haloalkane dehalogenase DhaA from *R.*

rhodochrous NCIMB 13064, the DhaA mutant DhaA31, and genes encoding haloalcohol dehalogenase HheC and epoxide hydrolase EchA from *A. radiobacter* AD1 on Duet vectors [22]. For the purpose of this study, *E. coli* BL21(DE3) degraders and host control strain BL21(DE3) with empty pCDF and pETDuet plasmids were prepared freshly by co-transforming plasmid constructs into competent cells using chemical transformation or electroporation. Cells were spread on LB agar plates with appropriate combination of antibiotics (25 $\mu g.ml^{-1}$ Sm and 75 $\mu g.ml^{-1}$ Amp for *E. coli* BL21(DE3) with empty pCDF and pETDuet, degWT, degWTopt, and deg31; 25 $\mu g.ml^{-1}$ Sm and 20 $\mu g.ml^{-1}$ Cm for deg31opt) and grown overnight at 37 °C. A single colony was used for preparation of overnight cultures in 10 ml of LB medium with the corresponding antibiotics. Precultures (250 µl) were used to inoculate fresh LB medium (25 ml) and cultures were grown at 37 °C with shaking until the cell density reached an optical density measured at 600 nm (OD_{600}) of 1 when the expression of recombinant genes was induced with IPTG or lactose. Induced cells were incubated overnight at 20 °C and biomass was collected at late exponential phase by centrifugation at 4000*g* for 10 min (4 °C). Pellets were washed with ice-cold sterile filtered sodium phosphate buffer (50 mM, pH 7.0) and dissolved in the same buffer for further use.

Incubation of resting cells with TCP and assessment of viability by plating

Cell suspensions of pre-induced or non-induced *E. coli* degraders and host cell controls were diluted with sterile filtered sodium phosphate buffer to a final OD_{600} of 7.0. TCP (4 mM) was dissolved in 5 ml of the same buffer in glass vials (25 ml) with a screw cap mininert valve for 1 h at 37 °C with shaking. The incubation was initiated by mixing 5 ml of the cell suspension with 5 ml of buffer with TCP. The final TCP concentration in 10 ml of the cell suspension of OD_{600} of 3.5 was 2 mM. In the cases indicated, 1 mM TBHP was added instead of TCP. In case of *E. coli* degraders, cell suspension samples (0.5 ml) were quenched in 0.5 ml acetone with hexan-1-ol, vortexed for 15 s and centrifuged at 18,000*g* for 2 min. The concentration of metabolites from the TCP pathway in the supernatant was analyzed using gas chromatography. The presence of GLY in the cell suspension was verified using a Free Glycerol Assay Kit (BioVision Inc.). Viability of degraders and controls before and after 2.5 or 5 h incubation with TCP was tested by plating 100 µl of cell suspension serially diluted with ice-cold sterile phosphate-buffered saline (PBS, 8 mM Na_2HPO_4, 1.5 mM KH_2PO_4, 3 mM KCl, and 137 mM NaCl, pH = 7.0) onto Plate Count Agar plates. Plates were incubated for 24 h at 37 °C and grown colonies were counted as $CFU.ml^{-1}.OD_{600}^{-1}$.

End-point flow cytometry

At the end of incubation of *E. coli* degraders and host cell controls with or without TCP, 0.1 ml of each cell suspension was added to 0.9 ml of filtered PBS. The suspensions were added with carboxy-H_2DCFDA (prepared as 4 mM stock solution in DMSO) to the final concentration of 40 µM and with PI (prepared as 200 $\mu g.ml^{-1}$ stock solution in PBS) to the final concentration of 2 $\mu g.ml^{-1}$ or with 1 µM $DiBAC_4(3)$ (prepared as 25 $\mu g.ml^{-1}$ stock solution in 4 mM EDTA). The suspensions with fluorescent probes were mixed by inverting and incubated in the dark for 10 min at room temperature. Flow cytometry analysis of fluorescence levels was performed using BD FACSAria II Sorp (BD Biosciences Co., San José, CA, USA) equipped with an argon-ion laser of 100 mW at 488 nm and a solid state 100 mW 561 nm as the excitation source. The carboxy-H_2DCFDA or $DiBAC_4(3)$ fluorescence emission at 525 or 516 nm, respectively, was detected using a 525/50-nm band pass filter array. The PI fluorescence emission at 617 nm was detected using a 610/20-nm band pass filter array. Data for at least 50,000 cells per single sample per experiment were collected. Size-related forward scatter signals gathered by the cytometer were used by the Cyflogic 1.2.1 software (CyFlo Ltd., Finland) to gate fluorescence data from bacteria in the stream. Percentage of cells positive for carboxy-H_2DCFDA, PI and $DiBAC_4(3)$ was calculated using the statistics module of the software.

Time-resolved flow cytometry

For time-resolved measurements of ROS accumulation in *E. coli* constructs, 0.5 ml of the cell suspension in sterile filtered sodium phosphate buffer (50 mM, pH 7.0) of OD_{600} of 7.0 was added with carboxy-H_2DCFDA at 40 µM. The suspension was mixed by inverting and incubated in the dark for 10 min at room temperature. The measurement was started immediately after mixing the cell suspension with 0.5 ml of sodium phosphate buffer with dissolved TCP in 5 ml polypropylene tube with snap cap. The final TCP concentration in 1 ml of cell suspension of OD_{600} of 3.5 was 2 mM. Flow cytometry analysis of fluorescence levels in time was performed on the same FACS instrumentation as above. FCS data were recorded in BD FACS Diva (v. 6.1.3, BD Biosciences Co.) at constant flow rate of 20,000 events. s^{-1} for intervals of 5 s every 30 min for total period of 5 h. Analysis of data was further performed in FlowJo LLC. (v. 7.6.5, Ashland, OR, USA) and exported to Excel (Office 2013, Microsoft Corp., Redmond, WA, USA) for graphical output. Time-resolved profiles of membrane depolarization in *E. coli* constructs during 5 h incubation in phosphate buffer were compiled from end-point measurements performed at time 0, 1, 2, 3, 4,

and 5 h following the procedure described in previous section.

Determination of expression levels of the TCP pathway enzymes

The expression levels of DhaA31, HheC, and EchA were determined in *E. coli* constructs induced with diverse concentrations of IPTG or 1 mM lactose. Washed cells were resuspended in sodium phosphate buffer and cell density was adjusted to OD_{600} of 7.0. 1 U of DNaseI per 1 ml of cell suspension was added. Cells were disrupted with One Shot cell disruptor (Constant Systems Ltd., UK) using 1.5 kbar shot. The cell lysate was centrifuged for 45 min at 18,000*g* at 4 °C and the resulting cell-free extract (CFE) was decanted. The concentration of total protein in CFEs was determined using Bradford reagent (Sigma Aldrich Co., USA). Samples of CFE containing 5 μg of total protein were separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). CFE prepared from *E. coli* BL21 (DE3) cells without plasmids were used as controls. Gels were stained with Coomassie Brilliant Blue R-250 (Fluka/Sigma-Aldrich Co., Switzerland) and analyzed using a GS-800 Calibrated Imaging Densitometer (Bio-Rad Laboraories Inc., USA). The expression levels of DhaA31, HheC and EchA and the relative ratios of the enzymes in degWT, deg31, and deg31opt were estimated from trace densities of corresponding bands using the software Quantity One 4.6.9 (Bio-Rad Laboratories Inc., USA).

Electron microscopy

To observe morphological changes, cells of *E. coli* deg31 induced with 0.2 mM IPTG or non-induced were incubated in presence or absence of 2 mM TCP in sodium phosphate buffer (50 mM, pH 7.0, OD_{600} of cell suspension of 3.5) at 37 °C for 5 h and then processed for transmission electron microscopy. Briefly, cells were harvested and pelleted by centrifugation (5000*g*, 5 min). Pellets were washed and resuspended in phosphate buffer and fixed in 3 % glutaraldehyde solution in the same buffer for 1 h at room temperature. Cells were then post-fixed in 1 % osmium tetroxide in phosphate buffer and after addition of 2.5 % agar and dehydration in a graded series of ethanol, the blocks of cells were impregnated with Durcupan (Sigma-Aldrich Co., USA) and embedded in silicone embedding moulds. Polymerization occurred for 3 days at 60–80 °C. Ultrathin sections prepared with a diamond knife on an Ultramicrotome EM UC6 (Leica Microsystems GmbH, Germany) were placed on copper grids, stained with 2.5 % uranyl acetate for 10 min and Reynolds lead citrate solution for 3 min and observed with a Morgagni 268D (FEI Co., The Netherlands) transmission electron microscope.

Gas chromatography and mass spectrometry measurements

A Gas Chromatograph 6890 N with a flame ionisation detector (GC-FID) and Gas Chromatograph 7890A and Mass Spectrometer (GC–MS) 5975C MSD (Agilent Technologies Inc., USA), both with the capillary column ZB-FFAP 30 m × 0.25 mm × 0.25 μm (Phenomenex Inc., USA) were used for routine analysis and quantification of TCP and its metabolites or for verification of the presence of individual metabolites from the TCP pathway in selected samples, respectively. Samples (2 μl) were injected into the GC with an inlet temperature of 250 °C and split ratio 20:1. The operational conditions for the column were: helium carrier gas with an initial flow of 0.6 $ml.min^{-1}$ for 1 min, followed by a flow gradient from 0.6 to 1.8 $ml.min^{-1}$ (ramp 0.2 $ml.min^{-1}$), temperature program set to give an initial column temperature of 50 °C for 1 min, followed by a temperature gradient from 50 to 220 °C hold for 2 min (ramp 25 $°C.min^{-1}$). The temperature of the detector was 250 °C. MS scan speed was 6.9 s^{-1}. This method was used for all GC analyses. For that purpose, calibration curve of 0–5 mM of TCP, DCP, ECH, CPD and GDL with internal standard hexan-1-ol was prepared. Detection limits calculated using the software OriginPro v8 (OriginLab Corporation, USA) were 3, 5, 6, 186 and 22 μM for TCP, DCP, ECH, CPD and GDL, respectively.

Statistical analysis

All experiments were independently repeated at least three times (number of repetitions is specified in figure legends) and the mean value of the corresponding parameter ± standard deviation is presented. Statistical significance was assessed using Student's *t* test with two-tailed hypothesis available in Microsoft Excel 2013 (Microsoft Corp., USA). The difference in between two independent data sets was considered statistically significant for $P < 0.05$.

Additional file

Additional file 1. Supplementary material.

Abbreviations

IPTG: isopropyl β-D-1-thiogalactopyranoside; TCP: 1,2,3-trichloropropane; PI: propidium iodide; carboxy-H_2DCFDA: 6-carboxy-2′,7′-dichlorodihydrofluorescein diacetate; $DiBAC_4(3)$: *bis*-(1,3-dibutylbarbituric acid) trimethine oxonol; TBHP: *tert*-butyl hydroperoxide.

Authors' contributions

PD, LC, PIN, RC, ZP, KS, VDL, and JD designed the experiments. PD and LC prepared the *E. coli* constructs and performed cultivations. LC carried out plating. PD, LC, RF, and PIN performed flow cytometry experiments. MS performed electron microscopy. PD and LC drafted the manuscript. ZP, RC, KS, PIN, VDL and JD finalized the manuscript. All authors read and approved the final manuscript.

Author details
[1] Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic. [2] International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, 656 91 Brno, Czech Republic. [3] Systems and Synthetic Biology Program, Centro Nacional de Biotecnología CNB-CSIC, Cantoblanco, 28049 Madrid, Spain. [4] Institute of Biophysics, Academy of Sciences of the Czech Republic, v.v.i., Kralovopolska 135, 612 65 Brno, Czech Republic. [5] Department of Histology and Embryology, Faculty of Medicine, Masaryk University, 625 00 Brno, Czech Republic. [6] Department of Experimental Biology, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic.

Acknowledgements
This work was funded by the Grant Agency of the Czech Republic (P503/12/0572), the European Regional Development Fund (CZ.1.05/1.1.00/02.0123), the Ministry of Education of the Czech Republic (LO1214, LM2011028) and the European Union (REGPOT 316345).

Competing interests
Jiri Damborsky and Zbynek Prokop are founders of Enantis Ltd., a biotechnology spin-off company from Masaryk University.

Received: 25 July 2015 Accepted: 5 December 2015
Published online: 21 December 2015

References

1. Marisch K, Bayer K, Cserjan-Puschmann M, Luchner M, Striedner G. Evaluation of three industrial *Escherichia coli* strains in fed-batch cultivations during high-level SOD protein production. Microb Cell Fact. 2013;12:58.
2. Studier FW, Moffatt BA. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol. 1986;189:113–30.
3. Choi JH, Keum KC, Lee SY. Production of recombinant proteins by high cell density culture of *Escherichia coli*. Chem Eng Sci. 2006;61(3):876–85.
4. Balzer S, Kucharova V, Megerle J, Lale R, Brautaset T, Valla S. A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression in *Escherichia coli*. Microb Cell Fact. 2013;12:26.
5. Tolia NH, Joshua-Tor L. Strategies for protein coexpression in *Escherichia coli*. Nat Methods. 2006;3:55–64.
6. Xu P, Vansiri A, Bhan N, Koffas MAG. ePathBrick: a synthetic biology platform for engineering metabolic pathways in *E. coli*. ACS Synth Biol. 2012;1(7):256–66.
7. Akhtar MK, Jones PR. Construction of a synthetic YdbK-dependent pyruvate:H_2 pathway in *Escherichia coli* BL21(DE3). Metab Eng. 2009;11:139–47.
8. Wu J, Du G, Zhou J, Che J. Metabolic engineering of *Escherichia coli* for (2S)-pinocembrin production from glucose by a modular metabolic strategy. Metab Eng. 2013;16:48–55.
9. Xu P, Gu Q, Wang W, Wong L, Bower AGW, Collins CH, Koffas MAG. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. Nat Commun. 2013;4:1409.
10. Fang MY, Zhang C, Yang S, Cui JY, Jiang PX, Lou K, Wachi M, Xing XH. High crude violacein production from glucose by *Escherichia coli* engineered with interactive control of tryptophan pathway and violacein biosynthesis pathway. Microb Cell Fact. 2015;14:8.
11. Glick BR. Metabolic load and heterologous gene expression. Biotechnol Adv. 1995;13:247–61.
12. Ramchuran SO, Holst O, Karlsson EN. Effect of postinduction nutrient feed composition and of lactose as inducer during production of thermostable xylanase in *Escherichia coli* glucose-limited fed-batch cultivations. J Biosci Bioeng. 2005;99:477–84.
13. Jones KL, Kim SW, Keasling JD. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. Metab Eng. 2000;2:328–38.
14. Silva F, Queiroz JA, Domingues FC. Evaluating metabolic stress and plasmid stability in plasmid DNA production by *Escherichia coli*. Biotechnol Adv. 2012;30:691–708.
15. Mairhofer J, Scharl T, Marisch K, Cserjan-Puschmann M, Striedner G. Comparative transcriptomic profiling and in-depth characterization of plasmid-based and plasmid-free *Escherichia coli* expression systems under production conditions. Appl Environ Microbiol. 2013;79:3802–12.
16. Kosinski MJ, Rinas U, Bailey JE. Isopropyl-β-D-thiogalactopyranoside influences the metabolism of *Escherichia coli*. Appl Microbiol Biotechnol. 1992;36:782–4.
17. Perfeito L, Ghozzi S, Berg J, Schnetz K, Lässig M. Nonlinear fitness landscape of a molecular pathway. PLoS Genet. 2011;7:e1002160.
18. Van Hylckama Vlieg JET, Poelarends GJ, Mars AE, Janssen DB. Detoxification of reactive intermediates during microbial metabolism of halogenated compounds. Curr Opin Microbiol. 2000;3:257–62.
19. Haro MA, de Lorenzo V. Metabolic engineering of bacteria for environmental applications: construction of *Pseudomonas* strains for biodegradation of 2-chlorotoluene. J Biotechnol. 2001;85:103–13.
20. Bosma T, Damborsky J, Stucki G, Janssen DB. Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. Appl Environ Microbiol. 2002;68:3582–7.
21. Yadid H, Rudolph J, Hlouchova K, Copley SD. Sequestration of a highly reactive intermediate in an evolving pathway for degradation of pentachlorophenol. Proc Natl Acad Sci U S A. 2013;110:E2182–90.
22. Kurumbang NP, Dvorak P, Bendl J, Brezovsky J, Prokop Z, Damborsky J. Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. ACS Synth Biol. 2014;3:172–81.
23. Zhu MM, Skraly FA, Cameron DC. Accumulation of methylglyoxal in anaerobically grown *Escherichia coli* and its detoxification by expression of the *Pseudomonas putida* glyoxalase I gene. Metab Eng. 2001;3:218–25.
24. Pitera DJ, Paddon CJ, Newman JD, Keasling JD. Balancing a heterologous mevalonate pathway for improved isoprenoid production in *Escherichia coli*. Metab Eng. 2007;9:193–2007.
25. Ajikumar PK, Xiao WH, Tyo KEJ, Wang Y, Simeon F, Leonard E, Mucha O, Phon TH, Pfeifer B, Stephanopoulos G. Isoprenoid pathway optimization for taxol precursor overproduction in *Escherichia coli*. Science. 2010;330:70–4.
26. Dahl RH, Zhang F, Alonso-Gutierrez J, Baidoo E, Batth TS, Redding-Johanson AM, Petzold CJ, Mukhopadhyay A, Lee TS, Adams PD, Keasling JD. Engineering dynamic pathway regulation using stress-response promoters. Nat Biotechnol. 2013;31:1039–46.
27. Woodruff LBA, Boyle NR, Gill RT. Engineering improved ethanol production in *Escherichia coli* with a genome-wide approach. Metab Eng. 2013;17:1–11.
28. Jönsson LJ, Alriksson B, Nilvebrant NO. Bioconversion of lignocellulose: inhibitors and detoxification. Biotechnol Biofuels. 2013;6:16.
29. Nicolaou SA, Gaida SM, Papoutsakis ET. A comparative view of metabolite and substrate stress and tolerance in microbial bioprocessing: from biofules and chemicals, to biocatalysis and bioremediation. Metab Eng. 2010;12:307–31.
30. Dvorak P, Bidmanova S, Damborsky J, Prokop Z. Immobilized synthetic pathway for biodegradation of toxic recalcitrant pollutant 1,2,3-trichloropropane. Environ Sci Technol. 2014;48:6859–66.
31. Pavlova M, Klvana M, Prokop Z, Chaloupkova R, Banas P, Otyepka M, Wade RC, Tsuda M, Nagata Y, Damborsky J. Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. Nat Chem Biol. 2009;5:727–33.
32. Dvorak P, Kurumbang NP, Bendl J, Brezovsky J, Prokop Z, Damborsky J. Maximizing the efficiency of multienzyme proces by stoichiometry optimization. ChemBioChem. 2014;15:1891–5.
33. Yokota T, Omori T, Kodama T. Purification and properties of haloalkane dehalogenase from *Corynebacterium sp.* strain m15–3. J Bacteriol. 1987;169:4049–54.
34. van Hylckama Vlieg JE, Tang L, Lutje Spelberg JH, Smilda T, Poelarends GJ, Bosma T, van Merode AE, Fraaije MW, Janssen DB. Halohydrin dehalogenases are structurally and mechanistically related to short-chain dehydrogenases/reductases. J Bacteriol. 2001;183:5058–66.
35. Rink R, Fennema M, Smids M, Dehmel U, Janssen DB. Primary structure and catalytic mechanism of the epoxide hydrolase from *Agrobacterium radiobacter* AD1. J Biol Chem. 1997;272:14650–7.
36. Zingaro KA, Nicolaou SA, Papoutsakis ET. Dissecting the assays to assess microbial tolerance to toxic chemicals in bioprocessing. Trends Biotechnol. 2013;31:643–53.

37. Samin G, Pavlova M, Arif MI, Postema CP, Damborsky J, Janssen DB. A *Pseudomonas putida* strain genetically engineered for 1,2,3-trichloropropane bioremediation. Appl Environ Microbiol. 2014;80:5467–76.
38. Tracy BP, Gaida SM, Papoutsakis ET. Flow cytometry for bacteria: enabling metabolic engineering, synthetic biology and the elucidation of complex phenotypes. Curr Opin Biotechnol. 2010;21:85–99.
39. Amor KB, Breeuwer P, Verbaarschot P, Rombouts FM, Akkermans ADL, De Vos WM, Abee T. Multiparametric flow cytometry and cell sorting for the assessment of viable, injured, and dead *Bifidobacterium* cells during bile salt stress. Appl Environ Microbiol. 2002;68:5209–16.
40. Pinto D, Almeida V. Almeida Santos M, Chambel L: Resuscitation of *Escherichia coli* VBNC cells depends on a variety of environmental or chemical stimuli. J Appl Microbiol. 2011;110:1601–11.
41. Haidinger W, Szostak MP, Jechlinger W, Lubitz W. Online monitoring of *Escherichia coli* ghost production. Appl Environ Microbiol. 2003;69:468.
42. Nikel PI, Pérez-Pantoja D, de Lorenzo V. Why are chlorinated pollutants so difficult to degrade aerobically? Redox stress limits 1,3-dichloroprop-1-ene metabolism by *Pseudomonas pavonaceae*. Philos Trans R Soc Lond B Biol Sci. 2013;368:20120377.
43. Rui L, Kwon YM, Reardon KF, Wood TK. Metabolic pathway engineering to enhance aerobic degradation of chlorinated ethenes and to reduce their toxicity by cloning a novel glutatione *S*-transferase, an evolved toluene *o*-monooxygenase, and γ-glutamylcysteine synthetase. Environ Microbiol. 2004;6:491–500.
44. Inoue A, Horikoshi K. A *Pseudomonas* thrives in high concentrations of toluene. Nature. 1989;338:264–6.
45. Crebelli R, Andreoli C, Carere A, Conti L, Crochi B, Cotta-Ramusino M, Benigni R. Toxicology of halogenated aliphatic hydrocarbons: structural and molecular determinants for the disturbance of chromosome segregation and the induction of lipid peroxidation. Chem Biol Interact. 1995;98:113–29.
46. Laetz CA, Baldwin DH, Collier TK, Hebert V, Stark JD, Scholtz NL. The synergistic toxicity of pesticide mixtures: implications for risk assessment and the conservation of endangered pacific salmon. Environ Health Perspect. 2009;117:348–53.
47. Feng M, He Q, Meng L, Zhang X, Wang Z. Evaluation of single and joint toxicity of perfluorooctane sulfonate, perfluorooctanoic acid and copper to *Carassius auratus* using oxidative stress biomarkers. Aquat Toxicol. 2015;161:108–16.
48. Wah Chu K, Chow KL. Synergistic toxicity of multiple heavy metals is revealed by a biological assay using a nematode and its transgenic derivative. Aquat Toxicol. 2002;61:53–64.
49. Kaino T, Takagi H. Gene expression profiles and intracellular contents of stress protectants in *Saccharomyces cerevisiae* under ethanol and sorbitol stresse. Appl Microbiol Biotechnol. 2008;79:273–83.
50. Hayashi S, Aono R, Hanai T, Mori H, Kobayashi T, Honda H. Analysis of organic solvent tolerance in *Escherichia coli* using gene expression profiles from DNA microarrays. J Biosci Bioeng. 2003;95:379–83.
51. Aono R, Kobayashi H, Joblin KN, Horikoshi K. Effects of organic solvents on growth of *Escherichia coli* K-12. Biosci Biotechnol Biochem. 1994;58:11.
52. Delvigne F, Goffin P. Microbial heterogeneity affects bioprocess robustness: dynamic single-cell analysis contributes to understanding of microbial populations. Biotechnol J. 2014;9:61–72.
53. Kuhlman T, Zhang Z, Saier MH, Hwa T. Combinatorial transcriptional control of lactose operon of *Escherichia coli*. Proc Natl Acad Sci U S A. 2007;104:6043–8.
54. Gatti-Lafranconi P, Dijkman WP, Devenish SR, Hollfelder F. A single mutation in the core domain of the *lac* repressor reduces leakiness. Microb Cell Fact. 2013;12:67.
55. Fernández-Castané A, Camial G, López-Santín J. Direct measurements of IPTG enable analysis of the induction behavior of *E. coli* in high cell density culture. Microb Cell Fact. 2012;11:58.
56. Donovan RS, Robinson CW, Glick BR. Optimizing inducer and culture conditions for expression of foreign proteins under the control of the *lac* promoter. J Ind Microbiol. 1996;16:145–54.
57. Pan H, Xie Z, Bao W, Zhang J. Optimization of culture conditions to enhance *cis*-epoxysuccinate hydrolase production in *Escherichia coli* by response surface methodology. Biochem Eng J. 2008;42:133–8.
58. Neubauer P, Hofmann K, Holst O, Mattiasson B, Kruschke P. Maximizing the expression of a recombinant gene in *Escherichia coli* by manipulation of induction time using lactose as inducer. Appl Microbiol Biotechnol. 1992;36:739–44.
59. Kilikian BV, Suárez ID, Liria CW, Gombert AK. Process strategies to improve heterologous protein production in *Escherichia coli* under lactose or IPTG induction. Process Biochem. 2000;35:1019–25.
60. Andrews KJ, Hegeman GD. Selective disadvantage of non-functional protein synthesis in *Escherichia coli*. J Mol Evol. 1976;8:317–28.
61. Malakar P, Venkatesh KV. Effect of substrate and IPTG concentrations on the burden to growth of *Escherichia coli* on glycerol due to the expression of Lac proteins. Appl Microbiol Biotechnol. 2012;93:2543–9.
62. Marbach A, Bettenbrock K. *lac* operon induction in *Escherichia coli*: systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA. J Biotechnol. 2012;157:82–8.
63. Leonard E, Lim KH, Saw PN, Koffas MAG. Engineering central metabolic pathways for high-level flavonoid production in *Escherichia coli*. Appl Environ Microbiol. 2007;73:3877–86.
64. Anthony JR, Anthony LC, Nowroozi F, Kwon G, Newman JD, Keasling JD. Optimization of the mevalonate-based isoprenoid biosynthesis pathway in *Escherichia coli* for production of the anti-malarian drug precursor amorpha-4,11-diene. Metab Eng. 2009;11:13–9.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central

microorganisms

MDPI

Article

Computational Modelling of Metabolic Burden and Substrate Toxicity in *Escherichia coli* Carrying a Synthetic Metabolic Pathway

Martin Demko [1], Lukáš Chrást [2], Pavel Dvořák [3], Jiří Damborský [2,*] and David Šafránek [1,*]

[1] Systems Biology Laboratory, Faculty of Informatics, Masaryk University, Botanická 68a, 602 00 Brno, Czech Republic; 325073@mail.muni.cz

[2] Loschmidt Laboratories, Department of Experimental Biology and RECETOX, Faculty of Science, Masaryk University, Kamenice 5, Bld. A13, 625 00 Brno, Czech Republic; chr@mail.muni.cz

[3] Department of Experimental Biology, Faculty of Science, Masaryk University, Kamenice 5, Bld. A25, 625 00 Brno, Czech Republic; pdvorak@sci.muni.cz

* Correspondence: jiri@chemi.muni.cz (J.D.); safranek@fi.muni.cz (D.Š.)

Received: 30 September 2019; Accepted: 7 November 2019; Published: 11 November 2019

Abstract: In our previous work, we designed and implemented a synthetic metabolic pathway for 1,2,3-trichloropropane (TCP) biodegradation in *Escherichia coli*. Significant effects of metabolic burden and toxicity exacerbation were observed on single cell and population levels. Deeper understanding of mechanisms underlying these effects is extremely important for metabolic engineering of efficient microbial cell factories for biotechnological processes. In this paper, we present a novel mathematical model of the pathway. The model addresses for the first time the combined effects of toxicity exacerbation and metabolic burden in the context of bacterial population growth. The model is calibrated with respect to the real data obtained with our original synthetically modified *E. coli* strain. Using the model, we explore the dynamics of the population growth along with the outcome of the TCP biodegradation pathway considering the toxicity exacerbation and metabolic burden. On the methodological side, we introduce a unique computational workflow utilising algorithmic methods of computer science for the particular modelling problem.

Keywords: biodegradation; computational modelling; population growth; metabolic burden; environmental pollutants

1. Introduction

Escherichia coli strain BL21(DE3) is frequently used in synthetic biology with embedded protein expression (pET) vectors enabling heterologous protein expression [1–6]. Especially, it has recently found use as a cell factory for heterologous expression of entire biochemical pathways [7–11]. In spite of its common usage in metabolic engineering, the expression system in *E. coli* BL21(DE3) suffers from certain problems primarily caused by strong overexpression of recombinant proteins triggered by exposing the cognate $LacI^{Q}/P_{lacUV5}$-T7 expression system to the synthetic inducer isopropyl-beta-D-thiogalactopyranoside (IPTG) [12]. Such negative effects that result from prioritising high levels of protein production to normal metabolic capacities in host cells are known as the *metabolic burden* [13,14]. One of the known factors causing the burden is the energetically expensive maintenance and replication of plasmid vectors carrying heterologous genes [15–17]. Other factors are associated with the activities of the foreign proteins which may interact with the metabolic network of the cell and burdens linked to the components of the expression system itself, such as IPTG. The individual factors may not always be additive, but can sometimes potentiate each other, leading to significantly larger effects on the living cells—the so-called *exacerbation effect*.

In our previous work, we designed and implemented a synthetic metabolic pathway for TCP biodegradation in *E. coli* [18]. The pathway was assembled with enzymes introduced to the cell using the pETDuet plasmids with $LacI^{Q}/P_{lacUV5}$-T7 expression system inducible by IPTG. In relation to that work, we observed the effect of metabolic burden and toxicity exacerbation on *E. coli* population in [19]. However, the mechanisms behind these effects and their influence on cell growth have not been targeted. Deeper understanding of underlying mechanisms is extremely important for metabolic engineering of efficient microbial cell factories for biotechnological processes. To the best of our knowledge, the metabolic burden effect has not yet been fully-handled regarding the dynamics of affected metabolites and the influence on cell population growth. In [20], the authors reviewed existing fundamental frameworks for the kinetic modelling of microbial growth based on basic hypotheses about the underlying reaction systems. An integrated model presented in [21] couples the growth rate with the gene expression and the growth rate with the growing population of cells. To the best of our knowledge, existing population-level growth models neither consider metabolic interactions together with toxic effects caused by some metabolites nor describe a burden linked with using of engineered metabolic pathways. It is apparent that such effects cause a significant increase of the complexity of the underlying non-linear dynamics. To that end, a precise explanation of the resulting emergent behaviour remains a challenge.

The approach of computational systems biology gives a powerful tool allowing to study the dynamics of biological mechanisms holistically in terms of mathematical models. An important aspect of modern systems biology is the requirement to ground the models within the relevant genomic context of the explored organism [22]. It is obvious that utilising mathematical modelling and computational analysis based on the models makes an important starting point for successful experiment design in synthetic biology. To that end, the synthetic pathway we designed for conversion of the highly toxic TCP to glycerol (GLY) in *E. coli* [23] was optimised by using our original mathematical model [18] that allowed us to simulate the behaviour of the synthetic pathway *in silico*. The synthetic pathway is depicted in Figure 1. It is composed of a few toxic intermediates with harmless glycerol as a final product and it utilises enzymes from other bacterial species. These enzymes represent the engineered form of haloalkane dehalogenase (DhaA31, originally from *Rhodococcus rhodochrous* NCIMB 13064) and haloalcohol dehalogenase (HheC), and epoxide hydrolase (EchA) from *Agrobacterium radiobacter* AD1. They have a major role in this pathway; however, since they are heterologous proteins in *E. coli*, they have to be produced at the expense of other substances (causing a particular instance of the metabolic burden).

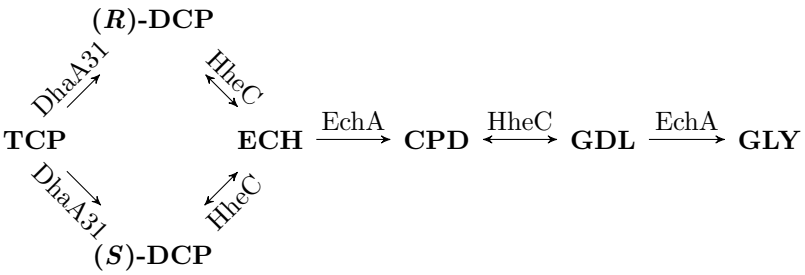


Figure 1. Model of metabolic pathway for biodegradation of TCP. A general scheme of an enzymatic metabolic pathway for biodegradation of TCP into GLY. Note that DhaA31 produces two different enantiomers of 2,3-dichloropropane-1-ol (DCP) with similar rate. However, enzyme HheC has notably different enantioselectivity with them. It is also worth noting that enzymes HheC and EchA are employed twice in the pathway. Other intermediates are epichlorohydrin (ECH), 3-chloropropane-1,2-diol (CPD), and glycidol (GDL).

In this paper, we propose a significant extension of the existing mathematical model of the pathway which addresses for the first time the combined effects of toxicity exacerbation and metabolic burden in the context of bacterial population growth. To calibrate and fine-tune the model with respect to our original synthetically modified *E. coli* strain, we conducted several dedicated wet-lab experiments. Based on the new model calibrated with respect to the real data, we explore the dynamics of the population growth along with the outcome of the TCP biodegradation pathway considering all the phenomena mentioned above. On the methodological side, we introduce a unique computational workflow utilising algorithmic methods of computer science. It allows us to fully exploit the expressive power of the model under parameter uncertainty.

2. Materials and Methods

2.1. Laboratory Methods

2.1.1. Bacterial Strains and Plasmids

Escherichia coli strain BL21(DE3) was used for this study with two modifications. Here, they are called deg31 (carrying synthetic metabolic pathway containing recombinant plasmids pCDF::*dhaA31* and pETDuet::*echA-hheC*) and host (carrying empty plasmids pCDF and pETDuet).

2.1.2. Preparation of Pre-Induced Cells

The *lysogeny broth* medium (i.e., LB medium) with volume of 10 mL containing respective antibiotic combination (75 µg/mL ampicillin, 25 µg/mL streptomycin) was inoculated with transformed cells from glycerol stock. The culture was incubated overnight at 37 °C with shaking (180 rpm) in incubator Innova 44 (New Brunswick Scientific, Edison, New Jersey, USA—this machine was used for all incubations). Then, 25 mL of fresh LB medium with respective antibiotics were inoculated with 250 µL of night culture and incubated at 37 °C with shaking (180 rpm) until OD_{600} reached 1. The cultures were induced with 0, 0.01, 0.05, 0.2 and 1 mM IPTG and incubated overnight at 20 °C. Cells were then collected at late exponential phase by centrifugation (4000× *g*) at 4 °C in centrifuge Sigma 6-16K (Merck, Darmstadt, Germany—this machine was used for all centrifugation) and washed with 50 mM sodium phosphate buffer (NaP buffer, pH 7). After washing and centrifugation, cells were resuspended in NaP buffer to final value of 7 OD_{600}.

2.1.3. Short Term Toxicity Test

Prior to the toxicity test, 4 mM TCP was diluted in 5 mL of NaP buffer in 25 mL sterile Reacti Flasks closed by screw caps and incubated for 1 h at 37 °C. The reaction was initiated by mixing preincubated buffer with TCP with 5 mL of cells in NaP buffer, resulting in final OD_{600} 3.5 and 2 mM concentration of TCP. Reaction suspension was incubated in water bath at 37 °C with constant shaking for 5 h.

Cell samples were taken from the reaction mixture at the beginning of the toxicity test and after 4 h of incubation. Cell suspension (100 µL) was aseptically withdrawn from the flask and serially diluted in 900 µL of PBS buffer (pH 7.4) up to the final dilution of 10^{-6} to 10^{-7}. Diluted cell suspensions (100 µL) was spread on Plate Count Agar (PCA) plates and incubated 24 h at 37 °C. After incubation, colonies on agar plates were manually counted and expressed as colony forming units per volume (CFU/mL).

2.1.4. Growth Test

Cells of *E. coli* BL21(DE3) strain carrying empty plasmids and degrading strain deg31 were pre-grown in LB medium containing respective antibiotic combination (75 µg/mL ampicillin and 25 µg/mL streptomycin) at 37 °C until the culture reached stationary phase. The cells were then centrifuged at 6000× *g* and resuspended in Synthetic Mineral Medium (SMM) [18]. SMM medium

(15 mL) with 10 mM glycerol in 25-mL Reacti Flask closed with screw cap was preincubated at 37 °C, growth test was initiated by addition of cells to 0.1 OD_{600}. Growth medium was supplemented with IPTG to the final concentrations 0, 0.01, 0.05, 0.2 and 1 mM. Cells were incubated at 37 °C with shaking (200 rpm). Optical density (OD_{600}) of the culture was measured periodically and cell dry weight (CDW) in g/L at given time intervals was determined by multiplying the OD_{600} values by 0.39 g/L [24]. Acute toxicity measurement was performed in the same conditions, and the medium for toxicity test contained respective concentrations of TCP pathway metabolites.

2.1.5. Glycerol Analysis

Samples withdrawn from incubated cell cultures were heated for 10 min at 95 °C, centrifuged and stored in the fridge prior to analysis. Free Glycerol Colorimetric/Fluorometric Assay kit (BioVision, USA) was used for analysis of glycerol content in the cells withdrawn from toxicity or growth tests following manufacturer's instructions.

2.2. Parameter Constraints

We consider a set of m unknown parameters $p_1, ..., p_m$ and the set $\mathbb{P} \subseteq \mathbb{R}^m_{\geq 0}$ denoting the so-called *parameter space* consisting of m-dimensional vectors of parameter values. For the purpose of this paper, we assume $\mathbb{P}$ to be bounded. The constraints on parameters could generally be of various types. In this paper, we consider linear constraints of the form:

$$a_1 \cdot p_1 + \cdots + a_m \cdot p_m \sim b \tag{1}$$

where $p_j \in \mathbb{R}_{\geq 0} : j \in \{1, \ldots, m\}$ is a parameter; $a_j, b \in \mathbb{R}$ stand for coefficients; and $\sim \in \{<, >, \leq, \geq, =, \neq\}$ represents (in)equalities.

2.3. Biochemical Model

A biochemical model, in this context, is a dynamical system given as a set of *ordinary differential equations* (ODEs) of the form

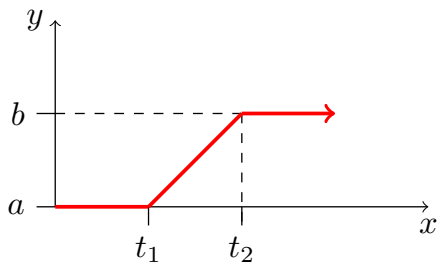
$$\dot{x}_i = f_i(\vec{x}, \vec{p}) : i \in \{1, \ldots, n\} \tag{2}$$

$$\vec{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n_{\geq 0} \tag{3}$$

$$\vec{p} = (p_1, \ldots, p_m) \in \mathbb{P} \tag{4}$$

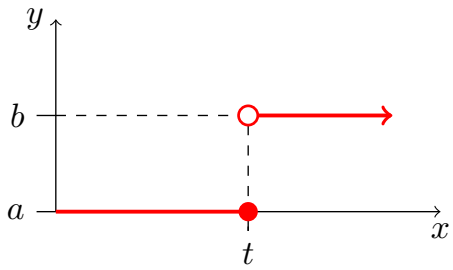
where $\vec{x}$ is a vector of n variables, $\vec{p}$ is a vector of m model parameters, and f_i is a kinetic function constructed as a sum of reaction rates where every sum member represents an *affine* or *bi-linear* function of x_i; a *ramp* or a *Heaviside step* function of x_i (Figures 2 and 3, respectively); or a function from a limited set of *non-linear* functions of x_i referred as sigmoidal functions; all of them possibly containing own set of internal parameters.

In general, our framework covers *mass action* kinetics ([25], Section 1.1) with stoichiometric coefficients not greater than one and all biologically relevant non-linear functions such as *Michaelis–Menten* [26] or *Hill* kinetics [27], and microbial growth kinetics such as *Monod* [28]. An additional requirement restricts the role of parameters in kinetic functions; in particular, we require f_i is affine in $\vec{p}$.



$$R(x, t_1, t_2, a, b) = \begin{cases} a & x < t_1 \\ a + (b - a) \cdot \frac{x - t_1}{t_2 - t_1} & t_1 \leq x \leq t_2 \\ b & x > t_2 \end{cases}$$

Figure 2. Definition of the ramp function. (**Right**) A mathematical definition of an increasing ramp function as used in our workflow. Parameters a and b are usually set to values 0 and 1, respectively; and vice versa for decreasing version. Values t_1 and t_2 typically represent some significant thresholds on x. (**Left**) Graphical description of the same function.



$$H(x, t, a, b) = \begin{cases} a & x \leq t \\ b & x > t \end{cases}$$

Figure 3. Definition of the Heaviside step function. (**Right**) A mathematical definition of an increasing Heaviside step function as used in our workflow. Parameters a and b are usually set to values 0 and 1, respectively; and vice versa for decreasing version. Value t typically represents an important threshold in the domain of x. (**Left**) Graphical description of the function—specifically, the increasing version.

2.4. Inverse Problem

An inverse problem in mathematics is a process of finding a good model (with parameters) reflecting given data; mathematically speaking:

$$\mathcal{M}(\pi) = d \tag{5}$$

where π is a vector of parameters of a model $\mathcal{M}$ and d represents the given data.

Intuitively, the problem is the following: "How to find the proper set of parameter values π given a model $\mathcal{M}$." In this paper, we distinguish two classes of the inverse problem—*parameter estimation* and *parameter synthesis*.

Parameter estimation is based on optimal fitting of the model parameters to observed experimental data. In the basic case, we assume the model to be a parameterised curve.

Parameter synthesis is a method that allows identification of satisfiable parameter values with respect to a given set of hypotheses restricting systems dynamics (described in a particular formalism) and a priori known parameter constraints (e.g., correlations of parameter values, constraints on production/degradation ratio, etc.). In this case, the parameters are assumed to be the model parameters (typically, the rate coefficients appearing in kinetic functions) [29–40].

2.4.1. Parameter Estimation and Regression

Parameter estimation is a well-known process for the identification of the model parameters from experimental data representing the observations of the system. The particular problem of finding a parameterised curve (i.e., function) that approximately fits a set of data is referred to as *regression*. A function is classified as either linear or non-linear concerning affinity of the fitted parameters. Consequently, the problem is classified as the *linear regression* or the *non-linear regression* according to the function class. In general, the linear regression is easier but of limited effectiveness and it is

preferred if we know the function. In this paper, almost all concerned functions (i.e., models) belong to the non-linear class. Hence, we assume only non-linear regression further referred to as regression or fitting (for simplicity) [41].

2.4.2. Parameter Synthesis and Robustness Monitoring

Both approaches rely upon a *temporal logic* for specification of the behaviour properties of dynamical systems (Section 2.3). Although several different temporal logics have been developed, they can all be determined in the context of paths (or runs) of the system, i.e., the infinite sequences of states describing the system dynamics in consecutive time events. Such sequences are encoded in temporal logic *formulae* [42].

Parameter synthesis performs comprehensive exploration of the continuous parameter space including the space of initial conditions. It provides a *qualitative* answer to the question "which settings of a model satisfy a given set of temporal properties". It is not as widely used as parameter estimation and therefore many of the developed tools are still in a prototype phase [29–40].

Robustness monitoring enables *quantitative* evaluation of the robustness of a model with respect to a temporal property under some perturbation of parameters (typically, reaction rate constants or initial conditions) [43–48].

2.5. Analysis Workflow

Mathematical models in synthetic biology are often represented in terms of the ODEs system. These models are quantitative and their functionality relies on particular parameter values. Parameter estimation is often the only solution to obtain proper parameter values that fit with experimental observations. In this section, we present our methodology step by step. The presented workflow is an extension of the workflow we introduced in [37].

2.5.1. General Assumptions

There are several preliminary assumptions upon which the workflow is formulated. First, the model must be a metabolic pathway (i.e., a linked series of chemical reactions catalysed by enzymes [49]). Next, we consider several assumptions limiting the experimental settings of the studied system:

1. We assume the total inducer concentration to be constant in the time frame of our interest. An inducer is supposed to have a function of an input parameter, and it would be an inadequate parameter should it be adjusted spontaneously over time. Otherwise, the inducer degradation rate would be needed either found in literature or extracted from experimental data.
2. The workflow is limited to protease-deficient bacterial strains (e.g., *E. coli* BL21). In particular, we assume the total concentration of every enzyme affecting the studied pathway is constant in the time frame of our interest. Moreover, no influx of the enzymes is permitted as a consequence of time-limited induction phase where the proteosynthesis takes place [50,51]. Additional synthetic processes are considered negligible in a microbial population stressed enough by the massive expression of (heterologous) genes during induction.
3. There occurs a metabolic burden effect caused by the heterologous genes expression during the induction process and possibly by the presence of an inducer itself which in a combined way affects the bacterial growth rate.
4. Finally, we assume the bacterial population is in the stationary phase after the induction process is finished.

These assumptions limit the use of the proposed workflow to studies of the synthetic metabolic pathway models describing the metabolite dynamics in a time frame of a few hours. Moreover, the workflow targets the protease-deficient bacterial strains assuming some of the enzymes are products of expression of heterologous genes initiated by a non-metabolisable inducer during the induction process. It is worth noting that these significant assumptions help to keep the mathematical model

computationally feasible in terms of the number of potentially unknown parameters (thus, minimising the risk of over-parameterisation). In particular, the degradation rates of enzymes can be neglected in such settings.

2.5.2. Workflow Description

In the previous section, we define a metabolic pathway as a chain of chemical reactions with metabolites as the inputs and outputs of the reactions catalysed by enzymes. This form of the system is assumed to be the *input* to the workflow. The mathematical vector of concentrations of all model enzymes can be understood as the specific point in the *enzymatic space*.

In *Step 1* of the workflow, we focus on the reduction of the enzymatic space. Assuming that some of the enzymes are products of the induction process, their concentration can be expressed as the set of functions with an inducer as the input. There is one such function for each enzyme. The internal parameters of these functions need to be extracted out of experimental data—measured in various initial concentrations of the inducer—possibly with the help of parameter estimation methods. Typically, there are more enzymes than inducers in biochemical reactions. This step eliminates the extent of dependency of the model on enzymes and it provides the reduction of the number of potential model parameters.

Up to this point, the model was considered without any effect on the bacterial population. We propose to monitor and possibly predict an impact of the pathway on the population—such as the metabolic burden if using plasmids carrying heterologous genes—instead of modelling entire bacterial system. Following this idea, in *Step 2* of the workflow, we extend the current model with new variables describing: (1) the bacterial population; and (2) a substrate used to feed the population. In general, there can be more sources of nutrition. The dynamics of the extension is defined by the *growth rate* and *death rate* functions. Each rate function must depend at least on the particular substrate and might depend on the inducer and any of the metabolites especially if they have a harmful effect on the population. To extract a proper rate function, fitting of the internal parameters to experimental data is needed. The traditionally used growth functions include Monod, Moser, Tessier, etc. [52]. They can also be used for diauxic growth ([53], Section 12.1.3). At the cost of expanding the model with a few new variables, our framework allows predicting the population development depending on settings of the model parameters.

To fine-tune the model, we can use the formal computational methods of parameter synthesis mentioned in Section 2.4.2. This approach is more efficient than iterative analysis of demanding laboratory experiments used for fitting. Thus, in *Step 3* of the workflow, we need to specify a set of relevant parameters—ideally, the coefficients that can be altered in an experiment design allowing validation of the model-based results. Besides that, we need to formulate the temporal properties tailored to the particular case study. In general, this requires having some a priori knowledge about the investigated system. Nevertheless, there exist several typical (template) properties which can be used universally (i.e., oscillation, bistability, etc.). By using parameter synthesis together with robustness monitoring, we can: (i) specify various hypothetical scenarios where experimental data are missing or even impossible to achieve in laboratory conditions; (ii) monitor the behaviour of the investigated model; and (iii) optimise a set of tunable model parameters.

2.6. Software Tools

For the fitting procedure, we use several tools. GraphPad Prism (https://www.graphpad.com/) (version 8.2.1 (441)) is used in Section 3.3 to fit the functions describing the stable concentration level of enzymes induced by IPTG. The tool GUI is easy to use for traditional functions such as Michaelis–Menten. However, to investigate and compare the usability of a comprehensive set of functions, we use a more advanced tool—the FME package (https://cran.r-project.org/web/packages/FME/index.html) (version 1.3.5) [54] of the R language [55]. It offers well-documented procedures: parameter identifiability; parameter sensitivity; parameter fitting; and more.

For the parameter synthesis procedure, we utilise our tool Pithya [40], providing a parameter synthesis algorithm working with an expressive temporal logic HUCTL_{P} [39,56]. Its main advantage is it enables a fully automatised investigation of the system dynamics including exploration of equilibria. The tool has GUI and CLI, both available on github (https://github.com/sybila/pithya-gui) and also online (http://pithya.ics.muni.cz).

For the robustness monitoring, we use our tool Parasim that implements the algorithm working with a signal temporal logic STL^* [46–48] expressing properties of continuous signals. STL^*operates with forward consecutive time events over real-valued signals (i.e., evaluated variables). Its strength is to identify various types of repetitive behaviour (i.e., oscillations). The tool is available on github (https://github.com/sybila/parasim) and makes part of the online toolset BioDivine (http://pithya.ics.muni.cz/galaxy).

Both tools were successfully used on various biological case studies but so far they have been employed only separately (Pithya [34,37,38,56–60] and Parasim [46,48]). In this paper, we combine their application on a single model for the first time. Using the input files available online in shared history notebooks in the BioDivine framework, the results conducted in Section 3.5 with Parasim (http://pithya.ics.muni.cz/galaxy/u/martin/h/parasim-on-tcp-model-public) and Pithya (http://pithya.ics.muni.cz/galaxy/u/martin/h/pithya-on-tcp-model-public) can be reproduced.

3. Results and Discussion

In this section, we apply the sequence of steps defined in Section 2.5 to the model of TCP metabolic pathway (Figure 1).

The outcome of the workflow is a novel model, the scheme of which is shown in Figure 4. It exhibits a modular structure, which was necessary to relieve the inverse modelling process (Section 2.4). We decided on this concept because the fitting of all new parts simultaneously would be practically impossible. Instead, only two to four parameters needed to be taken into account simultaneously. Due to this fact, the dedicated experimental data—capturing only partial behaviour—had to be obtained (Data S1–S3).

3.1. Extended Assumptions

Our model satisfies the general theoretical assumptions defined in Section 2.5.1: (1) the standard non-metabolisable IPTG inducer has been used [61]; (2) *E. coli* BL21 (DE3) is a protease deficient B strain, and all experiments were conducted in a closed environment with a limited amount of nutrients; (3) the core of the model is an expression of heterologous genes causing experimentally-provable metabolic burden [19]; and (4) all experiments were preceded by the induction phase after which the bacterial population was in stationary phase. Next, we assume the time frame of 5 h in which all metabolic pathway experiments and simulations were performed.

Next, we specify an extended set of assumptions characterising this case study:

1. We define a new variable called Bact standing for CDW (g/L) of *E. coli* population taken as 0.39 g/L = 1 OD_{600} [24].
2. We assume IPTG to be the only inducer for the synthetic pathway. Concentration of IPTG is considered to be constant in the given time frame.
3. Reversible reactions in the TCP-degradation pathway are considered negligible.
4. The initial concentration of substances (e.g., TCP_0, GLY_0, $IPTG_0$) and the population (i.e., $Bact_0$) determines the input for the system.
5. Dynamics of individual enzymes is approximated as a constant function of time in the given time frame. Moreover, enzymes dynamics is considered to be independent on the size of the bacterial population in the given time frame.
6. Total conversion of TCP into GLY is assumed to occur in a sufficiently long time reflecting the known behaviour of the pathway.
7. Viability of the bacterial population is given as the function of the pathway compounds toxicity, metabolic burden and the presence of nutrients (i.e., GLY).
8. Toxic effects of the pathway compounds are considered to be mutually independent.
9. Glycerol is the only assumed nutrient.
10. We assume natural degradation (death rate) of the bacterial population.

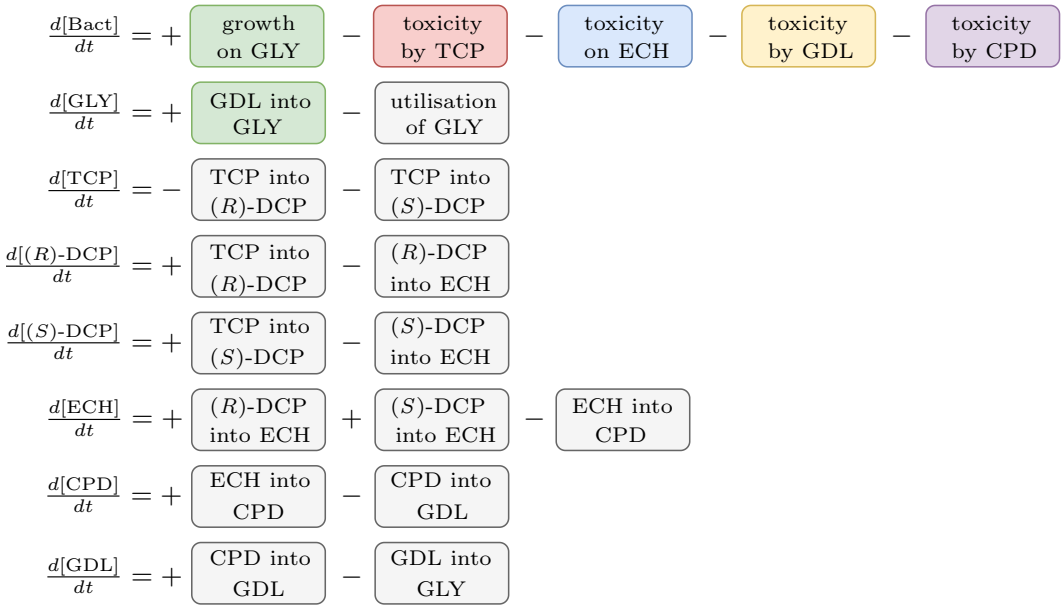


Figure 4. Proposed model scheme. A schematic description of the novel ODE model with highlighted extending partitions. Each partition (i.e., module) represents a linearly independent part of a particular equation. Each module has a unique semantic function. Note that some of the modules are used in more than one equation. The coloured modules represent entirely new parts of the final model and the grey modules were extended in this study. This modular feature was necessary to handle the fitting of such a complex model to the experimental data.

We are aware of the fact that Assumption (5) makes a significant approximation. The assumption reflects our former studies, in which we worked with pre-cultured pre-induced resting cells. In particular, estimated enzyme concentrations have been set as the endpoint values determined at a certain time interval after the overnight induction (Section 2.1.2).

3.2. Workflow Input: Synthetic TCP Degradation Pathway

Dvorak and co-workers [18] tested three different forms of the enzyme DhaA (two of them mutant) in order to improve the efficiency of the pathway. Nevertheless, we consider only the most effective form denoted DhaA31 (referred to here as DhaA for simplicity). This particular enzyme produces two different enantiomers of the DCP compound with a similar rate ($k_{\text{cat,TCP,}(R)\text{-DCP}} = 0.58\ (\text{s}^{-1})$ as the 55% of 1.05 (s^{-1}) for (*R*)-DCP and $k_{\text{cat,TCP,}(S)\text{-DCP}} = 0.47\ (\text{s}^{-1})$ as the 45% of 1.05 (s^{-1}) in favor of (*S*)-DCP). However, the enzyme HheC is enantioselective and prefers (*R*)-DCP ($k_{\text{cat,}(R)\text{-DCP}} = 1.81\ (\text{s}^{-1})$ vs. $k_{\text{cat,}(S)\text{-DCP}} = 0.08\ (\text{s}^{-1})$, where the greater means the better). As a consequence, (*S*)-DCP accumulates during the biodegradation process until (*R*)-DCP is consumed. Note that both enantiomers are supposed to be equally toxic to the host.

In general, reactions catalysed by HheC are reversible. However, in this particular case, the *catalytic efficiency* (i.e., $\frac{k_{\text{cat}}}{K_{\text{M}}}$) of EchA in turning ECH into CPD is much higher than the catalytic efficiency of HheC towards the reaction ECH $\rightarrow$ (*R*,*S*)-DCP. Due to this fact, the reverse reaction was removed from the mathematical model ([23], Chapter 2). HheC also catalyses the reaction CPD $\leftrightarrow$ GDL. The reverse reaction is compensated by a special kind of competitive version of Michaelis–Menten equation in the reaction GDL $\rightarrow$ GLY of the mathematical model. Thorough research can be found in [23]. However, we did a comparative test with several simulations of the model with competitive version of Michaelis–Menten from ([23], Chapter 2) and the simplified model containing only common Michaelis–Menten equations and, according to the results in Figure S1, we decided to use the simplified model instead (Figure S2).

3.3. Step 1: Enzymatic Space Settings and Reduction

Based on the simplifying assumptions mentioned above, the concentration levels of DhaA, HheC and EchA are represented as constants in the input mathematical model (Figure S2). In such settings, these constants can be understood as parameters. To extend the model with respect to IPTG concentration and in agreement with *Step 1* in Section 2.5.2, it is necessary to define functions describing the concentrations of enzymes depending on the concentration of IPTG. We extracted these functions by parameter fitting to experimental data obtained from densitometric analysis of enzyme cell-free extracts prepared from cells induced with different IPTG concentrations (more details in Appendix A). To that end, we employed Michaelis–Menten kinetics because the data showed a good agreement with its shape (Figure 5), although the MM is typically used for the description of a rate and not a concentration. The quantitative result of the fitting with statistical evaluation is shown in Table S1.

The initial growth of the functions is in perfect agreement with the *switch-like* influence when using a non-metabolisable inducer such as IPTG on $\text{LacI}^{\text{Q}}/P_{lacUV5}$-T7 expression system [62,63] which is employed in heterologous plasmids containing genes of enzymes DhaA, HheC and EchA. The resulting functions of stable concentrations for particular enzymes are shown in Equations (6)–(8), respectively:

$$\text{DhaA}_{total} = \frac{V_{\text{max,D}} \cdot [\text{IPTG}]}{K_{\text{M,D}} + [\text{IPTG}]} \tag{6}$$

$$\text{HheC}_{total} = \frac{V_{\text{max,H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \tag{7}$$

$$\text{EchA}_{total} = \frac{V_{\text{max,E}} \cdot [\text{IPTG}]}{K_{\text{M,E}} + [\text{IPTG}]} \tag{8}$$

with $V_{\text{max,D}} = 0.001904$, $V_{\text{max,H}} = 0.005391$, and $V_{\text{max,E}} = 0.004998$ being maximum rate constants (s^{-1}) and $K_{\text{M,D}} = 0.01749$, $K_{\text{M,H}} = 0.008255$, and $K_{\text{M,E}} = 0.004855$ being Michaelis constants (mM).

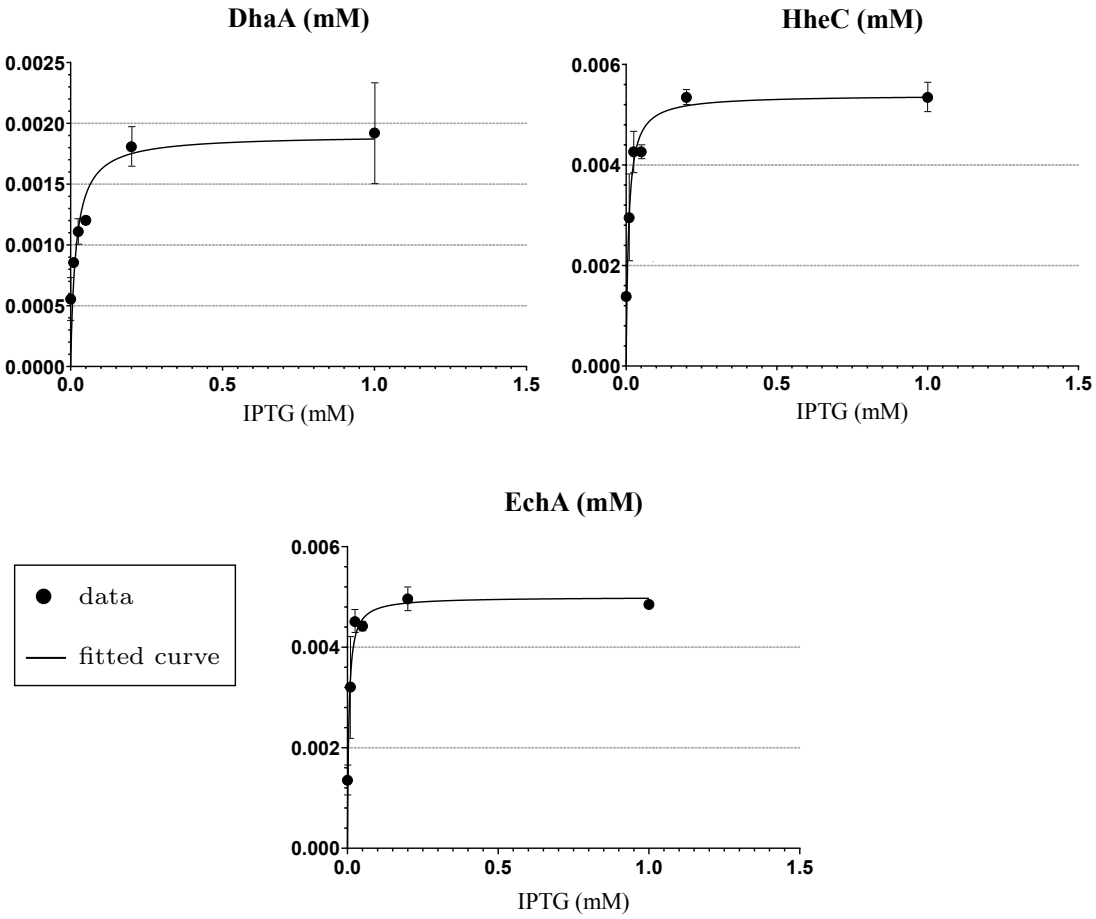


Figure 5. Results of fitting to enzymes concentration data. Three plots show individual results of fitting to concentration data measured in various starting concentration levels of the inducer ($IPTG_0$)—0.0, 0.01, 0.025, 0.05, 0.2 and 1.0 (mM)—for three different enzymes: DhaA (**top left**); HheC (**top right**); and EchA (**bottom**). The experimental data are pictured as points and the results—fitted curves—are pictured as lines. Both axes show a concentration level in mM, the inducer on the x-axis and the particular enzyme on the y-axis of the particular plot. Error bars represent standard deviation values calculated from two independent experiments.

3.4. Step 2: Integration with Population Growth

The original model does not take into account the bacterial population. Therefore, in agreement with *Step 2* in Section 2.5.2, we introduce a mechanism explaining the dependency of the substrate (GLY) and the inducer (IPTG) on the population. The Monod equation is commonly used to explain bacterial population growth rate μ, and, indeed, the results are in good agreement with experimental data (Figure 6). For the sake of completeness, we have exemplified several different alternatives (Tessier ([64], in French), Moser [65], Aiba–Edwards [66], Andrews [67], etc.). For the full list of considered functions, see the comparative table (Table S2). Nevertheless, according to Figure S3, the best results were observed by Monod growth model defined in Equations (9) and (10), where μ is the specific growth rate, μ_{max} is the maximum specific growth rate, K is the half-velocity constant, γ_{GLY} is the rate of substrate utilisation and Y is the yield coefficient.

$$\mu = \frac{\mu_{max} \cdot [\mathrm{GLY}]}{K + [\mathrm{GLY}]} \tag{9}$$

$$\gamma_{\mathrm{GLY}} = \frac{-1 \cdot \mu}{Y} \tag{10}$$

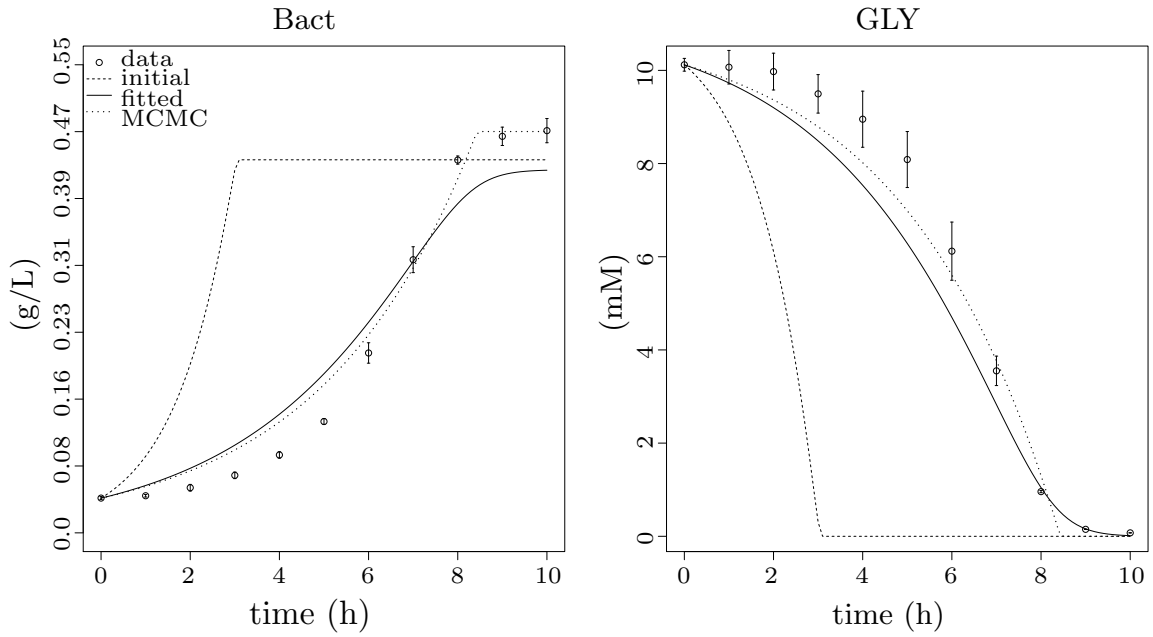


Figure 6. Results of fitting to population growth data. Two plots showing different results of fitting to population growth data from two points of view: (**left**) the particular results for CDW starting at 0.0429 and ending at 0.468 g/L after 10 h of growth; and (**right**) the result for the substrate utilisation only starting at 10.12 and ending at 0.07 mM after 10 h of growth. The experimental data are pictured as points with standard error bars, the dashed lines show simulation data for initial values of Monod function (i.e., initial point of fitting), the solid lines show the results of fitting (Section 2.4.1) and the dotted lines represent the final results optimised by MCMC method of the FME package (Section 2.4.1), which show the best agreement with the experimental data. The x-axes show time of experiment in hours; the y-axis of the right plot shows the concentration of GLY in mM; and the y-axis of the left plot shows CDW in g/L of bacterial population (Bact).

Traditional growth functions seem not to be sufficient to explain the effect of the metabolic burden caused by IPTG and heterologous genes expression, as displayed in Figure S4. Thus, considering the experimental data in Figure 7, we decided to use a specific function to model the gap between results for values $t_1 = 0.01$ and $t_2 = 0.05$ of mM IPTG_0. It seems that the Heaviside step function (Figure 3) is a straightforward option for the apparent step in-between resulting values. However, as we do not know the exact value of the breaking point, it is safer to assume the linear behaviour. Hence, we decided to employ the ramp function defined in Figure 2 in the way that the maximum specific growth rate constant μ_{max}—estimated above using Monod growth function (Figure 6)—is replaced with the ramp function defining an actual growth rate from the interval of the maximum (μ_{max}) and the minimum specific growth rate (μ_{min}) parameterised with IPTG_0. The best result of the fitting is shown in Figure 8 and represents evidence of the metabolic burden caused by IPTG with heterologous genes expression.

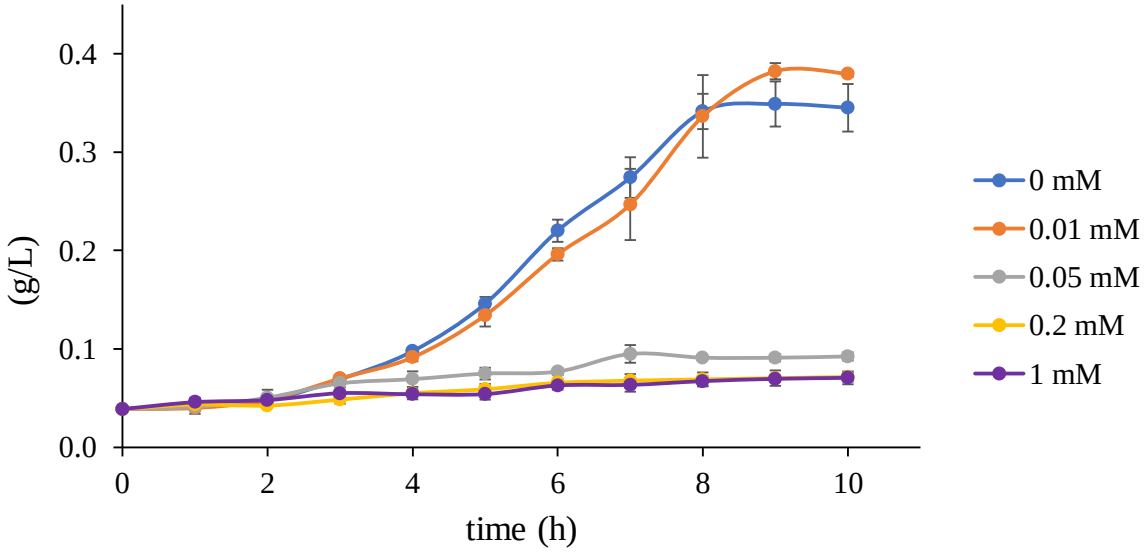


Figure 7. Bacterial population growth data reflecting concentration of inducer. The plot shows curves of population growth on GLY for cultures reflecting different concentrations of IPTG prepared according to Section 2.1.4. All cultures—carrying plasmids with heterologous metabolic pathway—started at the same value but ended with different size of population depending on the initial concentration of the inducer ($IPTG_0$): 0, 0.01, 0.05, 0.2, and 1 (mM). Note that the rising amount of IPTG led to progressive inhibition of bacterial growth. The most interesting is the big step from 0.01 to 0.05 of IPTG. This notable difference in the population on the relatively small interval of IPTG values and minimal changes in the population for the rest of IPTG concentrations shows the high sensitivity of the population to $IPTG_0$.

For the completeness, we also assume a death rate coefficient (γ_{Bact}), the value of which fits approximately in the range of $[3.876 \cdot 10^{-3}, 5.382 \cdot 10^{-4}]$ (h^{-1}). The uncertainty of these values comes from the fact that the death rate is usually not the main interest of microbiologists, and also some methods of measurement make the death phase hard to discern ([68], Section 3.1.4). The origin of the range values is explained in Appendix B and has the base in [69]. Although the range is an approximation, it is a good starting point for further investigation by parameter synthesis and robustness monitoring.

The decision of taking microbial population into account would not be complete without an understanding of the toxic effect of the pathway components. Here, the selection of the optimal function is not straightforward because no function has been made a standard for this purpose. Therefore, we investigated several functions (Appendix C). The best results are shown in Figures S5–S8 for TCP, ECH, CPD and GDL, respectively. Both DCP enantiomers were observed to have minimal effect on the population even for high doses (i.e., 4 mM).

With respect to the nature of the experimental data, some of the results were far from the best fit. Notably, the case of time-series data for TCP concentration of 2 mM was the worst (Figure S5). However, in the long time horizon, the population stabilises more or less on the same CDW (Figure S9). This fact indicates some delay in the cellular response to the presence of the toxic pollutant.

As we are interested in an explanation, or at least a mathematical description of the exacerbation effect, we introduced the Heaviside step function in the expression describing the TCP toxic effect on the bacterial population regarding of IPTG concentration. In other words, the purpose of the step function is to improve the TCP toxicity function according to an observation from the experimental data such that it simulates the exacerbation when TCP and IPTG effects are combined.

In one case, a simple evidence of the exacerbation phenomenon has been given in [19] where Dvorak and co-workers compared the results of similar experiments of pre-induced (IPTG +) or

non-induced (IPTG −) cells incubated in buffer with TCP (TCP +) or without (TCP −) (Figure 9 displays all four combinations: −−, −+, +−, and ++). The figure clearly shows some unexplained disruption amplifying the population extinction, although the data are not sufficiently convincing.

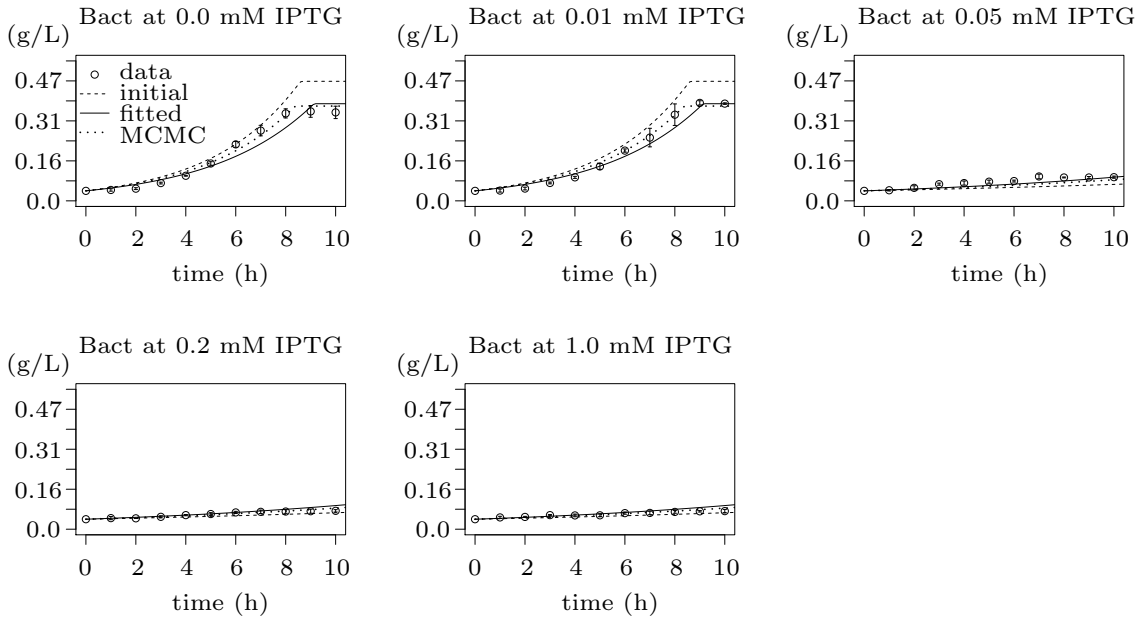


Figure 8. Results of fitting to population growth data reflecting metabolic burden caused by IPTG. The plot contains five figures, each showing fitting of the same model to the bacterial population growth data for different concentration of the inducer (IPTG) during 10 h long induction phase. The experimental data are pictured as points with standard error bars, the dashed lines show simulation data for initial values of the model function (i.e., initial point of fitting), the solid lines show the results of fitting and the dotted lines represent the final results optimised by MCMC method of the FME package (Section 2.6), which show the best agreement with the experimental data. The model with the best fit appears to be an enhanced Monod function where the maximum growth rate constant is substituted by the ramp function (defined in Figure 2) going from the maximum growth rate to the minimum growth rate reflecting the metabolic burden effect of the gradually-growing concentration of IPTG. The x-axes show the time of experiment in hours while the y-axes show CDW in g/L.

In another, more extensive case, two datasets for the same type of experiment with *E. coli* population degrading over time are compared for various initial concentrations of IPTG—$IPTG_0$ (Figure 10). The first dataset represents the individual effect of IPTG and the second one reflects the combined activity of IPTG and TCP. Remarkable is the fact that the exacerbation is more or less conserved for various values of $IPTG_0$ except for the case when $IPTG_0$ equals zero (displaying the toxic effect of TCP only). By acquiring a proportion of the median of the differences for various positive concentrations of $IPTG_0$ to the difference of the individual TCP effect (where $IPTG_0$ equals zero) we have obtained a value of 1.82. Due to the fact it is dimensionless, it is suitable for a wide variety of models regardless of the units used for determination of the bacterial population. Therefore, we incorporate this value as a potential exacerbation parameter ($ex_{on} = 1.82$, whereas $ex_{off} = 1$) in the next step of our workflow.

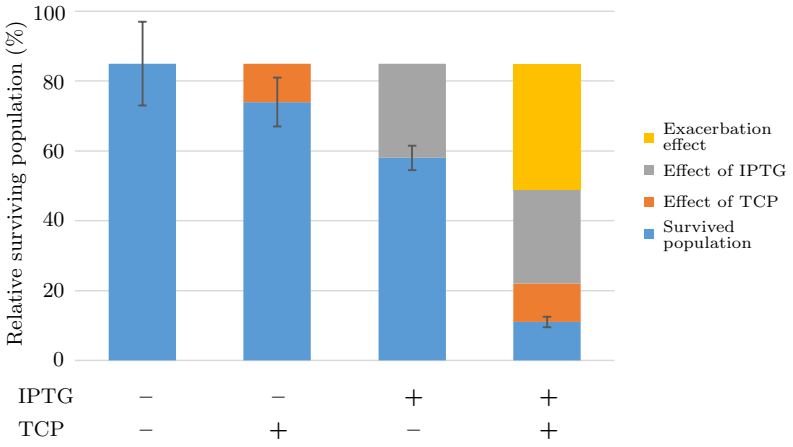


Figure 9. Evidence of exacerbation effect of IPTG on toxicity caused by TCP. Combined effect of metabolic burden caused by 0.2 mM IPTG and toxicity caused by TCP on *E. coli* BL21(DE3) cells carrying empty plasmids pCDF and pETDuet is displayed as the percentage of survived cells (blue bars) after induction in medium with or without 2 mM TCP. Pre-induced (IPTG +) or non-induced (IPTG −) cells were incubated in buffer with TCP (TCP +) or without (TCP −). The separate negative effects of TCP (orange), IPTG (gray), and the exacerbation of TCP toxicity in cells pre-induced with IPTG (yellow) are indicated. Error bars represent standard deviations calculated from at least five independent experiments. Note that experimental data come from [19].

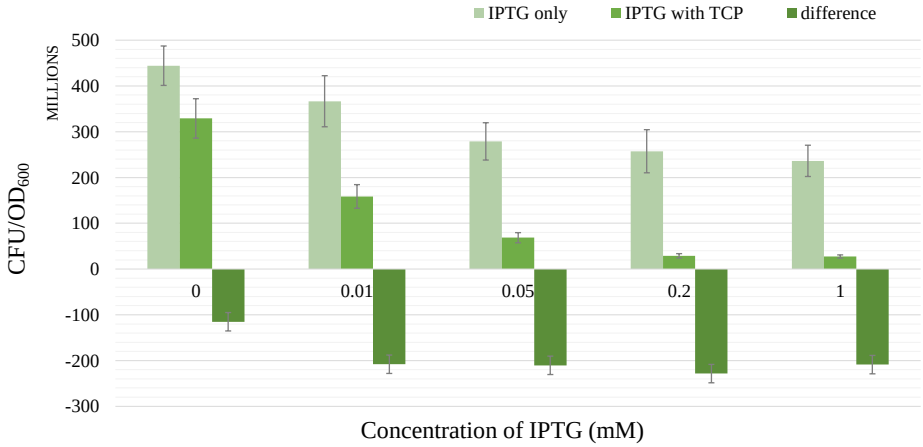


Figure 10. Exacerbation of TCP toxicity in *Escherichia coli* BL21(DE3) bearing synthetic TCP pathway by IPTG. The leftmost column in each dataset represents the population of cells pre-induced with various concentrations of IPTG. The middle column of each dataset shows the population of survived cells pre-induced with the same amount of IPTG after 5 h in the presence of 2 mM TCP. The rightmost column in each dataset shows the difference of the first and the second column (i.e., third = second − first). Note that the population in the first column of the first dataset is pre-induced with 0 mM IPTG and incubated in absence of TCP, which makes it a control group. The second column in the same dataset shows the sole effect of 2 mM TCP on the population. It is remarkable that the third columns in all other datasets seem to be in perfect match and approximately 1.82 times bigger than the same column in the first dataset. We explain this fact by the existence of the exacerbation effect. Different concentrations of IPTG were used for the induction of the TCP pathway expression from pCDF and pETDuet plasmids. Error bars represent standard deviations calculated from at least three independent experiments except for the rigthmost column in each dataset where error bars represent standard error of values in these columns.

3.5. Step 3: Model Dynamics Exploration

As the outcome of the previous part of the workflow, we obtained the extended ODE model (Figure 11). It makes an input into *Step 3* of the workflow (Section 2.5.2). For the purpose of the employed analysis techniques (parameter synthesis and robustness monitoring), we converted the model into two different formats compatible with the used software (the BIO (https://github.com/sybila/ode-generator/blob/master/README.md#model-syntax) format and the SBML (http://sbml.org/Documents/Specifications) format, respectively).

$$
\begin{aligned}
\frac{d[\text{Bact}]}{dt} &= [\text{Bact}] \cdot R([\text{IPTG}], t_1, t_2, \mu_{max}, \mu_{min}) \cdot \frac{[\text{GLY}]}{K + [\text{GLY}]} - [\text{Bact}] \cdot \gamma_{\text{Bact}} \\
&- [\text{Bact}] \cdot H([\text{IPTG}], t_1, ex_{\text{off}}, ex_{\text{on}}) \cdot \frac{k_{i,\text{TCP}} \cdot [\text{TCP}]^{n_{i,\text{TCP}}}}{K_{i,\text{TCP}} + [\text{TCP}]^{n_{i,\text{TCP}}}} \\
&- [\text{Bact}] \cdot k_{i,\text{ECH}} \cdot (1 - \exp(-\frac{[\text{ECH}]}{K_{i,\text{ECH}}})) - [\text{Bact}] \cdot k_{i,\text{CPD}} \cdot (1 - \exp(-\frac{[\text{CPD}]}{K_{i,\text{CPD}}})) \\
&- [\text{Bact}] \cdot k_{i,\text{GDL}} \cdot (1 - \exp(-\frac{[\text{GDL}]}{K_{i,\text{GDL}}}))
\end{aligned}
$$

$$
\frac{d[\text{TCP}]}{dt} = -\frac{V_{\max,\text{D}} \cdot [\text{IPTG}]}{K_{\text{M,D}} + [\text{IPTG}]} \cdot \frac{k_{1_R} \cdot [\text{TCP}]}{K_{\text{M},1} + [\text{TCP}]} - \frac{V_{\max,\text{D}} \cdot [\text{IPTG}]}{K_{\text{M,D}} + [\text{IPTG}]} \cdot \frac{k_{1_S} \cdot [\text{TCP}]}{K_{\text{M},1} + [\text{TCP}]}
$$

$$
\frac{d[(R)\text{-DCP}]}{dt} = \frac{V_{\max,\text{D}} \cdot [\text{IPTG}]}{K_{\text{M,D}} + [\text{IPTG}]} \cdot \frac{k_{1_R} \cdot [\text{TCP}]}{K_{\text{M},1} + [\text{TCP}]} - \frac{V_{\max,\text{H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \cdot \frac{k_{2_R} \cdot [(R)\text{-DCP}]}{K_{\text{M},2_R} + [(R)\text{-DCP}]}
$$

$$
\frac{d[(S)\text{-DCP}]}{dt} = \frac{V_{\max,\text{D}} \cdot [\text{IPTG}]}{K_{\text{M,D}} + [\text{IPTG}]} \cdot \frac{k_{1_S} \cdot [\text{TCP}]}{K_{\text{M},1} + [\text{TCP}]} - \frac{V_{\max,\text{H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \cdot \frac{k_{2_S} \cdot [(S)\text{-DCP}]}{K_{\text{M},2_S} + [(S)\text{-DCP}]}
$$

$$
\begin{aligned}
\frac{d[\text{ECH}]}{dt} &= \frac{V_{\max,\text{H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \cdot \frac{k_{2_R} \cdot [(R)\text{-DCP}]}{K_{\text{M},2_R} + [(R)\text{-DCP}]} - \frac{V_{\max,\text{E}} \cdot [\text{IPTG}]}{K_{\text{M,E}} + [\text{IPTG}]} \cdot \frac{k_3 \cdot [\text{ECH}]}{K_{\text{M},3} + [\text{ECH}]} \\
&+ \frac{V_{\max,\text{H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \cdot \frac{k_{2_S} \cdot [(S)\text{-DCP}]}{K_{\text{M},2_S} + [(S)\text{-DCP}]}
\end{aligned}
$$

$$
\frac{d[\text{CPD}]}{dt} = \frac{V_{\max,\text{E}} \cdot [\text{IPTG}]}{K_{\text{M,E}} + [\text{IPTG}]} \cdot \frac{k_3 \cdot [\text{ECH}]}{K_{\text{M},3} + [\text{ECH}]} - \frac{V_{\max,\text{H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \cdot \frac{k_4 \cdot [\text{CPD}]}{K_{\text{M},4} + [\text{CPD}]}
$$

$$
\frac{d[\text{GDL}]}{dt} = \frac{V_{\max,\text{H}} \cdot [\text{IPTG}]}{K_{\text{M,H}} + [\text{IPTG}]} \cdot \frac{k_4 \cdot [\text{CPD}]}{K_{\text{M},4} + [\text{CPD}]} - \frac{V_{\max,\text{E}} \cdot [\text{IPTG}]}{K_{\text{M,E}} + [\text{IPTG}]} \cdot \frac{k_5 \cdot [\text{GDL}]}{K_{\text{M},5} + [\text{GDL}]}
$$

$$
\begin{aligned}
\frac{d[\text{GLY}]}{dt} &= \frac{V_{\max,\text{E}} \cdot [\text{IPTG}]}{K_{\text{M,E}} + [\text{IPTG}]} \cdot \frac{k_5 \cdot [\text{GDL}]}{K_{\text{M},5} + [\text{GDL}]} \\
&- \frac{[\text{Bact}]}{Y} \cdot R([\text{IPTG}], t_1, t_2, \mu_{max}, \mu_{min}) \cdot \frac{[\text{GLY}]}{K + [\text{GLY}]}
\end{aligned}
$$

$$
\begin{gathered}
t_1 = 0.01,\ t_2 = 0.05,\ \mu_{min} = 0.0746,\ \mu_{max} = 0.2686,\ K = 2.7465 \cdot 10^{-5}, \\
\gamma_{\text{Bact}} = 0.0022,\ ex_{\text{off}} = 1,\ ex_{\text{on}} = 1.82,\ k_{i,\text{TCP}} = 15.131,\ n_{i,\text{TCP}} = 1.4256, \\
K_{i,\text{TCP}} = 523.9533,\ k_{i,\text{ECH}} = 0.2653,\ K_{i,\text{ECH}} = 0.3617,\ k_{i,\text{CPD}} = 0.05486, \\
K_{i,\text{CPD}} = 1.7178 \cdot 10^{-6},\ k_{i,\text{GDL}} = 0.4866,\ K_{i,\text{GDL}} = 81.2447,\ Y = 0.0838,\ k_{1_R} = 2088, \\
k_{1_S} = 1692,\ k_{2_R} = 6516,\ k_{2_S} = 288,\ k_3 = 51732,\ k_4 = 8568,\ k_5 = 14256,\ K_{\text{M},1} = 1.79, \\
K_{\text{M},2_R} = 2.49,\ K_{\text{M},2_S} = 3.33,\ K_{\text{M},3} = 0.09,\ K_{\text{M},4} = 0.86,\ K_{\text{M},5} = 3.54, \\
V_{\max,\text{D}} = 0.0019,\ K_{\text{M,D}} = 0.01749,\ V_{\max,\text{H}} = 0.005391,\ K_{\text{M,H}} = 0.008255, \\
V_{\max,\text{E}} = 0.004998,\ K_{\text{M,E}} = 0.004855
\end{gathered}
$$

Figure 11. An ODE model of the extended TCP metabolic pathway. The model represents a chain reaction for biodegradation of TCP into GLY reflecting the *E. coli* population. Note that the rate constants of the original ODE model were rescaled from seconds (s^{-1}) into hours (h^{-1}) because the data used for fitting were sampled every hour. It concerns the original rate constants (k_{1_R}, k_{1_S}, k_{2_R}, k_{2_S}, k_3, k_4, k_5). Units: $k_*, V_*, \mu_*, \gamma_*(h^{-1})$; $t_*, K_*(\text{mM})$; $ex_*, Y, n_*(\text{unitless})$.

Next, we specify the model parameters that make the objectives of further analysis. While the previous parts of the workflow targeted the internal parameters adapting the model for the particular cell population, in this part, we target the model parameters that can be perturbed and fine-tuned. In particular, we consider the following set of model parameters for tuning:

1. **Concentration of IPTG in mM**: IPTG is an obvious candidate for tuning because many aspects of the model depend on it and it can be controlled easily in the experimental environment (since its concentration is considered constant during the experiment, it can be referred by its initial concentration, denoted $\mathbf{IPTG}_0$).
2. **Size of bacterial population (Bact) in g/L**: The initial population size, denoted $\mathbf{Bact}_0$, makes the crucial input of the model and it affects the model output—the final population size (reached in a given time). In general, the initial population size can be controlled in experiments.
3. **Initial concentration of TCP ($\mathbf{TCP}_0$) in mM**: The key input of the model that must be set in order to make the modelled metabolic pathway work; it can be easily set to any arbitrary value during the experiments.
4. **Death rate of the population ($\gamma_{\mathbf{Bact}}$) in $\mathbf{h}^{-1}$**: The death rate is considered as a parameter because we are interested in the dynamics of the microbial culture and the effects affecting the growth.

Property 1. *The complete degradation of TCP as fast as possible with the least accumulated toxicity.*

In [37], we declared the desired property of the model dynamics verbally (stated as Property 1). The model used in that study did not concern the bacterial population. Therefore, the notion of toxicity was interpreted as an artificial accumulation of the inhibitory effect of the particular pathway's (intermediate) products. The inhibitory effect was experimentally measured and is traceable in [18,23].

Property 2. *The complete degradation of TCP as fast as possible with the most survived bacteria.*

In the extended model, we are able to investigate directly the effect of the particular pathway's products on the bacterial population. To that end, the desired property is lifted to the population level resulting in Property 2. Due to the very abstract character, this property serves as a theoretical concept and several adjustments have to be performed to use it with computational analysis methods.

First, the part "*the complete degradation of TCP*" is translated into "*the TCP concentration is close to zero or below a minimal threshold (e.g., 0.01 of mM)*". This is due to the possible errors of numerical solvers and the nature of the employed model approximation/abstraction algorithms.

Second, the terms such as "*as fast as possible*" and "*the most survived bacteria*" cannot be interpreted numerically, which is necessary for the computing. For that reason, we use various numerical thresholds to instantiate such abstract terms in the specified property.

As we addressed in Section 2.4.2, both considered approaches (parameter synthesis and robustness monitoring) employ a particular temporal logic for the specification of properties. We employ various properties compatible with both approaches and we utilise their specific advantages.

The results of parameter synthesis contain comprehensive information about parameter values for all possible initial settings of variables in considered ranges. In particular, not all *positive* results (parameter values satisfying the particular property) are automatically valid in all initial settings of model variables. The parameter synthesis method computes only positive results for which there exist some valid initial values of model variables. It is also worth noting that, due to the model overapproximation performed inside the parameter synthesis procedure, the complement to the set of positive results does not necessarily imply a valid negative result [37]. In other words, the so-called false-positives might exist.

Property 3. *The population will not drop below 0.08 g/L until TCP will degrade fully (drop below 0.01 mM).*

Addressing Property 2, we designed a suitable reformulation (Property 3), which is compatible with the parameter synthesis procedure. The results of parameter synthesis for this property are shown in Figure 12. The displayed parameters have been synthesised in given ranges. Every blue region depicts a set of values satisfying the stated property in at least one initial value of model variables. In this particular case, the blue regions make joint projections across all dimensions (i.e., parameters and variables) of the modelled system. Consequently, the property is confirmed to be satisfied by every combination of parameters (respectively, initial conditions) in the considered ranges.

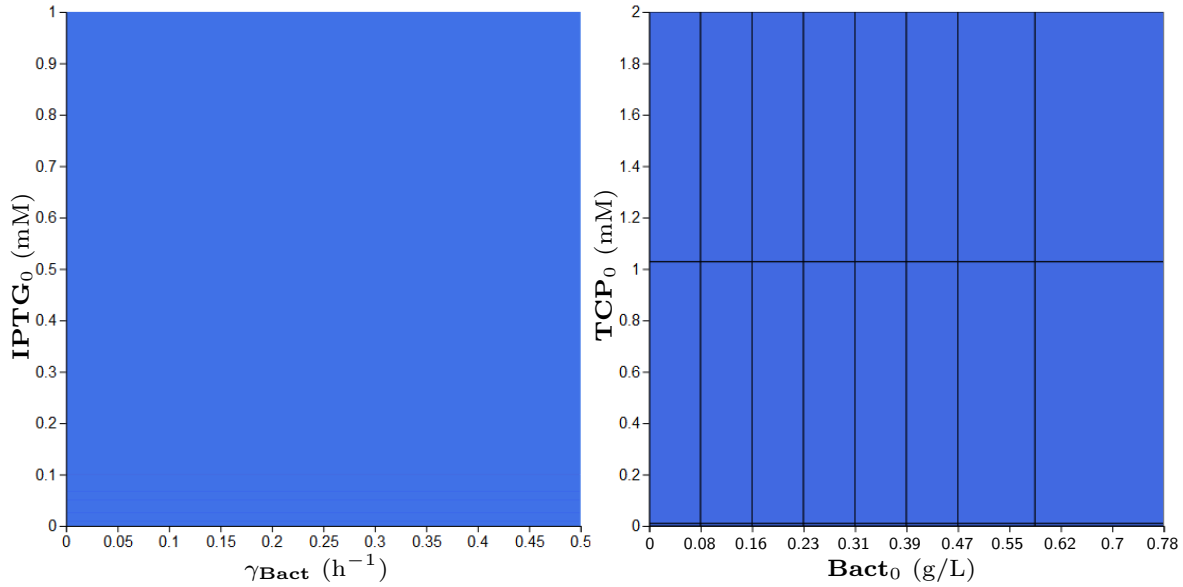


Figure 12. Results of parameter synthesis process for Property 3. Inside the figure, one can see two plots with blue regions. Both plots show a combination of parameters and (or) variables of the model where each point represents the particular evaluation of considered parameters (or variables). Every blue region represents a set of evaluations satisfying the stated property in at least one initial condition of the model. Here, the blue regions make joint projections across all non-displayed dimensions (i.e., parameters and variables). Consequently, the property holds in every combination of initial conditions (respectively, parameters) in the particular ranges.

Property 4. *Eventually, there will happen a situation where the population never exceeds a low value (stays below 0.08 g/L forever) and TCP concentration stays above 0.01 mM (never fully degrades).*

To support the above results, we decided to consider a property with the opposite meaning (Property 4). In particular, the analysis resulted with no positive results which shows us a certain agreement with the previous analysis. However, as we have already addressed, we cannot directly interpret an "empty result" due to potential existence of false-positives. Therefore, we decided to investigate Property 4 more deeply with the help of the robustness monitoring analysis (see the analysis of Property 7).

Property 5. *Eventually, there will happen a situation in which TCP concentration is currently above 0.01 mM and the population never exceeds a low value (stays below 0.08 g/L forever).*

An interesting fact appears in the analysis based on comparing Property 4 with its weaker form represented by Property 5. A difference between the two properties is only in the predicate about the TCP concentration. However, for the parameter synthesis method, the difference is significant; as in Property 5, we do not require the concentration of TCP to be above 0.01 mM forever. In Figure 13, it is shown that for the weaker property there exist positive results. However, they appear only for the population death rate (parameter γ_{Bact}) higher than 0.07 s^{-1}, which is an overrated value.

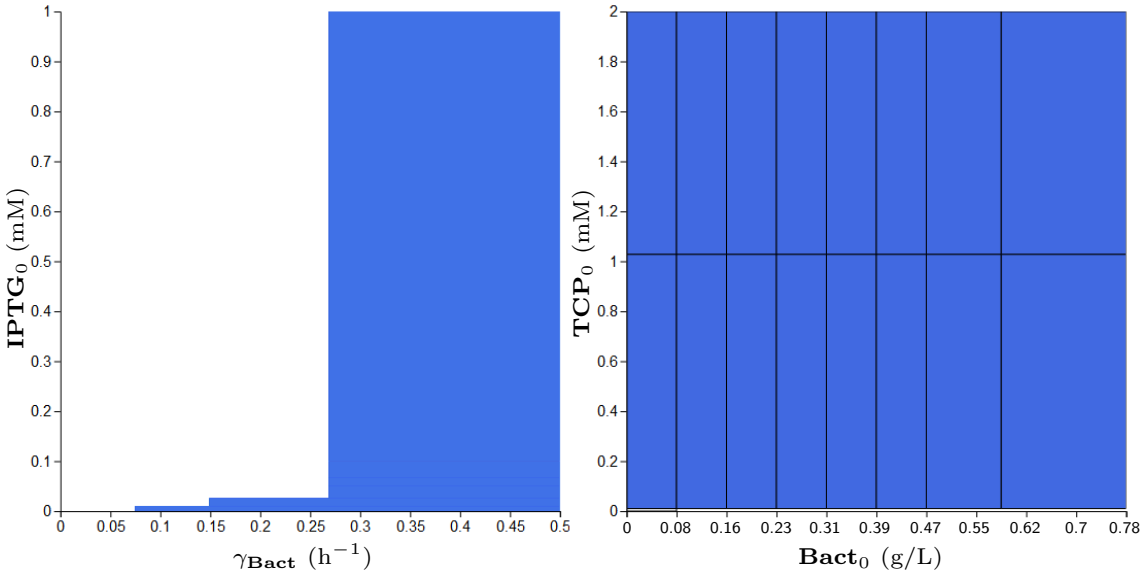


Figure 13. Results of parameter synthesis process for Property 5. Inside the figure, one can see two plots with blue regions. Both plots show a combination of parameters and (or) variables of the model where each point represents the particular evaluation of considered parameters (or variables). Every blue region represents a set of evaluations satisfying the stated property in at least one initial condition of the model. Here, the blue regions make joint projections across all non-displayed dimensions (i.e., parameters and variables). Consequently, the property holds in every combination of initial conditions (respectively, parameters) in the particular ranges.

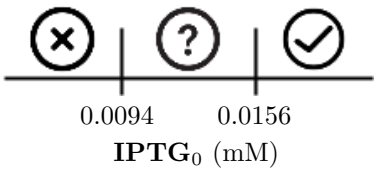
The parameter synthesis method works with abstractions of systems dynamics that do not consider time explicitly. However, it allows exploring patterns of model dynamics globally (regardless of the concrete settings of initial conditions). On the contrary, the robustness monitoring method considers time but works locally (i.e., dynamics is simulated in time for a given set of initial conditions). Therefore, there is a chance of missing an interesting behaviour occurring for an initial condition which is not included in the considered set.

Property 6. *The population will never drop below half of its initial value in the 5 h scope and TCP will degrade (drop below 0.01 mM) in the 2.5 h scope at the same time.*

Property 6 gives another reformulation of the desired property that now includes timing aspects. It requires the entire TCP degradation to be realised in a certain time limit. The particular results obtained with robustness monitoring are shown in Figure S10. For brevity, we present a simple diagram (Figure 14) showing that the change of the population death rate (γ_{Bact}) as well as the initial size of the population (Bact_0) have practically no influence on this property. In Figure S11, the range of the death rate coefficient is increased to the number of 0.1 (h^{-1}), which is a considerably overrated value. However, the influence is still considered negligible. More interesting appears to be the effect of

the TCP initial concentration (TCP_0) on the satisfiability of Property 6. The dependency between TCP_0 and $IPTG_0$ is non-linear (Figure 15).

(a)



(b)

$\mathbf{TCP}_0(\mathrm{mM})$	$\mathbf{Bact}_0(\mathrm{g/L})$	$\gamma_{\mathbf{Bact}}(\mathrm{h}^{-1})$
2.0	(0.024, 0.78)	(0.0, 0.01)

Figure 14. Diagram for satisfiability of Property 6 reflecting $IPTG_0$. (**a**) A simple diagram presenting the qualitative results of robustness monitoring for Property 6. It shows the influence of $IPTG_0$ on the satisfiability of the particular property. The two thresholds divide the area of influence into three sections. The left one which robustly violates the property, the right one—satisfying the property (robustly)—and the middle one where the result is not robust. This result holds for all combinations of the parameters (and variables) in the table (**b**).

Property 7. *The population dies eventually (drops below 0.01 g/L) while TCP does not degrade entirely (does not drop below 0.1 mM) in the 5 h horizon.*

To justify the results of parameter synthesis obtained for Property 4, we formulate a similar property that is compatible with the robustness monitoring method (Property 7). The results are shown in Figure S12—they contain only negative values (the property is definitely not satisfied in any sampled point). The summary of these results is available in a compact table (Table 1). For the reasonable range of the death rate coefficient, the results do not change significantly. However, there is some noticeable influence for the increased range of the death rate (up to 0.1 h^{-1}) shown in Figure S13. Nevertheless, the obtained negative results support the previous outcome of parameter synthesis and even improve it by adding the quantitative information.

Table 1. Results of robustness monitoring for Property 7. A simple table presents the relevant ranges of initial conditions (i.e., the concentration of variables and setting of parameters) which robustly violate Property 7.

$IPTG_0$ (mM)	TCP_0 (mM)	$Bact_0$ (g/L)	γ_{Bact} (h^{-1})
(0.0, 1.0)	(0.0, 4.0)	(0.024, 0.78)	(0.0, 0.01)

Robustness monitoring proves to be useful for this type of models. On the other hand, parameter synthesis approach is designed for models with a more complex relationship between variables where some interesting bifurcations can take place.

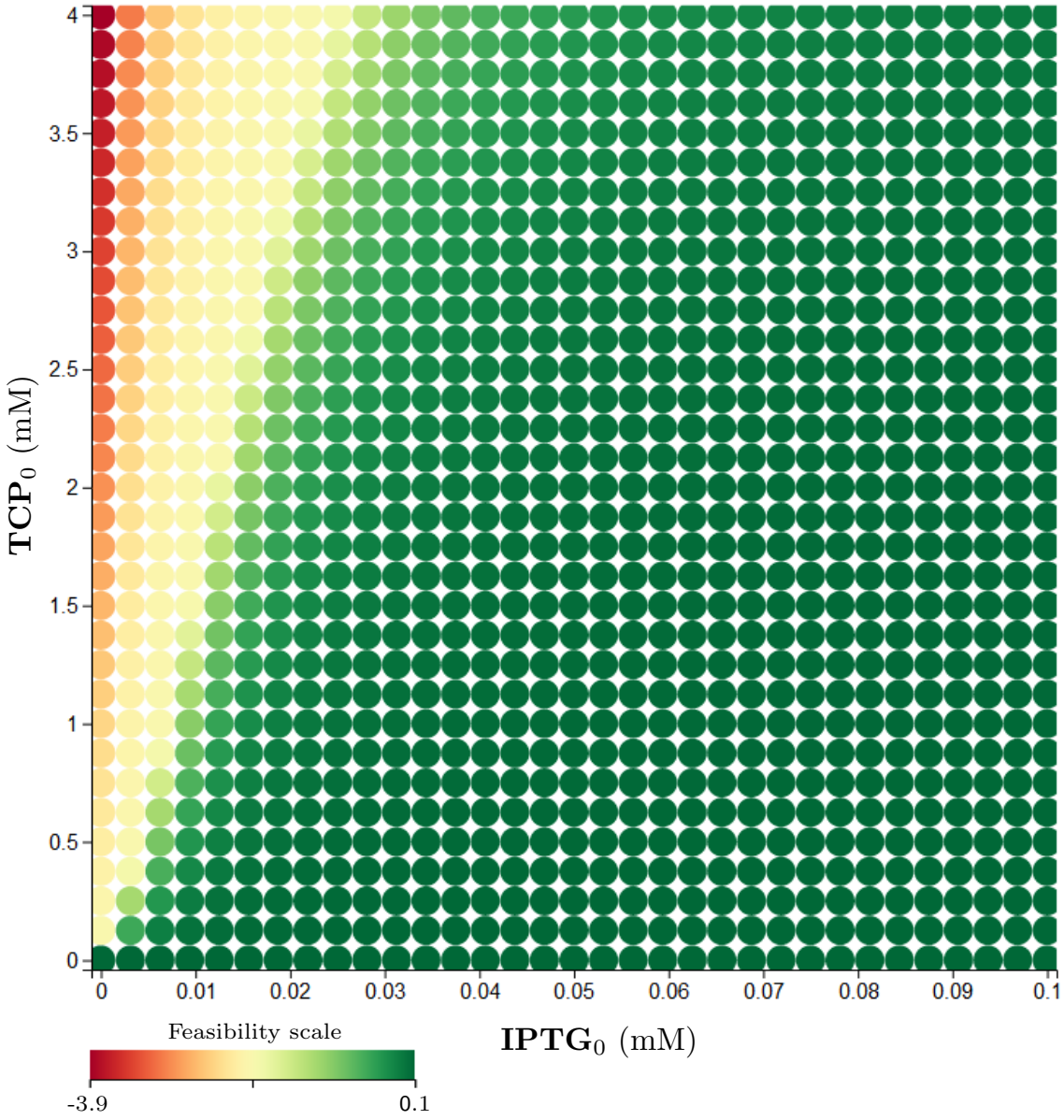


Figure 15. Results of robustness monitoring for Property 6 concerning TCP_0. The figure shows a two-dimensional plot with various coloured circles pointing by their centre to the particular setting of the plotted parameters (or variables). Initial values of variables and considered parameters (if not displayed in any axis) are: $Bact_0 = 0.487$ (g/L); GLY_0, (*R*)-DCP_0, (*S*)-DCP_0, ECH_0, CPD_0, GDL_0, $TCP_0 = 0$ (mM); $\gamma_{Bact} = 0.0022$ (h^{-1}). All the constants can be found in Figure 11. The shades of green colour imply a feasibility of the particular property in the particular initial setting while the shades of red imply a violation of the property—the darker the tone, the stronger the feasibility/violation. At the bottom of the plot, there is the feasibility scale mapped to real values. The plot represents a single layer of the entire parameter space.

4. Conclusions

The effects of metabolites on the fitness of bacterial strains carrying artificial pathways are of great interest for the community of metabolic engineers. The adverse effects of toxic metabolites and metabolic burden on the host cells need to be considered and ideally even quantitatively characterised by the computational modelling. Here, we present development of such a model for the *E. coli*

BL21(DE3) strain overexpressing the artificial pathway for degradation of highly toxic environmental pollutant 1,2,3-trichloropropane (TCP).

To assess the complex behaviour of the system, we created a unique mathematical model which combines population modeling with prediction of effects of metabolite toxicity and burden caused by application of the standard synthetic inducer IPTG triggering expression of heterologous biodegradation pathway. We were interested in conditions of the system in which the biodegradation is efficient while population survives and we investigated the system under these conditions using state-of-the-art computational methods.

First, we extended the original model of the metabolic pathway with a new model variable describing the bacterial population growth over time. The new model reflects the initial concentration of the inducer ($IPTG_0$). Second, we formalised the following features of the investigated system: (1) the metabolic burden effect caused by increased concentration of the inducer; (2) a highly toxic effect of TCP and other metabolites of the pathway; and (3) an exciting phenomenon of toxicity exacerbation occurred by the joint effect of the inducer (IPTG) and the pathway's substrate (TCP). Finally, using computational biology methods, we investigated the continuous parameter space of initial conditions and uncertain coefficients targeting the interesting properties of the model. We believe that the obtained results give a solid basis for further optimisation of the synthetic pathway in the considered strain.

Further refinement of the model is planned for future work. In particular, we aim at producing the data describing increasing enzyme concentrations in time. That will allow us to generalise the model by representing enzyme concentrations as model variables. Although the model itself is fine-tuned with respect to the considered *E. coli* BL21(DE3) strain, it can provide a modelling basis in different scenarios where a non-metabolisable inducer is used to control a closed cell population environment in stationary phase. On the computational side, the employed computational framework combining the carefully selected set of formal methods and simulation-based tools can be reused in any modelling case fitting the class of non-linear ODEs including the enzyme kinetics.

Supplementary Materials: The following are available online at http://www.mdpi.com/2076-2607/7/11/553/s1, Appendix A: The experimental data used for the fitting of functions explaining the enzymes concentration determined at respective IPTG concentrations; Appendix B: The origin of the death rate coefficient; Appendix C: The list of considered inhibition models; Data S1: IPTG growth curves; Data S2: Growth on 10 mM glycerol; Data S3: Toxicity on 10 mM glycerol; Figure S1: The comparative test of two mathematical models for biodegradation of TCP; Figure S2: The model of the metabolic pathway for biodegradation of TCP without reverse reactions; Figure S3: The comparative simulation of two bacterial growth functions; Figure S4: Evidence of need for proper function describing population growth reflecting metabolic burden caused by IPTG; Figure S5: The results of fitting to population growth data reflecting toxicity caused by TCP; Figure S6: The results of fitting to population growth data reflecting toxicity caused by ECH; Figure S7: The results of fitting to population growth data reflecting toxicity caused by CPD; Figure S8: The results of fitting to population growth data reflecting toxicity caused by GDL; Figure S9: The results of fitting to population growth data reflecting toxicity caused by TCP—prolonged simulation; Figure S10: The results of robustness monitoring for Property 6; Figure S11: The results of robustness monitoring for Property 6 with extended range of the death rate coefficient; Figure S12: The results of robustness monitoring for Property 7; Figure S13: The results of robustness monitoring for Property 7 with extended range of the death rate coefficient; Table S1: The quantitative results for fitting of the functions explaining the enzymes concentration reflecting IPTG; and Table S2: The results of fitting several growth functions to the same set of experimental data.

Author Contributions: Conceptualisation, J.D. and D.Š.; methodology, M.D. and D.Š.; software, M.D.; validation, L.C.; formal analysis, M.D.; investigation, M.D. and P.D.; resources, J.D. and D.Š.; data curation, L.C.; writing—original draft preparation, M.D. and L.C.; writing—review and editing, P.D., J.D. and D.Š.; visualisation, M.D.; supervision, J.D. and D.Š; project administration, J.D.; and funding acquisition, J.D. and D.Š.

Funding: This research was funded by the Czech Ministry of Education grants number CZ.02.1.01/0.0/0.0/1_026/0008451, CZ.02.1.01/0.0/0.0/16_019/0000868, LM2015047, and LM2015055, by Grant Agency of Czech Republic grant number GA18-00178S, and by EU grants Rafts4Biotech (720776) and Sinfonia (722610). Computational resources were provided by CERIT-SC (LM2015042) and Meta-Centrum (LM2015085).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

IPTG	isopropyl-beta-D-thiogalactopyranoside
TCP	1,2,3-trichloropropane
DCP	2,3-dichloropropane-1-ol
ECH	epichlorohydrin
CPD	3-chloropropane-1,2-diol
GDL	glycidol
GLY	glycerol
DhaA	haloalkane dehalogenase
EchA	epoxide hydrolase
HheC	haloalcohol dehalogenase
MM	Michaelis-Menten
CDW	cell dry weight
ODE	ordinary differential equations
MCMC	Markov chain Monte Carlo
GUI	graphical user interface
CLI	command-line interface

Appendix A. The Experimental Data Used for the Fitting of Functions Explaining the Enzymes Concentration Determined at Respective IPTG Concentrations

In Tables A1 and A2, the results of the experiments published in [19] are presented. These results contain data from two experiments, one of which is shown in Figure A1. The tables show relative portions of enzymes in total soluble protein content of cell-free extract for different starting concentrations of IPTG—the inducer used for induction of host cells (*E. coli* deg31) under conditions described in Laboratory Methods.

The total masses of enzymes $\left(\frac{\text{mg}}{10\ \text{mL}}\right)$ were calculated from value 4.02 (measured for 0.2 mM IPTG and coloured in blue) appropriately for the rest of IPTG concentrations reflecting different portions of total enzymes (Column 2 in Tables A1 and A2). The total mass values were then used for calculation of the enzymes mass in cell-free extract (Tables A3 and A4) with respect to the particular enzyme relative portion (Tables A1 and A2, respectively). Then, using atomic mass values in Table A5 and enzyme masses in cell-free extract from two different experiments, we calculated molar concentrations (mM) for all enzymes. Finally, the median and standard deviation were calculated from these values (Table A6) and used for fitting, as described in Section 3.3.

Table A1. Relative fractions of enzymes from Experiment 1.

IPTG (mM)	Content in Cell-Free Extract (%) *	DhaA31 (Relative Portion)	HheC (Relative Portion)	EchA (Relative Portion)	Total Mass $\left(\frac{\text{mg}}{10\ \text{mL}}\right)$
1.0	50	0.19	0.39	0.42	4.19
0.2	48	0.17	0.40	0.43	4.02
0.05	40	0.13	0.38	0.49	3.35
0.025	40	0.12	0.41	0.47	3.35
0.01	22	0.16	0.36	0.48	1.84
0.0	15	0.20	0.33	0.47	1.26

* Calculated as a portion of total soluble protein content.

Table A2. Relative fractions of enzymes from Experiment 2.

IPTG (mM)	Content in Cell-Free Extract (%) *	DhaA31 (Relative Portion)	HheC (Relative Portion)	EchA (Relative Portion)	Total Mass ($\frac{mg}{10\ mL}$)
1.0	60	0.15	0.39	0.46	3.890
0.2	62	0.15	0.39	0.46	4.020
0.05	50	0.13	0.38	0.49	3.242
0.025	50	0.11	0.37	0.52	3.242
0.01	42	0.11	0.37	0.52	2.723
0.0	15	0.17	0.40	0.44	0.973

* Calculated as a portion of total soluble protein content.

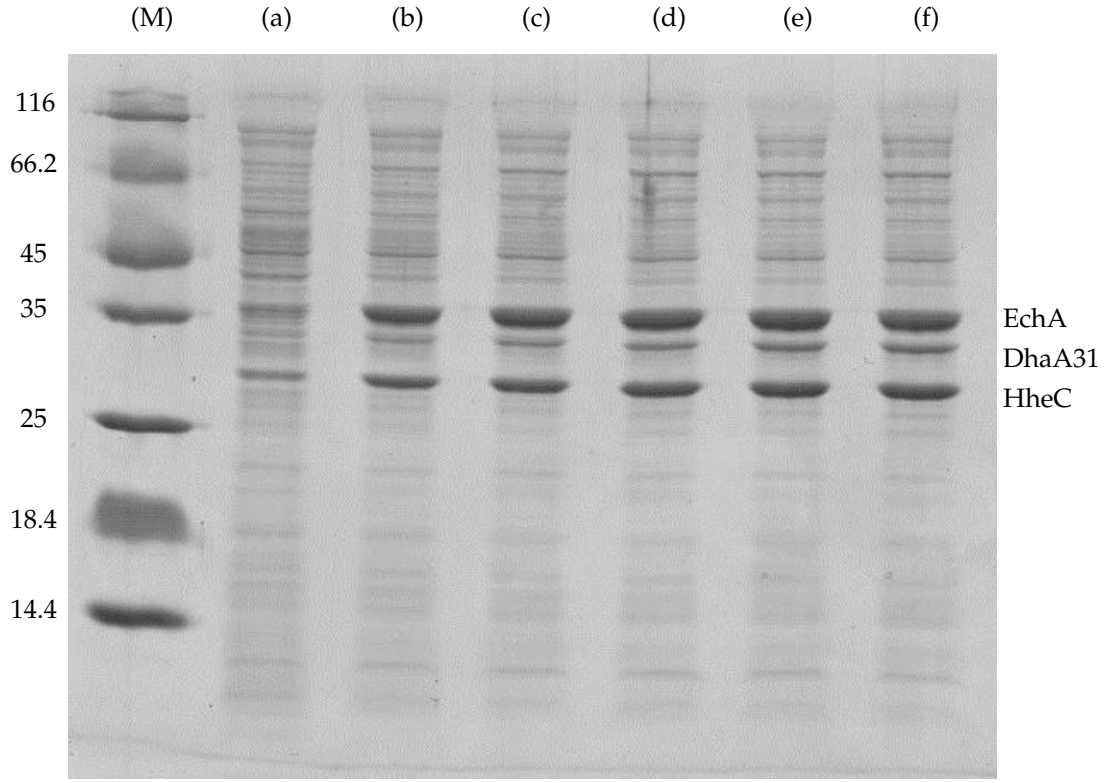


Figure A1. Sodium dodecyl sulfate polyacrylamide gel electrophoresis of cell-free extracts obtained from *E. coli* deg31 cells induced with various concentrations of IPTG ((**M**) protein marker (116, 66.2, 45, 35, 25, 18.4, and 14.4 kDa)): (**a**) 0.0 mM IPTG; (**b**) 0.01 mM IPTG; (**c**) 0.025 mM IPTG; (**d**) 0.05 mM IPTG; (**e**) 0.2 mM IPTG; and (**f**) 1.0 mM IPTG. Bands of DhaA31 (34 kDa), HheC (29 kDa) and EchA (35 kDa) are marked.

Table A3. Mass and molar concentration of enzymes in cell-free extract from Experiment 1.

IPTG (mM)	Total Mass ($\frac{mg}{10\ mL}$)	DhaA31 ($\frac{mg}{L}$) *	HheC ($\frac{mg}{L}$) *	EchA ($\frac{mg}{L}$) *	DhaA31 (mM) #	HheC (mM) #	EchA (mM) #
1.0	4.19	79.6	163.3	175.9	2.24×10^{-3}	5.57×10^{-3}	4.82×10^{-3}
0.2	4.02	68.3	160.8	172.9	1.92×10^{-3}	5.48×10^{-3}	4.74×10^{-3}
0.05	3.35	43.6	127.3	164.2	1.22×10^{-3}	4.34×10^{-3}	4.50×10^{-3}
0.025	3.35	40.2	137.4	157.5	1.13×10^{-3}	4.68×10^{-3}	4.32×10^{-3}
0.01	1.84	29.5	66.3	88.4	8.29×10^{-4}	2.26×10^{-3}	2.43×10^{-3}
0.0	1.26	25.1	41.5	59.0	7.06×10^{-4}	1.41×10^{-3}	1.62×10^{-3}

* Calculated as a relative portion of the total mass $\left(\frac{mg}{10\ mL}\right) \times 100$. # Calculated as $\left(\frac{mg}{Da \cdot L}\right)$.

Table A4. Mass and molar concentration of enzymes in cell-free extract from Experiment 2.

IPTG (mM)	Total Mass ($\frac{mg}{10\ mL}$)	DhaA31 ($\frac{mg}{L}$) *	HheC ($\frac{mg}{L}$) *	EchA ($\frac{mg}{L}$) *	DhaA31 (mM) #	HheC (mM) #	EchA (mM) #
1.0	3.89	58.4	151.7	179.0	1.64×10^{-3}	5.17×10^{-3}	4.91×10^{-3}
0.2	4.02	60.3	156.8	184.9	1.69×10^{-3}	5.34×10^{-3}	5.07×10^{-3}
0.05	3.24	42.1	123.2	158.9	1.18×10^{-3}	4.20×10^{-3}	4.36×10^{-3}
0.025	3.24	35.7	120.0	168.6	1.00×10^{-3}	4.09×10^{-3}	4.62×10^{-3}
0.01	2.72	30.0	100.8	141.6	8.42×10^{-4}	3.44×10^{-3}	3.88×10^{-3}
0.0	9.73×10^{-1}	16.5	38.9	42.8	4.65×10^{-4}	1.33×10^{-3}	1.17×10^{-3}

* Calculated as a relative portion of the total mass $\left(\frac{mg}{10\ mL}\right) \times 100$. # Calculated as $\left(\frac{mg}{Da \cdot L}\right)$.

Table A5. Mass values of enzymes.

DhaA31 (Da)	HheC (Da)	EchA (Da)
35576.37	29333.07	36465.11

Table A6. Molar concentration of enzymes in cell-free extract calculated as a median of the values in Tables A3 and A4 with standard deviation values (i.e., stdev columns).

IPTG	DhaA31 (mM)		HheC (mM)		EchA (mM)	
(mM)	Median	stdev	Median	stdev	Median	stdev
1.0	1.94×10^{-3}	4.22×10^{-4}	5.37×10^{-3}	2.79×10^{-4}	4.87×10^{-3}	5.97×10^{-5}
0.2	1.81×10^{-3}	1.60×10^{-4}	5.41×10^{-3}	9.69×10^{-5}	4.91×10^{-3}	2.34×10^{-4}
0.05	1.20×10^{-3}	2.79×10^{-5}	4.27×10^{-3}	9.90×10^{-5}	4.43×10^{-3}	1.03×10^{-4}
0.025	1.07×10^{-3}	9.02×10^{-5}	4.39×10^{-3}	4.19×10^{-4}	4.47×10^{-3}	2.16×10^{-4}
0.01	8.35×10^{-4}	9.45×10^{-6}	2.85×10^{-3}	8.30×10^{-4}	3.15×10^{-3}	1.03×10^{-3}
0.0	5.85×10^{-4}	1.71×10^{-4}	1.37×10^{-3}	6.15×10^{-5}	1.40×10^{-3}	3.15×10^{-4}

Appendix B. The Origin of the Death Rate Coefficient

We used the mean value of death rate in percentage per generation (% p.g.) from article *Robust growth of Escherichia coli* [68] for *E. coli* strain B/r, which is close to the BL21 we used in experiments. The particular value ranges from 0.5% p.g. to 1% p.g.. However, we need a death rate per hour (h^{-1}) in our model. To achieve that, we divided p.g. values with the generation (doubling) time of the population in our model. The growth rate (g) in our model is in the range $[0.07, 0.26]$ (h^{-1}) according to fitting of the model to the experimental data. Thus, the particular values of generation time (t_g) are between 2.58 h and 9.29 h for the doubling of the population. Both growth rate and doubling time were obtained as a median from at least three experiments. The final death-rate-per-hour values—limiting the range—are the maximum and minimum (emphasised numbers) from Table A7.

Table A7. The death rate (γ) coefficients evaluated from different growth rates (g) using the generation time (t_g) and percentage death rate per generation (% p.g.).

g (h^{-1})	t_g (h)	γ (h^{-1}) as 0.5% p.g.	γ (h^{-1}) as 1% p.g.
0.26	2.58	$1.938 \cdot 10^{-3}$	$\mathbf{3.876 \cdot 10^{-3}}$
0.07	9.29	$\mathbf{5.382 \cdot 10^{-4}}$	$1.076 \cdot 10^{-3}$

Appendix C. The List of Considered Inhibition Models

In this step of modelling, we found a good model describing the bacterial population growth on a substrate (e.g., glycerol). In this case study, the model involves traditional Monod equation [28] as the growth rate function:

$$\mu = \frac{\mu_{\max}[S]}{K_S + [S]} \tag{A1}$$

where μ is the specific growth rate, $\mu_{\max}$ is the maximum specific growth rate for the culture, $[S]$ is the substrate concentration, and K_S is the *half-saturation constant* (or the *affinity constant*). Both $\mu_{\max}$ and K_S reflect intrinsic physiological properties of a particular type of microorganism, the substrate utilisation and the temperature of growth ([67], Chapter 3).

The Monod equation can express the rate of change in the number of cells as well as the fluxion of the cell mass (both can be defined by $[B]$). Thus, the whole model of bacterial growth is:

$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} \tag{A2}$$

In general, we could use a different model as the growth rate function [52]. We only need to have some proper model and think of it as an abstract module describing population growth. In this way, we can use a different module to explain the inhibition of the population growth by a poisonous substance (e.g., a toxic water pollutant such as 1,2,3-trichloropropane, TCP). However, there is no universal inhibition model. Therefore, we conducted several simulations with various combinations of models to fit the experimental data properly. In the following list, $[B]$, $[S]$ and $[X]$ stand for the bacterial population, the substrate concentration and the toxic substance concentration, respectively. The majority of the following equations consist of a growth module (e.g., Monod) and some inhibition module (separated by minus character). Equations (11)–(16) adopt a different form (in general called *competitive inhibition* [25]) using one module (e.g., Monod) extended by different types of inhibition:

Monod + Hill [27]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot \frac{k[X]^n}{{K_i}^n + [X]^n} \tag{A3}$$

Monod + Aiba-Edward [65]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot \frac{k[X]}{K_X + [X]} \cdot \exp\left(\frac{-[X]}{K_i}\right) \tag{A4}$$

Monod + Andrews [66]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot \frac{k}{(1 + \frac{K_X}{[X]})(1 + \frac{[X]}{K_i})} \tag{A5}$$

Monod + Haldane-Andrews [66]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot \frac{k[X]}{K_X + [X] + \frac{[X]^2}{K_i}} \tag{A6}$$

Monod + Monod:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot \frac{k[X]}{K_X + [X]} \tag{A7}$$

Monod + Moser [64]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot \frac{k[X]^n}{K_X + [X]^n} \tag{A8}$$

Monod + Tessier [63]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot k\left(1 - \exp\left(\frac{-[X]}{K_i}\right)\right) \tag{A9}$$

Monod + Tessier-type [63]:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} - [B] \cdot k\left(\exp\left(\frac{-[X]}{K_i}\right) - \exp\left(\frac{-[X]}{K_X}\right)\right) \tag{A10}$$

competitive inhibition:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S(1 + \frac{[X]}{K_i}) + [S]} \tag{A11}$$

non-competitive inhibition:
$$\frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}}{1 + \frac{[X]}{K_i}} \cdot \frac{[S]}{K_S + [S]} \tag{A12}$$

$$\text{uncompetitive inhibition:} \quad \frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S](1 + \frac{[X]}{K_i})} \tag{A13}$$

$$\text{non-competitive inhibition using negative Hill:} \quad \frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} \cdot \frac{k \cdot {K_i}^n}{{K_i}^n + [X]^n} \tag{A14}$$

$$\text{non-competitive inhibition using negative Moser:} \quad \frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} \cdot \frac{k \cdot K_i}{K_i + [X]^n} \tag{A15}$$

$$\text{non-competitive exponential inhibition:} \quad \frac{d[B]}{dt} = [B] \cdot \frac{\mu_{\max}[S]}{K_S + [S]} \cdot \frac{k}{{K_i}^{[X]}} \tag{A16}$$

where K_i and K_X are inhibition constants explaining inhibitory effects of poisonous substance at higher concentrations; k is an inhibitory rate; and n is a coefficient explaining cooperativity (as defined in Hill equation) or another biochemical property, depending on the context.

References

1. Marisch, K.; Bayer, K.; Cserjan-Puschmann, M.; Luchner, M.; Striedner, G. Evaluation of three industrial Escherichia coli strains in fed-batch cultivations during high-level SOD protein production. *Microb. Cell Factories* **2013**, *12*, 58.
2. Zhang, Z.; Kuipers, G.; Niemiec, Ł.; Baumgarten, T.; Slotboom, D.J.; de Gier, J.W.; Hjelm, A. High-level production of membrane proteins in *E. coli* BL21 (DE3) by omitting the inducer IPTG. *Microb. Cell Factories* **2015**, *14*, 142.
3. Choi, J.H.; Keum, K.C.; Lee, S.Y. Production of recombinant proteins by high cell density culture of Escherichia coli. *Chem. Eng. Sci.* **2006**, *61*, 876–885.
4. Balzer, S.; Kucharova, V.; Megerle, J.; Lale, R.; Brautaset, T.; Valla, S. A comparative analysis of the properties of regulated promoter systems commonly used for recombinant gene expression in Escherichia coli. *Microb. Cell Factories* **2013**, *12*, 26.
5. Tolia, N.H.; Joshua-Tor, L. Strategies for protein coexpression in Escherichia coli. *Nat. Methods* **2006**, *3*, 55.
6. Xu, P.; Vansiri, A.; Bhan, N.; Koffas, M.A. ePathBrick: A synthetic biology platform for engineering metabolic pathways in E. coli. *ACS Synth. Biol.* **2012**, *1*, 256–266.
7. Wu, J.; Du, G.; Zhou, J.; Chen, J. Metabolic engineering of Escherichia coli for (2S)-pinocembrin production from glucose by a modular metabolic strategy. *Metab. Eng.* **2013**, *16*, 48–55.
8. Xu, P.; Gu, Q.; Wang, W.; Wong, L.; Bower, A.G.; Collins, C.H.; Koffas, M.A. Modular optimization of multi-gene pathways for fatty acids production in E. coli. *Nat. Commun.* **2013**, *4*, 1409.
9. Fang, M.Y.; Zhang, C.; Yang, S.; Cui, J.Y.; Jiang, P.X.; Lou, K.; Wachi, M.; Xing, X.H. High crude violacein production from glucose by Escherichia coli engineered with interactive control of tryptophan pathway and violacein biosynthetic pathway. *Microb. Cell Factories* **2015**, *14*, 8.
10. Liu, C.; Men, X.; Chen, H.; Li, M.; Ding, Z.; Chen, G.; Wang, F.; Liu, H.; Wang, Q.; Zhu, Y.; et al. A systematic optimization of styrene biosynthesis in Escherichia coli BL21 (DE3). *Biotechnol. Biofuels* **2018**, *11*, 14.
11. Fordjour, E.; Adipah, F.K.; Zhou, S.; Du, G.; Zhou, J. Metabolic engineering of Escherichia coli BL21 (DE3) for de novo production of l-DOPA from d-glucose. *Microb. Cell Factories* **2019**, *18*, 74.
12. National Center for Biotechnology Information. PubChem Compound Database; CID=656894. Available online: https://pubchem.ncbi.nlm.nih.gov/compound/656894 (accessed on 18 December 2018).
13. Glick, B.R. Metabolic load and heterologous gene expression. *Biotechnol. Adv.* **1995**, *13*, 247–261.
14. Wu, G.; Yan, Q.; Jones, J.A.; Tang, Y.J.; Fong, S.S.; Koffas, M.A. Metabolic burden: Cornerstones in synthetic biology and metabolic engineering applications. *Trends Biotechnol.* **2016**, *34*, 652–664.
15. Jones, K.L.; Kim, S.W.; Keasling, J. Low-copy plasmids can perform as well as or better than high-copy plasmids for metabolic engineering of bacteria. *Metab. Eng.* **2000**, *2*, 328–338.
16. Silva, F.; Queiroz, J.A.; Domingues, F.C. Evaluating metabolic stress and plasmid stability in plasmid DNA production by Escherichia coli. *Biotechnol. Adv.* **2012**, *30*, 691–708.

17. Mairhofer, J.; Scharl, T.; Marisch, K.; Cserjan-Puschmann, M.; Striedner, G. Comparative transcription profiling and in-depth characterization of plasmid-based and plasmid-free Escherichia coli expression systems under production conditions. *Appl. Environ. Microbiol.* **2013**, *79*, 3802–3812.
18. Kurumbang, N.P.; Dvorak, P.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. *ACS Synth. Biol.* **2013**, *3*, 172–181.
19. Dvořák, P.; Chrást, L.; Nikel, P.I.; Fedr, R.; Souček, K.; Sedláčková, M.; Chaloupková, R.; Lorenzo de, V.; Prokop, Z.; Damborský, J. Exacerbation of substrate toxicity by IPTG in Escherichia coli BL21 (DE3) carrying a synthetic metabolic pathway. *Microb. Cell Factories* **2015**, *14*, 201.
20. de Jong, H.; Casagranda, S.; Giordano, N.; Cinquemani, E.; Ropers, D.; Geiselmann, J.; Gouzé, J.L. Mathematical modelling of microbes: Metabolism, gene expression and growth. *J. R. Soc. Interface* **2017**, *14*, doi:10.1098/rsif.2017.0502.
21. Weiße, A.Y.; Oyarzún, D.A.; Danos, V.; Swain, P.S. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E1038–E1047.
22. Kitano, H. Systems biology: Towards system-level understanding of biological systems. *Found. Syst. Biol.* **2001**, 1.
23. Dvořák, P. Engineering of the Synthetic Metabolic Pathway for Biodegradation of an Environmental Pollutant. Ph.D. Theses, Masaryk University, Faculty of Science, Brno, Czech Republic, 2014.
24. Glazyrina, J.; Materne, E.M.; Dreher, T.; Storm, D.; Junne, S.; Adams, T.; Greller, G.; Neubauer, P. High cell density cultivation and recombinant protein production with Escherichia coli in a rocking-motion-type bioreactor. *Microb. Cell Factories* **2010**, *9*, 42.
25. Keener, J.P.; Sneyd, J. *Mathematical Physiology*; Springer: New York, NY, USA, 1998; Volume 1.
26. Michaelis, L.; Menten, M.L. Die kinetik der invertinwirkung. *Biochem. Ztg.* **1913**, *49*, 333–369.
27. Hill, A.V. The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *J. Physiol.* **1910**, *40*, 4–7.
28. Monod, J. *Recherches sur la Croissance des Cultures Bactériennes*; Hermann & Cie: Paris, France, 1942.
29. Yordanov, B.; Belta, C. Parameter synthesis for piecewise affine systems from temporal logic specifications. In *Hybrid Systems: Computation and Control*; Springer: New York, NY, USA, 2008; pp. 542–555.
30. Donzé, A.; Clermont, G.; Langmead, C.J. Parameter synthesis in nonlinear dynamical systems: Application to systems biology. *J. Comput. Biol.* **2010**, *17*, 325–336.
31. Donzé, A. Breach, a toolbox for verification and parameter synthesis of hybrid systems. In Proceedings of the Computer Aided Verification, Edinburgh, UK, 15–19 July 2010; Springer: New York, NY, USA, 2010; pp. 167–170.
32. Barnat, J.; Brim, L.; Krejčí, A.; Streck, A.; Šafránek, D.; Vejnár, M.; Vejpustek, T. On parameter synthesis by parallel model checking. *IEEE/ACM Trans. Comput. Biol. Bioinform. (TCBB)* **2012**, *9*, 693–705.
33. Češka, M.; Dannenberg, F.; Kwiatkowska, M.; Paoletti, N. Precise parameter synthesis for stochastic biochemical systems. In *Computational Methods in Systems Biology*; Springer: Cham, Switzerland, 2014; pp. 86–98.
34. Brim, L.; Češka, M.; Demko, M.; Pastva, S.; Šafránek, D. Parameter Synthesis by Parallel Coloured CTL Model Checking. In *Computational Methods in Systems Biology*; Springer: Cham, Switzerland, 2015; Volume 9308, pp. 251–263.
35. Bortolussi, L.; Milios, D.; Sanguinetti, G. U-Check: Model Checking and Parameter Synthesis Under Uncertainty. In Proceedings of the QEST'15, Madrid, Spain, 1–3 September 2015; Springer: New York, NY, USA, 2015; LNCS, Volume 9259, pp. 89–104.
36. Bogomolov, S.; Schilling, C.; Bartocci, E.; Batt, G.; Kong, H.; Grosu, R. Abstraction-Based Parameter Synthesis for Multiaffine Systems. In *Hardware and Software: Verification and Testing*; Springer: New York, NY, USA, 2015; pp. 19–35.
37. Demko, M.; Beneš, N.; Brim, L.; Pastva, S.; Šafránek, D. High-performance symbolic parameter synthesis of biological models: A case study. In *Computational Methods in Systems Biology*; Springer: Cham, Switzerland, 2016; pp. 82–97.

38. Beneš, N.; Brim, L.; Demko, M.; Pastva, S.; Šafránek, D. Parallel SMT-based parameter synthesis with application to piecewise multi-affine systems. In Proceedings of the International Symposium on Automated Technology for Verification and Analysis, Chiba, Japan, 17–20 October 2016; Springer: New York, NY, USA, 2016, Volume 9936, pp. 192–208.
39. Pastva, S. Parallel Parameter Synthesis from Hybrid Logic HUCTL Formulas. Master's Thesis, Masaryk University, Faculty of Informatics, Brno, Czech Republic, 2017.
40. Beneš, N.; Brim, L.; Demko, M.; Pastva, S.; Šafránek, D. Pithya: A Parallel Tool for Parameter Synthesis of Piecewise Multi-affine Dynamical Systems. In Proceedings of the International Conference on Computer Aided Verification (CAV), Heidelberg, Germany, 24–28 July 2017; Springer: New York, NY, USA, 2017; pp. 591–598.
41. Aster, R.C.; Borchers, B.; Thurber, C.H. *Parameter Estimation and Inverse Problems*; Elsevier: Amsterdam, The Netherlands, 2018.
42. Hughes, G.E.; Cresswell, M.J. *A New Introduction to Modal Logic*; Psychology Press: London, UK, 1996.
43. Rizk, A.; Batt, G.; Fages, F.; Soliman, S. A general computational method for robustness analysis with applications to synthetic gene networks. *Bioinformatics* **2009**, *25*, i169–i178.
44. Rizk, A.; Batt, G.; Fages, F.; Soliman, S. Continuous valuations of temporal logic specifications with applications to parameter optimization and robustness measures. *Theor. Comput. Sci.* **2011**, *412*, 2827–2839.
45. Donzé, A.; Fanchon, E.; Gattepaille, L.M.; Maler, O.; Tracqui, P. Robustness analysis and behavior discrimination in enzymatic reaction networks. *PLoS ONE* **2011**, *6*, e24246.
46. Brim, L.; Dluhoš, P.; Šafránek, D.; Vejpustek, T. STL*: Extending signal temporal logic with signal-value freezing operator. *Inf. Comput.* **2014**, *236*, 52–67.
47. Vejpustek, T. Robustness Analysis of Extended Signal Temporal Logic STL*. Master's Thesis, Masaryk University, Faculty of Informatics, Brno, Czech Republic, 2013.
48. Brim, L.; Vejpustek, T.; Šafránek, D.; Fabriková, J. Robustness Analysis for Value-Freezing Signal Temporal Logic. In HSB'13, EPTCS, 2013; Volume 125, pp. 20–36.
49. Nelson, D.L.; Lehninger, A.L.; Cox, M.M. *Lehninger Principles of Biochemistry*; Macmillan: New York, NY, USA, 2008.
50. Studier, F.W.; Moffatt, B.A. Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* **1986**, *189*, 113–130.
51. Ferrer-Miralles, N.; Saccardo, P.; Corchero, J.L.; Xu, Z.; García-Fruitós, E. General introduction: Recombinant protein production and purification of insoluble proteins. In *Insoluble Proteins*; Springer: New York, NY, USA, 2015; pp. 1–24.
52. Sadhukhan, S.; Villa, R.; Sarkar, U. Microbial production of succinic acid using crude and purified glycerol from a Crotalaria juncea based biorefinery. *Biotechnol. Rep.* **2016**, *10*, 84–93.
53. Kim, B.H.; Gadd, G.M. *Bacterial Physiology and Metabolism*; Cambridge University Press: Cambridge, UK, 2008.
54. Soetaert, K.; Petzoldt, T. Inverse Modelling, Sensitivity and Monte Carlo Analysis in R Using Package FME. *J. Stat. Softw. Artic.* **2010**, *33*, 1–28.
55. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2015.
56. Beneš, N.; Brim, L.; Demko, M.; Pastva, S.; Šafránek, D. A model checking approach to discrete bifurcation analysis. International Symposium on Formal Methods. Springer: New York, NY, USA, 2016; pp. 85–101.
57. Brim, L.; Demko, M.; Pastva, S.; Šafránek, D. High-Performance Discrete Bifurcation Analysis for Piecewise-Affine Dynamical Systems. In *Hybrid Systems Biology*; Springer: New York, NY, USA, 2015; pp. 58–74.
58. Hajnal, M.; Šafránek, D.; Demko, M.; Pastva, S.; Krejčí, P.; Brim, L. Toward Modelling and Analysis of Transient and Sustained Behaviour of Signalling Pathways. In Proceedings of the HSB 2016, Grenoble, France, 20–21 October 2016; Springer: New York, NY, USA, 2016; pp. 57–66.
59. Barnat, J.; Beneš, N.; Brim, L.; Demko, M.; Hajnal, M.; Pastva, S.; Šafránek, D. Detecting Attractors in Biological Models with Uncertain Parameters. In Proceedings of the International Conference on Computational Methods in Systems Biology (CMSB), Darmstadt, Germany, 27–29 September 2017; Springer: New York, NY, USA, 2017; pp. 40–56.

60. Beneš, N.; Brim, L.; Pastva, S.; Šafránek, D.; Troják, M.; Červeňy, J.; Šalagovič, J. Fully Automated Attractor Analysis of Cyanobacteria Models. In Proceedings of the 2018 22nd IEEE International Conference on System Theory, Control and Computing (ICSTCC), Sinaia, Romania, 10–12 October 2018; pp. 354–359.
61. Jia, B.; Jeon, C.O. High-throughput recombinant protein expression in *Escherichia coli*: current status and future perspectives. *Open Biol.* **2016**, *6*, 160196.
62. Vilar, J.M.; Guet, C.C.; Leibler, S. Modeling network dynamics: The lac operon, a case study. *J. Cell Biol.* **2003**, *161*, 471–476.
63. Santillán, M.; Mackey, M.; Zeron, E. Origin of bistability in the lac operon. *Biophys. J.* **2007**, *92*, 3830–3842.
64. Tessier, G. Les lois quantitatives de la croissance. *Ann. Physiol. Physiochim. Biol.* **1936**, *12*, 527–573.
65. Moser, H. *The Dynamics of Bacterial Populations Maintained in the Chemostat*; Carnegie Institution of Washington: Washington, DC, USA, 1958.
66. Edwards, V.H. The influence of high substrate concentrations on microbial kinetics. *Biotechnol. Bioeng.* **1970**, *12*, 679–712.
67. Andrews, J.F. A mathematical model for the continuous culture of microorganisms utilizing inhibitory substrates. *Biotechnol. Bioeng.* **1968**, *10*, 707–723.
68. Maier, R.M.; Pepper, I.L.; Gerba, C.P. *Environmental Microbiology*; Academic Press: New York, NY, USA, 2009; Volume 397.
69. Wang, P.; Robert, L.; Pelletier, J.; Dang, W.L.; Taddei, F.; Wright, A.; Jun, S. Robust growth of *Escherichia coli*. *Curr. Biol.* **2010**, *20*, 1099–1103.

© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).

Environmental Science & Technology

ACS Editors' Choice

Article

pubs.acs.org/est

Terms of Use

Immobilized Synthetic Pathway for Biodegradation of Toxic Recalcitrant Pollutant 1,2,3-Trichloropropane

Pavel Dvorak,[†,‡] Sarka Bidmanova,[†,‡,§] Jiri Damborsky,[†,‡,§] and Zbynek Prokop*[,†,‡,§]

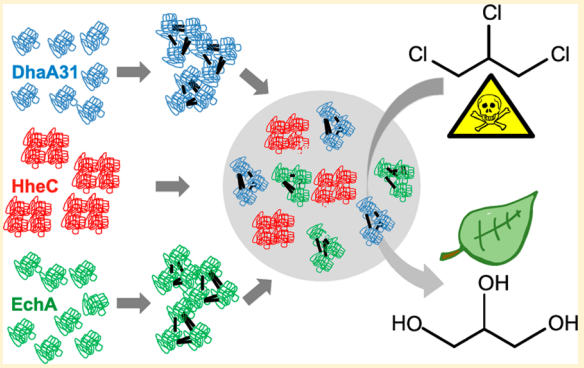
[†]Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic

[‡]International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, 656 91 Brno, Czech Republic

[§]Enantis, Ltd., Palackeho trida 1802/129, 612 00 Brno, Czech Republic

Supporting Information

ABSTRACT: The anthropogenic compound 1,2,3-trichloropropane (TCP) has recently drawn attention as an emerging groundwater contaminant. No living organism, natural or engineered, is capable of the efficient aerobic utilization of this toxic industrial waste product. We describe a novel biotechnology for transforming TCP based on an immobilized synthetic pathway. The pathway is composed of three enzymes from two different microorganisms: engineered haloalkane dehalogenase from *Rhodococcus rhodochrous* NCIMB 13064, and haloalcohol dehalogenase and epoxide hydrolase from *Agrobacterium radiobacter* AD1. Together, they catalyze consecutive reactions converting toxic TCP to harmless glycerol. The pathway was immobilized in the form of purified enzymes or cell-free extracts, and its performance was tested in batch and continuous systems. Using a packed bed reactor filled with the immobilized biocatalysts, 52.6 mmol of TCP was continuously converted into glycerol within 2.5 months of operation. The efficiency of the TCP conversion to the intermediates was 97%, and the efficiency of conversion to the final product glycerol was 78% during the operational period. Immobilized biocatalysts are suitable for removing TCP from contaminated water up to a 10 mM solubility limit, which is an order of magnitude higher than the concentration tolerated by living microorganisms.



INTRODUCTION

1,2,3-trichloropropane (TCP) is an anthropogenic compound recently recognized as an emerging contaminant of groundwater.[1,2] TCP is produced worldwide in quantities reaching 50 000 t annually and used by chemical companies as a solvent, precursor of soil fumigants, and building block for synthesis of other chemicals, e.g., dichloropropene or polysulfone liquid polymers.[3] TCP is also formed as a byproduct during the synthesis of epichlorohydrin. Due to its massive production, TCP can often be found at industrial and hazardous waste sites.[4] Recent incidents with drinking water sources in California polluted by TCP emphasized the need to develop efficient technologies for removing this toxic and carcinogenic compound from the environment.[5]

Conventional remediation techniques are relatively inefficient, due to the physical and chemical properties of the compound.[6] The exception is promising reductive conversion by zerovalent zinc.[6,7] In addition to abiotic transformations, biotransformations are extensively studied for their ability to decontaminate sites polluted with chemical contaminants.[8] Recently, isolated bacterial strains were found to transform TCP under anaerobic conditions.[9,10] However, these biotransformations often result in products toxic for surrounding environments, and due to the low efficiency can be applied only at limited (<1 mg·L^{-1}) TCP concentrations.[11] No organism capable of aerobic TCP biodegradation has yet been isolated, probably due to the anthropogenic nature of TCP and its recent introduction into the environment.

The absence of natural catabolic pathways for aerobic TCP utilization was addressed by Bosma and co-workers.[12,13] The authors assembled a synthetic metabolic pathway with the heterologous expression of haloalkane dehalogenase DhaA from *Rhodococcus* sp. in the natural host *Agrobacterium radiobacter* AD1, capable of utilization of haloalcohols. Engineered AD1 strain, expressing 5-times more active variant of haloalkane dehalogenase, showed a slow growth on 1 mM TCP in minimal medium. However, practical utility of this construct is limited by the toxicity of TCP for bacterial cells above 1 mM concentration, isufficient activity of engineered haloalkane dehalogenase, accumulation of toxic intermediates, and legislative barriers on the use of genetically modified microorganisms.[13] We recently addressed the problems of the low catalytic efficiency of the first dehalogenation step and

Received: January 23, 2014
Revised: April 30, 2014
Accepted: May 1, 2014
Published: May 1, 2014

Downloaded via MASARYK UNIV on September 2, 2021 at 20:39:00 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

 © 2014 American Chemical Society dx.doi.org/10.1021/es500396r | Environ. Sci. Technol. 2014, 48, 6859−6866

accumulation of intermediates limiting the productivity of the pathway in vivo by combination of protein and metabolic engineering.[14]

In this study, we report the application of an immobilized enzymatic pathway for a complete five-step degradation of TCP to the harmless product glycerol (GLY; Figure 1). We

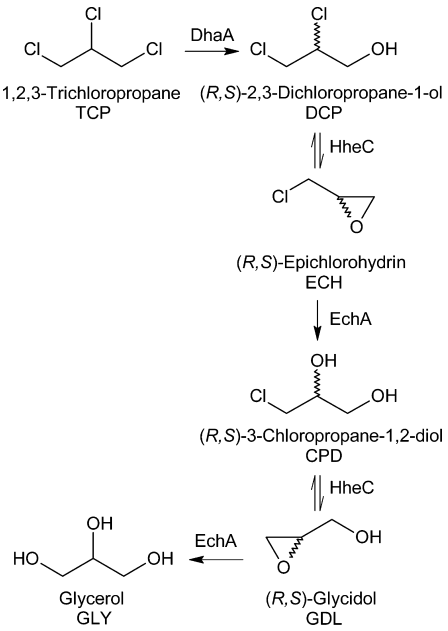


Figure 1. Scheme of the five-step synthetic metabolic pathway for biotransformation of 1,2,3-trichloropropane to glycerol. Abbreviations used for individual metabolites are shown. DhaA, haloalkane dehalogenase from *Rhodococcus rhodochrous* NCIMB 13064; HheC, haloalcohol dehalogenase; and EchA, epoxide hydrolase, both from *Agrobacterium radiobacter* AD1.

immobilized engineered 32-times more active haloalkane dehalogenase DhaA31 from *Rhodococcus rhodochrous* NCIMB 13064,[15] the wild-type haloalcohol dehalogenase HheC and the wild-type epoxide hydrolase EchA from *A. radiobacter* AD1 into cross-linked enzyme aggregates (CLEAs)[16] and poly(vinyl alcohol) (PVA) LentiKats lenses.[17] The immobilized biocatalysts were used to convert TCP into the desirable commodity chemical GLY in both batch and continuous systems. A comparison of the dynamic behavior of the complex multienzyme system in both soluble and immobilized forms is provided. To the best of our knowledge, this is the first report on the use of an immobilized synthetic metabolic pathway employing engineered enzyme for the biotransformation of an environmental pollutant. The established immobilization strategy is robust and suitable for scale-up. The developed biocatalyst is recyclable, resistant to biodegradation, compatible with high input loads of TCP, and can operate under mild nonsterile conditions. Moreover, the possibility to operate the process without the use of genetically modified microorganisms makes this biotechnology suitable for environmental applications.

EXPERIMENTAL SECTION

Reagents. TCP, 2,3-dichloropropane-1-ol (DCP), epichlorohydrin (ECH), 3-chloropropane-1,2-diol (CPD), glycidol (GDL), and GLY standards were purchased from Sigma-Aldrich (USA). All of the chemicals used in this study were of analytical grade. Bovine serum albumin (BSA) for preparation of CLEAs was purchased from Sigma-Aldrich (USA). PVA and poly(ethylene glycol) of M_W 1000 were provided by LentiKat's a.s. (Czech Republic). The solution of the cross-linker dextran polyaldehyde (DPA) was prepared according to a procedure described elsewhere.[18] A Free Glycerol Assay Kit was purchased from BioVision (USA). The work with toxic compounds was conducted in a fume hood and with protective equipment to minimize safety risks.

Preparation of Enzymes. The genes of DhaA, DhaA31, HheC, and EchA were commercially synthesized (Geneart, Germany). Synthetic genes were subcloned into pET21b (*dhaA*, *dhaA*31, and *hheC*), and pET28b (*echA*). The constructs were transformed into *E. coli* DH5α for plasmid propagation and into *E. coli* BL21(DE3) for enzyme expression. The recombinant His-tagged DhaA, DhaA31, and EchA were purified from cell-free extracts using 5 mL Ni-NTA Superflow column (Qiagen, Germany). Wild-type HheC was purified using Econo-Column (Bio-Rad, USA) packed with 25 mL of Q Sepharose Fast Flow (GE Healthcare, USA). Details of the expression and purification protocols are provided in Section 1 of the Supporting Information (SI).

Preparation of CLEAs. CLEAs of DhaA31 and EchA were prepared by dissolving 25 mg of enzyme and 25 mg of BSA in 1 mL of 10 mM phosphate buffer (pH 7.5). The solution was added to 9 mL of saturated ammonium sulfate (pH 8.0). After 1 h of incubation in an ice bath with stirring, 1.3 mL of DPA was added and cross-linking occurred for another 1 h. The resulting suspension was centrifuged at 4000*g* for 15 min at 4 °C. The supernatant was stored for the determination of residual enzymatic activity. CLEAs were resuspended in 20 mL of saturated sodium hydrogen carbonate and incubated with stirring in an ice bath for 30 min. The suspension was centrifuged at 4000*g* for 15 min at 4 °C and the supernatant was removed. CLEAs were washed with a 50 mM phosphate buffer (pH 7.5) and stored in 1 mL of this buffer at 4 °C before further use.

Encapsulation of Enzymes into PVA Particles Lenti-Kats. PVA (0.56 g) and poly(ethylene glycol) (0.34 g) were mixed with 3.7 mL of distilled water and heated at 98 °C until PVA dissolved completely. The liquid was cooled to 35 °C, and 10 mg of free HheC in a PD buffer (50 mM phosphate buffer of pH 7.5 with 2 mM dithiothreitol) or 1 g of CLEAs of EchA or DhaA31 was added and mixed thoroughly. Small droplets (3−4 mm) of mixture were dripped on plastic plates and incubated at 37 °C until LentiKats lost 70% of their initial wet weight. Dried LentiKats were soaked in 50 mL of 0.1 M sodium sulfate for 2 h to reswell. HheC LentiKats were washed with a PD buffer and stored in the same buffer at 4 °C. LentiKats of EchA and DhaA31 were washed with a 50 mM phosphate buffer (pH 7.5) and stored at 4 °C.

Enzyme Assays. Specific activities of soluble and immobilized DhaA31 with TCP and HheC with DCP and CPD were assayed in 10 mL of a 50 mM Tris-SO_4 buffer (pH 8.5) at 37 °C with a 10 mM substrate. The concentration of the reaction product (chloride ions) was measured by the method of Iwasaki[15,19] The specific activity of EchA was assayed by following the substrate depletion in 10 mL of a 50 mM Tris-SO_4 buffer (pH 8.5) at 37 °C with 5 mM ECH or 10 mM GDL. Samples of the reaction mixture were mixed with acetone (1:1) with internal standard hexanol and analyzed by gas chromatography (GC) as described in Section 11 of the SI. The same GC method was used for the quantitative analyses of

all metabolites of the TCP pathway except for GLY. The rates of abiotic conversions of all metabolites at selected time intervals were negligible.

Multienzyme Conversion of TCP in Batch System. The multienzyme conversion of TCP to GLY was assayed in 15 mL of a 50 mM Tris-SO_4 buffer (pH 8.5) in gastight glass vials (Sigma-Aldrich, USA) incubated at 37 °C. The reaction was initiated by adding 1 mg of DhaA31, HheC, and EchA in a soluble or immobilized form, into the reaction mixture with 5 mM TCP. The samples withdrawn from the reaction mixture (0.5 mL for soluble enzymes or CLEAs and 0.1 mL for LentiKats) were mixed with acetone (1:1) containing hexanol as an internal standard and analyzed by GC. Selected samples were analyzed by gas chromatograph and mass spectrometer (GC-MS) (Agilent, USA) for identification of metabolites otherwise routinely detected by GC (SI Section 11). The concentration of GLY was determined spectrophotometrically by the Free Glycerol Assay Kit. Samples of the reaction mixture (0.1 mL) were heated at 95 °C for 5 min, centrifuged at 18 000*g* for 1 min, diluted in an assay buffer, and analyzed according to the manufacturer's protocol. Concentrations of GLY were calculated from absorbance at 570 nm.

Multienzyme Conversion of TCP in Continuous System. Glass column 1 (28 cm high, 1.5 cm internal diameter, 50 mL working volume) with 100 mg of DhaA31 immobilized in 47 g of wet LentiKats and glass column 2 (25 cm high, 2.5 cm internal diameter, 100 mL working volume) with 100 mg of both HheC and EchA immobilized in 45 and 43 g of LentiKats, respectively, were used for the removal of TCP in a 10-week continuous operation of a packed bed reactor placed in fume hood at 22 ± 2 °C. A feed vessel contained TCP dissolved under stirring in 1 L of distilled water buffered with 0.1 M Tris-SO_4 (pH 8.2) to a final theoretical concentration of 5 mM (week 1) or 10 mM (weeks 2−10). The experimental concentrations of TCP, ranging from 2.25 to 7.97 mM, for evaluation of system efficiency were determined in the input of column 1 (SI Figure S8 and Table S6). The operation conditions for the packed bed reactor are summarized in SI Table S1. Influent and effluent lines were constructed from polytetrafluoroethylene tubing. Samples from the feed vessel, inlet of column 1, and effluent vessels 1 and 2 were withdrawn periodically and analyzed by GC and the Free Glycerol Assay Kit. A new cycle of operation was started whenever the content of the feed vessel was completely transferred to the effluent vessel 2. To evaluate possible leaching of the immobilized enzymes from the reactor, samples from effluent vessel 2 were withdrawn, concentrated 25 times using stirred ultrafiltration cells (Millipore, USA), and analyzed by SDS−polyacrylamide gel electrophoresis (SDS−PAGE).

RESULTS AND DISCUSSION

Multienzyme Conversion of TCP Using Free Enzymes. We initially tested the ability of an in vitro assembled pathway to fully convert 5 mM TCP to GLY in one-pot reaction at a time interval of 30 h. Purified DhaA, HheC, and EchA were mixed in the mass ratio of 1:1:1 mg and incubated with TCP. Because the molecular weights of DhaA, HheC, and EchA are similar (34, 29, and 35 kDa, respectively), the proposed 1:1:1 ratio corresponded closely with the same molar ratio of enzymes. The pH 8.5 and temperature 37 °C were selected to approach the reaction optima of all three enzymes.[20−22] Using wild-type enzymes, TCP and the intermediates DCP and GDL were degraded from 73% within 30 h of the reaction (Figure 2A and SI Table S2). The percentage of degradation correlated well with amount of GLY produced from TCP during the same time interval.

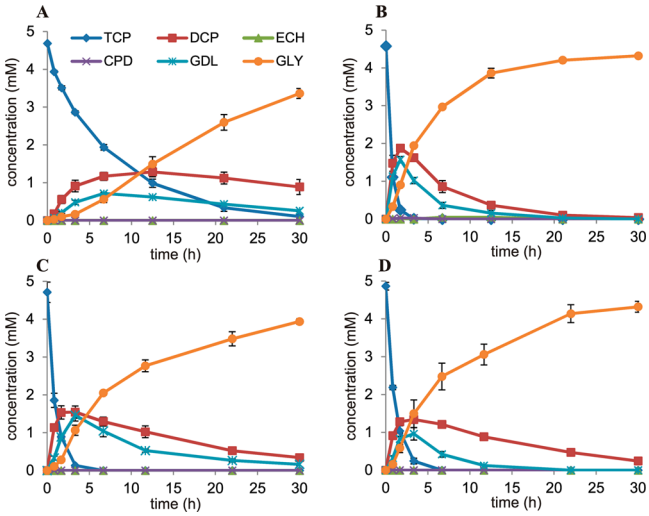


Figure 2. Time courses of multienzyme conversions of 1,2,3-trichloropropane with free and immobilized enzymes: (A) free purified enzymes DhaAwt, HheC, and EchA; (B) free purified enzymes DhaA31, HheC, and EchA; (C) immobilized purified enzymes DhaA31, HheC, and EchA; (D) immobilized cell-free extracts containing DhaA31, HheC, and EchA. All reactions were performed using enzymes of mass ratio of 1:1:1 mg in 15 mL of reaction mixture. TCP, 1,2,3-trichloropropane; DCP, 2,3-dichloropropane-1-ol; ECH, epichlorohydrin; CPD, 3-chloropropane-1,2-diol; GDL, glycidol; GLY, glycerol. Each data point represents the mean value ± standard deviation from three independent experiments.

The time course of the reaction showed the major bottleneck of the pathway: slow consumption of TCP by DhaA. Additionally, significant accumulations of two intermediates, DCP and GDL, were observed. The accumulation of DCP was caused by the high enantioselectivity ($E \geq 100$) of HheC.[23] Nonselective DhaA converts the prochiral TCP into both enantiomers of DCP in an almost equimolar ratio. Because HheC prefers (*R*)-DCP, (*S*)-DCP tends to accumulate (SI Section 2). The specific activities of EchA with ECH and GDL (29.5 and 6.5 μmol·min^{-1}·mg^{-1}, respectively) in combination with the previously reported kinetic parameters indicate that ECH is a much better substrate for EchA than GDL.[24] Therefore, the substrate specificity of EchA is probably the main cause of GDL accumulation during the multienzyme conversion of TCP.

To overcome the limitation of the first reaction step, the wild-type DhaA (k_{cat} = 0.04 s^{-1}, k_{cat}/K_m = 40 s^{-1}·M^{-1}) was substituted with the mutant DhaA31 (k_{cat} = 1.26 s^{-1}, k_{cat}/K_m = 1050 s^{-1}·M^{-1}), constructed in our laboratory using a computer-assisted design.[15] The selectivity of DhaA31 with TCP remained unchanged. The benefit of engineered DhaA31 in the multienzyme conversion of TCP was verified during the second experiment with free enzymes. The time course of the reaction shows that TCP was completely converted into its metabolites within the first 200 min of the reaction, and the degradation of TCP, DCP, and GDL reached 99% of the theoretical maximum within 30 h of the reaction (Figure 2B and SI Table S3). The conversion of TCP to GLY reached 95% of the theoretical maximum. We conclude that complete in

dx.doi.org/10.1021/es500396r | Environ. Sci. Technol. 2014, 48, 6859−6866

vitro biodegradation of TCP and its biotransformation to the final product GLY is possible despite the suboptimal stereochemistry and specificity in the pathway resulting in the accumulation of DCP and GDL in the initial phase of reaction.

Immobilization of DhaA31, HheC, and EchA. Application of free enzymes in biotransformation processes is not practical due to their complicated recycling, limited use in bioreactors, and low stability in harsh process conditions such as elevated temperatures or the presence of organic cosolvents.[25] We therefore immobilized DhaA31, HheC, and EchA to avoid such limitations. Various strategies for immobilization of individual biocatalysts are available.[26] Multienzyme systems also benefit from immobilization, but development of joint immobilization protocols for all employed catalysts is challenging. There is currently no protocol available for immobilization of DhaA, HheC, and EchA. Therefore, we tested enacapsulation of three enzymes into PVA, which has previously proven its utility in immobilization of haloalkane dehalogenase LinB, close relative of DhaA.[27]

The encapsulation of biocatalysts into PVA hydrogel is a widely used immobilization technique.[28,29] Lens-shaped PVA hydrogel particles, known as LentiKats, are promising matrices for biocatalysis due to their low cost, resistance to biodegradation, and good mechanical properties.[17] Their favorable geometry (thickness 200−400 μm) allows better mass transfer than spherical microbeads.[30,31] The biotechnology based on LentiKats has already been used in large scale applications, e.g., synthesis of beta-lactam antibiotics or denitrification in wastewater treatment. The size of LentiKats allows easy separation from the reaction mixture and is suitable for application in packed bed reactors, which can suffer from large pressure drops over the column when packed with particles of inadequate size.[32] LentiKats provide a hydrophilic environment with the pores sufficiently small to entrap whole cells, cross-linked enzymes,[30] cross-linked enzyme aggregates (CLEAs),[27,33] or free enzymes of molecular weights higher than 80 kDa.

In contrast to the HheC tetramer with a molecular weight of 116 kDa, smaller monomeric molecules of DhaA31 and EchA are not suitable for direct encapsulation in LentiKats. Therefore, aggregation and cross-linking of DhaA31 and EchA was carried out to produce CLEA particles. Immobilization in CLEAs is a straightforward technique which has been used for many enzymes, including hydrolases.[16,34,35] The protocols for preparation of CLEAs from LinB and for miscroscopic monitoring of immobilized biocatalyst were previously established in our laboratory.[27,36] Here we expand these protocols also for immobilization of DhaA31 and EchA. To enable detailed characterization of assembled pathway, immobilization was initially performed separately for each of the three purified enzymes. CLEAs of DhaA31 and EchA were prepared by mixing an enzyme solution with lyophilized bovine serum albumin, serving as a proteic feeder.[37] The protein mixture was precipitated using ammonium sulfate, which is widely used for preparing CLEAs due to its low cost and easy treatment.[38] The suspension of aggregated enzymes was cross-linked using DPA, which had been shown to have a less detrimental effect on enzymatic activity than the widely used glutaraldehyde.[39] CLEAs and free HheC were labeled by fluorescent dyes to allow microscopic monitoring of biocatalysts during the immobilization procedure (Figure 3). Free HheC and wet CLEAs of DhaA31 and EchA were separately encapsulated in LentiKats in mass ratio of 1:4 (enzyme or CLEAs:PVA gel).

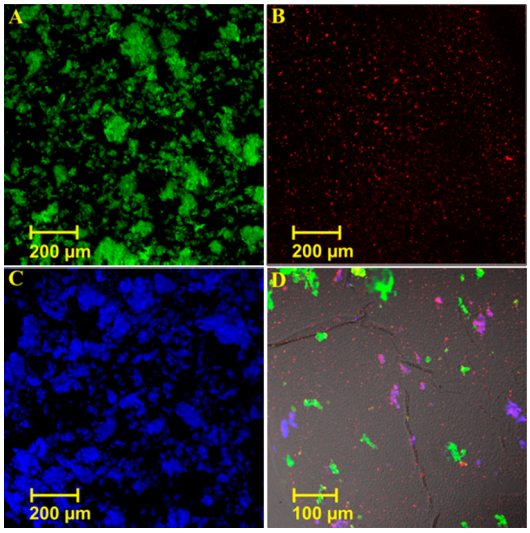


Figure 3. Confocal microscopy of immobilized enzymes from the synthetic pathway for biodegradation of 1,2,3-trichloropropane: (A) Cross-linked enzyme aggregates (CLEAs) of haloalkane dehalogenase DhaA31 labeled with fluorescein 5(6)-isothiocyanate; (B) free haloalcohol dehalogenase HheC labeled with X-rhodamine-5(6)-isothiocyanate; (C) CLEAs of epoxide hydrolase EchA labeled with Pacific Blue succinimidyl ester; (D) section through combi-LentiKat containing CLEAs of DhaA31, EchA, and free HheC. Visible fissures are caused by the fragility of PVA after freezing and appear after treatment with Leica Cryocut 1800 Cryostat (Leica Microsystems, Germany).

Characterization of Immobilized Enzymes. Immobilized enzymes were characterized in terms of their activity, storage and operational stability, and distribution in PVA matrix. Immobilization resulted in the decrease of specific activities of all three enzymes with their five corresponding substrates (SI Figure S2). The catalytic performance of CLEAs of DhaA31 was 80% of the initial activity of the soluble enzyme. CLEAs of EchA showed reduced activity with both GDL (50%) and ECH (37%). The encapsulation of free HheC and CLEAs of DhaA31 and EchA into PVA hydrogel led to activity retention of approximately 73/69% with DCP/CPD, 54% with TCP, and 17/36% with ECH/GDL, respectively. Observed activity retentions for immobilized DhaA and HheC are comparable with studies describing formation of CLEAs or LentiKats using other hydrolases, such as penicillin acylase, acetyl xylan esterase, or naringinase.[30,33,40] Some epoxide hydrolases have been immobilized with high activity retention using alternative methods, e.g., interaction with carrier through His-tag or covalent binding.[41,42] We verified these methods with EchA using Ni-NTA Agarose and activated CH Sepharose 4B as carriers, achieving 93% and 68% activity retention with ECH (SI Section 4). Nevertheless, high activity retention in these matrices is compromised by higher price and lower mechanical stability, which makes them less suitable for full-scale processes. The specific activity of soluble EchA with ECH (29.5 μmol·min^{-1}·mg^{-1}) is 1 order of magnitude higher than the specific activity of DhaA31 with TCP (1.1 μmol·min^{-1}·mg^{-1}) and HheC with DCP (1.6 μmol·min^{-1}·mg^{-1}) and CPD (3.0 μmol·min^{-1}·mg^{-1}). Despite the reduction of the catalytic performance of EchA after immobilization, the pattern of

dx.doi.org/10.1021/es500396r | Environ. Sci. Technol. 2014, 48, 6859−6866

specific activities of three enzymes remained unchanged (SI Figure S3). We therefore concluded that even 80% loss of the activity of EchA toward ECH after immobilization in CLEAs and LentiKats would not result in the accumulation of intermediates in the TCP pathway and the analogous immobilization strategy was maintained also for this enzyme.

No enzymatic activities were detected in the supernatants obtained during the preparation of CLEAs and LentiKats, besides low activity of HheC corresponding to approximately 10% loss of enzyme during immobilization. Therefore, we expect that the decreased activities of DhaA31, HheC, and EchA are predominantly due to the partial deactivation caused by aggregation, cross-linking, or encapsulation in PVA hydrogel, rather than due to the leaching of enzymes during immobilization process.

The storage stability of LentiKats of DhaA31 and EchA in a phosphate buffer without additives at 4 °C was assayed during a 3-month period and compared with the storage stability of free enzymes (SI Section 5). After this period, changes in the activity of immobilized and free enzymes correlated for both DhaA31 and EchA. Free and immobilized DhaA31 retained 93% and 86% of its initial activity with TCP, respectively. EchA showed 121% and 122% of its initial activity in free and immobilized form, respectively. The increase was statistically significant for free EchA (t test, $p < 0.05$). Similar moderate increases in specific activity during storage were reported also for some other enzymes.[27,30]

HheC showed significantly lower storage stability than DhaA31 and EchA. Free HheC stored in a phosphate buffer at 4 °C lost all its activity with DCP within 2 months (data not shown). Tang and co-workers previously suggested that inactivation of HheC is caused by the monomerization of the enzyme and intramolecular disulfide bond formation under oxidizing conditions.[43] Confirming this suggestion, the stability of HheC was significantly improved by its storage in a buffer containing a reducing agent (SI Figure S4A). Free and immobilized HheC retained 63% and 49% of its initial activity, respectively, after 2 months of storage in the presence of 2 mM dithiothreitol. The lower activity of HheC immobilized in LentiKats is due to partial leaching (SI Figure S4B). The storage stability of free enzymes was tested also at room temperature (22 ± 2 °C), which was later applied during the continuous removal of TCP in a packed bed reactor (SI Section 5). In contrast to DhaA31 and EchA, HheC showed better storage stability at room temperature than at 4 °C (SI Figures S4 and S5). The enzyme retained 70% of its initial activity after 10 weeks of storage at room temperature in phosphate buffer without any additive.

Multienzyme Conversion of TCP Using Immobilized Enzymes. Numerous studies describe the immobilization of individual enzymes or whole cells, but reports on the effects of immobilization on the performance of complete biochemical pathways are scarce.[44] For synthetic biodegradation pathways employing engineered enzymes such reports are to the best of our knowledge missing. We applied TCP pathway immobilized in LentiKats in a one-pot multienzyme reaction to study how the modified activities of individual immobilized enzymes affected the time course of TCP biodegradation. Immobilized enzymes were mixed in a ratio 1:1:1 mg and incubated with 5 mM TCP under the same conditions as had been used for soluble enzymes (Figure 2C and SI Table S4). Despite the fact that activity of individual enzymes after immobilization was reduced by 27−83%, the efficiency of the whole pathway was lower only about 11% when compared to the conversion with free enzymes. The conversion of TCP to GLY after 30 h reached 84% and conversion of all intermediates reached 89% of the theoretical maximum. We also studied the reusability of the immobilized TCP pathway in a batch system (SI Section 6). The immobilized pathway retained 77% of its initial efficiency of TCP conversion to GLY after 10 successive cycles of batch operation.

The application of purified enzymes immobilized individually provides better control over the reaction by enabling the tuning of enzyme stoichiometry and by compensating for the activity loss of individual enzymes. However, the purification and separate treatment of all three enzymes increases the cost of the biocatalyst. We therefore demonstrated that cell-free extracts obtained from *E. coli* BL21(DE3) cells expressing DhaA31, HheC, or EchA can be utilized. For the preparation of the LentiKats, cell-free extracts were mixed to provide a ratio of enzymes corresponding to 1:1:1 (SI Section 7). The combi-LentiKats containing approximately 1 mg of each of the three enzymes were used for the biodegradation of 5 mM TCP. The degradation of TCP and the intermediates DCP and GDL reached 95% of the theoretical maximum within 30 h of reaction (Figure 2D, SI Table S5). The overall reaction profile was very similar to the profile obtained with purified immobilized enzymes (Figure 2C). The conversion of TCP to GLY reached 89% of the theoretical maximum. Thus, the coimmobilization of enzymes and their improved proximity resulted in conversion comparable by efficiency with free enzymes. The proximity and homogeneous distribution of DhaA31, HheC, and EchA in combi-LentiKats was verified by confocal microscopy (Figure 3D).

Multienzyme Conversion of TCP Using Packed Bed Reactor. The performance of immobilized pathway was also tested under continuous operation using a packed bed reactor composed of two columns (SI Figure S8). Column 1 was packed with LentiKats of DhaA31, and column 2 was packed with a mixture of LentiKats of HheC and EchA. The total mass ratio of enzymes 1:1:1 corresponded to the previous batch experiments. DhaA31 was separated from HheC and EchA in order to prevent the inhibition of the last step in the pathway, the conversion of GLD to GLY by TCP (K_i = 0.21 mM). EchA was combined with HheC in column 2 in order to prevent a reverse reaction and push the reaction equilibrium during the conversion of DCP and CPD by HheC toward the product (SI Figure S1).[23] The separation of enzymes into two columns also enabled better control over the individual reaction steps and evaluation of their efficiency.

The packed bed reactor was operated at room temperature 22 ± 2 °C and pH 8.2 for 10 weeks under the conditions described in SI Table S1. During that period, the reactor with an effective volume of 0.15 L processed 11 L of contaminated water. Experimental concentrations of TCP for evaluating column 1 efficiency were determined in the input of the column 1 due to the leakage of TCP from the pumping system before reaching the column (SI Figure S8 and Table S6). The average leakage caused by hydrophobic nature of TCP (log P = 2.24) and its tendency to penetrate through rubber tubing of peristaltic pump was approximately 29%. The experimental concentrations of TCP in the column 1 input for each of 10 weeks of operation were 2.25, 3.03, 5.91, 6.54, 7.97, 6.31, 7.94, 6.65, 6.80, and 7.02 mM. The levels of residual TCP and produced DCP were determined in the effluent of column 1. The levels of residual DCP and produced GDL and GLY were

determined in the effluent of column 2. These values were used to evaluate the efficiency of both columns and the efficiency of GLY production from TCP (Figure 4). No leakage of intermediates was observed. Neither ECH nor CPD was detected in the system.

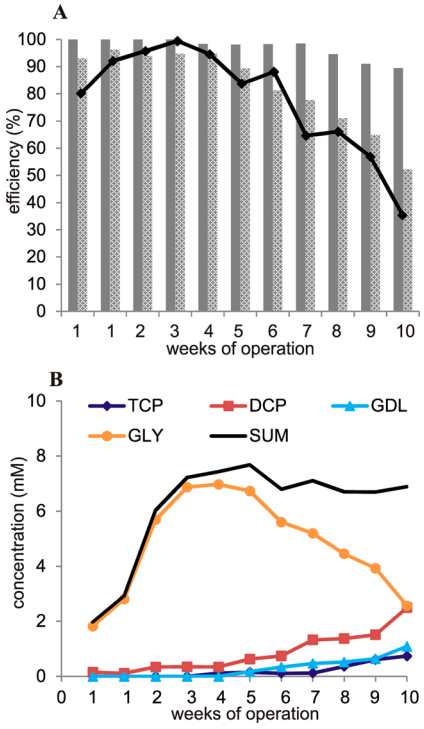


Figure 4. (A) Efficiency of column 1 (filled) and column 2 (cross-hatched) of the packed bed reactor during the continuous biodegradation of 1,2,3-trichloropropane (TCP). Black diamonds with a solid line show the overall efficiency of TCP conversion to glycerol by immobilized biocatalysts in the two-step packed bed reactor. Efficiency was calculated from the concentration of TCP measured in the inlet of column 1 and the concentration of glycerol measured in effluent vessel 2. (B) Residual concentrations of the metabolites from the synthetic pathway for biodegradation of TCP detected in effluent vessel 1 (TCP) and effluent vessel 2 (2,3-dichloropropane-1-ol, DCP; glycidol, GDL; glycerol, GLY) during 10 weeks of operating the packed bed reactor.

The efficiencies of column 1 and 2 were higher than 95% during the first 8 and 4 weeks of operation, respectively (Figure 4 and SI Table S6). The efficiency of the overall conversion of TCP to GLY decreased from values above 90% observed during the first 4 weeks to values lower than 60% at the end of 2 months of operation. The decreased reactor efficiency accompanied by an accumulation of unreacted TCP, DCP, and GDL in effluent vessels could be attributed to (i) slow enzyme inactivation caused by oxidation of HheC and thermal unfolding of DhaA31 and EchA and (ii) slow leaching of enzymes from the reactor, which resulted in about 10% loss of HheC activity and about 5% loss of DhaA31 and EchA activities during the operation (Figure 4B and SI Figure S9). In total, 65.5 mmol of the 67.7 mmol of TCP that entered the packed bed reactor during 10 weeks of operation was converted to intermediates (efficiency 97%), and 52.6 mmol of GLY was produced (efficiency 78%).

Perspectives on In Vitro Biotransformation of TCP. Recent studies on bacterial utilization of 1,3-dichloroprop-1-ene and chlorinated ethenes suggest that the biodegradation of chlorinated aliphatic pollutants is associated with the accumulation of reactive intermediates and strong oxidative stress, representing a significant barrier for the evolution of corresponding aerobic metabolic pathways in vivo.[45,46] The absence of natural microorganisms carrying aerobic pathways for the biodegradation of TCP and experiences gained during attempts to engineer microorganisms growing on this toxic compound seem to support this view.[11−14] Engineering synthetic pathway for converting toxic TCP to the harmless GLY in vitro is an alternative approach, which does not suffer from the limitations of engineered bacterium. It is now widely accepted that in vitro multienzyme systems represent an emerging field of biocatalysis due to their simplicity, predictability, and controllability.[47,48] This study shows that an in vitro assembly of natural or synthetic enzymatic pathways can be a promising concept for the biodegradation of environmental pollutants and can provide promising results especially when combined with protein engineering techniques.

Developed biotechnology requires further validation before it can be scaled-up and used in real conditions. The performance of the immobilized biocatalyst should be tested in real water samples contaminated with TCP. Pretreatment of the contaminated water by adjusting high salinity or eliminating possible enzyme-inhibiting constituents might be necessary. Conversion of TCP to GLY was demonstrated to be efficient at 22 ± 2 °C and is also possible in broad pH range 7−10 (SI Section 10). Yet, buffering of the contaminated water to pH close to the reaction optima of enzymes can be beneficial to achieve maximal efficiency of the process. The favorable features of the presented biotechnology are (i) degradation of TCP using immobilized cell-free extracts instead of purified enzymes is possible, (ii) material used for immobilization of enzymes is affordable, safe, and nonbiodegradable, (iii) protocol for disposal of used LentiKats by burning is well established, and (iv) the amount of enzymes necessary to degrade almost 10 g of TCP during the operation of reactor can be obtained from less than 2 L of cell culture yet without optimized cultivation conditions. An immobilized synthetic pathway works at TCP concentrations that are close to the water solubility limit of TCP (10 mM), which is 1 order of magnitude higher than concentrations tolerated by engineered microorganisms.[13] At the same time, the system has the potential to cope with significantly lower concentrations of TCP. Kinetic parameters of DhaA31 for TCP (k_{cat} = 1.26 s^{-1}, k_{cat}/K_m = 1050 $s^{-1}{\cdot}M^{-1}$) are of the same order of magnitude as those of the haloalkane dehalogenase DhlA for 1,2-dichloroethane (k_{cat} = 3.3 s^{-1}, k_{cat}/K_m = 6200 $s^{-1}{\cdot}M^{-1}$), which has already proven its utility in a groundwater purification plant treating micro- to nanomolar concentrations of the pollutant.[49]

The remaining bottlenecks of the pathway are (i) lower activity of HheC with nonpreferred (*S*)-DCP, (ii) accumulation of GDL due to the substrate preference of EchA, and (iii) gradual inactivation of HheC. These important, but not critical, limitations can be overcome by another round of protein engineering or modification of the immobilization protocol. The flux of intermediates through the immobilized pathway can be further tuned using kinetic modeling and optimization of enzyme stoichiometry, which is an objective of our future research.

■ ASSOCIATED CONTENT

Supporting Information

Details on the preparation of enzymes, the activity, storage stability, and recycling of immobilized enzymes, the preparation of combi-LentiKats from cell-free extracts, construction and operational conditions of packed bad reactor and the analysis of enzyme leaching, multienzyme conversion of TCP at pH 7 and 10 at 20 °C, and some analytical procedures (e.g., GC, chiral GC, and GC/MS) together with supporting tables. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*Phone: +420 549 496 667; fax: +420 549 4962 03; e-mail: zbynek@chemi.muni.cz; mail: Kamenice 5/A13, Brno 625 00, Czech Republic.

Notes

The authors declare the following competing financial interest(s): Jiri Damborsky and Zbynek Prokop are founders of Enantis s.r.o., a biotechnology spin-off company from Masaryk University.

■ ACKNOWLEDGMENTS

We thank Dr. Josef Jaros, Eva Hrdlickova, and Katarina Tvaruzkova (Masaryk University, Brno, Czech Republic) for their technical assistance during the confocal microscopy and enzyme preparation, and Dr. Radek Stloukal (LentiKat's a.s.) for stimulating discussions. This research was supported by the Grant Agency of the Czech Republic (P503/12/0572 to J.D.), Support of Study Stays of Czech Researchers Abroad III: Young Talent Incubator (CZ.1.07/2.3.00/20.0239 to P.D.), the European Regional Development Fund (CZ.1.05/2.1.00/01.0001 to Z.P., CZ.1.05/1.1.00/02.0123 to J.D.), the European Union Framework Programme (REGPOT 316345 to Z.P. and S.B.), and the Czech Ministry of Education (CZ.1.07/2.3.00/20.0239 to P.D. and CZ.1.07/2.3.00/20.0183 to S.B.).

■ REFERENCES

(1) *Toxicological Review of 1,2,3-Trichloropropane; CAS No. 96-18-4*; U.S. Environmental Protection Agency: Washington, DC, 2009; http://www.epa.gov/iris/toxreviews/0200tr.pdf.

(2) California Department of Public Health. *1,2,3-Trichloropropane*; 2009. http://cdph.ca.gov/certlic/drinkingwater/Pages/123tcp.aspx. (accessed 30 October 2013).

(3) *Product Safety Assessment: 1,2,3-Trichloropropane*; The Dow Chemical Company, 2008; http://msdssearch.dow.com/PublishedLiteratureDOWCOM/dh_031c/0901b8038031c3b0.pdf?filepath=productsafety/pdfs/noreg/233-00534.pdf&fromPage=GetDoc (accessed 30 October 2013).

(4) *Emerging Contaminant – 1,2,3-Trichloropropane (TCP)*; Fact Sheet; U.S. Environmental Protection Agency: Washington, DC, 2009.

(5) *Dow's Cancer-causing "Garbage" Chemical in Drinking Water*; http://www.panna.org/blog/dows-cancer-causing-garbage-chemical-drinking-water (accessed 30 October 2013).

(6) Sarathy, V.; Salter, A. J.; Nurmi, J. T.; O'Brien Johnson, G.; Johnson, R. L.; Tratnyek, P. G. Degradation of 1,2,3-trichloropropane (TCP): Hydrolysis, elimination, and reduction by iron and zinc. *Environ. Sci. Technol.* **2010**, *44*, 787−793.

(7) Salter-Blanc, A. J.; Tratnyek, P. G. Effects of solution chemistry on the dechlorination of 1,2,3-trichloropropane by zero-valent zinc. *Environ. Sci. Technol.* **2011**, *45*, 4073−4079.

(8) Copley, S. D. Evolution of efficient pathways for degradation of anthropogenic chemicals. *Nat. Chem. Biol.* **2009**, *5*, 559−566.

(9) Bowman, K. S.; Nobre, M. F.; da Costa, M. S.; Rainey, F. A.; Moe, W. M. *Dehalogenimonas alkenigignens* sp. nov., a chlorinated-alkane-dehalogenating bacterium isolated from groundwater. *Int. J. Syst. Evol. Microbiol.* **2013**, *63*, 1492−1498.

(10) Yan, J.; Rash, B. A.; Rainey, F. A.; Moe, W. M. Isolation of novel bacteria within the *Chloroflexi* capable of reductive dechlorination of 1,2,3-trichloropropane. *Environ. Microbiol.* **2009**, *11*, 833−843.

(11) Samin, G.; Janssen, D. B. Transformation and biodegradation of 1,2,3-trichloropropane (TCP). *Environ. Sci. Pollut. Res. Int.* **2012**, *19*, 3067−3078.

(12) Bosma, T.; Kruizinga, E.; de Bruin, E. J.; Poelarends, G. J.; Janssen, D. B. Utilization of trihalogenated propanes by *Agrobacterium radiobacter* AD1 through heterologous expression of the haloalkane dehalogenase from *Rhodococcus* sp. strain M15-3. *Appl. Environ. Microbiol.* **1999**, *65*, 4575−4581.

(13) Bosma, T.; Damborský, J.; Stucki, G.; Janssen, D. B. Biodegradation of 1,2,3-trichloropropane through directed evolution and heterologous expression of a haloalkane dehalogenase gene. *Appl. Environ. Microbiol.* **2002**, *68*, 3582−3587.

(14) Kurumbang, N. P.; Dvorak, P.; Bendl, J.; Brezovsky, J.; Prokop, Z.; Damborsky, J. Computer-assisted engineering of the synthetic pathway for biodegradation of a toxic persistent pollutant. *ACS Synth. Biol.* **2013**, *3*, 172−181.

(15) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. Redesigning dehalogenase access tunnels as a strategy for degrading an anthropogenic substrate. *Nat. Chem. Biol.* **2009**, *5*, 727−733.

(16) Cao, L.; van Rantwijk, F.; Sheldon, R. A. Cross-linked enzyme aggregates: A simple and effective method for the immobilization of penicillin acylase. *Org. Lett.* **2000**, *2*, 1361−1364.

(17) Schlieker, M.; Vorlop, K.-D. A novel immobilization method for entrapment: LentiKats®. In *Immobilization of Enzymes and Cells*; Guisan, J. M., Ed.; Humana Press: Totowa, NJ, 2006; Vol. *22*, pp 333−343.

(18) Drobchenko, S. N.; Isaeva-Ivanova, L. S.; Kleiner, A. R.; Lomakin, A. V.; Kolker, A. R.; Noskin, V. A. An investigation of the structure of periodate-oxidised dextran. *Carbohydr. Res.* **1993**, *241*, 189−199.

(19) Iwasaki, I.; Utsumi, S.; Ozawa, T. New colorimetric determination of chloride using mercuric thiocyanate and ferric ion. *Bull. Chem. Soc. Jpn.* **1952**, *25*, 226.

(20) Kulakova, A. N.; Larkin, M. J.; Kulakov, L. A. The plasmid-located haloalkane dehalogenase gene from *Rhodococcus rhodochrous* NCIMB 13064. *Microbiology* **1997**, *143*, 109−115.

(21) Van Hylckama Vlieg, J. E.; Tang, L.; Lutje Spelberg, J. H.; Smilda, T.; Poelarends, G. J.; Bosma, T.; van Merode, A. E.; Fraaije, M. W.; Janssen, D. B. Halohydrin dehalogenases are structurally and mechanistically related to short-chain dehydrogenases/reductases. *J. Bacteriol.* **2001**, *183*, 5058−5066.

(22) Jacobs, M. H.; Van den Wijngaard, A. J.; Pentenga, M.; Janssen, D. B. Characterization of the epoxide hydrolase from an epichlorohydrin-degrading *Pseudomonas* sp. *Eur. J. Biochem. FEBS* **1991**, *202*, 1217−1222.

(23) Spelberg, J. L. Enantioselective biocatalytic conversions of epoxides. Ph.D. Dissertation, University of Groningen: Groningen, 2003.

(24) Rink, R.; Fennema, M.; Smids, M.; Dehmel, U.; Janssen, D. B. Primary structure and catalytic mechanism of the epoxide hydrolase from *Agrobacterium radiobacter* AD1. *J. Biol. Chem.* **1997**, *272*, 14650−14657.

(25) Ba, S.; Arsenault, A.; Hassani, T.; Jones, J. P.; Cabana, H. Laccase immobilization and insolubilization: From fundamentals to applications for the elimination of emerging contaminants in wastewater treatment. *Crit. Rev. Biotechnol.* **2013**, *33*, 404−418.

(26) Guisán, J. M., Ed. *Methods in Biotechnology: Immobilization of Enzymes and Cells*, 2nd ed.; Humana Press Inc.: Totowa, NJ, 2006.

(27) Bidmanova, S.; Damborsky, J.; Prokop, Z. Immobilization of haloalkane dehalogenase LinB from *Sphingobium japonicum* UT26 for biotechnological applications. *J. Biocatal. Biotransform.* **2013**, *2*, 1−7.

(28) Jurchescu, I.-M.; Hamann, J.; Zhou, X.; Ortmann, T.; Kuenz, A.; Prüße, U.; Lang, S. Enhanced 2,3-butanediol production in fed-batch cultures of free and immobilized *Bacillus licheniformis* DSM 8785. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 6715−6723.
(29) Cárdenas-Fernández, M.; Neto, W.; López, C.; Álvaro, G.; Tufvesson, P.; Woodley, J. M. Immobilization of *Escherichia coli* containing ω-transaminase activity in LentiKats®. *Biotechnol. Prog.* **2012**, *28*, 693−698.
(30) Nunes, M. A. P.; Fernandes, P. C. B.; Ribeiro, M. H. L. High-affinity water-soluble system for efficient naringinase immobilization in polyvinyl alcohol-dimethyl sulfoxide lens-shaped particles. *J. Mol. Recognit.* **2012**, *25*, 580−594.
(31) Czichocki, G.; Dautzenberg, H.; Capan, E.; Vorlop, K. D. New and effective entrapment of polyelectrolyte-enzyme-complexes in LentiKats. *Biotechnol. Lett.* **2001**, *23*, 1303−1307.
(32) Sheldon, R. A. Characteristic features and biotechnological applications of cross-linked enzyme aggregates (CLEAs). *Appl. Microbiol. Biotechnol.* **2011**, *92*, 467−477.
(33) Wilson, L.; Illanes, A.; Pessela, B. C. C.; Abian, O.; Fernández-Lafuente, R.; Guisán, J. M. Encapsulation of crosslinked penicillin G acylase aggregates in lentikats: Evaluation of a novel biocatalyst in organic media. *Biotechnol. Bioeng.* **2004**, *86*, 558−562.
(34) Sheldon, R. A. Cross-linked enzyme aggregates (CLEAs): Stable and recyclable biocatalysts. *Biochem. Soc. Trans.* **2007**, *35*, 1583−1587.
(35) Tran, D. N.; Balkus, K. J. Perspective of recent progress in immobilization of enzymes. *ACS Catal.* **2011**, *1*, 956−968.
(36) Bidmanova, S.; Hrdlickova, E.; Jaros, J.; Ilkovics, L.; Hampl, A.; Damborsky, J.; Prokop, Z. Microscopic monitoring provides information on structure and properties during biocatalyst immobilization. *Biotechnol. J.* **2014**, DOI: 10.1002/biot.201300049.
(37) Shah, S.; Sharma, A.; Gupta, M. N. Preparation of cross-linked enzyme aggregates by using bovine serum albumin as a proteic feeder. *Anal. Biochem.* **2006**, *351*, 207−213.
(38) Schoevaart, R.; Wolbers, M. W.; Golubovic, M.; Ottens, M.; Kieboom, A. P. G.; van Rantwijk, F.; van der Wielen, L. A. M.; Sheldon, R. A. Preparation, optimization, and structures of cross-linked enzyme aggregates (CLEAs). *Biotechnol. Bioeng.* **2004**, *87*, 754−762.
(39) Mateo, C.; Palomo, J. M.; van Langen, L. M.; van Rantwijk, F.; Sheldon, R. A. A new, mild cross-linking methodology to prepare cross-linked enzyme aggregates. *Biotechnol. Bioeng.* **2004**, *86*, 273−276.
(40) Montoro-García, S.; Gil-Ortiz, F.; Navarro-Fernández, J.; Rubio, V.; García-Carmona, F.; Sánchez-Ferrer, Á. Improved cross-linked enzyme aggregates for the production of desacetyl ß-lactam antibiotics intermediates. *Bioresour. Technol.* **2010**, *101*, 331−336.
(41) Choi, S. H.; Kim, H. S.; Lee, I. S.; Lee, E. Y. Functional expression and magnetic nanoparticle-based immobilization of a protein-engineered marine fish epoxide hydrolase of *Mugil cephalus* for enantioselective hydrolysis of racemic styrene oxide. *Biotechnol. Lett.* **2010**, *32*, 1685−1691.
(42) Yildirim, D.; Tükel, S. S.; Alagöz, D.; Alptekin, Ö. Preparative-scale kinetic resolution of racemic styrene oxide by immobilized epoxide hydrolase. *Enzyme Microb. Technol.* **2011**, *49*, 555−559.
(43) Tang, L.; van Hylckama Vlieg, J. E. .; Lutje Spelberg, J. H.; Fraaije, M. W.; Janssen, D. B. Improved stability of halohydrin dehalogenase from *Agrobacterium radiobacter* AD1 by replacement of cysteine residues. *Enzyme Microb. Technol.* **2002**, *30*, 251−258.
(44) Scism, R. A.; Bachmann, B. O. Five-component cascade synthesis of nucleotide analogues in an engineered self-immobilized enzyme aggregate. *ChemBioChem* **2010**, *11*, 67−70.
(45) Nikel, P. I.; Pérez-Pantoja, D.; de Lorenzo, V. Why are chlorinated pollutants so difficult to degrade aerobically? Redox stress limits 1,3-dichloroprop-1-ene metabolism by *Pseudomonas pavonaceae*. *Philos. Trans. R. Soc. London B. Biol. Sci.* **2013**, *368*, 20120377.
(46) Rui, L.; Kwon, Y. M.; Reardon, K. F.; Wood, T. K. Metabolic pathway engineering to enhance aerobic degradation of chlorinated ethenes and to reduce their toxicity by cloning a novel glutathione S-transferase, an evolved toluene o-monooxygenase, and gamma-glutamylcysteine synthetase. *Environ. Microbiol.* **2004**, *6*, 491−500.
(47) Hodgman, C. E.; Jewett, M. C. Cell-free synthetic biology: Thinking outside the cell. *Metab. Eng.* **2012**, *14*, 261−269.
(48) Hold, C.; Panke, S. Towards the engineering of in vitro systems. *J. R. Soc. Interface R. Soc.* **2009**, *6* (Suppl 4), S507−521.
(49) Stucki, G.; Thueer, M. Experiences of a large-scale application of 1,2-dichloroethane degrading microorganisms for groundwater treatment. *Environ. Sci. Technol.* **1995**, *29*, 2339−2345.

ACS Catalysis

Research Article

pubs.acs.org/acscatalysis

Engineering a de Novo Transport Tunnel

Jan Brezovsky,[†,&] Petra Babkova,[†,&] Oksana Degtjarik,[‡,§] Andrea Fortova,[†] Artur Gora,[†,#] Iuliia Iermak,[‡,§] Pavlina Rezacova,[∥,⊥] Pavel Dvorak,[†] Ivana Kuta Smatanova,[‡,§] Zbynek Prokop,[†] Radka Chaloupkova,[†] and Jiri Damborsky*[,†]

[†]Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, Kamenice 5/A13, 625 00 Brno, Czech Republic

[‡]Faculty of Science, University of South Bohemia in Ceske Budejovice, Branisovska 1760, 37005 Ceske Budejovice, Czech Republic

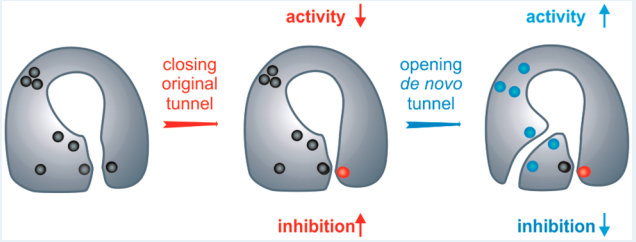
[§]Center for Nanobiology and Structural Biology ASCR, Zamek 136, 37333 Nove Hrady, Czech Republic

[∥]Institute of Organic Chemistry and Biochemistry ASCR, v.v.i. Flemingovo nam. 2, 166 10 Prague 6, Czech Republic

[⊥]Institute of Molecular Genetics of the ASCR, v.v.i. Videnska 1083, 142 20 Prague 4, Czech Republic

Supporting Information

ABSTRACT: Transport of ligands between buried active sites and bulk solvent is a key step in the catalytic cycle of many enzymes. The absence of evolutionary optimized transport tunnels is an important barrier limiting the efficiency of biocatalysts prepared by computational design. Creating a structurally defined and functional "hole" into the protein represents an engineering challenge. Here we describe the computational design and directed evolution of a de novo transport tunnel in haloalkane dehalogenase. Mutants with a blocked native tunnel and newly opened auxiliary tunnel in a distinct part of the structure showed dramatically modified properties. The mutants with blocked tunnels acquired specificity never observed with native family members: up to 32 times increased substrate inhibition and 17 times reduced catalytic rates. Opening of the auxiliary tunnel resulted in specificity and substrate inhibition similar to those of the native enzyme and the most proficient haloalkane dehalogenase reported to date (k_{cat} = 57 s^{-1} with 1,2-dibromoethane at 37 °C and pH 8.6). Crystallographic analysis and molecular dynamics simulations confirmed the successful introduction of a structurally defined and functional transport tunnel. Our study demonstrates that, whereas we can open the transport tunnels with reasonable proficiency, we cannot accurately predict the effects of such change on the catalytic properties. We propose that one way to increase efficiency of an enzyme is the direct its substrates and products into spatially distinct tunnels. The results clearly show the benefits of enzymes with de novo transport tunnels, and we anticipate that this engineering strategy will facilitate the creation of a wide range of useful biocatalysts.



KEYWORDS: *transport tunnel, protein engineering, protein design, activity, specificity, substrate inhibition, stability, substrate binding, product release, water dynamics*

Downloaded via MASARYK UNIV on September 2, 2021 at 20:45:34 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

Enzymes are very efficient and specific biocatalysts that have naturally evolved to convert a wide range of compounds. These reactions take place in the enzyme's active sites, which contain several essential catalytic residues.[1] The active site of many enzymes is buried inside the protein core and is connected to the bulk solvent through transport tunnels.[2,3] The anatomy, physicochemical properties, and dynamics of these tunnels have a large influence on enzymatic catalysis by determining the exchange rates of substrates or products between the active sites and solvent environment.[2,4,5] Tunnels may also serve more sophisticated functions by (i) allowing access of preferred substrates, while denying access of nonpreferred substrates,[6,7] (ii) avoiding damage of the enzymes containing the transition metals by their poisoning,[7,8] (iii) preventing damage of cellular organelles by toxic intermediates,[9] (iv) enabling reactions requiring the absence of water,[10] and (v) synchronizing reactions involving a number of substrates, intermediates, or cofactors.[11,12] Tunnels can contain one or more molecular gates, which fine-tune their dynamical properties and enable their open and closed states.[6,13−15]

Enzymes are commonly used for the production of chemicals, pharmaceuticals, food, agricultural additives, and biofuels.[16] Engineering of enzymes is often required to overcome their natural limitations. The importance of transport processes for enzymatic catalysis has prompted many engineering studies to focus on protein tunnels.[2] Such studies have provided enzymes with substantially altered activity,[4,17,18] specificity,[19,20] enantioselectivity,[21] and stability.[22−24] Investigation of de novo designed Kemp eliminase showed that,

Received: July 23, 2016
Revised: September 15, 2016
Published: September 23, 2016

 © 2016 American Chemical Society

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

although it is possible to introduce a novel catalytic function into an existing protein core, several rounds of random mutagenesis are necessary to enhance the desired activity to a sufficient level. Notably, a number of random mutations were found to open a ligand transport pathway.[25] This result suggests that de novo design of novel catalysts needs to consider not only introduction of a proper active site into the protein scaffold[26] but also creation of appropriate tunnels for efficient exchange of substrates and products between the designed active site and surrounding environment.

Despite the obvious importance of tunnels to enzymatic catalysis and effect of their engineering, we are not aware of any attempt to introduce an entirely new transport route into the existing protein scaffold. In this study, we address emerging questions regarding whether it is possible to design and introduce the ligand transport tunnel de novo. We closed the main transport tunnel of haloalkane dehalogenase LinB by a disulfide bridge and subsequently introduced a new tunnel into a different part of the protein by three substitutions selected using focused directed evolution (Figure 1). The de novo

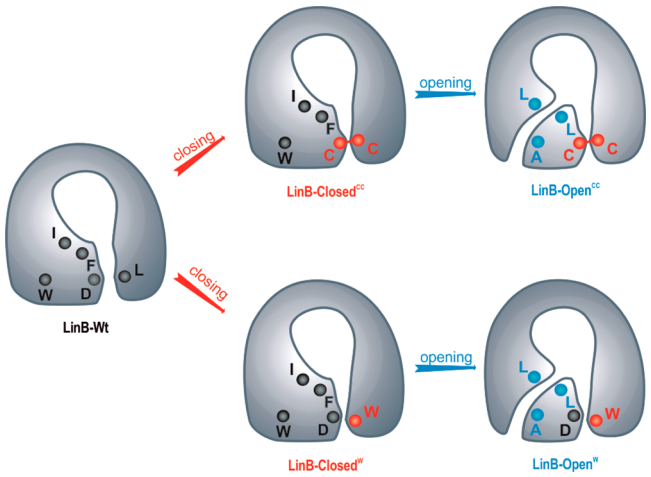


Figure 1. Strategy for engineering de novo transport tunnels. In the first step, the main access tunnel of LinB-Wt was blocked either by a disulfide bridge in LinB-Closed^{CC} or by bulky tryptophan in LinB-Closed^{W}. In the second step, a new de novo tunnel was introduced by directed evolution focused on three positions. Enzyme structures are shown in gray. Engineered residues are shown as black (wild-type), red (blocking the main tunnel), and blue (opening the de novo tunnel) spheres.

tunnel was then transplanted to a mutant with a main transport tunnel closed by bulky tryptophan to allow structural characterization of the constructs. The crystal structures of the mutants confirmed introduction of a new tunnel to the scaffold, and its functional relevance was shown by molecular dynamics simulations. The final mutant exhibited marked improvement in all chemical reaction steps, product release rates, and substrate inhibition, all contributing to its performance as the most active haloalkane dehalogenase known so far. Information gleaned from this study will help to guide the introduction of de novo tunnels to native and designed enzymes.

RESULTS

Closing the Main Tunnel by a Disulfide Bridge. Initially, we designed a mutant in which the main tunnel would be blocked by a disulfide bridge to markedly limit accessibility of the active site via this tunnel. Such a mutant would then serve as a useful "negative control" and template for further mutagenesis, owing to easier assessment of the effects of the new tunnel opening (Figure 1). Analysis of 12 crystal structures of the wild-type enzyme (LinB-Wt) using the in-house computational tool CAVER 3.0.2[27,28] revealed the presence of two tunnels, p1 and p2, with average bottleneck radii of over 1 Å. The former tunnel was further divided into two branches, p1a and p1b (Figure S1A and Table S1 in the Supporting Information). We selected the widest and shortest p1 tunnel for closure by a disulfide bridge. The use of the bridge for the blockade was motivated by the possibility of detecting its presence and in this way monitoring if the closure was maintained during further engineering of the de novo tunnel. Residue L177 divides by its side chain the p1 tunnel into two branches and was the obvious primary target for substitution to cysteine due to its importance to function of LinB-Wt shown by previous mutagenesis studies.[5,29] The remaining mutable residues in the bottleneck of p1a and p1b positioned across the tunnel from L177 were considered for introduction of the second cysteine (Table S2 and Figure S1B in the Supporting Information). Two pairs of mutations, D147C+L177C and L177C+A247C, with the shortest distance between sulfur atoms of the designed disulfide bridge were selected for experimental construction by site-directed mutagenesis.

To study the formation of the designed disulfide bridges, the mutants digested with trypsin and chymotrypsin were analyzed by MALDI-MS/MS. While no signal corresponding to formation of the designed disulfide bridge was detected for L177C+A247C mutant, peptides containing cysteines in positions 147 and 177 were observed after reduction and alkylation of the mutant D147C+L177C (Figure S2 in the Supporting Information). Conclusively, we confirmed the presence of a disulfide bridge between C147 and C177 by analysis of the pepsin digest (Figure S3 in the Supporting Information). The successful introduction of the designed disulfide bridge implied blockade of the p1 tunnel in the D147C+L177C mutant, which was denoted as LinB-Closed^{CC}.

De Novo Creation of a Transport Tunnel. The LinB-Closed^{CC} mutant was then used as a template for the de novo design of a novel transport tunnel. To identify a suitable location for introduction of the tunnel, we analyzed the ensemble of X-ray structures of LinB-Wt by CAVER 3.0.2 using a 0.6 Å probe. By using such a small probe, subatomic voids in proteins were explored for existence of continuous prospective pathways that were too small to transport even a single atom but could be enlarged to do so by mutagenesis. This analysis revealed nine potential pathways in addition to the three well-defined transport tunnels observed in the crystal structures (Figure S4A and Table S3 in the Supporting Information). The shortest of the pathways, denoted as p3, was orthogonal to the already presented p1a, p1b, and p2 tunnels and was selected as the target for de novo creation of the new transport tunnel. This pathway had two bottlenecks, which were formed by side-chain atoms of seven residues (Figure S4B–E and Table S4 in the Supporting Information).

Since there is only limited knowledge of the properties needed for an "optimal" transport tunnel, we decided to rely on a semirational approach for the creation of tunnel p3. We employed site-saturation mutagenesis of the selected bottleneck residues to open a new transport route. From the first bottleneck, closest to the active site, the two mutable residues

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597–7610

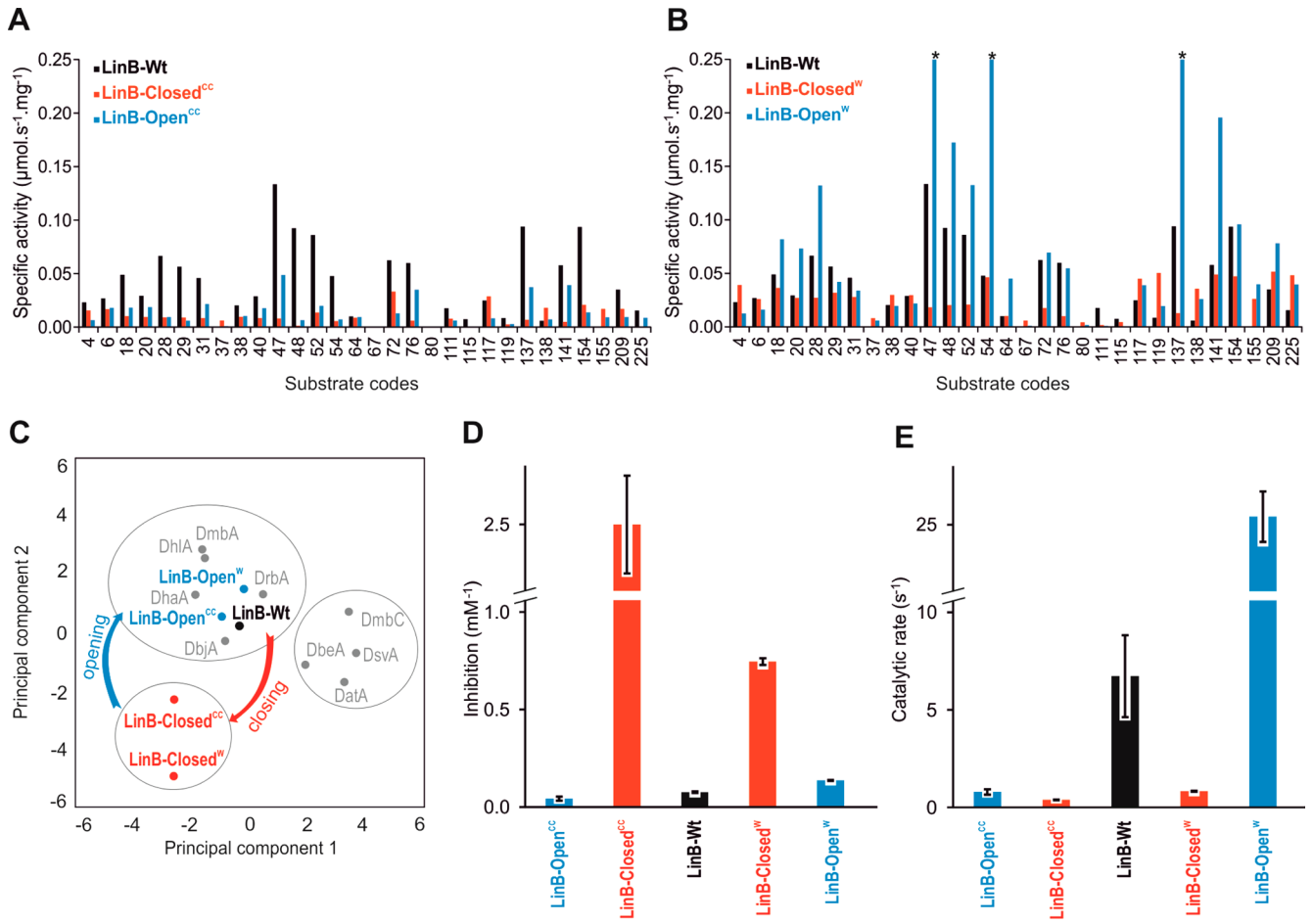


Figure 2. Substrate specificity and steady-state kinetics of enzymes with engineered transport tunnels. (A) Comparison of specific activities of LinB-Wt, LinB-Closed^{CC} with closed tunnel and its counterpart LinB-Open^{CC} with de novo opened tunnel. (B) Comparison of specific activities of LinB-Wt, LinB-Closed^{W} with closed tunnel, and its counterpart LinB-Open^{W} with de novo opened tunnel. The identities of substrates are provided in Methods and in Table S6 in the Supporting Information. Asterisks indicate activities above 0.25 μmol s^{-1} mg^{-1}. (C) Clustering of LinB variants according to their substrate specificities toward 30 halogenated compounds visualized using the score plot of the first two principal components (t_1/t_2) from the principal component analysis. Variants with closed tunnels (LinB-Closed^{CC} and LinB-Closed^{W}) were clustered in a substrate specificity group different from that of the corresponding variants with de novo opened tunnels (LinB-Open^{CC} and LinB-Open^{W}) and wild-type enzyme (LinB-Wt). Loadings plot p_1/p_2 showing the substrates that govern the clustering is shown in Figure S11 in the Supporting Information (D) Catalytic rates of enzymes measured with 1,2-dibromoethane at the highest common substrate concentration of 3.4 mM. (E) Inhibition of enzymes by substrate 1,2-dibromoethane expressed as an equilibrium constant for the formation of the inhibitory complex, $1/K_{si}$.

F143 and I211 with their side chains pointing toward p3 were selected for mutagenesis. From the second bottleneck, closest to the protein surface, the bulkiest residue W140 of this bottleneck was selected for mutagenesis. Simultaneous saturation of the selected amino acids without the need for creating an extensive library, was achieved by using the degenerate codons NBG, BNT, and NBA, designed by the CASTER program.[30] The codons enabled introduction of primarily smaller residues while allowing the original ones at given positions. Altogether, 4140 colonies were screened for their activity toward 1,2-dibromoethane, the best substrate of LinB-Wt. The genes of 36 positive candidates surpassing the activity of LinB-Closed^{CC} were sequenced, providing the 12 unique mutants M1−M12 with modified residues in the p3 pathway (Table S5 in the Supporting Information). We measured the activities of these mutants toward six substrates of different sizes and physicochemical properties to identify those with the most diverse specificities (Table S5 and Figure S5 in the Supporting Information). Three mutants, M2, M7, and M8, selected using principal component analysis, were analyzed by MALDI-MS/MS to check whether the disulfide bridges were preserved. Analysis of pepsin digests confirmed the presence of a disulfide bridge in M8 carrying mutations W140+F143L+I211L (Figures S6 and S7 in the Supporting Information). This mutant was denoted LinB-Open^{CC} and used in further experiments.

Characterization of the Closed and Open Mutants with Introduced Disulfide Bridge. Circular dichroism (CD) spectroscopy verified that both LinB-Closed^{CC} and LinB-Open^{CC} retained correct folding and thermal stability (Figure S8 in the Supporting Information). Measurements with a standard set of 30 substrates of dehalogenases[31] showed a systematic decrease in the specific activity for the majority of these substrates with LinB-Closed^{CC}, indicating successful blockade of the p1 tunnel. Similar activities among the structurally different substrates resulted in an unusually flat activity profile for this mutant (Figure 2A and Table S6 in the Supporting Information). The overall activity of LinB-Open^{CC} with these substrates was increased by ∼35% in comparison to LinB-Closed^{CC} (Figure 2A and Table S6). The flat specificity

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

Table 1. Burst Kinetics of LinB Variants with Substrate 1,2-Dibromoethane at 37 °C and pH 8.6[a]

enzyme	mutation	$k_{burst,Hal}$ (s^{-1})	$k_{burst,Alc}$ (s^{-1})	k_{ss} (s^{-1})
LinB-Wt		60 ± 30	40 ± 20	5.0 ± 0.5
LinB-Closed^{CC}	D147C+L177C	21 ± 6	17 ± 3	2.8 ± 0.2
LinB-Open^{CC}	D147C+L177C++W140A+F143L+I211L	50 ± 10	40 ± 10	1.5 ± 0.1
LinB-Closed^{W}	L177W	310 ± 70	100 ± 20	1.3 ± 0.1
LinB-Open^{W}	L177W+W140A+F143L + I211L	300 ± 100	240 ± 50	18.3 ± 0.4

[a]Definitions: $k_{burst\ Hal}$, rate constant of halide burst; $k_{burst,Alc}$, rate constant of alcohol burst; k_{ss}, rate at the steady-state phase of the studied reaction.

profile of LinB-Closed^{CC} suggests the protein forms a separate substrate specificity group aside from all native haloalkane dehalogenases (Figure 2C). The substrate specificity of LinB-Open^{CC} was comparable to that of LinB-Wt, resulting in the clustering of both proteins within the same group (Figure 2C). The steady-state kinetics determined for LinB-Closed^{CC} with 1,2-dibromoethane showed a decreased catalytic rate constant, k_{cat}, and unusually strong substrate inhibition constant, K_{si} (Figure 2D,E and Table S7 in the Supporting Information). The catalytic rate of LinB-Open^{CC} was slightly lower in comparison to LinB-Closed^{CC}, while the strong substrate inhibition was dramatically reduced beyond the level of both LinB-Closed^{CC} and LinB-Wt, making this mutant particularly efficient at higher substrate concentrations.

We observed a clear pre-steady-state burst of halide and alcohol product formation for all three enzymes upon rapid mixing with excess 1,2-dibromoethane (Table 1 and Figure S9A in the Supporting Information). The occurrence of halide and alcohol bursts during the transient phase of the reaction (k_{burst}) indicates that hydrolysis of the alkyl−enzyme intermediate and all steps before it were fast and after dissociation of formed products slowed down the reaction to the steady-state level (k_{ss}). The same kinetic pattern was observed for all three enzymes, implying that 1,2-dibromoethane conversion was limited by the release of products. The observed slowdown in this rate-limiting step in LinB-Closed^{CC} mutant indicated successful blockade of the engineered p1 tunnel. In comparison to LinB-Wt, also the rate of the halide burst was notably decreased in LinB-Closed^{CC}, suggesting that introduction of the disulfide bridge into the p1 tunnel might have reduced the rates of initial catalytic steps: substrate binding and/or cleavage of the carbon−halogen bond. We also noted the decreased rate of consequent hydrolysis of the alkyl−enzyme intermediate. In the case of LinB-Open^{CC}, the rates of the initial catalytic steps reached levels attainable with LinB-Wt (Table 1 and Figure S9A). Our attempts to obtain the crystal structures of both mutants were, however, unsuccessful.

Second-Generation Mutant with Open de Novo Tunnel. To enhance our chances of obtaining structural insights into the opening and closing of enzyme tunnels, we turned our attention to a previously published mutant (referred to as LinB-Closed^{W}) that has the p1 tunnel blocked by the bulky mutation L177W.[5] The mutations W140A+F143L +I211L responsible for opening the de novo tunnel p3 in LinB-Open^{CC} were transplanted to LinB-Closed^{W}, resulting in the mutant LinB-Open^{W}.

Both the LinB-Closed^{W} and LinB-Open^{W} mutants showed correct folding and only small deviations from the thermal stability of LinB-Wt (Figure S8 in the Supporting Information). The substrate specificity toward the 30 standard substrates followed the trends previously observed for the LinB-Closed^{CC} and LinB-Open^{CC} mutants, i.e. (i) decreased overall activity of LinB-Closed^{W} with relatively uniform distribution of activities and (ii) improved activities for the mutant with introduced p3 tunnel that even surpassed the activity of the wild type with 20 substrates (Figure 2B and Table S6 in the Supporting Information). These changes in the specificity profile resulted in clustering of LinB-Closed^{W} in the separate substrate specificity group together with LinB-Closed^{CC} but apart from other dehalogenases and both mutants with opened p3 tunnel (Figure 2C). The observed analogy also held for the steady-state kinetics: i.e., the LinB-Closed^{W} mutant showed strong substrate inhibition and lower catalytic rate. Substrate inhibition was notably reduced and the turnover number was improved for the mutant LinB-Open^{W} with the de novo opened p3 tunnel (Figure 2D,E; Table S7 in the Supporting Information).

Deeper insights into the catalytic properties of these mutants were obtained from burst experiments, which showed that the catalytic steps before and during hydrolysis were improved far beyond those of LinB-Wt (Table 1 and Figure S9B in the Supporting Information). The observed enhancements could be explained by altered solvation of the active site, reducing a competition between waters and the substrate for the nucleophile initiating nucleophilic substitution. A more restricted active site may also promote positioning of the substrate in the reactive complex.[4,32,33] The LinB-Closed^{W} mutant with altered p1 tunnel had markedly slower product release, comparable to that of the mutants carrying the disulfide bridge, whereas the LinB-Open^{W} mutant with engineered p3 tunnel showed large enhancement of the product release rates (Table 1 and Figure S9B).

Structural Analysis of the Second-Generation Mutants with Engineered Tunnels. The crystal structure of LinB-Closed^{W} (PDB ID 4WDQ) was determined to 1.6 Å resolution (Table S8 in the Supporting Information). Two crystal structures of LinB-Open^{W} mutant were solved, with resolutions of 2.5 Å (PDB ID 4WDR) and 1.3 Å (PDB ID 5LKA). Analysis of the tunnels in these structures by CAVER 3.0.2 revealed that the widths of both p1a and p1b were reduced by 0.5−0.6 Å in LinB-Closed^{W}, resulting in limited active-site accessibility of the enzyme (Table 2 and Figure 3). Importantly, when the mutations W140A+F143L+I211L were transplanted into LinB-Open^{W}, we observed widening of the radius of the p3 tunnel by 0.4 Å on average in comparison to LinB-Closed^{W} (Table 2 and Figure 3). Opening of the p3 tunnel was partially countered by backbone rearrangements, which brought the introduced L143 closer to L211 to compensate for interactions lost due to the mutagenesis (Figure S10A in the Supporting Information). With regard to the width of p3 tunnel introduced to LinB-Open^{W}, the p3 tunnel became comparable to the other three tunnels: the tunnels p1a and p1b remained less open and p2 similarly open as in LinB-Wt, providing evidence for the successful engineering of p3 tunnel geometry (Table 2 and Figure 3).

Interestingly, W177 adopted two different conformations in two chains present in the asymmetric unit of the crystal

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

Table 2. Geometric Parameters of Transport Tunnels and Potential Pathways in LinB Variants

param	tunnel/ pathway	enzyme[a] LinB-Wt	LinB-Closed^W	LinB-Open^W
bottleneck radius (Å)	p1b	1.4 ± 0.2	0.8	0.9 ± 0.2
	p1a	1.1 ± 0.2	0.6	0.9 ± 0.2
	p2	1.0 ± 0.1	1.5	1.1 ± 0.1
	p3	0.7 ± 0.1	0.6	1.0 ± 0.1
tunnel length (Å)	p1b	13 ± 1	13	16 ± 4
	p1a	16 ± 1	13	15 ± 1
	p2	19 ± 1	19	19 ± 2
	p3	16 ± 2	14	16 ± 2

[a]Data for LinB-Wt were averaged over 12 crystal structures (Table S1 in the Supporting Information). Analysis of tunnels in LinB-Closed^W was performed on a single available crystal structure (PDB ID 4WDQ). Data for LinB-Open^W were averaged over the three biological units available from the two crystal structures (PDB ID 4WDR and 5LKA; Table S9 in the Supporting Information).

structure of LinB-Open^W (PDB ID 4WDR), indicating that the mutation F143L might promote the conformational dynamics of W177 (Figure S10B in the Supporting Information). In contrast, in the high-resolution crystal structure of LinB-Open^W (PDB ID 5LKA) only one conformation of W177 was present, analogous to those found in the structure of LinB-Closed^W and the chain A of the 4WDR structure (Figure S10B), suggesting that the second conformation of W177 observed in chain B of the 4WDR structure might rather be due to a different crystal packing. There are two unique salt bridges formed by Asp149 and Arg148 from the chain B to Arg122 and Glu123 from the chain A. These additional interactions cause a shift in the position of Asp147 in chain B that is located in the vicinity of residues involved in these interactions. Since Asp147 forms the hydrogen bond to Trp177, the change in its position could be responsible for the unique conformation of Trp177 observed in the affected chain B.

Molecular Dynamics Simulations of the Second-Generation Mutants with Engineered Tunnels. The geometry of tunnels analyzed in the crystal structures indicated that there might be up to four functional tunnels in the LinB-Open^W mutant. We have performed three independent 200 ns

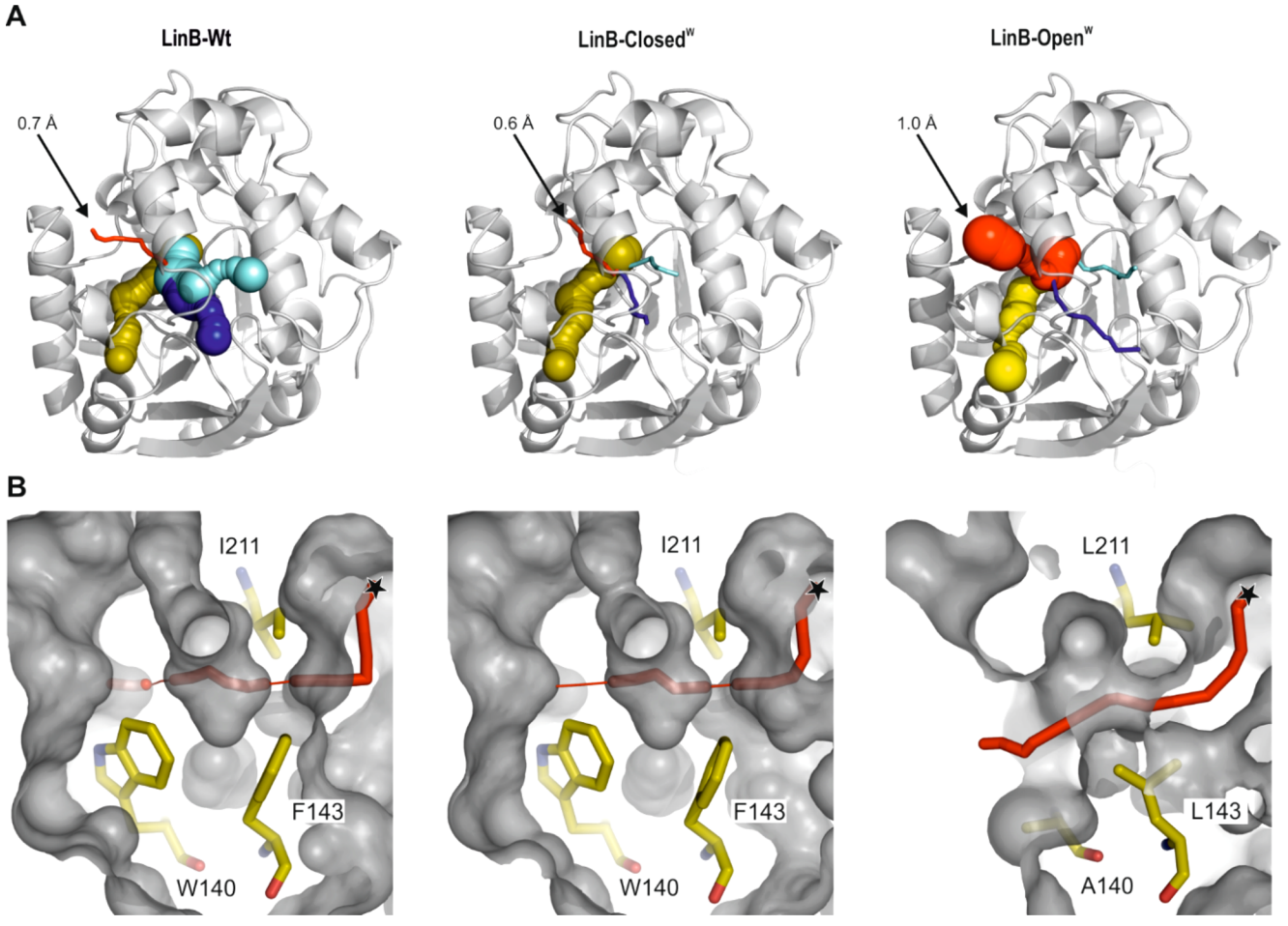


Figure 3. Crystal structures of enzymes with engineered transport tunnels. (A) Transport tunnels and potential pathways in LinB-Wt (PDB ID 1K63), LinB-Closed^W (PDB ID 4WDQ), and LinB-Open^W (PDB ID 5LKA). The p1b, p1a, p2, and p3 tunnels are shown as blue, cyan, dark yellow, and red spheres, respectively, whereas corresponding potential pathways of sub-angstrom width are shown as centerlines. Enzyme structures are shown as gray cartoons. The bottleneck radii of p3 tunnels are shown as numbers (see Table 2 for more details). (B) Detail of the geometry of p3 tunnels in LinB variants. Enzyme structures are shown as gray surfaces. Residues targeted during engineering of the p3 tunnel are shown as yellow sticks. The p3 transport tunnels or potential pathways are indicated by red centerlines with their thicknesses corresponding to the tunnel width. The black asterisk highlights the beginning of the tunnel in the active site.

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

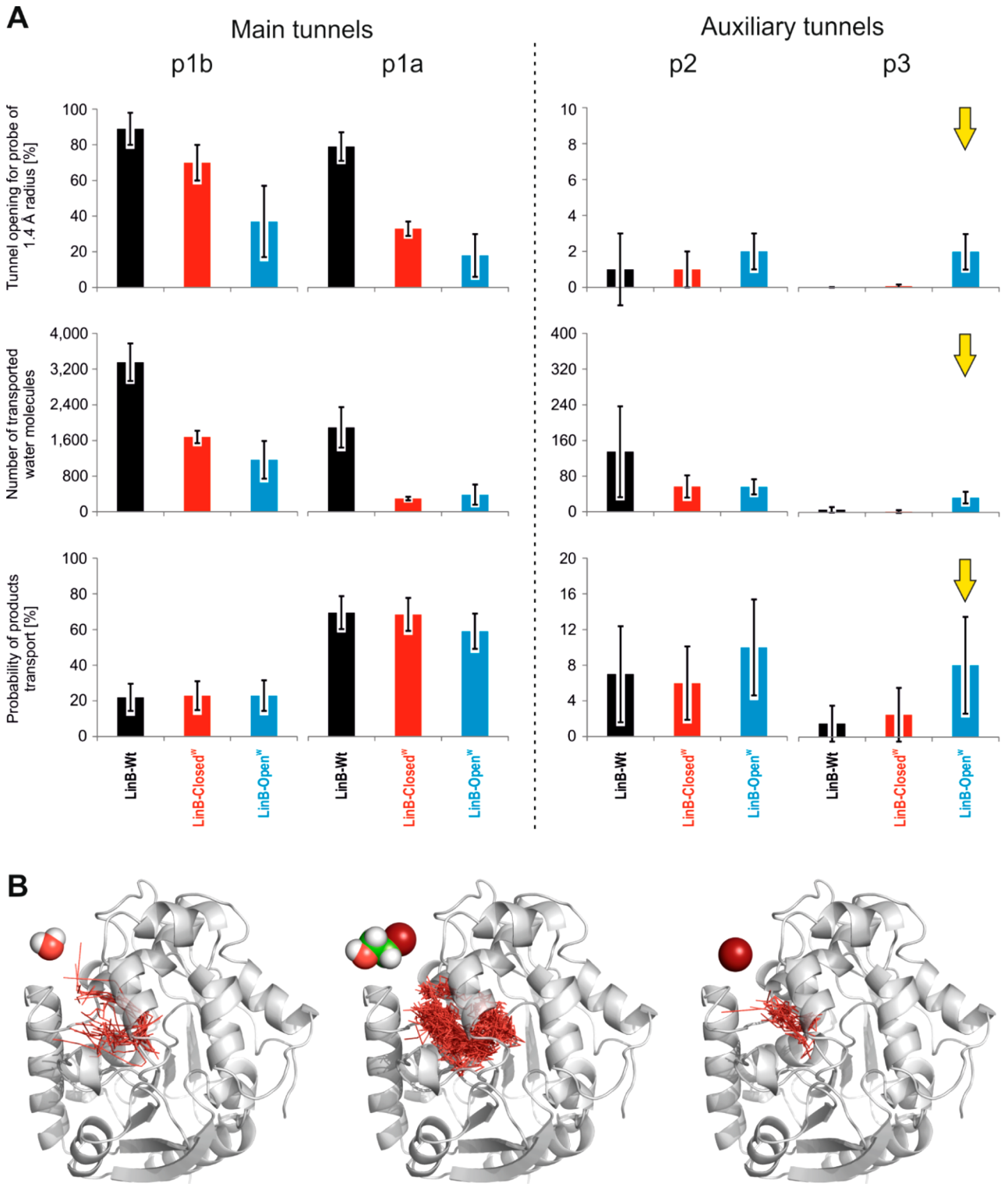


Figure 4. Functional relevance of transport tunnels. (A) Structurally and functionally relevant properties of tunnels obtained by molecular dynamics simulations. Systematic changes in the functional properties of the p3 tunnel are highlighted by the yellow arrows. Note the different scales of *Y* axis used for the main and auxiliary tunnels. (B) Utilization of p3 tunnel for transport of three important ligands (water, left; 2-bromoethanol, middle; bromide ion, right) in molecular dynamics simulations of LinB-Open^{W}. Enzyme structures are shown as gray cartoons, while traces of the ligand's center of mass during individual transport events are shown as red lines. The representative structures of ligands are shown as spheres.

long molecular dynamics (MD) simulations for LinB-Wt and LinB-Closed^{W} and two simulations for each of three starting structures of LinB-Open^{W} to discern the potential roles of these tunnels and to study the effect of mutations on tunnel dynamics. CAVER 3.0.2 was used to analyze 40000 snapshots from each simulation to identify the access tunnels, to quantify their open and close states, and to study time-resolved changes in their bottleneck radii. The p1b and p1a tunnels were clearly dominant in LinB-Wt, being open for a 1.4 Å probe in 89% and 79% of the simulation time, respectively (Figure 4A). The

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

mutation L177W in LinB-Closed^{W} resulted in average bottlenecks of p1b and p1a tunnels decreased by 0.3 and 0.2 Å (Table S10 in the Supporting Information), respectively, and we also observed a notable drop in the opening of the p1a tunnel by more than 40% (Figure 4A). Interestingly, both p1a and p1b tunnels were even more constricted in LinB-Open^{W}, as was also proposed by the high-resolution structure (PDB ID 5LKA) of this mutant (Table S9 in the Supporting Information). The p2 tunnel was open in 1% of the simulations only in all three LinB variants (Figure 4A). The p3 tunnel was not detected at all with a 0.7 Å probe during the simulations of LinB-Wt and LinB-Closed^{W}. In the case of LinB-Open^{W}, the engineered p3 tunnel was systematically open for 2% of the simulation time, confirming the opening of a new auxiliary tunnel in this mutant. The average bottleneck radius of the de novo p3 tunnel was 1.1 Å (Table S10), resulting in an average widening of at least 0.4 Å that is in excellent agreement with the observations made using the respective crystal structures.

To evaluate the functional relevance of tunnels in LinB variants, we examined their utilization for the transport of three chemical species relevant to the catalysis. First, we analyzed the transport of water molecules through these tunnels (Figure 4A and Table S11 in the Supporting Information). In all three enzymes, the majority of water molecules were transported via p1a and p1b tunnels, though the traffic through these tunnels was significantly decreased in both variants carrying the L177W mutation. The de novo engineered p3 tunnel in LinB-Open^{W} was newly used by water molecules (Figure 4), whereas no change of p2 permeability for water molecules was observed (Figure 4A and Table S11).

Finally, we focused on the egress of two products of the 1,2-dibromoethane conversion from the enzymes' active sites. Since the passage of larger molecules of products would require more pronounced opening of employed tunnels than was observed during MD simulations of ligand-free LinB variants, we performed a series of 100 independent accelerated MD simulations per starting structure. The length of the simulations varied between 4 and 200 ns to capture the entire release event of both products: i.e., bromide ion and 2-bromoethanol. These simulations confirmed the important role of both p1a and p1b tunnels for the release of products (Figure 4A and Table S12 in the Supporting Information). Interestingly, the p1a tunnel was preferentially used during the egress of both products, possibly due to the location of the p1a tunnel in the flexible cap domain and p1b at the interface of the cap domain and rigid main domain. Similarly to the previous findings, the use of the p2 tunnel for egress of products was not significantly altered in the three investigated variants. Most importantly, the de novo p3 tunnel in LinB-Open^{W} was systematically employed by both products during their releases (Figure 4), confirming its role in the exchange of ligands between the buried active site and surrounding environment. Due to variable simulation lengths, the effect of mutations could not be explained by observed transport probabilities alone.

To better understand the role of individual mutations on the proficiency of LinB variants, we analyzed the time required for both products to leave the active site through individual tunnels (Figure S12A in the Supporting Information). We observed more than 2-fold slowdown in the overall product release from Linb-Closed^{W}, which was mainly caused by slower transport through the preferred p1a tunnel blocked by L177W mutation. The transport rates were restored to the wild-type levels in LinB-Open^{W}. In this mutant, we noted improvements in product release rates via p1a and newly via p3 (Figure S12A). The penetrability of the engineered p3 tunnel reached a level comparable with that of other tunnels (p1a, p1b, and p2). A detailed analysis of 236 simulations capturing the release of at least one product via the p1a tunnel in LinB-Open^{W} revealed that the initial solvation of the products inside the active site was enabled nearly equally by p1a and p3 tunnels (Figure S12B). Solvation of products is a prerequisite for product release from the active site of dehalogenases, since the water molecules must disrupt hydrogen-bonding interactions of the bromide ion with the halide-stabilizing residues and a hydrogen-bonding interaction of 2-bromoethanol with the nucleophile.[34,35] The relevance of the p3 tunnel for the entry of water molecules to the active site cavity is underscored by a partial blockage of p1a and p1b tunnels by 2-bromoethanol product (Figure S12C).

The interference of the product molecule with solvation of the active site led us to probe the potential hindrance of rate-limiting product release by the substrate molecules located in the tunnels of LinB-Open^{W}. We performed 15 independent 200 ns long MD simulations with the presence of three substrate molecules in the solution surrounding the enzyme. We observed 16 events of the substrate entering the tunnels of LinB-Open^{W}, including also a binding of the multiple substrates (Figure S13A in the Supporting Information). Importantly, this analysis revealed that the p1a tunnel was dominantly used by the substrate molecules and that the release of products via this tunnel will in many instances suffer from the interference with substrate molecules (Figure S13B). Such notable limitation of product release via the p1a tunnel enhances the relative importance of other tunnels, including the newly introduced p3 tunnel, for the product release. This interference is especially relevant under conditions involving an excess of substrate molecules, similar to those used in our experiments.

DISCUSSION

Protein tunnels play a vital role in enzymatic function, and their modification can result in altered activity, substrate specificity, enantioselectivity, and stability.[2,17−22] Here, we demonstrate that introduction of a functional de novo tunnel into the protein structure is possible. On the basis of structural analysis and MD simulations, the p3 transport tunnel was opened in LinB-Open^{W} due to three mutations in the bottleneck residues. The new tunnel was utilized for transport of both alcohol and halide ion products of catalysis as well as water molecules, which act as a cosubstrate and facilitate product release.[35] The structural parameters of the de novo created p3 tunnel matched that of natively present auxiliary tunnel p2, which corresponds well with the fact that we have evolved only three out of seven bottleneck residues belonging to the p3 tunnel in this study. Further engineering of the p3 tunnel toward its role as the main transport tunnel might therefore require optimization of the remaining tunnel-lining residues.

What else have we found during this exercise? Introduction of the disulfide bridge to close the original main tunnel in LinB-Closed^{CC} turned out to be an efficient strategy during the initial screening for mutants with a newly opened tunnel, but it later hampered determination of the crystal structures of the variants carrying this disulfide bridge. Encountered problems could be connected with too large initial distances for creation of the covalent bond (Table S2 in the Supporting Information), possibly introducing a strain into the protein structures. A comparison of structures of LinB-Open^{W} with its template

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

LinB-Closed^{W} revealed that introducing primarily small amino acids may not be the most optimal approach for designing a tunnel opening because the protein structure tends to compensate for lost contacts by adopting more stable conformations that fill newly created empty spaces (Figure S10A in the Supporting Information). Future protocols for designing permanent tunnels will need to explicitly consider interactions among the residues surrounding the newly formed tunnels. Alternatively, designing tunnels equipped with molecular gates to allow both closed and open states[6,13−15] may represent an efficient strategy for producing active and stable enzymes.[24]

On the basis of its structural and functional parameters, the newly open p3 tunnel should be considered as an auxiliary tunnel. The dramatic improvement of enzymatic catalysis in LinB-Open^{W} could be attributed to the introduction of the p3 tunnel to the protein structure, since we did not obtain any evidence for improvements in structural or functional properties for other transport tunnels in LinB-Open^{W} from crystallographic analysis or molecular dynamic simulations. However, we cannot entirely rule out indirect effects of the mutations on enzyme catalysis via modulating enzyme dynamics. Functionally, we observed changes in the substrate specificity, kinetics, and substrate inhibition. A complete switch in the specificity profiles and even creation of mutants with substrate specificity unknown for native enzymes was achieved by the conducted engineering of tunnels. In the case of LinB-Open^{W}, this led to improvements in all of the catalytic steps, rendering this mutant the most proficient haloalkane dehalogenase ever isolated[31,36] or constructed in the laboratory.[4,24,29,37−39] In contrast, the mutant LinB-Open^{CC} showed improvements solely in the initial steps involving substrate and water access, whereas no improvement was observed for product release. Such differences in effects of adding the auxiliary tunnel into these two closed templates might be connected to the structural basis of p1 tunnel closure in LinB-Closed^{CC} and LinB-Closed^{W}. The different nature of p1 closure appeared to cause preferences for these ligands to shift to various degrees, which seems to correspond to the fact that the geometry and physicochemical properties of tunnels define their selectivity toward individual ligands: substrates, products, and solvent molecules. Moreover, the multiple tunnels in the investigated enzymes seemed to be interconnected. The auxiliary tunnels can promote the catalysis by efficiently solvating the released products (Figure S12B in the Supporting Information). They may also allow a simultaneous transport of substrates and products, avoiding the need for synchronization (Figure S13B in the Supporting Information). Such complexity could be responsible for mixed effects observed upon blockage and opening of individual tunnels. Interestingly, we observed clear trends in substrate inhibition, which was enhanced in both closed mutants and markedly reduced in both mutants with the de novo open tunnel. As similar effects traceable to tunnel modification have also been observed previously,[22,24,40] the re-engineering of protein tunnels could represent a novel avenue for the removal of substrate or product inhibition for biocatalysts in industrial biotechnology or synthetic biology.[41]

CONCLUSIONS

Our study demonstrates that, whereas we can open and close the transport tunnels with reasonable proficiency, we cannot accurately predict the effects of such changes on the catalytic properties at this stage. This is because of the complex effects of multiple tunnels on the selectivity toward different ligands. The way forward would be to quantitatively evaluate the effect of changes in physicochemical properties of engineered tunnels on the transport of substrates, products, and water solvent and consequently individual steps of the catalytic cycle. Our current observations suggest that one way to increase the efficiency of de novo designed catalysts is to direct their substrates and products into spatially distinct tunnels. Despite its relative infancy, we believe that the proposed strategy for opening transport tunnels could be useful for protein engineering of biocatalysts and, more specifically, offers an extension to available protocols for current de novo enzyme design focusing primarily on the optimization of active sites toward transition state stabilization.

METHODS

Site-Directed Mutagenesis and Gene Synthesis. The mutant recombinant genes *linB*(D147C+L177C) and *linB*(L177C+A247C) were constructed by inverse PCR using *Pfu* DNA polymerase (Finnzymes, Espoo, Finland) with the following oligonucleotides: *linB*(D147C)-f, 5′-GGA TTT TCC CGA ACA GTG TCG CGA TCT GTT TC-3′; linB(D147C)-r, 5′-GCC CAT TCG ATC GGC ATG GCG ATC GC-3′; linB(A247C)-f, 5′-CCG AGC CGG GAT GCC TGA CCA CGG GCC G-3′; linB(D147C)-r, 5′-CGT TGA TGA AGA GTT TCG GAA TCG GGC-3′. Plasmid pAQN-*linB*27 (L177C)[29] was used as a template for PCR. The resulting PCR product was treated with *Dpn*I restriction endonuclease to remove pAQN-*linB* template plasmid. The product was treated with T4 polynucleotide kinase (New England Biolabs, Ipswich, CT, USA), self-ligated, and transformed into *Escherichia coli* DH5*α*. Plasmids pAQN-*linB*(D147C+L177C) and pAQN-*linB*(L177C+A247C) from *E. coli* DH5*α* were isolated using a GeneJET plasmid miniprep kit (Fermentas, Burlington, VT, USA), and the presence of a *linB* gene variant in the plasmid was verified by double-digestion with *Eco*RI and *Hin*dIII. The nucleotide sequence of the mutant was confirmed by the dideoxy-chain termination sequencing method with an automated DNA sequencer ABI PRISM 310 genetic analyzer (Applied Biosystems, Foster City, CA, USA). The variant LinB-Closed^{W} was constructed by inverse PCR with the oligonucleotides 5′-GCG CAG GAT GTG TCC GGG GAG-3′ and 5′-AAC TTG TTC GAC AAA AAC-3′, as described previously.[29] The gene of LinB-Open^{W} was synthesized artificially (Entelechon, Regensburg, Germany) according to the sequence of LinB-Closed^{W}.

Construction of Library by Saturation Mutagenesis. Saturation mutagenesis was performed using a QuikChange Lightning multisite-directed mutagenesis kit (Stratagene, La Jolla, CA, USA) in all three selected positions simultaneously. The proximity of the amino acid residues W140 and F143 enabled degenerate codons mutating these positions to be included in one oligonucleotide. The following oligonucleotides with degenerated codons were used for saturation mutagenesis: linB(W140+F143)-sm, 5′-C GCC ATG CCG ATC GAA NBG GCG GAT BNT CCC GAA CAG TGT CGC G-3′; linB(I211)-sm, 5′-C CTG TCT TGG CCT CGC CAA NBA CCG ATC GCA GGC ACC-3′ (where N and B mean any nucleotide and any nucleotide except for A, respectively). All three positions were mutated simultaneously in one PCR reaction conducted according to the manufacturer's protocol. The plasmid pAQN-*linB*(D147C+L177C) was used as a template for the PCR reaction. The volume of the PCR

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

reaction mixture was 25 μL. The reaction mixture contained 2.5 μL of 10X QuikChange Lightning multi reaction buffer, 50 ng of template DNA, 100 ng of each primer, 1 μL of dNTP mix, and 1 μL of QuikChange Lightning multi enzyme blend. The PCR proceeded under the following conditions: 95 °C 2 min and 30 × (95 °C 20 s, 58 °C 30 s, 65 °C 2 min), with a final extension step at 65 °C for 5 min. The PCR products were then treated with methylation-dependent endonuclease *Dpn*I for 5 min at 37 °C. The generated plasmids were transformed into *E. coli* XJb(DE3) cells (ZymoResearch, Orange, CA, USA) using the standard electroporation protocol. The transformation mixture was placed in a Micropulser cuvette with a 0.1 cm gap (BTX, Holliston, MA, USA). A pulse of 1.8 kV was delivered by an ECM 399 Generator (BTX, Holliston, MA, USA). Candidates selected on the basis of screening on microtiter plates (MTPs) were used to propagate plasmid in 10 mL of LB medium containing ampicillin (100 $\mu g.ml^{-1}$), and their plasmid DNA was isolated by a GeneJET plasmid miniprep kit (Fermentas, Burlington, Canada). Plasmids were commercially sequenced together with selected mutants using the following pAQN vector specific primers: pAQN-f, 5′-GTT GAC AAT TAA TCA TCG GCT CG-3′; pAQN-r, 5′-GCC GCC AGG CAA ATT CTG-3′ (Macrogen, Amsterdam, The Netherlands).

Screening of Library Constructed by Site-Saturation Mutagenesis. Sterile MTP wells were filled with 150 μL of Luria−Bertani (LB) medium containing ampicillin at a final concentration of 100 $\mu g\ mL^{-1}$. Wells were inoculated with a single transformant using sterile toothpicks. Cultures were grown overnight at 37 °C. After 14 h of cultivation, 50 μL of each culture from the cultivation plate was added to 50 μL of 30% glycerol to generate a replica plate for storage. Next, 50 μL of fresh LB medium with ampicillin, L-arabinose at a final concentration of 3 mM, and isopropyl-β-D-thiogalactoside (IPTG) at a final concentration of 0.5 mM were added to each well of the cultivation plate. The MTPs were cultivated at 30 °C for 4 h. Cell cultures were harvested by centrifugation at 1600*g* for 20 min (Laborzentrifugen, Osterode am Harz, Germany), the supernatant was discarded, and the MTPs were frozen at −80 °C. *E. coli* XJb containing plasmid pAQN-*linB-Wt* was used as a positive control, and *E. coli* XJb containing plasmid pAQN-*linB*(D147C+L177C) was used as a negative control. Prior to use, the MTPs were defrosted and kept at room temperature for 10 min. Next, 50 μL of lysis buffer (1 mM HEPES, 20 mM Na_2SO_4, and 1 mM EDTA, pH 8.2) was added to each well. Cell debris was removed from the lysate by centrifugation at 1600*g* for 20 min (Laborzentrifugen, Osterode am Harz, Germany) after incubation at room temperature at 100 rpm for 1 h. The method described by Holloway et al.[42] was carried out with modifications. The assay relied on detection of protons produced during the dehalogenation reaction in addition to halides and alcohols. A 25 μL portion of lysate was transferred to each well of a new MTP, and 175 μL of assay buffer (1 mM HEPES, 20 mM Na_2SO_4, and 1 mM EDTA, pH 8.2) containing 1,2-dibromoethane was added. The substrate was incubated in the reaction buffer at 37 °C for 30 min before starting the reaction. The MTP plate was securely closed by the lid and Parafilm. The reaction mixture was diluted by buffer containing pH indicator phenol red (1 mM HEPES, 20 mM Na_2SO_4, and 1 mM EDTA, 50 $\mu g\ mL^{-1}$ phenol red, pH 8.2) for detection after 14 h of dehalogenation. The change in color of the pH indicator was estimated at 540 nm using a Sunrise spectrophotometer (Tecan, Mannedorf, Switzerland) as previously described by Holloway.

Overexpression and Purification. To overproduce LinB variants in *E. coli*, the corresponding genes were subcloned into the expression vectors pAQN (*linB-Closed*^{cc} and *linB-Closed*^{w}) and pET21b (*linBM1-M12*, *linB-Open*^{cc}, and *linB-Open*^{w}) under the control of the T7lac promoter (Novagen, Madison, WI, USA) and gene expression was induced by addition of IPTG. *E. coli* BL21 and *E. coli* BL21(DE3) cells containing the pAQN and pET21b plasmids, respectively, were cultured in 4 L of LB medium at 37 °C. When the culture reached an optical density of 0.5 at a wavelength of 600 nm, enzyme expression (at 20 °C) was induced by addition of IPTG to a final concentration of 0.5 mM. The cells were harvested, disrupted by sonication using an Ultrasonic UP200S processor (Hielscher, Teltow, Germany), and centrifuged for 1 h at 4 °C and 21000*g*. Enzymes were purified by metallo-affinity chromatography using a Ni-NTA Superflow column (Qiagen, Hilden, Germany) on Äkta FPLC (GE Healthcare, Uppsala, Sweden). His-tagged enzymes were bound to the resin in equilibrating buffer (20 mM potassium phosphate buffer, pH 7.5, containing 0.5 M sodium chloride, and 10 mM imidazole). Unbound and weakly bound fractions were washed out, and His-tagged enzymes were consequently eluted with purification buffer containing 300 mM imidazole. The purified proteins were pooled and dialyzed against 50 mM potassium phosphate buffer (pH 7.5) overnight at 4 °C. The purity of the purified proteins was checked by SDS-polyacrylamide gel electrophoresis in 15% polyacrylamide gels.

MALDI-MS/MS Analysis. Phosphate buffer used for purification of LinB variants with the designed disulfide bridge (LinB-D147C+L177C, L177C+A247C, M2, M7, and M8) was exchanged with glycine buffer (20 mM) using a HiTrap desalting column (Amersham Biosciences, Freiburg, Germany). Samples of LinB variants were independently digested with protease trypsin, chymotrypsin, and pepsin and analyzed by MALDI-MS/MS. The digested samples were treated with dithiothreitol (DTT) and iodacetamide. The modified samples were analyzed by MALDI-MS/MS. Mass spectra of the proteins were recorded on an Ultraflex III spectrometer (Bruker Daltonik, Bremen, Germany) and processed with XMASS 5.1.5 software (Bruker, Bremen, Germany).

Circular Dichroism Spectroscopy and Thermal Denaturation. To access the secondary structure and correct folding of LinB variants, circular dichroism (CD) spectra were recorded at room temperature using a Chirascan CD spectrometer (Applied Photophysics, Leatherhead, U.K.) equipped with a Peltier thermostat. Data were collected from 185 to 260 nm at a scan rate of 100 nm/min with a 1 s response time and 2 nm bandwidth using a 0.1 cm quartz cuvette containing the enzyme in 50 mM potassium phosphate buffer (pH 7.5). Each spectrum shown represents an average of five individual scans and has been corrected for absorbance caused by the buffer. CD data were expressed in terms of the mean residue ellipticity. Thermal unfolding of the LinB variants was followed by monitoring the ellipticity at 221 nm over the temperature range 20−80 °C with a resolution of 0.1 °C and heating rate of 1 °C/min. Recorded thermal denaturation curves were roughly normalized to represent signal changes between approximately 1 and 0 and fitted to sigmoidal curves using software Origin 8.0 (OriginLab, Northampton, MA, USA). Melting temperatures (T_m) were evaluated as the midpoint of the normalized thermal transition.

Specific Activity Measurements. The specific activities of LinB variants were assayed by a colorimetric method.[43] The specific activities of LinB-Closed^{CC}, LinB-Open^{CC}, LinB-

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

Closed^{W}, and LinB-Open^{W} were measured with a set of 30 halogenated substrates (**4**, 1-chlorobutane; **6**, 1-chlorohexane; **18**, 1-bromobutane; **20**, 1-bromohexane; **28**, 1-iodopropane; **29**, 1-iodobutane; **31**, 1-iodohexane; **37**, 1,2-dichloroethane; **38**, 1,3-dichloropropane; **40**, 1,5-dichloropentane; **47**, 1,2-dibromoethane; **48**, 1,3-dibromopropane; **52**, 1-bromo-3-chloropropane; **54**, 1,3-diiodopropane; **64**, 2-iodobutane; **67**, 1,2-dichloropropane; **72**, 1,2-dibromopropane; **76**, 2-bromo-1-chloropropane; **80**, 1,2,3-trichloropropane; **111**, bis(2-chloroethyl) ether; **115**, chlorocyclohexane; **117**, bromocyclohexane; **119**, 1-(bromomethyl)cyclohexane; **137**, 1-bromo-2-chloroethane; **138**, chlorocyclopentane; **141**, 4-bromobutyronitrile; **154**, 1,2,3-tribromopropane; **155**, 1,2-dibromo-3-chloropropane; **209**, 3-chloro-2-methylpropene; **225**, 2,3-dichloropropene), whereas the activities of LinB M1-M12 variants were measured toward a reduced set of 6 selected halogenated substrates (**18**, **20**, **29**, **47**, **48**, and **117**). The dehalogenation reaction was performed at 37 °C in 25 mL Reacti-flasks closed by Mininert valves. The reaction mixture contained 10 mL of glycine buffer (100 mM, pH 8.6) and 10 μL of halogenated substrate. The reaction was initiated by addition of the enzyme. The reaction was monitored by withdrawing 1 mL samples at periodic intervals from the reaction mixture and immediately mixing the samples with 0.1 mL of 35% nitric acid to terminate the reaction. Halide ions released during the reaction were detected spectrophotometrically at 460 nm with mercuric thiocyanate and ferric ammonium sulfate using a SUNRISE microplate reader (Tecan, Grödig/Salzburg, Switzerland). The dehalogenation activity was quantified as the rate of product formation with time.

Principal Component Analysis. A matrix containing the activity data for 10 wild-type HLDs and 4 LinB variants (LinB-Closed^{CC}, LinB-Open^{CC}, LinB-Closed^{W}, and LinB-Open^{W}) with 30 substrates was analyzed by principal component analysis (PCA)[44] using Statistica 10.0 (StatSoft, Tulsa, OK, USA) in an attempt to uncover relationships between individual HLDs on the basis of their activities toward the set of substrates following the methodology described elswhere.[31] In short, the raw data were log-transformed and weighted relative to the activity of the individual enzyme toward other substrates prior to performing the PCA in order to better discern the enzyme specificity profiles. These transformed data were used to identify substrate-specificity groups of enzymes exhibiting similar specificity profiles regardless of their overall specific activities. Here, such analysis provided two types of outcomes: (i) a score plot showing the projection of enzymes onto a lower-dimensional subspace formed by selected principal components according to their substrate specificity and (ii) a loadings plot quantifying the contributions of the activities with individual substrates to a particular principal component. PCA analysis of activity data of LinB mutants M1−M12 with six selected halogenated substrates was carried out by the same procedure as described above in order to identify LinB variants with the most diverse specificities.

Steady-State Kinetic Measurements. The catalytic properties of the enzymes were described by steady-state kinetic parameters determined with 1,2-dibromoethane as a substrate. The steady-state kinetics of LinB variants were measured using a VP-ITC isothermal titration microcalorimeter (MicroCal, Piscataway, NJ, USA) at 37 °C. The microcalorimeter reaction mixture vessel was filled with 1.4 mL of enzyme solution at a concentration of 0.004−0.01 mg/mL (100 mM glycine buffer, pH 8.6). The substrate solution was prepared in the same buffer by addition of 1,2-dibromoethane to a final concentration of 4 mM. The substrate concentration was verified by gas chromatography (Finnigen, San Jose, CA, USA). The enzyme was titrated at 150 s intervals in the reaction mixture vessel with increasing amounts of substrate while maintaining pseudo-first-order conditions. Each injection increased the substrate concentration, thereby increasing the enzyme reaction rate (increased heat generated), until the enzyme became saturated. A total of 28 injections were carried out during the titration. The reaction rates reached after every injection (in units of thermal power) were converted to enzyme turnover. The calculated enzyme turnover plotted against the actual concentration of the substrate after every injection was then fitted by nonlinear regression to kinetic models using Origin 8.0 (OriginLab, Northampton, MA, USA).

Pre-Steady-State Burst Analysis. Rapid quench flow experiments were performed at 37 °C in a glycine buffer at pH 8.6 using a QFM 400 rapid quench flow instrument (BioLogic, Grenoble, France). The reaction was started by rapid mixing of 75 μL of enzyme with 75 μL of substrate solution and quenched with 100 μL 0.8 M H_2SO_4 after time intervals ranging from 5 ms to 1.6 s. The quenched mixture was directly injected into 0.5 mL of ice-cold diethyl ether containing 1,2-dichloroethane as internal standard. After extraction, the diethyl ether layer containing noncovalently bound substrate and alcohol product was collected, dried on a short column containing anhydrous Na_2SO_4, and analyzed on an Agilent 7890 gas chromatograph (Agilent, Santa Clara, CA, USA) equipped with a DB-FFAP capillary column (30 m × 0.25 mm × 0.25 μm, Phenomenex) and connected to an Agilent 5975C mass spectrometer (Agilent, Santa Clara, CA, USA). The amount of halide in the water phase was measured using an 861 Advanced Compact ion chromatograph equipped with METROSEP A Supp 5 column (Metrohm, Herisau, Switzerland). The kinetic data were fitted to the following burst equation using the software Origin 6.1 (OriginLab, Northampton, MA, USA):

$$\frac{[P]}{[E]_0} = A_0(1 - e^{-k_{obs}t}) + k_{ss}t$$

Crystallization and Data Collection. Crystals of LinB-Closed^{W} (PDB ID 4WDQ) were prepared as described previously.[45] Briefly, 1 μL of protein at a concentration of 5 mg mL^{-1} was mixed with 1 μL of precipitant (0.1 M MES, pH 5.6, 0.2 M $MgCl_2$, 25% (w/v) PEG 3350) and equilibrated against 800 μL of precipitant solution in 24-well CombiClover plates (Molecular Dimensions Limited, Newmarket, U.K.). Crystals of LinB-Closed^{W} were grown within 4 days at room temperature and then frozen in liquid nitrogen without additional cryoprotection. Diffraction data were collected at 100 K at the EMBL Hamburg beamline X13 (DESY, Hamburg, Germany) using the MAR CCD 165 mm detector at a wavelength of 0.812 Å. The data were processed with HKL3000.[46]

LinB-Open^{W} at a concentration of 11 mg mL^{-1} was used to set up crystallization trials. The crystals of LinB-Open^{W} were formed after initial crystallization screening by the sitting drop vapor diffusion method in MRC 96-well plates (Hampton Research, Aliso Viejo, CA, USA) using a Gryphon crystallization robot (ARI, Sunnyvale, CA, USA). Trigonal crystals were grown at 277 K in a Morpheus crystallization screen (Molecular Dimensions Limited, Newmarket, U.K.) in a

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

mixture containing 50 mM MOPS buffer, 50 mM HEPES, pH 7.5, 0.3 M $CaCl_2$, 0.3 M $MgCl_2$, 13% (w/v) PEG550MME, and 7% (w/v) PEG 20000. Rodlike crystals belonging to an orthorhombic space group were grown at 277 K in PEGs Suite crystallization screen (Qiagen, Valencia, CA, USA) in a mixture containing 0.2 M sodium thiocyanate and 20% (w/v) PEG 3350. All crystals were frozen in liquid nitrogen without additional cryoprotection. Diffraction data for LinB-Open^W (PDB ID 4WDR) were collected at the ESRF ID-29 beamline (Grenoble, France[47]) equipped with a Pilatus 6 M pixel detector at a wavelength of 0.9724 Å. Diffraction data for LinB-Open^W (PDB ID 5LKA) were collected at the beamline 14.1 equipped with Pilatus 6 M pixel detector at a wavelength of 0.9184 Å at the BESSY II electron-storage ring (Helmholtz-Zentrum Berlin, Berlin-Adlershof, Germany[48]). The XDSAPP graphical user interface[49] for running *XDS*[50] was used for indexing, integration, and scaling of the diffraction data.

Structure Determination, Refinement, and Analysis. Crystal structures of LinB-Closed^W (PDB ID 4WDQ) and LinB-Open^W (PDB ID 4WDR) were solved by molecular replacement with the MOLREP program.[51] The haloalkane dehalogenase LinB-Wt (PDB ID 1CV2[52]) and refined structure of LinB-Closed^W were used as search models for LinB-Closed^W and LinB-Open^W, respectively. Model refinement was carried out using REFMAC 5.2,[53] part of the CCP4 software package,[54] alternated with cycles of model building using Coot.[55] After eight cycles of isotropic restrained refinement, each chain of LinB-Open^W (PDB ID 4WDR) was divided into four TLS (translation/libration/screw) groups suggested by the TLS motion determination server[56] and another eight cycles of TLS restrained refinement were performed.[57] The crystal structure of LinB-Open^W (PDB ID 5LKA) was solved with Phaser[58] incorporated in the PHENIX suite[59] using the refined structure of LinB-Open^W (PDB ID 4WDR) as a search model. LinB-Open^W (PDB ID 5LKA) was refined in PHENIX[60] with manual model building in Coot. The MolProbity online server[61] was used to analyze the structures. The data collection, processing, and refinement statistics for LinB-Closed^W and LinB-Open^W are summarized in Table S8 in the Supporting Information.

Tunnel Analysis in Crystal Structures. Twelve available crystal structures of LinB-Wt were obtained from the RSCB PDB database to analyze their tunnels. Newly obtained structures were used for LinB-Closed^W (PDB ID 4WDQ) and LinB-Open^W (PDB ID 4WDR and 5LKA). All ligands, water molecules, and protein atoms at alternative conformations were removed, and all structures were aligned to the 1MJ5 structure using PyMOL 1.7.[62] Tunnels were analyzed by the standalone version of CAVER 3.0.2.[28] Every atom was approximated by 13 spheres with radii corresponding to the smallest atom present in the analyzed protein structure. The starting point was specified with the following coordinates: (14.924; 35.851; 6.326) Å. To prevent collision of the starting point with other protein atoms, an automatic optimization of its coordinates was performed. Transport tunnels were identified using a probe radius of 1.0 Å, whereas potential pathways were identified using a probe radius of 0.6 Å. Redundant tunnels were automatically removed from each structure, and tunnels were clustered using the threshold of 4.5.

Molecular Dynamics Simulations of Free Enzymes. Crystal structures of LinB-Wt (PDB ID 1MJ5), and its mutants LinB-Closed^W (PDB ID 4WDQ) and LinB-Open^W (PDB ID 4WDR and 5LKA) were prepared for molecular dynamics (MD) simulations by removing all ligands and water molecules. In the case of LinB-Open^W we obtained three initial monomeric structures (two chains from 4WDR and one from 5LKA). Hydrogen atoms were then added to all protein structures using H++ server at pH 8.5 and a salinity of 0.1 M with internal and external dielectric constants of 10 and 80, respectively.[63] Initial water molecules were placed into the protein structures using 3D-RISM theory according to the Placevent algorithm.[64,65] Cl^- and Na^+ ions were then added to the final concentration of 0.1 M using the Tleap module of AMBERTools15.[66] Using the same module, a truncated octahedron of TIP3P water molecules[67] was added to the distance of 10 Å from any atom in all of the systems.

Energy minimization and MD simulations were carried out in the PMEMD.CUDA module[68,69] of AMBER14[66] using the ff14SB force field.[70−72] The investigated systems were minimized by 500 steps of steepest descent followed by 500 steps of conjugate gradient in five rounds of decreasing harmonic restraints. The restraints were applied as follows: 500 kcal mol^{-1} $Å^{-2}$ on all heavy atoms of protein and then 500, 125, 25, and 0 kcal mol^{-1} $Å^{-2}$ on backbone atoms only. The subsequent MD simulations employed periodic boundary conditions, the particle mesh Ewald method for treatment of electrostatics interactions beyond 10 Å cutoff,[73,74] and 2 ft time step with the SHAKE algorithm to fix all bonds containing the hydrogens.[75] Equilibration simulations consisted of two steps: (i) 2 ns of gradual heating from 0 to 310 K under a constant volume, using a Langevin thermostat[76] with collision frequency of 1.0 ps^{-1}, with harmonic restraints of 5.0 kcal mol^{-1} $Å^{-2}$ on the positions of all protein atoms; (ii) 22 ns of unrestrained simulation at 310 K using the Langevin thermostat, and the constant pressure of 1.0 bar using the pressure coupling constant of 1.0 ps. Finally, production MD simulations were run for 200 ns at 310 K using the weak-coupling thermostat,[77] and a constant pressure of 1.0 bar using a pressure coupling constant of 1.0 ps. In total, three separate MD simulations were performed for LinB-Wt and LinB-Closed^W, whereas two simulations were run for each of the three available starting structures of LinB-Open^W. Coordinates were saved in 5 ps intervals, and the trajectories were analyzed using the Cpptraj module[78] of AMBERTools15, Pymol 1.7, and VMD 1.9.1.[79]

The dynamic behavior of tunnel network was analyzed by CAVER 3.0.2. Every atom of the protein structure was approximated by 13 spheres with radii corresponding to the smallest atom present in the analyzed protein structure. The starting point was specified by the three residues Trp109, Asn38, and His272. To prevent collision of the starting point with other protein atoms, an automatic optimization of its coordinates was performed. The tunnels were searched in 40000 snapshots from each MD simulation using a probe radius of 0.7 Å and the default settings. The redundant tunnels were automatically removed from each snapshot. The clustering of 100000 randomly sampled tunnels was performed with Murtagh's algorithm using clustering thresholds of 3 and 4.5 for LinB-Wt and its mutants, respectively. The remaining tunnels were assigned to the 20 top-ranking tunnel clusters by using a *k*-nearest-neighbor classifier.

The transport of water molecules was investigated by first considering all waters farther than 16 Å from the CG atom of Asp108 as being outside a protein, whereas waters closer than 6 Å were inside the cavity. Next, only the events involving complete transport of a water molecule from one region to the other during the analyzed MD simulation were considered.

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

Finally, the identity of the molecular tunnel employed during such transport event was disclosed by calculating the shortest average distance between all points of the ligand trajectory during such an event and the closest points of all tunnels calculated by CAVER 3.0.2.

Molecular Dynamics Simulations of Substrate Entry and Product Releases. The structures of 2-bromoethanol product and 1,2-dibromoethane substrate were created with the Avogadro 1.0.3 program.[80] The geometry was energy minimized at the MP2/6-31G* level of theory using the Gaussian09 program, revision E.01.[81] The partial atomic charges of both molecules were obtained using RESP ESP charge derive (R.E.D.) server 2.0[82,83] at the HF/6-31G* level of theory using the Gaussian09 program utilized in the R.E.D. server. The charges on 2-bromoethanol and 1,2-dibromoethane were derived employing the RESP-A1A charge model, using a single-conformation multiorientation RESP fit. The atom types were derived in analogy with the force field of Cornell et al.[84]

Three substrate molecules were placed randomly around the protein beyond the solvent shell introduced by the Placevent algorithm. In total, five different initial positions of substrates were used for each of the three available starting structures of LinB-Open^{W}. Initial positions of products in the active site of LinB variants were obtained by molecular docking performed with AutoDock Vina 1.1.2 using an exhaustiveness of 50, a maximum of 20 generated binding modes, and maximum energy difference between the best and the worst binding modes of 1.5 kcal mol^{-1}.[85] Bromide ions were placed at positions corresponding to the coordinates of chloride ions bound between the halide-stabilizing residues Asn38 and Trp109 in the crystal structure of LinB-Wt (PDB ID 1MJ5). Subsequently, the region of the active site selected for molecular docking was set to 22.5 × 22.5 × 22.5 Å centered at the bromide product. AutoDock atom types were assigned to 2-bromoethanol product and protein structures. The input files were converted into an AutoDock compliant format with the AutoDockTools4 module of MGLTools 1.5.6.[86] For each system, four binding modes of 2-bromoethanol were found to form a hydrogen bond to the catalytic nucleophile Asp108. All of these modes were used as the input structures for following MD simulations.

For simulations of substrate entry, the preparation, minimization, and equilibration of the system during MD simulations employed the same protocol as described above for the simulations of free enzymes. The simulations of product release followed the same protocol with three exceptions: (i) the initial restraints of 500 kcal mol^{-1} Å^{-2} were applied also to both products; (ii) restraints of 5.0 kcal mol^{-1} Å^{-2} were applied to both products during the 2 ns long heating phase under the constant volume; (iii) three additional distance-based restraint potentials were applied to maintain two hydrogen bonds between the two halide stabilizing residues and the bound bromide ion (both of 0.5 kcal mol^{-1} Å^{-2}) and a hydrogen bond between Asp108 and the 2-bromoethanol (1 kcal mol^{-1} Å^{-2}) during 22 ns long equilibration MD simulations under constant pressure. These restraints were one-sided to enable shortening of the hydrogen bond distances without any penalty, and they were centered at the actual hydrogen bond distances obtained from the molecular docking calculations.

By the protocol described for simulations of free enzymes, production MD simulations of substrate entry were run for 200 ns for each of the three available starting structures of LinB-Open^{W}, reaching a total simulation time of 3 μs. In production simulations of slow product release, the dual-boosting approach was employed using separate torsional and total boost potentials to accelerate both internal and diffusive degrees of freedom of the product release simulations.[87] Parameters defining the dihedral (E^{dih}; α^{dih}) and total (E^{tot}; α^{tot}) boosts were calculated on the basis of a protocol employed by Pierce et al.,[88] which was modified as follows. Average dihedral energies (E_0^{dih}) and total potential energies (E_0^{tot}) of unbiased systems were obtained from the last 20 ns of the preceding equilibration MD simulations of respective systems. Using the information on the overall number of atoms (N^{atm}) and number of protein residues (N^{res}) in the systems, the boost parameters were calculated using the following relationships: $\alpha^{dih} = 0.2C^{res}N^{res}$; $E^{dih} = E_0^{dih} + C^{res}N^{res}$; $\alpha^{tot} = C^{atm}N^{atm}$; $E^{tot} = E_0^{tot} + \alpha^{tot}$, where C^{res} and C^{atm} are empirical constants of 3.5 kcal mol^{-1} residue^{-1} and of 0.2 kcal mol^{-1} atom^{-1}, respectively. The accelerated MD simulations were performed at 310 K using the weak-coupling thermostat[77] and a constant pressure of 1.0 bar using a pressure coupling constant of 1.0 ps. In total, 25 separate accelerated MD simulations were performed for each of four initial binding modes, providing in total 100 simulations of LinB-Wt and LinB-Closed^{W} and 300 simulations of LinB-Open^{W}. These simulations were run for variable times from 4 to 200 ns until both products left the enzyme, which was monitored by using a distance cutoff of 18 Å between the starting point of tunnels defined as the centers of mass of Trp109, Asn38, and His272 residues and centers of mass of each product. When no release of both products was observed, a simulation was discarded, providing the final sets of 100 simulations for LinB-Wt, 99 simulations for LinB-Closed^{W}, and 285 simulations for LinB-Open^{W}. Analyses of successful simulations were performed with the same tools as employed for the MD simulations of free enzymes. The uncertainties in the probability of product transport through a particular tunnel were estimated using a bootstrap method,[89] 10000 times resampling the data from a random subset of 50 simulations.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.6b02081.

Tunnel analysis, specific activities, kinetics data, CD spectra, MALDI-MS/MS spectra, information on diffraction data and refinement statistics, data on design of mutagenesis, and results from molecular dynamics simulations (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail for J.D.: jiri@chemi.muni.cz.

Present Address

[#]Biotechnology Centre, Silesian University of Technology, ul. Krzywoustego 8, 44-100 Gliwice, Poland.

Author Contributions

[&]J.B. and P.B. contributed equally.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors express many thanks to the technician Hana Moskalikova (Masaryk University, Brno, Czech Republic) for

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

help with protein purification and collection of data using rapid quench flow and to Prof. Dan S. Tawfik (Weizmann Institute of Science, Rehovot, Israel) and Prof. Uwe T. Bornscheuer (University Greifswald, Greifswald, Germany) for critical reading of the manuscript. The authors are grateful to the user support teams of the European Synchrotron Radiation Facility at Grenoble beamline ID29 and BESSY II electron storage ring beamline 14.1 for assistance during the data collection. The work was supported by the Grant Agency of the Czech Republic (P503/12/0572, P207/12/0775, and GA16-06096S), the Czech Ministry of Education of the Czech Republic (LO1214, LM2011028, and LH14027) and the Research Support Program of Masaryk University (MUNI/M/1888/2014). A.G. was supported by the National Science Centre, Poland (DEC-2015/18/M/NZ1/00427). Computational resources were provided by the CESNET LM2015042 and the CERIT Scientific Cloud LM2015085, provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures".

REFERENCES

(1) Bartlett, G. J.; Porter, C. T.; Borkakoti, N.; Thornton, J. M. *J. Mol. Biol.* **2002**, *324*, 105−121.
(2) Prokop, Z.; Gora, A.; Brezovsky, J.; Chaloupkova, R.; Stepankova, V.; Damborsky, J. In *Protein Engineering Handbook*; Lutz, S., Bornscheuer, U. T., Eds.; Wiley-VCH: Weinheim, Germany, 2012; Vol. 3, pp 421−464.
(3) Kingsley, L. J.; Lill, M. A. *Proteins: Struct., Funct., Genet.* **2015**, *83*, 599−611.
(4) Pavlova, M.; Klvana, M.; Prokop, Z.; Chaloupkova, R.; Banas, P.; Otyepka, M.; Wade, R. C.; Tsuda, M.; Nagata, Y.; Damborsky, J. *Nat. Chem. Biol.* **2009**, *5*, 727−733.
(5) Biedermannova, L.; Prokop, Z.; Gora, A.; Chovancova, E.; Kovacs, M.; Damborsky, J.; Wade, R. C. *J. Biol. Chem.* **2012**, *287*, 29062−29074.
(6) Zhou, H. X.; Wlodek, S. T.; McCammon, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 9280−9283.
(7) Liebgott, P.-P.; Leroux, F.; Burlat, B.; Dementin, S.; Baffert, C.; Lautier, T.; Fourmond, V.; Ceccaldi, P.; Cavazza, C.; Meynial-Salles, I.; Soucaille, P.; Fontecilla-Camps, J. C.; Guigliarelli, B.; Bertrand, P.; Rousset, M.; Léger, C. *Nat. Chem. Biol.* **2010**, *6*, 63−70.
(8) Volbeda, A.; Martin, L.; Cavazza, C.; Matho, M.; Faber, B. W.; Roseboom, W.; Albracht, S. P. J.; Garcin, E.; Rousset, M.; Fontecilla-Camps, J. C. *JBIC, J. Biol. Inorg. Chem.* **2005**, *10*, 239−249.
(9) Thoden, J. B.; Huang, X.; Raushel, F. M.; Holden, H. M. *J. Biol. Chem.* **2002**, *277*, 39722−39727.
(10) Amaro, R. E.; Myers, R. S.; Davisson, V. J.; Luthey-Schulten, Z. A. *Biophys. J.* **2005**, *89*, 475−487.
(11) Fan, Y.; Lund, L.; Shao, Q.; Gao, Y.-Q.; Raushel, F. M. *J. Am. Chem. Soc.* **2009**, *131*, 10211−10219.
(12) Zhou, H.-X.; McCammon, J. A. *Trends Biochem. Sci.* **2010**, *35*, 179−185.
(13) Gora, A.; Brezovsky, J.; Damborsky, J. *Chem. Rev.* **2013**, *113*, 5871−5923.
(14) McCammon, J. A.; Northrup, S. H. *Nature* **1981**, *293*, 316−317.
(15) Law, R. J.; Henchman, R. H.; McCammon, J. A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 6813−6818.
(16) Davids, T.; Schmidt, M.; Böttcher, D.; Bornscheuer, U. T. *Curr. Opin. Chem. Biol.* **2013**, *17*, 215−220.
(17) Tan, X.; Loke, H.-K.; Fitch, S.; Lindahl, P. A. *J. Am. Chem. Soc.* **2005**, *127*, 5833−5839.
(18) Frank, R. A. W.; Titman, C. M.; Pratap, J. V.; Luisi, B. F.; Perham, R. N. *Science* **2004**, *306*, 872−876.
(19) Fedorov, R.; Vasan, R.; Ghosh, D. K.; Schlichting, I. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 5892−5897.
(20) Jez, J. M.; Bowman, M. E.; Noel, J. P. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 5319−5324.
(21) Lafaquière, V.; Barbe, S.; Puech-Guenot, S.; Guieysse, D.; Cortés, J.; Monsan, P.; Siméon, T.; André, I.; Remaud-Siméon, M. *ChemBioChem* **2009**, *10*, 2760−2771.
(22) Koudelakova, T.; Chaloupkova, R.; Brezovsky, J.; Prokop, Z.; Sebestova, E.; Hesseler, M.; Khabiri, M.; Plevaka, M.; Kulik, D.; Kuta Smatanova, I.; Rezacova, P.; Ettrich, R.; Bornscheuer, U. T.; Damborsky, J. *Angew. Chem., Int. Ed.* **2013**, *52*, 1959−1963.
(23) Floor, R. J.; Wijma, H. J.; Colpa, D. I.; Ramos-Silva, A.; Jekel, P. A.; Szymański, W.; Feringa, B. L.; Marrink, S. J.; Janssen, D. B. *ChemBioChem* **2014**, *15*, 1660−1672.
(24) Liskova, V.; Bednar, D.; Prudnikova, T.; Rezacova, P.; Koudelakova, T.; Sebestova, E.; Kuta Smatanova, K. S.; Brezovsky, J.; Chaloupkova, R.; Damborsky, J. *ChemCatChem* **2015**, *7*, 648−659.
(25) Blomberg, R.; Kries, H.; Pinkas, D. M.; Mittl, P. R. E.; Grütter, M. G.; Privett, H. K.; Mayo, S. L.; Hilvert, D. *Nature* **2013**, *503*, 418−421.
(26) Bornscheuer, U. T.; Huisman, G. W.; Kazlauskas, R. J.; Lutz, S.; Moore, J. C.; Robins, K. *Nature* **2012**, *485*, 185−194.
(27) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. *PLoS Comput. Biol.* **2012**, *8*, e1002708.
(28) Pavelka, A.; Sebestova, E.; Kozlikova, B.; Brezovsky, J.; Sochor, J.; Damborsky, J. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2016**, *13*, 505−517.
(29) Chaloupkova, R.; Sykorova, J.; Prokop, Z.; Jesenska, A.; Monincova, M.; Pavlova, M.; Tsuda, M.; Nagata, Y.; Damborsky, J. *J. Biol. Chem.* **2003**, *278*, 52622−52628.
(30) Reetz, M. T.; Carballeira, J. D. *Nat. Protoc.* **2007**, *2*, 891−903.
(31) Koudelakova, T.; Chovancova, E.; Brezovsky, J.; Monincova, M.; Fortova, A.; Jarkovsky, J.; Damborsky, J. *Biochem. J.* **2011**, *435*, 345−354.
(32) Hur, S.; Kahn, K.; Bruice, T. C. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 2215−2219.
(33) Olsson, M. H. M.; Warshel, A. *J. Am. Chem. Soc.* **2004**, *126*, 15167−15179.
(34) Negri, A.; Marco, E.; Damborsky, J.; Gago, F. *J. Mol. Graphics Modell.* **2007**, *26*, 643−651.
(35) Klvana, M.; Pavlova, M.; Koudelakova, T.; Chaloupkova, R.; Dvorak, P.; Prokop, Z.; Stsiapanava, A.; Kuty, M.; Kuta-Smatanova, I.; Dohnalek, J.; Kulhanek, P.; Wade, R. C.; Damborsky, J. *J. Mol. Biol.* **2009**, *392*, 1339−1356.
(36) Koudelakova, T.; Bidmanova, S.; Dvorak, P.; Pavelka, A.; Chaloupkova, R.; Prokop, Z.; Damborsky, J. *Biotechnol. J.* **2013**, *8*, 32−45.
(37) Moriuchi, R.; Tanaka, H.; Nikawadori, Y.; Ishitsuka, M.; Ito, M.; Ohtsubo, Y.; Tsuda, M.; Damborsky, J.; Prokop, Z.; Nagata, Y. *AMB Express* **2014**, *4*, 72.
(38) Sykora, J.; Brezovsky, J.; Koudelakova, T.; Lahoda, M.; Fortova, A.; Chernovets, T.; Chaloupkova, R.; Stepankova, V.; Prokop, Z.; Smatanova, I. K.; Hof, M.; Damborsky, J. *Nat. Chem. Biol.* **2014**, *10*, 428−430.
(39) Hasan, K.; Gora, A.; Brezovsky, J.; Chaloupkova, R.; Moskalikova, H.; Fortova, A.; Nagata, Y.; Damborsky, J.; Prokop, Z. *FEBS J.* **2013**, *280*, 3149−3159.
(40) Lu, L.-Y.; Hsieh, Y.-C.; Liu, M.-Y.; Lin, Y.-H.; Chen, C.-J.; Yang, Y.-S. *Mol. Pharmacol.* **2007**, *73*, 660−668.
(41) Reed, M. C.; Lieb, A.; Nijhout, H. F. *BioEssays* **2010**, *32*, 422−429.
(42) Holloway, P.; Trevors, J. T.; Lee, H. *J. Microbiol. Methods* **1998**, *32*, 31−36.
(43) Iwasaki, I.; Utsumi, S.; Ozawa, T. *Bull. Chem. Soc. Jpn.* **1952**, *25*, 226−226.
(44) Wold, S.; Esbensen, K.; Geladi, P. *Chemom. Intell. Lab. Syst.* **1987**, *2*, 37−52.
(45) Degtjarik, O.; Chaloupkova, R.; Rezacova, P.; Kuty, M.; Damborsky, J.; Kuta Smatanova, I. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2013**, *69*, 284−287.

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

(46) Minor, W.; Cymborowski, M.; Otwinowski, Z.; Chruszcz, M. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006**, *62*, 859−866.
(47) Gabadinho, J.; Beteva, A.; Guijarro, M.; Rey-Bakaikoa, V.; Spruce, D.; Bowler, M. W.; Brockhauser, S.; Flot, D.; Gordon, E. J.; Hall, D. R.; Lavault, B.; McCarthy, A. A.; McCarthy, J.; Mitchell, E.; Monaco, S.; Mueller-Dieckmann, C.; Nurizzo, D.; Ravelli, R. B. G.; Thibault, X.; Walsh, M. A.; Leonard, G. A.; McSweeney, S. M. *J. Synchrotron Radiat.* **2010**, *17*, 700−707.
(48) Mueller, U.; Darowski, N.; Fuchs, M. R.; Förster, R.; Hellmig, M.; Paithankar, K. S.; Pühringer, S.; Steffien, M.; Zocher, G.; Weiss, M. S. *J. Synchrotron Radiat.* **2012**, *19*, 442−449.
(49) Krug, M.; Weiss, M. S.; Heinemann, U.; Mueller, U. *J. Appl. Crystallogr.* **2012**, *45*, 568−572.
(50) Kabsch, W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 125−132.
(51) Vagin, A.; Teplyakov, A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 22−25.
(52) Marek, J.; Vevodova, J.; Smatanova, I. K.; Nagata, Y.; Svensson, L. A.; Newman, J.; Takagi, M.; Damborsky, J. *Biochemistry* **2000**, *39*, 14082−14086.
(53) Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 355−367.
(54) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011**, *67*, 235−242.
(55) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 486−501.
(56) Painter, J.; Merritt, E. A. *J. Appl. Crystallogr.* **2006**, *39*, 109−111.
(57) Winn, M. D.; Isupov, M. N.; Murshudov, G. N. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 122−133.
(58) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(59) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L.-W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 213−221.
(60) Afonine, P. V.; Grosse-Kunstleve, R. W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012**, *68*, 352−367.
(61) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010**, *66*, 12−21.
(62) *The PyMOL molecular graphics system, version 1.7*; Schrödinger, LLC.
(63) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; Onufriev, A. *Nucleic Acids Res.* **2005**, *33*, W368−371.
(64) Sindhikara, D. J.; Hirata, F. *J. Phys. Chem. B* **2013**, *117*, 6718−6723.
(65) Sindhikara, D. J.; Yoshida, N.; Hirata, F. *J. Comput. Chem.* **2012**, *33*, 1536−1543.
(66) Case, D. A.; Darden, T. A., III; Simmerling, C. L.; Wang, J.; Duke, R. E.; Luo, R.; Walker, R. C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. *AMBER 14*; University of California: San Francisco, 2015.
(67) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926−935.
(68) Salomon-Ferrer, R.; Götz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. *J. Chem. Theory Comput.* **2013**, *9*, 3878−3888.
(69) Le Grand, S.; Götz, A. W.; Walker, R. C. *Comput. Phys. Commun.* **2013**, *184*, 374−380.
(70) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *J. Chem. Theory Comput.* **2015**, *11*, 3696−3713.
(71) Joung, I. S.; Cheatham, T. E., 3rd *J. Phys. Chem. B* **2009**, *113*, 13279−13290.
(72) Joung, I. S.; Cheatham, T. E., 3rd *J. Phys. Chem. B* **2008**, *112*, 9020−9041.
(73) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089−10092.
(74) Essmann, U.; Perera, L.; Berkowitz, M.; Darden, T.; Lee, H.; Pedersen, L. *J. Chem. Phys.* **1995**, *103*, 8577−8593.
(75) Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J. C. *J. Comput. Phys.* **1977**, *23*, 327−341.
(76) Zwanzig, R. *J. Stat. Phys.* **1973**, *9*, 215−220.
(77) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684−3690.
(78) Roe, D. R.; Cheatham, T. E., 3rd *J. Chem. Theory Comput.* **2013**, *9*, 3084−3095.
(79) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33−38.
(80) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. *J. Cheminf.* **2012**, *4*, 17.
(81) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, Ö.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09 Revision E.01*; Gaussian Inc., Wallingford, CT, USA, 2009.
(82) Vanquelef, E.; Simon, S.; Marquant, G.; Garcia, E.; Klimerak, G.; Delepine, J. C.; Cieplak, P.; Dupradeau, F.-Y. *Nucleic Acids Res.* **2011**, *39*, W511−517.
(83) Dupradeau, F.-Y.; Pigache, A.; Zaffran, T.; Savineau, C.; Lelong, R.; Grivel, N.; Lelong, D.; Rosanski, W.; Cieplak, P. *Phys. Chem. Chem. Phys.* **2010**, *12*, 7821−7839.
(84) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, *117*, 5179−5197.
(85) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455−461.
(86) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. *J. Comput. Chem.* **2009**, *30*, 2785−2791.
(87) Hamelberg, D.; de Oliveira, C. A. F.; McCammon, J. A. *J. Chem. Phys.* **2007**, *127*, 155102.
(88) Pierce, L. C. T.; Salomon-Ferrer, R.; Augusto, F.; de Oliveira, C.; McCammon, J. A.; Walker, R. C. *J. Chem. Theory Comput.* **2012**, *8*, 2997−3002.
(89) Efron, B.; Tibshirani, R. *Stat. Sci.* **1986**, *1*, 54−75.

DOI: 10.1021/acscatal.6b02081
ACS Catal. 2016, 6, 7597−7610

ACS Catalysis

Research Article

Cite This: ACS Catal. 2018, 8, 2402−2412

pubs.acs.org/acscatalysis

Exploration of Enzyme Diversity by Integrating Bioinformatics with Expression Analysis and Biochemical Characterization

Pavel Vanacek,[†,‡,∥] Eva Sebestova,[†,∥] Petra Babkova,[†,‡] Sarka Bidmanova,[†,‡] Lukas Daniel,[†,‡] Pavel Dvorak,[†] Veronika Stepankova,[†,‡,§] Radka Chaloupkova,[†,‡,§] Jan Brezovsky,[†,‡] Zbynek Prokop,*[,†,‡,§] and Jiri Damborsky*[,†,‡]

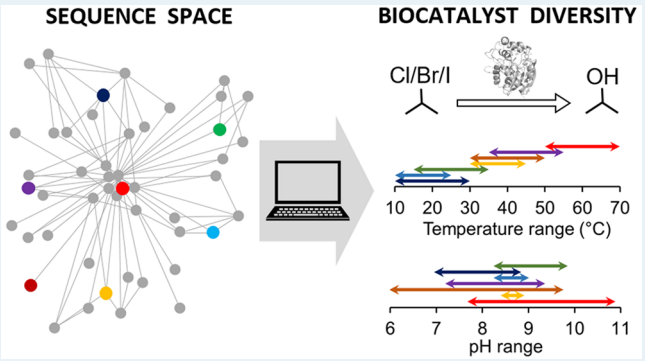
[†]Loschmidt Laboratories, Department of Experimental Biology and Research Centre for Toxic Compounds in the Environment RECETOX, Faculty of Science, Masaryk University, 625 00 Brno, Czech Republic

[‡]International Clinical Research Center, St. Anne's University Hospital, Pekarska 53, 656 91 Brno, Czech Republic

[§]Enantis Ltd., Biotechnology Incubator INBIT, Kamenice 34, 625 00 Brno, Czech Republic

Supporting Information

ABSTRACT: Millions of protein sequences are being discovered at an incredible pace, representing an inexhaustible source of biocatalysts. Here, we describe an integrated system for automated in silico screening and systematic characterization of diverse family members. The workflow consists of (i) identification and computational characterization of relevant genes by sequence/structural bioinformatics, (ii) expression analysis and activity screening of selected proteins, and (iii) complete biochemical/biophysical characterization and was validated against the haloalkane dehalogenase family. The sequence-based search identified 658 potential dehalogenases. The subsequent structural bioinformatics prioritized and selected 20 candidates for exploration of protein functional diversity. Out of these 20, the expression analysis and the robotic screening of enzymatic activity provided 8 soluble proteins with dehalogenase activity. The enzymes discovered originated from genetically unrelated Bacteria, Eukaryota, and also Archaea. Overall, the integrated system provided biocatalysts with broad catalytic diversity showing unique substrate specificity profiles, covering a wide range of optimal operational temperature from 20 to 70 °C and an unusually broad pH range from 5.7 to 10. We obtained the most catalytically proficient native haloalkane dehalogenase enzyme to date ($k_{cat}/K_{0.5}$ = 96.8 mM^{-1} s^{-1}), the most thermostable enzyme with melting temperature 71 °C, three different cold-adapted enzymes showing dehalogenase activity at near-to-zero temperatures, and a biocatalyst degrading the warfare chemical sulfur mustard. The established strategy can be adapted to other enzyme families for exploration of their biocatalytic diversity in a large sequence space continuously growing due to the use of next-generation sequencing technologies.



KEYWORDS: *diversity, sequence space, bioinformatics, biocatalyst, biochemical characterization, activity, substrate specificity, haloalkane dehalogenases*

Downloaded via MASARYK UNIV on September 2, 2021 at 20:51:03 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

INTRODUCTION

The postgenomic era is characterized by an exponential increase in the number of protein sequences,[1] which represent an immense treasure of novel enzyme catalysts with unexplored structural-functional diversity. Despite their enormous promise for biological and biotechnological discovery, experimental characterization has been performed on only a small fraction of the available sequences.[2] This "big data" problem is further extended by continuous genome and metagenome sequencing projects which employ powerful next-generation sequencing technologies.[3,4] Traditional biochemical techniques are time-demanding, cost-ineffective, and low-throughput, providing insufficient capacity for the exploitation of genetic diversity.[5] In response to these limitations, high-throughput experimental techniques employing miniaturization and automation have been developed in order to keep track of the ever-growing sequence information.[6−8] Although fluorescent biochemical assays implemented in microformats provide easily measurable signals, enzyme activities from determined end point measurements often differ from those obtained using native substrates ("You get what you screen for.")[6] These microformat techniques are powerful tools for the prefiltering of large libraries; however, they must be followed by additional assays with the target substrates. Robotic platforms using the microtiter plate format are therefore employed to provide quantitative kinetic data.[9] Despite these innovations, the

Received: October 16, 2017
Revised: January 28, 2018
Published: February 2, 2018

© 2018 American Chemical Society

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

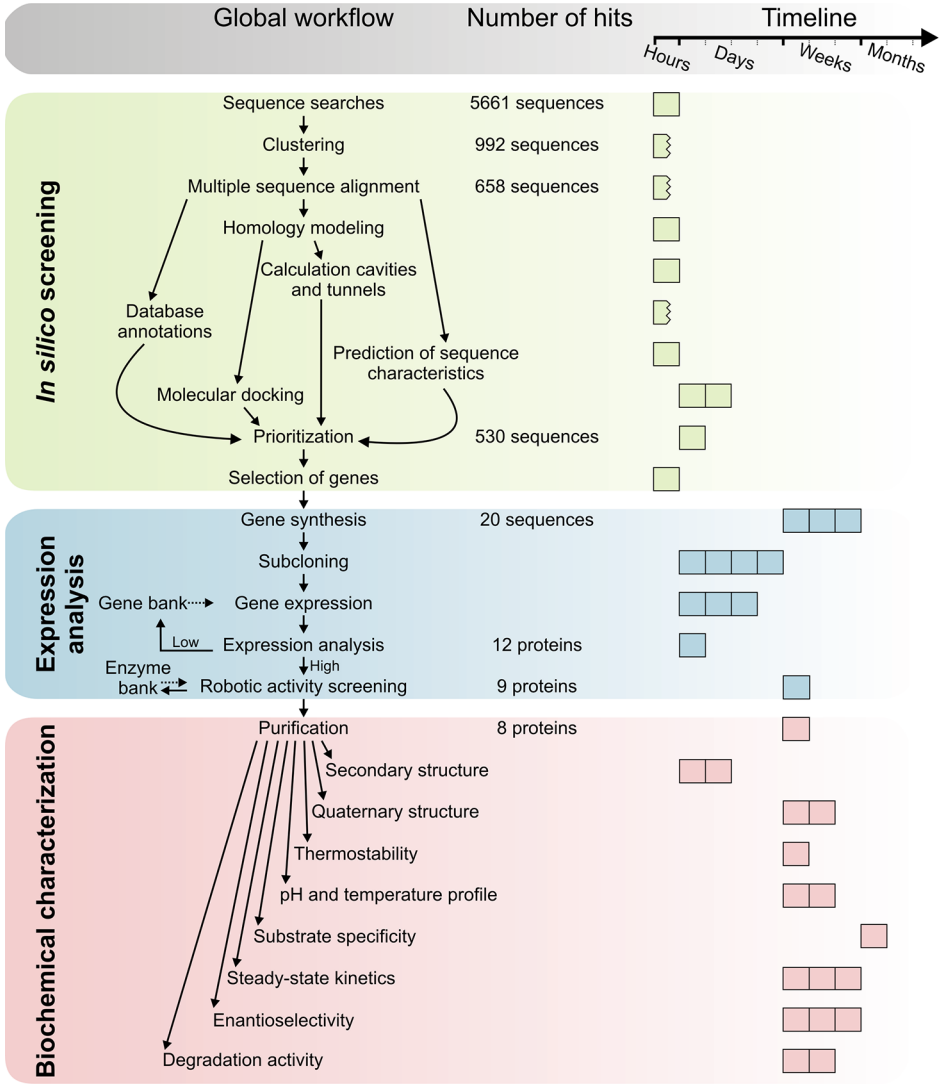


Figure 1. Workflow of an integrated system for the exploitation of the protein structural and functional diversity. Different colors highlight three distinct phases of the workflow: (i) automated sequence and structural bioinformatics (green), (ii) protein production and robotic activity screening (blue) and (iii) biochemical characterization (red). The timeline shows the periods required for data collection and analysis, divided into intervals of 4 h, 4 days, 3 weeks and 3 months column for clarity. A cut square indicates a time requirement of less than 1 h.

existing experimental methods do not provide sufficient capacity for the full biochemical/biophysical characterization of proteins spanning the ever-increasing sequence space, and new technical solutions are therefore required.

Computational approaches offer an adequate capacity for in silico screening of a large pool of sequence entries to facilitate the identification and rational selection of attractive targets for experimental testing. The recently published genomic mining strategy employing molecular modeling and structural bioinformatics has demonstrated the identification of enzymes catalyzing the targeted reaction in a synthetic pathway. This exciting approach led to the discovery of decarboxylases from more than 239 selected hits without expensive and laborious enzyme-engineering efforts.[10] The main benefit of this study lies in the effective sampling of particular enzymatic activity from sequence databases. Developing automated computational workflows and integrating them with experimental platforms is thus essential if the effective discovery of novel proteins is desired. In addition to their applicability in finding new members of protein families, they can identify a wealth of functional novelty when a homologue is found in an unexpected biological setting (such as a new species or environment) or co-occurring with other proteins.[11] Likely hotspots of functional novelty in sequence space may be uncovered either in under-sampled phyla from the tree of life or by finding functional shifts in sequence motifs or domain architecture.[12] Moreover, in silico analysis of structural and functional properties can reveal evolutionary changes in the enzymatic machinery.

Here, we describe an integrated system comprising automated in silico screening protocol and experimental procedures for characterization and the exploitation of the structural and functional diversity of an entire enzyme family (Figure 1). As proof of concept, we used this system to explore the diversity of microbial enzymes haloalkane dehalogenases (EC 3.8.1.5, HLDs). HLDs have been identified in a broad spectrum of microorganisms inhabiting soil, water, animal tissues, and symbiotic plants.[13−21] These α/β-hydrolase fold enzymes, which belong to one of the largest protein superfamilies (>100000 members), catalyze the hydrolytic

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

dehalogenation of a wide range of organohalogens. HLDs can be employed for the biocatalysis of optically pure building blocks,[22−24] the bioremediation of environmental pollutants,[25−27] the decontamination of chemical warfare agents,[28,29] the biosensing of pollutants,[30−32] and molecular imaging.[33−35] This diversity of reported practical applications is especially astonishing if one bears in mind that only two dozen HLDs have been biochemically characterized during the last 30 years, although the sequences of hundreds of putative HLDs are available in genomic databases.[20,36] The proposed integrated system effectively explored the sequence diversity and delivered eight novel biocatalyst possessing unique properties. Particularly, the most catalytically proficient HLD enzyme to date, the most thermostable biocatalyst and the extremophile-derived enzymes are promising for various biotechnological applications. The strategy was critically evaluated by validation of in silico predictions against experimentally verified results.

EXPERIMENTAL SECTION

In Silico Screening. The sequences of three experimentally characterized HLDs were used as queries for two iterations of PSI-BLAST v2.2.28+[37] searches against the NCBI nr database (version 25-9-2013)[38] with *E*-value thresholds of 10^{-20}. Information about the source organisms of all putative HLDs was collected from the NCBI Taxonomy and Bioproject databases.[38] A multiple sequence alignment of all putative full-length HLD sequences was constructed by Clustal Omega v1.2.0.[39] The homology modeling was performed using Modeler v9.11.[40] Pockets in each homology model were calculated and measured using the CASTp program[41,42] with a probe radius of 1.4 Å. The CAVER v. 3.01 program[43] was then used to calculate tunnels in the ensemble of all homology models. The three-dimensional structures of 34 halogenated compounds, which are environmental pollutants, artificial sweeteners, chemical warfare agents, or their surrogates and disinfectants, were constructed in Avogadro[44] and docked to the catalytic pockets using AutoDock 4.2.3. Each local search was based on the pseudo Solis and Wets algorithm with a maximum of 300 iterations per search.[45] The chance for soluble expression in *E. coli* of each protein was predicted based on the revised Wilkinson-Harrison solubility model.[46,47] Detailed bioinformatics protocols are described in the Supporting Information.

Expression Analysis and Activity Screening. Codon-optimized genes encoding 20 putative HLDs were designed and commercially synthesized. The synthetic genes were subcloned individually into the expression vector pET21b between the *Nde*I/*Xho*I restriction sites. *E. coli* BL21(DE3), *E. coli* ArcticExpress(DE3), and *E. coli* Rosetta-gami B(DE3) pLysS competent cells were transformed with DNA constructs using the heat-shock method and expressed in lysogeny broth (LB) or EnPresso B medium. Biomass was harvested at the end of the cultivation, washed, and disrupted using a homogenizer. The activity of cell-free extract toward 1-iodobutane, 1,2-dibromoethane, and 4-bromobutyronitrile substrates was robotically screened at 10, 37, and 55 °C. Detailed experimental protocols are described in the Supporting Information.

Biochemical and Biophysical Characterization. Enzymes were purified using single-step nickel affinity chromatography. Secondary structure was evaluated using circular dichroism spectroscopy at room temperature. Size-exclusion chromatography with static light scattering, refractive index, ultraviolet and differential viscometer detectors was used to analyze protein quaternary structure, molecular weights, hydrodynamic radius, and intrinsic viscosities. Thermal stability was analyzed by circular dichroism spectroscopy and robotic differential scanning calorimetry. The thermal unfolding was monitored by change in the ellipticity or heat capacity over the temperature range from 20 to 90 °C. The temperature profile was determined as an effect of temperature on enzymatic activity toward 1,3-diiodopropane at pH 8.6 over the temperature range from 5 to 80 °C. The pH profile was determined as an effect of pH on enzymatic activity toward 1,3-dibromopropane at the pH ranging from 4 to 12 at 10, 37, or 55 °C. Substrate specificity toward a set of 30 halogenated compounds was analyzed at 10, 37, or 55 °C. The specific activity data toward 30 substrates were analyzed by Principal Component Analysis (PCA). The steady-state kinetics of the novel HLDs toward 1,2-dibromoethane were measured using an isothermal titration calorimeter at either 10, 37, or 55 °C. Enantioselectivity was evaluated from kinetic resolution of 2-bromopentane or ethyl 2-bromopropionate at 20 °C. Enzymatic activity toward chemical warfare agent sulfur mustard was measured using fluorescent assay. The degradation of selected environmental pollutants, 1,3-dichloropropene, γ-hexachlorocyclohexane, and hexabromocyclododecane was analyzed using robotic GC/MS. Detailed experimental protocols are described in the Supporting Information section.

RESULTS

In Silico Screening. The developed automated workflow for the in silico identification and characterization of HLDs provides a useful tool for the selection of interesting proteins for experimental characterization. Sequence database searches led to the identification of 5661 sequences representing putative HLDs and their close relatives. In order to automatically distinguish between putative HLD sequences and sequences of proteins from other families, average-link hierarchical clustering based on pairwise sequence distances was applied. After removing 39 artificial sequences, 953 putative HLDs were retained (333 from HLD-I, 295 from HLD-II, 314 from HLD-III, and 11 from HLD-IIIb). On the basis of multiple sequence alignment, 117 incomplete and 178 degenerate sequences were excluded from the data set. The substitution of halide-stabilizing residues was the most common reason for exclusion (Table S2). The remaining 658 putative HLDs were subjected to homology modeling and in silico characterization. The following data were gathered for each putative HLD: (i) sequence annotations, (ii) the taxonomy of the source organism, (iii) extremophilic properties of the source organism, (iv) a list of highly similar proteins and the most closely related known HLDs, (v) the HLD subfamily, (vi) composition of the catalytic pentad, (vii) the domain composition, (viii) the predicted solubility, and if applicable, (ix) suitable template for homology modeling and (x) constructed homology model, (xi) the volume and area of the catalytic pocket, (xii) characteristics of access tunnels, and (xiii) structures of enzyme−substrate complexes.

The majority of sequences in HLD-I, HLD-II, and HLD-IIIb were correctly annotated (75%, 80%, and 71%, respectively), while this was the case for only 26% of the HLD-III sequences (Table S3). More than half of the sequences in HLD-III were annotated generally as α/β-hydrolase fold enzymes (45%) or hydrolases (10%). Due to the presence of multidomain proteins, 3% of the HLD-I sequences and 11% of the HLD-III sequences were annotated as CMP deaminases and acyl-

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

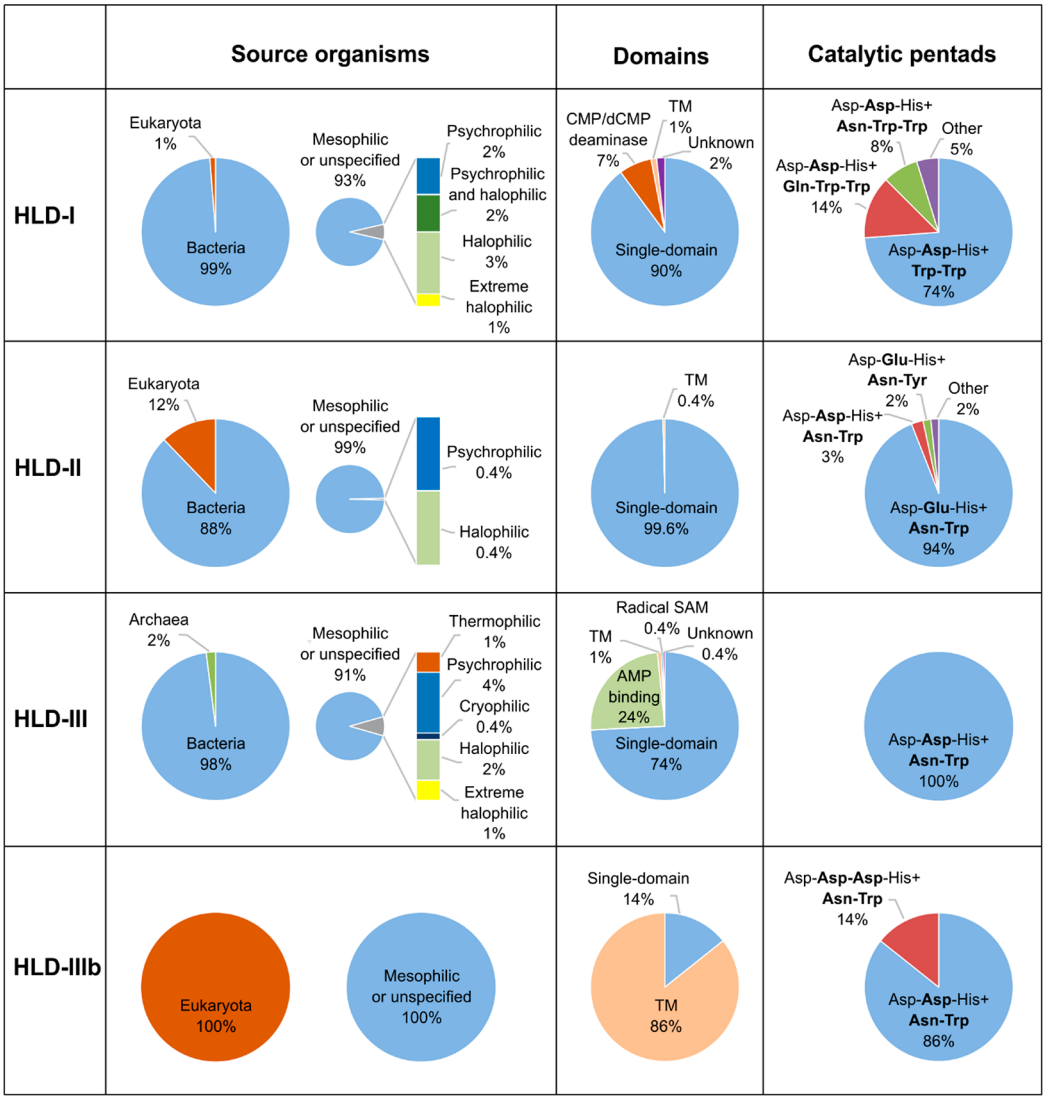


Figure 2. Overview of putative HLDs identified. Most putative HLD sequences are composed of one α/β-hydrolase domain (single-domain). Additionally, some sequences contain the N-terminal cytidine and deoxycytidylate deaminase domain (CMP/dCMP deaminase), the N-terminal radical SAM domain (radical SAM), the C-terminal AMP binding domain (AMP binding), or transmembrane helices (TM). The catalytic pentad in HLDs is composed of nucleophile-catalytic acid−base+halide-stabilizing residues. The nucleophile and catalytic base are conserved in all family members, whereas the catalytic acid and halide-stabilizing residues are variable (highlighted in bold).

CoA synthetases, respectively. Miss-annotations of single-domain proteins were rare (2 proteins). An overview of source organisms, catalytic pentads, and domain compositions of the putative HLDs identified is provided in Figure 2. The sequences of putative HLDs were identified in the genomes of organisms from all three domains of life. The source organisms included 3 thermophilic, 1 cryophilic, 13 psychrophilic, 3 psychrophilic-moderate halophilic, 12 moderate halophilic, and 4 extreme halophilic strains. The majority of the putative HLDs of extremophilic origin were found in subfamilies HLD-I and HLD-III. The prevalent compositions of the catalytic pentads agreed with those described previously.[36] Potential alternative catalytic pentad compositions were predicted for 26% of HLD-I, 6% of HLD-II, and 14% of HLD-IIIb sequences.

Even though all HLDs that have been experimentally characterized to date are single-domain proteins, we identified a number of multidomain proteins in the HLD-I and HLD-III subfamilies. The N-terminal cytidine and deoxycytidylate deaminase domain were detected in 12 HLD-I sequences, while 60 HLD-III sequences had a C-terminal AMP binding domain and one HLD-III sequence had the N-terminal radical SAM domain. N-Terminal transmembrane helices were predicted for 6 out of 7 HLD-IIIb sequences, while they were present in only a small proportion of sequences from the other three subfamilies. Since the majority of structurally characterized HLDs can be found in the HLD-II subfamily, the most reliable homology models could be constructed for members of this subfamily (Figure S1). No protein structures are currently available for members of the HLD-III/b subfamilies, limiting the possibility of homology modeling in these subfamilies (10% for HLD-III and none for HLD-IIIb; Figure S2). Overall, homology models were built and evaluated for 275 sequences.

The predicted volumes of catalytic pockets ranged from 126 Å^3 to 1981 Å^3 (Figure S3). Generally, the HLD-I subfamily members were predicted to have smaller pockets (median volume 554 Å^3) than HLD-II (716 Å^3) and HLD-III (777 Å^3). Tunnels were calculated for the ensemble of all superimposed

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

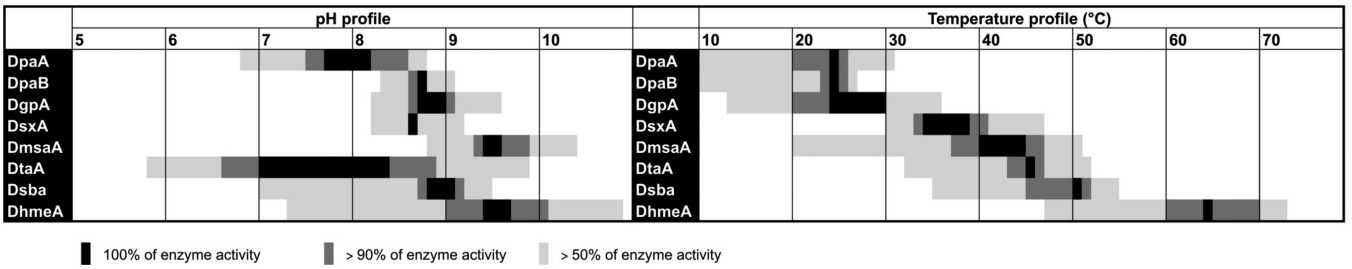
Table 1. Annotations and Predicted Properties of 20 Putative HLDs Selected for Experimental Characterization[a]

	source organism								
name	organism	taxonomy	extremophile	subfamily	predicted catalytic pentad	presence of additional protein domain	volume of active site cavity [Å^3]	radii of transport tunnels p1/p2/p3 [Å]	number of predicted active compounds
DacA	*Aspergillus clavatus* NRRL 1	eukaryota[b]		IIIb[b]	2 catalytic acids[b]	TM helices[b]	N/A	N/A	N/A
DssA	*Shewanella sediminis* HAW-EB3	bacteria	psychrophile[b]	I	3 halide-stabilizing residues[b]		N/A	N/A	N/A
DcaA	*Chloroflexus aurantiacus* J-10-fl	bacteria	thermophile[b]	III	standard		N/A	N/A	N/A
DdaA	*Desulfatibacillum alkenivorans* AK-01	bacteria		III	standard	radical SAM[b]	N/A	N/A	N/A
DamA	*Amycolatopsis mediterranei* S699	bacteria		II	standard		1507[b]	2.1/1.6/1.0	5
DmtA	*Microbacterium testaceum* StLB037	bacteria		III	standard	AMP-binding[b]	N/A	N/A	N/A
DtaA	*Trichoderma atroviride* IMI 206040	eukaryota[b]		II	standard		348	1.3/1.4/–	8[b]
DsbA	*Streptomyces bingchenggensis* BCW-1	bacteria		II	standard		796	2.2/1.3/–	9[b]
DhmeA	*Haloferax mediterranei* ATCC 33500	archaea[b]	moderate halophile[b]	III	standard		N/A	N/A	N/A
DadA	*Alcanivorax dieselolei* B5	bacteria		II	standard		1218	1.5/1.6/–	5[b]
DmgA	marine gamma proteobacterium HTCC2148	bacteria		I	3 halide-stabilizing residues[b]		N/A	N/A	N/A
DpaB	*Paraglaciecola agarilytica* NO_2	bacteria	moderate halophile + psychrophile[b]	I	standard		396	1.9/–/–	4
DpaA	*Paraglaciecola agarilytica* NO_2	bacteria	moderate halophile + psychrophile[b]	I	standard		943	1.9/–/1.4	1
DmsaA	*Marinobacter santoriniensis* NKSG1	bacteria	extreme halophile[b]	I	standard		928	1.9/1.2/1.2	3
DgpA	gamma proteobacterium NOR5-3	bacteria	psychrophile[b]	I	3 halide-stabilizing residues[b]		N/A	N/A	N/A
DcsA	*Caenispirillum salinarum* AK4	bacteria		II	standard		N/A	N/A	N/A
DsxA	*Sandarakinorhabdus* sp. AAP62	bacteria		I	standard		126[b]	1.8/–/–	7
DlxA	*Limnohabitans* sp. Rim47	bacteria		I	3 halide-stabilizing residues[b]	dCMP_cyt_deam[b]	456	2.1/1.1/1.1	7
DncA	*Nonomuraea coxensis*	bacteria		II	standard		1981[b]	2.5/1.6/1.6[b]	1
DmmaA	*Mycobacterium marinum* str. Europe	bacteria		II	standard		545	2.1/1.1/1.1	8[b]

[a]N/A, prediction not available due to lack of homology model. [b]Parameters used as selection criteria.

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

Table 2. Temperature and pH Profiles of 8 Biochemically Characterized HLDs



homology models and then clustered, enabling automated and direct comparison of the tunnels identified in different proteins. The three top-ranked tunnel clusters, corresponding to the p1, p2, and p3 tunnels, were further analyzed. Given that the probe radius used was 1 Å, the p1 and p2 tunnels were found in the majority of the models analyzed (99% and 84%, respectively), while the p3 tunnel was identified in only 50% of models (Figure S4). Finally, 34 potential HLD substrates were docked to homology models and evaluated (Table S4). Generally, the largest number of substrates in the reactive orientation was detected for HLD-II (Figure S5). Almost 15% of HLD-II members were found to have more than 5 substrates in the reactive orientation, compared to 5% and 4% of the HLD-I and HLD-III subfamily members, respectively. One or no substrates were identified for 67% of the HLD-III members, compared to 43% of the HLD-I and 43% of HLD-II members.

The subsequent semirational selection was employed to prioritize the computationally characterized set of 658 putative HLDs based on their sequence and structural characteristics as well as the annotations available. To further enrich the sequence diversity for this fold family, we excluded putative HLDs that had sequence identity lower than 90% to any that had been experimentally characterized. Gathered data for the remaining 530 sequentially distinct putative HLDs were compiled into the data set, on which the selection criteria were applied. The initial criterion was aimed to maximize the diversity of the target properties within the HLD family. Further we prioritized the putative HLDs predicted as soluble with high probability, those with lowest identity to known HLDs, or if the homology model was available, those predicted as likely active HLDs with high confidence. Simultaneously, we strived to filter out or minimize the candidates from HLD-III subfamily, since they are often difficult-to-express. This procedure was followed until the required number of hits with the given properties was received as illustrated in Table S5. It enabled sampling of putative HLDs with the most diverse or completely novel properties. The set of 20 selected genes encoding putative HLDs consisted of (i) bacterial, eukaryotic, and archaeal enzymes, (ii) single- and multidomain enzymes, (iii) enzymes originating from extremophilic organisms, (iv) enzymes belonging to four subfamilies, (v) enzymes with alternative composition of the catalytic pentads, (vi) enzymes with small and large active-site cavity, and (vii) enzymes possessing HLD activity with high confidence (Table 1, Table S6). Putative HLDs are denoted according to the previously established convention (first letter is "D" for dehalogenase; second and third letters are the initials of the source organism; and a last letter "A" or "B" refers to the first or second HLD from a single source organism).

Expression Analysis and Activity Screening. The expression analysis of putative HLD genes was performed in three different *E. coli* strains, with two types of cultivation media and under five expression conditions (Figure S6). In total, 15 (75%) out of 20 target genes were overexpressed and 12 (60%) genes provided proteins in a soluble form that enabled screening of enzymatic activity. We observed some agreement between the theoretically predicted and the experimentally determined solubility of the 12 HLDs that were successfully produced in *E. coli* (Table S7). A robotic platform employing a 96-well microtiter plate format was setup for fast screening of the HLD activity. Altogether, the activities of cell-free extracts of 12 soluble putative HLDs were screened against 3 diverse halogenated substrates (1-iodobutane, 1,2-dibromoethane, and 4-bromobutyronitrile) in an all-enzyme versus an all-substrate screen format at 3 temperatures regime (10, 37, and 55 °C) to maximize the chance of detecting HLD activity. Out of these 12 soluble HLDs, 9 exhibited hydrolytic activity toward at least one of the substrates tested (Table S8). Eight out of nine novel HLDs were successfully purified and subjected to detailed biochemical and biophysical characterization.

Biochemical Characterization. Biochemical and biophysical characterization included the (i) determination of folding and secondary structure, (ii) quaternary structure, (iii) thermostability, (iv) temperature and pH optima, (v) substrate specificity, (vi) steady-state kinetics, (vii) enantioselectivity, and (viii) activity toward selected environmental pollutants and warfare agents.

Circular dichroism spectroscopy was used to assess the correctness of folding and secondary structure composition. All enzymes exhibited CD spectra characteristic of α-helical content and proper folding[48] (Figure S7). The quaternary structure was examined by size-exclusion chromatography. DpaB, DsxA, DgpA, DtaA, and DsbA are monomeric proteins with calculated molecular weights (M_w) ranging between 31.2 and 36.9 kDa. DpaA is as a dimer with a determined M_w = 70.6 kDa and DhmeA is an oligomer with M_w > 2 MDa (Table S9). The oligomeric state of DmsaA could not be determined due to protein precipitation under the conditions tested. Thermally induced denaturation of novel HLDs was tested by monitoring ellipticity or heat capacity (Table S10). Three enzymes originating from psychrophilic organisms (DpaA, DpaB, and DgpA) exhibited significantly lower melting temperatures T_m = 35.2−38.2 °C in comparison to the enzymes of mesophilic origin with T_m = 47.2−54.6 °C. The highest melting temperature T_m = 70.6 °C was determined for the archaeal enzyme DhmeA, making this enzyme the most thermostable wild-type HLD ever reported.

The temperature and pH profiles were determined using a high-throughput robotic assay. Altogether, the novel HLDs exhibited high diversity of operational temperature optima ranging from 20 to 70 °C, while the majority of previously

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

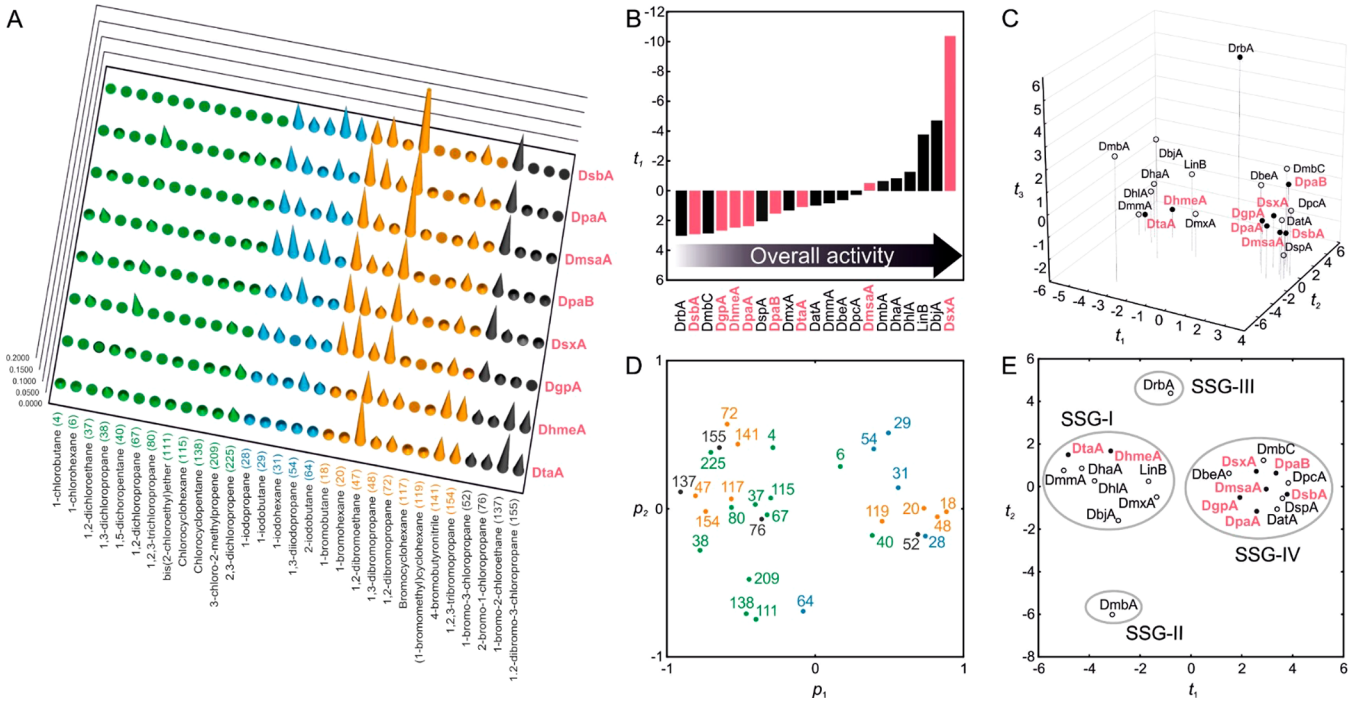


Figure 3. Comparison of the substrate specificities and overall catalytic activities of novel HLDs with previously characterized enzymes using multivariate statistics. (A) Substrate specificity profiles of newly discovered HLDs toward 30 halogenated substrates. (B) The score-contribution plot t_1 shows differences in overall activities of individual HLDs and explains 40% of the variance in the untransformed data set. (C) The score plot $t_1/t_2/t_3$ from PCA of the transformed data set, which suppressed differences in absolute activities and allowed substrate specificity profiles to be compared. The score plot, which describes 55.5% of the variance in the data set, shows enzymes clustered in individual substrate-specificity groups (SSGs). Objects (HLDs) with similar properties (specificity profiles) are colocalized. (D+E) The loading plot p_1/p_2 and the corresponding score plot t_1/t_2 from PCA of the transformed data set. The score plot describes 45.0% of the variance in the data set. The loading plot shows the main substrates for each SSG. Brominated substrates are shown in amber, chlorinated in green, iodinated in blue, and brominated and chlorinated in black.

characterized HLDs[35] possess a narrow range of temperature optima from 30 to 50 °C (Table 2). Moreover, DhmeA showed nearly 90% of its maximum activity even at 70 °C (Figure S8). In correspondence to previously characterized HLDs, the optimal pH conditions ranges between neutral to moderate alkaline condition for most of the novel HLDs variants (Figure S9). Remarkably wide pH tolerance was observed in the case of DhmeA which maintain significant activity in alkaline conditions to nearly pH 11 and DtaA covering unusually broad pH range between 5.7 and 10.0.

Substrate specificity was assessed with a panel of 30 halogenated hydrocarbons (Figure S10). All enzymes exhibited a preference for halogenated hydrocarbons in the following order: brominated > iodinated ≫ chlorinated (Table S11). The optimal length of the alkyl-chain for the substrate was between three and four carbon atoms (Figure 3A). The majority of enzymes possessed rather narrow substrate specificities, converting 12−22 out of the 30 halogenated hydrocarbons. The broadest substrate specificity profile was detected for DtaA, which showed activity toward 27 of the substrates, including the recalcitrant environmental pollutant 1,2,3-trichloropropane. PCA of the transformed data clustered the novel enzymes within substrate specificity groups (SSG) based on their overall substrate specificity profiles (Figure 3, panels C and E). The novel HLDs are spread across different SSGs, indicating significant diversity. DhmeA and DtaA were clustered in SSG-I, whereas the rest of enzymes in SSG-IV. PCA of untransformed specific activities compared the overall activities of HLDs (Figure 3B). Many of the new enzymes exhibited moderate or low activities, indicating that the natural substrates for the new family members are not adequately covered by the used set of 30 representative compounds. The striking exception is the enzyme DsxA, which showed outstanding activity toward most of the tested substrates. This enzyme exhibited the highest specific activity 493.7, 661.2, and 444.6 nmol s^{-1} mg^{-1} toward 1-bromobutane, 1,3-dibromopropane, and 1-bromo-3-chloropropane, respectively. This is 2−3 times higher than those recorded for any previously identified HLDs.

The catalytic properties of the novel enzymes were assessed by measuring steady-state kinetics toward the model substrate 1,2-dibromoethane (Table S12). With the exception of DmsaA, which provided classical hyperbolic kinetics, all novel enzymes exhibited positive cooperativity with Hill coefficients from 1.3 to 2.4. All newly discovered HLDs have shown high $K_{0.5}$ values with 1,2-dibromoethane (≥2.1 mM). The kinetics of DsxA was determined also with 1,3-dibromopropane. Significantly lower $K_{0.5}$ = 0.17 mM and high k_{cat} = 16.5 s^{-1} results in exceptionally high catalytic efficiency ($k_{cat}/K_{0.5}$ = 96.82 mM^{-1} s^{-1}), which is approximately 5.5-fold higher in comparison to the most efficient native HLD of today DadB.[17]

The enantioselectivity of novel HLDs was assayed by determining the kinetic resolution of model *β*-bromoalkane (2-bromopentane) and model *α*-bromoester (ethyl 2-bromopropionate). High enantioselectivity (*E*-value > 200) toward ethyl 2-bromopropionate was observed with DpaA, DpaB, DsxA, and DtaA (Figures S11 and S12), moderately enantioselective DmsaA and DhmeA provided *E*-value 54 and 79, and negligible enantioselectivity was shown by DgpA and DsbA (Table S13). In most cases, only weak enantioselectivity (*E*-value < 15) was recorded with 2-bromopentane. The only

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

exception was DpaA, which exhibited moderate enantioselectivity with *E*-value 85. All enzymes showed (*R*)-enantiopreference which is in correspondence with previously characterized HLDs.[22]

Degradation of halogenated environmental pollutants and warfare agents is one of the main applications of HLDs (Table S14). The activity of novel HLDs has been tested with 1,3-dichloropropene, hexabromocyclododecane, and γ-hexachlorocyclohexane. The majority of enzymes showed significant catalytic activities toward 1,3-dichloropropene (0.5−149.3 nmol s^{-1} mg^{-1}), a synthetic compound introduced into the environment through its use as a fumigant. Activity to other tested environmental pollutants has not been detected. Novel enzymes were also screened with the chemical weapon sulfur mustard. Sulfur mustard is a prominent warfare chemical which has been shown to be transformed to the nontoxic product through the action of HLDs.[28] The screening identified significant activity of DtaA toward sulfur mustard (1.46 nmol s^{-1} mg^{-1}), which makes this enzyme a potential candidate for use in decontamination mixtures or biodetection devices.

DISCUSSION

In this study, we bring our contribution to the big data challenge in the postgenomics era by the development of an automated in silico screening protocol for the exploitation of the protein functional diversity within an enzyme family. Since the rapidly growing genomic databases may contain vague sequence annotations and miss-annotations, our sequence-based search was employed to identify the new members of a protein family based on their evolutionary relationships to other known family members. The putative HLD sequences were automatically identified using global pairwise sequence identities and average-link hierarchical clustering. Furthermore, we cut the hierarchy of sequences at the level of individual HLD-subfamilies, this minimizing the risk of selecting non-HLD sequences.

The sequence analysis identified a number of putative HLDs whose catalytic pentads had alternative compositions: (i) three halide-stabilizing residues, (ii) two catalytic acids, (iii) the HLD-I/II members which have halide-stabilizing Gln/Tyr, (iv) and the HLD-II members with Asp serving as the catalytic acid. It is unclear whether proteins containing these abnormalities represent functional HLDs. Only 58% of the putative HLDs were functionally annotated. While miss-annotations were rare, many proteins were annotated as α/β-hydrolases or hypothetical proteins. The largest number of annotation problems occurred in the HLD-III subfamily, which contains only three experimentally characterized enzymes.[16,20] All 658 putative HLDs were characterized computationally to provide criteria for candidate selection exploring the diversity within the identified sequence space. The selection procedure was based on the mapping of the sequence and structural characteristics as well as annotations. The procedure yielded candidate proteins originating from new species, environments, and under-sampled phyla; proteins with novel domain combinations; proteins with alternative composition of the catalytic pentad; and proteins belonging to newly identified subfamily.

To critically evaluate the effectiveness and the main bottlenecks of the platform, we performed the validation of predictions against experimental data. The expression analysis was performed in three different *E. coli* strains, with two types of cultivation media under five different conditions. Although available solubility prediction tools have been employed during the selection of the candidates and a number of diverse expression conditions have been tested, we attained only 60% success rate for the production of soluble proteins. This is in an agreement with previously published large-scale expression trials demonstrating that 50−80% of bacterial proteins and 15−20% of nonbacterial proteins can be produced in *E. coli* in a soluble form.[49−51] The production of soluble proteins for experimental characterization remains a challenging, "hit-or-miss" affair, and currently represents the biggest bottleneck in studies of this type. With regard to time requirements and cost effectiveness, a more reasonable strategy is to apply expression screening to a larger number of candidates from protein databases rather than wasting time and resources on optimizing the production of "difficult-to-express" proteins. Robust expression systems must constitute an indispensable component of studies of this type.[50] Prediction of protein solubility using software tools represents an attractive future perspective with a challenge toward development of methods achieving higher reliability of estimates.

The initial robotic activity screening performed with cell-free extracts revealed 9 active enzymes out of 12 tested, implying a 75% success rate. Lack of activity in the case of 3 enzymes may have been due to low levels of HLD in the cell-free extract, requirements for specific conditions or preferences for unknown substrate. The latter may indicate that the previously identified "universal" substrates may not be preferred for all existing HLDs. In order to minimize the risk of missing interesting biocatalysts, the substrates used in the activity screening step should be carefully considered. Thanks to the miscellaneous origins and selection approach oriented to maximize diversity of selected candidates, the identified enzymes exhibited a wide range of characteristics with several unique properties. They originated from various phylogenetically unrelated organisms belonging to the domains of Bacteria, Eukaryota, and newly also Archaea. This first archaeal HLD with melting temperature 71 °C represents the most thermostable wild-type HLD known to date. On the contrary, DpaA, DpaB, and DgpA, originating from psychrophilic organisms with melting temperatures below 40 °C open the possibility for the operation at near-to-zero temperatures, which is attractive mainly for environmental applications (e.g., biodegradation or biosensing).[35] The eminent diversity of the novel variants with a wide range of optimal temperature from 20 to 70 °C and broad pH range from 6 to 11 offers expanding operational window to biotechnological applications. Observed extremophilic characteristics were reliably predicted by in silico protocol.

The majority of the newly isolated HLDs exhibited moderate or low enzymatic activities toward 30 halogenated compounds, suggesting that the currently used representative set of substrates for this enzyme family may lack some relevant native substrates. One notable exception is DsxA from *Sandarakinorhabdus* sp. AAP62, which showed exceptionally high activity toward many brominated substrates, particularly those with alkyl-chains containing 2−4 carbon atoms. This observation corresponds with the restricted volume of the active-site cavity predicted by homology modeling. Importantly, DsxA represents the most catalytically efficient member of this enzyme family ever isolated. Crystallographic and cryo-electron microscopy analysis of several newly isolated HLDs is ongoing in our laboratory. DhmeA from the HLD-III subfamily and DgpA with unique 30 amino acid insertion in the N-terminal part of the cap domain are particularly interesting targets.

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

Solving the very first structure from subfamily-III will provide a template for predicting the 3D structures of other subfamily members. In subsequent work, we will implement and release the computational part of the workflow as a user-friendly web tool. The experimental testing process will be extended by integration of miniaturized lab-on-chip assays, requiring only tiny fractions of a protein material and providing increase in the throughput.

CONCLUSIONS

In summary, we have demonstrated that the enormous wealth of genomic sequences available in public databases can be efficiently explored by in silico mapping of proteins structural and functional diversity within protein families. Integration of sequence/structural bioinformatics with experimental procedures enabled us to narrow down the number of enzyme candidates under consideration and allowed their catalytic properties to be explored with reasonable expenditures of time and effort. Although examples of sequence-mining platforms have previously been reported, here we describe integrated platform for the computational analysis of the structural and functional diversity of an entire enzyme family, coupled to full biochemical and biophysical characterization of the identified hits.[10,52] Using our platform, number of novel HLDs with potential practical uses have been identified, characterized, and made available to the community in industry and academia. Further application of our platform to other enzyme families will expand our knowledge in the field of enzymology and will lead to the discovery of novel biocatalysts for the biotechnological and pharmaceutical industries.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acscatal.7b03523.

Supporting Methods: in silico screening, expression analysis, biochemical and biophysical characterization; Supporting Tables: list of known HLDs used for definition of HLD clusters, catalytic residues in putative degenerated HLD sequences, annotations of putative HLD sequences, list of the chemical formulas of 34 halogenated compounds, overview of the selection process with respective criteria, accession numbers of putative HLDs, overview of expression analysis of 20 putative HLDs, robotic screening of enzymatic activity, the parameters related to size, shape, and oligomeric state of HLDs, melting temperatures of HLDs by CD spectroscopy and robotic DSC, specific activities of HLDs toward 30 halogenated substrates, steady-state kinetic parameters of HLDs, enantioselectivities of HLDs toward 2-bromopentane and ethyl 2-bromopropionate, activities of HLDs toward warfare agent and environmental pollutants; Supporting Figures: representative 3D structure of HLD, availability of homology modeling templates in individual HLD subfamilies, distributions of predicted volumes of catalytic pockets in HLD subfamilies, distributions of predicted bottleneck radii of p1, p2, and p3 tunnels in HLD subfamilies, distributions of mechanism-based geometric criteria for reactivity in HLD subfamilies, expression analysis of the set of 20 putative HLDs, far-UV circular dichroism spectra of HLDs, temperature profiles, pH profiles, the set of 30 halogenated substrates, kinetic resolution of 2-bromopentane, and kinetic resolution of ethyl-2-bromopropionate (PDF)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: zbynek@chemi.muni.cz. Tel: +420-5-4949-6667.
*E-mail: jiri@chemi.muni.cz. Tel: +420-5-4949-3467.

ORCID

Jiri Damborsky: 0000-0002-7848-8216

Author Contributions

‖P.V. and E.S. contributed equally.

Notes

The authors declare the following competing financial interest(s): Dr. Veronika Stepankova, Dr. Radka Chaloupkova and Dr. Zbynek Prokop work part-time at the University biotechnology spin-off company Enantis Ltd. Enzymes isolated throughout this study will be produced in sufficient quantities and distributed upon request to the community.

ACKNOWLEDGMENTS

The work was supported by the Grant Agency of the Czech Republic (GA16-06096S, GA16-07965S, and GA16-24223S), the Technological Agency of the Czech Republic (TH02010219), the Ministry of Education, Youth, and Sports of the Czech Republic (LQ1605 and LO1214) and the European Union (720776 and 722610). Computational resources were provided by CESNET (LM2015042) and the CERIT Scientific Cloud (LM2015085), as part of the "Projects of Large Research, Development, and Innovations Infrastructures" program (LM2015051, CZ.02.1.01/0.0/0.0/16_013/0001761, LM2015047, LM2015055).

REFERENCES

(1) The UniProt Consortium. UniProt: A Hub for Protein Information. *Nucleic Acids Res.* **2015**, *43*, D204−D212.10.1093/nar/gku989

(2) Schnoes, A. M.; Brown, S. D.; Dodevski, I.; Babbitt, P. C. Annotation Error in Public Databases: Misannotation of Molecular Function in Enzyme Superfamilies. *PLoS Comput. Biol.* **2009**, *5*, e1000605.

(3) Check Hayden, E. The $1,000 Genome 2006. *Nature* **2014**, *507*, 294−295.

(4) Khosla, C. Quo Vadis, Enzymology? *Nat. Chem. Biol.* **2015**, *11*, 438−441.

(5) Furnham, N.; Garavelli, J. S.; Apweiler, R.; Thornton, J. M. Missing in Action: Enzyme Functional Annotations in Biological Databases. *Nat. Chem. Biol.* **2009**, *5*, 521−525.

(6) Bornscheuer, U. T. Protein Engineering: Beating the Odds. *Nat. Chem. Biol.* **2016**, *12*, 54−55.

(7) Colin, P.-Y.; Kintses, B.; Gielen, F.; Miton, C. M.; Fischer, G.; Mohamed, M. F.; Hyvönen, M.; Morgavi, D. P.; Janssen, D. B.; Hollfelder, F. Ultrahigh-Throughput Discovery of Promiscuous Enzymes by Picodroplet Functional Metagenomics. *Nat. Commun.* **2015**, *6*, 10008.

(8) Chen, B.; Lim, S.; Kannan, A.; Alford, S. C.; Sunden, F.; Herschlag, D.; Dimov, I. K.; Baer, T. M.; Cochran, J. R. High-Throughput Analysis and Protein Engineering Using Microcapillary Arrays. *Nat. Chem. Biol.* **2016**, *12*, 76−81.

(9) Dörr, M.; Fibinger, M. P. C.; Last, D.; Schmidt, S.; Santos-Aberturas, J.; Böttcher, D.; Hummel, A.; Vickers, C.; Voss, M.; Bornscheuer, U. T. Fully Automatized High-Throughput Enzyme Library Screening Using a Robotic Platform. *Biotechnol. Bioeng.* **2016**, *113*, 1421−1432.

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

(10) Mak, W. S.; Tran, S.; Marcheschi, R.; Bertolani, S.; Thompson, J.; Baker, D.; Liao, J. C.; Siegel, J. B. Integrative Genomic Mining for Enzyme Function to Enable Engineering of a Non-Natural Biosynthetic Pathway. *Nat. Commun.* **2015**, *6*, 1−9.
(11) Lobb, B.; Doxey, A. C. Novel Function Discovery through Sequence and Structural Data Mining. *Curr. Opin. Struct. Biol.* **2016**, *38*, 53−61.
(12) Studer, R. A.; Dessailly, B. H.; Orengo, C. A. Residue Mutations and Their Impact on Protein Structure and Function: Detecting Beneficial and Pathogenic Changes. *Biochem. J.* **2013**, *449*, 581−594.
(13) Jesenska, A.; Pavlova, M.; Strouhal, M.; Chaloupkova, R.; Tesinska, I.; Monincova, M.; Prokop, Z.; Bartos, M.; Pavlik, I.; Rychlik, I.; Möbius, P.; Nagata, Y.; Damborsky, J. Cloning, Biochemical Properties, and Distribution of Mycobacterial Haloalkane Dehalogenases. *Appl. Environ. Microbiol.* **2005**, *71*, 6736−6745.
(14) Sato, Y.; Monincova, M.; Chaloupkova, R.; Prokop, Z.; Ohtsubo, Y.; Minamisawa, K.; Tsuda, M.; Damborsky, J.; Nagata, Y. Two Rhizobial Strains, *Mesorhizobium loti* MAFF303099 and *Bradyrhizobium japonicum* USDA110, Encode Haloalkane Dehalogenases with Novel Structures and Substrate Specificities. *Appl. Environ. Microbiol.* **2005**, *71*, 4372−4379.
(15) Jesenska, A.; Bartos, M.; Czernekova, V.; Rychlik, I.; Pavlik, I.; Damborsky, J. Cloning and Expression of the Haloalkane Dehalogenase Gene *dhmA* from *Mycobacterium avium* N85 and Preliminary Characterization of DhmA. *Appl. Environ. Microbiol.* **2002**, *68*, 3724−3730.
(16) Jesenska, A.; Monincova, M.; Koudelakova, T.; Hasan, K.; Chaloupkova, R.; Prokop, Z.; Geerlof, A.; Damborsky, J. Biochemical Characterization of Haloalkane Dehalogenases DrbA and DmbC, Representatives of a Novel Subfamily. *Appl. Environ. Microbiol.* **2009**, *75*, 5157−5160.
(17) Li, A.; Shao, Z. Biochemical Characterization of a Haloalkane Dehalogenase DadB from *Alcanivorax dieselolei* B-5. *PLoS One* **2014**, *9*, e89144−89153.
(18) Hesseler, M.; Bogdanović, X.; Hidalgo, A.; Berenguer, J.; Palm, G. J.; Hinrichs, W.; Bornscheuer, U. T. Cloning. Functional Expression, Biochemical Characterization, and Structural Analysis of a Haloalkane Dehalogenase from *Plesiocystis pacifica* SIR-1. *Appl. Microbiol. Biotechnol.* **2011**, *91*, 1049−1060.
(19) Hasan, K.; Fortova, A.; Koudelakova, T.; Chaloupkova, R.; Ishitsuka, M.; Nagata, Y.; Damborsky, J.; Prokop, Z. Biochemical Characteristics of the Novel Haloalkane Dehalogenase DatA, Isolated from the Plant Pathogen *Agrobacterium tumefaciens* C58. *Appl. Environ. Microbiol.* **2011**, *77*, 1881−1884.
(20) Fung, H. K. H.; Gadd, M. S.; Drury, T. a.; Cheung, S.; Guss, J. M.; Coleman, N. V.; Matthews, J. M. Biochemical and Biophysical Characterisation of Haloalkane Dehalogenases DmrA and DmrB in *Mycobacterium* Strain JS60 and Their Role in Growth on Haloalkanes. *Mol. Microbiol.* **2015**, *97*, 439−453.
(21) Fortova, A.; Sebestova, E.; Stepankova, V.; Koudelakova, T.; Palkova, L.; Damborsky, J.; Chaloupkova, R. DspA from *Strongylocentrotus purpuratus*: The First Biochemically Characterized Haloalkane Dehalogenase of Non-Microbial Origin. *Biochimie* **2013**, *95*, 2091−2096.
(22) Prokop, Z.; Sato, Y.; Brezovsky, J.; Mozga, T.; Chaloupkova, R.; Koudelakova, T.; Jerabek, P.; Stepankova, V.; Natsume, R.; Van Leeuwen, J. G. E.; Janssen, D. B.; Florian, J.; Nagata, Y.; Senda, T.; Damborsky, J. Enantioselectivity of Haloalkane Dehalogenases and Its Modulation by Surface Loop Engineering. *Angew. Chem.* **2010**, *122*, 6247−6251.
(23) Schober, M.; Faber, K. Inverting Hydrolases and Their Use in Enantioconvergent Biotransformations. *Trends Biotechnol.* **2013**, *31*, 468−478.
(24) Swanson, P. E. Dehalogenases Applied to Industrial-Scale Biocatalysis. *Curr. Opin. Biotechnol.* **1999**, *10*, 365−369.
(25) Dvorak, P.; Bidmanova, S.; Damborsky, J.; Prokop, Z. Immobilized Synthetic Pathway for Biodegradation of Toxic Recalcitrant Pollutant 1,2,3-Trichloropropane. *Environ. Sci. Technol.* **2014**, *48*, 6859−6866.
(26) Dravis, B. C.; Lejeune, K. E.; Hetro, A. D.; Russell, A. J. Enzymatic Dehalogenation of Gas Phase Substrates with Haloalkane Dehalogenase. *Biotechnol. Bioeng.* **2000**, *69*, 235−241.
(27) Lal, R.; Pandey, G.; Sharma, P.; Kumari, K.; Malhotra, S.; Pandey, R.; Raina, V.; Kohler, H. E.; Holliger, C.; Jackson, C.; Oakeshott, J. G. Biochemistry of Microbial Degradation of Hexachlorocyclohexane and Prospects for Bioremediation. *Microbiol. Mol. Biol. Rev.* **2010**, *74*, 58−80.
(28) Prokop, Z.; Oplustil, F.; DeFrank, J.; Damborsky, J. Enzymes Fight Chemical Weapons. *Biotechnol. J.* **2006**, *1*, 1370−1380.
(29) Bidmanova, S.; Steiner, M.-S.; Stepan, M.; Vymazalova, K.; Gruber, M. A.; Duerkop, A.; Damborsky, J.; Prokop, Z.; Wolfbeis, O. S. Enzyme-Based Test Strips for Visual or Photographic Detection and Quantitation of Gaseous Sulfur Mustard. *Anal. Chem.* **2016**, *88*, 6044−6049.
(30) Bidmanova, S.; Chaloupkova, R.; Damborsky, J.; Prokop, Z. Development of an Enzymatic Fiber-Optic Biosensor for Detection of Halogenated Hydrocarbons. *Anal. Bioanal. Chem.* **2010**, *398*, 1891−1898.
(31) Bidmanova, S.; Kotlanova, M.; Rataj, T.; Damborsky, J.; Trtilek, M.; Prokop, Z. Fluorescence-Based Biosensor for Monitoring of Environmental Pollutants: From Concept to Field Application. *Biosens. Bioelectron.* **2016**, *84*, 97−105.
(32) Campbell, D. W.; Müller, C.; Reardon, K. F. Development of a Fiber Optic Enzymatic Biosensor for 1,2-Dichloroethane. *Biotechnol. Lett.* **2006**, *28*, 883−887.
(33) Los, G. V.; Encell, L. P.; Mcdougall, M. G.; Hartzell, D. D.; Karassina, N.; Zimprich, C.; Wood, M. G.; Learish, R.; Ohana, R. F.; Urh, M.; Simpson, D.; Mendez, J.; Zimmerman, K.; Otto, P.; Vidugiris, G.; Zhu, J.; Darzins, A.; Klaubert, D. H.; Bulleit, R. F.; Wood, K. V. HaloTag: A Novel Protein Labeling Technology for Cell Imaging and Protein Analysis. *ACS Chem. Biol.* **2008**, *3*, 373−382.
(34) Ohana, R. F.; Encell, L. P.; Zhao, K.; Simpson, D.; Slater, M. R.; Urh, M.; Wood, K. V. HaloTag7: A Genetically Engineered Tag That Enhances Bacterial Expression of Soluble Proteins and Improves Protein Purification. *Protein Expression Purif.* **2009**, *68*, 110−120.
(35) Koudelakova, T.; Bidmanova, S.; Dvorak, P.; Pavelka, A.; Chaloupkova, R.; Prokop, Z.; Damborsky, J. Haloalkane Dehalogenases: Biotechnological Applications. *Biotechnol. J.* **2013**, *8*, 32−45.
(36) Chovancova, E.; Kosinski, J.; Bujnicki, J. M.; Damborsky, J. Phylogenetic Analysis of Haloalkane Dehalogenases. *Proteins: Struct., Funct., Genet.* **2007**, *67*, 305−316.
(37) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997**, *25*, 3389−3402.
(38) N.R. Coordinators. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2013**, *41*, D8−D20.
(39) Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T. J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; Thompson, J. D.; Higgins, D. G. Fast, Scalable Generation of High-Quality Protein Multiple Sequence Alignments Using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539.
(40) Sali, A.; Blundell, T. L. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J. Mol. Biol.* **1993**, *234*, 779−815.
(41) Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: Computed Atlas of Surface Topography of Proteins with Structural and Topographical Mapping of Functionally Annotated Residues. *Nucleic Acids Res.* **2006**, *34*, W116−W118.
(42) Liang, J.; Edelsbrunner, H.; Woodward, C. Anatomy of Protein Pockets and Cavities: Measurement of Binding Site Geometry and Implications for Ligand Design. *Protein Sci.* **1998**, *7*, 1884−1897.
(43) Chovancova, E.; Pavelka, A.; Benes, P.; Strnad, O.; Brezovsky, J.; Kozlikova, B.; Gora, A.; Sustr, V.; Klvana, M.; Medek, P.; Biedermannova, L.; Sochor, J.; Damborsky, J. CAVER 3.0: A Tool for the Analysis of Transport Pathways in Dynamic Protein Structures. *PLoS Comput. Biol.* **2012**, *8*, e1002708.
(44) Hanwell, M. D.; Curtis, D. E.; Lonie, D. C.; Vandermeersch, T.; Zurek, E.; Hutchison, G. R. Avogadro: An Advanced Semantic

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

Chemical Editor, Visualization, and Analysis Platform. *J. Cheminf.* **2012**, *4*, 17.
(45) Solis, F. J.; Wets, R. J.-B. Mnimization by Random Search Techniques. *Math. Oper. Res.* **1981**, *6*, 19−30.
(46) Harrison, R. G. Expression of Soluble Heterologous Proteins via Fusion with NusA Protein. *InNovations* **2000**, *11*, 4−7.
(47) Wilkinson, D. L.; Harrison, R. G. Predicting the Solubility of Recombinant Proteins in *Escherichia coli*. *Nat. Biotechnol.* **1991**, *9*, 443−448.
(48) Woody, R. W. *Circular Dichroism and the Conformational Analysis of Biomolecules*, 1st ed.; Fasman, G. D., Ed.; Plenum Press: New York, 1996.
(49) Gräslund, S.; Nordlund, P.; Weigelt, J.; Hallberg, B. M.; Bray, J.; Gileadi, O.; Knapp, S.; Oppermann, U.; Arrowsmith, C.; Hui, R.; Ming, J.; et al. Protein Production and Purification. *Nat. Methods* **2008**, *5*, 135−146.
(50) Braun, P.; LaBaer, J. High Throughput Protein Production for Functional Proteomics. *Trends Biotechnol.* **2003**, *21*, 383−388.
(51) Pacheco, B.; Crombet, L.; Loppnau, P.; Cossar, D. A Screening Strategy for Heterologous Protein Expression in *Escherichia coli* with the Highest Return of Investment. *Protein Expression Purif.* **2012**, *81*, 33−41.
(52) Bastard, K.; Smith, A. A. T.; Vergne-Vaxelaire, C.; Perret, A.; Zaparucha, A.; De Melo-Minardi, R.; Mariage, A.; Boutard, M.; Debard, A.; Lechaplais, C.; Pelle, C.; Pellouin, V.; Perchat, N.; Petit, J.-L.; Kreimeyer, A.; Medigue, C.; Weissenbach, J.; Artiguenave, F.; De Berardinis, V.; Vallenet, D.; Salanoubat, M. Revealing the Hidden Functional Diversity of an Enzyme Family. *Nat. Chem. Biol.* **2014**, *10*, 42−49.

DOI: 10.1021/acscatal.7b03523
ACS Catal. 2018, 8, 2402−2412

Articles from Section II: Engineering bacteria for utilization and valorization of substrates from lignocellulosic residues

[11] Dvořák, P. and de Lorenzo, V. (2018) Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. *Metabolic Engineering*. *48*, 94–108. (2018 IF = 8.115)

[12] Dvořák, P.*, Kováč, J., and de Lorenzo, V. (2020) Biotransformation of D-xylose to D-xylonate coupled to medium-chain-length polyhydroxyalkanoate production in cellobiose-grown *Pseudomonas putida* EM42. *Microbial Biotechnology*. *13*, 1273-83. (*corresponding author, 2020 IF = 4.181)

[13] Dvořák, P.*, Bayer, E.A., and de Lorenzo, V. (2020) Surface display of designer protein scaffolds on genome-reduced strains of *Pseudomonas putida*. *ACS Synthetic Biology*. *9*, 2749–64. (*co-corresponding author, 2020 IF = 5.229)

[14] Espeso, D.R., **Dvořák, P.,** Aparicio, T., and de Lorenzo, V. (2020) An automated DIY framework for experimental evolution of *Pseudomonas putida*. *Microbial Biotechnology*. *14,* 2679-2685. (2020 IF = 4.181)

[15] Dvořák, P., Alvarez-Carreño, C., Ciordia, S., Paradela, A., and de Lorenzo, V. (2021) An updated structural model of the A domain of the *Pseudomonas putida* XylR regulator poses an atypical interplay with aromatic effectors. *Environmental Microbiology*. *23*, 4418-4433. (2021 IF = 5.491)

Metabolic Engineering 48 (2018) 94–108

Contents lists available at ScienceDirect

Metabolic Engineering

journal homepage: www.elsevier.com/locate/meteng

Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose

Pavel Dvořák, Víctor de Lorenzo*

Systems and Synthetic Biology Program, Centro Nacional de Biotecnología CNB-CSIC, Cantoblanco, 28049 Madrid, Spain

ARTICLE INFO

Keywords:
Pseudomonas putida
Cellobiose
Xylose
Metabolism
bglC
xylABE

ABSTRACT

Given its capacity to tolerate stress, NAD(P)H/ NAD(P) balance, and increased ATP levels, the platform strain *Pseudomonas putida* EM42, a genome-edited derivative of the soil bacterium *P. putida* KT2440, can efficiently host a suite of harsh reactions of biotechnological interest. Because of the lifestyle of the original isolate, however, the nutritional repertoire of *P. putida* EM42 is centered largely on organic acids, aromatic compounds and some hexoses (glucose and fructose). To enlarge the biochemical network of *P. putida* EM42 to include disaccharides and pentoses, we implanted heterologous genetic modules for D-cellobiose and D-xylose metabolism into the enzymatic complement of this strain. Cellobiose was actively transported into the cells through the ABC complex formed by native proteins PP1015-PP1018. The knocked-in β-glucosidase BglC from *Thermobifida fusca* catalyzed intracellular cleavage of the disaccharide to D-glucose, which was then channelled to the default central metabolism. Xylose oxidation to the dead end product D-xylonate was prevented by deleting the *gcd* gene that encodes the broad substrate range quinone-dependent glucose dehydrogenase. Intracellular intake was then engineered by expressing the *Escherichia coli* proton-coupled symporter XylE. The sugar was further metabolized by the products of *E. coli xylA* (xylose isomerase) and *xylB* (xylulokinase) towards the pentose phosphate pathway. The resulting *P. putida* strain co-utilized xylose with glucose or cellobiose to complete depletion of the sugars. These results not only show the broadening of the metabolic capacity of a soil bacterium towards new substrates, but also promote *P. putida* EM42 as a platform for plug-in of new biochemical pathways for utilization and valorization of carbohydrate mixtures from lignocellulose processing.

1. Introduction

Due to the physicochemical stresses that prevail in the niches in which the soil bacterium *Pseudomonas putida* thrives (it is typically abundant in sites contaminated by industrial pollutants), this microorganism is endowed with a large number of traits desirable in hosts of harsh biotransformations of industrial interest (Dvořák et al., 2017; Nikel et al., 2014). The *P. putida* strain KT2440 is a saprophytic, non-pathogenic, GRAS-certified (Generally Recognized as Safe) bacterium; as the most thoroughly characterized laboratory pseudomonad, it has an expanding catalogue of available systems and synthetic biology tools (Aparicio et al., 2017; Elmore et al., 2017; Martínez-García and de Lorenzo, 2017, p.; Nogales et al., 2017). This bacterium is becoming a laboratory workhorse as well as a valued cell factory (Benedetti et al., 2016; Loeschcke and Thies, 2015; Nikel and de Lorenzo, 2013; Poblete-Castro et al., 2012). Its high resistance to endogenous and exogenous stresses makes it tolerant to industrially relevant chemicals (*e.g.,* ethanol, *p*-coumaric acid, toluene) and to by-products of biomass hydrolysis (furfural, 5-(hydroxymethyl)furfural, benzoate, acetic acid) at concentrations that are inhibitory to other microbial platforms, including *Escherichia coli* (Calero et al., 2017; Guarnieri et al., 2017; Johnson and Beckham, 2015; Nikel and de Lorenzo, 2014). The nutritional landscape of typical *P. putida* niches (plant rhizosphere, polluted soil) has pushed its metabolic specialization towards aromatic compounds (Jiménez et al., 2002) and organic acids (Dos Santos et al., 2004). The very few carbohydrates on which *P. putida* KT2440 can grow are confined to some hexoses (glucose and fructose), with an inability to metabolize disaccharides or 5-carbon sugars productively (Nogales et al., 2017; Puchałka et al., 2008; Rojo, 2010). This limits the options for its use as a platform to process the carbohydrate products of cellulosic and hemicellulosic wastes.

Lignocellulose can be decomposed to cellulose (25–55%), hemicellulose (11–50%), and lignin (10–40%), all of which can be further hydrolyzed enzymatically to shorter carbohydrate polymers and oligomers, monomeric sugars, and lignin-derived aromatics (Mosier et al., 2005). When standard commercial cellulosic cocktails are applied, D-

* Correspondence to: Systems and Synthetic Biology Program, Centro Nacional de Biotecnología (CNB-CSIC), Darwin 3, Campus de Cantoblanco, Madrid 28049, Spain.
E-mail address: vdlorenzo@cnb.csic.es (V. de Lorenzo).

https://doi.org/10.1016/j.ymben.2018.05.019
Received 21 February 2018; Received in revised form 15 May 2018; Accepted 31 May 2018
Available online 02 June 2018
1096-7176/ © 2018 Published by Elsevier Inc. on behalf of International Metabolic Engineering Society.

glucose and D-xylose are two major monomeric products of pretreated plant biomass hydrolysis (Taha et al., 2016). A number of microorganisms have been engineered to use these sugars as substrates for the biomanufacturing of value-added chemicals (Kawaguchi et al., 2016). Cellodextrins, including D-cellobiose, are the most abundant by-products of cellulose saccharification and predominate following partial hydrolysis (Chen, 2015; Singhania et al., 2013). Well-defined, industrially relevant microorganisms with efficient cellobiose metabolism are thus desirable (Kawaguchi et al., 2016; Parisutham et al., 2017; Taha et al., 2016). Attempts to engineer bacteria for cellobiose utilization and valorization included periplasmic or extracellular β-glucosidase expression (Chen et al., 2011; Rutter et al., 2013), or its surface display (Muñoz-Gutiérrez et al., 2014). This last approach was also recently used for *P. putida* KT2440, which does not utilize cellobiose, but surface display of three cellulases, including β-glucosidase BglA from *Clostridium thermocellum*, resulted in only trace conversion of cellulosic substrate into glucose and no growth (Tozakidis et al., 2016). Alternative strategies included intracellular assimilation of cellobiose *via* endogenous or exogenous β-glucosidase (Ha et al., 2011; Vinuselvi and Lee, 2011), and *via* phosphorolysis catalyzed by cellobiose phosphorylase (Shin et al., 2014). Despite the need for an efficient cellobiose transporter, intracellular assimilation is generally regarded as more efficient, because it circumvents carbon catabolite repression (CCR) and prevents cellobiose inhibition of extracellular cellulases in the bioreactor (Ha et al., 2011; Parisutham et al., 2017; Teugjas and Väljamäe, 2013).

Co-consumption of cellobiose with hemicellulose-derived xylose is even more appealing and has been established in several microorganisms (Lynd et al., 2002; Stephanopoulos, 2007; Ha et al., 2011; Lane et al., 2015; Lee et al., 2016; Shin et al., 2014) Some *Pseudomonas* species can use xylose as a C source for growth (Liu et al., 2015) but *P. putida* KT2440 can not (Le Meur et al., 2012; Nogales et al., 2017; Puchałka et al., 2008). Both KT2440 and *P. putida* S12, another pseudomonad with biotechnological potential, were nonetheless engineered successfully for xylose utilization by implementation of the isomerase pathway from *E. coli* (Le Meur et al., 2012; Meijnen et al., 2008). This route fuels the pentose phosphate pathway *via* the action of xylose isomerase (EC 5.3.1.5) and xylulokinase (EC 2.7.1.17), which convert D-xylose to D-xylulose-5-phosphate. However, *P. putida* KT2440 has not been engineered yet for co-consumption of glucose or cellobiose with xylose. In industrially relevant microorganisms such as *E. coli*, *S. cerevisiae* or *Z. mobilis*, co-utilization of lignocellulose-derived sugars is hindered by complex CCR mechanisms that prioritize glucose from other carbon sources (Görke and Stülke, 2008; Kayikci and Nielsen, 2015). These mechanisms must be circumvented by mutagenesis or introduction of heterologous catabolic routes and sugar transporters (Kim et al., 2015; Lawford and Rousseau, 2002). Glucose metabolism in *P. putida* is not as central as it is in *E. coli* (Rojo, 2010). We thus anticipated that *P. putida* recombinants empowered with suitable enzymes and transporters of exogenous origin could co-utilize cellobiose and glucose with xylose without restrictions.

Here we sought to engineer *P. putida* for efficient growth on cellobiose and to test its ability to co-metabolize this disaccharide with xylose. We combined a metabolic engineering approach (Fig. 1) with the *P. putida* KT2440-derived strain *P. putida* EM42 (Martínez-García et al., 2014b). This strain has a streamlined genome (300 genes, ~4.3% of the genome deleted) that results in improved physiological properties compared to *P. putida* KT2440 including lower sensitivity to oxidative stress, increased growth rates, and enhanced expression of heterologous genes (Lieder et al., 2015; Martínez-García et al., 2014b). We show that recruitment of one heterologous β-glucosidase for intracellular cellobiose hydrolysis, and the implementation of the *E. coli* xylose isomerase route and xylose transporter sufficed to cause the co-utilization of cellobiose and xylose by *P. putida* EM42 with inactivated glucose dehydrogenase while maintaining its ability to use glucose. We also demonstrate that *P. putida* metabolism generates more ATP when cells are grown on cellobiose instead of glucose. This study expands the catalytic scope of *P. putida* towards utilization of major components of all three lignocellulose-derived fractions. Moreover, given that the cellobiose is a by-product also of standard cellulose saccharification, which would remain untouched in the sugar mix (due to the inability of most microorganisms to assimilate it), our results demonstrate rational tailoring of an industrially relevant microbial host to achieve a specific step in such a biotechnological value chain.

2. Materials and methods

2.1. Bacterial strains, plasmids, and growth conditions

All bacterial strains and plasmids used in this study are listed in Table 1. *Escherichia coli* strains used for cloning or triparental mating were routinely grown in lysogeny broth (LB; 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl) with agitation (170 rpm) at 37 °C. Chloramphenicol (Cm, 30 μg mL^{-1}) was supplemented to the medium with *E. coli* helper strain HB101. *Pseudomonas putida* recombinants were routinely pre-cultured overnight in 2.5 mL of LB medium with agitation of 300 rpm (Heidolph Unimax 1010 and Heidolph Incubator 1000; Heidolph Instruments, Germany) at 30 °C. For initial tests of expression of heterologous genes in *P. putida*, cells were transferred to 25 mL of fresh LB medium in Erlenmeyer flask and cultivated as described in Section 2.4. For the growth experiments with different carbohydrates, overnight culture was spun by centrifugation (4000 g, RT, 5 min), washed with M9 minimal medium (per 1 L: 4.25 g Na_2HPO_4 $2H_2O$, 1.5 g KH_2PO_4, 0.25 g NaCl, 0.5 g NH_4Cl) added with $MgSO_4$ to the final concentration of 2 mM, and with 2.5 mL L^{-1} trace element solution (Abril et al., 1989). Thiamine HCl (1 mM) was added to the minimal medium for cultures with *E. coli* recombinants. Cells were resuspended to OD_{600} of 0.1 in 25 mL of the same medium with kanamycin (Km, 50 μg mL^{-1}), in case of recombinants with pSEVA2213 or pSEVA238 plasmid, or streptomycin (Sm, 60 μg mL^{-1}), in case of *P. putida* EM42 Δ*gcd bglC*, and with carbon source (glucose, xylose, or cellobiose) of concentration defined in the text or respective figure caption. All used solid media (LB and M9) contained 15 g L^{-1} agar. M9 solid media were prepared with 2 mM $MgSO_4$, 2.5 mL L^{-1} trace element solution (Abril et al., 1989) and 0.2% (w/v) citrate, 0.4% xylose, or 0.4% cellobiose used as a sole carbon source.

2.2. Plasmid and strain constructions

DNA was manipulated using standard laboratory protocols (Sambrook and Russell, 2001). Genomic DNA was isolated using GenElute bacterial genomic DNA kit (Sigma-Aldrich, USA). Plasmid DNA was isolated with QIAprep Spin Miniprep kit (Qiagen, USA). The oligonucleotide primers used in this study (Table S1) were purchased from Sigma-Aldrich (USA). The genes of interest were amplified by polymerase chain reaction (PCR) using Q5 high fidelity DNA polymerase (New England BioLabs, USA) according to the manufacturer's protocol. The reaction mixture (50 μL) further contained polymerase HF or GC buffer (New England BioLabs, USA), dNTPs mix (0.2 mM each; Roche, Switzerland), respective primers (0.5 mM each), water, template DNA, and DMSO. GC buffer and DMSO were used for amplification of genes from *P. putida*. PCR products were purified with NucleoSpin Gel and PCR Clean-up (Macherey-Nagel, Germany). DNA concentration was measured with NanoVue spetrophotometer (GE Healthcare, USA). Colony PCR was performed using 2 × PCR Master Mix solution of Taq DNA polymerase, dNTPs and reaction buffer (Promega, USA). All used restriction enzymes were from New England BioLabs (USA). Digested DNA fragments were ligated using Quick Ligation kit (New England BioLabs, USA). PCR products and digested plasmids separated by DNA electrophoresis with 0.8% (w/v) agarose gels were visualized using Molecular Imager VersaDoc (Bio-Rad, USA). Plasmid constructs were confirmed by DNA sequencing (Macrogen, South Korea).

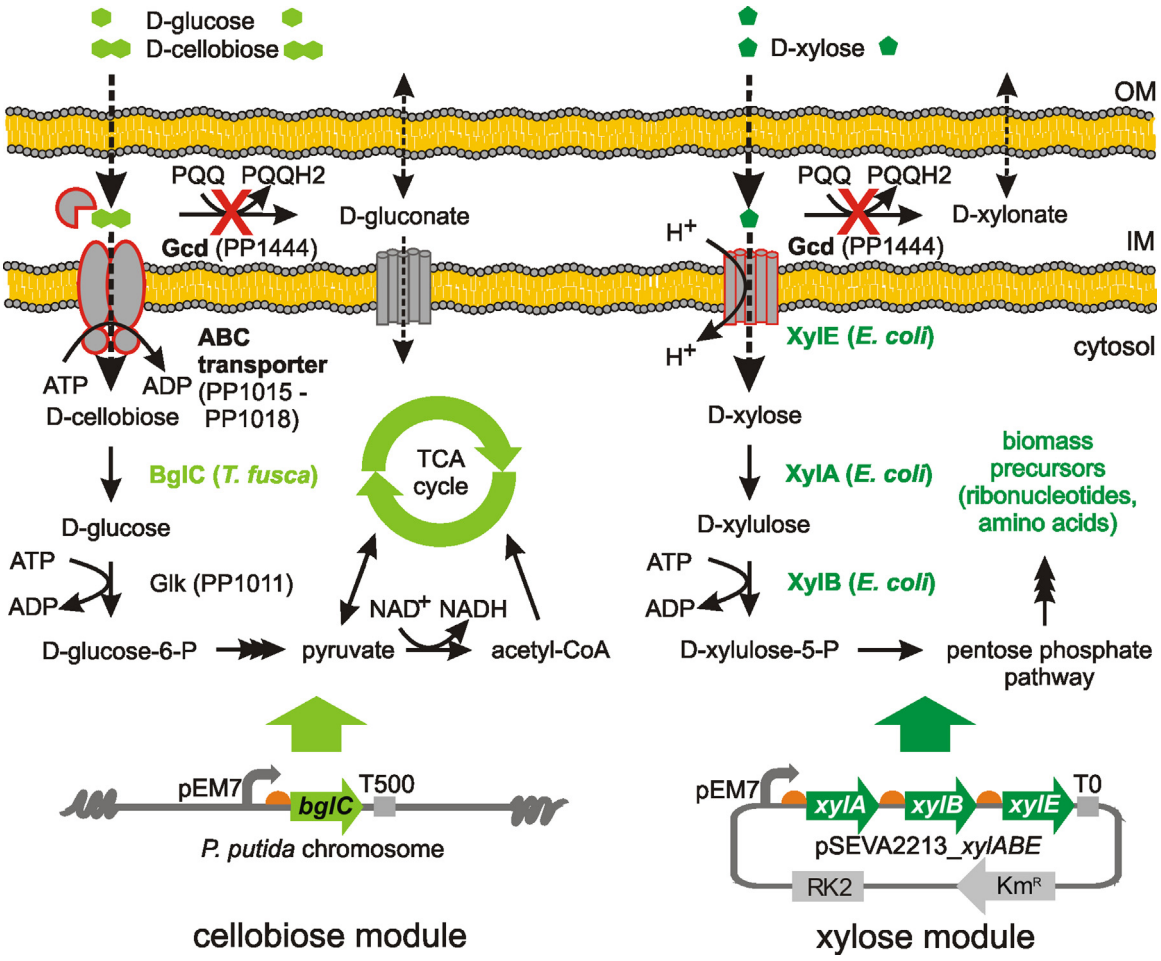


Fig. 1. Engineering *Pseudomonas putida* EM42 for co-utilization of D-cellobiose and D-glucose with D-xylose. Cellobiose metabolism in *P. putida* was established by implantating the β-glucosidase BglC from *Thermobifida fusca*. Experiments suggested that the transport pathways for glucose are important also for cellobiose uptake in *P. putida* cells. These pathways are (i) the ATP-dependent ABC transporter, and (ii) the peripheral oxidative route, which starts with periplasmic conversion of glucose to intermediate D-gluconate through the action of the membrane-bound glucose dehydrogenase Gcd. Xylose metabolism in *P. putida* was established by implantating of xylose isomerase XylA, xylulokinase XylB, and xylose-proton symporter XylE from *E. coli*. For the purpose of cellobiose and xylose co-utilization, an expression cassette with the *bglC* gene was inserted into the chromosome of *P. putida* EM42 Δ*gcd*, while the synthetic *xylABE* operon was expressed from the low copy pSEVA2213 plasmid under the constitutive pEM7 promoter. The EM42 Δ*gcd* mutant was used to avoid xylose conversion by glucose dehydrogenase to dead-end product D-xylonate. PQQ, pyrroloquinoline quinone; Glk, glucokinase; TCA cycle, tricarboxylic acid cycle; OM, (outer membrane); IM (inner membrane).

Chemocompetent *E. coli* Dh5α cells were transformed with ligation mixtures or complete plasmids and individual clones selected on LB agar plates with Km (50 μg mL^{-1}) were used for preparation of glycerol (20% w/v) stocks. Constructed plasmids were transferred from *E. coli* Dh5α donor to *P. putida* EM42 by tripartite mating, using *E. coli* HB101 helper strain with pRK600 plasmid (Table 1). Alternatively, electroporation (2.5 kV, 4 – 5 ms pulse) was used for transformation of *P. putida* cells with selected plasmids using a MicroPulser electroporator and Gene Pulser Cuvettes with 0.2 cm gap (Bio-Rad, USA). Preparation of *P. putida* electrocompetent cells and electroporation procedure was performed as described elsewhere (Aparicio et al., 2015). *P. putida* transconjugants or transformants were selected on M9 agar plates with citrate or LB agar plates, respectively, with Km (50 μg mL^{-1}) at 30 °C overnight.

Construction of cellobiose metabolism module. The *ccel_2454* gene encoding β-glucosidase (EC 3.2.1.21) from *Clostridium cellulolyticum* was synthesized together with consensus ribosome binding site (RBS; GeneArt/Thermo Fisher Scientific, Germany) and subcloned from delivery vector pMA_*ccel2454* into pSEVA238 upon digestion with *Kpn*I and *Pst*I resulting in pSEVA238_*ccel2454*. The β-glucosidase encoding *bglX* gene was PCR amplified from the genomic DNA of *P. putida* KT2440 using bglX fw and bglX rv primers, digested with *Nde*I and *Hind*III and cloned into corresponding restriction sites of modified pSEVA238 resulting in pSEVA238_*bglX*. The *Nde*I site and a consensus RBS were previously introduced into the standard SEVA polylinker of pSEVA238 (unpublished plasmid). The *bglC* gene encoding β-glucosidase from *Thermobifida fusca* with N-terminal 6xHis tag was subcloned from pET21a_*bglC* construct into *Nde*I and *Hind*III restriction sites of modified pSEVA238. The *bglC* gene with the RBS and His tag was subsequently PCR amplified using primers bglC fw and bglC rv, the PCR product was cut with *Sac*I and *Pst*I and subcloned into pSEVA2213 giving rise to pSEVA2213_*bglC*.

Insertion of bglC gene into P. putida chromosome. The *bglC* gene with consensus RBS was subcloned into *Sac*I and *Pst*I sites of mini-Tn5-vector pBAMD1–4 (Martínez-García et al., 2014a). Original pBAMD1–4 plasmid was endowed with pEM7 promoter subcloned into *Avr*II and *Eco*RI sites. *P. putida* EM42 Δgcd cells (100 μL) were electroporated with plasmid DNA (200 ng) and recovered for 7 h in 5 mL of modified Terrific Broth (TB) medium (yeast extract 24 g L^{-1}, tryptone 20 g L^{-1}, KH_2PO_4 0.017 M, K_2HPO_4 0.072 M) at 30 °C with shaking (170 rpm). Cells were collected by centrifugation (4000 rpm, 10 min) and re-suspended in 100 mL of selection M9 medium with 5 g L^{-1} cellobiose and streptomycin (50 μg mL^{-1}). After four days of incubation at 30 °C with shaking (170 rpm), cells were spun (4000 rpm, 15 min) and plated

Table 1
Strains and plasmids used in this study.

Strain or plasmid	Characteristics	Source or reference
Escherichia coli		
Dh5α	Cloning host: F-λ-*endA1 glnX44*(AS) *thiE1 recA1 relA1 spoT1 gyrA96*(NalR) *rfbC1 deoR nupG* Φ80(*lacZ*ΔM15) Δ(*argF-lac*)*U169 hsdR17*($r_K^- m_K^+$)	Grant et al. (1990)
CC118λ*pir*	Cloning host: *araD139* Δ(*ara-leu*)*7697* Δ*lacX74 galE galK phoA20 thi*− *1 rpsE rpoB*(Rif^R) *argE*(Am) *recA1*, λpir lysogen	Herrero et al. (1990)
HB101	Helper strain for tri-parental mating: F– λ– *hsdS20*(r_B^- m_B^-) *recA13 leuB6*(Am) *araC14* Δ(*gpt-proA*)62 *lacY1 galK2*(OC) *xyl*− *5 mtl*− *1 thiE1 rpsL20*(Sm^R) *glnX44*(AS)	Boyer and Roulland-Dussoix (1969)
Pseudomonas putida		
KT2440	Wild-type strain, spontaneous restriction-deficient derivative of strain mt − 2 cured of the TOL plasmid pWW0	Bagdasarian et al. (1981)
EM42	KT2440 derivative: Δprophages1,2,3,4 ΔTn7 Δ*endA1* Δ*endA2* Δ*hsdRMS* Δflagellum ΔTn*4652*	Martínez-García et al. (2014b)
EM42 Δ*gts*	EM42 with scarless deletion of *gtsABCD* operon (PP_1015-PP_1018) encoding glucose ABC transporter	This study
EM42 Δ*gcd*	EM42 with scarless deletion of *gcd* gene (PP_PP1444) encoding periplasmic glucose dehydrogenase	This study
EM42 Δ*gts* Δ*gcd*	EM42 with scarless deletions of *gtsABCD* and *gcd*	This study
EM42 Δ*gcd bglC*	EM42 Δ*gcd* with synthetic expression cassette (pEM7 *bglC aadA*, Sm^R) in chromosome	This study
Plasmids		
pRK600	Helper plasmid for tri-parental mating: *oriV*(*ColE1*) RK2 *tra*$^+$ *mob*$^+$, Cm^R	Kessler et al. (1994)
pSEVA238	Expression vector: *oriV*(pBBR1) *xylS-Pm neo*, Km^R	Silva-Rocha et al. (2013)
pSEVA238_*gfp*	pSEVA238 bearing the gene of monomeric superfolder GFP (msfGFP)	SEVA collection
pSEVA2213	Expression vector: *oriV*(RK2) pEM7 *neo*, Km^R	Silva-Rocha et al. (2013)
pBAMD1–4	Mini-Tn5 delivery plasmid: *ori*(R6K) *bla aadA*, Amp^R Sm^R/Sp^R	Martínez-García et al. (2014a)
pEMG	Plasmid for genome editing in Gram-negative bacteria: *ori* (R6K) *neo*, Km^R	Martínez-García and de Lorenzo (2012)
pSW-I	Expression vector with gene of I-SceI, a homing endonuclease from *Saccharomyces cerevisiae*: ori(RK2) *xylS-Pm bla*, Amp^R	Martínez-García and de Lorenzo (2012)
pET21a_*bglC*	Expression vector (Novagen) with *bglC* gene encoding β-glucosidase from *Thermobifida fusca*: *ori*(pBR322) *lacI* T7 *bla*, Amp^R	Moraïs et al. (2012)
pSEVA238_*gfp*C	pSEVA238 with *gfp* gene encoding monomeric superfolder green fluorescent protein (msfGFP) lacking RBS but with TAG STOP codon (*Hind*III/*Spe*I)	This study
pSEVA238_*bglC*	pSEVA238 with *bglC* gene (*Sac*I/*Pst*I)	This study
pSEVA238_*ccel2454*	pSEVA238 with *ccel2454* gene encoding β-glucosidase from *Clostridium cellulolyticum* (*Kpn*I/*Pst*I)	This study
pSEVA238_*bglX*	pSEVA238 with *bglX* gene encoding putative β-glucosidase from *P. putida* KT2440 (*Nde*I/*Hind*III)	This study
pSEVA2213_*bglC*	pSEVA2213 with *bglC* gene (*Sac*I/*Pst*I)	This study
pSEVA2213_*xylAB*	pSEVA2213 with *xylAB* part of *xyl* operon from *E. coli* encoding XylA xylose isomerase and xylB xylulokinase (*Eco*RI/*Bam*HI)	This study
pSEVA2213_*xylABE*	pSEVA2213_*xylAB with xylE* gene from *E. coli* encoding XylE xylose-proton symporter (*Bam*HI/*Hind*III)	This study
pSEVA238_*xylE-gfp*C	pSEVA238_*gfp*C with *xylE* gene subcloned in frame upstream *gfp* (*Bam*HI/*Hind*III); *xylE* lacks STOP codon but has own consensus RBS	This study
pEMG_*gtsABCD*	pEMG with the sequences (~500 bp) spanning upstream and downstream the *gtsABCD* operon in *P. putida KT2440* genome (*Eco*RI/*Bam*HI)	Dr. Alberto Sánchez-Pascuala
pEMG_*gcd*	pEMG with the sequences (~500 bp) spanning upstream and downstream the *gcd* gene in *P. putida KT2440* genome (*Eco*RI/*Bam*HI)	Dr. Alberto Sánchez-Pascuala

Abbreviations: RBS, ribosome binding site; Amp, ampicillin; Km, kanamycin; Sm, streptomycin; Sp, spektinomycin.

on selection M9 agar plates with 5 g L^{-1} cellobiose and streptomycin (50 μg mL^{-1}). Three fastest growing clones were re-streaked on fresh M9 agar plates with streptomycin or with streptomycin (50 μg mL^{-1}) and ampicillin (500 μg mL^{-1}) to rule out insertion of the whole pBAMD1–4 plasmid. The growth of three candidates in liquid minimal medium with cellobiose was verified. The insertion site of expression cassette (pEM7 promoter, *bglC* gene, T500 transcriptional terminator, and *aadA* gene) in chromosome of the fastest growing clone was determined by two-round arbitrary primed PCR with Arb6, Arb2, ME-O-Sm-Ext-F, and ME-O-Sm-Int-F primers (Table S1) following the protocol described before (Martínez-García et al., 2014a). Me-O-Sm-Int-F was used as a sequencing primer for PCR product. Position of *bglC* expression cassette in *P. putida* chromosome was reversly verified by colony PCR with bglC check fw and xerD check rv primers.

Construction of xylose metabolism module. The *xylAB* part of *E. coli xyl* operon encoding xylose isomerase (EC 5.3.1.5) XylA and xylulokinase (EC 2.7.1.17) XylB was amplified from genomic DNA of *E. coli* BL21 (DE3) using xylAB fw and xylAB rv primers. PCR product was digested with *Eco*RI and *Bam*HI and ligated into pSEVA2213 giving rise to pSEVA2213_*xylAB*. The gene of xylose-proton symporter (*xylE*) was amplified from the genomic DNA of *E.coli* BL21 (DE3) using two-step PCR protocol. In the first step, the gene was amplified using xylE fw 1 and xylE rv 1 primers. The sample of the reaction mixture with the PCR product (1 μL) was transferred into the second reaction with xylE fw 2 and xylE rv 2 primers. Final PCR product was digested with *Bam*HI and *Hind*III and cloned downstream *xylAB* operon in pSEVA2213_*xylAB* resulting in pSEVA2213_*xylABE*. For the purpose of construction of the plasmid allowing translational fusion of XylE to monomeric superfolded GFP (msfGFP), *gfp* gene was initially amplified without its own RBS but with STOP codon from pSEVA238_*gfp* plasmid (SEVA collection) using gfpC fw and gfpC rv primers. The PCR product was digested with *Hind*III and *Spe*I and ligated into pSEVA238, cut with the same pair of enzymes, giving rise to pSEVA238_gfpC. The *xylE* gene was amplified from pSEVA2213_*xylABE* with its synthetic RBS but without STOP codon using xylE-gfp fw and xylE-gfp rv primers. The PCR product was digested with *Bam*HI and *Hind*III and cloned upstream the *gfp* gene in pSEVA238_*gfp*C, resulting in pSEVA238_*xylE-gfp*C.

Preparation of deletion mutants of P. putida EM42. Deletion mutants were prepared using the protocol described previously (Aparicio et al., 2015). Briefly, the regions of approximately 500 bp upstream and downstream the *gtsABCD* genes (PP_1015 – PP_1018) were PCR amplified with TS1F-gtsABCD, TS1R-gtsABCD and TS2F-gtsABCD, TS2R-gtsABCD primers, respectively. TS1 and TS2 fragments were joined through SOEing-PCR (Horton et al., 1990), the PCR product was digested with *Eco*RI and *Bam*HI and cloned into non-replicative pEMG plasmid. The resulting pEMG_*gtsABCD* construct was propagated in *E. coli* CC118λ*pir* cells and the whole TS1-TS2 region was sequenced in several clones selected based on results of colony PCR with TS1F-gtsABCD and TS2R-gtsABCD primers (product of about 1 kb expected). The sequence verified pEMG_*gtsABCD* plasmid was transformed into

competent EM42 cells by electroporation. Transformants were selected on LB agar plates with Km and co-integrates were identified by colony PCR with TS1F-gtsABCD and TS2R-gtsABCD primers. The pSW-I plasmid was transformed into selected co-integrate by electroporation. Transformants were plated on LB agar plates with Km and Amp and expression of I-SceI in selected clone inoculated into 5 mL of LB was induced with 1 mM 3-methylbenzoic acid (3MB) overnight. Induced cells were plated on LB agar plates with Amp and the positive clone EM42 Δ*gts* with a loss of Km resistance marker and deletion was confirmed by colony PCR using check(-)gtsABCD fw and TS2R-gtsABCD primers. PCR product size in case of scarless deletion was 1250 bp. The quinoprotein glucose dehydrogenase gene *gcd* (PP_1444) was deleted accordingly in *P. putida* EM42, resulting in *P. putida* EM42 Δ*gcd*, and in *P. putida* EM42 Δ*gts*, resulting in *P. putida* EM42 Δ*gts* Δ*gcd*, using a set of TS primers listed in Table S1. Expression of I-SceI in selected co-integrates was induced with 1 mM 3MB for 6 h. Check(-)gcd fw and check (-)gcd rv primers were used to confirm deletion of *gcd* gene. PCR product size in case of deletion was 1500 bp. *P. putida* recombinants were cured of pSW-I plasmid after several passes in LB medium lacking Amp.

2.3. Calculations of dry cell weight and growth parameters

Biomass was determined as dry cell weight. Samples of cultures grown in M9 minimal medium with 5 g L^{-1} glucose were transferred into 2 mL pre-dried and pre-weighed Eppendorf tubes and pelleted at 13,000 g for 10 min. The pellets were washed twice with distilled water and dried at 80 °C for 48 h. Based on the prepared standard curve, one A_{600} unit is equivalent to 0.38 g L^{-1} of dry cell weight. Specific growth rate (μ) was determined during exponential growth as a slope of the data points obtained by plotting the natural logarithm of A_{600} values against time. Substrate consumption rate (*r*) was determined for initial 12 and 24 h of culture as r = (c substrate at t_0 - c substrate at t_1) / (t_1 - t_0). Biomass yield ($Y_{X/S}$) was determined 24 h after each culture started to grow exponentially as $Y_{X/S}$ = c biomass at t_1 / (c substrate at t_0 - c substrate at t_1). Specific carbon (C) consumption rate (q_s) was determined during exponential growth on glucose or cellobiose as q_s = (mmol C at t_0 - mmol C at t_1) / ((t_1 - t_0) * (g biomass at t_1 - g biomass at t_0)).

2.4. Enzyme activity assays

For initial screening of β-glucosidase activities, 25 mL of LB medium was inoculated from overnight cultures to A_{600} = 0.05 and cells were grown for 3 h at 30 °C with shaking (170 rpm). Expression of β-glucosidase genes from pSEVA238 plasmid was then induced with 1 mM 3MB. After induction, cells were grown in the same conditions for additional 5 h and then harvested by centrifugation (4000 g, 4 °C, 10 min). Cell pellets were lysed by adding 1 mL of BugBuster Protein Exctraction Reagent with 1 μL of Lysonase Bioprocessing Reagent (both from Merck Millipore, USA) for 15 min at RT with slow agitation. For later β-glucosidase activity measurement of BglC in EM42 recombinants, cell lysates were prepared by spinning (21,000 g, 4 °C, 2 min) 4 mL of cells growing in 25 mL of minimal medium with 5 g L^{-1} cellobiose (except for *P. putida* EM42 Δgts Δgcd pSEVA2213_*bglC* recombinant, which was grown in LB medium). Cells were collected in mid log phase (A_{600} = 1.0). Cell pellets were added with 200 μL of BugBuster Protein Exctraction Reagent and 0.2 μL of Lysonase Bioprocessing Reagent and lysed for 15 min at RT with slow agitation. Cell lysates for xylose isomerase and xylulokinase activity determination were prepared by sonicating concentrated cell solutions prepared by spinning (4000 g, 4 °C, 10 min) 25 mL of cells grown in LB medium to A_{600} = 1.0. Cell pellets were washed by 5 mL of ice-cold 50 mM Tris-Cl buffer (pH 7.5) resuspended in 1 mL of the same buffer, placed in ice bath and disrupted by sonication. In all cases, cell lysates were centrifuged at 21,000 g for 30 min at 4 °C and supernatants, termed here as cell-free extracts (CFE), were used for activity determination. Total protein concentration in CFE was measured using the method of Bradford (Bradford, 1976) with a commercial kit (Sigma-Aldrich, USA). Crystalline bovine serum albumin (Sigma-Aldrich, USA) was used as a protein standard.

β-glucosidase activity was measured using synthetic substrate *p*-nitrophenyl-*β*-D-glucopyranoside (pNPG; Sigma-Aldrich, USA). Reaction mixture of total volume = 600 μL contained 550 μL of 50 mM sodium phosphate buffer (pH 7.0), 30 μL of pNPG (final conc. 5 mM), and 20 μL of CFE. Reaction was started by adding CFE to the mixture of buffer and substrate in Eppendorf tube pre-incubated 10 min at 37 °C. CFE from *P. putida* cells producing BglC enzyme was diluted 100–200 times. Reaction was terminated after 15 min of incubation at 37 °C in thermoblock by adding 400 μL of 1 M Na_2CO_3. Linearity of the enzymatic reaction during 15 min time course was initially verified by periodical withdrawal of the samples from reaction mixture of total volume of 1800 μL. Absorbance of the mixture was measured at 405 nm with UV/Vis spectrophotometer Ultrospec 2100 (Biochrom, UK) and activity was calculated using calibration curve prepared with p-nitrophenol standard (Sigma-Aldrich, USA). β-glucosidase activity in culture supernatants was measured correspondingly with 166 μL of culture supernatant in 600 μL of reaction mixture.

Activity of xylose isomerase (XylA) was measured as described by Le Meur et al. (2012) in microtiter plate format. In this assay, activity of XylA is coupled to consumption of NADH by sorbitol dehydrogenase. The assay mixture of total volume of 200 μL contained 50 mM Tris-Cl buffer (pH 7.5), 1 mM triethanolamine, 0.2 mM NADH, 0.5 U of sorbitol dehydrogenase, 10 mM $MgSO_4$, and 50 mM xylose. Reaction at 30 °C was started by addition of 5 μL of CFE.

Xylulokinase (XylB) activity was determined using the assay described by Eliasson et al. (2000) in microtiter plate format. In this assay, XylB activity is coupled with activities of pyruvate kinase and lactate dehydrogenase leading to the consumption of NADH. The reaction mixture of total volume of 200 μL contained 50 mM Tris-Cl buffer (pH 7.5), 0.2 mM NADH, 2 mM ATP, 2 mM $MgCl_2$, 0.2 mM phosphoenolpyruvate, 10 U of pyruvate kinase and 10 U, lactate dehydrogenase, and 10 mM D-xylulose. Reaction at 30 °C was started by addition of 5 μL of 20-fold diluted CFE.

Both in xylose isomerase and xylulokinase assay, the depletion of NADH was measured spectrophotometrically at 340 nm with Victor^2 1420 Multilabel Counter (Perkin Elmer, USA). Molar extinction coefficient of 6.22 mM^{-1} cm^{-1} for NADH was used for activity calculations. 1 unit (U) of activity corresponds to 1 μmol of substrate (pNPG or NADH) converted by enzyme per 1 min.

Activity of glucose dehydrogenase (Gcd) in *P. putida* EM42 and *P. putida* EM42 Δgcd was determined by measuring conversion of 5 g L^{-1} xylose to xylonate by cells suspension of A_{600} = 0.55 in 25 mL of M9 medium at 30 °C. The time course of the reaction was 6 h. 1 U of enzyme activity in whole cells corresponds to 1 μmol of xylonate produced per 1 min.

2.5. SDS-PAGE and Western blot analyses

CFE for determination of expression levels of selected enzymes were prepared using cell pellets from cultures induced with 1 mM 3MB and lysed with BugBuster Protein Exctraction Reagent as described above. Samples of CFE containing 5 μg of total protein were added with 5 × Laemmli buffer, boiled at 95 °C for 5 min and separated by SDS-PAGE using 12% gels. CFE prepared from *P. putida* cells with empty pSEVA238 plasmid was used as control. Gels were stained with Coomassie Brilliant Blue R-250 (Fluka/Sigma-Aldrich, Switzerland).

The staining step was omitted for Western blotting. Instead, proteins were electrotransferred from gel onto Immobilon-P membrane (Merck Millipore, Germany) of pore size = 0.45 μm using Trans-Blot SD Semi-Dry Transfer Cell (Bio-Rad, USA). Transfer conditions were: constant electric current of 0.1 A per gel, voltage of 5–7 V, time of run 30 min. Membrane was blocked overnight at 4 °C in 3% (w/v) dry milk in PBS buffer with 1% (v/v) TWEEN 20 and then incubated with mouse anti-

6xHis tag monoclonal antibody-HRP conjugate (Clontech, USA) for 2 h at RT. Membrane was washed with PBS buffer with 1% (v/v) TWEEN 20 and the proteins were visualized after incubation with BM Chemiluminescence Blotting Substrate (POD; Roche, Switzerland) using Amersham Imager 600 (GE Healthcare Life Sciences, USA).

2.6. Determination of ATP levels in P. putida cells

The ATP content in *P. putida* recombinants growing in M9 minimal medium with 5 g L^{-1} glucose or cellobiose was determined as described previously by Lai et al. (2016). Briefly, 2 mL of cell culture of A_{600} = 0.5, representing ~0.38 mg CDW, was centrifuged (13,000 g, 4 °C 2 min). Pellets were resuspended in 250 µL of ice-cold 20 mM Tris-Cl buffer (pH 7.75) with 2 mM EDTA and ATP was extracted using the trichloroacetic acid (TCA) method (Lundin and Thore, 1975). Ice-cold 5% (w/v) TCA (250 µL) with 4 mM EDTA was added, the suspension was mixed by vortexing for 20 s and incubated on ice for 20 min. Then, the suspension was centrifuged (13,000 g, 4 °C, 10 min) and supernatant (10 µL) was diluted 20-fold with ice cold 20 mM Tris-Cl buffer (pH 7.75) with 2 mM EDTA. The solution (50 µL) was mixed with 50 µL of the reagent from bioluminescence based ATP determination kit (Biaffin, Germany) prepared according to the manufacturer's instructions. After 10 min of incubation in dark, the luminescence signal of diluted supernatants was read in white 96-well assay plate (Corning Incorporated, USA) using the microplate reader SpectraMax (Molecular Devices, USA). The luminescence was quantified using the calibration curve prepared with pure ATP (Sigma-Aldrich, USA).

2.7. Confocal microscopy

Localization of XylE transporter fused with msfGFP in the cell membrane was verified by confocal microscopy of cells expressing the chimeric gene from pSEVA238_*xylE-gfp*C. Cells bearing pSEVA238_*gfp* or empty pSEVA238 were used as controls. Cells were grown in LB medium (30 °C, 275 rpm) until A_{600} = 0.5. Expression was induced with 0.5 mM 3MB and the growth continued at the same conditions for another 2.5 h. The sample of cell culture (100 µL) was centrifuged (5000 g, 4 °C, 5 min). Cells were washed twice with 1.5 mL of ice cold phosphate buffer saline (PBS; per 1 L: 8 g NaCl, 0.2 g KCl, 1.44 g Na_2HPO_4, 0.24 g KH_2PO_4, pH adjusted to 7.4 with HCl) and finally resuspended in 1 mL of the same buffer. Cells (5 µL of suspension) were mounted on poly-L-lysine coated glass slides (Sigma-Aldrich, USA) for 20 min, covered with cover glass and the slides were analyzed using confocal multispectral miscroscope Leica TCS SP5 (Leica Microsystems, Germany).

2.8. Other analytical techniques

Cell growth was monitored by measuring absorbance of a cell suspension at 600 nm using UV/Vis spectrophotometer Ultrospec 2100 (Biochrom, UK). For determination of glucose, xylose, xylonate, and cellobiose concentrations, culture supernatant (0.5 mL) was centrifuged (21,000 g, 4 °C, 10 min), filtered using Whatman Puradisc 4 mm syringe filter with nylon membrane (GE Healthcare Life Sciences, USA) and stored at − 20 °C for following analyses. Samples were analyzed by HPLC-LS system 920LC with light scattering (PL-ELS) detector (Agilent Technologies, USA) equipped with Microsorb MV NH2 column (5 µm, 250 mm × 4.6 mm). MilliQ H_2O (A) and acetonitrile (B) were used as eluents at a flow rate of 1 mL min^{-1}. Column temperature was 30 °C. Chemicals were identified using pure compound standards. Glucose and xylose concentrations in culture supernatants were also determined by Glucose (GO) Assay Kit (Sigma-Aldrich, USA) and Xylose Assay Kit (Megazyme, Ireland), respectively. Xylonic acid was measured using the hydroxamate method (Lien, 1959). Samples of culture supernatants (75 µL) were mixed 1:1 with 1.3 M HCl and heated at 100 °C for 20 min to convert xylonate to xylono-γ-lactone. Samples were cooled down on ice and 50 µL were added to 100 µL of hydroxylamine reagent freshly prepared by mixing 2 M hydroxylamine HCl with 2 M NaOH (pH of the reagent should be between 7 and 8). After 2 min interval, 65 µL of 3.2 M HCl and subsequently 50 µL of $FeCl_3$ solution (10 g in 100 mL of 0.1 M HCl) were added. Absorbance at 550 nm was measured immediately using Victor^2 1420 Multilabel Counter (Perkin Elmer, USA). Xylonate concentrations were quantified with a standard curve prepared using the pure compound (Sigma-Aldrich, USA).

2.9. Statistical analyses

All the experiments reported here were repeated independently at least twice (number of repetitions is specified in figure and table legends). The mean values and corresponding standard deviations (SD) are presented. When appropriate, data were treated with a two-tailed Student's *t*-test in Microsoft Office Excel 2013 (Microsoft Corp., USA) and confidence intervals were calculated for given parameter to manifest a statistically significant difference in means between two experimental datasets.

3. Results and discussion

3.1. Engineering the cellobiose metabolism module

Initial tests with *P. putida* EM42 in minimal medium with cellobiose showed no growth of this platform strain on the disaccharide (Fig. 2A). The absence of growth on cellobiose implied either lack of transport through the cell membranes or missing enzymatic machinery for disaccharide hydrolysis or phosphorolysis. As we detected no β-glucosidase (EC 3.2.1.21) activity in culture supernatants or lysates prepared from cells incubated with cellobiose, we focused initially on the latter bottleneck.

We probed functional expression of two different intracellular β-glucosidases, Ccel_2454 from the Gram-positive *meso*philic bacterium *Clostridium cellulolyticum* and BglC from the Gram-positive thermophilic bacterium *Thermobifida fusca* (see Supporting information for nucleotide sequences of all enzymes used in this study). Both enzymes have been expressed successfully in *E. coli* and are compatible with moderate temperatures and neutral or slightly acidic pH (Desai et al., 2014; Fan et al., 2012; Spiridonov and Wilson, 2001). In addition, we overexpressed the EM42 endogenous *bglX* gene (PP_1403), which is annotated as periplasmic β-glucosidase in the Pseudomonas Genome Database (www.pseudomonas.com/). The enzyme has relatively high amino acid sequence identity (61%) with the well-characterized *E. coli* β-glucosidase BglX (Yang et al., 1996). Each of these genes was cloned into the pSEVA238 plasmid downstream of the inducible XylS/Pm promoter. We tested the effect of gene expression on EM42 viability, as well as soluble protein production and enzyme activity in cell-free extracts (CFE).

All three enzymes were produced in the soluble fraction of the *P. putida* chassis grown in LB medium (Fig. S1A), but only Ccel_2454 and BglC showed measurable β-glucosidase activity. No activity was detected in CFE containing endogenous BglX, whose overexpression also had a clear toxic effect on the host (Fig. S2). Sequence of the *bglX* gene cloned from our *P. putida* strain was identical to the sequences available in Pseudomonas Genome Database. More studies probing the cellulolytic activities in *P. putida* KT2440 exist (Tozakidis et al., 2016) but, according to our knowledge, there is currently no work which would prove ß-glucosidase activity of BglX in this strain. It is thus probable that the gene gained a loss-of-function mutation a long time ago, when the bacterium was not benefited by the enzyme (Hottes et al., 2013), and such mutation is now spread in used laboratory KT2440 strains and their derivatives worldwide. The Ccel_2454 activity (0.03 ± 0.01 U/1 mg of total protein in CFE) measured with the colorimetric substrate p-nitrophenol-beta-D-glucopyranoside at 37 °C (pH 7.0) was limited. In contrast, CFE of *P. putida* EM42 expressing the *bglC* gene showed high

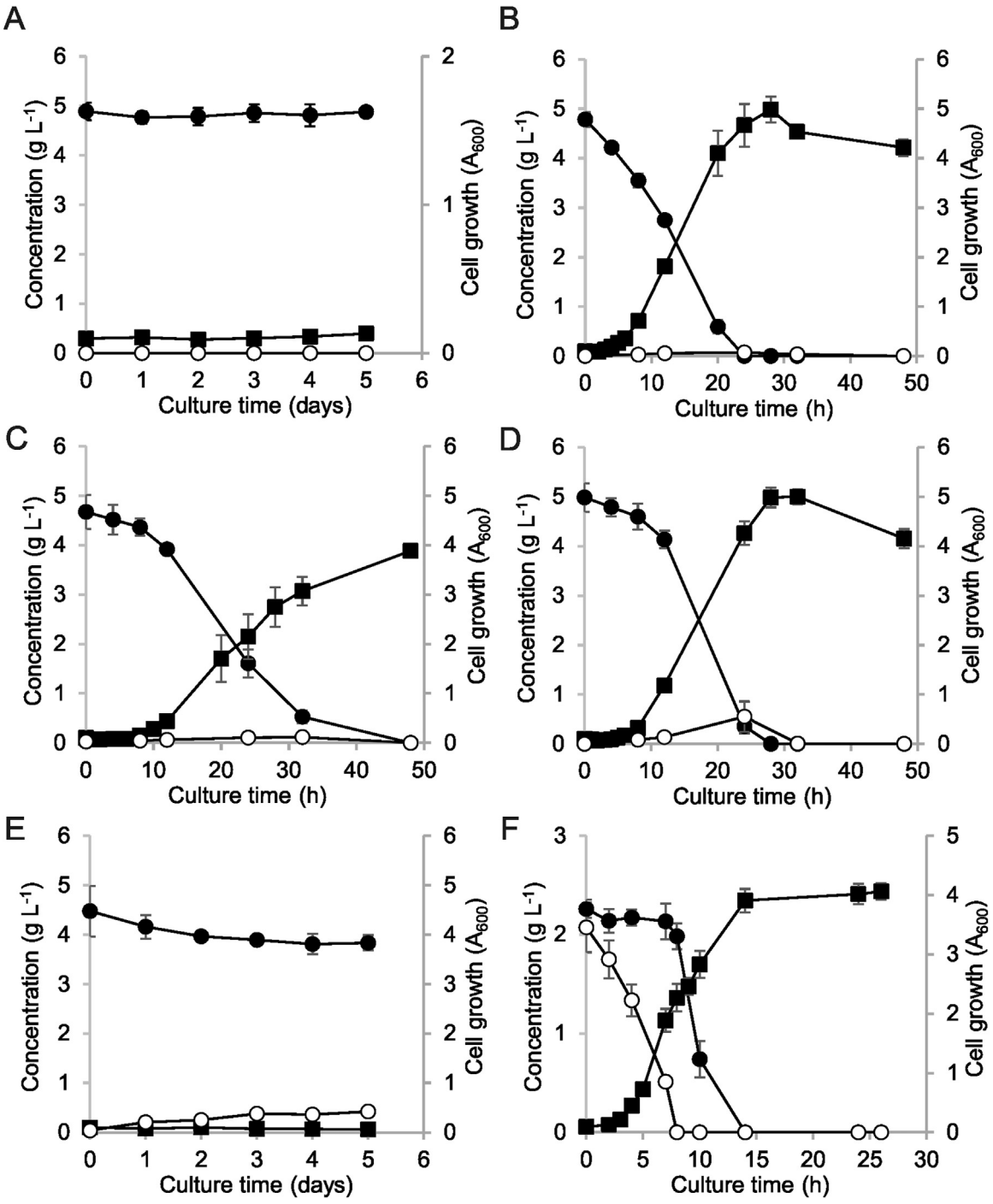


Fig. 2. Growth of *P. putida* EM42 and its recombinants in minimal medium with 5 g L^{-1} D-cellobiose. Experiments were carried out in shaken flasks (30 °C, 170 rpm). (A) *P. putida* EM42, (B) *P. putida* EM42 pSEVA2213_*bglC*, (C) *P. putida* EM42 Δ*gts* pSEVA2213_*bglC*, (D) *P. putida* EM42 Δ*gcd* pSEVA2213_*bglC*, (E) *P. putida* EM42 Δ*gts* Δ*gcd* pSEVA2213_*bglC*, (F) *P. putida* EM42 pSEVA2213_*bglC* in minimal medium with D-glucose and D-cellobiose (2 g L^{-1} each). D-cellobiose, filled circles (•); D-glucose, open circles (○); cell growth, filled squares (■). Data points shown as mean ± SD from two to three independent experiments.

activity (4.57 ± 0.50 U mg^{-1}). Trace BglC activity was also detected in culture supernatants (2.76 ± 0.24 U L^{-1}), which indicates that a small amount of overexpressed enzyme can either exit the cell during growth through an unknown mechanism or some quantity of the enzyme is released into the medium after the lysis of dead cells. Western blot analysis of a CFE sample containing His-tagged BglC and a sample of culture supernatant (Fig. S1B) nonetheless confirmed that the great majority of the protein was expressed intracellularly. After induction with 0.2 mM 3-methylbenzoate (3MB), strong BglC expression and activity enabled growth of the host strain in minimal medium with 5 g L^{-1} cellobiose as a sole carbon source. In this preliminary shake flask experiment, the OD_{600} at 48 h was 3.6 (Fig. S3).

The *bglC* gene was subsequently subcloned into the low copy plasmid pSEVA2213 with the constitutive promoter pEM7 that functions well in both *P. putida* and *E. coli* (Zobel et al., 2015). The *P. putida* EM42 bearing the pSEVA2213_*bglC* construct grew rapidly in minimal medium with 5 g L^{-1} cellobiose, with only a ~2 h adaptation period and a specific growth rate of 0.35 ± 0.02 h^{-1} (Fig. 2B, Table 2). The substrate consumption rate and biomass yield parameters were about 40% and 20% lower, respectively, than those of *P. putida* EM42

Table 2
Growth parameters for batch cultures of *Pseudomonas putida* EM42 recombinants carried out on cellobiose. Values shown as mean ± SD from two to three independent experiments.

Strain	μ (h^{-1})[a]	Substrate consumption rate at 12/24 h ($g\,L^{-1}\,h^{-1}$)[b]	$Y_{X/S}$ ($g\,g^{-1}$)[c]	BglC specific activity ($U\,mg^{-1}$)[d]
EM42 pSEVA2213_*bglC*	0.35 ± 0.02	0.17 ± 0.01/0.20 ± 0.01	0.37 ± 0.05	6.48 ± 0.75
EM42 Δ*gts* pSEVA2213_*bglC*	0.29 ± 0.04	0.06 ± 0.02/0.13 ± 0.03	0.31 ± 0.06	4.73 ± 0.36
EM42 Δ*gcd* pSEVA2213_*bglC*	0.30 ± 0.05	0.07 ± 0.01/0.19 ± 0.01	0.38 ± 0.00	5.18 ± 0.47
EM42 Δ*gts* Δ*gcd* pSEVA2213_*bglC*	n.d.	n.d.	n.d.	1.22 ± 0.20[e]
EM42 Δ*gcd bglC*	0.28 ± 0.03	0.26 ± 0.02/0.22 ± 0.00	0.29 ± 0.06	3.78 ± 0.35
EM42 pSEVA2213[f]	0.60 ± 0.03	0.23 ± 0.00/0.21 ± 0.00	0.48 ± 0.03	n.d.

[a] The specific growth rate (μ) was determined during exponential growth.
[b] The substrate consumption rate for cellobiose and glucose was determined for the initial 12 and 24 h of culture.
[c] The biomass yield on substrate ($Y_{X/S}$) was determined 24 h after each culture began to grow exponentially.
[d] Specific activities were determined in cell free extracts prepared from cells grown on cellobiose to A_{600} = 1.0.
[e] Specific activity of BglC in EM42 Δgts Δgcd pSEVA2213_*bglC* recombinant was determined in cell-free extracts obtained from cells grown in LB medium.
[f] Strain with empty pSEVA2213 plasmid cultured in minimal medium with 5 $g\,L^{-1}$ glucose as a sole carbon source.

pSEVA2213 cultured in minimal medium with glucose (Table 2). As all disaccharide was consumed within 24 h of culture, the *P. putida* strain outperformed the best engineered *E. coli* strain CP12CHBASC30, which utilized 3.3 $g\,L^{-1}$ of cellobiose after 32 h in comparable conditions (Vinuselvi and Lee, 2011). No growth was observed of *E. coli* Dh5α transformed with pSEVA2213_*bglC* and incubated in the same conditions (Fig. S4), despite the relatively high β-glucosidases activity (1.51 ± 0.12 $U\,mg^{-1}$) detected in these cells. This result highlights the excellent match in between the BglC cellulase selected and the *P. putida* host.

3.2. Cellobiose transport in P. putida EM42

To decipher transport pathways for cellobiose (a glucose dimer) in *P. putida*, we first focused on the well-described import of monomeric glucose (del Castillo et al., 2007). There are two routes for glucose uptake in *P. putida* KT2440. The first route encompasses direct translocation of a glucose molecule from periplasm to cytoplasm by the ATP-dependent ABC transporter, encoded by the *gtsABCD* operon (PP_1015-PP_1018), and subsequent phosphorylation of hexose to glucose-6-phosphate by glucokinase (del Castillo et al., 2007). The second is a periplasmic pathway formed by membrane-bound PQQ-dependent glucose dehydrogenase Gcd (PP_1444). This enzyme oxidizes D-glucose to D-glucono-1,5-lactone, which is hydrolyzed either spontaneously or by the action of gluconolactonase Gnl (PP_1170) to D-gluconate. Gluconate can be further oxidized to 2-keto-D-gluconate by periplasmic gluconate dehydrogenase (Gad). Both gluconate and 2-ketogluconate can pass through the outer membrane or are imported into the cytoplasm, where the oxidation pathway merges with the direct phosphorylation route at the level of 6-phospho-D-gluconate (del Castillo et al., 2007).

Uptake of cellobiose and other cellodextrins of varying lengths through the ABC-type transporters is common in cellulolytic bacteria due to the relatively broad specificity of these systems (Nataf et al., 2009; Parisutham et al., 2017). To test its relevance for cellobiose uptake in EM42 strain, we deleted the *gtsABCD* operon that encodes the ABC glucose transporter. The mutant transformed with the pSEVA2213_*bglC* plasmid could still grow on cellobiose but it showed a substantially prolonged (~7 h) adaptation phase (Fig. 2C) and three determined growth parameters were reduced when compared with *P. putida* EM42 pSEVA2213_*bglC* (Table 2). We thus assumed that the glucose ABC transporter played a role in cellobiose uptake, but was not the only access route for the disaccharide in *P. putida*.

Closure of the second glucose uptake route by deleting the glucose dehydrogenase gene *gcd* in *P. putida* EM42 had no notable effect on growth on cellobiose, but slightly prolonged the lag phase (~3 h; Fig. 2D). Substrate consumption during the initial 12 h of growth was nonetheless significantly reduced (Table 2). This suggests that also Gcd might participate in cellobiose uptake and the reduced consumption of substrate could then stem from slower initial expression of ABC transporter components (the remaining transport route for cellobiose in EM42 Δgcd mutant) that is normally induced by monomeric glucose (del Castillo et al., 2007). It is worth noting here that neither the genes which encode two out of three carbohydrate-selective porins (*oprB-1* and *oprB-2*) adjacent to *gtsABCD* and *gcd*, respectively, nor their regulatory sequences were affected by the scarless deletions. Growth on disaccharide was completely abolished when the deletions in the direct phosphorylation and oxidative routes were combined (Fig. 2E, Table 2). This seems to confirm the significance of glucose uptake pathways for metabolism of cellobiose in *P. putida* EM42. Nonetheless, currently available data do not provide insight into the detailed functioning of the peripheral oxidative route in upper cellobiose metabolism in *P. putida*, which thus remains to be elucidated by our future experiments.

From the acquired growth parameters of the *P. putida* mutants, it can be deduced that the direct phosphorylation route is important for cellobiose import. On the other hand, experimental evidence shows that glucose enters *P. putida* cells predominantly through the peripheral oxidative pathway (Nikel et al., 2015). Despite these opposing access route preferences, when *P. putida* EM42 pSEVA2213_*bglC* was exposed to a mixture of the two sugars (2 $g\,L^{-1}$ each), we observed diauxic growth (Fig. 2F). Glucose was utilized first during the initial 8 h of the experiment. When all hexose was removed from the medium, cellobiose was consumed rapidly during the next 6 h of the culture. To conclude, these experiments suggest that glucose and cellobiose share the same acess routes in *P. putida* and that monomeric hexose is a preferred substrate in the mixture of the two carbon sources.

3.3. Probing energetic benefit of cellobiose metabolism in engineered P. putida

ATP is a universal energy source and a major driving force for biochemical processes in microbial cell factories (Hara and Kondo, 2015). Due to its variant of glycolysis – the Entner-Doudoroff pathway - *P. putida* yields only one net ATP per one mole of assimilated glucose (Nikel et al., 2015). In the case, for instance, of *E. coli* with its characteristic Embden–Meyerhof–Parnas pathway, the ATP yield per molecule of glucose is twice as high. It is thought that environmental or engineered microorganisms that prefer to metabolize cellobiose instead of glucose are more energetic and robust than their glucose-utilizing counterparts (Chen, 2015; Lynd et al., 2002; Parisutham et al., 2017). The benefits were demonstrated in bacteria with a cellobiose-specific PEP-phosphotransferase system transporter, in which one mole of ATP is consumed per one mole of imported substrate; this was also apparent in microorganisms that metabolize cellobiose through phosphorolysis, in which only one ATP per disaccharide is needed to form two activated molecules of glucose (Kajikawa and Masaki, 1999; Shin et al., 2014;

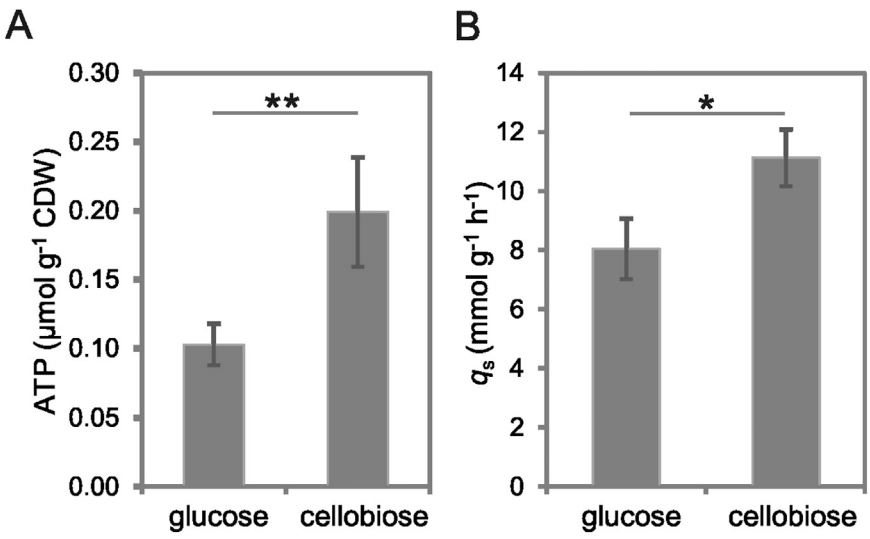


Fig. 3. Energetic benefit of *P. putida* EM42 cells growing on D-cellobiose. (A) Comparison of intracellular ATP concentrations in *P. putida* EM42 pSEVA2213 growing on 5 g L^{-1} D-glucose and in *P. putida* EM42 pSEVA2213_*bglC* growing in minimal medium with 5 g L^{-1} D-cellobiose. Measurements were performed with cells collected from cultures in exponential growth (A_{600} = 0.5). (B) Specific carbon consumption rate (q_s) for glucose (8.04 ± 1.02 mmol g^{-1} h^{-1}) and cellobiose (11.13 ± 0.96 mmol g^{-1} h^{-1}) determined in the exponential growth phase for *P. putida* EM42 pSEVA2213 and *P. putida* EM42 pSEVA2213_*bglC*, respectively. Data shown as mean ± SD from three independent experiments. Asterisks denote significance in difference in between two means at $P < 0.05$ (*) or $P < 0.01$ (**).

Thurston et al., 1993).

We determined ATP levels in EM42 pSEVA2213 and EM42 pSEVA2213_*bglC* strains grown on glucose and cellobiose, respectively, to evaluate the effect of altered substrate on *P. putida* energy status (Fig. 3A). Indeed, the ATP level in *P. putida* grown on cellobiose was almost double that of the cells cultured on glucose. ATP savings could partially stem from cellobiose transport. One ATP per molecule of glucose is theoretically saved when cellobiose enters the cell with the help of an ABC-type transporter, as the ATP cost is known to be constant per import event (Parisutham et al., 2017). The oxidative pathway is also thought to provide the cell with additional energy through the transfer of electrons from membrane-bound Gcd and Gad directly to the respiratory chain enzymes (Ebert et al., 2011). The higher ATP level on cellobiose was not accompanied by more efficient conversion of carbon into biomass (Table 2, Fig. 3B), which might be attributed to the higher respiration activity of the cells grown on cellobiose (Ebert et al., 2011). An extraordinary amount of ATP in the cell could also inhibit citrate synthase, which would slow down the TCA cycle as well as formation of biomass precursors (Smith and Williamson, 1971). To sum up, the ATP saved during the growth on a biotechnologically relevant substrate can be used by the cells *e.g.*, for fueling exogenous biosynthetic reactions, and will thus further increase the value of the engineered EM42 platform strain as a cell factory for bioproduction (Hara and Kondo, 2015).

3.4. Engineering a metabolic module for xylose utilization

We first verified the absence of xylose catabolism in *P. putida* EM42. Our bioinformatic analysis confirmed that the KT2440 genome has no genes that encode homologues of *E. coli* XylA or XylB proteins that form the isomerase pathway. The EM42 strain was then incubated in minimal medium with xylose, and the cell density and substrate concentration were measured for five consecutive days (Fig. 4A). No growth was observed, despite the fact that only 10% of the starting xylose concentration was detected in the culture medium after five days. Meijnen and co-workers (2008) described a similar phenomenon for engineered *P. putida* S12. In that case, the majority of D-xylose was oxidized to the dead-end product D-xylonate by the periplasmic glucose dehydrogenase Gcd. In fact, xylonate concentrations determined in samples from the five-day experiment with the strain EM42 suggested that all xylose was converted to the acid, which was not assimilated by the cells (Fig. 4A). Xylonate formation was accompanied by a decrease in the culture pH from 7.0 to 6.2. This initial experiment provided additional evidence of Gcd broad substrate specificity in *P. putida* KT2440 and its derivatives. The phenomenon of xylonate formation was not discussed in the study by Le Meur et al. (2012), who implanted the isomerase pathway in the KT2440 strain.

To avoid accumulation of an undesirable metabolite, we transplanted the *xylAB* fragment of the *xyl* operon from *E. coli* BL21 (DE3), which encodes xylose isomerase XylA and xylulokinase XylB, directly to *P. putida* EM42 Δ*gcd*. The gene absence was verified by measuring the whole-cell activity with xylose. While the Gcd activity of 2.87 ± 0.21 U per 1 g of dry cell mass was detected in *P. putida* EM42, no activity was detected in Δ*gcd* mutant.

The *xylAB* fragment was amplified as a whole. The *xylA* gene (SI sequences) was provided with a consensus RBS, with the native RBS maintained upstream of the *xylB* gene. The fragment was cloned into pSEVA2213 and *xylAB* expression was verified in the EM42 Δ*gcd* strain (Fig. S5). XylA and XylB activities determined in CFE were higher than those reported for engineered *P. putida* S12 growing on xylose (Table 3) (Meijnen et al., 2008). The recombinant EM42 cells nonetheless showed only limited growth and substrate uptake in minimal medium with 5 g L^{-1} xylose (Table 3, Fig. 4B). When xylonate accumulation no longer hindered efficient cell use of xylose, we found substrate transport to be another bottleneck to xylose metabolism in this host. This was not anticipated based on a previous study with the xylose-utilizing KT2440 strain, which reported only implantation of the XylAB metabolic module with no transport system (Le Meur et al., 2012). In another report, an upregulated glucose ABC transporter was nonetheless defined as one of the major changes that shaped laboratory-evolved *P. putida* S12 towards rapid growth on xylose (Meijnen et al., 2012).

The xylose-proton symporter XylE from *E. coli* is a relatively small (491 amino acids) single-gene transporter with a known tertiary structure and a well-described transport mechanism (Davis and Henderson, 1987; Wisedchaisri et al., 2014). In the recent study by Yim et al. (2016), XylE was selected as the best candidate among three tested pentose transporters that allowed growth of *Corynebacterium glutamicum* on xylose. The same transporter was also applied successfully in *Zymomonas mobilis* (Dunn and Rao, 2014). We probed XylE performance in *P. putida*. The gene was PCR-amplified from the genome of *E. coli* BL21(DE3) and supplied with consensus RBS to secure sufficient expression of *xylE* cloned downstream of the *xylAB* fragment in pSEVA2213_*xylAB*, to form the synthetic *xylABE* operon. The *xylE* was also simultaneously amplified without a stop codon and was cloned upstream of the *gfp* gene in the pSEVA238_*gfp* plasmid bearing the inducible XylS/Pm promoter to form translational fusion. The resulting construct was electroporated into *P. putida* EM42, and fluorescence microscopy confirmed targeting of XylE-GFP protein chimera in the cell membrane after 3MB induction (Fig. 5).

Once we confirmed correct XylE expression and localization in *P. putida*, the EM42 Δ*gcd* strain was transformed with the

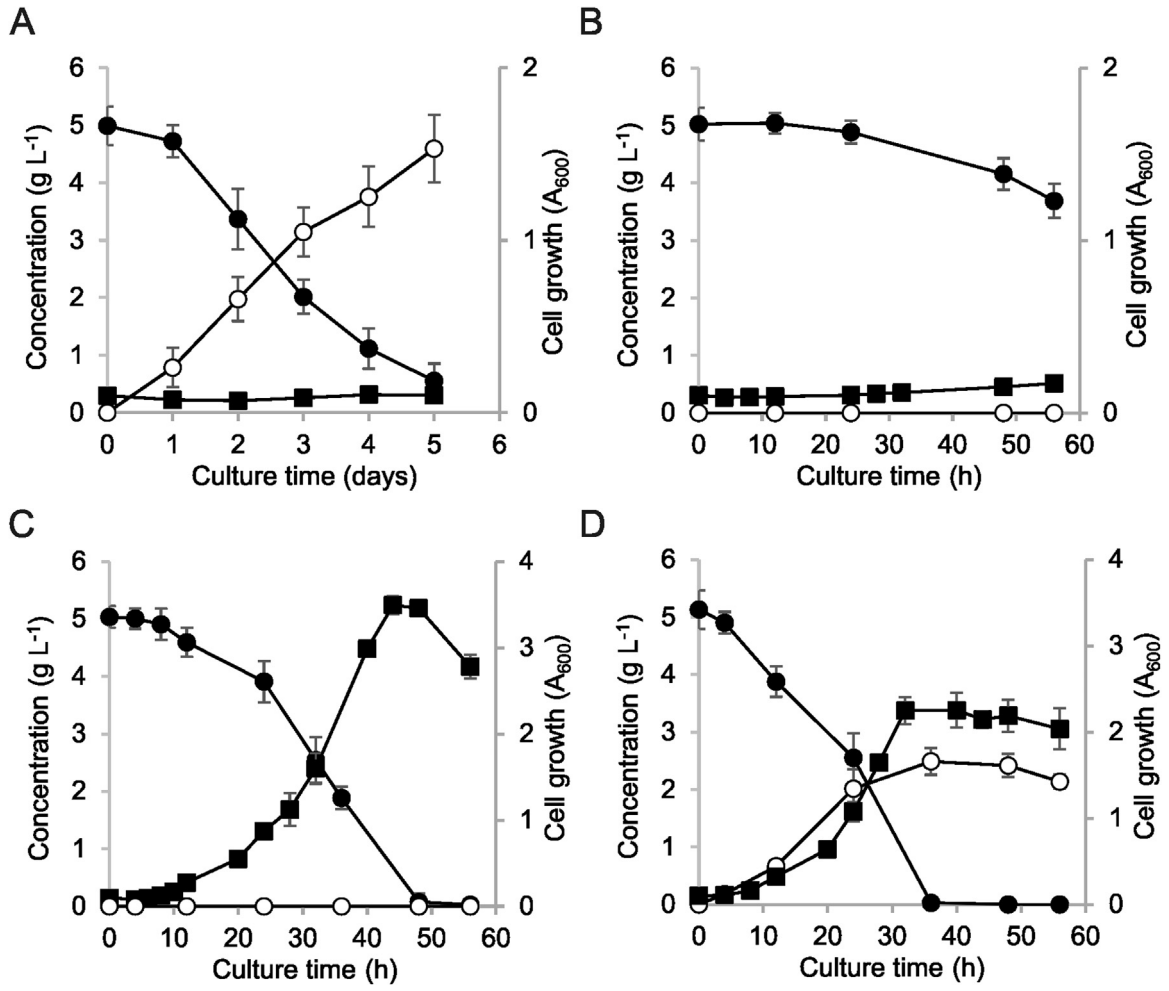


Fig. 4. Growth of *P. putida* EM42 and its recombinants in minimal medium with 5 g L^{-1} D-xylose. Experiments were carried out in shaken flasks at 30 °C and 170 rpm. (A) *P. putida* EM42, (B) *P. putida* EM42 Δ*gcd* pSEVA2213_*xylAB*, (C) *P. putida* EM42 Δ*gcd* pSEVA2213_*xylABE*, (D) *P. putida* EM42 pSEVA2213_*xylABE*. D-xylose, closed circles (•); D-xylonate, open circles (○); cell growth, closed squares (■). Data shown as mean ± SD from three independent experiments.

pSEVA2213_*xylABE* construct to test functioning of the whole synthetic operon. Co-expression of the exogenous transporter with xylose isomerase and xylulokinase genes improved the specific growth rate on xylose by 8-fold when compared with the recombinant without *xylE* (Fig. 4C and Table 2); improvement was also observed for other growth parameters (Table 2). The faster growth of the recombinant bearing the *xylABE* synthetic operon could not be attributed to changes in XylA or XylB activity (Table 2). We observed a > 2-fold decrease in XylB activity when *xylE* was subcloned downstream of *xylB* and the position of the gene in operon was changed. Finally, we verified the importance of the *gcd* deletion for complete xylose utilization by *P. putida* recombinants in the experiment with the EM42 pSEVA2213_*xylABE* strain (Fig. 4D). It is clear from the time course of the culture that ~50% of all uptaken xylose was still converted non-productively to xylonate in the bacterium with functional Gcd, despite the presence of the heterologous machinery that funnels the substrate to the pentose phosphate pathway.

We thus demonstrate that efficient xylose metabolism can be established in the platform strain *P. putida* EM42 when two major bottlenecks – the peripheral oxidative pathway and the missing transport system – are removed. Neither of these bottlenecks was rationally engineered in previous studies of pseudomonad metabolization of xylose (Le Meur et al., 2012; Meijnen et al., 2008, 2012). The specific growth and xylose consumption rates of the *xylABE*-bearing recombinant were lower than those we measured for *P. putida* EM42 utilizing glucose or cellobiose (Tables 2, 3). Both parameters can nonetheless be improved in co-utilization experiments in which host cell growth is supported by an additional carbon source (Ha et al., 2011; Lee et al., 2016).

Table 3
Growth parameters and specific activities of xylose isomerase and xylulokinase for batch cultures of *Pseudomonas putida* EM42 recombinants carried out on xylose. Data shown as mean ± SD from three independent experiments.

Strain	μ (h^{-1})[a]	Substrate consumption rate (g L^{-1} h^{-1})[b]	$Y_{X/S}$ (g g^{-1})[c]	XylA specific activity (U mg^{-1})[d]	XylB specific activity (U mg^{-1})[d]
EM42 Δ*gcd* pSEVA2213_*xylAB*	0.02 ± 0.00	0.01 ± 0.01	0.08 ± 0.02	0.12 ± 0.00	2.58 ± 0.12
EM42 Δ*gcd* pSEVA2213_*xylABE*	0.17 ± 0.02	0.05 ± 0.01	0.27 ± 0.08	0.12 ± 0.01	1.14 ± 0.12

[a] The specific growth rate (μ) was determined during exponential growth.
[b] The xylose consumption rate was determined for the initial 24 h of culture.
[c] The biomass yield on substrate ($Y_{X/S}$) was determined 24 h after each culture began exponentiall growth.
[d] Specific activities were determined in cell free extracts prepared from cells grown on xylose to OD_{600} = 1.0.

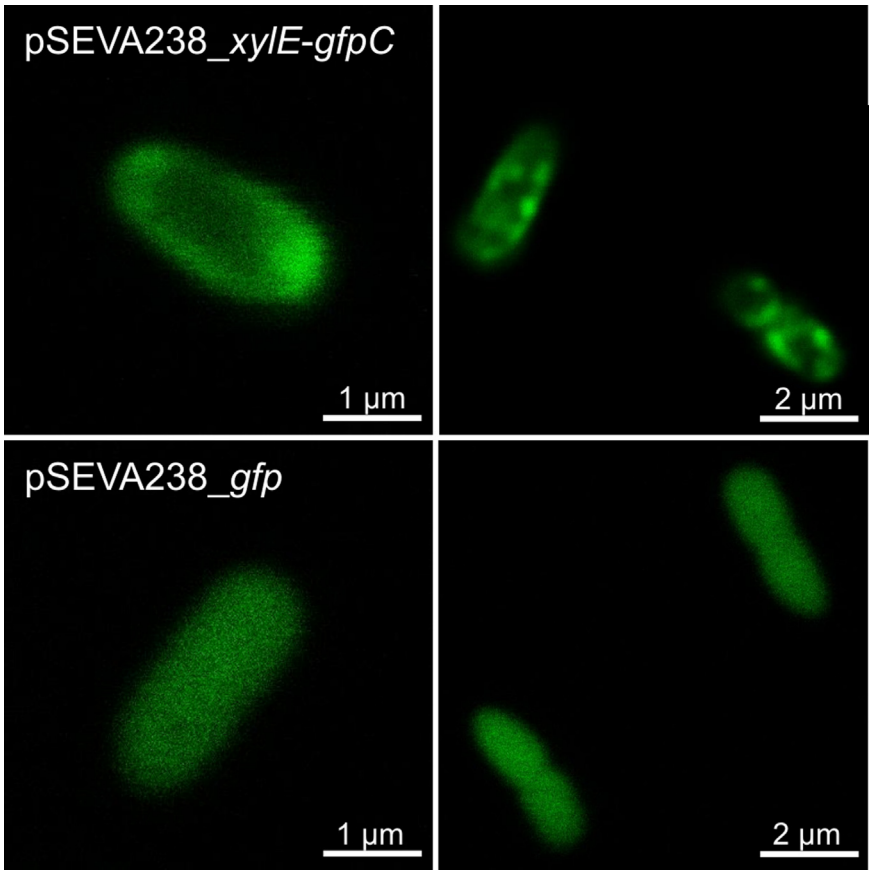


Fig. 5. Localization of the XylE-GFP chimera in the cell membrane of *P. putida* EM42. The experiment was performed with *P. putida* EM42 pSEVA238_*xylE-gfpC* cells with *xylE* and *gfp* genes cloned to form a translational fusion, and with *P. putida* EM42 pSEVA238_*gfp* (negative control). Expression of the chimeric gene or *gfp* only was induced for 2.5 h with 0.5 mM 3-methylbenzoate in lysogeny broth at 30 °C. GFP fluorescence in cells was then monitored with a confocal microscope. Note that GFP fluorescence in cells producing the XylE-GFP chimera is predominant in membrane regions.

3.5. Co-utilization of xylose with glucose and cellobiose by engineered *P. putida EM42*

Simultaneous uptake of carbohydrates is a desirable property in any microbial cell factory used in bioprocesses for valorization of lignocellulosic substrates (Lynd et al., 2002; Stephanopoulos, 2007). Co-utilization of biomass hydrolysis products increases the efficiency of the process and prevents accumulation of non-preferred sugars, usually pentoses, in batch and continuous fermentations (Jarmander et al., 2014; Kim et al., 2015).

We first sought to verify simultaneous utilization of glucose and xylose in *P. putida* EM42 Δ*gcd* pSEVA2213_*xylABE*. This experiment with two monomeric sugars was an essential prerequisite for co-utilization of xylose and cellobiose in engineered *P. putida*. The strain with the *gcd* deletion was used to avoid xylose oxidation to the dead-end by-product xylonate. As explained in Section 3.2., this deletion is not detrimental either for glucose or for cellobiose uptake in *P. putida*, and both molecules can enter the cell with the help of ABC transporter. Equal concentrations of monosacharides (2 g L^{-1}) were used to better visualize the differences in glucose and xylose consumption. In cultures of the negative control *P. putida* EM42 Δ*gcd* pSEVA2213 lacking the *xylABE* operon, the xylose concentration dropped by only 13% after 24 h (Fig. 6A). It is possible that some xylose entered the cells by non-specific transport routes. *Pseudomonas putida* EM42 Δ*gcd* pSEVA2213_*xylABE* assimilated glucose and xylose simultaneously, and no sugar was detected in culture supernatants after 24 h (Fig. 6B). The presence of the additional carbon source significantly accelerated xylose uptake by recombinant *P. putida* (Fig. 6B). On average, 2 g L^{-1} of xylose were consumed during the initial 24 h of co-utilization experiments, while < 1 g L^{-1} was utilized in cultures with pentose alone at a starting concentration of 5 g L^{-1} (Fig. 4C). The consumption of both sugars and cell growth could be accelerated by pre-culturing the cells in minimal medium with glucose and xylose (Fig. S6), which presumably led to pre-induction of the glucose ABC transporter components (del Castillo et al., 2007) and to faster adaptation to the carbohydrate carbon source. The substrate consumption rate of xylose was nonetheless still lower than that of glucose (Fig. 6B and S6). This experiment demonstrated the ability of engineered *P. putida* to co-utilize hexose and pentose without CCR.

For the cellobiose and xylose co-utilization experiments, new metabolic modules had to be combined in a single *P. putida* cell. As the synthetic *bglC-xylABE* operon borne on a single pSEVA2213 plasmid appeared to be unstable, we integrated the *bglC* gene directly into the *P. putida* EM42 Δ*gcd* chromosome. The gene with the pEM7 promoter and consensus RBS was subcloned into the mini-Tn5-vector pBAMD1–4, which allows for random chromosomal insertions and subsequent selection of the optimal phenotype from a broad expression landscape (Martínez-García et al., 2014a). The plasmid construct was electroporated into *P. putida* with a transformation efficiency of ~160,000 CFU μg^{-1} of DNA. The best candidates were selected in minimal medium with cellobiose (see Methods). The fastest growing clone with the *bglC* insertion in gene annotated as the tyrosine recombinase subunit *xerD* (PP_1468) was designated *P. putida* EM42 Δ*gcd bglC*. With another tyrosine recombinase XerC, functional XerD is pivotal for the process of chromosome segregation during bacterial cell division (Blakely et al., 2000). Insertion of the expression cassette with *bglC* into the *xerD* gene nonetheless did not notably affect host fitness. *Pseudomonas putida* EM42 Δ*gcd bglC* growth parameters on cellobiose were comparable with those of the *P. putida* EM42 Δ*gcd* recombinant bearing the *bglC* gene on the plasmid; substrate consumption rate was even faster (Table 2, Fig. S7). The insertion had no effect on cell viability either in rich LB medium or in minimal medium with citrate as the gluconeogenic carbon source (Fig. S8).

Hence, *P. putida* EM42 Δ*gcd bglC* was transformed with the pSEVA2213_*xylABE* construct, and the resulting recombinant used for co-utilization experiments. Cells were incubated in minimal medium

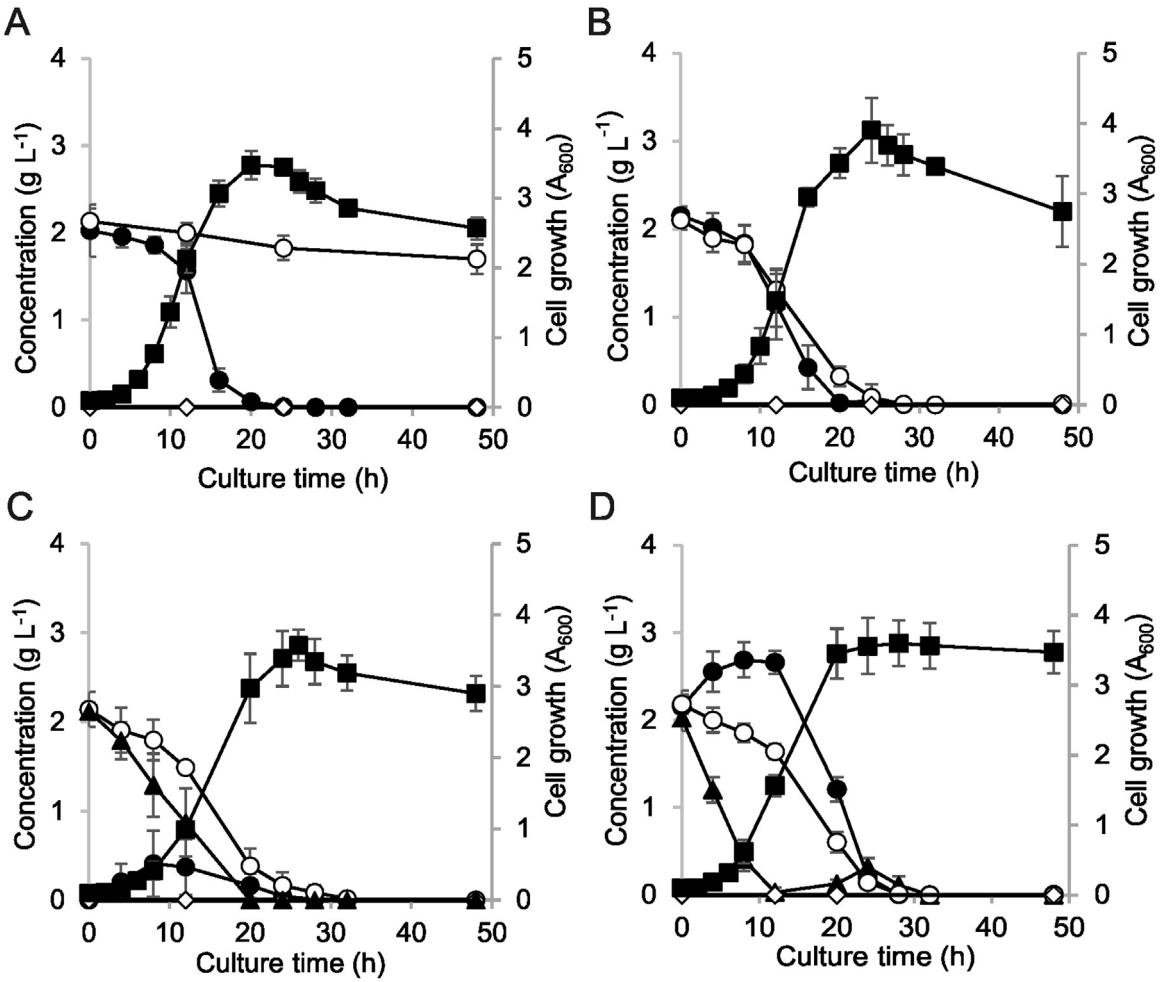


Fig. 6. Co-utilization of D-xylose with D-glucose or D-cellobiose by *P. putida* EM42 recombinants in minimal medium. Experiments were carried out in shaken flasks (30 °C, 170 rpm). (A) *P. putida* EM42 Δ*gcd* pSEVA2213 with glucose and xylose (2 g L^{-1} each), (B) *P. putida* EM42 Δ*gcd* pSEVA2213_*xylABE* with glucose and xylose (2 g L^{-1} each), (C) *P. putida* EM42 Δ*gcd bglC* pSEVA2213_*xylABE* with cellobiose and xylose (2 g L^{-1} each), (D) *P. putida* EM42 Δ*gcd bglC* pSEVA2213_*xylABE* with cellobiose, xylose, and glucose (2 g L^{-1} each). D-glucose, filled circles (•); D-xylose, open circles (○); D-cellobiose, filled triangles (▲); D-xylonate, open diamonds (◊); cell growth, filled squares (■). Data shown as mean ± SD from three independent experiments.

with cellobiose and xylose at equal concentrations (2 g L^{-1}). Once again, both substrates were co-utilized rapidly; no residual carbohydrate was detected in supernatants after 28 h culture (Fig. 6C). Xylose was consumed slower than cellobiose. We also tested performance of EM42 Δ*gcd bglC* pSEVA2213_*xylABE* strain in mixture of cellobiose, xylose, and glucose at 2 g L^{-1} concentration each (Fig. 6D). While xylose consumption remained unaffected by presence of glucose in medium, cellobiose uptake was notably accelerated. In contrast, glucose concentration in the culture supernatants increased until cellobiose was consumed and then all the hexose was co-utilized with remaining xylose. We argue that glucose, present in the medium already at the beginning of the culture, allowed faster cellobiose utilization by enhancing expression of the ABC transporter operon and adjacent *oprB* porin gene (del Castillo et al., 2007). We described in Section 3.2. that the direct phosporylation route appears to be the major uptake pathway for cellobiose, but up to 80–90% of glucose is assimilated through the oxidative route in *P. putida* KT2440 (Nikel et al., 2015). Such induced expression of the ABC transporter in *P. putida* Δ*gcd* mutant could thus lead to the inverted substrate preference observed in the last co-utilization experiment.

Small amounts of extracellular glucose were detected in all the cultures with *P. putida* EM42 Δ*gcd* mutant grown on cellobiose (Figs. 2D, 6C, 6D, and S7). This glucose accumulation was more pronounced in EM42 Δ*gcd bglC* strain which consumed cellobiose faster than *P. putida* EM42 Δ*gcd* with *bglC* gene on plasmid. We thus hypothesize that glucose accumulation in the EM42 Δ*gcd* mutant might stem from an imbalance between high β-glucosidase activity and *P. putida* glycolysis, namely its upper part encompassing glucokinase (PP_1011) and glucose-6-phosphate 1-dehydrogenase (PP_1022, PP_4042, PP_5351). The system must accommodate all the carbon in the form of D-glucose rather than in form of D-6-phosphogluconate, which predominantly enters the glycolysis in the cells with functional Gcd (Nikel et al., 2015). The imbalance could be reduced for instance by parallel modulation of expression of respective genes (Zhu et al., 2017).

4. Conclusions

Here we engineered *Pseudomonas putida* EM42, a robust platform strain derived from *P. putida* KT2440, to metabolize cellobiose and xylose, and to co-utilize these two carbohydrates. Within 24 h of culture in minimal medium, *P. putida* EM42 expressing the intracellular β-glucosidase BglC utilized 5 g L^{-1} cellobiose as a sole carbon source, thus outperforming the best cellobiose-utilizing *E. coli* strain constructed to date (Vinuselvi and Lee, 2011). This result highlights the need to select cellulases with good target host compatibility. We demonstrated that *P. putida* uses its native transport routes for cellobiose uptake, as no heterologous transporter had to be implanted into our

recombinants. This aligns *P. putida* KT2440 and its derivatives with several other microorganisms, such as *Clostridium thermocellum*, *Klebsiella oxytoca*, *Neurospora crassa*, and *Streptomyces spp*., which possess cellodextrin transport systems and can thus manage cellobiose-rich mixtures that result from partial hydrolysis of cellulosic materials (Lynd et al., 2002; Nataf et al., 2009; Zhou et al., 2001). While the function of glucose ABC transporter in cellobiose uptake appears to be clear and is supported by similar cases reported in the literature (Nataf et al., 2009; Parisutham et al., 2017), the role of Gcd and peripheral oxidative route is more difficult to explain. One could speculate that Gcd also has cellobiose dehydrogenase activity (EC 1.1.99.18) and converts cellobiose to cellobiono-1,5-lactone (Henriksson et al., 2000), hydrolyzed spontaneously to cellobionic acid, which would be transported to the cytoplasm. Intracellular β-glucosidase could then cleave cellobionate to glucose and gluconic acid (Li et al., 2015), two molecules easily metabolized by *P. putida*. Nonetheless, neither cellobionate formation nor its further metabolism in *P. putida* can be confirmed based on currently available experimental data and the role of Gcd in cellobiose transport in *P. putida* EM42 remains to be elucidated.

Metabolism of cellobiose generates more ATP in *P. putida* than when the same cells are cultured on glucose. The cellobiose-utilizing *P. putida* would thus be an even more robust host than the template strain for accommodating heterologous or engineering endogenous metabolic pathways for biosynthesis of value-added chemicals directly from the disaccharide or a co-substrate (Hara and Kondo, 2015). Finally, we identify the ability of *P. putida*, following introduction of the *xylABE* synthetic operon, to co-utilize cellobiose or glucose with pentose, with no need for further interventions in the regulatory mechanisms of central carbon metabolism. Xylose metabolism in the EM42 chassis was established based on the conclusions of previous reports of pentose-utilizing *P. putida* strains (Le Meur et al., 2012; Meijnen et al., 2008), whereas the need for an exogenous transporter and *gcd* deletion for complete xylose consumption and co-utilization with glucose in our *P. putida* KT2440 derivative is demonstrated in this study.

There is indeed room for improvement and further testing of the strains constructed here. Sugar consumption rates must be improved to secure that the *P. putida* recombinants will be industrially useful. We aim at addressing this challenge in our future work by altering sugar transport and fine-tuning expression of the relevant genes. We nonetheless argue that already this foundational study increases the value of *P. putida* for the biotechnological recycling of lignocellulosic feedstocks, specifically for processes that include partial hydrolysis of the input material. Using synthetic and systems biology approaches, carbon from new (hemic)cellulosic substrates -cellobiose and xylose- can be streamlined towards valuable chemicals whose production has been reported in *P. putida*, such as mcl-PHA (Poblete-Castro et al., 2013, p.), rhamnolipids (Tiso et al., 2017), terpenoids (Mi et al., 2014), coronatines (Gemperlein et al., 2017) and others (Loeschcke and Thies, 2015; Poblete-Castro et al., 2012). Given recent progress in *P. putida* KT2440 engineering for valorization of lignin (Johnson and Beckham, 2015; Linger et al., 2014), one could hypothesize that this effort could result in recombinant bacterial workhorses capable of simultaneous biotechnological processing of and adding value to all three lignocellulose-derived fractions.

Acknowledgements

We would like to thank Prof. Edward A. Bayer for providing us with pET21a_*bglC* plasmid, to Dr. Esteban Martínez-García for *P. putida* EM42 strain, and to Dr. Alberto Sánchez-Pascuala for pEMG_*gtsABCD* and pEMG_*gcd* plasmids and corresponding oligonucleotide primers. This work was funded by Projects HELIOS (BIO2015–66960-C3-2-R; MINECO/FEDER) and contracts of the European Union ARISYS (ERC-2012-ADG-322797), EmPowerPutida (EU-H2020-BIOTEC-2014-2015-6335536), MADONNA (H2020-FET-OPEN-RIA-2017-1 (766975) as well as the InGEMICS-CM (B2017/BMD-3691) contract of the Comunidad de Madrid (FSE, FECER). PD was the beneficiary of the EU's Horizon 2020 Marie Sklodowska-Curie grant agreement No 704410 (FUTURE).

Conflict of interest

Authors declare no conflict of interest.

Appendix A. Supplementary material

Dataset to this manuscript is available at: http://dx.doi.org/10.17632/j7ypmvfnvt.1

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.ymben.2018.05.019.

References

Abril, M.A., Michan, C., Timmis, K.N., Ramos, J.L., 1989. Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. J. Bacteriol. 171, 6782–6790.

Aparicio, T., de Lorenzo, V., Martínez-García, E., 2017. CRISPR/Cas9-based counterselection boosts recombineering efficiency in *Pseudomonas putida*. Biotechnol. J. http://dx.doi.org/10.1002/biot.201700161.

Aparicio, T., Lorenzo, V. de, Martínez-García, E., 2015. Broadening the SEVA plasmid repertoire to facilitate genomic editing of Gram-negative bacteria. In: Hydrocarbon and Lipid Microbiology Protocols, Springer Protocols Handbooks. Springer, Berlin, Heidelberg, pp. 9–27.

Bagdasarian, M., Lurz, R., Rückert, B., Franklin, F.C., Bagdasarian, M.M., Frey, J., Timmis, K.N., 1981. Specific-purpose plasmid cloning vectors. II. Broad host range, high copy number, RSF1010-derived vectors, and a host-vector system for gene cloning in *Pseudomonas*. Gene 16, 237–247.

Benedetti, I., de Lorenzo, V., Nikel, P.I., 2016. Genetic programming of catalytic *Pseudomonas putida* biofilms for boosting biodegradation of haloalkanes. Metab. Eng. 33, 109–118.

Blakely, G.W., Davidson, A.O., Sherratt, D.J., 2000. Sequential strand exchange by XerC and XerD during site-specific recombination at dif. J. Biol. Chem. 275, 9930–9936.

Boyer, H.W., Roulland-Dussoix, D., 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. 41, 459–472.

Bradford, M.M., 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. 72, 248–254.

Calero, P., Jensen, S.I., Bojanovič, K., Lennen, R.M., Koza, A., Nielsen, A.T., 2017. Genome-wide identification of tolerance mechanisms toward *p*-coumaric acid in *Pseudomonas putida*. Biotechnol. Bioeng. http://dx.doi.org/10.1002/bit.26495.

Chen, P., Fu, X., Ng, T.B., Ye, X.-Y., 2011. Expression of a secretory β-glucosidase from *Trichoderma reesei* in *Pichia pastoris* and its characterization. Biotechnol. Lett. 33, 2475–2479.

Chen, R., 2015. A paradigm shift in biomass technology from complete to partial cellulose hydrolysis: lessons learned from nature. Bioengineered 6, 69–72.

Davis, E.O., Henderson, P.J., 1987. The cloning and DNA sequence of the gene *xylE* for xylose-proton symport in *Escherichia coli* K12. J. Biol. Chem. 262, 13928–13932.

del Castillo, T., Ramos, J.L., Rodríguez-Herva, J.J., Fuhrer, T., Sauer, U., Duque, E., 2007. Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida:* genomic and flux analysis. J. Bacteriol. 189, 5142–5152.

Desai, S.H., Rabinovitch-Deere, C.A., Tashiro, Y., Atsumi, S., 2014. Isobutanol production from cellobiose in *Escherichia coli*. Appl. Microbiol. Biotechnol. 98, 3727–3736.

Dos Santos, V.A.P.M., Heim, S., Moore, E.R.B., Strätz, M., Timmis, K.N., 2004. Insights into the genomic basis of niche specificity of *Pseudomonas putida* KT2440. Environ. Microbiol. 6, 1264–1286.

Dunn, K.L., Rao, C.V., 2014. Expression of a xylose-specific transporter improves ethanol production by metabolically engineered *Zymomonas mobilis*. Appl. Microbiol. Biotechnol. 98, 6897–6905.

Dvořák, P., Nikel, P.I., Damborský, J., de Lorenzo, V., 2017. Bioremediation 3.0: engineering pollutant-removing bacteria in the times of systemic biology. Biotechnol. Adv. 35, 845–866.

Ebert, B.E., Kurth, F., Grund, M., Blank, L.M., Schmid, A., 2011. Response of *Pseudomonas putida* KT2440 to increased NADH and ATP demand. Appl. Environ. Microbiol. 77, 6597–6605.

Eliasson, A., Boles, E., Johansson, B., Osterberg, M., Thevelein, J.M., Spencer-Martins, I., Juhnke, H., Hahn-Hägerdal, B., 2000. Xylulose fermentation by mutant and wild-type strains of *Zygosaccharomyces* and *Saccharomyces cerevisiae*. Appl. Microbiol. Biotechnol. 53, 376–382.

Elmore, J.R., Furches, A., Wolff, G.N., Gorday, K., Guss, A.M., 2017. Development of a high efficiency integration system and promoter library for rapid modification of *Pseudomonas putida* KT2440. Metab. Eng. Commun. 5, 1–8.

Fan, L.-H., Zhang, Z.-J., Yu, X.-Y., Xue, Y.-X., Tan, T.-W., 2012. Self-surface assembly of

cellulosomes with two miniscaffoldins on *Saccharomyces cerevisiae* for cellulosic ethanol production. Proc. Natl. Acad. Sci. USA 109, 13260–13265.
Gemperlein, K., Hoffmann, M., Huo, L., Pilak, P., Petzke, L., Müller, R., Wenzel, S.C., 2017. Synthetic biology approaches to establish a heterologous production system for coronatines. Metab. Eng. 44, 213–222.
Görke, B., Stülke, J., 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat. Rev. Microbiol. 6, 613–624.
Grant, S.G., Jessee, J., Bloom, F.R., Hanahan, D., 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc. Natl. Acad. Sci. USA 87, 4645–4649.
Guarnieri, M.T., Ann Franden, M., Johnson, C.W., Beckham, G.T., 2017. Conversion and assimilation of furfural and 5-(hydroxymethyl)furfural by *Pseudomonas putida* KT2440. Metab. Eng. Commun. 4, 22–28.
Ha, S.-J., Galazka, J.M., Kim, S.R., Choi, J.-H., Yang, X., Seo, J.-H., Glass, N.L., Cate, J.H.D., Jin, Y.-S., 2011. Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. Proc. Natl. Acad. Sci. USA 108, 504–509.
Hara, K.Y., Kondo, A., 2015. ATP regulation in bioproduction. Microb. Cell Fact. 14, 198.
Henriksson, G., Johansson, G., Pettersson, G., 2000. A critical review of cellobiose dehydrogenases. J. Biotechnol. 78, 93–113.
Herrero, M., de Lorenzo, V., Timmis, K.N., 1990. Transposon vectors containing non-antibiotic resistance selection markers for cloning and stable chromosomal insertion of foreign genes in gram-negative bacteria. J. Bacteriol. 172, 6557–6567.
Horton, R.M., Cai, Z.L., Ho, S.N., Pease, L.R., 1990. Gene splicing by overlap extension: tailor-made genes using the polymerase chain reaction. BioTechniques 8, 528–535.
Hottes, A.K., Freddolino, P.L., Khare, A., Donnell, Z.N., Liu, J.C., Tavazoie, S., 2013. Bacterial adaptation through loss of function. PLoS Genet. 9. http://dx.doi.org/10.1371/journal.pgen.1003617.
Jarmander, J., Hallström, B.M., Larsson, G., 2014. Simultaneous uptake of lignocellulose-based monosaccharides by *Escherichia coli*. Biotechnol. Bioeng. 111, 1108–1115.
Jiménez, J.I., Miñambres, B., García, J.L., Díaz, E., 2002. Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. Environ. Microbiol. 4, 824–841.
Johnson, C.W., Beckham, G.T., 2015. Aromatic catabolic pathway selection for optimal production of pyruvate and lactate from lignin. Metab. Eng. 28, 240–247.
Kajikawa, H., Masaki, S., 1999. Cellobiose transport by mixed ruminal bacteria from a cow. Appl. Environ. Microbiol. 65, 2565–2569.
Kawaguchi, H., Hasunuma, T., Ogino, C., Kondo, A., 2016. Bioprocessing of bio-based chemicals produced from lignocellulosic feedstocks. Curr. Opin. Biotechnol. 42, 30–39.
Kayikci, Ö., Nielsen, J., 2015. Glucose repression in *Saccharomyces cerevisiae*. FEMS Yeast Res. 15. http://dx.doi.org/10.1093/femsyr/fov068.
Kessler, B., Herrero, M., Timmis, K.N., de Lorenzo, V., 1994. Genetic evidence that the XylS regulator of the *Pseudomonas* TOL meta operon controls the Pm promoter through weak DNA-protein interactions. J. Bacteriol. 176, 3171–3176.
Kim, S.M., Choi, B.Y., Ryu, Y.S., Jung, S.H., Park, J.M., Kim, G.-H., Lee, S.K., 2015. Simultaneous utilization of glucose and xylose via novel mechanisms in engineered *Escherichia coli*. Metab. Eng. 30, 141–148.
Lai, B., Yu, S., Bernhardt, P.V., Rabaey, K., Virdis, B., Krömer, J.O., 2016. Anoxic metabolism and biochemical production in *Pseudomonas putida* F1 driven by a bioelectrochemical system. Biotechnol. Biofuels 9, 39.
Lane, S., Zhang, S., Wei, N., Rao, C., Jin, Y.-S., 2015. Development and physiological characterization of cellobiose-consuming *Yarrowia lipolytica*. Biotechnol. Bioeng. 112, 1012–1022.
Lawford, H.G., Rousseau, J.D., 2002. Performance testing of *Zymomonas mobilis* metabolically engineered for cofermentation of glucose, xylose, and arabinose. Appl. Biochem. Biotechnol. 98–100, 429–448.
Le Meur, S., Zinn, M., Egli, T., Thöny-Meyer, L., Ren, Q., 2012. Production of medium-chain-length polyhydroxyalkanoates by sequential feeding of xylose and octanoic acid in engineered *Pseudomonas putida* KT2440. BMC Biotechnol. 12, 53.
Lee, J., Saddler, J.N., Um, Y., Woo, H.M., 2016. Adaptive evolution and metabolic engineering of a cellobiose- and xylose- negative *Corynebacterium glutamicum* that co-utilizes cellobiose and xylose. Microb. Cell Fact. 15, 20.
Li, X., Chomvong, K., Yu, V.Y., Liang, J.M., Lin, Y., Cate, J.H.D., 2015. Cellobionic acid utilization: from *Neurospora crassa* to *Saccharomyces cerevisiae*. Biotechnol. Biofuels 8, 120.
Lieder, S., Nikel, P.I., de Lorenzo, V., Takors, R., 2015. Genome reduction boosts heterologous gene expression in *Pseudomonas putida*. Microb. Cell Fact. 14, 23.
Lien, O.G., 1959. Determination of gluconolactone, galactonolactone, and their free acids by hydroxamate method. Anal. Chem. 31, 1363–1366.
Linger, J.G., Vardon, D.R., Guarnieri, M.T., Karp, E.M., Hunsinger, G.B., Franden, M.A., Johnson, C.W., Chupka, G., Strathmann, T.J., Pienkos, P.T., Beckham, G.T., 2014. Lignin valorization through integrated biological funneling and chemical catalysis. Proc. Natl. Acad. Sci. USA 111, 12013–12018.
Liu, Y., Rainey, P.B., Zhang, X.-X., 2015. Molecular mechanisms of xylose utilization by *Pseudomonas fluorescens*: overlapping genetic responses to xylose, xylulose, ribose and mannitol. Mol. Microbiol. 98, 553–570.
Loeschcke, A., Thies, S., 2015. *Pseudomonas putida*-a versatile host for the production of natural products. Appl. Microbiol. Biotechnol. 99, 6197–6214.
Lundin, A., Thore, A., 1975. Comparison of methods for extraction of bacterial adenine nucleotides determined by firefly assay. Appl. Microbiol. 30, 713–721.
Lynd, L.R., Weimer, P.J., van Zyl, W.H., Pretorius, I.S., 2002. Microbial cellulose utilization: fundamentals and biotechnology. Microbiol. Mol. Biol. Rev. MMBR 66, 506–577.
Martínez-García, E., Aparicio, T., de Lorenzo, V., Nikel, P.I., 2014a. New transposon tools tailored for metabolic engineering of gram-negative microbial cell factories. Front. Bioeng. Biotechnol. 2, 46.
Martínez-García, E., de Lorenzo, V., 2017. Molecular tools and emerging strategies for deep genetic/genomic refactoring of *Pseudomonas*. Curr. Opin. Biotechnol. 47, 120–132.
Martínez-García, E., de Lorenzo, V., 2012. Transposon-based and plasmid-based genetic tools for editing genomes of gram-negative bacteria. Methods Mol. Biol. 813, 267–283.
Martínez-García, E., Nikel, P.I., Aparicio, T., de Lorenzo, V., 2014b. *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. Microb. Cell Fact. 13, 159.
Meijnen, J.-P., de Winde, J.H., Ruijssenaars, H.J., 2012. Metabolic and regulatory rearrangements underlying efficient D-xylose utilization in engineered *Pseudomonas putida* S12. J. Biol. Chem. 287, 14606–14614.
Meijnen, J.-P., de Winde, J.H., Ruijssenaars, H.J., 2008. Engineering *Pseudomonas putida* S12 for efficient utilization of D-xylose and L-arabinose. Appl. Environ. Microbiol. 74, 5031–5037.
Mi, J., Becher, D., Lubuta, P., Dany, S., Tusch, K., Schewe, H., Buchhaupt, M., Schrader, J., 2014. *De novo* production of the monoterpenoid geranic acid by metabolically engineered *Pseudomonas putida*. Microb. Cell Fact. 13, 170.
Moraïs, S., Morag, E., Barak, Y., Goldman, D., Hadar, Y., Lamed, R., Shoham, Y., Wilson, D.B., Bayer, E.A., 2012. Deconstruction of lignocellulose into soluble sugars by native and designer cellulosomes. mBio 3, e00508–e00512.
Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., Ladisch, M., 2005. Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresour. Technol. 96, 673–686.
Muñoz-Gutiérrez, I., Moss-Acosta, C., Trujillo-Martinez, B., Gosset, G., Martinez, A., 2014. Ag43-mediated display of a thermostable β-glucosidase in *Escherichia coli* and its use for simultaneous saccharification and fermentation at high temperatures. Microb. Cell Fact. 13, 106.
Nataf, Y., Yaron, S., Stahl, F., Lamed, R., Bayer, E.A., Scheper, T.-H., Sonenshein, A.L., Shoham, Y., 2009. Cellodextrin and laminaribiose ABC transporters in *Clostridium thermocellum*. J. Bacteriol. 191, 203–209.
Nikel, P.I., Chavarría, M., Fuhrer, T., Sauer, U., de Lorenzo, V., 2015. *Pseudomonas putida* KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and pentose phosphate pathways. J. Biol. Chem. 290, 25920–25932.
Nikel, P.I., de Lorenzo, V., 2014. Robustness of *Pseudomonas putida* KT2440 as a host for ethanol biosynthesis. N. Biotechnol. 31, 562–571.
Nikel, P.I., de Lorenzo, V., 2013. Engineering an anaerobic metabolic regime in *Pseudomonas putida* KT2440 for the anoxic biodegradation of 1,3-dichloroprop-1-ene. Metab. Eng. 15, 98–112.
Nikel, P.I., Martínez-García, E., de Lorenzo, V., 2014. Biotechnological domestication of pseudomonads using synthetic biology. Nat. Rev. Microbiol. 12, 368–379.
Nogales, J., Gudmundsson, S., Duque, E., Ramos, J.L., Palsson, B.O., 2017. Expanding the computable reactome in *Pseudomonas putida* reveals metabolic cycles providing robustness. bioRxiv 139121. http://dx.doi.org/10.1101/139121.
Parisutham, V., Chandran, S.-P., Mukhopadhyay, A., Lee, S.K., Keasling, J.D., 2017. Intracellular cellobiose metabolism and its applications in lignocellulose-based biorefineries. Bioresour. Technol. 239, 496–506.
Poblete-Castro, I., Becker, J., Dohnt, K., dos Santos, V.M., Wittmann, C., 2012. Industrial biotechnology of *Pseudomonas putida* and related species. Appl. Microbiol. Biotechnol. 93, 2279–2290.
Poblete-Castro, I., Binger, D., Rodrigues, A., Becker, J., Martins Dos Santos, V.A.P., Wittmann, C., 2013. *In-silico*-driven metabolic engineering of *Pseudomonas putida* for enhanced production of poly-hydroxyalkanoates. Metab. Eng. 15, 113–123.
Puchałka, J., Oberhardt, M.A., Godinho, M., Bielecka, A., Regenhardt, D., Timmis, K.N., Papin, J.A., Martins dos Santos, V.A.P., 2008. Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. PLoS Comput. Biol. 4, e1000210.
Rojo, F., 2010. Carbon catabolite repression in *Pseudomonas*: optimizing metabolic versatility and interactions with the environment. FEMS Microbiol. Rev. 34, 658–684.
Rutter, C., Mao, Z., Chen, R., 2013. Periplasmic expression of a *Saccharophagus* cellodextrinase enables *E. coli* to ferment cellodextrin. Appl. Microbiol. Biotechnol. 97, 8129–8138.
Sambrook, J., Russell, D.W., 2001. Molecular Cloning: a Laboratory Manual. CSHL Press.
Shin, H.-D., Wu, J., Chen, R., 2014. Comparative engineering of *Escherichia coli* for cellobiose utilization: hydrolysis versus phosphorolysis. Metab. Eng. 24, 9–17.
Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., Páez-Espino, A.D., Durante-Rodríguez, G., Kim, J., Nikel, P.I., Platero, R., de Lorenzo, V., 2013. The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. Nucleic Acids Res. 41, D666–D675.
Singhania, R.R., Patel, A.K., Sukumaran, R.K., Larroche, C., Pandey, A., 2013. Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. Bioresour. Technol. 127, 500–507.
Smith, C.M., Williamson, J.R., 1971. Inhibition of citrate synthase by succinyl-CoA and other metabolites. FEBS Lett. 18, 35–38.
Spiridonov, N.A., Wilson, D.B., 2001. Cloning and biochemical characterization of BglC, a beta-glucosidase from the cellulolytic actinomycete *Thermobifida fusca*. Curr. Microbiol. 42, 295–301.
Stephanopoulos, G., 2007. Challenges in engineering microbes for biofuels production. Science 315, 801–804.
Taha, M., Foda, M., Shahsavari, E., Aburto-Medina, A., Adetutu, E., Ball, A., 2016. Commercial feasibility of lignocellulose biodegradation: possibilities and challenges. Curr. Opin. Biotechnol. 38, 190–197.
Teugjas, H., Väljamäe, P., 2013. Product inhibition of cellulases studied with ^{14}C-labeled cellulose substrates. Biotechnol. Biofuels 6, 104.

Thurston, B., Dawson, K.A., Strobel, H.J., 1993. Cellobiose versus glucose utilization by the ruminal bacterium *Ruminococcus albus*. Appl. Environ. Microbiol. 59, 2631–2637.
Tiso, T., Zauter, R., Tulke, H., Leuchtle, B., Li, W.-J., Behrens, B., Wittgens, A., Rosenau, F., Hayen, H., Blank, L.M., 2017. Designer rhamnolipids by reduction of congener diversity: production and characterization. Microb. Cell Fact. 16, 225.
Tozakidis, I.E.P., Brossette, T., Lenz, F., Maas, R.M., Jose, J., 2016. Proof of concept for the simplified breakdown of cellulose by combining *Pseudomonas putida* strains with surface displayed thermophilic endocellulase, exocellulase and β-glucosidase. Microb. Cell Fact. 15, 103.
Vinuselvi, P., Lee, S.K., 2011. Engineering *Escherichia coli* for efficient cellobiose utilization. Appl. Microbiol. Biotechnol. 92, 125–132.
Wisedchaisri, G., Park, M.-S., Iadanza, M.G., Zheng, H., Gonen, T., 2014. Proton-coupled sugar transport in the prototypical major facilitator superfamily protein XylE. Nat. Commun. 5, 4521.
Yang, M., Luoh, S.M., Goddard, A., Reilly, D., Henzel, W., Bass, S., 1996. The *bglX* gene located at 47.8 min on the *Escherichia coli* chromosome encodes a periplasmic beta-glucosidase. Microbiol. Read. Engl. 142, 1659–1665.
Yim, S.S., Choi, J.W., Lee, S.H., Jeong, K.J., 2016. Modular optimization of a hemicellulose-utilizing pathway in *Corynebacterium glutamicum* for consolidated bioprocessing of hemicellulosic biomass. ACS Synth. Biol. 5, 334–343.
Zhou, S., Davis, F.C., Ingram, L.O., 2001. Gene integration and expression and extracellular secretion of *Erwinia chrysanthemi* endoglucanase CelY (celY) and CelZ (celZ) in ethanologenic *Klebsiella oxytoca* P2. Appl. Environ. Microbiol. 67, 6–14.
Zhu, X., Zhao, D., Qiu, H., Fan, F., Man, S., Bi, C., Zhang, X., 2017. The CRISPR/Cas9-facilitated multiplex pathway optimization (CFPO) technique and its application to improve the *Escherichia coli* xylose utilization pathway. Metab. Eng. 43, 37–45.
Zobel, S., Benedetti, I., Eisenbach, L., de Lorenzo, V., Wierckx, N., Blank, L.M., 2015. Tn7-based device for calibrated heterologous gene expression in *Pseudomonas putida*. ACS Synth. Biol. 4, 1341–1351.

microbial biotechnology

Open Access

Brief report

Biotransformation of D-xylose to D-xylonate coupled to medium-chain-length polyhydroxyalkanoate production in cellobiose-grown *Pseudomonas putida* EM42

Pavel Dvořák[1]* Jozef Kováč[1] and Víctor de Lorenzo[2]

[1]*Department of Experimental Biology (Section of Microbiology), Faculty of Science, Masaryk University, Kamenice 753/5, 62500, Brno, Czech Republic.*

[2]*Systems and Synthetic Biology Program, Centro Nacional de Biotecnología CNB-CSIC, Cantoblanco, Darwin 3, 28049, Madrid, Spain.*

Summary

Co-production of two or more desirable compounds from low-cost substrates by a single microbial catalyst could greatly improve the economic competitiveness of many biotechnological processes. However, reports demonstrating the adoption of such co-production strategy are still scarce. In this study, the ability of genome-edited strain *Pseudomonas putida* EM42 to simultaneously valorize D-xylose and D-cellobiose – two important lignocellulosic carbohydrates – by converting them into the platform chemical D-xylonate and medium-chain-length polyhydroxyalkanoates, respectively, was investigated. Biotransformation experiments performed with *P. putida* resting cells showed that promiscuous periplasmic glucose oxidation route can efficiently generate extracellular xylonate with a high yield. Xylose oxidation was subsequently coupled to the growth of *P. putida* with cytoplasmic β-glucosidase BglC from *Thermobifida fusca* on D-cellobiose. This disaccharide turned out to be a better co-substrate for xylose-to-xylonate biotransformation than monomeric glucose. This was because unlike glucose, cellobiose did not block oxidation of the pentose by periplasmic glucose dehydrogenase Gcd, but, similarly to glucose, it was a suitable substrate for polyhydroxyalkanoate formation in *P. putida*. Co-production of extracellular xylose-born xylonate and intracellular cellobiose-born medium-chain-length polyhydroxyalkanoates was established in proof-of-concept experiments with *P. putida* grown on the disaccharide. This study highlights the potential of *P. putida* EM42 as a microbial platform for the production of xylonate, identifies cellobiose as a new substrate for mcl-PHA production, and proposes a fresh strategy for the simultaneous valorization of xylose and cellobiose.

Received 8 January, 2019; revised 5 March, 2020; accepted 25 March, 2020.
*For correspondence. E-mail pdvorak@sci.muni.cz; Tel. +420 549 493 396; Fax +420 541211214.
Microbial Biotechnology (2020) **0**(0), 1–11
doi:10.1111/1751-7915.13574

Funding information

This work was funded by the SETH Project of the Spanish Ministry of Science (RTI 2018-095584-B-C42), MADONNA (H2020-FET-OPEN-RIA-2017-1-766975), BioRoboost (H2020-NMBP-BIO-CSA-2018) and SYNBIO4FLAV (H2020-NMBP/0500), contracts of the European Union and the S2017/BMD-3691 InGEMICS-CM funded by the Comunidad de Madrid (European Structural and Investment Funds) as well as by the Czech Science Foundation (19-06511Y).

Introduction

Up to 220 million tonnes of lignocellulosic and cellulosic waste are potentially available for biotechnological purposes only in the EU every year (Searles and Malins, 2013). Lignocellulose can be decomposed by physical or chemical pre-treatment to cellulose, hemicellulose and lignin, and these fractions can be further hydrolysed enzymatically to monomeric sugars and lignin-derived aromatics serving as cheap substrates for microbial fermentations and biosynthesis of value-added chemicals (VAC; Mosier *et al.*, 2005; Kawaguchi *et al.*, 2016). Economics of these bioprocesses is regrettably still often unsatisfactory but can be significantly improved by parallel valorization of two or more lignocellulosic substrates. This is allowed by co-streaming of carbon from several sources into a single-valued compound or by simultaneous production of two or more VAC (Dumon *et al.*, 2012; Li *et al.*, 2017; Larroude *et al.*, 2018; Baral *et al.*, 2019; Wang *et al.*, 2019). Co-production of extracellular and intracellular biochemicals is desirable for facilitated downstream processing (Wang *et al.*, 2019). However, studies reporting such parallel biomanufacturing of two

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

VAC from the second-generation carbon sources are infrequent and well-defined cell factories that could efficiently perform these tasks are scarce.

The soil bacterium and growingly used robust platform strain *P. putida* KT2440 can naturally assimilate a spectrum of aromatic compounds and organic acids but only a few plant biomass-derived carbohydrates: glucose, mannose and fructose (Linger *et al*., 2014; Belda *et al*., 2016; Nikel and de Lorenzo, 2018; Jayakody *et al*., 2018). Its metabolism was engineered to reach out to other sugars, including carbohydrates typically produced upon (hemi)cellulose hydrolysis or pyrolysis (Meijnen *et al*., 2008; Linger *et al*., 2016; Löwe *et al*., 2018). In a recent work, *P. putida* EM42, a *P. putida* KT2440-derived strain with streamlined genome and better physiological properties (Martínez-García *et al*., 2014), was empowered with a xylose transporter and isomerase pathway from *Escherichia coli* along with a cytoplasmic β-glucosidase BglC from *Thermobifida fusca* (Dvořák and de Lorenzo, 2018). This allowed the resulting strain to co-utilize and grow on mixtures of D-glucose, D-cellobiose and D-xylose. However, the mix of carbohydrates was metabolized and converted into CO_2 and biomass without any other return.

There are various possibilities to use *P. putida* for VAC biomanufacturing from glucose and cellobiose (Poblete-Castro *et al*., 2012; Loeschcke and Thies, 2015). *P. putida* KT2440 has been traditionally employed as a model organism for the production of medium-chain-length polyhydroxyalkanoates (mcl-PHA), biodegradable polyesters applicable for manufacturing of packaging materials, textile or medical implants (Chen, 2009; Prieto *et al*., 2016; Li *et al*., 2017). The mcl-PHA have better elastomeric properties and broader application potential than short-chain-length PHA produced by *Cupriavidus necator* or recombinant *E. coli* (Chen, 2009). Synthesis of mcl-PHA was demonstrated from fatty acids and unrelated substrates such as acetate, ethanol, glycerol or some sugars including glucose (Prieto *et al*., 2016) but never from cellodextrins such as cellobiose. In a previous study, we also identified the ability of *P. putida* EM42 to oxidize D-xylose to D-xylonate, a platform molecule of considerable biotechnological interest (Werpy and Petersen, 2004; Toivari *et al*., 2012a,b; Mehtiö *et al*., 2016; Dvořák and de Lorenzo, 2018). D-xylonate was reported to be used as a complexing agent or chelator, as a precursor of polyesters, 1,2,4-butanetriol, ethylene glycol or glycolate, and it can serve as a cheap, non-food-derived alternative for D-gluconic acid (Toivari *et al*., 2012a,b). Xylonate is naturally formed in the first step of oxidative metabolism of xylose by some archaea, bacteria and fungi *via* the action of D-xylose or D-glucose dehydrogenases. Production of xylonate was reported for instance in *Gluconobacter oxydans*, in several *Pseudomonas* strains including *P. fragi, P. taiwanensis or P. putida S12,* or in *Klebsiella pneumoniae (*Buchert *et al*., 1988; Meijnen *et al*., 2008; Köhler *et al*., 2015; Wang *et al*., 2016). Several other microorganisms including *Escherichia coli* or *Saccharomyces cerevisiae* were engineered for xylonate production from xylose (Nygård *et al*., 2011; Liu *et al*., 2012; Toivari *et al*., 2012a,b; Gao *et al*., 2019). High production costs nonetheless hinder commercialization of both xylonate and mcl-PHA, and new solutions are appealing for easing the biomanufacture of these chemicals (Chen, 2009; Toivari *et al*., 2012a,b; Mehtiö *et al*., 2016; Li *et al*., 2017). Their co-production from the second-generation carbon sources can thus be a promising approach in this context.

We present below our efforts to merge the advantages of *P. putida* EM42 as a natural xylonate producer with the ability of an engineered variant to grow on cellulose-derived substrate. Our results confirm that *P. putida* EM42 can convert xylose to xylonate with a high yield with its periplasmic glucose oxidative pathway and release the acid in the medium (Fig. 1). Furthermore, we show that xylonate production is inhibited in the presence of glucose but does occur in the cellobiose-grown recombinant strain. Most importantly, we demonstrate that periplasmic production and release of xylonate by cellobiose-grown *P. putida* EM42 are accompanied by parallel accumulation of mcl-PHA in the cells.

Results and discussion

Biotransformation of xylose to xylonate by P. putida *EM42 resting cells*

Periplasmic xylose conversion to xylonate was previously identified as a competing reaction for xylose assimilation by recombinant *P. putida* EM42 during a five-day cultivation experiment (Dvořák and de Lorenzo, 2018). Periplasmic glucose dehydrogenase was shown to be a crucial component for xylose oxidation in our strain as well as in several xylonate-producing bacteria including *Klebsiella pneumoniae* and some other pseudomonads (Hardy *et al*., 1993; Meijnen *et al*., 2008; Köhler *et al*., 2015; Wang *et al*., 2016; Dvořák and de Lorenzo, 2018; Bator *et al*., 2020). In *P. putida* KT2440, and correspondingly also in strain EM42, membrane-bound glucose dehydrogenase Gcd (PP1444) oxidizes xylose to xylonolactone with pyrroloquinoline quinone (PQQ) as a cofactor. Lactone can then open spontaneously in the presence of water or might be converted to xylonate with the help of gluconolactonase Gnl (PP1170; Fig. 1). Neither xylose nor xylonate is utilized for biomass formation (Dvořák and de Lorenzo, 2018; Bator *et al*., 2020).

Here, we initially tested whether xylose can be oxidized to xylonate in a short time interval and with a high

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

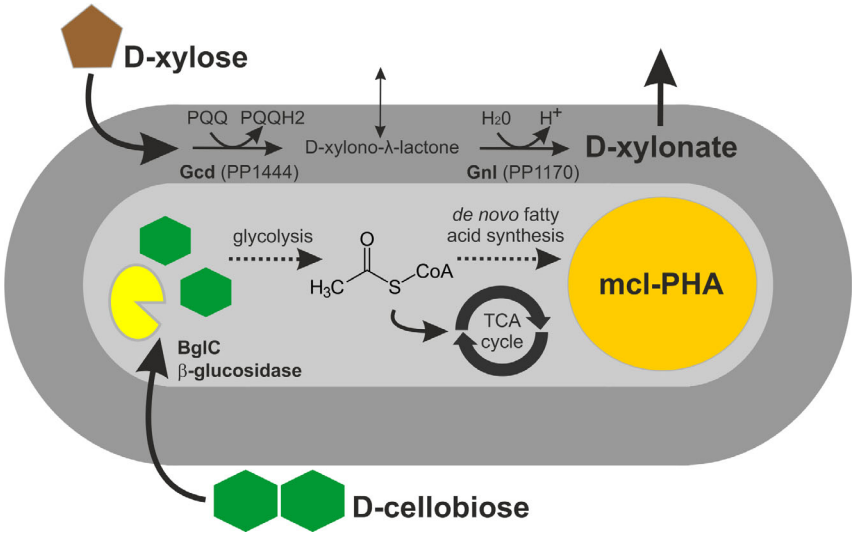


Fig. 1. Co-production of D-xylonate and medium-chain-length polyhydroxyalkanoates from D-xylose and D-cellobiose, respectively, in *bglC* + *P. putida* EM42. Innate periplasmic oxidative route and introduced cytoplasmic β-glucosidase BglC from *Thermobifida fusca* allow simultaneous valorization of D-xylose and D-cellobiose in *Pseudomonas putida* EM42. D-xylose is oxidized to platform chemical D-xylonate which is released into the medium. D-cellobiose, on the other hand, is transported into the cell, cleaved in two D-glucose molecules by BglC and gives rise to acetyl-CoA, a precursor molecule for the production of intracellular biopolymers (polyhydroxyalkanoates, PHA) *via de novo* fatty acid synthesis in nitrogen-limited conditions. Periplasmic space and cytoplasm are shown in dark and pale grey, respectively. Abbreviations: Gcd, glucose dehydrogenase; Gnl, gluconolactonase; PQQ and PQQH, pyrroloquinoline quinone and its reduced form, respectively; TCA cycle, tricarboxylic acid cycle; mcl-PHA, medium-chain-length polyhydroxyalkanoates.

yield by *P. putida* resting cells of defined optical density. Xylonolactone concentrations were newly determined in culture supernatants using the hydroxamate method (Lien, 1959), which allowed more precise quantification of xylonate than in our previous work (Dvořák and de Lorenzo, 2018). *P. putida* EM42 cells (strains and plasmids used in this study are listed in Table S1), pre-cultured in lysogeny broth (LB), washed and diluted to a starting $OD_{600} \sim 0.5$, were incubated for 48 h in M9 minimal medium with 5 g l^{-1} xylose (Fig. 2A). The yield of extracellular xylonate detected in the medium at the end of the incubation was 0.95 g per g of xylose which was 86% of the theoretical maximum 1.11 g g^{-1} (molar mass of D-xylose and D-xylonate is 150.13 and 166.13 g mol^{-1}, respectively). Lactone accumulated in small quantities (up to 0.45 g l^{-1}) in the medium during the initial phase of fast xylose conversion, but its concentration then declined to zero at the end of the experiment. The release of sugar acid was accompanied by a pH drop in the medium from the initial 7.00 ± 0.00 to 6.15 ± 0.04 at the end of the reaction. Neither lactone nor xylonate was detected in the identical experiment repeated with *P. putida* EM42 Δ*gcd* mutant lacking glucose dehydrogenase (Fig. 2B). These experiments confirmed the importance of Gcd for D-xylose oxidation to xylonate in *P. putida* EM42 and showed that xylonolactone intermediate is converted rapidly to xylonate which is released into the medium rather than utilized by the cells. In contrast, a study with *P. fragi* (the best-described pseudomonad in terms of xylonate production thus far) reported slow spontaneous hydrolysis and accumulation of inhibitory xylonolactone in this bacterium during the early phases of fermentation experiments (Buchert *et al*., 1986; Buchert and Viikari, 1988). Another well-characterized xylose-oxidizing pseudomonad, *P. taiwanensis* VLB120, uses xylonate for biomass formation (Köhler *et al*., 2015). *P. putida* thus represents an attractive addition to these strains for fast high-yield production of extracellular xylonate.

It is worth noting that the resting *P. putida* cells could be recycled and used repeatedly in five cycles of xylose (5 g l^{-1}) oxidation to xylonate (Fig. S1). The conversion reached 94% in the first cycle, then decreased and reached 60% in the last fifth cycle. As the optical density of the cells measured at the end of each cycle continuously decreased, the decline in productivity can be attributed mainly to the loss of the biomass in the reactions due to the centrifugation/re-suspension cycles and cell lysis (Fig. S1). Medium pH drop detected at the end of each cycle corresponded with the level of xylose-to-xylonate conversion (Fig. S1). This result indicates that the xylose oxidation in *P. putida* EM42 is not necessarily growth-dependent as reported with *P. fragi* (Buchert *et al*., 1986; Buchert and Viikari, 1988). It is noteworthy that a number of studies on microbial xylonate production have reported the association of xylose oxidation to

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

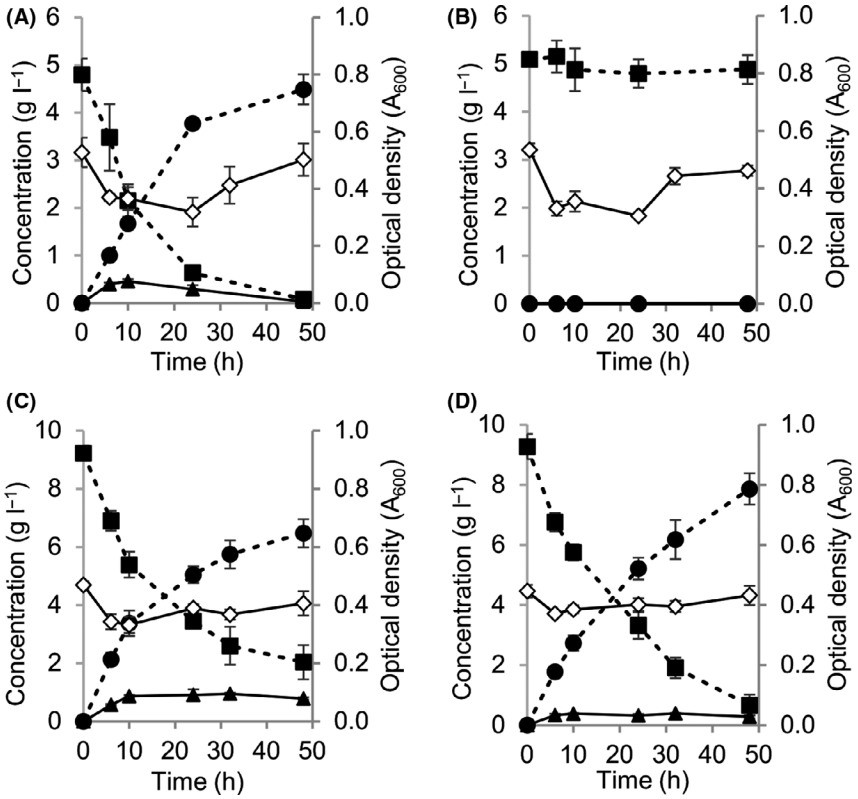


Fig. 2. Biotransformation of xylose to xylonate by *P. putida* EM42 resting cell. Incubation of resting cells of (A) *Pseudomonas putida* EM42 and (B) its deletion mutant *P. putida* EM42 Δ*gcd* in minimal medium with 5 g l^{-1} D-xylose. Experiments were carried out in 25 ml of M9 minimal medium in flasks shaken at 170 r.p.m. and 30°C. (C) Incubation of *P. putida* EM42 resting cells in 25 ml of minimal medium with 10 g l^{-1} D-xylose in flasks shaken at 170 r.p.m. and 30°C. (D) Incubation of *P. putida* EM42 resting cells in 25 ml of buffered M9 minimal medium (100 mM sodium phosphate buffer) with 10 g l^{-1} D-xylose in flasks shaken at 300 r.p.m. and 30°C. In all experiments, minimal medium was inoculated to the initial A_{600} of 0.5 using cells obtained from an overnight culture in lysogeny broth. D-xylose, filled squares (■); D-xylonate, filled circles (●); D-xylono-λ-lactone, filled triangles (▲); cell biomass, open diamonds (◊). Data points shown as mean ± SD of three biological replicates.

a host's growth (Toivari *et al*., 2012a,b; Köhler *et al*., 2015; Wang *et al*., 2016) but some have not. One example of the latter is recent work by Zhou *et al.* (2017) on *G. oxydans*, which could be used repeatedly for xylonate production in a bioreactor with an improved oxygen delivery system. Such cell recycling can be a promising strategy offering high xylonate yield and reduced process costs.

Since oxygen availability may become a bottleneck for the xylose-to-xylonate conversion, we next examined the effect of improved aeration through increased agitation of *P. putida* resting cells. In Zhou *et al.* (2017), the increase in agitation speed from 300 to 500 rpm enhanced the accumulation of xylonate by 25%. To check whether we could observe the same trend, we incubated resting cells in minimal medium with 5 g l^{-1} xylose at agitation of 170 or 300 r.p.m. and the level of xylose conversion to xylonate was determined after 48 h (Fig. S2A). Xylose oxidation to lactone and xylonate was 8% more efficient in flasks agitated at higher speed, but the increase was only marginal. Another variable tested was pH. Xylonate accumulation results in acidification of the medium, and low pH can inhibit the activity of glucose dehydrogenase, as shown previously for *P. fragi* (Buchert *et al*., 1986). To inspect the effect of pH, we increased the buffering capacity of the M9 medium by mixing it with 100 mM sodium phosphate buffer while escalating xylose concentration to 10 g l^{-1} to intensify acidification. In these conditions, EM42 cells gave rise to ~ 12% more oxidized product after 48 h than cells in non-buffered cultures (Fig. S2B). The final pH values determined in buffered and non-buffered cultures (5.92 ± 0.03 and 4.53 ± 0.12, respectively) proved that the sodium phosphate buffer of used 100 mM concentration could efficiently prevent excessive pH drop. The joint effect of modified reaction conditions is shown in Figure 2C and D which depict time courses of xylose conversion to xylonate by *P. putida* EM42 resting cells incubated in non-buffered minimal medium in flasks shaken at 170 r.p.m. (Fig. 2C) and in buffered medium in flasks shaken at 300 r.p.m. (Fig. 2D). In the latter case, the xylonate yield reached 0.85 ± 0.06 g g^{-1} xylose

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

after 48-h time interval (Table 1) while the yield in the reaction with unmodified conditions was only 0.70 ± 0.07 g g^{-1}. The major visible effects of intense agitation and medium buffering were reduced accumulation of xylonolactone and faster xylose-to-xylonate conversion especially during the last 24 h of the reaction in which a pronounced pH drop was prevented (the final pH values in buffered and non-buffered cultures were 4.45 ± 0.01 and 5.99 ± 0.08, respectively). These observations on culture conditions were considered for increasing the efficiency of xylose conversion to xylonate also in subsequent experiments with growing *P. putida* cells.

Xylose biotransformation to xylonate by P. putida *EM42 growing on glucose or cellobiose*

In none of the experiments mentioned above, xylose oxidation to xylonate was tested during growth. Instead, the transformation experiments were preceded by the production of whole-cell catalyst biomass. Similarly to other naturally occurring or recombinant xylonate producers (Nygård *et al*., 2011; Toivari *et al*., 2012a,b; Wang *et al*., 2016; Zhou *et al*., 2017; Gao *et al*., 2019), *P. putida* was grown in a medium rich in amino acids and vitamins, namely in LB (La Rosa *et al*., 2016). However, such complex media are expensive and thus unsuitable for large-scale bioprocesses. As an alternative, the growth of xylonate-producing microorganism on low-cost carbon source derived, e.g. from lignocellulosic materials, would be desirable. D-glucose is the most abundant monomeric sugar in lignocellulosic hydrolysates prepared by using commercial enzyme cocktails with endoglucanase, exoglucanase and β-glucosidase (Taha *et al*., 2016), and it is also a good growth substrate for *P. putida* (del Castillo *et al*., 2007; Nikel *et al*., 2015; Sasnow *et al*., 2016; Kukurugya *et al*., 2019; Kohlstedt and Wittmann, 2019). However, glucose is a preferred substrate for glucose dehydrogenase and might thus inhibit xylose oxidation by this enzyme. Figure 3A shows that this is exactly the case, *P. putida* EM42 cultured in minimal medium with 5 g l^{-1} glucose did not oxidize xylose during the first eight hours of the experiment, i.e. when glucose was consumed by the cells. As a consequence, the production of xylonate (which occurred concomitantly with growth) was postponed and less than 20% of xylose was oxidized to the acid at the end of the two-day culture (Table 1). Inhibition of xylose transformation to xylonate by glucose was confirmed in an additional experiment using an increased concentration of the hexose (10 g l^{-1}, Fig. S3).

We attempted to bypass this bottleneck by employing D-cellobiose as an alternative growth substrate for *P. putida*. D-cellobiose is a disaccharide composed of two β-glucose monomers linked by a β(1 → 4) bond. It is a by-product of cellulose saccharification with standard commercial mixtures of cellulases but becomes a predominant product when β-glucosidase is omitted from

Table 1. Parameters determined in the cultures with *Pseudomonas putida* EM42 resting cells and in the cultures with *P. putida* EM42 or *P. putida* EM42 pSEVA2213_*bglC* grown on D-glucose or D-cellobiose, respectively, and transforming D-xylose to D-xylonate.

P. putida strain and culture conditions	μ (h^{-1})[e]	Xylonate yield (g g^{-1} xylose)	Xylonate productivity (mg l^{-1} h^{-1})	CDW (g)	pH
EM42 resting cells opt[a]	n.a.	$0.56 \pm 0.06/0.85 \pm 0.06$	$217 \pm 15/164 \pm 11$	$0.15 \pm 0.01/0.16 \pm 0.01$	$6.30 \pm 0.01/5.99 \pm 0.08$
EM42 glucose	0.58 ± 0.02	$0.18 \pm 0.02/0.17 \pm 0.03$	$69 \pm 6/32 \pm 6$	$1.73 \pm 0.12/1.60 \pm 0.06$	$6.26 \pm 0.01/5.81 \pm 0.03$
bglC$^+$ EM42 cellobiose	0.27 ± 0.03	$0.30 \pm 0.06/0.48 \pm 0.09$	$114 \pm 18/93 \pm 13$	$1.46 \pm 0.13/1.88 \pm 0.05$	$6.16 \pm 0.06/5.19 \pm 0.03$
bglC$^+$ EM42 cellobiose optA[b]	0.28 ± 0.05	$0.34 \pm 0.05/0.54 \pm 0.10$	$138 \pm 21/107 \pm 17$	$2.06 \pm 0.07/1.85 \pm 0.03$	$6.23 \pm 0.07/5.88 \pm 0.05$
bglC$^+$ EM42 cellobiose optB[c]	0.30 ± 0.02	$0.35 \pm 0.02/0.50 \pm 0.01$	$144 \pm 8/102 \pm 3$	$2.41 \pm 0.11/2.18 \pm 0.07$	$6.27 \pm 0.08/6.00 \pm 0.04$
bglC$^+$ EM42 cellobiose PHA[d]	0.24 ± 0.01	$0.41 \pm 0.09/0.52 \pm 0.08$	$156 \pm 32/99 \pm 13$	$0.86 \pm 0.02/1.24 \pm 0.14$	$6.24 \pm 0.02/5.90 \pm 0.01$

Values represent the mean ± standard deviation of three biological replicates. Parameters (except for μ) were determined after 24 h/48 h of the culture. CDW, cell dry weight; n.a., not applicable.

a. Resting cells, pre-cultured in LB medium, were incubated in flasks with M9 minimal medium buffered with 100 mM sodium phosphate buffer and shaken at 300 r.p.m.

b. Cultures, inoculated from pre-cultures grown in LB medium, were carried out in flasks with M9 minimal medium buffered with 100 mM sodium phosphate buffer and shaken at 300 r.p.m.

c. Cultures, inoculated from pre-cultures grown in M9 medium with D-cellobiose, were carried out in flasks with M9 minimal medium buffered with 100 mM sodium phosphate buffer and shaken at 300 r.p.m.

d. Cultures were carried out in flasks with M9 minimal medium with reduced content of nitrogen, buffered with 100 mM sodium phosphate buffer and shaken at 300 r.p.m.

e. The specific growth rate (μ) was determined during exponential growth.

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

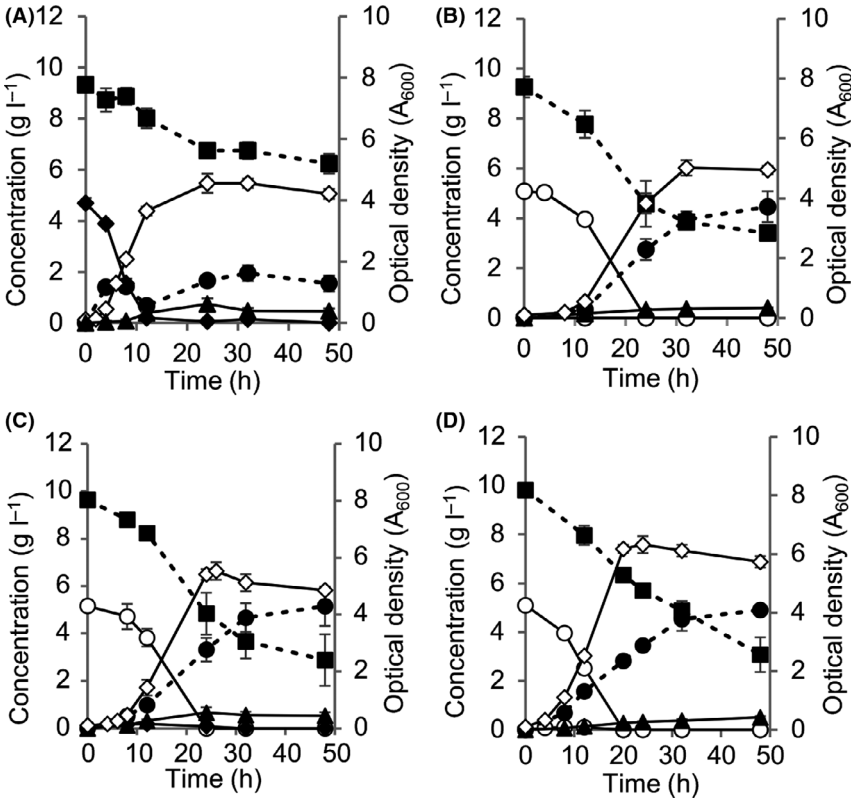


Fig. 3. Biotransformation of D-xylose to D-xylonate by *P. putida* EM42 growing on D-glucose or D-cellobiose. Two-day cultures of (A) *Pseudomonas putida* EM42 in minimal medium with 10 g l^{-1} D-xylose and 5 g l^{-1} D-glucose used as a sole carbon source for growth. (B,C,D) Cultures of *Pseudomonas putida* EM42 pSEVA2213_*bglC* in minimal medium with 10 g l^{-1} D-xylose and 5 g l^{-1} D-cellobiose used as a sole carbon source. Experiments (A) and (B) were carried out in 25 ml of minimal medium in flasks shaken at 170 r.p.m. and 30°C. Minimal medium was inoculated to the initial A_{600} of 0.1 using cells obtained from an overnight culture in lysogeny broth. Experiments (C) and (D) were performed in flask with 25 ml of minimal medium buffered with 100 mM sodium phosphate buffer and shaken at 300 r.p.m. (30°C). Cells used for inoculation of the main culture to the initial A_{600} of 0.1 were pre-grown overnight in lysogeny broth (C) or in minimal medium with 5 g l^{-1} D-cellobiose (D). D-xylose, filled squares (■); D-xylonate, filled circles (●); D-xylono-λ-lactone, filled triangles (▲); D-glucose, filled diamonds (♦); D-cellobiose, open circles (○); cell biomass, open diamonds (◇). Data points shown as mean ± SD of three biological replicates. Please note that the elevated xylonate concentrations detected after 4 and 8 h in the culture (A) do not reflect the real levels of the xylose oxidation product. Hydroxamate method (Lien, 1959) used here for xylonate quantification was originally designed for the detection of gluconate and its lactone, which temporarily accumulated in the culture medium during glucose utilization in (A). Accumulation of gluconate at the times 4 and 8 h was verified also by the specific D-Gluconic Acid/ D-Glucono-δ-lactone Assay Kit (Megazyme, data not shown).

the cocktail (Chen, 2015). Well-defined microbial hosts capable of efficient cellobiose utilization are therefore desirable because they can be applied in simultaneous saccharification and fermentation of cellulose for production of VAC while the process cost is reduced as addition of expensive β-glucosidase is not needed (Ha *et al*., 2011; Chen, 2015; Parisutham *et al*., 2017).

Previous work revealed that a recombinant *P. putida* EM42 derivative which expressed β-glucosidase gene *bglC* from *T. fusca* grew rapidly on D-cellobiose as a sole carbon source (Fig. 1; Dvořák and de Lorenzo, 2018). In this case, cellobiose enters *P. putida* cells through the glucose ABC transporter and it is then cleaved by BglC to two glucose molecules which are further processed in the cytoplasm. The peripheral glucose oxidative pathway probably does not play a role in cellobiose uptake. Hence, it was presumed that cellobiose could be used instead of glucose as a growth substrate for *P. putida* while xylose would be oxidized by non-occupied Gcd (Fig. 1). To test this hypothesis, we cultured *P. putida* EM42 pSEVA2213_*bglC* in minimal medium with 5 g l^{-1} cellobiose and 10 g l^{-1} xylose. Cellobiose was consumed within the initial 24 h of the culture under conditions described in the legend of Figure 3. No glucose was detected in the medium. During the same time interval, 2.75 ± 0.42 g l^{-1} of xylonate was produced from xylose with average volumetric productivity 114 mg l^{-1} h^{-1} which was 65% higher than in the culture on glucose (Table 1). Xylose oxidation was fastest during the initial 32 h of the exponential growth phase and then slowed down in the stationary phase. Xylonate yield at the end of the two-day experiment was 0.48 ± 0.09 g g^{-1} xylose. Minor quantities of xylonolactone were detected in supernatant during the whole

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

course of the culture (Fig. 3B). Xylonate production and cellular growth were accompanied by acidification of the medium: the pH decreased from 7.00 ± 0.00 to 6.16 ± 0.06 and 5.19 ± 0.03 after 24 and 48 h of culture, respectively (Table 1).

The xylonate productivity after initial 24 h of the exponential growth further increased 1.21-fold (to 138 mg l^{-1} h^{-1}) when the *bglC*$^{+}$ *P. putida* EM42 strain was cultured in the modified conditions used previously with resting cells (100 mM sodium phosphate buffer and 300 r.p.m.; Fig. 3C and Table 1). Then, the cells entered the stationary growth phase and xylonate production during the additional 24 h of culture was comparable with the former experiment with cells grown in standard M9 medium at 170 r.p.m. Pre-growing the cells in M9 minimal medium with 5 g l^{-1} cellobiose reduced lag phase of the main culture and improved biomass yield but did not help with increasing the xylonate yield and productivity (Fig. 3D, Table 1). We argue that suboptimal oxygen supply in shake flasks might be the limiting factor preventing efficient xylose oxidation by dense culture in the stationary period. In any case, these experiments indicate that cellobiose, an abundant cellulosic carbohydrate, does not inhibit xylose oxidation to xylonate in *P. putida* and can thus be used as a growth substrate for cells performing this biotransformation.

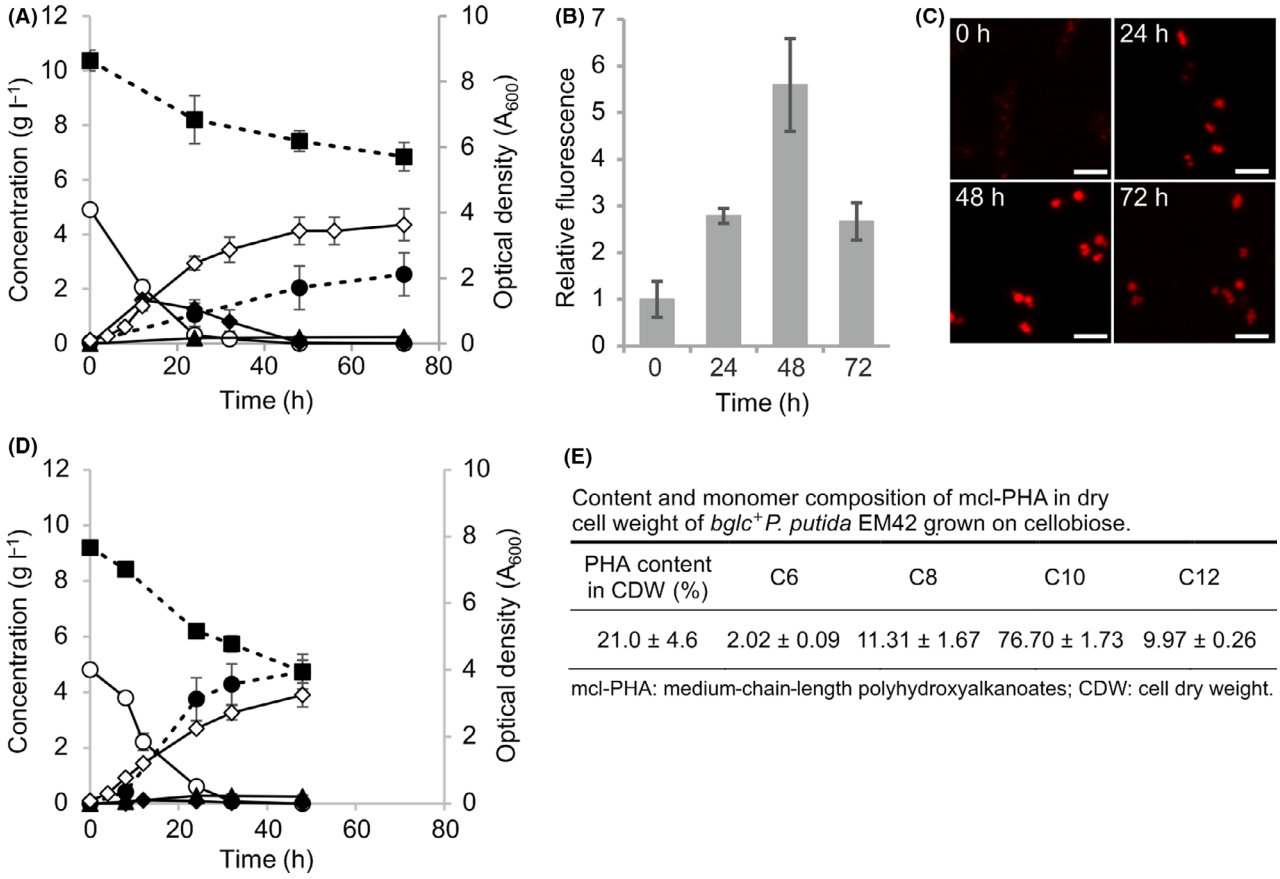
Co-production of xylonate and PHA by P. putida *EM42 grown on cellobiose*

The ability of *P. putida* to both metabolize cellobiose in the cytoplasm and oxidize xylose by the periplasmic pathway paved the way for parallel co-production of the two biotechnologically relevant compounds – xylonate and mcl-PHA. The mcl-PHA have been reported to be co-produced with alginate oligosaccharides from glucose or glycerol (Guo *et al*., 2011; Licciardello *et al*., 2017) or with rhamnolipids from fatty acids (Hori *et al*., 2011). Also, D-xylonate was generated simultaneously with xylitol or bioethanol from xylose and glucose (Wiebe *et al*., 2015; Zhu *et al*., 2019). However, the synthesis of mcl-PHA along with the release of xylonate has not yet been reported. To this end, we first examined the formation of PHA granules in cellobiose-grown *P. putida* cells. As shown in Figure S4, flow cytometry and confocal microscopy identified PHA in the bacteria (Experimental procedures and Results and discussion in Supporting information).

This simple test indicated that *P. putida* EM42 *bglC*$^{+}$ metabolized cellobiose to the monomeric glucose, then to acetyl-CoA and next channelled this metabolic intermediate towards the formation of the polymer. In order to verify that PHA could be generated along with xylonate production, the *bglC*$^{+}$ strain was pre-cultured in nitrogen-rich LB medium (to avoid any PHA accumulation) and then grown in nitrogen-limited M9 medium with 100 mM sodium phosphate buffer, 5 g l^{-1} cellobiose and 10 g l^{-1} xylose (Fig. 4). Sugar and xylonate concentrations were determined in culture supernatants while intracellular PHA formation was followed by flow cytometry and confocal microscopy. As shown in Figure 4A–C, xylonate and PHA were produced simultaneously during the initial 48 h of the three-day experiment. Cellular polymer content increased during the first two days and then declined towards the end of the experiment (Fig. 4B and C). This trend correlated with the presence of the carbon source (cellobiose and glucose) in the medium (Fig. 4A). As in previous experiments, cellobiose was almost completely consumed within the initial 24 h. However, uptake of the disaccharide was this time accompanied by the appearance of glucose in the medium, which reached its maximum concentration (1.61 ± 0.37 g l^{-1}) at 12 h of the culture. Minute quantities of extracellular gluconate were detected as well (≤ 0.1 g l^{-1}, data not shown). Under these circumstances, it became apparent that the secreted glucose affected xylose oxidation by Gcd; only ~ 25% of the pentose was converted to xylonate at the end of the experiment. Although we do not have a trivial explanation for such unexpected release of glucose, we speculate that it could be due to [i] slower growth (μ = 0.19 ± 0.01 h^{-1}) under nitrogen limitation as compared to the standard M9 medium (μ = 0.30 ± 0.02 h^{-1}; Fig. 3C) and/or [ii] an imbalance between the knocked-in BglC β-glucosidase and the innate Glk glucokinase (PP1011) activities stemming from the difference in composition of pre-culture (LB) and culture (M9 with cellobiose) medium (see a scrutiny of these possibilities in supplementary Results and discussion, Experimental procedures and Fig. S5).

To overcome this bottleneck, *P. putida* EM42 *bglC*$^{+}$ cells were pre-grown overnight in standard M9 medium with 5 g l^{-1} cellobiose instead of LB. Faster growth of the main cultures (μ = 0.24 ± 0.01 h^{-1}) in the nitrogen-limited M9 medium with cellobiose and xylose was then indeed observed, and only minute concentrations of glucose (up to 0.12 g l^{-1}) were detected in the supernatants during the first 24 h (Fig. 4D). As a consequence, the volumetric productivity of xylonate during this period increased 3.5-fold (from 44 ± 18 mg l^{-1} h^{-1} to 156 ± 32 mg l^{-1} h^{-1}) when compared with the previous experiment shown in Fig. 4A. Xylonate yield was 2.4 times higher and reached 0.52 ± 0.08 g g^{-1} xylose after 48 h of the culture (Table 1). Interestingly, the xylonate yield per gram of cell dry weight was 1.7-fold higher compared to the cells growing faster and reaching higher OD_{600} in M9 medium with standard nitrogen content (Fig. 3C, Table 1).

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.



(E)

Content and monomer composition of mcl-PHA in dry cell weight of *bglc*$^{+}$*P. putida* EM42 grown on cellobiose.

PHA content in CDW (%)	C6	C8	C10	C12
21.0 ± 4.6	2.02 ± 0.09	11.31 ± 1.67	76.70 ± 1.73	9.97 ± 0.26

mcl-PHA: medium-chain-length polyhydroxyalkanoates; CDW: cell dry weight.

Fig. 4. Co-production of D-xylonate and PHA from D-xylose and D-cellobiose, respectively, by cellobiose-grown *P. putida* EM42 pSEVA2213_*bglC*. (A) Initial culture inoculated from overnight pre-culture in lysogeny broth was carried out in 25 ml of nitrogen-limited M9 minimal medium with 100 mM sodium phosphate buffer, 5 g l^{-1} cellobiose and 10 g l^{-1} xylose in flasks shaken at 300 r.p.m. and 30°C. (B) Relative fluorescence of bacterial population analysed by flow cytometry every 24 h during the three-day culture. Cells were stained by Nile Red and processed as described in Supplementary Information. (C) Confocal microscopy of *P. putida* cells collected at denoted time intervals. Stained bacteria were processed as described in Supplementary Information. White scale bars show 2 μm distance. (D) Culture inoculated from overnight pre-cultures in M9 minimal medium with 5 g l^{-1} cellobiose was carried out in the same conditions as were described for (A). (E) Content and monomer composition of medium-chain-length polyhydroxyalkanoates in cell dry weight of *P. putida* EM42 pSEVA2213_*bglC* cells collected at the end of the two-day culture (graph D). D-xylose, filled squares (■); D-xylonate, filled circles (●); D-xylono-λ-lactone, filled triangles (▲); D-glucose, filled diamonds (♦); D-cellobiose, open circles (○); cell biomass, open diamonds (◇). Data points and columns in (A), (B) and (D) show mean ± SD of three biological replicates.

The same cultures were stopped after 48 h to quantify also PHA content within the cells which turned out to be 21 % (w/w) of cell dry weight. The biopolymer yield was 0.05 ± 0.01 g g^{-1} cellobiose. These values are close to those reported for *P. putida* KT2440 grown on glucose (Huijberts *et al*., 1992; Poblete-Castro *et al*., 2013). The PHA titre in the shake flask reached 0.26 ± 0.03 g l^{-1}. The monomer composition of the analysed biopolymer was also consistent with the previous reports on mcl-PHA production from glucose (Fig. 4E). The major fraction (> 75%) was formed by 3-hydroxydecanoate, followed by 3-hydroxyoctanoate, 3-hydroxydodecanoate and small amount of 3-hydroxyhexanoate. Taken together, the above experiments confirmed the co-production of two value-added molecules (xylonate and mcl-PHA) out of xylose and cellobiose in *P. putida*.

Conclusion

In this work, we have exploited the metabolic versatility of *P. putida* EM42, a robust derivative of *P. putida* KT2440, for prototyping the simultaneous conversion of xylose and cellobiose into xylonate and mcl-PHA. Periplasmic oxidation of D-xylose to D-xylonate was first assayed with recyclable *P. putida* EM42 resting cells. Rapid transformation of pentose into free xylonate with only minor accumulation of xylonolactone intermediate was observed. Such extracytoplasmic production and secretion are advantageous over intracellular xylose oxidation: cytoplasm acidification is avoided, the reaction of interest does not cross-interfere with the host's metabolism, and xylonate can be purified directly from the culture medium (Wang *et al*., 2016).

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

We then demonstrated that xylose conversion to xylonate can be efficiently catalysed also by recombinant *P. putida* EM42 *bglC*$^{+}$ growing on D-cellobiose. In contrast to monomeric glucose, which is a preferred substrate for glucose dehydrogenase in *P. putida*, the disaccharide did not compete with xylose for Gcd and was a better carbon source for growth-associated xylonate production. Importantly, cellobiose-grown *P. putida* was able to stream the carbon from disaccharide into the intracellular mcl-PHA and concomitantly oxidize xylose to xylonate. Resting cell system can be a preferable option for shake flask set-up if xylonate is the only product of interest because it offers higher productivity and yields than growing cells (Table 1). However, the benefit of co-valorization of the two substrates can be exploited only with growing cells. Both xylonate and PHA yields could be further increased not only through bioprocess design but also by additional genetic interventions in the host that are known to improve the two bioproductions separately. This includes, e.g., overexpression of *gcd* and PQQ biosynthesis genes which would deprive them of their natural regulation (An and Moe, 2016; Yu *et al*., 2018) and/or overexpression of pyruvate dehydrogenase subunit gene *acoA* (Borrero-de Acuña *et al*., 2014). These optimization efforts will be the subject of our further work. Bioprocesses based on microbial hosts capable of parallel production of two or more VAC from cheap abundant substrates are drawing considerable attention (Dumon *et al*., 2012; Li *et al*., 2017; Larroude *et al*., 2018; Baral *et al*., 2019; Wang *et al*., 2019). We argue that the study with recombinant *P. putida* EM42 expressing cytoplasmic β-glucosidase reported here represents a promising route for valorization of (hemi)cellulosic residues and an attractive alternative to the xylonate and mcl-PHA bioproductions reported thus far.

Acknowledgements

We thank Prof. Ivana Márová and Assoc. Prof. Stanislav Obruča from Brno University of Technology for the valuable discussions and provided know-how and infrastructure for PHA quantification and characterization. This work was funded by the SETH Project of the Spanish Ministry of Science (RTI 2018-095584-B-C42), MADONNA (H2020-FET-OPEN-RIA-2017-1-766975), BioRoboost (H2020-NMBP-BIO-CSA-2018) and SYNBIO4FLAV (H2020-NMBP/0500), contracts of the European Union and the S2017/BMD-3691 InGEMICS-CM funded by the Comunidad de Madrid (European Structural and Investment Funds) as well as by the Czech Science Foundation (19-06511Y).

Conflict of interest

The authors declare that they have no conflict of interest.

References

An, R., and Moe, L.A. (2016) Regulation of pyrroloquinoline quinone-dependent glucose dehydrogenase activity in the model rhizosphere-dwelling bacterium *Pseudomonas putida* KT2440. *Appl Environ Microbiol* **82:** 4955–4964.

Baral, N.R., Sundstrom, E.R., Das, L., Gladden, J., Eudes, A., Mortimer, J.C., *et al*. (2019) Approaches for more efficient biological conversion of Lignocellulosic feedstocks to biofuels and bioproducts. *ACS Sustainable Chem Eng* **7:** 9062–9079.

Bator, I., Wittgens, A., Rosenau, F., Tiso, T., and Blank, L.M. (2020) Comparison of three xylose pathways in *Pseudomonas putida* KT2440 for the synthesis of valuable products. *Front Bioeng Biotechnol* **7:** 480.

Belda, E., van Heck, R.G.A., José Lopez-Sanchez, M., Cruveiller, S., Barbe, V., Fraser, C., *et al*. (2016) The revisited genome of *Pseudomonas putida* KT2440 enlightens its value as a robust metabolic chassis. *Environ Microbiol* **18:** 3403–3424.

Borrero-de Acuña, J.M., Bielecka, A., Häussler, S., Schobert, M., Jahn, M., Wittmann, C., *et al*. (2014) Production of medium chain length polyhydroxyalkanoate in metabolic flux optimized *Pseudomonas putida*. *Microb Cell Fact* **13:** 88.

Buchert, J., and Viikari, L. (1988) The role of xylonolactone in xylonic acid production by *Pseudomonas fragi*. *Appl Microbiol Biotechnol* **27:** 333–336.

Buchert, J., Viikari, L., Linko, M., and Markkanen, P. (1986) Production of xylonic acid by *Pseudomonas fragi*. *Biotechnol Lett* **8:** 541–546.

Buchert, J., Puls, J., and Poutanen, K. (1988) Comparison of *Pseudomonas fragi* and *Gluconobacter oxydans* for production of xylonic acid from hemicellulose hydrolyzates. *Appl Microbiol Biotechnol* **28:** 367–372.

del Castillo, T., Ramos, J.L., Rodríguez-Herva, J.J., Fuhrer, T., Sauer, U., and Duque, E. (2007) Convergent peripheral pathways catalyze initial glucose catabolism in *Pseudomonas putida*: genomic and flux analysis. *J Bacteriol* **189:** 5142–5152.

Chen, G.-Q. (2009) A microbial polyhydroxyalkanoates (PHA) based bio- and materials industry. *Chem Soc Rev* **38:** 2434–2446.

Chen, R. (2015) A paradigm shift in biomass technology from complete to partial cellulose hydrolysis: lessons learned from nature. *Bioengineered* **6:** 69–72.

Dumon, C., Song, L., Bozonnet, S., Fauré, R., and O'Donohue, M.J. (2012) Progress and future prospects for pentose-specific biocatalysts in biorefining. *Process Biochem* **47:** 346–357.

Dvořák, P., and de Lorenzo, V. (2018) Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. *Metab Eng* **48:** 94–108.

Gao, C., Hou, J., Xu, P., Guo, L., Chen, X., Hu, G., *et al*. (2019) Programmable biomolecular switches for rewiring flux in *Escherichia coli*. *Nat Commun* **10:** 3751.

Guo, W., Song, C., Kong, M., Geng, W., Wang, Y., and Wang, S. (2011) Simultaneous production and characterization of medium-chain-length polyhydroxyalkanoates and

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

alginate oligosaccharides by *Pseudomonas mendocina* NK-01. *Appl Microbiol Biotechnol* **92:** 791–801.

Ha, S.-J., Galazka, J.M., Kim, S.R., Choi, J.-H., Yang, X., Seo, J.-H., *et al*. (2011) *Engineered Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc Natl Acad Sci USA* **108:** 504–509.

Hardy, G.P., Teixeira de Mattos, M.J., and Neijssel, O.M. (1993) Energy conservation by pyrroloquinoline quinol-linked xylose oxidation in *Pseudomonas putida* NCTC 10936 during carbon-limited growth in chemostat culture. *FEMS Microbiol Lett* **107:** 107–110.

Hori, K., Ichinohe, R., Unno, H., and Marsudi, S. (2011) Simultaneous syntheses of polyhydroxyalkanoates and rhamnolipids by *Pseudomonas aeruginosa* IFO3924 at various temperatures and from various fatty acids. *Biochem Eng* **53:** 196–202.

Huijberts, G.N., Eggink, G., de Waard, P., Huisman, G.W., and Witholt, B. (1992) *Pseudomonas putida* KT2442 cultivated on glucose accumulates poly(3-hydroxyalkanoates) consisting of saturated and unsaturated monomers. *Appl Environ Microbiol* **58:** 536–544.

Jayakody, L.N., Johnson, C.W., Whitham, J.M., Giannone, R.J., Black, B.A., Cleveland, N.S., *et al*. (2018) Thermochemical wastewater valorization via enhanced microbial toxicity tolerance. *Energy Environ Sci* **11:** 1625–1638.

Kawaguchi, H., Hasunuma, T., Ogino, C., and Kondo, A. (2016) Bioprocessing of bio-based chemicals produced from lignocellulosic feedstocks. *Curr Opin Biotechnol* **42:** 30–39.

Köhler, K.A.K., Blank, L.M., Frick, O., and Schmid, A. (2015) D-Xylose assimilation via the Weimberg pathway by solvent-tolerant *Pseudomonas taiwanensis* VLB120. *Environ Microbiol* **17:** 156–170.

Kohlstedt, M., and Wittmann, C. (2019) GC-MS-based ^{13}C metabolic flux analysis resolves the parallel and cyclic glucose metabolism of *Pseudomonas putida* KT2440 and *Pseudomonas aeruginosa* PAO1. *Metab Eng* **54:** 35–53.

Kukurugya, M.A., Mendonca, C.M., Solhtalab, M., Wilkes, R.A., Thannhauser, T.W., and Aristilde, L. (2019) Multi-omics analysis unravels a segregated metabolic flux network that tunes co-utilization of sugar and aromatic carbons in *Pseudomonas putida*. *J Biol Chem* **294:** 8464–8479.

La Rosa, R., Behrends, V., Williams, H.D., Bundy, J.G., and Rojo, F. (2016) Influence of the Crc regulator on the hierarchical use of carbon sources from a complete medium in *Pseudomonas*. *Environ Microbiol* **18:** 807–818.

Larroude, M., Celinska, E., Back, A., Thomas, S., Nicaud, J.-M., and Ledesma-Amaro, R. (2018) A synthetic biology approach to transform *Yarrowia lipolytica* into a competitive biotechnological producer of β-carotene. *Biotechnol Bioeng* **115:** 464–472.

Li, T., Elhadi, D., and Chen, G.-Q. (2017) Co-production of microbial polyhydroxyalkanoates with other chemicals. *Metab Eng* **43:** 29–36.

Licciardello, G., Ferraro, R., Russo, M., Strozzi, F., Catara, A.F., Bella, P., and Catara, V. (2017) Transcriptome analysis of *Pseudomonas mediterranea* and *P. corrugata* plant pathogens during accumulation of medium-chain-length PHAs by glycerol bioconversion. *N Biotechnol* **37:** 39–47.

Lien, O.G. (1959) Determination of gluconolactone, galactonolactone, and their free acids by hydroxamate method. *Anal Chem* **31:** 1363–1366.

Linger, J.G., Vardon, D.R., Guarnieri, M.T., Karp, E.M., Hunsinger, G.B., Franden, M.A., *et al*. (2014) Lignin valorization through integrated biological funneling and chemical catalysis. *Proc Natl Acad Sci USA* **111:** 12013–12018.

Linger, J.G., Hobdey, S.E., Franden, M.A., Fulk, E.M., and Beckham, G.T. (2016) Conversion of levoglucosan and cellobiosan by *Pseudomonas putida* KT2440. *Metab Eng Commun* **3:** 24–29.

Liu, H., Valdehuesa, K.N.G., Nisola, G.M., Ramos, K.R.M., and Chung, W.-J. (2012) High yield production of D-xylonic acid from D-xylose using engineered *Escherichia coli*. *Bioresour Technol* **115:** 244–248.

Loeschcke, A., and Thies, S. (2015) *Pseudomonas putida*-a versatile host for the production of natural products. *Appl Microbiol Biotechnol* **99:** 6197–6214.

Löwe, H., Sinner, P., Kremling, A., and Pflüger-Grau, K. (2018) Engineering sucrose metabolism in *Pseudomonas putida* highlights the importance of porins. *Microb Biotechnol* **13:** 97–106.

Martínez-García, E., Nikel, P.I., Aparicio, T., and de Lorenzo, V. (2014) Pseudomonas 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microb Cell Factories* **13:** 159.

Mehtiö, T., Toivari, M., Wiebe, M.G., Harlin, A., Penttilä, M., and Koivula, A. (2016) Production and applications of carbohydrate-derived sugar acids as generic biobased chemicals. *Crit Rev Biotechnol* **36:** 904–916.

Meijnen, J.-P., de Winde, J.H., and Ruijssenaars, H.J. (2008) Engineering *Pseudomonas putida* S12 for efficient utilization of D-xylose and L-arabinose. *Appl Environ Microbiol* **74:** 5031–5037.

Mosier, N., Wyman, C., Dale, B., Elander, R., Lee, Y.Y., Holtzapple, M., and Ladisch, M. (2005) Features of promising technologies for pretreatment of lignocellulosic biomass. *Bioresour Technol* **96:** 673–686.

Nikel, P.I., and de Lorenzo, V. (2018) *Pseudomonas putida* as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metab Eng* **50:** 142–155.

Nikel, P.I., Chavarría, M., Fuhrer, T., Sauer, U., and de Lorenzo, V. (2015) *Pseudomonas putida* KT2440 strain metabolizes glucose through a cycle formed by enzymes of the Entner-Doudoroff, Embden-Meyerhof-Parnas, and Pentose Phosphate Pathways. *J Biol Chem* **290:** 25920–32.

Nygård, Y., Toivari, M.H., Penttilä, M., Ruohonen, L., and Wiebe, M.G. (2011) Bioconversion of d-xylose to d-xylonate with *Kluyveromyces lactis*. *Metab Eng* **13:** 383–391.

Parisutham, V., Chandran, S.-P., Mukhopadhyay, A., Lee, S.K., and Keasling, J.D. (2017) Intracellular cellobiose metabolism and its applications in lignocellulose-based biorefineries. *Bioresour Technol* **239:** 496–506.

Poblete-Castro, I., Becker, J., Dohnt, K., dos Santos, V.M., and Wittmann, C. (2012) Industrial biotechnology of *Pseudomonas putida* and related species. *Appl Microbiol Biotechnol* **93:** 2279–2290.

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

Poblete-Castro, I., Binger, D., Rodrigues, A., Becker, J., Martins Dos Santos, V.A.P., and Wittmann, C. (2013) *In-silico*-driven metabolic engineering of *Pseudomonas putida* for enhanced production of poly-hydroxyalkanoates. *Metab Eng* **15:** 113–123.

Prieto, A., Escapa, I.F., Martínez, V., Dinjaski, N., Herencias, C., de la Peña, F., *et al*. (2016) A holistic view of polyhydroxyalkanoate metabolism in *Pseudomonas putida*. *Environ Microbiol* **18:** 341–357.

Sasnow, S.S., Wei, H., and Aristilde, L. (2016) Bypasses in intracellular glucose metabolism in iron-limited *Pseudomonas putida*. *Microbiologyopen* **5:** 3–20.

Searles, S., and Malins, C. (2013) ICCT White Paper. Availability of cellulosic residues and wastes in the EU. [WWW document]. URL https://www.theicct.org/sites/default/files/publications/ICCT_EUcellulosic-waste-residues_20131022.pdf.

Taha, M., Foda, M., Shahsavari, E., Aburto-Medina, A., Adetutu, E., and Ball, A. (2016) Commercial feasibility of lignocellulose biodegradation: possibilities and challenges. *Curr Opin Biotechnol* **38:** 190–197.

Toivari, M., Nygård, Y., Kumpula, E.-P., Vehkomäki, M.-L., Benčina, M., Valkonen, M., *et al*. (2012a) Metabolic engineering of *Saccharomyces cerevisiae* for bioconversion of D-xylose to D-xylonate. *Metab Eng* **14:** 427–436.

Toivari, M.H., Nygård, Y., Penttilä, M., Ruohonen, L., and Wiebe, M.G. (2012b) Microbial D-xylonate production. *Appl Microbiol Biotechnol* **96:** 1–8.

Wang, C., Wei, D., Zhang, Z., Wang, D., Shi, J., Kim, C.H., *et al*. (2016) Production of xylonic acid by *Klebsiella pneumoniae*. *Appl Microbiol Biotechnol* **100:** 10055–10063.

Wang, Y., Ling, C., Chen, Y., Jiang, X., and Chen, G.-Q. (2019) Microbial engineering for easy downstream processing. *Biotechnol Adv* **37:** 107365.

Werpy, T., and Petersen, G. (2004) Top value added chemicals from biomass: Volume I - Results of screening for potential candidates from sugars and synthesis gas. https://doi.org/10.2172/15008859.

Wiebe, M.G., Nygård, Y., Oja, M., Andberg, M., Ruohonen, L., Koivula, A., *et al*. (2015) A novel aldose-aldose oxidoreductase for co-production of D-xylonate and xylitol from D-xylose with *Saccharomyces cerevisiae*. *Appl Microbiol Biotechnol* **99:** 9439–9447.

Yu, S., Lai, B., Plan, M.R., Hodson, M.P., Lestari, E.A., Song, H., and Krömer, J.O. (2018) Improved performance of *Pseudomonas putida* in a bioelectrochemical system through overexpression of periplasmic glucose dehydrogenase. *Biotechnol Bioeng* **115:** 145–155.

Zhou, X., Zhou, X., and Xu, Y. (2017) Improvement of fermentation performance of *Gluconobacter oxydans* by combination of enhanced oxygen mass transfer in compressed-oxygen-supplied sealed system and cell-recycle technique. *Bioresour Technol* **244:** 1137–1141.

Zhu, M., Sun, L., Lu, X., Zong, H., and Zhuge, B. (2019) Establishment of a transient CRISPR-Cas9 genome editing system in *Candida glycerinogenes* for co-production of ethanol and xylonic acid. *J Biosci Bioeng* **128:** 283–289.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Experimental procedures

Results and discussion

Table S1. Strains and plasmids used in this study.

Fig. S1. Recycling of *P. putida* EM42 resting cells converting D-xylose to D-xylonate.

Fig. S2. Effect of agitation speed and medium pH on conversion of xylose to xylonate by *P. putida* EM42 resting cells.

Fig. S3. Two-day culture of *P. putida* EM42 in minimal medium with 10 g l^{-1} D-xylose and 10 g l^{-1} D-glucose used as a sole carbon source for growth.

Fig. S4. Polyhydroxyalkanoate (PHA) accumulation in *P. putida* EM42 cells grown on diverse carbon sources including D-cellobiose.

Fig. S5. Activities of β-glucosidase BglC and glucokinase Glk measured in cell-free extracts prepared from *bglC*$^+$ strain of *P. putida* EM42 grown in several different conditions.

© 2020 The Authors. *Microbial Biotechnology* published by John Wiley & Sons Ltd and Society for Applied Microbiology.

ACS SyntheticBiology

pubs.acs.org/synthbio Research Article

Surface Display of Designer Protein Scaffolds on Genome-Reduced Strains of *Pseudomonas putida*

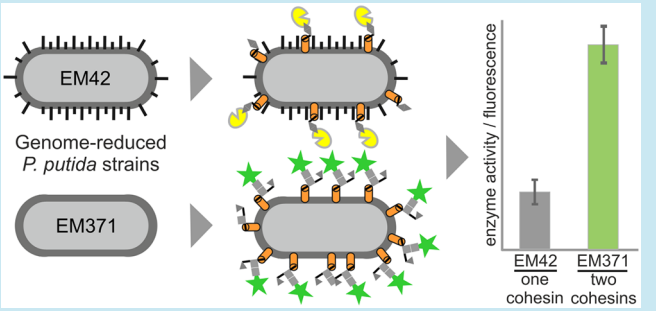
Pavel Dvořák,* Edward A. Bayer, and Víctor de Lorenzo*

Cite This: https://dx.doi.org/10.1021/acssynbio.0c00276 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

Downloaded via MASARYK UNIV on September 22, 2020 at 09:28:01 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

ABSTRACT: The bacterium *Pseudomonas putida* KT2440 is gaining considerable interest as a microbial platform for biotechnological valorization of polymeric organic materials, such as lignocellulosic residues or plastics. However, *P. putida* on its own cannot make much use of such complex substrates, mainly because it lacks an efficient extracellular depolymerizing apparatus. We seek to address this limitation by adopting a recombinant cellulosome strategy for this host. In this work, we report an essential step in this endeavor—a display of designer enzyme-anchoring protein "scaffoldins", encompassing cohesin binding domains from divergent cellulolytic bacterial species on the *P. putida* surface. Two *P. putida* chassis strains, EM42 and EM371, with streamlined genomes and differences in the composition of the outer membrane were employed in this study. Scaffoldin variants were optimally delivered to their surface with one of four tested autotransporter systems (Ag43 from *Escherichia coli*), and the efficient display was confirmed by extracellular attachment of chimeric β-glucosidase and fluorescent proteins. Our results not only highlight the value of cell surface engineering for presentation of recombinant proteins on the envelope of Gram-negative bacteria but also pave the way toward designer cellulosome strategies tailored for *P. putida*.



KEYWORDS: *Pseudomonas putida, cellulosome, designer scaffoldin, surface display, synthetic biology*

Polymeric organic materials such as lignocellulosic residues or plastics represent a potentially inexhaustible source of cheap carbon and energy for biotechnology and synthetic biology enterprises.[1,2] Adoption of these recalcitrant feedstocks for bioproduction of valuable chemicals requires the employment of microbial hosts with a suite of properties that would allow them to perform complex biocatalytic conversions efficiently even under harsh conditions of industrial processes. Such microorganisms are currently not available but can be obtained by engineering suitable robust platform strains.

Pseudomonas putida KT2440, a Gram-negative bacterial workhorse, fulfills some of the crucial criteria to become a host of choice for biotechnological upcycling of polymeric wastes. It has been recently engineered for utilization and valorization of several plant biomass-derived sugars,[3−5] lignin-born aromatic chemicals,[6,7] and even products of plastic degradation.[8] It was also demonstrated that this bacterium can process oligomeric carbohydrates[4,9,10] as well as coutilize hexose and pentose sugars—glucose and xylose—and consume simultaneously glucose and an aromatic substrate with a lack of diauxia.[4,11] These and other characteristics, including safety status,[12] rapid growth and low nutritional demand,[13] considerable resistance to inhibitory chemicals,[14,15] its employment in large-scale fermentations for production of value-added chemicals,[16] or its compliance to genetic manipulations, and the available palette of engineering tools,[17−19] make *P. putida* an attractive candidate for the demanding biotechnological task sketched above. However, *P. putida*, same as the majority of other domesticated microbial platforms, lacks an efficient extracellular depolymerizing apparatus and cannot degrade complex recalcitrant substrates alone.

The most efficient natural polymer degraders known to date, cellulolytic bacteria, such as *Clostridium thermocellum*, display on their surface cellulosomes—remarkable nanomachines composed of scaffoldin proteins that attach and orchestrate multiple carbohydrate-active enzymes on the cell surface.[20] The binding of cellulases to scaffoldins is mediated by strong ($K_D \sim 10^{-9} - 10^{-10}$ M) highly specific non-covalent interactions between cohesin and dockerin binding domains.[21] The firmness of cohesin−dockerin pairs is provided by extensive H-bond networks.[22] Concerted action of dockerin-tagged cellulases clustered on the cell surface through interactions with cohesins in scaffoldin proteins can enhance

Received: May 21, 2020
Published: September 2, 2020

© XXXX American Chemical Society

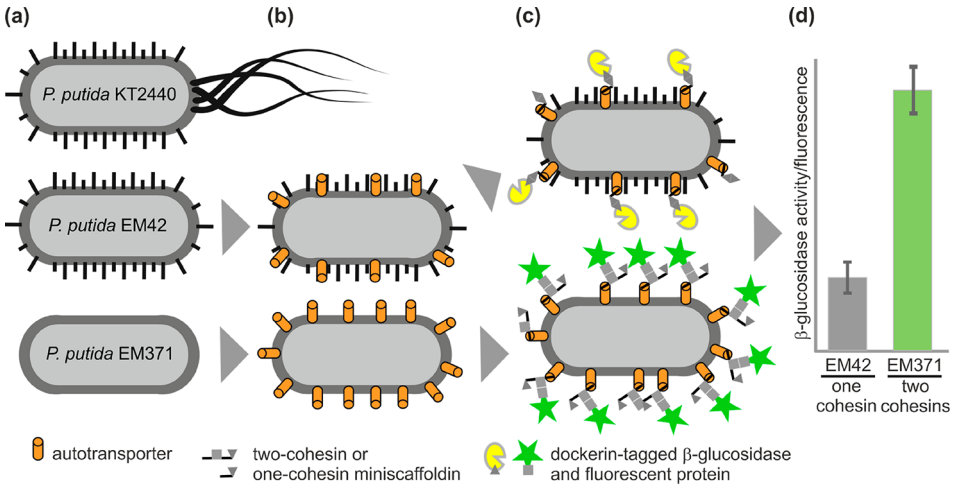


Figure 1. Engineering *Pseudomonas putida* for surface display of cohesin-containing designer protein scaffoldins: schematic overview. (a) Two derivatives of *P. putida* KT2440, strains EM42 and EM371, with streamlined genomes lacking destabilizing genetic elements and certain surface structures (represented here by black bars), respectively, employed in this study. (b) Four type V secretion systems were tested in the target host, and the best-performing autotransporter was selected for further work. (c and d) Single- or two-cohesin scaffoldins were displayed on the EM42 and EM371 surfaces, and the efficiency of the binding of dockerin-tagged recombinant proteins (*β*-glucosidase or fluorescent proteins) to the cellular surface was evaluated and quantified.

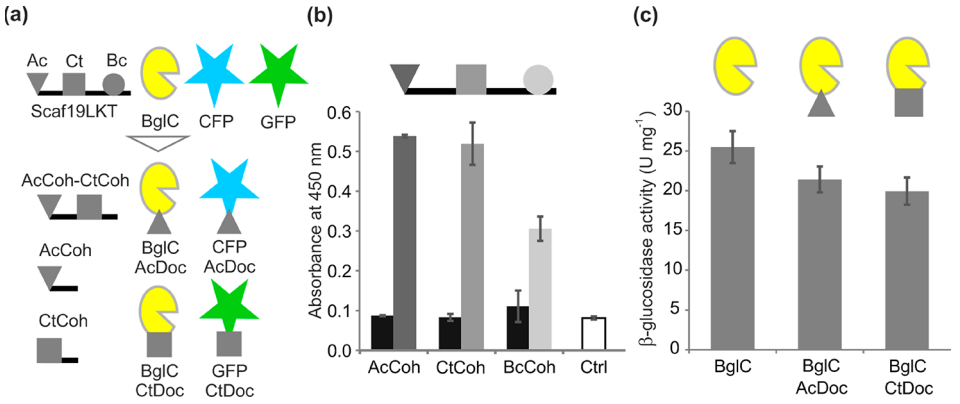


Figure 2. Recombinant proteins used in this study and initial evaluation of their function. (a) Three truncated variants of previously constructed synthetic scaffoldin gene *scaf19LKT*[47] (for detailed description see the Materials and Methods section) were prepared and adopted for this study: two variants with either single CtCoh (square symbol) or AcCoh (triangle symbol) cohesin from *Clostridium thermocellum* or *Acetivibrio cellulolyticus*, respectively, and a two-cohesin scaffoldin with both AcCoh and CtCoh interconnected with a 29 amino-acid-long linker. Furthermore, two chimeric variants of *β*-glucosidase BglC from *Thermobifida fusca* with CtDoc or AcDoc dockerin matching complementary cohesins were constructed, as well as a recombinant green fluorescent protein (GFP) with CtDoc and mCerulean fluorescent protein (CFP) with AcDoc. (b) ELISA-based verification of function of AcCoh, CtCoh, and BcCoh (cohesin from *Bacteroides cellulosolvens*) in the Scaf19LKT scaffoldin produced in *P. putida* EM42. Controls without scaffoldin but with added dockerin-tagged proteins and antibodies (for details see the Materials and Methods section) are shown as black bars. Ctrl, control with scaffoldin and antibodies but no dockerin-tagged protein. Data are shown as mean ± SD from two biological replicates, which were each taken as the mean value from two technical replicates. (c) Comparison of activities of free purified dockerin-tagged BglC variants produced in *P. putida* EM42 with activity of wild-type BglC. Data are shown as mean ± SD from three biological replicates.

cellulose degradation up to 50-fold when compared with free enzymes.[23]

Natural cellulosome producers (e.g., *C. thermocellum*, *Bacteroides cellulosolvens*, *Acetivibrio cellulolyticus*) are difficult to genetically manipulate or cultivate, and their cellulosomes are large and often very complex.[24] Hence, smaller synthetic *designer cellulosomes* or *minicellulosomes* have been assembled either *in vitro* using purified components[25−27] or *in vivo* on the surface of a suitable microbial host.[28−32] A fundamental prerequisite for successful minicellulosome assembly is a display of scaffoldin proteins on the surface of a target host. Truncated forms of native scaffoldins or designer hybrid scaffoldins with cohesin domains from diverse cellulolytic organisms were delivered to the cell surface of a recombinant *Saccharomyces cerevisiae*,[28,33] *Bacillus subtilis*,[31] *Clostridium acetobutylicum*,[34] *Lactobacillus plantarum*,[32] or *Lactococcus lactis*.[35] However, efficient expression and secretion of cellulosome components in a phylogenetically distant Gram-negative host with different codon usage, G+C content, and a two-layer structure of cell wall remain very challenging. Not surprisingly, surface display of scaffoldins and subsequent *in vivo* assembly of a designer cellulosome of any size has not yet been reported in a biotechnologically relevant Gram-negative bacterium. Thus far, display or secretion of individual depolymerizing enzymes has been achieved in engineered *E. coli*, allowing it to grow on cellooligosaccharides or produce

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

biofuels from plant biomass.[36−39] In the case of *P. putida*, single cellulases from *Ruminiclostridium thermocellum* and hemicellulases from *Bacillus subtilis* were displayed on the surface of recombinant strains mixed to form designer cocultures of resting cells.[40,41] The joint activities of these enzymes in cell suspensions of high cell densities resulted in the production of small quantities of glucose[40] or xylose[41] from filter paper or arabinoxylan, respectively, which were insufficient to support the growth of the host bacterium. Designer cellulosomes can result in more efficient saccharification of (hemi)cellulose compared to the strategy that utilizes cellulases displayed separately on several strains.[42] To test that hypothesis in a Gram-negative bacterial host, an efficient display of scaffoldins for docking of cellulases on its surface must first be achieved.

In this study, we aimed to display structurally distinct variants of a designer miniscaffoldin on the surface of genome-reduced *P. putida* strains designated EM42 and EM371 (Figure 1 and Figure S1). The two strains were adopted mainly due to the differences in the presence of outer membrane structures and complexity of the bacterial surface, which can impose structural constraints for displayed proteins.[43] Strain EM42 possesses 11 non-adjacent genomic deletions (300 genes, ∼4.3% of the whole genome) that were shown to improve expression of heterologous genes and enhanced biotechnological potential of this *P. putida* KT2440 derivative.[4,44,45] Except for missing flagellum, the cell surface of *P. putida* EM42 resembles that of the wild-type strain KT2440. In the strain EM371, on contrary, most of the non-essential outer membrane structures that are used by bacteria to coordinate motion (flagella) or to develop biofilms and interact with their surroundings (*e.g.*, fimbriae, pili, curli, adhesins, exopolysaccharides, lipopolysaccharides) were eliminated (230 genes, ∼4.7% of the entire genome).[46] The "shaved" surface endowed *P. putida* EM371 with properties potentially beneficial for extracellular recombinant protein production and facilitated downstream processing. Here, expression and display of scaffoldin variants on the surface of EM42 and EM371, *via* one of the four tested autotransporter systems, allowed for: (i) addressing important questions on how the character of secreted molecules and cell surface contribute to cohesin−dockerin interactions and (ii) comparison of the capability of the two *P. putida* strains to be used for *in vivo* assembly of surface-exposed designer catalytic nanomachines.

RESULTS AND DISCUSSION

Preparation of Cohesin-Containing Scaffoldin Variants and Dockerin-Tagged Chimeric Proteins. Scaffoldin Scaf19L with three cohesins interconnected with 29 and 35 amino acid (aa) long flexible linkers[47] was initially adopted for this study (Figure 2a, Materials and Methods section, and the Supporting Information). The scaffoldin gene was subcloned into the pSEVA238 plasmid with inducible XylS/*Pm* expression system,[48] and the construct was inserted into *P. putida* EM42. However, induction of the gene expression resulted in reduced host's fitness (Supporting Information Methods and Figure S2a). Western blot analysis of the cell lysate revealed limited Scaf19L solubility and its susceptibility to proteolytic cleavage (Supporting Information Methods and Figure S2b). The codon adaptation index (CAI), calculated for *scaf19L* gene and *P. putida* KT2440 host using the online tool JCat,[49] was very low (0.10; where CAI of 1.0 signifies the best match), which indicated that gene toxicity could have reflected codon bias. Hence, the *scaf19L* gene was synthesized and codon optimized for expression in *P. putida*. Heterologous expression of the resultant optimized gene *scaf19LKT* with increased CAI (0.51) had no negative effect on *P. putida* EM42 growth (Figure S2a). Most of the Scaf19LKT protein was produced in a soluble form (>75% compared to <50% before optimization), and proteolysis was not observed (Figure S2b).

The functionality of cohesins in Scaf19LKT produced in *P. putida* EM42 was then analyzed using an enzyme-linked immunosorbent assay (ELISA)-based binding assay.[47] Figure 2b shows that all three cohesins in Scaf19LKT produced in *P. putida* were able to bind their respective dockerins. AcCoh−AcDoc and CtCoh−CtDoc pairs provided almost twice as much signal as the BcCoh−BcDoc combination, which signified a possible tighter binding of the former two pairs. We therefore selected these pairs for experiments in the current study. Three truncated variants of *scaf19LKT* gene encoding scaffoldins AcCoh, CtCoh, and AcCoh−CtCoh with a single cohesin or with two cohesins were prepared for surface display on *P. putida* EM42 and EM371 strains (Figure 2a).

We then aimed at the assembly of dockerin-tagged reporter proteins that could be employed in a rapid robust assay for the detection of displayed scaffoldins on the surface of the target host cells. Displayed cohesins can be detected and quantified by ELISA-based protocols[32,47] or by assays with dockerin-tagged enzymes, such as β-glucuronidase.[35] The latter approach was adopted in this study. β-Glucosidase (EC 3.2.1.21) BglC from *Thermobifida fusca* is functionally expressed in *P. putida* EM42 up to 30% of the total soluble protein.[4] Measurement of its hydrolytic activity with synthetic *p*-nitrophenyl-β-D-glucopyranoside (pNPG) substrate is simple and fast. We modified this enzyme by fusing it on its C terminus with AcDoc or CtDoc dockerin (Figure 2a). Two chimeras and wild-type BglC with a polyhistidine tag were produced in *P. putida* EM42 and purified (Figure S3a−c). As shown in Figure 2c, specific activities of BglC−AcDoc and BglC−CtDoc reached 84% and 78% of the wild-type activity, respectively (activities were normalized to the estimated purities of BglC, BglC−AcDoc, and BglC−CtDoc, which were 80%, 70%, and 85%, respectively; see Supporting Information Methods and Figure S3a,b,c). Dockerin-tagged variants of cyan fluorescent protein mCerulean and monomeric superfolder green fluorescent protein, abbreviated here as CFP−AcDoc and GFP−CtDoc, respectively, were prepared in the same way as BglC chimeras for spectroscopic and microscopic confirmation of displayed scaffoldins (Figure 2a and Figure S3d,e).

Selection of an Optimal Autotransporter System for Display of Scaffoldins on the Surface of *P. putida* EM42 and EM371 Strains. Monomeric type V secretion pathway proteins known as autotransporters have been used for decorating surfaces of Gram-negative bacteria with recombinant proteins,[50,51] mainly due to their simplicity (a single gene encodes all three domains needed for display: a signal peptide, a surface-exposed passenger, and a transmembrane β-domain), high display efficiencies reaching in certain cases 10^4−10^5 enzyme molecules per cell, and relatively low toxicity of recombinant variants toward a bacterial host.[52,53] Three autotransporters, namely, EhaA from enterohemorrhagic *E. coli*,[54] EstP from *P. putida*,[55] and immunoglobulin A (IgA) protease from *Neisseria gonorrhoeae*,[56] were used with some success for a passenger export in *P. putida* KT2440, but in

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

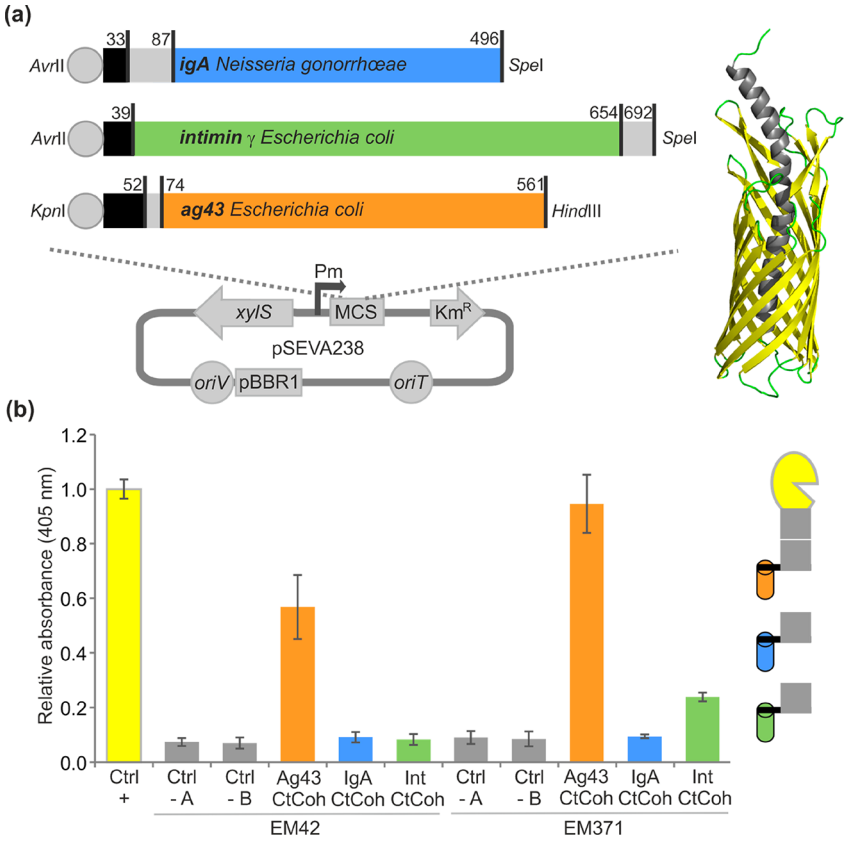


Figure 3. Selection of the type V secretion system for display of scaffoldins on the surface of *Pseudomonas putida* EM42 and EM371 strains. (a) Cloning of genes encoding translocation β-barrel domain and part of the α-helix passenger (model of EstP autotransporter prepared by iTasser[64] is shown as a representative example) of three compared autotransporter systems into the marked restriction sites of the pSEVA238 plasmid polylinker. The genes were preceded by a synthetic ribosome binding site (gray sphere), and contained a signal peptide sequence (in black) and their own multiple cloning site (in gray). The length of the individual segments in the amino acid sequence is shown. (b) Display test of CtCoh cohesin on the surface of EM42 and EM371 strains with the three autotransporters. The cells displaying CtCoh were mixed with purified BglC−CtDoc, washed and incubated with *p*-nitrophenyl-β-D-glucopyranoside substrate. End-point absorbance of the reaction product *p*-nitrophenol was determined and compared with the absorbance measured for purified BglC−CtDoc of 4 μg mL^{-1} concentration in the reaction mixture (Ctrl+). EM42 and EM371 cells with empty pSEVA238 plasmid mixed with BglC−CtDoc (Ctrl− A) and EM42 and EM371 pSEVA238_*ag43AT*−*ctCoh* cells mixed with wild-type BglC (Ctrl− B) were used as negative controls. Data are shown as mean ± SD from at least three independent experiments, each conducted in two technical replicates.

general, the reports on recombinant protein surface display in this host are scarce.

Because of the difficulty in predicting the functionality of a secretion system with a given passenger protein in a selected bacterial host,[57] we sought to evaluate four different autotransportes for display of designer scaffoldins in *P. putida*. Three previously constructed systems, including translocator domains (β-barrels and α-helix linkers) of the IgA protease from *N. gonorroheae*,[58] antigen 43 (Ag43) from *E. coli*,[38,59] and EaeA intimin (Int, inverse autotransporter) from *E. coli*,[60,61] were adopted for this study (Figure 3a, Supporting Information Table S1, and Materials and Methods section). Moreover, *P. putida* EstP esterase (PP_0418) translocator sequence, encoding the β-barrel domain with a spanning α-helical linker complemented by a synthetic multicloning site and a native signal peptide sequence, was synthesized and tested. IgA and intimin were already available in our laboratory; the genes were cloned in pSEVA238 plasmid. The *ag43* gene from the pAg43pol vector[38] and synthesized *estP* gene were each subcloned into pSEVA238 as well. The EstP autotransporter was nonetheless soon excluded from the list because its expression in *P. putida* EM42 appeared to be extremely toxic for the host (Figure S4 and Supporting Information Methods). Such a strong toxic effect might be attributed to the burden caused by overproduction of the native protein with efficient secretion signal and subsequent distortion of the cell membrane due to the high number of integrating β-barrels.[62] Overproduced autotransporter molecules can also exhaust secretion machinery, namely, Sec and BAM systems, required for the export of proteins necessary for cell growth and maintenance.[50,63] Induction of expression of the remaining three autotransporters had no or negligible effect on the host's viability (Figure S4). All these systems were therefore selected for further testing.

In the next step, the codon-optimized gene encoding CtCoh (Mw = 16.2 kDa) was subcloned into the polylinkers of *igAAT*, *ag43AT*, and *intAT*, and the three autotransporters were tested for display of the single-cohesin scaffoldin on the surface of EM42 and EM371 strains. Cells displaying scaffoldin were added with an excess of purified BglC−CtDoc (~1.0 × 10^5 molecules per cell, the value was determined as described in the Materials and Methods section) and washed, and their β-glucosidase activity was determined by measuring the end-point absorbance of the reaction product *p*-nitrophenol released after hydrolysis of the pNPG substrate (Figure 3b). The absorbance of the supernatant fluids from the reactions

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

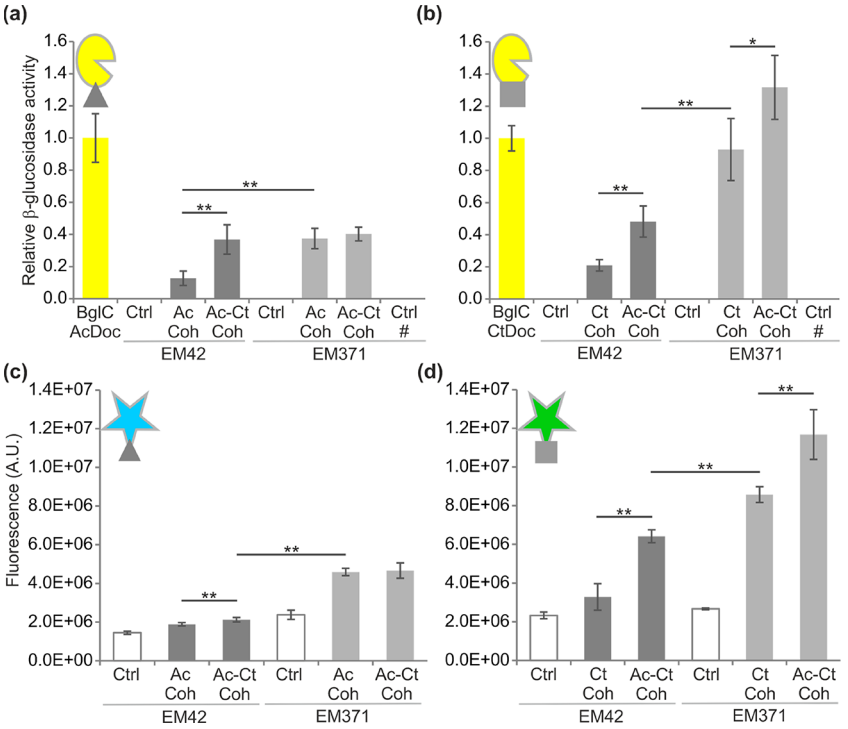


Figure 4. Binding of dockerin-tagged β-glucosidase and fluorescent proteins on *Pseudomonas putida* EM42 and EM371 cells displaying designer scaffoldins. (a and b) Binding of BglC−AcDoc and BglC−CtDoc, respectively, to EM42 and EM371 cells displaying AcCoh, CtCoh, or AcCoh−CtCoh scaffoldins. β-Glucosidase activities were measured with whole cells and related to the activity of purified BglC−AcDoc (4.64 ± 0.71 U) or BglC−CtDoc (4.91 ± 0.39 U) of 2 μg mL^{-1} concentration in the reaction mixture. EM42 and EM371 cells with the pSEVA238b_*ag43AT* plasmid were used as controls (Ctrl). EM371 cells with the pSEVA238b_*ag43AT−ctCoh* or *−acCoh* were mixed with BglC−AcDoc or BglC−CtDoc, respectively, to check cross-reactivity of used cohesin-dockerin pairs (Ctrl#). Data are shown as mean ± SD from at least two independent experiments, each conducted in two−three biological replicates. (c and d) Binding of CFP−AcDoc and GFP−CtDoc, respectively, to EM42 and EM371 cells displaying AcCoh, CtCoh, or AcCoh−CtCoh scaffoldins. EM42 and EM371 cells with pSEVA238b_*ag43AT* plasmid were used as controls (Ctrl). Data are shown as mean ± SD from three biological replicates, each conducted in two technical replicates. Asterisks denote significance in difference in between two means at $P < 0.05$ (*) or $P < 0.01$ (**).

with whole cells was related to the absorbance of the mixture with purified BglC−CtDoc, which served as a positive control. An obvious advantage of such an approach is that all measured activity can be attributed only to the chimeric enzyme molecules attached from outside to the surface of the intact cells that passed cycles of centrifugation and washing. This feature is especially valuable when considering the fact that whole-cell ELISA and proteinase accessibility assays, frequently used to evaluate the efficiency of recombinant protein display, are prone to false-positive results.[53,65]

Figure 3b reveals that only the cells expressing Ag43 autotransporter showed high β-glucosidase activity, which, in the case of the EM371 recombinant, reached the level of the pure enzyme control. Some activity was also detected with EM371 cells displaying scaffoldin *via* intimin, but their EM42 counterparts were not able to attach dockerin-tagged BglC to their surface. Absorbance measured in supernatant fractions from reactions with the EM42 pSEVA238_*intAT−CtCoh* recombinant did not surpass those of the negative controls (Figure 3b). The same was also true for both recombinants (either EM42 or EM371) expressing *igAAT*, which suggests that this autotransporter was the least efficient in CtCoh display out of the three tested candidates. SDS polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot analysis of the cell lysates (Supporting Information Methods and Figure S5a,b) confirmed the expression of the *ag43AT−ctCoh* gene in both EM42 and EM371 recombinants. The chimera comprised ∼3% of the total cellular protein. In contrast, expression of *intAT−ctCoh* was not detected on SDS-PAGE gel in either of the two lysates, and only a faint band was identified among the blotted proteins of the EM371 recombinant (Figure S5a,b). Hence, the malfunctioning of this autotransporter system can be attributed to the poor expression of the construct. The same conclusion cannot be made for IgAAT−CtCoh because its production was detectable in EM42 and EM371 lysates, both on the SDS-PAGE gel and on the blotting membrane. However, no signal was seen with whole preinduced EM42 and EM371 cells bearing the *igAAT−ctCoh* gene after their incubation with HRP-conjugated anti-6xHis tag antibody during the dot blot analysis (Figure S5c). Hence, one possible explanation of the IgAAT−CtCoh malfunctioning is that the chimera was expressed but it was not functionally displayed on the *P. putida* surface, perhaps due to improper folding during its transport and maturation. It is worth noting that *P. putida*, empowered with the IgA translocator construct identical to the one used here, was able to secrete detectable quantities of metallothioneins or eukaryotic leucine zippers in the former works of Valls et al.[66] and Martínez-García and co-workers,[46] respectively. Such discrepancy supports the repeated observation that the ability of a certain autotransporter to export a passenger of choice in a given host bacterium cannot be predicted with confidence prior to experimental verification.[53,57]

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Dot blot analysis of whole preinduced cells also confirmed the display of CtCoh *via* the Ag43 autotransporter in the EM371 strain (Supporting Information Methods and Figure S5c). None of the EM42 strains, including EM42 with the Ag43AT−CtCoh construct, showed a luminescence signal. We hypothesize that this could be ascribed to the location of the 6xHis tag between the CtCoh and Ag43AT molecules and its poor accessibility for antibody on the surface of the EM42 cells.[67] Taken together, the results discussed above identified Ag43 as a promising secretion system for surface display of cohesin binding domains in *P. putida*. The Ag43-based secretion systems have been repeatedly proven useful for display or secretion of recombinant proteins in *E. coli*,[37,38,68] but, to the best of our knowledge, the use of this autotransporter has not yet been reported for *P. putida*.

Native Ag43 (Uniprot ID: P39180) possesses an N-terminal signal peptide (aa 0−52), a surface-exposed N-proximal β-helical α-passenger domain (aa 53−551), and a C-terminal β-barrel domain (aa 552−1039) whose part is buried in the outer membrane.[69] Under normal circumstances, the α domain is cleaved during the secretion process by an internal protease motif and remains attached to the cell surface by non-covalent interactions.[68] The Ag43 autotransporter employed in this study lacks the whole α domain[38], and the passenger protein (miniscaffoldin in our case) thus remains covalently attached to the cell surface.

Binding of Dockerin-Tagged β-Glucosidase and Fluorescent Proteins on *P. putida* EM42 and EM371 Cells Displaying Scaffoldin Variants. Cellulase BglC−CtDoc was successfully attached to *P. putida* cells expressing Ag43 autotransporter with a single-cohesin scaffoldin. But how is such a directed interaction influenced by the character and size of the displayed scaffoldin? And how does the character of a host cell's surface contribute to such bonding? We aimed to answer these questions in the following part of our study. Surface displays of the structurally distinct AcCoh (Mw = 16.5 kDa), CtCoh, and AcCoh−CtCoh (Mw = 34.6 kDa) scaffoldins on the *P. putida* EM42 and EM371 recombinants were evaluated by attachment of dockerin-tagged β-glucosidase variants BglC−AcDoc and BglC−CtDoc and chimeric fluorophores CFP−AcDoc and GFP−CtDoc (Figure 4).

Whole-cell activities of β-glucosidase were quantitatively determined in defined time intervals and related to the activity of the corresponding purified recombinant enzyme (Figure 4a,b and Materials and Methods section). Fluorescence of *P. putida* cells with surface-attached CFP or GFP was measured in microtiter plate format (Figure 4c,d). The results confirmed that both cohesin domains were accessible on the cell surface and functional in binding their respective dockerins. The cross-reactivity controls verified exclusivity of AcCoh−AcDoc and CtCoh−CtDoc bonds (Figure 4a,b).

Similar trends in the binding efficiency of dockerin-tagged proteins to the cells with displayed cohesins were observed in both assay types. EM371 cells bound more dockerin-tagged proteins and showed significantly higher β-glucosidase activity and fluorescence ($P < 0.01$) than EM42 recombinants in almost all tested scenarios. This could not be ascribed to the better expression of autotransporter−scaffoldin chimeras in the EM371 strain. On the contrary, the expression levels of all three constructs were approximately 25% lower in EM371 than in EM42, as judged by densitometric analysis of the SDS-PAGE gel with samples of cell lysates (Figure S6 and Supporting Information Methods). Hence, we assume that better display or accessibility of the exported scaffoldins on the surface of EM371 strain is the correct explanation. Deletion mutants of *E. coli* BL21(DE3) lacking several abundant but non-essential outer membrane proteins were previously shown to be excellent hosts for overexpression of heterologous membrane β-barrel proteins, presumably because targeted knockouts relieved some of the burden on the Sec and BAM secretion machineries.[50] Moreover, removed surface structures leave more space for heterologous transporters and their passengers, which can make them better accessible for their binding partners.[70]

Out of the two tested cohesin-dockerin pairs, AcCoh and AcDoc showed lower binding affinity in both EM42 and EM371 strains (Figure 4). CtCoh−CtDoc showed better bonding in all assays, even though the position of CtCoh in the two-cohesin scaffoldin was theoretically less favorable for interaction with the respective dockerin (closer to the cell wall) than that of AcCoh. This observation does not match the former outcome of the ELISA assay, which indicated equally strong bonds in the two pairs (Figure 2b). However, the current experiments were performed with whole cells not cell-free extracts as in the case of ELISA, and it is possible that AcCoh architecture was affected during the secretion process in the non-native host. In the case of fluorescence measurements (Figure 4c,d), the dimmer signals from the cells decorated with AcCoh and CFP−AcDoc complexes could partially stem also from the lower brightness of the cyan fluorophore when compared with GFP.[71] However, the most probable cause of the observed phenomenon is that the cohesin−dockerin pair from the mesophilic *Acetivibrio cellulolyticus* is less stable and more prone to disruption during the experimental treatment (including several cycles of cell centrifugation and washing) than the binding domains from thermophilic *Clostridium thermocellum*. As shown recently by Gunnoo and co-workers,[22] who performed molecular dynamics simulations with the very same cohesin−dockerin pairs, the *C. thermocellum* system presents a stronger hydrogen bond network and higher binding affinities than the *A. cellulolyticus* complex even at ambient temperature. These theoretical calculations together with our experimental data indicate that the binding domains from thermophilic cellulolytic bacteria might be a better choice for assembly of stable scaffoldin−enzyme interactions also on the surface of mesophilic microbial hosts.

Another conclusion that deserves attention is that *P. putida* cells displaying larger AcCoh−CtCoh scaffoldins were able to bind more enzyme or fluorophore molecules than recombinants decorated with single cohesins (Figure 4). This trend was especially pronounced with *P. putida* EM42 pSEVA238_*AcCoh−CtCoh*. The strain showed 2.9-fold or 2.3-fold higher β-glucosidase activity (Figure 4a,b) and 1.5-fold or 4.3-fold enhanced fluorescence (Figure 4c,d) when compared with its AcCoh- or CtCoh-exporting counterparts, respectively. This somewhat counterintuitive observation is in agreement with the study of Wieczorek and Martin[35] who described the same phenomenon for the secretion of designer scaffoldins in the Gram-positive bacterium *Lactococcus lactis*. Larger molecules may make more space for themselves on the "bushy" surface of EM42 strain and are thus better accessible for binding partners than buried single cohesins. This hypothesis is supported by the fact that the difference between the display of single- and two-cohesin scaffoldin was smaller in case of EM371 recombinants (Figure 4). In these, both smaller and

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

larger scaffoldins were equally (Figure 4a,c) or similarly (Figure 4b,d, the difference is 1.4- and 1.5-fold) accessible for AcDoc- and CtDoc-tagged proteins, respectively. The higher accessibility of the two-cohesin scaffoldin in the EM42 recombinant was certainly not the outcome of its better expression. On the contrary, AcCoh−CtCoh was the least expressed construct from all three tested scaffoldin variants both in EM42 and EM371 (Figure S6). The fact that the larger two-cohesin scaffoldin was more accessible on the surfaces of the EM42 and EM371 recombinants than the individually displayed cohesins is promising for our further work. Only parallel assembly of two or more complementary enzymatic activities on the same scaffoldin can promote the substrate channeling effect and result in more efficient whole-cell biocatalysts compared to bacteria displaying or secreting single enzyme molecules.[42,72]

The aforementioned phenomena were reconfirmed by observing the cells with surface-docked fluorescent proteins in the confocal microscope (Figure 5). We could see that (i) *P. putida* EM371 recombinants showed brighter fluorescent signal than the EM42 strains, (ii) the fluorescence of cells with attached CFP−AcDoc was dimmer than the fluorescence of their GFP−CtDoc binding counterparts, and (iii) the fluorescence of the EM42 strain expressing the larger AcCoh−CtCoh scaffoldin was more visible than the fluorescence of EM42 cells with displayed single cohesins. The microscopy technique allowed us to also examine positioning of the secreted scaffoldins on the surface of the tested strains. The fluorescence signal was more concentrated in the poles of all EM42 and EM371 recombinants (Figure 5), which indicates asymmetric distribution of the Ag43 autotransporter in the cellular membrane. Such accumulation of overexpressed Ag43 at the polar edges was recently reported also for *E. coli*[68] and can be a consequence of a cell wall material pushing toward the poles through the continuous lateral insertion of new peptidoglycan building blocks over the rounds of growth and division.[73]

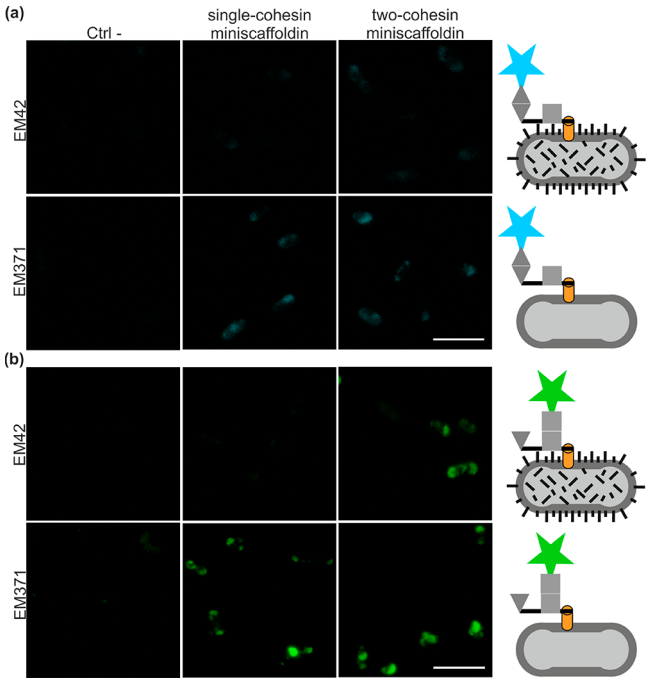


Figure 5. Confocal microscopy of dockerin-tagged fluorescent proteins bound to the *Pseudomonas putida* EM42 and EM371 cells displaying designer scaffoldins *via* the Ag43 autotransporter. (a) CFP−AcDoc binding to EM42 or EM371 cells displaying AcCoh or AcCoh−CtCoh scaffoldins. (b) GFP−CtDoc binding to EM42 or EM371 cells displaying CtCoh or AcCoh−CtCoh scaffoldins. Cells with the pSEVA_*ag43AT* plasmid were used as a negative control (Ctrl−). All figures are in the same scale, the white bar size is 4 μm.

Quantification of the Dockerin-Tagged Enzyme Molecules Bound to a Single *P. putida* Cell. The verified exclusivity of AcCoh−AcDoc and CtCoh−CtDoc bonds (Figure 4) together with the assumed 1:1 dockerin−cohesin binding ratio allowed us to estimate the number of enzyme molecules docked to the surface of a single *P. putida* cell. The calculated theoretical number of BglC−CtDoc molecules in the control reaction with 1 μg of the chimeric enzyme (Figure 4b) was $\sim 9.49 \times 10^{12}$, and the approximate number of *P. putida* cells in suspension used for the whole-cell β-glucosidase activity measurements was $\sim 1.14 \times 10^{9}$. Knowing these values, we estimated the counts of cohesin−dockerin bonds per bacterium to be $\sim 1.8 \times 10^{3}$ or $\sim 4.0 \times 10^{3}$ for the EM42 strain displaying the CtCoh or AcCoh−CtCoh scaffoldins, respectively, and $\sim 7.8 \times 10^{3}$ or $\sim 11.0 \times 10^{3}$ for the EM371 strain decorated with CtCoh or AcCoh−CtCoh, respectively. The amount of designer scaffoldins displayed on a single *P. putida* EM371 cell *via* an autotransporter system was similar to the number of comparable single- and two-cohesin scaffoldins exported to the surface of the Gram-positive bacterium *L. lactis* in the study of Wieczorek and Martin.[35] On the contrary, the number of recombinant protein molecules displayed on the EM42 strain did not exceed the values typically reported for Gram-negative hosts such as *E. coli*.[61,74] These calculations highlighted the superiority of the *P. putida* EM371 over the EM42 strain in terms of scaffoldin secretion efficiency.

Though up to tens of thousands of scaffoldins can be displayed on the surface of single EM371 cell, this system needs further optimization to anchor the number of extracellular enzymes that would be sufficient for degradation of lignocellulosic residues or other polymeric compounds. It is known that grams of free cellulases are required for efficient hydrolysis of hundreds of grams of lignocellulosic solids in industrial processes,[36,75] and perhaps hundreds of milligrams of clustered cellulosomal enzymes would be enough considering the higher activity of these complexes.[23,76]

With ∼5−10 mg of β-glucosidase bound to the surface of *P. putida* EM371 with two-cohesin scaffoldin in a hypothetical 1 L culture of OD_{600} of 10, we are still far from the needed values. However, further improvement of scaffoldin display in *P. putida* is possible because the outer membrane of a Gram-negative bacterium can accommodate up to hundreds of thousands of protein molecules,[77] and the secretion system presented here has not yet been optimized for the given host. More scaffoldins could be available on the cell surface also after, *e.g.*, optimization of *ag43AT* expression,[51] removal of periplasmic and outer membrane proteases that can cleave autotransporter or its passenger,[78] or addition of secretion-promoting CBM to the N-terminus of displayed cohesins.[35] Besides, the binding of enzymes to the displayed scaffoldins could be further enhanced by the adoption of cohesin−dockerin pairs from extreme thermophiles[22,79] or by increasing the length of linkers between individual cohesins.[80]

Viability Tests and Growth of EM42 and EM371 Recombinants with Surface-Docked β-Glucosidase in Minimal Medium with Cellobiose. As discussed above,

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

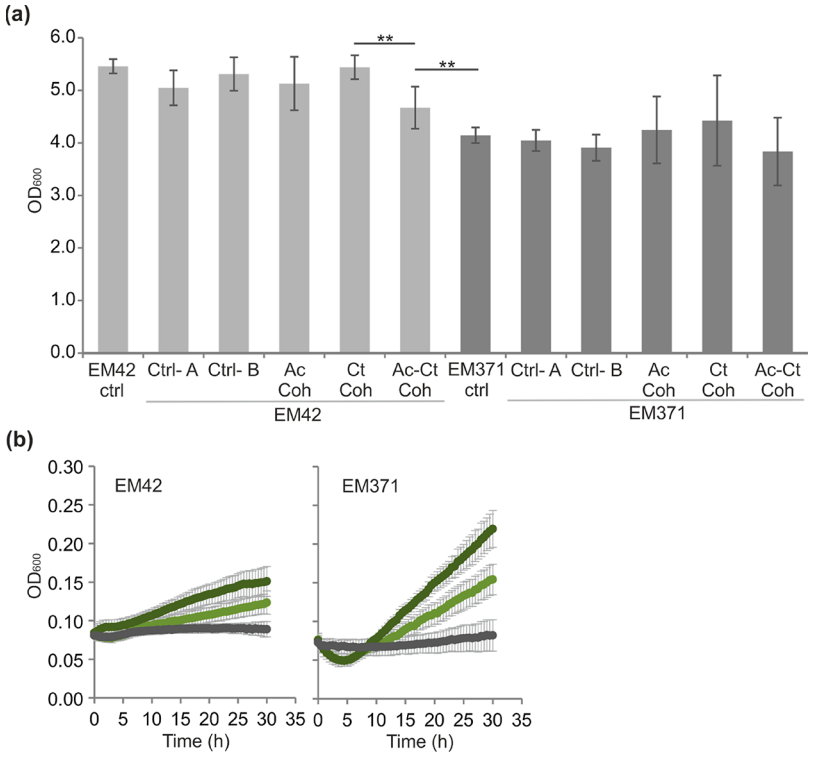


Figure 6. Viability of *Pseudomonas putida* EM42 and EM371 strains displaying designer scaffoldins *via* the Ag43 autotransporter. (a) Optical density of *P. putida* cultures measured after 5 h of induction with 0.5 mM 3-methylbenzoate. Used controls were as follows: EM42 ctrl and EM371 ctrl, plasmid-free EM42 and EM371 cells; Ctrl− A, cells with empty pSEVA238; Ctrl− B, cells with pSEVA238_*ag43AT*. Data are shown as mean ± SD from at least three independent experiments, each conducted in two biological replicates. Asterisks denote significance in the difference in between two means at $P < 0.01$ (**). (b) Growth of *P. putida* recombinants with *β*-glucosidase molecules attached to the displayed two-cohesin miniscaffoldin in minimal medium with cellobiose (5 g L^{-1}). EM42 and EM371 cells with displayed AcCoh−CtCoh were incubated with purified BglC−CtDoc only (pale green line) or with a mixture of BglC−CtDoc and BglC−AcDoc (dark green line), washed, and grown at 30 °C in wells of 96-well microtiter plates containing M9 minimal medium with 5 g L^{-1} D-cellobiose. EM42 and EM371 strains bearing pSEVA238b_*ag43AT* plasmid were used as controls (gray line). Error bars show standard deviations from three biological replicates, each conducted in two technical replicates.

autodisplay of recombinant proteins can affect the viability of a host, which might eventually hamper the applicability of a whole-cell biocatalyst. We compared the final optical densities of the cultures conducted to prepare *P. putida* recombinants for the aforementioned assays to evaluate the effect of scaffoldin expression on the viability of EM42 and EM371 strains (Figure 6). Absorbance of cultures with the EM42 recombinant secreting the two-cohesin scaffoldin was reduced by 9% ($P < 0.05$) or 14% ($P < 0.01$) when compared with cultures of EM42 pSEVA238_*ag43AT*−*acCoh* or EM42 pSEVA238_*ag43AT*−*ctCoh*, respectively. In the case of EM42 *acCoh*−*ctCoh*$^+$ cultures, the 14% decrease in final OD_{600} corresponded to an approximate reduction in 3.50×10^8 cells or 1.41×10^{12} BglC molecules per milliliter of culture. Nonetheless, given the 2.3-fold greater capacity of the AcCoh−CtCoh-decorated cells to bind BglC−CtDoc (Figures 4 and 5), the total theoretical number of cell-bound BglC molecules was still 2-fold higher in EM42 *acCoh*−*ctCoh*$^+$ cultures as compared to the one-scaffoldin counterpart. This underscores that the cell viability concessions observed were overshadowed by the greater capacity for AcCoh−CtCoh-decorated cells to bind dockerin-tagged proteins. There were no statistically significant differences in viability among EM371 recombinants with displayed scaffodins (Figure 6a).

It must be emphasized here that all EM371 recombinants and controls grew slower than their EM42 counterparts, and their final ODs were ∼20% lower ($P < 0.01$) (Figure 6a). This observation is in agreement with the previous study of Martínez-García and co-workers in 2020[46] and can be attributed to the reduced fitness of the EM371 strain. Reduced viability of EM371, when compared with the robust EM42 strain, is a concern. If scaffoldin-bearing *P. putida* strains should serve as enzyme carriers, they must provide both sufficiently high cell densities and a considerable capacity to attach recombinant proteins to their surface. The data presented in this work, nonetheless, indicate that the benefit of greater display efficiency of EM371 can compensate for lower vigor of this strain.

To verify this assumption and to test the two parameters simultaneously, we incubated the AcCoh−CtCoh-displaying EM42 and EM371 strains with either BglC−CtDoc only (we will call here these strains EM42+ and EM371+ for simplification) or with both BglC−CtDoc and BglC−AcDoc (EM42++ and EM371++ in the following text), and we let the washed cells grow in minimal medium with cellobiose used as a sole carbon source (Figure 6b). The assay confirmed that all four tested strains bound such a quantity of *β*-glucosidase molecules, which provided *P. putida* cells with an amount of glucose sufficient for growth. The observed growth was slow and linear, probably due to the constant limited amount of *β*-glucosidase molecules in the reaction. Importantly, EM371 recombinants clearly outperformed EM42 strains (Figure 6b).

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

The EM371++ strain grew ~ 3 times faster than EM42++ (μ = 0.064 and 0.023 h^{-1}, respectively), and the EM371+ strain grew ~ 4 times faster than EM42+ (μ = 0.048 and 0.013 h^{-1}, respectively). It is also noteworthy that EM42++ and EM371+ + grew significantly faster ($P < 0.01$) than the EM42+ and EM371+ strains. This indicated that both cohesins at least in some portion of displayed AcCoh−CtCoh scaffoldins were occupied by BglC−AcDoc and BglC−CtDoc at the same time, and there was an additive effect of the two BglC molecules. The specific growth rate of EM42+ was ~57% of the growth rate of EM42++ and the μ of EM371+ was ~75% of EM371+ +'s rate. This result was in agreement with the outcome of the above-described assays (Figure 4) and suggested once again that β-glucosidase was better attached to the cell surface through the CtDoc−CtCoh interaction than through the AcDoc−AcCoh binding pair.

The described experiment was also an addition to our previously published study in which we identified the growth of *P. putida* with cytoplasmic BglC on cellobiose.[4] Here, we demonstrated that this enzyme can be tagged with dockerin, overproduced in *P. putida*, purified, and turned into a cellulosomal mode on the same host, while its activity with a natural substrate is preserved. Intracellular BglC production is nonetheless a more attractive option for our future research. In such a scenario, only two enzymes—an endoglucanase and an exoglucanase or a cellobiohydrolase—need to be assembled in a functional minicellulosome on the surface of a *P. putida* recombinant to enable the decomposition of crystalline cellulose and its utilization for cellular growth. Moreover, well-expressed intracellular β-glucosidase would not become a bottleneck for efficient cellulose hydrolysis as is often reported[40,81] but, on the contrary, would secure rapid drain of cellobiose from the cellular surface and prevent inhibition of the remaining cellulases.[82]

Besides the limited number of displayed scaffoldin molecules on the *P. putida* surface, the supplementation of dockerin-tagged enzymes is a remaining challenge for *in vivo* assembly of functional designer cellulosome-like structures in the Gram-negative host. The addition of purified enzymes is an expensive option, and high-yield secretion of recombinant proteins in bacteria with two fosfolipid membranes is problematic.[40,57,83] This obstacle could be bypassed, *e.g.*, by the design of a co-operative single-species consortium in which a smaller part of the population overexpresses required enzymes and is sacrificed *via* the induced autolysis to supply biocatalysts to the production strain(s) bearing scaffoldins.[84] This and other strategies are being considered to complement the work described here and pave the way to *P. putida*-based cell factories for the valorization of polymeric waste feedstocks.

On the top of this, the developed autodisplay system in *P. putida* could be adapted for alternative applications such as screening of libraries of antimicrobial peptides,[85] development of whole-cell electrochemical biosensor,[86] or display of enzymes, antibodies, or metal-binding proteins for pollutant removal from the environment.[87]

CONCLUSIONS

We have evaluated the capacity of two biotechnologically relevant strains of the Gram-negative bacterium *P. putida*, EM42 and EM371 with differences in the complexity of the cellular surface, to display variants of designer scaffoldin proteins and serve as future carriers for cellulosome-like structures. Single- and two-cohesin scaffoldins were displayed *via* the Ag43 autotransporter on the surfaces of the *P. putida* EM42 and EM371 recombinants. The resulting complexes were not cellulolytic, but their investigation allowed insight into parameters affecting the assembly of such synthetic structures on the surface of a Gram-negative bacterial host. We confirmed both the nature of a displayed scaffoldin and the character of a host cell's surface matter.

Two-cohesin scaffoldin was more accessible on the surfaces of the EM42 and EM371 strains in the majority of tested scenarios, and the cohesin−dockerin pair from thermophilic *C. thermocellum* showed better bonding than the pair from mesophilic *A. cellulolyticus*. The EM371 strain with surface-bound β-glucosidase or fluorescent proteins outperformed the EM42 recombinant in all conducted assays including the growth on natural substrate cellobiose. This indicates that, despite its slightly compromised fitness, *P. putida* EM371 is a promising platform for attachment of designer catalytic scaffoldins and other applications that benefit from microbial secretion of recombinant proteins, and "surface shaving" represents a viable strategy for the design of new Gram-negative bacterial catalysts.

MATERIALS AND METHODS

Bacterial Strains, Media, and Growth Conditions. Bacterial strains used in this study are listed in Supporting Information Table S1. *Escherichia coli* strains employed for cloning or triparental mating and *Pseudomonas putida* strains used for heterologous gene expression were routinely grown in lysogeny broth (LB; 10 g L^{-1} tryptone, 5 g L^{-1} yeast extract, 5 g L^{-1} NaCl, pH 7.0) at 37 or 30 °C, respectively. All strains were routinely precultured overnight (15 h) in 2.5 mL of LB medium with agitation of 300 rpm (Heidolph Unimax 1010 and Heidolph Incubator 1000; Heidolph Instruments). All used solid media (LB or M9 salts minimal medium; per 1 L: 8.5 g of $Na_2HPO_4{\cdot}2H_2O$, 3.0 g of KH_2PO_4, 1.0 g of NH_4Cl, 0.5 g of NaCl, pH 7.0) contained 15 g L^{-1} agar. Solid M9 salts media were prepared with 2 mM $MgSO_4$, 2.5 mL L^{-1} trace element solution,[88] and 0.2% (w/v) citrate used as a sole carbon source for *P. putida* strains. Antibiotics of following final concentrations were added to the liquid and solid media to maintain used plasmids: kanamycin (Km) 50 μg mL^{-1}, chloramphenicol (Cm) 30 μg mL^{-1}, ampicillin (Amp) 150 μg mL^{-1}. Growth conditions specific for individual experiments are described in the following sections.

Plasmid and Strain Construction. All used and constructed plasmids from this study are listed in Supporting Information Table S1. Standard laboratory protocols[89] were used for DNA manipulations. Oligonucleotide primers used in this study (Supporting Information Table S2) were purchased from Sigma-Aldrich. Plasmid DNA was isolated with QIAprep Spin Miniprep kit (Qiagen). The genes of interest were amplified by polymerase chain reaction (PCR) using Q5 high fidelity DNA polymerase (New England BioLabs) according to the manufacturer's protocol. The reaction mixture (50 μL) consisted of polymerase HF buffer (New England BioLabs), water, dNTPs mix (0.2 mM each; Roche), template DNA, primers (0.5 mM each), and GC enhancer (in case of high GC content in an amplified gene; New England BioLabs). Two-step overlap extension (OE) PCR[90] with Q5 DNA polymerase was adopted for the preparation of chimeric genes. PCR products from the first reaction were purified with NucleoSpin Gel and PCR Clean-up (Macherey-Nagel) and used (5 ng) as templates in the second PCR round. NucleoSpin Gel and PCR

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Clean-up kit was also routinely used for purification of PCR products either directly from PCR mixture or agarose gels. Colony PCR was performed using NZY*Taq*II 2x Green Master Mix solution (NZYTech). DNA concentration was measured with a NanoVue spectrophotometer (GE Healthcare). All restriction enzymes and Quick Ligation kit used for ligation of digested fragments were from New England BioLabs. Digested plasmids and PCR products were separated by DNA electrophoresis with 0.8% (w/v) agarose gels and visualized using Molecular Imager VersaDoc (Bio-Rad). The flawlessness of PCR-amplified genes cloned into target plasmids was checked by DNA sequencing (Macrogen). Chemocompetent *E. coli* cells (CC118 or DH5α) transformed with plasmids or ligation mixtures were selected on LB agar plates with respective antibiotic, single clones were restreaked on new LB agar plates with an antibiotic, and grown cells were collected in 1 mL of LB with glycerol (20% w/v) and stored at −80 °C. Plasmid constructs were transferred from *E. coli* Dh5α or CC118 donor to *P. putida* EM42 or EM371 by triparental mating, using *E. coli* HB101 helper strain with pRK600 plasmid (Supporting Information Table S1). Alternatively, electroporation (2.5 kV, 4−5 ms pulse) was used for the transformation of *P. putida* cells with selected plasmids using a MicroPulser electroporator and Gene Pulser Cuvettes with a 0.2 cm gap (Bio-Rad). Preparation of *P. putida* electrocompetent cells and electroporation procedure itself was performed as described elsewhere.[91] *P. putida* transconjugants or transformants were selected on M9 agar plates with citrate or on LB agar plates, respectively, with respective antibiotic. All plasmid constructs inserted in *P. putida* were first isolated and rechecked by restriction digestion before the strain was used for further work.

Preparation of Scaffoldin Variants, Chimaeric Proteins, and Autotransporters. The original synthetic scaffoldin gene *scaf19L*[47] encodes carbohydrate-binding module CBM3a and cohesin CohCt A2 (named CtCoh in this study) of the cellulosomal-scaffolding protein CipA from *Clostridium thermocellum*, cohesin CohAc C3 (named AcCoh in this study) of the cellulosomal-scaffolding protein ScaC from *Acetivibrio cellulolyticus*, and cohesin CohBc B3 (named BcCoh in this study) of the cellulosomal-scaffolding protein ScaB from *Bacteroides cellulosolvens* interconnected with 27−35 aa long linkers. The *scaf19L* gene was PCR amplified using Q5 polymerase and primers Sca19L fw and Sca19L rv (Supporting Information Table S2) and subcloned into *Nde*I and *Pst*I restriction sites of the modified version of pSEVA238 expression plasmid, pSEVA238b, with synthetic ribosome binding site (RBS).[4] In parallel, a version of the *scaf19LKT* gene with synthetic RBS was synthesized and codon-optimized for expression in *P. putida* KT2440 (GeneCust). The synthetic gene was subcloned from delivery vector pUC57_*scaf19LKT* into *Sac*I and *Kpn*I sites of pSEVA238.

The *bglC* gene encoding β-glucosidase (EC 3.2.1.21) from *Thermobifida fusca* with an N-terminal 6x histidine tag and *ctDoc* gene (encoding *C. thermocellum* dockerin, named here as CtDoc, complementary to CtCoh) codon-optimized for expression in *P. putida* KT2440 (GeneCust) were separately amplified by PCR from pSEVA238b_*bglC*[4] and pUC57_*ctDoc*, respectively, using primer pairs BglC−CtDoc TS1F/BglC−CtDoc TS1R and BglC−CtDoc TS2F/BglC−CtDoc TS2R. The first round PCR products were sewed in the second round of OE PCR with TS1F and TS2R primers. The resulting *bglC*−*ctDoc* chimeric gene was cloned into *Nde*I and *Hind*III sites of pSEVA238b. The pSEVA238b_*bglC*−*acDoc* construct bearing *bglC* gene tagged with codon-optimized *A. cellulolyticus* ScaB dockerin, named here AcDoc, was prepared correspondingly using primer pairs BglC−CtDoc TS1F/BglC−AcDoc TS1R and BglC−AcDoc TS2F/BglC−AcDoc TS2R and template plasmids pSEVA238b_*bglC* and pUC57_*acDoc*.

For the purpose of construction of the plasmid allowing the translational fusion of CtDoc to monomeric superfolded GFP (msfGFP), the *gfp* gene was initially amplified with synthetic RBS but without a STOP codon from pSEVA238_*gfp* plasmid (SEVA collection) using GFP-N fw and GFP-N rv primers. The PCR product was digested with *Avr*II and *Eco*RI and ligated into pSEVA238, cut with the same pair of enzymes, giving rise to pSEVA238_*gfpN*. The *ctDoc* gene was amplified from pSEVA238b_*bglC*−*ctDoc* with its synthetic SGGGS Gly−Ser linker and with its TAA STOP codon using CtDoc fw and CtDoc rv primers. The PCR product was digested with *Sac*I and *Hind*III and cloned downstream of the *gfp* gene in pSEVA238_*gfpN*, resulting in pSEVA238_*gfp*−*ctDoc*. The *gfp*−*ctDoc* construct was then subcloned into *Nde*I and *Hind*III sites of pET21b for the purpose of gene overexpression and purification of the chimeric protein.

The pSEVA238b_*ag43AT* construct was prepared by subcloning the recombinant Ag43 autotransporter gene from pAg43pol[38] into *Nde*I and *Hind*III sites of the pSEVA238b vector with synthetic RBS. The *estPAT* gene encoding C-terminal part (331 AA) of EstP esterase autotransporter from *P. putida* KT2440 (PP_0418) with original 23 aa N-terminal leader sequence, E-tag, and polylinker was commercially synthesized (GeneCust) and then subcloned from the delivery vector pUC57 into *Nde*I and *Hind*III sites of pSEVA238b.

Single cohesin gene *ctCoh* was PCR amplified from pSEVA238_*scaf19LKT* for the purpose of subcloning into the polylinkers of three tested autotransporter systems using primer pair CtCoh fw and CtCoh rv1 (for cloning into *Xho*I and *Bam*HI sites of pSEVA238b_*ag43AT* polylinker) or CtCoh fw and CtCoh rv2 (for cloning into *Eco*RI and *Bam*HI sites of pSEVA238_*igAAT* and pSEVA238_*intAT*). These manipulations gave rise to the constructs pSEVA238_*igAAT*−*ctCoh*, pSEVA238_*intAT*−*ctCoh*, and pSEVA238b_*ag43AT*−*ctCoh*. The pSEVA238b_*ag43AT*−*acCoh* construct was prepared by subcloning of the *acCoh* gene, PCR amplified from pSEVA238_*scaf19LKT* with primers AcCoh fw and AcCoh rv, into *Xho*I and *Bam*HI sites of pSEVA238b_*ag43AT* polylinker. Similarly, the two-cohesin scaffoldin sequence *acCoh*−*ctCoh* was amplified from pSEVA238_*scaf19LKT* using the AcCoh−CtCoh fw and CtCoh rv1 primers and inserted into *Xho*I and *Bam*HI sites of pSEVA238b_*ag43AT*.

Analysis of Cohesin−Dockerin Interactions by Affinity-Based ELISA. Proper folding and ability of AcCoh, CtCoh, and BcCoh cohesins in Scaf19LKT scaffoldin produced in *P. putida* to bind respective dockerins was verified by affinity-based ELISA (enzyme-linked immunosorbent assay). *P. putida* EM42 pSEVA238_*scaf19LKT* was pregrown overnight in 2.5 mL of LB with Km and 10 mM $CaCl_2$. The cells were then used to inoculate 50 mL of LB medium with Km and 10 mM $CaCl_2$ in the main culture to the starting OD_{600} of 0.05. The cells were cultured for 3 h at 30 °C with shaking (200 rpm), and expression of scaffoldin gene was then induced with 1 mM 3MB. After 5 more hours of growth, the cells were collected by centrifugation (2500*g*, 4 °C, 15 min) and washed with ice-cold TBS buffer (25 mM Tris-Cl, 137 mM NaCl, 2.7 mM KCl, pH 7.2) with 10 mM $CaCl_2$ and

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

0.05% Tween 20, and the cell pellet was frozen at −80 °C overnight. The pellet was then melted and added with 3 mL of TBS buffer, 0.3 mL of PopCulture Reagent (Merck Millipore), 10 μL of DNase I (5 μg mL^{-1}; Sigma-Aldrich), and 10 μL of lysozyme (25 mg mL^{-1} stock; Sigma-Aldrich). The cell suspension was incubated for 30 min at room temperature and centrifuged (21 000*g*, 4 °C, 30 min), and the supernatant (cell-free extract, CFEs) was used for further work. The concentration of total protein in CFE with recombinant Scaf19LKT was determined by Bradford reagent (Sigma-Aldrich), CFE was diluted 1000 times in 0.1 M sodium carbonate coating buffer (pH 9.0), and 100 μL (the same volume was used for all following steps) was used for overnight coating of individual wells in a MaxiSorp high protein-binding capacity 96-well ELISA plate (Nunc) at 4 °C. CFE prepared from *P. putida* EM42 bearing empty pSEVA238 was used as a control. In the morning, CFE was discarded from the plate, blocking buffer (TBS with 10 mM $CaCl_2$ and 0.05% Tween 20, and 2% BSA) was added to the wells, and the plate was incubated at room temperature for 1 h (the same incubation conditions were preserved in the remaining steps of the protocol). Blocking buffer was then removed, and the wells were added with fresh blocking buffer containing 0.1 μg mL^{-1} concentration of one of three purified recombinant variants of *Geobacillus sp.* WBI xylanase Xyn−AcDoc, Xyn−CtDoc, or Xyn−BcDoc were tagged with a dockerin module from *A. cellulolyticus*, *C. thermocellum*, or *B. cellulosolvens*, respectively. After the incubation step, the blocking buffer with proteins was discarded and the wells were washed three times with wash buffer (TBS with 10 mM $CaCl_2$ and 0.05% Tween 20). Anti-Xyn primary antibody diluted 10 000× in the blocking buffer was then added to the wells. After incubation, the washing step was repeated (three washes) and the secondary antibody (HRP-labeled anti-rabbit) diluted 10 000× in blocking buffer was added into the wells. This was followed by incubation, wash steps (four washes), and final detection with 100 μL of 3,3′,5,5′-tetramethylbenzidine (TMB) and Substrate-Chromogen (Dako). The reaction occurred for 30 s and then was stopped by the addition of 50 μL of 1 M H_2SO_4 per well. The intensity of the resulting color was measured spectrophotometrically at 450 nm.

Protein Purification by Affinity Chromatography. His-tagged BglC, BglC−AcDoc, BglC−CtDoc, CFP−AcDoc, and GFP−CtDoc were purified from CFE prepared from *P. putida* EM42 or *E. coli* BL21-Gold (DE3) (Agilent Technologies) cells. Overnight cultures in 10 mL of LB with Km and 2 mM $CaCl_2$ were inoculated from single colonies of *P. putida* EM42 pSEVA238_*bglC*, *P. putida* EM42 pSEVA238_*bglC−AcDoc*, *P. putida* EM42 pSEVA238_*bglC−CtDoc*, or *E. coli* BL21-Gold (DE3) pET28a_*cfp−AcDoc* on LB agar plates. Medium with Amp was used for the *E. coli* BL21-Gold (DE3) pET21b_*gfp−ctDoc* recombinant. Overnight cultures were used for inoculation of 200 mL of fresh LB medium with Km or Amp and 2 mM $CaCl_2$ to the final OD_{600} of 0.05. Cells were grown with shaking (170 rpm) for 2.5 h, and expression of chimeric genes was induced by 1 mM 3MB (*P. putida* with pSEVA238 plasmids) or 50 μM isopropyl β-D-1-thiogalactopyranoside (*E. coli* with pET plasmids). After induction, the *P. putida* cells were further cultured under the same conditions for another 5 h, while the *E. coli* cultures were grown overnight at a reduced temperature of 20 °C. Cells were then pelleted by centrifugation (2000*g*, 15 min, 4 °C), washed with ice-cold purification buffer A (TBS with 10 mM $CaCl_2$, 0.05% Tween 20, and 5 mM imidazole, pH 7.2), centrifuged again, resuspended in 5 mL of the same buffer, and frozen at −80 °C. The next day, the cell suspension was melted, added with one-fourth of cOmplete EDTA-free protease inhibitor cocktail tablet (Roche, Switzerland) and 5 μL of Lysonase Bioprocessing Reagent (Merck Millipore), and sonicated 6 × 2 min on ice until the suspension became transparent. The cell lysate was centrifuged (21 000*g*, 4 °C, 30 min), and the resulting CFE was collected. Total protein concentration in CFE was determined by Bradford reagent (Sigma-Aldrich), and ∼55 mg of total protein in CFE was incubated for 60 min at 4 °C with 2 mL of Ni-NTA agarose (QIAGEN) equilibrated with purification buffer A. The slurry was applied to a 10 mL Poly-Prep chromatography column (Bio-Rad), which was then washed with 10 mL of purification buffer B (purification buffer A with 50 mM imidazole). His-tagged protein was eventually eluted with purification buffer C (purification buffer A with 500 mM imidazole) in 1.5 mL fractions. Fractions with the highest β-glucosidase activity or GFP or CFP fluorescence were pooled and applied to an Amicon Ultra centrifugal filter unit with 10 kDa cutoff (Merck Millipore) for protein concentration and buffer exchange (TBS with 10 mM $CaCl_2$, 0.05% Tween 20, and 10% glycerol). The concentration of purified proteins was determined by Bradford reagent, and proteins were stored at 4 °C or at −20 °C for further use.

β-Glucosidase Activity Assay. β-Glucosidase activity of purified BglC, BglC−CtDoc, and BglC−AcDoc was measured using the synthetic substrate *p*-nitrophenyl-β-D-glucopyranoside (pNPG; Sigma-Aldrich) following the previously described protocol[4] with some modifications. Briefly, the reaction mixture (total volume 1200 μL) contained 1138 μL of 50 mM sodium phosphate buffer (pH 7.0), 60 μL of pNPG (final concentration 5 mM), and 2 μL of enzyme of concentration adjusted to 0.1 mg mL^{-1}. The reaction was run in 1.5 mL test tubes at 37 °C. Samples (600 μL) were withdrawn at 10 and 20 min intervals and mixed with 400 μL of 1 M Na_2CO_3 to stop the reaction, and the absorbance of the mixture was measured in a cuvette at 405 nm with an UV/vis spectrophotometer Ultrospec 2100 (Biochrom). Linearity of the enzymatic reaction during the given time interval was checked prior to these measurements. The specific activity (U mg^{-1}) was calculated using the calibration curve, prepared with a *p*-nitrophenol standard (Sigma-Aldrich). One unit (U) of enzymatic activity corresponds to 1 μmol of *p*-nitrophenol produced per minute.

Evaluation of Scaffoldin Display and Cohesin−Dockerin Binding on the *P. putida* Cell Surface by β-Glucosidase Activity Assay. The following culture procedure was used for the preparation of cells for subsequent whole-cell assays as well as for evaluation of recombinant protein levels in derived *P. putida* CFEs using SDS-PAGE and Western blot analyses. *P. putida* EM42 or EM371 strains, bearing empty pSEVA238 plasmid (negative control) or pSEVA238 plasmid harboring an autotransporter gene with a single- or double-cohesin scaffoldin, were inoculated directly from glycerol stocks into 2.5 mL of LB medium with Km and 10 mM $CaCl_2$ and grown overnight with shaking. Overnight cultures were used to inoculate 10 mL of LB medium with Km and 10 mM $CaCl_2$ to a starting OD_{600} of 0.05. Cells were cultured for 2.5 h with shaking (250 rpm, Unimax 1010 shaker and 1000 Inkubator, Heidolph), and expression of the autotransporter−scaffoldin chimera was then induced with 0.5 mM 3MB. After 4.5 h of induction, the final OD_{600} was

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

measured and cells were collected by centrifugation (2000*g*, 4 °C, 10 min). Cells were washed with 10 mL of ice-cold TBS buffer with 10 mM $CaCl_2$ and 0.05% Tween 20, centrifuged again, and resuspended in 0.5 mL of the same buffer to an OD_{600} of 10.0. Then, 25 μL of purified BglC, BglC−CtDoc, or BglC−AcDoc at 1 mg mL^{-1} concentration (~1.0 × 10^5 molecules per cell) was added to the cell suspension, and the whole mixture was incubated in a 1.5 mL test tube for 14 h (overnight) at 4 °C with gentle rotation of the tube (SB2 Rotator, Stuart).

For comparison of the three distinct autotransporters displaying CtCoh on the surface of *P. putida* EM42 or EM371, the cell suspension (250 μL) was washed three times with 1 mL of ice-cold TBS buffer with $CaCl_2$ and Tween 20, resuspended in 250 μL of the same buffer (warmed to room temperature), and added with 5 μL of 100 mM pNPG. The same mixture with EM42 or EM371 cells bearing an empty pSEVA238 plasmid served as a negative control. The mixture in TBS buffer without cells but with purified BglC (1 μL of enzyme of 1 mg mL^{-1} concentration) served as a positive control. The mixtures were incubated for 60 min at 37 °C with shaking (300 rpm, Unimax 1010 shaker and 1000 Inkubator, Heidolph). Cells were then pelleted by brief centrifugation (2370*g*, room temperature, 3 min), supernatants (100 μL) were mixed with 50 μL of 1 M Na_2CO_3 in a 96-well plate, and absorbance at 405 nm was measured using a Victor2 1420 Multilabel Counter (PerkinElmer). End-point absorbance values, which corresponded with the activity of the cell surface-bound β-glucosidase molecules, were related to the absorbance measured for free β-glucosidase (positive control) and used for comparison of the displaying capacity of the tested autotransporters in the EM42 and EM371 strains. Possible cross-reactivity of AcDoc with CtCoh and CtDoc with AcCoh was tested using the aforementioned protocol applied on EM371 cells, which displayed Ag43AT−CtCoh or Ag43AT−AcCoh fusion and were incubated with purified BglC−AcDoc or BglC−CtDoc protein, respectively.

Alternatively, for the purpose of quantitative evaluation of BglC−CtDoc or BglC−AcDoc binding to a respective scaffoldin variant (CtCoh, AcCoh, or AcCoh−CtCoh) displayed on EM42 and EM371 cells with Ag43AT, a cell suspension after overnight incubation with complementary BglC−XDoc chimera was washed three times with ice-cold TBS buffer containing $CaCl_2$ and Tween 20 and diluted twice with the same buffer (OD_{600} = 5.0), and 0.5 mL was warmed in a 1.5 mL test tube for 10 min at 37 °C. Suspension with EM42 or EM371 cells bearing pSEVA238b_*ag43AT* served as negative control. Suspensions were then added with 10 μL of 100 mM pNPG to start the enzymatic reaction. A positive control reaction was run in TBS buffer with the addition of 1 μL of purified BglC−CtDoc or BglC−AcDoc at 1 mg mL^{-1} concentration. Samples (100 μL) were withdrawn every 5 min (total length of the reaction was 20 min) and mixed with 50 μL of 1 M Na_2CO_3, and the cells were removed by centrifugation. The supernatant (100 μL) was transferred to the 96-well plate, and absorbance at 405 nm was measured using a Victor2 1420 Multilabel Counter (PerkinElmer). In the case of low β-glucosidase activity detected, the time intervals of the sample withdrawal were extended to allow for accurate activity quantification. Activity (1 U = 1 μM min^{-1}) was calculated using a calibration curve prepared with a *p*-nitrophenol standard (Sigma-Aldrich). Activity of the enzyme bound to the surface of each of the tested recombinants was related to the calculated activity of free purified BglC−CtDoc or BglC−AcDoc and used for quantitative comparison of the enzyme-binding capacity of the EM42 and EM371 strains displaying single-cohesin or two-cohesin scaffoldin.

The number of enzyme molecules bound to the surface of a given EM42 or EM371 recombinant was estimated from the molecular weight of BglC−CtDoc (63.46 kDa) or BglC−AcDoc (63.94 kDa) calculated using the ExPASy Compute pI/Mw tool (https://web.expasy.org) and from the average number of *P. putida* cells present in a 0.5 mL suspension of OD_{600} of 5.0 (~1.1375 × 10^9), determined by counting the cells in samples diluted to OD_{600} of 0.2 in a Bürker chamber (Brand) using an Olympus BX50 microscope with a 100× oil immersion objective (Olympus). Assuming a dockerin/cohesin ratio of 1:1, the calculated amount of β-glucosidase molecules anchored to the cell corresponds to the number of scaffoldins displayed on the cell surface.

Evaluation of Scaffoldin Display and Cohesin−Dockerin Binding on the *P. putida* Cell Surface Using Fluorescent Proteins. EM42 and EM371 cells for quantitative evaluation of GFP−CtDoc or CFP−AcDoc binding to scaffoldin variants displayed with Ag43AT were prepared as described in the previous section. The cells (0.5 mL, OD_{600} = 1.0) were incubated overnight at 4 °C with 25 μL of purified dockerin-tagged fluorescent protein of 1 mg mL^{-1} stock concentration (~1.8 × 10^5 fluorophore molecules per cell). After washing and dilution with TBS buffer, the cell suspension was transferred (150 μL per well) to black, clear-bottom, 96-well microplates (Corning), and the optical density (600 nm) and GFP (475 nm excitation/515 nm emission) or CFP (440 nm excitation/500 nm emission) fluorescence were measured with a SpectraMax iD5 microplate reader (Molecular Devices). Recorded fluorescence was normalized by cell density.

Confocal Microscopy. Confocal microscopy was used to visually evaluate GFP−CtDoc or CFP−AcDoc binding to scaffoldin variants displayed with Ag43AT on the surface of EM42 and EM371 strains. Cells bearing pSEVA238b_*ag43AT* were used as a negative control. After incubation with a purified fluorescent protein, a cell suspension was washed and diluted in TBS buffer and 5 μL was dropped on poly-L-lysine coated glass slides (Sigma-Aldrich) and dried for 60 min at room temperature. Then, the cells were mounted with 5 μL of ProLong antifade reagent (Thermo Fisher Scientific) and covered with cover glass, and the slides were analyzed using an inverted confocal microscope Leica DMi8 S (Leica Microsystems) using a 100× oil immersion objective and 6× digital zoom.

Growth of *P. putida* Recombinants with Surface-Attached β-Glucosidase in Minimal Medium with Cellobiose. Preinduced *P. putida* EM42 and EM371 cells with displayed AcCoh−CtCoh scaffoldins were prepared as described above. EM42 and EM371 strains, bearing the pSEVA238b_*ag43AT* plasmid, were used as negative controls. The cells (0.5 mL, OD_{600} = 10.0) were incubated with 25 μL of purified BglC−CtDoc (1 mg mL^{-1}) or with a mixture (1:1) of BglC−CtDoc and BglC−AcDoc (25 μL each) at room temperature for 1 h. The suspension was washed three times with ice-cold TBS buffer with $CaCl_2$ and Tween 20, and the cells were used to inoculate wells of a 96-well microtiter plate containing M9 minimal medium with 2 mM $MgSO_4$, 100 μM $CaCl_2$, 20 μM $FeCl_3$, 2.5 mL L^{-1} trace element solution, and 5 g L^{-1} D-cellobiose (the only carbon source). The plate with a

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

lid was incubated in a Tecan Infinite 200 Pro reader (Tecan) at 30 °C with discontinuous linear shaking, and absorbance at 600 nm was measured at 30 min intervals.

Statistical Analyses. Experiments reported here were conducted at least in two biological replicates (the number of experiments and replicates is specified in figure legends). The mean values and corresponding standard deviations (SD) are presented. When appropriate, data were treated with a two-tailed Student's *t* test in Microsoft Office Excel 2013 (Microsoft Corp.), and confidence intervals were calculated for a given parameter to manifest a statistically significant difference in means between two experimental data sets.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acssynbio.0c00276.

Discussions of effect of scaffoldin and autotransporter expression on the viability of *P. putida* cells, SDS-PAGE and Western blot analyses, and dot blot analysis, tables of strains, plasmids, and oligonucleotide primers used, and figures of schematic reconstruction, effect of *scaf19L* and *scaf19LKT* expression, Western blot analysis, SDS-polyacrylamide gels, effect of an autotransporter expression, and supporting sequences (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Pavel Dvořák − *Department of Experimental Biology (Section of Microbiology), Faculty of Science, Masaryk University, 62500 Brno, Czech Republic;* orcid.org/0000-0002-3215-4763; Phone: +420 549 493 396; Email: pdvorak@sci.muni.cz

Víctor de Lorenzo − *Systems and Synthetic Biology Program, Centro Nacional de Biotecnología CNB-CSIC, 28049 Madrid, Spain;* orcid.org/0000-0002-6041-2731; Phone: +34 91 585 4536; Email: vdlorenzo@cnb.csic.es; Fax: +34 91 585 4506

Author

Edward A. Bayer − *Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot 76100, Israel;* orcid.org/0000-0001-7749-5150

Complete contact information is available at: https://pubs.acs.org/10.1021/acssynbio.0c00276

Author Contributions

P.D. and V.d.L. planned the experiments. P.D. did the practical work. All authors analyzed and discussed the data and contributed to the writing of the article.

Funding

This work was funded by the SETH Project of the Spanish Ministry of Science RTI 2018-095584-B-C42, the MADONNA (H2020-FET-OPEN-RIA-2017−1−766975), BioRoboost (H2020-NMBP-BIO-CSA-2018), SYNBIO4FLAV (H2020-NMBP/0500) and MIX-UP (H2020-Grant 870294) Contracts of the European Union, and the S2017/BMD-3691 In-GEMICS-CM Project of the Comunidad de Madrid (European Structural and Investment Funds) as well as by the Czech Science Foundation (19-06511Y).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

Authors thank Prof. Alfredo Martinez from Universidad Nacional Autónoma de México for provided pAg43pol construct, Dr. Sarah Morais and Dr. Johanna Stern from Weizmann Institute of Science, Israel for valuable discussions and help with the ELISA technique, Dr. Esteban Martínez-García from CNB-CSIC, Spain for providing *P. putida* EM371 strain and for valuable discussions, and the Microscopy department of CNB-CSIC for the help with confocal microscopy.

■ REFERENCES

(1) Blank, L. M., Narancic, T., Mampel, J., Tiso, T., and O'Connor, K. (2020) Biotechnological upcycling of plastic waste and other non-conventional feedstocks in a circular economy. *Curr. Opin. Biotechnol. 62*, 212−219.

(2) Baral, N. R., Sundstrom, E. R., Das, L., Gladden, J., Eudes, A., Mortimer, J. C., Singer, S. W., Mukhopadhyay, A., and Scown, C. D. (2019) Approaches for More Efficient Biological Conversion of Lignocellulosic Feedstocks to Biofuels and Bioproducts. *ACS Sustainable Chem. Eng. 7*, 9062−9079.

(3) Peabody, G. L., Elmore, J. R., Martinez-Baird, J., and Guss, A. M. (2019) Engineered *Pseudomonas putida* KT2440 co-utilizes galactose and glucose. *Biotechnol. Biofuels 12*, 295.

(4) Dvořák, P., and de Lorenzo, V. (2018) Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. *Metab. Eng. 48*, 94−108.

(5) Löwe, H., Sinner, P., Kremling, A., and Pflüger-Grau, K. (2020) Engineering sucrose metabolism in *Pseudomonas putida* highlights the importance of porins. *Microb. Biotechnol. 13*, 97−106.

(6) Salvachua, D., Rydzak, T., Auwae, R., De Capite, A., Black, B. A., Bouvier, J. T., Cleveland, N. S., Elmore, J. R., Furches, A., Huenemann, J. D., Katahira, R., Michener, W. E., Peterson, D. J., Rohrer, H., Vardon, D. R., Beckham, G. T., and Guss, A. M. (2020) Metabolic engineering of *Pseudomonas putida* for increased polyhydroxyalkanoate production from lignin. *Microb. Biotechnol. 13*, 290−298.

(7) Kohlstedt, M., Starck, S., Barton, N., Stolzenberger, J., Selzer, M., Mehlmann, K., Schneider, R., Pleissner, D., Rinkel, J., Dickschat, J. S., Venus, J., van Duuren, J. B. J. H., and Wittmann, C. (2018) From lignin to nylon: Cascaded chemical and biochemical conversion using metabolically engineered *Pseudomonas putida*. *Metab. Eng. 47*, 279−293.

(8) Franden, M. A., Jayakody, L. N., Li, W.-J., Wagner, N. J., Cleveland, N. S., Michener, W. E., Hauer, B., Blank, L. M., Wierckx, N., Klebensberger, J., and Beckham, G. T. (2018) Engineering *Pseudomonas putida* KT2440 for efficient ethylene glycol utilization. *Metab. Eng. 48*, 197−207.

(9) Molina, L., Rosa, R. L., Nogales, J., and Rojo, F. (2019) *Pseudomonas putida* KT2440 metabolism undergoes sequential modifications during exponential growth in a complete medium as compounds are gradually consumed. *Environ. Microbiol. 21*, 2375−2390.

(10) Dvořák, P., Kováč, J., and de Lorenzo, V. (2020) Biotransformation of d-xylose to d-xylonate coupled to medium-chain-length polyhydroxyalkanoate production in cellobiose-grown *Pseudomonas putida* EM42. *Microb. Biotechnol. 13*, 1273.

(11) Kukurugya, M. A., Mendonca, C. M., Solhtalab, M., Wilkes, R. A., Thannhauser, T. W., and Aristilde, L. (2019) Multi-omics analysis unravels a segregated metabolic flux network that tunes co-utilization of sugar and aromatic carbons in *Pseudomonas putida*. *J. Biol. Chem. 294*, 8464.

(12) Kampers, L. F. C., Volkers, R. J. M., and Martins dos Santos, V. A. P. (2019) *Pseudomonas putida* KT2440 is HV1 certified, not GRAS. *Microb. Biotechnol. 12*, 845−848.

(13) Nogales, J., Mueller, J., Gudmundsson, S., Canalejo, F. J., Duque, E., Monk, J., Feist, A. M., Ramos, J. L., Niu, W., and Palsson,

B. O. (2020) High-quality genome-scale metabolic modelling of *Pseudomonas putida* highlights its broad metabolic capabilities. *Environ. Microbiol. 22*, 255−269.
(14) Horlamus, F., Wang, Y., Steinbach, D., Vahidinasab, M., Wittgens, A., Rosenau, F., Henkel, M., and Hausmann, R. (2019) Potential of biotechnological conversion of lignocellulose hydrolyzates by *Pseudomonas putida* KT2440 as a model organism for a bio-based economy. *GCB Bioenergy 11*, 1421−1434.
(15) Jayakody, L. N., Johnson, C. W., Whitham, J. M., Giannone, R. J., Black, B. A., Cleveland, N. S., Klingeman, D. M., Michener, W. E., Olstad, J. L., Vardon, D. R., Brown, R. C., Brown, S. D., Hettich, R. L., Guss, A. M., and Beckham, G. T. (2018) Thermochemical wastewater valorization via enhanced microbial toxicity tolerance. *Energy Environ. Sci. 11*, 1625−1638.
(16) Poblete-Castro, I., Becker, J., Dohnt, K., dos Santos, V. M., and Wittmann, C. (2012) Industrial biotechnology of *Pseudomonas putida* and related species. *Appl. Microbiol. Biotechnol. 93*, 2279−2290.
(17) Sánchez-Pascuala, A., de Lorenzo, V., and Nikel, P. I. (2017) Refactoring the Embden−Meyerhof−Parnas Pathway as a Whole of Portable GlucoBricks for Implantation of Glycolytic Modules in Gram-Negative Bacteria. *ACS Synth. Biol. 6*, 793−805.
(18) Nikel, P. I., and de Lorenzo, V. (2018) *Pseudomonas putida* as a functional chassis for industrial biocatalysis: From native biochemistry to trans-metabolism. *Metab. Eng. 50*, 142−155.
(19) Martínez-García, E., and de Lorenzo, V. (2017) Molecular tools and emerging strategies for deep genetic/genomic refactoring of *Pseudomonas*. *Curr. Opin. Biotechnol. 47*, 120−132.
(20) Xu, Q., Resch, M. G., Podkaminer, K., Yang, S., Baker, J. O., Donohoe, B. S., Wilson, C., Klingeman, D. M., Olson, D. G., Decker, S. R., Giannone, R. J., Hettich, R. L., Brown, S. D., Lynd, L. R., Bayer, E. A., Himmel, M. E., and Bomble, Y. J. (2016) Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. *Sci. Adv. 2*, No. e1501254.
(21) Adams, J. J., Pal, G., Jia, Z., and Smith, S. P. (2006) Mechanism of bacterial cell-surface attachment revealed by the structure of cellulosomal type II cohesin-dockerin complex. *Proc. Natl. Acad. Sci. U. S. A. 103*, 305−310.
(22) Gunnoo, M., Cazade, P.-A., Bayer, E. A., and Thompson, D. (2018) Molecular simulations reveal that a short helical loop regulates thermal stability of type I cohesin-dockerin complexes. *Phys. Chem. Chem. Phys. 20*, 28445−28451.
(23) Demain, A. L., Newcomb, M., and Wu, J. H. D. (2005) Cellulase, clostridia, and ethanol. *Microbiol. Mol. Biol. Rev. 69*, 124−154.
(24) Artzi, L., Bayer, E. A., and Moraïs, S. (2017) Cellulosomes: bacterial nanomachines for dismantling plant polysaccharides. *Nat. Rev. Microbiol. 15*, 83−95.
(25) Fierobe, H.-P., Mingardon, F., Mechaly, A., Bélaïch, A., Rincon, M. T., Pagès, S., Lamed, R., Tardif, C., Bélaïch, J.-P., and Bayer, E. A. (2005) Action of Designer Cellulosomes on Homogeneous Versus Complex Substrates: Controlled Incorporation of Three Distinct Enzymes into a Defined Trifunctional Scaffoldin. *J. Biol. Chem. 280*, 16325−16334.
(26) Arfi, Y., Shamshoum, M., Rogachev, I., Peleg, Y., and Bayer, E. A. (2014) Integration of bacterial lytic polysaccharide monooxygenases into designer cellulosomes promotes enhanced cellulose degradation. *Proc. Natl. Acad. Sci. U. S. A. 111*, 9109−9114.
(27) You, C., and Zhang, Y.-H. P. (2014) Annexation of a high-activity enzyme in a synthetic three-enzyme complex greatly decreases the degree of substrate channeling. *ACS Synth. Biol. 3*, 380−386.
(28) Tsai, S.-L., DaSilva, N. A., and Chen, W. (2013) Functional display of complex cellulosomes on the yeast surface via adaptive assembly. *ACS Synth. Biol. 2*, 14−21.
(29) Cho, H.-Y., Yukawa, H., Inui, M., Doi, R. H., and Wong, S.-L. (2004) Production of minicellulosomes from *Clostridium cellulovorans* in *Bacillus subtilis* WB800. *Appl. Environ. Microbiol. 70*, 5704−5707.
(30) Wen, F., Sun, J., and Zhao, H. (2010) Yeast surface display of trifunctional minicellulosomes for simultaneous saccharification and fermentation of cellulose to ethanol. *Appl. Environ. Microbiol. 76*, 1251−1260.
(31) Arai, T., Matsuoka, S., Cho, H.-Y., Yukawa, H., Inui, M., Wong, S.-L., and Doi, R. H. (2007) Synthesis of *Clostridium cellulovorans* minicellulosomes by intercellular complementation. *Proc. Natl. Acad. Sci. U. S. A. 104*, 1456−1460.
(32) Stern, J., Moraïs, S., Ben-David, Y., Salama, R., Shamshoum, M., Lamed, R., Shoham, Y., Bayer, E. A., and Mizrahi, I. (2018) Assembly of Synthetic Functional Cellulosomal Structures onto the Cell Surface of *Lactobacillus plantarum*, a Potent Member of the Gut Microbiome. *Appl. Environ. Microbiol. 84*, e00282-18.
(33) Fan, L.-H., Zhang, Z.-J., Yu, X.-Y., Xue, Y.-X., and Tan, T.-W. (2012) Self-surface assembly of cellulosomes with two miniscaffoldins on *Saccharomyces cerevisiae* for cellulosic ethanol production. *Proc. Natl. Acad. Sci. U. S. A. 109*, 13260−13265.
(34) Perret, S., Casalot, L., Fierobe, H.-P., Tardif, C., Sabathe, F., Belaich, J.-P., and Belaich, A. (2004) Production of heterologous and chimeric scaffoldins by *Clostridium acetobutylicum* ATCC 824. *J. Bacteriol. 186*, 253−257.
(35) Wieczorek, A. S., and Martin, V. J. J. (2012) Effects of synthetic cohesin-containing scaffold protein architecture on binding dockerin-enzyme fusions on the surface of *Lactococcus lactis*. *Microb. Cell Fact. 11*, 160.
(36) Bokinsky, G., Peralta-Yahya, P. P., George, A., Holmes, B. M., Steen, E. J., Dietrich, J., Lee, T. S., Tullman-Ercek, D., Voigt, C. A., Simmons, B. A., and Keasling, J. D. (2011) Synthesis of three advanced biofuels from ionic liquid-pretreated switchgrass using engineered *Escherichia coli*. *Proc. Natl. Acad. Sci. U. S. A. 108*, 19949−19954.
(37) Wargacki, A. J., Leonard, E., Win, M. N., Regitsky, D. D., Santos, C. N. S., Kim, P. B., Cooper, S. R., Raisner, R. M., Herman, A., Sivitz, A. B., Lakshmanaswamy, A., Kashiyama, Y., Baker, D., and Yoshikuni, Y. (2012) An engineered microbial platform for direct biofuel production from brown macroalgae. *Science 335*, 308−313.
(38) Muñoz-Gutiérrez, I., Moss-Acosta, C., Trujillo-Martinez, B., Gosset, G., and Martinez, A. (2014) Ag43-mediated display of a thermostable β-glucosidase in *Escherichia coli* and its use for simultaneous saccharification and fermentation at high temperatures. *Microb. Cell Fact. 13*, 106.
(39) Tanaka, T., Hirata, Y., Nakano, M., Kawabata, H., and Kondo, A. (2014) Creation of cellobiose and xylooligosaccharides-coutilizing *Escherichia coli* displaying both β-glucosidase and β-xylosidase on its cell surface. *ACS Synth. Biol. 3*, 446−453.
(40) Tozakidis, I. E. P., Brossette, T., Lenz, F., Maas, R. M., and Jose, J. (2016) Proof of concept for the simplified breakdown of cellulose by combining *Pseudomonas putida* strains with surface displayed thermophilic endocellulase, exocellulase and β-glucosidase. *Microb. Cell Fact. 15*, 103.
(41) Schulte, M. F., Tozakidis, I. E. P., and Jose, J. (2017) Autotransporter-Based Surface Display of Hemicellulases on *Pseudomonas putida*: Whole-Cell Biocatalysts for the Degradation of Biomass. *ChemCatChem 9*, 3955−3964.
(42) Moraïs, S., Shterzer, N., Lamed, R., Bayer, E. A., and Mizrahi, I. (2014) A combined cell-consortium approach for lignocellulose degradation by specialized *Lactobacillus plantarum* cells. *Biotechnol. Biofuels 7*, 112.
(43) Lee, S. Y., Choi, J. H., and Xu, Z. (2003) Microbial cell-surface display. *Trends Biotechnol. 21*, 45−52.
(44) Martínez-García, E., Nikel, P. I., Aparicio, T., and de Lorenzo, V. (2014) *Pseudomonas* 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microb. Cell Fact. 13*, 159.
(45) Lieder, S., Nikel, P. I., de Lorenzo, V., and Takors, R. (2015) Genome reduction boosts heterologous gene expression in *Pseudomonas putida*. *Microb. Cell Fact. 14*, 23.
(46) Martínez-García, E., Fraile, S., Rodriguez-Espeso, D., Vecchietti, D., Bertoni, G., and de Lorenzo, V. (2020) The naked bacterium: emerging properties of a surfome-streamlined *Pseudomonas putida* strain. *ACS Synth. Biol.*, DOI: 10.1021/acssynbio.0c00272.

(47) Vazana, Y., Barak, Y., Unger, T., Peleg, Y., Shamshoum, M., Ben-Yehezkel, T., Mazor, Y., Shapiro, E., Lamed, R., and Bayer, E. A. (2013) A synthetic biology approach for evaluating the functional contribution of designer cellulosome components to deconstruction of cellulosic substrates. *Biotechnol. Biofuels 6*, 182.
(48) Ramos, J. L., Gonzalez-Carrero, M., and Timmis, K. N. (1988) Broad-host range expression vectors containing manipulated meta-cleavage pathway regulatory elements of the TOL plasmid. *FEBS Lett. 226*, 241−246.
(49) Grote, A., Hiller, K., Scheer, M., Münch, R., Nörtemann, B., Hempel, D. C., and Jahn, D. (2005) JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res. 33*, W526−W531.
(50) Meuskens, I., Michalik, M., Chauhan, N., Linke, D., and Leo, J. C. (2017) A New Strain Collection for Improved Expression of Outer Membrane Proteins. *Front. Cell. Infect. Microbiol. 7*, 464.
(51) Jong, W. S. P., Schillemans, M., ten Hagen-Jongman, C. M., Luirink, J., and van Ulsen, P. (2018) Comparing autotransporter β-domain configurations for their capacity to secrete heterologous proteins to the cell surface. *PLoS One 13*, No. e0191622.
(52) Rutherford, N., and Mourez, M. (2006) Surface display of proteins by gram-negative bacterial autotransporters. *Microb. Cell Fact. 5*, 22.
(53) Tozakidis, I. E. P., Sichwart, S., and Jose, J. (2015) Going beyond *E. coli*: autotransporter based surface display on alternative host organisms. *New Biotechnol. 32*, 644−650.
(54) Wells, T. J., Sherlock, O., Rivas, L., Mahajan, A., Beatson, S. A., Torpdahl, M., Webb, R. I., Allsopp, L. P., Gobius, K. S., Gally, D. L., and Schembri, M. A. (2008) EhaA is a novel autotransporter protein of enterohemorrhagic *Escherichia coli* O157:H7 that contributes to adhesion and biofilm formation. *Environ. Microbiol. 10*, 589−604.
(55) Leščić Ašler, I., Ivić, N., Kovačić, F., Schell, S., Knorr, J., Krauss, U., Wilhelm, S., Kojić-Prodić, B., and Jaeger, K.-E. (2010) Probing Enzyme Promiscuity of SGNH Hydrolases. *ChemBioChem 11*, 2158−2167.
(56) Klauser, T., Pohlner, J., and Meyer, T. F. (1990) Extracellular transport of cholera toxin B subunit using *Neisseria* IgA protease beta-domain: conformation-dependent outer membrane translocation. *EMBO J. 9*, 1991−1999.
(57) Freudl, R. (2018) Signal peptides for recombinant protein secretion in bacterial expression systems. *Microb. Cell Fact. 17*, 52.
(58) Veiga, E., de Lorenzo, V., and Fernández, L. A. (2003) Autotransporters as scaffolds for novel bacterial adhesins: surface properties of *Escherichia coli* cells displaying Jun/Fos dimerization domains. *J. Bacteriol. 185*, 5585−5590.
(59) Kjaergaard, K., Hasman, H., Schembri, M. A., and Klemm, P. (2002) Antigen 43-mediated autotransporter display, a versatile bacterial cell surface presentation system. *J. Bacteriol. 184*, 4197−4204.
(60) Wentzel, A., Christmann, A., Adams, T., and Kolmar, H. (2001) Display of passenger proteins on the surface of *Escherichia coli* K-12 by the enterohemorrhagic *E. coli* intimin EaeA. *J. Bacteriol. 183*, 7273−7284.
(61) Salema, V., Marín, E., Martínez-Arteaga, R., Ruano-Gallego, D., Fraile, S., Margolles, Y., Teira, X., Gutierrez, C., Bodelón, G., and Fernández, L. A. (2013) Selection of single domain antibodies from immune libraries displayed on the surface of *E. coli* cells with two β-domains of opposite topologies. *PLoS One 8*, No. e75126.
(62) Tozakidis, I. E. P., Lüken, L. M., Üffing, A., Meyers, A., and Jose, J. (2020) Improving the autotransporter-based surface display of enzymes in *Pseudomonas putida* KT2440. *Microb. Biotechnol. 13*, 176−184.
(63) Schlegel, S., Rujas, E., Ytterberg, A. J., Zubarev, R. A., Luirink, J., and de Gier, J.-W. (2013) Optimizing heterologous protein production in the periplasm of *E. coli* by regulating gene expression levels. *Microb. Cell Fact. 12*, 24.
(64) Yang, J., and Zhang, Y. (2015) I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res. 43*, W174−W181.
(65) Van Gerven, N., Sleutel, M., Deboeck, F., De Greve, H., and Hernalsteens, J.-P. (2009) Surface display of the receptor-binding domain of the F17a-G fimbrial adhesin through the autotransporter AIDA-I leads to permeability of bacterial cells. *Microbiology 155*, 468−476.
(66) Valls, M., de Lorenzo, V., Gonzàlez-Duarte, R., and Atrian, S. (2000) Engineering outer-membrane proteins in *Pseudomonas putida* for enhanced heavy-metal bioadsorption. *J. Inorg. Biochem. 79*, 219−223.
(67) Dhillon, J. K., Drew, P. D., and Porter, A. J. (1999) Bacterial surface display of an anti-pollutant antibody fragment. *Lett. Appl. Microbiol. 28*, 350−354.
(68) Ahan, R. E., Kırpat, B. M., Saltepe, B., and Şeker, U. Ö. Ş. (2019) A Self-Actuated Cellular Protein Delivery Machine. *ACS Synth. Biol. 8*, 686−696.
(69) van der Woude, M. W., and Henderson, I. R. (2008) Regulation and function of Ag43 (flu). *Annu. Rev. Microbiol. 62*, 153−169.
(70) Wagner, S., Bader, M. L., Drew, D., and de Gier, J.-W. (2006) Rationalizing membrane protein overexpression. *Trends Biotechnol. 24*, 364−371.
(71) Shaner, N. C., Steinbach, P. A., and Tsien, R. Y. (2005) A guide to choosing fluorescent proteins. *Nat. Methods 2*, 905−909.
(72) Zhang, Y.-H. P. (2011) Substrate channeling and enzyme complexes for biotechnological applications. *Biotechnol. Adv. 29*, 715−725.
(73) Laloux, G., and Jacobs-Wagner, C. (2014) How do bacteria localize proteins to the cell pole? *J. Cell Sci. 127*, 11−19.
(74) Hörnström, D., Larsson, G., van Maris, A. J. A., and Gustavsson, M. (2019) Molecular optimization of autotransporter-based tyrosinase surface display. *Biochim. Biophys. Acta, Biomembr. 1861*, 486−494.
(75) Roche, C. M., Dibble, C. J., Knutsen, J. S., Stickel, J. J., and Liberatore, M. W. (2009) Particle concentration and yield stress of biomass slurries during enzymatic hydrolysis at high-solids loadings. *Biotechnol. Bioeng. 104*, 290−300.
(76) Stern, J., Moraïs, S., Lamed, R., and Bayer, E. A. (2016) Adaptor Scaffoldins: An Original Strategy for Extended Designer Cellulosomes, Inspired from Nature. *mBio 7*, No. e00083.
(77) Smit, J., Kamio, Y., and Nikaido, H. (1975) Outer membrane of *Salmonella typhimurium*: chemical analysis and freeze-fracture studies with lipopolysaccharide mutants. *J. Bacteriol. 124*, 942−958.
(78) Jose, J., Bernhardt, R., and Hannemann, F. (2002) Cellular surface display of dimeric Adx and whole cell P450-mediated steroid synthesis on *E. coli*. *J. Biotechnol. 95*, 257−268.
(79) Kahn, A., Moraïs, S., Galanopoulou, A. P., Chung, D., Sarai, N. S., Hengge, N., Hatzinikolaou, D. G., Himmel, M. E., Bomble, Y. J., and Bayer, E. A. (2019) Creation of a functional hyperthermostable designer cellulosome. *Biotechnol. Biofuels 12*, 44.
(80) Hammel, M., Fierobe, H.-P., Czjzek, M., Kurkal, V., Smith, J. C., Bayer, E. A., Finet, S., and Receveur-Bréchot, V. (2005) Structural basis of cellulosome efficiency explored by small angle X-ray scattering. *J. Biol. Chem. 280*, 38562−38568.
(81) Sørensen, A., Lübeck, M., Lübeck, P. S., and Ahring, B. K. (2013) Fungal Beta-Glucosidases: A Bottleneck in Industrial Use of Lignocellulosic Materials. *Biomolecules 3*, 612−631.
(82) Gefen, G., Anbar, M., Morag, E., Lamed, R., and Bayer, E. A. (2012) Enhanced cellulose degradation by targeted integration of a cohesin-fused β-glucosidase into the *Clostridium thermocellum* cellulosome. *Proc. Natl. Acad. Sci. U. S. A. 109*, 10298−10303.
(83) Becker, S., Theile, S., Heppeler, N., Michalczyk, A., Wentzel, A., Wilhelm, S., Jaeger, K.-E., and Kolmar, H. (2005) A generic system for the *Escherichia coli* cell-surface display of lipolytic enzymes. *FEBS Lett. 579*, 1177−1182.
(84) Honjo, H., Iwasaki, K., Soma, Y., Tsuruno, K., Hamada, H., and Hanai, T. (2019) Synthetic microbial consortium with specific roles designated by genetic circuits for cooperative chemical production. *Metab. Eng. 55*, 268−275.

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

(85) Tucker, A. T., Leonard, S. P., DuBois, C. D., Knauf, G. A., Cunningham, A. L., Wilke, C. O., Trent, M. S., and Davies, B. W. (2018) Discovery of Next-Generation Antimicrobials through Bacterial Self-Screening of Surface-Displayed Peptide Libraries. *Cell 172*, 618−628.
(86) Zhang, Z., Liu, J., Fan, J., Wang, Z., and Li, L. (2018) Detection of catechol using an electrochemical biosensor based on engineered *Escherichia coli* cells that surface-display laccase. *Anal. Chim. Acta 1009*, 65−72.
(87) Dvořák, P., Nikel, P. I., Damborský, J., and de Lorenzo, V. (2017) Bioremediation 3.0: Engineering pollutant-removing bacteria in the times of systemic biology. *Biotechnol. Adv. 35*, 845−866.
(88) Abril, M. A., Michan, C., Timmis, K. N., and Ramos, J. L. (1989) Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *J. Bacteriol. 171*, 6782−6790.
(89) Green, M. R. (2012) *Molecular Cloning: A Laboratory Manual*, Fourth ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
(90) Horton, R. M., Cai, Z., Ho, S. N., and Pease, L. R. (1990) Gene splicing by overlap extension: Tailor-made genes using the polymerase chain reaction. *BioTechniques 8* (5), 528−535.
(91) Aparicio, T., de Lorenzo, V., and Martínez-García, E. (2017) Broadening the SEVA Plasmid Repertoire to Facilitate Genomic Editing of Gram-Negative Bacteria. In *Hydrocarbon and Lipid Microbiology Protocols: Genetic, Genomic and System Analyses of Pure Cultures* (McGenity, T. J., Timmis, K. N., and Nogales, B., Eds.), pp 9−27, Springer, Berlin, Heidelberg.

https://dx.doi.org/10.1021/acssynbio.0c00276
ACS Synth. Biol. XXXX, XXX, XXX−XXX

Check for updates

microbial biotechnology

Open Access

Brief Report

An automated DIY framework for experimental evolution of *Pseudomonas putida*

David R. Espeso,[1] Pavel Dvořák,[2] Tomás Aparicio[1] and Víctor deLorenzo[1,*]

[1]*Systems Biology Program, Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco, Madrid 28049, Spain.*

[2]*Department of Experimental Biology, Faculty of Science, Masaryk University, Brno 62500, Czech Republic.*

Summary

Adaptive laboratory evolution (ALE) is a general and effective strategy for optimizing the design of engineered genetic circuits and upgrading metabolic phenotypes. However, the specific characteristics of each microorganism typically ask for exclusive conditions that need to be adjusted to the biological chassis at stake. In this work, we have adopted a do-it-yourself (DIY) approach to implement a flexible and automated framework for performing ALE experiments with the environmental bacterium and metabolic engineering platform *Pseudomonas putida*. The setup includes a dual-chamber semi-continuous log-phase bioreactor design combined with an anti-biofilm layout to manage specific traits of this bacterium in long-term cultivation experiments. As a way of validation, the prototype was instrumental for selecting fast-growing variants of a *P. putida* strain engineered to metabolize D-xylose as sole carbon and energy source after running an automated 42 days protocol of iterative regrowth. Several genomic changes were identified in the evolved population that pinpointed the role of RNA polymerase in controlling overall physiological conditions during metabolism of the new carbon source.

Received 18 July, 2020; revised 12 September, 2020; accepted 22 September, 2020.

*For correspondence. E-mail vdlorenzo@cnb.csic.es; Tel. +34 91 585 4536; Fax +34 91 585 4506.

Microbial Biotechnology (2020) **0**(0), 1–7
doi:10.1111/1751-7915.13678

Funding information

This work was funded by the SETH Project of the Spanish Ministry of Science RTI 2018-095584-B-C42, the MADONNA (H2020-FET-OPEN-RIA-2017-1-766975), BioRoboost (H2020-NMBP-BIO-CSA-2018), SYNBIO4FLAV (H2020-NMBP/0500) and MIX-UP (H2020-Grant 870294) Contracts of the European Union and the S2017/BMD-3691 InGEMICS-CM Project of the Comunidad de Madrid (European Structural and Investment Funds) as well as by the Czech Science Foundation (19-06511Y).

Introduction

The development of do-it-yourself (DIY) technical solutions (Moe-Behrens *et al*., 2013; de Lorenzo and Schmidt, 2017) for performing adaptive laboratory evolution (ALE) experiments (Portnoy *et al*., 2011; Dragosits and Mattanovich, 2013; LaCroix *et al*., 2017) is expanding the capabilities of researchers to integrate this attractive technique in their regular laboratory workflows. Some examples include the development of automatic microbial cultivation platforms operating mini-chemostats (Amanullah *et al*., 2010; Bergenholm *et al*., 2019), turbidostats (Marlière *et al*., 2011; Wong *et al*., 2018; McGeachy *et al*., 2019) or segregostats (Sassi *et al*., 2019). Yet, in any circumstance ALE experiments have to be designed taking into account the biological constrains of the evolving microbe and the target to achieve. One of such microorganisms of interest is the soil bacterium *Pseudomonas putida* (in particular strain KT2440) which, because of its distinct management of oxidative stress, has emerged as a prime host of engineered redox reactions (Nikel *et al*., 2014, 2016; Nikel and de Lorenzo, 2018). On this background, we set out to design and implement a DIY framework specifically developed for applying flexible ALE protocols to this bacterium for the sake of increasing its performance as whole-cell catalyst.

The construction details and every step of the implementation of the evolutionary device are fully disclosed in the Supplementary Information (Materials, Equipment and other procedural features: see Figs S1–S25 and Tables S1–S3). The reader is encouraged to access such accompanying particulars for a more complete comprehension of the technical solution hereby presented. The experimental setup was inspired in the turbidostat scheme proposed by Marlière *et al*. (2011) but was redesigned considering a number of constrains

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd
This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

linked to the intrinsic biological features of the KT2440 strain of *P. putida*. One first consideration is that the specimen of interest belongs to a bacterial species that naturally sticks to surfaces and builds considerable amounts of biofilms (Auerbach *et al*., 2000; Espinosa-Urgel *et al*., 2000; Tolker-Nielsen *et al*., 2000; Espeso *et al*., 2018). Biofilm formation is operationally problematic, because it clogs culture conduits and selects for surface super-sticker variants. A second constraint is that *P. putida* KT2440 is strictly aerobic (Nikel and de Lorenzo, 2013; Kampers *et al*., 2019), and proper aeration is required to ensure culture viability and vitality during the long-term experiment. Furthermore, the evolutionary platform must ensure isolation of the manipulated culture to avoid contamination by microorganisms that may displace the template strain. Finally, growth media quality should be secured at all times for maximizing cell division and foster DNA replication – thereby increasing chances of mutations.

With these criteria in mind, an experimental setup was designed and assembled to execute a basic protocol for sustaining bacterial growth for long periods of time. Figure 1A shows the thereby implemented workflow process as a block diagram. The sketch illustrates a recurrent cycle in which a semi-continuous incubation of a culture is executed by a period of time defined by the user. The protocol included a control loop in the reactor incubation step (blue box) where the workflow was stalled in a periodic subroutine of incubation steps followed by optical density measurements at regular intervals ($t_{sampling}$). Such a recurrent sequence ended when an upper threshold OD_{600} value was reached, allowing the workflow in this manner to keep advancing. Figure 1B shows the fluidic layout implemented for succeeding with this protocol. The basic setup includes (i) a bioreactor coupled to a photodetector to obtain OD_{600} readings, (ii) an auxiliary chamber to allow biofilm cleaning with NaOH and H_2O, (iii) a rack of pumps to deliver the different chemicals and (iv) a group of valves to set the logic of the liquid transport through the circuit. The tubing is connected in a circular fashion with two independent waste outputs and venting connections to ensure an uninterrupted cell culture with sufficient aeration. The design was complemented with electronic and control layers, consistently designed to make possible the synchronized actuation of all these devices (Figs S1–S15). Additionally, this basic arrangement was complemented with the manufacturing of 3D-printed supports to spatially arrange pumps and valves (Figs S16–S18) and the assembly of an online optical density chamber to gather OD_{600} lectures (Figs. S19–S25).

To test the efficacy of the thereby constructed DIY platform, we used a derivative of *P. putida* KT2440 that had been engineered to grow on D-xylose, a pentose abundant in hydrolysates of lignocellulosic materials (Chen *et al*., 2017). The construct at stake (named *P. putida* mk-1, Table S3) bears a large number of genomic modifications for increasing stability, raising the intracellular levels of ATP and NAD(P)H (Martínez-García *et al*., 2014a,b) and avoiding misrouting of intermediates during D-xylose metabolism. Specifically, *P. putida* mk-1 lacks flagella and other energy-draining cellular devices and has a deletion of *gcd* (thereby lacking glucose dehydrogenase). In addition, the strain bears a chromosomal implant of a synthetic *xylABE* operon encoding XylA (xylose isomerase), XylB (xylulokinase) and XylE (xylose-proton symporter) from *Escherichia coli* (Dvořák and de Lorenzo, 2018). To this end, the DNA segment bearing *xylABE* was assembled in a mini-Tn*5* transposon vector (Martínez-García *et al*., 2014a,b) as described in the Supplementary information. During the construction of the test strain, the mobile element mini-Tn*5* Sm::[P_{EM7} → *xylABE*] was randomly inserted throughout the genome of strain *P. putida* EM42 Δ*gcd* (Table S3). The organization of the mini-Tn*5* transposon was such that the *xylABE* operon could be expressed from the synthetic P_{EM7} promoter engineered in the mobile element as well as from readthrough transcription of nearby promoters close to the site of insertion.

Selection of the best grower clone on D-xylose as sole carbon source yielded the aforementioned strain *P. putida* mk-1 with the business DNA segment inserted in the midst of the locus PP_2260 (a putative glycerol-phosphate ABC transporter ATP-binding protein; Fig. 2C). Whether there was a benefit in the interruption of that ORF is unknown, but insertion of the DNA segment with [P_{EM7} → *xylABE*] in the chromosome secured the stable inheritance of the knocked-in trait during the course of the bioreactor experiment (Fig. 2A).

Next, strain *P. putida* mk-1 was inoculated in an intermediate reactor chamber with an operative volume of 20 ml and containing around 10^9 cultivated cells with an OD_{600} bounded within the range [0.1–0.5]. During a 45 days period, cells were recurrently incubated and diluted (Fig. 2A) using M9 minimal medium supplemented with 0.2% (w/v) D-xylose and 60 μg ml^{-1} streptomycin. Under these simple conditions, the setup selects for faster growers which – in case of appearance – should bear mutations that increase overall physiological fitness and/or improve nesting of the implanted metabolic segment in the background biochemical network of *P. putida*. The progress of the experiment is shown in Figure 2A. At the end of the corresponding period of time, an increase in growth rate of the population present in the culture became clearly noticeable, same as the fact that beneficial changes occurred probably mainly during the initial phase of the experiment (Fig. 2A, Fig. S26). To examine the basis of such a change,

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd

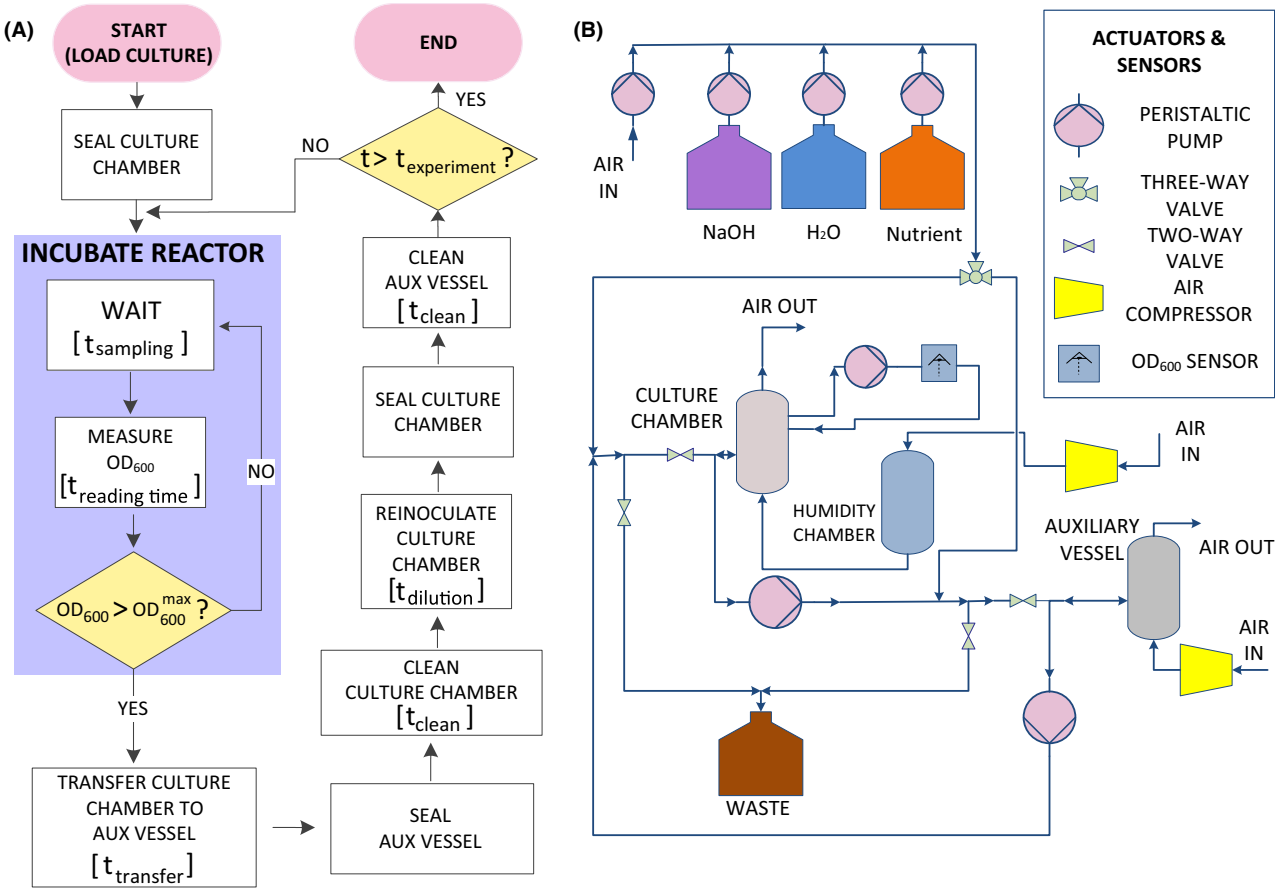


Fig. 1. Schematic representation of the DIY device for experimental evolution of *Pseudomonas putida*.
A. Block diagram showing the process workflow to implement conceptualized from the ALE protocol taken to analyse as example. The workflow shows the different high-level actions to perform, their relative order of execution, timing, recurrence loops (i.e. blue square) and decision taking points (on yellow).
B. Conceptual scheme showing the actuators, sensors and vessels used to design the fluidic layer of the ALE experimental device. A set of peristaltic pumps, compressors and valves is in charge of transporting different chemicals to clean/wash the vessels and feed a bacterial culture constantly monitored by an optical sensor reader.

samples were collected from the reactor and further inspected. First, the evolved sample (hereafter called mk-2) was verified as an authentic descendant of the original strain. For this, the mk-2 sample was plated on LB and M9 + 0.2% (w/v) citrate agar dishes and cells were streaked out to discard any contamination. Strain clonal identity was confirmed through PCR of the genome with primers 5′CTTCAGCTCTTCGCTGTACA3′ and 5′GCGTGCGCTACAACCTTAC3′ that amplify the region surrounding the deletion of the glucose dehydrogenase gene (PP_1444) present in the template strain and which acted as a diagnostic signature. Second, the growth rate of the evolved culture on D-xylose as the only C source was re-assessed in respect to the precursor strain performing independent growth curve assays in Erlenmeyer flasks. Regression slopes comparing the two (Fig. 2B) indicated that the evolved specimen grew a 60% faster than template strain. Finally, the genomes of the original *P. putida* mk-1 strain and the evolved counterpart mk-2 were sequenced to find mutations that could account for the observed shift in the growth phenotype.

While no modifications became apparent in the bacteria of reference, the faster-growing derivative bore 3 conspicuous changes in its chromosome. The first modification of the evolved genome was found in the *rpoC* gene of *P. putida*, which encodes the β′ subunit of RNA polymerase. The *rpoC* of gene of mk-2 had a point mutation C → T in codon 51 (cCt → cTt) causing a quite drastic change $Pro_{51}Leu$. The emergence of this modification acted in fact as a descriptor of the efficacy of the evolutionary experiment. This is because as a large number of *rpoC* mutations have been reported in the course of laboratory evolution studies aiming to increase *E. coli* growth rate (Cheng *et al.*, 2014; Wytock *et al.*, 2018; Kavvas *et al.*, 2020). Therefore, the

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd

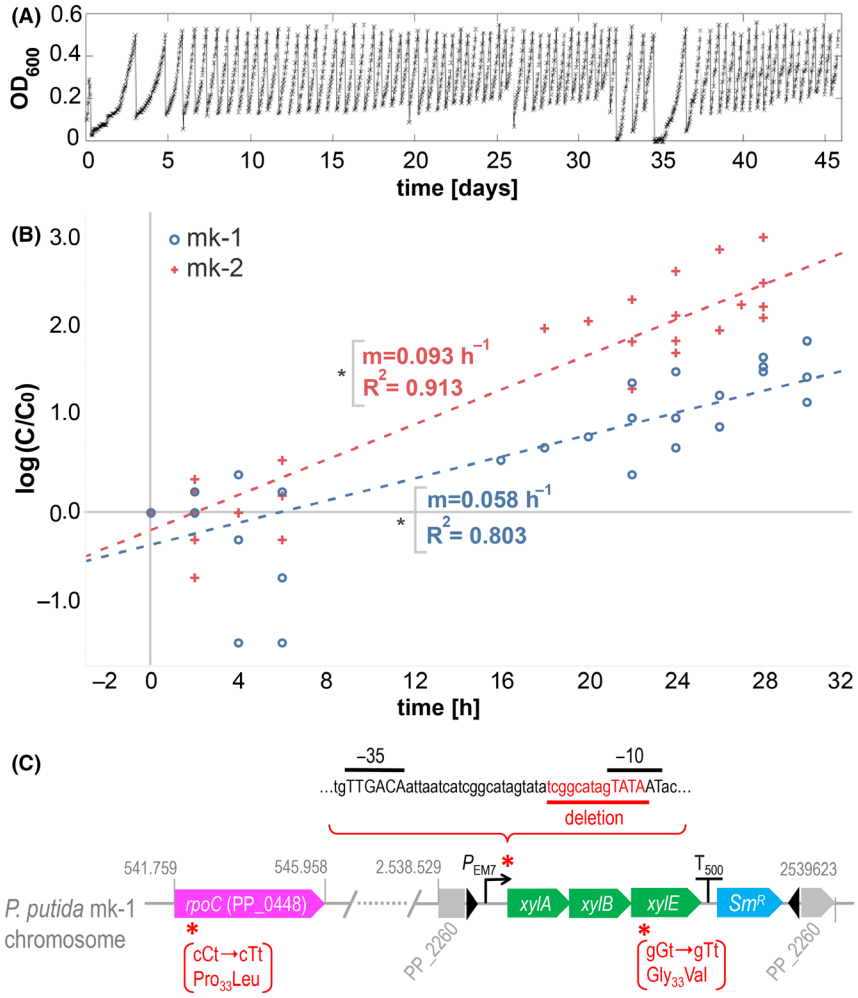


Fig. 2. Evolution of an engineered Xyl$^+$ strain of *P. putida* along a 42 days protocol of iterative regrowth.
A. Optical density evolution during the 45 days period of the ALE experiment. The sawtooth pattern of the graph corresponds to the culture dilution dynamic (semi-continuous culture) used by the device, programmed to hold the optical density within an exponential growth regime with optical densities within the range [0.1 – 0.5].
B. Independent ALE validation experiments. Growth curve assays using shake flasks were performed to estimate the growth rates of template (mk-1, blue) and evolved (mk-2, red). For the tested conditions, mk-2 sample exhibited a 60% increment respect to template strain. The plot shows a fitting of three independent biological replicates. Asterisks indicate that both regressions passed *t*-test at 5% confidence ($P < 0.05$).
C. Mutations detected by whole genome sequencing of the *P. putida* mk-2 sample. A scheme of the *P. putida* mk-1 chromosome is depicted showing relevant genes and genomic changes detected after the evolution procedure. Genomic coordinates of PP_2260 (locus of mini-Tn5 insertion) and *rpoC* refer to *P. putida* EM42 ancestral strain. Inverted repeats ME-I and ME-O, defining the edges of mini-Tn5, are also shown by black arrowheads with *xylABE* cluster in between. Locations of detected mutations are denoted by red asterisks. P_{EM7} sequence features −35 and −10 boxes in high case and underlined text, while deletion found in mk-2 genome appears underlined in red colour. Single nucleotide changes found in *rpoC* and *xylE* appear in brackets: wild type and mutated codon are depicted with mutated site in high case. The amino acid change and position in the polypeptide are also shown below.

$Pro_{51}Leu$ change plausibly reflects a similar adaptation in *P. putida*. The other two mutations were identified in the implanted xylose cluster. One of them involved a single base change G → T in the codon 33 of the *xylE* gene (gGt → gTt) which translates into a $Gly_{33}Val$ amino acid change. The *xylE* product is a xylose-proton symporter of *E. coli* composed by several transmembrane domains connected by periplasmic/cytoplasmic amino acid stretches (Davis and Henderson, 1987), and Gly_{33} is located in a periplasmic side close to the H^+ coupling site Asp_{27} (Madej *et al.*, 2014). While a number of loss-of-function mutations have been reported for *xylE* (Sun *et al.*, 2012), to the best of our knowledge no changes are known to enhance xylose transport. The $Gly_{33}Val$ observed in mk-2 could be one of them, an issue that deserves further studies. Finally, a 13-bp deletion removing part of the P_{EM7} −10 box was observed also in the faster-growing culture. As mentioned above, P_{EM7} is a strong synthetic promoter engineered for driving the expression of *xylABE* cluster in *P. putida* mk-1 strain.

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd

The loss of part of the −10 box (Collado-Vides *et al*., 1991; Meysman *et al*., 2014) is expected to reduce the promoter activity in the mk-2 cells. Selective pressure to curb P_{EM7} strength might be related to the fact that overproduction of the XylE transporter is toxic to *P. putida* cells (unpublished data). The coexistence in mk-2 of mutations anticipated to both decrease *xylE* transcription and improve XylE efficiency could reflect a solution to the conflict between the negative effects of overproducing a membrane protein and the need to secure a sufficient inflow of the carbon source for a faster growth. Note that – as discussed above – even complete elimination of the P_{EM7} promoter of the genomic implant [P_{EM7} → *xylABE*] could still deliver expression of the operon owing to readthrough transcription from promoter (s) outside the mini-Tn*5* insertion (Fig. 2C).

Whether the effects of these three mutations found in the evolved, faster-growing sample are additive, synergistic or altogether independent is beyond the scope of this technical note and will be the subject of subsequent studies. Correspondingly, further rational engineering cuts and ALE with the constructed system are being considered to remove additional metabolic bottleneck(s) that could have prevented achieving even faster growth of *P. putida* recombinant on the non-native substrate during the evolution experiment (Elmore *et al*., 2020). Yet, the data presented above accredits the power of the simple and affordable DIY setup described here to generate phenotypes of considerable biotechnological interest in the synthetic biology chassis and metabolic engineering platform *Pseudomonas putida*. Besides the enhancement of catabolic traits, the authors foresee the use of the bespoken device also for the evolution of biosynthetic pathways in *P. putida* and other bacterial cell factories. As additional modules, e.g., for absorbance or fluorescence quantification, can be easily integrated into the presented setup, we entertain the use of this framework also for the accelerated evolution of industrially relevant strains equipped with genetically encoded product-responsive biosensors (Mahr *et al*., 2015).

Acknowledgements

Authors are indebted to Alfonso Jaramillo, Rui Rodrigues and Philippe Marliere for fruitful discussions and advice on automation of protocols, hardware support and electronic device design, assembly and maintenance. Lee Cronin and Soichiro Tsuda are gratefully acknowledged for their valuable help with 3D printing technology and Arduino programming. This work was funded by the SETH Project of the Spanish Ministry of Science RTI 2018-095584-B-C42, the MADONNA (H2020-FET-OPEN-RIA-2017-1-766975), BioRoboost (H2020-NMBP-BIO-CSA-2018), SYNBIO4FLAV (H2020-NMBP/0500) and MIX-UP (H2020-Grant 870294) Contracts of the European Union and the S2017/BMD-3691 InGEMICS-CM Project of the Comunidad de Madrid (European Structural and Investment Funds) as well as by the Czech Science Foundation (19-06511Y).

Conflict of interest

The authors declare no competing financial interest.

References

Amanullah, A., Otero, J.M., Mikola, M., Hsu, A., Zhang, J., Aunins, J., *et al*. (2010) Novel micro-bioreactor high throughput technology for cell culture process development: Reproducibility and scalability assessment of fed-batch CHO cultures. *Biotechnol Bioeng* **106:** 57–67.

Auerbach, I.D., Sorensen, C., Hansma, H.G., and Holden, P.A. (2000) Physical morphology and surface properties of unsaturated *Pseudomonas putida* biofilms. *J Bacteriol* **182:** 3809–3815.

Bergenholm, D., Liu, G., Hansson, D., and Nielsen, J. (2019) Construction of mini-chemostats for high-throughput strain characterization. *Biotechnol Bioeng* **116:** 1029–1038.

Chen, H., Liu, J., Chang, X., Chen, D., Xue, Y., Liu, P., *et al*. (2017) A review on the pretreatment of lignocellulose for high-value chemicals. *Fuel Process Technol* **160:** 196–206.

Cheng, K.K., Lee, B.S., Masuda, T., Ito, T., Ikeda, K., Hirayama, A., *et al*. (2014) Global metabolic network reorganization by adaptive mutations allows fast growth of *Escherichia coli* on glycerol. *Nat Commun* **5:** 3233.

Collado-Vides, J., Magasanik, B., and Gralla, J.D. (1991) Control site location and transcriptional regulation in *Escherichia coli*. *Microbiol Mol Biol Rev* **55:** 371–394.

Davis, E.O., and Henderson, P.J. (1987) The cloning and DNA sequence of the gene xylE for xylose-proton symport in *Escherichia coli* K12. *J Biol Chem* **262:** 13928–13932.

Dragosits, M., and Mattanovich, D. (2013) Adaptive laboratory evolution–principles and applications for biotechnology. *Microb Cell Fact* **12:** 64.

Dvořák, P., and de Lorenzo, V. (2018) Refactoring the upper sugar metabolism of *Pseudomonas putida* for co-utilization of cellobiose, xylose, and glucose. *Metab Eng* **48:** 94–108.

Elmore, J.R., Dexter, G.N., Salvachúa, D., O'Brien, M., Klingeman, D.M., Gorday, K., *et al*. (2020) Engineered *Pseudomonas putida* simultaneously catabolizes five major components of corn stover lignocellulose: glucose, xylose, arabinose, p-coumaric acid and acetic acid. *Metab Eng* **62:** 62–71.

Espeso, D.R., Martínez-García, E., Carpio, A., and de Lorenzo, V. (2018) Dynamics of *Pseudomonas putida* biofilms in an upscale experimental framework. *J Ind Microbiol Biotechnol* **45:** 899–911.

Espinosa-Urgel, M., Salido, A., and Ramos, J.L. (2000) Genetic analysis of functions involved in adhesion of

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd

Pseudomonas putida to seeds. *J Bacteriol* **182:** 2363–2369.

Kampers, L.F.C., van Heck, R.G.A., Donati, S., Saccenti, E., Volkers, R.J.M., Schaap, P.J., *et al*. (2019) In silico-guided engineering of *Pseudomonas putida* towards growth under micro-oxic conditions. *Microb Cell Fact* **18:** 179.

Kavvas, E.S., Antoniewicz, M., Long, C., Ding, Y., Monk, J.M., Palsson, B.O., and Feist, A.M. (2020) Laboratory evolution of multiple E. coli strains reveals unifying principles of adaptation but diversity in driving genotypes. *bioRxiv* 2020.2005.2019.104992.

LaCroix, R.A., Palsson, B.O., and Feist, A.M. (2017) A model for designing adaptive laboratory evolution experiments. *Appl Environ Microbiol* **83:** e03115-03116.

de Lorenzo, V., and Schmidt, M. (2017) The do-it-yourself movement as a source of innovation in biotechnology - and much more. *Microb Biotechnol* **10:** 517–519.

Madej, M.G., Sun, L., Yan, N., and Kaback, H.R. (2014) Functional architecture of MFS D-glucose transporters. *Proc Natl Acad Sci USA* **111:** E719–E727.

Mahr, R., Gätgens, C., Gätgens, J., Polen, T., Kalinowski, J., and Frunzke, J. (2015) Biosensor-driven adaptive laboratory evolution of l-valine production in *Corynebacterium glutamicum*. *Metab Eng* **32:** 184–194.

Marlière, P., Patrouix, J., Döring, V., Herdewijn, P., Tricot, S., Cruveiller, S., *et al*. (2011) Chemical evolution of a bacterium's genome. *Angew Chem Int Ed Engl* **50:** 7109–7114.

Martínez-García, E., Aparicio, T., de Lorenzo, V., and Nikel, P.I. (2014a) New transposon tools tailored for metabolic engineering of Gram-negative microbial cell factories. *Front Bioeng Biotechnol* **2:** 46.

Martínez-García, E., Nikel, P.I., Aparicio, T., and de Lorenzo, V. (2014b) Pseudomonas 2.0: genetic upgrading of *P. putida* KT2440 as an enhanced host for heterologous gene expression. *Microb Cell Fact* **13:** 159.

McGeachy, A.M., Meacham, Z.A., and Ingolia, N.T. (2019) An accessible continuous-culture turbidostat for pooled analysis of complex libraries. *ACS Synth Biol* **8:** 844–856.

Meysman, P., Collado-Vides, J., Morett, E., Viola, R., Engelen, K., and Laukens, K. (2014) Structural properties of prokaryotic promoter regions correlate with functional features. *PLoS One* **9:** e88717.

Moe-Behrens, G.H., Davis, R., and Haynes, K.A. (2013) Preparing synthetic biology for the world. *Front Microbiol* **4:** 5.

Nikel, P.I., Chavarría, M., Danchin, A., and de Lorenzo, V. (2016) From dirt to industrial applications: *Pseudomonas putida* as a Synthetic Biology chassis for hosting harsh biochemical reactions. *Curr Opin Chem Biol* **34:** 20–29.

Nikel, P.I., and de Lorenzo, V. (2013) Engineering an anaerobic metabolic regime in *Pseudomonas putida* KT2440 for the anoxic biodegradation of 1, 3-dichloroprop-1-ene. *Metab Eng* **15:** 98–112.

Nikel, P.I., and de Lorenzo, V. (2018) *Pseudomonas putida* as a functional chassis for industrial biocatalysis: from native biochemistry to trans-metabolism. *Metab Eng* **50:** 142–155.

Nikel, P.I., Martínez-García, E., and de Lorenzo, V. (2014) Biotechnological domestication of pseudomonads using synthetic biology. *Nat Rev Microbiol* **12:** 368–379.

Portnoy, V.A., Bezdan, D., and Zengler, K. (2011) Adaptive laboratory evolution–harnessing the power of biology for metabolic engineering. *Curr Opin Biotechnol* **22:** 590–594.

Sassi, H., Nguyen, T.M., Telek, S., Gosset, G., Grünberger, A., and Delvigne, F. (2019) Segregostat: a novel concept to control phenotypic diversification dynamics on the example of Gram-negative bacteria. *Microb Biotechnol* **12:** 1064–1075.

Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., and Yan, N. (2012) Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490:** 361–366.

Tolker-Nielsen, T., Brinch, U.C., Ragas, P.C., Andersen, J.B., Jacobsen, C.S., and Molin, S. (2000) Development and dynamics of *Pseudomonas*sp. biofilms. *J Bacteriol* **182:** 6482–6489.

Wong, B.G., Mancuso, C.P., Kiriakov, S., Bashor, C.J., and Khalil, A.S. (2018) Precise, automated control of conditions for high-throughput growth of yeast and bacteria with eVOLVER. *Nat Biotechnol* **36:** 614–623.

Wytock, T.P., Fiebig, A., Willett, J.W., Herrou, J., Fergin, A., Motter, A.E., and Crosson, S. (2018) Experimental evolution of diverse *Escherichia coli* metabolic mutants identifies genetic loci for convergent adaptation of growth rate. *PLoS Genet*. **14:** e1007284.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. Wiring used in the electric circuit. Table shows the type of connection, which parts of the circuit connects, the type of used wire and the operation voltage.

Table S2. Printing parameters used to manufacture the 3D printed supports and equipment.

Table S3. Strains and plasmids used in this study.

Fig. S1. Device assembly composition scheme. (A) The device is divided in three layers (fluidic layer, electronic layer and control layer) that are also physically separated by design. Every layer is in charge (B) of managing different resources, performing different tasks and receive / deliver different feedback from the other layers.

Fig S2. Schematic illustration of how the different parts of the device are connected by the electronic layer. Fluidic equipment is wired to a switch board, connecting all actuators to a MOSFET Arduino PCB that converts the 5 V control signals delivered by the Arduino Card into effective modulation of 12 V power supply for actuators.

Fig. S3. Fluidic hardware control logic acting in each stage of the implemented protocol, and expected transport of liquids within the device.

Fig. S4. Fluidic layer assembly blueprint. Every connector, tubing, actuator and sensor participating in the implemented protocol is detailed to make more intuitive how the device was constructed.

Fig. S5. ULN2803A MOSFET transistor array scheme and operation regimes. This model contains 8 NPN logic level MOSFET (up) which essentially works as electric gates that regulate the flow of a current passing through them (in

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd

Source – Drain direction) by modulating the incoming voltage in the Gate pin, leading to off/on or varying loading regimes (down).

Fig. S6. MOSFET Arduino PCB blueprint. Lateral pin holes are placed to match with Arduino pins, allowing an easy fitting. Board size is 100 × 75 mm. Hole Pitch is 2.54 mm. Red and Blue lines depict tracks printed in the front and bottom of the PCB, respectively.

Fig. S7. Switch PCB blueprint. Small pin boxes correspond to actuator KK254 switches, which are connected to 16 –pin compacted pin headers host male IDC connectors in charge of linking the board with MOSFET input pins. Real size is 150 × 100 mm. Hole pitch is 2.54 mm. Red and blue lines depict tracks printed in the front and bottom of the PCB, respectively.

Fig. S8. Fuse PCB blueprint. Small pin boxes correspond to actuator KK254 switches. Resettable fuse is soldered in the rounded pinholes. PCB size is 35.5 × 12.7 mm. Hole pitch is 2.54 mm. Blue lines depict tracks printed in the bottom of the PCB.

Fig. S9. FT232H PCB blueprint. Small pin boxes correspond to actuator KK254 switches. Rounded pinholes are used to solder straight PCB sockets that allow a removable connection with FT232H. D0 and D1 labels define the orientation of the chip. PCB size is 60 × 40 mm. Hole pitch is 2.54 mm. Blue lines depict tracks printed in the bottom of the PCB.

Fig. S10. Power board PCB blueprint. Small pin boxes correspond to actuator KK254 switches. Large square labeled as "12V DC IN" is designed to assemble a 4-way MOLEX mini-fit switch used to host 12V ATX terminals (typically found in PC power supplies). Additional through holes are available to add a 47 μF and 100 nF condenser to reduce noise coming from the power supply. PCB size is 50 x 20 mm. Hole pitch for KK254 switches is 2.54 mm, and 4.2 mm for MOLEX mini-fit. Blue lines depict tracks printed in the bottom of the PCB.

Fig. S11. Relay board PCB blueprint. Small pin boxes correspond to actuator KK254 switches. Large square labeled is designed to assemble a 12VDC SPDT Relay used to activate or deactivate the air compressor unit, which works at 120/220 VAC. A diode (see arrow) should be used to damp current oscillations when switching ON /OFF the relay. PCB size is 60 × 20 mm. Hole pitch for KK254 switches is 2.54 mm. Blue lines depict tracks printed in the bottom of the PCB.

Fig. S12. Images of the assembled device. (A) General overview; (B) Electronic and Control layer; (C) Fluidic layer.

Fig. S13 Schematic representation of wiring and connector diagram. Each component of the device is connected with a labeled wire. Each wire has a wire type and two terminals that use different connectors.

Fig. S14. Example of coding structuring model used to program the Arduino card. Complex high-level instructions are divided in simpler actions using a cascading decomposition. The codification is then implemented by aggregating these point actions into functions, tested before being used, and then taken as blocks to create more complex routines. This methodology allows minimizing the debugging step and makes easier the code comprehension.

Fig. S15. Schematic representation of code structure used to program the Arduino. First an assignation of pins to physical actuators is performed. Next variables used to control the logic of the flow (times to wait, values of OD600, etc.) are included in the code. Then functions in charge of executing operations of increasing complexity to the fluid are defined one by one. Finally, the execution loop calls the highest-level functions to recurrently apply the required physical actions planned in the cyclic operation to be performed.

Fig. S16. Drawing showing for WPX1 pump support design.

Fig. S17. Drawing showing WPM pump support design.

Fig. S18. Drawing showing valve scaffold design.

Fig. S19. Design drawing of the wet chamber.

Fig. S20. Images of the optical reader set. Individual parts (down) are screwed with 4 × 25 mm M3 bolts and nuts to assemble a compact closed black box with the wet chamber enclosed in it (up). Only two side holes used to emit 610 nm light and to perform the light measurements.

Fig. S21. Design drawing of the chamber scaffold, part I.

Fig. S22. Design drawing of the chamber scaffold, part II.

Fig. S23. Design drawing of the chamber scaffold, part III.

Fig. S24. Calibration slope relating light reading vs optical density in a *P. putida* KT2440 culture grown in M9+0.2% (w/v) glucose using the propose custom-made optical reader.

Fig. S25. OD_{600} readings obtained during the overnight incubation of *P. putida* KT2440 culture grown in M9+0.2% (w/v) glucose using the device.

Fig. S26. Growth rate evolution during the experiment development. The sharp peak located at day 30 was a result of a numerical drift caused by the detection of arbitrarily larger values of optical density lectures due to a hardware failure. Once realigned, the optical sensor started working properly. Dashed lines mark the average value ± standard deviation.

Fig. S27. Comparison of growth of *P. putida* EM42 Δgcd (control) and mk-1 strain in rich growth medium and in minimal medium with carbon source. (A) Rich lysogeny broth medium; (B) M9 minimal medium with 5 g l^{-1} glucose. (C) M9 minimal medium with 5 g l^{-1} citrate. Experiment was carried out in microtiter plate (150 μl of medium per well) at 30°C. *P. putida* EM42 Δgcd, filled squares; mk-1, open squares. Data points shown as means of absorbance A600 of four biological replicates. Standard deviations were within 10% of the mean values.

Fig. S28. Stock loading system. By using a loading port (1) and coupling a sterile empty syringe (2), the tubes connecting both bottles are filled with liquid (3), which induce a liquid transfer from the new stock bottle (left) to the empty reactor container bottle (right) by height differences (4) with minimum risk of contamination.

© 2020 The Authors. *Microbial Biotechnology* published by Society for Applied Microbiology and John Wiley & Sons Ltd

environmental microbiology

Environmental Microbiology (2021) **23**(8), 4418–4433 doi:10.1111/1462-2920.15628

An updated structural model of the A domain of the *Pseudomonas putida* XylR regulator poses an atypical interplay with aromatic effectors

Pavel Dvořák,[1] **Carlos Alvarez-Carreño**,[2,3] **Sergio Ciordia**,[4] **Alberto Paradela**[4] **and Víctor de Lorenzo**[2*]

[1]*Department of Experimental Biology (Section of Microbiology), Faculty of Science, Masaryk University, Brno, Kamenice 753/5, 62500, Czech Republic.*

[2]*Systems Biology Department, Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco, Madrid, 28049, Spain.*

[3]*Centro Tecnológico José Lladó, División de Desarrollo de Tecnologías Propias, Técnicas Reunidas, Calle Sierra Nevada, 16, San Fernando de Henares, Madrid, 28830, Spain.*

[4]*Proteomics Core Facilit, Centro Nacional de Biotecnología-CSIC, Campus de Cantoblanco, Madrid, 28049, Spain.*

Summary

A revised model of the aromatic binding A domain of the σ^{54}-dependent regulator XylR of *Pseudomonas putida* mt-2 was produced based on the known 3D structures of homologous regulators PoxR, MopR and DmpR. The resulting frame was instrumental for mapping a number of mutations known to alter effector specificity, which were then reinterpreted under a dependable spatial reference. Some of these changes involved the predicted aromatic binding pocket but others occurred in distant locations, including dimerization interfaces and putative zinc binding site. The effector pocket was buried within the protein structure and accessible from the outside only through a narrow tunnel. Yet, several loop regions of the A domain could provide the flexibility required for widening such a tunnel for passage of aromatic ligands. The model was experimentally validated by treating the cells *in vivo* and the purified protein *in vitro* with benzyl bromide, which reacts with accessible nucleophilic residues on the protein surface. Structural and proteomic analyses confirmed the predicted in/out distribution of residues but also supported two additional possible scenarios of interaction of the A domain with aromatic effectors: a dynamic interaction of the fully structured yet flexible protein with the aromatic partner and/or inducer-assisted folding of the A domain.

Introduction

Operons encoding enzymatic routes typically found in environmental bacteria for biodegradation of aromatic environmental pollutants are often regulated by transcriptional factors (TFs) directly responsive to the pathway substrates themselves or to metabolic intermediates of the catabolic process (Shingler, 2003; Galvão and de Lorenzo, 2006). One conspicuous class of such factors belongs to the so-called NtrC superfamily of bacterial enhancer binding proteins (EBPs) that act at a distance on cognate promoters in concert with the σ^{54}-containing form of RNA polymerase (Weiss *et al*., 1992; North *et al*., 1993). The conserved structure of EBPs comprises three distinct domains (Fig. 1), the N-terminal of which (the so-called A domain) being the one that receives the environmental signal that turns on the protein to become an effective transcriptional activator. In a subset of EBPs, the A domain inhibits the interaction of the central domain of the regulator with the σ^{54}-dependent transcription initiation complex. Binding of the aromatic effector to the A domain of the EBP at stake relieves the intramolecular repression thereby triggering transcription initiation (Fig. 1D; Pérez-Martín and de Lorenzo, 1995). Archetypal TFs of this sort include the XylR protein, which regulates the two σ^{54}-dependent promoters found in the *xyl* operons for degradation of *m*-xylene borne by TOL plasmid pWW0 of *Pseudomonas putida* mt-2 (Fig. 1; Abril *et al*., 1989). That the interaction of the aromatic effector with the TF is localized in a small protein segment that can be swapped with homologue moieties with other specificities in similar EBPs had made the A moiety of XylR an appealing platform for developing biosensors for a variety of aromatic compounds (Galvão and

Received 14 January, 2021; revised 16 May, 2021; accepted 6 June, 2021. *For correspondence. E-mail vdlorenzo@cnb.csic.es; Tel. (+34) 91 585 45 36; Fax (+34) 91 585 45 06.

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

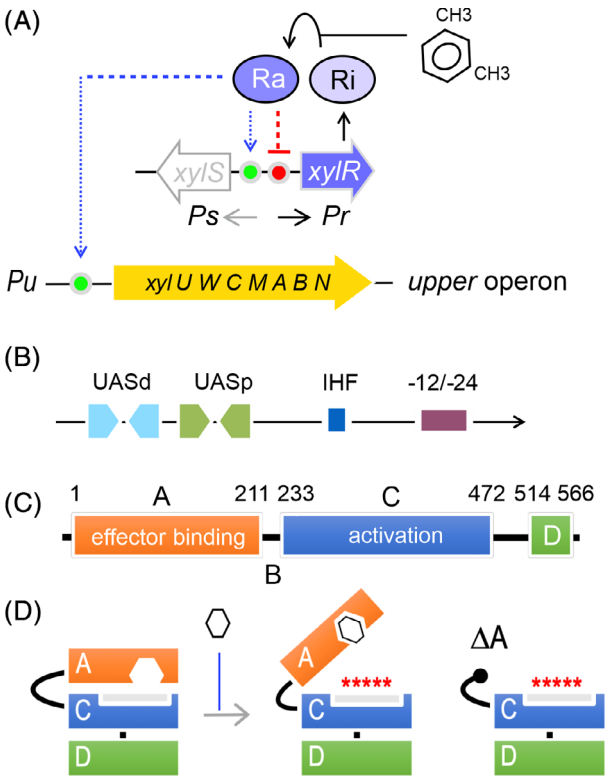


Fig 1. Biological and regulatory context of XylR and role of its A domain.

A. The products of the *upper* TOL pathway encoded by plasmid pWW0 transform *m*-xylene into 3-methylbenzoate, and the *lower* operon (not shown) that produces enzymes for further metabolism of this compound into TCA cycle intermediates. XylR and XylS are the transcriptional regulators that control the expression of either operon. The *xylR* is expressed from the *Pr* promoter and XylR produced in an inactive form (Ri) that, in the presence of the pathway substrate (*m*-xylene) changes to an active form (Ra). Ra activates both *Pu* and *Ps*, triggering expression of the *upper* pathway and XylS respectively. Ra also acts as a repressor of its own transcription. Operons and regulatory elements not to scale.

B. The *Pu* promoter region. The DNA segment of interest is expanded, showing the location of distal and proximal upstream binding sites for XylR (UASd and UASp), the −12/−24 motif recognized by σ^{54}-RNAP, and one integration host factor (IHF) binding site located in between.

C. Functional domains of XylR (C signs ATPase domain and D stands for DNA binding domain). Note that XylR binds the UAS of *Pu* regardless of inducer addition. The sketch indicates amino acid residues at the limits between the functional domains and the localization of the relevant functions within the protein sequence.

D. Activation of XylR by *m*-xylene. The TF folds such that the N-terminal A domain hinders an activation surface of the regulator. Effector binding to the A domain releases the inhibition and XylR is then able to activate σ^{54}-RNAP. Deletion of the whole A domain originates an effector-independent, constitutively active variant of XylRΔA. [Color figure can be viewed at wileyonlinelibrary.com]

de Lorenzo, 2006; Huang *et al*., 2008) including metabolites and explosives (Garmendia *et al*., 2008; de las Heras and de Lorenzo, 2011). Alas, this endeavour has been recurrently curbed by the lack of a reliable structural reference which has not only prevented directed mutagenesis of the residues involved in effector recognition, but also a mechanistic understanding of the many XylR mutants responsive to non-native inducers. This is because of the difficulty to purify the intact protein in a native form. ΔA truncated variants of XylR which activate constitutively the target σ^{54}-promoters of the TOL plasmid (Fig. 1) can be purified in large amounts. But to this day, purification of full-length XylR (or its separate A domain) in an active form has been challenging owing to the tendency of the product to form insoluble inclusion bodies (Pérez-Martin *et al*., 1997; Kim *et al*., 2005).

As a way to overcome these limitations, a structural model of XylR A domain was proposed by Devos and colleagues (2002) using the best bioinformatic tools for protein threading available at the time (Fig. S1). The model was based on (i) crystallographic data collected for catechol o-methyltransferase (COMT: PDB ID: 1VID; Vidgren *et al*., 1994), (ii) distribution of characteristic residues in sequences from related families (XylR, DmpR, or HbpR among others), nonpolar determinants and correlating mutations (multiple sequence alignment), and (iii) physico-chemical properties of the conserved amino acids in proteins from the family of σ^{54}-dependent EBPs and the COMT enzyme. The XylR structural model had 207 residues and had the shape of a typical Rossmann fold architecture with eight α helices and seven β strands. The binding cavity was proposed to be formed by loops and residues M37, F65, E140 and E172. The remaining residues of mostly hydrophobic character were proposed to stabilize aromatic ligand in the binding site and define the specificity of XylR (Delgado and Ramos, 1994; Shingler and Pavel, 1995; Garmendia *et al*., 2001). This model was for years the only one available to interpret XylR mutations and make sense of their phenotypes. Yet, it became obvious from the beginning that the structural basis of many mutants could not be easily recognized in the model and therefore a better one was badly needed.

Fortunately, in 2016, crystal structures were published of the A domains of the EBPs and XylR homologues PoxR and MopR (Patil *et al*., 2016; Ray *et al*., 2016) that activate σ^{54}-promoters for phenol-degradation operons of *Ralstonia eutropha* and *Acinetobacter calcoaceticus* respectively. The structure of the A domain of the archetypal dimethylphenol-responsive σ^{54}-activator DmpR from *P. putida* has recently been published as well (Park *et al*., 2020). Alignment of the N-terminal A-domains (211 amino acids) of these proteins (Fig. S2) shows a high sequence identity among the four proteins. Furthermore, that the A domains of DmpR and XylR can be swapped without any loss of function other than the exchange of effector specificity (i.e. *m*-xylene vs. phenol; Shingler and Moore, 1994) indicates the likeness of their respective tertiary structures. This state of affairs has enabled us to revise the structure of the A domain of XylR with the reliable frame of PoxR, MopR and DmpR.

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

In this work, we have developed and studied a model of the effector binding moiety of XylR that both overcomes the shortcomings of that of Devos and colleagues (2002) and provides a structural rationale for many of the incomprehensible phenotypes of mutants generated in the past on this TF. Furthermore, we provide genetic and biochemical evidence in support of the proposed tridimensional shape of the domain. Finally, we show data suggesting that aromatic effectors of XylR access the protein *in vivo* before it is fully folded into an otherwise non-receptive and transcriptionally dead form—rather than binding the mature TF *in vitro*.

Results

Modelling of XylR A domain structure by molecular threading

XylR A domain shows 40%, 41% and 46% sequence identity with ligand recognition domains of PoxR, MopR and DmpR respectively (Fig. S2). Indeed, crystal structures of PoxR (PDB ID: 5FRU and 5FRV), MopR (PDB ID: 5KBE) and DmpR (PDB ID: 6IY8) sensory domains (or whole EBP in case of DmpR) obtained in high resolution (1.85, 1.90, 2.50 and 3.42 Å respectively) appeared repeatedly among top 10 templates selected based on their significance from the LOMETS threading programs during I-TASSER calculations. The phenol-responsive sensory domain of PoxR (PDB ID: 5FRU) was, with 93% coverage and TM score of 0.91, structurally closest to the modelled XylR A domain. The confidence of each model generated by I-TASSER is quantitatively measured by C-score which is typically in the range of -5 to $+2$ (Yang and Zhang, 2015). The higher the value, the better is the model. C-score and RMSD (root-mean-square deviation of atomic positions) of the top-ranked model of XylR A domain were 0.93 and 3.6 $\pm$ 2.5 respectively, which signs highly accurate prediction.

The top-ranked model (Fig. 2), used for further work, shows typical structural features previously described for PoxR, MopR and DmpR proteins (Patil *et al*., 2016; Ray *et al*., 2016; Park *et al*., 2020). The XylR A domain is formed by a mixed α/β fold of seven α helices and seven β strands. N-terminal part consists of two α helices (α1 and α2) and three-stranded antiparallel β sheet between them. In PoxR, MopR and DmpR this part of the sensory domain together with helix α5 play a crucial role in protein dimerization. It can also be assumed that XylR A domain forms a dimer. The core part of the domain comprises a four-stranded antiparallel β sheet (β4–β7) and a bundle of three α helices (α3, α4 and α6). The binding pocket for aromatic ligands (Fig. 2), as predicted by I-TASSER and verified by CAVER web (Stourac *et al*., 2019), lies in between and is formed by the residues F93, G96, P97, Y100, V108, V124, A126, W128, Y155, A156, Y159, F170 and I185 which originate in all four β strands and in helices α4 and α6. The calculated volume of the pocket is 207 $\pm$ 3 Å^3. Seven (G96, P97, V108, W128, Y155, A156 and Y159) out of the 13, mostly hydrophobic, residues are conserved among XylR, PoxR, MopR and DmpR (Fig. S2). Position 124 is variable but in all three proteins is occupied by a small hydrophobic residue. Positions 100, 126 and 170 are occupied by histidine, phenylalanine and tyrosine respectively, in PoxR and MopR, or by histidine, methionine and phenylalanine respectively, in DmpR. Particularly size of the pocket and alterations in the binding site residues seem to be the key determinants of the ligand specificity of XylR and related bacterial enhancers (Patil *et al*., 2016; Ray *et al*., 2016, 2018). To verify the accuracy of the size of the modelled binding pocket, three cognate effectors (*m*-xylene, 3-methylbenzyl alcohol, toluene) and three bulky aromatic molecules that were previously shown to have no activation capacity for the wild-type XylR (2,4-dinitrotoluene, biphenyl and nonylbenzene) were docked in the cavity (Fig. S3) and sorted based on the average ΔG of their five top-ranked orientations (Abril *et al*., 1989; Galvão *et al*., 2007). All three XylR activators showed significantly ($P \leq 0.001$) lower binding energies (-7.12 ± 0.18, -6.66 ± 0.18 and $-\ 6.40 \pm 0.12$ kcal mol^{-1} for *m*-xylene, 3-methylbenzyl alcohol and toluene respectively) than 2,4-dinitrotoluene, biphenyl and nonylbenzene (-3.62 ± 0.76, -2.14 ± 0.05 and -1.16 ± 0.78 kcal mol^{-1} respectively). These docking experiments indicate that the binding cavity of XylR A domain is better suited for aromatic ligands with a single benzene ring and not too bulky substituents.

The inducer-recognition site subregion in the XylR A domain model is followed by a putative zinc binding pocket between α6 and the second antiparallel hairpin motif (β6, β7). As was proven for PoxR, MopR and DmpR, C149 from α6 and E172, C175 and C183 from the hairpin motif (XylR numbering is used) bind zinc atom which is important for structural integrity of the whole domain (Patil *et al*., 2016; Ray *et al*., 2016; Park *et al*., 2020). The XylR E172K mutant reported by Delgado and Ramos (Delgado and Ramos, 1994) showed a substantially reduced response to cognate effector molecules such as toluene or *m*-xylene. This fact together with the highly conserved nature of the four residues among NtrC family members (Fig. S2; Laitaoja *et al*., 2013) sign the importance of the site and possible metal binding also in XylR. C-terminal part of the XylR A domain consists of helix α7 which transmits the signal upon effector binding to the B linker region.

Prediction of inducer-access tunnels in the XylR A domain model

Because the predicted binding pocket is buried in the modelled A domain structure, a computational analysis was performed using CAVER web (Stourac *et al*., 2019) to reveal

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

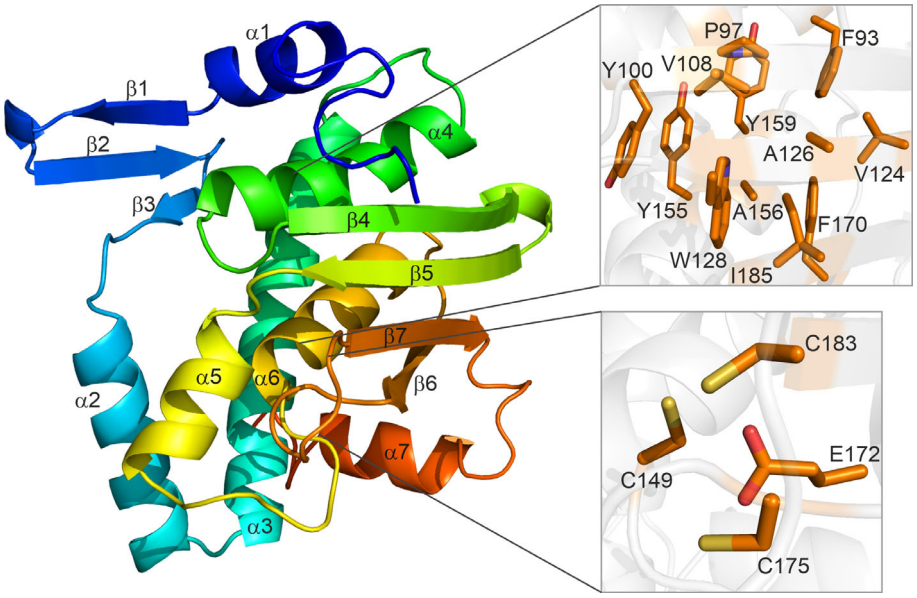


Fig 2. Tertiary structure of XylR effector binding domain A predicted by molecular threading with I-TASSER (Yang and Zhang, 2015) using 211 N-terminal amino acids of the transcription factor (Uniprot accession code: P06519). The structure backbone is coloured from N terminus (blue) to C terminus (orange). Secondary structure elements (seven β strands and seven α helices) are labelled. Binding pocket residues (upper zoomed-in window) suggested by CAVER web 1.0 (Stourac *et al*., 2019) and residues of a conserved zinc binding site (lower zoomed-in window) identified previously in crystal structures of PoxR, MopR and DmpR are highlighted as sticks. Only side chains of selected residues without hydrogens are visualized for better clarity (binding pocket residue G96 next to P97 is not shown). Dimerization interface (N motif) is supposedly formed by α helices α1 and α2, three-stranded antiparallel β sheet between them and helix α5. [Color figure can be viewed at wileyonlinelibrary.com]

possible entry site(s) and tunnel(s) that connect the cavity with the bulk solvent. Two hypothetical tunnels were predicted by the software – one going through the bundle of helices α3, α4 and α6 and the other one passing between β4 and α4 (Fig. 3A). Despite having a slightly narrower calculated bottleneck (0.7 vs. 0.9 Å), the latter tunnel seems to be the major passage for ligands. The main tunnel in PoxR sensory domain was proposed in a similar location and with bottleneck residues E96 and P112 (Patil *et al*., 2016) which correspond to L94 and L110 suggested, together with F93 and M113, by CAVER for XylR (Fig. 3B). Moreover, the spatial shifts of the whole hairpin motif β4 and β5, including bottleneck residue P112, observed in PoxR sensory domain upon binding of ligands with different dimensions indicated certain flexibility of the binding pocket and its *mouth* between β4 and α4 (Patil *et al*., 2016). This flexibility, possibly enhanced by the mobility of loop regions that interconnect secondary elements defining the binding pocket and its entry (Fig. 3B), might be another factor that shapes ligand specificity among related bacterial enhancers from the NtrC superfamily in diverse environmental conditions.

Mapping and reinterpretation of mutations in XylR A domain

The new model allowed re-interpretation of mutations in the ligand recognition domain reported during the last three decades in studies focused on the alteration of XylR effector specificity (Table 1, Fig. 4). The specificity of XylR was modulated by targeted or random mutagenesis toward phenols and their derivatives (Garmendia *et al*., 2001; Galvão *et al*., 2007), precursors of explosive chemicals such as nitrotoluenes (Delgado and Ramos, 1994; Garmendia *et al*., 2001; Galvão *et al*., 2007; de las Heras and de Lorenzo, 2011), or bulky effectors such as biphenyl (Garmendia *et al*., 2001). As shown here by molecular docking, these compounds can be theoretically accommodated in the binding pocket of XylR A domain but their binding energies are higher than those of natural effectors (Fig. S3). Hence, modification of the pocket's shape and physicochemical properties is a necessary prerequisite for the productive binding of these ligands.

In the former model, Devos and colleagues (2002) proposed a binding groove to be formed by four loops, and particularly four conserved residues in this site (M37, F65, E140 and E172) were expected to contribute to the effector specificity of XylR (Fig. S1). However, none of these amino acids makes part of the actual binding pocket nor are they in close proximity to it (Fig. 4A). M37, F65 and E140 are positioned in dimer interface (in β3, α3 and α5 respectively) while E172 is one of the residues forming putative zinc binding site. Also, the majority of the reported mutations have different locations and

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

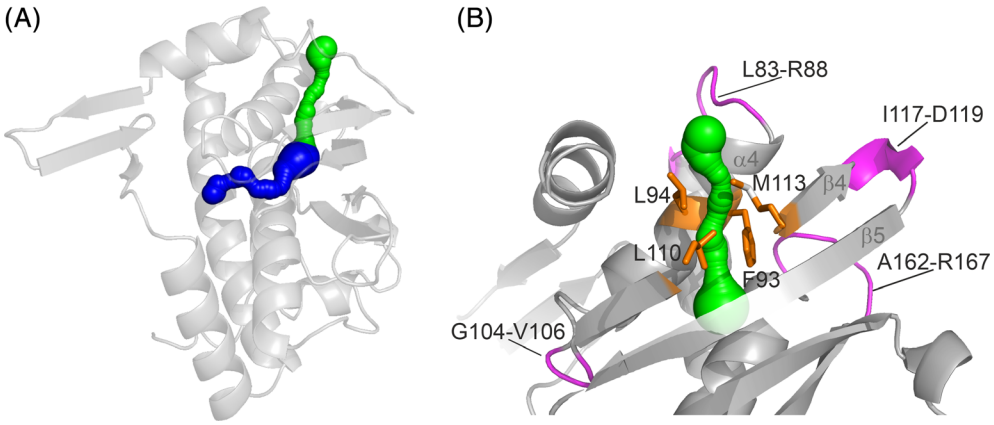


Fig 3. Protein tunnels for entry of effectors into the binding pocket of XylR identified by CAVER web 1.0 (Stourac *et al*., 2019).
A. Two identified tunnels that connect XylR binding pocket with bulk solvent are shown in blue and green.
B. The probable main tunnel (in green) passing between helix α4 and strand β4 is shown together with proposed bottleneck residues (orange sticks) and loop regions (in magenta) whose flexibility might affect the size of the tunnel and the binding pocket. [Color figure can be viewed at wileyonlinelibrary.com]

structural effects than previously suggested (Table 1, Fig. 4B). Only three amino acid substitutions (P97A, V124A and Y159) occurred directly in the binding pocket. All these mutations expanded the effector scope of XylR toward bulkier ligands such as xenobiotic chemical 2,4-dinitrotoluene or biphenyls (Table 1). Their effect can thus be attributed to the expansion or loosening of the pocket. For instance, V124A, located in the beta strand β5 in the new model, significantly ($P < 0.01$) increases the volume of the pocket from 207.0 $\pm$ 3.0 Å^3 to 288.7 $\pm$ 12.7 Å^3 as calculated by CAVER web and visualized in PyMOL (Fig. S4A). *In silico* docking experiments indicated better binding of bulky ligands in the enlarged cavity. Three previously tested non-native ligands of XylR – 2,4-dinitrotoluene, biphenyl and nonylbenzene – showed lower binding energies (-4.62 ± 0.92, -5.40 ± 0.00 and -3.64 ± 0.28 kcal mol^{-1} respectively) when docked in V124 mutant and compared with the previous docking in wild type (-3.62 ± 0.76, -2.14 ± 0.05 and -1.16 ± 0.78 kcal mol^{-1} respectively). The decrease in binding energies was statistically significant for the latter two molecules ($P < 0.01$).

Interestingly, more reported mutations were located along with the A domain dimer interface (F48I/T, F65L, D135N, I136T, shuffled region 46–50). The frequent occurrence of mutations in this region in protein variants with altered inducer specificity was observed also for the other mentioned EBPs (Ray *et al*., 2016). Several substitutions were present in the putative zinc binding site (E172K) or in close proximity to it (S174R, L184I). These might affect the folding of the whole domain and secondarily also the flexibility of the binding cavity. In the case of the aforementioned substitution E172K, this effect was already too pronounced and detrimental for the productive binding of most of the tested ligands including natural effectors of XylR (Table 1; Delgado and Ramos, 1994).

Last but not least, some reported mutations (P85S, shuffled region 161–166) can be found in the loops that connect secondary structure elements that shape the binding pocket and zinc binding site. For instance, the substitution of rigid proline in position 85 for serine resulted in XylR variant capable of transcription activation in the absence of effector (Delgado *et al*., 1995). This is not surprising from the current perspective, because a new A domain model locates the residue in the loop between helices α3 and α4 which form the bottom of the binding pocket and its entry tunnel. Similarly, shuffling in the loop region 161–166 improved response to nitrotoluenes, phenol and bulky biphenyl probably via altering the flexibility of helix α6 and beta sheet β6 that contribute to the binding pocket and zinc binding site. Also expanded effector scope toward 2,4-dinitrotoluene observed for mutant Y159F might be attributed to the loosening of the crucial parts of the A domain. Hydrogen from hydroxyl group of Y159 in helix α6 can form H-bond with oxygen from residue V92 in α4 and stabilize the bottom of the binding pocket (Fig. S4B). This bond is absent in mutant Y159F. Mutation L222R in B linker, which is responsible for signal transduction between A and ATPase domain, is worth mentioning too though it cannot be located in the current model. This substitution emerged repeatedly from error-prone PCR libraries together with mutations in A domain and gave rise to XylR variants responsive to 2,4-dinitrotoluene (Table 1; Galvão *et al*., 2007).

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

Table 1. Reinterpretation of mutations in XylR A domain.

Mutation	Former location and interpretation[a]	Updated location and interpretation	Effect	Reference
F48I/T	Helix α3 close to the C terminus: Modulation of signal transfer from A to B and C domains of XylR	Helix α2: Dimer interface	Novel response to 2,4-dinitrotoluene (DNT), 3-NT and chlorinated phenols; improved response to 2- and 3-NT	Galvão *et al*., 2007; de las Heras and de Lorenzo, 2011
F65L	Loop between β1 α4: Putative binding pocket	Helix α3: Dimer interface	Improved response to 2- and 4-NT and cognate effector 3-xylene; novel response to 3-NT	(Garmendia *et al*., 2001)
P85S	Strand β2	Loop between helices α3 and α4: Increased flexibility of the loop and binding pocket	Activation of transcription in absence of effector	Delgado *et al*., 1995
P97A	Helix α5: Interdomain contacts	Helix α4: Binding pocket, near access tunnel	Novel response to 2,4-DNT	Galvão *et al*., 2007
V124A	Helix α6: Interdomain contacts	Strand β5: Binding pocket	Stronger response to cognate effectors 3-xylene and benzene; improved response to NTs and biphenyls	Garmendia *et al*., 2001
D135N	Strand β4: Interdomain contacts	Helix α5: Dimer interface	Activation of transcription in absence of effector; improved response to 2-NT and novel response to 3-NT	Delgado *et al*., 1995; Salto *et al*., 1998
I136T	Strand β4: Near putative binding pocket	Helix α5: Dimer interface	Improved response to new effector 2,4-DNT	de las Heras and de Lorenzo, 2011
Y159F	Loop between α7 and β5	Helix α6: Binding pocket, stabilization of helices α6 and α4 forming bottom of binding pocket.	Novel response to 2,4-DNT and chlorinated phenols; improved response to NTs in combination with mutation in L222	Galvão *et al*., 2007
E172K	Loop between β5 and α8: Putative binding pocket	Zinc binding site	Lower response to native effectors; novel response to 3-NT	Delgado and Ramos, 1994
S174R	Loop between β5 and α8: Putative binding pocket	C terminus of strand β6: Near zinc binding site	Improved response to new effector 2,4-DNT	de las Heras and de Lorenzo, 2011
L184I	Helix α8: Interdomain contacts	N terminus of strand β7: Near zinc binding site	Improved response to 2- and 4-NT and benzene; novel response to 3-NT	Garmendia *et al*., 2001
L222R	B linker transmitting signal from domain A do domain C.	B linker transmitting signal from domain A do domain C.	Novel response to 2,4-DNT	Galvão *et al*., 2007; de las Heras and de Lorenzo, 2011
Shuffled region 46–50[b]	Helix α3	Helix α2: Dimer interface	Improved response to NTs, novel response to phenol and biphenyl	Garmendia *et al*., 2001
Shuffled region 161-166[b]	Strand β5	Loop between α6 and β6: Flexibility of the beta sheet forming part of binding pocket and zinc binding site; signal transmission to B linker	Improved response to NTs, novel response to 3-NT, phenol and biphenyl	Garmendia *et al*., 2001

[a]Based on previous structural model published by Devos and colleagues (2002).
[b]This region was modified by DNA shuffling between A domains of XylR and DmpR.

Taken together, the hereby presented model of XylR A domain allows more realistic explanation of the effects of new and previously reported mutations. Despite the fact that the interpretations based on the old model of Devos and colleagues (2002) were mostly incorrect, the major conclusions of the former studies remain valid. Effector specificity of XylR and related transcription factors is a result of complex, sometimes counterintuitive, interactions of several factors (de las Heras and de Lorenzo, 2011). It can be modulated by substitutions in the binding pocket (Ray *et al*., 2018) as well as targeted or random mutations occurring in distant locations, particularly in A domain dimer interface, in certain loops, or in regions responsible for signal transduction. The new

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

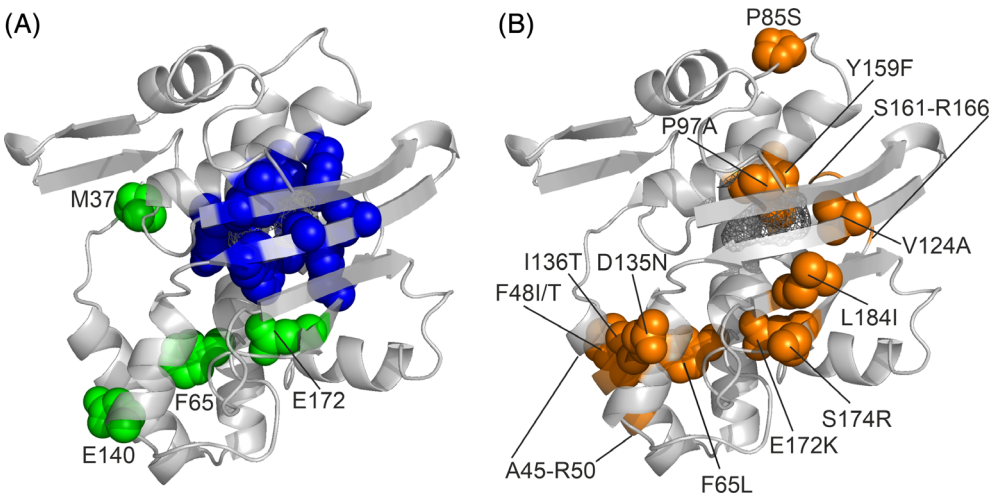


Fig 4. Mapping of revisited binding pocket residues and residues previously targeted by mutagenesis (listed in Table 1) on the predicted structure of XylR effector binding domain.
A. The residues proposed to interact with XylR effectors based on the structural model of Devos and colleagues (2002) are shown as green spheres and labelled. Side chains of the binding pocket residues proposed for the new model (F93, G96, P97, Y100, V108, V124, A126, W128, Y155, A156, Y159, F170, and I185) are shown as blue spheres.
B. Single residues reported in mutagenesis studies are shown as orange spheres and labelled. Shuffled regions (A45-R50 and S161-R166) are shown as orange parts of the cartoon. The surface of the binding pocket is shown as a grey wireframe. [Color figure can be viewed at wileyonlinelibrary.com]

structural model of XylR A domain can be now used in combination with advanced computational tools such as molecular dynamics simulations to further specify the effects of some of the discussed mutations (Klvana *et al*., 2009).

Rationale for experimental validation of the revised model

Next, we sought to verify the new model experimentally. One possible option was a chemical cross-linking with an effector molecule possessing a functional group that can form a covalent bond with certain amino acid moieties of the protein (Mattson *et al*., 1993). Such a technique in combination with mass spectrometry can be used to confirm amino acids exposed to the solvent, including those in the ligand binding pouch (Gingras *et al*., 2007). Benzyl bromide with highly reactive bromomethyl substituent was selected as a suitable structural analogue of natural aromatic effectors such as *m*-xylene or toluene. Binding pocket in XylR A domain can theoretically accommodate benzyl bromide (Fig. S3G), though the binding energy ΔG calculated during molecular docking of this ligand (-5.58 ± 0.24 kcal mol^{-1}) was somewhat higher than that of natural effectors *m*-xylene or toluene (-7.12 ± 0.18 and -6.40 ± 0.12 kcal mol^{-1} respectively). Benzyl bromide acts as a selective alkylator of sulfur nucleophiles such as methionine or cysteine but can react also with other nucleophilic amino acids (including tyrosine, aspartic acid, or lysine present in the binding pocket or on the surface of XylR) when applied in a higher amount and for a longer time interval (Rogers *et al*., 1976; Lang *et al*., 2006). We hypothesized that interactions of XylR with benzyl bromide would provide information on site(s) in the protein structure which naturally interact(s) with aromatic effectors and amino acid residues exposed on the protein surface, and, thus, would either support or disprove the validity of the new model.

In vivo evidence of interaction between XylR and benzyl bromide

Benzyl bromide interaction with XylR was initially studied using *Escherichia coli* strain CC118 *Pu-lacZ* (Table S1) bearing recombinant plasmid pCON916 with *xylR* gene under the control of its native *Pr* promoter (de Lorenzo *et al*., 1991; Garmendia *et al*., 2001). In this system, wild-type XylR activated by *m*-xylene induced expression of β-galactosidase whose activity was quantified (Fig. 5A and B; Miller, 1972). When benzyl bromide was used instead of *m*-xylene, no β-galactosidase activity (neither the background activity observed in the absence of effector) was detected (Fig. 5C). It is worth mentioning here that exposure to the vapours of *m*-xylene and benzyl bromide did not affect the viability of the bacterial strains used in this study. Hence, there were two possible explanations for the observed lack of β-galactosidase activity in cells exposed to benzyl bromide. Either benzyl bromide inactivated β-galactosidase or it interacted with XylR in a way that resulted in a protein form unable of *Pu-lacZ* induction. To test the first hypothesis, β-galactosidase

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

activity was measured in strain *E. coli* MC4100 [MAD2] which expressed the mutant *xylRΔA* gene which encodes a constitutive variant of the transcription factor lacking the signal recognition domain (Table S1). As expected, XylRΔA activated the *Pu-lacZ* fusion of *E. coli* MC4100 [MAD2] both in the presence and absence of the ligand. β-galactosidase activity was detectable in this strain even after the addition of benzyl bromide (Fig. 5C) therefore indicated that benzyl bromide did not affect the enzyme but rather interacted with XylR, most probably with its A domain. These results showed that the wild-type XylR is activated when it comes in contact with *m*-xylene and it loses its capacity to induce transcription upon interaction with benzyl bromide.

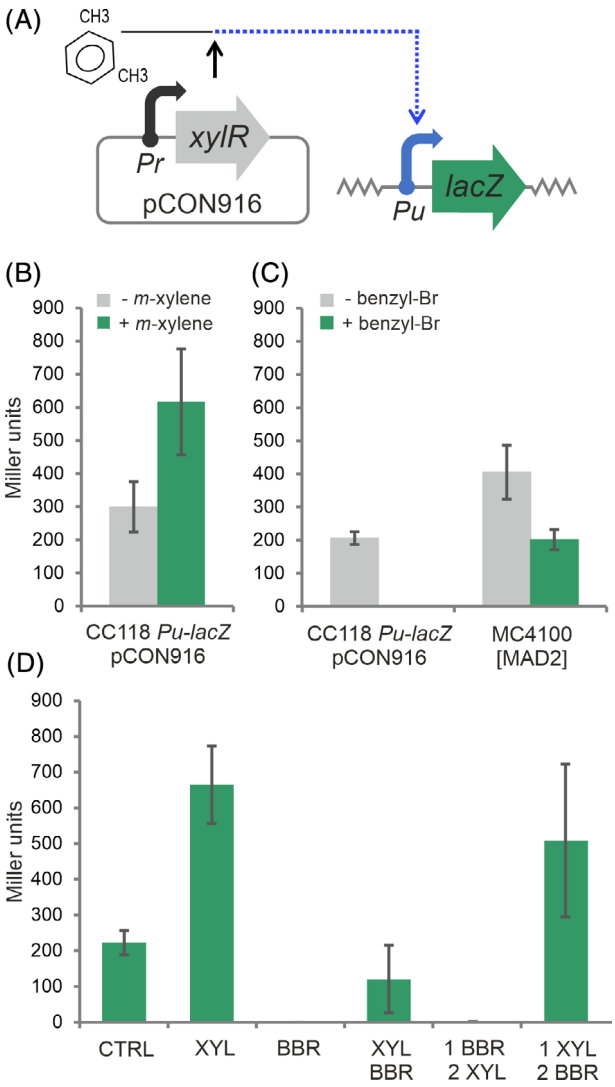


Fig 5. *In vivo* evidence of XylR interaction with *m*-xylene and benzyl bromide.
A. A scheme of the sensor device in *Escherichia coli* CC118 *Pu-lacZ* strain with pCON916 plasmid bearing wild-type *xylR*.
B. β-galactosidase activity of *E. coli* CC118 *Pu-lacZ* pCON916 in the absence or presence of *m*-xylene effector.
C. β-galactosidase activity of *E. coli* strain CC118 *Pu-lacZ* pCON916 and strain MC4100 [MAD2], bearing *xylR* variant with deleted A domain, in absence or presence of benzyl bromide.
D. β-galactosidase activity of *E. coli* CC118 *Pu-lacZ* pCON916 strain exposed to *m*-xylene and/or benzyl bromide in several specific conditions. Cells were grown to OD_{600} of 1.0 and then were: left uninduced (CTRL), exposed to saturated vapours of *m*-xylene (XYL), benzyl bromide (BBR), or of both chemicals in parallel (XYL/BBR), first exposed to benzyl bromide and then to *m*-xylene (1 BBR/2 XYL), or first exposed to *m*-xylene and then to benzyl bromide (1 XYL/2 BBR). Shown data represent means ± SD from two to three experiments. [Color figure can be viewed at wileyonlinelibrary.com]

Further tests were conducted again with *E. coli* CC118 *Pu-lacZ* pCON916. The strain was exposed to *m*-xylene and/or benzyl bromide in several different conditions to probe whether there was a competition between the two molecules for the same site in the structure of wild type XylR (Fig. 5D). Parallel exposure to *m*-xylene and benzyl bromide resulted in the significantly reduced β-galactosidase activity of the cells. Furthermore, XylR was completely unable to induce expression of *lacZ* when the cells were first exposed to benzyl bromide and then to *m*-xylene. However, the induction capacity of XylR was almost fully retained when the order of the two chemicals was reverse. These results suggested that benzyl bromide acted as an inhibitor of XylR activation by *m*-xylene and that such inhibition was probably competitive (i.e., inhibitor and effector molecules competed for the same binding site). To test this hypothesis, we turned to the biosensor strain *P. putida* BXPu-LUX, which bear chromosomal insertions of DNA encoding (i) the *xylR* gene expressed from its native promoter *Pr* and (ii) a *Pu-lux* transcriptional fusion (Fig. 6A; de las Heras *et al*., 2008). In this strain, XylR activated by an effector induces expression of the *luxCDABE* operon from *Photorhabdus luminescens* (Winson et al., 1998) and the specific bioluminescence of the whole cells can be detected non-disruptively in selected time intervals. On this background, *P. putida* BXPu-LUX cells were exposed to four different vapour pressures of *m*-xylene in the absence (Fig. 6B) or presence (Fig. 6C) of benzyl bromide and bioluminescence was recorded at fixed times. The inverse initial velocities of bioluminescence formation were plotted against the inverse relative vapour pressures in the double reciprocal plot (Fig. 6D). In the resulting graphics, the two dashed lines had the same Y-intercept, implying that benzyl bromide is a competitive inhibitor of XylR, i.e., it can enter and occupy non-productively the same binding site in the A domain.

Chemical cross-linking of purified XylR-His with benzyl bromide

The data above most plausibly indicated that benzyl bromide reaches out and interacts with the same target site

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

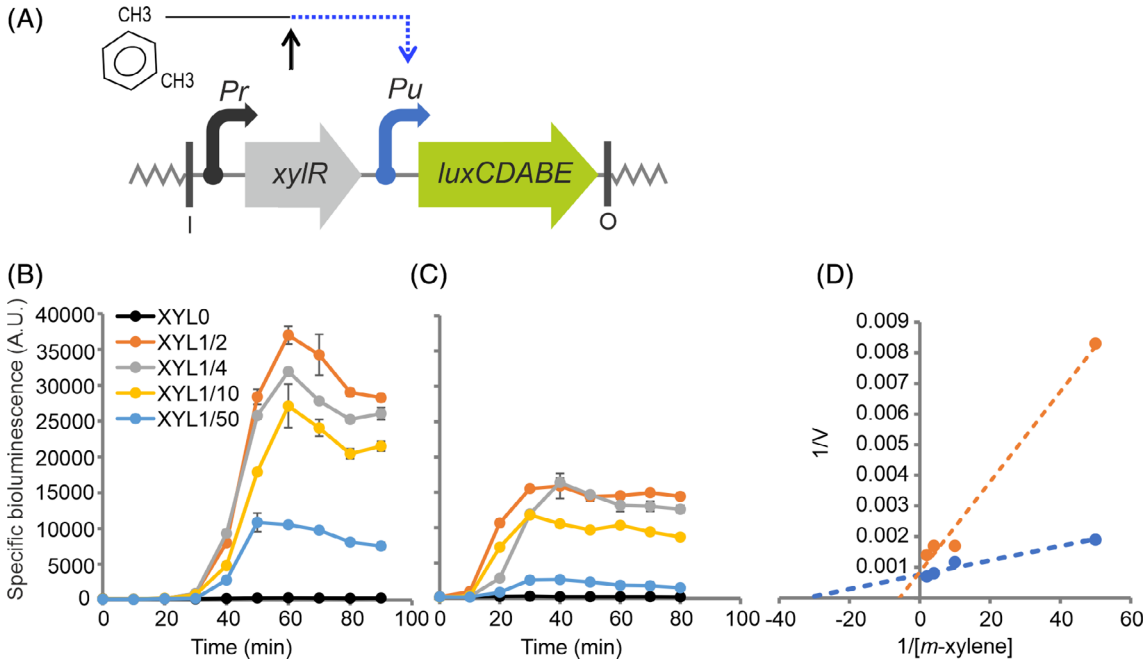


Fig 6. Determination of the competitive relation between *m*-xylene and benzyl bromide in XylR activation *in vivo*.
A. A scheme of the sensor device in *Pseudomonas putida* BXPu-LUX strain (de las Heras *et al*., 2008).
B. Detection of specific bioluminescence of *P. putida* BXPu-LUX exposed to four different vapour pressures of *m*-xylene (XYL). The fractions shown in the figure legend correspond to the portions of the effector mixed with dibutyl phthalate solvent in 50 μl of the total volume of the mixture used for the induction of the cells.
C. Detection of specific bioluminescence of *P. putida* BXPu-LUX exposed to four vapour pressures of *m*-xylene (identical to those used in the previous experiment) and constant vapour pressure of benzyl bromide. Benzyl bromide was mixed with dibutyl phthalate and *m*-xylene to form 1/20 of the total volume (50 μl) of the mixture.
D. A double reciprocal plot for the four vapour pressures of *m*-xylene without (blue dots) and with (orange dots) a constant vapour pressure of benzyl bromide plotted against the velocity of production of specific bioluminescence of the strain BXPu-LUX. The velocities were calculated for time intervals 30–50 min in Fig. 5A and 10–30 min in Fig. 5B. Shown data represent means ± SD from three biological replicates. [Color figure can be viewed at wileyonlinelibrary.com]

in the A domain of XylR, but it is unable to activate the protein – instead, it blocks the ability of *m*-xylene to act as a bona fide inducer. Given the chemical reactivity of benzyl bromide, we entertained that the treatment of purified XylR with this reagent could form covalent bonds with nucleophilic amino acids exposed on the surface of the transcription factor and those present in the accessible effector binding cavity of the A domain. Such bound amino acids could then be determined with mass spectrometry and mapped on the predicted protein structure – thereby validating or challenging the model. To this end, we generated a recombinant XylR variant with 6xHis tag on C terminus of the protein (see Experimental procedures for details). To verify its functionality, the DNA sequence of the His-tagged protein was first expressed from pCON1238 (Table S1) in *E. coli* CC118 *Pu-lacZ* and β-galactosidase activity of the cells was measured after induction with *m*-xylene and compared with *E. coli* CC118 *Pu-lacZ* (pCON916) encoding the wild type *xylR* gene (Fig. S5A and B). *E. coli* CC118 *Pu-lacZ* (pCON1238) cells were also exposed to *m*-xylene and/or benzyl bromide in the same conditions described before for *E. coli* CC118 *Pu-lacZ* (pCON916) as shown in Fig. S5C, Fig. 5D. Although the capacity of XylR-His to activate *Pu-lacZ* fusion was affected by the presence of the tag, the interaction pattern with *m*-xylene and benzyl bromide was almost identical to that observed for wild-type XylR.

Next, XylR-His was produced in soluble form in *Escherichia coli* BL21(DE3) pLysS transformed with pCON1238 plasmid and purified (> 90% purity) using immobilized metal affinity chromatography (Fig. S6, Experimental procedures). XylR was then mixed with the excess of benzyl bromide and the mixture incubated overnight in moderately alkali conditions to promote cross-linking with nucleophilic amino acids. The mixture was then loaded on SDS-PAGE gel and the XylR-His band was cut out for mass spectrometry analysis. XylR-His non-exposed to benzyl bromide was used as a control. Peptides with mass increments corresponding to benzylation were identified (Supporting Information File S2) and their amino acid sequence was determined. There were in total 16 modified amino acids in 15 different peptides (Table 2). These amino acids were distributed

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

Table 2. Peptides with amino acids cross-linked with benzyl bromide (highlighted) and their respective positions in the XylR-His protein.

No.	Sequence	Start-end	Position	Domain
1	MQHED**M**QDLSSQIR	9–22	14	A
2	I**W**LGEQR	30–36	31	A
3	EII**S**LIGVE**R**	51–60	60	A
4	LGYQSGL**M**DAELA**R**	68–81	75,81	A
5	LLTMDIA**R**	110–118	118	A
6	IIFQETS**C**R	168–176	175	A
7	**S** DPIVDER	205–212	205	A
8	**Y** ELQTQVANLR	213–223	213	B
9	QY**D**GQYYGIGHSPAYK	228–243	230	B
10	LIT**A**TNENLEEAV**K**	371–384	384	C
11	LNVFPV**H**IPPLR	396–407	402	C
12	AMEAC**L**HYQ**W**PGNIR	439–453	448	C
13	LEEESGDS**W**FR	496–506	504	C–D
14	QIIDQGVSL**E**DLEAGL**M**R	507–524	523	D
15	**C** GQNISQAAR	530–539	530	D

along the whole XylR sequence but, interestingly, eight of them (M14, W31, R60, M75, R81, R118, C175 and S205) were concentrated in A domain which represents only about one third of the whole protein size (Fig. 7A). The side chains of all amino acids that reacted with benzyl bromide (3x Met, 3x Arg, 3x Trp, 2x Cys, 1x Ser, 1x Tyr, 1x Asp, 1x Lys and 1x His) are potential nucleophiles in the deprotonated state with the following relative order of nucleophilicity of functional groups: $R–S^-$ > $R–NH_2$ > $R–COO^-$ = $R–O^-$ (Rogers *et al*., 1976; Lang *et al*., 2006; Bischoff and Schlüter, 2012). Heteroaromatic systems of Trp and His are less prone to benzylation but such reaction is possible due to the nature of *pi* electrons of indole and imidazole rings respectively, and was probably promoted by the longer reaction times of our experiment.

Mapping of the thereby modified residues on the structural model of the XylR A was done manually in PyMOL and verified using NetSurfP 2.0 server (Klausen *et al*., 2019). As shown in Fig. 7B and Fig. S7, the chains of seven of these amino acids (M14, W31, R60, M75, R81, R118 and S205), were exposed on the protein surface. The one exception was C175 that appeared *buried* in the protein structure. Note, however, that the nucleophilic thiol group of C175 is turned toward the protein surface (Fig. 7C) and it is the most accessible residue to benzylation out of the three cysteine residues that form a putative zinc binding site in this part of the signal recognition domain. It was remarkable, that none of the amino acids predicted to shape the effector binding pocket or the tunnel (Figs. 2 and 3) was crosslinked with benzyl bromide, although such a reaction could in principle be possible with residues Y100, M113, W128, Y155 and Y159. Taken together, these results suggest that benzyl bromide did react with amino acids accessible on the A domain surface but, under given conditions, could not penetrate the protein structure to reach the binding pocket. As discussed below, we believe this result is significant and can hint toward a possible distinct mechanism of activation of XylR.

Discussion

The data above calls for leaving behind the structural model of the XylR A domain proposed by Devos and colleagues (2002) and adopting an updated model prepared by the molecular threading of the cognate sequence within the available crystal structures of the A domains of related transcription regulators PoxR, MopR and DmpR. That the new model is way more reliable than the previous one is accredited by the results of treating the purified protein with benzyl bromide. This reagent is known to covalently bind highly nucleophilic methionines (Rogers *et al*., 1976; Lang *et al*., 2006), but due to the relatively long time of treatment, it can also react with less nucleophilic amino acids, including surface-exposed tryptophans or histidines. These experiments, which define precisely the interior and the exterior of the domain, delivered a virtually perfect match between the model and the results (Fig. 7).

The new, dependable structure of the XylR A domain, in particular the organization of the effector binding site, has enabled us to apply a wealth of structural analysis tools for the reinterpretation of abundant genetic data on this TF and it has provided hints for decoding its possible activation mechanism. Software for the detection of effector binding cavities in proteins (I-TASSER and Caver) was specially useful in this respect. For instance, 8 out of the 13 residues forming the predicted binding pocket overlapped the so called *minimal binding region* (amino acids 110–186) previously defined by mutagenesis of XylR (Pérez-Martín and de Lorenzo, 1996). Furthermore, when the modelled binding cavity was probed by molecular docking, the binding energies of the bound ligands

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

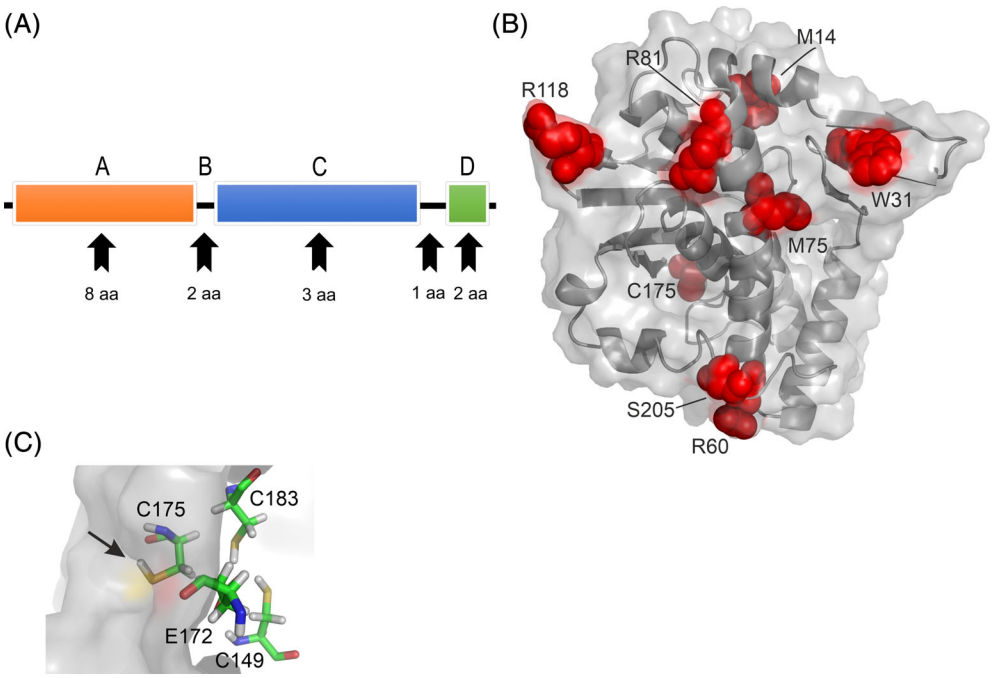


Fig 7. Mapping of the amino acid residues cross-linked with benzyl bromide in the XylR structure.
A. Modular scheme of XylR structure with all its domains and distribution of amino acids modified by benzyl bromide. The visualized XylR domains are: A domain which represents 37% of the whole XylR sequence, B linker (4%), C domain (ATPase domain, 42%) and D domain (DNA binding domain, 8%).
B. Mapping of the amino acid residues cross-linked with benzyl bromide in the predicted structure of XylR A domain. Modified residues are shown as red spheres in the XylR A domain structure depicted in cartoon and surface mode.
C. Detail of putative zinc binding site and thiol group of C175 (black arrow) in the A domain model. Zinc binding residues are shown as sticks coloured by element (carbon is green, nitrogen is blue, hydrogen is white, oxygen is red, sulfur is yellow), protein surface is depicted in grey. [Color figure can be viewed at wileyonlinelibrary.com]

matched well the actual effector preferences of the wild-type XylR (Abril *et al*., 1989; Galvão *et al*., 2007). Most of the residues in the predicted binding cavity are conserved among PoxR, MopR, DmpR and XylR. However, two residues in XylR pocket (Y100 and A126) differ from phenol binders PoxR, MopR and DmpR. Replacement of His102 (Y100 in XylR) for Tyr in PoxR made the TF insensitive to phenol (Patil *et al*., 2016). Single substitution H106Y (corresponds to Y100 in XylR) changed the specificity of MopR from phenols toward benzene ligands (Ray *et al*., 2018). Replacement of an additional two residues F132A (A126 in XylR) and Y176F (F170 in XylR) further enlarged the MopR cavity for bulkier benzene derivatives. These analyses provide a good picture of how the architecture of the aromatic binding pocket looks like and they pave the way for developing XylR-based biosensors with rationally re-designed effector specificities. Yet, note that effector specificity of XylR and TFs might be determined not only by the geometry of the cavity that accommodates the aromatic compound (Table 1) but also by other structural and mechanistic constraints (de las Heras and de Lorenzo, 2011). In fact, it has been observed that effector mutants can very often arise from changes in locations distant from the binding pocket and the predicted tunnel. In this respect, the updated model suggests the presence in the XylR A domain of a distinct segment – so called N motif (Fig. 2) – involved in dimerization of the equivalent section of PoxR, MopR and DmpR (Bush and Dixon, 2012; Patil *et al*., 2016; Ray *et al*., 2016; Park *et al*., 2020). In this segment, the majority of effector mutants of DmpR, MopR and XylR were mapped (Table 1; Patil *et al*., 2016; Ray *et al*., 2016). How can such dimerization interface determine agonist specificity? Perhaps such mutants affect intramolecular signal transmission downstream of the effector binding proper. In these cases, the mutation away from the binding site may upgrade a good binder but non-productive aromatic into an efficient inducer. Obviously, how this can happen deserves further studies.

Still, the most intriguing aspect of the A domain geometry is the fact that the ultimate binding site of the protein for the aromatic effector is well buried in the internal core of the protein structure and not readily accessible from the outside. The only apparent way to reach out such a site is by penetrating a narrow tunnel (0.7 Å at the narrower part) coated with apolar amino acids that can

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

hardly sustain the passage of molecules with a radius of ~7 Å like *m*-xylene (Li *et al*., 2020). We were puzzled that despite our finding that benzyl bromide can act as a competitive inhibitor of xylene-activated XylR *in vivo* (in both *E. coli* and *P. putida*, Figs. 5D and 6), no benzyl bromide crosslink occurred *in vitro* with any of the potentially nucleophilic amino acids of the access tunnel and binding pocket (candidates were Y100, M113, W128, Y155, or Y159, Figs. 2 and 3) of the purified A domain. In contrast, binding pockets of purified PoxR, MopR and DmpR with similarly narrow accession tunnels can accept aromatic ligands (Patil *et al*., 2016; Ray *et al*., 2016; Park *et al*., 2020). Patil and colleagues (2016) discussed in their work certain flexibility of the PoxR binding pocket and its mouth between β4 and α4 upon binding of ligands with different dimensions. This flexibility can be provided by the wealth of loop regions recognized in the resolved A domain structures of NtrC superfamily of EBPs including XylR (Fig. 3B). Because MopR sensory domain could be crystallized only in the presence of a ligand, Ray and colleagues (2016) have argued that it is quite flexible and exists in the open and closed forms. The switch from the open to the crystallisable compact form of A domain is proposed to be prompted by the binding of a ligand and zinc atom. Very recently, Park and colleagues (2020) reported that the phenol binding pocket in the resolved DmpR protomer structures (which showed stronger electron density and supposedly higher occupation by phenol) had a smaller volume than the pocket in protomers with weaker electron density. This result also implies the existence of open and compact forms of A domains of certain EBPs. Two possible explanations thus exist for the observed absence of benzyl bromide crosslink in the binding pocket of purified XylR. One is that the aromatic molecule could penetrate into the flexible binding cavity but the positioning or orientation of extant nucleophilic residues was not optimal for crosslink formation. An alternative explanation suggests a possible distinct yet complementary behaviour of XylR when compared with other three EBPs. Chances are that the A domain possesses an open ligand-accepting form *in vivo* but it cannot bind effectors under *in vitro* conditions when any potential port of entry to the binding cavity seems to be blocked. It could be that the inducer interacts with the protein while it is being produced and folded, not after it has matured. In this way, the constellation of contacts could occur while the A domain is still a partially structured domain and thus accessible to the aromatic agonist. While the relevance of the two proposed ligand binding scenarios needs to be verified with additional experiments, the latter would explain not only the benzyl bromide data above but also the recurrent failure to have a full-length XylR protein transcriptionally responsive to *m*-xylene in any *in vitro* test done thus far—while maintaining its full DNA binding ability (Pérez-Martin *et al*., 1997). Note that there may not be a clear-cut boundary between direct effector binding to the complete protein, binding a flexible form, or interacting during folding and it is imaginable that different EBPs have followed different evolutionary itineraries to the same end.

Experimental procedures

Strains, plasmids and growth conditions

Bacterial strains and plasmids used in this study are listed in Table S1. Nutrient-rich lysogeny broth (LB; (Sambrook *et al*., 1989) and defined M9 minimal medium (Miller, 1972) with 2 mM $MgSO_4$, 2 ‰ thiamine and 0.2% sodium citrate were used for growth of *E. coli* and *P. putida* strains. Bacteria were also grown in Petri dishes on LB agar (1.5%) solid medium or M9 agar (1.6%) medium with 2 mM $MgSO_4$, 2 ‰ thiamine and 0.2% sodium citrate. *E. coli* and *P. putida* strains were grown at 37°C and 30°C respectively, unless stated otherwise. If needed, liquid and solid media were supplemented with 100 or 500 μg ml^{-1} ampicillin (Ap, higher concentration was used for *P. putida*), 75 μg ml^{-1} kanamycin (Km), 30 μg ml^{-1} chloramphenicol (Cm), 10 μg ml^{-1} tetracycline (Tc), 1 mg ml^{-1} carbenicillin (Cb), or 10 μg ml^{-1} gentamicin (Gm) to select for bacteria with plasmid(s). The 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal; 40 μg ml^{-1}) was added to detect β-galactosidase activity. For the induction of liquid cultures with volatile compounds, 50 μl of *m*-xylene or benzyl bromide was dropped into the reservoir in the centre of a specially designed culture flask. In the case of induction of cells grown on solid media, *m*-xylene (50 μl) was pipetted in a 200 μl plastic tip. The tip was fixed in the centre of the lid of a Petri dish with adhesive tape. Volatile reagents of superior purity (> 99%) were purchased from Sigma Aldrich. Strains *E. coli* CC118 *Pu-lacZ* (pCON916) and *E. coli* CC118 *Pu-lacZ* (pCON1238) were freshly prepared by transforming chemocompetent *E. coli* cells (Sambrook *et al*., 1989) with respective plasmids.

β-Galactosidase and bioluminescence assays

For determination of β-galactosidase activity levels, *E. coli* strains CC118 *Pu-lacZ* (pCON916), CC118 *Pu-lacZ* (pCON1238), or MC4100[MAD2] were grown overnight in LB medium, in the morning diluted to OD_{600} of 0.05 in fresh LB and cultured till OD_{600} of 1.0. At this point, cells were exposed to saturated vapours of *m*-xylene (50 μl), benzyl bromide (50 μl), or 1:1 mixture of both chemicals (25 μl each) in conditions described in figure legends. After a 4 h interval, LacZ activity levels were

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

determined in cells permeabilized by chloroform and sodium dodecyl sulfate using the method of Miller (1972). Alternatively, cells were first exposed to one of the two chemicals (25 μl) for 2 h and then the second compound was added (25 μl) for the remaining 2 h. Bioluminescence assays with biosensor strain *P. putida* BXPu-LUX were performed accordingly. After 4 h exposure to aromatic effector(s), culture aliquots (200 μl) were placed in 96-well plate (NUNC) and emission of light was measured with Victor II 1420 Multilabel Counter (PerkinElmer). Recorded bioluminescence was normalized by the optical density of cells in each well. A double reciprocal plot (Fig. 6D) was constructed from values of used *m*-xylene volumes and velocities of bioluminescence formation in the time interval between 30 and 50 min for assay without benzyl bromide (Fig. 6B) and between 10 and 30 min for assay with benzyl bromide (Fig. 6C). Exposure to the volatile chemicals had no negative effect on the viability of the cells.

Purification of XylR-His

E. coli BL21(DE3) pLysS cells were transformed with pCON1238 plasmid bearing *xylR-His* gene, plated on LB agar plate with Ap and grown at 30°C. Following day, cells from the plate were collected in liquid LB with Ap and grown at 30°C (170 rpm) till reaching OD_{600} of 3.0. Cells were then diluted to OD_{600} of 0.15 in fresh LB with Ap and grown at 19°C. Expression of *xylR-His* was induced with 0.4 mM IPTG at the OD_{600} of 0.7 and culture continued at 19°C till reaching OD_{600} of 2.0. Cells were then centrifuged, re-suspended in buffer A (20 mM sodium phosphate, 1.0 M NaCl, 0.1% Triton X-100, 10% glycerol, 1 mM β-mercaptoethanol, pH 7.2) and frozen at −80°C. Melted cells, kept on ice, were lysed by sonication and the crude extract was separated from cell ballast by centrifugation. The crude extract was filtered through a 0.22 μm membrane filter (Merck Millipore) and loaded on a disposable purification column packed with TALON Superflow Resin (Clontech). The resin was washed with buffer A with an increasing concentration of imidazole (0–200 mM). Samples taken from individual fractions were analysed on denaturing SDS polyacrylamide gels (8%) stained with Coomassie Brilliant Blue (Bio-Rad). Fractions containing pure (> 90%) XylR-His protein (64.6 kDa) were pooled, concentrated in Amicon Ultra centrifugal filter (Merck Millipore), and dialysed against buffer B (20 mM sodium phosphate, 0.5 M NaCl, 0.1% Triton X-100, 30% glycerol, 1 mM β-mercaptoethanol, pH 7.5). Protein aliquots of concentration of 0.55 mg ml^{-1} were stored at −80°C for further use.

Exposure of purified XylR-His to benzyl bromide, mass spectrometry analysis of modified TF

Purified XylR-His (6.96 μM) in buffer B was mixed with benzyl bromide (74 mM) in a total volume of 20 μl in a microtube. The protein was left to react with the aromatic chemical overnight at room temperature with modest agitation. The whole volume of the reaction mixture was then loaded on denaturing SDS polyacrylamide gel (8%) and the XylR-His band was cut for mass spectrometry analysis. The mass spectrometry analysis of the peptides obtained after trypsin digest of the protein sample was performed by Proteomics Core Facility of the Spanish National Centre for Biotechnology (CNB-CSIC) in Madrid. Peptides were first separated by liquid chromatography using Ultimate 3000 nano LC system (Dionex) equipped with 75 mm I.D 100 mm reversed-phase column (300 nl min^{-1} flow) and then analysed using mass spectrometers 4800 MALDI TOF/TOF (Applied Biosystems) or HCT Ultra Ion-Trap (Bruker Daltonics) working in dynamic exclusion mode. These two sources of experimental evidence were complementary. In some cases, multiple reaction monitoring mode was used to isolate and fragment specific *m/z* values corresponding to putative benzyl bromide-labelled peptides. For protein identification, LC-ESI-MS/MS spectra were transferred to BioTools 2.0 interface (Bruker Daltonics) to search in the Uniprot database using a licensed version of Mascot v.2.2.04 search engine (Matrix Science). Search parameters were set as follows: carbamidomethyl cysteine as fixed modification by the treatment with iodoacetamide, oxidized methionines and benzylation as variable modifications, peptide mass tolerance of 0.5 Da for the parental mass and fragment masses and one missed cleavage site. In all protein identifications, the probability mowse scores were greater than the minimum score fixed as significant with a *P*-value minor than 0.05. Peptides obtained by fragmentation of non-modified XylR-His were used as a control. Analysis of signal intensities allowed the identification of peptides with mass increments of +90 or +91 Da caused by benzylation of amino acids with nucleophilic functional groups.

Multiple sequence alignment and prediction of XylR A domain structure by molecular threading

The amino acid sequences of transcriptional regulatory proteins XylR from *Pseudomonas putida* (UniProt ID: P06519), PoxR from *Ralstonia sp*. E2 (UniProt ID: O84957), MopR from *Acinetobacter guillouiae* (UniProt ID: Q43965) and DmpR (also known as CapR) from *P. putida* (UniProt ID: Q7WSM9) were retrieved in FASTA format and 211 amino acids of the respective ligand recognition domains were selected for multiple

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

sequence alignment ClustalW (Chenna *et al*., 2003) and prediction of secondary structure elements using and ESPript 3.0 (Robert and Gouet, 2014). The 211 amino acid sequence of XylR A domain was subsequently used for molecular threading by I-TASSER server (Iterative Threading ASSEmbly Refinement; Yang and Zhang, 2015) which represents a hierarchical approach for prediction of protein structure and function. I-TASSER has been repeatedly ranked as a top server in Community Wide Experiment on the Critical Assessment of Techniques for Protein 3D Structure Prediction (http://www.predictioncenter.org) and can be thus considered a reliable accurate tool for the given purpose. The structural model of XylR A domain was built based on multiple-threading alignments by LOMETS performed with templates from PDB and iterative TASSER assembly simulations. The model with the highest confidence (C) score was selected for further work.

Prediction of binding pocket and tunnels in the model of XylR A domain

Ligand binding site in the structural model of XylR A domain was first predicted by COFACTOR and COACH approaches on I-TASSER server (Yang and Zhang, 2015). The set of the residues forming the binding pocket was deduced from the top-ranked PDB of the homologous phenol-responsive sensory domain of PoxR (PDB ID: 5FRU). The binding pocket in XylR A domain PDB file prepared in PyMOL 1.6.0.0 (Schrödinger) was then predicted also using CAVER web 1.0 (Stourac *et al*., 2019). Pocket with the highest druggability score in CAVER (0.95 ± 0.00) corresponded to the one previously proposed by I-TASSER. CAVER web was also used to calculate the volume of the binding pocket in wild-type XylR A domain and in V124A mutant. The calculation was repeated three times for each structure and the mean values of pocket volume are presented with standard deviations. The data were treated with a two-tailed Student's *t* test in Microsoft Office Excel 2013 (Microsoft Corp., USA) and confidence intervals were calculated. Pocket residues suggested by both I-TASSER and CAVER and manually verified in the modelled structure (F93, G96, P97, Y100, V108, V124, A126, W128, Y155, A156, Y159, F170 and I185) are used in this work. Tunnels and bottleneck residues were calculated by CAVER using default program parameters (minimum probe radius 0.9 Å).

Molecular docking of aromatic ligands in XylR A domain model

The modelled structure of XylR A domain was prepared for molecular docking in PyMOL. The structure with added hydrogens was saved as PDB file and uploaded together with a ligand molecule in MOL2 format. The geometries of ligands were optimized using Avogadro 1.2. The binding pocket region of XylR was selected for the docking performed by PyMOL Autodock Vina Plugin for Windows 2.2 (Trott and Olson, 2010). The binding pocket region was defined by a grid box of $40 \times 40 \times 40$ Å. Ligand poses with the lowest binding energies were saved and visualized using PyMOL. The docking experiment was repeated three times for each ligand. The mean values of binding energies of five top-ranked orientations calculated for each of the tested molecules and corresponding standard deviations (SD) are presented. The data were treated with a two-tailed Student's *t* test in Microsoft Office Excel 2013 (Microsoft Corp., USA), and confidence intervals were calculated.

Acknowledgements

Authors are indebted to the Severo Ochoa Programme of Centres of Excellence. This work was funded by the SETH (RTI2018-095584-B-C42) (MINECO/FEDER), SyCoLiM (ERA-COBIOTECH 2018 - PCI2019-111859-2) Project of the Spanish Ministry of Science and Innovation. MADONNA (H2020-FET-OPEN-RIA-2017-1-766975), BioRoboost (H2020-NMBP-BIO-CSA-2018-820699), SynBio4Flav (H2020-NMBP-TR-IND/H2020-NMBP-BIO-2018-814650) and MIX-UP (MIX-UP H2020-BIO-CN-2019-870294) Contracts of the European Union and the InGEMICS-CM (S2017/BMD-3691) Project of the Comunidad de Madrid - European Structural and Investment Funds - (FSE, FECER). The grant 19-06511Y of the Czech Science Foundation to P.D. is also gratefully acknowledged.

References

Abril, M.A., Michan, C., Timmis, K.N., and Ramos, J.L. (1989) Regulator and enzyme specificities of the TOL plasmid-encoded upper pathway for degradation of aromatic hydrocarbons and expansion of the substrate range of the pathway. *J Bacteriol* **171**: 6782–6790.

Bischoff, R., and Schlüter, H. (2012) Amino acids: chemistry, functionality and selected non-enzymatic post-translational modifications. *J Proteomics* **75**: 2275–2296.

Bush, M., and Dixon, R. (2012) The role of bacterial enhancer binding proteins as specialized activators of σ54-dependent transcription. *Microbiol Mol Biol Rev MMBR* **76**: 497–529.

Chenna, R., Sugawara, H., Koike, T., Lopez, R., Gibson, T. J., Higgins, D.G., and Thompson, J.D. (2003) Multiple sequence alignment with the clustal series of programs. *Nucleic Acids Res* **31**: 3497–3500.

de las Heras, A., Carreño, C.A., and de Lorenzo, V. (2008) Stable implantation of orthogonal sensor circuits in Gram-negative bacteria for environmental release. *Environ Microbiol* **10**: 3305–3316.

de las Heras, A., and de Lorenzo, V. (2011) Cooperative amino acid changes shift the response of the σ^{54}-

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

dependent regulator XylR from natural *m*-xylene towards xenobiotic 2,4-dinitrotoluene. *Mol Microbiol* **79**: 1248–1259.

de Lorenzo, V., Herrero, M., Metzke, M., and Timmis, K.N. (1991) An upstream XylR- and IHF-induced nucleoprotein complex regulates the sigma 54-dependent Pu promoter of TOL plasmid. *EMBO J* **10**: 1159–1167.

Delgado, A., and Ramos, J.L. (1994) Genetic evidence for activation of the positive transcriptional regulator Xy1R, a member of the NtrC family of regulators, by effector binding. *J Biol Chem* **269**: 8059–8062.

Delgado, A., Salto, R., Marqués, S., and Ramos, J.L. (1995) Single amino acids changes in the signal receptor domain of XylR resulted in mutants that stimulate transcription in the absence of effectors. *J Biol Chem* **270**: 5144–5150.

Devos, D., Garmendia, J., de Lorenzo, V., and Valencia, A. (2002) Deciphering the action of aromatic effectors on the prokaryotic enhancer-binding protein XylR: a structural model of its N-terminal domain. *Environ Microbiol* **4**: 29–41.

Galvão, T.C., and de Lorenzo, V. (2006) Transcriptional regulators à la carte: engineering new effector specificities in bacterial regulatory proteins. *Curr Opin Biotechnol* **17**: 34–42.

Galvão, T.C., Mencía, M., and de Lorenzo, V. (2007) Emergence of novel functions in transcriptional regulators by regression to stem protein types. *Mol Microbiol* **65**: 907–919.

Garmendia, J., de las Heras, A., Galvão, T.C., and De Lorenzo, V. (2008) Tracing explosives in soil with transcriptional regulators of *Pseudomonas putida* evolved for responding to nitrotoluenes. *J Microbial Biotechnol* **1**: 236–246.

Garmendia, J., Devos, D., Valencia, A., and de Lorenzo, V. (2001) A la carte transcriptional regulators: unlocking responses of the prokaryotic enhancer-binding protein XylR to non-natural effectors. *Mol Microbiol* **42**: 47–59.

Gingras, A.-C., Gstaiger, M., Raught, B., and Aebersold, R. (2007) Analysis of protein complexes using mass spectrometry. *Nat Rev Mol Cell Biol* **8**: 645–654.

Huang, W.E., Singer, A.C., Spiers, A.J., Preston, G.M., and Whiteley, A.S. (2008) Characterizing the regulation of the *Pu* promoter in *Acinetobacter baylyi* ADP1. *Environ Microbiol* **10**: 1668–1680.

Kim, M.N., Park, H.V., Lim, W.K., and Shin, H.J. (2005) Construction and comparison of *Escherichia coli* whole-cell biosensors capable of detecting aromatic compounds. *J Microbiol Methods* **60**: 235–245.

Klausen, M.S., Jespersen, M.C., Nielsen, H., Jensen, K.K., Jurtz, V.I., Sønderby, C.K., *et al*. (2019) NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning. *Proteins Struct Funct Bioinform* **87**: 520–527.

Klvana, M., Pavlova, M., Koudelakova, T., Chaloupkova, R., Dvorak, P., Prokop, Z., *et al*. (2009) Pathways and mechanisms for product release in the engineered haloalkane dehalogenases explored using classical and random acceleration molecular dynamics simulations. *J Mol Biol* **392**: 1339–1356.

Laitaoja, M., Valjakka, J., and Jänis, J. (2013) Zinc coordination spheres in protein structures. *Inorg Chem* **52**: 10983–10991.

Lang, S., Spratt, D.E., Guillemette, J.G., and Palmer, M. (2006) Selective labeling of selenomethionine residues in proteins with a fluorescent derivative of benzyl bromide. *Anal Biochem* **359**: 253–258.

Li, X., Wang, J., Bai, N., Zhang, X., Han, X., da Silva, I., *et al*. (2020) Refinement of pore size at sub-angstrom precision in robust metal–organic frameworks for separation of xylenes. *Nat Commun* **11**: 4280.

Mattson, G., Conklin, E., Desai, S., Nielander, G., Savage, M.D., and Morgensen, S. (1993) A practical approach to crosslinking. *Mol Biol Rep* **17**: 167–183.

Miller, J.H. (1972) Experiments in molecular genetics, [Cold Spring Harbor, N.Y.]: Cold Spring Harbor Laboratory.

North, A.K., Klose, K.E., Stedman, K.M., and Kustu, S. (1993) Prokaryotic enhancer-binding proteins reflect eukaryote-like modularity: the puzzle of nitrogen regulatory protein C. *J Bacteriol* **175**: 4267–4273.

Park, K.-H., Kim, S., Lee, S.-J., Cho, J.-E., Patil, V.V., Dumbrepatil, A.B., *et al*. (2020) Tetrameric architecture of an active phenol-bound form of the AAA + transcriptional regulator DmpR. *Nat Commun* **11**: 2728.

Patil, V.V., Park, K.-H., Lee, S.-G., and Woo, E. (2016) Structural analysis of the phenol-responsive sensory domain of the transcription activator PoxR. *Struct Lond Engl 1993* **24**: 624–630.

Pérez-Martin, J., Cases, I., and de Lorenzo, V. (1997) Design of a solubilization pathway for recombinant polypeptides in vivo through processing of a bi-protein with a viral protease. *Protein Eng* **10**: 725–730.

Pérez-Martín, J., and De Lorenzo, V. (1995) The amino-terminal domain of the prokaryotic enhancer-binding protein XylR is a specific intramolecular repressor. *Proc Natl Acad Sci U S A* **92**: 9392–9396.

Pérez-Martín, J., and de Lorenzo, V. (1996) ATP binding to the sigma 54-dependent activator XylR triggers a protein multimerization cycle catalyzed by UAS DNA. *Cell* **86**: 331–339.

Ray, S., Gunzburg, M.J., Wilce, M., Panjikar, S., and Anand, R. (2016) Structural basis of selective aromatic pollutant sensing by the effector binding domain of MopR, an NtrC family transcriptional regulator. *ACS Chem Biol* **11**: 2357–2365.

Ray, S., Panjikar, S., and Anand, R. (2018) Design of protein-based biosensors for selective detection of benzene groups of pollutants. *ACS Sens* **3**: 1632–1638.

Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res* **42**: W320–W324.

Rogers, G.A., Shaltiel, N., and Boyer, P.D. (1976) Facile alkylation of methionine by benzyl bromide and demonstration of fumarase inactivation accompanied by alkylation of a methionine residue. *J Biol Chem* **251**: 5711–5717.

Salto, R., Delgado, A., Michán, C., Marqués, S., and Ramos, J.L. (1998) Modulation of the function of the signal receptor domain of XylR, a member of a family of prokaryotic enhancer-like positive regulators. *J Bacteriol* **180**: 600–604.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor, N. Y: Cold Spring Harbor Laboratory.

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,

Shingler, V. (2003) Integrated regulation in response to aromatic compounds: from signal sensing to attractive behaviour. *Environ Microbiol* **5**: 1226–1241.

Shingler, V., and Moore, T. (1994) Sensing of aromatic compounds by the DmpR transcriptional activator of phenol-catabolizing *Pseudomonas sp*. strain CF600. *J Bacteriol* **176**: 1555–1560.

Shingler, V., and Pavel, H. (1995) Direct regulation of the ATPase activity of the transcriptional activator DmpR by aromatic compounds. *Mol Microbiol* **17**: 505–513.

Stourac, J., Vavra, O., Kokkonen, P., Filipovic, J., Pinto, G., Brezovsky, J., *et al*. (2019) Caver Web 1.0: identification of tunnels and channels in proteins and analysis of ligand transport. *Nucleic Acids Res* **47**: W414–W422.

Trott, O., and Olson, A.J. (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J Comput Chem* **31**: 455–461.

Vidgren, J., Svensson, L.A., and Liljas, A. (1994) Crystal structure of catechol O-methyltransferase. *Nature* **368**: 354–358.

Weiss, V., Claverie-Martin, F., and Magasanik, B. (1992) Phosphorylation of nitrogen regulator I of *Escherichia coli* induces strong cooperative binding to DNA essential for activation of transcription. *Proc Natl Acad Sci U S A* **89**: 5088–5092.

Winson, M.K., Swift, S., Hill, P.J., Sims, C.M., Griesmayr, G., Bycroft, B.W., *et al*. (1998) Engineering the luxCDABE genes from *Photorhabdus luminescens* to provide a bioluminescent reporter for constitutive and promoter probe plasmids and mini-Tn5 constructs. *FEMS Microbiol Lett* **163**: 193–202.

Yang, J., and Zhang, Y. (2015) I-TASSER server: new development for protein structure and function predictions. *Nucleic Acids Res* **43**: W174–W181.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1: Table S1 Strains and plasmids used in this study.

Fig. S1 Original structural model of XylR A domain proposed by Devos and colleagues (20022002).

Fig. S2 Multiple sequence alignment of effector binding domain A of XylR, PoxR, MopR and DmpR from *Pseudomonas putida* mt-2, *Ralstonia sp*. E2, *Acinetobacter guillouiae* and *P. putida* KCTC 1452 respectively

Fig. S3 Docking of ligands in predicted binding pocket of XylR A domain.

Fig. S4 The effects of mutations V124A and Y159F in the modelled structure of XylR A domain.

Fig. S5 *In vivo* evidence of XylR-His interaction with *m*-xylene and benzyl bromide.

Fig. S6 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (8% gel) of purified XylR-His protein.

Fig. S7 Surface accessibility of amino acids in XylR A domain predicted by NetSurfP 2.0 server.

Appendix S2: Mass spectrometry analyses for identification of XylR peptides with nucleophilic amino acids cross-linked with benzyl bromide.

© 2021 The Authors. *Environmental Microbiology* published by Society for Applied Microbiology and John Wiley & Sons Ltd.,