Novel carboxylic ester hydrolases from marine hydrocarbonoclastic bacteria – insights into organic solvent tolerance, substrate promiscuity and polyester hydrolysis

Inaugural-Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Heinrich-Heine-Universität Düsseldorf

Vorgelegt von

Alexander Bollinger

aus Köln

Düsseldorf, März 2020

aus dem Institut für Molekulare Enzymtechnologie
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der

Mathematisch-Naturwissenschaftlichen Fakultät der

Heinrich-Heine-Universität Düsseldorf

Referent:	Prof. Dr. Karl-Erich Jaeger
Korreferent:	Prof. Dr. Michael Feldbrügge
Tag der mündlichen Prüfung:	10.08.2020

I. Abstract

Biotechnological application of enzymes is an important tool for the transformation of chemical processes toward a bio-based economy. Therefore, the number of newly identified enzymes is growing each year, however this is not appropriately reflected by their industrial application. The low number of enzymes used in industrial processes compared to the available number of enzyme sequences is mainly due to insufficient characterization and performance of the biocatalysts under process conditions.

Particular bottlenecks for the effective and efficient application of enzymes are for example (I) insufficient stability of the biocatalysts under process conditions, which often means in the presence of organic solvents, (II) a narrow range of use due to high substrate specificity, and (III) no activity with synthetic polymeric substrates. To cope with these challenges not only the identification of candidate enzymes but the molecular characterization for respective features is needed.

In this thesis, the portfolio of available biocatalysts of the carboxylic ester hydrolase family was broadened by the identification of 25 enzymes (CE02 to CE26), most of them completely new, from two marine hydrocarbonoclastic bacteria, *Alcanivorax borkumensis* SK2, and *Pseudomonas aestusnigri* VGXO14. These enzymes were systematically assayed for relevant characteristics: substrate promiscuity, organic solvent tolerance, and hydrolysis of synthetic polyester substrates. For each characteristic, at least one outstanding example was found, and the molecular mechanism of the respective feature was investigated.

The highest substrate promiscuity was found for CE07 from *A. borkumensis*. In a comparative analysis of more than 100 diverse ester hydrolases, CE07 ranked among the 10 most substrate promiscuous enzymes. Furthermore, the enzyme accepts complex water-insoluble esters as common in pharmaceutical biotechnology. Interestingly, whereas prominent substrate promiscuity was found to be negatively correlated with enantioselectivity for most carboxylic ester hydrolases, this is not true for CE07, which is selective for two out of ten chiral esters. Molecular docking computation with a homology model of CE07 and (*R*/*S*)-Menthyl acetate as ligand suggests, that the observed selectivity is rather based on differences in the catalytic rate of the reaction than on preferred binding of one enantiomer.

To assess the organic solvent tolerance, a novel assay was devised and established which allowed to determine this characteristic directly from whole cell lysate, facilitating high throughput enzyme characterization. This way, the novel biocatalyst CE13 from *P. aestusnigri* was identified to be highly organic solvent tolerant, retaining substantial activity after incubation in 80 % acetonitrile for several hours. The molecular basis of the organic solvent tolerance of CE13 is connected to a large number of charged amino acid residues and stabilizing disulfide bridges, whereas it is not associated with high thermal stability.

Furthermore, a screening strategy to identify polyester hydrolytic enzymes was established using the anionic aliphatic polyester-polyurethane Impranil DLN as a substrate. Of all enzymes tested, one biocatalyst originating from *P. aestusnigri* (CE16, also named PE-H) was found to show distinct polyester hydrolytic activity. Moreover, biochemical characterization of CE16 revealed evidence for amorphous PET film hydrolysis at 30 °C, which was improved by rational mutagenesis to allow activity with PET from a commercial single use bottle. The crystal structure of CE16 and of the improved variant was solved at high resolution (1.09 Å and 1.35 Å respectively) representing the first crystal structure of a type IIa PET hydrolytic enzyme. The enzyme structures were subsequently used to rationalize the enhanced activity by a more accessible active site and suggest the molecular mechanism of polymer binding by molecular docking computations.

With the knowledge gained on substrate promiscuity, organic solvent tolerance and polyester hydrolysis, the possibility to combine multiple features in a single biocatalyst was evaluated. While substrate promiscuity and organic solvent tolerance were found to complement well, molecular mechanisms for polyester hydrolysis, like a surface exposed and flexible active site, appear contradicting to substrate promiscuity and organic solvent tolerance. Hence, one biocatalyst to comply with all demands remains utopic, but an approximation to this future biocatalyst is conceivable.

In conclusion, the presented thesis took part in deepening the understanding of highly important molecular characteristics for industrial application of biocatalysts, gave rise to a set of new and comprehensively characterized carboxylic ester hydrolases, and provided hints for the design of a next generation biocatalyst, which combines most of the respective features.

II. Zusammenfassung

Die biotechnologische Anwendung von Enzymen ist ein wichtiges Werkzeug für den Wandel von klassischen chemischen Prozessen hin zu einer bio-basierten Ökonomie. In Folge dessen nimmt die Zahl der jährlich neuentdeckten Enzyme stetig zu, was jedoch nicht direkt durch eine im gleichen Maße gesteigerte industrielle Anwendung wiedergespiegelt wird. Die geringe Anzahl der industriell genutzten Enzyme im Vergleich zu der hohen Anzahl an verfügbaren Enzymsequenzen liegt zu einem guten Teil an einer unzureichenden Charakterisierung und folglich auch Leistung unter Prozessbedingungen.

Zu den Gründen, die einer effektiven und effizienten Anwendung von Enzymen entgegenstehen, gehören unter anderem (I) eine unzureichende Stabilität der Biokatalysatoren unter Prozessbedingungen, oftmals bedingt durch die Nutzung organischer Lösungsmittel, (II) eine geringe Bandbreite an Einsatzmöglichkeiten für ein spezifisches Enzym, aufgrund dessen hoher Substratspezifität und (III) eine geringe Aktivität des Enzyms mit synthetischen, polymeren Substraten. Um diese Herausforderungen zu bewältigen ist es nicht ausreichend, neue Enzymkandidaten zu identifizieren, sondern vielmehr nötig eine detaillierte molekulare Charakterisierung hinsichtlich der relevanten Eigenschaften vorzunehmen.

Im Rahmen dieser Arbeit wurden 25 (CE02 bis CE26) zuvor teils unbekannte Carboxylester-Hydrolasen aus zwei marinen hydrocarbonoklastischen Bakterien, *Alcanivorax borkumensis* SK2 und *Pseudomonas aestusnigri* VGXO14, identifiziert. Darüber hinaus wurden die Enzyme hinsichtlich ihrer Substratpromiskuität, Toleranz für organische Lösungsmittel und Fähigkeit zur Polyesterhydrolyse systematisch untersucht. Für jede dieser Charakteristika wurde ein herausragendes Beispiel gefunden und die zugrundeliegenden molekularen Eigenschaften näher untersucht.

Die höchste Substratpromiskuität zeigte das Enzym CE07 aus *A. borkumensis*. In einer vergleichenden Studie mit mehr als 100 verschiedenen Ester Hydrolasen konnte CE07 unter den 10 Enzymen mit der größten Substratpromiskuität eingeordnet werden. Darüber hinaus ist das Enzym in der Lage besonders komplexe, nicht wasserlösliche Ester, wie sie in der pharmazeutischen Biokatalyse häufig vorkommen, als Substrate anzunehmen. Interessanterweise zeigte CE07 dabei nicht die für hoch promiske Carboxylester-Hydrolasen typische Abwesenheit von Enantioselektivität,

sondern eine deutliche Selektivität für zwei von zehn getesteten chiralen Substraten. Mittels *molecular-docking* Berechnungen für ein Homologiemodell von CE07 mit (*R*/*S*)-Menthylacetat als Ligand konnte die Vermutung unterstützt werden, dass die beobachtete Selektivität durch unterschiedliche Reaktionsraten und nicht durch bevorzugte Bindung eines Enantiomers bedingt ist.

Zur Bestimmung der Lösungsmitteltoleranz wurde ein neues hochdurchsatzfähiges Verfahren entwickelt und angewandt, welches direkte Messungen mit Zelllysat ermöglichte. So wurde das neue Enzym CE13 aus *P. aestusnigri* als hochgradig tolerant gegenüber verschiedenen Lösungsmitteln identifiziert; es zeigte selbst nach mehrstündiger Inkubation in 80 % Acetonitril deutliche Restaktivität. Der zugrundeliegende molekulare Mechanismus der Lösungsmittelstabilität von CE13 konnte auf eine hohe Anzahl geladener Aminosäurereste, sowie mehrere Disulfidbrücken zurückgeführt werden, war jedoch nicht mit einer hohen Temperaturstabilität des Enzyms assoziiert.

Des Weiteren wurde eine Strategie zum *Screening* polyesterhydrolytischer Enzyme etabliert, die auf der Verwendung des anionischen aliphatischen Polyester-Polyurethan Impranil DLN als Substrat beruht. Unter allen getesteten Enzymen konnte eines gefunden werden, das deutliche Aktivität mit diesem Substrat zeigte, die Polyesterhydrolase CE16 (ebenfalls PE-H genannt) aus *P. aestusnigri*. Durch eine nähere biochemische Charakterisierung konnte die Hydrolyse von amorpher PET Folie bei 30 °C nachgewiesen werden. Im Weiteren konnte die Aktivität des Enzyms, mittels einer gezielten Mutagenese, auf PET Folie einer Einwegflasche als Substrat erweitert werden. Die Kristallstruktur des Enzyms und der verbesserten Variante konnte in hoher Auflösung aufgeklärt werden (1.09 Å und 1.35 Å) und wurde als die erste Struktur eines PET hydrolysierenden Enzyms des Typs IIa identifiziert. Basierend auf diesen Enzymstrukturen konnte sowohl die Verbesserung der Aktivität durch einen deutlich besseren Zugang zum aktiven Zentrum des Enzyms begründet, als auch der Bindemechanismus des Polymers nachvollzogen werden.

Durch die neugewonnenen Erkenntnisse zur Substratpromiskuität, Lösungsmitteltoleranz und Aktivität gegenüber Polyestern konnte die Möglichkeit evaluiert werden, mehrere dieser Eigenschaften in einem Biokatalysator zu vereinen. Hierbei zeigte sich, dass sich die Eigenschaften Substratpromiskuität und Lösungsmitteltoleranz wahrscheinlich gut vereinen lassen, jedoch molekulare

Grundlagen der Polyesterhydrolyse, wie die Positionierung und die Flexibilität des aktiven Zentrums, nicht mit den anderen Charakteristika vereinbar sind. Folglich ist der Versuch, alle Eigenschaften in einem Biokatalysator zu vereinen, utopisch, jedoch ist eine Annäherung an solch einen Biokatalysator in der Zukunft denkbar.

Die vorliegende Arbeit trägt somit zu einem tieferen Verständnis von für die industrielle Anwendung relevanten molekularen Eigenschaften von Enzymen aus der Klasse der Carboxylester-Hydrolasen bei, vergrößert die Anzahl der bekannten und charakterisierten Enzyme dieser Klasse und liefert Hinweise für die Entwicklung zukünftiger Biokatalysatoren, die möglichst viele der relevanten Eigenschaften auf sich vereinigen.

III. Table of contents

IV. Abbreviations

(v/v)	volume per volume
(w/v)	weight per volume
3D	three dimensional
ABC	ATP binding cassette
ATP	adenosine triphosphate
BHET	bis(2-hydroxyethyl) terephthalate
BY	by attribution
CC	creative commons
CEH / CEHs	carboxylic ester hydrolase / carboxylic ester hydrolases
CTAC	cetrimonium chloride
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
e. g.	*exempli gratia* (Latin): for example
ed / eds	editor / editors
et al.	*et alia* (Latin): and others
FASTA	a file format for DNA and protein sequences
i. e.	*id est* (Latin): in other words
IPTG	Isopropyl-β-D-1-thiogalactopyranoside
IT	information technology
kbp	kilo base pairs
k_{cat}	enzyme turnover number
K_m	Michaelis-Menten constant
LB	Luria-Bertani
Mbp	Mega base pairs
MHET	mono(2-hydroxyethyl) terephthalate
NC	non-commercial
PCR	polymerase chain reaction
PDB	protein data bank
PET	polyethylene terephthalate
rmsd	root-mean-square deviation
SDS	sodium dodecyl sulfate
Sec	general secretion (pathway)

Tat	twin-arginine translocation (pathway)

V. List of publications

Conference contributions

Molitor R, Bollinger A, Turkes L, Thies S, Jaeger K-E (2020) Identification of novel polyester hydrolases from marine Pseudomonads. *Annual Meeting 2020 of the Association for General and Applied Microbiology (VAAM)*. Leipzig, Germany 2020-03-08-11 (poster presentation)

Kubicki S, Bator I, Bollinger A, Molitor R, Linden S, Tiso T, Loeschcke A, Jaeger K-E, Thies S (2019). Recombinant bacterial cell factories for discovery, tailoring, and production of biosurfactants. *2nd Synthetic biology for natural products conference*. Puerto Vallarta, Mexico 2019-06-02-05 (oral presentation)

Thies S, Bollinger A, Molitor R, Jaeger K-E (2018). Finding Enzymes – straightforward strategies to assess hydrolytic capabilities of *Pseudomonas* sp. *Pseudomonas grassroots meeting.* Frankfurt a. M. Germany, 2018-11-08-09 (poster presentation)

Bollinger A, Thies S, Coscolín C, Katzke N, Söltl S, Navarro-Fernández J, Golyshin P, Ferrer M, Jaeger K-E (2018). Marine bacteria as a prolific source for novel biocatalysts, *CLIB International Conference.* Düsseldorf Germany, 2018-01-17-18 (poster presentation)

Bollinger A, Thies S, Raauf M, Kuklinski A, Coscolín C, Ferrer M, Jaeger K-E (2018). Exploring marine *Pseudomonas* species for polyester hydrolases. *9th International Congress on Biocatalysis*. Hamburg Germany, 2018-08-26-30 (poster presentation)

Thies S, Bollinger A , Katzke N, Domröse A, Klein A, Loeschcke A, Drepper T,Coscolín C, Ferrer M, Pietruszka J, Jaeger, K-E (2017). Lessons from (meta)genomes – novel biocatalysts and compounds. *7th European Conference on Prokaryotic and Fungal Genomics ProkaGenomics*. Göttingen, Germany 2017-09-22 (poster presentation)

Bollinger A, Thies S, Coscolín C, Katzke N, Söltl S, Navarro-Fernández J, Golyshin P, Ferrer M, Jaeger K-E (2017). Marine bacteria as a prolific source for novel biocatalysts. *FEMS 7th Congress of European Microbiologists*. Valencia 2017-07-09-13 (poster presentation)

Publications

Bollinger A, Molitor R, Thies S, Koch R, Coscolín C, Ferrer M, Jaeger K-E (2020) Identification of organic solvent tolerant carboxylic ester hydrolases for organic synthesis. Appl Environ Microbiol. doi: 10.1128/AEM.00106-20

Bollinger A, Thies S, Knieps-Grünhagen E, Gertzen C, Kobus S, Höppner A, Ferrer M, Gohlke H, Smits SHJ, Jaeger K-E (2020) A novel polyester hydrolase from the

marine bacterium *Pseudomonas aestusnigri* – structural and functional insights. Front. Microbiol. 11:114. doi: 10.3389/fmicb.2020.00114

Bollinger A, Thies S, Katzke N, Jaeger K-E (2020) The biotechnological potential of marine bacteria in the novel lineage of *Pseudomonas pertucinogena*. Microb Biotechnol. doi: 10.1111/1751-7915.13288

Molitor R, Bollinger A, Kubicki S, Loeschcke A, Jaeger K, Thies S (2020) Agar plate-based screening methods for the identification of polyester hydrolysis by *Pseudomonas* species. Microb Biotechnol 1751-7915.13418. doi: 10.1111/1751-7915.13418

Kubicki S, Bollinger A, Katzke N, Jaeger K, Loeschcke A, Thies S (2019) Marine biosurfactants: biosynthesis, structural diversity and biotechnological applications. Mar Drugs 17:1–30 . doi: 10.3390/md17070408

Coscolín C, Katzke N, García-Moyano A, Navarro-Fernández J, Almendral D, Martínez-Martínez M, Bollinger A, Bargiela R, Gertler C, Chernikova TN, Rojo D, Barbas C, Tran H, Golyshina O V., Koch R, Yakimov MM, Bjerga GEK, Golyshin PN, Jaeger K-E, Ferrer M (2019) Bioprospecting reveals class III ω-transaminases converting bulky ketones and environmentally relevant polyamines. Appl Environ Microbiol 85:1–20 . doi: 10.1128/AEM.02404-18

Coscolín C, Martínez-Martínez M, Chow J, Bargiela R, García-Moyano A, Bjerga G, Bollinger A, Stokke R, Steen I, Golyshina O, Yakimov M, Jaeger K-E, Yakunin A, Streit W, Golyshin P, Ferrer M (2018) Relationships between substrate promiscuity and chiral selectivity of esterases from phylogenetically and environmentally diverse microorganisms. Catalysts 8:10 . doi: 10.3390/catal8010010

Coscolín C, Bargiela R, Martínez-Martínez M, Alonso S, Bollinger A, Thies S, Chernikova TN, Hai T, Golyshina O V., Jaeger K-E, Yakimov MM, Golyshin PN, Ferrer M (2018) Hydrocarbon-degrading microbes as sources of new biocatalysts. In: McGenity TJ (ed) Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes. Springer, Cham, Cham, pp 1–21. doi: 10.1007/978-3-319-60053-6_13-1

Poehlein A, Daniel R, Thürmer A, Bollinger A, Thies S, Katzke N, Jaeger K-E (2017) First insights into the genome sequence of *Pseudomonas oleovorans* DSM 1045. Genome Announc 5:e00774-17 . doi: 10.1128/genomeA.00774-17

Martínez-Martínez M, Coscolín C, Santiago G, Chow J, Stogios PJ, Bargiela R, Gertler C, Navarro-Fernández J, Bollinger A, Thies S, Méndez-García C, Popovic A, Brown G, Chernikova TN, García-Moyano A, Bjerga GEKK, Pérez-García P, Hai T, Del Pozo M V., Stokke R, Steen IH, Cui H, Xu X, Nocek BP, Alcaide M, Distaso M, Mesa V, Peláez AI, Sánchez J, Buchholz PCFF, Pleiss J, Fernández-Guerra A, Glöckner FO, Golyshina O V., Yakimov MM, Savchenko A, Jaeger K-EE, Yakunin AF, Streit WR, Golyshin PN, Guallar V, Ferrer M, The INMARE Consortium (2017) Determinants and prediction of esterase substrate promiscuity patterns. ACS Chem Biol 13:acschembio.7b00996 . doi: 10.1021/acschembio.7b00996

Patents

EP3572523A1 „Novel carboxyesterhydrolases“

European patent application, date of filing 25.05.2018

1 Introduction

1.1 Blue biotechnology

The effort to utilize living organisms for their own benefit has characterized humankind ever since. Long before the common era, humans actively started to domesticate animals, grow, select, and evolve plants, and (unconsciously) made use of microorganisms via fermentation of fruits, grain or milk. These efforts are often contributed to biotechnology since they make use of living beings in a controlled manner. The first reported use of the term biotechnology is attributed to Hungarian engineer Karl Ereky 100 years ago (Ereky, 1919; Fári and Kralovánszky, 2006). Today, various definitions of the term biotechnology can be found (Bhatia, 2018), from extensive, describing all technological usage of living organisms in general, to specific meanings, like the application of genetically modified organisms. Biotechnology can be classified into different colors according to the practical field of its application. Hence, red biotechnology takes part in the medical field, yellow biotechnology deals with insects, green biotechnology with plants, white biotechnology with industrial production and blue biotechnology with marine organisms (Figure 1-1).

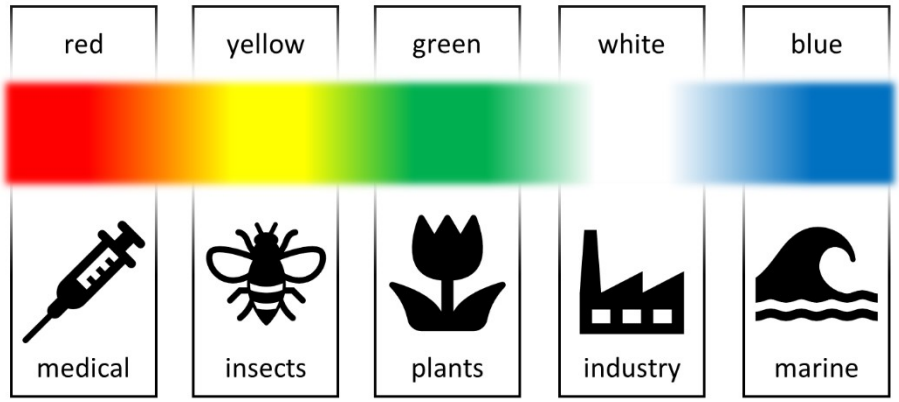


Figure 1-1 The colors of biotechnology according to their field of application.

A few of the most prominent examples from different sectors of biotechnology are (I) the recombinant production of human Insulin (Johnson, 1983) and different antibiotics starting with penicillin (Fleming, 1929) for the treatment of diabetes and infectious diseases, (II) generation of improved plants like golden rice (Beyer *et al.*, 2002) which is meant to fight vitamin A deficiency in developing countries, (III) addition of enzymes

like proteases, lipases, amylases, and cellulases in laundry detergent to enhance cleaning performance as well as for saving water and energy (Olsen and Falholt, 1998), or (IV) isolation of anti-tumor compounds from marine endosymbiotic bacteria (Schofield *et al.*, 2015) which are already approved for therapeutic usage.

Blue biotechnology, also known as marine biotechnology, is a rather young field of biotechnology and deals with marine organisms and products thereof for technical application. It can be defined as “The application of science and technology to living organisms from marine resources, as well as parts, products and models thereof, to alter living or non-living materials for the production of knowledge, goods and services.” according to the Organization for Economic Co-operation and Development (OECD, 2017). This field of biotechnology is especially interesting, because most part of our planet is covered by the oceans that include diverse environments with regions of extreme temperature, as present in the arctic or near volcanic hydrothermal vents, high salinity in underwater brine pools, or extreme pressure at the bottom of a deep-sea trench (Ferrer *et al.*, 2019). As it is for terrestrial habitats, microorganisms and in particular bacteria outnumber every other species in the oceans, with a high chance to harbor compounds humans can take advantage of. However, the marine environment was believed to be a limited source for novel microorganisms for a long time. This changed rapidly with the development of modern sequencing techniques and culture independent identification of microorganisms such as metagenomics. With Craig Venters sequencing project in the Sargasso Sea (Venter *et al.*, 2004) at the latest, the oceans were recognized as a prolific source for diverse microorganisms. This was proven by different bioprospecting studies, greatly broadening the sequence space for marine species and providing functional studies to novel genes and organisms found in the ocean. The Tara Oceans circumvention expedition for example contributed more than 2600 novel draft genome sequences (Tully *et al.*, 2018) and the INMARE H2020 project (http://www.inmare-h2020.eu/) about one thousand active enzymes with one third fully characterized (Ferrer *et al.*, 2019). Hence, blue biotechnology holds a great chance for the discovery of completely new organisms, proteins and compounds with high potential benefit for mankind (Table 1-1).

Table 1-1 Discoveries from the field of blue biotechnology.

description	reference
green fluorescent protein (GFP) from *Aequorea victoria*, used in numerous biotechnological applications	(Shimomura *et al.*, 1962; Shimomura, 2005)
Yondelis (Trabectedin), a chemotherapeutic compound from a tunicate endosymbiotic bacterium	(Schofield *et al.*, 2015)
sorbicillactone A & B, alkaloids from a marine sponge associated fungus, active against leukemia cells	(Bringmann *et al.*, 2005, 2007)
ArcticExpress (DE3), a recombinant *Escherichia coli* BL21(DE3) strain for protein production at low temperatures	(Ferrer *et al.*, 2003)
Ara-A, an antiviral nucleoside from a marine sponge	(Sagar *et al.*, 2010)
cephalosporin C, an example for antibiotics from marine fungi	(Silber *et al.*, 2016)

1.2 Marine hydrocarbonoclastic bacteria

The oceans comprise a seemingly unlimited source of microorganisms to discover, from generalists to niche adapted specialists. Among this broad diversity, highly specialized marine bacteria appear to be a prolific source for similarly highly specialized biocatalysts. The adaptation of a microorganism to specific environmental challenges evolves enzymes with characteristics enabling the survival of the microbe in the particular environment, for example cold and saline marine environments facilitate the evolution of cold active and salt tolerant enzymes (Tchigvintsev *et al.*, 2015). Thus, the rational selection of a niche adapted microbe can help to identify biomolecules with the required biochemical and biophysical characteristics.

The group of hydrocarbonoclastic (crude oil-degrading) bacteria represents such adapted microorganisms, which have evolved to degrade and assimilate a wide range of different alkanes, the major constituents of crude oil (Head *et al.*, 2006; Yakimov *et al.*, 2007; Brooijmans *et al.*, 2009). Alkane degrading bacteria, also known as hydrocarbonoclastic bacteria (HCB), can be further specified according to the range of carbon sources they can feed on; where many HCB can grow on other substrates than crude oil, like sugars and amino acids as well, the group of obligate hydrocarbonoclastic bacteria (OHCB), or obligate marine hydrocarbonoclastic bacteria (OMHCB), preferentially use crude oil or other alkanes as sole carbon source and can use only a limited number of organic acids such as pyruvate, or acetate for growth

(Golyshin *et al.*, 2003). These bacteria are found in low number over almost all marine environments (Head *et al.*, 2006), but in case of an oil spill, the bacteria bloom and rapidly outnumber other species (Yakimov *et al.*, 2007).

The group of OMCHB comprise representatives of the genera *Alcanivorax, Cycloclasticus, Marinobacter, Neptunomonas, Oceanobacter, Oleiphilus, Oleispira,* and *Thalassolituus,* with *Alcanivorax borkumensis* being the most prominent example (Yakimov *et al.*, 2007; Cafaro *et al.*, 2013). This marine γ-proteobacterium was isolated in 1991 near the Isle of Borkum, North Germany (Schulz *et al.*, 1991), and first described as novel species in 1998 (Yakimov *et al.*, 1998). It can degrade and assimilate a large range of linear alkanes up to C32 in length, facilitated by a set of genes essential for alkane degradation (Schneiker *et al.*, 2006). The general aerobic alkane degradation pathway of bacteria is shown in Figure 1-2. The first step for the metabolization of alkanes is catalyzed by an alkane hydroxylase, which activates the hydrocarbons by terminal hydroxylation, this is followed by the action of an alcohol dehydrogenase, an aldehyde dehydrogenase, and an acyl-CoA synthetase to feed the activated hydrocarbon in the β-oxidation pathway. Alternatively, the fatty acid can be further modified by an ω-fatty acid monooxygenase, alcohol dehydrogenase, and aldehyde dehydrogenase to yield a dicarboxylic acid. Besides terminal alkane oxidation a second pathway is known, the subterminal oxidation of alkanes, which can generate fatty acids by the formation of an ester bond within the alkane chain. This pathway comprises an alkane hydroxylase, an alcohol dehydrogenase and a Baeyer-Villiger monooxygenase to produce an ester, which is subsequently hydrolyzed by a carboxylic ester hydrolase. The fatty acid can be directed to β-oxidation or, in case of acetate, to the citric acid cycle and the remaining terminal hydroxylated alkane can be further metabolized via the terminal oxidation pathway (Beilen *et al.*, 2003; Rojo, 2009; Ji *et al.*, 2013).

This genetic repertoire is of high interest from the perspective of blue biotechnology, comprising genes for mono- and dioxygenases, dehydrogenases, and carboxylic ester hydrolases, as well as biosynthesis pathways for biosurfactants and storage lipids (Schneiker *et al.*, 2006; Kubicki *et al.*, 2019). Moreover, their lifestyle might have led to the development of enzymes with features typical for marine enzymes like cold-adaptation, salt and pH tolerance, or barophilicity (Ferrer *et al.*, 2005; Trincone, 2011; Tchigvintsev *et al.*, 2015).

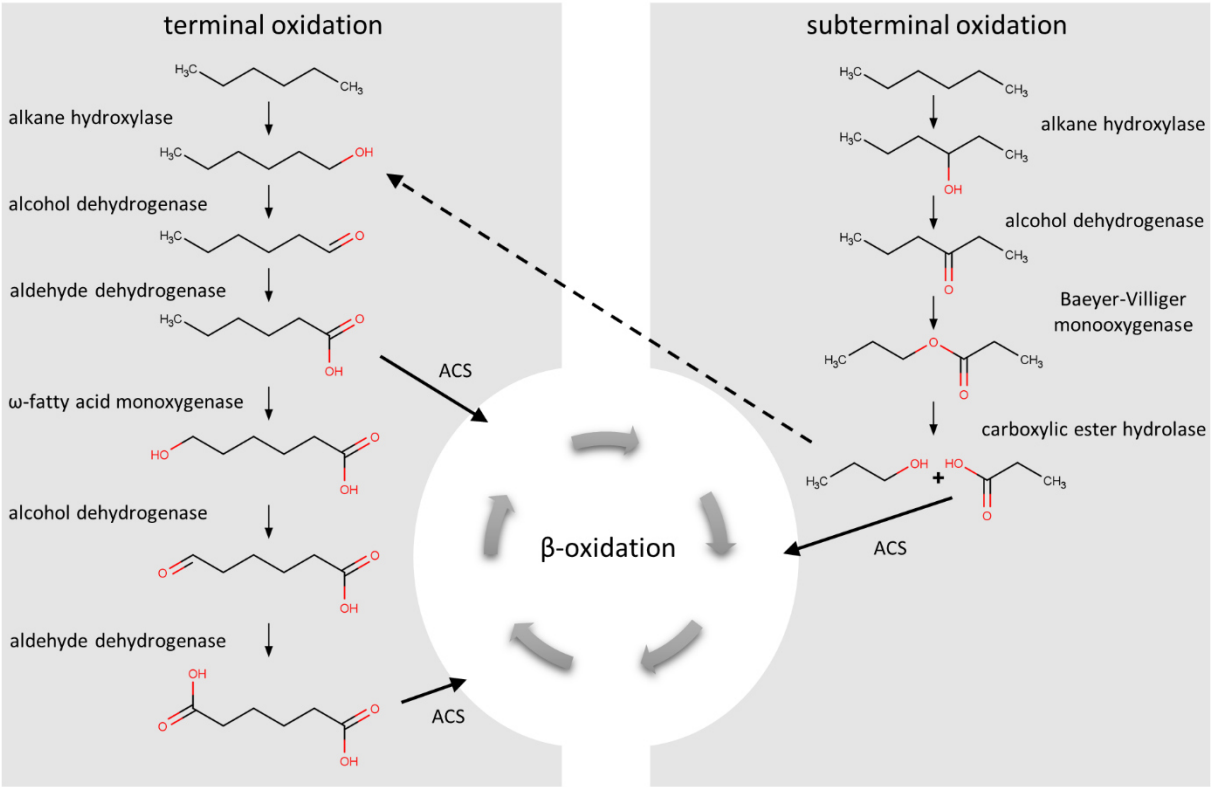


Figure 1-2 Bacterial aerobic alkane oxidation pathways. Oxidative alkane catabolism uses several enzymatic steps to produce fatty acids or diacids via terminal oxidation (left), or subterminal oxidation (right). The products are activated by an acyl-CoA synthetase (ACS) before further use for energy generation by β-oxidation. The pathway is shown for hexane but applies in general for aliphatic alkanes.

1.3 Bioprospecting for novel enzymes

The discovery of novel enzymes was for long restricted to microbial isolates cultivable with standard microbiological methods in the laboratory. However, only a tiny fraction of all prokaryotic organisms, often below 1 % depending on the sample's origin (Amann *et al.*, 1995), can be cultivated this way, thus the majority of biodiversity remained out of reach. This changed with the development of culture independent methodologies, the metagenomics, allowing to directly access the DNA of all organisms from an environmental sample. Since then, metagenomic bioprospecting studies have significantly contributed to the identification of novel enzymes (Madhavan *et al.*, 2017) and secondary metabolites (Lorenz and Eck, 2005; Thies *et al.*, 2016), and have proven their importance in both academia and industry, which is represented by a number of excellent reviews in this field (Streit and Schmitz, 2004; Lorenz and Eck, 2005; Sabree *et al.*, 2009; Ferrer, Martínez-Martínez, *et al.*, 2015).

The discovery of novel enzymes from single organisms or metagenomic sources comprises different steps (Figure 1-3), each with specific challenges to consider. First, DNA is extracted from either genomic or metagenomic sources. Here, the characteristics of the organisms' habitat directly affect the nature of their harbored biocatalysts. For example, marine sampling sites can be characterized by low temperature and high salinity, thus selecting for salt tolerant and cold active enzymes (Tchigvintsev *et al.*, 2015).

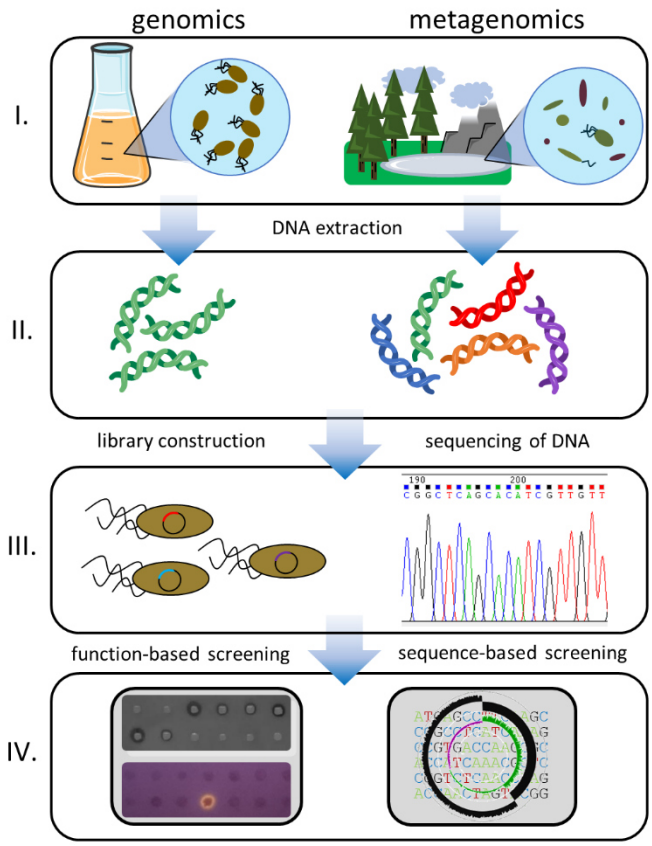


Figure 1-3 Steps in the bioprospecting for novel enzymes. **(I.)** Bioprospecting can use either samples from single organisms (genomics) or the entirety of organisms isolated from a specific spot (metagenomics). **(II.)** Genomic DNA (gDNA) from a single organism or environmental DNA (eDNA) from a metagenomic source is extracted. **(III.)** The DNA can be directly used for sequencing or to construct libraries of DNA fragments in appropriate host organisms. **(IV.)** Libraries can be used for function-based screening to directly assess enzymatic activities (*e. g.* by enzyme activity indicating agar plates) or the DNA sequence information can be mined for homologs of known enzymes. Pictures of the DNA helices and the Erlenmeyer flask were retrieved from servier medical art (https://smart.servier.com/), licensed under Creative Commons Attribution 3.0 (CC BY).

A rational selection of the sample site can therefore guide the screening process and help to identify biocatalysts with the desired characteristics. In a next step, the extracted DNA can be used for the construction of genomic or metagenomic libraries,

or can be sequenced directly. The library again, can either be subjected to function-based or sequence-based screening to identify novel enzymes. Function-based screening offers the opportunity to identify completely new biocatalysts, but suffers from a strong dependence on the availability of a suitable and effective screening assay for the desired trait (Ferrer, Martínez-Martínez, *et al.*, 2015) and the genetic compatibility of the original organisms of the library and the host organism (Liebl *et al.*, 2014). To address this, broad host range expression vectors were developed to enable function-based screening in host organisms additional to *Escherichia coli*, for example *Bacillus subtilis*, *Pseudomonas putida*, or *Streptomyces lividans* (Martinez *et al.*, 2004; Troeschel *et al.*, 2012). Sequence-based screening on the other hand, is independent from these restrictions, and, with modern IT infrastructure and bioinformatics tools at hand, can be faster than function-based screening. The mining of metagenomic DNA sequences using DNA probes, specific or degenerated PCR primers, overall sequence homology or hidden-Markov-models has led to the identification of numerous enzymes (Simon and Daniel, 2011; Danso *et al.*, 2018). However, sequence-based screening relies on known sequence-to-function relationships, which restricts this method to the identification of enzymes similar to already known biocatalysts (Madhavan *et al.*, 2017).

1.4 Carboxylic ester hydrolases

The technological aspect of blue biotechnology is the industrial application of marine derived biomolecules. Besides the widely-applied green fluorescent protein from the jellyfish *Aequorea victoria* and compounds like chemo-therapeutics introduced before (Table 1-1), enzymes and biocatalytic processes are a major product of studies related to blue biotechnology.

In the industrial context, different enzymes are used for manifold applications, including the dairy, baking, beverage, animal feed, pulp and paper, polymer, detergent, cosmetic, medical, chemical, pharmaceutical or waste management industry (Singh *et al.*, 2016). Enzymes are classified into 7 classes and several subclasses, defined by the Enzyme Commission (EC) as part of the International Union of Biochemistry and Molecular Biology (IUBMB), according to the chemical reaction they catalyze. Recognized classes are oxidoreductases (EC 1), transferases (EC 2), hydrolases (EC 3), lyases (EC 4), isomerases (EC 5), ligases (EC 6), and translocases (EC 7). They

catalyze oxidation/reduction reactions, the transfer of functional groups *e. g.* amino groups, water dependent cleavage/synthesis reactions, water independent addition/removal of functional groups, the molecular rearrangement of a substrate molecule, the energy (ATP) dependent synthesis of a molecule, or the translocation of a molecule, respectively.

Among the most important enzyme classes for industrial biocatalysis are hydrolases (EC 3), which are subdivided in 13 subclasses depending on the chemical group which is hydrolyzed, with enzymes acting on ester bonds (EC 3.1) being the most diverse group consisting again of more than 20 subgroups. Among them, carboxylic ester hydrolases (EC 3.1.1) can be further specified according to their preferred substrate leading to more than 100 accepted names for carboxylic ester hydrolase enzymes, with growing number each year. For example, lipases (EC 3.1.1.3) are discriminated from carboxylesterases (EC 3.1.1.1), sometimes referred to as esterases, by differences in their substrate preference: where lipases can hydrolyze long chain triacylglycerols, esterases prefer water miscible short chain triacylglycerols (Bornscheuer, 2002). Lipases, as a prominent example for carboxylic ester hydrolases, fulfill important prerequisites for biotechnological applications. They allow reactions at low temperatures, have exquisite high stereo- and regioselectivity, do not require cofactors or catalyze side reactions (Casas-Godoy *et al.*, 2012). In addition, lipases can be produced recombinantly at high yields and are generally well studied enzymes which allow for rational protein engineering (Jaeger and Eggert, 2002). Lipases are often reported to be organic solvent tolerant, resistant to harsh reaction conditions including acidic or alkaline pH, tolerant for high concentrations of metals or salts, and tolerant for detergents (Salihu and Alam, 2015).

On the basis of their amino acid sequences, bacterial lipolytic enzymes can be classified into 19 different families (Kovacic *et al.*, 2019), most of them consisting of multiple characterized examples. Due to the large number of solved crystal structures today, a structural classification of enzymes from the group of carboxylic ester hydrolases is possible too, leading to five clans (A to E) of carboxylic ester hydrolases with several subfamilies (Chen *et al.*, 2016). Besides typical α/β-hydrolases, β-propeller fold enzymes, and enzymes with an α-helix bundle structure are known (Figure 1-4). Nevertheless, most enzymes from the group of carboxylic ester hydrolases adapt a three dimensional structure of the α/β-hydrolase fold (Chen *et al.*,

2016) with a catalytic triad consisting of a nucleophile, a histidine and an acid in common (Holmquist, 2000).

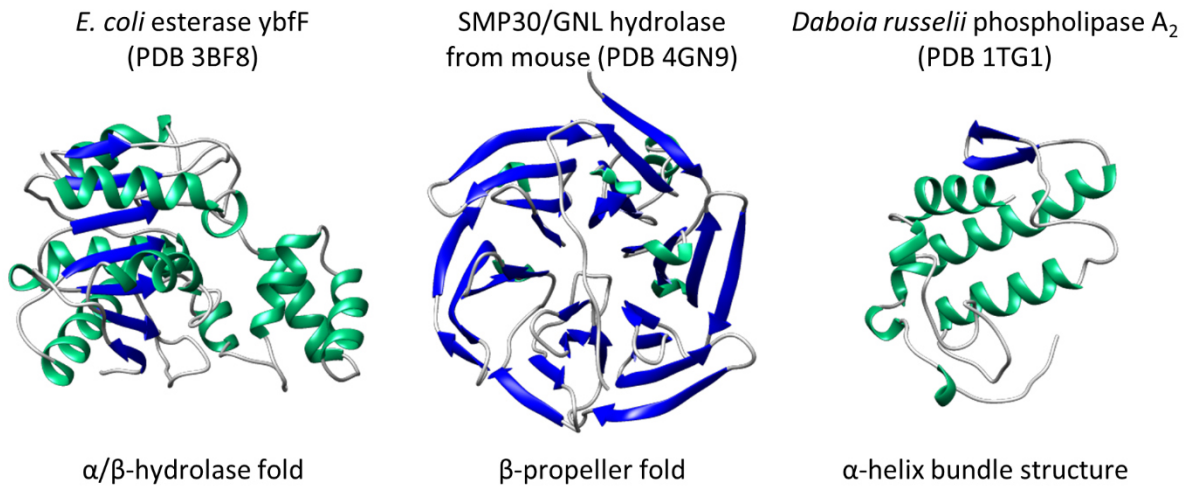


Figure 1-4 Examples for different structural folds from the group of carboxylic ester hydrolases, adapted from (Chen *et al.*, 2016). Crystal structures are shown in ribbon representation with β-sheets colored in blue and α-helices in green.

The α/β-hydrolase fold carboxylic ester hydrolases show a common reaction mechanism. It consist of an acylation and deacylation step, including two tetrahedral intermediate states (Casas-Godoy *et al.*, 2012). First, a serine residue is activated by proton transfer from its hydroxyl group to the imidazole functionality of a close histidine residue which is polarized by a near glutamic- or aspartic acid residue (Figure 1-5). The serine now acts as nucleophile and attacks the carbonyl carbon of the ester bond, leading to a tetrahedral intermediate with a partial negative charge at the carbonyl group which is stabilized by hydrogen bonds with amino groups from the backbone of two amino acids of the so-called oxyanion hole. The alcohol moiety of the former ester is released, and the carboxylic acid remains covalently attached to the serine residue, forming the acyl-enzyme complex. For deacylation, a water molecule is activated the same way the catalytic serine was activated before and functions as nucleophile to attack the carbonyl group of the acyl-enzyme complex. A second tetrahedral intermediate is formed, from which the carboxylic acid moiety of the former ester is released, and the hydroxyl group of the catalytic serine is regenerated.

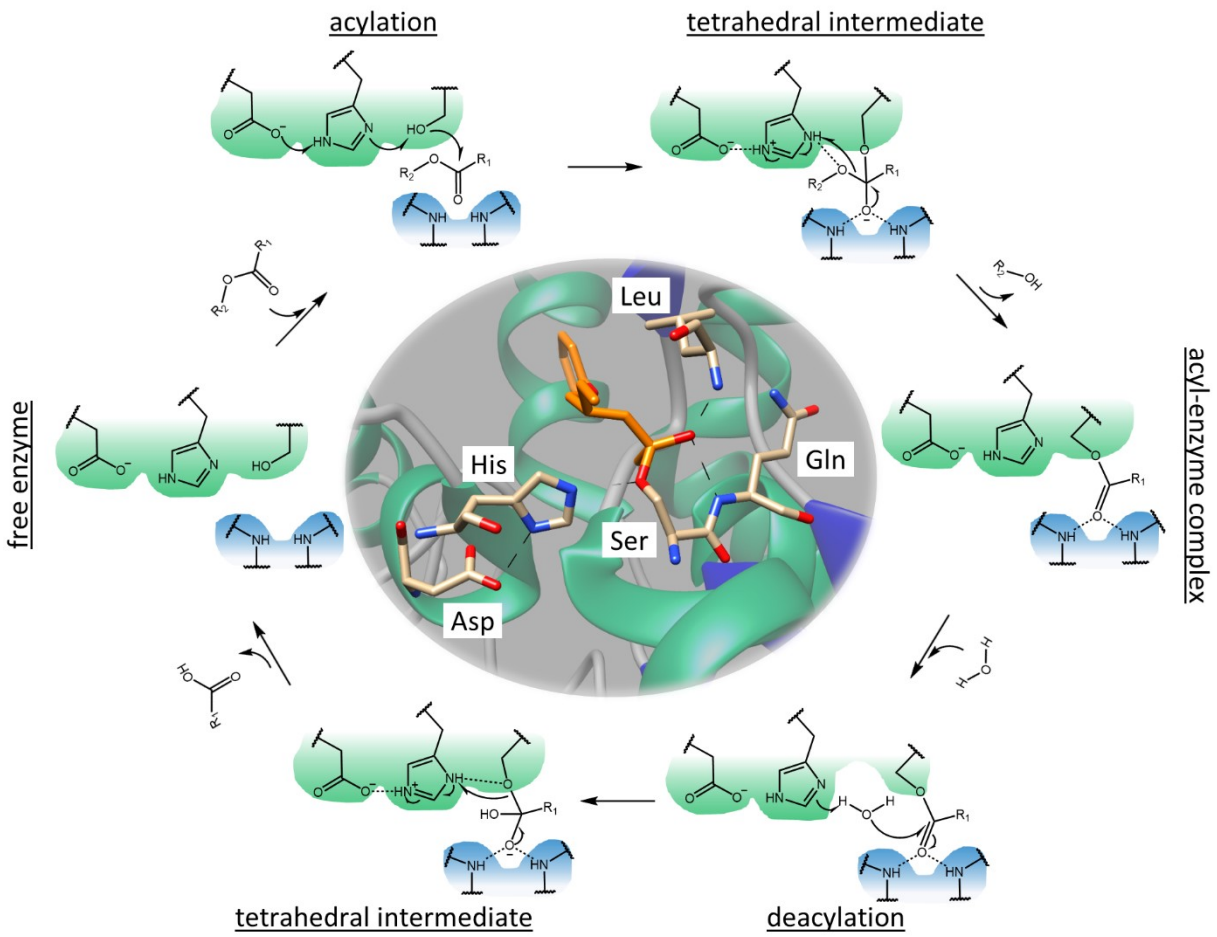


Figure 1-5 Catalytic mechanism of carboxylic ester hydrolases. **Outer circle:** relevant structures at each step in the catalytic cycle are shown, the ground state is depicted on the left (free enzyme). The catalytic triad with amino acid residues of aspartic acid, histidine, and serine is depicted with green background. The oxyanion hole with backbone amino groups of two amino acids is shown with blue background. **Inner circle:** close-up view of the catalytic center of the *Burkholderia cepacia* lipase crystal structure in complex with a substrate analogue (PDB 2NW6), visualized with the program UCSF Chimera. In the crystal structure, the catalytic site residues and amino acids of the oxyanion hole are shown as sticks, superimposed on the ribbon representation of the structure. The substrate analogue (*S*)-1-phenoxybutan-2-yl methylphosphonochloridate is colored orange. The amino acids of the catalytic triad and the oxyanion hole are abbreviated in 3-letter code, hydrogen bonds are indicated by dashed lines.

The detailed understanding of the molecular reaction mechanism, the wealth of structural information, the high production yield, selectivity, catalytic efficiency, chemical tolerance and the simple application without the need for cofactors have contributed to the success of carboxylic ester hydrolases in academic research and industrial applications. However, some relevant characteristics are not fully understood today and need to be addressed to facilitate efficient and effective application of these enzymes: the ability to convert structurally diverse substrates (substrate promiscuity),

the tolerance for typical reaction media applied in chemical and pharmaceutical industry (tolerance for organic solvents), and the reaction with polymeric substrates (polyester hydrolysis) as common in all modern plastic materials.

1.4.1 Substrate promiscuity and enantioselectivity

The principle of enzyme-substrate interaction is often figuratively introduced in the first place as lock and key, meaning one enzyme binds one specific substrate and catalyzes the reaction to a product (Fischer, 1894). This seems logic, in particular when realizing how many different compounds are present within a living cell and how important it is to find the right substrate for an enzyme. For some highly specific enzymes, this analogy applies very well, but for many known biocatalysts it was proven wrong. In contrast to the lock and key hypothesis, most enzymes have been described to be able to use more than one specific substrate (Jensen, 1976; Copley, 2003, 2017; Hult and Berglund, 2007); this feature is called substrate promiscuity.

The benefit of a promiscuous enzyme for an organism lies in the rapid evolution of novel enzymatic functions (Jensen, 1976; O'Brien and Herschlag, 1999; Nobeli *et al.*, 2009). Driven by selective pressure, an organism can thus evolve an enzyme by starting from a low side activity of an existing protein scaffold, to generate an effective novel function. This goes often in hand with gene duplication followed by diversification and specialization, which can lead to the development of completely novel functions (Hughes, 1994). The adaptation to a new but similar substrate might also be accomplished without duplication events when the primary function can be maintained. Such a promiscuous side activity of an enzyme can prove important when environmental conditions suddenly change, for example in case of the exchange of a nutrition source. This was recently shown for the adaptation of *E. coli* to metabolize non-native substrates (*e. g.* D-lyxose) in a laboratory evolution experiment (Guzmán *et al.*, 2019). The authors showed that the population was able to adapt to the new substrate in as few as 20 generations by modifying substrate turnover rates of promiscuous enzymes towards the new substrate. Some enzymes are known, which even are specifically developed to be substrate promiscuous, for example cytochrome P450 enzymes that detoxify a range of xenobiotics in the human liver (Zanger and Schwab, 2013), or plant sesquiterpene synthases that are able to generate many diverse products (*i. e.* 52 different products by a single enzyme) as a reaction to plant

pathogens (Steele *et al.*, 1998). For biotechnological applications of carboxylic ester hydrolases for example, substrate promiscuity is of special interest, giving the opportunity to use an established biocatalytic process with a single highly promiscuous enzyme to produce a range of different products.

Besides substrate promiscuity, selectivity is highly important for the application of enzymes in biocatalytic synthesis routes. In particular, the discrimination of enantiomers by enzymes facilitated the success of biocatalysts in the field of chemistry. Enantiomers, meaning two chemically identical molecules which constitute non-superposable mirror-images, play an extraordinary role in biological systems. This is best illustrated by the fact that natural proteins are made of L-amino acids only. The mirror-image D-amino acids are very rare, for example, they can be found in some antimicrobial peptides (Agrawal *et al.*, 2017). Where it is generally challenging to discriminate enantiomers in chemical synthesis, it is an inherent feature of most biocatalysts to accept only one enantiomer as a substrate. This is of particular importance for pharmaceutical applications, because many active pharmaceutical ingredients are chiral and often only one enantiomer shows the desired effect (Nguyen *et al.*, 2006). Therefore, enzymes, including many examples from the class of carboxylic ester hydrolases, have been successfully applied for kinetic resolution of racemic compounds and were extensively engineered to improve their enantioselectivity (Jaeger and Eggert, 2004; Li and Reetz, 2016).

Both features, substrate promiscuity and chiral selectivity are without question highly desired for many enzyme applications. However, the rather unspecific reaction of promiscuous enzymes with diverse substrates might be a contradiction to the distinct differentiation between enantiomers.

1.4.2 Organic solvent tolerance

Living organisms have evolved in general to thrive in an aqueous environment and so have their proteins and biocatalysts. Thus, most biochemical reactions in a living cell act with water as the bulk solvent. Direct contact to organic solvents in contrast is rare in nature and normally leads to the disruption of the cell membrane of an organism and to the degeneration of its proteins, leading eventually to cell death (Isken and Heipieper, 2003). The tolerance of enzymes towards decent amounts of organic solvent is on the other hand often a prerequisite for their biotechnological application in synthetic chemistry (Figure 1-6). Many important chemical reactions can only be carried out in the absence of water to shift the reaction equilibria towards synthesis, to allow high substrate load, to tune selectivity, to reduce side reactions, or to solubilize hydrophobic reactants (Klibanov, 2001; Kumar *et al.*, 2016). Thus, enzymes which can withstand the presence of organic solvents are important tools for industrial biotechnology.

class	name	log *P*	log *P*	name	class
sulfoxide	dimethyl sulfoxide	-1.3	1.6	triethylamine	amines
alcohols	methanol	-0.76	2	chloroform	halogenated hydrocarbons
esters	ethyl acetate	0.68	2.5	toluene	aromatic hydrocarbons
ethers	diethylether	0.85	3.2	cyclohexane	cyclic hydrocarbons
ketones	hexanone	1.3	3.5	*n*-hexane	aliphatic hydrocarbons

Figure 1-6 Examples for organic solvents commonly used in organic chemistry. The name and classification of each organic solvent is shown along with the partitioning coefficient in octanol:water (log *P*) (data taken from Laane *et al.*, 1987) and the chemical structure. The picture of the Erlenmeyer flask was retrieved from servier medical art (https://smart.servier.com/), licensed under Creative Commons Attribution 3.0 (CC BY).

Organic solvent tolerant enzymes can be found in many different organisms, they are often reported for extremophiles, or organisms living in (organic solvent) polluted environments. Some examples for naturally organic solvent tolerant enzymes from the class of carboxylic ester hydrolases are found within microbial lipases (Jaeger and

Eggert, 2002), but with considerable difference in the degree of tolerance among this group of enzymes. Furthermore, the tolerance of enzymes towards different organic solvents is always related to the chemical solvent properties, like polarity, denaturation capacity, and hydrophobicity (Kumar *et al.*, 2016). To compare among different enzymes, the organic solvent and the reaction conditions must be identical. Most studies feature multiple tested organic solvents for enzyme stability assessment, overlapping in the common polar organic solvent methanol. However, the individual experimental conditions are considerably different, thus, a classification often becomes difficult. Some examples of bacterial organic solvent tolerant carboxylic ester hydrolases are shown in Table 1-2 along with their reported performance in methanol.

Table 1-2 Examples of organic solvent tolerant carboxylic ester hydrolases and reported data on methanol resistance.

enzyme	methanol	performance*	reference
HZ lipase from *Aneurinibacillus thermoaeropilus* HZ	25 %	146 %, 0.5 h	(Masomian *et al.*, 2013)
lipase from *Bacillus licheniformis* MTCC 6824	50 %	71 %, 0.5 h	(Chakraborty and Raj, 2008)
lipase from *Bacillus thermoleovorans* CCR11	70 %	90 %, 1 h	(Castro-Ochoa *et al.*, 2005)
S31 lipase from *Burkholderia cepacia*	25 %	100 %, 6 h	(Lu *et al.*, 2009)
LST-03 lipase from *Pseudomonas aeruginosa* LST-03	25 %	50 %, 276 h	(Ogino *et al.*, 2000)
RL74 from *Ralstonia sp*. CS274	25 %	61 %, 24 h	(Yoo *et al.*, 2011)
SCL from *Staphylococcus capitis*	15 %	70 %, 0.5 h	(Rmili *et al.*, 2019)
Est2 from a compost soil metagenome	60 %	100 %, 2 h	(Lu *et al.*, 2019)

* the performance shows the residual activity (%) after incubation in methanol for a specified amount of time (h) as stated in the reference.

The molecular basis of organic solvent tolerance is not understood to the last detail yet, however, structural, computational, and mutagenesis studies have shed some light on some principle mechanisms. Comparative molecular dynamics (MD) simulation of human lysozyme, *Burkholderia cepacia* lipase (BCL) and *Trametes versicolor* laccase revealed effects on the structure of the proteins in the presence of polar and non-polar organic solvent (Mohtashami *et al.*, 2018). The denaturation was mainly caused by the

disruption of hydrophobic interactions, which are known to be the most important stabilizing forces for protein structure (Pace *et al.*, 2011). The initiation of this process is different for polar and non-polar organic solvents but is promoted by the penetration of organic solvent in the hydrophobic core of the protein for both cases. In a study of Mohtashami and coworkers, it could be demonstrated that *n*-hexane, as an example for a non-polar organic solvent, first caused surface denaturation of human lysozyme, followed by the penetration into the protein core which led to a structural collapse, whereas methanol, as an example for a polar organic solvent, tended to penetrate faster into the protein, rapidly broke the tertiary structure, followed by denaturation of β-sheet secondary structural elements (Mohtashami *et al.*, 2018). Furthermore, it was found that BCL is more tolerant for organic solvents than the laccase or lysozyme and the tolerance was specified to be higher for non-polar over polar organic solvents. The higher tolerance for non-polar over polar organic solvents was reported for a number of different enzyme classes (Stepankova *et al.*, 2013) and is connected with the stripping of essential water by polar organic solvents from the protein surface (Zaks and Klibanov, 1988). Obviously, the tolerance of proteins for organic solvent relates to their three-dimensional structure. As an example, α/β-hydrolase fold proteins, which are common among carboxylic ester hydrolases, comprise helical structures around the β-sheet core of the protein (Figure 1-4), the helical structures are more stable in polar organic solvent, shielding the less stable β-sheet elements (Mohtashami *et al.*, 2018).

A further factor directly affecting protein stability in organic solvent is the concentration of organic solvent in aqueous solution. It seems likely, that the more solvent is present the less stable a protein will be. However, it was reported that organic solvent tolerant enzymes are often more stable in pure organic solvent than in highly concentrated mixtures with water or even pure water. This effect was attributed to enhanced rigidity or loss of flexibility in the absence of water which functions as a lubricant (Klibanov, 2001); the tradeoff of reduced flexibility however, was calculated to be a drastically reduced catalysis rate (Mohtashami *et al.*, 2018). In contrast, small amounts of organic solvent can drive catalysis by lowering the activation free energy of the reaction (Mohtashami *et al.*, 2018), forcing the enzyme in an open confirmation (Salihu and Alam, 2015), or disrupting enzyme aggregates (Almarsson and Klibanov, 1996).

Nevertheless, the activation effect of organic solvent on some enzymes cannot be attributed to the polarity of the solvents alone. The effect of different polar organic

solvents was studied regarding the enzyme activity and structure for the *B. subtilis* lipase variant B6 (Kamal *et al.*, 2013). Here, a rapid loss of activity was observed for acetone, acetonitrile, and dimethylformamide, whereas an activation effect was observed for dimethyl sulfoxide, 2-propanol, and methanol up to 30 to 40% organic solvent concentration, then the activity was completely abolished. However, these effects were not supported by observations on a structural level, because the protein retained its native structure over all organic solvent concentrations up to 60 %. Hence, effects on enzyme activity can be assessed before structural effects become visible. This was also shown for the hydrolysis of *rac*-1-phenylpropyl by BCL, which showed an increased activation free energy in the presence of 50 % methanol, whereas the enzyme did not show structural denaturation at even higher methanol concentrations (Mohtashami *et al.*, 2018).

Thus, polar organic solvents are more harmful to proteins than non-polar are and high concentrations of polar organic solvents in aqueous solution are more harmful than pure organic solvent, whereas low concentrations of organic solvent can accelerate enzymatic reactions. Because of the high importance of organic solvent tolerance for industrial application and the low number of naturally highly organic solvent tolerant enzymes, extensive studies were conducted to improve enzyme performance in organic solvents as comprehensively reviewed before (Stepankova *et al.*, 2013). Briefly summarized, the authors present that enhanced stability of the enzyme can be generated by immobilization, reversed micellar preparation, chemical modification, use of alternative solvent systems like ionic liquids, stabilizing additives, or genetic engineering. The latter can be achieved by rational design or directed evolution and has been applied successfully many times (Arnold, 1990; Kawata and Ogino, 2009; Reetz *et al.*, 2010; Korman *et al.*, 2013; Monsef Shokri *et al.*, 2014). The enhancement of protein stability is here produced either by stabilization of the native protein form or by destabilization of the unfolded state (Kazlauskas, 2018). Another benefit of the genetic engineering approach is to gain deeper insight into the underlying molecular mechanisms of organic solvent tolerance. An outstanding mutagenesis study of lipase A from *B. subtilis*, covering all possible mutations of the enzyme, showed for example the contribution of amino acid substitutions to detergent resistance (Fulton *et al.*, 2015), stability in ionic liquids (Frauenkron-Machedjou *et al.*, 2015), and organic solvent tolerance (Frauenkron-Machedjou *et al.*, 2018). For organic solvent tolerance, the authors reported substitutions to charged residues at the surface and substitutions

to polar residues in buried regions of the protein to be beneficial; in general, they found substitutions to chemically different amino acid residues to be more beneficial than to chemically similar ones (Frauenkron-Machedjou *et al.*, 2018).

Taken together, organic solvent tolerant enzymes are important for biotechnological applications, but the feature itself is not very common and hard to predict, thus classical screening for these enzymes is needed. To stress the organic solvent tolerance of an enzyme, the most harmful condition can be tested, which is high concentration of polar organic solvent in an aqueous system. If organic solvent tolerant enzymes are found, their stability can be further improved by different techniques, maybe most importantly genetic engineering.

1.4.3 Polyester hydrolysis

The modern world greatly relies on synthetic polymers, many of which have replaced their natural counterparts, like glass, paper, and latex for example, because of their superior durability, mechanical properties, and light weight (Wei and Zimmermann, 2017a). Synthetic polymers are almost exclusively made from fossil feedstocks, are cheap to produce, easy to handle in manufacturing processes, very robust in form of their final products, and not prone to natural decomposing, thus, are the top preferred materials for innumerable applications. They are present in everyday life in form of packaging for food or beverages, textile fibers, or isolative coverings of electronic items, as well as high-tech applications in the medical context, aviation, or information technology. Hence, synthetic polymers are not to be missed in today's society. This is represented by an annual production of approximately 350 million tons of plastic worldwide (PlasticsEurope, 2018).

With all the advantages comes one major drawback, which is the recalcitrance of the synthetic material against decomposition in the environment. Since these polymers do not occur in nature, they are in consequence not biodegradable, thus they require special waste treatment (Wei and Zimmermann, 2017a). Without proper recycling and municipal waste management, the material accumulates in the environment and often ends up in the ocean, where a part of the plastic pollution becomes visible by the formation of large floating garbage patches (Lebreton *et al.*, 2018). When accumulated in the environment, plastic waste can cause hazardous damage to inhabiting wildlife,

for example by ingestion of small plastic particles or entanglement in larger sized fragments (Webb *et al.*, 2013).

Where many industrial countries have installed an efficient waste management system, developing countries often lack the same and in addition, have a high demand for plastics at the same time. Even in Europe, with modern waste management at hand only about 30 % of plastic waste is allocated to recycling; the majority is used for energy recovery or disposed in landfills (PlasticsEurope, 2018).

This pollution problem has gained much public attention in the last years and has driven the research in the field of bioremediation and biosynthesis of synthetic polymers. Olefin based polymers like polyethylene, and polypropylene are accounting for the majority of plastic waste in total (PlasticsEurope, 2018). Those homoatomic polymers are connected by carbon-carbon bonds exclusively (Figure 1-7A) and are highly resistant to degradation in nature (Wei and Zimmermann, 2017a). Heteroatomic polymers like polyesters or polyamides comprise more reactive ester or amide linkages, rendering them more easy to degrade in comparison to olefins; still most synthetic polyesters are not biodegradable (Wei and Zimmermann, 2017a). The most abundant polyester in packaging waste is polyethylene terephthalate (PET), which consists of both aromatic and aliphatic constituents, terephthalic acid and ethylene glycol, respectively. The packaging material is rapidly disposed after a single use in most cases and was long thought to be not biodegradable.

In search for biological decomposition mechanisms of polyesters like PET, research has focused on natural polyester hydrolytic enzymes, namely cutinases. The class of cutinases belongs to the group of carboxylic ester hydrolases with cutin, a plant derived (aliphatic) polyester, as their natural substrate. Cutinase enzymes were historically identified in plant pathogenic microorganisms where they function as extracellular hydrolytic enzymes to cleave the hydrophobic wax cuticle of the plant and facilitate the invasion by the pathogen (Nikolaivits *et al.*, 2018). Therefore, they have evolved to act on hydrophobic ester linked polymers, which also occur in synthetic polyesters. Polyester hydrolytic cutinases were found to be active on a range of synthetic polyesters (Nikolaivits *et al.*, 2018), some of them are even able to cleave the ester linkage in PET (Taniguchi *et al.*, 2019) (Figure 1-7B).

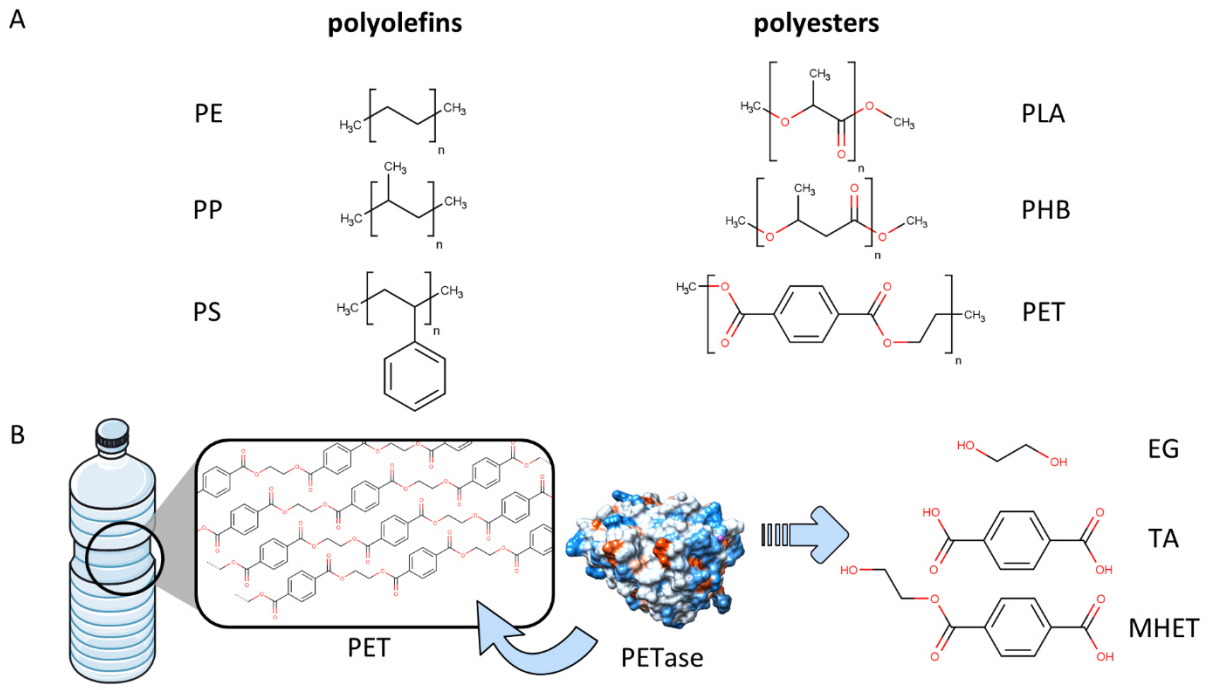


Figure 1-7 (A) Examples of polyolefins and polyesters**.** Chemical structures of polyethylene (PE), polypropylene (PP), polystyrene (PS), polylactic acid (PLA), polyhydroxybutyrate (PHB), and polyethylene terephthalate (PET) are shown. **(B)** Schematic representation of enzymatic depolymerization of PET. The PET hydrolytic enzyme, here *Ideonella sakaiensis* PETase molecular structure is shown (PDB 6EQE), acts on polymeric PET from a PET bottle to produce monomeric ethylene glycol (EG), terephthalic acid (TA), and mono(2-hydroxyethyl) terephthalate (MHET). The picture of the plastic bottle was retrieved from servier medical art (https://smart.servier.com/), licensed under Creative Commons Attribution 3.0 (CC BY).

In 2016, the first report of a bacterium able to degrade and assimilate PET was published including the identification of corresponding catalysts, a polyester hydrolase called PETase and a second hydrolytic enzyme called MHETase. The first enzyme was classified as the first member of a novel group of carboxylic ester hydrolases, the polyethylene terephthalate hydrolases, according to their preference for PET over other tested esters (Yoshida *et al.*, 2016). Nevertheless, enzymatic activity of known examples of polyester hydrolases on PET is rather low, therefore novel enzymes are needed to be discovered in search for better performing biocatalysts (Wei and Zimmermann, 2017b). Recently, a huge potential in the marine environment, which was underestimated before, for polyester hydrolases and PETase like enzymes was shown in a large sequence-based bioprospecting study (Danso *et al.*, 2018). In fact, most polyester hydrolases known today originate from terrestrial sources, suggesting a huge untapped potential in the marine environment.

1.5 Outline of the thesis

The availability of robust biocatalysts with broad applicability is important to drive the progress of enzyme application in industrial biotechnology. The aim of this thesis was to find novel carboxylic ester hydrolases which match demands common for industrial applications; in particular organic solvent tolerance, a broad substrate spectrum and activity with polyester substrates. Thus, the wealth of marine microbes was tapped to uncover biocatalysts showing the demanded features and gain insights into the underlying molecular mechanisms.

Therefore, the chance to find diverse relevant carboxylic ester hydrolases was evaluated for different marine crude oil-associated bacteria (chapter 2.1 & 2.2). Subsequently, the identification, production and characterization of 25 carboxylic ester hydrolases from two marine hydrocarbonoclastic bacteria, namely *Alcanivorax borkumensis* and *Pseudomonas aestusnigri*, was undertaken:

(I) The substrate promiscuity was investigated and ranked among diverse carboxylic ester hydrolases (chapter 2.3) with special emphasis on enantioselectivity (chapter 2.4).

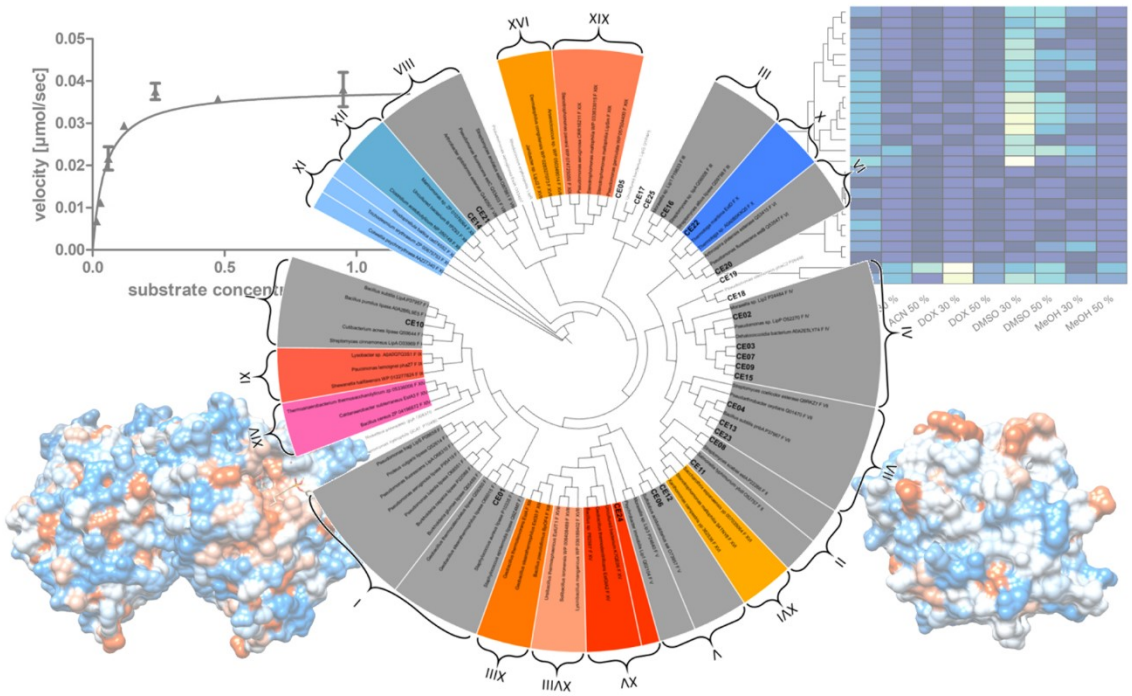
(II) A novel, fast, and easy screening system was developed to determine the tolerance for polar organic solvents and identify extraordinary stable carboxylic ester hydrolases (chapter 2.5).

(III) The hydrolysis of polyester substrates was demonstrated with different easy to use, agar plate-based methods for *P. aestusnigri* and other Pseudomonads (chapter 2.6). Furthermore, the responsible biocatalyst of *P. aestusnigri* was identified, and its activity was proven for the complex polyester polyethylene terephthalate (PET) including insights into the underlying structure-function relationship (chapter 2.7).

This way, novel promising carboxylic ester hydrolases for industrial biotechnological applications were discovered and likewise, principal understandings of the molecular mechanisms of respective features were gained. This can pave the way to future studies on the combination of multiple desired features in a single biocatalyst.

2 Results

The following section consists of seven manuscripts, six are published in peer-reviewed scientific journals and one is published as a book chapter, building together the framework of this thesis. This thesis was conducted in frame of the INMARE H2020 project (http://www.inmare-h2020.eu/). Hence, the work outlined below include also results of joint publications in close collaboration with partners of the INMARE consortium, in particular with the group of Manuel Ferrer at the Institute of Catalysis, Consejo Superior de Investigaciones Científicas (CSIC) in Madrid, Spain and the group of Sander Smits at the Center for Structural Studies (CSS), Heinrich Heine University, Duesseldorf, Germany. Each publication is covered by a statement on the own contribution.



from screening to structure - a bioprospecting story

2.1 Hydrocarbon-degrading microbes as sources of new biocatalysts

Cristina Coscolín, Rafael Bargiela, Mónica Martínez-Martínez, Sandra Alonso, Alexander Bollinger, Stephan Thies, Tatyana N. Chernikova, Tran Hai, Olga V. Golyshina, Karl-Erich Jaeger, Michail M. Yakimov, Peter N. Golyshin, and Manuel Ferrer

In: McGenity T. (eds) Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes. Handbook of Hydrocarbon and Lipid Microbiology. Springer, Cham. (2018)

Available online:

https://doi.org/10.1007/978-3-319-60053-6_13-1

Own contribution:

Investigation: construction and screening of a genomic library from *Alcanivorax borkumensis* SK2 and identification of novel carboxylic ester hydrolases thereof, suggestion of the genetic background for the subterminal alkane oxidation by *A. borkumensis*. Participation in review and editing of the book chapter.

Hydrocarbon-Degrading Microbes as Sources of New Biocatalysts

Cristina Coscolín, Rafael Bargiela, Mónica Martínez-Martínez, Sandra Alonso, Alexander Bollinger, Stephan Thies, Tatyana N. Chernikova, Tran Hai, Olga V. Golyshina, Karl-Erich Jaeger, Michail M. Yakimov, Peter N. Golyshin, and Manuel Ferrer

Contents

C. Coscolín · M. Martínez-Martínez · S. Alonso · M. Ferrer (✉)
Institute of Catalysis, Consejo Superior de Investigaciones Científicas (CSIC), Madrid, Spain
e-mail: cristina.coscolin@csic.es; m.martinez@csic.es; sandra.alonso@csic.es; mferrer@icp.csic.es

R. Bargiela · T. N. Chernikova · T. Hai
School of Natural Sciences, Bangor University, Bangor, UK
e-mail: f.bargiela@bangor.ac.uk; t.chernikova@bangor.ac.uk; t.hai@bangor.ac.uk

A. Bollinger · S. Thies · K.-E. Jaeger
Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf and Forschungszentrum Jülich GmbH, Jülich, Germany
e-mail: a.bollinger@fz-juelich.de; s.thies@fz-juelich.de; k.-e.jaeger@fz-juelich.de

O. V. Golyshina · P. N. Golyshin
School of Natural Sciences, Bangor University, Bangor, UK

Centre for Environmental Biotechnology, Bangor University, Bangor, UK
e-mail: o.golyshina@bangor.ac.uk; p.golyshin@bangor.ac.uk

M. M. Yakimov
Institute for Coastal Marine Environment, Consiglio Nazionale delle Ricerche, Messina, Italy

Immanuel Kant Baltic Federal University, Kaliningrad, Russia
e-mail: michail.yakimov@iamc.cnr.it

T. J. McGenity (ed.), *Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes*, Handbook of Hydrocarbon and Lipid Microbiology,
https://doi.org/10.1007/978-3-319-60053-6_13-1

Abstract

Petroleum hydrocarbons, including those discharged to the marine environment, are metabolized through different catabolic pathways by a number of microorganisms. Each hydrocarbon-degrading microorganism produces interesting enzymes for degrading alkanes and/or aromatic compounds that allow them to be used as sources of carbon and energy, and thus, these microbes occupy hydrocarbon-rich ecological niches. Their diversity and hydrocarbon-degrading metabolic abilities have been extensively examined in multiple environmental and phylogenetic contexts. Genes encoding enzymes involved in degradation, such as alkane hydroxylases and other monooxygenases, P450 cytochromes, rubredoxin reductases, and ferredoxin reductases, have been examined by genome analysis, and a number of them have been successfully cloned, expressed, purified, and their activities confirmed. However, in these microorganisms, the accumulated information regarding other types of enzymes, particularly those most used at industrial level, is limited. Here, we compile information about the accumulated enzymatic knowledge of obligate marine hydrocarbonoclastic bacteria (OMHCB), key players in bioremediation of hydrocarbons in contaminated marine ecosystems. We focused on bacteria of the genera *Cycloclasticus*, *Alcanivorax*, *Oleispira*, *Thalassolituus*, and *Oleiphilus*. Enzymatic data of these representative OMHCB members are restricted to enzymes of the class hydroxylases, cytochrome P450, dioxygenases, synthases, dehalogenases, ligases, and mostly for hydrolases with a typical α/β hydrolase fold. Despite the limited information reported, the available data suggest that these organisms may be important sources of industrial biocatalysts, the analysis of which may deserve deeper investigation. Comparative information is provided regarding the occurrence of key biotechnologically relevant ester-hydrolases in the genomes of OMHCB and suggesting which of the OMHCB may potentially have higher promise as a source of biocatalysts. We also discuss how the properties of these enzymes could be biologicallly important for these bacteria, as some of them can convert a broad range of compounds.

1 Introduction

Hydrocarbons are generally toxic and persistent molecules and their release can seriously harm the environment. Hydrocarbons are degraded by indigenous microorganisms that can use these compounds as carbon source thanks to their specialized metabolism (Cappello et al. 2007; Ron and Rosenberg 2014). Blooms of

hydrocarbon-degrading bacteria (HDB) occur after crude-oil spills or hydrocarbon release (Yakimov et al. 2005). Such blooms may occur frequently as oil spills represent around 10% of global marine oil pollution (GESAMP 2007). In pristine sites, these microorganisms typically comprise less than 1% of the total bacterial population, but they proliferate rapidly when oil contamination occurs, sometimes accounting for up to 90% of the microbial community (Kasai et al. 2002; Cappello et al. 2007; Manilla-Pérez et al. 2010; Bargiela et al. 2015). In all cases, the first step in hydrocarbon biodegradation is the molecule activation presumably between two carbon atoms. The resulting compounds are then transformed through specific catabolic pathways that converge in common intermediates (Díaz et al. 2013; Boll et al. 2014; Bargiela et al. 2017).

1.1 Principal Hydrocarbon-Degrading Bacteria

The different groups of aerobic hydrocarbon-degrading bacteria are specialized in the degradation of a particular range of compounds. Each microorganism possesses different metabolic abilities, allowing the degradation of different types of substrates. Among the HDB is the group of obligate marine hydrocarbonoclastic bacteria (OMHCB), which can only use hydrocarbons as sole carbon and energy source, in addition to some organic acids, such as acetate and pyruvate (Head et al. 2006; McGenity et al. 2012; Joye et al. 2016). To our knowledge, the marine environment is the only place known to date where obligate hydrocarbon degraders have been found. Most representative OMHCB belong to the class *Gammaproteobacteria*, one of the most physiologically and phylogenetically diverse class of Bacteria, consisting of 15 orders (Gutierrez 2017). Within OMHCB we focus on the genera *Alcanivorax*, *Cycloclasticus*, *Oleispira*, *Thalassolitus*, and *Oleiphilus* (Harayama et al. 2004; Yakimov et al. 2007; Manilla-Pérez et al. 2010; Liu and Liu 2013), which are almost undetectable in pristine marine environments, but are among the most abundant bacteria in polluted marine sites.

Alcanivorax is one of the most studied genera, which currently comprises 11 recognized and one proposed species Ca. *A. indicus*. The vast majority of them (*A. borkumensis*, *A. dieselolei*, *A. gelatiniphagus, A. hongdengensis*, *A. jadensis*, *A. venustensis*, and Ca. "*A. indicus*") grow in presence of *n*-alkanes and branched alkanes, with the inability to use any carbohydrate or amino acid as carbon source. The genome of *A. borkumensis* SK2 was the first sequenced OMHCB genome (Yakimov et al. 1998, 2007; Golyshin et al. 2003). Bacteria of the genera *Thalassolituus* (Yakimov et al. 2004), *Oleiphilus* (Golyshin et al. 2002), and *Oleispira* (Yakimov et al. 2003) also have a high specificity for aliphatic alkanes (Yakimov et al. 2004). Members of the genus *Cycloclasticus* (Dyksterhouse et al. 1995) grow on minimal medium supplemented with poly-aromatic hydrocarbons (PAHs) like naphthalene, phenanthrene, or anthracene as sole carbon source (Harayama et al. 2004; Messina et al. 2016). Within the *Gammaproteobacteria*, other recent obligate polycyclic aromatic hydrocarbon (PAH)-degrading specialists have been described, which almost exclusively use PAHs as sole source of carbon

and energy and live in association with marine phytoplankton, including bacteria from the genera *Polycyclovorans*, *Algiphilus*, and *Porticoccus* (Gutierrez et al. 2012, 2013, 2015a, b).

Beyond the OMHCB, other bacterial groups are able to metabolize hydrocarbons that have been recently reviewed by Prince et al. (2018). In this chapter, we focus on the accumulated enzymatic knowledge of OMHCB, key players in the bioremediation of hydrocarbons in contaminated marine ecosystems, particularly during crude-oil spills (Kasai et al. 2002; Cappello et al. 2007; Yakimov et al. 2007; Manilla-Pérez et al. 2010; Bargiela et al. 2015).

1.2 Genomes and Genes Encoding Degrading Enzymes from OMHCB

Imagine the microbial communities responding to hydrocarbon-uptake and how variable this can be (Kostka et al. 2011; Gutierrez et al. 2013). The most obvious response will be production of enzymes supporting the degradation of such hydrocarbons to intermediates feeding into central metabolism (Lu et al. 2011; Mason et al. 2012; Kimes et al. 2013; Mason et al. 2014). These microorganisms may contain versatile enzymes that not only allow microorganisms to metabolize multiple molecules in nature, but also produce enzymes that are potentially valuable to multiple biotechnological processes. The enzyme arsenal of OMHCB has been examined mostly in the context of alkane and aromatics degradation, revealing their diversity and versatility to facilitate pollutant degradation (Kasai et al. 2003; Schneiker et al. 2006; Teimoori et al. 2011, 2012; Golyshin et al. 2013; Kube et al. 2013; Naing et al. 2013; Messina et al. 2016).

Genomes of representative cultivable OMHCB capable of degrading a number of crude-oil components have been reported (Table 1) and their genomic basis established. They include genomes from alkane degraders such as *Alcanivorax* (Golyshin et al. 2003; Schneiker et al. 2006; Lai et al. 2012a; Lai and Shao 2012a, b; Luan et al. 2014; Miura et al. 2014; Barbato et al. 2015; Zhang et al. 2016), *Oleispira* (Kube et al. 2013), *Oleiphilus* (Toshchakov et al. 2017), *Thalassolituus* (Golyshin et al. 2013; Dong et al. 2014) and from PAH-specializing degraders such as *Cycloclasticus* (Messina et al. 2016; Toshchakov et al. 2017). Currently (February 2018), the genomes of seven species of the genera *Alcanivorax* are available (Lai et al. 2012a; Lai and Shao 2012a, b; Parks et al. 2017; Fu et al. 2018), with that of *A. borkumensis* SK2 being the first to be sequenced (Schneiker et al. 2006) (Table 1). Genomes of *Cycloclasticus,* including those of *Cycloclasticus* sp. 78-ME (Messina et al. 2016), *Cycloclasticus* sp. strain PY97M (Cui et al. 2013), and *Cycloclasticus* sp. strain P1 (Lai et al. 2012b), are available. Within the genera *Oleispira, Thalassolituus*, and *Oleiphilus*, only the genomes of *Oleispira antarctica* RB-8 (Kube et al. 2013), *Thalassolituus oleivorans* MIL-1, and *Oleiphilus messinensis* ME102 (Toshchakov et al. 2017) have been sequenced and are available (Table 1).

The production and characterization of some of the key catabolic enzymes participating in degradation steps in some of the above OMHCB members have

been examined. Thus, examples exist for the successful cloning, expression, production, and characterization of alkane hydroxylases (Hara et al. 2004; van Beilen et al. 2004; Miri et al. 2010; Naing et al. 2013) and other monooxygenases (Wang and Shao 2012), P450 cytochromes (Jung et al. 2016), rubredoxin reductases (Teimoori et al. 2011), and ferredoxin reductases (Teimoori et al. 2012) from *Alcanivorax* strains. The biochemical information of such enzymes for *Cycloclasticus* strains is limited to few PAH dioxygenases (Kasai et al. 2003; Shindo et al. 2011) and cytochrome P450 (Misawa et al. 2011). For other OMHCB, no examples have been described reporting the successful cloning, purification, and characterization of such enzymes.

2 Biotechnologically Relevant Enzymes from OMHCB

The accumulated level of information regarding the characteristics and potential of other types of enzymes present in the genomes of OMHCB is limited (Table 1). This limited knowledge is particularly noticeable for enzymes that are most commonly used at an industrial level, aldo-keto reductases, transaminases, and serine ester-hydrolases from the structural superfamily of α/β-hydrolases (Ferrer et al. 2015; Martínez-Martínez et al. 2017). Below, we compile the reported cases in which enzymes others that alkane hydroxylases, monooxygenases, P450 cytochromes, rubredoxin reductases, and ferredoxin reductases have been reported. Some of these enzymes have been used for the production of valuable molecules.

2.1 Biotechnologically Relevant Enzymes from *Alcanivorax* Strains

Until February 2018, and to the best of our knowledge, a total of 10 enzymes with potential biotechnological uses have been reported from bacteria of the *Alcanivorax* genus, the most explored species being *A. borkumensis* SK2 (Table 1). A recent example is the production of ω-hydroxy palmitic acid from palmitic acid via *A. borkumensis* SK2 fatty acid ω-hydroxylase (CYP153A13) from the P450 (CYP)153 family (Jung et al. 2016). A high product yield of 4.6 g/L of ω-hydroxy palmitic acid was achieved in fed-batch resting cell reactors when *Escherichia coli* cells expressing CYP153A13 were confronted with 5.1 g/L palmitic acid for 30 h, at 30 °C. Production and characterization of other enzymes from *A. borkumensis* SK2 have been reported. They include two ester-hydrolases for oil degradation (Kadri et al. 2018), a succinate coenzyme A ligase with a broad substrate range and useful for the formation of succinate analogues in vitro (Nolte et al. 2014), and a TesB-like thioesterase for the production of extracellular polyhydroxyalkanoates (Sabirova et al. 2006). Recently, three hydrolases from *A. borkumensis* SK2 were found to be efficient for either hydrolyzing poly-(DL-lactic acid) polyester (ABO2449) (Hajighasemi et al. 2016) or esters containing *p*-nitrophenyl and α-naphthyl blocks, with acyl chains ranging from acetate to

Table 1 Current (February 2018) protein-coding genes, theoretical and with activity experimentally confirmed, present in the genomes of representative OMHCB

	No. of proteins in the genome	Accession number of genome sequence	No. of proteins with confirmed enzymatic activity[a]
A.borkumensis[T]	2,750	NC_008260	19[b]
A. diselolei	4,362	CP003466	1[c]
A. jadensis	3,266	ARXU01000000	n.d.
A. hondengengensis	3,416	AMRJ00000000	n.d.
A. nanhaiticus	3,778	ARXV00000000	n.d.
A. pacificus	3,669	CP004387	n.d.
A. xenomutans	4,275	CP012331	n.d.
Cycloclasticus sp. 78-ME	2,585	FO203512	7[d]
Cycloclasticus sp. strain PY97M	2,264	ASHL00000000	n.d.
Cycloclasticus sp. strain P1	2,249	CP003230	n.d.
O. antarctica RB-8	3,919	FO203512	15[e]
T. oleivorans MIL-1	3,603	HF680312	n.d.
O. messinensis ME102	5,502	PRJNA362330	n.d.

[a]Enzymes others than those involved in biodegradation steps such as alkane hydroxylases and other monooxygenases, P450 cytochromes, rubredoxin reductases, ferredoxin reductases, and dioxygenases. Abbreviation: *n.d.*, not described

[b]Enzymes with confirmed activity included one fatty acid ω-hydroxylase from the P450 (CYP)153 family, one succinate coenzyme A ligase, one TesB-like thioesterase, and 16 ester-hydrolases

[c]The enzyme with confirmed activity included a haloalkane dehalogenase

[d]Enzymes with confirmed activity included seven ester-hydrolases, four of them belonging to the MCP family. Structure is available for one enzyme with dual ester-hydrolase: MCP hydrolase activity (PDB 4I3F)

[e]Enzymes with confirmed activity included one pyrophosphatase, one phoshonoacetaldehyde hydrolase, two fumarylacetoacetate isomerase/hydrolase, two 2-keto-3-deoxy-6-phosphogluconate aldolase, one amidohydrolase, one isochorismatase hydrolase, one transaldolase, one glycerophosphodiesterase, one dihydroorotate oxidase, two ester-hydrolases, and two chaperones. The crystal structures for 12 of these enzymes (3I4Q, 3IRU, 3L53, 3LAB, 3LNP, 3LQY, 3M16, 3QVM, 3QVQ, 3V77, 3VCR, 3I6Y) and one protein with unknown function (3LMB) are available

laurate, as well as tri-acyl-glycerols (preferably tributyrin) and other short-chain esters (Tchigvintsev et al. 2015). Notably, polyester-degrading activity of ABO2449 was found in emulsified and solid poly-(DL-lactic acid), with the capacity to degrade solid material being of high industrial interest. ABO2449 had highest esterase activity at 30–37 °C and retained 32% of its maximal activity at 4 °C, suggesting that it is a cold-adapted esterase, which is consistent with the ability of *A. borkumensis* to grow at 4 °C. Cold-active esterase has been also reported in *A. dieselolei* (Zhang et al. 2014a); the enzyme did show high stability in the presence of solvents. A haloalkane dehalogenase from *A. dieselolei* B-5 with

potential use for biocatalysis and bioremediation applications has been also reported (Li and Shao 2014). An activity assay with 46 halogenated substrates indicated this enzyme possessed broad substrate range, with preference for brominated substrates and chlorinated alkenes. The dehalogenase was most active in the range from 20 °C to 50 °C, with an optimal at 50 °C. Also, a 5-enolpyruvyl-shikimate-3-phosphate synthase for the biosynthesis of aromatic amino acids has been reported from the deep-sea bacterium *Alcanivorax* sp. L27 through screening the genomic library (Zhang et al. 2014b). Its optimal temperature was 50 °C, and it retained 20% of its activity at 0 °C. To the best of our knowledge, no enzyme with confirmed activity has been described for any of the genomes of the other five species of the genera *Alcanivorax* with genome sequences available (Table 1).

Taken together, the available biochemical data revealed that broad substrate range and optimal temperatures of up to 20–50 °C and capacity to retain some activity at 4 °C are common features of enzymes from bacteria of the genus *Alcanivorax*.

2.2 Biotechnologically Relevant Enzymes from *Cycloclasticus* Strains

Until February 2018, and to the best of our knowledge, a total of six enzymes with potential biotechnological uses have been reported from bacteria of the genus *Cycloclasticus* (Table 1). Bioconversions in which hydroxyl groups were regio- and stereo-specifically introduced into various substituted naphthalenes, such as 1-methoxy- and 1-ethoxy-naphthalenes, methylnaphthalenes, dimethylnaphthalenes, and naphthalenecarboxylic acid methyl esters, and β-eudesmol, have been reported. Thus, the conversion was performed by recombinant *E. coli* cells expressing an aromatic dihydroxylating dioxygenase from *Cycloclasticus* strain A5 (Shindo et al. 2011); 10 novel prenyl naphthalen-ols were produced by combinatorial bioconversion when cells expressing the dioxygenase were allowed to react with a number of aromatic substrates at 25 °C. Similarly, *E. coli* cells expressing cytochrome P450 BM3 variant from *Cycloclasticus* strain A5 were capable of introducing at 25 °C a hydroxyl group regio- and stereo-specifically into a sesquiterpene β-eudesmol (Misawa et al. 2011). Also, four hydrolases from the α/β hydrolase family of *Cycloclasticus* sp. ME7 have also been reported with the unusual capacity to efficiently hydrolase C–O and C–C bonds in a broad spectrum of substrates (Alcaide et al. 2013). These enzymes were most active at 40–55 °C, although they retained high activity at temperatures as low as 4–20 °C.

Taken together, as for enzymes from *Alcanivorax*, the available biochemical data suggest that some of the enzymes from *Cycloclasticus* are also characterized by an unusually broad substrate range and are capable of performing well at temperatures of up to 25–55 °C, but also capable of retaining activity at temperatures as low as 4 °C. The only crystal structure reported is from the dual ester-hydrolase:meta-cleavage product (MCP) hydrolase from *Cycloclasticus* sp. ME7 (PDB 4I3F; Alcaide et al. 2013).

2.3 Biotechnologically Relevant Enzymes from *Oleispira* Strains and Other OMHCB

Bacteria of the *Oleispira* genus are the OMHCB for which the most industrially relevant enzymes have been reported currently (February, 2018) (Table 1). The analysis of 15 proteins (including, chaperones, hydrolases, and oxidases) from *O. antarctica* RB-8 revealed that most of its enzymes are functioning suboptimally at temperatures close to that being optimal for bacterial growth. However, their activities at 4 °C are nevertheless sufficient to facilitate the active growth of this bacterium in polar (and deep sea) waters (Ferrer et al. 2003; Goral et al. 2012; Lemak et al. 2012; Kube et al. 2013). Enzymes with confirmed activity included a pyrophosphatase (OLEAN_C30460), a glycerophosphodiesterase (OLEAN_C34790), a dihydroorotate oxidase (OLEAN_C16020), and three ester-hydrolases (OLEI01171, OLEAN_C09750, and OLEAN_C31070) from *O. antarctica* RB-8 (Lemak et al. 2012; Kube et al. 2013). The structures of four of these enzymes have been reported (PDB 3I4Q, 3QVM, 3I6Y, and 3QVQ; Lemak et al. 2012; Kube et al. 2013). These enzymes have in common showed salt-stimulation and optimum temperatures of 15–25 °C, although some of them were found to be most active at 35–50 °C (OLEAN_C34790) and some retained high activity level at 4 °C (OLEAN_C09750); however, their substrate profile was not evaluated in a broad context, so that their biotechnological potential remains to be established. Additionally, although experimental confirmations are needed, the activity of one phoshonoacetaldehyde hydrolase, two fumarylacetoacetate isomerase/hydrolase, two 2-keto-3-deoxy-6-phosphogluconate aldolase, one amidohydrolase, one isochorismatase hydrolase, and one transaldolase were putatively suggested by their X-ray structural determinations (3IRU; 3LAB; 3 L53; 3LNP; 3LQY; 3 M16; 3 V77; 3VCR). The structure of a protein with unknown function from this bacterium is also available (3LMB; Kube et al. 2013). Finally, two proteins from the strain RB-8, which do not have enzymatic activity but chaperone activities, were published (Ferrer et al. 2003) and patented, and their expression in mesophilic bacteria such as *E. coli* allows growth at temperatures as low as 4–10 °C.

Taken together, *Oleispira antarctica* RB-8 is the bacterium from the OMHCB for which more extensive biochemical and structural data are available, and the accumulated data suggest its enzymes as being most active at 15–25 °C, but some of them are capable of performing well at 50 °C, and others retain more than 80% activity at temperatures as low as 4 °C. Until February 2018, no examples were found that report enzymes from bacteria of the genera *Thalassolituus* and *Oleiphilus* (Table 1).

3 Ester-hydrolases from OMHCB

Ester-hydrolases are a class of enzymes widely distributed in the environment and with important physiological functions; this is why at least one can be found in each bacterial genome (Ferrer et al. 2015). They include hydrolases that are among the most important industrial biocatalysts, and extensive biochemical knowledge has

been accumulated from many microorganisms (Ferrer et al. 2015). As mentioned above, currently (February 2018), five ester-hydrolases from the α/β hydrolase family from *A. borkumensis* SK2, four from *Cycloclasticus* sp. ME7, and three from *O. antarctica* RB-8 have been isolated and characterized. By using both genomic and metagenomics approaches, the latter to avoid cultivation and thus access the widest possible genomic variability (Distaso et al. 2017), a wide set of 145 ester-hydrolases from environmentally and phylogenetically diverse origin have been isolated and their substrate specificity reported (Martínez-Martínez et al. 2018). A subset of 25 of them were isolated from OMHCB, particularly from bacteria of the genera *Alcanivorax* (11 in total), *Cycloclasticus* (12), *Thalassolituus* (1), and *Oleiphilus* (1). The source of the 25 enzymes was the genomes of three cultivable OMHCB, *A. borkumensis* SK2, *Cycloclasticus* sp. 78-ME, *O. messinensis* ME102^{T}, and clone libraries created from chronically polluted seawater samples from Milazzo harbor (Sicily, Italy), Messina harbor (Sicily, Italy), Bizerte lagoon (Tunisia), Ancona harbor (Ancona, Italy), and crude oil enrichment from coastal seawater (Kolguev, Russia). The description of this new set of enzymes from OMHCB allows us to further investigate their features and relevance, which will be discussed below.

3.1 Diversity and Divergence at the Sequence Level

The sequences of the 25 ester-hydrolases from OMHCB recently described by Martínez-Martínez et al. (2018) were compared with the sequences available in the NCBI nonredundant public database. Protein sequences were 51.3–99.7% similar to noncharacterized homologous proteins in the database. The pairwise amino acid sequence identity ranged from 0.5% to 99.6%. The pairwise amino acid sequence identity for all 11 ester-hydrolases from *A. borkumensis* SK2 ranged from 2.8% to 23.3%. The pairwise amino acid sequence identity for all 12 ester-hydrolases from *Cycloclasticus* sp. 78-ME ranged from 7.0% to 65.8%, except two enzymes which share 99.6% identity at the amino acid sequence level. The only ester-hydrolase from *Thalassolituus* shows a pairwise comparison of up to 23% to all other ester-hydrolases, and the one from *O. messinensis* ME102T up to 23.4%. BLAST searches were performed for all query sequences by running NCBI BLASTP against the current version of the Lipase Engineering Database using an E-value threshold of 10^{-10}. All, but four, were unambiguously assigned to some of the 14 existing families (F) of the Arpigny and Jaeger classification (Arpigny and Jaeger 1999; Ferrer et al. 2015), which are defined based on amino acid sequence similarity and the presence of specific sequence motifs. These included sequences with a typical α/β hydrolase fold and conserved G-X-S-X-G (FI: 3, FIV: 3, FV: 5, FVI: 1, and FVII: 1) or G-X-S-(L) (FII: 1) motifs and sequences with a serine beta-lactamase-like modular (non α/β hydrolase fold) architecture and a conserved S-X-X-K motif (FVIII: 4). An additional set of seven sequences were assigned to the MCP hydrolase family. Noticeably, all seven sequences encoding hydrolases from the MCP family have been isolated from *Cycloclasticus*; this observation is consistent with the fact that this enzyme participates in the degradation of aromatics such as catechol and biphenyl,

for which this bacterium is specialized compared to *Alcanivorax*, *Thalassolituus*, and *Oleiphilus*, which are alkane degraders.

Taken together, the primary sequence analysis suggests that the diversity of polypeptides encoding ester-hydrolases from OMHCB is not dominated by a particular type of protein or highly similar protein clusters but consists of diverse nonredundant sequences assigned to multiple folds and subfamilies, which are distantly related to known homologs in many cases. It also revealed the extensive divergence at the sequence level between and within ester-hydrolases from OMHCB.

3.2 Biotechnologically Relevant Characteristics

Experimental data on substrate conversion (i.e., units g^{-1}) followed for 24 h at pH 8.0 and 30 °C were reported for each of the 25 ester-hydrolases from OMHCB for 96 distinct esters. These include esters with variation in size of acyl and alcohol groups and with growing residues (aromatic, aliphatic, branched, and unbranched) at both sides, leading to more challenging substrates because a larger group adjacent to the ester bond increases the difficulty of conversion. Halogenated, chiral, and sugar esters, lactones and an alkyl di-ester, were also included. The substrate profiles of all 25 ester-hydrolases, when tested with a set of 96 chemically and structurally distinct esters, are summarized in Fig. 1. The enzymes had a substrate spectrum that ranged from 62 to 4 substrates. According to recently established criteria (Martínez-Martínez et al. 2018), 2 out of 25 would fall into the category of ester-hydrolases with prominent substrate promiscuity (capable of hydrolyzing 30 or more esters) and 16 into the category of ester-hydrolases with moderate substrate promiscuity (capable of hydrolyzing from 10 to 29 esters), and only seven hydrolyzed nine or fewer esters. The percentage of ester-hydrolases showing moderate-to-prominent promiscuity accounted for around 73% when considering all 145 enzymes previously reported by Martínez-Martínez et al. (2018) and which assigned to at least 10 phyla and 40 genera. This percentage is similar to that found for ester-hydrolases of OMHCB (18 out of 25; or 72%), which suggests that, at least in terms of substrate scope, the hydrolytic potential of ester-hydrolases from these OMHCB is as good as that of other bacterial groups. This is important from a practical point of view, as, along with requirements of a technical nature such as selectivity, scalability, and robustness, a narrow substrate spectrum is one of the most frequent problems for industrial enzyme applications (Martínez-Martínez et al. 2018). A consensus exists that "the more substrates an enzyme converts the better," opening application ranges with consequent reduction of the production cost of multiple enzymes.

As shown in Fig. 1, each of the 25 enzymes is characterized by distinct substrate spectra in agreement with the differences at the sequence level. The two ester-hydrolases from *Oleiphilus* and *Thalassolituus* were characterized by a narrow substrate scope, but the low number of enzymes characterized so far from these bacteria does not allow conclusions to be drawn about whether this is a common feature among ester-hydrolases from these bacteria. However, as the number of

ester-hydrolases from *Alcanivorax* (11 in total) and *Cycloclasticus* (12 in total) was in the same range, the differences in substrate scope could be analyzed. As shown in Fig. 1, we could observe that the capability to convert large aromatic and sterically hindered esters such as benzyl (*R*)-(+)-2-hydroxy-3-phenylpropionate, benzoic acid-4-formyl-phenylmethyl ester, 2,4-dichlorophenyl 2,4-dichlorobenzoate, 2,4-dichlorophenyl 2,4-dichlorobenzoate, and diethyl-2,6-dimethyl 4-phenyl-1,4-dihydro pyridine-3,5-dicarboxylate was most frequent among ester-hydrolases from *Cycloclasticus* (8 out 12 enzymes), compared to enzymes from *Alcanivorax* (1 out 11 enzymes). This was also noticed when examining the hydrolytic capacity towards mono-aromatic esters, such as benzoate esters that were also hydrolyzed by most *Cycloclasticus* esterases, compared with *Alcanivorax* ester-hydrolases. This agrees with the fact that members of the genus *Alcanivorax* grow commonly in the presence of *n*-alkanes and branched alkanes, with the inability to use, in the majority of the cases, PAHs, whereas members of the genus *Cycloclasticus* commonly grow on PAHs. It is thus plausible that the ability of *Cycloclasticus* ester-hydrolases to convert hindered esters is a reflection of the adaptation of this bacterium to manage complex aromatic molecules. Ester-hydrolases capable of converting polyaromatic esters have been identified also by metagenomic-based mining of a naphthalene-enriched community (Martínez-Martínez et al. 2014), suggesting that genomes and metagenomes from PAH-polluted environments may be a source of enzymes with a higher probability to convert hindered molecules.

The biotechnological potential of ester-hydrolases can be evaluated not only by their substrate scope, but also by their ability to convert important molecules in synthetic organic chemistry such as *p*-hydroxybenzoic acid (paraben) esters and related esters (2- or 3-hydroxybenzoates). Two *Alcanivorax* ester-hydrolases and six from *Cycloclasticus* had the capacity to hydrolyze these chemicals, which demonstrates that these organisms' enzymes have the potential to perform conversions of industrially relevant molecules.

Interestingly, ester-hydrolases from all four OMHCB were able to release free acid from methyl esters of hydroxycinnamic acids, such as ferulic and cinnamic esters. This finding was unexpected since cinnamoyl/feruroyl ester hydrolases are abundant in microbes involved in fiber breakdown processes (for review see Wong 2006), but not in marine microorganisms. In addition, one of the *Alcanivorax* ester-hydrolases hydrolyzed glucose esters. The physiological role of this type of activity is normally associated with the cleavage of ester bonds to remove the acetyl moieties from complex polymers (Wong 2006). The deacetylation process increases biodegradability and renders complex polysaccharides more accessible to the attack of other polysaccharide hydrolytic enzymes (Biely et al. 1985; Grohman et al. 1989; Blum et al. 1999; Christov and Prior 1993). Taking this into account, it appears that ester-hydrolases from OMHCB, i.e., bacteria widespread in oceans after petroleum spills, might have the potential capacity to contribute to the degradation of complex carbohydrates or polymers (i.e., coming from algae) present in the oceans.

Finding versatile biocatalysts that are chirally selective is of interest for the pharmaceutical and chemical industry (Coscolín et al. 2018). The experimental data on substrate conversion against a number of chiral esters (including (*R*) and

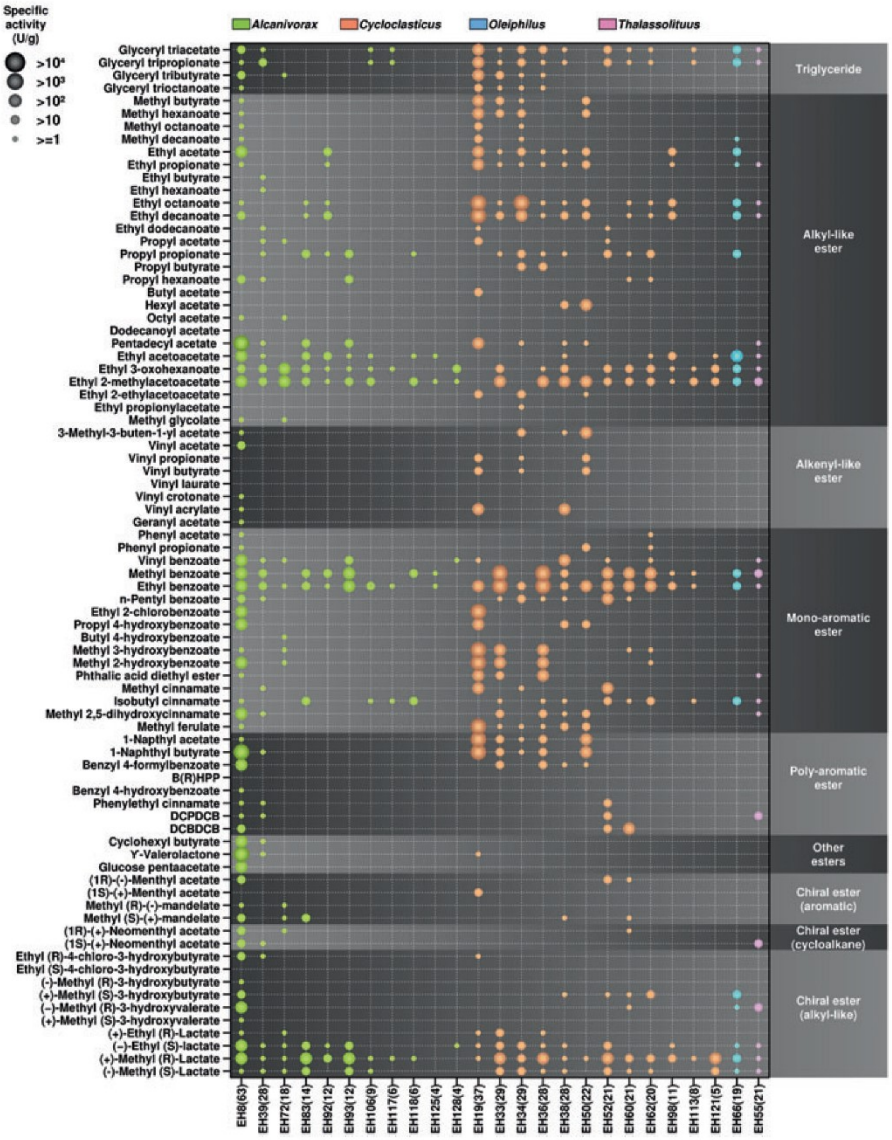


Fig. 1 Substrate ranges of ester-hydrolases from OMHCB Ester-hydrolases. The ID and characteristic of each of the ester-hydrolases is reported elsewhere (Martínez-Martínez et al. 2018). The ID code representing each ester-hydrolase is given at the bottom. Each hydrolase is named based on the code "EH," which means ester-hydrolase, followed by an arbitrary number for the most to least promiscuous enzyme, following the criteria reported by Martínez-Martínez et al. (2018). The number in bracket indicates the number of esters hydrolyzed by each enzyme, which are listed in the left side. The figure was created with the R language console using information about the specific activity in units/g protein (measurable by the size of the circle; see legend on the left side) at 30 °C and pH 8.0 of the analyzed enzymes against a set of 96 substrates tested by Martínez-Martínez et al. (2018). Only the esters for which activity was detected for any of the ester-hydrolases

(*S*) enantiomers) have been also reported for the 25 ester-hydrolases from OMHCB (Martínez-Martínez et al. 2018; Coscolín et al. 2018). The reported data revealed the distinct capacity of ester-hydrolases from OMHCB to hydrolyze 16 chiral esters that included: (*R*) and/or (*S*) enantiomers of menthyl acetate, methyl mandelate, neo-menthyl acetate, ethyl 4-chloro-3-hydroxybutyrate, methyl 3-hydroxybutyrate, methyl 3-hydroxyvalerate, and methyl and ethyl lactate. Preferred enzymes are those that show stringent selectivity, although it is commonly considered that enzymes with a selectivity factor of 25 or above begin to have commercial value (Coscolín et al. 2018). As shown in Fig. 1, a number of ester-hydrolases were capable of hydrolyzing only one of the enantiomers for some of the esters, namely, 9 out 11 from *Alcanivorax*, 8 out of 12 from *Cycloclasticus*, 1 from *Oleiphilus*, and 1 from *Thalassolituus*.

3.3 Promiscuous Character of Ester-hydrolases from OMHCB: Physiological Implication

The relevance of substrate promiscuity is indisputable as the operating basis for biological processes and cell function (Martínez-Martínez et al. 2018). For example, the evolutionary progress of enzymes from lower to higher substrate specificity allows the recruitment of alternate pathways for carbon cycling and innovations across metabolic subsystems and the tree of life by maximizing the growth rate and growth efficiency. In general, promiscuous enzymes enable an energetically more favorable lifestyle for a cell than specialized enzymes, and therefore, the cell does not require many different enzymes to take up substrates, favoring genome minimization and streamlining. In addition, the extension of the substrate scope without compromising primary or ancestral enzyme activities is a major driver of microbial adaption to extreme habitats.

Out of the 11 ester-hydrolases from *Alcanivorax* reported by Martínez-Martínez et al. (2018), 10 were retrieved from the genome of *A. borkumensis* SK2 (Schneiker et al. 2006). One additional ester-hydrolase (named EH92) was isolated from a clone library created from chronically polluted seawater samples from Ancona harbor (Ancona, Italy); taxonomic binning of the DNA fragment containing the gene encoding this ester-hydrolase associated it to *Alcanivorax*, although its assignation to strain SK2 could not be unambiguously confirmed. Additionally, Tchigvintsev et al. (2015) and Hajighasemi et al. (2016) reported three ester-hydrolases, one capable of hydrolyzing poly-(DL-lactic acid) polyester, from the genome of *A. borkumensis* SK2. Finally, the presence of a lipase and an esterase in protein extracts from strain

Fig. 1 (continued) from OMHCB (81 esters out of 96 tested) are listed in the figure. Abbreviations as follows: DCBDCB, 2,4-dichlorobenzyl 2,4-dichlorobenzoate; DCPDCB, 2,4-dichlorophenyl 2,4-dichlorobenzoate; B(R)HPP, benzyl (R)-(+)-2-hydroxy-3-phenylpropionate

SK2 was also confirmed by Kadri et al. (2018), although their identities and sequences were not confirmed. The genome of strain SK2 consists of a circular chromosome with 2,755 predicted coding sequences. Thus, genes encoding ester-hydrolases with confirmed activity represent about 0.55% of the total genes (15 out of 2,755 genes). Out of the ester-hydrolases from the genome of SK2, one (named EH8 following the nomenclature by Martínez-Martínez et al. 2018) showed a prominent promiscuity level being capable of converting 63 chemically and structurally diverse esters out of 96 tested, whereas the other nine were only capable of converting fewer than 28 esters (Fig. 1). Promiscuous ester-hydrolases thus represent a minor percentage in the genome of *A. borkumensis* SK2, and possibly other OMHCB. It is thus plausible that *Alcanivorax* developed a limited number of highly promiscuous ester-hydrolases, which may have a major biological role that is yet to be established. This enzyme was capable of degrading acetyl moieties from molecules that are known to be included in complex polymers, and thus, its role in microbial survival by expanding the set of potential substrates that can be used as carbon source cannot be ruled out. An additional function of esterases in bacteria such as *Alcanivorax* is supporting degradation of linear alkanes. Here, a specific aerobic degradation mechanism designated as subterminal oxidation pathway exists as an alternative path to the well-explored terminal degradation. This pathway requires a monooxygenase (MO), an alcohol dehydrogenase (DH), a Baeyer Villiger monooxygenase (BMVO), and finally an esterase. The genetic background of this degradation pathway has presently not been described; but loci encoding putative BMVO, DH, and esterase can be identified in the genome of *A. borkumensis* suggesting an operon organization of the respective genes. It is plausible that some of the esterases described to date from *Alcanivorax strains* may be related to or identical with the enzymes which cleave alkanes producing acetates.

Out of the 12 ester-hydrolases from *Cycloclasticus*, seven were retrieved from the genome of *Cycloclasticus* sp. 78-ME (Alcaide et al. 2013; Messina et al. 2016). The genome of strain 78-ME consists of a circular chromosome and a plasmid with 2,585 predicted coding sequences. Thus, genes encoding ester-hydrolases with confirmed activity represents about 0.3% of the total genes (seven out of 2,585). The remaining five ester-hydrolases were isolated from clone libraries created from chronically polluted seawater samples from Milazzo harbor (Sicily, Italy), Messina harbor (Sicily, Italy), and Bizerte lagoon (Tunisia); taxonomic binning of the DNA fragment containing the gene encoding these ester-hydrolases associated them to *Cycloclasticus*. The seven ester-hydrolases from the genome of strain 78-ME showed a capacity to hydrolyze from five to 37 esters. Four of them belong to the MCP family, two showed a typical α/β hydrolase fold with a conserved G-X-S-X-G or G-X-S-(L) motif, and one was not classified into existing families. Hydrolases from the MCP family hydrolases are known to help in the degradation of aromatic pollutants such as catechol and biphenyl (Alcaide et al. 2013). However, the ability of such enzymes to convert molecules other than fission products of bicyclic (i.e., 2-hydroxy-6-oxo-6-phenylhexa-2,4-dienoate) and monocyclic (i.e., 2-hydroxy-6-oxohepta-2,4-dienoate) aromatics is uncommon. The available biochemical data suggest that *Cycloclasticus* has developed an arsenal of hydrolases with broad

substrate range that not only participate in the degradation of pollutants, but also expand the pool of substrates capable of being used by this bacterium. Hence, these types of enzymes may contribute to the global carbon cycling processes and for channeling complex substrates into the common catabolic pathways, including recalcitrant organic pollutants.

4 Concluding Remarks and Research Needs

Crude oil pollution and the chemical diversity of its components, in combination with environmental constraints such as depth, oxygen concentration, temperature, nutrient input, and other physical and chemical factors, may distinctly influence microbial populations and the biodegradation processes they mediate in response to accidental crude-oil spills in seawater and seawater sediments. A particular group of marine bacteria, so-called OMHCB, are known to be among those that increase most in response to oil spills. Their analysis by cultivation approaches in combination with genomic and metagenomics analysis has confirmed that they are widespread in marine environments. Their versatile ability to degrade a number of pollutants through a series of catabolic enzymes has also been extensively investigated. Accumulated information exists about their taxonomic diversity, genomes, and genes. However, this understanding has not been as well developed for investigating enzymes with potential biotechnological use. This is particularly noticeably given OMHCB are widespread microorganisms in marine environments, particularly after oil spills, cultivable members of all known genera are available, and genomes are available. Recent investigations of a number of enzymes from key OMHCB members have revealed they contain enzymes, not only supporting the degradation of pollutants to be used as carbon sources, but also other enzymes capable of expanding the potential carbon sources to be used by those bacteria and also capable of converting a wide portfolio of molecules. Accumulated knowledge revealed that the properties of these enzymes, particularly their substrate scope, may be a reflection of the microbial metabolism and its adaptation to the respective environment. The ability of OMHCB enzymes to convert multiple substrates could relevant to their survival in nature as few of them can convert a large number of compounds, normally present simultaneously in natural environment where oil spills may occur. Whatever the biological meaning of the broad substrate scope of some of the OMHCB enzymes, the available data suggest that OMHCB bacteria contain in their genomes multiple ester-hydrolases (at least 15 confirmed in the genome of *A. borkumensis* SK2, and seven in the genome of *Cycloclasticus* sp. 78-ME) capable of converting multiple molecules, while at the same time being stereo-selective, two features highly desired in industrial settings. Recent studies support this suggestion, since a functional genome analysis of the recently described marine bacterium *Pseudomonas aestusnigri* revealed not only the presence of a highly specialized metabolism, but also of 22 different putative esterase genes which represent 0.62% of all predicted coding sequences (Gomila et al. 2017). Actually, there is a need to generate more information that helps to decipher not only the industrial potential of

available enzymes from OMHCB, but also to recover new ones by genome mining of already available genomes (Table 1) or metagenomic mining in contaminated marine sites and to investigate their distribution in the environment and their biochemical and structural properties and biotechnological potential.

Acknowledgments This project received funding from the European Union's Horizon 2020 research and innovation program [Blue Growth: Unlocking the potential of Seas and Oceans] under grant agreement no. [634486] (project acronym INMARE). This research was also supported by the grants PCIN-2014-107 (within ERA NET IB2 grant no. ERA-IB-14-030 - MetaCat), PCIN-2017-078 (within the ERA-MarineBiotech grant ProBone), BIO2014-54494-R, and BIO2017-85522-R from the Spanish Ministry of Economy, Industry and Competitiveness (actually, Ministry of Science, Innovation and Universities). P.N.G. gratefully acknowledges funding from the UK Biotechnology and Biological Sciences Research Council (grant no. BB/M029085/1). R.B. and P. N.G. acknowledge the support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund (ERDF) via Welsh Government. P.N.G. acknowledges the support of the Centre of Environmental Biotechnology Project funded by the European Regional Development Fund (ERDF) through Welsh Government. The authors gratefully acknowledge financial support provided by the European Regional Development Fund (ERDF). C. Coscolín thanks the Spanish Ministry of Economy, Industry and Competitiveness for a PhD fellowship (Grant BES-2015-073829).

References

Alcaide M, Tornés J, Stogios PJ, Xu X, Gertler C, Di Leo R, Bargiela R, Lafraya A, Guazzaroni ME, López-Cortés N, Chernikova TN, Golyshina OV, Nechitaylo TY, Plumeier I, Pieper DH, Yakimov MM, Savchenko A, Golyshin PN, Ferrer M (2013) Single residues dictate the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family. Biochem J 454:157–166

Arpigny JL, Jaeger KE (1999) Bacterial lipolytic enzymes: classification and properties. Biochem J 343:177–183

Barbato M, Mapelli F, Chouaia B, Crotti E, Daffonchio D, Borin S (2015) Draft genome sequence of the hydrocarbon-degrading bacterium *Alcanivorax dieselolei* KS-293 isolated from surface seawater in the Eastern Mediterranean Sea. Genome Announc 3:e01424–e01415

Bargiela R, Mapelli F, Rojo D, Chouaia B, Tornés J, Borin S, Richter M, Del Pozo MV, Cappello S, Gertler C, Genovese M, Denaro R, Martínez-Martínez M, Fodelianakis S, Amer RA, Bigazzi D, Han X, Chen J, Chernikova TN, Golyshina OV, Mahjoubi M, Jaouanil A, Benzha F, Magagnini M, Hussein E, Al-Horani F, Cherif A, Blaghen M, Abdel-Fattah YR, Kalogerakis N, Barbas C, Malkawi HI, Golyshin PN, Yakimov MM, Daffonchio D, Ferrer M (2015) Bacterial population and biodegradation potential in chronically crude oil-contaminated marine sediments are strongly linked to temperature. Sci Rep 5:11651

Bargiela R, Yakimov MM, Golyshin PN, Ferrer M (2017) Distribution of hydrocarbon degradation pathways in the sea. In: TJ MG, Timmis KN, Nogales B (eds) Consequences of microbial interactions with hydrocarbons, oils, and lipids: production of fuels and chemicals. Springer International Publishing, Cham, pp 1–23

Biely P, Puls J, Schneider H (1985) Acetylxylan esterases in fungal cellulolytic systems. FEMS Lett 186:80–84

Blum DL, Li X, Chen H, Ljungdahl LG (1999) Characterization of an acetyl xylan esterase from anaerobic fungus *Orpinomyces* sp. strain PC-2. Appl Environ Microbiol 65:3990–3995

Boll M, Löffler C, Morris BEL, Kung JW (2014) Anaerobic degradation of homocyclic aromatic compounds via arylcarboxyl-coenzyme A esters: organisms, strategies and key enzymes. Environ Microbiol 16:612–627

Cappello S, Caruso G, Zampino D, Monticelli LS, Maimone G, Denaro R, Tripodo B, Troussellier M, Yakimov MM, Giuliano L (2007) Microbial community dynamics during assays of harbour oil spill bioremediation: a microscale simulation study. J Appl Microbiol 102:184–194

Christov LP, Prior BA (1993) Esterases of xylan-degrading microorganisms: production, properties, and significance. Enzyme Microb Technol 15:460–475

Coscolín C, Martínez-Martínez M, Chow J, Bargiela R, García-Moyano A, Bjerga GEK, Bollinger A, Stokke R, Steen IH, Golyshina OV, Yakimov MM, Jaeger K-E, Yakunin AF, Streit WR, Golyshin PN, Ferrer M (2018) Relationships between substrate promiscuity and chiral selectivity of esterases from phylogenetically and environmentally diverse microorganisms. Catalysts 8:10

Cui Z, Xu G, Li Q, Gao W, Zheng L (2013) Genome sequence of the pyrene- and fluoranthene-degrading bacterium *Cycloclasticus* sp. strain PY97M. Genome Announc 1:e00536–e00513

Díaz E, Jiménez JI, Nogales J (2013) Aerobic degradation of aromatic compounds. Curr Opin Biotechnol 24:431–442

Distaso MA, Tran H, Ferrer M, Golyshin PN (2017) Metagenomic mining of enzyme diversity. In: TJ MG, Timmis KN, Nogales B (eds) Consequences of microbial interactions with hydrocarbons, oils, and lipids: production of fuels and chemicals. Springer International Publishing, Cham, pp 1–25, 1–23

Dong C, Chen X, Xie Y, Lai Q, Shao Z (2014) Complete genome sequence of *Thalassolituus oleivorans* R6-15, an obligate hydrocarbonoclastic marine bacterium from the Arctic Ocean. Stand Genomic Sci 9:893–901

Dyksterhouse SE, Gray JP, Herwig RP, Lara JC, Staley JT (1995) *Cycloclasticus pugetii* gen. nov., sp. nov., an aromatic hydrocarbon-degrading bacterium from marine sediments. Int J Syst Bacteriol 45:116–123

Ferrer M, Chernikova TN, Yakimov MM, Golyshin PN, Timmis KN (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures. Nat Biotechnol 21:1266–1267

Ferrer M, Bargiela R, Martínez-Martínez M, Mir J, Koch R, Golyshina OV, Golyshin PN (2015) Biodiversity for biocatalysis: a review of the α/β-hydrolase fold superfamily of ester-hydrolases-lipases discovered in metagenomes. Biocatal Biotransformation 33:235–249

Fu X, Lai Q, Dong C, Wang W, Shao Z (2018) Complete genome sequence of *Alcanivorax xenomutans* P40, an alkane-degrading bacterium isolated from deep seawater. Mar Genomics 38:1–4

GESAMP (2007) Estimates of oil entering the marine environment from sea-based activities. Journal series GESAM reports and studies, vol 75. IMO Publisher, London, 96 pp

Golyshin PN, Chernikova TN, Abraham W-R, Lünsdorf H, Timmis KN, Yakimov MM (2002) Oleiphilaceae fam. nov., to include *Oleiphilus messinensis* gen. nov., sp. nov., a novel marine bacterium that obligately utilizes hydrocarbons. Int J Syst Evol Microbiol 52:901–911

Golyshin PN, Martins Dos Santos VA, Kaiser O, Ferrer M, Sabirova YS, Lünsdorf H, Chernikova TN, Golyshina OV, Yakimov MM, Pühler A, Timmis KN (2003) Genome sequence completed of *Alcanivorax borkumensis*, a hydrocarbon-degrading bacterium that plays a global role in oil removal from marine systems. J Biotechnol 106:215–220

Golyshin PN, Werner J, Chernikova TN, Tran H, Ferrer M, Yakimov MM, Teeling H, Golyshina OV, MAMBA Scientific Consortium (2013) Genome sequence of *Thalassolituus oleivorans* MIL-1 (DSM 14913T). Genome Announc 1:e0014113

Gomila M, Mulet M, Lalucat J, García-Valdés E (2017) Draft genome sequence of the marine bacterium *Pseudomonas aestusnigri* VGXO14 T. Genome Announc 5:e00765–e00717

Goral AM, Tkaczuk KL, Chruszcz M, Kagan O, Savchenko A, Minor W (2012) Crystal structure of a putative isochorismatase hydrolase from *Oleispira antarctica*. J Struct Funct Genomics 13:27–36

Grohman K, Mitchell PJ, Himmel ME, Sale BE, Schroeder HA (1989) The role of ester groups in resistance of plant cell wall polysaccharides to enzymatic hydrolysis. Appl Biochem Biotechnol 20:45–61

Gutierrez T (2017) Aerobic hydrocarbon-degrading Gammaproteobacteria – *Porticoccus*. In: McGenity TJ, Prince R (eds) Handbook of hydrocarbon and lipid microbiology, vol 6. Springer, Heidelberg, Germany

Gutierrez T, Green DH, Whitman WB, Nichols PD, Semple KT, Aitken MD (2012) *Algiphilus aromaticivorans* gen. nov., sp. nov., an aromatic hydrocarbon-degrading bacterium isolated from a culture of the marine dinoflagellate *Lingulodinium polyedrum*, and proposal of Algiphilaceae fam. nov. Int J Syst Evol Microbiol 62:2743–2749

Gutierrez T, Singleton DR, Berry D, Yang T, Aitken MD, Teske A (2013) Hydrocarbon-degrading bacteria enriched by the Deepwater Horizon oil spill identified by cultivation and DNA-SIP. ISME J 7:2091–2104

Gutierrez T, Thompson HF, Angelova A, Whitman WB, Huntemann M, Copeland A, Chen A, Kyrpides N, Markowitz V, Palaniappan K, Ivanova N, Mikhailova N, Ovchinnikova G, Andersen E, Pati A, Stamatis D, Reddy TB, Ngan CY, Chovatia M, Daum C, Shapiro N, Cantor MN, Woyke T (2015a) Genome sequence of *Polycyclovorans algicola* strain TG408, an obligate polycyclic aromatic hydrocarbon-degrading bacterium associated with marine eukaryotic phytoplankton. Genome Announc 3:e00207–e00215

Gutierrez T, Whitman WB, Huntemann M, Copeland A, Chen A, Kyrpides N, Markowitz V, Pillay M, Ivanova N, Mikhailova N, Ovchinnikova G, Andersen E, Pati A, Stamatis D, Reddy TB, Ngan CY, Chovatia M, Daum C, Shapiro N, Cantor MN, Woyke T (2015b) Genome sequence of *Porticoccus hydrocarbonoclasticus* strain MCTG13d, an obligate polycyclic aromatic hydrocarbon-degrading bacterium associated with Marine eukaryotic phytoplankton. Genome Announc 3:e00672–e00615

Hajighasemi M, Nocek BP, Tchigvintsev A, Brown G, Flick R, Xu X, Cui H, Hai T, Joachimiak A, Golyshin PN, Savchenko A, Edwards EA, Yakunin AF (2016) Biochemical and structural insights into enzymatic depolymerization of polylactic acid and other polyesters by microbial carboxylester-hydrolases. Biomacromolecules 17:2027–2039

Hara A, Baik SH, Syutsubo K, Misawa N, Smits TH, van Beilen JB, Harayama S (2004) Cloning and functional analysis of alkB genes in *Alcanivorax borkumensis* SK2. Environ Microbiol 6:191–197

Harayama S, Kasai Y, Hara A (2004) Microbial communities in oil-contaminated seawater. Curr Opin Biotechnol 15:205–214

Head IM, Jones DM, Röling WF (2006) Marine microorganisms make a meal of oil. Nat Rev Microbiol 4(3):173–182

Joye SB, Kleindienst S, Gilbert JA, Handley KM, Weisenhorn P, Overholt WA, Kostka JE (2016) Responses of microbial communities to hydro-carbon exposures. Oceanography 23:136–149

Jung E, Park BG, Ahsan MM, Kim J, Yun H, Choi KY, Kim BG (2016) Production of ω-hydroxy palmitic acid using CYP153A35 and comparison of cytochrome P450 electron transfer system in vivo. Appl Microbiol Biotechnol 100:10375–10384

Kadri T, Rouissi T, Magdouli S, Brar SK, Hegde K, Khiari Z, Daghrir R, Lauzon JM (2018) Production and characterization of novel hydrocarbon degrading enzymes from *Alcanivorax borkumensis*. Int J Biol Macromol 112:230–240

Kasai Y, Kishira H, Harayama S (2002) Bacteria belonging to the genus *Cycloclasticus* play a primary role in the degradation of aromatic hydrocarbons released in a marine environment. Appl Environ Microbiol 68:5625–5633

Kasai Y, Shindo K, Harayama S, Misawa N (2003) Molecular characterization and substrate preference of a polycyclic aromatic hydrocarbon dioxygenase from *Cycloclasticus* sp. strain A5. Appl Environ Microbiol 69:6688–6697

Kimes NE, Callaghan AV, Aktas DF, Smith WL, Sunner J, Golding B, Drozdowska M, Morris PJ (2013) Metagenomic analysis and metabolite profiling of deep-sea sediments from the Gulf of Mexico following the Deepwater Horizon oil spill. Front Microbiol 4:50

Kostka JE, Prakash O, Overholt WA, Green SJ, Freyer G, Canion A, Delgardio J, Norton N, Hazen TC, Huettel M (2011) Hydrocarbon-degrading bacteria and the bacterial community response in

Gulf of Mexico beach sands impacted by the Deepwater Horizon oil spill. Appl Environ Microbiol 77:7962–7974

Kube M, Chernikova TN, Al-Ramahi Y, Beloqui A, Lopez-Cortez N, Guazzaroni ME, Heipieper HJ, Klages S, Kotsyurbenko OR, Langer I, Nechitaylo TY, Lünsdorf H, Fernández M, Juárez S, Ciordia S, Singer A, Kagan O, Egorova O, Petit PA, Stogios P, Kim Y, Tchigvintsev A, Flick R, Denaro R, Genovese M, Albar JP, Reva ON, Martínez-Gomariz M, Tran H, Ferrer M, Savchenko A, Yakunin AF, Yakimov MM, Golyshina OV, Reinhardt R, Golyshin PN (2013) Genome sequence and functional genomic analysis of the oil-degrading bacterium *Oleispira antarctica*. Nat Commun 4:2156

Lai Q, Shao Z (2012a) Genome sequence of an alkane-degrading bacterium, *Alcanivorax pacificus* type strain W11-5, isolated from deep sea sediment. J Bacteriol 194:6936

Lai Q, Shao Z (2012b) Genome sequence of the alkane-degrading bacterium *Alcanivorax hongdengensis* type strain A-11-3. J Bacteriol 194:6972

Lai Q, Li W, Shao Z (2012a) Complete genome sequence of *Alcanivorax dieselolei* type strain B5. J Bacteriol 194:6674

Lai Q, Li W, Wang B, Yu Z, Shao Z (2012b) Complete genome sequence of the pyrene-degrading bacterium *Cycloclasticus* sp. strain P1. J Bacteriol 194:6677

Li A, Shao Z (2014) Biochemical characterization of a haloalkane dehalogenase DadB from *Alcanivorax dieselolei* B-5. PLoS One 9:e89144

Liu Z, Liu J (2013) Evaluating bacterial community structures in oil collected from the sea surface and sediment in the northern Gulf of Mexico after the Deepwater Horizon oil spill. Microbiology 2:492–504

Lu Z, Deng Y, Van Nostrand JD, He Z, Voordeckers J, Zhou A, Lee Y-J, Mason OU, Dubinky EA, Chavarria KL, Tom LM, Fortney JL, Lamendella R, Jansson JK, D'haeseller P, Hazen TC, Zhou J (2011) Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume. ISME J 6:451–460

Luan X, Cui Z, Gao W, Li Q, Yin X, Zheng L (2014) Genome sequence of the petroleum hydrocarbon-degrading bacterium *Alcanivorax* sp. strain 97CO-5. Genome Announc 2: e01277–e01214

Manilla-Pérez E, Lange AB, Hetzler S, Steinbüchel A (2010) Occurrence, production, and export of lipophilic compounds by hydrocarbonoclastic marine bacteria and their potential use to produce bulk chemicals from hydrocarbons. Appl Microbiol Biotechnol 86:1693–1706

Martínez-Martínez M, Lores I, Peña-García C, Bargiela R, Reyes-Duarte D, Guazzaroni ME, Peláez AI, Sánchez J, Ferrer M (2014) Biochemical studies on a versatile esterase that is most catalytically active with polyaromatic esters. Microb Biotechnol 7:184–191

Martínez-Martínez M, Bargiela RM, Coscolín C, Navarro J, Golyshin PN, Ferrer M (2017) Functionalization and modification of hydrocarbon-like molecules guided by metagenomics: esterases and lipases from the α/β-hydrolase fold superfamily and transaminases as study cases. In: McGenity TJ, Timmis KN, Nogales B (eds) Consequences of microbial interactions with hydrocarbons, oils, and lipids: production of fuels and chemicals. Springer International Publishing, Cham, pp 1–21. 978-3-319-31421-1

Martínez-Martínez M, Coscolín C, Santiago G, Chow J, Stogios PJ, Bargiela R, Gertler C, Navarro-Fernández J, Bollinger A, Thies S, Méndez-García C, Popovic A, Brown G, Chernikova TN, García-Moyano A, Bjerga GEK, Pérez-García P, Hai T, Del Pozo MV, Stokke R, Steen IH, Cui H, Xu X, Nocek BP, Alcaide M, Distaso M, Mesa V, Peláez AI, Sánchez J, Buchholz PCF, Pleiss J, Fernández-Guerra A, Glöckner FO, Golyshina OV, Yakimov MM, Savchenko A, Jaeger KE, Yakunin AF, Streit WR, Golyshin PN, Guallar V, Ferrer M, The Inmare Consortium (2018) Determinants and prediction of esterase substrate promiscuity patterns. ACS Chem Biol 13:225–234

Mason OU, Hazen TC, Borglin S, Chain PSG, Dubinsky EA, Fortney JL, Han J, Holman H-YN, Hultman J, Lamendella R, Mackelprag R, Malfatti S, Tom LM, Tringe SG, Woyke T, Zhou J, Rubin EM, Jansson JK (2012) Metagenome, metatranscriptome and single-cell sequencing reveal microbial response to Deepwater Horizon oil spill. ISME J 6:1715–1727

Mason OU, Scott NM, Gonzalez A, Robbins-Pianka A, Baelum J, Kimbrel J, Bouskill NJ, Prestat E, Borglin S, Joyner DC, Fortney JL, Jurelevicius D, Stringfellow WT, Álvarez-Cohen L, Hazen TC, Knight R, Gilbert JA, Jansson JK (2014) Metagenomics reveals sediment microbial community response to Deepwater Horizon oil spill. ISME J 8:1464–1475

McGenity TJ, Folwell BD, McKew BA, Sanni GO (2012) Marine crude-oil biodegradation: a central role for interspecies interactions. Aquat Biosyst 8:10

Messina E, Denaro R, Crisafi F, Smedile F, Cappello S, Genovese M, Genovese L, Giuliano L, Russo D, Ferrer M, Golyshin PN, Yakimov MM (2016) Genome sequence of obligate marine polycyclic aromatic hydrocarbons-degrading bacterium *Cycloclasticus* sp. 78-ME, isolated from petroleum deposits of the sunken tanker Amoco Milford Haven, Mediterranean Sea. Mar Genomics 25:11–13

Miri M, Bambai B, Tabandeh F, Sadeghizadeh M, Kamali N (2010) Production of a recombinant alkane hydroxylase (AlkB2) from *Alcanivorax borkumensis*. Biotechnol Lett 32:497–502

Misawa N, Nodate M, Otomatsu T, Shimizu K, Kaido C, Kikuta M, Ideno A, Ikenaga H, Ogawa J, Shimizu S, Shindo K (2011) Bioconversion of substituted naphthalenes and β-eudesmol with the cytochrome P450 BM3 variant F87V. Appl Microbiol Biotechnol 90:147–157

Miura T, Tsuchikane K, Numata M, Hashimoto M, Hosoyama A, Ohji S, Yamazoe A, Fujita N (2014) complete genome sequence of an alkane degrader, *Alcanivorax* sp. Strain NBRC 101098. Genome Announc 2:e00766–e00714

Naing SH, Parvez S, Pender-Cudlip M, Groves JT, Austin RN (2013) Substrate specificity and reaction mechanism of purified alkane hydroxylase from the hydrocarbonoclastic bacterium *Alcanivorax borkumensis* (AbAlkB). J Inorg Biochem 121:46–52

Nolte JC, Schürmann M, Schepers CL, Vogel E, Wübbeler JH, Steinbüchel A (2014) Novel characteristics of succinate coenzyme A (Succinate-CoA) ligases: conversion of malate to malyl-CoA and CoA-thioester formation of succinate analogues in vitro. Appl Environ Microbiol 80:166–176

Parks DH, Rinke C, Chuvochina M, Chaumeil PA, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW (2017) Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. Nat Microbiol 2:1533–1542

Prince RC, Amande TJ, McGenity TJ (2018) Prokaryotic hydrocarbon degraders. In: McGenity TJ (ed) Taxonomy, genomics and ecophysiology of hydrocarbon-degrading microbes: handbook of hydrocarbon and lipid microbiology, 2nd edn. Springer, Cham

Ron EZ, Rosenberg E (2014) Enhanced bioremediation of oil spills in the sea. Curr Opin Biotechnol 27:191–194

Sabirova JS, Ferrer M, Lünsdorf H, Wray V, Kalscheuer R, Steinbüchel A, Timmis KN, Golyshin PN (2006) Mutation in a "tesB-like" hydroxyacyl-coenzyme A-specific thioesterase gene causes hyperproduction of extracellular polyhydroxyalkanoates by *Alcanivorax borkumensis* SK2. J Bacteriol 188:8452–8459

Schneiker S, Martins dos Santos VA, Bartels D, Bekel T, Brecht M, Buhrmester J, Chernikova TN, Denaro R, Ferrer M, Gertler C, Goesmann A, Golyshina OV, Kaminski F, Khachane AN, Lang S, Linke B, McHardy AC, Meyer F, Nechitaylo T, Pühler A, Regenhardt D, Rupp O, Sabirova JS, Selbitschka W, Yakimov MM, Timmis KN, Vorhölter FJ, Weidner S, Kaiser O, Golyshin PN (2006) Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. Nat Biotechnol 24:997–1004

Shindo K, Tachibana A, Tanaka A, Toba S, Yuki E, Ozaki T, Kumano T, Nishiyama M, Misawa N, Kuzuyama T (2011) Production of novel antioxidative prenyl naphthalen-ols by combinational bioconversion with dioxygenase PhnA1A2A3A4 and prenyltransferase NphB or SCO7190. Biosci Biotechnol Biochem 75:505–510

Tchigvintsev A, Tran H, Popovic A, Kovacic F, Brown G, Flick R, Hajighasemi M, Egorova O, Somody JC, Tchigvintsev D, Khusnutdinova A, Chernikova TN, Golyshina OV, Yakimov MM, Savchenko A, Golyshin PN, Jaeger KE, Yakunin AF (2015) The environment shapes microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine metagenomes. Appl Microbiol Biotechnol 99:2165–2178

Teimoori A, Ahmadian S, Madadkar-Sobhani A, Bambai B (2011) Rubredoxin reductase from *Alcanivorax borkumensis*: expression and characterization. Biotechnol Prog 27:1383–1389

Teimoori A, Ahmadian S, Madadkar-Sobhani A (2012) Biochemical characterization of two recombinant ferredoxin reductases from *Alcanivorax borkumensis* SK2. Biotechnol Appl Biochem 59:457–464

Toshchakov SV, Korzhenkov AA, Chernikova TN, Ferrer M, Golyshina OV, Yakimov MM, Golyshin PN (2017) The genome analysis of *Oleiphilus messinensis* ME102 (DSM 13489^T) reveals backgrounds of its obligate alkane-devouring marine lifestyle. Mar Genomics 36:41–47

van Beilen JB, Marín MM, Smits TH, Röthlisberger M, Franchini AG, Witholt B, Rojo F (2004) Characterization of two alkane hydroxylase genes from the marine hydrocarbonoclastic bacterium *Alcanivorax borkumensis*. Environ Microbiol 6:264–273

Wang W, Shao Z (2012) Diversity of flavin-binding monooxygenase genes (*alm*A) in marine bacteria capable of degradation long-chain alkanes. FEMS Microbiol Ecol 80:523–533

Wong DW (2006) Feruloyl esterase: a key enzyme in biomass degradation. Appl Biochem Biotechnol 133:87–112

Yakimov MM, Golyshin PN, Lang S, Moore ERB, Abraham W-R, Lünsdorf H, Timmis KN (1998) *Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing marine bacterium. Int J Syst Bacteriol 48:339–348

Yakimov MM, Giuliano L, Gentile G, Crisafi E, Chernikova TN, Abraham W-R, Lünsdorf H, Timmis KN, Golyshin PN (2003) *Oleispira antarctica* gen. nov., sp. nov., a novel hydrocarbonoclastic marine bacterium isolated from Antarctic coastal sea water. Int J Syst Evol Microbiol 53:779–785

Yakimov MM, Giuliano L, Denaro R, Crisafi E, Chernikova TN, Abraham W-R, Lünsdorf H, Timmis KN, Golyshin PN (2004) *Thalassolituus oleivorans* gen. nov., sp. nov., a novel marine bacterium that obligately utilizes hydrocarbons. Int J Syst Evol Microbiol 54:141–148

Yakimov MM, Denaro R, Genovese M, Cappello S, D'Auria G, Chernikova TN, Timmis KN, Golyshin PN, Giluliano L (2005) Natural microbial diversity in superficial sediments of Milazzo Harbor (Sicily) and community successions during microcosm enrichment with various hydrocarbons. Environ Microbiol 7:1426–1441

Yakimov MM, Timmis KN, Golyshin PN (2007) Obligate oil-degrading marine bacteria. Curr Opin Biotechnol 18:257–266

Zhang S, Wu G, Liu Z, Shao Z, Liu Z (2014a) Characterization of EstB, a novel cold-active and organic solvent-tolerant esterase from marine microorganism *Alcanivorax dieselolei* B-5(T). Extremophiles 18:251–259

Zhang Y, Yi L, Lin Y, Zhang L, Shao Z, Liu Z (2014b) Characterization and site-directed mutagenesis of a novel class II 5-enopyruvylshikimate-3-phosphate (EPSP) synthase from the deep-sea bacterium *Alcanivorax* sp. L27. Enzym Microb Technol 63:64–70

Zhang H, Liu R, Wang M, Wang H, Gao Q, Hou Z, Gao D, Wang L (2016) Draft genome sequence of *Alcanivorax* sp. strain KX64203 isolated from deep-sea sediments of Iheya North, Okinawa Trough. Genome Announc 4:e00872–e00816

2.2 The biotechnological potential of marine bacteria in the novel lineage of *Pseudomonas pertucinogena*

Alexander Bollinger, Stephan Thies, Nadine Katzke, and Karl-Erich Jaeger

Microbial Biotechnology (2020) 13, 19–31

Available online:

https://doi.org/10.1111/1751-7915.13288

Own contribution:

Participation in conceptualization and writing of the original manuscript, including introduction, discussion and the chapter on polyester hydrolases. Investigation of relevant literature, sequence analysis, data curation and formal analysis. Design of the figure and tables. Review and editing of the final manuscript.

microbial biotechnology

Open Access

Minireview

The biotechnological potential of marine bacteria in the novel lineage of *Pseudomonas pertucinogena*

Alexander Bollinger,[1] Stephan Thies,[1] Nadine Katzke[1] and Karl-Erich Jaeger[1,2,*]

[1] *Institute of Molecular Enzyme Technology, Heinrich-Heine-University Düsseldorf, Forschungszentrum Jülich, D-52425 Jülich, Germany.*

[2] *Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany.*

Summary

Marine habitats represent a prolific source for molecules of biotechnological interest. In particular, marine bacteria have attracted attention and were successfully exploited for industrial applications. Recently, a group of *Pseudomonas* species isolated from extreme habitats or living in association with algae or sponges were clustered in the newly established *Pseudomonas pertucinogena* lineage. Remarkably for the predominantly terrestrial genus *Pseudomonas*, more than half (9) of currently 16 species within this lineage were isolated from marine or saline habitats. Unlike other *Pseudomonas* species, they seem to have in common a highly specialized metabolism. Furthermore, the marine members apparently possess the capacity to produce biomolecules of biotechnological interest (e.g. dehalogenases, polyester hydrolases, transaminases). Here, we summarize the knowledge regarding the enzymatic endowment of the marine *Pseudomonas pertucinogena* bacteria and report on a genomic analysis focusing on the presence of genes encoding esterases, dehalogenases, transaminases and secondary metabolites including carbon storage compounds.

Received 10 May, 2018; revised 24 May, 2018; accepted 25 May, 2018.

*For correspondence. E-mail karl-erich.jaeger@fz-juelich.de; Tel. +49 2461 613716; Fax +49 2461 612490.

Microbial Biotechnology (2018) **0**(0), 1–13
doi:10.1111/1751-7915.13288

Funding Information
The authors receive funding from the European Union's Horizon 2020 research and innovation program (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project 'INMARE' under grant agreement No. 634486 and ERA-IB 5 'METACAT'. ST is financially supported by the Ministry of Culture and Science of the German State of North Rhine-Westphalia within in the framework of the NRW Strategieprojekt BioSC (No. 313/323-400-00213).

Introduction

The oceans cover the largest part of the earth's surface and represent one of the most diverse environments on our planet (Venter *et al.*, 2004; Armbrust and Palumbi, 2015; Tully *et al.*, 2018). Researchers have started to explore this diversity by identification and isolation of biocatalysts and secondary metabolites produced by marine organisms, opening up a novel branch of research and application termed blue biotechnology. Examples comprise not only the famous green fluorescent protein produced by the jellyfish *Aequorea victoria* and its use in innumerable applications including clinical diagnostics and therapeutics (Ohba *et al.*, 2013; Enterina *et al.*, 2015; Hoffman, 2015), but also antiviral and anticancer compounds isolated from marine sponges (Calcabrini *et al.*, 2017), biocatalysts with potential application for the production of pharmaceutical building blocks like solketal (Ferrer *et al.*, 2005), and trabectedin (supplied as Yondelis by PharmaMar S.A.) as an example for a chemotherapeutic compound produced by putative endosymbiotic bacteria of the sea squirt *Ecteinascidia turbinate* (Schofield *et al.*, 2015). Marine metagenomics and the identification and characterization of isolated marine organisms have led to a huge set of gene and genome sequences, as recently shown by the assembly of more than 2600 draft genomes from data collected during the Tara Oceans circumnavigation expedition (Tully *et al.*, 2018), as well as experimental data enabling insights into the biochemical potential of marine habitats and the respective microorganisms (Li and Qin, 2005; Kennedy *et al.*, 2008; Popovic *et al.*, 2015).

A number of marine bacteria hold great potential for biotechnological applications, for example the marine bacterium *Alcanivorax borkumensis,* which is known to play a key role in bioremediation of oil spills (Schneiker *et al.*, 2006). The strain produces a biosurfactant (Yakimov *et al.*, 1998) and possesses several genes

encoding esterases and monooxygenases of high biotechnological interest (Tchigvintsev *et al.*, 2015). Several marine *Acinetobacter sp.* were also shown to produce biosurfactants (Mnif and Ghribi, 2015), for example glycolipoproteins with useful surface-active properties (Peele *et al.*, 2016). An *Enterobacter* species isolated from a shark jaw produces rather uncommon medium-chain-length polyhydroxyalkanoates which may have biomedical applications (Wecker *et al.*, 2015). Furthermore, a range of novel antibiotics was identified from marine bacteria of diverse sources (Eom *et al.*, 2013).

Members of the genus *Pseudomonas* which belongs to the γ-proteobacteria can colonize diverse habitats and produce useful biomolecules including lipases (Jaeger *et al.*, 1996; Liu *et al.*, 2017), fluorescent proteins (Torra *et al.*, 2015), degradation pathway enzymes (Poblete-Castro *et al.*, 2012), rhamnolipids (Chong and Li, 2017), phenazines (Bilal *et al.*, 2017) and a number of heterologous secondary metabolites (Loeschcke and Thies, 2015). The main fraction of *Pseudomonas* species described so far is assigned to terrestrial habitats (Romanenko *et al.*, 2005); however, there are also reports of species isolated from marine environments (Baumann *et al.*, 1983), e.g. *P. marincola* (Romanenko *et al.*, 2008), *P. aeruginosa* (Manwar *et al.*, 2004) or *P. glareae* (Romanenko *et al.*, 2015).

The huge number of *Pseudomonas* species was phylogenetically distributed into three lineages comprising 13 groups. One and by far the smallest of these lineages, which was only recently established, consists of a single so-called *Pseudomonas pertucinogena* group. A small number of newly described *Pseudomonas* species with remarkable properties cluster within this lineage (García-Valdés and Lalucat, 2016; Peix *et al.*, 2018). Originally, it consisted of only two species, namely *P. pertucinogena* and *P. denitrificans* (Anzai *et al.*, 2000). As the classification of *P. denitrificans* is known to be ambiguous (Doudoroff *et al.*, 1974), it is difficult to assign respective studies to the correct genus and species; these studies are therefore not considered within this review. The original *P. pertucinogena* group was recently classified as a separate lineage within the genus *Pseudomonas* (Peix *et al.*, 2018) comprising nine marine (including the salt lake isolate *P. salina*) and seven non-marine members (Table 1).

Remarkably, enzymes from these bacteria were mentioned in different studies focusing on the bioinformatics identification of novel biocatalysts relevant for biotechnology applications, although most of the relevant species or genome sequences have been described only recently. We summarize here reports on biotechnologically relevant biomolecules produced by marine specimens of the *P. pertucinogena* lineage and describe their respective habitats with prevailing harsh environmental conditions. Based on the findings reported, we have examined the genetic capabilities of this group of bacteria to potentially produce biotechnologically relevant enzymes including polyester hydrolases, rare dehalogenases, ω-transaminases as well as secondary metabolites and carbon storage compounds. Apparently, bacteria of the *P. pertucinogena* lineage have the potential to produce such biotechnologically relevant biomolecules; however, this has for each case to be experimentally validated.

The *Pseudomonas pertucinogena* lineage

Representatives of the genus *Pseudomonas* are generally equipped with a versatile metabolism that is reflected by a rather large genome with sizes ranging from 4.1 (for some *P. stutzeri* strains) to more than 6 Mbp (e.g. for *P. syringae*, *P. aeruginosa*, *P. putida*, *P. protegens* and *P. fluorescens*). Hilker *et al.* reported a core genome consisting of more than 4000 open reading frames after analysing the genome sequences of 20 different *P. aeruginosa* strains (Hilker *et al.*, 2015). Similar dimensions were shown for *P. putida*; the core genome consists of at least 3386 genes and the average genome size is reported with about 6 Mbp (Udaondo *et al.*, 2016; Lopes *et al.*, 2018). An analysis of 76 newly isolated fluorescent *Pseudomonas* strains from tropical soil showed more than 5500 coding sequences per genome and an average genome size of more than 6 Mbp for the newly assembled genomes (Lopes *et al.*, 2018). In contrast, all members of the *P. pertucinogena* lineage possess, as far as known, a comparably small genome of less than 4 Mbp (Table 1, column 6), coding for about 3500 genes. These bacteria were isolated from diverse habitats, for example in association with marine sponges (Romanenko *et al.*, 2005), from the air (Azhar *et al.*, 2017), deep-sea sediments (Wang and Sun, 2016) or heavy metal contaminated soil (Zhang *et al.*, 2011) and they are distributed over a large geographical area (Fig. 1). Currently, only 16 species are assigned to this lineage, with about half of them attributed to marine environments (Table 1, marked with superscript f). Most of these species were first described during the last 10 years, but the lineage is likely to be further extended in the near future, e.g. by 16 metal resistant endophytic bacteria which appear to be near relatives of *P. sabulinigri* isolated from marshlands (Rocha *et al.*, 2016).

Many of the marine species of the *P. pertucinogena* lineage are adapted to cold environments, with reported growth at temperatures below 15°C and tolerance for moderate salt concentrations, surely related to their marine living conditions. Some species were described to live in association with aquatic plants, algae or sponges. As an example for such a symbiotic relationship,

Table 1. Bacteria belonging to the *P. pertucinogena* lineage.

Species	Habitat[a]	Origin[a]	Temperature range[b]	Salinity range[b]	Accession No[c]	Reference[d]
P. pertucinogena	Not recorded, deposit of the ATCC		n.d	n.d		Kawai and Yabuichi (1975)
P. bauzanensis	Soil from an industrial site	Bozen, South Tyrol, Italy	5–30°C	0–10%	NZ_FOGN00000000.1 NZ_FOUA00000000.1	Zhang *et al.* (2011)
P. formosensis	Food-waste compost	Taiwan	20–50°C	0–5.0%	NZ_FOYD00000000.1	Lin *et al.* (2013)
P. populi	Stems of *Populus euphratica* tree	Khiyik River, China 40°43′22″ N 85°19′18″ E	4–45°C	1–3%		Anwar *et al.* (2016)
P. saudimassiliensis	Air samples in an urban environment	Makkah, Saudi Arabia	37°C	n.d.	LM997413.1	Azhar *et al.* (2017)
P. xiamenensis	Activated sludge in sewage treatment	Xiamen, China	10–45°C	0–8%		Lai and Shao, 2008;
P. xinjiangensis	Desert sand	Xinjiang, China	4–42°C	0–6%	NZ_LT629736.1	Liu *et al.* (2009)
P. salina[f]	Salt lake	Xiaochaidan, China 37°28′53″N 95°30′19″E	4–35°C	0–12.0%		Zhong *et al.* (2015)
P. aestusnigri[f]	Crude oil-contaminated intertidal sand samples	Spain 42°46′ 29.27″ N 9°7′27.08″ W	18–37°C	2–12.5%	NZ_NBYK00000000.1	Sánchez *et al.* (2014), Gomila *et al.* (2017b)
P. litoralis[f]	Mediterranean seawater	Spain 40°27′24″N 0°31′36″E	15–37°C	0–15%	NZ_LT629748.1	Pascual *et al.* (2012)
P. ocean[f]	Deep-sea (1350 m)	Okinawa Trough, Pacific Ocean	4–41°C	0–10%	NZ_PPSK00000000.1	Wang and Sun (2016), García-Valdés *et al.* (2018)
P. pachastrellae[f]	Sponge *Pachastrella*	Philippine Sea	7–41°C	0–10%	NZ_MUBC00000000.1	Romanenko *et al.* (2005), Gomila *et al.* (2017a)
P. pelagia[f] CL-AP6 (type strain)	Antarctic green algae *Pyramimonas gelidicola* co-culture	Antarctic Ocean	4–33°C	0.5–8%	NZ_AROI00000000.1	Hwang *et al.* (2009), Koh *et al.* (2013)
58	Artic fjord	Norway, Ny Alesund			NZ_NWMT00000000.1	
P. sabulinigri[f]	Black beach sand	Soesoggak, Jeju Island, Korea	4–37°C	0–10%	NZ_LT629763.1	Kim *et al.* (2009)
P. salegens[f]	Aquatic plants of saline wetland	Gomishan saline wetland, Iran 37°03′N 54°01′E	4–35°C	1–10%	NZ_LT629787.1	Amoozegar *et al.* (2014)
P. profundi[f]	Deep-sea (1000 m)	Pacific Ocean, Mariana Trench 11°23.152′N 142° 29.062′E	4–40°C	0–10%		Sun *et al.* (2018)

a. Environment from which the species was isolated (Habitat) and geographical origin of the sample (Origin) as stated in the type strain description.
b. As stated in the respective type strain description.
c. Accession numbers of GenBank/RefSeq entries for the genomes or, in cases of draft genome sequences, the accession number of the respective master entry.
d. References for original descriptions and, if applicable, genome announcements.
e. S. Thies and A. Bollinger, unpublished data.
f. Marine isolates.

P. pelagia is discussed to depend on its host's protection against freezing to survive under the harsh conditions in the Antarctic Ocean, as antifreeze activity of the bacterium itself was not observed (Koh *et al.*, 2013). Aside from the colonization of algae, plants and sponge surfaces, contaminated environments seem to be a preferred habitat for these marine bacteria, including endophytes isolated from a heavy metal accumulating plant at a contaminated salt marshland (Rocha *et al.*, 2016) and crude oil-contamination sites (Lamendella *et al.*, 2014 and Sánchez *et al.*, 2014). The ability to degrade different hydrocarbons, the major constituents

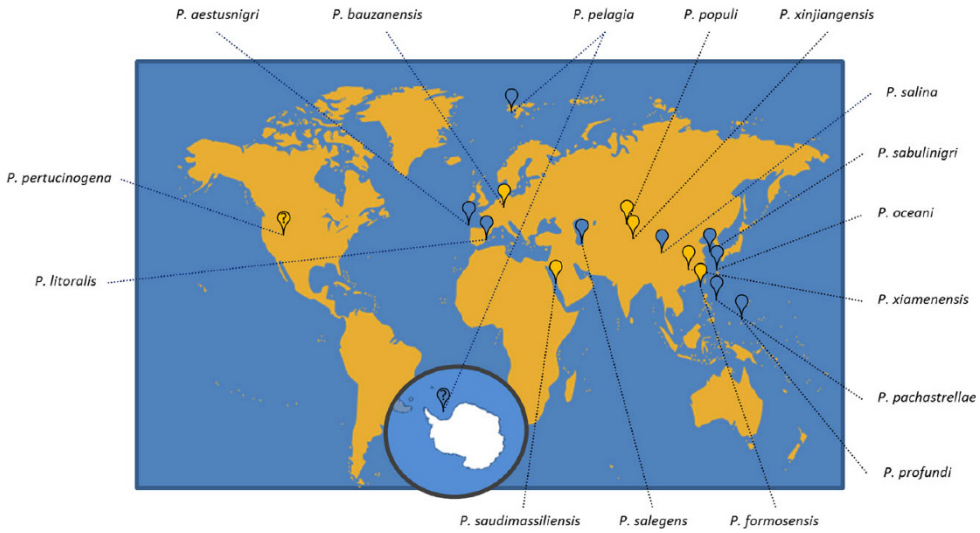


Fig. 1. Geographical distribution of 16 known bacterial species clustering in the *P. pertucinogena* lineage. Marine habitats are indicated by blue symbols, ambiguous description of the sampling site is indicated by a question mark. The Antarctic continent is pictured in a circle.

of crude oil, was proposed for *P. aestusnigri*, which possesses genes encoding phenol hydroxylase for degradation of aromatic hydrocarbons as well as an alkane-1-monooxygenase for aliphatic alkane degradation (Gomila *et al.*, 2017b).

Bacteria belonging to the genus *Pseudomonas* are generally well known for their versatile metabolism. The metabolic flexibility of Pseudomonads is reflected by their ability to use for growth a range of different carbon sources including carbohydrates, organic acids, alcohols, alkanes and most amino acids (Palleroni, 1984; Daniels *et al.*, 2010). A comprehensive phenomics analysis of *P. putida* strain DOT-T1E revealed a complex hierarchical network regulating the utilization of different carbon, nitrogen or sulfur sources (Daniels *et al.*, 2010). In contrast, members of the novel *P. pertucinogena* lineage seem to be rather limited with respect to the spectrum of utilizable carbon sources. Sánchez *et al.* described eight species of the *P. pertucinogena* group which utilize short chain carbonic acids, e.g. lactate or propionate and few amino acids, whereas they fail to utilize glucose as well as 75–80% of all tested carbon sources provided by standard phenotyping assays (Sánchez *et al.*, 2014). Undoubtedly, additional studies are needed to explore the metabolism of these bacteria in more detail. However, the currently available data suggest that strains of the *P. pertucinogena* group may constitute an exception of niche-adapted specialists within the genus *Pseudomonas*, as already suggested by the specific sites of isolation and further corroborated by their small genomes and their limited metabolic flexibility. The psychrophilic and moderate halophilic habitats indeed suggest potential for a variety of biotechnological applications of the bacteria themselves, but also their enzymes (Cavicchioli *et al.*, 2011; Cafaro *et al.*, 2013; Yin *et al.*, 2015; Danso *et al.*, 2018).

Enzymes produced by bacteria of the *P. pertucinogena* lineage have been mentioned by several studies, in particular those useful for polymer degradation and synthesis of chiral molecules (Schallmey *et al.*, 2014; Haernvall *et al.*, 2017, 2018). Enzyme coding genes were identified by mining of sequence data and subsequently expressed in heterologous hosts to prove their functionality. The results indicate that *in silico* mining of *P. pertucinogenae* genomes indeed constitutes a suitable strategy to assess the biotechnological potential of this currently still small group of *Pseudomonas* species. Hence, we have analysed the genomes of seven species of the marine *P. pertucinogena* lineage with respect to their capacity to produce biotechnologically relevant enzymes and compounds. In general, BLASTP (Altschul *et al.*, 1997) was applied to identify homologues of proteins reported in literature taking at least 40% identity to the query sequence and query coverage of at least 85 % as lower borders for selection. However, as a number of closely related species were analysed, the sequence homology of homologous proteins was significantly higher, as stated in the respective paragraphs.

These analyses revealed a considerable number of genes encoding useful enzymes as well as the potential to synthesize secondary metabolites and storage compounds (Table 2).

Biocatalysts

Enzymes are applied for a large number of different biotechnological applications, but increasingly also as (often enantioselective) biocatalysts in synthetic organic chemistry driving the development of green and sustainable processes (Sheldon and Woodley, 2018). Here, enzymes to be obtained from bacteria of the *P. pertucinogena* lineage can significantly contribute, in particular polyester hydrolases, dehalogenases and transaminases.

Polyester hydrolases

Carboxylic ester hydrolases (EC 3.1.1.) represent an important group of biocatalysts for industrial applications in a wide range of different sectors, like the pulp and paper, the pharmaceutical and the food industry (Singh *et al.*, 2016). They catalyse both the hydrolysis and the synthesis of esters, often with high enantioselectivity (Casas-Godoy *et al.*, 2012). During the last decades, the degradation of polyester compounds became more and more important due to the increasing environmental pollution with non-biodegradable polyesters such as polyethyleneterephthalate (PET) (Narancic and O'Connor, 2017). As this synthetic polyester cannot be properly recycled, more than 70% of the total plastic packaging waste may ultimately enter the food chain through inadequately treated waste water, the oceans and subsequently marine micro- and macroorganisms (Wei and Zimmermann, 2017).

The potential of *P. pelagia*, a member of the *P. pertucinogena* lineage, to effectively degrade ionic phthalic acid-based polyesters was shown recently (Haernvall *et al.*, 2017). The respective biocatalyst is a putatively secreted lipase designated as PpelaLip, which was identified by a homology guided sequence search of different extracellular hydrolases from *Pseudomonas sp.* using as a template the amino acid sequence of the *Thermobifida cellulosilytica* cutinase (Thc_Cut1), an enzyme known to efficiently hydrolyse different polyesters. After successful recombinant production and purification, the activity of the enzyme was experimentally demonstrated with different polyester substrates (Haernvall *et al.*, 2017). In a successive study, the applicability of this biocatalyst for wastewater treatment was shown (Haernvall *et al.*, 2018). We also performed a homology search with BLASTP (Altschul *et al.*, 1997) using *P. pelagia* lipase PpelaLip as a query against all published genome sequences from the *P. pertucinogena* lineage (Table 1, column 6) and identified putative proteins with 70–80% sequence identity. Interestingly, these PpelaLip homologous enzymes are not unique within the marine bacteria of the *P. pertucinogena* lineage, as we have identified homologs in all *P. pertucinogena* species (Tables 2 and 3). Apparently, representative species of the *P. aeruginosa* and *P. fluorescens* lineages also encode such enzymes; however, the overall similarity is low indicating that this type of putative polymer-degrading enzyme represents a distinct characteristic of the *P. pertucinogena* bacteria (Table 3). In a recent study of a PETase from *Ideonella sakaiensis* (Yoshida *et al.*, 2016), the classification of PETases in three groups was suggested based on amino acid sequence alignments and putative enzymes from *P. sabulinigri*, *P. pachastrellae* and *P. litoralis* were grouped into type IIa of PET-degrading enzymes (Joo *et al.*, 2018).

Dehalogenases

Dehalogenases catalyse the cleavage of carbon–halogen bonds and have potential applications in the chemical industry and for detoxification (Kurihara and Esaki, 2008). They are used for organic synthesis of optically pure building blocks, recycling of by-products from chemical processes, bioremediation and biosensing of toxic pollutants (Koudelakova *et al.*, 2013). Halohydrin dehalogenases (HHDH) represent a particularly interesting type of dehalogenases which naturally catalyse the dehalogenation of haloalcohols with the formation of the corresponding epoxides. In the reverse reaction, i.e. opening of the epoxide ring, they behave promiscuous accepting a wide range of nucleophiles such as azide, cyanide or nitrite enabling the synthesis of a wide range of chiral molecules (Schallmey *et al.*, 2014; Koopmeiners *et al.*, 2016). Schallmey *et al.* recently developed a bioinformatics pipeline to uncover these rare enzymes within sequence data sets and successfully accessed selected hits by heterologous expression and subsequent demonstration of HHDH activity. During this study, an HHDH was identified in the *P. pelagia* CL-AP6 genome and assigned to subgroup D, designated as HheD12 (Schallmey *et al.*, 2014). This same study unveiled that other *Pseudomonas* species did not encode HHDHs; and only five of 43 reported enzymes originate from γ-proteobacteria. A BLASTP analysis of the available genome data within the *P. pertucinogena* lineage with the *P. pelagia* CL-AP6 enzyme (WP_022962804.1) as a query revealed hits with identities between 91% (*P. pelagia* strain 58) and 68% (*P. salegens*) in all marine species with the exception of *P. litoralis* (Table 2). Within the terrestrial isolates, only the *P. xinjiangensis* genome encoded such an enzyme

Table 2. The catalytic and biosynthetic potential of marine *P. pertucinogena* bacteria. Genome sequences were analysed with different bioinformatics tools for the presence of polyester hydrolases (PE hydrolase), halohydrin dehalogenases (HHDH and HheD12), ω-transaminases (ω-TA), flavin-binding fluorescent proteins (FbFPp), polyhydroxyalkanoates (PHA) and ectoin synthesis clusters (Ectoin).

Species	Strain	PE hydrolase[a]	HHDH[b]	ω-TA[c]	FbFP[d]	PHA[e]	Ectoin[e]
P. aestusnigri[f]	VGXO14	WP_088276085.1	WP_088273591.1	WP_088276225.1 WP_088273722.1 WP_088276261.1	WP_088273209.1	Yes	Yes
P. litoralis[f]	2SM5	WP_090272969.1	–	WP_090274676.1 WP_090275926.1	–	No	Yes
P. pachastrellae[f]	CCUG 46540	WP_083724990.1	WP_083723433.1	WP_083728130.1	WP_083728464.1	Yes	Yes
P. pelagia[f]	58	WP_096345769.1	WP_096348266.1	WP_096346315.1 WP_096346382.1	WP_096345677.1	Yes	Yes
	CL-AP6	WP_022964382.1	WP_022962804.1 (HheD12)	WP_022961575.1 WP_022964449.1 WP_022963892.1	WP_022961159.1	Yes	Yes
P. sabulinigri[f]	JCM 14963	WP_092287377.1	WP_092284942.1	WP_092286338.1 WP_092286396.1	WP_092288528.1	Yes	Yes
P. salegens[f]	CECT 8338	WP_092388080.1	WP_092387787.1	WP_092388656.1 WP_092389204.1	WP_092383819.1	Yes	Yes
P. oceani[f]	DSM 100277	WP_104736494.1	WP_104737909.1	WP_104738025.1 WP_104739904.1 WP_104737746.1	WP_104739045.1	Yes	Yes

a. Proteins with at least 70% identity to the polyester hydrolase PpelaLip from *P. pelagia* (Haernvall *et al*., 2017).
b. Proteins with high similarity to *P. pelagia* HheD12 (Schallmey *et al*., 2014).
c. Proteins with at least 40% identity to selected known ω-TAs as query sequence and a query coverage of at least 90%.
d. Proteins with identities >60% to PpSB1-LOV (NP_746738.1), identified by BLASTP (Altschul *et al*., 1997).
e. Presence of a complete metabolite synthesis cluster predicted by the antiSMASH pipeline (Weber *et al*., 2015; Blin *et al*., 2017).
f. Marine isolates.

(identity 77%). This suggests an important role for these enzymes especially in marine environments.

ω-Transaminases

Chiral amines are valuable building blocks for a variety of compounds produced by the chemical and pharmaceutical industries. For pharmaceuticals, an estimated share of 40% contains at least one amine functionality (Kelly *et al*., 2018). While several options for enzymatic chiral amine production exist, asymmetric synthesis by ω-transaminases (ω-TA) is greatly preferred as the theoretical yield is 100% (Cassimjee *et al*., 2010; Koszelewski *et al*., 2010). Most TAs need pyridoxal-5′-phosphate as a cofactor and catalyse the asymmetric synthesis of chiral amines by transferring an amino group from an aminated donor to various carbonyl compounds (Savile *et al*., 2010; Börner *et al*., 2017; Guo and Berglund, 2017). A diverse array of reactions is reported towards the synthesis of pharmaceuticals or pharmaceutical intermediates involving ω-TAs by both asymmetric synthesis and kinetic resolution (Kelly *et al*., 2018). Thus, the identification of novel enzymes with ω-transaminase activity is of high importance for both science and industry.

Pseudomonas sp. appear to be a promising source for ω-transaminases (Wilding *et al*., 2015; Poehlein *et al*., 2017; Wu *et al*., 2017). Recently, an ω-transaminase which most probably originates from *P. sabulinigri* was discovered during an activity-based screening of a metagenomic library which originated from a polluted harbour site in Ancona, Italy (M. Ferrer, personal communication). While *P. sabulinigri* was originally identified in samples from Korean beach (Table 1), another report on closely related specimen discovered in a metalloid polluted salt marsh in the northwest coast of Portugal (Rocha *et al*., 2016) hints at the potential of this *Pseudomonas* strain to thrive in contaminated environments.

Based on this finding, a homology search with the BLASTP tool was performed to gain insights into the potential of marine representatives of the *P. pertucinogena* lineage for the production of ω-transaminases following a strategy reported earlier (Mathew and Yun, 2012). After searching the genome sequences of *P. aestusnigri*, *P. litoralis*, *P. sabulinigri*, *P. salegens*, *P. pelagia* strain 58, *P. pelagia* strain CL-AP6, *P. pachastrellae* and *P. oceani*, the results were filtered for hits with at least 40% identity to the query sequence and query coverage of at least 90%.

The amino acid sequences of the four recently characterized (*S*)-selective ω-transaminases from *P. putida* (Wu *et al*., 2017) were used as the main set of queries. The respective enzymes belong to differing families of transaminases, namely the 4-aminobutyrate pyruvate aminotransferase family (EC 2.6.1.96, BAN53958.1), beta-alanine pyruvate transaminase family (EC 2.6.1.18, BAN52522.1), aspartate aminotransferase family (EC 2.6.1.1, BAN55495.1) and putrescine-pyruvate aminotransferase (EC 2.6.1.113, BAN57107.1). BAN52522.1,

Table 3. Amino acid sequence homology expressed as identity in percentage to the known polyester hydrolase PpelaLip (Haernvall *et al.*, 2017) from *P. pelagia* strain CLAP6 as identified using BLASTP (Altschul *et al.*, 1997).

	No.	Organism	Strain	Identity in %	Protein ID
Marine	1	*P. pelagia*	CL-AP6	100	WP_022964382
	2	*P. pelagia*	58	80	WP_096345769
	3	*P. aestusnigri*	VGXO14	74	WP_088276085
	4	*P. litoralis*	2SM5	73	WP_090272969
	5	*P. pachastrellae*	JCM 12285	74	WP_083724990
	6	*P. sabulinigri*	JCM 14963	72	WP_092287377
	7	*P. sabulinigri*	JCM 14963	71	WP_092287378
	8	*P. salegens*	CECT 8338	72	WP_092388080
	9	*P. salegens*	CECT 8338	70	WP_092388077
	10	*P. oceani*	DSM 100277	74	WP_104736494
Not marine	11	*P. formosensis*	JCM 18415	73	WP_090538641
	12	*P. saudiamassiliensis*	12M76	73	WP_044499735
	13	*P. xinjiangensis*	NRRL B-51270	72	SDS09569
	14	*P. xinjiangensis*	NRRL B-51270	76	WP_093397383
Other *Pseudomonas* species	15	*P. syringae*	ICMP13650	31	KPW53696
	16	*P. aeruginosa*	PA01	28	WP_003143191
	17	*P. putida*	ATH-43	27	WP_046786320
	18	*P. protegens*	4	24	WP_102863500
	19	*P. stutzeri*	28a39	52	WP_102852227
	20	*P. fluorescens*	C3	29	WP_046049461

which is identical to the enzyme used for resolving a crystal structure (PDB 3A8U), delivered one hit with *P. litoralis*, *P. sabulinigri*, *P. salegens*, *P. pelagia* (both strains) and *P. pachastrellae*, and two hits with *P. aestusnigri* and *P. oceani*. Additional searches with sequences of BAN53958.1, BAN55495.1 and BAN57107.1 each returned the same enzyme sequence in every examined genome, as well as a second hit in the genome of *P. pelagia* CL-AP6. All hits have been annotated as aspartate aminotransferase family proteins according to the NCBI database.

Additional homology searches with sequences of (*S*)-selective ω-transaminases from *Ruegeria pomeroyi*, *Vibrio fluvialis* and *Chromobacterium violaceum,* all attributed to the aspartate aminotransferase family, returned hits with the same amino acid sequences as the aforementioned (*S*)-selective transaminases from *P. putida*, with identities up to 60%. Homology searches with (*R*)-selective transaminases from *Aspergillus fumigatus*, *Aspergillus fisheri*, *Fusarium graminearum* and *Arthrobacter* sp. according to Pavlidis *et al.* (2016) delivered no results within the search parameters. Conclusively, every searched genome contains at least two structurally different putative ω-transaminases, *P. aestusnigri*, *P. oceani* and *P. pelagia* CL-AP6 contain three such candidate enzymes (Table 2).

Flavin-binding fluorescent proteins

Flavin-binding fluorescent proteins (FbFPs) were developed as reporter proteins which constitute an oxygen-independent alternative to the family of green fluorescent proteins. They are derived from blue light photoreceptors of the L(ight)-O(xygen)-V(oltage) domain family (Drepper *et al.*, 2013; Buckley *et al.*, 2015). Besides their O_2 independence, FbFPs are small proteins (M_r: 12–16 kDa) and exhibit fast folding kinetics; thus, they are valuable reporter proteins for quantitative real-time analysis of different bio(techno)logic processes (Potzkei *et al.*, 2012; Rupprecht *et al.*, 2017). Commonly, LOV photoreceptors consist of LOV domains fused to various effector domains which are activated by conformational changes of the LOV domain in response to a light stimulus. However, predominantly among bacteria, the so-called short LOV proteins only consisting of the light-perceiving LOV receptor domain have been identified (Losi and Gärtner, 2008) with PpSB1-LOV originating from *P. putida* representing a well-studied example (Drepper *et al.*, 2007; Wingen *et al.*, 2014). As it is known that marine environments are a rich source for LOV proteins (Pathak *et al.*, 2012), we investigated the genomes of the marine *P. pertucinogena* bacteria for occurrence of short LOV proteins using a BLASTP search for homologues of PpSB1-LOV (NP_746738.1). We identified proteins with >63% identity in every marine strain with exception of *P. litoralis*. The highest identity observed was 75% for a *P. aestusnigri* protein. While it is reported that short LOV proteins occur in 10% of all *Pseudomonas* species generally (Rani *et al.*, 2013), this seems to be a remarkable frequency. Hence, the *P. pertucinogena* lineage may represent a promising source for novel FbFPs with unique properties given the fact that these bacteria can thrive in dark, cold and toxic environments.

Secondary metabolites and storage compounds

Pseudomonas species, in general, produce a number of well-studied secondary metabolites, for example rhamnolipids, phenazines, pyoverdines or syringafactins (Laursen and Nielsen, 2004; Visca *et al*., 2007; Burch *et al*., 2014; Tiso *et al*., 2017). In contrast, the secondary metabolism of the *P. pertucinogena* members still remains undiscovered, despite the eponymous pertucin produced by *P. pertucinogena* that was described to be active against phase I *Bordetella pertussis* (Kawai and Yabuichi, 1975). We therefore mined the genome data available *via* Genbank (Table 1, column 6) of the marine *P. pertucinogena* lineage organisms with respect to secondary metabolite production pathways applying the antiSMASH pipeline with enabled cluster finder algorithm (Weber *et al*., 2015; Blin *et al*., 2017). Not surprising, large gene clusters encoding modular polyketide or non-ribosomal peptide synthetases (NRPS) were very rarely detected within these comparably small-sized genomes. Only in the genome sequences of *P. sabulinigri,* an NRPS cluster was predicted that contains four adenylation domains suggesting the synthesis of a hitherto undescribed tetrapeptide derivative. Furthermore, *P. sabulingri* as well as *P. salegens* and *P. pelagia* strain 58 possess putative biosynthesis pathways for aryl-polyenes, a widespread class of antioxidant natural products (Schöner *et al*., 2016), which, however, seems to be of less importance for biotechnological applications.

The aforementioned capability of *P. pertucinogena* for pertucin production suggests that bacteriocin production might also be found as a feature of these bacteria. Bacteriocins are ribosomally produced peptides which are post-translationally processed to become antimicrobial peptides and may thus be applied by the pharma or food industries (Hassan *et al*., 2014; Yang *et al*., 2014). Bacteriocin clusters were indeed predicted in the genomes of several species (*P. sabulinigri. P. aestusnigri, P. pachastrellae, P. pelagia* CL-AP6, *P. oceani*), but they do not appear as a common feature.

In contrast, gene clusters coding for the biosynthesis of the osmoprotectant ectoin are common among the *P. pertucinogena* bacteria and were predicted in all genomes, regardless of whether marine or soil origin. Ectoines are biotechnologically produced, e.g. with *Halomonas spec*., and used as moisturizing ingredients in cosmetics (Yin *et al*., 2015; Bownik and Stępniewska, 2016). Ectoin synthesis is widespread among marine or halophilic bacteria (Yin *et al*., 2015) as it helps to cope with high salt concentrations allowing growth at salt concentrations up to 8%, and in some cases even 15%, as reported for *P. pertucinogena* lineage bacteria (Table 1, column 5).

Notably, several of the investigated genomes seem to contain elements similar to the emulsan biosynthetic pathways. Emulsan is a surface-active polymeric bioemulsifier best known from *Acinetobacter* species. Bioemulsifiers and biosurfactants are considered as interesting natural products for biotechnological applications as detergents or emulsifiers in consumer products, pharmaceutical or environmental applications (Rosenberg and Ron, 1997; Fracchia *et al*., 2014; Gudiña *et al*., 2016). Biosurfactant or bioemulsifier production would fit the observed surface activity in cultures of *P. pachastrellae* (Antoniou *et al*., 2015) and is furthermore known among bacteria living in oil-contaminated environments (Satpute *et al*., 2010; Cafaro *et al*., 2013). A MultiGeneBlast using the *Acinetobacter iwoffii* emulsan cluster (Acc. No. AJ243431.1) as input sequence (Medema *et al*., 2013) revealed that none of the marine strains contains a complete cluster; furthermore, all species lack a protein homologous to the respective polymerizing enzyme Wzy. These strains also lack lipopeptide-related non-ribosomal peptide synthetase (NRPS) clusters as well as operons with homology to the rhamnolipid synthesis genes *rhlAB* from *P. aeruginosa*; hence, the biosurfactant production capacities of the *P. pertucinogena* bacteria remain undiscovered.

Bacterial carbon storage compounds, namely triacylglycerols or wax esters and polyhydroxyalkanoates (PHA), are also of interest for biotechnology (Alvarez and Steinbüchel, 2002; Steinbüchel and Lütke-Eversloh, 2003) with the latter compounds discussed as a naturally produced alternative to common petroleum-derived polyester materials (Narancic and O'Connor, 2017). Gene loci encoding for PHA production are present in all available genomes of the marine *P. pertucinogena* bacteria according to a MultiGeneBlast analysis using the PHA locus of *P. putida* KT2440 from AE015451 as input sequence. This result confirms microscopical observations of PHA granulae within the cells reported for some of the species (Liu *et al*., 2009). Remarkably, no homologous gene clusters were identified in all terrestrial species tested. Noteworthy, all marine species contained homologs to the wax ester or triacylglyceride synthases of *Alcanivorax borkumensis* and *Marinobacter hydrocarbonoclasticus* (Alvarez, 2016). Marine *P. pertucinogena* species may thus be able to adapt their carbon storage metabolite production to the respective environmental conditions.

Conclusions

Bacterial species belonging to the recently established *P. pertucinogena* lineage are barely explored until today; nevertheless, it appears that they clearly diverge from other *Pseudomonas* species with respect

to their metabolism, genome size and, not least, environmental conditions. Notably, by applying bioinformatics tools for genome mining, we discovered that these bacteria hold a high potential for a variety of biotechnological applications. Presumably, these findings will be corroborated by further approaches of whole genome sequencing, *in silico* genome data mining, gene synthesis and expression in established hosts, which will further expand the still limited set of enzymes from already reported but also from other relevant enzyme classes, e.g. keto-reductases. In addition, the bacteria themselves may increasingly be used for biotechnological applications, in particular, psychro- and halophilic as well as hydrocarbonoclastic and heavy metal tolerant bacteria (Margesin and Feller, 2010; Cavicchioli *et al*., 2011; Cafaro *et al*., 2013; Yin *et al*., 2015). The current reports on characterized enzymes are limited to only a few marine species, but it is tempting to speculate that future studies on terrestrial species of the *P. pertucinogena* lineage may uncover such features as well; especially in species as *P. bauzanensis* which can cope with toxic contaminations. *P. pertucinogena* bacteria are easy to cultivate, at least in complex media, and can biosynthesize natural products such as PHAs, ectoin or bioemulsifiers, even with hydrocarbon pollutions or human-made polymers as alternative carbon sources, thus contributing to the saving of natural resources (Wierckx *et al*., 2015).

Acknowledgements

The authors receive funding from the European Union's Horizon 2020 research and innovation program (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project 'INMARE' under grant agreement No. 634486 and ERA-IB 5 'METACAT'. ST is financially supported by the Ministry of Culture and Science of the German State of North Rhine-Westphalia within in the framework of the NRW Strategieprojekt BioSC (No. 313/323-400-00213).

Conflict of interest

None declared.

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* **25:** 3389–3402.

Alvarez, H.M. (2016) Triacylglycerol and wax ester-accumulating machinery in prokaryotes. *Biochimie* **120:** 28–39.

Alvarez, H.M., and Steinbüchel, A. (2002) Triacylglycerols in prokaryotic microorganisms. *Appl Microbiol Biotechnol* **60:** 367–376.

Amoozegar, M.A., Shahinpei, A., Sepahy, A.A., Makhdoumi-Kakhki, A., Seyedmahdi, S.S., Schumann, P., and Ventosa, A. (2014) *Pseudomonas salegens* sp. nov., a halophilic member of the genus *Pseudomonas* isolated from a wetland. *Int J Syst Evol Microbiol* **64:** 3565–3570.

Antoniou, E., Fodelianakis, S., Korkakaki, E., and Kalogerakis, N. (2015) Biosurfactant production from marine hydrocarbon-degrading consortia and pure bacterial strains using crude oil as carbon source. *Front Microbiol* **6:** 274.

Anwar, N., Abaydulla, G., Zayadan, B., Abdurahman, M., Hamood, B., Erkin, R., *et al.* (2016) *Pseudomonas populi* sp. nov., an endophytic bacterium isolated from *Populus euphratica*. *Int J Syst Evol Microbiol* **66:** 1419–1425.

Anzai, Y., Kim, H., Park, J.Y., Wakabayashi, H., and Oyaizu, H. (2000) Phylogenetic affiliation of the pseudomonads based on 16S rRNA sequence. *Int J Syst Evol Microbiol* **50(Pt 4):** 1563–1589.

Armbrust, E.V., and Palumbi, S.R. (2015) Marine biology. Uncovering hidden worlds of ocean biodiversity. *Science* **348:** 865–867.

Azhar, E.I., Papadioti, A., Bibi, F., Ashshi, A.M., Raoult, D., and Angelakis, E. (2017) "*Pseudomonas saudimassiliensis*" sp. nov. a new bacterial species isolated from air samples in the urban environment of Makkah, Saudi Arabia. *New Microbes New Infect* **16:** 43–44.

Baumann, P., Bowditch, R.D., Baumann, L., and Beaman, B. (1983) Taxonomy of marine *pseudomonas* species: *P. stanieri* sp. nov.; *P. perfectomarina* sp. nov., nom. rev.; *P. nautica*; and *P. doudoroffii*. *Int J Syst Bacteriol* **33:** 857–865.

Bilal, M., Guo, S., Iqbal, H.M.N., Hu, H., Wang, W., and Zhang, X. (2017) Engineering *Pseudomonas* for phenazine biosynthesis, regulation, and biotechnological applications: a review. *World J Microbiol Biotechnol* **33:** 191.

Blin, K., Wolf, T., Chevrette, M.G., Lu, X., Schwalen, C.J., Kautsar, S.A., *et al.* (2017) antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* **45:** W36–W41.

Börner, T., Rämisch, S., Bartsch, S., Vogel, A., Adlercreutz, P., and Grey, C. (2017) Three in one: temperature, solvent and catalytic stability by engineering the cofactor-binding element of amine transaminase. *ChemBioChem* **18:** 1482–1486.

Bownik, A., and Stępniewska, Z. (2016) Ectoine as a promising protective agent in humans and animals. *Arch Ind Hyg Toxicol* **67:** 260–265.

Buckley, A.M., Petersen, J., Roe, A.J., Douce, G.R., and Christie, J.M. (2015) LOV-based reporters for fluorescence imaging. *Curr Opin Chem Biol* **27:** 39–45.

Burch, A.Y., Zeisler, V., Yokota, K., Schreiber, L., and Lindow, S.E. (2014) The hygroscopic biosurfactant syringafactin produced by *Pseudomonas syringae* enhances fitness on leaf surfaces during fluctuating humidity. *Environ Microbiol* **16:** 2086–2098.

Cafaro, V., Izzo, V., Notomista, E., and Di Donato, A. (2013) Marine hydrocarbonoclastic bacteria. In *Marine*

Enzymes for Biocatalysis. Trincone, A. (ed). Cambridge, UK: Woodhead Publishing, pp. 373–402.

Calcabrini, C., Catanzaro, E., Bishayee, A., Turrini, E., and Fimognari, C. (2017) Marine sponge natural products with anticancer potential: an updated review. *Mar Drugs* **15:** 310.

Casas-Godoy, L., Duquesne, S., Bordes, F., Sandoval, G., and Marty, A. (2012) Lipases: an overview. In *Lipases and Phospholipases. Methods in Molecular Biology (Methods and Protocols)*. Sandoval, G. (ed). New York, NY: Humana Press, pp. 3–30.

Cassimjee, K.E., Branneby, C., Abedi, V., Wells, A., and Berglund, P. (2010) Transaminations with isopropyl amine: equilibrium displacement with yeast alcohol dehydrogenase coupled to in situ cofactor regeneration. *Chem Commun* **46:** 5569.

Cavicchioli, R., Charlton, T., Ertan, H., Omar, S.M., Siddiqui, K.S., and Williams, T.J. (2011) Biotechnological uses of enzymes from psychrophiles. *Microb Biotechnol* **4:** 449–460.

Chong, H., and Li, Q. (2017) Microbial production of rhamnolipids: opportunities, challenges and strategies. *Microb Cell Fact* **16:** 137.

Daniels, C., Godoy, P., Duque, E., Molina-Henares, M.A., de la Torre, J., Del Arco, J.M., *et al.* (2010) Global regulation of food supply by *Pseudomonas putida* DOT-T1E. *J Bacteriol* **192:** 2169–2181.

Danso, D., Schmeisser, C., Chow, J., Zimmermann, W., Wei, R., Leggewie, C., *et al.* (2018) New insights into the function and global distribution of polyethylene terephthalate (PET) degrading bacteria and enzymes in marine and terrestrial metagenomes. *Appl Environ Microbiol* **84:** e02773-17.

Doudoroff, M., Contopoulou, R., Kunisawa, R., and Palleroni, N.J. (1974) Taxonomic validity of *Pseudomonas denitrificans* (Christensen) Bergey *et al.* request for an opinion. *Int J Syst Bacteriol* **24:** 294–300.

Drepper, T., Eggert, T., Circolone, F., Heck, A., Krauß, U., Guterl, J.-K., *et al.* (2007) Reporter proteins for *in vivo* fluorescence without oxygen. *Nat Biotechnol* **25:** 443–445.

Drepper, T., Gensch, T., and Pohl, M. (2013) Advanced *in vivo* applications of blue light photoreceptors as alternative fluorescent proteins. *Photochem Photobiol Sci* **12:** 1125–1134.

Enterina, J.R., Wu, L., and Campbell, R.E. (2015) Emerging fluorescent protein technologies. *Curr Opin Chem Biol* **27:** 10–17.

Eom, S.-H., Kim, Y.-M., and Kim, S.-K. (2013) Marine bacteria: potential sources for compounds to overcome antibiotic resistance. *Appl Microbiol Biotechnol* **97:** 4763–4773.

Ferrer, M., Golyshina, O.V., Chernikova, T.N., Khachane, A.N., Martins dos Santos, V.A.P., Yakimov, M.M., *et al.* (2005) Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. *Chem Biol* **12:** 895–904.

Fracchia, L., Ceresa, C., Franzetti, A., Cavallo, M., Gandolfi, I., Hamme, J.Van., *et al.* (2014) Industrial applications of biosurfactants. In *Biosurfactants: Production and Utilization—Processes, Technologies, and Economics*. Kosaric, N., and Sukan, F. (eds). Boca Raton, FL: CRC Press Taylor & Francis Group, pp. 245–268.

García-Valdés, E., and Lalucat, J. (2016) *Pseudomonas*: molecular phylogeny and current taxonomy. In *Pseudomonas: Molecular and Applied Biology*. Kahlon, R. S. (ed). Cham, Switzerland: Springer International Publishing, pp. 1–23.

García-Valdés, E., Gomila, M., Mulet, M., and Lalucat, J. (2018) Draft genome sequence of *Pseudomonas oceani* DSM 100277^T, a deep-sea bacterium. *Genome Announc* **6:** e00254-18.

Gomila, M., Mulet, M., Lalucat, J., and García-Valdés, E. (2017a) Draft genome sequence of *Pseudomonas pachastrellae* strain CCUG 46540^T, a deep-sea bacterium. *Genome Announc* **5:** e00136-17.

Gomila, M., Mulet, M., Lalucat, J., and García-Valdés, E. (2017b) Draft genome sequence of the marine bacterium *Pseudomonas aestusnigri* VGXO14^T. *Genome Announc* **5:** e00765-17.

Gudiña, E.J., Teixeira, J.A., and Rodrigues, L.R. (2016) Biosurfactants produced by marine microorganisms with therapeutic applications. *Mar Drugs* **14:** 38.

Guo, F., and Berglund, P. (2017) Transaminase biocatalysis: optimization and application. *Green Chem* **19:** 333–360.

Haernvall, K., Zitzenbacher, S., Wallig, K., Yamamoto, M., Schick, M.B., Ribitsch, D., and Guebitz, G.M. (2017) Hydrolysis of ionic phthalic acid based polyesters by wastewater microorganisms and their enzymes. *Environ Sci Technol* **51:** 4596–4605.

Haernvall, K., Zitzenbacher, S., Biundo, A., Yamamoto, M., Schick, M.B., Ribitsch, D., and Guebitz, G.M. (2018) Enzymes as enhancers for the biodegradation of synthetic polymers in wastewater. *ChemBioChem* **19:** 317–325.

Hassan, M., Kjos, M., and Nes, I. (2014) Antimicrobial peptides from prokaryotes. In *Novel Antimicrobial Agents and Strategies*. Phoenix, D.A., Harris, F., and Dennison, S.R. (eds). Weinheim, Germany: Wiley-VCH Verlag GmbH & Co. KGaA.

Hilker, R., Munder, A., Klockgether, J., Losada, P.M., Chouvarine, P., Cramer, N., *et al.* (2015) Interclonal gradient of virulence in the *Pseudomonas aeruginosa* pangenome from disease and environment. *Environ Microbiol* **17:** 29–46.

Hoffman, R.M. (2015) Application of GFP imaging in cancer. *Lab Invest* **95:** 432–452.

Hwang, C.Y., Zhang, G.I., Kang, S.-H., Kim, H.J., and Cho, B.C. (2009) *Pseudomonas pelagia* sp. nov., isolated from a culture of the Antarctic green alga *Pyramimonas gelidicola*. *Int J Syst Evol Microbiol* **59:** 3019–3024.

Jaeger, K.-E., Liebeton, K., Zonta, A., Schimossek, K., and Reetz, M.T. (1996) Biotechnological application of *Pseudomonas aeruginosa* lipase: efficient kinetic resolution of amines and alcohols. *Appl Microbiol Biotechnol* **46:** 99–105.

Joo, S., Cho, I.J., Seo, H., Son, H.F., Sagong, H.-Y., Shin, T.J., *et al.* (2018) Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nat Commun* **9:** 382.

Kawai, Y., and Yabuichi, E. (1975) *Pseudomonas pertucinogena* sp. nov., an organism previously misidentified as *Bordetella pertussis*. *Int J Syst Bacteriol* **25:** 317–323.

Kelly, S.A., Pohle, S., Wharry, S., Mix, S., Allen, C.C.R., Moody, T.S., and Gilmore, B.F. (2018) Application of ω-

transaminases in the pharmaceutical industry. *Chem Rev* **118:** 349–367.

Kennedy, J., Marchesi, J.R., and Dobson, A.D. (2008) Marine metagenomics: strategies for the discovery of novel enzymes with biotechnological applications from marine environments. *Microb Cell Fact* **7:** 27.

Kim, K.-H., Roh, S.W., Chang, H.-W., Nam, Y.-D., Yoon, J.-H., Jeon, C.O., *et al.* (2009) *Pseudomonas sabulinigri* sp. nov., isolated from black beach sand. *Int J Syst Evol Microbiol* **59:** 38–41.

Koh, H.Y., Jung, W., Do, H., Lee, S.G., Lee, J.H., and Kim, H.J. (2013) Draft genome sequence of *Pseudomonas pelagia* CL-AP6, a psychrotolerant bacterium isolated from culture of Antarctic green alga *Pyramimonas gelidicola*. *Genome Announc* **1:** e00699-13.

Koopmeiners, J., Halmschlag, B., Schallmey, M., and Schallmey, A. (2016) Biochemical and biocatalytic characterization of 17 novel halohydrin dehalogenases. *Appl Microbiol Biotechnol* **100:** 7517–7527.

Koszelewski, D., Tauber, K., Faber, K., and Kroutil, W. (2010) ω-Transaminases for the synthesis of non-racemic α-chiral primary amines. *Trends Biotechnol* **28:** 324–332.

Koudelakova, T., Bidmanova, S., Dvorak, P., Pavelka, A., Chaloupkova, R., Prokop, Z., and Damborsky, J. (2013) Haloalkane dehalogenases: biotechnological applications. *Biotechnol J* **8:** 32–45.

Kurihara, T., and Esaki, N. (2008) Bacterial hydrolytic dehalogenases and related enzymes: occurrences, reaction mechanisms, and applications. *Chem Rec* **8:** 67–74.

Lai, Q., and Shao, Z. (2008) *Pseudomonas xiamenensis* sp. nov., a denitrifying bacterium isolated from activated sludge. *Int J Syst Evol Microbiol* **58:** 1911–1915.

Lamendella, R., Strutt, S., Borglin, S., Chakraborty, R., Tas, N., Mason, O.U., *et al.* (2014) Assessment of the Deepwater Horizon oil spill impact on Gulf coast microbial communities. *Front Microbiol* **5:** 130.

Laursen, J.B., and Nielsen, J. (2004) Phenazine natural products: biosynthesis, synthetic analogues, and biological activity. *Chem Rev* **104:** 1663–1685.

Li, X., and Qin, L. (2005) Metagenomics-based drug discovery and marine microbial diversity. *Trends Biotechnol* **23:** 539–543.

Lin, S.-Y., Hameed, A., Liu, Y.-C., Hsu, Y.-H., Lai, W.-A., and Young, C.-C. (2013) *Pseudomonas formosensis* sp. nov., a gamma-proteobacteria isolated from food-waste compost in Taiwan. *Int J Syst Evol Microbiol* **63:** 3168–3174.

Liu, M., Luo, X., Zhang, L., Dai, J., Wang, Y., Tang, Y., *et al.* (2009) *Pseudomonas xinjiangensis* sp. nov., a moderately thermotolerant bacterium isolated from desert sand. *Int J Syst Evol Microbiol* **59:** 1286–1289.

Liu, W., Li, M., and Yan, Y. (2017) Heterologous expression and characterization of a new lipase from *Pseudomonas fluorescens* Pf0-1 and used for biodiesel production. *Sci Rep* **7:** 15711.

Loeschcke, A., and Thies, S. (2015) *Pseudomonas putida* – a versatile host for the production of natural products. *Appl Microbiol Biotechnol* **99:** 6197–6214.

Lopes, L.D., Davis, E.W., Pereira E de Silva, M.C., Weisberg, A.J., Bresciani, L., Chang, J.H., *et al.* (2018) Tropical soils are a reservoir for fluorescent *Pseudomonas* spp. biodiversity. *Environ Microbiol* **20:** 62–74.

Losi, A., and Gärtner, W. (2008) Bacterial bilin- and flavin-binding photoreceptors. *Photochem Photobiol Sci* **7:** 1168.

Manwar, A.V., Khandelwal, S.R., Chaudhari, B.L., Meyer, J.M., and Chincholkar, S.B. (2004) Siderophore production by a marine *Pseudomonas aeruginosa* and its antagonistic action against phytopathogenic fungi. *Appl Biochem Biotechnol* **118:** 243–251.

Margesin, R., and Feller, G. (2010) Biotechnological applications of psychrophiles. *Environ Technol* **31:** 835–844.

Mathew, S., and Yun, H. (2012) ω-Transaminases for the production of optically pure amines and unnatural amino acids. *ACS Catal* **2:** 993–1001.

Medema, M.H., Takano, E., and Breitling, R. (2013) Detecting sequence homology at the gene cluster level with multigeneblast. *Mol Biol Evol* **30:** 1218–1223.

Mnif, I., and Ghribi, D. (2015) High molecular weight bioemulsifiers, main properties and potential environmental and biomedical applications. *World J Microbiol Biotechnol* **31:** 691–706.

Narancic, T., and O'Connor, K.E. (2017) Microbial biotechnology addressing the plastic waste disaster. *Microb Biotechnol* **10:** 1232–1235.

Ohba, Y., Fujioka, Y., Nakada, S., and Tsuda, M. (2013) Fluorescent protein-based biosensors and their clinical applications. *Prog Mol Biol Transl Sci* **113:** 313–348.

Palleroni, N.J. (1984) *Pseudomonas*. In *Bergey's manual of systematic bacteriology*. Bergey, D.H., David, H., Krieg, N.R., and Holt, J.G. (eds). Baltimore, MD: Williams & Wilkins, pp. 141–199.

Pascual, J., Lucena, T., Ruvira, M.A., Giordano, A., Gambacorta, A., Garay, E., *et al.* (2012) *Pseudomonas litoralis* sp. nov., isolated from Mediterranean seawater. *Int J Syst Evol Microbiol* **62:** 438–444.

Pathak, G.P., Losi, A., and Gärtner, W. (2012) Metagenome-based screening reveals worldwide distribution of LOV-domain proteins. *Photochem Photobiol* **88:** 107–118.

Pavlidis, I.V., Weiß, M.S., Genz, M., Spurr, P., Hanlon, S.P., Wirz, B., *et al.* (2016) Identification of (*S*)-selective transaminases for the asymmetric synthesis of bulky chiral amines. *Nat Chem* **8:** 1076–1082.

Peele, K.A., Ch, V.R.T., and Kodali, V.P. (2016) Emulsifying activity of a biosurfactant produced by a marine bacterium. *3 Biotech* **6:** 177.

Peix, A., Ramírez-Bahena, M.H., and Velázquez, E. (2018) The current status on the taxonomy of *Pseudomonas* revisited: an update. *Infect Genet Evol* **57:** 106–116.

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.

Poehlein, A., Daniel, R., Thürmer, A., Bollinger, A., Thies, S., Katzke, N., and Jaeger, K.-E. (2017) First insights into the genome sequence of *Pseudomonas oleovorans* DSM 1045. *Genome Announc* **5:** e00774-17.

Popovic, A., Tchigvintsev, A., Tran, H., Chernikova, T.N., Golyshina, O.V., Yakimov, M.M., *et al.* (2015) Metagenomics as a tool for enzyme discovery: hydrolytic enzymes from marine-related metagenomes. In *Advances*

in Experimental Medicine and Biology. Krogan, N.J., and Babu, M. (eds). Cham, Switzerland: Springer, pp. 1–20.

Potzkei, J., Kunze, M., Drepper, T., Gensch, T., Jaeger, K.-E., and Buechs, J. (2012) Real-time determination of intracellular oxygen in bacteria using a genetically encoded FRET-based biosensor. *BMC Biol* **10:** 28.

Rani, R., Jentzsch, K., Lecher, J., Hartmann, R., Willbold, D., Jaeger, K.-E., and Krauss, U. (2013) Conservation of dark recovery kinetic parameters and structural features in the *Pseudomonadaceae* "short" light, oxygen, voltage (LOV) protein family: implications for the design of LOV-based optogenetic tools. *Biochemistry* **52:** 4460–4473.

Rocha, J., Tacão, M., Fidalgo, C., Alves, A., and Henriques, I. (2016) Diversity of endophytic *Pseudomonas* in *Halimione portulacoides* from metal(loid)-polluted salt marshes. *Environ Sci Pollut Res* **23:** 13255–13267.

Romanenko, L.A., Uchino, M., Falsen, E., Frolova, G.M., Zhukova, N.V., and Mikhailov, V.V. (2005) *Pseudomonas pachastrellae* sp. nov., isolated from a marine sponge. *Int J Syst Evol Microbiol* **55:** 919–924.

Romanenko, L.A., Uchino, M., Tebo, B.M., Tanaka, N., Frolova, G.M., and Mikhailov, V.V. (2008) *Pseudomonas marincola* sp. nov., isolated from marine environments. *Int J Syst Evol Microbiol* **58:** 706–710.

Romanenko, L.A., Tanaka, N., Svetashev, V.I., and Mikhailov, V.V. (2015) *Pseudomonas glareae* sp. nov., a marine sediment-derived bacterium with antagonistic activity. *Arch Microbiol* **197:** 693–699.

Rosenberg, E., and Ron, E.Z. (1997) Bioemulsans: microbial polymeric emulsifiers. *Curr Opin Biotechnol* **8:** 313–316.

Rupprecht, C., Wingen, M., Potzkei, J., Gensch, T., Jaeger, K.-E., and Drepper, T. (2017) A novel FbFP-based biosensor toolbox for sensitive *in vivo* determination of intracellular pH. *J Biotechnol* **258:** 25–32.

Sánchez, D., Mulet, M., Rodríguez, A.C., David, Z., Lalucat, J., and García-valdés, E. (2014) *Pseudomonas aestusnigri* sp. nov., isolated from crude oil-contaminated intertidal sand samples after the Prestige oil spill. *Syst Appl Microbiol* **37:** 89–94.

Satpute, S.K., Banat, I.M., Dhakephalkar, P.K., Banpurkar, A.G., and Chopade, B.A. (2010) Biosurfactants, bioemulsifiers and exopolysaccharides from marine microorganisms. *Biotechnol Adv* **28:** 436–450.

Savile, C.K., Janey, J.M., Mundorff, E.C., Moore, J.C., Tam, S., Jarvis, W.R., *et al.* (2010) Biocatalytic asymmetric synthesis of chiral amines from ketones applied to sitagliptin manufacture. *Science* **329:** 305–309.

Schallmey, M., Koopmeiners, J., Wells, E., Wardenga, R., and Schallmey, A. (2014) Expanding the halohydrin dehalogenase enzyme family: identification of novel enzymes by database mining. *Appl Environ Microbiol* **80:** 7303–7315.

Schneiker, S., Martins dos Santos, V.A.P., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., *et al.* (2006) Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat Biotechnol* **24:** 997–1004.

Schofield, M.M., Jain, S., Porat, D., Dick, G.J., and Sherman, D.H. (2015) Identification and analysis of the bacterial endosymbiont specialized for production of the chemotherapeutic natural product ET-743. *Environ Microbiol* **17:** 3964–3975.

Schöner, T.A., Gassel, S., Osawa, A., Tobias, N.J., Okuno, Y., Sakakibara, Y., *et al.* (2016) Aryl polyenes, a highly abundant class of bacterial natural products, are functionally related to antioxidative carotenoids. *ChemBioChem* **17:** 247–253.

Sheldon, R.A., and Woodley, J.M. (2018) Role of biocatalysis in sustainable chemistry. *Chem Rev* **118:** 801–838.

Singh, R., Kumar, M., Mittal, A., and Kumar, P. (2016) Microbial enzymes: industrial progress in 21st century. *3 Biotech* **6:** 174.

Steinbüchel, A., and Lütke-Eversloh, T. (2003) Metabolic engineering and pathway construction for biotechnological production of relevant polyhydroxyalkanoates in microorganisms. *Biochem Eng J* **16:** 81–96.

Sun, J., Wang, W., Ying, Y., Zhu, X., Liu, J., and Hao, J. (2018) *Pseudomonas profundi* sp. nov., isolated from deep-sea water. *Int J Syst Evol Microbiol* **68:** 1776–1780.

Tchigvintsev, A., Tran, H., Popovic, A., Kovacic, F., Brown, G., Flick, R., *et al.* (2015) The environment shapes microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine metagenomes. *Appl Microbiol Biotechnol* **99:** 2165–2178.

Tiso, T., Thies, S., Müller, M., Tsvetanova, L., Carraresi, L., Bröring, S., *et al.* (2017) Rhamnolipids: production, performance, and application. In *Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals*. Lee, S.Y. (ed). Cham, Switzerland: Springer International Publishing, pp. 1–37.

Torra, J., Burgos-Caminal, A., Endres, S., Wingen, M., Drepper, T., Gensch, T., *et al.* (2015) Singlet oxygen photosensitisation by the fluorescent protein Pp2FbFP L30M, a novel derivative of *Pseudomonas putida* flavin-binding Pp2FbFP. *Photochem Photobiol Sci* **14:** 280–287.

Tully, B.J., Graham, E.D., and Heidelberg, J.F. (2018) The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Sci. Data* **5:** 170203.

Udaondo, Z., Molina, L., Segura, A., Duque, E., and Ramos, J.L. (2016) Analysis of the core genome and pangenome of *Pseudomonas putida*. *Environ Microbiol* **18:** 3268–3283.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al.* (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304:** 66–74.

Visca, P., Imperi, F., and Lamont, I.L. (2007) Pyoverdine siderophores: from biogenesis to biosignificance. *Trends Microbiol* **15:** 22–30.

Wang, M., and Sun, L. (2016) *Pseudomonas oceani* sp. nov., isolated from deep seawater. *Int J Syst Evol Microbiol* **66:** 4250–4255.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., *et al.* (2015) antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43:** W237–W243.

Wecker, P., Moppert, X., Simon-Colin, C., Costa, B., and Berteaux-Lecellier, V. (2015) Discovery of a mcl-PHA with unexpected biotechnical properties: the marine environment of French Polynesia as a source for PHA-producing bacteria. *AMB Express* **5:** 74.

Wei, R., and Zimmermann, W. (2017) Biocatalysis as a green route for recycling the recalcitrant plastic polyethylene terephthalate. *Microb Biotechnol* **10:** 1302–1307.

Wierckx, N., Prieto, M.A., Pomposiello, P., de Lorenzo, V., O'Connor, K., and Blank, L.M. (2015) Plastic waste as a novel substrate for industrial biotechnology. *Microb Biotechnol* **8:** 900–903.

Wilding, M., Walsh, E.F.A., Dorrian, S.J., and Scott, C. (2015) Identification of novel transaminases from a 12-aminododecanoic acid-metabolizing *Pseudomonas* strain. *Microb Biotechnol* **8:** 665–672.

Wingen, M., Potzkei, J., Endres, S., Casini, G., Rupprecht, C., Fahlke, C., *et al.* (2014) The photophysics of LOV-based fluorescent proteins – new tools for cell biology. *Photochem Photobiol Sci* **13:** 875–883.

Wu, H.-L., Zhang, J.-D., Zhang, C.-F., Fan, X.-J., Chang, H.-H., and Wei, W.-L. (2017) Characterization of four new distinct ω-transaminases from *Pseudomonas putida* nbrc 14164 for kinetic resolution of racemic amines and amino alcohols. *Appl Biochem Biotechnol* **181:** 972–985.

Yakimov, M.M., Golyshin, P.N., Lang, S., Moore, E.R., Abraham, W.R., Lünsdorf, H., and Timmis, K.N. (1998) *Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing marine bacterium. *Int J Syst Bacteriol* **48(Pt 2):** 339–348.

Yang, S.-C., Lin, C.-H., Sung, C.T., and Fang, J.-Y. (2014) Antibacterial activities of bacteriocins: application in foods and pharmaceuticals. *Front Microbiol* **5:** 241.

Yin, J., Chen, J.-C., Wu, Q., and Chen, G.-Q. (2015) Halophiles, coming stars for industrial biotechnology. *Biotechnol Adv* **33:** 1433–1442.

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., *et al.* (2016) A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **351:** 1–5.

Zhang, D.-C., Liu, H.-C., Zhou, Y.-G., Schinner, F., and Margesin, R. (2011) *Pseudomonas bauzanensis* sp. nov., isolated from soil. *Int J Syst Evol Microbiol* **61:** 2333–2337.

Zhong, Z.-P., Liu, H.-C., Zhou, Y.-G., Hou, T.-T., Liu, Z.-P., Liu, Y., and Wang, F. (2015) *Pseudomonas salina* sp. nov., isolated from a salt lake. *Int J Syst Evol Microbiol* **65:** 2846–2851.

2.3 Determinants and prediction of esterase substrate promiscuity patterns

Mónica Martínez-Martínez, Cristina Coscolín, Gerard Santiago, Jennifer Chow, Peter J. Stogios, Rafael Bargiela, Christoph Gertler, JoséNavarro-Fernández, Alexander Bollinger, Stephan Thies, Celia Méndez-García, Ana Popovic, Greg Brown, Tatyana N. Chernikova, Antonio García-Moyano, Gro E. K. Bjerga, Pablo Pérez-García, Tran Hai, Mercedes V. Del Pozo, Runar Stokke, Ida H. Steen, Hong Cui, Xiaohui Xu, Boguslaw P. Nocek, María Alcaide, Marco Distaso, Victoria Mesa, Ana I. Peláez, Jesús Sánchez, Patrick C. F. Buchholz, Jürgen Pleiss, Antonio Fernández-Guerra, Frank O. Glöckner, Olga V. Golyshina, Michail M. Yakimov, Alexei Savchenko, Karl-Erich Jaeger, Alexander F. Yakunin, Wolfgang R. Streit, Peter N. Golyshin, Víctor Guallar, Manuel Ferrer, and The INMARE Consortium

ACS Chem. Biol. (2018), 13, 225−234

Available online:

https://doi.org/10.1021/acschembio.7b00996

Own contribution:

Investigation: construction and screening of two genomic libraries originating from *Alcanivorax borkumensis* SK2 and *Pseudomonas oleovorans* DSM1045. Identification, cloning, expression and verification of 8 novel esterase coding genes. Writing: reviewing and editing the manuscript.

ACS chemical biology

Articles

Cite This: ACS Chem. Biol. 2018, 13, 225−234

Determinants and Prediction of Esterase Substrate Promiscuity Patterns

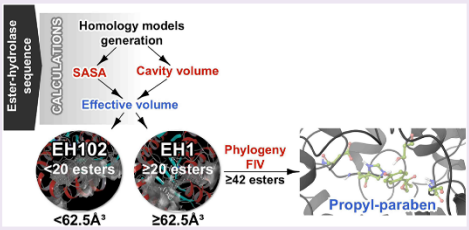
Mónica Martínez-Martínez,[†,ζ] Cristina Coscolín,[†,ζ] Gerard Santiago,[‡,ζ] Jennifer Chow,[§] Peter J. Stogios,[∥] Rafael Bargiela,[†,γ] Christoph Gertler,[⊥,δ] José Navarro-Fernández,[÷] Alexander Bollinger,[#] Stephan Thies,[#] Celia Méndez-García,[∇,ε] Ana Popovic,[∥] Greg Brown,[∥] Tatyana N. Chernikova,[⊥] Antonio García-Moyano,[○] Gro E. K. Bjerga,[○] Pablo Pérez-García,[§] Tran Hai,[⊥] Mercedes V. Del Pozo,[÷] Runar Stokke,[◆] Ida H. Steen,[◆] Hong Cui,[∥] Xiaohui Xu,[∥] Boguslaw P. Nocek,[¶] María Alcaide,[†] Marco Distaso,[⊥] Victoria Mesa,[∇] Ana I. Peláez,[∇] Jesús Sánchez,[∇] Patrick C. F. Buchholz,[%] Jürgen Pleiss,[%] Antonio Fernández-Guerra,[$,●,■] Frank O. Glöckner,[$,●] Olga V. Golyshina,[⊥] Michail M. Yakimov,[⌀,★] Alexei Savchenko,[∥] Karl-Erich Jaeger,[#,α] Alexander F. Yakunin,[∥,ψ] Wolfgang R. Streit,[§,ψ] Peter N. Golyshin,[⊥,ψ] Víctor Guallar,[*,‡,β,ψ] Manuel Ferrer,[*,†,ψ] and The INMARE Consortium

[†]Institute of Catalysis, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain
[‡]Barcelona Supercomputing Center (BSC), 08034 Barcelona, Spain
[§]Biozentrum Klein Flottbek, Mikrobiologie & Biotechnologie, Universität Hamburg, 22609 Hamburg, Germany
[∥]Department of Chemical Engineering and Applied Chemistry, University of Toronto, M5S 3E5 Toronto, Ontario, Canada
[⊥]School of Biological Sciences, Bangor University, LL57 2UW Bangor, United Kingdom
[#]Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Düsseldorf, 52425 Jülich, Germany
[∇]Department of Functional Biology-IUBA, Universidad de Oviedo, 33006 Oviedo, Spain
[○]Uni Research AS, Center for Applied Biotechnology, 5006 Bergen, Norway
[◆]Department of Biology and KG Jebsen Centre for Deep Sea Research, University of Bergen, 5020 Bergen, Norway
[¶]Structural Biology Center, Biosciences Division, Argonne National Laboratory, Argonne, 60439 Illinois, United States
[%]Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, 70569 Stuttgart, Germany
[$]Jacobs University Bremen gGmbH, Bremen, Germany
[●]Max Planck Institute for Marine Microbiology, 28359 Bremen, Germany
[■]University of Oxford, Oxford e-Research Centre, Oxford, United Kingdom
[⌀]Institute for Coastal Marine Environment, Consiglio Nazionale delle Ricerche, 98122 Messina, Italy
[★]Immanuel Kant Baltic Federal University, 236041 Kaliningrad, Russia
[α]Institute for Bio- and Geosciences IBG-1: Biotechnology, Forschunsgzentrum Jülich GmbH, 52425 Jülich, Germany
[β]Institució Catalana de Recerca i Estudis Avançats (ICREA), 08010 Barcelona, Spain

Supporting Information

ABSTRACT: Esterases receive special attention because of their wide distribution in biological systems and environments and their importance for physiology and chemical synthesis. The prediction of esterases' substrate promiscuity level from sequence data and the molecular reasons why certain such enzymes are more promiscuous than others remain to be elucidated. This limits the surveillance of the sequence space for esterases potentially leading to new versatile biocatalysts and new insights into their role in cellular function. Here, we performed an extensive analysis of the substrate spectra of

continued...



Received: November 21, 2017
Accepted: November 28, 2017
Published: November 28, 2017

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

145 phylogenetically and environmentally diverse microbial esterases, when tested with 96 diverse esters. We determined the primary factors shaping their substrate range by analyzing substrate range patterns in combination with structural analysis and protein–ligand simulations. We found a structural parameter that helps rank (classify) the promiscuity level of esterases from sequence data at 94% accuracy. This parameter, the active site effective volume, exemplifies the topology of the catalytic environment by measuring the active site cavity volume corrected by the relative solvent accessible surface area (SASA) of the catalytic triad. Sequences encoding esterases with active site effective volumes (cavity volume/SASA) above a threshold show greater substrate spectra, which can be further extended in combination with phylogenetic data. This measure provides also a valuable tool for interrogating substrates capable of being converted. This measure, found to be transferred to phosphatases of the haloalkanoic acid dehalogenase superfamily and possibly other enzymatic systems, represents a powerful tool for low-cost bioprospecting for esterases with broad substrate ranges, in large scale sequence data sets.

Enzymes with outstanding properties in biological systems and the conditions favoring their positive selection are difficult to predict. One of these properties is substrate promiscuity, which typically refers to a broad substrate spectrum and acceptance of larger substrates. This phenomenon is important from environmental,[1] evolutionary,[2−5] structural,[6−8] and biotechnological[9,10] points of view. The relevance of substrate promiscuity is indisputable as the operating basis for biological processes and cell function. As an example, the evolutionary progress of enzymes from lower to higher substrate specificity allows the recruitment of alternate pathways for carbon cycling and innovations across metabolic subsystems and the tree of life by maximizing the growth rate and growth efficiency.[11] Promiscuous enzymes are energetically more favorable than specialized enzymes,[4] and therefore, the cell does not require many different enzymes to take up substrates, favoring genome minimization and streamlining.[12] In addition, the acquisition of new specificities without compromising primary or ancestral ones is a major driver of microbial adaptation to extreme habitats.[13] From a more practical standpoint, along with requirements of a technical nature such as selectivity, scalability and robustness, a narrow substrate spectrum is one of the most frequent problems for industrial enzyme applications.[14] A consensus exists that "the more substrates an enzyme converts the better," opening application ranges with consequent reduction of the production cost of multiple enzymes.[10,14,15]

Enzymes with wide substrate ranges occur naturally, as systematically investigated for halo-alkane dehalogenases,[16] phosphatases,[1] beta-lactamases,[2,17] and hydroxyl-nitrile lyases.[5] Some enzymes are more promiscuous than others simply due to their fold or degree of plasticity or the presence of structural elements or mutations occurring under selection in the proximity of the active-site cavity and access tunnels favoring promiscuity. However, the general explanation, if any, by which an enzyme binds and converts multiple substrates is unknown, although molecular insights have been reported for single enzymes.[18] A tool that can clearly distinguish promiscuous versus non-promiscuous enzymes and suggest substrates potentially being converted or not by them might therefore be valuable in applying low-cost sequencing in discovery platforms in any biological context.

In an ideal scenario, functional characterization of enzymes with genomics[19] and metagenomics[10,20] techniques using a large library of substrates would guide the analysis of sequence-to-promiscuity relationships and explore the mechanistic basis of promiscuity. In addition, such studies may help identify a new generation of highly promiscuous microbial biocatalysts. However, extensive bioprospecting and biochemical studies are rare,[10] despite the growing number of sequences available through low-cost sequencing efforts[21] and the growing number of enzymes that are typically characterized with limited substrate sets.[14] To address this knowledge gap, we functionally assessed the substrate specificity of a set of 145 phylogenetically, environmentally, and structurally diverse microbial esterases (herein referred to as "EH," which means Ester Hydrolase) against a customized library of 96 different substrates to find predictive markers of substrate promiscuity rather than discrete determinants of substrate specificity that may differ from protein to protein. EHs were selected for an analysis of substrate promiscuity because they typically have specific definitions of molecular function, can be easily screened in genomes and metagenomes compared with many other classes of proteins, are among the most important groups of biocatalysts for chemical synthesis, and are widely distributed in nature, with at least one EH per genome.[14]

Our work adds important insights and empirical, structural, and computational data to facilitate the elucidation of the molecular basis of substrate promiscuity in EHs, which was further extended to phosphatases from the haloalkanoic acid dehalogenase (HAD) superfamily. This was achieved by deciphering what we consider a predictive structural marker of substrate promiscuity and by establishing the reasons why certain such enzymes are more promiscuous than others and can convert substrates that others cannot. This study does not pretend to generate a quantitative measure to predict the number of compounds that an enzyme will hydrolyze but a tool and a parameter that will help in ranking (classifying) promiscuity level. Following on from that, we propose in this work the first molecular classification method of this kind derived from first principle molecular simulations and with clear physical/structural interpretation. This work also provides an example of the utility of this parameter to screen the sequence space for highly promiscuous EHs that may compete with best commercial EH preparations. We also provide first preliminary evidence of a number of underexplored microbial phylogenetic lineages containing EHs with a prominent substrate range.

■ RESULTS AND DISCUSSION

The Substrate Range of 145 Diverse EHs. A total of 145 EHs were investigated. Extensive details of the sources and screen methods are provided in the Supporting Information Methods and Table S1. In an environmental context, the source of enzymes was highly diverse because they were isolated from bacteria from 28 geographically distinct sites (125 EHs in total) and from six marine bacterial genomes (20 EHs; Supporting Information Figure S1). A phylogenetic analysis also indicated that sequences belong to bacteria distributed across the entire phylogenetic tree (Supporting Information Results and Figure S2).

The 145 putative proteins exhibited maximum amino acid sequence identities (Supporting Information Table S1) ranging from 29.1 to 99.9% to uncharacterized homologous proteins in public databases, with an average value (reported as %, with the

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

interquartile range (IQR) in parentheses) of 74.3% (40.3%). The pairwise amino acid sequence identity for all EHs ranged from 0.2 to 99.7% (Supporting Information Table S2), with an average value of 13.7% (7.6%). BLAST searches were performed for all query sequences by running NCBI BLASTP against the current version of the Lipase Engineering Database[22] using an E-value threshold of 10^{-10} and were successful for all but nine candidates. A total of 120 EH sequences were unambiguously assigned to some of the 14 existing families (F) of the Arpigny and Jaeger classification, which are defined based on amino acid sequence similarity and the presence of specific sequence motifs.[14,23] These EHs included sequences with a typical α/β hydrolase fold and conserved G-X-S-X-G (FI, 20; FIV, 36; FV, 33; FVI, 5; and FVII, 6) or G-X-S-(L) (FII, 9) motifs and sequences with a serine beta-lactamase-like modular (non-α/β hydrolase fold) architecture and a conserved S-X-X-K motif (FVIII, 11). An additional set of nine sequences were assigned to the *meta*-cleavage product (MCP) hydrolase family[24] and six to the so-called carbohydrate esterase family,[25] both with typical α/β hydrolase folds. Finally, one was a cyclase-like protein from the amido-hydrolase superfamily.[26] Sequences-to-family assignments are summarized in the Supporting Information Table S1. Taken together, the primary sequence analysis suggests that the diversity of polypeptides is not dominated by a particular type of protein or highly similar protein clusters but consists of diverse nonredundant sequences assigned to multiple folds and subfamilies, which are distantly related to known homologues in many cases.

The substrate profiles of all EHs were examined using a set of 96 chemically and structurally distinct esters (Supporting Information Table S3). We are aware that the number of compounds hydrolyzed may be an ambiguous indicator of promiscuity, because the size and composition of the library may influence the results. For this reason, the composition of the library was not random but based on including esters with variation in size of acyl and alcohol groups and with growing residues (aromatic, aliphatic, branched, and unbranched) at both sides, leading to more challenging substrates because a larger group adjacent to the ester bond increases the difficulty of conversion. Halogenated, chiral, and sugar esters, lactones, and an alkyl diester were also included. Esters with nitro substituents were not included. We used the partitioning coefficient (log P value) to indicate the chemical variability of the esters because this parameter reflects electronic and steric effects and hydrophobic and hydrophilic characteristics. Log P was determined with the software ACD/ChemSketch 2015.2.5. Log P values (Supporting Information, Table S3) ranged from −1.07 (for methyl glycolate) to 23.71 (for triolein), with an average value (IQR in parentheses) of 3.13 (2.86), which indicates that the ester library used in this study had broad chemical and structural variability. Nevertheless, adding new substrates could surely help (and even change) the ranking of the EHs herein analyzed. The dynamic range of the assay may also influence the results. For this reason, to detect enzyme−substrate pairs for a given EH, the ester library was screened with each of the 145 EHs in a kinetic pH indicator assay in 384-well plates,[24,27,28] which unambiguously allow quantifying specific activities at pH 8.0 and 30 °C, using a substrate concentration above 0.5 mM (see Supporting Information, Results). Two commercial lipases, CalA and CalB from *Pseudozyma aphidis* (formerly *Candida antarctica*), were included in the assays for comparison. Using this data set, we linked the biocatalytic data to the sequence information for the respective enzyme. In this study, sequence information meant any sequence that encoded an EH of interest. Biocatalytic data meant experimental data on substrate conversion (i.e., units g^{-1} or U g^{-1}) followed for 24 h.

We determined the probability of finding an EH with a broad substrate profile by plotting the number of esters that were hydrolyzed by all preparations. Figure 1 shows that the number

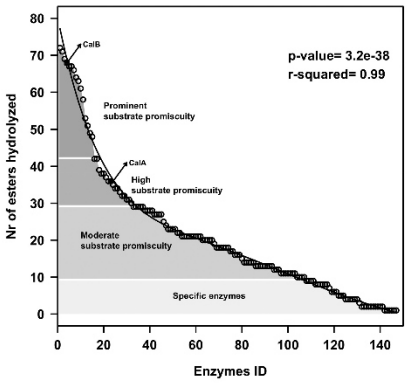


Figure 1. Number of ester substrates hydrolyzed by each of the 145 EHs investigated in this study. The commercial preparations CalA and CalB (marked with filled square) are also included. This figure is created from data in the Supporting Information Table S1. The activity protocol established and used to identify the esters hydrolyzed by each EH was based on a 550 nm follow-up pH indicator assay described in the Supporting Information Methods. The list of the 96 structurally different esters tested is shown in Figure 2. Full details of the activity protocol are provided in the Supporting Information Methods. The trend line shows a not-single exponential fit of the experimental data. The fit was obtained using R script and the "lm" function, to extract a polynomial regression of degree 6 with the following line "model ← lm(MM[,1] ~ poly(MM[,2],6,raw = TRUE))", where MM[,1] corresponds to the number of esters hydrolyzed, and MM[,2] the position in the *x* axis (from 1 to 147).

of esters hydrolyzed by all 147 EHs (including CalA/B) fits to an exponential distribution (r^2 = 0.99; *p* value $3.2e^{-38}$; Pearson's correlation coefficient) with a median of 18 substrates per enzyme, nine hits at the 25th percentile, and 29 hits at the 75th percentile. On the basis of this distribution and a previously established criterion,[1] we considered an enzyme specific if it used nine esters or fewer (27% of the total), as showing moderate substrate promiscuity if it used between 10 and 29 esters (51% of the total), and as showing high-to-prominent promiscuity if it used 30 or more esters (22% of the total). This criterion indicated a percentage of EHs with a prominent substrate range similar to that found for HAD phosphatases (24%).[1]

Phylogeny Is a Predictive Marker of Substrate Promiscuity. Hierarchical clustering was performed to evaluate the differences in substrate range patterns (Figure 2). For the sake of simplicity, clustering was performed for those EHs that hydrolyzed 10 or more esters (i.e., 107 total EHs). We first observed a large percentage of enzymes with presumptive broad active site environments that accommodated large aromatic and sterically hindered esters such as benzyl (*R*)-(+)-2-hydroxy-3-phenylpropionate (49% of the total), benzoic acid-4-formyl-phenylmethyl ester (27%), 2,4-dichlorophenyl 2,4-dichlorobenzoate (~8%), 2,4-dichlorophenyl 2,4-dichlorobenzoate (~5%), and diethyl-2,6-dimethyl 4-phenyl-1,4-dihydro

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

ACS Chemical Biology Articles

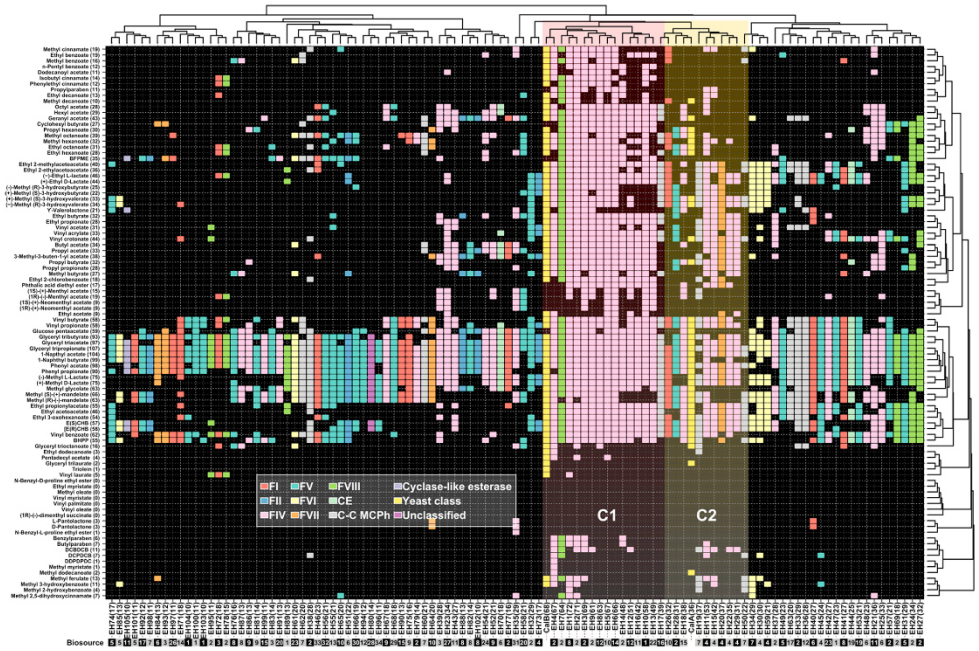


Figure 2. Hierarchical clustering of the substrate ranges of the EHs. Only EHs that hydrolyzed 10 or more esters were considered (107 in total, including CalA/B). This figure is created from data in the Supporting Information Table S3. The specific activities of the EHs for each of the 96 esters were determined as described in Figure 1. The list of the 96 esters tested and the frequency of each ester considered as a hit (in brackets) are shown on the left side. The ID code representing each EH is given at the bottom. Each hydrolase is named based on the code "EH," which means Ester Hydrolase, followed by an arbitrary number from 1 to 145 for the most to least promiscuous enzyme. The number in brackets indicates the number of esters hydrolyzed by each enzyme. The biosource of each EH is indicated at the bottom with a number in white or black squares that follows the nomenclature in the Supporting Information Figure S1. The figure was created with the R language console using a binomial table with information about the activity/inactivity (1/0) of the analyzed enzymes against the 96 substrates as a starting point. For the central graphic, which shows the data in Supporting Information Table S3, we used the drawing tools provided by the basic core packages of R. The hierarchical clusters of the enzymes (shown at the top) and substrates (shown on the right side) were generated by calculating a distance matrix using a "binomial" method and the hclust function to generate the tree. Using the functions as .phylo and plot.phylo from the ape package, the clusters were added to the top and right of the figure. A combination of the Set1 palette from the R package RColorBrewer and colors from the basic palette from R were used as the color palette for sequences assigned to each family (F; see inset), including FI to FVII, carbohydrate esterase (CE), and carbon−carbon *meta*-cleavage product hydrolase (C−C MCP) families, all with a typical α/β hydrolase fold, FVIII serine beta-lactamase with non α/β hydrolase fold, and cyclase-like protein from the amido-hydrolase superfamily. Sequences that were not unambiguously ascribed to existing families were referred to as "Unclassified," and those of yeast origin were assigned to "yeast class." The two "clusters" C1 and C2 that contained the most substrate-promiscuous EHs are color-coded under a shadowed background.

pyridine-3,5-dicarboxylate (~1%). Therefore, even though the EHs in this study were identified by a selection process based on the utilization of short esters (see Supporting Information Methods), the isolation of EHs with ample substrate spectra and the ability to hydrolyze very large substrates was not compromised.

We detected drastic shifts in substrate specificity (Figure 2), with glyceryl tripropionate as the only substrate hydrolyzed by all EHs. This is consistent with the high sequence variability within EHs, with an average pairwise identity of 13.74%. We then sought to determine the primary factors shaping the substrate range and thus defined different functional clusters. First, we observed that global sequence identity was of limited relevance for inferring the substrate range because no correlation was found (r^2 = 0.25) between the differences in identity and the number of esters that were hydrolyzed (Supporting Information Tables S1 and S2). Second, comparisons of the substrate range and the hydrolysis rate (U g^{-1} for the best substrates) were performed (Supporting Information Table S1). No correlation existed (r^2 = 0.073), suggesting that our assay conditions allow evaluating the promiscuity level whatever the hydrolytic rate of the EH is. In addition to the low correlation values, no threshold above or below which one could qualitatively classify the substrate range was observed in both cases, so that sequence identity and hydrolytic rate are neither predictive nor classification parameters of promiscuity. Additionally, no link between substrate range and habitat was found because EHs from the same biosource fell into separate clusters (Figure 2). Phylogeny-substrate spectrum relationships were further examined. Figure 2 indicates that the broad substrate-spectrum EHs did not cluster in a single phylogenetic branch, yet substrate promiscuity was mostly found for members of one of 10 subfamilies covered. Indeed, 67% of the EHs that could hydrolyze 30 or more esters

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

(mostly located in clusters C1 and C2 in Figure 2) were assigned to FIV,[14,23] and this percentage increased to 84% when considering only those EHs that could hydrolyze 42 to 72 esters (Figure 2; cluster C1). In addition to FIV members, a FVIII serine beta-lactamase showed prominent substrate spectra (see cluster 1). Members of both families (FIV, 8; FVIII, 1; see cluster C1) hydrolyzed as many esters (from 61 to 72) as the yeast family member CalB (68 esters), the most promiscuous commercially available lipase preparation used for the production of fine chemicals.[29]

Phylogeny was thus indicated as a predictive marker of the substrate range of EHs, as although a broad substrate scope was assigned to several sequence clusters, this feature was prevalent in members of FIV. A query sequence that matched FIV could be easily identified by means of the consensus motif GDSAGG around the catalytic serine; this family is also called the hormone-sensitive lipase (HSL) family because a number of FIV EHs display a striking similarity to the mammalian HSL.[14,23] Noticeably, the location of some FIV members in functional clusters with narrow substrate spectra (Figure 2) suggests that factors other than phylogeny contribute to the substrate spectra of EHs.

The Active Site Effective Volume Is a Prominent Marker of EH Promiscuity. Structural-to-substrate spectrum relationships were further examined by protein−ligand simulations to find additional markers of promiscuity. Crystals from recombinant EH1,[28] the protein with the broadest substrate range under our assay conditions, were obtained as described in the Supporting Information Methods. The enzyme with the widest substrate range was considered the best candidate for understanding the nature of promiscuity. This enzyme seems to have a wide active site environment as, under our assay conditions, it accepted 72 esters ranging from short (*e.g.*, vinyl acetate) to large (*e.g.*, 2,4-dichlorobenzyl-2,4-dichlorobenzoate; Figure 2). We also obtained crystals of recombinant EH102, which was isolated from the same habitat[28] but had a restricted substrate range, hydrolyzing only 10 of the 96 esters tested (Figure 2). Crystallographic data and refinement statistics for the two structures are given in Supporting Information Table S4.

To rationalize the substrate range shown by EH1 and EH102, we performed substrate migration studies using the software Protein Energy Landscape Exploration (PELE), which is an excellent tool to map ligand migration and binding, as shown in studies with diverse applications.[30−32] To map the tendency of a substrate to remain close to the catalytic triad, the substrate was placed in a catalytic position, within a proton abstraction distance from the catalytic serine, and allowed to freely explore the exit from the active site. The PELE results for both proteins and glyceryl triacetate are shown in Figure 3a. Clearly, EH1 has a significantly better binding profile, with an overall lower binding energy and a better funnel shape, whereas EH102 had a qualitatively unproductive binding-energy profile. This difference in the binding mechanism can be explained by the catalytic triad environment. EH1 has a somewhat wide but buried active site, whereas EH102 has a surface-exposed catalytic triad (Figure 4a). These structural differences translate into significant changes in the active site volume, as defined using Fpocket; the active site cavity of EH1 is 3-fold larger than that of EH102. Moreover, important changes are observed when inspecting the solvent exposure of the cavity. Figure 3b shows the relative solvent accessible surface area (SASA) for the substrate along the exploration of PELE, computed as a (dimensionless) percentage (0−1) of the ligand SASA in solution. Even at catalytic positions (distance Ser(O)−substrate(C) ∼ 3−4 Å), in EH102 we observe that ∼40% of the surface of the substrate is accessible to the solvent, which greatly destabilizes the substrate and facilitates escape to the bulk solvent. By contrast, EH1 has a larger but almost fully occluded site, with relative SASA values of approximately 0−10%, which can better stabilize the substrate.

After defining key points underlying the promiscuity of EH1, i.e., a larger active site volume and a lower SASA (Figure 4a), we extended the analysis to other EHs. First, we collected all 11 available crystal structures (Supporting Information Table S1) and computed the active site volume and relative SASA of the catalytic triad (Figure 5, square symbols). We next extended the analysis to the rest of the EHs using homology modeling (using the 11 crystals available) and produced a structural model for 84 additional enzymes. The missing ones were those with sequence identities of less than 25% (to an existing crystal) or those for which the catalytic triad could not be unambiguously identified (*i.e.*, not suitable alignments). Figure 5 (circle symbols) shows the active site effective volume data for all structural models. The analysis indicated a ratio threshold of 62.5 $Å^3$ for qualitatively classifying substrate promiscuity. Note that the relative SASA of the catalytic triad (derived from the GetArea server, see Supporting Information Methods) adopts values of 0−100; the actual value of the effective volume threshold will depend on the chosen range. We observed that values equal to or higher

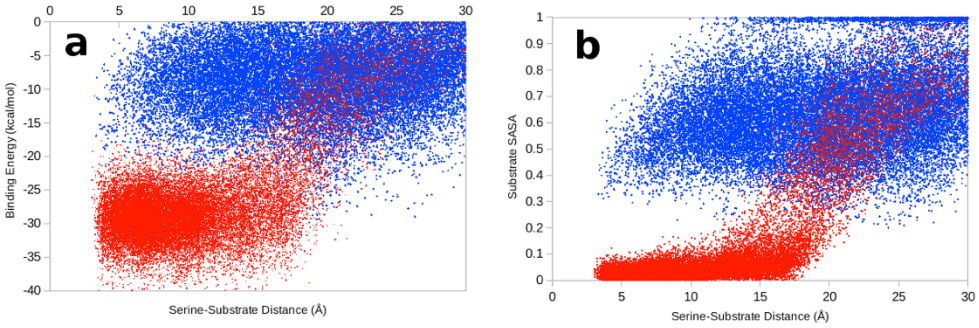


Figure 3. Protein Energy Landscape Exploration (PELE) analysis. Panel a shows the protein−substrate interaction plots for EH1 (red) and EH102 (blue). Panel b shows the relative SASA for glyceryl triacetate in EH1 (red) and EH102 (blue) computed as a dimensionless ratio (0−1) using PELE.

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

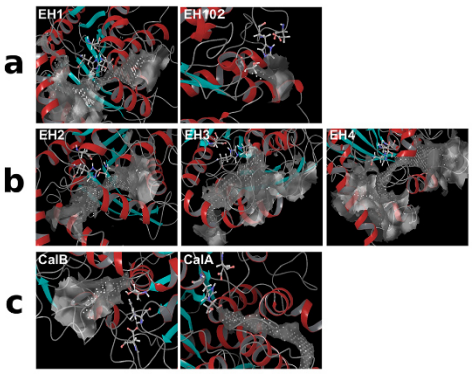


Figure 4. Catalytic triad exposure of selected EHs with the broadest and lowest substrate ranges. (a) The catalytic triad (ball-and-sticks) and the main adjacent cavity (gray clouds) as detected by SiteMap are underlined to demonstrate the differences between a promiscuous (EH1) and nonpromiscuous (EH102) EHs. EH1 can hydrolyze 72 esters and has a defined hidden binding cavity (effective volume: 166.7 Å^3). EH102, by contrast, can hydrolyze only 10 esters and has a surface-exposed triad (high SASA) and an almost negligible binding cavity (38.5 Å^3). The three top EHs with the broadest substrate ranges (b), positioned in the ranking after EH1, and the commercial CalB and CalA lipases (c), are also represented. On each panel, we highlight the catalytic triad and the main adjacent cavity as detected by SiteMap, demonstrating the differences in active site topology. EH2, EH3, and EH4, all assigned to FIV (as EH1), hydrolyzed 71, 69, and 67 esters and have defined but distinct hidden binding cavities (500, 200, and 200 Å^3, in the same order), as EH1. CalB, which was capable of hydrolyzing 68 esters, has a binding cavity (200 Å^3) that is also hidden but highly different from those of the other EHs. CalA, by contrast, hydrolyzed only 36 esters and has a low surface-exposed triad (SASA), with restrictive access to the catalytic triad (1000 Å^3).

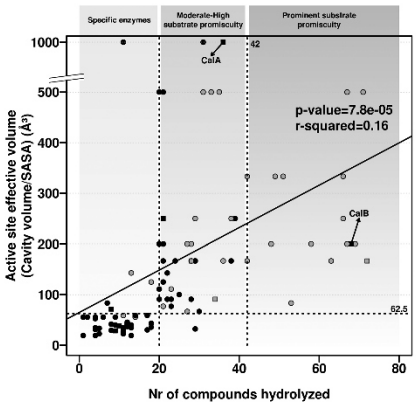


Figure 5. Defining of the substrate range of the EH by topology of the catalytic environment. The figure shows the relationships between the active site effective volume (in Å^3) and enzyme promiscuity (number of substrates hydrolyzed). Note that the presented data were obtained using the active site cavity volume computed in Å^3 and SASA as a dimensionless ratio from 0 to 100 using the GetArea server (http://curie.utmb.edu/getarea.html). The panel contains information for EHs for which crystal structures (square) and homology models (circles) could be unambiguously established (sequence identity ≥25%) and the catalytic triad identified. Gray circles and squares indicate the EHs assigned to FIV. The analysis indicated a threshold ratio (indicated by a horizontal dashed gray line) at which it is possible to qualitatively classify substrate promiscuity based on hydrolysis of at least 20 substrates. Phylogenetic analysis further extended the substrate spectra to ≥42 esters, as only enzymes assigned to FIV and conforming to the 62.5 Å^3 threshold, together with CalB, were capable of converting such a high number of esters. The positioning for the commercial CalA and CalB lipases is indicated.

than 62.5 Å^3 corresponded to EHs with activity for 20 or more of the 96 substrates tested and the opposite. There were only six outliers out of 95 EHs that did not follow this rule. Thus, the performance is of excellent (with 94%) accuracy if used as a classifier. The effective volume, however, does not have quantitative predictions for the exact number of esters hydrolyzed (r^2 = 0.16 for data in Figure 5), most likely because above the 62.5 Å^3 threshold, the capability to hydrolyze more or less substrates may specifically depend on the topology of the catalytic environment (Figure 4a−c), which may differ within families. Particularly, none of the different family members that conformed to the ≥62.5 Å^3 threshold, except those from FIV (*i.e.*, at least 50% of its members as shown in Figure 5, gray circle symbols) and CalB, could hydrolyze 42 or more esters. Therefore, the classification potential of the effective volume measure increased when combined with phylogenetic data. Noticeably, we observed that the predictive capacity of cavity volume/SASA is not influenced by the presence of flexible elements in the structure (Supporting Information Results).

The Active Site Effective Volume Is Also Indicative of Molecules Being Accepted As Substrates. We further used the active site cavity volume/SASA to also dissect its role in substrate specificity. We restricted the analysis to the 96 EHs for which this value could be unambiguously calculated (see above). The analysis indicated that the conversion of 34 esters was only observed for EHs conforming to the ≥62.5 Å^3 threshold (Supporting Information Figure S3). All but two (vinyl crotonate and ethyl acetate) could be considered large alkyl or hindered aromatic esters and included important molecules in synthetic organic chemistry such as paraben esters. This suggests that active sites with larger volume and a lower SASA (i.e., cavity less exposed to the surface) will most likely support hydrolysis of these esters. Therefore, the effective volume measure could be used to some extent as an indicator of substrates that may or may not be hydrolyzed by EHs. However, not all EHs fitting the ≥62.5 Å^3 threshold could convert all 34 of these esters, implying that this measure does not allow deepening into substrate specificity, which may depend on the topology of the catalytic environments as mentioned previously (Figure 4a−c). However, we found that the probability that benzyl-, butyl-, and propyl-paraben esters, major intermediates in chemical synthesis, are converted by members of the FIV with an effective volume ≥ 62.5 Å^3 is significantly higher (~35%) than that of EHs from FIV with a volume < 62.5 Å^3 and EHs from other families, whatever the value of the effective volume (approaching zero percent in our study); for those EHs for which effective volume could not be calculated, this probability is as low as 1.9% (Supporting Information Figure S4). This again exemplifies that the effective volume measure, when combined with phylogenetic information, is not only indicative of a promiscuity level but also can be used to predict the capacity to hydrolyze esters such as paraben esters. Screen programs to find EHs capable of converting

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

paraben esters should most likely be directed to find those assigned to FIV and with cavity volume/SASA ≥ 62.5 Å^3.

The Effective Volume Is Also a Marker of Substrate Promiscuity in Proteins Other than EHs. In order to evaluate the possibility that the active site effective volume may be a marker of substrate promiscuity in other enzymes, substrate spectra-effective volume relationships should be investigated in other protein families. In this line, Huang *et al.*[1] recently performed a systematic analysis of the substrate spectra of 200 phosphatases of the HAD superfamily, when tested against a set of 167 substrates. We collected the available crystal structures of each of the HAD phosphatases (Supporting Information Table S5) and computed the active site effective volume. We restricted the analysis to C2 cap members as they were reported to have a broader substrate spectrum,[1] and crystal structures with low to high effective volume are available. Interestingly, we observed that the effective volume (using the two conserved aspartic catalytic residues as the corrective SASA factor) was highly correlated (r^2 = 0.92) with the substrate range (Figure 6). Thus, the effective volume can be used as a molecular classification parameter of substrate promiscuity of phosphatases of the HAD superfamily when crystal structures are available. When this analysis was extended to the rest of the enzymes using homology modeling, we observed a similar trend to that of EHs (Supporting Information Figure S5). That is, no correlation existed (r^2 = 0.043), but still the effective volume can be used as a classifier of the substrate range as for EHs. Indeed, although a threshold could not be unambiguously established, sequences with the top 10 effective volumes belong to moderate-high to high promiscuity enzymes.

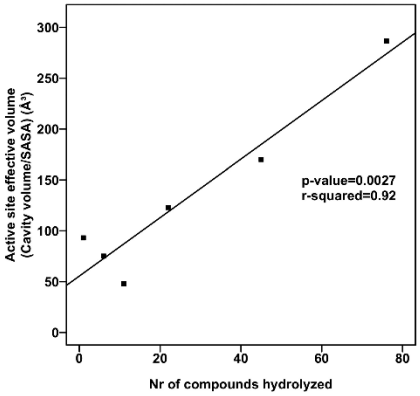


Figure 6. Relationships between the active site effective volume (in Å^3) and enzyme promiscuity (number of substrates hydrolyzed) of C2 members of HAD phosphatases. The number of substrates converted by each HAD phosphatase was obtained from Huang *et al.*[1] and is summarized in Supporting Information Table S5. The panels contain information for HAD phosphatases for which crystal structures were available and the catalytic residues identified. The active site effective volume (in Å^3) was calculated as described in Figure 5.

In conclusion, we found that the topology around the catalytic position, by means of an active site effective volume (cavity volume/SASA) threshold, is a dominant criterion of substrate promiscuity in EHs, which can be further extended by adding phylogenetic analysis. The rationale behind this parameter is as follows. Large volumes increase promiscuity until a certain value at which the cavity becomes too exposed and is not capable of properly accommodating and, importantly, retaining the substrate in specific catalytic binding interactions. This point is well captured by the SASA percentage of the catalytic triad, a dimensionless ratio that corrects for large volume measures in exposed sites. Importantly, the parameters of active site volume and relative SASA can be easily transferred to other systems. Indeed, the fact that the EHs investigated herein have different folds and that this parameter was also a marker of substrate spectra for phosphatases of the HAD superfamily opens the possibility of applying the effective volume measure to other enzymes requiring substrate anchoring. In all cases, the effective volume threshold-to-substrate relationships must be established. We would like to make note that the active site volume is not a static property, as the active site will breathe, depending on how flexible the protein is. In addition to that, the 62.5 Å^3 threshold for qualitatively classifying substrate promiscuity is based on the analysis of 147 EHs when tested against 96 esters. Although increasing the number of EHs and esters may influence this threshold and increase accuracy, it will not affect the fact that the measurement of the effective volume (cavity volume/SASA) can be used as the first molecular classification method of substrate promiscuity in EHs.

Our measurement is not a quantitative one, but rather a qualitative ranking (classification) procedure that will allow, for example, selecting sequences in databases for expression, particularly those encoding promiscuous enzymes capable of converting multiple substrates. This will substantially reduce reagent and labor costs compared to methods requiring the extensive cloning of all genes, and the expression and characterization of all enzymes in databases to later find those being promiscuous.[33] This possibility was herein examined by successfully mapping the open reading frames from the TARA Oceans project assemblies[31] and by identifying a high number of sequences encoding EHs with presumptive prominent substrate promiscuity (Supporting Information Results, Figures S6 and S7). Application of the effective volume measure to examine the sequences daily generated or deposited in databases requires having some crystals or X-ray structures for the model production. This limitation prevents predicting promiscuity from sequences lacking any structural information. Indeed, 36% of the EHs in this study (52 of the 147, including CalA/B) could not be included in the correlation because no calculation was possible. Accumulation of structural information and design and application of better modeling algorithms in the future will help solving this limitation.[35] Future studies might also explore molecular dynamics (MD) simulations to measure also the flexibility of the active site and not just the size of the cavity. By using this strategy, it was recently reported that the broad promiscuity of the members of the alkaline phosphatase superfamily arises from cooperative electrostatic interactions in the active site, allowing each enzyme to adapt to the electrostatic needs of different substrates.[36] In the particular case of EH phylogeny, a marker which does not require a three-dimensional structure was also suggested as a predictive classification marker of the substrate range. Indeed, this study suggests that in case of an unknown EH for which a crystal structure is not available or a homology model could not be established, then its assignation to family IV[14,23] increases the likelihood that this EH is promiscuous.

The present study not only provides clear evidence that substrate promiscuity in EHs has evolved from different core structural domains fitting an effective volume around the active

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

site, albeit with a bias toward that occurring in FIV members, but also from different phylogenetic lineages, many of which remain underexplored to date (Supporting Information Results and Figure S2). These are new findings as it was previously thought that the substrate range in a superfamily increased from a single ancestral core domain,[1] and because the identities of some microbial groups containing promiscuous enzymes, herein EHs, were previously unknown. Finally, this study also enabled the selection of a set of EH candidates that can compete with best commercial EHs such as CalB, as they show a broader substrate profile and specific activities up to 3-fold higher (Supporting Information Table S6). Their sequences can be used to search databases for similar promiscuous EHs. Further investigations should also determine the occurrence of other types of promiscuous EH phenotypes with broader substrate ranges than those identified in this study. For example, at least the stability of substrate-promiscuous EHs at different temperatures and with various solvents, along with the occurrence and evolution of secondary reactions, should be investigated in terms of condition and catalytic promiscuity.

METHODS

Protein Samples. Two main sources of EHs were used in the present study, all of them isolated *via* naïve and sequence-based screens in genomes and metagenomes. A first set of samples was EHs previously reported, as in the bibliography (69 in total), and that were herein substrate-profiled for first time. A second set was EHs (77) that are herein reported for first time. The extensive details of the source, cloning, expression, and purification of each of the active and soluble EHs are provided in the Supporting Information Methods and Table S1.

Ester Bond Hydrolysis Activity Assessment: Substrate Profiling Tests with 96 Esters. Hydrolytic activity was assayed at 550 nm using 96 structurally diverse esters in 384-well plates as previously described.[24,27,28] Before the assay, a concentrated stock solution of the esters was prepared at a concentration of 100 mg mL^{-1} in acetonitrile and dimethyl sulfoxide (DMSO). The assays were conducted according to the following steps. First, a 384-well plate (Molecular Devices, LLC, CA, USA) was filled with 20 μL of 5 mM *N*-(2-hydroxyethyl)piperazine-*N*′-(3-propanesulfonic acid (EPPS) buffer, at pH 8.0, using a QFill3 microplate filler (Molecular Devices, LLC, CA, USA). Second, 2 μL of each ester stock solution was added to each well using a PRIMADIAG liquid-handling robot (EYOWN TECHNOLOGIES SL, Madrid, Spain). The ester was dispensed in replicates. After adding the esters, the 384-well plate was filled with 20 μL of 5 mM EPPS buffer, at pH 8.0, containing 0.912 mM Phenol Red (used as a pH indicator) using a QFill3 microplate filler. The final ester concentration of the ester in each well was 1.14 mg mL^{-1}, and the final concentration of Phenol Red was 0.45 mM. A total of 2 μL of protein extract (containing 1−5 mg mL^{-1} pure protein or 200 mg mL^{-1} wet cells expressing proteins) was immediately added to each well using an Eppendorf Repeater M4 pipet (Eppendorf, Hamburg, Germany) or a PRIMADIAG liquid-handling robot. Accordingly, the total reaction volume was 44 μL, with 4.5% (v/v) acetonitrile or DMSO in the reaction mixture. After incubation at 30 °C in a Synergy HT Multi-Mode Microplate Reader, ester hydrolysis was measured spectrophotometrically in continuous mode at 550 nm for a total time of 24 h. Commercially available CALA L and CALB L (Novozymes A/S, Bagsvaerd, Denmark) were diluted 10-fold with 5 mM EPPS buffer, at pH 8.0, and 2 μL of this solution was used immediately for reaction tests under the conditions described before. In all cases, specific activities (in U g^{-1} protein) were determined. One unit (U) of enzyme activity was defined as the amount of wet cells expressing EHs or pure EHs required to transform 1 μmol of substrate in 1 min under the assay conditions using the reported extinction coefficient ($\varepsilon_{Phenol\text{-}red}$ at 550 nm = 8450 M^{-1} cm^{-1}). All values were corrected for nonenzymatic transformation (i.e., the background rate) and for the background signal using *E. coli* cells that did not express any target protein (control cells included empty vectors). Note that a positive reaction was indicated by the restrictive criterion of a change greater than 6-fold above the background signal. Specific activity determinations (in U g^{-1}) for wet cells expressing each of the selected EHs or pure or commercial proteins are available in Supporting Information Tables S3 and S6, respectively.

Structural Determinations and Homology Modeling. The proteins EH1 and EH102 were expressed, purified, and crystallized using the sitting-drop method in Intelliplate 96-well plates and a Mosquito liquid-handling robot (TTP LabTech) according to previously described procedures.[37] For EHs for which crystal structures were not available, homology models were developed using Prime software from Schrödinger. Prime uses BLAST (with BLOSUM62 matrix) for homology search and alignment and refines the results using the Pfam database and pairwise alignment with ClustalW.

Protein Energy Landscape Exploration (PELE) Simulations. We used Protein Energy Landscape Exploration (PELE) software to sample the binding modes of glyceryl triacetate with EH1 and EH102.[38,39] The initial structures were taken from the coordinates of the EH1 and EH102 crystal structures (PDB codes: 5JD4 and 5JD3, respectively). The protonation state of titratable residues was estimated with the Protein Preparation Wizard (PROPKA)[40] and the H++ server (http://biophysics.cs.vt.edu/H++) followed by visible inspection. At pH 8 (the pH at which the activity assays were performed), the catalytic triad histidine residues were δ-protonated, and the catalytic triad aspartic acid residues were deprotonated, resulting in the formation of a histidine-serine and histidine-aspartic hydrogen-bonding network. The glyceryl triacetate structure was fully optimized with Jaguar[41] in an implicit solvent, and the electrostatic potential charges were computed with the density functional M06 at the 6-31G* level of theory. The ligand parameters were extracted from these for the classic simulations.

Cavity Volume and Solvent Accessible Surface Area (SASA) calculation. The relative Solvent Accessible Surface Area (SASA) for a residue was obtained using the GetArea Web server.[42] Cavity volumes were computed with Fpocket,[43] a very fast open-source protein pocket (cavity) detection algorithm based on Voronoi tessellation. Fpocket includes two other programs (dpocket and tpocket) that allow the extraction of pocket descriptors and the testing of owned scoring functions, respectively.

For extensive details of the methods, see Supporting Information Methods.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acschembio.7b00996.

Supporting Results, Methods, Figures S1−S7, and Table S4 (PDF)

Tables S1−S3, S5, and S6 (XLS)

AUTHOR INFORMATION

Corresponding Authors

*E-mail: victor.guallar@bsc.es.

*E-mail: mferrer@icp.csic.es.

ORCID

Gerard Santiago: 0000-0002-0506-3049

Jürgen Pleiss: 0000-0003-1045-8202

Alexander F. Yakunin: 0000-0003-0813-6490

Víctor Guallar: 0000-0002-4580-1114

Manuel Ferrer: 0000-0003-4962-4714

Present Addresses

[γ]Current address: School of Chemistry, Bangor University, LLS7 2UW Bangor, UK.

[δ]Current address: Lehrstuhl für Biotechnologie, RWTH Aachen University, Aachen, Germany.

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

[ζ]Current address: Carl R. Woese Institute for Genomic Biology, Urbana, USA.

Author Contributions

[ζ]These authors contributed equally to this work.

Author Contributions

[ψ]These authors contributed equally in coordinating activities.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

C.C. thanks the Spanish Ministry of Economy, Industry and Competitiveness for a Ph.D. fellowship (Grant BES-2015-073829). V.M. thanks the Francisco José de Caldas Scholarship Program (Administrative Department of Science, Technology and Innovation, COLCIENCIAS). The authors acknowledge the members of the MAMBA, MAGICPAH, ULIXES, KILLSPILL and INMARE Consortia for their support in sample collection. David Rojo is also acknowledged for his valuable help with log P calculations. This project received funding from the European Union's Horizon 2020 research and innovation program [Blue Growth: Unlocking the potential of Seas and Oceans] under grant agreement no. 634486 (project acronym INMARE). This research was also supported by the European Community Projects MAGICPAH (FP7-KBBE-2009-245226), ULIXES (FP7-KBBE-2010-266473), and KILLSPILL (FP7-KBBE-2012-312139) and grants BIO2011-25012, PCIN-2014-107, BIO2014-54494-R, and CTQ2016-79138-R from the Spanish Ministry of Economy, Industry and Competitiveness. The present investigation was also funded by the Spanish Ministry of Economy, Industry and Competitiveness within the ERA NET IB2, grant no. ERA-IB-14-030 (MetaCat), the UK Biotechnology and Biological Sciences Research Council (BBSRC), grant no. BB/M029085/1, and the German Research Foundation (FOR1296). R.B. and P.N.G. acknowledge the support of the Supercomputing Wales project, which is part-funded by the European Regional Development Fund (ERDF) *via* the Welsh Government. O.V.G. and P.N.G. acknowledge the support of the Centre of Environmental Biotechnology Project funded by the European Regional Development Fund (ERDF) through the Welsh Government. A.Y. and A.S. gratefully acknowledge funding from Genome Canada (2009-OGI-ABC-1405) and the NSERC Strategic Network grant IBN. A.I.P. was supported by the Counseling of Economy and Employment of the Principality of Asturias, Spain (Grant FC-15-GRUPIN14-107). V.G. acknowledges the joint BSC-CRG-IRB Research Program in Computational Biology. The authors gratefully acknowledge financial support provided by the European Regional Development Fund (ERDF).

■ REFERENCES

(1) Huang, H., Pandya, C., Liu, C., Al-Obaidi, N. F., Wang, M., Zheng, L., Toews Keating, S., Aono, M., Love, J. D., Evans, B., Seidel, R. D., Hillerich, B. S., Garforth, S. J., Almo, S. C., Mariano, P. S., Dunaway-Mariano, D., Allen, K. N., and Farelli, J. D. (2015) Panoramic view of a superfamily of phosphatases through substrate profiling. *Proc. Natl. Acad. Sci. U. S. A. 112*, E1974−1983.

(2) Huang, R., Hippauf, F., Rohrbeck, D., Haustein, M., Wenke, K., Feike, J., Sorrelle, N., Piechulla, B., and Barkman, T. J. (2012) Enzyme functional evolution through improved catalysis of ancestrally non-preferred substrates. *Proc. Natl. Acad. Sci. U. S. A. 109*, 2966−2971.

(3) Yip, S. H., and Matsumura, I. (2013) Substrate ambiguous enzymes within the *Escherichia coli* proteome offer different evolutionary solutions to the same problem. *Mol. Biol. Evol. 30*, 2001−2012.

(4) Price, D. R., and Wilson, A. C. (2014) Substrate ambiguous enzyme facilitates genome reduction in an intracellular symbiont. *BMC Biol. 12*, 110.

(5) Devamani, T., Rauwerdink, A. M., Lunzer, M., Jones, B. J., Mooney, J. L., Tan, M. A., Zhang, Z. J., Xu, J. H., Dean, A. M., and Kazlauskas, R. J. (2016) Catalytic promiscuity of ancestral esterases and hydroxynitrile lyases. *J. Am. Chem. Soc. 138*, 1046−1056.

(6) Hult, K., and Berglund, P. (2007) Enzyme promiscuity: mechanism and applications. *Trends Biotechnol. 25*, 231−238.

(7) Copley, S. D. (2015) An evolutionary biochemist's perspective on promiscuity. *Trends Biochem. Sci. 40*, 72−78.

(8) London, N., Farelli, J. D., Brown, S. D., Liu, C., Huang, H., Korczynska, M., Al-Obaidi, N. F., Babbitt, P. C., Almo, S. C., Allen, K. N., and Shoichet, B. K. (2015) Covalent docking predicts substrates for haloalkanoate dehalogenase superfamily phosphatases. *Biochemistry 54*, 528−537.

(9) Nobeli, I., Favia, A. D., and Thornton, J. M. (2009) Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol. 27*, 157−167.

(10) Ferrer, M., Martínez-Martínez, M., Bargiela, R., Streit, W. R., Golyshina, O. V., and Golyshin, P. N. (2016) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. *Microb. Biotechnol. 9*, 22−34.

(11) Braakman, R., and Smith, E. (2014) Metabolic evolution of a deep-branching hyperthermophilic chemoautotrophic bacterium. *PLoS One 9*, e87950.

(12) Giovannoni, S. J., Cameron Thrash, J., and Temperton, B. (2014) Implications of streamlining theory for microbial ecology. *ISME J. 8*, 1553−1565.

(13) Lan, T., Wang, X. R., and Zeng, Q. Y. (2013) Structural and functional evolution of positively selected sites in pine glutathione S-transferase enzyme family. *J. Biol. Chem. 288*, 24441−24451.

(14) Ferrer, M., Bargiela, R., Martínez-Martínez, M., Mir, J., Koch, R., Golyshina, O. V., and Golyshin, P. N. (2015) Biodiversity for biocatalysis: A review of the α/β-hydrolase fold superfamily of esterases-lipases discovered in metagenomes. *Biocatal. Biotransform. 33*, 235−249.

(15) Schmid, A., Dordick, J. S., Hauer, B., Kiener, A., Wubbolts, M., and Witholt, B. (2001) Industrial biocatalysis today and tomorrow. *Nature 409*, 258−268.

(16) Koudelakova, T., Chovancova, E., Brezovsky, J., Monincova, M., Fortova, A., Jarkovsky, J., and Damborsky, J. (2011) Substrate specificity of haloalkane dehalogenases. *Biochem. J. 435*, 345−354.

(17) Risso, V. A., Gavira, J. A., Mejia-Carmona, D. F., Gaucher, E. A., and Sanchez-Ruiz, J. M. (2013) Hyperstability and substrate promiscuity in laboratory resurrections of Precambrian β-lactamases. *J. Am. Chem. Soc. 135*, 2899−2902.

(18) Amin, S. R., Erdin, S., Ward, R. M., Lua, R. C., and Lichtarge, O. (2013) Prediction and experimental validation of enzyme substrate specificity in protein structures. *Proc. Natl. Acad. Sci. U. S. A. 110*, E4195−4202.

(19) Anton, B. P., Chang, Y. C., Brown, P., Choi, H. P., Faller, L. L., Guleria, J., Hu, Z., Klitgord, N., Levy-Moonshine, A., Maksad, A., Mazumdar, V., McGettrick, M., Osmani, L., Pokrzywa, R., Rachlin, J., Swaminathan, R., Allen, B., Housman, G., Monahan, C., Rochussen, K., Tao, K., Bhagwat, A. S., Brenner, S. E., Columbus, L., de Crécy-Lagard, V., Ferguson, D., Fomenkov, A., Gadda, G., Morgan, R. D., Osterman, A. L., Rodionov, D. A., Rodionova, I. A., Rudd, K. E., Söll, D., Spain, J., Xu, S. Y., Bateman, A., Blumenthal, R. M., Bollinger, J. M., Chang, W. S., Ferrer, M., Friedberg, I., Galperin, M. Y., et al. (2013) The COMBREX project: design, methodology, and initial results. *PLoS Biol. 11*, e1001638.

(20) Colin, P. Y., Kintses, B., Gielen, F., Miton, C. M., Fischer, G., Mohamed, M. F., Hyvönen, M., Morgavi, D. P., Janssen, D. B., and Hollfelder, F. (2015) Ultrahigh-throughput discovery of promiscuous enzymes by picodroplet functional metagenomics. *Nat. Commun. 6*, 10008.

(21) Chen, C., Huang, H., and Wu, C. H. (2017) Protein bioinformatics databases and resources. *Methods Mol. Biol. 1558*, 3−39.

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

(22) Fischer, M., and Pleiss, J. (2003) The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res. 31*, 319−321.
(23) Arpigny, J. L., and Jaeger, K. E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem. J. 343*, 177−183.
(24) Alcaide, M., Tornés, J., Stogios, P. J., Xu, X., Gertler, C., Di Leo, R., Bargiela, R., Lafraya, A., Guazzaroni, M. E., López-Cortés, N., Chernikova, T. N., Golyshina, O. V., Nechitaylo, T. Y., Plumeier, I., Pieper, D. H., Yakimov, M. M., Savchenko, A., Golyshin, P. N., and Ferrer, M. (2013) Single residues dictate the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family. *Biochem. J. 454*, 157−166.
(25) Lombard, V., Bernard, T., Rancurel, C., Brumer, H., Coutinho, P. M., and Henrissat, B. (2010) A hierarchical classification of polysaccharide lyases for glycogenomics. *Biochem. J. 432*, 437−444.
(26) Popovic, A., Hai, T., Tchigvintsev, A., Hajighasemi, M., Nocek, B., Khusnutdinova, A. N., Brown, G., Glinos, J., Flick, R., Skarina, T., Chernikova, T. N., Yim, V., Brüls, T., Paslier, D. L., Yakimov, M. M., Joachimiak, A., Ferrer, M., Golyshina, O. V., Savchenko, A., Golyshin, P. N., and Yakunin, A. F. (2017) Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. *Sci. Rep. 7*, 44103.
(27) Janes, L. E., Löwendahl, C., and Kazlauskas, R. J. (1998) Rapid quantitative screening of hydrolases using pH indicators. Finding enantioselective hydrolases. *Chem. Eur. J. 4*, 2317−2324.
(28) Martínez-Martínez, M., Alcaide, M., Tchigvintsev, A., Reva, O., Polaina, J., Bargiela, R., Guazzaroni, M. E., Chicote, A., Canet, A., Valero, F., Rico Eguizabal, E., Guerrero, Mdel C., Yakunin, A. F., and Ferrer, M. (2013) Biochemical diversity of carboxyl esterases and lipases from Lake Arreo (Spain): a metagenomic approach. *Appl. Environ. Microbiol. 79*, 3553−3562.
(29) Daiha, K. d. G., Angeli, R., de Oliveira, S. D., and Almeida, R. V. (2015) Are lipases still important biocatalysts? A study of scientific publications and patents for technological forecasting. *PLoS One 10*, e0131624.
(30) Borrelli, K. W., Cossins, B., and Guallar, V. (2009) Exploring hierarchical refinement techniques for induced fit docking with protein and ligand flexibility. *J. Comput. Chem. 31*, 1224−1235.
(31) Hernández-Ortega, A., Borrelli, K., Ferreira, P., Medina, M., Martínez, A. T., and Guallar, V. (2011) Substrate diffusion and oxidation in GMC oxidoreductases: an experimental and computational study on fungal aryl-alcohol oxidase. *Biochem. J. 436*, 341−350.
(32) Santiago, G., de Salas, F., Lucas, F., Monza, E., Acebes, S., Martínez, A., Camarero, S., and Guallar, V. (2016) Computer-aided laccase engineering: toward biological oxidation of arylamines. *ACS Catal. 6*, 5415−5423.
(33) Barak, Y., Nov, Y., Ackerley, D. F., and Matin, A. (2008) Enzyme improvement in the absence of structural knowledge: a novel statistical approach. *ISME J. 2*, 171−179.
(34) Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., d'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B. T., Royo-Llonch, M., Sarmento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S., Bowler, C., de Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M. B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S. G., Bork, P., et al. (2015) Structure and function of the global ocean microbiome. *Science 348*, 1261359.
(35) Moult, J., Fidelis, K., Kryshtafovych, A., Schwede, T., and Tramontano, A. (2016) Critical assessment of methods of protein structure prediction: Progress and new directions in round XI. *Proteins: Struct., Funct., Genet. 84* (Suppl 1), 4−14.
(36) Barrozo, A., Duarte, F., Bauer, P., Carvalho, A. T. P., and Kamerlin, S. C. L. (2015) Cooperative electrostatic interactions drive functional evolution in the alkaline phosphatase superfamily. *J. Am. Chem. Soc. 137*, 9061−9076.
(37) Alcaide, M., Stogios, P. J., Lafraya, Á., Tchigvintsev, A., Flick, R., Bargiela, R., Chernikova, T. N., Reva, O. N., Hai, T., Leggewie, C. C., Katzke, N., La Cono, V., Matesanz, R., Jebbar, M., Jaeger, K. E., Yakimov, M. M., Yakunin, A. F., Golyshin, P. N., Golyshina, O. V., Savchenko, A., and Ferrer, M. (2015) Pressure adaptation is linked to thermal adaptation in salt-saturated marine habitats. *Environ. Microbiol. 17*, 332−345.
(38) Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B 105*, 6474−6487.
(39) Borrelli, K. W., Vitalis, A., Alcantara, R., and Guallar, V. (2005) PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. *J. Chem. Theory Comput. 1*, 1304−1311.
(40) Madhavi Sastry, G., Adzhigirey, M., Day, T., Annabhimoju, R., and Sherman, W. (2013) Protein and ligand preparation: parameters, protocols, and influence on virtual screening enrichments. *J. Comput.-Aided Mol. Des. 27*, 221−234.
(41) Bochevarov, A. D., Harder, E., Hughes, T. F., Greenwood, J. R., Braden, D. A., Philipp, D. M., Rinaldo, D., Halls, M. D., Zhang, J., and Friesner, R. A. (2013) Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum Chem. 113*, 2110−2142.
(42) Fraczkiewicz, R., and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem. 19*, 319.
(43) Le Guilloux, V., Schmidtke, P., and Tuffery, P. (2009) Fpocket: An open source platform for ligand pocket detection. *BMC Bioinf. 10*, 168.

DOI: 10.1021/acschembio.7b00996
ACS Chem. Biol. 2018, 13, 225−234

SUPPORTING INFORMATION

Determinants and prediction of esterase substrate promiscuity patterns

Mónica Martínez-Martínez, Cristina Coscolín, Gerard Santiago, Jennifer Chow, Peter J. Stogios, Rafael Bargiela, Christoph Gertler, José Navarro-Fernández, Alexander Bollinger, Stephan Thies, Celia Méndez-García, Ana Popovic, Greg Brown, Tatyana N. Chernikova, Antonio García-Moyano, Gro E.K. Bjerga, Pablo Pérez-García, Tran Hai, Mercedes V. Del Pozo, Runar Stokke, Ida H. Steen, Hong Cui, Xiaohui Xu, Boguslaw P. Nocek, María Alcaide, Marco Distaso, Victoria Mesa, Ana I. Peláez, Jesús Sánchez, Patrick C. F. Buchholz, Jürgen Pleiss, Antonio Fernández-Guerra, Frank O. Glöckner, Olga V. Golyshina, Michail M. Yakimov, Alexei Savchenko, Karl-Erich Jaeger, Alexander F. Yakunin, Wolfgang R. Streit, Peter N. Golyshin, Víctor Guallar, Manuel Ferrer. The INMARE Consortium

Table of Contents……S1

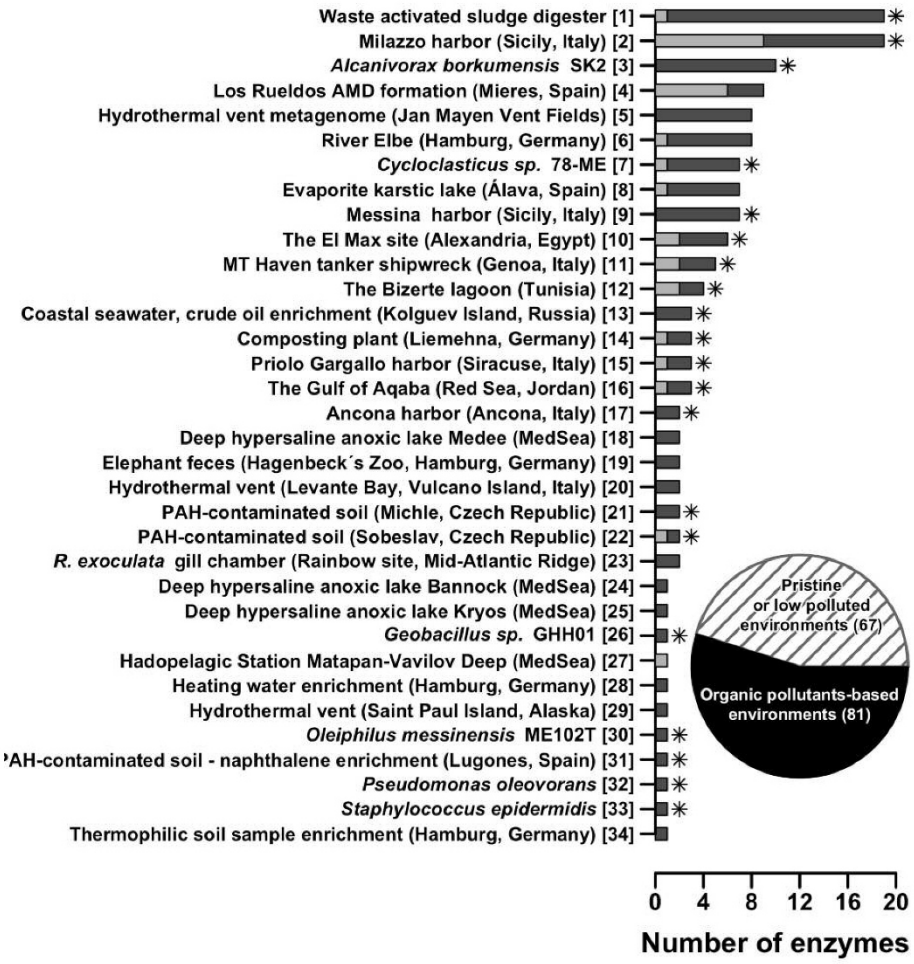


Figure S1. Bio-resources (habitats or microbes) of all 145 EHs in this study. This figure is generated from data in Supporting Information, Tables S1. The bars summarized the number of EHs characterized per each of the bio-resources screened. The number of highly promiscuous hits (i.e., using 30 or more esters) is shown by gray color in the bars. Two major bio-sources of enzymes were covered, namely organically polluted environments or bacteria inhabiting them, and pristine or low-pollution environments. The number of each of each of these bio-sources is summarized in the circular inset. The bio-sources defined as polluted are specifically marked with an asterisk in the bars. Each of the bio-resources is assigned an arbitrary number in square brackets, which is further used as ID in Fig. 2.

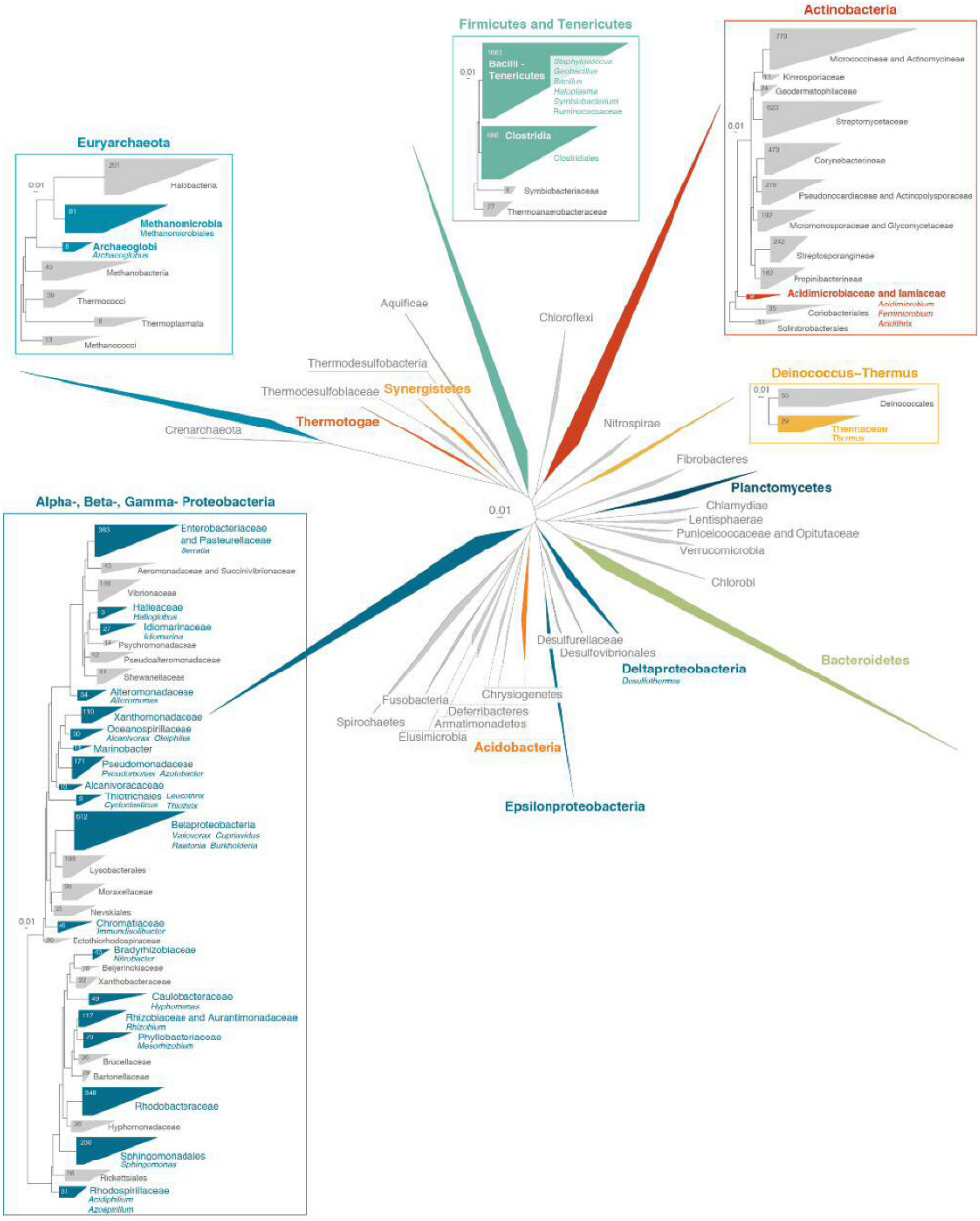


Figure S2. The associations of EHs to cultured and uncultured bacteria from multiple phylogenetic lineages. This figure is generated from data in Supporting Information, Tables S1. Maximum likelihood phylogenetic analysis of the 16S rRNA genes of all species present in the most recent

release of the Living Tree Project database. The groups in which enzymes have been characterized are presented with a color code. Microbial clades not covered in this study are colored in gray. The subtrees represent in detail the groups that contain EHs with broad substrate spectra (i.e., using 30 esters or more) described at the family and genus levels. The numbers in each collapsed clade from the subtrees designate the number of 16S rRNA gene sequences included in each group. Only major taxonomic groups with more than 5 type species present in the database were included in the phylogeny. The tree scales represent the nucleotide substitutions per site. Note that the phylogenetic binning of sequences encoding EHs (for details see Supporting Information, Table S1) was performed using a genome linguistics approach. Briefly, metagenomic fragments were searched for oligonucleotide compositional similarity (frequencies of tetranucleotides) against all sequenced bacterial chromosomes, plasmids and phages using the GOHTAM web tool.[1] For short DNA fragments, compositional analysis could not be performed, and a comprehensive analysis of the TBLASTX results was conducted. Both methods have been proven successful for suggesting the origin of metagenomics sequences.[1,2] However, we are aware that deep assignations (including at the species level) cannot be obtained for short DNA fragments and that this tool may not be appropriate for phylogenetic analysis of sequences with no homology in databases.

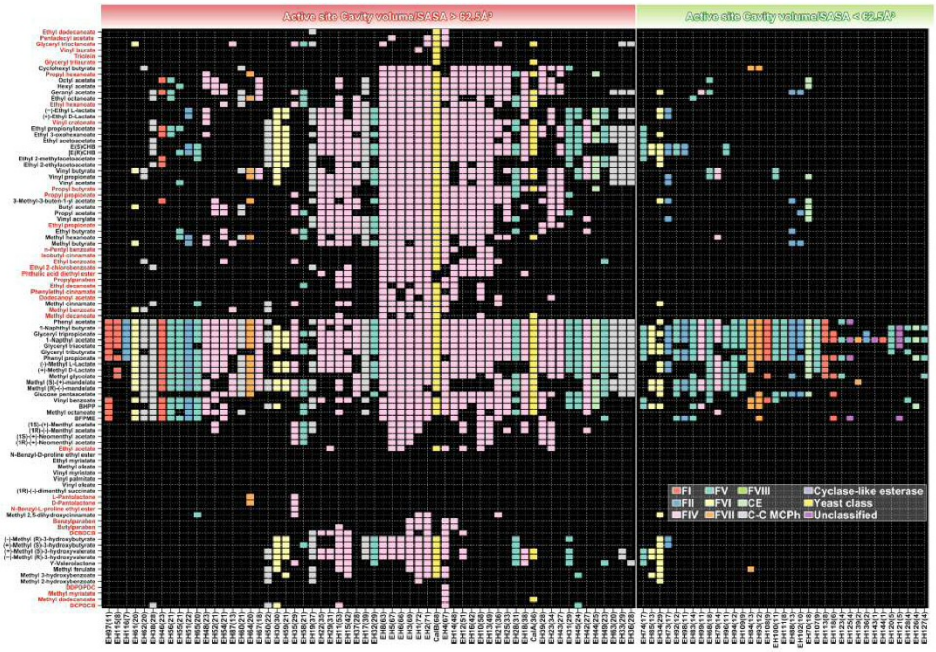


Figure S3. Active site effective volume and phylogeny may be also indicative of specific conversions. This figure is generated from data in Supporting Information, Tables S1 and S3. The figure is constructed as described in Fig. 2, although distance matrixes are not shown. Note that in Fig. 2 only EHs capable of converting 10 or more esters are given whatever their active site effective volume. In this case, the figure summarizes the substrate spectra of EHs from which active site effective volume could be unambiguously identified (96 in total) independently of the number of esters they can hydrolyze. The figure summarizes the substrate spectra of EHs with effective volumes ≥62.5 $Å^3$ (on the left side) or < 62.5 $Å^3$ (on the right side). The 34 esters found to be hydrolyzed under our assay conditions by EHs with effective volumes ≥62.5 $Å^3$ are indicated in the left side with a red color. Note that in all cases the threshold of 62.5 $Å^3$ was defined when the active site cavity volume was computed in $Å^3$ and SASA as a dimensionless ratio from 0 to 100 using the GetArea server (http://curie.utmb.edu/getarea.html).

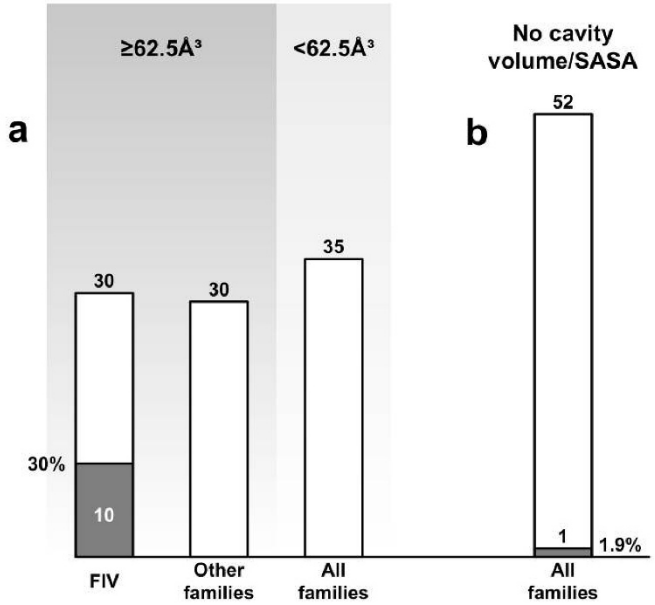


Figure S4. The active site effective volume and phylogeny are predictive markers of the capacity to convert benzyl-, butyl- and propyl-paraben esters. This figure is generated from data in Supporting Information, Tables S3 and S6. (a) Numbers (shown on the top of the bars) of EHs, assigned to FIV or other families, conforming to the ≥62.5 $Å^3$-threshold, or to all families and <62.5 $Å^3$-threshold. (b) Numbers of EHs (shown on the top of the bars) for which effective volume could not be unambiguously measured because sequence identities of less than 25% (to an existing crystal) or because not suitable alignments. In all cases, the number and percentage of EHs capable of converting paraben esters is shown by gray color in the bars. The threshold of 62.5 $Å^3$ was defined when the active site cavity volume was computed in $Å^3$ and SASA as a dimensionless ratio from 0 to 100 using the GetArea server (http://curie.utmb.edu/getarea.html).

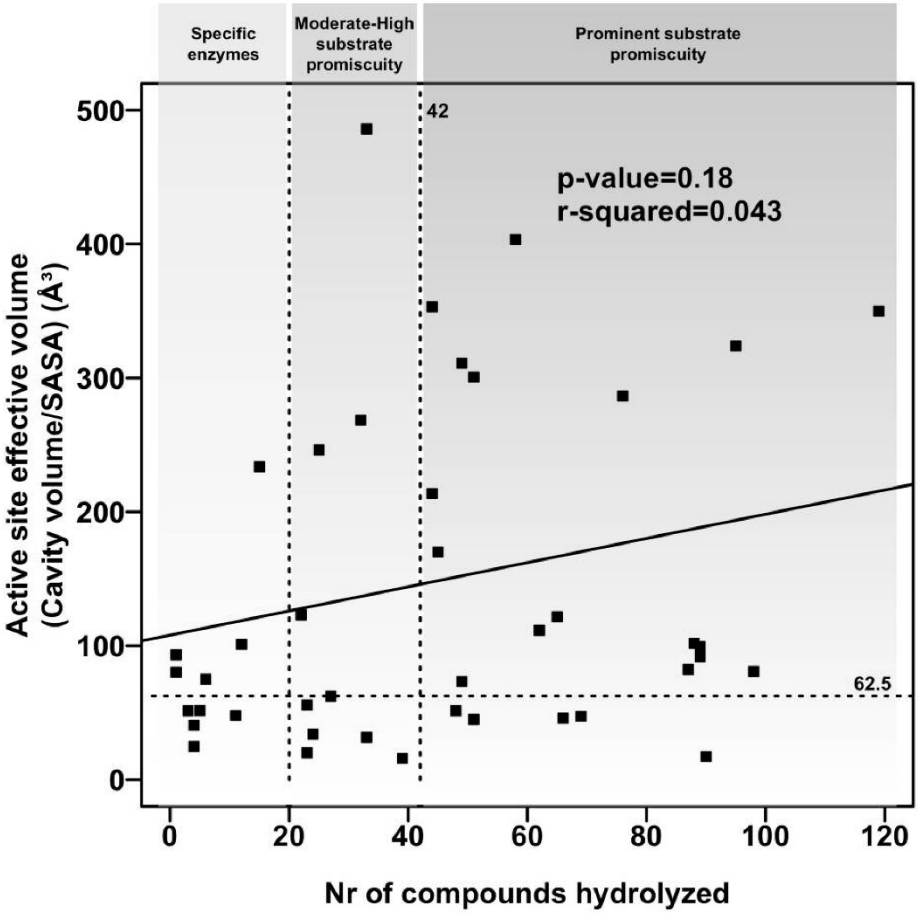


Figure S5. Relationships between the active site effective volume (in Å^3) and enzyme promiscuity (number of substrates hydrolyzed) of C2 members of HAD phosphatases. The number of substrates converted by each HAD phosphatase was obtained from Huang *et al.*[6] and is summarized in Supporting Information, Table S5. The panels contain information for HAD phosphatases for both crystal structures and homology models were available or could be unambiguously established (sequence identity ≥25%) and the catalytic residues identified. The threshold of substrates being converted, exemplifying the level of promiscuity as for EHs, is indicated. The active site effective volume (in Å^3) was calculated as described in Fig. 5.

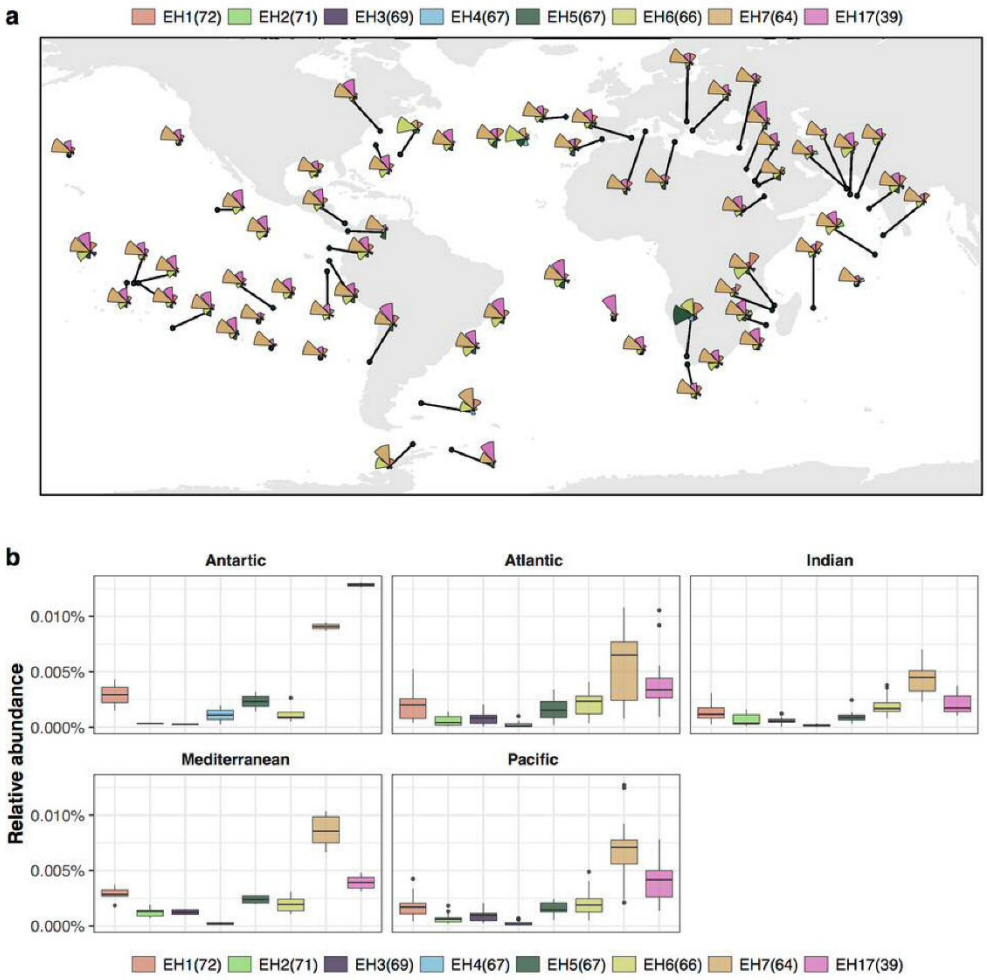


Figure S6. Exploring the sequence space for sequences that encode the most promiscuous EHs. The number in brackets close to the enzyme ID indicates the number of esters hydrolyzed by each enzyme, as described in the legend of Fig. 2. (a) Geographic distribution of the 8 selected promiscuous EHs in the 242 TARA Oceans samples. (b) Relative abundance of the 8 promiscuous EHs in five regions.

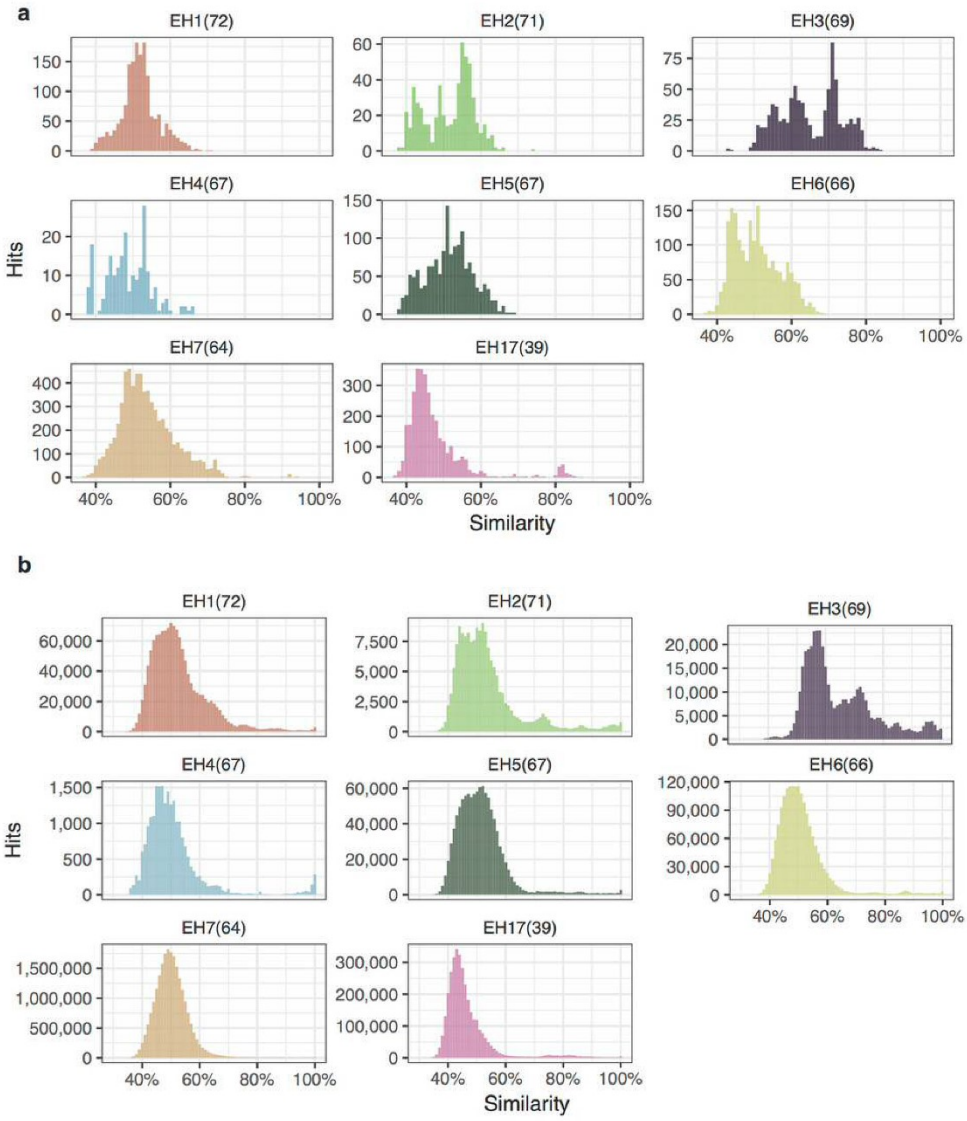


Figure S7. Distribution of the similarity values. The number in the bracket close to the enzyme ID indicates the number of esters hydrolyzed by each enzyme, as described in the legend of Fig. 2. (a) Blastp similarity distribution for every hit against each of the 8 selected promiscuous EHs. (b) All-vs-all blastp similarity distributions for all EH homologs found in the TARA Oceans samples.

Table S1. General information about all 145 EHs and the 2 commercial preparations investigated in this study. The database is an Excel table that provides the following information: ID code used in this study, accession number, PDB code if crystals were available, family to which each sequence was assigned, number of esters hydrolyzed (as found in Supporting Information, Table S3), active site cavity volume/SASA, average maximum specific activity (in U g^{-1} wet cells) (see Supporting Information, Table S3), sequence information, screening technique (sequence or naïve screen), enzyme source (with a full description and short description as in Figure 1), habitat type (as described in Figure 1 inset), cloning vector and expression host, expression conditions (i.e., antibiotic, inductor and temperature), a reference describing the identification, cloning and expression, the top hit in NCBI and the sequence identity, the theoretical molecular weight (ranging from 22 to 103 kDa) and the isoelectric point (ranging from 3.81 to 11.02). *Note 1*: To unambiguously assign sequences to families, a phylogenetic tree was generated using the FastTree v2.1.7 algorithm[3] implemented with the Shimodaira-Hasegawa test. Reference sequences that were unambiguously assigned to each of the esterase/lipase families[4,5] were used to help classify each of the 145 EHs. The final sequence alignments and the tree are available from the authors upon request. *Note 2*: The source organisms of selected polypeptides were identified by a search of oligonucleotide patterns against the GOHTAM database and TBLASTX to reveal compositional similarities between the gene sequences and/or DNA fragments containing the genes that encoded EHs and all analyzed sequenced bacterial chromosomes, plasmids and phages.[1,2] Due to the extensive size, the table is submitted as a separate file in Excel format.

Table S2. Pairwise sequence similarities for all 147 EHs (including CalA/B) as calculated using Needleman-Wunsch alignments performed against all the other candidates ("all-vs.-all"). Due to the extensive size, the table is submitted as a separate file in Excel format.

Table S3. The specific activity is given as U g^{-1} wet cells expressing esterases tested against a set of 96 structurally different esters. The table is an Excel table in which the following information is provided: name of ester, compound chemical class, log P value with standard deviation (calculated using the ACD/ChemSketch 2015.2.5 software), SMILES code, source or brand (with link to the source), molecular mass (g/mol), reaction conditions (including assay method, substrate and pH indicator concentration, buffer, shaking, temperature [30 °C] and pH [8.0], reaction time, additives and concentrations, cell amount), and average specific activity values for all 145 EH preparations in whole-cell assays given in U g^{-1} wet cell pellet. For the Cal A and Cal B preparations, the specific activity values are given in U g^{-1} total protein. The assays were performed as replicates, with the

average value given, and the standard deviation was less than 1% in all cases. Due to the extensive size, the table is submitted as a separate file in Excel format.

Table S4. X-ray diffraction data collection and refinement statistics.

Protein	EH1	EH102
PDB code	5JD4	5DJ3
Data collection		
Space group	P 1	$P2_1$
Cell dimensions		
a, *b*, *c* (Å)	90.28, 90.10, 110.76	66.59, 129.45,
α, β, γ (°)	68.02, 79.60, 67.57	129.85
		90, 97.49, 90
Resolution, Å	25.00 – 2.05	29.73 – 2.30
R_{merge}[a]	0.072 (0.606)[b]	0.125 (0.543)
$I / \sigma(I)$	21.66 (2.52)	9.86 (2.50)
Completeness, %	94.9 (91.2)	96.9 (93.4)
Redundancy	3.0 (2.9)	3.6 (3.6)
Refinement		
Resolution, Å	24.95 – 2.05	29.73 – 2.30
No. of unique reflections:	16392, 1875	86908, 4375
working, test		
R-factor/free *R*-factor[c]	16.9/21.1 (27.9/30.0)	18.4/23.6
		(22.4/28.1)
No. of refined atoms,		
molecules	19 262, 8	13 721, 8
Protein	324	23
Solvent	2463	1572
Water		
B-factors		
Protein	37.1	37.1
Solvent	63.1	60.5

Water	48.8	47.0
r.m.s.d.		
Bond lengths, Å	0.003	0.004
Bond angles, °	0.730	0.684

[a]$R_{sym} = \Sigma_h\Sigma_i|I_i(h) - \langle I(h)\rangle/\Sigma_h\Sigma_i I_i(h)$, in which $I_i(h)$ and $\langle I(h)\rangle$ are the *i*th and mean measurement of the intensity of reflection *h*.

[b]The values for the outer shells of the data are indicated in parentheses.

[c]$R = \Sigma|F_p^{obs} - F_p^{calc}|/\Sigma F_p^{obs}$, in which F_p^{obs} and F_p^{calc} are the observed and calculated structure factor amplitudes, respectively.

* = molecules in the active site cleft.

Table S5. General information about all HAD phosphatases analyzed in this study. The database is an Excel table that provides the following information: ID code, Uniprot ID and number of substrates being converted as reported,[6] and active site cavity volume/SASA. In case of protein for which a crystal structure was not available, the homology to the PDB template (in %) is given. Due to the extensive size, the table is submitted as a separate file in Excel format.

Table S6. The specific activity is given as U g^{-1} pure protein tested against a set of 96 structurally different esters. The data for the Cal A and Cal B commercial preparations are also included. The table is an Excel table in which the following information is provided: name of ester, compound chemical class, log P value with standard deviation (calculated using the ACD/ChemSketch 2015.2.5 software), SMILES code, source or brand (with link to the source), molecular mass (g/mol), reaction conditions (including assay method, substrate and pH indicator concentration, buffer, shaking, temperature [30°C] and pH [8.0], reaction time, additives and concentrations, enzyme amount), and average specific activity values for all tested EHs in pure protein assays given in U g^{-1} protein. The assays were performed as replicates, with the average value given, and the standard deviation was less than 0.5% in all cases. It should be noted that all preparations were tested in pure form as His-tagged proteins, but for the sake of simplicity, the specific activities of only the most promiscuous EHs (i.e., those that reacted with 42 or more esters) are shown. All datasets are available upon request from the authors. Due to the extensive size, the table is submitted as a separate file in Excel format.

RESULTS

Source of EH. A search of oligonucleotide patterns against the GOHTAM database[1,2] and TBLASTX analysis indicated that all 145 EHs were distributed across the entire phylogenetic tree (i.e., in at least 10 phyla and 40 genera), although a bias (70%) toward proteobacterial EH was noted (Supporting Information, Table S1, Fig. S2). No clear affiliation other than *Bacteria* was observed for 9 EHs, and an additional set of 43 EHs could not be assigned at the genus level. Within the 31 highly promiscuous EHs (i.e., using 30 or more esters; excluding CalA/B), an unambiguous affiliation was found for 21 EHs that were associated with at least 2 phyla (the most abundant was Proteobacteria (94%)) and 15 genera. These genera included *Sphingomonas*, *Rhizobium*, *Pseudomonas*, *Alteromonas* and *Acidiphilium*, which have been well-explored with respect to their enzymatic contents, as well as genera that have been largely neglected, such as *Acidithrix, Acidimicrobium*/*Ferrimicrobium*, *Alcanivorax*, *Cycloclasticus*, *Immundisolibacter*, *Idiomarina*, *Hyphomonas,* and *Halioglobus* (Supporting Information, Fig. S2). Of the known 24 bacterial and 5 archaeal phyla with cultured representatives,[7] the present study covered EHs from 10 and 1 phyla, respectively (Supporting Information, Fig. S2); much diversity remains to be uncovered in the coming years. The reason that enzyme diversity is so biased is intriguing, even though a very broad diversity was sampled in this study. Note that because many microbial lineages others than those herein analyzed may contain promiscuous EHs a statistical significance of the enrichment of substrate-promiscuous EHs in the above bacterial species cannot be assessed, which is out of the scope of the present study. Whatever the case, we have provided first preliminary evidences that a number of underexplored microbial phylogenetic lineages contained EHs with prominent substrate range. Similarly, we noticed a high percentage of EHs with prominent substrate promiscuity in the chronically polluted seashore area of Milazzo harbor in Sicily (Italy) compared to other sites (Supporting Information, Fig. S1), but again the statistical significance cannot be obtained because the number of enzymes examined was over-represented in this site and because habitats others than those herein examined remains to be explored.

Substrate profiling: general considerations and dynamic range of the assay. Identifying EHs with a broad substrate profile remains a major bottleneck for biocatalysis and biotechnological processes in general.[5] Rapid screening methods will facilitate the identification of such enzymes. Medium-to-high throughput colorimetric protocols have been previously developed that enable EH hydrolytic activity testing with pure proteins.[1,8-11] However, protein purification on a large scale is a very time-consuming and expensive process.[12,13] Consequently, the methods in the present study were adapted to use whole cells, which allowed broad sampling and reduced the time, effort and

cost of identifying enzymes with ample or restricted substrate spectra. Prior to substrate profiling, all clones were tested for activity on agar plates using the three model esters 1-naphthyl acetate, glyceryl tri-acetate and tri-propionate, and their activity was confirmed (see Supporting Information, Methods).

The activity protocol established and used in this study included growing *E. coli* cultures expressing the enzyme of interest overnight in the presence of the appropriate antibiotics and an inducer, followed by activity analysis of the pellet fractions against 96 esters via a pH indicator assay in 384-well plates, at pH 8.0 and 30 °C. Acid is produced after ester bond cleavage by the action of the hydrolytic enzyme contained in the cell pellet,[8] which induces a color change of the pH indicator that can be measured spectrophotometrically at 550 nm. Following the recommendation of Janes *et al*.[8] the concentrations of the pH indicator (0.45 mM) and each of the esters (from 1.3 to 13.2 mM; see concentrations in Methods) were chosen to maximize the accuracy and sensitivity. The number of cells should be as high as possible to maximize sensitivity and to ensure accurate "activity data." In our assay, activity data refers to experimental "time course data" of substrate conversion using 0.4 mg of wet cells per ester, in which time course data might indicate specific activity (in U per g wet cells) from time-course conversion after a maximum of 24 h. This amount of cells (0.4 mg per ester) and time frame (up to 24 h) were found to ensure sensitivity and detection of all esters hydrolyzed by any given EH, as determined by studying the frequency of each ester considered as a hit using different numbers of cells and different time frames and comparing the data from whole-cell assays with pure proteins. Thus, the substrate utilization for all pure proteins supported the validity of the whole-cell screening assay in this study. Under our assay conditions, all esters hydrolyzed by whole cells (Supporting Information, Table S3) were also hydrolyzed when using pure proteins and *vice versa* (Supporting Information, Table S6), which demonstrated that substrate limitation problems using whole cells were not anticipated and that the rapid assay used in this study can be applied to screen substrate ambiguity for any type of cells containing EHs (i.e., fosmid clones, expression clones, microbial cells, etc.). However, we cannot rule out the possibility that unambiguous detection of the conversion of a substrate at a significantly low rate may not be possible using whole cells.

The rapid whole-cell assay used in this study to analyze the substrate profile may be of interest in future genomic and metagenomics screening programs to help identify and prioritize clones with a broad substrate range, whether they contain one or more sequences that encode EHs in a cloned DNA fragment. To screen for promiscuous enzymes, cells from pure cultures, enrichments or clones that express a particular DNA fragment can be screened with standard substrates. Those found to be active against a common standard ester (e.g., 1-napthyl-acetate or glyceryl tri-butyrate)

but do not show a restricted promiscuous profile in a subsequent substrate profiling test will not merit *a priori* sequencing and/or cloning efforts if the objective of the screening program is to find substrate-promiscuous enzymes. By contrast, clones active against a broad range of esters at a high rate merit further sequencing, cloning, expression and characterization. This will substantially reduce the reagent and labor cost while using modest resources. If the activity tests, which can be extended to a broader set of chemical blocks of industrial interest, are also directly performed under conditions resembling industrial requirements, for example in the presence of solvents or at high temperature, efforts can be made toward identifying new, versatile and robust biocatalysts fulfilling industry criteria. Indeed, preliminary tests indicated that the assay method used in this study can be performed in the presence of solvents such as methanol, ethanol, acetonitrile and DMSO at final concentrations of up to 30% (*v*/*v*).

Assays were performed at pH 8.0 and 30 °C in the absence of any chemical other than the ester and a small amount (4.5% v/v) of solvent (DMSO or acetonitrile) needed to dissolve the ester. Therefore, no surfactant was added. Although this may compromise the solubility of very hydrophobic molecules, the majority of large hydrophobic molecules, such as triolein, 2,4-dichlorophenyl 2,4-dichlorobenzoate, 2,4-dichlorobenzyl 2,4-dichlorobenzoate, and diethyl-2,6-dimethyl 4-phenyl-1,4-dihydro pyridine-3,5-dicarboxylate, were hydrolyzed by some EHs under our assay conditions, suggesting that our assay conditions did not introduce any bias in the detection of substrate-enzyme pairs.

Reactions were performed at pH 8.0 and 30 °C. Because we sampled a very broad diversity of habitats from moderately cold to thermophilic, we decided to use an assay temperature of 30 °C, the temperature at which we anticipated most EHs would show appreciable activity. A pH of 8.0 was selected because it was the pH required for the Phenol Red pH indicator used, although pH 7.0 with *p*-nitrophenol as the pH indicator may also be used.[8] We used pH 8.0 because most EHs have been reported to show neutral-to-slightly alkaline pH optima. We are aware that we also included samples from environments ranging from sea-water like habitats to acid mine drainage and that the enzymes from these environments may show different pH optima. Consequently, in many cases (not shown), pH and temperature values for optimal activity were compiled (from previous studies or after the evaluation of optimal parameters for the enzymes investigated herein for the first time) to ensure that the EHs were most active at 30 °C and pH 8.0. Having said that the assay herein used can be extended to pH 7.0 (using *p*-nitrophenol as pH indicator) and at any temperature, preferably below 70ºC to avoid evaporation; a minimum concentration of 0.5 mM substrate is recommended to ensure detection of pH shift.[8] The concentration of buffer (5 mM EPPS) and pH indicator

(Phenol red 0.45 mM) were empirically determined as being optimal to ensure that changes in pH shift during reaction gives linear changes in absorbance.[8]

The predictive capacity of cavity volume/SASA is not influenced by the presence of flexible elements in the structure. The predictive capacity of the active site effective volume may also sensitively depend on the conformational state of flexible elements such as the lid.[5] Indeed, some EH proteins contain a lid domain that covers the active site, preventing substrate binding and requiring structural rearrangement to attain an open conformation. To evaluate whether the homology model-derived active site effective volume value of an EH differs if the template is in the open or closed state, we used the crystal structures of CalA and CalB in closed and open conformations (3GUU/2VEO and 4K6G/5A71). The variations were minimal. Since CalA and CalB do not have the classic lid, which consists of one or two helices that move considerably as rigid bodies, we extended the applicability of this measure to sequences of typical-lid fungal-like lipases such as lipases from *Candida rugosa* (open/close PDB: 1CRL/1THR) and *Thermomyces lanuginosa* (1DTE/1DT5). In all cases, the effective volume ($\geq$76.9 $Å^3$) was above the 62.5 $Å^3$-threshold, which is indicative of high substrate promiscuity, with minimal variations between the open and closed conformations. Therefore, the presence of lid domains does not have an observable effect on the substrate promiscuity level, which is mostly defined by the sequence-defined topology of the active site environment, herein exemplified by the active site effective volume.

An example of the potential of the effective volume measure for bioprospecting of EH with prominent substrate promiscuity. This study also produced a subset of EHs with large substrate ranges, competitive with the best industrial prototypes. Their sequences can be used as targets for bioprospecting similar sequences in large-scale datasets. As example, we used MMSEQS2[14] to screen the occurrence of promiscuous EHs against the predicted open reading frames from the TARA Oceans project assemblies,[15] which was used as a study case. All hits with an E-value threshold of 10^{-10} and sequence coverage $\geq$0.6 were selected (see Supporting Information, Methods). For simplicity, the search was restricted to sequences encoding EH1 to EH6 (assigned to FIV), EH7 (assigned to FVIII), and EH17 (assigned to MCP hydrolase family). All of these EHs were among the most substrate-promiscuous EHs and encompassed phylogenetically and environmentally diverse sequences (see Clusters 1 and 2 in Fig. 2). The data presented in Supporting Information, Fig. S6a reveal that not all sequences encoding promiscuous EHs were equally abundant. The homologues to FVIII serine beta-lactamase EH7 and MCP hydrolase EH17 were the most abundant in the majority of the sites, which is most noticeably when examining their accumulative abundances in all 5 oceanic regions (Supporting Information, Fig. S6b). At this stage we would like to notice that obtaining statistically significant differences to environmental

metadata and to global distributions, albeit of interest, is out the scope of the present study. Rather, we studied the degree of novelty of the EH homologs found in the TARA Oceans samples, by performing an all-vs-all blastp search.[16] Supporting Information, Fig. S7a shows the distribution of the similarity values of every hit against our reference sequences, and Supporting Information, Fig. S7b shows the values for all comparisons. The results revealed that the sequences of EH herein reported are distantly related to those in the TARA Oceans Project, which are also quite diverse. The active site effective volume calculation for all homologs (sequence identity >40%) in Tara Ocean samples indicated that all have volumes ranging from approx. 500 to 66.7 $Å^3$, which are above the 62.5 $Å^3$-throshold, and indicative of a broad substrate range, yet to be elucidated.

METHODS

Protein samples. Of the 145 EHs, the screening, cloning and expression of 69 were previously reported. The expression systems and purification conditions used in this study were as reported previously, in order to proceed with their characterization. These 69 EHs were identified by screening meta-genome clone libraries with the short esters α-naphthyl acetate and/or glyceryl tri-butyrate or, in one case, *p*-nitrophenyl-octanoate. Full details of these enzymes are extensively provided in Supporting Information, Table S1. The remaining 77 sequences encoding EHs are reported in this study for the first time and include 15 from Milazzo harbor (Sicily, Italy),[11,17] 9 from the Los Rueldos acid mine (Mieres, Spain),[18] 8 from a hydrothermal vent metagenome (Jan Mayen Vent Fields) (the meta-sequences are available from the National Center for Biotechnology Information (NCBI) database under the ID PRJNA296938, SAMN04111445), 8 from the River Elbe (Hamburg, Germany),[19] 6 from the El Max site (Alexandria, Egypt),[17] 4 from the Bizerte lagoon (Tunisia),[17] 3 from Messina harbor (Sicily, Italy),[17] 3 from the Gulf of Aqaba (Red Sea, Jordan),[17] 2 from Ancona harbor (Ancona, Italy),[20] 2 from elephant feces (Hagenbeck's Zoo, Hamburg, Germany),[21] 1 from Priolo Gargallo harbor (Syracuse, Italy),[17] 6 from *A. borkumensis* SK2,[22] 5 from *Cycloclasticus* sp. 78-ME,[23] 1 from *Oleiphilus messinensis* ME102T,[24] 1 from *Staphylococcus epidermidis*, 1 from *Pseudomonas oleovorans* DSM-1045 (genome not yet published), and 1 from *Geobacillus* sp. GHH01 (the genome sequence has been deposited in GenBank under accession no. CP004008).[25] In all cases, DNA extraction from the corresponding material,[11,17-25] preparation of pCCFOS1 libraries and naïve screening using α-naphthyl acetate and glyceryl tri-butyrate were performed as described elsewhere.[26] Positive clones containing presumptive EHs were selected, and their DNA inserts were sequenced using a MiSeq Sequencing System (Illumina, San Diego, USA) with a 2 × 150-bp sequencing kit. After sequencing, the reads

were quality-filtered and assembled to generate non-redundant meta-sequences, and genes were predicted and annotated as described previously.[27] The meta-sequences used for *in silico* screening are available from the NCBI database.[17,18,20] Protein-coding genes from metagenomes (sequence-based mining) and from the DNA inserts from positive clones (naïve screening) were screened (score > 45; e-value < $10e^{-3}$) using BLASTP and PSI-BLAST searches[28] for enzymes of interest against the ESTHER (*ESTerases and alpha/beta-Hydrolase Enzymes and Relatives*) and LED (*Lipase Engineering*) databases.[29,30]

As summarized in Supporting Information, Table S1, 128 of 145 genes encoding EHs were available or cloned in this study in common expression vectors (Ek/LIC 46, pET21a, pET22b, pVLT31, pCR-XL-TOPO, and p15Tv-L, among others) using a PCR-based approach and appropriate DNA samples as templates. The remaining set of 18 genes (accession codes KY483640-KY483649 and KY203030-KY203037) were synthesized by GenScript (Hong Kong) Limited in expression vectors. Of these 18 genes, the genes KY483640–KY483649 were synthesized in the pCDFDuet expression vector. The genes KY203030 to KY203037 were cloned using a recently developed vector suite that facilitates sub-cloning based on fragment exchange (FX) into multiple expression vectors.[31,32] Proteins were His-tagged at the C-terminus and purified as previously described (see Supporting Information, Table S1) with slight modifications. Briefly, selected *E. coli* clones that expressed each protein were grown at 37 °C on solid Luria Bertani (LB) agar medium supplemented with the appropriate antibiotics (Table S1), and one colony was picked and used to inoculate 10 mL of LB broth plus antibiotic in a 0.25-L flask. The cultures were then incubated at 37 °C and 200 rpm overnight. Afterward, 10 mL of this culture was used to inoculate 0.5 L of LB medium, which was then incubated to an OD_{600nm} of approximately 0.7 (ranging from 0.55 to 0.75) at 37 °C. Protein expression was induced by adding the appropriate inducer to a final concentration of approx. 1 mM (see Supporting Information, Table S1), followed by incubation for 16 h at 16°C. The cells were harvested by centrifugation at 5000 × *g* for 15 min to yield a pellet of 2-3 g/L pellet (wet weight). The wet cell pellet was frozen at -86 °C overnight, thawed and resuspended in 15 mL of 40 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.0. Lysonase Bioprocessing Reagent (Novagen, Darmstadt, Germany) was added (4 µL/g wet cells) and incubated for 60 min on ice with rotating mixing. The cell suspension was sonicated for a total of 5 min and centrifuged at 15000 × *g* for 15 min at 4°C, and the supernatant was retained. The His-tagged proteins were purified at 4 °C after binding to a Ni-NTA His·Bind resin (Sigma Chemical Co. (St Louis, MO, USA)), followed by extensive dialysis of the protein solutions against 20 mM HEPES buffer (pH 7.0) by ultra-filtration through low-adsorption hydrophilic 10,000 nominal molecular weight limit cutoff membranes (regenerated cellulose, Amicon) and storage at -86°C.

Purity was assessed as >98% using SDS-PAGE analysis[33] in a Mini PROTEAN electrophoresis system (Bio-Rad). The protein concentration was determined according to Bradford with bovine serum albumin as the standard.[34]

Quick tests for the production of active proteins and preparation of protein samples for hydrolysis activity assessment. Prior to substrate profiling, an activity test was performed to verify that each protein was active when expressed in *E. coli*. Protein activity was followed by monitoring the hydrolytic activity directly in the induced cells. Selected *E. coli* clones that expressed each protein were grown at 37°C on solid LB agar media supplemented with the appropriate antibiotics (Supporting Information, Table S1), and one colony was picked and used to inoculate 500 µL of LB broth plus antibiotic in a 2-mL Eppendorf tube. The culture was then incubated at 37 °C and 700 rpm in a Thermomixer (Eppendorf, Hamburg, Germany) for 7 h. To obtain uniform colonial growth, 5 µL of each culture was spotted on the surface of an LB agar plate (90 cm ø) supplemented with the appropriate antibiotic and expression inducer (Supporting Information, Table S1). The plates were incubated overnight at 37°C, and the agar surface was then covered with three different esters, 1-naphthyl acetate, glyceryl tri-acetate and tri-propionate, which are common EH substrates. Because EHs are commonly active at neutral or slightly alkaline pH, the activity tests were performed at pH 7.0. Briefly, the agar surface was covered with a layer of 20 mL of 1-naphthyl acetate/Fast Blue RR salt in HEPES buffer, pH 7.0, containing 0.4% (w/v) agar.[26] The hydrolysis of 1-naphthyl acetate was monitored by following the formation of an intense brown precipitate because of the release of naphthol, which was further oxidized. In parallel, a second plate was covered with a layer of 20 mL of 5 mM *N*-(2-hydroxyethyl)piperazine-*N*'-(3-propanesulfonic acid (EPPS) buffer supplemented with 0.45 mM Phenol Red as a pH indicator, 1 mL of a glyceryl tri-acetate and tri-propionate stock solution (200 mg/mL in acetonitrile each), and 0.4% (w/v) agar. If EH activity occurred, a yellow halo was evident around the colony due to acid formation. Clones containing all 145 EHs investigated in this study tested positive in in at least one of the two screening methods, indicating the production of soluble active proteins, as further confirmed by SDS-PAGE analysis in a Mini PROTEAN electrophoresis system (Bio-Rad).[33]

Selected *E. coli* clones that expressed active and soluble proteins were grown at 37 °C on solid LB agar supplemented with the appropriate antibiotics (Supporting Information, Table S1), and one colony was picked and used to inoculate 500 µL of LB broth plus antibiotic in a 2-mL Eppendorf tube. The culture was then incubated at 37 °C and 700 rpm in a Thermomixer (Eppendorf, Hamburg, Germany) for 7 h, after which 300 µL of each culture was used to seed LB agar petri dishes (90 cm ø) supplemented with the appropriate antibiotic and expression inducer (i.e., isopropyl β-D-1-thiogalactopyranoside or arabinose; see Table S1). The 300-µL culture was spread

on the plate to obtain uniform growth that covered the entire surface of the plate. The plates were incubated overnight at 37 °C. After incubation, 5 mL of 40 mM HEPES buffer, pH 7.0, was added to the surface of each plate. The bacterial cells were detached using sterile disposable Drigalsky spatulas, and the cell suspensions were transferred to a 5-mL Eppendorf tube and pelleted by centrifugation at 8000 rpm for 10 min at 4 °C. The supernatant was discarded, and the pellets were washed and centrifuged twice in the same buffer. The washed, wet pellets were weighed. For the activity tests, 100 mg of wet intact cells was re-suspended in 0.5 mL of 5 mM EPPS buffer, pH 8.0. This suspension was mixed by vortexing for 1 min and 2 µL of these suspensions were used immediately for the activity test as described in Methods. The activity was further confirmed using purified His-tagged proteins. Prior to use, stock solutions of 1-5 mg mL^{-1} protein in 5 mM EPPS buffer, pH 8.0, were prepared after extensive dialysis against this buffer by ultra-filtration through low-adsorption hydrophilic 10,000 nominal molecular weight limit cutoff membranes (regenerated cellulose, Amicon) and 2 µL of protein solutions (from 1.0-17 µg of total protein depending on the specific activity of the particular EH) were immediately used for the activity tests.

Structural determinations. The proteins EH1 and EH102 were expressed and purified according to previously described procedures.[10] The His_6 tags were removed by TEV protease cleavage. The proteins were crystallized using the sitting-drop method in Intelliplate 96-well plates and a Mosquito liquid-handling robot (TTP LabTech), which mixed 0.5 µL of 20-25 mg mL^{-1} protein with 0.5 µL of the following reservoir solutions: EH1 - 0.1 M Tris pH 8.5, 0.2 M ammonium sulfate, 25% (w/v) PEG3350, EH102 - 0.1 M Tris pH 8.5, 0.2 M ammonium sulfate, 25% (w/v) PEG3350, 1/70 units of chymotrypsin. The crystal was cryoprotected with the reservoir solution supplemented with Paratone-N oil prior to flash freezing in an Oxford Cryosystems Cryostream. For the EH1 crystal, diffraction data were collected at 100 K and the Cu K° emission wavelength using a Rigaku HF-0007 rotating anode with a Rigaku R-AXIS IV++ detector. For the EH102 crystal, diffraction data were collected at 100 K and the Se absorption edge wavelength at the Structural Biology Center, Advanced Photon Source, beamline 19-ID using an ADSC Quantum 315R CCD detector. All diffraction data were reduced with HKL3000.[35] The structures were determined by molecular replacement using the structures from the Est2 Protein Databank (PDB) codes 1EVQ and 3RJT for EH1 and EH102, respectively, and the software Phenix.phaser.[36] Refinement was completed with Phenix.refine.[37] All B-factors were refined as isotropic with TLS parameterization. All geometries were verified using Phenix and the wwPDB server.

Protein Energy Landscape Exploration (PELE) sampling. We used Protein Energy Landscape Exploration (PELE) software to sample the binding modes of glyceryl tri-acetate with EH1 and EH102.[38,39] PELE is a Monte Carlo algorithm composed of a sequence of perturbation,

relaxation, and Metropolis acceptance tests. In the first step, the ligand is subjected to random rotations and translations, while the protein is perturbed based on the anisotropic network model (ANM).[40] The maximum allowed translation for ligand perturbation was 1.5 Å, and the maximum rotation was 20°. During the protein perturbation, all atoms were displaced by a maximum of 0.5 Å by moving the α-carbons following a random linear combination of the 6 lowest eigenvectors obtained in the ANM model. The relaxation step included the repositioning of all amino acid side chains within 6 Å of the ligand and the 5 side chains with the highest energy increase along the previous ANM step. The relaxation stage ended with a truncated Newton minimization using the OPLS all-atom force field and an implicit surface-generalized Born continuum solvent. The new proposed minima were then accepted or rejected based on a Metropolis test. The substrate binding plots contained all accepted conformations for three 12-h simulations using 200 processors.

Homology modeling. Homology models were developed using Prime software from Schrödinger. Prime uses BLAST (with BLOSUM62 matrix) for homology search and alignment and refines the results using the Pfam database and pairwise alignment with ClustalW.

Cavity Volume and Solvent Accessible Surface Area (SASA) calculation. The relative Solvent Accessible Surface Area (SASA) for a residue was obtained using the GetArea web server.[41] This service allows a user to submit a PDB file and retrieve the relative SASA or the solvation energy in a variety of formats. This server has several options and allows the user to compute individual residue exposure under the "Select desired level of output" option. Thus, the "exposure" of the active site can be computed using the catalytic amino acids, e.g. the conserved catalytic Ser/Asp/His triad in esterases and the two conserved aspartic catalytic residues in HAD phosphatases. Cavity volumes were computed with Fpocket,[42] a very fast open-source protein pocket (cavity) detection algorithm based on Voronoi tessellation. Fpocket includes two other programs (dpocket and tpocket) that allow the extraction of pocket descriptors and the testing of owned scoring functions, respectively.

SUPPORTING INFORMATION REFERENCES

(1) Ménigaud, S., Mallet, L., Picord, G., Churlaud, C., Borrel, A., Deschavanne, P. (2012) GOHTAM: a website for 'Genomic Origin of Horizontal Transfers, Alignment and Metagenomics'. *Bioinformatics 28*, 1270-1271.

(2) Martínez-Martínez, M., Alcaide, M., Tchigvintsev, A., Reva, O., Polaina, J., Bargiela, R., Guazzaroni, M. E., Chicote, A., Canet, A., Valero, F., Rico Eguizabal, E., Guerrero, Mdel C., Yakunin, A. F., and Ferrer, M. (2013) Biochemical diversity of carboxyl esterases and lipases from Lake Arreo (Spain): a metagenomic approach. *Appl. Environ. Microbiol.* 79, 3553-3562.

(3) Price, M.N., Dehal, P.S., Arkin, A.P. (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5, e9490.

(4) Arpigny, J. L., and Jaeger, K. E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem. J. 343*, 177-183.

(5) Ferrer, M., Bargiela, R., Martínez-Martínez, M., Mir, J., Koch, R., Golyshina, O. V., and Golyshin, P. N. (2015) Biodiversity for biocatalysis: A review of the α/β-hydrolase fold superfamily of esterases-lipases discovered in metagenomes. *Biocatal. Biotransform. 33*, 235-249.

(6) Huang, H., Pandya, C., Liu, C., Al-Obaidi, N.F., Wang, M., Zheng, L., Toews Keating, S., Aono, M., Love, J. D., Evans, B., Seidel, R. D., Hillerich, B. S., Garforth, S. J., Almo, S. C., Mariano, P. S., Dunaway-Mariano, D., Allen, K. N., and Farelli, J. D. (2015) Panoramic view of a superfamily of phosphatases through substrate profiling. *Proc. Natl. Acad. Sci. USA 112,* E1974-1983.

(7) Yarza, P., Yilmaz, P., Pruesse, E., Glöckner, F.O., Ludwig, W., Schleifer, K.H., Whitman, W.B., Euzéby, J., Amann, R., Rosselló-Móra, R. (2014) Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol. 12*, 635-645.

(8) Janes, L. E., Löwendahl, C., and Kazlauskas, R. J. (1998) Rapid quantitative screening of hydrolases using pH indicators. Finding enantioselective hydrolases. *Chem. Eur. J. 4*, 2317-2324.

(9) Alcaide, M., Tornés, J., Stogios, P.J., Xu, X., Gertler, C., Di Leo, R., Bargiela, R., Lafraya, A., Guazzaroni, M. E., López-Cortés, N., Chernikova, T. N., Golyshina, O. V., Nechitaylo, T.Y ., Plumeier, I., Pieper, D. H., Yakimov, M.M., Savchenko, A., Golyshin, P. N., and Ferrer, M. (2013) Single residues dictate the co-evolution of dual esterases: MCP hydrolases from the α/β hydrolase family. *Biochem. J. 454*,157-166.

(10) Alcaide, M., Stogios, P. J., Lafraya, Á., Tchigvintsev, A., Flick, R., Bargiela, R., Chernikova, T. N., Reva, O. N., Hai, T., Leggewie, C.C., Katzke, N., La Cono, V., Matesanz, R., Jebbar, M., Jaeger, K. E., Yakimov, M. M., Yakunin, A. F., Golyshin, P. N., Golyshina, O. V., Savchenko, A., Ferrer, M. (2015) Pressure adaptation is linked to thermal adaptation in salt-saturated marine habitats. *Environ. Microbiol. 17*, 332-345.

(11) Popovic, A., Hai, T., Tchigvintsev, A., Hajighasemi, M., Nocek, B., Khusnutdinova, A. N., Brown, G., Glinos, J., Flick, R., Skarina, T., Chernikova, T. N., Yim, V., Brüls, T., Paslier, D.L., Yakimov, M. M., Joachimiak, A., Ferrer, M., Golyshina, O. V., Savchenko, A., Golyshin, P. N., and Yakunin, A. F. (2017) Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. *Sci. Rep. 7*, 44103.

(12) Barak, Y., Nov, Y., Ackerley, D. F., and Matin, A. (2008) Enzyme improvement in the absence of structural knowledge: a novel statistical approach. *ISME J. 2*,171-179.

(13) Ferrer, M., Martínez-Martínez, M., Bargiela, R., Streit, W. R., Golyshina, O. V., and Golyshin, P. N. (2016) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. *Microb. Biotechnol. 9*, 22-34

(14) Steinegger, M., Soeding, J. (2016) Sensitive protein sequence searching for the analysis of massive

data sets. *bioRxiv* 079681.

(15) Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., d'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., Kokoszka, F., Lepoivre, C., Lima-Mendez, G., Poulain, J., Poulos, B. T., Royo-Llonch, M., Sarmento, H., Vieira-Silva, S., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S.; Tara Oceans coordinators, Bowler, C., de Vargas, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Jaillon, O., Not, F., Ogata, H., Pesant, S., Speich, S., Stemmann, L., Sullivan, M. B., Weissenbach, J., Wincker, P., Karsenti, E., Raes, J., Acinas, S. G., and Bork P. (2015) Structure and function of the global ocean microbiome. *Science 348*, 1261359.

(16) Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T. L (2009) BLAST+: architecture and applications. *BMC Bioinformatics 10*, 421.

(17) Bargiela, R., Mapelli, F., Rojo, D., Chouaia, B., Tornés, J., Borin, S., Richter, M., Del Pozo, M. V., Cappello, S., Gertler, C., Genovese, M., Denaro, R., Martínez-Martínez, M., Fodelianakis, S., Amer, R. A., Bigazzi, D., Han, X., Chen, J., Chernikova, T. N., Golyshina, O. V., Mahjoubi, M., Jaouanil, A., Benzha, F., Magagnini, M., Hussein, E., Al-Horani, F., Cherif, A., Blaghen, M., Abdel-Fattah, Y. R., Kalogerakis, N., Barbas, C., Malkawi, H. I., Golyshin, P. N., Yakimov, M. M., Daffonchio, D., and Ferrer, M. (2015) Bacterial population and biodegradation potential in chronically crude oil-contaminated marine sediments are strongly linked to temperature. *Sci Rep 5*, 11651.

(18) Méndez-García, C., Mesa, V., Sprenger, R. R., Richter, M., Diez, M. S., Solano, J., Bargiela, R., Golyshina, O. V., Manteca, Á., Ramos, J. L., Gallego, J. R., Llorente, I., Martins dos Santos, V. A., Jensen, O. N., Peláez, A. I., Sánchez, J., and Ferrer, M. (2014) Microbial stratification in low pH oxic and suboxic macroscopic growths along an acid mine drainage. *ISME J. 8*, 1259-1274.

(19) Rabausch, U., Juergensen, J., Ilmberger, N., Böhnke, S., Fischer, S., Schubach, B., Schulte, M., Streit, W. R. (2013) Functional screening of metagenome and genome libraries for detection of novel flavonoid-modifying enzymes. *Appl. Environ. Microbiol. 79*, 4551-4563.

(20) Bargiela, R., Gertler, C., Magagnini, M., Mapelli, F., Chen, J., Daffonchio, D., Golyshin, P. N., and Ferrer, M. (2015) Degradation network reconstruction in uric acid and ammonium amendments in oil-degrading marine microcosms guided by metagenomic data. *Front. Microbiol. 6*, 1270.

(21) Güllert, S., Fischer, M. A., Turaev, D., Noebauer, B., Ilmberger, N., Wemheuer, B., Alawi, M., Rattei, T., Daniel, R., Schmitz, R. A., Grundhoff, A., and Streit, W. R. (2016) Deep metagenome and metatranscriptome analyses of microbial communities affiliated with an industrial biogas fermenter, a cow rumen, and elephant feces reveal major differences in carbohydrate hydrolysis strategies. *Biotechnol. Biofuels 9*, 121.

(22) Schneiker, S., Martins dos Santos, V. A., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., Chernikova, T. N., Denaro, R., Ferrer, M., Gertler, C., Goesmann, A., Golyshina, O. V., Kaminski, F., Khachane, A. N., Lang, S., Linke, B., McHardy, A. C., Meyer, F., Nechitaylo, T., Pühler, A., Regenhardt, D., Rupp, O., Sabirova, J. S., Selbitschka, W., Yakimov, M. M., Timmis, K. N., Vorhölter, F. J., Weidner, S.,

Kaiser, O., and Golyshin, P. N. (2006) Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat. Biotechnol. 24*, 997-1004.

(23) Messina, E., Denaro, R., Crisafi, F., Smedile, F., Cappello, S., and Genovese, M. (2016) Genome sequence of obligate marine polycyclic aromatic hydrocarbons-degrading bacterium *Cycloclasticus* sp. 78-ME, isolated from petroleum deposits of the sunken tanker Amoco Milford Haven, Mediterranean Sea. *Mar. Genomics* 25, 11-13.

(24) Golyshin, P. N., Chernikova, T. N., Abraham, W. R., Lünsdorf, H., Timmis, K. N., Yakimov, M. M. (2002) Oleiphilaceae fam. nov., to include *Oleiphilus messinensis* gen. nov., sp. nov., a novel marine bacterium that obligately utilizes hydrocarbons. *Int. J. Syst. Evol. Microbiol. 52*, 901-911.

(25) Wiegand, S., Rabausch, U., Chow, J., Daniel, R., Streit, W. R., Liesegang, H. (2013) Complete genome sequence of *Geobacillus* sp. strain GHH01, a thermophilic lipase-secreting bacterium. *Genome Announc.* 1, e0009213.

(26) Reyes-Duarte, D., Ferrer, M., and García-Arellano, H. (2012) Functional-based screening methods for lipases, esterases, and phospholipases in metagenomic libraries. *Methods Mol Biol 861*, 101-113.

(27) Placido, A., Hai, T., Ferrer, M., Chernikova, T. N., Distaso, M., Armstrong, D., Yakunin, A. F., Toshchakov, S. V., Yakimov, M. M., Kublanov, I. V., Golyshina, O. V., Pesole, G., Ceci, L. R., Golyshin, P. N. (2015) Diversity of hydrolases from hydrothermal vent sediments of the Levante Bay, Vulcano Island (Aeolian archipelago) identified by activity-based metagenomics and biochemical characterization of new esterases and an arabinopyranosidase. *Appl. Microbiol. Biotechnol. 99*, 10031-10046.

(28) Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res. 25*, 3389-3402.

(29) Fischer, M., and Pleiss, J. (2003) The Lipase Engineering Database: a navigation and analysis tool for protein families. *Nucleic Acids Res. 31*, 319-321.

(30) Barth, S., Fischer, M., Schmid, R. D., and Pleiss J. (2004) The database of epoxide hydrolases and haloalkane dehalogenases: one structure, many functions. *Bioinformatics 20*, 2845-2847.

(31) Geertsma, E. R., and Dutzler, R. (2011) A versatile and efficient high-throughput cloning tool for structural biology. *Biochemistry 50*, 3272-3278.

(32) Bjerga, G. E. K., Arsın, H., Larsen, Ø., Puntervoll, P., Kleivdal, H. T. (2016) A rapid solubility-optimized screening procedure for recombinant subtilisins in *E. coli*. *J. Biotechnol*. 20, 38-46.

(33) Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*, 680-685.

(34) Bradford, M. M. (1976) A rapid and sensitive method for the quantification of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem. 72*, 248-254.

(35) Minor, W., Cymborowski, M., Otwinowski, Z., and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution--from diffraction images to an initial model in minutes. *Acta Crystallogr. D Biol. Crystallogr. 62*, 859-866.

(36) 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., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D: Biol. Crystallogr. 66*, 213-221.

(37) Emsley, P., and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D: Biol. Crystallogr. 60*, 2126-2132.

(38) Kaminski, G. A., Friesner, R. A., Tirado-Rives, J., and Jorgensen, W. L. (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J. Phys. Chem. B 105*, 6474-6487.

(39) Borrelli, K. W., Vitalis, A., Alcantara, R., and Guallar, V. (2005) PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. *Chem. Theory Comput. 1*, 1304-1311.

(40) Bochevarov, A. D., Harder, E., Hughes, T. F., Greenwood, J. R., Braden, D. A., Philipp, D. M., Rinaldo, D., Hall, M. D., Zhang, J., and Friesner, R. A. (2013) Jaguar: A high-performance quantum chemistry software program with strengths in life and materials sciences. *Int. J. Quantum. Chem. 113*, 2110-2142.

(41) Fraczkiewicz, R., and Braun, W. (1998) Exact and efficient analytical calculation of the accessible surface areas and their gradients for macromolecules. *J. Comput. Chem. 19*, 319.

(42) Guilloux, V. L., Schmidtke, P., and Tuffery, P. (2009) Fpocket: An open source platform for ligand pocket detection. *BMC Bioinformatics 10*, 168.

2.4 Relationships between substrate promiscuity and chiral selectivity of esterases from phylogenetically and environmentally diverse microorganisms

Cristina Coscolín, Mónica Martínez-Martínez, Jennifer Chow, Rafael Bargiela, Antonio García-Moyano, Gro E. K. Bjerga, Alexander Bollinger, Runar Stokke, Ida H. Steen, Olga V. Golyshina, Michail M. Yakimov, Karl-Erich Jaeger, Alexander F. Yakunin, Wolfgang R. Streit, Peter N. Golyshin, and Manuel Ferrer

Catalysts (2018), 8, 10

Available online:

https://doi.org/10.3390/catal8010010

Own contribution:

Investigation: contribution to the enzyme collection by the identification, production and verification of esterase coding genes from *Alcanivorax borkumensis* SK2 and *Pseudomonas oleovorans* DSM1045. Writing: reviewing and editing the manuscript.

catalysts

MDPI

Communication

Relationships between Substrate Promiscuity and Chiral Selectivity of Esterases from Phylogenetically and Environmentally Diverse Microorganisms

Cristina Coscolín [1,†], Mónica Martínez-Martínez [1,†], Jennifer Chow [2], Rafael Bargiela [1,3], Antonio García-Moyano [4], Gro E. K. Bjerga [4], Alexander Bollinger [5], Runar Stokke [6], Ida H. Steen [6], Olga V. Golyshina [7,8], Michail M. Yakimov [9,10], Karl-Erich Jaeger [5,11], Alexander F. Yakunin [12], Wolfgang R. Streit [2], Peter N. Golyshin [7,8] and Manuel Ferrer [1,*]

1 Institute of Catalysis, Consejo Superior de Investigaciones Científicas, 28049 Madrid, Spain; cristina.coscolin@csic.es (C.C.); m.martinez@csic.es (M.M.-M.); rafaelbb@icp.csic.es (R.B.)
2 Biozentrum Klein Flottbek, Mikrobiologie & Biotechnologie, Universität Hamburg, 22609 Hamburg, Germany; jennifer.chow@uni-hamburg.de (J.C.); wolfgang.streit@uni-hamburg.de (W.R.S.)
3 School of Chemistry, Bangor University, Bangor LL57 2UW, UK; f.bargiela@bangor.ac.uk
4 Uni Research AS, Center for Applied Biotechnology, 5006 Bergen, Norway; antonio.Moyano@uib.no (A.G.-M.); Gro.Bjerga@uni.no (G.E.K.B.)
5 Institute of Molecular Enzyme Technology, Heinrich-Heine-University Düsseldorf, 52426 Jülich, Germany; a.bollinger@fz-juelich.de (A.B.); k.-e.jaeger@fz-juelich.de (K.-E.J.)
6 Department of Biology and KG Jebsen Centre for Deep Sea Research, University of Bergen, 5020 Bergen, Norway; Runar.Stokke@uib.no (R.S.); Ida.Steen@uib.no (I.H.S.)
7 School of Biological Sciences, Bangor University, LL57 2UW Bangor, UK; o.golyshina@bangor.ac.uk (O.V.G.); p.golyshin@bangor.ac.uk (P.N.G.)
8 Centre for Environmental Biotechnology, Bangor University, Bangor LL57 2UW, UK
9 Institute for Coastal Marine Environment, Consiglio Nazionale delle Ricerche, 98122 Messina, Italy; michail.yakimov@iamc.cnr.it
10 Immanuel Kant Baltic Federal University, 236040 Kaliningrad, Russia
11 Institute for Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, 52426 Jülich, Germany
12 Department of Chemical Engineering and Applied Chemistry, University of Toronto, Toronto, ON M5S 3E5, Canada; a.iakounine@utoronto.ca
* Correspondence: mferrer@icp.csic.es; Tel.: +34-91-5854872
† These authors contributed equally to this work.

Received: 22 December 2017; Accepted: 3 January 2018; Published: 5 January 2018

Abstract: Substrate specificity and selectivity of a biocatalyst are determined by the protein sequence and structure of its active site. Finding versatile biocatalysts acting against multiple substrates while at the same time being chiral selective is of interest for the pharmaceutical and chemical industry. However, the relationships between these two properties in natural microbial enzymes remain underexplored. Here, we performed an experimental analysis of substrate promiscuity and chiral selectivity in a set of 145 purified esterases from phylogenetically and environmentally diverse microorganisms, which were assayed against 96 diverse esters, 20 of which were enantiomers. Our results revealed a negative correlation between substrate promiscuity and chiral selectivity in the evaluated enzymes. Esterases displaying prominent substrate promiscuity and large catalytic environments are characterized by low chiral selectivity, a feature that has limited commercial value. Although a low level of substrate promiscuity does not guarantee high chiral selectivity, the probability that esterases with smaller active sites possess chiral selectivity factors of interest for industry (>25) is significantly higher than for promiscuous enzymes. Together, the present study unambiguously demonstrates that promiscuous and selective esterases appear to be rare in nature and that substrate promiscuity can be used as an indicator of the chiral selectivity level of esterases, and vice versa.

Keywords: esterase; metagenomics; promiscuity; selectivity

1. Introduction

Presently, there is a great need for suitable biocatalysts with high process performance as greener alternatives to chemical synthesis [1,2]. It is expected that up to 40% of bulk chemical synthesis processes could be substituted by enzymatic catalysis by 2020 [1]. Along with requirements of a technical nature, such as process development and optimization, it is however widely recognized that the establishment of enzymatic processes is mainly a problem of finding, optimizing, or designing new and/or better performing enzymes. Nature is a rich reservoir from where enzymes can be isolated [3,4], because they are continuously evolving as a consequence of natural selection. Promiscuous enzymes are effective for converting multiple substrates, thus, they are industrially relevant [4–6]. Enzymes need to also be robust and, preferably, chiral selective to reduce raw material costs in the synthesis of pure chiral compounds [1,2,4,7]. That is, they need to be able to cleave preferentially only one chiral ester when offered a racemic mixture. Is it possible to find versatile enzymes displaying prominent substrate range and stringent chiral selectivity? Evaluating this possibility was the starting point of the present study.

In this study, we are interested in investigating as model enzymes serine ester hydrolases, hereafter referred to as esterases, from the structural superfamily of α/β-hydrolases. The activity of these esterases relies mainly on a catalytic triad usually formed by Ser, Asp/Glu, and His [8]. This enzyme class was selected for a number of reasons: it is widely distributed in the environment, it has important physiological functions, it includes hydrolases that are among the most important industrial biocatalysts, and extensive biochemical knowledge has been accumulated [4,5,7].

Just focusing on those from uncultivated microorganisms discovered through metagenomic approaches, esterases with prominent chiral selectivity have been identified and their use in the kinetic resolution of a number of esters is reported. Recent examples include those preferably hydrolyzing one of the chiral esters in racemic mixtures of ibuprofen esters [9,10]; ketoprofen esters [11–14]; solketal esters [15]; esters of phenylalkyl carboxylic acids, 1,1,1-trifluoro-2-phenylbut-3-yn-2-yl acetate and 3,7-dimethyl-1,6-octadien-3-yl acetate [16,17]; methyl 3-phenylglycidate [18]; 1-phenylethyl acetate [19,20]; ofloxacin butyl ester [21]; 1-octin-3-ol, 3-chlor-1-phenyl-1-propanol, trimethylsilylbutinol, *cis*/*trans*-1,2-cyclohexanediol, and isopropylidenglycerol acetate [22]; glycidyl butyrate [23]; methyl-mandelate, glycidyl-4-nitrobenzoate, methyl-3-bromo-2-methyl propionate, methyl lactate, menthyl acetate, neomenthyl acetate, pantolactone, and methyl 3-hydroxybutyrate [22,24,25]; 1-octin-3-ol, 3-chlor-1-phenyl-1-propanol, and trimethylsilylbutinol [22]; methyl-3-hydroxy-2-methylpropionate [26]; and esters of secondary alcohols [27,28], to cite some. The advances in metagenomics techniques and screening methods have allowed the discovery of these and other selective esterases [29]. These studies exemplify that esterases with selective character occur naturally, and that their chiral preference depend on structural factors in the proximity of the active-site. However, whether the selective character of these esterases, and many others, is linked to a broad or a narrow substrate spectrum has not been investigated, due to limited substrate sets employed.

Here, we investigate the relationships between the level of substrate promiscuity and chiral selectivity of a large set of 145 phylogenetically and environmentally diverse microbial esterases, whose specific activity against 96 distinct esters that included 20 chiral esters have been recently reported [5]. We provide unambiguous experimental evidence suggesting a negative association between substrate specificity and chiral selectivity in native esterases.

2. Results and Discussion

2.1. Relationships between Substrate Promiscuity and Chiral Selectivity

We have recently described an extensive analysis of the substrate spectra of 145 phylogenetically and environmentally diverse microbial esterases [5]. Experimental data on substrate conversion (i.e., units g^{-1} or $U \cdot g^{-1}$) followed for 24 h, at pH 8.0 and 30 °C was reported for 96 distinct esters. They included esters with variation in size of acyl and alcohol groups and with growing residues (aromatic, aliphatic, branched, and unbranched), halogenated esters, sugar esters, lactones, an alkyl

di-ester, and 20 chiral esters (including (*R*) and (*S*) enantiomers of menthyl acetate, *N*-benzyl-proline ethyl ester, methyl mandelate, ethyl 4-chloro-3-hydroxybutyrate, methyl 3-hydroxybutyrate, methyl 3-hydroxyvalerate, neomenthyl acetate, methyl and ethyl lactate, and pantolactone). By meaning of the partitioning coefficient (log *p* value), which reflects electronic and steric effects and hydrophobic and hydrophilic characteristics, the 96 esters do show a broad chemical and structural variability [5]. This chemical variability also characterized the chiral esters tested (Figure 1).

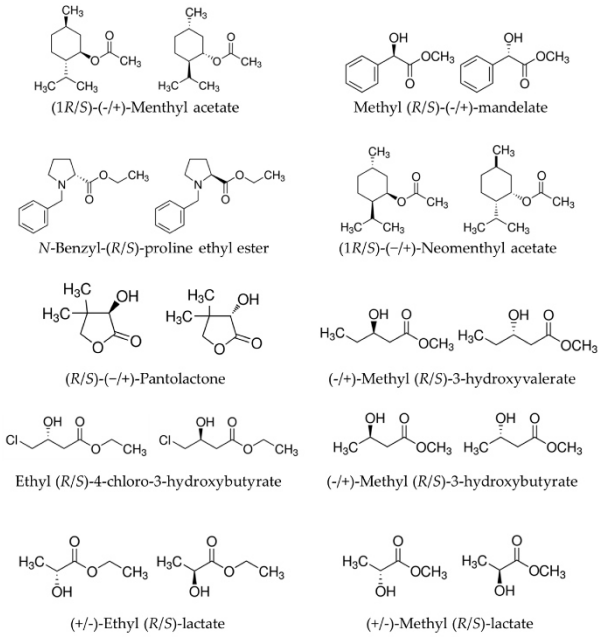


Figure 1. Representative chemical structures of 20 chiral esters used to evaluate chiral selectivity.

To find the relationships between substrate promiscuity and chiral selectivity we calculated the chiral selectivity factor for each of the 145 esterases and the 20 chiral esters tested (i.e., two enantiomers per pair) that were included in the 96-ester library (Figure 1). Selectivity factor was calculated as the ratio of specific activity ($U \cdot g^{-1}$) of the preferred over the non-preferred chiral ester when both esters were tested separately [30] (see Materials and Methods). These calculations were extracted from datasets reported previously [5]. It should be mentioned that these apparent values may not correspond to true selectivity or enantiomeric factors calculated when the enzyme is confronted to a racemic mixture, because the rates of hydrolysis of the enantiomers were measured separately [30]; nevertheless, recent studies have clearly demonstrated that apparent and true selectivity values closely match each other [15]. These values were plotted against the number of esters hydrolyzed by each of the esterases (Figure 2), previously reported for each of the esterases [5].

From the 145 esterases, 40 did not show appreciable activity under assay conditions for any of the chiral esters tested. From those being active against at least one of the chiral esters (105 in total), 80 esterases were characterized by selectivity factors below a threshold of 25. Although esterases with stringent selectivity are preferred, it is commonly considered that enzymes with selectivity factor of 25 or above begin to have commercial value [31]. On the other hand, we found 25 chiral selective esterases, as judged by a selectivity factor above 25 (Figures 2 and 3). Ten of them showed stringent selectivity, that is they were capable of hydrolyzing only one of the enantiomer (Figure 3). Twelve of

them were characterized by selectivity factors ranging from 25.9 to 59.3, and three did show prominent selectivity factors ranging from 219 to 686 (Figures 2 and 3).

As shown in Figure 2, we found a negative association between the level of substrate promiscuity, by meaning of the number of esters hydrolyzed, and the chiral selectivity factor. In further detail, according to criteria previously established [5] we considered an esterase 'specific' if it hydrolyses 9 esters or fewer, as 'moderate-to-highly promiscuous' if it hydrolyses between 10 and 42 esters, and as 'prominently promiscuous' if it hydrolyses 42 or more esters. None of the 25 hydrolases which showed a selectivity factor ≥25 were prominently promiscuous. Rather, they were capable of accepting 36 or fewer substrates. However, not all hydrolases converting 36 or fewer esters and acting against chiral esters were selective according to the 25-selective factor threshold. Indeed, only 25 out of 85 in total (or 29%) were selective, with different selectivity factors and chiral preferences (Figure 3). This is most likely due to the fact that the ability to selectively hydrolyze an enantiomer in a racemate may depend on the topology of the catalytic environment [5].

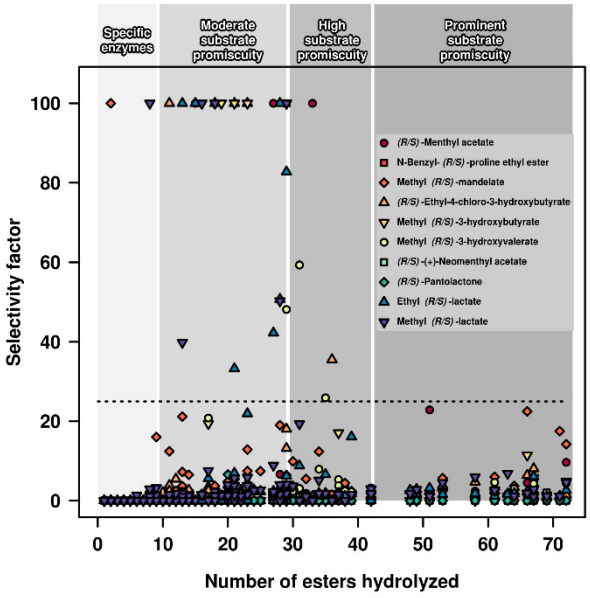


Figure 2. Chiral selectivity factor vs. number of esters hydrolyzed per each of the 145 hydrolases tested. Selectivity factor was calculated per each pair of enantiomers as the ratio of specific activity ($U \cdot g^{-1}$) of the preferred over the non-preferred chiral ester when each of the chiral esters was tested separately. Chiral esters are color coded. The value 100 was arbitrarily given to represent those esterases capable of hydrolyzing, under our assay conditions, only one of the enantiomers (100% selective) and those with selectivity factors higher than 100. These data are based on the data reported previously [5], using conditions described in Materials and Methods. The level of promiscuity, according to criteria previously established [5], is marked under a shadowed grey background. The 25-selectivity factor threshold at which an esterase started to have commercial value is indicated by a horizontal dashed gray line.

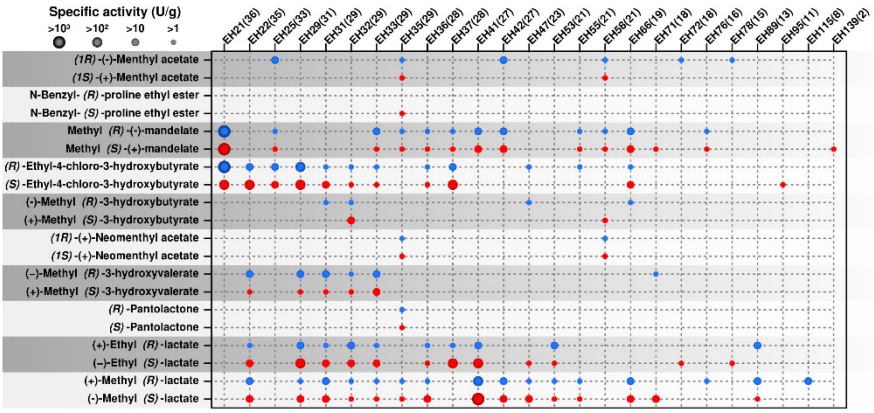


Figure 3. Chiral preferences of 25 hydrolases which were found to be selective for at least one chiral ester according to the 25-selective factor threshold. The figure illustrates the specific activity ($U \cdot g^{-1}$; represented by the size of the circles) of each esterase per each of the 20 chiral esters tested. The ID code for each esterase (for full description see ref. [5]) is shown on the top; the number of esters (out of 96 tested) hydrolyzed by each esterase is shown in brackets. The Figure was created with the R language console from data previously reported [5]. The list of the 20 chiral esters tested is shown on the left, with (*R*)-enantiomer in blue and (*S*)-enantiomer in red color. The protocol established and used to identify the esters hydrolyzed by each esterase is described in Materials and Methods.

2.2. Occurrence of Multi Selective Esterases

Figure 3 summarizes the chiral preference of esterases fitting to the 25-selective factor threshold for each of the 20 chiral esters tested. As can be seen in Figure 3, esterases showed different level of promiscuity and preference for (*R*) or (*S*) methyl acetate, menthyl mandelate, methyl 3-hydroxybutyrate, *N*-benzyl-proline ethyl ester, ethyl-4-chloro-3-hydroxybutyrate, and (m)ethyl lactate. Esterases selective for (*R*) or (*S*)-pantolactone, *N*-benzyl-proline ethyl ester, and neomenthyl acetate were the least abundant, suggesting these chiral esters are less preferred substrates. As shown in Figure 3, we also found that 5 out of 25 esterases fitting to the 25-selective threshold did show stringent selectivity or selectivity factor higher than 25 for several chiral esters differing in chemical and structural nature. They include one being selective for methyl 3-hydroxybutyrate (*R*-selective) and ethyl lactate (*S*-preference) (EH31); one for ethyl-4-chloro-3-hydroxybutyrate (*R*-selective), methyl 3-hydroxybutyrate (*R*-selective), and ethyl lactate (*S*-selective) (EH47); one for methyl mandelate (*S*-selective), methyl 3-hydroxyvalerate (*S*-selective) and methyl lactate (*S*-selective) (EH71); and two for menthyl acetate (*R*-selective) and methyl lactate (*S*-selective) (EH72 and EH78). The other 20 esterases did show the capacity to preferentially hydrolyze only one enantiomer (Figure 3). This suggests that multi selective esterases may have a lower abundancy.

3. Materials and Methods

3.1. Source of Chemicals, Enzymes, and Datasets

All chemicals for which activity data are reported were of the purest grade available and were purchased as reported [5]. The present study used datasets of hydrolytic activity ($U \cdot g^{-1}$) for 145 esterases assayed at 550 nm using 96-structurally diverse esters in 384-well plates. Reactions were followed for 24 h, at pH 8.0 and 30 °C. Datasets are available elsewhere [5].

3.2. Selectivity Factor Calculation

The chiral selectivity factor is defined as the ratio of the specific activity [30] for each enantiomer, measured separately as described previously [5]. Briefly, reaction mixture contains 5 mM *N*-(2-hydroxyethyl)piperazine-*N'*-(3-propanesulfonic acid buffer, pH 8.0, 4.5% (*v*/*v*) acetonitrile or dimethyl sulfoxide, 0.45 mM Phenol Red (used as a pH indicator), a concentration of each of the esters of 1.14 mg·mL^{-1}, and 2 µg of proteins. Reactions were allowed to proceed kinetically at 30 °C and hydrolytic activity (U·g^{-1}) calculated followed for 24 h [5]. Selectivity factor was calculated considering the preferred over the non-preferred chiral ester, whatever the preferred (*R*) or (*S*) ester.

4. Conclusions

Herein, we show the value of the systematic investigation of enzyme activity to deepen our understanding of the relationships between substrate promiscuity and chiral selectivity. By comparing the number of esters that 145 diverse esterases hydrolyze as an indicator of the substrate promiscuity level and their selectivity factors as an indicator of enantio-selectivity, we found unambiguous evidence that esterases with broad substrate spectra do commonly show low selectivity for chiral molecules. In this study, the proportion of esterases with both prominent promiscuity and selectivity approaches zero percent. By contrast, the proportion of esterases with low to moderate promiscuity but prominent selectivity was as high as 29%. This suggests that the substrate promiscuity may be used as an indicator of the selective character of esterases. Promiscuous esterases acting against multiple substrates, while at the same time being enantio-selective, appear to be rare in nature, or at least in the habitats from where the esterases herein described were isolated [5]. As these enzymes are of interest for application purposes [1–6,32], protein engineering and rational design may be needed to obtain esterases being promiscuous and selective for industrial applications. We anticipate that the possibility to transform a promiscuous but not selective esterase into an efficient enantio-selective biocatalyst would require less engineering effort because increasing the selectivity for an enantiomer may involve a reduced number of contacts close to the active sites (for a recent example see reference [33]). Conversely, increasing the substrate spectra of a selective non-promiscuous esterase would require large rearrangement of the catalytic environment which may, at the same time, result in significant reduction or even loss of enantio-selectivity. This is because non-promiscuous esterases are characterized by catalytic environments that are highly exposed and have small volumes, while an esterase for being promiscuous requires a large active site volume and lower relative solvent accessible surface area [5], that are difficult to be designed through few mutations.

Acknowledgments: This project has received funding from the European Union's Horizon 2020 research and innovation program (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project 'INMARE' under grant agreement No. 634486 and ERA-IB 5 'METACAT'. This work was further funded by grants PCIN-2014-107 (within ERA NET IB2 grant nr. ERA-IB-14-030—MetaCat), PCIN-2017-078 (within the ERA-MarineBiotech grant ProBone), BIO2014-54494-R and BIO2017-85522-R from the Spanish Ministry of Economy and Competitiveness. The present investigation was also funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC), grant nr. BB/M029085/1. P.N.G. acknowledges the support of the Supercomputing Wales project, which is partly funded by the European Regional Development Fund (ERDF) via the Welsh Government. O.V.G. and P.N.G. acknowledge the support of the Centre of Environmental Biotechnology Project funded by the European Regional Development Fund (ERDF) through the Welsh Government. A.F.Y. gratefully acknowledges funding from Genome Canada (2009-OGI-ABC-1405) and the NSERC Strategic Network grant IBN. The authors gratefully acknowledge financial support provided by the European Regional Development Fund (ERDF). C. Coscolín thanks the Spanish Ministry of Economy, Industry and Competitiveness for a PhD fellowship (Grant BES-2015-073829).

Author Contributions: M.F., P.N.G., and W.R.S. conceived the study; C.C., M.M.-M., J.C., R.B., A.G.-M., G.E.K.B., A.B., R.S., I.H.S., O.V.G., M.M.Y., K.-E.J. and A.F.Y. contributed to enzyme collection and data analysis; M.F. drafted the manuscript which was revised by all co-authors.

Conflicts of Interest: The authors declare no conflict of interest. The founding sponsors had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Martínez-Martínez, M.; Bargiela, R.; Ferrer, M. Metagenomics and the search for industrial enzymes. In *Biotechnology of Microbial Enzymes*, 1st ed.; Brahmachari, G., Demain, A.L., Adrio, J.L., Eds.; Academic Press: Chennai, India, 2015; pp. 167–184.
2. Martínez-Martínez, M.; Bargiela, R.; Coscolín, C.; Navarro-Fernández, J.; Golyshin, P.N.; Ferrer, M. Functionalization and modification of hydrocarbon-like molecules guided by metagenomics: Enzymes most requested at the industrial scale for chemical synthesis as study cases. In *Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Production of Fuels and Chemicals*; Lee, S.Y., Ed.; Springer International Publishing AG: Cham, Switzerland, 2016; pp. 1–26.
3. Yarza, P.; Yilmaz, P.; Pruesse, E.; Glöckner, F.O.; Ludwig, W.; Schleifer, K.H.; Whitman, W.B.; Euzéby, J.; Amann, R.; Rosselló-Móra, R. Uniting the classification of cultured and uncultured bacteria and archaea using 16S rRNA gene sequences. *Nat. Rev. Microbiol.* **2014**, *12*, 635–645. [CrossRef] [PubMed]
4. Ferrer, M.; Martínez-Martínez, M.; Bargiela, R.; Streit, W.R.; Golyshina, O.V.; Golyshin, P.N. Estimating the success of enzyme bioprospecting through metagenomics: Current status and future trends. *Microb. Biotechnol.* **2016**, *9*, 22–34. [CrossRef] [PubMed]
5. Martínez-Martínez, M.; Coscolín, C.; Santiago, G.; Chow, J.; Stogios, P.; Bargiela, R.; Gertler, C.; Navarro-Fernández, J.; Bollinger, A.; Thies, S.; et al. Determinants and prediction of esterase substrate promiscuity patterns. *ACS Chem. Biol.* **2017**. [CrossRef] [PubMed]
6. Schmid, A.; Dordick, J.S.; Hauer, B.; Kiener, A.; Wubbolts, M.; Witholt, B. Industrial biocatalysis today and tomorrow. *Nature* **2001**, *409*, 258–268. [CrossRef] [PubMed]
7. Ferrer, M.; Bargiela, R.; Martínez-Martínez, M.; Mir, J.; Koch, R.; Golyshina, O.V.; Golyshin, P.N. Biodiversity for Biocatalysis: A review for α/β-hydrolases of the esterase-lipase superfamily as case of Study. *Biocatal. Biotransform.* **2015**, *33*, 235–249. [CrossRef]
8. Aranda, J.; Cerqueira, N.M.; Fernandes, P.A.; Roca, M.; Tuñon, I.; Ramos, M.J. The Catalytic Mechanism of Carboxylesterases: A Computational Study. *Biochemistry* **2014**, *53*, 5820–5829. [PubMed]
9. Elend, C.; Schmeisser, C.; Hoebenreich, H.; Steele, H.L.; Streit, W.R. Isolation and characterization of a metagenome-derived and cold-active lipase with high stereospecificity for (*R*)-ibuprofen esters. *J. Biotechnol.* **2007**, *130*, 370–377. [CrossRef] [PubMed]
10. Chow, J.; Kovacic, F.; Dall Antonia, Y.; Krauss, U.; Fersini, F.; Schmeisser, C.; Lauinger, B.; Bongen, P.; Pietruszka, J.; Schmidt, M.; et al. The Metagenome-derived enzymes LipS and LipT increase the diversity of known lipases. *PLoS ONE* **2012**, *7*, e47665. [CrossRef] [PubMed]
11. Kim, Y.J.; Choi, G.S.; Kim, S.B.; Yoon, G.S.; Kim, Y.S.; Ryu, Y.W. Screening and characterization of a novel esterase from a metagenomic library. *Protein Expr. Purif.* **2006**, *45*, 315–323. [CrossRef] [PubMed]
12. Yoon, S.; Kim, S.; Ryu, Y.; Kim, T.D. Identification and characterization of a novel (*S*)-ketoprofen-specific esterase. *Int. J. Biol. Macromol.* **2007**, *41*, 1–7. [CrossRef] [PubMed]
13. Ngo, T.D.; Ryu, B.H.; Ju, H.; Jang, E.J.; Kim, K.K.; Kim, T.D. Crystallographic analysis and biochemical applications of a novel penicillin-binding protein/beta-lactamase homologue from a metagenomic library. *Acta Crystallogr. D Biol. Crystallogr.* **2014**, *70*, 2455–2466. [CrossRef] [PubMed]
14. Kim, J.; Seok, S.H.; Hong, E.; Yoo, T.H.; Seo, M.D.; Ryu, Y. Crystal structure and characterization of esterase Est25 mutants reveal improved enantioselectivity toward (*S*)-ketoprofen ethyl ester. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 2333–2342. [CrossRef] [PubMed]
15. Ferrer, M.; Golyshina, O.V.; Chernikova, T.N.; Khachane, A.N.; Martins dos Santos, V.A.P.; Yakimov, M.M.; Timmis, K.N.; Golyshin, P.N. Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. *Chem. Biol.* **2005**, *12*, 895–904. [CrossRef] [PubMed]
16. Kourist, R.; Hari Krishna, S.; Patel, J.S.; Bartnek, F.; Hitchman, T.S.; Weiner, D.P.; Bornscheuer, U.T. Identification of a metagenome-derived esterase with high enantioselectivity in the kinetic resolution of arylaliphatic tertiary alcohols. *Org. Biomol. Chem.* **2007**, *5*, 3310–3313. [CrossRef] [PubMed]
17. Fernández-Álvaro, E.; Kourist, R.; Winter, J.; Böttcher, D.; Liebeton, K.; Naumer, C.; Eck, J.; Leggewie, C.; Jaeger, K.E.; Streit, W.; et al. Enantioselective kinetic resolution of phenylalkyl carboxylic acids using metagenome-derived esterases. *Microb. Biotechnol.* **2010**, *3*, 59–64. [CrossRef] [PubMed]
18. Ouyang, L.M.; Liu, J.Y.; Qiao, M.; Xu, J.H. Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl. Biochem. Biotechnol.* **2013**, *169*, 15–28. [CrossRef] [PubMed]

19. Martini, V.; Glogauer, A.; Muller-Santos, M.; Lulek, J.; de Souza, E.; Mitchell, D.; Pedrosa, F.; Krieger, N. First co-expression of a lipase and its specific foldase obtained by metagenomics. *Microb. Cell Factories* **2014**, *13*, 171. [CrossRef] [PubMed]
20. Alnoch, R.C.; Martini, V.P.; Glogauer, A.; Costa, A.C.; Piovan, L.; Muller-Santos, M.; de Souza, E.M.; de Oliveira Pedrosa, F.; Mitchell, D.A.; Krieger, N. Immobilization and characterization of a new regioselective and enantioselective lipase obtained from a metagenomic library. *PLoS ONE* **2015**, *10*, e0114945. [CrossRef] [PubMed]
21. Jeon, J.H.; Kim, J.T.; Kang, S.G.; Lee, J.H.; Kim, S.J. Characterization and its potential application of two esterases derived from the arctic sediment metagenome. *Mar. Biotechnol.* **2009**, *11*, 307–316. [CrossRef] [PubMed]
22. Elend, C.; Schmeisser, C.; Leggewie, C.; Babiak, P.; Carballeira, J.D.; Steele, H.L.; Reymond, J.L.; Jaeger, K.E.; Streit, W.R. Isolation and biochemical characterization of two novel metagenome-derived esterases. *Appl. Environ. Microbiol.* **2006**, *72*, 3637–3645. [CrossRef] [PubMed]
23. Martínez-Martínez, M.; Alcaide, M.; Tchigvintsev, A.; Reva, O.; Polaina, J.; Bargiela, R.; Guazzaroni, M.-E.; Chicote, Á.; Canet, A.; Valero, F.; et al. Biochemical diversity of carboxyl esterases and lipases from Lake Arreo (Spain): A metagenomic approach. *Appl. Environ. Microbiol.* **2013**, *79*, 3553–3562. [CrossRef] [PubMed]
24. Alcaide, M.; Tchigvintsev, A.; Martinez-Martinez, M.; Popovic, A.; Reva, O.N.; Lafraya, A.; Bargiela, R.; Nechitaylo, T.Y.; Matesanz, R.; Cambon-Bonavita, M.A.; et al. Identification and characterization of carboxyl esterases of gill chamber-associated microbiota in the deep-sea shrimp *Rimicaris exoculata* by using functional metagenomics. *Appl. Environ. Microbiol.* **2015**, *81*, 2125–2136. [CrossRef] [PubMed]
25. Placido, A.; Hai, T.; Ferrer, M.; Chernikova, T.N.; Distaso, M.; Armstrong, D.; Yakunin, A.F.; Toshchakov, S.V.; Yakimov, M.M.; Kublanov, I.V.; et al. Diversity of hydrolases from hydrothermal vent sediments of the Levante Bay.; Vulcano Island (*Aeolian archipelago*) identified by activity-based metagenomics and biochemical characterization of new esterases and an arabinopyranosidase. *Appl. Microbiol. Biotechnol.* **2015**, *99*, 10031–10046. [CrossRef] [PubMed]
26. Lee, H.W.; Jung, W.K.; Kim, Y.H.; Ryu, B.H.; Kim, T.D.; Kim, J.; Kim, H. Characterization of a novel alkaline family VIII esterase with *S*-enantiomer preference from a compost metagenomic library. *J. Microbiol. Biotechnol.* **2016**, *26*, 315–325. [CrossRef] [PubMed]
27. Gao, W.; Fan, H.; Chen, L.; Wang, H.; Wei, D. Efficient kinetic resolution of secondary alcohols using an organic solvent-tolerant esterase in non-aqueous medium. *Biotechnol. Lett.* **2016**, *38*, 1165–1171. [CrossRef] [PubMed]
28. Kumar, R.; Banoth, L.; Banerjee, U.C.; Kaur, J. Enantiomeric separation of pharmaceutically important drug intermediates using a metagenomic lipase and optimization of its large scale production. *Int. J. Biol. Macromol.* **2017**, *95*, 995–1003. [CrossRef] [PubMed]
29. Böttcher, D.; Zägel, P.; Schmidt, M.; Bornscheuer, U.T. A microtiter plate-based assay to screen for active and stereoselective hydrolytic enzymes in enzyme libraries. *Methods Mol. Biol.* **2017**, *1539*, 197–204. [PubMed]
30. Gawley, R.E. Do the terms "% ee" and "% de" make sense as expressions of stereoisomer composition or stereoselectivity? *J. Org. Chem.* **2006**, *71*, 2411–2416. [CrossRef] [PubMed]
31. Reetz, M.T. (Ed.) Introduction to directed evolution. In *Directed Evolution of Selective Enzymes: Catalysts for Organic Chemistry and Biotechnology*; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, Germany, 2016; pp. 1–16.
32. Romano, D.; Bonomi, F.; de Mattos, M.C.; de Sousa Fonseca, T.; de Oliveira Mda, C.; Molinari, F. Esterases as stereoselective biocatalysts. *Biotechnol. Adv.* **2015**, *33*, 547–565. [CrossRef] [PubMed]
33. Wikmark, Y.; Svedendahl Humble, M.; Bäckvall, J.E. Combinatorial library based engineering of *Candida antarctica* lipase A for enantioselective transacylation of sec-alcohols in organic solvent. *Angew. Chem. Int. Ed. Engl.* **2015**, *54*, 4284–4288. [CrossRef] [PubMed]

2.5 Identification of organic solvent tolerant carboxylic ester hydrolases for organic synthesis

Alexander Bollinger, Rebecka Molitor, Stephan Thies, Rainhard Koch, Cristina Coscolín, Manuel Ferrer, and Karl-Erich Jaeger

Appl. Env. Microb. (2020), in print

Available online:

https://doi.org/10.1128/AEM.00106-20

Own contribution:

Conceptualization of the study. Investigation: construction and screening of two genomic libraries originating from *Alcanivorax borkumensis* SK2 and *Pseudomonas aestusnigri* VGXO14. Identification, cloning, and verification of the tested enzymes. Establishing of the pH indicator screening method and conducting amino acid sequence comparison. Supervision: oversight of data collection, evaluation, and interpretation. Writing: participation in figure and tables preparation; writing of the original draft manuscript, participation in review and editing of the final manuscript.

AEM Accepted Manuscript Posted Online 28 February 2020
Appl. Environ. Microbiol. doi:10.1128/AEM.00106-20

Identification of organic solvent tolerant carboxylic ester hydrolases for organic synthesis

Alexander Bollinger[1a], Rebecka Molitor[1a], Stephan Thies[1], Rainhard Koch[2], Cristina Coscolín[3], Manuel Ferrer[3] and Karl-Erich Jaeger[1,4*]

[1] Institute of Molecular Enzyme Technology, Heinrich Heine University Duesseldorf, Juelich, Germany

[2] Bayer AG, Leverkusen, Germany

[3] Institute of Catalysis, Consejo Superior de Investigaciones Científicas, Madrid, Spain

[4] Institute for Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Juelich, Germany

[a] AB and RM contributed equally to this work. Authors are listed in alphabetical order.

*Address for correspondence: e-mail: karl-erich.jaeger@fz-juelich.de, Tel +49 2461 613716 Fax: +49 2461 612490

Running title: organic solvent tolerant carboxylic ester hydrolases

Keywords: high throughput screening, polar organic solvent, carboxylic ester hydrolases, *Alcanivorax borkumensis*, *Pseudomonas aestusnigri*

Abstract

Biocatalysis has emerged as an important tool in synthetic organic chemistry enabling the chemical industry to execute reactions with high regio- or enantioselectivity and under usually mild reaction conditions while avoiding toxic waste. Target substrates and products of reactions catalyzed by carboxylic ester hydrolases are often poorly water-soluble and require organic solvents, enzymes are evolved by nature to be active in cells, i.e. in aqueous rather than organic solvents. Therefore, biocatalysts withstanding organic solvents are urgently needed. Current strategies to identify such enzymes rely on laborious tests carried out by incubation in different organic solvents and determination of residual activities. Here, we describe a simple assay useful to screen large libraries of carboxylic ester hydrolases for resistance and activity in water miscible organic solvents. We have screened a set of 26 enzymes, most of them identified in this study, with four different water miscible organic solvents. The triglyceride tributyrin was used as a substrate and fatty acids released by enzymatic hydrolysis were detected by a pH shift indicated by the indicator dye nitrazine yellow. With this strategy, we succeeded to identify a novel highly organic solvent tolerant esterase from *Pseudomonas aestusnigri*. In addition, the newly identified enzymes were tested with sterically demanding substrates, as common in pharmaceutical intermediates, and two enzymes from *Alcanivorax borkumensis* were identified which outcompeted the gold standard ester hydrolase CalB from *Candida antarctica*.

Importance

Major challenges hampering biotechnological applications of esterases include the requirement to accept non-natural and chemically demanding substrates and tolerance of the enzymes towards organic solvents which are often required to solubilize such substrates. We describe here a high throughput screening strategy to identify novel organic solvent tolerant carboxylic ester hydrolases (CEs). Among these enzymes, CEs active against water-insoluble bulky substrates were identified. Our results thus contribute to fostering the identification and biotechnological application of CEs.

50 **Introduction**

51 Enzymes are frequently used in biotechnology and are of high interest for many commercial
52 applications (1–3). Besides the detergent, dairy, or baking industry, they are successfully applied
53 in the fine chemical and pharma sector because of their superior stereo- and regioselectivity (1,
54 2, 4). This is reflected by a steadily growing market for enzymes and products thereof, as well as
55 by industrial attempts to protect intellectual property in this field (5, 6). Indeed, the high
56 demand has contributed to the fact that 2018 was named the Year of Biotechnology (7)
57 according to the fact that private biotech companies raised more money in 2018 than in any
58 previous year.

59 In the last years, the combination of metagenomics and next generation sequencing has
60 resulted in a massive accumulation of sequence data and, as a consequence, *in silico* predictions
61 of numerous novel biocatalysts (8, 9). However, the vast majority of this sequence information
62 is not validated experimentally in terms of confirmation of a proposed function and therefore is
63 of limited use (10).

64 Hydrolases (EC 3) represent one of the most important class of enzymes for biocatalytic
65 applications catalyzing a wealth of different hydrolysis reactions, amidation, kinetic resolution,
66 esterification, polycondensation and many others (11). Among the hydrolases, carboxylic ester
67 hydrolases (CEs) (EC 3.1.1), which catalyze the reversible hydrolysis of carboxylic ester bonds,
68 have found wide applications. This is why novel CEs are targets of screen programs, in which
69 they are identified by different high throughput screening systems including halo formation on
70 agar plates, chromogenic and fluorimetric methods, pH shift detection, fluorescence activated
71 cell sorting (FACS) techniques, microfluidic systems, mass spectroscopic analysis, and others (12,
72 13).

73 Enzymes of this class can be found in every living organism; however, marine
74 hydrocarbonoclastic bacteria, also known as marine crude oil-degrading bacteria, have been
75 shown to be a prolific source for biotechnologically relevant CEs (14). These bacteria live in close
76 contact to alkanes (15), their preferred carbon and energy source, some of which are organic
77 solvents. Hence, it is reasonable to assume that crude oil-degrading bacteria may encode and
78 produce organic solvent tolerant enzymes. The best studied example from this group of bacteria

is *Alcanivorax borkumensis* SK2, with at least 12 different CEs with experimentally proven activity (16–19). In contrast, the crude oil-associated bacterium *Pseudomonas aestusnigri* VGXO14 (20) is almost unexplored with respect to CE activities, but its genome sequence hints at a number of CE-encoding genes (21).

Biotechnological applications of CEs and enzymes in general often require the biocatalyst to operate under non-natural reaction conditions and accept artificial substrates rendering substrate promiscuity and enzyme tolerance for extreme pH, salt, and organic solvents a prerequisite for applications. In organic synthesis in particular, substrates and/or products are usually not water-soluble thus requiring the presence of water miscible organic solvents. Whereas a broad substrate specificity can (at least to a certain extend) be predicted from primary sequence information (17), it is still very difficult to predict solvent tolerance exclusively from primary sequence information. Furthermore, experimental data on solvent tolerance are usually obtained by measuring residual enzyme activities in buffer solutions after prior incubation in organic solvents. Preferably, both incubation and activity measurements should be performed in the presence of organic solvents.

In the present study, we describe a set of 25 CEs, 15 newly identified in this study, from *A. borkumensis* and *P. aestusnigri*. Using a simple high throughput assay, organic solvent tolerant CEs were found and tested for their ability to hydrolyze water-insoluble substrates. As a result, we report on novel CEs with broad substrate promiscuity and high organic solvent tolerance.

101 **Results**

102

103 **Cloning and expression of carboxylic ester hydrolases**

104 Mineral oil-degrading bacteria haven proven as a prolific source for lipolytic enzymes (9, 22, 23).
105 In this study we focused on two marine hydrocarbonoclastic bacteria, namely *Alcanivorax*
106 *borkumensis* and *Pseudomonas aestusnigri*, and screened them for CEs. In total, we constructed
107 a set of 26 different CEs (Table 1, supplementary Table S1) belonging to different families of
108 bacterial lipolytic enzymes (24, 25) and showing an overall low sequence identity
109 (supplementary Fig. S1). Eight of these CEs were first described by Martínez-Martínez et al. (17),
110 one was identified by Hajighasemi et al. (18), one recently by Bollinger et al. (28), and 15 were
111 newly identified in this study. Of the CEs used in this study, 16 were recovered from genome
112 libraries after naïve screening; additionally, 9 were identified through genome sequence
113 searches. All CEs identified from genome sequences and 6 of the CEs recovered from genome
114 libraries were cloned into pET-22b(+) high level expression vectors (Table 1). The remaining 10
115 CEs obtained from library screens were cloned as genomic fragments into pCR-XL-TOPO vectors
116 resulting in mediocre expression levels (Table 1, Table S1). The set was completed by HZ lipase
117 from *Aneurinibacillus thermoaerophilus*, which was previously described as organic solvent
118 tolerant and thermostable (26, 27) and was thus used as benchmark enzyme. In all cases, the
119 presence of enzymatically active CEs was confirmed by hydrolysis of the substrate 4-nitrophenyl
120 butyrate after heterologous expression in *E. coli* BL21(DE3) (data not shown).

121

122 **Screening of CEs for organic solvent tolerance**

123 Organic solvent tolerance of enzymes is usually determined by incubation at a defined solvent
124 concentration for a limited time period (e.g. 30 minutes) and subsequent activity measurement
125 in a buffer without solvent. Determination of enzymatic activity in the absence of organic
126 solvent may give rise to false positive results. We found that a pH indicator-based assay using
127 nitrazine yellow (29) (Fig. 1 A) yields reliable results in the presence of up to 50 % (v/v) organic
128 solvents (Fig. 1 B). Four different water miscible organic solvents were chosen based on their
129 relevance for synthetic organic chemistry (30), namely methanol, acetonitrile, dimethyl

130 sulfoxide, and 1,4-dioxane and tested at two concentrations, 30 % and 50 % (v/v). To exclude
131 enzymes, which are only active for a short time in the presence of organic solvents, a 2 h pre-
132 incubation step was introduced before substrate (tributyrin) was added to the reaction. After
133 addition of the substrate and incubation for 18 h the ratio of the indicator absorptions at 450
134 and 600 nm, respectively, was determined. The absorption of a reaction mixture without
135 substrate was subtracted. This is important when whole cell extracts are used as in this study,
136 which may contain intrinsic CEs active towards membrane lipids. Cloned CEs were expressed
137 either at high levels from promoter P_{T7} in pET22b(+) or at low level from their native promoters
138 in pCR-TOPO-XL based genomic library. *E. coli* cells were perforated by treatment with
139 polymyxin B and cell lysates were transferred into assays plates semiautomatically using a 96-
140 channel pipette (Platemaster, Gilson).

141 The activity data was plotted as a heat map and enzymes were hierarchically clustered
142 according to their activity in different solvent systems, visualized by a row dendrogram (Fig. 2).
143 Three groups of enzymes could be distinguished based on their tolerance towards organic
144 solvents: (I) tolerant enzymes with prominent activity under almost all tested conditions, (IIa)
145 medium tolerant enzymes displaying high activity when a low concentration of DMSO was
146 present, and (IIb) sensitive enzymes showing decreased activity.

147 As expected, the benchmark enzyme HZ lipase (CE01) proved to be tolerant showing activity
148 under all tested conditions. Interestingly, CE13 from *P. aestusnigri* was found to be similarly
149 tolerant exhibiting even higher activity in the presence of 50 % acetonitrile, which was the most
150 disruptive reaction condition tested here. Remarkably, this enzyme did not show prominent
151 activity without solvent added, indicating an activation by organic solvents. Two enzymes were
152 found to be active in all organic solvents except 50 % acetonitrile: CE16 and CE20, with CE20
153 showing higher activity in the presence of 30 % acetonitrile than CE16. Moreover, the majority
154 of enzymes was active at high concentrations of methanol and dimethyl sulfoxide but not 1,4-
155 dioxane or acetonitrile. Activity was detected at 50 %, but not at 30 % (v/v) organic solvent
156 concentration for CE03, CE09, and CE13 with methanol, for CE08 and CE19 with acetonitrile, for
157 CE17 and CE12 with dimethyl sulfoxide, and for CE05, CE21 and CE24 with acetonitrile and 1,4-

158 dioxane. This observation might reflect an activating effect of the organic solvent as described
159 for different enzymes including lipases (31–34).

160 Based on these results, we selected enzymes CE13, CE20 and the benchmark enzyme CE01 for
161 further characterization. The respective cell lysates were incubated with 80 % (v/v) of organic
162 solvents, since most enzymes are rapidly inactivated at concentrations above 70 % (v/v) (35),
163 and the residual activity was determined after 3 h and 24 h (Fig. 2). Under these conditions, the
164 activity of CE01 rapidly decreased during incubation in acetonitrile, 1,4-dioxane and methanol
165 (Fig. 3, A-C); about 38 % of residual activity was retained after 3 h and 21 % after 24 h
166 incubation in dimethyl sulfoxide (Fig. 3, D). The newly identified esterase CE13, which we
167 proposed to be highly solvent tolerant, showed 33 %, 58 %, and 64 % residual activity after 3 h
168 of incubation in acetonitrile, 1,4-dioxane, and methanol, respectively (Fig. 3, A-C). After 24 h of
169 incubation, the residual activity further decreased to less than 10 %. Remarkably, an increased
170 activity was observed in the presence of dimethyl sulfoxide, which resulted in about 264 %
171 activity after 24 h (Fig. 3, D). CE20 appeared less resistant and showed a complete loss of
172 activity when methanol was present and a rapid deactivation by 1,4-dioxane (Fig. 3, B-C). When
173 dimethyl sulfoxide was present for 3 h, less than 10 % residual activity was measured, however,
174 at extended incubation time residual activity was determined as about 18 % (Fig. 3, D). Notably,
175 about 50 % residual activity was detected for CE20 after 3 h of incubation with acetonitrile
176 whereas no activity was left after 24 h (Fig. 3, A). In contrast, no activity was observed with the
177 nitrazine yellow assay in the presence of 50 % acetonitrile indicating that CE20 may at least
178 partly be re-activated when the enzyme is transferred from organic to aqueous solvent.

179 The observation of increased enzyme activity upon incubation in organic solvent was previously
180 connected to a temperature significantly below the enzyme's half-inactivation temperature
181 (34). The half-inactivation temperature (T_{50}) of CE01, CE13 and CE20 was determined as 58°C,
182 56°C, and 57°C, respectively (supplementary Fig. S2). These values do not differ in a range large
183 enough to explain the observation that only CE13 was "activated" upon incubation in DMSO at
184 an assay temperature of 30°C.

185

186 **Screening of CEs active towards substrates with poor water solubility**

187 The ability of CEs to accept multiple substrates is an important property for biocatalytic
188 applications; however, many industrially relevant compounds are poorly soluble in water. We
189 therefore decided to test the newly expressed CEs for their ability to hydrolyze sterically
190 demanding substrates of low solubility in water using the nitrazine yellow assay and 30 % (v/v)
191 dimethyl sulfoxide as co-solvent. Four different substrates of increasing complexity were used
192 which all represent esters of 2-chlorobenzoic acid (CBA), namely with ethanol (#1), xylenol (#2),
193 3-(quinazolin-4-ylamino)phenol (#3) and 3-(4-methoxyphenoxy)-4-oxo-2-(trifluoromethyl)-4*H*-
194 chromen-7-ol (#4) (Fig. 4A). The latter two compounds mimic precursor for approved drugs like
195 the tyrosine-kinase inhibitor gefitinib, used in lung cancer treatment (36), or novel compounds
196 promising for the treatment of different types of cancer (37, 38). CBA is a strong carbonic acid
197 thus enabling the detection also of enzymes with low activities, which may represent potential
198 candidates for enzyme engineering. Remarkably, CE07 hydrolyzed all four substrates (Fig. 4B)
199 and CE03 hydrolyzed substrates #1, #2, and #4 whereas most of the CEs tested could not
200 hydrolyze substrates #3 or #4 (Fig. 4B, supplementary Fig. S2). Substrate #3 was not completely
201 soluble in 30% (v/v) DMSO, however, enzyme activity could be determined by measuring the
202 ratio of absorptions at 450 and 600 nm. These results were confirmed by repeating the
203 reactions with 5 U each of CE07 and CE03 (determined with 4-nitrophenyl butyrate as the
204 substrate) and detection of the products by HPLC (supplementary Fig. S3). A commercial
205 preparation of CalB was included as a reference enzyme known to accept many structurally
206 diverse ester substrates (17). Both CE03 and CE07 hydrolyzed all substrates, whereas CalB
207 hydrolyzed only substrates #1 and #3 (Table 2). In this assay, in contrast to the nitrazine yellow
208 assay, hydrolysis of compound #3 by CE03 could also be demonstrated. CE07 hydrolyzed all 4
209 substrates and was the best performing enzyme with substrate #3.

210 In addition to these results, we also studied the CE substrate specificity with a set of 96
211 chemically and structurally different ester substrates as described recently (17, 39). In this assay
212 system, CE07 was also identified as highly substrate promiscuous accepting 65 different ester
213 substrates, but CE03 exhibited only medium promiscuity hydrolyzing 25 esters. In contrast,
214 some enzymes proving highly substrate promiscuous in this assay system, e.g. CE13 which

hydrolyzed 51 different esters and did show a prominent activity in the presence of organic solvents, were inactive against substrates #3 or #4 (supplementary Table S2).

Discussion

Tolerance against various organic solvents and acceptance of diverse synthetic substrates are both required for applications of CEs in industrial biocatalysis. Substrate promiscuity has recently been investigated in detail (17), but tolerance against organic solvents has not been systematically investigated with a larger set of enzymes. Tolerance against organic solvents is often determined by measuring residual activity of an enzyme after incubation, but not in the presence of a solvent (40–47). The accuracy of this approach is improved by following the time-dependent decrease in enzyme activity over a longer period of time, a method that is not suitable for high throughput screening approaches. On the other hand, a variety of pH shift assays is available allowing to determine enzyme activities also at high throughput (12, 39, 48). To the best of our knowledge, organic solvent tolerance was not systematically investigated using a pH shift assay. pH indicators such as 4-nitrophenol (used at pH 7.0) and phenol red (used at pH 8.0) (17, 48) support concentrations of solvents lower than 30 %. Some indicator compounds such as anilines are known to tolerate high concentrations of organic solvents, e.g. acetonitrile (49), however, they are not suitable to detect shifts from physiological pH. In this study, we observed that the indicator dye nitrazine yellow undergoes a color shift below pH 7 (29, 50) and can thus be used for the determination of pH changes in the presence of different water-miscible organic solvents of up to 50 % (v/v) concentration. Notably, this approach is limited to testing of water miscible organic solvents; non-polar organic solvents form two-phase systems, which are difficult to read out with colorimetric microtiter plate (MTP) scale assays.

Here, we have described an assay applicable for the fast and simple determination of solvent tolerant CEs at high throughput, which was applied to a benchmark CE and 25 CEs from *A. borkumensis* and *P. aestusnigri*, two marine oil-degrading bacteria that were shown to represent a prolific, and, in case of the latter (20), nearly unexplored, source of this class of enzymes. This observation indicates that preferably oil-degrading bacteria may contain the genetic information to produce organic solvent tolerant enzymes. We have identified a number

244 of organic solvent tolerant CEs also active against water-insoluble substrates mimicking
245 industrial relevant compounds.

246 More in details, the method allowed the identification of CEs with outstanding performance in
247 the presence of organic solvents, commonly very harmful to the activity of these enzymes.
248 Particularly, in comparison to other reported organic solvent tolerant CEs, CE13 identified here
249 by the nitrazine yellow assay can withstand organic solvents even at higher concentrations
250 retaining about 30 % residual activity after 3 h incubation in 80 % (v/v) acetonitrile. For
251 comparison, the organic solvent tolerant ARM lipase from *Geobacillus* sp. strain ARM showed a
252 near to complete inactivation after 30 minutes incubation in 30 % (v/v) acetonitrile (47), the
253 lipase from *Staphylococcus saprophyticus* M36 displayed 27 % residual activity after 30 minutes
254 incubation in 25 % (v/v) acetonitrile (42), and the cold-adapted lipase LipP from *Pseudomonas*
255 sp. strain B11-1 was completely inactivated after 1 h incubation in 30 % (v/v) acetonitrile (51).
256 Nevertheless, there are enzymes with reported tolerance against acetonitrile, for example a
257 lipase from *Burkholderia ambifaria* YCJ01 which retained full activity after 24 h of incubation in
258 25 % (v/v) acetonitrile and about 60 % residual activity after 60 days under these conditions
259 (52). Not only this stability in presence of acetonitrile is outstanding, but also an about 3-fold
260 activation after 24 h in the presence of dimethyl sulfoxide at concentrations as high as 80 %.
261 This solvent is very deleterious for CEs because of its highly polar character and, to the best of
262 our knowledge, no example of a CE that shows such activation level has been reported to date.
263 The general phenomenon of enzyme activation upon incubation in increasing concentrations of
264 organic solvents was reported to be connected to a significant difference between the assay
265 temperature and the enzymes thermal half-inactivation point (T_{50}) (34). This might point to a
266 limitation of our approach such that enzymes with a thermal half-inactivation point significantly
267 above 30°C were identified as organic solvent tolerant. However, none of these enzymes was
268 found to be completely inactive, suggesting that all are stable and active at 30°C (note that
269 substrate was added after 2 h incubation at 30°C). Moreover, the T_{50} of enzymes CE01, CE13,
270 and CE20 was determined to differ by 2°C only, suggesting that the observed differences in
271 organic solvent tolerance were not caused by differences between assay temperature and T_{50}.
272 The screening strategy described here can thus speed up the detection of CEs with prominent

273 organic solvent tolerance which is regarded as an important feature for biotechnological
274 applications of CEs.

275 At the same time, the method can facilitate the identification of CEs active against substrates
276 that, because of their poor water solubility, require the addition of high concentration of
277 deleterious solvents. The enzymes CE03 and CE07 can serve as examples, they were found to
278 accept sterically demanding ester substrates often present in pharmaceutically relevant
279 compounds.

280 In conclusion, we have examined 26 CEs, of which the isolation of 11 has been previously
281 reported and 15, to the best of our knowledge, have not been reported previously. Among them CE13
282 from *P. aestusnigri* show high organic solvent tolerance and CE03 and CE07 from *A. borkumensis*
283 exhibit a broad substrate specificity and activity towards complex ester substrates mimicking
284 pharmaceutical building blocks. Furthermore, a screening method with the indicator dye
285 nitrazine yellow was established which allows the fast and simple identification of novel organic
286 solvent tolerant CEs.

287 **Materials and Methods**

288

289 **Construction of genomic libraries**

290 Small insert genomic libraries were constructed with genomic DNA extracted from cells of
291 *Pseudomonas aestusnigri* and *Alcanivorax borkumensis* as described (53). Freeze dried cells of
292 *P. aestusnigri* VGXO14 (DSM 103065) and *A. borkumensis* SK2 (DSM 11573) strains were
293 purchased from the German Collection of Microorganisms and Cell Cultures (DSMZ,
294 Braunschweig, Germany). *P. aestusnigri* was grown in LB-medium Luria/Miller (Carl Roth,
295 Karlsruhe, Germany) and *A. borkumensis* in marine broth 2216 (BD Difco, Heidelberg, Germany)
296 supplemented with 1 % (w/v) sodium pyruvate at 30 °C for two days or until sufficient cell
297 growth was observed. Cells were collected by centrifugation and genomic DNA was extracted by
298 chemical lysis and phenol-chloroform extraction as described earlier (54). The genomic DNA was
299 fragmented by sonication and DNA fragments of 5-10 kB were recovered by extraction from an
300 agarose gel with the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel, Düren, Germany).
301 The DNA fragments were end-repaired by with T4 DNA polymerase (Thermo Fisher Scientific,
302 Darmstadt, Germany) and Klenow fragment (Thermo Fisher Scientific, Darmstadt, Germany),
303 terminal phosphates were cleaved by FastAP (Thermo Fisher Scientific, Darmstadt, Germany),
304 and adenine overhangs were introduced by *Taq* DNA polymerase (Thermo Fisher Scientific,
305 Darmstadt, Germany). Subsequently, the DNA fragments were cloned with the TOPO XL PCR
306 Cloning Kit (Invitrogen, Solingen, Germany) as recommended by the manufacturer. Competent
307 *E. coli* TOP10 cells (Thermo Fisher Scientific, Darmstadt, Germany) were transformed with the
308 recombinant pCR-XL-TOPO plasmid library into by electroporation. Recombinant *E. coli* TOP10
309 were cultivated in LB-medium Luria/Miller (Carl Roth, Karlsruhe, Germany) and auto induction
310 medium (20 g/l tryptone from casein, 5 g/l NaCl, 5 g/l yeast extract, 6 g/l Na_2HPO_4, 3 g/l KH_2PO_4,
311 0.6 % glycerol, 0.2 % lactose, 0.05 % glucose) ((55), modified according to
312 https://openwetware.org/wiki/Lidstrom:Autoinduction_Media) at 37 and 30 °C for DNA
313 replication and protein production, respectively.

314 **Activity-based screening for carboxylic ester hydrolases**

315 Genomic libraries from *P. aestusnigri* and *A. borkumensis* were screened using *E. coli* TOP10 as a
316 host, pCR-XL-TOPO as vector, and tributyrin containing agar plates for the identification of
317 esterase producing clones as described earlier (56). The clone libraries were plated on agar
318 plates (LB-media, 1.5 % (w/v) agar, 50 µg/ml kanamycin, 1.5 % (v/v) tributyrin, and 1.5 g/l gum
319 arabic) and incubated for one day at 37 °C, following incubation for up to one week at room
320 temperature. Clones showing halo formation were collected, grown overnight at 37 °C, 150 rpm
321 in a 100 ml Erlenmeyer flask filled with 10 ml LB-medium Luria/Miller (Carl Roth, Karlsruhe,
322 Germany), supplemented with 50 µg/ml kanamycin. Esterase activity was confirmed as
323 described (57), using 4-nitrophenyl butyrate (*p*NPB) as substrate. Plasmid DNA was extracted
324 from active clones with the innuPREP Plasmid Mini Kit 2.0 (Analytik Jena, Jena, Germany). The
325 size of the inserted DNA fragment was determined by hydrolysis with *Eco*RI (Thermo Fisher
326 Scientific, Darmstadt, Germany), followed by agarose gel electrophoresis. The terminal ends of
327 the insert DNA were Sanger sequenced (by Eurofins Genomics, Ebersberg, Germany) using the
328 oligonucleotides included in the TOPO XL PCR Cloning Kit (Invitrogen, Solingen Germany). The
329 resulting sequences were mapped to the genomes of *P. aestusnigri* (RefSeq:
330 NZ_NBYK00000000.1) or *A. borkumensis* (RefSeq: NC_008260.1) to identify the complete insert
331 sequence of the corresponding DNA fragment. To identify CE encoding genes, insert DNA
332 sequences were analyzed by searching GenBank and using the tools ORF finder (58) and BASys
333 annotation (59). The gene encoding the HZ lipase from *Aneurinibacillus thermoaerophilus* strain
334 HZ (designated as CE01) was amplified from a metagenomic library clone (Bollinger,
335 unpublished). For high level expression of selected CEs, genes were PCR amplified with Phusion
336 High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Darmstadt, Germany) following the
337 manufacturer's recommendations using specific oligonucleotides (Table 3) and subsequently
338 cloned into pET-22b(+) vector (Novagen, Darmstadt, Germany) by sequence and ligase
339 independent cloning (60) or directional cloning using restriction and ligation (61) with
340 endonucleases *Nde*I, or *Xba*I in combination with *Xho*I, or *Hin*dIII (Thermo Fisher Scientific,
341 Darmstadt, Germany).

342 **Sequence-based screening and cloning of esterases**

343 In addition to CE genes identified by activity-based screening, CE genes were identified by a text
344 search (search terms: lipase, carboxylesterase or esterase) of the GenBank file containing the
345 reference sequences for *P. aestusnigri* (RefSeq: NZ_NBYK00000000.1) and *A. borkumensis*
346 (RefSeq: NC_008260.1). The respective genes were cloned into pET-22b(+) (Novagen,
347 Darmstadt, Germany) as described above, using specific oligonucleotides (Table 4).

348 **Expression of carboxylic ester hydrolases**

349 CE producing strains *E. coli* BL21(DE3) (62) carrying pET-22b(+) and *E.* coli TOP10 carrying pCR-
350 XL-TOPO were grown in triplicate for 24 h at 37 °C and 800 rpm in deep well plates with 1 ml LB-
351 medium Luria/Miller (Carl Roth, Karlsruhe, Germany) supplemented with the appropriate
352 antibiotic and 0.5 % glucose. 20 µl of these cultures were used to inoculate expression cultures
353 in 980 µl autoinduction medium (20 g/l tryptone from casein, 5 g/l NaCl, 5 g/l yeast extract, 6 g/l
354 Na_2HPO_4, 3 g/l KH_2PO_4, 0.6 % glycerol, 0.2 % lactose, 0.05 % glucose) ((55), modified according
355 to https://openwetware.org/wiki/Lidstrom:Autoinduction_Media) with the respective
356 antibiotic, which were incubated for 20 h at 30 °C under agitation with 800 rpm. The cultures
357 were harvested by centrifugation, the supernatants discarded, the cells suspended in 100 µl cell
358 lysis solution containing polymyxin B (10 mM potassium phosphate buffer pH 7.2, 0.1 mg/ml
359 polymyxin B), and incubated for 1 h at 37 °C.

360 **Amino acid sequence analysis of carboxylic ester hydrolases**

361 Amino acid sequences of CEs used in this study were aligned with a set of enzymes representing
362 known examples of each family of bacterial lipolytic enzymes (25). The alignment was
363 performed using Clustal Omega (63), the phylogenetic tree was constructed with IQ-TREE (64)
364 under default conditions and the graphical representation was done using iTOL (65). The global
365 sequence identity matrix was obtained using Clustal Omega multiple sequence alignment with
366 the amino acid sequences of CEs used in this study.

367 **Nitrazine yellow assay to determine organic solvent tolerance**

368 Organic solvent tolerance of CEs was determined by mixing 100 µl of the CE containing cell
369 extracts with 100 µl of the respective solvent in a microtiter plate (MTP) to reach final solvent
370 concentrations of 0 %, 30 %, and 50 % (v/v) and incubation for 2 h at 30 °C. During incubation,

371 the MTP lid was sealed with organic solvent stable tape to prevent evaporation. After pre-
372 incubation with organic solvents, 10 µl of the sample were combined with 180 µl nitrazine
373 yellow containing assay buffer (5 mM potassium phosphate buffer pH 7.2, 20 µg/ml nitrazine
374 yellow, and 0 %, 30 %, or 50 % (v/v) of the respective organic solvent) and 10 µl of substrate
375 solution (200 mM tributyrin in acetonitrile) or 10 µl acetonitrile for the control. In case of a color
376 shift after addition of the organic solvent, the pH was titrated to neutral (blue color) with
377 potassium hydroxide solution. The reaction mixture was incubated for 18 h at 30 °C and
378 afterwards measured for pH change. The activities were measured using a TECAN infinite
379 M1000 Pro photometer at the absorption maxima of the indicator dye λ = 450 and 600 nm. The
380 quotient of the absorption values determined at both wavelengths was used to measure the pH
381 shift. Each value was corrected by subtraction of the control which did not contain substrate
382 before calculation of mean values and standard deviations. To reduce false positives, values in
383 the range of the standard deviation of the empty vector control were considered as not active
384 (NA).

385 **Heatmap plot**

386 The language R and the package gplots function were used to write a script allowing to plot the
387 activity data obtained from the nitrazine yellow assay in form of a heatmap. The code for
388 generating the heatmap is given as supporting method in the supplementary material.

389 **Determination of organic solvent tolerance**

390 The CE producing *E. coli* BL21(DE3) cells carrying the pET-22b(+) vector and *E.* coli TOP10 cells
391 carrying the pCR-XL-TOPO vector were grown for 24 h at 37 °C and 150 rpm in 100 ml
392 Erlenmeyer flasks with 10 ml LB medium supplemented with appropriate antibiotic and 0.5 %
393 glucose. The expression cultures were inoculated in 250 ml Erlenmeyer flasks with 25 ml auto
394 induction medium (20 g/l tryptone from casein, 5 g/l NaCl, 5 g/l yeast extract, 6 g/l Na_2HPO_4,
395 3 g/l KH_2PO_4, 0.6 % glycerol, 0.2 % lactose, 0.05 % glucose) ((55), modified according to
396 https://openwetware.org/wiki/Lidstrom:Autoinduction_Media) with antibiotic to an optical
397 density at λ = 580 nm of 0.05 and incubated for 20 h at 30 °C under agitation of 160 rpm. The
398 main cultures were collected by centrifugation, the supernatant was discarded, and the cells
399 were suspended in 1/10 of the original volume with 100 mM potassium phosphate buffer

400 pH 7.2. Cells were lysed by sonication, the cell suspension was tested for esterase activity using
401 4-nitrophenyl butyrate as the substrate (57) and diluted accordingly. The cell suspension was
402 mixed with 80 % (v/v) of organic solvents 1,4-dioxan, acetonitrile, methanol or dimethyl
403 sulfoxide and incubated at 30 °C. After 0, 3, and 24 h of incubation, 20 µl of the solution was
404 mixed with 180 µl assay solution (100 mM potassium phosphate buffer pH 7.2, 1 mM 4-
405 nitrophenyl butyrate, 5 % (v/v) acetonitrile) and esterase activity was determined at λ = 410 nm,
406 30 °C for 10 minutes using a TECAN infinite M1000 Pro photometer.

407 **Determination of activity towards water insoluble substrates**

408 CEs were produced as mentioned above and tested with the nitrazine yellow assay as described
409 above with the following modifications: the enzymes were tested in the presence of 30 % (v/v)
410 DMSO without pre-incubation and the substrates were #1: ethyl 2-chlorobenzoate; #2: 3,5-
411 dimethylphenyl 2-chlorobenzoate; #3: 3-(quinazolin-4-ylamino)phenyl 2-chlorobenzoate; #4: 3-
412 (4-methoxyphenoxy)-4-oxo-2-(trifluoromethyl)-4*H*-chromen-7-yl 2-chlorobenzoate (kindly
413 provided by Bayer AG, Leverkusen, Germany). The heatmap was calculated and plotted as
414 described above.

415 **Measurement of half-inactivation temperature**

416 The thermostability of CE01, CE13, and CE20 was investigated by measuring the enzyme half-
417 inactivation temperatures (T_{50}). The enzymes were produced with *E. coli* LOBSTR cells (66)
418 carrying the respective recombinant pET-22b(+) vector. The expression cultures were inoculated
419 from precultures in 5000 ml Erlenmeyer flasks with 500 ml auto induction medium as described
420 above and incubated for 24 h at 30 °C (CE01), 25 °C (CE13), or 37 °C (CE20) at 160 rpm. The
421 cultures were collected by centrifugation (30 minutes at 6,000 × g, 4 °C), the supernatant was
422 discarded, and the cells were stored at -20 °C.

423 For protein purification, cells were suspended in purification buffer (20 mM Na_2HPO_4 pH 7.4,
424 500 mM NaCl, 10 mM imidazole) at 10 % (w/v) and lysed with a high-pressure homogenizer
425 (EmulsiFlex-C5, AVESTIN Europe, GmbH) with three passages at 8,000 psi. The soluble protein
426 fraction was obtained by centrifugation (30 minutes, 4 °C, 36,000 × g) and passed through
427 2.5 ml equilibrated Ni-NTA matrix (Ni-NTA Superflow, Qiagen GmbH) by gravity flow. After
428 washing with at least 10 column volumes (CV) of purification buffer, bound proteins were eluted

429 with 8 ml of elution buffer (20 mM Na_2HPO_4 pH 7.4, 500 mM NaCl, 500 mM imidazole). The
430 elution fraction was concentrated by centrifugal ultrafiltration (Vivaspin 20, 10,000 MWCO,
431 Satorius AG) prior to the buffer exchange to 100 mM potassium phosphate buffer pH 7.2,
432 100 mM NaCl by using PD-10 desalting columns (GE Healthcare) according to the
433 manufacturer's recommendation. The purified protein fractions were stored at -20°C.

434 The enzyme half-inactivation temperatures were determined with enzyme solutions diluted with
435 100 mM potassium phosphate buffer pH 7.2 to an activity of about 1 U/ml measured with 4-
436 nitrophenyl butyrate as the substrate, incubated in a PCR plate, sealed with adhesive aluminum
437 foil and incubated at various temperatures (40 – 80 °C) for 1 h using a gradient thermocycler
438 Biometra TAdvanced (Analytik Jena, Jena). Subsequently residual enzyme activity was measured
439 with 4-nitrophenyl butyrate as the substrate (57). The data obtained from three reactions was
440 plotted (mean and standard deviation) using GraphPad Prism (GraphPad Software, Inc., USA). A
441 nonlinear fit (Boltzmann sigmoidal) was used to calculate the half-inactivation temperature.

442 **Detection of 2-chlorobenzoic acid by HPLC**

443 After determination of esterase activity with *p*NPB as described (57), 5 U of the respective
444 enzyme was mixed with substrate solution to give a final concentration of 5 mM of the
445 compounds #1 to #4, 70 mM potassium phosphate buffer pH 7.2, and 30 % (v/v) dimethyl
446 sulfoxide as co-solvent in PTFE capped glass vials. The reaction mixtures were incubated for 18 h
447 at 30 °C. Subsequently, the mixes were filtered through 0.22 µm pore size PTFE filters and
448 analyzed for 2-CBA by HPLC performed as described (67) using a Accucore™ C18 LC-column
449 (100 mm x 2.1 mm, 2.6 µm particle size, 80 Å pore size, Thermo Scientific) on a LC10-Ai LC
450 system (Shimadzu, Duisburg, Germany), with a gradient of water/acetonitrile (solvent A is water
451 with 0.1 % formic acid; solvent B is acetonitrile with 0.1 % formic acid, start at 5 % B; hold at 5 %
452 B for 1.5 min; gradient from 5 % B to 98 % B in 5.5 min; hold at 98 % B for 2 min; from 98 % B to
453 5 % B in 0.5 min and hold at 5 % B for 2 min to reequilibrate) at a flow rate of 1 ml/min. The
454 retention time of 2-CBA was determined as 4.78 minutes using a pure standard. The integral of
455 the respective signal was used to quantify the amount of 2-CBA released from the substrates
456 based on the calibration line from a log serial dilution of 2-CBA.

457 **Determination of substrate specificity**

458 Aside esters #1 to #4, an additional set of 96 esters with different degree of solubility, were also
459 tested to evaluate the degree of substrate promiscuity. The specific activity (units mg^{-1})
460 determinations were assayed at 550 nm using a pH indicator (phenol red; $\varepsilon_{550\ nm}$ = 8450 M^{-1} cm^{-}
461 ^{1}) assay at 550 nm in 384-well plates as previously described (17, 39). Briefly, cells were grown
462 overnight at 37 °C on solid agar medium containing inducer and antibiotics. Cells were washed
463 from the plates, collected by centrifugation and lysed by sonication after mixing in a vortex for
464 1 min in 5 mM N-(2-hydroxyethyl) piperazine N0-(3-propanesulfonic acid) buffer (EPPS buffer),
465 adjusted to pH 8.0 with NaOH. The lysed cells were combined with 96 different esters as
466 substrates and phenol red as pH indicator in 384-well plates giving a final concentration of
467 1.14 mg/ml of the respective ester, 0.45 mM phenol red, 4.5 % acetonitrile and about 1 mg/ml
468 of the lysed cells in 44 µl EPPS buffer pH 8.0. Reactions were incubated at 30 °C and the
469 hydrolysis was followed at 550 nm for 24 h to calculate specific enzyme activities. Calculations
470 were performed in triplicates and corrected for non-enzymatic transformation.

471

472 **Acknowledgements**

473 The authors received funding from the European Union's Horizon 2020 research and innovation
474 program (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project
475 'INMARE' under grant agreement No. 634486. ST is financially supported by the Ministry of
476 Culture and Science of the German State of North Rhine-Westphalia within in the framework of
477 the NRW Strategieprojekt BioSC (No. 313/323-400-00213). MF acknowledges grant PCIN-2017-
478 078 (within the Marine Biotechnology ERA-NET) and BIO2017-85522-R from the Ministry of
479 Science, Innovation and Universities with the cofunds of the ERDF and the Agencia Estatal de
480 Investigación (AEI). C. Coscolín thanks the Spanish Ministry of Economy, Industry and
481 Competitiveness for a PhD fellowship (Grant BES-2015-073829).

482

483 References
484

485 1. Singh R, Kumar M, Mittal A, Mehta PK. 2016. Microbial enzymes: industrial progress in
486 21st century. 3 Biotech 6:174.

487 2. Reetz MT. 2013. Biocatalysis in organic chemistry and biotechnology: past, present, and
488 future. J Am Chem Soc 135:12480–12496.

489 3. Sheldon RA, Woodley JM. 2018. Role of biocatalysis in sustainable chemistry. Chem Rev
490 118:801–838.

491 4. Patel RN. 2018. Biocatalysis for synthesis of pharmaceuticals. Bioorg Med Chem 26:1252–
492 1274.

493 5. De Godoy Daiha K, Angeli R, De Oliveira SD, Almeida RV. 2015. Are lipases still important
494 biocatalysts? A study of scientific publications and patents for technological forecasting.
495 PLoS One 10:1–20.

496 6. Carlson R. 2016. Estimating the biotech sector's contribution to the US economy. Nat
497 Biotechnol 34:247–255.

498 7. Hodgson J. 2019. Biotech's baby boom. Nat Biotechnol 37:502–512.

499 8. Tully BJ, Graham ED, Heidelberg JF. 2018. The reconstruction of 2,631 draft metagenome-
500 assembled genomes from the global oceans. Sci Data 5:170203.

501 9. Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina O V., Golyshin PN. 2015.
502 Estimating the success of enzyme bioprospecting through metagenomics: current status
503 and future trends. Microb Biotechnol 9:22–34.

504 10. Ferrer M, Méndez-García C, Bargiela R, Chow J, Alonso S, García-Moyano A, Bjerga GEK,
505 Steen IH, Schwabe T, Blom C, Vester J, Weckbecker A, Shahgaldian P, de Carvalho CCCR,
506 Meskys R, Zanaroli G, Glöckner FO, Fernández-Guerra A, Thambisetty S, de la Calle F,
507 Golyshina O V, Yakimov MM, Jaeger K-E, Yakunin AF, Streit WR, McMeel O, Calewaert J-B,
508 Tonné N, Golyshin PN, INMARE Consortium. 2019. Decoding the ocean's microbiological
509 secrets for marine enzyme biodiscovery. FEMS Microbiol Lett 366.

510 11. Goswami A, Van Lanen SG. 2015. Enzymatic strategies and biocatalysts for amide bond
511 formation: Tricks of the trade outside of the ribosome. Mol Biosyst 11:338–353.

512 12. Peña-García C, Martínez-Martínez M, Reyes-Duarte D, Ferrer M. 2016. High throughput
513 screening of esterases, lipases and phospholipases in mutant and metagenomic libraries:
514 A review. Comb Chem High Throughput Screen 19:605–615.

515 13. Jaeger K-E, Kovacic F. 2014. Determination of lipolytic enzyme activities, p. 111–134. *In*
516 Filloux, A, Ramos, JL (eds.), *Pseudomonas* Methods and Protocols. Methods in Molecular
517 Biology (Methods and Protocols). Humana Press, New York, NY.

518 14. Coscolín C, Bargiela R, Martínez-Martínez M, Alonso S, Bollinger A, Thies S, Chernikova
519 TN, Hai T, Golyshina O V., Jaeger K-E, Yakimov MM, Golyshin PN, Ferrer M. 2018.
520 Hydrocarbon-degrading microbes as sources of new biocatalysts, p. 353–373. *In*
521 McGenity, TJ (ed.), Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading
522 Microbes. Springer, Cham, Cham.

523 15. Brooijmans RJW, Pastink MI, Siezen RJ. 2009. Genomics update hydrocarbon-degrading
524 bacteria: the oil-spill clean-up crew 2:587–594.

525 16. Tchigvintsev A, Tran H, Popovic A, Kovacic F, Brown G, Flick R, Hajighasemi M, Egorova O,
526 Somody JC, Tchigvintsev D, Khusnutdinova A, Chernikova TN, Golyshina O V., Yakimov
527 MM, Savchenko A, Golyshin PN, Jaeger K-E, Yakunin AF. 2015. The environment shapes
528 microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine
529 metagenomes. Appl Microbiol Biotechnol 99:2165–2178.

530 17. Martínez-Martínez M, Coscolín C, Santiago G, Chow J, Stogios PJ, Bargiela R, Gertler C,
531 Navarro-Fernández J, Bollinger A, Thies S, Méndez-García C, Popovic A, Brown G,
532 Chernikova TN, García-Moyano A, Bjerga GEKK, Pérez-García P, Hai T, Del Pozo M V.,
533 Stokke R, Steen IH, Cui H, Xu X, Nocek BP, Alcaide M, Distaso M, Mesa V, Peláez AI,
534 Sánchez J, Buchholz PCFF, Pleiss J, Fernández-Guerra A, Glöckner FO, Golyshina O V.,
535 Yakimov MM, Savchenko A, Jaeger K-E, Yakunin AF, Streit WR, Golyshin PN, Guallar V,
536 Ferrer M, The Inmare Consortium. 2018. Determinants and prediction of esterase
537 substrate promiscuity patterns. ACS Chem Biol 13:225–234.

538 18. Hajighasemi M, Tchigvintsev A, Nocek B, Flick R, Popovic A, Hai T, Khusnutdinova AN,
539 Brown G, Xu X, Cui H, Anstett J, Chernikova TN, Brüls T, Le Paslier D, Yakimov MM,
540 Joachimiak A, Golyshina O V., Savchenko A, Golyshin PN, Edwards EA, Yakunin AF. 2018.
541 Screening and characterization of novel polyesterases from environmental metagenomes
542 with high hydrolytic activity against synthetic polyesters. Environ Sci Technol 52:12388–
543 12401.

544 19. Hajighasemi M, Nocek BP, Tchigvintsev A, Brown G, Flick R, Xu X, Cui H, Hai T, Joachimiak
545 A, Golyshin PN, Savchenko A, Edwards EA, Yakunin AF. 2016. Biochemical and structural
546 insights into enzymatic depolymerization of polylactic acid and other polyesters by

547 microbial carboxylesterases. Biomacromolecules 17:2027–2039.

548 20. Sánchez D, Mulet M, Rodríguez AC, David Z, Lalucat J, García-valdés E. 2014.
549 *Pseudomonas aestusnigri* sp. nov., isolated from crude oil-contaminated intertidal sand
550 samples after the Prestige oil spill. Syst Appl Microbiol 37:89–94.

551 21. Gomila M, Mulet M, Lalucat J, García-Valdés E. 2017. Draft genome sequence of the
552 marine bacterium *Pseudomonas aestusnigri* VGXO14^{T}. Genome Announc 5:e00765-17.

553 22. Cafaro V, Izzo V, Notomista E, Di Donato A. 2013. Marine hydrocarbonoclastic bacteria, p.
554 373–402. *In* Trincone, A (ed.), Marine Enzymes for Biocatalysis. Woodhead Publishing.

555 23. Popovic A, Hai T, Tchigvintsev A, Hajighasemi M, Nocek B, Khusnutdinova AN, Brown G,
556 Glinos J, Flick R, Skarina T, Chernikova TN, Yim V, Brüls T, Paslier D Le, Yakimov MM,
557 Joachimiak A, Ferrer M, Golyshina O V., Savchenko A, Golyshin PN, Yakunin AF. 2017.
558 Activity screening of environmental metagenomic libraries reveals novel carboxylesterase
559 families. Sci Rep 7:44103.

560 24. Arpigny JL, Jaeger KE. 1999. Bacterial lipolytic enzymes: classification and properties.
561 Biochem J 343 Pt 1:177–183.

562 25. Kovacic F, Babic N, Krauss U, Jaeger K-E. 2019. Classification of lipolytic enzymes from
563 bacteria, p. 1–35. *In* Rojo, F (ed.), Aerobic Utilization of Hydrocarbons, Oils and Lipids.
564 Springer International Publishing, Cham.

565 26. Masomian M, Rahman RNZRA, Salleh AB, Basri M. 2013. A new thermostable and organic
566 solvent-tolerant lipase from *Aneurinibacillus thermoaerophilus* strain HZ. Process
567 Biochem 48:169–175.

568 27. Masomian M, Abd Rahman RNZR, Salleh AB, Basri M. 2016. Analysis of comparative
569 sequence and genomic data to verify phylogenetic relationship and explore a new
570 subfamily of bacterial lipases. PLoS One 11:1–20.

571 28. Bollinger A, Thies S, Knieps-Grünhagen E, Kobus S, Höppner A, Smits SH, Ferrer M, Jaeger
572 K-E. 2020. A novel polyester hydrolase from the marine bacterium *Pseudomonas*
573 *aestusnigri* - structural and functional insights. Front Microbiol 11:114.

574 29. Griswold KE. 2003. pH sensing agar plate assays for esterolytic enzyme activity. Methods
575 Mol Biol 230:203–211.

576 30. Ferrer M, Bargiela R, Martínez-Martínez M, Mir J, Koch R, Golyshina O V., Golyshin PN.
577 2015. Biodiversity for biocatalysis: A review of the α/β-hydrolase fold superfamily of
578 esterases-lipases discovered in metagenomes. Biocatal Biotransformation 33:235–249.

579 31. Torres C, Otero C. 1996. Influence of the organic solvents on the activity in water and the
580 conformation of *Candida rugosa* lipase: Description of a lipase-activating pretreatment.
581 Enzyme Microb Technol 19:594–600.

582 32. Kamal MZ, Yedavalli P, Deshmukh M V., Rao NM. 2013. Lipase in aqueous-polar organic
583 solvents: activity, structure, and stability. Protein Sci 22:904–15.

584 33. Kumar A, Dhar K, Kanwar SS, Arora PK. 2016. Lipase catalysis in organic solvents:
585 advantages and applications. Biol Proced Online 18:2.

586 34. Cerdobbel A, De Winter K, Aerts D, Kuipers R, Joosten HJ, Soetaert W, Desmet T. 2011.
587 Increasing the thermostability of sucrose phosphorylase by a combination of sequence-
588 and structure-based mutagenesis. Protein Eng Des Sel 24:829–834.

589 35. Stepankova V, Bidmanova S, Koudelakova T, Prokop Z, Chaloupkova R, Damborsky J.
590 2013. Strategies for stabilization of enzymes in organic solvents. ACS Catal 3:2823–2836.

591 36. Kazandjian D, Blumenthal GM, Yuan W, He K, Keegan P, Pazdur R. 2016. FDA approval of
592 Gefitinib for the treatment of patients with metastatic EGFR mutation-positive non-small
593 cell lung cancer. Clin Cancer Res 22:1307–1312.

594 37. Jafari E, Khajouei MR, Hassanzadeh F, Hakimelahi GH, Khodarahmi GA. 2016.
595 Quinazolinone and quinazoline derivatives: recent structures with potent antimicrobial
596 and cytotoxic activities. Res Pharm Sci 11:1–14.

597 38. Puppala M, Zhao X, Casemore D, Zhou B, Aridoss G, Narayanapillai S, Xing C. 2016. 4*H*-
598 Chromene-based anticancer agents towards multi-drug resistant HL60/MX2 human
599 leukemia: SAR at the 4th and 6th positions. Bioorg Med Chem 24:1292–7.

600 39. Reyes-Duarte D, Coscolín C, Martínez-Martínez M, Ferrer M, García-Arellano H. 2018.
601 Functional-based screening methods for detecting esterase and lipase activity against
602 multiple substrates, p. 109–117. *In* Sandoval, G (ed.), Lipases and Phopholipases.
603 Methods in Molecular Biology, 2nd ed. Humana Press.

604 40. Lima VM., Krieger N, Mitchell D., Fontana J. 2004. Activity and stability of a crude lipase
605 from *Penicillium aurantiogriseum* in aqueous media and organic solvents. Biochem Eng J
606 18:65–71.

607 41. Rahman RNZRA, Baharum SN, Basri M, Salleh AB. 2005. High-yield purification of an
608 organic solvent-tolerant lipase from *Pseudomonas* sp. strain S5. Anal Biochem 341:267–
609 274.

610 42. Fang Y, Lu Z, Lv F, Bie X, Liu S, Ding Z, Xu W. 2006. A newly isolated organic solvent
611 tolerant *Staphylococcus saprophyticus* M36 produced organic solvent-stable lipase. Curr
612 Microbiol 53:510–515.

613 43. Yan J, Yang J, Xu L, Yan Y. 2007. Gene cloning, overexpression and characterization of a
614 novel organic solvent tolerant and thermostable lipase from *Galactomyces geotrichum*
615 Y05. J Mol Catal B Enzym 49:28–35.

616 44. Zhao L-L, Xu J-H, Zhao J, Pan J, Wang Z-L. 2008. Biochemical properties and potential
617 applications of an organic solvent-tolerant lipase isolated from *Serratia marcescens*
618 ECU1010. Process Biochem 43:626–633.

619 45. Zhang A, Gao R, Diao N, Xie G, Gao G, Cao S. 2009. Cloning, expression and
620 characterization of an organic solvent tolerant lipase from *Pseudomonas fluorescens*
621 JCM5963. J Mol Catal B Enzym 56:78–84.

622 46. Ahmed EH, Raghavendra T, Madamwar D. 2010. An alkaline lipase from organic solvent
623 tolerant *Acinetobacter* sp. EH28: Application for ethyl caprylate synthesis. Bioresour
624 Technol 101:3628–3634.

625 47. Ebrahimpour A, Rahman RNZRA, Basri M, Salleh AB. 2011. High level expression and
626 characterization of a novel thermostable, organic solvent tolerant, 1,3-regioselective
627 lipase from *Geobacillus* sp. strain ARM. Bioresour Technol 102:6972–6981.

628 48. Janes LE, Löwendahl AC, Kazlauskas RJ. 1998. Quantitative screening of hydrolase libraries
629 using pH indicators: Identifying active and enantioselective hydrolases. Chem - A Eur J
630 4:2324–2331.

631 49. Jordan F. 1973. Acidity scales in mixed water-acetonitrile buffer solutions. J Phys Chem
632 77:2681–2683.

633 50. Kirkwood J, Wilson J, O'keefe S, Hargreaves D. 2014. A high-throughput colourimetric
634 method for the determination of pH in crystallization screens. Acta Crystallogr Sect D Biol
635 Crystallogr 70:2367–2375.

636 51. Choo DW, Kurihara T, Suzuki T, Soda K, Esaki N. 1998. A cold-adapted lipase of an Alaskan
637 psychrotroph, *Pseudomonas* sp. strain B11-1: gene cloning and enzyme purification and
638 characterization. Appl Environ Microbiol 64:486–491.

639 52. Yao C, Cao Y, Wu S, Li S, He B. 2013. An organic solvent and thermally stable lipase from
640 *Burkholderia ambifaria* YCJ01: Purification, characteristics and application for chiral
641 resolution of mandelic acid. J Mol Catal B Enzym 85–86:105–110.

642 53. Nacke H, Will C, Herzog S, Nowka B, Engelhaupt M, Daniel R. 2011. Identification of novel
643 lipolytic genes and gene families by screening of metagenomic libraries derived from soil
644 samples of the German Biodiversity Exploratories. FEMS Microbiol Ecol 78:188–201.

645 54. Troeschel SC, Drepper T, Leggewie C, Streit WR, Jaeger K-E. 2010. Novel tools for the
646 functional expression of metagenomic DNA, p. 117–139. *In* Streit, W, Daniel, R (eds.),
647 Metagenomics. Methods in Molecular Biology (Methods and Protocols). Humana Press.

648 55. Studier FW. 2005. Protein production by auto-induction in high density shaking cultures.
649 Protein Expr Purif 41:207–234.

650 56. Reyes-Duarte D, Ferrer M, García-Arellano H. 2012. Functional-based screening methods
651 for lipases, esterases, and phospholipases in metagenomic libraries, p. 101–113. *In*

652 Sandoval, G (ed.), Lipases and Phospholipases. Methods in Molecular Biology (Methods
653 and Protocols). Humana Press.

654 57. Nolasco-Soria H, Moyano-López F, Vega-Villasante F, del Monte-Martínez A, Espinosa-
655 Chaurand D, Gisbert E, Nolasco-Alzaga HR. 2018. Lipase and phospholipase activity
656 methods for marine organisms, p. 139–167. *In* Sandoval, G (ed.), Lipases and
657 Phospholipases. Methods in Molecular Biology (Methods and Protocols). Humana Press.

658 58. Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD,
659 Schriml LM, Sequeira E, Tatusova TA, Wagner L. 2003. Database resources of the National
660 Center for Biotechnology. Nucleic Acids Res 31:28–33.

661 59. Van Domselaar GH, Stothard P, Shrivastava S, Cruz JA, Guo A, Dong X, Lu P, Szafron D,
662 Greiner R, Wishart DS. 2005. BASys: a web server for automated bacterial genome
663 annotation. Nucleic Acids Res 33:W455-459.

664 60. Jeong JY, Yim HS, Ryu JY, Lee HS, Lee JH, Seen DS, Kang SG. 2012. One-step sequence-and
665 ligation-independent cloning as a rapid and versatile cloning method for functional
666 genomics Studies. Appl Environ Microbiol 78:5440–5443.

667 61. Green MR, Sambrook J. 2012. Molecular cloning: a laboratory manual, 4th ed. Cold Spring
668 Harbor Laboratory Press, Cold Spring Harbor, N.Y.

669 62. Studier FW, Moffatt BA. 1986. Use of bacteriophage T7 RNA polymerase to direct
670 selective high-level expression of cloned genes. J Mol Biol 189:113–130.

671 63. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert
672 M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality

673 protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.

674 64. Trifinopoulos J, Nguyen L-T, von Haeseler A, Minh BQ. 2016. W-IQ-TREE: a fast online
675 phylogenetic tool for maximum likelihood analysis. Nucleic Acids Res 44:W232–W235.

676 65. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and
677 annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242–W245.

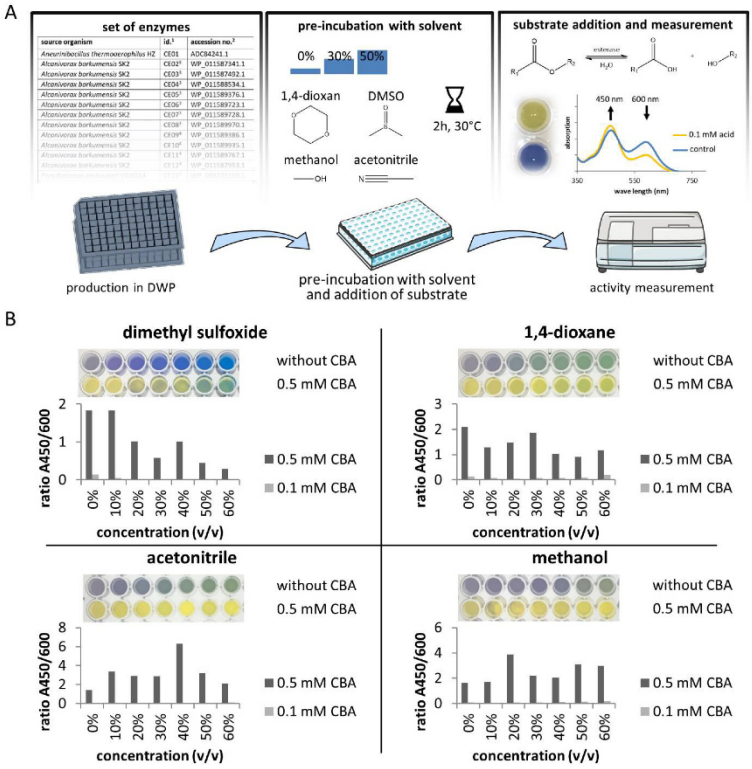
678 66. Andersen KR, Leksa NC, Schwartz TU. 2013. Optimized *E. coli* expression strain LOBSTR
679 eliminates common contaminants from His-tag purification. Proteins 81:1857–1861.

680 67. Domröse A, Weihmann R, Thies S, Jaeger K-E, Drepper T, Loeschcke A. 2017. Rapid
681 generation of recombinant *Pseudomonas putida* secondary metabolite producers using
682 yTREX. Synth Syst Biotechnol 2:310–319.

683

684 **Figures and Tables**

685

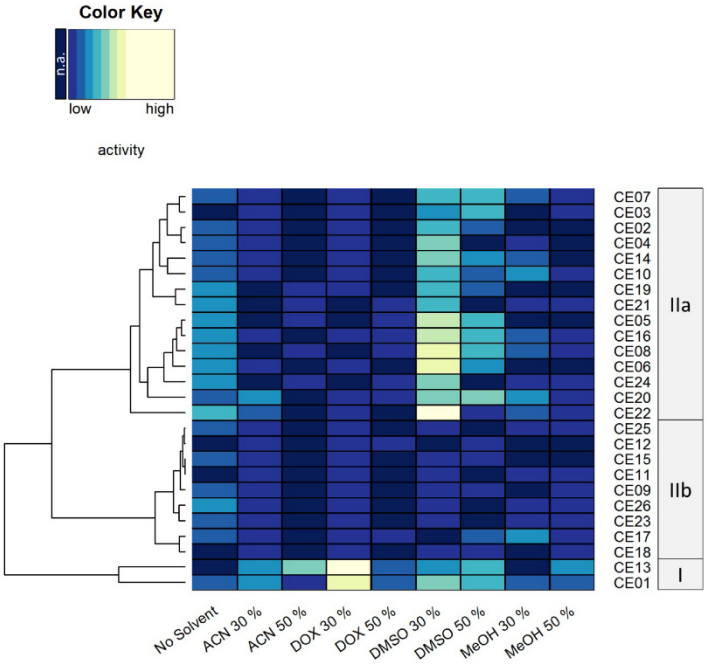


686

687 **Figure 1** High throughput identification of organic solvent tolerant CEs using the nitrazine yellow
688 assay. (A) Workflow of the nitrazine yellow assay. Pictures of the 96-well plate and the plate
689 reader were retrieved from servier medical art (https://smart.servier.com/), licensed under
690 Creative Commons Attribution 3.0 (CC BY). (B) Nitrazine yellow (20 µg/ml) was mixed with
691 different concentrations of organic solvent and potassium phosphate buffer (5 mM) titrated

692 with potassium hydroxide solution (10 mM) until a neutral pH was reached (indicated by a light
693 green to blue color) After addition of 2-chlorobenzoate (CBA), the pH shift was measured
694 photometrically by determining the ratio of absorbances at 450 and 600 nm compared to a
695 control without CBA.

696

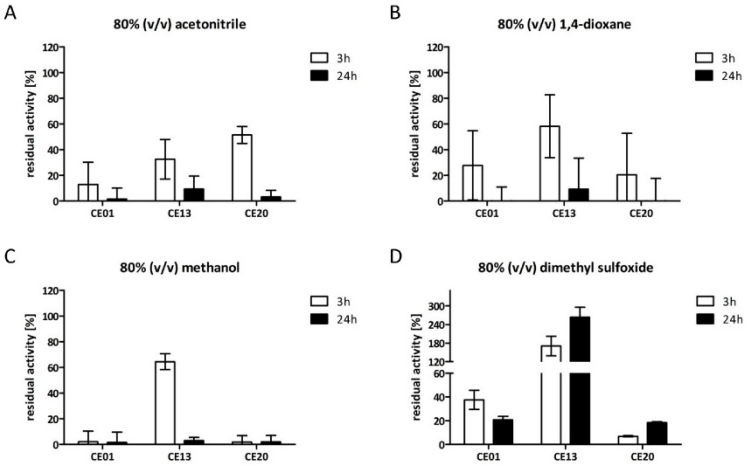


697

698 **Figure 2** Heatmap representation of CE activity in the presence of different water miscible
699 organic solvents. Each row represents an individual enzyme, with the enzyme identifier depicted
700 on the right side. Columns stand for respective organic solvents indicated at the bottom. The
701 dendrogram on the left side indicates a hierarchical clustering of CEs based on their activity in
702 the presence of different organic solvents. CE classes of different solvent tolerance are indicated

703 in grey boxes on the right. The activity data are visualized with dark blue (not active, n.a.) to
704 yellow (highly active) colors indicated by the color key. Conditions tested were without addition
705 of organic solvent (no solvent), acetonitrile (ACN), 1,4-dioxane (DOX), dimethyl sulfoxide
706 (DMSO), or methanol (MeOH) at 30 % or 50 % (v/v) concentration. Reactions were carried out at
707 30°C, 18 h, 5 mM potassium phosphate buffer pH 7.2 containing 20 µg/ml nitrazine yellow.

708



709

710 **Figure 3** Residual CE activity after incubation in the presence of organic solvents. Enzymes were
711 incubated for 3 h and 24 h in 80 % (v/v) concentration of acetonitrile (A), 1,4-dioxane (B),
712 methanol (C), or dimethyl sulfoxide (D). Residual activity was determined with 4-nitrophenyl
713 butyrate as the substrate and calculated relative to the initial activity of the respective enzyme
714 set as 100 %. Error bars indicate standard deviations of three separate experiments. Reactions
715 were conducted at 30°C, 100 mM potassium phosphate buffer pH 7.2.

716

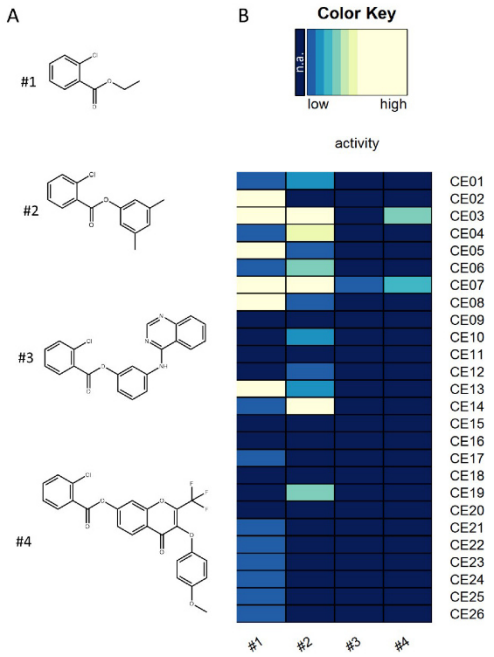


Figure 4 Hydrolysis of 2-chlorobenzoic acid esters in the presence of 30 % (v/v) dimethyl sulfoxide determined by the nitrazine yellow assay. (A) Structural formulas of tested compounds #1 - #4, (B) heatmap plot of enzyme activities. Substrates are #1: ethyl 2-chlorobenzoate. #2: 3,5-dimethylphenyl 2-chlorobenzoate. #3: 3-(quinazolin-4-ylamino)phenyl 2-chlorobenzoate. #4: 3-(4-methoxyphenoxy)-4-oxo-2-(trifluoromethyl)-4*H*-chromen-7-yl 2-chlorobenzoate. Each row of the heatmap represents an individual enzyme with the enzyme identifier indicated on the right side. Each column represents a different substrate. The activity data are visualized from dark blue (not active, n.a.) to yellow (highly active) as indicated by the color key. Reaction conditions were 18 h incubation at 30°C in 5 mM potassium phosphate buffer pH 7.2 containing 20 µg/ml nitrazine yellow, 30 % (v/v) dimethyl sulfoxide, 5 % (v/v) acetonitrile and 10 mM of the substrates #1 - #4.

729 **Table 1** Carboxylic ester hydrolases cloned and expressed in this study.

source organism	id.[1]	accession no.[2]	vector	reference
Aneurinibacillus thermoaerophilus HZ	CE01	ADC84241.1	pET-22b(+)	(26)
Alcanivorax borkumensis SK2	CE02[3]	WP_011587341.1	pET-22b(+)	(17)
Alcanivorax borkumensis SK2	CE03[3]	WP_011587492.1	pET-22b(+)	(17)
Alcanivorax borkumensis SK2	CE04[3]	WP_011588534.1	pCR-XL-TOPO	(17)
Alcanivorax borkumensis SK2	CE05[3]	WP_011589376.1	pCR-XL-TOPO	(17)
Alcanivorax borkumensis SK2	CE06[3]	WP_011589723.1	pCR-XL-TOPO	(17)
Alcanivorax borkumensis SK2	CE07[3]	WP_011589728.1	pET-22b(+)	this study
Alcanivorax borkumensis SK2	CE08[3]	WP_011589970.1	pCR-XL-TOPO	(17)
Alcanivorax borkumensis SK2	CE09[4]	WP_011589386.1	pET-22b(+)	this study
Alcanivorax borkumensis SK2	CE10[4]	WP_011589935.1	pET-22b(+)	(17)
Alcanivorax borkumensis SK2	CE11[4]	WP_011589767.1	pET-22b(+)	(17)
Alcanivorax borkumensis SK2	CE12[4]	WP_011587953.1	pET-22b(+)	(19)
Pseudomonas aestusnigri VGXO14	CE13[3]	WP_088275369.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE14[3]	WP_088277870.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE15[3]	WP_088277153.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE16[4]	WP_088276085.1	pET-22b(+)	(28)
Pseudomonas aestusnigri VGXO14	CE17[4]	WP_088276582.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE18[4]	WP_088273225.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE19[4]	WP_088277509.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE20[4]	WP_088273217.1	pET-22b(+)	this study
Pseudomonas aestusnigri VGXO14	CE21[3]	WP_088273788.1	pCR-XL-TOPO	this study
Pseudomonas aestusnigri VGXO14	CE22[3]	SEG59772.1	pCR-XL-TOPO	this study
Pseudomonas aestusnigri VGXO14	CE23[3]	WP_088274564.1	pCR-XL-TOPO	this study
Pseudomonas aestusnigri VGXO14	CE24[3]	WP_088275865.1	pCR-XL-TOPO	this study
Pseudomonas aestusnigri VGXO14	CE25[3]	WP_088273867.1	pCR-XL-TOPO	this study
Pseudomonas aestusnigri VGXO14	CE26[3]	n.d.[5]	pCR-XL-TOPO	this study

730 [1]Enzyme identifier used in this study.

731 [2] NCBI-accession number of the respective protein.

732 [3] CEs identified by naïve screening.

733 [4]CEs identified by genome mining.

734 [5]The coding sequence of this esterase was not determined; the DNA fragment carried by the library clone was
735 identical to *Pseudomonas aestusnigri* VGXO14 scaffold00001 NBYK01000001.1 position 282473 to 286927 (see also
736 supplementary Table S1).

737

738 **Table 2** Enzyme activity of CE03, CE07 and CalB towards substrates #1 - #4. The hydrolysis
739 product 2-chlorobenzoic acid (CBA) was detected by HPLC and is given as mean concentration
740 from three independent reactions including standard deviations (±). A reaction mix without
741 addition of enzyme served as a control; CBA concentrations in the range of the control reaction
742 indicated no activity (n.a.). Substrates were #1: ethyl 2-chlorobenzoate. #2: 3,5-dimethylphenyl
743 2-chlorobenzoate. #3: 3-(quinazolin-4-ylamino)phenyl 2-chlorobenzoate. #4: 3-(4-
744 methoxyphenoxy)-4-oxo-2-(trifluoromethyl)-4*H*-chromen-7-yl 2-chlorobenzoate. Reaction
745 conditions were 30°C, 18 h, 5 U of enzyme, 5 mM substrate, 30 % DMSO, 70 mM potassium
746 phosphate buffer pH 7.2.

	CE03		CE07		CalB		no enzyme	
substrate	mean [mM]	±	mean [mM]	±	mean [mM]	±	mean [mM]	±
#1	4.340	1.124	3.593	1.290	5.177	0.408	0.020	0.000
#2	4.823	0.667	3.297	2.013	n.a.		0.033	0.019
#3	1.007	0.295	4.563	0.034	0.143	0.005	0.053	0.009
#4	3.740	0.120	1.983	0.581	n.a.		0.033	0.017

747

748 **Table 3** List of oligonucleotides for PCR amplification and cloning of CEs identified by naïve
749 screening. Restriction endonuclease sites used for directional cloning are underlined;
750 oligonucleotides without marked restriction site were used for sequence and ligase
751 independent cloning (SLIC).

id.[1]	forward (5' -> 3')	reverse (5' -> 3')
CE01	CTTTAAGAAGGAGATATACATATGCAAAAGGAAAGAAAAAATC	CAGTGGTGGTGGTGGTGGTGCTCTCTCACAGATAATGAACC
CE02	GCTCATATGAATCCTGCCGTTATTGAG	TACCTCGAGCAACCGCCGCTTGGTCTCAAC
CE03	CTTTAAGAAGGAGATATACATATGGCTTCTATTCCCGCAC	GTGGTGGTGGTGGTGGTGCTCTGACGATATCTCCGGGATTG
CE07	GTCCATATGAGCCTTCAAGCCCG	TACCTCGAGTGCTTCTTTAATGAATGCGACAATC
CE13	GCGCATATGCCTCAATCTTTTAAAC	CTTCTCGAGGGGCAATACCAGCGGCG
CE14	CTTTAAGAAGGAGATATACATATGAGCGGACTCAACCGG	CAGTGGTGGTGGTGGTGGTGCTCGCTGAGCGTCGGCACCAG
CE15	GCGCATATGTCCAGGTACGTTGATG	CGCCTCGAGGCTTACCGAGTCGGCCTG

752 [1]Enzyme identifier as used in this study.

753

754 **Table 4** List of oligonucleotides for PCR amplification and cloning of CEs identified by genome
755 sequence search. Restriction endonuclease sites used for directional cloning are underlined.

id.[1]	forward (5' -> 3')	reverse (5' -> 3')
CE09	GAGCATATGAGCCTGTTTGTTGATCGCATCAG	GCGAAGCTTTCATGCGTGAGCGTCCTCTTC
CE10	CGCATATGGATCTGATCATTTTTCTGC	CGGAAGCTTGTTGCAGATCAATATTTAC
CE11	ATACATATGCCGGTCCCCGAAAC	GACAAGCTTTCAGGCGTGTATTTCAATC
CE12	GCGCATATGGAACCACTTGAACTTGAGGAC	GCGAAGCTTCTATTCACTCAGGTAGCTGAGCACAAC
CE16	AGGTCTAGATGGAGGCTACACCTCATG	GTGCTCGAGGTACGGGCAGTTGCCGCGATAATC
CE17	GCGCATATGCACACTCTGTTCAAACG	GCGAAGCTTTCAGTCCAAGGCCTGC
CE18	GCGCATATGAATAACCTTACGTTACTGCCC	GACAAGCTTCGCTTGCGCTTCCAGCC
CE19	GCGCATATGGTGGTCAATCTCTTTCAGC	GACAAGCTTCGCTTTTTCCCAACCGCGTG
CE20	GCGCATATGTCACCGCAC	GACAAGCTTCGCAAGTCCGAGGCGTTC

756 [1]Enzyme identifier as used in this study.
757

Supplementary Material

Identification of organic solvent tolerant carboxylic ester hydrolases for organic synthesis

Alexander Bollinger[1a], Rebecka Molitor[1a], Stephan Thies[1], Rainhard Koch[2], Cristina Coscolín[3], Manuel Ferrer[3] and Karl-Erich Jaeger[1,4*]

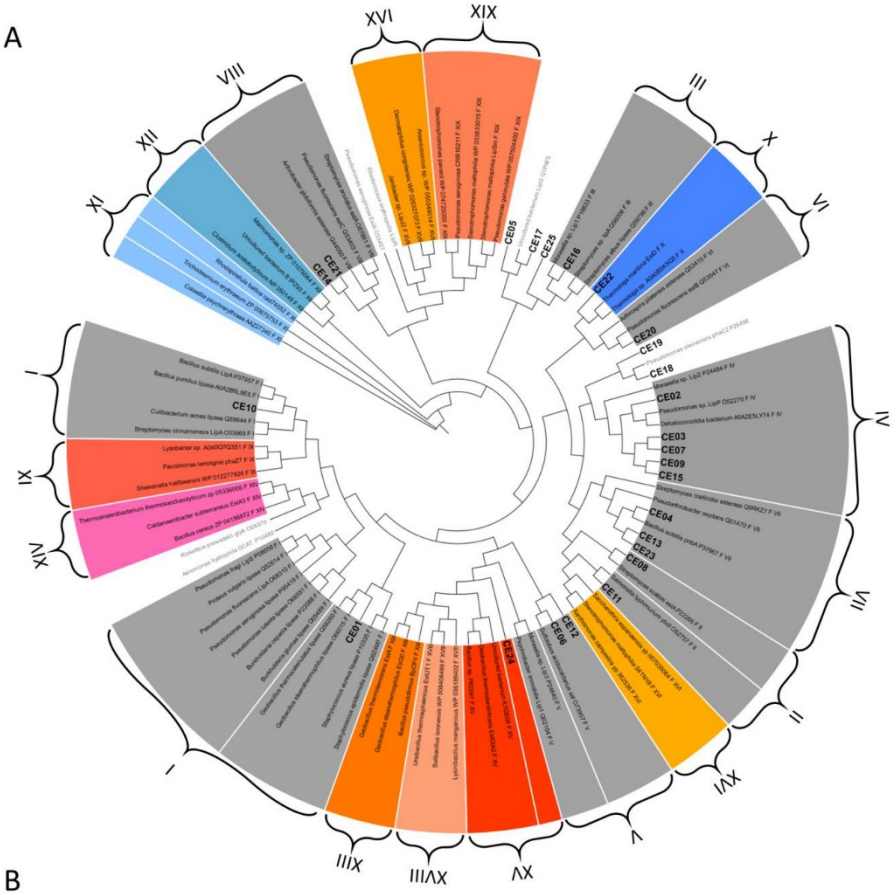
[1] Institute of Molecular Enzyme Technology, Heinrich Heine University Duesseldorf, Juelich, Germany

[2] Bayer AG, Leverkusen, Germany

[3] Institute of Catalysis, Consejo Superior de Investigaciones Científicas, Madrid, Spain

[4] Institute for Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Juelich, Germany

A



B

	CE01	CE02	CE03	CE04	CE05	CE06	CE07	CE08	CE09	CE10	CE11	CE12	CE13	CE14	CE15	CE16	CE17	CE18	CE19	CE20	CE21	CE22	CE23	CE24	CE25
CE01	100	12	12	17	13	20	13	15	17	17	13	12	18	14	16	17	11	14	15	11	10	9	19	10	18
CE02	12	100	26	14	12	12	24	10	27	11	14	11	12	6	29	11	16	14	16	20	6	15	12	12	16
CE03	12	26	100	12	15	8	46	13	31	13	13	11	10	5	35	9	10	16	16	21	9	13	11	14	13
CE04	17	14	12	100	13	19	14	16	12	12	21	15	31	17	18	13	11	13	17	17	15	13	31	11	15
CE05	13	12	15	13	100	10	15	14	13	12	10	12	12	13	15	11	26	13	16	11	12	17	16	13	14
CE06	20	12	8	19	10	100	13	15	12	11	16	14	16	21	11	15	9	6	6	13	15	11	18	11	13
CE07	13	24	46	14	15	13	100	11	29	14	16	12	9	4	29	11	10	17	16	21	6	15	13	15	14
CE08	15	10	13	16	14	15	11	100	8	10	23	17	18	14	9	9	11	7	10	7	14	13	18	4	12
CE09	17	27	31	12	13	12	29	8	100	18	18	7	11	8	43	14	10	15	15	23	8	14	13	14	11
CE10	17	11	13	12	12	11	14	10	18	100	16	14	15	7	16	17	9	12	18	16	6	21	14	18	17
CE11	13	14	13	21	10	16	16	23	18	16	100	29	14	12	22	22	4	10	13	17	11	9	16	13	11
CE12	12	11	11	15	12	14	12	17	7	14	29	100	12	7	9	19	14	8	16	14	12	9	15	5	18
CE13	18	12	10	31	12	16	9	18	11	15	14	12	100	14	14	19	13	11	13	11	15	14	41	11	16
CE14	14	6	5	17	13	21	4	14	8	7	12	7	14	100	8	13	10	12	9	3	38	9	14	6	16
CE15	16	29	35	18	15	11	29	9	43	16	22	9	14	8	100	13	11	16	18	24	11	15	14	13	13
CE16	17	11	9	13	11	15	11	9	14	17	22	19	19	13	13	100	7	12	18	10	12	13	16	13	30
CE17	11	16	10	11	26	9	10	11	10	9	4	14	13	10	11	7	100	16	20	13	7	10	14	5	18
CE18	14	14	16	13	13	6	17	7	15	12	10	8	11	12	16	12	16	100	13	17	11	14	9	14	11
CE19	15	16	16	17	16	6	16	10	15	18	13	16	13	9	18	18	20	13	100	17	11	20	14	17	14
CE20	11	20	21	17	11	13	21	7	23	16	17	14	11	3	24	10	13	17	17	100	4	15	15	15	16
CE21	10	6	9	15	12	15	6	14	8	6	11	12	15	38	11	12	7	11	11	4	100	10	13	7	12
CE22	9	15	13	13	17	11	15	13	14	21	9	9	14	9	15	13	10	14	20	15	10	100	12	15	12
CE23	19	12	11	31	16	18	13	18	13	14	16	15	41	14	14	16	14	9	14	15	13	12	100	12	16
CE24	10	12	14	11	13	11	15	4	14	18	13	5	11	6	13	13	5	14	17	15	7	15	12	100	8
CE25	18	16	13	15	14	13	14	12	11	17	11	18	16	16	13	30	18	11	14	16	12	12	16	8	100

16 **Figure S1** Classification of CEs used in this study into the currently 19 families of bacterial
17 lipolytic enzymes as shown by the unrooted phylogenetic tree (A). Enzymes used in this study
18 (CE01 to CE25) are printed in bold, families I to VIII are shown in grey, whereas families IX to XIX
19 are shown in different colors. The global sequence identity matrix of amino acid sequences of
20 the enzymes used in this study is given as a table (B). Maximum identity (100 %) is shown by
21 dark grey, high identity (>30 %) is shown by light grey background color.

22

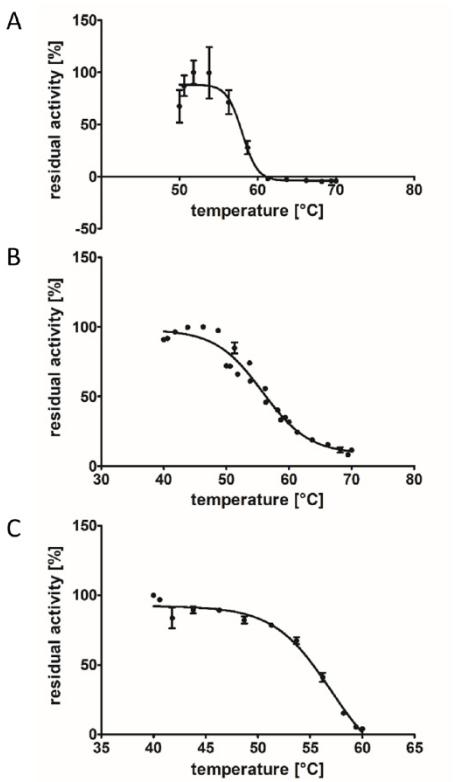


Figure S2 Thermal inactivation curves for CE01 (A), CE13 (B), and CE20 (C). Purified enzymes (about 1 U/ml in 100 mM potassium phosphate buffer pH 7.2) were incubated at various temperatures for 1 h using a PCR gradient cycler. Subsequently, enzyme activity was measured using 4-nitrophenyl butyrate as a substrate. The activity data was plotted relative to the highest activity and fitted nonlinear (Boltzmann sigmoidal). The enzymes half-inactivation temperature (T_{50}) is found at the temperature where the enzyme shows 50 % residual activity. The mean value and standard deviation of three separate reactions is shown.

	1	2	3	4	5	6	7	8	9	10	11	12
A	CE01	CE02	CE03	CE09	CE07	CE10	CE11	CE12	CE13	CE14	CE15	CE16
B	CE17	CE18	CE19	CE20	CE01	CE02	CE03	CE09	CE07	CE10	CE11	CE12
C	CE13	CE14	CE15	CE16	CE17	CE18	CE19	CE20	CE01	CE02	CE03	CE09
D	CE07	CE10	CE11	CE12	CE13	CE14	CE15	CE16	CE17	CE18	CE19	CE20
E	control	control	control									
F	CE21	CE14	CE22	CE23	CE24	CE26	CE25	CE04	CE06	CE05	CE08	CE21
G	CE14	CE22	CE23	CE24	CE26	CE25	CE04	CE06	CE05	CE08	CE21	CE14
H	CE22	CE23	CE24	CE26	CE25	CE04	CE06	CE05	CE08			

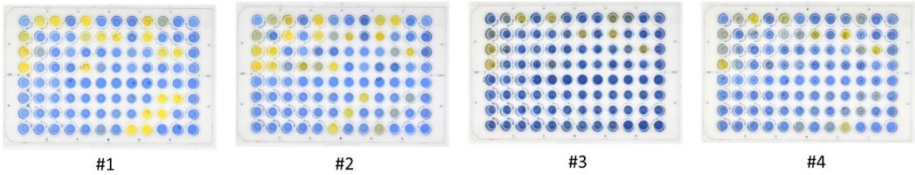
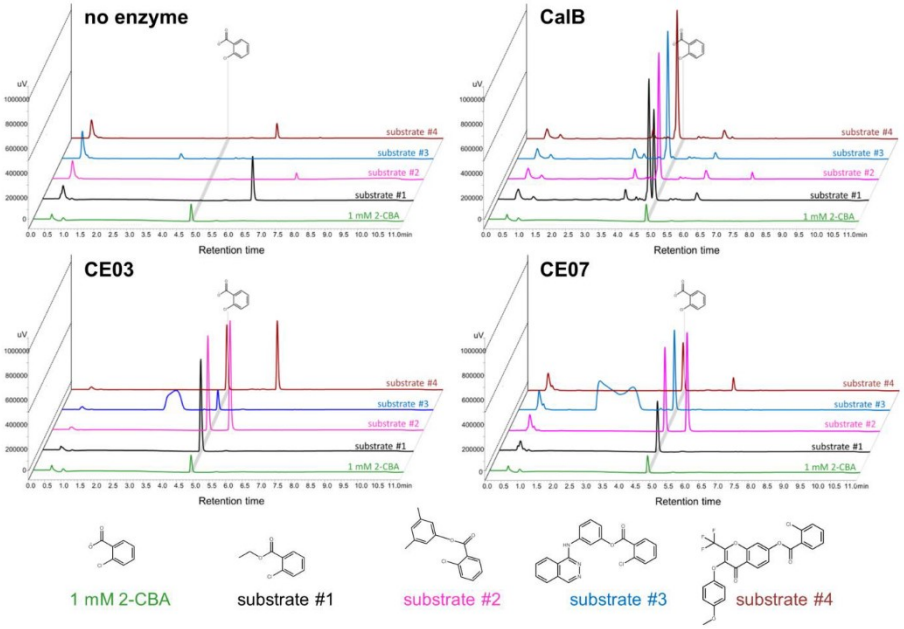


Figure S3 Layout of the assay plates and photographs after 18 h incubation at 30°C in 5 mM potassium phosphate buffer pH 7.2 containing 20 µg/ml nitrazine yellow, 30 % (v/v) dimethyl sulfoxide, 5 % (v/v) acetonitrile and 10 mM of the substrates #1 - #4. Substrates are #1: ethyl 2-chlorobenzoate. #2: 3,5-dimethylphenyl 2-chlorobenzoate. #3: 3-(quinazolin-4-ylamino)phenyl 2-chlorobenzoate. #4: 3-(4-methoxyphenoxy)-4-oxo-2-(trifluoromethyl)-4*H*-chromen-7-yl 2-chlorobenzoate.



	Ret. Time	Area	Height	calc. conc. (mM)	Accuracy[%]
2-CBA Standard 0 mM	-----	-----	-----	-----	-----
2-CBA Standard 0.01 mM	4.79	4496	1335	0.01	101.8
2-CBA Standard 0.1 mM	4.79	50195	14052	0.10	99.8
2-CBA Standard 1 mM	4.79	509201	138318	1.00	100
no enzyme control #1	4.79 ±0.00	9336 ±955	2674 ±276	0.02 ±0.00	
no enzyme control #2	4.80 ±0.00	8623 ±1348	1636 ±28	0.02 ±0.00	
no enzyme control #3	4.79 ±0.00	29606 ±232	6976 ±16	0.06 ±0.00	
no enzyme control #4	4.79 ±0.00	22000 ±195	5216 ±49	0.04 ±0.00	
CalB #1	4.80 ±0.00	2638222 ±207048	714066 ±54703	5.18 ±0.41	
CalB #2	-----	-----	-----	-----	
CalB #3	4.79 ±0.01	71727 ±410876	18989 ±101753	0.14 ±0.81	
CalB #4	-----	-----	-----	-----	
CE03 #1	4.79 ±0.00	2211501 ±571791	604228 ±154686	4.34 ±1.12	
CE03 #2	4.79 ±0.00	2458861 ±340285	680460 ±91858	4.82 ±0.67	
CE03 #3	4.79 ±0.00	511791 ±150794	135939 ±38089	1.01 ±0.29	
CE03 #4	4.78 ±0.00	1907004 ±62544	534513 ±15852	3.74 ±0.12	
CE07 #1	4.79 ±0.00	1832990 ±657330	495309 ±174573	3.59 ±1.29	
CE07 #2	4.79 ±0.01	1678834 ±1026990	460839 ±280794	3.30 ±2.01	
CE07 #3	4.79 ±0.00	2326591 ±18145	619761 ±16245	4.56 ±0.03	
CE07 #4	4.78 ±0.00	1009735 ±269201	282854 ±82030	1.98 ±058	
Average (N=44)	4.789 ±0.0055				
%RSD (N=44)	0.115796				
Maximum	4.801				
Minimum	4.776				

40

6

41 **Figure S4** Quantification of 2-Chlorobenzoic acid (2-CBA) released from different ester
42 compounds upon enzymatic hydrolysis by CE03, CE07, or CalB. Substrates were #1: ethyl 2-
43 chlorobenzoate. #2: 3,5-dimethylphenyl 2-chlorobenzoate. #3: 3-(quinazolin-4-ylamino)phenyl
44 2-chlorobenzoate. #4: 3-(4-methoxyphenoxy)-4-oxo-2-(trifluoromethyl)-4*H*-chromen-7-yl 2-
45 chlorobenzoate. Reaction conditions were 30°C, 18 h, 5 U of enzyme, 5 mM substrate, 30 %
46 DMSO, 70 mM potassium phosphate buffer pH 7.2.

47

48

49

50

51 **Table S1** Specification of the genome fragment carried by the recombinant pCR-XL-TOPO
52 plasmids of the genomic library clones used in this study with given start and end position of the
53 fragment on respective contig (genbank id), enzyme identifier (id.), and protein accession
54 number (acc.).

id.	acc.	vector	genbank id	start pos.	end pos.
CE04	WP_011588534.1	pCR-XL-TOPO	NC_008260.1	1424579	1428511
CE05	WP_011589376.1	pCR-XL-TOPO	NC_008260.1	2402770	2406029
CE06	WP_011589723.1	pCR-XL-TOPO	NC_008260.1	2771474	2777305
CE08	WP_011589970.1	pCR-XL-TOPO	NC_008260.1	3048294	3054618
CE21	WP_088273788.1	pCR-XL-TOPO	NBYK01000001.1	742906	734589
CE22	SEG59772.1	pCR-XL-TOPO	NBYK01000011.1	96693	107328
CE23	WP_088274564.1	pCR-XL-TOPO	NBYK01000003.1	149164	155945
CE24	WP_088275865.1	pCR-XL-TOPO	NBYK01000006.1	120599	125419
CE25	WP_088273867.1	pCR-XL-TOPO	NBYK01000001.1	836589	847776
CE26	n.d.	pCR-XL-TOPO	NBYK01000001.1	282473	286927

55

56

57

58 **Table S2** Substrate promiscuity data of CEs used in this study assessed with 96 different ester
59 substrates. Activity data are given in U/g of wet cell weight.

60 The dimensions of Table S2 did not allow it to be included here; Table S2 can be found as
61 separate file in excel format.

62

63

Supporting method

The source code used to plot the heatmap in the language R is given below:

```
heatmap.2(
      matrix,
      Rowv=TRUE,
      Colv=NULL,
      col=rev(c("#FFFFDD","#FFFFDD","#FFFFDD","#FFFFDD","#FFFFDD","#FFFFDD","#
      EDF8B1","#C7E9B4","#7FCDBB","#41B6C4", "#1D91C0","#225EA8","#253494")),
      scale="none",
      margins=c(6,6),
      symm=FALSE,
      vline=NULL,
      hline=NULL,
      trace="none",
      density.info="none",
      srtCol=35,
      cexCol=0.90,
      symkey=FALSE,
      symbreaks=FALSE,
      colsep=c(1:50),
      rowsep=c(1:50),
      sepcolor="black",
      sepwidth=c(0.01,0.01),
      na.color="#081D58",
      key.xlab="activity",
      key.xtickfun=function()
      {cex<-par("cex")*par("cex.axis")
              side<-1
              line<-0
              col<-par("col.axis")
              font<-par("font.axis")
              mtext("low", side=side, at=0, adj=0,
              line=line, cex=cex, col=col, font=font)
              mtext("high", side=side, at=1, adj=1,
              line=line, cex=cex, col=col, font=font)
              return(list(labels=FALSE, tick=FALSE))})
```

2.6 Agar plate-based screening methods for the identification of marine *Pseudomonas* sp. capable of polyester hydrolysis

Rebecka Molitor, Alexander Bollinger, Sonja Kubicki, Anita Loeschcke, Karl-Erich Jaeger, and Stephan Thies

Microbial Biotechnology (2020) 13, 274–284

Available online:

https://doi.org/10.1111/1751-7915.13418

Own contribution:

Participation in conceptualization of the study. Investigation: initial tests of the enzymatic activity of *Pseudomonas aestusnigri* VGXO14 including polyester hydrolase activity. Supervision: oversight of the tests of different agar plate-based assays for polyester hydrolase activity. Writing: contribution to writing of the original draft and to reviewing and editing the final manuscript.

microbial biotechnology

Open Access

Brief Report

Agar plate-based screening methods for the identification of polyester hydrolysis by *Pseudomonas* species

Rebecka Molitor,[1] Alexander Bollinger,[1] Sonja Kubicki,[1] Anita Loeschcke,[1] Karl-Erich Jaeger[1,2] and Stephan Thies[1]*

[1] *Institute of Molecular Enzyme Technology, Heinrich-Heine-University Düsseldorf, Forschungszentrum Jülich, D-52425, Jülich, Germany.*

[2] *Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, D-52425, Jülich, Germany.*

Summary

Hydrolases acting on polyesters like cutin, polycaprolactone or polyethylene terephthalate (PET) are of interest for several biotechnological applications like waste treatment, biocatalysis and sustainable polymer modifications. Recent studies suggest that a large variety of such enzymes are still to be identified and explored in a variety of microorganisms, including bacteria of the genus *Pseudomonas*. For activity-based screening, methods have been established using agar plates which contain nanoparticles of polycaprolactone or PET prepared by solvent precipitation and evaporation. In this protocol article, we describe a straightforward agar plate-based method using emulsifiable artificial polyesters as substrates, namely Impranil® DLN and liquid polycaprolactone diol (PLD). Thereby, the currently quite narrow set of screening substrates is expanded. We also suggest optional pre-screening with short-chain and middle-chain-length triglycerides as substrates to identify enzymes with lipolytic activity to be further tested for polyesterase activity. We applied these assays to experimentally demonstrate polyesterase activity in bacteria from the *P. pertucinogena* lineage originating from contaminated soils and diverse marine habitats.

Received 23 January, 2019; revised 5 April, 2019; accepted 8 April, 2019.

*For correspondence. E-mail s.thies@fz-juelich.de; Tel. +49 2461 613790; Fax + 49 2461 612461.

Microbial Biotechnology (2019) **0**(0), 1–11
doi:10.1111/1751-7915.13418

Funding Information
The authors received funding from the European Union's Horizon 2020 research and innovation programme (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project 'INMARE' under grant agreement No. 634486. ST, SK and AL are financially supported by the ministry of Culture and Science of the German State of North Rhine-Westphalia within in the framework of the NRW Strategieprojekt BioSC (No. 313/323-400-00213).

Introduction

Recent attention of both the scientific community and the public was drawn to microorganisms with enzymatic capabilities to degrade the plastic polymer polyethylene terephthalate (PET) (Wei *et al.*, 2016; Wierckx *et al.*, 2018) that was assumed to be biologically inert for a long time (Moharir and Kumar, 2019). Probably the most prominent example is the β-proteobacterium *Ideonella sakaiensis* isolated from a plastic-polluted site (Yoshida *et al.*, 2016) which produces an enzyme named *Is*PE-Tase (Austin *et al.*, 2018; Gong *et al.*, 2018; Joo *et al.*, 2018) that was shown to be responsible for the biodegradation of PET. Crystallographic studies revealed that this enzyme shows a cutinase-like structure (Joo *et al.*, 2018) which is in line with other studies on enzymatic degradation of PET by enzymes that were initially described as cutinases (Nikolaivits *et al.*, 2018).

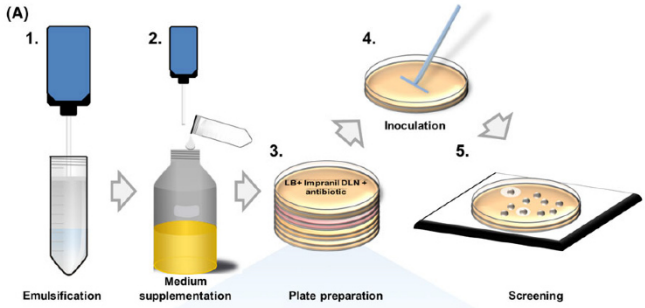
Cutinases are lipolytic enzymes and thus primarily active on carboxylic ester bonds (EC 3.1.1) but defined by activity on polyesters like the plant surface material cutin (Nikolaivits *et al.*, 2018) and, as a consequence, were assigned to a distinct enzyme subclass (EC 3.1.1.74). Cutinases are now spotlighted in the development of new strategies to deal with man-made plastic pollution: Most studies attempting to hydrolyze artificial polyesters are conducted applying such cutinase-like enzymes (Korpecka *et al.*, 2010; Austin *et al.*, 2018; Nikolaivits *et al.*, 2018). However, lipolytic enzymes clustering within other families (Arpigny and Jaeger, 1999), e.g. family VIII (β-lactamase like), were likewise associated with polyesterase activity very recently (Biundo *et al.*, 2017; Müller *et al.*, 2017; Hajighasemi *et al.*, 2018). Besides biodegradation of artificial polyesters like PET, cutinases/

polyesterases are discussed for different biotechnological applications, e.g. sustainable polymerization and polymer modification processes or biocatalytic transesterification and ester synthesis reactions (Nikolaivits *et al.*, 2018). Most polyesterases known today are secreted into the extracellular medium, potentially facilitating industrial production with either wild-type or suitable recombinant host strains.

Currently, high-throughput activity-based screening assays are a frequently applied method to identify novel biocatalysts within environmental isolates or metagenomic libraries (Popovic *et al.*, 2015; Martin *et al.*, 2016; Peña-García *et al.*, 2016; Thies *et al.*, 2016). These assays are of key importance for reducing the experimental workload to allow assigning of an activity of interest to an individual clone which can then be further characterized. To this end, agar plate-based activity assays are typically applied. Here, clear or coloured zones which are formed around the bacterial colonies indicate the production of a catalytically active enzyme. Suchlike approaches to identify organisms or clones with polyesterase activity currently mostly rely on clearance of media containing polycaprolactone (PCL) or PET nanoparticles prepared by solvent precipitation and evaporation techniques (Jarrett *et al.*, 1984; Nishida and Tokiwa, 1993; Wei *et al.*, 2014). Notably, these assays imply safety hazards and the production of organic solvent waste. In this protocol article, we describe water-emulsifiable polyesters (Fig. S1) as substrates for rapid and straightforward agar plate-based screening assays as an alternative or at least complementary strategy to identify polyesterase activity in bacterial clones, here exemplified by the identification of such enzymatic activities exhibited by yet unexplored *Pseudomonas* species. These assays generally allow for high-throughput identification of relevant clones, e.g. in metagenomic or genomic libraries (Fig. 1).

Step-by-step protocols for agar plate preparation and polyesterase activity screening

Polyesterases are lipolytic enzymes and are thus detected by non-specific esterase assays like an agar plate-based screening with the substrate tributyrin. The use of this universal substrate with a short-chain fatty acid triglyceride will also detect activities of esterases, true lipases,



(B)

Plate substrate	Composition	Target enzymes
Pre-screening		
Tributyrin	Triacylglycerides (C4)	Polyesterases, esterases, lipases, Acyl transferases, other hydrolases
Coconut oil	Triacylglycerides (~60% C12-C14)	Polyesterases, lipases
Polyesterase screening		
Polycaprolactone	Aliphatic polyester	Polyesterases
Polycaprolactone diol (M_n530)	Aliphatic polyester	Polyesterases
Impranil® DLN	Aliphatic polyester polyurethane	Polyesterases, polyurethanase

Fig. 1. Workflow for agar plate-based screening for polyesterase active clones.
A. Steps of plate preparation and screening: 1. Prepare an emulsion/suspension with the respective substrate (if necessary). 2. Combine substrate emulsion/suspension and molten agar-containing nutrient medium. 3. Pour the warm medium into suitable Petri dishes and let the agar solidify. Suitable supplements for induction of gene expression or selection may be included as well. 4. Plate bacteria either by transfer of single colonies using autoclaved toothpicks, 96 pin replicators or a robotic colony picker, or spread appropriate cell suspensions with glass beads or a Drigalski spatula. Incubate for at least 16 h at a temperature optimal for the applied organism. 5. Document the appearance of halos and/or fluorescence if applicable.
B. Overview on the described substrates (including the chain lengths of the dominant fatty acid for the triacylglycerides) and the enzymatic activities that can be identified with the respective screening plates.

phospholipases or even peptidases and acyl transferases. The use of triglycerides with long-chain fatty acids (FA) like olive oil instead is more selective for lipases because activity towards substrates with fatty acid chains > C10 is a characteristic of these enzymes (Kouker and Jaeger, 1987). However, cutinases have been categorized between esterases and true lipases because they are reported to have higher affinities for short-chain to middle-chain FA ester substrates with chain lengths up to C8 or C12 (Nikolaivits *et al.*, 2018); as a result of this substrate specificity, established lipase-specific screenings with long-chain plant oils like olive oil (Kouker and Jaeger, 1987) may miss lipolytic enzymes with additional polyesterase activity. The application of coconut oil that contains, in contrast, a large portion of C6-C14 FA esters (Sankararaman and Sferra, 2018), may bridge the gap between too universal and too lipase-specific substrates used for screening (tributyrin and olive oil respectively). Here, we suggest using the substrates tributyrin and coconut oil for an optional pre-screening to identify lipolytic activity because both substrates are inexpensive and easily available. As a second step, we describe the utilization of easy-to-emulsify polyesters which can serve as appropriate substrates, i.e. Impranil® DLN, an anionic aliphatic polyester polyurethane, and polycaprolactone diol PCD_{Mn530} as a polycaprolactone derivative of lower molecular weight. Impranil® DLN emulsion was already described as a substrate in agar plates for polyurethanase screening (Howard *et al.*, 2001). PCD_{Mn530} constitutes a viscous liquid which can be emulsified in liquid media in contrast to amorphous or crystalline solids like the commonly applied polycaprolactone.

General remarks

Media preparation. For the plate assays, we used autoclaved (121°C, 20 min) LB medium (Carl Roth, Karlsruhe, Germany), consisting of 10 g l^{-1} tryptone (peptone from casein), 5 g l^{-1} yeast extract and 10 g l^{-1} NaCl solubilized in deionized water supplemented with 15 g l^{-1} agar–agar (Carl Roth) as the growth medium because it proved suitable for growth of the selected *Pseudomonas* strains despite their partial marine origin and *Escherichia coli* that was included as negative control. However, other growth media or agar plates supplemented with antibiotics or expression inductors may also be tested, if required. As examples, polyesterase screening plates supplemented with the here introduced polyesterase substrates polycaprolactone diol and Impranil® DLN based on MME minimal medium (Vogel and Bonner, 1956) as well as artificial seawater medium (Passeri *et al.*, 1992) with regard to the sea-born strains of the *P. pertucinogena* lineage (Fig. S2). The agar was melted just before plate preparation or, alternatively, applied immediately after autoclaving. An Ultra Turrax T25 basic (IKA Labortechnik, Staufen, Germany), previously rinsed with 70% (v/v) ethanol, was applied with 16 000 rpm for both the preparation of substrate emulsions in sterile deionized water (if applicable) and their homogeneous emulsification into molten LB agar (cooled to a temperature of about 60–70°C). Emulsification using an ultrasonic emulsifier according to manufacturer's instructions is also possible. Here, it was in particular applied for smaller volumes (≤ 1 ml). The emulsion of the substrates in hot agar is recommended to maintain sterility of the plates.

Bacterial clones. In the presented examples, single colonies of *Pseudomonas* sp. and *E. coli* BL21(DE3), respectively, are transferred from a master plate to the indicator plates using sterile toothpicks. In general, it should also be applicable to directly plate (meta-)genomic libraries prepared in *E. coli* (Katzke *et al.*, 2017) using commercially available kits for TopoTA-cloning (Thermo Scientific, Waltham, MA, USA) or CopyControl™ Fosmid Library Production (epicentre, Madison, WI, USA), or mutagenesis libraries of specific genes in expression vectors prepared by standard molecular cloning methods. Agar plate-based screening assays are typically suitable for that application. As an example, tributyrin plates are an established tool for metagenomic library screenings (Peña-García *et al.*, 2016). Plating of dilutions of environmental samples to isolate species of interest might also be tested. However, it has to be kept in mind that the applied growth medium and incubation conditions will in general select for a subpopulation of the plated microbial strains. Hence, a good part of the natural diversity may be lost.

In the here presented example, plates were incubated at the optimal growth temperature of the used bacterial strains to allow colony formation and afterwards incubated at 4°C. At low temperatures, halo formation proceeds, whereas bacterial growth is slowed down. Thereby, the perception of activity is often facilitated without the danger of overgrowing (see below, section 'example'). Hence, prolonged incubation of screening plates at a lower temperature is a common strategy to detect poor activities in activity-based metagenomic library screenings (Popovic *et al.*, 2015, 2017; Thies *et al.*, 2016). Incubation temperature, incubation time to establish growth and the necessity for prolonged incubation at 4°C depend of course on the investigated organisms and enzymes.

Universal screening for lipolytic enzymes

Tributyrin assay.

(i) Prepare a 50% (v/v) tributyrin (Applichem, Darmstadt, Germany) emulsion in sterile distilled water and add

50 g l^{-1} gum arabic (Carl Roth) (Jaeger and Kovacic, 2014). Gum arabic powder is used as an emulsifying agent for the triglyceride. Homogenize the mixture for at least 1 min to yield a stable emulsion, e.g. using an Ultra Turrax (see general remarks).

Note: Add the gum arabic powder to the respective volume of water. Filling water into a tube or a bottle with a layer of the powder at the bottom should be avoided because it will result in a hard-to-dissolve clot of gum.

(ii) Add 30 ml of tributyrin emulsion per 1 l of molten LB agar (see general remarks) and mix thoroughly, e.g. using an Ultra Turrax (see general remarks).
(iii) Pour 25 ml medium portions into appropriate Petri dishes and allow the agar to solidify for at least 15 min.
(iv) Plate bacterial clones (see general remarks) and incubate at optimal growth temperature for the specific organism for at least 16 h.
(v) Positive clones are identified by a clearing halo after overnight growth or after prolonged (2–7 days) incubation at 4°C for clones expressing low amounts or less active enzymes. Photodocument agar plates (e.g. with a digital camera) and further proceed with selected clones, which were identified as lipolytically active, as appropriate.

Coconut oil assay.

(i) Melt coconut oil (Biozentrale Naturprodukte, Wittibreut – Ulbering, Germany) by incubation at 30–37°C. Pre-heat sterile distilled water to 60°C. Heating the water in advance should avoid a drop of temperature below 30°C during the preparation of the emulsion in the next step and therefore prevent a partial hardening of the coconut oil which will hamper successful emulsification.
(ii) Prepare a 50% (v/v) coconut oil emulsion in the preheated water containing 50 g l^{-1} gum arabic (Carl Roth) and 0.35 g l^{-1} rhodamine B (Sigma-Aldrich/Merck, Darmstadt, Germany). Homogenize the mixture for at least 1 min to yield a stable emulsion.

Note: Add the gum arabic powder to the respective volume of water. Filling water into a tube or a bottle with a layer of the powder at the bottom should be avoided because it will result in a hard-to-dissolve clot of gum.

(iii) Add 20 ml of coconut oil emulsion per 1 l of molten LB agar (see general remarks) and mix thoroughly, e.g. using an Ultra Turrax (see general remarks).
(iv) Pour 25 ml medium portions into appropriate Petri dishes and allow the agar to solidify for at least 15 min.
(v) Plate bacterial clones (see general remarks) and incubate at optimal growth temperature for the specific organism for at least 16 h.
(vi) Positive clones are identified by a fluorescent halo after overnight growth or after prolonged (2–7 days) incubation at 4°C for clones expressing low amounts or less active enzymes. Because of the low solubility of middle- and long-chain fatty acids in aqueous media, clearing halos are barely formed. Hence, esterase/lipase activity is detected by fluorescent complexes that are formed between the cationic rhodamine B and free fatty acids released from the substrate lead to yellow fluorescing colonies and/or halos around active colonies. These can be visualized by irradiation of the plate with UV light, e.g. at 254 nm, for example on a UV table for visualization of ethidium bromide-labelled DNA after gel electrophoresis. Photodocument agar plates (e.g. with a digital camera) and further proceed with selected clones, which were identified as lipolytically active, as appropriate.

Note: If the propagation of the colonies in further experiments is planned, apply UV radiation only for a short period of time (a few seconds) to prevent damaging effects of the UV light. Alternatively, blue light can be used for excitation, e.g. by NGFG15-FastGene Blue/Green LED Gel TransIlluminator XL (460–530 nm). However, background fluorescence of rhodamine B (excitation maximum 580 nm) increases (Fig. S3).

Note: Agar plates containing oils and rhodamine B constitute a frequently applied robust assay to detect lipase activities in different bacteria (Kouker and Jaeger, 1987; Jaeger and Kovacic, 2014). However, the production of fluorescent pigments may interfere with the rhodamine B fluorescence. The fluorescent siderophore pyoverdine leads to a bright fluorescence of many *Pseudomonas* strains under UV light exposure. Production of fluorescent siderophores should not be an issue for libraries within *E. coli* or the strains of the *P. pertucinogena* lineage investigated here because of the absence of respective gene clusters (Bollinger *et al.*, 2018). However, if pyoverdine producers are the strains of interest, supplementing additional iron to the medium decreases the siderophore production and therefore the autofluorescence (Fig. S4). The necessary amount is probably dependent on the physiology of the investigated strain and the applied growth medium. Concentrations described in protocols for appropriate mineral media for the respective strain may offer a suitable starting point.

Polyesterase screenings

Impranil® DLN assay.

(i) Add 4 ml of Impranil® DLN-SD emulsion (COVESTRO, Leverkusen, Germany) per 1 l of sterile molten LB agar (see general remarks) and mix thoroughly, e.g. using an Ultra Turrax (see general remarks).

(ii) Pour 25 ml medium portions into appropriate Petri dishes and allow the agar to solidify for at least 15 min.
(iii) Plate bacterial clones (see general remarks) and incubate at optimal growth temperature for the specific organism for at least 16 h.
(iv) Positive clones are identified by a clearing halo after overnight growth or after prolonged (2–7 days) incubation at 4°C for clones expressing low amounts or less active enzymes. Photodocument agar plates (e.g. with a digital camera) and further proceed with selected clones, which were identified as active, as appropriate.

Note: Impranil® DLN-SD emulsion contains isothiazolones as biocidal supplements to prevent spoilage. In the concentrations used here, we observed no impaired growth of the investigated bacteria.

Note: The anionic Impranil® DLN-SD may become difficult to emulsify into the agar in our experience when a salt-rich growth medium is applied.

Note: Impranil® DLN is also described as a useful substrate to uncover polyurethanase activities. Although this activity, like polyester hydrolysis, is not widespread, the verification of hits from an Impranil® DLN-based screening by determination of sequence homology or additional esterase activity assays is suggested (not described in this article).

Polycaprolactone diol (PCD_{Mn530}) assay.

(i) Prepare a 50% (v/v) PCD (average M_n 530 Da) emulsion: Mix the PCD_{Mn530} (Sigma-Aldrich/Merck) and 50 g l^{-1} gum arabic with sterile distilled water. Homogenize the mixture for at least 1 min to yield a stable emulsion, e.g. using an Ultra Turrax (see general remarks).

Note: Add the gum arabic powder to the respective volume of water. Filling water into a tube or a bottle with a layer of the powder at the bottom should be avoided because it will result in a hard-to-dissolve clot of gum.

(ii) Add 30 ml of PCD_{Mn530} emulsion per 1 l of LB agar and mix thoroughly, e.g. using an Ultra Turrax (see general remarks).
(iii) Pour 25 ml medium portions into appropriate Petri dishes and allow the agar to solidify for at least 15 min.
(iv) Plate bacterial clones (see general remarks) and incubate at optimal growth temperature for the specific organism for at least 16 h.
(v) Positive clones are identified by a distinct halo after overnight growth or after prolonged (2–7 days) incubation at 4°C for clones expressing low amounts or less active enzymes. Notably, the formation of clearing halos by enzymatic activity on PCD agar plates is often accompanied by a grainy accumulation of apparent hydrolysis or transesterification products at the edges of the halo, which is not observed on plates supplemented with tributyrin or Impranil® DLN. However, this even enhances the perceptibility of the halo (Fig. S4). Photodocument agar plates (e.g. with a digital camera) and further proceed with selected clones, which were identified as active on polyesters, as appropriate.

Assay with polycaprolactone (PCL) nanoparticle plates (common polyesterase assay, protocol derived from Jarrett et al. *(1984)).*

(i) Prepare a 5 g l^{-1} PCL solution by completely solving PCL (average M_n ~10 000 by GPC, density 1.146 g ml^{-1}, Sigma-Aldrich/Merck) in pre-heated acetone at 50°C under continuous stirring. Pre-heat an appropriate volume of sterile water likewise to 50°C for the next step.
(ii) Prepare a PCL particle suspension by slowly pouring the PCL solution drop by drop under continuous stirring into the water until a final acetone percentage of ca. 10–15% is reached.

Note: A turbid dispersion should be formed. Pour carefully, because too fast supplementation of PCL solution easily leads to the formation of tiny globular plastic particles instead of a homogenous suspension.

(iii) Add 100 ml of the warm PCL suspension per 1 l of LB agar (see general remarks) and mix thoroughly, e.g. using an Ultra Turrax (see general remarks).
(iv) Pour 25 ml medium into appropriate Petri dishes and allow the agar to solidify for at least 15 min.
(v) Plate bacterial clones (see general remarks) and incubate at optimal growth temperature for the specific organism for at least 16 h.
(vi) Positive clones are identified by a clearing halo on slightly turbid plates after overnight growth or after prolonged (2–7 days) incubation at 4°C for clones expressing low amounts or less active enzymes. Photodocument agar plates (e.g. with a digital camera) and further proceed with selected clones, which were identified as active on polyesters, as appropriate.

Example: Identification of polyesterase activity among members of the *P. pertucinogena* group

The recently established *Pseudomonas pertucinogena* lineage (Peix *et al.*, 2018) consists of several species barely explored until today. The group appears especially interesting for its distinct characteristics with respect to metabolism, genome size and, not least,

habitats with very specific conditions including cold, high-salt and chemically contaminated environments (Bollinger *et al.*, 2018). Remarkably for the predominantly terrestrial genus *Pseudomonas*, most of the species within this lineage were isolated from marine or saline habitats. Unlike other *Pseudomonas* species, which are well known for their versatile metabolism, bacteria of the *P. pertucinogena* lineage seem to have a more niche-adapted metabolism in common (Bollinger *et al.*, 2018). This is indicated by a comparably small genome and a limited spectrum of utilizable carbon sources. However, the current knowledge about the specific ecological and physiological properties of these species is very limited. An *in silico* search for cutinase homologous proteins uncovered a lipase of *Pseudomonas pelagia* (PpelaLip) which was recombinantly expressed in *E. coli* and proven to hydrolyze different artificial aromatic polyesters, among them poly(oxyethylene terephthalate) (Haernvall *et al.*, 2017a; Haernvall *et al.*, 2018). The strain itself exhibited likewise activity on the polyesters (Haernvall *et al.*, 2017a). The occurrence of genes encoding closely related proteins to PpelaLip appeared to be a common feature of this lineage of *Pseudomonas* sp. (Bollinger *et al.*, 2018). Therefore, we investigated one terrestrial and four species from different marine habitats differing in temperature, type of contamination and water depth for polyester hydrolyzing properties (Table 1). All strains were purchased from DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen, Braunschweig) and included *P. pelagia* CL-AP6^T as described producer of PpelaLip as a positive control. We further included *P. bauzanensis* BZ93^T isolated from contaminated soil and the marine species *P. litoralis* 2SM5^T, *P. aestusnigri* VGXO14^T and *P. oceani* KX 20^T. *P. putida* KT2440 (Belda *et al.*, 2016) was also included as a well-established and frequently applied member of the fluorescent Pseudomonads (Loeschcke and Thies, 2015) with a versatile metabolism, but without previously described polyesterase activity. In addition, we comprised *E. coli* BL21(DE3) (Studier and Moffatt, 1986) as a negative control because *E. coli* is applied as a standard host for metagenomic library screenings and recombinant esterase production, respectively, with negligible background activity.

All strains were streaked on LB agar plates and incubated at 30°C for 24 h. Distinct colonies of each strain were transferred to the indicator plates using sterile toothpicks and grown for 24 h at 30°C. Plates were photodocumented (Fig. 2), incubated further 24 h at 30°C and afterwards stored at 4°C for 4 days before the final photodocumentation (Fig. S5). All strains showed activity on the indicator plates, except *E. coli* BL21(DE3) and *P. putida* KT2440 which appeared as polyesterase-negative. The production of the fluorescent siderophore pyoverdine leads to bright fluorescence of the latter strain under UV light exposure. As halos were formed not only on Impranil® DLN, which may also result from other enzymatic activities (Fig. 1), but also on the two other polyester substrates, the tested strains of the *P. pertucinogena* lineage can be assumed to produce lipolytic and/or polyester hydrolyzing enzymes. Hydrolysis of the applied substrates by polyesterases was furthermore confirmed by clearing halos exhibited by *E. coli* with pEBP18_Cut (Troeschel *et al.*, 2012). This strain is able to express the cutinase gene from *Fusarium solani* f.sp. *pisi* which constitutes a well-characterized enzyme known for its polyesterase activity (Wei *et al.*, 2016; Wierckx *et al.*, 2018), from a shuttle vector applicable to metagenomic library screenings (Thies *et al.*, 2016) (Fig. S6).

In conclusion, polyesterase activity that was suggested by the previous identification of respective genes by sequence homology searches (Bollinger *et al.*, 2018) could be experimentally confirmed for these strains. The

Table 1. *Pseudomonas* strains analysed for polyesterase activity.

Species	DSMZ No.	Habitat[a]	Origin[a]	References[b]
P. aestusnigri VGXO14^T	103 065	Crude oil-contaminated intertidal sand samples	Spain 42°46′ 29.27″ N 9°7′27.08″ W	Sánchez *et al.* (2014); Gomila *et al.* (2017)
P. bauzanensis BZ93^T	22 558	Soil from an industrial site	Bozen, South Tyrol, Italy	Zhang *et al.* (2011)
P. litoralis 2SM5^T	26 168	Seawater of the Mediterranean coast	Spain 40° 27′ 24″ N 0° 31′ 36″ E	Pascual *et al.* (2012)
P. oceani KX 20^T	100 277	Deep-sea (1350 m)	Okinawa Trough, Pacific Ocean	Wang and Sun (2016); García-Valdés *et al.* (2018)
P. pelagia CL-AP6^T	25 163	Antarctic green algae co-culture	Antarctic Ocean	Hwang *et al.* (2009); Koh *et al.* (2013)
P. putida KT2440[c]	6125	Plasmid free derivative of *P. putida* mt-2, isolated from soil in Japan		Nakazawa (2002); Belda *et al.* (2016)

a. Environment from which the species was isolated (habitat) and geographical origin of the sample (origin) as stated in the type strain description.
b. References for original descriptions and, if applicable, genome announcements.
c. *P. putida* was included as an established representative of the fluorescent Pseudomonads.

comparison of the halo sizes as an indicator for the enzymatic activity revealed remarkable differences: (i) Closely related species exhibited very different strengths of activity. *P. litoralis* and *P. oceani* showed large hydrolysis halos already after one night of incubation, whereas the activity of *P. pelagia* became clearly visible only after growth for 48 h and several further days at 4°C (Fig. S5). (ii) The activity on polyester substrates appeared more prominent than on tributyrin and coconut oil while a polyesterase itself should be able to hydrolyze the triglyceride substrates likewise very well (A. Bollinger, unpublished). Both observations may be caused by variances in the specific activities of the different enzymes or attributable to differentially regulated polyesterase production and secretion by the different bacteria in reaction to the substrates. Further studies are necessary to assess whether the enzyme biochemistry or the bacterial physiology is the dominant factor behind the apparently massive differences in polyesterase activity.

Discussion

Lipolytic enzymes with activities on polyesters are already highly interesting for a variety of industrial applications and may become even more important in the future for plastic waste and microplastic removal (Wei and Zimmermann, 2017; Urbanek *et al.*, 2018). Currently, many scientific studies on the detection of bacterial polyesterase enzyme production rely on a first screening step utilizing the clearance of media from polymer nanoparticles, often polycaprolactone as a simple aliphatic polyester or PET as prominent industrially applied polyester (Jarrett *et al.*, 1984; Nishida and Tokiwa, 1993; Wei *et al.*, 2014). These procedures are associated with certain disadvantages as they imply solving of the solid polymer in hazardous organic solvents like acetone, dichloromethane or 1,1,1,3,3,3-hexafluoro-2-propanol and subsequent precipitation by careful addition to heated water or immediately to molten agar.

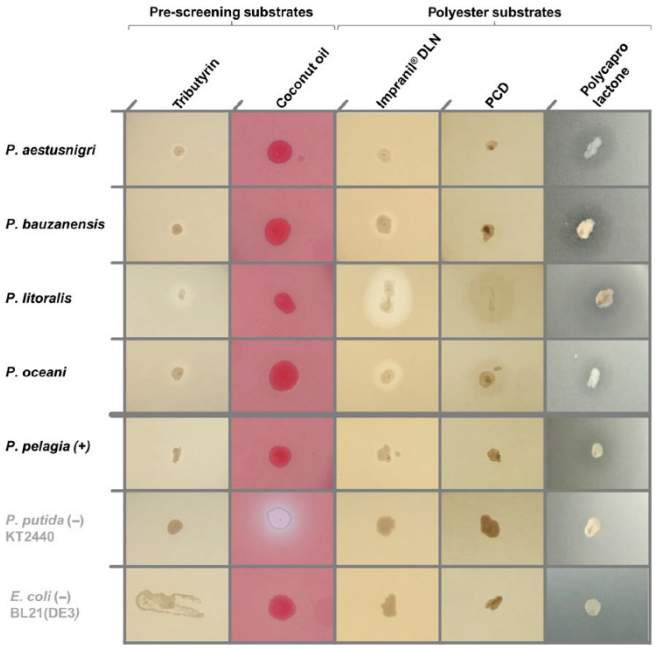


Fig. 2. Polyesterase activities exhibited by *Pseudomonas* species. The colonies were grown for 24 h at 30°C on LB agar plates supplemented with different substrates: Tributyrin (esterase activity); coconut oil + rhodamine B (mid-chain-length hydrolyzing esterase); Impranil® DLN (synthetic polyester polyurethane, polyesterase activity); PCD_{Mn530}, polycaprolactone diol (synthetic polyester, polyesterase activity); and polycaprolactone nanoparticles (current standard for polyesterase screens, polyesterase activity). *P. putida* as an example for a fluorescent Pseudomonad and *E. coli* as a negative control are indicated by grey letters. The white halo around *P. putida* relies on the fluorescence of the siderophore pyoverdine and does not indicate clearance of the substrate. All plates were photodocumented under white light, except coconut oil + rhodamine B-supplemented plates which were exposed to UV light (λ = 254 nm). Shown are exemplary colonies of a set of at least three colonies for each combination on independent plates. Halo formation of the depicted colony is representative for all replicates.

In both cases, a temperature just below the boiling point of the applied solvents is necessary. Hence, the boiling temperature may easily be exceeded and the risk of a sudden evaporation of hot solvents is immanent. Connected to this is the reassembly of larger globular polymer particles because of the sudden exposition of the insoluble polymer to water. Finally, the solvent has to be evaporated by heat or ultrasonication to avoid detrimental effects to the cells which are to be investigated. In addition, this step is necessary for the exchange of solvent shells around the plastic nanoparticles against water which makes them accessible for hydrolases. In our experience, the named procedures not only bear safety hazards but also require considerable handling practice to obtain reproducible results. The application of emulsifiable polyesters like Impranil® DLN or lower molecular weight polycaprolactone derivatives instead appeared in our hands to be a more rapid and straightforward procedure. While PCD–agar was to our knowledge not described to prepare screening plates before, Impranil® DLN-supplemented agar has previously been described and applied to identify and characterize polyurethanases, e.g. in biofilms that degrade coatings in military aircrafts (Howard *et al.*, 2001; Biffinger *et al.*, 2015, 2018; Hung *et al.*, 2016). In our experiences, this substrate is also perfectly suited to identify polyesterases. This observation is in line with studies using this substrate to assess cutinase activities in turbidometric experiments (Schmidt *et al.*, 2017). It is further supported by the fact that *F. solani* f.sp. *pisi* cutinase producing recombinant *E. coli* that are able to hydrolyze the polyester polyurethane (Fig. S6). However, Impranil® DLN screening may yield false positive hits constituting protease- or amidase-like enzymes rather than esterases. Hence, hits from these screenings should be verified using esterase activity assays based on the hydrolysis of, e.g., triglycerides or *p*-nitrophenol esters (Jaeger and Kovacic, 2014). Generally, application of inexpensive and easily available triglyceride substrates like tributyrin and coconut oil for screening may be highly useful to pre-select esterolytic organisms or clones in a library in advance to specific polyesterase assays (Fig. S7). PCD_{Mn530} is near twice as expensive as the applied polycaprolactone (source: Sigma-Aldrich); however, the small amounts necessary to prepare one litre medium for screening approaches render this substrate also affordable according to our experience. Impranil® DLN-SD emulsion is conventionally purchased as a bulk product for industrial coating applications. Hence, conditions to obtain small scaled product samples for the laboratory application have to be enquired on an individual basis. The applicability of a two-step strategy combining pre-selection and subsequent polyesterase activity assay was shown by the identification of novel types of polyesterases within a set of hydrolases from metagenomic libraries that were identified by their lipolytic activity in previous studies (Hajighasemi *et al.*, 2018).

The halo formation on the opaque white or yellowish Impranil® DLN agar (in dependence of the used medium) and the dark framed halos on PCD plates appeared to facilitate visual recognition of poor activities in comparison with semi-transparent nanoparticle plates. This straightforward readout might also be useful in applications using cutinases as reporter proteins in high-throughput approaches. Examples include transcriptional fusions confirming the successful transcription of target operons to identify promising expression strains (Domröse *et al.*, 2017), or as a model protein for studies on protein secretion, e.g. using signal peptide libraries (Knapp *et al.*, 2017). In addition, both substrates generally expand the set of polymers applicable to screenings. They may be used in combination with nanoparticle-based screenings to increase hit rates and to detect a broad variety of enzymatic activities in mixed samples as it can be assumed that different enzymes are differentially active on diverse unnatural substrates. Certainly, the aim of the screening was an important determinant for the selection of the substrate. Both assays described here apply aliphatic polyesters (Fig. S1), whereas many of the widely used polyesters like PET contain aromatic building blocks. However, previous studies showed that a large portion of enzymes is active on both types of substrates (Wei *et al.*, 2014; Danso *et al.*, 2018). This suggests that aliphatic polyesters might still serve as a useful substrate to pre-select candidates for further investigation even if aromatic polyester hydrolysis is the activity of interest. However, evolutive development of respective specificities towards a separation of both activities is discussed (Austin *et al.*, 2018); the presented assays are probably not suitable to indicate activity of such enzymes that are selective for aromatic polyesters.

In conclusion, the presented assays are suitable for high-throughput screening applications and may not completely replace but functionally complement the existing nanoparticle-based activity assays to exploit novel organisms and biocatalysts with polyesterase activity. For optimal results, these methods need to be interlinked with appropriate *in silico* strategies to exploit the available DNA sequence information. By using a hidden Markov Model-based search strategy to screen sequence data sets, Danso and co-authors showed that a surprisingly large variety of potential polyesterases is still to be discovered, in particular in bacteria which are currently not considered as a prime source for cutinases (Danso *et al.*, 2018). *Pseudomonas* species may constitute an example; in the context of polymer hydrolysis,

they appeared as a source for enzymes hydrolyzing polyurethane (Wilkes and Aristilde, 2017) for a long time, but some very recent reports by the Guebitz group indicated also polyesterase activity in Pseudomonads (Haernvall *et al.*, 2017a,b; Wallace *et al.*, 2017). The here reported confirmation of polyesterase activity of bacteria from the *P. pertucinogena* lineage, that was already suggested by sequence homology searches (Bollinger *et al.*, 2018), underlines the biotechnological potential of this group of bacteria. The predominantly marine Pseudomonas lineage, which includes psychrophilic, halophilic, as well as hydrocarbonoclastic, and heavy metal-tolerant species, may harbour many more intriguing biocatalysts with extraordinary properties.

Agar plate-based assays are a frequently applied tool for the activity-based screenings of metagenomic libraries, in particular for lipolytic enzymes (Popovic *et al.*, 2015, 2017; Peña-García *et al.*, 2016; Thies *et al.*, 2016), also with special emphasis on pollutant degrading enzymes (Ufarté *et al.*, 2015). The functionality of this assay with the typical host for metagenomic libraries, *E. coli*, expressing a cutinase encoding gene was indicated here (Fig. S6). In this light, the here presented assays may also prove useful to identify polyester-hydrolytic biocatalysts within metagenomic libraries containing DNA, e.g. from microplastic-polluted habitats. In the future, this may contribute to the exploitation of novel biocatalysts for biotechnological and environmental applications and shed light on natural plastic degradation processes in microbial communities.

Acknowledgements

The authors received funding from the European Union's Horizon 2020 research and innovation programme (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project 'INMARE' under grant agreement No. 634486. ST, SK and AL are financially supported by the ministry of Culture and Science of the German State of North Rhine-Westphalia within in the framework of the NRW Strategieprojekt BioSC (No. 313/323-400-00213).

Conflict of interest

None declared.

References

Arpigny, J.L., and Jaeger, K.-E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem J* **343(Pt 1):** 177–183.

Austin, H.P., Allen, M.D., Donohoe, B.S., Rorrer, N.A., Kearns, F.L., Silveira, R.L., *et al.* (2018) Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proc Natl Acad Sci USA* **115:** E4350–E4357.

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.

Biffinger, J.C., Barlow, D.E., Cockrell, A.L., Cusick, K.D., Hervey, W.J., Fitzgerald, L.A., *et al.* (2015) The applicability of Impranil®DLN for gauging the biodegradation of polyurethanes. *Polym Degrad Stab* **120:** 178–185.

Biffinger, J.C., Crookes-Goodson, W.J. and Barlow, D.E. (2018) Assignment of direct vs. indirect mechanisms used by fungi for polyurethane coating degradation. SERDP Final Report for SEED WP-2745. [WWW document] URL: https://www.serdp-estcp.org/content/download/47939/456696/file/WP-2745%20Final%20Report.pdf.

Biundo, A., Ribitsch, D., Steinkellner, G., Gruber, K., and Guebitz, G.M. (2017) Polyester hydrolysis is enhanced by a truncated esterase: less is more. *Biotechnol J* **12:** 8.

Bollinger, A., Thies, S., Katzke, N. and Jaeger, K.-E. (2018) The biotechnological potential of marine bacteria in the novel lineage of *Pseudomonas pertucinogena*. *Microb Biotechnol*. https://doi.org/10.1111/1751-7915.13288. [Epub ahead of print].

Danso, D., Schmeisser, C., Chow, J., Zimmermann, W., Wei, R., Leggewie, C., *et al.* (2018) New insights into the function and global distribution of polyethylene terephthalate (PET)-degrading bacteria and enzymes in marine and terrestrial metagenomes. *Appl Environ Microbiol* **84:** AEM.02773-17.

Domröse, A., Weihmann, R., Thies, S., Jaeger, K.-E., Drepper, T., and Loeschcke, A. (2017) Rapid generation of recombinant *Pseudomonas putida* secondary metabolite producers using yTREX. *Synth Syst Biotechnol* **2:** 310–319.

García-Valdés, E., Gomila, M., Mulet, M., and Lalucat, J. (2018) Draft genome sequence of *Pseudomonas oceani* dsm 100277^T, a Deep-sea bacterium. *Genome Announc* **6:** e00254-18.

Gomila, M., Mulet, M., Lalucat, J., and García-Valdés, E. (2017) Draft genome sequence of the marine bacterium *Pseudomonas aestusnigri* VGXO14^T. *Genome Announc* **5:** e00765-17.

Gong, J., Kong, T., Li, Y., Li, Q., Li, Z., Zhang, J., *et al.* (2018) Biodegradation of microplastic derived from poly (ethylene terephthalate) with bacterial whole-cell biocatalysts. *Polymers (Basel)* **10:** 1326.

Haernvall, K., Zitzenbacher, S., Wallig, K., Yamamoto, M., Schick, M.B., Ribitsch, D., and Guebitz, G.M. (2017a) Hydrolysis of ionic phthalic acid based polyesters by wastewater microorganisms and their enzymes. *Environ Sci Technol* **51:** 4596–4605.

Haernvall, K., Zitzenbacher, S., Yamamoto, M., Schick, M.B., Ribitsch, D., and Guebitz, G.M. (2017b) A new arylesterase from *Pseudomonas pseudoalcaligenes* can hydrolyze ionic phthalic polyesters. *J Biotechnol* **257:** 70–77.

Haernvall, K., Zitzenbacher, S., Biundo, A., Yamamoto, M., Schick, M.B., Ribitsch, D., and Guebitz, G.M. (2018) Enzymes as enhancers for the biodegradation of synthetic polymers in wastewater. *ChemBioChem* **19:** 317–325.

Hajighasemi, M., Tchigvintsev, A., Nocek, B., Flick, R., Popovic, A., Hai, T., *et al.* (2018) Screening and characterization of novel polyesterases from environmental metagenomes with high hydrolytic activity against synthetic polyesters. *Environ Sci Technol* **52:** 12388–12401.

Howard, G.T., Vicknair, J., and MacKie, R.I. (2001) Sensitive plate assay for screening and detection of bacterial polyurethanase activity. *Lett Appl Microbiol* **32:** 211–214.

Hung, C.-S., Zingarelli, S., Nadeau, L.J., Biffinger, J.C., Drake, C.A., Crouch, A.L., *et al.* (2016) Carbon catabolite repression and impranil polyurethane degradation in *Pseudomonas protegens* strain Pf-5. *Appl Environ Microbiol* **82:** 6080–6090.

Hwang, C.Y., Zhang, G.I., Kang, S.-H., Kim, H.J., and Cho, B.C. (2009) *Pseudomonas pelagia* sp. nov., isolated from a culture of the Antarctic green alga *Pyramimonas gelidicola*. *Int J Syst Evol Microbiol* **59:** 3019–3024.

Jaeger, K.-E., and Kovacic, F. (2014) Determination of lipolytic enzyme activities. *Methods Mol Biol* **1149:** 111–134.

Jarrett, P., Benedict, C.V., Bell, J.P., Cameron, J.A., and Huang, S.J. (1984) Mechanism of the Biodegradation of polycaprolactone. In *Polymers as Biomaterials*. Boston, MA: Springer US, pp. 181–192.

Joo, S., Cho, I.J., Seo, H., Son, H.F., Sagong, H.-Y., Shin, T.J., *et al.* (2018) Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nat Commun* **9:** 382.

Katzke, N., Knapp, A., Loeschcke, A., Drepper, T., and Jaeger, K.-E. (2017) Novel tools for the functional expression of metagenomic DNA. In *Metagenomics-Methods and Protocols*. Streit, W.R., and Daniel, R. (eds). New York: Springer, New York, pp. 159–196.

Knapp, A., Ripphahn, M., Volkenborn, K., Skoczinski, P., and Jaeger, K.-E. (2017) Activity-independent screening of secreted proteins using split GFP. *J Biotechnol* **258:** 110–116.

Koh, H.Y., Jung, W., Do, H., Lee, S.G., Lee, J.H., and Kim, H.J. (2013) Draft genome sequence of *Pseudomonas pelagia* CL-AP6, a psychrotolerant bacterium isolated from culture of Antarctic green alga *Pyramimonas gelidicola*. *Genome Announc* **1:** e00699-13.

Korpecka, J., Heumann, S., Billig, S., Zimmermann, W., Zinn, M., Ihssen, J., *et al.* (2010) Hydrolysis of cutin by PET-hydrolases. *Macromol Symp* **296:** 342–346.

Kouker, G., and Jaeger, K.-E. (1987) Specific and sensitive plate assay for bacterial lipases. *Appl Env Microbiol* **53:** 211–213.

Loeschcke, A., and Thies, S. (2015) *Pseudomonas putida*-a versatile host for the production of natural products. *Appl Microbiol Biotechnol* **99:** 6197–6214.

Martin, M., Vandermies, M., Joyeux, C., Martin, R., Barbeyron, T., Michel, G., and Vandenbol, M. (2016) Discovering novel enzymes by functional screening of plurigenomic libraries from alga-associated Flavobacteriia and Gammaproteobacteria. *Microbiol Res* **186–187:** 52–61.

Moharir, R.V., and Kumar, S. (2019) Challenges associated with plastic waste disposal and allied microbial routes for its effective degradation: a comprehensive review. *J Clean Prod* **208:** 65–76.

Müller, C.A., Perz, V., Provasnek, C., Quartinello, F., Guebitz, G.M., and Berg, G. (2017) Discovery of polyesterases from moss-associated microorganisms. *Appl Environ Microbiol* **83:** e02641-16.

Nakazawa, T. (2002) Travels of a *Pseudomonas*, from Japan around the world. *Environ Microbiol* **4:** 782–786.

Nikolaivits, E., Kanelli, M., Dimarogona, M., Topakas, E., Nikolaivits, E., Kanelli, M., *et al.* (2018) A middle-aged enzyme still in its prime: recent advances in the field of cutinases. *Catalysts* **8:** 612.

Nishida, H., and Tokiwa, Y. (1993) Distribution of poly(β-hydroxybutyrate) and poly(ε-caprolactone) aerobic degrading microorganisms in different environments. *J Environ Polym Degrad* **1:** 227–233.

Pascual, J., Lucena, T., Ruvira, M.A., Giordano, A., Gambacorta, A., Garay, E., *et al.* (2012) *Pseudomonas litoralis* sp. nov., isolated from Mediterranean seawater. *Int J Syst Evol Microbiol* **62:** 438–444.

Passeri, A., Schmidt, M., Haffner, T., Wray, V., Lang, S., and Wagner, F. (1992) Marine biosurfactants. IV. Production, characterization and biosynthesis of an anionic glucose lipid from the marine bacterial strain MM1. *Appl Microbiol Biotechnol* **37:** 281–286.

Peix, A., Ramírez-Bahena, M.-H., and Velázquez, E. (2018) The current status on the taxonomy of *Pseudomonas* revisited: an update. *Infect Genet Evol* **57:** 106–116.

Peña-García, C., Martínez-Martínez, M., Reyes-Duarte, D., and Ferrer, M. (2016) High throughput screening of esterases, lipases and phospholipases in mutant and metagenomic libraries: a review. *Comb Chem High Throughput Screen* **19:** 605–615.

Popovic, A., Tchigvintsev, A., Tran, H., Chernikova, T.N., Golyshina, O.V., Yakimov, M.M., *et al.* (2015) Metagenomics as a tool for enzyme discovery: hydrolytic enzymes from marine-related metagenomes. *Adv Exp Med Biol* **883:** 1–20.

Popovic, A., Hai, T., Tchigvintsev, A., Hajighasemi, M., Nocek, B., Khusnutdinova, A.N., *et al.* (2017) Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. *Sci Rep* **7:** 44103.

Sánchez, D., Mulet, M., Rodríguez, A.C., David, Z., Lalucat, J., and García-Valdés, E. (2014) *Pseudomonas aestusnigri* sp. nov., isolated from crude oil-contaminated intertidal sand samples after the Prestige oil spill. *Syst Appl Microbiol* **37:** 89–94.

Sankararaman, S., and Sferra, T.J. (2018) Are we going nuts on coconut oil? *Curr Nutr Rep* **7:** 107–115.

Schmidt, J., Wei, R., Oeser, T., Dedavid e Silva, L., Breite, D., Schulze, A. and Zimmermann, W. (2017) Degradation of polyester polyurethane by bacterial polyester hydrolases. *Polymers (Basel)* **9:** 65.

Studier, F.W., and Moffatt, B.A. (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J Mol Biol* **189:** 113–130.

Thies, S., Rausch, S.C., Kovacic, F., Schmidt-Thaler, A., Wilhelm, S., Rosenau, F., *et al.* (2016) Metagenomic discovery of novel enzymes and biosurfactants in a slaughterhouse biofilm microbial community. *Sci Rep* **6:** 27035.

Troeschel, S.C., Thies, S., Link, O., Real, C.I., Knops, K., Wilhelm, S., *et al.* (2012) Novel broad host range shuttle vectors for expression in *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*. *J Biotechnol* **161:** 71–79.

Ufarté, L., Laville, É., Duquesne, S., and Potocki-Veronese, G. (2015) Metagenomics for the discovery of pollutant degrading enzymes. *Biotechnol Adv* **33:** 1845–1854.

Urbanek, A.K., Rymowicz, W., and Mirończuk, A.M. (2018) Degradation of plastics and plastic-degrading bacteria in cold marine habitats. *Appl Microbiol Biotechnol* **102:** 7669–7678.

Vogel, H.J., and Bonner, D.M. (1956) Acetylornithinase of *Escherichia coli*: partial purification and some properties. *J Biol Chem* **218:** 97–106.

Wallace, P.W., Haernvall, K., Ribitsch, D., Zitzenbacher, S., Schittmayer, M., Steinkellner, G., *et al.* (2017) PpEst is a novel PBAT degrading polyesterase identified by proteomic screening of *Pseudomonas pseudoalcaligenes*. *Appl Microbiol Biotechnol* **101:** 2291–2303.

Wang, M., and Sun, L. (2016) *Pseudomonas oceani* sp. nov., isolated from deep seawater. *Int J Syst Evol Microbiol* **66:** 4250–4255.

Wei, R., and Zimmermann, W. (2017) Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: how far are we? *Microb Biotechnol* **10:** 1308–1322.

Wei, R., Oeser, T., Then, J., Kühn, N., Barth, M., Schmidt, J., and Zimmermann, W. (2014) Functional characterization and structural modeling of synthetic polyester-degrading hydrolases from *Thermomonospora curvata*. *AMB Express* **4:** 44.

Wei, R., Oeser, T., Schmidt, J., Meier, R., Barth, M., Then, J., and Zimmermann, W. (2016) Engineered bacterial polyester hydrolases efficiently degrade polyethylene terephthalate due to relieved product inhibition. *Biotechnol Bioeng* **113:** 1658–1665.

Wierckx, N., Narancic, T., Eberlein, C., Wei, R., Drzyzga, O., Magnin, A., *et al.* (2018) Plastic biodegradation: Challenges and opportunities. In *Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Biodegradation and Bioremediation. Handbook of Hydrocarbon and Lipid Microbiology*. Steffan, R. (ed). Cham: Springer International Publishing, pp. 1–29.

Wilkes, R.A., and Aristilde, L. (2017) Degradation and metabolism of synthetic plastics and associated products by *Pseudomonas* sp.: capabilities and challenges. *J Appl Microbiol* **123:** 582–593.

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., *et al.* (2016) A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **351:** 1196–1199.

Zhang, D.-C., Liu, H.-C., Zhou, Y.-G., Schinner, F., and Margesin, R. (2011) *Pseudomonas bauzanensis* sp. nov., isolated from soil. *Int J Syst Evol Microbiol* **61:** 2333–2337.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Molecular structures of the discussed substrates.
Fig. S2. Growth and polyesterase activities exhibited by *Pseudomonas* species on agar plates based on artificial sea medium and MME minimal medium.
Fig. S3. Coconut oil/rhodamine B agar plates exposed to different light conditions.
Fig. S4. Effect of the additional supplementation of Fe^{2+} on the autofluorescence of *P. putida* on coconut oil +rhodamine B agar plates.
Fig. S5. Polyesterase activities exhibited by *Pseudomonas* species after prolonged incubation at 4°C.
Fig. S6. Polyesterase exhibited by *E. coli* BL21(DE3) expressing the *F. solani* f.sp *pisi* cutinase gene.
Fig. S7. Schematic workflow for agar plate-based screening for polyesterase active clones within a (meta-)genomic library.

Supporting material

Agar plate-based screening methods for the identification of polyester hydrolysis by *Pseudomonas* species

Rebecka Molitor[1], Alexander Bollinger[1], Sonja Kubicki[1], Anita Loeschcke[1], Karl-Erich Jaeger[1,2], Stephan Thies[1]

[1]Institute of Molecular Enzyme Technology, Heinrich-Heine-University Düsseldorf, Forschungszentrum Jülich, D-52425 Jülich, Germany

[2]Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany

Address for correspondence: e-mail: s.thies@fz-juelich.de, Tel +49 2461 613790 Fax: +49 2461 612490

Content:

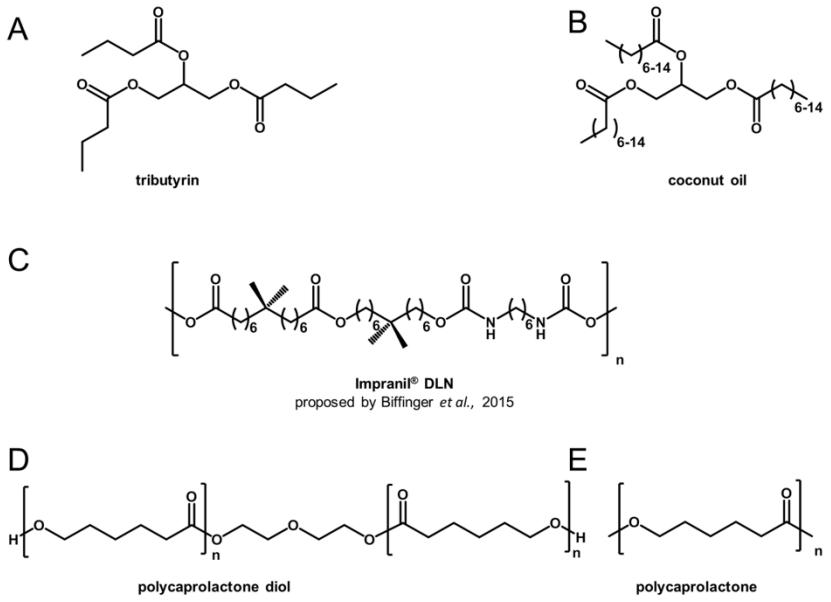


Fig. S1. **Molecular structures of the discussed substrates. A** Tributyrin (triacylglyceride with three butyric acid residues). **B** Coconut oil (triacylglyceride containing predominantly saturated fatty acid chains with 8-16 carbon atoms). **C** Proposed structure of Impranil® DLN according to Biffinger *et al.,* 2015. The actual structure is not publicly available. **D** Polycaprolactone diol (aliphatic polyester of C6 chains). **E** Polycaprolactone (aliphatic polyester of C6 chains).

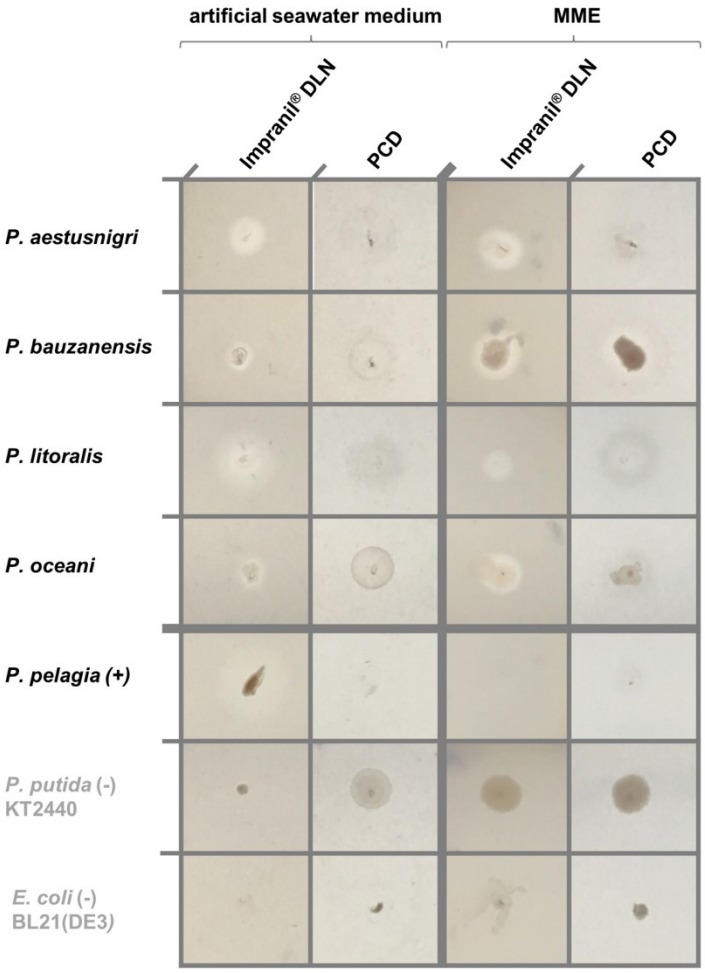


Fig. S2. Growth and polyesterase activities exhibited by *Pseudomonas* species and *E. coli* on agar plates based on artificial sea medium and MME minimal medium. The colonies were grown for 48 h at 30 °C on mineral medium agar plates supplemented with the two polyesterase screening substrates: Impranil® DLN (synthetic polyester polyurethane); PCD_{Mn530}, polycaprolactone diol (synthetic polyester, *P. putida* as an example for a fluorescent Pseudomonad and *E. coli* as negative control are indicated by grey letters. All plates were photo-documented under white light.

Artificial seawater as previously described (Passeri *et al.* 1992), consisting of 23 g/L NaCl, 0.75 g/L KCl, 1.47 g/L $CaCl_2$ x $2H_2O$, 5.08 g/L $MgCl_2$ x $6H_2O$, 6.16 g/L MgSO4 x 7H20, 0.89 g/L Na_2HPO_4x2 H_20, 6.0 g/L KNO_3, 0.03 g/L $FeSO_4$x7 H_2O, pH 7.0, was modified by addition of 15 g/L agar-agar and immediately before plate preparation 5 g/L sterile filtrated sodium pyruvate and the respective substrate emulsion.

MME agar was prepared by mixing autoclaved 2 fold concentrated modified MME medium as previously described (Vogel and Bonner, 1956) with an autoclaved 30 g/L agar-agar solution in water (ratio 1:1) and the addition of KNO_3 and sodium pyruvate as additional nitrogen and carbon sources, respectively, to enable growth of the species of the *P. pertucinogena* lineage. The final composition contained 2 g/L $MgSO_4$ $7H_2O$, 2 g/L citric acid x H_2O, 10 g/L K_2HPO_4 anhydrous, 3.5 g/L $NaNH_4HPO_4$ x $4H_2O$, 5 g/L KNO_3, 5 g/L sodium pyruvate, 15 g/L agar-agar, and the respective polyester.

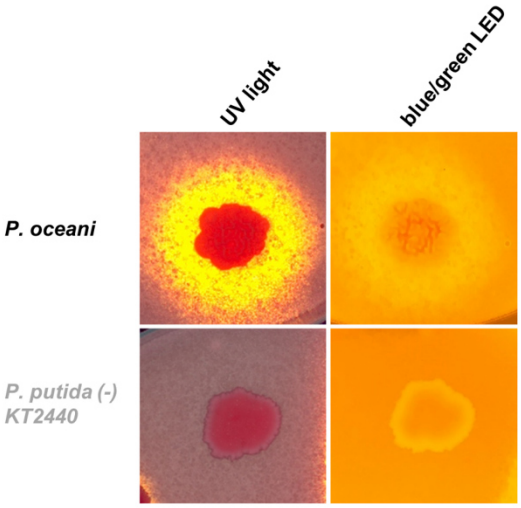


Fig. S3. Coconut oil/rhodamine B agar plates exposed to different light conditions. Colonies of *P. putida* and *P. oceani* on LB agar supplemented with coconut oil + rhodamine B, exposed to UV light (λ = 254 nm) and illuminated with a NGFG15 - FastGene Blue/Green LED Gel TransIlluminator XL light table (λ = 460 nm - 530 nm). Lipolytic activity is indicated by a fluorescent halo. The pictures were taken after storing the plates for 4 weeks at 4 °C succeeding incubation for 46 h at 30 °C for colony formation.

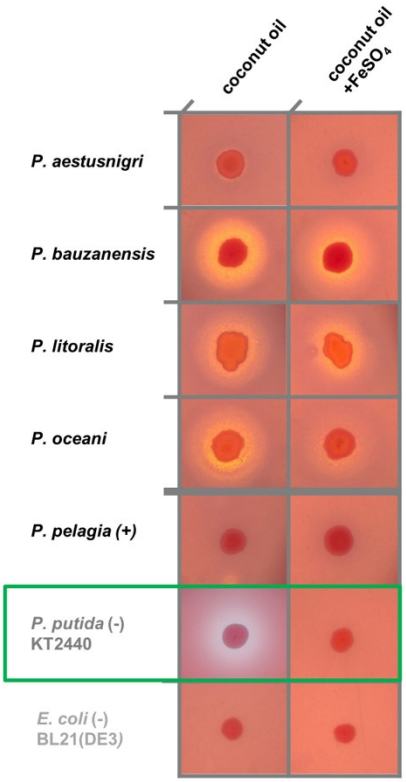


Fig. S4. Effect of the additional supplementation of Fe^{2+} on the autofluorescence of *P. putida* on coconut oil + rhodamine B agar plates. Colonies of *P. putida*, *E. coli*, and the strains of the *P. pertucinogena* lineage on LB agar supplemented with coconut oil + rhodamine B, exposed to UV light (λ = 254 nm) without and with additional supplementation of 12 mg/L $FeSO_4$. This represents the twofold concentration of the in-house established M9 minimal medium protocol. Lipolytic activity is indicated by a fluorescent halo. The pictures were taken after incubation for 46 h at 30 °C and subsequent storage for 96 h at 4 °C.

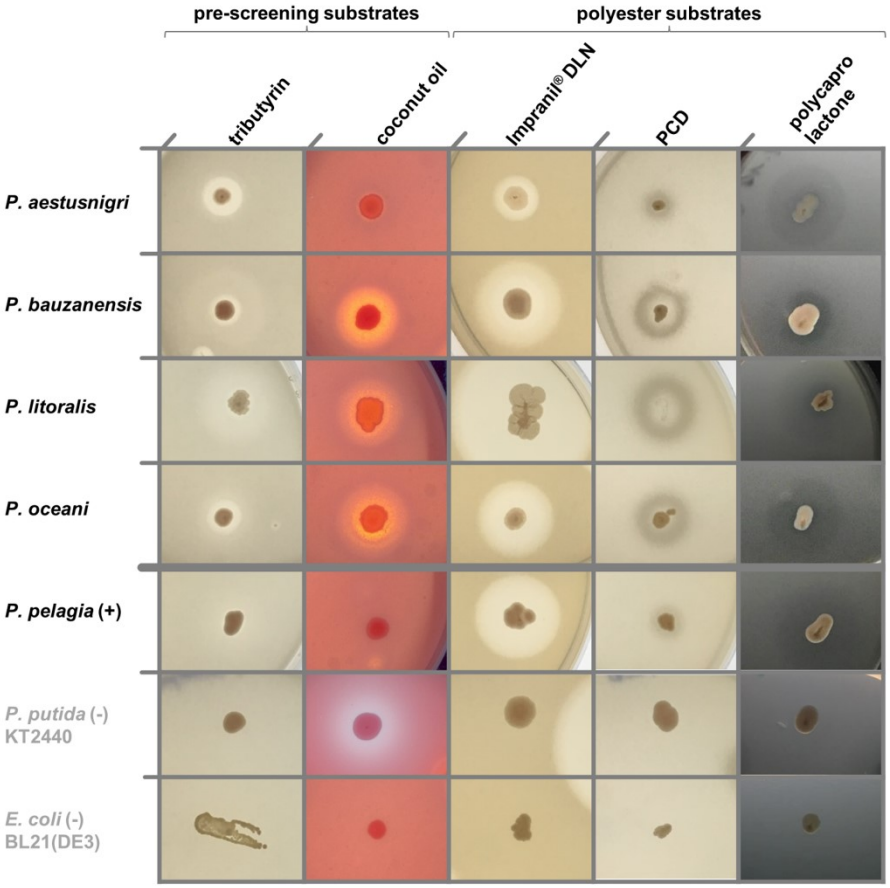


Fig. S5. Polyesterase activities produced by *Pseudomonas* species. Colonies were grown for 48 h at 30 °C and stored for 96 h at 4 °C on LB agar plates supplemented with different substrates: Tributyrin (esterase activity); coconut oil + rhodamine B (mid-chain length hydrolysing esterase); Impranil® DLN (synthetic polyester polyurethane, polyesterase activity); PCD_{Mn530}, polycaprolactone diol (synthetic polyester, polyesterase activity); polycaprolactone nanoparticles (current standard for polyesterase screens, polyesterase activity). *P. putida* as an example for a fluorescent Pseudomonad and *E. coli* as negative control are indicated by grey letters. The white halo around *P. putida* relies on the fluorescence of the siderophore pyoverdin and does not indicate clearance of the substrate. All plates were photodocumented under white light, except coconut oil + rhodamine B-supplemented plates which were exposed to UV light (λ = 254 nm). Shown are exemplary colonies of a set of at least 3 colonies for each combination on independent plates. Halo formation of the respectively depicted colony is representative for all replicates. Pictures of *P. litoralis* showing activity on Impranil® DLN and polycaprolactone were scaled down by 25% to depict the complete clearing halo.

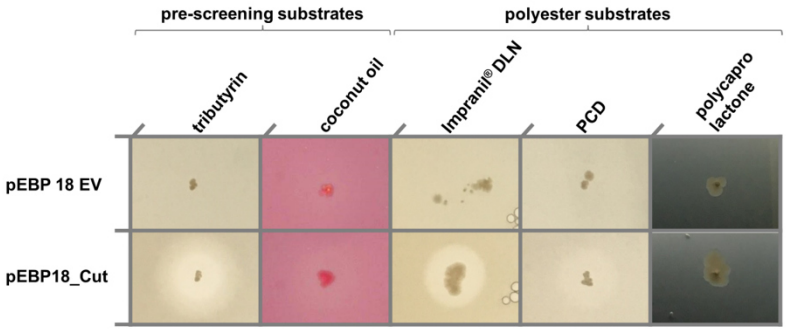


Fig. S6. Polyesterase exhibited by *E. coli* BL21(DE3) expressing the *F. solani* f.sp *pisi* cutinase gene. *E. coli* BL21(DE3) with the cutinase expression vector pEBP18_Cut and an empty vector as control, respectively, (Troeschel *et al.*, 2012) were grown on LB agar plates containing kanamycin (50 mg/L) and isopropyl-β-D-thiogalactopyranosid (IPTG) 95 mg/L) for the induction of target gene expression, supplemented with different substrates: Tributyrin (esterase activity); coconut oil + rhodamine B (mid-chain length hydrolysing esterase); Impranil® DLN (synthetic polyester polyurethane, polyesterase activity); PCD_{Mn530}, polycaprolactone diol (synthetic polyester, polyesterase activity); polycaprolactone nanoparticles (current standard for polyesterase screens, polyesterase activity).
All plates were photo-documented after incubation for 48 h at 37 °C under white light, except coconut oil + rhodamine B-supplemented plates which were exposed to UV light (λ = 254 nm).

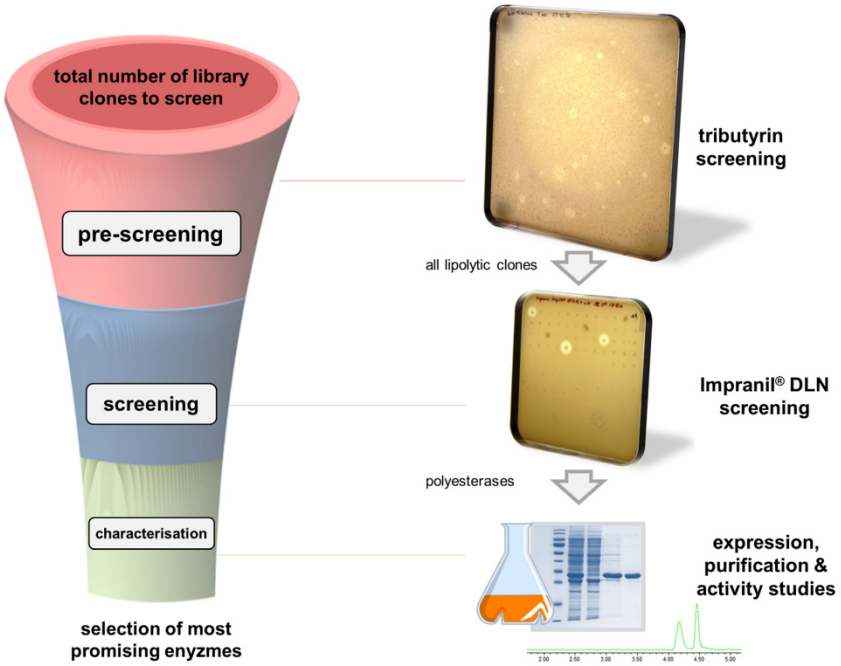


Fig. S7 Schematic workflow for agar plate-based screening for polyesterase active clones within a (meta-)genomic library to narrow down the clone number and reduce the workload for polyesterase characterisation**.** The whole library is pre-screened on inexpensive esterase substrates like tributyrin to reduce the number of clones for polyesterase activity screenings. All clones that produce lipoplytic enzymes are identified here. All of those are screened on polyester screening substrates like Impranil® DLN. Enzymes from clones with polyesterase activity can be investigated in further experiments concerning production, purification, and biochemical and structural characterisation to identify the most promising enzyme for the desired biotechnological application.

2.7 A novel polyester hydrolase from the marine bacterium *Pseudomonas aestusnigri* - structural and functional insights

Alexander Bollinger, Stephan Thies, Esther Knieps-Grünhagen, Christoph Gertzen, Stefanie Kobus, Astrid Höppner, Manuel Ferrer, Holger Gohlke, Sander H. J. Smits, and Karl-Erich Jaeger

Front. Microbiol. (2020) 11, 114

Available online:

https://doi.org/10.3389/fmicb.2020.00114

Own contribution:

Contribution in the conceptualization of the study. Investigation: identification, cloning, purification, and characterization of the polyester hydrolase PE-H. Sequence based analyses and structure guided mutational study of single and multiple amino acid substitutions. Quantification of PET hydrolysis. Writing of the original draft and contribution to reviewing and editing the final manuscript.

frontiers in Microbiology

ORIGINAL RESEARCH
published: 13 February 2020
doi: 10.3389/fmicb.2020.00114

A Novel Polyester Hydrolase From the Marine Bacterium *Pseudomonas aestusnigri* – Structural and Functional Insights

*Alexander Bollinger[1], Stephan Thies[1], Esther Knieps-Grünhagen[1], Christoph Gertzen[2,3], Stefanie Kobus[2], Astrid Höppner[2], Manuel Ferrer[4], Holger Gohlke[3,5], Sander H. J. Smits[2,6] and Karl-Erich Jaeger[1,7]**

[1] Institute of Molecular Enzyme Technology, Heinrich Heine University Düsseldorf, Jülich, Germany, [2] Center for Structural Studies, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, [3] Institute for Pharmaceutical and Medicinal Chemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, [4] Institute of Catalysis, Consejo Superior de Investigaciones Cientificas, Madrid, Spain, [5] Institute of Biological Information Processing (IBI-7: Structural Biochemistry), John von Neumann Institute for Computing and Jülich Supercomputing Centre, Forschungszentrum Jülich GmbH, Jülich, Germany, [6] Institute of Biochemistry, Heinrich Heine University Düsseldorf, Düsseldorf, Germany, [7] Institute of Bio- and Geosciences IBG-1: Biotechnology, Forschungszentrum Jülich GmbH, Jülich, Germany

OPEN ACCESS

Edited by:
Ren Wei,
University of Greifswald, Germany

Reviewed by:
Gert Weber,
Helmholtz-Gemeinschaft Deutscher Forschungszentren (HZ), Germany
Weidong Liu,
Tianjin Institute of Industrial Biotechnology (CAS), China

***Correspondence:**
Karl-Erich Jaeger
karl-erich.jaeger@fz-juelich.de

Specialty section:
This article was submitted to Microbiotechnology, Ecotoxicology and Bioremediation, a section of the journal Frontiers in Microbiology

Received: 07 November 2019
Accepted: 17 January 2020
Published: 13 February 2020

Citation:
Bollinger A, Thies S, Knieps-Grünhagen E, Gertzen C, Kobus S, Höppner A, Ferrer M, Gohlke H, Smits SHJ and Jaeger K-E (2020) A Novel Polyester Hydrolase From the Marine Bacterium Pseudomonas aestusnigri – Structural and Functional Insights. Front. Microbiol. 11:114. doi: 10.3389/fmicb.2020.00114

Biodegradation of synthetic polymers, in particular polyethylene terephthalate (PET), is of great importance, since environmental pollution with PET and other plastics has become a severe global problem. Here, we report on the polyester degrading ability of a novel carboxylic ester hydrolase identified in the genome of the marine hydrocarbonoclastic bacterium *Pseudomonas aestusnigri* VGXO14^T. The enzyme, designated PE-H, belongs to the type IIa family of PET hydrolytic enzymes as indicated by amino acid sequence homology. It was produced in *Escherichia coli*, purified and its crystal structure was solved at 1.09 Å resolution representing the first structure of a type IIa PET hydrolytic enzyme. The structure shows a typical α/β-hydrolase fold and high structural homology to known polyester hydrolases. PET hydrolysis was detected at 30°C with amorphous PET film (PETa), but not with PET film from a commercial PET bottle (PETb). A rational mutagenesis study to improve the PET degrading potential of PE-H yielded variant PE-H (Y250S) which showed improved activity, ultimately also allowing the hydrolysis of PETb. The crystal structure of this variant solved at 1.35 Å resolution allowed to rationalize the improvement of enzymatic activity. A PET oligomer binding model was proposed by molecular docking computations. Our results indicate a significant potential of the marine bacterium *P. aestusnigri* for PET degradation.

Keywords: ***Pseudomonas aestusnigri*, marine bacteria, polyester degradation, polyethylene terephthalate, PET, crystal structure**

INTRODUCTION

The modern society depends on the production and use of synthetic polymers which are uniformly present in both, basic and high-tech applications. The low production costs for plastic made from fossil feedstock and the high durability of the material are major advantages but have become a burden for the global ecosystem. Plastic waste is produced at a much faster rate than it is recycled (Moharir and Kumar, 2019); hence, it is disposed in landfills at a large extend where it can take

centuries to degrade completly. Morever, small plastic particles, so-called microplastics, usually evade municipal waste collection, being directly released into waste water and spread easily around the globe (Rochman, 2018). Hence, plastic waste accumulates in the environment to a large extent with very slow biodegradation to occur (Lebreton et al., 2018).

Where most plastics are inert polyolefins, consisting of carbon-carbon bonds, heteroatomic plastics like polyamides, polyurethanes and polyesters provide chemical groups of higher reactivity and thus are more easily degraded biologically (Wei and Zimmermann, 2017). The most abundant polyester plastic, present for example in packaging waste, is polyethylene terephthalate (PET) (Adrados et al., 2012). In the European Union, five million metric tons of this polyester were used for the production of plastics in 2017 (PlasticsEurope, 2018).

Enzymes catalyzing the degradation of polyesters such as polycaprolactone (PCL), polylactic acid (PLA), or PET are found within the class of carboxylic ester hydrolases (E.C. 3.1.1), most of them are classified as cutinases (E.C. 3.1.1.74), enzymes naturally adapted to act on polymeric ester substrates, e.g., the wax cuticle of plants (Nikolaivits et al., 2018). Studies on the identification of polyester degrading enzymes have shown that only a small fraction of carboxylic ester hydrolases is able to degrade synthetic polyester substrates. In a comprehensive metagenomics screening study, a subset of 23 carboxylesterases were tested for PLA hydrolysis, yielding seven positive hits (Popovic et al., 2017). More recently, by screening of over 200 different purified hydrolases for activity on synthetic polyesters, 36 positive enzymes were identified of which 10 enzymes showed high activity on multiple polyester substrates (Hajighasemi et al., 2018). For PET degradation, comprehensive activity-based screening studies are missing. However, a bioinformatics study using a hidden Markov model succeeded to identify PET hydrolase genes using the UniProtKB database and more than 100 metagenome datasets, many of which originated from marine sources (Danso et al., 2018). The reported frequency of PET hydrolases was, dependent on the origin of the metagenomic sample, between 0.0001 and 1.5 hits per megabases of sequence, with the highest hit rate in a metagenome obtained from an oil polluted environment (Danso et al., 2018). In contrast to the marine origin of many predicted PET hydrolases, most of the PET degrading enzymes studied so far originate from terrestrial sources, with Cut190 from *Saccharomonospora viridis* (Kawai et al., 2014), Tha_Cut1 from *Thermobifida alba* (Ribitsch et al., 2012), Thc_Cut1 and Thc_Cut2 from *T. cellulosilytica* (Ribitsch et al., 2012), Tfu_0883, Tfu_0882 and TfCut2 from *T. fusca* (Chen et al., 2010; Roth et al., 2014) and LCC identified from a leaf-branch compost metagenome (Sulaiman et al., 2012) for example. All these enzymes share a characteristic thermostability which is in line with the lifestyle of their thermophilic host organism or the respective environment. This feature is beneficial for the degradation of solid PET, since the glass transition temperature of PET, i.e., the temperature where the polymer becomes flexible and thus more accessible to enzymatic degradation, is about 75°C (Wei et al., 2019a). However, biodegradation of PET can also occur at lower temperatures, as demonstrated with PETase from *Ideonella sakaiensis*, the first such enzyme originating from a mesophilic organism (Yoshida et al., 2016). This enzyme outcompetes other cutinases for the hydrolysis of crystalline PET at 30°C as demonstrated in a comparative study with the leaf-branch compost cutinase (LCC) and a cutinase from the *Thermobifida* group (Yoshida et al., 2016). The elucidation of PETase three-dimensional structures by different groups (Han et al., 2017; Austin et al., 2018; Chen et al., 2018; Joo et al., 2018; Liu B. et al., 2018; Liu C. et al., 2018; Palm et al., 2019), lead to a proposal for the degradation mechanism and structural hallmarks responsible for superior activity as reviewed by Taniguchi et al. (2019). Structural features compared to other cutinase structures include an additional disulfide bond for improved stability at the position of the active site histidine, allowing for increased flexibility of the adjacent extended loop region (Fecker et al., 2018), thus facilitating the interaction with the polymer (Joo et al., 2018). Based on sequence and structural information, Joo et al. defined different types of PET degrading enzymes: Most known cutinases were assigned to type I, and enzymes possessing an additional disulfide bond and an extended loop region were assigned to type II, which was subdivided into types IIa and IIb based on the amino acid composition of respective regions (Joo et al., 2018). Crystal structures are published for several representatives of type I (Roth et al., 2014; Sulaiman et al., 2014; Miyakawa et al., 2015; Ribitsch et al., 2017), for type IIb just one enzyme with solved crystal structures exists (Han et al., 2017; Austin et al., 2018; Fecker et al., 2018; Joo et al., 2018; Liu B. et al., 2018; Liu C. et al., 2018; Palm et al., 2019), and, to the best of our knowledge, no crystal structure is known for a type IIa enzyme.

Recently, we observed that the marine bacterium *Pseudomonas aestusnigri* showed polyester degrading activity (Molitor et al., 2020). In this study, we identified the polyester hydrolase named PE-H which belongs to type IIa of PET hydrolases and demonstrated its activity toward PET as a substrate. We also report on the first crystal structure of a type IIa PET hydrolase. By a site-directed mutagenesis approach, inspired by known PETase structural features, we obtained a PE-H variant with significantly improved activity. The crystal structure of this variant was solved as well allowing us to rationalize our biochemical findings.

MATERIALS AND METHODS

Enzyme Production and Purification

Construction of the Expression Plasmid

The gene coding for the enzyme PE-H (locus tag B7O88_RS11490 of NCBI Reference Sequence NZ_NBYK01000007.1) was cloned into expression vector pET-22b(+) (Novagen) in frame with the vector-encoded hexa histidine tag utilizing *Xba*I and *Xho*I endonuclease restriction sites (Green and Sambrook, 2012). The gene was amplified by polymerase chain reaction (PCR) with Phusion High-Fidelity DNA Polymerase (Thermo Scientific) following the manufacturer's recommendations. Genomic DNA from *Pseudomonas aestusnigri* was isolated with the DNeasy, Blood and Tissue Kit (Qiagen GmbH) according to the manufacturer's

protocol and used as template with oligonucleotides PE-H_fw (AGGTCTAGATGGAGGCTACACCTCATG) and PE-H_rv (GTGCTCGAGGTACGGGCAGTTGCCGCGATAATC). The resulting recombinant plasmid pET22b_PE-H_{c6H} was used to transform chemical competent *E. coli* DH5α cells (Woodcock et al., 1989) for replication and *E. coli* BL21(DE3) cells (Hanahan, 1983) for T7 DNA polymerase driven expression (Studier and Moffatt, 1986).

Recombinant Protein Production

Protein production was carried out in Erlenmeyer flasks filled to 1/10 of the maximal volume with auto induction media (20 g/l tryptone from casein, 5 g/l NaCl, 5 g/l yeast extract, 6 g/l Na_2HPO_4, 3 g/l KH_2PO_4, 0.6% glycerol, 0.2% lactose, 0.05% glucose) (Studier, 2005) modified as described in[1] supplemented with 100 μg/ml ampicillin, for 24 h at 30°C with shaking (160 rpm). The culture was inoculated to an optical density of 0.05 (λ = 580 nm) from a culture grown overnight in LB media (Luria/Miller, Carl Roth GmbH & Co. KG) supplemented with 0.5% glucose and 100 μg/ml ampicillin. After the designated production time cells were collected by centrifugation for 30 min at 6,000 × *g*, 4°C, the supernatant was discarded, and cell pellets were stored at −20°C or used subsequently.

Protein Purification

Purification of PE-H was performed by immobilized metal ion affinity chromatography (IMAC) and size exclusion chromatography (SEC). Cell pellets were resuspended in lysis buffer (20 mM Na_2HPO_4 pH 7.4, 500 mM NaCl, 10 mM imidazole) at 10% (w/v) and disrupted using a high-pressure homogenizer (EmulsiFlex-C5, AVESTIN Europe, GmbH) with three passages at about 8,000 psi. Cell debris and insoluble aggregates were removed by centrifugation (30 min, 4°C, 36,000 × *g*), soluble proteins were mixed with about 5 ml Ni-NTA matrix (Ni-NTA Superflow, Qiagen GmbH) per liter of culture, the matrix was washed and equilibrated with lysis buffer prior to this, and incubated for 30 min at 4°C. The matrix was filled into a gravity flow column and washed with at least 10 column volumes (CV) of washing buffer (20 mM Na_2HPO_4 pH 7.4, 500 mM NaCl, 30 mM imidazole) before elution with 3 CV of elution buffer (20 mM Na_2HPO_4 pH 7.4, 500 mM NaCl, 500 mM imidazole). Eluted proteins were concentrated by centrifugal ultrafiltration (Vivaspin 20, 10,000 MWCO, Satorius AG) and desalted using 100 mM potassium phosphate buffer pH 7.4 with PD-10 desalting columns (GE Healthcare) according to the manufacturer's recommendation. Prior to protein crystallization studies, SEC was applied to further improve protein purity. An ÄKTA Explorer system (GE Healthcare) equipped with a HiLoad 16/60 Superdex 200 prep grade column (GE Healthcare) was used and proteins were eluted with 1.5 CV of 10 mM potassium phosphate buffer pH 7.4 as the mobile phase at a flow rate of 1 ml/min. Fractions of 5 ml volume were collected and tested individually for esterase activity with 4-nitrophenyl butyrate as the substrate. Protein containing fractions, as determined by absorption at λ = 280 nm, with esterase activity were pooled, concentrated by centrifugal ultrafiltration, and analyzed or stored at 4°C.

[1]https://openwetware.org/wiki/Lidstrom:Autoinduction_Media

Determination of Protein Concentration

Protein concentrations were determined using a micro-volume spectrophotometer (NanoDrop, Thermo Fisher Scientific) using protein specific molecular weight (32,308 Da) and extinction coefficient (48,610 $M^{-1}*cm^{-1}$) as calculated using the ProtParam web service (Gasteiger et al., 2005).

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS-PAGE)

SDS-PAGE analysis was carried out according to Laemmli (1970). Protein containing samples were mixed with sample buffer [50 mM Tris-HCl pH 6.8, 0.03% (w/v) bromphenol blue, 10% (v/v) glycerol, 4% (w/v) SDS, 2% (v/v) 2-mercaptoethanol], boiled for 5–10 min at 98°C, and applied to a 12% polyacrylamide gel using the Mini-PROTEAN system (Biorad GmbH) with Laemmli buffer [25 mM Tris-HCl pH 8.8, 192 mM glycine, 0.1% (w/v) SDS]. After separation for 15 min at 100 V and 40 min at 200 V, staining of the gel with Coomassie solution [10% (w/v) ammonium sulfate; 1.2% (v/v) phosphoric acid (85% aqueous solution), 0.1% Coomassie Brilliant Blue R250, 20% methanol] was applied (Blakesley and Boezi, 1977). Visual documentation was done using an Advanced Imager system (INTAS Science Imaging Instruments GmbH).

Biochemical Characterization

Qualitative Determination of Polyester Hydrolase Activity

Rapid qualitative determination of polyester hydrolase activity was carried out on agar plates containing Impranil DLN-SD (Covestro AG) as the substrate as described earlier (Molitor et al., 2020). 1.5% (w/v) agar-agar (Carl Roth GmbH & Co. KG) was added to either LB media (Luria/Miller, Carl Roth GmbH & Co. KG) for assessment of polyester hydrolase activity of bacteria, or to 100 mM potassium phosphate buffer pH 7.4, for assessment of polyester hydrolase activity of purified enzymes, and sterilized by autoclaving (20 min at 121°C). The molten agar media were allowed to cool down (about 60°C) before addition of heat labile compounds (e.g., antibiotics) and 1% (v/v) Impranil DLN-SD. The media was mixed with a magnetic stirrer, poured into Petri dishes, and dried for 15–30 min under a sterile laminar flow hood (Herasafe KS, Thermo Fisher Scientific). To assess activity, bacterial cells were transferred with a sterile toothpick to the plates and incubated at the respective optimal growth temperature for 1–3 days, or purified enzymes were dissolved in 100 mM potassium phosphate buffer pH 7.4 and directly applied to the plates.

Quantitative Determination of Esterase Activity

Esterase activity of purified enzymes was quantified using the substrates 4-nitrophenyl butyrate (*p*NPB) or hexanoate (*p*NPH) as described earlier (Nolasco-Soria et al., 2018). Briefly, 10 μl of enzyme solution was combined with 190 μl substrate solution (1 mM 4-nitrophenyl ester, 5% acetonitrile, 100 mM potassium

phosphate buffer pH 7.4) in a flat bottom 96-well microtiter plate and the reaction was followed at 30°C in a microplate reader (SpectraMax i3x, Molecular Devices, LLC) at λ = 410 nm. Initial reaction velocity corrected by a control reaction without enzyme was used to calculate the release of 4-nitrophenol per minute using formula [1].

$$\frac{V[min^{-1}] * v_{mtp} * F}{d[cm] * \varepsilon\ [mM^{-1} * cm^{-1}] * v_{enz} * c\ [mg * ml^{-1}]} = A\ [U * mg^{-1}] \quad (1)$$

V is the initial reaction velocity, v_{mtp} the volume in the well of the microtiter plate, F the dilution factor, d the path length, ε the extinction coefficient of 4-nitrophenol at pH 7.4, v_{enz} the volume of the enzyme sample, c the enzyme concentration and A the specific enzyme activity. One unit (U) was defined as the amount of enzyme needed to release 1 μmol of 4-nitrophenol per minute.

Determination of Protein Thermal Melting Point

Protein melting curves were measured by nano differential scanning fluorimetry (nanoDSF) using a Prometheus device (NanoTemper Technologies, Inc.), according to the manufacturer's recommendation. Briefly, purified enzyme (protein concentration 4–8 mg/ml) in 20 mM Tris pH 8 buffer was loaded into NanoTemper capillary tubes and applied to the Prometheus device for a melting scan at 10% excitation power, from 20 to 95°C at a heating rate of 1°C per minute.

Enzymatic Hydrolysis of BHET and PET and Quantification of Reaction Products

For hydrolysis of BHET and PET films, the enzymatic reaction was set up as described earlier (Yoshida et al., 2016) with minor modifications. The reaction mixture in a total volume of 300 μl was composed of 500 nM purified enzyme in 20 mM potassium phosphate buffer pH 7.4 with 20% (v/v) dimethyl sulfoxide (DMSO) and either 0.75 μl 400 mM BHET (95% purity, Sigma Aldrich) dissolved in DMSO or a circular piece of PET film (6 mm diameter). The PET pieces were produced from either amorphous PET film (0.25 mm thickness, Goodfellow Cambridge, Ltd.) or PET film derived from a commercial single use PET water bottle (trademark "Gut und Günstig," EDEKA) using a puncher, were washed with ethanol p.A., and were dried under a sterile laminar flow hood prior to use. The reaction mixtures were incubated for 24 h for BHET or 48 h for PET film at 30°C. BHET hydrolysis was stopped by removing the enzymes using ultrafiltration with centrifugal filters with a molecular weight cutoff (MWCO) of 10,000 Da (VWR International GmbH). PET film hydrolysis was stopped by heat inactivation of the enzymes for 20 min at 85°C and subsequent filtration with polyamide syringe filters of 0.2 μm pore size. The reaction filtrates were analyzed with an UPLC System (Acquity UPLC, Waters GmbH) equipped with an Acquity UPLC BEH C18 column (1.7 μm particle size) adapted from a published method (Yoshida et al., 2016). The mobile phase consisted of (A) 20 mM Na_2HPO_4 pH 2.5 (pH adjusted with H_2SO_4) and (B) methanol, the effluent was monitored at λ = 240 nm. The column was kept at constant temperature of 35°C and a flow rate of 0.208 ml/min. The program was 75% (A) and 25% (B) for 1.28 min, followed by a linear gradient to 100% (B) in 2 min, hold 100% (B) for 3 min, linear gradient from 100 to 25% (B) in 1 min and hold 25% (B) until minute 8.28 was reached. For terephthalic acid (TA) and BHET, commercially available standards were used to calculate amounts from calibration curves. For MHET no commercial standard was available. Therefore, a series of enzymatic reactions with BHET as substrate and MHET as major product was performed as described above and the percental distribution of the BHET and the MHET peak areas in combination with the known amount of the substrate BHET was used to calculate a calibration curve for enzymatic hydrolysis of MHET.

Site Directed Mutagenesis

To introduce single and multiple amino acid substitutions to PE-H, site directed mutagenesis was carried out. Therefore, QuikChange PCR was applied as described earlier (Edelheit et al., 2009) with mutagenic primer pairs (**Table 1**) and the recombinant plasmid pET22b_PE-H_{c6H} as a template. *E. coli* DH5α (Woodcock et al., 1989) cells were transformed with the recombinant plasmids by heat-shock (Hanahan, 1983) for vector replication, plasmid DNA was isolated with innuPREP Plasmid Mini Kit 2.0 (Analytic Jena AG), and mutations were verified by Sanger sequencing (eurofins genomics GmbH and LGC genomics GmbH). Cloning was simulated and sequence analysis was carried out using Clone Manager software (Sci-Ed Software).

TABLE 1 | Oligonucleotide sequences of primers used for site directed mutagenesis of PE-H.

Name	Sequence (5'->3')
S171A	GGCGTCATTGGCTGGGCGATGGGCGGTGGCGGC
	GCCGCCACCGCCCATCGCCCAGCCAATGACGCC
D217A	CTTTGCCTGTGAGTCGGCGGTGATCGCGCCGGTC
	GACCGGCGCGATCACCGCCGACTCACAGGCAAAG
H249A	CAATGGTGGCAGCGCGTACTGCGGTAATGGC
	GCCATTACCGCAGTACGCGCTGCCACCATTG
G254S	CACTACTGCGGTAATAGCGGCAGCATCTACAAC
	GTTGTAGATGCTGCCGCTATTACCGCAGTAGTG
S256N	GCGGTAATGGCGGCAACATCTACAACGATGTG
	CACATCGTTGTAGATGTTGCCGCCATTACCGC
I257S	GGTAATGGCGGCAGCAGCTACAACGATGTGCTG
	CAGCACATCGTTGTAGCTGCTGCCGCCATTACC
Y258N	GTAATGGCGGCAGCATCAACAACGATGTGCTGAGC
	GCTCAGCACATCGTTGTTGATGCTGCCGCCATTAC
N259Q	GGCGGCAGCATCTACCAGGATGTGCTGAGCCGG
	CCGGCTCAGCACATCCTGGTAGATGCTGCCGCC
ext.loop	GCAGCCACTACTGCGGTAATTCGGGCAACTCGAATCAGGATG
	CCGAACCGGCTCAGCACATCCTGATTCGAGTTGCCCGAATTAC
Y250S	AATGGTGGCAGCCACTCCTGCGGTAATGGCGGC
	GCCGCCATTACCGCAGGAGTGGCTGCCACCATT
Q294A	CACACTTCCGACTCTGCCATCTCCGATTATCGC
	GCGATAATCGGAGATGGCAGAGTCGGAAGTGTG
I219Y	GTGAGTCGGATGTGTACGCGCCGGTCCTCCAG
	CTGGAGGACCGGCGCGTACACATCCGACTCACA

Crystallization

Wild Type Enzyme PE-H

Several crystals were observed by using commercial kits from NeXtal (Qiagen, Hilden, Germany) and Molecular Dimensions (Suffolk, England) for initial screening. 0.1 μL homogenous protein PE-H (11 mg/ml, in 10 mM potassium phosphate buffer pH 7.4) was mixed with 0.1 μL reservoir solution and equilibrated against 40 μL reservoir solution in sitting drop MRC3 plates (Swissci) at 12°C. Crystals or needles appeared with this vapor diffusion method after a few days.

Variation of one of these conditions [0.1 M sodium acetate pH 4.5, 16% (w/v) PEG 3000] via grid screen (sitting drop, 1 μl + 1 μl over 300 μl reservoir at 12°C) resulted in well diffracting crystals with a maximum size of 120 × 30 × 20 μm after 1 week in an optimized condition composed of 0.1 M sodium acetate pH 4.5, 16% (w/v) PEG 3000, 0.036 mM LysoFos Choline14.

Enzyme Variant PE-H Y250S

Initial screening was performed as for the wild type (WT) PE-H (14 mg/ml of PE-H $Y_{250}S$). Within 3 weeks, rod shaped crystals appeared and reached their maximum size of 75 × 35 × 20 μm in 0.2 M lithium sulfate, 0.1 M sodium citrate pH 3.5 and 28% (v/v) PEG 400.

To cryoprotect the crystals, all drops were overlaid with 2 μl mineral oil before the crystals were harvested and flash frozen in liquid nitrogen.

Data Collection and Structure Determination

Data sets of a single crystals of the wild type enzyme PE-H were collected at the ID29 at ESRF (Grenoble, France) at 100K equipped with a Dectris Pilatus 6M detector. Data sets for enzyme variant PE-H Y250S were collected at P13 at DESY (Hamburg, Germany) at 100K equipped with a Dectris Pilatus 6M detector.

Data sets were processed with XDS (Kabsch, 2010). For PE-H, the XDS_ASCII.HKL-file together with the protein sequence was used as input in autorickshaw webservice[2] to obtain initial phases via molecular replacement. These output files were directly used in ARP/wARP webservice[3] for further model building and phase improvement. Subsequently, the model was further built and refined manually using COOT (Emsley et al., 2010) software followed by REFMAC5 from the ccp4 suite (Collaborative Computational Project, 1994). For PE-H Y250S, the refined wild type structure served as search model in molecular replacement. Further model building and refinement was performed as already described. Structures were deposited in the protein data bank[4] under the accession codes 6SBN (WT PE-H) and 6SCD (PE-H Y250S). All structure related figures were prepared with PyMOL (Schrödinger, LLC, United States)[5], for the structure based alignments we used the PDBeFold webserver[6].

[2]http://www.embl-hamburg.de/Auto-Rickshaw/
[3]https://arpwarp.embl-hamburg.de
[4]https://www.rcsb.org
[5]https://pymol.org/2/
[6]http://www.ebi.ac.uk/msd-srv/ssm/cgi-bin/ssmserver

Molecular Docking Computations

For the molecular docking, ligands BHET, MHET, and 2-HE(MHET)$_4$ were drawn and converted into a 3D structure with Maestro (Schrödinger, LLC, New York). The ligands and proteins were protonated according to pH 7.4 using the Epik routine in Maestro. The ligands were subsequently docked into the binding pocket of the respective enzymes using a combination of AutoDock as a docking engine and the DrugScore^{2018} distance-dependent pair-potentials as an objective function (Goodsell et al., 1996; Sotriffer et al., 2002; Dittrich et al., 2019). In the docking, default parameters were used, with the exception of the clustering RMSD cutoff, which was set to 2.0 Å. Binding modes were considered valid, if they were part of a cluster that comprised at least 20% of all docking poses.

Bioinformatic Tools and Software

Multiple sequence alignment was carried out with Clustal Omega using default settings (Sievers et al., 2011), basic local alignment searches (BLAST) were done using the web service of the National Center for Biotechnology Information (NCBI) (Altschul et al., 1990; Wheeler et al., 2003), for data visualization, GraphPad Prism (GraphPad Software, Inc., United States) and OriginLab (OriginLab Corporation, United States) were used. For the identification and description of molecular cavities, the MOLE 2.5 software (Sehnal et al., 2013) was used employing default options with an 8.0 probe radius.

RESULTS

PE-H From *Pseudomonas aestusnigri* Is a Polyester Hydrolase

Pseudomonas aestusnigri, a bacterium belonging to the *P. pertucinogena* phylogenetic lineage, shows hydrolytic activity on different polyester substrates (Molitor et al., 2020). Additionally, we observed hydrolytic activity indicated by formation of clear halos upon growth of *P. aestusnigri* on agar plates containing Impranil DLN, an anionic aliphatic polyester-polyurethane used for surface coating of textiles (**Figure 1A**). Bioinformatic analysis of the *P. aestusnigri* genome sequence (Gomila et al., 2017) led us to predict the existence of a polyester hydrolase coding gene (Bollinger et al., 2020). The respective gene was cloned into a pET-22b(+) expression vector and produced by expression in *E. coli* BL21(DE3). The recombinant bacteria showed significant hydrolytic activity when grown on Impranil DLN containing solid media (**Figure 1B**). Hence, the predicted polyester hydrolase gene indeed codes for a functional polyester hydrolase which we named PE-H. The respective gene (locus tag: B7O88_RS11490) codes for a protein of 304 amino acids (protein id: WP_088276085.1) comprising a signal peptide of 25 amino acids for Sec-dependent translocation as predicted by SignalP (Almagro Armenteros et al., 2019). For further characterization, the protein was produced in soluble form along with its native signal peptide and fused to the vector-encoded hexa-histidine tag, which allowed a one-step purification by immobilized metal ion chromatography (IMAC) (**Figure 1D**). The purified protein has a

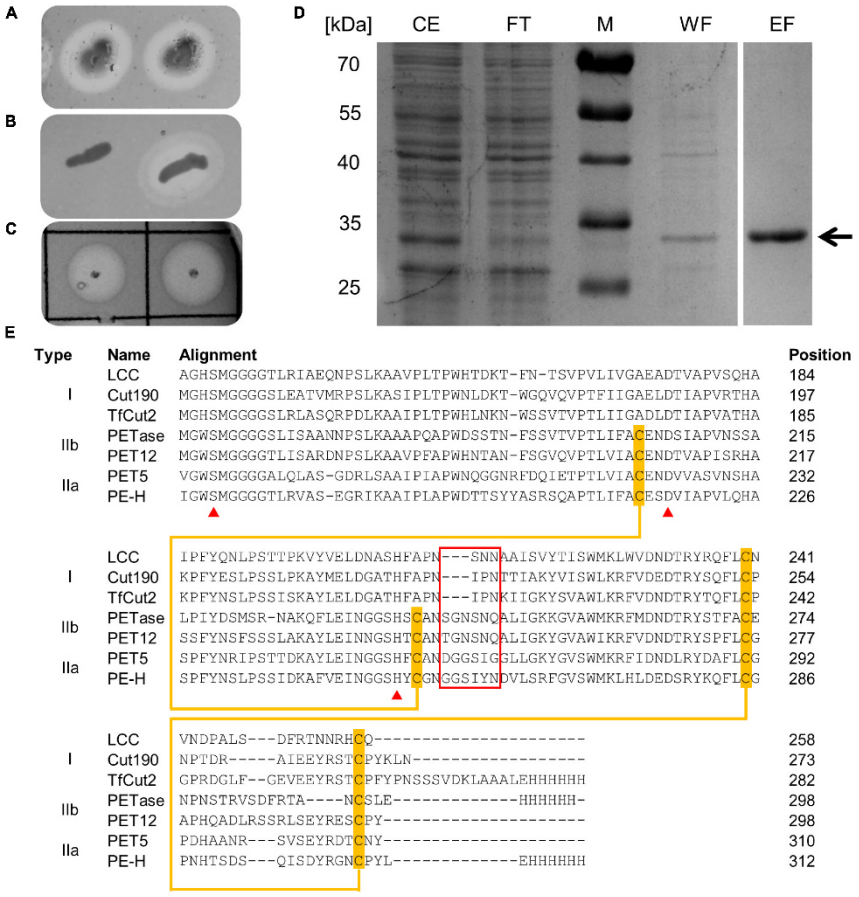


FIGURE 1 | **(A)** Colonies of *P. aestusnigri* grown on an Impranil DLN containing agar plate show clearing halos, indicating polyester hydrolase activity. **(B)** Hydrolytic activity of *E. coli* BL21(DE3) empty vector control (left) and PE-H production strain (right) on Impranil DLN containing agar plate. **(C)** Hydrolytic activity of purified PE-H on Impranil DLN containing agar plate. **(D)** Coomassie brilliant blue stained gel after SDS-PAGE of cell extracts of *E. coli* BL21(DE3) containing plasmid pET22b_PE-H$_{c6H}$ before and after purification by IMAC; the position of PE-H is indicated by an arrow. Lanes contained cell extract (CE), flow through (FT), washing step (WF), and eluted protein (EF); the size of the molecular weight standards (M) are indicated on the left. **(E)** Part of a multiple sequence alignment of PE-H with amino acid sequences of different PET hydrolytic enzymes using the program Clustal Omega. The full length alignment can be found in the **Supplementary Figure S1**. The enzymes were assigned to different types of polyester hydrolases (Joo et al., 2018). Amino acid residues of the catalytic triad are marked by a red triangle, disulfide forming cysteine residues are highlighted in orange and connected by an orange line. Amino acids forming the extended loop region which is specific for type II PET hydrolytic enzymes are framed in red. Abbreviations are: leaf-branch compost metagenome cutinase (LCC); *Saccharomonospora viridis* cutinase (Cut190); *Thermobifida fusca* cutinase (TfCut2); *Ideonella sakaiensis* PET hydrolase (PETase); *Polyangium brachysporum* PET hydrolase (PET12); *Oleispira antarctica* PET hydrolase (PET5).

molecular weight of about 32 kDa and prominent activity toward the polymer substrate Impranil DLN (**Figure 1C**). Furthermore, a basic local alignment search using BLAST (Altschul et al., 1990; Wheeler et al., 2003) with the protein sequences in the protein database (PDB) as search set revealed similar sequences of cutinases and PET hydrolytic enzymes originating from *S. viridis*, *Thermobifida* sp. and *I. sakaiensis* with 48–51% identity at more than 80% query coverage.

The relation of PE-H to cutinases and other PET hydrolytic enzymes was analyzed by multiple sequence alignments of the amino acid sequence of PE-H with representative examples of each type of proven PET hydrolytic enzymes as proposed by

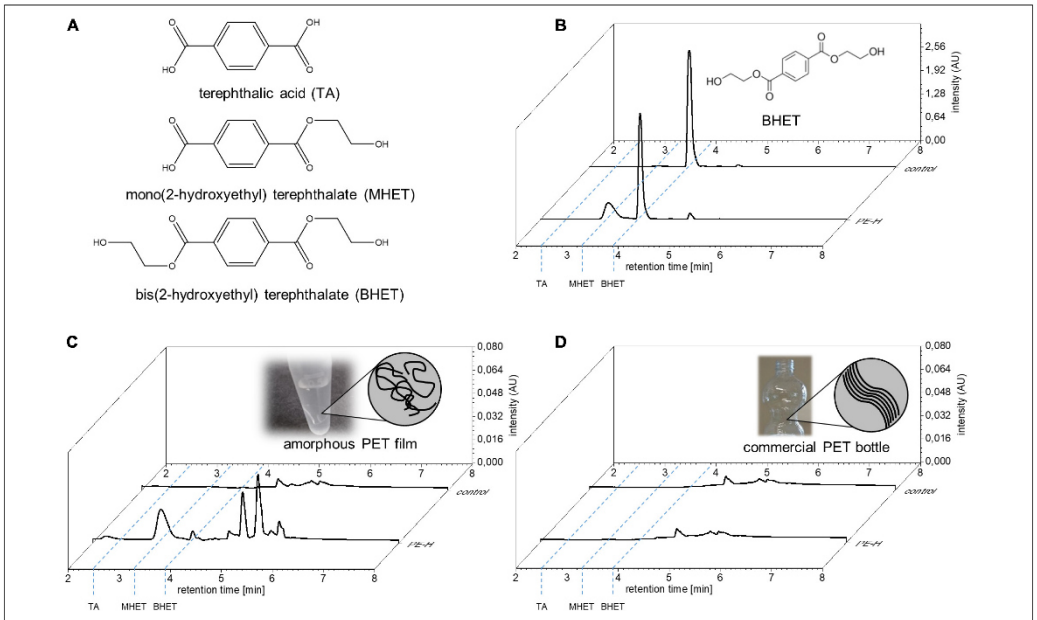


FIGURE 2 | Hydrolysis of PET by PE-H. **(A)** Structural formula of aromatic reaction products obtained by hydrolysis of PET or BHET. **(B)** UPLC analysis of reaction products obtained with PE-H and the substrates bis(2-hydroxyethyl) terephthalate (BHET), **(C)** amorphous PET film (PETa) and **(D)** PET film derived from a commercial single use PET bottle (PETb). The UPLC trace at 240 nm wavelength is shown for the control reaction without addition of the enzyme (trace in the back, labeled as control), and the enzyme reaction (trace in front, labeled PE-H). The retention times of the standard compounds terephthalic acid (TA), and bis(2-hydroxyethyl) terephthalate (BHET) as well as of the reaction product mono(2-hydroxyethyl) terephthalate (MHET) are indicated by blue dashed lines. The respective substrate of the reaction is depicted in the back of each diagram. For BHET, the structural formula is given; for PET films, a cartoon representation is shown to illustrate the different arrangement of PET fibers in amorphous and crystalline films.

Joo et al. (2018) was done using Clustal Omega. For type I, sequences, the leaf-branch compost metagenome cutinase LCC (Sulaiman et al., 2012), the cutinase Cut190 from *S. viridis* (Kawai et al., 2014), and the cutinase TfCut2 from *T. fusca* (Roth et al., 2014) were used. PETase from *I. sakaiensis* (Yoshida et al., 2016) and PET12 from *Polyangium brachysporum* (Danso et al., 2018) served as examples for type IIb and PET5 from *Oleispira antarctica* (Danso et al., 2018) for type IIa, as demonstrated before (Taniguchi et al., 2019). The alignment revealed a clear discrimination of PE-H from cutinases of type I, owing to an additional disulfide bond and additional amino acids close to the catalytically active histidine (**Figure 1E** and **Supplementary Figure S1**). Both characteristics are typical for PET hydrolytic enzymes of type II (Joo et al., 2018). Furthermore, due to the amino acid composition of the region constituting additional amino acids, known as extended loop in case of PETase (Joo et al., 2018), PE-H can be classified as a type IIa PET hydrolytic enzyme together with PET5 from *Oleispira antarctica* (Danso et al., 2018). In fact, PE-H is a close homolog of the enzymes encoded by *Pseudomonas sabulinigri*, *P. pachastrellae*, and *P. litoralis* [all belong to the phylogenetic lineage of *P. pertucinogena* (Peix et al., 2018)], which were proposed as PET-degrading enzymes of type IIa by *in silico* sequence comparison (Joo et al., 2018).

PE-H Degrades PET

The classification of PE-H as a type IIa PET degrading enzyme suggested PET degrading activity. Therefore, we tested the enzymatic activity of PE-H with the substrates monomeric bis(2-hydroxyethyl) terephthalate (BHET), amorphous PET film (PETa) and PET film derived from a commercial single use PET bottle (PETb). The PET polymer consists of esterified terephthalic acid (TA) and ethylene glycol (EG), allowing for TA, EB, and esters of both compounds with different degree of polymerization as degradation products, for example BHET or mono(2-hydroxyethyl) terephthalate (MHET) (**Figure 2A**). The enzyme hydrolyzed both BHET (**Figure 2B**) and PETa (**Figure 2C**), releasing mono(2-hydroxyethyl) terephthalate (MHET), but no terephthalic acid (TA). When PETb was used as a substrate, no hydrolysis product was detected (**Figure 2D**).

The determination of the crystal structure of PETase from *I. sakaiensis* (Yoshida et al., 2016) by different groups in 2017/2018 (Han et al., 2017; Austin et al., 2018; Chen et al., 2018; Fecker et al., 2018; Joo et al., 2018; Liu B. et al., 2018;

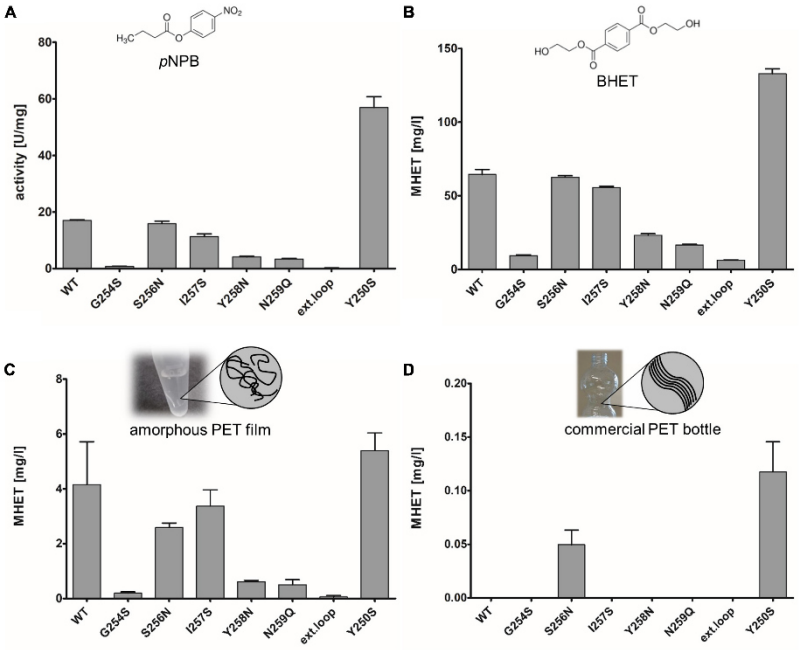


FIGURE 3 | Enzymatic activity of PE-H and different variants constructed by site directed mutagenesis. Substrates were **(A)** 4-nitrophenyl butyrate (*p*NPB), **(B)** bis(2-hydroxyethyl) terephthalate (BHET), **(C)** amorphous PET film, and **(D)** PET from a commercial single use bottle. Standard deviations of three individual reactions are shown as error bars. The respective substrate is depicted above each diagram. For *p*NPB and BHET, structural formulas are shown; for amorphous and commercial PET a cartoon representation is shown to illustrate the different arrangement of PET fibers.

Liu C. et al., 2018; Palm et al., 2019) allowed for detailed insights into the structure-function relationship of PETase with regard to PET degradation. By structural comparison to cutinases, the active site cleft of PETase was shown to be wider and shallower which seemed to be important for the PET hydrolytic activity of PETase (Austin et al., 2018; Liu B. et al., 2018). Moreover, a number of regions on the protein surface were proposed to be important for the efficient PET hydrolysis, among them (i) serine residue S238 which showed an important contribution to the enzyme activity (Joo et al., 2018), (ii) tryptophan residue W159 with proposed contribution in π-stacking interaction with terephthalic acid moiety of PET (Han et al., 2017; Chen et al., 2018), and (iii) a conserved extended loop region consisting of six amino acids connecting β8-α6 (Joo et al., 2018), which was proposed to mediate substrate binding. The amino acid composition of PE-H differs at the corresponding positions (**Supplementary Table S1**) prompting us to introduce a series of single amino acid substitutions into PE-H by site directed mutagenesis and subsequently evaluate the activity of the respective enzyme variants with different substrates.

The substitution of active site amino acids by alanine led to inactive variants as expected (data not shown). The extended loop region can be found for PE-H as well, with an amino acid composition representative for polyester hydrolases of type IIa (Joo et al., 2018). Both, single substitutions at each position of the extended loop to those present in PETase, and the exchange of the entire extended loop region to that of PETase were constructed and the resulting variants were produced, purified, and tested for activity against the substrates 4-nitrophenyl butyrate (*p*NPB), BHET, and PET film. In comparison to the wild type enzyme PE-H, neither the single substitution variants nor the loop exchange variant, designed to match the amino acid residues at the corresponding positions of PETase (**Supplementary Table S1**), showed a higher specific activity toward BHET or PET film. Of these variants, G254S, Y258N, N259Q, and the variant with the combined mutations also showed a significantly decreased esterase activity determined with *p*NPB as the substrate (**Figure 3A**). The reduced activity of the variants was accompanied by a decrease in melting temperature of about 5–10°C (**Supplementary Table S2**) in comparison to the wild type enzyme (T_m at about 51°C), indicating a destabilizing effect of the mutations. Variant PE-H S256N, I257S, and Y250S showed a less drastic decrease in melting temperature of about 1–3°C.

Interestingly, the variant Y250S of PE-H showed a more than threefold increase in specific activity for *p*NPB, assuming a general importance of this amino acid position for enzymatic activity. The same trend was observed with BHET as the substrate (**Figure 3B**), showing decreased activity of all variants compared to the wild type except for variant Y250S. Similar results were obtained with PETa as the substrate with variant Y250S producing more MHET than the wild type enzyme or any other variant tested (**Figure 3C**). Nevertheless, the effect was less clear compared to the soluble substrates BHET and *p*NPB. The activity of the enzyme on PET derived from a commercial single use bottle as determined by MHET release appeared to be in general low. Only two variants, Y250S and S256N, led to detectable formation of MHET from PETb, but only at low amounts (**Figure 3D**).

PE-H Crystal Structure Confirms Similarity to PET Degrading Enzymes

In order to gain insight into the molecular basis of the polyester hydrolytic activity of PE-H and its variant Y250S, we solved the crystal structures of both enzymes. The PE-H structure was solved at 1.09 Å resolution, containing one monomer in the asymmetric unit, with 10.7% for R_{work} and 13.7% for R_{free}. Although the electron density was of extremely good quality, the N-terminal part (aa 1–37) and one short stretch (aa 286–291) were not visible. The structure of variant PE-H Y250S was solved at 1.35 Å resolution; here, only the N-terminus (aa 1–39) is missing. Electron densities around the active site of PE-H and variant Y250S are shown in **Supplementary Figure S2**. The crystal structure of PE-H Y250S contained a PEG molecule bound to the protein surface (**Supplementary Figure S3**). Data collection and structure refinement statistics are given in the **Supplementary Material** (**Supplementary Table S3**). The PE-H protein shows a canonical α/β-fold consisting of a central twisted β-sheet composed of 9 β-strands flanked by 7 α-helices on both sides (**Figure 4**), as already reported for homologous structures, i.e., cutinases and PETases (Han et al., 2017; Numoto et al., 2018). Two disulfide bonds are present in PE-H connecting C214-C251 and C285-C302 as is common for type II PET degrading enzymes.

Both structures reported here display a nearly identical overall fold (WT PE-H to PE-H Y250S: rmsd 0.188 Å over 217 Cα atoms) with main differences in two loop regions: the loop connecting β3-α2 (aa 98–104) adopts a "close" conformation in WT PE-H with regard to the active site cleft, with a loop connecting β4-α3 (aa 123–128) positioned parallel to it, while both are shifted against each other in the Y250S mutant (**Figure 5A**), thereby creating more space in the catalytic site. The highly conserved residues S171, D217 and H249 build the catalytic triad which is located closely below the surface with the S171 position known as the "nucleophilic elbow". The oxyanion hole is constituted by the backbone NH groups of M172 and F98. The loop arrangement narrowing the active site cleft in the wild type enzyme is stabilized by a polar contact between the hydroxyl group of Y250 and the backbone amine of E102 (**Figure 5B**). This structural rearrangement further leads to an increased active site cavity volume from 153 $Å^3$ for WT PE-H to 362 $Å^3$ for variant Y250S as determined with the program MOLE2.5. A comparison of the active site cleft molecular surface of WT PE-H and variant Y250S shows the altered topology affecting the observed increase in active site cavity volume. In WT PE-H, residues of the loop connecting β3-α2 (F98, V99, S100) together with D129 and I219 are arranged in a way that access to the active site is limited, whereas PE-H variant Y250S shows a much deeper cleft (**Figure 5C**).

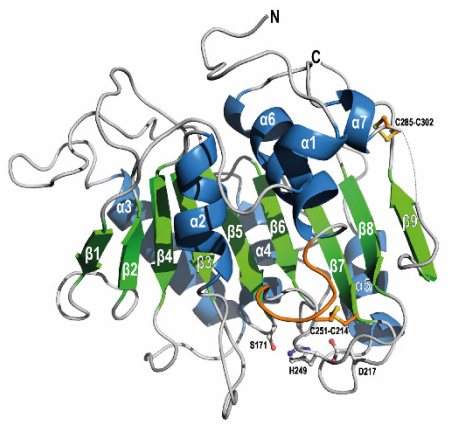


FIGURE 4 | Overall fold of WT PE-H (PDB code 6SBN) in cartoon representation. The short stretch (aa 286–291) is not visible (gray dashed line). The α/β-fold consists of a central twisted β-sheet composed of 9 β-strands flanked by 7 α-helices on both sides. The extended loop region and the Cys residues (ball-and-sticks) forming the disulfide bonds are highlighted in orange. Residues building the catalytic triad are shown as gray ball-and-sticks with labels. Rotations between the structure depictions are indicated in **Supplementary Figure S4**.

To identify structural homologs of PE-H, we performed structural alignments of both PE-H structures independently against the PDB. At maximum, we found 60 similar protein chains with rmsd values in the range of 1.1 to 2.7 Å for Cα atoms and 53% sequence identity at maximum (data not shown). For PE-H, the Cut190 triple mutant (TM) S176A/S226P/R228S (PDB 5ZRR) from *S. viridis* is the most similar (rmsd 1.18 Å); for variant Y250S, it is the PETase double mutant (DM) R103G/S131A (PDB 5XH3) from *I. sakaiensis* (rmsd 1.17 Å). The 10 protein chains most similar to both PE-H structures described here are listed in **Supplementary Table S4**. To further analyze the different architecture of both PE-H variants reported here, we compared their surfaces with those of their structurally most similar homolog (**Supplementary Figure S5**). All four molecules share a similar pattern of surface charge which is dominated by larger patches of either positive or negative charge or hydrophobic areas, respectively. With regard to the active site cleft, Cut190 TM shows a similar narrowed active site as does WT PE-H, whereas PETase DM has a larger and deeper cleft as is the case for variant Y250S.

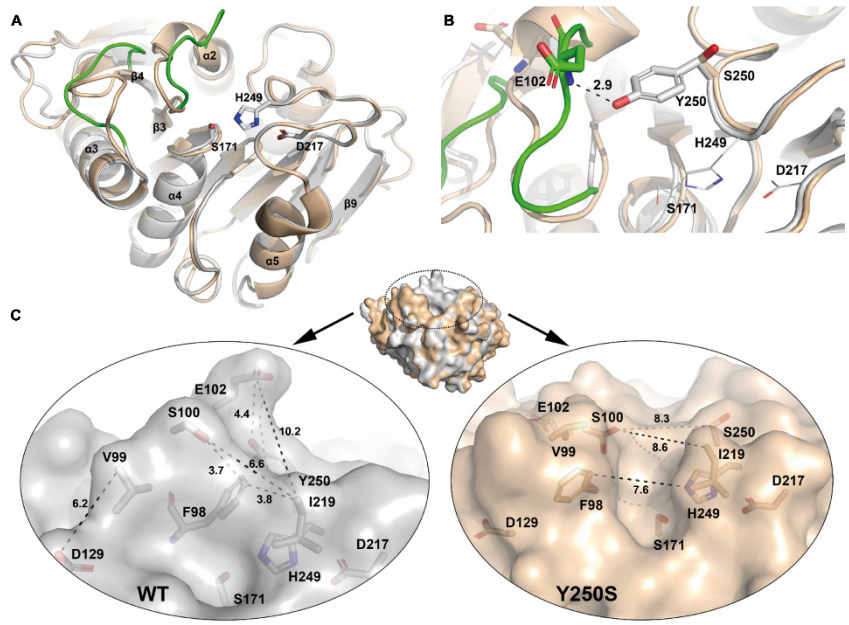


FIGURE 5 | Superposition of PE-H (PDB code 6SBN, silver) and its variant Y250S (PDB code 6SCD, wheat). **(A)** The top view on the active site and the different orientation of the two loops is shown. Residues of the catalytic triad are depicted as sticks with labels, loops of PE-H are colored in green to highlight the structural differences. **(B)** A zoom in the active site cleft with the catalytic residues depicted as lines. Remarkably, Y250 (gray sticks) makes a polar contact with E102 (green sticks) in the wild type protein, the distance is shown as a dashed line and given in Å, whereas in variant Y250S, residue E102 (wheat sticks) is far apart [color code as in **(A)**]. **(C)** Surface representation of WT PE-H (silver) and variant Y250S (wheat), on top, both structures are superposed with the catalytic cleft indicated by a dashed circle. In the panels, the differences in both structures are displayed in detail; with WT PE-H in the left panel and variant Y250S in the right panel. Residues confining the active site cleft are depicted as sticks, labeled, and connected by dotted lines, with the corresponding distances given in Å. Rotations between the different structure depictions are indicated in **Supplementary Figure S4**.

Superimposition of the PE-H structure and the structures of Cut190 (**Figure 6A**) and PETase (**Figure 6B**) revealed a high structural identity. PE-H possesses a disulfide bond linking residues C285 and C302 which is also present in Cut190, and a second disulfide bond located close to the active site (**Figure 6A**, C251-C214); both these disulfide bonds are also present in PETase (**Figure 6B**). These three structures which represent PET hydrolytic enzymes of types I, IIa, and IIb differ mainly within five loop regions (**Figures 6C,D**) comprising PE-H amino acid positions 68–71, 84–87, 97–102, 124–130, and 254–259. Three of these regions at positions 97–102, 124–130, and 254–259 are located in the vicinity of the active site indicating a putative contribution to substrate binding.

Molecular Docking Computations Suggest PE-H Substrate Binding Mode

Substrate binding was analyzed by docking the PET tetramer 2-HE(MHET)$_4$, BHET, and MHET to wild type PE-H and variant Y205S using predicted protonation states at pH 7.4 for both the ligands and proteins. Lowest-energy configurations of the ligands in the proteins from the largest cluster were taken as binding poses as done previously (Diedrich et al., 2016; Krieger et al., 2017). Binding modes were considered valid, if at least 20% of all poses are contained in this cluster. During the docking, ligands were allowed to explore the whole protein to ensure an unbiased sampling of potential binding poses.

The predicted binding poses provide mechanistic insights into the function of PE-H and highlight the differences between wild type PE-H and variant Y250S. In PE-H, MHET and BHET are predicted to bind adjacent to the catalytic site (**Figure 7A**). In the *apo* crystal structure, the catalytic site of PE-H is apparently too narrow to allow favorable substrate binding (**Figure 5C**). This suggests that in PE-H a conformational change is necessary to accommodate a substrate. BHET and MHET bind with the phenyl rings to a hydrophobic groove and are additionally stabilized via hydrogen-bonding interactions to S103, D106, S248, and S256 (**Figure 7B**). Although no valid binding mode was identified for 2-HE(MHET)$_4$, as no cluster contained more than 4% of all binding poses, the lowest-energy pose found in

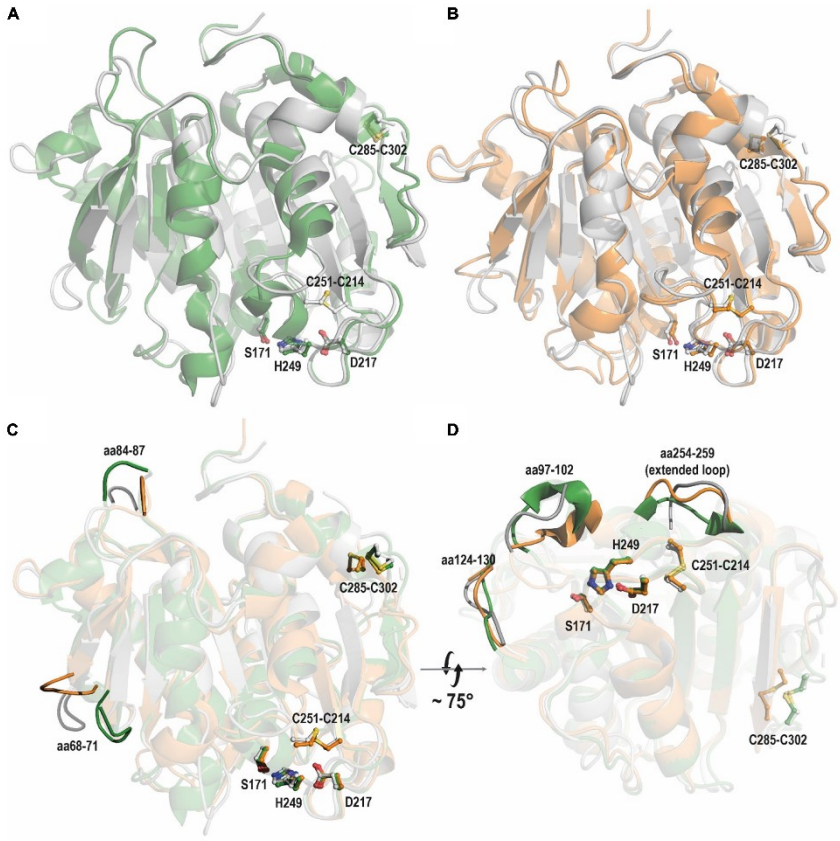


FIGURE 6 | Overlay of PE-H WT (PDB code 6SBN, gray) with **(A)** Cut190 (PDB code 4WFI, green) and **(B)** PETase (PDB code 6ANE, orange). Residues of the active site and disulfide bonds are represented as sticks with labels (according to PE-H numbering). **(C)** Superimposition of PE-H (PDB code 6SBN, gray), PETase (PDB code 6ANE, orange), and Cut190 (PDB code 4WFI, green). Five regions of the superimposition are represented in full tone on fade-out color to highlight the differences of the structures. In **(D)** the structures are turned from bottom to top.

this docking is located in the adjacent groove like BHET and MHET (**Figure 7B**), which hints at the groove's importance for the function of PE-H.

In the variant Y250S, the catalytic site has twice the volume than in wild type PE-H. As a result, the predicted binding pose of MHET is located in the catalytic site, whereas BHET still binds to the same groove as in PE-H (**Figure 7C**). MHET is stabilized by hydrophobic interactions to F98, V99, M172, and I219, and the ester carbonyl carbon is placed at an optimal distance to be attacked by S171 (**Figure 7D**). A second binding pose of BHET, which is 0.43 $kcal \cdot mol^{-1}$ less favorable than the lowest-energy one, is found in the catalytic site similar to MHET (**Figure 7E**) and is thereby in an optimal position to be hydrolyzed by S171. Again, no valid binding modes were identified for 2-HE(MHET)$_4$, likely owing to the size of the ligand. The presence of binding poses for BHET adjacent to the active site suggests a possible polymer binding mode, where one polymer unit binds to the groove adjacent to the catalytic site, a second unit bridges the distance to the catalytic site, and a third unit is cleaved from the polymer chain (**Figure 7F**).

DISCUSSION

Enzymes acting on PET are almost exclusively homologs of cutinases with many of them originating from thermophilic actinobacteria (Nikolaivits et al., 2018). PETase from *I. sakaiensis* was the first enzyme isolated from a mesophilic host and

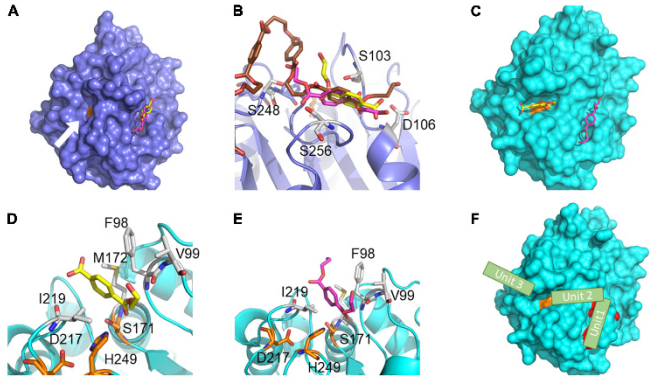


FIGURE 7 | Predicted ligand binding modes in wild type PE-H and variant Y250S. The predicted binding poses of BHET (magenta), MHET (yellow), and 2-HE(MHET)$_4$ (brown) in WT PE-H (navy) and the variant Y250S (cyan). In **(A,C,F)** S171 is shown in orange, and in **(B,D,E)** the catalytic triad (S171, D217, and H249) is shown in orange and interacting residues are shown in white. **(A)** In wild type PE-H, BHET and MHET bind to a groove adjacent to the catalytic site (white arrow). **(B)** BHET and MHET bind to the hydrophobic groove and are stabilized by hydrogen bonding interactions with S103, D106, S248, and S256. 2-HE(MHET)$_4$ binds similarly to BHET and MHET in the groove adjacent to the catalytic site. **(C)** In the variant Y250S, MHET binds to the catalytic site, while BHET occupies the hydrophobic groove. **(D)** MHET binds to the catalytic site and is stabilized by hydrophobic interactions to F98, V99, M172, and I219 such that S171 can attack the carbonyl carbon for ester hydrolysis. **(E)** A second binding pose of BHET binds similar to MHET. **(F)** Proposed mechanism of PET polymer interaction. Residues G254, Y258, and N259, which when substituted decrease esterase activity, are shown in red. One polymer unit (stylized green rectangle) binds to the groove adjacent to the catalytic site, a second unit bridges the distance to the catalytic site, and a third unit is cleaved from the polymer chain at the catalytic unit.

catalyzing the biodegradation of PET at mild temperature of 30°C (Yoshida et al., 2016). Structural studies showed characteristic features discriminating PETase from cutinases with PET hydrolytic activity (Joo et al., 2018) and lead to the proposal of three subclasses of PET hydrolyzing enzymes. With this information, novel PET hydrolytic enzymes were identified, among them enzymes from marine origin (Danso et al., 2018). This finding is of considerable interest because the oceans are known to be a sink for the worldwide plastic waste (Lebreton et al., 2018).

A Novel PET Degrading Enzyme From the Marine Bacterium *P. aestusnigri*

In this study, we have identified the protein PE-H, a novel polyester hydrolase from the marine bacterium *P. aestusnigri* which hydrolyses different polyethylene terephthalate substrates with mono(2-hydroxyethyl) terephthalate (MHET) as the major hydrolysis product. The formation of MHET, but not TA as hydrolysis product was previously described for other PET degrading enzymes, for example TfCut2 from *T. fusca* (Barth et al., 2015; Wei et al., 2016), Thc_Cut2 from *T. cellulosilytica* (Herrero Acero et al., 2011), and PETase from *I. sakaiensis* (Yoshida et al., 2016). However, in case of TfCut2, MHET was almost completely hydrolyzed to TA after 24 h of reaction (Barth et al., 2015), and the cutinases Thc_Cut1 from *T. cellulosilytica* and Thf42_Cut1 from *T. fusca* were reported to release more TA than MHET from PET film after a reaction time of 120 h (Herrero Acero et al., 2011). Hence, PE-H accumulating MHET from PET film appears to be common feature among PET hydrolyzing enzymes, but the almost complete absence of TA after 48 h of reaction was not observed before.

Hydrolysis of PET by PE-H yielded 4.2 (±1.6) mg/L MHET after a reaction time of 48 h at 30°C, variant PE-H Y250S produced 5.4 (±0.6) mg/L MHET under the same conditions. This amount is considerably lower than reported for the *I. sakaiensis* PETase with more than 150 mg/L after 96 h reaction time at 30°C (Austin et al., 2018). However, it should be mentioned that activities are difficult to compare because different studies use PET substrates which differ among each other, e.g., by the degree of crystallinity. Obviously, a comprehensive comparative study of different PET degrading enzymes would be required. In case of PE-H, the marine origin of the producer *P. aestusnigri* suggests that natural substrates may be aliphatic polyesters produced by algae and plants, because marine bacteria of the *P. pertucinogena* lineage have been found associated with these organisms (Bollinger et al., 2020).

Rational Mutagenesis Resulted in PE-H Variant Y250S With Improved Activity

Wild type PE-H was unable to hydrolyze a PET film substrate obtained from a commercial PET bottle (PETb). This observation prompted us to identify differences in the amino acid sequences between PE-H and the *I. sakaiensis* PETase and analyze them by a site directed mutagenesis study. As the result, a single amino acid substitution (Y250S) was identified which led to a significant increase in enzymatic activity of PE-H toward different substrates, including PETb. In comparison to the results obtained with PETa, the best performing variant PE-H

Y250S showed a 45-fold reduced amount of MHET produced with PETb as the substrate. This observation can be explained by the molecular arrangement of the PET fibers in PETb which is known to comprise more crystalline regions and thus, is more recalcitrant to enzymatic degradation (Wei and Zimmermann, 2017). The highly ordered structure of crystalline PET hampers the enzymes access to single PET fibers, particularly at temperatures below the glass transition temperature of the polymer. Nevertheless, two single amino acid substitution variants, S256N and Y250S, produced detectable amounts of MHET from PETb while the wild type enzyme did not.

The Crystal Structure of PE-H Is the First Structure of Type IIa PET Hydrolytic Enzyme

In an attempt to rationalize the PE-H activity toward PET as a substrate and to explain the effect of the Y250S substitution, we report here the first crystal structure of a type IIa PET hydrolytic enzyme and its comparison with the crystal structure of the improved variant PE-H Y250S. Interestingly the comparison of both structures showed that the substitution Y250S replacing the aromatic residue tyrosine located next to the active site histidine by the small residue serine significantly improved the enzymatic activity toward different substrates. The relevance of this serine residue has also been observed with PETase variant S238F which showed an about 40% reduced enzymatic activity (Joo et al., 2018). Thus, a small, polar, and uncharged amino acid at the position next to the active histidine seems to be an important characteristic of PET hydrolytic enzymes of type IIb, whereas enzymes of type I and type IIa possess large aromatic residues at the corresponding position (see **Figure 1E** and **Supplementary Figure S1**). This finding is also supported by a very recent study which reported a significant increase in activity of the cutinase TfCut2 from *Thermobifida fusca* obtained by substitution of F209 (the corresponding position in PE-H is Y250) with serine and alanine (Furukawa et al., 2019). The authors tested additional substitutions at this position introducing amino acids with chemically different side-chains and concluded that the increase in activity correlated with the size of the respective side-chain. The structural analysis of PE-H and the variant Y250S showed a rearrangement of the loop connecting α2-β3, creating space which results in a better accessible active site. The role of the active site cleft size was also found to be important for PETase, which was reported to be three times wider at its widest point compared to *T. fusca* cutinase TfCut2 (Austin et al., 2018) and also wider compared to *Fusarium solani pisi* cutinase (Liu B. et al., 2018). Hence, the authors proposed that the wider active side cleft enables PETase to accommodate larger substrates like PET. The enhanced activity of PETase at lower temperatures compared to cutinases was attributed to enhanced flexibility within the active center whereby the active site proximal disulfide bridge diminishes loss of stability that is to be expected as tradeoff for enhanced flexibility (Fecker et al., 2018). This disulfide bridge is a common feature of all type II PET hydrolytic enzymes and is in consequence also found in PE-H. The reshaping of the active site cleft of PE-H may therefore be the main reason for the enhanced catalytic activity observed in this study, enabling the enzyme to accommodate larger substrates as seen with PET from a commercial single use PET bottle. In PE-H, besides the steric effects of a larger amino acid, a polar contact between Y250 and E102 was observed which may contribute to narrowing the active site. However, most cutinases carry a phenylalanine at the corresponding position (Joo et al., 2018) which does not allow for a polar contact. This observation therefore does not explain the differences between PET hydrolytic enzymes of type I and II regarding PET hydrolysis. To reveal further differences between PET hydrolytic enzymes of types I, IIa, and IIb, the structures of one enzyme of each type (namely PE-H, Cut190, and PETase) were superimposed and three loop regions located close to the active site were identified as differing. Some amino acids of the corresponding regions in PETase were already reported to be involved in substrate interaction, for example Y87 (F98 in PE-H) and N246 (Y258 in PE-H) which were assigned to substrate binding subsites I and IIc (Joo et al., 2018), respectively.

The PET Binding Mode of PE-H

To gain insight into protein substrate interaction of PE-H with PET, molecular docking with substrates MHET, BHET, and a previously reported PET tetramer (Joo et al., 2018) was conducted. Interestingly, wild type PE-H was not found to viably bind the tested substrates in the active site, but substrate molecules accumulated in an adjacent groove. For variant Y250S, viable binding modes for MHET and BHET, but not the PET tetramer were found. However, biochemical data revealed catalytic activity of PE-H and variant Y250S and PET film as a substrate. This observation might be connected to the flexibility of the active site of PE-H, suggesting that PE-H can rearrange its active site cleft to accommodate respective substrates. In line with that, flexibility of the active site was shown for PETase and reported to be crucial for the enzyme activity at room temperature (Fecker et al., 2018). Furthermore, MHET was accommodated in a catalytically viable position in PE-H Y250S; thus, the observed absence of MHET hydrolysis cannot be attributed to low affinity or preferred binding in a detrimental position for hydrolysis. Molecular reasons for that substrate preference might be addressed in further studies.

The two distinct binding poses observed for variant PE-H Y250S hint at a mechanism in which one polymer unit binds to the groove adjacent to the catalytic site, a second unit bridges the distance to the catalytic site, and a third unit is cleaved from the polymer chain. This mechanism with the adjacent groove acting as an anchor point for the polymer could increase the processivity of PE-H, similar to the adjacent substrate binding site in Hyal lyases (Rigden and Jedrzejas, 2003; Stern and Jedrzejas, 2008). A similar PET binding mechanism that involves a binding grove that stabilizes different units of a polymeric substrate was proposed for PETase binding elements of 2-HE(MHET)$_4$ in subsites I, IIa, IIb, and IIc, facilitating the binding of the PET polymer (Joo et al., 2018). The presence of an adjacent substrate binding site may also contribute to the significant decrease in esterase activity of PE-H variants G254S, Y258N, and N259Q and the variant with the combined mutations. All three substitutions are located at the adjacent groove (**Figure 7F**)

and alter their shape and molecular recognition properties, presumably hampering substrate binding that way.

As we did not obtain valid docking poses for 2-HE(MHET)$_4$, we cannot comment on the *gauche*-to-*trans* ratio of the substrate's OC–CO torsion angle when bound to the catalytic site, which has been discussed recently (Joo et al., 2018; Seo et al., 2019; Wei et al., 2019b). Our lowest-energy pose of 2-HE(MHET)$_4$ (**Figure 7B**), which is located in the adjacent groove, shows all-*gauche* torsion angles. This result is in line with low *trans*-to-*gauche* ratios of (9 – 14): (91 – 86) experimentally determined for amorphous PET (Schmidt-Rohr et al., 1998; Wei et al., 2019b). We note, though, that even an increase in the *trans* content by a factor of ~3 as discussed (Joo et al., 2018; Seo et al., 2019; Wei et al., 2019b) relates to changes in the conformational free energy of the substrate on the order of the thermal energy and, thus, may be compensated by favorable interactions with the protein.

CONCLUSION

Polyester hydrolases are of considerable interest for a variety of biotechnological applications (Nikolaivits et al., 2018), for example the removal of cyclic PET oligomers from polyester fibers in the textile industry (Riegels et al., 1997). As far as biocatalytic degradation of PET in an industrial context is concerned, elevated temperatures close to the glass transition temperature of PET are desirable (Wei and Zimmermann, 2017). Hence, protein engineering of PE-H would be required to design a PE-H variant for efficient PET hydrolysis at elevated temperatures as has been demonstrated for PETase (Son et al., 2019). However, additional applications appear feasible for PE-H as it has been demonstrated that the PE-H homologous enzyme PpelaLip originating from the closely related genus *Pseudomonas pelagia* can be used for the biodegradation of different synthetic polyesters and may thus be applicable for waste water treatment (Haernvall et al., 2017). In this case, the activity of the respective biocatalyst at low temperatures is an advantage allowing for the hydrolysis of synthetic polymers at 15°C (Haernvall et al., 2018).

We have described here the identification, biochemical and structural characterization of the polyester hydrolase PE-H from the marine mesophilic bacterium *P. aestusnigri*. The crystal structure of PE-H represents the first structure of a type IIa PET hydrolytic enzyme thus closing the gap between PET hydrolytic enzymes of type I and IIb. We furthermore succeeded to significantly increase the enzymatic activity of PE-H toward different substrates by introducing a single amino acid substitution, namely Y250S. This substitution resulted in a rearrangement of the active site conformation, favored by prevention of a polar contact of Y250 located next to the catalytic active histidine, and a loop region of the active site cleft. Furthermore, a PET polymer binding mechanism was proposed based on molecular docking computations. Our results thus provide important information regarding structural features required for efficient polyester degradation and indicate that marine bacteria such as *P. aestusnigri* may prove as a prolific source for such enzymes.

DATA AVAILABILITY STATEMENT

Structures were deposited in the protein data bank (https://www.rcsb.org) under the accession codes 6SBN (WT PE-H) and 6SCD (PE-H Y250S).

AUTHOR CONTRIBUTIONS

K-EJ conceived the research concept. AB, ST, MF, and SS designed the experiments. AB, ST, and EK-G performed the experimental work. SK, AH, and SS performed the crystallization trials and structure solution. CG and HG conducted the molecular docking computations, discussed the results, and wrote the corresponding parts of the manuscript. AB, ST, and AH analyzed the data and wrote the draft manuscript. K-EJ and MF revised the manuscript. All authors read and approved the final manuscript.

FUNDING

The authors received funding from the European Union's Horizon 2020 Research and Innovation Program (Blue Growth: Unlocking the potential of Seas and Oceans) through the Project "INMARE" under grant agreement no. 634486. ST is financially supported by the Ministry of Culture and Science of the German State of North Rhine-Westphalia within in the framework of the NRW Strategieprojekt BioSC (No. 313/323-400-00213). MF acknowledges grant BIO2017-85522-R, from the Ministry of Science, Innovation and Universities, with the cofunds of the ERDF and the Agencia Estatal de Investigación (AEI). The Center for Structural Studies is funded by the Deutsche Forschungsgemeinschaft (DFG Grant number 417919780). The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data or in writing the manuscript.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the funding by the organizations mentioned in the Funding section. The authors thank the beamline scientists of the ID29 at ESRF (Grenoble, France) and at P13 at DESY (Hamburg, Germany) and Maha Raauf for excellent experimental support with PE-H mutagenesis. Computational support by the "Zentrum für Informations und Medientechnologie" at the Heinrich-Heine-Universität Düsseldorf and the computing time provided by the John von Neumann Institute for Computing (NIC) to HG on the supercomputer JUWELS at Jülich Supercomputing Centre (JSC) (user ID: HKF7) are gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.00114/full#supplementary-material

REFERENCES

Adrados, A., de Marco, I., Caballero, B. M., López, A., Laresgoiti, M. F., and Torres, A. (2012). Pyrolysis of plastic packaging waste: a comparison of plastic residuals from material recovery facilities with simulated plastic waste. *Waste Manag.* 32, 826–832. doi: 10.1016/j.wasman.2011.06.016

Almagro Armenteros, J. J., Tsirigos, K. D., Sønderby, C. K., Petersen, T. N., Winther, O., Brunak, S., et al. (2019). SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* 37, 420–423. doi: 10.1038/s41587-019-0036-z

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2

Austin, H. P., Allen, M. D., Donohoe, B. S., Rorrer, N. A., Kearns, F. L., Silveira, R. L., et al. (2018). Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proc. Natl. Acad. Sci. U.S.A.* 115, E4350–E4357. doi: 10.1073/pnas.1718804115

Barth, M., Oeser, T., Wei, R., Then, J., Schmidt, J., and Zimmermann, W. (2015). Effect of hydrolysis products on the enzymatic degradation of polyethylene terephthalate nanoparticles by a polyester hydrolase from *Thermobifida fusca*. *Biochem. Eng. J.* 93, 222–228. doi: 10.1016/J.BEJ.2014.10.012

Blakesley, R. W., and Boezi, J. A. (1977). A new staining technique for proteins in polyacrylamide gels using coomassie brilliant blue G250. *Anal. Biochem.* 82, 580–582. doi: 10.1016/0003-2697(77)90197-X

Bollinger, A., Thies, S., Katzke, N., and Jaeger, K.-E. (2020). The biotechnological potential of marine bacteria in the novel lineage of *Pseudomonas pertucinogena*. *Microb. Biotechnol.* 13, 19–31. doi: 10.1111/1751-7915.13288

Chen, C.-C., Han, X., Ko, T.-P., Liu, W., and Guo, R.-T. (2018). Structural studies reveal the molecular mechanism of PETase. *FEBS J.* 285, 3717–3723. doi: 10.1111/febs.14612

Chen, S., Su, L., Billig, S., Zimmermann, W., Chen, J., and Wu, J. (2010). Biochemical characterization of the cutinases from *Thermobifida fusca*. *J. Mol. Catal. B Enzym.* 63, 121–127. doi: 10.1016/j.molcatb.2010.01.001

Collaborative Computational Project. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. Sect. D* 50, 760–763. doi: 10.1107/S0907444994003112

Danso, D., Schmeisser, C., Chow, J., Zimmermann, W., Wei, R., and Leggewie, C. (2018). New insights into the function and global distribution of polyethylene terephthalate (PET)-degrading bacteria and enzymes in marine and terrestrial metagenomes. *Appl. Environ. Microbiol.* 84:e02773-17. doi: 10.1128/AEM.02773-17

Diedrich, D., Hamacher, A., Gertzen, C. G. W., Alves Avelar, L. A., Reiss, G. J., Kurz, T., et al. (2016). Rational design and diversity-oriented synthesis of peptoid-based selective HDAC6 inhibitors. *Chem. Commun.* 52, 3219–3222. doi: 10.1039/c5cc10301k

Dittrich, J., Schmidt, D., Pfleger, C., and Gohlke, H. (2019). Converging a knowledge-based scoring function: DrugScore^{2018}. *J. Chem. Inf. Model.* 59, 509–521. doi: 10.1021/acs.jcim.8b00582

Edelheit, O., Hanukoglu, A., and Hanukoglu, I. (2009). Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnol.* 9:61. doi: 10.1186/1472-6750-9-61

Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. Sect. D* 66, 486–501. doi: 10.1107/S0907444910007493

Fecker, T., Galaz-Davison, P., Engelberger, F., Narui, Y., Sotomayor, M., Parra, L. P., et al. (2018). Active site flexibility as a hallmark for efficient PET degradation by *I. sakaiensis* PETase. *Biophys. J.* 114, 1302–1312. doi: 10.1016/j.bpj.2018.02.005

Furukawa, M., Kawakami, N., Tomizawa, A., and Miyamoto, K. (2019). Efficient degradation of poly(ethylene terephthalate) with *Thermobifida fusca* cutinase exhibiting improved catalytic activity generated using mutagenesis and additive-based approaches. *Sci. Rep.* 9:16038. doi: 10.1038/s41598-019-52379-z

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., et al (2005). "Protein identification and analysis tools on the ExPASy server," in *The Proteomics Protocols Handbook*, ed. J. M. Walker, (Totowa, NJ: Humana Press), 571–607. doi: 10.1385/1-59259-890-0:571

Gomila, M., Mulet, M., Lalucat, J., and García-Valdés, E. (2017). Draft genome sequence of the marine bacterium *Pseudomonas aestusnigri* VGXO14T. *Genome Announc.* 5:e00765-17. doi: 10.1128/genomeA.00765-17

Goodsell, D. S., Morris, G. M., and Olson, A. J. (1996). Automated docking of flexible ligands: applications of AutoDock. *J. Mol. Recognit* 9, 1–5. doi: 10.1002/(sici)1099-1352(199601)9:1<1::aid-jmr241<3.0.co;2-6

Green, M. R., and Sambrook, J. (2012). *Molecular Cloning: A Laboratory Manual*, 4th Edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.

Haernvall, K., Zitzenbacher, S., Biundo, A., Yamamoto, M., Schick, M. B., Ribitsch, D., et al. (2018). Enzymes as enhancers for the biodegradation of synthetic polymers in wastewater. *ChemBioChem* 19, 317–325. doi: 10.1002/cbic.201700364

Haernvall, K., Zitzenbacher, S., Wallig, K., Yamamoto, M., Schick, M. B., Ribitsch, D., et al. (2017). Hydrolysis of ionic phthalic acid based polyesters by wastewater microorganisms and their enzymes. *Environ. Sci. Technol.* 51, 4596–4605. doi: 10.1021/acs.est.7b00062

Hajighasemi, M., Tchigvintsev, A., Nocek, B. P., Flick, R., Popovic, A., Hai, T., et al. (2018). Screening and characterization of novel polyesterases from environmental metagenomes with high hydrolytic activity against synthetic polyesters. *Environ. Sci. Technol.* 52, 12388–12401. doi: 10.1021/acs.est.8b04252

Han, X., Liu, W., Huang, J.-W., Ma, J., Zheng, Y., Ko, T.-P., et al. (2017). Structural insight into catalytic mechanism of PET hydrolase. *Nat. Commun.* 8:2106. doi: 10.1038/s41467-017-02255-z

Hanahan, D. (1983). Studies on transformation of *Escherichia coli* with plasmids. *J. Mol. Biol.* 166, 557–580.

Herrero Acero, E., Ribitsch, D., Steinkellner, G., Gruber, K., Greimel, K., Eiteljoerg, I., et al. (2011). Enzymatic surface hydrolysis of PET: effect of structural diversity on kinetic properties of cutinases from *Thermobifida*. *Macromolecules* 44, 4632–4640. doi: 10.1021/ma200949p

Joo, S., Cho, I. J., Seo, H., Son, H. F., Sagong, H.-Y., Shin, T. J., et al. (2018). Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nat. Commun.* 9:382. doi: 10.1038/s41467-018-02881-1

Kabsch, W. (2010). XDS. *Acta Crystallogr. Sect. D* 66, 125–132. doi: 10.1107/S0907444909047337

Kawai, F., Oda, M., Tamashiro, T., Waku, T., Tanaka, N., Yamamoto, M., et al. (2014). A novel Ca^{2+}-activated, thermostabilized polyesterase capable of hydrolyzing polyethylene terephthalate from *Saccharomonospora viridis* AHK190. *Appl. Microbiol. Biotechnol.* 98, 10053–10064. doi: 10.1007/s00253-014-5860-y

Krieger, V., Hamacher, A., Gertzen, C. G. W., Senger, J., Zwinderman, M. R. H., Marek, M., et al. (2017). Design, multicomponent synthesis, and anticancer activity of a focused histone deacetylase (HDAC) inhibitor library with peptoid-based cap groups. *J. Med. Chem.* 60, 5493–5506. doi: 10.1021/acs.jmedchem.7b00197

Laemmli, U. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685. doi: 10.1038/227680a0

Lebreton, L., Slat, B., Ferrari, F., Sainte-Rose, B., Aitken, J., Marthouse, R., et al. (2018). Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. *Sci. Rep.* 8:4666. doi: 10.1038/s41598-018-22939-w

Liu, B., He, L., Wang, L., Li, T., Li, C., Liu, H., et al. (2018). Protein crystallography and site-direct mutagenesis analysis of the poly(ethylene terephthalate) hydrolase PETase from *Ideonella sakaiensis*. *Chembiochem* 19, 1471–1475. doi: 10.1002/cbic.201800097

Liu, C., Shi, C., Zhu, S., Wei, R., and Yin, C.-C. (2018). Structural and functional characterization of polyethylene terephthalate hydrolase from Ideonella sakaiensis. *Biochem. Biophys. Res. Commun.* 508, 289–294. doi: 10.1016/j.bbrc.2018.11.148

Miyakawa, T., Mizushima, H., Ohtsuka, J., Oda, M., Kawai, F., and Tanokura, M. (2015). Structural basis for the Ca^{2+}-enhanced thermostability and activity of PET-degrading cutinase-like enzyme from *Saccharomonospora viridis* AHK190. *Appl. Microbiol. Biotechnol.* 99, 4297–4307. doi: 10.1007/s00253-014-6272-8

Moharir, R. V., and Kumar, S. (2019). Challenges associated with plastic waste disposal and allied microbial routes for its effective degradation: a comprehensive review. *J. Clean. Prod.* 208, 65–76. doi: 10.1016/j.jclepro.2018.10.059

Molitor, R., Bollinger, A., Kubicki, S., Loeschcke, A., Jaeger, K., and Thies, S. (2020). Agar plate-based screening methods for the identification of polyester

hydrolysis by *Pseudomonas* species. *Microb. Biotechnol.* 13, 274–284. doi: 10.1111/1751-7915.13418

Nikolaivits, E., Kanelli, M., Dimarogona, M., and Topakas, E. (2018). A middle-aged enzyme still in its prime: recent advances in the field of cutinases. *Catalysts* 8:612. doi: 10.3390/CATAL8120612

Nolasco-Soria, H., Moyano-López, F., Vega-Villasante, F., del Monte-Martínez, A., Espinosa-Chaurand, D., Gisbert, E., et al. (2018). "Lipase and phospholipase activity methods for marine organisms," in *Lipases and Phospholipases. Methods in Molecular Biology (Methods and Protocols)*, ed. G. Sandoval, (New York, NY: Humana Press), 139–167. doi: 10.1007/978-1-4939-8672-9_7

Numoto, N., Kamiya, N., Bekker, G.-J., Yamagami, Y., Inaba, S., Ishii, K., et al. (2018). Structural dynamics of the PET-degrading cutinase-like enzyme from *Saccharomonospora viridis* AHK190 in substrate-bound states elucidates the Ca^{2+}-driven catalytic cycle. *Biochemistry* 57, 5289–5300. doi: 10.1021/acs.biochem.8b00624

Palm, G. J., Reisky, L., Böttcher, D., Müller, H., Michels, E. A. P., Walczak, M. C., et al. (2019). Structure of the plastic-degrading *Ideonella sakaiensis* MHETase bound to a substrate. *Nat. Commun.* 10:1717. doi: 10.1038/s41467-019-09326-3

Peix, A., Ramírez-Bahena, M. H., and Velázquez, E. (2018). The current status on the taxonomy of *Pseudomonas* revisited: an update. *Infect. Genet. Evol.* 57, 106–116. doi: 10.1016/j.meegid.2017.10.026

PlasticsEurope. (2018). *Plastics – the Facts.* Brussels: PlasticsEurope. Available at: https://www.plasticseurope.org/application/files/6315/4510/9658/Plastics_the_facts_2018_AF_web.pdf

Popovic, A., Hai, T., Tchigvintsev, A., Hajighasemi, M., Nocek, B., Khusnutdinova, A. N., et al. (2017). Activity screening of environmental metagenomic libraries reveals novel carboxylesterase families. *Sci. Rep.* 7:44103. doi: 10.1038/srep44103

Ribitsch, D., Acero, E. H., Greimel, K., Eiteljoerg, I., Trotscha, E., Freddi, G., et al. (2012). Characterization of a new cutinase from *Thermobifida alba* for PET-surface hydrolysis. *Biocatal. Biotransformation* 30, 2–9. doi: 10.3109/10242422.2012.644435

Ribitsch, D., Hromic, A., Zitzenbacher, S., Zartl, B., Gamerith, C., Pellis, A., et al. (2017). Small cause, large effect: structural characterization of cutinases from *Thermobifida cellulosilytica*. *Biotechnol. Bioeng.* 114, 2481–2488. doi: 10.1002/bit.26372

Riegels, M., Koch, R., Pedersen, L., and Lund, H. (1997). Enzymatic Hydrolysis of Cyclic oligomers. Patent No. WO 97/27237. Geneva: World Intellectual Property Organization.

Rigden, D. J., and Jedrzejas, M. J. (2003). Structures of *Streptococcus pneumoniae* hyaluronate lyase in complex with chondroitin and chondroitin sulfate disaccharides. Insights into specificity and mechanism of action. *J. Biol. Chem.* 278, 50596–50606. doi: 10.1074/jbc.M307596200

Rochman, C. M. (2018). Microplastics research-from sink to source. *Science* 360, 28–29. doi: 10.1126/science.aar7734

Roth, C., Wei, R., Oeser, T., Then, J., Föllner, C., Zimmermann, W., et al. (2014). Structural and functional studies on a thermostable polyethylene terephthalate degrading hydrolase from *Thermobifida fusca*. *Appl. Microbiol. Biotechnol.* 98, 7815–7823. doi: 10.1007/s00253-014-5672-0

Schmidt-Rohr, K., Hu, W., and Zumbulyadis, N. (1998). Elucidation of the chain conformation in a glassy polyester, PET, by two-dimensional NMR. *Science* 280, 714–717. doi: 10.1126/science.280.5364.714

Sehnal, D., Svobodová Vařeková, R., Berka, K., Pravda, L., Navrátilová, V., Banáš, P., et al. (2013). MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J. Cheminform.* 5:39. doi: 10.1186/1758-2946-5-39

Seo, H., Cho, I. J., Joo, S., Son, H. F., Sagong, H., Choi, S. Y., et al. (2019). Reply to "Conformational fitting of a flexible oligomeric substrate does not explain the enzymatic PET degradation.". *Nat. Commun.* 10:5582. doi: 10.1038/s41467-019-13493-8

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7:539. doi: 10.1038/msb.2011.75

Son, H. F., Cho, I. J., Joo, S., Seo, H., Sagong, H.-Y., Choi, S. Y., et al. (2019). Rational protein engineering of thermo-stable PETase from *Ideonella sakaiensis* for highly efficient PET degradation. *ACS Catal.* 9, 3519–3526. doi: 10.1021/acscatal.9b00568

Sotriffer, C. A., Gohlke, H., and Klebe, G. (2002). Docking into knowledge-based potential fields: a comparative evaluation of DrugScore. *J. Med. Chem.* 45, 1967–1970. doi: 10.1021/jm025507u

Stern, R., and Jedrzejas, M. J. (2008). Carbohydrate polymers at the center of life's origins: the importance of molecular processivity. *Chem. Rev.* 108, 5061–5085. doi: 10.1021/cr078240l

Studier, F. W. (2005). Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.* 41, 207–234. doi: 10.1016/j.pep.2005.01.016

Studier, F. W., and Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *J. Mol. Biol.* 189, 113–130. doi: 10.1016/0022-2836(86)90385-2

Sulaiman, S., Yamato, S., Kanaya, E., Kim, J., Koga, Y., Takano, K., et al. (2012). Isolation of a novel cutinase homolog with polyethylene terephthalate-degrading activity from leaf-branch compost by using a metagenomic approach. *Appl. Environ. Microbiol.* 78, 1556–1562. doi: 10.1128/AEM.06725-11

Sulaiman, S., You, D.-J., Kanaya, E., Koga, Y., and Kanaya, S. (2014). Crystal structure and thermodynamic and kinetic stability of metagenome-derived LC-cutinase. *Biochemistry* 53, 1858–1869. doi: 10.1021/bi401561p

Taniguchi, I., Yoshida, S., Hiraga, K., Miyamoto, K., Kimura, Y., and Oda, K. (2019). Biodegradation of PET: current status and application aspects. *ACS Catal.* 9, 4089–4105. doi: 10.1021/acscatal.8b05171

Wei, R., Breite, D., Song, C., Gräsing, D., Ploss, T., Hille, P., et al. (2019a). Biocatalytic degradation efficiency of postconsumer polyethylene terephthalate packaging determined by their polymer microstructures. *Adv. Sci.* 6:1900491. doi: 10.1002/advs.201900491

Wei, R., Oeser, T., Schmidt, J., Meier, R., Barth, M., Then, J., et al. (2016). Engineered bacterial polyester hydrolases efficiently degrade polyethylene terephthalate due to relieved product inhibition. *Biotechnol. Bioeng.* 113, 1658–1665. doi: 10.1002/bit.25941

Wei, R., Song, C., Gräsing, D., Schneider, T., Bielytskyi, P., Böttcher, D., et al. (2019b). Conformational fitting of a flexible oligomeric substrate does not explain the enzymatic PET degradation. *Nat. Commun.* 10:5581. doi: 10.1038/s41467-019-13492-9

Wei, R., and Zimmermann, W. (2017). Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: how far are we? *Microb. Biotechnol.* 10, 1308–1322. doi: 10.1111/1751-7915.12710

Wheeler, D. L., Church, D. M., Federhen, S., Lash, A. E., Madden, T. L., Pontius, J. U., et al. (2003). Database resources of the National Center for Biotechnology. *Nucleic Acids Res.* 31, 28–33. doi: 10.1093/nar/gkv1290

Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., et al. (1989). Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucleic Acids Res.* 17, 3469–3478. doi: 10.1093/nar/17.9.3469

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., et al. (2016). A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* 351, 1196–1199. doi: 10.1126/science.aad6359

Conflict of Interest: HG and K-EJ are employed by Forschungszentrum Jülich GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers

Supplementary Material

```
Type  Name     Alignment                                                  Position
      LCC      ------------------------------------------------SNPYQRGPNPT  11
I     Cut190   ----------------------------------------MRGSHHHHHHGSNPYERGPDPT  22
      TfCut2   ------------------------------------------------ANPYERGPNPT  11
IIb   PETase   --MNFPRASRLMQAAVLGGLMAVSA-------------------AATAQTNPYARGPNPT  39
      PET12    MPPDCVLPRRLAAAALLASATLVPL-------------------SAAAQTNPYQRGPDPT  41
IIa   PET5     --MN---KSILKKLSFGTSVLLVSMNALSWTPSPTPNPDPIPDPTPCQDDCDFTRGPNPT  55
      PE-H     MPFNK--KSVLA--LCGAGALLFSMSALANNPAP-------TDPGDSGGGSAYQRGPDPS  49

      LCC      RSALT-ADGPFSVATYTVSRLSVSGFGGGVIYYPTGTS-LTFGGIAMSPGYTADASSLAW  69
I     Cut190   EDSIEAIRGPFSVATERVSSF-ASGFGGGTIYYPRETDEGTFGAVAVAPGFTASQGSMSW  81
      TfCut2   DALLEARSGPFSVSEENVSRLSASGFGGGTIYYPREN--NTYGAVAISPGYTGTEASIAW  69
IIb   PETase   AASLEASAGPFTVRSFTVS-R-PSGYGAGTVYYPTNAG-GTVGAIAIVPGYTARQSSIKW  96
      PET12    TRDLEDSRGPFRYASTNVR-S-PSGYGAGTIYYPTDVS-GSVGAVAVVPGYLARQSSIRW  98
IIa   PET5     PSSLEASTGPYSVATRSVASS-VSGFGGGTLHYPTNTT-GTMGAIAVVPGFLLQESSIDF  113
      PE-H     VSFLEADRGQYSVRSSRVSSL-VSGFGGGTIYYPTGTT-GTMGAVVVIPGFVSAESSIDW  107

      LCC      LGRRLASHGFVVLVINTNSRFDYPDSRASQLSAALNYLRT---SSPSAVRARLDANRLAV  126
I     Cut190   YGERVASQGFIVFTIDTNTRLDQPGQRGRQLLAALDYLVE---RSDRKVRERLDPNRLAV  138
      TfCut2   LGERIASHGFVVITIDTITTLDQPDSRAEQLNAALNHMIN---RASSTVRSRIDSSRLAV  126
IIb   PETase   WGPRLASHGFVVITIDTNSTLDQPSSRSSQQMAALRQVASLNGTSSSPIYGKVDTARMGV  156
      PET12    WGPRLASHGFVVITLDTRSTSDQPASRSAQQMAALRQVVALSETRSSPIYGKVDPNRLAV  158
IIa   PET5     WGPKLASHGFVVITISANSGFDQPASRATQLGRALDYVINQSNGSNSPISGMVDTTRLGV  173
      PE-H     WGPKLASYGFVVMTIDTNTGFDQPPSRARQINNALDYLVSQNSRSSSPVRGMIDTNRLGV  167

      LCC      AGHSMGGGGTLRIAEQNPSLKAAVPLTPWHTDKT-FN-TSVPVLIVGAEADTVAPVSQHA  184
I     Cut190   MGHSMGGGGSLEATVMRPSLKASIPLTPWNLDKT-WGQVQVPTFIIGAELDTIAPVRTHA  197
      TfCut2   MGHSMGGGGSLRLASQRPDLKAAIPLTPWHLNKN-WSSVTVPTLIIGADLDTIAPVATHA  185
IIb   PETase   MGWSMGGGGSLISAANNPSLKAAAPQAPWDSSTN-FSSVTVPTLIFACENDSIAPVNSSA  215
      PET12    MGWSMGGGGTLISARDNPSLKAAVPFAPWHNTAN-FSGVQVPTLVIACENDTVAPISRHA  217
IIa   PET5     VGWSMGGGGALQLAS-GDRLSAAIPIAPWNQGGNRFDQIETPTLVIACENDVVASVNSHA  232
      PE-H     IGWSMGGGGTLRVAS-EGRIKAAIPLAPWDTTSYYASRSQAPTLIFACESDVIAPVLQHA  226

      LCC      IPFYQNLPSTTPKVYVELDNASHFAPN---SNNAAISVYTISWMKLWVDNDTRYRQFLCN  241
I     Cut190   KPFYESLPSSLPKAYMELDGATHFAPN---IPNTTIAKYVISWLKRFVDEDTRYSQFLCP  254
      TfCut2   KPFYNSLPSSISKAYLELDGATHFAPN---IPNKIIGKYSVAWLKRFVDNDTRYTQFLCP  242
IIb   PETase   LPIYDSMSR-NAKQFLEINGGSHSCANSGNSNQALIGKKGVAWMKRFMDNDTRYSTFACE  274
      PET12    SSFYNSFSSSLAKAYLEINNGSHTCANTGNSNQALIGKYGVAWIKRFVDNDTRYSPFLCG  277
IIa   PET5     SPFYNRIPSTTDKAYLEINGGSHFCANDGGSIGGLLGKYGVSWMKRFIDNDLRYDAFLCG  292
      PE-H     SPFYNSLPSSIDKAFVEINGGSHYCGNGGSIYNDVLSRFGVSWMKLHLDEDSRYKQFLCG  286

      LCC      VNDPALS---DFRTNNRHCQ----------------------                   258
I     Cut190   NPTDR-----AIEEYRSTCPYKLN------------------                   273
      TfCut2   GPRDGLF--GEVEEYRSTCPFYPNSSSVDKLAAALEHHHHHH                   282
IIb   PETase   NPNSTRVSDFRTA----NCSLE-------------HHHHHH-                   298
      PET12    APHQADLRSSRLSEYRESCPY---------------------                   298
IIa   PET5     PDHAANR---SVSEYRDTCNY---------------------                   310
      PE-H     PNHTSDS---QISDYRGNCPYL-------------EHHHHHH                   312
```

Supplementary Material

Figure S1 Multiple sequence alignment of PE-H with amino acid sequences of different cutinases and PET hydrolases using the program Clustal Omega. The enzymes were assigned to different types of polyester hydrolases. Amino acid residues of the catalytic triad are marked by a red triangle, disulfide forming cysteine residues are highlighted in orange and connected by an orange line. Amino acids of the extended loop region, specific for type II PET hydrolases, are framed in red. Abbreviations are: leaf-branch compost metagenome cutinase (LCC); *Saccharomonospora viridis* cutinase (Cut190); *Thermobifida fusca* cutinase (TfCut2); *Ideonella sakaiensis* PET hydrolase (PETase); *Polyangium brachysporum* PET hydrolase (PET12); *Oleispira antarctica* PET hydrolase (PET5).

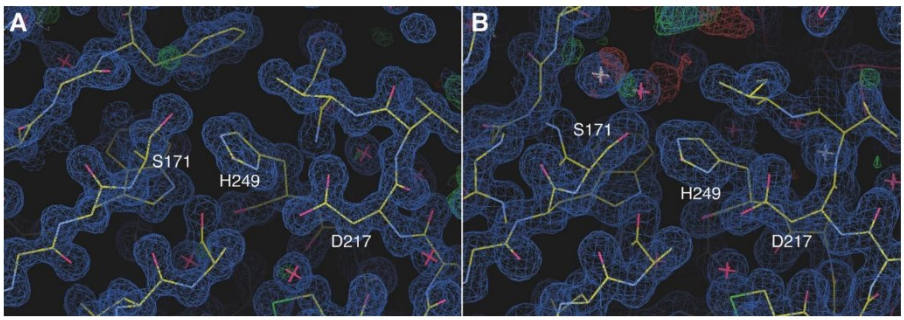


Figure S2 Electron densities around the active site of **(A)** PE-H WT (PDB code 6SBN) and **(B)** PE-H Y250S (PDB code 6SCD) with residues of the catalytic triad highlighted with labels. 2FoFc contoured at 1 σ.

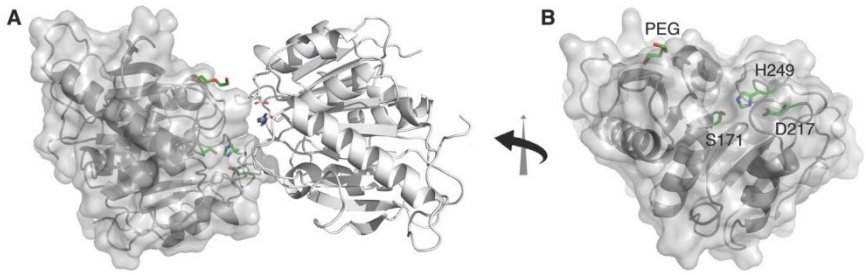
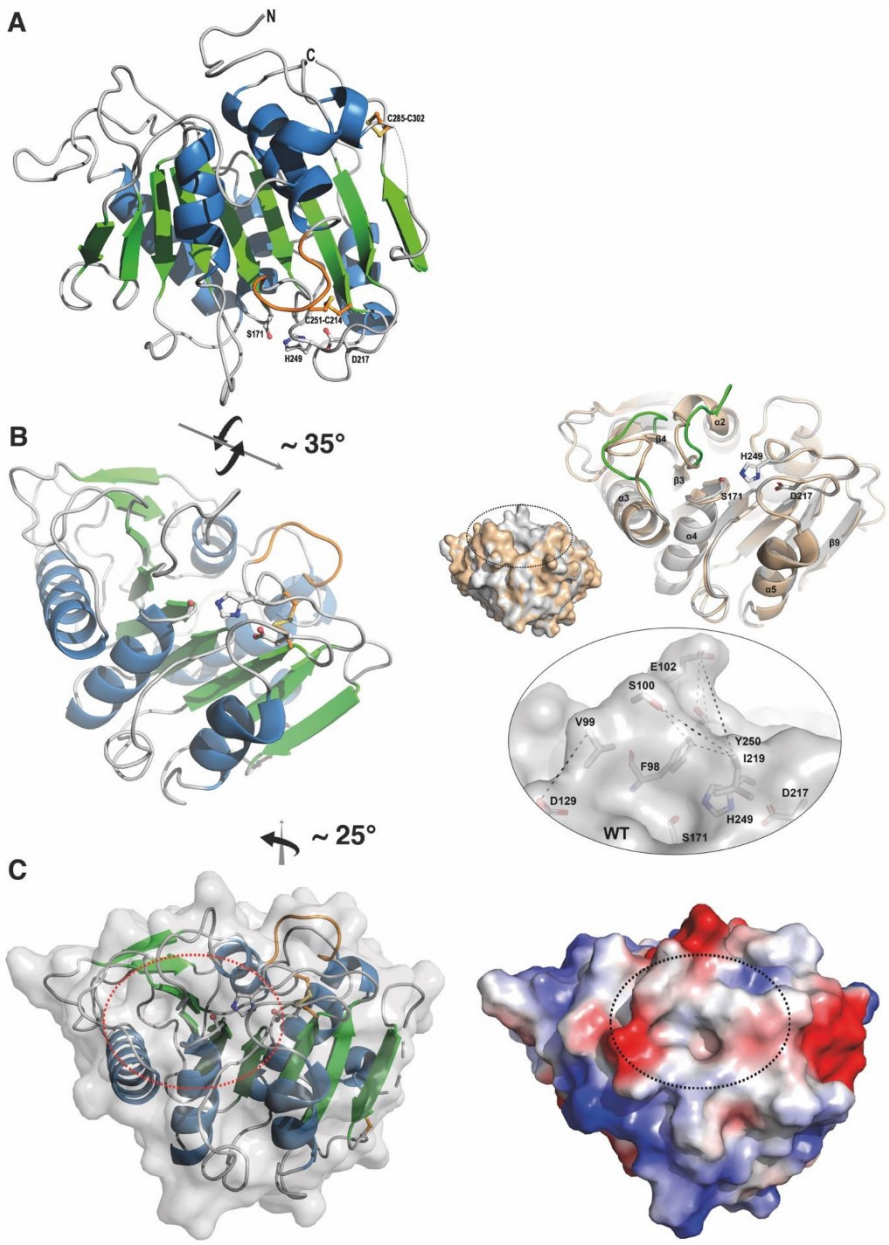


Figure S3 (**A**) Asymmetric unit containing two molecules of PE-H Y250S (PDB code 6SCD) depicted as cartoon showing the monomer with bound PEG molecule also in surface representation. In (**B**) the left molecule is rotated counter-clockwise with the residues of the catalytic triad as well as the PEG molecule shown as green sticks.

2


A
N
C
C285-C302
C251-C214
S171
H249
D217
B
~ 35°
α2
β4
H249
β3
α3
S171
D217
α4
β9
α5
E102
S100
V99
Y250
I219
F98
D217
D129
H249
WT
S171
~ 25°
C

3

Figure S4 (**A**) Orientation of PE-H as reference to Figure 4. (**B**) PE-H is rotated for around 35° from bottom to top to zoom in the active site (as in Figure 5). In (**C**) PE-H is additionally rotated counter-clockwise for around 25° for a better view on the active site cleft (as in Figure S5).

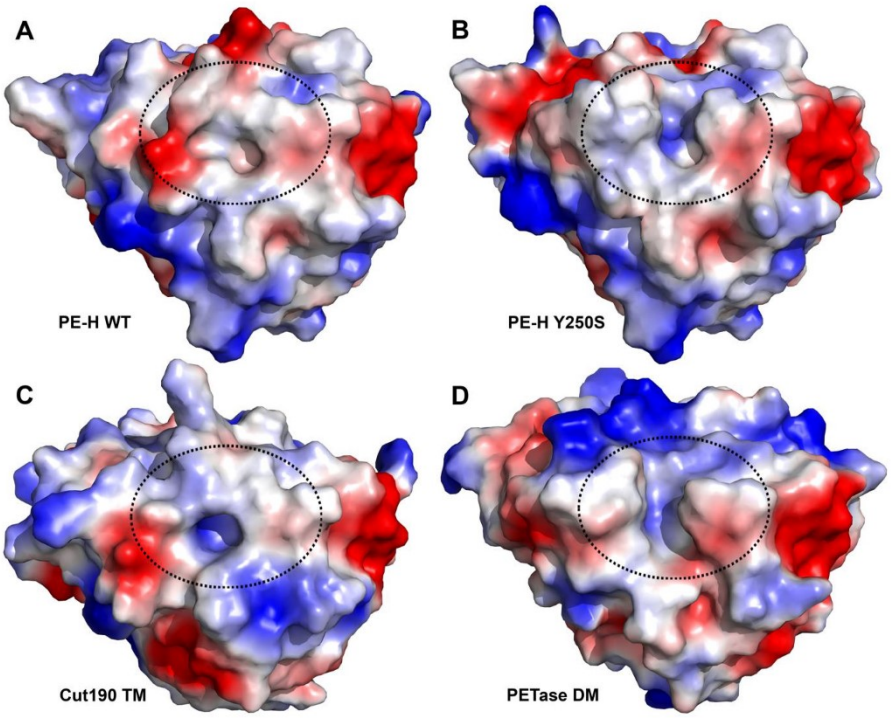


Figure S5 Surface representation of PE-H variants and two structural homologues. **(A)** shows wild type PE-H (PDB code 6SBN), **(B)** PE-H Y250S (PDB code 6SCD), **(C)** Cut190 triple mutant (TM) from *S. viridis* (PDB code 5ZRR) and **(D)** PETase double mutant (DM) from *I. sakaiensis* (PDB code 5XH3). Color code of the electrostatic surface is blue for positive, red for negative charge. All molecules are shown in identical orientation with the active site cleft positioned in the middle and marked by a dashed line.

Table S1 Amino acid composition of PETase (6EQE) and PE-H (6SBN) at selected positions as marked in the main text (i to iii).

	active site			(i)	(ii)	(iii)					
PETase	S160	D206	H237	S238	W159	S242	G243	N244	S245	N246	Q247
PE-H	S171	D217	H249	Y250	W170	G254	G255	S256	I257	Y258	N259

Table S2 Thermal melting points (Tm) of PE-H (WT) and different variants determined by nano differential scanning fluorimetry. The top row shows the respective single amino acid substitutions. Mutations G254S, S256N, I257S, Y258N, and N259Q were combined (ext.) to resemble the amino acid combination of the extended loop region of PETase.

	WT	G254S	S256N	I257S	Y258N	N259Q	ext.	Y250S
Tm [°C]	50.8	39.7	49.5	48.5	43.8	45.2	39.4	49.8

Table S3 Data collection and refinement statistics of PE-H

	PE-H	**PE-H Y250S**
Beamline	ID29, ESRF, Grenoble	P13, DESY, Hamburg
Crystal parameters		
Space group	C 2 2 21	I 21 21 21
Unit cell parameters:		
a, b, c (Å)	68.81, 80.01, 88.91	94.53, 98.27, 121.23
α, β, γ (°)	90, 90, 90	90, 90, 90
Data collection and processing		
Wavelength	0.9762	0.9505
Resolution (Å)	50.00 – 1.09 (1.15 – 1.09)	77.34 – 1.35 (1.42-1.35)
Total reflections	1293506 (198125)	494296 (78069)
Unique reflections	101774 (16134)	122029 (19634)
Multiplicity	12.7 (12.3)	4.1 (4.0)
Completeness (%)	99.5 (98.4)	99.11 (99.26)
I/σ(I)	16.20 (4.70)	14.98 (2.52)
Wilson B-factor (Å^2)	10.4	16.7
R-merge	0.076 (0.429)	0.043 (0.439)
R-meas	0.080 (0.448)	0.050 (0.505)
CC 1/2	0.999 (0.970)	0.998 (0.849)
Refinement statistics		
Reflections used in refinement	99673	116011
Reflections used for R-free	2057	5715
R work (%)	10.69	11.3
R free (%)	13.70	15.2
Number of non-hydrogen atoms	2401	4590
macromolecules	2028	4048
ligands	5	49
solvent	368	493
RMS (bonds)	0.027	0.026
RMS (angles)	2.206	2.430
Ramachandran plot:		
favoured (%)	96.5	97.3
Allowed (%)	2.7	2.5
outliers (%)	0.8	0.2
Average B-factor (Å^2)	14.0	22.0
Model content		
Monomers ASU	1	2
Protein residues	263 (38-285; 292-306)	532 (40-305)
Ligand	1 ACT, 1 NA	1 SO4, 3 PO4, 1 ACT, 1 PEG, 2 GOL, 3 CL, 3 NA
Waters	368	495
PDB code	6SBN	6SCD

(Statistics for the highest resolution shell are shown in parentheses)

Table S4: Structure based alignment of wild type PE-H and PE-H Y250S against pdb entries. Top 10 of the most similar structures are given with their corresponding alignment quality (Q-score), root mean square deviation (RMSD), number of aligned residues (N_{algn}), their sequence identity (Seq-%), their PDB number (Target PDB), and description of the PDB entry (Protein).

Wild type PE-H (pdb-code: 6SBN)						
##	**Q-score**	**RMSD**	**N_{algn}**	**Seq-%**	**Target PDB**	**Protein**
1	0.76	1.18	246	51	5zrr:A	**cutinase** Cut190 S176A/S226P/R228S mutant (*Saccharomonospora viridis*)
2	0.75	1.20	245	51	5xg0:B	*1* **PET hydrolase** *(Ideonella sakaiensis)*
3	0.75	1.18	244	51	6ane:A	*2* **PET hydrolase** *(Ideonella sakaiensis)*
4	0.75	1.24	245	51	4cg2:A	*3* **PET hydrolase** *(Thermobifida fusca)*
5	0.75	1.25	245	51	4cg1:A	*4* **PET hydrolase** *(Thermobifida fusca)*
6	0.74	1.22	240	53	5luj:A	*5* **cutinase 2** *(Thermobifida cellulosilytica)*
7	0.74	1.25	244	52	5lul:A	*6* **cutinase 2** R19S/R29N/A30V mutant *(Thermobifida cellulosilytica)*
8	0.74	1.32	246	52	5luk:A	*7* **cutinase 2** R29N/A30V mutant *(Thermobifida cellulosilytica)*
9	0.73	1.28	245	51	5lui:A	**cutinase 1** (*Thermobifida cellulosilytica*)
10	0.73	1.25	242	52	5zoa:A	*8* **cutinase** *(Thermobifida fusca)*
PE-H Y_{250}S chain A (pdb-code 6SCD)						
1	0.84	1.17	259	51	5xh3:A	**9** **PET hydrolase** R103G/S131A mutant (HEMT complex) *(Ideonella sakaiensis)*
2	0.84	1.17	259	51	5xfz:A	**10** **PET hydrolase** R103G/S131A mutant *(Ideonella sakaiensis)*
3	0.83	1.11	257	51	6qgc:A	**11** **PETase** *(Ideonella sakaiensis)*
4	0.83	1.14	257	51	5xfy:A	**12** **PET hydrolase** S131A mutant *(Ideonella sakaiensis)*
5	0.83	1.16	257	51	5xh2:A	**13** **PET hydrolase** R103G/S131A mutant (pNP complex) *(Ideonella sakaiensis)*
6	0.83	1.10	256	52	6eqg:C	**14** **PET hydrolase** *(Ideonella sakaiensis)*
7	0.83	1.08	253	52	4wfk:A	**15** **cutinase** Cut190 S226P mutant (Ca(2+)-bound state) *(Saccharomonospora viridis)*
8	0.83	1.08	253	52	4wfj:A	**16** **cutinase** Cut190 S226P mutant (Ca(2+)-bound state) *(Saccharomonospora viridis)*
9	0.83	1.08	255	52	6eqg:B	**17** **PET hydrolase** *(Ideonella sakaiensis)*
10	0.83	1.07	255	52	6eqh:A	**18** **PET hydrolase** *(Ideonella sakaiensis)*

3 General Discussion

In this thesis, novel carboxylic ester hydrolases from marine hydrocarbonoclastic bacteria, namely *Alcanivorax borkumensis* and *Pseudomonas aestusnigri*, were identified and characterized with emphasis on important criteria for industrial applications: a broad substrate spectrum, tolerance for polar organic solvents, and the hydrolysis of synthetic polyesters. The identification of candidate enzymes was facilitated by respective screening strategies established in this thesis.

First, the chance to identify relevant carboxylic ester hydrolases from the named bacteria was evaluated (chapter 2.1 & 2.2). Then, activity- and sequence-based screening for novel CEHs of respective organisms was conducted and substrate promiscuity was systematically investigated together with numerous CEHs from diverse microorganisms (chapter 2.3 & 2.4). For the identification of organic solvent tolerant CEHs, a novel screening strategy was developed and applied to yield CEHs active in the presence of high concentrations of polar organic solvents (chapter 2.5). With regard to the bacteria's potential for synthetic polyester degradation, screening for polyester hydrolases was applied, the responsible biocatalyst was identified, its performance was increased by protein engineering, and its structure-function relationship was studied (chapter 2.6 & 2.7).

In the following chapter, the investigated characteristics of the novel CEHs will be summed up. The gained insights with regard to the key characteristics organic solvent tolerance, substrate promiscuity and polyester hydrolysis will be discussed based on outstanding examples from the newly identified CEHs (chapter 3.1). The design of a hypothetical Swiss-army knife biocatalyst, meaning the combination of multiple industrial demanded features within one single enzyme, is envisaged and discussed in the light of the findings of this thesis and the literature (chapter 3.2). Lastly, open questions that arose from this work and the current state of science are outlined to guide future studies (chapter 3.3).

3.1 Novel CEHs from *A. borkumensis* and *P. aestusnigri* – general overview and outstanding biocatalysts

Today, much effort is put in the transition of traditional chemical processes toward more sustainable technologies. Biotechnology can fulfill these needs by applying enzymes as green biocatalysts, thus reducing toxic byproducts, the waste of energy, and the use of fossil resources (Sheldon and Woodley, 2018). To facilitate industrial application, biocatalysts must fulfill industrial demands, in particular stability in the presence of organic solvents and reactivity with the desired substrate (Ferrer, Bargiela, *et al.*, 2015). To find novel CEHs, matching this requirements in the best case, the genomes of the marine hydrocarbonoclastic bacteria *A. borkumensis* and *P. aestusnigri* were screened sequence- and function-based (Figure 3-1). Therefore, genomic libraries were constructed and screened in the heterologous host *E. coli* for esterase activity, to identify CEHs that are actively expressed in this established protein production host. Complementary, the complete or draft genome sequence of both bacteria was searched for putative CEHs. CEHs not covered by the function-based screening were then cloned into expression vectors, checked for activity and used for further characterization. In summary, for *A. borkumensis* 11, and for *P. aestusnigri* 14 CEHs were identified and characterized in this thesis (Table 3-1, see appendix for gene (chapter 5.2) and protein (chapter 5.3) sequences). For the nearly unexplored *P. aestusnigri*, the complete set of CEHs was previously undescribed. Although intensively studied for years as a hydrocarbonoclastic model organism, only 2 of 11 CEHs of *A. borkumensis* identified here were characterized before, ABO1251 (Tchigvintsev *et al.*, 2015) and ABO2449 (Hajighasemi *et al.*, 2016). Four CEHs from *A. borkumensis* with proven activity, ABO1197 (Tchigvintsev *et al.*, 2015), ABO0116, ABO1483, ABO1895 (Hajighasemi *et al.*, 2018), were not covered by the screening strategy applied here.

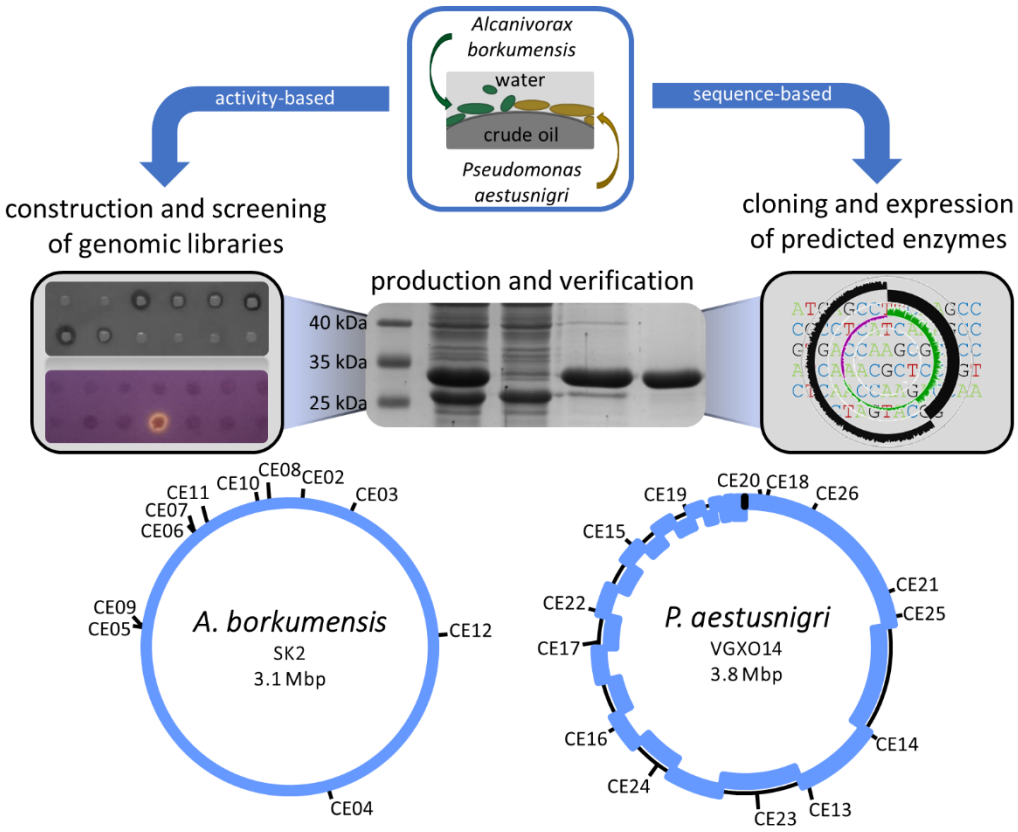


Figure 3-1 Schematic representation of the screening for novel CEHs from *A. borkumensis* and *P. aestusnigri*. The genomes of both bacteria were used to construct genomic libraries, followed by agar plate-based screening, which was complemented by sequence-based identification, cloning and production of CEHs. Novel CEHs identified in this thesis were marked on the circular representation of the whole genome of *A. borkumensis* SK2 (Schneiker *et al.*, 2006) and the draft genome of *P. aestusnigri* VGXO14 (Gomila *et al.*, 2017). Each contig of the draft genome sequence is represented by a blue box.

Table 3-1 Overview of organic solvent stability and substrate promiscuity of carboxylic ester hydrolases from *A. borkumensis* and *P. aestusnigri* characterized in this study. The enzyme identifier is given according to chapter 2.3 (ch. 2.3) and 2.5 (ch. 2.5). The accession number of the respective biocatalyst (acc. no.) is given along with information on its isolation (origin) and classification to the family of bacterial lipolytic enzymes (family) (Arpigny and Jaeger, 1999; Kovacic *et al.*, 2019); “n. d.” means not determined. The number of hydrolyzed substrates out of 96 as a marker for substrate promiscuity (promiscuity) is shown according to chapter 2.3 and 2.5; polyester hydrolytic activity is indicated by a plus “(+)”. The organic solvent tolerance (OST) is indicated by “+” (tolerant), “+++” (highly tolerant), or “o” (low tolerance), according to chapter 2.5. The reference for the first characterization of each enzyme is shown; publications that are part of this thesis are marked by an asterisk “*”.

origin	ch. 2.5	ch. 2.3	acc. no.	family	promiscuity	OST	reference
A. borkumensis SK2	CE02	EH8	WP_011587341.1	IV	63	+	(Martínez-Martínez *et al.*, 2018)*
	CE03	EH39	WP_011587492.1	IV	25	+	(Martínez-Martínez *et al.*, 2018)*
	CE04	EH93	WP_011588534.1	VII	12	+	(Tchigvintsev *et al.*, 2015)
	CE05	EH106	WP_011589376.1	-	9	+	(Martínez-Martínez *et al.*, 2018)*
	CE06	EH83	WP_011589723.1	V	14	+	(Hajighasemi *et al.*, 2016)
	CE07	-	WP_011589728.1	IV	65	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE08	EH117	WP_011589970.1	II	6	+	(Martínez-Martínez *et al.*, 2018)*
	CE09	-	WP_011589386.1	IV	9	o	(Bollinger, Molitor, *et al.*, 2020)*
	CE10	EH118	WP_011589935.1	I	6	+	(Martínez-Martínez *et al.*, 2018)*
	CE11	EH125	WP_011589767.1	II	4	o	(Martínez-Martínez *et al.*, 2018)*
	CE12	EH92	WP_011587953.1	V	12	o	(Martínez-Martínez *et al.*, 2018)*
P. aestusnigri VGXO14	CE13	-	WP_088275369.1	VII	51	+++	(Bollinger, Molitor, *et al.*, 2020)*
	CE14	-	WP_088277870.1	VIII	51	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE15	-	WP_088277153.1	IV	34	o	(Bollinger, Molitor, *et al.*, 2020)*
	CE16	-	WP_088276085.1	III	31(+)	+	(Bollinger, Thies, Knieps-Grünhagen, *et al.*, 2020)*
	CE17	-	WP_088276582.1	-	-	o	(Bollinger, Molitor, *et al.*, 2020)*
	CE18	-	WP_088273225.1	-	-	o	(Bollinger, Molitor, *et al.*, 2020)*
	CE19	-	WP_088277509.1	-	-	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE20	-	WP_088273217.1	VI	-	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE21	-	WP_088273788.1	VIII	22	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE22	-	SEG59772.1	X	12	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE23	-	WP_088274564.1	VII	9	o	(Bollinger, Molitor, *et al.*, 2020)*
	CE24	-	WP_088275865.1	XV	7	+	(Bollinger, Molitor, *et al.*, 2020)*
	CE25	-	WP_088273867.1	-	1	o	(Bollinger, Molitor, *et al.*, 2020)*
	CE26	-	n. d.	n. d.	2	o	(Bollinger, Molitor, *et al.*, 2020)*

Since both bacteria investigated in this thesis were isolated from marine crude oil-associated environments, the observed occurrence of substrate promiscuous and organic solvent tolerant CEHs might be connected to their lifestyle. The genome of both organisms is small compared to their near relatives, *A. borkumensis* features a 3.1 Mbp genome whereas *Alcanivorax* sp. often show a genome size around 4 Mbp. Likewise, *P. aestusnigri* comprises a genome of 3.8 Mbp, while other *Pseudomonas* species are often reported with a genome size of around 6 Mbp (Table 3-2). The high number of active CEHs identified in this study account for about 0.4 % of all the proteins coded in the genome of each bacterium. If taken previously reported active CEHs for *A. borkumensis* into account, 0.55 % of all genes from this bacterium codes for CEHs (chapter 2.1; (Coscolín, Bargiela, *et al.*, 2018)). Thus, both organisms are prolific sources for CEHs. In general, esterases and lipases are frequently found in many organisms but at frequencies of 0.05 % to 0.35 % only (Ferrer, Martínez-Martínez, *et al.*, 2015). Thus, an impact of the environment on the probability of finding CEHs can be supposed. This is furthermore supported by a higher number of enzymes identified from polluted over pristine environments (see chapter 2.3, Figure S1; (Martínez-Martínez *et al.*, 2018)). The higher percentage of CEHs found for oil dwellers like *A. borkumensis* and *P. aestusnigri* might reflect an adaptation to crude oil-polluted sites. However, a connection between crude oil-contamination and a high number of CEH genes is not obvious, because the degradation of crude oil alkanes relies on alkane oxidative enzymes in the first place, but not on CEHs. However, CEHs might participate here via the subterminal alkane degradation pathway (Beilen *et al.*, 2003; Rojo, 2009; Ji *et al.*, 2013) (chapter 1.2, Figure 1-2). The subterminal degradation of alkanes comprises the formation of an ester bond by the consecutive action of an alkane hydroxylase, an alcohol dehydrogenase, and a Baeyer-Villiger monooxygenase, followed by the cleavage of the ester by a CEH to produce an alcohol and a fatty acid. The fatty acid can be directly metabolized via the β-oxidation pathway and the alcohol can be modified via the terminal alkane oxidation pathway to be routed to the β-oxidation pathway.

Table 3-2 Comparison of the genome size among different representatives of *Alcanivorax* and *Pseudomonas* species. The species and strain names are given along with their reported genome size and the NCBI accession number for the respective reference sequence. The final row shows the average (Ø) genome size along with the standard deviation (±) among either *Alcanivorax* sp. (left) or *Pseudomonas* sp. (right) of the strains mentioned in the table.

***Alcanivorax* sp.**	**genome size**	***Pseudomonas* sp.**	**genome size**
A. borkumensis SK2	3.1 Mbp NC_008260.1	*P. aestusnigri* VGXO14	3.8 Mbp NZ_NBYK00000000.1
A. jadensis T9	3.6 Mbp NZ_ARXU00000000.1	*P. aeruginosa* PAO1	6.3 Mbp NC_002516.2
A. hongdengensis A-11-3	3.7 Mbp NZ_AMRJ00000000.1	*P. putida* KT2440	6.2 Mbp NC_002947.4
A. pacificus W11-5	4.2 Mbp NZ_CP004387.1	*P. fluorescens* F113	6.8 Mbp NC_016830.1
A. dieselolei B5	4.9 Mbp NC_018691.1	*P. protegens* CHA0	6.9 Mbp NC_021237.1
A. xenomutans P40	4.7 Mbp NZ_CP012331.1	*P. syringae pv. syringae* B728a	6.1 Mbp NC_007005.1
A. profundi MTEO17	3.7 Mbp NZ_QYYA00000000.1	*P. stutzeri* CGMCC 1.1803	4.5 Mbp NC_015740.1
A. gelatiniphagus MEBiC 08158	4.2 Mbp NZ_VCQT00000000.1	*P. oleovorans* DSM 1045	4.9 Mbp NZ_BDAL00000000.1
A. indicus SW127	3.4 Mbp NZ_QGMP00000000.1	*P. luteola* NBRC 103146	5.4 Mbp NZ_BDAE00000000.1
A. mobilis MT13131	4.1 Mbp NZ_NMQZ00000000.1	*P. straminea* JCM 2783	5.0 Mbp NZ_FOMO00000000.1
A. nanhaiticus 19-m-6	4.1 Mbp NZ_ARXV00000000.1	*P. lutea* DSM 17257	5.7 Mbp NZ_JRMB00000000.1
Ø	4.0 (± 0.5) Mbp	Ø	5.6 (± 0.9) Mbp

A second idea is based on the consideration, that crude oil is a mixture of hydrocarbons also including esters by nature (Marshall and Rodgers, 2004). Moreover, crude oil constituents greatly differ in their number of atoms, double bonds, and elements like nitrogen, sulfur, and oxygen (Marshall and Rodgers, 2004). The complexity of crude oil, which constitutes over 17,000 different compounds and varies significantly in its composition depending on the source of the oil (Marshall and Rodgers, 2004; Head *et al.*, 2006), can thus be a driving force for the development of biocatalysts with a broad substrate range or a large set of degradative enzymes, to detoxify cell-harming ingredients and exploit the wealth of compounds as nutrient source.

This might be illustrated by a recent study, reporting examples for both evolutionary strategies (Wright *et al.*, 2020). Here, the metabolization of plasticizers, chemicals that are used in plastic manufacturing which constitute ester comprising aromatic or

aliphatic hydrocarbons in the first place, by two marine microbes was investigated. The combined analysis of the metabolome, genome, and proteome of the two bacterial isolates showed the first step of the catabolic pathway for dibutyl phthalate and acetyl tributyl citrate are catalyzed by at least three CEHs in *Mycobacterium* sp. DBP42 or by a single CEH in *Halomonas* sp. ATBC28. Therefore, the authors proposed, the two bacteria might display different evolutionary strategies to cope with the uncommon plasticizer substrates; *Mycobacterium* sp. DBP42 uses a range of different CEHs whereas *Halomonas* sp. ATBC28 employs a promiscuous enzyme.

The occurrence of organic solvent tolerant CEHs could be related to the crude oil-associated lifestyle too. As crude oil consists of alkanes, which are organic solvents by nature (chapter 1.4.2, Figure 1-6), enzymes produced by hydrocarbonoclastic bacteria have a high chance to come into contact with these compounds. The influence of an organism's environment may directly favor the development of organic solvent tolerant enzymes, speaking in particular of extracellularly located enzymes. This was for example reported for the identification of an organic solvent tolerant lipase (Ogino *et al.*, 2000) from a likewise organic solvent tolerant bacterium (Ogino *et al.*, 1994). For *A. borkumensis*, two protein secretion machineries have been identified in the genome sequence, the type II and type IV secretion systems. In addition, the authors report on five genes, which encode proteins similar to the HlyD family, suggesting the ability for type I secretion (Schneiker *et al.*, 2006). For *P. aestusnigri*, the draft genome sequence revealed genes for a type VI secretion system (Gomila *et al.*, 2017) as well as the general secretion pathway (gsp) and type II secretion system. Hence, both bacteria are likely to secrete proteins into the extracellular space.

Out of the enzymes investigated in this thesis, six were predicted to entail signal sequences for Sec-dependent translocation to the periplasmic space (Table 3-3), all originating from *P. aestusnigri*. However, predicted secretion did not correlate with the observation of organic solvent tolerance for most CEHs in this thesis. In case of the highly organic solvent tolerant enzyme CE13, secretion via the Sec-secretion pathway was predicted by the occurrence of a characteristic N-terminal signal peptide, comprising a signal peptidase II cleavage site (Table 3-3). This qualifies the protein as lipoprotein, showing an N-terminal lipobox sequence, Leu-Ser-Ala-Cys in case of CE13, with the cleavage site after Ala, and Cys being target of acylation.

Table 3-3 Prediction of secretion of the CEHs investigated in this thesis by amino acid sequence analysis with the program SignalP 5.0 (Almagro Armenteros *et al.*, 2019). For *A. borkumensis* CEHs, no secretion signal was detected, hence, the results for *P. aestusnigri* CEHs are shown only. The enzyme identifier is given according to chapter 2.5 (ch. 2.5). The prediction identifies proteins with a Sec-signal peptide, comprising signal peptidase cleavage site I (Sec/SPI) or II (Sec/SPII), a Tat-signal peptide, comprising signal peptidase cleavage site I (Tat/SPI), or other proteins, meaning the absence of detectable signal peptides (OTHER). The organic solvent tolerance (OST) is indicated as in Table 3-1.

ch. 2.5	SignalP 5.0	OST
CE13	Sec/SPII	+++
CE14	OTHER	+
CE15	OTHER	o
CE16	Sec/SPI	+
CE17	Sec/SPI	o
CE18	OTHER	o
CE19	OTHER	+
CE20	OTHER	+
CE21	Sec/SPI	+
CE22	OTHER	+
CE23	Sec/SPII	o
CE24	OTHER	+
CE25	Sec/SPI	o

For a long time, it was believed that bacterial lipoproteins are exclusively found inside the cell, anchored either in the inner or the outer membrane and facing into the periplasm, but recent results show lipoproteins can be located outside the cell as well (Wilson and Bernstein, 2016). The pathway to direct lipoproteins to the inner side of the outer membrane, the Lol-pathway, comprises five proteins LolABCDE, where LolA functions as a shuttle for the lipoprotein, LolB as a receptor, and LolCDE resembles an ABC transporter (Okuda and Tokuda, 2011). Furthermore, secretion of lipoproteins across the outer membrane was reported to employ the type II or type V secretion pathway. However, by investigating the localization of lipoproteins in *Borrelia burgdorferi*, which lacks all numbered secretion pathways found in Gram-negative bacteria and lacks LolB, surface localization of many lipoproteins was detected. It was supposed, that LolA directs the respective lipoproteins to an undiscovered localization pathway, probably involving a specific flippase (Chen and Zückert, 2011). *P. aestusnigri* possesses homologs of LolA (WP_088274250.1) and LolCDE (WP_088276468.1; WP_088276466.1; WP_088276468.1) of the Lol-pathway but the draft genome sequence lacks a LolB homolog. However, due to the presence of different secretion machineries in the genome of *P. aestusnigri*, an export of CE13 to the cell surface or the extracellular space is possible. Hence, organic solvent tolerance of this enzyme may constitute an adaptation to the crude oil-associated lifestyle of the organism and a putative surface exposed localization of CE13.

In general, enzymes from organic solvent tolerant bacteria are frequently reported to be tolerant for organic solvents (Gupta and Khare, 2009), thus this might be an inherent feature of enzymes from these microbes. In their review, Gupta and Khare listed about

20 such enzymes, however, they only found a few reports on the structural basis for organic solvent tolerance (Gupta and Khare, 2009). Studies which described the basis of organic solvent tolerance for the respective biocatalyst, assert that disulfide bonds and surface hydrophobicity relate to the organic solvent tolerance.

The ability to hydrolyze polyester substrates is in general uncommon among carboxylic ester hydrolases. For example, in a comprehensive study of more than 200 purified CEHs, the authors found 36 enzymes active on polyester substrates, 12 of which were active with a PET trimer (Hajighasemi *et al.*, 2018). CEHs active with large PET polymers or PET film are even less abundant, making PET hydrolases particular rare enzymes (Danso *et al.*, 2018). For *A. borkumensis*, six enzymes were described before showing polyester hydrolytic activity (Hajighasemi *et al.*, 2018), two of these were part of this thesis, CE04 (ABO1251) and CE06 (ABO2449). However, no polyester hydrolytic activity was observed for any *A. borkumensis* CEH used in this thesis with the substrate Impranil DLN (appendix, Figure 5-1). This difference can be reasoned by (I) the nature of the used substrates (Impranil DLN is a synthetic polyester-polyurethane and the study by Hajighasemi *et al*. used polylactic acid derivatives mainly (Hajighasemi *et al.*, 2018)) and (II) the use of purified enzymes by Hajighasemi *et al.* in contrast to the use of living cells in this thesis. The screening applied in this thesis might therefore be limited by common bottlenecks of function-based screenings, as low expression of the target gene, low stability of the enzyme in the host organism, or low enzyme activity under the chosen laboratory condition. Thus, a further investigation of enzymes identified in this thesis in form of purified protein can afford the opportunity to observe enzyme activities, which were missed so far.

The investigation of Pseudomonads of the *P. pertucinogena* lineage, however, led straightforward to the identification of polyester hydrolases, which were found in every genome sequence available for bacteria of this phylogenetic lineage (chapter 2.2; (Bollinger, Thies, Katzke, *et al.*, 2020)). Further investigation gave prove to a polyester hydrolytic phenotype of the respective organisms (chapter 2.6; (Molitor *et al.*, 2020)). For one member of this lineage, *P. aestusnigri*, one enzyme was identified which was able to degrade the polyester-polyurethane Impranil DLN (CE16, appendix Figure 5-1) as well as PET film (chapter 2.7; (Bollinger, Thies, Knieps-Grünhagen, *et al.*, 2020)). However, a causality of the occurrence of PET hydrolases related to the hydrocarbonoclastic lifestyle of this bacterium is not obvious. Nevertheless, there is indication that hydrocarbonoclastic bacteria might be enriched for this class of CEHs,

given by the fact, that in a large sequence-based bioprospecting study, the highest hit rate for PET hydrolases was found in a metagenome from a sample of a crude-oil contaminated site (Danso *et al.*, 2018).

For the novel CEHs recovered from the hydrocarbonoclastic bacteria in this thesis (Table 3-1), characteristics important for biotechnological applications, were investigated in more detail. In the following, the outstanding specimens CE07, CE13, and CE16 (also named PE-H in chapter 2.7) for substrate promiscuity, organic solvent tolerance, and polyester hydrolysis, are discussed in the light of the current knowledge to illustrate underlying principles of these features (Figure 3-2).

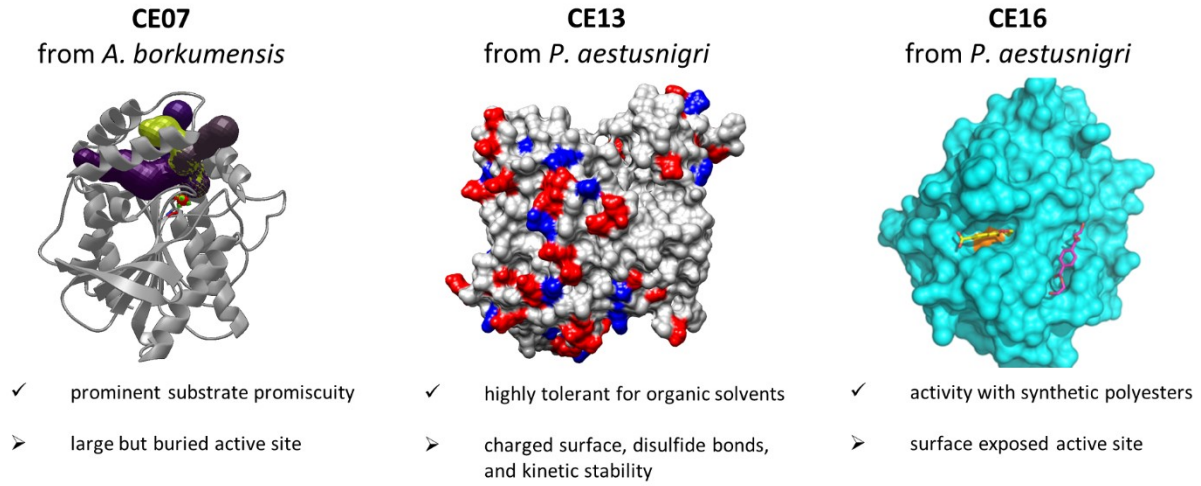


Figure 3-2 Overview of the three CEHs CE07, CE13, and CE16, which were investigated in detail because of their outstanding features. The exceptional characteristic (✓) and its molecular basis (➢), as discussed below (chapter 3.1.1, 3.1.2, and 3.1.3), are given. For the substrate promiscuous CE07 from *A. borkumensis*, a ribbon representation of a homology model (Mulnaes and Gohlke, unpublished) is shown, with the active site serine depicted as stick. The large size of the active site is illustrated by four different colored tunnels, computed with MOLE 2.0 (Sehnal *et al.*, 2013), which conjoin next to the active site serine. For the organic solvent tolerant CE13 from *P. aestusnigri*, a molecular surface representation of a homology model (Mulnaes and Gohlke, unpublished) is shown, with basic amino acid residues colored blue and acidic amino acid residues colored red. For the polyester hydrolytic CE16 from *P. aestusnigri*, a molecular surface representation of the crystal structure of the engineered variant PE-H Y250S (PDB code 6SCD) is shown. Substrate molecules were computed to bind to the surface exposed active site (MHET, yellow) and to a hydrophobic grove next to the active site (BHET, magenta) (Bollinger, Thies, Knieps-Grünhagen, *et al.*, 2020).

3.1.1 High substrate promiscuity does not forbid chiral selectivity

To measure the substrate promiscuity of CEHs, Martínez-Martínez and coworkers assembled a set of 96 esters of high structural diversity (Martínez-Martínez *et al.*, 2018). The number of esters hydrolyzed was used in this study as a basis for the classification of more than 145 CEHs according to their substrate promiscuity. This analysis of CE07 showed, that it was able to hydrolyze 65 out of 96 structurally diverse esters (chapter 2.5; (Bollinger, Molitor, *et al.*, 2020)). Ranked within the enzyme set investigated in chapter 2.3 (Martínez-Martínez *et al.*, 2018), this renders the enzyme one of the ten most promiscuous CEHs tested so far with this methodology (Figure 3-3). In addition, CE07 was shown to react with demanding, water insoluble esters of relevant chemical building blocks (chapter 2.5; (Bollinger, Molitor, *et al.*, 2020)). The enzymatic hydrolysis of these compounds was not shown before and points out great potential for an application of CE07 in pharmaceutical biotechnology, because these building blocks occur in the synthesis routes of many active pharmaceutical ingredients (Hameed *et al.*, 2018). A biotechnological application of CE07 is therefore conceivable for the hydrolysis of complex, aromatic, halogenated, and water insoluble esters, as it moreover tolerates different organic solvents.

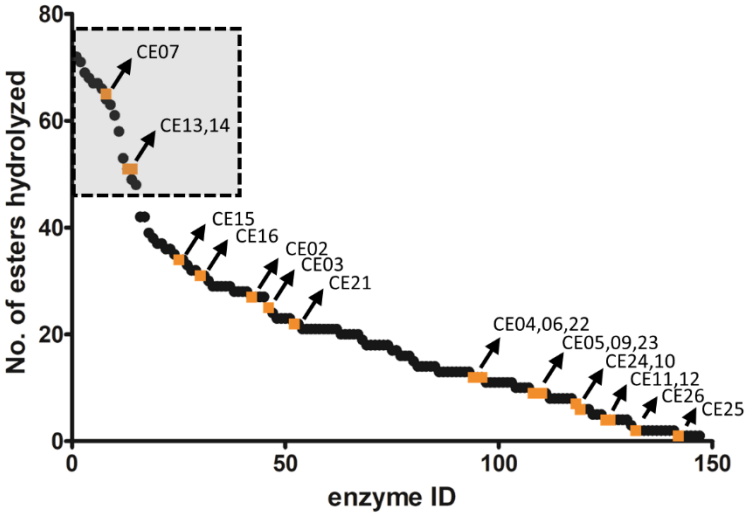


Figure 3-3 Ranking of the substrate promiscuity for CEHs identified in this thesis among 145 diverse ester hydrolases. Data taken from chapter 2.3 (Martínez-Martínez *et al.*, 2018) and 2.5 (Bollinger, Molitor, *et al.*, 2020). The number of esters hydrolyzed out of 96 as a marker for substrate promiscuity is plotted against the enzyme identifier as defined in chapter 2.3 (black circles). Enzymes identified in this thesis are labeled according to chapter 2.5 (orange rectangles, black arrow). Enzymes showing prominent substrate promiscuity are highlighted (faded grey rectangle with dashed black line).

In accordance with the results shown in chapter 2.5, CE07 can cope with decent concentrations of different organic solvents. As shown in the bachelor thesis by Simone Söltl (Söltl, 2017), which was done in association with this thesis, CE07 (named LipD by Söltl 2017) was tolerant for 30 to 40 % (v/v) of DMSO, methanol, ethanol, acetone, and for acetonitrile or isopropanol at concentrations up to 20 % (v/v). Thus, the usage of many organic solvents as cosolvents with CE07 is possible.

A major drawback of many substrate promiscuous CEHs is their low chiral selectivity (chapter 2.4; (Coscolín, Martínez-Martínez, *et al.*, 2018)). Interestingly, CE07 appears to be an exception and was strictly selective for the (*R*) enantiomer of Menthyl acetate and Methyl 3-hydroxybutyrate, when the (*R*) and (*S*) enantiomer of both esters was tested separately (*i. e.* not as racemic mixture) (Figure 3-4, left). A homology model of CE07 was calculated by the group of Holger Gohlke at the Heinrich Heine University Duesseldorf (unpublished), using TOPModel and TOPScore (Mulnaes and Gohlke, 2018), which allowed molecular docking computation with (*R*/*S*)-Menthyl acetate as ligand using UCSF Chimera (Pettersen *et al.*, 2004) with autodock VINA (Trott and Olson, 2009). The ligand structures were retrieved from ZINC database (Sterling and Irwin, 2015). Both enantiomers can be found in mirror-image orientation, with similar predicted free energies of binding (-6.1 kcal/mol for the *R* and -6.2 kcal/mol for the *S* enantiomer), in a reasonable distance to the catalytic serine (about 4 Å from the serine hydroxyl to the carbonyl carbon of the substrate) and in hydrogen bond distance (2.3 Å) between the ligand carbonyl oxygen and the amide backbone of Gly77 of the oxyanion hole (Figure 3-4, right). Hence, CE07 might be able to bind the two enantiomers and for that reason the selectivity of CE07 must be rationalized differently. In contrast to the selectivity for the hydrolysis of (*R*)-Menthyl acetate and Methyl (*R*)-3-hydroxybutyrate, selectivity was low for other chiral esters such as Methyl (*R*)-3-hydroxyvalerate, which differs in only one carbon atom of the main chain length from Methyl (*R*)-3-hydroxybutyrate, suggesting that distinct structural characteristics of the active site environment of CE07 define its selectivity.

Various decisive characteristics of enzyme-substrate interactions are known with regard to enantioselectivity. A well investigated example is the lipase from *Burkholderia cepacia* (BCL), formerly known as *Pseudomonas cepacia*, which is enantioselective for multiple substrates. In this case, the enantioselectivity of the enzyme was rationalized by a better binding of the preferred enantiomer (Lang *et al.*, 1998; Tuomi and Kazlauskas, 1999), a shorter hydrogen bond distance between the hydroxyl group

of the catalytic serine and the carbonyl carbon of the preferred substrate (Tomić *et al.*, 2001), or a higher rate of catalysis for the preferred enantiomer (Mezzetti *et al.*, 2005). However, it was also shown that the molecular mechanism of enantioselectivity of the same enzyme for different substrates can differ considerably. For example, Lang *et al.* showed that the preference for the (*R*) enantiomer of a chiral triacylglyceride substrate analog is based on unfavorable binding of the (*S*) enantiomer to BCL (Lang *et al.*, 1998), whereas Mezzetti *et al.* discovered no preference in binding of two enantiomers of a chiral primary alcohol substrate analog (nearly the same K_m), but a 100-fold higher reaction rate (k_{cat}) for the (*S*) enantiomer with BCL (Mezzetti *et al.*, 2005). Thus, for a detailed characterization of CE07 enantioselectivity, enzyme kinetic studies and the test of racemic substrates are needed.

substrate	**SF**
(*R*/*S*)-Menthyl acetate	100 (*R*)
N-Benzyl-(*R*/*S*)-proline ethyl ester	n. a.
Methyl-(*R*/*S*)-mandelate	8 (*S*)
(*R*/*S*)-Ethyl-4-chloro-3-hydroxybutyrate	2 (*S*)
Methyl-(*R*/*S*)-3-hydroxybutyrate	100 (*R*)
Methyl-(*R*/*S*)-3-hydroxyvalerate	6 (*R*)
(*R*/*S*)-Neomenthyl acetate	n. a.
(*R*/*S*)-Pantolactone	n. a.
Ethyl-(*R*/*S*)-lactate	1
Methyl-(*R*/*S*)-lactate	2 (*R*)

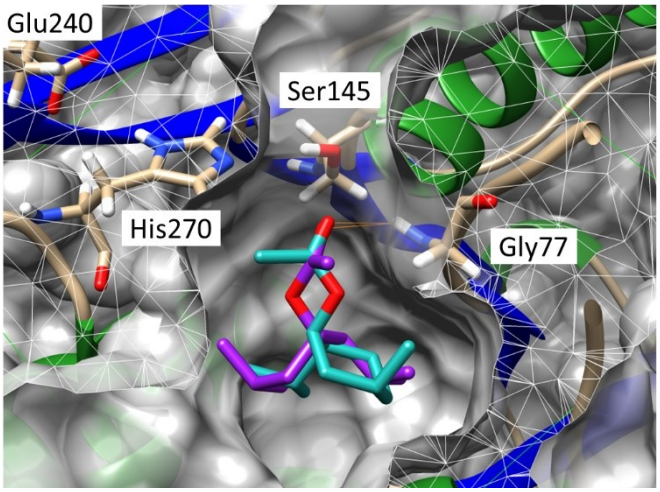


Figure 3-4 Chiral selectivity of CE07. **Left:** the chiral selectivity factor (SF), expressed as ratio of the specific activity for the preferred enantiomer over the non-preferred, is given for 10 chiral substrates; "n. a." means not active. The specific activity was determined for each enantiomer individually. Data from chapter 2.5, Table S2 (Bollinger, Molitor, *et al.*, 2020). **Right:** zoom into the active site of CE07 homology model with docked substrates (*R*)-Menthyl acetate (cyan) and (*S*)-Menthyl acetate (magenta). Amino acids of the catalytic triad and the oxyanion hole are shown as sticks with labels. Molecular surface (grey) was capped (white mesh) to allow insight into the buried active site. The lowest energy pose for each enantiomer, that allowed hydrogen bonding with the oxyanion hole (orange line), is shown. Molecular docking computation was done using UCSF Chimera along with autodock VINA.

However, the fact that CE07 displayed high substrate promiscuity and chiral selectivity appears interesting for applications in the pharmaceutical industry for example, where large molecules are common as substrate for biocatalysis, and enantiomerically pure products are desired. The selectivity of CE07 for some chiral esters exemplifies the general ability of this enzyme for enantioselective hydrolysis and provides the basis for

an adjustment of selectivity toward further desired substrates by protein engineering. The engineering of enantioselectivity has been successfully performed in the past, *e. g.* by directed evolution, both with and without the knowledge of the protein structure, for many different enzyme classes (Jaeger and Eggert, 2004; Li and Reetz, 2016).

A deeper insight into the molecular structure of CE07 would surely promote protein engineering approaches, for example a structure in complex with a substrate analog can help to reveal interaction sites for the tetrahedral intermediate and to compute models for enantioselectivity. For CE07 however, a crystal structure was not solved so far, mainly due to limitations in the purification of the protein. In a bachelor thesis by Deborah Weide (Weide, 2018), associated with the presented doctoral thesis, a simple partial purification procedure for CE07 was established, for a direct precipitation of the protein from cell extracts. The method based on precipitation of CE07 with ammonium sulfate allowed to avoid costly chromatographic purification methods, it thus might offer an economically feasible application for the enzyme. In addition, proteins which can be precipitated in an active form by the use of ammonium sulfate can be further processed with glutaraldehyde to form cross-linked enzyme aggregates (CLEAs) (Schoevaart *et al.*, 2004; Roy *et al.*, 2017). This immobilization technique avoids expensive immobilization supports, thus the economical application of CE07 can probably be increased by formulation as CLEA.

3.1.2 Mechanistic insights into the organic solvent tolerance of CE13

An extraordinary example for an organic solvent tolerant esterase was identified with CE13 (chapter 2.5; (Bollinger, Molitor, *et al.*, 2020)). The enzyme was active in the presence of different polar organic solvents, for example 50 % (v/v) acetonitrile, and retained significant activity after several hours of incubation in 80 % (v/v) methanol, acetonitrile, 1,4-dioxane, and dimethyl sulfoxide. Although it is generally accepted that lipases and esterases have a high probability to show tolerance for organic solvents (Jaeger and Eggert, 2002), examples which are tolerant for such high concentrations as reported for CE13 are scarce. As an example, the lipase from *Burkholderia ambifaria* was reported to retain 60 % residual activity after incubation for 60 days at 30 °C in 25 % of acetonitrile (Yao *et al.*, 2013), which was revealed to be the most disruptive organic solvent tested in this study as well as in chapter 2.5 of this thesis. CE13 was not tested under the same condition, but at significantly higher acetonitrile

concentration (80 %) where it retained 33 % residual activity after 3 h incubation at 30 °C. The molecular mechanism of the high organic solvent tolerance of CE13 is therefore very interesting.

In general, the basis for organic solvent tolerance of enzymes is most probably composed of different cooperative molecular characteristics. To name a few examples, structural rigidity was reported as a hallmark for organic solvent tolerance, because it slows down the penetration of organic solvent molecules into the hydrophobic protein core (Mohtashami *et al.*, 2018). Furthermore, the presence of charged amino acid residues on the protein surface as well as polar residues in the protein core was identified to significantly contribute to a proteins tolerance for organic solvents (Frauenkron-Machedjou *et al.*, 2018). Likewise, thermostability has been suggested to be positively correlated with organic solvent tolerance (Doukyu and Ogino, 2010; Kumar *et al.*, 2016), however without proof of causality so far.

The organic solvent tolerant CE13 from the marine mesophilic bacterium *P. aestusnigri* shows two important features which relate to organic solvent tolerance. First, the protein comprises a high number of negatively charged amino acid residues, leading to a theoretical pI of 4.1 as calculated with the program ProtParam (Gasteiger *et al.*, 2005). Second, the amino acid sequence shows six Cys residues, which might be involved in disulfide bond formation and thus, stabilizing the protein structure. These features were investigated in greater detail during a master thesis conducted by Marvin Bulka within the framework of this thesis, complementing the presented studies on CE13 (Bulka, 2019). By enzyme activity measurements Bulka showed, that CE13 is not thermostable, exhibiting a half-inactivation point at about 50 °C. Furthermore, Bulka observed extraordinary tolerance of CE13 for different organic solvents and chemicals, including harsh denaturants like SDS or urea. In absence of a crystal structure for CE13, a homology model was computed by the group of Holger Gohlke at the Heinrich Heine University Duesseldorf (unpublished) using TOPModel and TOPScore (Mulnaes and Gohlke, 2018), which was used to compare structural features with a homology model of EstDL30, computed with Phyre2 (Kelley *et al.*, 2015). Est DL30 is a structural homolog of CE13 (rmsd about 0.7 Å) that was reported to be less tolerant for organic solvents and denaturation agents (Tao *et al.*, 2011), thus might reveal structural determinants for CE13 stability. Based on this comparison, Bulka suggested that the negative surface charge of the protein contributes to the resistance for the anionic detergent SDS and further strengthened this hypothesis by

enzyme inactivation measurements with cationic detergents and divalent cations (Bulka, 2019). Moreover, the author studied the stability of CE13 by nano differential scanning fluorimetry upon the addition of a reducing agent, an anionic, and a cationic detergent (Figure 3-5).

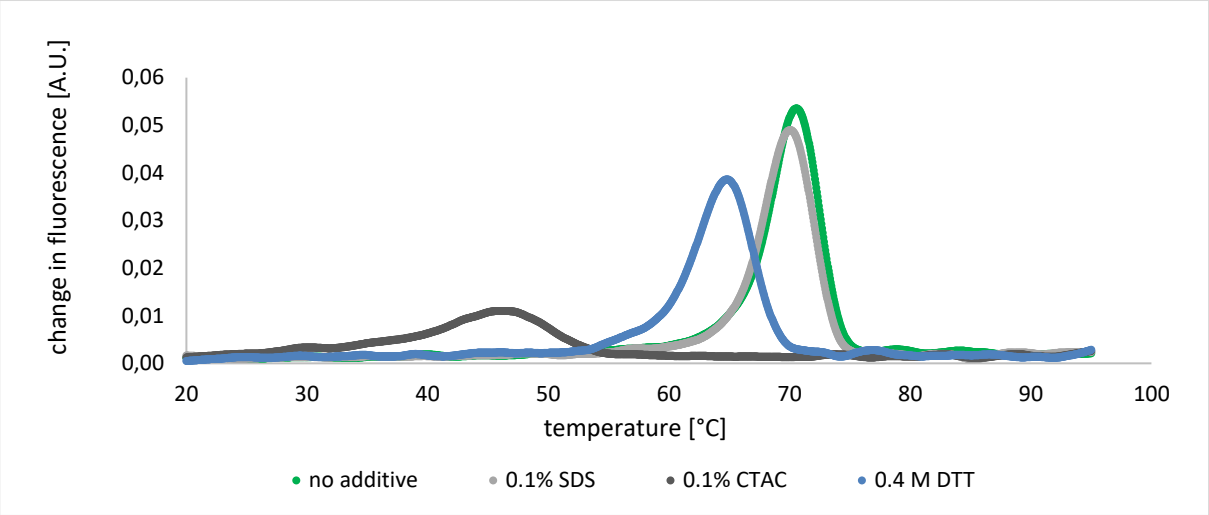


Figure 3-5 Nano differential scanning fluorimetry measurement of the thermal melting point of CE13 upon incubation with different additives. Data taken from (Bulka, 2019). The thermal melting point (T_m) is located at the local maximum of the curve. Either 0.1 % of anionic detergent sodium dodecyl sulfate (SDS, light grey), 0.1 % of cationic detergent cetrimonium chloride (CTAC, dark grey), 0.4 M of reducing agent dithiothreitol (DTT, blue), or no additive (no additive, green) were added to CE13. T_m were 70.4 °C (no additive), 69.9 °C (SDS), 64.6 °C (DTT) and 46.0 °C (CTAC).

The results indicate reduced stability of the protein with the reducing agent DTT present and a severely lower melting point upon the addition of the cationic detergent CTAC. Addition of anionic detergent SDS however, did not reduce the melting point of CE13. These observations are in good agreement with the hypothesized relevance of negative surface charge and disulfide bonds for the high tolerance of CE13 for organic solvents and chemical denaturants.

Moreover, studies of the SDS tolerance of CE13 gave the first evidence for a high kinetic stability of CE13, which may represent one more feature related to the high organic solvent tolerance of the enzyme. In general, proteins exist in an equilibrium between native and unfolded state, which are characterized by a difference in Gibbs free energy (ΔG) (Figure 3-6). Kinetically stable proteins show a high activation energy of unfolding ($\Delta G^{\ddagger}$), which traps the protein in a native conformation (Manning and Colón, 2004). Therefore, kinetically stable proteins have a slow rate of unfolding in the

presence of chemical denaturants, show resistance to proteases, and exhibit a long half-life (Colón *et al.*, 2017).

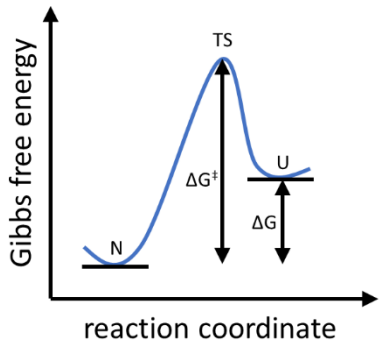


Figure 3-6 Reaction coordinate diagram for protein unfolding. The native state (N) of a protein shows lower Gibbs free energy (ΔG) than the unfolded state (U); therefore, the native state is thermodynamically favored. To unfold a protein, energy is required to reach the transition state (TS) of the reaction on the way toward the unfolded state. The energy difference between the native state and the transition state is called activation energy (ΔG‡). Kinetically stable proteins are trapped in the native state by a high activation energy for protein unfolding (Manning and Colón, 2004).

However, the basis of the observed kinetic stability of CE13 remains to be proven. To gain deeper insights into the molecular mechanism of organic solvent tolerance of CE13, molecular dynamics simulations of the interaction of the protein with organic solvent could be carried out as it was already successful for the analysis of organic solvent tolerance of the *B. cepacia* lipase BCL and other enzymes (Mohtashami *et al.*, 2018).

3.1.3 Molecular characteristics of CE16 suggest possible applications

For the degradation of synthetic polyesters, including the abundant waste polymer polyethylene terephthalate (PET), CE16 (named PE-H in chapter 2.7) was identified as one of the first examples of PET degrading enzymes from a marine mesophilic bacterium. Its crystal structure was shown to represent the first example for a PET hydrolytic enzyme of type IIa (chapter 2.7; (Bollinger, Thies, Knieps-Grünhagen, *et al.*, 2020)). High structural similarity to other described PET hydrolases was shown, with differences in the orientation of five loop regions. A drastic effect of the orientation of two of these loops, the first connecting β3-α2, and the second connecting β4-α3, was shown by a mutational analysis. This difference was connected to a specific amino acid position, where a substitution from tyrosine to serine led to a structural

rearrangement, and to enhanced hydrolysis of various substrates. Similar biochemical effects were reported for a homologous cutinase, TfCut2 from *Thermobifida fusca*, mutated at the same position (Furukawa *et al.*, 2019), but structural evidence explaining these results was not included in the respective study. Because of the high structural homology among this class of enzymes, a similar mechanism like it was shown for CE16 might apply to this enzyme too.

With such deep insights into the molecular mechanism of this biocatalyst, a rational evaluation of future applications is possible. For PET hydrolytic enzymes like cutinases, various applications were suggested before (Figure 3-7).

(I) The bioremediation of plastic waste, which was released in the environment, might be the most obvious field of application for PET hydrolyzing enzymes. Hence, it was suggested by different studies dealing with microbial plastic degradation in the past (Webb *et al.*, 2013; Kumar *et al.*, 2016; Wei and Zimmermann, 2017b, 2017a; Danso *et al.*, 2019; Hiraga *et al.*, 2019; Kawai *et al.*, 2019; Salvador *et al.*, 2019). A special challenge is the low temperature in most polluted sites and the fact that many efficient PET degrading enzymes originate from thermophilic organisms. Thus, their activity at low temperature is not sufficient. The first known and best studied example for an enzyme, which is active with PET at ambient temperature, is PETase from *Ideonella sakaiensis* (Yoshida *et al.*, 2016). The bacterium *I. sakaiensis* can degrade and assimilate PET, by employing the two enzymes PETase and MHETase. Therefore, a direct conversion of the recalcitrant polyester to bacterial biomass is possible and paves the way to various recycling and upcycling processes. Since its identification in 2016, the application of PETase was exemplified for biocatalytic PET degradation (Yoshida *et al.*, 2016; Han *et al.*, 2017; Austin *et al.*, 2018; Joo *et al.*, 2018), and suggested to contribute to a circular PET economy (Wierckx *et al.*, 2018; Hiraga *et al.*, 2019; Taniguchi *et al.*, 2019). Moreover, PETase was reported to be actively expressed in marine microalgae, opening the possibility to develop photobioreactors for sun-light driven PET degradation processes, or enabling bioremediation strategies in marine habitats to clean PET-polluted sites (Moog *et al.*, 2019).

(II) The application of thermostable cutinases like TfCut2 from *Thermobifida fusca* can facilitate biotechnological PET recycling processes. In particular the combination of a reaction temperature close to the glass transition temperature of PET (around 80 °C), at which the amorphous regions of the polymer become more flexible and thus more

accessible for enzymatic attack, and the addition of an ionic detergent to facilitate the binding of the catalyst to the polymer was shown to enhance the catalytic activity significantly toward PET (Furukawa *et al.*, 2019). The authors further suggest an industrial application in combination with a flow reactor.

(III) The treatment of wastewater for polyester particles is an example where PET hydrolytic enzymes from mesophilic bacteria are considered promising. An enzyme (PpelaLip) originating from *Pseudomonas pelagia*, a close relative of *P. aestusnigri*, was successfully applied for poly(oxyethylene terephthalate)-based polymer hydrolysis under simulated wastewater treatment plant conditions (Haernvall *et al.*, 2018).

(IV) In the textile industry, polyester fibers are used due to their low price, robustness, and easy manufacturing, but some features of the material can be cumbersome, for example the hydrophobicity of the textiles surface. Treatment with cutinases can increase the hydrophilicity of these surfaces by producing free hydroxyl and carbonic acid groups (Nikolaivits *et al.*, 2018). This procedure was recognized by the Novozymes company in a patent in 1997 (Riegels *et al.*, 1997).

(V) Besides the hydrolysis of polyesters, the synthesis reaction of such is catalyzed by cutinases too (Nikolaivits *et al.*, 2018). The enzymes thus may also find application in the synthesis of polyester materials. The cutinase from *Humicola insolens* (HiC) was used for the polycondensation of different diacids and diols (Hunsen *et al.*, 2007). However, the use of bulk organic solvents and high temperature restricts the application to cutinases able to cope with these harsh reaction conditions. Another example for an application in a synthesis reaction, avoiding bulk solvent and high temperature, was shown with PETase in a double transesterification reaction to yield the diester of 2,5-furandicarboxylate and 1,4-butanediol, a precursor of the bioplastic polybutylene adipate-cobutylene 2,5-furandicarboxylate (PBAF) (Parisi *et al.*, 2019). Remarkably, the transesterification was carried out in the presence of 50 % aqueous solution.

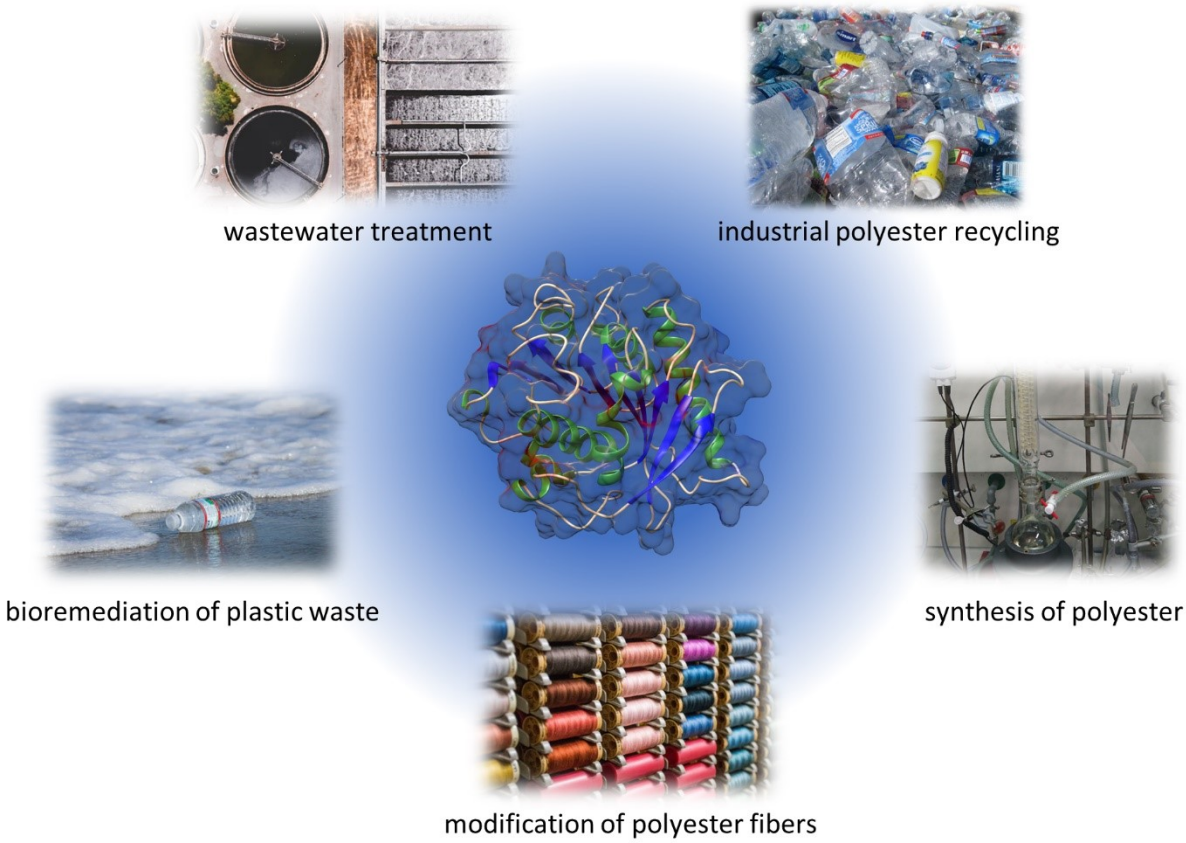


Figure 3-7 Possible applications for polyester hydrolytic enzymes. The structure of CE16 mutant Y250S (PDB code 6SCD) is depicted in the center with pictures of examples for applications shown around. Pictures of the plastic bottle on a seashore (by Brian Yurasits), blue labeled plastic bottles (by Tanvi Sharma), wastewater treatment plant (by Ivan Bandura), and assorted colored thread lot (by Héctor J. Rivas) were retrieved from https://unsplash.com licensed under public domain (CC0).

Of the above listed applications of PET hydrolytic enzymes, CE16 may be suited as wastewater treatment enzyme (III), for the modification of polyester fibers (IV), or the synthesis of (poly-)esters (V). Because of its high enzymatic activity at low temperature – CE16 showed the highest activity at around 20 °C and retained about 50 % activity at around 10 °C (Turkes, 2019) – the enzyme may directly be applied to cleave polyester substrates in wastewater, or used as a laundry detergent enzyme for washing at low temperature, which goes along with tremendous energy savings. However, the use of CE16 as a washing agent enzyme might be hampered due to stability issues and might require protein engineering to enhance the enzyme's tolerance for detergents. In a bachelor thesis by Lejla Turkes, which was conducted in association with this thesis, CE16 was found to be rapidly inactivated by low concentrations of surfactants, exemplified by rhamnolipids (Turkes, 2019), biological surfactants applicable in detergents. In a first attempt to use CE16 in a synthesis reaction, the synthesis of different aliphatic wax esters was demonstrated (Figure 3-8).

The synthesis of polyesters with CE16 was not investigated so far, but an application similar to the transesterification of 2,5-furandicarboxylate and 1,4-butanediol shown for PETase (Parisi *et al.*, 2019) is conceivable, with respect to the observed synthetic activity of CE16.

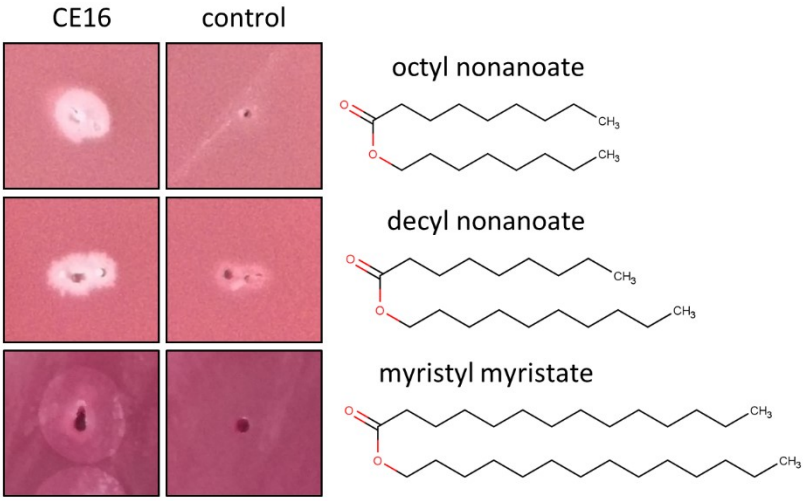


Figure 3-8 Assessment of the synthesis of different aliphatic wax esters by CE16, using the method described by Sandoval and Marty (Sandoval and Marty, 2007). The enzyme (CE16), or buffer without enzyme (control) was applied to agar plates (left side) containing an aliphatic alcohol and a fatty acid (7.5 g/l nonanoate or myristate; 7.5 g/l 1-octanol, 1-decanol, or 1-myristol; 1.5 g/l gum arabic; 10 mg/l rhodamine B; 1.5 % (w/v) agar-agar in LB medium). Pictures were taken after 48 h incubation at 30 °C. The white substance around the application spot of the enzyme shows the produced wax ester upon esterification activity of the enzyme. Structural formulas of the ester products are shown (right side).

Otherwise, applications like bioremediation of waste PET (I), or industrial PET recycling (II) are probably not feasible with CE16. For practical bioremediation of PET, application of microorganisms able to metabolize the polyester is advantageous over the application of pure enzyme. The originating organism of CE16, *P. aestusnigri,* however is most probably not able to fully degrade and assimilate PET, since a homolog to the second key enzyme in the PET catabolic pathway identified in *I. sakaiensis*, named MHETase (Yoshida *et al.*, 2016), is missing. Similarly, the low thermostability of CE16 is a limit for its application in an industrial PET recycling process. The highest enzymatic degradation rate for PET can be achieved close to the glass transition temperature of PET (around 80 °C). Thus, the efficiency of thermostable enzymes from thermophiles is in general superior to PET hydrolyzing enzymes from mesophiles (Wei *et al.*, 2019). For CE16, immobilization experiments were carried out to increase its thermostability. In a bachelor thesis conducted by Alina Kuklinski, in the frame of this doctoral thesis, first immobilization trials were done

(Kuklinski, 2017). The immobilized CE16 showed improved thermal stability, retaining 20 % of its initial activity at 70 °C, and reusability of the immobilized enzyme formulation. Thus, the usage of CE16 in other forms than free enzyme is in principal possible and might allow an application at elevated temperature.

3.2 Protein engineering towards the dream biocatalyst – combining industrially important features

For industrial application of biocatalysts, process transfer from small to large scale is one of the most challenging steps. This is made more difficult by the need to change the biocatalyst depending on the reaction to be catalyzed or the substrate to be used. Every biocatalyst-substrate combination comes with different requirements, which must be optimized individually. With highly promiscuous enzymes, however, the chance is high to be able to reuse established industrial process conditions and avoid upscaling bottlenecks to a large part. In theory, such promiscuous “dream” biocatalyst can catalyze reactions with a wealth of substrates at varying process conditions. In reality, this might not be achievable with one single protein, but with some requirements fulfilled by a given enzyme, an approximation to this utopic biocatalyst is possible.

The here investigated features, substrate promiscuity, organic solvent tolerance, and polyester hydrolysis are desirable in combination to obtain a robust and simultaneously versatile biocatalyst. The investigated features come with specific biochemical and biophysical characteristics of which some may combine well, and others may exclude each other (Figure 3-9). For example, high substrate promiscuity of an enzyme usually comes along with decreased enantioselectivity (chapter 2.4; (Coscolín, Martínez-Martínez, *et al.*, 2018)). However, enzymes exist which display enantioselectivity and substrate promiscuity at the same time (CE07, chapter 2.5; (Bollinger, Molitor, *et al.*, 2020)) and methods are at hand to engineer such important features in a straightforward and reliable way (Li and Reetz, 2016; Sun *et al.*, 2016). Thus, a substrate promiscuous enzyme may be used as a chassis, which can be modified to fit specific tasks, and therefore enabling to benefit from an optimized process for this biocatalyst for different products. An inherent feature of most highly substrate promiscuous CEHs is a large and buried active site, to provide on the one hand enough space for the substrate to bind and on the other hand to prevent unfavorable contact of the substrate with the surrounding solvent (chapter 2.3; (Martínez-Martínez *et al.*,

2018)). In contrast, polyester degrading enzymes feature solvent accessible active sites to enable binding of the hydrophobic polymer chain (chapter 2.7; (Bollinger, Thies, Knieps-Grünhagen, *et al.*, 2020)). Hence, the feature of polyester hydrolytic activity and broad substrate spectra for non-polymeric substrates is contradicting.

The organic solvent tolerance of proteins is often connected to the rigidity of their three-dimensional structure, because denaturation of proteins by organic solvents occurs due to penetration of organic solvent molecules into the hydrophobic protein core. The hydrophobic interaction between amino acids of the protein core, which is the main driving force of protein stability (Pace *et al.*, 2011), is then disturbed by the organic solvent molecules, leading to a loss of structure. Characteristics which decelerate this process effect a tolerance for organic solvents.

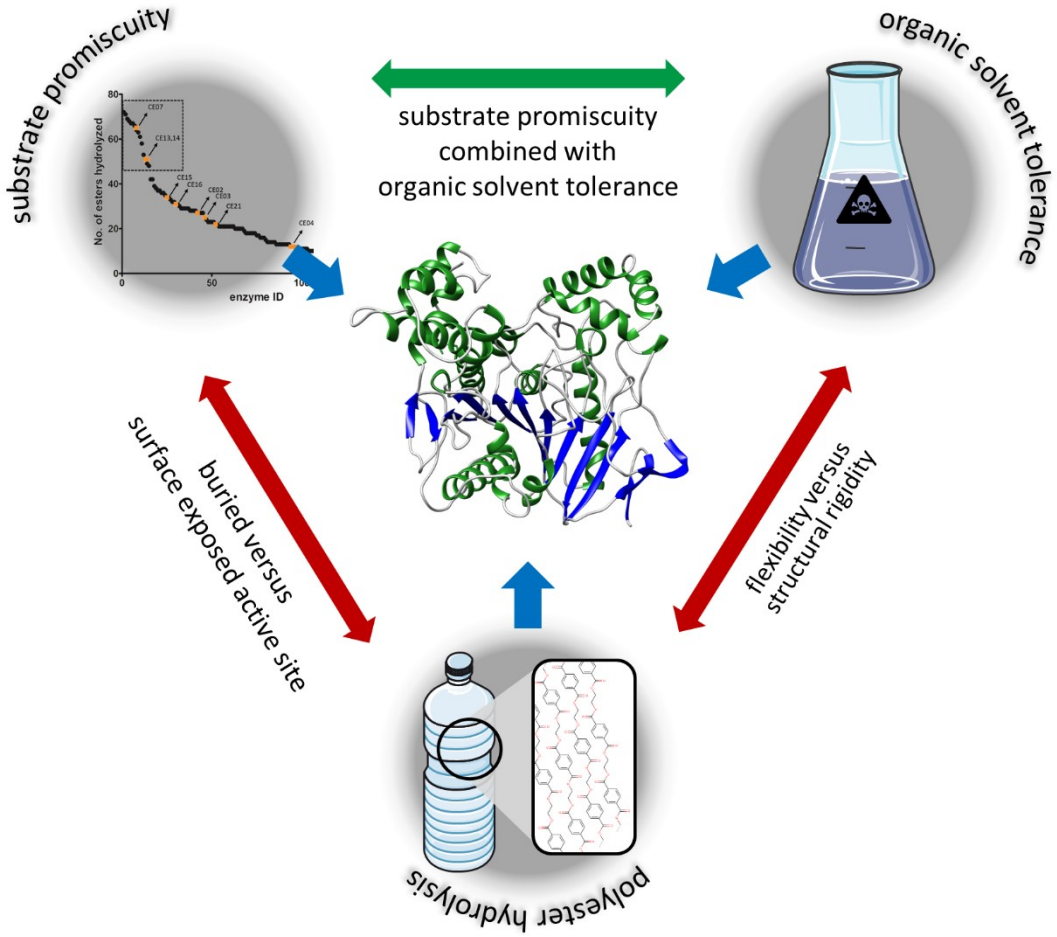


Figure 3-9 Interdependence of substrate promiscuity, organic solvent tolerance, and polyester hydrolytic activity of CEHs. Features which combine well are connected by a green arrow, contradicting characteristics are connected by a red arrow. In the center, the lipolytic biocatalyst is exemplified by a predicted structure of CE13 (Mulnaes and Gohlke, unpublished). Pictures of the plastic bottle and the Erlenmeyer flask were retrieved from servier medical art (https://smart.servier.com/), licensed under Creative Commons Attribution 3.0 (CC BY).

Hence, mutations increasing the organic solvent tolerance of *Bacillus subtilis* lipase A (BSLA) were reported for substitutions to amino acids with polar residues at buried positions along with substitutions to charged amino acid residues at surface exposed positions of the protein (Frauenkron-Machedjou *et al.*, 2018). Polar residues thereby support hydrogen bond formation in the protein inner core to stabilize the folded state (or the other way around, to destabilize the unfolded state), and charged residues at the protein surface contribute to maintain the hydration shell. Thus, regions contributing to enhanced protein stability can be located apart from the active site of an enzyme and therefore, amino acid substitutions which increase the organic solvent tolerance, and which alter the substrate specificity of a CEH, might be introduced to the same protein. Conclusively, the combination of substrate promiscuity and organic solvent tolerance might be possible in a single enzyme.

For polyester hydrolytic enzymes, organic solvent tolerance does not appear mandatory, but the underlying molecular mechanism of organic solvent tolerance might facilitate their application. In case of the organic solvent tolerant CE13, kinetic stability was suggested to be the basis of this feature (chapter 2.5 and 3.1.2), which also caused resistance to chemical denaturants like SDS, and is known to cause a long enzyme half-life due to structural rigidity (Colón *et al.*, 2017). For polyester hydrolysis at low temperature, as shown for PET hydrolytic enzymes from mesophilic organisms *e. g.* PETase or CE16, durability is of high value, because the decomposition of synthetic polymers like PET usually occurs at a slow rate and thus takes much time. For example, the PET degrading microbial consortia “no. 46”, which was the source of the PET degrading bacterium *I. sakaiensis*, degrades PET film at 30 °C at a rate of 0.13 mg cm^{-2} day^{-1} (Yoshida *et al.*, 2016). An acceleration of the PET decomposition was shown for example by the addition of SDS for the *I. sakaiensis* PETase (Furukawa *et al.*, 2018). Therefore, enzymes with a long half-life and tolerance for chemicals like SDS, as seen for CE13 (chapter 2.5 and 3.1.2), are of considerable relevance for polyester hydrolysis under these conditions. The molecular characteristics for the kinetic stability of CE13 are connected to the negative surface charge of the enzyme and stabilizing disulfide bonds (chapter 2.5 & 3.1.2). However, a combination of molecular characteristics for polyester hydrolysis and kinetic stability might not be feasible. This is mainly due to the contradiction of structural rigidity as a hallmark for kinetic stability and structural flexibility as a hallmark for polyester hydrolysis, as seen by the surface exposed and flexible active site of PETase (Fecker *et al.*, 2018).

Mutations decreasing the protein flexibility by addition of intramolecular contacts, as reported for additional polar interaction in BSLA (Frauenkron-Machedjou *et al.*, 2018), might therefore also reduce the PET hydrolytic activity of PETase. Moreover, the modification of the surface charge might likewise not be feasible for polyester hydrolases, as demonstrated by the substitution of four cationic surface exposed amino acids of PETase to glutamic acid, which led to a clearly decreased PET degradation rate (Furukawa *et al.*, 2018).

The combination of multiple mutations with beneficial impact for a specific characteristic, *e. g.* enantioselectivity, within one enzyme was demonstrated several times (Li and Reetz, 2016), however the combination of mutations affecting different characteristics, like stability and selectivity, was not investigated in detail so far. Yet it was shown that most mutations which introduce a new function to an enzyme cause destabilization to the protein (Tokuriki *et al.*, 2008). Hence, naturally stable proteins are generally thought to tolerate more destabilizing mutations as a tradeoff for enhanced catalytic characteristics (Socha and Tokuriki, 2013). With regard to this, the basis for a biocatalyst which is meant to be extensively modified by protein engineering to fulfill different industrially relevant functions, should be extraordinary stable by nature.

In conclusion, the combination of all three features, organic solvent tolerance, substrate promiscuity, and polyester hydrolysis in a single biocatalyst is difficult due to conflicting molecular characteristics of the features among each other. To design a CEH, close to the ideal biocatalyst, which shows as much of these features as possible, the combination of substrate promiscuity and organic solvent tolerance can be recommended. For example, CE13 identified in this thesis could be used as a chassis for substrate specificity engineering, because of its high structural stability by nature (CE13 showed 60 % residual activity after incubation in 80 % methanol for 3 h) and its likewise broad substrate range (51 out of 96 substrates hydrolyzed) (chapter 2.5; (Bollinger, Molitor, *et al.*, 2020)). As discussed above, both features are not contradicting and can be further improved by protein engineering approaches. A combination of polyester degradation with the other two desired features, organic solvent tolerance as an effect of kinetic stability, and substrate promiscuity, is not easily possible, but the combination of the latter two in one single biocatalyst is realistic.

3.3 Future perspectives and research needs

To promote the change from chemical to biochemical processes and enable a bio-based economy, the demand for novel biocatalysts is high. Nowadays the availability of putative enzyme candidates is not limiting, but robust data on biochemical and biophysical properties of the enzymes are. In this thesis, novel CEHs from marine hydrocarbonoclastic bacteria were identified, isolated, and characterized regarding their potential for polyester hydrolysis, as well as their substrate spectra, and organic solvent tolerance.

To accomplish this, an assay was established to identify CEH activity in the presence of polar organic solvents. By this strategy, 16 organic solvent tolerant CEHs were identified, including one highly tolerant enzyme (CE13). Further protein stability studies of this enzyme point to kinetic stability or structural rigidity as the basis of tolerance for organic solvents as well as detergents. However, without a known structure of the protein, rationalization of the observed stability is complicated. Future studies should therefore aim to solve the 3D structure of CE13; with this at hand, molecular dynamics simulations can be used to unveil stability determinants as it was shown for different proteins before (Mohtashami *et al.*, 2018). By comparison to a homologous protein, which does not show high stability under the same conditions, regions of the protein or positions in the amino acid sequence of the polypeptide relevant for stability can be identified. The most similar enzyme in terms of protein sequence identity, which was characterized for organic solvent tolerance is EstDL30 (33 % global sequence identity to CE13; complete inactivation after 40 minutes with 30 % acetonitrile), a CEH derived from a soil metagenomic library (Tao *et al.*, 2011). However, no structural data of this enzyme is available, thus structural elucidation of this CE13 homolog can help to understand the high organic solvent tolerance of CE13.

Among the CEHs investigated in this thesis, multiple enzymes with a broad substrate spectrum were identified. This feature is on the one hand beneficial for a broad applicability of the biocatalysts, but on the other hand comes along with low chiral selectivity, which is crucial for applications where enantiopure products are desired. Interestingly, CE07 showed selectivity for some chiral esters when the enantiomers were provided separately. Real applications however use racemic mixtures of substrates for kinetic resolution. Thus, CE07 must be tested with racemic substrates in the future to evaluate potential application for kinetic resolution. Moreover, a

mutagenesis of the enzyme to increase selectivity is possible with well-established methodologies of protein engineering (Li and Reetz, 2016). In future studies, a crystal structure of CE07, best in complex with a chiral substrate analog, will be of great help to understand the substrate specificity of this enzyme in detail.

In case of the polyester hydrolytic enzyme CE16 (named PE-H in chapter 2.7), different possible applications could be suggested based on biochemical characteristics and knowledge of the three-dimensional structure. Further studies can now evaluate the enzymes performance under real and simulated application conditions. In addition, the diversity of CE16 homologs among bacteria of the *P. pertucinogena* lineage (chapter 2.2; (Bollinger, Thies, Katzke, *et al.*, 2020)) has not yet been investigated in detail, but significant differences in the polyester hydrolytic activity of the organisms were observed (chapter 2.6; (Molitor *et al.*, 2020)), which suggest differences in the respective biocatalysts. Revealing these differences can further deepen our understanding of the polyester hydrolytic mechanism of this class of enzymes. The investigation of a number of polyester hydrolytic enzymes is especially important, since no structure of such an enzyme in complex with its polymeric substrate exists yet, thus a set of different enzymes' structures, which are able to bind the same polymeric substrate, can help to develop a reliable model for polyester hydrolysis.

Another strategy that is promising for the future design of a next generation biocatalyst is the introduction of additional active sites to an existing enzyme scaffold (Santiago *et al.*, 2018). These "plurizymes" have been shown to significantly alter an enzyme's substrate spectrum and even allow to equip an enzyme with completely new catalytic mechanisms, as exemplified by introduction of a chemocatalytic site for Friedel-Crafts alkylation into an artificial esterase active site (Alonso *et al.*, 2020). The highly stable biocatalysts identified in this thesis are best suited as scaffolds for such modifications, since stable proteins allow the alteration of a higher number of amino acid residues (Socha and Tokuriki, 2013).

4 References

Agrawal, S., Acharya, D., Adholeya, A., Barrow, C.J., and Deshmukh, S.K. (2017) Nonribosomal peptides from marine microbes and their antimicrobial and anticancer potential. *Front. Pharmacol.* **8**: 828.

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., *et al*. (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat. Biotechnol.* **37**: 420–423.

Almarsson, O. and Klibanov, A.M. (1996) Remarkable activation of enzymes in nonaqueous media by denaturing organic cosolvents. *Biotechnol. Bioeng.* **49**: 87–92.

Alonso, S., Santiago, G., Cea-Rama, I., Fernandez-Lopez, L., Coscolín, C., Modregger, J., *et al*. (2020) Genetically engineered proteins with two active sites for enhanced biocatalysis and synergistic chemo- and biocatalysis. *Nat. Catal.* **3**: 319–328.

Amann, R.I., Ludwig, W., and Schleifer, K.H. (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**: 143–169.

Arnold, F.H. (1990) Engineering enzymes for non-aqueous solvents. *Trends Biotechnol.* **8**: 244–249.

Arpigny, J.L. and Jaeger, K.E. (1999) Bacterial lipolytic enzymes: classification and properties. *Biochem. J.* **343 Pt 1**: 177–183.

Austin, H.P., Allen, M.D., Donohoe, B.S., Rorrer, N.A., Kearns, F.L., Silveira, R.L., *et al*. (2018) Characterization and engineering of a plastic-degrading aromatic polyesterase. *Proc. Natl. Acad. Sci. U. S. A.* **115**: E4350–E4357.

Beilen, J.B. Van, Li, Z., and Duetz, W.A. (2003) Diversity of alkane hydroxylase systems in the environment. **58**: 427–440.

Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., and Potrykus, I. (2002) Golden rice: introducing the β-carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin A deficiency. *J. Nutr.* **132**: 506S–510S.

Bhatia, S. (2018) History, scope and development of biotechnology. In, *Introduction to Pharmaceutical Biotechnology*. IOP Publishing, pp. 1–61.

Bollinger, A., Molitor, R., Thies, S., Koch, R., Coscolín, C., Ferrer, M., and Jaeger, K. (2020) Identification of organic solvent tolerant carboxylic ester hydrolases for organic synthesis. *Appl. Environ. Microbiol.* in press.

Bollinger, A., Thies, S., Katzke, N., and Jaeger, K.-E. (2020) The biotechnological potential of marine bacteria in the novel lineage of *Pseudomonas pertucinogena*. *Microb. Biotechnol.* **13**: 19–31.

Bollinger, A., Thies, S., Knieps-Grünhagen, E., Gertzen, C., Kobus, S., Höppner, A., *et al*. (2020) A novel polyester hydrolase from the marine bacterium *Pseudomonas aestusnigri* - structural and functional insights. *Front. Microbiol.* **11**: 114.

Bornscheuer, U.T. (2002) Microbial carboxyl esterases: classification, properties and application in biocatalysis.

FEMS Microbiol. Rev. **26**: 73–81.

Bringmann, G., Gulder, T., Lang, G., Schmitt, S., Stöhr, R., Wiese, J., *et al*. (2007) Large-scale biotechnological production of the antileukemic marine natural product sorbicillactone A. *Mar. Drugs* **5**: 23–30.

Bringmann, G., Lang, G., Gulder, T.A.M., Tsuruta, H., Mühlbacher, J., Maksimenka, K., *et al*. (2005) The first sorbicillinoid alkaloids, the antileukemic sorbicillactones A and B, from a sponge-derived *Penicillium chrysogenum* strain. *Tetrahedron* **61**: 7252–7265.

Brooijmans, R.J.W., Pastink, M.I., and Siezen, R.J. (2009) Genomics update hydrocarbon-degrading bacteria : the oil-spill clean-up crew. **2**: 587–594.

Bulka, M. (2019) Reinigung und Charakterisierung einer Esterase aus *Pseudomonas aestusnigri* für biotechnologische Anwendungen. Master thesis, Heinrich Heine University Düsseldorf, Germany.

Cafaro, V., Izzo, V., Notomista, E., and Di Donato, A. (2013) Marine hydrocarbonoclastic bacteria. In, Trincone,A. (ed), *Marine Enzymes for Biocatalysis*. Woodhead Publishing, pp. 373–402.

Casas-Godoy, L., Duquesne, S., Bordes, F., Sandoval, G., and Marty, A. (2012) Lipases: An Overview. In, Sandoval, G. (ed), *Lipases and Phospholipases. Methods in Molecular Biology (Methods and Protocols)*. Humana Press, pp. 3–30.

Castro-Ochoa, L.D., Rodríguez-Gómez, C., Valerio-Alfaro, G., and Oliart Ros, R. (2005) Screening, purification and characterization of the thermoalkalophilic lipase produced by *Bacillus thermoleovorans* CCR11. *Enzyme Microb. Technol.* **37**: 648–654.

Chakraborty, K. and Raj, R.P. (2008) An extra-cellular alkaline metallolipase from *Bacillus licheniformis* MTCC 6824: purification and biochemical characterization. *Food Chem.* **109**: 727–736.

Chen, S. and Zückert, W.R. (2011) Probing the *Borrelia burgdorferi* surface lipoprotein secretion pathway using a conditionally folding protein domain. *J. Bacteriol.* **193**: 6724–6732.

Chen, Y., Black, D.S., and Reilly, P.J. (2016) Carboxylic ester hydrolases: classification and database derived from their primary, secondary, and tertiary structures. *Protein Sci.* **25**: 1942–1953.

Colón, W., Church, J., Sen, J., Thibeault, J., Trasatti, H., and Xia, K. (2017) Biological roles of protein kinetic stability. *Biochemistry* **56**: 6179–6186.

Copley, S.D. (2003) Enzymes with extra talents: moonlighting functions and catalytic promiscuity. *Curr. Opin. Chem. Biol.* **7**: 265–272.

Copley, S.D. (2017) Shining a light on enzyme promiscuity. *Curr. Opin. Struct. Biol.* **47**: 167–175.

Coscolín, C., Bargiela, R., Martínez-Martínez, M., Alonso, S., Bollinger, A., Thies, S., *et al*. (2018) Hydrocarbon-degrading microbes as sources of new biocatalysts. In, McGenity, T.J. (ed), *Taxonomy, Genomics and Ecophysiology of Hydrocarbon-Degrading Microbes*. Springer, Cham, Cham, pp. 353–373.

Coscolín, C., Martínez-Martínez, M., Chow, J., Bargiela, R., García-Moyano, A., Bjerga, G., *et al*. (2018) Relationships between substrate promiscuity and chiral selectivity of esterases from phylogenetically and

environmentally diverse microorganisms. *Catalysts* **8**: 10.

Danso, D., Chow, J., and Streita, W.R. (2019) Plastics: environmental and biotechnological perspectives on microbial degradation. *Appl. Environ. Microbiol.* **85**: 1–14.

Danso, D., Schmeisser, C., Chow, J., Zimmermann, W., Wei, R., Leggewie, C., *et al*. (2018) New insights into the function and global distribution of polyethylene terephthalate (PET)-degrading bacteria and enzymes in marine and terrestrial metagenomes. *Appl. Environ. Microbiol.* **84**: e02773-17.

Doukyu, N. and Ogino, H. (2010) Organic solvent-tolerant enzymes. *Biochem. Eng. J.* **48**: 270–282.

Ereky, K. (1919) Biotechnologie der Fleisch-, Fett-, und Milcherzeugung im landwirtschaftlichen Grossbetriebe: für naturwissenschaftlich gebildete Landwirte verfasst. P. Parey, Berlin, Germany.

Fári, M.G. and Kralovánszky, U.P. (2006) The founding father of biotechnology: Károly (Karl) Ereky. *Int. J. Hortic. Sci.* **12**: 9–12.

Fecker, T., Galaz-Davison, P., Engelberger, F., Narui, Y., Sotomayor, M., Parra, L.P., and Ramírez-Sarmiento, C.A. (2018) Active site flexibility as a hallmark for efficient PET degradation by *I. sakaiensis* PETase. *Biophys. J.* **114**: 1302–1312.

Ferrer, M., Bargiela, R., Martínez-Martínez, M., Mir, J., Koch, R., Golyshina, O. V., and Golyshin, P.N. (2015) Biodiversity for biocatalysis: a review of the α/β-hydrolase fold superfamily of esterases-lipases discovered in metagenomes. *Biocatal. Biotransformation* **33**: 235–249.

Ferrer, M., Chernikova, T.N., Yakimov, M.M., Golyshin, P.N., and Timmis, K.N. (2003) Chaperonins govern growth of *Escherichia coli* at low temperatures. *Nat. Biotechnol.* **21**: 1266–1267.

Ferrer, M., Golyshina, O. V., Chernikova, T.N., Khachane, A.N., Martins dos Santos, V.A.P., Yakimov, M.M., *et al*. (2005) Microbial enzymes mined from the Urania deep-sea hypersaline anoxic basin. *Chem. Biol.* **12**: 895–904.

Ferrer, M., Martínez-Martínez, M., Bargiela, R., Streit, W.R., Golyshina, O. V., and Golyshin, P.N. (2015) Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. *Microb. Biotechnol.* **9**: 22–34.

Ferrer, M., Méndez-García, C., Bargiela, R., Chow, J., Alonso, S., García-Moyano, A., *et al*. (2019) Decoding the ocean's microbiological secrets for marine enzyme biodiscovery. *FEMS Microbiol. Lett.* **366**: fny285.

Fischer, E. (1894) Einfluss der Configuration auf die Wirkung der Enzyme. *Berichte der Dtsch. Chem. Gesellschaft* **27**: 2985–2993.

Fleming, A. (1929) On the antibacterial action of cultures of a *Penicillium*, with special reference to their use in the isolation of *B. influenzæ*. *Br. J. Exp. Pathol.* **10**: 226–236.

Frauenkron-Machedjou, V.J., Fulton, A., Zhao, J., Weber, L., Jaeger, K.-E., Schwaneberg, U., and Zhu, L. (2018) Exploring the full natural diversity of single amino acid exchange reveals that 40 – 60 % of BSLA positions improve organic solvents resistance. *Bioresour. Bioprocess.* **5**: 2.

Frauenkron-Machedjou, V.J., Fulton, A., Zhu, L., Anker, C., Bocola, M., Jaeger, K.E., and Schwaneberg, U. (2015) Towards understanding directed evolution: more than half of all amino acid positions contribute to ionic liquid resistance of *Bacillus subtilis* lipase A. *ChemBioChem* **16**: 937–945.

Fulton, A., Frauenkron-Machedjou, V.J., Skoczinski, P., Wilhelm, S., Zhu, L., Schwaneberg, U., and Jaeger, K.E. (2015) Exploring the protein stability landscape: *Bacillus subtilis* lipase A as a model for detergent tolerance. *ChemBioChem* **16**: 930–936.

Furukawa, M., Kawakami, N., Oda, K., and Miyamoto, K. (2018) Acceleration of enzymatic degradation of poly(ethylene terephthalate) by surface coating with anionic surfactants. *ChemSusChem*. **11:** 4018.

Furukawa, M., Kawakami, N., Tomizawa, A., and Miyamoto, K. (2019) Efficient degradation of poly(ethylene terephthalate) with *Thermobifida fusca* cutinase exhibiting improved catalytic activity generated using mutagenesis and additive-based approaches. *Sci. Rep.* **9**: 16038.

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. In, Walker, J.M. (ed), *The Proteomics Protocols Handbook*. Humana Press, Totowa, NJ, pp. 571–607.

Golyshin, P.N., Martins Dos Santos, V.A.P., Kaiser, O., Ferrer, M., Sabirova, Y.S., Lünsdorf, H., *et al*. (2003) Genome sequence completed of *Alcanivorax borkumensis*, a hydrocarbon-degrading bacterium that plays a global role in oil removal from marine systems. *J. Biotechnol.* **106**: 215–220.

Gomila, M., Mulet, M., Lalucat, J., and García-Valdés, E. (2017) Draft genome sequence of the marine bacterium *Pseudomonas aestusnigri* VGXO14^{T}. *Genome Announc.* **5**: e00765-17.

Gupta, A. and Khare, S.K. (2009) Enzymes from solvent-tolerant microbes: useful biocatalysts for non-aqueous enzymology. *Crit. Rev. Biotechnol.* **29**: 44–54.

Guzmán, G.I., Sandberg, T.E., LaCroix, R.A., Nyerges, Á., Papp, H., de Raad, M., *et al*. (2019) Enzyme promiscuity shapes adaptation to novel growth substrates. *Mol. Syst. Biol.* **15**: 1–14.

Haernvall, K., Zitzenbacher, S., Biundo, A., Yamamoto, M., Schick, M.B., Ribitsch, D., and Guebitz, G.M. (2018) Enzymes as enhancers for the biodegradation of synthetic polymers in wastewater. *Chembiochem* **19**: 317–325.

Hajighasemi, M., Nocek, B.P., Tchigvintsev, A., Brown, G., Flick, R., Xu, X., *et al*. (2016) Biochemical and structural insights into enzymatic depolymerization of polylactic acid and other polyesters by microbial carboxylesterases. *Biomacromolecules* **17**: 2027–2039.

Hajighasemi, M., Tchigvintsev, A., Nocek, B., Flick, R., Popovic, A., Hai, T., *et al*. (2018) Screening and characterization of novel polyesterases from environmental metagenomes with high hydrolytic activity against synthetic polyesters. *Environ. Sci. Technol.* **52**: 12388–12401.

Hameed, A., Al-Rashida, M., Uroos, M., Ali, S.A., Arshia, Ishtiaq, M., and Khan, K.M. (2018) Quinazoline and quinazolinone as important medicinal scaffolds: a comparative patent review (2011-2016). *Expert Opin. Ther. Pat.* **28**: 281–297.

Han, X., Liu, W., Huang, J.-W., Ma, J., Zheng, Y., Ko, T.-P., *et al*. (2017) Structural insight into catalytic mechanism of PET hydrolase. *Nat. Commun.* **8**: 2106.

Head, I.M., Jones, D.M., and Röling, W.F.M. (2006) Marine microorganisms make a meal of oil. *Nat. Rev. Microbiol.* **4**: 173–182.

Hiraga, K., Taniguchi, I., Yoshida, S., Kimura, Y., and Oda, K. (2019) Biodegradation of waste PET: a sustainable solution for dealing with plastic pollution. *EMBO Rep.* **20**: e49365.

Holmquist, M. (2000) Alpha/Beta-hydrolase fold enzymes: structures, functions and mechanisms. *Curr. Protein Pept. Sci.* **1**: 209–235.

Hughes, A.L. (1994) The evolution of functionally novel proteins after gene duplication. *Proceedings. Biol. Sci.* **256**: 119–124.

Hult, K. and Berglund, P. (2007) Enzyme promiscuity : mechanism and applications. **25**: 231–238.

Hunsen, M., Azim, A., Mang, H., Wallner, S.R., Ronkvist, A., Wenchun, X., and Gross, R.A. (2007) A cutinase with polyester synthesis activity. *Macromolecules* **40**: 148–150.

Isken, S. and Heipieper, H.J. (2003) Toxicity of organic solvents to microorganisms. In, Bitton,G. (ed), *Encyclopedia of Environmental Microbiology*. American Cancer Society, pp. 3147–3155.

Jaeger, K.-E. and Eggert, T. (2002) Lipases for biotechnology. *Curr. Opin. Biotechnol.* **13**: 390–397.

Jaeger, K.E. and Eggert, T. (2004) Enantioselective biocatalysis optimized by directed evolution. *Curr. Opin. Biotechnol.* **15**: 305–313.

Jensen, R.A. (1976) Enzyme recruitment in evolution of new function. *Annu. Rev. Microbiol.* **30**: 409–425.

Ji, Y., Mao, G., Wang, Y., and Bartlam, M. (2013) Structural insights into diversity and *n*-alkane biodegradation mechanisms of alkane hydroxylases. *Front. Microbiol.* **4**: 58.

Johnson, I.S. (1983) Human insulin from recombinant DNA technology. *Science* **219**: 632–637.

Joo, S., Cho, I.J., Seo, H., Son, H.F., Sagong, H.-Y., Shin, T.J., *et al*. (2018) Structural insight into molecular mechanism of poly(ethylene terephthalate) degradation. *Nat. Commun.* **9**: 382.

Kamal, M.Z., Yedavalli, P., Deshmukh, M. V., and Rao, N.M. (2013) Lipase in aqueous-polar organic solvents: activity, structure, and stability. *Protein Sci.* **22**: 904–915.

Kawai, F., Kawabata, T., and Oda, M. (2019) Current knowledge on enzymatic PET degradation and its possible application to waste stream management and other fields. *Appl. Microbiol. Biotechnol.* **103**: 4253–4268.

Kawata, T. and Ogino, H. (2009) Enhancement of the organic solvent-stability of the LST-03 lipase by directed evolution. *Biotechnol. Prog.* **25**: 1605–1611.

Kazlauskas, R. (2018) Engineering more stable proteins. *Chem. Soc. Rev.* **47**: 9026–9045.

Kelley, L.A., Mezulis, S., Yates, C.M., Wass, M.N., and Sternberg, M.J.E. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**: 845–858.

Klibanov, A.M. (2001) Improving enzymes by using them in organic solvents. *Nature* **409**: 241–246.

Korman, T.P., Sahachartsiri, B., Charbonneau, D.M., Huang, G.L., Beauregard, M., and Bowie, J.U. (2013) Dieselzymes: development of a stable and methanol tolerant lipase for biodiesel production by directed evolution. *Biotechnol. Biofuels* **6**: 70.

Kovacic, F., Babic, N., Krauss, U., and Jaeger, K.-E. (2019) Classification of lipolytic enzymes from bacteria. In, Rojo,F. (ed), *Aerobic Utilization of Hydrocarbons, Oils and Lipids*. Springer International Publishing, Cham, pp. 1–35.

Kubicki, S., Bollinger, A., Katzke, N., Jaeger, K., Loeschcke, A., and Thies, S. (2019) Marine biosurfactants: biosynthesis, structural diversity and biotechnological applications. *Mar. Drugs* **17**: 1–30.

Kuklinski, A. (2017) Identifizierung neuer Esterasen aus *Pseudomonas aestusnigri* und Evaluierung für biotechnologische Anwendungen. Bachelor thesis, Heinrich Heine University Düsseldorf, Germany.

Kumar, A., Dhar, K., Kanwar, S.S., and Arora, P.K. (2016) Lipase catalysis in organic solvents: advantages and applications. *Biol. Proced. Online* **18**: 2.

Laane, C., Boeren, S., Vos, K., and Veeger, C. (1987) Rules for optimization of biocatalysis in organic solvents. *Biotechnol. Bioeng.* **30**: 81–87.

Lang, D.A., Mannesse, M.L.M., De Haas, G.H., Verheij, H.M., and Dijkstra, B.W. (1998) Structural basis of the chiral selectivity of *Pseudomonas cepacia* lipase. *Eur. J. Biochem.* **254**: 333–340.

Lebreton, L., Slat, B., Ferrari, F., Sainte-Rose, B., Aitken, J., Marthouse, R., *et al*. (2018) Evidence that the Great Pacific Garbage Patch is rapidly accumulating plastic. *Sci. Rep.* **8**: 4666.

Li, G. and Reetz, M.T. (2016) Learning lessons from directed evolution of stereoselective enzymes. *Org. Chem. Front.* **3**: 1350–1358.

Liebl, W., Angelov, A., Juergensen, J., Chow, J., Loeschcke, A., Drepper, T., *et al*. (2014) Alternative hosts for functional (meta)genome analysis. *Appl. Microbiol. Biotechnol.* **98**: 8099–8109.

Lorenz, P. and Eck, J. (2005) Metagenomics and industrial applications. *Nat. Rev. Microbiol.* **3**: 510–516.

Lu, M., Dukunde, A., and Daniel, R. (2019) Biochemical profiles of two thermostable and organic solvent-tolerant esterases derived from a compost metagenome. *Appl. Microbiol. Biotechnol.*

Lu, Y., Lu, F., Wang, X., Bie, X., Sun, H., Wuyundalai, and Lu, Z. (2009) Identification of bacteria producing a thermophilic lipase with positional non-specificity and characterization of the lipase. *Ann. Microbiol.* **59**: 565–571.

Madhavan, A., Sindhu, R., Parameswaran, B., Sukumaran, R.K., and Pandey, A. (2017) Metagenome analysis: a powerful tool for enzyme bioprospecting. *Appl. Biochem. Biotechnol.* 1–16.

Manning, M. and Colón, W. (2004) Structural basis of protein kinetic stability: resistance to sodium dodecyl sulfate suggests a central role for rigidity and a bias toward beta-sheet structure. *Biochemistry* **43**: 11248–11254.

Marshall, A.G. and Rodgers, R.P. (2004) Petroleomics: the next grand challenge for chemical analysis. *Acc. Chem. Res.* **37**: 53–59.

Martínez-Martínez, M., Coscolín, C., Santiago, G., Chow, J., Stogios, P.J., Bargiela, R., *et al*. (2018) Determinants and prediction of esterase substrate promiscuity patterns. *ACS Chem. Biol.* **13**: 225–234.

Martinez, A., Kolvek, S.J., Yip, C.L.T., Hopke, J., Brown, K.A., MacNeil, I.A., and Osburne, M.S. (2004) Genetically modified bcterial strains and novel bacterial artificial chromosome shuttle vectors for constructing environmental libraries and detecting heterologous natural products in multiple expression hosts. *Appl. Environ. Microbiol.* **70**: 2452–2463.

Masomian, M., Rahman, R.N.Z.R.A., Salleh, A.B., and Basri, M. (2013) A new thermostable and organic solvent-tolerant lipase from *Aneurinibacillus thermoaerophilus* strain HZ. *Process Biochem.* **48**: 169–175.

Mezzetti, A., Schrag, J.D., Cheong, C.S., and Kazlauskas, R.J. (2005) Mirror-image packing in enantiomer discrimination: molecular basis for the enantioselectivity of *B. cepacia* lipase toward 2-methyl-3-phenyl-1-propanol. *Chem. Biol.* **12**: 427–437.

Mohtashami, M., Fooladi, J., Haddad-Mashadrizeh, A., Housaindokht, M.R., and Monhemi, H. (2018) Molecular mechanism of enzyme tolerance against organic solvents: insights from molecular dynamics simulation. *Int. J. Biol. Macromol.* **122**: 914–923.

Molitor, R., Bollinger, A., Kubicki, S., Loeschcke, A., Jaeger, K., and Thies, S. (2020) Agar plate-based screening methods for the identification of polyester hydrolysis by *Pseudomonas* species. *Microb. Biotechnol.* **13**: 274–284.

Monsef Shokri, M., Ahmadian, S., Akbari, N., and Khajeh, K. (2014) Hydrophobic substitution of surface residues affects lipase stability in organic solvents. *Mol. Biotechnol.* **56**: 360–368.

Moog, D., Schmitt, J., Senger, J., Zarzycki, J., Rexer, K.-H., Linne, U., *et al*. (2019) Using a marine microalga as a chassis for polyethylene terephthalate (PET) degradation. *Microb. Cell Fact.* **18**: 171.

Mulnaes, D. and Gohlke, H. (2018) TopScore: using deep neural networks and large diverse data sets for accurate protein model quality assessment. *J. Chem. Theory Comput.* **14**: 6117–6126.

Nguyen, L.A., He, H., and Pham-Huy, C. (2006) Chiral drugs: an overview. *Int. J. Biomed. Sci.* **2**: 85–100.

Nikolaivits, E., Kanelli, M., Dimarogona, M., and Topakas, E. (2018) A middle-aged enzyme still in its prime: recent advances in the field of cutinases. *Catalysts* **8**: 612.

Nobeli, I., Favia, A.D., and Thornton, J.M. (2009) Protein promiscuity and its implications for biotechnology. *Nat. Biotechnol.* **27**: 157–167.

O'Brien, P.J. and Herschlag, D. (1999) Catalytic promiscuity and the evolution of new enzymatic activities. *Chem. Biol.* **6**: R91–R105.

OECD (2017) Marine biotechnology. *OECD Science, Technology and Industry Policy Papers* **43:** 1–51.

Ogino, H., Miyamoto, K., and Ishikawa, H. (1994) Organic-solvent-tolerant bacterium which secretes organic-

solvent-stable lipolytic enzyme. *Appl. Environ. Microbiol.* **60**: 3884–3886.

Ogino, H., Nakagawa, S., Shinya, K., Muto, T., Fujimura, N., Yasuda, M., and Ishikawa, H. (2000) Purification and characterization of organic solvent-stable lipase from organic solvent-tolerant *Pseudomonas aeruginosa* LST-03. *J. Biosci. Bioeng.* **89**: 451–457.

Okuda, S. and Tokuda, H. (2011) Lipoprotein sorting in bacteria. *Annu. Rev. Microbiol.* **65**: 239–259.

Olsen, H.S. and Falholt, P. (1998) The role of enzymes in modern detergency. *J. Surfactants Deterg.* **1**: 555–567.

Pace, C.N., Fu, H., Fryar, K.L., Landua, J., Trevino, S.R., Shirley, B.A., *et al*. (2011) Contribution of hydrophobic interactions to protein stability. *J. Mol. Biol.* **408**: 514–528.

Parisi, D., Riley, C., Srivastava, A.S., McCue, H. V., Johnson, J.R., and Carnell, A.J. (2019) PET hydrolysing enzymes catalyse bioplastics precursor synthesis under aqueous conditions. *Green Chem.* **21:** 3827-3833.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004) UCSF Chimera — a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**: 1605–1612.

PlasticsEurope (2018) Plastics – the facts Brussels, Belgium. Available online at: https://www.plasticseurope.org/application/files/6315/4510/9658/Plastics_the_facts_2018_AF_web.pdf

Reetz, M.T., Soni, P., Fernández, L., Gumulya, Y., and Carballeira, J.D. (2010) Increasing the stability of an enzyme toward hostile organic solvents by directed evolution based on iterative saturation mutagenesis using the B-FIT method. *Chem. Commun.* **46**: 8657.

Riegels, M., Koch, R., Pedersen, L., and Lund, H. (1997) Enzymatic hydrolysis of cyclic oligomers. *World Intellect. Prop. Organ.* WO 97/27237.

Rmili, F., Achouri, N., Smichi, N., Krayem, N., Bayoudh, A., Gargouri, Y., *et al*. (2019) Purification and biochemical characterization of an organic solvent-tolerant and detergent-stable lipase from *Staphylococcus capitis*. *Biotechnol. Prog.* **35:** e2833.

Rojo, F. (2009) Degradation of alkanes by bacteria. *Environ. Microbiol.* **11**: 2477–2490.

Roy, I., Mukherjee, J., and Gupta, M.N. (2017) Cross-linked enzyme aggregates for applications in aqueous and nonaqueous media. *Methods Mol. Biol.* **1504**: 109–123.

Sabree, Z.L., Rondon, M.R., and Handelsman, J. (2009) Metagenomics. In, Schaechter,M. (ed), *Encyclopedia of Microbiology*. Academic Press, pp. 622–632.

Sagar, S., Kaur, M., and Minneman, K.P. (2010) Antiviral lead compounds from marine sponges. *Mar. Drugs* **8**: 2619–2638.

Salihu, A. and Alam, M.Z. (2015) Solvent tolerant lipases: A review. *Process Biochem.* **50**: 86–96.

Salvador, M., Abdulmutalib, U., Gonzalez, J., Kim, J., Smith, A.A., Faulon, J.L., *et al*. (2019) Microbial genes for a circular and sustainable bio-PET economy. *Genes* **10**: 373.

Sandoval, G. and Marty, A. (2007) Screening methods for synthetic activity of lipases. *Enzyme Microb. Technol.*

40: 390–393.

Santiago, G., Martínez-Martínez, M., Alonso, S., Bargiela, R., Coscolín, C., Golyshin, P.N., *et al*. (2018) Rational engineering of multiple active sites in an ester hydrolase. *Biochemistry* **57**: 2245–2255.

Schneiker, S., Martins dos Santos, V.A.P., Bartels, D., Bekel, T., Brecht, M., Buhrmester, J., *et al*. (2006) Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat. Biotechnol.* **24**: 997–1004.

Schoevaart, R., Wolbers, M.W., Golubovic, M., Ottens, M., Kieboom, A.P.G., Van Rantwijk, F., *et al*. (2004) Preparation, optimization, and structures, of cross-linked enzyme aggregates (CLEAs). *Biotechnol. Bioeng.* **87**: 754–762.

Schofield, M.M., Jain, S., Porat, D., Dick, G.J., and Sherman, D.H. (2015) Identification and analysis of the bacterial endosymbiont specialized for production of the chemotherapeutic natural product ET-743. *Environ. Microbiol.* **17**: 3964–3975.

Schulz, D., Passeri, A., Schmidt, M., Lang, S., Wagner, F., Wray, V., and Gunkel, W. (1991) Marine Biosurfactants, I. Screening for biosurfactants among crude oil degrading marine microorganisms from the North Sea. *Zeitschrift für Naturforsch. C* **46**: 197–203.

Sehnal, D., Svobodová Vařeková, R., Berka, K., Pravda, L., Navrátilová, V., Banáš, P., *et al*. (2013) MOLE 2.0: advanced approach for analysis of biomacromolecular channels. *J. Cheminform.* **5**: 39.

Sheldon, R.A. and Woodley, J.M. (2018) Role of biocatalysis in sustainable chemistry. *Chem. Rev.* **118**: 801–838.

Shimomura, O. (2005) The discovery of aequorin and green fluorescent protein. *J. Microsc.* **217**: 1–15.

Shimomura, O., Johnson, F.H., and Saiga, Y. (1962) Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, *Aequorea*. *J. Cell. Comp. Physiol.* **59**: 223–239.

Silber, J., Kramer, A., Labes, A., and Tasdemir, D. (2016) From discovery to production: biotechnology of marine fungi for the production of new antibiotics. *Mar. Drugs* **14**: 137.

Simon, C. and Daniel, R. (2011) Metagenomic analyses: past and future trends. *Appl. Environ. Microbiol.* **77**: 1153–1161.

Singh, R., Kumar, M., Mittal, A., and Mehta, P.K. (2016) Microbial enzymes: industrial progress in 21st century. *3 Biotech* **6**: 174.

Socha, R.D. and Tokuriki, N. (2013) Modulating protein stability - directed evolution strategies for improved protein function. *FEBS J.* **280**: 5582–5595.

Söltl, S. (2017) Produktion, Reinigung und biochemische Charakterisierung neuer Esterasen aus dem hydrocarbonoklastischen Bakterium *Alcanivorax borkumensis*. Bachelor thesis, FH Aachen University of Applied Sciences, Germany.

Steele, C.L., Crock, J., Bohlmann, J., and Croteau, R. (1998) Sesquiterpene synthases from grand fir (*Abies grandis*). Comparison of constitutive and wound-induced activities, and cDNA isolation, characterization, and

bacterial expression of delta-selinene synthase and gamma-humulene synthase. *J. Biol. Chem.* **273**: 2078–2089.

Stepankova, V., Bidmanova, S., Koudelakova, T., Prokop, Z., Chaloupkova, R., and Damborsky, J. (2013) Strategies for stabilization of enzymes in organic solvents. *ACS Catal.* **3**: 2823–2836.

Sterling, T. and Irwin, J.J. (2015) ZINC 15 - ligand discovery for everyone. *J. Chem. Inf. Model.* **55**: 2324–2337.

Streit, W.R. and Schmitz, R.A. (2004) Metagenomics - the key to the uncultured microbes. *Curr. Opin. Microbiol.* **7**: 492–498.

Sun, Z., Lonsdale, R., Wu, L., Li, G., Li, A., Wang, J., *et al*. (2016) Structure-guided triple-code saturation mutagenesis: efficient tuning of the stereoselectivity of an epoxide hydrolase. *ACS Catal.* **6**: 1590–1597.

Taniguchi, I., Yoshida, S., Hiraga, K., Miyamoto, K., Kimura, Y., and Oda, K. (2019) Biodegradation of PET: current status and application aspects. *ACS Catal.* **9**: 4089–4105.

Tao, W., Lee, M.H., Wu, J., Kim, N.H., and Lee, S.-W. (2011) Isolation and characterization of a family VII esterase derived from alluvial soil metagenomic library. *J. Microbiol.* **49**: 178–185.

Tchigvintsev, A., Tran, H., Popovic, A., Kovacic, F., Brown, G., Flick, R., *et al*. (2015) The environment shapes microbial enzymes: five cold-active and salt-resistant carboxylesterases from marine metagenomes. *Appl. Microbiol. Biotechnol.* **99**: 2165–2178.

Thies, S., Rausch, S.C., Kovacic, F., Schmidt-thaler, A., Wilhelm, S., Rosenau, F., *et al*. (2016) Metagenomic discovery of novel enzymes and biosurfactants in a slaughterhouse biofilm microbial community. *Sci. Rep.* **6**: 27035.

Tokuriki, N., Stricher, F., Serrano, L., and Tawfik, D.S. (2008) How protein stability and new functions trade off. *PLoS Comput. Biol.* **4**: 35–37.

Tomić, S., Dobovičnik, V., Šunjić, V., and Kojić-Prodić, B. (2001) Enantioselectivity of *Pseudomonas cepacia* lipase towards 2-methyl-3(or 4)-arylalkanols: an approach based on the stereoelectronic theory and molecular modeling. *Croat. Chem. Acta* **74**: 343–357.

Trincone, A. (2011) Marine biocatalysts: enzymatic features and applications. *Mar. Drugs* **9**: 478–499.

Troeschel, S.C., Thies, S., Link, O., Real, C.I., Knops, K., Wilhelm, S., *et al*. (2012) Novel broad host range shuttle vectors for expression in *Escherichia coli*, *Bacillus subtilis* and *Pseudomonas putida*. *J. Biotechnol.* **161**: 71–79.

Trott, O. and Olson, A.J. (2009) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**: 455–461.

Tully, B.J., Graham, E.D., and Heidelberg, J.F. (2018) The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Sci. Data* **5**: 170203.

Tuomi, W.V. and Kazlauskas, R.J. (1999) Molecular basis for enantioselectivity of lipase from *Pseudomonas cepacia* toward primary alcohols. Modeling, kinetics, and chemical modification of Tyr29 to increase or

decrease enantioselectivity. *J. Org. Chem.* **64**: 2638–2647.

Turkes, L. (2019) Expression, Reinigung und vergleichende Charakterisierung von Polyesterasevarianten aus marinen *Pseudomonas* species. Bachelor thesis, Heinrich Heine University Düsseldorf, Germany.

Venter, J.C., Remington, K., Heidelberg, J.F., Halpern, A.L., Rusch, D., Eisen, J.A., *et al*. (2004) Environmental genome shotgun sequencing of the Sargasso Sea. *Science* **304**: 66–74.

Webb, H.K., Arnott, J., Crawford, R.J., and Ivanova, E.P. (2013) Plastic degradation and its environmental implications with special reference to poly(ethylene terephthalate). *Polymers* **5**: 1–18.

Wei, R., Breite, D., Song, C., Gräsing, D., Ploss, T., Hille, P., *et al*. (2019) Biocatalytic degradation efficiency of postconsumer polyethylene terephthalate packaging determined by their polymer microstructures. *Adv. Sci.* **6**: 1900491.

Wei, R. and Zimmermann, W. (2017a) Biocatalysis as a green route for recycling the recalcitrant plastic polyethylene terephthalate. *Microb. Biotechnol.* **10**: 1302–1307.

Wei, R. and Zimmermann, W. (2017b) Microbial enzymes for the recycling of recalcitrant petroleum-based plastics: how far are we? *Microb. Biotechnol.* **10**: 1308–1322.

Weide, D. (2018) Reinigung einer Esterase mit breitem Substratspektrum aus *Alcanivorax borkumensis*. Bachelor thesis, Heinrich Heine University Düsseldorf, Germany.

Wierckx, N., Narancic, T., Eberlein, C., Wei, R., Drzyzga, O., Magnin, A., *et al*. (2018) Plastic Biodegradation: challenges and opportunities. In, Steffan,R. (ed), *Consequences of Microbial Interactions with Hydrocarbons, Oils, and Lipids: Biodegradation and Bioremediation. Handbook of Hydrocarbon and Lipid Microbiology.* Springer, Cham, pp. 1–29.

Wilson, M.M. and Bernstein, H.D. (2016) Surface-exposed lipoproteins: an emerging secretion phenomenon in Gram-negative bacteria. *Trends Microbiol.* **24**: 198–208.

Wright, R.J., Bosch, R., Gibson, M.I., and Christie-Oleza, J.A. (2020) Plasticizer degradation by marine bacterial isolates: a proteogenomic and metabolomic characterization. *Environ. Sci. Technol*. **54**: 2244–2256.

Yakimov, M.M., Golyshin, P.N., Lang, S., Moore, E.R., Abraham, W.R., Lünsdorf, H., and Timmis, K.N. (1998) *Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing marine bacterium. *Int. J. Syst. Bacteriol.* **48 Pt 2**: 339–348.

Yakimov, M.M., Timmis, K.N., and Golyshin, P.N. (2007) Obligate oil-degrading marine bacteria. *Curr. Opin. Biotechnol.* **18**: 257–266.

Yao, C., Cao, Y., Wu, S., Li, S., and He, B. (2013) An organic solvent and thermally stable lipase from *Burkholderia ambifaria* YCJ01: purification, characteristics and application for chiral resolution of mandelic acid. *J. Mol. Catal. B Enzym.* **85–86**: 105–110.

Yoo, H.Y., Simkhada, J.R., Cho, S.S., Park, D.H., Kim, S.W., Seong, C.N., and Yoo, J.C. (2011) A novel alkaline lipase from *Ralstonia* with potential application in biodiesel production. *Bioresour. Technol.* **102**: 6104–6111.

Yoshida, S., Hiraga, K., Takehana, T., Taniguchi, I., Yamaji, H., Maeda, Y., *et al*. (2016) A bacterium that degrades and assimilates poly(ethylene terephthalate). *Science* **351**: 1196–1199.

Zaks, A. and Klibanov, A.M. (1988) Enzymatic catalysis in nonaqueous solvents. *J. Biol. Chem.* **263**: 3194–201.

Zanger, U.M. and Schwab, M. (2013) Cytochrome P450 enzymes in drug metabolism: regulation of gene expression, enzyme activities, and impact of genetic variation. *Pharmacol. Ther.* **138**: 103–141.

5 Appendix

5.1 Polyester hydrolysis of CEHs used in this thesis

CE01	CE02	CE03	CE09	CE07	CE10	CE11	CE12	CE13	CE14	CE15	CE16	pET-22b(+)
CE17	CE18	CE19	CE20	CE01	CE02	CE03	CE09	CE07	CE10	CE11	CE12	
CE13	CE14	CE15	CE16	CE17	CE18	CE19	CE20	CE01	CE02	CE03	CE09	
CE07	CE10	CE11	CE12	CE13	CE14	CE15	CE16	CE17	CE18	CE19	CE20	
control	control	control										
CE21	CE14	CE22	CE23	CE24	CE26	CE25	CE04	CE06	CE05	CE08	CE21	pCR-XL-TOPO
CE14	CE22	CE23	CE24	CE26	CE25	CE04	CE06	CE05	CE08	CE21	CE14	
CE22	CE23	CE24	CE26	CE25	CE04	CE06	CE05	CE08				

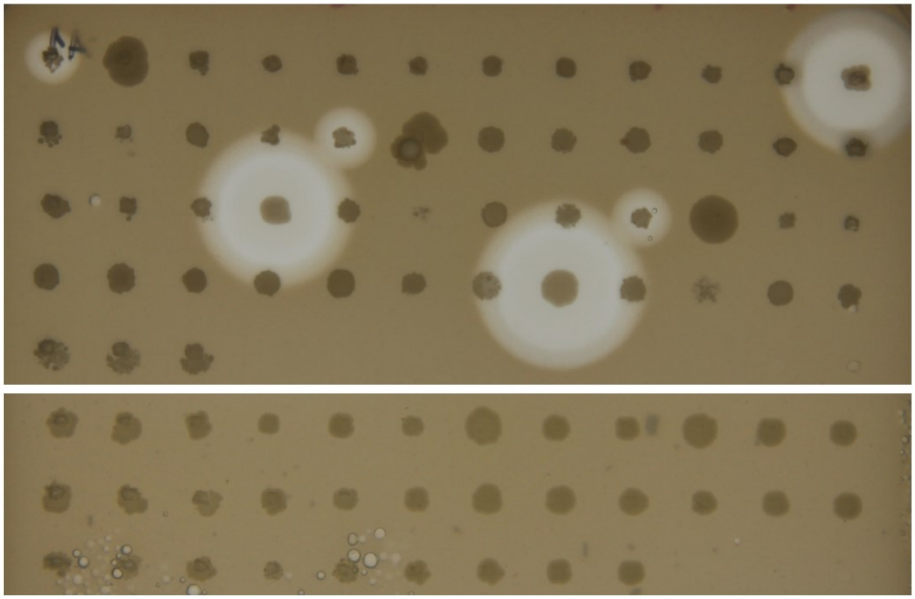


Figure 5-1 Activity of CEHs used in this thesis with Impranil DLN as a substrate. The distribution of the CEH producing *E. coli* cells is shown on the top. The first four rows apply the pET-22b(+) expression vector with *E.* coli strain BL21(DE), the last three rows apply the pCR-XL-TOPO vector with *E.* coli strain TOP10. Either single genes (pET-22b(+) vector) or genome fragments of 5 kbp at average (pCR-XL-TOPO vector) were cloned. *E.* coli BL21(DE) cells transformed with the unmodified pET-22b(+) vector were used as a control. Agar plates with 0.4 % (v/v) Impranil DLN suspension as substrate are shown below the table. Pictures were taken after 48 h incubation at 30 °C; clearing halos around bacterial colonies indicate polyester hydrolase activity. LB medium supplemented with either 100 µg/ml ampicillin or 50 µg/ml kanamycin was used to prepare the agar plates. Plates prepared with ampicillin were supplemented with 0.4 mM IPTG as inducer. CE01 was the HZ lipase from *Aneurinibacillus thermoaerophilus* strain HZ, CE02 to CE12 originate from *A. borkumensis*, and CE13 to CE26 originate from *P. aestusnigri*; respective protein identifier can be found in chapter 2.5 and **Table 3-1**.

5.2 Gene sequences of CE01 to CE25 in FASTA file format

```
>CE01
ATGCAAAAGGAAAGAAAAAATCAATATCCAATCGTTCTAGTTCATGGGTTTGCAGGTTGGGGAAGAGA
TGAAATGCTGGGTGTAAAGTACTGGGGTGGAATGCATGACATTCAAGAGGATTTGAAACAGTATGGTT
ACGAAACACACACTGCGGTAGTAGGACCGTTTTCAAGTAACTGGGATCGTGCATGCGAATTATATGCT
CAACTTGTTGGTGGAACAGTGGATTATGGTGCTGCACACGCTGAAAAATATGGACATGACCGGTTTGG
TCGAACCTATCCTGGGCTTTTGAAGAATTGGGATGGAGAACACAAAATCCATTTAATCGGACATAGCA
TGGGTGGACAGACGGTTCGTGTGTTAACGCAATTGTTAAAAGAGGGAAGCCAGGAAGAACGAGAGTAT
GCGAAAAAGCATGGGGTGCAATTGTCTCCGCTATTTGAAGGTGGGAAGTCGTGGGTTCACAGTGTAAC
AACGATTGCAACGCCGAACGATGGTACAACGTTGGCTGATGTAGTGACACAACTCATTCCAGCAGCAC
AACAAATTATGGGACTGGCTGCTGCGGTGTCAGGCAATACAAATGTACCGGTTTATGACTTCAAACTC
GATCAGTGGGGATTGAAAAGAAAAGCAGGTGAATCCTTTGTGCACTATGCGGATCGCGTATGGAATAG
TGGAATTTGGACAAATACAAAAGATATTAGTGCATGGGACCTCAAACCAGAAGGAGCCAAAGAGCTGA
ATAACTGGGTAAAAGCGCAGCCGGATGTCTATTATTTCTCATACAGTGGAGAAGCTACATTCAGAAGT
CTGATCACGGGACATCATCTCCCAGATCTAACAATGAATAAATTGATTACTCCATTTGGTATTTTCTT
AGGCTGCTACGGCAGTGATGAAAAGTGGTGGCAGAATGATGGGATTGTAAACACAATTTCGATGAATG
GGCCGAAGCTTGGCTCAACAGATGAGATTGTCCCATATGATGGAACACCTAAAATCGGAAAATGGAAT
GACATGGGGATTCAGGAAAACTGGGATCATGCTGATTATATTGGGCTCAGTCTTTCATATGTTTTGGG
AATAGAAAAAATTGAAGATTTCTACCGTGGCGTTGCAGATATGCTTGGTTCATTATCTGTGAGATAA

>CE02
ATGAATCCTGCCGTTATTGAGCGGGCCACGGTACGTGCCCTGATGTCTTTGCCGGGCCCGGTTTTGGA
GCGTTTGGCTGCGGGTTTGGAAACCCATAGTCGTCCGCATTTGGATTCGCGTCTGCGCTTTCTGCTTG
CGCTCAGTGGGGCCAAGCCCACCCTGGATTCAGGGACGGTGGAGCAGGCTCGGCAGATTTATCGGAGC
ACGCTGGCATTGCTAGATATGGCACCGGTTAGCCTTCCTGTGGTTGTTGATCATCAGGTTAGCATGGA
GGACGGTAGCCAAATTCTAGTGCGGCGTTATCGCCCGGCCGATGCTCCGCTGGTGTCACCGGCTATTA
TGTTTTTTCACGGTGGAGGATTCACTATTGGCGGTGTTGAAGAGTATGACCGGTTGTGTCGCTATATA
GCGAAGCGTACCAATGCGGTGGTGCTGAGTGTGGATTACCGGTTGGCGCCGGAACACCCAGCCCCTGC
CGGCATGGATGATGCGTTGGAGGCTTGGCGCTGGTTGCTGAATAACACCGCGCAGCTGGGGCTTGATC
CGAACCGCTTGGCGGTGATGGGCGATAGCGCTGGCGGCTGCATGAGCGCAGTGGTGTCACAACAGGCC
AAGCTGGCCGGCCTGGCGTTGCCAGCTTTACAGGTGCTGATCTACCCCACCACTGACGCCGCCCTGGC
CCACCCTTCCGTGCAAACGTTGGGGCAAGGGTTCGGCTTGGATATACCGCTGCTGACTTGGTTCCGAG
GACACTTTGTTCAAGACCCAGCGGTGATCGAGGACTACCGCGTTTCGCCCCTGCGCAACCCTGATTTA
ACTGGCCTGCCTGAGGCGATCGTTATCACGGCTACGGATCCCTTGCGGGACGAAGGCTTGGAGTACGC
CCAAAAATTACGTGAGGCAGGAAACACGGTGACCTCACTGGATTATCCCGAACTGATACACGGCTTTA
TTTCCATGGGGGGTGTTGTTCCGGCGGCGCGTAAGGCAATCAATGACATTTGCGTTGAGACCAAGCGG
CGGTTGTGA

>CE03
ATGGCTTCTATTCCCGCACACCTCATGAAGTTGCTTCTTCGTGCCGGGGTCAAGCGCGATATACGCGA
CCCGGACAAACTGGTAAAGCACTTACGCAGAGCAATGAACGCTCCTTTGGCGCCCTCACCGCTTCCAC
GCGGAATACGGTTACAGCGAGGCAAGGTTGCCGGGACGGCAGGCCACTGGCTCAGCCCCACCGACCCA
CAAACAACCATACTCTACCTGCATGGTGGCGCGTTTATTGGTGGCCGCCTTGCAACCTATCACAACTT
TTGTGGCCATCTGGCCCGCACGCTTAACGCGCGGGTTTTTCTTCCAGATTACAGACTCGCGCCAGAGC
ATCCGTTCCCGGCAGCGACGGATGATGCCTTTAATGTTTATCGTGAACTAATGGCTGATCCCCGCCCT
ATAGTCATTGCTGGAGACTCCGCAGGGGGCAACCTGACTCTGGTCACGCTACTACGCGCCCGAGATCA
CAAGCTGAGAATGCCGGCCTGTGCCGTAGCCATCTCCCCAGCATCAGATGCCCGAGGCAATTTGATGT
CCCGACAGGCGAATAGCGACAGTGATGCCATGCTATCTCACTGCATGATTGAGGTCGCAACGGATGTC
TACTTGGCAGGCGCTGATCCCGCCCACCCATATGCCTCACCAATCACCCAAGACTTTACCGGCTTGCC
ACCTCTACTTTTCACTGTAAGCAGCGAAGAGTGCCTGCGCGACGATGCCTATGCAGCCGCTCACTGCG
CCCGACAAGCCGGCGTGCCTGTGCAGCTTCTGGAACGTAAAGATATGCCACACGTCTGGCCCGTATTC
ACTTTCCTGCTGCCGGAGGCAAAACAAGACCTCCCCACAATCGTCCGTTTTCTTCGCAAATATCTTGC
CACCACAGACGCCCACGAGGAGGCATTTCACACTGCAAACGAGAGCGACACGACAATCCCGGAGATAT
CGTCATGA
```

>CE04
ATGACAGCAATAATTCGTCAGGGGCGTTATCAAGGACTTAGCAGCAAGGGCGTAACGGAGTATCGAGG
TATTCCGTTTGCCAAGGCGCCACTGGGTGAGTGGCGCTTTAAAGCCCCTCAACCGCTACCGGACAGCG
AAGACTGTGTTAACGCTGACCGTTACCCCCTGGCGTCATTGCAGCCACGCAATCCTATAATGGGAATT
CAGGAATCCGGTGAAGATTGCCTTTATCTGAATATCTGGGCCCCTGAAGGTGAGGGCCCGTTTCCGGT
GATGGTGTGGTTCCATGGCGGAGGCTATATGGCAGGCTCCACGTCCCAGGCTCTATATAATGGCGCAG
AGTTAGCACGCAGCCAAAAAGTGGTAGTGGTTAATGCGGCATACCGTTTGGGCGCCATGGGCTTTGCG
GATTTTTCGGCGGTGGCGCCAGAGTTGGATGCCGATACCAATCTGGGCCTGCGCGATCAGTTGGCGGC
GCTACAATGGGTACAAGAGAATATTGCCGCTTTTGCTGGCGACGATAAGCAGGTGACGATCTTTGGTG
AGTCTGCTGGTGGTTTCAGCGTCTGCTCATTACTGGCCTGCCCGCAGGCTGATGAGTTATTCCAAGCG
GCCATTGTGCAGTCAGGGGGAGCGGATTTTGTGCTAGCCCCGGATCAGGTACGCAAGGTCACTAATGC
GTTTGTGGCAGCCTTGCCGGGCGACGGCAGCGCCGCCGAAAAGCTGCTCAGTGCGGATAACAAGGGCT
GGATTAAGGCCCAAAATGCGGCGGTGAAAGTGTTGGTGGATCGAGGCTTGCGTACCACCACACCTCAA
TTTGCCATGAACTTCTTGCCCATGGTGGATGGTGACGTGCTGCCGCAATTGCCGGTTGATGCCATCGC
TGCAGGCGCTGCTGCCAACAAGCGTGTAATGGCTGGCGTATGTCGTGATGAATTCAACTTCTTCCAGT
ACGCTGGCGTGCTGGCTGGCACAACCACCATGGACGCCCTGAGGGAAATCAGCGATGAGGAAATTGTG
AGCCGATTTGAGCGTGCGCTGCCCGGTAATGGCCGCCGTGCTTTTGATTATTACCAGACAGCGGTGGA
GCCGGATGCACGGCGCAGCCGCCTGGACTGGTTAGCGGCTATGGAATCCGATCGTCTGTTTCGGGTGC
CCACAGTGCGCCTACTTGATGCCCAGAGCCAGCATGCTCAGTGCTGGGGGTTTCAGTTTACCTGGCCC
AGCGAGCCCTTCGGTGTGCCACTGGGGGCCTGCCATGTGGTGGACGTGCCCTTTGTCTTCGGTGTGAC
AGATACCCCAGCGGGAATGTACTTCACCGGAGGTACGAGTGAAGCGCGGGCGCTGTCTCATCAGGTGC
AGGCAGCGTGGGGCACTTTTGCGCGGGGCGATGCGCCAGGCTGGAATGCTTGGCAATCGGACCGTCAG
GTGTGCCAGCTGGGCCCGGGTGAGACCATGGCTTCCTTGTTGGATGAAAGTGGCGAACAGCTCTGGCG
AGACATCATCCCGGTGGTTTGA

>CE05
GTGCAAATGGCGACTTATAAAGCCCCCCTCGAGGACATGCGTTTTGTTCTCAACGATGTGTTCAAAGC
GGATCAACTCTGGGCCTCCATGCCGGCCACCGCAGAGGTGACTCAGGATTTGTCCGATGCGATTCTGG
AAGAAGCGGGCAAAATGACCGAAGGCTTGCTGTTCCCGCTGAATCGTAACAGCGATGAGCAGGGCTGC
ACTTGGACCGATGGGGCGGTGACTACCCCGGATGGTTTCAAGGAAGCATTCAAGACCTTTGCCGAAAA
TGGTTGGAGCGCATTTTCCGGTAACCCGGAGTTTGGTGGCCAGGGTATGCCCAAGTCATTGGCGGTGC
TTTTTGAAGAAATGATGCACAGCGCTTGTTCTTCGTTTGCCTTGTATCCGGCACTAACCAGTGGTGCG
TGCCTAGCGGTGGATGCCCACGCTTGCGAAGAACTCAAGTCGCAGTACTTACCCAAGCTTTATAGCGG
TGAGTGGAGTGGCACCATGTGCCTGACTGAGCCGCATTCCGGGACAGACCTGGGTATTCTACGGACCA
AGGCTGTGCCTAACGATGATGGCTCGTTCAACATTACCGGCACCAAGATTTTTATCACCGGTGGTGAA
CATGACCTGACCGGCAACCATGTCCATCTGGTGCTGGCTAAGTTGCCGGATGCGCCAGCAGGCTCCAA
AGGGATCTCCCTGTTCCTGGTGCCCAAGTTTTTGCCTGACGCGGACAACAATCCGGGTGAGGCCAACG
GTGTCACCTGTGGTTCCATCGAACATAAGATGGGGATCAAGGGCTCTGCCACCTGTGTTATGAACTTC
GACGATGCCAAAGGCTGGATTATCGGTGAGCCGAACCAGGGGCTGGCGTGCATGTTCACTATGATGAA
CTACGAGCGACTGTCCATCGGGCTACAGGGCTTGGGGTTGGGTGAAGTTAGCTACCAGAGTGCTGTGG
AATATGCGCGAGAGCGCCTACAAGGTCGCAGCGCCACCGGTGCGAAAAACCCGCAAGGACCGGCGGAT
CCGATCATTGTGCACGGTGATGTGCGTCGCATGCTGATGAACATGCGCGCCATCAATGAAGGAGGCCG
TGCGCTGGCAGCCTATGTGGGCATGCAGCTGGATACGTCTAAGTTCAGCGAAGACGCGGAGGCCAAGA
AGAAAGCGGAGGATCGGGTGGCGCTGTTGACGCCGGTCGCTAAAGCTTTCTTTACCGATCGAGGCCTG
GACACCACCATTACCGGGCAGCAGGTATTTGGGGGCCACGGTTATATTCGTGAATGGGGTATGGAGCA
GTTTGTTCGCGATTGCCGTATTTCCCAGATCTACGAAGGCACTAACGGTATTCAAGCGTTGGACTTAG
CTGGCCGTAAAGTGGTCCGCAACGGTGGTAAATCCGTGGATGCTTTCCTGGCCGATGCCCAGGCCTGG
GTGGATGCTAATGCCGATAATGCGCAGCTAGCCGCAGTAAAAGACGACCTGCAAGCGGCTTTGCTATT
GCTGAAATCTTCCACCGATACACTGCTGGCCCAGGCCGGCAACAACCCGGATGCGATCAGCGCTGCTG
CGGTGGAGTACCTGGATATCTTTGGTTATGTGCTCTACGCATGGCTGTGGGCGCAGATGCTAGCGGCC
ACCGATGACCGTGATGATGACTTCGCTAAAGCCAAACGTATCACCGGGCAATACTACTTTCAGCGTGT
ACTGCCAAAGGCCCAATCACTGGCTGCCCAGCTGAATAGCGGTGCGGATGCCATGATGAGCCTGGACG
CGGAACAGTTCTGA

>CE06

ATGCAGTCAGGTACGGTCGCGAGTAATGGCATCGAGCTTTTTTATGAGAGCCGTGGTCCGGAAAACGG
GGAGCCCATGGTCTTCGTCATGGGGCTGTCGGCGCAGATGGTGTTCTGGCCGGATACCTTGCTGGATG
CGTTGGCGGCAAAAGGCTATCGGGTGATCCGGTTTGATAACCGTGATGTGGGTAAATCCACCCAAATC
CGCAAGCCGATAAAGCAGGGGCCGGTGTCGGCGATCCTGCGCCGCATAATCGGTCTGCCGGTGGAGAG
CCCTTACACCTTGCACGATATGGTTGCCGATACGGTGGGCCTGCTAGATGCATTAAATATTGAGCGGG
CGCATTTTGTTGGTGCCTCCATGGGCGGGATGATCAGCCAGTTGATGGCGGGCACTCACCCGGAGCGG
GTGCTGTCCCTCACCTCAATCATGTCCAGTAATAACAGCTCATTACTGCCACCGCCGAAACCCTCGGC
GTTAAGAGTTTTAATTGCCCCGCGCGTGAAAGTGGAAACCGAAGAGCAGTTCGTTACCTTCGGCTTGG
AAATGATGAGTAAGCTGGCCGGTACCTTGCCCCAGGGTAAGGAGGAGCTGGCGGCCATGTATCGGGCT
GCCTGGGCGCGAGGCATTAACCCGAGAGGTATCCGCAACCAGTTTCTGGCGATCACGGCCACCGGTTC
TTTAAGCAAAACCTTGAAGCAGATCCAATGCCCGACCACGGTGATCCATGGCGGCGCTGACCCGTTAA
TTCGTCCTGCCGGTGGCAAGGCATCTGCCCGCGCTATTCGCGGTGCCAAGCTGGTTATTATTCCCGGC
ATGGGCCATGATTTTCCGCCATCGGTGATTGATCGCATCGGTGAGCTGATTGCGGAAACTGCCGGGCG
AGCCAATTCCGTAGTGCCGCCGGCGGTAGGGTGA

>CE07
ATGAGCCTTCAAGCCCGCCTCATCAAAGCCGTGACCAAGCGCACCATCAAACGCTCCGGTCTCAACCA
AGACCAACTAGTACGGCATCTGCGTAAGGTATTCAACGAAACCCCTGTTCTGACCCTATTACCCCGTG
GAGTAAAACTCAGTCGCGTGGAGCACCCCGCTTTTACCGGTGATCGAATCAGCGTGCAGCGTCCAGAA
ATGGCGGTGCTTTACCTGCATGGCGGCGCCTACATTGGCGGCATCACCAAGACCTACCACAACTTAGC
CGGGCGGTTGGCAAAAAAATTAAACGCTGAGGTATTTTTACCCGTCTATCCGTTTGCCCCAGAGCATC
CTTACCCGGCCGCAGTGAATCGAGTCATGGAGGCCTACGAATACCTATTAAGCCTAGGCAAGCAACCG
CAGGATATCGTTATTGCCGGCGATTCCGCCGGTGGTGGGCTCACCTTGGCCACGCTGCTGCATATCCG
CGACAAAGGCCTGGAACAACCCCGCTGCGCGGTGACATTTTCGCCCGCCAGCAACGCCTTTCCTGACG
ACAGTATCCTGGAGGCCCTAGACCCCAGCGATGCGATGCTGTCTGCAGACATCATTCGCACTGCTATC
GAGATCTATACCCCCAACCCAGAAGACCGCAGCCAGCCCTACGCCTCCCCCTGCCTAGGAGACTACAC
CAACATCTGCCCGCTACTGATCACCGCCAGCACCGACGAACTGCTTTATGCGGATGGCAAGCGAGTGA
AGCGCGTGGCGGAAAAGGCGGGCGTCAAGGTGACATGGATAGAGCGCCCCGGCGTATTCCATGTATGG
CCGGTCATGGTGCCCTTCCTGCCAGAGGCGAACAAAGACCTGAAACGGATTGTCGCATTCATTAAAGA
AGCATAG

>CE08
ATGAGCCGAGCAACATCAGGCTCCGCACTTACACCGGTATCGTCTTTTTCTCTGGTGGCAGATAGCGA
TCAGGCTTTTACGCTGCGTGTTACTTTGGCGCGTTTGGCCACTCGCACACTGGATATGCAGTATTACA
TCTGGGATGACGACACGACCGGAAAGCTGCTCATCTACCGAGTGTTGGAGGCAGCGCGGCGGGGCGTG
AGGGTGCGCATGTTGTTGGATCATGCCAATCAGCTGGGCCGGGATGTGAAGTGGGCGGTGCTGGATGC
GCACCCAAATATTGAGGTGCGCCTGTTCAATCCGTTCAAGGGGCGCTACAAGCATTTTCTGCAATGGC
TCTACTATGCTCCGCGATTAAATCACCGCATGCATAACAAAGCTTGGATTATGGACGGCGAACGAGCG
TTGGTGGGTGGGCGTAATGTTGCGGATCACTATTTCGGCGTAAATCCATCCACTAATTTCCGTGATCT
GGATCTGTATGCCCATGGCGCCATCGTGGCCGACACGCAAATGGCTTTTGATGCCTTTTGGGATAGCC
CGTTAGCGGTGTCTATGAAAAGTTACCGTCATCGCACGGAGGCGACAGCGGATCGGGCCTGGCGTTGG
TTGGGGCGCTGGCGCACTTCATTGCAAGGCTACCCTTACCTCTTTCCCCAGCATGAAGCATTTTTCAC
CGATTACTTGGCAAAACAAGAGCAGCAATGTGTGCAGGCTCCGGCAGCGTTGCTGTTTGATAGCCCTG
ATAAGGCCAGTGGTGCGAAGCAGACGTTGATGGGAGAGCAACTGGCGCGTTTGCTGGGCAGGCAAGAC
AATCGGGAGTTATTGATGGAGGCCAGTTATTTTATTCCCGGTGATGCTTTTGTTGAGGCGCTGGGAGG
GTTCCAGCAGCGCGGCGGACGCGCGGCGGTGCTGACTAACAGTTTGGCCACTAATGATGTGATTGCCG
CTCATGGTGCTTATGCGCGTTACCGTGCAGCACTGTTGGCTGCAGGGGTTGAGCTGCATGAATTACAG
CCCAATGCCCGGGCACTGCACCGGCAGGTGCGGTTGTTCAAAGGGCGCTCCCAGGCGAGCCTGCATAC
CAAAGCGATGGTGTTGGATCGACGCGAAGTGTTTGTCGGTTCTTTCAACATCGATCCGCGCTCGGTGC
ATTTGAATACCGAAATGGGATACTACGTAGTGTCCGCTGAACTGGGGCAGCAGGTGGCGGCATTTATT
GAAGAAGGGTTAGCGCCGGAAAACAGTTACCGTCTTGAAAAGCAGGAAGGCGAATTATTGTGGGGCGC
GGATGGATATGACGGTAAGCCGGTCTTGTATTCCCGTGAACCCAAGGCCACCGTTTGGCGACGCATGC
TGGCATGGTTATTGTCGTTATTGCCGATTGAGCGGATGTTGTAA

>CE09

ATGAGCCTGTTTGTTGATCGCATCAGCCGTAGCCAGTCGACACTGCAACAAGTACTGGCAGGCACGCT
ACGCCAATCCCTGGCACTACTATTCAAACCCGTTGCCAATCCACGCTTCAGCTTTACTGCACAGCGGC
GCTGGCTAGACATTATGGCCCACAGTACCCTCAAAGCCCGAGGCATCGGCAGCCACGATAACGCGTTA
GCCAATGTACCTTGCCGCCACTACCAGCCGTTGCATGCACAAACTACCGTGCTCTATTTACATGGCGG
TGGTTACGTGGCTGGGAGTCCGGACAGCCACAAGTCCATTACCTCGCACCTAGCCAAATTCGCCAATG
CCCATGTGGTGGTGCCGGATTACCGGCTCGCTCCGGAACACCCCTGCCCTGCCGCCATCGAAGACGCC
GTCGCGGTCTACCTAGCTCTGCTAGAGAGCGGCGTAAGCCCCGGCTCTCTAACGTTGATGGGCGACTC
CGCTGGAGGCGGATTAGCACTGGCCACGCTGCAAGCGTTAAAAGCCAGCGGTAAGCCGCTACCCAGCA
GCGTTATTCTTTTTCCCCCTGGGTTGACCTAACCTTGGGACAGCTTTACGACACCGACCGAGATATC
ATGCTAAGCCATAGCTGGCTAGCCAGCGCGGCTGACGCCTATGCGGGTAATGATACCCGCAATGTCCA
GTGCTCCCCGCTGTTCGGCGAACTAAGTGACTTGCCCCCCACGCTAATCCAAGCCGGCGGCGATGAAA
TTCTGCTTAATGATAGCCACCGTCTGTGCGCTGCTTTAAACGAAGCGGGCACCCCAACACGGCTACAG
ATTCACCCGCAACGCTGGCACGATTTTCAATTACACGCGGGTGTACTTGCAGACTCGGATCACGCTCT
GATGACCTGCGCACGCTTTATTCACCAACATGCCATCCATGACCATGAAGAGGACGCTCACGCATGA

>CE10
ATGGATCTGATCATTTTTCTGCTGTCGCTGCTAGTGATGAGCACGGTTACCCTGACACTGGCGTTTTA
TCTGTGCTGGTACTACGACCGATGCAGCTTTCCCGAGCAATCTGAGCTGCCCGGTGAACCCCCTTTAC
GGCTGTTGCCGACGCTACTTGGCATGGCCAAGGAGGCCGCCGCCCTGACAGTGCTAGTGCTTAGTTAT
CCGTTGCGCCTGATCCACGATTCTTCACCGGTACGTAGCCGCCATCACGGCGAGCCCCCGGTGATTCT
GGTCCACGGTTACGGCGGCAACAGTGCCAACTTTTTATTTATGCAGTGGCGCCTGAAATGGCGCGGCT
GGTCCAATGTGTATTCGGTGAGTTATACCCCGCCGCATATCAATGCCCGCAAACTATCCCAGCAGGTG
GTTGATCATGTAGAGCGTATCCTGGCGACCACAGGCGCGGAAAAAGCCCACTTGGTTTGCCACTCCAT
GGGCGGCCCGTTAACCCGCTATGCGCTCAAGAATTTAGGGCTGGCTGGCAAGGTAGACCGGGTGATTA
CCTTGGGTAGCCCGCATTATGGAACGCGAGTTGCCGGCCTGTTTCCGCCATTGGGGGCTGCCGCCCAA
CTGCGCTACCAGAGCCCTTTTATCCGAGAGCTGGCCACCGAGGCAACCTGCCCAGGGGGCGCGCGGTA
CTTCTCCATCTTCTCCAACATGGATAATTTTGTGCTGCCGGTGTCCGCGGCGGTATTGCACGGTGCGG
AAGACAATATCCATGTTCCTTATCTGGGGCACTGCTCGCTGCTCTACAGTAGCCGGGTTATGGATCAA
GTGGAGCGTTGTCTTTTGGCCCCGCCCAAGGCGGCGGAGTAA

>CE11
ATGCCGGTCCCCGAAACGTCTCTTATTTATATCCACGGTTTCAATAGCTCACCAGCATCCCATAAAGC
CGGGTTGTTGCGATCAGTGTTTGAGCAAGCGGGCTGCCCGGAACGGTTGGTCGTGCCCGCATTGCCGC
CTTCGCCAGTGCTAGCCATGGCGGAATTGGATGCAGCCATTGCTGGGGCTGGCCCGGTGGCCCTAGTG
GGATCGTCGTTGGGCGGCTTCTACGCCACCTGGCTGGCGGAGCACCATGGTTTGAAGGCTGCAATCGT
AAATCCTGCGGTAAGGCCGTGGCGACTACTGGACAAATACACAGGTATCCAGCATAACTACCATACTG
GCGAAGCCTATCAGTTCGATCCGGCTTGGCTGGAGCAGTTGCGCCGCTATGAGGTAGCAGAGCTCACC
CAGCCGGAAAATCTATTGTTGCTGACCCAGACGGGTGACGAAACCCTGGACTGGGAAGATGGCTGGGA
CTATTACGGCGACTGTCACCGTTATTGTGGCTTGGGTGGCAGTCACGGCTTTGATAACTTTGACGCCT
TCATTCCGCTGATGCTGCGGTTTTGTGGGATTGAAATACACGCCTGA

>CE12
ATGGAACCACTTGAACTTGAGGACCAATTTGTTACTACCGAAGCCGGCTACCAGCTGCACTACCAAAG
CGCGGGTACTGGCGAGCCGCTGATTTTCCTCCACGGTGGTGGCCCCGGAGCCACCAGCTTCGGCAACT
TCTACTACAACGCACCGGCGTTCCTGGAACAGTACCAGTGTTTCTTCTACAACATGCCCGGCTACGGC
CAATCCAGCAAGTTAGTGGTTGAGGCGCCGATGTACAGCTTCCACGCCAGCATGCTCGCTGAATTCAT
GGACCTGGTCGGCATCCCGCGAGCGCACCTGGTCTGTCAGTCCTTCGGCGGCTGTGCGGCCATCAAAC
TGGCGATCGACCACCCTGAGCGCGTCAATAAACTGGTATTGATGGGCGCCCAGCCGATGTTTGGCGGG
GTTATTGATCCATTGAAACTGATGAGCAAACACGCGGCCAACATTATCCTCGACTACTACGGGGGCGA
AGGTCCGACCCCAGAAAAGATGCGACGTTTACTGGCCGATTACGAGATGCACGATGACAGCAAGCTCA
CTGACTGGACGGTGAACGATCGCTTTGCCAACAGCACAGACCCGGAACTGCTTGAAGTCGCCCGCACT
CCCGGTTTCATGGGGCAACCCGAATCGCTGCTCGGTGACTTGCACCGCAATGTCGTCCCCACCCTTTG
CCTGTGGGGTCTGCACGACTGGTTTGGCGGCCCGGACGTGCCCCTGCTGTTTCTCAACCAGTTTGCAA
ACGCTCAATTGTTTATCGAAGGACAGGGCGCTCACCACTGGCAGACCGAACTACCCGAGCGCTTCAAT
CGCGTTGTGCTCAGCTACCTGAGTGAATAG

>CE13
ATGCCTCAATCTTTTAAACAGTACAGGACAACAATAATGAGTACTGCTCGCCTGACGCCAACCACGGC
ATCAACACTCAAGCTGCTTTCCGTCGCCGTCGCGGCCCTCACGCTTTCCGCCTGCCTGTCAGGTGGCG
GCGGTGGCCGTGCCGACGGCCCCAACCCCCTGGCGATTGAAACCAGCGAGGGTAAAGTCGTCGGTATC
AGCAACGACGGCATCCGCGTCTTCCGCGGCATTCCCTACGCCGCACCGCCGGTTGGCGACCTGCGCCT
CGCCCCGCCGCAGCCACCCGCCAGCAGAAGTGAAACACTCAGACTGAGCGAAGAGTTCGGCAACAGCT
GCCCGCAGTCCGATCTGACCACTGGCCAGCAGGTAGGCAACGAGGACTGCCTGTACCTGAACGTTTAT
GCGCCCGCTGAAGCGGAGGATCTGCCGGTAATGGTCTGGATCCATGGTGGCGCCTTCGTGTTTGGTAA
CGGCGGCGGTGAGTACGACCCGACGCGCCTGGTCGAACAGGACGTTATCGTCGTGACACTCAATTACC
GCCTTGGCAATCTCGGTTTTCTGGCCCACCCTGCCCTGGAGTCCGATGCCGGCAACTTTGCCCTGATG
GACCAGCAGCTCGCGCTGGCCTGGGTCAAGGAAAACATCGCTGCCTTTGGCGGTGATCCGGCCAATGT
CACAATCTTCGGCGAGTCCGCCGGTGGCCATAGCGTGATGAGCCACATTGTCTCGCCCCGGGCTGAAG
AGGCAGATCTGTTCCAGCGCGCCATTGTCCAGAGCGGCTCCTACGCGCCCTTCCAGATGCCCAAGGCA
ACAGCGCAGTTCCTCGGCACATCGGTAGCCAATGGTCTCGGCTGCACCGACCCGGAGACCGCTGCAAG
CTGCCTGCGCTCCCTGCCGGTATCCGCGTTCCTGGCGGCCCAGGGCTCGCAGTCAATCCCGGTTGTGG
ATCCGGACGACGACCTGCTACCCAAGTCCATCCAGCAGGCACTGGCCGACGGCGACTTCAACAGCTCG
CTCGATATCATGATCGGCAGCAACCAGAACGAGGGCACCCTGTTCGTCGCGCTGGATGAAGTAGGAGG
CGATCCTATTGATGACGAGGCGGAATACCGTGAGCGCGTTGCCGAGTTCTTCCAGCCCTACCAGGCGA
GCATCCCCTTCGACGACGACCAGATCGCCACTGATTATCTGGACTTTGTTGACGGTGCGGCAAAACCC
TTCGAGGCAGCCCTGTCCGGAATCTGGACCGACTTCATGTTCGCCTGCAACGCCTACTCGCAGGCCAG
CACCTTTGCCGGCGCCAGCATGAATACCTTCCAGTACTGGTTCCGCGACGAGGACGCCCCCTGGACCC
TGGTGCCGCCATTTGCAGTGAGCTTCCCGCTGGGCGCCACCCACGCTGGGGAAATCCCCTACGTACTC
TACCCACAGGCCATCATGGAGCAGCGCTATACCGGCGATCCGGATGACCTGAACAGTCTGGCCGGCGA
GATGGTCGACTACTGGACTCAATTCGCCAAAACCGGCGATCCCAACACCACCGACGGGGTCGCAGCAG
CCTGGCAGCAGGCGGCAACCGGGAACCTGTTGACGCTGGACGTACCCAACGCAAGCAACGCAAATACG
CTTGGCTTCCTGGGCTATCACCACTGCAGCTACTGGGCTGACCCGCCGCTGGTATTGCCCTGA

>CE14
ATGGCAAACAGCAACTGCATCGCCGGCGTTGACCTGGAACGGCTCTCACGCATCGGCACTCACCTGCA
ATCGGCCTATATTGACGCAGGCAAGCTGCCCGGCGCTCTGACTCTGGTGGCGCGCCGTGGCCAGGTTG
TCTATTGCGAAGCACAAGGGCTGAGAGACGTGGAGCGCGGCCTGCCGGTCGAGCGCGATACGCTGTTC
CGCATCTACTCGATGACCAAGCCGGTCACCTCTGTTGCCCTGATGCAGCTGTACGAACAAGGTCGCTT
TCTGCTGGATGACCCGGTCCACAAGTACATCCCCAGCTGGAGCAACCTGCGCGTCTACAAGACCGGCA
GCCATCCGCAGATTCTGACCACCCCCTGCGAGCGCCCCATGACCATCCGTGACCTGCTGACCCACCAG
TCCGGTCTGACCTATGGCTTCATGAATCGCACCAACGTCGACGCCGCCTACCGGGCGCTCAAACTCGA
CGGCGGCCCCGGCCTGACCCTCGACCGTCTGATCGACGAACTGTCACGGCTGCCGCTTGAGTTCTCGC
CCGGCACGGCCTGGAACTACTCGGTGGCGACAGACGTCTGCGGTTATCTGGTGCAAGTACTGTCGGGC
ATGAGCCTGGACGACTACTTCAAGCGCCATATCTTCGACCCGCTGGAGATGAGCGACACCTTCTTCAC
CGTCCCGGCCGATCGCATCCCGCGCTTCGCCGCCTGCTATCAGTTCCAGCCCGGCGGCAGCTTCAGTC
TGCAGGATGACCCTCAGGATTCCAGCTTCACCAAGGCCCACGGCTTTCTCTCCGGCGGCGGCGGTCTG
GTCTCGACCATCGACGACTACTACCGCTTTGCCCAGGCACTGGCCAACGGTGGTGAACTGAACGGCGC
CCGCATCATCGGCCGCAAGACGCTTGAATTCATGCGCATGAACCACCTGCCCGGCAACCAGGATCTGC
CCGGTGTATCCGTCGGTTCCTTCAGTGAAACGCCCTACGAGGGCAGCGGATTCGGTCTCGGCTTTTCA
GTCAAGATCGACGTCGCCAAATCGCAGACCAACGGCTCCGTCGGCGAATACGGCTGGGGCGGCATGGC
CAGTACCAACTTCTTCGTTGATCCGGTTGAGGATCTGTTGATGGTGTTTATGACCCAGCTGATTCCGT
CGTCGAGTTATCCGGTAAGGCAGGAGTTGCGGGCGATCATCAATGGCGCGTTGGTGGATTGA

>CE15
ATGTCCAGGTACGTTGATGAGCTGTATCGCAATCCGGGGCAGCCCGGGTTGCGCGCCCTTCTGCGCGG
TATGCTGAAACTGCTGTTTCGCGGTTTGATCCGCCCGCCCGTACCCTTTGCCGTGCAGGCGCTGGTGC
TGCGTCTGCTGACCCTCGGCATGCCGCTGGCCAAAGGCGTGACCCGCAGCGCCGAGCAGATCGCCGGA
CGGTCCTGCATGTGGCACCGCCCGGCCGCTGGGGGCAACGGTCGCGTGCTGCTGTACCTGCATGGCGG
CGCCTTCGTCATCGGCTCCCCGCAGACCCATCGGGGCATCTGCTCGGCGCTGGCCAGCCGTGGCCAGT
TTGATGTCTGCGCACTCGATTACCGACTGGCGCCGGCGCACCCGGCACCAGCGGCCTGTGACGATGCG
GTCGCCGCCTACCAGGCGCTGCTGCAGCGAGGCTATACGCCCGCGCAGATCACCCTGGTCGGCGATTC
GGCGGGCGGCAACCTGGTACTGGTGACGGCGCAAAAGCTGGCCGCGCTCAAGCTGCCGCTGCCCGCCT

CACTGGTCTGCTTTTCACCGGTCACCGACATGACCGCCGAGCAGCTGCACGCGCCTGCGGCAGGCGAT
CCACTGCTGCATCCGTCCTGGCTCGACAGCGCCCGCGACGCCTACTGCCCGGCCGGGCTGGACCGCGC
CGATCCGTTGGTGTCGCCGCTGTTCGGCCAGCTCAAGGGCTTGCCGCCATTGCTGCTGCAGGTCGGTG
AAGACGAGGTGCTGCTGAACGACAGCCTGCGGCTGGCCGAAGCCGCGCGCGCAGCCGACGTTGCCGTA
CGGCTGGAGCGCTACACGGACCTGTGGCATGTATTTCAGGCCCACACCGGGCTGCTGCACAGTGCTGA
TGCGGCGCTGCAGCGGGTGGTGGAGTTTGTGAACAGCGCTCAGGCCGACTCGGTAAGCTGA

>CE16
ATGCCATTTAACAAGAAAAGCGTTCTTGCCCTCTGTGGCGCTGGCGCGCTGCTGTTCTCGATGTCTGC
CCTGGCCAACAACCCGGCACCGACCGATCCGGGTGATTCGGGCGGTGGTTCCGCCTATCAGCGCGGTC
CGGATCCGTCCGTCAGCTTCCTCGAAGCCGATCGCGGTCAGTACAGCGTCCGTTCCAGCCGTGTCTCC
AGCCTGGTCAGCGGCTTTGGCGGCGGCACCATTTACTATCCGACAGGCACAACCGGCACTATGGGTGC
GGTCGTTGTAATTCCCGGCTTCGTTTCTGCCGAGTCTTCGATTGACTGGTGGGGCCCGAAACTGGCGT
CCTATGGTTTCGTGGTCATGACCATTGACACCAACACCGGCTTTGATCAGCCGCCGAGCCGTGCGCGT
CAGATCAACAACGCACTGGATTATCTGGTTAGCCAGAACAGCCGCAGCTCCAGCCCGGTTCGCGGCAT
GATCGACACCAATCGTCTGGGCGTCATTGGCTGGTCGATGGGCGGTGGCGGCACCCTGCGTGTTGCCA
GTGAAGGTCGTATCAAGGCAGCCATTCCGCTGGCGCCCTGGGATACAACCAGCTACTACGCCAGCCGC
TCGCAGGCACCGACTCTGATCTTTGCCTGTGAGTCGGATGTGATCGCGCCGGTCCTCCAGCATGCCTC
ACCGTTCTACAATTCCCTGCCTTCCAGCATCGACAAGGCCTTTGTGGAAATCAATGGTGGCAGCCACT
ACTGCGGTAATGGCGGCAGCATCTACAACGATGTGCTGAGCCGGTTCGGGGTGTCCTGGATGAAACTG
CATCTGGATGAAGACAGCCGCTACAAGCAGTTCCTCTGCGGACCGAACCACACTTCCGACTCTCAGAT
CTCCGATTATCGCGGCAACTGCCCGTACTAA

>CE17
ATGCCCATGCACACTCTGTTCAAACGCGGTCTGGCCGCACTCGCCCTCAGCACCCTGGTCAGCCTGCC
GGCCATGGCCAGCAACCCCGGCATCAGCAGCCCGGACACCATGATCCTTGGCGATTCGATCTTTGCCC
TGTCCGGTGATATTCACGAGAACCTGGAAGCCGATCTGGACGAGAACATCGACACCTACGCCCGCTCC
GGCTGCCAGCTGACCGGCGGTAACGTTCTCTGTTCACGGCTGTATTCGGTAGAAAACCAGTATGCCCG
GGCGGACAAGAGCGGCATTCGTACGGTGATCTTCAATGGCGGCGGTAATGACATTCAGCTCAACAGCT
GCCGTCCGTCGCTGAGCGCCTGCATGCCGCTGCTCAACGAGCTGGAAGATCGCATCGCCACCCTCGTG
CAAAAGATGCGCAATGACGGTATCGAAGAGATCATCTTCCTCGGCTACTACAACGCCGCGGGCAGCGC
CGAGAACCTGCAGGACATCAACAACTACAGCATGAACTACAAGGCTGCCGCCTACCCGGGCATGGGTG
TGAAGTTTATCGATGTGCGCGCCGACTTCGCCGGCCGCGAGTCGATTTACATCACCAGTGACGGCATT
CATCCGACCGCAGCCGGCTCGCGTGTGCTGTCCAACCGCATCCTGCAGGCCTTGGACTGA

>CE18
ATGAATAACCTTACGTTACTGCCCGGCTGGGCGCTGGCCGCCAGCAACCTGCTGCCGTTGCAGCAGGC
GTTGCACGAGCGGCTGCCCGACCTGACTGTGCAGCTGGCGGAGTTGCCGCCGCGGCTGCAGATGTCGA
CCCTCGAGCCGGACCTGACGGCGCTGGCCGAGGCCTTGCCGGCTGGCTGGCTGGCCGGCTGGTCGCTG
GGCGGCATGCTGGCAGTGCAGCTGCAGCGCCGCTTCCCGGAGCGCTTCTTTGGCGTCATCACCATCGC
CAGCAATGCCTGCTTTGCGGCGCGCAATGGCTGGCCTCATGCGATGCCGGTCGATACCTTCAAGGCCT
TCCTGGCTGATGCCCGCGAGCAACCCGAGCGCATTCTCAAGCGCTTTTCCCTGCTGGTAACGCAAGGT
GGCGAGAACGCCCGGGAACTGAGCAAGCAGTTGCAGTGGAGCGATGCCGATCCGCTGCAACGTCTCAA
CCAGCTGGCGCTGCTCGGCGTTATTGATAACCGCATCCCTCTGCAGCGTTGTGCGCAACCGGTACTGC
ACTGCCTCGGTGGTAACGACGCGCTGGTGCCTGCGGCCGTGGAATCGGACCTGCTGGCGCTCAATGGC
AATGCCCGGGTGAGGATGCTGCCGCAGGGCAGTCATGCCCTGCCGCTGGAAGCGCCTCTCTGGCTGGC
GGGTGAAATCGCCGAGTGGCTGGAAGCGCAATGA

>CE19
ATGGTGGTCAATCTCTTTCAGCGTGGCAGGGGCAAGGCGCTGCTGCAGTCGATGGAAAAGGCAGAAAG
CTATGAGCAATGGAGTGAGCTGGCTGCTGCCTGGGATCACGAGGAGGGACTGGACGACTGGAAGCAGG
ATGACGCGTGCGAGAGCTATGACTACCGTTCGATTCGTCAGCGCCTGGATGTTCTTCGCGATCTGCGT
TTTCGCAAGGATTACCACCAATTGCTGTTTGTGCTTAACGAGGGCGTTCATGGCAACCTCGGCGGCAT
GGGCAAGCCGTCGCTGTATAACAAGGCCCGGCTTGGGACCAAGAATCTCATTACCCGCTACATTGACG
AACTGTGCGGCGCGCTTAACGACTTGAACGAGGTGGATGACAAGATCATCTCGTTGGCTGAAAAGCAG
GATTTCTTCCTGCGTGCCAGCCACTGCAACGGGCGGACCGCGCTGATGCTGAGTGGCGGCGCCGTACT

TGGCTTCTTCCACGCCGGTGTGCTCAAGGCGCTGTTCGATCAGGGGCTGCTGCCCGAAATCATCTCCG
GCAGCAGCGCCGGTTCTATCCTTGCCGCCACCGCCTGTACCCACGTTGATGAGGAGCTGCGCGAGCGG
CTGAGCCTCGACAACCTGCATCATGAAGTAGAACAGACCAGCGCGATTCGGCCTTCTCTCAGCCTGTT
TGGTGGCGGTCGGGCGATGGATGCCGACAACCTGCGTGACTACCTGGCCAAGATCATCCCGGATCTGA
CCTTTCAGGAAGCGTTTGAGCTGACCGGTCGCAAGCTCAATATCACTGTTACCGGGCTGTCGACCCAG
CAGGCCCCGCGTCTGCTTAACGCTATCACCTCGCCCAACGTGCTGATTCGCTCTGCGGTTACAGCCTC
ATGTGCGATCTACGGCATCTACCCGCCGGTAACGCTGATGTGCCGCAACGCAGCGGGCGAGACGGTAC
CCTACCTGCCGGGTATGAAATGGATCGACGGCTCCTTTGCCGATGATCTGCCGGCCAAGCGTCTGGCA
CGCCTGTACGGCGTCAATCACTTCATCTCCAGCATGACCAACCCGGCGGCGCTTGCCATCACGCCCGA
CCCGGATGCGCCGCGCAACCGGTTGCGCAATCTGGCTCATTTTCAGGCGCGCTTCCTGAAAATCGGTG
CCGCCGAAGCGATGCGCTTCAGCCGCGATAACGTCCGAATCAAGTCGCCGGTGCTCAGTCTGATGCAA
CATCTGACCTACGGCGTGCTGGCACAGGAATATACGGCAGATATCAATATCTTCCTGCGCAACCGCTG
GGATCACCCGCTGCGGCTGCTGGCACCGCCAAGCCGCGAAGCGATGCACCGCCTGATCCATGAAGGCG
AGCGCTCCACCTGGGAGAAGATCGAAATGGTCCGCAACTGCACCGCCGTCAGCCGGACGCTGGATGCC
ATCCTGCACACACGCGGTTGGGAAAAATAG

>CE20
ATGTCACCGCACATCATCGAACCCGCCACGCCCGCCACCCGCAGCGTCATCTGGCTGCACGGTCTGGG
CGCCGACTGCTACGATTTCGTGCCGGTAGTCGAGGCCCTGCAACTGCCCGCCGATCATGCCATCCGCT
TCATTTTTCCACAGGCACCGACCCGGCCGGTCACCATCAACGGCGGGTTCCCGATGCCCTGCTGGTAC
GACATCCTCGGCATGTCGCCAGCCCGGGCAATCAACCAGAGTCATCTGGATGAGGCCGTTGCCCTGGT
ACGTCAGCTGATCGACGAACAATGCGCCCAGGGCATTGCGCTGGAAAACATCATTCTCGCCGGGTTCT
CCCAGGGCGGCGCGGTTGTGCTCTGCACCGCTGTCAGCAGCGAGCTGCCGCTTGGCGGTGTGATGGCG
CTATCCACCTACGGGCCGGGACTGGACCTGCTGCTCGAGCAACACCCGGCCCGCCAACCGCTGGAGCT
GTTCTGCGCACACGGCCGTTTTGATGACGTGCTGCCATTGGCGATGGGCCGCGAAGCCCATGACCTGA
TGCAGGCCGCCGGCCACCAGACCCGCTGGTACGAATACCCGATGGCGCACGAGGTCTGCATGGACGAA
ATCGCCGACATCCGTCGCTGGTTGGTCGAACGCCTCGGACTTTGA

>CE21
ATGATCGGATTACGGATACGTACCCTGTTGGGCGGGATGTTGCTCGCCGGCGGCCTCGTCTCTGGCAG
TTGGGCTGACAGCGCGCCCGAGGCGTTCTCGCAGCAGCGTCTGGATGCCTTTGGTGCGACGATGGCCG
AGGAAGTCCAGCAGGGGCGAATCGCGGGCATTGCAACGCTGGTTTACCAAAACGGTGAGGTGGTACAG
CGTGGCCAATACGGGTTTGCGGATCGCGAGCAGGGCAAACCACTGACGCCCGACAGCCTCTACAAGAT
TTTCTCGCTGACCAAGCTGGTGACGGGCACTGCGCTGCTGACGCTTTACGACGAGGGTCGTTTCGAGC
TGGATGATCCGGTCGGTAAATACATTCCCGAGCTGCAAGATCTTCAGGTTGCCATCGCCGATGGCCCG
GACGGTATGCCCGAGACGGAGCCGTCAGCGCACCCGGTGACCATTCGTGAGCTGATGACCCATACCGG
CGGCTTCACCTATGGTCGCTTCGGCAATACCCAGGTCGATCAGCTGTATGTCAAGGCGGATATTCAGA
ACCCTGATTCCACCTTCGCCGACATGATGGCCAAGCTGCAGCACATTCCGCTGCTCTATCAGCCGGGC
ACCGTATGGAACTACAGCATCTCGGTTGATATCCAGGCCTATCTGGTTGAGGTGCTGTCGGGCAAATC
GTTTGATGTCTATCTGCAGGAGCACCTGTTCGATCCGCTGAAGATGGTCGATACCGACTTCTATGCGC
CGGCAGGCAAGGCCGAGCGACTGGCGCTGTCCTACAAGCCACAGGAAGACGGTTCGCTGGCGCCGTTG
CCCAATACACCATTTCTCAGCAAGCCGCGGTTTCTCAATGGCGGTGGCGGGCTGATCTCGTCAATGGA
CGACTATCTGCGGTTTGCGCGCATGTTGCTGGCTGGCGGCGAGCTGGACGGGGTGCGGATTCTCAAGC
CAGAAACGGTCGAGTTGATGCGTCAGAACCAGCTGCCGGAGGGCGTGGAGAATATCGGCCCGTTCTTT
CCCGGTAACCAGTTTGGCCTGAACGTGGCGGTGGTCAACGACTCGCCCGCTGCGGGTTATTTGCCCGA
AGGCACTTATTGGTGGTGGGGAATTCAAGGAGCCTGGTGCTGGATTGATCCGGCCAATCAGGCGATTG
TGATCGGCATGATGCAGAATACCGACTATCGTCTGTCACGGATGATTCACGGCAAGGCGAGTCGTGCG
CTGTATGGGCCGGCTGCGGGGCAGTGA

>CE22
ATGACAAGGCTGAAAAAGACGGCGTTGCTGACCATGGTGATCAGCAGCCTGACGCTGACAGGTTGTCT
GAGTGGGGGTGGCGGTGGATCGTCATCCAACGATCAGGCCGACAATCGGGTCCGGTTGCAGGACCAAC
GTGTGGAGGAAGGCTTTTTTAACGTGAATGAAGCCGGCTTGCCATTCGATGCCTTGCCGGAATACAGC
GACTCAAGCCGCTGGACAGGCGTACTGAACGGTGCAGGCTACCGAATTGAAGTGCCGGCAAACTGGAA
CGGTATGCTGGTGATGTACACCCACGGTTTCCGCATCAACAACATCGACCAGTTGACAGTCGACACTC
CGCCCATGCGCCGTTATTTGCTCGAAGAGGGCTACGCTTGGGCAGCCTCCAGTTACAGCGCCAACTTC

TACGATGTGCGTGCAGGTCTGGAAGATACCAACGCTTTGGCCCTGGCGTTCACTGATATTGCCGCCGA
AAACGGCCGCACACTGGATACGCCGGATAAATACTACATTACCGGTGTTTCCATGGGTGGTCACATTG
CTGCGGCGGCGGTAGAGCGGGAGACGCTGGCGACAGCCAATAATGTCGTCAATTACTCCGGTTCCGCA
CCGATGTGCGGCGTCGTCGGAGATACCGAGCTCTTCAATTTCTTTGCAGGCTACACCATCGCGCTTGC
AACACTATCCGGTAACCCTATTGATGAATTCCCTTTGACCCAAGAAGAAGCTGCAACGATCGTAGCCA
ACGCCCGCACTGTGCTCTGGAACGACTACACCACGAACAAGTTTCCGAGTGGCCTGACACAGCAAGGG
TTGGCGTTCTATCAGACGCTGAAGAACCTCAGTGGCGGTGAGCGACCGATGTATCCAATCAGCTTTGG
CCTGTTCCAGGACCTGCTGCAGGGCTTTGCCGGATCAGACGGCACTGTCGACGGCATTTTGCTGGATA
ACGTGGTCGACACCAACGCCATTACCTATCGCTTTGAAAGCGAGTTGGGCGAGCCCTTAACCACAGAC
GAAATCAATTTCAACAGCATGATCGTCAAGGCCACAGCAACAGAGGGTGCCAACGATCTGCGCGATGA
CGGCCTGCGCTGGATTCCGAAGGTCAATGGCGAGTTTGATGTACCCGTTGTCACCGCACATACCATTG
GTGATCTGTTTGTGCCGATTCTGATGGAACAGGAATACGCGCGGCGCGCAGCAGAAAATGGCAGTGAC
GACCTGCTGGTACAGCGCGCTATACGGGCTCCCGGCCACTGCGACTTTACCGAAGCAGAGTTCACTGC
CACCCTGGCGGCCATGCTTGACTGGGAGCAAGGCGGGCCGAAACCAGGCGGCGATGACTGGCTGACGC
CCGCCACCGTGGCGGCAGAAGACTTTGGCTGCACTTACACTATTGACGGACCGGTCGCAGAAGGGAAC
TATCCGCGCGCGGGCCTGCCATCCTGCACACCTGACTGA

>CE23
ATGGAGACGACGCCTGTGCAGACATTGCTTTCGAATTTTCCCCCCGTGCGGCTGGCACTGGCAGCTGT
GGTGACGACGTCCCTTTCCCTGACCGGCTGCCTGTCCGGTGGCGGTGGCTCATCATCTGATGGTGCCG
CAACGGTGGAAGAGCCAACCGATTTGCGCACCACCCTCTCTGGTGATGTGCGCGGGGTTGAGCAAGAG
GGCTATTTCGAGTTCTTGGGAATTCCGTACGCTGCAGCACCGGTCGGGGATTTGCGCTTCGCTGCGCC
ACAGGCAGCAGCTGGCTGGGACGAGGAGCGCGCTGCAGATGCCTACGGCAGTGCCTGTCCCCAGGCGG
GGCAGACCTTCGACGAGGCCGAGGACTGTCTGTACCTGAACGTGTTTACCCCCAAAAACCGGGGCCCG
CATCCGGTGATGGTGTGGTTCCATGGTGGCGCCTTTGTGTTTGGTTCAGGTGGTGGCACCTATGTACC
CCCGCGGCTGGTTGCGGAGGATATTGTTGTCGTGACGGTAAATTACCGCCTGGGCAAGCTGGGCTTTA
CCGCTCATCCGGGTTTGACTGAGGAGCAGGGCGCGTCCGGTAGTTACGGCATTCTGGATCAGCAGATG
GCGCTTCAGTGGGTGCAGGATAATATCGAAGCCTTTGCTGGTGACCCTGCCAATGTCACCATTTTTGG
TGAGTCGGCTGGCGGCTTGAGCGTTCTTAGCCATCTGGTGTCACCTGCCTCTGCCGGGCTGTTTGCCA
AGGCGATTGTGCAGAGTGGCTCCTACAGCAGGGTGCAGGCGTCAATGGCGGTGGCCGAAAGTGCCGGA
GAGGGCTTCGCTACCCTTCTTGGCTGTGATGCAGCGACAGCTGCGCAAGAGGTTGAGTGCATGCGCAG
CAAAAGTGTTGCGCAGATTATGGCTATCGCACCCGGGACGGTGACACCGACATTGCGGCCTGATGTAT
TGCCCGCCTCGGTCAATCAAGGCCTGGCAGCAGGCGAATTTAATGATGTGCCGTTGCTGATGGGAACC
AACTCGGATGAATGGAGTTATTTTCTGGCCAGCCGCGGGGAGTTGAATCCGATCACTCAGGCGAATTA
CCAGTTTCTGCTCAACAATAGCGTTGGACCGGTGGAAACGCCGAATGTTGTCGCGCTCTATCCGGCGT
CTGACTTCAGCGATGATTACGCGGCGTTGATGACTGCGGTGGGCACTGATGGCAATTTTGCCTGTAAT
GCCAGCGTGCAGGCAGCCAGCGTTGCGGAGAACGGCAGCCAGCCGCTATTCGTCTACGAGTTTGCCGA
TCGCGACGCGCCGTCGCCAGGTATTGTCGCTCCGAGTTGGCTGACGCTGGGGGCAACGCACGCCTCTG
AGCTGGCCTACATCTTCGGTACCGACGACTCGTTCCGAGTGCGCGGCGCAACCGACGATCAGGTCGCT
CTGGCCGATATCATGAGCCGTGCCTGGACGGCCTTCGCCCGGTCCGGTAATCCAAGCCATGACGACCT
GACCTGGGTAGATTACGGTGATAGTGCCGGTGGCATGGTCTCGTTCGAGACACCTGAGGTACAGCCTC
TTTCCCGTTCGGTGTTTCGCAGCGTGCATCGCTGTGACTACTGGACTCCTGAGGTCTGA

>CE24
ATGCAACTGCTGATAGGTCTGATCGGCCTGTTCGTTCTGATCGCGTTCAGCCTGCGTCGCTGGCTGCT
GCGCCGCGAGTCACCGCAGGTGCAGGCAGTCGATTTCGACGGTGAACTGTATCGAATCGGACAGGCCT
GCGTGGCCCGCCGCGGCCCGCAGGATGCCGAACAGACCGTTCTGGTGATGCCCGGGTTTGTGGAGAAT
TTTCTCTACTTCACCGAGCACTATGAGAATCCATCCCTGCAATTGATCCTGCTCACCAGTGCCGACTA
CCATGTGCCGGTCAACCGGCCCCGCTTCAGCGAGGCCGACTGGATCAGCCAGCCGAAGGCCCGCGCCG
GCACCATTGCCTACGACGCAGAGGTACTCAACCAGGCCCTGCAGCACCTGGCGACCGGCAAACAGATA
CGCGTGCATGGCCACTCACGCGGCGGCGCTGTAGTACTGGAGTCGGCGCGCCAGCGCCCCGACCTGTT
CACGCGGGTTGAGGTCATACTGGAGGCGCCGGTACTGCCCCAGGGCAAGCCCTATCGCGATGTACCGC
GATTCTTCCGCTGGTTTCTGCCCTTCTATCTGTTTGCCTGGCAACAGCAGCCCATCTCGGCCGCCAAC
CGCAGCCTGTTCGGCCCACTCGACGATCCGCGCAAGCGCGAGCTGATCATGGCGCTGCCATTCAACCC
GAAACGCGGCAGCACCTTCGTCCACAACATTCTGGATCTGGCCAGCTGGATGCCGACCAATACCCCGG
ACATCTATCAGCACGTCTGCCGCGGGGCGATTCTGGTTCCCAGCGCGGATCGTATTCTGGAGCCCCGT

GCGATGCTGCACAGCGCCCGCCAGGCAGAACGCCTGCAAATCATCGAAATCGAGGGCTGCAGCCACCT
GATCACCGCCGATCGCCCGGACTGCATTCCCCCGCTGACGACCTTGACCCCATGA

>CE25
ATGAATGCACAGATGTTCAAACTGGCCACACTCGCGCTCGGCTTGTCGCTTTCCAGCATCGCGCTGGC
CACCAACCCCGGTGGTGGTGGCGGCACCGGCAACCCGGCTACCGGGACCGGTTTCCCCGGCGTGAGCA
GCTTCAGTGCCGATGGCCCATTCGCGACGACCAGCGGCAGCGCGGGCAGCAGCTGCACGGTGTTTCGC
CCGTCCTCCCTCGGCGAGAACAACCGCAAACACCCGATCATCGTCTGGGGCAACGGCACCACTGCCTC
GCCCAGCACCTACGCAGCCCTGCTTGAGCACTGGGCCTCGCACGGTTTCGTCGTGATCGCAGCCAATA
CCTCCAACGCCGGCACCGGCCAGGAAATGCTGGGTTGCGTCGACTACCTGACCACCCAGAACAACCGC
AGCAGCGGCACCTATGCGAACAAGCTGGACCTCAACCGCATTGGCGCGGCCGGGCACTCCCAAGGCGG
TGGCGGCACCATCATGGCGGGGCAGGACTACCGCATCAAAGTCACTGCGCCGTTCCAGCCGTACACTA
TTGGTCTGGGCCACAACAGCAGCTCGCAATCCAATCAGAATGGTCCCATGTTCCTGATGACCGGCAGC
GCCGACACCATCGCCAGCCCGACCCTGAACGCCTTGCCAGTCTACAACCGCGCTAACGTGCCGGTATT
CTGGGGTGAACTGTCCAGAGCCAGCCATTTCGAGCCGGTCGGCAATGCGGGCGACTACCGTGGCCCCT
CAACCGCCTGGTTCCGCTACCATCTGATGGACGACGCCAGCGCCGAAGACACGTTCTACGGCAGCAAT
TGTGACCTGTGCAGCGACCGCGACTGGGATGTACGTCGCAAGGGCATCAACTGA

5.3 Protein sequences of CE01 to CE25 in FASTA file format

>CE01
MQKERKNQYPIVLVHGFAGWGRDEMLGVKYWGGMHDIQEDLKQYGYETHTAVVGPFSSNWDRACELYA
QLVGGTVDYGAAHAEKYGHDRFGRTYPGLLKNWDGEHKIHLIGHSMGGQTVRVLTQLLKEGSQEEREY
AKKHGVQLSPLFEGGKSWVHSVTTIATPNDGTTLADVVTQLIPAAQQIMGLAAAVSGNTNVPVYDFKL
DQWGLKRKAGESFVHYADRVWNSGIWTNTKDISAWDLKPEGAKELNNWVKAQPDVYYFSYSGEATFRS
LITGHHLPDLTMNKLITPFGIFLGCYGSDEKWWQNDGIVNTISMNGPKLGSTDEIVPYDGTPKIGKWN
DMGIQENWDHADYIGLSLSYVLGIEKIEDFYRGVADMLGSLSVR
>CE02
MNPAVIERATVRALMSLPGPVLERLAAGLETHSRPHLDSRLRFLLALSGAKPTLDSGTVEQARQIYRS
TLALLDMAPVSLPVVVDHQVSMEDGSQILVRRYRPADAPLVSPAIMFFHGGGFTIGGVEEYDRLCRYI
AKRTNAVVLSVDYRLAPEHPAPAGMDDALEAWRWLLNNTAQLGLDPNRLAVMGDSAGGCMSAVVSQQA
KLAGLALPALQVLIYPTTDAALAHPSVQTLGQGFGLDIPLLTWFRGHFVQDPAVIEDYRVSPLRNPDL
TGLPEAIVITATDPLRDEGLEYAQKLREAGNTVTSLDYPELIHGFISMGGVVPAARKAINDICVETKR
RL
>CE03
MASIPAHLMKLLLRAGVKRDIRDPDKLVKHLRRAMNAPLAPSPLPRGIRLQRGKVAGTAGHWLSPTDP
QTTILYLHGGAFIGGRLATYHNFCGHLARTLNARVFLPDYRLAPEHPFPAATDDAFNVYRELMADPRP
IVIAGDSAGGNLTLVTLLRARDHKLRMPACAVAISPASDARGNLMSRQANSDSDAMLSHCMIEVATDV
YLAGADPAHPYASPITQDFTGLPPLLFTVSSEECLRDDAYAAAHCARQAGVPVQLLERKDMPHVWPVF
TFLLPEAKQDLPTIVRFLRKYLATTDAHEEAFHTANESDTTIPEISSE
>CE04
MTAIIRQGRYQGLSSKGVTEYRGIPFAKAPLGEWRFKAPQPLPDSEDCVNADRYPLASLQPRNPIMGI
QESGEDCLYLNIWAPEGEGPFPVMVWFHGGGYMAGSTSQALYNGAELARSQKVVVVNAAYRLGAMGFA
DFSAVAPELDADTNLGLRDQLAALQWVQENIAAFAGDDKQVTIFGESAGGFSVCSLLACPQADELFQA
AIVQSGGADFVLAPDQVRKVTNAFVAALPGDGSAAEKLLSADNKGWIKAQNAAVKVLVDRGLRTTTPQ
FAMNFLPMVDGDVLPQLPVDAIAAGAAANKRVMAGVCRDEFNFFQYAGVLAGTTTMDALREISDEEIV
SRFERALPGNGRRAFDYYQTAVEPDARRSRLDWLAAMESDRLFRVPTVRLLDAQSQHAQCWGFQFTWP
SEPFGVPLGACHVVDVPFVFGVTDTPAGMYFTGGTSEARALSHQVQAAWGTFARGDAPGWNAWQSDRQ
VCQLGPGETMASLLDESGEQLWRDIIPVV
>CE05
MATYKAPLEDMRFVLNDVFKADQLWASMPATAEVTQDLSDAILEEAGKMTEGLLFPLNRNSDEQGCTW
TDGAVTTPDGFKEAFKTFAENGWSAFSGNPEFGGQGMPKSLAVLFEEMMHSACSSFALYPALTSGACL
AVDAHACEELKSQYLPKLYSGEWSGTMCLTEPHSGTDLGILRTKAVPNDDGSFNITGTKIFITGGEHD
LTGNHVHLVLAKLPDAPAGSKGISLFLVPKFLPDADNNPGEANGVTCGSIEHKMGIKGSATCVMNFDD

AKGWIIGEPNQGLACMFTMMNYERLSIGLQGLGLGEVSYQSAVEYARERLQGRSATGAKNPQGPADPI
IVHGDVRRMLMNMRAINEGGRALAAYVGMQLDTSKFSEDAEAKKKAEDRVALLTPVAKAFFTDRGLDT
TITGQQVFGGHGYIREWGMEQFVRDCRISQIYEGTNGIQALDLAGRKVVRNGGKSVDAFLADAQAWVD
ANADNAQLAAVKDDLQAALLLLKSSTDTLLAQAGNNPDAISAAAVEYLDIFGYVLYAWLWAQMLAATD
DRDDDFAKAKRITGQYYFQRVLPKAQSLAAQLNSGADAMMSLDAEQF
>CE06
MQSGTVASNGIELFYESRGPENGEPMVFVMGLSAQMVFWPDTLLDALAAKGYRVIRFDNRDVGKSTQI
RKPIKQGPVSAILRRIIGLPVESPYTLHDMVADTVGLLDALNIERAHFVGASMGGMISQLMAGTHPER
VLSLTSIMSSNNSSLLPPPKPSALRVLIAPRVKVETEEQFVTFGLEMMSKLAGTLPQGKEELAAMYRA
AWARGINPRGIRNQFLAITATGSLSKTLKQIQCPTTVIHGGADPLIRPAGGKASARAIRGAKLVIIPG
MGHDFPPSVIDRIGELIAETAGRANSVVPPAVG
>CE07
MSLQARLIKAVTKRTIKRSGLNQDQLVRHLRKVFNETPVLTLLPRGVKLSRVEHPAFTGDRISVQRPE
MAVLYLHGGAYIGGITKTYHNLAGRLAKKLNAEVFLPVYPFAPEHPYPAAVNRVMEAYEYLLSLGKQP
QDIVIAGDSAGGGLTLATLLHIRDKGLEQPRCAVTFSPASNAFPDDSILEALDPSDAMLSADIIRTAI
EIYTPNPEDRSQPYASPCLGDYTNICPLLITASTDELLYADGKRVKRVAEKAGVKVTWIERPGVFHVW
PVMVPFLPEANKDLKRIVAFIKEA
>CE08
MSRATSGSALTPVSSFSLVADSDQAFTLRVTLARLATRTLDMQYYIWDDDTTGKLLIYRVLEAARRGV
RVRMLLDHANQLGRDVKWAVLDAHPNIEVRLFNPFKGRYKHFLQWLYYAPRLNHRMHNKAWIMDGERA
LVGGRNVADHYFGVNPSTNFRDLDLYAHGAIVADTQMAFDAFWDSPLAVSMKSYRHRTEATADRAWRW
LGRWRTSLQGYPYLFPQHEAFFTDYLAKQEQQCVQAPAALLFDSPDKASGAKQTLMGEQLARLLGRQD
NRELLMEASYFIPGDAFVEALGGFQQRGGRAAVLTNSLATNDVIAAHGAYARYRAALLAAGVELHELQ
PNARALHRQVRLFKGRSQASLHTKAMVLDRREVFVGSFNIDPRSVHLNTEMGYYVVSAELGQQVAAFI
EEGLAPENSYRLEKQEGELLWGADGYDGKPVLYSREPKATVWRRMLAWLLSLLPIERML
>CE09
MSLFVDRISRSQSTLQQVLAGTLRQSLALLFKPVANPRFSFTAQRRWLDIMAHSTLKARGIGSHDNAL
ANVPCRHYQPLHAQTTVLYLHGGGYVAGSPDSHKSITSHLAKFANAHVVVPDYRLAPEHPCPAAIEDA
VAVYLALLESGVSPGSLTLMGDSAGGGLALATLQALKASGKPLPSSVILFSPWVDLTLGQLYDTDRDI
MLSHSWLASAADAYAGNDTRNVQCSPLFGELSDLPPTLIQAGGDEILLNDSHRLCAALNEAGTPTRLQ
IHPQRWHDFQLHAGVLADSDHALMTCARFIHQHAIHDHEEDAHA
>CE10
MDLIIFLLSLLVMSTVTLTLAFYLCWYYDRCSFPEQSELPGEPPLRLLPTLLGMAKEAAALTVLVLSY
PLRLIHDSSPVRSRHHGEPPVILVHGYGGNSANFLFMQWRLKWRGWSNVYSVSYTPPHINARKLSQQV
VDHVERILATTGAEKAHLVCHSMGGPLTRYALKNLGLAGKVDRVITLGSPHYGTRVAGLFPPLGAAAQ
LRYQSPFIRELATEATCPGGARYFSIFSNMDNFVLPVSAAVLHGAEDNIHVPYLGHCSLLYSSRVMDQ
VERCLLAPPKAAE
>CE11
MPVPETSLIYIHGFNSSPASHKAGLLRSVFEQAGCPERLVVPALPPSPVLAMAELDAAIAGAGPVALV
GSSLGGFYATWLAEHHGLKAAIVNPAVRPWRLLDKYTGIQHNYHTGEAYQFDPAWLEQLRRYEVAELT
QPENLLLLTQTGDETLDWEDGWDYYGDCHRYCGLGGSHGFDNFDAFIPLMLRFCGIEIHA
>CE12
MEPLELEDQFVTTEAGYQLHYQSAGTGEPLIFLHGGGPGATSFGNFYYNAPAFLEQYQCFFYNMPGYG
QSSKLVVEAPMYSFHASMLAEFMDLVGIPRAHLVCQSFGGCAAIKLAIDHPERVNKLVLMGAQPMFGG
VIDPLKLMSKHAANIILDYYGGEGPTPEKMRRLLADYEMHDDSKLTDWTVNDRFANSTDPELLEVART
PGFMGQPESLLGDLHRNVVPTLCLWGLHDWFGGPDVPLLFLNQFANAQLFIEGQGAHHWQTELPERFN
RVVLSYLSE
>CE13
MSTARLTPTTASTLKLLSVAVAALTLSACLSGGGGGRADGPNPLAIETSEGKVVGISNDGIRVFRGIP
YAAPPVGDLRLAPPQPPASRSETLRLSEEFGNSCPQSDLTTGQQVGNEDCLYLNVYAPAEAEDLPVMV
WIHGGAFVFGNGGGEYDPTRLVEQDVIVVTLNYRLGNLGFLAHPALESDAGNFALMDQQLALAWVKEN
IAAFGGDPANVTIFGESAGGHSVMSHIVSPRAEEADLFQRAIVQSGSYAPFQMPKATAQFLGTSVANG
LGCTDPETAASCLRSLPVSAFLAAQGSQSIPVVDPDDDLLPKSIQQALADGDFNSSLDIMIGSNQNEG
TLFVALDEVGGDPIDDEAEYRERVAEFFQPYQASIPFDDDQIATDYLDFVDGAAKPFEAALSGIWTDF
MFACNAYSQASTFAGASMNTFQYWFRDEDAPWTLVPPFAVSFPLGATHAGEIPYVLYPQAIMEQRYTG

DPDDLNSLAGEMVDYWTQFAKTGDPNTTDGVAAAWQQAATGNLLTLDVPNASNANTLGFLGYHHCSYW
ADPPLVLP
>CE14
MANSNCIAGVDLERLSRIGTHLQSAYIDAGKLPGALTLVARRGQVVYCEAQGLRDVERGLPVERDTLF
RIYSMTKPVTSVALMQLYEQGRFLLDDPVHKYIPSWSNLRVYKTGSHPQILTTPCERPMTIRDLLTHQ
SGLTYGFMNRTNVDAAYRALKLDGGPGLTLDRLIDELSRLPLEFSPGTAWNYSVATDVCGYLVQVLSG
MSLDDYFKRHIFDPLEMSDTFFTVPADRIPRFAACYQFQPGGSFSLQDDPQDSSFTKAHGFLSGGGGL
VSTIDDYYRFAQALANGGELNGARIIGRKTLEFMRMNHLPGNQDLPGVSVGSFSETPYEGSGFGLGFS
VKIDVAKSQTNGSVGEYGWGGMASTNFFVDPVEDLLMVFMTQLIPSSSYPVRQELRAIINGALVDE
>CE15
MSRYVDELYRNPGQPGLRALLRGMLKLLFRGLIRPPVPFAVQALVLRLLTLGMPLAKGVTRSAEQIAG
RSCMWHRPAAGGNGRVLLYLHGGAFVIGSPQTHRGICSALASRGQFDVCALDYRLAPAHPAPAACDDA
VAAYQALLQRGYTPAQITLVGDSAGGNLVLVTAQKLAALKLPLPASLVCFSPVTDMTAEQLHAPAAGD
PLLHPSWLDSARDAYCPAGLDRADPLVSPLFGQLKGLPPLLLQVGEDEVLLNDSLRLAEAARAADVAV
RLERYTDLWHVFQAHTGLLHSADAALQRVVEFVNSAQADSVS
>CE16
MPFNKKSVLALCGAGALLFSMSALANNPAPTDPGDSGGGSAYQRGPDPSVSFLEADRGQYSVRSSRVS
SLVSGFGGGTIYYPTGTTGTMGAVVVIPGFVSAESSIDWWGPKLASYGFVVMTIDTNTGFDQPPSRAR
QINNALDYLVSQNSRSSSPVRGMIDTNRLGVIGWSMGGGGTLRVASEGRIKAAIPLAPWDTTSYYASR
SQAPTLIFACESDVIAPVLQHASPFYNSLPSSIDKAFVEINGGSHYCGNGGSIYNDVLSRFGVSWMKL
HLDEDSRYKQFLCGPNHTSDSQISDYRGNCPYLE
>CE17
MHTLFKRGLAALALSTLVSLPAMASNPGISSPDTMILGDSIFALSGDIHENLEADLDENIDTYARSGC
QLTGGNVLCSRLYSVENQYARADKSGIRTVIFNGGGNDIQLNSCRPSLSACMPLLNELEDRIATLVQK
MRNDGIEEIIFLGYYNAAGSAENLQDINNYSMNYKAAAYPGMGVKFIDVRADFAGRESIYITSDGIHP
TAAGSRVLSNRILQALD
>CE18
MNNLTLLPGWALAASNLLPLQQALHERLPDLTVQLAELPPRLQMSTLEPDLTALAEALPAGWLAGWSL
GGMLAVQLQRRFPERFFGVITIASNACFAARNGWPHAMPVDTFKAFLADAREQPERILKRFSLLVTQG
GENARELSKQLQWSDADPLQRLNQLALLGVIDNRIPLQRCAQPVLHCLGGNDALVPAAVESDLLALNG
NARVRMLPQGSHALPLEAPLWLAGEIAEWLEAQ
>CE19
MVVNLFQRGRGKALLQSMEKAESYEQWSELAAAWDHEEGLDDWKQDDACESYDYRSIRQRLDVLRDLR
FRKDYHQLLFVLNEGVHGNLGGMGKPSLYNKARLGTKNLITRYIDELCGALNDLNEVDDKIISLAEKQ
DFFLRASHCNGRTALMLSGGAVLGFFHAGVLKALFDQGLLPEIISGSSAGSILAATACTHVDEELRER
LSLDNLHHEVEQTSAIRPSLSLFGGGRAMDADNLRDYLAKIIPDLTFQEAFELTGRKLNITVTGLSTQ
QAPRLLNAITSPNVLIRSAVTASCAIYGIYPPVTLMCRNAAGETVPYLPGMKWIDGSFADDLPAKRLA
RLYGVNHFISSMTNPAALAITPDPDAPRNRLRNLAHFQARFLKIGAAEAMRFSRDNVRIKSPVLSLMQ
HLTYGVLAQEYTADINIFLRNRWDHPLRLLAPPSREAMHRLIHEGERSTWEKIEMVRNCTAVSRTLDA
ILHTRGWEK
>CE20
MSPHIIEPATPATRSVIWLHGLGADCYDFVPVVEALQLPADHAIRFIFPQAPTRPVTINGGFPMPCWY
DILGMSPARAINQSHLDEAVALVRQLIDEQCAQGIALENIILAGFSQGGAVVLCTAVSSELPLGGVMA
LSTYGPGLDLLLEQHPARQPLELFCAHGRFDDVLPLAMGREAHDLMQAAGHQTRWYEYPMAHEVCMDE
IADIRRWLVERLGL
>CE21
MSRTTMIGLRIRTLLGGMLLAGGLVSGSWADSAPEAFSQQRLDAFGATMAEEVQQGRIAGIATLVYQN
GEVVQRGQYGFADREQGKPLTPDSLYKIFSLTKLVTGTALLTLYDEGRFELDDPVGKYIPELQDLQVA
IADGPDGMPETEPSAHPVTIRELMTHTGGFTYGRFGNTQVDQLYVKADIQNPDSTFADMMAKLQHIPL
LYQPGTVWNYSISVDIQAYLVEVLSGKSFDVYLQEHLFDPLKMVDTDFYAPAGKAERLALSYKPQEDG
SLAPLPNTPFLSKPRFLNGGGGLISSMDDYLRFARMLLAGGELDGVRILKPETVELMRQNQLPEGVEN
IGPFFPGNQFGLNVAVVNDSPAAGYLPEGTYWWWGIQGAWCWIDPANQAIVIGMMQNTDYRLSRMIHG
KASRALYGPAAGQ
>CE22
MRIHTRENVSKRDTRYLPFASSHPLLKTARTNKQQSTCAWRTTRMTRLKKTALLTMVISSLTLTGCLS
GGGGGSSSNDQADNRVRLQDQRVEEGFFNVNEAGLPFDALPEYSDSSRWTGVLNGAGYRIEVPANWNG

MLVMYTHGFRINNIDQLTVDTPPMRRYLLEEGYAWAASSYSANFYDVRAGLEDTNALALAFTDIAAEN
GRTLDTPDKYYITGVSMGGHIAAAAVERETLATANNVVNYSGSAPMCGVVGDTELFNFFAGYTIALAT
LSGNPIDEFPLTQEEAATIVANARTVLWNDYTTNKFPSGLTQQGLAFYQTLKNLSGGERPMYPISFGL
FQDLLQGFAGSDGTVDGILLDNVVDTNAITYRFESELGEPLTTDEINFNSMIVKATATEGANDLRDDG
LRWIPKVNGEFDVPVVTAHTIGDLFVPILMEQEYARRAAENGSDDLLVQRAIRAPGHCDFTEAEFTAT
LAAMLDWEQGGPKPGGDDWLTPATVAAEDFGCTYTIDGPVAEGNYPRAGLPSCTPD
>CE23
MQTLLSNFPPVRLALAAVVTTSLSLTGCLSGGGGSSSDGAATVEEPTDLRTTLSGDVRGVEQEGYFEF
LGIPYAAAPVGDLRFAAPQAAAGWDEERAADAYGSACPQAGQTFDEAEDCLYLNVFTPKNRGPHPVMV
WFHGGAFVFGSGGGTYVPPRLVAEDIVVVTVNYRLGKLGFTAHPGLTEEQGASGSYGILDQQMALQWV
QDNIEAFAGDPANVTIFGESAGGLSVLSHLVSPASAGLFAKAIVQSGSYSRVQASMAVAESAGEGFAT
LLGCDAATAAQEVECMRSKSVAQIMAIAPGTVTPTLRPDVLPASVNQGLAAGEFNDVPLLMGTNSDEW
SYFLASRGELNPITQANYQFLLNNSVGPVETPNVVALYPASDFSDDYAALMTAVGTDGNFACNASVQA
ASVAENGSQPLFVYEFADRDAPSPGIVAPSWLTLGATHASELAYIFGTDDSFRVRGATDDQVALADIM
SRAWTAFARSGNPSHDDLTWVDYGDSAGGMVSFETPEVQPLSRSVFRSVHRCDYWTPEV
>CE24
MQLLIGLIGLFVLIAFSLRRWLLRRESPQVQAVDFDGELYRIGQACVARRGPQDAEQTVLVMPGFVEN
FLYFTEHYENPSLQLILLTSADYHVPVNRPRFSEADWISQPKARAGTIAYDAEVLNQALQHLATGKQI
RVHGHSRGGAVVLESARQRPDLFTRVEVILEAPVLPQGKPYRDVPRFFRWFLPFYLFAWQQQPISAAN
RSLFGPLDDPRKRELIMALPFNPKRGSTFVHNILDLASWMPTNTPDIYQHVCRGAILVPSADRILEPR
AMLHSARQAERLQIIEIEGCSHLITADRPDCIPPLTTLTP
>CE25
MIRATLPYTQAPYREPSAHLAQPFQQQDKRTHRMNAQMFKLATLALGLSLSSIALATNPGGGGGTGNP
ATGTGFPGVSSFSADGPFATTSGSAGSSCTVFRPSSLGENNRKHPIIVWGNGTTASPSTYAALLEHWA
SHGFVVIAANTSNAGTGQEMLGCVDYLTTQNNRSSGTYANKLDLNRIGAAGHSQGGGGTIMAGQDYRI
KVTAPFQPYTIGLGHNSSSQSNQNGPMFLMTGSADTIASPTLNALPVYNRANVPVFWGELSRASHFEP
VGNAGDYRGPSTAWFRYHLMDDASAEDTFYGSNCDLCSDRDWDVRRKGIN

6 Acknowledgement

The presented thesis was conducted in frame of the EU H2020 project “Industrial Applications of Marine Enzymes” (INMARE), whereby many of the presented studies were facilitated by numerous fruitful collaborations with partners of the INMARE consortium. I am deeply grateful for the opportunity to contribute to this fascinating project and for the chance to meet so many amazing people. Thank you, Carla, Cristina, Manolo, Rainhard, and the complete INMARE consortium.

For the excellent supervision, for considering me to join the INMARE team, for all his valuable advice over the past years, and for his continuous confidence in my work I want to thank my supervisor Prof. Karl-Erich Jaeger. Likewise, I want to thank Prof. Michael Feldbrügge for his mentorship and the co-examination of my thesis.

For the excellent cooperation on the protein crystallization studies I want to thank Astrid, Stefanie and Sander from the Center for Structural Studies (CSS) at the Heinrich Heine University Düsseldorf.

Furthermore, I gratefully acknowledge the dedicated work of all highly motivated students, I could supervise in the lab. Thank you, Alina, Debby, Lesley, Lejla, Maha, Marvin, Simone, and Vivien. To work together with you has driven me throughout the thesis and I really enjoyed accompanying you part of your way to become scientists.

The Institute for molecular enzyme technology has become a second home during my thesis for sure, thus I want to thank the complete IMET-staff for their continuous support, their unreserved kindness, and the perfectly comfortable atmosphere every day. A special thanks to my colleague Rebecka, for her commitment to our study on organic solvent tolerant enzymes and her motivation and passion for our joint projects. I want to acknowledge especially Anita and Tom for the helpful as well as motivating discussions during my thesis. Just as well, I want to thank Stephan for his supervision, many hours of fruitful discussions, his company in the lab, his unique appreciative kind of collaboration, and his guidance throughout my thesis.

Finally, I want to thank my dear family and wife, who did not get tired of supporting me throughout the past years of studying and completing my doctoral thesis.

7 Erklärung

Hiermit versichere ich, dass die vorliegende Dissertation von mir selbstständig ohne unerlaubte Hilfe verfasst wurde und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet wurden. Diese Dissertationsschrift wurde in der vorliegenden oder einer ähnlichen Form noch bei keiner anderen Fakultät eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den _____ _____

Alexander Bollinger