Catalysis Science & Technology

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/catalysis

Catalysis Science & Technology

Manuscript ID CC-COM-11-2014-008906

PAPER

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

Substrate channel evolution of an esterase for the synthesis of Cilastatin

Zheng-Jiao Luan,[a] Fu-Long Li,[a] Shuai Dou,[a] Qi Chen,[a] Xu-Dong Kong,[a] Jiahai Zhou,[b] Hui-Lei Yu*[a] and Jian-He Xu*[a]

Received (in XXX, XXX) Xth XXXXXXXXX 20XX, Accepted Xth XXXXXXXXX 20XX

DOI: 10.1039/b000000x

The esterase *Rh*Est1 from *Rhodococcus* sp. ECU1013 has been reported for the enantioselective hydrolysis of ethyl (*S*)-(+)-2,2-dimethylcyclopropane carboxylate, producing the building block of Cilastatin. In this work, error-prone PCR and site-directed saturation mutagenesis were applied to *Rh*Est1 for activity improvement, with pH-indicator assay as high-throughput screening method. As a result, *Rh*Est1$_{A147I/V148F/G254A}$, with mutations surrounding the substrate access channel, showed a 5-fold increase in its specific activity compared with the native enzyme, as well as a 4-fold increase in protein solubility. Combined with the determination of protein structure and computational analysis, this work shows that the amino acids around the substrate channel play a more important role in the activity evolution of *Rh*Est1 than those in the active site.

1. Introduction

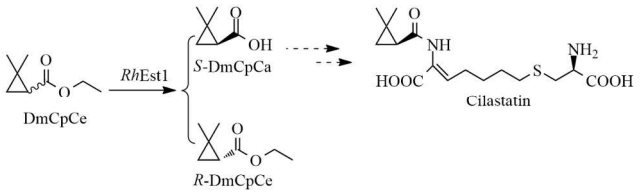
Cyclopropanes, with unique reactivity and unusual spectroscopic properties, can be found in a wide range of naturally occurring compounds and pharmaceutical agents.[1] Cilastatin is a small molecule containing a cyclopropyl moiety that can be used to prevent the breakdown of the carbapenem antibiotic Imipenem and increase its urinary recovery from approximately 35% to 60–75%. Cilastatin can also protect the kidneys against potential toxic effects from higher doses of Imipenem. During the past three decades, Imipenem and Cilastatin have been used in combination as a 1:1 mixture, known as Tienam, which showed a broad spectrum activity and fewer side effects for the treatment of severe infections.[2]

(*S*)-(+)-2,2-Dimethylcyclopropane carboxylic acid [(*S*)-DmCpCa] or (*S*)-(+)-2,2-dimethylcyclopropane carboxamide is a chiral building block for the synthesis of Cilastatin. During the past years, considerable chemical methods have been developed for the asymmetric synthesis of this important building block.[3] Recently, research efforts have been directed towards the development of biotechnological processes for the synthesis of advanced chiral intermediates because they require fewer unit operations and afford higher levels of optical purity than traditional chemical processes.[4] Wang *et al.*[4a] reported the enzymatic resolution of ethyl (±)-2,2-dimethylcyclopropane carboxylate [(±)-DmCpCe, 65 mM] using Novozyme 435 with a load of 16 g L^{-1}. In our previous work[5], a new esterase (*Rh*Est1) was discovered from *Rhodococcus* sp. ECU1013, which shows excellent enantioselectivity towards the hydrolysis of (*S*)-DmCpCe (Scheme 1). This work aims to improve the activity of *Rh*Est1 for the efficient and economic synthesis of (*S*)-DmCpCa, which is the key precursor for Cilastatin.

Directed evolution is a powerful tool in enzyme engineering and a series of techniques have been developed for the creation of diverse protein libraries.[6] Error-prone PCR offers entirely random mutagenesis for a wide range of engineering goals. Through sequential rounds of random mutation and efficient high-through screening method, error-prone PCR may uncover extraordinary mutations. In a seminal study published by Frances Arnold in 1993, a subtilisin E variant containing 10 amino acid substitutions with increased activity in DMF was generated by sequential rounds of random mutagenesis.[7a] Later the approach was developed with novel selection strategies and faster techniques for enzyme evolution.[7b,7c] As more protein structures are solved and various computational approaches are developed, structure-guided protein designs are becoming popular means to construct focused libraries.[8] The active sites of the enzyme are often taken into consideration for their significant influences on the specificity and performance of the enzyme. Yet, for enzymes whose active sites are buried in the protein core, the substrate access channel is especially more essential for enzyme evolution. Experimental examples have shown clearly that the modification of access pathways represents a successful strategy for proteins from various enzyme classes.[9] Jiri Damborsky group reported mutants of haloalkane dehalogenase as engineered by directed evolution around the access tunnels, with up to 32-fold higher activity than wild type towards 1,2,3-trichloropropane.[10]

In this study, we employed error-prone PCR and site-directed mutagenesis arround substrate channel in a combinatorial manner to improve the activity of *Rh*Est1 towards (±)-DmCpCe. A pH indicator (phenol red) was used in the current study for high-

throughput screening because the bioreaction produced acid, which could lead to a shift in the pH of the reaction medium.[11] The outcomes of this study showed that the strategy offered another patten for the activity evolution of enzymes.



Scheme 1 Biocatalytic resolution of (±)-DmCpCe for the synthesis of Cilastatin precursor mediated by a novel esterase *Rh*Est1.

2. Experimental

2.1. Reagents

Racemic ethyl (±)-2,2-dimethylcyclopropane carboxylate was kindly donated by Hisoar Pharmaceutical Co., Ltd (Zhejiang, China). Unless otherwise stated, all other chemicals were obtained commercially and were of analytical grade. Tryptone and yeast extract were obtained from Oxoid (Shanghai, China). r*Taq* polymerase, restriction endonucleases (*Dpn*I, *Eco*RI, *Hin*d III), T4 DNA ligase and PrimeSTAR^{TM} HS were all purchased from Takara Biotechnology Co., Ltd (Dalian, China) and stored at −20°C. Primers for site-directed saturation mutagenesis were synthesized by Generay Biotech Co., Ltd. (Shanghai, China) (Table S1). The pET28a(+) expression vector was purchased from Novagen (Shanghai, China). *E. coli* strain BL21 (DE3) from Tiangen (Shanghai, China) was used as the host strain for gene cloning and expression.[12]

2.2. Protein expression and purification

Cultivation of the recombinant *E. coli* cells expressing *Rh*Est1 was performed as described previously.[5b] After cultivation, the cells were harvested by centrifugation (6000×*g*) and the cell pellet was resuspended in lysis buffer (20 mM Tris–HCl, 500 mM NaCl, 5 mM 2-mercaptoethanol, 20 mM imidazole, pH 8.0), disrupted by sonification and the cell lysate was centrifuged (30000×*g*) at 4°C for 40 min. The resulting supernatant was loaded onto a Ni–NTA column (5 mL) equilibrated with lysis buffer, and the retained protein was eluted with an increasing gradient from 10 to 500 mM imidazole in elution buffer at a flow rate of 5 mL/min. Then gel filtration was performed by chromatography on a Superdex 75 column (GE Healthcare) and eluted with Tris–HCl (25 mM, pH 8.0) containing 150 mM NaCl. SDS–PAGE analysis of the eluted protein revealed over 95% purity of the target protein, which was concentrated to around 20 mg/mL, and aliquots were immediately frozen in liquid nitrogen and stored at −80°C.

2.3. Protein crystallization, data collection and structure determination

Crystals of *Rh*Est1 were obtained by the sitting-drop vapor diffusion method at 18°C by mixing the purified protein (14 mg/mL) with an equal volume of reservoir solution containing 3 M Na/K Phosphate, 0.1 M Tris pH 8.5. The selenium-substituted *Rh*Est1 was crystallized in the condition of 20% (w/v) PEG4000, 0.1 M citrate, pH 5.5, 10% (v/v) 2-propanol at 18°C. Prior to data collection, crystals were soaked in the cryoprotectant solution composed of their mother liquid and 10% (v/v) glycerol and then flash cooled in liquid nitrogen.

Diffraction data of native *Rh*Est1 was collected on a RaxisIV++ imaging plate detector (Rigaku, TX, USA) using an in-house Rigaku MicroMax-007 HF rotating-anode X-ray generator (1.5418 Å) operating at 40 kV and 30 mA. Diffraction data of *Se-Rh*Est1 was collected at the wavelength of 0.98 Å using an ADSC Quantum 315r detector at beamline BL17U of Shanghai Synchrotron Radiation Facility (SSRF). The intensity sets of the native-*Rh*Est1 crystals and *Se-Rh*Est1 crystals were indexed, integrated and scaled with the HKL2000 package.[13]

The single-wavelength anomalous diffraction (SAD) phases were calculated with Autosol.[14] A model of *Rh*Est1 was built automatically by ARP/wARP[15] and manually adjusted using COOT[16]. The 1.9 Å resolution structure of *Rh*Est1 was solved by molecular replacement method using Phaser[17] in the CCP4 crystallographic suite with this model as the template, rounds of automated refinement were performed with PHENIX.[14] Data collection and refinement statistics are summarized in Table S2. Figure 1 shows the overall structure of *Rh*Est1.

2.4. Error-prone PCR

Random mutagenesis was carried out by error-prone PCR. Plasmid pET28a(+) containing the wild-type *Rh*Est1 gene was used as the template for the first generation of random mutagenesis. Primers 5'-CCG**GAATTC**ATGTCTATTCGTGAAGCCGTC-3' and 5'-CTCGAGTGCGGCCGC**AAGCTT**TTAACCGAGGCTCGAGATGAA-3' were used as forward and reverse primers, respectively. For the 850 *bp* target gene, 100 μM Mn^{2+} was used to obtain the desired level of mutagenesis rate. PCR-mutated genes were digested with *Eco*RI and *Hin*dIII, and ligated into the plasmid pET28a(+) digested with the same two restriction enzymes. *E. coli* BL21 (DE3) cells were transformed with the resultant mutant plasmids and plated onto LB agar medium supplemented with 50 μg/mL of kanamycin.

2.5. Site-directed saturation mutagenesis

The site-directed mutagenesis was performed by using PrimeSTAR^{TM} HS. The reaction mixture contained 100 ng template DNA (plasmid containing the variant *Rh*Est1$_{A147V/G254A}$ gene), 1.25 U PrimeSTAR^{TM} HS, 0.2 μM of each primer, sterilized distilled water up to 50 μL. The extension reaction was initiated by preheating the reaction mixture to 98°C for 10 s, annealing at 55–65°C for 30 s according to the melting temperature of the primer pair, followed by elongation at 68°C for 6.5 min. The template DNA was digested with 10 U *Dpn*I for 1 h at 37°C. Five microliters of the plasmids containing the mutated gene was transformed into 100 μL competent cells of *E. coli* BL21 (DE3).

2.6. Library screening

Considering the production of acid, a colorimetric and pH-responsive method was used as the high-throughput screening method. Because of the optimum pH of *Rh*Est1 was 8.0, phenol red was chosen (pK_a = 8.00, pH ranges from 6.6 to 8.0, the color changes from yellow to red) as the pH indicator. The reaction mixture contained 100 μL enzyme solution and 10 mM (±)-

DmCpCe (10% v/v DMSO), and the plates were shaken for 30 min at 30°C. Then 100 μL indicator mixture was added into each well, which contained 20 μL 5 mg/L phenol red and 80 μL 100 mM $CaCl_2$. The mutants with higher hydrolase activity would show the bigger difference value at A_{430} and A_{560}.[11e]

2.7. Activity assay of the variants towards (±)-DmCpCe

The enzyme activity assay was performed in 500 μL reaction mixture, containing a certain purified enzyme and 10 mM (±)-DmCpCe (10% v/v DMSO), conducted at 30°C, 1000 rpm for a certain time. The product was acidified with 20% H_2SO_4, extracted with 500 μL ethyl acetate (1 mM dodecane as the internal standard) and detected by GC (SHIMADZU GC-2014). The detailed program for detection has been reported previously.[5a] One unit of enzyme activity was defined as the amount of enzyme releasing 1.0 μmol of (*S*)-DmCpCa per minute under the assay conditions.

2.8. Kinetic parameters determination

The kinetic parameters of the purified variants on the substrate (±)-DmCpCe were determined by measuring the activity under the varied substrate concentration (0.1–10 mM). The Michaelis-Menten (K_m) constant and the maximal reaction rate (V_{max}) of the enzyme were calculated from Lineweaver-Burk plots.

2.9. Enzymatic resolution of (±)-DmCpCe by *Rh*Est1, *Rh*Est1$_{A147V/G254A}$ and *Rh*Est1$_{A147I/V148F/G254A}$

The 10 mL reactions was performed with magnetic stirring at 25°C, containing 200 mM pH 8.0 Tris–HCl buffer, 5 g/L lyophilized cell-free extracts and 100 mM (±)-DmCpCe (0.5% w/v Tween-80). The specific activity of *Rh*Est1, the variants *Rh*Est1$_{A147V/G254A}$ and *Rh*Est1$_{A147I/V148F/G254A}$ towards *p*-nitrophenyl acetate was 5.06, 27.4, 58.3 U/mg $_{\text{lyophilized enzyme powder}}$, respectively. Samples were taken periodically to measure the conversions and ee_p.

For large scale preparations, (±)-DmCpCe (50 mmol or 100 mmol) was dissolved in 100 mL Tris–HCl buffer (100 mM, pH 8.0). Lyophilized cell-free extracts (0.5 g or 2.5 g) were added to initiate the reaction. The pH was automatically maintained at 8.0 by titrating 1 M NaOH. After the reaction was terminated, the reaction mixture was extracted with dichloromethane in alkaline conditions to remove the substrate, then the reaction mixture was acidified to pH < 2.0. The mixture was extracted with 100 mL dichloromethane for three times. The mixed organic layer was dried over anhydrous sodium sulfate and then evaporated in vacuum. The isolated product was validated based on ^{1}H NMR spectra and the data was also given in Fig. S1.

3. Results and discussion

3.1. Structure of *Rh*Est1

The crystal structure of *Rh*Est1 was resolved at 1.9 Å resolution (PDB ID: 4RNC). Consideration of the *Rh*Est1 structure revealed that it was consisted of a core domain with the α/β hydrolase fold and a V-shaped cap domain (Fig. 1A). Through sequential and structural alignment with the reported esterase structures, the catalytic triad of *Rh*Est1 was identified as Ser101, Asp225 and His253. Ser101 was located in a canonical nucleophile elbow in the conserved Gly-X_1-Ser-X_2-Gly sequence of α/β hydrolase. Asp225 and His253 are in positions consistent with a hydrolytic mechanism. An oxyanion hole was also identified, which was composed of the main chain nitrogen atoms of Trp33 and Tyr102.[18]

The structure of *Rh*Est1 was found to be very similar (with an r.m.s.d. of 1.249 Å) to that of the esterase PFE from *Pseudomonas fluorescens* (PDB ID: 1VA4), which favours esters with short chain acyl groups.[19] Though the sequence alignment of the two enzymes was only 27%, the overlap of *Rh*Est1 and PFE showed that the catalytic triad and oxyanion hole were almost in the same orientation. The largest deviation of the two enzymes is the shape of the cap domain, which might greatly influence the substrate channel of the enzyme (Fig. 1B).

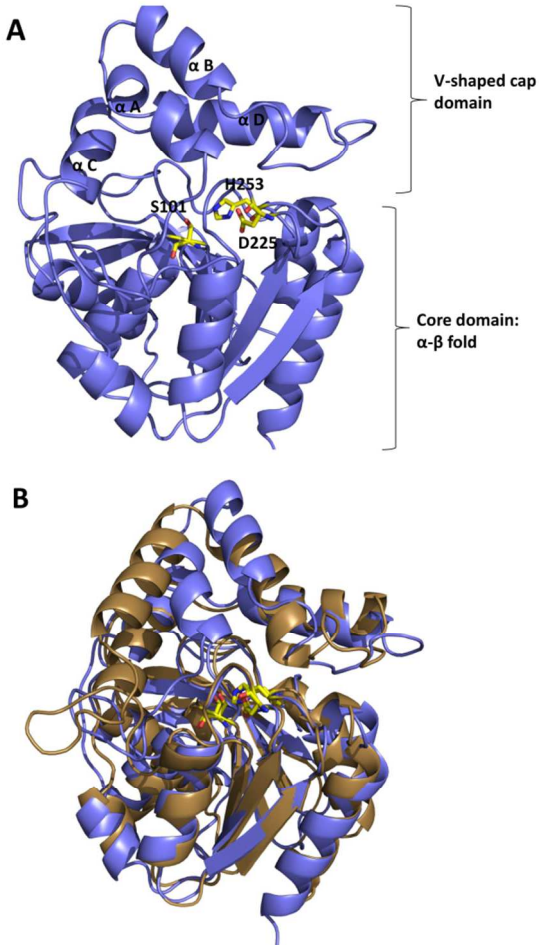


Fig. 1 A. Overall structure of *Rh*Est1, the catalytic triad are shown in yellow sticks. The cap domain is consisted of four α helices, α A and α B helixes are almost parallel, while α C and α D form a V-shaped opening with great angle. B. Superposition of *Rh*Est1 (slate) and PFE from *Pseudomonas fluorescens* (sand) in a stereo view. The cap domain is most obvious difference between *Rh*Est1 and PFE.

3.2. Error-prone PCR mutagenesis for activity improvement

Based on the heterologous expression system *E. coli*-pET28a(+)-*Rh*Est1, error-prone PCR was applied firstly to generate a random library for activity improvement of *Rh*Est1 towards (±)-DmCpCe. After screening over 10,000 variants, *Rh*Est1$_{A147V/G254A}$ showed a 2.5-fold higher activity than the wild-type enzyme, as well as a 4-fold increase in its solubility. To investigate the effect of individual amino acid replacement on *Rh*Est1 activity, single

mutations of A147V and G254A were constructed. It showed that A147V mutation played a key role in enhancing the activity of the enzyme and G254A mutation, which was positioned next to the His253 residue of the catalytic triad, mainly contributed to the observed increase in the solubility (Table 1 and Fig. 2).

Table 1 Expression and specific activity of *Rh*Est1 wild-type and variants obtained by directed evolution.

Enzyme	Total activity (U)[a]	Specific activity (U/mg protein)[b]	Total enzyme (mg)[c]
*Rh*Est1	0.45	0.17 ± 0.01	2.6
*Rh*Est1$_{A147V}$	1.5	0.48 ± 0.02	3.1
*Rh*Est1$_{G254A}$	2.3	0.13 ± 0.02	18
*Rh*Est1$_{A147V/G254A}$	5.3	0.42 ± 0.02	13

[a] Total activity of the recombinant *E. coli* cell towards (±)-DmCpCe cultivated in the 100 mL flask. The recombinant *E. coli* cell was induced with 0.1 mM IPTG at 25°C for 12 h.

[b] Specific activity of the purified enzyme towards (±)-DmCpCe.

[c] Total enzyme (mg) = Total activity / specific activity.

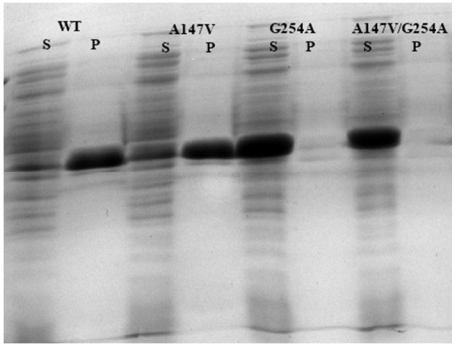


Fig. 2 SDS-PAGE for the protein expression of the wild-type *Rh*Est1 and the variants. S: soluble fraction; P: insoluble fraction.

3.3. Site-directed mutagenesis around the substrate channel

Structure-guided designs are intended to be more functionally applied for biocatalyst mutations, because more information is presented and the total numbers of variants that must be created and screened are greatly reduced. This approach is increasingly relying on computational methods, which helps to identify potential sites for mutagenesis and estimate mutant properties in silico. CAVER 3.0 was used to identify the tunnels existed in *Rh*Est1 (http://www.caver.cz) (Fig. 3A).[9d] For further understanding the relationships between substrate and enzyme, (*S*)-DmCpCe was docked into the binding pocket of *Rh*Est1 using the Autodock Vina software (http://vina.scripps.edu/)[20], and the most reasonable substrate-binding position matching the catalytic mechanism was selected. This docking analysis revealed that the substrate was lying in the middle of a long and narrow substrate channel, buried between the cap domain and the core domain.

It was reported that the size, shape, physico-chemical properties and dynamics of the protein channels (otherwise known as the *Keyhole* in the *Keyhole-Lock-Key* model of enzyme catalysis) can determine the kinetics and equilibria of substrate entry and/or product exit.[9a] In this work, amino acids surrounding the substrate channel were chosen for site-directed mutagenesis experiments (Fig. 3B). Two regions in particular were targeted, including (1) the amino acid residues surrounding the acyl-binding pocket (**Library A**); and (2) the amino acid residues surrounding the alcohol-binding site (**Library B**). Mutant A147V/G254A was used as the parent for the second round of site-directed mutation experiments.

Table 2 lists the libraries constructed from the site-directed saturation mutagenesis experiments as well as the screening results of the mutagenesis. Subsequent screening of the two libraries revealed that the activity of *Rh*Est1$_{A147I/V148F/G254A}$ and *Rh*Est1$_{V140L/A147V/G254A}$ from **Library A** was nearly 2-fold, 1.5-fold greater than that of *Rh*Est1$_{A147V/G254A}$, respectively, meanwhile the expression levels of the two variants remained the same as *Rh*Est1$_{A147V/G254A}$. It was observed that the positive mutations in **Library A** were all located around the substrate entrance channel of the structure. This site might be the gate of the channel, which controlled the mobility of substrate from the media surrounding the protein into the active site. For **Library B**, there was no positive mutation obtained with higher activity. This phenomenon might be explained that the amino acids in **Library B**, which consist of the alcohol-binding site, were not the limited factor for activity improvement.

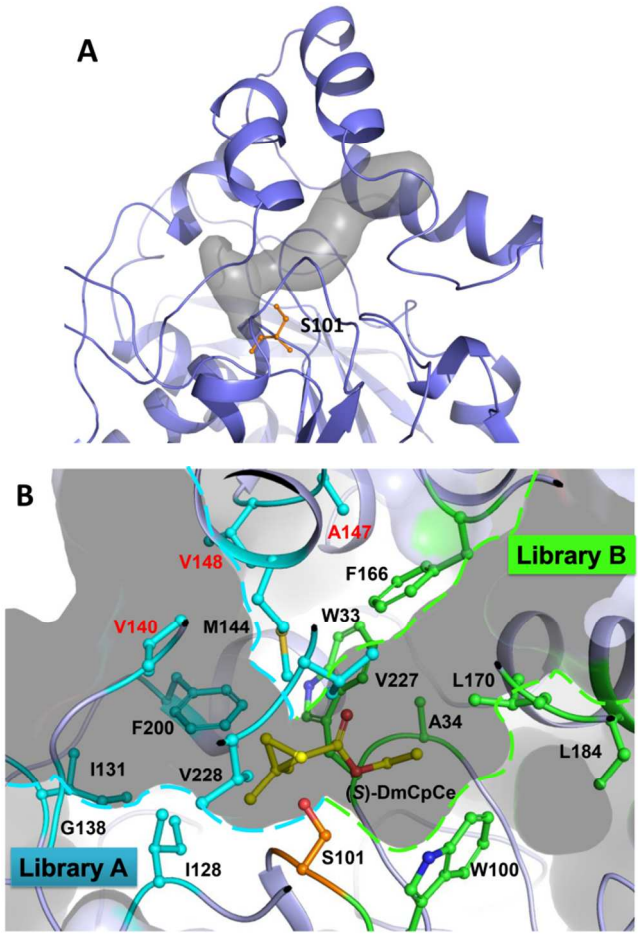


Fig. 3 A. *Rh*Est1 channels calculated by CAVER 3.0 are shown in grey surface. B. Molecular docking of (*S*)-DmCpCe (yellow sticks) into *Rh*Est1. The nucleophile **Ser101** is shown as an orange stick model. The amino acids of **Libraries A** and **B** are shown as cyan and green stick models, respectively. The two parts of the substrate channel surrounded by **Libraries A** and **B** were labelled with cyan and green dashed lines, respectively.

Table 2 Results of site-directed saturation mutagenesis experiments.[a]

Library	Mutation site	Codon[b]	Enzyme	Specific activity (U/mg protein)
			*Rh*Est1	0.17
			*Rh*Est1$_{A147V/G254A}$	0.42
Library A				
	Ile128	NNK (94)	n.d.[c]	
	Ile131	NNK (94)	n.d.	
	Gly138	NNK (94)	n.d.	
	Val140	NNK (94)	*Rh*Est1$_{V140L/A147V/G254A}$	0.62
	Met144	NNK (94)	n.d.	
	Ala147/Val148	NDT (460)	*Rh*Est1$_{A147I/V148F/G254A}$	0.78
	Phe200	NNK (94)	n.d.	
	Val227/Val228	NDT (460)	n.d.	
Library B				
	Trp33/Ala34	NDT (460)	n.d.	
	Trp100	NNK (94)	n.d.	
	Phe166	NNK (94)	n.d.	
	Leu170	NNK (94)	n.d.	
	Leu184	NNK (94)	n.d.	

[a] *Rh*Est1$_{A147V/G254A}$ was used as the parent for site-directed mutation.

[b] The amino acids adjacent to each other were simultaneously randomized by the degenerated codon NDT, or otherwise randomized by NNK (N = A, G, C, T; K=G, T; D=A, G, T. NDT involves 12 codons for 12 amino acids and NNK involves 32 codons for 20 amino acids). The number of transformants screened is shown in parentheses.

[c] n.d.: variants with improved activity were not detected in these sites mutation.

3.4. Kinetic parameters for the variants

The steady-state kinetic parameters for (±)-DmCpCe indicated that the catalytic efficiencies of all the constructed mutants were significantly higher than that of the wild-type *Rh*Est1 (Table 3). For the variant *Rh*Est1$_{A147V/G254A}$ or *Rh*Est1$_{V140L/A147V/G254A}$, the kinetic efficiency k_{cat}/K_m was about 3-fold greater than that of the wild-type enzyme. For the variant *Rh*Est1$_{A147I/V148F/G254A}$, a 5.6-fold decrease in K_m combined with a 2.4-fold increase in k_{cat} contributed to a total 13-fold increase in k_{cat}/K_m.

Table 3 Kinetic parameters for wild-type and the variants.[a]

Enzyme	K_m (mM)	k_{cat} (min^{-1})	k_{cat}/K_m ($min^{-1}mM^{-1}$)	ee_p (%)
*Rh*Est1[5b]	0.43	8.0	19	97
*Rh*Est1$_{A147V/G254A}$	0.15 ± 0.04	10 ± 0.6	64	97
*Rh*Est1$_{V140L/A147V/G254A}$	0.7 ± 0.1	42.5 ± 2.2	64	64
*Rh*Est1$_{A147I/V148F/G254A}$	0.077 ± 0.020	19 ± 0.6	240	92

[a] Kinetic parameters of the purified wild-type *Rh*Est1 enzyme and the variants towards (±)-DmCpCe were determined by measuring their activity under varied substrate concentrations. Reaction conditions: Tris-HCl buffer, 100 mM, pH 8.0 at 30°C, 1000 rpm.

3.5. Catalytic performance of the variants

The enzymatic resolution of (±)-DmCpCe (100 mM) was conducted comparatively using the wild-type and two variants *Rh*Est1$_{A147V/G254A}$ or *Rh*Est1$_{A147I/V148F/G254A}$, with 5 g L^{-1} of lyophilized cell-free extract. As shown in Fig. 4, the reactions catalyzed by the *Rh*Est1$_{A147V/G254A}$ and *Rh*Est1$_{A147I/V148F/G254A}$ variants achieved 45% conversion after 3.0 and 1.25 h, respectively. In contrast, the wild-type enzyme required 9.0 h to reach a similar level of conversion. The catalytic activities of the two variants were therefore much higher than that of *Rh*Est1, and the enantioselectivity of the *Rh*Est1$_{A147V/G254A}$ variant was similar to that of the parent (97% ee_p) through the entire reaction process. It is noteworthy that the ee_p value of the *Rh*Est1$_{A147I/V148F/G254A}$ variant has dropped. However, considering the whole process involved in the production of Cilastatin, the resulting (*S*)-(+)-2,2-dimethylcyclopropane carboxamide precursor could be readily recrystallized to make up for this drop in the ee_p.[21]

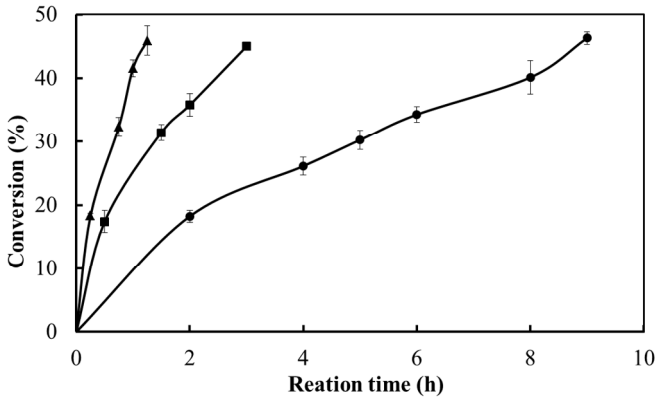


Fig. 4 Enzymatic resolution of (±)-DmCpCe at 100 mM by *Rh*Est1 (circle), *Rh*Est1$_{A147V/G254A}$ (square) or *Rh*Est1$_{A147I/V148F/G254A}$ (triangle) at a catalyst loading of 5 g L^{-1} of lyophilized cell-free extract. The reaction was conducted in Tris-HCl buffer (pH 8.0, 200 mM) at 25°C.

To confirm the feasibility of the biotransformation process on a large scale, (*S*)-DmCpCa was prepared at the 100 mL scale by the wild-type *Rh*Est1 and the variant *Rh*Est1$_{A147I/V148F/G254A}$. As shown in Table 4, for the biocatalytic resolution of (±)-DmCpCe (0.5 M), the catalyst loading of the variant *Rh*Est1$_{A147I/V148F/G254A}$ was 5-fold decreased than the wild-type *Rh*Est1.

Table 4 Preparation of (*S*)-DmCpCa by *Rh*Est1 and *Rh*Est1$_{A147I/V148F/G254A}$.[a]

Enzyme	[S] (mM)	Catalyst loadng (g L^{-1})	Time (h)	Conv. (Yield) (%)	ee_p (%)	E
*Rh*Est1	0.5	25	13	36.4 (30.7)	97	114
*Rh*Est1$_{A147I/V148F/G254A}$	0.5	5	12	41.5 (33.8)	90	36
*Rh*Est1$_{A147I/V148F/G254A}$	1.0	5	24	31.7 (30.1)	92	36

[a] The enzymatic resolutions were conducted at a substrate concentration of 0.5–1.0 M in the 200 mM Tris-HCl buffer at 100 mL scale. The pH was automatically maintained at 8.0 by titrating with 1 M NaOH. The conversion and ee_p value were analyzed by GC as described in the Experimental Section.

3.6. Structural analysis of the evolution

The results of this study revealed that mutations surrounding the substrate channel were much more amenable to protein engineering rather than those in the active site. The channel played an important role in substrate entry and/or product release, as well as affecting the activity, specificity and enantioselectivity of the enzyme. In this study, positive mutations were mainly located on the surface of the *Rh*Est1 structure, with the variants in **Library A** containing mutations at the entrance to the whole substrate channel. However, mutations situated directly at the entrance to the channel would have a significant impact on the enantioselectivity of the enzyme. For example, the *Rh*Est1$_{V140L/A147V/G254A}$ variant in **Library A** showed an improvement in its activity compared with *Rh*Est1$_{A147V/G254A}$,

however, the ee_p value for the bioresolution dropped to 64% when the conversion reached 43% under the same reaction conditions. As shown in Fig. 5, steric hindrance resulting from the Val140 residue at the narrowest section of the entrance would determine the direction in which the substrate entered the channel and affect the enantioselectivity of the enzyme. The Ala147 and Val148 residues surrounding the entrance to the substrate channel are relatively far from the active site. Mutations at these residues would allow for an improvement in the activity that would be compatible with the enantioselectivity of the enzyme. The activity of the *Rh*Est1$_{A147I/V148F/G254A}$ variant was 5-fold greater than that of the wild-type and did not suffered a significant loss in its enantioselectivity towards the hydrolysis of (*S*)-DmCpCe.

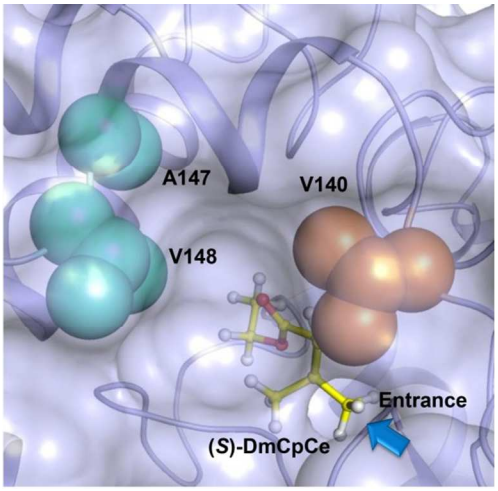


Fig. 5 Close up view of hot spots selected from **Library A**. (*S*)-DmCpCe is shown as a yellow stick model; **Ala147/Val148** and **Val140** are shown as cyan and orange spheres, respectively. The entrance to the substrate channel has been indicated with a blue arrow.

4. Conclusion

In summary, *Rh*Est1$_{A147I/V148F/G254A}$ has been identified as an efficient mutant for the enantioselective hydrolysis of (*S*)-DmCpCe following two rounds of evolution involving a combination of error-prone PCR and site-directed saturation mutagenesis. To the best of our knowledge, this study represents the first reported example of an improvement in the activity of an esterase towards a substrate containing a cyclopropyl group. In this study, mutations surrounding the entrance to the substrate channel allowed for an effective balance between improvements in the activity and a high level of enantioselectivity, which is a big challenge for improving the activity of *Rh*Est1 towards ethyl 2,2-dimethylcyclopropane carboxylate.

Acknowledgements

We are grateful for access to beamline BL17U1 at Shanghai Synchrotron Radiation Facility and thank the beamline staff for technical support. This work was financially supported by NSFC (No. 21276082), MOST (Nos. 2011CB710800 and 2011AA02A210), and SKLBE (No. 2060204).

References

[a] *State Key Laboratory of Bioreactor Engineering and Shanghai Collaborative Innovation Centre for Biomanufacturing, East China University of Science and Technology, Shanghai 200237, China.*
E-mail: huileiyu@ecust.edu.cn or jianhexu@ecust.edu.cn;
Fax: +86-21-6425-0840; Tel: +86-21-6425-2498.
[b] *State Key Laboratory of Bio-organic and Natural Products Chemistry, Shanghai Institute of Organic Chemistry, Chinese Academy of Sciences, Shanghai 200032, China.*

‡ Electronic Supplementary Information (ESI) available: Table containing sequences of primers for PCR-based site-directed mutagenesis; Table of crystal structure data collection and refinement for *Rh*Est1; the ^{1}H NMR spectrum data of (*S*)-DmCpCa.

1 a) A. de Meijere, *Angew. Chem. Int. Ed.*, 1979, **18**, 809; b) J. Salaün, *Chem. Rev.*, 1989, **89**, 1247; c) X. J. Wang, S. Takami, M. Kubo and A. Miyamoto, *Chem. Phys.*, 2002, **279**, 7; d) Y. Katsuda, *Pestic. Sci.*, 1999, **55**, 775; e) K. H. Park and M. J. Kurth, *Tetrahedron*, 2002, **58**, 8629; f) A. Gagnon, M. Duplessis and L. Fader, *Org. Prep. Proced. Int.*, 2010, **42**, 1; g) M. Baranac-Stojanović and M. Stojanović, *J. Org. Chem.*, 2013, **78**, 1504; h) T. F. Schneider, J. Kaschel and D. B. Werz, *Angew. Chem. Int. Ed.*, 2014, **53**, 5504.

2 a) D. W. Graham, W. T. Ashton, L. Barash, J. E. Brown, R. D. Brown, L. F. Canning, A. Chen, J. P. Springer and E. F. Rogers, *J. Med. Chem.*, 1987, **30**, 1074; b) A. C. Rodloff, E. J. C. Goldstein and A. Torres, *J. Antimicrob. Chemoth.*, 2006, **58**, 916; c) J. Hoffman, J. Trimble and G. M. Brophy, *Neurocrit. Care*, 2009, **10**, 403; d) K. M. Papp-Wallace, A. Endimiani, M. A. Taracila and R. A. Bonomo, *Antimicrob. Agents Chemother.*, 2011, **55**, 4943; e) Y. X. Wu, K. Chen and Z. H. Shi, *Curr. Pharm. Biotechnol.*, 2014, **15**, 685.

3 a) W. A. Donaldson, *Tetrahedron*, 2001, **57**, 8589; b) D. Y. K. Chen, R. H. Pouwer and J. A. Richard, *Chem. Soc. Rev.*, 2012, **41**, 4631.

4 a) F. Ren, P Wang, J. Huang and J. Y. He, *Biotechnol. Biotec. Eq.*, 2012, **26**, 3412; b) F. Y. Liang, J. Huang, J. Y. He and P. Wang, *Biotechnol. Bioproc. Eng.*, 2012, **17**, 952; c) N. M. Shaw, K. T. Robins and A. Kiener, *Adv. Synth. Catal.*, 2003, **345**, 425; d) R. C. Zheng, Z. Y. Yang, C. C. Li, Y. G. Zheng and Y. C. Shen, *J. Mol. Catal. B-Enzym.,* 2014, **102**, 161; e) S. J. Jin, R. C. Zheng, Y. G. Zheng and Y. C. Shen, *J. Appl. Microbiol.*, 2008, **105**, 1150; f) M. X. Wang and G. Q. Feng, *J. Mol. Catal. B-Enzym.,* 2002, **18**, 267; g) S. J. Yeom, H. J. Kim, D. K. Oh, *Enzym. Microb. Technol.*, 2007, **41**, 842.

5 a) C. H. Liu, J. Pan, Q. Ye and J. H. Xu, *Appl. Microbiol. Biotechnol.*, 2013, **97**, 7659; b) Y. Zhang, J. Pan, Z. J. Luan, G. C. Xu, S. H. Park and J. H. Xu, *Appl. Environ. Microbiol.,* 2014, **80**, 7348.

6 a) U. T. Bornscheuer and M. Pohl, *Curr. Opin. Chem. Biol.*, 2001, **5**, 137; b) J. D. Bloom, M. M. Meyer, P. Meinhold, C. R. Otey, D. MacMillan and F. H. Arnold, *Curr. Opin. Struc. Biol.*, 2005, **15**, 447; c) N. J. Turner, *Nat. Chem. Biol.*, 2009, **5**, 567; d) M. B. Quin and C. Schmidt-Dannert, *ACS Catal.*, 2011, **1**, 1017; e) H. Jochens, M. Hesseler, K. Stiba, S. K. Padhi, R. J. Kazlauskas and U. T. Bornscheuer, *ChemBioChem*, 2011, **12**, 1508; f) M. Wang, T. Si and H. M. Zhao, *Bioresour. Technol.*, 2012, **115**, 117; g) M. T. Reetz, *J. Am. Chem. Soc.*, 2013, **135**, 12480; h) M. D. Lane and B. Seelig, *Curr. Opin. Chem. Biol.*, 2014, **22**, 129.

7 a) K. Chen and F. H. Arnold, *Proc. Natl. Acad. Sci. U.S.A.*, 1993, **90**, 5618; b) T. H. Yoo, M. Pogson, B. L. Iverso and G. Georgiou, *ChemBioChem*, 2012, **13**, 649; c) S. Fujii, T. Matsuura, T. Sunami, Y. Kazuta and T. Yomo, *Proc. Natl. Acad. Sci. U.S.A.*, 2013, **110**, 16796.

8 a) J. R. Cherry, M. H. Lamsa, P. Schneider, J. Vind, A. Svendsen, A. Jones and A. H. Pedersen, *Nat. Biotechnol.*, 1999, **17**, 379; b) M. T. Reetz, L. W. Wang and M. Bocola, *Angew. Chem. Int. Ed.,* 2006, **118**, 1258; c) S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz and J. P. Acevedo, *ACS Synth. Biol.*, 2013, **2**, 83; d) T. Davids, M. Schmidt, D. Böttcher and U. T. Bornscheuer, *Curr. Opin. Chem. Biol.*, 2013, **17**, 215; e) Y. Li, P. C. Cirino, *Biotechnol. Bioeng.*, 2014, **111**, 1273; f) J. Damborsky and J. Brezovsky, *Curr. Opin. Chem. Biol.*, 2014, **19**, 8.

9 a) Z. Prokop, A. Gora, J. Brezovsky, R. Chaloupkova, V. Stepankova and J. Damborsky, in *Protein Engineering Handbook*, ed. S. Lutz and U. T. Bornscheuer, Wiley-VCH Verlag GmbH & Co. KGaA, 1st edn., 2012, vol. 3, pp. 421–463; b) A. Gora, J. Brezovsky and J.

Damborsky, *Chem. Rev.*, 2013, **113**, 5871; c) H. Morita, K. Wanibuchi, H. Nii, R. Kato, S. Sugio and I. Abe, *Proc. Natl. Acad. Sci. U.S.A.*, 2010, **107**, 19778; d) E. Chovancova, A. Pavelka, P. Benes, O. Strnad, J. Brezovsky, B. Kozlikova, A. Gora, V. Sustr, M. Klvana, P. Medek, L. Biedermannova, J. Sochor and J. Damborsky, *PLOS Comput. Biol.*, 2012, **8**, 1; e) T. Koudelakova, R. Chaloupkova, J. Brezovsky, Z. Prokop, E. Sebestova, M. Hesseler, M. Khabiri, M. Plevaka, D. Kulik, I. K. Smatanova, P. Rezacova, R. Ettrich, U. T. Bornscheuer and J. Damborsky, *Angew. Chem. Int. Ed.*, 2013, **52**, 1959; f) X. D. Kong, S. G. Yuan, L. Li, S. Chen, J. H. Xu and J. Zhou, *Proc. Natl. Acad. Sci. U.S.A.*, 2014, **111**, 15717; (g) X. D. Kong, Q. Ma, J. Zhou, B. B. Zeng and J. H. Xu, *Angew. Chem. Int. Ed.*, 2014, **53**, 6641.

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

11 a) G. T. John and E. Heinzle, *Biotechnol. Bioeng.*, 2001, **72**, 620; b) J. P. Goddard and J. L. Reymond, *Curr. Opin. Biotechnol.*, 2004, **15**, 314; c) U. T. Bornscheuer, *Eng. Life Sci.*, 2004, **4**, 539; d) M. Schmidt and U. T. Bornscheuer, *Biomol. Eng.*, 2005, **22**, 51; e) Y. Gong, G. C. Xu, G. W. Zheng, C. X. Li and J. H. Xu, *J. Mol. Catal. B-Enzym.*, 2014, **109**, 1.

12 a) F. L. Du, H. L. Yu, J. H. Xu and C. X. Li, *Bioresour. Bioproc.* 2014, **1**, 10; b) W. X. Zhang, G. C. Xu, L. Huang, J. Pan, H. L. Yu and J. H. Xu, *Org. Lett.*, 2013, **15**, 4917.

13 Z. Otwinowski and W. Minor, in *Method Enzymol*, ed. C. W. Carter, Jr. Academic Press, 1997, pp. 307.

14 P. D. Adams, R. W. Grosse-Kunstleve, L. W. Hung, T. R. Ioerger, A. J. Mccoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter and T. C. Terwilliger, *Acta Crystallogr. D,* 2002, **58**, 1948.

15 W. T. M. Mooij, S. X. Cohen, K. Joosten, G. N. Murshudov and A. Perrakis, *Structure*, 2009, **17**, 183.

16 P. Emsley and K. Cowtan, *Acta Crystallogr. D*, 2004, **60**, 2126.

17 A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.,* 2007, **40**, 658.

18 a) M. Nardini and B. W. Dijkstra, *Curr. Opin. Struc. Biol.*, 1999, **9**, 732; b) C. C. Akoh, G. C. Lee, Y. C. Liaw, T. H. Huang and J. F. Shaw, *Prog. Lipid Res.*, 2004, **43**, 534; c) P. Liu, Y. F. Wang, H. E. Ewis, A. T. Abdelal, C. D. Lu, R. W. Harrison and I. T. Weber, *J. Mol. Biol.*, 2004, **342**, 551.

19 a) K. K. Kim, H. K. Song, D. H. Shin, K. Y. Hwang, S. Choe, O. J. Yoo and S. W. Suh, *Structure*, 1997, **5**, 1571; b) J. D. Cheeseman, A. Tocilj, S. Park, J. D. Schrag and R. J. Kazlauskas, *Acta Crystallgr. D*, 2004, **60**, 1237; c) A. M. F. Liu, N. A. Somers, R. J. Kazlauskas, T. S. Brush, F. Zocher, M. M. Enzelberger, U. T. Bornscheuer, G. P. Horsman, A. Mezzetti, C. Schimidt-Dannert and R. D. Schmid, *Tetrahedron Asymmetr.*, 2001, **12**, 545.

20 a) O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455−461; b) D. Seeliger and B. L. de Groot, *J. Comput. Aid. Mol. Des.*, 2010, **24**, 417.

21 M. Minai, T. Katsura and Y. Ueda, Eur. Pat., 155779B1, 1985.

Graphical Abstract

Error-prone PCR and site-directed mutagenesis around substrate channel were combinatorially employed for improving an esterase (*Rh*Est1) activity towards Cilastatin building block. *Rh*Est1$_{A147I/V148F/G254A}$, with mutations surrounding the entrance to substrate channel, showed 5 times higher activity than the native enzyme and additional 4-fold increase in protein solubility.