Protein Engineering, Design & Selection vol. 22 no. 7 pp. 413–420, 2009
Published online June 9, 2009 doi:10.1093/protein/gzp019

Directed evolution of *Candida antarctica* lipase A using an episomaly replicating yeast plasmid†

Anders G. Sandström, Karin Engström, Jonas Nyhlén, Alex Kasrayan and J.-E. Bäckvall[1]

Department of Organic Chemistry, Arrhenius Laboratory, Stockholm University, SE-106 91 Stockholm, Sweden

†Dedicated to Professor Karl Hult on the occasion of his 65th birthday.

[1]To whom correspondence should be addressed.
E-mail: jeb@organ.su.se

We herein report the first directed evolution of *Candida antarctica* lipase A (CalA), employing a combinatorial active-site saturation test (CAST). Wild-type CalA has a modest *E*-value of 5.1 in kinetic resolution of 4-nitrophenyl 2-methylheptanoate. Enzyme variants were expressed in *Pichia pastoris* by using the episomal vector pBGP1 which allowed efficient secretory expression of the lipase. Iterative rounds of CASTing yielded variants with good selectivity toward both the (*S*)- and the (*R*)-enantiomer. The best obtained enzyme variants had *E*-values of 52 (*S*) and 27 (*R*).

Keywords: *Candida antarctica*/directed evolution/hydrolases/*Pichia pastoris*/yeast vector

Introduction

A major part of the drugs manufactured today contains enantiopure molecules (Beck, 2002); hence, highly enantioselective reactions for the production of enantiopure building blocks are of great industrial importance. Enzymes are an attractive class of catalysts often used in the synthesis of enantiopure compounds. They usually exhibit high enantioselectivity, operate under mild reaction conditions and have a large substrate scope (Faber, 2000; Kirk *et al*., 2002). However, with substrates that are very different from the natural substrate, enzymes can display low enantioselectivity and/or poor reactivity. A solution to this problem is to genetically modify the enzyme and thereby increase the substrate acceptance or enantioselectivity.

Directed evolution of enzymes has attracted considerable interest and during the past years, several powerful methods have been developed (Reetz *et al*., 1997; Brackmann and Johnsson, 2002; Arnold and Georgiou, 2003; Otten and Quax, 2005). Combinatorial active-site saturation test (CAST) is a method based on semi-rational selection of two or more amino acids surrounding the catalytic site and simultaneous randomisation of these residues. Residues that are located in close sequence proximity are selected, so that only one primer pair is needed for each site-directed mutagenesis (Reetz *et al*., 2005; Reetz and Carbelleira, 2007). By iteratively saturating different parts of the active site area, synergistic effects can be anticipated (Reetz and Sanchis, 2008). The CASTing methodology has previously proven to be successful in the directed evolution of *Pseudomonas aeruginosa* lipase and *Aspergillus niger* epoxide hydrolase (Reetz *et al*., 2006a; Carballeira *et al*., 2007). The bottleneck of directed evolution is often the time-demanding screening of enzyme variants (Zhao and Arnold, 1997), but since the CASTing method is concentrated to a limited set of amino acids around the active site, it has a higher hits-to-samples ratio in comparison to other methods, e.g. error prone PCR. The library sizes can therefore be reduced (Reetz *et al*., 2006b).

Candida antarctica lipase A (CalA) is an eukaryotic lipase derived from the yeast *C.antarctica* (Høegh *et al*., 1995). CalA has shown excellent activity towards large substrates (Kirk and Christensen, 2002), and has also displayed enantioselectivity towards tertiary alcohols (Henke *et al*., 2002). CalA has been used for the preparation of β-amino acids with excellent enantioselectivity (Gedey *et al*., 2001; Liljeblad and Kanerva, 2006). Furthermore, this lipase is highly thermostable and has intriguing sn-2 preference towards triglycerides (Dominguez de Maria *et al*., 2005).

CalA is a serine hydrolase, belonging to the α/β-hydrolase family, with a typical catalytic triad. The catalytic triad is composed of the nucleophilic Ser184, His366 and Asp334 [The authors would like to point out in the sake of clarity, that in a couple of papers (Kasrayan *et al*., 2007; Ericsson *et al*., 2008), the numbering of amino acid residues is numbered +10 in comparison to the original Novozyme cloning publication (Høegh *et al*., 1995). The +10 numbering of residues is also used in this article.]. An uncommon oxyanion hole stabilizer, the acidic residue Asp95, is also observed in the X-ray structure of CalA (Ericsson *et al*., 2008).

CalA has previously been expressed in *A.oryzae*, *Escherichia coli*, as well as in *Pichia pastoris* using the pPICZ vector (Høegh *et al*., 1995, Pfeffer *et al*., 2006; 2007; Kasrayan *et al*., 2007). While *E.coli* is a capable expression system, as the growth is quick and overproduction methods are well established, it has drawbacks when used to express eukaryotic proteins; the inability to handle post-translational processes such as glycosylation and signal peptide cleavage can often pose a problem (Cregg *et al*., 2000). Another drawback of *E.coli* is the laborious and often activity-reducing lysis protocols that have to be applied in order to liberate the enzymes.

The yeast *P.pastoris* is steadily becoming more attractive as an expression system, especially for eukaryotic proteins (Daly and Hearn, 2005). The yeast can grow to high cell densities under proper conditions, which allows high enzyme production. The ability to secrete the protein by utilizing the α-mating signal from *Saccharomyces cerevisiae* is highly advantageous; this significantly facilitates the means of harvesting the proteins (Cereghino *et al*., 2002). The advantages of *P.pastoris* for high-throughput screening have been further enhanced by the design of the constitutively expressing episomaly replicating plasmid pBGP1 (Lee *et al*., 2005). The pBGP1 plasmid contains a GAP promoter, an α-mating signal and the PARS1 sequence for episomal replication (Cregg *et al*., 1985; Waterham *et al*., 1997). Compared to AOX1-based genomically integrating vectors such as pPICZ,

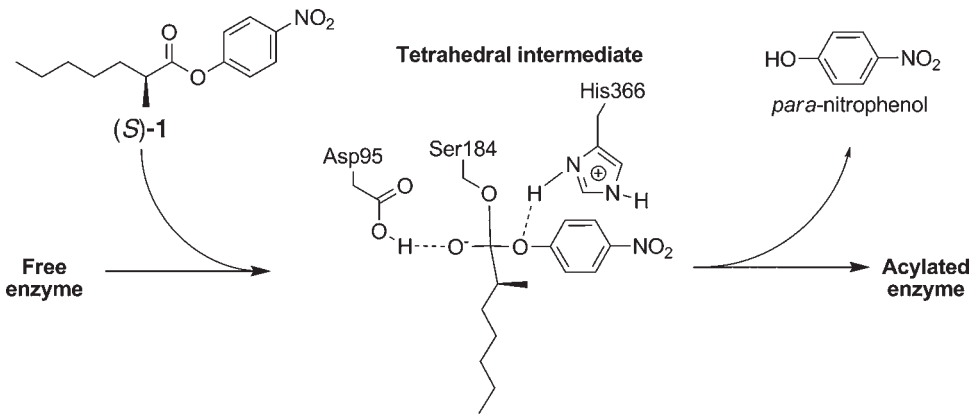

Fig. 1. An abbreviated reaction mechanism for the formation of the tetrahedral intermediate of (*S*)-**1.** The model used for the selection of the amino acid residues for CASTing had this intermediate docked in the active site. The release of *para*-nitrophenol can be monitored spectrophotometrically at 410 nm.

the episomaly replicating plasmids have less variable expression levels, which is important for well-to-well reproducibility. Liu *et al*. recently showed the straightforwardness of directly integrating mutagenic library PCR-products, although unwanted integration effects were observed (Liu *et al*., 2008). Episomaly replicating plasmids are in comparison conveniently recovered by commercially available yeast plasmid preparation kits. The drawback of episomaly replicating yeast plasmids is that continuous antibiotic selection is necessary as PARS1-containing plasmids lack long-term replication stability (Hong *et al*., 2006).

In this paper, we report a directed evolution study of CalA towards 4-nitrophenyl 2-methylheptanoate (**1**) (Fig. 1). The wild-type CalA was found to have a low enantiopreference for the (*S*)-enantiomer of **1,** the *E*-value being 5.1. CASTing and the episomaly replicating yeast vector pBGP1 was used for the undemanding creation of the expression libraries. The release of *para*-nitrophenolate from the respective enantiomer was followed spectrophotometrically to screen for interesting enzyme variants.

Experimental section

General equipment and chemicals

All chemicals, unless otherwise mentioned, were purchased from Sigma-Aldrich. 2-Methylheptanoic acid was purchased from Alfa Aesar. All enzymes, unless otherwise mentioned, were purchased from Fermentas. Phusion DNA polymerase was purchased from Finnzymes. Plasmid preparation kits, Cycle-pure kit and Yeast plasmid kit was purchased from Omega-BioTek. Amicon centrifugal filter devices were purchased from Millipore. The used electroporator was a Bio-Rad Micropulser. Microtiter plates from Corning were used. Zeocin was purchased from Invivogen. Media ingredients came from Becton, Dickinson and Company. Spectrophotometric readings were performed on a Synergy 2 from BioTek. ^{1}H and ^{13}C NMR spectra were recorded at 400 MHz. GC analyses were performed on a Varian GC 3900 using an IVADEX-1 chiral column from IVA analysentechnik.

Media recepies

LB medium (10 g l^{-1} tryptone, 5 g l^{-1} yeast extract, 5 g l^{-1} NaCl, pH 7.0) was used for bacterial cultivation. YPD medium (10 g l^{-1} yeast extract, 20 g l^{-1} peptone, 20 g l^{-1} dextrose) was used for *P.pastoris* cultivation and expression. YPDS (10 g l^{-1} yeast extract, 20 g l^{-1} peptone, 20 g l^{-1} dextrose, 1 M sorbitol) was used for recently transformed *P.pastoris*.

pBGP1-CalA vector construction

The construction of the pPICZα-CalA vector containing two potential signal peptide cleavage sites has been described previously (Kasrayan *et al*., 2007; Ericsson *et al*., 2008). The pPICZα-CalA plasmid and the pBGP1 plasmid were double digested with FastDigest XbaI & FastDigest XhoI restriction enzymes. The pBGP1 fragment was dephosphorylated, and the fragments of pBGP1 and CalA were gel-purified, ligated and transformed into chemo-competent *E.coli* DH5α. The vector sequence was validated by sequencing (sequence is included in the Supplementary data, available at *PEDS* online). The pBGP1 shuttle vector was a kind gift of Charles C. Lee (Lee *et al*., 2005).

Preparation of combinatorial libraries

Amino acid residues surrounding the nucleophilic Ser184 were evaluated to see whether they could influence the formation of the tetrahedral intermediate (Fig. 1). The selection of the targeted amino acid residues is explained in Fig. 2. All mutagenic libraries were created by site-directed mutagenesis. The PCRs were performed with the following protocol: a pre-incubation was performed at 95°C for 10 min, after which DNA polymerase was added. The cycle of denaturation at 98°C for 10 s, annealing at 58–68°C for 30 s and elongation at 72°C for 3 min was repeated 30 times. A 10 min elongation at 72°C ended the reaction. The PCR products were purified using Cycle-pure kit followed by DpnI treatment to remove wild-type template. The nicked plasmid was transformed into DH5α and a large fraction of the transformed cells were grown overnight for a medium-sized plasmid preparation. In parallel was a fraction of the transformed cells plated, to prepare them for a sequencing assay. Sequencing was performed on 6–10 colonies per library, to verify expected diversity (Eurofins MWG Operon). All pBGP1-carrying DH5α were selected using carbenicillin (50 μg ml^{-1}) in LB and LB-agar plates.

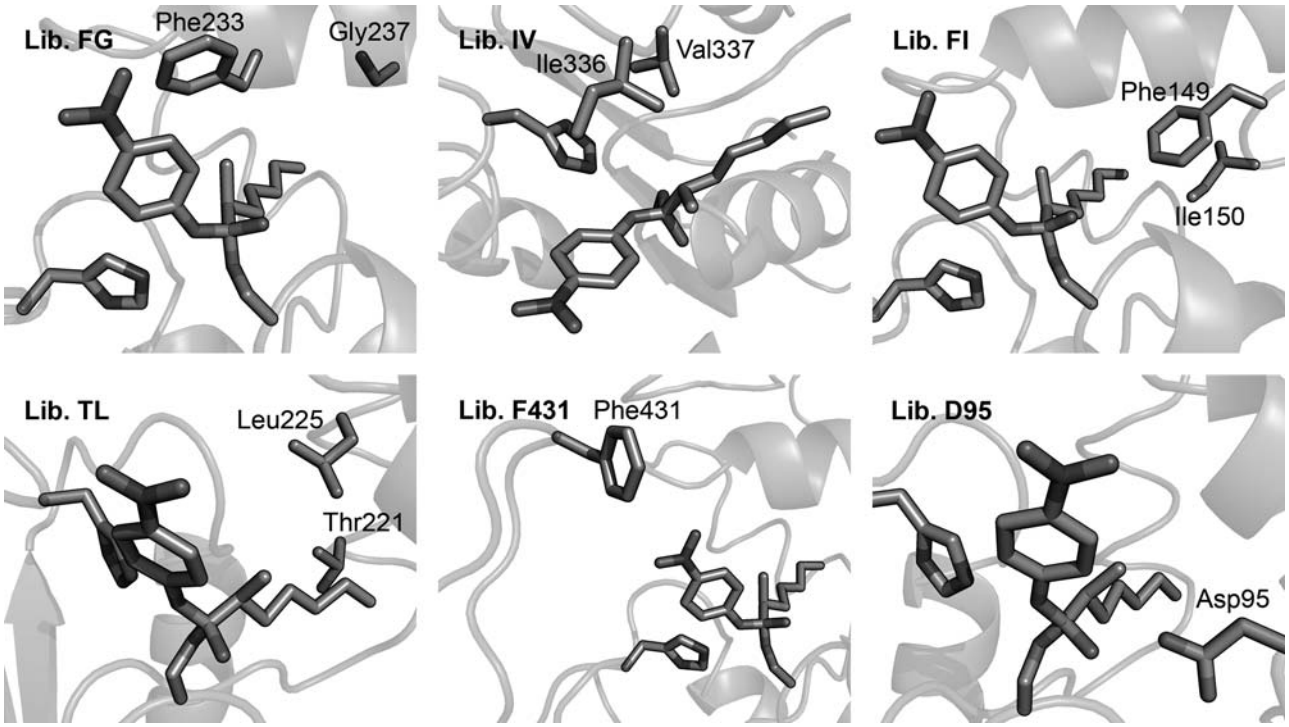

Fig. 2. The active site with the catalytic residues His366 and Ser184 is displayed in each panel. The nucleophilic Ser184 is bound to the tetrahedral intermediate form of (*S*)-**1.** The initially constructed libraries were directed at the following areas; **Library FG,** the sterically interesting tandem pair, Phe233/Gly237, which is lining the substrate cavity entrance (Gly237 lack side chain, so the backbone is displayed for clarity); **Library IV,** the conserved, unpolar Ile336/Val337, which are located on an inward bent loop, pointing in the direction of the α-methyl group of (*R*)-**1**; **Library FI,** the sterically large, hydrophobic Phe149/Ile150, which point in the direction of the α-methyl group of (*S*)-**1**; **Library TL,** an area with more outward pointing residues, Thr221/Leu225; **Library F431,** the phenylalanine that is located on the lid-like C-terminal flap, Phe431, which is assumed to play a crucial role for interfacial activation; **Library D95,** Asp95, a residue suggested to contribute to the oxyanion hole stabilisation. The model was generated by molecular dynamics (see Supplementary data, available at *PEDS* online). A colour version of this figure is published as Supplementary data, available at *PEDS* online.

Primers

Primers were ordered from Eurofins MWG Operon.

The following primers were used for the mutagenesis (All are written 5′–3′):

Library FI (F149NDTI150NDT):
FI2_fw GGCTTCAAAGCCGCCNDTNDTGCTGGCTACGAAG
FI2_rv CTCTTCGTAGCCAGCAHNAHNGGCGGCTTTGAAG
Library FG (F233NDT/G237NDT):
LibFG_fw GACCCTTCGCCGGCNDTGCCCTGGCGNDTGTTTCGGGTC
LibFG_rv GAGAGACCCGAAACAHNCGCCAGGGCAHNGCCGGCGAAG
Library IV (I336NDT/V337NDT):
LibIV_fw GCGATCCCCGACGAGNDTNDTCCGTACCAGCC
LibIV_rv CAGGCTGGTACGGAHNAHNCTCGTCGGGGATC
Library TL (T221NYY, L225NYY):
LibTL_fw GTGAGCGCCAAGGACNYYTTTACATTCNYYAACGGCGGACCC
LibTL_rv GAAGGGTCCGCCGTTRRNGAATGTAAARRNGTCCTTGGCGCT
Library F431 (F431NNK):
LibF431_fw ACGGCAAGCAGAGTGCGNNKGGCAAGCCC
LibF431_rv AGGGCTTGCCMNNCGCACTCTGCTTGCCGTTGAGGC
Library D95 (D95NNK):
LibD95_fw
TACCAGGTCTACGAGNNKGCCACGGCGC
LibD95_rv AGCGCCGTGGCMNNCTCGTAGACCTGGTACG
Tyr93Phe:
Y93F_fw TCGTACCAGGTCTTTGAGGATGC
Y93F_rv CCGTGGCATCCTCAAAGACCTGG

Transformation of Pichia pastoris

Pichia pastoris X33 (Invitrogen) was made electrocompetent by adhering strictly to the manufacturer's protocol (BioRad Micropulser protocol). Until employed, cells were stored in 40 μl aliquots at −80°C without any treatment. Library plasmid preparation was mixed with thawed cells, and electroporated. Cells were incubated at 30°C with YPDS (1 ml) for 2 h followed by plating on YPDS-agar plates containing zeocin (100 μg ml^{-1}) and carbenicillin (100 μg ml^{-1}). Colonies usually appeared after 3 days of incubation at 30°C.

Cultivation of yeast libraries

Single colonies were picked and inoculated in conical deep 96-well plates. Each well contained YPD (800 μl), zeocin (100 μg ml^{-1}) and carbenicillin (100 μg ml^{-1}). The deep well plates were shaken at 250 rpm for 96 h at 29°C, leaning approximately 30°. The shaking amplitude was 2.5 cm. The

yeast was pelleted after overexpression by centrifugation of the deep well plates. The supernatant was harvested by aspiration and used directly for optical screening or stored at 4°C. Master plates containing the pelleted cells were stored at −80°C until further analysis or cultivation.

For the comparison of expression and growth of enzyme variants, over night cultures of strains were grown in YPD containing zeocin (100 mg ml^{-1}) and carbenicillin (100 mg ml^{-1}). These cultures were diluted to an OD_{600} of 0.1, in 50 ml YPD containing zeocin (100 mg ml^{-1}) and carbenicillin (100 mg ml^{-1}) and were grown in 250 ml Ehrlenmayer flask with a cotton-fibre stopper and shaken at 250 rpm (shaking amplitude 2.5 cm) at 29°C during 96 h. Samples were taken each 24 h. OD_{600} was measured, and cells were pelleted and supernatant was aspirated to quantify secreted expression levels. Bio-Rad Protein Assay was used to quantify protein yield, using BSA as a standard. All measurements were made in doublets.

Optical screening

Screening buffer (165 μl, 100 mM potassium phosphate, 10% v/v acetone, 4% v/v Triton X-100, 0.2% w/v gum arabic, pH 8.0) (Schultz *et al*., 2007) and (*R* or *S*)-**1** (10 μl, 2 mg ml^{-1} in acetonitrile) were pre-mixed before dispension. Yeast library supernatant (25 μl) and the pre-mixed buffer and substrate solution (175 μl) were dispensed into a microtiter plate, and absorbance was measured at 410 nm for 15 min. The ratio between the initial rates of (*S*)- and the (*R*)-enantiomer was calculated and plotted against the V_{max} of the fast reacting enantiomer. Variants that showed high ratio and acceptable activity were chosen for a kinetic resolution experiment. All samples were measured in duplicates.

Kinetic resolution

Potassium phosphate-buffer (850 μl, 100 mM, pH 8.0), *rac*-1 (125 μl, 2 mg ml^{-1} in acetonitrile) and enzyme supernatant (25 μl) were mixed in a 2 ml glass vial. The reaction was shaken at 30°C, until the reaction had reached 10–30% conversion. The reaction mixture was acidified to a pH below 3 by addition of HCl (10%) and the products were extracted with CH_2Cl_2. The solvent was evaporated and the product was dissolved in toluene (20 μl). The enantiomeric excess was determined by GC. The reported *E*-values were based on an average of three independent reactions. (Experimental data is included as Supplementary data, available at *PEDS* online.)

Kinetic measurements

Fresh enzyme batches were purified by hydrophobic interaction chromatography as previously described (Kasrayan *et al*., 2007). The enzymes were concentrated in Tris–HCl buffer (25 μl, 100 mM, pH 7.5) by a Amicon Ultra 15 ml 10 K centrifugal column. Substrate solutions of enantiomerically enriched **1** ((*S*)-**1,** 88% ee; (*R*)-**1,** 98% ee) were prepared in the range 0.1–20 mg ml^{-1} in acetonitrile. Substrate solution (20 μl) was mixed with screening buffer (155 μl) (*vide supra*) in a microtiter plate. Reaction was started by addition of enzyme solution (25 μl), and absorbance was measured at 410 nm. Blanks were performed using Tris–HCl buffer (25 μl, 100 mM, pH 7.5) instead of enzyme solution. Reaction progressions were monitored at 30°C, for at least 30 min, and initial rates were calculated. The absorbance was correlated to a standard curve for *para*-nitrophenol which was performed under identical assay conditions. The apparent k_{cat} and K_m values were calculated by a direct fit to the Michaelis–Menten plot based on the experimental initial rates.

Sequencing

Interesting enzyme variants from the screening were sequenced. Pelleted cells from master plates were used to inoculate YPD containing zeocin (100 mg ml^{-1}) and carbenicillin (100 mg ml^{-1}), and cultures were shaken at 30°C, usually for 48 h or until high optical density was attained. The plasmid was then extracted using a Yeast plasmid kit, and subsequently transformed into DH5α to obtain a higher plasmid yield. The bacterial cells were cultivated, and plasmid preparations were produced and sequenced using appropriate primers.

General procedure for synthesis of 4-nitrophenyl 2-methylheptanoate

2-Methylheptanoic acid (1.00 ml, 6.42 mmol), DMAP (78 mg, 0.64 mmol) and Et_3N (0.94 ml, 6.74 mmol) were dissolved in dry CH_2Cl_2 (10 ml) under argon and stirred at 0°C for 15 min. 4-Nitrophenyl chloroformate (1.28 g, 6.36 mmol) dissolved in CH_2Cl_2 (2 ml) was added to the reaction mixture and was stirred at 0°C for 2 h. Reaction mixture was diluted with CH_2Cl_2, then extracted with HCl (0.1 M), $NaHCO_3$ (1 M) and brine, thereafter dried over $MgSO_4$ and concentrated. Purification was performed by flash column chromatography (Pentane/EtOAc 15:1) to give the product in yields between 61 and 92%. ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 0.91$ (t, $J = 7.0$ Hz, 3H), 1.31 (d, $J = 6.8$ Hz, 3H), 1.31–1.46 (m, 6H), 1.52–1.62 (m, 1H), 1.76–1.86 (m, 1H), 2.72 (m, 1H), 7.26 (m, 2H), 8.27 (m, 2H) ppm. ^{13}C NMR (400 MHz, $CDCl_3$): $\delta = 14.1$, 17.0, 22.6, 27.0, 31.8, 33.7, 39.9, 122.5, 125.3, 145.4, 155.9, 174.5 ppm.

Resolution of 2-methylheptanoic acid

2-Methylheptanoic acid (6.48 g, 45 mmol), 1-pentanol (3.7 ml, 33.75 mmol) and Na_2SO_4 (19.2 g, 135 mmol) were dissolved in cyclohexane (150 ml) at room temperature. After 15 min, *C.rugosa* lipase (1.0 g) was added and the reaction mixture was stirred at room temperature for 3 days. Solid material was removed by filtration. The filtrate was extracted five times with 1 M Na_2CO_3 and once with brine, dried over $MgSO_4$ and concentrated to give (*S*)-pentyl 2-methylheptanoate (2.95 g, 13.8 mmol). The obtained ester was dissolved in ethanol (20 ml) together with 2 M sodium hydroxide (6.9 ml, 13.8 mmol) and heated at reflux for 4 h. Reaction mixture was then acidified to pH 1 and extracted with CH_2Cl_2. The combined organic phases were washed with brine, dried with $MgSO_4$ and concentrated to give (*S*)-2-methylheptanoic acid (1.80 g, 12.5 mmol, 88% ee).

The combined Na_2CO_3 phases obtained above were acidified to pH 1 with 6 M HCl while cooled in an ice-bath and thereafter extracted with CH_2Cl_2. The combined organic phases were washed with brine, dried over $MgSO_4$ and concentrated to give (*R*)-2-methylheptanoic acid (4.40 g, 30.5 mmol, 40% ee). To improve the enantiomeric excess, the obtained acid was used in a second kinetic resolution, but now running it beyond 50% conversion. The procedure was

the same as above, but with the following work-up protocol. Solid material was removed by filtration. Filtrate was extracted five times with 1 M Na_2CO_3. Combined Na_2CO_3 phases were acidified to pH 1 with 6 M HCl while cooled in an ice-bath, and then extracted with CH_2Cl_2. The combined organic phases were washed with brine, dried over $MgSO_4$ and concentrated to give (*R*)-2-methylheptanoic acid (1.99 g, 13.8 mmol, 98% ee).

Results and discussion

Design of the pBGP1-CalA vector

The CalA gene was derived from a previously constructed pPICZ-CalA vector (Ericsson *et al*., 2008), and was transferred to the pBGP1 shuttle vector by standard molecular biology techniques. The pBGP1 plasmid was convenient to use in the library expression process, as restreaks of *P.pastoris* transformed with this plasmid was not needed prior to overexpression. Protein expression levels were less variable between mutant clones when this plasmid was used. The possibility to extract the plasmid from the yeast for subsequent rounds of mutagenesis also made it a convincing choice for enzyme expression.

Combinatorial active-site saturation test (CAST)

The selections of amino acids for the saturation mutagenesis were based on the X-ray structure of CalA (Ericsson *et al*., 2008). A model of the tetrahedral intermediate using *(S)*-**1** as substrate was created (Fig. 1). Amino acid residues surrounding the nucleophilic serine (Ser184) were selected for the CASTing (Fig. 2). The rationale for selecting these residues is explained in Fig. 2. Some regions further into the acyl chain pocket were also chosen. An approach of targeting two amino acid positions in tandem was used. Reduced libraries were employed in order to decrease the theoretical number of enzyme variants. NDT degeneracy, which codes for 12 amino acids (CDFHGILNRSVY), was chosen for the majority of the created libraries. This decreased the colony sampling requirements to a more manageable level (Reetz *et al*., 2008; Clouthier *et al*., 2006). The number of picked colonies was enough to obtain at least 95% theoretical codon coverage; 600 colonies were picked for each of the tandem libraries. For a couple of single-site saturation tests, NNK degeneracy (all 20 amino acids) was used.

Site-directed mutagenesis was performed using the pBGP1-CalA vector as template to obtain combinatorial DNA libraries. These libraries were amplified in *E.coli*, and larger plasmid preparations were obtained. Diversity of the plasmid preparations was validated by sequencing of *E.coli* colonies that had been transformed in parallel. The prepared plasmid preparations were transformed into *P.pastoris* by electroporation. Colonies were picked and grown in 96-deep well plates. The enzyme variants were expressed and harvested from the yeast supernatant.

Screening of lipase variants

As a first part of the screening, an optical pre-screen was applied to find enzyme variants with improved enantioselectivity. Separated enantiomers of (*R*)-**1** and (*S*)-**1** were hydrolysed by the enzyme in parallel reactions, and the increase of *para*-nitrophenolate (at pH 8, the anionic form of *para*-nitrophenol) was monitored spectrophotometrically (Fig. 1). The ratio of the initial rates between the two enantiomers was calculated. The enzyme variants with the highest ratio were selected for a kinetic resolution of *rac*-**1,** and the reactions were analysed by gas chromatography to obtain *E*-values. (Sih and Wu, 1989).

The most interesting variants found in the initial round originated from Library FG. This library yielded several selective variants towards both *(R)*-**1** and *(S)*-**1.** These results are assumed to be caused by the modification of the sterical space by exchanging this interesting tandem pair. The best *(R)*-**1** selective enzyme variant had an *E*-value of 27, and had the amino acid exchanges Phe233Leu and Gly237Tyr. The best (*S*)-**1** selective CalA variant found in the initial round had three altered amino acids; Thr64Met, Phe233Asn and Gly237Leu. This variant had an *E*-value of 19. The Thr64Met substitution is assumed to have been caused by a misincorporation during the PCR step. The Met64 is located at the surface, pointing outwards, at a distance of 22 Å away from the active site. Thus, it is assumed that this residue exchange has no influence on the enantioselectivity of the enzyme variant.

The best *(S)*-**1** selective enzyme was used in a second round of CASTing, using the same primers as for the initial libraries. After two rounds of CASTing, we could obtain an (*S*)**-1** selective enzyme with an *E*-value of 52. This variant had the following amino acid substitutions; Thr64Met/Phe149Ser/Ile150Asp/Phe233Asn/Gly237Leu. This variant was obtained using the primers for Library FI. In the initial round, Library FI gave quite modest effects on enantioselectivity indicating a synergistic effect caused by the amino acid exchanges in the second generation variant (Table I). In contrast, the second generation variants based on *(R)*-**1** selective F233L/G237Y did not show any further improvement of the enantioselectivity.

Several attempts to generate functional variants based on Library IV yielded only weakly active enzymes. By examining the model of the substrate-bound enzyme, it appears that both these amino acid residues point in the direction of the α-methyl group of (*R*)-**1** (Fig. 2). It was therefore assumed that this library would have an impact on enantioselectivity. We identified that the targeted amino acids Ile336 and Val337 are exceptionally conserved (Ericsson *et al*., 2008). These residues are located close in sequence space to Asp334, the acid component of the catalytic triad. The strict conservation of these amino acids is an indication of their structural importance.

Table I. Key CalA variants developed in this work

Designation	Library origin	*E*-value
Wild type	–	5.1 (*S*)
J3E7	LibFI	7.8 (*S*)
F431:1E2[a]	LibF431	9.9 (*S*)
F431:2E3[a]	LibF431	9.1 (*S*)
F431:2B12[a]	LibF431	10 (*S*)
F233L/G237Y	LibFG	27 (*R*)
T64M/F233N/G237L	LibFG	19 (*S*)
T64M/F149S/I150D/F233N/G237L	LibFG/LibFI	52 (*S*)

[a]Sequences included in Supplementary data, available at *PEDS* online.

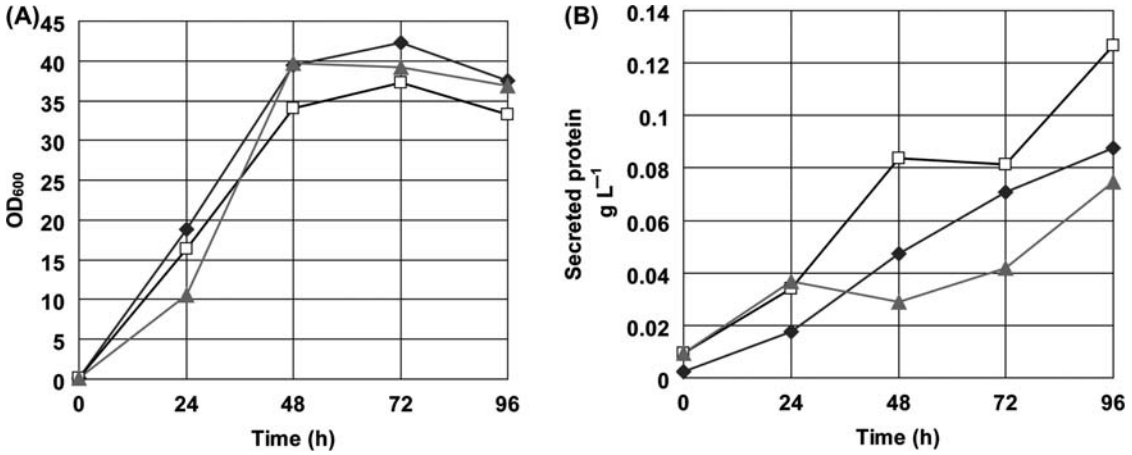

Fig. 3. Growth and expression for CalA variants. (**A**) The OD600 of the growing cultures, and (**B**) the secreted protein expression in cleared supernatants. CalA-Wild type (diamonds), CalA-T64M/F149S/I150D/F233 N/G237L (open squares) and CalA-F233L/G237Y (triangles) were grown for 96 h, and samples were taken every 24 h.

Growth and expression

To examine whether the most selective enzyme variants affected enzyme expression and growth of yeast cultures, growth and expression assays were performed in 50 ml YPD media (Fig. 3.). *Pichia pastoris* pBGP1 CalA expression levels ($\sim$0.1 g l^{-1}, 96 h) are definitely lower compared to pPICZ-based CalA expression systems ($\sim$0.6 g l^{-1}, 5 days, in shake flasks) (Pfeffer *et al*., 2006). The advantage with the pBGP1 vector is that it requires no extra feeding and that it can be extracted and easily reused for new rounds of mutagenesis. The selective enzyme variants cause apparently no impairment on growth. The expression levels are continuously increasing during growth, with some variation between the enzyme variants. The enzyme expression in 96-deep well plates could be expected to reach similar concentrations as in 50 ml cultures, yet approximately 0.2 g l^{-1} enzyme was the norm after 96 h of cultivation, to the large extent caused by evaporation of water from the supernatant.

Kinetic measurements and structural analysis of enantioselective variants

CalA wild-type and selective variants were further analysed by measurement of their apparent kinetic constants. The initial rates were used as a template for a non-linear regression of the kinetic constants to fit into the Michaelis–Menten equation (Table II). The selective variants have similar turnover rate as the wild type for their respective preferred enantiomer. They also show increased substrate specificity for the preferred enantiomer, via a lowering of K_m. Apparently, the modifications of the active site do not disturb the fine catalytic machinery in either selective variant. The catalytic efficiency ratio of the two enantiomers ($(k_{cat}/K_m)_{fast}/(k_{cat}/K_m)_{slow}$) correlates with the *E*-values obtained from kinetic resolution (Table I).

The amino acid residue substitutions introduced in the (*S*)-**1** selective T64M/F149S/I150D/F233N/G237L variant have caused the acyl chain pocket to turn into a more spacious cavity (Fig. 4). The most important contribution comes from the Phe149Ser and Phe233Asn substitutions for the widening of the substrate cavity. Molecular dynamics simulations indicate that a hydrogen bond is formed between Asn233 and Asp150, as well as between Ser149 and Asp150. The cavity has also become distinctly more hydrophilic. The α-methyl group of (*R*)-**1** might be repelled by the polarity of the introduced residues, while (*S*)-**1** can align itself in the substrate pocket without being repelled. The substrate binding of (*S*)-**1** is particularly favoured in this variant, whereas the turnover rate is actually slightly lower than that for the wild type.

Table II. Apparent kinetic constants for CalA variants with enantiomers of **1** as substrates

Enzyme variant	Substrate	k_{cat} (s^{-1})	K_m (μM)	k_{cat}/K_m (s^{-1} M^{-1})	$(k_{cat}/K_m)_{fast}/(k_{cat}/K_m)_{slow}$
CalA-WT	**1**-(*S*)	1.83	2460	744	4.3
	1-(*R*)	0.48	2700	174	
CalA-T64M/F149S/	**1**-(*S*)	1.43	1060	1350	28.1
I150D/F233 N/G237L	**1**-(*R*)	0.13	2690	48	
CalA-F233L/G237Y	**1**-(*S*)	0.34	4200	80	20
	1-(*R*)	2.60	1630	1600	

In the (*R*)-**1** selective variant, F233L/G237Y, it can be observed that the acyl chain is taking a slightly twisted turn due to the partial blockage of the original acyl chain pocket by the large Tyr237. The inversion of enantiopreference is caused both by a relaxation of substrate binding (higher K_m) and decreased turnover rate for (*S*)-**1,** and at the same time a tighter substrate binding for (*R*)-**1.** For both the (*S*)- and the (*R*)-selective CalA variants, catalytic efficiency has increased by synergistic effects as both catalytic rate and substrate binding have changed.

Single site saturation test

The conserved residue Asp95 has been further studied, and appears to be the oxyanion hole stabilizer (Ericsson *et al*., 2008). A complete saturation test led to not a single functional enzyme variant, except for recovered wild-type variants. This shows that Asp95 is critical for the catalytic activity. Based on structure comparison (Ericsson *et al*., 2008), it was indicated that the active site of CalA had some similarities with peptidase DPP-IV, which is suggested to

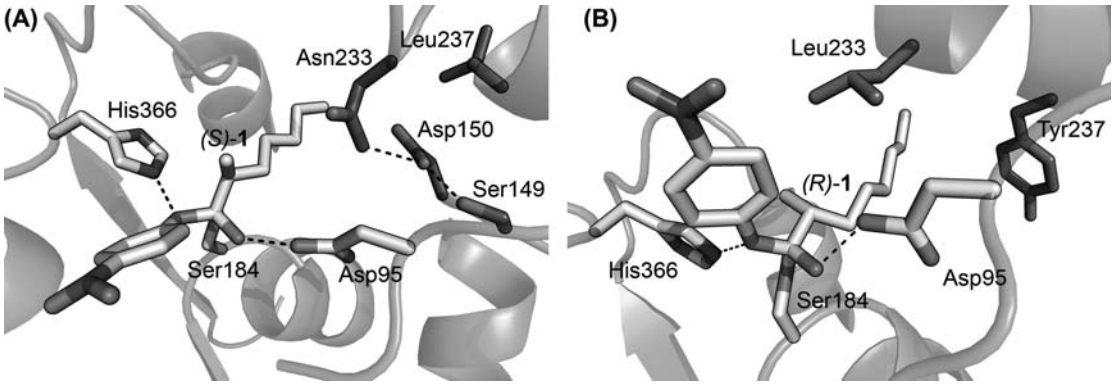

Fig. 4. Models displaying the active site of (**A**) the (*S*)-**1** selective, and (**B**) the (*R*)-**1** selective CalA variants. The catalytic serine (Ser184), the catalytic triad participant His366 and the oxyanion hole stabilizer Asp95 are displayed in each panel. The preferred enantiomer of **1,** bound to Ser184, is displayed in its tetrahedral intermediate form. The new sets of amino acid residues are displayed in dark grey. The models were generated by molecular dynamics (see Supplementary data are available at *PEDS* online). A colour version of this figure is published as Supplementary data, available at *PEDS* online.

have a tyrosine coordinating to the oxyanion of the tetrahedral intermediate (Bjelke *et al*., 2004). The mutation Tyr93Phe was carried out to confirm that Tyr93 does not contribute to the catalytic activity. The Tyr93Phe variant demonstrated fully retained activity, thus indicating that it is not a participant in the catalytic mechanism (results not shown).

The conserved, hydrophobic Phe431 is located on the C-terminal flap (Ericsson *et al*., 2008). This flap is implied to be the flexible part which contributes to the observed interfacial activation of CalA (Martinelle *et al*., 1995). The affinity for this flap to swing out and make contact to external hydrophobic environments may depend to a large degree on Phe431. This residue may also be of importance to the catalytic activity, as it contributes with hydrophobicity to the active site during substrate acceptance. The F431 library was generally inactive, which indicates that Phe431 is indeed a crucial component for the activity of the enzyme.

Intriguingly, even though most of the variants from the F431 library were dysfunctional, a few variants with increased enantioselectivity toward *(S)*-**1** were found. All of these had PCR-induced abnormalities in the gene. Two of the sequences had longer open reading frames compared with the wild type, and one of the sequences had a premature termination caused by the introduction of a stop codon. The activities of these variants were comparable to that of the wild type (sequences included in Supplementary data, available at *PEDS* online).

Conclusions

We have performed a directed evolution of the thermostable enzyme *C.antarctica* lipase A directed at the acyl chain binding pocket. Enzyme variants with high enantioselectivity towards both *(R)*- and *(S)*-4-nitrophenyl 2-methylheptanoate have been found. Similar α-methyl acid may form basis for many other more interesting substrates with similar structural features but with functionalisations such as unsaturations on the acyl chain. The developed variants have shown their usefulness as expression and growth is not compromised, and as well show retained or increased catalytic efficiency (k_{cat}/K_m) for the desired enantiomers. The developed enzyme variants may also form starting points for further developments of enzymes with increased enantioselectivity towards other more sterically hindered chiral carboxylic acids. Furthermore, we demonstrate the advantage of using the pBGP1 vector in *P.pastoris* for heterologous expression in directed evolution experiments. The constitutive expression facilitates the handling and downstream processing of medium-throughput assays. Reduced amino acid sets were utilized to keep the library sizes small. All these factors contribute to an overall undemanding directed evolution approach of eukaryotic proteins.

Acknowledgements

We thank Professor Charles Lee for his invaluable gift of the pBGP1-plasmid, and Dr José Daniel Carballeira and Dr Marco Bocola for support and suggestions. The Department of Biochemistry and Biophysics, Stockholm University is gratefully acknowledged for allowing us to use some of their equipment.

Funding

This work was supported by the Swedish Research Council (VR), the Swedish Foundation for Strategic Research (SSF) and the K&A Wallenberg Foundation.

Supplementary data

Supplementary data are available at *PEDS* online.

References

Arnold,F.H. and Georgiou,G. (eds) (2003) *Methods in Molecular Biology*. Humana Press.

Beck,G. (2002) *Synlett*, **6**, 837–850.

Bjelke,J.R., Christensen,J., Branner,S., Wagtmann,N., Olsen,C., Kanstrup,A.B. and Rasmussen,H.B. (2004) *J. Biol. Chem.*, **279**, 34691–34697.

Brackmann,S. and Johnsson,K. (eds) (2002) *Directed Molecular Evolution of Proteins: Or How to Improve Enzymes for Biocatalysis*. Wiley-VCH, Weinheim.

Carballeira,J.D., Krumlinde,P., Bocola,M., Vogel,A., Reetz,M.T. and Bäckvall,J.-E. (2007) *Chem. Commun. (Camb.)*, **19**, 1913–1915.

Cereghino,G.P., Cereghino,J.L., Ilgen,C. and Cregg,J.M. (2002) *Curr. Opin. Biotechnol.*, **13**, 329–332.

Clouthier,C.M., Kayser,M.M. and Reetz,M.T. (2006) *J. Org. Chem.*, **71**, 8431–8437.

Cregg,J.M., Barringer,K.J., Hessler,A.Y. and Madden,K.R. (1985) *Mol. Cell. Biol.*, **5**, 3376–3385.

Cregg,J.M., Cereghino,J.L., Shi,J. and Higgins,D.R. (2000) *Mol. Biotechnol.*, **16**, 23–52.
Daly,R. and Hearn,M.T. (2005) *J. Mol. Recognit.*, **18**, 119–138.
Dominguez de Maria,P., Carboni-Oerlemans,C., Tuin,B., Bargeman,G., van der Meer,A. and van Gemert,R. (2005) *J. Mol. Catal. B Enzym.*, **37**, 36–46.
Ericsson,D.J., Kasrayan,A., Johansson,P., Bergfors,T., Sandström,A.G., Bäckvall,J.-E. and Mowbray,S.L. (2008) *J. Mol. Biol.*, **376**, 109–119.
Faber,K. (2000) *Biotransformations in Organic Chemistry*, 4th edn. Springer-Verlag, Berlin.
Gedey,S., Liljeblad,A., Lázár,L., Fülöp,F. and Kanerva,L.T. (2001) *Tetrahedron: Asymmetry*, **12**, 105–110.
Henke,E., Pleiss,J. and Bornscheuer,U.T. (2002) *Angew. Chem., Int. Ed.*, **41**, 3211–3213.
Høegh,I., Patkar,S., Halkier,T. and Hansen,M.T. (1995) *Can. J. Bot.*, **73**, 869–875.
Hong,I.-P., Anderson,S. and Choi,S.-G. (2006) *J. Microbiol. Biotechnol.*, **16**, 1362–1368.
Kasrayan,A., Bocola,M., Sandström,A.G., Lavén,G. and Bäckvall,J.-E. (2007) *ChemBioChem*, **8**, 1409–1415.
Kirk,O. and Christensen,M.W. (2002) *Org. Process Res. Dev.*, **6**, 446–451.
Kirk,O., Borchert,T.V. and Fuglsang,C.C. (2002) *Curr. Opin. Biotechnol.*, **13**, 345–351.
Lee,C.C., Williams,T.G., Wong,D.W. and Robertson,G.H. (2005) *Plasmid*, **54**, 80–85.
Liljeblad,A. and Kanerva,L.T. (2006) *Tetrahedron*, **62**, 5831–5854.
Liu,Z., Pscheidt,B., Avi,M., Gaisberger,R., Hartner,F.S., Schuster,C., Skranc,W., Gruber,K. and Glieder,A. (2008) *ChemBioChem*, **9**, 58–61.
Martinelle,M., Holmquist,M. and Hult,K. (1995) *Biochim. Biophys. Acta*, **1258**, 272–276.
Otten,L.G. and Quax,L.G. (2005) *Biomol. Eng.*, **22**, 1–9.
Pfeffer,J., Richter,S., Nieveler,J., Hansen,C.E., Rhlid,R.B., Schmid,R.D. and Rusnak,M. (2006) *Appl. Microbiol. Biotechnol.*, **72**, 931–938.
Pfeffer,J., Rusnak,M., Hansen,C.-E., Rhlid,R.B., Schmid,R.D. and Maurer,S.C. (2007) *J. Mol. Catal. B Enzym.*, **45**, 62–67.
Reetz,M.T. and Carballeira,J.D. (2007) *Nat. Protoc.*, **2**, 891–903.
Reetz,M.T. and Sanchis,J. (2008) *ChemBioChem*, **9**, 2260–2267.
Reetz,M.T., Zonta,A., Schimossek,K., Jaeger,K.-E. and Liebeton,K. (1997) *Angew. Chem. Int. Ed.*, **36**, 2830–2832.
Reetz,M.T., Bocola,M., Carballeira,J.D., Zha,D. and Vogel,A. (2005) *Angew. Chem. Int. Ed.*, **44**, 4192–4196.
Reetz,M.T., Wang,L.W. and Bocola,M. (2006a) *Angew. Chem. Int. Ed.*, **45**, 2494.
Reetz,M.T., Carballeira,J.D., Peyralans,J., Höbenreich,H., Maichele,A. and Vogel,A. (2006b) *Chem. Eur. J.*, **12**, 6031–6038.
Reetz,M.T., Kahakeaw,D. and Lohmer,R. (2008) *ChemBioChem*, **9**, 1797–1804.
Schultz,N., Hobley,T.J. and Syldatk,C. (2007) *Biotechnol. Lett.*, **29**, 365–371.
Sih,C.J. and Wu,S.H. (1989) *Topics Stereochem.*, **19**, 63–125.
Waterham,H.R., Digan,M.E., Koutz,P.J., Lair,S.V. and Cregg,J.M. (1997) *Gene*, **186**, 37–44.
Zhao,H. and Arnold,F.H. (1997) *Curr. Opin. Struct. Biol.*, **7**, 480–485.

Received April 8, 2009; revised May 11, 2009; accepted May 11, 2009

Edited by Bauke Dijkstra