Advanced Synthesis & Catalysis

Accepted Article

Title: Highly Focused Library-Based Engineering of Candida antarctica Lipase B with (S)-Selectivity Towards sec-Alcohols

Authors: Yixin Cen, Danyang Li, Jian Xu, Qiong-Si Wu, Qi Wu, and Xian-Fu Lin

To be cited as: *Adv. Synth. Catal.* 10.1002/adsc.201800711

Link to VoR: http://dx.doi.org/10.1002/adsc.201800711

10.1002/adsc.201800711

FULL PAPER

DOI: 10.1002/adsc.201((will be filled in by the editorial staff))

Highly Focused Library-Based Engineering of *Candida antarctica* Lipase B with (*S*)-Selectivity Towards *sec*-Alcohols

Yixin Cen,[+a] Danyang Li, [+a] Jian Xu,[a] Qiongsi Wu,[a] Qi Wu[a,*] and Xianfu Lin[a,*]

[a] Department of Chemistry, Zhejiang University, Hangzhou 310027, People's Republic of China
Fax: (+86)-571-8795-2618; e-mail: llc123@zju.edu.cn or xflin@zju.edu.cn
[+] These authors contributed equally to this work.

Received: ((will be filled in by the editorial staff))

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/adsc.201######.((Please delete if not appropriate))

Abstract. *Candida antarctica* lipase B (CALB) is one of the most extensively used biocatalysts in both academia and industry and exhibits remarkable (*R*)-enantioselectivity for various chiral *sec*-alcohols. Considering the significance of tailor-made stereoselectivity in organic synthesis, a discovery of enantiocomplementary lipase mutants with high (*R*)- and (*S*)-selectivity is valuable and highly desired. Herein, we report a highly efficient directed evolution strategy, using only 4 representative amino acids, namely, alanine (A), leucine (L), lysine (K), tryptophan (W) at each mutated site to create an extremely small library of CALB variants requiring notably less screening. The obtained best mutant with three mutations W104V/A281L/A282K displayed highly reversed (S)-selectivity towards a series of *sec*-alcohol with *E* values up to 115 (conv. 50%, ee 94%). Compared with the previously reported (*S*)-selective CALB variant, W104A, a single mutation provided less selectivity, while the synergistic effects of three mutations in the best variant endow better (*S*)-selectivity and a broader substrate scope than the W104A variant. Structural analysis and molecular dynamics simulation unveiled the source of reversed enantioselectivity.

Keywords: *Candida antarctica* lipase B; Directed evolution; Focused library; *sec*-Alcohols; Hydrolysis

Introduction

Chiral *sec*-alcohols belong to a type of important intermediates in the pharmaceutical industry for the production of active pharmaceutical ingredients to treat health conditions such as cardiovascular diseases and hypertension.[1-3] For example, (*R*)- and (*S*)-enantiomers of 1-phenylethanol are significant optically active substances.[4,5] (*S*)-1-phenylethanol can be used to synthesize sertraline to treat depression and can also be a synthetic precursor to treat asthma and enhance the immune system.[6] On the other hand, (*R*)-1-phenylethanol can be used to synthesize drugs that inhibit cholesterol absorption.[7,8] Hence, the tailor-made preparation of optically pure chiral *sec*-alcohols are particularly important.

Kinetic resolution (KR) is a key tool in the synthesis of optically active *sec*-alcohols. The common hydrolases show preferential (*R*)-stereoselectivity for chiral *sec*-alcohol in the hydrolysis or acylation reactions according to the "Kazlauskas rule", [9] since large substituents bind to the large hydrophobic pocket and small ones to the medium pocket. So far only few hydrolases showed (*S*)-stereoselectivity towards *sec*-alcohols, as in the case of subtilisin, which can selectively acylate (*S*)-configurational *sec*-alcohols or hydrolyze their esters.[10,11]

To improve and alter the catalytic properties of enzymes, protein engineering has emerged as a powerful means during the past two decades.[12] Directed evolution and rational design are two main strategies for protein engineering.[13] ISM (iterative saturation mutagenesis) based on CAST (combinatorial active site saturation test) proposed by Reetz and coworkers have been successfully applied to increase thermostability, activity and selectivity of various enzymes.[14]

Candida antarctica lipase B (CALB) is one of the most widely used lipases and has a broad range of applications in kinetic resolution of racemic alcohols and amines in both academic and industrial laboratories on account of its extraordinary catalytic performance.[15,16] Structurally, CALB has an active site composed of the catalytic triad Ser105/ His224/ Asp187, an oxyanion hole, and two binding pockets for the acyl-moiety of the ester and the alcohol parts.[17-19] Consequently, protein engineering aiming to improve or reverse the stereoselectivity of CALB for racemic acids or *sec*-alcohol substrates mainly focused on the corresponding acyl or alcohol-binding pockets, respectively.[20,21] Considering the generally equivalent significance of two enantiomers of resolved substrates, discovering

10.1002/adsc.201800711

enantiocomplementary mutants with high (*R*)- and (*S*)-selectivities is valuable and highly desired. For example, we have previously demonstrated that CALB acid-binding pocket can be modified through directed evolution to create enantiocomplementary mutants for the kinetic resolution of racemic acids.[20b] Concerning the inversion of the stereoselectivity of CALB towards certain *sec*-alcohols, Hult and coworkers reported an important CALB mutation, W104A, which was obtained by a rational design and being situated in the alcohol-binding pocket, was able to accommodate much larger groups than the wild-type (WT) CALB, and this change transformed the highly (*R*)-selective WT into an (*S*)-selective mutant.[22] However, the enantioselectivity of W104A was unsatisfactory, the *E* values were low ($2<E<12$) for some typical *sec*-alcohols, such as phenylethanol, phenylpropanol and 2-hexanol, in their acylation reactions, although these *E* values could be increased by an appropriate choice of solvents in combination with a higher temperature and longer alkyl chains of *sec*-alcohols.[21-23] Thus, it is highly desirable to explore other mutations at the W104 hot spot or supplementary mutations at other important sites surrounding the alcohol-binding pocket to obtain CALB mutants with higher (*S*)-selectivity than W104A. On the other hand, all previous studies concerning the (*S*)-selective mutant W104A were performed for the acylation reactions in organic solvents, which leaves uncertainty about the selectivity of this (*S*)-selective mutant for the hydrolytic kinetic resolution of *sec*-alcohols.

Considering that large gene libraries created by random mutagenesis often show low hit ratios, herein we constructed a highly focused library containing less than 20 variants using a proposed new strategy. This strategy was related to iterative saturation mutagenesis (ISM) at CAST sites proposed by Reetz and coworkers, [14] while using only 4 representative amino acids, namely, alanine (A), leucine (L), lysine (K) and tryptophan (W) at each site to create an extremely small library requiring notably less screening. The best mutant obtained in this study contained a triple W104V/A281L/A282K mutation and displayed better (*S*)-selectivity for the hydrolytic kinetic resolution of a series of *sec*-alcohol esters than the previously reported W104A variant.

Results and Discussion

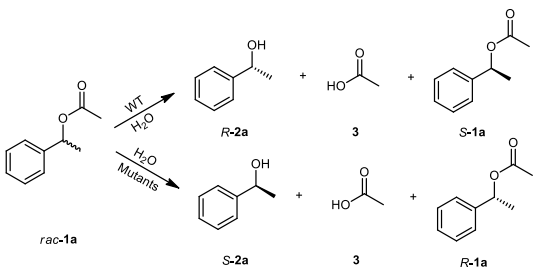


Figure 1. Stereoselective hydrolysis of 1-phenylethyl acetate catalyzed by wide type (WT) CALB and mutants.

The hydrolytic kinetic resolution of racemic 1-phenylethyl acetate (*rac*-**1a**) was chosen as the model reaction (Figure 1). As mentioned above, WT CALB had a preference to (*R*)-*sec*-alcohol, and the resolution result of *rac*-**1a** was the 99% ee value with 50% yield.

To obtain (*S*)-selective mutants for *rac*-**1a**, which meant inverting the Kazlauskas rule, the stereoselective pocket of alcohols should be changed to accommodate the phenylalkanol substituents, which are larger than the ethyl group. An early study [21-23] has identified W104 as the most important hot spot in the alcohol-binding pocket of CALB for the reversed (*S*)-selectivity of *sec*-alcohols, and W104 should be replaced by smaller amino acids to enlarge the volume of the stereoselectivity pocket. According to the X-ray structure of CALB (PDB code 1TCA,[17] Figure 2A), we selected four important hydrophobic residues W104, L278, A281 and A282 (sites A, C, and B as shown in Figure 2B) surrounding the alcohol-binding pocket as mutagenesis targets. It is well-known that highly efficient screening is one of the biggest bottlenecks in directed evolution of enantioselective enzymes, especially for multiple-site saturation mutagenesis. Thus, we proposed to construct one highly focused library composed of only a dozen variants. Considering the importance of the space of the alcohol stereoselective pocket, the principle of amino acid exchange is based on the different side chain volumes of the targeting amino acids.

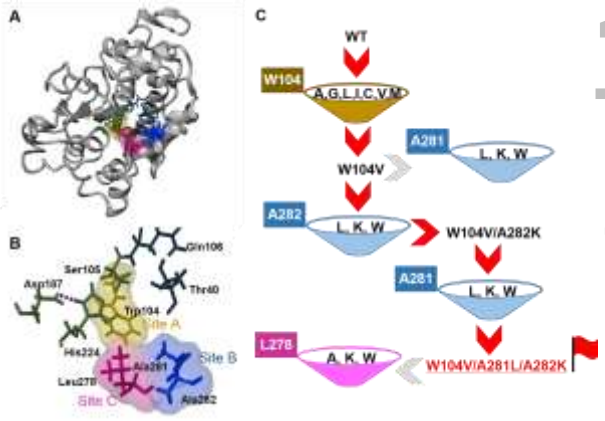


Figure 2. (**A**) X-ray structure of WT CALB (PDB: 1TCA)[17] used as a guide in directed evolution. (**B**) Mutagenesis sites A (Trp104), B (Ala282/ Ala282), C (Leu278). (The catalytic triad S105-H224-D187 and the oxyanion hole T40-Q106 are shown with black sticks, and the mutant site A is shown with yellow sticks, site B is blue, site C is pink.) (**C**) Strategy for the construction of highly focused CALB library.

First, we chose seven residues (alanine (A), glycine (G), leucine (L), isoleucine (I), cysteine (C), valine (V) and methionine (M)) for an amino acid exchange of W104, because of their smaller volume than that of tryptophan (W), which made up library A. Based on

the best hit obtained from the "subsaturation" mutagenesis at site A (W104), three other amino acid residues at sites B and C were varied iteratively to A, L, K and W to create a CALB library containing less than 20 variants (Figure 2C, Table 1). Alanine was the smallest amino acid except glycine, and glycine exchange usually impacted the secondary structure of lipases leading the loss of activity. Leucine and tryptophan were chosen as amino acids with medium-to-large side chains, and lysine was a functional residue with a long carbon chain. This strategy for selecting amino acids can remarkably reduce the size of libraries, only require dozens of screening, and have a high percentage of positive hits. All the screening results are shown in Table 1.

In library A, most of the W104 variants containing small side chains showed reversed (*S*)-selectivity with moderate E-values, as expected. It is noteworthy that the replacement with smaller amino acids did not give rise to better (*S*)-selectivity of 1-phenylethyl acetate. For example, the W104G mutant (WB3) was expectedly inactive, probably due to the disrupted catalytic structure of the active site of this mutant or too large of a space in the active site to efficiently bind the substrate. The W104A mutant (WB2) only showed moderate (*S*)-selectivity for *rac*-**1a** with E = 18 (ee_p 78%, conv. 49%). Interestingly, the best mutant in library A was W104V showing an acceptable E value (E = 21) with the reversed (*S*)-selectivity (WB7, entry 7 in Table 1). The valine side chain was larger than in alanine and glycine, but smaller than in other large amino acids, probably having the most suitable space to accommodate the phenyl group of *rac*-**1a**. The best mutant WB7 was then used to perform A/L/K/W mutagenesis at site B (A281 and A282). This library provided an improved variant W104V/A282K (E = 29) (entry 10, Table 1). Further mutagenesis at A281 of this variant resulted in the best variant WB15 with W104V/A281L/A282K mutation, providing reversed 94% ee upon 46% conversion (E(*S*) = 80, entry 15, Table 1). An attempt to continue to the third-round evolution at site C (L278) failed to induce further improvement. The total screening effort involved the evaluation of only 19 variants.

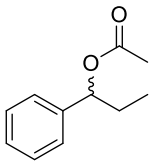
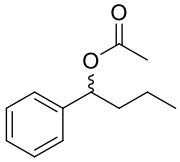
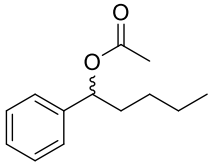
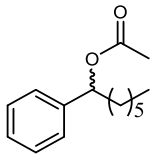
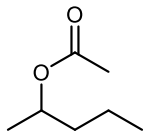
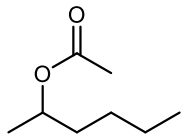
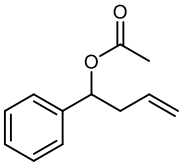
Table 1. CALB mutants catalyzed kinetic hydrolysis of *rac*-**1a**.[a]

Entry	Library	Mutant	Sequence	Conversion[b] (%)	ee_p[b] (%)	E[c] (*R*)	E[c] (*S*)
1	-	WT	-	50	99 (*R*)	>200	
2	A	WB2	W104A	49	78 (*S*)		18
3		WB3	W104G	1	0		
4		WB4	W104L	9	63 (*S*)		5
5		WB5	W104I	44	65 (*S*)		8
6		WB6	W104C	37	35 (*S*)		3
7		WB7	W104V	38	85 (*S*)		21
8		WB8	W104M	48	44 (*S*)		4
9	B	WB9	W104V+A282L	40	70 (*S*)		9
10		WB10	W104V+A282K	48	85 (*S*)		29
11		WB11	W104V+A282W	14	85 (*R*)	14	
12		WB12	W104V+A281L	20	50 (*S*)		3
13		WB13	W104V+A281K	14	35 (*S*)		2
14		WB14	W104V+A281W	3	65 (*S*)		5
15		WB15	W104V+A281L+A282K	46	94 (*S*)		80
16		WB16	W104V+A281K+A282K	25	30 (*S*)		2
17		WB17	W104V+A281W+A282K	37	60 (*S*)		6
18	C	WB18	W104V+L278A+A281L+A282K	29	87 (*S*)		20
19		WB19	W104V+L278K+A281L+A282K	9	93 (*R*)	30	
20		WB20	W104V+L278W+A281L+A282K	43	54 (*R*)	5	
21	-	WB21	A281L+A282K	49	99 (*R*)	>200	
22	-	WB22	A282K	49	99 (*R*)	>200	
23	-	WB23	A281L	49	99 (*R*)	>200	

[a] Reactions were performed with *rac*-**1a** (0.0125 mmol), acetonitrile (50 μL) and enzymes in PBS (950 μL, 50 mM, pH = 7.5) at 37 °C for several hours. [b] Determined by chiral GC analysis using Agilent CP-Chirasil-Dex CB column and dodecane as an internal standard. [c] Enantiomeric ratio calculated according to the equation, $E = \ln[1-C^*(1+ee_p)] / \ln[1-C^*(1-ee_p)]$, the estimated error of the measurement is less than 3%.

10.1002/adsc.201800711

Table 2. Hydrolytic kinetic resolution of different substrates using CALB mutants originally evolved for *rac*-**1**.[a]

Entry	Substrate	Enzymes	Time (h)	Conversion[b] (%)	ee_p.[c] (%)	*E* (*R*)	*E* (*S*)
1	 *rac*-**1b**	WT	8	46(42)	>99	>200	
2		WB2	12	49(39)	92.3 ± 0.1		73.1 ± 1.2
3		WB15	24	50(41)	94.0 ± 0.2		115.1 ± 5.1
4	 *rac*-**1c**	WT	48	7(-)	>99	>200	
5		WB2	15	48(39)	73.6 ± 1.5		13.3 ± 1.1
6		WB15	24	23(19)	96.0 ± 0.8		65.1 ± 12.5
7	 *rac*-**1d**	WT	48	0(-)	0	-	-
8		WB2	30	45(41)	54.3 ± 5.2		5.3 ± 1.0
9		WB15	36	21(15)	77.3 ± 2.3		9.2 ± 1.7
10	 *rac*-**1e**	WT	48	0(-)	0	-	-
11		WB2	30	32(25)	81.0 ± 3.1		14.2 ± 2.8
12		WB15	36	27(20)	80.1 ± 2.2		12.2 ± 1.6
13	 *rac*-**1f**	WT	5	49(40)	>99	>200	
14		WB2	8	43(35)	39 ± 3		<4
15		WB15	14	30(20)	95 ± 2		69.8 ± 30.1
16	 *rac*-**1g**	WT	5.5	49(37)	>99	>200	
17		WB2	10	30(18)	35.2 ± 0.4		<3
18		WB15	14	24(18)	91 ± 3		31.5 ± 11.5
19	 *rac*-**1h**	WT	8	49(41)	93.0 ± 0.2	83.5 ± 2.9	
20		WB2	13	46(37)	84.5 ± 3.2		27.0 ± 7.1
21		WB15	30	38(26)	96.6 ± 0.6		109.2 ± 20.1

[a] Reactions were performed with different substrates (0.0125 mmol), acetonitrile (50 μL) and enzymes in PBS (950 μL, 50 mM, pH = 7.5) at 37 °C for several hours. [b] Reaction yields were determined by chiral GC analysis using an Agilent CP-Chirasil-Dex CB column or chiral HPLC analysis using a ChiralPak OJ-H column. The isolation yields (the actual yields) are given in parentheses (For convenience of isolation, the reaction scale was increased 40 times, thus, the used substrate were 0.5 mmol). [c] The mean value of 2-3 measurements.

10.1002/adsc.201800711

Deconvolution was further performed to explore the effect of each mutation in the WB15 variant on the reversed (*S*)-selectivity of *rac*-**1a**. Three single-point mutants (W104V, A281L, A282K) and three double-point mutants (W104V/A281L, W104V/A282K, and A281L/A282K) were compared (entries 7, 23, 22, 12, 10 and 21, Table 1). A281L, A282K and A281L/A282K mutants showed similar (*R*)-selectivity to the WT implying the crucial role played by W104V mutation for reversing stereoselectivity. W104V (*E*(*S*) = 21), W104V/A281L (*E*(*S*) = 3) and W104V/A282K (*E*(*S*) = 29) exert synergistic effects in the best variant of WB15 (*E*(*S*) = 80).

Next, the substrate range of the best (*S*)-selective mutant WB15 evolved for the model substrate *rac*-**1a** was determined. Moreover, considering that all previous studies on (*S*)-selectivity of W104A (WB2) were carried out for the acylation reaction of racemic phenylalkanols in nonaqueous media,[21-23] it is valuable to determine the stereoselectivity of W104A in the hydrolytic kinetic resolution of phenylalkanol esters. The comparison results of WT, WB2 and the best WB15 variants as the catalysts for hydrolytic kinetic resolution of a series of phenylalkanol esters are listed in Table 2. As pointed by Hult and coworkers,[21-23] the largest alkyl group of *sec*-alcohols that can smoothly fit into the alcohol-binding pocket of WT CALB is ethyl. Due to this constraint, the substrates **1c**–**1e** are not acceptable for WT CALB, and a drastic drop in both activity and selectivity is observed (entries 4, 7, 10, Table 2). W104A (WB2) variant displayed reversed (*S*)-selectivity for almost all the tested substrates with moderate-to-good *E* values (3-73). Upon employing the best mutant WB15, remarkably enhanced selectivity was observed in many cases when compared with WB2. For example, higher *E* values for (*S*)-selectivity of WB15 were obtained in the KR of *rac*-**1b**-**1d** and *rac*-**1f**-**1h** than WB2 (entries 1-9, 13-21, Table 2). A CALB mutant with a much higher reversed (*S*)-selectivity than the well-known W104A mutant was successfully obtained.

The length of the substrate alkyl substituent showed an interesting influence on the enantioselectivity of the WB2 and WB15. When the phenylalkanol alkyl chain changed from methyl to ethyl, the (*S*)-selectivity of WB2 and WB15 were improved remarkably, and *E* values increased from 18 to 73 (WB2) and from 80 to 115 (WB15), respectively. However, a further increase in the alkyl chain length of phenylalkanol from ethyl to butyl resulted in a drastic drop in the selectivity of WB2 and WB15. Although the enlarged stereoselectivity pockets of the (*S*)-selective variants WB2 and WB15 were mainly occupied by the phenyl substituents of these substrates, the increased alkyl chain length of phenylalkanols probably increased the hydrophobicity of alkyl substituents, thus, easily leading the competitive orientation of alkyl substituents down into the stereoselectivity pocket, and further resulting in the decreasing selectivity. When the alkyl chain length was hexyl, the (*S*)-selectivities of both WB2 and WB15 were slightly improved, which was consistent with the report by Hult on the influence of the substrate alkyl substituent length on the enantioselectivity of W104A-catalyzed transesterifications in cyclohexane.[23] Interestingly, when phenyl was substituted by propyl and butyl, namely, straight chain alcohols, WB15 showed much higher (*S*)-selectivity than WB2 mutant (entries 14-15, 17-18, Table 2).

As shown in Table 2, the reaction time of different substrates with WB15 was much longer than WT, this finding was in full accord with the result of the kinetic study (Table S1 in the Supporting Information). To evaluate the thermostability of WB15, the T_{50}^{30} values were measured by the 50% residual activity after a heat treatment for 30 min. The results showed a slightly decreased thermostability of WB15 (47 °C) when compared with WT CALB (55 °C). We hypothesized that the decrease in thermostability of WB15 may be caused by the introduction of point mutations, which affected the structural stability of the enzymes.

In addition, the isolation yields were lower than the GC or HPLC yields in Table 2, the loss of products may be partly caused by the absorption of samples on silica gel and the loss of impure products during column chromatography purification. This situation may become less pronounced in a scaled-up process. Therefore, we next performed the scaled-up kinetic resolution of the substrate *rac*-**1b** using crude enzyme WB15 for the preparation of (*S*)-**2b**. The reaction mixture containing 0.5 g substrate was shaken at 37 °C until the reaction had reached 40-50% conversion. Then the crude products were extracted with MTBE. By using flash column chromatography, 174 mg of (*S*)-**2b** (94% ee) and 220 mg of (*R*)-**1b** (90% ee) were obtained. The isolated yields reached up to 45% in this scaled-up process, which was similar to the yield in the screening reaction determined by GC (47% conversion). This result documented the synthetic utility of the reversed stereoselective kinetic resolution catalyzed by the mutant WB15. To expand the application range, the acyl-transfer reactions were also tested for catalysis with WT and the best mutant WB15 in organic phase. *Rac*-**2b** (2.0 mg) and vinyl acetate (3.0 eq) were dissolved in 1.0 mL naphthane with 25 mg immobilized CALB. After shaking at 200 rpm and 50 °C for 12 hours, the reactions were analyzed by GC. WT CALB provided a good result with 45% conversion and 99% ee, as expected. To our delight, the WB15 also exhibited good enantioselectivity of 99% ee with 10% conversion. The low conversion might be improved by optimizing the immobilization methods and reaction solvents in further studies.

To obtain some insight into the source of complementary stereoselectivity from WT and WB15 CALB, molecular dynamics (MD) computations were further carried out using the modeling suite YASARA.[24-28] All WB15 mutations were introduced

by swapping the specific amino acid residues of WT CALB, followed by energy minimizations.

The tetrahedral intermediates (TI) of (*R*)- and (*S*)-**1a** were built on the side-chain O atom of the catalytic Ser105 in the WT CALB and WB15 variants, respectively. In MD simulations of (*R*)- and (*S*)-TIs, 100 simulation frames (100 ps to 10 ns) were evaluated and additionally minimized to derive statistical averages and properties of the corresponding local minima. Finally, the distances and bond angles for some important H-bonds were also compared (Table 3). The energy-minimized structures of (*R*)- and (*S*)-TIs of WT CALB and mutant WB15 are shown in Figure 3.

For the WT CALB, the space in the alcohol-binding pocket is very small because tryptophan occupies position 104, and the largest alkyl group of *sec*-alcohols that can be accommodated is ethyl.[21-23] Thus, as shown in Figure 3A and 3B, both (*R*)-TI-CALB complex and (*S*)-TI-CALB complex have similar orientations of methyl group located inside the alcohol-binding pocket. However, they have completely different H-bonds formed and interactions between the substrate and active sites. All necessary H-bonds for catalysis were found in the optimized pose of the fast-reacting (*R*)-TI-CALB complex (Figure 3A), while the most important H-bond between H224 (HE2 atom) and the bound (*S*)-TI oxygen (O1 atom) was lost in the nonproductive (*S*)-TI-CALB complex (Figure 3B, black line), leading to no nucleophilic attack towards the ester bond of **1a**. Additionally, (*R*)-TI-CALB complex had a shorter distance between HE2 and O1 atoms (1.812 Å vs. 2.341Å) and larger bond-angles of NE2--HE2--O1 (150.3° vs. 134.4°) (Table 3) than that of nonproductive (*S*)-TI-CALB complex. These results confirmed the high (*R*)-selectivity of WT CALB and were also consistent with the previous calculation studies by Hult.[22]

For the (*S*)-selective mutant WB15, the W104V mutation greatly increased the space in the alcohol-binding pocket, thus, both methyl and phenyl can be accommodated efficiently, and two binding modes for the orientations of the alcohol moiety in the active site of WB15 are possible for (*R*)-**1a** or (*S*)-**1a**, respectively.[29] Using (*R*)-**1a** as an example, which is shown in Figure S2 (Supporting Information), Mode I placed the large substituent (phenyl group) into the alcohol-binding pocket and the medium substituent (methyl group) towards the active site entrance (Figure S2A and S2C in the Supporting Information). Exchanging the positions of large and medium substituents represented a starting conformation of Mode II (Figure S2B and S2D in the Supporting Information). After MD simulations of these two modes, we found that the configuration constructed by Mode I was more dominant for the complex of (*R*)-**1a** with WB15 mutant. The formation energy of the energy-minimized structure of (*R*)-TIs in Mode I was much lower than that in Mode II (-640.6 KJ/mol *vs.* -605.7 KJ/mol), probably due to the stronger hydrophobic interaction of phenyl than of methyl in the large alcohol-binding pocket surrounded by hydrophobic residues V104, L278, L281, and I285. Thus, we chose the Mode I as the binding model for nonproductive (*R*)-TI-WB15 complex (Figure 3C). The binding mode for the productive (*S*)-TI-WB15 complex is similar as that of (*R*)-TI-WB15 complex (Figure 3D). However, their H-bond networks between the substrate and active sites are different. The productive (*S*)-TI-WB15 complex has all necessary H-bonds for catalysis (Figure 3D), while the most important H-bond between H224 (HE2 atom) and the bound (*R*)-TI oxygen (O1 atom) was lost in the nonproductive (*R*)-TI-WB15 complex (Figure 3C, black line). Moreover, the bond distances (*d*) between HE2 and O1 atoms and the corresponding bond angles (α) of NE2--HE2--O1 in (*S*)-TI-WB15 complex are also more favorable for the nucleophilic attack towards the ester bond of **1a** than (*R*)-TI-WB15 complex (Table 3). These results clearly implied the reversed (*S*)-selectivity of WB15 mutant towards *sec*-alcohols.

Table 3. The hydrogen bond distances (*d*) between His 224 (HE2 atom) and the bound (*R*)- or (*S*)-tetrahedral intermediate (TI) oxygen (O1 atom), the corresponding bond angles (α) of NE2--HE2--O1, and the possibility of hydrogen bond formation (P_1 and P_2) between TI (O atom of carbonyl) and the oxyanion hole residues (Gln106 and Thr40).

TI-Enzymes complex	*d* (Å) [a]	α (°) [a]	P_1 [b]	P_2 [b]
(*R*-**1a**)-TI-WT	1.812	150.3	0.90	1.00
(*S*-**1a**)-TI-WT	2.341	134.4	- [c]	- [c]
(*R*-**1a**)-TI-WB15	3.745	142.0	- [c]	- [c]
(*S*-**1a**)-TI-WB15	1.798	174.1	0.91	1.00

[a] Data were obtained from the energy-minimized structure of the (*R/S*-**1a**)-TI-enzyme complex of WT CALB or variant WB15. [b] Data were the average values obtained from 100 snapshots of the TI-enzyme complex of WT CALB or variant WB15 during 10 ns MD. [c] Not detected.

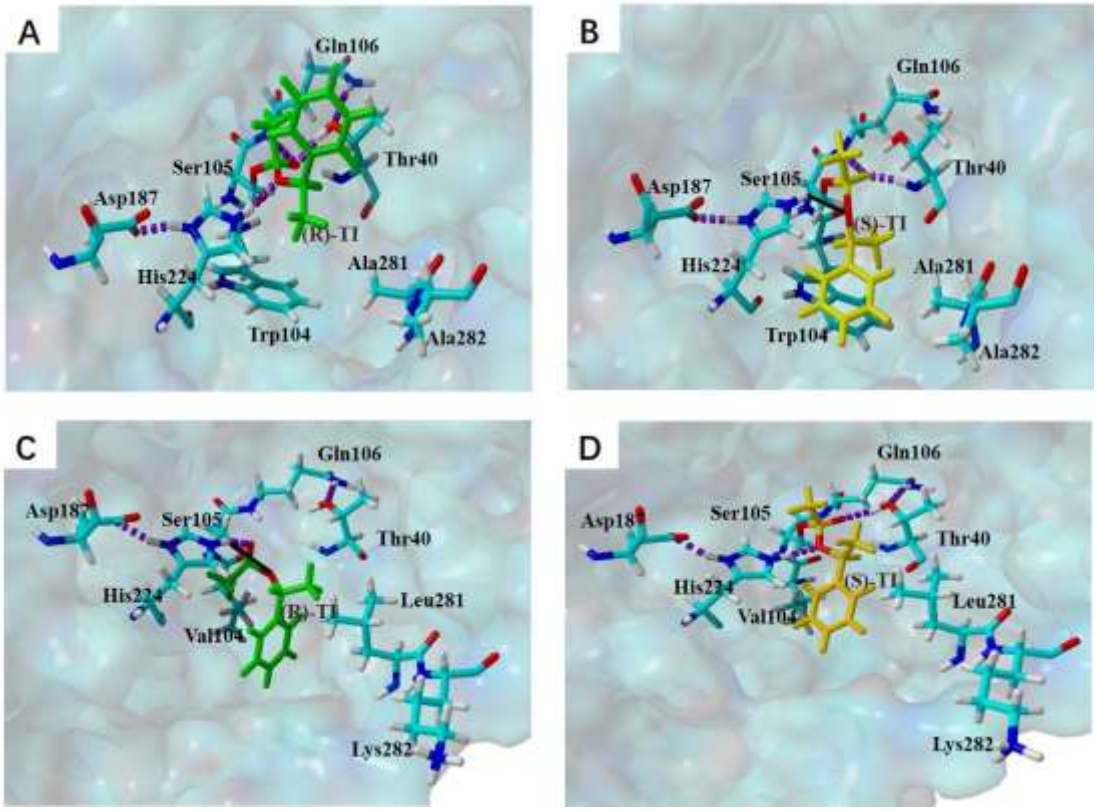


Figure 3. (**A**) A productive binding mode and a hydrogen bond networks of the (*R*)-**1a** complex with WT CALB having Ser105-His224-Asp187 as the catalytic triad. The H-bond networks for nucleophilic attack and stabilizing the oxyanion are indicated by purple dotted lines. (**B**) A nonproductive orientation found for the less reactive (*S*)-**1a**-CALB complex, revealing that the H-bond between H224 (HE2 atom) and the bound ester oxygen no longer exists. (**C**) and (**D**) are the optimized binding modes of the less reactive (*R*)-**1a** and the preferential (*S*)-**1a**, respectively, in (*S*)-selective variant WB15.

Conclusion

To reverse the inherent (*R*)-enantioselectivity of wild-type *Candida antarctica* lipase B (CALB) towards chiral *sec*-alcohol, an efficient directed evolution strategy was proposed to construct a highly focused library containing only 20 variants, which drastically reduced the screening effort. W104 was selected as the first targeting site to perform "subsaturation" mutagenesis by introducing seven residues with a volume smaller than that of Trp, as amino acids exchange. The variant W104V displayed better reversed (*S*)-selectivity than the W104A mutant, and it was found that an exchange with smaller amino acids at the W104 site is not necessarily better. Then, further iteration mutagenesis at A281 and A282 sites by using only 4 representative amino acids of A, L, K and W was adopted. To our great delight, we finally have found that WB15 (W104V/A281L/A282K) has a good *E* value (*E* = 80) with (*S*)-selectivity in the hydrolysis of model substrate and displays broad substrate scope of various phenylalkanol esters. Deconvolution experiments implied the crucial role played by W104V mutation for reversing stereoselectivity and the synergistic effects of these mutations W104V, W104V/A281L and W104V/A282K in the best variant WB15. Structural analysis and molecular dynamics simulation unveiled the source of reversed enantioselectivity. The enlargement of alcohol-binding pocket caused by point mutation at W104 position and the modification of an active site entrance at A281/A282 positions allow (*S*)-configurational *sec*-alcohol to form a productive complex. The construction strategy of highly focused library demonstrated herein enables us to engineer various lipases more conveniently and efficiently to obtain the tailor-made biocatalysts and optically pure products in further investigation.

Experimental Section

Materials and Methods

Unless otherwise specified, all reagents were obtained from commercial sources and used without further purification. Pfu polymerase and Dpn I were purchased from Thermo Fischer Scientific Inc. The E.Z.N.A.® Plasmid Midi Kit was supplied by Omega-Bio Tek. The ^{1}H NMR and ^{13}C NMR spectra were recorded with a Bruker AMX400 MHz spectrometer using TMS as an internal standard. All spectra were recorded at room temperature in $CDCl_3$, the following abbreviations were used to describe peak splitting patterns when appropriate: s = singlet, d = doublet, t = triplet, q = quartet, m = multiplet. Coupling constants were reported in Hertz (Hz). The

High-Resolution Mass Spectrometer was obtained from the WATERS GCT Premier with Electron Impact Ion Source.

Chiral GC analysis was performed with a Shimadzu GC-2014C instrument (Shimadzu, Japan) equipped with an FID detector. Samples were analyzed on an Agilent CP-Chirasil-Dex CB column (25 m × 0.25 mm × 0.1 μm, Teknokroma, Spain), under the following conditions: carrier gas: N_2; detector temperature, 220 °C. Chiral HPLC was performed with a Chiralpak OJ-H column (250 mm×4.6 mm, n-hexane/2-propanol as the mobile phase) and a UV detector (220nm). Detailed chiral separation conditions are shown in Supporting Information.

Library Generation

WT-CALB plasmid (pETM11-CALB) was used as DNA template, and various primers are shown in Table S2 in the Supporting Information. The 50 μL PCR mix contained ddH_2O (29 μL), Pfu 10X buffer (5 μL), dNTP (4 μL, 2.5 mM each), forward primers (5 μL, 2.5 μM each), silent reverse primer (5 μL, 2.5 μM), template plasmid (1 μL, 100 ng/μL) and 1 μL of Pfu polymerase. PCR reactions were performed at 94 °C for 5 min; followed by denaturation at 94 °C for 1 min, annealing at 60 °C for 1 min, and elongation at 72 °C for 14 min repeated 30 times. A final extension at 72 °C for 10 min ended the reaction. To ensure elimination of the circular methylated template plasmid, 20 μL of PCR reaction mixture were mixed with 2 μL DpnI (10 U/μL) and incubated overnight at 37 °C. The Dpn I-digested product was purified with an Omega PCR purification spin-column, and the plasmid (6 μL) was used to transform 80 μL of E. coli Origami2 electrocompetent cells (containing chaperone plasmid pGro7, Takara, Japan). The transformation mixture was incubated with 1 mL of LB medium at 37 °C with shaking at 200 rpm and spread on LB-agar plates containing kanamycin (34 μg/mL) and chloramphenicol (34 μg/mL).

Expression of CALB Mutants

Colonies appeared after cultivation for 12-16 h at 37 °C and were picked into 5.0 mL LB medium containing kanamycin (34 μg/mL) and chloramphenicol (34 μg/mL), and then incubated at 37 °C under shaking at 200 rpm overnight. A fresh 200 mL of TB media was added to 2mL preculture containing 1.0 mg/mL L-arabinose as the inducer for expression of chaperone pGro7, kanamycin (34 μg/mL) and chloramphenicol (34 μg/mL). The cultures were shaken at 37 °C until the optical density at 600 nm reached 0.6, and then cooled to 4°C for 1h. Then, 1.0 mM isopropyl β-thiogalactopyranoside (IPTG) added into the cultures to induce CALB expression for 48 h at 18 °C. The tubes were centrifuged at 4000 rpm and 4 °C for 25 min, and then the supernatants were discarded. The cell pellet of each tube was resuspended in 50 mM Tris-HCl (pH 7.5) and lysed by sonification (Bandelin, 15×10 sec with 10 sec intervals, at 40% pulse, in a water-ice bath). The cell debris was removed by centrifugation for 25 min at 4 °C. The supernatant was stored at -78 °C.

Molecular Modeling and MD Simulation

The molecular modeling and MD simulation were performed using YASARA structure (version 15.5.31).[25] The AMBER03 force field[26] with default settings was used for the protein and AutoSMILES force field assignment for the substrates and the tetrahedral intermediates.[27] Energy minimization in the water environment was performed using the steepest descent and simulated annealing simulations. The MD simulation was performed at 298 K using a time-step of 1.25 fs for inter- and intramolecular forces over 10 ns in an NPT ensemble using PME.[28]

Models of the enzyme variants with bound **1a** were based on the crystal structure (PDB ID: 1TCA). The (*R*)- and (*S*)-**1a** substrates were allowed to equilibrate for 10 ns by a molecular dynamics simulation in WT and WB15, respectively.

Acknowledgements

The financial support from National Natural Science Foundation of China (21472169, 21574113) and the Fundamental Research Funds for the Central Universities (2018QNA3010) is gratefully acknowledged.

References

[1] X. Wu, J. Xiao, *Chem. Commun.* **2007**, *38*, 2449-2466.

[2] O. Jurcek, M. Wimmerova, Z. Wimmer, *Coord. Chem. Rev*. **2008**, *252*, 767-781.

[3] G. W. Huisman, J. Liang, A. Krebber, *Curr. Opin. Chem. Biol*. **2010**, *14*, 122-129.

[4] S. H. Schöfer, N. Kaftzik, P. Wasserscheid, U. Kragl, *Chem. Commun*. **2001**, *5*, 425-426.

[5] W. D. Fessner, T. Anthonsen, Modern biocatalysis: stereoselective and environmentally friendly reactions, Wiley-VCH, Weinheim, **2009**.

[6] R. N. Patel, *Stereoselective Biocatalysis*, Chemical Industry Press, Beijing, **2004**.

[7] L. S. Chua, M. R. Sarmidi, *J. Mol. Catal. B: Enzym.* **2004**, *28*, 111-119.

[8] A. P. de los Ríos, F. V. Rantwijk, R. A. Sheldon, *Green Chem*. **2012**, *14*, 1584-1588.

[9] a) R. J. Kazlauskas, A. N. E. Weissfloch, A. T. Rappaport, L. A. Cuccia, *J. Org. Chem.* **1991**, *56*, 2656-2665; b) K. E. Jaeger, T. Eggert, *Curr. Opin. Biotechnol.* **2002**, *13*, 390–397; c) A. S. de Miranda, L.

S. Miranda, R. O. de Souza, *Biotechnol. Adv.* **2015**, *33*, 372-393.

[10] M. J. Kim, Y. I. Chung, Y. K. Choi, H. K. Lee, D. Kim, J. Park, *J. Am. Chem. Soc.* **2003**, *125*, 11494–11495.

[11] L. Borén, B. Martin-Matute, Y. Xu, A. Córdova, J. E. Bäckvall, *Chem. Eur. J.* **2005**, *12*, 225-232.

[12] a) G. A. Strohmeier, H. Pichler, O. May, M. Gruber-Khadjawi, *Chem. Rev.* **2011**, *111*, 4141–4164; b) L. G. Otten, F. Hollmann, I. W. C. E. Arends, *Trends Biotechnol.* **2010**, *28*, 46-54; c) A. Díaz-Rodríguez, B. G. Davis, *Curr. Opin. Chem. Biol.* **2011**, *15*, 211–219; d) A. S. Bommarius, J. K. Blum, M. J. Abrahamson, *Curr. Opin. Chem. Biol.* **2011**, *15*, 194–200; e) U. T. Bornscheuer, G. W. Huisman, R. J. Kazlauskas, S. Lutz, J. C. Moore, K. Robins, *Nature* **2012**, *485*, 185-194; f) C. K. Prier, F. H. Arnold, *J. Am. Chem. Soc.* **2015**, *137*, 13992−14006.

[13] Recent reviews of directed evolution: a) N. J. Turner, *Nat. Chem. Biol.* **2009**, *5*, 567–573; b) C. Jäckel, D. Hilvert, *Curr. Opin. Biotechnol.* **2010**, *21*, 753–759; c) M. T. Reetz, *Angew. Chem., Int. Ed.* **2011**, *50*, 138–174; d) E. M. Brustad, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2011**, *15*, 201–210; e) A. Currin, N. Swainston, P. J. Day, D. B. Kell, *Chem. Soc. Rev.* **2015**, *44*, 1172–1239; f) C. A. Denard, H. Ren, H. Zhao, *Curr. Opin. Chem. Biol.* **2015**, *25*, 55–64; g) H. Renata, Z. J. Wang, F. H. Arnold, *Angew. Chem., Int. Ed.* **2015**, *54*, 3351–3367; h) A. S. Bommarius, *Annu. Rev. Chem. Biomol. Eng.* **2015**, *6*, 319–345; i) S. C. Hammer, A. M. Knight, F. H. Arnold, *Curr. Opin. Green Sustain. Chem.* **2017**, *7*, 23-30.

[14] M. T. Reetz, J. D. Carballerira, *Nat. Protoc.* **2007**, *2*, 891-903.

[15] a) U. T. Bornscheuer, R. J. Kazlauskas, Hydrolases in Organic Synthesis: Regio- and Stereoselective Biotransformations, Wiley-VCH, Weinheim, **1999**; b) K. E. Jaeger, B. W. Dijkstra, M. T. Reetz, *Annu. Rev. Microbiol.* **1999**, *53*, 315−351; c) O. Kirk, M. W. Christensen, *Org. Prog. Res. Dev.* **2002**, *6*, 446−451; d) A. Baldessari, L. E. Iglesias, *Methods Mol. Biol.* **2012**, *861*, 457−469.

[16] S. Naik, A. Basu, R. Saikia, B. Madan, P. Paul, R. Chaterjee, J. Brask, A. Svendsen, *J. Mol. Catal. B: Enzym.* **2010**, *65*, 18−23.

[17] J. Uppenberg, M. T. Hansen, S. Patkar, T. A. Jones. *Structure*. **1994**, *2*, 293–308.

[18] J. Uppenberg, N. M. Oehrner, K. Hult, G. J. Kleywegt, S. Patkar, *Biochemistry* **1995**, *34*, 16838−16851.

[19] R. T. Otto, H. Scheib, U. T. Bornscheuer, J. Pleiss, C. Syldatk, R. D. Schmid. *J. Mol. Catal. B: Enzym.* **2000**, *8*, 201−211.

[20] a) P. B. Juhl, K. Doderer, F. Hollman, O. Thum, J. Pleiss, *J. Biotechnol.* **2010**, *150*, 474-480. b) Q. Wu, P. Soni, M. T. Reetz, *J. Am. Chem. Soc.* **2013**, *135*, 1872–1881.

[21] A. O. Magnusson, J. C. Rotticcimulder, A. Santagostino, K. Hult, *Chembiochem* **2005**, *6*, 1051-1056.

[22] A. O. Magnusson, M. Takwa, A. Hamberg, K. Hult, *Angew. Chem. Int. Ed.* **2005**, *44*, 4582−4585.

[23] M. Vallin, P. Syrén, K. Hult, *ChemBioChem* **2010**, *11*, 411-416.

[24] E. Krieger, G. Koraimann, G. Vriend, *Proteins: Struct., Funct., Bioinf.* **2002**, *47*, 393-402.

[25] E. Krieger, T. Darden, S. B. Nabuurs, A. Finkelstein, G. Vriend, *Proteins: Struct., Funct., Bioinf.* **2004**, *57*, 678-683.

[26] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. Kollman, *J. Comput. Chem.* **2003**, *24*, 1999-2012.

[27] A. Jakalian, D. B. Jack, C. I. Bayly, *J. Comput. Chem* **2002**, *23*, 1623-1641.

[28] U. Essmann, L. Perera, M. L. Berkowitz, *J. Chem. Phys.* **1995**, *103*, 8577-8593.

[29] C. Orrenius, F. Hbffner, D. Rotticci, N. öhrner, T. Norin, K. Hult, *Biocatal. Biotransform.* **1998**, *16*, 1-15.

10.1002/adsc.201800711

FULL PAPER

Highly Focused Library-Based Engineering of *Candida antarctica* Lipase B with (*S*)-Selectivity Towards *sec*-Alcohols

Adv. Synth. Catal. **2018**, *Volume*, Page – Page

Yixin Cen,[+a] Danyang Li,[+a] Jian Xu,[a] Qiongsi Wu,[a] Qi Wu[a,*] and Xianfu Lin[a,*]