Supporting Information

Sandström et al. 10.1073/pnas.1111537108

SI Text

General Equipment and Chemicals. Unless otherwise mentioned, all chemicals were purchased from Sigma-Aldrich and all enzymes were purchased from Fermentas. Phusion DNA polymerase was obtained from Finnzymes. Plasmid preparation kits and Gel and Cycle-Pure kits were purchased from Promega. Yeast plasmid kit was obtained from Omega-BioTek. Amicon centrifugal filter devices were purchased from Millipore. BioRad Micropulser was used for electroporation. Zeocin was purchased from InvivoGen. A Synergy 2 from BioTek was used for spectrophotometric assays. GC analyses were performed on a Varian GC 3900 using an IVA-DEX-1 chiral column from IVA Analysentechnik.

Media Recipes. 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. Yeast extract/peptone/dextrose (YPD) medium (10 $g\,L^{-1}$ yeast extract, 20 $g\,L^{-1}$ peptone, 20 $g\,L^{-1}$ dextrose) was used for *Pichia pastoris* cultivation and expression.

Vector Construction. Construction of template vector pBGP1-*Candida antarctica* lipase A (CALA) has been described previously (1).

Primers. Table S1 lists the primers used for the mutagenesis (all are written 5′–3′).

Preparation of Combinatorial Library Gene Fragment. Production of overlapping fragments. In five tubes adjusted to 100 μL each, 1× GC buffer, 10 ng pBGP1-CALA, 50 pmol of each primer pair, 200 μM of each dNTP and 1 U Phusion DNA polymerase was mixed and cycled with the following program; 95 °C for 2 min, then 24 cycles of 98 °C for 10 s, 72 °C (−0.5°/cycle) for 30 s, and 72 °C for 45 s, followed by 10 cycles of 98 °C for 10 s, 58 °C for 30 s, 72 °C for 2 min, and a final elongation at 72 °C for 10 min. All fragments were purified by a Wizard SV PCR Clean-up kit and the concentration was determined via Abs_{260} quantification and agarose gel densitometry. The primer pairs used were as followed: Fragment 1, Amplification_f and Group1_r; Fragment 2, Group1_f and Group2_r; Fragment 3, Group2_f and Group3_r; Fragment 4, Group3_f and Group4_r; Fragment 5, Group4_f and Amplification_r.

Combinatorial Library Gene Fragment Assembly. Overlap extension-PCR. In a tube adjusted to 100 μL, 1× GC buffer, 30 fmol of every fragment, 200 μM of each dNTP, and 1 U Phusion Hot Start DNA polymerase was mixed and cycled with the following program; 98 °C for 30 s, then 16 cycles of 98 °C for 10 s, 68 °C (−0.5°/cycle) for 30 s, 72 °C for 1 min and then hold at 4 °C for the addition of amplification primers (50 pmol each), whereafter the program continues with 98 °C for 30 s, then 16 cycles of 98 °C for 30 s, 68 °C (−0.5°/cycle) for 30 s, 72 °C for 1 min, and then a final elongation of 72 °C for 10 min.

Preparation of Partial pBGP1-CALA Vector Fragment. In a tube adjusted to 100 μL, 1× GC buffer, 10 ng pBGP1-CALA, 50 pmol of each Group1_r and Group4_f primer, 200 μM of each dNTP and 1 U Phusion DNA polymerase was mixed and cycled with the following program; 95 °C for 2 min, then 30 cycles of 98 °C for 10 s, 63 °C for 30 s, 72 °C for 3 min followed by final elongation at 72 °C for 10 min. The fragment was purified by Promega Wizard SV PCR Clean-up kit and concentration was determined via Abs_{260} quantification and densitometry.

Transformation of *P. pastoris*. *P. 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. Combinatorial library gene fragment (200 ng) were mixed with the partial pBGP1-CALA vector fragment (200 ng) and dehydrated. Thawed X33 cells were mixed with the dried DNA, and electroporated (0.2 mm cuvettes, 1.5 kV, 1 pulse). Cells were incubated at 30 °C with YPDS (YPD with 1 M D-sorbitol, 1 mL) for 2 h followed by plating on YPD-agar plates containing zeocin (100 $\mu g\,mL^{-1}$) and carbenicillin (100 $\mu g\,mL^{-1}$). Colonies usually appeared after 3 d of incubation at 30 °C.

Library Expression in *P. pastoris*. Single colonies were picked and inoculated in conical deep 96-well plates. Yeast library was expressed in 800 μL YPD containing 100 $\mu g\,mL^{-1}$ carbenicillin, 100 $\mu g\,mL^{-1}$ zeocin, at 29 °C at 350 rpm in an Infors Unitron (2.5 cm shaking amplitude) equipped with an M-tray. The yeast was pelleted by centrifugation, after 5 d of growth, and the supernatant was harvested by aspiration. The supernatant was used directly for optical screening or stored at 4 °C. The protein concentration of the harvested supernatants was generally 200 $\mu g\,mL^{-1}$. No change in protein expression was observed upon mutation. Master plates containing the pelleted cells were stored at −80 °C until further analysis or cultivation. BioRad Protein Assay was used to quantify protein yield, using BSA as a standard. All measurements were made in doublets.

Optical Screening for Improved Activity. Screening buffer (165 μL, 100 mM potassium phosphate, 10% vol/vol acetone, 4% vol/vol Triton X-100, 0.2% wt/vol gum arabic, pH 8.0) (2) and *rac*-**1** (10 μL, 2 $mg\,mL^{-1}$ in acetonitrile) were premixed before dispension. Yeast library supernatant (25 μL) and the premixed buffer and substrate solution (175 μL) were dispensed into a microtiter plate, and absorbance was measured at 410 nm for 15 min. Wild-type CALA was used as reference. Variants that showed high activity were selected for a kinetic resolution experiment to determine the enantioselectivity (*E*). For the assessment of functional diversity *rac*-**3** and *rac*-**4** was used, under identical assay conditions.

Enzyme Purification Using Hydrophobic Interaction Chromatography (HIC). The X33 supernatant displayed unspecific background activity toward ester **1** (also observed as a strong blank in Fig. 4) thereby necessitating purification of CALA variants for proper analysis of specific activity and potential kinetic investigations. Ammonium sulfate was added to 0.2-μm-filtered supernatants from 50-mL CALA variant cultures to final concentration of 1 M before loading the sample onto a HiPrep Butyl FF 16/10 (GE Healthcare). The column was preequilibrated in a starting buffer [1 M $(NH_4)_2SO_4$, 50 mM potassium phosphate, pH 7.0]. Samples were loaded onto the column, and eluted by a gradient of 100% starting buffer to 100% elution buffer (ddH_2O). Fractions containing the CALA variant were concentrated using Amicon Ultra-15 10 kD columns and the buffer was changed (potassium phosphate 100 mM, pH 8.0). BioRad Protein Assay was used to quantify protein yield, using BSA as a standard.

Specific Activity. Identical assay conditions as for the optical screening were used, but with HIC-purified enzyme. Initial rates were correlated to a 4-nitrophenol standard curve. Standard error of the mean is reported. For comparison, a typical activity of the

wild type toward a preferred substrate, 4-nitrophenyl hexadecanoate, is around 140 μmol min^{-1} mg^{-1}.

Kinetic Resolution. Potassium phosphate buffer (5 mL, 100 mM, pH 8.0), *rac*-**1** (650 μL, 2 mg mL^{-1} in acetonitrile) and HIC-purified enzyme (20 μg, in potassium phosphate buffer) were mixed. The reaction was shaken at 30 °C, until the reaction had reached 10–30% conversion. The reaction mixture was acidified by addition of HCl (1 mL, 1 M), extracted twice with dichloromethane (DCM) and concentrated. Conversion was determined by ^{1}H NMR, the enantiomeric excess was determined by chiral GC. Optical rotation was measured and absolute stereochemistry was determined by comparison with literature values.

Sequencing. Pelleted cells from master plates were used to inoculate YPD containing zeocin (100 μg mL^{-1}) and carbenicillin (100 μg 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 Yeast plasmid kit, and subsequently transformed into *Escherichia coli* DH5α to obtain a higher plasmid yield. The bacterial cells were cultivated, and plasmids were extracted and sequenced using appropriate primers.

Preparation of 4-Nitrophenyl Esters. The corresponding acids of esters **1**–**4** were purchased from commercial sources. Acid (6.42 mmol), 4-dimethylaminopyridine (78 mg, 0.64 mmol), and Et_3N (0.94 mL, 6.74 mmol) were dissolved in dry DCM (10 mL) under argon and stirred at 0 °C for 15 min. A solution of 4-nitrophenyl chloroformate (1.28 g, 6.36 mmol) in DCM (2 mL) was added to the reaction mixture that was then stirred at 0 °C for 2 h. The reaction mixture was diluted with DCM and extracted with HCl (0.1 M), $NaHCO_3$ (1 M), and finally with brine. The organic phase was dried over $MgSO_4$ and concentrated. Purification by flash column chromatography (Pentane/EtOAc) gave esters **1**–**4**, which were characterized by ^{1}H NMR spectroscopy.

Substrate Walking Libraries. Several residues were targeted pairwise and subjected to saturation mutagenesis, using NDT degeneracy (coding for 12 amino acids). A total of six libraries were created. These libraries targeted three sites; Phe(or Tyr)149NDT/Ile(or Asn)150NDT, Thr221NDT/Leu225NDT, and Phe(or Gly)233NDT/Gly237NDT, using CALA-WT as well as CALA-YNG as template. No significantly improved variants toward hydrolysis of **1** were found.

Residue Substitution Suggestions. The following residue substitution suggestions might serve as an inspiration for a combinatorial reshaping of the substrate pocket using binary substitution sets. First, a model has to be generated with the desired substrate docked or covalently bound in the substrate pocket/active site. A comparison has to be made, comparing the desired substrate with the enzyme's preferred substrate(s). If the substrate or a moiety of that substrate is larger than the natural substrate, we recommend using the degeneracy found in the expanding set (Table S2). If the desired substrate is smaller than the enzyme's preferred substrate, decreasing set is recommended (Table S3). It is highly recommended to analyze the degree of conservation of the residue sites, as described in *Experimental Methods*.

Expanding the Substrate Pocket Cavity. Using PyMol, (3) select all atoms of the desired substrate that are not part of the reactive functional group/moiety. Expand the atom selection by 4 Å around. All amino acids with some part of the side chain (or C_α) within 4 Å that became selected, becomes part of the set.

Decreasing the Substrate Pocket Cavity. Using PyMol, (3) select the atoms of the enzyme's preferred substrate that does not overlap with the desired substrate. Expand the atom selection by 5 Å around. All amino acids with some part of the side chain (or C_α) within 5 Å that became selected, becomes part of the set. These sites are recommended to use the decreasing set (Table S3). If, for example not enough residues became selected (thereby contributing to an unsatisfactory small library), include more distant residues: Select the same atoms of the desired substrate, as previously mentioned. Expand the selection by 6 Å atoms around. All amino acids with some part of the side chain (or C_α) within 6 Å that became selected, are added to the set. More radical modifications can be used. We recommend using the highly decreasing set (Table S3) for these sites.

Of course, we recommend removing residues from the set that the experimentalist knows is crucial for catalysis, or residues that are highly important for cofactor binding and coordination. The size of the library is also highly dependent on the screening capabilities, and of course has to be determined from case to case.

1. Sandström AG, Engström K, Nyhlén J, Kasrayan A, Bäckvall J-E (2009) Directed evolution of *Candida antarctica* lipase A using an episomally replicating yeast plasmid. *Protein Eng Des Sel* 22:413–420.
2. Schultz N, Hobley TJ, Syldatk C (2007) Spectrophotometric assay for online measurement of the activity of lipase immobilized on micro-magnetic particles. *Biotechnol Lett* 29:365–371.
3. DeLano WL (2002) The PyMOL molecular graphics system. Available at http://www.pymol.org.

Table S1. Primers used for mutagenesis (all are written 5′–3′)

	Primers
Group1_f	GGCTTCAAAGCCGCCTWCAWCGCTGGCTACGAAG
Group2_f	GGCACGSCAGTGAGCGCCAAGGACASCTTTACATTCKTAAACGGCGG
Group3_f	GCCGGCKKCGSACTGGCGGSAGTTTCGGGTC
Group4_f	CAGAGTGCGKTCGGCAAGCCCTTTG
Group1_r	CTCTTCGTAGCCAGCGWTGWAGGCGGCTTTGAAG
Group2_r	CGTTTAMGAATGTAAAGSTGTCCTTGGCGCTCACTGSCGTGCCGCC
Group3_r	CGAAACTSCCGCCAGTSCGMMGCCGGCGAAG
Group4_r	GGGCTTGCCGAMCGCACTCTGCTTG
Amplification_f	GACGGCGGTGCTCGACACGCCC
Amplification_r	CTAAGGTGGTGTGATGGGGCCAAAGGGCTT

Degenerate nucleotide designation; K = T,G; M = A,C; S = G,C; W = A,T. The mutagenic codons are underlined. Primers were ordered from Eurofins MWG Operon.

Table S2. Suggested residues for expanding the substrate pocket

WT residue	Expanding set		Highly expanding set*	
	Alternative residue	Codon degeneracy (sense)	Alternative residue	Codon degeneracy (sense)
Ala	Gly	GSA	—	—
Arg	Lys	ARA	—	—
Asn	Ser	ARC	—	—
Asp	—	—	Ala	GMC
Cys	Ser	TSC	—	—
Gln	—	—	Leu	CWA
Glu	Asp	GAK	—	—
Gly	—	—	—	—
His	—	—	Pro	CMC
Ile	Val	RTA	—	—
Leu	Ile	MTA	Val	KTA
Lys	Thr	AMA	—	—
Met	Ile	ATK	Val	RTG
Phe	Leu	TTK	Val	KTC
Pro	Ala	SCA	—	—
Ser	Gly	RGC	—	—
Thr	Ser	ASC	—	—
Trp	Leu	TKG	Gly	KGG
Tyr	Asn	WAC	Ser	TMC

*It is suggested that the column titled "Highly expanding set" should (only) be used for residues that appear to clash significantly with the potential substrate. If the backbone at the region appears to be displaced, and radical modification has to be undertaken, these suggested residues could definitely be used. Also, the electrostatic properties deviates more from the WT residue in this set, compared to the column titled "Expanding set."

Table S3. Suggested residues for reducing the substrate pocket cavity

WT residue	Decreasing set		Highly decreasing set*	
	Alternative residue	Codon degeneracy (sense)	Alternative residue	Codon degeneracy (sense)
Ala	Val	GYA	—	—
Arg	—	—	Trp	WGG
Asn	—	—	Tyr	ASN
Asp	Glu	GAK	—	—
Cys	—	—	Tyr	TRC
Gln	—	—	Arg	CRA
Glu	—	—	Gln	SAA
Gly	—	—	—	—
His	Arg	CRC	—	—
Ile	Leu	MTA	Phe	WTC
Leu	Ile	MTA	Phe	TTK
Lys	Arg	ARA	—	—
Met	Leu	MTG	—	—
Phe	—	—	—	—
Pro	Leu	CYA	—	—
Ser	Thr	ASC	Tyr	TMC
Thr	—	—	Asn	AMC
Trp	—	—	—	—
Tyr	—	WAC	Ser	TMC

*It is suggested that the column titled "Highly decreasing set" should (only) be used for residues that are further (~6 Å) from the potential substrate, and should foremost be used to expand the diversity of the library, if too few residues are close-range interacting (<5 Å) with the desired modeled substrate. Also, the electrostatic properties deviates more from the WT residue in this set, compared to the column titled "Decreasing set."