박 성 순 교수지도

석사학위 청구논문

Substrate and protein engineering of *Candida antarctica* lipase B for improving enantioselectivity toward alkyl tetrahydrofuran-2-carboxylates

기질 및 단백질 변형을 통한 *Candida antarctica* lipase B 의 입체선택성 증가에 대한 연구

2011

성신여자대학교 대학원

화 학 과

박 지 민

Substrate and protein engineering of *Candida antarctica* lipase B for improving enantioselectivity toward alkyl tetrahydrofuran-2-carboxylates

기질 및 단백질 변형을 통한 *Candida antarctica* lipase B의
입체선택성 증가에 대한 연구

박 성 순 교수지도

이 논문을 석사학위 논문으로 제출함

2011년 11월

성신여자대학교 대학원
화 학 과
박 지 민

인 준 서

박 지 민의 석사학위논문을 인준함

심사위원 _____인

심사위원 _____인

심사위원 _____인

성신여자대학교 대학원

논문개요

효소는 높은 입체 선택성과 위치 선택성이 있는 효율적인 촉매이다. 또한 광범위한 기질 특이성을 가지므로 다양한 종류의 반응에 응용될 수 있으며, 반응 조건이 온화하고 환경 친화적인 장점을 가지고 있다. 특히 지방 가수분해 효소인 리파아제는 동력학적 분리 (kinetic resolution)를 통해 의학 산업에 핵심요인이 되는 광학적으로 순수한 화합물의 제조에 활용 가능하므로, 유기합성 촉매로 가장 많이 사용되고 있다.

효소 재조합 기술은 높은 반응성과 선택성을 얻기 위하거나 chemocatalytic 반응에 사용하기 위한 기술이었다. 최근에는 라세믹 혼합물로부터 광학적으로 순수한 화합물을 얻기 위한 생촉매 역할까지 확장되어 사용되고 있다. 화학적 혹은 분자생물학적 접근방법을 통해 효소를 재조합 하는 방법은 기존 효소의 입체 선택성을 향상시키거나 전환시키는데 중요한 역할을 한다.

광학적으로 순수한 tetrahydrofuran-2-carboxylate (THFC)는 의약 산업 및 연구에 있어서 유용한 중간체이다. 특히 (*R*)-THFC는 임상적으로 효과적인 비천연 *β*-lactam 항생제 furopenem을 만드는 출발물질이다. 하지만, 전통적인 합성방법을 통해 광학적으로 순수한 THFC 얻는 효과적인 방법은 아직 잘 알려져 있지 않다.

본 논문에서는 효소의 kinetic resolution을 통해 광학적으로 순수한 THFC을 얻는 연구를 다루었다. *Candida antarctica*에서 추출한 lipase B (CAL-B)를 효소로 사용하여 다양한 THFC 에스터 화합물을 대상으로 동력학적 분리를 적용하였다. Wild-type의 CAL-B는 THFC에 대해 낮은 입체선택성을 보이기 때문에 이를 향상시키기 위해 기질과 효소를 재조합하였다. 기질 조작을 통하여 THFC 에스터의 알코올 부분의 알킬기를 길게 만들고 효소 재조합을 통하여 active site 주변의 아

미노산들을 큰 아미노산으로 변형시켜 THFC의 회전에 제한을 주었다. 에스터의 알코올 부분의 알킬기가 에탄올에서 부탄올, 이소부탄올, 이소펜탄올로 길어지고 입체적으로 커짐으로서 wild-type의 (*S*)-입체선호도가 (*R*)-입체선호도로 전환되었고 변형체들 또한 향상된 입체선택성을 보였다. 특히 I189Q (*E* = 13.5)와 V190I (*E* = 11.3)와 같은 변형체들의 입체선택성은 wild-type (*E* = 5.6)보다 2.5배 향상되었다.

CONTENTS

LIST OF TABLE

CHAPTER 1

CHAPTER 2

CHAPTER 3

LIST OF FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 1. Introduction

Enzymes are natural catalysts, performing chemical transformation on organic compounds. They are used for thousands of years in a variety of biotechnological processes and for many commercial purposes, such as the dairy, textile, chemical, detergent, food, pharmaceutical, and brewing industries.

Utilization of an enzyme in organic synthesis contains both advantages and disadvantages (Table 1).[1)] Enzymes are very efficient catalysts. Compare to nonenzymatic reactions, they accelerate reaction by a factor of 10^8-10^{10}. And enzyme can catalyze a broad spectrum of reactions. Enzymes also use often more environmental friendly conditions such as lower reaction temperatures, no need of protecting groups, less waste produced and less energy input required. They also show high selectivity including chemoselectivity, regioselectivity, diastereoselectiviy and enantioselectivity. On the other hand, the disadvantages of using enzymes are exist. High costs, poor stability, product inhibition, need of cofactors and especially low reaction rate in organic solvents.

Table 1. Advantages and disadvantages of using enzymes[1].

Advantage	Disadvantage
• Enzymes are very efficient catalysts • Enzymes are environmentally acceptable. • Enzymes act under mild conditions. • Enzymes are compatible with each other. • Enzymes can catalyze a broad spectrum of reactions. • Enzyme display three major types of selectivities. - Chemoselectivity - Regioselectivity and diastereoselectiviy - Enantioselectivity	• Enzymes are provided by nature in only one enantiomeric form. • Enzymes require narrow operation parameters. • Enzymes display their highest catalytic activity in water. • Enzymes are bound to their natural cofactors. • Enzymes are prone to inhibition phenomena. • Enzymes may cause allergies.

To overcome these disadvantages the system of enzymatic reaction needs to be modified.[2] Modification of enzymes can improve the stability of the wild-type enzyme and the performance towards non-natural reactions, or alter the natural characteristics. This modification is called protein engineering and it can be divided into rational design and directed evolution. When modification happens to substrate, it is called as substrate engineering. Substrate engineering are exploited to improve the enzyme activity and enantioselectivity. And it also influences substrate acceptance and the regioselectivity and stereoselectivity.

In this thesis, substrate and protein engineering were studied to improve the enantioselectivity of *Candida antarctica* lipase B (CAL-B) toward ester of tetrahydrofuran-2-carboxylates. In substrate engineering, the alcohol part of the ester was substituted and the influence was studied. Altering residue near the active site of CAL-B was employed to enhance the enantioselectivity.

1.1. Lipases (EC 3.1.1.3)

Lipases (EC 3.1.1.3) are glycerol ester hydrolases that catalyze the hydrolysis of triacylglycerols. They found in a wide range of organisms including animals, plants, fungi, and bacteria. They belong to the family of serine hydrolases and α/β-fold family. They contain the catalytic triad, Asp/Glu-His-Ser, which is similar to that of serine proteases.[3)]

Lipases are stable in organic solvents, do not require cofactors, possess a broad substrate specificity, and exhibit a high enantioselectivity. Because of these reasons, lipases play an important role in biotechnology. Therefore, lipases are employed a variety of applications including lipid hydrolyzers in laundry industry, flavor production in the dairy industry, pitch control in the pulp and paper industry, lipid biotransformation in the oleochemical industry, and biocatalysis in synthetic reaction.[4)]

Lipases are often classified under esterases because lipases and esterases both catalyze hydrolysis of esters.[5)] However, if they are classified based on substrate preference, esterases usually accept water-soluble esters of short-chain fatty acid triglycerides like tributyrin. But, lipases prefer water-insoluble esters or triglycerides composed of long chain fatty acids. Lipases also can catalyze the reverse reaction of hydrolysis and other reactions such as transesterification, aminolysis, and lactone ring-opening reaction (Figure 1).

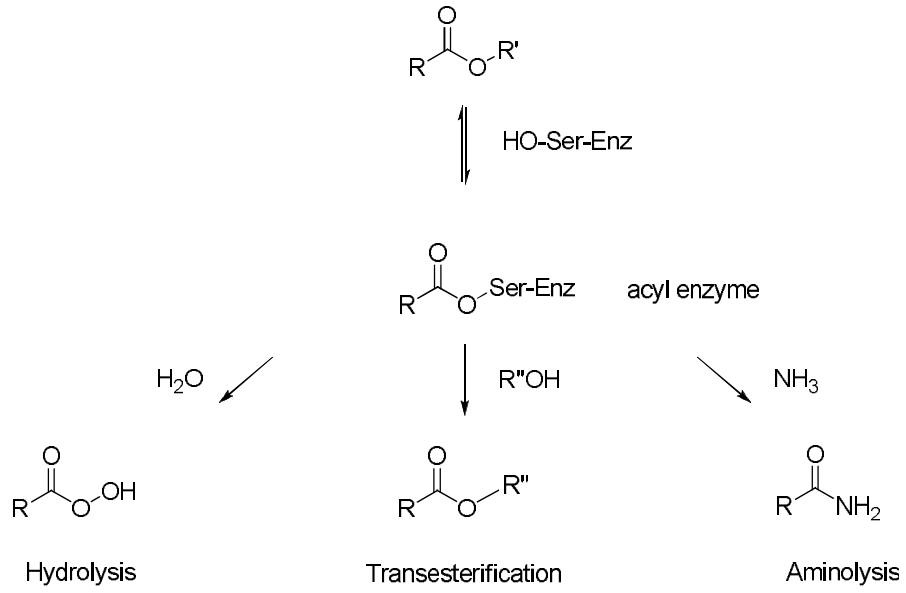


Figure 1. Reactions catalyzed by lipases or esterases

An interesting characteristic of lipases is that they have a lid that covering the active site, and require a water-oil interface in reaction media (*Candida antarctica* lipase B shows an exception because it contains small or no lid). Since the active site of most lipases is buried underneath a lid, their conformations need to be changed from closed to open to perform productive reaction.[6] Without the conformational change, the lid prevents substrate from the active site. When lipase exposures to a lipid interface, it changes a conformation of the lid and turns to the open conformation. Then the catalytic machinery is exposed to the substrate. This process is called as the interfacial activation of lipases (Figure 2).[7] And it explains why lipases typically show better activity towards in soluble substrates in aqueous media.

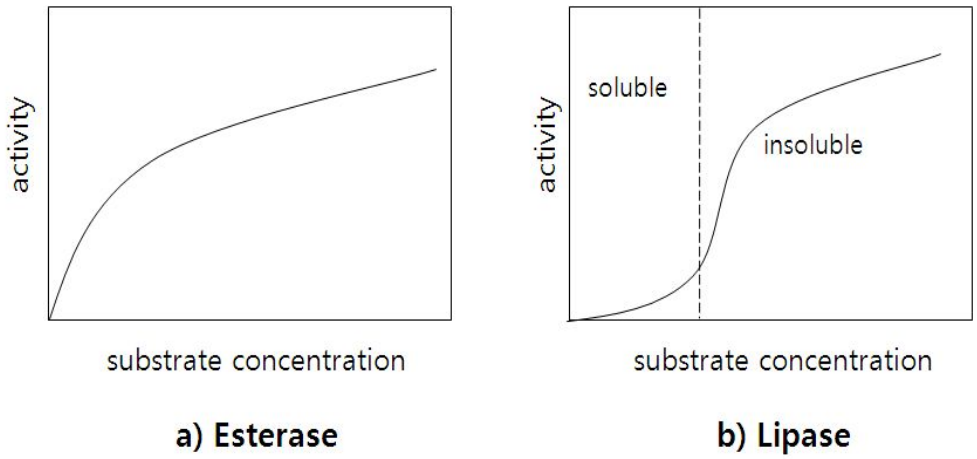


Figure 2. a) Esterase kinetics following normal Michaelis-Menten kinetics and b) Lipase kinetics where interfacial activation is observed.

Lipases are classified by protein sequence alignments, which is also consistent with the 3-D structures of lipases and is more reliable (Table 2).[8] The most useful lipases for organic synthesis are poricine pancreatic lipase (PPL), lipase from *Bukholeria cepacia* (BCL), lipase from *Candida rugosa* (CRL), and lipase B from *Candida antarctica* (CAL-B).

Table 2. Classification of commercial lipases according to similarities in protein sequence.

Classification	Characteristics	Examples
Mammalian (pancreatic) lipases	50 kDa	PPL
Fungal lipases		
Candida rugosa family	60-65 kDa	CRL, GCL, CE
Rhizomucor family	30-35 kDa	CAL-B, RML, ROL, HLL, PcamL
Unclassified		ANL, CAL-A, CLL
Bacterial lipases		
Pseudomonas family	30-35 kDa	PCL, PFL, CVL
Staphylococcus family	40-45 kDa	BTL2

abbreviation: PPL: porcine pancreas lipase; CRL: *Candida rugosa* lipase; GCL: *Geotrichum candidum* lipase; CE: pancreatic cholesterol esterase; CAL-B: *Candida antarctica* lipase B; RML: *Rhizomucor miehei* lipase: ROL: *Rhizopus oryzae* lipase; HLL: *Humicola lanuginosa* lipase; PcamL: *Penicillium camembertii* lipase; ANL: *Aspergillus niger* lipase; CAL-A: *Candida antarctica* lipase A; CLL: Candida lipolytica lipase; PCL: Pseudomonas cepacia lipase; PFL: *Pseudomonas fluorescens* lipase; CVL: *Chromobacterium viscosum* lipase; BTL2: *Bacillus thermocatenulatus* lipase II

1.1.1. Lipase from *Candida antarctica* (CAL-B)

Candida antarctica lipase B (CAL-B) is one of the most used lipase in industry. CAL-B has high thermal stability and activity in a wide range of nonaqueous solvents. And also shows high activity and enantioselectivity toward a wide range of alcohols, while its enantioselectivity toward carboxylic acids is usually low.[9] It is a fungal lipase with 317 amino acid length and 33 kDa weight. It also belongs to α/β-hydrolase fold hydrolase family. The interfacial activation, a common feature of lipases, is not shown by CAL-B. CAL-B is active toward water-soluble substrate and also active toward water-insoluble substrates. In 1994, the structure of CAL-B was determined by X-ray crystallography (Figure 3.)[10].

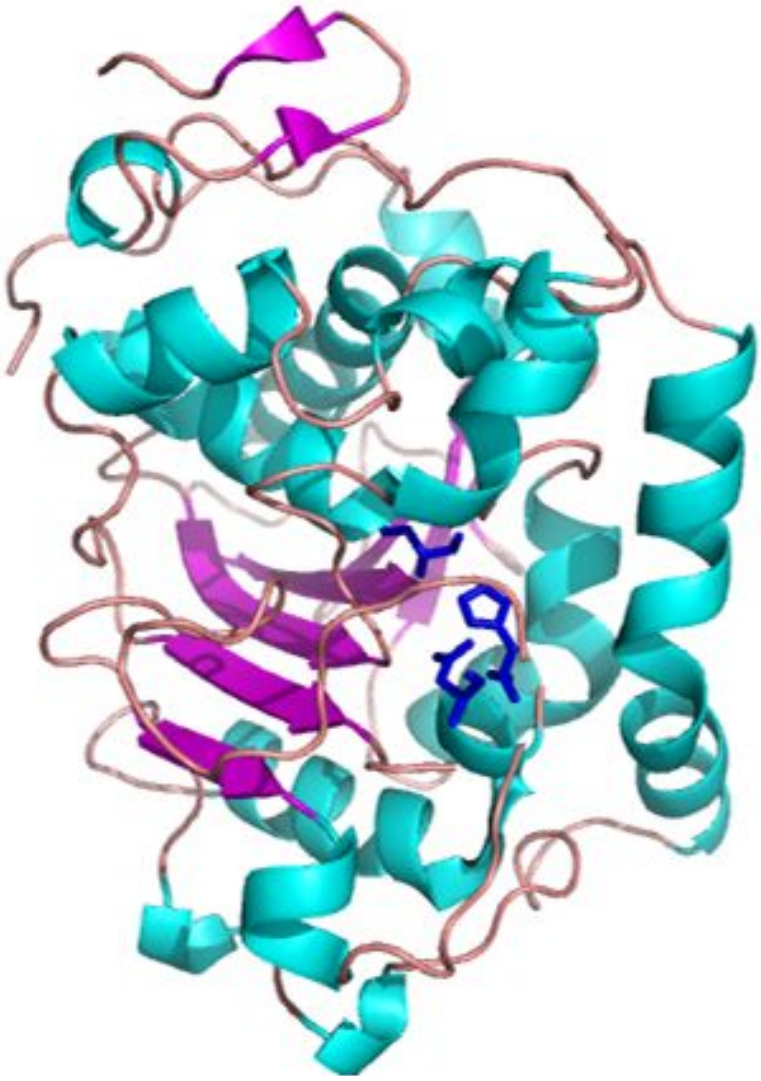


Figure 3. Structure of *Candida antarctica* lipase B (1TCA). The catalytic triad consists of Ser105, Asp187, and His224 and is represented by sticks.

The catalytic triad of CAL-B (*Candida antarctica* lipase B) is Ser105-His224-Asp187. The active site is buried at the bottom of the tunnel-shaped binding site with the oxyanion hold. They are responsible for the catalytic action with the bi-bi mechanism (Figure 4). The mechanism can be divided into two steps by a covalent acyl-enzyme intermediate.[9] The oxyanion hole is composed of the backbone amide proton of Gln106 and the backbone amide proton and the side chain hydroxyl proton of Thr40. The protons stabilize the oxyanion at the transition state by hydrogen bond. A charge relay system involving the catalytic triad of residues forms the basis of the mechanism. First, the serine attacks the ester to generate the first tetrahedral intermediate (T_d1). An acyl enzyme intermediate is formed from the release of the alcohol and then the attack by second substrate forms the second tetrahedral intermediate (T_d2). Release of the second product regenerates the free enzyme. Alternatively, another nucleophile such as an alcohol can attack the acyl enzyme and thereby yielding a new ester (a transesterification reaction).

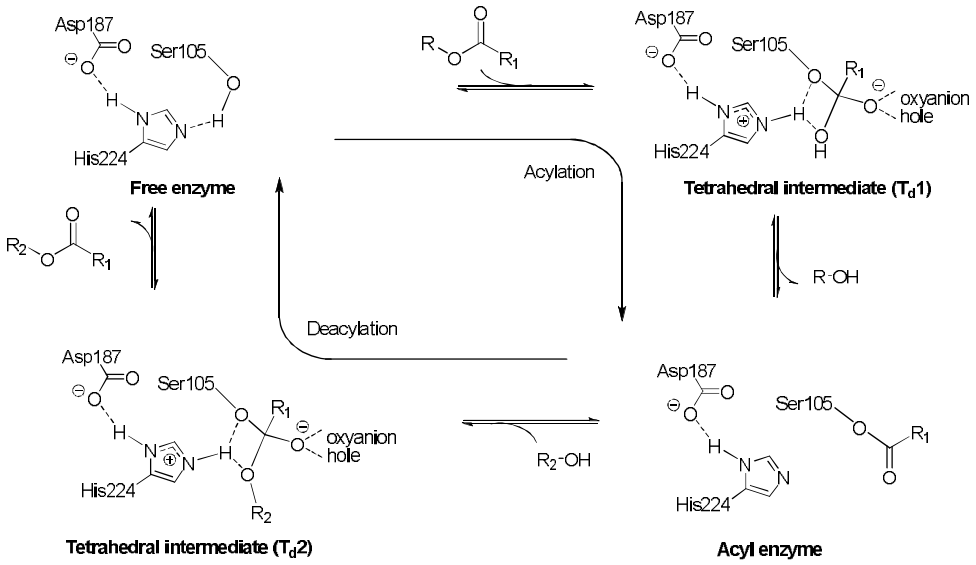


Figure 4. Catalytic mechanism of a *Candida antarctica* lipase B catalyzed hydrolysis or transesterification.

1.2. Protein expeession system

Protein expression systems are widely used in the life science, biotechnology, and medical research. It is one of the subcomponents of gene expression. Over the past decades a variety of hosts and vector systems for recombinant protein expression have been reported. Typical expression vectors contain promoters that direct synthesis of the large amount of mRNA corresponding to the gene. They may also include sequences that encode genetic traits allowing cells containing the vectors to be selected, and sequences that increase the efficiency with which the mRNA is translated. From expression systems enormous number of enzymes and proteins have been used for molecular biological research; DNA polymerase for PCR, reverse transcriptase for RNA analysis, and restriction endonucleases for cloning.

Six different expression systems have been used for manufacture of biopharmaceutics: bacteria, yeast, insect cells, mammalian cells, transgenic animals, and transgenic plants. Researchers commonly select mammalian, insect, yeast, and prokaryotic hosts (Figure 5).

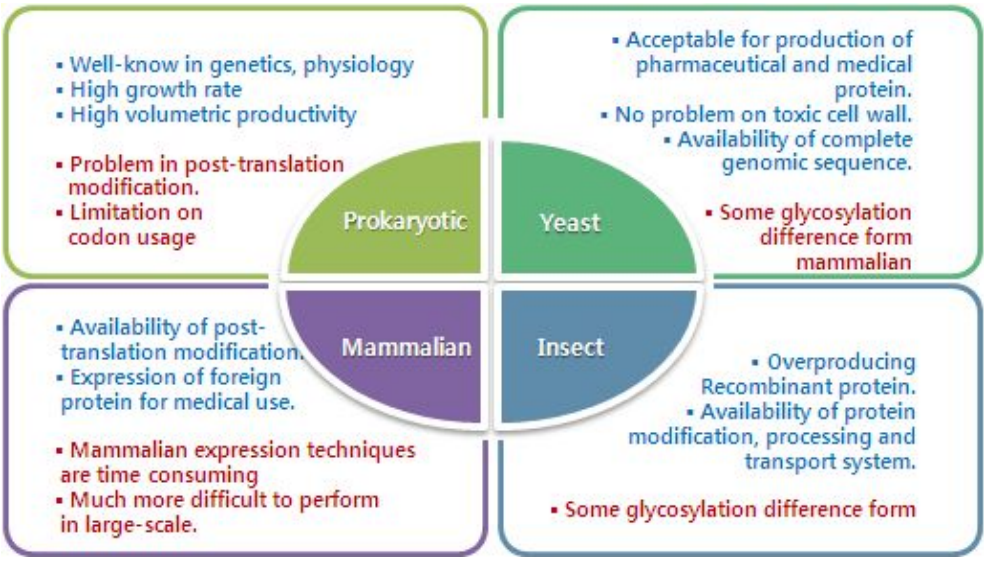


Figure 5. Gene expression systems.

1.2.1. Post-translational modification

A number of post-translational modifications of proteins are known, but their function is not exactly known. In prokaryotes post glycosylation is not performed. But in eukaryotes a variety of post-translational modifications are performed, including methylation, sulfation, lipid addition, phosphorylation, and glycosylation. Such modifications may be of critical importance to the function of an expressed protein. Secreted proteins, membrane proteins, and proteins targeted to vesicles or certain intracellular organelles are likely to be glycosylated. The oligosaccharides can divided into *N*-linked and *O*-linked by depending on the amino acid they are attached to. The most common and best study is *N*-linked glycosylation, where oligosaccharides are uniquely added to asparagine found in Asn-X-Ser/Thr recognition sequences in proteins (X can be any amino acid except proline).[9] *O*-linked glycosylation function is not clearly known but it is thought to play a role in protein secretion and cell anchoring.[11]

Glycosylation can also functionalize the natural proteins. It can prevent the protein from degradation by proteases by sterically hindering access to the backbone. Glycosylation can help to dissolve hydrophobic proteins with its hydrophilic sugars and it is sometimes important for activity of enzymes.[12]

1.2.2. Prokaryotic expression systems

Prokaryrotic expression systems is most widely used. Because of wide knowledge about these systems, many advantages are known. The availability of commercial vectors and strains, easy to transform, grows quickly (20 min for E. *coli* and 2 hours vs. *P. pastoris*)[9] in simple media, and requires inexpensive equipment for growth and storage. And in most cases, they can be made to produce adequate amounts of protein suitable for the intended application. A disadvantage of prokaryotic expression systems are that post-translational modifications are not performed.

There are two general classes: gram negative bacteria and gram positive bacteria. The gram negative bacterium *Escherichia coli* remains one of the most attractive expression hosts because of its rapid growth with high density on inexpensive substrates, its well-characterized genetics, and availability of many cloning vectors and various mutants of host strains.

Their expression vectors are plasmids. They have an origin of replication that controls the copy number, a selectable marker that not only insures plasmid stability in the cell but also facilitates selection for the transformed cells, a promoter and a transcription terminator.[13]

Gram positive bacteria have a long history in applications for the food industry and the agricultural and medicinal biotechnology.[14] Unlike gram negative bacteria, gram positive bacteria do not have a periplasmic space and can secrete proteins to the culture medium. Therefore proteins with a leader sequence are required. The general characters of gram-positive bacteriums show cytoplasmic lipid membrane and thick

peptidoglycan layer.

Gram positive bacteria can be divided into two groups by their GC (guanine and cytosine) contents. One of the high GC content groups are *Streptomyces* that can secrete heterologous proteins in the culture medium at high concentration. It has potential as production host for therapeutic proteins in large scale. *Bacilus*, *Clostridium*, *Lactococcus*, *Lactobacillus*, *Staphylococcus* and *Streptococcus* belong to the low GC content group. The codon usage of these species is influenced by bacteria GC content.

1.2.3. Eukaryotic expression

Eukaryotic expression systems is often the only workable way to the large-scale production of authentic, post-translationally modified proteins. It is becoming a suitable system to overexpress virtually any gene products. Mammalian, insect, and lower eukaryotic host are commercially available systems and they show a wide range of possibilities for expression.

1.2.3.1. Yeast

Yeasts are unicellular eukaryotic micro-organisms. The most well known yeasts is *Saccharomyces cerevisiae*. And another kind of yeast, *Pichia pastoris,* is widely used. *Saccharomyces cerevisiae* has been used as a model system to study genetics, providing a many knowledge about its genetics and cultivation because it is unicellular and eukaryotic.

The expression systems offer high efficiency, high cell densities, high yields and low fermentation costs. Yeast is easy to grow in large scale with simple nutritional demands. Different from *E. coli*, expression in yeast provides correctly folded proteins directly to the medium. Other advantages of yeast are that they are safer than working with mammalian tissue or cell lines, host cell proteins are secreted low level, and no specialized bioreactor is required.

1.2.3.2. Mammalian cells

Over other expression systems, the most advantages of protein production in mammalian cell lines or mammals is that they can express protein as its native form. Expression vectors are commercially available and well defined regulatory track is recorded. The protein can be secreted by the mammary glands and it can be harvested by milking.[15)]

A great disadvantage of the use of mammalian cells is their requirement for expensive and complex medium. And the culture media might contain bovine proteins and causes allergies.

1.2.3.3. Plants

Whole plants or plant cell suspension cultures are used in protein produce. They can perform post-translational modification. The advantages of plant system are low production of biomass, potential for large scale use, and low risk for pathogenic contaminants. Disadvantages are that there can be allergic reactions.

1.2.3.4. Insect cells

Insect cells are suitable for protein expression when proper post-translational modifications is required. Their expression system is safe, convenient to scale up, and reasonable in the expression levels. And also they use inexpensive serum free media, and are rapidly developed to a cell line. Even though all these things are possible, the insect cell technology are not very used in the biotech industry.

1.3. Protein engineering

Although enzyme catalyzes a wide range of chemical reactions with high specificity and selectivity, naturally occurring enzymes often looses their optimal function (activity, stability, specificity, and selectivity) under process conditions like high temperature, organic solvents, and extreme pH. To overcome this limitation, protein engineering can be employed for altering enzyme properties by DNA modification.[2)] Protein engineering can be categorized into two general strategies: rational design and directed evolution (Figure 6).

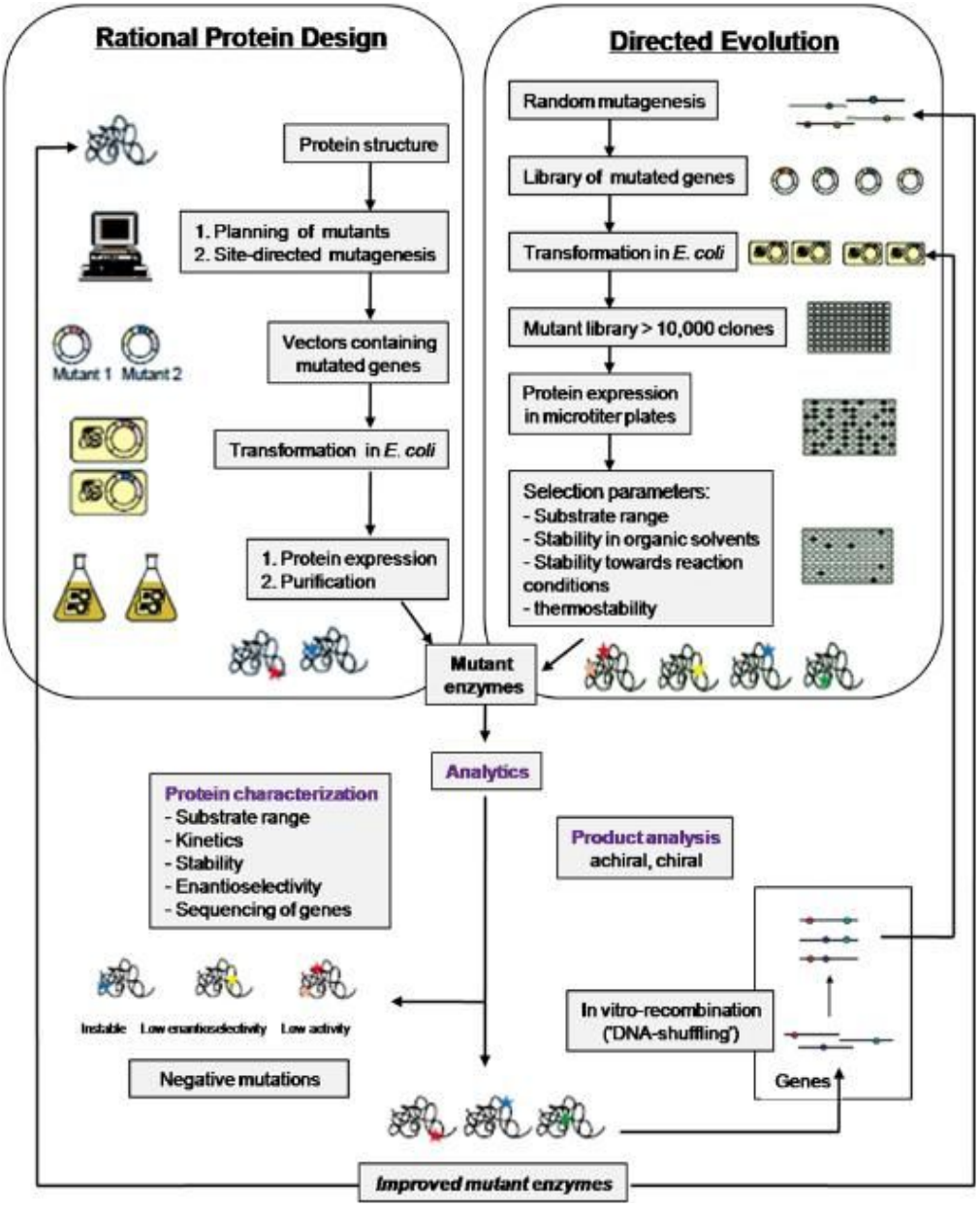


Figure 6. Comparison of rational protein design and directed evolution. (The figure is adapted from Bornscheuer, U. T.; Pohl, M. *Curr. Opin. Chem. Biol.* **2001**, *5*, 137-142)

1.3.1. Rational design

Rational design usually requires detailed knowledge of enzyme structure, function, and catalytic mechanism.[16] The three-dimensional protein structure is obtained by X-ray crystallography or NMR method. It involves rational alterations of selected residues in a protein to improve the enzyme selectivity or stability. In many cases, the rational design method can make a change close to the active site to improve the catalytic performance. However, it is hard to predict the effect of distant mutation.

The most effective and widely used strategy of rational design approach is site-directed mutagenesis. A few PCR-based mutagenesis methods have been reported[17] and the overlap extension PCR mutagenesis method (Figure 7) is the most widely used.[18]

Complementary primers are used to generate two fragments having overlapping ends. These fragments are purified and combined for another round of PCR without additional primers. This reaction is called overlap extension because the 3' end of the first PCR product is complimentary to the 5' end of the other first PCR product. The result product is amplified further by PCR. Specific alterations in the amino acid can be introduced by incorporating nucleotide changes into the overlapping primers. This method represents a significant improvement of site-directed mutagenesis because it is much faster, simpler and approaches 100% efficiency in the generation of mutant product.

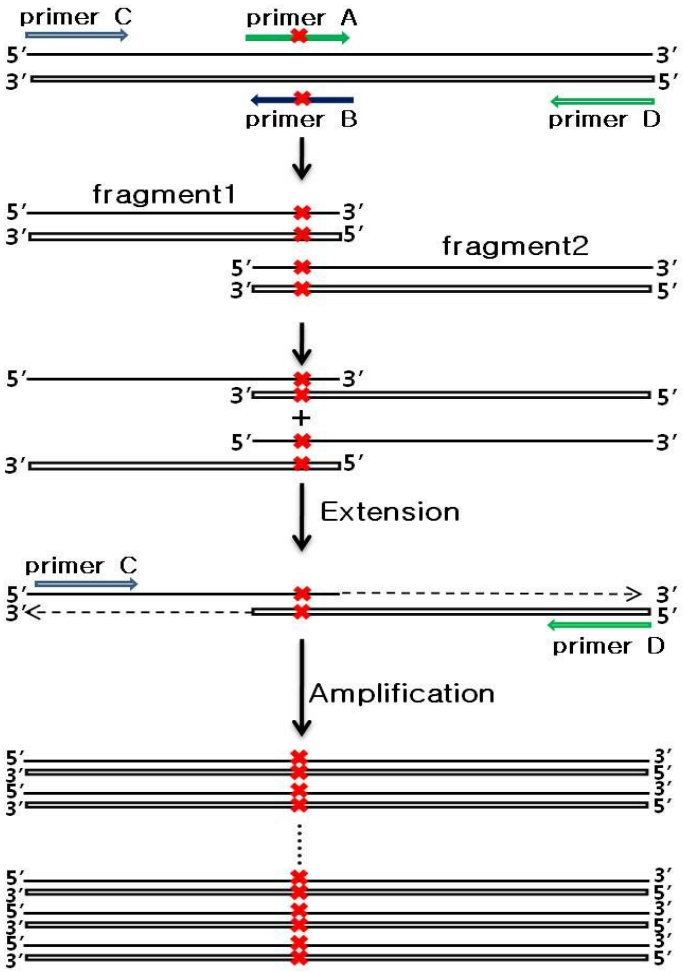


Figure 7. Overlap extension PCR to introduce site-directed mutagenesis. Four different primers (primer A, B, C, D) are used.

A different method to perform site-directed mutagenesis uses the mismatched oligonucleotides.[19] Two different primers are used to introduce the mutations. One primer containing the desired mutation in the gene and the other primer containing a mutation in the plasmid that will eliminate a unique restriction site. The primers annealed to the same strand of the plasmid and extended by using DNA polymerase. The newly synthesized strands are joined by DNA ligase and a mutated strand of the plasmid is now formed. The plasmids are subjected to a cleavage of the restriction site that was deleted by the mutagenesis, in order to distinguish the parental plasmids from the plasmids containing the mutation. The linearized parental plasmids have a much lower transformation frequency than the intact mutated plasmids. The advantage of using this method is the ligation step is not needed and disadvantage is when plasmids are larger than 10 kb, its fidelity of the polymerase decreases significantly[20].

1.3.2. Directed evolution

Directed evolution (sometimes called irrational design) mimics the natural evolution process in the laboratory and involves repeated cycles of generating a library of different protein variants.[21] It only relies on a very simple algorithm that nature has successfully been using for eons. And it dose not require any detailed structural and functional understanding of the target enzymes. For these reasons, directed evolution is necessary strategy, even though rational designed is an effective protein strategy.

The general scheme of directed evolution is shown in Figure 8. Genetic diversity is first introduced into a target gene through random mutagenesis or recombination. The library of mutant genes in then transformed into host cells in which the mutant genes are converted into their corresponding proteins. Functionally improved mutant proteins are identified through an appropriate selection or screening strategy. The same process will be repeated until the goal is achieved or no further improvement is possible.

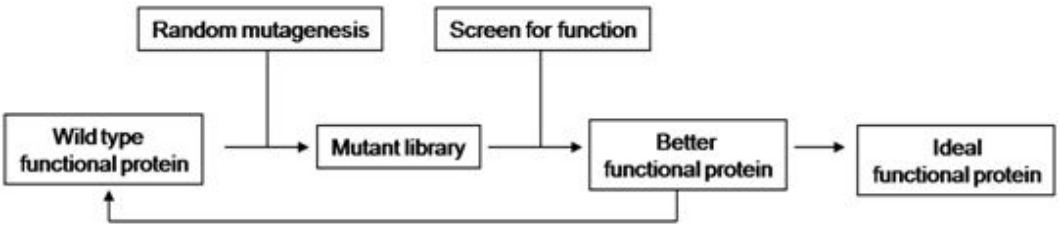


Figure 8. The general scheme of directed evolution.

The directed evolution, such as error-prone PCR and DNA shuffling, has become a powerful tool for tailoring various enzyme or protein functions. As a consequence, a molecular approach through genetic engineering becomes a good alternative to achieve better functional protein via directed evolution. Error-prone PCR introduces random copying errors by imposing imperfect and thus mutagenic, reaction conditions[22]. The mutation rate can be increased by several techniques. The use of magnganese instead of magnesium as the divalent cation greatly increases mutation frequency of DNA polymerase[23]. Biasing the ratios of natural deoxyribonucleotide triphosphate (dNTP) pools also increases the frequency of misincorporations and frameshifts.[24]

1.4. Enzyme enantioselectivity

Enantioselectivity is the ratio the specificity constant of the enzyme for the two enantiomers. It can be expressed as the enatiomeric ratio, *E*. This is defined as the ratio of specificity constants (k_{cat}/K_M) for both enantiomers.[25] By this value, enantiomer preference of enzyme can be evaluated.

$$E = \frac{(\frac{k_{cat}}{K_m})_R}{(\frac{k_{cat}}{K_M})_S}$$

Enantioselectivity can be explained by the difference of the activation energy (E_a) for the two enantiomers. The ground states of two enantiomers and that of product have same energy level each other (Figure 10). The only difference for the enzymatic conversion of the substrate to the product is the activation energy.

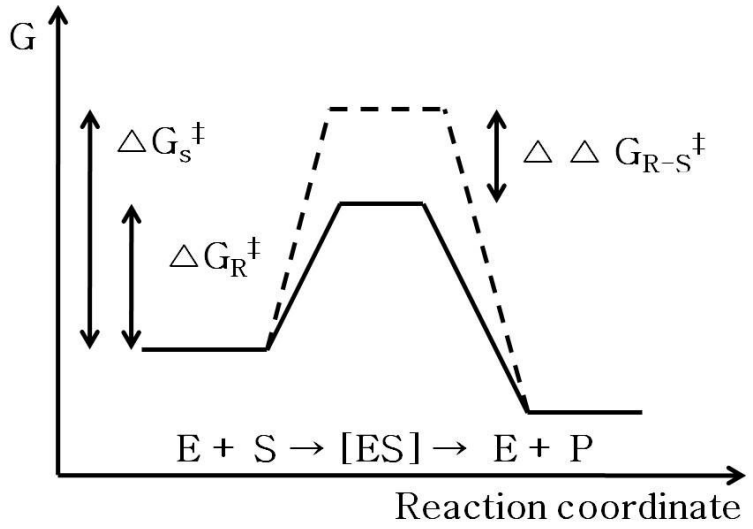


Figure 9. The free energy profiled of an enantioselective enzyme catalyzed reaction.

Enantiomeric ratio is related to the difference in the Gibbs free energy of the transition states of the enantiomers ($^{\triangle\triangle}G^{\ddagger}$).[26)]

$$\triangle\triangle G^{\ddagger} = -RTln\frac{V_R}{V_S} = -RTlnE$$

(R: gas constant, 8.31 $Jmol^{-1}K^{-1}$, T: temperature)

1.5. Investigation in this thesis

Enantiopure tetrahydrofuran-2-carboxylate (THFC) is an important intermediate as well as a building block in the pharmaceutical industry and research. Especially (*R*)-tetrahydrofuran-2-carboxylate is the first compound that is incorporated into a penem skeleton to give furopenem, a clinically efficient non-natural *β*-lactam antibiotic (Figure 10).[27)]

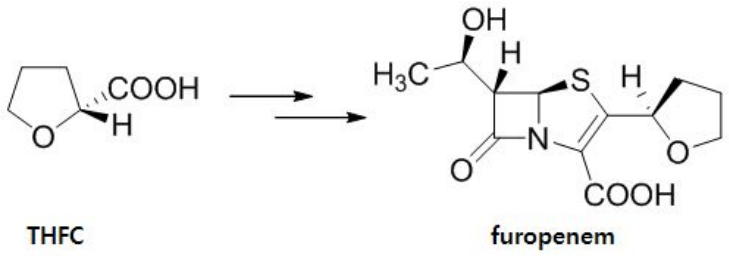


Figure 10. Structures of (*R*)-tetrahydrofuran-2-carboxylate and furopenem.

In this thesis, an enzymatic kinetic resolution method is used to produce enantiopure THFC. Hydrolysis reaction toward a series of esters of tetrahydrofuran-2-carboxylate was performed by *Candida antarctica* lipase B (CAL-B). However, parental enantioselectivity of CAL-B is not very high enough (*E*=3.41). CAL-B prefer long chain fatty acid and when substrates get into the active site they shaped as puckered shape (Figure 11). Based on these facts, enzyme engineering approach of using site-directed mutagenesis and substrate engineering are employed to improve enantioselectivity.

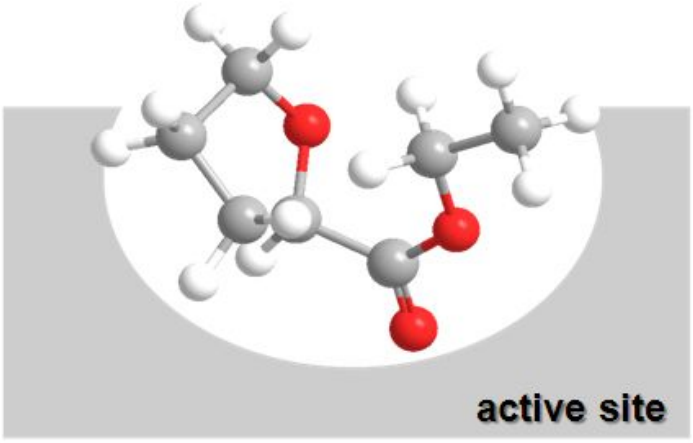


Figure 11. Puckered shape of substrate in the active site of enzyme.

1.6. Previous study

In the previous study, several sites of CAL-B were selected in considering hydrogen bonding and repulsive interaction with the oxygen atom of the substrate, ethyl tetrahydorfuran-2-carboxylate (Figure 12).[28)]

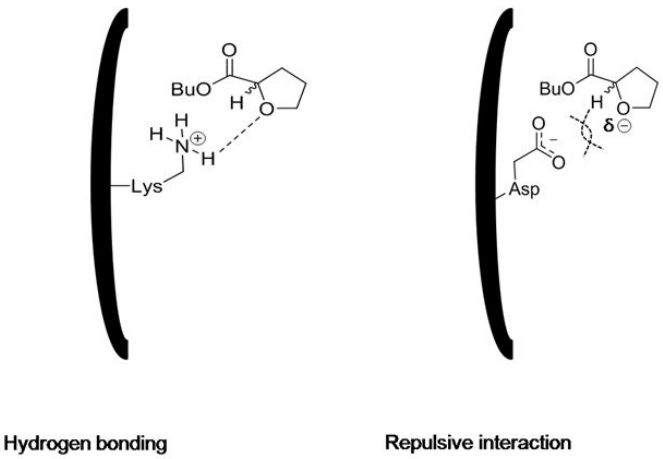


Figure 12. The effects of construction of hydrogen bond and repulsive interaction in kinetic resolution of THFC.

Based on the crystal structure, three residues near the active site, D134, Q157, and V190, were selected and changed into lysine, valine, glutamate, and glycine (Table 3).

Table 3. The list of mutant enzymes.

Entry	Enzyme
1	D134K V190G
2	D134V V190G
3	D134Q V190G
4	Q157K

The wild-type and all mutants showed (*S*)-prefernence (Table 4). Unfortunately, all mutants showed lower enantioselectivity than wild-type.

Table 4. Enantioselectivity values determined for the wild-type CAL-B and the mutants toward the racemic ethyl tetrahydrofuran-2-carboxylate.[a)]

Entry	Enzyme	*E*	Enantiopreference
1	wild-type	3.4	*S*
2	D134K V190I	1.8	*S*
3	D134V V190I	1.7	*S*
4	D134Q V190I	1.6	*S*
5	Q57K	2.0	*S*

[a)]This data is adapted from reference 28.

CHAPTER 2. Experimental section

General Methods. Chemicals and buffers were purchased from Sigma-Aldrich. LB-broth was purchased from Merk. *Pfu* DNA polymerase, Taq DNA polymerase and restriction enzymes (*Nco* I and *Sal* I) were purchased from Enzynomics (Daejeon, Korea). DNA oligomers were obtained from Sigma-Proligo (Singapore). The vector (*p*BADgIIIa) was purchased from Merck biosciences. DNA sequencing was performed by Solgent Co. (Daejeon, Korea). ^{1}H-NMR spectrum was recorded from D_2O solutions on a Varian 500 MHz spectrometer. Gas Chromatography (Agilent 6890N) analyzed by chiral capillary column (Cyclosil-B 30 m × 0.25 mm).

CAL-B gene cloning. The optimized gene of lipase B from *Candida antarctica* (opt2CAL-B) was synthesized and amplified amplification using the primers opt2CAL-B_F1_*Nco* I (5'-TAGCACCATGGCTCTGCCGTCT-3') and opt2CAL-B_R1_*Sal* I (5'-TGATGGTCGACCGGAGTTA-3'). The forward and reverse primers contain a *Nco* I site at 3'-end and a *Sal* I site at 5'-end, respectively. The sequence of opt2CAL-B gene is shown in Table 5. The PCR product was restricted with *Nco* I and *Sal* I and ligated to the vector *p*BADgIIIa (Invirogen) that also restricted with same enzymes (Figure 13).

The resulting plasmid constructed (*p*BADgIIIa opt2_CAL-B_5D) was transformed to *E. coli* (Top10) and transformants were selected on LB-ampicillin plates. By PCR colony screening, correctly subcloned genes were selected.

Table 5. Codon-optimized sequences of CAL-B.

Sequence (opt2 CAL-B_5D)
CTG CCG TCT GGT TCC GAT CCG GCT TTC TCC CAG CCG
AAA TCC GTG CTG GAC GCG GGT CTG ACC TGT CAG GGT
GCT TCT CCA AGC AGC GTG TCT AAA CCG ATC CTG CTG
GTA CCG GGC ACC GGT ACC ACT GGC CCG CAG TCT TTC
GAC AGC AAC TGG ATT CCA CTG TCC ACC CAA CTC GGT
TAT ACT CCT TGC TGG ATC TCT CCG CCG CCG TTT ATG
CTG AAC GAT ACT CAG GTA AAC ACT GAA TAC ATG GTA
AAC GCT ATC ACC GCT CTG TAC GCA GGT TCT GGT AAC
AAC AAA CTG CCA GTG CTG ACC TGG TCC CAG GGT GGT
CTG GTT GCA CAA TGG GGC CTG ACT TTC TTC CCG TCT
ATC CGT TCT AAA GTG GAC CGT CTG ATG GCA TTC GCT
CCG GAC TAC AAA GGT ACT GTG CTG GCT GGC CCG CTG
GAT GCA CTG GCT GTA TCT GCG CCA TCC GTG TGG CAG
CAG ACC ACT GGT TCT GCG CTG ACC ACT GCA CTG CGT
AAC GCT GGT GGT CTG ACC CAG ATC GTT CCG ACT ACT
AAC CTG TAC AGC GCA ACC GAT GAG ATC GTT CAG CCG
CAG GTA TCT AAC TCC CCG CTG GAT TCT TCT TAC CTG
TTC AAC GGT AAG AAC GTT CAG GCT CAG GCT GTT TGT
GGC CCG CTG TTC GTT ATC GAT CAC GCA GGT TCC CTG
ACC TCC CAG TTC AGC TAT GTG GTT GGC CGC TCT GCT
CTG CGC TCC ACC ACT GGT CAA GCG CGC TCT GCT GAC
TAC GGC ATC ACC GAC TGC AAC CCG CTG CCG GCG AAC
GAC TTA ACC CCG GAA CAG AAG GTT GCA GCT GCG GCT
CTG CTG GCA CCG GCT GCA GCT GCA ATT GTT GCG GGC
CCG AAA CAG AAC TGC GAA CCG GAC CTG ATG CCG TAC
GCT CGT CCG TTC GCG GTT GGT AAA CGC ACT TGT TCT
GGC ATC GTA ACT CCG

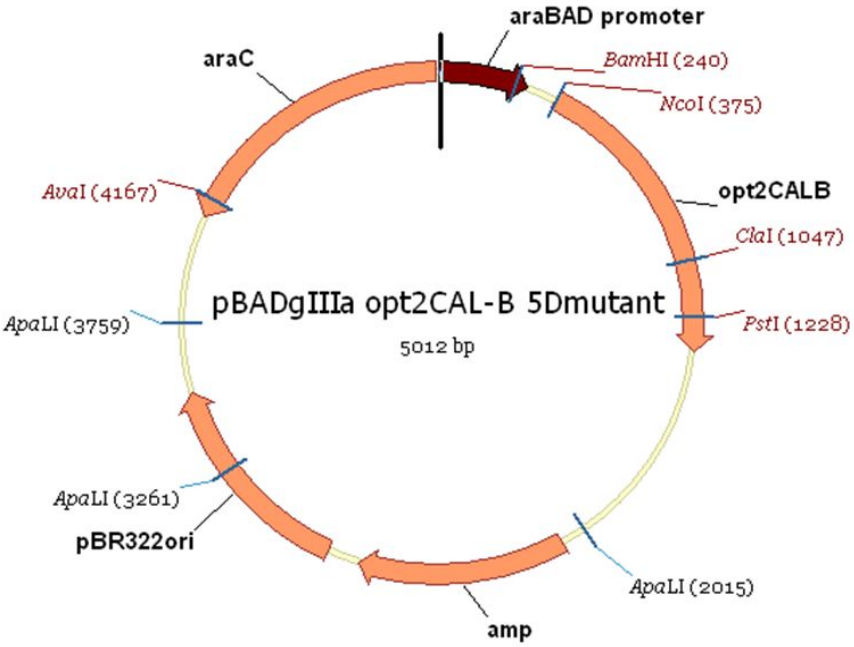


Figure 13. The map of vector *p*BADgIIIa coding opt2 CAL-B_5D.

Site-directed mutagensis. The overlap extension PCR mutagenesis method was for site-directed mutagenesis. opt2_CAL-B_5D was used as template. The four different primers are required. opt2CAL-B_F1_*Nco* I and opt2_CAL-B_ R1_*Sal* I were used for introducing restriction site to 3' and 5' ends. And other 2 primers cover the region where a point mutation is desired. The primers that are used for mutation are listed in Table 6.

Table 6. CAL-B mutagenesis primers.

Mutagenesis primers
opt2_CAL-B_F1_NcoI
5'-tag cac cat ggc tct gcc gtc t-3'
opt2_CAL-B_R1_SalI
5'-tga tgg tcg acc gga gtt acg atg-3'
F_CAL-B_5D_D134E
5'-gca ttc gct ccg gaa tac aaa ggt act gtg-3'
R_CAL-B_5D_D134E
5'-agt acc ttt gta ttc cgg agc gaa tgc cat-3'
F_CAL-B_5D_D134Q
5'-gca ttc gct ccg cag tac aaa ggt act gtg-3'
R_CAL-B_5D_D134Q
5'-agt acc ttt gta ctg cgg agc gaa tgc cat-3'
F_CAL-B_5D_T139I
5'-gac tac aaa ggt att gtg ctg gct ggc ccg-3'
R_CAL-B_5D_T139I
5'-gcc agc cag cac aat acc ttt gta gtc cgg-3'
F_CAL-B_5D_T139L
5'-gac tac aaa ggt ctg gtg ctg gct ggc ccg-3'
R_CAL-B_5D_T139L
5'-gcc agc cag cac cag acc ttt gta gtc cgg-3'
F_mu_CALB_I189D
5'-agc gca acc gat gag gac gtt cag ccg cag gta tc-3'

R_mu_CALB_I189D

5'-gat acc tgc ggc tga acg tcc tca tcg gtt gcg ct-3'

F_mu_CAL-B_I189E

5 - cag cgcg aac cga tga gga agt tca gcc gca ggt atc taa c-3'

R_mu_CALB_I189E

5'-gtt aga tac ctg cgg ctg aac ttc ctc atc ggt tgc gct g-3'

F_mu_CALB_I189M

5'-agc gca acc gat gag atg gtt cag ccg cag gta tc-3'

R-mu_CALB-I189M

5'-gat acc tgc ggc tga acc atc tca tcg gtt gcg ct-3'

F_mu_CALB_I189Q

5'-gcg caa ccg atg agc agg ttc agc cgc ag-3'

R_mu_CALB_I189Q

5'-ctg cgg ctg aac ctg ctc atc ggt tgc gc-3'

F_CAL-B_5D_V190I

5'-acc gat gag atc att cag ccg cag gta tct-3'

R_CAL-B_5D_V190I

5'-tac ctg cgg ctg aat gat ctc atc ggt tgc-3'

F_CAL-B_5D_V190L

5'-acc gat gag atc ctg cag ccg cag gta tct-3'

R_CAL-B_5D_V190L

5'-tac ctg cgg ctg cag gat ctc atc ggt-3'

F_CAL-B_I189E V190I

5'-gca acc gat gag gaa att cag ccg cag-3'

R_CAL-B_I190E V190I

5'-ctg cgg ctg aat ttc ctc atc ggt tgc-3'

F_CAL-B_I189M V190I

5'-gca acc gat gag atg att cag ccg cag-3'

R_CAL-B_I190M V190I

5'-ctg cgg ctg aat cat ctc atc ggt tgc-3'

F_CAL-B_I189Q V190I

5'-gca acc gat gag cag att cag ccg cag-3'

R_CAL-B_V190I_I289Q

5'-ctg cgg ctg aat ctg ctc atc ggt tgc-3'

Protein expression The *E.coli* (Top10) containing the CAL-B gene was incubated in 15 mL of LB medium containing 15 μL of ampicillin (100 mg/ml) at 37 ℃ for 16 h. And 1 mL of *E. coli* overnight culture was added to 100 mL of LB medium, and incubated at 37 °C and 200 rpm to an OD_{600} of 0.5. Protein expression was induced by addition of 1 mL of an arabinose solution (2% w/v) and the expression culture was incubated for 6 h at 25 °C and 200 rpm. When the OD_{600} reached 1.5, the cells were harvested by centrifugation (15 min, 3,800 *xg*, 4 °C) and the supernatant was discarded. The cell pellet (~0.8 g) was resuspended in 4 ml of the lysis buffer (NaH_2PO_4, 50 mM; NaCl, 300 mM; imidazole, 10 mM; pH 8.0 adjusted with NaOH) and then sonication (4 kHz, 5 times of 20 sec pulse with 30 sec interval) was applied. After centrifugation (10 min, 10,000 *g*, 4 °C), the cell debris and the supernatant was separated for further experiments. The cell debris was dissolved in a 4 mL of urea solution ($(NH_2)_2CO$, 8 mM, containing 1 mM dithiothreitol) for SDS-PAGE analysis.

Synthesis of various ester of tetrahydrofuran-2-carboxylate. 10mL of tetrahydro-2-furoic acid (0.104 mol) was added to 40 ml of various ethanols (ethanol 0.1713 mol; butanol 0.4382 mol; 2-methyl-1-propanol 0.4437 mol; 2-methyl-1-butanol 0.1751 mol) and ten drops of sulfuric acid were added. The reaction mixture was refuluxed for 24 h. Ethanol was evaporated by a rotavap and the concentrated was diluted with 10mL of diethyl ether. The ether solution was washed twice with a 40 mL of saturated sodium carbonate aqueous solution. The ether layer was dried over anhydrous sodium sulfate and concentrated by rotovap.

(±)-ethyl tetrahydrofuran-2-carboxylate (yield = 8.73 g)

^{1}H NMR (500 MHz, $CDCl_3$): δ 4.59-4.56 (1H, t); 4.36-4.30 (2H, m); 4.17-4.02 (2H, m); 2.39-2.11 (2H, m); 2.10-2.03 (2H, m); 1.42-1.40 (3H, t).; ^{13}C NMR (125 MHz, $CDCl_3$, δ): 173.463, 77.363-76.855, 69.384, 60.950, 30.214, 25.267, 14.249.; MS (ESI) m/z = 145.0 $(M+H)^+$

(±)-butyl tetrahydrofuran-2-carboxylate (yield = 8.13 g)

^{1}H NMR (500 MHz, $CDCl_3$): δ 4.50-4.48 (1H, dd); 4.21-4.16 (2H, t); 4.08-3.94 (2H, q); 2.31-2.27 (2H. m); 2.07-1.94 (2H, m); 1.71-1.58 (2H, m); 1.46-1.39 (2H, m); 1.00-0.96 (3H, t); ^{13}C NMR (125 MHz, $CDCl_3$, δ): 173.596, 77.351-76.844, 69.391, 64.815, 30.669-30.584, 25.279, 19.125, 13.146.; MS (ESI) m/z = 172.0 $(M+H)^+$

(±)–isobutyl tetrahydrofuran-2-carboxylate (yield = 8.96g)

^{1}H NMR (500 MHz, $CDCl_3$): δ 4.49-4.46 (1H, t); 4.05-4.01 (2H, d); 3.95-3.91 (2H, m); 2.29-2.22 (2H, m); 2.06-1.99 (1H, m); 1.97-1.87 (2H, m); 0.93-0.91 (3H, dd); ^{13}C NMR (125 MHz, $CDCl_3$, δ): 173.589, 77.363-16.852, 70.921, 69.395, 30.332, 27.781, 25.290, 19.081.; MS (ESI) m/z = 171.0 $(M+H)^+$

(±)–isopentyl tetrahydrofuran-2-carboxylate (yield = 9.52 g)

^{1}H NMR (500 MHz, $CDCl_3$): δ 4.46-4.43 (1H, t); 4.19-4.11 (2H, t); 4.04-3.66 (2H, m); 2.28-1.98 (2H, m); 1.97-1.88 (2H, m); 1.72-1.64 (1H, m); 1.56-1.47 (2H, q), 0.95-0.88 (3H, dd); ^{13}C NMR (125 MHz, $CDCl_3$, δ): 173.604, 77.366-76.855, 69.414, 63.645, 37.322, 30.288, 25.301-25.105, 22.520.; MS (ESI) m/z = 187.0 $(M+H)^+$

Measurement of hydrolytic activity toward hydrolysis of *p*-nitrophenyl acetate of the CAL-B mutant enzyme. Hydrolytic activity of the enzymes was measured by following the hydrolysis of *p*-nitrophenyl acetate at 404 nm. The assay solution was prepared by mixing *p*-nitrophenyl acetate (20 μL, 200 mM in acetonitrile), acetonitrile (870 μL), and BES buffer (5 mM, pH 7.2, 11.110 mL). The absorbance change was measured at 405 nm for 10min after mixing the assay solution (100 μL) with the enzyme solution (5 μL).

Determination of enantioselectivity for hydrolysis of (±)–ethyl tetrahydro furan-2-carboxylate and (±)–isopentyl tetrahydrofuran-2 -carboxylate. The hydrolysis reaction was monitored using a pH-stat. The reaction started by an addition of 100 μl of the crude enzyme solution (in 5 mM BES buffer, pH 7.2) to a solution of 200 μL of (±)–ethyl tetrahydrofuran-2-carboxylate (1 M in acetonitrile) and 500 μl of acetonitrile in 9300 μl of BES buffer (1 mM, pH 7.2). The reaction was maintained at pH 7.2 by controlled addition of 0.1 N NaOH. Reaction was stopped at conversion around 40 % by extraction with 10 mL of diethyl acetate twice. To protonate the product, the buffer layer's pH was down to around pH 2 by adding three drops of HCl (1 M). And the product was extracted with 10 mL of diethyl ether twice. Both diethyl acetate and diethyl ether were dried by anhydrous sodium sulfate. Diethyl ether was concentrated by rotavap. The organic layers (diethyl acetate and diethyl ether) analyzed by GC with chiral capillary column: initial column

temperature 100 ℃ for 3 min, ramp up to 130 ℃ at a rate of 2.0 ℃/min and then held at 130 ℃ for 20 min. Enantioselectivity was calculated using the method of Chen *et al*. from enantiomeric excesses of both starting ester and acid product.[29)]

CHAPTER 3. Result and Discussion

Selection of mutation sites. To improve the enantioselectivity of CAL-B for kinetic resolution of tetrahydrofuran-2-carboxylate (THFC), three residues were selected (I189, V190, and T138). These residues are placed near the catalyric His224, and could influence the substrate rotation. To prevent the rotation of the substrate, these residues were changed into the amino acids containing bulky side chain. The residue I189 was mutated to aspartic acid, glutamic acid, methionine, and glutamine. The residue V190 and T138 were replaced with bigger sized amino acid such as leucine and isoleucine (Table 7 and Figure 14).

Table 7. The list of mutant enzymes.

Entry	Enzymes
1	wild-type
2	I189D
3	I189E
4	I189M
5	I189Q
6	V190I
7	V190L
8	T138I
9	T138L
10	I189Q V190L
11	1I89M V190L
12	I189E V190L

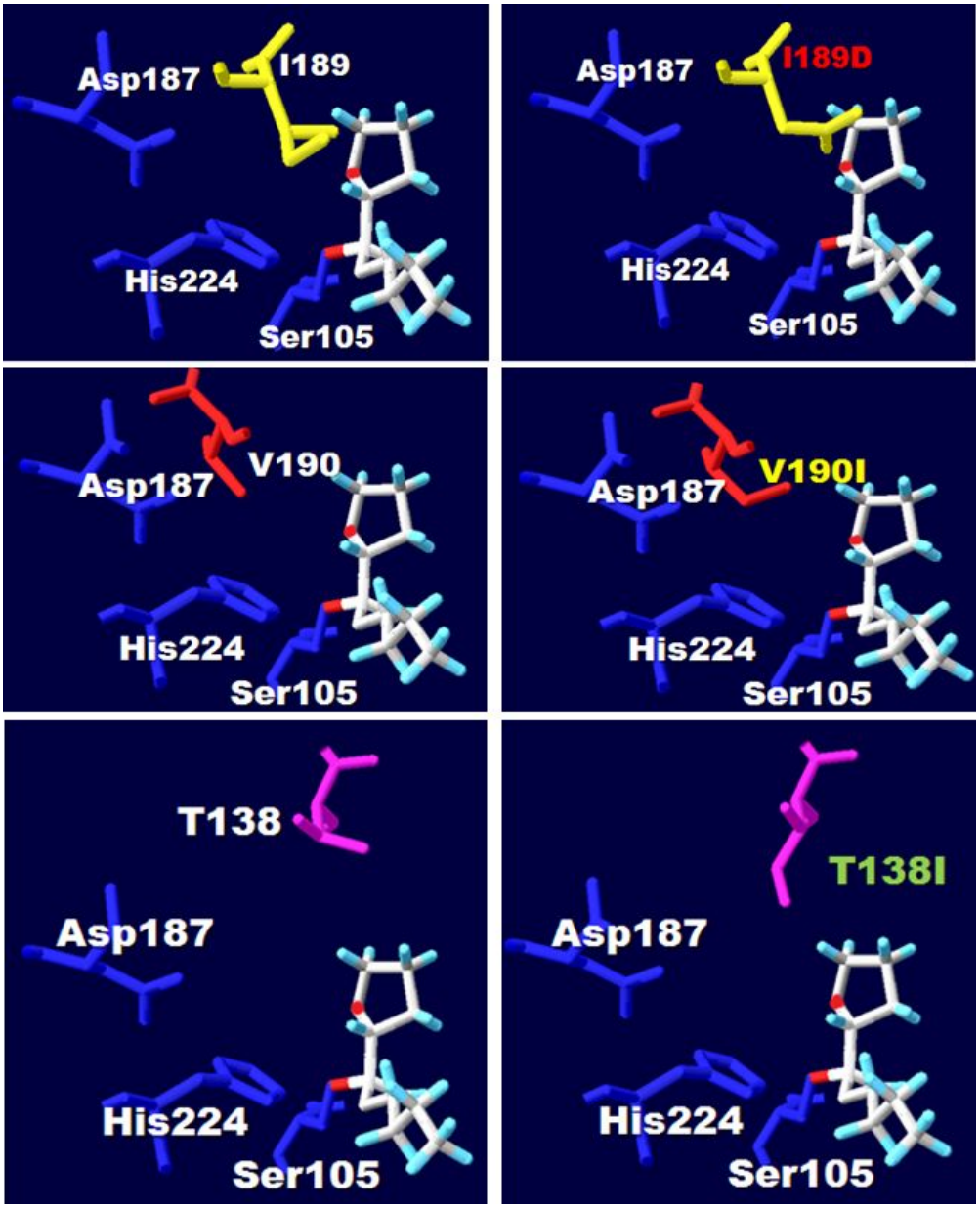


Figure 14. A docking structure of THFC in the CAL-B active site. Three mutation sites, I189, V190, and T138, are located near the active site and they are mutated to bulky amino acids. (a) I189 mutate to aspartic acid; (b) V190 mutate to isoleucine; (c) T138 mutate to isoleucine.

Protein expression. The mutant CAL-B enzymes were expressed same as the wild type enzyme. As shown in SDS-PAGE (Figure 15), CAL-B enzymes appeared at 33 kDa.

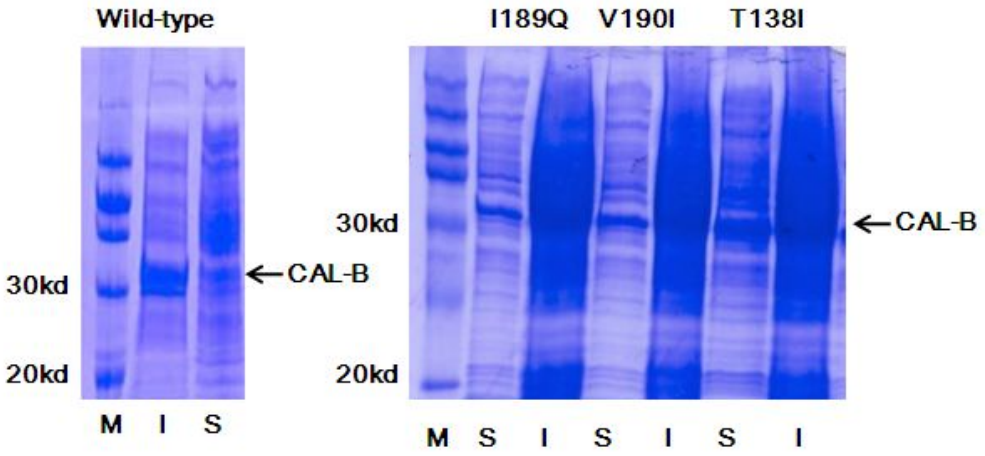


Figure 15. SDS-PAGE analysis of CAL-B produced in *E.coli* (Top10). SDS-PAGE was performed on a 12% polyacrylamide gel and stained using the Coomassie brilliant blue. M, molecular weight marker; I, insoluble fraction; S, soluble fraction.

Hydrolysis activity toward *p*-nitrophenyl acetate. To conform that mutation does not influence the active site, hydrolysis activity toward *p*-nitrophenyl acetate was performed (Figure 16).

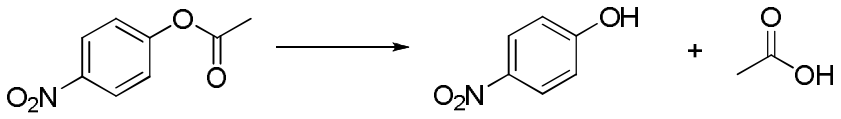


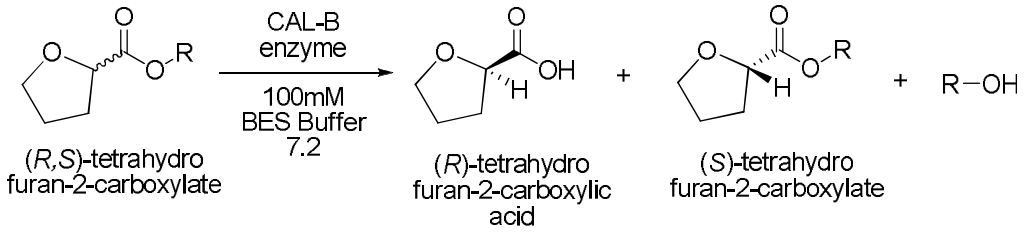
Figure 16. Hydrolysis activity toward *p*-nitrophenyl acetate.

Although most mutants showed decreased activity than wild-type, all mutants still sustain hydrolytic activity(Table 8). I189Q, T138I, and T138L showed higher activity then wild-type.

Table 8. Specific activity of CAL-B and mutant enzymes

Entry	enzyme	specific activity for *p*-nitrophenyl acetate (μmol $min^{-1}mg^{-1}$)
1	wild-type	7.30×10^{-4}
2	I189D	3.80×10^{-4}
3	I189E	3.79×10^{-4}
4	I189M	6.87×10^{-5}
5	**I189Q**	**8.44×10^{-3}**
6	V190I	3.33×10^{-4}
7	V190L	4.92×10^{-4}
8	**T138I**	**9.97×10^{-4}**
9	**T138L**	**1.15×10^{-3}**
10	I189E V190I	7.56×10^{-5}
11	I189M V190I	2.23×10^{-4}
12	I189Q V190I	1.12×10^{-4}

Hydrolysis reaction. Hydrolysis reaction was performed with four different tetrahydrofuran-2-carboxylate esters (Figure 17 and Table 9). Crude extracts were used to determine enantioselectivity.



R : (1) CH_2CH_3 (2) $CH_2CH_2CH_2CH_3$ (3) $CH_2CH(CH_3)_2$ (4) $CH_2CH_2CH(CH_3)_2$

Figure 17. Hydrolytic resolution of (*R,S*)-tetrahydrofuran-2-carboxylate in buffer media.

Table 9. Substrates of hydrolytic resolution.

Entry	Substrate
1	 ethyl tetrahydrofuran-2-carboxylate
2	 butyl tetrahydrofuran-2-carboxylate
3	 isobutyl tetrahydrofuran-2-carboxylate
4	 isopentyl tetrahydrofuran-2-carboxylate

First, hydrolysis reaction performed with ethyl tetrahydrofuran-2-carboxylate and three single mutant enzymes. As shown in Table 10, the wild-type and the single mutants at position 189 had limited effect on enzyme enantioselectivity. But all mutants showed reversed enantiopreference from the *S* to the *R* substrate.

Table 10. Enantioselectivity values determined for the wild-type CAL-B and the single mutants toward the racemic ethyl tetrahydrofuran-2-carboxylate.

Entry	Enzyme	Conversion(%)	*E*	Enantiopreference
1	wild-type	44.61	3.41	*S*
2	I189D	39.23	1.57	*R*
3	I189E	43.69	2.03	*R*
4	I189M	44.76	2.16	*R*

From these results, limitation of the rotation of the ring by introducing larger of negatively charged residues can cause to alter the enantiopreference. Nextly, rotation of furan ring was also limited by substrate engineering. The alkyl moiety of the ester was changed into longer alkyl groups such as butyl group. Without introducing bulky amino acid to the active site, the enantiopreference of wild-type showed (*S*)-preference. And enantioselectivity of all mutants were improved about 2 folds (Table 11). These results prove that not only protein engineering but also substrate engineering is needed to influence the enantiopreference and enantioselectivity

Table 11. Enantioselectivity values determined for the wild-type CAL-B and the single mutants toward the racemic butyl tetrahydrofuran-2-carboxylate.

Entry	Enzyme	Conversion(%)	*E*	Enantiopreference
1	wild-type	47.09	2.09	*R*
2	I189E	50.92	4.33	*R*
3	I189M	52.52	4.12	*R*

To get more improved enantioselectivity, 5 mutants are additionally made (I189Q, V190I, V190L, T190I, and T190L) and substrate was change into isopentyl tetrahydrofuran-2-carboxylate. As expected, this reaction showed significant effects on the enantiopreference (Table 12). First, enantiomeric ratio of wild-type was increased up to two folds than that of ethyl tetrahydrofuran-2-carboxylate. And compared to the parental wild-type enzyme, three single mutants (I189D, E, M) showed an increased enantioselectivity by up to two folds. From five new mutants, I189Q and V190I showed the best enantiomeric ratio (*E* = 13.51 and *E* = 11.26).

Table 12. Enantioselectivity values determined for the wild-type CAL-B and the single mutants toward isopentyl tetrahydrofuran-2-carboxylate.

Entry	Enzyme	Conversion (%)	*E*	Enantiopreference
1	wild-type	45.45	5.58	*R*
2	I189D	47.60	6.02	*R*
3	I189E	40.04	12.76	*R*
4	I189M	41.14	10.94	*R*
5	**I189Q**	**43.47**	**13.51**	***R***
6	**V190I**	**40.79**	**11.26**	***R***
7	V190L	36.58	5.91	*R*
8	T138I	35.14	8.57	*R*
9	T138L	34.43	7.65	*R*

Double mutants of I189 and V190 were made to obtain more increased enantiomeric ratio. We expected increased *E* value up to 20, but these mutants do not showed increased enantiomeric ratio.

Table 13. Enantioselectivity values determined for the double mutants toward isopentyl tetrahydrofuran-2-carboxylate.

Entry	Enzyme	Conversion (%)	*E*	Enantiopreference
1	I189E V190I	37.99	7.25	*R*
2	I189M V190I	33.01	10.48	*R*
3	I189Q V190I	42.39	13.78	*R*

CHAPTER 4. Conclusion

Enzymes are highly efficient catalysts with extraordinary enantio- and regioselectivity, and they can accept a wide range of complex molecules as substrates. Obtaining enantiomerically pure compounds is important because they are key factor of pharmaceutical industrial fields.

In this thesis, poor parental enantioselectivity of CAL-B were improved by substrate engineering and enzyme engineering approaches using site-directed mutagenesis. Through substrate engineering, enantiopreference of CAL-B wild-type is changed from (*S*)-selective to (*R*)-selective. And mutant enzymes such as I189Q (E = 13.51) and V190I (E = 11.26) showed higher enantioselectivity than wild-type (E = 5.58).

Although enantioselectivity is improved up to 13, these value is not enough for synthetic applications. Therefore, further experiment may be done by double mutagenesis; for example combination of V190I and T138I. And new enzymes produced by combination of site-directed mutagenesis and random mutagenesis are expected to show improved enantioselectivity.

Reference

1) Jonhannes, T.; Simurdiak, M. R.; Zhao, H. *Encyclo. Chem. Proces*. **2005**, *30*, 101-110.

2) Beilen, J. B. V.; Li, Zhi. *Curr. Opin. Biotechnol*. **2002**, *13*, 338-344.

3) Mugnusson, A. Rational redesign of *Candida antarctica* lipase B. Ph. D. Thesis, KTH, Sweden, 2005.

4) Jaeger K. E.; Reetz, M. T. *Trends Biotechnol*. **1998**, *16*, 396-403.

5) Faber, K. Biotransformations in Organic Chemistry, 5th ed.; Springer-Verlag: Berlin, Germany, 2000.

6) Grochulski, P.; Li, Y.; Schrag, J. D.; Bouthillier, F.; Smith, P.; Harrison, D.; Rubin, B.; Cygler, M. *J. Biol. Chem.* **1993**, *268,* 12843-12947.

7) a) Fersht, A. Enzyme Structure and Mechism, 2nd ed.; W. H. Freeman and Co.: New York, 1985, pp 98-103. b) Verger, R. *Trends Biotechnol*. **1997**, 15, pp 32-38.

8) Bornsheuer, U. T.; Kazlauskas, R. J. Hydrolases in Organic Synthesis: Regio- and Stereoselctive Biotransformations; Wiley-VCH: Weinheim, Germany, pp 20-21, 1999.

9) a) Bornscheuer, U. T.; Kazlauskas, R. J. Hydrolases in organic synthesis. 2nd ed.; Wiley-VCH Weinheim, **1999**. b) Doson, G.; Wlodawer, A. *Trends Biochem. Soc*. **1998**, *23*, 347-352. c) Hedstrom, L. *Chem. Rev*. **2002**, *102*, 4501-4523.

10) Rutticci-Mulder, J. C. Expression and mutagenesis studies of *Candida antarctica* lipase B. Ph. D. Thesis, KTH, Sweden, **2003**.

11) Strahl-Bolsinger, S.; Gentzsch, M.; Tanner, W. *Biochimeca et Biophysica Acta*. **1999**, *1426*, 297-307.

12) a) Brocca, S.; Persson, M.; Wehtje, E.; Adlercreutz, P.; Alberghinsa, L.; Lotti, M. *Protein Science.* **2002**, *9*, 985-990.

13) Georgiou, G. Expression of protein in bacteria. In "Protein Engineering, principles and practice". Wiley , New York, USA, 1996, pp 101-127

14) Binnie,; C. Cossar, J. D.; Steward, D. I. H. *Trends Biotechnol*. **1997**, *15*, 315-320.

15) a) Houdebine, L. M. *J. Biotech*. **1994**, *34*, 269-287.

16) Chen, Z.; Zhao, H. Encyclopedia of chemical processing. 2006, pp 2467-2477.

17) Carter, P. *Biochem. J*. **1986**, *237,* 1-7.

18) Hoston, R. M.; Cai, Z. L.; Ho, S. N.; Pease, L.R. *Biotechniques* **1990**, *8*, 528-535.

19) Carter, P. *Biochem. J*. **1986**, *237*, 1-7.

20) van den Burg. B.; Vriend, G.; Veltman, O. R.; Venema, G.; Eijsink, V. G. H. *Proc. Natl. Acad. Sci. USA*. **1998**, *95*, 2056-2060.

21) Arnold, F. H. *Acc. Chem. Res*. **1998**, *31*, 125-131.

22) a) Cadwell, R.; Joyce, G.; *PCRMeth.Appl*. **1991**, *2*, 28–33. b) Leung, D.; Chen, E.; Goeddel, D. *Technique*, **1989**, *1*, 11–15.

23) Bechman, R.A.; Mildvan, A.S.; Loeb, L.A. *Biochemistry*, **1985**, 24, 5810-5817.

24) Bebenek, K.; Roberts, J.D.; Kunkel, T.A *J.Biol.Chem*. **1992**, 267, 3589-3596.

25) Raza, S.; Fransson, L.; Hult. K. *Protein Science* **2001**, *10*, 329-338.

26) Phillips, R. S. *Enzyme Microb. Technol*. **1992**, *14*, 417-419.

27) Fujima, Y.; Hirayama, Y.; Ikunaka, M.; Nishimoto, Y. *Tetrahedron: Asymmetry* **2003**, *14*, 1385-1391.

28) Park, S. Rational approach for improving enantioselectivity and altering reaction mechnism of hydrolases. AD. Thesis, SSWU, Korea, **2010**.

29) Shin, C. J.; Chen, C. S.; Fujimoto. Y.; Girdaukas, G. *J. Am. Chem. Soc*. **1982**, *104*, 7294-7299.

ABSTRACT

Substrate and protein engineering of *Candida antarctica* lipase B for improving enantioselectivity toward alkyl tetrahydrofuran-2-carboxylates

Jeemin Park

Department of chemistry

Graduate School of

Sungshin Women's University

Enzymes are highly efficient catalysts with extraordinary enantio- and regioselectivity, and they can accept a wide range of complex molecules as substrates. In particular, lipases are the most employed catalysts in organic synthesis to yield optically pure compounds through kinetic resolution. Obtaining enantiomerically pure compounds are important because they are a key factor in pharmaceutical industrial fields.

Recently, advance in tailoring enzymes for improving activity and selectivity and combined use of enzymes with chemocatalytic reactions have expanded the role of biocatalysis to produce enantiopure compounds from racemic mixtures. Directed evolution or rational design techniques have proven to be successful for the development of enzymes with either enhanced or inverted enantioselectivity compared to their parental enzymes.

Enantiopure tetrahydrofuran-2-carboxylate (THFC) is an important intermediate as well as a building block in the pharmaceutical industry and research. Especially (*R*)-tetrahydrofuran-2-carboxylate is the first compound that is incorporated into a penem skeleton to give furopenem, a clinically efficient non-natural *β*-lactam antibiotic.

In this thesis, enzymatic kinetic resolution method is used to produce enantiopure tetrahydrofuran-2-carboxylate. Hydrolysis reaction of various esters of tetrahydrofuran-2-carboxylate (THFC) was performed by *Candida antarctica* lipase B (CAL-B). However, parental enantioselectivity of CAL-B is not high enough (E = 3.4) to produce enantiopure THFC. To improve enantioselectivity of CAL-B, enzyme engineering approach using site-directed mutagenesis is employed with substrate engineering. Through substrate engineering, enantiopreference of CAL-B wild-type is changed from (*S*)-selective to (*R*)-selective. And mutanted enzymes like I189Q (E = 13.5) and V190I (E = 11.3) showed higher enantioselectivity than wild-type (E = 5.6).