

In presenting this dissertation/thesis as a partial fulfillment of the requirements for an advanced degree from Emory University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations, governing materials of this type. I agree that permission to copy from, or to publish, this thesis/dissertation may be granted by the professor under whose direction it was written, or, in his/her absence, by the Dean of the Graduate School when such copying or publication is solely for scholarly purposes and does not involve potential financial gain. It is understood that any copying from, or publication of, this thesis/dissertation which involves potential financial gain will not be allowed without written permission.

**Engineering *Candida antarctica* lipase B by circular
permutation and incremental truncation**

By

Zhen Qian
Doctor of Philosophy

Department of Chemistry

Dr. Stefan Lutz
Adviser

Dr. Dale E. Edmondson
Committee Member

Dr. Vince Conticello
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the Graduate School

Date

**Engineering *Candida antarctica* Lipase B by Circular
Permutation and Incremental Truncation**

By

Zhen Qian

B.S., Sichuan University, 1997

M.S., East China University of Science and Technology, 2000

Advisor: Stefan Lutz, Ph.D.

An Abstract of

A dissertation submitted to the Faculty of the Graduate
School of Emory University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Chemistry

2007

Abstract

The lipase B from *Candida antarctica* (CALB) is a versatile biocatalyst with broad applications in industry and in organic synthesis. In recent years, a growing interest has been put on CALB engineering to broaden its applications in asymmetric synthesis. In this dissertation a combined engineering strategy – circular permutation and incremental truncation – was employed, which yielded variants with greatly enhanced catalytic performance.

A random circularly permuted CALB library was generated and subsequent library screening identified 63 unique functional variants, whose termini constitutes approximately 20% positions of the whole polypeptide chain. Particularly interesting from the catalysis standpoint are permutations in the enzyme's putative lid and in the long helix $\alpha 10$ which flanks the active site pocket. Selected library members were characterized by kinetic, circular dichroism, and enantioselectivity analysis. The data demonstrated that circular permutation could greatly improve the catalytic performance of CALB, while the variants' overall three-dimensional structure and high enantioselectivity are retained.

Secondly, incremental truncation was employed to increase the thermostability of the most active permutant cp283 (circularly permuted CALB whose N-terminus starts at amino acid 283 of the wild-type sequence). An incremental truncation library was generated by truncating an extended loop formed by the original termini and the linker. Library screening identified functional variants with up to 11 amino acid truncation. Characterization of selected library members showed that the variants maintain their high activity as the parental cp283, and have improved thermostability. Both biochemical

analysis and crystallographic study verified that thermostabilization is conferred by protein dimerization. Crystal structure of the best variant cp283- Δ 7 was solved, which revealed a domain-swapped dimer structure with the N-terminal segment (Ala283~Val315) swapped between two subunits, which to our knowledge is the first example of a dimeric α/β hydrolase. The crystal structure also provides explanations to experimental results obtained on other circularly permuted CALB variants.

Overall, the results from this dissertation have shown that circular permutation could greatly improve the catalytic efficiency of CALB, and circular permutation in combination with incremental truncation has yielded better CALB variants with high activity, high enantioselectivity and moderate thermostability, which would be advantageous for future applications in organic synthesis.

**Engineering *Candida antarctica* Lipase B by Circular
Permutation and Incremental Truncation**

By

Zhen Qian

B.S., Sichuan University, 1997

M.S., East China University of Science and Technology, 2000

Advisor: Stefan Lutz, Ph.D.

A dissertation submitted to the Faculty of the Graduate
School of Emory University in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

Department of Chemistry

2007

Acknowledgement

First of all, I would like to thank my advisor Dr. Stefan Lutz for his support to my graduate study. It was a great opportunity and experience for me to work in his lab, and I really appreciate his help, encouragement and guidance during the last four years.

I would like to thank my committee members Dr. Dale E. Edmondson and Dr. Vincent Conticello for their insightful and critical suggestions, and their technical and instrumental assistance to my research project, which allowed it to proceed smoothly. I would also like to thank Dr. Min Li and Jin Wang in Dr. Edmondson's lab for providing me with the *P. Pastoris* strains and assisting me in yeast fermentation.

Part of my dissertation is in collaboration with Dr. Xiaodong Chen's lab at Emory University. I'm very grateful to Dr. Xiaodong Cheng, Dr. John R. Horton, and Dr. Da Jia for their help in preparing the protein crystals and in solving the crystal structures.

I would like to thank all past and present group members, whose friendship and help made my graduate study a wonderful memory. My special thanks go to Dr. Wayne Patrick, who introduced me into the world of biochemistry and taught me fundamental experimental techniques in molecular biology. Thanks to Christina J. Fields who synthesized several substrates for my project.

No words could express how thankful I am to my husband Zhongbo Fei, my parents Yongfa Qian and Wanzhu Zheng. Their endless love, support and encouragement are the sources of strength for me throughout these years. Especially my husband Zhongbo Fei, whose accompany, love and support are treasures in my life.

Table of contents

Chapter 1 Introduction	1
1.1 A general introduction to lipases	1
1.1.1 Reactions catalyzed by lipases.....	1
1.1.2 Applications of lipases.....	1
1.1.3 The fold of lipases.....	3
1.2 <i>Candida antarctica</i> lipase B.....	4
1.2.1 Structural features and catalytic mechanism of CALB.....	4
1.2.2 Thermostability, substrate specificity, and enantioselectivity of CALB	11
1.2.3 Overexpression of CALB	14
1.2.4 CALB engineering.....	15
1.3 Circular permutation of proteins.....	19
1.3.1 Naturally occurring circular permutations in proteins.....	19
1.3.1.1 Mechanisms of circular permutation in nature	19
1.3.1.2 Functional importance of circular permutation in nature	19
1.3.1.3 Circularly permuted α/β hydrolase-fold protein in nature	22
1.3.2 Circular permutation as a genetic engineering approach.....	25
1.3.2.1 Rational design of circularly permuted proteins.....	25
1.3.2.2 Random circular permutation	27

Chapter 2 Improving the catalytic activity of *Candida antarctica* lipase B

by circular permutation	30
2.1 Introduction.....	30
2.2 Materials and methods	31
2.2.1 Construction of wt-CALB expression vector	31
2.2.2 Random circulation permutation of CALB gene	32
2.2.3 Creation of the pPIC9-cp-CALB library.....	33
2.2.3.1 Direct library cloning.....	33
2.2.3.2 Two-step library cloning.....	35
2.2.4 Creation of the (His) ₆ -tag free pPIC9-cp-CALB library.....	37
2.2.5 Library screening	38
2.2.6 Protein expression, purification, and activity assays	40
2.3 Results and discussions.....	40
2.3.1 Creation of the cp-CALB library	40
2.3.2 Library screening and analysis.....	41
2.3.3 Selected library members for kinetic characterization.....	45
2.4 Conclusions.....	48

Chapter 3 Characterization of circularly permuted CALB variants .49

3.1 Introduction.....	49
3.2 Materials and methods	50
3.2.1 Construction of His-free protein expression vector	50
3.2.2 Protein expression, purification, and activity assays	52

3.2.3 Circular dichroism analysis.....	52
3.2.4 <i>P. pastoris</i> fermentation.....	52
3.2.5 Protein immobilization.....	52
3.2.6 Active site titration of immobilized lipase.....	53
3.2.7 Kinetic analysis of lipase catalyzed esterification reactions.....	53
3.3 Results and discussions.....	55
3.3.1 Kinetic analysis of (His) ₆ -tag and C-terminal extension free circularly permuted CALB variants.....	59
3.3.2 Circular dichroism analysis of selected variants.....	62
3.3.3 Enantioselectivity comparison of cp283 with wt-CALB.....	64
3.4 Conclusions.....	67

Chapter 4 Secondary engineering and thermostabilization of circularly permuted CALB68

4.1 Introduction.....	68
4.2 Materials and methods	71
4.2.1 Creation of the cp283 incremental truncation library	71
4.2.2 Creation of the C-terminal incremental truncation library	72
4.2.3 Protein expression, purification, and activity assays	73
4.2.4 Circular dichroism analysis and gel filtration.....	73
4.3 Results and Discussions.....	74
4.3.1 Creation of the incremental truncation libraries	74
4.3.1.1 Creation of the cp283 incremental truncation library	74
4.3.1.2 Analysis of the cp283 incremental truncation library.....	76

4.3.1.3 Creation and analysis of the C-terminal incremental truncation library...	79
4.3.2 Characterization of selected variants	80
4.3.2.1 Kinetic analysis on selected variants	83
4.3.2.2 Circular dichroism analysis on selected variants	83
4.3.2.3 Gel filtration analysis on selected variants	84
4.3.3 Stabilization mechanism of the cp283 truncation variants	89
4.4 Conclusions.....	91

Chapter 5 The crystal structure of cp283- Δ 7: circular permutation and domain swapping92

5.1 Introduction.....	92
5.2 Materials and methods	96
5.2.1 Crystallization of cp283- Δ 7	96
5.2.1.1 cp283- Δ 7 expression and purification	96
5.2.1.2 cp283- Δ 7 crystallization	96
5.2.1.3 Crystal data collection.....	97
5.2.2 Construction and characterization of the truncation variants	99
5.2.2.1 Construction of the truncation variants.....	99
5.2.2.2 Protein expression, purification and activity assays	99
5.2.2.3 Circular dichroism analysis of the variants.....	99
5.3 Results and discussions.....	100
5.3.1 The crystal structure of cp283- Δ 7	100
5.3.1.1 Crystal structure of cp283- Δ 7	100
5.3.1.2 Energetic contributions to domain swapping and thermostabilization ...	102

5.3.1.3 Comparison of structures of wt-CALB with cp283- Δ 7	104
5.3.1.4 Explanations for experimental results on circularly permuted variants..	104
5.3.2 Truncation study on cp283- Δ 7 and cp283	107
5.3.2.1 Creation of the truncation variants.....	107
5.3.2.2 Kinetic analysis of the truncation variants.....	109
5.3.2.3 Circular dichroism analysis of the truncation variants	109
5.4 Conclusions.....	112
Chapter 6 Conclusions and perspectives.....	113
Appendices.....	116
A Materials and media	116
A1 Chemicals.....	116
A2 Enzymes	117
A3 Strains and plasmids.....	118
A4 Buffers and Media.....	118
B General Methods	119
B1 Protein expression in <i>P. pastoris</i>	119
B2 Protein purification by Ni-NTA column	119
B3 Protein purification by HIC and gel filtration	120
B4 Activity assays.....	121
B5 <i>P. pastoris</i> fermentation	121
B6 Circular dichroism analysis.....	123
References.....	124

List of figures

Figure 1-1 Reactions catalyzed by lipases	2
Figure 1-2 Secondary structure diagram of the canonical α/β hydrolase fold	5
Figure 1-3 Three dimensional structure of CALB	6
Figure 1-4 Secondary structure diagram of CALB	7
Figure 1-5 Reaction mechanism of CALB	9
Figure 1-6 The active site pocket of CALB	10
Figure 1-7 Model for predicting the fast enantiomer of secondary alcohols	13
Figure 1-8 Schematic diagram of protein circular permutation	20
Figure 1-9 Schematic illustration of the mechanisms of circular permutation	21
Figure 1-10 Examples of circularly permuted proteins	23
Figure 1-11 Topology diagram of the PHB depolymerase (a) in comparison to the canonical α/β hydrolase fold (b)	24
Figure 1-12 Creation of a random circular permutation library	29
Fig 2-1 Schematic overview of the direct library cloning	34
Figure 2-2 Schematic overview of the two-step library cloning	36
Figure 2-3 Structures of the tertiary alcohol esters used in library screening	39
Figure 2-4 Circular permutation of CALB	42
Figure 2-5 Library screening on tributyrin plates	43
Figure 2-6 New termini mapped on wild-type CALB (1tca)	44
Figure 3-1 Selected ten variants for characterization	57

Figure 3-2 SDS-PAGE gel example of purified proteins.....	58
Figure 3-3 Far-UV circular dichroism (CD) spectra and T_m of selected variants.....	63
Figure 4-1 Schematic overview of the structures of wt-CALB and cp283.....	70
Figure 4-2 Ribbon diagram of the WT α -spectrin SH3 domain and the circular permutants	70
Figure 4-3 Graphic illustration of the sequences of wt-CALB and cp283	75
Figure 4-4 Creation of the incremental truncation library	75
Figure 4-5 Incremental truncation library screening on tributyrin plates	77
Figure 4-6 Sequences of variants picked from the native library sorted by the length of truncation	78
Figure 4-7 The first step in the creation of the C-terminal truncation library	81
Figure 4-8 Structure model of IT301	81
Figure 4-9 Selected cp283 truncation variants for characterization	82
Figure 4-10 Circular dichroism analysis of selected variants	86
Figure 4-11 The Ellman's assay on selected variants	87
Figure 4-12 Gel filtration analysis of selected variants.....	88
Figure 4-13 Circular dichroism analysis of $\Delta 4$ -monomer and $\Delta 4$ -dimer	90
Figure 5-1 Graphic illustration of the sequences of wt-CALB, cp283 and cp283- $\Delta 7$	94
Figure 5-2 Model of cp283- $\Delta 7$ dimerization	94
Figure 5-3 Domain-swapped dimer structure of staphylococcal nuclease	95
Figure 5-4 Crystal structure of the cp283- $\Delta 7$ dimer	101
Figure 5-5 Hinge loop truncation that leads to domain swapping.....	103
Figure 5-6 Interface between two subunits of the dimer.....	103

Figure 5-7 Structure alignment of wt-CALB with one subunit of cp283- Δ 7 dimer	105
Figure 5-8 Gel filtration analysis of the truncation variants.....	108
Figure 5-9 Far-UV circular dichroism (CD) spectra and T_m of the truncation variants..	111

List of tables

Table 1-1 Industrial applications of lipases	2
Table 1-2 Precedents of CALB engineering	18
Table 2-1 Apparent kinetic constants for selected CALB variants	47
Table 3-1 Primers used for construction of His-tag free variants	51
Table 3-2 Apparent kinetic constants for selected CALB variants	61
Table 3-3 Kinetic and enantioselectivity analysis on lipases (wt-CALB and cp283) catalyzed transesterification reaction with vinyl acetate	66
Table 3-4 Kinetic and enantioselectivity analysis on lipases (wt-CALB and cp283) catalyzed esterification reactions with n-propanol	66
Table 4-1 Kinetic data of p-nitrophenyl butyrate hydrolysis for selected variants	85
Table 4-2 Kinetic analysis of Δ 4-monomer and Δ 4-dimer on p-NB hydrolysis	90
Table 5-1 Crystallization condition comparison of wt-CALB and cp283- Δ 7	98
Table 5-2 Comparison of crystal parameters	98
Table 5-3 Sequences of the truncation variants	108
Table 5-4 Apparent kinetic constants for truncation variants	110

Chapter 1 Introduction

1.1 A general introduction to lipases

1.1.1 Reactions catalyzed by lipases

Lipases (EC 3.1.1.3, triacylglycerol hydrolases) are ubiquitous enzymes with important physiological functions and considerable industrial potential (Sharma, Chisti et al. 2001; Hasan, Shah et al. 2006). Physiologically, lipases hydrolyze triglycerides into diglycerides, monoglycerides, fatty acids and glycerol. In addition to their natural function of hydrolyzing carboxylic ester bonds, lipases also catalyze the reverse esterification, amidation or transesterification reactions in aqueous and anhydrous organic solvents (Bornscheuer and Kazlauskas 1999) (Figure 1-1). This diversity, in combination with their broad substrate specificity, stability in organic solvents and usually high regio- and/or stereo- selectivity, makes lipases the enzymes of choice for a broad range of applications in industries as well as in academic organic synthesis (Anderson, Larsson et al. 1998; Krishna and Karanth 2002).

1.1.2 Applications of lipases

As shown in Table 1-1, lipases constitute an important group of enzymes in industry with wide applications in the processing of fats and oils, in the manufacture of detergents, food and paper, and in the production of cosmetics and pharmaceuticals (Sharma, Chisti et al. 2001).

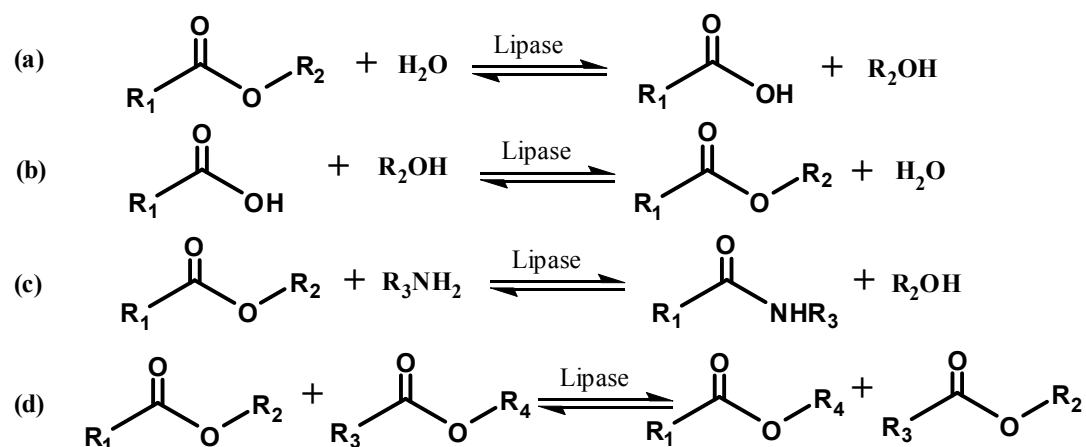


Figure 1-1 Reactions catalyzed by lipases (a) Ester hydrolysis; (b) Esterification; (c) Amidation; (d) Transesterification.

Industry	Action	Product or application
Detergents	Hydrolysis of fats	Removal of oil stains from fabrics
Dairy foods	Hydrolysis of milk fat, cheese ripening, modification of butter fat	Development of flavoring agents in milk, cheese, and butter
Bakery foods	Flavor improvement	Shelf-life prolongation
Beverages	Improved aroma	Beverages
Food dressings	Quality improvement	Mayonnaise, dressings, and whippings
Health foods	Transesterification	Health foods
Meat and fish	Flavor development	Meat and fish products, fat removal
Fats and oils	Transesterification; hydrolysis	Cocoa butter, margarine, fatty acids, glycerol, mono-, and diglycerides
Chemicals	Enantioselectivity, synthesis	Chiral building blocks, chemicals
Pharmaceuticals	Transesterification; hydrolysis	Specialty lipids, digestive aids
Cosmetics	Synthesis	Emulsifiers, moisturizers
Leather	Hydrolysis	Leather products
Paper	Hydrolysis	Paper with improved quality
Cleaning	Hydrolysis	Removal of fats

Table 1-1 Industrial applications of lipases (Sharma, Chisti et al. 2001)

Specifically, in recent years a growing interest has been put on the utilization of lipases for the preparation of chiral synthons. Chirality is a key factor in the efficacy of many drugs and a major obstacle in the synthesis of various chiral building blocks. The high regio- and enantioselectivity of lipases and their relatively mild reaction conditions provide an appealing alternative for the resolution of racemic mixtures and the synthesis of chiral intermediates (Anderson, Larsson et al. 1998; Krishna and Karanth 2002; Hasan, Shah et al. 2006). Currently lipases are used by many pharmaceutical companies worldwide for the preparation of optically active intermediates on a kilogram scale, for example the muscle relaxant baclofen and the anti-inflammatory drug flurbiprofen. Meanwhile, the high enantioselectivity of lipases is well-recognized by organic chemists and their role in organic synthesis has dramatically increased in the last decade (Gotor-Fernandez, Busto et al. 2006).

1.1.3 The fold of lipases

Lipases are members of the α/β hydrolase-fold family – one of the largest groups of structurally related enzymes with diverse catalytic functions (Holmquist 2000). The α/β hydrolase fold was identified in 1992 by comparing structures of five hydrolytic enzymes including diene lactone hydrolase, haloalkane dehalogenase, wheat serine carboxypeptidase II, acetylcholinesterase and the lipase from *Geotrichum candidum* (Ollis, Cheah et al. 1992; Nardini and Dijkstra 1999). Since then, a lot of enzymes with different catalytic functions have been found to belong to the same superfamily, such as thioester hydrolase, peptide hydrolase, haloperoxidase, epoxide hydrolase and C-C bond breaking enzymes (Holmquist 2000). A common structural feature of the α/β hydrolase-fold enzymes is a core composed of mostly parallel β sheets flanked on both sides by α

helices (Jaeger, Dijkstra et al. 1999) (Figure 1-2). The positions of the helices in three dimensions vary considerably from one enzyme to another, but all the members in the family, including lipases, contain a catalytic triad at the C-terminal ends of the β sheets composed of a nucleophile (serine, cysteine or aspartic acid), an acidic residue, and an absolutely conserved histidine residue. The linear sequence of the catalytic triad is always nucleophile-acid-His in α/β hydrolase-fold enzymes, which is different from a linear sequence of His-Asp-Ser for the trypsin family, Cys-His-Asn for the papain family, and Asp-His-Ser for the subtilisin family (Schrag and Cygler 1997).

1.2 *Candida antarctica* lipase B

Candida antarctica lipase B (CALB) is a member of the lipase family. As suggested by its name, CALB originates from the yeast *Candida antarctica*. The strain was originally isolated from Antarctica in a search for enzymes with extreme properties (Anderson, Larsson et al. 1998). Two different lipases are produced by the strain, lipase A (CALA) and lipase B (CALB). Both lipases have been purified and characterized (Patkar, Bjoerking et al. 1993; Hoegh, Patkar et al. 1995).

1.2.1 Structural features and catalytic mechanism of CALB

CALB is composed of 317 amino acids with a molecular weight 33 kDa. Its crystal structure was solved in 1994 by the Uppenberg group, which revealed a globular α/β hydrolase-fold structure with seven central strands of β sheets connected by loops and α helices (Uppenberg, Hansen et al. 1994) (Figure 1-3, 1-4). The active site of CALB is located at the bottom of a pocket on top of the central β sheets. Same as other members in the α/β hydrolase-fold family, catalysis is carried out by a catalytic triad Ser105-Asp187-

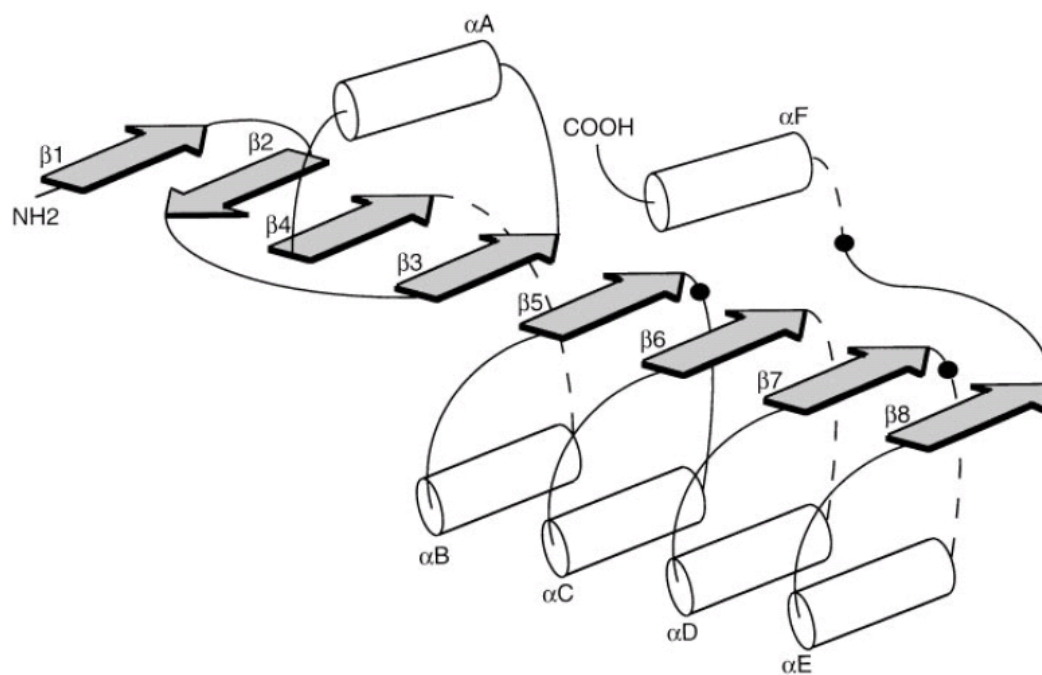


Figure 1-2 Secondary structure diagram of the canonical α/β hydrolase fold. α helices and β strands are indicated by cylinders and shaded arrows respectively. The location of the catalytic triad is indicated by black dots. The location of possible insertions are indicated by dashed lines (Jaeger, Dijkstra et al. 1999; Nardini and Dijkstra 1999).

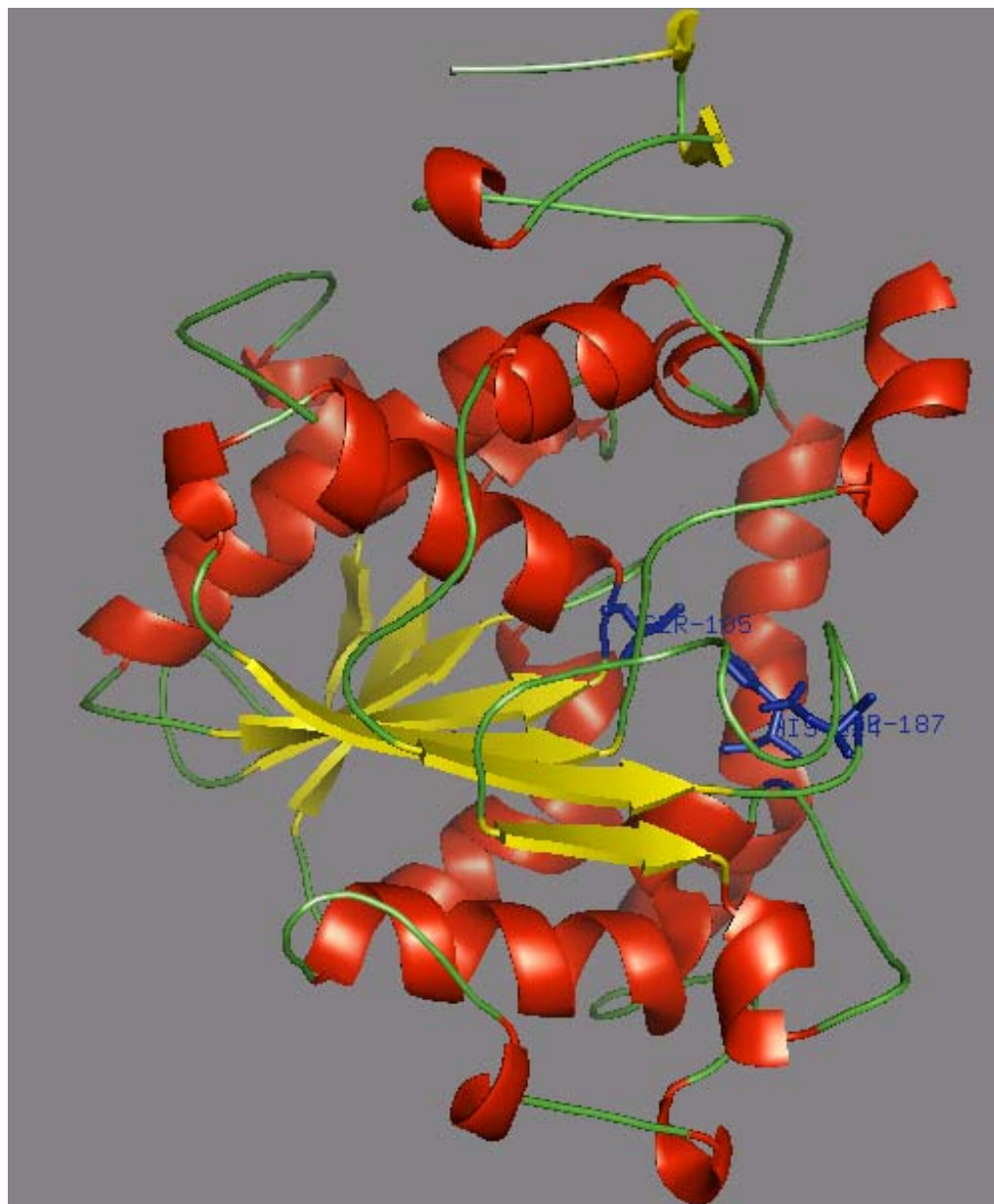


Figure 1-3 Three dimensional structure of CALB: Seven central β sheets are shown in yellow, helices are shown in red, and the catalytic triad is shown in sticks in blue (Uppenberg, Hansen et al. 1994).

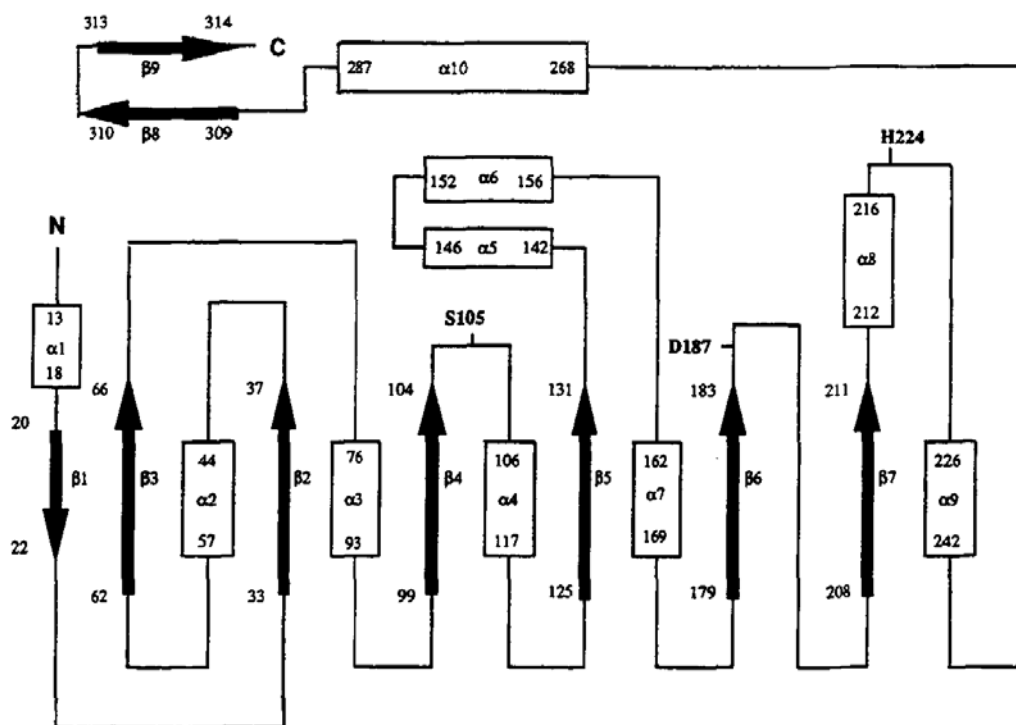


Figure 1-4 Secondary structure diagram of CALB. α helices are shown in cylinders and β sheets are shown in black arrows. (Uppenberg, Hansen et al. 1994)

His224 following a bi-bi ping-pong mechanism (Martinelle and Hult 1995; Rotticci 2000; Rotticci-Mulder 2003).

As illustrated in Figure 1-5, the catalytic triad Ser105-His224-Asp187 forms a charge relay which greatly enhances the nucleophilicity of Ser105. In the first step of the reaction, the enzyme is acylated by the first substrate and forms a tetrahedral intermediate. The oxyanion in the intermediate is stabilized by the oxyanion hole, which is a spatial arrangement of three hydrogen bond donors with one from the side chain of Thr40 and the other two from the backbone amides of Thr40 and Gln106 (Magnusson 2005). After the release of the first product, the second substrate enters the active site and the second tetrahedral intermediate is formed. The enzyme is then deacylated and ready for the next catalytic cycle (Rotticci 2000; Rotticci-Mulder 2003).

Three helices, helix 5, 6, and 10 make up most of the wall of the funnel shaped active site pocket (Figure 1-6) (Uppenberg, Hansen et al. 1994). Upon substrate binding, the acyl moiety of the substrate lies in the acyl binding pocket above the catalytic triad, while the alcohol leaving group/nucleophile lies in the alcohol binding pocket below the triad (Gotor-Fernandez, Busto et al. 2006).

One structural feature that distinguishes lipase from esterase is the presence of a lid that covers the active site, along with a phenomenon called interfacial activation. Interfacial activation occurs when the conformation of the lid rearranges from a closed to an open state at the lipid-water interface, which results in a burst of substrate turnover (Brzozowski, Derewenda et al. 1991). However, exceptions exist with lipases either having a lid but not exhibiting interfacial activation or showing interfacial activation but only with limited substrates (Verger 1997). In the crystal structure of CALB, a short helix

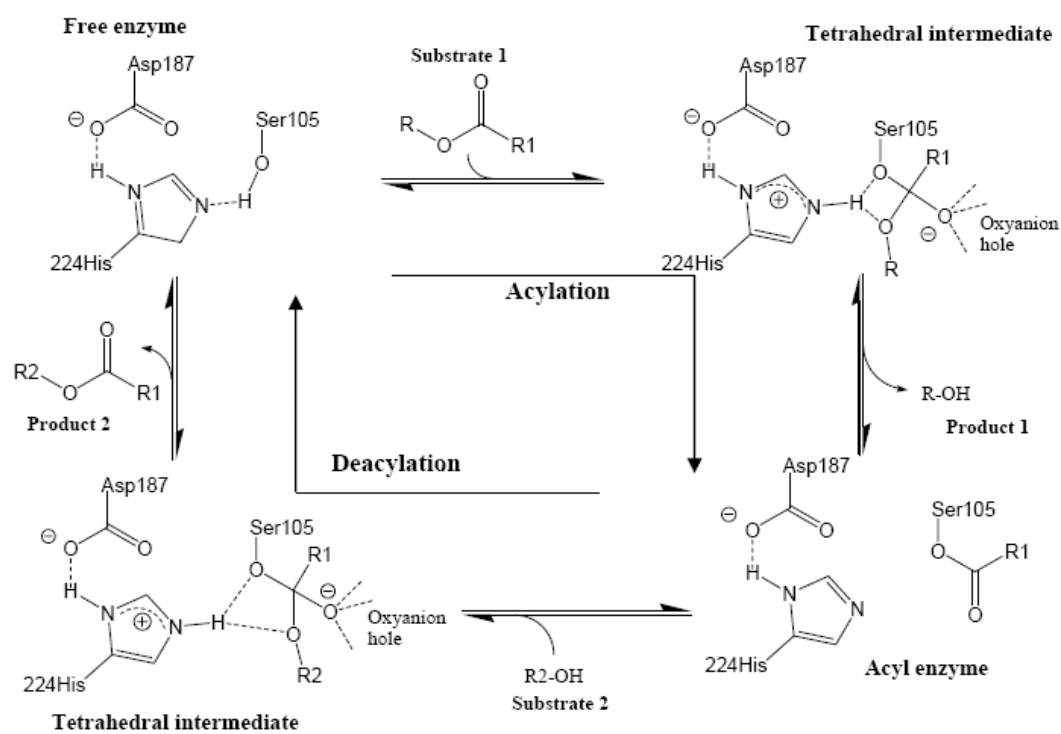


Figure 1-5 Reaction mechanism of CALB (Rotticci 2000; Rotticci-Mulder 2003)

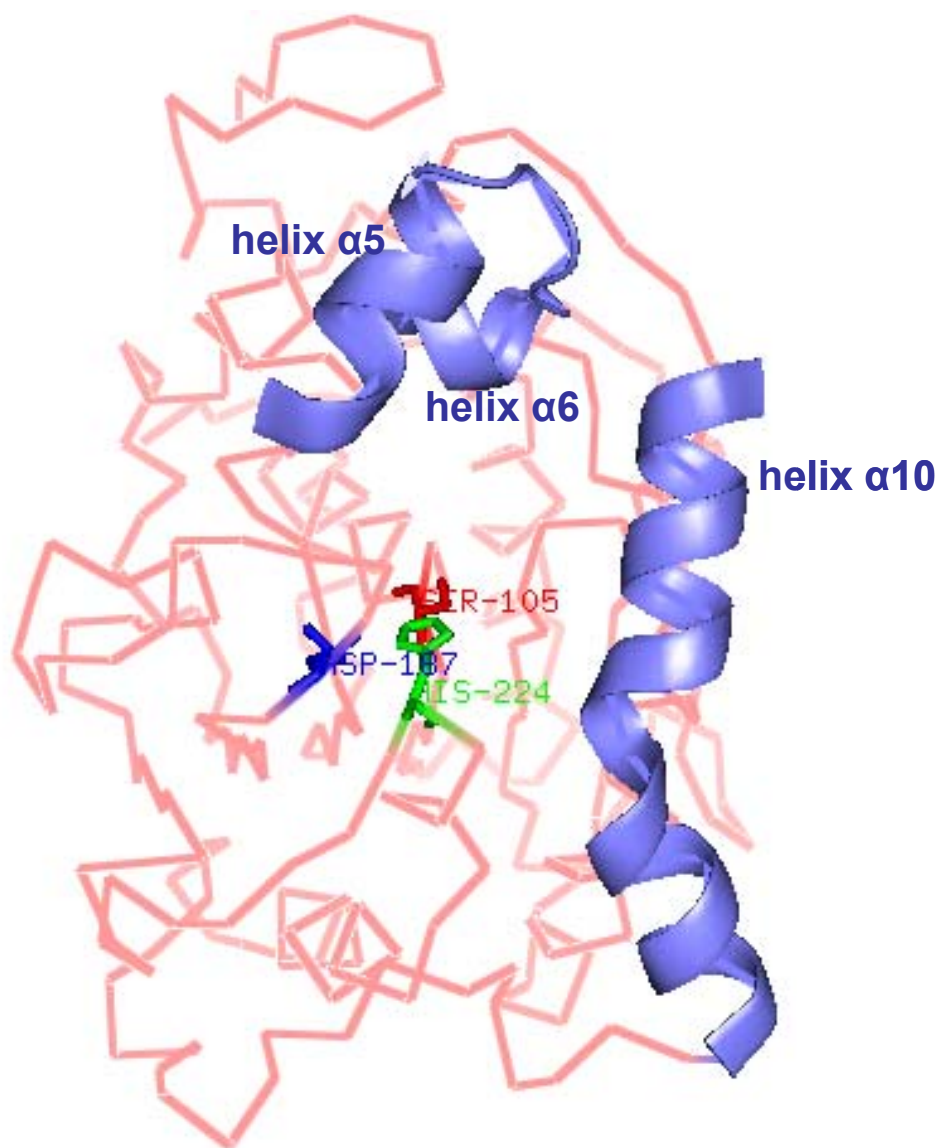


Figure 1-6 The active site pocket of CALB. Helices 5, 6, and 10 are shown in blue. The catalytic triad is shown in sticks.

- helix 5 (residue 142-146, GPLDA) shows high mobility. As helix 5 forms part of the active site pocket, its high mobility seems to suggest that it might be the lid of CALB that controls the access of substrates to the active site. However, no interfacial activation was observed for CALB catalysis. As suggested by the Hult group, this helix might instead function as part of the lipid anchoring surface and helps in the binding of the lipid (Martinelle, Holmquist et al. 1995).

1.2.2 Thermostability, substrate specificity, and enantioselectivity of CALB

CALB is a robust enzyme in view of its high stability and solvent tolerability (Anderson, Larsson et al. 1998). The thermal denaturation temperature of CALB is between 50 °C and 60 °C depending on the pH of the buffer used (Anderson, Larsson et al. 1998; Kirk and Christensen 2002). Although its optimum pH for catalysis is 7, CALB is stable in aqueous media with the pH ranging from 3.5 to 9.5. In addition, CALB retains its activity in organic solvents, and its stability can be further increased by immobilization onto solid supports. Once immobilized, CALB is highly thermostable and can be used continuously at elevated temperatures in organic solvents over extended period of time without any significant loss of activity (Anderson, Larsson et al. 1998; Kirk and Christensen 2002).

Although the physiological function of a lipase is the hydrolysis of triglycerides, it can usually accept a broad range of substrates dramatically different from the ones it is associated with in nature (Theil 1995). The substrate specificity of CALB in correlation to its structural features has been studied extensively. According to its catalytic

mechanism, the substrate specificity of CALB can be dissected into two parts, the acyl donor and the alcohol moiety. Previous study has shown that CALB has a rather broad tolerance for fatty acid length and structure (Anderson, Larsson et al. 1998). Compared to lipases from *Mucor miehei*, *Humicola*, and *Pseudomonas* that prefers fatty acids longer than octanoic acid, CALB showed better activity towards short-chain fatty acids, which is related to its relatively short fatty acid binding pocket (Kirk, Bjoerkling et al. 1992; Pleiss, Fischer et al. 1998).

For the alcohol moiety, CALB is best known for the resolution of secondary alcohols (Haeffner and Norin 1999). The enantio-preference of CALB toward secondary alcohols is consistent with the Kazlauskas rule, predicting that the enantioselectivity is higher when two substituents on the chiral carbon of secondary alcohols are different in size rather than being approximately the same size (Kazlauskas, Weissfloch et al. 1991). High enantioselectivity can be achieved when the small substituent is smaller than *n*-propyl and the large moiety larger than *n*-propyl (Rotticci, Haeffner et al. 1998). The size requirement is conveyed by the restricted volume of the alcohol binding pocket of CALB, which is only large enough to accommodate substituent smaller than propyl. Therefore, the faster enantiomer is able to place its small substituent into the alcohol binding pocket with the right orientation, while the slower enantiomer has to put its large substituent into the pocket, steric clashes prevent proper orientation of the substrate for effective catalysis (Figure 1-7).

The electrostatic interaction between the enzyme and the substrate is another factor that could contribute to the enantioselectivity of lipases. The crystal structure of *C.*

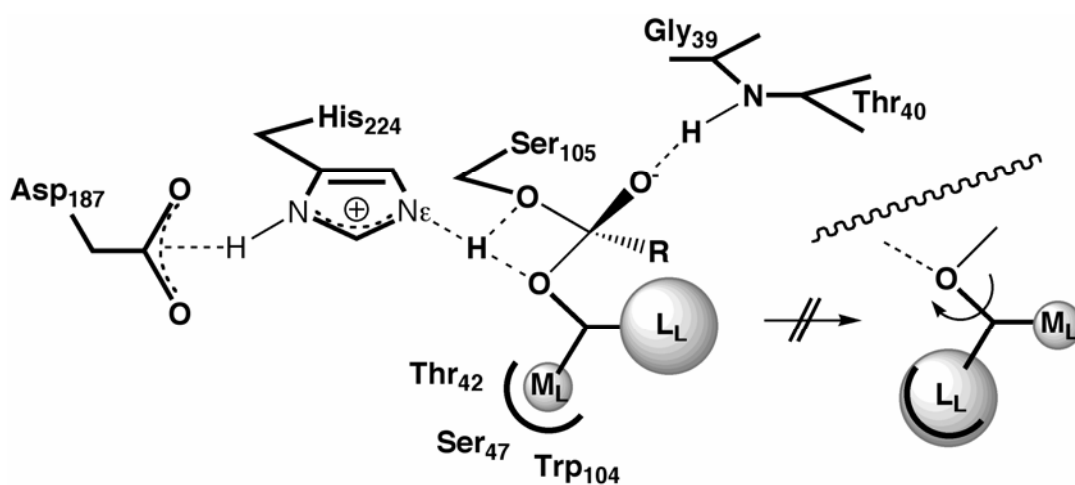


Figure 1-7 Model for predicting the fast enantiomer of secondary alcohols. L_L and M_L represent the large and small substituents on the chiral carbon respectively.

rugosa lipase in complex with the analogue to the fast reacting (*R*)-menthyl ester shows that the catalytic histidine forms a favorable hydrogen bond with the oxygen atom of the menthyl group, which is missing for the slow reacting enantiomer (Uppenberg, Ohrner et al. 1995). Although *C. rugosa* lipase and CALB share low sequence/structure identity, and the formation of a similar hydrogen bond in CALB has not been verified, previous studies have shown that electrostatic interaction does play a role on CALB's enantioselectivity. In these studies, a dramatic enantioselectivity difference was observed between a halogenated substrate and a substrate without halogen-substitution but of the same size. The differences in the electrostatic potential of the substrates and the electrostatic interaction between the electron-withdrawing halogen and the stereoselectivity pocket were believed to account for the observed change in enantioselectivity (Rotticci, Haeffner et al. 1998).

Overall, both the size limitation of the active site pocket and the electrostatic interactions between the enzyme and the substrate contribute to the enantioselectivity of CALB toward secondary alcohols.

1.2.3 Overexpression of CALB

A prerequisite to the industrial/academic applications of CALB is an efficient and reliable expression system. Different host systems have been tested for CALB expression, including *E. coli*, fungi and yeast. Although *E. coli* is the most popular and simplest host for recombinant protein expression, the yield of CALB expression in *E. coli* is very low (Rotticci-Mulder 2003; Blank, Morfill et al. 2006; Liu, Schmid et al. 2006), most probably due to the codon usage bias between eukaryotes and prokaryotes, and the lack of post-translational modification ability in *E. coli*. Successful host systems reported

include the fungi *Aspergillus oryzae* and the yeasts *Pichia pastoris* and *Saccharomyces cerevisiae* (Hoegh, Patkar et al. 1995; Rotticci-Mulder, Gustavsson et al. 2001; Zhang, Suen et al. 2003). The yeast *P. pastoris* expression system has received much attention in recent years in view of the rapid growth rate of yeast cells, ease of handling, and the capability of post-translational modifications. Many heterologous proteins, including CALB, have been successfully expressed in *P. pastoris* with high yield (Romanos 1995; Rotticci-Mulder, Gustavsson et al. 2001; Sherman 2002; Macauley-Patrick, Fazenda et al. 2005). In our study, the *P. pastoris* expression system was adopted for overexpression of all the CALB variants generated.

1.2.4 CALB engineering

Rational and random protein engineering approaches are currently commonly used for two overlapping goals, one is to expand the function or improve the properties of a protein, and the other is to test our fundamental understanding about protein folding and function (Poole and Ranganathan 2006). Lipases are attractive engineering targets mostly because of their potential applications in the production of enantiomerically pure compounds. As reported, the worldwide sales volume for chiral drugs is estimated to be US \$200 billion in 2006, and the overall need of the world market for chiral fine chemicals, pharmaceuticals, agrochemicals and flavor compounds is steadily increasing every year (Kaur and Sharma 2006). In view of the application value of CALB in asymmetric synthesis, especially after the availability of its crystal structure in 1994, a growing interest has been put on CALB engineering to further tailor its properties (Lutz 2004). The main focus of these studies was either to increase its thermostability (Zhang, Suen et al. 2003; Chodorge, Fourage et al. 2005), its activity or enantioselectivity toward

certain substrates (Ottosson, Rotticci-Mulder et al. 2001; Rotticci, Rotticci-Mulder et al. 2001; Suen, Zhang et al. 2004; Magnusson, Rotticci-Mulder et al. 2005; Magnusson, Takwa et al. 2005), or to introduce a new reaction mechanism into its scaffold (Carlqvist, Svedendahl et al. 2005; Svedendahl, Hult et al. 2005).

As shown in Table 1-2, both rational and random engineering approaches have been employed for CALB engineering, including site-directed mutagenesis, error-prone PCR and DNA shuffling.

The residues in CALB targeted by site-directed mutagenesis include T42, S47, M72, T103, W104, and S105. These residues are either directly involved in catalysis (S105) or positionally important in defining the active site pocket of CALB (T42, S47, M72 and W104). The size of the stereo- specificity pocket of CALB is confined by the backbone atoms and the side chains of T42, S47, and W104. Mutations S47A or T42V-S47A increased the enantioselectivity of the enzyme toward halohydrins by two-fold, but the specific activities of the mutants were lowered by two- to five-fold (Ottosson, Rotticci-Mulder et al. 2001; Rotticci, Rotticci-Mulder et al. 2001). Mutating W104 to alanine created a larger stereo- specificity pocket that was able to accommodate the phenyl group of 1-phenylethanol, which resulted in an inversed enantioselectivity toward (*S*)-enantiomer of the substrate (Magnusson, Rotticci-Mulder et al. 2005; Magnusson, Takwa et al. 2005). Mutating M72 to leucine increased the oxidation stability of the enzyme, verified by higher residual activity after exposure to peroxyoctanoic acid. However, the thermostability of the enzyme was significantly decreased (Patkar, Vind et al. 1998). In 2005, the Hult group showed that by mutating the nucleophile Ser105 to alanine, instead of the original hydrolytic activity, the enzyme could be tailored to

catalyze the Michael addition reactions of thiols to α,β -unsaturated carbonyl compounds (Carlqvist, Svedendahl et al. 2005; Svedendahl, Hult et al. 2005). All these examples suggested that the activity, function, and enantioselectivity of an enzyme can be optimized or modified by rational engineering effectively based on the knowledge about the mechanism and structure of the enzyme.

Other than structure-based rational design, directed evolution methods such as error-prone PCR and DNA shuffling were applied to CALB engineering. Several variants (V210I-A281E, V210I-A281E-V221D, N292Y) were identified from the error-prone PCR library with an up to 30-fold improved thermo-resistance (Zhang, Suen et al. 2003; Chodorge, Fourage et al. 2005). The mutation A281E was found to be the key residue responsible for the observed increase in thermo-resistance in the first two variants. The residue is part of a hydrophobic surface region, and mutating it to glutamate disrupted a possible nucleation site for aggregation and facilitated the refolding of the enzyme. The residue N317 is also located at the surface of the protein. It was reasoned that mutating it to tyrosine eliminated its potential to be involved in the deamidation process which would cause a serious disruption of the protein structure.

Instead of traditional directed evolution methods, circular permutation was employed in our study for CALB engineering. It was hypothesized that the internal relocation of a protein's N- and C-termini in or near the active site can increase chain flexibility and active site accessibility, which could translate into higher activity for structurally more demanding substrates.

Year	Engineering method	Purpose of engineering	Variants created	Properties of the variants	References
1997	Site-directed mutagenesis	To examine the role of T103 when mutated it back to a consensus G in other lipases	T103G	Increased thermostability, half- reduced activity, and unchanged specificity and enantioselectivity	(Patkar, Svendsen et al. 1997)
1998	Site-directed mutagenesis	To create more space for the active site	W104H	Reduced thermostability, activity, and enantioselectivity	(Patkar, Vind et al. 1998)
		To increase the oxidation stability of CALB	M72L	Increased stability towards oxidation, retained activity, and significantly reduced thermostability	
2001	Site-directed mutagenesis	To improve enantioselectivity toward halohydrins	S47A; T42V-S47A	Increased enantioselectivity toward halohydrins to around two fold, and reduced specific activity	(Ottosson, Rotticci-Mulder et al. 2001; Rotticci, Rotticci-Mulder et al. 2001)
2003	Error-prone PCR	To improve resistance towards irreversible thermal inactivation	V210I-A281E; V210I-A281E-V221D	A 20-fold increase in half-life at 70°C, and increased activity	(Zhang, Suen et al. 2003)
2004	DNA shuffling	To improve activity toward the hydrolysis of diethyl 3-(3',4'-dichlorophenyl)glutarate (DDG)	Chimeric protein	20-fold improved activity toward DDG, and increased thermostability	(Suen, Zhang et al. 2004)
2004	Site-directed mutagenesis	To catalyze Michael-type additions of thiols to α,β -unsaturated carbonyl compounds	S105A	Improved activity to Michael-type additions	(Carlqvist, Svedendahl et al. 2005; Svedendahl, Hult et al. 2005)
2005	Site-directed mutagenesis	To increase activity for large secondary alcohols	W104A	Increased activity for acylation of heptan-4-ol and nonan-5-ol, and inversed enantioselectivity to 1-phenylethanol	(Magnusson, Rotticci-Mulder et al. 2005; Magnusson, Takwa et al. 2005)
2005	Error-prone PCR	To improve the thermo-resistance	N292Y	Improved thermo-resistance at 90°C	(Chodorge, Fourage et al. 2005)

Table 1-2 Precedents of CALB engineering

1.3 Circular permutation of proteins

The concept of circularly permuted proteins is shown in Figure 1-8. The original termini of the parental protein are connected by a linker, and new termini are generated by breaking at another site of the polypeptide backbone (Figure 1-8).

1.3.1 Naturally occurring circular permutations in proteins

1.3.1.1 Mechanisms of circular permutation in nature

Naturally occurring circularly permuted proteins have been identified in various organisms, including viruses, bacteria, plants, and higher animals (Pan and Uhlenbeck 1993). They are speculated to derive either from posttranslational modification, gene duplication or from exon shuffling events (Figure 1-9) (Lindqvist and Schntider 1997; Uliel, Fliess et al. 2001; Weiner and Bornberg-Bauer 2006). The plant aspartic proteinase insert swaposin is a typical example of a circularly permuted protein derived from the gene duplication/deletion events (Russell 1995). Differently the first reported eukaryotic circularly permuted protein, concanavalin A, is the product of the post-translational transposition and ligation within initial polypeptide of the parental protein favin (Cunningham, Hemperly et al. 1979; Carrington, Auffret et al. 1985; Pan and Uhlenbeck 1993).

1.3.1.2 Functional importance of circular permutation in nature

Circular permutation of natural proteins may have functional importance. It was hypothesized that in the case of swaposin, the movement of the termini facilitates the insertion of the swaposin domain within the aspartic proteinase, taking advantage of the

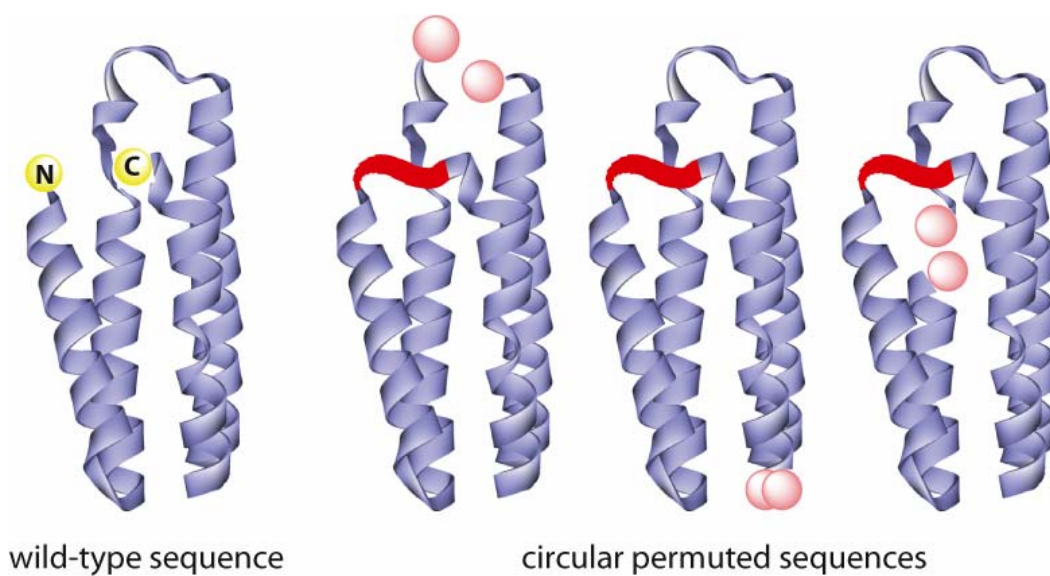


Figure 1-8 Schematic diagram of protein circular permutation. The positions of natural termini are shown in yellow; the linker used to connect the natural termini is shown in red ribbons; the positions of new termini are shown in pink.

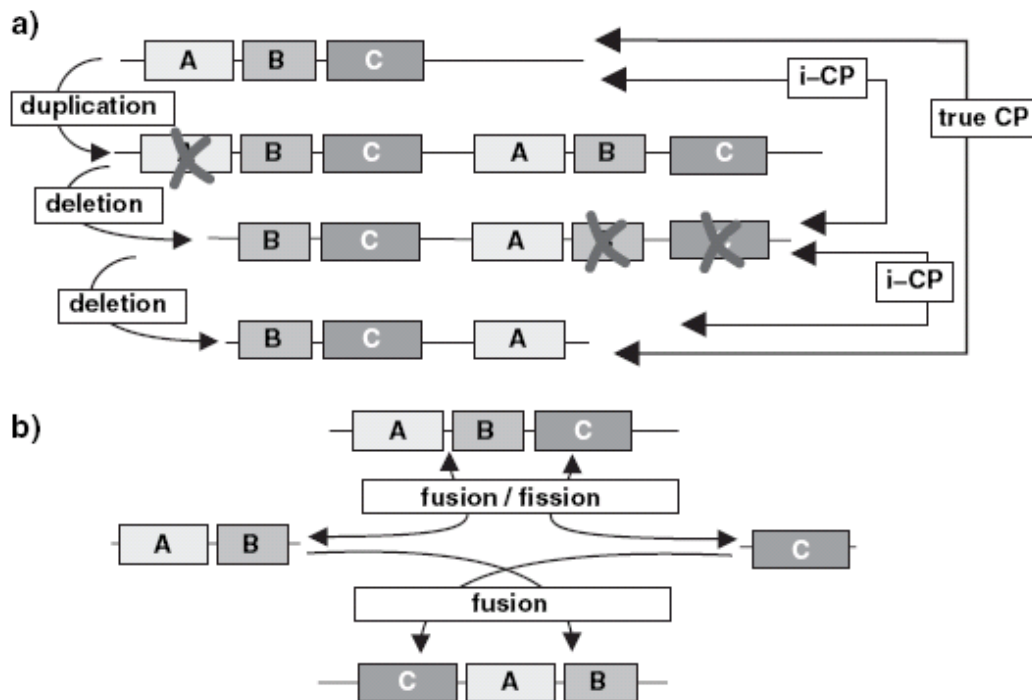


Figure 1-9 Schematic illustration of the mechanisms of circular permutation (a) gene duplication/deletion (b) exon shuffling or fusion/fission A, B, and C denote protein domains. X: domain deletion; i-CP: intermediate CP, a result of the duplication/deletion mechanism (Weiner and Bornberg-Bauer 2006).

orientation difference between swaposin and unpermuted saposin domain (Figure 1-10 (a)) (Russell 1995). For members of the aldolase superfamily that share a common TIM (triosephosphate isomerase) barrel fold containing eight $\beta\alpha$ motifs assembled in a circular arrangement, the structural character may assist the occurrence of circular permutation and results in a shift of the active site locations in sequence. It is proposed that this active site flexibility facilitates further adaptation of the protein for new functions, which accounts in part for the functional diversity of the TIM barrel proteins (Figure 1-10 (b)) (Todd, Orengo et al. 2002).

1.3.1.3 Circularly permuted α/β hydrolase-fold protein in nature

The only example of naturally occurring circularly permuted α/β hydrolase-fold protein is the polyhydroxybutyrate (PHB) depolymerase from *Penicillium funiculosum*. Different from the canonical α/β hydrolase-fold protein which contains a continuous polypeptide linkage for the central β -sheets from one edge to the other edge (Figure 1-11 (b)), the PHB depolymerase has its termini located in the middle of the central β -sheets (Figure 1-11 (a)) (Hisano, Kasuya et al. 2006). It was hypothesized that the enzyme might evolve from the canonical α/β hydrolase fold through gene duplication/deletion events. This presents an excellent example showing that discontinuity in the central β -sheets doesn't disrupt the overall α/β hydrolase fold and the function of the enzyme. And the fact that circular permutation is a strategy used by nature on α/β hydrolase-fold proteins provides us strong support on applying the same strategy artificially on CALB.

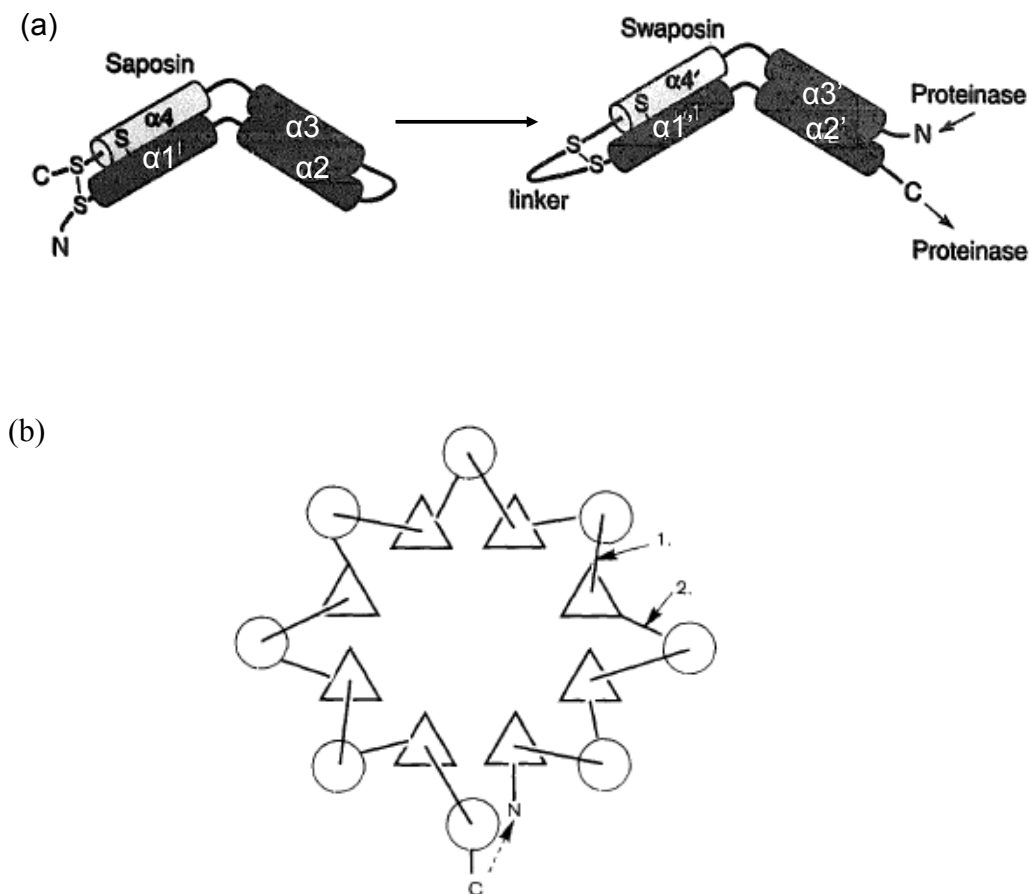
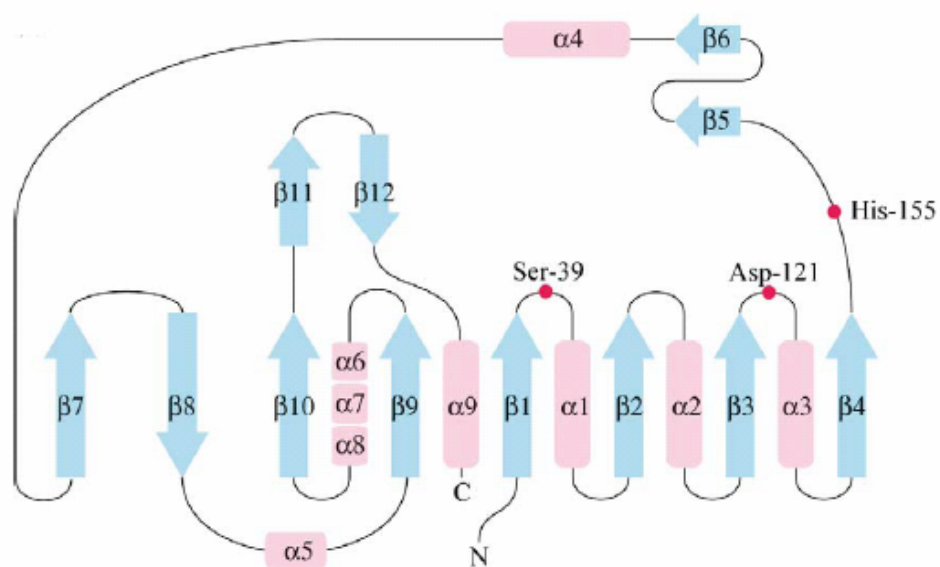


Figure 1-10 Examples of circularly permuted proteins (a) A hypothetical saposin fold and analogous swaposin fold (Russell 1995) (b) Topology diagram of the TIM-barrel fold. β strands are symbolized by triangles, α helices by spheres. Arrow 1 shows where the new termini are formed in the circularly permuted variant of glucosyltransferase with respect to α -amylase. Arrow 2 shows where the new termini are formed in transaldolase with respect to class I aldolases. The dotted arrow shows the new connection formed in both enzymes (Lindqvist and Schtider 1997).

(a)



(b)

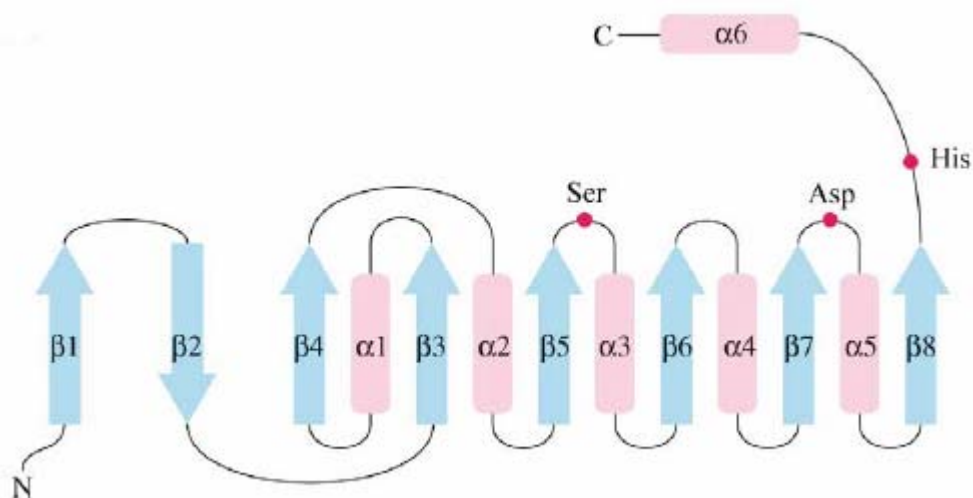


Figure 1-11 Topology diagram of the PHB depolymerase (a) in comparison to the canonical α/β hydrolase fold (b). Boxes and arrows represent α -helices and β -strands, respectively. Positions of the catalytic residues are indicated as red circles (Hisano, Kasuya et al. 2006).

1.3.2 Circular permutation as a genetic engineering approach

1.3.2.1 Rational design of circularly permuted proteins

Circular permutation was first performed artificially on bovine pancreatic trypsin inhibitor by Goldenberg and Creighton in 1983 through chemical condensation (Goldenberg and Creighton 1983). In 1989, the Kirschner group pioneered the use of genetic engineering in designing circularly permuted phosphoribosyl anthranilate isomerase (Luger, Hommel et al. 1989). Since then, a number of proteins have been studied by circular permutation, including T4 lysozyme (Zhang, Bertelsen et al. 1993; Sagermann, Baase et al. 2004), T1 ribonuclease (Mullins, Wesseling et al. 1994; Johnson and Raushel 1996), DHFR (Protasova, Kireeva et al. 1994; Uversky, Kutysenko et al. 1996; Nakamura and Iwakura 1999; Iwakura, Nakamura et al. 2000; Smith and Matthews 2001; Arai, Maki et al. 2003; Svensson, Zitzewitz et al. 2006), green fluorescent protein (Topell, Hennecke et al. 1999), beta-lactamase (Osuna, Perez-Blancas et al. 2002; Gebhard, Risso et al. 2006), aspartate transcarbamoylase (Zhang and Schachman 1996), beta-glucanase (Hahn, Piotukh et al. 1994; Ay, Hahn et al. 1998), 5-aminolevulinate synthase (Cheltsov, Guida et al. 2003), cyanovirin-N (Barrientos, Louis et al. 2002), myoglobin (Ribeiro and Ramos 2005), and streptavidin (Chu, Freitag et al. 1998).

The study on artificial circularly permuted proteins may afford valuable information about the importance of natural termini for the folding process and maintenance of tertiary structure or the biological function. It could also help in elucidating critical structures for catalysis, as a break in the peptide backbone that is functionally important presumably would significantly reduce the catalytic efficiency of the enzyme (Pan and Uhlenbeck 1993). For example, the Iwakura group employed

circular permutation in mapping the functional elements of the active site M20 loop of *E. coli* dihydrofolate reductase (DHFR). They found that one region inside the loop (L8~M16) is crucial for the folding of the protein, while two others (M16~M20, L24~D27) are vital to the catalytic efficiency of the enzyme (Nakamura and Iwakura 1999). These regions were found to be involved in substrate binding, which is in good accordance with the circular permutation results (Nakamura and Iwakura 1999). In another example, by studying a circularly permuted variant of human immunodeficiency virus (HIV)-inactivating protein cyanovirin-N (CV-N), the Gronenborn group found that although the circularly permuted variant adopts the identical fold as the wild type enzyme, its activity is dramatically lower (Barrientos, Louis et al. 2002). This finding together with subsequent mutagenesis and ^1H - ^{15}N HSQC studies revealed the importance of the original C-terminus for the anti-HIV activity of the enzyme, which provides valuable information for further design of virucides (Barrientos, Louis et al. 2002).

Besides its applications in identifying functional elements, circular permutation is often used in studying the folding pathways of proteins, as it alters the connectivity between secondary structures and thus may result in a change in the folding transition state, folding nucleus, and folding speed of the protein (Clementi, Jennings et al. 2001; Li and Shakhnovich 2001; Miller, Fischer et al. 2002; Weikl and Dill 2003; Bulaj, Koehn et al. 2004; Lindberg, Haglund et al. 2006). A computational study showed that an incision at the folding nucleus of the wild-type protein model changes the position of folding nucleus of the permuted variant dramatically and lowers its folding speed, while cutting at originally unstructured sites retains the original folding nucleus. Both changes in

entropy and in topology may contribute to the folding variation introduced by circular permutation (Li and Shakhnovich 2001).

Two other noteworthy examples that illustrate the benefit of circular permutation as a genetic engineering approach include its applications in creating the functional interleukin4-toxin fusion protein, and in the crystallization of the signal recognition particle receptor β subunit (Kreitman, Puri et al. 1994; Schwartz, Walczak et al. 2004). The Kreitman group showed that direct fusion of the natural interleukin 4 (IL4) with *Pseudomonas* exotoxin reduced its activity to 1% that of the free protein. The creation of a circularly permuted IL4-toxin fusion recovered the IL4 activity to 10% that of the free protein, significantly higher than the results obtained by the employment of various linkers to connect the natural IL4 and toxin (Kreitman, Puri et al. 1994). In the other example, although the crystal structure of the signal recognition particle receptor β subunit (SR β) in complex with the interaction domain SRX of the α subunit of SR was solved, the SR β alone failed the crystallization attempts. The Schwartz group speculated that an internal highly flexible loop of around 30 amino acids might be causing the problems with crystallization. By creation of a circularly permuted SR β variant by connecting the natural termini with a heptapeptide linker and removing the whole flexible loop, the permuted variant is functional and readily crystallized (Schwartz, Walczak et al. 2004). Both examples provide convincing evidence of the value of circular permutation, and similar approaches should also be applicable to the study on other proteins.

1.3.2.2 Random circular permutation

Compared to the rational design of circular permutants, random circular permutation stands for a more comprehensive approach in studying the tolerance of

shifting protein termini throughout the whole protein sequence, and in identifying central elements for maintaining the protein assembly, stability, and function.

Experimentally, random circular permutation can be achieved at the gene level by connecting the natural termini with a proper linker and digesting with *DNaseI* to generate a pool of circularly permuted genes, then the transformation and expression of the gene pool yields a library of circularly permuted proteins (Figure 1-12) (Graf and Schachman 1996; Ostermeier and Benkovic 2001).

The strategy has been applied to proteins including aspartate transcarbamoylase (ATCase) (Graf and Schachman 1996; Beernink, Yang et al. 2001; Ni and Schachman 2001), 5-aminolevulinate synthase (Cheltsov, Barber et al. 2001), disulfide oxidoreductase DsbA (Hennecke, Sebbel et al. 1999), and green fluorescent protein (GFP) (Baird, Zacharias et al. 1999). These studies suggest that the extent to which a protein can tolerate the change in chain connectivity and disruption of regular secondary structures is case specific. 51 out of 189 positions of DsbA can tolerate circular permutation and give correctly folded and functional variants, while the ratios for ATCase (~20 out of 316), 5-aminolevulinate synthase (21 out of 509) and GFP (~10 out of 238) are much lower (Graf and Schachman 1996; Baird, Zacharias et al. 1999; Hennecke, Sebbel et al. 1999; Beernink, Yang et al. 2001; Cheltsov, Barber et al. 2001; Ni and Schachman 2001). In all the cases, the new termini of functional variants are clustered, indicating certain regions of a protein are more tolerant to disruptions than others. As expected, structurally less-defined positions such as loops are more tolerant of new termini. However, these studies suggested that a protein can still withstand disruption of certain α -helices and β -sheets, while those intact secondary structures might

be essential for the proper folding or function of the enzyme (Graf and Schachman 1996; Baird, Zacharias et al. 1999; Hennecke, Sebbel et al. 1999; Beernink, Yang et al. 2001; Cheltsov, Barber et al. 2001; Ni and Schachman 2001).

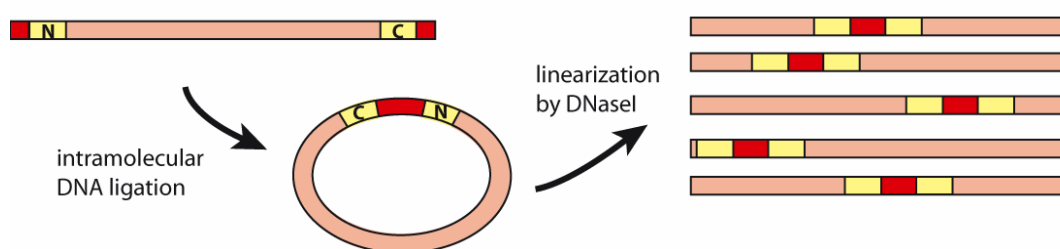


Figure 1-12 Creation of a random circular permutation library. The gene is circularized by intramolecular DNA ligation and relinearized by *DNase I* digestion to generate a library of circularly permuted genes.

Chapter 2 Improving the catalytic activity of *Candida antarctica* lipase B by circular permutation

2.1 Introduction

The broad substrate specificity and high enantioselectivity of CALB renders it a valuable biocatalyst in industrial and synthetic transformations. However, the performance of wild type CALB may not fulfill all the requirements for desired reactions. For example, its stability may be too low in a specific solvent at the temperature required by the reaction, its activity and enantioselectivity may not be ideal for the substrate in the reaction, or it does not possess the catalytic ability for certain substrates like tertiary alcohols. Different approaches have been employed to optimize the performance of CALB. For instance, new immobilization materials and methods were designed to stabilize the enzyme (Siddiqui and Cavicchioli 2005; Mateo, Palomo et al. 2006), and various protein engineering techniques were used to modify its substrate specificity and optimize its activity. In the latter case, rational protein design, random mutagenesis, and DNA shuffling have generated CALB variants with altered specificity, selectivity, and stability (Rotticci, Rotticci-Mulder et al. 2001; Lutz 2004; Suen, Zhang et al. 2004).

In our study, rather than substituting amino acids, circular permutation was employed to explore the effects of termini relocation on CALB's catalytic performance. Circular permutation as a genetic engineering approach is usually applied in studying the functional elements in a protein structure or protein folding, but little attention has been

paid to its effect on an enzyme's catalytic performance. We hypothesized that the internal relocation of a protein's N- and C-termini in or near the active site can increase chain flexibility and active site accessibility, which could translate into higher activity for structurally more demanding substrates. Based on the hypothesis, a random circularly permuted CALB library was generated and functional variants were identified via library screening. Selected library members were overexpressed and their kinetic properties were studied on multiple substrates. Our data demonstrate that circular permutation of CALB results in numerous functional variants with improved catalytic activity.

2.2 Materials and methods

2.2.1 Construction of wt-CALB expression vector

The wild type CALB (wt-CALB) gene was isolated from *Candida antarctica* (ATCC strain # 32657) by a two-step PCR amplification using the primers ZQ_CALBfor1 (5'-GAGGCTGAAGCTCATCATCATCATCATAGCAGCGGCCTTGTTCCACGTCTACCTTCCGGTTCGGACCCT-3'), ZQ_CALBfor2 (5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCATCATCATCATCAT-3'), and ZQ_CALBrev (5'-CGCGCGGCCGCTTAGGGGGTGACGATGCCGGAGCA-3'). The amplified gene included a (His)₆ tag followed by a thrombin cleavage site at the N-terminus of the lipase gene. Restriction enzyme recognition sites *Xho*I and *Not*I were also introduced into the 5' and the 3' ends respectively (recognition sequence underlined). The PCR product was digested with *Xho*I and *Not*I and ligated to the vector pPIC9 digested with the same restriction enzymes. This construct (pPIC9-CALB) brought the CALB gene under the

control of the methanol inducible alcohol oxidase promoter (AOX1) and in frame with the α -factor secretion signal peptide of *Saccharomyces cerevisiae*.

2.2.2 Random circulation permutation of CALB gene

The wt-CALB was amplified by PCR using primers ZQ_cpCALB_for (5'-GGTACTAGTGGTGGCCTACCTTCCGGTTCGGACCCT-3') and ZQ_cpCALB_rev (5'-CGCACTAGTACCGCCGGGGGTGACGATGCCGGAGCA-3') harboring a *SpeI* site at both ends (underlined). After digestion with *SpeI*, 5 μ g PCR fragment was circularized at a concentration of 2.5 ng/ μ l with 90 Weiss units T4 DNA ligase overnight at 16 °C. This construct generated a circular CALB with an 18-bp linker that encodes Gly-Gly-Thr-Ser-Gly-Gly joining the natural N- and C-terminals. After ethanol precipitation, the DNA was subjected to *ExonucleaseIII* (120 units/ μ g DNA) digestion at 37 °C for 30 min to remove remaining linear DNA. The *ExonucleaseIII* was inactivated by heating at 72 °C for 20 min. The DNA was purified by QIAquick columns and eluted with 50 μ l EB buffer.

Random relinearization of the circularized gene was performed by limited digestion with *DNaseI* (*RNaseI*-free; 0.5 milliunits/ μ g DNA) in 50 mM Tris-HCl, pH 7.5, 1 mM MnCl₂, DNA (5 μ g/ml) at room temperature for 15 min. The reaction was stopped by adding 10 μ l 0.5 M EDTA, and desalted by QIAquick columns (Qiagen, Valencia, CA) into elution buffer (10 mM Tris-HCl, pH 8.5). The linearized DNA was repaired using T4 DNA polymerase (1 unit/ μ g DNA) and T4 ligase (2 Weiss units/ μ g DNA) at room temperature for 1 h in T4 ligase buffer with the addition of dNTPs to a final concentration of 150 μ M. The linearized and cured DNA was recovered by agarose gel electrophoresis.

2.2.3 Creation of the pPIC9-cp-CALB library

2.2.3.1 Direct library cloning

The original library cloning strategy was to directly incorporate the circularly permuted CALB (cp-CALB) gene into the expression vector pPIC9 by blunt-end ligation (Figure 2-1). For that purpose the vector pPIC9 was prepared blunt-ended in a few steps. Firstly, a unique *SapI* site in pPIC9 was removed by *NdeI/PciI* double digestion, and the overhangs generated by digestion were filled by Klenow polymerase and the vector was intramolecularly ligated back to the circular form. The vector was then made blunt-ended at desired position for cp-CALB gene insertion, which was achieved by using the *Drosophila melanogaster* kinase gene DmdNK. DmdNK was PCR amplified with primers ZQ_CALB_for2, ZQ_CALB_for3 (GAGGCTGAAGCTCATCATCATCATCATCATAGCAGCGGCCTTGTTCACGTCGAAGAGCGCCGAGGGCACCCAGCCC), and ZQ_CALB_rev2 (CGCGCGGCCGCTCAAGAAGAGCCACCAGAACCGGCGAGGGCTG), which incorporated *XhoI*, *NotI*, and *SapI* sites at the ends of the gene (underlined). The PCR product was then digested with *XhoI* and *NotI* and ligated into the modified pPIC9 vector digested with the same restriction enzymes. The following *SapI* digestion and Klenow polymerase treatment generated blunt-ended vector with the α -secretion factor sequence and a (His)₆ tag at one end and a stop codon at the other end, which was ready for the incorporation of cp-CALB gene.

However, the direct blunt-end ligation of cp-CALB into the expression vector turned out to be problematic due to the following several reasons: (1) The size of the vector is large (8 kb), which increased the difficulty of manipulation. (2) The restriction enzyme *SapI* is inefficient; undigested and single-digested vector will contaminate the

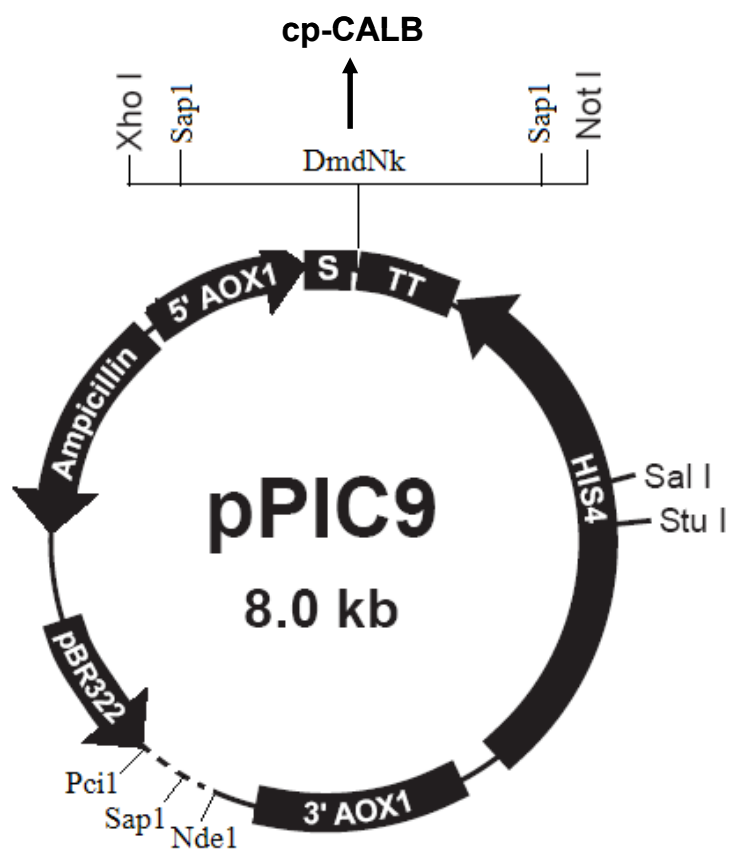


Fig 2-1 Schematic overview of the direct library cloning

ligation reaction with the cp-CALB gene. (3) Klenow polymerase was used to fill in the overhangs of *SapI* digestion. Due to the 3' → 5' exonuclease activity of the Klenow polymerase, the fill-in step is challenging. The remaining staggered end of the vector limited the size of the library that can be obtained in the following ligation reaction with cp-CALB gene.

2.2.3.2 Two-step library cloning

Successful library integration was instead achieved by a two-step library cloning strategy using pAMB-CAT as a shuttle vector (Figure 2-2). In preparation for library cloning, pAMB-CAT was modified to carry the N-terminal extensions (His-tag, Thrombin cleavage site) upstream from the CALB cloning site plus a stop codon immediately following the site of insertion. Therefore, PCR-amplified wild type CALB (primers: ZQ_CALBfor1, ZQ_CALBfor2, ZQ_CALBrev) was digested with *NotI/XhoI* and ligated to the vector pAMB-CAT digested with the same restriction enzymes. The resulting vector pAMB-CALB was amplified using primers ZQ_pAMBfor (5'-CCGGATATCAGGCCT TGGAACAAGGCCGCTGCTATG-3') and ZQ_pAMBrev (5'-CCGGATATCTTATAA GCGGCCGCAAGCTTGTCG-3'), which harbored a *StuI* and a *PsiI* site (underlined) as well as *EcoRV* sites (dashed lines) flanking both ends. The amplified vector was digested with *EcoRV* and ligated with a segment generated from *EcoRV* digestion of pET-16b vector to increase the size of the insert. This enabled subsequent digests to be monitored. Finally, the vector was digested with *StuI* and *PsiI* and the cp-CALB library was incorporated into the vector by blunt-end ligation. Transformation of the plasmid into electro-competent *E. coli* DH5 α -E cells generated the pAMB-cp-CALB library (~5 x 10⁵ members). The colonies were harvested and the

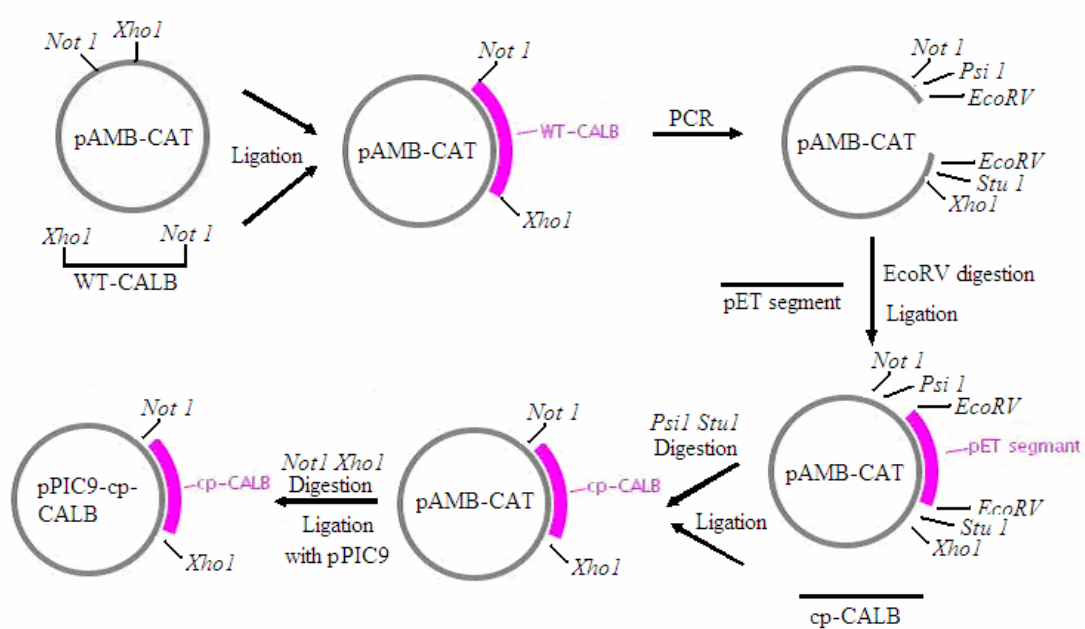


Figure 2-2 Schematic overview of the two-step library cloning

plasmid was isolated by QIAprep Spin Miniprep kit. In a second cloning step, the cp-CALB library was integrated in pPIC9. Purified pAMB-cp-CALB was digested with *NotI/XhoI* and the segment containing the cp-CALB library was ligated to the pPIC9 vector digested with the same enzymes. Approximately 1.5×10^6 colonies were obtained after transformation into electro-competent *E. coli* DH5 α -E cells. The transformants were harvested and the plasmid was isolated using the QIAprep Spin Miniprep kit.

2.2.4 Creation of the (His)₆-tag free pPIC9-cp-CALB library

Addressing concerns that the His-tag may interfere with the enzyme function, a second library without the affinity tag was created. Again the two-step library cloning strategy was used. The vector pAMB-CALB (created in the first library) was PCR amplified with primers ZQ_pAMBfor2 (5'-CCGGATATCGCCGGCTTCAGCCTCTCTTTTCTCGAG-3') and ZQ_pAMBrev (5'-CCGGATATCTTATAAGCGGCCGCAAGCTTGTCG-3'), which harbored a *NaeI* and a *PsiI* site (underlined) as well as *EcoRV* sites (dashed lines) flanking both ends. The amplified vector was digested with *EcoRV* and ligated with a segment generated from *EcoRV* digestion of pET-16b vector to increase the size of the insert. Then the vector was digested with *NaeI* and *PsiI* and the cp-CALB library was incorporated into the vector by blunt-end ligation. Transformation of the plasmid into electro-competent *E. coli* DH5 α -E cells generated the pAMB-cp-CALB2 library ($\sim 3.2 \times 10^4$ members). The colonies were harvested and the plasmid was isolated by QIAprep Spin Miniprep kit. Purified pAMB-cp-CALB2 was digested with *NotI/XhoI* and the segment containing the cp-CALB library was ligated to the pPIC9 vector digested with the same enzymes. Approximately 0.5×10^6 colonies were obtained

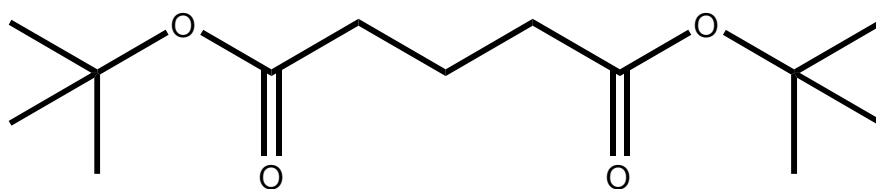
after transformation into electro-competent *E. coli* DH5 α -E cells. The transformants were harvested and the plasmid was isolated using the QIAprep Spin Miniprep kit.

2.2.5 Library screening

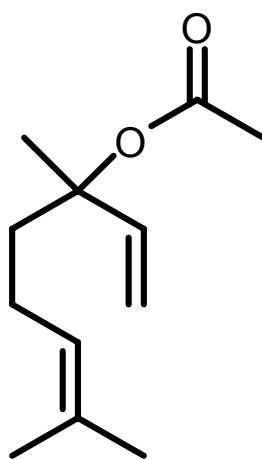
After digestion with *SacI* and ethanol precipitation, the pPIC9-cp-CALB libraries with/without the His-tag were transformed into electro-competent *P. pastoris* strain GS115 (Wu and Letchworth 2004) and plated on MM-tributylin plates. Colonies appeared after four days of incubation at 30 °C. Active cp-CALBs were identified by the formation of a clear halo surrounding the respective host colony (Gupta, Rathi et al. 2003). These colonies were picked and replated on MD and MM-trybutyrin plates to verify the lipase activity. After confirmation of the lipase activity, the sequences of the corresponding cp-CALB genes were obtained by colony PCR and DNA sequencing using primers ZQ-pPIC9-for (5'-TACTATTGCCAGCATTGCTGC-3') and ZQ-pPIC9-rev (5'-GCAAATGGCATTCT GACATCC-3').

The library was also screened for hydrolytic activity on tertiary-alcohol esters. Instead of tributyrin, a tertiary-alcohol ester glutaric acid *bis*-(*t*-butyl) ester (Figure 2-3) (synthesized by Christina Fields) was added to the plate media. Similar to the tributyrin plates, active cp-CALBs could be identified by the formation of a halo around the host colony. However, no hit came out of the library screening.

In addition, a plate assay for the screening of the hydrolytic activity toward linalyl acetate (Figure2-3) was tested. The assay is based on the fact that acetate released from the hydrolysis of linalyl acetate will lower the pH of the media around host colonies, which can be detected by the color change of a pH-indicator bromothymol blue (pH range 6.0-7.6) from blue to yellow. The assay was reported to be successful for the



glutaric acid bis-(t-butyl) ester



linalyl acetate

Figure 2-3 Structures of the tertiary alcohol esters used in library screening

screening of bacteria esterases (Schlacher, Stanzer et al. 1998). However, the growth of yeast cells is much slower than bacteria, and the materials secreted by yeast cells acidify the media over time. In combination with the shallow pH range of the indicator, the assay could not be adapted to our yeast expression system.

2.2.6 Protein expression, purification, and activity assays

The overexpression and purification of wild type CALB and selected circularly permuted variants was performed as described in Appendices B1~B3. The two-step purification by hydrophobic interaction chromatography (Rotticci-Mulder, Gustavsson et al. 2001) in combination with size exclusion chromatography enables the rapid isolation of lipase variants whose His-tag is not accessible (circular permutants with termini in the protein's interior region) or has been removed all together.

The activity of circularly permuted CALB variants on hydrolysis of DiFMU octanoate and p-NB was determined as described in Appendix B4.

2.3 Results and discussions

2.3.1 Creation of the cp-CALB library

Given the difficulty of identifying suitable permutation sites by rational design, a comprehensive, combinatorial library of randomly permuted CALB variants was generated. Starting with the wild-type CALB gene, a flanking oligonucleotide sequence which encodes for the flexible six-amino acid linker (-GGTSGG-) was introduced to bridge the ~ 17 Å distance between the original termini. After intramolecular ligation, the circular DNA was linearized in random positions using *DNaseI* (Graf and Schachman 1996; Baird, Zacharias et al. 1999; Beernink, Yang et al. 2001). Reaction conditions were

chosen such that, on average, only a single cut per DNA strand was introduced. The resulting library of CALB permutants was then cloned into pPIC9 and transformed into *Pichia pastoris* GS115 for protein expression. DNA sequence analysis of 96 randomly chosen members in the naive library ($\sim 5 \times 10^5$ colonies) confirmed the unbiased distribution of new termini over the entire length of the protein sequence (Figure 2-4).

2.3.2 Library screening and analysis

Functional variants in the CALB library were identified by colony screening on tributyrin plates (Figure 2-5). The DNA sequence analysis of functional members identified 63 unique protein sequences with termini in positions other than wild type (Figure 2-4). Addressing concerns that the His-tag may interfere with the enzyme function, a second selection experiment of the same CALB library without the affinity tag was performed. The DNA sequence analysis of functional candidates indicated that the location and distribution of permutation sites in functional CALB variants was the same as shown in Figure 2-4.

The data indicate that CALB tolerates permutations in numerous positions over the entire length of the protein. When mapped on the wild-type CALB structure, the new termini not only coincide with surface loops but also interrupt secondary structure elements on the enzyme's surface and interior regions (Figure 2-6). Most noticeable is the concentration of functional permutations from amino acids 243 to 317. This sequence is largely surface-exposed, wrapping around the front of the α/β -hydrolase core and forming the alcohol-binding portion of the active site pocket ($\alpha 10$) (Uppenberg, Hansen et al. 1994; Uppenberg, Ohrner et al. 1995). Two additional regions tolerant to permutation stand out: first, amino acids 44 and 47, which are located in close proximity to the

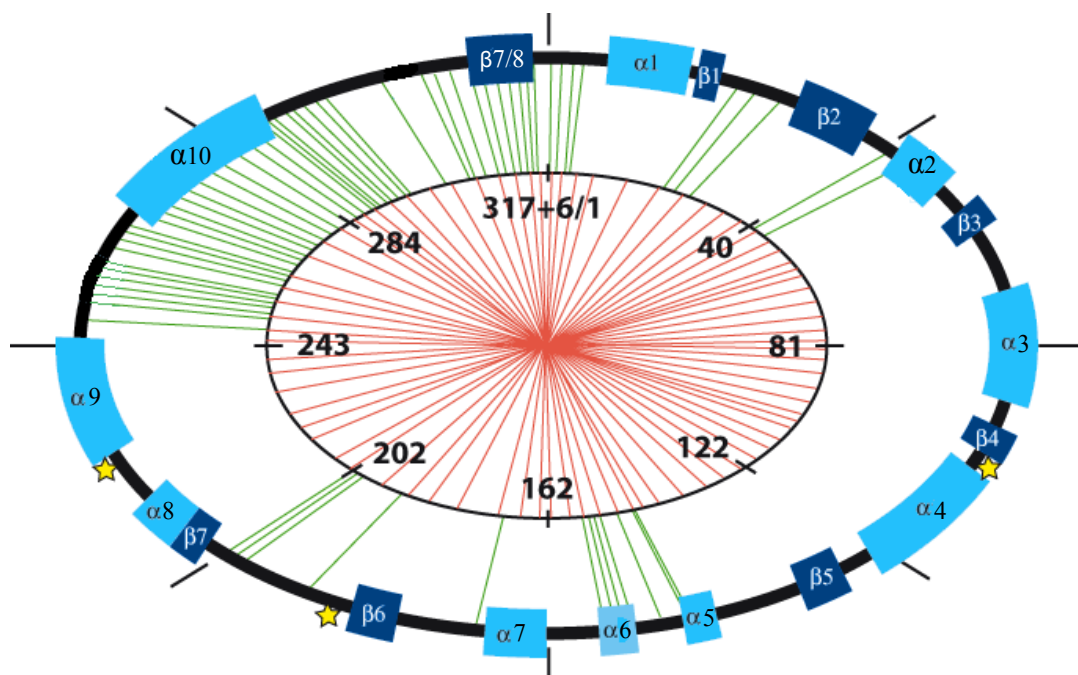


Figure 2-4 Circular permutation of CALB. Unbiased distribution of permutation sites over the entire length of CALB (317 amino acids) and the six amino acids linker in the native library (inner circle, red lines) and permutation sites of variants selected for tributyrin hydrolysis (outer circle, green lines). Secondary structures are shown in shades of blue; the three active site residues are marked by yellow stars.

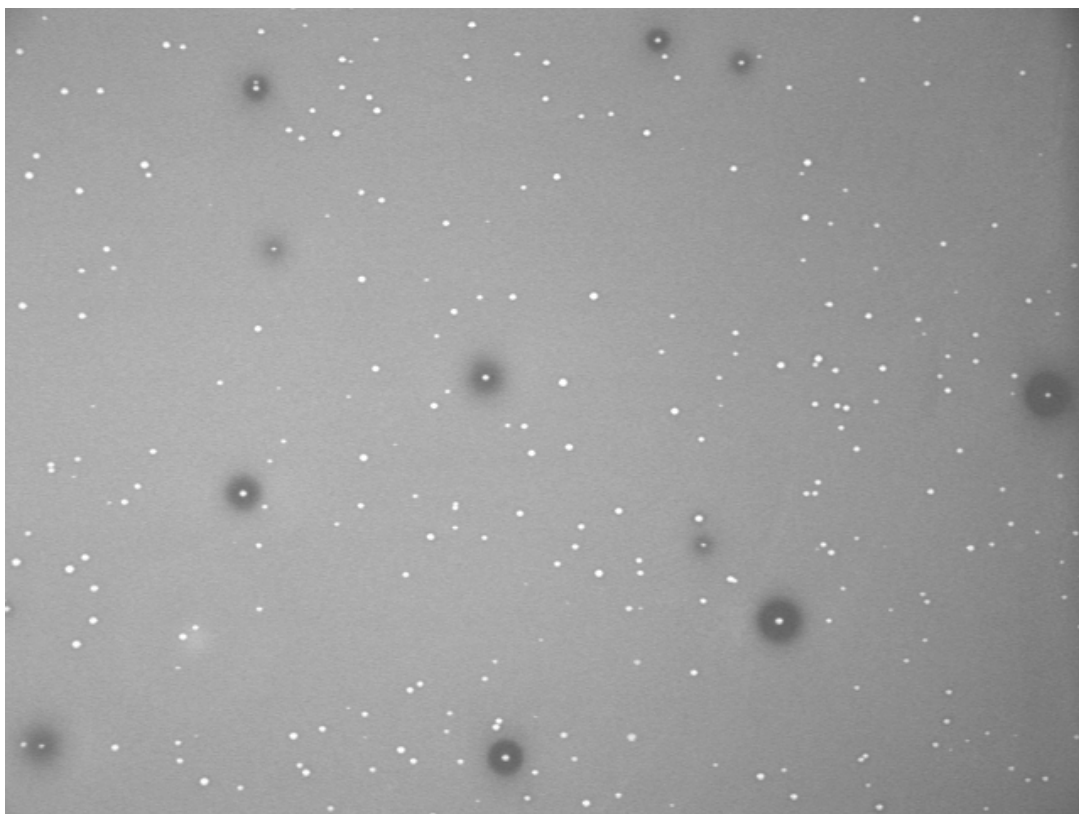


Figure 2-5 Library screening on tributyrin plates. Functional variants were indicated by halos around colonies.

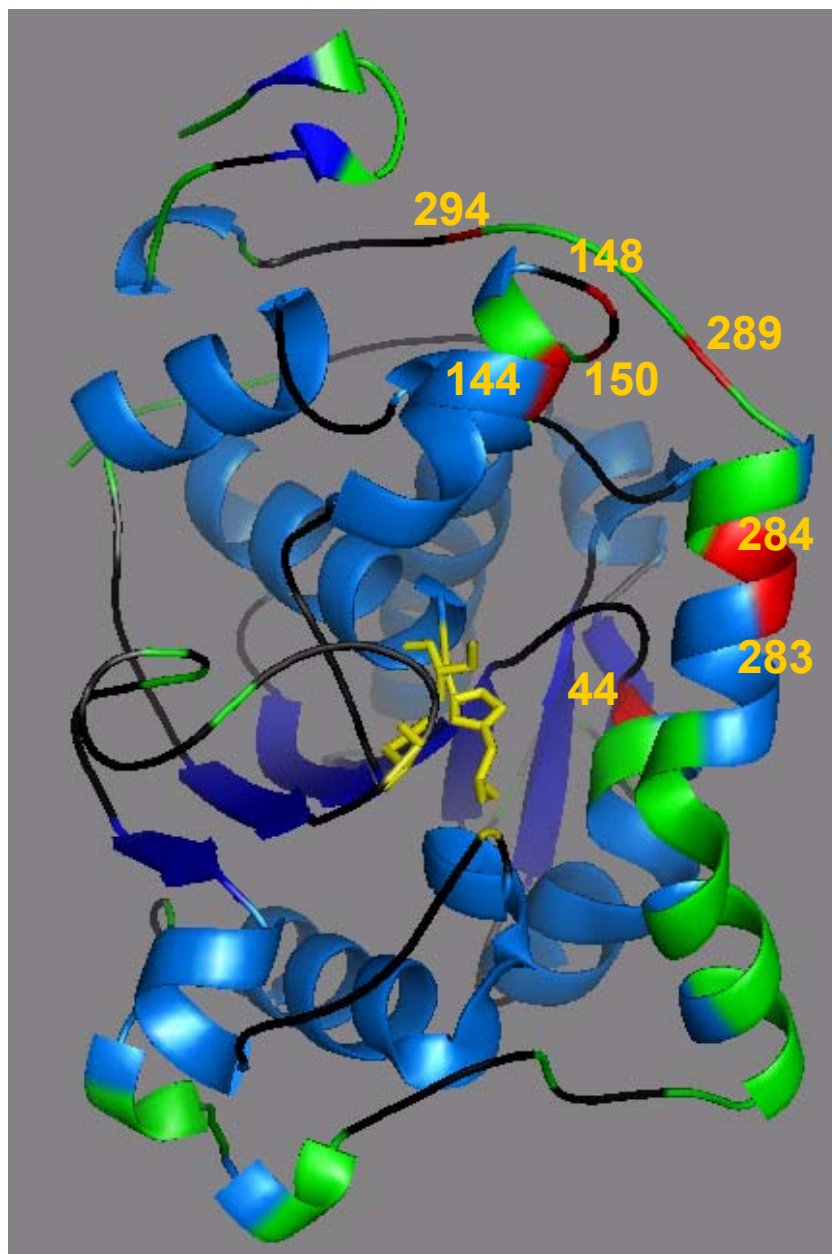


Figure 2-6 New termini mapped on wild-type CALB (1tca) (Uppenberg, Hansen et al. 1994). The locations of permissible permutation sites are indicated in green, and candidates selected for subsequent *in vitro* characterization are labeled in red. The residues of the catalytic triad are shown in yellow.

oxyanion stabilizing residues; second, a cluster of permutations in $\alpha 5/6$ (amino acids 135-155), which constitutes the enzyme's lid region. We also identified two protein segments (residues 48-143 and 204-246) with no functional permutation. These regions make up the core of the α/β -hydrolase fold and include residues S105 and H224 of the catalytic triad. We speculate that the absence of functional permutation near these residues, as well as the presence of only a single site proximal to the triad's third amino acid (D187), reflects this region's importance to catalysis and possibly its relevance to protein folding.

2.3.3 Selected library members for kinetic characterization

To examine the impact of circular permutation on catalysis, we selected eight functional CALB variants with termini in or near the active site for detailed kinetic characterization (Figure 2-6). Following overexpression in *P. pastoris*, the proteins were purified to homogeneity, and kinetic data for these variants were measured on two standard lipase substrates, p-nitrophenol butyrate (p-NB) and 6.8-difluoro-4-methylumbelliferyl (DiFMU) octanoate (Table 2-1).

The kinetic analysis confirmed that circular permutation has a significant impact on CALB's catalytic performance. The most substantial improvements in enzymatic activity over wild-type CALB were observed upon termini relocation into $\alpha 10$. Three of the four variants (cp283, cp284, cp289) show a consistent 10-fold improvement in their apparent k_{cat} values for p-NB and up to 60- fold increases in DiFMU octanoate turnover. The impact of backbone cleavage in $\alpha 10$ on active site accessibility and the rate-limiting deacylation step in wild-type CALB were investigated subsequently (Chapter 3-5). In contrast, a removal of the entire protein fragment (amino acids 284-293) in cp294 proves detrimental to catalysis. Whether the deletion dismantles the active site pocket,

preventing productive substrate binding, or affects protein stability as the disulfide bond-forming C293 is eliminated remains unclear.

The backbone cleavage in the lid region (cp144, cp148, cp150) showed only moderate effects on hydrolysis of test substrates. Both K_M and k_{cat} for all three variants stay within 2-fold of the parent enzyme under the described assay conditions. Structure models predict close interactions of this protein region with the substrate's acyl portion (Uppenberg, Hansen et al. 1994). Thus, future experiments may explore the turnover of bulkier carboxylates by these CALB variants. Furthermore, circular permutation of the lid region may alter the enzyme's response to changes in the reaction medium. The latter can affect lipase activity by modulating conformational changes in the lid region.

Finally, the kinetic data for cp44 show a 10 to 100-fold reduction in relative specificity, compared to that of wild-type CALB. The close proximity of the permutation site to the oxyanion binding pocket likely results in the topological misalignment of the active site residues. Consistent with our observation of permutation-free protein segments, we hypothesize that protein permutation does increase local backbone flexibility. While such flexibility seems detrimental at positions in proximity to active site residues, the relaxation effects can be beneficial when applied to protein regions which contribute to the active site topology but do not directly carry a side chain involved in catalysis.

Table 2-1 Apparent kinetic constants for selected CALB variants with p-nitrophenyl butyrate and DiFMU octanoate as substrates

enzyme variants		p-nitrophenyl butyrate				DiFMU octanoate			
Name ^a	Sequence ^b	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1}\text{min}^{-1}$)	relative specificity ^c	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1}\text{min}^{-1}$)	relative specificity ^c
Wild type	L1/P317	410 ± 40	305 ± 10	0.74	1.0	2.6 ± 0.3	2 ± 0.1	0.8	1.0
cp44	G44/T43	690 ± 90	6 ± 0.5	0.01	0.01	5.6 ± 0.8	0.5 ± 0.05	0.1	0.13
cp144	L144/A141	550 ± 50	178 ± 7	0.32	0.4	2.0 ± 0.5	1 ± 0.1	0.5	0.6
cp148	A148/L147	500 ± 30	171 ± 4	0.34	0.5	3.5 ± 0.5	1.5 ± 0.2	0.4	0.4
cp150	S150/V149	510 ± 90	520 ± 45	1.02	1.4	2.7 ± 0.8	2.1 ± 0.2	0.8	1.0
cp283	A283/A283-KRPRINSP	280 ± 50	2971 ± 180	10.61	14.3	2.5 ± 0.5	25 ± 1.4	10.9	13.6
cp284	A284/A287-KRPRINSP	550 ± 70	2980 ± 200	5.42	7.3	8.8 ± 1.0	34 ± 4	3.8	4.8
cp289	P289/A284-KRPRINSP	260 ± 30	3258 ± 215	12.53	16.9	5.5 ± 1.0	120 ± 7	23	28.8
cp294	E294/A283	310 ± 40	73 ± 4	0.23	0.3	9.5 ± 2.0	2.6 ± 0.3	0.3	0.4

Note: a. CALB nomenclature: cp44 = circular permuted protein whose N-terminus starts at amino acid 44 of the wild-type sequence. b. N- and C-terminal amino acids (all in single-letter code) are listed. Small variations in chain length of individual permutants are caused by reading frame shifts and staggered ends upon *DNaseI* digestion. c. Relative specificity = k_{cat}/K_M (variant) / k_{cat}/K_M (wild type).

2.4 Conclusions

In summary, CALB engineering by circular permutation has generated 63 new, unnatural lipase variants. Kinetic analysis confirmed that these protein variants can have sustained or improved catalytic function on multiple substrates over wild-type, mutant, and shuffled CALBs (Rotticci, Rotticci-Mulder et al. 2001; Zhang, Suen et al. 2003; Suen, Zhang et al. 2004). The observed rate enhancements are believed to result from improved active site accessibility and increased local protein backbone flexibility.

Beyond exploring the tolerance of CALB toward circular permutation and the activity of circularly permuted variants, studies are needed to address the substrate specificity and enantioselectivity of these lipase variants, and also the impact of permutations on CALB's structural integrity and stability on the molecular level, which will be discussed in the next chapter.

Chapter 3 Characterization of circularly permuted CALB variants

3.1 Introduction

Circular permutation as a genetic engineering approach has often been used in mapping the important structural elements of proteins and in studying protein folding, but little attention has been paid to its effects on the catalytic performance of enzymes. In Chapter 2 we have shown that, when performing random circular permutation on CALB, around 20% of the positions of whole peptide sequence can tolerate new termini, and the location of new termini is directly coupled to the catalytic performance of the enzyme. Relocation of the termini to helix $\alpha 10$ results in an increased activity for hydrolysis of test substrates p-NB and DiFMU octanoate, while the new termini in the putative lid region have only moderate effect on catalysis. The focus of this Chapter is to gain more understanding about the impact of permutations on CALB's structural integrity and stability at the molecular level, and to study how circular permutation impacts the substrate specificity and enantioselectivity of the variants.

Most studies on circularly permuted proteins, such as T4 lysozyme, ribonuclease T1, aspartate transcarbamoylase, myoglobin, and DsbA have shown that the variants maintain the overall native-like structures with only small variations, especially in regions around the joined original termini and the new termini (Zhang, Bertelsen et al. 1993; Mullins, Wesseling et al. 1994; Ni and Schachman 2001; Fishburn, Keeffe et al.

2002; Manjasetty, Hennecke et al. 2004; Sagermann, Baase et al. 2004). Although these structural variations are small, they usually lead to a decrease in thermostability (Zhang, Bertelsen et al. 1993; Mullins, Wesseling et al. 1994; Ni and Schachman 2001; Fishburn, Keeffe et al. 2002; Manjasetty, Hennecke et al. 2004; Sagermann, Baase et al. 2004). Here, circular dichroism (CD) analysis on selected CALB variants showed that although the variants maintain the overall structure as the wt-CALB, their thermostability is compromised to a different extent. Enantioselectivity analysis on the most promising variant cp283 and wt-CALB showed that the variant retains or exhibits higher enantioselectivity towards test substrates. Our results suggested that circular permutation does not alter the overall 3-D structure of CALB variants, instead increased local chain flexibility near new termini impacts catalysis.

3.2 Materials and methods

3.2.1 Construction of His-free protein expression vector

The N-terminal (His)₆-tag and the C-terminal extension (an artifact caused by reading frame shifts and staggered ends upon *DNaseI* digestion) on circularly permuted variants were removed by PCR amplification with primers listed in Table 3-1.

The PCR products were digested with *XhoI* and *NotI* and ligated to the vector pPIC9 digested with the same restriction enzymes. The vector was then linearized by *SacI* digestion and transformed into *P. pastoris* strain GS115.

Primer name	Primer sequence
pPIC9 FOR	5'-TACTATTGCCAGCATTGCTGC-3'
pPIC9 REV	5'-GCAAAT GGCATTCTGACATCC-3'
CP144 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCTCGATGCACTCGCGGTTAGT-3'
CP148 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTGCGGTTAGTGCACCCTCCGTA-3'
CP150 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTAGTGCACCCTCCGTATGGCAG-3'
CP283 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTGCAGCCATCGTGGCGGGTCCA-3'
CP284 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTGCCATCGTGGCGGGTCCAAAG-3'
CP289 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCCAAAGCAGAACTGCGAGCCC-3'
CP193 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCAGGTGTCCAACTCGCCACTC-3'
CP277 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCTCCTGGCGCCGGCGGCTGCA-3'
CP278 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCTGGCGCCGGCGGCTGCAGCC-3'
CP268 FOR	5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCCCGAGCAAAGGTCGCCGCG-3'
CP283 REV	5'-GCGGCGGCCGCTTAAGCCGCCGGCGCCAGGAG-3'
CP284 REV	5'-CGCGCGGCCGCTTATGCAGCCGCCGGCGCCAG-3'
CP289 REV	5'-CGCGCGGCCGCTTAACCCGCCACGATGGCTGC-3'
CP193 REV	5'-CGCGCGGCCGCTTAAGGCTGAACGATCTCGTCGG-3'
CP277 REV	5'-CGCGCGGCCGCTTACGCAGCCGCGGCGACCTT-3'
CP278 REV	5'-CGCGCGGCCGCTTAGAGCGCAGCCGCGGCGAC-3'
CP268 REV	5'-CGCGCGGCCGCTTAAGTCAGATCATTGGCGGG-3'

Table 3-1 Primers used for construction of His-tag free variants

3.2.2 Protein expression, purification, and activity assays

Protein expression and purification was performed as described in Appendices B1 and B3. The activity assays on p-NB and DiFMU octanoate hydrolysis for all the variants were performed as described in Appendix B4.

3.2.3 Circular dichroism analysis

Circular dichroism analysis was performed as described in Appendix B6.

3.2.4 *P. pastoris* fermentation

Both cp283 and wt-CALB were expressed in large quantities in *P. pastoris* by fermentation. The procedures were described in Appendix B5. After fermentation, the supernatant was collected and filtered through a Millipore 0.22 μm filter.

3.2.5 Protein immobilization

Ion exchange resin Lewatit VP OC 1600 (SYBRON Chemicals Inc.) was used for lipase immobilization. Resin was pre-washed with ethanol and dried under vacuum. For immobilization, 1 g resin was added to 50 ml filtered fermentation supernatant and was incubated in a head-over-head shaker at 4 °C for a day. Residual activity in the solution was then tested using p-NB as a substrate (residual activity should be less than 10%). The supernatant was decanted, and fresh 50 ml supernatant was added and the incubation continued for another 24 hours. The process was repeated until the resin was saturated (indicated by the high residual activity of the solution). The resin was washed with 50 mM potassium buffer (pH 6.0) for several times, dried under vacuum, and stored at 4 °C.

3.2.6 Active site titration of immobilized lipase

The amount of active lipase immobilized on resin was determined by an active site titration assay. The active-site inhibitor methyl 4-methylumbelliferyl hexylphosphonate was synthesized by Christina Fields. Stock solution of the inhibitor (0.3 mM) was prepared in acetonitrile and stored at -20 °C. For a typical titration assay, 30 µl inhibitor stock solution was diluted in 970 µl acetonitrile, and six different amounts (1-15 mg) of immobilized enzyme was added. The inhibition reactions were incubated on a head-over-head shaker at room temperature for a week until the daily fluorescence reading remained unchanged. A 50 µl aliquot of the reaction mixture was taken and mixed with 450 µl Tris-HCl buffer (50 mM, pH 8.0) and the concentration of released 4-MU was determined by measuring the fluorescence intensity on a Synergy-HT microtiterplate reader (Bio-Tek Instruments, Winooski, VT) (excitation 360 nm, emission 460 nm). The amount of active lipase on resin was calibrated and calculated by the linear relationship of the amount of immobilized enzyme with the amount of 4-MU released.

3.2.7 Kinetic analysis of lipase catalyzed esterification reactions

Transesterification of 6-methyl-5-hepten-2-ol with vinyl acetate was performed in cyclohexane at room temperature (23 °C). Each 2 ml reaction mixture contains 1-10 mg immobilized enzyme, 50 mM internal GC standard (6-methyl-5-hepten-2-one) and varying amounts of 6-methyl-5-hepten-2-ol (25-1500 mM). The mixture was incubated for 30 min, and the reaction was initiated by the addition of vinyl acetate. Samples at different time points (1-6 min) were taken regularly to determine the initial reaction rates. For each reaction at least five samples were taken, and the overall conversions were controlled within 5%. The samples were analyzed by gas chromatography G6850

(Agilent Technologies) installed with a Cyclosil-B column (length 30 m, i.d. 0.32 mm, film 0.25 mm, Agilent) connected to a flame ionization detector. Hydrogen was used as the carrier gas, and the temperature was 75 °C for 30 min. The retention time was 15.8 min for (*S*)-6-methyl-5-hepten-2-ol and 17.1 min for its (*R*)-enantiomer.

Transesterification of 3-hydroxy tetrahydrofuran with vinyl acetate was performed the same way except that acetonitrile was used as the solvent, and the concentration of the internal GC standard was lowered to 10 mM. The temperature program for GC analysis was: 65 °C for 5 min, 2 °C/min to 90 °C, then 10 °C/min to 120 °C. The retention times for (*R*)- and (*S*)- products are 15.1 and 16.9 min.

Transesterification of α -methyl-1-naphthalenemethanol with vinyl acetate was performed in acetonitrile at room temperature. Benzophenone was used as internal standard. Each 2 ml reaction mixture contains 1-10 mg immobilized enzyme, 10 mM internal standard and varying amount of substrate (25-1500 mM). The temperature program for GC analysis was 160 °C for 40 min. The retention time was 16.5 min for (*S*)-enantiomer, 18.3 min for (*R*)-enantiomer, and 11.9 min for (*R*)-product.

Esterification of (*R*)- and (*S*)-2-fluoro- α -methyl-4-biphenylacetic acid (flurbiprofen) with 1-propanol was performed in 4-methyl-2-pentanone at 50 °C. Each 2 ml reaction mixture contains 10-40 mg immobilized enzyme, 5 mM internal standard benzophenone and varying amount of pure enantiomers of flurbiprofen (50-600 mM). The mixture was incubated at 50 °C for 30 min, and then the reaction was initiated by the addition of 1-propanol. The molar ratio of flurbiprofen to 1-propanol is 1:1. Samples at different time points were taken regularly to determine the initial reaction rates. For each reaction at least five samples were taken, and the overall conversions were controlled

within 5%. The samples were analyzed by gas chromatography G6850 (Agilent Technologies) installed with a HP-1 column (length 30 m, i.d. 0.32 mm, film 0.25 mm, Agilent) connected to a flame ionization detector. Hydrogen was used as the carrier gas, and the temperature was set at 170 °C for 30 min. The retention time was 10.4 min for flurbiprofen and 12.8 min for the product.

Esterification of (*R*)- and (*S*)-phenylpropionic acid with 1-propanol was performed as described for flurbiprofen, except that 4'-methoxy acetophenone (final concentration 10 mM) was used as an internal standard. The temperature program for GC analysis was 110 °C for 10 min. The retention time was 5.4 min for phenylpropionic acid and 7.4 min for the product.

3.3 Results and discussions

Previous kinetic study on selected circularly permuted CALB variants showed that variants with new termini relocated in or near helix $\alpha 10$ (amino acid P268-A287) show the most substantial improvement in catalysis. As helix $\alpha 10$ forms part of the substrate binding pocket, it was reasoned that the introduction of new termini in that region increases local chain flexibility and substrate accessibility, which in turn translates into higher activity. The three promising variants studied include cp283, cp284 and cp289, with the termini positioned either at the upper-middle portion of the helix (cp283 and cp284) or at the C-terminal end of the helix (cp289). As functional variants were identified from library screening with new termini all along the long helix $\alpha 10$, it would be interesting to see how termini at other positions of the helix impact on enzymatic activity and whether a better variant with greater rate enhancement could be identified. For that purpose, three more variants were chosen for further study, including cp268 with

new termini at the N-terminus of the helix, as well as cp277 and cp278 with new termini near the bent region of the helix $\alpha 10$ and also in a position closest to the catalytic triad (Figure 3-1).

Another interesting candidate is cp193 with new termini generated between amino acid P192/Q193, only six amino acids away from the catalytic aspartic acid D187. With a break in the peptide chain close to the catalytic machinery of the enzyme, an effect on activity is likely in either a positive or negative way.

In our previous design for the generation of the random circular permutation CALB library, a (His)₆-tag was attached to the N-terminus of the variants to simplify the protein purification process. Although we have shown that libraries with/without the affinity tag have the same location and distribution of permutation sites in functional CALB variants, it turned out later that some variants failed to be purified by the affinity tag, most probably because the His-tag is not accessible upon relocation of the termini to the protein's interior region. Instead, a two-step purification procedure using hydrophobic interaction chromatography and gel filtration worked well for all the variants. In addition, small variations in chain length of individual permutants were found, which were artifacts caused by reading frame shifts and staggered ends upon *DNaseI* digestion. The unwanted (His)₆-tag and extra extensions at the C-terminus were corrected by PCR for all the variants selected, including three variants with termini at the putative lid region (Figure 3-1). All ten variants were expressed in yeast *P. pastoris* and purified to homogeneity. An example of the SDS-PAGE gel after protein purification was shown in Figure 3-2.

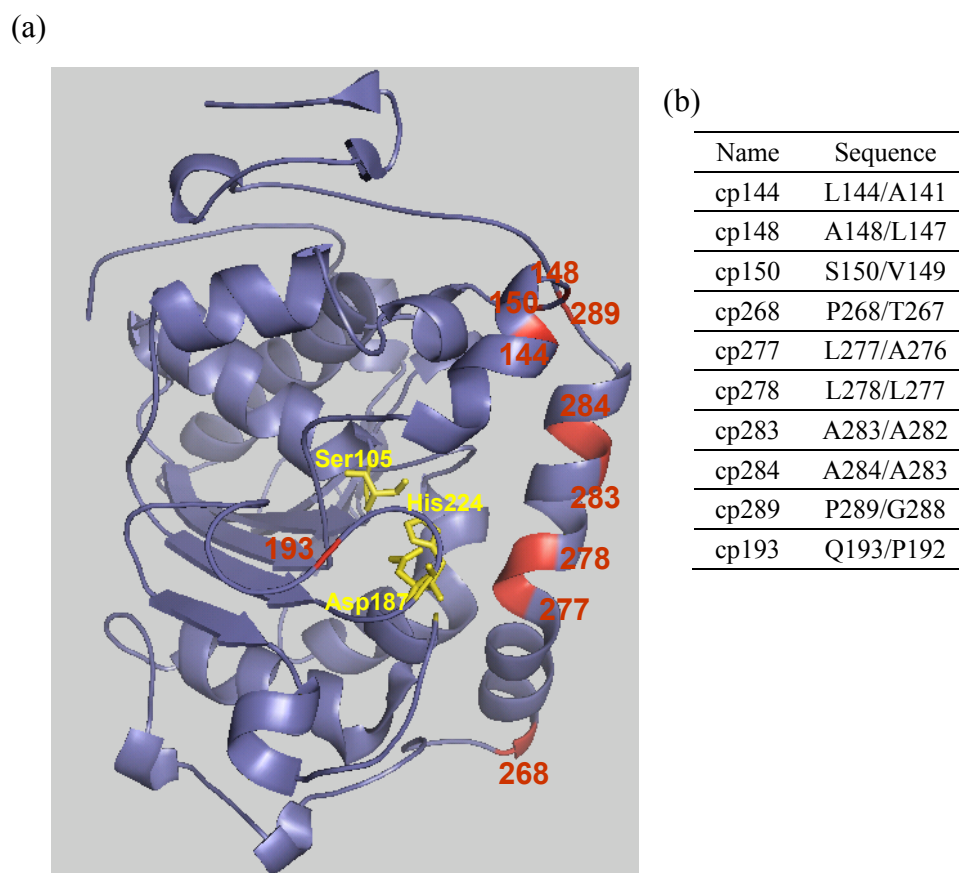


Figure 3-1 Selected ten variants for characterization (a) Candidates selected for subsequent *in vitro* characterization are labeled in red. The residues of the catalytic triad are shown in yellow. (b) Name and sequence of selected variants. Nomenclature: cp144 = circularly permuted protein whose N-terminus starts at amino acid 144 of the wild-type sequence. Sequence lists the N- and C-terminal amino acids (all in single-letter code) of the variant.

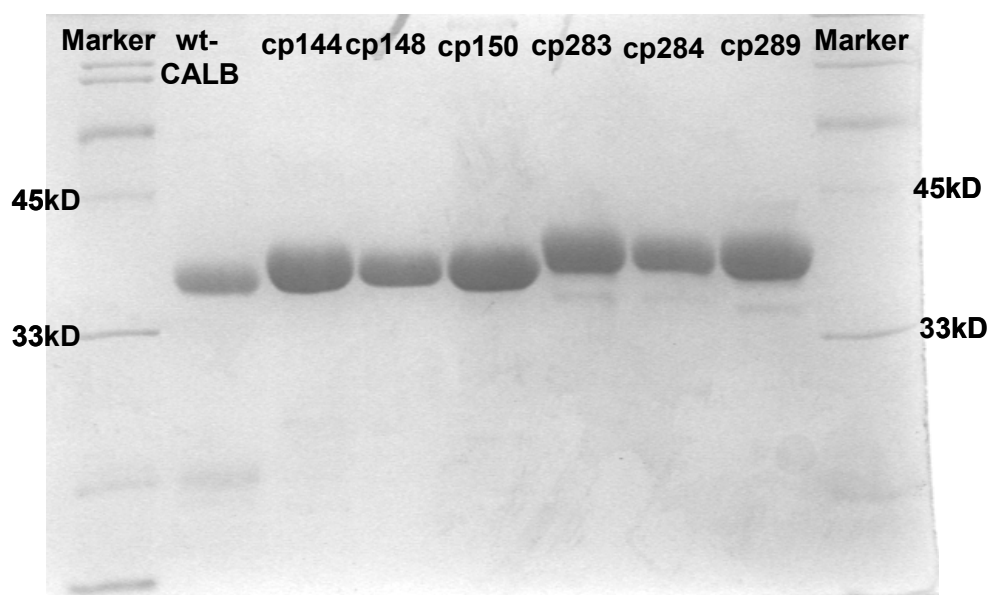


Figure 3-2 SDS-PAGE gel example of purified proteins

3.3.1 Kinetic analysis of (His)₆-tag and C-terminal extension free circularly permuted CALB variants

The hydrolytic activity of the variants toward substrates p-nitrophenyl butyrate (p-NB) and DiFMU octanoate were determined (Table 3-2). Kinetic analysis confirmed that the presence/absence of a (His)₆-tag doesn't change the activity of three lid region variants (cp144, cp148, cp150). Compared to the wild type CALB, their k_{cat} and K_M values for test substrates are within two-fold that of the wild type CALB. As suggested by the Hult group, this putative lid region of CALB may help in anchoring the long acyl chain of the lipid (Martinelle, Holmquist et al. 1995). Therefore, future study may involve exploring the activity of these variants toward substrates with an increasing acyl chain length.

Removal of the (His)₆-tag and C-terminal extension from three helix α 10 variants (cp283, cp284, cp289) has marginal effect on their k_{cat}/K_M values for p-NB hydrolysis. Compared to wild type CALB, the rate enhancement for all three variants is around 10-fold. However, a substantial improvement was observed on DiFMU octanoate hydrolysis. For both cp283 and cp284, after removing the terminal extensions, the rate enhancement is over 100-fold compared to wild type CALB. The improvement comes from the faster turn-over rate of the substrate, as reflected in their over 100-fold increase in the k_{cat} value and unchanged K_M . In addition, the position of new termini is relevant to the hydrolytic activity of the variant on this specific substrate, as cp283 shows the greatest rate improvement (174-fold), cp284 has slightly lower activity than cp283 (141-fold increase), while cp289 only shows a 37-fold activity increase. As helix α 10 is involved in the binding of the alcohol portion of the substrate, and DiFMU octanoate has a bulkier

alcohol portion compared to p-NB, the result supports our hypothesis that new termini near the active site increases local chain flexibility which is advantageous for structurally more demanding substrates.

The kinetic analysis on three more variants (cp268, cp277, and cp278) provided a better picture about the relationship between the position of new termini on helix $\alpha 10$ and the extent of rate enhancement. Although amino acids Leu277 and Leu278 are in the middle of the helix $\alpha 10$ and right next to the active site pocket, cp277 and cp278 didn't show high activity as expected. Their activity for p-NB hydrolysis is only 2~4-fold higher than the wild type CALB and 6~12-fold higher for DiFMU octanoate hydrolysis. Noticeable is their almost 2-fold higher K_M values and much lower k_{cat} s compared to those of cp283. Maybe the flexibility in the peptide backbone close to the active site is detrimental for efficient binding of the substrate. The activity of cp268, a variant with termini at the N-terminal end of the helix $\alpha 10$, also shows only moderate rate enhancement for both substrates (around 10-fold).

Finally, the activity of cp193 shows a 10-fold decrease for both substrates compared to the wild type CALB. Since the new termini are in close proximity to the catalytic Asp187, the flexibility of the termini may result in a slight fluctuation of the position of Asp187, which interferes with the proper orientation of the charge relay for catalysis. Combining the results above, kinetic analysis on ten selected circularly permuted variants suggests that efficient catalysis of the enzyme prefers flexibility of residues contribute to topology of the active site pocket rather than those directly involved in catalysis.

Table 3-2 Apparent kinetic constants for selected CALB variants with p-nitrophenyl butyrate and DiFMU octanoate as substrates

enzyme variant		p-nitrophenyl butyrate				DiFMU octanoate			
Name ^a	sequence ^b	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1}\text{min}^{-1}$)	rel. speci. ^c	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1}\text{min}^{-1}$)	rel. speci. ^c
Wild type	L1/P317	410 ± 40	305 ± 10	0.74	1	2.6 ± 0.3	2 ± 0.1	0.8	1
cp144	L144/A141	820 ± 100	435 ± 28	0.53	0.7	3.0 ± 0.3	1.5 ± 0.1	0.49	0.6
cp148	A148/L147	550 ± 70	481 ± 29	0.87	1.2	6.1 ± 0.7	3.5 ± 0.2	0.57	0.7
cp150	S150/V149	425 ± 50	347 ± 16	0.82	1.1	5.5 ± 0.9	2.2 ± 0.2	0.41	0.5
cp268	P268/T267	580 ± 90	3051 ± 229	5.26	7	2.3 ± 0.3	28.8 ± 1.1	12.52	16
cp277	L277/A276	820 ± 100	1356 ± 94	1.65	2	3.4 ± 0.5	16.3 ± 0.9	4.87	6
cp278	L278/L277	1180 ± 160	3117 ± 269	2.64	4	5.2 ± 0.8	49.9 ± 3.2	9.65	12
cp283	A283/A282	410 ± 60	3251 ± 206	7.93	11	2.4 ± 0.4	340.2 ± 16.9	139.51	174
cp284	A284/A283	520 ± 80	4380 ± 298	8.42	11	2.2 ± 0.2	242.5 ± 7.6	112.62	141
cp289	P289/G288	790 ± 85	8055 ± 455	10.17	14	5.1 ± 1.0	149.8 ± 12.9	29.52	37
cp193	Q193/P192	1200 ± 170	80 ± 6	0.07	0.1	1.9 ± 0.3	0.2 ± 0.01	0.10	0.1

Note: a. CALB nomenclature: cp144 = circular permuted protein whose N-terminus starts at amino acid 144 of the wild-type sequence. b. N- and C-terminal amino acids (all in single-letter code) are listed. c. Relative specificity = k_{cat}/K_M (variant)/ k_{cat}/K_M (wild type).

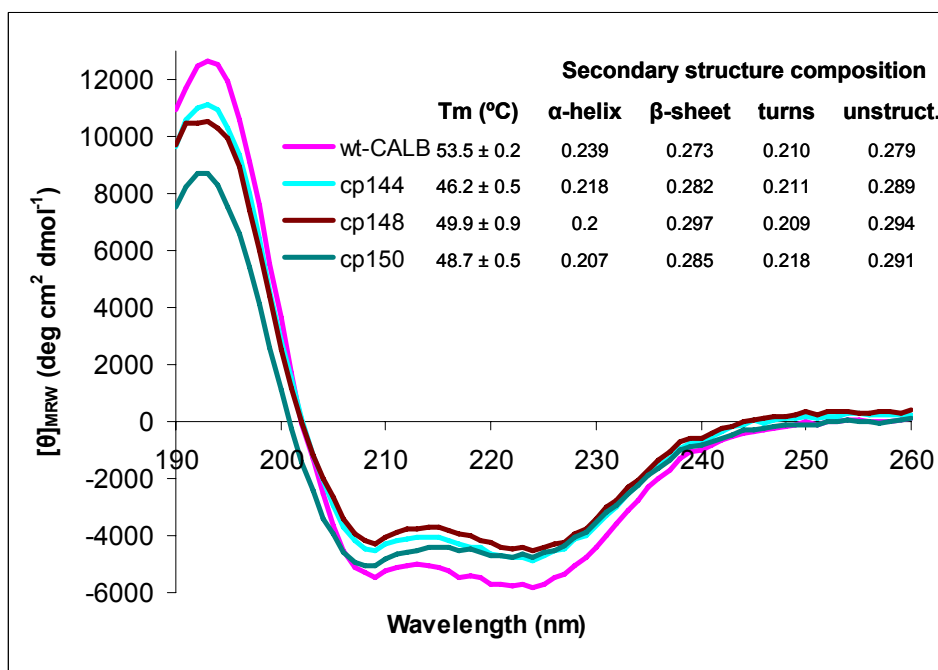
3.3.2 Circular dichroism analysis of selected variants

Far-UV circular dichroism (CD) spectra of selected variants were determined to examine how circular permutation would impact on the integrity of the enzyme (Figure 3-3). The spectrum of wt-CALB has negative peak maxima at 208 nm and 222 nm, characteristic of the α -helical content. The spectra of the variants have similar shape as that of the wild type enzyme, suggesting that the variants maintain their overall secondary structure as wt-CALB. However, reduced ellipticity at wavelength around 190 nm, and at 208 and 222 nm suggests a partial loss of the helical content, which is confirmed by calculations of the secondary structure percentage by using the software CDPro (Sreerama and Woody 2000) (CONTINLL, data set: SMP56).

Compared to the wild type CALB, circularly permuted variants have decreased temperature of unfolding according to the thermo-denaturation temperature (T_m) measurement. In comparison, the reduction in thermostability for the putative lid region variants (cp144, cp148 and cp150) is less severe than that of the helix α 10 variants. Wild type CALB has a T_m around 54 °C, and T_{ms} of three lid region variants fall between 46~50 °C, while T_{ms} of all the helix α 10 region variants are lower than 42 °C (Figure 3-3).

As mentioned, most other studies on circularly permuted proteins have shown that the variants keep the overall structures as the wild type enzymes with only small local variations in some regions, especially around the linker used to tether the original termini and near the new termini generated (Zhang, Bertelsen et al. 1993; Mullins, Wesseling et al. 1994; Ni and Schachman 2001; Fishburn, Keeffe et al. 2002; Manjasetty, Hennecke et al. 2004; Sagermann, Baase et al. 2004). Furthermore in all these studies the stability of

(a)



(b)

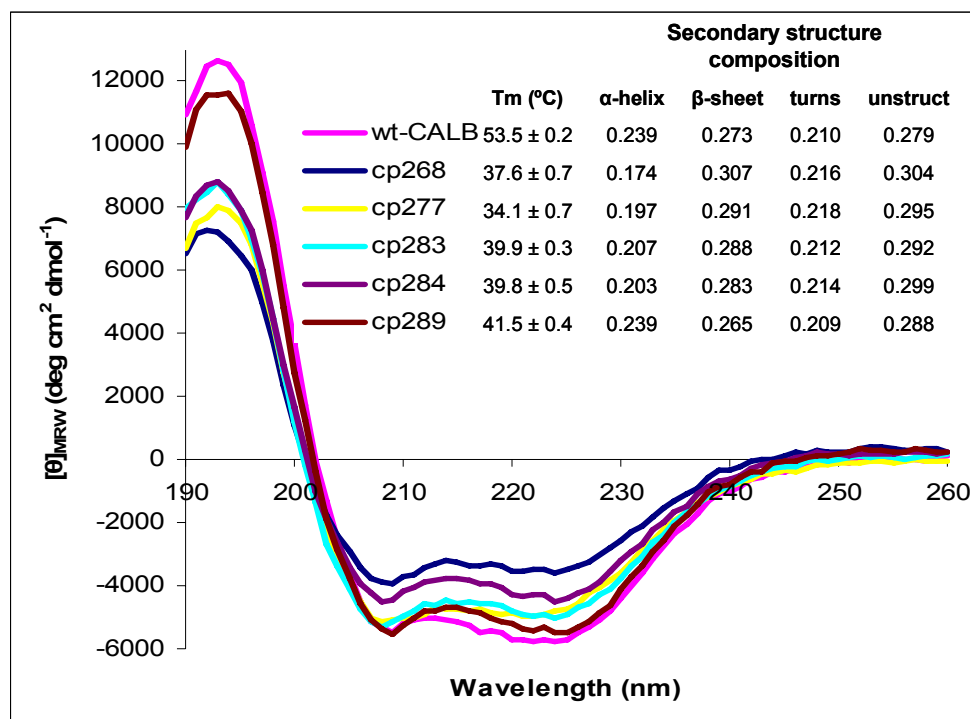


Figure 3-3 Far-UV circular dichroism (CD) spectra and T_m of selected variants. Secondary structure composition was calculated using CDPro (CONTINLL, data set: SMP56) (a) variants with new termini in/near helix α17 (b) variants with new termini in the putative lid region

circularly permuted variants is consistently lower than the wild type enzyme. Owing to the overall structural similarity, the reduced thermostability is suggested to be caused by these small structural variations. Our result is consistent with previous studies, although the reason responsible for reduced thermostability is unclear.

3.3.3 Enantioselectivity comparison of cp283 with wt-CALB

Wild type CALB shows high enantioselectivity toward esterification of secondary alcohols. The impact of circular permutation on the substrate specificity and enantioselectivity of the variants was studied by using cp283 as a model in view of its high hydrolytic activity on test substrates. Both cp283 and wt-CALB were expressed in large quantities by *P. pastoris* fermentation and immobilized on the ion exchange resin Lewatit VP OC 1600 for reactions in organic solvents. The amount of active enzyme on the resin was quantified by an active-site titration assay using a phosphonate inhibitor – methyl 4-methylumbelliferyl hexylphosphonate as described in Materials and Methods section.

The kinetic constants for transesterification of three secondary alcohol substrates (6-methyl-5-hepten-2-ol, 3-hydroxytetrahydrofuran, and 1-(1-naphthyl) ethanol) with vinyl acetate, catalyzed by the immobilized enzymes, were determined and the enantioselectivity of the enzyme was calculated. Overall the results showed that circular permutation doesn't change the enantiomeric preference of the enzyme (Table 3-3). Same as wt-CALB, cp283 shows high enantioselectivity for substrate 6-methyl-5-hepten-2-ol and 1-(1-naphthyl) ethanol. In both cases, the (*R*)-enantiomers of the substrates are preferred and no esterification products of (*S*)-enantiomers were detected above the GC detection limit over the length of the assay. In addition, cp283 has a 3-fold activity

increase for transesterification of (*R*)-6-methyl-5-hepten-2-ol compared to wt-CALB, which comes from an increase in k_{cat} , indicating a faster turnover rate of the substrate. Both enzymes show similar activity for transesterification of (*R*)-1-(1-naphthyl) ethanol, but due to a bulkier size of the naphthyl substitution compared to that in 6-methyl-5-hepten-2-ol, the $k_{\text{cat}}/K_{\text{M}}$ value drops to $\sim 2 \text{ min}^{-1}\text{mM}^{-1}$.

Both cp283 and wt-CALB have poor enantioselectivity toward 3-hydroxytetrahydrofuran with an E value ~ 2 in favor of the (*S*)-enantiomer. According to previous studies on CALB, one requirement for high enantioselectivity is that the size of two substituents on the chiral carbon should be significantly different with a small substituent smaller than *n*-propyl (Rotticci, Haeffner et al. 1998). The special ring structure of 3-hydroxytetrahydrofuran certainly does not fulfill the requirement. It was also noticed that although cp283 shows similar enantioselectivity as the wt-CALB for transesterification of 3-hydroxytetrahydrofuran, its $k_{\text{cat}}/K_{\text{M}}$ value is slightly reduced. The reduction is due to an almost doubled K_{M} value, indicating a poor binding of the substrate, which might be caused by the increased local flexibility of new termini in helix $\alpha 10$.

Besides secondary alcohol substrates, the enantioselectivity of cp283 toward chiral carboxylic acids was also tested via the esterification of flurbiprofen and phenylpropionic acid with *n*-propanol (Table 3-4). Both wt-CALB and cp283 prefer the same enantiomers – the (*R*)-enantiomer of the substrates. In addition, cp283 exhibits 4-fold higher activity and almost 2-fold greater enantioselectivity compared to wt-CALB.



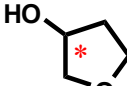
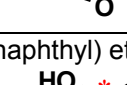
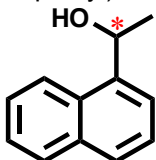
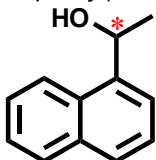
Enzyme		Substrate	K_M (mM^{-1})	$k_{cat} \times 10^{-3}$ (min^{-1})	k_{cat}/K_M ($min^{-1} mM^{-1}$)	rel. speci.	E
WT	R-	6-methyl-5-hepten-2-ol 	240±36	8.9 ± 0.6	37	1.0	>40
	S-		ND	ND	ND	ND	
cp283	R-		260±39	31.3 ± 2.0	121	3.2	>40
	S-		ND	ND	ND	ND	
WT	R-	3-hydroxytetrahydrofuran 	201±19	0.6 ± 0.02	3.1	1.0	1.8
	S-		251±23	1.4 ± 0.05	5.6	1.0	
cp283	R-		408±50	0.8 ± 0.04	2.0	0.6	2.0
	S-		457±49	1.8 ± 0.08	4.0	0.7	
WT	R-	1-(1-naphthyl) ethanol 	845±103	1.6 ± 0.1	1.9	1.0	>15
	S-		ND	ND	ND	ND	
cp283	R-		946±75	2.4 ± 0.1	2.5	1.3	>15
	S-		ND	ND	ND	ND	

Table 3-3 Kinetic and enantioselectivity analysis on lipases (wt-CALB and cp283) catalyzed transesterification reaction with vinyl acetate. $E = k_{cat}/K_M$ (fast enantiomer) / k_{cat}/K_M (slow enantiomer), Relative specificity = k_{cat}/K_M (cp283) / k_{cat}/K_M (wt-CALB)

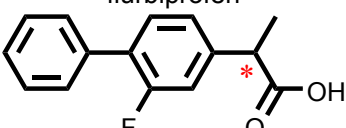
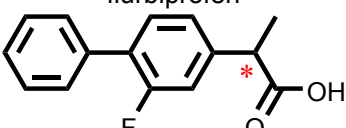
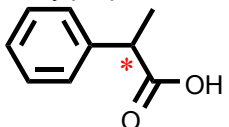
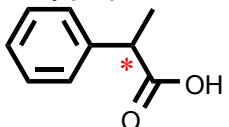
Enzyme		Substrate	K_M (mM^{-1})	k_{cat} (min^{-1})	k_{cat}/K_M ($min^{-1} mM^{-1}$)	rel. speci.	E
WT	R-	flurbiprofen 	182±27	17.4 ± 1.2	0.10	1.0	26
	S-		393±50	1.4 ± 0.1	0.004	1.0	
cp283	R-		163±33	64.3 ± 5.9	0.39	3.9	42
	S-		402±35	3.8 ± 0.2	0.01	2.5	
WT	R-	phenylpropionic acid 	238±45	70 ± 6	0.29	1.0	5
	S-		412±37	24 ± 1	0.06	1.0	
cp283	R-		125±27	86 ± 6	0.69	2.4	9
	S-		354±21	27 ± 1	0.08	1.3	

Table 3-4 Kinetic and enantioselectivity analysis on lipases (wt-CALB and cp283) catalyzed esterification reactions with n-propanol. $E = k_{cat}/K_M$ (fast enantiomer) / k_{cat}/K_M (slow enantiomer). Relative specificity = k_{cat}/K_M (cp283) / k_{cat}/K_M (wt-CALB)

Overall the retained enantiomeric preference and similar or better esterification activity of cp283 again suggest that the variant maintains its three dimensional structure and the active-site pocket topology as wt-CALB, and the increased local chain flexibility of the new termini impacts the binding of substrate and its catalytic efficiency.

3.4 Conclusions

In summary, circular dichroism and enantioselectivity analysis suggested that the circularly permuted CALB variants maintain the overall three dimensional structures as the wild type enzyme in view of the same shape of their far-UV circular dichroism spectra and preference for enantiomers. The variant cp283 shows the greatest rate enhancement (174-fold) for DiFMU octanoate hydrolysis, and similar or better enantioselectivity toward esterification of test substrates.

However, one drawback of circular permutation is that it usually results in variants with decreased thermostability, which is also the case for our circularly permuted CALB variants. To improve the thermostability, the variant cp283 was subjected to a secondary generation engineering which is discussed in the next chapter.

Chapter 4 Secondary engineering and thermostabilization of circularly permuted CALB

4.1 Introduction

In previous chapters, we have shown that circular permutation has a significant effect on CALB's catalytic performance. The best variant cp283 exhibits a 174-fold activity increase on DiFMU-octanoate hydrolysis and similar or better enantioselectivity for a number of substrates tested compared to wt-CALB. Circular dichroism analysis suggested that relocation of termini didn't affect the overall structure of the variant. Instead, the increased local chain flexibility near the new termini impacts the catalytic efficiency of the variant (Chapter 3).

One side effect of circular permutation on CALB is the reduction in the variants' thermostability. As determined by circular dichroism, the thermal denaturation temperature of cp283 drops from 53 °C for wt-CALB to 40 °C. The stability loss was also observed for other circularly permuted proteins such as T4 lysozyme, ribonuclease T1, aspartate transcarbamoylase, myoglobin, and DsbA (Zhang, Bertelsen et al. 1993; Mullins, Wesseling et al. 1994; Ni and Schachman 2001; Fishburn, Keefe et al. 2002; Manjasetty, Hennecke et al. 2004; Sagermann, Baase et al. 2004). It was reasoned that destabilization is caused by the structural differences between the wild type enzyme and the variant. However, most studies showed that circularly permuted variants maintain the native-like overall structure with only small variations, mainly in regions around the

connected old termini and near the new termini (Hahn, Piotukh et al. 1994; Ay, Hahn et al. 1998; Chu, Freitag et al. 1998; Rojas, Garcia-Vallve et al. 1999; Fishburn, Keeffe et al. 2002; Tougard, Bizebard et al. 2002; Manjasetty, Hennecke et al. 2004; Sagermann, Baase et al. 2004). Therefore, structural changes in these two areas are the most likely source of destabilization. One example of proof is the study on T4 lysozyme. Although the circular permutants assume the same overall structure as the wild type T4 lysozyme, the composition and length of the linker used to connect the natural termini was found to result in small structural variations near that region and was directly coupled to the degree of destabilization (Sagermann, Baase et al. 2004).

In our study, a six amino acid glycine-rich linker (-GGTSGG-) was used to connect the natural termini. The flexible linker bridges the distance between the original termini (17 Å) without causing strain, yet forms an extended loop together with the original termini (Figure 4-1), whose flexibility is entropically unfavorable and might become a factor of destabilization.

A study by the Regan group showed that, replacing a two residue loop in a four-helix-bundle protein with increasing lengths of glycine linkers destabilizes the protein because the closure of longer loops is entropically unfavorable (Nagi and Regan 1997). Similar results were obtained in the study on circularly permuted α -spectrin SH3 domains, in which lengthening the linker by increasing the number of glycines from two to ten accelerates unfolding and slows down the folding rate of the protein (Figure 4-2) (Viguera and Serrano 1997). Therefore, it was reasoned that if the linker used in our study is longer than necessary, the flexibility of the loop formed would destabilize the enzyme, while partial truncation of the loop could recover its stability.

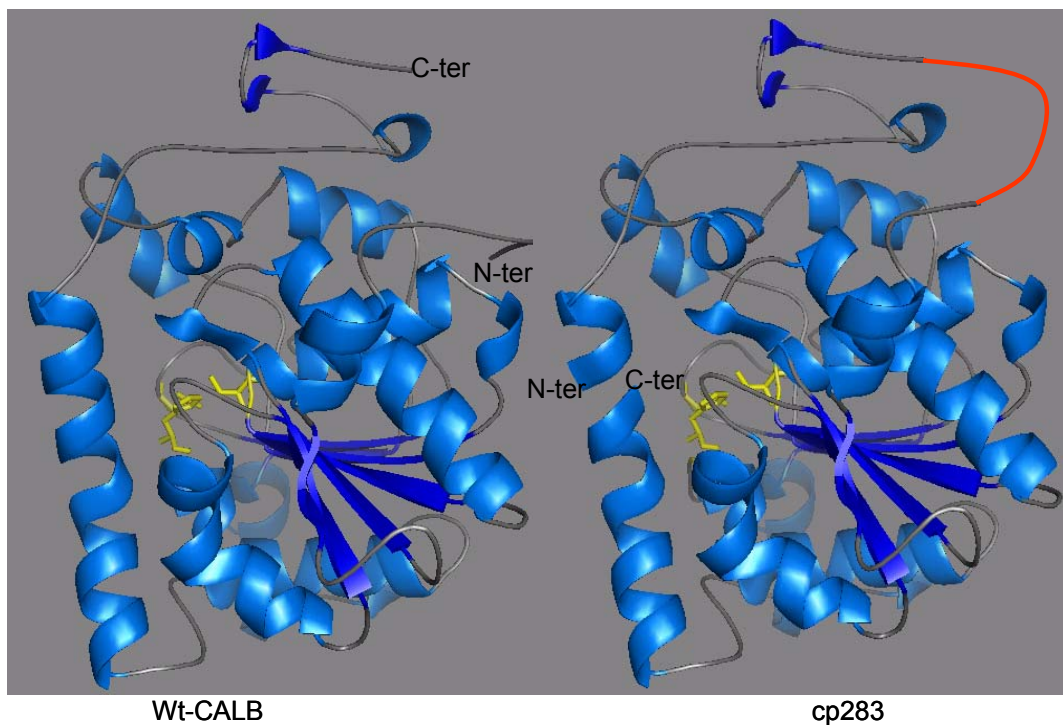


Figure 4-1 Schematic overview of the structures of wt-CALB and cp283. The linker used to connect the natural termini is shown in red. Active site residues are shown in yellow sticks. The N- and C-termini are indicated.

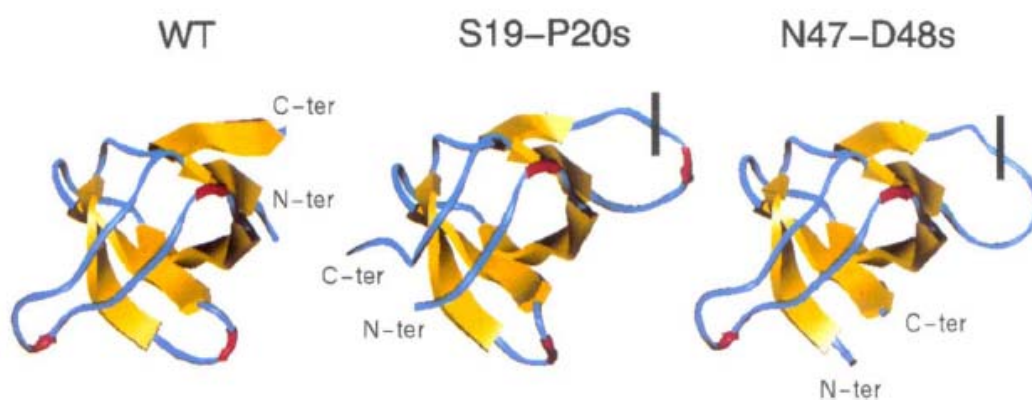


Figure 4-2 Ribbon diagram of the WT α -spectrin SH3 domain and the circular permutants S19-P20s and N47-D48s. the black bar shows the point in which the glycine residues have been inserted (Viguera and Serrano 1997).

To test the hypothesis and increase the thermostability of cp283, incremental truncation was employed to randomly shorten the loop formed by the linker and the native termini. Following generation of the incremental truncation library, functional variants were identified by the previous used tributyrin screen. Variants with truncation up to eleven amino acids (~25% of the loop) were found.

4.2 Materials and methods

4.2.1 Creation of the cp283 incremental truncation library

The cp283 gene was incorporated into vector pAMB-CAT using *NotI* and *XhoI* restriction sites. The plasmid was then linearized by *SpeI* digestion. The incremental truncation library was generated following the protocol of Marc Ostermeier and Stefan Lutz (Ostermeier and Lutz 2003). In detail, the linearized plasmid was amplified by *Taq* DNA polymerase using primers ZQ_cpCALB_for (5'-GGTACTAGTGGTGGCCTACCTTCCGGTTCGGACCCT-3'), ZQ_cpCALB_rev (5'-CGCACTAGTACCGCCGGGGGTGACGATGCGGGAGCA-3'), and spiked dNTPs (dNTP : α S-dNTP = 7 : 1). After purification with Qiagen's QIAquick PCR purification kit, the PCR product was digested by *ExonucleaseIII* (120 units/ μ g DNA) at 37 °C for 30 min. The reaction was quenched by the addition of 5 volumes of PB buffer and purified by QIAquick PCR purification kit. The 5'-overhang was removed by incubation with Mung bean nuclease (2.5 units/ μ g DNA, DNA concentration 0.1 μ g/ μ l) at 30 °C for 30 min, and the DNA was purified by Qiagen spin columns. The purified DNA was treated by Klenow polymerase to repair the sticky ends (1 units/ μ g DNA, DNA concentration 0.1 μ g/ μ l, 25 °C for 15 min and 75 °C for 20 min). After purification by Qiagen's QIAquick spin column, intramolecular

ligation was performed (DNA concentration: 2.5 ng/ μ l, 16 °C overnight). The ligation mixture was concentrated by ethanol precipitation and electroporated into electro-competent *E. coli* DH5 α cells. About 3 million colonies were obtained. Purified plasmid was subjected to *NotI* and *XhoI* digestion, and the DNA fragment between 750 and 1000 bp was extracted from agarose gel. The fragment was ligated into the vector pPIC9 digested with the same restriction enzymes. The ligation mixture was again transformed into DH5 α cells and a library of $\sim 4.5 \times 10^5$ colonies was obtained. The library was harvested and the plasmid was purified from cells, digested with *SacI*, transformed into *Pichia pastoris* strain GS115 and plated on MM-tributyryn plates. Active library members were visualized by halos around the colonies. Those colonies were picked and submitted for DNA sequencing.

4.2.2 Creation of the C-terminal incremental truncation library

The wild type CALB gene was PCR-amplified using primers CALB_for_hisfree (5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTCTACCTTCCGGTTCGGACCCTGC C-3') and ZQ_CALB_rev (5'-CGCGCGGCCGCTTAGGGGGTGACGATGCCGGAGC A-3'). The PCR product was digested with *NotI* and *XhoI* and ligated into the vector pAMB-CAT digested with the same restriction enzymes. The plasmid was linearized by *EcoRI* digestion. Linearized plasmid was amplified by *Taq* DNA polymerase using primers Trunc_for (5'-GAGCTCCGTCGACAAGCTTGCGG-3') and Trunc_rev (5'-GGATGAGCATTTCATCAGGCGGGCA-3') and with spiked dNTPs. The following procedure was the same as that of the cp283 library, except that after Klenow polymerase treatment and purification by Qiagen spin column, the DNA was digested with *XhoI*, and

size selection (fraction between 750 bp and 1 kb) was performed afterwards by gel extraction.

The extracted DNA was ligated into a modified vector pAMB-pET digested with *PsiI* and *XhoI*, and transformed into *E. coli* DH5 α cells. Around 1.5×10^5 colonies were obtained. The cells were harvested and the plasmid was purified by Qiagen miniprep kit. After digestion of the plasmid with *NotI* and *XhoI*, the fraction containing CALB gene fragments was extracted and ligated into the plasmid pPIC9 digested with the same enzymes. The ligation mixture was again transformed into *E. coli* DH5 α cells and a library of 1.2×10^6 colonies was obtained. The plasmid was purified, digested with *SacI*, transformed into *Pichia pastoris* strain GS115 and screened on MM-tributyryn plates.

4.2.3 Protein expression, purification, and activity assays

Protein expression and purification was performed as described in Appendices B1 and B3. The activity assays on p-NB and DiFMU octanoate hydrolysis for all the variants were performed as described in Appendix B4.

4.2.4 Circular dichroism analysis and gel filtration

Circular dichroism was performed as described in Appendix B6. Gel filtration was performed on a Superdex-200 10/300 GL column (Amersham Biosciences, Piscataway, NJ), using 50 mM K-phosphate buffer (pH 7.0) containing 150 mM NaCl. Bio-Rad's gel filtration standard (Catalog # 151-1901) was used for molecular weight calculations.

4.3 Results and Discussions

4.3.1 Creation of the incremental truncation libraries

4.3.1.1 Creation of the cp283 incremental truncation library

Different from the wild type CALB, circularly permuted variant cp283 begins with sequence corresponding to amino acid 283~317 in wt-CALB, followed by the linker (GGTSGG), and then the amino acid 1~282 (Figure 4-3).

Creation of the cp283 incremental truncation library was performed following the ITCHY protocol of Ostermeier and Lutz (Ostermeier and Lutz 2003). Briefly, cp283 gene was incorporated into the pAMB-CAT vector, and the vector was linearized by utilizing the unique *SpeI* site in the middle of the linker sequence, which displayed the sequences of the natural termini on ends of linearized vector (Figure 4-4a). The following PCR amplification with spiked dNTPs (dNTP : α S-dNTP = 7 : 1) randomly incorporated α S-dNMPs into PCR products. Taking advantage of the observation that thio-phosphate linkers between two nucleosides can not be cleaved by *ExonucleaseIII*, the enzyme hydrolyzes the digested PCR product from 5' to 3' until the first thio-phosphate group is encountered. The single-stranded 3' to 5' overhang is then removed by Mung bean nuclease and the ends are made blunt-ended by Klenow polymerase treatment. The DNA product after all the treatments appeared as a smear when expanded on an agarose gel (Figure 4-4b). The following intramolecular ligation and transforming electro-competent *E. coli* DH5 α cells generated a pAMB-CAT-IT283 library of ~3 million colonies.

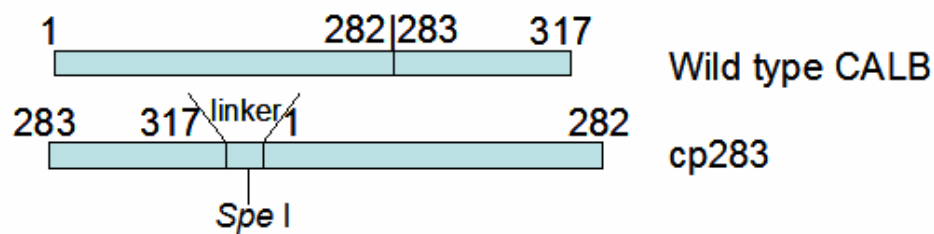


Figure 4-3 Graphic illustration of the sequences of wt-CALB and cp283. The number corresponds to the amino acid number in wt-CALB. The unique *SpeI* site in the linker was used in the generation of the incremental truncation cp283 library

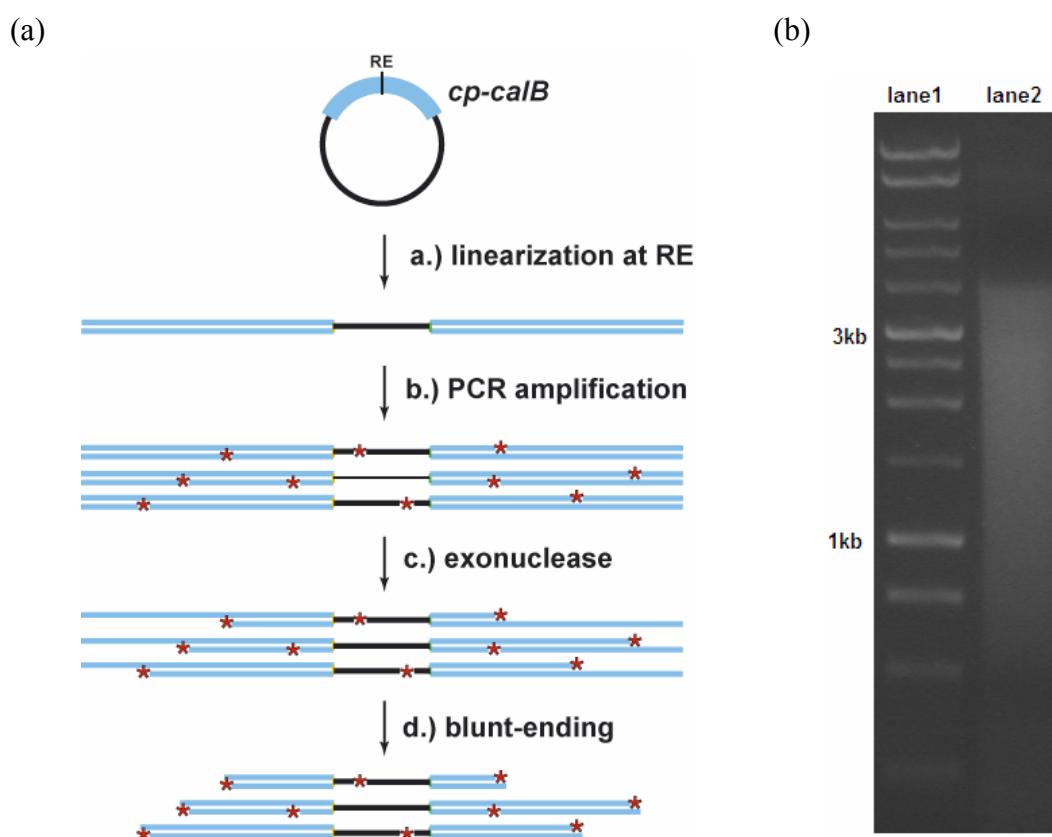


Figure 4-4 Creation of the incremental truncation library (a) Schematic overview of the creation of the cp283 incremental truncation library. (b) The truncation product ran on 0.7% agarose gel (lane 1, size marker; lane 2 truncation product)

Wild type CALB and its circularly permuted variants were successfully expressed in yeast *P. pastoris* (Rotticci-Mulder, Gustavsson et al. 2001; Qian and Lutz 2005). To utilize the yeast expression system and screen for functional variants, truncated cp283 genes were recovered from pAMB-CAT-IT283 plasmid by digestion with *XhoI* and *NotI*, and religated into the *P. pastoris* shuttle vector pPIC9 digested with the same enzymes. The reason behind the two-step subcloning strategy is the big size of pPIC9 (8 kb), which is problematic in manipulation. Before religation into pPIC9, a size selection of truncated cp283 gene length between 750-1000 base pairs was performed, eliminating a substantial part of smaller, presumably nonfunctional variants. The subsequent transformation of *E. coli* DH5 α cells yielded a library of ~0.45 million colonies.

The plasmid was purified from the library and electroporated into *P. pastoris* and colonies were screened by tributyrin plates. Functional variants were identified based on halo formation around colonies indicative of tributyrin hydrolysis (Figure 4-5).

4.3.1.2 Analysis of the cp283 incremental truncation library

Sequencing of 44 randomly picked variants showed that both original termini were truncated without visible bias, and size selection controlled the length of truncated gene in a reasonable range (750-1000bp) (Figure 4-6).

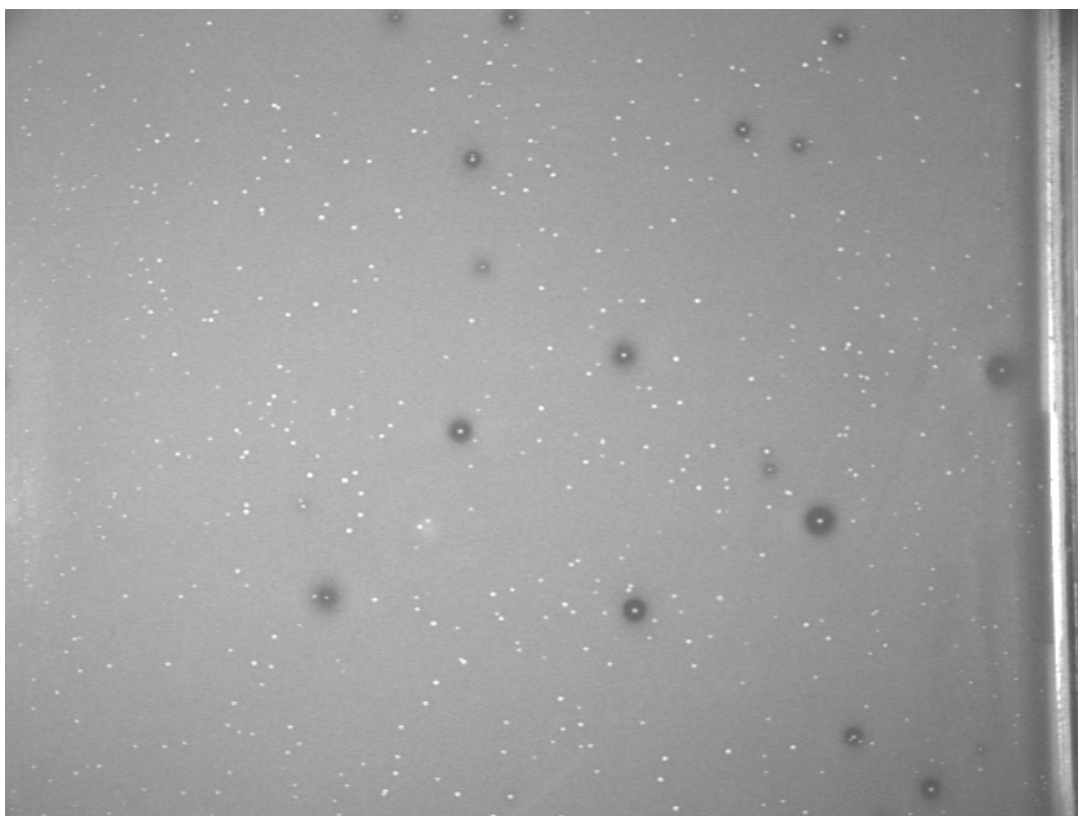


Figure 4-5 Incremental truncation library screening on tributyrin plates. Functional variants were indicated by halos around the colonies.

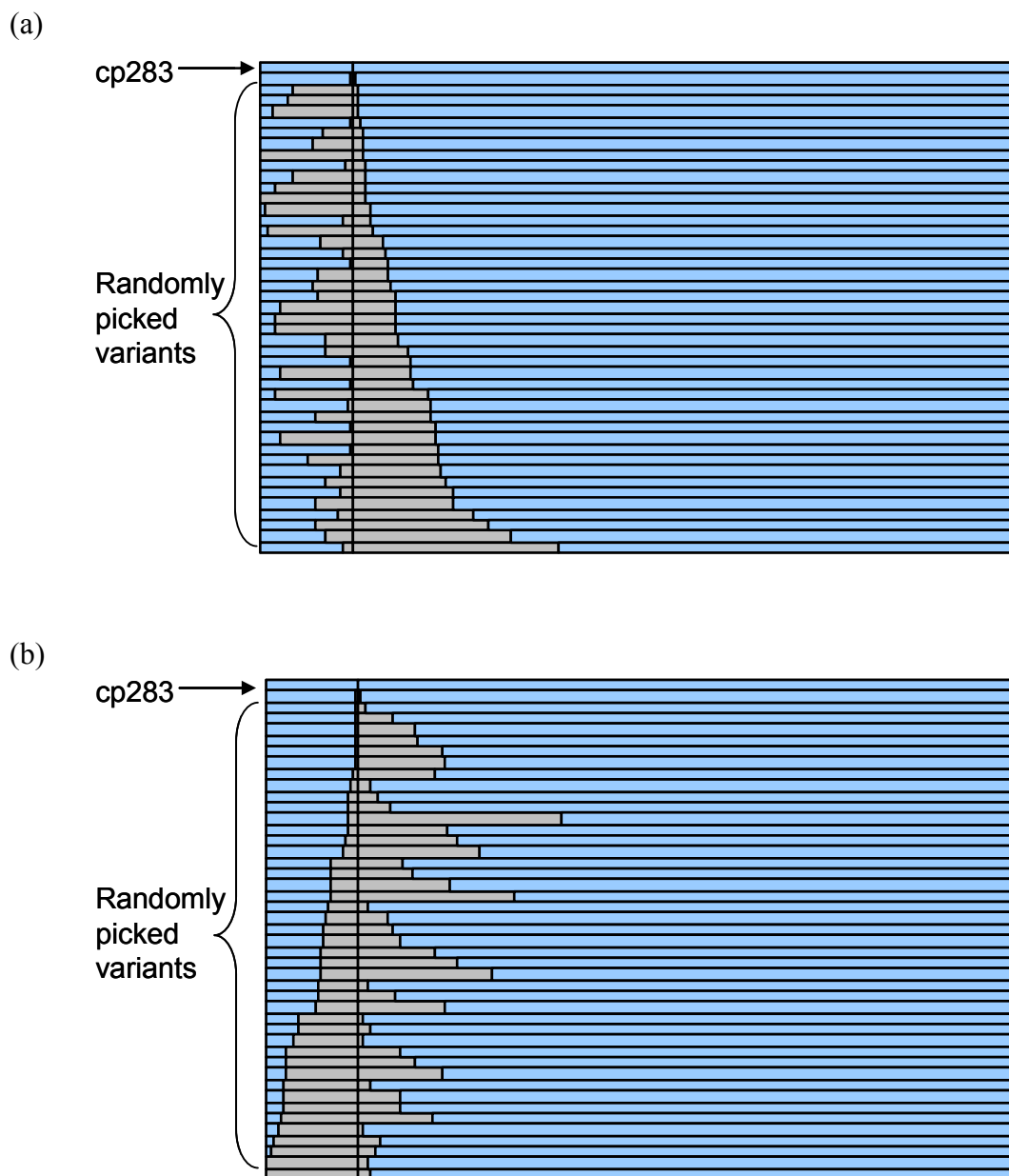


Figure 4-6 Sequences of variants picked from the native library sorted by the length of truncation on (a) the original N-terminus or (b) the original C-terminus. Shown in gray are the truncated portion, shown in blue are the leftover sequences of the variants.

DNA sequencing of ~100 functional variants revealed 31 different truncation patterns with deletions up to 11 amino acids compared to the parental cp283. As shown in Figure 4-4a, linearization of the plasmid split the cp283 gene in the middle of the linker sequence, followed by the sequences corresponds to the original N- and C- termini in wt-CALB. Ideally both termini have equal chance of truncation. However, library screening only identified functional variants with truncation at the original C-terminus, while the N-terminus was largely unchanged, suggesting the importance of N-terminus for either maintaining the correct folding or the function of the enzyme.

4.3.1.3 Creation and analysis of the C-terminal incremental truncation library

To examine how truncation of the C-terminus would impact the performance of wt-CALB, a C-terminal incremental truncation library of wt-CALB was created separately. Different from the generation of the cp283 library, after incorporation of wt-CALB into the vector pAMB-CAT, the vector was linearized by *EcoRI* digestion at the C-terminal end of CALB gene (Figure 4-7). Then α S-dNTPs incorporation and *ExonucleaseIII* digestion generated a library with truncations at the C-terminus of wt-CALB. The following steps were the same as that for the creation of the cp283 library (as described in Materials and Methods section).

After transformation of the library into *P. pastoris* strain GS115, the library was screened on tributyrin plates. DNA sequencing of functional variants identified truncations up to 16 amino acids at the C-terminus. However, the expression level of IT301 (in which A301 becomes the last amino acid instead of P317) is much lower, verified not only by the smaller halo size on tributyrin plate, but also by several separate overexpression attempts in shaking flasks. The yield of the protein after purification was

very low (< 0.5 mg/L), which made further characterization of the enzyme difficult. As the C-terminal region of wt-CALB contains a disulfide bond formed between residues C311 and C293, we speculated that the absence of the disulfide bond in IT301 destabilizes the enzyme.

Several nonfunctional variants were also analyzed by DNA sequencing and a variant with 21 amino acids' truncation was identified (in which D296 becomes the C-terminal residue). Therefore, although the C-terminus of wt-CALB is far from the active site and has very little secondary structures as shown in the crystal structure, it is indispensable for either ensuring the correct folding of the enzyme or maintaining its function and stability (Figure 4-8).

4.3.2 Characterization of selected variants

Eight cp283 variants with 2 to 11 amino acids' truncations were selected for further characterization (Figure 4-9). For comparison, one variant IT309 from the C-terminal truncation library was also selected. All the variants were overexpressed in *P. pastoris* strain GS115 and purified to homogeneity. The expression levels of cp283- $\Delta 2$, $\Delta 4$ and $\Delta 7$ were higher than those of the $\Delta 8$, $\Delta 10s$, $\Delta 11s$ and IT309. Purified proteins were studied by kinetic assay, circular dichroism, and size exclusion chromatography.

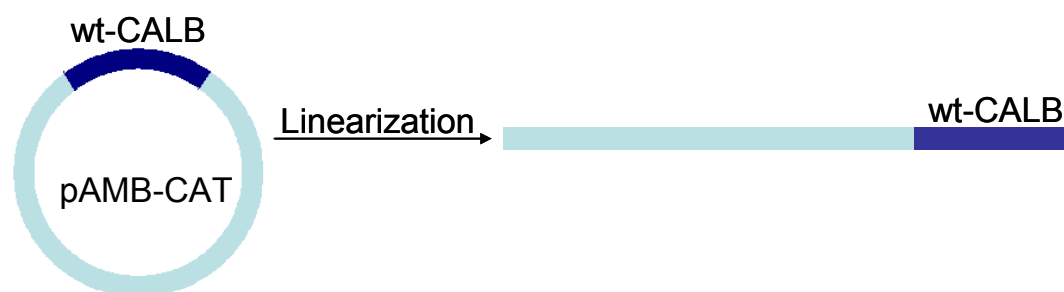


Figure 4-7 The first step in the creation of the C-terminal truncation library. wt-CALB was incorporated into the vector pAMB-CAT and was linearized at the C-terminal end of the CALB gene by *EcoRI* digestion.

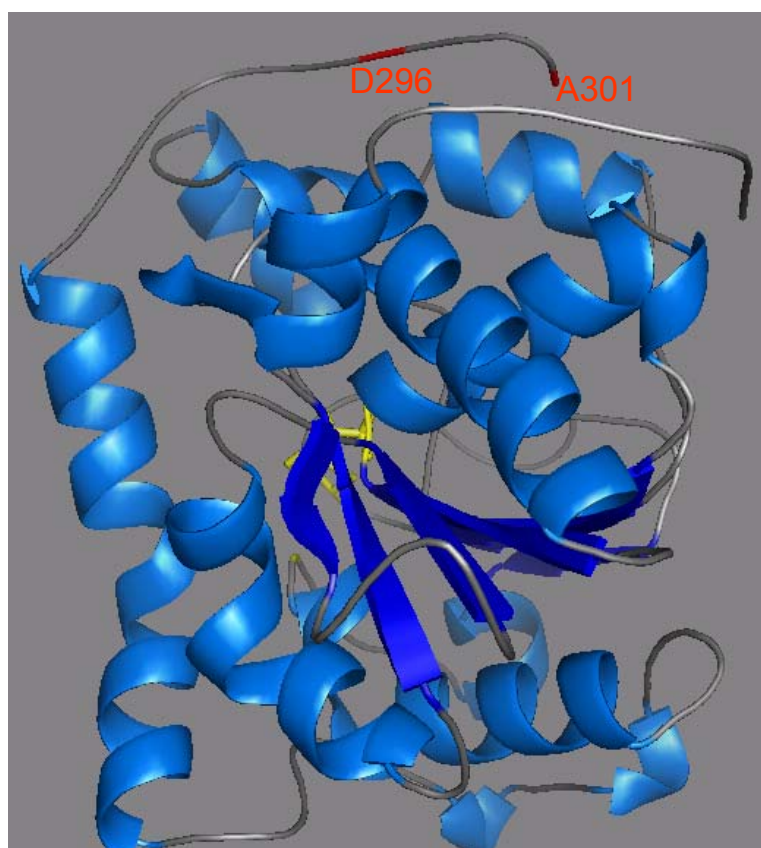


Figure 4-8 Structure model of IT301. The C-terminal residues A301 and D296 of a nonfunctional variant are shown in red.

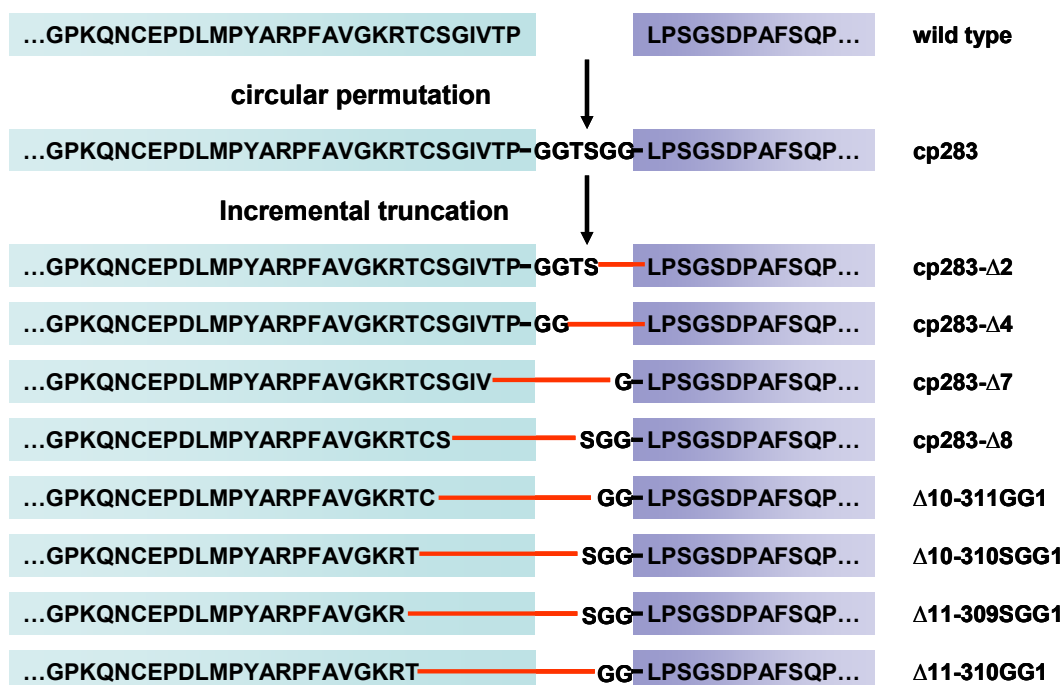


Figure 4-9 Selected cp283 truncation variants for characterization. The sequences of the variants were compared to wt-CALB and cp283. The names of the variants were shown on the right side.

4.3.2.1 Kinetic analysis on selected variants

The hydrolytic activity of the variants on p-nitrophenyl butyrate (p-NB) was determined. As shown in Table 4-1, all truncated cp283 variants show similar activity as cp283, and compared to wt-CALB the rate enhancement is around 6-12 fold, which suggested that truncation did not significantly affect the active site topology of the enzyme. The C-terminal truncation variant IT309 is 3-fold less active than wt-CALB, which is due to a decrease in the k_{cat} value.

4.3.2.2 Circular dichroism analysis on selected variants

Far-UV circular dichroism (CD) spectrum of wt-CALB has negative peaks at 208 nm and 222 nm, characteristic of the α -helical content (Figure 4-10). The spectrum of cp283 has similar shape as that of wt-CALB but with reduced ellipticity at 208, 222 and 190 nm, suggesting a partial loss of the helical content. As shown in Figure 4-10, the CD spectra of cp283- Δ 2, Δ 4 and Δ 7 show a gradually re-gain in ellipticity, and a steady increased in T_m from 40 °C to 45 °C. The increase reached a maximum in cp283- Δ 7, and further truncation rapidly destabilizes the enzyme, possibly due to the strain induced into the structure by excessive truncation. In addition, the abrupt T_m drop in cp283- Δ 8, cp283- Δ 10s and cp283- Δ 11s could be caused by other structural changes, in particular the breakage of the disulfide bond involving C311 and C294. Wild type CALB contains six cysteines which form three disulfide bonds. Excessive truncation might increase the distance between C311 and C294 so that the disulfide bond could not form, or has C311 truncated which eliminates the possibility for the disulfide bond formation. In both cases, the lack of the disulfide bond would result in a decrease in the stability of the enzyme.

Two selected cp283- Δ 10 variants with/without C311 show similar T_m s and CD spectra, suggesting that the particular disulfide bond doesn't form in both variants. The Ellman's assay indicated the existence of free cysteine in variants cp283- Δ 8 and cp283- Δ 10 in comparison to wt-CALB and cp283- Δ 7 (Figure 4-11). However, the Ellman's assay is limited to quantitate unhindered, solvent-accessible free sulfhydryl groups (Petach 1994; King, Elkins et al. 1999). The number of free cysteines in the truncation variants was not able to be quantitated by the assay possibly due to their inadequate solvent accessibility.

4.3.2.3 Gel filtration analysis on selected variants

Combining the results above, the engineering strategy proves to be effective in generating an improved variant cp283- Δ 7 with high activity and increased thermostability. Although we hypothesized that the stability improvement was conferred by the reduced loop flexibility upon truncation, size exclusion chromatography revealed an unexpected dimerization mechanism.

The wild type CALB and cp283 are both monomers. In contrast, size exclusion chromatography of the truncated cp283 revealed a gradually increasing dimer percentage from cp283- Δ 2, Δ 4 to Δ 7 (Figure 4-12). As calculated by peak area integration, the dimer percentage is about 20% in Δ 2 and 50% in Δ 4, while Δ 7 almost completely dimerizes. Further truncation seems to disrupt the dimer interface as cp283- Δ 8 shows to be a complicated mixture of monomer, dimer and higher-order oligomers, while cp283- Δ 10 and cp283- Δ 11 return to the monomeric state.

variants	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\text{min}^{-1}\mu\text{M}^{-1}$)	Relative specificity
Wild type	410 ± 40	305 ± 10	0.74	1.0
cp283	410 ± 60	3251 ± 206	7.93	10.7
IT309	570 ± 80	110 ± 7	0.19	0.3
cp283- $\Delta 2$	375 ± 40	3100 ± 150	8.27	11.2
cp283- $\Delta 4$	160 ± 25	1390 ± 60	8.69	11.7
cp283- $\Delta 7$	240 ± 30	1400 ± 65	5.83	7.9
cp283- $\Delta 8$	293 ± 36	2645 ± 117	9.03	12.2
$\Delta 10$ -311GG1	315 ± 32	1409 ± 52	4.47	6.0
$\Delta 10$ -310SGG1	214 ± 27	1858 ± 68	8.68	11.7
$\Delta 11$ -309SGG1	245 ± 22	1485 ± 42	6.06	8.2
$\Delta 11$ -310GG1	221 ± 28	1243 ± 50	5.62	7.6

Table 4-1 Kinetic data of p-nitrophenyl butyrate hydrolysis for selected variants.
 Relative specificity = k_{cat}/K_M (variant)/ k_{cat}/K_M (wild type)

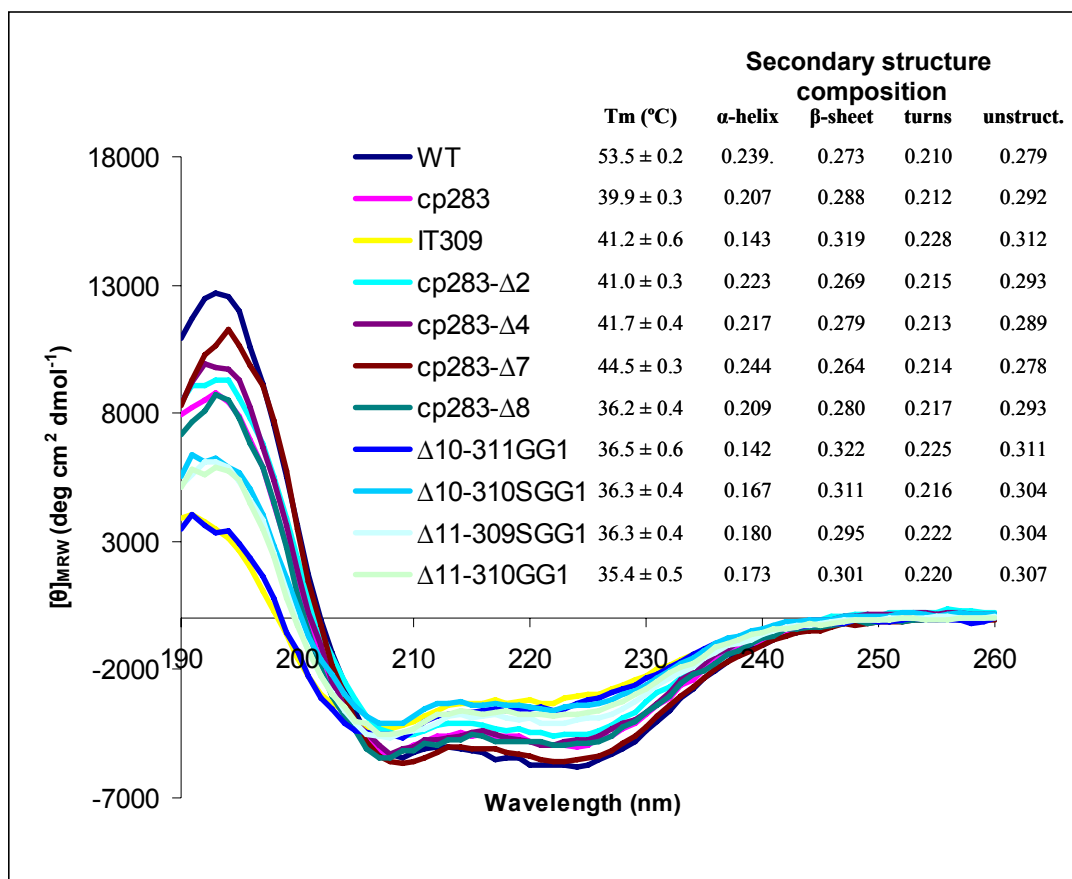


Figure 4-10 Circular dichroism analysis of selected variants. Secondary structure composition was calculated using CDPro (CONTINLL, data set: SMP56).

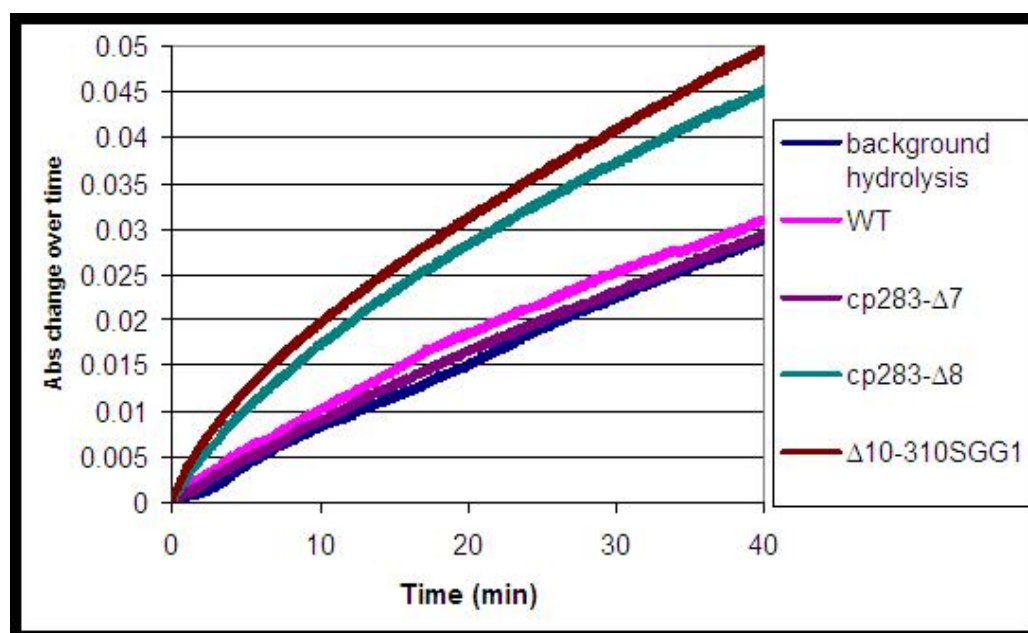
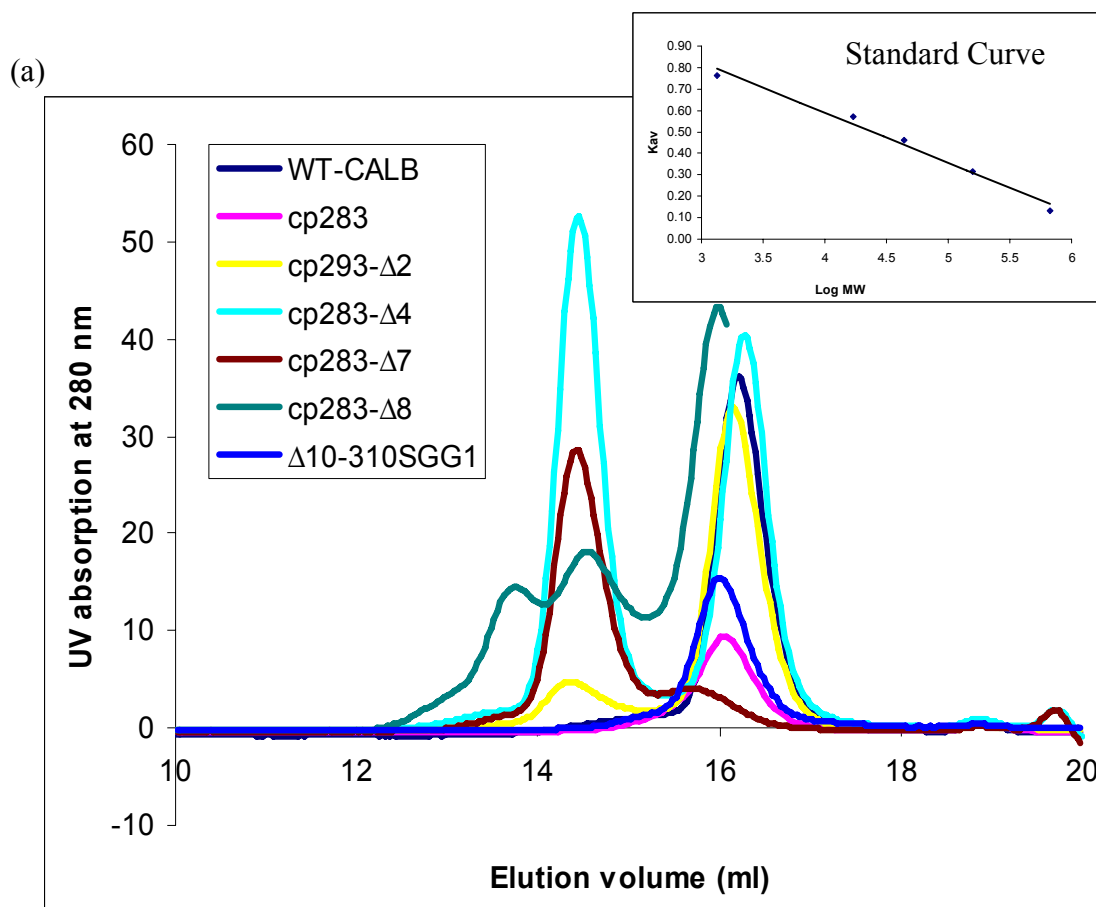


Figure 4-11 The Ellman's assay on selected variants. The absorbance at 412 nm was monitored.



(b)

Elution volume (ml)	Molecular weight (MW)
16	28kD
14.4	68kD
13.7	97kD

Figure 4-12 Gel filtration analysis of selected variants (a) Gel filtration curves of the variants. Standard curve generated by using Bio-Rad's gel filtration standard, curve was fit using the Origin® software (version 7; OriginLab Corporation). (b) Molecular weight of each peak calculated by the standard curve.

4.3.3 Stabilization mechanism of the cp283 truncation variants

To explore whether both oligomeric forms are active and to evaluate their individual stability, the dimer and the monomer fractions of cp283- Δ 4 were separated by gel filtration. After separation, no transition from monomer to dimer or visa versa was detected after incubation at 4 °C for up to seven days, suggesting both forms exist in a stable state. Both fractions show similar activity toward p-NB hydrolysis (Table 4-2), and circular dichroism analysis confirmed that the increased stability is conferred exclusively by the dimeric form of the enzyme, as the monomer has the same T_m as cp283 while the dimer fraction shows a 2 °C increase in T_m (Figure 4-13).

Reinforcement of a higher oligomerization state is known to be one of the mechanisms of thermostabilization (Eijsink, Gaseidnes et al. 2005). An extreme example is the study on the *E. coli lac* repressor. A single mutation (K84L) at the dimer interface increases the thermostability of the protein by 40 °C due to improved hydrophobic packing of the dimer interface (Gerk, Leven et al. 2000). Homo-oligomerization is also a common stabilization strategy used by thermophiles and hyperthermophiles, supported by the evidence that some enzymes from thermophiles are higher-order oligomers than their counterparts in mesophiles (Hess, Kruger et al. 1995; Dams and Jaenicke 1999; Thoma, Hennig et al. 2000). When the dimer association of the phosphoribosylanthranilate isomerase (PRAI) from the hyperthermophile *Thermotoga maritima* is weakened, the resulting monomer differs in crystal structure from that of the parental subunit only in the restructured dimer interface but is far less thermostable than the parental dimer (Thoma, Hennig et al. 2000).

variants	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\text{min}^{-1}\mu\text{M}^{-1}$)	Relative specificity
cp283- $\Delta 4$	160 ± 25	1390 ± 60	8.69	11.7
$\Delta 4$ -monomer	230 ± 22	1671 ± 59	7.27	9.8
$\Delta 4$ -dimer	204 ± 20	2203 ± 72	10.8	14.6

Table 4-2 Kinetic analysis of $\Delta 4$ -monomer and $\Delta 4$ -dimer on p-NB hydrolysis.
Relative specificity = k_{cat}/K_M (variant) / k_{cat}/K_M (wild type)

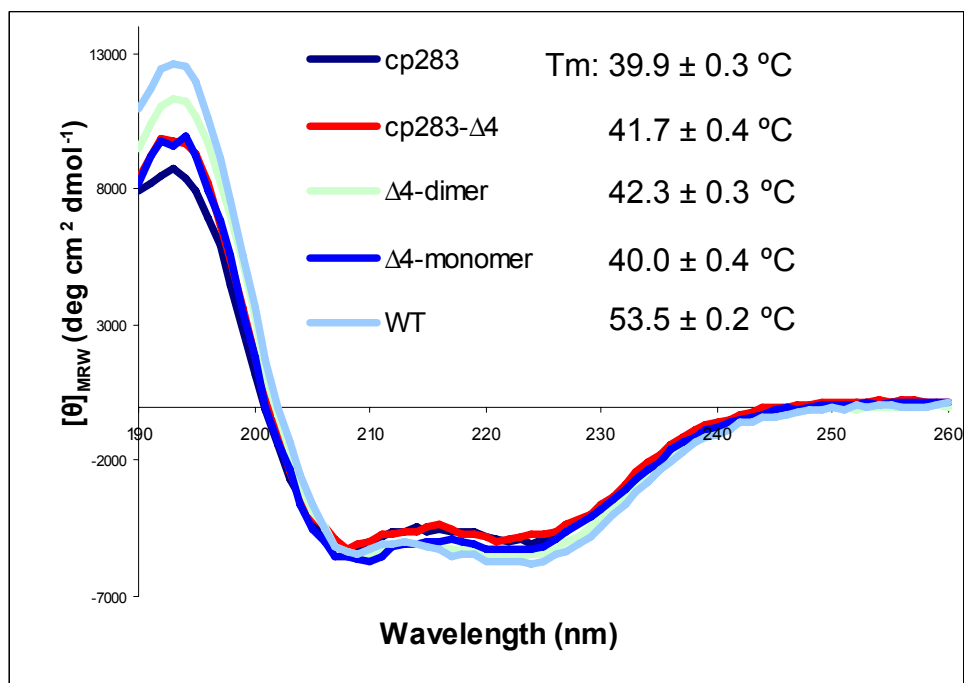


Figure 4-13 Circular dichroism analysis of $\Delta 4$ -monomer and $\Delta 4$ -dimer. Secondary structure composition was calculated using CDPro (CONTINLL, data set: SMP56).

As both wt-CALB and the parental cp283 are monomers, it was hypothesized that the dimer might form by swapping the structural elements between two subunits. The definition of domain swap and its origin will be discussed in detail in the next chapter.

4.4 Conclusions

In summary, the thermostability of cp283 was improved by incrementally truncating the loop formed by connecting the natural termini with a six-amino acid linker during circular permutation. The best variant cp283- Δ 7 has a seven-amino acid truncation in the loop and shows a 5 °C increase in T_m compared to the parental cp283. Thermostabilization of the variants was revealed to be coupled with a change in the enzyme's oligomeric state from monomer to dimer. Further study would involve studying the dimer interface and the dimerization mechanism of the variants, which will be discussed in the next chapter.

Chapter 5 The crystal structure of cp283- Δ 7: circular permutation and domain swapping

5.1 Introduction

In the last three chapters, we have shown step by step that, by circular permutation, a better CALB variant cp283 with greatly improved activity and retained or improved enantioselectivity was obtained. Subsequent incremental truncation partially restored the thermostability of cp283, raising the temperature of unfolding from 40 °C to 45 °C. The increase in thermostability was originally assumed to be conferred by favorable entropy for the closure of a shorter loop in the protein structure as reported by the Regan group (Nagi and Regan 1997). Gel filtration study instead suggested that dimerization of the variants endows them higher stability. The best variant cp283- Δ 7 with a seven-amino acids truncation in the loop was shown to be a dimer by gel filtration, while the wild type CALB and the parental cp283 are monomers (Figure 4-12a, chapter 4).

Compared to the parental cp283, seven amino acids (including two original C-terminal residues T316 and P317, and five linker residues GGTSG) were truncated in cp283- Δ 7 (Figure 5-1).

It was hypothesized that truncation might introduce strain into the protein structure which finally strips the N-terminal portion of cp283- Δ 7 (amino acid 283~315)

from its normal position in one monomer and incorporates it into the equivalent position on another monomer, thereby forming a domain-swapped dimer structure (Figure 5-2).

Three-dimensional domain swapping was defined by Eisenberg and colleagues as substituting of a globular domain (or one or more secondary structure elements) of a protein molecule with an identical portion from a second molecule, with the swapped domain having an environment essentially the same as that of the original domain in a protein monomer (Bennett, Choe et al. 1994; Bennett, Schlunegger et al. 1995). Shortening the hinge loop connecting two fragments of a protein is known to facilitate domain swapping when the closed monomer structure is no longer sterically possible after loop truncation, while a domain-swapped dimer structure restores the interactions between two fragments and is therefore favored (Green, Gittis et al. 1995; Hakansson and Linse 2002; Newcomer 2002; Rousseau, Schymkowitz et al. 2003). One example was the study on staphylococcal nuclease. A six-amino acids deletion in the surface loop of the protein led to the formation of a domain-swapped dimer (Figure 5-3) (Green, Gittis et al. 1995). A similar phenomenon was observed with single-chain antibody fragment Fv, consisting of only the variable light chain (V_L) and variable heavy chain (V_H) domains covalently connected by a polypeptide linker. With a linker length shorter than 15 amino acids, the protein is largely dimeric or forms higher oligomers. In contrast, when the linker is longer than 15 amino acids, the protein becomes dominantly monomeric (Raag and Whitlow 1995).

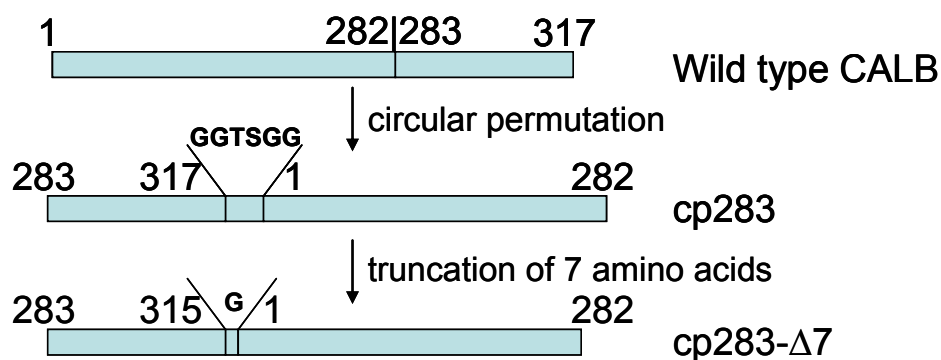


Figure 5-1 Graphic illustration of the sequences of wt-CALB, cp283 and cp283-Δ7. The number corresponds to the amino acid number in wt-CALB.

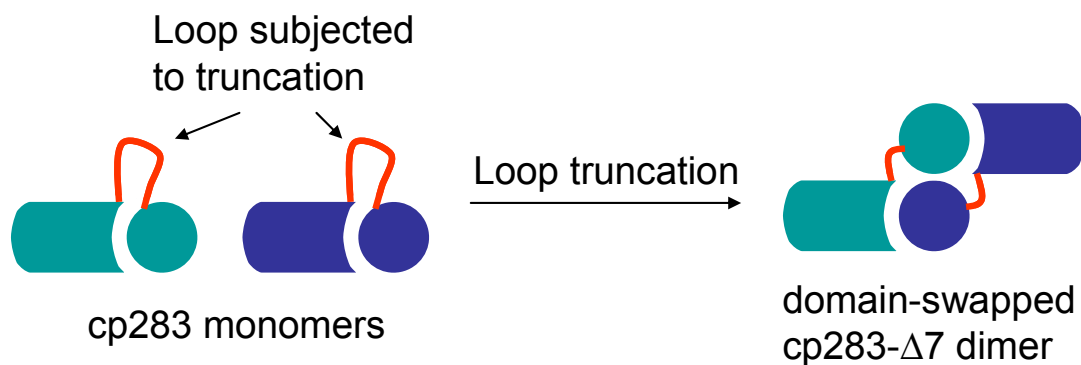


Figure 5-2 Model of cp283-Δ7 dimerization. Loop truncation made it impossible to form a closed monomer, instead the domain-swapped dimer is favored.

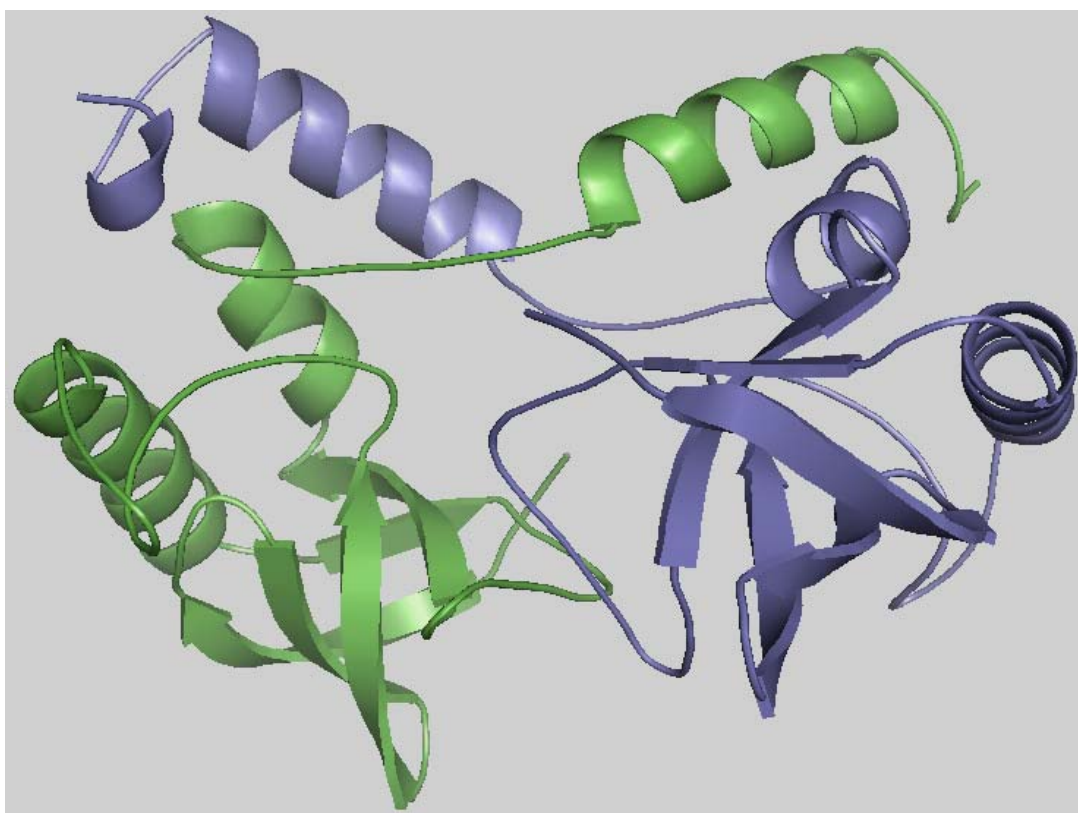


Figure 5-3 Domain-swapped dimer structure of staphylococcal nuclease. Two subunits are shown in green and slate. The C-terminal helices swap places in the dimer structure (Green, Gittis et al. 1995).

To test our hypothesis, the crystal structure of cp283- Δ 7 was solved, revealing an intertwined dimer structure consistent with our model.

In addition, the crystal structure helps to explain previous experimental results on circularly permuted CALB variants. Same as in the parental cp283, the new termini of cp283- Δ 7 dissect the helix α 10 into two fragments. The absence of electron density for those helical fragments indicates their high flexibility. To evaluate their role on catalysis, the two fragments were either separately or simultaneously truncated. The truncation didn't affect the oligomeric state, but did impact on the catalytic performance of the variants.

5.2 Materials and methods

5.2.1 Crystallization of cp283- Δ 7

5.2.1.1 cp283- Δ 7 expression and purification

The variant cp283- Δ 7 was expressed in large quantity by *P. pastoris* fermentation as described in Appendix B5. The protein was purified from supernatant by hydrophobic interaction chromatography and gel filtration as described in Appendix B3. Purified protein was buffer exchanged into the Tris·HCl buffer (pH 7.0, 50 mM) and concentrated to ~25 mg/ml by ultrafiltration (Amicon Ultra-15 centrifugal filter unit; Millipore, Bedford, MA).

5.2.1.2 cp283- Δ 7 crystallization

Crystallization conditions were screened using Crystal Screen B main, C main, and Index (Hampton) (288 different conditions) by the sitting drop method at 16 °C, and

further optimized by adjusting the concentrations of the precipitants with the hanging drop method. Crystals grew in several days and were cryoprotected in mother liquor supplemented with additional 25% ethylene glycol before plunging into liquid nitrogen.

Besides apo-protein, crystals of cp283- Δ 7 in complex with the suicide inhibitor methyl 4-methylumbelliferyl hexylphosphonate were also prepared. Instead of incubating cp283- Δ 7 with the inhibitor before crystallization which yielded crystals with poor quality, the apo-protein crystals were soaked in the inhibitor solution for 1 hour and then cryoprotected as described above.

Compared to the crystallization conditions reported for wt-CALB, the crystallization conditions for cp283- Δ 7 are different (Table 5-1) (Uppenberg, Hansen et al. 1994; Uppenberg, Patkar et al. 1994; Uppenberg, Ohrner et al. 1995).

5.2.1.3 Crystal data collection

The crystal data set was collected by Dr. John R. Horton from Dr. Xiaodong Cheng's lab. The crystals belong to different space groups compared to those of the wt-CALB reported (Table 5-2). The resolutions for apo-protein and protein-inhibitor-complex crystals are 1.5 Å and 1.7 Å respectively. The crystal structure of cp283- Δ 7 was solved by John R. Horton through molecular replacement based on the crystal structure of wt-CALB (PDB code: 1TCA). The structure of cp283- Δ 7/inhibitor complex was solved thereafter in reference to that of the apo-protein. No obvious difference was observed between the structures of apo-protein and protein-inhibitor complex.

Enzyme	Crystallization condition
wt-CALB	20% polyethylene glycol 4000, 50 mM sodium acetate buffer, pH 3.6 and 10% isopropanol.
wt-CALB /Tween 80	1 M ammonium sulfate, 10% dioxane, 0.1 M sodium citrate/citric acid buffer (pH 4.0)
wt-CALB /phosphanate inhibitor	1.3 M ammonium sulfate, 0.1 M sodium citrate/citric acid buffer (pH 4.0), 10% dioxane
cp283- Δ 7	25% PEG 3350, 0.1 M BisTris pH 5.5, 0.2 M NaCl
cp283- Δ 7 /inhibitor	25% PEG 3350, 0.2 M NaH ₂ PO ₄

Table 5-1 Crystallization condition comparison of wt-CALB and cp283- Δ 7 (Uppenberg, Hansen et al. 1994; Uppenberg, Patkar et al. 1994; Uppenberg, Ohrner et al. 1995)

Variant	PDB Code	Space Group	a(Å)	b(Å)	c(Å)	α (°)	β (°)	γ (°)	Resolution (Å)
wt-CALB	1LBS	C2	229.50	95.60	86.80	90	90	90	2.6
	1LBT	P21	95.10	50.20	99.50	90	90.6	90	2.5
	1TCA	P212121	62.10	46.70	92.10	90	90	90	1.5
	1TCB	P21	69.20	50.50	86.70	90	101.5	90	2.1
	1TCC	P21	67.00	50.50	86.70	90	100.1	90	2.5
cp283- Δ 7	/	P3221	111.70	111.70	54.70	90	90	120	1.5, 1.7

Table 5-2 Comparison of crystal parameters

5.2.2 Construction and characterization of the truncation variants

5.2.2.1 Construction of the truncation variants

The truncation variants were PCR amplified using cp283 and cp283-Δ7 as templates and primers cp288-for (5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTGGTCCAAAGCAGAACTGCGAG – 3'), cp265-rev (5'-CGCGCGGCCGCTTAATCATTGGCGGGAAGAGGGTTGC-3'), cp283-for (5'-CGCCTCGAGAAAAGAGAGGCTGAAGCTGCAGCCATCGTGGCGGGTCCA-3'), and pPIC9-rev (5'-GCAAATGGCATTCTGACATCC-3'). The PCR products were digested by *Xho*I and *Not*I, and ligated into the vector pPIC9 digested with the same enzymes. The ligation mixture was used to transform *E. coli* DH5α cells and the sequence of each variant was verified by DNA sequencing. The vector harboring the correct variant gene was linearized by *Sac*I digestion and electroporated into *P. pastoris* strain GS115.

5.2.2.2 Protein expression, purification and activity assays

The expression and purification of all the truncation variants were performed as described in Appendices B1 and B3. The hydrolytic activity of the variants was tested on p-NB and DiFMU octanoate as described in Appendix B4.

5.2.2.3 Circular dichroism analysis of the variants

Far-UV circular dichroism spectra of the variants and their thermal denaturation temperatures were determined as described in Appendix B4

5.3 Results and discussions

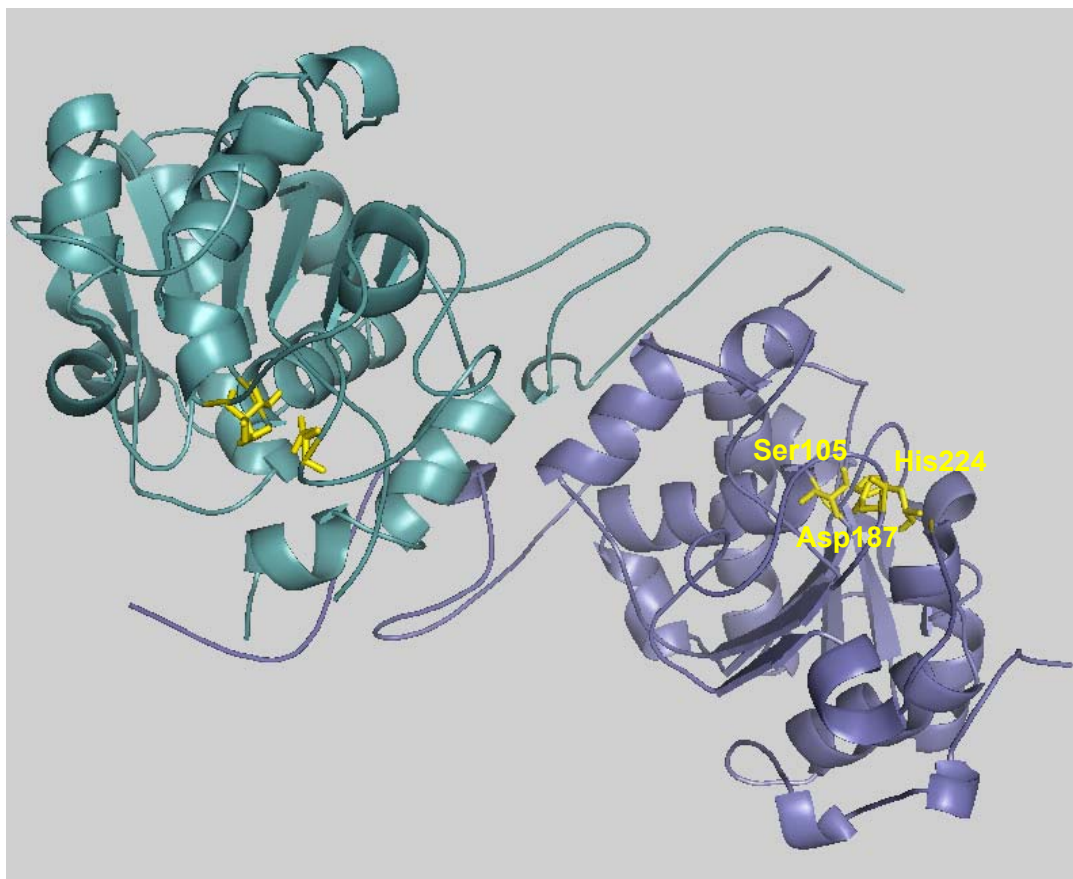
5.3.1 The crystal structure of cp283- Δ 7

X-ray crystallography is a valuable tool for interpreting the experimental data on structure and activity of engineered proteins and in providing explanations for the differences between the variant and the parental protein. From the cp283 incremental truncation library, variants with 2 to 11 amino acid deletions in the loop connecting the natural termini were identified. For variants cp283- Δ 2, cp283- Δ 4 and cp283- Δ 7 with 2 to 7-amino acid truncation, the dimer percentage increases gradually and cp283- Δ 7 completely dimerizes. X-ray crystallography was employed to study the dimer structure and dimerization mechanism. To simplify the protein purification procedure, the variant cp283- Δ 7 was picked for the study

5.3.1.1 Crystal structure of cp283- Δ 7

As shown in Figure 5-4, two subunits in the cp283- Δ 7 dimer are arranged in a head-to-head orientation. As hypothesized, the dimer is formed by domain swapping – the N-terminal portion of one subunit (A283~V315) is incorporated into the equivalent position of the other subunit.

(a)



(b)



Figure 5-4 Crystal structure of the cp283-Δ7 dimer. (a) Two subunits are shown in purple and cyan respectively, and the catalytic triads are shown in yellow. (b) Sequence illustration of cp283-Δ7, swapped region is shown in red.

5.3.1.2 Energetic contributions to domain swapping and thermostabilization

According to previous studies, three dimensional domain swapping by hinge loop truncation would occur when the loop is too tight to close the monomeric conformation and maintain normal interactions between two domains, and the open monomeric form is unstable due to exposure of residues normally buried in the closed interface. Instead the domain-swapped dimer restores the stable domain-interface and is therefore favored (Figure 5-5) (Bennett, Schlunegger et al. 1995; Rousseau, Schymkowitz et al. 2003). The formation of cp283- Δ 7 dimer would probably follow the same pathway.

Considering the higher thermostability of cp283- Δ 7 compared to the parental cp283, the question is what are the factors contribute to stabilization. As mentioned by Eisenberg and colleagues, the monomer and domain-swapped dimer generally have comparable energies as the same interfaces are formed by the swapped domain in its monomeric and dimeric molecules. Small energy differences may come from the conformational and environmental difference for the hinge loop that leads to dimerization, the formation of extra interface between two subunits in the dimer that doesn't exist in monomer structure, and the loss of rotational and translational entropy by dimer formation (Bennett, Schlunegger et al. 1995). In cp283- Δ 7 dimer, an association interface of 1969 Å² is created between two subunits, including not only the original interface that exists in monomeric wt-CALB, but also a secondary interface unique in the dimer structure near amino acid Asn169 (Figure 5-6). Noticeably the Asn169s from two subunits form strong hydrogen bonds with each other. Therefore, favorable interactions including hydrogen bonding and hydrophobic interactions within the secondary interface might afford cp283- Δ 7 higher thermostability.

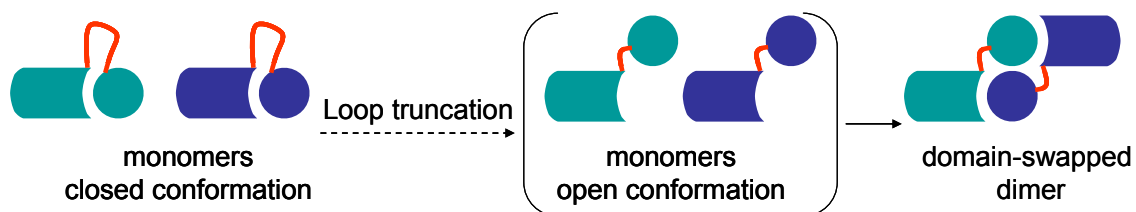


Figure 5-5 Hinge loop truncation that leads to domain swapping. The protein is composed of two domains connected by a loop (Kirk, Bjoerkling et al.). After truncation the loop is too short to form the closed conformation and the open conformation is unstable, and the domain-swapped dimer is favored.

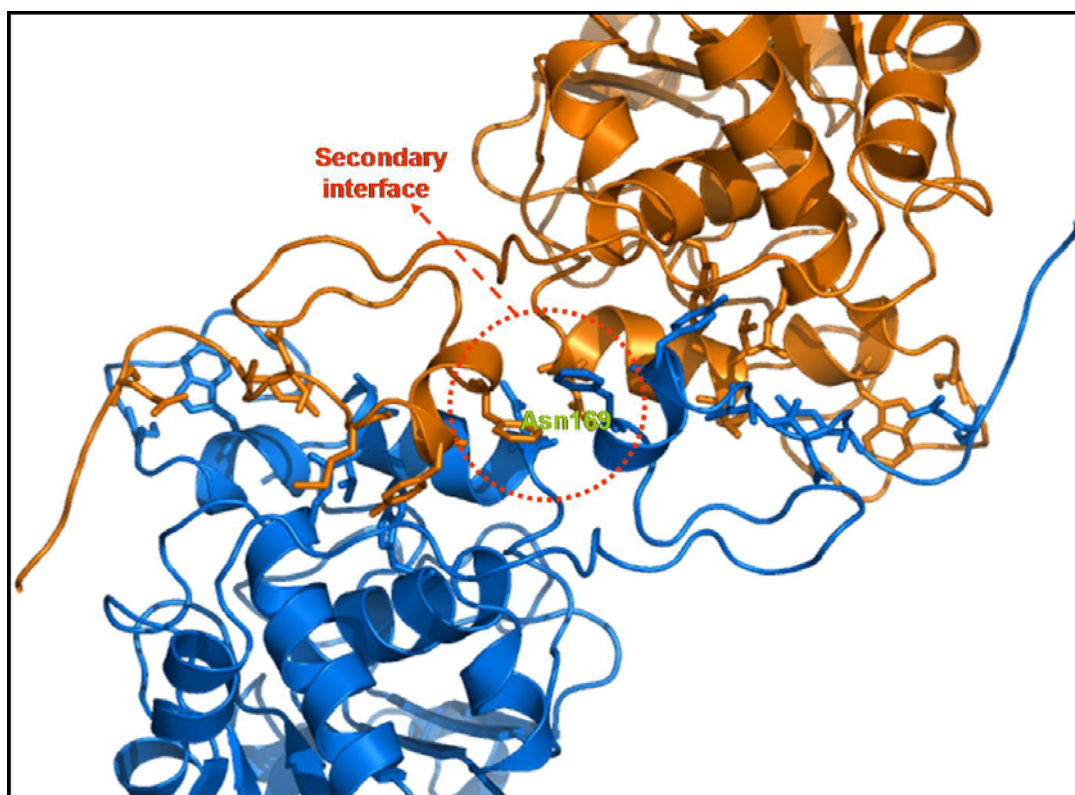


Figure 5-6 Interface between two subunits of the dimer. The secondary interface between two subunits is marked with red circles. Asn169 is labeled in green

5.3.1.3 Comparison of structures of wt-CALB with cp283- Δ 7

When the crystal structure of wt-CALB is aligned with that of cp283- Δ 7, apparently the variant maintains an identical core structure as the wild type enzyme including the central β -sheets and the active site pocket (Figure 5-7). The main difference lies in the structure after amino acid Leu261 in wt-CALB, including helix α 10 and the C-terminus of wt-CALB. There is also a minor shift in position for amino acids Leu1~Gly4 which are at the beginning of the N-terminus in wt-CALB. Besides these differences, the two structures are superimposable between amino acid Ser5 and Leu261.

5.3.1.4 Explanations for experimental results on circularly permuted variants

As an engineering product of the circularly permuted variant cp283, the crystal structure of cp283- Δ 7 would help in understanding the experimental results obtained on cp283 and other circularly permuted CALB variants.

Regardless of the domain-swap, cp283- Δ 7 maintains the overall structure of wt-CALB except for regions around the old termini and new termini, which is consistent with the results obtained on other circularly permuted proteins such as T4 lysozyme, ribonuclease T1, aspartate transcarbamoylase, myoglobin, and DsbA (Zhang, Bertelsen et al. 1993; Mullins, Wesseling et al. 1994; Ni and Schachman 2001; Fishburn, Keefe et al. 2002; Manjasetty, Hennecke et al. 2004; Sagermann, Baase et al. 2004), and is also in good accordance with previous circular dichroism analysis on circularly permuted CALB variants. We would expect that the parental protein cp283 has the same structure as cp283- Δ 7 except for the domain-swap. The similar overall structure would explain the retained enantiomeric preference and high enantioselectivity of cp283 in comparison to

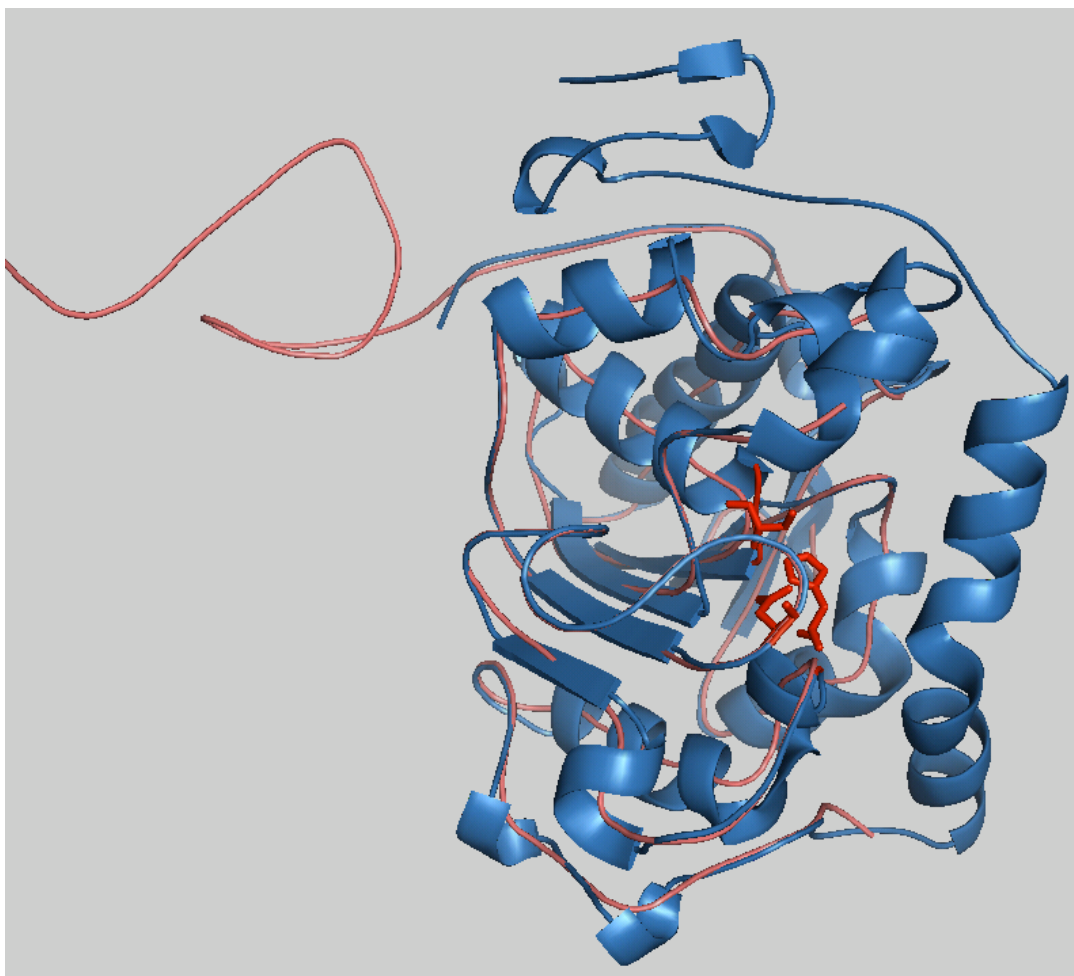


Figure 5-7 Structure alignment of wt-CALB with one subunit of cp283- Δ 7 dimer. The wt-CALB is shown in cartoon in blue, and the cp283- Δ 7 subunit is shown in salmon lines. The catalytic triad is shown as sticks in red.

wt-CALB, as the key element for enantioselectivity – the alignment of the active-site residues – is unchanged.

Noteworthy in the crystal structure of cp283- Δ 7 is the missing of electron density in two regions: the putative lid (amino acid Gly142~Val149) and the new termini (amino acids Leu266~Ala288), indicating their high flexibility. The flexibility of the lid regions residues was also observed by the Uppenberg group when solving the wt-CALB crystal structure (Uppenberg, Hansen et al. 1994). Although no interfacial activation phenomenon was detected, the lid might help in anchoring the acyl chain of the lipid (Martinelle, Holmquist et al. 1995).

The new termini of cp283- Δ 7 dissected the original helix α 10 between Ala282/Ala283. The lack of electron density for residues Leu266~Ala282 and Ala283~Ala288 indicates their high flexibility, which supports our original hypothesis that circular permutation would increase local chain flexibility near new termini. As helix α 10 forms part of the active site binding pocket, we would expect that its flexibility impacts on the catalytic performance of the enzyme. The observed 8 and 141-fold activity increase of cp283- Δ 7 on p-NB and DIFMU-octanoate hydrolysis might result from the increased substrate accessibility to the active site. In addition, helix α 10 is involved in substrate binding in wt-CALB, the high flexibility of new termini generated inside the helix might result in a looser binding and an increased K_M value, as was observed on esterification of 3-hydroxytetrahydrofuran catalyzed by cp283.

5.3.2 Truncation study on cp283- Δ 7 and cp283

5.3.2.1 Creation of the truncation variants

Without electron density in the crystal structure of cp283- Δ 7 for the terminal segments (Leu266~Ala 282, Ala283~Ala288), their secondary structures are unknown. It was hypothesized that deletion of these two fragments from cp283- Δ 7 would have little effect on the overall structure, but would expose the active site and convert a tunnel-shaped active site pocket into a cleft and consequently would impact on the catalytic performance of the enzyme and be favorable for accepting structurally more demanding substrates. To verify the hypothesis and also investigate the role of these two flexible segments on the stability and catalysis of the enzyme, three truncation variants with either one or both fragments deleted were constructed (Table 5-3). The same truncations were introduced in the parental cp283 for comparison.

All six variants were expressed in *P. pastoris* and purified to homogeneity. Their oligomeric states were determined by gel filtration, which showed that truncation doesn't affect their oligomeric state as all the cp283- Δ 7 variants are dimeric and all the cp283 variants are monomeric (Figure 5-8).

Variant	Sequence
$\Delta 7-\Delta 5$	Truncating Ala283~Ala288 from cp283- $\Delta 7$
$\Delta 7-\Delta 17$	Truncating Leu266~Ala 282 from cp283- $\Delta 7$
$\Delta 7-\Delta 22$	Truncating Leu266~Ala 282 and Ala283~Ala288 from cp283- $\Delta 7$
cp283- $\Delta 5$	Truncating Ala283~Ala288 from cp283
cp283- $\Delta 17$	Truncating Leu266~Ala 282 from cp283
cp283- $\Delta 22$	Truncating Leu266~Ala 282 and Ala283~Ala288 from cp283

Table 5-3 Sequences of the truncation variants

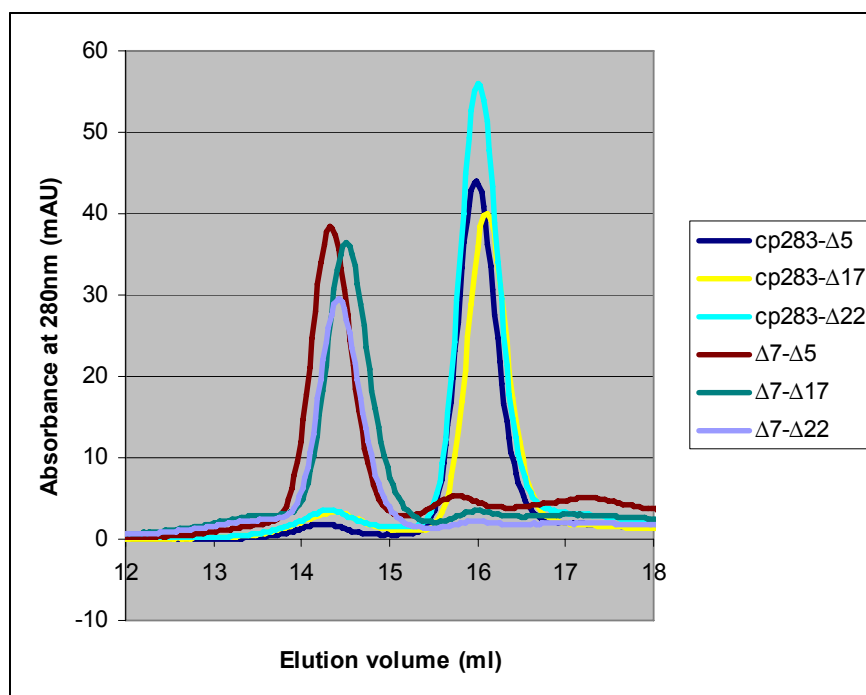


Figure 5-8 Gel filtration analysis of the truncation variants. Molecular weight was calculated using the standard curve in Figure 4-12. Elution volume 16 ml corresponds to a molecular weight of 28kD, Elution volume 14.4 ml corresponds to a molecular weight of 68kD.

5.3.2.2 Kinetic analysis of the truncation variants

The activity of the variants on p-NB and DiFMU octanoate hydrolysis was determined and compared to those of the parental proteins (Table 5-4). The results for cp283 and cp283- Δ 7 truncation variants are consistent in that both variants with the segment Ala283~Ala288 deletion show a similar hydrolytic activity as their parental proteins. The rate enhancement is around 8-fold on p-NB hydrolysis and over 100-fold on DiFMU octanoate hydrolysis. However, truncating the segment Leu266~Ala282 from the protein lowers its hydrolytic activity. All four variants (Δ 7- Δ 17, Δ 7- Δ 22, cp283- Δ 17 and cp283- Δ 22) show an approximately 4-fold activity decrease for p-NB hydrolysis and more than 10-fold decrease in DiFMU octanoate hydrolysis compared to the parental proteins. However, their activities are still higher than the wt-CALB, suggesting the segment Leu266~Ala282 is not indispensable but it facilitates catalysis. In addition, removing all 22 amino acids (Leu266~Ala288) exposes the active site of the enzyme to the solvent, so future study might involve investigating the activity of these variants toward substrates with bulkier or long-chain substituents.

5.3.2.3 Circular dichroism analysis of the truncation variants

Far-UV circular dichroism (CD) spectra of the variants suggest that all the variants fold properly with similar secondary structures (Figure 5-9). Truncating amino acids Ala283~Ala288 doesn't affect the thermal denaturation temperature (T_m) of the variants, and further deletion of amino acids Leu266~Ala282 only leads to a slight decrease in T_m .

Table 5-4 Apparent kinetic constants for truncation variants with p-nitrophenyl butyrate and DiFMU octanoate as substrates

enzyme variant			p-nitrophenyl butyrate				DiFMU octanoate			
Name ^a	sequence ^b	oligmeric state ^d	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1}\text{min}^{-1}$)	rel. speci. _c	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($\mu\text{M}^{-1}\text{min}^{-1}$)	rel. speci. _c
Wild type	L1/P317	monomer	410 ± 40	305 ± 10	0.74	1.0	2.6 ± 0.3	2 ± 0.1	0.8	1.0
cp283- Δ 7	A283/A282	dimer	240 ± 30	1400 ± 65	5.83	7.9	2.2 ± 0.2	249.2 ± 2.9	113.3	141.6
Δ 7- Δ 5	G288/A282	dimer	237 ± 29	1413 ± 55	5.96	8.1	1.7 ± 0.2	180.2 ± 5.3	104.4	130.5
Δ 7- Δ 17	A283/D265	dimer	1007 ± 88	1143 ± 61	1.14	1.5	1.8 ± 0.3	14.1 ± 0.6	7.7	9.7
Δ 7- Δ 22	G288/D265	dimer	548 ± 80	732 ± 52	1.34	1.8	0.8 ± 0.06	6.1 ± 0.1	7.7	9.6
cp283	A283/A282	monomer	410 ± 60	3251 ± 206	7.93	10.7	2.4 ± 0.4	340.2 ± 16.9	139.5	174.4
cp283- Δ 5	G288/A282	monomer	257 ± 38	1614 ± 77	6.28	8.5	3.2 ± 0.7	331.2 ± 27.7	102.3	127.9
cp283- Δ 17	A283/D265	monomer	927 ± 174	1746 ± 187	1.88	2.5	4.6 ± 0.5	49.5 ± 2.5	10.8	13.5
cp283- Δ 22	G288/D265	monomer	806 ± 146	1234 ± 126	1.53	2.1	2.1 ± 0.3	20.2 ± 0.9	9.5	11.9

Note: a. CALB nomenclature: cp283 = circular permuted protein whose N-terminus starts at amino acid 283 of the wild-type sequence. b. N- and C-terminal amino acids (all in single-letter code) are listed. c. Relative specificity = k_{cat}/K_M (variant) / k_{cat}/K_M (wild type). d. Oligmeric state of the enzyme was determined by gel filtration.

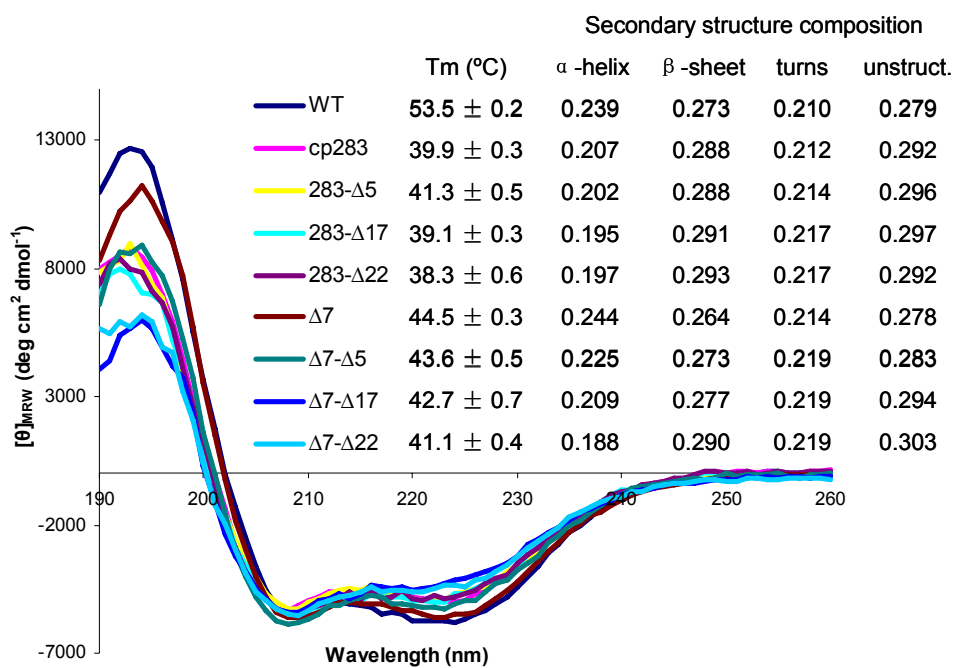


Figure 5-9 Far-UV circular dichroism (CD) spectra and T_m of the truncation variants. Secondary structure composition was calculated using CDPro (Sreerama and Woody 2000) (CONTINLL, data set: SMP56).

5.4 Conclusions

In summary, crystallization of cp283- Δ 7 revealed a domain-swapped dimer structure, which is the first example of a dimeric α/β hydrolase to our knowledge. The crystal structure not only explains the stability endowed by dimerization, but also provides valuable information for interpreting the experimental results obtained on the circularly permuted variants.

Overall, the engineering strategy circular permutation in combination with incremental truncation generated CALB variants with greatly enhanced catalytic efficiency, high enantioselectivity and moderate thermostability, which might turn out to be advantages for their future applications in organic synthesis.

Chapter 6 Conclusions and perspectives

Lipases are ubiquitous and versatile biocatalysts that have broad applications in industry and in organic synthesis. The lipase B from *Candida antarctica* (CALB) has shown to be an excellent candidate in enantiospecific transformations with broad substrate specificity and high enantioselectivity. In recent years, the interest on CALB engineering is increasing with an aim to broaden its application in asymmetric synthesis. In this dissertation the engineering strategy – circular permutation in combination with incremental truncation was applied to CALB engineering, which yielded variants with greatly enhanced catalytic performance.

Circular permutation as a genetic engineering approach has often been used in studying protein folding, but little attention has been paid to its effects on an enzyme's catalytic performance. We hypothesized that internal relocation of a protein's N- and C-termini in or near the active site can increase chain flexibility and active site accessibility, which could translate into higher activity for structurally more demanding substrates. Based on this hypothesis, a random circularly permuted CALB library was generated and subsequent screening identified 63 unique functional variants. Particularly interesting from a catalysis standpoint are permutations in the long helix $\alpha 10$ which flanks the active site pocket, and variants with new termini in the enzyme's putative lid region. Characterization of selected library members by kinetic, circular dichroism, and enantioselectivity analysis demonstrated that relocation of termini has a significant impact on the catalytic performance of CALB, especially for variants with new termini in

helix $\alpha 10$, where the hydrolytic activity increased up to a 174-fold for the test substrate DiFMU-octanoate.

One side effect of circular permutation is the negative impact on the thermostability of the variants, which is also the case for the CALB variants. Therefore, secondary engineering by incremental truncation was performed on the most active circular permutant cp283. An incremental truncation library was generated by truncating an extended loop connecting the natural termini, and library screening identified functional variants with deletion of up to 11 residues which corresponds to ~25% of the entire loop region. The characterization of selected library members showed that the variants maintain the high activity of the parental cp283 and have an up to 5 °C increase in thermo-denaturation temperature. Both experimental evidence and crystallography confirmed that thermostabilization is conferred via dimerization of the protein. The crystal structure of the best variant, cp283- $\Delta 7$, was solved, revealing a dimer structure with its N-terminal portion (Ala283~Val305) swapped between two subunits, which to our knowledge is the first example of a dimeric α/β hydrolase.

Overall, the results from this dissertation have shown that circular permutation could improve the catalytic efficiency of CALB for over 100-fold, and circular permutation in combination with incremental truncation have yielded better CALB variants with high activity, enantioselectivity and moderate thermostability, which would be advantageous for future applications in organic synthesis.

Future study may involve studying the substrate specificity of circularly permuted CALB variants in comparison to wt-CALB. The substrate specificity profile of the

variants would be beneficial for a better understanding about the impact of circular permutation on catalysis, and guide future applications in asymmetric synthesis.

In addition, the absence of electron density for the lid region residues in cp283- Δ 7 indicates their high flexibility. Further studies are required to explore the function of the lid in relation to catalysis. The crystal structure of cp283- Δ 7 also revealed the flexibility of the new termini. Truncation of termini residues exposes the active site and converts the tunnel-shaped active site pocket into a cleft. Future study may involve studying the activity of the truncation variants for substrates with bulky or long-chain substituents.

Finally based on the crystal structure of cp283- Δ 7, the activity and enantioselectivity of the circularly permuted CALB variants could be fine-tuned by rational mutagenesis to fit specific applications.

Appendices

A Materials and media

A1 Chemicals

Antifoam Breox FMT 30	Cognis Co.
Benzophenone	Aldrich
Bromothymol blue	Aldrich
6,8-Difluoro-4-methylumbelliferyl octanoate (DiFMU octanoate)	Molecular Probes (Eugene, OR)
6,8-Difluoro-7-hydroxy-4-methylcoumarin (DiFMU)	Molecular Probes (Eugene, OR)
(±)-Flurbiprofen	Aldrich
Glutaric acid bis-(t-butyl) ester	Synthesized by Christina Fields
(±)-3-Hydroxy tetrahydrofuran	Aldrich
Ion exchange resin Lewatit VP OC 1600	SYBRON Chemicals Inc.
Linalyl acetate	Aldrich
(±)-6-Methyl-5-hepten-2-ol	Aldrich
6-Methyl-5-hepten-2-one	Aldrich

4-Methyl-2-pentanone	Aldrich
(±)- α -Methyl-1-naphthalene methanol	Aldrich
Methyl 4-methylumbelliferyl hexylphosphonate	Synthesized by Christina Fields
4'-Methoxy acetophenone	Aldrich
p-Nitrophenyl butyrate (p-NB)	Sigma (St. Louis, MO)
(±)-Phenylpropionic acid	Aldrich
Tributylin	Aldrich
Vinyl acetate	Aldrich

A2 Enzymes

Alkaline phosphatase	Roche, Indianapolis, IN
<i>DNase</i> I	Roche, Indianapolis, IN
<i>Exonuclease</i> III	Promega, Madison, WI
Klenow polymerase	New England Biolabs
Mung bean nuclease	New England Biolabs
<i>Pfx</i> DNA polymerase	Invitrogen, Carlsbad, CA
Restriction enzymes	New England Biolabs
<i>Taq</i> DNA polymerase	New England Biolabs
T4 DNA ligase	Promega, Madison, WI

T4 DNA polymerase Promega, Madison, WI

A3 Strains and plasmids

E. coli strain DH5 α -E Invitrogen, Carlsbad, CA

Pichia pastoris strain GS115 (his4) Invitrogen, Carlsbad, CA

Pichia pastoris strain KM71 (his4) Invitrogen, Carlsbad, CA

pPIC9 Invitrogen, Carlsbad, CA

pAMB-CAT Ambion, Austin, TX

pET-16b Novagen, Madison, WI

A4 Buffers and Media

BMGY medium 10 g yeast extract, 20 g peptone, 13.4 g yeast nitrogen base, 0.4 mg biotin, 10 ml glycerol, 100 ml 1 M potassium phosphate buffer, pH 6.0 per liter

BMMY medium 10 g yeast extract, 20 g peptone, 13.4 g yeast nitrogen base, 0.4 mg biotin, 5 ml methanol, and 100 ml 1 M potassium phosphate buffer, pH 6.0 per liter

YPG medium/plates 10 g yeast extract, 20 g bacto peptone,
20 g glucose per liter
Add 20 g agar if making plates

MD His⁻ plates 13.4 g yeast nitrogen base, 0.4 mg biotin, 20 g dextrose, 15 g

agar per liter

MM-tributylin plates 13.4 g yeast nitrogen base, 0.4 mg biotin, 5 ml methanol, 10 ml tributyrin, 15 g agar per liter

B General Methods

B1 Protein expression in *P. pastoris*

The overexpression and purification of wild type CALB was performed as previously described (Rotticci-Mulder, Gustavsson et al. 2001). The same protocol was adopted for the isolation of circular permutation variants of CALB. Briefly, the pPIC9 vector containing the lipase gene was linearized by *SacI* digestion and electroporated into *P. pastoris* cells (GS115). Aliquots were plated on MD His⁻ plates and incubated at 30 °C. Colonies appeared on plates after 2 days of incubation. A single colony was picked to inoculate 25 ml BMGY medium and the culture was incubated at 30 °C until it reached an OD₆₀₀ of 2-6. The cells were harvested and resuspended in BMMY medium to an OD₆₀₀ of 1. Protein expression was induced by addition of methanol to a final concentration of 0.5% (v/v) every 24 hours. After 4 days of incubation, the culture medium containing the lipase was separated from the cells by centrifugation (1500 g, 4 °C, 10 min).

B2 Protein purification by Ni-NTA column

The His-tagged CALB and circularly permuted variants were isolated from the clear supernatant via affinity chromatography on Ni-NTA agarose (Qiagen, Valencia,

CA) using 2.5 ml resin per 100 ml supernatant. The column was washed with two column volumes of buffer 1 (20 mM imidazole, 300 mM NaCl, 50 mM NaH₂PO₄, pH 8.0) and enzyme was eluted in two column volumes of buffer 2 (250 mM imidazole, 300 mM NaCl, 50 mM NaH₂PO₄, pH 8.0). All fractions were analyzed by SDS-PAGE and product-containing aliquots were pooled. Purified protein was exchanged into storage buffer (150 mM NaCl, 50 mM K-phosphate, pH 7.0) by ultrafiltration (Amicon Ultra-4 centrifugal filter unit; Millipore, Bedford, MA), and stored at 4 °C. The protein concentration was determined spectrophotometrically at 280 nm ($\epsilon = 3.3 \times 10^4 \text{ M}^{-1}\text{cm}^{-1}$) (Rotticci, Norin et al. 2000).

B3 Protein purification by HIC and gel filtration

Alternatively, hydrophobic interaction chromatography in combination with size exclusion chromatography was employed to purify the protein to homogeneity (Rotticci-Mulder, Gustavsson et al. 2001). For the HIC purification route, the clear culture supernatant was mixed with 2 M (NH₄)₂SO₄ solution and 1 M K-phosphate buffer (pH 7.0) to a final concentration of 1 M and 50 mM respectively. The protein samples were then loaded on a HIC column (7 ml butyl-sepharose 4 resin (AmershamBiosciences, Piscataway, NJ) pre-equilibrated with 1 M (NH₄)₂SO₄, 50 mM K-phosphate buffer (pH 7.0) (buffer 4). The column was rinsed with 4 volumes of buffer 4, followed by a stepwise reduction of (NH₄)₂SO₄ in the phosphate buffer (0.2 M increments, 4 column volumes per step). Lipase activity in the eluant was monitored via p-NB hydrolysis (see below) and fractions containing the desired activity were pooled and concentrated by ultrafiltration (Amicon Ultra-15 centrifugal filter unit; Millipore, Bedford, MA). According to SDS-PAGE, the eluted protein has >85% purity. Further removal of

contaminants was possible by gel filtration on a Superdex-200 10/300 GL column (AmershamBiosciences, Piscataway, NJ), using 50 mM K-phosphate buffer (pH 7.0) containing 150 mM NaCl. SDS-PAGE analysis of the final product showed >95% purity.

B4 Activity assays

Lipase activity was determined by measuring the initial hydrolysis rate of p-NB and DiFMU octanoate at room temperature on a Synergy-HT microtiterplate reader (Bio-Tek Instruments, Winooski, VT). Stock solutions of p-NB (200 mM) and DiFMU octanoate (3 mM) were prepared in DMSO. p-NB hydrolysis over a substrate range of 0 – 1.6 mM was measured in 50 mM K-phosphate buffer (pH 7.5) at 400 nm (ϵ for p-NB = $13260 \text{ M}^{-1} \text{ cm}^{-1}$) (Bender and Marshall 1968). The rate of DiFMU octanoate hydrolysis was determined by measuring the DiFMU formation over a substrate range of 0 – 12 μM in 50 mM K-phosphate buffer (pH 7.0) at an excitation/emission wavelength 360/460 nm. Kinetic constants were calculated by fitting the initial rates to the Michaelis-Menten equation using the Origin® software (version 7; OriginLab Corporation).

B5 *P. pastoris* fermentation

Fermentation was performed in a 5-liter BioFlo 3000 fermenter (New Brunswick Scientific Co., Inc.), based on the Invitrogen fermentation protocol. Breox FMT 30 (Cognis Co.) antifoam was used in the fermentation process when needed. At the beginning the glass fermentation vessel containing half the working volume of fermentation basal salts medium (2.67% of 85% phosphoric acid, 0.093% calcium sulfate, 1.82% potassium sulfate, 1.49% magnesium sulfate-7H₂O, 0.413% potassium hydroxide, and 4% glycerol) was sterilized. Following sterilization, the pH of the

medium was aseptically adjusted to 5.0 with 28% ammonium hydroxide. PTM₁ trace salts solution (0.6% cupric sulfate-5H₂O, 0.008% sodium iodide, 0.3% manganese sulfate-H₂O, 0.02% sodium molybdate-2H₂O, 0.002% boric acid, 0.05% cobalt chloride, 2% zinc chloride, 6.5% ferrous sulfate-7H₂O, 0.02% biotin, and 0.5% sulfuric acid) was aseptically added to a ratio of 4.35 ml PTM₁ trace salts/liter of fermentation basal salts medium. Fermentation conditions for Mut⁺ GS115 transformants were set as following: 30 °C, 500 rpm agitation, DO 30%, pH 5.0. The fermenter was inoculated with 10% of the initial fermentation volume of a culture grown in minimal glycerol medium (1% yeast extract, 2% peptone, 100 mM potassium phosphate, pH 6.0, 1.34% YNB, 0.00004% biotin, and 1% glycerol), and the first stage of the fermentation process was running until all the glycerol in the medium was completely consumed. Then a glycerol fed-batch phase was initiated by feeding 50% glycerol containing 12 ml/L trace salts at a rate of 12 ml/hr/liter initial fermentation volume until the cell wet-weight reached 180 g/liter. The pH of the culture was let to drop to 4.5 by yeast's metabolism and maintained by the addition of 28% ammonium hydroxide. After complete consumption of glycerol, casamino acids were added to 1% of the initial fermenter volume and the culture was induced by initiating a 100% methanol feed containing 12 ml/l trace salts. The fed rate was set at 3 ml/hr/liter initial fermentation volume, and was gradually increased to 9ml/hr/liter and was kept throughout the remainder of the fermentation. During the fermentation process, 5 ml culture samples were collected at different time points for later analysis. After induction for 2 days, the cells were harvested and the supernatant was collected and filtered using Millipore 0.22 µm filter.

B6 Circular dichroism analysis

CD spectra were determined using a JASCO 810 (JASCO, Easton, MD) spectropolarimeter. Purified proteins were exchanged into 50 mM K-phosphate buffer (pH 7.0) by ultrafiltration (Amicon Ultra-4 centrifugal filter unit; Millipore, Bedford, MA) and adjusted to 1 mg/ml as measured by the UV absorbance at 280 nm using $\epsilon = 33000 \text{ M}^{-1}\text{cm}^{-1}$. Far-UV CD spectra (190~260 nm) were recorded at 10 °C with a 0.1 mm cuvette, and the scanning speed was set at 20 nm/min with a response time 4 second and band width 2 nm. Five spectra were acquired and averaged for each sample.

Thermal denaturation was monitored by following the ellipticity at 222 nm at a 1 °C/min heating rate from 10 to 80 °C.

References

- Anderson, E. M., K. M. Larsson, et al. (1998). "One biocatalyst-many applications: the use of *Candida antarctica* B-lipase in organic synthesis." *Biocatalysis and Biotransformation* **16**: 181-204.
- Arai, M., K. Maki, et al. (2003). "Testing the relationship between foldability and the early folding events of dihydrofolate reductase from *Escherichia coli*." *Journal of Molecular Biology* **328**(1): 273-288.
- Ay, J., M. Hahn, et al. (1998). "Crystal structures and properties of de novo circularly permuted 1,3-1,4-beta-glucanases." *Proteins-Structure Function and Genetics* **30**(2): 155-167.
- Baird, G. S., D. A. Zacharias, et al. (1999). "Circular permutation and receptor insertion within green fluorescent proteins." *Proceedings of the National Academy of Sciences of the United States of America* **96**(20): 11241-11246.
- Barrientos, L. G., J. M. Louis, et al. (2002). "Design and initial characterization of a circularly permuted variant of the potent HIV-inactivating protein cyanovirin-N." *Proteins-Structure Function and Genetics* **46**(2): 153-160.
- Beernink, P. T., Y. R. Yang, et al. (2001). "Random circular permutation leading to chain disruption within and near alpha helices in the catalytic chains of aspartate transcarbamoylase: Effects on assembly, stability, and function." *Protein Science* **10**(3): 528-537.
- Bender, M. L. and T. H. Marshall (1968). "The elastase-catalyzed hydrolysis of p-nitrophenyl trimethylacetate." *Journal of the American Chemical Society* **90**(1): 201-207.
- Bennett, M. J., S. Choe, et al. (1994). "Domain Swapping - Entangling Alliances between Proteins." *Proceedings of the National Academy of Sciences of the United States of America* **91**(8): 3127-3131.
- Bennett, M. J., M. P. Schlunegger, et al. (1995). "3d Domain Swapping - a Mechanism for Oligomer Assembly." *Protein Science* **4**(12): 2455-2468.
- Blank, K., J. Morfill, et al. (2006). "Functional expression of *Candida antarctica* lipase B in *Escherichia coli*." *Journal of Biotechnology* **125**(4): 474-483.

- Bornscheuer, U. T. and R. J. Kazlauskas (1999). "Hydrolases in organic synthesis - regio- and stereoselective biotransformations."
- Brzozowski, A. M., U. Derewenda, et al. (1991). "A Model for Interfacial Activation in Lipases from the Structure of a Fungal Lipase-Inhibitor Complex." Nature **351**(6326): 491-494.
- Bulaj, G., R. E. Koehn, et al. (2004). "Alteration of the disulfide-coupled folding pathway of BPTI by circular permutation." Protein Science **13**(5): 1182-1196.
- Carlqvist, P., M. Svedendahl, et al. (2005). "Exploring the active-site of a rationally redesigned lipase for catalysis of Michael-type additions." ChemBiochem **6**(2): 331-336.
- Carrington, D. M., A. Auffret, et al. (1985). "Polypeptide Ligation Occurs during Post-Translational Modification of Concanavalin-A." Nature **313**(5997): 64-67.
- Cheltsov, A. V., M. J. Barber, et al. (2001). "Circular permutation of 5-aminolevulinate synthase - Mapping the polypeptide chain to its function." Journal of Biological Chemistry **276**(22): 19141-19149.
- Cheltsov, A. V., W. C. Guida, et al. (2003). "Circular permutation of 5-aminolevulinate synthase - Effect on folding, conformational stability, and structure." Journal of Biological Chemistry **278**(30): 27945-27955.
- Chodorge, M., L. Fourage, et al. (2005). "Rational strategies for directed evolution of biocatalysts - Application to *Candida antarctica* lipase B (CALB)." Advanced Synthesis & Catalysis **347**(7-8): 1022-1026.
- Chu, V., S. Freitag, et al. (1998). "Thermodynamic and structural consequences of flexible loop deletion by circular permutation in the streptavidin-biotin system." Protein Science **7**(4): 848-859.
- Clementi, C., P. A. Jennings, et al. (2001). "Prediction of folding mechanism for circular-permuted proteins." J. Mol. Biol. **311**: 879-890.
- Cunningham, B. A., J. J. Hemperly, et al. (1979). "Favin Versus Concanavalin-a - Circularly Permuted Amino-Acid Sequences." Proceedings of the National Academy of Sciences of the United States of America **76**(7): 3218-3222.
- Dams, T. and R. Jaenicke (1999). "Stability and folding of dihydrofolate reductase from the hyperthermophilic bacterium *Thermotoga maritima*." Biochemistry **38**(28): 9169-9178.
- Eijsink, V. G. H., S. Gaseidnes, et al. (2005). "Directed evolution of enzyme stability." Biomolecular Engineering **22**(1-3): 21-30.

- Fishburn, A. L., J. R. Keefe, et al. (2002). "A circularly permuted myoglobin possesses a folded structure and ligand binding similar to those of the wild-type protein but with a reduced thermodynamic stability." Biochemistry **41**(44): 13318-13327.
- Gebhard, L. G., V. A. Risso, et al. (2006). "Mapping the distribution of conformational information throughout a protein sequence." Journal of Molecular Biology **358**(1): 280-288.
- Gerk, L. P., O. Leven, et al. (2000). "Strengthening the dimerisation interface of Lac repressor increases its thermostability by 40 deg. C." Journal of Molecular Biology **299**(3): 805-812.
- Goldenberg, D. P. and T. E. Creighton (1983). "Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor." J Mol Biol. **165**: 407-413.
- Gotor-Fernandez, V., E. Busto, et al. (2006). "Candida antarctica lipase B: An ideal biocatalyst for the preparation of nitrogenated organic compounds." Advanced Synthesis & Catalysis **348**(7-8): 797-812.
- Graf, R. and H. K. Schachman (1996). "Random circular permutation of genes and expressed polypeptide chains: Application of the method to the catalytic chains of aspartate transcarbamoylase." Proceedings of the National Academy of Sciences of the United States of America **93**(21): 11591-11596.
- Green, S. M., A. G. Gittis, et al. (1995). "One-Step Evolution of a Dimer from a Monomeric Protein." Nature Structural Biology **2**(9): 746-751.
- Gupta, R., P. Rathi, et al. (2003). "Lipase assays for conventional and molecular screening: an overview." Biotechnol. Appl. Biochem. **37**: 63-71.
- Haeffner, F. and T. Norin (1999). "Molecular modelling of lipase catalysed reactions. Prediction of enantioselectivities." Chemical & Pharmaceutical Bulletin **47**(5): 591-600.
- Hahn, M., K. Piotukh, et al. (1994). "Native-Like in-Vivo Folding of a Circularly Permuted Jellyroll Protein Shown by Crystal-Structure Analysis." Proceedings of the National Academy of Sciences of the United States of America **91**(22): 10417-10421.
- Hakansson, M. and S. Linse (2002). "Protein reconstitution and 3D domain swapping." Current Protein & Peptide Science **3**(6): 629-642.
- Hasan, F., A. A. Shah, et al. (2006). "Industrial applications of microbial lipases." Enzyme and Microbial Technology **39**(2): 235-251.
- Hennecke, J., P. Sebbel, et al. (1999). "Random circular permutation of DsbA reveals segments that are essential for protein folding and stability." Journal of Molecular Biology **286**(4): 1197-1215.

- Hess, D., K. Kruger, et al. (1995). "Dimeric 3-Phosphoglycerate Kinases from Hyperthermophilic Archaea - Cloning, Sequencing and Expression of the 3-Phosphoglycerate Kinase Gene of *Pyrococcus-Woesei* in *Escherichia-Coli* and Characterization of the Protein - Structural and Functional Comparison with the 3-Phosphoglycerate Kinase of *Methanothermus-Fervidus*." European Journal of Biochemistry **233**(1): 227-237.
- Hisano, T., K. I. Kasuya, et al. (2006). "The crystal structure of polyhydroxybutyrate depolymerase from *Penicillium funiculosum* provides insights into the recognition and degradation of biopolyesters." Journal of Molecular Biology **356**(4): 993-1004.
- Hoegh, I., S. Patkar, et al. (1995). "Two lipases from *Candida antarctica*: cloning and expression in *Aspergillus oryzae*." Can. J. Bot. **73**: S869.
- Holmquist, M. (2000). "Alpha/Beta-Hydrolase Fold Enzymes: Structures, Functions and Mechanisms." Current Protein and Peptide Science **1**: 209-235.
- Iwakura, M., T. Nakamura, et al. (2000). "Systematic circular permutation of an entire protein reveals essential folding elements." Nature Structural Biology **7**(7): 580-585.
- Jaeger, K. E., B. W. Dijkstra, et al. (1999). "Bacterial biocatalysts: Molecular biology, three-dimensional structures, and biotechnological applications of lipases." Annual Review of Microbiology **53**: 315-+.
- Johnson, J. L. and F. M. Raushel (1996). "Influence of primary sequence transpositions on the folding pathways of ribonuclease T1." Biochemistry **35**(31): 10223-10233.
- Kaur, J. and R. Sharma (2006). "Directed evolution: An approach to engineer enzymes." Critical Reviews in Biotechnology **26**(3): 165-199.
- Kazlauskas, R. J., A. N. E. Weissfloch, et al. (1991). "A Rule to Predict Which Enantiomer of a Secondary Alcohol Reacts Faster in Reactions Catalyzed by Cholesterol Esterase, Lipase from *Pseudomonas-Cepacia*, and Lipase from *Candida-Rugosa*." Journal of Organic Chemistry **56**(8): 2656-2665.
- King, N. M., K. M. Elkins, et al. (1999). "Reactivity of the invariant cysteine of silver hake parvalbumin (Isoform B) with dithionitrobenzoate (DTNB) and the effect of differing buffer species on reactivity." Journal of Inorganic Biochemistry **76**(3-4): 175-185.
- Kirk, O., F. Bjoerkling, et al. (1992). "Fatty acid specificity in lipase-catalyzed synthesis of glucoside esters." Biocatalysis **6**(2): 127-134.
- Kirk, O. and M. W. Christensen (2002). "Lipases from *Candida antarctica*: Unique biocatalysts from a unique origin." Organic Process Research & Development **6**(4): 446-451.

- Kreitman, R. J., R. K. Puri, et al. (1994). "A circularly permuted recombinant interleukin 4 toxin with increased activity." Proc. Natl. Acad. Sci. **91**: 6889-6893.
- Krishna, S. H. and N. G. Karanth (2002). "Lipases and lipase-catalyzed esterification reactions in nonaqueous media." Catalysis Reviews-Science and Engineering **44**(4): 499-591.
- Li, L. and E. I. Shakhnovich (2001). "Different circular permutations produced different folding nuclei in proteins: A computational study." Journal of Molecular Biology **306**(1): 121-132.
- Lindberg, M. O., E. Haglund, et al. (2006). "Identification of the minimal protein-folding nucleus through loop-entropy perturbations." Proceedings of the National Academy of Sciences of the United States of America **103**(11): 4083-4088.
- Lindqvist, Y. and G. Schtider (1997). "Circular permutation of natural protein sequences: Structural evidence." Curr. Opin. Struct. Biol. **7**: 422-427.
- Liu, D., R. D. Schmid, et al. (2006). "Functional expression of *Candida antarctica* lipase B in the *Escherichia coli* cytoplasm - a screening system for a frequently used biocatalyst." Applied Microbiology and Biotechnology **72**(5): 1024-1032.
- Luger, K., U. Hommel, et al. (1989). "Correct folding of circularly permuted variants of a β -barrel enzyme in vivo." Science **243**: 206-210.
- Lutz, S. (2004). "Engineering lipase B from *Candida antarctica*." Tetrahedron-Asymmetry **15**(18): 2743-2748.
- Macauley-Patrick, S., M. L. Fazenda, et al. (2005). "Heterologous protein production using the *Pichia pastoris* expression system." Yeast **22**(4): 249-270.
- Magnusson, A. (2005). "Rational design of *Candida antarctica* lipase B."
- Magnusson, A. O., J. C. Rotticci-Mulder, et al. (2005). "Creating space for large secondary alcohols by rational redesign of *Candida antarctica* lipase B." ChemBiochem **6**(6): 1051-1056.
- Magnusson, A. O., M. Takwa, et al. (2005). "An S-selective lipase was created by rational redesign and the enantioselectivity increased with temperature." Angewandte Chemie-International Edition **44**(29): 4582-4585.
- Manjasetty, B. A., J. Hennecke, et al. (2004). "Structure of circularly permuted DsbA(Q100T99): preserved global fold and local structural adjustments." Acta Crystallographica Section D-Biological Crystallography **60**: 304-309.
- Martinelle, M., M. Holmquist, et al. (1995). "On the interfacial activation of *Candida antarctica* lipase A and B as compared with *Humicola lanuginosa* lipase." Biochimica et Biophysica Acta **1258**: 272-276.

- Martinelle, M. and K. Hult (1995). "Kinetics of Acyl Transfer-Reactions in Organic Media Catalyzed by Candida-Antarctica Lipase-B." Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology **1251**(2): 191-197.
- Mateo, U., J. M. Palomo, et al. (2006). "Glyoxyl agarose: A fully inert and hydrophilic support for immobilization and high stabilization of proteins." Enzyme and Microbial Technology **39**(2): 274-280.
- Miller, E. J., K. F. Fischer, et al. (2002). "Experimental evaluation of topological parameters determining protein-folding rates." Proceedings of the National Academy of Sciences of the United States of America **99**(16): 10359-10363.
- Mullins, L. S., K. Wesseling, et al. (1994). "Transposition of Protein Sequences - Circular Permutation of Ribonuclease-T1." Journal of the American Chemical Society **116**(13): 5529-5533.
- Nagi, A. D. and L. Regan (1997). "An inverse correlation between loop length and stability in a four-helix-bundle protein." Folding & Design **2**(1): 67-75.
- Nakamura, T. and M. Iwakura (1999). "Circular permutation analysis as a method for distinction of functional elements in the M20 loop of Escherichia coli dihydrofolate reductase." Journal of Biological Chemistry **274**(27): 19041-19047.
- Nardini, M. and B. W. Dijkstra (1999). "alpha/beta hydrolase fold enzymes: the family keeps growing." Current Opinion in Structural Biology **9**(6): 732-737.
- Newcomer, M. E. (2002). "Protein folding and three-dimensional domain swapping: a strained relationship?" Current Opinion in Structural Biology **12**(1): 48-53.
- Ni, X. H. and H. K. Schachman (2001). "In vivo assembly of aspartate transcarbamoylase from fragmented and circularly permuted catalytic polypeptide chains." Protein Science **10**(3): 519-527.
- Ollis, D. L., E. Cheah, et al. (1992). "The Alpha/Beta-Hydrolase Fold." Protein Engineering **5**(3): 197-211.
- Ostermeier, M. and S. J. Benkovic (2001). "Construction of hybrid gene libraries involving the circular permutation of DNA." Biotechnology Letters **23**(4): 303-310.
- Ostermeier, M. and S. Lutz (2003). "The creation of ITCHY hybrid protein libraries." Methods Mol Biol. **231**: 129-141.
- Osuna, J., A. Perez-Blancas, et al. (2002). "Improving a circularly permuted TEM-1 beta-lactamase by directed evolution." Protein Engineering **15**(6): 463-470.
- Ottosson, J., J. C. Rotticci-Mulder, et al. (2001). "Rational design of enantio selective enzymes requires considerations of entropy." Protein Science **10**(9): 1769-1774.

- Pan, T. and O. C. Uhlenbeck (1993). "Circularly Permuted DNA, Rna and Proteins - a Review." Gene **125**(2): 111-114.
- Patkar, S., J. Vind, et al. (1998). "Effect of mutations in *Candida antarctica* B lipase." Chemistry and Physics of Lipids **93**(1-2): 95-101.
- Patkar, S. A., F. Bjoerking, et al. (1993). "Purification of two lipases from *Candida antarctica* and their inhibition by various inhibitors." Indian Journal of Chemistry, Section B: Organic Chemistry Including Medicinal Chemistry **32**(B)(1): 76-80.
- Patkar, S. A., A. Svendsen, et al. (1997). "Effect of mutation in non-consensus sequence Thr-X-Ser-X-Gly of *Candida antarctica* lipase B on lipase specificity, specific activity and thermostability." Journal of Molecular Catalysis B-Enzymatic **3**(1-4): 51-54.
- Petach, H. H. (1994). "Reactivity of Protein Sulfhydryl-Groups with Disulfides." Tetrahedron Letters **35**(10): 1587-1588.
- Pleiss, J., M. Fischer, et al. (1998). "Anatomy of lipase binding sites: the scissile fatty acid binding site." Chemistry and Physics of Lipids **93**(1-2): 67-80.
- Poole, A. M. and R. Ranganathan (2006). "Knowledge-based potentials in protein design." Current Opinion in Structural Biology **16**(4): 508-513.
- Protasova, N. Y., M. L. Kireeva, et al. (1994). "Circularly Permuted Dihydrofolate-Reductase of *Escherichia-Coli* Has Functional-Activity and a Destabilized Tertiary Structure." Protein Engineering **7**(11): 1373-1377.
- Qian, Z. and S. Lutz (2005). "Improving the catalytic activity of *Candida antarctica* lipase B by circular permutation." Journal of the American Chemical Society **127**(39): 13466-13467.
- Raag, R. and M. Whitlow (1995). "Single-Chain Fvs." Faseb Journal **9**(1): 73-80.
- Ribeiro, E. A. and C. H. I. Ramos (2005). "Circular permutation and deletion studies of myoglobin indicate that the correct position of its N-terminus is required for native stability and solubility but not for native-like heme binding and folding." Biochemistry **44**(12): 4699-4709.
- Rojas, A., S. Garcia-Vallve, et al. (1999). "Circular permutations in proteins." Biologia **54**(3): 255-277.
- Romanos, M. (1995). "Advances in the Use of *Pichia-Pastoris* for High-Level Gene-Expression." Current Opinion in Biotechnology **6**(5): 527-533.
- Rotticci-Mulder, J. C. (2003). "Expression and mutagenesis study of *Candida antarctica* lipase B." Dissertation.

- Rotticci-Mulder, J. C., M. Gustavsson, et al. (2001). "Expression in *Pichia pastoris* of *Candida antarctica* lipase B and lipase B fused to a cellulose-binding domain." Protein Expression and Purification **21**(3): 386-392.
- Rotticci, D. (2000). "Understanding and engineering the enantioselectivity of *Candida antarctica* lipase B towards sec-alcohols." Dissertation.
- Rotticci, D., F. Haeffner, et al. (1998). "Molecular recognition of sec-alcohol enantiomers by *Candida antarctica* lipase B." Journal of Molecular Catalysis B-Enzymatic **5**(1-4): 267-272.
- Rotticci, D., T. Norin, et al. (2000). "An active-site titration method for lipases." Biochim Biophys Acta **1483**(1): 132-140.
- Rotticci, D., J. C. Rotticci-Mulder, et al. (2001). "Improved enantioselectivity of a lipase by rational protein engineering." ChemBiochem **2**(10): 766-770.
- Rousseau, F., J. W. H. Schymkowitz, et al. (2003). "The unfolding story of three-dimensional domain swapping." Structure **11**(3): 243-251.
- Russell, R. B. (1995). "Swaposins: circular permutations within genes encoding saposin homologues." Protein sequence motifs.
- Sagermann, M., W. A. Baase, et al. (2004). "Relocation or duplication of the helix a sequence of T4 lysozyme causes only modest changes in structure but can increase or decrease the rate of folding." Biochemistry **43**(5): 1296-1301.
- Schlacher, A., T. Stanzer, et al. (1998). "Detection of a new enzyme for stereoselective hydrolysis of linalyl acetate using simple plate assays for the characterization of cloned esterases from *Burkholderia gladioli*." Journal of Biotechnology **62**(1): 47-54.
- Schrag, J. D. and M. Cygler (1997). "Lipases and alpha/beta hydrolase fold." Lipases, Part A **284**: 85-107.
- Schwartz, T. U., R. Walczak, et al. (2004). "Circular permutation as a tool to reduce surface entropy triggers crystallization of the signal recognition particle receptor beta subunit." Protein Science **13**(10): 2814-2818.
- Sharma, R., Y. Chisti, et al. (2001). "Production, purification, characterization, and applications of lipases." Biotechnology Advances **19**(8): 627-662.
- Sherman, F. (2002). "Getting started with yeast." Methods Enzymol. **350**: 3-41.
- Siddiqui, K. S. and R. Cavicchioli (2005). "Improved thermal stability and activity in the cold-adapted lipase B from *Candida antarctica* following chemical modification with oxidized polysaccharides." Extremophiles **9**(6): 471-476.

- Smith, V. F. and C. R. Matthews (2001). "Testing the role of chain connectivity on the stability and structure of dihydrofolate reductase from E-coli: Fragment complementation and circular permutation reveal stable, alternatively folded forms." Protein Science **10**(1): 116-128.
- Sreerama, N. and R. W. Woody (2000). "Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set." Analytical Biochemistry **287**(2): 252-260.
- Suen, W. C., N. Y. Zhang, et al. (2004). "Improved activity and thermostability of *Candida antarctica* lipase B by DNA family shuffling." Protein Engineering Design & Selection **17**(2): 133-140.
- Svedendahl, M., K. Hult, et al. (2005). "Fast carbon-carbon bond formation by a promiscuous lipase." Journal of the American Chemical Society **127**(51): 17988-17989.
- Svensson, A. K. E., J. A. Zitzewitz, et al. (2006). "The relationship between chain connectivity and domain stability in the equilibrium and kinetic folding mechanisms of dihydrofolate reductase from E.coli." Protein Engineering Design & Selection **19**(4): 175-185.
- Theil, F. (1995). "Lipase-Supported Synthesis of Biologically-Active Compounds." Chemical Reviews **95**(6): 2203-2227.
- Thoma, R., M. Hennig, et al. (2000). "Structure and function of mutationally generated monomers of dimeric phosphoribosylanthranilate isomerase from *Thermotoga maritima*." Structure with Folding & Design **8**(3): 265-276.
- Todd, A. E., C. A. Orengo, et al. (2002). "Plasticity of enzyme active sites." Trends in Biochemical Sciences **27**(8): 419-426.
- Topell, S., J. Hennecke, et al. (1999). "Circularly permuted variants of the green fluorescent protein." Febs Letters **457**(2): 283-289.
- Tougaard, P., T. Bizebard, et al. (2002). "Structure of a circularly permuted phosphoglycerate kinase." Acta Crystallographica Section D-Biological Crystallography **58**: 2018-2023.
- Uliel, S., A. Fliess, et al. (2001). "Naturally occurring circular permutations in proteins." Protein Engineering **14**(8): 533-542.
- Uppenberg, J., M. T. Hansen, et al. (1994). "Sequence, Crystal-Structure Determination and Refinement of 2 Crystal Forms of Lipase-B from *Candida-Antarctica*." Structure **2**(4): 293-308.

- Uppenberg, J., N. Ohrner, et al. (1995). "Crystallographic and molecular-modeling studies of lipase B from *Candida antarctica* reveal a stereospecificity pocket for secondary alcohols." Biochemistry **34**(51): 16838-16851.
- Uppenberg, J., S. Patkar, et al. (1994). "Crystallization and Preliminary-X-Ray Studies of Lipase-B from *Candida-Antarctica*." Journal of Molecular Biology **235**(2): 790-792.
- Uversky, V. N., V. P. Kutysenko, et al. (1996). "Circularly permuted dihydrofolate reductase possesses all the properties of the molten globule state, but can resume functional tertiary structure by interaction with its ligands." Protein Science **5**(9): 1844-1851.
- Verger, R. (1997). "'Interfacial activation' of lipases: Facts and artifacts." Trends in Biotechnology **15**(1): 32-38.
- Viguera, A. R. and L. Serrano (1997). "Loop length, intramolecular diffusion and protein folding." Nature Structural Biology **4**(11): 939-946.
- Weikl, T. R. and K. A. Dill (2003). "Folding kinetics of two-state proteins: Effect of circularization, permutation, and crosslinks." Journal of Molecular Biology **332**(4): 953-963.
- Weiner, J. and E. Bornberg-Bauer (2006). "Evolution of circular permutations in multidomain proteins." Molecular Biology and Evolution **23**(4): 734-743.
- Wu, S. X. and G. J. Letchworth (2004). "High efficiency transformation by electroporation of *Pichia pastoris* pretreated with lithium acetate and dithiothreitol." Biotechniques **36**(1): 152-154.
- Zhang, N. Y., W. C. Suen, et al. (2003). "Improving tolerance of *Candida antarctica* lipase B towards irreversible thermal inactivation through directed evolution." Protein Engineering **16**(8): 599-605.
- Zhang, P. H. and H. K. Schachman (1996). "In vivo formation of allosteric aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains: Implications for protein folding and assembly." Protein Science **5**(7): 1290-1300.
- Zhang, T., E. Bertelsen, et al. (1993). "Circular Permutation of T4-Lysozyme." Biochemistry **32**(46): 12311-12318.