Novel biocatalysts for technical applications by evolution and design

Habilitationsschrift zur Erlangung

der

Venia legendi

für das Fachgebiet

Molekulare Mikrobiologie

an der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Heinrich-Heine-Universität Düsseldorf

von

Dr. rer. nat. Thorsten Eggert

Dezember 2006

Für Barbara und Laila

Table of contents

Chapter 1:

Wer uns vor nutzlosen Wegen warnt,
leistet uns einen ebenso guten Dienst wie derjenige,
der uns den rechten Weg anzeigt.

Heinrich Heine

Zusammenfassung

Die Nutzung von Enzymen als biologische Katalysatoren in der organischen Synthese hat gerade in den letzten Jahren enorm an Bedeutung hinzugewonnen, da deren hohe katalytische Effizienz und Selektivität Vorteile bietet. Einige Verbindungen für den Einsatz in der pharmazeutischen Industrie können aufgrund ihrer Komplexität ausschließlich auf dem Wege der Biokatalyse hergestellt werden. Ferner ermöglicht die Fähigkeit von Biokatalysatoren, unter milden Reaktionsbedingungen (neutraler pH-Wert, Temperatur von 20-30°C, Normaldruck) aktiv zu sein den kostenreduzierenden Einsatz durch geringeren Energieaufwand beim Einstellen der optimalen Reaktionsbedingungen. Darüber hinaus kommen Umweltschutzaspekte in Betracht, da oft weniger toxische Nebenprodukte während der Reaktion entstehen. Aus den vorgenannten Gründen spielt die industrielle, oder auch **Weiße Biotechnologie** wie sie heute genannt wird, eine bedeutende Rolle neben der klassischen chemischen Synthese. Eine Vielzahl von erfolgreichen Beispielen zeigt die Leistungsfähigkeit von isolierten Enzymen oder ganzen mikrobiellen Zellen bei der Katalyse zur Herstellung organischer Verbindungen. So ist es beispielsweise möglich, durch Biokatalyse chirale Bausteine herzustellen, die mit klassischen (Chemo-)Katalysatoren nur schlecht oder gar nicht zugänglich sind.

Nicht zuletzt aufgrund dieser Leistungsfähigkeit bekannter Enzyme steigt die Nachfrage an geeigneten neuen Biokatalysatoren in den letzten Jahren rapide an. Eine Herausforderung an die Weiße Biotechnologie in den kommenden Jahren wird daher die Bereitstellung neuer, aktiver und stabiler Enzyme für den technischen Einsatz sein. Auch wenn diese Aussage zunächst trivial erscheint, müssen hierzu neue wissenschaftlich fundierte und technisch umsetzbare Lösungswege entwickelt werden.

In der vorliegenden Habilitationsschrift werden neue Strategien beschrieben für drei Teilaspekte zur Bereitstellung neuer Biokatalysatoren: (i) die Identifizierung neuer Enzyme, (ii) ihre Produktion in mikrobiellen Expressionswirten (iii) sowie ihre Optimierung durch molekulares Engineering. Das Ziel ist hierbei, ausgehend von einem gewünschten Produkt, geeignete Biokatalysatoren für die entsprechende Reaktion zu finden, die später sogar in technische Prozesse eingesetzt werden können (Abb. 1.1).
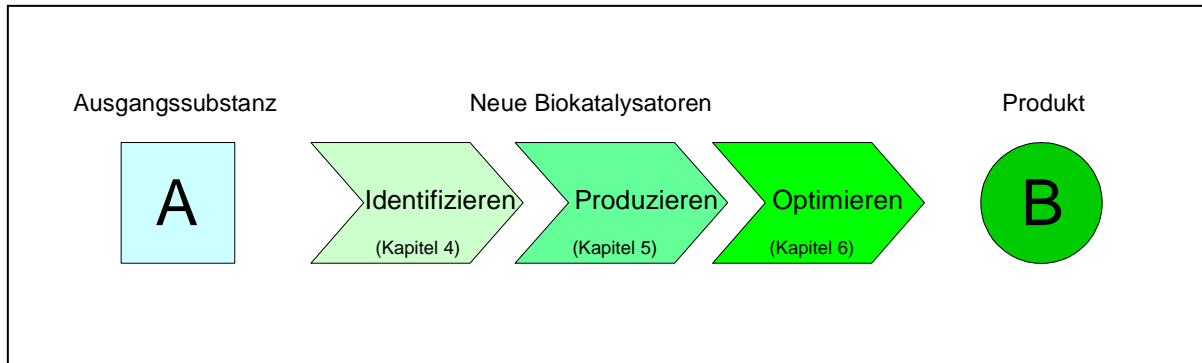
Abb. 1.1: Der Weg von einer gewünschten chemischen Reaktion zum biokatalytischen Prozess erfordert mehrere Teilschritte. Diese Teilschritte umfassen die Identifizierung, die Produktion sowie die Optimierung der Biokatalysatoren, wie es in den Kapiteln 4 bis 5 in dieser Arbeit beispielhaft an ausgesuchten Modellenzymen und organismen beschrieben wird.

In **Kapitel 4** der Arbeit werden die Probleme bei der Identifizierung neuer enzymkodierender Gene thematisiert. Am Beispiel von drei verschiedenen Reaktionstypen wird die Entwicklung von Hochdurchsatz-fähigen Screening Methoden beschrieben. Ein gekoppelter Enzymtest, bei dem enantioselektive (*R*)- und (*S*)-spezifische Alkoholdehydrogenasen sowie Diaphorase verwendet wird, konnte zum Nachweis chiraler Alkohole für einen automatisierten Hochdurchsatz-Assay entwickelt werden. Reproduzierbare Messungen der Enantiomerenüberschüsse (*ee*) konnten bei der durch Lipase katalysierten Esterhydrolyse bis in den μM-Bereich erfolgen. Weiterhin wurden spektrophotometrische Assays für die Identifizierung neuer Hydroxynitril Lyasen und Benzoylformiat Decarboxylasen entwickelt. Beide Enzymklassen katalysieren für die organische Synthese interessante C-C-Verknüpfungs-Reaktionen. Darüber hinaus wird ein Wuchsselektions-System auf Basis eines *Pseudomonas putida* Stammes beschrieben, das zur Durchmusterung sehr umfangreicher Expressions-Bibliotheken (Genom-, Metagenom-, Zufallsmutagenese-Banken) eingesetzt werden kann, um neue Benzoylformiat Decarboxylasen sowie andere Enzymklassen, deren Aktivität zur auf Freisetzung von Benzaldehyd führt, zu identifizieren.

Der zweite Themenkomplex, der sich an die Identifizierung neuer Enzym-Gene anschließt, ist die Produktion von Biokatalysatoren in mikrobiellen Wirten. Im **Kapitel 5** wird dieser Schwerpunkt aufgegriffen und am Beispiel der Enzymproduktion mit anschließender Sekretion im Gram-positiven Modellorganismus *Bacillus subtilis* erörtert. Neben der Entwicklung neuer Vektoren für die Überexpression und Sekretion in *B. subtilis* steht vor allem die Optimierung der Sekretion im Mittelpunkt. Die dargestellten Arbeiten untersuchen zum ersten Mal systematisch die Bedeutung des N-terminalen Signalpeptides in Bezug auf die Sekretionseffizienz heterologer Proteine im Modellorganismus. Anhand einer kompletten Bibliothek der natürlichen *B. subtilis* Signalpeptide sowie unter Einsatz von Methoden der gerichteten Evolution wird in der vorliegenden Arbeit eine effiziente Strategie für die

Identifizierung geeigneter Sekretionssignale vorgestellt. Ferner wird anhand eines abgeleiteten Modells die Hypothese eines „balancierten Gleichgewichts" aller beteiligten Komponenten der Sekretionsmaschinerie aufgestellt, wonach je nach Faltungseffizienz des Targets auf der *trans*-Seite der Membran zum Teil eine verringerte Translokation / Prozessierung zu einer Steigerung der Sekretionseffizienz führt. Bisher wurde eine effiziente Translokation / Prozessierung immer mit einer hohen Sekretionsleistung für das heterologe Protein gleichgesetzt.

Das dritte Arbeitsfeld, dargestellt in **Kapitel 6**, umfasst die problembezogene Optimierung von Biokatalysatoren. Wenn geeignete Enzyme erst einmal identifiziert und schließlich auch in ausreichender Menge produziert werden können, schließt sich oft die Frage nach einer dem geplanten biokatalytischen Prozess angemessene Optimierung an. Stabilitäts- oder Selektivitätsansprüche stehen hierbei meist im Vordergrund. Die gerichtete Evolution, also die Auslese im Reagenzglas der am besten angepassten Variante nach dem Darwinschen Prinzip, zählt inzwischen neben dem rationalen Design zu den Standardtechnologien in der molekularen Optimierung von Biokatalysatoren. In dieser Arbeit werden neue Mutagenesemethoden vorgestellt, die an ausgewählten Modellenzymen erprobt und den klassischen Strategien der sog. *error prone polymerase chain reaction* (epPCR) und dem DNA shuffling gegenübergestellt wurden. Die komplette Sättigungsmutagenese, gezeigt am Beispiel der extrazellulären Lipase A aus *B. subtilis* (BSLA), wird im Rahmen dieser Arbeit als zuverlässige Mutagenesestrategie für eine erste Generation in der gerichteten Evolution vorgeschlagen. Für eine zweite Generation, bei der üblicherweise rekombinative Ansätze zum Einsatz kommen, wird eine technisch einfache und effektive neuentwickelte Methode auf Basis der Multiplex-PCR vorgestellt und ebenfalls am Beispiel der BSLA auf ihre Leistungsfähig geprüft.

Zum Abschluss werden zwei Projekte zur Theorie-unterstützten *in vitro* Evolution beschrieben, die in Zukunft sicherlich weiter an Bedeutung gewinnen wird, um die molekulare Optimierung von Enzymen signifikant zu beschleunigen. Rationales Design und gerichtete Evolution wurden kombiniert, um in die von Natur aus minimale α/β-Hydrolase BSLA zusätzliche Domänen einzubringen, ohne das allgemeine Faltungsprinzip zu zerstören. Durch Computer-Modelling wurden geeignete Positionen im Enzym ausgewählt und neue Aminosäure-Sequenzen eingefügt. Im Anschluss daran wurden durch gerichtete Evolution das BSLA-Rückgrat und die neu eingebrachten Domänen aneinander angepasst. Diese gewissermaßen „rationalisierte" Evolution erweiterte effektiv die katalytischen Eigenschaften des zu optimierenden Enzyms. Des Weiteren konnte gezeigt werdem, dass theoretische Methoden wie QM/MM-Rechnungen in der Lage sind, aufwendige Labor-Experimente durch Computersimulation zu ersetzen, wodurch Methoden der Zufallsmutagenese zielgerichteter einsetzt werden können. Ziel ist die Begrenzung des

Umfanges der Variantenbibliotheken, was wiederum den Screeningaufwand signifikant reduzieren hilft. Schließlich führt eine gesteigerte Qualität bei der Erstellung von Zufallsmutanten-Bibliotheken zu einer Verbesserung der Ausbeute in Optimierungs-prozessen.

Chapter 2:

*An investment in knowledge
always pays the best interest.*

Benjamin Franklin

General introduction

Thirty years ago, on 7th of April 1976, Herbert W. Boyer and Robert A. Swanson founded the first biotech company named Genentech Inc. and by this it is believed today that they have founded the biotechnology industry. Nowadays, the industrial biotechnology – also known as **White Biotechnology** – becomes more and more important as a complement to the classical chemical industry. An increasing number of successful examples demonstrate the power of utilising natural (bio-)catalysts in industrial processes. Both isolated enzymes and whole (microbial) cells are used by the chemical industries in various market fields like nutrition, wellness, pharmacy, agro and fine chemical production. The production of L-amino acids by Degussa (Düsseldorf, Germany) or acrylamide by Nitto Chemical Industry Co. Ltd. (Tokyo, Japan) represent prominent examples of whole cell biotransformation, whereas the lipase catalyzed production of an intermediate in the synthesis of diltiazem by Tanabe Seiyaku Co. Ltd. (Osaka, Japan) or the lipase catalyzed production of chiral amines and alcohols by BASF (Ludwigshafen, Germany) are biocatalytic processes using isolated enzymes.
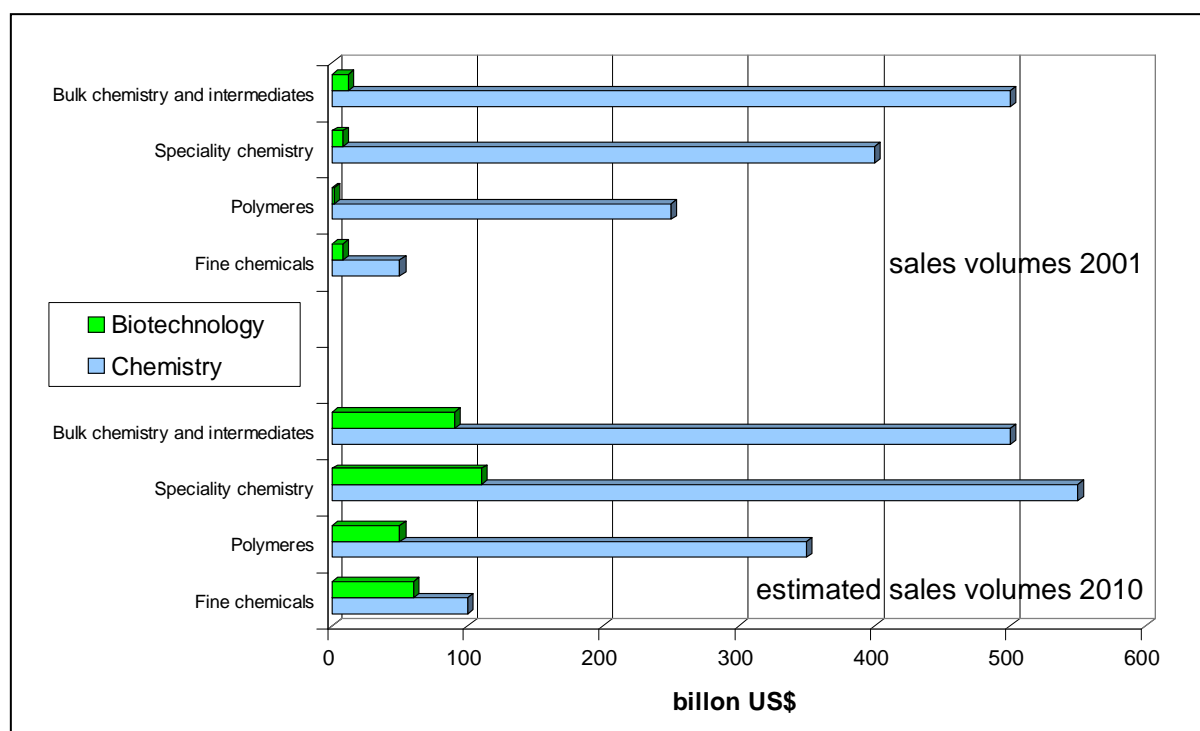
Fig. 2.1: Development of biotechnological processes for the production of organic compounds in different sectors of chemical industry. In all fields of chemistry (bulk chemistry and intermediates, speciality chemistry, polymers and fine chemicals) the percentage of biotechnological processes will increase in the next years. The highest substitution rate of classical chemistry is expected to occur in the field of fine chemicals. Here, about 60 % of all products might use biotechnology in their production process. Source: Festel Capital.

Nevertheless, considering the chemical industry as a whole, the implementation of biocatalytic processes is still a niche. However, this niche will expand tremendously over the next years as predicted by various market research institutions (Fig. 2.1). The greatest impact of industrial biotechnology, beside the fuel sector, is expected in the fine chemicals segment, where by 2010 up to 60 % of products may use biotechnology in their production processes. A key driver here is the growth of biological pharmaceuticals for which no traditional chemical synthesis exists. This outlook is given by the chemical industry in the European Union organized in the European Technology Platform for sustainable chemistry called SusChem [3] indicating the importance and prospects of White Biotechnology in the near future. Major advantages of biocatalysis compared to conventional (chemo)catalysis is the high chemo-, regio- and enantioselectivities towards complex substrates achieved by the use of microbial cells or isolated enzymes. Furthermore, energy consumption and toxic waste products can be reduced because biocatalysts often work in aqueous solutions at moderate temperatures. So far not for every chemical reaction a corresponding biocatalytic equivalent exists; however, biocatalysis experts state the theoretical existence of enzymes (i.e. naturally occurring or molecular engineered) catalyzing any kind of reaction which is possible by conventional (chemo)catalysis [4].

This indicates the necessity of identifying novel active and stable enzymes as a prerequisite to successfully carry out biocatalysis. This statement appears to be trivial; however, the identification of new enzyme coding genes and their overexpression in microbial hosts to provide sufficient amounts of biocatalyst is the major challenge of White Biotechnology in the forthcoming decade. In order to establish biocatalytic processes for technical applications from scratch three topics are of major importance: (i) the identification of novel enzyme coding genes; (ii) the efficient production of biocatalysts by overexpression and secretion in appropriate microbial hosts. (iii) Biocatalysts isolated from nature might not fulfil the high standards of chemical processes, which makes optimization necessary, with respect to process parameters (e.g. thermostability, pH-optimum, solvent stability, enantioselectivity, etc.). Here, directed evolution and/or rational design approaches might push a promising (natural) enzyme to a (high-tech) biocatalyst efficiently applied in industry.

Chapter 3:

Scope and objectives

The objective of this habilitation thesis is the development of general concepts to make novel biocatalysts available for technical applications. As outlined in figure 3.1 different embarrassments have to be overcome to establish a new competitive biocatalytic process. The focus of chapter 4 is the development of high-throughput screening and selection systems for various enzyme activities, which is the key technology either in identification of novel biocatalysts or improved enzyme variants in directed evolution approaches. In chapter 5 strategies for the efficient production and subsequent secretion of biocatalysts in the expression host *Bacillus subtilis* are presented. New methods and strategies for biocatalyst improvement are introduced in chapter 6 mainly using bacterial lipases as model enzymes.
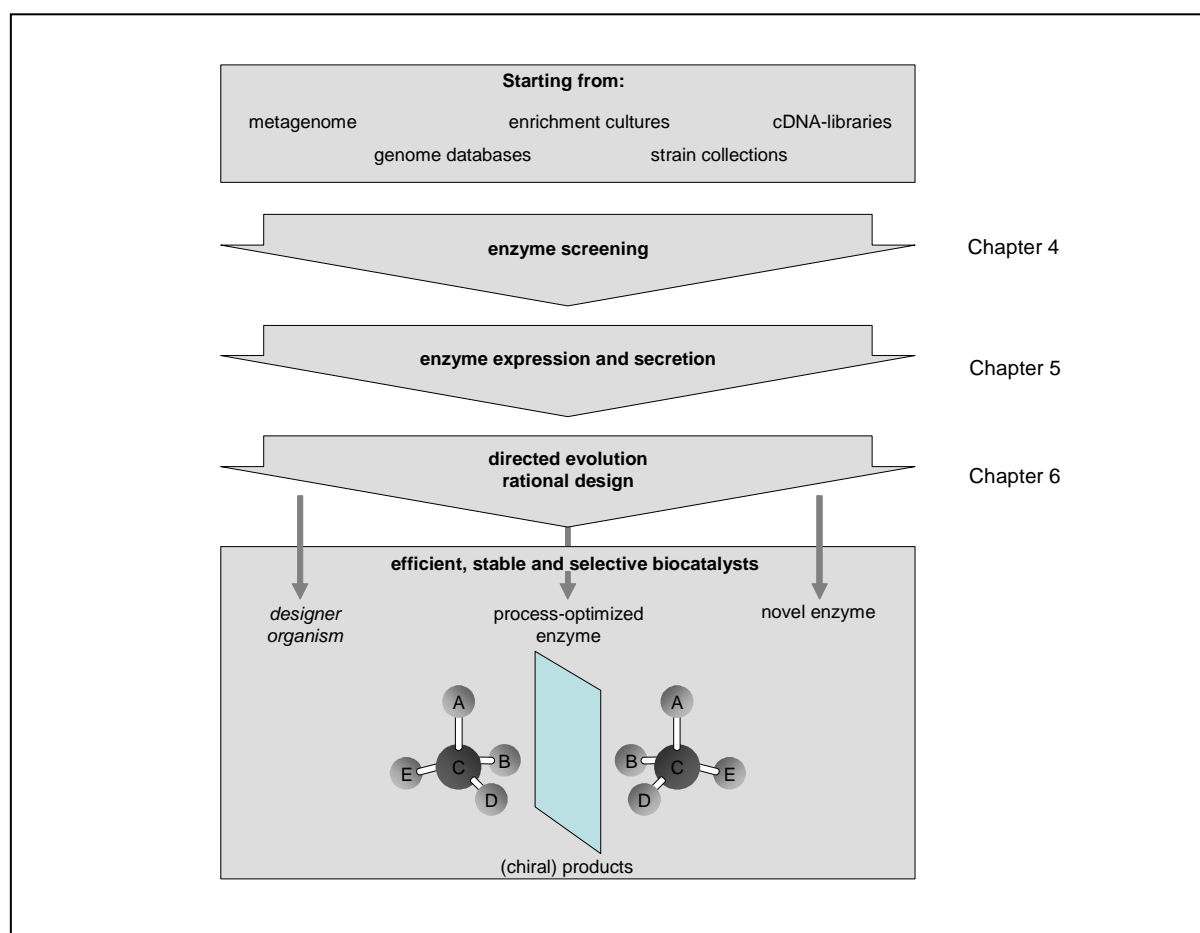

Fig. 3.1: Graphic overview of topics investigated within this work. The overall objective constitutes the development of strategies and novel methods to provide competitive biocatalytic processes.

Chapter 4:

You get what you screen for.

Biocatalyst identification by high-throughput screening and selection systems

Andexer, J., Guterl, J.-K., Pohl, M. and **Eggert, T.** (2006) A High-throughput screening assay for hydroxynitrile lyase activity. *Chem. Commun.* **40**: 4201-4203.

Bustos-Jaimes, I., **Eggert, T.**, Bogo, E., Puls, M., Hummel, W. and Jaeger, K.-E. (2006) New enzymatic method for high-throughput screening for enantiomeric excess of chiral alcohols and its application in molecular evolution. *ChemBioChem submitted for publication.*

Henning, H., Leggewie, C., Pohl, M., Müller, M., **Eggert, T.** and Jaeger, K.-E. (2006) Identification of novel biocatalysts displaying benzoylformate-decarboxylase activity by selection. *Appl. Environmental. Microbiol.* **72**: 7510-7517.

Wendorff, M., **Eggert,T.**, Pohl, M., Dresen, C., Müller, M. and Jaeger, K.-E. (2006) Directed evolution to increase the substrate range of benzoylformate decarboxylase from *Pseudomonas putida*. In: Asymmetric synthesis with chemical and biological methods. Enders, D. and Jaeger, K.-E. (eds.) Wiley-VCH, Weinheim; 298-311.

Jaeger, K.-E. and **Eggert, T.** (2004) Enantioselective biocatalysis optimized by directed evolution. *Curr. Opin. Biotechnol.* **15**: 305-313.

4.1 Introduction

Enzymes are superior (bio)catalysts enabling the production of a wide variety of different products in the field of fine chemicals [5,6]. Therefore, the demand for appropriate enzymes and their corresponding genes increases rapidly. Nowadays, in principle two different strategies to identify novel enzyme genes are state-of-the-art technology. On the one hand the **sequence-based screening** is used in public or commercial DNA-databases such as the National Center for Biotechnology Information (NCBI, www.ncbi.nlm.nih.gov) or the ERGO[TM] bioinformatics suite by Integrated Genomics Inc. (http://ergo.integratedgenomics.com/ERGO/) [7]. These databases constitute a steadily expanding source of gene sequences easily screened by computer tools like the BLAST-search algorithm [8,9]. However, there is one major drawback of this strategy: it is only possible to identify genes homologous to already known sequences; therefore, it is impossible to identify completely unknown biocatalysts having novel structural backbones by this so-called database-mining approach.

On the other hand **activity-based screening** in culture collections, genome- or cDNA-libraries, and, by far the most extensive gene source, the metagenome is possible. Here, huge libraries of wild type strains or recombinant expression strains are screened for desired enzyme activities. Finding the desired clone is not trivial; therefore, powerful <u>**high-throughput screening (HTS)**</u> or **selection systems** must be available to simultaneously assay the biocatalytic activity of 10^4 to 10^7 individual variants. Genetic selection is by far the most elegant and powerful way to identify the "one in a million". However, a microbial system

must be established in which the catalytic activity provides a growth advantage. Often the substrate of interest is provided as sole carbon or nitrogen source, whereby hydrolysis or modification of the compound enables the cells to grow. Other selection systems are based on *in vivo* or *in vitro* display technologies: the most popular one being phage display originally developed by Smith [10]. Here, the members of a variant enzyme library are displayed on the surface of the filamentous phage fd as a fusion to the N-terminus of the minor coat protein, also referred to as the gene-3-protein (g3p), thereby physically linking the phenotype and the genotype of the biocatalyst. Enzymes showing desired binding properties can be selected from a pool of randomly mutagenized variants. Successful examples of phage display selection have been reported to identify enzyme variants with improved biophysical properties and / or enhanced catalytic activities [11-14]. Furthermore, promising preliminary results show covalent and selective binding of phage-bound lipase to a chiral phosphonate inhibitor [15,16]. The selection of enzyme variants showing improved enantioselectivities using these chiral suicide inhibitors have been shown with respect to enantioselective hydrolysis of 1,2-*o*-isopropylidene-*sn*-glycerol (IPG) esters catalyzed by *B. subtilis* lipase A [17]; however, the general applicability to identify enantioselective enzymes remains to be demonstrated. In addition to phage display, powerful bacterial surface display systems are available [18] which can be screened in ultra high-throughput by using fluorescence-activated cell sorting (FACS) [19].

Unfortunately, in the majority of cases growth selection or surface display is not practical to identify a particular enzyme, making clone separation and individual assaying necessary. This can be performed in microtiter plates using HTS assays (Fig. 4.1A) or on agar plates using indicator media, like tributyrin for detection of esterolytic and lipolytic activity (Fig. 4.1B). The screening capacity of these systems is enlarged by using automation technology and standard procedures prepared by pipetting and colony transfer robots (Fig. 4.1C). Depending on the enzyme property a wide range of spectrophotometric assays using chromogenic / fluorogenic substrates or substrate analogues have been developed [20].

These assays are mainly based on three different strategies: (i) chromogenic or fluorogenic substrates, (ii) staining of the product or (iii) the use of indicator dyes. An overview of frequently used spectrophotometric and fluorometric assays is given in Table 4.1. For a comprehensive summary of the state-of-the-art high-throughput screening technology recent review articles [20-22] and books on assay development [23,24] describe these methods in detail. In addition to the spectrophotometric and fluorometric assays, sophisticated HTS-systems have been developed mainly in the group of Prof. Reetz (Max-Planck Institut für Kohlenforschung, Mülheim an der Ruhr, Germany) for the identification of enantioselective biocatalysts [25-27].
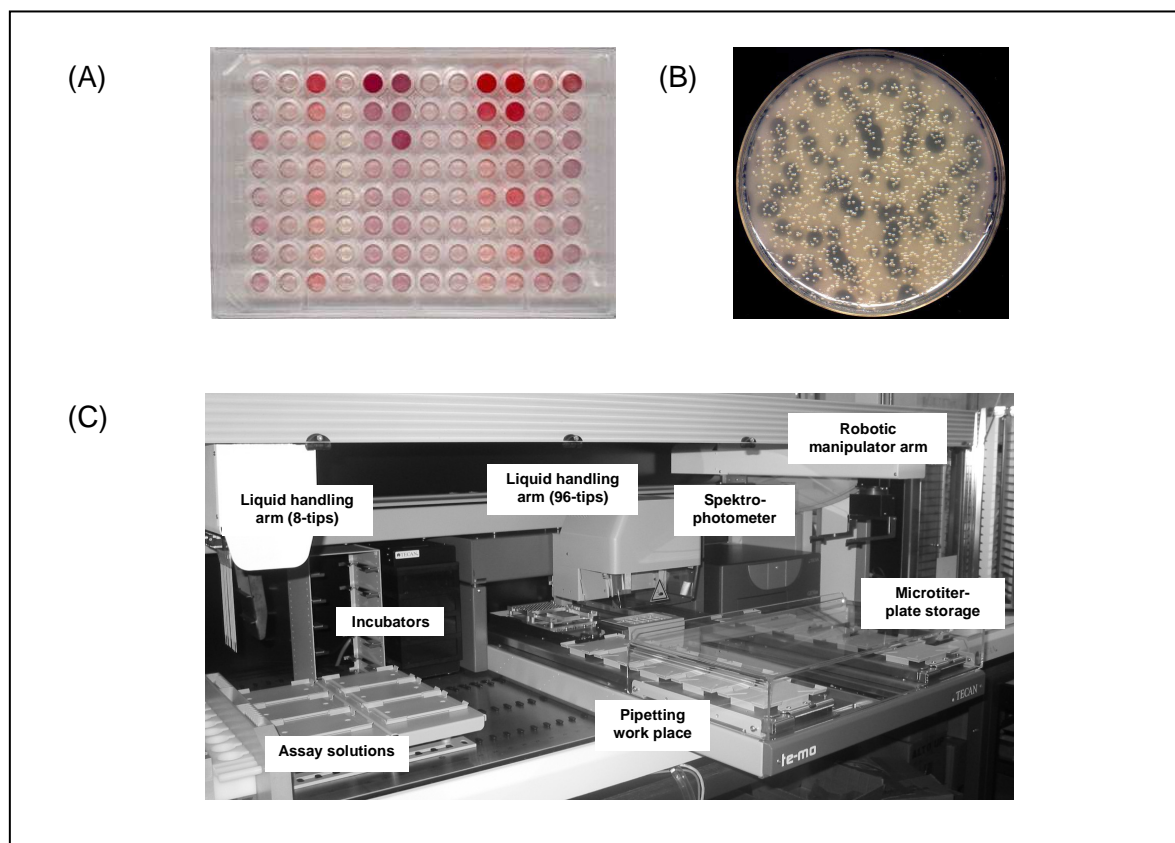
Fig. 4.1: High-throughput screening systems. (A) Spectrophotometric screening on carboligation activity and **(B)** agar-plate screening to identify esterolytic or lipolytic biocatalysts. Activity towards the substrate is indicated by red color formation or clear halos surrounding the bacterial colonies, respectively. **(C)** The screening capacity of these systems is enlarged by using automation technology and standard procedures prepared by pipeting or colony transfer robots.

Although during the last years a vast number of novel high-throughput screening assays have been developed, in the context of biocatalyst improvement, the screening is still considered as the bottleneck in enzyme identification and improvement. In the future, miniaturization and automation of variant library screening must be continued.

Novel screening and selection systems have been developed as part of this habilitation thesis. As described in chapter 4.2 an easy to perform spectrophotometric high-throughput screening assay for chiral alcohols was developed, as demonstrated by detecting enantioselective lipase catalyzed ester hydrolysis. Another spectrophotometric high-throughput assay has been established to identify novel or improved hydroxynitrile lyases (chapter 4.3), a class of highly versatile biocatalysts able to synthesize chiral cyanohydrins. Finally, in chapter 4.4 a spectrophotometric assay and a growth selection system based on a *Pseudomonas putida* strain are described in order to identify improved or novel C-C-bond forming enzymes like benzoylformate decarboxylases and benzaldehyde lyases.

Table 4.1: Spectrophotometric and fluorometric assays for high-throughput screening (HTS) of biocatalyst libraries.

Method[a] / Coloring substance	Enzyme	Reference
1) chromogenic or fluorogenic substrates		
p-nitrophenol	lipase, esterase, protease, monooxygenase	28,29
Umbelliferone	lipase, esterase, protease, phosphatase, epoxidhydrolase,	30,31
	transaldolase, transketolase	32,33
4-(p-nitrobenzyl)pyridine (NBP)	epoxidhydrolase	34
9, 11, 13, 15-octadecatetraenoic acid ester (Parinaric acid ester)	lipase	35
resorufin ester (e.g. 1,2-o-dilauryl-rac-glycero-3-glutaric acid-resorufin ester)	lipase	36
2) product-staining or -conversion		
4-nitro-7-chloro-benzo-2-oxa-1,3-diazole (NBD-Cl)	amidase	37
4-hydrazino-7-nitro-2,1,3-benzoxadiazole (NBD-H)	lipase	38
o-phthaldialdehyde-2-mercaptoethanol	nitrilase	39
NAD(P)H accumulation by alcohol dehydrogenase activity	lipase, esterase	40
3) product detection using indicator dyes		
pH-indicators (e.g. bromothymol blue, phenol red)	lipase, esterase, amidase, haloalkane dehalogenase	41-43
2,3,5-triphenyltetrazolium chloride (Tetrazolium red)	pyruvate decarboxylase, benzoylformate decarboxylase	44,45
Fuchsin	epoxidhydrolase	46
6-methoxy-N-(3-sulfopropyl)quinolinium (SPQ)	dehalogenase	47

[a]This list of methods gives an overview of widely used screening assays but it is not exhaustive. Further information can be found in recent reviews [20,22,48].

4.2 High-throughput screening of chiral alcohols

Chiral alcohols constitute valuable building blocks in the production of fine chemicals used in pharmaceutical and agricultural industries as well as in the food and lifestyle market as potent flavours and fragrances [49]. In order to discover new enantioselective catalysts, huge libraries of various chemical- or biocatalysts must be screened. Therefore, during the last decade many research groups have developed highly sophisticated screening technologies to identify enantioselective catalysts as summarized in a recent review article by Reetz [27]. However, some of these assays are highly specific for one target molecule or necessitate

expensive screening facilities; thereby, hampering a broader applicability. The development of smart and inexpensive assays useful in screening enantioselective catalysts for the production of chiral building blocks still remains a challenging task for the future.

Spectrophotometric screening assays based on coupled enzyme reactions have several advantages, like high sensitivities because of the signal amplification effect [40]. We established a sensitive assay based on the enantioselective oxidation of alcohols by using two different alcohol dehydrogenases (ADHs) namely the (R)-specific ADH from *Lactobacillus kefir* (LKADH) and the (S)-specific ADH from *Rhodococcus erythropolis* (READH), whose enantioselectivities and value as catalysts have been previously reported [50,51]. The oxidation of either (R)-**1** or (S)-**1** produces NAD(P)H, which is then again oxidized to NAD(P) by diaphorase from *Clostridium kluyveri* with the concomitant reduction of 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl-2H-tetrazolium (INT) **3** to its corresponding formazan red-violet dye **4** (Fig. 4.2). The formation of this dye can be easily followed at 492 nm.
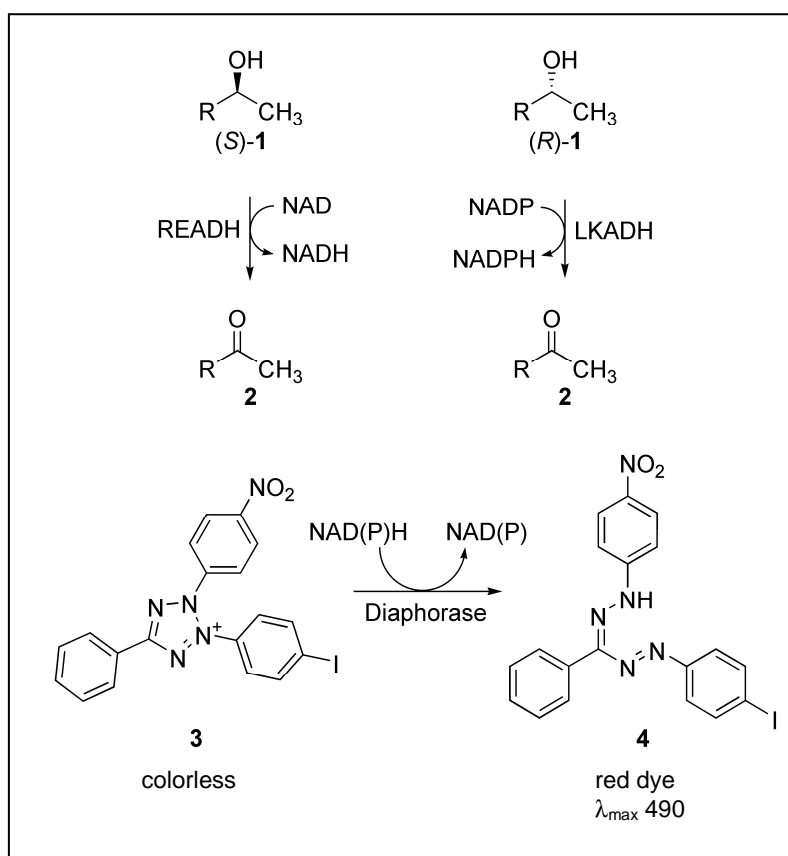
Fig. 4.2: Enantioselective reactions catalyzed by READH and LKADH and its coupling to diaphorase redox system. Oxidation of either alcohol (S)-**1** or (R)-**1** correspondingly produces a molecule of coenzyme, NADH or NADPH, which in turn are oxidized by diaphorase through the reduction of **3** to produce the corresponding red formazan dye **4**.

The addition of the coupling enzyme to measure the reduced coenzymes, either NADH or NADPH, increased the sensitivity of the previously reported enzymatic methods [40]. This is because the extinction coefficient (ε) at λ_{max} 490 nm for compound (**4**) is 2.4 fold higher than

ε at 340 nm for NAD(P)H. Moreover, this wavelength (490 nm) does not interfere with aromatic compounds and proteins in the measurement of the reduced coenzymes. Under the conditions reported here we were able to reliably measure alcohol concentrations as low as 25 µM, keeping the error close to 10%.

In practice, we have used this assay to screen site-saturated mutant libraries of BSLA [52,53]. During the screening procedure, we were able to detect BSLA mutants which showed enantioselectivities different from the wild type enzyme. Additionally, the same assay was used to screen a mutant library produced by error-prone PCR of a BSLA-variant which was previously rationally modified by inserting an artificial lid-motif [54] (see also chapter 6.4). The assay was scaled down to 60 µL per sample enabling the screening in 384-well microtiter plates automated by a TECAN workstation 200 pipetting robot (Fig. 4.3).
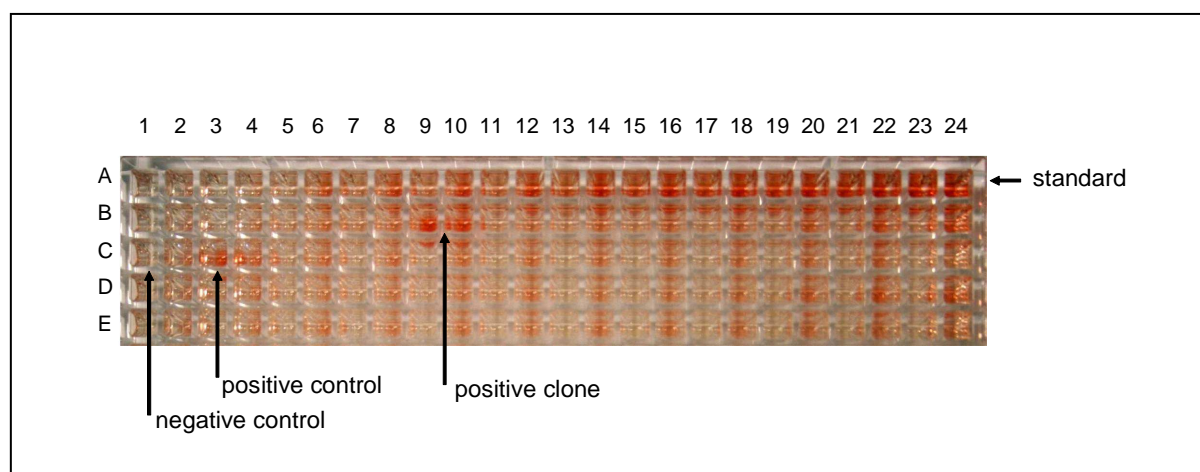
Fig. 4.3: Screening for activity and enantioselectivity towards the lipase catalyzed ester hydolysis producing chiral alcohols. The image shows the color development in a 384-well microtiter plate after 5 minutes of incubation at 30℃. The first row contains, from left to right, increasing concentrations of the used substrate (here 1-phenyletyl acetate). In the second row a sample with high substrate concentration is visible; this sample comes from the very active mutant MPB7. On the left side of the third row, a sample from the wild type BSLA is used as the positive control.

4.3 Spectrophotometric screening assay for hydroxynitrile lyase activity

Hydroxynitrile lyases (HNLs) form a heterogeneous group of enzymes which is characterized by the ability to catalyze enantioselectively the cleavage of cyanohydrins into HCN and the corresponding carbonyl compound (cyanogenesis). The natural function of HNLs in plants is the defense against herbivores by releasing toxic amounts of HCN [55]. Since the reverse reaction is also catalyzed enantioselectvely by HNLs, these enzymes are valuable catalysts for the synthesis of cyanohydrins, which are versatile chiral building blocks in pharmaceutical and agrochemical industries [56].

Until now, only five HNLs are biochemically characterized in detail concerning substrate specifity and enantioselectivity. Two of them, the HNLs from *Prunus amygdalus* (bitter almond) and *Linum usitatissimum* (flax) show (*R*)-selectivity [5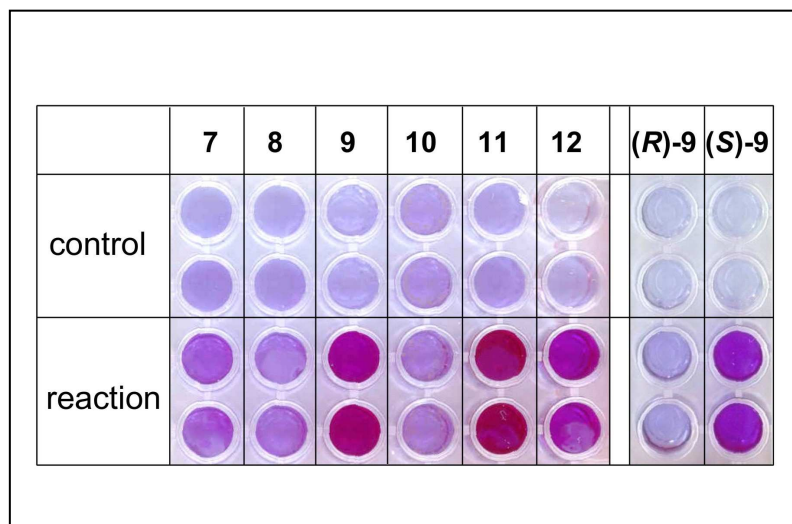7,58], whereas the other HNLs from *Sorghum bicolor* (millet), *Hevea brasiliensis* (para rubber tree) and *Manihot esculenta* (cassava) show (*S*)-selectivity towards different cyanohydrins [59,60].

In order to identify novel HNLs and / or to optimize recombinant HNLs by directed evolution, a high-throughput screening assay useful to detect HNL-activity theoretically towards any cyanohydrin substrate was developed. The assay consists of two parts, as outlined in figure 4.4.

Fig. 4.4: Schematic overview of the assay system. **(A)** Biotransformation step. The substrate cyanohydrin **5** is enzymatically converted to a carbonyl compound **6** and HCN. Six different cyanohydrins (**7**: acetaldehyde cyanohydrin, **8**: propionaldehyde cyanohydrin, **9**: benzaldehyde cyanohydrin, **10**: 3-phenoxybenzaldehyde cyanohydrin, **11**: acetone cyanohydrin, **12**: cyclohexanone cyanohydrin) were tested with MeHNL. **(B)** Cyanide-determination step. Cyanide anions are oxidized by N-chlorosuccinimide **13** (stabilized with succinimide **14**) to cyanide cations, which react with isonicotinic acid **15** forming a dialdehyde **16**, which is coupled to two molecules barbituric acid **17** to form the dye **18** measured spectrophotometrically at 600 nm.

In the first step the enzymatic cyanohydrin cleavage liberates HCN, which is subsequently detected spectrophotometrically at 600 nm wavelength in the second step (Fig. 4.5). The assay is useful for both, preparing enzyme fingerprints, and screening large variant libraries generated in metagenome or directed evolution approaches. The assay is highly sensitive; at least 5 ng of purified *Manihot esculenta* HNL representing 1 mU of enzyme activity was reliably detectable in the assay. Furthermore, the assay can be used to determine the enantioselectivity of HNLs or variants in HNL-libraries towards chiral substrates which are available in as single enantiomers.

Fig. 4.5: Spectrophotometric determination of hydroxynitrile lyase activity in 96-well microtiter plates with different substrates **7 – 12** (see figure 4.4). Control: autolysis of the respective cyanohydrin (without enzyme). For reactions 10 µL of *E. coli* crude cell extracts containing over-expressed MeHNL were used. Faint blue to purple color represents increasing amount of cyanide. The application of enantiomerically pure substrates can be used to estimate the enantioselectivity of the biocatalyst as demonstrated in the case of enantiomerically pure (*R*)- and (*S*)-benzaldehyde cyanohydrin **9**.

4.4 Screening and selection systems for C-C-bond forming enzymes

A third group of biocatalysts catalyzing important reactions used to obtain chiral organic building blocks belong to the family of thiamine diphosphate (ThDP)-dependent enzymes. The interesting feature of these enzymes is their ability to catalyze the cleavage and the formation of carbon-carbon (C-C) bonds in a stereoselective way. In comparison to the biocatalytic strategy, the state-of-the-art organic chemistry methods in this field is often hampered by the necessity of using time and cost extensive protection and deprotection cycles. Until now several ThDP-dependent enzymes have been identified and characterized in detail converting a number of donor and acceptor substrates into a huge variety of different chiral products (Fig. 4.6) as summarized in recent review articles [61-63]. However, there is still a growing demand of these kinds of enzymes, because of missing alternative strategies by organic chemistry.
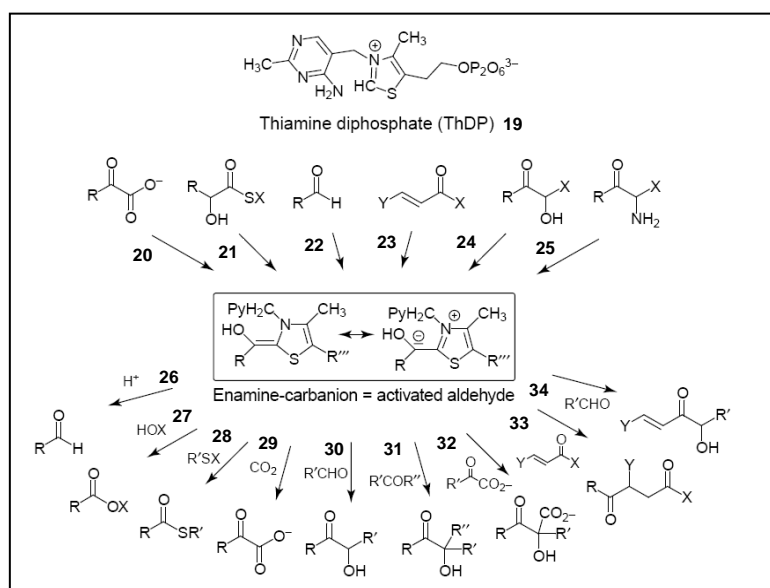

Fig. 4.6: Overview of different donor and acceptor molecules converted by ThDP-dependent enzymes into various chiral products. All kinds of reactions lead to the activated intermediate (middle) by addition of donor molecules (**20-25**) to the C2 atom of the ThDP-cofactor (**19**), which are subsequently transferred to one of the acceptor molecules (**26-34**) [62].

Screening for improved benzoylformate decarboxylases

One representative member of the ThDP-dependent enzymes is the class of benzoylformate decarboxylases (BFD; EC 4.1.1.7) naturally catalyzing the non-oxidative decarboxylation of benzoylformate to benzaldehyde and carbon dioxide. This reaction takes place in the mandelate pathway, in which mandelate is degraded to benzoic acid which is further metabolized in the β-ketoadipate pathway to succiny-CoA and acetyl-CoA, both substrates for the citric acid cycle.
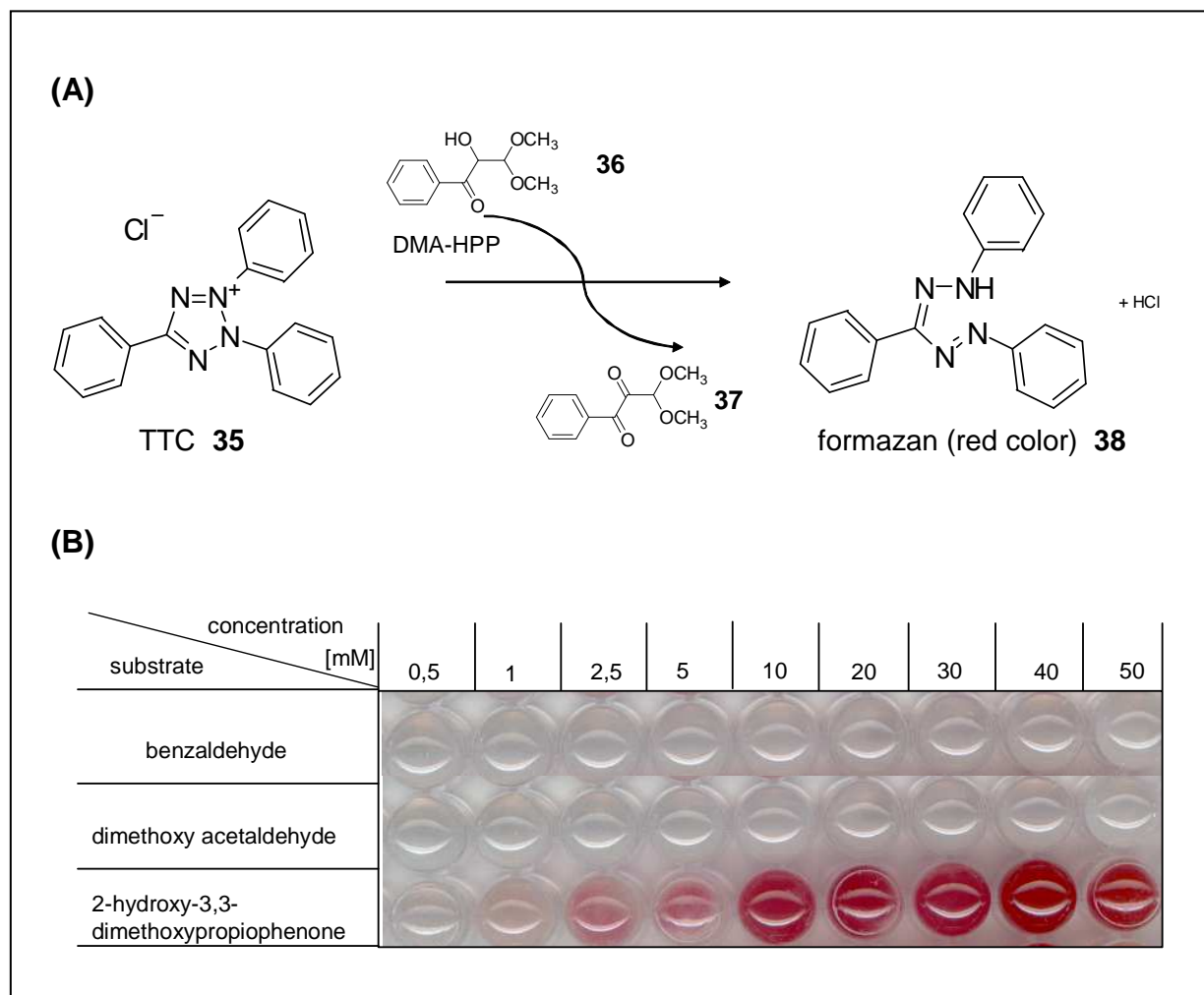

Fig. 4.7: Colorimetric assay for carboligase activity. (A) The screening product 2-hydroxy-3,3-dimethoxypropiophenone (DMA-HPP) **36** reduces 2,3,5-triphenyltetrazolium chloride (TTC) **35** to the respective formazane **38** which has an intense red color. **(B)** Formation of the red formazane dye **38** can be observed only in the presence of the product **36**, both substrates dimethoxy acetaldehyd and benzaldehyde do not cause any change in color.

The BFD from *Pseudomonas putida* has been characterized in great detail with respect to its biochemical properties [64,65] and 3D structure [66,67]. As an interesting side reaction, these enzymes catalyze the carboligation of aldehydes to form chiral 2-hydroxy ketones (Fig. 4.6; donor **3**, acceptor **11**) [64,68]. The physiological role of this additional enzymatic activity is still

unknown; nevertheless, by this the carboligation of benzaldehyde, benzaldehyde derivatives, and acetaldehyde is possible to synthesize 2-hydroxy propiophenone derivatives with an enantiomeric excess (*ee*) of 82-94% for 2-HPP, depending on the benzaldehyde concentration and on the reaction temperature [64,65]. With regard to 2-HPP formation the related ThDP-dependent enzyme benzaldehyde lyase (BAL) catalyzes the same carboligation reaction but with reverse stereoselectivity [69,70]. By using these two enantiocomplementary enzymes, many 2-HPP analogues can be synthesized in both enantiomeric forms. The formation of (*R*)-benzoin has also been reported as a ligation product of this reaction which is catalyzed by BFD albeit with negligible activity [64].

In order to screen for improved BFDs having a broader substrate spectrum a modified high-throughput screening assay was established based on the results already published by Breuer et al. [44]. This color assay is based on the reduction of the 2,3,5-triphenyltetrazolium chloride (TTC) to the corresponding formazane which has an intense red color (Fig. 4.7). This assay was proved to be useful by screening an epPCR library of about 8.000 clones identifying one BFD-variant with an optimized acceptor aldehyde spectrum.

Growth selection to identify novel benzoylformate decarboxylases

A novel growth selection system was developed to identify biocatalysts exhibiting BFD-activity using the *Pseudomonas putida* strain KT2440 [71,72] which is able to grow on benzaldehyde as the sole carbon source. These bacteria presumably metabolize benzaldehyde *via* the β-ketoadipate-pathway; nevertheless, the strain KT2440 was unable to grow in selective media with benzoylformate as sole carbon source. Other *P. putida* strains like ATCC12633 [73,74] are able to grow both on mandelate and benzoylformate as sole carbon source, because of the presence of benzoylformate decarboxylase activity.
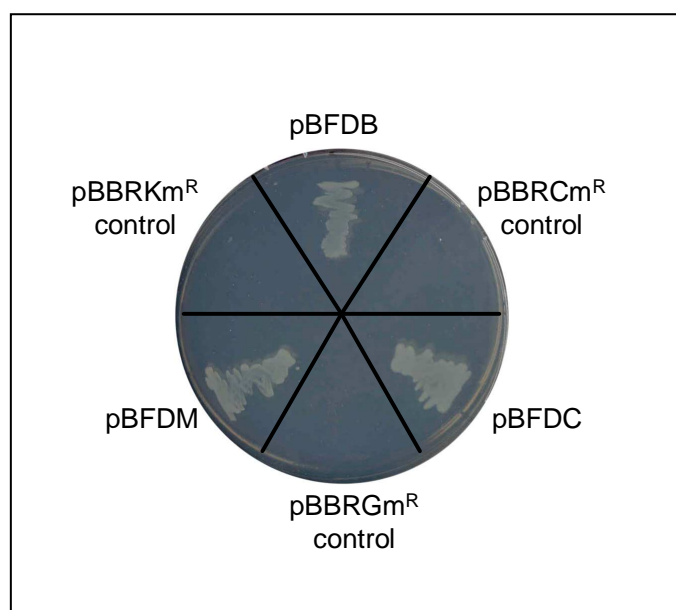

Fig. 4.8: Growth of recombinant *P. putida* KT2440 displaying benzoylformate decarboxylase activity on an agar plate containing benzoylformate selective medium. *P. putida* KT2440 possessing the plasmids pBFDB (coding for BfdB from the genomic DNA-library of *P. putida* ATCC12633), pBFDC (coding for BfdC from the genomic DNA-library of *P. putida* ATCC12633Δ*mdlC*), and pBFDM (coding for BfdM from a metagenomic library) expressed active BFD-enzymes enabling growth by converting benzoylformate. *P. putida* strains harboring the corresponding plasmids pBBRCm[R], pBBRGm[R] and pBBRKm[R] served as negative controls.

When *P. putida* KT2440 was grown on benzoylformate-containing selective media the growth deficiency could be restored by expression *in trans* of genes encoding benzoylformate decarboxylases (Fig. 4.8). The functionality of the selection system was proven by the isolation of three novel benzoylformate decarboxylases, one of them originating from a metagenomic library. These novel enzymes were active towards the substrate benzoylformate and they shared only low sequence similarity to presently known BFDs.

4.5 Discussion

It is obvious from the literature that many sophisticated high-throughput screening (HTS) assays exist today (for details see Table 4.1). However, there is still a bottleneck in the identification of novel biocatalysts, because many HTS-systems are highly specific for one substrate or for a substrate surrogate. Other screenings, mainly in the field of assaying the enantiomeric purity of chiral compounds need expensive equipment like HPLC, ESI-MS or NMR [27,75,76] that are not general available in molecular biology laboratories. Therefore, novel screening assays which are easy to perform and applicable towards various substrates are still needed in the field of biocatalyst identification and engineering.
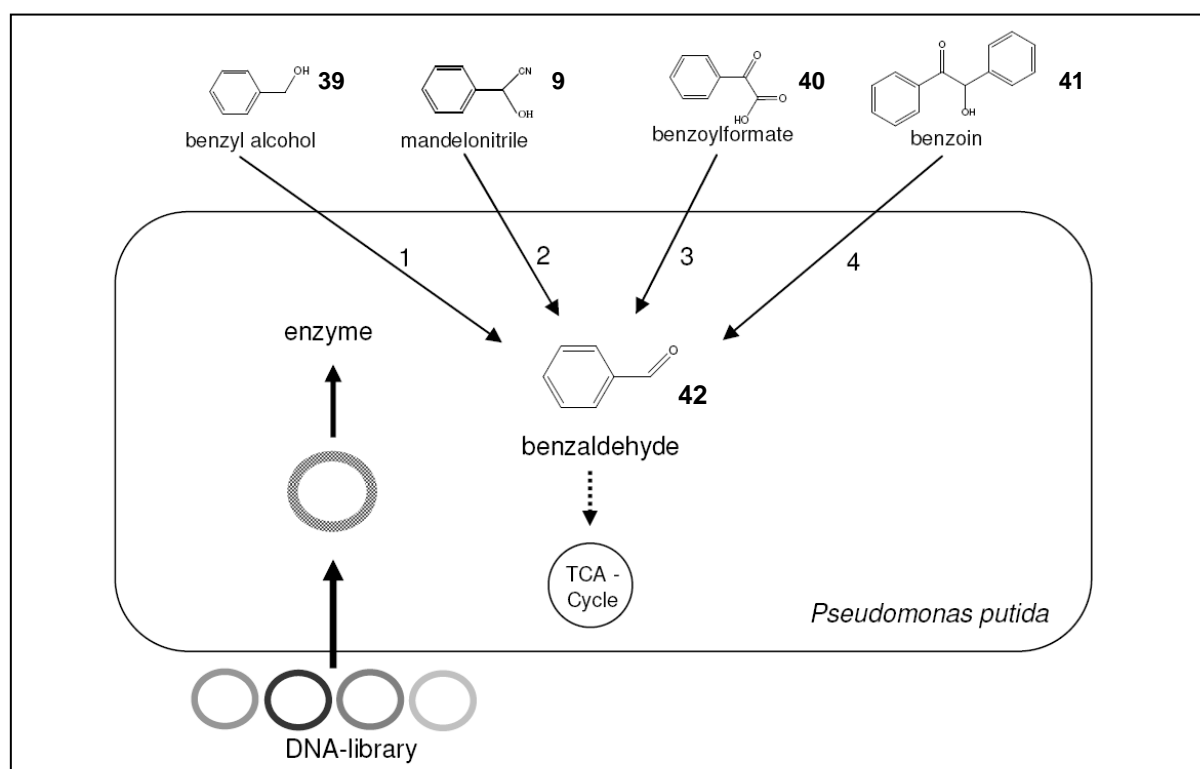

Fig. 4.9: Growth selection on benzaldehyde to identify industrially important biocatalysts from genomic- or metagenomic DNA-libraries. Accessible enzymes may include 1: alcohol dehydrogenase (ADH), 2: hydroxynitrile lyase (HNL), 3: benzoylformate decarboxylase (BFD), and 4: benzaldehyde lyase (BAL).

Here, three novel systems for HTS and selection have been established. In addition, a well established screening system for benzoylformate decarboxylases has been adapted to a broader substrate spectrum. All systems show high sensitivity without loss of reproducibility. The enzyme-coupled assay (chapter 4.2) detects µM amounts of chiral alcohols as a reaction product of enantioselective ester hydrolysis. Nevertheless, this assay should also be useful in the detection of redox reactions leading to chiral alcohols as well.

In order to screen larger libraries selection systems are essential. The selection system described in chapter 4.4 (Fig. 4.8) proved to be useful in the identification of benzoylformate decarboxylase should in principle be suitable in selecting other enzyme activities based on the carbon source in the selective medium. In ongoing studies new areas of application will be investigated as outlined in figure 4.9.

Chapter 5:

Learning is like rowing against the current.
As soon as you stop, you drift back.

Benjamin Britten

Biocatalyst production in *Bacillus subtilis*

Brockmeier, U., Caspers, M., Freudl, R., Jockwer, A., Noll, T. and **Eggert, T.** (2006) General strategy to improve heterologous protein secretion in the Gram-positive bacterium *Bacillus subtilis* by signal sequence screening and directed evolution. *J. Mol. Biol.* **362**: 393-402.

Brockmeier, U., Wendorff, M. and **Eggert, T.** (2006) Versatile expression and secretion vectors for *Bacillus subtilis. Curr. Microbiol.* **52**: 143-148.

Detry, J., Rosenbaum, T., Lütz, S., Hahn, D., Jaeger, K.-E., Müller, M. and **Eggert, T.** (2006) Biocatalytic production of enantiopure cyclohexane-1,2-diol on a preparative scale using extracellular lipases from *Bacillus subtilis. Appl. Microbiol. Biotechnol.* **72**: 1107-1116.

5.1 Introduction

The Gram-positive bacterium *Bacillus subtilis* has developed into one of the most important hosts for the production and secretion of homologous and heterologous proteins. Major adventages of *B. subtilis* includes: (i) the strain is classified as a GRAS – generally recognized as safe – organism free of any endotoxin. (ii) Genetic manipulation tools, like plasmids, transformation protocols, construction of deletion and/or insertion mutants as well as the preparation of stock cultures are well established for *B. subtilis.* (iii) Furthermore, the strain is well known with respect to its genome [77-79], proteome [80], transcriptome [81] and nowadays also to its secretome [82]. (iv) Furthermore, bacilli are widely used in large scale fermentation processes, growing fast and unproblematic on inexpensive media up to high cell-densities [83]. (v) In contrast to Gram-negative bacteria, *B. subtilis* and related bacilli contain only one membrane – the cytoplasmic membrane – to be passed by the exported proteins. *B. subtilis* offers efficient secretion machineries – the Sec-apparatus being the most important – guiding the expressed protein directly into the culture supernatant (Fig. 5.1); thereby, bypassing the time-consuming cell disruption which makes the downstream processing of the protein much easier. In addition, in case of efficient secretion the formation of inclusion bodies in the cytoplasm is reduced, leading to higher amounts of properly folded and active enzymes. The protein production process can be run continuously, performing the downstream processing in parallel.

On the basis of these advantages *B. subtilis, B. licheniformis* and other related bacilli have developed to important expression strains frequently used in industrial fermentations as well as fundamental research. For homologous proteins like proteases, amylases, pectate lyases, and glucanases the production and secretion in *B. subtilis* can be highly efficient [83]; however, the overexpression and secretion of heterologous proteins show major drawbacks due to different bottlenecks in the secretory pathway [84,85].
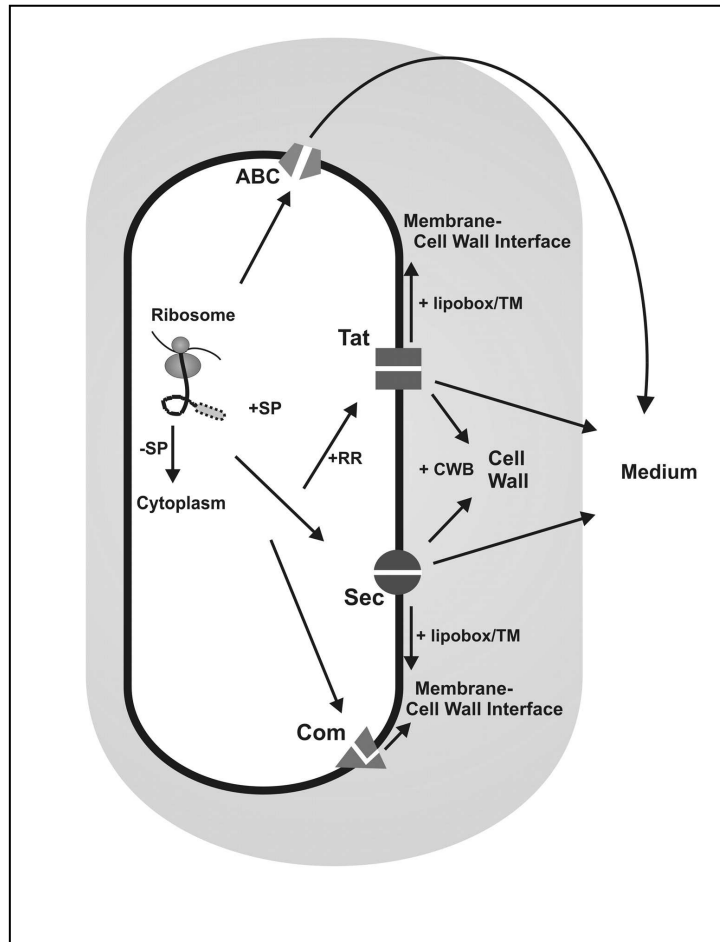
Fig. 5.1: Protein export pathways in _B. subtilis_ according to _Tjalsma et al._[82] . The major secretion pathways are classified as ABC-transporter (ABC), Sec-dependent translocation (Sec), Twin-arginine translocation pathway (Tat), and the pseudopilin export (Com). Abbreviations: +SP, containing signal peptide; –SP, without an N-terminal signal peptide; TM, transmembrane segment; +lipobox, lipid modification; +CWB, cell wall-binding repeats; +RR, signal peptide containing the Twin-arginine motif.

Protein secretion using N-terminal signal peptides

Selective protein targeting across cell-membranes is one of the most important processes in all kinds of living cells, because correct protein function is closely associated with correct positioning within the cell. The involvement of so-called signal peptides at the N-terminus of the nascent polypeptide chain as a "label" for targeting was first proposed in 1971 as signal hypothesis by Blobel and Sabatini for the protein transport across the endoplasmic reticulum (ER) membrane [86]. Later it became obvious that translocation processes across the bacterial plasma membrane share this common feature [87,88]. Since that time continuous progress has been made in revealing the mechanism of protein transport across the bacterial cytoplasmic membrane. The membrane-located secretion machinery as well as many accessory proteins located either in the cytoplasm, periplasm or membrane-associated have been identified and characterized in more detail [89-92].

In addition to the components of the secretion apparatus, the signal peptides (SPs) of the secreted proteins play an important role in the efficiency of translocation across the membrane. Therefore, SPs were investigated intensively with respect to their amino acid composition and to their role in membrane translocation of exported proteins [93-97]. Signal

peptides share some common characteristic features, conserved in different organisms. Most SPs are composed of three distinct regions: (i) the positively charged N-domain, (ii) the hydrophobic core region, the so-called H-domain, and (iii) the signal peptidase (SPase) recognition-site termed C-domain (Fig. 5.2). Based on these criteria, many signal peptide prediction tools have been developed [98] with SignalP being the most popular and user-friendly program [99]. However, no prediction program is able to estimate reliable secretion efficiencies of export targets. The prediction of secretion efficiency becomes even more complicated when heterologous proteins are fused to an SP instead to its natural secretion partner. As it became obvious that the signal peptide and secreted protein constitute a unique unit, where the N-terminus of the mature protein, the so-called signal peptide mature junction, plays an important role in the secretion efficiency, the prediction of "good" SPs for heterologous protein secretion became impossible because it also depends on the N-terminus of the mature part of the secreted protein. Therefore, the high yield secretory production of industrially interesting proteins in heterologous fermentation hosts is often limited by the secretion efficiency, because the "wrong" SP was fused to the secretion partner.
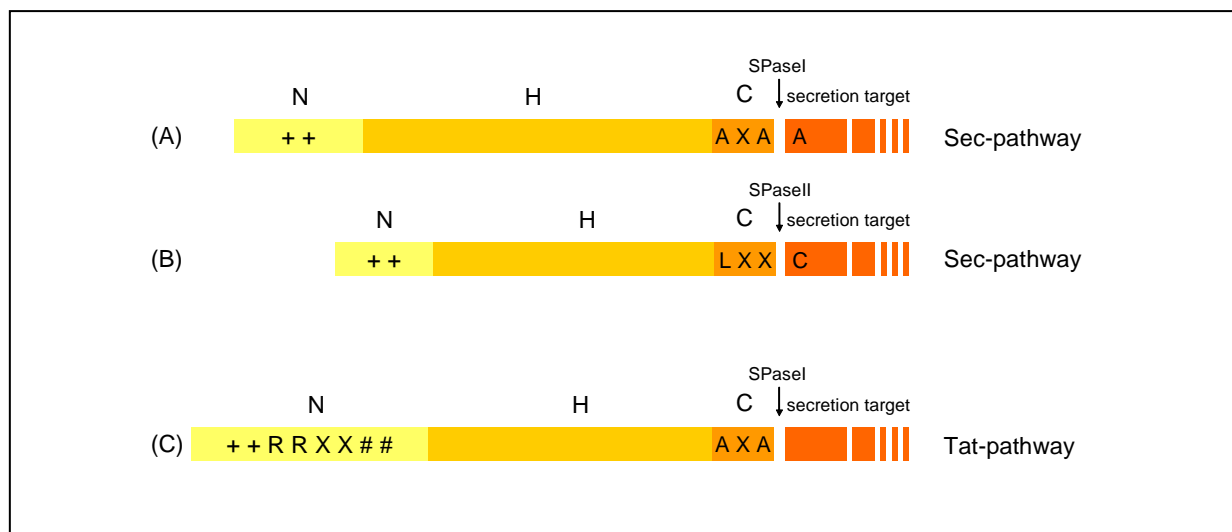

Fig. 5.2: Schematic illustration of typical N-terminal signal peptides (SPs). (A) The overall structure of typical Sec-type SPs, **(B)** Sec-dependent Lipoprotein SPs and **(C)** Twin-arginine SPs contain a positively charged N-region, a hydrophobic H-region and a signal peptidase (SPase) recognition site called C-region. Amino acids are given in the one-letter-code, + is a positively charged amino acid, x can be any amino acid and # is a hydrophobic residue.

In order to develop a general strategy to improve protein secretion efficiencies in *B. subtilis* we investigated the possibility to find the optimal SP out of a library of all naturally occurring SPs from the secretion host (chapter 5.2). In a second approach we have applied directed evolution techniques to optimize the amount of secreted target proteins (chapter 5.3).

5.2 Optimization of heterologous protein secretion by signal sequence screening

At the beginning of the secretion optimization project two prerequisites had to be fulfilled: (i) a stable and self replicating (i.e. extra chromosomal) expression plasmid with sufficient promoter sequence and multiple cloning site must be available, and (ii) a high-throughput screening system for quantification of secretion efficiencies in *B. subtilis* must be established.

Versatile expression and secretion vectors for Bacillus subtilis

Different kinds of secretion vectors have been published already for *B. subtilis* [100,101]; however, until now, the choice of available expression plasmids combining similar properties like *E. coli*-systems is still limited. Therefore, we constructed a series of multi-copy expression vectors named pBSMuL based on the *E. coli - B. subtilis* shuttle plasmid pMA5 with a pUB110 ori for replication in *B. subtilis* [102,103]. These plasmids offer several advantages with respect to biotechnological applications: downstream of strong constitutive promoter(s) an artificial DNA-fragment was inserted containing (i) a *Bacillus* ribosome binding site (rbs), (ii) a multiple cloning site (MCS) composed of 13 frequently used restriction sites, (iii) the possibility of in-frame fusion to a hexa-histidine-tag (6xHis) for convenient one-step purification by immobilized-metal affinity chromatography and (iv) the intracellular protein expression using the *Nde*I restriction site.

Two vector plasmids, pBSMuL1 and pBSMuL2, were constructed to clone any gene of interest as a translational fusion of the *B. subtilis* SP obtained from the extracellular lipase LipA. Furthermore, a third vector plasmid pBSMuL3 also allows the convenient exchange of the SP in front of the target gene. Using pBSMuL3 should allow to fuse any SP with any secretion target protein. The versatility of the pBSMuL vector-series with respect to replication stability, segregational stability and target expression was proved using the cutinase from *Fusarium solani pisi* as the model secretion target cloned into pBSMuL1 and pBsMuL2.

High-throughput screening of secretion efficiency

In our project lipolytic enzymes were used as model secretion targets. Therefore, as the secretion host the lipase deletion mutant *B. subtilis* TEB1030 has been chosen [104]. In cooperation with the group of Prof. Büchs (Rheinisch Westfälische Technische Hochschule (RWTH) Aachen, Germany) the cell growth in microtiter plates (96 deep well) was optimized with respect to growth medium compositions, culture volumes, growth temperature and shaking intensity in order to minimize growth fluctuations (Fig. 5.3A). After culturing the cells, the quantitative detection of the lipolytic activity in the culture supernatant was performed using a spectrophotometric assay as described by Winkler and Stuckmann [105]. As the

substrates p-nitrophenyl-palmitate (pNPP) or p-nitrophenyl-caproat (pNPC) were used in an automated high-throughput screening assay (Fig. 5.3B) that was realized in microtiter plates (96-well-plates) using the TECAN workstation Genesis 200 Freedom.
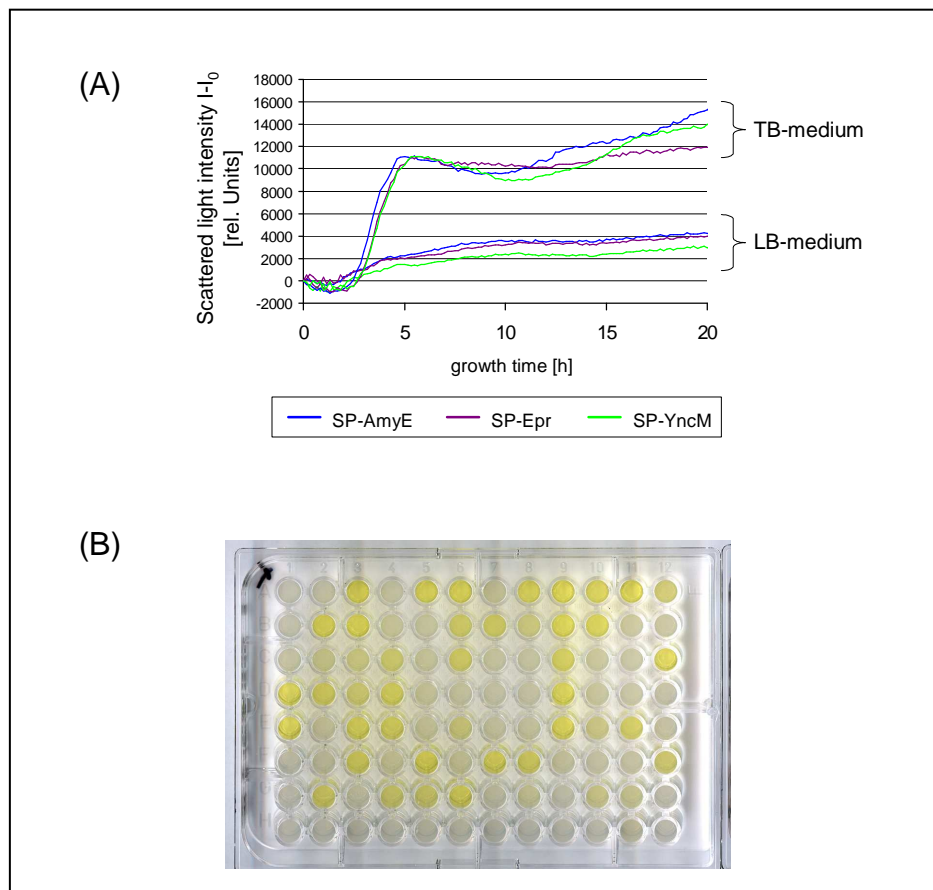

Fig. 5.3: Development of a microtiter plate based high-throughput screening assay for the detection of secretion efficiencies in *B. subtilis*. (A) In order to minimize growth fluctuations of the expression and secretion strain the medium composition, culture volumes, growth temperature and shaking intensity was optimized in cooperation with the group of Prof. Büchs (RWTH Aachen, Germany) using a Quantitative Micro-Reactor Cultivation (QMRC) system which can be analyzed online. *B. subtilis* strains expressing cutinase from *Fusarium solani pisi* fused to different SPs (SP-AmyE, SP-Epr and SP-YncM) leading to different secretion efficiencies, show comparable cell densities when optimized growth conditions were chosen. **(B)** The automated activity assay was spectrophotometrically analyzed indicating high secretion levels by increasing yellow colors.

High-throughput screening of all Sec-type signal sequences fused to cutinase

All Sec-dependent, non-lipoprotein signal peptides as identified by Tjalsma *et al* [82] were fused to the model secretion target cutinase from *Fusarium solani pisi* and expressed and secreted in *B. subtilis* TEB1030. 25 of the 173 SP-cutinase fusions could not be expressed in *B. subtilis*. Here, the plasmids constructed in *E. coli* could not be successfully transferred into the expression and secretion strain, indicating a lethal effect. However, the remaining 148 SP-cutinase fusions could be successfully transferred into the expression host. The screening of this signal peptide library revealed strong differences in lipolytic activity of the culture supernatants, ranging from no secreted cutinase up to 4.7 U/mL (Fig. 5.4), corresponding to 35 mg/L as verified by Western blotting experiments.
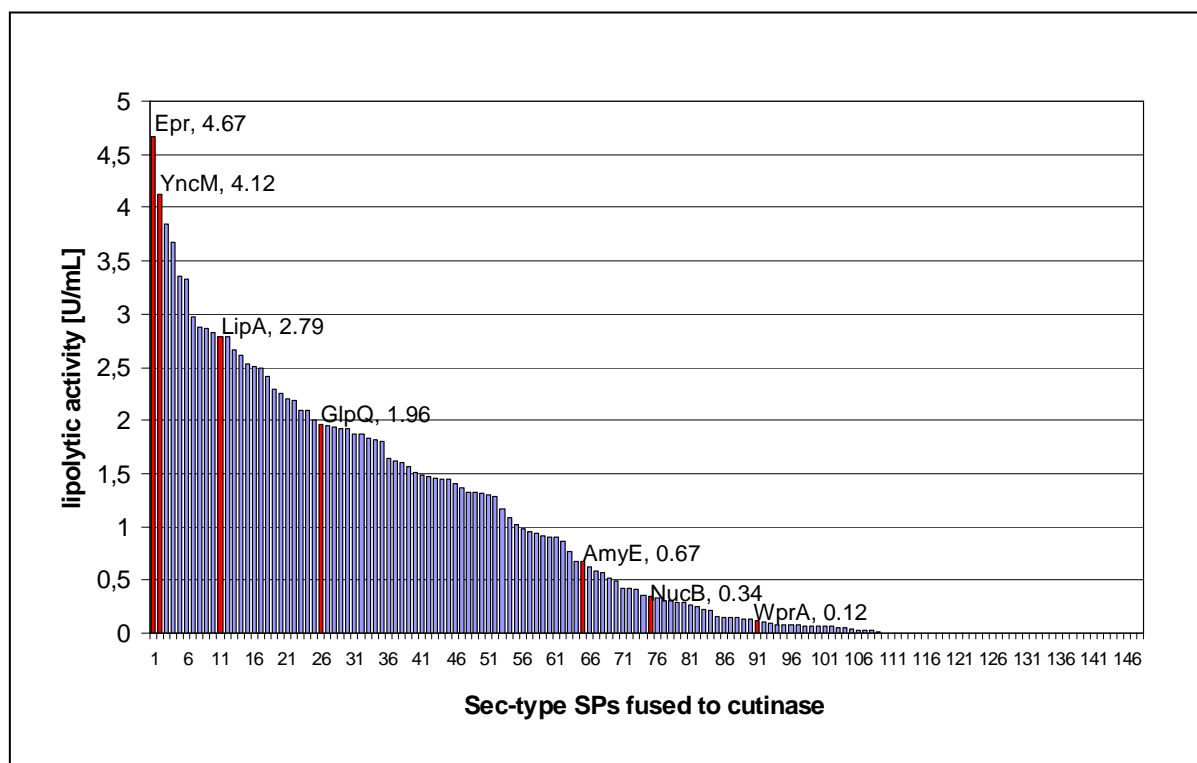
Fig. 5.4: Comparison of all screened signal peptides (SPs) used for export of heterologous cutinase in *B. subtilis*. The SPs of Epr, YncM, LipA, GlpQ, AmyE, NucB and WprA, which have been analyzed further by Western-blotting and pulse-chase experiments, are highlighted in red. The results represent data from 12 independent experiments.

The "SP-ranking" summarized in figure 5.4 contains some astonishing results, which were not expected. (i) Only a few SPs lead to high secretion efficiencies with a clear preference for Epr being the best with respect to cutinase secretion. (ii) Well known SPs secreting there natural target highly efficient into the culture supernatant like the proteases NucB and WprA[82] lead to improper secretion values when fused to cutinase. (iii) When charges, length and hydrophobicities of the SPs shown in figure 5.4 were analyzed in more detail, no clear pattern was obvious which would indicate any rule to predict the best SPs. (iv) If the varying amounts of cutinase in the supernatant for the different signal peptide cutinase fusion proteins is solely due to differences in the efficiencies of protein translocation across the cytoplasmic membrane and SP processing, those fusions with the highest amounts of protein in the supernatant should exhibit the fastest processing kinetics. However, this is clearly not the case as demonstrated by pulse-chase experiments. While the translocation and processing of the Epr fusion protein, which shows the highest lipolytic acitivity and the largest amount of cutinase protein in the supernatant, is indeed very fast, the second-"best" signal peptide with respect to the amounts of cutinase protein and activity in the supernatant, YncM, shows the slowest translocation and processing of all the proteins investigated in the pulse chase experiment. Furthermore, the AmyE fusion, which shows the lowest protein amounts

and lipolytic activity in the supernatant of the five proteins analyzed in the pulse-chase, is processed very rapidly.

Individual screening of signal peptides for every secretion target is necessary

After screening all Sec-type signal peptides with respect to their efficiency in cutinase secretion the question arose whether the SP-ranking also applies for other heterologous secretion target proteins. Therefore, another target protein, an intracellular esterase form metagenomic origin, was fused to the pool of SPs available in our lab. In this case the SP-ranking of figure 5.4 was not reproduced. The three most efficient SPs in the secretion of the metagenome esterase lead to poor (SP-YwmC and SP-YpjP) or no (SP-YojL) secretion of cutinase (Fig. 5.5) In addition, the best SP with respect to cutinase secretion (SP-Epr) was fused directly to the metagenome esterase, showing that the most efficient SP for one target results in unsatisfying low secretion amounts when fused to another export target. This direct comparison of secretion efficiencies of the same SPs fused to different target proteins underlines the secretion target specificity of the "best" SP. Futhermore, like in cutinase secretion, no clear correlation of secretion efficiencies of the target protein and parameters like charges, length or hydrophobicities were obvious.
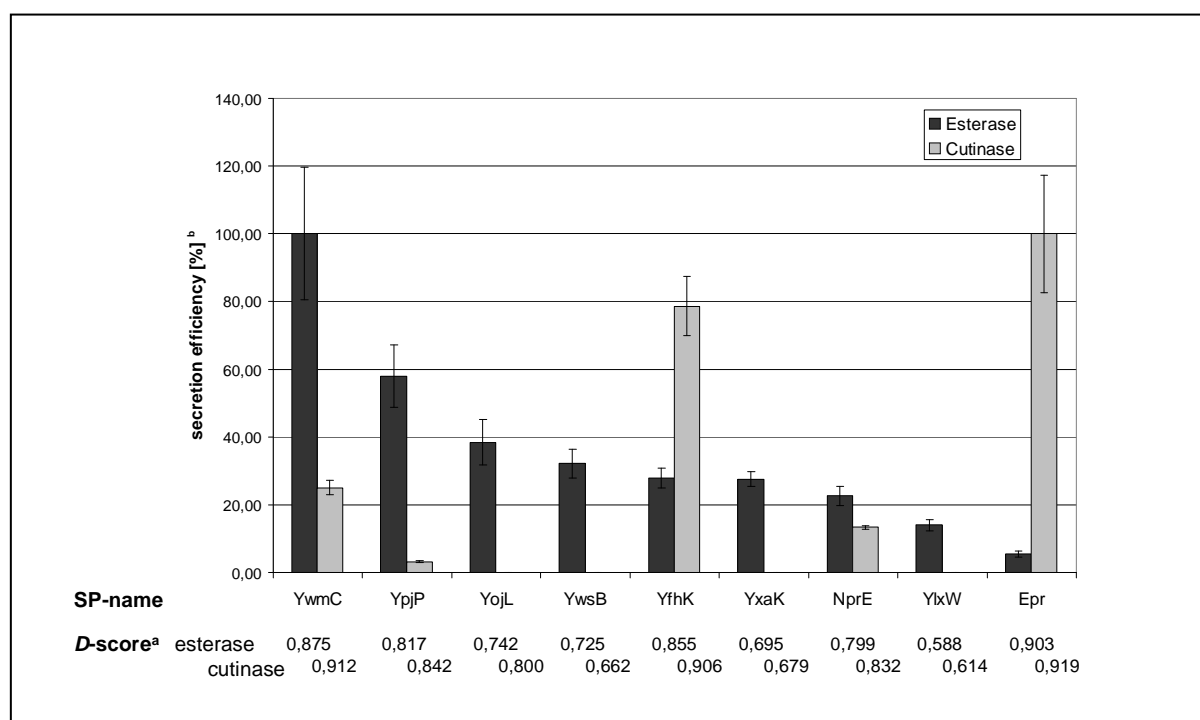

SP-name	YwmC	YpjP	YojL	YwsB	YfhK	YxaK	NprE	YlxW	Epr
D-score[a] esterase	0,875	0,817	0,742	0,725	0,855	0,695	0,799	0,588	0,903
cutinase	0,912	0,842	0,800	0,662	0,906	0,679	0,832	0,614	0,919

Fig. 5.5 Identification of the most efficient signal peptides in secretion of the heterologous lipolytic enzymes cutinase from *Fusarium solani pisi* and esterase EstCL1 from metagenomic origin in *B. subtilis*. Indices: [a]calculated for the first 70 amino acids of hydrid proteins SP-esterase EstCL1 and SP-cutinase, [b]secretion efficiencies of YwmC-esterase (1,5 U/mL) and Epr-cutinase (4,67 U/mL) were set to 100 %.

5.3 Optimization of heterologous protein secretion by directed evolution

One important result of the screening of all Sec type signal peptides (SPs) from *B. subtilis* described above in chapter 5.2 was the unpredictability of efficient SPs for heterologous protein secretion. As it is used in enzyme engineering when nothing is known about the structure-function relationships, directed evolution should provide the solution to identify (non-natural) SPs for heterologous protein secretion. We have chosen the SP of AmyE leading to a moderate secretion of cutinase in *B. subtilis* (0.67 U/mL) for further optimization by directed evolution strategies, using saturation mutagenesis in the N- as well as in the C-domain (Fig. 5.6A). Both regions are known to have high impact on the secretion level of the fusion partner [90,95]. Screening of saturation libraries in the C-domain revealed no further enhancement of the secretion level; however, screening of the saturation libraries in the N-region of SP-AmyE show four better variants leading to increased cutinase secretion (Fig. 5.6B, C).
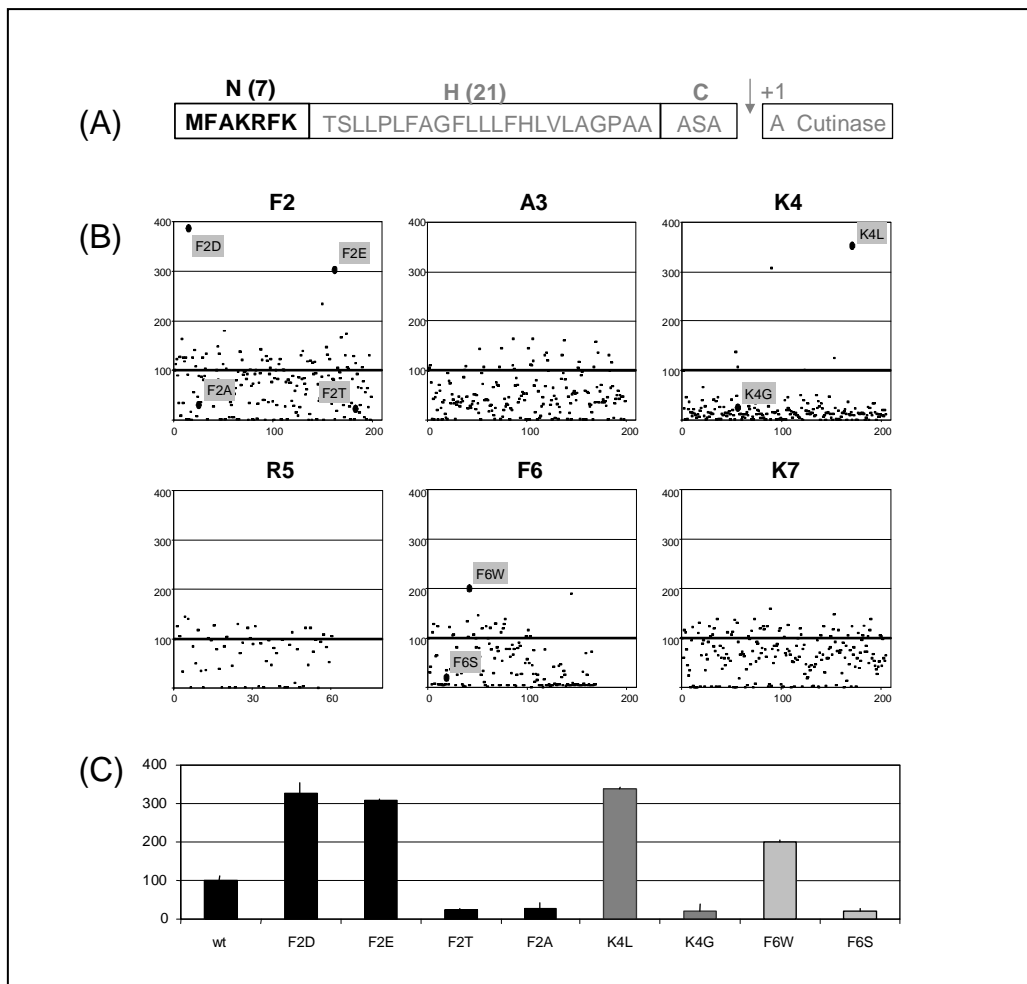

Fig. 5.6: Directed evolution of signal peptide SP-AmyE with respect to cutinase secretion efficiency in *B. subtilis*. (A) The N-region of SP-AmyE was randomly mutated by saturation mutagenesis. **(B)** At each position about 60-200 randomly chosen variants were screened for cutinase secretion yields. The wild type secretion level was set to 100 % as indicated by the bold line. **(C)** Finally, four better performing SP-variants and four variants showing decreased secretion efficiency were sequenced and investigated in more detail.

The four best SP-variants were sequenced, thereby identifying the befinicial substitutions as indicated in Figure 5.6. Interestingly, at position Phe2 of the N-domain the better performing variants have acquired exchanges (Phe→Asp and Phe→Glu) both resulting in a decrease of the net charge from +3 to +2, which was believed to be contradictory to secretion efficiency. In contrast, a significant reduction in secretion occurred when the signal peptide showed uncharged amino acids like alanine or threonine at this position. Other important positions of the SP-AmyE N-region in terms of cutinase secretion were identified to be Lys4 and Phe6. Again, like in the screening approach of all naturally occurring Sec-dependent SPs, the result of this directed evolution approach was not predictable. Pulse-chase experiments of the Phe→Asp and Phe→Glu SP-AmyE variants also indicated a reduced translocation and processing of the 3-fold more efficient signal peptides fused to cutinase (Fig. 5.7). Like the experiments presented in chapter 5.2, these results indicate a well balanced system to achieve high secretion efficiencies, which sometimes needs a slowed down translocation and processing in order to increase the overall secretion amount. Based on these results we speculate about a complex network of components affecting the overall secretion yields in *B. subtilis* as discussed below.
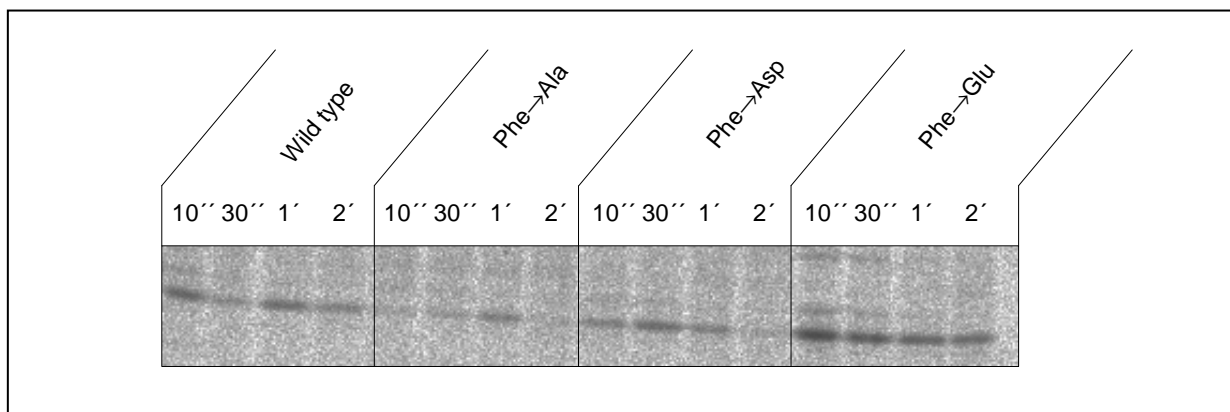

Fig. 5.7: Pulse-chase experiments showing the translocation and procession kinetics of cutinase fused to the SP-AmyE (wild type) and its SP-variants Phe$_2$→Ala, Phe$_2$→Asp and Phe$_2$→Glu. Translocation and processing of variant Phe$_2$→Glu is slowed down in comparison to the wild type. However, the overall secretion yield of this variant is three times higher than the wild type as shown in figure 5.6. The pulse chase experiments have been performed in cooperation to the research group of Prof. Freudl (Institute of Biotechnology I, Forschungszentrum Jülich, Germany).

5.4 Discussion

Nowadays, enzymes are widely used in industry for bulk applications like detergent additives in washing powder as well as specialized biocatalysts in the production of fine chemicals. The sufficient production, also referred to as (over)expression, of the enzymes in microbial hosts is an important prerequisite to carry out biocatalysis. Many enzymes are produced by *Bacillus* and *Aspergillus* strains because those expression hosts are able to secrete high amounts of enzymes into the culture medium. By this means, the enzyme isolation is rather

simple, reducing the overall costs of the catalyst. Unfortunately, secretion efficiencies of about 20 g/L culture medium are only achieved when homologous enzymes are produced (Table 5.1).

In order to improve heterologous protein secretion in *B. subtilis* and to elucidate the bottlenecks in heterologous protein secretion, we have systematically investigated the role of the N-terminal signal peptides (SPs) with respect to secretion efficiencies of heterologous proteins. Many secretion systems are known in Gram-negative [106-111] and Gram-positive (Fig. 5.1) [82,112] bacteria; however, the Sec pathway translocating the enzymes as unfolded polypeptide chains was chosen as being the most promising pathway with respect to high yield secretion of heterologous proteins.

Table 5.1: Industrial application of *Bacillus* spp. enzymes and their exclusive production in the homologous expression host.

Product	Application	Origin	Production host
alkaline protease (subtilisin)	Detergents	*B. subtilis / B. licheniformis*	*B. subtilis / B. licheniformis*
alkaline amylase	Detergents	*B. licheniformis*	*B. licheniformis*
pullulanase	starch degradation	*B. halodurans*	*B. halodurans*
glucose isomerase	starch processing	*B. coagulans*	*B. coagulans*
CGTase	starch conversion to cyclodextrins	*B. firmus*	*B. firmus*
β-glucanase	glucan modification	*B. subtilis*	*B. subtilis*
xylanase	food processing	*B. subtilis*	*B. subtilis*
cellulase	cellulose degradation	*B. subtilis*	*B. subtilis*
chitinase	chitin degradation	*B. thuringiensis*	*B. thuringiensis*
levansucrase	hydrolase/transferase	*B. circulans*	*B. circulans*
esterase	lipolytic degradation	*B. circulans*	*B. circulans*

In summary, the best SP for the secretion of one target protein is not automatically the best, or even a sufficient SP, for the secretion of another different target protein. Furthermore, efficient translocation and processing, such as the one mediated by the signal peptide from AmyE does not automatically result in high amounts of secreted protein, nor does a slow translocation and processing necessarily result in low amounts of secreted protein.

Instead, it appears that the secretion efficiency for a given target protein in *B. subtilis* is determined by a complex pattern of events. For example, the fusion of the secretion target to different SPs might result in different mRNA stabilities of the corresponding transcripts and, therefore, in different amounts of precursor proteins synthesized. Furthermore, after targeting to the membrane, translocation across the membrane, and processing, the availability of cell-associated chaperones, such as PrsA, and the presence of cell-associated and/or secreted proteases most likely also significantly influence the amounts of protein that accumulate in the culture supernatant.

Depending on the folding efficiency at the extracellular site of the cytoplasmatic membrane of *B. subtilis*, improperly folded protein accumulates in the cell wall to a greater or lesser extent. This folding efficiency is essentially determined by the availability of extracytosolic folding catalysts like PrsA [113]. In the case of PrsA overloading due to highly efficient targeting, translocation, and processing of the secretion target (e.g. as observed for cutinase fused to the SP-AmyE), posttranslocational folding could become a limiting step and improperly folded protein might accumulate in the cell wall. As a consequence of this secretion stress signal, cell-associated proteases like HtrA and HtrB are upregulated via the CssR/CssS two-component regulatory system [114,115].
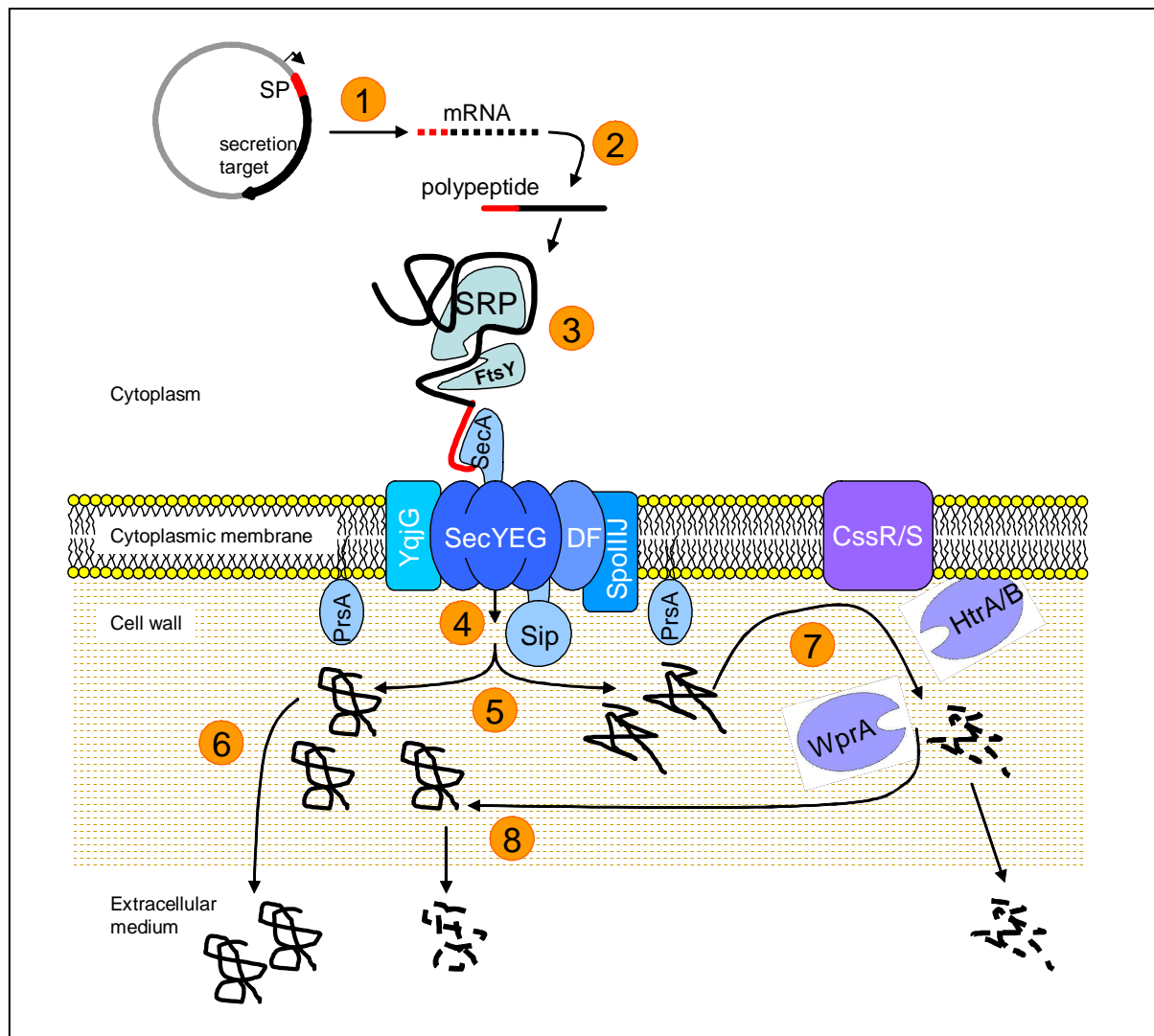

Fig. 5.8: Components affecting the overall secretion efficiency of heterologous proteins in *B. subtilis*. (1) Transcription and transcript stability, (2) translation, (3) targeting to the membrane, (4) translocation across the membrane and processing, (5) protein folding, (6) correctly folded proteins are released into the culture medium, (7) misfolding leads to secretion stress which induces proteases which degrade the incorrectly folded proteins and (8) potentially also some correctly folded proteins. Therefore, depending on the folding capacity of the heterologous secretion target within the cell wall sometimes slower translocation and processing leads to a higher overall yield in secretion. The system must be well balanced to channel most of the translocated proteins to the "left side" (correctly folded and released into the culture medium) instead of the "right side" which probably induces cell wall located proteases.

These proteases are mainly thought to degrade unfolded proteins. However, the increased HtrA/HtrB amounts might also result in an increased degradation of the pool of properly folded foreign proteins that are temporarily present in the cell wall, the extent of which depends on the individual sensitivity of the foreign protein towards these proteases. Both the degradation of the unfolded and properly folded foreign proteins will finally result in a reduction of the overall secretion efficiency of the respective heterologous secretion target. Consequently, in order to achieve high quantities of correctly folded heterologous target protein in the extracellular medium, the amount of translocated proteins emerging from the trans-side of the membrane (influenced by: transcription, translation, targeting, translocation and processing efficiencies) and the extracytosolic folding capacity (influenced by: specific folding properties of the secretion target and availability of extracytosolic folding catalysts) must be well balanced as visualized in figure 5.8.

In addition, accessory proteins like the Sec-motor protein SecA or chaperones like PrsA do also play a significant role in the translocation efficiency of heterologous proteins as already shown by different research groups (for overview see [82,85] and references therein). Nevertheless, directed evolution of these proteins is not published so far in literature. Therefore, in an ongoing project FtsH, PrsA and SecA have been subjected to random mutagenesis experiments and subsequent screening of the cutinase secretion amount. Preliminary results show high potential in further increases in secretion efficiencies; however, again indications of a balanced system appear. At the end a SecA variant will be perfectly optimized for one SP fused to a particular secretion target but the optimized SecA motor protein must not be the best with other SP secretion target fusions.

Chapter 6:

Novel biocatalysts by evolution
and design
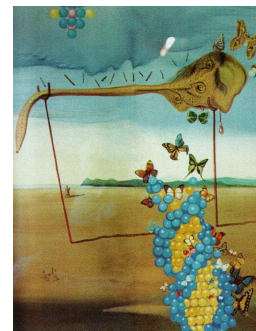
Salvador Dali – Butterfly landscape, 1957-58

Reetz, M.T., Puls, M., Carballeira, J.D., Vogel, A., Jaeger, K.-E., **Eggert, T.**, Thiel, W., Bocola, M. and Otte, N. (2006) Learning from directed evolution: further lessons from theoretical investigations into kooperative mutations in lipase enantioselectivity. *ChemBioChem.* **8**: 106-112*.*

Secundo, F., Carrea, G., Tarabiono, C., Gatti-Lafranconi, P., Brocca, S., Lotti, M., Jaeger, K.-E., Puls, M. and **Eggert, T.** (2006) The lid is a structural and functional determinant of lipase activity and selectivity. *J. Mol. Catal. B Enzym.* **39**: 166-170.

Krauss, U. and **Eggert, T.** (2005) *In.silico.mutagenesis*: a primer selection tool designed for sequence scanning applications used in directed evolution experiments. *BioTechniques* **39**: 679-682.

Funke, S.A., Otte, N., **Eggert, T.**, Bocola, M., Jaeger, K.-E. and Thiel, W.* (2005) Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. *Protein Eng. Des. Sel.* **18**: 509-514.

Eggert, T., Funke, S.A., Rao, N.M., Acharya, P., Krumm, H., Reetz, M.T., and Jaeger, K.-E. (2005) Multiplex-PCR-based recombination as a novel high fidelity method for directed evolution. *ChemBioChem.* **6**: 1062-1067.

Eggert, T., Leggewie, C., Puls, M., Streit, W., van Pouderoyen, G., Dijkstra, B.W. and Jaeger, K.-E. (2004) Novel biocatalysts by identification and design. *Biocal. Biotrans.* **22**: 139-144.

Funke, S.A., Eipper, A., Reetz, M.T., Otte, N., Thiel, W., van Pouderoyen, G., Dijkstra, B.W., Jaeger, K.-E. and **Eggert, T.** (2003) Directed evolution of an enantioselective *Bacillus subtilis* lipase. *Biocal. Biotrans.* **21**: 67-73.

6.1 Introduction

Today´s enzymes are the product of biological evolution which has taken several millions of years. They usually catalyze a given reaction with high specificity and enantioselectivity. However, since they are adjusted perfectly to their physiological role, their activity and stability are often far away from what organic chemists need. This is true for the stability of enzymes in organic solvents and particularly for substrate specificity and enantioselectivity of reactions yielding industrially important compounds.

Nature itself appears to provide a solution for this apparent dilemma: natural evolution produces a large number of variants by mutation and subsequently selects the "fittest" variant. This process can be mimicked in the test tube by using modern molecular biology methods of mutation and recombination. This collection of methods has been termed "directed" or "*in vitro*" evolution and provides a powerful tool for the development of biocatalysts with novel properties, without requiring knowledge on enzyme structures or catalytic mechanisms [116-119]. The general strategy for isolating enzymes with novel properties by directed evolution is outlined in figure 6.1.

In addition to the "random" strategy directed evolution, another state-of-the-art technology in molecular enzyme engineering follows a more rational way by analyzing the enzyme's structure and reaction mechanism to predict advantageous modifications. Therefore, this approach is called **rational design**. Both techniques, **directed evolution** and rational design, depend on enzyme modification at the DNA-level by introducing mutations into the gene; however, in both cases different additional information about the enzyme and molecular methods must be available as outlined below.
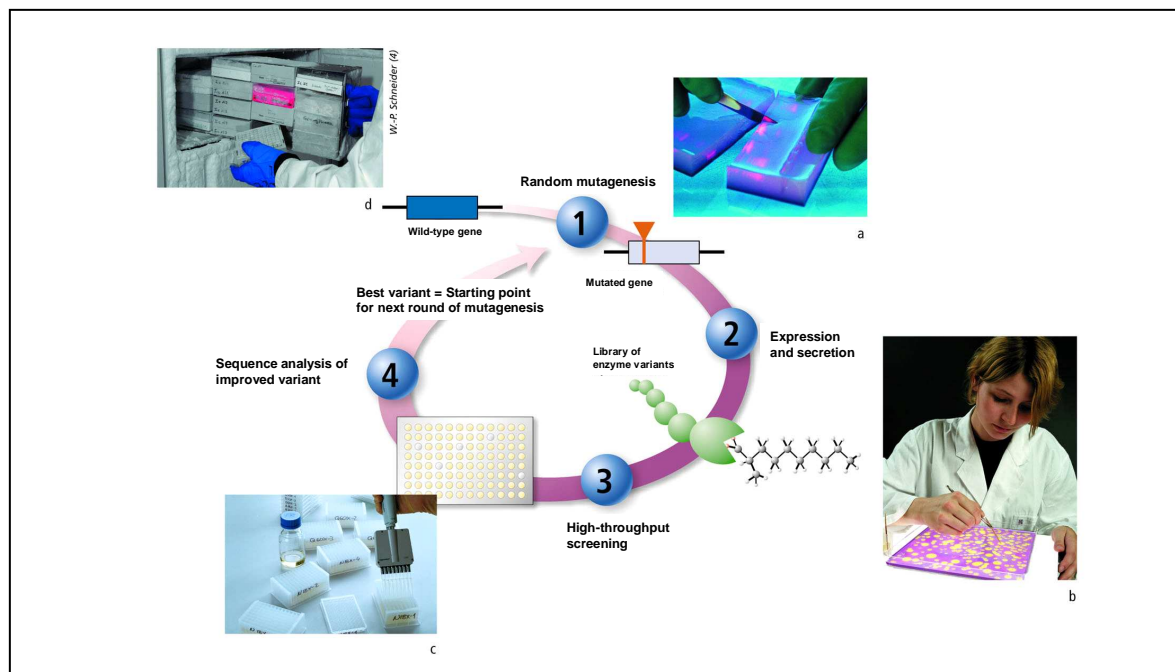

Fig. 6.1: Schematic overview of directed evolution experiments. (1) Variant gene libraries are generated by *in vitro* random mutagenesis using non-recombinative (introducing point mutations) or recombinative methods. (2) Gene libraries are cloned into expression vectors and the corresponding biocatalyst libraries are produced *in vivo* by microbial host strains. (3) Biocatalysts showing the desired properties are identified by high-throughput screening or selection systems. (4) The genes of best performing enzyme variants are isolated and sequenced to parent the next round of evolution until the desired properties are reached.

Rational enzyme design

In 1958 the first three-dimensional structure of a protein, namely myoglobin, has been reported by Kendrew [120,121], who won the Nobel Prize four years later in 1962 together with Perutz for their studies on the crystal structures of globular proteins. Since that time the number of novel protein and peptide structures solved by X-ray diffraction and, more recently, by NMR spectroscopy has reached 39.300 (October 2006); which, as a consequence, improved our knowledge on enzyme architecture and functionality considerably. From 1971 onwards, these biological macromolecular structures have been deposited in the Protein Data Bank (PDB) at Brookhaven National Laboratories (http://www.rcsb.org/pdb/) [122].

In addition, methods in the field of recombinant DNA are becoming more and more state-of-the-art technology and widely used in natural sciences. In particular, the polymerase chain reaction (PCR) developed by Mullis (Nobel Prize in 1993) simplifies the amplification of enzyme coding genes significantly and also the introduction of site directed mutations [123,124].

Using the novel techniques of recombinant DNA and the structural knowledge of the phage T4 lysozyme, Matsumura and co-workers published their ground breaking results on enzyme stabilization by molecular engineering in 1988. In site-directed mutagenesis experiments, two, four or six amino acid residues, spatially close to each other on the surface of the native enzyme were changed to cysteine residues. Consequently, the variant enzymes contained one, two or three disulfide bonds, hampering the thermal unfolding of the native enzyme structure. The triple-disulfide variant unfolds at a temperature 23.4°C higher than the wild type lysozyme [125-127].

Many examples using the rational design approach to increase enzyme stability (temperature, pH, organic solvents) or specific activity followed [117]. Successful examples of rational enzyme design to improve or invert enantioselectivity are relatively rare. However, Pleiss and co-workers (Institute of Technical Biochemistry, University of Stuttgart, Germany) have reported the improvement in enantioselective hydrolysis of linalyl acetate by *B. subtilis* *p*-nitrobenzyl esterase variants, as predicted by computer simulations. Furthermore, an inverted enantio-preference using 2-phenyl-3-butin-2-yl-acetate as a model substrate was achieved [128]. Other successful examples of improved enantioselectivities by rational protein engineering have been presented by the group of Hult (Department of Biotechnology, Royal Institute of Technology, Stockholm, Sweden) working on *Candida antarctica* lipase B (CALB) [129-132] and the group of Raushel (Department of Chemistry, Texas A&M University, USA) on phosphotriesterase from *Pseudomonas diminuta* [133,134]. Here at least two variants were created showing a million-fold difference in enantioselectivity towards the substrate ethyl phenyl *p*-nitrophenyl phosphate [134].

Nevertheless, there are still fundamental problems when applying rational enzyme design: (i) the three-dimensional structure of the enzyme and (ii) ideas with respect to the molecular functions of certain amino acid side chains must be available when rational design has to be applied. An alternative might be the development of a reliable structural model based on related enzymes. (iii) In general it is not possible to predict exactly the final structure of a variant enzyme using computer simulations; however, the methods continuously improve, and include theoretical methods using combined quantum mechanical and molecular mechanical calculations (QM/MM) [135]; however, in all simulations the effect of enzyme dynamics are neglected. (iv) Furthermore, solid state structures derived from crystallography could be different from protein structures in solution.

Directed evolution

Owing to the major difficulties encountered with rational protein design when creating biocatalysts that perform better towards interesting non-natural substrates, scientific researchers have established a collection of methods termed "directed" or "*in vitro*" evolution over the last decade. In this way they have provided a powerful tool-box for the development of biocatalysts with novel properties without requiring knowledge of the enzyme structures or catalytic mechanisms. This strategy mimicks natural evolution in the test tube; however, it reduces the time scale from millions of years to several months or even weeks. Like nature itself, the evolution in the test tube has been successfully applied in engineering enzymes with optimal catalyst performance with respect to specific activities, enzyme stabilities (pH, temperature, solvent), enantioselectivities, regioselectivities, substrate and cofactor specificities as well as expression rates and enzyme solubilities [136]. Molecular diversity is created by random mutagenesis and/ or recombination of a target gene or a set of related genes (i.e. gene family) using one of the methods summarized in Table 6.1.

Table 6.1: Molecular methods for directed evolution. The random mutagenesis methods are divided into four subclasses: (1) random point mutations, (2) insertion and deletion, (3) homology dependent *in vitro* recombination and (4) homology independent *in vitro* recombination. Recent review articles are recommended to obtain a detailed overview and critical comparisons of the methods listed here [137-139].

Method	Reference
1) random point mutations	
mutator strains (e.g. *E. coli* XL1-Red)	[140]
error-prone polymerase chain reaction (epPCR)	[141-144]
saturation mutagenesis	[145]
sequence saturation mutagenesis (SeSaM)	[146,147]
2) insertion and deletion	
random elongation mutagenesis (REM)	[148]
random insertion / deletion mutagenesis (RID)	[149]
random deletions and repeats	[150]
codon shuffling	[151]
codon-based random deletion (COBARDE)	[152,153]
frame shuffling	[154]
3) *in vitro* recombination (homology-dependent)	
DNA shuffling	[155-157]
family shuffling	[158]
staggered extension process (StEP)	[159]
random priming recombination (RPR)	[160]
heteroduplex recombination	[161]
ssDNA-family shuffling	[162,163]
degenerate oligonucleotide gene shuffling (DOGS)	[164]
random chimeragenesis on transient templates (RACHITT)	[165]
mutagenic and unidirectional reassembly (MURA)	[166]
synthetic shuffling	[167]
structure-based combinatorial protein engineering (SCOPE)	[168]
recombination-dependent exponential amplification PCR (RDA-PCR)	[169]
assembly of designed oligonucleotides (ADO)	[170]
recombined extension on truncated templates (RETT)	[171]
multiplex-PCR-based recombination (MUPREC)	[172]
4) *in vitro* recombination (homology-independent)	
incremental truncation for the creation of hybrid enzymes (ITCHY)	[173,174]
sequence homology independent protein recombination (SHIPREC)	[175]
combination of ITCHY and DNA-Shuffling (SCRATCHY)	[176,177]
exon shuffling	[178]
random multi-recombination PCR (RM-PCR)	[179]
sequence-independent site-directed chimeragenesis (SISDC)	[180]

The standard mutagenesis technique in directed evolution is still the easy to handle error prone polymerase chain reaction (epPCR) randomly introducing point mutations into a target DNA-sequence. The diversity of an enzyme library generated by epPCR is usually calculated by correlating the basepair-substitutions introduced per gene to the amino acid exchanges introduced per enzyme molecule. Afterwards, the overall size of a variant library can be calculated by a combinatorial algorithm [181] as shown in Table 6.2. This algorithm is based on the assumption that all 19 remaining amino acids can be introduced at a single position (E = 19). Unfortunately, this is not true for the case of epPCR, because the event of two or even three basepair exchanges per codon is highly unlikely. At best, one nucleotide of a given codon will be exchanged, thereby leading to just 9 (instead of 64 possible) different codons encoding 4 - 7 (instead of 20) different amino acids. In reality, the number of amino acid exchanges to be achieved depends on the type of the original codon as illustrated in figure 6.2. Silent mutations (i.e. those which do not result in an amino acid exchange) are more likely for some types of codons (e.g. CGA coding for arginine) than for other types (e.g. AAC encoding asparagine).

Table 6.2: Theoretical number of enzyme variants in a library obtained for an enzyme consisting of X=181 amino acids (e.g. lipase A from *Bacillus subtilis*) with one to five amino acid exchanges per molecule.

Number of amino acid exchanges [M]	number of variants[a] [N]
1	3.439
2	5.880.690
3	6.666.742.230
4	5.636.730.555.465
5	3.791.264.971.605.760

[a] values calculated with $E = 19$ using the algorithm:

$$N = \frac{E^M X!}{(X-M)!M!}$$

where N = number of variants at maximum size of diversity; E = number of amino acids exchanged per position; M = total number of amino acid exchanges per enzyme molecule; X = number of amino acids per enzyme molecule.

The mutational bias of DNA-polymerases poses yet another restriction to epPCR leading to a further lowered diversity of the mutant libraries. In most of the published studies that used *Taq*-polymerase in MnCl$_2$-containing buffer the enzyme preferentially introduced A→T, T→A transversions and A→G, T→C transitions. A→C and T→G transversions as well as G→A and C→T transitions were also introduced, but at much lower frequencies. The frequencies of transversions G→T and C→A were very low and G→C and C→G transversions hardly ever happened [182]. If this mutational bias was also taken into account (Fig. 6.2), the calculated library sizes represent only about 20 % of the theoretical sizes. Therefore, better and unbiased mutagenesis methods as reliable basis in directed evolution experiments must be established. Complete saturation mutagenesis is one strategy to solve this problem as outlined in chapter 6.2.

Fig. 6.2: Mutational bias of epPCR. The substitution of one nucleotide per codon results in nine new triplets which may encode four to seven different amino acids depending on the type of codon. **(A)** The example shows that the mutation of the codon AAC coding for asparagine can yield a maximum of seven different amino acids, whereas the mutation of the codon CGA coding for arginine can yield a maximum of four different amino acids. **(B)** Low frequencies of transversions G→T, C→A, G→C and C→G result in a further decrease of diversity: for codon AAC six different amino acid exchanges may occur, and for the GC-rich codon CGA just a single new amino acid exchange is expected. Background color coding: green shows codons which encode new amino acids, orange indicates silent mutations, yellow indicates the formation of stop codons, and red shows codons which would require for formation an unfavored base pair exchange (G→T, C→A, G→C or C→G). The blue letters indicate nucleotides exchanged by epPCR.

In addition to the methods introducing random point mutations, at least 21 *in vitro* recombination techniques have been published which are summarized in Table 6.2 and described in detail in recent review articles [137-139]. Among these approaches DNA-shuffling is still the method of choice for most directed evolution experiments. Other recombination methods use different experimental strategies to ensure exchange of DNA fragments from different variants. All these methods result in a significant improvement of the efficiency to create novel enzymes by directed evolution. However, they also have major drawbacks including (i) a recombinational bias depending on the target DNA and (ii) the creation of additional diversity by introducing novel point mutations during recombination, a process which may result in a library size far too large to be screened by available methods. Therefore, novel developments in directed evolution methodology focus on improving library quality instead of quantity [137].

A major strategy to reduce the size of a library is based on increasing the fidelity of the recombination process. The original DNA shuffling protocol led to introduce an average of seven novel point mutations per kb resulting in extra diversity [155,156]. This effect is favoured when screening capacity is not a limiting factor as for powerful selection systems like growth-defect complementation (see chapter 4.4), phage display [11] or fluorescence-activated cell sorting (FACS) [19]. Unfortunately, such systems are not available as yet to select for enzyme properties like enantioselectivity. Zhao and Arnold modified the DNA shuffling protocol to reduce the rate of newly introduced point mutations by using different DNA polymerases during fragment reassembly [157]. The creation of DNA fragments by using restriction endonucleases also reduced the number of novel point mutations; however, it also increased the bias of recombination [162]. Nevertheless, the methods based on DNase I digestion all have in common that large amounts of DNA are needed and the frequency of recombination is very low for neighbouring mutations.

As a consequence to this, still novel mutagenesis methods and even more important novel general strategies are necessary in directed evolution. In this chapter some solutions are evaluated: (i) complete saturation mutagenesis to overcome the mutational bias (chapter 6.2), (ii) multiplex-PCR-based recombination as a novel high fidelity method for *in vitro* recombination (chapter 6.3) and (iii) theory assisted semi-random mutagenesis as an intelligent way to narrow down library sizes (chapter 6.4).

6.2 First generation of directed evolution: complete saturation mutagenesis

As directed evolution experiments comprise iterative cycles of mutation and identification of improved variants by screening or selection, the first generation library must provide the starting material also for further rounds of evolution. The diversity of the first generation library is therefore very important, because some members of this library will parent all following variants. Different directed evolution methods may be considered to generate this library. The drawbacks of many random mutagenesis methods as outlined above, forced efforts to develop alternative techniques for preparing a reliable first generation library.

Site-specific saturation mutagenesis is a method to generate all possible variants of a protein at each amino acid position as calculated by the algorithm shown in Table 6.2. This method introduces all possible base tripletts at a given codon position; thereby, resulting in the formation of all 20 amino acids at this position of the protein. Complete mutagenesis at all positions finally yields a library comprising all single-site sequence variants of the corresponding protein.

Fig. 6.3: User interface of the *insilico.mutagenesis* primer prediction tool. (1) First, the program requires the input of a target nucleotide sequence, including flanking vector sequences in Fasta-format. It is not necessary to include the complete vector-sequence; about 40 base-pairs up- and downstream of the gene of interest is enough to enable primer design for whole gene saturation or scanning mutagenesis. (2) Second, a unique sequence identifier (sequence name) must be provided for data processing purposes. (3) Third, the mutagenesis codon must be selected from a pull-down menu. (4) Next, the program requires the input of the start- and stop-position of the target gene (or intergenic region) within the overall sequence. (5) Also, the region that should be mutated must be specified, and (6) in practice, too long megaprimer might be inefficiently elongated by the polymerase in the second round of PCR - probably due to the formation of secondary structures - it has been proven practical to design mutagenesis primers in a way that the megaprimer does not exceed a certain length [183]. Therefore, the input of a so-called oligo-switch position is necessary.

At present, three examples exist of complete saturation mutagenesis libraries being used as a starting point for directed evolution experiments. In one case, a haloalkane dehalogenase from *Rhodococcus rhodochrous* was mutagenized resulting in a significant increase in thermostability of this enzyme. A total of eight single-site variants with improved thermostability were identified in the library, and the combination of these eight mutations in a single second generation variant further improved the half-life of the enzyme by a factor of 30.000 [184]. The other case concerns the creation of a nitrilase variant which showed improved enantioselectivity in the hydrolysis of 3-hydroxyglutaryl nitrile at high substrate concentrations [185,186]. The third example of complete saturation mutagenesis was generated in our group in cooperation with the organic chemistry group of Prof. Reetz (Max-Planck Institut für Kohlenforschung, Mülheim a.d. Ruhr, Germany) with the objective to enhance the enantioselectivity of lipase A from *B. subtilis* (BSLA). In this project, directed evolution by the use of epPCR revealed several improved variants; however, further rounds of epPCR did not result in further improvement of enantioselectivity. In order to overcome the problems, a BSLA saturation library has been constructed by means of megaprimer-PCR, introducing mutations by using degenerated oligonucleotides in the first PCR [183]. In order to speed up the process of complete saturation mutagenesis a novel primer design tool called

insilico.mutagenesis has been developed accessible via the internet homepage www.insilico.uni-duesseldorf.de (Fig. 6.3).
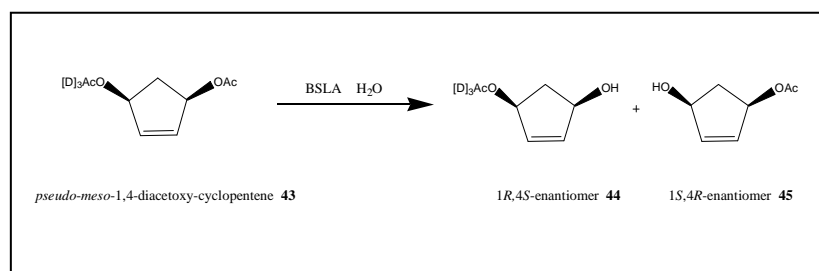

Fig. 6.4: The asymmetric hydrolysis of the model compound *meso*-1,4 diacetoxy-2-cyclopentene. The substrate *pseudo-meso*-1,4 diacetoxy-2-cyclopentene was deuterium labeled to follow the formation of chiral alcohols by differences in their mass spectrum.

Finally, a library consisting of about 70.000 clones (384 clones at each of 181 amino acid positions) of BSLA was screened for variants showing increased enantioselectivity in the asymmetric hydrolysis of the model substrate *pseudo-meso*-1,4-diacetoxy-cyclopentene with formation of the (1*S*,4*R*)- and (1*R*,4*S*)-enantiomers (Fig. 6.4) by using an ESI-MS-based assay [75]. Excess sampling is necessary in order to ensure that all of the 19 variants theoretically possible at each position have been generated and screened. Figure 6.5 exemplary shows typical screening results: saturation at position Gly30 produced many inactive enzyme variants, whereas at position Ser127, many clones with wild type properties were found showing a substrate conversion of 90-100 % and an enantioselectivity of about 40 % *ee*. The inactive variants (conversion rate ≤ 5 %) identified by high-throughput ESI-MS screening were not further investigated. The reasons for producing those variants include low or missing gene expression caused by the incorporation of rare or stop codons, as well relegation during cloning might result in clones which do not contain the lipase gene.

Identification of hot spot positions affecting enantioselectivity

For most amino acid positions, the results obtained from screening of the saturation library were similar to those shown for amino acids Gly30 and Ser127 (Fig. 6.5, upper panel). However, saturation at amino acid positions Asn18 and Tyr49 produced several variants with clearly altered enantioselectivities (Fig. 6.5, lower panel). The saturation library at position Tyr49 revealed a single interesting variant showing an inverted enantioselectivity of 40 % *ee*. The underlying mutation was identified as Tyr49Ile. The variants obtained from screening of the saturation library at position Asn18 showed reversed enantioselectivities while retaining high conversion rates. Therefore, six clones of this library showing high enantioselectivities were randomly chosen and their DNA sequence was determined; thereby, confirming that they all contained a BSLA gene carrying the desired mutation at the target codon-position no. 18. This result led us to conclude that the amino acid at position no. 18 represents a "hot spot" with respect to the enantioselective hydrolysis of the model substrate.
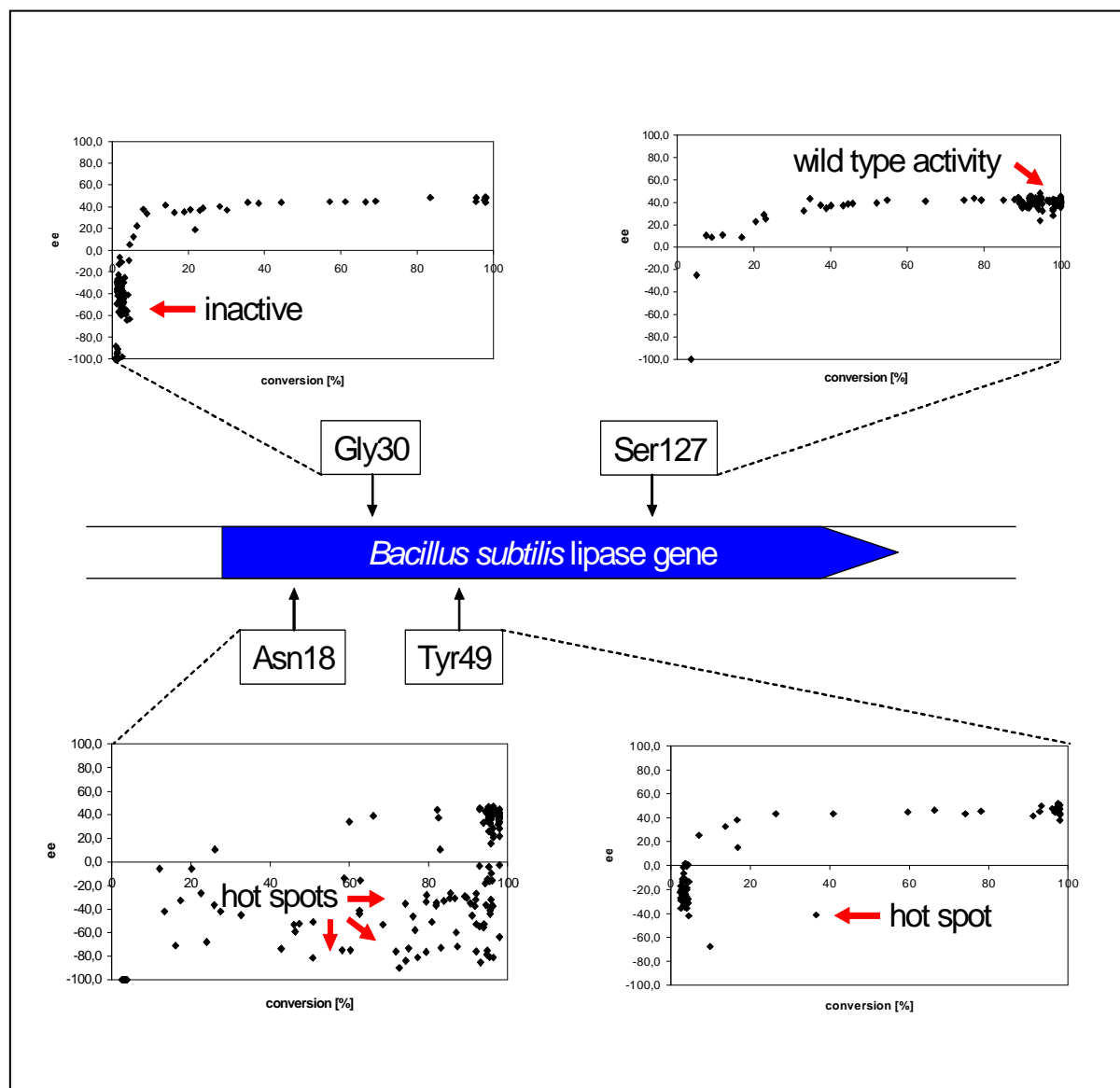
Fig. 6.5: Enantioselectivities and activities of selected variants obtained from the complete saturation mutagenesis library of BSLA by high-throughput screening using ESI-MS. Saturation at amino acid Gly30 revealed many inactive enzyme variants, whereas most variants at Ser127 showed wild type activity (90-100% substrate conversion) and enantioselectivity (*ee* = 40 %) as indicated with red arrows. Amino acid positions Asn18 and Tyr49 were identified as "hot spots" with respect to the enantioselectivity of BSLA.

Consequently, we have constructed separately all 19 variants replacing Asn18 by site directed mutagenesis. These variants together with wild type BSLA were overexpressed intracellularly in *E. coli* BL21 (DE3) as N-terminal His-tagged fusion proteins. The expression levels differed between the variants, but nevertheless, they were all catalytically active as determined by an agar plate assay using tributyrin as the substrate and spectrophotometrically with *p*-nitrophenyl-palmitate as the substrate.

Asparagine 18 determines the enantiopreference of BSLA towards the model-substrate meso-1,4-diacetoxy-2-cyclopentene

Interestingly, all 19 BSLA-variants replacing Asn18 displayed reversed enantioselectivities towards the model substrate. The wild type BSLA hydrolyzed *pseudo-meso*-1,4-diacetoxy-cyclopentene with an *ee*-value of about 40 % producing the (1*R*, 4*S*)-enantiomer. Variants Asn18Asp, Asn18Pro, Asn18Val and Asn18Ala did not exhibit marked enantioselectivities with *ee*-values ranging from 2-15%. Nine variants, namely Asn18Gly, Asn18Thr, Asn18Phe, Asn18Tyr, Asn18Ile, Asn18Lys, Asn18Trp, Asn18Arg and Asn18His, converted the substrate with moderate but inverted enantioselectivities (about 30 – 55 % *ee*) producing the (1*S*, 4*R*)-enantiomer whereas variants Asn18Glu, Asn18Met, Asn18Leu, Asn18Cys, Asn18Gln and Asn18Ser hydrolyzed the model substrate at moderate to high enantioselectivities of about 75 – 88 % *ee*, with Asn18Ser as the best-performing variant resulting in an *ee* = 88% (Fig. 6.6). The importance of amino acid position number 18 in BSLA for the enantioselective recognition of chiral substrates has recently been confirmed independently. An Asn18Ile variant was identified which hydrolyzed butyrate esters of *rac*-1,2-*o*-isopropylidene-*sn*-glycerol (IPG) with inverted enantioselectivity as compared to the wild type enzyme. Here, the variant library was constructed by a cassette mutagenesis approach mutating several amino acid residues surrounding the active site. The library was screened with a phage display selection system using immobilized IPG-phosphonate inhibitors [15,17].
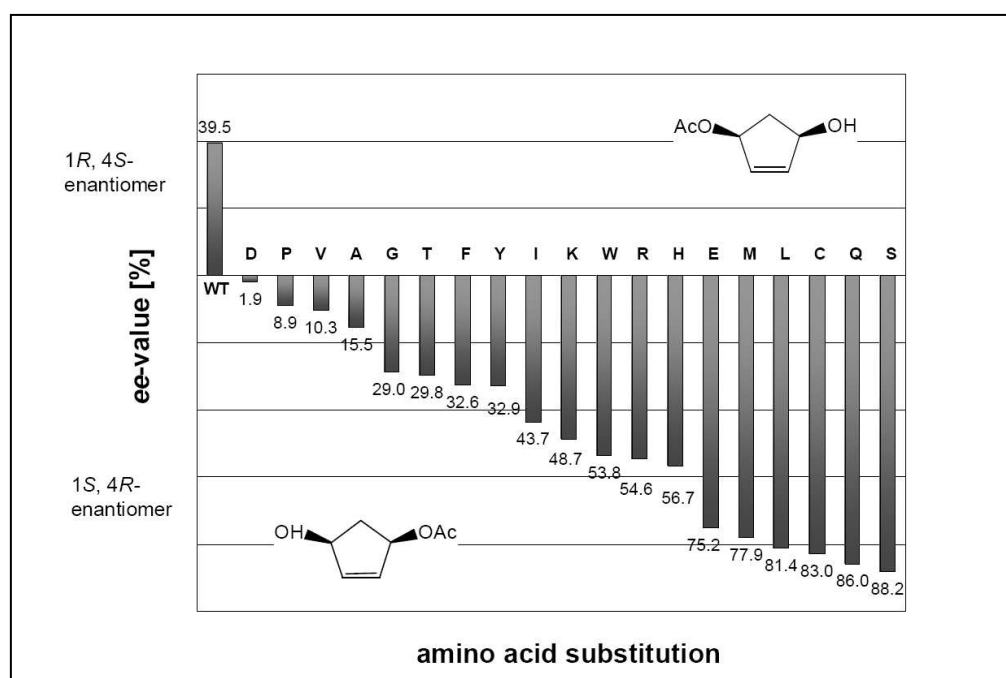

Fig. 6.6: Amino acid position Asn18 determines the enantioselectivity of BSLA. Enantioselective hydrolysis of *meso*-1,4-diacetoxy-2-cyclopentene catalyzed by wild type BSLA (wt) and 19 variants of Asn18 (amino acid substitutions are shown in the one-letter code). The *ee*-values were obtained by chiral GC from five independent experiments with standard deviations being lower than 5 %.

6.3 Second generation of directed evolution: *in vitro* gene-shuffling using multiplex-PCR based recombination

Usually, directed evolution is an iterative process of random mutagenesis and selection of better performing variants which can be used as basis for another round of mutagenesis and selection. As presented in chapter 6.2 the first generation of random mutagenesis libraries must be created carefully with respect to diversity; nevertheless, also for the following rounds of random mutagenesis the library construction is of great importance. In many directed evolution approaches performed by various groups it was beneficial to prepare the second generation by methods of *in vitro* recombination. Thereby, advantageous mutations can be combined and deleterious ones can be eliminated in a process called back-crossing. As obvious from Table 6.1 many recombination techniques are available today; however, they also have major drawbacks as discussed above.

Multiplex-PCR-based recombination (*MUPREC*) was developed to overcome some of these limitations of the already published *in vitro* recombination techniques, in order to combine beneficial mutations identified in a first round of mutagenesis, randomly without introducing new base substitutions. Such a method should be able to combine a large number of single point mutations without creating gene libraries that exceed the capacity of high-throughput screening systems. The random recombination of 15 point mutations will generate a library consisting of about 32.700 variants (see formula in Table 6.3) which is believed to be screened by most high-throughput screening methods. Therefore, such a method will generate high quality libraries having appropriate diversity while keeping library sizes in a screenable range.

Table 6.3: Theoretical library sizes generated by randomly recombined point mutations.

number of single point mutations[a]	number of recombinants[b]
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1.024
11	2.048
12	4.096
13	8.192
14	16.384
15	32.768

[a] number of single point mutations to be recombined
[b] number of true recombinants without any new point mutation,

calculated using the formula $\sum_{k=0}^{n}\binom{n}{k}=2^n$

n = number of single point mutations to be recombined
k = overall number of point mutations present in the variant protein

MUPREC is a two-step procedure for (i) amplification of DNA fragments that carry the selected point mutations for recombination and (ii) the random combination of those fragments to result in the full size gene (Fig. 6.7; left panel). In principle both steps have to occur separately; however, when the template gene is used with two different flanking regions and specific flanking PCR-primers are used, the *MUPREC* protocol can be reduced to an all-in-one reaction (Fig. 6.7; right panel).
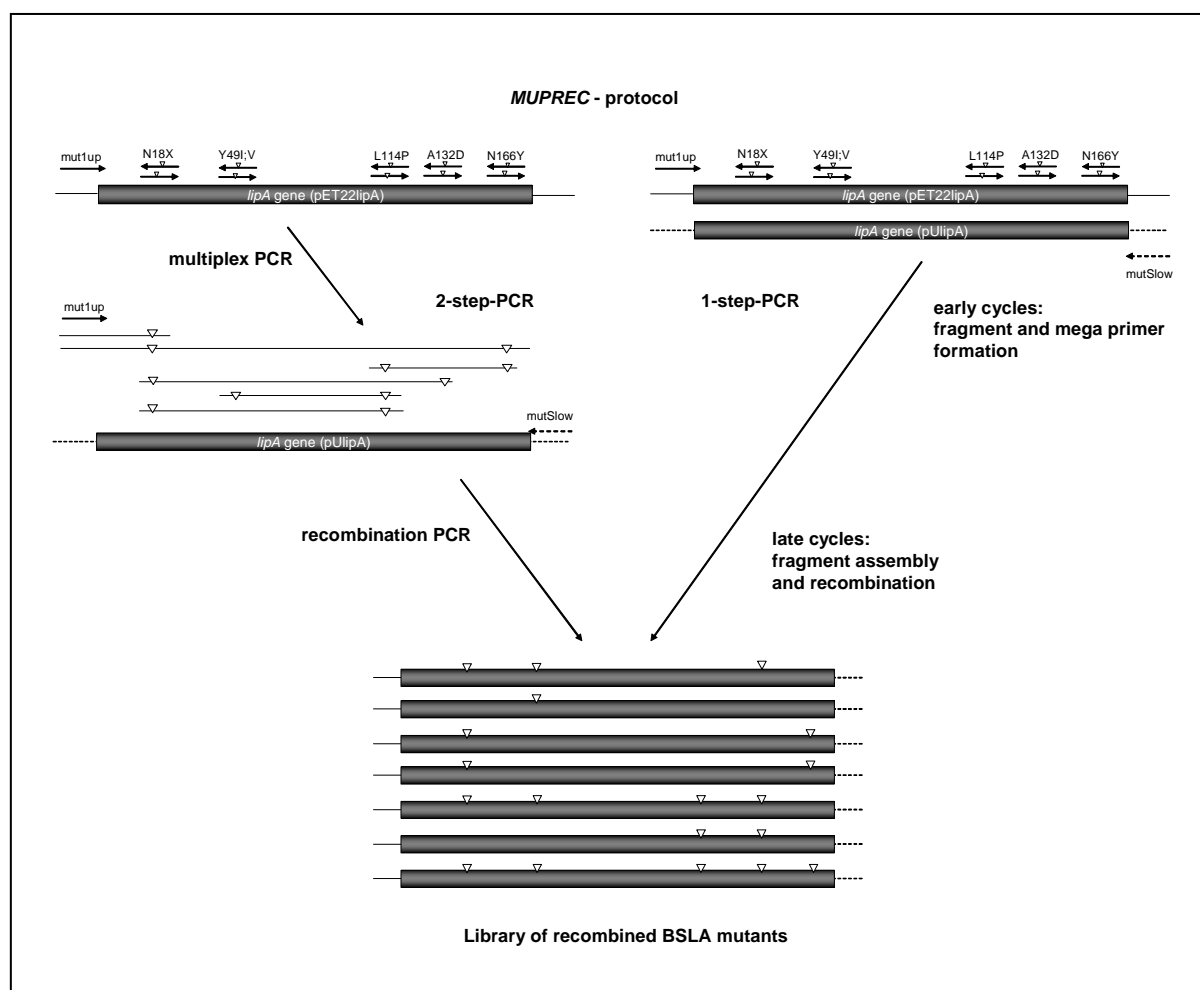

Fig. 6.7: The multiplex-PCR-based recombination (*MUPREC*) process. Mixtures of upper and lower primer-pairs that carry the point mutations to be recombined are used in a multiplex-PCR to amplify gene fragments, which are recombined in a second PCR. The efficiencies of fragment formation during the multiplex-PCR which are mainly determined by the melting temperature of the mutagenesis primers, can directly be monitored using the 2-step method. Alternatively, the one-step protocol can be applied for convenient and high fidelity recombination. Triangles indicate point mutations.

The method has been proven to be useful in molecular evolution by recombining and subsequent screening of point mutations identified in epPCR and saturation mutagenesis (see chapter 6.2) experiments in order to improve the enantioselectivity and thermostability of *B. subtilis* lipase A (BSLA). Our results clearly show that *MUPREC* allows the directed recombination of previously identified point mutations without introducing a significant

number of novel and possibly unwanted mutations which is in contrast to other homology-dependent recombination protocols like DNA-shuffling (7 additional point mutations per kb) or StEP (0.6 additional point mutations per kb). Therefore, the application of *MUPREC* may help to significantly reduce the screening efforts. As the consequence the overall time necessary for biocatalyst engineering projects will be reduced either.

6.4 Theory-assisted evolution of enzyme properties

Rational design and directed evolution seem to be two completely different techniques based on contradicting scientific philosophies. Rational design has a much longer history and is used in molecular biology to improve enzyme properties. Directed evolution was for the first time proved to be useful in the improvement of enzyme properties, this strategy was used in parallel to rational design. Either one <u>or</u> the other method was applied by different research groups, splitting the scientific community into two distinct sections: one believing in rational design and the other trusting directed evolution as being the more efficient strategy for molecular engineering.

Nowadays, rational design and directed evolution are still distinguished from each other; however, researchers see the beneficial aspects in both techniques and start to combine them in order to speed up the engineering process [187-189]. In my group two different approaches were investigated to speed up molecular enzyme engineering by combining aspects of rational and random approaches. On the one hand we have designed and inserted novel domains onto existing enzyme backbones and have subsequently optimized these enzymes by directed evolution. In a second project, we evaluated the strategy of a computational pre-screening using combined quantum mechanical and molecular mechanical (QM/MM) methods in cooperation with the group of Prof. Thiel at the Max-Planck Institut für Kohlenforschung (Mülheim a.d. Ruhr, Germany).

Construction of lipase variants containing artificial domains or domain deletions

Lipases of human, plant, and microbial origin were intensively investigated during the last decades to understand their kinetic properties and structure-function-relationships. Several three-dimensional lipase structures have revealed the presence of a lid-domain covering the active site which moves away upon contact with lipidic substrates allowing the substrates to enter the active site. This "induced fit" of the enzyme was in perfect agreement with the previous experimental observation of interfacial activation which is defined as a sharp increase in lipase activity when the substrate concentration reaches saturation. Human pancreatic and a fungal lipase were the first lipases whose three-dimensional structures were solved [190,191] providing this elegant explanation for the phenomenon of interfacial activation.

Both lipases possess α-helical polypeptide chains which cover the catalytically active site. Therefore, the active site residues were inaccessible to substrate molecules unless the lipase bound to a lipid interface. Conformational reorientations resulting in a movement of the so-called lid-structure made the active site residues accessible for the substrate. This hypothesis was later confirmed by elucidating the X-ray structures of inhibitor-bound lipases [192,193]. Furthermore, the lid-domain has been identified as being important for substrate recognition. Catalytic activity, substrate specificity [194,195], and enantioselectivity [196] of lipases are all influenced by the lid-domain. Also, the lid plays a key role for lipases catalyzing reactions in organic media. Here, lipases can be activated by using a procedure called interfacial activation-based molecular (bio)imprinting (IAMI), which "traps" the enzyme in an active form after detergent or fatty acid treatment [197,198]. This activated lipase should be in a permanently open conformation and therefore be highly active also in apolar organic solvents. In summary, the lid-like structural elements constitute important domains of lipases. During the last few years, three-dimensional structures of about 29 lipases, including 10 lipases from bacterial origin, have been elucidated (Table 6.4).

Table. 6.4: Overview of three-dimensional structures of lipases. A lid-domain was missing in the structures of lipases marked with an asterisk. All structures are available via the homepage of the Protein Data Bank (PDB) at Brookhaven National Laboratories (http://www.rcsb.org/pdb/).

Eucaryotic	
mammals / pancreatic lipases	*Bos taurus* (cattle)
	Canis familiaris (dog)
	Cavia porcellus (guinea pig)
	Equus caballus (horse)
	Homo sapiens (human)
	Rattus norvegicus (rat)
	Sus scrofa (pig)
mammals / gastric lipases	*Canis familiaris* (dog)
	Homo sapiens (human)
Yeast and Fungi	*Candida antarctica*
	Candida rugosa
	*Fusarium solani pisi
	Geotrichum candidum
	Penicillium camembertii
	Rhizomucor miehei
	Rhizopus delemar
	Rhizopus niveus
	Saccharomyces cerevisiae
	Thermomyces lanuginosa
Procaryotic	
Cyanobacteria	*Nostoc spec.*
Beta-Proteobacteria	*Burkholderia cepacia*
	Burkholderia glumae
	Chromobacterium viscosum
Gamma-Proteobacteria	*Pseudomonas aeruginosa*
Firmicutes	**Bacillus subtilis*
	Enterococcus faecalis
	Geobacillus stearothermophilus L1
	Geobacillus stearothermophilus P1
Actinobacteria	*Streptomyces exfoliates*

Surprisingly, not all of these structures did reveal the presence of a lid-domain. Cutinase from *Fusarium solani pisi* [199] and lipase from *Bacillus subtilis* [200] are catalytically active on long chain lipids without having a lid-domain. Furthermore, several examples are known of lipases which do not show interfacial activation with the lipases from *P. glumae, P. aeruginosa* and *Candida antarctica* (lipase B) being the best characterized examples. All these enzymes do contain the characteristic lid-domain which has to move to ensure catalysis as demonstrated for the *P. aeruginosa* lipase.

Therefore, rational design and directed evolution techniques were combined in the present work in order to study the role of the lid-domain for lipase-catalyzed reactions. Two model enzymes, namely the lid-containing lipase from *P. aeruginosa* (PAL) and the naturally lid-less *B. subtilis* lipase A (BSLA) were chosen for investigation. BSLA represents the minimal α/β-hydrolase structure and it has neither a lid covering the active site nor a lid-like loop located close enough to the active site to function as a lid. Therefore, this enzyme served as the backbone to engineer various artificial lid-domains close to the active site. Simultaneously, the lid-containing enzyme PAL should be engineered to isolate lid-less variants still showing enzymatic activity.

The pronounced lid-structure in PAL, indicated in green color in figure 6.8, was eliminated in different ways in order to restore the overall α/β-hydrolase fold: in the first approach the lid-domain (amino acids 117-163) was deleted and the flanking amino acids D116 and S164 were (i) directly connecting, (ii) connected by an additional glycine residue and (iii) by a random amino acid (construction of a saturation library containing 19 different variants). In a second approach based on homology comparisons to the lipase B from *Candida antarctica*, the amino acids 122-152 containing the lid-structure of PAL were deleted (iv) by connecting the hinge residues I121 and L153, and in another approach (v) by connecting the hinge residues I121 and L153 by randomizing simultaneously the neighbouring amino acids 153, 154 and 156 (construction of a saturation library of 8.000 different variants). Unfortunately, neither a single variant nor members of the variant libraries exhibited lipolytic activity. In ongoing random mutagenesis approaches still lid-less PAL-variants are screened. In this case the limits of rational evolution became obvious. A lid-less PAL-variant maybe difficult to create because of the complex folding assistance necessary in the wild type strain *P. aeruginosa*, which includes the specific folding chaperone LipH, the disulfide bond forming enzymes of the Dsb-system and the secretion machinery [201-203]. Therefore, interesting fundamentals of PAL-folding can be revealed once a lid-less variant can be generated. However, the missing active PAL-variant with solvent exposed active site underlines the importance of lid-structures in lipases in general and for *Pseudomonas* lipases in particular in order to reveal a correctly folded enzyme. Recent structural investigations by solving the

X-ray structure of *Burkholderia glumae* lipase co-crystallized with its corresponding chaperone indicate the direct contact of the folding catalyst to the lid-domain of the lipase. Therefore, in future evolution studies the chaperone must also be mutagenized and thereby adapted to the new PAL without lid-domain.
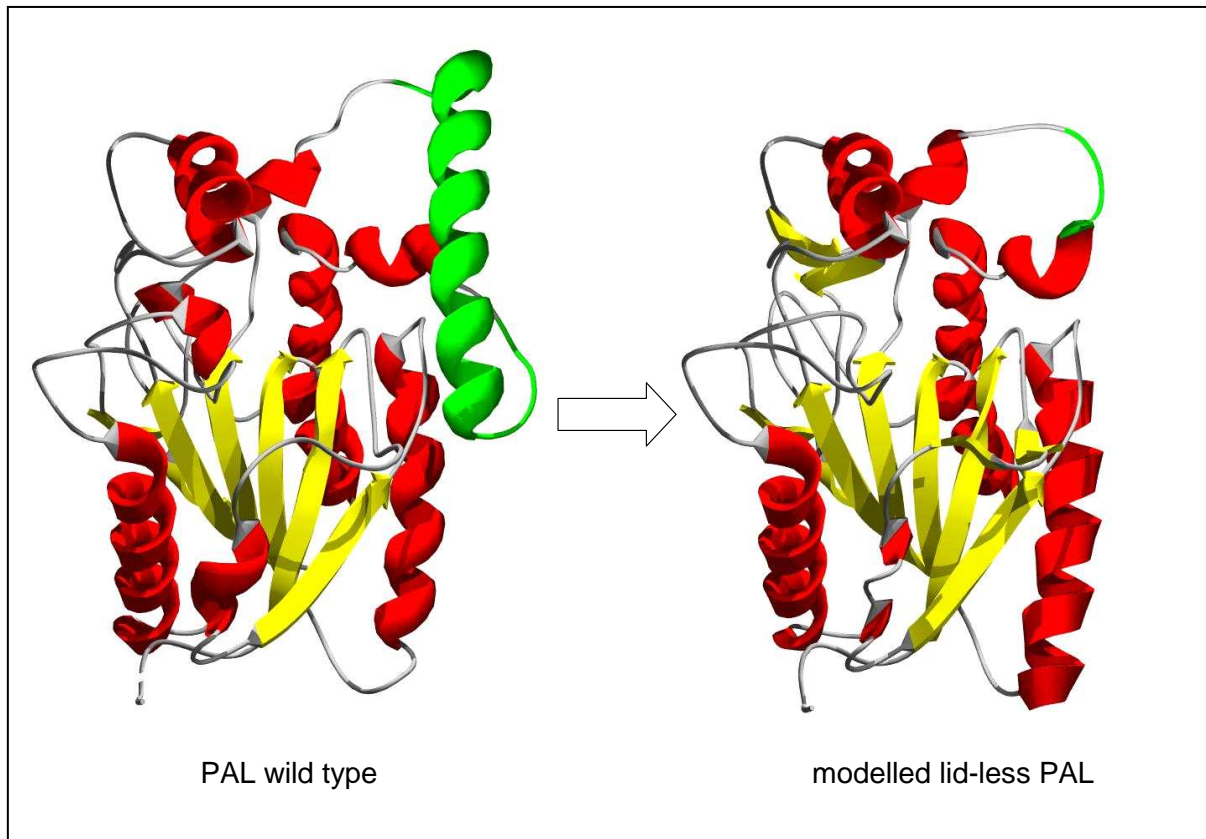

PAL wild type   modelled lid-less PAL

Fig. 6.8: Modelling of lid-less PAL-variants. The pronounced lid-structure in PAL wild type (left) is colored in green. The α/β-hydrolase backbone is shown in ribbon style with α-helices in red and the central β-sheet in yellow.

In computer-based studies the three lipolytic enzymes cutinase from *Fusarium solani pisi*, acetylxylanesterase from *Penicillium purpurogenum,* and human pancreatic lipase were compared with respect to the domains which are located close to the respective active site. These enzymes were chosen because their three-dimensional structures show a high homology to the structure of BSLA. Furthermore, lids or lid-like-domains were present in all three X-ray structures but are missing in BSLA. Subsequently, these lids were modeled into the structure of BSLA indicating experimental options to engineer these lids into BSLA without disturbing the core α/β-hydrolase fold (Fig. 6.9). This work has been done in close cooperation with Prof. Bauke W. Dijkstra (University of Groningen, The Netherlands) and Dr. Frédéric Carrière (CNRS-Marseille, France).
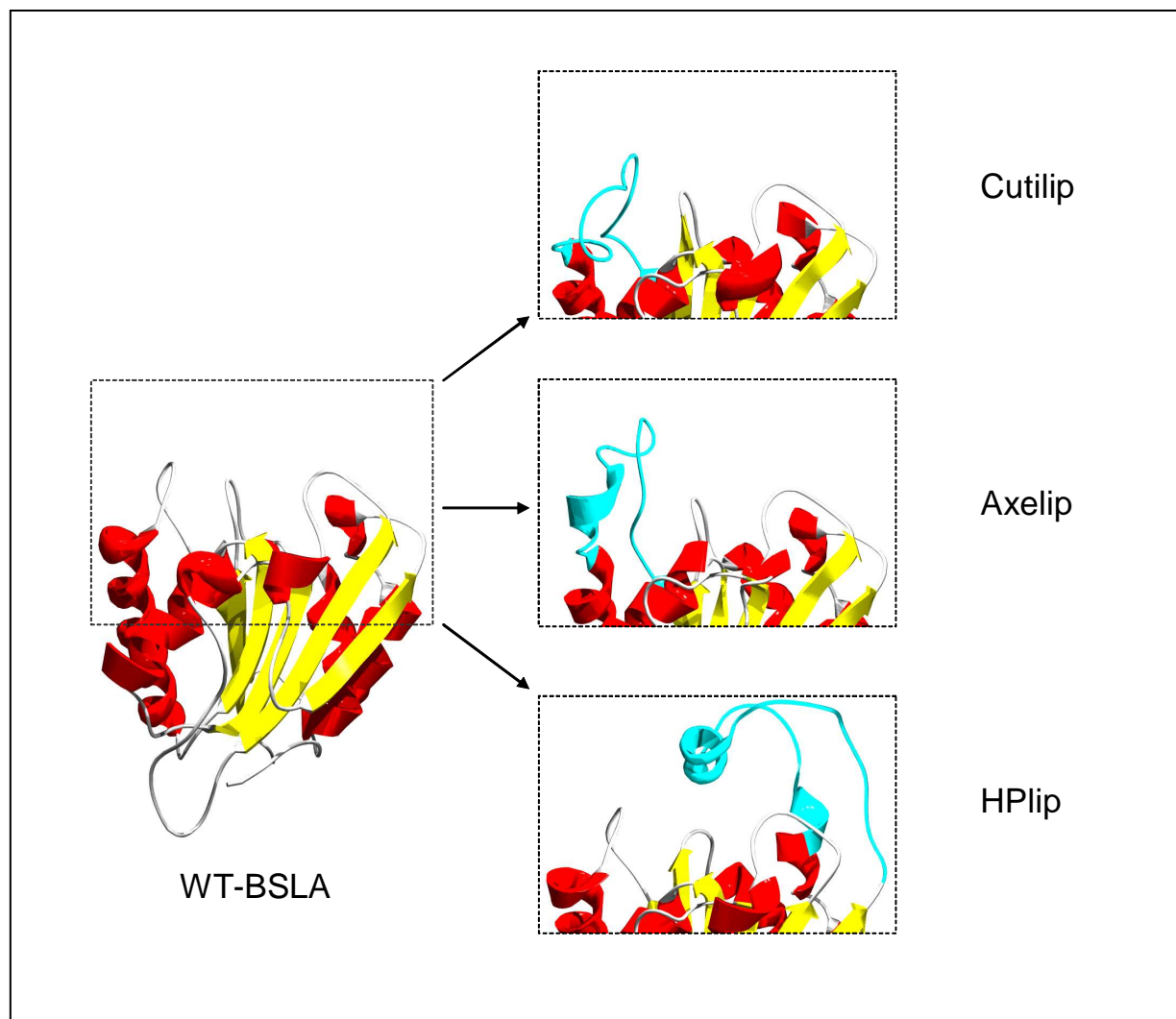
Fig. 6.9: Homology-based construction of BSLA-variants containing artificial lid-structures. The wild type X-ray structure of BSLA (left) is shown completely with the central β-sheet in yellow and the surrounding α-helices of the α/β-hydrolase fold in red. The solvent exposed active site is surrounded by a dashed-line box. The artificial lid-variants (right) contain the additional lid-cassette which is highlighted in light blue. In the BSLA-variant Cutilip the surface located loop at position V39-N51 was replaced by a cassette of 22 amino acids homologous to cutinase from *Fusarium solani pisi*. The variant Axelip was constructed in the same way by replacing the loop by a cassette of 24 amino acids homologous to acetylxylan esterase from *Penicillium purpurogenum*. The third BSLA-variant HPlip was constructed by replacement of the amino acids G153-G155 by a cassette of 28 amino acids homologous to the pronounced lid-domain of human pancreatic lipase.

The resulting BSLA-variants were overexpressed in high amounts in *E. coli* and tested for activity with the substrates tributyrin (pH-stat assay) and *p*-nitrophenyl palmitate (spectrophotometric assay). All variants show activity, at least against one of the substrates. However, all variants containing an artificial lid-domain exhibited a lower activity than the wild type. The maintainance of activity of all three of the lid-variants had not been taken for granted. However, the success of the approach verifies the preliminary considerations that the α/β-hydrolase fold is indeed a suitable scaffold for the introduction of artificial structural elements such as loops as functional modules while maintaining catalytic activity. Yet the resulting variants were less active than the wild type (Table 6.5), which does not come as a

surprise, since the wild type enzyme is a product of a long process of natural evolution, whereas the lid-structures were modelled onto the BSLA-scaffold at once without any selective pressure. The decrease in activity compared to the wild type enzyme was ascribed to the fact that the insertions did not yet fit properly to the backbone of the BSLA-scaffold. This is not so much valid for Axelip, but mostly for Cutilip and HPlip. Therefore, these variants (Cutilip and HPlip) were subjected to a directed evolution approach with epPCR as the method for random mutagenesis to find variants in which the lid-structures and the BSLA-backbone are better adapted to each other.

Table 6.5: Screening results of the first generation of Cutilip and HPlip variants showing increased lipolytic activity in comparison to the unmodified artificial lid-variants. The wild type BSLA without additional lid-domains show an activity of 327 ± 59 U/mL towards the used substrate.
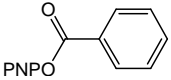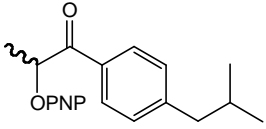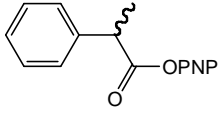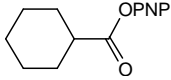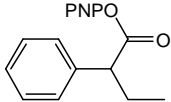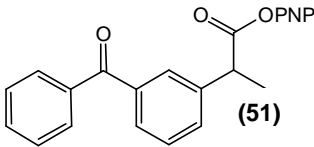
variant name	base substitution[a]	amino acid substitution	position of aa substitution	lipase activity [U/mL][b]	
Cutilip	-	-	-	24 ±	3.8
C8D12	aac→gac	N91D	α/β-scaffold	240 ±	13.7
C12F1	tac→cac	N91S	α/β-scaffold	49 ±	4.1
C9G4	aac→agc	Y148H	α/β-scaffold	40 ±	10.1
C14E9	ctt→cct	L123P	α/β-scaffold	63 ±	8.9
HPlip	-	-	-	6 ±	2
H1H2	ttc→tcc, aat→gat	F17S, N94D	α/β-scaffold	37 ±	20
H4G7	ttc→tac	F17Y	α/β-scaffold	20 ±	7
H13C8	aag→gag	K156E	artificial lid	36 ±	14
H13D1	aac→gac	K157D	artificial lid	10 ±	1.6

[a] base substitutions are highlighted in red
[b] activity towards *p*-nitrophenyl palmitate; the assay conditions were as follows: 0.8 mM substrate dissolved in isopropanol was used in Sørensen Phosphate buffer (47.3 mM Na_2HPO_4 x $2H_2O$, 2.7 mM K_2HPO_4, 5mM sodium dodecylsulfate, 1mg/mL gummi arabicum, pH 8.0)

A first generation of about 5000 clones of each lid-variant was screened for increased activity with at least 4 different mutants of each lid-variant having significantly higher catalytic activities (Table 6.5.). The combination of rational design and random mutagenesis has now produced highly active lipase variants containing additional structural elements, namely lids or lid-like structures covering or shielding the active site, respectively. The evolutionaly adaptation was necessary to "relax" the system, which makes the investigation of catalytic properties like interfacial activation or bio-imprinting possible. As well, enzyme properties like the regio- and stereoselectivity are modulated by the additional spatial hindrance close to the active site introduced by the artificial lid-structures. First experiments using samples of the evolved BSLA lid-variants indicate changes in substrate range and chiral discrimination of the substrates (Table. 6.6); thereby, a new field of mutagenesis was opened up without disturbing the structural core of the enzyme.

Table 6.6: Evolved BSLA lid-variants show altered substrate specificities. The catalytic activity of *B. subtilis* lipase A (wtBSLA), the three lid-variants of BSLA, which were created by rational design (Cutilip, HPlip, Axelip), and two lipase lid-variants further improved by directed evolution (C8D12, H1H2; see Table 6.5) were tested towards different *p*-nitrophenyl ester substrates (**46-51**). The first generation variant C8D12 of HPlip (see Fig. 6.9) and the rationally designed lid-variant Axelip (see Fig. 6.9) show broad substrate specificities.

Substrates[a]	Catalytic activity [U/mL][b]					
	wtBSLA	Cutilip	HPlip	C8D12	H1H2	Axelip
(46)	160,9 ± 28,6	23,5 ± 5,7	27,1 ± 7,4	65,9 ± 2,2	74,8 ± 16,4	39,4 ± 3,9
(47)	4,2 ± 0,7	-	-	25,2 ± 1,8	4,1 ± 0,5	19,0 ± 0,3
(48)	4,8 ± 1,3	1,5 ± 1,7	-	36,5 ± 0,5	-	28,7 ± 0,4
(49)	59,0 ± 5,3	20,9 ±1,8	16,5 ± 1,8	171,6 ± 1,9	57,1 ± 1,2	148,0 ± 2,4
(50)	-	-	-	13,8 ± 0,6	-	10,0 ± 1,1
(51)	-	-	-	38,3 ± 3,0	-	24,9 ± 2,1

[a] the substrates used in this study are *p*-nitrophenyl palmitate (**46**), *p*-nitrophenyl 2-(4-isobutyl-phenyl)-propanoate [ibuprofen ester] (**47**), *p*-nitrophenyl 2-phenylpropanoate (**48**), *p*-nitrophenyl cyclohexancarboxylate (**49**), *p*-nitrophenyl 2-phenylbutanoate (**50**), and *p*-nitrophenyl 3 benzoylbenzate [ketoprofen ester] (51). These substrates were provided by the group of Prof. Reetz (Max-Planck Institut für Kohlenforschung, Mülheim a.d. Ruhr, Germany).
[b] the enzyme was provided as culture supernatants; the assay was repeated five times using in independently grown cultures; the assay conditions were as follows: 1mg/mL substrate dissolved in acetonitrile was used in 100 mM Tris-HCl-buffer (pH 7.5

Computational equivalent of alanine scanning mutagenesis as library pre-screening tool

As described in chapter 6.2, a complete saturation mutagenesis library that contains all single-site amino acid substitutions of *B. subtilis* lipase A (BSLA) was available in my research group. In order to prove how useful this library was, we decided to screen the 3.439 BSLA-variants again towards a different substrate. As a model compound *rac*-1-(2-naphthyl)-

ethyl-acetate *rac*-**52** was chosen, which was hydrolyzed highly enantioselective by BSLA wild type leading to (*R*)-1-(2-naphthyl)-ethanol (*R*)-**54** with an *ee*-value of 99 % (Fig. 6.10A). An indicator agar plate assay was developed to identify BSLA-variants showing an inverted enantioselectivity of the reaction (Fig. 6.10B).
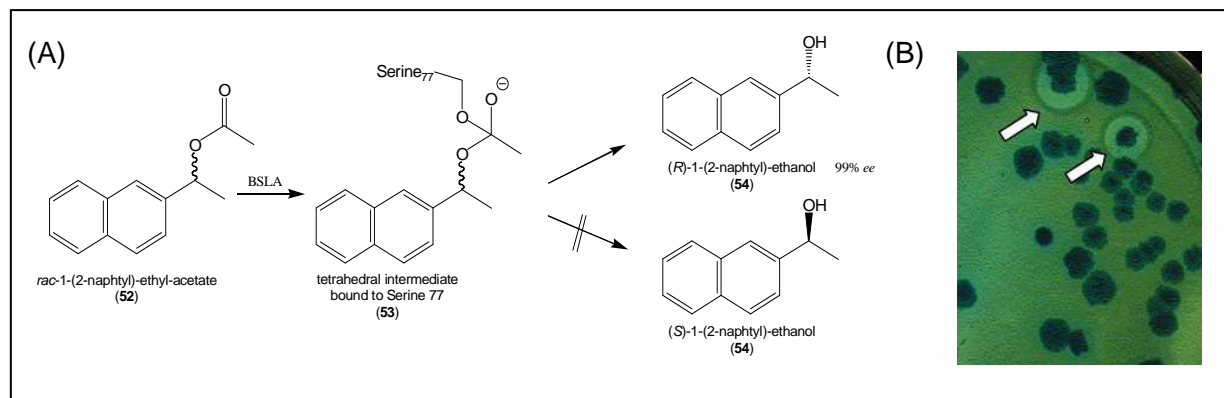

Fig. 6.10: Model reaction for the evaluation of computational pre-screening and high-throughput screening of the saturation mutagenesis library by indicator agar plates. (A) Reaction scheme of the enantioselective hydrolysis of the model compound *rac*-1-(2-naphthyl)-ethyl-acetate **52** by wild type BSLA leading to (*R*)-1-(2-naphthyl)-ethanol (*R*)-**54** with an *ee*-value of 99 %. **(B)** High-throughput screening on indicator agar plates for the hydrolysis of (*S*)-1-(2-naphthyl)-ethyl-acetate (*S*)-**54**. Activity is indicated by clear halos surrounding the bacterial colonies. Two clones that produce (*S*)-selective lipase variants are marked by arrows.

A total number of 10.500 clones were screened towards (*S*)-1-(2-naphthyl)-ethyl-acetate (*S*)-**52**, which represents a theoretical oversampling by a factor of three, thereby ensuring a complete coverage of the entire saturation library. It turned out that only one position, amino acid His76, was important with respect to the enantioselectivity of BSLA towards the model ester. Again the screening of the complete saturation mutagenesis library was successful to identify better performing enzyme variants. In this particular case, the screening of an epPCR library would have been failed due to the mutational bias of the method [139]. All identified variants needed two or three base exchanges in the same codon, which is impossible to achieve when epPCR is used for random mutagenesis. However, nowadays the creation of complete saturation mutagenesis libraries is still a tedious and time consuming strategy, and therefore not practical in all directed evolution projects. Hopefully, the improvements made in generating synthetic gene sequences, commercialized for example by the company Geneart (Regensburg, Germany), will make complete saturation mutagenesis a state-of-the-art technique in the near future, like sequencing and oligonucleotide synthesis today.

Nevertheless, in cooperation with the Theoretical Chemistry Group of Prof. Thiel (Max-Planck Institut für Kohlenforschung, Mülheim a.d. Ruhr, Germany) we found that a computational equivalent of alanine scanning mutagenesis based on QM/MM methodology can be applied to identify amino acid positions important for the activity of an enzyme. The electrostatic influence of all amino acid side chains in BSLA on the rate-determining reaction

barrier for hydrolysis of the model ester was estimated, highlighting the residue His76 being the only one seriously involved in the hydrolysis of (*S*)-**52** (Fig. 6.11).
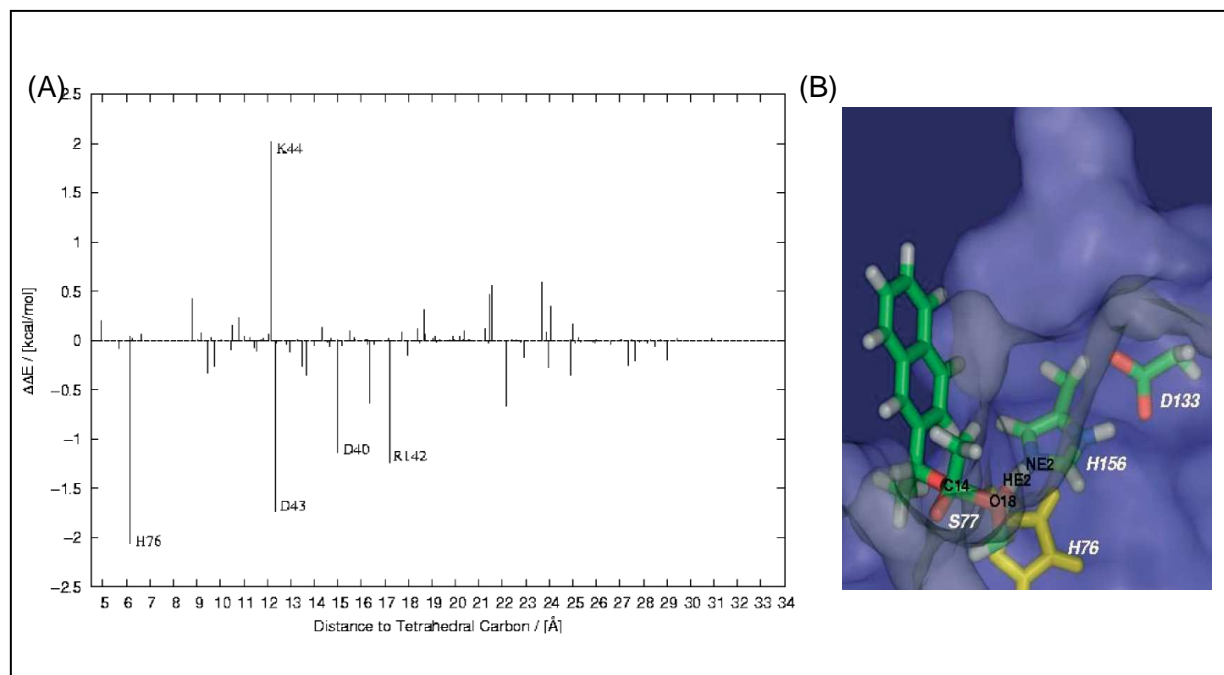

Fig. 6.11: Computational equivalent of alanine scanning mutagenesis based on QM/MM methodology. (A) Modulation of the reaction barrier height by electrostatic perturbation of the environment (deletion of charges on side chains). The distance is measured from the tetrahedral carbon in the tetrahedral intermediate (see Fig. 6.11B) to the geometric center of the individual amino acid side chains. Large contributions are labelled with amino acid positions. The five amino acid positions that have a pronounced effect (>1 kcal/mol) on the reaction barrier were futher investigated. Four of these (Lys44, Asp43, Asp40 and Arg142) represent ionizable groups located on the protein surface. We have added counterions close to the charged sites of the groups above and reevaluated the barrier. We found that the contributions drop below 1 kcal/mol for each group, and consequently, we do not consider them as hot spots. The remaining position identified in the QM/MM-scan was residue His76 which is located below the active serine. **(B)** Structural view of BSLA and its active site pocket. Amino acids of the proposed catalytic triad and His76 are displayed as stick models. Ser77 and (*R*)-**52** form a tetrahedral intermediate (see **53** Fig. 6.11A).

6.5 Discussion

Enzymes are cellular biocatalysts which were adapted by natural evolution to their natural environment, which is normally an aqueous solution at neutral pH, normal pressure and moderate temperature. However, the use of biocatalysts for White Biotechnology applications often requires conversions of non-natural substrates and under non-natural reaction conditions. An example is the chiral synthesis of cyanohydrines: the hydroxynitrile lyase catalyzed addition of hydrocyanic acid to aldehydes or ketones must be carried out in an acidic environment (pH value < 5.5) to avoid the non-catalyzed side reaction. Hence, the pH stability of the enzymes has to be adjusted accordingly. The lack of solubility of aromatic substrates and products in water comprises another frequently occurring problem: a biocatalytic reaction must be carried out in or in the presence of organic solvent, but most natural enzymes are inactive under these conditions. As a consequence, finally the creation

of tailor-made biocatalysts which are perfectly adjusted to particular process requirements will be the last step to establish a robust and competitive biocatalytic process. Therefore, the process-engineering will always include the molecular engineering of the biocatalyst itself (i.e. tailor-made enzymes) or when the whole cell is used as the biocatalsyst so-called "designer organisms" will be created.

The two state-of-the-art molecular engineering methodologies rational design and directed evolution constitute elegant tools to meet these requirements. However, as demonstrated in this work, novel and improved protocols must be developed to reach the objectives straightforwardly. Another strategy might be a compromise by the means of combining beneficial aspects of both technologies. In my opinion this **"rationalized evolution"** which sounds like a contradiction in terms will open a new field of molecular engineering, which might be more efficient than rational design or directed evolution alone. Promising results as presented here (chapter 5.2, 6.4) but also successful examples published by other research groups [187-189] indicate high potentials because of high quality mutant libraries containing many beneficial mutants.

Chapter 7:

*To achieve the possible,
one must attempt the impossible again and again.*

Hermann Hesse

Outlook

The production of chiral organic compounds in sufficient quality and quantity for the use in various fields of our daily life will be achieved more and more by the use of biocatalysts in addition to the classical chemical ways as outlined in the general introduction (chapter 2). In the scientific work presented here novel techniques and general strategies were developed in order to understand the functional principle of enzymes in more detail. Novel routes to identify biocatalysts from nature have been developed as well as new strategies in the field of enzyme production and secretion. Insights of enantioselective biocatalysis were provided by the application of directed evolution, i.e. the use of complete saturation mutagenesis and *in vitro* recombination identify the possibility to invert stereoselectivity of enzymes by exchanging one single amino acid side chain. In summary, contributions were made in all three sections as outlined in figure 7.1 in order to understand the conversion of a substrate "A" into a product "B" by the use of biocatalysis.
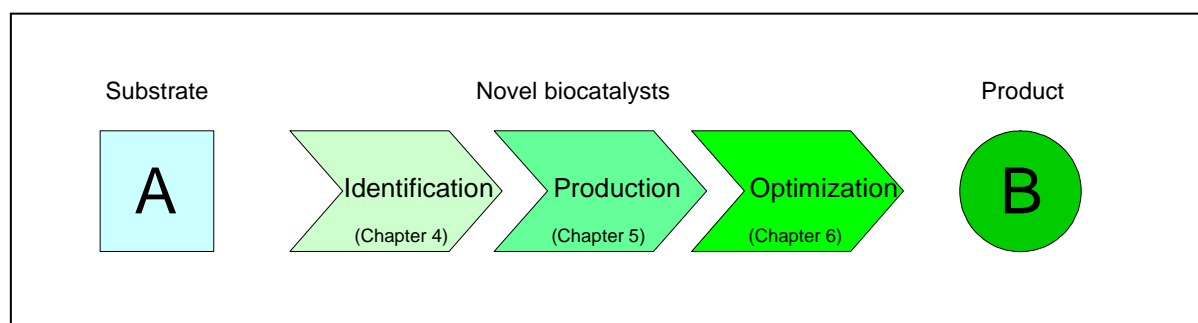

Fig. 7.1: The way from a desired chemical reaction to the technical biocatalytic process. First a new enzyme catalysing the target reaction must be identified from sources like enrichment cultures, strain collections, genome databases, metagenome or cDNA libraries. Once a promising biocatalyst is identified, it must be produced in microbial hosts in sufficient amounts. Finally, the enzyme, which is optimal adjusted over millions of natural evolution to its physiological function must be adapted to the technical process parameters.

In the future the development of appropriate biocatalytic processes will include the microbial cell as one unit. Systems biology will unravel more and more the secrets hidden by the complexity of microorganisms, which will open the door to create optimized designer organisms efficiently synthesising complex organic compounds, even with multiple chiral centers. Nowadays, the "omics-technologies" including flux and metabolom analysis as well as the molecular evolution and design methodology are well established; therefore, the combination of those disciplines with respect to develop tailor made process-microorganisms starts to become realistic. Thereby, a competitive sustainable chemistry will be established

using bio-renewables instead of petroleum as starting material. The industrial biocatalysis is now at a turning point to leave its niche position in chemical synthesis behind and to develop into a serious and competitive strategy for the production of fuel and energy (e.g. bio-ethanol, bio-H_2, etc.) as well as fine- and in parts also bulk-chemicals. Nevertheless, in order to reach these objectives the fundamental and applied research in the field of biocatalysis is worth to be intensified further.

Chapter 8:

References

1.   Fairley, J. (2003). Revisiting natural products. *Beyond Borders: The global biotechnology report 2003* Ernst & Young.

2.   Jarvis, L. M. (2006). Lure of biologics. *Chem Eng News* **84**, 22-23.

3.   SusChem. (2005). European Technology Platform for SUSTAINABLE CHEMISTRY, Sustainable chemistry strategic research agenda 2005.

4.   Griengl, H. (2006). Enzymkatalysierte Aldolreaktionen – gibt es hier neue Entwicklungen. Kolloquium Zentrum für Mikrobielle Biotechnologie.

5.   Liese, A., Seelbach, K. & Wandrey, C. (2006). *Industrial Biotransformations*. 2nd edit, Wiley-VCH, Weinheim.

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

7.   Overbeek, R., Larsen, N., Walunas, T., D'Souza, M., Pusch, G., Selkov, E., Jr., Liolios, K., Joukov, V., Kaznadzey, D., Anderson, I., Bhattacharyya, A., Burd, H., Gardner, W., Hanke, P., Kapatral, V., Mikhailova, N., Vasieva, O., Osterman, A., Vonstein, V., Fonstein, M., Ivanova, N. & Kyrpides, N. (2003). The ERGO genome analysis and discovery system. *Nucleic Acids Res* **31**, 164-171.

8.   Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-410.

9.   McGinnis, S. & Madden, T. L. (2004). BLAST: at the core of a powerful and diverse set of sequence analysis tools. *Nucleic Acids Res* **32**, W20-W25.

10.   Smith, G. P. (1985). Filamentous fusion phage - novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315-1317.

11.   Fernandez-Gacio, A., Uguen, M. & Fastrez, J. (2003). Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol* **21**, 408-414.

12.   Lin, H. N. & Cornish, V. W. (2002). Screening and selection methods for large-scale analysis of protein function. *Angew Chem Int Ed Engl* **41**, 4403-4425.

13.   Sieber, V., Plückthun, A. & Schmid, F. X. (1998). Selecting proteins with improved stability by a phage-based method. *Nat Biotechnol* **16**, 955-960.

14.   Verhaert, R. M. D., Beekwilder, J., Olsthoorn, R., van Duin, J. & Quax, W. J. (2002). Phage display selects for amylases with improved low pH starch-binding. *J Biotechnol* **96**, 103-118.

15.   Dröge, M. J., Rüggeberg, C. J., van der Sloot, A. M., Schimmel, J., Dijkstra, D. S., Verhaert, R. M. D., Reetz, M. T. & Quax, W. J. (2003). Binding of phage displayed *Bacillus subtilis* lipase A to a phosphonate suicide inhibitor. *J Biotechnol* **101**, 19-28.

16.   Reetz, M. T., Rüggeberg, C. J., Dröge, M. J. & Quax, W. J. (2002). Immobilization of chiral enzyme inhibitors on solid supports by amide-forming coupling and olefin metathesis. *Tetrahedron* **58**, 8465-8473.

17.   Dröge, M. J., Boersma, Y. L., van Pouderoyen, G., Vrenken, T. E., Rüggeberg, C. J., Reetz, M. T., Dijkstra, B. W. & Quax, W. J. (2006). Directed evolution of *Bacillus subtilis* lipase A by use of enantiomeric phosphonate inhibitors: crystal structures and phage display selection. *ChemBioChem* **7**, 149-157.

18.   Jose, J. (2006). Autodisplay: efficient bacterial surface display of recombinant proteins. *Appl Microbiol Biotechnol* **69**, 607-614.

19.   Becker, S., Schmoldt, H. U., Adams, T. M., Wilhelm, S. & Kolmar, H. (2004). Ultra-high-throughput screening based on cell-surface display and fluorescence-activated cell sorting for the identification of novel biocatalysts. *Curr Opin Biotechnol* **15**, 323-329.

20.   Goddard, J.-P. & Reymond, J.-L. (2004). Enzyme assays for high-throughput screening. *Curr Opin Biotechnol* **15**, 314-322.

21.   Reymond, J. L. & Wahler, D. (2002). Substrate arrays as enzyme fingerprinting tools. *ChemBioChem* **3**, 701-708.

22.   Wahler, D. & Reymond, J. L. (2001). High-throughput screening for biocatalysts. *Curr Opin Biotechnol* **12**, 535-544.

23.   Arnold, F. H. & Georgiou, G. (2003). *Directed enzyme evolution : screening and selection methods.* Methods in Molecular Biology, 230, Humana press, Totowa, New Jersey.

24.   Eisenthal, R. & Danson, M. (2002). *Enzyme assays: a practical approach*, Oxford University Press, Oxford.

25.   Reetz, M. T. (2002). New methods for the high-throughput screening of enantioselective catalysts and biocatalysts. *Angew Chem Int Ed Engl* **41**, 1335-1338.

26.   Reetz, M. T. (2003). An overview of high-throughput screening systems for enantioselective enzymatic transformations. In *Directed enzyme evolution: screening and selection methods* (Arnold, F. H. & Georgiou, G., eds.), Vol. 230, pp. 259-282. Humana Press, Totowa, New Jersey.

27.   Reetz, M. T. (2006). Directed evolution of enantioselective enzymes as catalysts for organic synthesis. *Adv Catal* **49**, 1-69.

28.   Farinas, E. T. (2003). Colorimetric screen for aliphatic hydroxylation by cytochrome P450 using *p*-nitrophenyl-substituted alkanes. In *Directed enzyme evolution: screening and selection methods* (Arnold, F. H. & Georgiou, G., eds.), Vol. 230, pp. 149-155. Humana Press, Totowa, New Jersey.

29.   Reetz, M. T., Zonta, A., Schimossek, K., Liebeton, K. & Jaeger, K.-E. (1997). Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angew Chem Int Ed Engl* **36**, 2830-2832.

30.   Badalassi, F., Wahler, D., Klein, G., Crotti, P. & Reymond, J. L. (2000). A versatile periodate-coupled fluorogenic assay for hydrolytic enzymes. *Angew Chem Int Ed Engl* **39**, 4067-4070.

31.   Leroy, E., Bensel, N. & Reymond, J. L. (2003). Fluorogenic cyanohydrin esters as chiral probes for esterase and lipase activity. *Adv Synth Catal* **345**, 859-865.

32.   Gonzalez-Garcia, E., Helaine, V., Klein, G., Schuermann, M., Sprenger, G. A., Fessner, W. D. & Reymond, J. L. (2003). Fluorogenic stereochemical probes for transaldolases. *Chem Eur J* **9**, 893-899.

33.   Sevestre, A., Helaine, V., Guyot, G., Martin, C. & Hecquet, L. (2003). A fluorogenic assay for transketolase from *Saccharomyces cerevisiae*. *Tetrahedron Lett* **44**, 827-830.

34.   Zocher, F., Enzelberger, M. M., Bornscheuer, U. T., Hauer, B. & Schmid, R. D. (1999). A colorimetric assay suitable for screening epoxide hydrolase activity. *Anal Chim Acta* **391**, 345-351.

35.   Beisson, F., Ferte, N., Nari, J., Noat, G., Arondel, V. & Verger, R. (1999). Use of naturally fluorescent triacylglycerols from *Parinari glaberrimum* to detect low lipase activities from *Arabidopsis thaliana* seedlings. *J Lipid Res* **40**, 2313-2321.

36.   Beisson, F., Tiss, A., Riviere, C. & Verger, R. (2000). Methods for lipase detection and assay: a critical review. *Eur J Lipid Sci technol* **102**, 133-153.

37.   Henke, E. & Bornscheuer, U. T. (2003). Fluorophoric assay for the high-throughput determination of amidase activity. *Anal Chem* **75**, 255-260.

38.   Konarzycka-Bessler, M. & Bornscheuer, U. T. (2003). A high-throughput-screening method for determining the synthetic activity of hydrolases. *Angew Chem Int Ed Engl* **42**, 1418-1420.

39.   Banerjee, A., Sharma, R. & Banerjee, U. C. (2003). A rapid and sensitive fluorometric assay method for the determination of nitrilase activity. *Biotech Appl Biochem* **37**, 289-293.

40.   Li, Z., Bütikofer, L. & Witholt, B. (2004). High-throughput measurement of the enantiomeric excess of chiral alcohols by using two enzymes. *Angew Chem Int Ed Engl* **43**, 1698-1702.

41.   Janes, L. E., Löwendahl, A. C. & Kazlauskas, R. J. (1998). Quantitative screening of hydrolase libraries using pH indicators: Identifying active and enantioselective hydrolases. *Chem Eur J* **4**, 2324-2331.

42.   Moris-Varas, F., Shah, A., Aikens, J., Nadkarni, N. P., Rozzell, J. D. & Demirjian, D. C. (1999). Visualization of enzyme-catalyzed reactions using pH indicators: Rapid screening of hydrolase libraries and estimation of the enantioselectivity. *Bioorg Med Chem* **7**, 2183-2188.

43. Zhao, H. (2003). A pH-indicator-based screen for hydrolytic haloalkane dehalogenase. In *Directed enzyme evolution: screening and selection methods* (Arnold, F. H. & Georgiou, G., eds.), Vol. 230, pp. 213-221. Humana Press, Totowa, New Jersey.

44. Breuer, M., Pohl, M., Hauer, B. & Lingen, B. (2002). High-throughput assay of (*R*)-phenylacetylcarbinol synthesized by pyruvate decarboxylase. *Anal Bioanal Chem* **374**, 1069-1073.

45. Lingen, B., Kolter-Jung, D., Dünkelmann, P., Feldmann, R., Grötzinger, J., Pohl, M. & Müller, M. (2003). Alteration of the substrate specificity of benzoylformate decarboxylase from *Pseudomonas putida* by directed evolution. *ChemBioChem* **4**, 721-726.

46. Doderer, K., Lutz-Wahl, S., Hauer, B. & Schmid, R. D. (2003). Spectrophotometric assay for epoxide hydrolase activity toward any epoxide. *Anal Biochem* **321**, 131-134.

47. Marchesi, J. R. (2003). A microplate fluorimetric assay for measuring dehalogenase activity. *J Microbiol Meth* **55**, 325-329.

48. Wahler, D. & Reymond, J. L. (2001). Novel methods for biocatalyst screening. *Curr Opin Chem Biol* **5**, 152-8.

49. Serra, S., Fuganti, C. & Brenna, E. (2005). Biocatalytic preparation of natural flavours and fragrances. *Trends Biotechnol* **23**, 193-198.

50. Abokitse, K. & Hummel, W. (2003). Cloning, sequence analysis, and heterologous expression of the gene encoding a (*S*)-specific alcohol dehydrogenase from *Rhodococcus erythropolis* DSM 43297. *Appl Microbiol Biotechnol* **62**, 380-386.

51. Weckbecker, A. & Hummel, W. (2004). Improved synthesis of chiral alcohols with *Escherichia coli* cells co-expressing pyridine nucleotide transhydrogenase, NADP$^{(+)}$-dependent alcohol dehydrogenase and NAD$^{(+)}$-dependent formate dehydrogenase. *Biotechnol Lett* **26**, 1739-1744.

52. Funke, S. A., Eipper, A., Reetz, M. T., Otte, N., Thiel, W., van Pouderoyen, G., Dijkstra, B. W., Jaeger, K.-E. & Eggert, T. (2003). Directed evolution of an enantioselective *Bacillus subtilis* lipase. *Biocatal Biotrans* **21**, 67-73.

53. Funke, S. A., Otte, N., Eggert, T., Bocola, M., Jaeger, K.-E. & Thiel, W. (2005). Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. *Protein Eng Des Sel* **18**, 509-514.

54. Eggert, T., Leggewie, C., Puls, M., Streit, W., van Pouderoyen, G., Dijkstra, B. W. & Jaeger, K.-E. (2004). Novel biocatalysts by identification and design. *Biocatal Biotrans* **22**, 139-144.

55. Sharma, M., Sharma, N. N. & Bhalla, T. C. (2005). Hydroxynitrile lyases: At the interface of biology and chemistry. *Enzyme Microb Technol* **37**, 279-294.

56. Schmidt, M. & Griengl, H. (1999). Oxynitrilases: From cyanogenesis to asymmetric synthesis. In *Biocatalysis - from Discovery to Application*, Vol. 200, pp. 193-226.

57. Breithaupt, H., Pohl, M., Bönig, W., Heim, P., Schimz, K.-L. & Kula, M.-R. (1999). Cloning and expression of (*R*)-hydroxynitrile lysae from *Linum usitatissimum* (flax). *J Mol Catal B: Enzym* **6**, 315-332.

58. Weis, R., Poechlauer, P., Bona, R., Skranc, W., Luiten, R., Wubbolts, M., Schwab, H. & Glieder, A. (2004). Biocatalytic conversion of unnatural substrates by recombinant almond R-HNL isoenzyme 5. *J Mol Catal B: Enzym* **29**, 211-218.

59. Förster, S., Roos, J., Effenberger, F., Wajant, H. & Sprauer, A. (1996). The first recombinant hydroxynitrile lyase and its application in the synthesis of (*S*)-cyanohydrins. *Angew Chem Int Ed Engl* **35**, 437-439.

60. Hasslacher, M., Schall, M., Hayn, M., Griengl, H., Kohlwein, S. D. & Schwab, H. (1996). Molecular cloning of the full-length cDNA of (*S*)-hydroxynitrile lyase from *Hevea brasiliensis*. Functional expression in *Escherichia coli* and *Saccharomyces cerevisiae* and identification of an active site residue. *J Biol Chem* **271**, 5884-5891.

61. Müller, M. & Sprenger, G. A. (2004). Thiamine-dependent enzymes as catalysts of C-C bonding reactions: the role of `orphan` enzymes. In *Thiamine: Catalytic mechanisms in normal and disease states* (Jordan, F. & Patel, M. S., eds.), pp. 77-92. Marcel Dekker, New York.

62. Pohl, M., Sprenger, G. A. & Müller, M. (2004). A new perspective on thiamine catalysis. *Curr Opin Biotechnol* **15**, 335-342.

63.	Samland, A. K. & Sprenger, G. A. (2006). Microbial aldolases as C-C bonding enzymes-unknown treasures and new developments. *Appl Microbiol Biotechnol* **71**, 253-264.

64.	Iding, H., Dünnwald, T., Greiner, L., Liese, A., Müller, M., Siegert, P., Grötzinger, J., Demir, A. S. & Pohl, M. (2000). Benzoylformate decarboxylase from *Pseudomonas putida* as stable catalyst for the synthesis of chiral 2-hydroxy ketones. *Chem Eur J* **6**, 1483-1495.

65.	Siegert, P., McLeish, M. J., Baumann, M., Iding, H., Kneen, M. M., Kenyon, G. L. & Pohl, M. (2005). Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. *Protein Eng Des Sel* **18**, 345-357.

66.	Hasson, M. S., Muscate, A., McLeish, M. J., Polovnikova, L. S., Gerlt, J. A., Kenyon, G. L., Petsko, G. A. & Ringe, D. (1998). The crystal structure of benzoylformate decarboxylase at 1.6 Å resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. *Biochemistry* **37**, 9918-9930.

67.	Polovnikova, E. S., McLeish, M. J., Sergienko, E. A., Burgner, J. T., Anderson, N. L., Bera, A. K., Jordan, F., Kenyon, G. L. & Hasson, M. S. (2003). Structural and kinetic analysis of catalysis by a thiamin diphosphate-dependent enzyme, benzoylformate decarboxylase. *Biochemistry* **42**, 1820-1830.

68.	Wilcocks, R., Ward, O. P., Collins, S., Dewdney, N. J., Hong, Y. P. & Prosen, E. (1992). Acyloin formation by benzoylformate decarboxylase from *Pseudomonas putida*. *Appl Environ Microbiol* **58**, 1699-1704.

69.	Demir, A. S., Pohl, M., Janzen, E. & Müller, M. (2001). Enantioselective synthesis of hydroxy ketones through cleavage and formation of acyloin linkage. Enzymatic kinetic resolution via C-C bond cleavage. *J Chem Soc Perkin Trans* **1**, 633-635.

70.	Demir, A. S., Sesenoglu, O., Eren, E., Hosrik, B., Pohl, M., Janzen, E., Kolter, D., Feldmann, R., Dünkelmann, P. & Müller, M. (2002). Enantioselective synthesis of alpha-hydroxy ketones via benzaldehyde lyase-catalyzed C-C bond formation reaction. *Adv Synth Catal* **344**, 96-103.

71.	Nelson, K. E., Weinel, C., Paulsen, I. T., Dodson, R. J., Hilbert, H., dos Santos, V. A. P. M., Fouts, D. E., Gill, S. R., Pop, M., Holmes, M., Brinkac, L., Beanan, M., DeBoy, R. T., Daugherty, S., Kolonay, J., Madupu, R., Nelson, W., White, O., Peterson, J., Khouri, H., Hance, I., Lee, P. C., Holtzapple, E., Scanlan, D., Tran, K., Moazzez, A., Utterback, T., Rizzo, M., Lee, K., Kosack, D., Moestl, D., Wedler, H., Lauber, J., Stjepandic, D., Hoheisel, J., Straetz, M., Heim, S., Kiewitz, C., Eisen, J., Timmis, K. N., Düsterhoft, A., Tümmler, B. & Fraser, C. M. (2002). Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. *Environ Microbiol* **4**, 799-808.

72.	Regenhardt, D., Heuer, H., Heim, S., Fernandez, D. U., Strompl, C., Moore, E. R. B. & Timmis, K. N. (2002). Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. *Environ Microbiol* **4**, 912-915.

73.	Stanier, R. Y. (1947). Acetic acid production from ethanol by fluorescent Pseudomonads. *J Bacteriol* **54**, 191-194.

74.	Stanier, R. Y. (1947). Simultaneous adaptation - a new technique for the study of metabolic pathways. *J Bacteriol* **54**, 339-348.

75.	Reetz, M. T., Becker, M. H., Klein, H. W. & Stöckigt, D. (1999). A method for high-throughput screening of enantioselective catalysts. *Angew Chem Int Ed Engl* **38**, 1758-1761.

76.	Reetz, M. T., Tielmann, P., Eipper, A., Ross, A. & Schlotterbeck, G. (2004). A high-throughput NMR-based ee-assay using chemical shift imaging. *Chem Commun*, 1366-1367.

77.	Kobayashi, K., Ehrlich, S. D., Albertini, A., Amati, G., Andersen, K. K., Arnaud, M., Asai, K., Ashikaga, S., Aymerich, S., Bessieres, P., Boland, F., Brignell, S. C., Bron, S., Bunai, K., Chapuis, J., Christiansen, L. C., Danchin, A., Debarbouille, M., Dervyn, E., Deuerling, E., Devine, K., Devine, S. K., Dreesen, O., Errington, J., Fillinger, S., Foster, S. J., Fujita, Y., Galizzi, A., Gardan, R., Eschevins, C., Fukushima, T., Haga, K., Harwood, C. R., Hecker, M., Hosoya, D., Hullo, M. F., Kakeshita, H., Karamata, D., Kasahara, Y., Kawamura, F., Koga, K., Koski, P., Kuwana, R., Imamura, D., Ishimaru, M., Ishikawa, S., Ishio, I., Le Coq, D., Masson, A., Mauel, C., Meima, R., Mellado, R. P., Moir, A., Moriya, S., Nagakawa, E., Nanamiya, H., Nakai, S., Nygaard, P., Ogura, M., Ohanan, T., O'Reilly, M., O'Rourke, M., Pragai, Z., Pooley, H. M., Rapoport, G., Rawlins, J. P., Rivas, L. A., Rivolta, C., Sadaie, A., Sadaie, Y., Sarvas, M., Sato, T., Saxild, H. H., Scanlan, E., Schumann, W., Seegers, J. F., Sekiguchi, J., Sekowska, A., Seror, S. J., Simon, M., Stragier, P., Studer, R., Takamatsu, H., Tanaka, T.,

Takeuchi, M., Thomaides, H. B., Vagner, V., van Dijl, J. M., Watabe, K., Wipat, A., Yamamoto, H., Yamamoto, M., Yamamoto, Y., Yamane, K., Yata, K., Yoshida, K., Yoshikawa, H., Zuber, U. & Ogasawara, N. (2003). Essential *Bacillus subtilis* genes. *Proc Natl Acad Sci U S A* **100**, 4678-4683.

78.   Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Cummings, N. J., Daniel, R. A., Denizot, F., Devine, K. M., Dusterhoft, A., Ehrlich, S. D., Emmerson, P. T., Entian, K. D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S. Y., Glaser, P., Goffeau, A., Golightly, E. J., Grandi, G., Guiseppi, G., Guy, B. J., Haga, K., Haiech, J., Harwood, C. R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M. F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., KlaerrBlanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S. M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R. P., Mizuno, M., Moestl, D., Nakai, S., Noback, M., Noone, D., Oreilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S. H., Parro, V., Pohl, T. M., Portetelle, D., Porwollik, S., Prescott, A. M., Presecan, E., Pujic, P., Purnelle, B., et al. (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature* **390**, 249-256.

79.   Westers, H., Dorenbos, R., van Dijl, J. M., Kabel, J., Flanagan, T., Devine, K. M., Jude, F., Seror, S. J., Beekman, A. C., Darmon, E., Eschevins, C., de Jong, A., Bron, S., Kuipers, O. P., Albertini, A. M., Antelmann, H., Hecker, M., Zamboni, N., Sauer, U., Bruand, C., Ehrlich, D. S., Alonso, J. C., Salas, M. & Quax, W. J. (2003). Genome engineering reveals large dispensable regions in *Bacillus subtilis*. *Mol Biol Evol* **20**, 2076-2090.

80.   Hecker, M. (2003). A proteomic view of cell physiology of *Bacillus subtilis* – bringing the genome sequence to life. In *Proteomics of microorganisms - Fundamental aspects and application* (Hecker, M. & Müllner, S., eds.), Vol. 83, pp. 57-92. Springer Verlag, Heidelberg.

81.   Mostertz, J., Scharf, C., Hecker, M. & Homuth, G. (2004). Transcriptome and proteome analysis of *Bacillus subtilis* gene expression in response to superoxide and peroxide stress. *Microbiology* **150**, 497-512.

82.   Tjalsma, H., Antelmann, H., Jongbloed, J. D., Braun, P. G., Darmon, E., Dorenbos, R., Dubois, J. Y., Westers, H., Zanen, G., Quax, W. J., Kuipers, O. P., Bron, S., Hecker, M. & van Dijl, J. M. (2004). Proteomics of protein secretion by *Bacillus subtilis*: separating the "secrets" of the secretome. *Microbiol Mol Biol Rev* **68**, 207-233.

83.   Schallmey, M., Singh, A. & Ward, O. P. (2004). Developments in the use of *Bacillus* species for industrial production. *Can. J. Microbiol.* **50**, 1-17.

84.   Bolhuis, A., Tjalsma, H., Smith, H. E., de Jong, A., Meima, R., Venema, G., Bron, S. & van Dijl, J. M. (1999). Evaluation of bottlenecks in the late stages of protein secretion in *Bacillus subtilis*. *Appl Environ Microbiol* **65**, 2934-2941.

85.   Braun, P., Gerritse, G., van Dijl, J. M. & Quax, W. J. (1999). Improving protein secretion by engineering components of the bacterial translocation machinery. *Curr Opin Biotechnol* **10**, 376-381.

86.   Blobel, G. & Sabatini, D. D. (1971). Ribosome-membrane interactions in eukaryotic cells. In *Biomembranes* (Manson, L. A., ed.), Vol. 2, pp. 193-195. Plenum Publishing Corp., New York.

87.   Blobel, G., Walter, P., Chang, C. N., Goldman, B. M., Erickson, A. H. & Lingappa, V. R. (1979). Translocation of proteins across membranes: the signal hypothesis and beyond. *Symp. Soc. Exp. Biol.* **33**, 3-36.

88.   Michaelis, S. & Beckwith, J. (1982). Mechanism of incoporation of cell envelope proteins in *Escherichia coli*. *Annu Rev Microbiol* **36**, 435-465.

89.   Berks, B. C., Palmer, T. & Sargent, F. (2005). Protein targeting by the bacterial twin-arginine translocation (Tat) pathway. *Curr Opin Microbiol* **8**, 174-181.

90.   Pugsley, A. P. (1993). The complete general secretory pathway in Gram-negative bacteria. *Microbiol Rev* **57**, 50-108.

91.   Pugsley, A. P., Francetic, O., Driessen, A. J. & de Lorenzo, V. (2004). Getting out: protein traffic in prokaryotes. *Mol Microbiol* **52**, 3-11.

92.   van Wely, K. H. M., Swaving, J., Freudl, R. & Driessen, A. J. M. (2001). Translocation of proteins across the cell envelope of Gram-positive bacteria. *FEMS Microbiol Rev* **25**, 437-454.

93.   Borchert, T. V. & Nagarajan, V. (1991). Effect of signal sequence alterations on export of levansucrase in *Bacillus subtilis*. *J Bacteriol* **173**, 276-282.

94.   Chen, M. & Nagarajan, V. (1994). Effect of alteration of charged residues at the N-termini of signal peptides on protein export in *Bacillus subtilis*. *J Bacteriol* **176**, 5796-5801.

95.   Gierasch, L. M. (1989). Signal Sequences. *Biochemistry* **28**, 923-930.

96.   von Heijne, G. (1985). Signal sequences. The limits of variation. *J Mol Biol* **184**, 99-105.

97.   von Heijne, G. (1990). The signal peptide. *J Membr Biol* **115**, 195-201.

98.   Menne, K. M. L., Hermjakob, H. & Apweiler, R. (2000). A comparison of signal sequence prediction methods using a test set of signal peptides. *Bioinformatics* **16**, 741-742.

99.   Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **340**, 783-795.

100.   Lam, K. H., Chow, K. C. & Wong, W. K. (1998). Construction of an efficient *Bacillus subtilis* system for extracellular production of heterologous proteins. *J Biotechnol* **63**, 167-177.

101.   Nagarajan, V., Albertson, H., Chen, M. & Ribbe, J. (1992). Modular expression and secretion vectors for *Bacillus subtilis*. *Gene* **114**, 121-126.

102.   Dartois, V., Coppee, J. Y., Colson, C. & Baulard, A. (1994). Genetic analysis and overexpression of lipolytic activity in *Bacillus subtilis*. *Appl Environ Microbiol* **60**, 1670-1673.

103.   Zyprian, E. & Matzura, H. (1986). Characterization of signals promoting gene expression on the *Staphylococcus aureus* plasmid pUB110 and development of a Gram-positive expression system. *DNA* **5**, 219-225.

104.   Eggert, T., Brockmeier, U., Dröge, M. J., Quax, W. J. & Jaeger, K.-E. (2003). Extracellular lipases from *Bacillus subtilis*: regulation of gene expression and enzyme activity by amino acid supply and external pH. *FEMS Microbiol Lett* **225**, 319-324.

105.   Winkler, U. K. & Stuckmann, M. (1979). Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*. *J Bacteriol* **138**, 663-670.

106.   Backert, S. & Meyer, T. F. (2006). Type IV secretion systems and their effectors in bacterial pathogenesis. *Curr Opin Microbiol* **9**, 207-217.

107.   Delepelaire, P. (2004). Type I secretion in gram-negative bacteria. *Biochim Biophys Acta* **1694**, 149-161.

108.   Desvaux, M., Hebraud, M., Henderson, I. R. & Pallen, M. J. (2006). Type III secretion: what's in a name? *Trends Microbiol* **14**, 157-160.

109.   Filloux, A. (2004). The underlying mechanisms of type II protein secretion. *Biochim Biophys Acta* **1694**, 163-179.

110.   Henderson, I. R., Navarro-Garcia, F., Desvaux, M., Fernandez, R. C. & Ala'Aldeen, D. (2004). Type V protein secretion pathway: the autotransporter story. *Microbiol Mol Biol Rev* **68**, 692-744.

111.   Sandkvist, M. (2001). Biology of type II secretion. *Mol Microbiol* **40**, 271-283.

112.   Simonen, M. & Palva, I. (1993). Protein secretion in *Bacillus* species. *Microbiol Rev* **57**, 109-137.

113.   Kontinen, V. P. & Sarvas, M. (1993). The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Mol Microbiol* **8**, 727-737.

114.   Darmon, E., Noone, D., Masson, A., Bron, S., Kuipers, O. P., Devine, K. M. & van Dijl, J. M. (2002). A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. *J Bacteriol* **184**, 5661-71.

115.   Hyyrylainen, H. L., Bolhuis, A., Darmon, E., Muukkonen, L., Koski, P., Vitikainen, M., Sarvas, M., Pragai, Z., Bron, S., van Dijl, J. M. & Kontinen, V. P. (2001). A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. *Mol Microbiol* **41**, 1159-1172.

116.   Arnold, F. H., Wintrode, P. L., Miyazaki, K. & Gershenson, A. (2001). How enzymes adapt: lessons from directed evolution. *Trends Biochem Sci* **26**, 100-106.

117.   Bornscheuer, U. T. & Pohl, M. (2001). Improved biocatalysts by directed evolution and rational protein design. *Curr Opin Chem Biol* **5**, 137-143.

118.   Reetz, M. T. (2004). Controlling the enantioselectivity of enzymes by directed evolution: practical and theoretical ramifications. *Proc Natl Acad Sci U S A* **101**, 5716-5722.

119.   Jaeger, K.-E., Eggert, T., Eipper, A. & Reetz, M. T. (2001). Directed evolution and the creation of enantioselective biocatalysts. *Appl Microbiol Biotechnol* **55**, 519-530.

120.   Kendrew, J. C., Bodo, G., Dintzis, H. M., Parrish, R. G., Wyckoff, H. & Phillips, D. C. (1958). 3-dimensional model of the myoglobin molecule obtained by X-ray analysis. *Nature* **181**, 662-666.

121.   Kendrew, J. C., Dickerson, R. E., Strandberg, B. E., Hart, R. G., Davies, D. R., Phillips, D. C. & Shore, V. C. (1960). Structure of myoglobin - 3-dimensional fourier synthesis at 2 Å resolution. *Nature* **185**, 422-427.

122.   Westbrook, J., Feng, Z. K., Chen, L., Yang, H. W. & Berman, H. M. (2003). The protein data bank and structural genomics. *Nucleic Acids Res* **31**, 489-491.

123.   Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G. & Erlich, H. (1986). Specific enzymatic amplification of DNA *in vitro* - the polymerase chain-reaction. *Cold Spring Harbor Symposia on Quantitative Biology* **51**, 263-273.

124.   Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA-polymerase. *Science* **239**, 487-491.

125.   Matsumura, M., Becktel, W. J., Levitt, M. & Matthews, B. W. (1989). Stabilization of phage-T4 lysozyme by engineered disulfide bonds. *Proc Nat Acad Sci USA* **86**, 6562-6566.

126.   Matsumura, M., Becktel, W. J. & Matthews, B. W. (1988). Hydrophobic stabilization in T4 lysozyme determined directly by multiple substitutions of Ile-3. *Nature* **334**, 406-410.

127.   Matsumura, M., Signor, G. & Matthews, B. W. (1989). Substantial increase of protein stability by multiple disulfide bonds. *Nature* **342**, 291-293.

128.   Henke, E., Bornscheuer, U. T., Schmid, R. D. & Pleiss, J. (2003). A molecular mechanism of enantiorecognition of tertiary alcohols by carboxylesterases. *ChemBioChem* **4**, 485-493.

129.   Magnusson, A., Hult, K. & Holmquist, M. (2001). Creation of an enantioselective hydrolase by engineered substrate-assisted catalysis. *J Am Chem Soc* **123**, 4354-4355.

130.   Orrenius, C., Haeffner, F., Rotticci, D., Ohrner, N., Norin, T. & Hult, K. (1998). Chiral recognition of alcohol enantiomers in acyl transfer reactions catalysed by *Candida antarctica* lipase B. *Biocatal Biotrans* **16**, 1-15.

131.   Rotticci, D., Haeffner, F., Orrenius, C., Norin, T. & Hult, K. (1998). Molecular recognition of sec-alcohol enantiomers by *Candida antarctica* lipase B. *J Mol Catal B: Enzym* **5**, 267-272.

132.   Rotticci, D., Rotticci-Mulder, J. C., Denman, S., Norin, T. & Hult, K. (2001). Improved enantioselectivity of a lipase by rational protein engineering. *ChemBioChem* **2**, 766-770.

133.   Chen-Goodspeed, M., Sogorb, M. A., Wu, F. Y., Hong, S. B. & Raushel, F. M. (2001). Structural determinants of the substrate and stereochemical specificity of phosphotriesterase. *Biochemistry* **40**, 1325-1331.

134.   Chen-Goodspeed, M., Sogorb, M. A., Wu, F. Y. & Raushel, F. M. (2001). Enhancement, relaxation, and reversal of the stereoselectivity for phosphotriesterase by rational evolution of active site residues. *Biochemistry* **40**, 1332-1339.

135.   Schöneboom, J. C., Lin, H., Reuter, N., Thiel, W., Cohen, S., Ogliaro, F. & Shaik, S. (2002). The elusive oxidant species of cytochrome P450 enzymes: Characterization by combined quantum mechanical/molecular mechanical (QM/MM) calculations. *J Am Chem Soc* **124**, 8142-8151.

136.   Rubin-Pitel, S. B. & Zhao, H. (2006). Recent advances in biocatalysis by directed enzyme evolution. *Comb Chem High Throughput Screen* **9**, 247-257.

137.   Lutz, S. & Patrick, W. M. (2004). Novel methods for directed evolution of enzymes: quality, not quantity. *Curr Opin Biotechnol* **15**, 291-297.

138.   Neylon, C. (2004). Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res* **32**, 1448-1459.

139. Wong, T. S., Zhurina, D. & Schwaneberg, U. (2006). The diversity challenge in directed protein evolution. *Comb Chem High Throughput Screen* **9**, 271-288.

140. Nguyen, A. W. & Daugherty, P. S. (2003). Production of randomly mutated plasmid libraries using mutator strains. In *Methods in Molecular Biology: Directed evolution library creation: methods and protocols* (Arnold, F. H. & Georgiou, G., eds.), Vol. 231, pp. 39-44. Humana press, Totowa, New Jersey.

141. Cadwell, R. C. & Joyce, G. F. (1992). Randomization of genes by PCR mutagenesis. *PCR Methods Appl* **2**, 28-33.

142. Cadwell, R. C. & Joyce, G. F. (1995). Mutagenic PCR. In *PCR primer: a laboratory manual* (Dieffenbach, C. H. & Dveksler, G. S., eds.), pp. 583. CSHL Press, Cold Spring Harbor.

143. Zaccolo, M., Williams, D. M., Brown, D. M. & Gherardi, E. (1996). An approach to random mutagenesis of DNA using mixtures of triphosphate derivatives of nucleoside analogues. *J Mol Biol* **255**, 589-603.

144. Zhou, Y. H., Zhang, X. P. & Ebright, R. H. (1991). Random mutagenesis of gene-sized DNA-molecules by use of PCR with Taq DNA-polymerase. *Nucleic Acids Res* **19**, 6052-6052.

145. Hughes, M. D., Nagel, D. A., Santos, A. F., Sutherland, A. J. & Hine, A. V. (2003). Removing the redundancy from randomised gene libraries. *J Mol Biol* **331**, 973-979.

146. Wong, T. S., Tee, K. L., Hauer, B. & Schwaneberg, U. (2004). Sequence saturation mutagenesis (SeSaM): a novel method for directed evolution. *Nucleic Acids Res* **32**, e26.

147. Wong, T. S., Tee, K. L., Hauer, B. & Schwaneberg, U. (2005). Sequence saturation mutagenesis with tunable mutation frequencies. *Anal Biochem* **341**, 187-189.

148. Matsuura, T., Miyai, K., Trakulnaleamsai, S., Yomo, T., Shima, Y., Miki, S., Yamamoto, K. & Urabe, I. (1999). Evolutionary molecular engineering by random elongation mutagenesis. *Nat Biotechnol* **17**, 58-61.

149. Murakami, H., Hohsaka, T. & Sisido, M. (2002). Random insertion and deletion of arbitrary number of bases for codon-based random mutation of DNAs. *Nat Biotechnol* **20**, 76-81.

150. Pikkemaat, M. G. & Janssen, D. B. (2002). Generating segmental mutations in haloalkane dehalogenase: a novel part in the directed evolution toolbox. *Nucleic Acids Res* **30**.

151. Chopra, S. & Ranganathan, A. (2003). Protein evolution by "codon shuffling": A novel method for generating highly variant mutant libraries by assembly of hexamer DNA duplexes. *Chem Biol* **10**, 917-926.

152. Osuna, J., Yanez, J., Soberon, X. & Gaytan, P. (2004). Protein evolution by codon-based random deletions. *Nucleic Acids Res* **32**, e136.

153. Yanez, J., Arguello, M., Osuna, J., Soberon, X. & Gaytan, P. (2004). Combinatorial codon-based amino acid substitutions. *Nucleic Acids Res* **32**, e158.

154. Kashiwagi, K., Isogai, Y., Nishiguchi, K. I. & Shiba, K. (2006). Frame shuffling: a novel method for *in vitro* protein evolution. *Protein Eng Des Sel* **19**, 135-140.

155. Stemmer, W. P. C. (1994). DNA-shuffling by random fragmentation and reassembly - *in vitro* recombination for molecular evolution. *Proc Nat Acad Sci USA* **91**, 10747-10751.

156. Stemmer, W. P. C. (1994). Rapid evolution of a protein *in vitro* by DNA-shuffling. *Nature* **370**, 389-391.

157. Zhao, H. & Arnold, F. H. (1997). Optimization of DNA shuffling for high fidelity recombination. *Nucleic Acids Res* **25**, 1307-1308.

158. Crameri, A., Raillard, S. A., Bermudez, E. & Stemmer, W. P. C. (1998). DNA shuffling of a family of genes from diverse species accelerates directed evolution. *Nature* **391**, 288-291.

159. Zhao, H., Giver, L., Shao, Z., Affholter, J. A. & Arnold, F. H. (1998). Molecular evolution by staggered extension process (StEP) in vitro recombination. *Nat Biotechnol* **16**, 258-261.

160. Shao, Z., Zhao, H., Giver, L. & Arnold, F. H. (1998). Random-priming *in vitro* recombination: an effective tool for directed evolution. *Nucleic Acids Res* **26**, 681-683.

161. Volkov, A. A., Shao, Z. & Arnold, F. H. (1999). Recombination and chimeragenesis by *in vitro* heteroduplex formation and *in vivo* repair. *Nucleic Acids Res* **27**, e18.

162. Kikuchi, M., Ohnishi, K. & Harayama, S. (1999). Novel family shuffling methods for the *in vitro* evolution of enzymes. *Gene* **236**, 159-167.

163.   Kikuchi, M., Ohnishi, K. & Harayama, S. (2000). An effective family shuffling method using single-stranded DNA. *Gene* **243**, 133-137.

164.   Gibbs, M. D., Nevalainen, K. M. & Bergquist, P. L. (2001). Degenerate oligonucleotide gene shuffling (DOGS): a method for enhancing the frequency of recombination with family shuffling. *Gene* **271**, 13-20.

165.   Coco, W. M., Levinson, W. E., Crist, M. J., Hektor, H. J., Darzins, A., Pienkos, P. T., Squires, C. H. & Monticello, D. J. (2001). DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat Biotechnol* **19**, 354-359.

166.   Song, J. K., Chung, B., Oh, Y. H. & Rhee, J. S. (2002). Construction of DNA-shuffled and incrementally truncated libraries by a mutagenic and unidirectional reassembly method: changing from a substrate specificity of phospholipase to that of lipase. *Appl Environ Microbiol* **68**, 6146-6151.

167.   Ness, J. E., Kim, S., Gottman, A., Pak, R., Krebber, A., Borchert, T. V., Govindarajan, S., Mundorff, E. C. & Minshull, J. (2002). Synthetic shuffling expands functional protein diversity by allowing amino acids to recombine independently. *Nat Biotechnol* **20**, 1251-1255.

168.   O'Maille, P. E., Bakhtina, M. & Tsai, M. D. (2002). Structure-based combinatorial protein engineering (SCOPE). *J Mol Biol* **321**, 677-691.

169.   Ikeuchi, A., Kawarasaki, Y., Shinbata, T. & Yamane, T. (2003). Chimeric gene library construction by a simple and highly versatile method using recombination-dependent exponential amplification. *Biotechnol Prog* **19**, 1460-1467.

170.   Zha, D. X., Eipper, A. & Reetz, M. T. (2003). Assembly of designed oligonucleotides as an efficient method for gene recombination: A new tool in directed evolution. *ChemBioChem* **4**, 34-39.

171.   Lee, S. H., Ryu, E. J., Kang, M. J., Wang, E. S., Piao, Z., Choi, Y. J., Jung, K. H., Jeon, J. Y. J. & Shin, Y. C. (2003). A new approach to directed gene evolution by recombined extension on truncated templates (RETT). *J Mol Catal B: Enzym* **26**, 119-129.

172.   Eggert, T., Funke, S. A., Rao, N. M., Acharya, P., Krumm, H., Reetz, M. T. & Jaeger, K.-E. (2005). Multiplex-PCR-based recombination as a novel high-fidelity method for directed evolution. *ChemBioChem* **6**, 1062-1067.

173.   Ostermeier, M., Nixon, A. E., Shim, J. H. & Benkovic, S. J. (1999). Combinatorial protein engineering by incremental truncation. *Proc Nat Acad Sci USA* **96**, 3562-3567.

174.   Ostermeier, M., Shim, J. H. & Benkovic, S. J. (1999). A combinatorial approach to hybrid enzymes independent of DNA homology. *Nat Biotechnol* **17**, 1205-1209.

175.   Sieber, V., Martinez, C. A. & Arnold, F. H. (2001). Libraries of hybrid proteins from distantly related sequences. *Nat Biotechnol* **19**, 456-460.

176.   Kawarasaki, Y., Griswold, K. E., Stevenson, J. D., Selzer, T., Benkovic, S. J., Iverson, B. L. & Georgiou, G. (2003). Enhanced crossover SCRATCHY: construction and high-throughput screening of a combinatorial library containing multiple non-homologous crossovers. *Nucleic Acids Res* **31**, e126.

177.   Lutz, S., Ostermeier, M., Moore, G. L., Maranas, C. D. & Benkovic, S. J. (2001). Creating multiple-crossover DNA libraries independent of sequence identity. *Proc Nat Acad Sci USA* **98**, 11248-11253.

178.   Kolkman, J. A. & Stemmer, W. P. C. (2001). Directed evolution of proteins by exon shuffling. *Nat Biotechnol* **19**, 423-428.

179.   Tsuji, T., Onimaru, M. & Yanagawa, H. (2001). Random multi-recombinant PCR for the construction of combinatorial protein libraries. *Nucleic Acids Res* **29**, e97.

180.   Hiraga, K. & Arnold, F. H. (2003). General method for sequence-independent site-directed chimeragenesis. *J Mol Biol* **330**, 287-296.

181.   Arnold, F. H. (1996). Directed evolution: creating biocatalysts for the future. *Chem Eng Sci* **51**, 5091-5102.

182.   Wong, T. S., Roccatano, D., Zacharias, M. & Schwaneberg, U. (2006). A statistical analysis of random mutagenesis methods used for directed protein evolution. *J Mol Biol* **355**, 858-871.

183.   Barettino, D., Feigenbutz, M., Valcarcel, R. & Stunnenberg, H. G. (1994). Improved method for PCR-mediated site-directed mutagenesis. *Nucleic Acids Res* **22**, 541-542.

184. Gray, K. A., Richardson, T. H., Kretz, K., Short, J. M., Bartnek, F., Knowles, R., Kan, L., Swanson, P. E. & Robertson, D. E. (2001). Rapid evolution of reversible denaturation and elevated melting temperature in a microbial haloalkane dehalogenase. *Adv Synth Catal* **343**, 607-617.

185. DeSantis, G., Wong, K., Farwell, B., Chatman, K., Zhu, Z. L., Tomlinson, G., Huang, H. J., Tan, X. Q., Bibbs, L., Chen, P., Kretz, K. & Burk, M. J. (2003). Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). *J Am Chem Soc* **125**, 11476-11477.

186. Short, J. M. (2001). Saturation mutagenesis in directed evolution. Diversa Corporation US Patent 6,171,820.

187. Reetz, M. T., Bocola, M., Carballeira, J. D., Zha, D. & Vogel, A. (2005). Expanding the range of substrate acceptance of enzymes: combinatorial active-site saturation test. *Angew Chem Int Ed Engl* **44**, 4192-4196.

188. Silberg, J. J., Endelman, J. B. & Arnold, F. H. (2004). SCHEMA-guided protein recombination. *Protein Eng* **388**, 35-42.

189. Voigt, C. A., Martinez, C., Wang, Z. G., Mayo, S. L. & Arnold, F. H. (2002). Protein building blocks preserved by recombination. *Nat Struct Biol* **9**, 553-558.

190. Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Hugejensen, B., Norskov, L., Thim, L. & Menge, U. (1990). A serine protease triad forms the catalytic center of a triacylglycerol lipase. *Nature* **343**, 767-770.

191. Winkler, F. K., Darcy, A. & Hunziker, W. (1990). Structure of human pancreatic lipase. *Nature* **343**, 771-774.

192. Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Hugejensen, B., Patkar, S. A. & Thim, L. (1991). A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex. *Nature* **351**, 491-494.

193. Nardini, M., Lang, D. A., Liebeton, K., Jaeger, K.-E. & Dijkstra, B. W. (2000). Crystal structure of *Pseudomonas aeruginosa* lipase in the open conformation - The prototype for family I.1 of bacterial lipases. *J Biol Chem* **275**, 31219-31225.

194. Carriere, F., Withers-Martinez, C., van Tilberugh, H., Roussel, A., Cambillau, C. & Verger, R. (1998). Structural basis for the substrate selectivity of pancreatic lipases and some related proteins. *Biochim Biophys Acta Rev Biomem* **1376**, 417-432.

195. Chahinian, H., Bezzine, S., Ferrato, F., Ivanova, M. G., Perez, B., Lowe, M. E. & Carriere, F. (2002). The beta 5 ' loop of the pancreatic lipase C2-like domain plays a critical role in the lipase-lipid interactions. *Biochemistry* **41**, 13725-13735.

196. Liebeton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B. W., Reetz, M. T. & Jaeger, K.-E. (2000). Directed evolution of an enantioselective lipase. *Chem Biol* **7**, 709-718.

197. Fishman, A. & Cogan, U. (2003). Bio-imprinting of lipases with fatty acids. *J Mol Catal B: Enzym* **22**, 193-202.

198. Mingarro, I., Abad, C. & Braco, L. (1995). Interfacial activation-based molecular bioimprinting of lipolytic enzymes. *Proc Nat Acad Sci USA* **92**, 3308-3312.

199. Martinez, C., Degeus, P., Lauwereys, M., Matthyssens, G. & Cambillau, C. (1992). *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. *Nature* **356**, 615-618.

200. van Pouderoyen, G., Eggert, T., Jaeger, K.-E. & Dijkstra, B. W. (2001). The crystal structure of *Bacillus subtilis* lipase: A minimal alpha/beta hydrolase fold enzyme. *J Mol Biol* **309**, 215-226.

201. Rosenau, F. & Jaeger, K.-E. (2000). Bacterial lipases from *Pseudomonas*: Regulation of gene expression and mechanisms of secretion. *Biochimie* **82**, 1023-1032.

202. Rosenau, F., Tommassen, J. & Jaeger, K.-E. (2004). Lipase-specific foldases. *ChemBioChem* **5**, 153-161.

203. Urban, A., Leipelt, M., Eggert, T. & Jaeger, K.-E. (2001). DsbA and DsbC affect extracellular enzyme formation in *Pseudomonas aeruginosa*. *J Bacteriol* **183**, 587-596.

Chapter 9:   *You´ll never walk alone*

<div align="right">Oscar Hammerstein II</div>

Acknowledgements

I would like to thank Prof. Dr. Karl-Erich Jaeger for many stimulating discussions on my work. He provided perfect "growing conditions" in his institute not only for *Escherichia coli* and *Bacillus subtilis*, but also for young research groups like the "directed evolution" group. Many thanks for the freedom to realize my research in the way I wanted to do.

Furthermore, I would like to acknowledge all members of the directed evolution group. Some of them are still working in the group, whereas the others have already left the group and work in other scientific institutes or in industry. Thanks to Dipl. Biol. Andexer, M.Sc. Bogo, M.Sc. Brissos, M.Sc. Casiraghi, Dr. Bertram, Dr. Brockmeier, Dr. Bustos-Jaimes, Dipl. Biol. Degering, Dr. Funke, Dipl. Biol. Franken, Dipl. Biol. Guterl, Dr. Konarzycka-Bessler, Dipl. Nat. Krauß, Ms. Nukicic, Ms. Piqueray, Dipl. Nat. Puls, Ms. Richter, Dipl. Ing. Rosenbaum, M.Sc. Sibilla, Dipl. Biol. Spinrath, Dipl. Biol. Wendorff and Ms. Wirtz for your teamwork and for your enthusiastic work in the lab.

I would like to thank all members of the Institute of Molecular Enzyme Technology (IMET) for the good collaboration during my stay in the institute. In particular, I would like to thank the groups headed by Dr. Leggewie, PD Dr. Pohl and Prof. Dr. Hummel for fruitful co-operations as documented by joint publications. Furthermore, thanks to Dipl. Ing. Gieren and Dipl. Ing. Goldbaum, you both did and do an excellent job in high-throughput screening by keeping our picking and pipetting robots constantly running. And, of course, special thanks go to Ms. Jansen for uncountable assistance in overcoming administrative obstacles.

My interdisciplinary work in the field of enzyme technology at the IMET was possible only because of many co-operations within the scientific community. Therefore, I would like to thank especially the following people for close collaboration: Prof. Dr. Freudl, Dr. Hubbuch, Dr. Lütz, Prof. Dr. Müller (now University of Freiburg), Prof. Dr. Noll (now University of Bielefeld), Prof. Dr. Sahm, Prof. Dr. Wandrey at the Institute of Biotechnology (Forschungszentrum Jülich GmbH); Prof. Dr. Pietruszka at the neighbouring Institute of Bioorganic Chemistry (Heinrich-Heine University Düsseldorf); Prof. Dr. Reetz, Prof. Dr. Thiel, and their co-workers Dr. Bocola, Dr. Krumm, Dr. Otte from the Max-Planck Institut für Kohlenforschung (Mülheim a.d. Ruhr); Prof. Dr. Liese at the Institute of Technical Biocatalysis (Technical University of Hamburg-Harburg), Prof. Dr. Büchs and Dipl. Ing.

Kinsey (RWTH Aachen), Prof. Dr. Kragl and Dipl. Chem. Schumacher (University of Rostock). I would like to thank my co-operation partners at the Rheinisch Westfälische Technische Hochschule (RWTH) Aachen within the Graduiertenkolleg 1166 "BioNoCo" Dr. Ansorge-Schumacher, Prof. Dr. Büchs, Dr. Greiner, Prof. Dr. Hartmeier, Prof. Dr. Leitner and Dr. Spieß.

Furthermore, I would like to thank my international cooperation partners Prof. Dr. Dijkstra, Prof. Dr. Quax and their co-workers Dr. Boersma, Dr. Dröge, Dipl. Ing. Pijning and Dr. van Pouderoyen (University of Groningen, The Netherlands), Prof. Dr. Carrière (CNRS Marseille, France), Dr. Secundo (CNR Milano, Italy) and Dr. Rao (Centre for Cellular and Molecular Biology Hyderabad, India).

Finally, special thanks go to my family for their continuous support and encouragement always to look forward.

Chapter 10:

Appendix

Publications of **Chapter 4 "Biocatalyst identification by high-throughput screening and selection systems"**:

A high-throughput screening assay for hydroxynitrile lyase activity

Jennifer Andexer, Jan-Karl Guterl, Martina Pohl and Thorsten Eggert*

A high-throughput screening assay for hydroxynitrile lyase activity accepting a wide range of HNL-substrates is presented, which is useful either for enzyme fingerprinting or screening of huge variant libraries generated in metagenome or directed evolution approaches.

Hydroxynitrile lyases (HNLs) naturally occur in plants integrated in microbial and herbivore defense mechanisms by HCN-release due to cyanohydrin cleavage. Some HNLs are biochemically characterized in detail concerning substrate specificity and enantioselectivity. The HNLs from *Prunus* species, like *P. amygdalus* (bitter almond)[1a] and *P. mume* (Japanese apricot)[1b] as well as the HNL from *Linum usitatissimum* (flax)[1c] show (*R*)-selectivity, whereas the other HNLs from *Sorghum bicolor* (millet),[1d] *Hevea brasiliensis* (para rubber tree)[1e] and *Manihot esculenta* (cassava)[1f] show (*S*)-selectivity towards different cyano-hydrins. Since the reverse reaction is also catalyzed by HNLs, these enzymes are valuable catalysts for the synthesis of cyanohydrins, which are versatile chiral building blocks in the pharmaceutical and agrochemical industries.[1] Apart from their availability, the application of HNLs is often restricted by their substrate range and low stability under technical conditions.

Novel or improved HNLs can be found by screening plants for appropriate enzymes,[2] by directed evolution, rational design, or by metagenomic approaches.[3] However, one major necessity for all these strategies is a simple and powerful high-throughput screening (HTS) assay to identify potential novel or better performing HNLs. So far HNL-activity has usually been determined by GC- or HPLC-analysis,[2a] which is not practical in high-throughput. In addition, a spectrophotometric assay based on HCN-detection using the well-known König reaction[4,5] has been described. However, this assay was applied in a total reaction volume of 10 mL which is not suitable for high-throughput screening approaches. Furthermore, only acetone cyanohydrin has been used so far as a substrate.[6] Another spectrophotometric assay for activity towards benzaldehyde cyanohydrin is available, detecting the increase in absorption of the released benzaldehyde at 280 nm wavelength.[7] Although this assay is applicable for high-throughput, it is restricted to aromatic substrates only and requires microtiter plates (MTPs) which are transparent in the UV. Furthermore, recently a colony assay, based on NADH-fluorescence has been developed to detect HNL-activity towards benzaldehyde cyanohydrin.[8]

Here, we describe an HNL assay in MTP format (200 μL), which allows screening in high-throughput using automated pipetting workstations. The broad applicability of the assay is demonstrated using different aliphatic and aromatic cyanohydrins (Fig. 1); two of them were achiral, four of them were used as a racemic mixture and furthermore benzaldehyde cyanohydrin was applied as (*R*)- and (*S*)-enantiomers to detect the enantioselectivity of recombinant HNL from *Manihot esculenta* (MeHNL) which was used as a model.

The assay consists of two parts, first the biotransformation step yielding HCN by cyanohydrin cleavage; subsequently HCN is detected spectrophotometrically at 600 nm wavelength in MTPs (Fig. 1). The crucial parameter in the cyanohydrin cleavage reaction is the pH, because many cyanohydrins decompose at pH > 6.0, whereas on the other hand the enzyme activity and stability
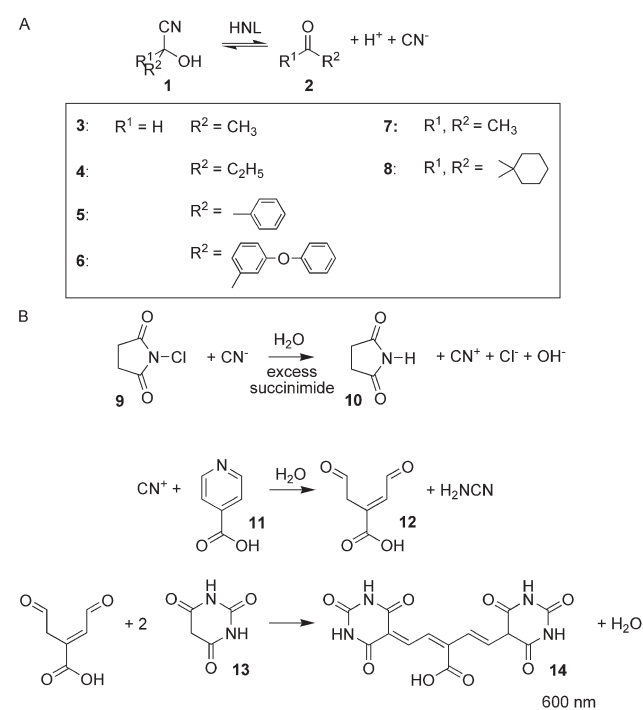

Fig. 1 Schematic overview of the assay system. A: Biotransformation step. The cyanohydrin **1** is enzymatically converted to a carbonyl compound **2** and HCN. Six different cyanohydrins (**3**: acetaldehyde cyanohydrin, **4**: propionaldehyde cyanohydrin, **5**: benzaldehyde cyanohydrin, **6**: 3-phenoxybenzaldehyde cyanohydrin, **7**: acetone cyanohydrin, **8**: cyclohexanone cyanohydrin) were tested with MeHNL. B: Cyanide-determination step (modified according to Markley *et al.*[12]): cyanide anions are oxidized by *N*-chlorosuccinimide **9** (stabilized with succinimide **10**) to cyanide cations, which react with isonicotinic acid **11** forming a dialdehyde **12**, which is coupled to two molecules of barbituric acid **13** to form the dye **14** which is measured spectrophotometrically at 600 nm.

Institute of Molecular Enzyme Technology, Heinrich-Heine University Düsseldorf, Forschungszentrum Jülich, 52426, Jülich, Germany. E-mail: t.eggert@fz-juelich.de; Fax: +49 (2461) 612490; Tel: +49 (2461) 612939

are significantly impaired at pH < 5.0. Therefore, assaying cyanohydrin cleavage at pH 5.0 to 5.5 is a good compromise.[6,7]

The assay is sufficiently sensitive to screen HNL-libraries using crude cell extracts. For the purpose of library screening, first the enzymatic reaction using *E. coli* crude cell extracts containing over-expressed MeHNL is performed, thereby a certain amount of cyanide is liberated from the cyanohydrin substrate. Therefore, 140 μL citrate–phosphate buffer pH 5.0, 10 μL of HNL containing crude cell extracts and 10 μL cyanohydrin solution (final concentration 15 mM) are mixed and incubated at room temperature for 5 min. By addition of 10 μL of mix I (*N*-chlorosuccinimide **9**/succinimide **10**) the biotransformation step is stopped,[9] thereby oxidizing the liberated CN^- to CN^+. After 2 min the colorimetric detection step is started by adding 30 μL of mix II (isonicotinic acid **11**/barbituric acid **13**).[10] Subsequently, the rate of color formation is measured spectrophotometrically over 20 min at 600 nm using a microtiter plate reader. The dye **14** is stable for at least 2 hours. Barbituric acid **13** is applied instead of the alternatively used dye compound 3-methyl-1-phenyl-5-pyrazolone in HCN-detection,[11,12] because the latter is unsuitable at pH-values below 7.[13] Isonicotinic acid is a well suited alternative to the widely used pyridine.[6]

Instead of measuring the spectrophotometric properties of the resulting aldehyde or ketone, the major advantage of this assay is the possibility of analyzing HNL-activity towards virtually any cyanohydrin by detecting the liberated HCN. This makes the assay suitable for a vast substrate screening as well as for detection of new or improved activities in enzyme libraries obtained by rational design or directed evolution. We have used our HTS-assay to determine MeHNL-activity towards six different aromatic and aliphatic cyanohydrins. All substrates were converted by MeHNL with **5** and **7** being the best substrates. The high selectivity of MeHNL towards the (*S*)-enantiomer of benzaldehyde cyanohydrin **5** is obvious when (*R*)-**5** and (*S*)-**5** are used separately in the assay (Fig. 2).

Furthermore, the assay allows calculation of specific enzymatic activity, since color development in the HCN-detection step is proportional to the amount of cyanide in the solution. Time dependent cyanohydrin conversion was calculated based on a
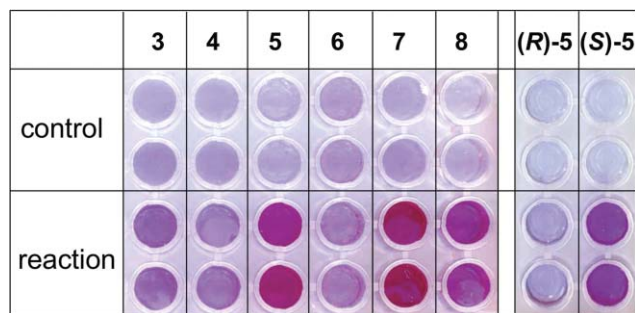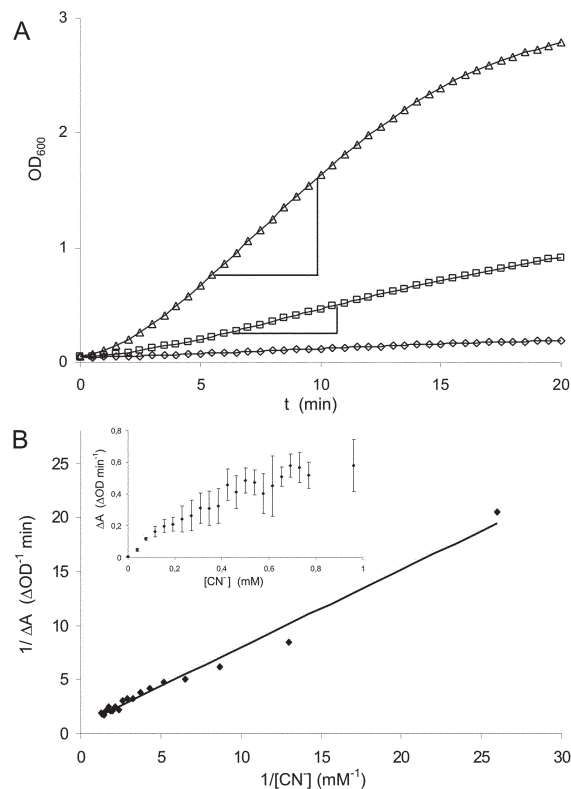

Fig. 3 Spectrophotometric detection of acetone cyanohydrin **7** cleavage using different amounts of purified MeHNL (–□– 50 ng, –△– 250 ng). A: The increase in absorbance at 600 nm over 20 min is shown. The amount of liberated cyanide is calculated from the linear part of the curve. Autolysis of the respective cyanohydrin is detected in a control without enzyme (–◇–) and subtracted from the slope values of the samples. B: Hyperbolic cyanide calibration curve and linearization by double reciprocal presentation.

cyanide standard curve ($K_2[Zn(CN)_4]$) (Fig. 3A) by correlating the rate of color formation at 600 nm with the cyanide concentration. For this purpose the hyperbolic standard curve was linearized in a double reciprocal diagram (Fig. 3B). For purified MeHNL[14] the calculated specific activity for acetone cyanohydrin **7** was 130 ± 30 U/mg, which is consistent with data from the literature, giving values between 92 U/mg and 260 U/mg[15] depending on the assay conditions. Comparison of the specific activities towards different substrates in Table 1 clearly demonstrates the highest catalytic activity of MeHNL towards the natural substrate acetone cyanohydrin. However, it must be taken into account that substrates **3**–**6** were applied as racemic mixtures containing 50% of the non-favored enantiomer, whereas substrates **7** and **8** are achiral.

Fig. 2 Spectrophotometric detection of hydroxynitrile lyase activity. Microtiter plate with different substrates **3**–**8** (see Fig. 1). Control: autolysis of the respective cyanohydrin (without enzyme). For reactions 10 μL of *E. coli* crude cell extracts containing over-expressed MeHNL were used. Faint blue to purple color represents an increasing amount of cyanide. The application of enantiomerically pure substrates can be used to estimate the enantioselectivity of the biocatalyst as demonstrated in the case of enantiomerically pure (*R*)- and (*S*)-benzaldehyde cyanohydrin **5**.

Table 1 Results of cyanohydrin cleavage catalyzed by MeHNL. Substrates **3**–**8** (15 mM, see Fig. 1) were incubated with purified MeHNL[14]

Substrate	Specific activity [U/mg]
3	1.3 (±0.4)
4	0.4 (±0.2)
5	19.1 (±4.9)
6	0.1 (±0.04)
7	130.0 (±30.0)
8	1.0 (±0.4)

In summary, a novel HCN-based high-throughput screening assay for HNL activity was developed. The assay is useful to detect activity and enantioselectivity of HNLs theoretically towards any cyanohydrin substrate. Limitations might occur in the case of hydrophobic substrates due to poor water solubility. This problem can be overcome by the use of emulsifying agents like gum arabic. As tested, the increased turbidity has no influence on the formation and spectrophotometric detection of the dye (data not shown). Therefore, the assay is useful for both preparing enzyme fingerprints and screening large variant libraries generated in metagenome or directed evolution approaches. The assay is highly sensitive; at least 5 ng of purified MeHNL representing 1 mU of enzyme activity was reliably detectable in the assay. Furthermore, the assay is robust and easy to handle without the necessity of expensive equipment; however, it is possible to automate the test by using pipetting robots in order to increase the sample throughput.

Notes and references

1 (a) A. Glieder, R. Weis, W. Skranc, P. Poechlauer, I. Dreveny, S. Majer, M. Wubbolts, H. Schwab and K. Gruber, *Angew. Chem., Int. Ed.*, 2003, **42**, 4815–4818; (b) S. Nanda, Y. Kato and Y. Asano, *Tetrahedron*, 2005, **61**, 10908–10916; (c) H. Breithaupt, M. Pohl, W. Bönigk, P. Heim, K. L. Schimz and M. R. Kula, *J. Mol. Catal. B: Enzym.*, 1999, **6**, 315–332; (d) H. Lauble, B. Miehlich, S. Förster, H. Wajant and F. Effenberger, *Biochemistry*, 2002, **41**, 12043–12050; (e) M. Hasslacher, C. Kratky, H. Griengl, H. Schwab and S. D. Kohlwein, *Proteins: Struct., Funct., Genet.*, 1997, **27**, 438–449; (f) H. Bühler, F. Effenberger, S. Förster, J. Roos and H. Wajant, *ChemBioChem*, 2003, **4**, 211–216; (g) H. Griengl, H. Schwab and M. Fechter, *Trends Biotechnol.*, 2000, **18**, 252–256; (h) M. Sharma, N. N. Sharma and T. C. Bhalla, *Enzyme Microb. Technol.*, 2005, **37**, 279–294; (i) F. Effenberger, S. Förster and H. Wajant, *Curr. Opin. Biotechnol.*, 2000, **11**, 532–539.
2 (a) Y. Asano, K. Tamura, N. Doi, T. Ueatrongchit, A. H.-Kittikun and T. Ohmiya, *Biosci., Biotechnol., Biochem.*, 2005, **69**, 2349–2357; (b) A. Hickel, G. Heinrich, H. Schwab and H. Griengl, *Biotechnol. Tech.*, 1997, **11**, 55–58.
3 (a) K.-E. Jaeger, T. Eggert, A. Eipper and M. T. Reetz, *Appl. Microbiol. Biotechnol.*, 2001, **55**, 519–530; (b) P. Berglund and S. Park, *Curr. Org. Chem.*, 2005, **9**, 325–336; (c) P. Lorenz and J. Eck, *Nat. Rev. Microbiol.*, 2005, **3**, 510–516.
4 W. König, *J. Prakt. Chem.*, 1904, **69**, 105–137.
5 A. J. A. Essers, *Acta Hortic.*, 1994, **375**, 97–104.
6 D. Selmar, F. J. Carvalho and E. E. Conn, *Anal. Biochem.*, 1987, **166**, 208–211.
7 M. Bauer, H. Griengl and W. Steiner, *Biotechnol. Bioeng.*, 1999, **62**, 20–29.
8 C. Reisinger, F. van Assema, M. Schuermann, Z. Hussain, P. Remler and H. Schwab, *J. Mol. Catal. B: Enzym.*, 2006, **39**, 149–155.
9 (a) M. A. Lischwe and M. T. Sung, *J. Biol. Chem.*, 1977, **252**, 4976–4980; (b) Incubation of MeHNL with 5mM *N*-chlorosuccinimide shows complete inactivation after 10 min.
10 Assay solutions: citrate–phosphate buffer: 24.3 mL 0.1 M citric acid, 25.7 mL 0.2 M K_2HPO_4 and 100 mL H_2O; Substrate solution: 300 mM cyanohydrin in 0.1 M citric acid. Insoluble substrates were emulsified by adding 20 µg/mL gum arabic; Mix I: 100 mM *N*-chlorosuccinimide with 10-fold excess of succinimide (w/w); Mix II: 65 mM isonicotinic acid, 125 mM barbituric acid in 0.2 M NaOH.
11 J. Epstein, *Anal. Chem.*, 1947, **19**, 272–274.
12 B. Markley, C. E. Meloan, J. L. Lambert and Y. C. Chiang, *Anal. Lett.*, 1987, **20**, 1225–1236.
13 J. H. Bradbury, M. G. Bradbury and S. V. Egan, *Acta Hortic.*, 1994, **375**, 87–96.
14 Enriched MeHNL solution (provided by Julich Chiral Solutions GmbH) was further purified by anion exchange chromatography on a Q-sepharose FF column at pH 5.7 using a modified protocol from S. Förster, J. Roos, F. Effenberger, H. Wajant and A. Sprauer, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 437–438.
15 (a) H. Wajant, S. Förster, H. Böttinger, F. Effenberger and K. Pfizenmaier, *Plant Sci.*, 1995, **108**, 1–11; (b) J. Hughes, J. H. Lakey and M. A. Hughes, *Biotechnol. Bioeng.*, 1997, **53**, 332–338.

A High-Throughput-Screening Method for the Enantiomeric Excess of Chiral Alcohols and its Application in Molecular Evolution[**]

Ismael Bustos-Jaimes, Thorsten Eggert, Eliane Bogo, Michael Puls, Andrea Weckbecker, Werner Hummel, Karl-Erich Jaeger[*]

Chiral alcohols are valuable intermediates in the synthesis of pharmaceutical, agricultural and fine chemicals.[1] They could be produced either by hydrocarbon oxidation, ketone reduction or ester hydrolysis. Nevertheless, these reactions usually produce non-enantiopure compounds and for this reason several methods for the enantioselective synthesis of alcohols have been developed, ranging from the synthesis of catalysts by combinatorial chemistry to the in vitro directed evolution of enzymes.[2] In any case high-throughput methods need to be applied in order to measure the enantiomeric excess (*ee*) or enantiopurity of the produced alcohols of a large number of samples. Several chemical methods for high-throughput screening of *ee* have been reported, including electrospray ionization coupled to mass spectrometry, HPLC coupled to circular dichroism, FTIR spectroscopy[3] and enzymatic methods.[4] Some of these sophisticated methods require, however, isotopically labelled *pseudo*-enantiomers for the assay and occasionally expensive equipment. Herein, we report a new colorimetric method for the evaluation of the *ee* value of alcohols based on enantioselective alcohol dehydrogenases (ADHs) coupled to a NAD(P)H oxidase (diaphorase) and its successful application in directed evolution for the screening of mutant libraries of lipases for enantioselective ester hydrolysis.

The assay is based on the enantioselective oxidation of alcohols by two different ADHs. The (*R*)-specific ADH from *Lactobacillus kefir* (LKADH) and the (*S*)-specific ADH from *Rhodococcus erythropolis* (READH), whose enantioselectivities and value as catalysts have been previously reported.[5] The oxidation of either (*R*)-**1** or (*S*)-**1** produces NAD(P)H, which is then again oxidized to NAD(P) by diaphorase from *Clostridium kluyveri* with the concomitant reduction of 2-(4-iodophenyl)-3-(4-nitrophenyl)-5-phenyl-2H-tetrazolium (INT) (**3**) to its corresponding formazan red-violet dye (**4**) (Scheme 1). The formation of this dye can be easily

---

[*] Dr. I. Bustos-Jaimes, Dr. T. Eggert, E. Bogo, M. Puls, Dr. A. Weckbecker, Prof. W. Hummel, Prof. K.-E. Jaeger
Institut für Molekulare Enzymtechnologie
Heinrich-Heine Universität Düsseldorf
Forschungzentrum Jülich, D-52425 Jülich, Germany
Fax: +49-2461-61-2490
E-mail: ((Correspondence Author))

Dr. I. Bustos-Jaimes
Departamento de Bioquímica
Fac. de Medicina, Universidad Nacional Autónoma de México
P.O. Box 70-159, C.U., México DF 04510, México

followed at 492 nm. The reaction is carried out within a time span of five minutes, during which the slope of colour development over time is linear. The regeneration of the oxidized form of the coenzyme also ensures the high reaction rates of ADHs. As in any other coupled assay, the amount of diaphorase, the coupling enzyme, was kept in excess relative to the ADH enzymes in order to follow first order kinetics.[6]
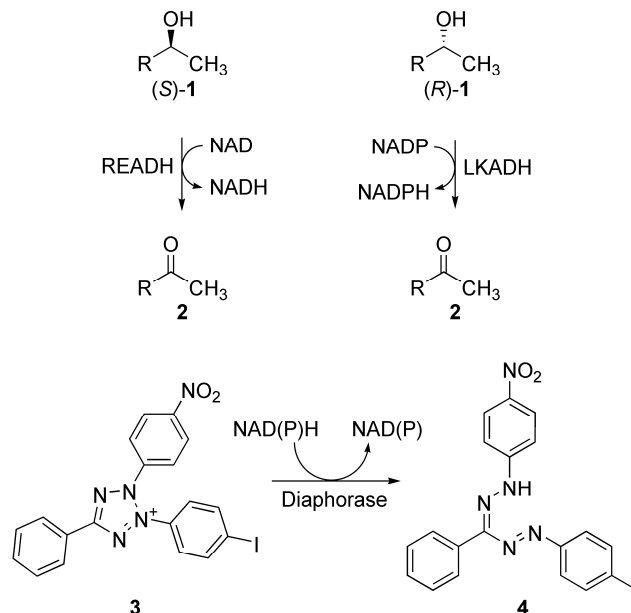
Scheme 1. Enantioselective reactions catalysed by READH and LKADH and its coupling to diaphorase redox system. Oxidation of either alcohol (*S*)-**1** or (*R*)-**1** correspondingly produces a molecule of coenzyme, NADH or NADPH, which in turn are oxidised by diaphorase through the reduction of **3** to produce the corresponding red formazan dye (**4**).

Previous studies using one or two enzymes to measure *ee* have been reported and their accuracy was well demonstrated. The mathematical framework for the analysis of these data has been developed based on the Michaelis-Menten equation.[4] Nevertheless, enzymes may display different kinetic behaviours in the presence of diverse compounds.[7] In our study, neither LKADH nor READH display hyperbolic initial-rate curves. Instead, both enzymes displayed sigmoidally shaped kinetic curves (Figure 1). This behaviour was also found when the activity was followed by measuring the formation of NAD(P)H at 340 nm. The origin of this positive cooperativity is unknown and deserves further research. Nevertheless it is possible to fit the experimental data to Equation (1), which is a form of the Michaelis-Menten equation including a term for the possible inhibitory effect of the non substrate enantiomer. In this equation we also included the Hill coefficient as a correction for the displayed positive cooperativity. With the estimated parameters is possible to calculate the concentrations of each enantiomer from the experimental initial rate for any sample.

$$v_0 = \frac{V_{max}[S]^h}{\left(K_m\left(1 + \dfrac{[I]}{K_I}\right)\right)^h + [S]^h} \qquad \text{Eq. (1)}$$

In Equation (1) v_0 is the initial rate, V_{max} is the maximum rate, [*S*] is the concentration of the substrate enantiomer, K_m is the Michaelis-Menten constant, [*I*] is the concentration of the non-substrate enantiomer, K_I is the competitive inhibition constant for *I* and *h* is the Hill coefficient. All these values rely on temperature, media and the nature of the substrate.
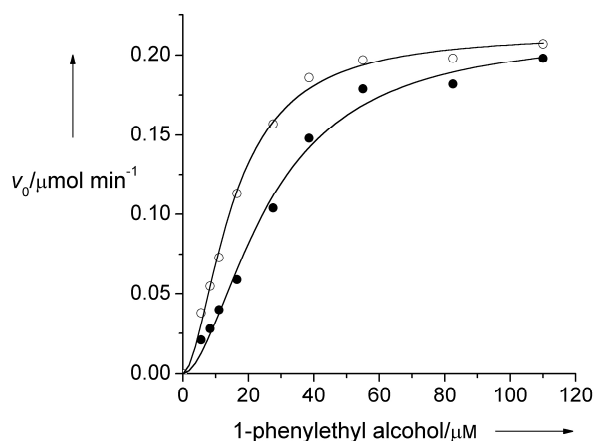

Figure 1. Kinetic behaviour of LKAHD (full dots) and READH (empty dots) for the (*R*) and (*S*) enantiomers of 1-phenylethyl alcohol (**5**), in that order, coupled to diaphorase at pH 7.0 and 30°C.

The addition of the coupling enzyme to measure the reduced coenzymes, either NADH or NADPH, increased the sensitivity of the previously reported enzymatic methods.[4] This is because λ_{max} (ε)=490 for compound 4 is 2.4 fold higher than λ_{max} (ε)=340 for NAD(P)H. Moreover, this wavelength is away from the region in which aromatic compounds and proteins absorb light and produce interferences in the measurement of the reduced coenzymes. Under the conditions reported here we were able to reliably measure alcohol concentrations as low as 25 μM, keeping the error close to 10%.

As in other enzymatic methods, a negative aspect is that samples coming from culture media have additional compounds that might interfere with the reactions of ADHs or diaphorase, varying thus the colour development. The intensity of the interferences strongly depends on the composition of the media. However, this negative aspect can be circumvented by introducing the adequate background controls.

The nature of the alcohol to be tested is also important for the success of the method. We assayed different alcohols displaying different structural properties (Scheme 2) and found that their structures are critical for the oxidation by the ADHs. Whereas 1-methyl alcohols (**5-11**) were good substrates for LKADH and READH, 1-ethyl alcohols (**12** and **13**) performed poorly as substrates and alcohol **14**, whose chain mobility is restricted, was not oxidised at all (Table 1). Nevertheless, is possible to apply this method in other alcohols by using different ADHs displaying activity toward the desired alcohols.

The method was tested using alcohol **5** as a model substrate. This compound was produced from the hydrolysis of the corresponding racemic acetate by the *Bacillus subtilis* lipase LipA (BSLA).[8] The wild-type form of this enzyme is able to hydrolyse the acetic ester of **5** producing the corresponding alcohol with a very high *ee* towards the (*R*)-enantiomer. While the *ee* measured by gas chromatography on chiral stationary phase is 92%, the *ee* value for

the same sample by our enzymatic method was 95%. This indicates that the method is quite suitable for high-throughput screening, even if it showed deviations at very high and very low substrate concentrations. These deviations occur due to 1) the inherent error generated by the ADHs behaviour at high substrate concentrations, where high variations in substrate produces only slight variations in initial rates, 2) the error associated to very low substrate concentrations, which then produces very low signals, and 3) the effect of the non substrate enantiomer, which almost certainly will act as a competitive inhibitor of the enzyme. Nevertheless these are not unsolvable problems. First a high-throughput method does not have to be analytical. Second, changing the reaction conditions in which the alcohol is produced can control the amount of alcohol present in the sample. Finally, the negative effect of the non-substrate enantiomer is taken into account by the corresponding competitive-inhibition term in the equation 1. Of course, a minimal kinetic characterization, or any other technique, should be used to measure the dissociation constants for each enantiomer to each ADH.
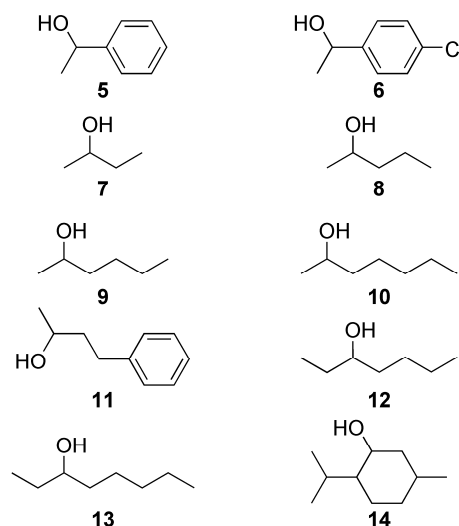
Scheme 2. Racemic alcohols tested for oxidation with READH and LKADH. **5**, 1-phenylethyl alcohol; **6**, 1-(4-chlorophenyl)ethanol; **7**, 2-butanol; **8**, 2-pentanol; **9**, 2-hexanol; **10**, 2-heptanol; **11**, 4-phenyl-2-butanol; **12**, 3-heptanol; **13**, 3-octanol; **14**, menthol.

A very important consideration for the development of this method was its applicability to a high-throughput format. In our conditions it was possible to analyse a 192 samples in less than 20 minutes, it is about 14000 samples per day. Thus this screening technique provides the possibility to perform a primary selection among a large collection of samples in short times, dispaying a very high sensitivity. The next step was the application of this new method for the selection of catalysts producing alcohols showing different *ee* values.

In the practice we used this assay to screen site-saturated mutant libraries of BSLA.[9] During the screening procedure, we were able to detect BSLA mutants with modified enantioselectivity in reference to the wild-type enzyme. We screened about 2000 variants and, as expected, the wild-type activity and high *ee* value (92%) was negatively affected in all the variants. For example the mutant EBAn showed activity against the acetic-ester of **5** and its estimated *ee* value by this screening method was 5% towards the *R*-alcohol, while by gas chromatography on chiral stationary phase it was 1%

towards the same alcohol. This demonstrates that the method is highly suitable for measuring the *ee* value of an alcohol mixture. Additionally, the same screening procedure was used to screen a mutant library produced by error-prone PCR of a BSLA variant which previously was rationally modified by inserting an artificial lid-motif.[10] In this approach, we were able to identify the mutant MPB7, whose *ee* value is 70% towards (*R*)-**5**. This *ee* value is not very impressive by itself if we take into account that the wild-type BSLA is even more enantioselective; nonetheless, this mutant was selected because it is about 115-fold and 14-fold more active than its parental form towards the (*R*)- and (*S*)-esters, respectively. Therefore this method enables the selection of the most active and the most selective catalyst in one step.

Table 1. Relative activities of LKADH and READH toward different alcohols.

Alcohol	LKADH Relative activity / %[a]	READH Relative activity / %[a]
5	100	100
6	97	96
7	88	35
8	97	48
9	110	96
10	100	110
11	97	100
12	64	10
13	48	39
14	0	0

[a] Relative to substrate **5**

The method keeps all the advantages of the previously described enzymatic methods for *ee* analysis and displays new remarkable properties: 1) the addition of a coupling enzyme allows a more sensitive detection of alcohols; 2) the previously used equations were adjusted to deal with unexpected enzyme kinetics; 3) the method is highly suitable for high-throughput screening, as demonstrated by its application in measuring the *ee* of alcohol mixtures produced by lipase mutants. Moreover, the inherent problems associated to the use of complex media in coupled enzyme reactions were circumvented.

Experimental Section

Diaphorase was purchased from Sigma-Aldrich as lyophilized powder. Recombinant LKADH and READH were purified as described previously[5] and their activities were measured for the reduction of acetophenone at 30℃, pH 7.0 in 50 mM triethanolamine buffer, with the concomitant production of NAD(P)H. Alcohols were purchased from ACROS Organics or Sigma-Aldrich. For high-throughput screening, a solution containing 1.0 U mL^{-1} of LKADH or 0.5 U mL^{-1}

of READH, 1.0 U mL^{-1} of diaphorase, 0.22 mM INT and 0.1% TritonX100 in phosphate buffer 200 mM pH 7.0 was prepared. 40 μL of this solution was mixed in 384-well microtiter plates with 20 μL the samples containing the alcohols to be analysed. Finally, to start the reaction, 20 μL of a 4.0 mM solution of the coenzymes, NAD for READH or NADP for LKADH, were added and mixed in the well. The mixture was then incubated 5 min at 30℃ and absorbance measurements at 492 nm were recorded. The samples containing the alcohols are composed of 100 μL of 0.5 M phosphate buffer pH 7.0, 400 μL of an emulsion of the substrate 0.2% (v/v) in presence of 0.1% of Arabic gum, and 500 μL of supernatants from cultures of *E. coli* BL21(DE3) harbouring the plasmid pET22b+ with the gene for the BSLA variants. This mixture was incubated at 37℃ until the desired advance of the hydrolysis reaction was reached. As negative controls we used the supernatants of the same strain of *E. coli* harbouring the same plasmid without the gene for BSLA, grown at the same conditions as all the other cultures. To construct a standard curve, we added defined concentrations of the alcohols, from 10 to 150 μM, to the negative-control supernatants. All liquid handling, incubation and microplate reading was carried out in a Tecan Workstation 200 pipetting robot equipped with a microplate reader. Site-specific saturation mutagenesis was performed as described in reference 9 and the epPCR library construction was carried out as described in reference 10.

[1] a) B. Plietker *Angew. Chem.* **2006**, *118*, 196-198; *Angew. Chem. Int. Ed.* **2006**, *45*, 190-192; b) W. S. Knowles, *Angew. Chem.* **2002**, *114*, 2096-2107; *Angew. Chem. Int. Ed.* **2002**, *41*, 1998-2007; c) R. Noyori, *Angew. Chem.* **2002**, *114*, 2108-2123; *Angew. Chem. Int. Ed.* **2002**, *41*, 2008-2022; d) K. B. Sharpless, *Angew. Chem.* **2002**, *114*, 2096-2107; *Angew. Chem. Int. Ed.* **2002**, *41*, 2126-2135.

[2] a) T. Bein, *Angew. Chem.* **1999**, *111*, 335-338; *Angew. Chem. Int. Ed.* **1999**, *38*, 323-326; b) M. Reetz, *Angew. Chem.* **2001**, *113*, 292-320; *Angew. Chem. Int. Ed.* **2001**, *40*, 284-310; c) M. Reetz, K.-E. Jaeger, *Chem. Eur. J.* **2000**, *6*, 407-412; d) K.-E. Jaeger, T. Eggert, A. Eipper, M. Reetz, *Appl. Microbiol. Biotechnol.* **2001**, *55*, 519-530.

[3] a) P. Tielmann, M. Boese, M. Luft, M. Reetz, *Chem. Eur. J.* **2003**, *9*, 3882-3887; b) M. Reetz, K. M. Kühling, H. Hinrichs, A. Deege, *Chirality* **2002**, *12*, 479-482; c) M. Reetz, *Angew. Chem.* **2002**, *114*, 1391-1394; *Angew. Chem. Int. Ed.* **2002**, *41*, 1335-1338; d) J. Guo, J. Wu, G. Siuzdak, M. G. Finn, *Angew. Chem.* **1999**, *111*, 1868-1871; *Angew. Chem. Int. Ed.* **1999**, *38*, 1755-1758.

[4] a) Z. Li, L. Bütikofer, B. Witholt *Angew. Chem.* **2004**, *116*, 1730-1734; *Angew. Chem. Int. Ed.* **2004**, *43*, 1698-1702; b) M. B. Onaran, C. T. Seto, *J. Org. Chem.* **2003**, *68*, 8136-8141; c) M. Baumann, R. Stürmer, U. T. Bornscheuer *Angew. Chem.* **2001**, *113*, 4329-4333; *Angew. Chem. Int. Ed.* **2001**, *40*, 4201-4204; d) P. Abato, T. Seto, *J. Am. Chem. Soc.* **2001**, *123*, 9206-9207.

[5] a) A. Weckbecker, W. Hummel, *Biotechnol. Lett.* **2004**, *26*, 1739-1744; b) K. Abokitse, W. Hummel, *Appl. Microbiol. Biotechnol.* **2003**, *62*, 380-386; c) C. W. Bradshaw, W. Hummel, C.-H. Wong, *J. Org. Chem.* **1992**, *57*, 1532-1536.

[6] a) J. S. Easterby, *Biochem. J.* **1984**, *219*, 834-847; b) A. C. Storer, A. Cornish-Bowden, *Biochem. J.* **1974**, *141*, 205-209.

[7] A. Cornish-Bowden, *Fundamentals of Enzyme Kinetics*, Portland Press, London, 2004, pp. 277-316.

[8] T. Eggert, G. Pancreac´h, I. Douchet, R. Verger, K.-E. Jaeger *Eur. J. Biochem.* **2000**, *267*, 6459-6469.

[9] S. A. Funke, A. Eipper, M. Reetz, N. Otte, W. Thiel, G. van Pouderoyen, B. W. Dijkstra, K.-E. Jaeger, T. Eggert, *Biocatal. Biotransform.* **2003**, *21*, 67-73.

[10] T. Eggert, C. Leggewie, M. Puls, W. Streit, G. van Pouderoyen, B. W. Dijkstra, K.-E. Jaeger, *Biocatal. Biotransform.* **2004**, *22*, 139-144.

Supplementary Material

The linearity of colour development was studied at different substrate concentrations. In Figure 2 and 3 is clear that the method displays a high linearity during the first ten minutes. Although, is preferred to measure the colour formation for shorter times because the reaction reaches its equilibrium at short times when low concentrations of the alcohol are present, this certainly increased the consistency of data.
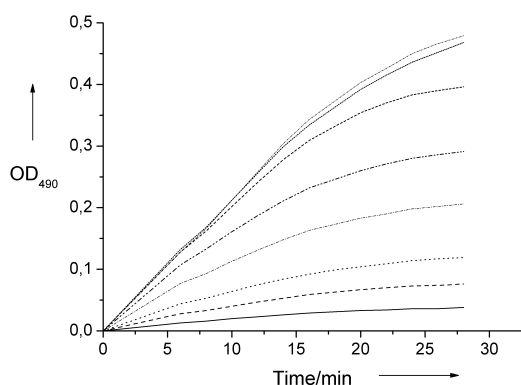

Figure 2. Colour development by the coupled reactions of LKADH and diaphorase with different concentrations of substrate **5** (5.0, 11.0, 16.5, 27.5, 38.5, 55.0, 82.5 and 110 µM) at pH 7 and 30°C.

Figure 3. Colour development by the coupled reactions of READH and diaphorase with different concentrations of substrate **5** (5.0, 11.0, 16.5, 27.5, 38.5, 55.0, 82.5 and 110 µM) at pH 7 and 30°C.

The presence of culture media in the coupled assay produces a background colour development, mainly in the presence of the READH. This background can be seen in Figure 4, in which the samples to be analysed by the two enzymes are positioned side by side. Nevertheless, when the controls are taken into account, the presence of amounts of alcohol as low as 5.0 µM can be easily detected. Higher concentrations can be seen by naked eye as shown in Figure 4.
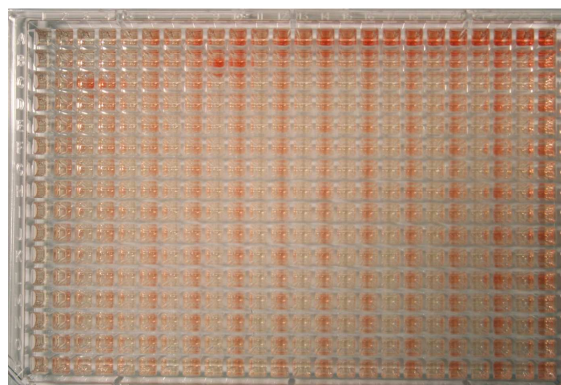
Figure 4. Screening for activity and *ee* in samples from an epPCR-generated library of a BSLA variant including an artificial lid motif. The image shows the colour development in a 384-well microtiter plate after 5 minutes of incubation at 30°C. The first row contains, from left to right, increasing concentrations of substrate **5**. In the second row a sample with high concentration of substrate is visible; this sample comes from the very active mutant MPB7. In the left side of the third row, a sample from the wild-type BSLA is present.

Identification of Novel Benzoylformate Decarboxylases by Growth Selection[∇][†]

Helge Henning,[1][‡] Christian Leggewie,[1][‡] Martina Pohl,[1] Michael Müller,[2]
Thorsten Eggert,[1] and Karl-Erich Jaeger[1]*

*Institute of Molecular Enzyme Technology, Heinrich Heine University Duesseldorf, Research Centre Juelich, D-52426 Juelich,
Germany,[1] and Institute of Pharmaceutical Sciences, Albert Ludwigs University, D-79104 Freiburg, Germany[2]*

A growth selection system was established using *Pseudomonas putida*, which can grow on benzaldehyde as the sole carbon source. These bacteria presumably metabolize benzaldehyde via the β-ketoadipate pathway and were unable to grow in benzoylformate-containing selective medium, but the growth deficiency could be restored by expression in *trans* of genes encoding benzoylformate decarboxylases. The selection system was used to identify three novel benzoylformate decarboxylases, two of them originating from a chromosomal library of *P. putida* ATCC 12633 and the third from an environmental-DNA library. The novel *P. putida* enzymes BfdB and BfdC exhibited 83% homology to the benzoylformate decarboxylase from *P. aeruginosa* and 63% to the enzyme MdlC from *P. putida* ATCC 12633, whereas the metagenomic BfdM exhibited 72% homology to a putative benzoylformate decarboxylase from *Polaromonas naphthalenivorans*. BfdC was overexpressed in *Escherichia coli*, and the enzymatic activity was determined to be 22 U/ml using benzoylformate as the substrate. Our results clearly demonstrate that *P. putida* KT2440 is an appropriate selection host strain suitable to identify novel benzoylformate decarboxylase-encoding genes. In principle, this system is also applicable to identify a broad range of different industrially important enzymes, such as benzaldehyde lyases, benzoylformate decarboxylases, and hydroxynitrile lyases, which all catalyze the formation of benzaldehyde.

White biotechnology uses microorganisms and (microbial) enzymes to manufacture a wide variety of different chemicals, including polymers, bulk chemicals, agrochemicals, and pharmaceuticals (18). The biocatalytic production of these chemicals offers a number of advantages, including usually mild reaction conditions, the avoidance of toxic wastes, and, in particular, access to a variety of enantiopure compounds due to the high substrate specificity and enantioselectivity of many enzymes. However, the identification of novel biocatalysts that possess desired properties is still a challenge. Classical strain-screening approaches that are used for the isolation of microorganisms are laborious and often unsuccessful. Recently, the metagenome approach was developed to access useful genetic information encoded by environmental DNA, which is directly isolated from various habitats, cloned, and expressed in appropriate microbial host strains (37). However, the identification of a clone producing a biocatalyst of interest within a large metagenomic library requires high-throughput screening technologies, which are difficult to set up and usually expensive to run. Selection provides an elegant solution to the problem. Here, the desired enzyme activity is linked to the survival of the respective clone. Novel selection strategies that use natural or modified transcriptional regulators that bind to the product of

an enzymatic reaction have recently been developed (22). Transcriptional control can also be accomplished by a modified eukaryotic nuclear receptor (39) or by a "riboswitch" RNA (3). Enzymes producing compounds such as prephenate (7, 28), pyruvate (8), ammonia (36), and 2-hydroxybenzaldehyde (48) and 1,2,4-trichlorobenzene (29) were identified. A similar approach, termed chemical selection or complementation (1), is based on the detection of small molecules in an in vivo assay of lyase- and ligase-type reactions with a modified yeast three-hybrid system.

In the present work, we report on the construction of a selection system useful to identify novel enzymes producing benzaldehyde as the reaction product, which can be used by the bacterial selection host as the sole carbon source. The respective strain must possess the enzyme benzaldehyde dehydrogenase and the β-ketoadipate pathway to convert benzoate into the tricarboxylic acid cycle intermediates succinyl-coenzyme A (CoA) and acetyl-CoA (Fig. 1). The genes encoding enzymes of the β-ketoadipate central pathway were detected in many bacteria, mainly belonging to the genera *Acinetobacter* and *Pseudomonas* (10). An example is *Pseudomonas putida* ATCC 12633, which is able to grow on aromatic compounds as the sole carbon source (13–15). In this strain, mandelate is converted to acetyl-CoA via the mandelate/β-ketoadipate pathway (Fig. 1), and the enzyme benzoylformate decarboxylase MdlC (synonym, BFD) (E.C. 4.1.1.7) catalyzes the formation of benzaldehyde from benzoylformate by decarboxylation. The structure of BFD was solved in the absence (11) and the presence (34) of mandelic acid as an inhibitor, confirming that the enzyme acts as a tetramer, and active-site residues have been elucidated by site-directed mutants (34, 40) and directed-evolution studies (26, 27).

* Corresponding author. Mailing address: Institute of Molecular Enzyme Technology, Heinrich Heine University Duesseldorf, Research Centre Juelich, D-52426 Jülich, Germany. Phone: (49)-2461-613716. Fax: (49)-2461-612490. E-mail: karl-erich.jaeger@fz-juelich.de.
† This paper is dedicated to Maria-Regina Kula on the occasion of her 70th birthday.
‡ H.H. and C.L. contributed equally to this work.
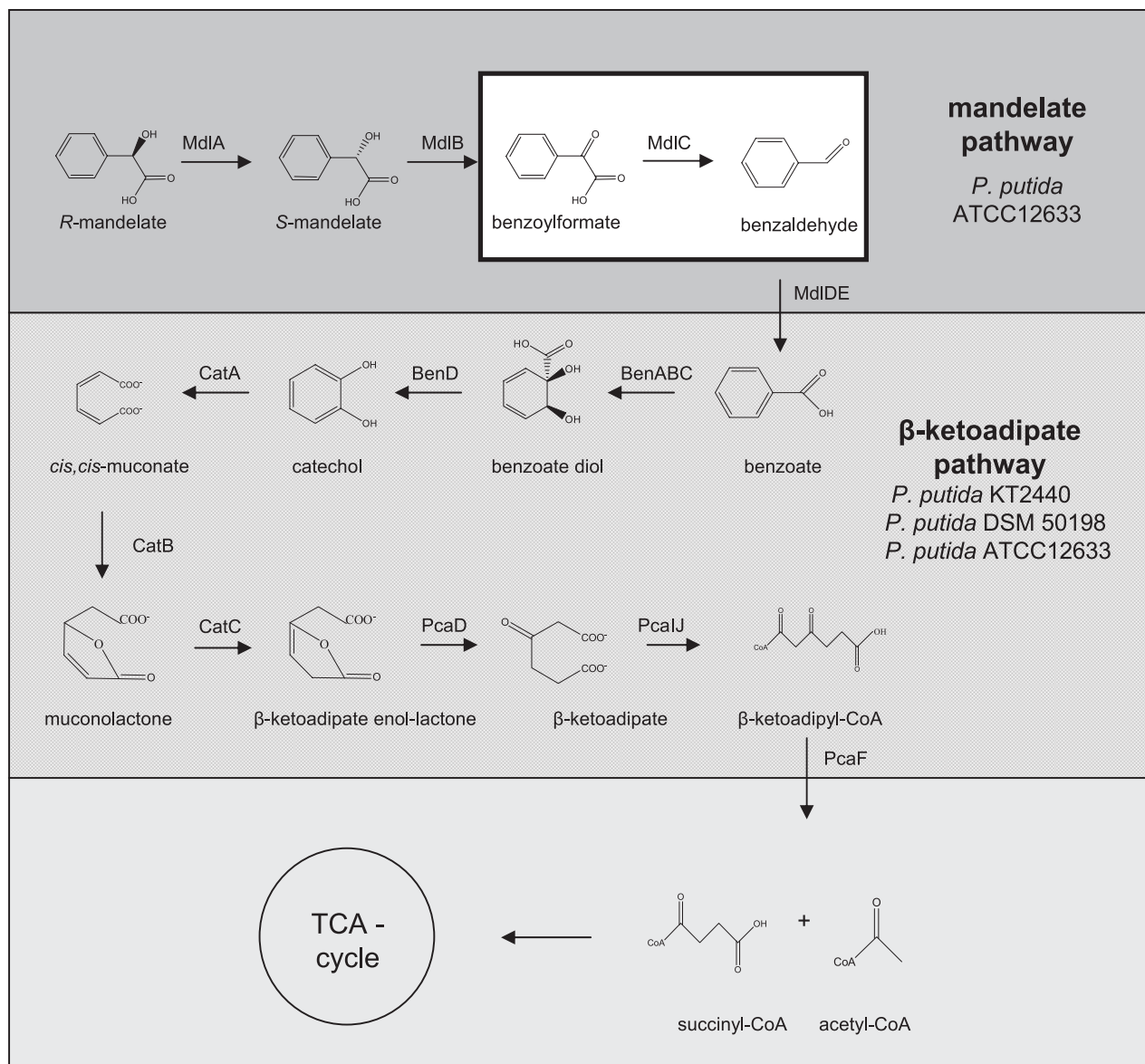∇ Published ahead of print on 29 September 2006.

FIG. 1. Mandelate and β-ketoadipate pathways in *P. putida*. The mandelate pathway in *Pseudomonas putida* ATCC 12633 is shown in dark gray, and the conversion of benzoylformate to benzaldehyde by a benzoylformate decarboxylase (MdlC) is highlighted. The β-ketoadipate pathway (gray) also exists in the majority of all pseudomonads. Additionally, *P. putida* KT2440 and *P. putida* DSM50198 possess a benzaldehyde dehydrogenase of unknown function. The metabolization of succinyl-CoA and acetyl-CoA via the tricarboxylic acid cycle is shown in light gray. MdlA, mandelate racemase; MdlB, *S*-mandelate dehydrogenase; MdlDE, NAD⁺- and NADP⁺-benzaldehyde dehydrogenases; BenABC, benzoate dioxygenase; BenD, 2-hydro-1,2-dihydroxybenzoate dehydrogenase; CatA, catechol-1,2-dioxygenase; CatB, *cis,cis*-muconate lactonizing enzyme (cycloisomerase); CatC, muconolactone isomerase; PcaD, β-ketoadipate enolactone hydrolase I; PcaIJ, β-ketoadipate succinyl-CoA transferase subunit; PcaF, β-ketoadipyl CoA thiolase; TCA, tricarboxylic acid.

We have newly constructed a *Pseudomonas*-based selection system to identify BFDs. These enzymes need thiamine diphosphate as a cofactor and have been identified in bacteria such as *Pseudomonas putida*, *Acinetobacter calcoaceticus*, and *Pseudomonas aeruginosa* (2, 2a, 12). They catalyze the decarboxylation of benzoylformate to benzaldehyde, but they also exhibit a carboligase side activity enabling the conversion of aldehydes to chiral 2-hydroxy ketones, which are versatile building blocks for a variety of different fine chemicals (17, 33).

The functionality of the selection system was proven by the isolation of three novel benzoylformate decarboxylases, one of them originating from a metagenomic library. These novel enzymes showed only low sequence similarity to presently known BFDs.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. The bacterial strains and plasmids used in this study are listed in Table 1. The *Escherichia coli* strain DH5α was used as a host for cloning and construction of recombinant plasmids. The strain *E. coli* DH5α carrying the plasmid pRK2013 with the *tra* genes was used

TABLE 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics[a]	Reference or source
Pseudomonas strains		
P. putida ATCC 12633	Wild type; catabolizes mandelate	ATCC (42, 43)
P. aeruginosa PAO1	Wild type	16
P. putida KT2440	Wild type; toluene-degrading *P. putida* mt-2	30, 35
P. putida DSM50198	Wild type; lack of mandelate pathway	DSMZ
E. coli strains		
DH5α	λ⁻ φ80d*lacZ*Δ*M15* Δ(*lacZYA-argF*)*U169 recA1 endA1 hsdR17*(rK⁻ mK⁻) *supE44 thi-1 gyrA relA1*	9
BL21(DE3)	F⁻ *ompT hsdS$_B$*(r$_B$⁻ m$_B$⁻) *gal dcm* (DE3)	45, 46
Plasmids		
pBBR1MCS-1	*rep mob lacZ'* Cm^r	24
pBBR1MCS-2	*rep mob lacZ'* Km^r	23
pBBR1MCS-5	*rep mob lacZ'* Gm^r	23
pSUP202	pBR325 derivative; Amp^r Cm^r Tc^r *mob*	41
pBluescript SK	P$_{T7}$ P$_{T3}$ P$_{lac}$ *lacZ'* Cm^r ColE1	Stratagene
pWKR202	pACY177 derivative; Gm^r	4
pRK2013	*tra*(RK2)$^+$ Km^r ColE1	6
pET16b	P$_{T7}$ P$_{lac}$ Amp^r ColE1; His tag	Novagen
pET22b	P$_{T7}$ P$_{lac}$ Amp^r ColE1 *pelB* signal sequence; His tag	Novagen
pG-TF2	P$_{Tc}$ *groES groEL tig* Cm^r	Takara Bio Inc. (31)
pBBRBFDPp	Gm^r; pBBR1MCS-5 derivative carrying the *mdlC* gene from *P. putida* ATCC 12633	This study
pBBRBFDPa	Gm^r; pBBR1MCS-5 derivative carrying the *mdlC* gene from *P. aeruginosa* PAO1	This study
pSUPΔMDLC	Plasmid based on pSUP202 for deletion of *mdlC* from *P. putida* ATCC 12633; Amp^r Cm^r Tc^r Gm^r	This study

[a] Abbreviations for antibiotics: Ap, ampicillin; Cm, chloramphenicol; Km, kanamycin; Gm, gentamicin; Tc, tetracycline.

for triparental conjugation. *E. coli* strains were cultured at 37°C and *Pseudomonas* strains at 30°C in solid or liquid Luria-Bertani medium. *Pseudomonas* strains displaying benzoylformate decarboxylase activity were selected by growth in liquid or solid minimal medium M9 (38) without glucose supplemented with 10 mM benzoylformate (purchased from Sigma-Aldrich) as the sole carbon source. The bacteria were incubated for at least 2 days at 30°C. When necessary, antibiotics were added at the following concentrations: chloramphenicol, 600 μg/ml; kanamycin, 50 μg/ml; and gentamicin, 30 μg/ml for *Pseudomonas* strains, and chloramphenicol, 50 μg/ml; kanamycin, 25 μg/ml; gentamicin, 10 μg/ml; and ampicillin, 100 μg/ml for *E. coli* strains.

Construction of the deletion mutant *P. putida* Δ*mdlC*. Genomic DNA from *P. putida* ATCC 12633 was used as a template for the amplification of a 500-bp fragment located at either the 5′ (*mdlC′*) or the 3′ (*′mdlC*) end of the 1,584-bp *mdlC* gene coding for benzoylformate decarboxylase. The following primers were used: BFDanfup (5′-ATA TGG ATC CAA GCT TAC ATG GCG ATC AAA AAG GTG G-3′), BFDanfdw (5′-ATA TCT CGA GGT CGA CGG CAC CAC ATA CGA ACT CTT-3′), BFDendup (5′-ATA TCC TAG AGT CGA CGG CCT TTG GCA GAA AGC GCT T-3′), and BFDenddw (5′-ATA TGG ATC CCC CGC GAA GGT TGA CCA AGA CGC TGG C-3′).

A gentamicin resistance cassette from plasmid pWKR202 was inserted between these two 500-bp fragments of *mdlC* by using the restriction site HindIII. The cloning steps were carried out in *E. coli* DH5α by using the plasmid pBluescript SK(+). The mutagenesis cassette consisted of three different fragments in the order *mdlC′*, gentamicin resistance cassette, and *′mdlC*. Subsequently, the mutagenesis cassette was cloned into the suicide vector pSUP202, and the resulting plasmid, pSUPΔMDLC, was transferred via triparental conjugation (6) into *P. putida* ATCC 12633. The benzoylformate decarboxylase gene, *mdlC*, was interrupted by homologous recombination, and the correct insertion of the mutagenesis cassette was confirmed by PCR analysis.

Construction of a genomic-DNA library from *P. putida* ATCC 12633 and *P. putida* ATCC 12633 Δ*mdlC*. The genomic DNAs from *P. putida* ATCC 12633 and *P. putida* ATCC 12633 Δ*mdlC* were isolated by a standard procedure (38) and partially digested with the restriction endonuclease Sau3AI. The broad-host-range vector pBBR1MCS was restricted with BamHI and used for cloning of genomic-DNA fragments of 2 to 8 kb. Finally, the genomic-DNA library was transferred into *P. putida* KT2440 by conjugation. The average insert size and the

frequency of clones harboring only vector DNA within each library were determined by restriction analysis of 24 clones.

Cloning and expression of the *mdlC* genes from *P. putida* ATCC 12633 and *P. aeruginosa* PAO1. Both *mdlC* genes were amplified using genomic DNA from *P. putida* ATCC 12633 and *P. aeruginosa*, respectively. The following primers were used: BFDPpup (5′-ATA TCCATG GCT TCG GTA CAC GGC ACC ACA TAC-3′), BFDPpdw (5′-ATA TCT CGA GTC ACT TCA CCG GGC TTA CGG TGC TTA C-3′), BFDPaup (5′-ATA TCA TAT GAA AAC CGT CCA TTC CGC G-3′), and BFDPadw (5′-ATA TAA GCT TTC AGG GTT CGA TGG TTT GCG-3′). The primers were designed for cloning into plasmid pBBR1MCS, which was used for subsequent screening of genomic or meta-genomic libraries. Both PCR products were ligated into SmaI-restricted pBBR1MCS-2, conferring kanamycin resistance. The plasmids were designated pBBRBFDPp for the *mdlC* gene from *P. putida* ATCC 12633 and pBBRBFDPa for the *mdlC* gene from *P. aeruginosa*. The correct integration of the genes was confirmed by DNA sequencing. Subsequently, the recombinant plasmids were transferred to *P. putida* KT2440 and analyzed for growth on agar plates containing 10 mM benzoylformate as the sole carbon source.

Cloning and expression of novel *bfd* genes from *P. putida* ATCC 12633. Putative *bfd* genes were amplified using genomic DNA from *P. putida* ATCC 12633. The following primers for cloning into expression plasmids pET16b and pET22b (Novagen) were used: BFD3NdepET (5′-ATA TCA TAT GAA AAC TGT TCA CGG CGC CAC-3′), BFD3BampET (5′-ATA TGG ATC CGG GCT CGA TGG TCT GGG TCG-3′), and BFD2BamHIoSTC (5′-TGG CCT TGA GGA TCC GCG GCT GCT-3′). The PCR products were cloned into the NdeI/BamHI-digested expression vectors pET16b and pET22b (Novagen). The correct integration of *bfdB* and *bfdC* was confirmed by DNA sequencing. The heterologous expression of both genes was accomplished in *E. coli* BL21(DE3), with coexpression of chaperones from plasmid pG-TF2 (31), as recommended by the manufacturer (Takara Bio Inc.).

Isolation and cloning of DNA from soil. A soil sample was collected from a meadow near Juelich, Germany, and the metagenomic DNA was isolated from 5 g soil based on the direct-lysis method of Zhou et al. (49). Additionally, the isolated metagenomic DNA was purified using the DNeasy Tissue Kit (QIAGEN), starting the protocol with the washing step. The purified metagenomic DNA was partially digested with the restriction endonuclease Sau3AI.

Fragments of 2 to 9 kb were isolated and ligated into the BamHI-restricted broad-host-range vector pBBR1MCS. The resulting metagenomic-DNA library was transferred to *P. putida* KT2440 by conjugation. The average insert size and the frequency of clones harboring only vector DNA within each library were determined by restriction analysis of 24 plasmids isolated from randomly chosen clones.

Cloning and expression of the novel BFD gene *bfdM* from a metagenomic library. The plasmid DNA from a clone exhibiting decarboxylase activity was sequenced by genome walking, and the following primers were used for the amplification of the novel *bfdM* gene and subsequent cloning into the expression vector pET22b (Novagen): BFDM1Nde (5′-ATA TCA TAT GCA AGA GAC AAC CCC CCA GAA T-3′) and BFDM1BamH (5′-ATA GGA TCC GGC CAC TTC GAC GAG CAC GGG C-3′). The PCR product was cloned into the NdeI/BamHI-digested expression vector pET22b (Novagen), and the correct integration of *bfdM* was confirmed by DNA sequencing. The heterologous expression was accomplished as described for *bfdB* and *bfdC*.

Coupled decarboxylase assay. The decarboxylase activity toward benzoylformate was determined as described previously (17). In brief, the assay mixture was prepared from the following stock solutions in standard buffer: benzoylformate solution (100 μl, 50 mM, adjusted to pH 6.0), NADH (100 μl, 3.5 mM), horse liver alcohol dehydrogenase (50 μl, 10 U; Sigma), and potassium phosphate buffer (700 μl, 50 mM, pH 6.0). The components were mixed in a 1.7-ml cuvette and incubated for 3 min at 30°C, and the reaction was started by the addition of sonicated cell extracts from overexpression cultures. These extracts were prepared as follows: 100 μl (equal to an optical density of 20 at 580 nm) supernatant of extracts centrifuged at 13,000 rpm for 20 min was sonicated twice for 3 min each time (50 cycles; 80% power; Sonoplus HD2070; Bandelin, Berlin, Germany). The slope was calculated from the linear part of the descending absorption curve determined at 340 nm and within a reaction time of 15 to 90 s.

Sequence analysis. The similarity of newly identified genes to known DNA or amino acid sequences was determined by Blastn or Blastp (http://www.ncbi.nlm.nih.gov/BLAST/). The sequence alignment was performed with ClustalW (http://www.ebi.ac.uk/clustalw/).

Nucleotide sequence accession numbers. The sequences of the three novel benzoylformate decarboxylase genes described here have been deposited in the EMBL database. They are AM284966 for *bfdB*, AM284967 for *bfdC*, and AM284968 for *bfdM*.

RESULTS AND DISCUSSION

Construction of a *Pseudomonas putida* selection strain. The elimination of a functional *mdlC* gene in *P. putida* ATCC 12633 interrupts the mandelate pathway and prevents the corresponding knockout mutant, *P. putida* ATCC 12633 Δ*mdlC*, from growing on mandelate as the sole carbon source. Consequently, this strain should be usable as a selection host to identify novel benzoylformate decarboxylase genes that can be introduced in *trans* and that may originate from a genomic, a metagenomic, or a directed-evolution library. The deletion mutant *P. putida* ATCC 12633 Δ*mdlC* was constructed by the insertion of a gentamicin resistance cassette into the functional gene *mdlC* and was cultivated on selective medium containing benzoylformate as the sole carbon source, enabling growth of the recombinant bacteria only upon expression of a functional plasmid-encoded BFD. Surprisingly, the mutant *P. putida* ATCC 12633 Δ*mdlC* was still able to grow on benzoylformate-containing medium. Therefore, this strain could not be used as a selection host; however, this result clearly indicated either the presence of so far unknown genes encoding enzymes with BFD activity or an unknown pathway for the degradation of benzoylformate.

Thus, we searched for bacterial strains having the ability to grow with benzaldehyde as the sole carbon source but lacking the enzymes catalyzing the reactions from (*R*)-mandelate to benzoylformate (Fig. 1). These bacteria are expected to possess a benzaldehyde dehydrogenase and to metabolize benzal-
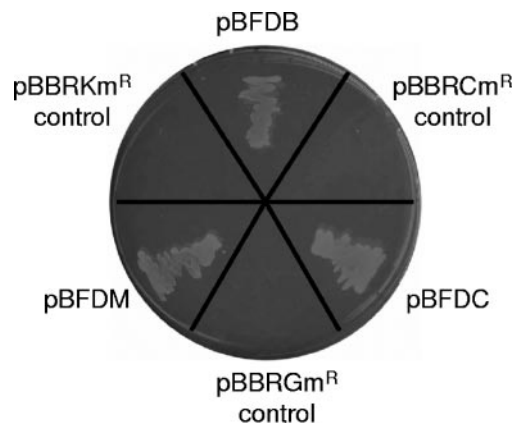

FIG. 2. Growth of recombinant *P. putida* KT2440 displaying benzoylformate decarboxylase activity on an agar plate containing benzoylformate selective medium. *P. putida* KT2440 possessing the plasmids pBFDB (coding for BfdB from the genomic-DNA library of *P. putida* ATCC 12633), pBFDC (coding for BfdC from the genomic-DNA library of *P. putida* ATCC 12633 Δ*mdlC*), and pBFDM (coding for BfdM from a metagenomic library) expressed active BFD enzymes enabling growth by converting benzoylformate. *P. putida* strains harboring the corresponding plasmids, pBBRCm^R, pBBRGm^R, and pBBRKm^R, served as controls.

dehyde via the β-ketoadipate pathway to yield the products acetyl-CoA and succinyl-CoA. A database analysis revealed that the strains *P. putida* KT2440 and *P. putida* DSM50198 should exhibit these features (21, 30). Furthermore, Tsou et al. described the presence of a β-ketoadipate pathway in *P. putida* DSM50198, which lacks genes of the mandelate pathway (47). First, we confirmed that both strains were unable to grow in selective medium containing benzoylformate as the sole carbon source. As a functional test, we then cloned into the broad-host-range vector pBBR1MCS the *mdlC* genes encoding BFDs from both *P. aeruginosa* PAO1 and *P. putida* ATCC 12633. The corresponding plasmids were transferred into the putative selection host strains, *P. putida* KT2440 and *P. putida* DSM50198, and both recombinant strains were able to grow on benzoylformate-containing medium, whereas the control strains, which lack the BFD genes, did not grow at all (data not shown). These results clearly indicated that both *P. putida* strains were useful as selection hosts.

Identification of two novel BFD-encoding genes in the genome of *P. putida* ATCC 12633. Our results suggested that *P. putida* ATCC 12633 may possess at least one so far unknown gene encoding a BFD activity. We therefore constructed DNA libraries from chromosomal DNAs of *P. putida* ATCC 12633 and the deletion mutant *P. putida* ATCC 12633 Δ*mdlC* using the plasmid pBBR1MCS. The resulting genomic-DNA libraries consisted of 4,500 and 4,000 clones for the *P. putida* wild type and the *mdlC* mutant, respectively, both with an average insert size of 5 kb, thereby reaching a fourfold coverage of the *P. putida* genome. Initially, these libraries were constructed in *E. coli* and then transferred to *P. putida* KT2440 by triparental mating. We identified several clones that grew on benzoylformate selective media. The corresponding plasmids were isolated, and restriction analysis revealed two different fragment patterns. Subsequent DNA sequencing identified two novel genes, which were designated *bfdB* and *bfdC*. The two genes

TABLE 2. Plasmids encoding putative benzoylformate decarboxylases, other ORFs identified, and their closest homologues
as identified from databases[a]

Plasmid name (insert size [kb])	ORF	Closest database similarity		
		Protein (accession no.)	Organism	Identity (%)
pBFDB (6.234)	BkdA1	Ketoacid dehydrogenase E1 α-subunit (AAA 65614)	*Pseudomonas putida*	100
	BkdR	Transcription regulator (AAA 65613)	*Pseudomonas putida*	100
	Orf3	α/β Hydrolase (ZP_00898945)	*Pseudomonas putida*	95
	Orf4	Glutamin synthetase (ZP_00900555)	*Pseudomonas putida*	95
	BenC	Benzoate dioxygenase (AAF 63450)	*Pseudomonas putida*	85
	BfdB	Benzoylformate decarboxylase (NP_253588)	*Pseudomonas aeruginosa*	83
pBFDC (2.952)	Orf1′	Putative LysR regulator (AAL 27559)	*Pseudomonas putida*	100
	BfdC	Benzoylformate decarboxylase (NP_253588)	*Pseudomonas aeruginosa*	83
	Orf3′	Probable MFS[b] transporter (NP-253587)	*Pseudomonas aeruginosa*	82
pBFDM (6.586)	Orf1″	Short-chain dehydrogenase (ZP_00243984)	*Rubrivivax gelatinosus*	65
	Orf2″	Transcriptional regulatory protein (ZP-01022169)	*Polaromonas naphthalenivorans*	60
	BfdM	Benzoylformate decarboxylase (ZP-01022170)	*Polaromonas naphthalenivorans*	72
	Orf4″	Hypothetical protein (ZP_01022174)	*Polaromonas naphthalenivorans*	60
	Orf5″	Sigma e factor sigma 24 (ZP_01022480)	*Polaromonas naphthalenivorans*	60

[a] The diagram shows the organization of *P. putida* genomic (pBFDB and pBFDC) and environmental (pBFDM) DNAs present in plasmids that conferred benzoylformate decarboxylase activity on *P. putida*. ORFs encoding benzoylformate decarboxylase activity are shown as black arrows; partial arrows represent incomplete ORFs.

[b] MFS, major facilitator superfamily.

share identical DNA sequences, except for those nucleotides encoding the C-terminal nine amino acids. The clones expressing these genes were able to grow on benzoylformate selective media, whereas the respective control strain showed no growth at all (Fig. 2). The two BFD-encoding genes are located in different regions of the *P. putida* ATCC 12633 genome, as deduced from the sequences of the respective flanking regions (Table 2). The genes up- and downstream of *bfdC* resemble the genomic region surrounding *mdlC* in *P. aeruginosa*, where a putative benzoate transporter of the major facilitator superfamily and a putative transcriptional regulator are encoded in the same orientation. Thus, the *bfdC* gene seems to be part of a gene cluster encoding proteins for the degradation of aromatic compounds, as suggested for MdlC of *P. aeruginosa* (44). On the other hand, *bfdB* seems more likely to have originated from a gene transfer. To our knowledge, the region surrounding this gene does not show any homology to known regions adjacent to other BFD genes identified in published bacterial genomes. The enzymes BfdB and BfdC exhibit 83% identity to the benzoylformate decarboxylase from *P. aeruginosa* and 63% to the previously described enzyme MdlC from *P. putida* ATCC 12633. These results demonstrate that (i) *P. putida*

KT2440 is an appropriate selection host strain and (ii) *P. putida* ATCC 12633 possesses two additional hitherto-unknown BFD-encoding genes, *bfdB* and *bfdC*, one of which was most likely duplicated during evolution.

Construction and screening of a metagenomic library. The successful identification of novel biocatalysts exhibiting carboligase activities from environmental DNA would provide additional support for a more general applicability of the growth selection system we have constructed. Therefore, we isolated metagenomic DNA directly from soil and constructed plasmid-based libraries in *P. putida* which were subsequently screened for BFD activity. The randomly digested metagenomic DNA was cloned into pBBR1MCS, transferred into *E. coli*, and finally propagated in *P. putida* KT2440. The metagenome library consisted of about 14,000 clones, with an average insert size of 2 to 10 kb. Plating of *P. putida* KT2440 on benzoylformate selective medium revealed one clone which was able to grow after 2 days of incubation (Fig. 2). The respective plasmid contained an insert with a size of 7 kb. DNA sequence determination revealed an open reading frame (ORF), which was designated *bfdM*. The deduced protein exhibits 72% identity to
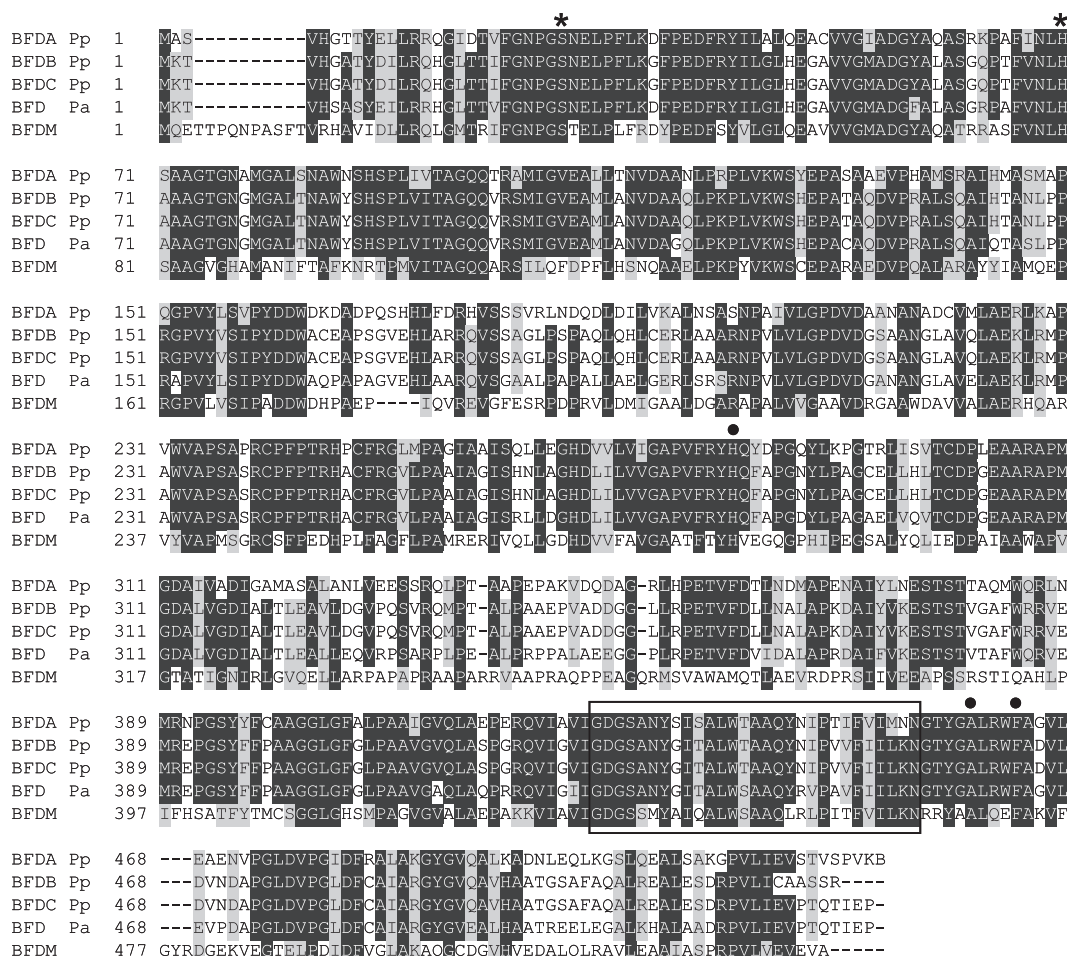
FIG. 3. Sequence alignment of MdlC (BfdA), BfdB, and BfdC from *P. putida* ATCC 12633 (Pp), BFD from *Pseudomonas aeruginosa* (Pa), and BfdM isolated from the metagenome. Highly conserved regions are highlighted in black, and similar amino acid residues are shown in gray. Amino acids forming the catalytic site are indicated by an asterisk. Residues tagged with a dot are involved in substrate binding (34, 40), and the boxed residues are involved in thiamine diphosphate cofactor and metal binding.

the amino acid sequence of a putative benzoylformate decarboxylase from *Polaromonas naphthalenivorans* CJ2T (Table 2). We further identified additional DNA regions upstream and downstream of *bfdM* which also showed similarity to genes from this gram-negative bacterium, including a putative transcriptional regulatory protein and a putative sigma factor. Recently, *P. naphthalenivorans* CJ2T was isolated from a coal tar-contaminated freshwater sediment, and it was demonstrated that this strain could grow on the polycyclic aromatic hydrocarbon naphthalene as the sole carbon and energy source (20). These findings suggest that *bfdM* may originate from an organism that shares the ability to grow on aromatic compounds.

Overexpression and characterization of the novel BFDs. The primary structures of the three novel BFDs we have identified were analyzed with respect to conserved regions, including active and cofactor binding sites. Figure 3 shows a sequence alignment of the novel enzymes with BFDs from *P. putida* KT244 and *P. aeruginosa* (MdlC), which shows the highest identity to BfdB and BfdC. All amino acid side chains forming the active site, the substrate binding site, and

the thiamine diphosphate binding site are highly conserved in the new enzymes. This also holds for the metagenome-derived enzyme BfdM, indicating that the reaction mechanism of benzoylformate decarboxylases is well conserved.

All three enzymes were detected by an activity screen toward the substrate benzoylformate upon expression in the homologous host *P. putida* KT2440; however, in all cases, the amounts of active enzyme produced were very low. The overexpression of the three putative *bfd* genes in *E. coli* BL21(DE3) under the control of the T7 promoter did not result in the formation of enzymatically active enzymes, presumably because inclusion bodies were formed. Therefore, we coexpressed the chaperones GroEL, GroES, and trigger factor from plasmid pG-TF2, trying to prevent the formation of inclusion bodies as described previously (31). While BfdB and BfdM remained inactive, BfdC now displayed an enzymatic activity of 22 U/ml with benzoylformate as the substrate. Interestingly, BfdC and BfdB are nearly identical, except for the number and type of the carboxy-terminal amino acids, which therefore seem to determine the functional expression of benzoylformate decarboxylases (Fig. 3).
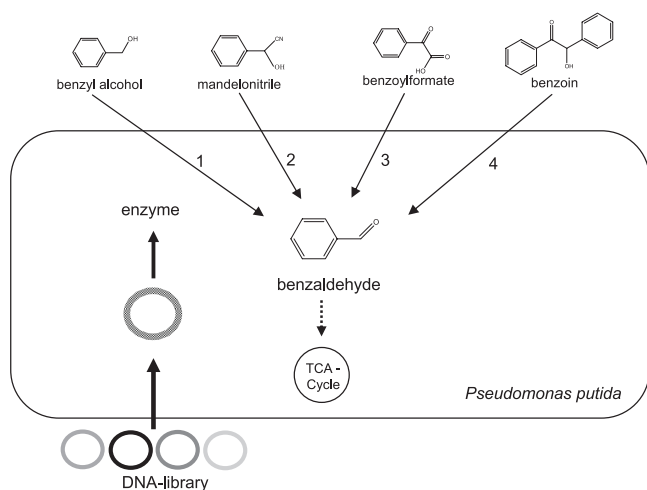
FIG. 4. Growth selection on benzaldehyde to identify industrially important biocatalysts from genomic- or metagenomic-DNA libraries. Accessible enzymes include (i) alcohol dehydrogenase, (ii) hydroxynitrile lyase, (iii) benzoylformate decarboxylase, and (iv) benzaldehyde lyase.

Conclusions. We have described here an efficient bacterial selection system which is based on a *P. putida* strain growing with benzaldehyde as the sole carbon and energy source. The selective capacity of this system was demonstrated by the successful identification of three novel BFDs, using benzoylformate as a model substrate. Apart from identifying BFDs, this system is more generally applicable, because benzaldehyde is not only part of the mandelate pathway but is also formed during phenylalanine degradation (25), in the veratryl alcohol pathway (19), from toluene degradation (32) and by cyanogenesis in plants (5). Therefore, the system we have constructed can be used to identify a broad range of different enzymes, including the industrially important enzymes benzaldehyde lyase, benzoylformate decarboxylase (33), hydroxynitrile lyase (5), and alcohol dehydrogenase, which all produce benzaldehyde by conversion of benzoin, benzoylformate, mandelonitrile, or benzyl alcohol, respectively (Fig. 4).

Furthermore, as most enzymes also catalyze reverse reactions, this method can presumably be employed to identify enzymes that use benzaldehyde as the substrate. This compound serves as a precursor to synthesize important chiral building blocks for organic chemistry, e.g., mandelonitrile or ephedrine. If relaxed substrate specificities and putative moonlighting activities of enzymes are taken into account, a rather broad synthetic-substrate spectrum should be accessible by the biocatalysts newly discovered with this selection method. Furthermore, the method is directly linked to the activity of the respective enzyme, thereby constituting a powerful tool to (i) identify new enzymes from very large libraries and (ii) provide experimental proof of enzymatic activities of putative enzyme-encoding ORFs identified within numerous genome and metagenome sequencing projects.

REFERENCES

1. **Baker, K., C. Bleczinski, H. Lin, G. Salzar-Jiminez, D. Sengupta, S. Kranse, and V. W. Cornish.** 2002. Chemical complementation: a reaction-independent genetic assay for enzyme catalysis. Proc. Natl. Acad. Sci. USA **99:** 16537–16542.
2. **Barrowman, M. M., and C. A. Fewson.** 1985. Phenylglyoxylate decarboxylase and phenylpyruvate decarboxylase from *Acinetobacter calcoaceticus*. Curr. Microbiol. **12:**235–239.
2a. **Barrowman, M. M., W. Harnett, A. J. Scott, C. A. Fewson, and J. R. Kusel.** 1986. Immunological comparison of microbial TPP-dependent nonoxidative alpha-keto acid decarboxylases. FEMS Microbiol. Lett. **34:**57–60.
3. **Buskirk, A. R., A. Landrigan, and D. R. Liu.** 2004. Engineering a ligand-dependent RNA transcriptional activator. Chem. Biol. **11:**1157–1163.
4. **Drepper, T., S. Gross, A. F. Yakunin, P. C. Hallenbeck, B. Masepohl, and W. Klipp.** 2003. Role of GlnB and GlnK in ammonium control of both nitrogenase systems in the phototrophic bacterium *Rhodobacter capsulatus*. Microbiology **149:**2203–2212.
5. **Fechter, M. H., and H. Griengl.** 2004. Hydroxynitrile lyases: biological sources and application as biocatalysts. Food Technol. Biotechnol. **42:**287–294.
6. **Figurski, D. H., and D. R. Helinski.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in *trans*. Proc. Natl. Acad. Sci. USA **76:**1648–1652.
7. **Gamper, M., D. Hilvert, and P. Kast.** 2000. Probing the role of the C-terminus of *Bacillus subtilis* chorismate mutase by a novel random protein-termination strategy. Biochemistry **39:**14087–14094.
8. **Griffiths, J. S., M. Cheriyan, J. B. Corbell, L. Pocivavsek, C. A. Fierke, and E. J. Toone.** 2004. A bacterial selection for the directed evolution of pyruvate aldolases. Bioorg. Med. Chem. **12:**4067–4074.
9. **Hanahan, D.** 1983. Studies on transformation of *Escherichia coli* with plasmids. J. Mol. Biol. **166:**557–580.
10. **Harwood, C. S., and R. E. Parales.** 1996. The β-ketoadipate pathway and the biology of self-identity. Annu. Rev. Microbiol. **50:**553–590.
11. **Hasson, M. S., A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A. Gerlt, G. L. Kenyon, G. A. Petsko, and D. Ringe.** 1998. The crystal structure of benzoylformate decarboxylase at 1.6 Å resolution: diversity of catalytic residues in thiamin diphosphate-dependent enzymes. Biochemistry **37:**9918–9930.
12. **Hegeman, G. D.** 1970. Methods Enzymol. **17:**674–678.
13. **Hegeman, G. D.** 1966. Synthesis of the enzymes of the mandelate pathway by *Pseudomonas putida*. III. Isolation and properties of constitutive mutants. J. Bacteriol. **91:**1161–1167.
14. **Hegeman, G. D.** 1966. Synthesis of the enzymes of the mandelate pathway by *Pseudomonas putida*. I. Synthesis of enzymes by the wild type. J. Bacteriol. **91:**1140–1154.
15. **Hegeman, G. D.** 1966. Synthesis of the enzymes of the mandelate pathway by *Pseudomonas putida*. II. Isolation and properties of blocked mutants. J. Bacteriol. **91:**1155–1160.
16. **Holloway, B. W., V. Krishnapillai, and A. F. Morgan.** 1979. Chromosomal genetics of *Pseudomonas*. Microbiol. Rev. **43:**73–102.
17. **Iding, H., T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir, and M. Pohl.** 2000. Benzoylformate decarboxylase from *Pseudomonas putida* as stable catalyst for the synthesis of chiral 2-hydroxy ketones. Chem. Eur. J. **6:**1483–1495.
18. **Jaeger, K. E.** 2004. Protein technologies and commercial enzymes: white is the hype—biocatalysts on the move. Curr. Opin. Biotechnol. **15:**269–271.
19. **Jensen, K. A., K. M. Evans, T. K. Kirk, and K. E. Hammel.** 1994. Biosynthetic pathway for veratryl alcohol in the ligninolytic fungus *Phanerochaete chrysosporium*. Appl. Environ. Microbiol. **60:**709–714.
20. **Jeon, C. O., W. Park, W. C. Ghiorse, and E. L. Madsen.** 2004. *Polaromonas naphthalenivorans* sp. nov., a naphthalene-degrading bacterium from naphthalene-contaminated sediment. Int. J. Syst. Evol. Microbiol. **54:**93–97.
21. **Jimenez, J. I., B. Minambres, J. L. Garcia, and E. Diaz.** 2002. Genomic analysis of the aromatic catabolic pathways from *Pseudomonas putida* KT2440. Environ. Microbiol. **4:**824–841.
22. **Konarzycka-Bessler, M., and K. E. Jaeger.** 2006. Select the best: novel biocatalysts for industrial applications. Trends Biotechnol. **24:**248–250.
23. **Kovach, M. E., P. H. Elzer, D. S. Hill, G. T. Robertson, M. A. Farris, R. M. Roop II, and K. M. Peterson.** 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. Gene **166:**175–176.
24. **Kovach, M. E., R. W. Phillips, P. H. Elzer, R. M. Roop II, and K. M. Peterson.** 1994. pBBR1MCS: a broad-host-range cloning vector. BioTechniques **16:**800–802.
25. **Lapadatescu, C., C. Ginies, J. L. Le Quere, and P. Bonnarme.** 2000. Novel

scheme for biosynthesis of aryl metabolites from L-phenylalanine in the fungus *Bjerkandera adusta*. Appl. Environ. Microbiol. **66:**1517–1522.

26. **Lingen, B., J. Grötzinger, D. Kolter, M. R. Kula, and M. Pohl.** 2002. Improving the carboligase activity of benzoylformate decarboxylase from *Pseudomonas putida* by a combination of directed evolution and site-directed mutagenesis. Protein Eng. **15:**585–593.

27. **Lingen, B., D. Kolter-Jung, P. Duenkelmann, R. Feldmann, J. Grötzinger, M. Pohl, and M. Müller.** 2003. Alteration of the substrate specificity of benzoylformate decarboxylase from *Pseudomonas putida* by directed evolution. Chembiochemistry **4:**721–726.

28. **MacBeath, G., P. Kast, and D. Hilvert.** 1998. Redesigning enzyme topology by directed evolution. Science **279:**1958–1961.

29. **Mohn, W. W., J. Garmendia, T. C. Galvao, and V. de Lorenzo.** 2006. Surveying biotransformations with a la carte genetic traps: translating dehydrochlorination of lindane (gamma-hexachlorocyclohexane) into *lacZ*-based phenotypes. Environ. Microbiol. **8:**546–555.

30. **Nelson, K. E., C. Weinel, I. T. Paulsen, R. J. Dodson, H. Hilbert, V. A. Martins dos Santos, D. E. Fouts, S. R. Gill, M. Pop, M. Holmes, L. Brinkac, M. Beanan, R. T. DeBoy, S. Daugherty, J. Kolonay, R. Madupu, W. Nelson, O. White, J. Peterson, H. Khouri, I. Hance, P. Chris Lee, E. Holtzapple, D. Scanlan, K. Tran, A. Moazzez, T. Utterback, M. Rizzo, K. Lee, D. Kosack, D. Moestl, H. Wedler, J. Lauber, D. Stjepandic, J. Hoheisel, M. Straetz, S. Heim, C. Kiewitz, J. A. Eisen, K. N. Timmis, A. Düsterhöft, B. Tümmler, and C. M. Fraser.** 2002. Complete genome sequence and comparative analysis of the metabolically versatile *Pseudomonas putida* KT2440. Environ. Microbiol. **4:**799–808.

31. **Nishihara, K., M. Kanemori, H. Yanagi, and T. Yura.** 2000. Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. Appl. Environ. Microbiol. **66:**884–889.

32. **Panke, S., J. M. Sanchez-Romero, and V. de Lorenzo.** 1998. Engineering of quasi-natural *Pseudomonas putida* strains for toluene metabolism through an *ortho*-cleavage degradation pathway. Appl. Environ. Microbiol. **64:**748–751.

33. **Pohl, M., B. Lingen, and M. Müller.** 2002. Thiamin-diphosphate-dependent enzymes: new aspects of asymmetric C-C bond formation. Chem. Eur. J. **8:**5288–5295.

34. **Polovnikova, E. S., M. J. McLeish, E. A. Sergienko, J. T. Burgner, N. L. Anderson, A. K. Bera, F. Jordan, G. L. Kenyon, and M. S. Hasson.** 2003. Structural and kinetic analysis of catalysis by a thiamin diphosphate-dependent enzyme, benzoylformate decarboxylase. Biochemistry **42:**1820–1830.

35. **Regenhardt, D., H. Heuer, S. Heim, D. U. Fernandez, C. Strompl, E. R. Moore, and K. N. Timmis.** 2002. Pedigree and taxonomic credentials of *Pseudomonas putida* strain KT2440. Environ. Microbiol. **4:**912–915.

36. **Robertson, D. E., J. A. Chaplin, G. DeSantis, M. Podar, M. Madden, E. Chi, T. Richardson, A. Milan, M. Miller, D. P. Weiner, K. Wong, J. McQuaid, B. Farwell, L. A. Preston, X. Tan, M. A. Snead, M. Keller, E. Mathur, P. L. Kretz, M. J. Burk, and J. M. Short.** 2004. Exploring nitrilase sequence space for enantioselective catalysis. Appl. Environ. Microbiol. **70:**2429–2436.

37. **Rondon, M. R., P. R. August, A. D. Bettermann, S. F. Brady, T. H. Grossman, M. R. Liles, K. A. Loiacono, B. A. Lynch, I. A. MacNeil, C. Minor, C. L. Tiong, M. Gilman, M. S. Osburne, J. Clardy, J. Handelsman, and R. M. Goodman.** 2000. Cloning the soil metagenome: a strategy for accessing the genetic and functional diversity of uncultured microorganisms. Appl. Environ. Microbiol. **66:**2541–2547.

38. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Habor Laboratory Press, Cold Spring Habor, N.Y.

39. **Schwimmer, L. J., P. Rohatgi, B. Azizi, K. L. Seley, and D. F. Doyle.** 2004. Creation and discovery of ligand-receptor pairs for transcriptional control with small molecules. Proc. Natl. Acad. Sci. USA **101:**14707–14712.

40. **Siegert, P., M. J. McLeish, M. Baumann, H. Iding, M. M. Kneen, G. L. Kenyon, and M. Pohl.** 2005. Exchanging the substrate specificities of pyruvate decarboxylase from *Zymomonas mobilis* and benzoylformate decarboxylase from *Pseudomonas putida*. Protein Eng. Des. Sel. **18:**345–357.

41. **Simon, R., U. Priefer, and A. Pühler.** 1983. A broad host range mobilization system for *in vivo* genetic-engineering—transposon mutagenesis in gram-negative bacteria. Biotechnology **1:**784–791.

42. **Stanier, R. Y.** 1947. Acetic acid production from ethanol by fluorescent pseudomonads. J. Bacteriol. **54:**191–194.

43. **Stanier, R. Y.** 1947. Simultaneous adaptation: a new technique for the study of metabolic pathways. J. Bacteriol. **54:**339–348.

44. **Stover, C. K., X. Q. Pham, A. L. Erwin, S. D. Mizoguchi, P. Warrener, M. J. Hickey, F. S. Brinkman, W. O. Hufnagle, D. J. Kowalik, M. Lagrou, R. L. Garber, L. Goltry, E. Tolentino, S. Westbrock-Wadman, Y. Yuan, L. L. Brody, S. N. Coulter, K. R. Folger, A. Kas, K. Larbig, R. Lim, K. Smith, D. Spencer, G. K. Wong, Z. Wu, I. T. Paulsen, J. Reizer, M. H. Saier, R. E. Hancock, S. Lory, and M. V. Olson.** 2000. Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature **406:**959–964.

45. **Studier, F. W.** 1991. Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system. J. Mol. Biol. **219:**37–47.

46. **Studier, F. W., and B. A. Moffatt.** 1986. Use of bacteriophage T7 RNA polymerase to direct selective high level expression of cloned genes. J. Mol. Biol. **189:**113–130.

47. **Tsou, A. Y., S. C. Ransom, J. A. Gerlt, D. D. Buechter, P. C. Babbitt, and G. L. Kenyon.** 1990. Mandelate pathway of *Pseudomonas putida*: sequence relationships involving mandelate racemase, (*S*)-mandelate dehydrogenase, and benzoylformate decarboxylase and expression of benzoylformate decarboxylase in *Escherichia coli*. Biochemistry **29:**9856–9862.

48. **van Sint Fiet, S., J. B. van Beilen, and B. Witholt.** 2006. Selection of biocatalysts for chemical synthesis. Proc. Natl. Acad. Sci. USA **103:**1693–1698.

49. **Zhou, J., M. A. Bruns, and J. M. Tiedje.** 1996. DNA recovery from soils of diverse composition. Appl. Environ. Microbiol. **62:**316–322.

2.2
Biological Methods

2.2.1
Directed Evolution to Increase the Substrate Range of Benzoylformate
Decarboxylase from *Pseudomonas putida*

Marion Wendorff, Thorsten Eggert, Martina Pohl, Carola Dresen, Michael Müller, and Karl-Erich Jaeger

2.2.1.1 Introduction

Benzoylformate decarboxylase (BFD; EC 4.1.1.7) belongs to the class of thiamine diphosphate (ThDP)-dependent enzymes. ThDP is the cofactor for a large number of enzymes, including pyruvate decarboxylase (PDC), benzaldehyde lyase (BAL), cyclohexane-1,2-dione hydrolase (CDH), acetohydroxyacid synthase (AHAS), and (1*R*,6*R*)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase (SHCHC), which all catalyze the cleavage and formation of C–C bonds [1]. The underlying catalytic mechanism is summarized elsewhere [2] (see also Chapter 2.2.3).

The mandelate pathway, in which mandelate **1** is degraded to benzoic acid **5**, is known in *Pseudomonas* and *Acinetobacter* species enabling these microorganisms to grow on mandelate as a sole carbon source. BFD is the third enzyme in the mandelate pathway (Scheme 2.2.1.1) and catalyzes the nonoxidative decarboxylation of benzoylformate **3** to benzaldehyde **4** and carbon dioxide. Benzaldehyde **4** is then oxidized to benzoic acid, which is further metabolized in the β-ketoadipate pathway and the citric acid cycle. BFD encoded by the gene *mdlC* is located within the tricistronic operon *mdlCBA* together with the genes *mdlA* encoding mandelate racemase and *mdlB* encoding (*S*)-mandelate dehydrogenase [3].

BFD from *Pseudomonas putida* has been characterized in detail with respect to its biochemical properties [4, 5] and 3D structure [6, 7]. Like other enzymes of this class, BFD is a homotetramer with a subunit size of about 56 kDa. The four active sites are formed at the interfaces of two subunits. The structure was published in 2003 [7] and contains the competitive inhibitor (*R*)-mandelate bound to the active sites, allowing model-based predictions about the interactions between active site residues and the substrate.

As a side reaction, BFD catalyzes the carboligation of aldehydes to form chiral 2-hydroxy ketones (Scheme 2.2.1.2) [4, 8]. The physiological role of this additional enzymatic activity is still unknown. The carboligation of benzaldehyde **4**, benzaldehyde derivatives, and acetaldehyde **6** was studied in detail [4]. BFD accepts benzaldehyde **4** as a donor substrate and acetaldehyde **6** as the acceptor substrate. The ligation product 2-(*S*)-hydroxypropiophenone (2-HPP) **7** is formed with an enantiomeric excess (*ee*) of 82–94%, depending on the benzaldehyde concentration and on the reaction temperature [4, 5]. With regard to 2-HPP formation, the related ThDP-dependent enzyme benzaldehyde lyase (BAL) catalyzes the same carboligation reaction but with reverse stereoselectivity [9, 10]. By using these two
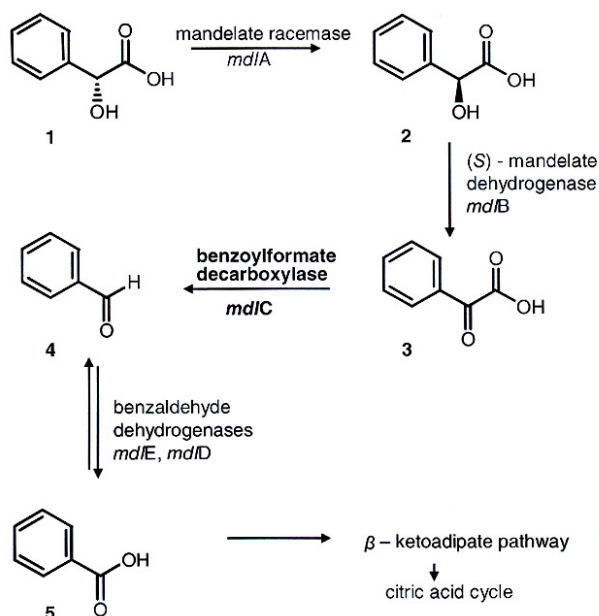
Scheme 2.2.1.1 Mandelate pathway: benzoylformate decarboxylase (BFD), encoded by the gene *mdl*C, catalyzes the conversion of benzoylformate **3** to benzaldehyde **4** and carbon dioxide.
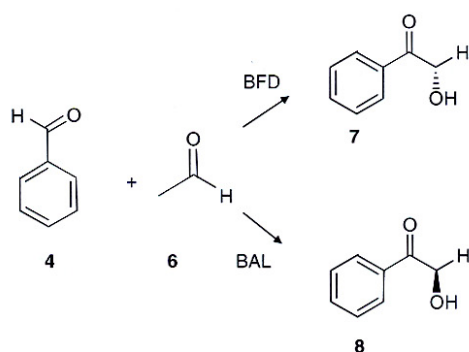
Scheme 2.2.1.2 BFD and BAL as enantiocomplementary catalysts: BFD and BAL act enantiocomplementarily in the formation of 2-HPP **7** and its derivatives, giving access to many 2-HPP analogues in either enantiomeric form.

enantiocomplementary enzymes, many 2-HPP analogues can be synthesized in both enantiomeric forms. The formation of (*R*)-benzoin (*R*)-**12** has also been reported as a ligation product of this reaction which is catalyzed by BFD, albeit with negligible activity [4].

The substrate range of BFD toward 2-substituted benzaldehydes could be increased by directed evolution, yielding a variant with a single amino acid
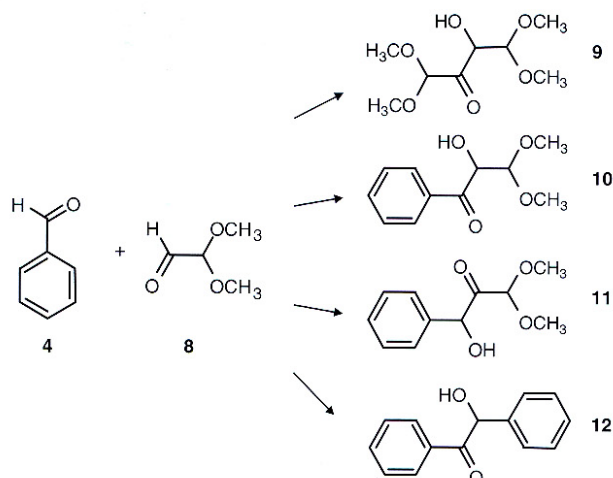
Scheme 2.2.1.3 Screening for novel carboligation activity: four enantiomeric products **9–12** are shown which can be formed starting from the substrates benzaldehyde **4** and dimethoxyacetaldehyde **8**.

exchange, L476Q. This variant also showed a higher stability in the presence of organic solvents [11]. By computer modeling studies amino acid L476 was identified as part of a surface-located loop close to the active site cavity. It was suggested that this residue may play a crucial role for substrate acceptance by exerting a gatekeeper-like function for the BFD active site [11]. At present, acetaldehyde is the only acceptor aldehyde known to be used by BFD variant L476Q.

Here we describe the results of an approach aimed to increase the substrate range of BFD with respect to more complex acceptor aldehydes. Firstly, a high-throughput screening assay had to be established making it possible to screen large libraries of BFD variants generated by directed evolution. Since we were interested in hydroxy ketones with additional functional side-chains, dimethoxyacetaldehyde **8** was used as an acceptor aldehyde which could lead to the carboligation products **9–12** shown in Scheme 2.2.1.3.

2.2.1.2 Materials and Methods

2.2.1.2.1 Reagents
Chemicals used in this work were purchased from Sigma-Aldrich (München, Germany). Restriction enzymes and T4 ligase were purchased from Eurogentec or Fermentas and used as recommended by the manufacturer. Oligonucleotides were obtained from Thermo Electron (Ulm, Germany).

2.2.1.2.2 Construction of Strains for Heterologous Expression of BFD and BAL
The gene encoding BFD variant L476Q was amplified by standard PCR using the upstream 33 bp primer PpBFD_*Nco*I, the downstream 34 bp primer PpBFD_*Hin-*

Table 2.2.1.1 Oligonucleotides used in this study.[a]

Oligoneclotide	Nucleotide sequence (5′ → 3′)	Modification
PpBFD_*Nco*I	ATAT**CCATGG**CTTCGGTACACGGCACCACATAC	*Nco*I
PpBFD_*Hind*III	ATAT**AAGCTT**CTTCACCGGGCTTACGGTGCTTAC	*Hind*III
MP_up_1.PCR	TTGTGGTGACCGT**CCATGG**CGATGATTAC	*Nco*I
MP_down_1.PCR	GGGCGCCGCCATC̲GCGACC	G → C
MP_up_2.PCR	TTGTGGTGACCGTCC	
MP_down_2.PCR	ATAT**CTCGAG**TGCGAAGGGGTCCATGC	*Xho*I
pKK233-2for	CACACAGGAAACAGA**CCATGG**	*Nco*I
pKK233-2rev	TCCGCCAAAACAGCC**AAGCTT**	*Hind*III

    a Restriction sites are highlighted in bold letters, and base substitutions are underlined.

dIII (Table 2.2.1.1), and plasmid pKKBFDL476Q as the template. The unique *Nco*I and *Hind*III restriction sites located in the PCR primers were used for cloning the resulting 1632 bp PCR product into the corresponding restriction sites of the expression plasmid pET28a (Table 2.2.1.2) leading to an in-frame fusion of a C-terminal 6xHis-Tag. The resulting plasmid was named pETBFDL476Q (Table 2.2.1.2).

The gene *bznB* encoding BAL was mutated by site-directed mutagenesis in order to modify an internal *Nco*I restriction site by a silent base exchange. The resulting gene was also cloned into the expression vector pET28a using the restriction sites *Nco*I and *Xho*I, resulting in plasmid pETBAL_*Nco*I. Both plasmids were transformed into the expression host *E. coli* BL21 (DE3), resulting in the expression strains *E.coli* pETBFDL476Q and *E.coli* pETBAL (Table 2.2.1.2).

2.2.1.2.3 Polymerase Chain Reactions

Standard PCR Amplification of DNA fragments was performed in a 50 µL reaction mix containing 1 ng of plasmid DNA as the template, 25 pmol of each primer, 0.2 mM dNTPs, 2.5 U of *Pfu* polymerase (Stratagene, Heidelberg, Germany). The buffer was used as recommended by the manufacturer. Conditions for PCR were as follows: $1 \times$ (5 min, 98 °C); $35 \times$ (1 min, 95 °C; 1 min, 58 °C; 1.5 min, 72 °C); and $1 \times$ (7 min, 72 °C). The PCR reaction was performed using a Mastercycler Gradient (Eppendorf, Hamburg, Germany).

Site-directed mutagenesis Site-directed mutagenesis was performed using megaprimer PCR in order to eliminate the internal *Nco*I restriction site located within the *bznB* gene encoding BAL. Two consecutive PCR reactions had to be carried out. In the first reaction, the primers MP_up_1.PCR and MP_down_1. PCR were used with MP_down_1.PCR carrying a single base mutation resulting in a silent base exchange. MP_up_1.PCR carried an artificial extension upstream of the start codon, producing a docking sequence for the upper primer MP_up_ 2.PCR in the second PCR reaction and thereby avoiding a template switch. The

Table 2.2.1.2 Bacterial strains and plasmids used in this study.

Strain or plasmids	Genotype or description	Reference
Strains		
E.coli BL21 (DE3)	F^- *omp T hsdS$_B$(r_B^- m_B^-) gal dcm (λclts857 ind1 Sam7 nin5 lacUV5-T7 gene 1)*	Novagen, Madison, USA
E.coli BAL	E.coli BL21 (DE3) carrying the plasmid pETBAL_Ncol	this study
E.coli BFDL476Q	E.coli BL21 (DE3) carrying the plasmid pETBFDL476Q	this study
E.coli pET28a	E.coli BL21 (DE3) carrying the empty vector pET28a	this study
Plasmids pET28a	ColE1 P_{T7lac} Kan^r	Novagen, Madison, USA
pKKBFDL476Q	*mdl*CL476Q encoding BFD variant L476Q cloned by *Ncol/Hind*lll into plasmid pKK233-2	ref. [11]
pETBFDL476Q	*mdl*CL476Q cloned by *Ncol/Hind*lll into pET28a, fusion of a C-terminal 6x-His-tag	this study
pETBAL	*bznB* from *Pseudomonas fluorescens* encoding BAL cloned by *Ndel/Xho*l into pET22b	M. Wendorff, unpublished results
pETBAL_Ncol	*bznB* without internal *Ncol* restriction site encoding BAL cloned by *Ncol/Xho*l to pET28a, fusion of a C-terminal H6 tag	this study

second PCR reaction was performed using the megaprimer produced in the first PCR reaction together with MP_up_2.PCR as the upper and MP_down_2.PCR as the lower primer. This PCR reaction yielded as the full length product the *bznB* gene without an internal *Nco*I restriction site. Plasmid pETBAL was used as the template for both PCRs, which were carried out under standard conditions. The annealing temperatures were 60 °C in the first and 55 °C in the second PCR reaction. All the primers are listed in Table 2.2.1.1.

2.2.1.2.4 Generation of a BFD Variant Library by Random Mutagenesis

Random mutagenesis of the gene encoding BFD variant L476Q was performed with the error-prone polymerase chain reaction (epPCR) [11, 14]. Oligonucleotides pKK233-2for and pKK233-2rev (Table 2.2.1.1) were used as primers. An error rate of two to four base substitutions per gene was achieved by using a reaction mix comprising: 5 pmol of each primer, 75 mM Tris/HCl buffer (pH 8.8), 20 mM $(NH_4)_2SO_4$, 1.5 mM $MgCl_2$, 0.5 mM $MnCl_2$, 0.2 mM dNTPs, 0.01% Tween 20, 1 ng template DNA (pKKBFDL476Q) (Table 2.2.1.2), and 10 U Goldstar *Taq*-poly-

merase (Eurogentec, Seraing, Belgium). The PCR reaction was performed as described above (annealing temperature 52 °C) using a Mastercycler Gradient (Eppendorf, Hamburg, Germany). A higher diversity of mutated BFD genes was achieved by mixing mutated genes obtained from three to five different epPCRs and subsequent ligation. The plasmids carrying the mutated DNA fragments were then transformed into the expression host *E.coli* BL21 (DE3) and the bacteria were plated on selective LB medium containing kanamycin (50 μg/ml).

2.2.1.2.5 High-Throughput Screening for Carboligation Activity with the Substrates Benzaldehyde and Dimethoxyacetaldehyde

The mutated BFD genes were cloned into the vector pET28a (Table 2.2.1.2) and the resulting plasmids were transformed into *E.coli* BL21 (DE3) (Table 2.2.1.2). The clones were transferred into 96 deep-well microtiter plates filled with 100 μL LB medium (10 g NaCl, 5 g yeast extract, 10 g trypton) supplemented with 50 μg mL^{-1} kanamycin. After overnight shaking (600 rpm) at 37 °C, 600 μL LB medium containing kanamycin (50 μg mL^{-1}) were added. After an additional 2.5 h of shaking (600 rpm) at 37°C the expression was induced by adding isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 0.8 mM. The induced culture was grown for another 24 h at 37 °C. For storage, a masterplate was generated by transferring the clones on a fresh LB agarplate supplemented with kanamycin (50 μg mL^{-1}). The cells were separated from the culture supernatant by centrifugation at 4000 rpm for 20 min. The supernatant was removed and 300 μL of potassium phosphate buffer (50 mM KP$_i$, pH 7) containing cofactors (1 mM ThDP, 5 mM MgSO$_4$) was added. Cell lysis was achieved by adding 1 mg mL^{-1} lysozyme and incubating at 30 °C for 30 min on a shaker. A 100 μL aliquot of each crude extract was transferred into a fresh 96-well plate, and 100 μL of substrate solution (50 mM KP$_i$, pH 7, 40 mM benzaldehyde, 120 mM dimethoxyacetaldehyde) was added. After 24 h of incubation at 30 °C, the reaction mixture was loaded into a new microtiter plate containing 10 μL of 2,3,5-triphenyltetrazolium chloride (0.4% in EtOH). The color reaction was started by adding 30 μL 3 M NaOH [16].

2.2.1.2.6 Expression and Purification of BFD Variants

An overnight culture of *E. coli* pBFDL476Q (Table 2.2.1.2) containing pBFD55E4 was diluted 1 : 100 to inoculate 3 × 1 L TB medium (12 g tryptone, 24 g yeast extract, 17 mM KH$_2$PO$_4$, 72 mM K$_2$HPO4, 4 mL glycerin) supplemented with 50 μg mL^{-1} kanamycin in a 5 L Erlenmeyer flask. At an optical density of OD$_{580}$ = 1, IPTG was added to a final concentration of 1 mM. The culture was incubated for 24 h at 37 °C. The cells were harvested by centrifugation (Hettich Rotina 35R, 5000 rpm, 20 min), and potassium buffer (50 mM KP$_i$ pH 6, 2.5 mM MgSO$_4$, 0.5 mM ThDP) was added to prepare a 30% w/v solution. Cells were disrupted by ultrasonication at 140 W for 10 × 60 s and the insoluble components were separated by centrifugation (Sorvall RC-5B, SS34, 16 000 rpm, 30 min). The BFD variants were purified from the crude extracts using Ni-NTA (Qiagen, Hilden, Germany) and subsequent gel filtration chromatography (Amersham

Biosciences, Uppsala, Sweden). The purification was performed as described elsewhere [11]. Finally, the purified enzyme variants were lyophilized and stored at −20 °C.

2.2.1.2.7 Protein Analysis Methods

Protein determination Protein concentrations were determined according to Bradford [12] using BSA as a standard.

SDS-polyacrylamide gel electrophoresis Protein samples were analyzed under denaturing conditions in a discontinuous gel system as described by Laemmli [13] using a 5% stacking gel and a 12% separating gel in a vertical gel system (BioRad, Mini Protean II, CA, USA).

2.2.1.2.8 Enzyme Activity Assays

Decarboxylase activity The decarboxylation of benzoylformate was studied using a coupled enzyme test as described previously [4].

Carboligase activity toward benzaldehyde and dimethoxyacetaldehyde Two different assays were performed: (1) 0.6 mg of purified BFD variant L476Q or 0.5 mg of purified BFD variant 55E4 were incubated in 1.2 mL of 50 mM KP_i (pH 7.0), containing 0.1 mM ThDP, 2.5 mM $MgSO_4$ with 20% DMSO in the presence of 5 or 20 mM benzaldehyde at 30 °C and shaking at 300 rpm. After 19 h the reaction was stopped by extraction with 300 µL ethyl acetate. (2) 0.6 mg purified BFD variant L476Q or 0.5 mg BFD variant 55E4 were incubated in 1.2 mL of 50 mM KP_i (pH 7.0) containing 0.1 mM ThDP, 2.5 mM $MgSO_4$ with 20% DMSO in the presence of 5 or 20 mM benzaldehyde and 60 or 500 mM dimethoxyacetaldehyde at 30 °C and shaking at 300 rpm. After four days the reaction was stopped by extraction with 200 µL ethyl acetate.

The concentrations of benzaldehyde, the mixed product, and benzoin and the enantiomeric excesses (*ee* values) were determined by chiral-phase HPLC with a photodiode array detector. Chiral-phase HPLC was performed on a Chiracel OD-H (Daicel, Düsseldorf, Germany) using isohexane/isopropanol (90 : 10) as eluent, a flow rate of 0.5 mL min^{-1} and a column oven at 40 °C. The retention time for benzaldehyde was 10.2 min, for the mixed product DMA–HPP 15.1 and 16.4 min, for the (*S*)-benzoin 20.3 min, and for the (*R*)-benzoin 28.5 min.

2.2.1.3 Results and Discussion

2.2.1.3.1 Overexpression of BFD in *Escherichia coli*

In previous studies on benzoylformate decarboxylase (BFD) from *Pseudomonas putida*, the enzyme was expressed using the pKK233-3 vector system and *E. coli* SG13009 as the expression host. Here, expression was controlled by two plasmids,

namely the expression plasmid pKK233-2 and plasmid pRep4 [4]. In an attempt to avoid experimental problems occurring in directed evolution and high-through-put screening, we have constructed novel expression strains based on the pET system, thereby superseding the presence of a second plasmid which might cause problems during the isolation of the expression plasmid.

The gene encoding BFD variant L476Q was cloned into the *NcoI*/*Hind*III sites of pET28a leading to plasmid pETBFDL476Q, which was transferred into *E. coli* BL21(DE3) by transformation giving the parental clone for the random mutagenesis experiment. Benzaldehyde lyase (BAL) from *Pseudomonas fluorescens* is known to catalyze the carboligation of benzaldehyde and dimethoxyacetaldehyde; hence, a BAL-expressing *E. coli* BL21 (DE3) strain was constructed (Table 2.2.1.2) to serve as a positive control for the high-throughput screening assay. For BAL being cloned into pET28a, an internal *NcoI* restriction site had to be eliminated first by site-directed mutagenesis. Afterward the modified gene was cloned into pET28a by flanking *NcoI* and *XhoI* restriction sites.

2.2.1.3.2 Random Mutagenesis of BFD Variant L476Q

BFD variant L476Q was chosen as the parental enzyme for optimization by directed evolution. As compared to wild-type BFD, variant L467Q showed an increased substrate range toward 2-substituted benzaldehydes [14]. Error-prone PCR (epPCR) was used for the creation of a random mutant library. The entire gene encoding BFD variant L476Q was subjected to mutation, as previous studies had demonstrated that amino acids which affect given enzyme properties may not be located exclusively in the vicinity of the active site but occur scattered over the entire enzyme [15]. The reaction conditions for epPCR were adjusted by using increased $MgCl_2$ and $MnCl_2$ concentrations, resulting in one to three amino acid exchanges per BFD subunit.

2.2.1.3.3 Development of a High-Throughput Screening Assay for Carboligase Activity

Growth conditions in deep-well microtiter plates were optimized with respect to optimal expression of active enzymes (Fig. 2.2.1.1). The best results were obtained with an expression time of 20 h at 37 °C (Fig. 2.2.1.1, lanes 7–9). Subsequently, *E. coli* cells were enzymatically disrupted by lysozyme treatment, and the carboligase activity was monitored by a modified tetrazolium salt color assay [16]. This color assay is based on the reduction of the 2,3,5-triphenyltetrazolium chloride (TTC) **13** to the corresponding formazan **15**, which has an intense red color (Fig. 2.2.1.2A). Before screening of a BFD variant library, substrates and products were tested in the color assay. Neither substrate, benzaldehyde **4** nor dimethoxyacetaldehyde **8**, reduced TTC **13**; however, the product 2-hydroxy-3,3-dimethoxy-propiophenone **10** already caused color formation at low concentrations of 2.5–10 mM (Fig. 2.2.1.2B). Benzoin **12** as the product also gave a color change at a similar concentration (data not shown).
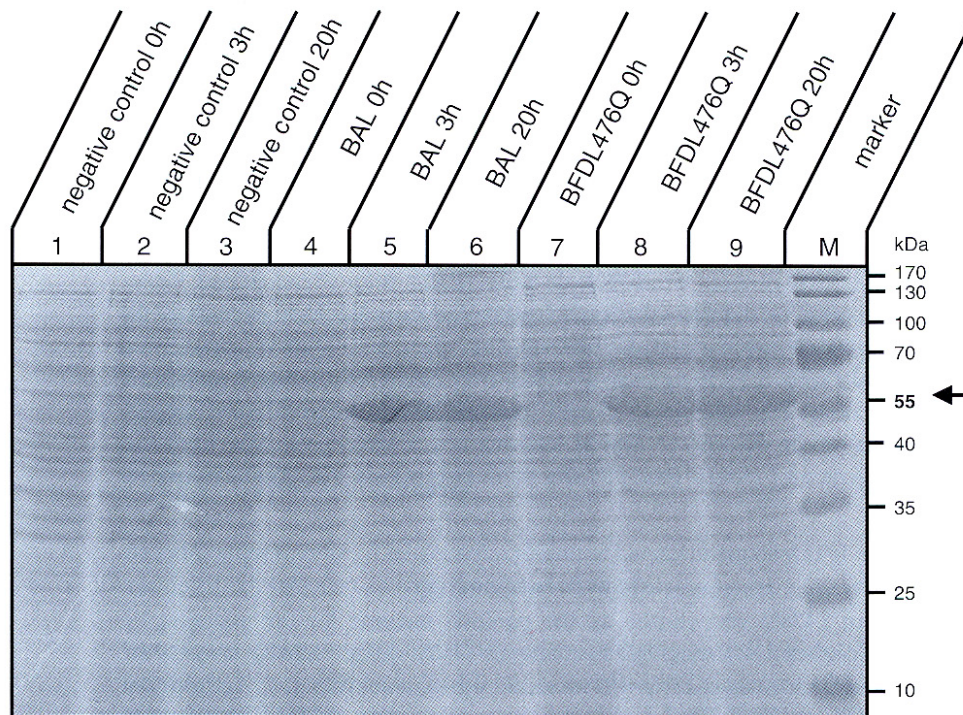
Fig. 2.2.1.1 SDS-PAGE analysis of cell extracts from *E. coli* overexpressing BAL and BFD. The expression cultures were induced (0 h) and grown for 3 and 20 h. Cell extracts (corresponding to $OD_{580} = 0.15$ of the bacterial culture) were harvested by centrifugation, and the proteins were separated on a 12% (w/v) polyacrylamide gel, and stained with Coomassie brilliant blue. The protein bands representing either BFD or BAL are marked with an arrow. Cell extracts from *E. coli* pET28, which served as a negative control, did not show a prominent protein band whereas the expression strains *E. coli* BAL and *E. coli* BFDL476Q showed a protein band of similar size after expression times of 3 h and 20 h.

Additionally, we tested cell extracts in the colorimetric assay: both negative control strains *E. coli* pET28a containing the empty vector and *E. coli* BFDL476Q (Table 2.2.1.2) did not show any enzyme activity in the TTC assay, whereas *E. coli* BAL (Table 2.2.1.2) led to the formation of an intense red color (data not shown).

2.2.1.3.4 Identification of a BFD Variant with an Optimized Acceptor Aldehyde Spectrum

The BFD mutant library generated by epPCR was expressed in microtiter plates and 8000 clones were subjected to the carboligation assay with benzaldehyde **4** and dimethoxyacetaldehyde **8** as the substrates. The reaction was incubated for 24 h at 30 °C as suggested from experiments using the positive control strain

(A)

(B)
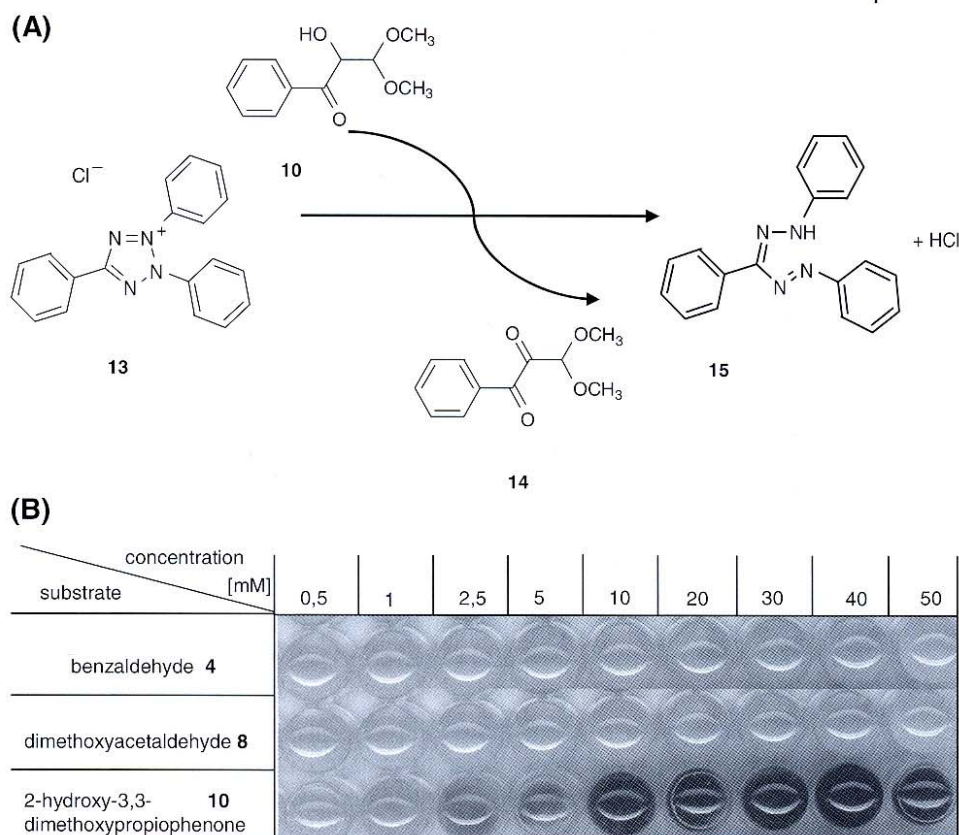
Fig. 2.2.1.2 Colorimetric assay for carboligase activity. A)
DMA–HPP **10** reduces 2,3,5-triphenyltetrazolium **13** chloride
to the respective formazan **15**, which has an intense red
color. B) Formation of the red formazan **15** dye can be
observed only in the presence of DMA–HPP **10**; neither
substrate dimethoxyacetaldehyde **8** nor benzaldehyde **4**
causes any change in color.

E.coli pETBAL (Table 2.2.1.2). Clone 55E4 showed activity with these substrates;
therefore, the respective DNA insert was sequenced, revealing four base ex-
changes in addition to the CTG to CAG mutation which was present in the pa-
rental BFD variant L476Q. Two of the mutations resulted in amino acid exchanges:
(1) the base triplet GCT at amino acid position 255 was mutated to GGT, now
encoding glycine instead of alanine; and (ii) base triplet ATT at amino acid posi-
tion 260 was mutated to ACT, exchanging isoleucine for threonine. In addi-
tion, two silent mutations were located at codons CTG (position 387) and GGT
(position 491).

Both amino acid exchanges are located in the β-domain of the subunits close to the active site channel. The Ala$_{255}$Gly exchange is found right at the beginning of an α-helix, whereas the Ile$_{260}$Thr exchange is in the middle of the same α-helix. Both amino acids are located close to the substrate channel leading to the active site of BFD, so a direct effect on the active site functionality is therefore feasible. Modeling studies initiated in collaboration with Jürgen Pleiss's group (Institute of Technical Biochemistry, University of Stuttgart, Germany) will give more detailed insights into potential effects of these mutations.

2.2.1.3.5 Biochemical Characterization of the BFD Variants

Both BFD variants L476Q and 55E4 were overexpressed in *E. coli* BL21 (DE3) as His-tag fusion proteins and purified to electrophoretic homogeneity as detected by SDS-PAGE by metal chelate affinity chromatography using Ni-NTA (data not shown).

2.2.1.3.6 Decreased Benzoyl Formate Decarboxylation Activity of Variant 55E4

The influence of the mutations on the decarboxylation activity was investigated using the natural substrate benzoylformate **3** (Scheme 2.2.1.1). BFD variant L476Q showed a 1.4-fold higher decarboxylase activity (324 U mg^{-1}) toward benzoylformate **3** than variant 55E4 (233 U mg^{-1}).

2.2.1.3.7 Formation of 2-Hydroxy-3,3-dimethoxypropiophenone and Benzoin

A carboligation reaction with two different aldehydes as substrates may result in the formation of four different 2-hydroxy ketones (Scheme 2.2.1.3). Since the catalysis of benzoin **12** formation is known to be a weak side-activity of BFD variant L476Q, dimethoxyacetaldehyde was applied in excess (1 : 3 and 1 : 100) in order to suppress any possible benzoin formation. First investigations concerning the carboligase potential of variant 55E4 were performed in 1.2 mL batch reactions using the same substrate mixture as in the high-throughput screening assay (20 mM benzaldehyde, 60 mM dimethoxyacetaldehyde), and the reaction products were analyzed after four days by HPLC. The results showed that both DMA–HPP **10** and benzoin **12** were formed in the biotransformations and variant 55E4 showed higher enzyme activity. An estimate of the product yields revealed a 20-fold excess of DMA–HPP **10** obtained with variant 55E4 (Fig. 2.2.1.3).

Further studies, with the substrate ratio altered from 1 : 3 to 1 : 100, were performed in order to suppress the formation of benzoin **12** and increase the yield of DMA–HPP **10**. Although this strategy was successful, the results showed that even a 100-fold excess of dimethoxyacetaldehyde **8** (500 mM) relative to benzaldehyde **4** (5 mM) could not completely supress benzoin **12** formation catalyzed by variant 55E4. Furthermore, the formation of the mixed product DMA–HPP **10** was decreased by a factor of 2 relative to the 1 : 3 substrate mixture. Variant 55E4 showed a 55-fold higher productivity with respect to the formation of DMA–HPP **10** under these conditions as compared to variant L476Q. The overall carboligation
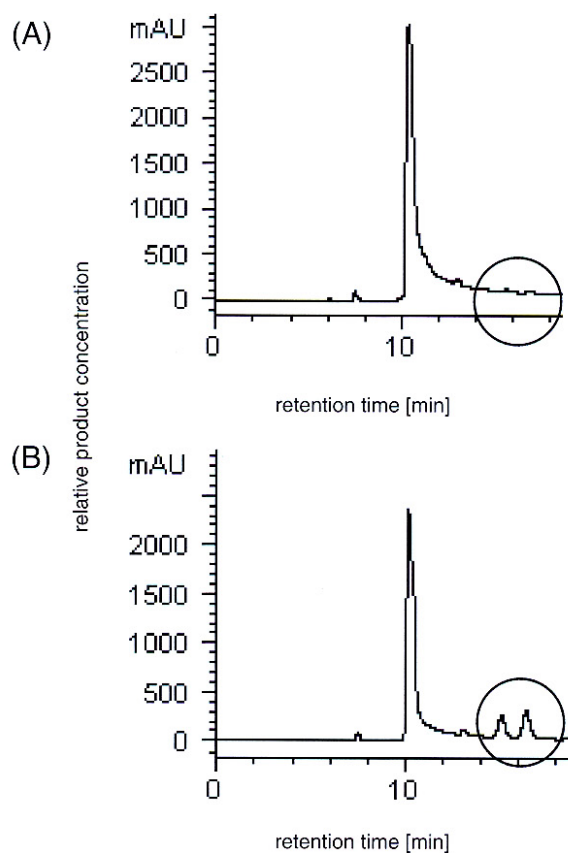
Fig. 2.2.1.3 HPLC diagram showing the formation of DMA–HPP catalyzed by variant BFD55E4. BFD variant enzymes were incubated with benzaldehyde **4** (20 mM) and dimethoxyacetaldehyde **8** (60 mM) for four days, and product formation was analyzed by chiral HPLC (see circles). A) BFD variant L476Q catalyzed the substrate conversion to DMA–HPP **10** only at a neglegible rate. B) BFD variant 55E4 catalyzed the formation of both enantiomers of DMA–HPP **10** with retention times of approx. 15.1 min for the (*S*)- and approx. 16.4 min for the (*R*)-enantiomer.

activity of BFD variant L476Q was hardly detectable. Also, the stability of both variants decreased in the presence of high aldehyde concentrations.

Benzoin formation was also investigated with benzaldehyde **4** as the substrate (at concentrations of 5 and 20 mM, respectively). Again, variant 55E4 showed a significantly higher benzoin-forming activity than variant L476Q (Table 2.2.1.3), which was even more pronounced at low benzaldehyde concentrations (5 mM), suggesting that the variant 55E4 has a higher affinity for benzaldehyde **4** than BFD variant L476Q. This observation coincides with the fact that benzoin forma-

Table 2.2.1.3 Synthesis of (*R*)-benzoin catalyzed by benzoylformate decarboxylase variants.

Substrates	BFD variant 55E4		BFD variant L476Q	
	Yield[a] [%]	ee[a] [%]	Yield[a] [%]	ee[a] [%]
5 mM benzaldehyde	48	99	3	>99
20 mM benzaldehyde	60	97	15	>99
20 mM benzaldehyde + 60 mM dimethoxyacetaldehyde	70	85	12	>99
5 mM benzaldehyde + 500 mM dimethoxyacetaldehyde	12	76	n.d.[b]	n.d.

a Benzaldehyde concentrations and the enantiomeric excesses (*ee*) of (*R*)-benzoin were determined by chiral-phase HPLC. Yields were calculated from amounts of benzaldehyde converted. b n.d.: the amount of benzoin formed was too low to calculate yield or *ee*.

tion is still catalyzed by variant 55E4 even in the presence of excess dimethoxy-acetaldehyde **8** (see above).

2.2.1.3.8 Enantioselectivity of the Carboligation Reaction

Variant 55E4 proved to be more sensitive toward increasing aldehyde concentrations than variant L476Q in terms of the enantioselectivity of the (*R*)-**12** formation. The enantiomeric excess (*ee*) of the (*R*)-**12** products are summarized in Table 2.2.1.3. At the lowest benzaldehyde concentration (5 mM), variant 55E4 catalyzed the formation of enantiomerically pure (>99%) (*R*)-**12**. With increasing total aldehyde concentration, the *ee* decreased continuously to 76% at a total concentration of 505 mM (5 mM benzaldehyde and 500 mM acetaldehyde). Remarkably, increasing aldehyde concentrations had no effect on the enantioselectivity of the parental variant L476Q yielding enantiopure (*R*)-**12** independently of the reaction conditions. Surprisingly, variant 55E4 catalyzed the formation of both (*R*)- and (*S*)-DMA–HPP **10** in equal amounts as determined after four days of reaction time. As the enantioselectivity of (*R*)-**12** formation also decreased under these conditions (see above), it is at present difficult to explain the lack of enantioselectivity toward the product DMA–HPP **10**; it may represent an intrinsic property of variant 55E4 or it may result from destabilizing effects caused by the aldehyde which may finally lead to a distortion of the active site.

Effects of high concentrations of benzaldehyde **4** on the enantioselectivity of the carboligation reaction have been described previously [4, 5]. Siegert et al. described the increase of (*R*)-**7** for a BFD variant by increasing the concentration of benzaldehyde **4** from 1 to 10 mM, and Iding et al. were able to increase the enantioselectivity of wild-type BFD by keeping the concentration of benzaldehyde **4** at a low level. Surprisingly, acetaldehyde **6** showed no such effect. More recently, we

have observed that the substrate dimethoxyacetaldehyde **8** affects the enantiose-
lectivity of variant 55E4 in a similar way.

2.2.1.4 **Conclusions**

The new BFD variant 55E4 showed an increased enzyme activity and substrate
range for acceptor aldehydes. Dimethoxyacetaldehyde **8** was accepted in carboliga-
tion reactions with benzaldehyde **4** as the donor aldehyde. The formation of
DMA–HPP **10** was enhanced by a factor of 55 as compared to BFD variant L476Q,
which was used as the parental enzyme for directed evolution. Furthermore, our
studies clearly indicated that the enantioselectivity of the reaction was influenced
by the total aldehyde concentration in the reaction and thus can be increased by
alteration of the reaction conditions. At present, we are trying to correlate by mo-
lecular modeling the observed effects with the structural consequences of the
amino acid exchanges that were randomly introduced into the BFD variant 55E4
by directed evolution.

References

1 M. Pohl, G. A. Sprenger, M. Müller, *Curr. Opin. Biotechnol.* **2004**, *15*, 335–342.

2 F. Jordan, *Nat. Prod. Rep.* **2003**, *20*, 184–201.

3 A. Y. Tsou, S. C. Ransom, J. A. Gerlt, D. D. Buechter, P. C. Babbitt, G. L. Kenyon, *Biochemistry* **1990**, *29*, 9856–9862.

4 H. Iding, T. Dünnwald, L. Greiner, A. Liese, M. Müller, P. Siegert, J. Grötzinger, A. S. Demir, M. Pohl, *Chem. Eur. J.* **2000**, *6*, 1483–1495.

5 P. Siegert, M. J. McLeish, M. Baumann, H. Iding, M. M. Kneen, G. L. Kenyon, M. Pohl, *Protein Eng. Des. Sel.* **2005**, *18*, 345–357.

6 M. S. Hasson, A. Muscate, M. J. McLeish, L. S. Polovnikova, J. A. Gerlt, G. L. Kenyon, G. A. Petsko, D. Ringe, *Biochemistry* **1998**, *37*, 9918–9930.

7 E. S. Polovnikova, M. J. McLeish, E. A. Sergienko, J. T. Burgner, N. L. Anderson, A. K. Bera, F. Jordan, G. L. Kenyon, M. S. Hasson, *Biochemistry* **2003**, *42*, 1820–1830.

8 R. Wilcocks, O. P. Ward, S. Collins, N. J. Dewdney, Y. Hong, E. Prosen, *Appl.* *Environ. Microbiol.* **1992**, *58*, 1699–1704.

9 A. S. Demir, M. Pohl, E. Janzen, M. Müller, *J. Chem. Soc. Perkin. Trans.* **2001**, *1*, 633–635.

10 A. S. Demir, O. Sesenoglu, E. Eren, B. Hosrik, M. Pohl, E. Janzen, D. Kolter, R. Feldmann, P. Dünkelmann, M. Müller, *Adv. Synth. Catal.* **2002**, *344*, 96–103.

11 B. Lingen, J. Grötzinger, D. Kolter, M. R. Kula, M. Pohl, *Protein Eng.* **2002**, *15*, 585–593.

12 M. M. Bradford, *Anal. Biochem.* **1976**, *72*, 248–254.

13 U. K. Laemmli, *Nature* **1970**, *227*, 680.

14 B. Lingen, D. Kolter-Jung, P. Dünkelmann, R. Feldmann, J. Grötzinger, M. Pohl, M. Müller, *ChemBioChem* **2003**, *4*, 721–726.

15 M. Bocola, N. Otte, K. E. Jaeger, M. T. Reetz, W. Thiel, *ChemBioChem* **2004**, *5*, 214–223.

16 M. Breuer, M. Pohl, B. Hauer, B. Lingen, *Anal. Bioanal. Chem.* **2002**, *374*, 1069–1073.

ELSEVIER

Enantioselective biocatalysis optimized by directed evolution
Karl-Erich Jaeger and Thorsten Eggert

Directed evolution methods are now widely used for the optimization of diverse enzyme properties, which include biotechnologically relevant characteristics like stability, regioselectivity and, in particular, enantioselectivity. In principle, three different approaches are followed to optimize enantioselective reactions: the development of whole-cell biocatalysts through the creation of designer organisms; the optimization of enzymes with existing enantioselectivity for process conditions; and the evolution of novel enantioselective biocatalysts starting from non-selective wild-type enzymes.

Addresses
Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Duesseldorf, Forschungszentrum Juelich, D-52426 Juelich, Germany
e-mail: karl-erich.jaeger@fz-juelich.de

Abbreviations
BFD	benzoylformate decarboxylase
BSLA	*Bacillus subtilis* lipase A
ee	enantiomeric excess
epPCR	error-prone polymerase chain reaction
ESI-MS	electrospray ionisation mass spectrometry
L-KDO	3-deoxy-L-manno-oct-2-ulosonic acid

Introduction
Directed evolution has matured to a standard technology in the field of molecular enzyme engineering. This development was triggered by rapid and significant improvements in the available novel random mutagenesis methods to generate large enzyme libraries [1•] and by progress in high-throughput screening or selection methods to identify better enzyme variants [2,3•,4]. During the past decade, about 1100 scientific articles dealing with directed enzyme evolution have been published (Figure 1), accompanied by several hundred patents reflecting the biotechnological relevance of the newly created enzyme variants.

Researchers in academia and industry have evolved enzyme characteristics as varied as activity, stability and substrate specificity [5,6]. Directed evolution meth-ods have also been applied to improve more difficult enzyme properties, like enantioselectivity. The stereo-selective synthesis of chiral organic compounds is of immense academic and industrial interest. The 'chiral market' for enantiomerically pure or enriched organic compounds continues to expand rapidly, and total sales of chiral pharmaceuticals exceeded US $100 billion in 2000. Therefore, enantioselective biocatalysis is a key issue in the production of fine chemicals for use in pharmaceuticals, agrochemicals or cosmetics. In 1997, in collaboration with the group of Reetz (Max-Planck Institut für Kohlenforschung, Germany), our group demonstrated the principle of creating an enantioselec-tive enzyme by directed evolution [7]. A lipase from the Gram-negative bacterium *Pseudomonas aeruginosa*, which was essentially non-selective (enantiomeric ratio of the reaction $E = 1.1$) towards the chiral model substrate 2-methyldecanoic acid p-nitrophenyl ester, was subjected to various methods of random mutagenesis and a high-throughput spectrophotometric screening assay was used to identify several variants that were highly enantiose-lective towards both the (S)- ($E = 51$) and the (R)-enantiomer ($E = 30$) of the substrate [8–10]. The X-ray structure of this *P. aeruginosa* lipase was solved in parallel [11] and used to rationalize by molecular modelling the amino acid substitutions that led to increased enantio-selectivity [9,12]. More recently, molecular modeling with classical force fields has been used to study several enantioselective *P. aeruginosa* lipase variants by subject-ing them to 1 ns molecular dynamics simulations. The results allowed rationalization of our previous experimen-tal finding that amino acid substitutions at positions 155 and 162 constituted 'hot spots' for lipase enantioselec-tivity [13••]. Hence, *P. aeruginosa* lipase represents the most thoroughly studied model enzyme to date with respect to evolving enantioselectivity by directed evolu-tion [12,14–17], as outlined in detail in a recent review article [18•].

The number of publications that describe the successful improvement of different enzymes by various directed evolution methods has increased continuously during the past decade, including attempts to evolve enantioselec-tive enzymes (Figure 1). In this article, we summarize recent work carried out to improve enantioselective bio-catalytic processes using directed evolution.

Strategies to improve enantioselective biocatalysis
Many different strategies are used to control chirality dur-ing biocatalysis, including substrate or solvent engineering, changes of reaction conditions, and a combination of

Figure 1
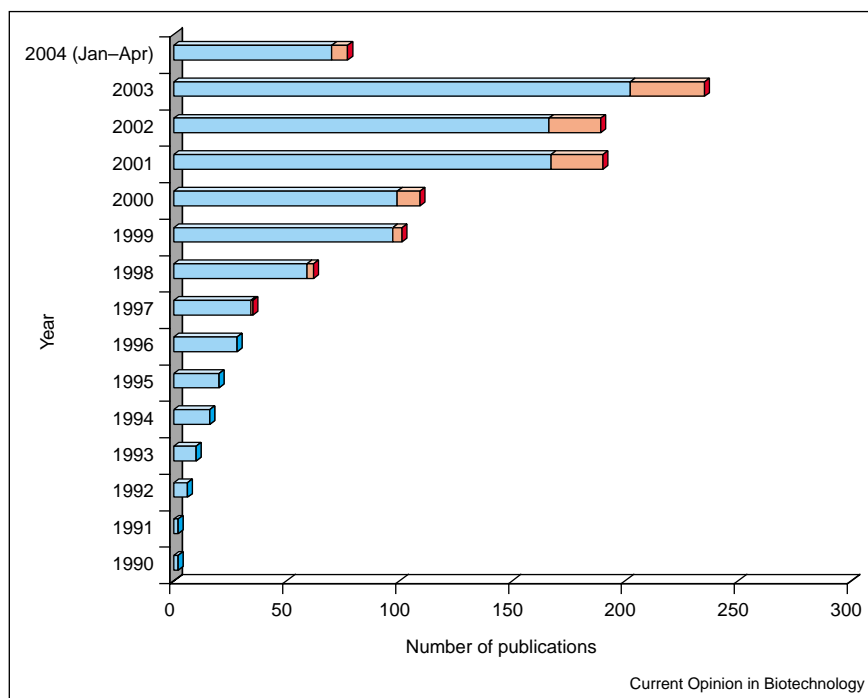
Numbers of scientific publications dealing with enzyme optimization by directed evolution and enantioselectivity. The number of publications per year in the field of directed evolution (blue) and directed evolution used to improve enantioselectivity (red) were retrieved from online databases on 'ISI web of science' (http://isi1.isiknowledge.com). The search terms 'directed evolution' or '*in vitro* evolution' were used to retrieve the number of publications on directed evolution, and, in addition, the search criterea 'enantio*' or 'stereo*' were used to identify publications dealing with enantiomers, enantioselectivity and stereoselectivity.

chemocatalysis and biocatalysis (as summarized in recent review articles [19–21]). All these approaches have in common that the biocatalyst itself is not modified. Instead, the modification of the solvent properties (e.g. polarity, hydrophobicity) or the reaction conditions (e.g. temperature, pH, pressure) results in altered enzyme–substrate interactions, which could sometimes also alter the active-site geometry of the enzyme. By contrast, genetic engineering methods (i.e. site-directed mutagenesis or directed evolution) directly alter at least the primary structure of an enzyme, and often alter the secondary and tertiary structure as well. In the following paragraphs, we will discuss three main strategies (Figure 2) that employ directed evolution as a means to improve biocatalytic routes for the production of enantiopure chemical compounds.

Evolution of whole-cell biocatalysts by creating designer organisms

Since the beginning of industrial biocatalysis, classical strain improvement has been used to create better microorganisms for commercial processes. Cells were treated with physical or chemical mutagens, such as UV light, ethylmethanesulfonate or *N*-methyl-*N′*-nitro-*N*-nitroso-guanidine, and better-performing mutants were identified by screening. Although this strategy is still widely

used in industry, reports describing classical strain improvement with respect to enantioselective biotransformations are rare. Wu *et al.* [22] have used this strategy to evolve the yeast strain *Trichosporon brassicae* with respect to its ability to catalyze the enantioselective hydrolysis of ketoprofen ethyl ester (Figure 3a). The improved mutant strain had an activity 1.8-fold higher than the wild-type in producing the desired (*S*)-enantiomer of ketoprofen. Undoubtedly, a major drawback of classical whole-cell mutagenesis is the unwanted accumulation of mutations, of which many may turn out to be deleterious. Therefore, the Biotech Companies Maxygen Inc. (Redwood City, CA, USA) and its subsidiary Codexis Inc. (Redwood City, CA, USA) have recently introduced a new strategy to speed up classical strain improvement by using whole-genome shuffling [23••,24]. This technique will diminish the accumulation of numerous unwanted point mutations; however, success depends on applying sophisticated screening or selection methods, which allow newly created strains to be identified that fulfil the needs of a given industrial process. Also, at least in Europe, legal restrictions could hamper the construction of such genetically modified organisms. It remains to be demonstrated whether genome shuffling will also prove useful to improve whole-cell biocatalysts for enantioselective biotransformations.

Figure 2
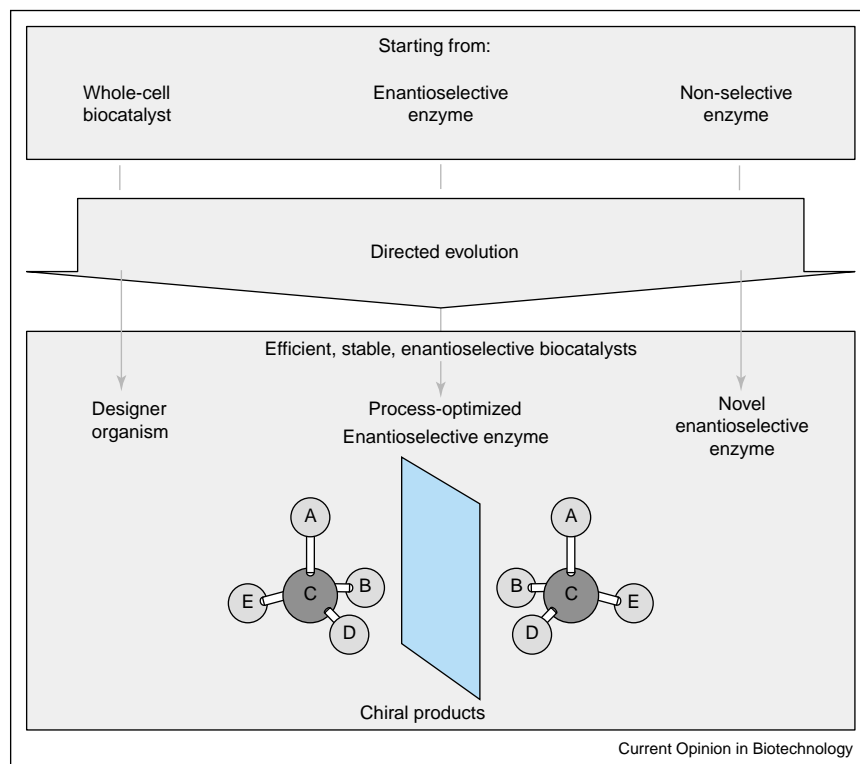
Strategies to optimize enantioselective biocatalytic processes by directed evolution.

Enantioselective whole-cell biocatalysis using designer organisms has been successfully performed through the cloning and expression of single enzyme genes or multiple genes encoding whole pathways in heterologous microbial host strains. By co-expression of the NADPH-dependent aldehyde reductase from *Sporobolomyces salmonicolor* and the cofactor-regenerating enzyme glucose dehydrogenase from *Bacillus megaterium*, an *Escherichia coli* strain was constructed for the stereoselective reduction of ethyl-4-chloro-3-oxobutanoate, a chiral

Figure 3
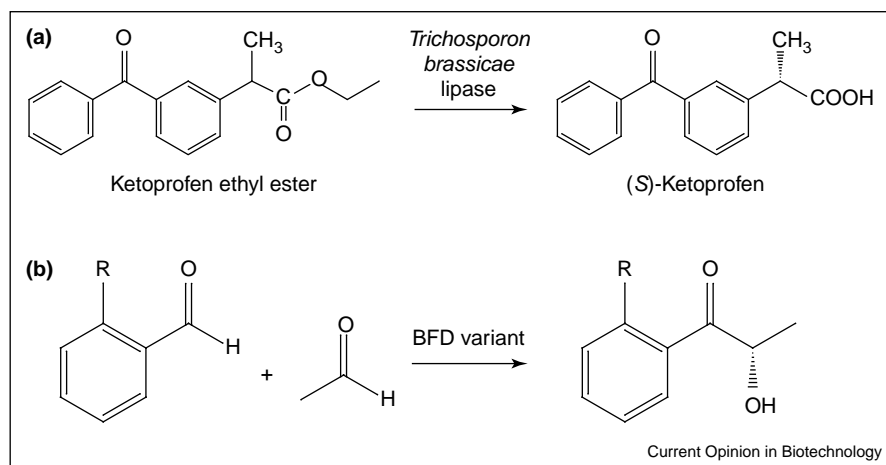
Selected examples of enantioselective biotransformations optimized by directed evolution. **(a)** Lipase-catalyzed asymmetric hydrolysis of ketoprofen ethyl ester to yield (*S*)-ketoprofen using the yeast *Trichosporon brassicae* as whole-cell biocatalyst. **(b)** Formation of enantiopure (*S*)-2-hydroxy-1-(2-methylphenyl)propane-1-one derivatives catalyzed by benzoylformate decarboxylase (BFD) variants Leu476Gln and Met365Leu-Leu461Ser from *P. putida*. R = CH₃, OMe, F, CL or Br.

building block for the chemical synthesis of various compounds [25]. Engineered *E. coli* cells were also used for the synthesis of enantiopure L- and D-amino acids from α-keto acids [26]. The construction of a novel *E. coli* whole-cell catalyst for the production of L-methionine constitutes an elegant example for the creation of a designer organism by combining directed evolution and heterologous gene expression. The key pathway enzyme, D-hydantoinase from *Arthrobacter*, was converted into an L-selective enzyme that also exhibited a fivefold increased activity. The evolved hydantoinase was co-expressed in *E. coli* with an L-*N*-carbamoylase and a hydantoin racemase, resulting in a fivefold increased productivity at 90% substrate conversion [27•].

Evolution of enzymes for process conditions

Enzymes can catalyze reactions with non-natural substrates, sometimes with high enantioselectivities. Unfortunately, these enzymes, which are adapted to their natural niches, often show low activities towards non-natural substrates or low stabilities under process conditions. Therefore, a second strategy to evolve efficient enantioselective biocatalytic processes aims to improve the stability and specific activity of an enzyme for a given chiral reaction under harsh process conditions. The specific activity of a benzoylformate decarboxylase (BFD) from *Pseudomonas putida* was evolved towards *ortho*-substituted benzaldehyde derivatives [28]. Whereas the wild-type enzyme hardly accepts *ortho*-substituted benzaldehyde derivatives like 2-methylbenzaldehyde, two BFD-variants generated by the error-prone polymerase chain reaction (epPCR) were identified in a colorimetric high-throughput screening assay which selectively catalyzed the formation of (*S*)-2-hydroxy-1-(2-methylphenyl)propan-1-one and derivatives thereof (Figure 3b). The localization of amino acid substitutions found in the BFD variants provided first insights into the mechanisms of substrate acceptance and enantioselective carbon–carbon bond formation by thiamine diphosphate-dependent enzymes [28].

Cytochrome P450 BM-3 originating from the Gram-positive bacterium *Bacillus megaterium* was evolved to hydroxylate linear alkanes enantioselectively using dioxygen as an oxidant [29]. This work demonstrated the strategy of evolving complex properties in a two-step process. First, highly active enzyme variants for alkane hydroxylation were generated in five rounds of directed evolution, creating a novel biocatalyst with 11 amino acid substitutions as compared with the wild-type enzyme [30]. In a second round of evolution, this variant was further optimized with respect to its regioselectivity and enantioselectivity [29]. The authors of this study claim that cytochrome P450 BM-3 from *B. megaterium* can be regarded as a 'one-enzyme-fits-all' oxidation catalyst. The backbone of this enzyme seems to be flexible enough to allow the introduction of many amino acid substitutions while still retaining its overall fold.

Evolution of enantioselective enzymes

The creation of an enantioselective enzyme by directed evolution usually starts from a non-enantioselective or moderately enantioselective wild-type enzyme, thus attempting to find a novel biocatalyst. Although neither knowledge of the three-dimensional structure nor of the catalytic mechanism is required, the analysis of enantioselective enzyme variants obtained from directed evolution experiments can also reveal fundamental structure–function relationships, particularly if the three-dimensional structure of the respective enzyme is known. We discuss here recent reports on evolving enantioselective enzymes (see Table 1) mainly covering the time period from January 2003 to April 2004.

Hydrolases

Hydrolases are by far the most prominent enzymes used for industrial biotransformations [31•,32]. Thus, most of the enzymes evolved for enantioselective reactions are hydrolases (Table 1) with three of them belonging to the lipase and esterase family (EC 3.1.1.x), one being an epoxide hydrolase (EC 3.3.2.3) and another a nitrilase (EC 3.5.5.1).

In an ongoing collaboration project with the group of Reetz, we have used directed evolution to improve the enantioselectivity of the *Bacillus subtilis* lipase A (BSLA) in the asymmetric hydrolysis of *meso*-1,4-diacetoxy-2-cyclopentene leading to the formation of chiral alcohols. This reaction does not constitute a kinetic resolution and can thus be carried out to 100% conversion; the wild-type enzyme leads to an enantiomeric excess (*ee*) value of only 38%, favouring the (1*R*,4*S*) enantiomer. Screening is carried out by electrospray ionisation mass spectrometry (ESI-MS) using the deuterium-labelled *pseudo-meso* substrate [33]. A complete site saturation library was constructed in which every single amino acid position in BSLA was exchanged for all possible 19 remaining amino acids, thereby generating a library that contains all theoretically possible variants carrying a single amino acid exchange. In the case of BSLA (181 amino acids) this library includes 5792 mutant genes encoding 3440 different enzyme variants. High-throughput ESI-MS screening identified several different variants with improved (*ee* value of 65% in favour of 1*S*,4*R*) and also inverted (*ee* value of 56% in favour of 1*R*,4*S*) enantioselectivities towards the model substrate [34•]. Recently, a BSLA variant was identified showing a reversed enantioselectivity of *ee* = 85% (SA Funke *et al.*, unpublished), clearly demonstrating that a complete saturation library constitutes a reliable basis for the first round of a directed evolution approach to evolve an enantioselective enzyme. New data from our laboratory also indicate that such a library can successfully be screened towards additional

Table 1

Enantioselective biotransformations performed with enzyme variants created by directed evolution.
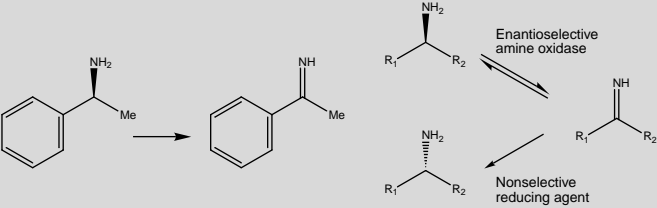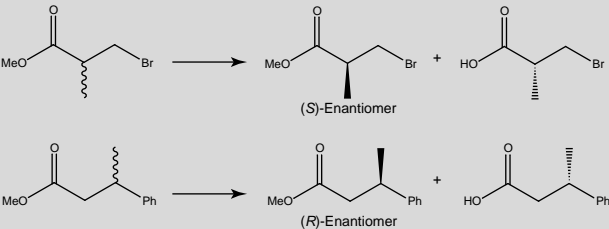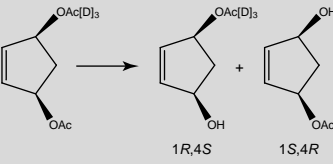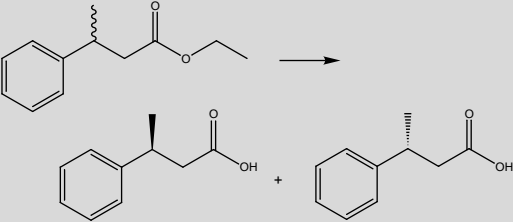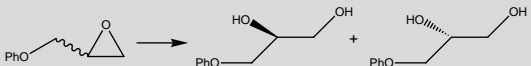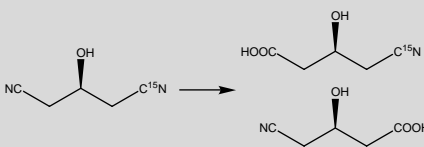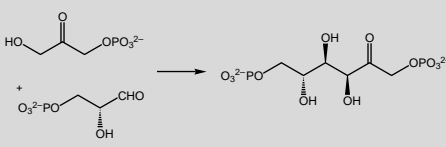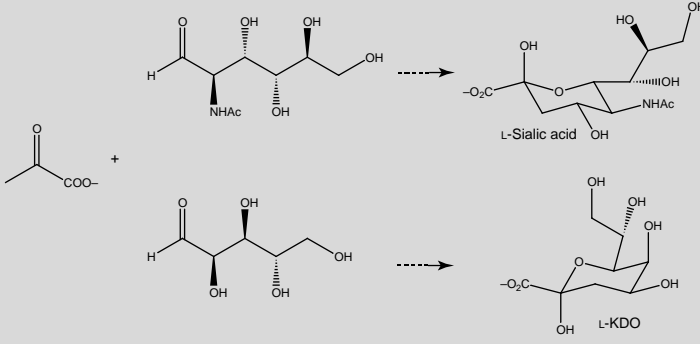
EC	Enzyme/source	Mutagenesis method	Substrate name/enzyme reaction	Screening system	Improvement	Ref.
1.4.3.4	Monoamine oxidase/ *Aspergillus niger*	Mutator strain *E. coli* XL1-red	α-Methylbenzyl amine 100% conversion is possible when using a nonselective reducing agent in addition	Colorimetric	Sixfold increase up to $E = 199$	[44,45,48]
3.1.1.1	Esterase/ *Pseudomonas fluorescens*	Mutator strain *E. coli* XL1-red	3-Bromo-2-methylpropanoate and ethyl 3-phenylbutyrate	Spectrophotometric using the Quick E-screening system [49]	$E = 12$ to $E = 19$ and $E = 3.5$ to $E = 12$	[40]
3.1.1.3	Lipase/ *Bacillus subtilis*	epPCR and complete site saturation mutagenesis	*Pseudo-meso* 1,4-diacetoxy-cyclopentene	ESI-MS	Inversion of enantioselectivity from $ee = 38\%$ ($1S,4R$) to $ee = 56\%$ ($1R,4S$)	[34•]
3.1.1.3	Lipase/ *Burkholderia cepacia*	Combinatorial mutagenesis using SIMPLEX (single-molecule-PCR-linked *in vitro* expression)	(*RS*)-Ethyl 3-phenylbutyrate	Spectrophotometric using the substrate analogue *p*-nitrophenyl 3-phenylbutyrate	$E_S = 33$ $E_R = 38$ (inversion)	[39••]

Table 1 (Continued)

EC	Enzyme/source	Mutagenesis method	Substrate name/enzyme reaction	Screening system	Improvement	Ref.
3.3.2.3	Epoxide hydrolase/ *Aspergillus niger*	epPCR	Glycidyl phenyl ether	ESI-MS	E = 4.6 to E = 10.8	[41]
3.5.5.1	Nitrilase/ metagenome	GSSM (gene site saturation mutagenesis)	^{15}N-(*R*)-3-hydroxyglutaryl nitrile	MS	From *ee* = 87.8 to 98.1 (at high substrate concentrations)	[37•]
4.1.2.40	Tagatose-1, 6-bisphosphate aldolase/ *E. coli*	DNA shuffling	D-Fructose-1,6-bisphosphate as the product	Spectrophotometric	Reversed stereochemistry during C–C bond-forming step. Producing D-fructose-1, 6-bisphosphate instead of D-tagatose-1, 6-bisphosphate	[47••]
4.1.3.3	*N*-Acetylneuraminic acid aldolase/*E. coli*	epPCR, DNA shuffling and saturation mutagenesis	Sialic acid and 3-deoxy-L-manno-oct-2-ulosonic acid (KDO) as the products	Fluorescence screen (NADH, excitation at 340 nm and emission at 450 nm)	Production of L-sialic acid and L-KDO	[46]

substrates, suggesting that the strategy of complete saturation mutagenesis may serve as an 'all-purpose' starting point for directed evolution experiments. The method, which was termed 'gene site saturation mutagenesis' by the Company Diversa [35], offers the unique advantage of unbiased access to all codon variations at every position [36,37•]. For successive rounds of directed evolution, recombinative methods like DNA-shuffling or staggered extension process might prove useful to achieve further improvements by combining previous positive mutations [1•]. Another example of successfully using complete saturation mutagenesis pertains to the evolution of a nitrilase for the enantioselective synthesis of (R)-4-cyano-3-hydroxybutyric acid [37•]. The enzyme, which was identified in a metagenome screening [38], showed an enantioselectivity of ee = 87.8%, which was improved to ee = 98.1% under high substrate conditions.

A new method to create an enantioselective enzyme was applied to a lipase originating from the bacterium *Burkholderia cepacia* [39••]. In a semi-random approach, four amino acid positions located close to the catalytic triad were substituted combinatorially with a limited variety of hydrophobic amino acids, leading to a library theoretically consisting of 2401 variants. Subsequently, single-molecule PCR was performed, followed by *in vitro* transcription and translation to produce the variant enzymes in sufficient amounts, while circumventing time-consuming cloning and transformation steps. This approach may help to overcome problems related to the heterologous expression of enzyme variants that are frequently encountered by many researchers trying to set up directed evolution experiments.

Another lipolytic enzyme, an esterase from *Pseudomonas fluorescens*, was evolved for the enantioselective hydrolysis of the chiral building blocks (S)-3-bromo-2-methyl-propanoate and ethyl-(R)-3-phenylbutanoate using the commercially available mutator strain *E. coli* XL1-red (Stratagene, Germany) and subsequent saturation mutagenesis [40]. Although the mutations identified here were located remote from the active site, the authors suggest that it would be more effective to focus mutagenesis experiments on active-site residues [40].

The enantioselectivity of an epoxide hydrolase from *Aspergillus niger* in the hydrolytic kinetic resolution of glycidyl phenyl ether was doubled from $E = 4.6$ to $E = 10.8$ by just one round of epPCR [41]. The authors describe the lack of an efficient expression system as a major drawback which prevented rapid and significant progress in their directed evolution experiments, thereby pointing to a problem of general importance. The construction of a potent and genetically stable overexpression system working in an appropriate host strain constitutes a necessary prerequisite for devising an efficient directed evolution strategy [42,43].

Oxidoreductases

The substrate specificity of a monoamine oxidase from *A. niger* was broadened and, in parallel, its enantioselectivity was improved. [44,45]. The best-performing variant Asn336Ser was created using the commercially available mutagenesis strain *E. coli* XL1-red in combination with high-throughput screening on agar plates. The use of the enantioselective amine oxidase in combination with a non-selective imine-reducing agent (see Table 1) can lead to 100% conversion and to the production of enantiomerically pure amines starting from racemic mixtures.

Lyases

An *N*-acetylneuraminic acid aldolase (Neu5Ac) from *E. coli* was subjected to directed evolution [46]. The $(\alpha/\beta)_8$-barrel enzyme was altered to completely reverse its enantioselectivity to produce the unnatural *N*-acetyl-L-neuraminic acid (L-sialic acid) and 3-deoxy-L-manno-oct-2-ulosonic acid (L-KDO); both compounds are of significant importance for the pharmaceutical industry. All mutations that affected the catalytic properties of the enzyme were located outside the $(\alpha/\beta)_8$-barrel active site. Interestingly, the crystal structure of a better-performing variant obtained from the second round of random mutagenesis was solved, but significant structural changes were not observed [46].

Another challenging directed evolution project was recently described, which attempted to modify the stereochemistry of an aldolase-catalyzed carbon–carbon bond formation [47••]. In three rounds of directed evolution using DNA shuffling, the *agaY* gene from *E. coli* coding for a tagatose-1,6-bisphosphate aldolase (EC 4.1.2.40) was evolved to a fructose-1,6-bisphosphate aldolase (EC 4.1.2.13), thereby creating a new enzyme that exhibited reversed stereospecificity. This newly evolved property did not correspond to an altered acceptance of unnatural substrates, instead, the stereochemical course of carbon–carbon bond formation was changed starting from the same substrates [47••].

Conclusions

Enantioselective biotransformations are increasingly being considered to manufacture a wide range of chiral intermediates and products. Thus, several strategies are followed in parallel trying to optimize existing enantioselective reactions or to create novel enantioselective biocatalysts. In this process, directed evolution has emerged as a key methodology. Convincing successes have been achieved; however, our ability to understand or ultimately direct and control enantioselective biocatalysis is still very limited, mainly because data are missing that elucidate the structure–function relationship of enantioselective enzymes. Presently, existing data suggest that subtle changes of the active-site geometry of an enzyme might be sufficient to convert a virtually non-selective into a highly enantioselective enzyme. Such changes may

be caused by substitutions of amino acids located either remote from the active site or in contact with the substrate. In any case, the resulting effects may not become obvious in crystal structures obtained from enantioselective enzyme variants, and time-resolved methods based on NMR or Fourier-transform infrared (FTIR) spectroscopy must be employed to unravel the underlying mechanisms. In addition, molecular dynamics simulations as well as quantum mechanical and molecular mechanical calculations will significantly improve our knowledge on the molecular basis of enantioselectivity. There is no doubt that the potential of directed evolution to create novel and enantioselective enzymes will not only trigger a basic understanding of enzyme enantioselectivity, but will also allow designer enzymes to be generated with precisely defined specifications.

Acknowledgements

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

* • of special interest
* •• of outstanding interest

1. Arnold FH, Georgiou G: Directed Evolution Library Creation:
•   Methods and Protocols, vol 231. Totowa, New Jersey: Humana press; 2003.
Excellent collection of recent methods used for directed enzyme evolution described in detailed bench protocols.

2. Reetz MT: **New methods for the high-throughput screening of enantioselective catalysts and biocatalysts**. *Angew Chem Int Ed Engl* 2002, **41**:1335-1338.

3. Arnold FH, Georgiou G: *Directed Enzyme Evolution: Screening and
•   Selection Methods*, vol 230. Totowa, New Jersey: Humana press; 2003.
A very useful overview of high-throughput screening methods available for diverse directed evolution approaches.

4. Goddard JP, Reymond JL: **Enzyme assays for high-throughput screening**. *Curr Opin Biotechnol* 2004, **15**: in press.

5. Tao HY, Cornish VW: **Milestones in directed enzyme evolution**. *Curr Opin Chem Biol* 2002, **6**:858-864.

6. Powell KA, Ramer SW, del Cardayre SB, Stemmer WPC, Tobin MB, Longchamp PF, Huisman GW: **Directed evolution and biocatalysis**. *Angew Chem Int Ed Engl* 2001, **40**:3948-3959.

7. Reetz MT, Zonta A, Schimossek K, Liebeton K, Jaeger KE: **Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution**. *Angew Chem Int Ed Engl* 1997, **36**:2830-2832.

8. Reetz MT, Wilensek S, Zha DX, Jaeger KE: **Directed evolution of an enantioselective enzyme through combinatorial multiple-cassette mutagenesis**. *Angew Chem Int Ed Engl* 2001, **40**:3589-3591.

9. Liebeton K, Zonta A, Schimossek K, Nardini M, Lang D, Dijkstra BW, Reetz MT, Jaeger KE: **Directed evolution of an enantioselective lipase**. *Chem Biol* 2000, **7**:709-718.

10. Zha DX, Wilensek S, Hermes M, Jaeger KE, Reetz MT: **Complete reversal of enantioselectivity of an enzyme-catalyzed reaction by directed evolution**. *Chem Commun* 2001: 2664-2665.

11. Nardini M, Lang DA, Liebeton K, Jaeger KE, Dijkstra BM: **Crystal structure of *Pseudomonas aeruginosa* lipase in the open conformation – the prototype for family I.1 of bacterial lipases**. *J Biol Chem* 2000, **275**:31219-31225.

12. Jaeger KE, Dijkstra BW, Reetz MT: **Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases**. *Annu Rev Microbiol* 1999, **53**:315-351.

13. Bocola M, Otte N, Jaeger KE, Reetz MT, Thiel W: **Learning from
•• directed evolution: theoretical investigations into cooperative mutations in lipase enantioselectivity**. *ChemBioChem* 2004, **5**:214-223.
This paper provides insight into remote and cooperative effects of mutations found in highly (S)-enantioselective variants of the *P. aeruginosa* lipase by using a combination of quantum mechanical force field calculations and molecular dynamics simulations. Steric relay effects were identified and the creation of a new binding pocket for the (S)-enantiomer was also observed.

14. Jaeger KE, Reetz MT: **Microbial lipases form versatile tools for biotechnology**. *Trends Biotechnol* 1998, **16**:396-403.

15. Jaeger KE, Reetz MT: **Directed evolution of enantioselective enzymes for organic chemistry**. *Curr Opin Chem Biol* 2000, **4**:68-73.

16. Jaeger KE, Eggert T, Eipper A, Reetz MT: **Directed evolution and the creation of enantioselective biocatalysts**. *Appl Microbiol Biotechnol* 2001, **55**:519-530.

17. Reetz MT, Jaeger KE: **Directed evolution as a means to create enantioselective enzymes for use in organic chemistry**. In *Directed Molecular Evolution of Proteins: or How to Improve Enzymes for Biocatalysis*. Edited by Johnsson K. Wiley-VCH; 2002: 245-279.

18. Reetz MT: **Controlling the enantioselectivity of enzymes by
•   directed evolution: practical and theoretical ramifications**. *Proc Natl Acad Sci USA* 2004, **101**:5716-5722.
The current status of creating enantioselective enzymes and hybrid catalysts by directed evolution is described with special emphasis on evolving enantioselective lipases originating from *P. aeruginosa* and *B. subtilis* (see [7–12,13••,14–17,34•,36]).

19. Hult K, Berglund P: **Engineered enzymes for improved organic synthesis**. *Curr Opin Biotechnol* 2003, **14**:395-400.

20. Bornscheuer UT: **Methods to increase enantioselectivity of lipases and esterases**. *Curr Opin Biotechnol* 2002, **13**:543-547.

21. Turner NJ: **Controlling chirality**. *Curr Opin Biotechnol* 2003, **14**:401-406.

22. Wu HY, Xu JH, Shen D, Xin Q: **Improved production of (S)-ketoprofen ester hydrolase by a mutant of *Trichosporon brassicae* CGMCC 0574**. *J Ind Microbiol Biotechnol* 2003, **30**:357-361.

23. Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WPC,
•• del Cardayre SB: **Genome shuffling leads to rapid phenotypic improvement in bacteria**. *Nature* 2002, **415**:644-646.
This novel method will have a high impact on classical strain improvement. The time course to develop production strains will be significantly reduced, as demonstrated for tylosin-producing *Streptomyces fragiae*.

24. Patnaik R, Louie S, Gavrilovic V, Perry K, Stemmer WPC, Ryan CM, del Cardayre S: **Genome shuffling of *Lactobacillus* for improved acid tolerance**. *Nat Biotechnol* 2002, **20**:707-712.

25. Kataoka M, Yamamoto K, Kawabata H, Wada M, Kita K, Yanase H, Shimizu S: **Stereoselective reduction of ethyl 4-chloro-3-oxobutanoate by *Escherichia coli* transformant cells coexpressing the aldehyde reductase and glucose dehydrogenase genes**. *Appl Microbiol Biotechnol* 1999, **51**:486-490.

26. Galkin A, Kulakova L, Yoshimura T, Soda K, Esaki N: **Synthesis of optically active amino acids from α-keto acids with *Escherichia coli* cells expressing heterologous genes**. *Appl Environ Microbiol* 1997, **63**:4651-4656.

27. May O, Nguyen PT, Arnold FH: **Inverting enantioselectivity by
•   directed evolution of hydantoinase for improved production of L-methionine**. *Nat Biotechnol* 2000, **18**:317-320.

A whole-cell biotransformation was optimized to produce enantiomerically pure L-methionine by evolving the key pathway enzyme D-hydantoinase.

28. Lingen B, Kolter-Jung D, Dünkelmann P, Feldmann R, Grotzinger J, Pohl M, Müller M: **Alteration of the substrate specificity of benzoylformate decarboxylase from *Pseudomonas putida* by directed evolution**. *ChemBioChem* 2003, **4**:721-726.

29. Peters MW, Meinhold P, Glieder A, Arnold FH: **Regio- and enantioselective alkane hydroxylation with engineered cytochromes P450 BM-3**. *J Am Chem Soc* 2003, **125**:13442-13450.

30. Glieder A, Farinas ET, Arnold FH: **Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase**. *Nat Biotechnol* 2002, **20**:1135-1139.

31. Liese A, Seelbach K, Wandrey C: *Industrial Biotransformations*.
•   Weinheim: Wiley-VCH; 2000.
Excellent summary of industrially relevant synthesis processes using enzymes.

32. Straathof AJJ, Panke S, Schmid A: **The production of fine chemicals by biotransformations**. *Curr Opin Biotechnol* 2002, **13**:548-556.

33. Reetz MT, Becker MH, Klein HW, Stöckigt D: **A method for high-throughput screening of enantioselective catalysts**. *Angew Chem Int Ed Engl* 1999, **38**:1758-1761.

34. Funke SA, Eipper A, Reetz MT, Otte N, Thiel W, Van Pouderoyen G,
•   Dijkstra BW, Jaeger KE, Eggert T: **Directed evolution of an enantioselective *Bacillus subtilis* lipase**. *Biocatal Biotransform* 2003, **21**:67-73.
Here, and in the work described in [37•], complete saturation mutagenesis was used to create a starting pool of enantioselective enzyme variants suitable for further optimization by directed evolution.

35. Short JM: **Saturation mutagenesis in directed evolution**. US Patent No. 6.171.820; 2001.

36. Eggert T, Reetz MT, Jaeger KE: **Directed evolution by random mutagenesis: a critical evaluation**. In *Enzyme Functionality: Design, Engineering, and Screening.* Edited by Svendsen A. New York: Marcel Dekker; 2004:375-390.

37. DeSantis G, Wong K, Farwell B, Chatman K, Zhu ZL, Tomlinson G,
•   Huang HJ, Tan XQ, Bibbs L, Chen P *et al.*: **Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM)**. *J Am Chem Soc* 2003, **125**:11476-11477.
See annotation for [34•].

38. DeSantis G, Zhu ZL, Greenberg WA, Wong K, Chaplin J, Hanson SR, Farwell B, Nicholson LW, Rand CL, Weiner DP *et al.*: **An enzyme library approach to biocatalysis: development of nitrilases for enantioselective production of carboxylic acid derivatives**. *J Am Chem Soc* 2002, **124**:9024-9025.

39. Koga Y, Kato K, Nakano H, Yamane T: **Inverting
•• enantioselectivity of *Burkholderia cepacia* KWI-56 lipase by combinatorial mutation and high-throughput screening using single-molecule PCR and *in vitro* expression**. *J Mol Biol* 2003, **331**:585-592.
The authors present an elegant solution to overcome problems related to the heterologous expression of variant libraries that often hamper directed evolution experiments.

40. Horsman GP, Liu AMF, Henke E, Bornscheuer UT, Kazlauskas RJ: **Mutations in distant residues moderately increase the enantioselectivity of *Pseudomonas fluorescens* esterase towards methyl 3-bromo-2-methylpropanoate and ethyl 3-phenylbutyrate**. *Chemistry* 2003, **9**:1933-1939.

41. Reetz MT, Torre C, Eipper A, Lohmer R, Hermes M, Brunner B, Maichele A, Bocola M, Arand M, Cronin A *et al.*: **Enhancing the enantioselectivity of an epoxide hydrolase by directed evolution**. *Org Lett* 2004, **6**:177-180.

42. Rosenau F, Jaeger KE: **Overexpression and secretion of biocatalysts in *Pseudomonas***. In *Enzyme Functionality: Design, Engineering, and Screening.* Edited by Svendsen A. New York: Marcel Dekker; 2004:617-631.

43. Rosenau F, Jaeger KE: **Overexpression and secretion of *Pseudomonas* lipases**. In *Pseudomons Volume 3.* Edited by Ramos JL. New York: Kluwer Academic/Plenum Publishers; 2004: 491-508.

44. Alexeeva M, Enright A, Dawson MJ, Mahmoudian M, Turner NJ: **Deracemization of α-methylbenzylamine using an enzyme obtained by *in vitro* evolution**. *Angew Chem Int Ed Engl* 2002, **41**:3177-3180.

45. Carr R, Alexeeva M, Enright A, Eve TSC, Dawson MJ, Turner NJ: **Directed evolution of an amine oxidase possessing both broad substrate specificity and high enantioselectivity**. *Angew Chem Int Ed Engl* 2003, **42**:4807-4810.

46. Wada M, Hsu CC, Franke D, Mitchell M, Heine A, Wilson I, Wong CH: **Directed evolution of *N*-acetylneuraminic acid aldolase to catalyze enantiomeric aldol reactions**. *Bioorg Med Chem* 2003, **11**:2091-2098.

47. Williams GJ, Domann S, Nelson A, Berry A: **Modifying the
•• stereochemistry of an enzyme-catalyzed reaction by directed evolution**. *Proc Natl Acad Sci USA* 2003, **100**:3143-3148.
Directed evolution was used to change the stereochemical course of a reaction leading to carbon–carbon bond formation catalyzed by tagatose-1,6-bisphosphate aldolase. Starting from the same educts, the wild-type enzyme formed tagatose-1,6-bisphosphate, whereas the evolved variant preferentially catalyzed the formation of its diastereoisomer fructose-1,6-bisphosphate.

48. Alexeeva M, Carr R, Turner NJ: **Directed evolution of enzymes: new biocatalysts for asymmetric synthesis**. *Org Biomol Chem* 2003, **1**:4133-4137.

49. Janes LE, Kazlauskas RJ, Quick E: **A fast spectrophotometric method to measure the enantioselectivity of hydrolases**. *J Org Chem* 1997, **62**:4560-4561.

Publications of **Chapter 5 "Biocatalyst production in *Bacillus subtilis*"**:

JMB

ELSEVIER

Systematic Screening of All Signal Peptides from *Bacillus subtilis*: A Powerful Strategy in Optimizing Heterologous Protein Secretion in Gram-positive Bacteria

Ulf Brockmeier[1], Michael Caspers[2], Roland Freudl[2] Alexander Jockwer[3], Thomas Noll[3] and Thorsten Eggert[1]*

[1]*Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Düsseldorf im Forschungszentrum Jülich D-52426 Jülich, Germany*

[2]*Institut für Biotechnologie 1 Forschungszentrum Jülich GmbH, D-52425 Jülich Germany*

[3]*Institut für Biotechnologie 2 Forschungszentrum Jülich GmbH, D-52425 Jülich Germany*

*Corresponding author

Efficient protein secretion is very important in biotechnology as it provides active and stable enzymes, which are an essential prerequisite for successful biocatalysis. Therefore, optimizing enzyme-producing bacterial strains is a major challenge in the field of biotechnology and protein production. In this study, the Gram-positive model bacterium *Bacillus subtilis* was optimized for heterologous protein secretion using a novel approach. Two lipolytic enzymes, cutinase from *Fusarium solani pisi* and a cytoplasmatic esterase of metagenomic origin, were chosen as reporters for heterologous protein secretion. In a systematic screening approach, all naturally occurring (non-lipoprotein) Sec-type signal peptides (SPs) from *B. subtilis* were characterized for their potential in heterologous protein secretion. Surprisingly, optimal SPs in cutinase secretion were inefficient in esterase secretion and *vice versa*, indicating the importance of an optimal fit between the SP and the respective mature part of the desired secretion target proteins. These results highlight the need for individually optimal signal peptides for every heterologous secretion target. Therefore, the SP library generated in this study represents a powerful tool for secretion optimization in Gram-positive expression hosts.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: protein secretion; signal peptides; secretion efficiency; cutinase; *Bacillus subtilis*

Introduction

Efficient protein secretion is very important in biotechnology as it provides active and stable enzymes, which are an essential prerequisite for successful biocatalysis. Therefore, optimizing enzyme-producing bacterial strains is a major challenge in the field of industrial biotechnology, also known as White Biotechnology. As one of the most important industrial fermentation hosts, the Gram-positive bacterium *Bacillus subtilis* has been

investigated intensively over the past few decades with respect to its potential for the secretion of heterologous proteins. In addition to the components of the secretion apparatus, the signal peptides (SPs) that channel the export proteins into the secretion machinery play a key role in the translocation across the membrane. Therefore, SPs were investigated in detail with respect to their amino acid composition and to their role in membrane translocation of exported proteins.[1–5]

Signal peptides share some common characteristic features, conserved in different organisms. Most SPs are composed of three distinct regions: (i) the positively charged N domain; (ii) the hydrophobic core region, the so-called H domain; and (iii) the hydrophilic signal peptidase (SPase) recognition site, termed the C domain. Based on these criteria, many signal peptide prediction tools have been developed[6] with SignalP being the most popular and user-friendly programm. SignalP is freely

accessible *via* the world wide web†.[7,8] The SignalP 3.0 server is able to predict the likelihood of a particular amino acid sequence whether it might act as a signal peptide or not by calculating a discrimination score, termed the *D*-score. Sequences leading to values above 0.5 are classified as signal peptides. Sequences showing calculated *D*-scores of >0.7 have a high probability of being signal peptides. Therefore, by using the SignalP prediction tool, Tjalsma *et al.* were able to predict essentially all secreted proteins in the Gram-positive bacterium *B. subtilis*.[9] Nevertheless, high *D*-scores calculated by SignalP do not concurrently indicate high translocation efficiencies of the secreted protein nor do low *D*-scores around 0.5 indicate inefficient translocation. The prediction of secretion efficiency becomes even more complicated when heterologous proteins are fused to an SP instead to its natural secretion partner. As it became obvious that the signal peptide and secreted protein constitute a unique unit, where the N terminus of the mature protein, the so-called signal-mature junction, plays an important role in the secretion efficiency, the prediction of "good" SPs for heterologous protein secretion became impossible because it also depends on the N terminus of the mature part of the secreted protein. Therefore, the high yield secretory production of industrially interesting proteins in heterologous fermentation hosts is often limited by the secretion efficiency, because the "wrong" SP was fused to the secretion partner.

In order to investigate and improve secretion efficiencies, site-specific and random mutagenesis experiments were performed to identify important residues or regions of SPs, which affect secretion efficiencies.[1,2,4,10] However, so far no systematic approach has been described for identifying the optimal SP with respect to its secretion partner. Here, we present the optimization of the secretion efficiencies of two heterologous model enzymes in *B. subtilis* by screening the optimal SP out of a library of all naturally occuring SPs of *B. subtilis*. As a consequence, we developed a simple system for the identification of the best natural *B. subtilis* signal peptide for the secretion of a given heterologous protein which can be subsequently fine-tuned by rational design or directed evolution.

Results

Library construction and high-throughput screening of all Sec-type signal sequences fused to cutinase

All Sec-dependent, non-lipoprotein signal peptides (SPs) as identified by Tjalsma *et al.*[9] were amplified by standard PCR from genomic DNA of *B. subtilis* 168 using specific primer pairs (see Supplementary Data). All 173 PCR products were separately analyzed according to size, purified and cloned in front of the secretion target cutinase from the fungus *Fusarium solani pisi* (Figure 1). According to our cloning strategy (HindIII/BamHI), three additional amino acids (Ala-Glu-Phe) were introduced at the N terminus of the secretion target protein. This meant that the +1 position of the signal peptidase cleavage site was always an alanine residue (Figure 1(b)).

Cutinase was chosen as a model target for heterologous protein secretion in *B. subtilis*, because of its eukaryotic origin and the availability of easy spectrophotometrical activity assays. Furthermore, cutinase itself and, more importantly, the class of enzymes it belongs to, the so-called α/β-hydrolases, are of high biotechnological importance.[11,12] For example the protease (subtilisin), present in washing powder, also belongs to the family of α/β-hydrolases and shows high structural similarities to the cutinase used in this study.[12–14]

The expression system was based on the vector plasmid pBSMuL1[15] which contains a strong constitutive promoter (P_{hpaII}), a *Bacillus*-specific ribosome binding site (rbs) and an ATG start codon in front of the SP-cutinase fusion (Figure 1). The resulting plasmid library, which contains all 173 *B. subtilis* Sec-signal sequences fused to the cutinase, was subsequently transformed into *B. subtilis* TEB1030, a strain that is deficient in both extracellular lipases and therefore lacks any disturbing extracellular lipolytic activity.[16] A *B. subtilis* expression library was therefore created for testing all native Sec-dependent signal peptides with respect to their efficiency in cutinase secretion. Spectrophotometrical screening of the SP library with respect to cutinase activity in the culture supernatant resulted in a clear ranking of all SPs with respect to cutinase secretion efficiency (Figure 2). The detailed results are given as Supplementary Data.

The efficiency of heterologous cutinase secretion is highly dependent on the signal peptide without correlation to the *D*-score

One general strategy for the qualification of signal peptides *ab initio* using computer tools is the *D*-score calculation by SignalP 3.0†.[7,8] The value calculated by the SignalP server indicates the probability of an amino acid sequence functioning as a signal peptide. Therefore, high *D*-scores (0.8–1.0) usually indicate that the corresponding sequence does indeed represent a signal peptide. In contrast, sequences possessing *D*-scores below 0.7 (0.5–0.7) have a lower chance of functioning as a signal peptide, and sequences possessing values below 0.5 are commonly classified as not being functional signal peptides. Using the complete set of native signal peptides from *B. subtilis*, we aimed to test whether a correlation exists between the predicted *D*-scores and the secretion efficiencies of heterologous reporter proteins.

Twenty-five of the 173 SP-cutinase fusions could not be expressed in *B. subtilis* (for details see Table 2

---

† http://www.cbs.dtu.dk/services/SignalP/

Figure 1. Cloning strategy of all Sec-type signal peptides (SPs) of *B. subtilis* as in-frame fusions with *Fusarium solani pisi* cutinase or metagenomic esterase. (a) Schematic overview of the expression vector used in this study. The signal peptide (SP) secretion target fusion was HindIII/BamHI cloned into the multiple cloning site (MCS) downstream of the strong constitutive HpaII promoter (P_{hpaII}). (b) Sequence details of the SP secretion target fusion. The *Bacillus* ribosome binding site (rbs) and the spacer in front of the ATG start codon ensure efficient mRNA translation of the signal sequence gene fusion. The signal peptides are divided into three domains: positively charged N region, hydrophobic H region and C region coding the SPaseI recognition site. All SP-cutinase fusions contain the junction built from the amino acids alanine, glutamate and phenylalanine. The HindIII (AAGCTT) and EcoRI (GAACTT) restriction sites for cloning are underlined.

Figure 2. Comparison of all screened signal peptides (SPs) used for export of heterologous cutinase in *B. subtilis*. The SPs of Epr, YncM, LipA, GlpQ, AmyE, NucB and WprA, which have been analyzed further by Western blotting and pulse–chase experiments, are highlighted in red. The results represent data from 12 independent experiments.
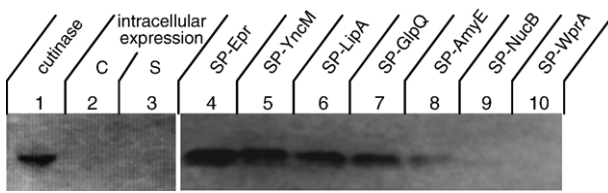
Figure 3. Immunodetection of cutinase in culture supernatants of different *B. subtilis* expression strains. Comparable amounts of culture supernatant proteins (20 μg) were blotted onto a membrane and the cutinase was immunodetected. Lane 1, as positive control, 1.5 μg of purified cutinase were loaded. Lanes 2 and 3, intracellular expression of cutinase using vector pBSMuL3-Cutintra; cellular fraction (C) and supernatant (S) were separated by centrifugation indicating rapid intracellular degradation of cutinase. Lanes 4–10, culture supernatants isolated from *B. subtilis* strains expressing various SP-cutinase fusions (SP-Epr, SP-YncM, SP-LipA, SP-GlpQ, SP-AmyE, SP-NucB and SP-WprA).

could be transferred successfully into the expression host. The screening of this signal peptide library revealed strong differences in lipolytic activity of the culture supernatants, ranging from no secreted cutinase up to 4.7 units/ml, corresponding to 35 mg/l as summarized in Figure 2.

This activity assay-based method for the determination of cutinase export was verified by immunoblotting experiments using cutinase-specific antibodies. Exemplarily, seven different SP-cutinase fusions were analyzed by Western blotting, with the SP-fusions containing the signal peptides from protease Epr (4.67 units/ml), unknown protein YncM (4.12 units/ml), lipase LipA (2.79 units/ml) and phosphodiesterase GlpQ (1.96 units/ml), were selected because of their high cutinase activity in the culture supernatant. Furthermore, fusions containing the signal peptide from α-amylase AmyE (0.67 units/ml) leading to moderate cutinase activity and nuclease NucB (0.34 units/ml), as well as protease WprA (0.12 units/ml), showing low cutinase activity in the culture supernatant, were analyzed. This direct cutinase quantification on protein level (Figure 3) was in accordance with our activity-based quantification of cutinase in the culture medium (Table 1). The exclusive occurence

in the Supplementary Data). Here, the plasmids constructed in *Escherichia coli* could not be successfully transferred into the expression and secretion strain *B. subtilis* TEB1030, indicating a lethal effect. However, the remaining 148 SP-cutinase fusions

Table 1. Comparison of all screened signal sequences used for export of heterologous cutinase in *B. subtilis*

		Signal peptides (SPs)	Secreted cutinase (U/mL)	Charge N region[a]	Hydrophobicity (%)[b]	D-score[c]
No.	Name	Amino acid sequence				
1	Epr	MKNMSCKLVVSVTLFFSFLTIGPLAHA	4.67	2	62.96	0.919
2	YncM	MAKPLSKGGILVKKVLIAGAVGTAVLFGTLSSGIPGLPAADA	4.12	4	76.19	0.507
3	YjfA	MKRLFMKASLVLFAVVFVFAVKGAPAKA	3.84	3	78.57	0.924
4	YfhK	MKKKQVMLALTAAAGLGLTALHSAPAAKA	3.67	3	68.97	0.906
5	Csn	MKISMQKADFWKKAAISLLVFTMFFTLMMSETVFA	3.35	3	62.86	0.689
6	LytD	MKKRLIAPMLLSAASLAFFAMSGSAQA	3.33	3	70.37	0.87
7	Bpr	MRKKTKNRLISSVLSTVVISSLLFPGAAGA	2.97	5	56.67	0.936
8	WapA	MKKRKRRNFK RFIAAFLVLALMISLVPADVLA	2.88	8	65.63	0.918
9	BglC	MKRSISIFITCLLITLLTMGGMIASPASA	2.87	2	65.52	0.839
10	LytB	MKSCKQLIVCSLAAILLLIPSVSFA	2.83	2	64.00	0.916
11	LipA	MKFVKRRIIALVTILMLSVTSLFALQPSAKA	2.79	4	64.52	0.874
...						
26	GlpQ	MRKNRILALFVLSLGLLSFMVTPVSA	1.96	3	69.23	0.921
...						
65	AmyE	MFAKRFKTSLLPLFAGFLLLFHLVLAGPAAASA	0.67	3	78.79	0.904
...						
75	NucB	MKKWMAGLFLAAAVLLCLMVPQQIQGASS	0.34	2	72.41	0.746
...						
88	DacF	MKRLLSTLLIGIMLLTFAPSAFA	0.14	2	73.91	0.909
89	TyrA	MNQMKDTILLAGLGLIGGSIALA	0.13	0	73.91	0.466
90	LytF	MKKKLAAGLTASAIVGTTLVVTPAEA	0.13	3	65.38	0.744
91	WprA	MKRRKFSSVVAAVLIFALIFSLFSPGTKAAA	0.12	4	67.74	0.941
...						
114	Ynza	MELSFTKILVILFVGFLVFGPDKLPALG	0	0	78.57	0.468
115	YobV	MKLERLLAMVVLLISKKQVQA	0	1	61.9	0.623
116	YocH	MKKTIMSFVAVAALSTTAFGAHA	0	2	65.22	0.88
117	YodV	MKVPKTMLLSTAAGLLLSLTATSVSA	0	2	61.54	0.927
...						

The results shown in this Table represent selected data from 12 independent experiments. The standard deviation was below 25%. The exhaustive data on all 148 SPs are summarized in Supplementary Data. The SP names highlighted in red were analyzed further by Western blotting and pulse–chase experiments. (For interpretation of the references to colour in this table legend, the reader is referred to the web version of this article.)

[a] The netto charge of the N region was calculated with amino acids aspartate and glutamate defined as –1, arginine and lysine defined as +1 and any other amino acid defined as 0.

[b] The percentage of hydrophobic amino acids in each signal sequence was calculated with amino acids G, A, V, L, I, M, F, W and P, which were defined as hydrophobic, and any other amino acid characterized as hydrophilic.

[c] D-score calculated by SignalP 3.0 (http://www.cbs.dtu.dk/services/SignalP/).

of processed protein in *B. subtilis* culture supernatants strongly indicates secretion *via* the Sec apparatus and makes cell lysis unlikely as an alternative explanation.

Our results showed that there seemed to be no correlation between high secretion efficiencies and *D*-scores. Among the ten best performing SPs identified in the screening, six show *D*-scores above 0.9 (e.g. Bpr, 0.936; YjfA, 0.924 and Epr, 0.919) as might have been expected; nevertheless, SPs with *D*-scores around 0.5–0.6 were also found (e.g. YncM, 0.507 and Csn, 0.684) to be highly efficient with regard to cutinase secretion (Table 1).

Furthermore, the screening revealed many SPs with high *D*-scores showing moderate (e.g. GlpQ), low (e.g. DacF and AmyE) or no secretion at all (e.g. YodV), as summarized in Table 1. In addition, other SP characteristics like the number of charged amino acids in the N region, as well as the overall hydrophobicity do not correlate with secretion efficiencies, indicating no clear pattern or rule for the prediction of the optimal SP with respect to a heterologous secretion partner.

Translocation and signal peptide processing efficiencies are not solely responsible for the varying amounts of cutinase in the culture supernatants of *B. subtilis*

The varying amounts of cutinase in the supernatants of strains expressing the different signal peptide cutinase fusion proteins could be due to several potential reasons. One of these reasons is that the different signal peptides mediate cutinase translocation across the cytoplasmic membrane *via* the Sec-translocase with different efficiencies. Furthermore, the processing efficiencies might vary between different SP-cutinase fusion proteins. Therefore, the kinetics of membrane translocation and signal peptide processing of selected cutinase fusion proteins were analyzed *via* pulse–chase experiments. For these experiments, fusion proteins with significant differences concerning total protein amounts of cutinase and hence, varying lipolytic activities, in the supernatant were selected.

Of the five signal peptide fusions examined by pulse–chase experiments, two were the most efficient in terms of cutinase protein amounts and lipolytic activities in the supernatants of the corresponding *B. subtilis* strains (Epr, 4.67 units/ml and YncM, 4.12 units/ml), two fusions possessed significantly lower protein amounts and lipolytic activities but were still considered to belong to the group of "efficient" signal peptides with respect to the lipolytic activity in the supernatant (LipA, 2.79 units/ml and GlpQ, 1.96 units/ml) and the fifth showed a comparably low amount of cutinase protein and enzymatic activity in the supernatant (AmyE, 0.67 unit/ml).

In Figure 4, the translocation and processing of cutinase precursor proteins into the mature form by removal of the signal peptide are shown. A strain carrying only the empty vector pBSMuL3 shows no
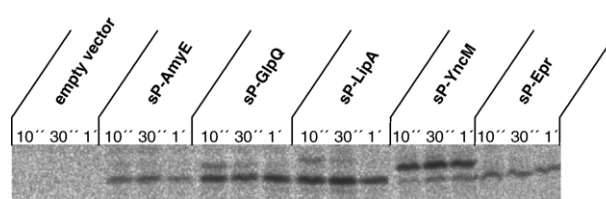

Figure 4. Processing kinetics of cutinase precursor proteins with different signal peptides in *B. subtilis* TEB1030. Kinetics were measured for strains with either the empty vector pBSMuL3 or cutinase fused to signal peptides SP-AmyE, SP-GlpQ, SP-LipA, SP-YncM, and SP-Epr on plasmid pBSMuL3Cut-ss. The *Bacillus* strains were pre-grown in S7 minimal medium containing methionine at 37 °C to an $A_{60}0$ nm of 0.8, washed twice, and starved in S7 minimal medium without methionine for 60 min at 33 °C. Subsequently, cells (2.5 ml) were labelled with 150 µCi of [^{35}S]methionine for 1 min, followed by chasing with a vast excess of non-radioactive methionine. Samples (600 µl) of each strain were withdrawn at certain time points (10 s, 30 s and 1 min), precipitated with ice-cold trichloroacetic acid and further processed as described by van Dijl *et al.*[40] Sample volumes were equalized to 7500 cpm.

protein bands at all, demonstrating that the bands observed in the other samples are indeed specific to the cutinase precursor and the mature form. The kinetics vary dramatically between the different signal peptides, with the Epr and AmyE signal peptides mediating the fastest translocation and processing.

If the varying amounts of cutinase in the supernatant for the different signal peptide cutinase fusion proteins is solely due to differences in the efficiencies of protein translocation across the cytoplasmic membrane and SP processing, those fusions with the highest amounts of protein in the supernatant should exhibit the fastest processing kinetics. However, this is clearly not the case. While the translocation and processing of the Epr fusion protein, which shows the highest lipolytic acitivity and the largest amount of cutinase protein in the supernatant, is indeed very fast, the second-"best" signal peptide with respect to the amounts of cutinase protein and activity in the supernatant, YncM, shows the slowest translocation and processing of all the proteins investigated in the pulse–chase experiment. Furthermore, the AmyE fusion, which shows the lowest protein amounts and lipolytic activity in the supernatant of the five proteins analyzed in the pulse–chase, is translocated and processed very rapidly.

Hence, translocation and SP processing efficiencies are not the sole determinants responsible for varying amounts of secreted cutinase present in the culture supernatant of the *B. subtilis* cells expressing different SP-cutinase fusion proteins.

One-step optimization for secretion of heterologous proteins in *B. subtilis*

In the first approach described above, cutinase was used as a heterologous secretion reporter. In this

experiment, every SP-cutinase fusion was ligated and cloned seperately to ensure that the signal peptide fusion library was complete. However, such a procedure is tedious, time-consuming and costly. Therefore, in another approach we tested whether the SP library could be used to find the optimal fusion partner in a standardized one-step procedure of cloning and screening.

In order to test the feasibility of the one-step optimization procedure, a mixture of all PCR-amplified SPs were cloned in front of the cutinase. The ligation mix, theoretically containing all SP-cutinase fusions, was transferred into *B. subtilis* TEB1030 and subsequently 930 clones were chosen randomly and analyzed with respect to cutinase activity in the culture supernatant. The 12 clones that performed best with secretion efficiencies of 4–5 units/ml of cutinase were isolated and sequenced. As the result, the eight best-performing SPs according to Table 1 were identified again, with YncM identified three times as well as LytD and WapA both being identified twice. It was thereby demonstrated that the fast random approach was suitable for identifying the same SPs as being highly efficient in cutinase secretion, as those that were characterized before using the one-by-one approach.

In order to demonstrate the general applicability of this system, a second heterologous protein was chosen for secretion optimization in *B. subtilis*. The esterase EstCL1 of metagenomic origin was amplified from plasmid p11EstCL1[17] and the resulting 974 bp PCR product was cloned into the MCS of

vector pBSMuL3. This metagenomic esterase was chosen because, as a member of the family IV of bacterial lipolytic enzymes,[11] it is an intracellular enzyme that is usually not secreted in its native host. Therefore, it should be a model reporter for naturally unsecreted proteins.

Again, an equimolar mixture of all PCR-amplified SPs was cloned in front of the secretion partner and then transferred into the expression host *B. subtilis* TEB1030. The culture supernatants of about 1000 clones were tested in an automated high-throughput screening assay using the standard esterase substrate *p*-nitrophenyl-caproate (*p*-NPC). Plasmid DNA of ten clones showing the highest secretion efficiencies of EstCL1 were isolated and analyzed by DNA sequencing. Signal peptides NtrE and YfhK were identified twice; therefore, eight different SPs selected in this screening were characterized in more detail (Figure 5), with SP-YwmC being the most efficient with respect to EstCL1 secretion. The highest secretion level of 1.5 units/ml (±0.1 unit/ml) corresponds to 15–20 mg/l being fortunately a lot for an intracellular enzyme of heterologous origin.[18] Again, like in cutinase secretion, *D*-scores do not correlate with secretion efficiencies of the target protein. Furthermore, the best-performing SPs in EstCL1 secretion do not correlate with the SP ranking of cutinase secretion. The three most efficient SPs in the secretion of the metagenome esterase lead to poor (SP-YwmC and SP-YpjP) or no (SP-YojL) secretion of cutinase (Table 1; Figure 5). In addition, we fused the best SP with respect to

SP-name	YwmC	YpjP	YojL	YwsB	YfhK	YxaK	NprE	YlxW	Epr
D-score[a] esterase	0,875	0,817	0,742	0,725	0,855	0,695	0,799	0,588	0,903
cutinase	0,912	0,842	0,800	0,662	0,906	0,679	0,832	0,614	0,919

[a] calculated for the first 70 amino acids of hydrid proteins SP-esterase EstCL1 and SP-cutinase
[b] secretion efficiencies of YwmC-esterase (1,5 U/mL) and Epr-cutinase (4,67 U/mL) were set to 100 %
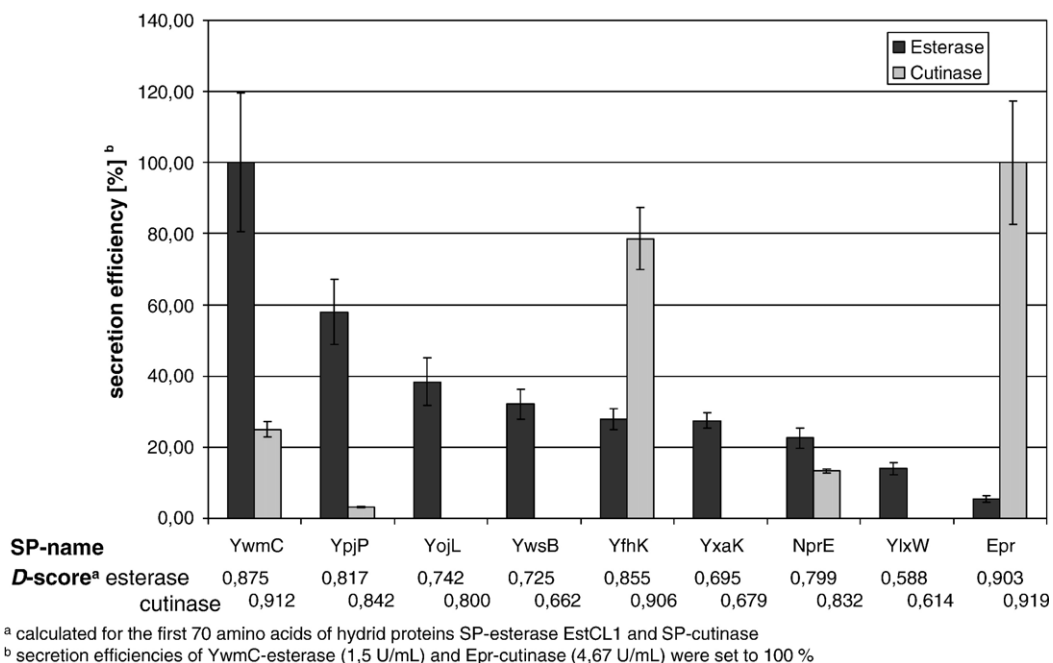
Figure 5. Identification of the most efficient signal peptide in secretion of the heterologous esterase EstCL1 in *B. subtilis*. The SP library containing a mixture of all Sec-type signal peptides fused to the metagenomic esterase EstCL1 was transferred into *B. subtilis* TEB1030. Culture supernatants of about 1000 transformants were screened towards esterase activity. Plasmid DNA of the ten clones that showed highest activity was sequenced. Signal peptides NtrE and YfhK were identified twice. For a direct comparison the efficiencies in cutinase secretion were included (see Supplementary Data). The SP-Epr fusion with EstCL1 was not identified in the screening approach and therefore was fused site-directed to compare the two best SPs with respect to cutinase and esterase secretion.

cutinase secretion (SP-Epr) directly to the metagenome esterase. SP-Epr fused to the metagenome esterase resulted in a secretion efficiency of 0.076 unit/ml (±0.01 units/ml) which is about 5% of the secretion level of the best SP-YwmC (1.5(±0.1) units/ml). This direct comparison of secretion efficiencies of the same SPs fused to different target proteins underlines the secretion target specificity of the best SP.

Discussion

For many years, *B. subtilis* and related Bacilli such as *Bacillus licheniformis* have been used as expression and secretion hosts. Their ability to express and secrete enzymes in high amounts of up to 20 g/l into the culture medium, as well as their classification as GRAS organisms, i.e. generally recognized as safe, free of any endotoxin, are the reasons for their wide use in industrial fermentation processes.[19] However, their high secretion efficiencies are realized only with homologous proteins from the *Bacillus* species fused to their cognate signal peptides (SPs). Heterologous enzyme secretion often results in inefficient and unsatisfying low yields. Secretion levels of at least 1–15 mg/l start to become useful for commercial purposes.[18,20] Therefore, fundamental research on the secretion mechanisms of *Bacillus* strains was performed to identify bottlenecks in the protein secretion pathway.[21]

Here, we have established a new strategy for the optimization of heterologous protein secretion in *B. subtilis* by screening a library of all natural SPs of the strain. Our systematic approach using two different lipolytic enzymes as secretion targets clearly demonstrate the inability of currently available computer tools to predict an optimal SP for heterologous protein secretion. One algorithm used to indicate the probability of a given amino acid sequence functioning as a signal peptide is employed by the SignalP server to calculate the so-called *D*-score, and is commonly used to predict extracellular proteins containing SPs.[7,8] However, as shown here, no correlation exists between the calculated *D*-score and the secretion efficiency of the heterologous secretion target protein.

Since it was observed that a defined alteration of net charge, hydrophobicity or the length of the SP can change secretion efficiency,[1,22] we analyzed these criteria in our SP library without clear correlation to the secretion amounts. Only the overall composition of SPs, positively charged N region, hydrophobic core region and suitable cleavage site, was obvious, and furthermore the SP library showed a strong heterogeneity in length, charge and hydrophobicity without correlation to the secretion efficiency of the chosen heterologous model reporter proteins. This was in the line of experiments analyzing the secretion efficiencies of SPs fused to the secretion target protein after modulation by means of site-directed mutagenesis. Here, the effects were not reproduced even when

beneficial modifications were transferred to other SPs fused to the same secretion partner.[1,2,22,23] Furthermore, secretion experiments in *E. coli* reported an influence of the N-terminal part of the mature protein in secretion efficiency, indicating the importance of an optimal interaction between signal peptide and secretion target.[24–27]

The best SP for the secretion of one target protein is not automatically the best, or even a sufficient SP, for the secretion of another different target protein. As demonstrated here, efficient translocation and processing, such as the one mediated by the signal peptide from AmyE fused to cutinase, does not automatically result in high amounts of secreted protein, nor does a slow translocation and processing necessarily result in low amounts of secreted protein.

Instead, it appears that the secretion efficiency for a given target protein in *B. subtilis* is determined by a complex pattern of events. For example, the fusion of the secretion target to different SPs might result in different mRNA stabilities of the corresponding transcripts and, therefore, in different amounts of precursor proteins synthesized. Furthermore, after targeting to the membrane, translocation across the membrane, and processing, the availability of cell-associated chaperones, such as PrsA, and the presence of cell-associated and/or secreted proteases most likely also influence significantly the amounts of protein that accumulate in the culture supernatant.

Depending on the folding efficiency at the extracellular site of the cytoplasmatic membrane of *B. subtilis*, improperly folded protein accumulates in the cell wall to a greater or lesser extent. This folding efficiency is essentially determined by the availability of extracytosolic folding catalysts like PrsA.[28] In the case of PrsA overloading due to highly efficient targeting, translocation, and processing of the secretion target (e.g. as observed for cutinase fused to the SP-AmyE), post-translocational folding could become a limiting step and improperly folded protein will accumulate in the cell wall. As a consequence of this secretion stress signal, cell-associated proteases like HtrA and HtrB are up-regulated *via* the CssR/CssS two-component regulatory system.[29,30] These proteases are mainly thought to degrade unfolded proteins. However, the increased HtrA/HtrB amounts might also result in an increased degradation of the pool of properly folded foreign proteins that are temporarily present in the cell wall, the extent of which depends on the individual sensitivity of the foreign protein towards these proteases. Both the degradation of the unfolded and properly folded foreign proteins will finally result in a reduction of the overall secretion efficiency of the respective heterologous secretion target. Consequently, in order to achieve high quantities of correctly folded heterologous target protein in the extracellular medium, the amount of translocated proteins emerging from the *trans*-side of the membrane (influenced by transcription, translation, targeting, translocation and processing

efficiencies) and the extracytosolic folding capacity (influenced by specific folding properties of the secretion target and availability of extracytosolic folding catalysts) must be well balanced.

With respect to the complexity of the heterologous protein secretion and because of the many interactions involved in the determination of the overall secretion efficiency, our strategy of screening a complete SP library fused to the secretion target represents a powerful tool for the quick and easy optimization of the export of virtually any target protein. In future work, a greater variety of secretion targets will be fused to our SP library and screened with respect to secretion efficiency. It is hoped that novel computer tools can be developed based on these results and used to identify the complex interaction patterns between SP and the mature part of the desired secretion target protein.

Methods

Bacterial strains and plasmids

The strains and plasmids used in this study are listed in Table 2. *B. subtilis* TEB1030 was used as the expression host for plasmid-encoded cutinase. *E. coli* JM109 served as a host for cloning and plasmid preparation.

Media, growth conditions and preparation of culture supernatants

E. coli and *B. subtilis* cells were grown in LB medium at 37 °C supplemented with 100 μg/ml of ampicillin and 50 μg/ml of kanamycin, respectively. For the screening experiments, *B. subtilis* cultures were grown in deep-well microtiter plates (96 wells, 2 ml per well; Greiner Bio-One,

Frickenhausen, Germany) using a microplate shaker (600 rpm, TiMix 5; Edmund Bühler GmbH, Hechingen, Germany) for 16 h. After this incubation period the cell cultures had reached optical densities of $A_{580} = 3.4(\pm 0.25)$ in the stationary phase. Culture supernatants were prepared by centrifugation (30 min at 5000g, 4 °C). Both lipolytic enzymes used in this study were stable in the culture supernatant over 24 h without significant loss of activity when stored at 4 °C; however, in order to minimize the risk of proteolytic degradation, the samples were used immediately after centrifugation.

For the pulse–chase labelling experiments, *B. subtilis* TEB1030 was grown in modified S7 medium,[31] in which Mops buffer was replaced with 20 mM potassium phosphate buffer and glucose was replaced with 1% (w/v) ribose.

Transformation of DNA

E. coli strain JM109 was used for electroporation in a MicroPulser (BioRad, München, Germany). The preparation of electrocompetent cells and the electrotransformation was performed following the manufacturer's protocol for the electrotransformation of *E. coli* cells. Competent cells of *B. subtilis* were prepared by the two-step procedure and transformed as described.[32]

Polymerase chain reaction (PCR) conditions

Amplification of DNA fragments was performed under standard PCR conditions in 25 μl reaction volumes as described.[15] The oligonucleotides used as PCR primers are summarized in Supplementary Data.

Lipolytic activity assays

The quantitative detection of the lipolytic activity of cutinase was performed using a spectrophotometric assay as described by Winkler & Stuckmann.[33] As the

Table 2. Bacterial strains and plasmids used in this study

Strains or plasmids	Genotype/characteristics	Source or reference
A. Strains		
E. coli JM109	e14-(McrA-) recA1 endA1 gyrA96 thi-1 hsdR17(rK- mK+)supE44 relA1 D(lac-proAB) [F' traD36 proAB lacI^{q}ZDM15]	Stratagene, Heidelberg
B. subtilis 168	trpC2	41
B. subtilis DB430	his nprE aprE bpf ispI	42
B. subtilis TEB1030	*B. subtilis* DB430 lipA lipB	16
B. Plasmids		
pBSMuL1	*B. subtilis* – *E. coli* shuttle vector for overexpression (P$_{HpaII}$, ribosome binding site) and secretion (sslipA); ColE1 repB Km^r Amp^r	15
pBSMuL3	pBSMuL1 derivative without ribosome binding site and without sslipA; inverted MCS	This study
pBSMuL3-Cutintra	pBSMuL3 containing a 655 bp HindIII-BamHI fragment of *Fusarium solani pisi* cutinase without signal sequence	This study
pBSMuL3-Cut	pBSMuL3 containing a 655 bp EcoRI-BamHI fragment of *Fusarium solani pisi* cutinase without start codon and signal sequence	This study
pBSMuL3-Cut-sp	pBSMuL3-Cut containing an in-frame fusion of all Sec-dependent signal peptides of *B. subtilis*; sp is an abbreviation for the particular signal peptide; according to the 173 different signal peptides of *B. subtilis*, 173 different pBSMuL3-Cut-sp plasmids were constructed	This study
p11EstCL1	pET11a expression vector (Novagen, Madison, USA) containing a 1191 bp NdeI/BamHI fragment of metagenomic esterase EstCL1	17
pBSMuL3-EstCL1	pBSMuL3 containing a 964 bp EcoRI-BamHI fragment of a metagenome esterase (EstCL1) without start codon signal sequence (comparable to pBSMuL3-Cut)	This study
pBSMuL3-EstCL1-spMix	pBSMuL3-Cut containing an in-frame fusion of all Sec-dependent signal sequences of *B. subtilis* (comparable to pBSMuL3-Cut-sp)	This study

substrate, 30 mg of *p*-nitrophenyl-palmitate (*p*NPP) was dissolved in 10 ml of isopropanol and mixed with 90 ml of Sørensen phosphate buffer (pH 8), supplemented with sodium deoxycholic acid (207 mg) and gum arabic (100 mg). The final concentration of the substrate was 0.8 mM. The enzymatic activity was calculated using a molar absorbtion coefficient of 15,000 M^{-1} cm^{-1}.

For the quantitative detection of esterase activity, 23.7 mg of *p*-nitrophenyl-caproat (*p*NPC) was dissolved in 5 ml of ethanol and added to 95 ml of 100 mM potassium phosphate buffer (pH 7.2) containing 10 mM $MgSO_4$ to yield a final concentration of 1 mM *p*NPC. Five–20 μl of culture supernatant was added to the substrate solution to give a final volume of 1 ml, and the ΔA_{410} was recorded over 15–30 min at 37 °C. The enzymatic activity was calculated using a molar absorbtion coefficient of 12,500 M^{-1} cm^{-1}.

The realization of high-throughput screening towards lipase (*p*NPP) and esterase (*p*NPC) activity was achieved in microtiter plates (96 –well plates) as described.[34,35]

Preparation of cutinase antibodies

Polyclonal anti-cutinase antibodies were produced by hybridoma cells resulting from the fusion of B-lymphocytes from immunized Balb/c mice with SP2/0 myeloma cells[36] and secreting these into the culture media. Cells were routinely cultured in serum-free hybridoma media (Ex-Cell 610–HSF; JRH biosciences, Lenexa, Kansas, USA). Batch cultivations for the production of antibodies were all carried out in spinner flasks (100 – 300 ml working volume; Techne, Cambridge, UK) with headspace aeration at an agitation speed of 40 rpm. Cultures were incubated at 37 °C in an atmosphere of 95% air saturation, 5% CO_2 and 99% humidity (Heraeus Instruments, Hanau, Germany). Initial cell densities were 0.2×10^6/ml for all cultivations and cell stock holding were appropriately split to keep cell densities below 1.5×10^6/ml. Supernatants containing anti-cutinase antibodies were recovered by centrifugation (10 min at 200*g* 4 °C) and kept refrigerated (4 °C) for short-term usage or frozen (–80 °C) for long-term storage, respectively.

Protein analysis and immunodetection of cutinase

The protein concentration was measured spectrophotometically at 595 nm in accordance with Bradford's method[37] using bovine serum albumin (BSA) as the standard. The SDS/polyacrylamide gel electrophoresis (SDS/PAGE) was performed using a 5% (w/v) stacking gel and a 15% separating gel.[38]

For immunodetection of cutinase, the proteins were concentrated using trichloroacetic acid (TCA) precipitation, separated by SDS/PAGE and Western-blotted onto a polyvinylidene difluoride membrane as described by Dunn.[39] Cutinase protein was detected using the 1:1000 diluted polyclonal antiserum. For the detection of cutinase, horseradish peroxidase-labelled goat anti-mouse antibody (BioRad, München, Germany) was used as the second antibody.

Pulse–chase protein labelling and immunoprecipitaion

Pulse–chase protein labelling, immunoprecipitation and SDS/PAGE were carried out as described by van Dijl *et al.*[40]

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2006.07.034

References

1. Borchert, T. V. & Nagarajan, V. (1991). Effect of signal sequence alterations on export of levansucrase in *Bacillus subtilis*. *J. Bacteriol.* **173**, 276–282.
2. Chen, M. & Nagarajan, V. (1994). Effect of alteration of charged residues at the N termini of signal peptides on protein export in *Bacillus subtilis*. *J. Bacteriol.* **176**, 5796–5801.
3. Gierasch, L. M. (1989). Signal Sequences. *Biochemistry*, **28**, 923–930.
4. von Heijne, G. (1985). Signal sequences. The limits of variation. *J. Mol. Biol.* **184**, 99–105.
5. von Heijne, G. (1990). The signal peptide. *J. Membr. Biol.* **115**, 195–201.
6. Menne, K. M. L., Hermjakob, H. & Apweiler, R. (2000). A comparison of signal sequence prediction methods using a test set of signal peptides. *Bioinformatics*, **16**, 741–742.
7. Nielsen, H., Engelbrecht, J., Brunak, S. & vonHeijne, G. (1997). Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng.* **10**, 1–6.
8. Bendtsen, J. D., Nielsen, H., von Heijne, G. & Brunak, S. (2004). Improved prediction of signal peptides: SignalP 3.0. *J. Mol. Biol.* **340**, 783–795.
9. Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S. & van Dijl, J. M. (2000). Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* **64**, 515–547.
10. Kaderbhai, M. A., Davey, H. M. & Kaderbhai, N. N. (2004). A directed evolution strategy for optimized export of recombinant proteins reveals critical determinants for preprotein discharge. *Protein Sci.* **13**, 2458–2469.
11. Jaeger, K. E., Dijkstra, B. W. & Reetz, M. T. (1999). Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. *Annu. Rev. Microbiol.* **53**, 315–351.
12. Maurer, K. H. (2004). Detergent proteases. *Curr. Opin. Biotechnol.* **15**, 330–334.
13. Martinez, C., Degeus, P., Lauwereys, M., Matthyssens, G. & Cambillau, C. (1992). *Fusarium solani* cutinase is a lipolytic enzyme with a catalytic serine accessible to solvent. *Nature*, **356**, 615–618.

14. Longhi, S., Czjzek, M., Lamzin, V., Nicolas, A. & Cambillau, C. (1997). Atomic resolution (1.0 angstrom) crystal structure of *Fusarium solani* cutinase: stereochemical analysis. *J. Mol. Biol.* **268**, 779–799.

15. Brockmeier, U., Wendorff, M. & Eggert, T. (2006). Versatile expression and secretion vectors for *Bacillus subtilis*. *Curr. Mibrobiol.* **52**, 143–148.

16. Eggert, T., Brockmeier, U., Dröge, M. J., Quax, W. J. & Jaeger, K. E. (2003). Extracellular lipases from *Bacillus subtilis*: regulation of gene expression and enzyme activity by amino acid supply and external pH. *FEMS Microbiol. Letters*, **225**, 319–324.

17. Leggewie, C. (2005). Novel biocatalysts from the metagenome: expression, identification and biochemical characterization, PhD Thesis, Heinrich Heine University, Düsseldorf.

18. Georgiou, G. & Segatori, L. (2005). Preparative expression of secreted proteins in bacteria: status report and future prospects. *Curr. Opin. Biotechnol.* **16**, 538–545.

19. Schallmey, M., Singh, A. & Ward, O. P. (2004). Developments in the use of *Bacillus* species for industrial production. *Can. J. Microbiol.* **50**, 1–17.

20. Majander, K., Anton, L., Antikainen, J., Lang, H., Brummer, M., Korhonen, T. K. & Westerlund-Wikstrom, B. (2005). Extracellular secretion of polypeptides using a modified Escherichia coli flagellar secretion apparatus. *Nature Biotechnol.* **23**, 475–481.

21. Braun, P., Gerritse, G., van Dijl, J. M. & Quax, W. J. (1999). Improving protein secretion by engineering components of the bacterial translocation machinery. *Curr. Opin. Biotechnol.* **10**, 376–381.

22. Sakakibara, Y., Tsutsumi, K., Nakamura, K. & Yamane, K. (1993). Structural requirements of *Bacillus subtilis* alpha-amylase signal peptide for efficient processing: *in vivo* pulse-chase experiments with mutant signal peptides. *J. Bacteriol.* **175**, 4203–4212.

23. Ravn, P., Arnau, J., Madsen, S. M., Vrang, A. & Israelsen, H. (2003). Optimization of signal peptide SP310 for heterologous protein production in *Lactococcus lactis*. *Microbiology*, **149**, 2193–2201.

24. Lehnhardt, S., Pollitt, S. & Inouye, M. (1987). The differential effect on 2 hybrid proteins of deletion mutations within the hydrophobic region of the *Escherichia coli* OmpA signal peptide. *J .Biol. Chem.* **262**, 1716–1719.

25. Li, P., Beckwith, J. & Inouye, H. (1988). Alteration of the amino terminus of the mature sequence of a periplasmic protein can severely affect protein export in *Escherichia coli*. *Proc. Natl Acad. Sci. USA*, **85**, 7685–7689.

26. MacIntyre, S., Eschbach, M.-L. & Mutschler, B. (1990). Export incompatibility of N-terminal basic residues in a mature polypeptide of *Escherichia coli* can be alleviated by optimising the signal peptide. *Mol. Gen. Genet.* **221**, 466–474.

27. Rasmussen, B. A. & Silhavy, T. J. (1987). The 1st 28 amino acids of mature LamB are required for rapid and efficient export from the cytoplasm. *Genes Dev.* **1**, 185–196.

28. Kontinen, V. P. & Sarvas, M. (1993). The PrsA lipoprotein is essential for protein secretion in *Bacillus subtilis* and sets a limit for high-level secretion. *Mol. Microbiol.* **8**, 727–737.

29. Darmon, E., Noone, D., Masson, A., Bron, S., Kuipers, O. P., Devine, K. M. & van Dijl, J. M. (2002). A novel class of heat and secretion stress-responsive genes is controlled by the autoregulated CssRS two-component system of *Bacillus subtilis*. *J. Bacteriol.* **184**, 5661–5671.

30. Hyyrylainen, H. L., Bolhuis, A., Darmon, E., Muukkonen, L., Koski, P., Vitikainen, M. *et al.* (2001). A novel two-component regulatory system in *Bacillus subtilis* for the survival of severe secretion stress. *Mol. Microbiol.* **41**, 1159–1172.

31. Vasantha, N. & Freese, E. (1980). Enzyme changes during *Bacillus subtilis* sporulation caused by deprivation of guanine nucleotides. *J. Bacteriol.* **144**, 1119–1125.

32. Kunst, F. & Rapoport, G. (1995). Salt stress is an environmental signal affecting degradative enzyme synthesis in *Bacillus subtilis*. *J. Bacteriol.* **177**, 2403–2407.

33. Winkler, U. K. & Stuckmann, M. (1979). Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*. *J. Bacteriol.* **138**, 663–670.

34. Reetz, M. T. (2004). Controlling the enantioselectivity of enzymes by directed evolution: Practical and theoretical ramifications. *Proc. Natl Acad. Sci. USA*, **101**, 5716–5722.

35. Reetz, M. T., Zonta, A., Schimossek, K., Liebeton, K. & Jaeger, K. E. (1997). Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution. *Angew Chem. Int. Ed. Engl.* **36**, 2830–2832.

36. Roque, A. C. A., Lowe, C. R. & Taipa, M. A. (2004). Antibodies and genetically engineered related molecules: production and purification. *Biotechnol. Prog.* **20**, 639–654.

37. Bradford, M. M. (1976). Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254.

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

39. Dunn, S. D. (1986). Effects of the modification of transfer buffer composition and the renaturation of proteins in gels on the recognition of proteins on Western blots by monoclonal antibodies. *Anal. Biochem.* **157**, 144–153.

40. van Dijl, J. M., de Jong, A., Smith, H. E., Bron, S. & Venema, G. (1991). Non-functional expression of *Escherichia coli* signal peptidase I in *Bacillus subtilis*. *J. Gen. Microbiol.* **137**, 2073–2083.

41. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V. *et al.* (1997). The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. *Nature*, **390**, 249–256.

42. Doi, R. H., Wong, S. L. & Kawamura, F. (1986). Potential use of *Bacillus subtilis* for secretion and production of foreign proteins. *Trends Biotechnol.* **4**, 232–235.

Versatile Expression and Secretion Vectors for *Bacillus subtilis*

Ulf Brockmeier, Marion Wendorff, Thorsten Eggert

Institut für Molekulare Enzymtechnologie, Heinrich-Heine Universität Düsseldorf, Forschungszentrum Jülich, D-52426 Jülich, Germany

Abstract. Most expression systems are based on *Escherichia coli* as the host strain because of the large availability of all kinds of vector plasmids. However, aside from the obvious advantages of *E. coli* systems, serious problems can occur during the process of heterologous gene expression and purification. Therefore, low expression rates, formation of inclusion bodies, improper protein-folding, and/or toxicity problems might enforce changing the expression host. Here we describe the construction of two new vectors, pBSMuL1 and pBSMuL2, for overexpression and secretion of heterologous proteins in *Bacillus subtilis* as an alternative expression host. The new plasmids combine several advantages in comparison to available *Bacillus* expression systems: an appropriate multiple cloning site consisting of 13 unique restriction sites, one (pBSMuL1) or two (pBSMuL2) strong constitutive promoters, a high efficient signal sequence for protein secretion, and the possibility to express proteins as His-tagged fusions for easy detection and purification. We have demonstrated the applicability of the novel vector plasmids for the production and purification of the heterologous cutinase from *Fusarium solani pisi*.

A large pool of cloning and expression plasmids is presently available for the Gram-negative bacterium *Escherichia coli* (e.g., pUC-vectors or the pET-vectors series commercialized by Novagen, Madison, WI). Therefore, most protein expression strategies in microbiologic research focus on this organism. However, aside from the obvious advantages of *E. coli* systems, serious problems can occur during the process of heterologous gene expression and purification: (1) low expression rates, (2) formation of inclusion bodies, (3) improper protein-folding, and/or (4) toxicity problems that require changing the expression host [1].

An alternative expression host also for large-scale production of foreign proteins is the Gram-positive bacterium *Bacillus subtilis*. One major advantage of the strain is its classification as a GRAS, i.e., generally recognized as safe, organism free of any endotoxin. Furthermore, compared with *E. coli*, the Gram-positive bacterium *B. subtilis* offers an efficient secretion apparatus that guides the expressed protein directly into the culture supernatant [13], thereby bypassing the time-consuming cell disruption that makes subsequent protein

Correspondence to: Thorsten Eggert; *email:* t.eggert@fz-juelich.de

purification much easier. In addition, in case of efficient secretion, the formation of inclusion bodies in the cytoplasm is decreased, leading to higher amounts of properly folded and active enzymes. For all of these reasons, *B. subtilis* has developed into an important expression strain frequently used in industrial fermentations in past decades [9].

Different kinds of secretion vectors have been published already for *B. subtilis* [7, 8]; however, until now the choice of available expression plasmids combining similar properties as *E. coli* systems is still limited. Therefore, we constructed the multicopy vectors pBSMuL1 and pBSMuL2 (Fig. 1A) based on the *E. coli–B. subtilis* shuttle plasmid pMA5 with a pUB110 ori for replication in *B. subtilis* [3, 15]. Both plasmids offer several advantages with respect to biotechnologic applications: Downstream of strong constitutive promoter(s), an artificial DNA fragment was inserted containing (1) a *Bacillus* ribosome binding site, (2) a *B. subtilis* signal sequence for efficient secretion of the recombinant protein (*sslipA*), (3) a multiple cloning site (MCS) composed of 13 frequently used restriction sites, (4) the possibility of in-frame fusion to a hexa-histidine-tag (6xHis) for convenient one-step purification by
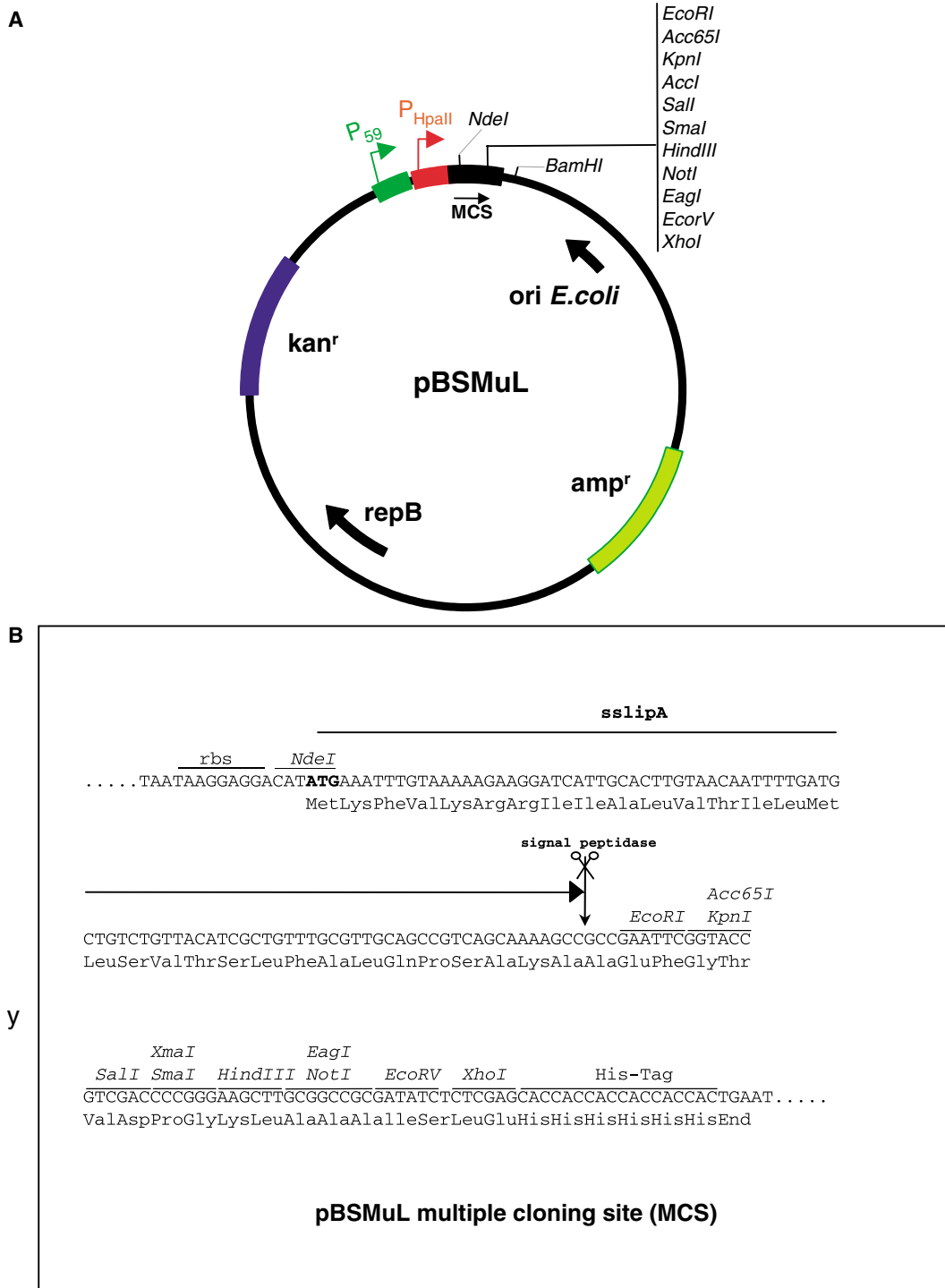
Fig. 1. Plasmid map of vectors pBSMuL1 (7494 bp) and pBSMuL2 (7559 bp). (A) Overview of relevant elements: P_{HpaII}, constitutive Gram-positive promoter; P_{59}, second constitutive Gram-positive promoter present only in pBSMuL2; MCS, multiple cloning site; ori *E.coli*, gene required for replication in *E. coli*; *amp*^{r}, β-lactamase gene conferring ampicillin resistance in *E.coli*; *repB*, gene required for replication in *B. subtilis*; *kan*^{r}, kanamycin resistance gene. (B) Properties of the artificial MCS: rbs, ribosome binding site (recognized in *B. subtilis*); ss*lipA*, *lipA* signal sequence of *B. subtilis* for efficient protein secretion into the culture medium (the signal peptidase cleavage site (Ala-Lys-Ala) is marked with an arrow); His-Tag, C-terminal in-frame fusion to six histidine residues for protein purification.

Table 1. Oligonucleotides and PCR primers

Primer	Sequence 5′ → 3′	Restriction site
U1	TAATAT**ATTAAT**AAGGAGGACATATGAAATTTGTAAAAAGAAGGATCATTGCACTTGTAACAATTTTGAT	*Vsp*I
U2	GCTGTCTGTTACATCGCTGTTTGCGTTGCAGCCGTCAGCAAAAGCCGCCGAATTCGGTACCGTCGACCCC	–
U3	GGGAAGCTTGCGGCCGCGATATCTCTCGAGCACCACCACCACCACCACTGA**ATTAAT**ATAT	*Vsp*I
L1	TGCAATGATCCTTCTTTTTACAAATTTCATATGTCCTCCTT**ATTAAT**ATAT	*Vsp*I
L2	GGCTGCAACGCAAACAGCGATGTAACAGACAGCATCAAAATTGTTACAAG	–
L3	CCGCAAGCTTCCCGGGGTCGACGGTACCGAATTCGGCGGCTTTTGCTGAC	–
L4	TAATAT**ATTAAT**TCAGTGGTGGTGGTGGTGGTGCTCGAGAGATATCGCGG	*Vsp*I
cuti1	ATAT**GAATT**CGCGCCTACTAGTAACCCTGCT	*Eco*RI
cuti2	ATAT**CTCGAG**AGCAGAACCACGGACAGCCCG	*Xho*I
P59-up	ATGGCTTGACAGGGAGAGATAGGTTTGATAGAATATAATAGTTGTCGCGGAAGCCATCCAT**GAAG**	*Bst*XI
P59-low	**CTTC**TACCGAACTGTCCCTCTCTATCCAAACTATCTTATATTATCAACAGCGCCTTCGGTAGGTA	*Bst*XI

[a]Restriction sites are highlighted in bold letters.
PCR, polymerase chain reaction.

immobilized-metal affinity chromatography; and (5) the possibility of intracellular protein expression by using the *Nde*I restriction site (Fig. 1B). Here we demonstrate the convenient cloning, gene expression, secretion, and subsequent one-step purification of cutinase from *Fusarium solani pisi* as C-terminal His-tagged fusion-protein from 1 L *B. subtilis* culture supernatant using the pBSMuL vectors.

Results and Discussion

The vector plasmids were constructed by excision of the original cloning site (*Eco*RV, *Kpn*I, and *Hin*dIII) of pMA5 [3] by *Eco*RI and *Hin*dIII double-digestion, blunting using the Klenow fragment, and subsequent religation, resulting in an intermediate plasmid named pMA5ΔMCS. The DNA fragment containing the new MCS of both vectors was constructed *in vitro* using seven oligonucleotides (U1 to U3 and L1 to L4; Table 1). A pool of all these oligonucleotides (5 pmol each dissolved in *A. dest*) with a final volume of 14 μL was incubated at 95°C to 100°C for 10 minutes and cooled down to room temperature afterwards. This sample was directly used as a template in a standard polymerase chain reaction (PCR) reaction (95°C for 30 seconds, 60°C for 30 seconds, and 72°C for 30 seconds; total of 35 cycles) using an additional 50 pmol of oligonucleotides U1 and L4 (Table 1) as flanking primers. The expected 201-bp DNA fragment was amplified, flanked by *Vsp*I restriction sites (ATTAAT). After gel extraction, this product was digested using *Vsp*I and ligated into the *Nde*I-linearized pMA5ΔMCS plasmid, thereby eliminating the original *Nde*I site of the pMA5 vector. The resulting plasmid was sequenced and named pBSMuL1. Compared with the original pMA5 vector, the cloning site is directly downstream of the strong *Hpa*II promoter [3];

therefore, no additional cloning steps are necessary to get the target gene under control of the promoter. To achieve an even higher expression and secretion level, we decided to insert a second constitutive promoter P_{59} from *Streptococcus cremoris* [12] upstream of the *Hpa*II promoter. The P_{59} promoter sequence (5′-CTTGA-CAGGGAGAGATAGGTTTGATAGAATATAATAGT TGTCGCG-3′) was synthesized *in vitro* as previously described by using the oligonucleotides P59-up and P59-low (Table 1). The final 65-bp-long DNA fragment containing the P_{59} promoter was cloned directly into the unique *Bst*XI site of pBSMuL1. The resulting plasmid was sequenced and named pBSMuL2 (Fig. 1).

To confirm the functionality of these vectors, we chose cutinase from the fungus *F. solani pisi* [5] as a heterologous enzyme for expression in *B. subtilis*. The cutinase gene was amplified from *E. coli* expression plasmid pMac5-8 [10] using the primer pair cuti1/cuti2 (Table 1) containing restriction sites for *Eco*RI and *Xho*I without including the cutinase stop codon at the 3′-end. The cutinase gene was amplified using the standard PCR protocol (95°C for 30 seconds, 60°C for 30 seconds, and 72°C 30 seconds; total of 35 cycles). The 0.7-kb PCR product was cloned into the MCS of pBSMuL1 and pBSMuL2 to obtain an in-frame fusion of the cutinase gene and the 6x-His-sequence. The resulting plasmid, pBSMuL1-Cut, was used to transform the lipase deficient *B. subtilis* strain TEB1030 [4], which lacks both genes of the extracellular lipolytic enzymes LipA and LipB as well as the extracellular proteases NprE and AprE. Consequently, the strain TEB1030 shows no disturbing lipolytic background activity and less proteolytic degradation in the supernatant. The resulting *Bacillus* strain expressing the cutinase was cultured in Luria-Bertani (LB) medium supplemented with kanamycin (50 μg/mL) for 20 hours at 37°C. The cul-
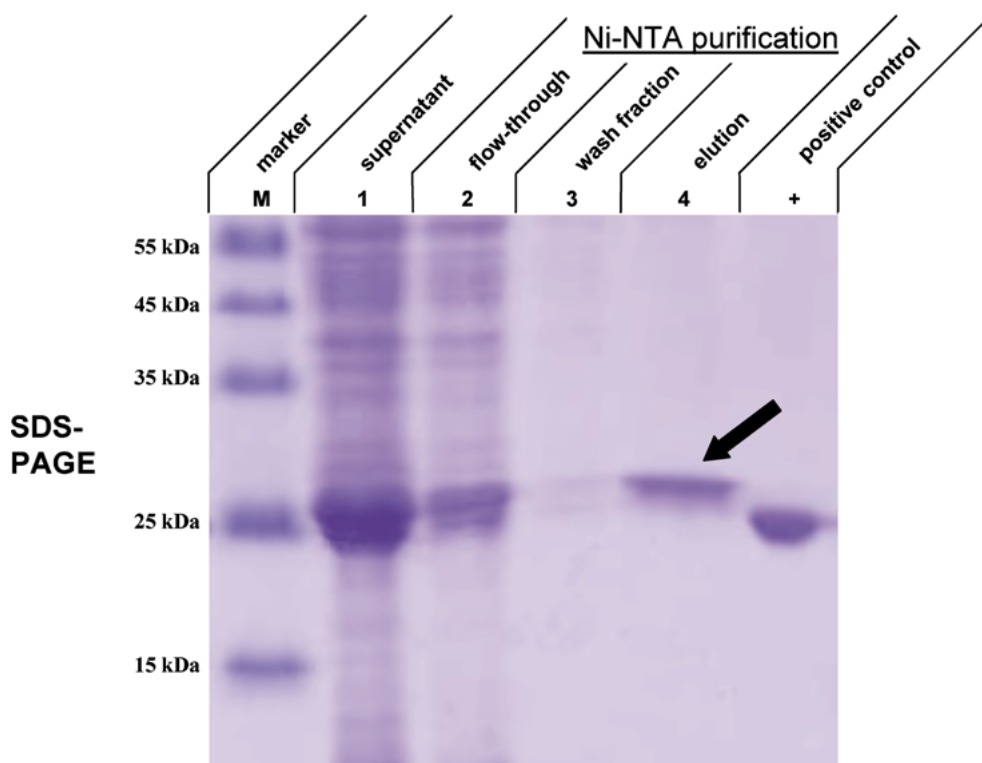
Fig. 2. Purification of His-Tag cutinase from the *B. subtilis* culture supernatant. The expression culture was grown for 20 hours at 37°C in LB medium supplemented with kanamycin (50 μg/mL). 600 μL (pH 8.2) of culture supernatant was loaded on a Ni-NTA Spin Column (Qiagen, Hilden, Germany), and the His-tagged cutinase was purified as recommended by the manufacturer's standard protocol. Protein of the supernatant, 25 μg (Lane 1), and comparable volumes of flow-through (Lane 2), wash fraction (Lane 3), and elution (Lane 4) were concentrated using 70% trichloride acid precipitation and separated on a protein gel (12% SDS-PAGE) stained with Coomassie brilliant blue. The protein band of the His-Tag cutinase (23.6 kDa) is marked by an arrow. As positive control, 2 μg purified native cutinase expressed in *E. coli* (22.2 kDa) was loaded (lane +).

ture supernatant was prepared by 5 minutes of centrifugation at 10,000g and microfiltration (0.22-μm membrane filter pore size; Millipore, Billerica, MA). Approximately 25 μg total protein was separated using sodium dodecylsulfate–polyacrylamide gel electrophoresis (SDS–PAGE) showing the successful overexpression and efficient secretion of 20 mg cutinase/L (Fig. 2, Lane 1). Using the Ni-NTA Spin Column (Qiagen, Hilden, Germany) as recommended by the manufacturer's standard protocol, the recombinant cutinase was purified to electrophoretic homogeneity as detected by SDS-PAGE (Fig. 2, lane 4). The enzyme was expressed and secreted in catalytically active conformation as determined spectrophotometrically using *p*-nitrophenyl-palmitate (pNPP) as the substrate [14]. The specific activity of the His-tagged cutinase towards pNPP was 200 U/mg and thus similar to the native cutinase expressed and secreted without His-tag in *B. subtilis*. Additional immunoblot analysis revealed no precursor form of the cutinase in the medium or in the cytoplasm (data not shown). Thus, the protein is completely se-

creted, processed, and released in an active form using the signal peptide of LipA. However, this signal sequence shows a twin arginine motif and was postulated to be recognized by the Tat pathway in *B. subtilis* [11]; however, recent studies revealed only a Sec-dependent secretion [6]. Furthermore, compared with other signal sequences, the LipA signal peptide was selected because of its high secretion efficiency.

For higher production yields, the second plasmid pBSMuL2 was tested in cutinase overexpression and secretion. Because of the additional constitutive promoter P_{59}, we were able to increase the secreted amount of mature cutinase three times. Recombinant cutinase, 60 mg/L, was detected in the culture supernatant by activity determination using pNPP as the substrate (Fig. 3).

As demonstrated here, the combination of two constitutive promoters leads to high expression and secretion levels without the necessity of induction. Therefore, the pBSMuL vector system can be applied in large-scale fermentations to avoid cost-intensive induc-
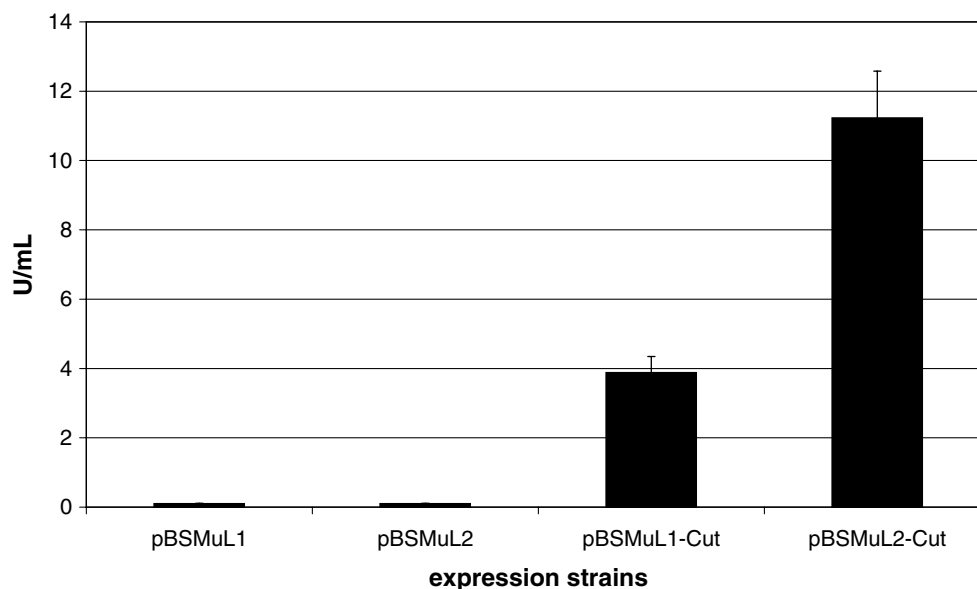
Fig. 3. Expression and secretion efficiency of cutinase in *B. subtilis* using vectors pBSMuL1 and pBSMuL2. The lipase-deficient strain *B. subtilis* TEB1030 [4] was transformed with the cutinase overexpression plasmids pBSMuL1-Cut and pBSMuL2-Cut. The strains were grown in LB medium for 20 hours at 37°C until the culture supernatants were isolated by centrifugation. As negative controls, *B. subtilis* TEB1030 was transformed with the empty vectors pBSMuL1 and pBSMuL2. All culture supernatants were tested for lipolytic activity using the spectrophotometric pNPP assay [14]. The results represent data from five independent experiments. LB, Luria-Bertani.

ers (e.g., isopropyl-β-D-thiogalactoside). Furthermore, it can be used for high throughput expression of variant libraries used in directed evolution experiments. Here, the omission of induction usually helps to avoid high fluctuations in expression rates when cultures are grown in microtiter plates (culture volumes ≤ 1 mL). Nevertheless, the use of constitutive promoter(s) might cause problems in expression of highly toxic proteins. However, protein export using the Sec pathway keeps the exoprotein in an unfolded transport competent and therefore most likely in an inactive conformation until it has passed the cell membrane.

Both pBSMuL-plasmids multiply by so-called rolling-circle–type replication (repB from *Staphylococcus aureus*), which might lead to the disadvantage of segregational instability [2]. Therefore, we tested the stability of both plasmids under selective and nonselective conditions. The strains were cultured for up to 24 hours in LB medium at 37°C with and without antibiotic selection. No decrease in cutinase secretion was detected under nonselective growth conditions compared with the control *Bacillus* strain, which was cultured under antibiotic-selection pressure. Even after inoculation of fresh LB medium and another 24 hours of growth at 37°C, no differences were obvious. However, after another round of inoculation and cultivation in fresh medium, the expression strain started to lose cutinase activity because of plasmid instability as described for repB-containing

vectors [2]. Therefore, even batch cultures grown under nonselective conditions inoculated from selective starter cultures will lead to sufficient protein production and secretion as demonstrated.

In summary two multicopy *B. subtilis* expression vectors for cloning of target genes under control of one or two strong constitutive promoter(s) have been constructed. The novel plasmids enable convenient cloning into a designed multiple-cloning site and high-level overexpression without the necessity of induction and efficient secretion of heterologous proteins in the Gram-positive host *B. subtilis*. This system might be useful as an alternative in case of serious problems concerning the well-established *E. coli* expression system.

Literature Cited

1. Baneyx F (1999) Recombinant protein expression in *Escherichia coli*. Curr Opin Biotechnol 10:411–421
2. Bron S, Luxen E (1985) Segregational instability of pUB110-derived recombinant plasmids in *Bacillus subtilis*. Plasmid 14:235–244
3. Dartois V, Coppée JY, Colson C, Baulard A (1994) Genetic analysis and overexpression of lipolytic activity in *Bacillus subtilis*. Appl Environ Microbiol 60:1670–1673

4. Eggert T, Brockmeier U, Dröge MJ, Quax WJ, Jaeger KE (2003) Extracellular lipases from *Bacillus subtilis*: Regulation of gene expression and enzyme activity by amino acid supply and external pH. FEMS Microbiol Lett 225:319–324

5. Egmond MR, de Vlieg J (2000) *Fusarium solani pisi* cutinase. Biochimie 82:1015–1021

6. Jongbloed JDH, Antelmann H, Hecker M, Nijland R, Bron S, Airaksinen U, et al. (2002) Selective contribution of the twin-arginine translocation pathway to protein secretion in *Bacillus subtilis*. J Biol Chem 277:44068–44078

7. Lam KH, Chow KC, Wong WK (1998) Construction of an efficient *Bacillus subtilis* system for extracellular production of heterologous proteins. J Biotechnol 63:167–177

8. Nagarajan V, Albertson H, Chen M, Ribbe J (1992) Modular expression and secretion vectors for *Bacillus subtilis*. Gene 114:121–126

9. Schallmey M, Singh A, Ward OP (2004) Developments in the use of *Bacillus* species for industrial production. Can J Microbiol 50:1–17

10. Schmid RD, Alberghina L, Verger R (1991) Lipases-structure, function and genetic engineering. GBF Monographs VCH Weinheim 16:243–251

11. Tjalsma H, Bolhuis A, Jongbloed JDH, Bron S, van DijL JM (2000) Signal peptide-dependent protein transport in *Bacillus subtilis*: A genome-based survey of the secretome. Microbiol Mol Biol Rev 64:515–547

12. van der Vossen JMBM, van der Lelie D, Venema G (1987) Isolation and characterization of *Streptococcus cremoris* Wg2-specific promoters. Appl Environ Microbiol 53:2452–2457

13. van Wely KHM, Swaving J, Freudl R, Driessen AJM (2001) Translocation of proteins across the cell envelope of Gram-positive bacteria. FEMS Microbiol Rev 25:437–454

14. Winkler UK, Stuckmann M (1979) Glycogen, hyaluronate, and some other polysaccharides greatly enhance the formation of exolipase by *Serratia marcescens*. J Bacteriol 138:663–670

15. Zyprian E, Matzura H (1986) Characterization of signals promoting gene expression on the *Staphylococcus aureus* plasmid pUB110 and development of a gram-positive expression system. DNA 5:219–225

BIOTECHNOLOGICAL PRODUCTS AND PROCESS ENGINEERING

Jean Detry · Thorsten Rosenbaum ·
Stephan Lütz · Doris Hahn · Karl-Erich Jaeger ·
Michael Müller · Thorsten Eggert

Biocatalytic production of enantiopure cyclohexane-*trans*-1,2-diol using extracellular lipases from *Bacillus subtilis*

Abstract Two extracellular lipases from *Bacillus subtilis*, *B. subtilis* lipase A and lipase B, have been expressed in the heterologous host *Escherichia coli*, biochemically characterized and used for the kinetic resolution of (*rac*)-*trans*-1,2-diacetoxycyclohexane. Both enzymes were selectively acting on the (*R,R*)-enantiomer of the racemic substrate, highly specifically hydrolyzing only one of the two ester groups present, thus allowing the preparation of enantiopure (*R,R*)- and (*S,S*)-cyclohexane-*trans*-1,2-diol. The reaction conditions for the use of purified enzyme and crude cell lyophilizate were optimized and reactions in batch and repetitive batch modes were carried out on a preparative scale to yield enantiopure product (>99% enantiomeric excess).

Introduction

Lipases (triacylglycerol hydrolases, EC 3.1.1.3) are widely distributed throughout animals, plants, and microorganisms. For many years, these enzymes have attracted enormous attention because of their biotechnological potential, i.e., their ability to catalyze both hydrolysis and synthesis

Jean Detry and Thorsten Rosenbaum have equally contributed to this work.

J. Detry · S. Lütz · D. Hahn · M. Müller
Institut für Biotechnologie 2, Forschungszentrum Jülich GmbH,
D-52425 Jülich, Germany

T. Rosenbaum · K.-E. Jaeger · T. Eggert (✉)
Institut für Molekulare Enzymtechnologie,
Heinrich-Heine-Universität Düsseldorf,
Forschungszentrum Jülich GmbH,
D-52426 Jülich, Germany
e-mail: t.eggert@fz-juelich.de
Tel.: +49-2461-612939
Fax: +49-2461-612490

M. Müller
Institut für Pharmazeutische Wissenschaften,
Albert-Ludwigs-Universität Freiburg,
D-79104 Freiburg, Germany

reactions with high regio- and enantioselectivity. Lipolytic enzymes are therefore the most widely used class of enzymes in organic chemistry for the synthesis of fine chemicals in the field of therapeutics, agrochemicals, and cosmetics (Jaeger and Eggert 2002; Liese and Lütz 2004; Buchholz et al. 2005). Presently, more than 800 lipase genes have been identified (Fischer and Pleiss 2003), and in many cases, the corresponding enzymes were overexpressed in homologous or heterologous expression hosts for characterization (Jaeger et al. 1999; Jaeger and Rosenau 2004).

Nowadays, more than 20 lipolytic enzymes have been identified and characterized from mesophilic and thermophilic *Bacillus* species, including lipases and esterases from *B. subtilis* (Dartois et al. 1992; Eggert et al. 2000; Dröge et al. 2005), *B. licheniformis* (Nthangeni et al. 2001), *B. pumilus* (Möller et al. 1991; Kim et al. 2002; Rasool et al. 2005), *B. megaterium* (Ruiz et al. 2002; Sekhon et al. 2005), *B. sphaericus* (Rahman et al. 2003), *B. circulans* (Kademi et al. 2000) *Geobacillus thermocatenulatus* (Schmidt-Dannert et al. 1994, 1996, 1997), *G. stearothermophilus* (Kim et al. 1998; Tyndall et al. 2002; Ewis et al. 2004), and *Bacillus* sp. (Dharmsthiti and Luchai 1999; Ruiz et al. 2003; Karpushova et al. 2005). Most of them share high sequence homology and therefore belong to one group, the family I.4 of true lipases according to the classification of bacterial lipolytic enzymes suggested by Arpigny and Jaeger (Arpigny and Jaeger 1999; Jaeger and Eggert 2002). So far, the two secreted lipolytic enzymes from *B. subtilis*, *B. subtilis* lipase A (BSLA) and lipase B (BSLB), constitute the best characterized members of family I.4 with respect to gene regulation (Dartois et al. 1992; Eggert et al. 2001, 2003), 3D-structure (van Pouderoyen et al. 2001), substrate specificity (Lesuisse et al. 1993; Eggert et al. 2000), and enantioselectivity (Dröge et al. 2003; Funke et al. 2003). However, their potential as useful biocatalysts on a preparative scale has not been demonstrated yet.

Enantiopure *trans*-cyclohexane-1,2-diols are useful as chiral intermediates for the synthesis of pharmaceuticals, agrochemicals, or crown ethers. The resolution of a racemic

mixture of *trans*-cyclohexane-1,2-diol is expensive and so far, biocatalysts remain the most efficient tool to obtain enantiopure (*R,R*)- or (*S,S*)-*trans*-cyclohexane-1,2-diol. A recent Japanese patent reported the biotransformation of cyclohexene oxide in (*R,R*)-*trans*-cyclohexane-1,2-diol (Nagai et al. 2003); however, only the (*R,R*)-enantiomer is accessible by this transformation. The use of hydrolases for the selective hydrolysis of one enantiomer of the diacetylated *trans*-cyclohexane-1,2-diol makes it possible to recover both enantiomers with high enantiomeric purity if the reaction is complete and enantioselective toward one enantiomeric form of the substrate.

Kawai et al. were the first to report a whole-cell biocatalytic hydrolysis of racemic diacetylated *trans*-1,2-cyclohexanediol (*rac*-**2**). They used *Rhizopus nigricans* and reported the formation of (*R,R*)-**1** with an enantiomeric excess (ee) of 50% (Kawai et al. 1981). Afterwards, the group of Schneider studied the hydrolysis of the diacetylated derivatives of (*rac*)-*trans*-1,2-cyclohexanediol by pig liver esterase (PLE) (Crout et al. 1986) and by a lipase from *Pseudomonas sp.* (SAM II) (Laumen et al. 1989; Seemayer and Schneider 1991). Crout et al. obtained both (*R,R*)-*trans*-1,2-cyclohexanediol [(*R,R*)-**1**] and (*S,S*)-**2** with an ee superior to 95%, and with 50% conversion. These authors used 33–67 U/mmol of substrate (0.3 mg PLE/mmol substrate), but precise reaction times were not reported (Crout et al. 1986). The same group described later that this transformation was unsatisfactory on a preparative scale because large quantities of racemic monoacetate were obtained (Laumen et al. 1989). Laumen et al. and Seemayer et al. obtained the (*R,R*)-monoacetylated diol [(*R,R*)-**3**] and the (*S,S*)-diacetylated diol [(*S,S*)-**2**] with ees of 96 and 97%, respectively, using *Pseudomonas* sp. (SAM II) lipase. The *E*-ratio was ≥100, the conversion reached 25%, and the reaction was stopped after 22 h (Laumen et al. 1989; Seemayer and Schneider 1991).

Kinetic resolution of *rac*-**2** was also performed using *Pseudomonas cepacia* lipase. Starting from 10 g of the racemic substrate (40 mg lipase/mmol substrate), Caron and Kazlauskas obtained 42% of (*R,R*)-*trans*-1,2-cyclohexanediol [(*R,R*)-**1**] and 38% of (*S,S*)-**2** after 7 days, both with ees >99% (Caron and Kazlauskas 1991).

In an enzyme screening approach, Reymond and coworkers characterized the activity of different hydrolases using the two enantiomers of the diacetylated *trans*-1,2-cyclohexanediol **2** among other diols. They showed that these enzymes failed to hydrolyze (*S,S*)-**2** and only three of the hydrolases exhibited activity toward (*R,R*)-**2**. Those enzymes were *Electrophorus electricus* acetylcholinesterase (Fluka F-01022), PLE (Fluka F-46058), and the so-called Proteus esterase number 7, a thermophilic esterase isolated from the culture collection of the French company Proteus. (Wahler et al. 2004).

The group of Poppe resolved *rac*-**2** on a preparative scale (500 mg of substrate) with lipase AK from Amano. They applied more than 1,200 U/mmol substrate (60 mg lipase/mmol substrate), and 21% of (*R,R*)-**3** was formed after 24 h at room temperature with an ee of 99% (Bodai et al. 2003).

In this study, the extracellular lipases from *B. subtilis* BSLA and BSLB were cloned and expressed in the heterologous host *Escherichia coli* BL21(DE3) as His-tagged fusion proteins, thereby overcoming productivity limitations observed with *B. subtilis* as the expression host (Lesuisse et al. 1993). In high cell density fed-batch fermentations (30 L-volume), we were able to produce 8–12 g each of lipases BSLA and BSLB. These enzyme preparations were used to investigate the potentials of both lipases as efficient biocatalysts in the chiral resolution of *trans*-1,2-cyclohexanediol (Scheme 1). After optimization of the reaction conditions, the reaction was performed on a preparative scale using purified enzyme and crude cell extracts. The enzymatic hydrolysis was also characterized and applied in a repetitive batch reaction. We demonstrate that both enzymes have a high hydrolytic activity toward (*R,R*)-**2** and can be applied advantageously on a preparative scale.

Materials and methods

Chemicals

(*rac*)-*trans*-1,2-Cyclohexanediol [(*rac*)-**1**] and silica gel 60 (0.040–0.063 mm) were purchased from Merck, (*R,R*)- and (*S,S*)-*trans*-1,2-cyclohexanediol [(*R,R*)- and (*S,S*)-**1**] were purchased from Fluka. Acetylated compounds, as depicted in Table 4, were synthesized by standard acetylation procedures starting from commercially available (Fluka, Aldrich) alcohols.

The substrate (*rac*)-*trans*-1,2-diacetoxy-cyclohexane [(*rac*)-**2**] was synthesized by chemical acetylation of (*rac*)-*trans*-1,2-cyclohexanediol [(*rac*)-**1**]. A solution of 500 mg (*rac*)-**1** (4.3 mmol), dissolved in 8.2 mL of acetic acid anhydride (86 mmol), was mixed with 8.2 mL pyridine and a catalytic amount of 4-dimethylaminopyridine. After 12 h of incubation at room temperature, 10 mL of water was added to the reaction medium. The reaction product was extracted three times using 7 mL ethyl acetate (NaCl was added to get a better phase separation). The combined organic layers were washed with 7 mL of a saturated solution of $NaHCO_3$ and with 7 mL of brine and then dried with anhydrous Na_2SO_4, filtered, and evaporated. The reaction product was purified by flash column chromatography (silica gel, petrol ether/ethyl acetate 1:1).
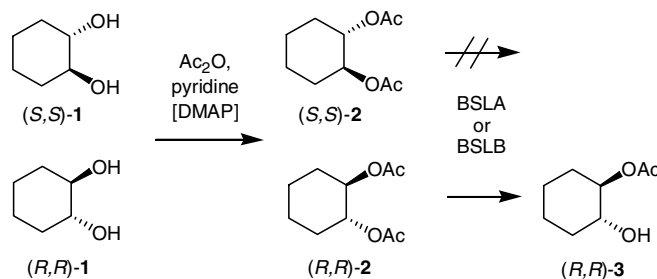
Scheme 1 *Bacillus subtilis* lipase (BSLA or BSLB) catalyzed kinetic resolution of 1,2-diacetoxycyclohexane

GC- and NMR-analysis

GC–mass spectrometry (MS) analysis was performed on a GC HP 6890 series (Agilent) combined with a mass spectrometer detector HP 5973 (injector at 250 °C, detector at 200 °C). The column used for the separation was a HP 19091S-433 (30 m×0.25 mm×0.25 μm film of HP-5MS 5% Phenyl Methyl Siloxane). GC quantification of the hydrolysis yield of the (rac)-trans-1,2-diacetoxycyclohexane was performed on a GC Chrompack CP9002 equipped with a flame ionization detector. The column used for the separation was a FS-Cyclodex β-I/P (CS 32324-3) (50 m×0.32 mm×0.46 μm film). Chiral GC analysis was performed on a Shimadzu GC-17A with a CHROMPACK WCOT fused silica column (25 m×0.25 mm×0.25 μm film of P Chirasil-DEX CB DF).

^{1}H-NMR spectra were recorded on a Bruker AMX 300 spectrometer (300 MHz, CDCl$_3$, 20 °C). The abbreviations presented have the following meanings: s for singlet, d for doublet, t for triplet, q for quartet, and m for multiplet. The analytical characteristics of the substances encountered in this study are as follows:

(rac)-trans-1,2-diacetoxy-cyclohexane [(rac)-**2**]	
Yellow oil	
^{1}H-NMR	(300 MHz, CDCl$_3$, 20 °C): δ [ppm]=1.30–1.42 (m, 4H), 1.71–1.74 (m, 2H), 2.05 (s, 6H, OAc), 2.07 (m, 2H), 4.78–4.82 (m, 2H)
(R,R)-trans-1-acetoxy-cyclohexan-2-ol [(R,R)-**3**]	
White crystals	
Mp	65.5 °C
$[\alpha]_D^{23}$	−45 (c=0.8 in CHCl$_3$)
^{1}H-NMR	(300 MHz, CDCl$_3$, 20 °C): δ [ppm]=1.23–1.32 (m, 4H), 1.69–1.73 (m, 2H), 2.01–2.05 (m, 2H), 2.09 (s, 3H, OAc), 2.21 (s, 1H, OH), 3.51–3.59 (m, 1H), 4.53–4.60 (m, 1H)
^{13}C-NMR	(75 MHz, CDCl$_3$, 20 °C): δ [ppm]=21.29 (CH$_3$), 23.74 (CH$_2$), 23.85 (CH$_2$), 29.93 (CH$_2$), 33.03 (CH$_2$), 72.77 (CH), 78.27 (CH), 171.32 (C=O)

Bacterial strains, plasmids and growth conditions

The strains and plasmids used in this study are listed in Table 1. *Bacillus subtilis* TEB 1030 (Eggert et al. 2003) was used as the homologous expression host for plasmid-encoded lipases. *Escherichia coli* XL1-blue *MRF'* was used as a host for cloning, and *E. coli* BL21(DE3) was used as a heterologous expression host for plasmid-encoded enzymes. *Escherichia coli* and *B. subtilis* were grown overnight in 5 mL Luria-Bertani (LB) medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl) in glass tubes at 37 °C. Expression cultures for protein purification were grown in 1 L LB/M9 medium (10 g/L tryptone, 5 g/L yeast extract, 5.5 g/L NaCl, 4 g/L glucose, 0.25 g/L MgSO$_4$×7-H$_2$O, 0.02 g/L CaCl$_2$, 7 g/L Na$_2$HPO$_4$×2H$_2$O, 3 g/L KH$_2$PO$_4$, 1 g/L NH$_4$Cl) in 5-L Erlenmeyer flasks. These cultures were inoculated at a starting cell density of OD$_{580}$=0.05 with 10 mL of overnight cultures. Plasmid-carrying *B. subtilis* and *E. coli* cells were selected with 50 μg/mL kanamycin and 100 μg/mL carbenicillin, respectively.

Lipase expression in the heterologous host *Escherichia coli*

Plasmid construction The *lipA* gene lacking its native signal sequence was PCR-amplified using the primer pair upA19–22/downA19 and upA19–22/downA22 (Table 2) for *Nde*I/*Xho*I cloning in vectors pET19b and pET22b(+), respectively. Depending on an internal *Nde*I restriction site in *lipB*, we first introduced a silent TCA to TCT mutation at codon position 58. For this purpose, we designed the mutagenesis primer BSLBNdeX (Table 2) and performed site-directed mutagenesis by using the megaprimer PCR-method (Barettino et al. 1994) as described before (Eggert et al. 2000). Afterwards, the *lipB* gene was amplified and cloned in the same way as *lipA* using the primer pair upB19–22/downB19 and upB19–22/downB22 (Table 2). The expression plasmids derived from pET19b resulting in N-terminal 10×His-tag fusions were named pTlipA_N10H and pTlipB_N10H (Table 1). The plasmids derived from cloning into pET22b(+) resulting in C-terminal 6×His-tag fusions were named pTlipA_C6H and pTlipB_C6H (Table 1).

Small-scale expression conditions The expression plasmids were transformed into *E. coli* BL21(DE3). The resulting expression strains were grown at 37 °C in LB/M9 medium inoculated with overnight culture until they reached an optical density of OD$_{580}$=0.6. Recombinant lipase expression was induced by adding isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 0.6 mM. Four hours after induction, the cells were harvested by centrifugation and resuspended in lysis buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 30 mM imidazole, pH 8.0), yielding a 20% (w/v) cell suspension. The cells were lysed by adding 200 μL lysozyme (100 mg/mL) and by ultrasonication. The cell extracts were clarified by centrifugation and subsequent filtration (5 μm pore size) and stored at −20 °C or used immediately for purification.

Large-scale expression conditions For large-scale production of BSLA and BSLB, high cell density fed-batch fermentations in 30-L Infors fermenters were performed as described previously (Korz et al. 1995; Ansorge and Kula 2000).

Purification of His-tagged *B. subtilis* lipases

Lipase purification from 100 g of cell material was performed by immobilized-metal affinity chromatography (IMAC) (Porath et al. 1975; Hochuli et al. 1987) using a 30-mL Ni-nitrilo-triacetic acid (NTA) superflow column (QIAGEN, Hilden, Germany). The column was equilibrated with lysis buffer (five column volumes) and the cell extract was applied at a flow rate of 0.5 mL/min. The

column was washed with washing buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 50 mM imidazole, pH 8.0) until UV-absorption (280 nm) reached the base-line again. His-tagged BSLA and BSLB were eluted with elution buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 250 mM imidazole, pH 8.0). In the final step, imidazole was removed by gel filtration chromatography (G-25 column, Amersham Pharmacia Biotech) using 2 mM glycine/NaOH buffer (pH 11.0). Enzyme samples were concentrated to 1 mg/mL by ultrafiltration using a 10-kDa cut-off membrane (polyethersulfone membrane in Vivacell250, Viva Science, Hannover, Germany).

Protein analysis methods

Protein concentration was measured at 595 nm according to the method of Bradford (1976) using bovine serum albumin (BSA) as the standard. *Lipolytic activity* of BSLA and BSLB was determined spectrophotometrically using *p*-nitrophenyl-palmitate (pNPP) as the substrate. The assay was performed as described (Eggert et al. 2000) using a Shimadzu UV1601 photometer and a reaction time of 5 min at 37 °C. *Sodium dodecyl sulfate-polyacrylamide gel electrophoresis* (SDS-PAGE) was performed using a 5% stacking gel and a 15% separating gel (Laemmli 1970).

Characterization of the hydrolytic activity and selectivity of *Bacillus subtilis* lipases toward different substrates

Five microliters of purified lipase solution was added to 895 μL 1 M Tris–HCl buffer (pH 7.5) and mixed with 100 μL of the acetylated substrate in DMSO (0.1 M). The samples were incubated for 24 h at room temperature. The reaction products and substrates were separated from the enzyme by solvent extraction. Next, 250 μL of the reaction medium was extracted with 250 μL of ethyl acetate. The phases were separated by centrifugation (13,000 rpm, for 5 min at room temperature) and the organic phase was analyzed by chiral GC or GC–MS.

Bacillus lipase stability and activity in lyophilized crude cell extracts

Preparation of lyophilized crude cell extracts from BSLB-overexpressing Escherichia coli. The *E. coli* BSLB expression strain [*E. coli* BL21 (DE3)+pTlipB_N10H; Table 1] was grown as described above in a 30-L fermenter under high cell density fed-batch conditions. One hundred grams of cell material was lyophilized overnight in a freeze-dryer (Lyovac GT2, STERIS, Mentor, OH, US).

Storage-stability of BSLB The lyophilized crude cell extract (18.7 μg BSLB per mg of lyophylisate) was dissolved in 50 mM sodium phosphate buffer (pH 8.0) and 50 mM glycine buffer (pH 11.0) (0.1 mg lyophylisate per mL), and was stored at −18, 4, 20, and 40 °C. At distinct time points, the remaining lipase activity of BSLB was determined spectrophotometrically using pNPP as the substrate.

Influence of temperature and pH on BSLB-activity The lyophilized crude cell extract was dissolved in 1 M sodium phosphate buffer (pH 8.0) (2 mg of cells per mL) and incubated at temperatures from 10 to 50 °C. The lipolytic activity at a given temperature was determined using the spectrophotometric assay. The pH-dependence of the BSLB-activity was tested using glycine (200 mM) or sodium phosphate buffer (200 mM) as indicated in Fig. 2a. Lyophilized cells were dissolved in the buffer (2 mg per mL) and incubated with (*rac*)-**2** (200 mM). The hydrolysis of **2** was determined by GC, with one unit of enzyme activity representing 1 μmol of substrate **2** hydrolyzed per minute and per milligram of lyophilized cells.

Hydrolysis of (*rac*)-*trans*-1,2-diacetoxy cyclohexane [(*rac*)-**2**] catalyzed by *Bacillus* lipase

Reaction performed on a 100-mg scale Two hundred fifty microliters of purified BSLA (1 mg/mL dissolved in 2 mM glycine/NaOH, pH 11.0) and 45 mL phosphate buffer 1 M (pH 7.0) were added to 100 mg of the acetylated substrate (*rac*)-**2** dissolved in 5 mL DMSO. The samples were stirred at 25 °C for 24 h. The reaction products were extracted with 3×10 mL ethyl acetate. The organic phases were recovered, combined, and concentrated (40 °C, 150 mbar). The products were separated by flash column chromatography (silica gel, petrol ether/ethyl acetate 1:1). The first product eluted was *trans*-1,2-diacetoxycyclohexane (**2**), essentially, the (*S,S*)-enantiomer, and the second was (*R,R*)-*trans*-1-acetoxycyclohexan-2-ol [(*R,R*)-**3**].

Reaction performed on a 1-g scale Purified BSLB (8.4 mg dissolved in 2 mM glycine/NaOH, pH 11.0) and 200 mL phosphate buffer (1 M, pH 7.0) were added to 1.0 g of the acetylated substrate (*rac*)-**2** dissolved in 10 mL DMSO. The samples were stirred at 25 °C for 24 h. The reaction products were extracted with 3×50 mL ethyl acetate. The organic phases were recovered, combined, and concentrated (40 °C, 150 mbar). The products were separated by flash column chromatography using isohexane/acetone 10:1, followed by ethyl acetate as the eluent.

Repetitive batch reaction The reaction was carried out in a 10-mL Amicon ultrafiltration cell with a YM3 membrane [pretreated in DMSO for swelling and then coated with protein by an aqueous solution of BSA (1 mg/mL)]. Twenty milligrams of lyophilized crude cell extract containing BSLB was dissolved in 5 mL reaction solution [1 M sodium phosphate buffer (pH 8.0) and DMSO 1:1].

Table 1 Bacterial strains and plasmids used in this study

Strain or plasmid	Genotype or description	Source or reference
Strain		
E. coli BL21(DE3)	F^- *ompT hsdS*$_B$(r_B^- m_B^-) *gal dcm* (λcIts857 ind1 Sam7 nin5 lacUV5-T7 gene1)	(Studier and Moffatt 1986) Novagen, Madison, USA
E. coli XL1-Blue *MRF'*	*recA1 endA1 gyrA96 thi1 hsdR17 supE44 relA1 lac*$^-$ [F' proAB lac^{q}ZΔM15 Tn10 (tet^r)]	(Bullock et al. 1987)
B. subtilis TEB1030	*His nprE aprE bpf ispI ΔlipA ΔlipB*	(Eggert et al. 2003)
Plasmids		
pET19b	ColE1 P$_{T7/lac}$ Amp^r	Novagen, Madison, USA
pET22b	ColE1 P$_{T7/lac}$ Amp^r	Novagen, Madison, USA
pTlipA_N10H	pET19b containing the *lipA* gene *NdeI/XhoI*	This study
pTlipA_C6H	pET22b containing the *lipA* gene *NdeI/XhoI*	This study
pTlipB_N10H	pET19b containing the *lipB* gene *NdeI/XhoI*	This study
pTlipB_C6H	pET22b containing the *lipB* gene *NdeI/XhoI*	This study

For each reaction, 200 mg of (*rac*)-**2** was added. After the completion of the reaction, the solution was ultrafiltrated and new substrate solution was added to the retentate. This was repeated four times. The collected product solution was purified using column chromatography (silica gel, hexane/ethyl acetate 1:1). Two hundred twenty-five milligrams of (*S,S*)-**2** (ee >99%) and 285 mg of (*R,R*)-**3** (ee >99%) were recovered.

Results

Expression and purification of BSLA and BSLB as poly-histidine fusion proteins

Four different expression plasmids were constructed to establish an efficient one-step purification procedure for BSLA and BSLB by affinity chromatography using IMAC (Porath et al. 1975; Hochuli et al. 1987). Recombinant plasmids were constructed by ligating the PCR-amplified lipase genes into the *NdeI/XhoI* restriction sites of vectors pET19b and pET22b to obtain in-frame fusions with an N-terminal 10×His-tag or a C-terminal 6×His-tag, respectively (Table 1). Expression plasmids were transformed into *E. coli* BL21(DE3), and lipase activity was detected after IPTG induction. All expression strains produced about one to three units of lipase activity per milliliter of culture suspension, and an additional protein band of M$_r$ 20 kDa was detected by SDS-PAGE analysis of whole-cell extracts (Fig. 1a).

All His-tagged lipases were purified from 100 mL expression cultures by IMAC using a 3-mL Ni-NTA superflow column (QIAGEN, Hilden, Germany). Here, purified samples from C-terminal 6×His-fusions of both BSLA and BSLB still contained residual protein contaminations. We suspect that the recombinant proteins exhibited a low binding affinity to the Ni-NTA matrix because they were eluted from the column already at low imidazole concentrations of 40 mM (data not shown). In contrast, the 10×His-tag N-terminal fusion lipases tightly bound to the Ni-NTA matrix and imidazole concentrations of 150 mM were required for elution, thereby allowing their one-step purification to electrophoretic homogeneity (Fig. 1b). The final yield of pure enzyme was 95 mg for BSLA-10×His and 22 mg for BSLB-10×His obtained from 1 L of expression culture (Table 3). The amounts of recombinant protein were further increased by high cell density fed-batch fermentation. In a final volume of 15 L fermentation broth, 1.5 kg of cells (wet weight) was produced containing $1.65×10^6$ units of BSLB. The enzy-

Table 2 Oligonucleotides used in this study

Primer	Nucleotide sequence (5'→3')	Modification
upA19–22	ATAT <u>CAT ATG</u> GCT GAA CAC AAT CCA GTC GTT ATG	*NdeI*
downA19	ATAT <u>CTC GAG</u> TCA **TTA** ATT CGT ATT CTG GCC CCC	*XhoI*, STOP
downA22	ATAT <u>CTC GAG</u> ATT CGT ATT CTG GCC CCC GCC G	*XhoI*
upB19–22	ATAT <u>CAT ATG</u> GAG TCA GTA CAT AAT CCT GTC G	*NdeI*
downB19	ATAT <u>CTC GAG</u>**TTA** ATT TGT ATT GAG GCC TCC GCC	*XhoI*, STOP
downB22	ATAT <u>CTC GAG</u> ATT TGT ATT GAG GCC TCC GCC ATT C	*XhoI*
BSLBNdeX	C CCG CAG CTT GCT <u>TCT</u> TAT GTT GAC CGT GTT	TCA→TCT
mut1low	GCCGCAAGCTTGTCGACGAGCTCTCATTA	
BSLB3	TATACCATGGAGTCAGTACATAATCCTGTCGTTCTT	

Inserted modifications are underlined and stop codons are printed in bold letters

Fig. 1 SDS-PAGE to analyze BSLA expression in the heterologous host *E. coli* (**a**) and its purification by metal chelate affinity chromatography using Ni-NTA (**b**). **a** *Lane M*, molecular marker; *lane 0*, cell extract prior to induction of lipase gene expression; *lanes 1–3*, cell extracts 1, 2, and 3 h after induction with ITPG. **b** *Lane 1*, cell extract obtained 3 h after induction; *lane 2*, flow-through of Ni-NTA column; *lane M*, molecular marker; *lanes 3 and 4*, fractions containing BSLA obtained by elution with 250 mM imidazole
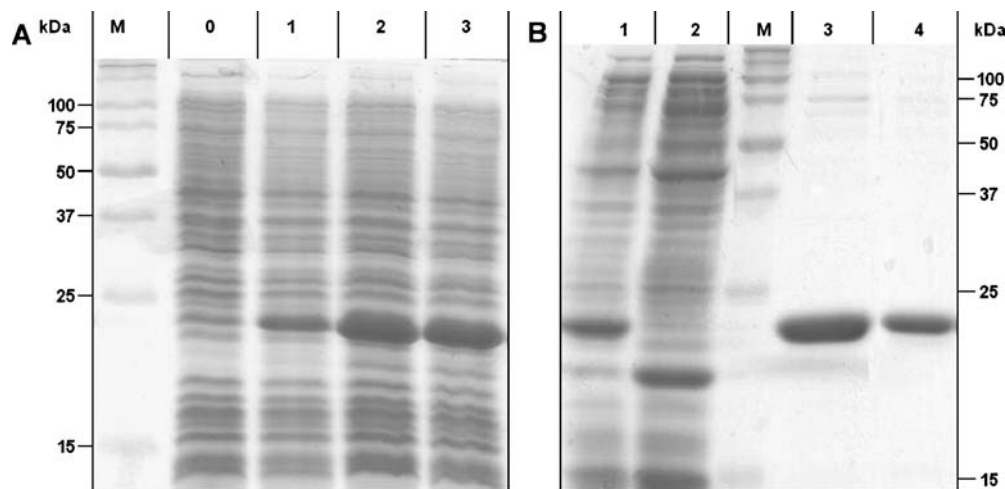
matic activity was determined spectrophotometrically using pNPP as the substrate. The final enzyme concentration was estimated to be about 6.5 mg per gram of cell wet weight, with a total yield of 10 g BSLB.

Properties of recombinant *Bacillus* lipases

Substrate specificities of purified enzymes The substrate specificities of recombinant enzymes BSLA-10×His and BSLB-10×His were compared to the wild-type enzymes purified from *B. subtilis* expresssion cultures (Eggert et al. 2000). Enzyme kinetics determined spectrophotometrically using the substrate *p*-nitrophenyl palmitate showed comparable results for wild-type and recombinant Histagged lipases. Native BSLA and BSLA-10×His had a maximum activity of 105±10 U/mg, and native BSLB and BSLB-10×His 162±15 U/mg of purified enzyme. The K_M-values were determined as 0.1 and 0.2 mM for recombinant BSLA and BSLB, respectively. Additionally, both *B. subtilis* lipases were tested for the hydrolysis of different chiral acetates of secondary alcohols as summarized in Table 4.

pH and temperature stabilities of purified enzymes Previous studies with BSLA and BSLB wild-type enzymes revealed a high stability at alkaline pH, with a maximum stability at pH 11 and a rapid loss of activity upon incubation below pH 5 (Eggert et al. 2000). The pH and temperature stability profiles for the recombinant enzymes BSLA- and BSLB-10×His were very similiar to those of the wild-type enzymes purified from the homologous host *B. subtilis* (data not shown).

Enantioselective hydrolysis of (*rac*)-*trans*-1,2-diacetoxycyclohexane [(*rac*)-**2**]

After 24 h of reaction, the GC–MS analyses showed that the lipases BSLA and BSLB were both active toward the (*R,R*)-*trans*-1,2-diacetoxycyclohexane [(*R,R*)-**2**] and did not accept (*S,S*)-**2** as a substrate (Scheme 1). It is noteworthy to report that in this case, as well as in the examples of the literature concerning the enzymatic resolution of (*rac*)-**2**, the hydrolases tested always showed hydrolytic activity toward the (*R,R*)-enantiomer preferentially, suggesting a structure homology in the active site of the different lipases used (Crout et al. 1986; Laumen et al. 1989; Caron and Kazlauskas 1991; Seemayer and Schneider 1991; Bodai et al. 2003).

The GC–MS analyses of the reaction products in the preliminary tests (pH 7.5) showed a substrate conversion rate of about 50% for BSLA and 20% for BSLB. The hydrolysis of (*R,R*)-**2** at 100% conversion would allow the separation of the enantiomers of the *trans*-1,2-cyclohexanediol (*rac*)-**1**. Therefore, two parameters were investigated in an attempt to improve the hydrolysis yields: the pH of the aqueous phase and the reaction temperature.

Table 3 Purification of recombinant BSLA and BSLB from cell extract

Enzyme	Purification step	Protein (mg)	Activity (U)	Specific activity (U/mg)	Recovery (%)
BSLA	Cell extract	950	35,450	37.3	100
	IMAC	95	9,975	105	28.1
BSLB	Cell extract	950	12,250	12.9	100
	IMAC	22	3,568	162	29.1

Table 4 Catalytic activities of BSLA- and BSLB-10×His towards unnatural substrates

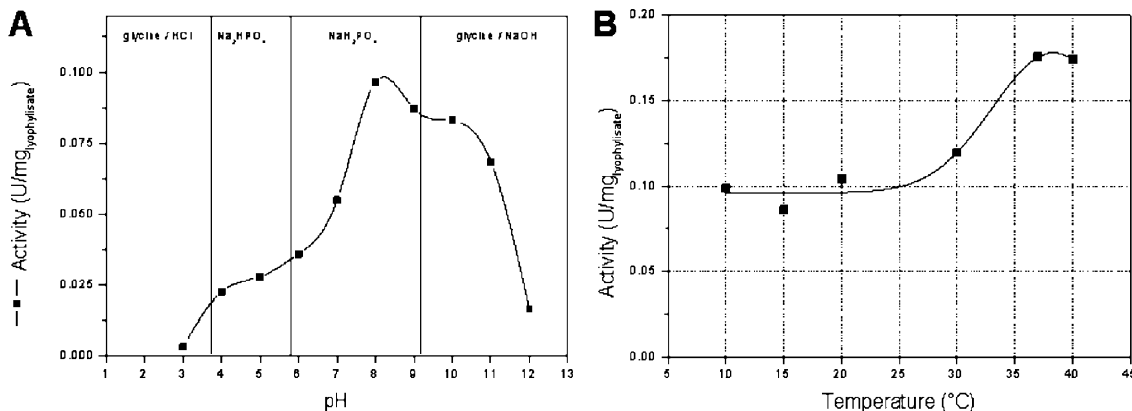| Acetylated substrates	Conversion (%)	
	BSLA-10×His	BSLB-10×His
(*rac*)-1-Octin-3-ol	46	28
(*rac*)-Trimethylsilylbutinol	41	5
(*R*)-3-Chloro-1-phenyl-1-propanol	75	67
cis-1,2-Cyclohexanediol	n.d.	n.d.
(*R,R*)-*trans*-1,2-Cyclohexanediol	90	54

n.d. no activity detectable

Fig. 2 a Influence of pH on BSLB activity of lyophilized cells incubated in 200 mM glycine or 200 mM sodium phosphate buffer. **b** Influence of temperature on BSLB activity of lyophilized cells incubated in 1 M sodium phosphate buffer (pH 8.0). A 2-mg crude enzyme preparation was used per milliliter reaction volume

Influence of pH and temperature of the reaction medium on the hydrolytic activity

The lipases BSLA and BSLB are highly tolerant to basic pH and have their optimum activity in the pH range of 8.0–10.0 (Eggert et al. 2001; van Pouderoyen et al. 2001). This suggested higher hydrolysis yields if the reaction would be performed at a pH higher than 7.5. Therefore, the reaction yields were determined with the isolated enzymes at pH 7.0 (phosphate buffer), pH 7.5 (phosphate buffer), and pH 9.0 (Tris–HCl), and the hydrolysis yields were determined by GC with (R,R)-2 as the substrate. The experiments were performed at 25 °C for 24 h and the reaction medium was stirred at 300 rpm. One blank sample was prepared and each reaction was repeated three times. Nonselective chemical hydrolysis occurred in the blank sample at pH 7.5 and increased at pH 9.0, forming the substrate diol. Therefore, the experiment was performed at pH 7.0 where the chemical substrate hydrolysis was negligible. At this pH, the best hydrolysis yields were obtained with BSLB (88.3%), and consequently, this pH was selected for the following experiments.

The influence of temperature was investigated at 25, 30, and 37 °C, with enantiopure (R,R)-2 as substrate. One

blank sample was prepared and each hydrolase reaction was repeated three times. The product yields obtained with BSLB did not change over the range of temperatures tested. However, the chemical hydrolysis rate increased with increasing temperature; therefore, the following experiments were carried out at 25 °C. The results obtained at 30 and 37 °C suggested that BSLA is more active than BSLB toward this substrate. The yields obtained with BSLA at 30 °C were reproducible (relative standard deviation<5%) and exceeded those obtained with BSLB by 10%. Additionally, it was found that purified BSLB tended to precipitate from the stock solution, resulting in lower yields than for BSLA-catalyzed reactions. Stirring and homogenization of the enzyme solution prior to starting the reactions increased the yield to the range obtained with BSLA.

Properties of lyophilized *Escherichia coli* cells overexpressing BSLB

Having used purified enzymes, we decided to additionally investigate whether the biocatalytic reaction could also be performed by using lyophilized cells of *E. coli* over-
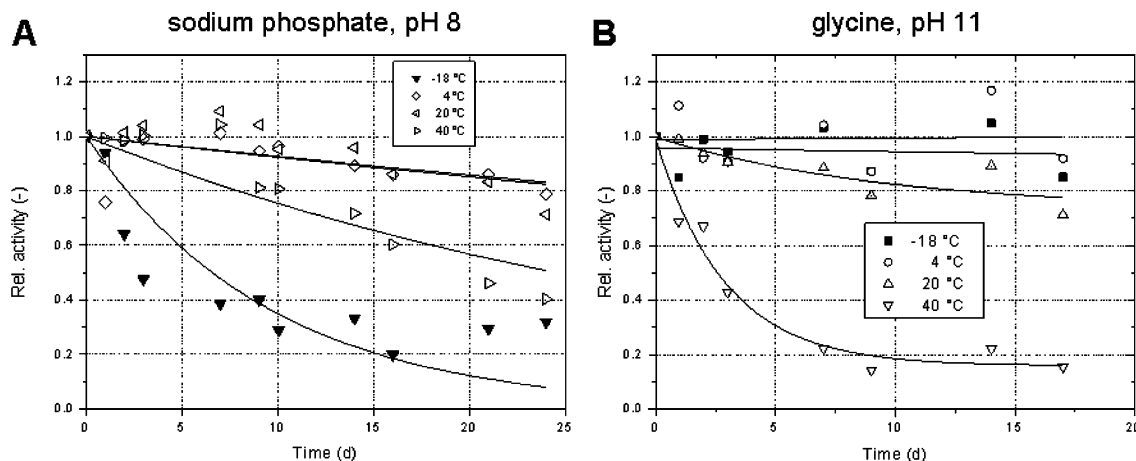

Fig. 3 Storage stability at different temperatures of lyophilized cells overexpressing BSLB and incubated in **a** 50 mM sodium phosphate (pH 8.0) and **b** 50 mM glycine buffer (pH 11). A 0.1-mg crude enzyme preparation was used per milliliter reaction volume

expressing BSLB. The lyophilized cell preparation contained 18.7 μg of His-tagged BSLB per milligram of biomass, with a specific activity 96 U/mg protein; this activity was in the same range as the activity of the native enzyme purified from *B. subtilis* (100 U/mg). In contrast to purified BSLB, the highest activity was found at pH 8.0 (Fig. 2a); therefore, all further experiments were carried out in sodium phosphate buffer (pH 8.0). Also, the temperature profile differed from that of the purified enzyme showing a significantly higher activity at 37 °C, as compared to 20 or 30 °C (Fig. 2b). On the other hand, the substrate was unstable at elevated temperatures under these conditions; therefore, the hydrolysis reactions were carried out at room temperature.

An interesting observation was made upon incubating the lyophilized cells in sodium phosphate and glycine buffer at different temperatures (Fig. 3). Only 25% of the initial enzymatic activity could be recovered after 3 weeks of storage at −18 °C in sodium phosphate buffer (pH 8.0), while there was no significant drop in activity observed after storage in glycine buffer (pH 11.0). The lyophilized cells could be stored at 4 and 20 °C in both buffers for more than 20 days without a significant loss of activity.

Control of the optimized reaction conditions on the analytical scale

The reactions performed at 25 °C, 300 rpm, and pH 7.0 gave the best hydrolysis yields for the substrate with BSLA and BSLB. Those reactions were performed with pure (R, R)-**2** as the substrate. The enzyme activity toward the enantiomer (S,S)-**2** and the racemic substrate mixture was tested under the same conditions and a blank sample was run, as well as three repetitive hydrolysis reactions. The experiments confirmed the selectivity of BSLA and BSLB toward (R,R)-**2**: Both enzymes did not hydrolyze the (S,S)-enantiomer. Furthermore, the yields were obtained with BSLA and BSLB in the range of 90% conversion for the racemic mixture and for the pure (R,R)-enantiomer.

Scale-up of the hydrolysis reaction

BSLA was chosen for the kinetic resolution of (rac)-**2** on a 100-mg scale. The ees of the two products were determined by chiral GC. Product (R,R)-**3** was isolated in 73.6% yield after flash column chromatography. The chemical purities of the recovered products were 92.4 and 88.2% for (S,S)-**2** and (R,R)-**3**, respectively. The ee of (R,R)-**3** was determined as >99% (GC); however, the separation on the silica gel column (eluent of petrol ether/ethyl acetate 1:1) was not optimal: The product fraction contained *trans*-1,2-diacetoxycyclohexane as an impurity.

Kinetic resolution on a gram scale was performed with BSLB as the catalyst. To reach complete hydrolysis, a higher ratio of enzyme to substrate (rac)-**2** was chosen (1.68 mg/mmol substrate). The proportions of DMSO and of phosphate buffer to the substrate were reduced in

Table 5 Product separation using isohexane/acetone 10:1 as eluent

Order of elution	Name	Quantity (mg)	Isolated molar yield (%)	ee (%)
1	(S,S)-*trans*-1,2-Diacetoxycyclohexane	190.3	36.7	>99
2	(R,R)-*trans*-1-Acetoxycyclohexan-2-ol	164.5	36.9	>99

comparison to the previous experiments. Chiral GC analysis performed on the raw product after extraction showed that the hydrolysis was total: Only the (S,S) *trans*-1,2-diacetoxycyclohexane and the (R,R) *trans*-1-acetoxycyclohexan-2-ol were found in the sample, both with ees >99.5%. The products were separated by flash column chromatography with isohexane/acetone 10:1 followed by ethyl acetate as the eluent (Bodai et al. 2003), resulting in a complete separation of (S,S)-**2** and (R,R)-**3**. The results are shown in Table 5.

Hydrolysis using lyophilized cells in a repetitive batch

The hydrolysis of (rac)-**2** was also carried out in a repetitive batch reaction (Fig. 4). In all five reactions, the enantioselectivity obtained for **3** was above ee 99%, proving that there were no unselective hydrolytic side activities towards the substrate present in the crude enzyme preparation. The maximal conversion of 50% was reached in all reactions in this series, although reaction time until completion lengthened significantly. The first reaction took only 23 h, as compared to more than 70 h in the last reaction, indicating a loss of enzyme activity under reaction conditions. From the five reactions in this series, total amounts of 225 mg of (S,S)-**2** (ee >99%) and 285 mg of (R, R)-**3** (ee >99%) were recovered by column chromatography.
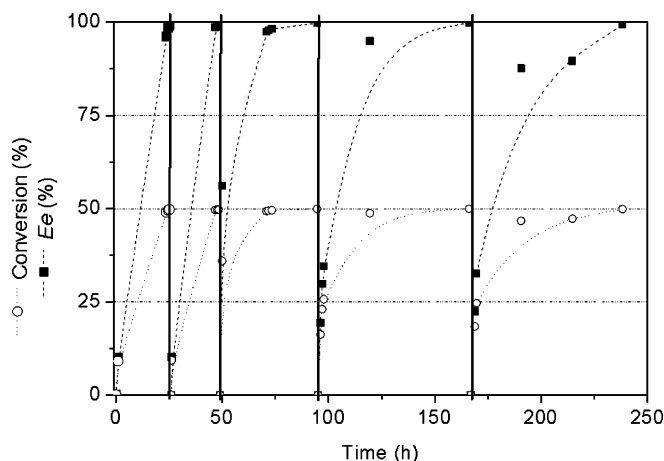

Fig. 4 Resolution of (rac)-**2** by lyophilized cells overexpressing BSLB in a repetitive batch [*lines* as optical aid, V=5 mL (1M sodium phosphate (pH 8.0) and DMSO 1:1)], Amicon stirrer cell (YM3 membrane), T=25 °C, 20 mg of crude BSLB preparation, 200 mg of rac-(**2**) per batch

Discussion

We showed in the present study that the recombinant lipases BSLA and BSLB from *B. subtilis* are highly interesting biocatalysts for biotechnological applications. Both enzymes can easily be produced and purified in high yield as His-tagged fusion proteins, and the reaction can also be carried out by using a cell lyophilisate, which shows high activity and enantioselectivity even on a preparative scale. The crude enzyme preparation could be stored for more than 20 days without significant loss of activity, but further investigations on the long-term stability still have to be carried out. The increasing loss of enzyme activity during the repetitive batch experiment, as identified by prolonged reaction times (Fig. 4), was higher than expected from the stability studies, which could be due to destabilizing effects on the enzyme of either the substrate or the product. Nevertheless, the lyophilized *E. coli* cells containing either BSLA or BSLB proved to be highly efficient biocatalysts due to the ease of their preparation and application.
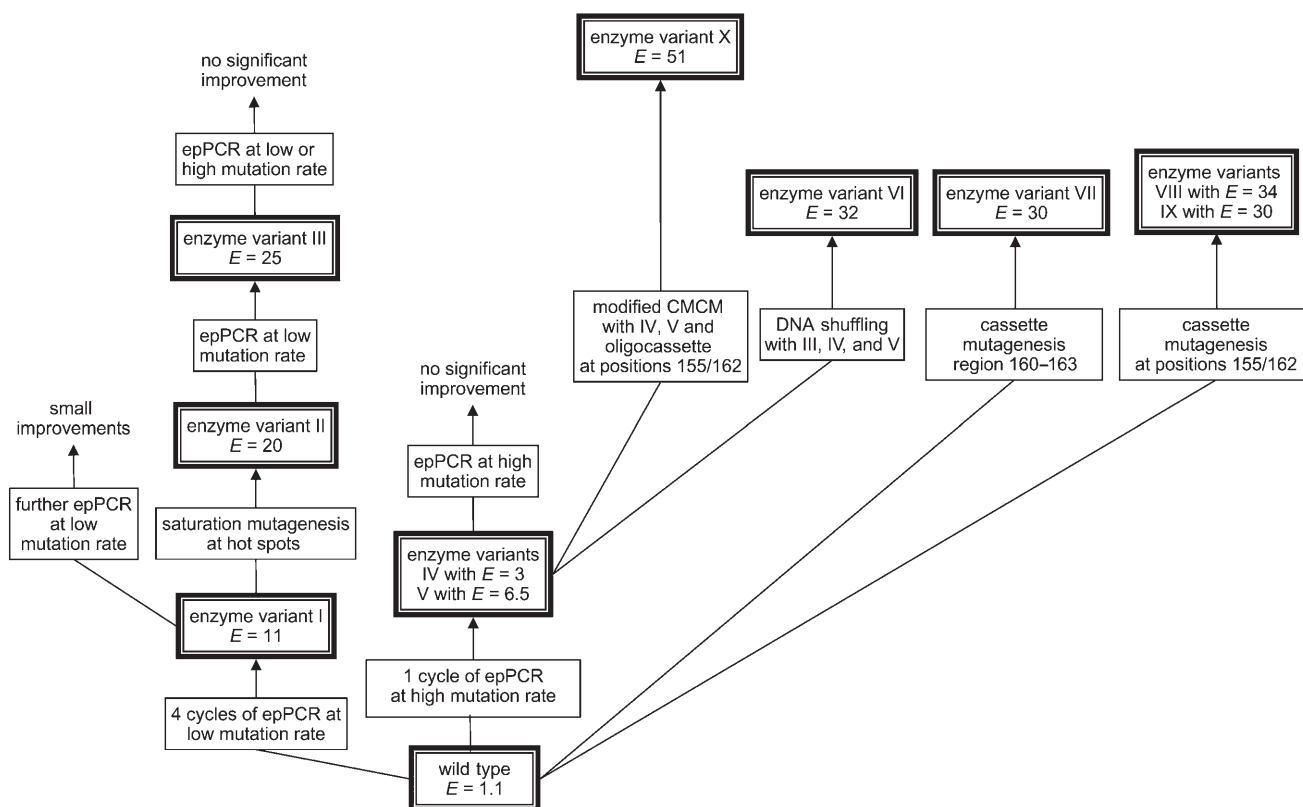
The lipases BSLA and BSLB catalyzed very efficiently the enantioselective hydrolysis of the bisacetylated (*R,R*)-enantiomer of cyclohexanediol. The reaction conditions were optimized so that (*R,R*)-*trans*-1,2-diacetoxycyclohexane was hydrolyzed quantitatively and both enantiomers could be separated with ee >99% on a preparative scale. Thus, the lipases BSLA and BSLB from *B. subtilis* were used here for the first time on a preparative scale, and an efficient enzymatic resolution of the substrate (*rac*)-*trans*-cyclohexane-1,2-diol was developed. In addition, these enzymes are particularly attractive because they are the smallest lipases known, with a Mr of about 19.5 kDa, and their three-dimensional structures were solved (BSLA) or have been modeled (van Pouderoyen et al. 2001; Eggert et al. 2001). In contrast to other lipases, their active sites are located close to the enzyme's surface, enabling easy access for a variety of different substrates. Consequently, their potential to hydrolyze various substrates with high enantioselectivity has been demonstrated (Funke et al. 2005). Furthermore, a complete saturation library of BSLA is available, allowing us to separately assess the role of each amino acid for a given reaction (Funke et al. 2003, 2005). The results described here, in combination with existing knowledge, make the extracellular lipases A and B of *B. subtilis* very interesting biocatalysts for further investigations, as well as for a variety of biotechnological applications.

References

Ansorge MB, Kula MR (2000) Investigating expression systems for the stable large-scale production of recombinant L-leucine-dehydrogenase from *Bacillus cereus* in *Escherichia coli*. Appl Microbiol Biotechnol 53:668–673

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

Barettino D, Feigenbutz M, Valcarcel R, Stunnenberg HG (1994) Improved method for PCR-mediated site-directed mutagenesis. Nucleic Acids Res 22:541–542

Bodai V, Orovecz O, Szakacs G, Novak L, Poppe L (2003) Kinetic resolution of *trans*-2-acetoxycycloalkan-1-ols by lipase-catalysed enantiomerically selective acylation. Tetrahedron Asym 14:2605–2612

Bradford MM (1976) Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding. Anal Biochem 72:248–254

Buchholz K, Kasche V, Bornscheuer UT (2005) Biocatalysts and enzyme technology. Wiley–VCH, Weinheim

Bullock WO, Fernandez JM, Short JM (1987) Xl1-Blue — a high-efficiency plasmid transforming RecA *Escherichia coli* strain with beta-galactosidase selection. Biotechniques 5:376–378

Caron G, Kazlauskas RJ (1991) An optimized sequential kinetic resolution of *trans*-1,2-cyclohexanediol. J Org Chem 56:7251–7256

Crout DHG, Gaudet VSB, Laumen K, Schneider MP (1986) Enzymatic hydrolysis of (+/−)-*trans*-1,2-diacetoxycycloalkanes — a facile route to optically active cycloalkane-1,2-diols. J Chem Soc Chem Commun 10:808–810

Dartois V, Baulard A, Schanck K, Colson C (1992) Cloning, nucleotide-sequence and expression in *Escherichia coli* of a lipase gene from *Bacillus subtilis* 168. Biochim Biophys Acta 1131:253–260

Dharmsthiti S, Luchai S (1999) Production, purification and characterization of thermophilic lipase from *Bacillus* sp. THL027. FEMS Microbiol Lett 179:241–246

Dröge MJ et al (2003) Binding of phage displayed *Bacillus subtilis* lipase A to a phosphonate suicide inhibitor. J Biotechnol 101:19–28

Dröge MJ, Bos R, Boersma YL, Quax WJ (2005) Comparison and functional characterisation of three homologous intracellular carboxylesterases of *Bacillus subtilis*. J Mol Catal B Enzym 32:261–270

Eggert T, Pencreac'h G, Douchet I, Verger R, Jaeger KE (2000) A novel extracellular esterase from *Bacillus subtilis* and its conversion to a monoacylglycerol hydrolase. Eur J Biochem 267:6459–6469

Eggert T, van Pouderoyen G, Dijkstra BW, Jaeger KE (2001) Lipolytic enzymes LipA and LipB from *Bacillus subtilis* differ in regulation of gene expression, biochemical properties, and three-dimensional structure. FEBS Lett 502:89–92

Eggert T, Brockmeier U, Dröge MJ, Quax WJ, Jaeger KE (2003) Extracellular lipases from *Bacillus subtilis*: regulation of gene expression and enzyme activity by amino acid supply and external pH. FEMS Microbiol Lett 225:319–324

Ewis HE, Abdelal AT, Lu CD (2004) Molecular cloning and characterization of two thermostable carboxyl esterases from *Geobacillus stearothermophilus*. Gene 329:187–195

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

Funke SA et al (2003) Directed evolution of an enantioselective *Bacillus subtilis* lipase. Biocatal Biotrans 21:67–73

Funke SA, Otte N, Eggert T, Bocola M, Jaeger KE, Thiel W (2005) Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. Protein Eng Des Sel 18:509–514

Hochuli E, Dobeli H, Schacher A (1987) New metal chelate adsorbent selective for proteins and peptides containing neighboring histidine residues. J Chromatogr 411:177–184

Jaeger KE, Eggert T (2002) Lipases for biotechnology. Curr Opin Biotechnol 13:390–397

Jaeger KE, Rosenau F (2004) Overexpression and secretion of *Pseudomonas* lipases. In: Ramos J-L (ed) *Pseudomonas* — biosynthesis of macromolecules and molecular metabolism. Kluwer/Plenum, New York, pp 491–508

Jaeger KE, Dijkstra BW, Reetz MT (1999) Bacterial biocatalysts: molecular biology, three-dimensional structures, and biotechnological applications of lipases. Annu Rev Microbiol 53:315–351

Kademi A, Ait-Abdelkader N, Fakhreddine L, Baratti J (2000) Purification and characterization of a thermostable esterase from the moderate thermophile *Bacillus circulans*. Appl Microbiol Biotechnol 54:173–179

Karpushova A, Brummer F, Barth S, Lange S, Schmid RD (2005) Cloning, recombinant expression and biochemical characterisation of novel esterases from *Bacillus* sp associated with the marine sponge *Aplysina aerophoba*. Appl Biochem Biotechnol 67:59–69

Kawai K, Imuta M, Ziffer H (1981) Microbially mediated enantioselective hydrolysis of racemic acetates. Tetrahedron Lett 22:2527–2530

Kim HK, Park SY, Lee JK, Oh TK (1998) Gene cloning and characterization of thermostable lipase from *Bacillus stearothermophilus* L1. Biosci Biotechnol Biochem 62:66–71

Kim HK, Choi HJ, Kim MH, Sohn CB, Oh TK (2002) Expression and characterization of Ca^{2+}-independent lipase from *Bacillus pumilus* B26. Biochim Biophys Acta 1583:205–212

Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer WD (1995) Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. J Biotechnol 39:59–65

Laemmli UK (1970) Cleavage of structural proteins during assembly of head of bacteriophage-T4. Nature 227:680–685

Laumen K, Breitgoff D, Seemayer R, Schneider MP (1989) Enantiomerically pure cyclohexanols and cyclohexane-1,2-diol derivatives — chiral auxiliaries and substitutes for (−)-8-phenylmenthol — a facile enzymatic route. J Chem Soc Chem Commun 3:148–150

Lesuisse E, Schanck K, Colson C (1993) Purification and preliminary characterization of the extracellular lipase of *Bacillus subtilis* 168, an extremely basic pH-tolerant enzyme. Eur J Biochem 216:155–160

Liese A, Lütz S (2004) Nonimmobilized biocatalysts in industrial fine chemical synthesis. In: Ullmann's encyclopedia of industrial chemistry, electronic release, 7th edn. Wiley–VCH, Weinheim

Möller B, Vetter R, Wilke D, Foillois B (1991) Alkaline *Bacillus* lipases, coding DNA sequences therfor and bacilli which produce these lipases. In: WO 91/16422

Nagai H, Tsuchida T, Dobashi K, Isshiki K (2003) Method for producing optically active 1,2-*trans*-cyclohexanediol. In: JP2003125793, Japan

Nthangeni MB, Patterton H, van Tonder A, Vergeer WP, Litthauer D (2001) Over-expression and properties of a purified recombinant *Bacillus licheniformis* lipase: a comparative report on *Bacillus* lipases. Enzyme Microb Technol 28:705–712

Porath J, Carlsson J, Olsson I, Belfrage G (1975) Metal chelate affinity chromatography, a new approach to protein fractionation. Nature 258:598–599

Rahman R, Chin JH, Salleh AB, Basri M (2003) Cloning and expression of a novel lipase gene from *Bacillus sphaericus* 205y. Mol Genet Genomics 269:252–260

Rasool S et al (2005) Molecular cloning of enantio selective ester hydrolase from *Bacillus pumilus* DBRL-191. FEMS Microbiol Lett 249:113–120

Ruiz C, Blanco A, Pastor FI, Diaz P (2002) Analysis of *Bacillus megaterium* lipolytic system and cloning of LipA, a novel subfamily I.4 bacterial lipase. FEMS Microbiol Lett 217:263–267

Ruiz C, Javier Pastor FI, Diaz P (2003) Isolation and characterization of *Bacillus* sp. BP-6 LipA, a ubiquitous lipase among mesophilic *Bacillus* species. Lett Appl Microbiol 37:354–359

Schmidt-Dannert C, Sztajer H, Stocklein W, Menge U, Schmid RD (1994) Screening, purification and properties of a thermophilic lipase from *Bacillus thermocatenulatus*. Biochim Biophys Acta 1214:43–53

Schmidt-Dannert C, Rua ML, Atomi H, Schmid RD (1996) Thermoalkalophilic lipase of *Bacillus thermocatenulatus*. I. Molecular cloning, nucleotide sequence, purification and some properties. Biochim Biophys Acta 1301:105–114

Schmidt-Dannert C, Rua ML, Schmid RD (1997) Two novel lipases from thermophile *Bacillus thermocatenulatus*: screening, purification, cloning, overexpression, and properties. Methods Enzymol 284:194–220

Seemayer R, Schneider MP (1991) Enzymatic preparation of optically pure *trans*-1,2-cycloalkanediols. J Chem Soc Chem Commun 1:49–50

Sekhon A, Dahiya N, Tiwari RP, Hoondal GS (2005) Properties of a thermostable extracellular lipase from *Bacillus megaterium* AKG-1. J Basic Microbiol 45:147–154

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

Tyndall JDA, Sinchaikul S, Fothergill-Gilmore LA, Taylor P, Walkinshaw MD (2002) Crystal structure of a thermostable lipase from *Bacillus stearothermophilus* P1. J Mol Biol 323:859–869

van Pouderoyen G, Eggert T, Jaeger KE, Dijkstra BW (2001) The crystal structure of *Bacillus subtilis* lipase: a minimal alpha/beta hydrolase fold enzyme. J Mol Biol 309:215–226

Wahler D, Boujard O, Lefevre F, Reymond JL (2004) Adrenaline profiling of lipases and esterases with 1,2-diol and carbohydrate acetates. Tetrahedron 60:703–710

Publications of **Chapter 6 "Novel biocatalysts by evolution and design"**:

Learning from Directed Evolution: Further Lessons from Theoretical Investigations into Cooperative Mutations in Lipase Enantioselectivity

Manfred T. Reetz,*[a] Michael Puls,[b] José Daniel Carballeira,[a] Andreas Vogel,[a] Karl-Erich Jaeger,*[b] Thorsten Eggert,[b] Walter Thiel,*[a] Marco Bocola,[a] and Nikolaj Otte[a]

An earlier experimental study, which involved the directed evolution of enantioselective lipase variants from Pseudomonas aeruginosa as catalysts in the hydrolytic kinetic resolution of 2-methyl-decanoic acid p-nitrophenyl ester, provided a mutant with six mutations. Consequently, the selectivity factor was found to increase from E = 1.1 for the wild-type to E = 51 for the best mutant. Only one of the amino acid exchanges in this mutant was found to occur next to the binding pocket, the other mutations being remote. Our previous theoretical analysis with molec-ular-dynamics simulations helped to unveil the source of en-hanced enantioselectivity: a relay mechanism that involves two of the six mutations was shown to induce strong cooperativity. In this investigation, single, double, and triple mutants were con-structed and tested as enantioselective catalysts. This study sup-ports our original postulate regarding the relay mechanism, offers further mechanistic insight into the role of individual muta-tions, and provides mutants that display even higher enantiose-lectivity (E of up to 64).

Introduction

We previously demonstrated that the method of directed evo-lution[1] can be applied successfully in the quest to create enantioselective enzymes for use in synthetic organic chemis-try.[2] The system that was studied most systematically con-cerns the hydrolytic kinetic resolution of rac-2-methyl-decanoic acid p-nitrophenyl ester (MDA ester) catalyzed by mutants of the lipase from Pseudomonas aeruginosa (Scheme 1).[3] The wild-type (WT) lipase has a selectivity factor of only E = 1.1, in slight favor of (S)-2.

The early papers in this ongoing project focused on the development of strategies for efficient probing of protein-se-quence space; these efforts are summarized in Scheme 2. The first approach utilized four consecutive cycles of error-prone polymerase chain reaction (epPCR) at low mutation rate (one amino acid exchange per enzyme), which resulted in a variant (I) with a moderate increase in S selectivity (E = 11).[3a] Never-theless, the proof-of-principle of this new approach to asym-metric catalysis had been achieved. Subsequently, other muta-genesis methods were tested. Saturation mutagenesis at se-lected amino acid positions ("hot spots") that were identified in four cycles of epPCR, doubled the degree of enantioselectiv-ity (variant II, E = 20).[3b] However, DNA shuffling of the corre-sponding genes failed to provide any significant enhancement of enantioselectivity.[3c] Therefore, several other strategies were probed, including the generation of a focused library that ori-ginated from randomization of amino acids 160–163 next to the binding pocket. This cassette mutagenesis resulted in the identification of an improved quadruple variant (VII) character-ized by the mutations E160A/S161D/L162G/N163F (E = 30).[3c]

The optimal strategy[3c] was found to involve the use of epPCR at considerably higher mutation rates combined with

[a] Prof. Dr. M. T. Reetz, Dr. J. D. Carballeira, Dr. A. Vogel, Prof. Dr. W. Thiel, Dr. M. Bocola, Dr. N. Otte
Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1, 45470 Mülheim/Ruhr (Germany)
Fax: (+49) 208-306-2985
E-mail: reetz@mpi-muelheim.mpg.de
          thiel@mpi-muelheim.mpg.de

[b] Dipl.-Nat. M. Puls, Prof. Dr. K.-E. Jaeger, Dr. T. Eggert
Institut für Molekulare Enzymtechnologie
Heinrich-Heine-Universität Düsseldorf
Forschungszentrum Jülich, 52426 Jülich (Germany)
Fax: (+49) 2461-612490
E-mail: karl-erich.jaeger@fz-juelich.de

R = n-C$_8$H$_{17}$
R' = p-NO$_2$C$_6$H$_4$

Scheme 1. Kinetic resolution of rac-2-methyldecanoic acid p-nitrophenyl ester (rac-1) catalyzed by the lipase from P. aeruginosa.

Scheme 2. Schematic representation of directed evolution of enantioselective enzymes (lipase variants) that catalyze the hydrolytic kinetic resolution of ester **1**.

an adapted form of Stemmer's combinatorial multiple-cassette mutagenesis (CMCM).[4] This was accomplished by first applying epPCR at a mutation rate averaging three amino acid exchange events per enzyme, which provided two improved triple variants, IV and V, that harbored mutations S53P/C180T/G272A and D20N/S161P/T234S, respectively. Since positions 155 and 162 had been shown to be hot spots in other experiments,[3b] DNA shuffling was performed with the genes that encoded these two mutants and an appropriate oligocassette; this induced simultaneous saturation mutagenesis at positions 155 and 162 (modified CMCM).[3c] This strategy provided the most selective mutant (X) to date, which displays a selectivity factor of $E = 51$. Scheme 2 summarizes the essential results of all of the approaches taken thus far to probe protein-sequence space. A total of about 40 000 clones were screened.

The most selective enzyme, variant X, is characterized by six mutations, namely D20N, S53P, S155M, L162G, T180I, and T234S. Unfortunately, it was not possible to obtain crystals suitable for X-ray structure analysis of this mutant. We therefore used the crystal structure of the WT lipase[5] to locate the spatial positions of the six mutations (Figure 1). To our initial surprise, only one mutation (L162G) was located directly next to the binding pocket, the others were remote.[3] Whereas remote (distal) mutations had been known to influence such enzyme properties as thermal stability or activity,[6] our results constituted the first example of stereoselectivity. This unusual observation called for a thorough theoretical analysis.[7]
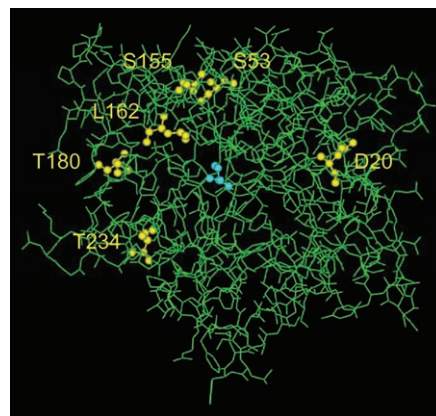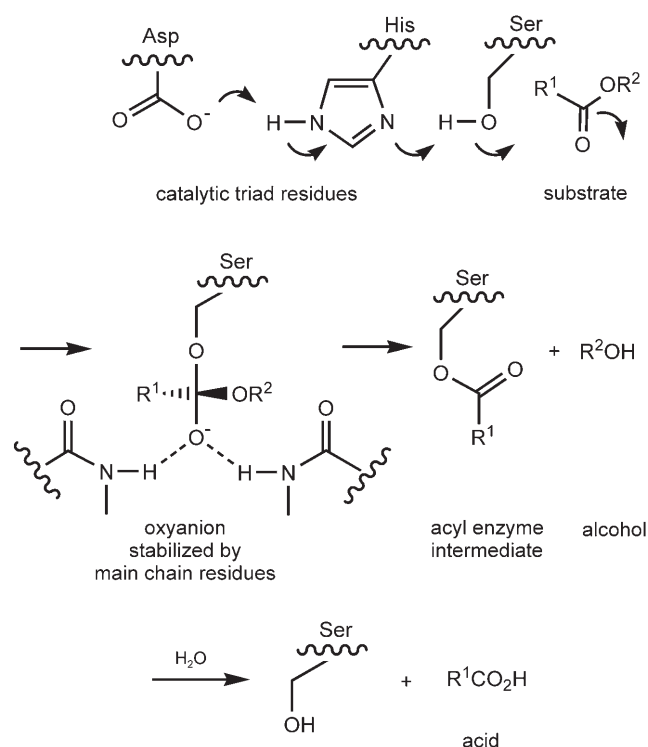

Figure 1. X-ray structure of the WT lipase from *P. aeruginosa* (green)[5] showing the active-site serine at position 82 (blue) and the six amino acids that were substituted for the generation of mutant X,[3c] namely D20, S53, S155, L162, T180, and T234.

The mechanism of lipase-catalyzed ester hydrolysis is known to involve a catalytic triad comprising aspartate, histidine, and serine. The latter amino acid undergoes rate-determining nucleophilic addition to the carbonyl function to form the so-called oxyanion, which is stabilized by H-bonds that originate from two amino acids in the protein environment (Scheme 3).[8]

Figure 2. A) Oxyanion derived from *rac*-**1** in the WT enzyme; B) oxyanion derived from the favored (*S*)-**1** (green) and the disfavored (*R*)-**1** (purple) in the double mutant S53P/L162G (mutations that also occur in variant X); dotted green lines indicate H-bonds; asterisks indicate sterrogenic centres.

Scheme 3. General mechanism of lipase-catalyzed hydrolysis of esters.[8]

In order to illuminate the source of the enantioselectivity of the best previous variant (X) a theoretical study was initiated, the initial results of which were published in 2004.[7] Following quantum-mechanical (QM) determination of the force fields associated with the oxyanion, comprehensive molecular-mechanical (MM) calculations were performed by using force-field molecular dynamics (MD). Two results of this study are of particular significance. Firstly, no appreciable difference in energy was found upon binding (*S*)- or (*R*)-**1** in the form of its respective Michaelis–Menten complex. Secondly, the main source of enantioselectivity was traced to the cooperative influence of two of the six mutations, namely L162G and S53P. The close mutation L162G enlarges the binding pocket to accommodate bulky α-chiral esters. At the same time, the remote mutation S53P triggers a relay mechanism that results in a strong cooperative effect. This unusual phenomenon can be described as follows. In the WT lipase, amino acids S53, H83, and S161 form a hydrogen bond network (Figure 2 A). Mutation S53P in variant X or other variants, such as the hypothetical double mutant S53P/L162G proposed in our earlier study,[7] causes this network to break up and allows the side chain of H83 to reposition itself. In the case of (*S*)-**1** this leads to an additional H-bond with the oxyanion (Figure 2B). Such stabilizing interaction is not possible in the case of the oxyanion that originates from enantiomer (*R*)-**1** due to steric repulsion of the methyl group at the stereogenic center. This model predicts that variant X should be more active than the WT enzyme, which is indeed the case.[9] On going from the WT lipase to variant X, the k_{cat}/K_m value increases by a factor of 240.[3e] The analysis also implies that mutations L162G and S53P are decisive for
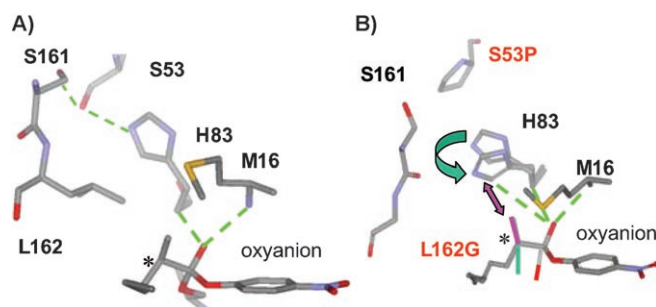
the stereoselectivity of variant X, while the other (remote) substitutions, D20N, S155M, T180I, and T234S are less relevant in this regard. Consequently, the hypothetical double mutant S53P/L162G was predicted to be an excellent catalyst.[7]

Earlier experimental results had shown position 155 to be an important hot spot.[3] For example, upon subjecting the mutant gene that encides an earlier variant (variant I; S149G/S155L/V47G and F259L[3a] obtained from four cycles of low-error epPCR) to saturation mutagenesis at position 155, a new variant (variant II; S149G/S155F/V47G/F259L) was identified in which leucine is exchanged for phenylalanine. This causes the selectivity factor to double from $E = 11$ to $E = 20$.[3b] Previous MD calculations indicated that in the case of (*R*)-**1** this substitution enhances *S* selectivity due to steric strain introduced by the bulky side chain of phenylalanine—an effect that slows down the reaction with respect to the transformation of the *S* enantiomer.[7] However, in the most selective mutant (X) position 155 harbors methionine and not phenylalanine.

The purpose of this study was to deepen our understanding of the catalyst enantiorecognition by designing further experiments as suggested by earlier theoretical analysis.[7] It was of particular interest to prepare and test six single mutants that correspond to the six mutations in variant X, another single mutant characterized by S155F, a double mutant, S53P/L162G (Figure 2B), and finally two triple mutants, S53P/S155M/L162G and S53P/S155F/L162G. The experimental results regarding enantioselectivity in the model reaction **1**→**2** were expected to provide clues on potential cooperative effects and/or the role of individual mutations that act alone. Moreover, since our original model[7] had shown that the histidine at position 83 forms an additional H-bond with the oxyanion in the reaction with the *S* enantiomer of **1** (Figure 2), several mutants were designed and prepared in which H83 is replaced by other amino acids.

Results and Discussion

Ten new variants M1–M10, designed by deconvoluting the best previous variant (X), were prepared by standard molecular biological methods. Since the WT and most variants of the lipase from *P. aeruginosa* are somewhat unstable in pure form, it is difficult to isolate the enzymes prior to testing their cata-

lytic profiles.[9, 10] Therefore, we proceeded as in previous studies by using the supernatants in which the WT enzyme and variants are perfectly stable.[3] Moreover, in order to make experimentally sound comparisons, the protocol of the hydrolytic kinetic resolution was standardized. The *E* values were then calculated by using the formula published by Sih and co-workers.[11] It should be noted that under these standardized conditions variant X has an $E = 50$ which is, within the experimental error ($\pm 15\%$), essentially identical to the value ($E = 51$) originally reported.[3c, 7] The catalytic profiles are summarized in Table 1. Precise kinetic studies were not performed because the enzyme could not be obtained in pure form nor were expression levels measured.

Table 1. Enantioselectivity in the hydrolytic kinetic resolution of *rac*-**1** catalyzed by single, double, and triple mutants of the lipase from *P. aeruginosa*.

Mutant	Mutation(s)	Conversion [%]	*ee* [%]	*E* value
M1	D20N	19	11	1.3
M2	S53P	17	48	3.1
M3	S155M	20	19	1.5
M4	S155F	15	33	2.1
M5	L162G	42	89	33
M6	T180I	24	15	1.4
M7	T234S	21	11	1.3
M8	S53P/L162G	18	96	64
M9	S53P/S155M/L162G	40	92	42
M10	S53P/S155F/L162G	34	95	58
X	D20N/S53P/S155M/L162G/T180I/T234S	20	95	50
WT	–	44	6.8	1.2

Due to the crucial role of H83, which was uncovered by the previous theoretical study,[7] several variants were prepared in which this histidine is replaced by other amino acids by using site-specific mutagenesis. The catalytic profiles of these mutants are summarized in Table 2.

The experimental results from this study are striking in several respects. Firstly, the double mutant M8 (S53P/L162G), which was predicted to have a high selectivity factor, is even more enantioselective ($E = 64$) than the previous best variant (X, $E = 50$), with six mutations including S53P/L162G. As delineated above and described in detail in our earlier theoretical study,[7] it is the combination of these two mutations in variant X that is crucial for the observed high enantioselectivity. This prediction is further corroborated by the observation that all of the single mutants that correspond to the six mutations in variant X, except for variant M5 characterized by L162G, show essentially no enhancement of stereoselectivity relative to the WT lipase. The selectivity factor observed with variant M5 amounts to $E = 33$, which is significant (Table 1). As pointed out previously, position 162 is directly next to the binding pocket in the region where branching (α-methyl group) of the acid moiety of the bound ester **1** occurs. The binding pocket has thus been expanded to accommodate the methyl group at the stereogenic center. This applies both to the *S* and *R* esters

Table 2. Enantioselectivity in the hydrolytic kinetic resolution of *rac*-**1** catalyzed by variants of the lipase from *P. aeruginosa*, which were prepared by site-directed mutagenesis at position 83.

Mutant	Mutation(s)	Conversion [%]	*ee* [%]	*E* value
M11	H83A	22	13	1.4
M12	H83L	11	28	1.8
M13	H83F	14	6	1.1
M14	H83A/L162G	42	93	54
M15	H83L/L162G	20	94	41
M16	H83F/L162G	14	74	7.5
M17	S53P/H83A/L162G	44	93	60
M18	S53P/H83L/L162G	33	94	48
M19	S53P/H83F/L162G	26	63	5.4
M20	S53P/H83F/S155F/L162G	17	70	6.4
M21	S53P/H83L/S155F/L162G	28	95	53
M22	S53P/H83A/S155F/L162G	21	92	31
M23	S53P/H83F/S155M/L162G	38	90	33
M24	S53P/H83A/S155M/L162G	35	91	32
M25	S53P/H83L/S155M/L162G	30	90	26

of **1**. In contrast to the direct effect of the single mutation L162G, the other mutation in the double mutant M8, namely S53P, has no significant influence on its own, as demonstrated by the catalytic profile of variant M2 ($E = 3.1$). Therefore, the cooperative effect of the two mutations, S53P/L162G, that characterize variant M8 ($E = 64$) and occur as two of the six mutations in variant X ($E = 50$), is clearly demonstrated.

As summarized in Table 1, the two triple mutants M9 (S53P/S155M/L162G) and M10 (S52P/S155F/L162G) also lead to pronounced degrees of *S* selectivity relative to the WT enzyme—although the *E*-factors are somewhat lower (42 and 58, respectively) than that of the best new variant, M8. The side chain of the amino acid at position 155 has some influence on enantioselectivity: the positive effect in the triple mutants decreases in the series Ser > Phe > Met. The crystal structure of the WT enzyme[5] shows that S155 is located under the lid above the active site; this stabilizes the lid domain by a direct H-bond between the serine side chain and the lid backbone. Mutations in this location can influence the steric strain on the substrate in the active site by a domino effect, and might also alter the lid position in a way that cannot be easily predicted.

Finally, the role of the histidine at position 83 was examined, although none of the random mutagenesis experiments (epPCR or DNA shuffling) had shown this position to be a hot spot (Table 2). Our model calls for additional stabilizing H-bonds between this histidine and the oxyanion in the reaction with *S*-configured substrate **1**, but not with the *R* enantiomer due to steric inhibition (Figure 2). Thus, one might expect poor enantioselectivity upon introducing mutations such as H83A, H83L, or H83F in the previous double or triple mutants. However, as shown in Table 2, this is not observed in all cases. In the case of the original single mutant M5 (L162G), which is characterized by a selectivity factor of $E = 33$, substituting histidine by alanine or leucine actually *increases* enantioselectivity (variant M14: $E = 54$; variant M15: $E = 41$). This result was unexpected, but it can be explained on the basis of our original model and a plausible conjecture. Replacing H83 by smaller

residues could lead to the stabilization of the oxyanion because incorporation of water molecules at position 83 is then possible, and results in the formation of H-bonds with the oxyanion derived from (*S*)- but not (*R*)-**1**. We have checked this conjecture by force-field calculations using the same methodology and setup as in our previous study.[7] H83 was manually replaced by alanine in an available theoretical structure of the tetrahedral intermediate of mutant M5,[7] and local geometry optimizations were carried out with water molecules manually added into the empty space formerly occupied by the histidine side chain. In the case of the *S* enantiomer, a local minimum was found with two bound water molecules that provide additional stabilization through H-bonds (Figure 3). No such mini-
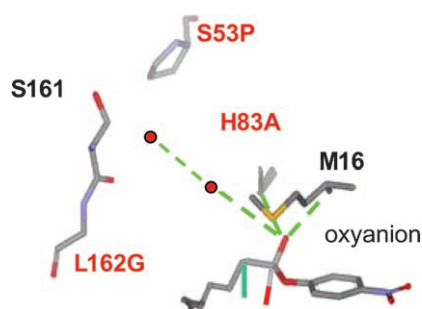
Figure 3. A model of the possible role of two bound water molecules (red dots) in mutant M14, which shows a selectivity factor of *E* = 54 in the model reaction (Table 2); dotted green lines indicate H-bonds.

mum could be located for the *R* enantiomer because of the steric constraints in the oxyanion hole (analogous to Figure 2). Both systems were then subjected to MD simulations, which involved an initial re-equilibration and a subsequent production run of 1 ns by using the same conventions as before.[7] During these MD simulations, the two additional water molecules in the active site of the *S* enantiomer turned out to be rather mobile, but at least one of them spent a substantial part of the time in the oxyanion hole in typical H-bond geometries. By contrast, water was never observed to enter the oxyanion hole of the *R* enantiomer during the 1 ns simulation. Hence, both the geometry optimization studies and MD simulations provide qualitative support for the notion that water selectively stabilizes the *S* enantiomer of mutant M14 (H83A/L162G) in the relevant part of the potential energy surface around the tetrahedral intermediate.

It should be noted that the introduction of phenylalanine (H83F) in variant M16 lowers enantioselectivity drastically (*E* = 7.5). This is likewise in line with our model, since the benzyl side chain of phenylalanine has no H-bonding capability, but resembles the side chain of histidine sterically, which leaves no space for water to be incorporated. Therefore, additional stabilization of the oxyanion by H-bonds that originate from water is not possible in this case. Our current model is thus able to explain why position 83, which was never a site of mutation in former directed-evolution experiments,[3] is crucial for enantiorecognition in most of the new cases studied herein.

Similar lines of thought pertain to mutational changes at position 83 in the case of the double and triple mutants (Table 2). For example, the triple mutant M17, which is characterized by S53P/H83A/L162G, is an excellent catalyst (*E* = 60). Here, again, an enzyme variant was created that is even more enantioselective than the best previous variant (X). The interpretation regarding the quadruple mutants M20, M21, M22, M23, M24, and M25 (Table 2) is more difficult since additional subtle effects might be operating.

In addition to clearing up the crucial molecular phenomena involved in the enhancement of enantioselectivity, this study also raises important questions regarding the efficiency of the various strategies used in probing protein-sequence space (Scheme 2). Due to the Darwinian nature of directed evolution,[1] application of this type of protein engineering can always be expected to provide positive results, irrespective of the particular strategy used. Following the ground-breaking work of the 1990s,[1,3] the real question currently concerns the challenge regarding efficiency. As noted earlier (Scheme 2), the evolutionary pathway from the WT *P. aeruginosa* lipase to variant X involves several steps beginning with epPCR at a relatively high mutagenesis rate with simultaneous introduction of three mutations. This was followed by DNA shuffling of two of the obtained mutants while simultaneously saturating at two positions shown earlier to be hot spots. The latter had been identified on the basis of multiple rounds of low-error epPCR (Scheme 2). The overall process, although successful, entails a fair amount of molecular biology and screening, and also picks up mutations that might be superfluous.

It is clear that mutations D20N, T180I, and T234S do not influence enantioselectivity in the reaction with *rac*-**1**, as shown experimentally by this study and supported by theory, which raises the question as to why they appeared in the original process of directed evolution. One explanation for superfluous mutations is simply a statistical one. Indeed, this phenomenon is not unusual in directed evolution,[1] although systematic studies that uncover elements of "inefficiency" when probing protein-sequence space are rare.[1,12] Of course, other effects along a given evolutionary pathway could occur. One of several lessons learned in this study is to make use of theory at all stages of a directed evolution project, which is a viable alternative to the strict use of purely rational design.[13] Indeed, our new approach to the directed evolution of functional proteins, namely iterative saturation mutagenesis, is based on the combination of amino acid randomization and rational design—iterative CASTing[14] is one embodiment of this general method. It will be of interest to test whether iterative CASTing in the present enzyme system is indeed more efficient than our previous approaches (Scheme 2).

Conclusions

We have shown in this study that directed evolution[1,2,3] coupled with appropriate theoretical analyses can lead to important lessons in enzyme mechanism(s). In the previous theoretical investigation of the experimental results obtained from directed evolution of enantioselective lipase variants from *P. aer-*

uginosa, a model was proposed that explains the observed remote effects.[7] This also provided predictions regarding the role of individual mutations, which called for renewed experimental work. As a result, new enantioselective mutants were discovered that support the model. On the practical side, they are even more selective than the best mutant previously evolved. Finally, theory suggested control experiments that involve additional mutational changes at a site not considered in earlier studies. This resulted in novel mutants that also display high enantioselectivity. Thus, the close intertwinement of experiment and theory is rewarding.

Experimental Section

Computational methods—model building and MD simulations: The same methods were applied as in our previous paper.[7] Briefly, the wild-type and mutant enzyme structures were built from the X-ray crystal structure of *P. aeruginosa* lipase complexed with R_C-($R_P S_P$)-1,2-dioctylcarbamoyl-glycero-3-*O*-*p*-nitrophenyl octylphosphonate[5] from the RCSB Protein Data Bank (ID: 1EX9). The tetrahedral inhibitor covalently bound to S82 was excised and replaced by the tetrahedral intermediate of the MDA ester. The system was solvated with a pre-equilibrated water droplet (20 Å) around the active site (substrate atom C9). Protein residues more than 20 Å away from the substrate stereocenter (substrate atom C1) in the active site were fixed. The 1 ns simulations were performed for each enantiomer with the Charmm 29b2 suite of programs[16] by using the CHARMM 22/27[15] parameters derived for the charged tetrahedral intermediate, as described before.[7]

Molecular biology methods: Mutants were constructed with PCR techniques such as QuikChange-Mutagenesis[17] or Megaprimer-method[18] by using vector-specific primers or site-specific mutagenesis primers (Table 3); the latter were used in combination to obtain double, triple, or quadruple mutants. The mutant genes were cloned into the vector pUCPL6A (derived from pUCPKS[19]) which is able to replicate in both *E. coli* and *P. aeruginosa*, transformed into *E. coli*, and after confirmation of the DNA-sequences, transformed into *P. aeruginosa* strain PABST7.1[20]—a lipase-deficient mutant of strain PAO1—for expression of the lipase variants in the homologous host.

Bacterial strains and growth conditions: Plasmids were transformed into *P. aeruginosa* PABST7.1 competent cells as previously described.[20] Single colonies were picked from Luria–Bertani (LB) agar plates (supplemented with the appropriate amount of antibiotics: 100 µg mL^{-1} carbenicillin and 50 µg mL^{-1} tetracycline) and used to inoculate 2× liquid LB media (15 mL) in an Erlenmeyer flask (100 mL). This preculture was grown in an orbital shaker at 30 °C and 300 rpm, overnight, after which 1 mL was used to inoculate a new Erlenmeyer flask with 2× LB media (15 mL). After 5 h at 30 °C and 300 rpm, lipase expression was induced by the addition of isopropyl-β-D-thiogalactopyranoside (IPTG; final concentration 0.1 mM) and the culture was incubated for a further 5 h. Lipase-containing supernatants were recovered by centrifugation at 8000 rpm for 45 min, and stored overnight at 4 °C.

Reaction conditions: Supernatant (50 µL), Tris-HCl buffer pH 7.5 (300 µL of a 100 mM solution) and substrate solution (25 µL; 10 mg mL^{-1} in acetonitrile) were used as reaction mixture. Reactions were allowed to reach a conversion of between 15–45 %. Extraction of products was initiated by adding CH$_2$Cl$_2$ (300 µL) to the reaction mixture, which was then mixed (vortex) for 30 s. HCl (30 µL of 10 % solution) was then added to protonate the remaining acid, and the mixture was stirred for an additional 30 s. The CH$_2$Cl$_2$ extract was then transferred to a 96-well glass plate for analysis by GC.

Analytical conditions: The hydrolysis of *p*-nitrophenyl 2-methyl decanoate (**1**) was monitored with GC analysis by using an IVADEX-1 (25 m, 0.15 µm, 0.25 mm) chiral column (IVA Analysentechnik, Meerbusch, Germany); carrier: nitrogen; flow: 1.4 mL min^{-1}; pressure 120 kPa; average velocity: 40 cm s^{-1}; injector and detector temperatures: 250 °C. The initial column temperature was at 140 °C for 21 min, ramp at 10 °C min^{-1} to 200 °C for 30 min; total time 57 min. Retention times were (*S*)-**2**: 18.6 min; (*R*)-**2**: 20.3 min; *p*-nitrophenol: 30 min; *rac*-**1**: 52.3 min.

Table 3. Primers used in this study. The respective mutant codons are indicated in bold type.

Name	Mutation	Primer sequence (5′→3′)
S53Pfor	S53P	ACGCAGTTGGACACC**CCG**GAAGTCCGCGGCGAG
S53Pev		TGCGTCAACCTGTGG**GGC**CTTCAGGCGCCGCTC
S155Ffor	S155F	ACCGGTACGCAGAAT**TTT**CTGGGCTCGCTGGAG
S155Frev		TGGCCATGCGTCTTA**AAA**GACCCGAGCGACCTC
S155Mfor	S155M	ACCGGTACGCAGAAT**ATG**CTGGGCTCGCTGGAG
S155Mrev		TGGCCATGCGTCTTA**TAC**GACCCGAGCGACCTC
L162Gfor	L162G	GGCTCGCTGGAGTCG**GGG**AACAGCGAGGGTGCC
L162Grev		CCGAGCGACCTCAGC**CCC**TTGTCGCTCCCACGG
H83Afor	H83A	CTGATCGGCCACAGC**CTG**GGCGGGCCGACCATC
H83Arev		GATGGTCGGCCCGCC**CAG**GCTGTGGCCGATCAG
H83Ffor	H83F	CTGATCGGCCACAGC**TTC**GGCGGGCCGACCATC
H83Frev		GATGGTCGGCCCGCC**GAA**GCTGTGGCCGATCAG
H83Lfor	H83L	CTGATCGGCCACAGC**GCC**GGCGGGCCGACCATC
H83Lrev		GATGGTCGGCCCGCC**GGC**GCTGTGGCCGATCAG
ppetDWN	–	GCTAACCAGTAAGGCAACCCCGCCAGCCTAGCC
plipAUP2	–	TAAAACGACGGCCAGTGAGCGCGCAATTAACC

[1] Reviews: a) *Methods in Molecular Biology, Vol. 231: Directed Evolution Library Creation: Methods and Protocols* (Eds.: F. H. Arnold, G. Georgiou), Humana, Totowa, **2003**; b) *Methods in Molecular Biology, Vol. 230: Directed Enzyme Evolution: Screening and Selection Methods* (Eds.: F. H. Arnold, G. Georgiou), Humana, Totowa, **2003**; c) *Directed Molecular Evolution of Proteins* (Eds.: S. Brakmann, K. Johnsson), Wiley-VCH, Weinheim, **2002**; d) *Enzyme Functionality—Design, Engineering, and Screening* (Ed.: A. Svendsen), Marcel Dekker, New York, **2004**; e) K. A. Powell, S. W. Ramer, S. B. del Cardayré, W. P. C. Stemmer, M. B. Tobin, P. F. Longchamp, G. W. Huisman, *Angew. Chem.* **2001**, *113*, 4068–4080; *Angew. Chem. Int. Ed.* **2001**, *40*, 3948–3959; f) M. T. Reetz, K.-E. Jaeger, *Top. Curr. Chem.* **1999**, *200*, 31–57; g) S. V. Taylor, P. Kast, D. Hilvert, *Angew. Chem.* **2001**, *113*, 3408–3436; *Angew. Chem. Int. Ed.* **2001**, *40*, 3310–3335; h) N. J. Turner, *Trends Biotechnol.* **2003**, *21*, 474–478; i) S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291–297.

[2] Comprehensive review of directed evolution of enantioselective enzymes: M. T. Reetz in *Advances in Catalysis, Vol. 49* (Eds.: B. C. Gates, H. Knözinger), Elsevier, San Diego, **2006**, pp. 1–69.

[3] a) M. T. Reetz, A. Zonta, K. Schimossek, K. Liebeton, K.-E. Jaeger, *Angew. Chem.* **1997**, *109*, 2961–2963; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2830–2832; b) K. Liebeton, A. Zonta, K. Schimossek, M. Nardini, D. Lang, B. W. Dijkstra, M. T. Reetz, K.-E. Jaeger, *Chem. Biol.* **2000**, *7*, 709–718; c) M. T. Reetz, S. Wilensek, D. Zha, K.-E. Jaeger, *Angew. Chem.* **2001**, *113*, 3701–3703; *Angew. Chem. Int. Ed.* **2001**, *40*, 3589–3591; d) D. Zha, S. Wilensek, M. Hermes, K.-E. Jaeger, M. T. Reetz, *Chem. Commun.* **2001**, 2664–2665; e) M. T. Reetz, *Tetrahedron* **2002**, *58*, 6595–6602; f) M. T. Reetz, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5716–5722.

[4] A. Crameri, W. P. C. Stemmer, *BioTechniques* **1995**, *18*, 194–196.

[5] M. Nardini, D. A. Lang, K. Liebeton, K.-E. Jaeger, B. W. Dijkstra, *J. Biol. Chem.* **2000**, *275*, 31219–31225.

[6] Examples of remote effects: a) S. Oue, A. Okamoto, T. Yano, H. Kagamiyama, *J. Biol. Chem.* **1999**, *274*, 2344–2349; b) H. Zhao, F. H. Arnold, *Protein Eng.* **1999**, *12*, 47–53; c) A. Iffland, P. Tafelmeyer, C. Saudan, K. Johnsson, *Biochemistry* **2000**, *39*, 10790–10798; d) M. Kumar, K. K. Kannan, M. V. Hosur, N. S. Bhavesh, A. Chatterjee, R. Mittal, R. V. Hosur, *Biochem. Biophys. Res. Commun.* **2002**, *294*, 395–401; e) P. K. Agarwal, S. R. Billeter, P. T. R. Rajagopalan, S. J. Benkovic, S. Hammes-Schiffer, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 2794–2799; f) F. H. Arnold, P. L. Wintrode, K. Miyazaki, A. Gershenson, *Trends Biochem. Sci.* **2001**, *26*, 100–106; g) G. P. Horsman, A. M. F. Liu, E. Henke, U. T. Bornscheuer, R. J. Kazlauskas, *Chem. Eur. J.* **2003**, *9*, 1933–1939.

[7] M. Bocola, N. Otte, K.-E. Jaeger, M. T. Reetz, W. Thiel, *ChemBioChem* **2004**, *5*, 214–223.

[8] a) R. D. Schmid, R. Verger, *Angew. Chem.* **1998**, *110*, 1694–1720; *Angew. Chem. Int. Ed.* **1998**, *37*, 1608–1633; b) *Enzyme Catalysis in Organic Synthesis: A Comprehensive Handbook, Vols. I–III* (Eds.: K. Drauz, H. Waldmann), 2nd ed., Wiley-VCH, Weinheim, **2002**; c) *Biotransformations in Organic Chemistry* (Ed.: K. Faber), 4th ed., Springer, Berlin, **2000**; d) S. H. Krishna, N. G. Karanth, *Catal. Rev. Sci Eng.* **2002**, *44*, 499–591; e) P. Berglund, K. Hult in *Stereoselective Biocatalysis* (Eds.: R. N Patel, N. Ramesh), Marcel Dekker, New York, **2000**, pp. 633–657; f) M. T. Reetz, *Curr. Opin. Chem. Biol.* **2002**, *6*, 145–150; g) QM-Studies of serine hydrolases: C.-H. Hu, T. Brinck, K. Hult, *Int. J. Quantum Chem.* **1997**, *89*–103; h) J. Pleiss in *Enzyme Functionality* (Ed.: A. Svendsen), Marcel Dekker, New York, **2003**, pp. 59–77; i) Mechanism of serine proteases: L. Hedstrom, *Chem. Rev.* **2002**, *102*, 4501–4523.

[9] S. Wilensek, PhD Dissertation, Ruhr-Universität Bochum (Germany), **2001**.

[10] M. T. Reetz, M. Bocola, J. D. Carballeira, D. Zha, A. Vogel, *Angew. Chem.* **2005**, *117*, 4264–4268; *Angew. Chem. Int. Ed.* **2005**, *44*, 4192–4196.

[11] All *E* values were derived by applying the formula of Sih et al.: C.-S. Chen, Y. Fujimoto, G. Girdaukas, C. J. Sih, *J. Am. Chem. Soc.* **1982**, *104*, 7294–7299.

[12] a) M. R. Parikh, I. Matsumura, *J. Mol. Biol.* **2005**, *352*, 621–628; b) W. M. Coco, W. E. Levinson, M. J. Crist, H. J. Hektor, A. Darzins, P. T. Pienkos, C. H. Squires, D. J. Monticello, *Nat. Biotechnol.* **2001**, *19*, 354–359; c) M. Kikuchi, K. Ohnishi, S. Harayama, *Gene* **2000**, *243*, 133–137; d) T. Eggert, M. T. Reetz, K.-E. Jaeger in *Enzyme Functionality—Design, Engineering, and Screening* (Ed.: A. Svendsen), Marcel Dekker, New York, **2004**, pp. 375–390; e) M. Kikuchi, K. Ohnishi, S. Harayama, *Gene* **2000**, *243*, 133–137; f) D. A. Drummond, B. L. Iverson, G. Georgiou, F. H. Arnold, *J. Mol. Biol.* **2005**, *350*, 806–816; g) M. Zaccolo, E. Gherardi, *J. Mol. Biol.* **1999**, *285*, 775–783; h) T. S. Wong, D. Roccatano, M. Zacharias, U. Schwaneberg, *J. Mol. Biol.* **2006**, *355*, 858–871; i) F. H. Arnold, *Nat. Biotechnol.* **2006**, *24*, 328–330.

[13] Recent example of rational design in improving the enantioselectivity of a lipase: a) T. Ema, T. Fujii, M. Ozaki, T. Korenaga, T. Sakai, *Chem. Commun.* **2005**, 4650–4651. For reviews of rational design of enzymes based on site-specific mutagenesis, see: b) F. Cedrone, A. Ménez, E. Quéméneur, *Curr. Opin. Struct. Biol.* **2000**, *10*, 405–410; c) J. L. Harris, C. S. Craik, *Curr. Opin. Chem. Biol.* **1998**, *2*, 127–132; d) R. J. Kazlauskas, *Curr. Opin. Chem. Biol.* **2000**, *4*, 81–88; e) Q.-S. Li, U. Schwaneberg, M. Fischer, J. Schmitt, J. Pleiss, S. Lutz-Wahl, R. D. Schmid, *Biochim. Biophys. Acta* **2001**, *1545*, 114–121.

[14] M. T. Reetz, L.-W. Wang, M. Bocola, *Angew. Chem.* **2006**, *118*, 1258–1263; *Erratum*, 2556; *Angew. Chem. Int. Ed.* **2006**, *45*, 1236–1241; *Erratum*, 2494.

[15] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher III, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin, M. Karplus, *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

[16] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, M. Karplus, *J. Comput. Chem.* **1983**, *4*, 187–217.

[17] The widely used QuikChange protocol (QuikChange Site-Directed Mutagenesis Kit, Instruction Manual, **2003**, Stratagene, La Jolla, CA) is based on a number of previous publications. These and other developments have been reviewed in: a) C. N. Dominy, D. W. Andrews in *Methods in Molecular Biology, Vol. 235: E. coli Plasmid Vectors* (Eds.: N. Casali, A. Preston), Humana, Totowa, **2003**, pp. 209–223; see also: b) M. A. Vandeyar, M. P. Weiner, C. J. Hutton, C. A. Batt, *Gene* **1988**, *65*, 129–133; c) R. D. Kirsch, E. Joly, *Nucleic Acids Res.* **1998**, *26*, 1848–1850.

[18] G. Sarkar, S. S. Sommer, *BioTechniques* **1990**, *8*, 404–407.

[19] A. A. Watson, R. A. Alm, J. S. Mattick, *Gene* **1996**, *172*, 163–164.

[20] K.-E. Jaeger, B. Schneidinger, F. Rosenau, M. Werner, D. Lang, B. W. Dijkstra, K. Schimossek, A. Zonta, M. T. Reetz, *J. Mol. Catal. B* **1997**, *3*, 3–12.

ELSEVIER

The lid is a structural and functional determinant of lipase activity and selectivity

Francesco Secundo [a,*], Giacomo Carrea [a], Chiara Tarabiono [a], Pietro Gatti-Lafranconi [b], Stefania Brocca [b], Marina Lotti [b], Karl-Erich Jaeger [c], Michael Puls [c], Thorsten Eggert [c]

[a] *Istituto di Chimica del Riconoscimento Molecolare–CNR, Via Mario Bianco 9, 20131 Milano, Italy*
[b] *Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Piazza della Scienza 2, 20126 Milano, Italy*
[c] *Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Düsseldorf, Forschungszentrum Jülich, Stetternicher Forst, D-52426 Jülich, Germany*

Available online 28 February 2006

Abstract

In several lipases access to the enzyme active site is regulated by the position of a mobile structure named the lid. The role of this region in modulating lipase function is reviewed in this paper analysing the results obtained with three different recombinant lipases modified in the lid sequence: *Candida rugosa* lipase isoform 1 (CRL1), *Pseudomonas fragi* lipase (PFL) and *Bacillus subtilis* lipase A (BSLA). A CRL chimera enzyme obtained by replacing its lid with that of another *C. rugosa* lipase isoform (CRL1LID3) was found to be affected in both activity and enantioselectivity in organic solvent. Variants of the PFL protein in which three polar lid residues were replaced with amino acids strictly conserved in homologous lipases displayed altered chain length preference profile and increased thermostability. On the other hand, insertion of lid structures from structurally homologous enzymes into BSLA, a lipase that naturally does not possess such a lid structure, caused a reduction in the enzyme activity and an altered substrate specificity. These results strongly support the concept that the lid plays an important role in modulating not only activity but also specifity, enantioselectivity and stability of lipase enzymes.
© 2006 Elsevier B.V. All rights reserved.

Keywords: *Candida rugosa* lipase isoform 1; *Bacillus subtilis* lipase A; *Pseudomonas fragi* lipase; Specificity; Temperature stability; Enantioselectivity; Site-directed mutagenesis; Domain swapping

1. Introduction

Lipases are the enzymes with the broadest use in biocatalysis [1–6]. Their application for the preparation of chiral building blocks, especially by kinetic resolution of racemic mixtures, is of particular interest for the pharmaceutical, agrochemical and food industries. In addition, a reason that contributes to the usefulness of lipases is their high activity in non-aqueous media (organic solvents, ionic liquids, solvent-free systems), in which the synthetic reaction is favored over hydrolysis. To improve the efficiency of lipases in biocatalysis, different approaches such as medium, substrate and protein engineering have been exploited [7,8]. In this frame, the ability to recognize and modify the molecular determinants of enzyme function might provide an effective starting point towards the optimization of biocatalytic processes. From a structural point of view, besides the

substrate binding region, there is evidence that suggest that the region of the lid – a mobile amphipatic structure which covers the catalytic active site of most lipases and whose length and complexity depend on the enzyme – is involved in modulating activity and selectivity of lipases [9–14]. In order to point out the importance of this structural element, herein we review our previous findings (and add some new results on the activity of BSLA and its mutant in organic solvents) on the effects of lid-related modifications and on the enzymatic properties of lipases from fungal and bacterial sources, differing from each other in sequence, presence/absence of a lid structure, molecular and biochemical features. Despite a marked divergence in sequence, all lipases feature the α/β-hydrolase fold [15,16], which makes approaches of rational mutagenesis and domain swapping possible. Targets of this comparative studies were *Candida rugosa* lipase isoform 1 (CRL1), *Pseudomonas fragi* lipase (PFL) and *Bacillus subtilis* lipase A (BSLA). Enzymes were subjected to three different approaches targeting their lids: domain exchange, sequence-based site-directed mutagenesis and domain insertion, based on the rationale discussed in the following. CRL1 belongs

* Corresponding author. Tel.: +39 02 28500029; fax: +39 02 28901239.
E-mail address: francesco.secundo@icrm.cnr.it (F. Secundo).

to a family of isoenzymes related in sequence but not identical in catalytic properties. The lid of CRL1 was substituted with the corresponding region of isoform 3 obtaining a chimera protein that was tested for activity and enantioselectivity [9,14]. PFL, a lipase endowed with activity at low temperature and an unusual specificity towards short-chain triglycerides, was modified by rational mutagenesis based on sequence comparison with homologous lipases showing higher temperature stability and specificity for medium- or long-chain triglycerides [17,18]. In the case of BSLA, since the wild type protein does not have a lid, the enzyme was modified by swapping lid sequences of the structurally related enzymes cutinase, acetylxylanesterase and human pancreatic lipase [19]. All modifications affected activity, specificity and, in the case of PFL, also stability of the proteins providing support to a new view of the lid structure as a key structural and functional element of lipases.

2. Lid sequence in wtCRL1, wtPFL, wtBSLA and their mutants

CRL isoenzymes 1 and 3 are related by 89% sequence identity on the overall peptide chain and the sequence of their lids differs for six amino acids as shown in Fig. 1a. Swapping of the CRL3 lid on the CRL1 scaffold therefore produces a chimera differing from the wtCRL1 in only 6 out of 534 residues. Fig. 1b compares the sequence of the PFL lid with those of other lipases from *Pseudomonas/Burkholderia* species related to each other and to PFL by approximately 40% sequence identity but differing from each other in temperature stability and substrate preference profile. Residues strictly conserved in the reference lipases at positions 137, 138 and 141 (valine, asparagines and glycine) are substituted in PFL by threonine (137 and 138) and serine (141).

The BSLA naturally does not possess a lid structure. In computer-based studies, three lipolytic enzymes that do exhibit a lid or lid-like structure, cutinase from *Fusarium solani pisi* [20], acetylxylanesterase from *Penicillium purpurogenum* [21] and the human pancreatic lipase [22], were chosen based on their structural homology to BSLA (Fig. 1c). The surroundings

of the active sites were compared in silico. The lids and lid-like structures of these three enzymes were modeled into the structure of BSLA indicating experimental options to engineer these lids into BSLA without disturbing the core of the α/β-hydrolase fold. The resulting variants were named CUTIlip, AXElip and HPlip, respectively.

3. Activity and enantioselectivity of wtCRL1 and CRL1LID3

Activity of the chimeric enzyme in aqueous medium was studied in a previous work, where it was shown that lid swapping conferred to CRL1 the ability to hydrolyze cholesterol esters typical of isoform 3. Moreover, wt and chimeric lipases displayed a different sensitivity to non-ionic detergents in the reaction mixture [9]. Activity and selectivity was further investigated in organic solvents [14]. Table 1 shows the results obtained in a model reaction, the alcoholysis of chloro ethyl 2-hydroxy hexanoate with methanol in hexane and isooctane. Specific activity was higher in the case of wtCRL1 at any tested water activity value. Highest transesterification activity was obtained with both enzymes at $a_w = 0.53$, thus ruling out effects related to different

Table 1
Transesterification rate of wtCRL1 and CRL1LID3 in organic solvents at different a_w-values

Enzyme	Rate (μmol/h per mg of lipase)					
	Hexane			Isooctane		
	a_w 0.06	a_w 0.53	a_w 0.84	a_w 0.06	a_w 0.53	a_w 0.84
wtCRL1	0.06	1	0.09	0.08	1.49	0.1
CRL1LID3	0.05	0.4	0.05	0.05	0.52	0.05

Transesterification was carried out with 0.5 mg of lipase, using as model reaction the alcoholysis of chloro ethyl 2-hydroxy hexanoate (0.013 M) with methanol (0.25 M). The mixture was shaken at 150 rpm and at 25 °C. The reaction progress was monitored by GLC (column: dimethylpentyl, β-cyclodextrin 25 m, 0.25 mm ID, MEGA) with an oven temperature from 90 to 130 °C with heating rate 2.5 °C/min.

```
        wtCRL1    EGTYEENLPKAALDLVMQSKVFEAVSPS   93
(a)     CRL1LID3  EGTFEENLGKTALDLVMQSKVFQAVLPQ   93


                              lid region
                  ┌─────────────────────────────────────────┐
        PFL       GPNHGSELADRLRL--AFVPGRLGETVAAALTTSFSAFLSALSGHPRLPQNALNALNALT   165
        BCL       TPHRGSEFADFVQDVLAYDPTGLSSSVIAAFVNVFGILTSSSHN---TNQDALAALQTLT   168
        PAL       APHKGSDTADFLR---QIPPGSAGEAVLSGLVNSLGALISFLSSGSTGTQNSLGSLESLN   168
(b)     BGL       TPHRGSEFADFVQDVLKTDPTGLSSTVIAAFVNVFGTLVSSSHN---TDQDALAALRTLT   163


        BSLA      31  wsrdklyaVD FWDKTGTNYN N---------- gpvlsrfvq kvldetgakk   70
        CUTIlip   31  wsrdklyaVG GAYRATLGDN ALPRGTSSAA- gpvlsrfvq kvldetgakk   79
        AXElip    31  wsrdklyaIN YPACGGQSSCGGASYSSSVAQG gpvlsrfvq kvldetgakk   80


        BSLA      141  srldgarnvq ihGVG----- ---------- ---------- higllyssq   164
(c)     HPlip     141  srldgarnvq ihMPGCLLNI LSQIVDIDGI WEGTRDFAAC higllyssq   189
```

Fig. 1. (a) Lid sequence of lipases from wtCRL1 and CRL1LID3, where residues differing in the two enzymes are highlighted in bold. (b) Comparison of the sequences of the lids of lipases from *Pseudomonas fragi* (PFL), *Burkholderia cepacia* (BCL), *Pseudomonas aeruginosa* (PAL) and *Burkholderia glumae* (BGL). Aminoacids subjected to mutagenesis in PFL are in bold. (c) Alignment of the structurally homologous sequences of *Bacillus subtilis* lipase A (BSLA) with the lid regions of its variants, generated by insertion of the lid sequences of *Fusarium solani pisi* cutinase (CUTIlip), *Penicillium purpurogenum* acetylxylane esterase (AXElip) and human pancreatic lipase (HPlip). The lid regions of CUTIlip and AXElip as well as the corresponding amino acids in BSLA are indicated in capital bold characters. In the same manner, the corresponding amino acids of BSLA and HPlip are indicated.

Table 2
Enhancement of activity of CRL1 and CRL1LID3 after bioimprinting with *N*-octyl-β-D-glucopyranoside

Enzyme	Enhancement factor[a]
wtCRL1	11.4
CRL1LID3	2.4

[a] Transesterification activity was determined at $a_w = 0.53$ using chloro ethyl 2-hydroxy hexanoate as the substrate, methanol as the nucleophile, petroleum ether as the solvent, under reaction conditions identical to those employed for the non-treated enzyme. For the bioimprinting procedure, see Ref. [14].

hydration states of the two proteins and pointing to the lid as involved in the lower catalytic activity of the chimera.

Several authors have attributed the lower activity of lipases in organic solvents than in aqueous media to the possibility that the enzymes are in the conformation with the lid closed (inactive). According to this hypothesis, the lower activity of CRL1LID3 could be ascribed to a smaller fraction of enzyme molecules trapped in the open (active) conformation. This point was investigated by applying bioimprinting [23], an approach that relies on treatment of lipase molecules with detergents that favor the open enzyme conformation. Such conformation is retained upon lyophilization giving rise to a population of activated enzymes,
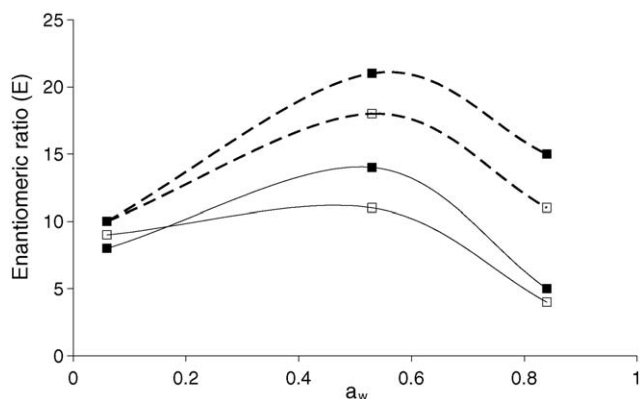
Fig. 2. Effect of water activity on the enantiomeric ratio (*E*) of wtCRL1 (dashed line) and CRL1LID3 in the kinetic resolution of a racemic mixture of 2-hydroxycaproic acid chloroethylester with methanol as nucleophile, in isooctane (□) and hexane (■). For reaction conditions, see Table 1. The *E*-values were calculated according to Ref. [25]. The conversion degree and the enantiomeric excess for *E* calculation were measured by GLC chromatography using the same conditions described in Table 1 legend.

when the preparation is resuspended in organic solvent. The lower enhancement factor observed for CRL1LID3 (Table 2) imprinted with *N*-octyl-β-D-glucopyranoside, is likely to be a consequence of a lower propensity of the chimera enzyme to shift from the closed to the open form.

Besides activity, also enantioselectivity was found to be affected by lid replacement when tested with the racemic substrate choro ethyl 2-hydroxy hexanoate in both isooctane and hexane at different a_w-values. In particular, CRL1LID3 showed a lower *E*-value than wtCRL1 in all tested conditions (Fig. 2). Interestingly, Colton et al. [24] have suggested that the increase of enantioselectivity of *C. rugosa* lipase observed upon treatment with 2-propanol might be due to the conversion of the enzyme conformation from the closed to the open form. This agrees with our hypothesis of a higher fraction of enzyme molecules in the open form in the case of CRL1 (that also shows higher *E*-values) compared to CRL1LID3.

4. Activity and stability of PLF and its lid mutants

The lipase from *P. fragi* is a cold-active enzyme as it retains 59% of its activity at 10 °C. At the same time, it is unstable at moderate temperatures with a half life of ca. 4 h at 29 °C. A further peculiar feature of this enzyme is a marked selectivity for short chain substrates [17]. Targets for mutagenesis were selected in the lid region based on the comparison with the sequences of homologous lipases differing from PFL both in thermostability and substrate preference. The alignment shown in Fig. 1b suggested to address residues conserved in the reference lipases, in particular at positions 137 and 138 where in PFL a polar threonine residue substitutes valine and asparagine, respectively, and position 141, where serine substitutes a glycine residue conserved in other lipases. Substrate specificity of PFL variants was assayed on triglycerides of growing chain length, i.e. tributyrin (C4), tricaprylin (C8) and trilaurin (C12). We noticed that substitutions T137V and T138N increased the relative activity on C8 substrates whereas no specific effects were produced by the substitution S141G [18]. The same proteins were further assayed for stability at 29 and 37 °C. Cold-adapted enzymes in fact often display high conformational flexibility as a molecular adaptation allowing them to counteract the decrease of catalytic efficiency at low temperature through low-energy cost interactions between the enzyme active site and the substrates [25]. Such flexibility causes low temperature stability.
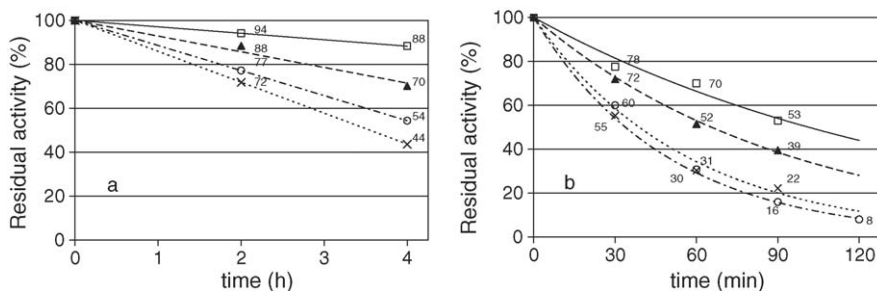

Fig. 3. Temperature-dependent loss of activity of wtPFL and mutants as measured in the hydrolysis of tricaprylin at 29 °C (a) and 37 °C (b). (○) wt; (□) T137V; (▲) T138N; (×) S141G.
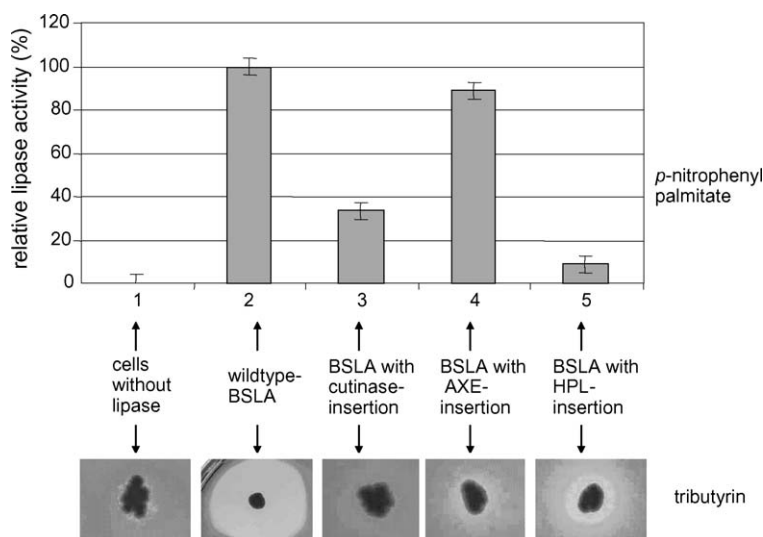
Fig. 4. Activity of BSLA and its lid-variants with the substrates *p*-nitrophenyl palmitate and tributyrin [17].

wtPFL and mutants were preincubated at either 29 or 37 °C and then assayed for activity at 29 °C, the T_{opt} of the enzyme. After 4 h incubation at 29 °C activity of wtPFL on tricaprylin was reduced by 50%, whereas mutants T137V and T138N retained 90 and 70% activity, respectively. At the less permissive temperature of 37 °C, differences become still more obvious (Fig. 3). Substitution S141G is destabilizing, possibly because it increases protein flexibility.

5. Activity and enantioselectivity of wtBSLA and its lid mutants

The BSLA-variants CUTIlip, AXElip and HPlip (Fig. 1c) were first assayed for enzymatic activity with the substrates *p*-nitrophenyl palmitate and tributyrin (Fig. 4). All of them showed enzymatic activity, at least against one of the substrates, although the variants containing artificial lid-structures exhibited a lower activity than the wild type [19].

BSLA is fairly active in organic solvent (petroleum ether). Instead the mutants have a modest activity in the transesterification reaction between sulcatol and vinylacetate and are inactive in the alcoholysis of chloro ethyl 2-hydroxy hexanoate with methanol (Table 3).

In the transesterification reaction between sulcatol and vinylacetate, carried out with the enzymes co-lyophilized with PEG (Table 3), after 260 h the degree of conversion (product formation) in the case of BSLA was about 30% and the enantiomeric ratio (*E*-value) was 10.2. The highest conversion degrees obtained in the case of BSLA lid-mutants were 14, 14 and 10% for CUTIlip, AXElip and HPlip, respectively; the negative control was 4%. The different conversion yields observed after 260 h should depend on different enzyme activities and stabilities. The results obtained with BSLA mutants with the substrate sulcatol are at least in part due to the mutated enzymes since both the reaction rates (Table 3) and the final conversion degree were higher than that observed in the case of the negative control, where an endogenous hydrolase activity might be

Table 3
Transesterification rate in petroleum ether of wtBSLA and BSLA lid-mutants

Enzyme	Rate (nmol/min)	
	Sulcatol[a]	Chloro ethyl 2-hydroxy hexanoate[b]
BSLA	122 (19)	566
HPlip	14 (3)	2
AXElip	16 (3)	3
CUTIlip	18 (0.5)	2
Negative control	9 (0.4)	3

Forty milligrams of enzyme were lyophilized with 5 mg of PEG and added to a final reaction volume of 1 ml. The reaction mixture was shaken at 150 rpm and at 25 °C. The reaction progress was monitored by GLC using the same conditions described in Table 1 legend. In parentheses, the rate obtained using the same amount of enzyme lyophilized without PEG.

[a] Rate of transesterification at $a_w = 0.11$ of sulcatol (0.013 M) with vinyl acetate (0.11 M) as acyl donor.

[b] Rate of transesterification at $a_w = 0.11$ of chloro ethyl 2-hydroxy hexanoate (0.013 M) with methanol (0.25 M).

present in the host microorganism. Because of the interference of the transesterification activity of the host microorganism and of the low activity values, a precise estimation of the BSLA lid-mutants enantioselectivity was not possible. It is interesting to note that the BSLA is more active in the alcoholysis of chloro ethyl 2-hydroxy hexanoate (the *E*-value with this substrate was 4.0) than in the transesterification reaction with sulcatol. This is just the opposite of what was observed with the mutants. Co-lyophilization of all enzyme preparations with methoxy poly(ethylene glycol) (PEG), an additive usually used to preserve the protein native structure during freeze-drying, led to a substantial increase of activity (Table 3) [26].

6. Conclusions

The data we present in this work indicate that the lid is involved in the modulation of lipase's functional properties as all changes introduced affected the behaviour of the target

enzymes. Activity and enantioselectivity of wtCRL1 in organic solvents were reduced by mutations in the lid sequence. On the other hand, the chimeric enzyme displayed novel specificities in aqueous solution as, for example, the ability of hydrolysing cholesterol esters. In the case of PFL, substitutions at T137 and T138 altered the chain length preference profile and increased the temperature stability of the enzyme. Insertion of lid fragments into BSLA yielded variants active in aqueous media and also in organic solvent although the activity in the latter was only 4–7% of the wild type enzyme. These results showed that the insertion of different lids into the same core-protein indeed result in different properties of the variants concerning substrate specificity. The overall decrease in activity of the variants compared to wtBSLA can be accounted by the fact that the foreign lid-structures are not as adapted to the core-protein of BSLA as they are to their host-protein structure.

In conclusion, the lid appears to play a very complex role in lipases activity, specificity and conformational stability. We have shown that protein engineering of lid structures might provide enzymes with new properties but often causes a reduction in the enzyme activity. This effect might be counteracted by adapting the modified lid to the protein context, i.e. through an approach of directed evolution.

References

[1] K. Drauz, H. Waldmann, Enzyme Catalysis in Organic Synthesis: A Comprehensive Handbook, vols. I–III, second ed., Wiley–VCH, Weinheim, 2002.
[2] K.E. Jaeger, T. Eggert, Curr. Opin. Biotechnol. 13 (2002) 390–397.
[3] M.T. Reetz, Curr. Opin. Chem. Biol. 6 (2002) 145–150.
[4] K.E. Jaeger, S. Ransac, B.W. Dijkstra, C. Colson, M. van Heuvel, O. Misset, FEMS Microbiol. Rev. 15 (1994) 29–63.
[5] R.D. Schmid, R. Verger, Angew. Chem. Int. Ed. Engl. 37 (1998) 1608–1633.
[6] R.J. Kazlauskas, U.T. Bornscheuer, in: H.J. Rehm, G. Reed, A. Pühler, P.J.W. Stadler, D.R. Kelly (Eds.), Biotechnology, Biotransformations with Lipases, vol. 8a, Wiley–VCH, Weinheim, 1998, pp. 37–191.
[7] A. Svendsen, Biochim. Biophys. Acta 1543 (2000) 223–238.
[8] P. Villeneuve, J.M. Muderhwa, J. Graille, M.J. Haas, J. Mol. Catal. B: Enzym. 9 (2000) 113–148.
[9] S. Brocca, F. Secundo, M. Ossola, L. Alberghina, G. Carrea, M. Lotti, Protein Sci. 12 (2003), 2312–2319 (and references therein).
[10] K.A. Dugi, H.L. Dichek, S. Santamarina-Fojo, J. Biol. Chem. 270 (1995) 25396–25401.
[11] M. Holmquist, I.G. Clausen, S. Patkar, A. Svendsen, K. Hult, J. Protein Chem. 14 (1995) 217–224.
[12] M. Holmquist, M. Martinelle, P. Berglund, I.G. Clausen, S. Patkar, A. Svendsen, K. Hult, J. Protein Chem. 12 (1995) 749–757.
[13] M. Martinelle, M. Holmquist, I.G. Clausen, S. Patkar, A. Svendsen, K. Hult, Protein Eng. 9 (1996) 519–524.
[14] F. Secundo, G. Carrea, C. Tarabiono, S. Brocca, M. Lotti, Biotechnol. Bioeng. 86 (2004) 236–240.
[15] D.L. Ollis, E. Cheah, B. Dijkstra, F. Frolow, S.M. Franken, M. Harel, S.J. Remington, I. Silman, J. Schrag, J.L. Sussman, K.H.G. Verscheuren, A. Goldman, Protein Eng. 5 (1992) 197–211.
[16] M. Nardini, B.W. Dijkstra, Curr. Opin. Struct. Biol. 9 (1999) 732–737.
[17] C. Alquati, L. DeGioia, G. Santarossa, L. Alberghina, P. Fantucci, M. Lotti, Eur. J. Biochem. 269 (2002) 3321–3328.
[18] G. Santarossa, P. Gatti Lafranconi, C. Alquati, L. De Gioia, L. Alberghina, P. Fantucci, M. Lotti, FEBS Lett. 579 (2005) 2383–2386.
[19] T. Eggert, C. Leggewie, M. Puls, W. Streit, G. Van Pouderoyen, B.W. Dijkstra, K.-E. Jaeger, Biocatal. Biotransform. 22 (2004) 139–144.
[20] C. Martinez, P. De Geus, M. Lauwereys, G. Matthyssens, C. Cambillau, Nature 356 (1992) 615–618.
[21] D. Ghosh, M. Erman, M. Sawicki, L. Puloma, D.R. Weeks, L. Naiyin, W. Pangborn, D.J. Thiel, R. Gutiérrez, J. Eyzaguirre, Acta Crystallogr. D 55 (1999) 779–784.
[22] F.K. Winkler, A. D'Arcy, W. Hunziker, Nature 343 (1990) 771–774.
[23] H. Gonzalez-Navarro, M.C. Bano, C. Abad, Biochemistry 40 (2001) 3174–3183.
[24] I.J. Colton, N. Ahmed, R.J. Kazlauskas, J. Org. Chem. 60 (1995) 212–217.
[25] G. Feller, CMLS Cell. Mol. Life Sci. 60 (2003) 648–662.
[26] F. Secundo, G. Carrea, Chem. Eur. J. 9 (2003) 3194–3199.

insilico.mutagenesis: a primer selection tool designed for sequence scanning applications used in directed evolution experiments

Ulrich Krauss and Thorsten Eggert

Heinrich-Heine-Universität Düsseldorf, Jülich, Germany

*Several primer prediction programs have been developed for a variety of applications. However, none of these tools allows the prediction of a large set of primers for whole gene site-directed mutagenesis experiments using the megaprimer method. We report a novel primer prediction tool (*insilico.mutagenesis*), accessible at www.insilico.uni-duesseldorf.de, developed for the application to high-throughput mutagenesis used in directed evolution or structure-function dependency projects, which involve the subsequent mutagenesis of a large number of amino acid positions (e.g., in whole gene saturation or gene scanning mutagenesis experiments). Furthermore, the program is suitable for all site-directed (saturation) mutagenesis approaches, such as saturation mutagenesis of promoter sequences and other types of untranslated intergenic regions. In anticipation of downstream cloning steps, the primer design tool also includes a restriction site control feature alerting the user if unwanted restriction sites have been introduced within the mutagenesis primer. The use of our tool promises to speed up the process of site-directed mutagenesis, as it instantly allows predicting a large set of primers.*

INTRODUCTION

The design of PCR primers (i.e., single-stranded DNA oligonucleotides complementary to a target sequence) has become an essential procedure in molecular biology for a variety of applications (e.g., gene amplification, sequencing, and site-directed mutagenesis). Thus, numerous primer prediction and analysis programs have been developed (1–4). However, none of the programs currently available facilitate the design of primers for whole gene site saturation and sequence-scanning mutagenesis experiments using the megaprimer PCR method (5). The manual design of oligonucleotide PCR primers for these approaches is a laborious task, due to the large set of mutagenesis primers to be designed. In order to reduce the working time and to standardize the design procedure, we have developed a web-based primer prediction program termed *insilico. mutagenesis*. The program takes nucleotide sequences of target regions (open reading frames or intergenic regions

like promoter sequences) as input and predicts mutagenesis primers that can be directly used in a megaprimer PCR-based mutagenesis approach. The *insilico.mutagenesis* tool is entirely written in Perl (6), uses MySQL tables for easy data storage, and possesses an HTML-based user interface. The tool can be accessed via the *insilico* web site (www.insilico.uni-duesseldorf.de).

Directed evolution has been proven to be a successful strategy to improve enzyme properties such as specific activities, substrate specificities, thermostabilities, or enantioselectivities (7–10). In most cases, single base mutations are introduced by means of error-prone PCR (epPCR) in a random manner. However, epPCR only results in a limited number of amino acid exchanges; therefore, only a small part of the total sequence space is accessible to mutagenesis. As a consequence, alternative techniques must be applied to generate a first generation library of high diversity (11). Complete saturation mutagenesis, also referred to as Gene Site Saturation Mutagenesis™

(GSSM™), is a novel technology for rapid in vitro evolution of proteins that can be used to circumvent this problem (12–14). Here, all possible base triplets are introduced at a given codon position, thereby resulting in the formation of a library containing all 20 amino acid exchanges at the target position. This is achieved at the genetic level by using degenerate mutagenesis primers. Subsequent use of in vitro PCR amplification generates a library of genes possessing all codon variations required for complete saturation of the original gene. DeSantis and coworkers applied this technique to generate a highly enantioselective nitrilase (13). Furthermore, the technique of complete saturation mutagenesis has been used in our institute to generate a variant *Bacillus subtilis* lipase A (BLSA) showing improved enantioselectivity toward different model substrates (14,15). Sequence-scanning mutagenesis techniques, like alanine- or tryptophan-scanning mutagenesis, can be applied to investigate the functional role of specific amino acid residues with respect to catalytic mechanism, substrate binding, or signal transduction (16,17).

Both complete saturation mutagenesis and scanning-mutagenesis techniques require the sequential saturation/substitution of numerous amino acid residues, depending on the size of the target protein or the region to be investigated. For the complete saturation of a regular protein consisting of 300 amino acids, 300 single codon exchanges (i.e., 300 megaprimer PCRs) must be performed. One step in such a challenging approach that is easily amenable to automation without the necessity of expensive robotic equipment, is the primer design using a personal computer. Therefore, we developed the program *insilico.mutagenesis* to automate the prediction of oligonucleotides that can be used directly in a megaprimer PCR approach.

REQUIRED INPUT

A schematic overview of the data-processing by *insilico.mutagenesis* is given in Figure 1. First, the program

requires the input of a target nucleotide sequence, including flanking vector sequences (plain sequence), to which we will refer as vectorA-template. It is not necessary to include the complete vector-sequence; about 40 bp up- and downstream of the gene of interest is enough to enable primer design for whole gene saturation or scanning mutagenesis. Second, a unique sequence identifier (sequence name) must be provided for data processing purposes. Third, the mutagenesis codon must be selected from a pull-down menu, taking into account the codon-usage of the desired expression host. Next, the program requires the input of the start and stop position of the target gene (or intergenic region) within the overall sequence. Also, the region that should be mutated must be specified, and because, in practice, a too long megaprimer might be inefficiently elongated by the polymerase in the second round of PCR—probably due to the formation of secondary structures—it has been proven practical to design mutagenesis primers in a way that the megaprimer does not exceed a certain length (5). Therefore, the input of a so-called oligo-switch position is necessary, as explained in more detail later.

PROGRAM ALGORITHM AND DATA PROCESSING

As one example, we designed all mutagenesis primers in the complete saturation mutagenesis of a *B. subtilis* lipase (14). The gene was 543 bp in length, consequently having 181 coding triplets and a TAA stop codon. The first 90 codon-exchanges (corresponding to 270 bp) are achieved by the design of reverse mutagenesis primers (as shown in Figure 2B), which are used together with a vectorA-specific forward primer. Accordingly, the last 91 codon-exchanges are introduced using a forward mutagenesis primer together with a vectorA-specific reverse primer. As a consequence, the amplified megaprimers do not exceed 273 bp in size. Therefore, these DNA fragments are well suited for the second PCR of megaprimer mutagenesis (5,18). The position at which the "switch" from
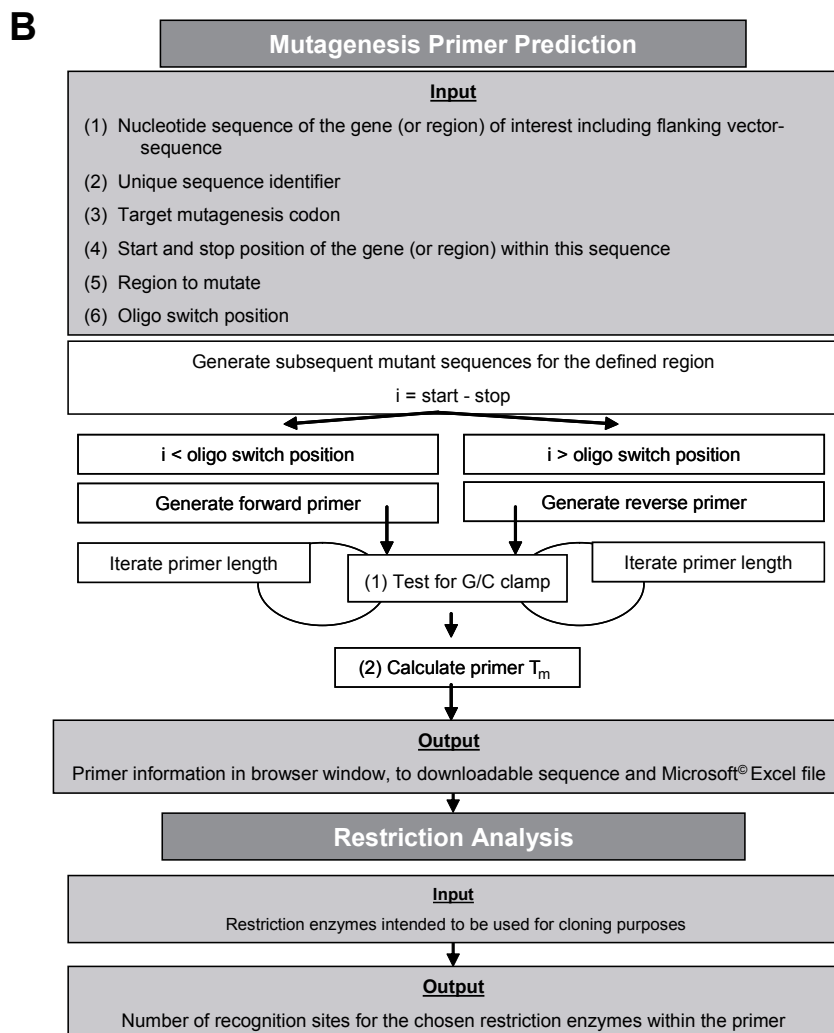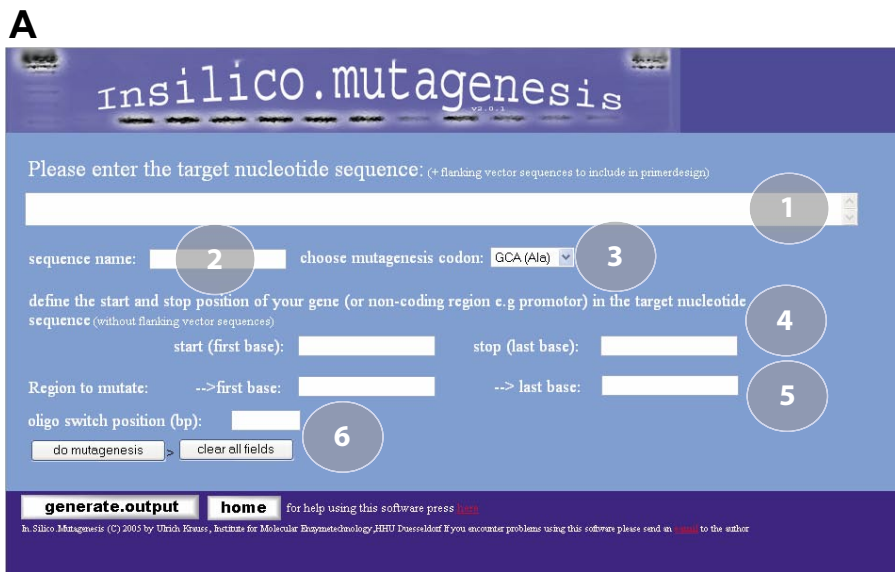
Figure 1. User interface and flowchart of data processing of the *insilico.mutagenesis* primer prediction tool. (A) Screenshot of the tool's data input user interface. (B) Flowchart of the tool's data processing and data generation algorithm. T_m, melting temperature.

a reverse to a forward mutagenesis primer occurs is referred to as the oligo-switch position. Usually, as in our lipase example, the position halfway along the gene of interest is used.

Using this input (Figures 1 and 2) the program generates mutant nucleotide sequences with subsequent single codon-exchanges within the region the user has defined. For each of the mutant sequences, a mutagenesis primer is predicted either as a forward or reverse primer, depending on the position of the desired mutation with respect to the oligo-switch position. The generated mutagenesis primers are checked for the ability to form the so-called GC clamps at the 3′ end, since Watson-Crick bonds between G and C will facilitate the initiation of complementary strand formation by the polymerase at the 3′ end of the hybridized primer (19). If no GC clamps can be formed due to the lack of G or C bases at the 3′ end of the primer, the program extends the oligonucleotide until a G or C is found

at its 3′ end. The maximum length of the primer is set at 40 bp. Finally, the mutagenesis primer data are stored in a MySQL database and displayed in form of an HTML table (Figure 1). In addition, the program calculates the melting temperature of every single primer based on the equation of Breslauer et al. (20) and the nearest neighbor thermodynamic parameter set as described by Allawi and SantaLucia (21). Furthermore, the predicted primer sequences can be viewed as FASTA-formatted text output in the browser window or can be downloaded as a Microsoft® Excel® spreadsheet.

ADDITIONAL ANALYSES

The program enables the user to check each predicted mutagenesis primer with respect to additional restriction endonuclease recognition sites, which might interfere with the intended cloning strategy for the

amplified full-length PCR product. Therefore, the program asks the user to supply the names of two restriction enzymes that will be used in the subsequent cloning steps. By using the BioPerl (22) module Bio::Restriction:: Analysis, *insilico.mutagenesis* indicates the number of recognition sites of those enzymes within each oligonucleotide. In case the program has predicted a mutagenesis primer whose sequence interferes with the desired cloning strategy, the oligonucleotide can be redesigned easily.

CONCLUSIONS

In summary, we have presented a novel primer design tool (*insilico. mutagenesis*) specifically developed for high-throughput mutagenesis primer prediction, useful in complete saturation and whole gene scanning mutagenesis experiments. Thus, *insilico.mutagenesis* is designed to speed up the process of directed evolution or structure-function dependency projects. Furthermore, the primer design tool includes a restriction site control feature alerting the user in case of introducing unwanted restriction sites within the mutagenesis primer anticipating the cloning strategy.

COMPETING INTERESTS STATEMENT
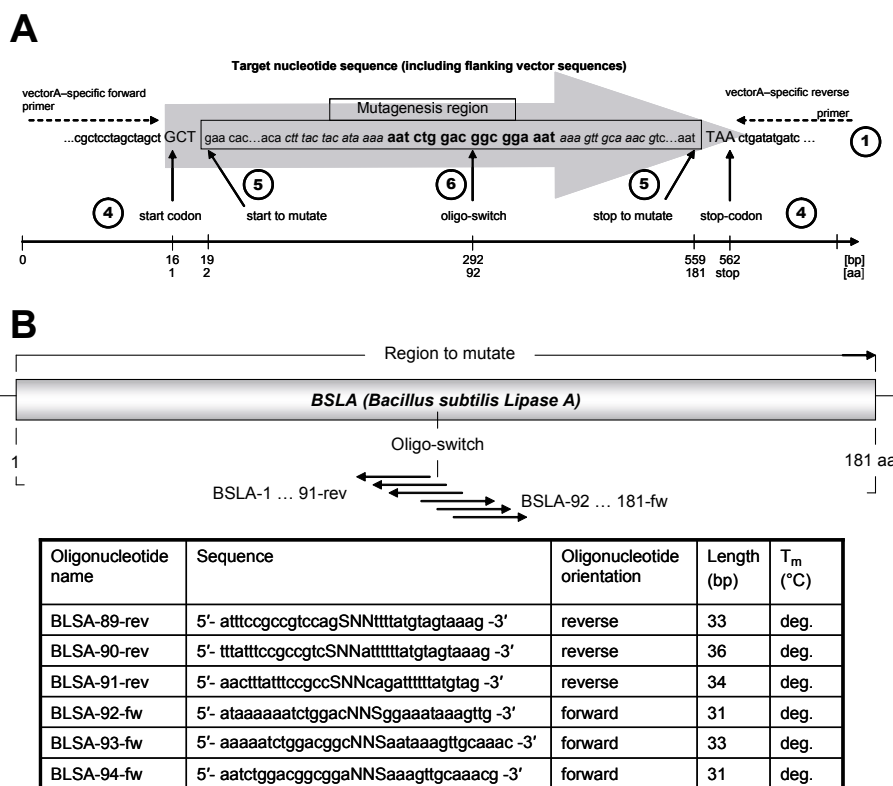
The authors declare no competing interests.

Figure 2. Application of *insilico.mutagenesis* to design primers for site-directed mutagenesis. (A) Visualization of the *Bacillus subtilis* Lipase A (BSLA) nucleotide sequence, explaining the input requirements of the *insilico.mutagenesis* tool. Codons 89–94 of the BSLA gene sequence are printed in bold face. (B) Orientation of mutagenesis primers around the oligo-switch position in the BSLA gene and sequences of the oligonucleotides computed by the *insilico.mutagenesis* tool. deg.,degenerate; T_m, melting temperature.

Oligonucleotide name	Sequence	Oligonucleotide orientation	Length (bp)	T_m (°C)
BLSA-89-rev	5′- atttccgccgtccagSNNttttatgtagtaaag -3′	reverse	33	deg.
BLSA-90-rev	5′- tttatttccgccgtcSNNatttttatgtagtaaag -3′	reverse	36	deg.
BLSA-91-rev	5′- aactttatttccgccSNNcagatttttatgtag -3′	reverse	34	deg.
BLSA-92-fw	5′- ataaaaaatctggacNNSggaaataaagttg -3′	forward	31	deg.
BLSA-93-fw	5′- aaaaatctggacggcNNSaatasaagttgcaaac -3′	forward	33	deg.
BLSA-94-fw	5′- aatctggacggcggaNNSaaagttgcaaacg -3′	forward	31	deg.

REFERENCES

1. **Rozen, S. and H. Skaletsky.** 2000. Primer3 on the WWW for general users and for biologist programmers. Methods Mol. Biol. *132*:365-386.
2. **Lu, G., M. Hallet, S. Pollock, and D. Thomas.** 2003. DePie: designing primers

for protein interaction experiments. Nucleic Acids Res. *31*:3755-3757.

3. **Turchin, A. and J.F. Lawler, Jr.** 1999. The Primer Generator: a program that facilitates the selection of oligonucleotides for site-directed mutagenesis. BioTechniques *26*:672-676.

4. **Canaves, J.M., A. Morse, and B. West.** 2004. PCR primer selection tool optimized for high-throughput proteomics and structural genomics. BioTechniques *36*:1040-1042.

5. **Barettino, D., M. Eigenbutz, R. Valcarcel, and H.G. Stunnernberg.** 1994. Improved method for PCR-mediated site-directed mutagenesis. Nucleic Acids Res. *22*:541-542.

6. **Moorhouse, M. and P. Barry.** 2004. Bioinformatics, Biocomputing and Perl. John Wiley & Sons, New York.

7. **Petrounia, I.P. and F. Arnold.** 2000. Designed evolution of enzymatic properties. Curr. Opin. Biotechnol. *11*:325-330.

8. **Brakmann, S.** 2001. Discovery of superior enzymes by directed molecular evolution. ChemBioChem. *2*:865-871.

9. **Reetz, M.T.** 2004. Controlling the enantioselectivity of enzymes by directed evolution: practical and theoretical ramifications. Proc. Natl. Acad. Sci. USA *101*:5716-5722.

10. **Jaeger, K.-E. and T. Eggert.** 2004. Enantioselective biocatalysis optimized by directed evolution. Curr. Opin. Biotechnol. *15*:305-313.

11. **Eggert, T., M.T. Reetz, and K.-E. Jaeger.** 2004. Directed evolution by random mutagenesis: a critical evaluation, p. 375-390. *In* A. Svendsen (Ed.), Enzyme Functionality: Design, Engineering, and Screening. Marcel Dekker, New York.

12. **Short, J.M., inventor.** Diversa Corporation, assignee. 2001. Saturation mutagenesis in directed evolution. U.S. Patent No. 6.171.820.

13. **DeSantis, G., K. Wong, B. Farwell, K. Chatman, Z. Zhu, G. Tomlinson, H. Huang, X. Tan, et al.** 2003. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). J. Am. Chem. Soc. *125*:11476-11477.

14. **Funke, S.A., A. Eipper, M.T. Reetz, N. Otte, W. Thiel, G. van Pouderoyen, B.W. Dijkstra, K.-E. Jaeger, and T. Eggert.** 2003. Directed evolution of an enantioselective *Bacillus subtilis* lipase. Biocatal. Biotransformation *21*:67-73.

15. **Funke, S.A., N. Otte, T. Eggert, M. Bocola, K.-E. Jaeger, and W. Thiel.** Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. Protein Eng. Des. Sel. (In press).

16. **Mahan, S.D., G.C. Ireton, B.L. Stoddard, and M.E. Black.** 2004. Alanine-scanning mutagenesis reveals a cytosine deaminase mutant with altered substrate preference. Biochemistry *43*:8957-8964.

17. **Santiago, J., G.R. Guzman, K. Torruellas, L.V. Rojas, and J.A. Lasalde-Dominicci.** 2004. Tryptophan scanning mutagenesis in the TM3 domain of the *Torpedo califonica* acetylcholine receptor beta subunit reveals an alpha-helical structure. Biochemistry *43*:10064-10070.

18. **Ling, M.M. and B.H. Robinson.** 1997. Approaches to DNA mutagenesis: an overview. Anal. Biochem. *254*:157-178.

19. **Lowe, T.M., J. Sharefkin, S.Q. Yang, and C.W. Dieffenbach.** 1990. A computer program for selection of oligonucleotide primers for the polymerase chain reaction. Nucleic Acids Res. *18*:1757-1761.

20. **Breslauer, K.J., R. Frank, H. Blocker, and L.A. Marky.** 1986. Predicting DNA duplex stability from the base sequence. Proc. Natl. Acad. Sci. USA *83*:3746-3750.

21. **Allawi, H.T. and J. SantaLucia, Jr.** 1997. Thermodynamics and NMR of internal G.T mismatches in DNA. Biochemistry *36*:10581-10594.

22. **Stajich, J.E., D. Block, K. Boulez, S.E. Brenner, S.A. Chervitz, C. Dagdigian, G. Fuellen, J.G. Gilbert, et al.** 2002. The bioperl toolkit: Perl modules for live sciences. Genome Res. *12*:1611-1618.

To purchase reprints

of this article, contact

Reprints@BioTechniques.com

Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution

Susanne Aileen Funke[1,3], Nikolaj Otte[2,3], Thorsten Eggert[1,4], Marco Bocola[2], Karl-Erich Jaeger[1] and Walter Thiel[2,4]

[1]Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Düsseldorf, Forschungszentrum Jülich, D-52426 Jülich and
[2]Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany

[3]These authors contributed equally to this paper.

[4]To whom correspondence should be addressed.
E-mail: t.eggert@fz-juelich.de; thiel@mpi-muelheim.mpg.de

Chiral compounds can be produced efficiently by using biocatalysts. However, wild-type enzymes often do not meet the requirements of a production process, making optimization by rational design or directed evolution necessary. Here, we studied the lipase-catalyzed hydrolysis of the model substrate 1-(2-naphthyl)ethyl acetate both theoretically and experimentally. We found that a computational equivalent of alanine scanning mutagenesis based on QM/MM methodology can be applied to identify amino acid positions important for the activity of the enzyme. The theoretical results are consistent with concomitant experimental work using complete saturation mutagenesis and high-throughput screening of the target biocatalyst, a lipase from *Bacillus subtilis*. Both QM/MM-based calculations and molecular biology experiments identify histidine 76 as a residue that strongly affects the catalytic activity. The experiments demonstrate its important influence on enantioselectivity.
Keywords: directed evolution/enantioselectivity/molecular modeling/QM/MM calculation/saturation mutagenesis

Introduction

The use of enzymes as natural catalysts for chemical processes, also referred to as 'white biotechnology', is a rapidly expanding field (Liese *et al*., 2000; Patel, 2003; Jaeger, 2004; Panke *et al*., 2004). The increasing demand to find useful biocatalysts has prompted the development of novel methods to identify new genes and isolate the corresponding biocatalyst proteins (Lorenz *et al*., 2002; Eggert *et al*., 2004a; Streit *et al*., 2004). However, enzyme properties normally do not fit the needs of a chemical process and, therefore, an array of molecular biological methods have been developed for enzyme optimization, with the most successful being directed evolution, which allows for the improvement of enzyme properties such as specific activity, substrate specificity and stability (Petrounia and Arnold, 2000; Cherry and Fidantsef, 2003; Robertson and Steer, 2004). Importantly, directed evolution can also be used to create enantioselective biocatalysts starting from non-selective wild-type enzymes (Jaeger and Eggert, 2004; Reetz, 2004). An effective directed evolution strategy requires the combination of different mutagenesis methods with

efficient high-throughput screening or selection techniques (Reetz and Jaeger, 2002; Jaeger and Eggert, 2004). In particular, the quality of the first-generation mutagenesis library is of key importance, because its variants usually parent all subsequent generations. Therefore, molecular biology methods generating mutant libraries of high diversity such as error-prone PCR or DNA shuffling are widely used, but, unfortunately, such techniques produce libraries which may consist of up to 10^{12} individual clones, far exceeding current screening capacities. Complete saturation mutagenesis represents an alternative and targeted strategy which generates a library containing all possible single amino acid exchanges of a target enzyme. This method creates mutant libraries of high diversity consisting of 10^3–10^4 individual variants, allowing scans of the entire sequence space of a given protein for important amino acid positions (Eggert *et al*., 2004b). Complete saturation mutagenesis, also referred to as gene site saturation mutagenesis (GSSM) (Short, 2001), has proved to be useful in the evolution of a dehalogenase from *Rhodococcus* (Gray *et al*., 2001), of a nitrilase isolated from the metagenome (DeSantis *et al*., 2003) and of a lipase from the Gram-positive bacterium *Bacillus subtilis* (Funke *et al*., 2003). Alternatively, alanine scanning mutagenesis can be used to identify 'hot spot' positions in a given enzyme (Cunningham and Wells, 1989; Weiss *et al*., 2000). Here, single alanine substitutions are introduced at each amino acid position of the respective enzyme to investigate the contribution of every side chain to a particular property. Still, complete saturation and alanine scanning mutagenesis remain labor- and cost-intensive strategies and alternative but complementary methods are urgently needed to allow the prediction of important amino acids which would narrow down the number of positions to be saturated.

Combined quantum mechanical and molecular mechanical (QM/MM) methods provide a realistic approach to compute the influence of individual amino acids on a given enzymatic reaction. These methods allow the study of chemical reactions in their native surroundings, where the reacting groups in the active site are treated at the QM level and the protein environment is simulated at the MM level (Sherwood *et al*., 2003). We chose this approach to estimate the electrostatic influence of all amino acid side chains in *Bacillus subtilis* lipase A (BSLA) on the rate-determining reaction barrier for ester hydrolysis.

Materials and methods

1-(2-Naphthyl)ethyl acetate (NEA) agar plate assay

The indicator agar plates were prepared as described previously for detecting cutin hydrolysis (Kolattukudy *et al*., 1981). The enantiomerically pure substrates (*R*)- and (*S*)-NEA (0.25 mg), kindly provided by Professor M.T.Reetz (Max-Planck-Institut für Kohlenforschung, Mülheim a.d. Ruhr,

Germany) were dissolved in 5 ml of dichloromethane containing the detergent Triton X-100 (300 mg). After evaporating the organic solvent (12 h at room temperature), the substrate mixture was emulsified in 6 ml of distilled water using ultrasonication, mixed with 50 ml of sterilized LB-Agar (10 g/l tryptone, 10 g/l NaCl, 5 g/l yeast extract) and vigorously homogenized using a high-speed mixer (Ultra-Turrax). The addition to the agar of Solvent Blue 38 (100 mg of dye per 50 ml of agar) significantly increased the contrast of clear halos formed by substrate hydrolysis.

Gas chromatographic (GC) analysis

For substrate conversion, 10 mM *rac*-NEA in 100 mM Tris–HCl buffer (pH 7.5) was incubated for 48 h at 20°C after adding lipase-containing cell extracts. The reaction products were isolated by extraction using ethyl acetate. GC analysis was performed on a Shimadzu GC-17A gas chromatograph. To separate both enantiomers of NEA, the following conditions were used: column, CP-Chirasil-DEX CB, 25 m × 0.25 mm i.d.; carrier gas, helium; temperature program, 5 min at 60°C, increased from 60 to 195°C at 5°C/min.

Computational QM/MM strategy

All computational studies were based on the crystal structure of BSLA (van Pouderoyen *et al.*, 2001). We started from a structure with an isopropylidene glycol phosphonate inhibitor bound to the active serine (Ser77) (PDB accession code 1R4Z, 1R5O). We replaced the inhibitor with (*R*)-NEA in a tetrahedral configuration and hydrated the active site with a sphere of water (radius 25 Å). An iterative procedure of relaxation, rehydration and molecular dynamics was performed at the MM level to obtain a sensible model for our QM/MM calculations (Bocola *et al.*, 2004). All pure MM calculations were done with the CHARMM27 (MacKerell *et al.*, 1998) force field as implemented in the Charmm program (Brooks *et al.*, 1983) (version c28b2). The QM/MM calculations employed Chemshell (Sherwood *et al.*, 2003), which is a modular package that allows the combination of several QM and MM codes. Here we used Turbomole (Ahlrichs *et al.*, 1989) at the BLYP (Becke, 1988; Lee *et al.*, 1988)/6–31+G* (Hehre *et al.*, 1972; Hariharan and Pople, 1973; Clark *et al.*, 1983) level for the QM part and DL_POLY (Smith and Forester, 1996) as driver of the CHARMM27 force field for the MM part. The QM region contained 34 atoms (colored atoms in Figure 1, excluding the naphthyl ring and H76) and open valencies at the QM/MM boundary were satisfied with hydrogen link atoms. We performed geometry optimizations with the HDLC optimizer (Billeter *et al.*, 2001) and located educt, transition state and product geometries for the first step of the ester hydrolysis reaction, the nucleophilic attack of the serine side chain (Ser77) on the planar ester carbon of NEA, yielding the tetrahedral intermediate (Scheme 1). The electrostatic impact on reaction barriers was estimated by a scan over all amino acid residues. In this procedure, we successively deleted the MM partial charges on the side chains of individual amino acids. Each of the charge sets thus obtained was used to re-evaluate the electronic energies of the tetrahedral intermediate, the transition state and the Michaelis complex. The electron densities were allowed to relax in the modified charge field. The scan was done on all 175 amino acid side chains resolved in the X-ray structure of the enzyme, excluding Ser77, Asp133 and His156, which belong to the catalytic triad and are within the
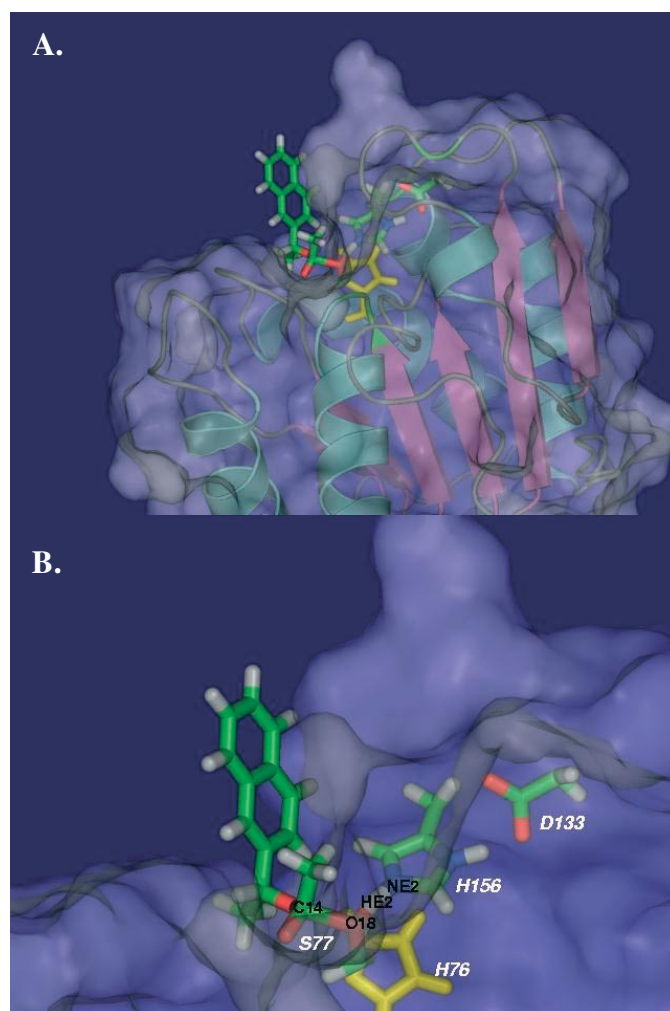

Fig. 1. Structural view of BSLA (**A**) and its active site pocket (**B**). Amino acids of the proposed catalytic triad and His76 are displayed as stick models. Ser77 and (*R*)-NEA form a tetrahedral intermediate (see **2** in Scheme 1).

QM region. Further computational details are given in the Supplementary data, available at *PEDS* Online.

Computational results

Bacterial lipases play an important role in biotechnology, mainly owing to their potential for catalyzing ester hydrolysis and also synthesis reactions, which often proceed with high specificity and enantioselectivity (Jaeger and Eggert, 2004). Among bacterial lipases, BSLA is a unique enzyme because it represents a minimal α/β-hydrolase (van Pouderoyen *et al.*, 2001). Furthermore, this enzyme hydrolyzes several acetic acid esters of secondary alcohols with high enantioselectivities, namely menthyl acetate, 1-phenylethyl acetate and 1-(2-naphthyl)ethyl acetate (NEA) giving *E* values of 22 (for the 1*R*,2*S*,5*R*-enantiomer), >100 (*R*) and >140 (*R*), respectively. In the hydrolysis of NEA, this lipase accepts virtually no (*S*)-enantiomer, resulting in an *ee* value of >99%. Therefore, we chose to study by QM/MM calculations the ester hydrolysis of NEA catalyzed by BSLA.

The mechanism of this reaction involves as the rate-determining catalytic step (Scheme 1) a nucleophilic attack of the Ser77 side-chain oxygen on the carbonyl carbon
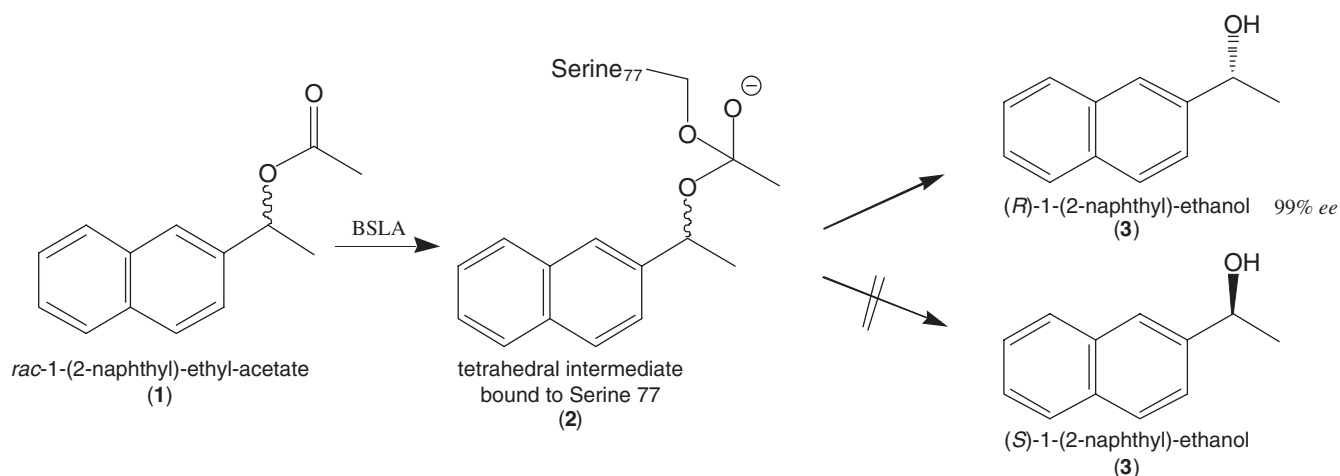
Scheme I. Reaction scheme for substrate hydrolysis by wild-type BSLA which shows an enantioselectivity of $E > 140$ towards (*R*)-NEA.
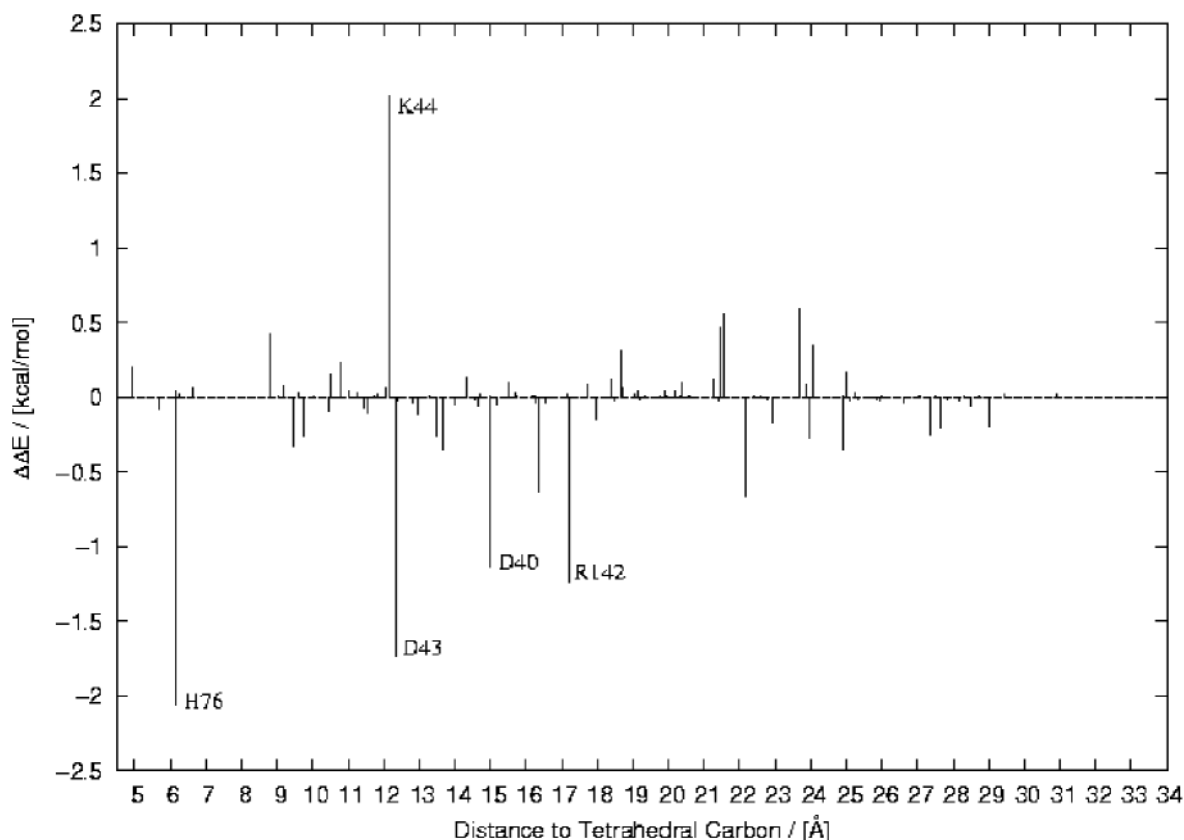

Fig. 2. Modulation of the reaction barrier height by electrostatic perturbation of the environment (deletion of charges on side chains). The distance is measured from the tetrahedral carbon in the tetrahedral intermediate (C14 in Figure 1) to the geometric center of the individual amino acid side chains. Large contributions are labeled.

under general base catalysis of His156, which transforms the ester **1** into the tetrahedral intermediate **2**, as also proposed for other serine hydrolases.

Geometry optimization at the QM/MM level leads to a react-ant conformation which resembles a possible Michaelis com-plex, with the ester oriented in such a way that facilitates the nucleophilic attack. The carbonyl oxygen is preoriented to enter the oxyanion hole formed by the backbone of Ile12 and Met78. We located a transition state connecting the Michaelis complex **1** and the tetrahedral intermediate **2** (Figure 1) and found a barrier of about 10 kcal/mol (for details see Supporting Information). The impact of a given mutation in the enzyme on the reaction barrier was estimated by setting to

zero the MM charges on the corresponding side chain and recalculation of the energy of the three structures (Michaelis–Menten complex, transition state and tetrahedral intermediate) corresponding to stationary points (Bash *et al.*, 1991; Dinner *et al.*, 2001). This procedure was repeated for all residues of the enzyme. In this sense, we performed an *in silico* electrostatic equivalent of an alanine scanning mutagenesis (Morrison and Weiss, 2001).

The calculations identified five amino acid positions that have a pronounced effect (>1 kcal/mol) on the reaction barrier (Figure 2). Four of these (Lys44, Asp43, Asp40 and Arg142) represent ionizable groups located on the protein surface. Shielding of these charges, e.g. by counterions from the

surrounding solution under physiological conditions, should diminish the influence of those residues. To test this hypothesis we added counterions close to the charged sites of the groups above and re-evaluated the barrier. We found that the contributions drop below 1 kcal/mol for each group and, consequently, we do not consider them as hot spots.

The remaining position identified in the QM/MM scan was residue His76, which is located below the active serine (Ser77) (Figure 1) and can form a hydrogen bond to the backbone oxygen of the active-site histidine (His156). This histidine is singly protonated in our model setup and its overall charge is therefore zero. Owing to its position, orientation and polarity, it may exert an important role during ester hydrolysis. According to the QM/MM calculations, the electronic effect of His76 is to raise the energy barrier for the nucleophilic attack.

Experimental results

Parallel to the *in silico* scanning experiments shown in Figure 2, we scanned experimentally the complete BSLA sequence space by saturation mutagenesis and performed an activity screening using enantiomerically pure (*R*)- and (*S*)-NEA as the model substrates. A total of 181 saturation mutagenesis experiments were carried out, covering all residues from Ala1 to Asn181. Accordingly, 181 mutagenesis primers were synthesized which contained one randomized codon (NNS; N = all nucleobases, S = guanine or cytosine) (see Supporting Information). Thus, a library consisting of 32 different mutant genes at every single codon position was generated using the mega-primer PCR method for site-specific mutagenesis (Barettino *et al.*, 1994; Funke *et al.*, 2003). Subsequent cloning of these genes into expression vectors and transformation into the expression host *Escherichia coli* BL21 (DE3) resulted in a library containing all possible single-site enzyme variants of BSLA, which represents a diversity of 5792 different mutant genes corresponding to 3439 variant proteins.

This library was screened for the enantioselective hydrolysis of NEA by using a high-throughput assay which allowed us to identify visually clones producing active lipases by clearing zones surrounding the bacterial colonies (Figure 3). *Escherichia coli* transformands expressing the BSLA saturation variants were plated out directly onto indicator plates which contained either (*R*)- or (*S*)-NEA as the substrate and the plates were incubated at 37°C for 48 h. A total of 21 000 clones were screened with 10 500 variants plated on each indicator medium containing one enantiomer, which represents a theoretical oversampling by a factor of three, thereby ensuring a complete coverage of the entire saturation library. As expected from previous experiments, about one-third of the BSLA variants were found to be inactive towards (*R*)-NEA, presumably because of deleterious mutations. The remaining variants were enzymatically active, as indicated by the formation of clearing zones of various sizes. In contrast, most of those colonies growing on agar plates with (*S*)-NEA as the substrate did not show any enzymatic activity, essentially as observed for wild-type BSLA. However, five colonies were identified which formed clear halos on (*S*)-NEA (Figure 3), indicating that they produced BSLA variants which had acquired the ability to convert the (*S*)-enantiomer of the substrate, indicating a changed enantioselectivity of the enzyme. All these variants were mapped at position His76, which had also been identified by
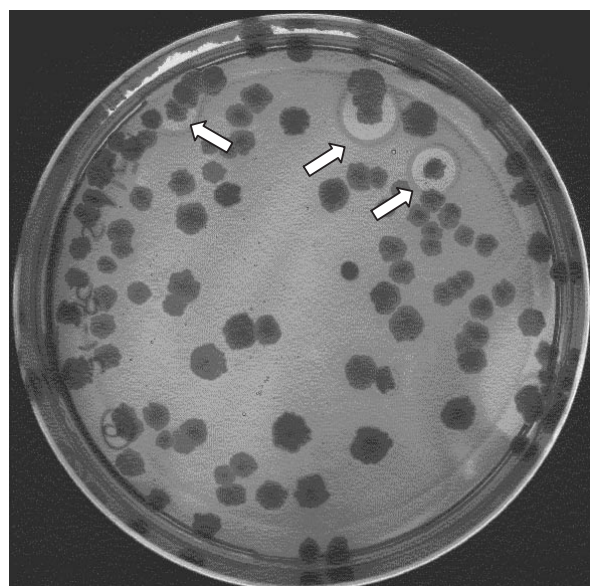
Fig. 3. High-throughput screening on indicator agar plates for the hydrolysis of (*S*)-NEA. Activity is indicated by clear halos surrounding the bacterial colonies (marked by arrows).

Table I. Enantioselectivities of *Bacillus subtilis* lipase variants identified by screening a complete saturation mutagenesis library for the hydrolysis of (*S*)-NEA

Variant No.	Substitution		*E* value
	Amino acid	Base	
wt	–	–	156 (*R*)
NEA1	His76Leu	CAC → CTG	6.8 (*R*)
NEA2	His76Ala	CAC → GCC	8.5 (*S*)
NEA3	His76Ala	CAC → GCG	n.d.[a]

Amino acid substitutions and base exchanges are given.
[a]n.d.: not determined owing to very low protein expression.

computational scanning. At least three different BSLA mutants were identified by DNA sequencing (Table I), named NEA1–NEA3, for further biochemical analysis.

The enzymatic activity and enantioselectivity of these BSLA variants were confirmed by chiral GC. Variant NEA1 (carrying the amino acid substitution His76Leu) showed a high conversion of (*S*)-NEA; however, it still hydrolyzed the (*R*)-enantiomer with an *E* value of 6.8. Interestingly, variant NEA2 (His76Ala) showed an inverted enantioselectivity of *ee* = 80% for the (*S*)-enantiomer of NEA, corresponding to an *E* value of 8.5 (Table I). Much to our surprise, we found hardly any activity for variant NEA3, although it contained the same amino acid exchange as variant NEA2, namely His76Ala. Closer inspection revealed that the two variants differed at the DNA level. In NEA2, the histidine to alanine exchange at position 76 was encoded by the codon 'GCC', whereas it was encoded by codon 'GCG' in NEA3. Consequently, we investigated the amounts of variant proteins produced by the respective clones using SDS–PAGE analysis and found that only a small amount of BSLA variant protein NEA3 was produced (Figure 4). This clearly indicates that codon usage can significantly influence the outcome of site saturation mutagenesis experiments. Previously published methods to saturate amino acid positions
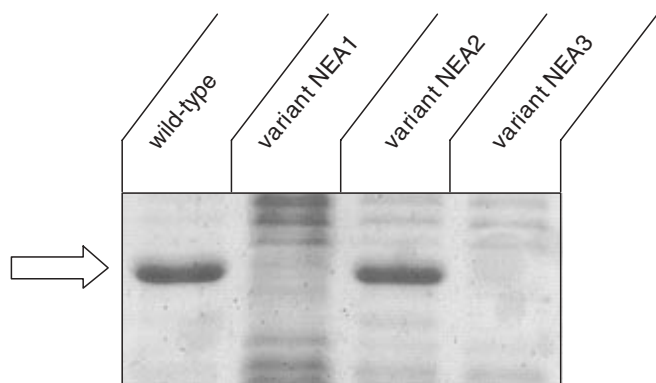
Fig. 4. SDS-PAGE to analyze the amount of wild-type (wt) or variant (NEA1–3) *Bacillus subtilis* lipase A (BSLA) produced by *E.coli*. The gel was stained with Coomassie Brilliant Blue; the band representing BSLA is marked by an arrow.

have used NNK- (DeSantis *et al.*, 2003), NNS- (Funke *et al.*, 2003) or MAX-codon mixtures (Hughes *et al.*, 2003), but our results clearly support saturation mutagenesis with substitution of a given codon triplet by all possible 64 codons (NNN).

Discussion and conclusion

Our computational procedure uses QM/MM-optimized geometries for the Michaelis complex, the transition state and the tetrahedral intermediate of the substrate (*R*)-NEA in the wild-type enzyme to estimate the electrostatic influence of each residue on the relevant barriers. This approach is motivated by the commonly accepted view that electrostatics is a dominant factor for biocatalytic activity in general (Warshel, 2003; Garcia-Viloca *et al.*, 2004); this should be a sensible assumption also in the present case since the calculated (QM/MM) transition state is a negatively charged oxyanion (see Supporting Information for details). The chosen procedure is fast enough for a qualitatitive prescreening of all residues in the enzyme.

It should be stressed that our simple electrostatic approach has several important limitations. First, it does not capture steric effects, which are assumed to be of minor importance; this is not strictly true, of course, as can already be seen from the present experimental result that the replacements His76Leu and His76Ala lead to different changes in enantioselectivity (see Table I). Second, it does not take into account the structural relaxations and rearrangements that occur after a mutation; it is intuitively clear that these will affect both activity and enantioselectivity and we have indeed confirmed in a recent molecular dynamics study (Bocola *et al.*, 2004) that such structural changes can rationalize remote and cooperative effects of mutations on the enantioselectivity observed for lipase-catalyzed ester hydrolysis in *Pseudomonas aerigunosa*. Third, our simple approach does not differentiate properly between (*R*)- and (*S*)-substrates: test calculations show that the electrostatic influence of His76 on the barriers is almost the same when using (*S*)-NEA rather than (*R*)-NEA as substrate in our procedure (–2.3 vs –2.1 kcal/mol, respectively), which implies that structural relaxations and possibly also non-electrostatic interactions need to be considered for proper prediction of enantioselectivity. Finally, entropic effects are also neglected.

A more quantitative theoretical modeling would involve the initial replacement of a given amino acid in the wild-type

structure followed by classical molecular dynamics runs to re-equilibrate the resulting mutant structure, which can already give detailed insight into the structural consequences of the mutation (Bocola *et al.*, 2004). QM/MM geometry optimizations of educt, transition state and product are then required to determine the barriers for a given mutant and substrate, while QM/MM molecular dynamics runs along the reaction path need to be performed to include entropic effects and derive free energy barriers (Ottosson *et al.*, 2001; Senn *et al.*, 2005). Following this protocol for all possible mutations and both enantiomeric substrates would, however, constitute an immense computational effort that is far beyond current capabilities.

Given this situation, we view our computational procedure as a simple and practical QM/MM-based tool that may identify promising sites of mutation by locating residues that exert a strong electrostatic influence on the computed barrier. A replacement of such a residue should then change the barrier appreciably and there should be a reasonable chance that this change may be different for the two enantiomeric substrates (more so than in cases where the barrier remains unaffected by the replacement). In this manner, promising sites of mutation to generate more enantioselective mutants may be suggested without actually addressing the demanding task of predicting enantioselectivities theoretically.

This strategy has been successful in the present case study. The QM/MM-based analysis shows strong electrostatic effects of His76 and experimental screening of the complete mutagenesis library of BSLA indicates a decisive role of this residue: only mutations involving His76 produce BSLA variants which convert the (*S*)-enantiomer of NEA and thus exhibit a changed enantioselectivity. These findings support the hypothesis that our simple QM/MM-based prescreening procedure may be applied as a tool to find amino acid positions important for enantioselectivity. This raises the prospect that enzyme optimization by directed evolution may be accelerated by the combination of computational prescreening and experimental library construction.

References

Ahlrichs,R., Bar,M., Haser,M., Horn,H. and Kolmel,C. (1989) *Chem. Phys. Lett.*, **162**, 165–169.

Barettino,D., Feigenbutz,M., Valcarcel,R. and Stunnenberg,H.G. (1994) *Nucleic Acids Res.*, **22**, 541–542.

Bash,P.A., Field,M.J., Davenport,R.C., Petrsko,G.A., Ringe,D. and Karplus,M. (1991) *Biochemistry*, **30**, 5826–5832.

Becke,A.D. (1988) *Phys. Rev. A*, **38**, 3098–3100.

Billeter,S.R., Turner,A.J. and Thiel,W. (2001) *Phys. Chem. Chem. Phys.*, **2**, 2177–2186.

Bocola,M., Otte,N., Jaeger,K.E., Reetz,M.T. and Thiel,W. (2004) *ChemBioChem*, **5**, 214–223.

Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) *J. Comput. Chem.*, **4**, 187–217.

Cherry,J.R. and Fidantsef,A.L. (2003) *Curr. Opin. Biotechnol.*, **14**, 438–443.

Clark,T., Chandrasekhar,J., Spitznagel,G.W. and Schleyer,P.V. (1983) *J. Comput. Chem.*, **4**, 294–301.

Cunningham,B.C. and Wells,J.A. (1989) *Science*, **244**, 1081–1085.

DeSantis,G. *et al.* (2003) *J. Am. Chem. Soc.*, **125**, 11476–11477.

Dinner,A., Blackburn,G.M. and Karplus,M. (2001) *Nature*, **413**, 752–755.

Eggert,T., Leggewie,C., Puls,M., Streit,W., van Pouderoyen,G., Dijkstra,B.W. and Jaeger,K.E. (2004a) *Biocatal. Biotrans.*, **22**, 139–144.

Eggert,T., Reetz,M.T. and Jaeger,K.E. (2004b) In Svendsen,A. (ed.), *Enzyme Functionality: Design, Engineering and Screening.* Marcel Dekker, New York, pp. 375–390.

Funke,S.A., Eipper,A., Reetz,M.T., Otte,N., Thiel,W., van Pouderoyen,G., Dijkstra,B.W., Jaeger,K.E. and Eggert,T. (2003) *Biocatal. Biotrans.*, **21**, 67–73.

Garcia-Viloca,M., Gao,J., Karplus,M. and Truhlar,D.G. (2004) *Science*, **303**, 186–195.

Gray,K.A., Richardson,T.H., Kretz,K., Short,J.M., Bartnek,F., Knowles,R., Kan,L., Swanson,P.E. and Robertson,D.E. (2001) *Adv. Synth. Catal.*, **343**, 607–617.

Hariharan,P.C. and Pople,J.A. (1973) *Theor. Chim. Acta*, **28**, 213–222.

Hehre,W.J., Ditchfield,R. and Pople,J.A. (1972) *J. Chem. Phys.*, **56**, 2257.

Hughes,M.D., Nagel,D.A., Santos,A.F., Sutherland,A.J. and Hine,A.V. (2003) *J. Mol. Biol.*, **331**, 973–979.

Jaeger,K.E. (2004) *Curr. Opin. Biotechnol.*, **15**, 269–271.

Jaeger,K.E. and Eggert,T. (2004) *Curr. Opin. Biotechnol.*, **15**, 305–313.

Kolattukudy,P.E., Purdy,R.E. and Maiti,I.B. (1981) *Methods Enzymol.*, **71**, 652–664.

Lee,C., Yang,W. and Parr,R.G. (1988) *Phys. Rev. B*, **37**, 785–789.

Liese,A., Seelbach,K. and Wandrey,C. (2000) *Industrial Biotransformations.* Wiley-VCH, Weinheim.

Lorenz,P., Liebeton,K., Niehaus,F. and Eck,J. (2002) *Curr. Opin. Biotechnol.*, **13**, 572–577.

MacKerell,A.D. *et al.* (1998) *J. Phys. Chem. B*, **102**, 3586–3616.

Morrison,K.L. and Weiss,G.A. (2001) *Curr. Opin. Chem. Biol.*, **5**, 302–307.

Ottosson,J., Rotticci-Mulder,J.C., Rotticci,D. and Hult,K. (2001) *Protein Sci.*, **10**, 1769–1774.

Panke,S., Held,M. and Wubbolts,M. (2004) *Curr. Opin. Biotechnol.*, **15**, 272–279.

Patel,R.N. (2003) *Curr. Opin. Drug Discov.*, **6**, 902–920.

Petrounia,I.P. and Arnold,F.H. (2000) *Curr. Opin. Biotechnol.*, **11**, 325–330.

Reetz,M.T. (2004) *Proc. Natl Acad. Sci. USA*, **101**, 5716–5722.

Reetz,M.T. and Jaeger,K.E. (2002) In Brakmann,S. and Johnsson,K. (eds), *Directed Molecular Evolution of Proteins.* Wiley-VCH, Weinheim, pp. 245–279.

Robertson,D.E. and Steer,B.A. (2004) *Curr. Opin. Chem. Biol.*, **8**, 141–149.

Senn,H.M., Thiel,S. and Thiel,W. (2005) *J. Chem. Theory Comput.*, **1**, 494–505.

Sherwood,P. *et al.* (2003) *Theochem*, **632**, 1–28.

Short,J.M. (2001) US Patent 6 171 820.

Smith,W. and Forester,T. (1996) *J. Mol. Graph.*, **14**, 136–141.

Streit,W.R., Daniel,R. and Jaeger,K.E. (2004) *Curr. Opin. Biotechnol.*, **15**, 285–290.

van Pouderoyen,G., Eggert,T., Jaeger,K.E. and Dijkstra,B.W. (2001) *J. Mol. Biol.*, **309**, 215–226.

Warshel,A. (2003) *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 425–443.

Weiss,G.A., Watanabe,C.K., Zhong,A., Goddard,A. and Sidhu,S.S. (2000) *Proc. Natl Acad. Sci. USA*, **97**, 8950–8954.

Supporting Information

Contents

A. Computational results

Key distances (Å) in BLYP/CHARMM optimised structure:
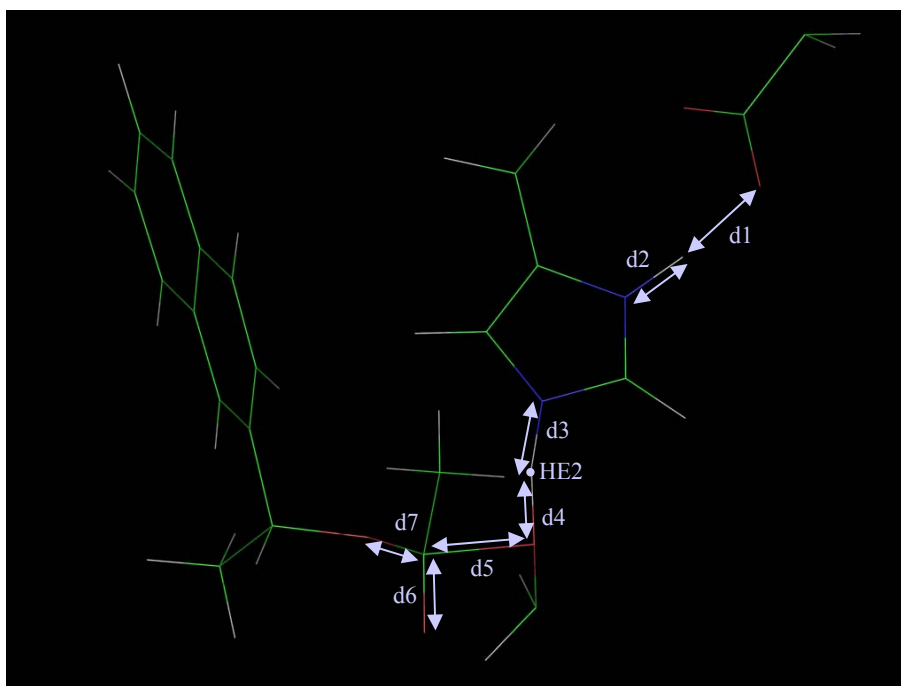
	Tetrahedral Intermediate (TI)	Transition State (TS)	Michaelis Complex (MC)
d1 (O-H)	1.56	1.59	1.66
d2 (N-H)	1.08	1.07	1.06
d3 (N-H)	1.06	1.22	1.70
d4 (O-H)	1.66	1.30	1.02
d5 (C-O)	1.60	1.94	2.62
d6 (C-O)	1.30	1.27	1.25
d7 (C-O)	1.50	1.43	1.37

BLYP/CHARMM reaction barriers in kcal/mol:

	TI → TS	*MC → TS*
QM	2.7	9.6
MM	-0.3	-3.9
QM/MM	2.4	5.7

Setup of the system. The model system was built from the crystal structure 1R50, monomer B. We replaced the inhibitor with a 1-(R)-β-naphthylethylacetate (NEA) substrate tetrahedral intermediate and added hydrogen atoms using the Charmm c28b2 program. The resulting structure was relaxed by a series of energy minimizations and molecular dynamics (MD) simulations described below using the CHARMM forcefield as implemented in the CHARMM program. No residue definitions are available for the tetrahedral intermediate in the standard topology files supplied with the CHARMM program. Therefore we used a set of custom parameters that we successfully applied in a previous study (M.Bocola et al., *ChemBioChem* 5, 214-223, **2004**).

MM calculations. The substrate was minimised (100 steps ABNR) and further relaxed in a dynamics run (1ps at 600K) while the rest of the enzyme was held fixed. Then we hydrated the active site with a preequilibrated water sphere (r =25Å) and minimised this additional volume (500 steps SD and 500 steps ABNR). Water molecules were always restrained by a quartic central potential when working at the MM-level to keep the shape of the volume spherical and to prevent evaporation of water in dynamics simulations. Thereafter we minimised (500 steps SD and 500 steps ABNR) the entire system, and ran a heating dynamics simulation (50000 steps) starting at 50 K and ending at 300 K. We rehydrated the system and repeated the minimisation and dynamics steps seven times. During these runs we used a set of fixed atoms which consisted of all residues beyond a spherical 12.5 Å cutoff, counted from the origin. The origin was defined to be at the initial position of the hydrogen atom HE2, shown in the drawing above. Harmonic positional restraints were applied to all non-substrate heavy atoms which had not been fixed. SHAKE was used in all dynamics calculations to constrain bonds to hydrogen atoms. Finally a 1.1 ns molecular dynamics simulation was run where all restraints on non-fixed atoms were released. We harvested snapshots for our QM/MM calculations from that trajectory. Our results are based on the coordinates at 900 ps. We energy minimised the system (500 steps SD, then 10000 steps ABNR), subjecting all residues within a sphere of 12.5 Å around the origin plus all water molecules within 9.0 Å of any atom of the substrate to the optimization. The same subset definition was applied in the QM/MM optimisations below.

QM/MM calculations. The entire system consists of 7658 atoms, the optimised region forms a subset of 1446 atoms. The finally adopted QM region (*QMA*) comprised 29 atoms, i.e. the sidechains of Ser77, His156 and Asp133 as well as the substrate in tetrahedral intermediate configuration (attached to Ser77) excluding the naphthyl ring. Open valencies at the QM/MM border were saturated using hydrogen link atoms. We used an electrostatic embedding scheme (D. Bakowies, W. Thiel, *J. Phys. Chem.*, 100, 10580-10594, **1996**), i.e. the fixed MM charges were introduced into the one-electron Hamiltonian of the QM calculation and the QM/MM electrostatic interactions were evaluated from the QM electrostatic potential and the MM atomic charges. No electrostatic cutoffs were used.

For three initial QM/MM refinement steps (described in the following) we chose a smaller QM region (*QMB*) where the aspartate was excluded. We energy minimised the system to ensure that a stable tetrahedral intermediate could

be located. This was done because the tetrahedral intermediate is known to be rather instable, but dissociation of covalently bound species cannot occur in the preceding MM calculations. In some cases we picked snapshots where the intermediate dissociated when switching to QM/MM methods (thus giving it the freedom to break down). From test calculations we knew the approximate location of the transition state in terms of a set of distances (d3,d4,d5 in the figure above). A harmonic restraint k(d4-d3-d5-s)^2 was applied with k set to 3 a.u and s to -1.8 Å, and a minimisation executed. We then used the transition state search algorithm implemented in the HDLC optimiser to further refine our structure. The reaction core was defined to consist of the atoms of d3, d4 and d5. A transition state was located and confirmed by a frequency analysis applied to all QM-atoms (in the field of the MM-atoms). We then switched to the larger QM-region (*QMA*) and relocated the transition state using the same approach as in the preceding step. Another frequency calculation was executed for confirmation. The normal mode vector was scaled by a factor of 0.75, and added to and subtracted from the transition state geometry. The resulting structures were already reminiscent of the Michaelis complex and the tetrahedral intermediate, indicating that the transition state connects these two species. The two structures were minimised again and yielded the Michaelis complex and the tetrahedral intermediate that were subsequently used in the perturbation calculation (and for geometry evaluation, see above).

Notes on QM/MM geometry optimisations. All QM/MM calculations were done with the HDLC (S.R. Billeter, A.J. Turner, W. Thiel, *Phys. Chem. Chem. Phys,* 2, 2177-2186, **2000**) optimiser implemented in the ChemShell program (*P.* Sherwood et al*., J. Mol. Struct.,* 632, 1-28, **2003**). We used hybrid delocalised coordinates in all optimisations. Two sets of convergence criteria were applied (see table below) for energy minimisations and transition state optimisations, respectively. In the latter case separate convergence criteria are used for the predefined reaction core (see above), and for the remainder (surrounding of the reaction core) of the optimised atoms.

Convergence criteria of the QM/MM geometry optimisations expressed in atomic units (angles in rad):

	Energy Minimisation	*TS optimisation*	
		Reaction core	Surrounding
maximum step component	0.0054	0.0054	1.0
RMS of step vector	0.0036	0.0036	1.0
maximum gradient component	0.00135	0.00135	0.00045
RMS of the gradient	0.0009	0.0009	0.0003

Notes on reaction barriers. All effects on barriers were evaluated between QM-energies, which were calculated in the field of point charges of the MM-atoms. For further validation, we recalculated barriers with B3LYP/CHARMM at the converged BLYP/CHARMM structures, and obtained 4.3 kcal/mol for TI → TS and 7.5 kcal/mol for MC → TS.

The results presented here refer to one particular snapshot (900ps) of the MD simulation. We have optimised the stationary points also for another snapshot. We found only minor fluctuations with regard to the key distances

presented above, the QM/MM barriers are 3.3 kcal/mol (TI → TS) and 8.3 kcal/mol (MC → TS) at the BLYP/CHARMM level.

B. Experimental Data

Oligonucleotides used for complete saturation mutagenesis:

name	DNA-sequence (5'→ 3')
mut1up	CTCCTCGCTGCCCAGCCGGCGATG
mutS3low	ATATAAGCTTCAGCAAACAGCTATGACCATGATTACGAATTC
A1X	CTCCTCGCTGCCCAGCCGGCGATGGCCATGNNSGAACACAATCCA
E2X	CTCCTCGCTGCCCAGCCGGCGATGGCCATGGCTNNSCACAATCCAGTC
H3X	CTCCTCGCTGCCCAGCCGGCGATGGCCATGGCTGAANNSAATCCAGTCGTT
N4X	CTCCTCGCTGCCCAGCCGGCGATGGCCATGGCTGAACACNNSCCAGTCGTTATG
P5X	CTCCTCGCTGCCCAGCCGGCGATGGCCATGGCTGAACACAATNNSGTCGTTATGGTTC
V6X	ACCGTGAACCATAACSNNTGGATTGTGTTCAGC
V7X	AATACCGTGAACCATSNNGACTGGATTGTGTTC
M8X	TCCAATACCGTGAACSNNAACGACTGGATTGTG
V9X	CCCTCCAATACCGTCSNNCATAACGACTGGATTG
H10X	TGCCCCTCCAATACCSNNAACCATAACGACTGG
G11X	TGATGCCCCTCCAATSNNGTGAACCATAACGAC
I12X	GAATGATGCCCCTCCSNNACCGTGAACCATAAC
G13X	ATTGAATGATGCCCCSNNAATACCGTGAACCATAAC
G14X	AAAATTGAATGATGCSNNTCCAATACCGTGAAC
A15X	CGCAAAATTGAATGASNNCCCTCCAATACCGTG
S16X	TCCCGCAAAATTGAASNNTGCCCCTCCAATACC
F17X	AATTCCCGCAAAATTSNNTGATGCCCCTCCAATAC
N18X	CTTAATTCCCGCAAASNNGAATGATGCCCCTCC
F19X	GCTCTTAATTCCCGCSNNATTGAATGATGCCCC
A20X	GCTCTTAATTCCSNNAAAATTGAATGATGC
G21X	ATAGCTCTTAATSNNCGCAAAATTGAATGATGC
I22X	TACGAGATAGCTCTTSNNTCCCGCAAAAAT
I23X	AGATACGAGATAGCTSNNAATTCCCGCAAAATTG
K24X	CTGAGATACGAGATASNNCTTAATTCCCGC
Y25X	GCCCTGAGATACGAGSNNGCTCTTAATTCCCGC
L26X	CCAGCCCTGAGATACSNNATAGCTCTTAATTCC
V27X	CGACCAGCCCTGAGASNNGAGATAGCTCTTAATTC
S28X	CCGCGACCAGCCCTGSNNTACGAGATAGCTC
Q29X	GTCCCGCGACCAGCCSNNAGATACGAGATAGCT
G30X	CTTGTCCCTCGACCASNNCTGAGATACGAGATAG
W31X	CAGCTTGTCCCGCGASNNGCCCTGAGATACGAG
S32X	ATACAGCTTGTCCCGSNNCCAGCCCTGAGATAC
R33X	TGCATACAGCTTGTCSNNCGACCAGCCCTGAG
D34X	AACTGCATACAGCTTSNNCCGCGACCAGCCCTG
K35X	ATCAACTGCATACAGSNNGTCCCGCGACCAGCC
L36X	AAAATCAACTGCATASNNCTTGTCCCGCGACCAG
Y37X	CCAAAAATCAACTGCSNNCAGCTTGTCCCGCGAC
A38X	GTCCCAAAAATCAACSNNATACAGCTTGTCCCG
V39X	CTTGTCCCAAAAATCSNNTGCATACAGCTTGTC
D40X	TGTCTTGTCCCAAAASNNAACTGCATACAGCTT
F41X	GCCTGTCTTGTCCCSNNATCAACTGCATACAG
W42X	TGTGCCTGTCTTGTCSNNAAAATCAACTGCATA
D43X	ATTTGTGCCTGTCTTSNNCCAAAAATCAACTGC
K44X	ATAATTTGTGCCTGTSNNGTCCCAAAAATCAAC
T45X	GTTATAATTTGTGCCSNNCTTGTCCCAAAAATC
G46X	ATTGTTATAATTTGTGSNNTGTCTTGTCCCAAAAATC
T47X	TCCATTGTTATAATTSNNGCCTGTCTTGTCCCAAAAATC

4

```
N48X        CGGTCCATTGTTATASNNTGTGCCTGTCTTGTC
Y49X        TACCGGTCCATTGTTSNNATTTTGTGCCTGTCTT
N50X        TAATACCGGTCCATTSNNATAATTTGTGCCTGT
N51X        TGATAATACCGGTCCSNNGTTATAATTTGTGCC
G52X        TCGTGATAATACCGGSNNATTGTTATAATTTGTG
P53X        AAATCGTGATAATACSNNTCCATTGTTATAATT
V54X        CACAAATCGTGATAASNNCGGTCCATTGTTATA
L55X        TTGCACAAATCGTGASNNTACCGGTCCATTG
S56X        CTTTTGCACAAATCGSNNTAATACCGGTCCATTG
R57X        AACCTTTTGCACAAASNNTGATAATACCGGTCC
F58X        TAAAACCTTTTGCACSNNTCGTGATAATACCGG
V59X        ATCTAAAACCTTTTGSNNAAATCGTGATAATAC
Q60X        TTCATCTAAAACCTTSNNCACAAATCGATAA
K61X        CGTTCCATCTAAAACSNNTTGCACAAATCGTG
V62X        ACCCGTTTCATCTAASNNCTTTTGCACAAATCG
L63X        CGCACCCGTTTCATCSNNAACCTTTTGCACAAATC
D64X        TTTCGCACCCGTTTCSNNTAAAACCTTTTGCAC
E65X        TTTTTTCGCACCCGTSNNATCTAAAACCTTTTG
T66X        CACTTTTTTCGCACCSNNTTCATCTAAAACC
G67X        ATCCACTTTTTTCGCSNNCGTTTCATCTAAAAC
A68X        AATATCCACTTTTTTSNNACCCGTTTCATCTAAAAC
K69X        GACAATATCCACTTTSNNCGCACCCGTTTCATC
K70X        AGCGACAATATCCACSNNTTTCGCACCCGTTTC
V71X        GTGAGCGACAATATCSNNTTTTTTCGCACCCGT
D72X        GCTGTGAGCGACAATSNNCACTTTTTTCGCACC
I73X        CATGCTGTGAGCGACSNNATCCACTTTTTTCGC
V74X        CCCCATGCTGTGAGCSNNAATATCCACTTTTTTC
A75X        GCCCCCCATGCTGTGSNNGACAATATCCACTTTTTTC
H76X        CGCGCCCCCCATGCTSNNAGCGACAATATCCAC
S77X        GTTCGCGCCCCCCATSNNGTGAGCGACAATATC
M78X        TGTGTTCGCGCCCCCSNNGCTGTGAGCGACAATATC
G79X        AAGTGTGTTCGCGCCSNNCATGCTGTGAGCGAC
G80X        GTAAAGTGTGTTCGCSNNCCCCATGCTGTGAGC
A81X        GTAGTAAAGTGTGTTSNNGCCCCCCATGCTGTG
N82X        TATGTAGTAAAGTGTSNNCGCGCCCCCCATGCTG
T83X        TTTTATGTAGTAAAGSNNGTTCGCGCCCCCCATG
L84X        ATTTTTTATGTAGTASNNTATGTTCGCGCCCCC
Y85X        CAGATTTTTTATGTASNNAAGTGTGTTCGCGCC
Y86X        GTCCAGATTTTTTATSNNGTAAAGTGTGTTCGC
I87X        GCCGTCCAGATTTTTSNNGTAGTAAAGTGTGTTC
K88X        TCCGCCGTCCAGATTSNNTATGTAGTAAAGTGTG
N89X        ATTTCCGCCGTCCAGSNNTTTTATGTAGTAAAG
L90X        TTTATTTCCGCCGTCSNNATTTTTTATGTAGTAAAG
D91X        AACTTTATTTCCGCCSNNCAGATTTTTTATGTAG
G92X        TGCAACTTTATTTCCSNNGTCCAGATTTTTTATG
G93X        GTTTGCAACTTTATTSNNGCCGTCCAGATTTTTTATG
N94X        GACGTTTGCAACTTTSNNTCCGCCGTCCAGATTTTTTATG
K95X        CACGACGTTTGCAACSNNATTTCCGCCGTCCAG
V96X        CGTCACGACGTTTGCSNNTTTATTTCCGCCGTC
A97X        AAGCGTCACGACGTTSNNAACTTTATTCCTGCC
N98X        GCCAAGCGTCACGACSNNTGCAACTTTATTTCC
V99X        GCCGCCAAGCGTCACSNNGTTTGCAACTTTATTTC
V100X       CGCGCCGCCAAGCGTSNNGACGTTTGCAAC
T101X       GTTCGCGCCGCCAAGSNNCACGACGTTTGCAAC
L102X       ACGGTTCGCGCCGCCSNNCGTCACGACGTTTGC
G103X       TAAACGGTTCGCGCCSNNAAGCGTCACGACGTT
G104X       CGTTAAACGGTTCGCSNNGCCAAGCGTCACGAC
A105X       TGTCGTTAAACGGTTSNNGCCGCCAAGCGTCAC
N106X       GCCTGTCGTTAAACGSNNCGCGCCGCCAAGCGTC
R107X       CTTGCCTGTCGTCAASNNGTTCGCGCCGCCAGG
L108X       CGCCTTGCCTGTCGTSNNACGGTTCGCGCCGCC
T109X       AAGCGCCTTGCCTGTSNNCAAACGGTTCGCGCC
T110X       CCGAAGCGCCTTGCCSNNCGTCAAACGGTTCGC
```

5

```
G111X       TCCCGAAGCGCCTTSNNTGTCGTCAAACGGTTC
K112X       TGTTCCCCGAAGCGCSNNGCCTGTCGTCAAACG
A113X       ATCTGTTCCCGGAAGSNNCTTGCCTGTCGTTAAAC
L114X       TGGATCTGTTCCCGGSNNCGCCTTGCCTGTCG
P115X       ATTTGGATCTGTTCCSNNAAGCGCCTTGCCTGTC
G116X       TTGATTTGGATCTGTSNNCGGAAGCGCCTTGCC
T117X       CTTTTGATTTGGATCSNNTCCCGGAAGCGCCTTG
D118X       AATCTTTTGATTTGGSNNTGTTCCCGGAAGCGC
P119X       TAAAATCTTTTGATTSNNATCTGTTCCCGGAAG
N120X       GTATAAAATCTTTTGSNNTGGATCTGTTCCCGG
Q121X       TGTGTATAAAATCTTSNNATTTGGATCTGTTCC
K122X       GGATGTGTATAAAATSNNTTGATTTGGATCTG
I123X       AATGGATGTGTATAASNNCTTTTGATTTGGATC
L124X       GTAAATGGATGTGTASNNATTCTTTTGATTTGG
Y125X       GCTGTAAATGGATGTSNNTAAAATCTTTTGATTTG
T126X       ACTGCTGTAAATGGASNNGTATAAAATCTTTTG
S127X       GGCACTGCTGTAAATSNNTGTGTATAAAATCTT
I128X       ATCGGCACTGCTGTASNNGGATGTGTATAAAATC
Y129X       CATATCGGCACTGCTSNNAATGGATGTGTATAAAATC
S130X       AATCATATCGGCACTSNNGTAAATGGATGTG
S131X       GACAATCATATCGGCSNNGCTGTAAATGGATG
A132X       CATGACAATCATATCSNNACTGCTGTAAATGG
D133X       ATTCATGACAATCATSNNGGCACTGCTGTAAATG
M134X       GTAATTCATGACAATSNNATCGGCACTGCTGTAAATG
I135X       TAAGTAATTCATGACSNNCATATCGGCACTGCTG
V136X       TGATAAGTAATTCATSNNAATCATATCGGCACTG
M137X       TCTTGATAAGTAATTSNNGACAATCATATCGGC
N138X       TAATCTTGATAAGTASNNCATGACAATCATATC
Y139X       ATCTAATCTTGATAASNNATTCATGACAATCATATCGGC
L140X       ACCATCTAATCTTGASNNGTAATTCATGACAATC
S141X       AGCACCATCTAATCTSNNTAAGTAATTCATGAC
R142X       TCTAGCACCATCTAASNNTGATAAGTAATTCATG
L143X       GTTTCTAGCACCATCSNNTCTTGATAAGTAATTC
144X        AACGTTTCTAGCACCSNNTAATCTTGATAAGTAATT
G145X       TTGAACGTTTCTAGCSNNATCTAATCTTGATAAG
A146X       GATTTGAACGTTTCTSNNACCATCTAATCTTG
R147X       ATGGATTTGAACGTTSNNAGCACCATCTAATCTTG
N148X       GCCATGGATTTGAACSNNTCTAGCACCATCTAATC
V149X       AACGCCATCGATTTGSNNGTTTCTAGCACCATC
Q150X       TCCAACGCCATGGATSNNAACGTTTCTAGCACC
I151X       GTGTCCAACGCCATGSNNTTGAACGTTTCTAGC
H152X       GATGTGTCCAACGCCSNNGATTTGAACGTTTCTAG
G153X       GCCGATGTGTCCAACSNNATGGATTTGAACGTTTC
V154X       AAGGCCGATGTGTCCSNNGCCATGGATTTGAAC
G155X       CAGAAGGCCGATGTGSNNAACGCCATGGATTTG
H156X       GTACAGAAGGCCGATSNNTCCAACGCCATGGATTTG
I157X       GCTGTACAGAAGGCCSNNGTGTCCAACGCCATG
G158X       GCTGCTGTACAGAAGSNNGATGTGTCCAACGCC
L159X       TTGGCTGCTGTACAGSNNGCCGATGTGTCCAAC
L160X       GACTTGGCTGCTGTASNNAAGGCCGATGTGTCC
Y161X       GTTGACTTGGCTGCTSNNCAGAAGGCCGATGTG
S162X       GCTGTTGACTTGGCTSNNGTACAGAAGGCCGATG
S163X       CAGGCTGTTGACTTGSNNGCTGTACAGAAGGCC
Q164X       AATCAGGCTGTTGACSNNGCTGCTGTACAGAAG
V165X       TTTAATCAGGCTGTTSNNTTGGCTGCTGTACAG
N166X       TTCTTTAATCAGGCTSNNGACTTGGCTGCTGTAC
S167X       CCCTCCTTTAATCAGSNNGTTGACTTGGCTGCTG
L168X       CAGCCCTTCTTTAATSNNGCTGTTGACTTGGCTGCTG
I169X       GTTCAGCCCTTCTTTSNNCAGGCTGTTGACTTG
K170X       GCCGTTCAGCCCTTCSNNAATCAGGCTGTTGAC
E171X       CCCGCCGTTCAGCCCSNNTTTAATCAGGCTGTT
G172X       ATTCGTATTCTCGCCCCCGCCGTTCAGSNNTTC
L173X       ATTCGTATTCTCGCCCCCGCCGTTSNNCCCTTC
```

6

```
N174X      ATTCGTATTCTCGCCCCCGCCSNNCAGCCCTTC
G175X      ATTCGTATTCTCGCCCCCSNNGTTCAGCCCTTC
G176X      ATTCGTATTCTCGCCSNNGCCGTTCAGCCCTTC
G177X      ATTCGTATTCTCSNNCCCGTTCAGCCCTTC
Q178X      GCCGCAAGCTTGTCGACGGAGCTCATtCGTATTSNNGCC
N179X      GCCGCAAGCTTGTCGACGGAGCTCATTCGTSNNCTC
T180X      GCCGCAAGCTTGTCGACGGAGCTCATTSNNATT
N181X      GCCGCAAGCTTGTCGACGGACTCSNNCGT
```

Multiplex-PCR-Based Recombination as a Novel High-Fidelity Method for Directed Evolution

Thorsten Eggert,*[a] Susanne Aileen Funke,[a] Nalam M. Rao,[b] Priyamvada Acharya,[b] Holger Krumm,[c] Manfred T. Reetz,[c] and Karl-Erich Jaeger[a]

A new and convenient method for the in vitro recombination of single point mutations is presented. This method efficiently reduces the introduction of novel point mutations, which usually occur during recombination processes. A multiplex polymerase chain reaction (multiplex-PCR) generates gene fragments that contain preformed point mutations. These fragments are subsequently assembled into full-length genes by a recombination-PCR step. The process of multiplex-PCR-based recombination (MUPREC) does not require DNase I digestion for gene-fragmentation and is therefore easy to perform, even with small amounts of target DNA. The protocol yields high frequencies of recombina-

tion without creating a wild-type background. Furthermore, the low error rate results in high-quality variant libraries of true recombinants, thereby minimizing the screening efforts and saving time and money. The MUPREC method was used in the directed evolution of a Bacillus subtilis *lipase that can catalyse the enantioselective hydrolysis of a model* meso-*compound. Thereby, the method was proved to be useful in producing a reliable second-generation library of true recombinants from which better performing variants were identified by using a high-throughput electrospray ionization mass spectrometry (ESI-MS) screening system.*

Introduction

Directed evolution has matured during the last decade to become a key technology in the field of molecular enzyme engineering, in particular, when neither the 3D structures nor the catalytic mechanisms of the enzymes are known. However, even if crystal structures are available and reaction mechanisms are well understood, directed evolution often provides alternative solutions in comparison to rational-design experiments.[1–3]

The creation of diversity is a crucial step in each directed-evolution experiment. Diversity can either be directly retrieved from nature by isolation of homologous but not identical genes or artificially generated by introducing random point mutations into a target gene. Moreover, subsequent recombination of this diversity has proved to be a very effective strategy for combining advantageous mutations and separating out deleterious ones. Today, at least twelve in vitro recombination methods have been published, which are summarized in two excellent review articles.[2,4] Among these approaches DNA-shuffling is still the method of choice for most directed-evolution experiments. Other methods, which include staggered extension process (StEP),[5] random priming recombination (RPR),[6] heteroduplex recombination,[7] ssDNA-family shuffling,[8] degenerate oligonucleotide gene shuffling (DOGS),[9] random chimeragenesis on transient templates (RACHITT),[10] mutagenic and unidirectional reassembly (MURA),[11] synthetic shuffling,[12] assembly of designed oligonucleotides (ADO)[13] and recombined extension on truncated templates (RETT)[14] use different experimental strategies to ensure the exchange of DNA fragments between different variants. Slight variations in these methods have been published by different groups.[15–17] All these methods result in a significant improve-

ment in the efficiency to create novel enzymes by directed evolution. However, they also have major drawbacks, including i) a recombinational bias depending on the target DNA and ii) the creation of additional diversity by introducing novel point mutations during recombination, a process that could result in a library far too large to be screened by available methods. Therefore, novel developments in directed-evolution methodology focus on improving library quality instead of quantity.[2]

A major strategy to reduce the size of a library is based on increasing the fidelity of the recombination process. The original DNA-shuffling protocol led to the introduction of an average of seven novel point mutations per kilobase (kb), which results in extra diversity.[18,19] This effect is favoured when screening capacity is not a limiting factor as it is for powerful selection systems like phage display[20] or fluorescence-activated cell sorting (FACS).[21] Unfortunately, such systems are not available as yet to select for enzyme properties like enantioselectivity. Zhao and Arnold modified the DNA-shuffling protocol to reduce the rate of newly introduced point mutations by using

[a] Dr. T. Eggert, Dr. S. A. Funke, Prof. Dr. K.-E. Jaeger
Institut für Molekulare Enzymtechnologie
Heinrich-Heine-Universität Düsseldorf, Forschungszentrum Jülich
52426 Jülich (Germany)
Fax: (+49) 2461-61-2490
E-mail: t.eggert@fz-juelich.de

[b] Dr. N. M. Rao, Dr. P. Acharya
Centre for Cellular and Molecular Biology
Uppal Road, Hyderabad 500007 (India)

[c] Dr. H. Krumm, Prof. Dr. M. T. Reetz
Max-Planck Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr (Germany)

different DNA polymerases during fragment reassembly.[22] The creation of DNA fragments by using restriction endonucleases also reduced the number of novel point mutations; however, it also increased the bias of recombination.[8] Nevertheless, the methods based on DNase I digestion have in common the facts that large amounts of DNA are needed and that the frequency of recombination is very low for neighbouring mutations.

Here, we describe a high-fidelity method for the recombination of point mutations that introduces a single, novel point mutation per 10 kb (mutation rate 1×10^{-4}) but results in a high frequency of recombination independent of the amino-acid positions to be recombined. Furthermore, the protocol for this multiplex-PCR-based recombination method is simple and generally applicable. The versatility of this method was tested by the recombination of point mutations that had been introduced into the *Bacillus subtilis* lipase A (BSLA) gene and subsequent screening for the enantioselective hydrolytic desymmetrization of a model *meso*-compound.

Results and Discussion

Multiplex-PCR-based recombination (MUPREC)

A protocol was developed for the efficient recombination of single point mutations that are generated by directed evolution methods. This protocol is based on multiplex-PCR for the amplification of those fragments that carry point mutations for recombination.

As a starting point we used two multiplex-PCR reactions that were performed simultaneously by using two different template plasmids of the target gene. In one reaction, a set of mutagenesis primers that were designed as lower primers were applied together with a universal upper primer (Table 1). This resulted in the formation of a mixture of different megaprimers, each containing a single point mutation. The other multiplex-PCR reaction produces the complementary megaprimers by applying a set of mutagenesis primers that are designed as upper primers and amplify gene fragments along with a universal lower primer. In a third PCR reaction, these megaprimers were used together with the flanking primers (mut1-up and mutS-low; Table 1) to produce the full-length gene that carried the desired point mutations. The high concentration of megaprimers in comparison to flanking primers resulted in megaprimer overlaps and subsequent elongations, which led to the random recombination of the desired point mutations (Figure 1). Theoretically, template switching occurs during megaprimer-annealing and -elongation processes and results in the formation of all possible combinations of point mutations. In practice, however, we have observed an accumulation of mutants that carry two or three point mutations (data not shown). Fortunately, these recombinants were randomly generated, nevertheless, recombinants with more than four point mutations were relatively rare.

Table 1. Oligonucleotides used in this study.

Primer	Sequence[a]	Modifications
mut1-up	5′-ctcctcgctgcccagccggcgatggccatg-3′	*Mls*I
mutS-low	5′-atataagcttcagcaaacagctatgaccatgattacgaattc-3′	*Hind*III
N18X-up	5′-ggaggggcatcatcatt<u>nns</u>tttgcgggaattaag-3′	N18 saturation primer
N18X-low	5′-cttaattcccgcaaa<u>snn</u>gaatgatgatgcccctcc-3′	N18 saturation primer
I22T-up	5′-ttcaattttgcgggga<u>act</u>aagagctatctcg-3′	I22T mutation
I22T-low	5′-cgagatagctctt<u>agt</u>tcccgcaaaattg-3′	I22T mutation
Y49C-up	5′-aagacaggcacaaatt<u>gt</u>aacaatggaccggta-3′	Y49C mutation
Y49C-low	5′-taccggtccattgtt<u>aca</u>atttgtgcctgtctt-3′	Y49C mutation
Y49I-up	5′-aagacaggcacaaat<u>atc</u>aacaatggaccggta-3′	Y49I mutation
Y49I-low	5′-taccggtccattgtt<u>gat</u>atttgtgcctgtctt-3′	Y49I mutation
Y49V-up	5′-aagacaggcacaaat<u>gtc</u>aacaatggaccggta-3′	Y49V mutation
Y49V-low	5′-taccggtccattgtt<u>gac</u>atttgtgcctgtctt-3′	Y49V mutation
N50S-up	5′-acaggcacaaattat<u>agc</u>aatggaccggtattatc-3′	D50S mutation
N50S-low	5′-taataccggtccatt<u>gct</u>ataatttgtgcctgt-3′	N50S mutation
F58L-up	5′-ccggtattatcacga<u>ctt</u>gtgcaaaaggttttag-3′	F58L mutation
F58L-low	5′-taaaaccttttgcaca<u>aag</u>tcgtgataataccgg-3′	F58L mutation
Q60L-up	5′-ttatcacgatttgt<u>gtt</u>gaaggttttagatgaa-3′	Q60L mutation
Q60L-low	5′-catctaaaaccttc<u>aac</u>acacaaatcgtgataa-3′	Q60L mutation
Q60N-up	5′-ttatcacgatttgt<u>gaa</u>caaggttttagatg-3′	Q60N mutation
Q60N-low	5′-catctaaaaccttc<u>gtt</u>cacaaatcgtgataa-3′	Q60N mutation
L114P-up	5′-acgacaggcaaggc<u>gcc</u>tccgggaacagatcc-3′	L114P mutation
L114P-low	5′-tggatctgttcccgga<u>ggc</u>gccttgcctgtcg-3′	L114P mutation
C124S-up	5′-ccaaatcaaaagatt<u>tca</u>tacacatccatttac-3′	C124S mutation
C124S-low	5′-gtaaatggatgtgta<u>tga</u>aatcttttgatttgg-3′	C124S mutation
A132D-up	5′-tccatttacagcagt<u>gac</u>gatatgattgtcatg-3′	A132D mutation
A132D-low	5′-catgacaatcatatc<u>gtc</u>actgctgtaaatgg-3′	A132D mutation
I157N-up	5′-caaatccatggcgttgga<u>gac</u>gccttctgtacagc-3′	I157N mutation
I157N-low	5′-gctgtacagaaggcc<u>gtc</u>tccaacgccatggatttg-3′	I157N mutation
N166Y-up	5′-tacagcagccaagtc<u>tac</u>agcctgattaaagaag-3′	N166Y mutation
N166Y-low	5′-ttctttaatcaggct<u>gta</u>gacttggctgctgtac-3′	N166Y mutation

[a] Mutated codon underlined.

This result forced us to change the protocol so that a larger number of small fragments would be formed in the first multiplex-PCR reaction. Here, a single, universal, upper primer (mut1-up) was used together with sense and antisense mutagenesis primers, which again produced megaprimer fragments. In addition, small fragments that were complementary to the middle of the gene and carried two point mutations were amplified. This fragment mixture together with a new gene template, which has different flanking regions, was then used along with a universal lower primer in a second so-called recombination-PCR reaction (Figure 2). The modification of the original protocol (Figure 1) resulted in the formation of an increased number of recombinants that carried multiple point mutations. This protocol is recommended for five or more point mutations that are to be recombined. The results shown in Table 2 indicate that up to eleven point mutations can be randomly recombined. This suggests that this method could also allow the recombination of an even higher number of point mutations. The limitation of this method is set solely by the size of the created library. As an example, the recombination of 20 point mutations will generate a library consisting of 1.05×10^6 different variants (see formula in Table 3), a number that exceeds the capacity of most high-throughput screening methods.

The correct formation of DNA fragments during a multiplex-PCR reaction, as shown by agarose gel electrophoresis (Figure 3), indicated that the protocol could be further simpli-
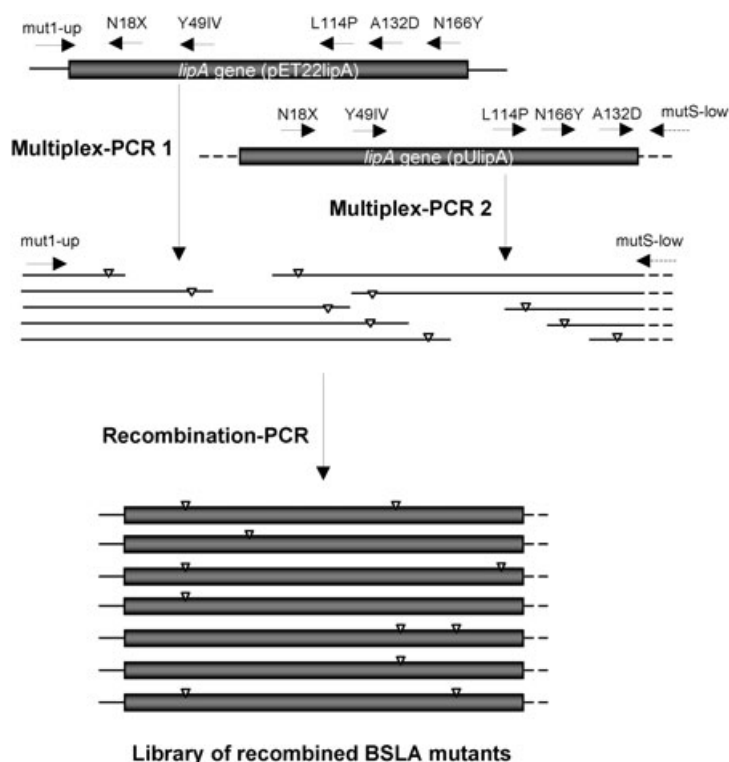
Figure 1. Initial experimental approach for the directed recombination of existing point mutations. This protocol comprised three independent PCRs and led to recombinant genes that contained combinations of up to three mutations at maximum. Triangles indicate point mutations.

Table 3. Theoretical library sizes generated by randomly recombined point mutations.	
Number of single point mutations[a]	Number of recombinants[b]
2	4
3	8
4	16
5	32
6	64
7	128
8	256
9	512
10	1024

[a] Number of single point mutations to be recombined. [b] Number of true recombinants without any new point mutation, calculated by using the formula $\sum_{k=0}^{n}\binom{n}{k}=2^n$ where n = number of single point mutations to be recombined and k = overall number of point mutations present in a variant protein.

fied by performing just a single PCR reaction. In this procedure (an all-in-one PCR), the amplification of wild-type sequences was excluded by using a universal upper primer, mut1-up, that only hybridized to pET22 lipA1 (template 1) and a universal lower primer, mutS-low, that only hybridized to pUlipA (template 2; Figure 2). We call these primers "universal" due to their sequence independence towards the gene to be mutated. Primer mut1-up ("universal" upper primer) hybridizes to the vector sequence of pET22b upstream of the target gene, whereas mutS-low ("universal" lower primer) anneals down-

Table 2. Mutations and amino-acid substitutions found in randomly chosen variants[a] that were generated by the MUPREC process.	
Variant no.	Mutation[b]
	Recombination of point-mutations I22T (att-act), Y49C (tat-tgt), N50S (aac-agc), F58L (ttt-ctt), Q60N (caa-aac), Q60L (caa-ttg), L114P (ctt-cct), C124S (tta-tca), A132D (gcc-gac), I157N (atc-gac), N166Y (aac-tac)
1	I22T, Y49C, N50S, L114P, C124S
2	I22T, N50S, L114P, C124S, **S24I (agc-atg)**
3	I22T, N50S
4	N50S, F58L, L114P
5	I22T, N50S, Q60N, L114P, C124S, I157N
6	I22T, Y49C, F58L, L114P
7	Y49C, L114P

[a] All variants showed lipolytic activity towards the substrate tributyrin. [b] Recombined point mutations are given as amino-acid exchanges; newly generated point mutations are indicated in bold; base substitutions are written in brackets.

stream of the target gene in the pUC18 vector. By using the two "universal" primers together with two different template vectors, the amplification of unwanted wild-type DNA is efficiently prevented for site-directed mutagenesis by using the megaprimer PCR technique.[23–25]

It should be noted that the mutagenesis and the "universal" primers must be carefully designed so that they have comparable melting temperatures, and equimolar concentrations of primers should be used in the reaction mixture. If a specific mutation is to be favoured, then the corresponding mutagenesis primer can be used in higher molar concentrations and will therefore be incorporated into the respective fragments at a statistically higher rate.

Application of MUPREC to evolve enantioselective lipase variants

Extracellular BSLA was optimized by directed evolution so that it catalyzed the enantioselective hydrolytic desymmetrization of 1,4-diacetoxycyclopentene (Scheme 1). Variant libraries were generated by error-prone PCR (epPCR), and by complete saturation mutagenesis.[26] During this project, we observed that several newly isolated enantioselective lipase variants showed a reduced thermostability. Therefore, we chose to recombine several mutations that lead to higher enantioselectivity with others previously shown to increase the thermostability of BSLA.[27] L114P, A132D and N166Y amino-acid substitutions were chosen for recombination as they resulted in increased thermostability, and several different substitutions at positions N18 and Y49 were chosen since they all resulted in increased enantioselectivity.[28] For position N18, a primer mix was used that encoded all 20 amino acids, and at position 49, the substitutions Y49I and Y49V were chosen. These had been identified during previous screenings (data not shown).
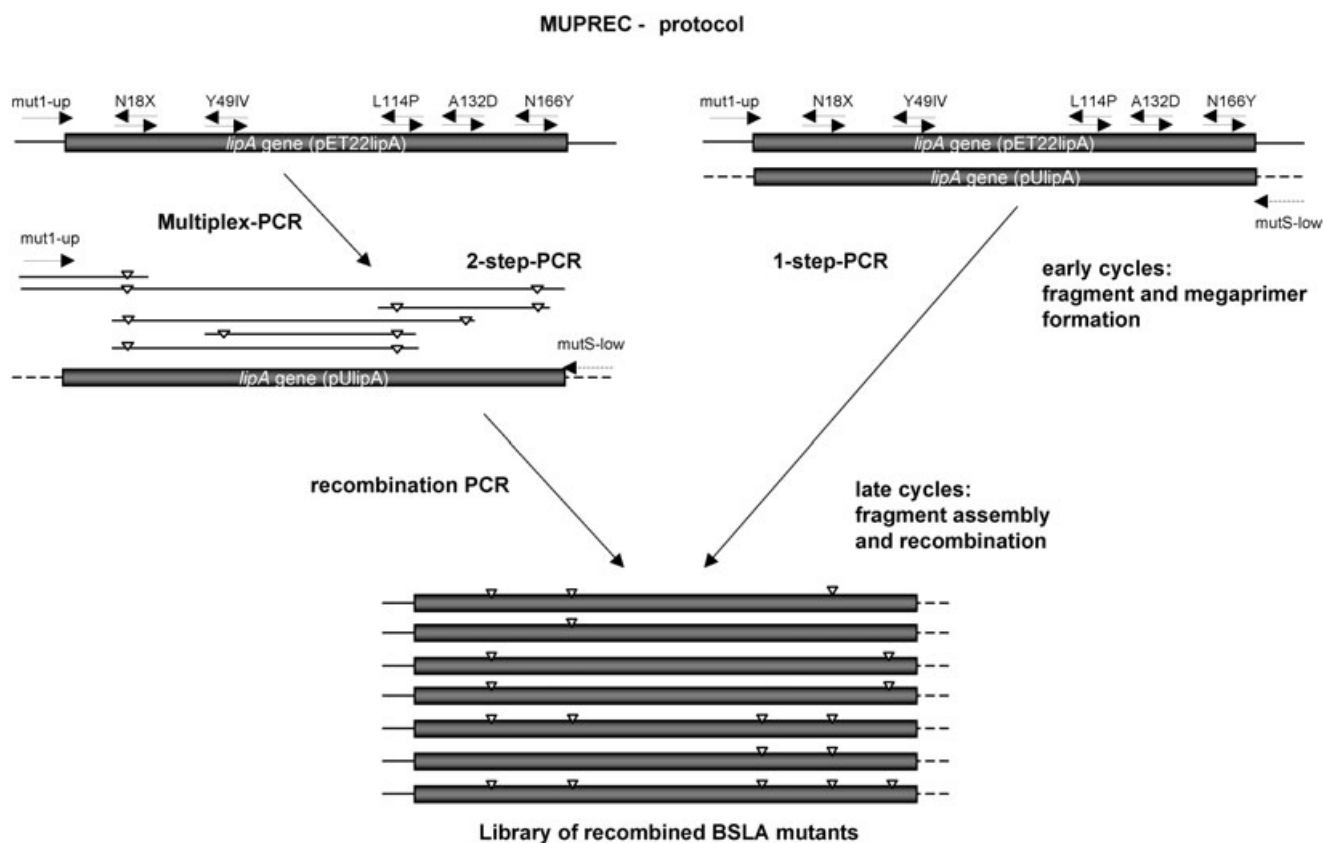
Figure 2. The MUPREC process. Mixtures of upper and lower primer pairs that carry the point mutations to be recombined are used in a multiplex PCR to amplify gene fragments, which are recombined in a second PCR. The efficiencies of fragment formation during the multiplex-PCR, which are mainly determined by the melting temperature of the mutagenesis primers, can be directly monitored by using the two-step method. Alternatively, the one-step protocol can be applied for convenient and high-fidelity recombination. More experimental details are given in the text. Triangles indicate point mutations.
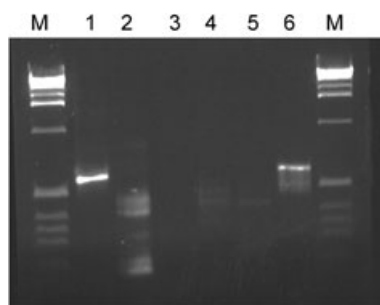
Figure 3. Gel electrophoretic analysis of a MUPREC experiment. Lane 1: Full-length BSLA gene amplified by standard PCR (positive control); lane 2: fragment mixture after multiplex PCR that contained megaprimer and internal PCR fragments that carried the point mutations to be recombined; lane 3: negative control for the recombination-PCR reaction using pUlipA as the template along with primers mut1-up and mutS-low; lane 4: negative control for the recombination PCR with pUlipA as the template and the fragment mixture shown in lane 2, but omitting the universal primers mut1-up and mutS-low; lane 5: negative control for the recombination PCR by using the fragment mixture shown in lane 2 and the universal primers mut1-up and mutS-low, but omitting the template pUlipA; lane 6: full-length PCR product after the recombination PCR.

The two-step MUPREC protocol described above (Figure 2, left) was used to monitor whether the correct formation of fragments and megaprimers occurred after the multiplex-PCR

reaction or not (Figure 3). Afterwards, the one-step MUPREC protocol (Figure 2, right) was used to essentially yield the same results. The amplified full-length genes were cloned into the expression vector pET22b by using the unique restriction sites *Mls*I and *Hin*dIII, which were introduced into the fragments during the recombination-PCR reaction. After over-expression in *E. coli*, a library of about 390 enzymatically active lipase variants was created and screened for enantioselectivity by using ESI-MS.[29] Nine BSLA variants were identified that showed inverse enantioselectivities to the wild-type enzyme (Table 4) and of which variant 37-01-G5 (N18Q, Y49V) also showed a much higher enzymatic activity when grown on tributyrin-indicator plates (Figure 4). Interestingly, for all variants, the increase in enantioselectivity was accompanied by a decrease in thermostability, although amino-acid substitutions were incorporated that were previously shown to increase the thermostability of BSLA. At present, the number of screened variants is still too low to conclude that a general incompatibility exists for combining thermostability and enantioselectivity in this lipase.

The efficiency of the MUPREC method

The efficiency of the MUPREC method was analysed by determination of the DNA sequences from randomly chosen re-

[D]₃-AcO────OAc → [D]₃-AcO────OH + OH────OAc (over BSLA H₂O)

| pseudo-*meso*-1,4-diacetoxy-cyclopentene | 1*R*,4*S*-enantiomer | 1*S*,4*R*-enantiomer |
| **(1)** | **(2)** | **(3)** |
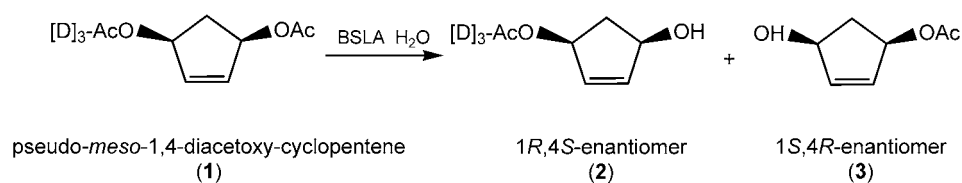
Scheme 1. The model reaction used to identify enantioselective variants of BSLA. The asymmetric hydrolysis of the model compound *meso*-1,4 diacetoxy-2-cyclopentene, was determined by a high-throughput ESI-MS screening system. The deuterium-labelled substrate pseudo-*meso*-1,4 diacetoxy-2-cyclopentene allows the formation of chiral alcohol products (**2**) and (**3**) to be identified by their mass differences.

combinants that showed lipase activity. All randomly chosen clones had acquired up to six different point mutations by recombination (Table 2); this indicates a high diversity of the library (Table 3). In total, 39 sequences of randomly chosen recombinants have been sequenced; this gives an estimated error rate of 1×10^{-4} for the MUPREC method, which is comparable to the rate observed for *Taq*-DNA-polymerases. Our results clearly show that MUPREC allows the directed recombination of previously identified point mutations without introducing a significant number of novel and possibly unwanted mutations, in contrast to other homology-dependent recombination protocols like DNA-shuffling (seven additional point mutations per kb) or StEP (0.6 additional point mutations per kb). Therefore, the application of MUPREC could help to significantly facilitate screening efforts. Furthermore, MUPREC did not produce any wild-type genes, as determined by screening for enantioselectivity or by DNA-sequencing.

Table 4. BSLA variants with improved enantioselectivity			
Variant	amino-acid exchanges	*ee* [%][a]	conversion [%][a]
wild-type	–	45 (1*R*,4*S*)	100
thermostable	L114P, A132D, N166Y	52 (1*R*,4*S*)	100
First generation (complete saturation mutagenesis library)			
144-F7	N18I	14 (1*S*,4*R*)	90
133A6	N18A	21 (1*S*,4*R*)	100
195-E8	N18L	65 (1*S*,4*R*)	75
22-N18C	N18C	72 (1*S*,4*R*)	85
145-F4	N18Q	82 (1*S*,4*R*)	75
133-H12	N18S	83 (1*S*,4*R*)	50
196-C2	Y49I	16 (1*S*,4*R*)	5
Second generation (MUPREC library)			
16-02-D1	N18L, N166Y	68 (1*S*,4*R*)	n.d.
16-02-B1	N18S, L114P, N166Y	23 (1*S*,4*R*)	n.d.
37-02-F2	N18H, N166Y	61 (1*S*,4*R*)	n.d.
37-02-B12	N18Q, Y49V	82 (1*S*,4*R*)	85
16-02-F1	N18Q, L114P	85 (1*S*,4*R*)	n.d.
16-02-G1	N18Q, L114P, A132D; N166Y	85 (1*S*,4*R*)	n.d.
37-03-A3	N18S, Y49I, L114P	30 (1*S*,4*R*)	n.d.
37-01-G5	N18Q, Y49V	82 (1*S*,4*R*)	85
37-02-E2	N18Q, Y49I	87 (1*S*,4*R*)	n.d.

[a] Enantioselectivity and conversion rate of the substrate pseudo-*meso*-1,4 diacetoxy-2-cyclopentene were determined by ESI-MS. n.d. = not determined.

Conclusion

We have described here a novel in vitro recombination method for application in directed-evolution experiments. The MUPREC process can be used to recombine single point mutations previously generated in directed-evolution experiments. This method avoids the amplification of wild-type genes and effectively prevents the formation of novel base substitutions. Therefore, the size of a recombination library is minimized, thereby enabling a complete library screen. The method does not require DNase I digestion for gene fragmentation and can therefore be easily carried out in a single PCR-step. Thus, MUPREC should prove useful in optimizing directed-evolution protocols based on libraries created by epPCR or complete saturation mutagenesis.

Experimental Section

Bacterial strains and growth conditions: Plasmids were constructed and transformed into *E. coli* strains XL1-blue or DH5α. *E. coli* cells were grown overnight in Luria–Bertani (LB) medium (5 mL) in glass tubes at 37°C and in the presence of appropriate amounts of ampicillin (100 µg mL⁻¹). The heterologous expression of BSLA and its variants was performed with *E. coli* BL21(DE3) in the presence of carbenicillin (100 µg mL⁻¹; Serva, Heidelberg, Germany).

General DNA techniques and plasmids: Plasmid DNA was prepared by using the plasmid purification midi-kit from QIAGEN (Hilden, Germany). Genomic DNA from *B. subtilis* 168 (obtained from the *Bacillus* Genetic Stock Center, Ohio, USA) was prepared by using the DNeasy Tissue Kit (QIAGEN, Hilden, Germany). Recombinant DNA techniques were performed as described by Sambrook et al.[30] Restriction digestion reactions and ligations were performed with enzymes from Fermentas (St. Leon-Rot, Germany) under conditions recommended by the manufacturer.
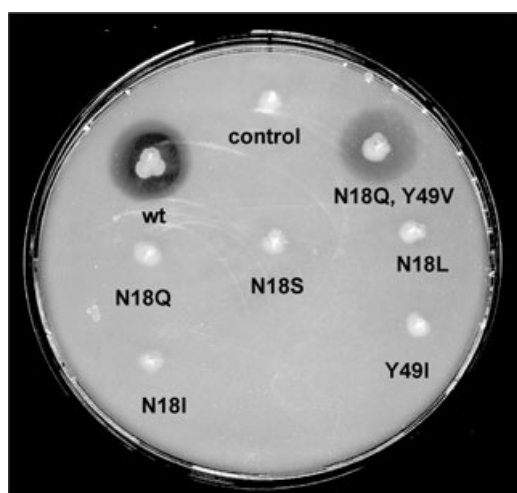

Figure 4. Lipolytic activities of wild-type BSLA and its enantioselective variants. *E. coli* clones expressing BSLA variants were plated on tributyrin-indicator plates and incubated for 24 h at 37°C. Variant 37-01-G5, which contains amino-acid substitutions N18Q and Y49V, shows wild-type (wt) activity indicated by the size of the clear halo around the colony.

The plasmids pET22 lipA or pUlipA were used as templates in PCR reactions. Both plasmids contain the BSLA gene with different up and downstream regions. The construction of the plasmid pET22 lipA has been described elsewere.[26] A BSLA-gene-containing plasmid with up- and downstream regions different from pET22 lipA was constructed by amplification of the lipase gene by using a standard PCR reaction with the 30 bp upper primer BSLA1 5′-ATAT-GATATCGCTGAACACAATCCAGTCGT-3′ and the 29 bp lower primer BSLA2 5′-TATAGAGCTCTCATTAATTCGTATTCTGG-3′. Genomic DNA from *B. subtilis* 168 (10 ng) was used as the template. The resulting 557 bp PCR product was cloned, blunt-end, into a *Hinc*II-digested pUC18 vector (Stratagene, Heidelberg, Germany) to result in plasmid pUlipA.

Standard-PCR conditions: Amplification of DNA fragments was performed in a 50 μL reaction mixture with plasmid (1 ng) or genomic DNA (10 ng) as template, primers (each 25 pmol), dNTPs (0.2 mm), *Taq* (2.5 U, Eurogentec, Seraing, Belgium) or *Pfu* polymerase (2.5 U, Stratagene, Heidelberg, Germany). Buffers containing $MgCl_2$ or $MgSO_4$ were used as recommended by the manufacturers. Conditions for PCR reactions were: 1× (3 min at 98 °C); 35× (1 min at 95 °C; 2 min at 58 °C, 1 min at 72 °C) and 1× (7 min at 72 °C). The PCR reactions were performed by using a Mastercycler Gradient (Eppendorf, Hamburg, Germany).

Multiplex-PCR conditions: Multiplex-PCR reactions were performed in 50 μL reaction mixtures, as described above for standard PCR, by using *Pfu* polymerase. The primers used for mutagenesis in this study are summarized in Table 1. To meet optimal annealing temperatures for every primer within the sample, the PCR conditions used were as follows: 1× (3 min at 98 °C); 35× (1 min at 95 °C; 2 min gradient from 70°–50°; 1 min at 72 °C) and 1× (7 min at 72 °C). The multiplex-PCR reactions were also performed by using a Mastercycler Gradient (Eppendorf, Hamburg, Germany). After identifying 65 °C to be the most efficient annealing temperature, we used this temperature in all following multiplex-PCR reactions: 1× (3 min at 98 °C); 35× (1 min at 95 °C; 2 min at 65 °C, 1 min at 72 °C) and 1× (7 min at 72 °C).

High-throughput screening for enantioselectivity: The recombined BSLA genes were cloned into the expression vector pET22b (Novagen, Madison, USA) as in-frame fusions to the *pelB*-signal sequence; this enables Sec-dependent protein secretion. The resulting plasmids were used to transform *E. coli* BL21(DE3) (Novagen, Madison, USA). The clones were cultured at 37 °C in 96-deep-well microtiter plates that were filled with LB/M9 medium (1 mL; 10 g L⁻¹ tryptone, 5 g L⁻¹ yeast extract, 5.5 g L⁻¹ NaCl, 4 g L⁻¹ glucose, 0.25 g L⁻¹ $MgSO_4 \cdot 7H_2O$, 0.02 g L⁻¹ $CaCl_2$, 7 g L⁻¹ $Na_2HPO_4 \cdot 2-H_2O$, 3 g L⁻¹ KH_2PO_4, 1 g L⁻¹ NH_4Cl) supplemented with carbenicillin (100 μg mL⁻¹). After 6 h of shaking at 37 °C ($OD_{580}=0.5$–0.7), lipase expression was induced by adding isopropyl-β-D-thio-galactopyranoside (final concentration 0.3 mm). The induced culture was grown at 37 °C, and the cells were separated from the medium by centrifugation at 5000 g for 10 min. An aliquot of 100 μL from the culture supernatant was taken from each well and pipetted into another 96-deep-well microtiter plate that contained $Na_2HPO_4/$ KH_2PO_4 buffer (800 μL; 10 mm, pH 7.5) and the substrate (100 μL) dissolved in dimethylsulfoxide (100 mm). After 24 h shaking at RT, the reaction solution was extracted with ethyl acetate and screened by electrospray ionization mass spectroscopy (ESI-MS).[29]

DNA sequence analysis: DNA sequence analysis of the mutant genes was performed by SequiServe (Vaterstetten, Germany) by using standard T7-promoter and T7-terminator primers.

[1] K.-E. Jaeger, T. Eggert, *Curr. Opin. Biotechnol.* **2004**, *15*, 305–313.
[2] S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291–297.
[3] M. T. Reetz, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5716–5722.
[4] C. Neylon, *Nucleic Acids Res.* **2004**, *32*, 1448–1459.
[5] H. Zhao, L. Giver, Z. Shao, J. A. Affholter, F. H. Arnold, *Nat. Biotechnol.* **1998**, *16*, 258–261.
[6] Z. Shao, H. Zhao, L. Giver, F. H. Arnold, *Nucleic Acids Res.* **1998**, *26*, 681–683.
[7] A. A. Volkov, Z. Shao, F. H. Arnold, *Nucleic Acids Res.* **1999**, *27*, e18.
[8] M. Kikuchi, K. Ohnishi, S. Harayama, *Gene* **1999**, *236*, 159–167.
[9] M. D. Gibbs, K. M. Nevalainen, P. L. Bergquist, *Gene* **2001**, *271*, 13–20.
[10] W. M. Coco, W. E. Levinson, M. J. Crist, H. J. Hektor, A. Darzins, P. T. Pienkos, C. H. Squires, D. J. Monticello, *Nat. Biotechnol.* **2001**, *19*, 354–359.
[11] J. K. Song, B. Chung, Y. H. Oh, J. S. Rhee, *Appl. Environ. Microbiol.* **2002**, *68*, 6146–6151.
[12] J. E. Ness, S. Kim, A. Gottman, R. Pak, A. Krebber, T. V. Borchert, S. Govindarajan, E. C. Mundorff, J. Minshull, *Nat. Biotechnol.* **2002**, *20*, 1251–1255.
[13] D. X. Zha, A. Eipper, M. T. Reetz, *ChemBioChem* **2003**, *4*, 34–39.
[14] S. H. Lee, E. J. Ryu, M. J. Kang, E. S. Wang, Z. Piao, Y. J. Choi, K. H. Jung, J. Y. J. Jeon, Y. C. Shin, *J. Mol. Catal. B* **2003**, *26*, 119–129.
[15] K. Miyazaki, *Nucleic Acids Res.* **2002**, *30*, e139.
[16] M. Kikuchi, K. Ohnishi, S. Harayama, *Gene* **2000**, *243*, 133–137.
[17] V. Abecassis, D. Pompon, G. Truan, *Nucleic Acids Res.* **2000**, *28*, E88.
[18] W. P. C. Stemmer, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747–10751.
[19] W. P. C. Stemmer, *Nature* **1994**, *370*, 389–391.
[20] A. Fernandez-Gacio, M. Uguen, J. Fastrez, *Trends Biotechnol.* **2003**, *21*, 408–414.
[21] S. Becker, H. U. Schmoldt, T. M. Adams, S. Wilhelm, H. Kolmar, *Curr. Opin. Biotechnol.* **2004**, *15*, 323–329.
[22] H. Zhao, F. H. Arnold, *Nucleic Acids Res.* **1997**, *25*, 1307–1308.
[23] D. Barettino, M. Feigenbutz, R. Valcarcel, H. G. Stunnenberg, *Nucleic Acids Res.* **1994**, *22*, 541–542.
[24] V. Picard, E. Ersdal-Badju, A. Lu, S. C. Bock, *Nucleic Acids Res.* **1994**, *22*, 2587–2591.
[25] A. Urban, S. Neukirchen, K.-E. Jaeger, *Nucleic Acids Res.* **1997**, *25*, 2227–2228.
[26] S. A. Funke, A. Eipper, M. T. Reetz, N. Otte, W. Thiel, G. Van Pouderoyen, B. W. Dijkstra, K.-E. Jaeger, T. Eggert, *Biocatal. Biotransform.* **2003**, *21*, 67–73.
[27] P. Acharya, E. Rajakumara, R. Sankaranarayanan, N. M. Rao, *J. Mol. Biol.* **2004**, *341*, 1271–1281.
[28] S. A. Funke, A. Eipper, H. Krumm, M. T. Reetz, K.-E. Jaeger, T. Eggert, unpublished results.
[29] M. T. Reetz, M. H. Becker, H. W. Klein, D. Stöckigt, *Angew. Chem.* **1999**, *111*, 1872–1875; *Angew. Chem. Int. Ed.* **1999**, *38*, 1758–1761.
[30] J. Sambrook, E. F. Fritsch, T. Maniatis, *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, New York, **1989**.

Taylor & Francis
Taylor & Francis Group

Novel Biocatalysts by Identification and Design

THORSTEN EGGERT[1], CHRISTIAN LEGGEWIE[1], MICHAEL PULS[1], WOLFGANG STREIT[2], GERTIE VAN POUDEROYEN[3], BAUKE W. DIJKSTRA[3] and KARL-ERICH JAEGER[1]*

[1] *Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Düsseldorf, Forschungszentrum Jülich, Stetternicher Forst, D-52426 Jülich, Germany;* [2] *Institut für Grenzflächen-Biotechnologie, Universität Duisburg-Essen, Geibelstr. 41, D-47057 Duisburg, Germany;* [3] *Laboratory of Biophysical Chemistry, Rijksuniversiteit Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands*

Enzymes produced from bacteria and eukaryotic organisms are presently being used for a large variety of different biotechnological applications. The rapidly increasing demand for enzymes which are active towards novel and often non-natural substrates has triggered the development of novel molecular biological methods of enzyme isolation and design. The metagenome approach is a cultivation-independent method which allows the direct cloning and expression of environmental DNA thereby providing access to a wealth of so-far unknown biocatalysts. Additionally, newly identified or existing biocatalysts can be further optimized by different methods of directed evolution. Here, the principle of the metagenome approach is outlined and a strategy is presented for the optimization of a bacterial lipase using a combination of rational design and directed evolution.

Keywords: Bacillus subtilis lipase; biodiversity; directed evolution; metagenome

INTRODUCTION

More than five thousand years ago, fermentation was discovered as a process for the production of alcohol. Also, without having any knowledge of the existence of enzymes and microorganisms, the Egyptian civilisation used yeast for baking bread, a technique which later became known as whole cell biocatalysis (Liese *et al.*, 2000). A breakthrough for using enzymes to catalyze chemical reactions occured in 1858, when Louis Pasteur succeeded in separating the $(+)$ and $(-)$ enantiomers of racemic tartaric acid using the fungus *Penicillium glaucum*. This experiment constitutes the first successful biocatalytic kinetic resolution (Pasteur, 1858).

Nowadays, biocatalysis using whole cells, crude cell extracts or purified enzymes has achieved a position of steadily increasing importance for the biotechnological production of food additives, agrochemicals, cosmetics and flavours, and, in particular, for pharmaceuticals. The rapidly growing demand for these compounds results in a pressing need to identify biocatalysts with novel and desired properties. Therefore, extended programs aim to collect novel microorganisms, plants or animals from all over the world to use them as a source for the identification of novel enzymes. However, natural evolution has adjusted today's enzymes to perfectly fit into their respective physiological niches. As a consequence, their biochemical properties like stability, activity, and enantioselectivity normally do not fulfill the needs of a chemical process.

In order to overcome these difficulties, a repertoire of tools for enzyme engineering was developed with most of them operating in a rational way. Firstly, the 3D-structure of a given enzyme is solved to allow the identification of important amino acids. Then, enzyme variants are constructed by classical site-directed mutagenesis based on predictions derived from the analysis of the 3D-structure and finally, these variants are biochemically characterized. More recently, a set of new methods was developed which are summarized as *"in vitro"* or "directed" evolution and provide a powerful tool for the creation of novel biocatalysts whithout requiring any knowledge of the enzyme structure or its catalytic mechanism (Cherry and Fidantsef, 2003; Farinas *et al.*, 2001; Jaeger *et al.*, 2001; Powell *et al.*, 2001; Tao and Cornish, 2002).

In this article we will describe promising novel approaches to (i) identify biocatalyst genes and (ii) optimize an enzyme by directed evolution.

* Corresponding author. Tel.: $+$ 49-2461-61-3716. Fax: $+$ 49-2461-61-2490. E-mail: karl-erich.jaeger@fz-juelich.de

THE METAGENOME APPROACH FOR MINING NEW BIOCATALYST GENES

Starting about 3.8 billion years ago, natural evolution has created prokaryotes capable of coping with adverse living conditions and inhabiting almost every ecological niche by metabolizing virtually all known substrates. Therefore, it is not surprising that prokaryotes became the dominant form of life and may represent the largest constituent of total biomass on earth (Whitman *et al.*, 1998). Despite their abundance, current estimates indicate that more than 99% of the prokaryotes present in natural environments like soil, water, sediments, or plant surfaces are not readily culturable in the laboratory by standard techniques and therefore remain inaccessible for biotechnological applications (Amann *et al.*, 1995). This fact is illustrated by the observation that one gram of soil may contain thousands of different species, however, only about 10–100 different species will be caught by known cultivation methods (Torsvik *et al.*, 2002). Obviously, these uncultivated bacteria not only significantly contribute to the ecology of a bacterial community (Torsvik and Ovreas, 2002) but they also represent an enormous biotechnological potential (Cowan, 2000). The so-called metagenome approach (Handelsman *et al.*, 1998) can overcome the cultivation problem by direct isolation and cloning of environmental DNA (eDNA) resulting in metagenome libraries which represent the genomes of all microorganisms present in a given sample independent of their culturability. The microbial diversity present in such a library can be analysed by 16S-rRNA sequencing and these libraries also constitute the starting material to identify novel biocatalyst-encoding genes by using high-throughput screening or selection methods (Fig. 1).

Nevertheless, several problems exist in constructing metagenome libraries. Isolated eDNA derived from soil-samples may be contaminated with phenolic compounds or humic acids which inhibit the following cloning steps, namely the digestion catalyzed by restriction endonucleases, the ligation, or the eDNA-amplification using the polymerase chain reaction (PCR). Several promising approaches were described to circumvent these problems (Rochelle, 2001). Different isolation strategies are suitable to recover the spectrum of prokaryotic diversity present in a respective sample: (1) The *ex situ* methods in which cells are isolated and concentrated from soil prior to their lysis, and (2) the *in situ* methods in which cells are lysed directly within the soil material (Courtois *et al.*, 2001). (3) A third strategy uses the enrichment of microorganisms for a desired enzyme activity prior to the isolation of metagenomic DNA. This strategy proved successful for isolating complete biotin operons and many different biocatalyst-encoding genes (Borneman, 1999; Entcheva *et al.*, 2001; Radajewski *et al.*, 2000; Schmeisser *et al.*, 2003; Voget *et al.*, 2003). (4) A new and elegant approach starts with an amplification of specific partial gene sequences using conserved and degenerate oligonucleotides called metagenome sequence tags (MST's). Subsequently, shuffling of the cloned fragments and PCR-amplification generates biocatalyst genes of increased diversity as shown for dehalogenases and haloperoxidases (Lorenz *et al.*, 2002). Undoubtedly, the metagenome approach will quickly generate an enormous amount of novel enzyme genes, however, novel tools have to be developed to increase the efficiency of cloning and expression.

DIRECTED EVOLUTION TO ENGINEER NOVEL BIOCATALYSTS

Many different enzymes have been subjected to optimization by directed evolution including proteases, amylases, laccases, phytases, and cellulases (Cherry and Fidantsef, 2003). Substrate specificity, thermal stability, and organic solvent resistance, but also more difficult properties such as cofactor-independence or enantioselectivity were evolved (Funke *et al.*, 2003; Lingen *et al.*, 2003; May *et al.*, 2000; Moore and Arnold, 1996; Wong *et al.*, 2004, Zhao and Arnold, 1999). In our group, a variety of different directed evolution methods have been used to evolve lipases which represent the most important class of enzymes for organic chemistry (Jaeger and Eggert, 2002; Jaeger and Reetz, 1998). In particular, we have extensively characterized the bacterial lipases from *Pseudomonas aeruginosa* and *Bacillus subtilis* and have studied their structure-function relationships by site-directed (Eggert *et al.*, 2000; Liebeton *et al.*, 2001) and random mutagenesis methods (Funke *et al.*, 2003; Liebeton *et al.*, 2000; Reetz *et al.*, 1997).

Lipases have been studied for many years, but there is still a debate over a general definition for a lipase and whether structural features can be identified in lipases which govern the lipase reaction or determine its specificity (Chahinian *et al.*, 2002b; Verger, 1997). In contrast to esterases, lipases show almost no activity as long as the substrate is present in the monomeric state. However, when the solubility limit of the substrate is exceeded and an emulsion is formed, a sharp increase in lipase activity occurs, a phenomenon termed interfacial activation (Sarda and Desnuelle, 1958). The first lipase 3D structures which were solved for human pancreatic and a fungal lipase (Brady *et al.*, 1990; Winkler *et al.*, 1990) provided an elegant explanation for this phenomenon. Both lipases were found to
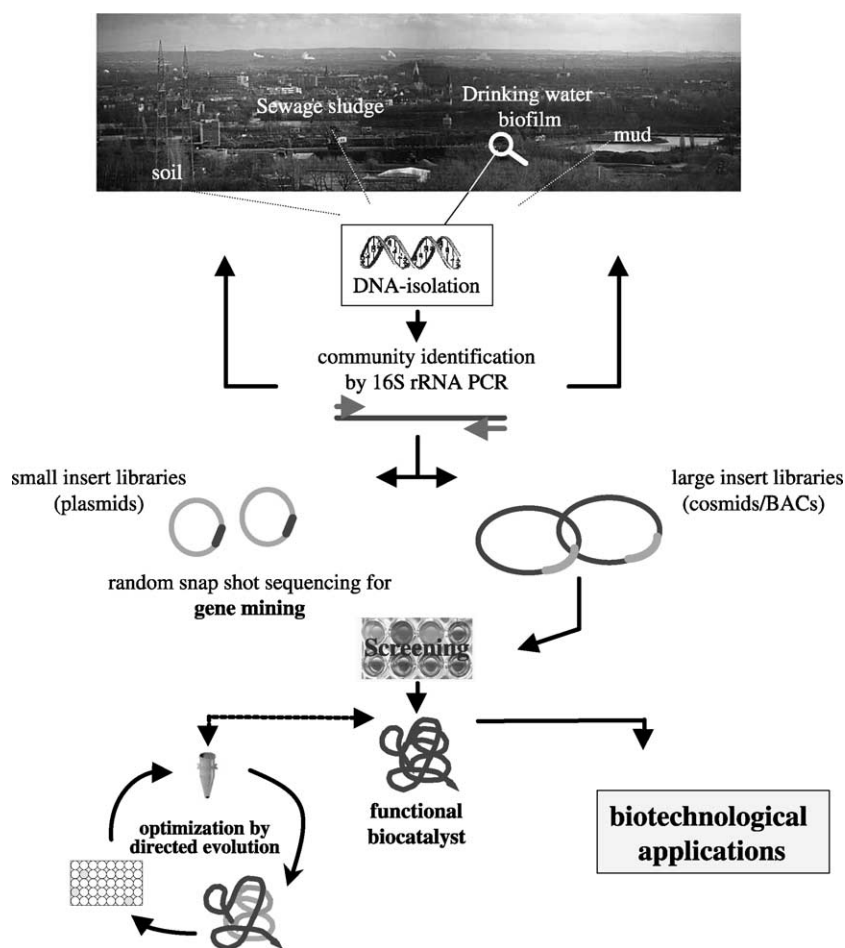
FIGURE 1 Schematic outline of the metagenome approach to isolate novel biocatalysts. DNA representing the collective genomes of a given habitat is PCR-amplified for prospecting biodiversity by 16S rRNA sequencing and cloned into appropriate vectors used for identification by DNA sequencing, expression cloning, and subsequent determination of enzyme activities, e.g. by high-throughput screening, the novel biocatalysts may further be optimized by directed evolution.

possess α-helical polypeptide chains capable of covering the active site thereby making the catalytic amino acid residues inaccessible to substrate molecules. The α-helices which were called "lid"-structures would move upon contact with the substrate interface resulting in conformational reorientations which render the active site residues accessible for the substrate. This "lid-hypothesis" was later confirmed by elucidating the X-ray structures of several other inhibitor-bound lipases (Brzozowski *et al.*, 1991; Nardini *et al.*, 2000) and has since been used to discriminate between 'true' lipases and esterases (Jaeger *et al.*, 1999; Verger, 1997). In addition, the lid-domain of lipases has also been identified as being important for substrate recognition, catalytic activity, substrate specificity, (Carriere *et al.*, 1998; Chahinian *et al.*, 2002a; Brocca *et al.*, 2003), enantio-selectivity (Liebeton *et al.*, 2000), and activity in organic solvents (Fishman and Cogan, 2003; Mingarro *et al.*, 1995). In summary, lid-like structural elements constitute the most important structural elements of lipases.

During the last few years, three-dimensional structures of 26 lipases, including 8 lipases from bacterial origin, have been elucidated (Table I). Surprisingly, not all of these structures revealed the presence of a lid-domain. The lipases from *Bacillus subtilis* (Dartois *et al.*, 1992; Eggert *et al.*, 2000; Lesuisse *et al.*, 1993) are catalytically active on long chain lipids although they do not possess a lid-domain. *B. subtilis* lipase A (BSLA) is the smallest lipase (181 amino acids; 19.3 kDa) known so far which has a solvent-exposed active site located at the bottom of a small cleft between two loops consisting of residues 10–15 and 131–137. The BSLA active site contains a preformed oxyanion hole so that conformational changes are not required to ensure the formation of the transition state. Thus, its small size and the absence of a lid make BSLA the minimal α/β-hydrolase-fold enzyme (Fig. 2) (van Pouderoyen *et al.*, 2001). Therefore, we have chosen this enzyme as a starting point to engineer by directed evolution new lid structures on top of the existing minimal α/β-hydrolase core.

TABLE I Overview of solved three-dimensional structures of lipases. Lipases lacking a lid are marked with an asterisk.

Eucaryotic	
Mammals/pancreatic lipases	*Homo sapiens* (human)
	Equus caballus (horse)
	Cavia porcellus (guinea pig)
	Sus scrofa (pig)
	Bos taurus (cattle)
	Canis familiaris (dog)
	Rattus norvegicus (rat)
Mammals/gastric lipases	*Homo sapiens* (human)
	Canis familiaris (dog)
Fungi	*Rhizomucor* miehei
	Geotrichum candidum
	*Fusarium solani
	Candida rugosa
	Penicillium camembertii
	Rhizopus delemar
	Thermomyces lanuginosa
	Candida antarctica
	Rhizopus niveus
Procaryotic	
Gram-negative	*Burkholderia glumae*
	Chromobacterium viscosum
	Burkholderia cepacia
	Pseudomonas aeruginosa
Gram-positive	*Streptomyces exfoliatus*
	*Bacillus subtilis
	Bacillus stearothermophilus L1
	Bacillus stearothermophilus P1

Construction of *B. subtilis* Lipase with an Artificial Lid-domain

In computer-based studies the lipolytic enzymes cutinase from *Fusarium solani pisi*, acetylxylanesterase from *Penicillium purpurogenum*, and human pancreatic lipase were compared with respect to the domains located close to their active sites. These enzymes were chosen because their three-dimensional structures

show a high homology to the structure of BSLA. Furthermore, lids or lid-like-domains were present in all three X-ray structures but are missing in the *Bacillus* lipase. Subsequently, these lids were modeled into the structure of BSLA indicating experimental options to engineer these lids into BSLA without disturbing the core α/β-hydrolase fold (Fig. 3).

BSLA-variants Carrying an Artificial Lid-domain Show Hydrolytic Activity

The lid-sequences were engineered into the BSLA-gene by using a modified two-step megaprimer PCR method. The resulting BSLA-variants (Fig. 3) were overexpressed in *E. coli* and tested for enzymatic activity with the substrates *p*-nitrophenyl palmitate (spectrophotometric assay, Fig. 4A) and tributyrin (agar plate assay, Fig. 4B). All variants showed enzymatic activity, at least against one of the substrates, although the variants containing artificial lid domains exhibited a lower activity than the wild-type. Therefore, we are currently trying to optimize the artificial lid domains by subjecting the corresponding DNA-fragments to random mutagenesis using ep-PCR. Preliminary results with a high-throughput screening assay on tributyrin indicator plates indicated that we have created several new variants which exhibit wild-type activities.

OUTLOOK

Enzyme-based biocatalysis provides a means to carry out chemical processes efficiently and economically. This fact is increasingly recognized as reflected by a rapidly growing enzyme market which was valued at approximately $1.5 billion already in 2000 (Cherry and Fidantsef, 2003) and is expected to increase by an average annual growth rate of at least 10%. The future success of enzyme technology will depend on the development of efficient and cost-effective processes for the production and downstream processing of enzymes. Even more important will be the identification of novel enzyme genes from natural sources, their high-level and functional expression, as well as their optimization for desired properties by directed evolution. The combination of these methods will undoubtedly result in a major breakthrough for enzyme technology, also in entirely new areas of technical applications.
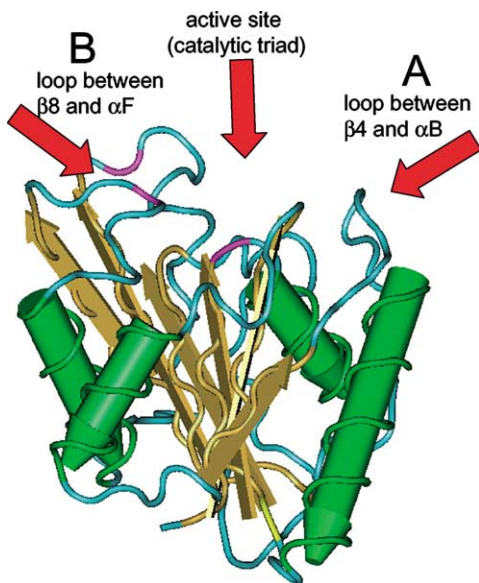
FIGURE 2 Crystal structure of *B. subtilis* lipase A (BSLA). The positions of catalytic triad residues Ser77, His156 and Asp133 are indicated in purple, and the two loop regions surrounding the active site are labelled with A (β4 and αB) and B (β8 and αF), respectively.
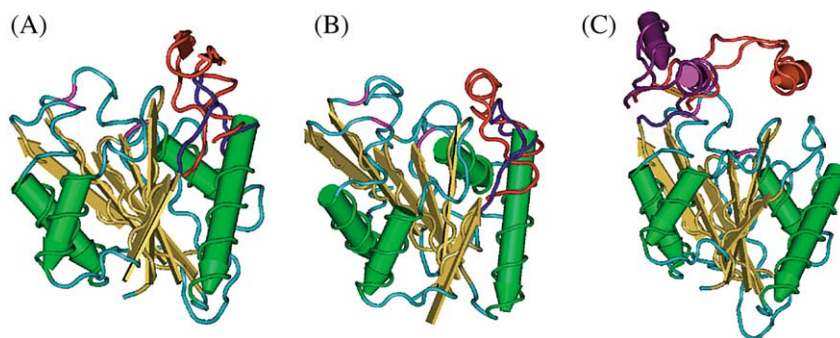
FIGURE 3 Structural models of BSLA-variants carrying artificial lid-domains. Lids engineered into BSLA at the given amino acid positions were from (A) acetylxylanesterase at position 39–51, (B) cutinase at position 39–51 and (C) human pancreatic lipase at position 153–155. The artificial lid domains are shown in red; the human pancreatic lipase lid is modeled both in the closed (red) and in the open (purple) conformation. The positions of the catalytic triad residues Ser77, His156, and Asp133 are indicated in purple.
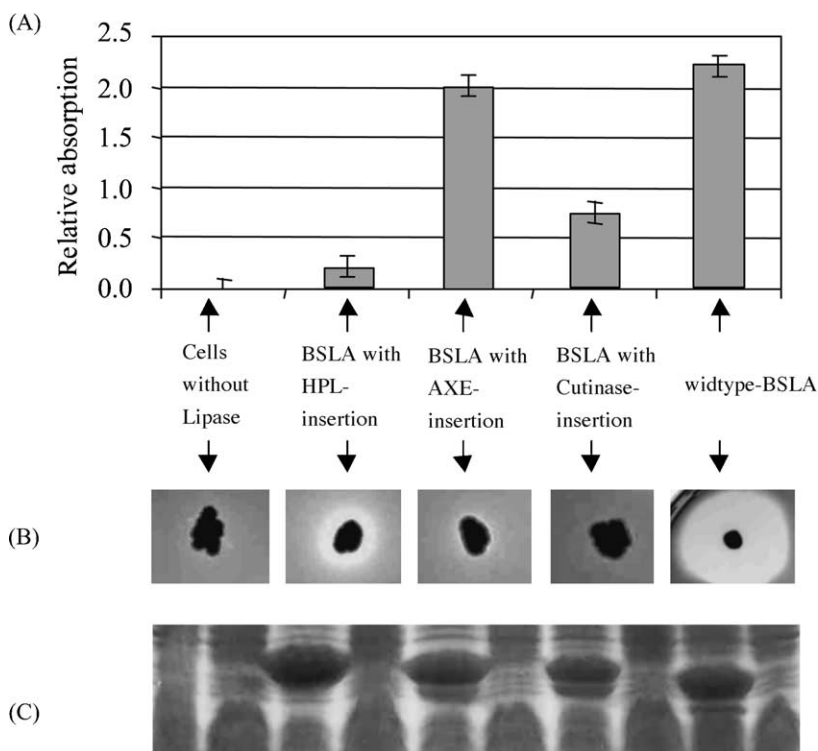

FIGURE 4 Catalytic activity of newly constructed BSLA-variants carrying artificial lid domains obtained from human pancreatic lipase (HPL), acetylxylanesterase (AXE) and cutinase. The catalytic activity towards (A) p-nitrophenyl palmitate and (B) tributyrin were tested in a spectrophotometric (Eggert et al., 2000) and an indicator plate assay, respectively. (C) Protein overexpression was detected by SDS-polyacrylamide gel electrophoresis of whole cell-lysates obtained from E. coli overexpressing the BSLA variants and subsequent staining with Coomassie Brilliant Blue. The lanes show cell extracts isolated before (right lanes) and after induction of lipase gene expression by addition of 0.3 mM isopropyl-β-D-thio-galactopyranoside (left lanes).

References

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.

Borneman, J. (1999) ''Culture-independent identification of micro-organisms that respond to specified stimuli'', *Appl. Environ. Microbiol.* **65**, 3398–3400.

Brady, L., Brzozowski, A. M., Derewenda, Z. S., Dodson, E., Dodson, G., Tolley, S., Turkenburg, J. P., Christiansen, L., Hugejensen, B., Norskov, L., Thim, L. and Menge, U. (1990) ''A serine protease triad forms the catalytic center of a triacylglycerol lipase'', *Nature* **343**, 767–770.

Brocca, S., Secundo, F., Ossola, M., Alberghina, L., Carrea, G. and Lotti, M. (2003) ''Sequence of the lid affects activity and specificity of *Candida rugosa* lipase isoenzymes'', *Prot. Sci.* **12**, 2312–2319.

Brzozowski, A. M., Derewenda, U., Derewenda, Z. S., Dodson, G. G., Lawson, D. M., Turkenburg, J. P., Bjorkling, F., Hugejensen, B., Patkar, S. A. and Thim, L. (1991) ''A model for interfacial activation in lipases from the structure of a fungal lipase-inhibitor complex'', *Nature* **351**, 491–494.

Carriere, F., Withers-Martinez, C., van Tilberugh, H., Roussel, A., Cambillau, C. and Verger, R. (1998) "Structural basis for the substrate selectivity of pancreatic lipases and some related proteins", *Biochim. Biophys. Acta* **1376**, 417–432.

Chahinian, H., Bezzine, S., Ferrato, F., Ivanova, M. G., Perez, B., Lowe, M. E. and Carriere, F. (2002a) "The beta 5′ loop of the pancreatic lipase C2-like domain plays a critical role in the lipase-lipid interactions", *Biochemistry* **41**, 13725–13735.

Chahinian, H., Nini, L., Boitard, E., Dubes, J. P., Comeau, L. C. and Sarda, L. (2002b) "Distinction between esterases and lipases: A kinetic study with vinyl esters and TAG", *Lipids* **37**, 653–662.

Cherry, J. R. and Fidantsef, A. L. (2003) "Directed evolution of industrial enzymes: an update", *Curr. Opin. Biotechnol.* **14**, 438–443.

Courtois, S., Frostegard, A., Goransson, P., Depret, G., Jeannin, P. and Simonet, P. (2001) "Quantification of bacterial subgroups in soil: comparison of DNA extracted directly from soil or from cells previously released by density gradient centrifugation", *Environ. Microbiol.* **3**, 431–439.

Cowan, D. A. (2000) "Microbial genomes–the untapped resource", *Trends Biotechnol.* **18**, 14–16.

Dartois, V., Baulard, A., Schanck, K. and Colson, C. (1992) "Cloning, nucleotide-sequence and expression in Escherichia-coli of a lipase gene from Bacillus-subtilis 168", *Biochim. Biophys. Acta* **1131**, 253–260.

Eggert, T., Pencreac'h, G., Douchet, I., Verger, R. and Jaeger, K. E. (2000) "A novel extracellular esterase from *Bacillus subtilis* and its conversion to a monoacylglycerol hydrolase", *Eur. J. Biochem.* **267**, 6459–6469.

Entcheva, P., Liebl, W., Johann, A., Hartsch, T. and Streit, W. R. (2001) "Direct cloning from enrichment cultures, a reliable strategy for isolation of complete operons and genes from microbial consortia", *Appl. Environ. Microbiol.* **67**, 89–99.

Farinas, E. T., Bulter, T. and Arnold, F. H. (2001) "Directed enzyme evolution", *Curr. Opin. Biotechnol.* **12**, 545–551.

Fishman, A. and Cogan, U. (2003) "Bio-imprinting of lipases with fatty acids", *J. Mol. Catal. B-Enzymatic* **22**, 193–202.

Funke, S. A., Eipper, A., Reetz, M. T., Otte, N., Thiel, W., Van Pouderoyen, G., Dijkstra, B. W., Jaeger, K. E. and Eggert, T. (2003) "Directed evolution of an enantioselective *Bacillus subtilis* lipase", *Biocatal. Biotrans.* **21**, 67–73.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. and Goodman, R. M. (1998) "Molecular biological access to the chemistry of unknown soil microbes: A new frontier for natural products", *Chem. Biol.* **5**, R245–R249.

Jaeger, K. E., Dijkstra, B. W. and Reetz, M. T. (1999) "Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases", *Annu. Rev. Microbiol.* **53**, 315–351.

Jaeger, K. E. and Eggert, T. (2002) "Lipases for biotechnology", *Curr. Opin. Biotechnol.* **13**, 390–397.

Jaeger, K. E., Eggert, T., Eipper, A. and Reetz, M. T. (2001) "Directed evolution and the creation of enantioselective biocatalysts", *Appl. Microbiol. Biotechnol.* **55**, 519–530.

Jaeger, K. E. and Reetz, M. T. (1998) "Microbial lipases form versatile tools for biotechnology", *Trends Biotechnol.* **16**, 396–403.

Lesuisse, E., Schanck, K. and Colson, C. (1993) "Purification and preliminary characterization of the extracellular lipase of *Bacillus subtilis* 168, an extremely basic pH-tolerant enzyme", *Eur. J. Biochem.* **216**, 155–160.

Liebeton, K., Zacharias, A. and Jaeger, K. E. (2001) "Disulfide bond in Pseudomonas aeruginosa lipase stabilizes the structure but is not required for interaction with its foldase", *J. Bacteriol.* **183**, 597–603.

Liebeton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B. W., Reetz, M. T. and Jaeger, K. E. (2000) "Directed evolution of an enantioselective lipase", *Chem. Biol.* **7**, 709–718.

Liese, A., Seelbach, K. and Wandrey, C. (2000) *Industrial Biotransformations* (Wiley-VCH, Weinheim).

Lingen, B., Kolter-Jung, D., Dünkelmann, P., Feldmann, R., Grotzinger, J., Pohl, M. and Müller, M. (2003) "Alteration of the substrate specificity of benzoylformate decarboxylase from Pseudomonas putida by directed evolution", *Chembiochem* **4**, 721–726.

Lorenz, P., Liebeton, K., Niehaus, F. and Eck, J. (2002) "Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space", *Curr. Opin. Biotechnol.* **13**, 572–577.

May, O., Nguyen, P. T. and Arnold, F. H. (2000) "Inverting enantioselectivity by directed evolution of hydantoinase for improved production of L-methionine", *Nature Biotechnol.* **18**, 317–320.

Mingarro, I., Abad, C. and Braco, L. (1995) "Interfacial activation-based molecular bioimprinting of lipolytic enzymes", *Proc. Natl Acad. Sci. USA.* **92**, 3308–3312.

Moore, J. C. and Arnold, F. H. (1996) "Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents", *Nature Biotechnol.* **14**, 458–467.

Nardini, M., Lang, D. A., Liebeton, K., Jaeger, K. E. and Dijkstra, B. M. (2000) "Crystal structure of Pseudomonas aeruginosa lipase in the open conformation – The prototype for family I.1 of bacterial lipases", *J. Biol. Chem.* **275**, 31219–31225.

Pasteur, L. (1858) "Mémoire sur la fermentation de l'acide tartarique", *C. R. Acad. Sci. (Paris)* **46**, 615–618.

Powell, K. A., Ramer, S. W., del Cardayre, S. B., Stemmer, W. P. C., Tobin, M. B., Longchamp, P. F. and Huisman, G. W. (2001) "Directed evolution and biocatalysis", *Angew. Chem. Int. Ed.* **40**, 3948–3959.

Radajewski, S., Ineson, P., Parekh, N. R. and Murrell, J. C. (2000) "Stable-isotope probing as a tool in microbial ecology", *Nature* **403**, 646–649.

Reetz, M. T., Zonta, A., Schimossek, K., Liebeton, K. and Jaeger, K. E. (1997) "Creation of enantioselective biocatalysts for organic chemistry by in vitro evolution", *Angew. Chem. Int. Ed.* **36**, 2830–2832.

Rochelle, P. A. (2001) "Extraction of nucleic acids from environmental samples", In: Rochelle, P. A., ed, *Environmental Molecular Microbiology: Protocols and Applications* (Horizon Scientific Press, Wymondham, UK).

Sarda, L. and Desnuelle, P. (1958) "Action de la lipase pancreatique sur les esters en emulsion", *Biochim. Biophys. Acta* **30**, 513–521.

Schmeisser, C., Stöckigt, C., Raasch, C., Wingender, J., Timmis, K. N., Wenderoth, D. F., Flemming, H. C., Liesegang, H., Schmitz, R. A., Jaeger, K. E. and Streit, W. R. (2003) "Metagenome survey of biofilms in drinking-water networks", *Appl. Environ. Microbiol.* **69**, 7298–7309.

Tao, H. Y. and Cornish, V. W. (2002) "Milestones in directed enzyme evolution", *Curr. Opin. Chem. Biol.* **6**, 858–864.

Torsvik, V. and Ovreas, L. (2002) "Microbial diversity and function in soil: from genes to ecosystems", *Curr. Opin. Microbiol.* **5**, 240–245.

Torsvik, V., Ovreas, L. and Thingstad, T. F. (2002) "Prokaryotic diversity – Magnitude, dynamics, and controlling factors", *Science* **296**, 1064–1066.

van Pouderoyen, G., Eggert, T., Jaeger, K. E. and Dijkstra, B. W. (2001) "The crystal structure of *Bacillus subtilis* lipase: A minimal alpha/beta hydrolase fold enzyme", *J. Mol. Biol.* **309**, 215–226.

Verger, R. (1997) "Interfacial activation' of lipases: Facts and artifacts", *Trends Biotechnol.* **15**, 32–38.

Voget, S., Leggewie, C., Uesbeck, A., Raasch, C., Jaeger, K. E. and Streit, W. R. (2003) "Prospecting for novel biocatalysts in a soil metagenome", *Appl. Environ. Microbiol.* **69**, 6235–6242.

Whitman, W. B., Coleman, D. C. and Wiebe, W. J. (1998) "Prokaryotes: The unseen majority", *Proc. Natl Acad. Sci. USA.* **95**, 6578–6583.

Winkler, F. K., Darcy, A. and Hunziker, W. (1990) "Structure of human pancreatic lipase", *Nature* **343**, 771–774.

Wong, T. S., Arnold, F. H. and Schwaneberg, U. (2004) "Laboratory evolution of cytochrome P450BM-3 monooxygenase for organic cosolvents", *Biotechnol. Bioeng.* **85**, 351–358.

Zhao, H. M. and Arnold, F. H. (1999) "Directed evolution converts subtilisin E into a functional equivalent of thermitase", *Prot. Eng.* **12**, 47–53.

Taylor & Francis
Taylor & Francis Group

Directed Evolution of an Enantioselective *Bacillus subtilis* Lipase

SUSANNE AILEEN FUNKE[a], ANDREAS EIPPER[§b], MANFRED T. REETZ[b], NIKOLAJ OTTE[c], WALTER THIEL[c], GERTIE VAN POUDEROYEN[d], BAUKE W. DIJKSTRA[d], KARL-ERICH JAEGER[a] and THORSTEN EGGERT[a]*

[a] *Institut für Molekulare Enzymtechnologie, Heinrich-Heine-Universität Düsseldorf, D-52426 Jülich, Germany;* [b] *Synthetische Organische Chemie, Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany;* [c] *Theoretische Chemie, Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany;* [d] *Laboratory of Biophysical Chemistry, Rijksuniversiteit Groningen, Nijenborgh 4, NL-9747 AG Groningen, The Netherlands*

Chiral compounds are of steadily increasing importance to the chemical industry, in particular for the production of pharmaceuticals. Where do these compounds come from? Apart from natural resources, two synthetic strategies are available: asymmetric chemical catalysis using transition metal catalysts and biocatalysis using enzymes. In the latter case, screening programs have identified a number of enzymes. However, their enantioselectivity is often not high enough for a desired reaction. This problem can be solved by applying directed evolution to create enantioselective enzymes as shown here for a lipase from *Bacillus subtilis*. The reaction studied was the asymmetric hydrolysis of *meso*-1,4-diacetoxy-2-cyclopentene with the formation of chiral alcohols which were detected by electrospray ionization mass spectrometry. Iterative cycles of random mutagenesis and screening allowed the identification of several variants with improved enantioselectivities. In parallel, we have started to use X-ray structural data to simulate the *Bacillus subtilis* lipase A-catalyzed substrate hydrolysis by using quantum mechanical and molecular mechanical calculations. This combined approach should finally enable us to devise more efficient strategies for the directed evolution of enantioselective enzymes.

Keywords: Directed evolution; Enantioselectivity; Lipase; Esterase; *Bacillus subtilis*; Saturation mutagenesis

INTRODUCTION

The world market for enantiomerically pure compounds is increasing, with the worldwide sales volume for single enantiomers reaching \$6.63 billion in 2001 (Stinson, 2001). The necessity to use enantiomerically pure compounds as active ingredients in pharmaceuticals will result in a further sales volume increase for enantiomerically pure building blocks to reach \$15.1 billion in 2005 (Rouhi, 2002). When looking at the chiral drugs market, the numbers are even more impressive: in 2001, single enantiomer drug sales reached \$147 billion, which corresponds to a market share of 36% for the worldwide pharmaceutical products (Rouhi, 2002).

Single enantiomers become available either by isolation from natural sources (the so-called chiral pool) or by asymmetric chemical catalysis using transition metal catalysts. In addition, biocatalysis has evolved as an attractive alternative, with lipases being the most widely used enzymes which work in aqueous as well as organic solvents and catalyze a large variety of different reactions with a high substrate specificity and stereoselectivity (Jaeger and Eggert, 2002; Reetz, 2002). However, biocatalysts generally suffer from the disadvantage that for a given synthetic transformation of interest, $A \rightarrow B$, enantioselectivity may well be poor. This problem can be solved by improving the enantioselectivity of enzymes using directed evolution (for reviews see Bornscheuer and Pohl, 2001; Jaeger *et al.*, 2001; Reetz and Jaeger, 2002). This strategy includes (i) the generation of mutant enzyme genes using different mutagenesis methods like error-prone polymerase chain reaction, saturation mutagenesis and *in vitro* recombination (Stemmer, 1994a,b; Zhao *et al.*, 1998; Shao *et al.*, 2000; Lutz *et al.*, 2001), (ii) followed by gene expression to generate the corresponding en-

§ present address: BASF AG, Carl-Bosch-Straße 38, 67056 Ludwigshafen, Germany

* Corresponding author. Tel.: + 49-2461-61-2939. Fax: + 49-2461-61-2490. E-mail: t.eggert@fz-juelich.de

zyme variants (Rosenau and Jaeger, 2003) and (iii)
high-throughput screening for enantioselectivity
(Reetz, 2001).

One general problem in evolving (enantioselec-
tive) enzymes concerns the choice of an appropriate
mutagenesis method to be used for generating a first
generation library. In most cases, epPCR is used
followed by recombinative methods and/or satura-
tion mutagenesis. However, single base substitutions
introduced by epPCR result in a limited number of
amino acid exchanges introduced into the enzyme.
In other words, only a small part of the total
sequence space is susceptible to mutagenesis and
alternative methods must be applied to generate a
first generation library of high diversity (Eggert and
Jaeger, 2003).

We have used directed evolution to improve the
enantioselectivity of *Bacillus subtilis* lipase A (BSLA)
in the asymmetric hydrolysis of *meso*-1,4-diacetoxy-
2-cyclopentene with the formation of chiral alcohols
(Fig. 1). Different directed evolution strategies were
compared to evaluate their potential for the efficient
creation of enantioselelective BSLA variants.

MATERIALS AND METHODS

Heterologous Expression of BSLA in *Escherichia coli*

The lipase gene *lip* A lacking a 93 bp fragment at the
5′-end which encodes its signal sequence was
amplified by standard PCR using the upstream 30
bp primer BSLA1 5′-ATATGATATCGCTGAACA-
CAATCCAGTCGT-3′, the downstream 29 bp-primer
BSLA2   5′-TATAGAGCTCTCATTAATTCG-
TATTCTGG-3′ and genomic DNA from *B. subtilis*
168 (obtained from the *Bacillus* Genetic Stock Center,
Ohio, USA) as a template. Unique *Eco*RV and *Sac*I
restriction sites were introduced and used to clone
the resulting 557 bp PCR product into the corre-
sponding restriction sites of the *Escherichia coli*
expression plasmid pET22b (Novagen, Madison,
USA) giving pET22lipA.

Standard PCR Conditions

Amplification of DNA fragments was performed in
a 50 µl reaction mix containing 1 ng plasmid- or 10
ng genomic-DNA as the template, 25 pmol of each
primer, 0.2 mM dNTP's, 2.5 U *Taq*- (Eurogentec,
Seraing, Belgium) or *Pfu*- (Stratagene, Heidelberg,
Germany) polymerase. Buffers containing $MgCl_2$ or
$MgSO_4$ were used as recommended by the manu-
facturers. Conditions for PCR were as follows: $1 \times$ (3
min 98°C); $35 \times$ (1 min 95°C; 2 min 58°C, 1 min 72°C)
and $1 \times$ (7 min 72°C). The PCR reaction was per-
formed using a Mastercycler Gradient (Eppendorf,
Hamburg, Germany).

Mutagenesis Methods

Random Mutagenesis (epPCR)

Random mutagenesis of the whole *B. subtilis* lipase
gene was performed by error-prone polymerase
chain reaction (epPCR) (Zhou *et al.*, 1991). The upper
30   bp   primer   mut1up   5′-CTCCTC-
GCTGCCCAGCCGGCGATGGCCATG-3′ and the
lower 29 bp primer mut1low 5′-GCCGCAAGCTT-
GTCGACGAGCTCTCATTA-3′ were used in a stan-
dard PCR reaction to introduce unique *Mls*I and
*Hind*III restriction sites for direct cloning into the
expression vector pET22b. An error rate of 2–4 base
substitutions per gene was achieved by using the
following reaction mix: 5 pmol of each primer, 75
mM Tris/HCl buffer (pH 8.8), 20 mM $(NH_4)_2SO_4$, 6.0
mM $MgCl_2$, 0.15 mM $MnCl_2$, 0.2 mM dNTPs, 0.1%
Tween 20, 1 ng template DNA (pET22lipA plasmid)
and 2.5 U *Taq*-polymerase (Eurogentec, Seraing,
Belgium). For epPCR with lower or higher error
rates, the concentration of $MnCl_2$ was varied as
described elsewhere (Jaeger *et al.*, 2001).

Single- and Multiple-site Saturation Mutagenesis

Saturation mutagenesis was performed using a
megaprimer PCR mutagenesis method as described
by Barettino *et al.* (1994). In the first PCR reaction,
the megaprimer harboring the desired point muta-
tion(s) was amplified by using a mutagenesis primer

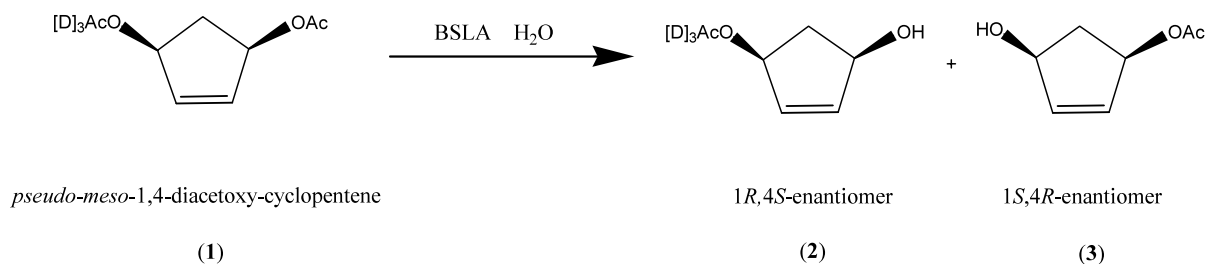|		
pseudo-meso-1,4-diacetoxy-cyclopentene	1*R*,4*S*-enantiomer	1*S*,4*R*-enantiomer
(1)	**(2)**	**(3)**

FIGURE 1   The asymmetric hydrolysis of the model compound *meso*-1,4 diacetoxy-2-cyclopentene was determined by using a high-
throughput ESI-MS system. The substrate (*pseudo-meso*-1,4-diacetoxy-2-cyclopentene) was deuterium labeled to follow the formation of
chiral alcohols (2) and (3) by differences in their mass spectrum.

carrying a randomized codon at the desired position and the upstream primer mut1up. In the second PCR reaction, the full-length gene was amplified using the megaprimer together with mut1up and an additional downstream primer mutSlow 5′-ATA-TAAGCTTCAGCAAACAGCTATGACCATGATAT-CGAATTC-3′.

For multiple-site saturation mutagenesis (MSSM), the megaprimer method was used with mutagenesis primers simultaneously saturating three amino acid positions: MSSM1 5′-AATCAGGCTGTTGACSNN-GCTGCTGTASNNAAGGCCSNNGTGTCCAACGC-CATG-3′ and MSSM2 5′-TTCATCTAAAACCTT-SNNCACSNNTCGTGATAATACCGGTCC ATTSN-NATAATTTGTGCCTGT-3′ (mutated codons are underlined). MSSM1 introduced mutations at amino acid positions 50, 58 and 60; and MSSM2 at amino acid positions 157, 160 and 164.

Expression and Screening

The PCR-amplified *lipA* mutant genes were cloned into vector pET22b (Novagen, Madison, USA) and the resulting plasmids were transformed into *E. coli* BL21(DE3) (Novagen, Madison, USA). The clones were cultured at 37°C in 96 deep well microtiter plates filled with 1 ml of LB/M9 medium (10 g/l tryptone, 5 g/l yeast extract, 5,5 g/l NaCl, 4 g/l glucose, 0.25 g/l $MgSO_4 \times 7H_2O$, 0.02 g/l $CaCl_2$, 7 g/l $Na_2HPO_4 \times 2H_2O$, 3 g/l KH_2PO_4, 1 g/l NH_4Cl) supplemented with 100 μg/ml carbenicillin. After 6 hours of shaking at 37°C ($OD_{580} = 0.5 - 0.7$) lipase expression was induced by adding isopropyl-β-D-thiogalactoside (IPTG) to a final concentration of 0.3 mM. The induced culture was grown at 37°C and the cells were separated from the culture by centrifugation at 5000 g for 10 min. An aliquot of 100 μl from the culture supernatant was taken from each well and pipetted into another 96 deep well microtiter plate containing 800 μl of 10 mM Na_2HPO_4/KH_2PO_4 buffer (pH 7.5) and 100 μl of the substrate dissolved in dimethylsulfoxide (100 mM). After 24 h of shaking at room temperature, the reaction solution was extracted with ethyl acetate and screened by electrospray ionization mass spectroscopy (ESI-MS) (Reetz *et al.*, 1999). Variants showing an improved enantioselectivity were reanalyzed independently by chiral gas chromatography. The corresponding mutant genes were sequenced using standard T7-promoter and T7-terminator primers (Medigenomix, Martinsried, Germany).

Stability of Wild Type and Variant Lipases

The thermal stability of the enzymes was determined after expression of the wild type protein and different variants in *E. coli* BL21(DE3) by incubation of the enzyme-containing supernatants for 2 h at 45°C. The residual enzymatic activities were determined spectrophotometrically using *p*-nitrophenyl-caprylate as the substrate (Eggert *et al.*, 2000).

RESULTS AND DISCUSSION

Directed Evolution of BSLA using Iterative Cycles of epPCR

Directed evolution was used to improve the enantioselectivity of BSLA in the asymmetric hydrolysis of *meso*-1,4-diacetoxy-2-cyclopentene with the formation of chiral alcohols (Fig. 1). This reaction does not constitute a kinetic resolution and can thus be carried out to 100% conversion. Screening was done by electrospray ionisation mass spectrometry (ESI-MS) using the deuterium labeled *pseudo-meso* substrate. The wild-type enzyme leads to an *ee*-value of only 38% in favor of the (1*R*, 4*S*) enantiomer (Fig. 1).

The initial strategy was based on random mutagenesis of the lipase gene using epPCR with a low mutagenesis frequency resulting in one amino acid exchange per variant and subsequent high throughput screening based on ESI-MS. In the first generation created by low error rate epPCR, about five improved mutants were identified, the best one resulting in an *ee*-value of 48% in the test reaction. In parallel, the screening of a first generation prepared by high error rate epPCR corresponding to 2−3 amino acid exchanges per variant resulted in five variants with *ee*-values up to 58% *ee* (Table I). These variants parented a second generation of mutants created by epPCR with high error rate. The total number of second-generation libraries contained 15 000 variants. Four enantioselective lipase variants with enantioselectivities up to *ee*-values of 69% were identified (Table I).

The best variants of the second generation libraries were used for a third round of epPCR. Although 9000 variants were screened, variants with a further improved enantioselectivity could not be identified (Table II).

The DNA sequence analysis of the lipase variants listed in Table I allowed us to identify 'hot-spot positions' important for the enantioselectivity of the enzyme. Further enzyme optimization was achieved by saturation mutagenesis at the following amino acid positions: Ile22, Tyr49, Asn50, Gln60, Leu124 and Gln164. Additionally, multiple site saturation mutagenesis (MSSM) was performed at two regions previously identified as being important for enantioselectivity comprising amino acid positions Asn50, Phe58 and Gln60 (MSSM1) and Ile157, Leu160 and Gln164 (MSSM2). This round of mutagenesis and screening lead to the identification of a novel BSLA variant with inverted enantioselectivity

TABLE I Enzyme variants with improved enantioselectivity towards the model substrate

| Generation (mutagenesis method)	Variant	*ee* [%]	Amino acid exchange
	wild-type	38	–
1st (low error rate epPCR)	–	40 – 50	n.d.
1st (high error rate epPCR)	G1-II-2H12	58	I22V Q164H
	G1-II-3-H6	57	I22T L114P
	G1-II-2-A7	54	L124S
	G1-III-10-C10	56	N50S
	G1-III-13-G9	54	Q60R
2nd (high error rate epPCR)	G2-II-1-C12	64	I22V L160Q Q164H
	G2-II-9-E1	69	I22T F58L L114P
	G2-II-5-E4	65	L124S I157N
	G2-II-5-E10	64	M78T L124S S130G
(saturation mutagenesis)	Q60N	65	Q60N
	Q60L	64	Q60L
(multiple site saturation mutagenesis)	MSSM1	38 – 54	n.d
	MSSM2	15 rev.	n.d.

n.d. = not determined.
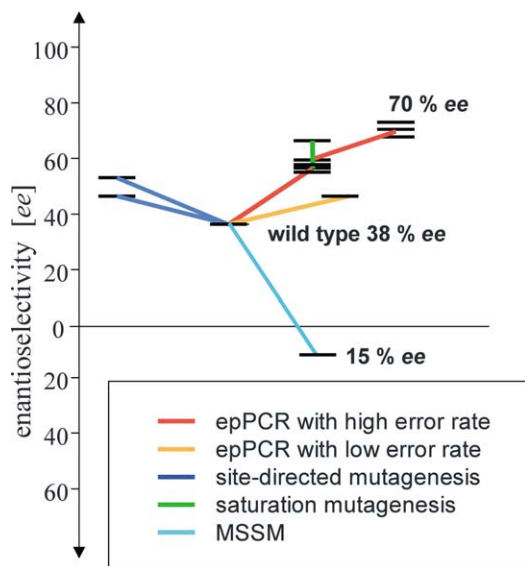
FIGURE 2 Directed evolution of enantioselective *B. subtilis* lipase. The mutagenesis methods used to evolve enantioselective variants are indicated by different colors.

(15% *ee*, rev.). The results of this directed evolution approach are summarized in Fig. 2.

Biochemical Characterization of Enantioselective Lipase Variants

In total, 14 enantioselective BSLA variants in the asymmetric hydrolysis of the model substrate were identified (Table I). However, after a third round of epPCR mutagenesis, improved variants were not obtained although about 9000 variants were screened. This result suggested that the improvement of BSLA enantioselectivity was either impossible or was caused by otherwise deleterious mutations. Therefore, we decided to compare enzyme activities and stabilities of the variants to those of the wild-type enzyme.

The specific activities of improved BSLA variants purified by phenyl sepharose chromatography (Eggert *et al.*, 2000) were different from the wild-type enzyme. However, all improved variants still

TABLE II Distribution of active variants in the different epPCR generations

Generation	epPCR method	Active variants [%][1]	Screened variants[2]	Max. *ee* [%][3]
1st	low error rate	85	1000	50
1st	high error rate	60	4000	58
2nd	high error rate	45	5 × 3000	69
3rd	high error rate	10	3 × 3000	65 – 69

[1]percentage of lipolytically active BSLA variants tested in a pre-screening on tributyrin agar plates.
[2]number of active enzyme variants screened for enantioselectivity in the asymmetric hydrolysis of *meso*-1,4-diacetoxy-2-cyclopentene.
[3]maximum *ee*-values identified from screening.

showed enzymatic activity towards the standard lipase substrate *p*-nitrophenyl caprylate indicating that lipolytic activity itself was retained also after the third round of random mutagenesis.

Additionally, the protein stability of selected variants and the wild-type lipase was compared by determination of their residual lipolytic activity after incubation for two hours at 45°C. The wild-type lipase was stable under these conditions, but most variants showed a decrease in lipolytic activity over time, indicating a decreased stability at elevated temperatures as compared to the wild-type enzyme. A plot of the enantioselectivities of the lipase variants versus their respective (thermo)stabilities revealed a negative correlation: the higher the enantioselectivities of the enzymes the lower were their stabilities (Fig. 3). Variants G2-II-1C12, G2-II-9E1 and G2-II-5E10 (see Table I) each containing three amino acid substitutions were selected to parent the third generation. Obviously, the mutations present in these variants not only lead to an increased enantioselectivity but also negatively affected their stability.

These results suggest that, in contrast to larger proteins of $M_r \geq 30$ kDa, small enzymes like BSLA (181 amino acids, M_r 19.4 kDa) are generally more susceptible to amino acid substitutions exerting a negative effect on stability and activity. Enzymes of smaller size have naturally been evolved as compact structures as suggested by the crystal structure of BSLA (van Pouderoyen *et al.*, 2001). A larger number of amino acid substitutions therefore dramatically increase the chance of causing negative effects on their overall structure. As a consequence, directed evolution of such enzymes should start with the preparation of a high diversity first generation library carrying a low number of mutations (1−2 amino acid exchanges per protein molecule). Accordingly, we observed that saturation mutagenesis seemed to have a better potential to improve the enzyme's enantioselectivity against the model substrate while maintaining its stability and activity as seen for variants I22V, Q60L, Q60R, Y49C and N50S (bold black letters in Fig. 3).

Directed Evolution of BSLA using Complete Saturation Mutagenesis

Single base mutations as introduced by epPCR cannot result in the creation of all theoretically possible amino acid substitutions. The construction of a mutant library containing all possible single amino acid exchanges would be necessary to achieve this ambitious goal. We have saturated every single amino acid position in BSLA by separate saturation mutagenesis of the codons encoding amino acids 1−181 of the native enzyme. The mutant genes were cloned and the enzymes were overexpressed in *E. coli*. In this ongoing project, several different variants with improved (*ee*-value of 65% in favor of 1*S*,4*R*) and inverted (*ee*-value of 56% in favor of 1*R*,4*S*) enantioselectivities towards the model substrate (Fig. 1) were identified. Several BSLA variants with improved specific activities were identified as well. Interestingly, most of these variants proved to be stable, presumably due to the low number of amino acid exchanges per protein molecule (Funke *et al.*, 2003).
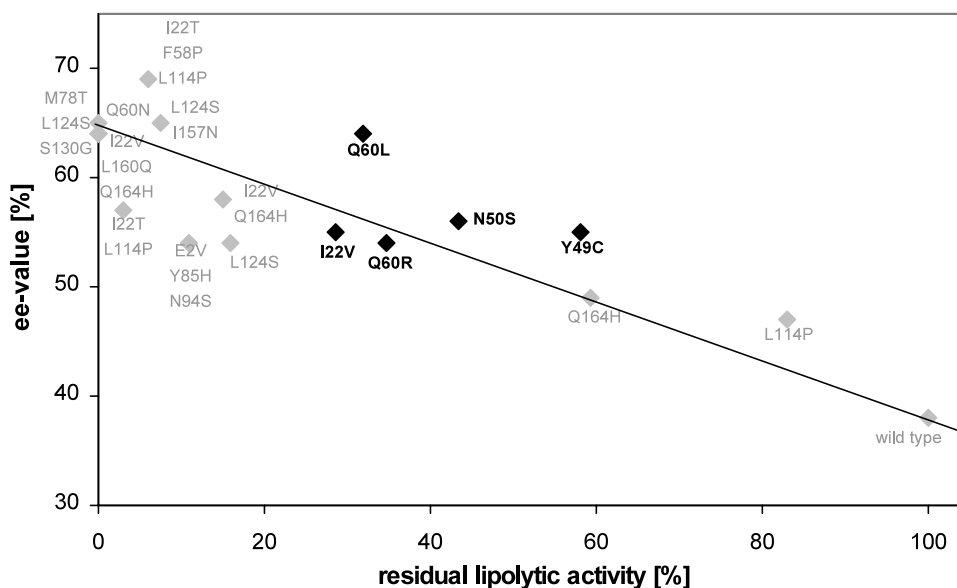

FIGURE 3 Thermostability of wild-type and variant lipases. The residual lipolytic activities were determined after 2 h of incubation at 45°C. The wild-type LipA protein was (thermo)stable under these conditions and its activity was defined as 100%. BSLA-variants showing improved enantioselectivity and maintaining their stability and activity are highlighted in bold black letters.

CONCLUSIONS AND FUTURE PROSPECTS

Directed evolution is a successful strategy to improve enzyme properties such as specific activities, substrate specificities or stabilities and optimize biocatalytic processes (Petrounia and Arnold, 2000; Zhao *et al.*, 2002). Additionally, it is a powerful tool to create enantioselective biocatalysts (Reetz *et al.*, 1997; Jaeger *et al.*, 1999; Liebeton *et al.*, 2000; Jaeger *et al.*, 2001; Jaeger and Eggert, 2002; Reetz and Jaeger, 2002). Success depends on the effective combination of different mutagenesis methods with efficient screening or selection procedures. The quality of the first generation library is of key importance because its variants usually parent all subsequent generations.

Although directed evolution works even without any knowledge of an enzyme's structure or reaction mechanism (Petrounia and Arnold, 2000), a given project may nevertheless turn out to be time-consuming and cumbersome. According to our experience, the knowledge of a three-dimensional protein structure can significantly speed up a directed evolution approach because it may allow the size of the sequence space to be narrowed down. Despite its value in identifying the spacial positions of important amino acid residues, a crystal structure itself does not reflect the time-resolved catalytic reaction mechanism. Therefore, we have recently started to simulate the BSLA-catalyzed substrate hydrolysis by using combined quantum mechanical and molecular mechanical calculations (QM/MM) (Fig. 4). These computer calculations describe a dynamic reaction by treating the active site residues quantum mechanically, while the surrounding en-vironment is simulated by a force field method (Schöneboom *et al.*, 2002). BSLA is an ideal candidate enzyme for QM/MM calculations because it is a small enzyme with a known crystal structure (van Pouderoyen *et al.*, 2001).

Hopefully, this approach will result in a more detailed understanding of the enzyme's topology and its changes in the course of the enantioselective reaction. In the end, a structure and theory assisted analysis of a directed evolution approach will facilitate the understanding of an enantioselective enzyme reaction and, at the same time, help to accelerate the creation of enantioselective lipases useful for industrial applications.
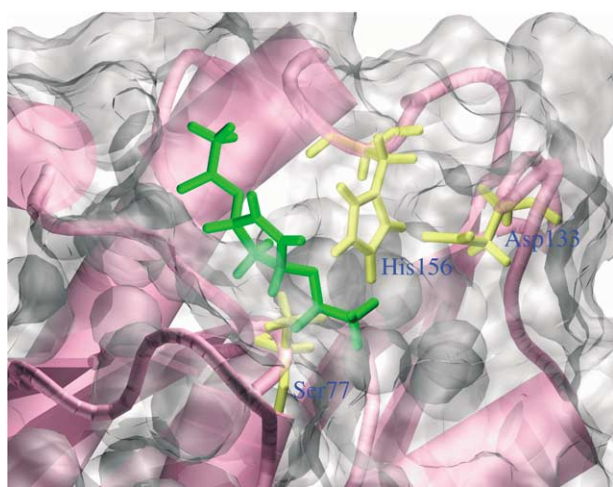

FIGURE 4   QM/MM simulation of *meso*-1,4-diacetoxy-2-cyclopentene hydrolysis catalyzed by BSLA calculated on the basis of the crystal structure. Quantum mechanical calculations were applied for the substrate (green) and the side chains of the catalytic triad residues Ser77, Asp133, and His156 (yellow). The remaining part of the enzyme was treated by a force field method. For clarity, the 20 Å water sphere capping the active site is not displayed.

References

Barettino, D., Feigenbutz, M., Valcárcel, R. and Stunnenberg, H.G. (1994) ''Improved method for PCR-mediated site-directed mutagenesis'', *Nucleic Acids Res.* **22**, 541–542.

Bornscheuer, U.T. and Pohl, M. (2001) ''Improved biocatalysts by directed evolution and rational protein design'', *Curr. Opin. Chem. Biol.* **5**, 137–143.

Eggert, T. and Jaeger, K.-E. (2003) ''Directed evolution by random mutagenesis: a critical evaluation'', In: Svendsen, A., ed, *Enzyme functionality: Design, Engineering and Screening* (Marcel Dekker, New York) (in press).

Eggert, T., Pencreac'h, G., Douchet, I., Verger, R. and Jaeger, K.-E. (2000) ''A novel extracellular esterase from *Bacillus subtilis* and its conversion to a monoacylglycerol hydrolase'', *Eur. J. Biochem.* **267**, 6459–6469.

Funke, S.A., Eipper, A., Eggert, T., Jaeger, K.-E., Reetz, M.T. (submitted for publication).

Jaeger, K.-E. and Eggert, T. (2002) ''Lipases for biotechnology'', *Curr. Opin. Biotechnol.* **13**, 390–397.

Jaeger, K.-E., Dijkstra, B.W. and Reetz, M.T. (1999) ''Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases'', *Annu. Rev. Microbiol.* **53**, 315–351.

Jaeger, K.-E., Eggert, T., Eipper, A. and Reetz, M.T. (2001) ''Directed evolution and the creation of enantioselective biocatalysts'', *Appl. Microbiol. Biotechnol.* **55**, 519–530.

Liebeton, K., Zonta, A., Schimossek, K., Nardini, M., Lang, D., Dijkstra, B.W., Reetz, M.T. and Jaeger, K.-E. (2000) ''Directed evolution of an enantioselective lipase'', *Chem. Biol.* **7**, 209–218.

Lutz, S., Ostermeier, M., Moore, G.I., Maranas, C.D. and Benkovic, S.J. (2001) ''Creating multiple-crossover libraries independent of sequence identity'', *Proc. Natl. Acad. Sci.* **98**, 11248–11253.

Petrounia, I.P. and Arnold, F. (2000) ''Designed evolution of enzymatic properties'', *Curr. Opin. Biotechnol.* **11**, 325–330.

Reetz, M.T. (2001) ''Combinatorial and evolution-based methods in the creation of enantioselective catalysts'', *Angew. Chem. Int. Ed.* **40**, 284–310.

Reetz, M.T. (2002) ''Lipases as practical biocatalysts'', *Curr. Opin. Chem. Biol.* **6**, 145–150.

Reetz, M.T. and Jaeger, K.-E. (2002) ''Directed evolution as a means to create enantioselective enzymes for use in organic chemistry'', In: Brakman, S. and Johnsson, K., eds, *Directed Molecular Evolution of Proteins or How to Improve Enzymes for Biocatalysis* (WILEY-VCH, Weinheim, Germany), pp 245–279.

Reetz, M.T., Zonta, A., Schimossek, K., Liebeton, K. and Jaeger, K.-E. (1997) ''Creation of enantioselective biocatalysts for organic chemistry by *in vitro* evolution'', *Angew. Chem. Int. Ed.* **36**, 2830–2832.

Reetz, M.T., Becker, M.H., Klein, H.W. and Stöckigt, D. (1999) ''A method for high-throughput screening of enantioselective catalysts'', *Angew. Chem. Int. Ed.* **38**, 1758–1761.

Rosenau, F. and Jaeger, K.-E. (2003) ''Overexpression of biocatalysts in *Pseudomonas*'', In: Svendsen, A., ed, *Enzyme functionality: Design, Engineering and Screening* (Marcel Dekker, New York) (in press).

Rouhi, A.M. (2002) ''Chiral roundup'', *Chem. Eng. News* **80**(23), 43–50.

Shao, Z., Zhao, H., Giver, L. and Arnold, F.H. (2000) ''Random-priming *in vitro* recombination: an effective tool for directed evolution'', *Nucleic Acids Res.* **26**, 681–683.

Schöneboom, J.C., Lin, H., Reuter, N., Thiel, W., Cohen, S., Ogliaro, F. and Shaik, S. (2002) ''The elusive oxidant species of cytochrome p450 enzymes: characterization by combined quantum mechanical/ molecular mechanical (QM/MM) calculations'', *J. Am. Chem. Soc.* **124**, 8142–8151.

Stemmer, W.P.C. (1994) ''Rapid evolution of a protein *in vitro* by DNA shuffling'', *Nature* **370**, 389–391.

Stemmer, W.P.C. (1994) ''DNA shuffling by random fragmentation and reassembly: *In vitro* recombination for molecular evolution'', *Proc. Natl. Acad. Sci.* **91**, 10747–10751.

Stinson, S.C. (2001) ''Chiral pharmaceuticals'', *Chem. Eng. News.* **79**(40), 79–97.

van Pouderoyen, G., Eggert, T., Jaeger, K.-E. and Dijkstra, B.W. (2001) ''The crystal structure of *Bacillus subtilis* lipase: a minimal α/β-hydrolase fold enzyme'', *J. Mol. Biol.* **309**, 215–226.

Zhao, H., Giver, L., Shao, Z., Affholter, J.A. and Arnold, F. (1998) ''Molecular evolution by staggered extension process (StEP) *in vitro* recombination'', *Nat. Biotechnol.* **16**, 258–261.

Zhao, H., Chockalingam, K. and Chen, Z. (2002) ''Directed evolution of enzymes and pathways for industrial biocatalysis'', *Curr. Opin. Biotechnol.* **13**, 104–110.

Zhou, Y.H., Zhang, X.P. and Ebright, R.H. (1991) ''Random mutagenesis of gene-sized DNA molecules by use of PCR with *Taq* DNA polymerase'', *Nucleic Acids Res.* **19**, 6052.

Curriculum Vitae

Dr. Thorsten Eggert

Personal data Born July 21st, 1972 in Essen, Germany

Married since November 27th, 1998 to Barbara Eggert nee Schmidt

One daughter Laila Henrike, born September 25th, 2005

University education

1993 – 1998 Biology studies at Ruhr-Universität Bochum

Diploma thesis: "Cloning, expression and characterization of a
disulfate-oxidoreductase from *Pseudomonas aeruginosa*"

1998 – 2001 Institut of Microbiology with Prof. Dr. Karl-Erich Jaeger

PhD thesis: „The lipolytic enzymes LipA and LipB from *Bacillus
subtilis*: characterization and optimization by directed evolution"

Professional career

since 2002 Wissenschaftlicher Assistent at the Institute of Molecular
Enzyme Technology, Heinrich-Heine Univertität Düsseldorf,
Head of Directed Evolution research group

Fellowships / Awards

2000 - 2001 PhD-Scholar-ship funded by the Wilhelm und Günter Esser-
Stiftung of the Ruhr-Universität Bochum

2002 Ruth-Massenberg-Promotionspreis 2002 awarded by the Ruhr-
Universität Bochum

Publicationlist

Dr. Thorsten Eggert

Publications in scientific journals

(1)   Eggert, T., Pencreac´h, G., Douchet, I., Verger, R. and Jaeger, K.-E. (2000) A novel extracellular esterase from *Bacillus subtilis* and its conversion to a monoacylglycerol hydrolase. *Eur. J. Biochem.* **267**: 6459-6469.

(2)   Urban, A., Leipelt, M., Eggert, T. and Jaeger, K.-E. (2001) DsbA and DsbC affect extracellular enzyme formation in *Pseudomonas aeruginosa*. *J. Bacteriol.* **183**: 587-596.

(3)   Eggert, T., van Pouderoyen, G., Dijkstra, B.W. and Jaeger, K.-E. (2001) Lipolytic enzymes LipA and LipB from *Bacillus subtilis* differ in regulation of gene expression, biochemical properties and three-dimensional structure. *FEBS Lett.* **502**: 89-92.

(4)   van Pouderoyen, G., Eggert, T., Jaeger, K.-E. and Dijkstra, B.W. (2001) The crystal structure of *Bacillus subtilis* lipase: a minimal α/β hydrolase fold enzyme. *J. Mol. Biol.* **309**: 215-226.

(5)   Jaeger, K.-E., Eggert, T., Eipper, A. and Reetz, M.T. (2001) Directed evolution and the creation of enantioselective biocatalysts. *Appl. Microbiol. Biotechnol.* **55**: 519-530.

(6)   Eggert, T., Rosenau, F. and Jaeger, K.-E. (2002) Bakterielle Lipasen: Bedeutsame Enzyme für Grundlagenforschung und Anwendung. *BIOforum* **25**: 573-576.

(7)   Eggert, T., van Pouderoyen, G., Pencreac´h, G., Douchet, I., Verger, R., Dijkstra, B.W. and Jaeger, K.-E. (2002) Biochemical properties and three-dimensional structures of two extracellular lipolytic enzymes from *Bacillus subtilis*. *Colloids and Surfaces B: Biointerfaces* **26**: 37-46.

(8)   Jaeger, K.-E. and Eggert, T. (2002) Lipases for Biotechnology. *Curr. Opin. Biotechnol.* **13**: 390-397.

(9)   Funke, S.A., Eipper, A., Reetz, M.T., Otte, N., Thiel, W., van Pouderoyen, G., Dijkstra, B.W., Jaeger, K.-E. and Eggert, T. (2003) Directed evolution of an enantioselective *Bacillus subtilis* lipase. *Biocatal. Biotrans.* **21**: 67-73.

(10)   Eggert, T., Brockmeier, U., Dröge, M., Quax, W.J. and Jaeger, K.-E. (2003) Extracellular lipases from *Bacillus subtilis*: regulation of gene expression and enzyme activity by amino acid supply and external pH. *FEMS Microbiol. Lett.* **225**: 319-324.

(11)   Eggert, T., Leggewie, C., Puls, M., Streit, W., van Pouderoyen, G., Dijkstra, B.W. and Jaeger, K.-E. (2004) Novel biocatalysts by identification and design. *Biocatal. Biotrans.* **22**: 139-144.

(12)   Jaeger, K.-E. and Eggert, T. (2004) Enantioselective biocatalysis optimized by directed evolution. *Curr. Opin. Biotechnol.* **15**: 305-313.

(13)   Eggert, T., Funke, S.A., Rao, N.M., Acharya, P., Krumm, H., Reetz, M.T. and Jaeger, K.-E. (2005) Multiplex-PCR-based recombination as a novel high fidelity method for directed evolution. *ChemBioChem.* **6**: 1062-1067.

(14)   Krauss, U., Losi, A., Gärtner, W., Jaeger, K.-E. and Eggert, T. (2005) Initial characterization of a blue-light sensing, phototropin-related protein from *Pseudomonas putida*: a paradigm for an extended LOV construct. *Phys. Chem. Chem. Phys.* **7**: 2804-2811.

(15)   Funke, S.A., Otte, N., Eggert, T., Bocola, M., Jaeger, K.-E. and Thiel, W. (2005) Combination of computational prescreening and experimental library construction can accelerate enzyme optimization by directed evolution. *Protein Eng. Des. Sel.* **18**: 509-514.

(16)   Krauss, U. and Eggert, T. (2005) *In.silico.mutagenesis*: a primer selection tool designed for sequence scanning applications used in directed evolution experiments. *BioTechniques.* **39**: 679-682.

(17)   Brockmeier, U., Wendorff, M. and Eggert, T. (2006) Versatile expression and secretion vectors for *Bacillus subtilis. Curr. Microbiol.* **52**: 143-148.

(18)   Drepper, T., Eggert, T., Hummel, W., Leggewie, C., Pohl, M., Rosenau, F. and Jaeger, K.-E. (2006) Neue Biokatalysatoren für die Weiße Biotechnologie. *Chem. Ing. Tech.* **78**: 239-248.

(19)   Eggert, T. (2006) Bionik im Reagenzglas – Biokatalysator-Design nach dem Vorbild der Natur durch Gerichtete Evolution. *Chem. Ing. Tech.* **78**: 453-459.

(20)   Detry, J., Rosenbaum, T., Lütz, S., Hahn, D., Jaeger, K.-E., Müller, M. and Eggert, T. (2006) Biocatalytic production of enantiopure cyclohexane-1,2-diol on a preparative scale using extracellular lipases from *Bacillus subtilis. Appl. Microbiol. Biotechnol.* **72**: 1107-1116.

(21)   Secundo, F., Carrea, G., Tarabiono, C., Gatti-Lafranconi, P., Brocca, S., Lotti, M., Jaeger, K.-E., Puls, M. and Eggert, T. (2006) The lid is a structural and functional determinant of lipase activity and selectivity. *J. Mol. Catal. B Enzym.* **39**: 166-170.

(22)   Drepper, T., Eggert, T., Hummel, W., Leggewie, C., Pohl, M., Rosenau, F. and Jaeger, K.-E. (2006) Novel biocatalysts for White Biotechnology. *Biotechnol. J.* **1**: 777-786.

(23)   Brockmeier, U., Caspers, M., Freudl, R., Jockwer, A., Noll, T. and Eggert, T. (2006) Systematic screening of all signal peptides from *Bacillus subtilis*: a powerful strategy in optimizing heterologous protein secretion in Gram-positive bacteria. *J. Mol. Biol.* **362**: 393-402.

(24)   Andexer, J., Guterl, J.-K., Pohl, M. and Eggert, T. (2006) A High-throughput screening assay for hydroxynitrile lyase activity. *Chem. Commun.* **40**: 4201-4203.

(25)   Henning, H., Leggewie, C., Pohl, M., Müller, M., Eggert, T. and Jaeger, K.-E. (2006) Identification of novel biocatalysts displaying benzoylformate-decarboxylase activity by selection. *Appl. Environmental. Microbiol.* **72**: 7510-7517.

(26)   Buttani, V., Losi, A., Eggert, T., Krauss, U., Jaeger, K.-E., Cao, Z. and Gärtner, W. (2006) Conformational analysis of the blue-light sensing protein YtvA reveals a competitive interface for LOV-LOV dimerization and interdomain interactions. *Photochem. Photobiol. Sci.* **6**: 41-49.

(27)   Reetz, M.T., Puls, M., Carballeira, J.D., Vogel, A., Jaeger, K.-E., **Eggert, T.**, Thiel, W., Bocola, M. and Otte, N. (2006) Learning from directed evolution: further lessons from theoretical investigations into kooperative mutations in lipase enantioselectivity. *ChemBioChem.* **8**: 106-112.

(28)   Drepper, T., **Eggert, T.**, Circolone, F., Heck, A., Krauß, U., Guterl, J.-K., Wendorff, M., Losi, A, Gärtner, W. and Jaeger, K.-E. (2007) *In vivo* fluorescence without oxygen: Novel fluorescence reporters for anaerobic systems. *Nat. Biotechnol.* **25**: 443-445.

Book chapters

(29)   Eggert, T., Reetz, M.T. and Jaeger, K.-E. (2004) Directed evolution by random mutagenesis: a critical evaluation. In: Enzyme functionality: design, engineering and screening. Svendsen, A. (ed.). New York: Marcel Dekker; 375-390.

(30)   Eggert, T. (2006) Optimization of industrial enzymes by molecular engineering. In: Industrial Biotransformations, 2nd edition. Liese, A., Seelbach, K., Wandrey, C. (eds.). Wiley-VCH, Weinheim; 93-114.

(31)   Wendorff, M., Eggert, T., Pohl, M., Dresen, C., Müller, M. and Jaeger, K.-E. (2006) Evolving the substrate specificity of benzoylformate decarboxylase from *Pseudomonas putida*. In: Asymmetric synthesis with chemical and biological methods. Enders, D. and Jaeger, K.-E. (eds.) Wiley-VCH, Weinheim. 298-311.

Patent applications

(32)   Michels, A., Pütz, A., Maurer, K.-H., Eggert, T. and Jaeger, K.-E. (2005) Verwendung von Esterasen zur Spaltung von Kunststoffen. Patent Application DE 10 2005 037 659.2

(33)   Eggert, T., Drepper, T., Guterl, J.-K., Heck, A., Krauß, U. and Jaeger, K.-E. (2005) Fluoreszenzmarker und dessen Verwendung. Patent Application DE 10 2005 048 828.5