In compliance with the
Canadian Privacy Legislation
some supporting forms
may have been removed from
this dissertation.

While these forms may be included
in the document page count,
their removal does not represent
any loss of content from the dissertation.

Studies of Selective Chemical Catalysis by Hydrolases

Harjap Singh Grewal
McGill University
Department of Chemistry

A thesis submitted to the Faculty of Graduate Studies and Research of McGill University in partial fulfillment of the degree of Master of Science

Submitted August, 2002

Your file Votre référence
ISBN: 0-612-88205-5
Our file Notre référence
ISBN: 0-612-88205-5

"It is a miracle that curiosity survives formal education."

Albert Einstein

Acknowledgments

I would first like to thank my supervisor, Dr. Romas Kazlauskas, for all his help, support and guidance during my studies. I would also like to thank my parents for all of their support. I would like to thank the people that I have collaborated with during my time here, namely Seongsoon Park, Krista Morley and Eniko Forro. I would also like to thank those who worked previously on my projects, namely Miguel Ferrero, Vicente Gotor and their research group in Spain, Neil Somers, Geoff Horsman and Dave Dietrich. Finally I would like to thank all of the people in the Kazlauskas research group that I have met and in particular Seongsoon Park for all his help.

Abstract

Hydrolase catalyzed reactions are used for selective chemical catalysis. With directed evolution, rational mutations and molecular modeling the selectivity of these hydrolases can be increased and the origin of selectivity determined. This study investigates four selective hydrolase catalyzed reactions. Using molecular modeling the unusual regioselectivity of *Pseudomonas cepacia* lipase (PCL) and the selectivity of *Candida antarctica B* lipase (CAL-B) in nucleoside acylation reactions has been attributed to the binding of the nucleoside base within the active site. The enantio-selectivity of *Pseudomonas fluorescens* esterase (PFE) has been improved using a rational approach to directed evolution and models to explain the origin of improved mutants has been produced. The enantioselectivity of the β-lactam ring opening reaction by CAL-B has been attributed to unfavorable steric interaction of the substrate with Ile189 and a model has been proposed for an alcohol bridge between the catalytic histidine (His224) and the lactam amine. Finally, high acetyl selectivity of ThermoGen esterase E018b has been demonstrated and reaction conditions optimized.

Collaborators

All research for this thesis was carried out under the supervision and with the collaboration of Dr. Romas Kazlauskas.

Chapter 2

Vincente Gotor and Miguel Ferrero.

Chapter 3

Seongsoon Park, Krista Morley and Geoff Horsman.

Chapter 4

Enikö Forró, Seongsoon Park and Ferenc Fülöp.

Chapter 5

Neil Somers and Dave Dietrich.

Table of Contents

Chapter 4. Modeling of β-lactam Hydrolysis by CAL-B

Chapter 5. Acetyl Selective Hydrolysis by ThermoGen Esterase E018b

Chapter 1: Introduction

Hydrolases

In biological systems hydrolases are essential enzymes for digestion of a variety of compounds. Sugars, lipids and proteins are all digested by specific hydrolases by cleavage through the addition of water, or hydrolysis. Apart from their essential role in biological systems, hydrolases have also been adapted for commercial and industrial tasks. Although naturally their catalytic activity most often involves hydrolysis of substrates, because of their ability to hydrolyze a variety of substrates, they have been well adapted for alternate reactions in industry. For example hydrolases have been used in the production of cheese, in detergents, for synthesis of drugs and other industrial molecules.[1]

Hydrolases are a large class of enzymes found in both prokaryotic and eukaryotic organisms and displaying a wide variety of reactions. The simplest way to classify hydrolases is by the reaction that they catalyze. Lipases and esterases hydrolyze ester bonds. These enzymes conform most with the structural details outlined above. Lipases catalyze the hydrolysis of large lipid esters, while esterases catalyze the hydrolysis of small water soluble carboxylic esters. Proteases hydrolyze peptide bonds of either small peptides or of large globular proteins, creating smaller subunits. Other hydrolases include glycosidases (which hydrolyze sugars), epoxide hydrolases and nitrile hydrolases. This study focuses on serine hydrolases, namely lipases and esterases that employ a nucleophilic serine as a catalytic residue.

Many hydrolases (particularly lipases) catalyze the reverse condensation reactions as well. And because hydrolases are quite flexible with regards to accepting a large variety of substrates they have also demonstrated the ability to carry out a variety of condensation reactions such as alcoholysis or aminolysis, Figure 1.

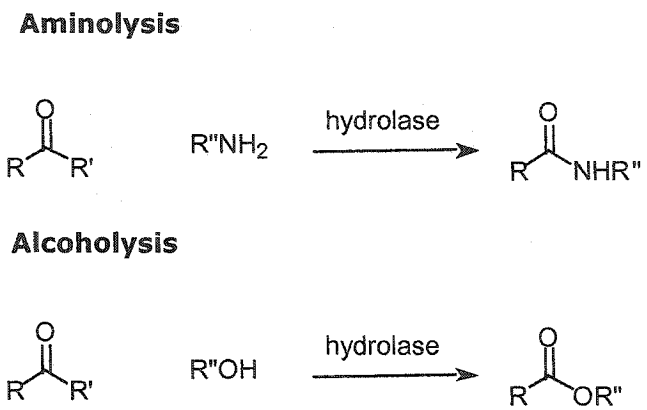


Figure 1: Aminolysis and alcoholysis catalyzed by hydrolases

In general, for the natural reaction of serine hydrolases, water is added across the bond, which is cleaved by the enzyme releasing the substrate from the catalytic residue the enzyme. Along with catalyzing reactions in with a similar mechanism, many of the structural features of serine hydrolases are also conserved, particularly among lipases and esterases. The 3-D structure of these hydrolases is a α/β-hydrolase fold, Figure 2. Within this fold, the locations of the key catalytic residues are conserved. The oxyanion stabilizing residues and the catalytic triad, composed of the nucleophilic residue, stabilizing histidine, and acid, are all located on the same or similar loops of the α/β-hydrolase fold. The catalytic triad itself is also similar among many hydrolases. Lipases

and esterases in particular all contain a nucleophilic Ser residue stabilized by His and an acid (either Asp or Glu). The enzyme folds in such a manner to orient these residues close to one another, which partially explains the generic fold. In addition, the oxyanion stabilizing residues must also be nearby. These residues form the oxyanion hole within the active site of the enzyme, within which the oxyanion is stabilized by hydrogen bonding with backbone amides and/or the side chain of the amino acids.

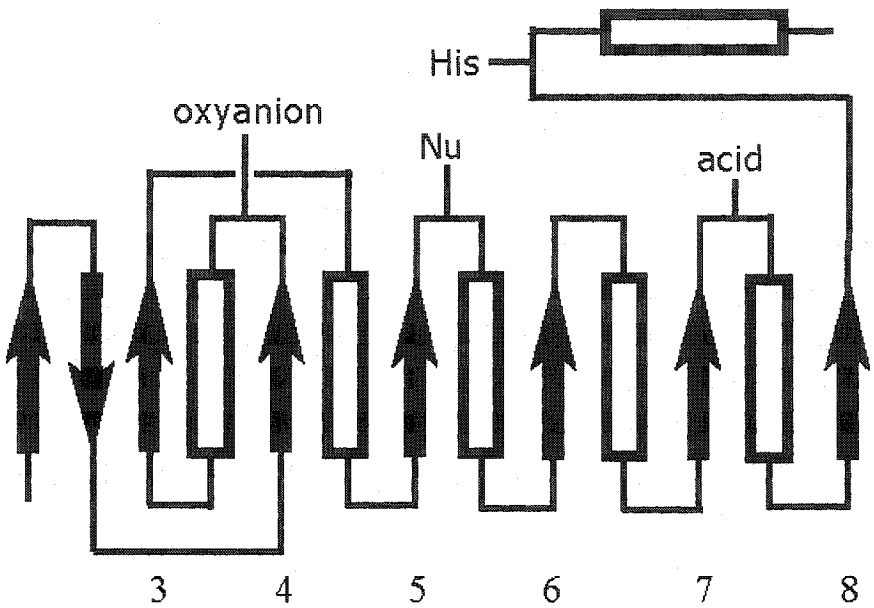


Figure 2: α/β-hydrolase fold. Basic folding of serine (Nu = Ser) hydrolases is conserved and key catalytic residues are also located on the same loop regions shown above.

Lipases have very hydrophobic active sites, which are normally covered by a lid structure when the lipase is in solution. Once the lipase comes in contact with a hydrophobic or lipid interface the active site of the enzyme is exposed by a

conformational change of the lid structure. By exposing the active site, the activity of the lipase is increased. This phenomena is referred to as interfacial activation.[2]

The reaction mechanism, Figure 3, proceeds through an initial nucleophilic attack by the catalytic residue (Ser) of the triad on to the carbonyl carbon at the point of hydrolysis of the substrate. This process is mediated by the catalytic acid and the His residue which abstracts the Ser proton in concerted process with the nucleophilic attack. The proton is then donated to the leaving group (alcohol, amine) of the hydrolyzed substrate. The hydrolase is then in an acyl-enzyme intermediate complex, which is hydrolyzed by the addition of water allowing for the release of the acyl moiety of the substrate as an acid. Hydrophobic regions and polar regions of the active site accommodate the specific substrates of the enzyme with specific binding regions for each group of the substrate.

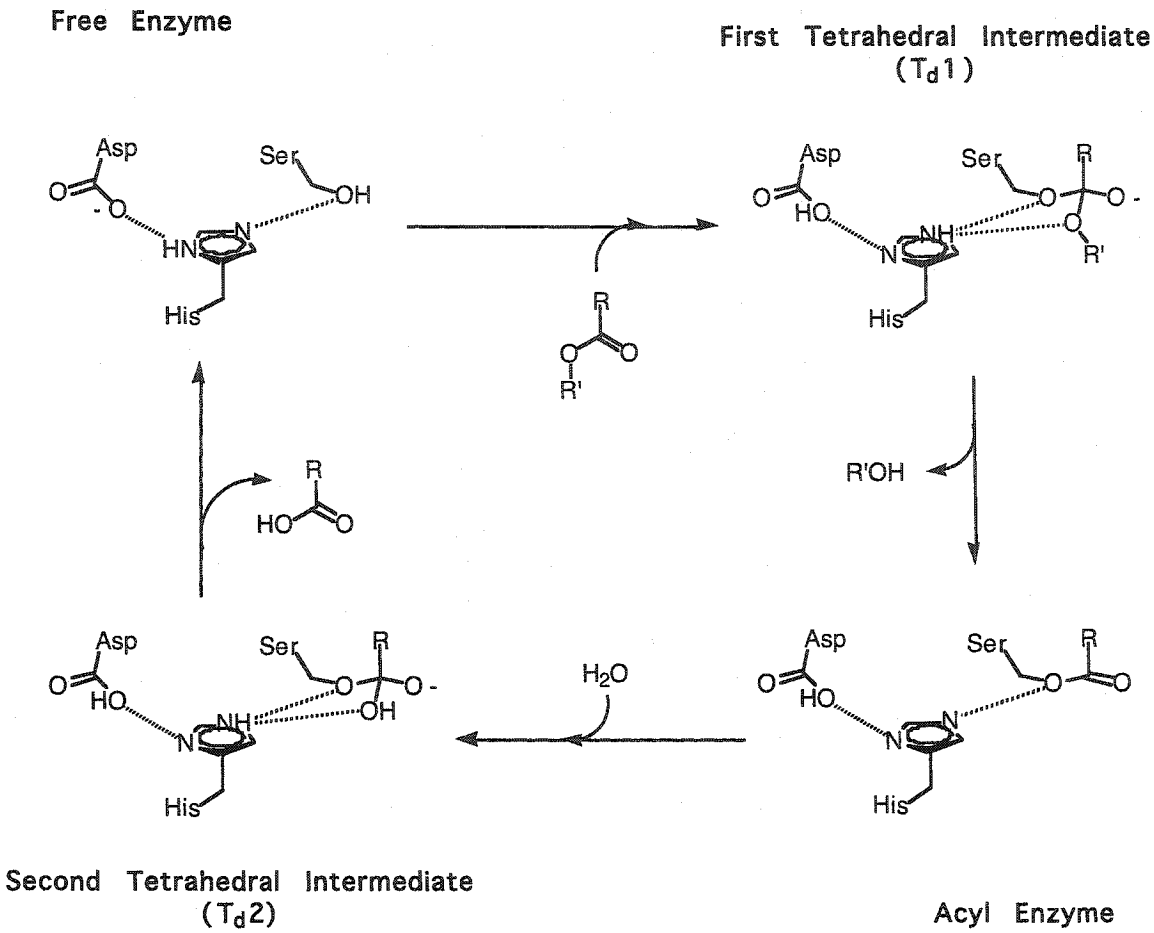


Figure 3: General reaction mechanism for hydrolases. The mechanism of most hydrolases is very similar. It is initiated by nucleophilic attack by the catalytic site (Ser) creating the first tetrahedral intermediate. The collapse of this intermediate produces the acyl enzyme. Nucleophilic attack by a water molecule creates the second tetrahedral intermediate, the collapse of which provides the free enzyme. Both nucleophilic attacks are assisted by hydrogen bonding within the catalytic triad.

Although, with their flexibility for substrates, serine hydrolases can be used for many reactions, their specific catalytic structures allow them to catalyze these reactions with some enantioselectivity or regioselectivity.

Hydrolases in this Study

Pseudomonas cepacia Lipase (PCL)

PCL (also known as *Burkholderia cepacia* lipase) is a bacterial lipase, which has been cloned and expressed from 4 different strains. The enzyme is 320 amino acids long and weighs 33 kDa. Several x-ray crystal structures are available of the lipase and they show that the lipase has a typical α/β-hydrolase fold.[3] The enzyme has a large substrate-binding site with two large hydrophobic regions. The catalytic triad of the enzyme is Ser87-His286-Asp264 and the backbone amides of Gln88 and Leu17 contribute to the stabilization of the oxyanion. There is a large lid structure that covers the large hydrophobic active site, a structure that is common to many lipases. Naturally catalyzing the hydrolysis of lipids, the active site is understandably hydrophobic. All lipase substrate-binding regions are quite hydrophobic and exhibit an increase in catalytic activity through interfacial activation. PCL shows a ~25 fold increase in activity when bound to a water-organic interface.

Candida antarctica B Lipase (CAL-B)

CAL-B is a fungal lipase that is 317 amino acids long and weighs 33 kDa.. Several x-ray crystal structures with a variety of substrates have been obtained of CAL-B which show a α/β-hydrolase fold.[4] The enzyme has one large hydrophobic region in the active site with a restricted alcohol binding region. The catalytic triad of the lipase is a Ser105-His224-Asp187 structure. The oxyanion hole has Gln106 hydrogen contributing a backbone amide and Thr40 contributing a backbone amide and side chain hydroxyl for hydrogen bonding with the oxyanion. CAL-B does not have a large lid and does not

show interfacial activation, which is unusual for a lipase. With no large lid structure and no interfacial activation CAL-B is similar to many esterases.

Pseudomonas fluorescens Esterase (PFE)

PFE is a bacterial esterase and is classified as an arylesterase. The esterase is 272 amino acids long and weighs 29.5 kDa. Presently there are no x-ray crystal structures available of the enzyme, however, a homology model of PFE was obtained from the structure of non-heme haloperoxidases.[5] These haloperoxidases have a α/β-hydrolase fold and exhibited 46-51% sequence identity with PFE. Models obtained with this degree of identity are generally quite good, for example, x-ray crystallographers use enzymes of similar identity as a scaffold for new enzyme structures. The active site of the enzyme is very restricted and has little exposure to the solvent. There are a large number of hydrophobic residues around the active site, the most common being Phe residues, the size of them being partially responsible for steric restrictions around the active site. The catalytic triad is also a Ser95-His252-Asp223 and the oxyanion hole contains two stabilizing backbone amides of Met96 and Trp29 hydrogen bonding with the oxyanion. There is, however, uncertainty in the structural analysis of PFE as the current structure available is a homology model.

ThermoGen Esterase E018b

Limited information is available about this enzyme and the sequence has not yet been determined. E018b is from a library of esterases of thermophiles.[6]

Enzymes for Selective Catalysis

Enantioselectivity

Nature provides numerous examples of enantiomers, which are two objects that are mirror images of one another and cannot be super-imposed on one another. We can find examples from our hands to molecules within us. However, nature does not provide many examples where both of the enantiomers are present, usually only one. This is especially true for the key nutrients used in biological system, such as amino acids and sugars. The use of specific enantiomers in these cases is vital for biological processes, such that if the opposite enantiomer is provided the organism cannot function. It is quite amazing that the natural processes of biology can discriminate between compounds that are identical with regards to construction and chemical properties so accurately. Often times, not only will the opposite enantiomer not be incorporated into the biological system, but will cause adverse side effects. Such discrimination not only demonstrates the complexity and intricacy of biological systems, but also suggests a reason for concern when humans develop new unnatural compounds, particularly for biological use.

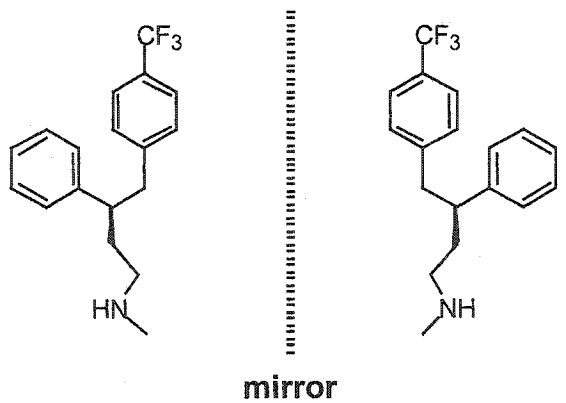


Figure 4: Fluoxetine enantiomers. The above molecules are examples of enantiomers,

which are mirror images of identical molecules that cannot be superimposed. Only the (*R*)-enantiomer (right hand side above) is the active form of the drug Prozac®, which is a racemic mixture of both enantiomers.

Many drugs that have been developed have demonstrated to have one enantiomer, which not only is inactive, but has adverse effects as well. For example, Prozac®, chemically known as fluoxetine, is a drug which is a racemate, that is a mixture of the two enantiomers, Figure 4. The (*R*)-fluoxetine enantiomer of the drug is the active form, while the other enantiomer is assumed to be the cause of many of the unpleasant side effects of the drug.[7] Due to these adverse effects the FDA now requires any drug having one enantiomer as it's active form to be produced as a single pure enantiomer.

Since the requirement of such compounds by the FDA there has been a large amount of research devoted to new and efficient ways of producing enantiopure compounds. One of the most effective methods for obtaining many enantiopure compounds is to isolate them from nature. This process allows one to develop a collection of various chiral compounds, or a chiral pool, of potential drug candidates or chiral building blocks for drug candidates. However, natural products cannot always be directly used as drugs, they do not afford all the chiral molecules that are of synthetic interest and often the process is not economically feasible. As is common among many human initiatives, it is very difficult to mimic nature, especially without its assistance. But due to the constant pursuit by chemical and pharmaceutical companies advances have been made.

Chemical methods have historically required a multi-step processes to arrive at enantiopure and biologically active compounds. However, with rapid advances in the field of chemical asymmetric synthesis[8], many successful reactions have been obtained.

Organometallic asymmetric synthesis involves chiral agents linked to a metal complex. The conjugate metal ion acts as the catalyst of the reaction while the organic chiral having favored interaction with one of the enantiomers of the substrate. However, these processes are still limited with regards to accommodating a large variety of substrates.

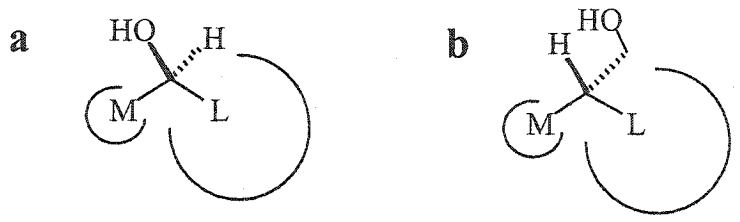
Another chemical synthesis approach for obtaining enantiopure compounds is by use chiral starting materials and performing reactions, which conserve or reverse the chirality of the compound. This method is limited by the availability of chiral starting material often provided by a chiral pool of natural compounds.

Separating two enantiomers is another method to obtain enantiopure compounds. Techniques include crystallization by using enantiopure seed crystals, diastereomer formation and chiral chromatography. These techniques are limited by chemical properties of the substrate and availability of chiral auxiliary agents.

Kinetic resolution is a technique used to increase the enantiomeric purity using enzymatic or chemical catalysis. Such a resolution is employed for reactions, which are selective, but not necessarily highly selective, for one of the enantiomers. During these reactions, one enantiomer of the substrate will react faster than the other causing the product of one to be in higher concentration than the other. One must then sacrifice their yield and stop the reaction before completion, in order to prevent both enantiomers from reacting to completion. As the reaction precedes the product of the fast enantiomer will accumulate. The ratio of the products of the fast enantiomer and slow enantiomer will reach a peak before the reaction is complete. Ideally, the reaction is stopped at this peak, often around 40 to 50% conversion, and extracted. The resolutions can be repeated to increase the enantiomeric purity of the product. [9]

Enzymes are often good catalysts for enantioselective synthesis and resolutions. Hydrolases, having already been used successfully in many industrial applications, have also become of interest as alternatives for chiral biocatalysis. The natural chiral selectivity of biological agents serves to be a favorable feature of hydrolases. In addition to their flexibility for a variety of substrates, there is a very large variety of hydrolases, which are available commercially and they can perform a variety of reactions. And possibly one of the best characteristics is that enzymes are environmentally friendly, unlike their organometallic counter parts.

Hydrolases are the most common enzymes used for resolutions and among them lipases, esterases and proteases are most often used. Lipases in particular have been used and studied extensively.[10] A set of empirical rules has been developed for lipases, including the two in this study, in order to be able to predict the enantio-preference of each lipase and which would be best to employ for a particular reaction. These empirical rules have been developed for primary alcohols, secondary alcohols, and carboxylic acids.[11] The rules are based on the binding of large and small substituents of primary and secondary alcohols within conserved regions of the active site. Along with CAL-B and PCL, *Candida rugosa* lipase, *Pseudomonas aeruginosa* lipase, *Pseudomonas fluorescens* lipase and *Rhizomucor miehei* lipase all display similar binding domains. They contain a large hydrophobic binding region, the natural binding domain of the lipid chain, and a medium pocket near the catalytic histidine. Residues at the chiral center of the substrate bind, according to their size, in accommodating binding regions of the lipase, hence determining the selectivity of the lipase for a particular enantiomer, Figure 5. The rules apply best for the secondary alcohol.



secondary alcohols

primary alcohols
(no O at stereocentre)

Figure 5: Empirical rules for lipase selectivity. The above diagrams are used to predict the fast acting enantiomer of primary and secondary alcohols for PCL. a) shows the preferred binding enantiomer for secondary alcohols. b) shows the preferred enantiomer of the primary alcohol. The rule does not work as well for the primary alcohols, particularly if there is an oxygen atom at the stereocenter.

Regioselectivity

Another problem facing synthetic chemists is that of selectively modifying a single functional group on a molecule containing multiple functional groups. The problem becomes even more difficult when the same functional groups exist within a molecule, the only difference being their location on the molecule. This is most evident in sugar chemistry, with multiple alcohol groups present within these molecules. The best and most effective method to over come this problem chemically has been the use of protecting groups. However, selective protection of these groups can be at times difficult by chemical means. Furthermore, many protection and deprotection synthetic protocols are very laborious and result in low yields.

The problem can be addressed by the use of enzymatic catalysts. Lipases are most frequently used in regio or chemo selective reactions, but there are some examples of esterases being employed for this purpose as well.[12] Though lipases generally only increase the chemical selectivity of the reaction many synthetically useful applications of

enzymes in such reaction have been reviewed[13] and the area of research is still progressing. In this study we investigate two reactions, which have unusually high chemo and regio-selectivity. The first is the acylation of nucleosides by PCL and the second is ester hydrolysis using ThermoGen esterase E018b.

Screening

Even with the empirical rules for lipases and with extensive data available for other hydrolases it remains difficult to predict the ideal enzyme for new substrates. Under these circumstances, a fast and reliable screening process of an enzyme library is required. Furthermore, such a screening process assists in the screening of multiple substrates for a single enzyme, allowing one to determine the variety of substrates for which the enzyme can be used (substrate mapping of the enzyme). High-throughput screening has developed into a very good option for identifying both enzymes for substrates and substrates for the enzymes. There have been many reliable developments in this field with regards to technique and instrumentation.[14]

One simple method developed to measure the activity of an enzyme towards a substrate is the use of high-throughput colorimetric screening. The technique uses a chromaphore containing compound as a substrate so that the progress of the reaction can be monitored via a color change. Such reactions can be performed in 96 well plates and observed using a micro-plate reader, but only yield and estimated enantioselectivity. This technique can be applied for measuring the enantioselectivity of enzymes as well, by measuring reaction rates of the two enantiomers in separate wells. Estimated enantioselectivities have been measured by using IR thermographic imaging to monitor

the reaction rates of the two enantiomers in separate wells. [15] Although this technique provides some indication of the enantioselectivity, or enantiomeric ratio (E or E value), of the enzyme, the values obtained are a good estimate at best. This is because the screening process neglects competition between enantiomers for the binding site of the enzyme by having them in separate wells.

The endpoint method is a more accurate measure of the E value.[16] This method involves having both enantiomers in the same reaction vessel. The reaction is stopped at approximately 40 % conversion, at which point the accumulation of the product of the fast reacting enantiomer is sufficient and the reaction has not progressed so far that most of the slow enantiomer has also reacted. Although this technique provides a very accurate E value, it generally is very laborious and cannot be used for the screening of a large number of samples at one time. The Reetz group has, however, recently developed some high throughput screening techniques for the endpoint method. [17, 18]

Quick E

Quick E[19] provides both a fast and reliable method to measure the E of a particular reaction using the true substrate, Figure 6. The technique is a high-throughput colorimetric screening method, similar to that of the Reetz group. The key difference is that this method adopts a second chromophore containing compound, which reacts with the enzyme. This compound acts as competition for the two enantiomers in their separate wells.

For this screening method *p*-nitrophenol is used as the pH sensitive indicator to measure substrate hydrolysis, the color change of which can be measured at 404 nm. A second chromophore, the reference compound, competes with the substrate for the active

site of the enzyme, and after hydrolysis, its color change can be measured at a different wavelength (574 nm for resorufin acetate). With the presence of this reference compound the rate of the two enantiomers can be compared to calculate a more accurate value of the enantioselectivity.

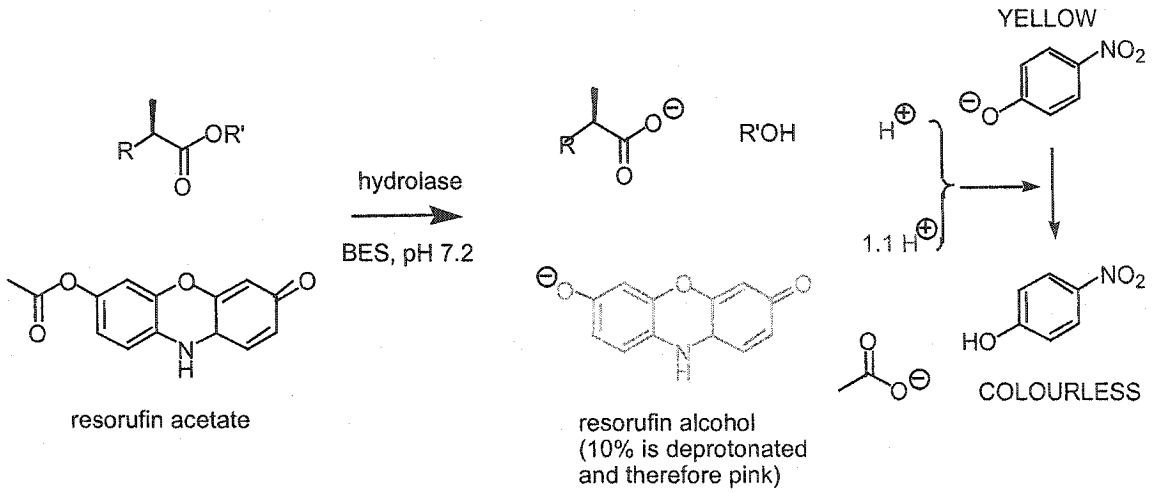


Figure 6: Quick E method for measuring enantio-selectivity. A pH sensitive indicator monitors the hydrolysis of the substrate. The color change of *p*-nitrophenol from yellow in the ionised form to colourless in the protonated form can be monitored at 404 nm. The reference compound, resorufin acetate, changes from an orange to pink colour and can be monitored at 574 nm. By having the reference compound, competition for the active site is introduced and the rate of reaction of two different enantiomers, relative to the reference compound, can be compared.

Increasing Hydrolase Enantioselectivity

Directed Evolution

Although hydrolases catalyze reactions of non-natural substrates, the selectivity is often poor. Researchers need high selectivity to obtain a single, very pure product after one reaction. An enantioselective catalyst, for example, requires an E value above 30 to be synthetically useful. E values above 100 would be considered to be very good. However, after screening new compounds one often finds E values well below 30.

To improve these E values directed evolution of the enzyme has offered a good solution. Naturally enzymes have developed by evolution to carry out their specific reactions. There is significant sequence identity among many hydrolases. The enzymes have adapted over time, with environmental pressure, mutations in the sequence and the natural selection process of these mutations, to carry out their specific reactions with high efficiency. Following the example of nature, researchers now use directed evolution to not only mimic this process, but also do it quickly.

Directed evolution[20] involves the introduction of mutations to the gene of a hydrolase. Random mutagenesis is used to introduce variation in the amino acid sequence of the hydrolase, creating a library of new mutant hydrolases. By limiting mutations to the gene of the hydrolase, researchers can identify the best mutant for a specific reaction using high-throughput screening. This mutant is then used as the parent gene for the next cycle of mutagenesis, effectively mimicking the selection process of nature and allowing for continued improvement until a satisfactory stereo or regio-selectivity is obtained. With the availability of high-throughput screening and sequencing, it is possible to identify improved mutants and characterize them quite rapidly.

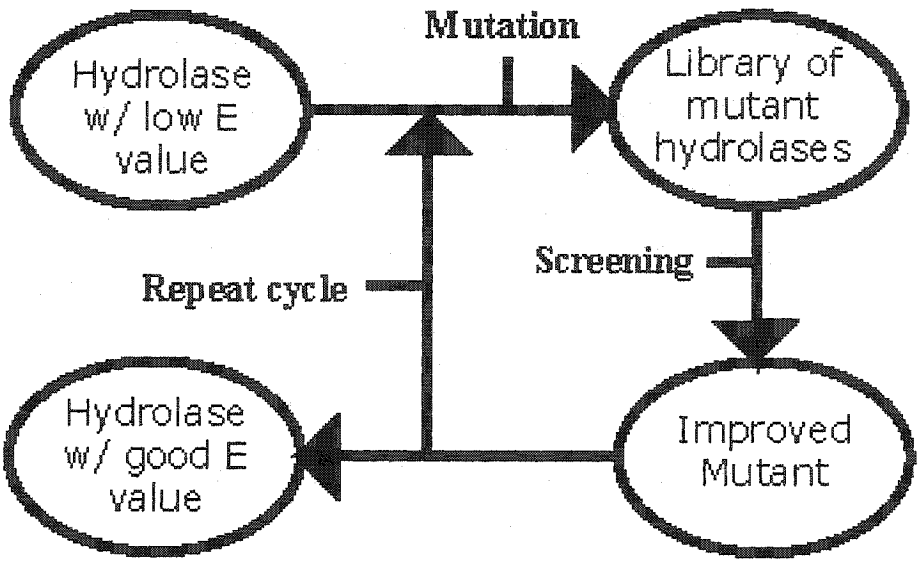


Figure 7: Process of directed evolution for a more enantioselective hydrolase. The process of directed evolution uses repeated mutation and screening of a hydrolase to produce hydrolases with improved enantioselectivity. The process can be repeated with improved mutants until an acceptable enantioselectivity is obtained.

Because this process uses random mutagenesis, it is very useful for manipulating enzymes for which little to no structural information is available. However, mutations occur throughout the gene and there are 20 possible amino acid that can be incorporated into each position. In order to be certain that all possible mutants have been tested the number of mutants that must be screened becomes quite high. If more than one mutation occurs in a single mutant enzyme, the screening numbers become even larger and it becomes difficult to determine which mutation is causing the improvement in selectivity. Therefore it is best to limit mutations to one amino acid substitution at a time in order to be able to easily characterize the mutants.

Rational Approach

In the case of hydrolases whose 3D structure is available, a rational approach to directed evolution can be employed. By rationally deciding what amino acids to target for mutagenesis, a limited number of mutants are created in the library and the amount of screening required is decreased. Furthermore, by targeting amino acids around the active site of the enzyme the chances of observing changes in the selectivity of the enzyme are increased due to the proximity of the residues to the substrate. Similar to the directed evolution approach, the improved mutants are identified and used as the parent gene for another round of rational mutagenesis.

The rational approach involves computer aided modeling of the hydrolase structure with the substrate. By looking at the substrate bound as a reaction intermediate, one can identify amino acid residues that are likely to effect the interaction of the substrate with the enzyme. With the rational approach one of these specific amino acids is randomly mutated and the library is then screened to identify improved mutants. Researchers often try to predict what specific amino acid would likely increase the selectivity of the enzyme. This process, protein engineering, involves mutation of the target amino acid to a single residue. The rational approach is a hybrid of the protein engineering and directed evolution strategies.

Molecular Modeling

Computer aided molecular modeling is becoming a more popular and more powerful tool for investigating chemical reaction and interactions. With the development of new computer hardware and software there is an ability now to do very complex and

large molecular calculations and simulations. Modeling has allowed researchers to investigate molecular interactions, vibrations and orientations, as well as enzymatic and chemical transition states. Modeling provides a new avenue to investigate problems or queries in chemistry that are impossible to investigate by chemical methods.[21]

One good example of how developments in computer modeling have helped chemists is in its simplest application, the visualization of 3D structures. Structural models of molecules are constantly formed to investigate intramolecular and intermolecular interaction. Like conventional, hand-held ball and stick models, which are used, computer models can now be used to visualize small molecules in free space, rotate them and manipulate bond angles. One advantage offered by computer modeling is in the depiction of macromolecules, such as proteins and DNA. Using conventional hand built models for such structures presents a problem in both the work required to build them and in the storage of such structures. With computer modeling, the building of large molecules has become easier and storage is now negligible.

With protein structures, many programs have been developed that allow for various visual displays of the molecules, from simple displays representing domains of the protein as β-sheets or α-helices, to detailed structures showing individual atoms, partial charges and hydrogen bonds.

With the integration of experimental data with these simple computer models, a much more powerful tool has developed. Besides allowing researchers to visualize models of molecules, the bond angles and bond lengths of the molecule can be predicted, molecular surfaces can be displayed, and the natural vibration and free movement of molecules can be mimicked. Another key function of this technology is to calculate the

energy of these molecules. The energy can be calculated for single molecules, assemblies of molecules and for reaction between molecules.

Molecular mechanics is one method with which energy values are calculated. These calculations use a hypothetical molecule with optimal structural characteristics (such as bond lengths and angles). Attributes of the molecule are based on experimental data and the molecules are portrayed as balls on springs. These calculations, however, are limited to comparing different conformations of the same molecular structures and complexes. More advanced quantum mechanics calculations can afford values for dipole moments and heat of formation along with energy values.

Many different fields of research have used molecular modeling to help advance their knowledge of chemical and biochemical processes. Simple molecular mechanics calculations can be used to determine molecular conformations, to determine targets for rational engineering of proteins and to model pharmacophores (spatial representation of target binding moieties of a substrate) for drug design. Besides providing possible protein modifications to change the selectivity of enzymatic reactions, these calculations can suggest modification for the substrate and reaction conditions. More advanced quantum mechanics help researchers investigate transition states of chemical reactions.

Drug design is one of the most intriguing applications of molecular modeling. Without structural information large amounts of screening of a variety of different compounds is required to find a "hit". Even with the large, time consuming amounts of screening the library of substrates is not often sufficient to provide a lead compound. Furthermore, many molecules have been patented and are not available for screening,

limiting the number of possible lead compounds available. Molecular modeling provides a solution using rational drug design.

By mapping the active site of an enzyme one can develop a pharmacophore for the particular enzymatic target. With this model, different functional groups can be added to the model and optimized in order to provide candidates for lead compounds. This process minimizes the amount of screening required by narrowing the number a candidates for potential drug molecules. The first drug to be developed by rational design was Captopril[22], an angiotensin converting enzyme inhibitor.

X-ray Crystal Structures of Hydrolases

Structures of proteins have been difficult to obtain by NMR. Although two and three dimensional NMR studies are being used currently to resolve enzyme structures, one of the best techniques available is x-ray crystallography.[23] The technique requires highly structured crystals of the protein and is the limiting factor in the determination of most crystal structures. Once the crystals are obtained they are showered with x-rays, the diffraction pattern of which is used to determine the structure of the protein with computer assistance. The structure determination is often accelerated by using the coordinates of similar proteins (having some sequence homology) of a known structure to help guide the process.

Crystal structures are also used to study the binding of substrates and to investigate the catalytic structures of serine hydrolases. Phosphonate inhibitors attach themselves covalently to the nucleophilic serine residue to inhibit the enzyme. By crystallizing a lipase with a phosphonate inhibitor attached, a crystal structure with a bound substrate and solvent exposed active site (open conformation of the lid) is

obtained. By manipulating the structure of the inhibitor, the binding of different substrates can be mimicked. However, crystallization of the hydrolase with an inhibitor can sometimes be difficult. Furthermore, insufficient amounts of solvent and the tight stacking of hydrolases within a crystal may lead to a poor structural representation of a hydrolase in solution. Also, the orientation of the inhibitor can effected by the stacking of the hydrolase with one another. For these reasons, the structures obtained are not always good representations of the catalytic complex of the substrate and hydrolase in solution. However, there are experimental results that can provide some insight with regards to the binding of the substrate.[24]

A library of crystal structures is available at a web-based protein databank (www.rcsb.org/pdb). The files are all in PDB format (extension *.pdb) and can be viewed using most macromolecular modeling programs. PDB files use Cartesian coordinates (XYZ format).

Transition States and Phosphonate Analogs

The transition state of a chemical and biochemical reaction determines the selectivity of the reaction. During the transition state of a reaction, substrates and catalysts must be in a fixed and restricted position in order to interact with one another. Other steric and electronegative interactions between the hydrolase and substrate away from the reaction center are critical during the transition state as well. Particular conformations and structures away from the reaction center will interact well with one another, while others will have poor interactions. These interactions will determine the selectivity of the reaction.

The interactions, or binding, of the substrate during the transition state of the reaction determine the selectivity of the enzyme. For instance, during the transition state, a single enantiomer of a chiral substrate may bind well to a particular hydrolase due to the chirality present in the structures (binding domains) of the active site. However, the modeling of transition states is not possible with simple molecular mechanics calculations. Quantum mechanics calculations can be used, but due to the difficult and complex structures of the enzyme substrate complex, the calculation would be very large, even for some computers today.

In order to use molecular mechanics, tetrahedral intermediates of the catalytic reaction of hydrolases are modeled. The first tetrahedral intermediate (T_d1) is preferably used because the complete substrate is still attached to the enzyme complex. In the second tetrahedral intermediate (T_d2) the alcohol or amine leaving group has already been displaced. The problem that is presented by the tetrahedral intermediate is that the common molecular mechanics force fields do not recognize a carbon-oxyanion moiety.

To model these tetrahedral intermediates, phosphonate analogs are used to mimic these structures. The covalently linked phosphonate-hydrolase complexes have the same tetrahedral geometry of the substrate in a T_d1 state. Furthermore, atomic partial charges of the phosphonate can be adjusted to mimic those of a carbonyl tetrahedral intermediate. Hence, phosphonates are considered to be good analogs of the T_d1 structure[25], and are used to model these intermediate.

Structure Optimization

One of the most powerful functions that molecular modeling offers is the ability to optimize the geometry of a substrate in a particular conformation. The optimization, or

minimization, functions search for and identify a local energy minima for molecules and enzyme complexes. The process involves repeated angle manipulation and bond stretching of all chemical entities selected for optimization, while calculating the energy for each conformation. With enough manipulation (or iterations of the calculation) the process identifies a low energy conformation, the local minima.

For small and simple molecules, such as an ethane molecule in free space, the process works well, finding all different conformation and bond lengths and eventually the lowest energy conformation of the molecule, the global minima. For molecular complexes and large molecules the process becomes more complicated. Many programs cannot overcome eclipsing interactions within a substrate except for the eclipsing of hydrogen atoms. Therefore, many calculations, without assistance, can only find the local energy minima rather than the global energy minima for the molecule or complex. In order to obtain a global minima, computer aided or manual conformational searching is required.

When optimizing the geometry of enzyme-substrate complexes a step-wise optimization is ideal. First the complete enzyme, with hydrogens and solvent molecules, must be relaxed step-wise from crystallographic coordinates in order to avoid drastic changes in the enzyme structure caused by unusual crystalline conformations or errors within the structure. To avoid such drastic changes, the amino acid side chains are first optimized keeping the enzyme backbone fixed. Similarly, when introducing a substrate, an unusual or high energy conformation of the substrate may cause enzyme conformation to change drastically. Again, to avoid this, the substrate is first optimized with the enzyme fixed, then the amino acid side chains are optimized. After taking these

precautions and obtaining a relatively low energy complex, the entire enzyme-substrate complex can be optimized.

Force Fields

Force fields are sets of atom types, parameters and equations, which allow software to develop simplified models of molecules. They provide the software with atom types, experimental data associated with these atom types, parameters that associate energy values to molecular structures and sets of equations used to calculate the energy with the data provided. Force fields are derived from experimental data and are created to most effectively represent particular types of molecules. For example, MM2[26] force field is for organic molecules while AMBER[27] is optimal for proteins.

The reason why AMBER is popular with enzyme structures, though it was originally developed for nucleic acids, is because the force field is more accurate in its treatment of partial charges on atoms and electrostatic effects. This is especially important for enzymatic models because, along with Van der Waals forces and hydrophobic interactions, electrostatic interaction and hydrogen bonding interactions determine the fold of the enzyme. More importantly, these interaction are a critical components of the active site and the catalytic activity of the enzyme. CVFF[28] and CFF91[29] are two other force fields developed for proteins and peptides and incorporate electrostatic interaction.

Conformational Searching

Conformational searching is often required during geometry optimization in order to investigate all possible conformation of a molecule and find the global minima. The process of conformational searching involves manipulation of dihedral angles to produce

all possible staggered conformation along every bond having free rotation. This can be accomplished either by manual manipulation of the angles or by random generation of conformations by a software protocol (Monte Carlo method[30]).

Manual conformational searching often requires a lot of time, particularly with long flexible molecules. The simplest way to obtain all possible conformations is to independently adjust each dihedral angle to three staggered conformations 120° apart. However, from personal experience, when investigating molecular conformations of a substrate within an active site of an enzyme the 120° rotations are often not sufficient. Besides the eclipsing energy barriers within the substrate, additional energy barriers are created by the interaction of the substrate with the convoluted surface of the enzyme. Again, these barriers can not be overcome by geometry optimization functions, and therefore, require a more rigorous investigation of substrate conformations. Further more, with long flexible substrates, manipulation of a dihedral angle 4 or 5 bond lengths away from a particular group of the substrate causes large changes in the spatial orientation of that group. Hence, even small manipulations of these distant angles can "fit" the particular group into another pocket or binding domain within the enzyme. Again, because of these compounded effects of an enzyme substrate complex, a rigorous approach to conformational searching is important. This does not necessarily produce an overwhelming amount of conformations as many of the conformations can be discarded due to obvious unfavorable steric contact between the enzyme and the substrate.

References

1 Faber, K. *Biotransformations in Organic Chemistry*, 2nd ed., Springer-Verlag, New York, **1995**.

2 Bornscheuer, U. T., Kazlauskas, R. J. *Hydrolases in Organic Synthesis*, Wiley-VCH, Weinheim, **1999**.

3 Noble, M. E. M., Cleasby, A., Johnson, L. N., Egmond, M. R., Frenken, L. G. J. *FEBS Lett.*, **1993**, 331, 123-128.

4 Uppenberg, J., Hansen, M. T., Patkar, S., Jones, T. A. *Structure*, **1994**, 2, 293-308.

5 Horsman, G. P. *Master of Science Thesis* McGill University, Department of Chemistry, **2001**.

6 ThermoGen Inc., Chicago Technology Park 2201 West Cambell Park Drive, Chicago, Illinois, 60612.

7 Robertson, D. W., Krushinski, J. H., Fuller, R. W., Leander, J. D. *J. Med. Chem.*, **1988**, 31, 171-180.

8 Berrisford, D. J., Bolm, C., Sharpless, K. B., *Angew. Chim. Int. Ed. Engl.*, **1995**, 34, 1059-1070.

9 Chen, C. S., Fujimoto, Y., Girdaukas, G., Sih, C. J. *J. Am. Chem. Soc.*, **1982**, 104, 7294-7299

10 Schmid, R. D., Verger, R. *Angew. Chim. Int.Ed.*, **1998**, 37, 1608-1633.

11 Weissfloch, A. N. E., Kazlauskas, R. J. *J. Org. Chem.*, **1995**, 60, 6959-6969; Bornscheuer, U. T., Kazlauskas, R. J. *Hydrolases in Organic Synthesis*, Wiley-VCH, Weinheim, **1999**.

12 Nagele, E., Schelhaas, M., Kuder, N., Waldmann, H. *J. Am. Chem. Soc.*, **1998**, 120, 6889-6902.

13 Gotor, V. *Bioorg. Med. Chem.*, **1999**, 10, 2189-2197.

14 Reetz, M. T. *Angew. Chim. Int. Ed.*, **2001**, 40, 284-310.

15 Reetz, M. T., Becker, M. H., Kulling, K. M., Holzwarth, A. *Angew. Chim. Int. Ed.*, **1998**, 37, 2647-2650.

[16] Chen, C. S., Fujimoto, Y., Girdaukas, G., Sih, C. J. *J. Am. Chem. Soc.*, **1982**, 104, 7294-7299.

[17] Reetz, M. T., Becker, M. H., Klein, H. W., Stockigt, D. *Angew. Chem. Int. Ed.*, **1999**, 38, 1758-1761.

[18] Reetz, M. T., Kuhling, K. M., Deege, A., Hinrichs, H., Belder, D., *Angew. Chem. Int. Ed.*, **2000**, 39, 3891-3893.

[19] Janes, L. E., Kazlauskas, R. J. *J. Org. Chem.*, **1997**, 62, 4560-4561.

[20] Reetz, M. T. *Tetrahedron*, **2002**, 58, 6595-6602.

[21] Kazlauskas, R., *Science*, **2001**, 293, 2277-2279.

[22] Bodor, N., Buchwald, P. *Medicinal Research Reviews*, **2000**, 20, 58-101.

[23] Baur, W. H., Kassner, D., *Acta Cryst.*, **1992**, B48, 356-369.

[24] Uppenberg, J., Oehrner, N., Norin, M., Hult, K., Kleywegt, G. J., Patkar, S., Waagen, V., Anthonsen, T., Jones, T. A. *Biochemistry* **1995**, *34*, 16838-16851.

[25] Kazlauskas, R. J. *Tibtech*, **1994**, 12, 464-472.

[26] Burkert, U., Allinger, N. L., *Molecular Mechanics*, ACS; Washington D.C., **1982**.

[27] Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Weiner, P. *J. Am. Chem. Soc.* **1984**, *106*, 765-784.

[28] Dauber-Osguthorpe, P., Roberts, V. A., Osguthorpe, D. J., Wolff, J., Genest, M., Hagler, A. T. *Proteins: Structure, Function and Genetics*, **1988**, 4, 31-47.

[29] Maple, J.R., Hwang, M.-J., Jalkanen, K.J., Stockfisch, T.P., Hagler, A.T., *J. Comp. Chem.*, **1998**, 19, 430-458.

[30] Kalos, M., Whitlock, P. *Monte Carlo Methods*, Wiley-Interscience, New York, **1986**.

Chapter 2: Molecular Modeling of Nucleoside Acylation by Lipases CAL-B and PCL

Abstract

Although the lipase from *Pseudomonas cepacia* (PCL) catalyzes the acylation of both primary and secondary alcohols, it shows unusual regioselectivity for the more hindered secondary alcohol of nucleosides. The lipase from *Candida antarctica* B (CAL-B) is selective for the primary alcohol of nucleosides. Computer aided molecular modeling was used to explain the selectivity of these enzymes. A phosphonate is used to mimick the tetrahedral intermediate. Only the tetrahedral intermediate model for the primary alcohol allowed for binding of the thymidine ring in a hydrophobic pocket of CAL-B. In PCL, only the secondary alcohol allowed for the thymidine ring to bind in a hydrophobic pocket. We propose that the binding of the base ring in a hydrophobic pocket of the lipase determines the selectivity. All structures were judged by steric interactions and presence of key transition state interactions.

Introduction

Nucleoside analogs have been of great interest as medicinal canidates for a variety of diseases. 3' modified nucleosides are used as inhibitors of DNA synthesis and are of interest as anti-viral and anti-tumor agents.[1] However, attaining such molecules often requires multistep synthetic approaches involving protection of the 5' alcohol, which can be difficult. Enzymatic catalysis provides a simpler method of producing such derivatives.[2]

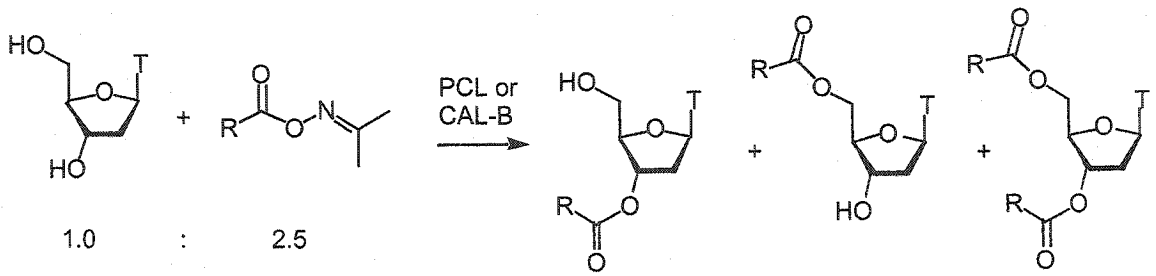
The use of enzymes has been demonstrated to be effective in asymmetric synthesis. *Pseudomonas cepacia* lipase (PCL) has shown to stereoselectivily acylate alcohols.[3] PCL has also shows unusual regioselectivity in the acylation of 2'-deoxynucleosides.[4] While PCL catalyzes the acylation of the more hindered 3' position, a similar lipase, *Candida antarctica* B (CAL-B), prefers the less hindered alcohol at the 5' position. Though the use of enzymes in synthesis is becoming common, the basis of their selectivity is not always well understood. Computer aided molecular modeling is used to provide an explanation for the unusual regioselectivity of PCL and to determine the basis of selectivity of CAL-B. Furthermore, the effect of different acyl groups on the kinetics of the reaction is considered, providing insight into the reactions the enzymes could effectively catalyze.

This work was started as an undergraduate honors project and continued during graduate work.

Results

Lipases CAL-B and immobilized PCL catalyze the acylation of 2'-deoxy nucleosides with different regioselectivity. CAL-B favors the chemically more reactive primary alcohol at the 5' position while PCL favors the more hindered secondary alcohol at the 3' position, eq. 1.[2] Previously 5' selective reactions with CAL-B and 3' selective reactions with PCL have shown good selectivity using oxime esters. A totally regioselective acylation at the 3' position nucleosides with oxime esters and PCL is best performed at 60 °C in pyridine.[5] 5' favored acylation of nucleosides with CAL-B is best performed in THF at 0 °C and yields the desired product in >20:1 ratio.[6] The nucleoside base used in

our study was thymine, however, similar kinetic results have been observed with several other bases.



1.0 : 2.5

T = Thymine

3' acylation 5' acylation 3' & 5' acylation

(1)

Though regioselectivity of CAL-B and PCL towards different nucleosides and different acyl donors has been previously documented.[7] Tables 1 and 2 provide a direct comparison using the same conditions, for the acylation reaction in Equation 1. The rates of reaction and degree of regioselectivity vary with the nature of the acyl group of the oxime ester (acetate, butanoate, decanoate, benzoate, phenylacetate). With both CAL-B and PCL the acyl groups had no effect on the regio-preference of the enzymes. PCL showed 3' regioselectivity in ratios of 40:1 and 98:1 with propyl and nonyl acyl chains respectively while CAL-B yielded 8:1 and 4:1 respective ratios of the 5' product. The regioselectivity of PCL presents an interesting reaction as it favors the more hindered secondary alcohol position. Reaction rates were about 50% lower with benzyl and phenyl groups for both enzymes but the regioselectivity remained quite high. For both enzymes the propyl and nonyl chains yielded the highest rates and selectivities.

Table 1. Regioselectivity of CAL-B catalyzed acylation of thymidine with oxime esters.[a]

R	Time (h)	Conv. (%)	3' (%)	5' (%)	3', 5' (%)	5' Selectivity[b]
Me	4	98	17	68	13	2.7:1
Pr	6.5	94	11	81	2	6.4:1
Non	53.5	99	20	79	0	4:1
Ph	39	12	0	12	0	>12:1
Bn	37	45	0	45	0	>45:1

[a] ~add experimental details~ [b]Selectivity calculated as a ratio 3' and 5' yields with double acylation yields added to each. Data collected by Vincente Gotor and Miguel Ferrero and their research group.

Table 2. Regioselectivity of PCL catalyzed acylation of thymidine with oxime esters.[a]

R	Time(h)	Conv.(%)	3' (%)	5' (%)	3', 5' (%)	3' Selectivity[b]
Me	1	92	44	30	18	1.3:1
Pr	1.58	92	79	2	11	6.9:1
Non	2.75	98	98	0	0	>98:1
Ph	89	28	24	4	0	6:1
Bn	37	42	42	0	0	>42:1

[a] ~add experimental details~ [b]Selectivity calculated as a ratio 3' and 5' yields with double acylation yields added to each. Data collected by Vincente Gotor and Miguel Ferrero and their research group.

To explain the different regioselectivities of the two lipases and especially the unusual preference of PCL for the secondary alcohol at the 3'-position, we used computer modeling. Starting with the x-ray crystal structures of CAL-B and PCL, we modeled analogs of tetrahedral intermediates in the butyrylation of thymidine at both the 5' and 3' positions. These models also rationalize the slow reactions when the acyl group on the oxime is benzoyl (R = Ph) and phenylacetyl (R = Bn). Although the transition state determines the rate and selectivity of a reaction, we focused on the tetrahedral intermediate as an analog of the transition state, Figure 1. Binding of the tetrahedral intermediate with in the active site of CAL-B and PCL and all catalytic hydrogen bonds are shown in Figure 1a and 1b. Further, we modeled a phosphonate analog of this tetrahedral intermediate because it dramatically simplified modeling using molecular mechanics force fields, while still resembling the transition state.[8]

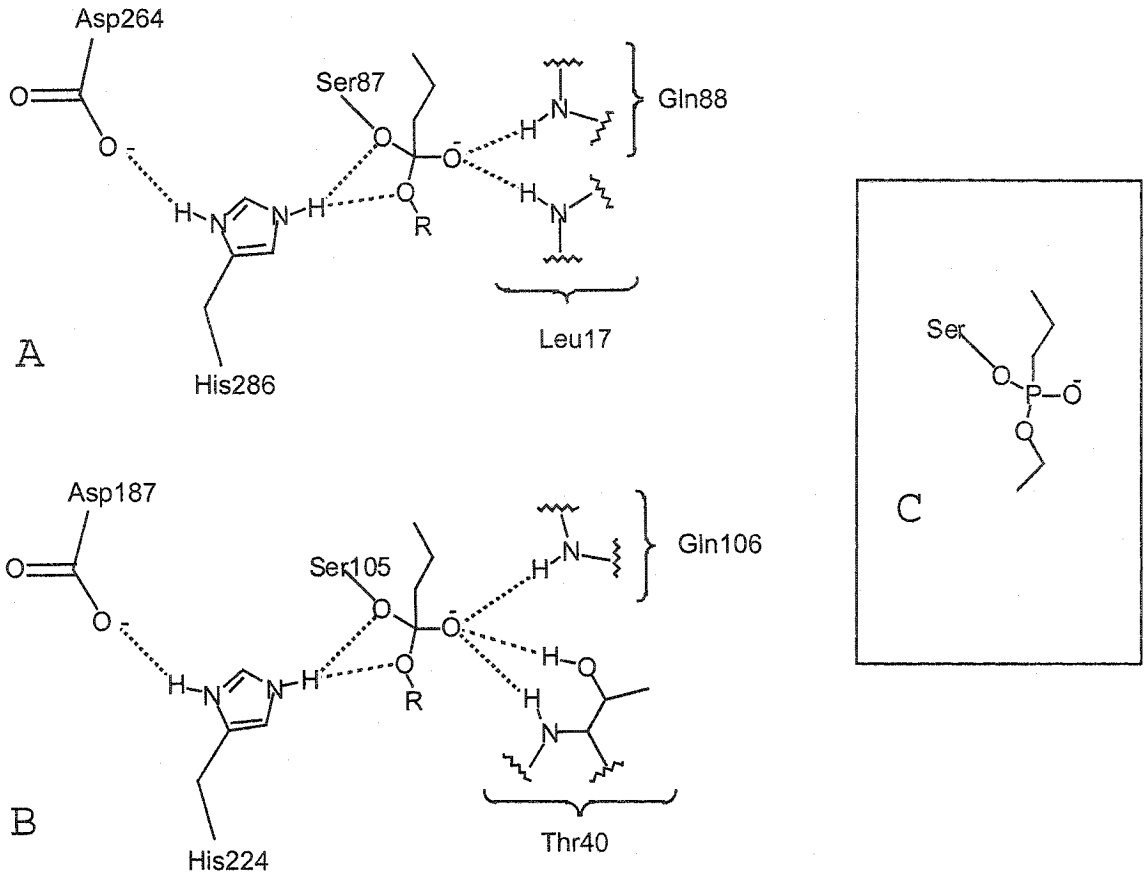


Figure 1: Key hydrogen bonds around the tetrahedral intermediates in CAL-B and PCL. A) In PCL there is one hydrogen bond between Asp264 and His286. There are two H-bonds His286 forms with hydroxyl oxygens of Ser87 and the substrate. Finally, there are two amide hydrogen bonds to stabilize the oxyanion. R is the alcohol leaving group, the nucleoside. B) CAL-B has a very similar hydrogen bond structure, but contains 3 hydrogen bonds with the oxyanion. Aside from two amide bonds, the oxyanion also hydrogen bonds with the hydroxyl group of Thr40. Again, R is the alcohol leaving group, the nucleoside. C) The tetrahedral phosphonate precursor used in the modeling to obtain structures containing key hydrogen bonds in both PCL and CAL-B shown in A and B.

We built the models starting with a phosphonate analog for butyrylation of ethanol, a simple substrate, Figure 1a. These simple phosphonates were constructed for both CAL-B and PCL. After geometry optimization the phosphonate structures contained all key hydrogen bonds required for catalysis, Figure 1b and 1c. Next, we attached and geometry optimized the ribose ring, followed by the thymine ring. We

systematically searched different conformations of these more complex structures and focused on those most likely to mimic catalytically productive transition states. Three criteria determined whether a conformation was catalytically productive: a) the presence of the key hydrogen bonds within the active site, b) lack of obvious steric clashes between the phosphonate and the lipase, and c) lack of unfavorable intramolecular interaction within the phosphonate.

In both lipases, the arrangement of the catalytic residues and the binding sites surrounding them are similar, Figure 2. Viewed with the catalytic triad Asp-His-Ser oriented from left to right, both lipases contain a large hydrophobic pocket above the Asp-His-Ser triad and a medium size pocket below it. The acyl moiety of the substrate lies in the large hydrophobic pocket, while the alcohol moiety points toward the solvent but parts of it may also bind in the medium and/or large hydrophobic pockets. In PCL the large hydrophobic pocket is flanked by residues Val266 and Val267 on the left and by Leu167 on the right, Phe119 at the top and Pro113 deep in the middle of the pocket. The medium sized pocket exists adjacent the catalytic His286 and Leu287. In addition to the large and medium hydrophobic pockets, PCL also contains an additional pocket, an alternate large hydrophobic pocket next to the medium pocket, which is narrow and lies below the catalytic triad. This alternate large pocket, between Ile290, Leu287, Thr18 and Tyr29 is situated to the right of the medium pocket below the catalytic site. The large hydrophobic pocket of CAL-B is wider than that of PCL. This pocket is lined by Ile189 and Val190 on the left, Val154 on the far right, as well as Leu140 and Leu144 at the top of the pocket. Deep in this pocket, Asp134 is on the right and Gln157 on the left. Below the active site of CAL-B there only exists a medium pocket and there is little room for

large substituents. The area below the catalytic Ser105 is crowded by Trp104 below it and the Leu278-Ala287 helix to the right. There is no alternate large pocket below the site and the binding space seems to be small between the catalytic His224, Leu278 and Trp104.

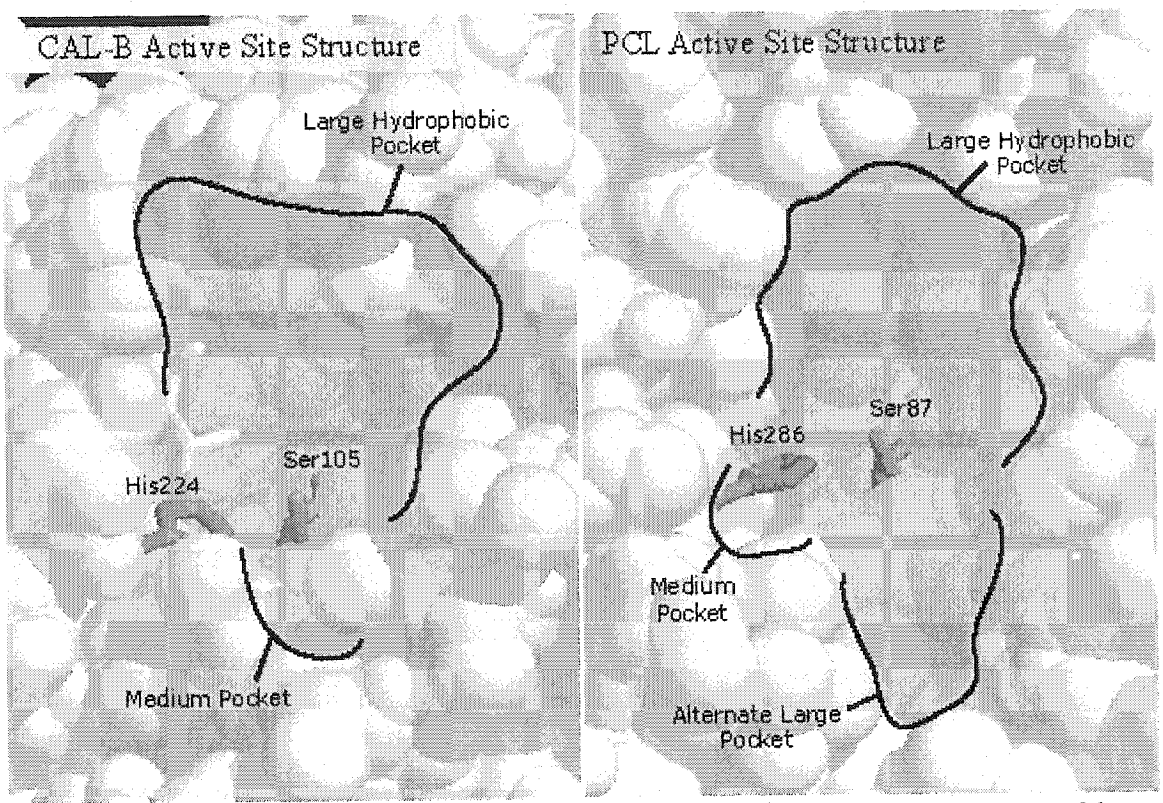


Figure 2. CAL-B and PCL active site structures from x-ray crystallography. The acyl group of the substrate binds first in the large hydrophobic pocket for both enzymes. This situates the substrate with the acyl group above the catalytic triad and the leaving alcohol, or the nucleoside, below the catalytic triad. The active site of CAL-B contains a large hydrophobic pocket above the catalytic residues and a medium sized pocket below. There is very little room for large substituents below the active site of CAL-B. The above image of CAL-B displays Ile189 in a stick representation to allow a better view of the large pocket of the lipase. PCL contains a large hydrophobic pocket (smaller than that of CAL-B) above the the catalytic residues. In addition to a medium pocket below the active site PCL contains a narrow alternate binding pocket.

In general, for both enzymes, active sites below the catalytic triad restrict the binding of the alcohol substrate such that only a few conformations were obtained with

the catalytic hydrogen bonds present. Binding of the thymine group proved to be most critical in both enzymes when determining the most productive substrate conformation.

Table 3: Conformations of phosphonate tetrahedral intermediates of nucleoside acylation with lipases PCL and CAL-B obtained by molecular modeling.[a]

Lipase	Reaction	Structure	Key H-Bonds	Thymine Ring Binding	Unfavourable Intramolecular Interactions
PCL	5' acylation	PCL-5'-A*	5 of 5	Un-bound	None
PCL	5' acylation	PCL-5'-B	5 of 5	Un-bound	Eclipsing[b]
PCL	5' acylation	PCL-5'-C	5 of 5	Un-bound	Eclipsing[b]
PCL	5' acylation	PCL-5'-D	5 of 5	Un-bound	Eclipsing[b]
PCL	3' acylation	PCL-3'-A*	5 of 5	Alternate Pocket	None
PCL	3' acylation	PCL-3'-B	3 of 5	Un-bound	None
CAL-B	5' acylation	CAL-B-5'-A*	6 of 6	Large Pocket (right)[d]	None
CAL-B	5' acylation	CAL-B-5'-B	5 of 6	Large Pocket	Syn-pentane[c]
CAL-B	5' acylation	CAL-B-5'-C	6 of 6	Un-bound	None
CAL-B	5' acylation	CAL-B-5'-D	6 of 6	Un-bound	None
CAL-B	5' acylation	CAL-B-5'-E	6 of 6	Un-bound	None
CAL-B	3' acylation	CAL-B-3'-A*	6 of 6	Near Large Pocket[e]	None
CAL-B	3' acylation	CAL-B-3'-B	6 of 6	Un-bound	None

[a] All models are constructed with a propyl acyl group. All structures are derived from the same phosphonate precursor containing all key hydrogen bonds. Hydrogen bonds are calculated with angles $>120^{O}$ and distances of <3 Å [b] Eclipsing interactions are observed within the substrate conformation. The interaction is seen between the atoms at the 5' position ($O\text{-}CH_2\text{-}C\text{-}R$) [c] Syn-pentane interaction was observed between the 5' hyroxyl and 1' ribose ring carbon [d] The large hydrophobic pocket of CAL-B is quite large and extends to the right of the catalytic triad. [e] The CAL-B-3'-A conformation has the thymine ring in close contact with the enzyme, however does not interact in a hydrophobic area of the enzyme. The ring is situated at the edge of the large hydrophobic pocket. * Indicates most productive conformation obtained for each model.

5' butyrylation catalyzed by PCL (not favored)

The best model of the tetrahedral intermediate, PCL-5'-A, was found to have no intramolecular problems, no obvious steric problems and maintained the 5 key hydrogen bonds. However, this model of the tetrahedral intermediate had the thymine unbound and projecting out of the active site. The 3' secondary alcohol in this structure is facing in the medium sized pocket near the catalytic histidine. Different conformations of the transition state were continually for the presence of all key hydrogen bonds and other steric or intramolecular problems. Manual searching identified four potentially productive conformations for PCL catalyzed butyrylation of thymidine at the 5' position, Table 3. However, none of these conformations contained a thymine ring bound in either the alternate pocket of the large pocket. Orienting the thymine ring in the narrow alternate hydrophobic pocket caused obvious steric contact with Ile290 or Tyr29. Orienting the thymine ring in the direction of the large hydrophobic pocket caused unfavorable intramolecular interactions within the substrate. In two of the structures, PCL-5'-B and PCL-5'-C, the ribose ring atoms eclipsed with the oxygen of the reacting alcohol. The other structure, PCL-5'-D contained a syn-pentane interaction between the 4' carbon of the ribose ring and the alpha carbon of the acyl chain.

3' butyrylation catalyzed by PCL (favored)

The best model of the tetrahedral intermediate, PCL-3'-A, has the thymidine ring situated in the alternate large hydrophobic pocket with with no steric problems. The 5' primary in this structure alcohol is situated in the medium sized hydrophobic pocket. A hydrogen bond between the N1 nitrogen of the thymine and Thr18, which lines the right side of the alternate pocket may be possible. Having only two adjustable dihedral angles

(refer to experimental section) a limited number of conformations were obtained, with only three of them with no steric problems. Of the three sterically unhindered conformation only one, PCL-3'-A, retained all five hydrogen bonds and had the thymine ring bound. The steric problems most common were interactions with amino acids in the narrow restricted region below the active site consisting of the alternate and medium pockets. Due to the trans orientation of the 3'-hydroxyl group relative to the thymine ring, the thymine cannot be situated in the large pocket. With the ribose ring below the active site the 3'hydroxyl group must point up towards the catalytic serine and the thymine ring is forced to face down, being trans to the hydroxyl group. This forces the thyimine ring to be situated below the active site.

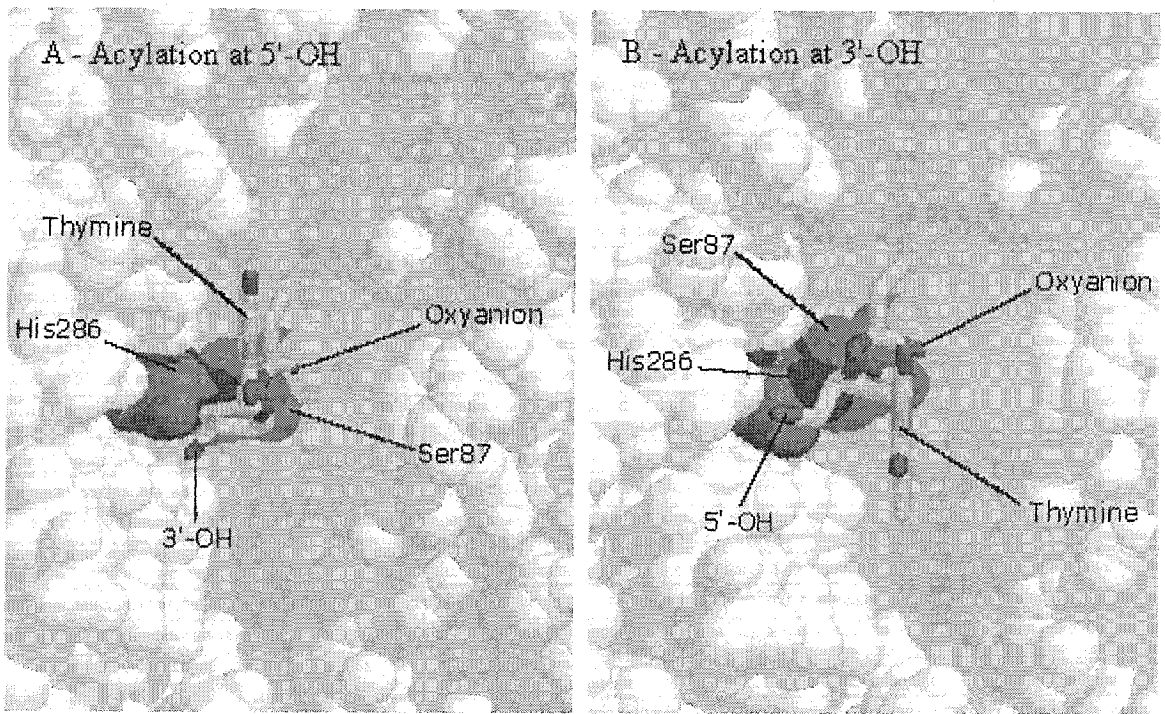


Figure 3. Best models of phosphonates in PCL. The best conformations of both 3' and 5' reactions with PCL are shown. The tetrahedral intermediate of the 5' acylation reaction (A) shows an un-bound thymine ring. The tetrahedral intermediate for acylation at the 3' postion (B) meets all the criteria for a productive conformation. The thymine ring is bound in the alternate hydrophobic pocket below the catalytic triad, suggesting an explanation to the selectivity of the enzyme.

Comparing the structures of 3' and 5' butyrylation in PCL (Figure 3a), the 3' tetrahedral intermediate PCL-3'-A best meets our criteria for a productive catalytic structure. Binding of the thymine ring appears to be most critical. The PCL-3'-A structure has the thymine ring bound in the alternate large pocket of PCL. Thus, the presence of the alternate large pocket in PCL allows for unusual regioselectivity of 3' acylations of 2-deoxy nucleosides.

5' butyrylation catalyzed by CAL-B (favored)

CAL-B-5'-A was selected as the best model of the tetrahedral intermediate having the thymine bound in the large hydrophobic pocket of CAL-B with all six hydrogen bonds and no steric or intramolecular problems. The thymine is bound on the right side of the large pocket, above Gln157. The large hydrophobic pocket of CAL-B is much wider than that of PCL and it allows for the thymine to bind in the upper large pocket without any unfavourable intermolecular interactions within the substrate. Trp 104 restricts the lower region of the CAL-B active site preventing binding of large substituents. There is no large hydrophobic binding pocket below the active site of CAL-B and hence no room for the thymine to bind. Of the large number of conformations explored (>50), only four other conformations were obtained with five or six of the catalytic hydrogen bonds present, not including CAL-B-5'-A. From these four structures with hydrogen bonding, CAL-B-5'-B is discarded due to intramolecular syn-pentane interaction while the others did not have the thymine ring bound within the enzyme active site.

3' butyrylation catalyzed by CAL-B (not favored)

Although no structures were obtained with the thymine ring bound directly in a hydrophobic region of the active site, CAL-B-3'_A was chosen to be the best model, having some Van der Waal contact with the enzyme. Again the nucleoside could not be

situated below the active site because of the limited room. The trans orientation of the 3' alcohol and nucleoside base also prevented the base ring from being situated in the large pocket. Due to these limitations, no structures were found to bind the base ring in a hydrophobic region of the catalytic site. Two conformations were found to have all six catalytic hydrogen bonds, the best of which (CAL-B-3'-A), places the thymine ring at the lower right edge of the large hydrophobic pocket, next to Ile285. The primary alcohol of this model rests in the small area below the catalytic Ser105, between His224, Trp104 and Leu278. This structure was accepted as the best model of the tetrahedral intermediate because the thymine ring does not remain unbound as is the case with CAL-B-3'-B.

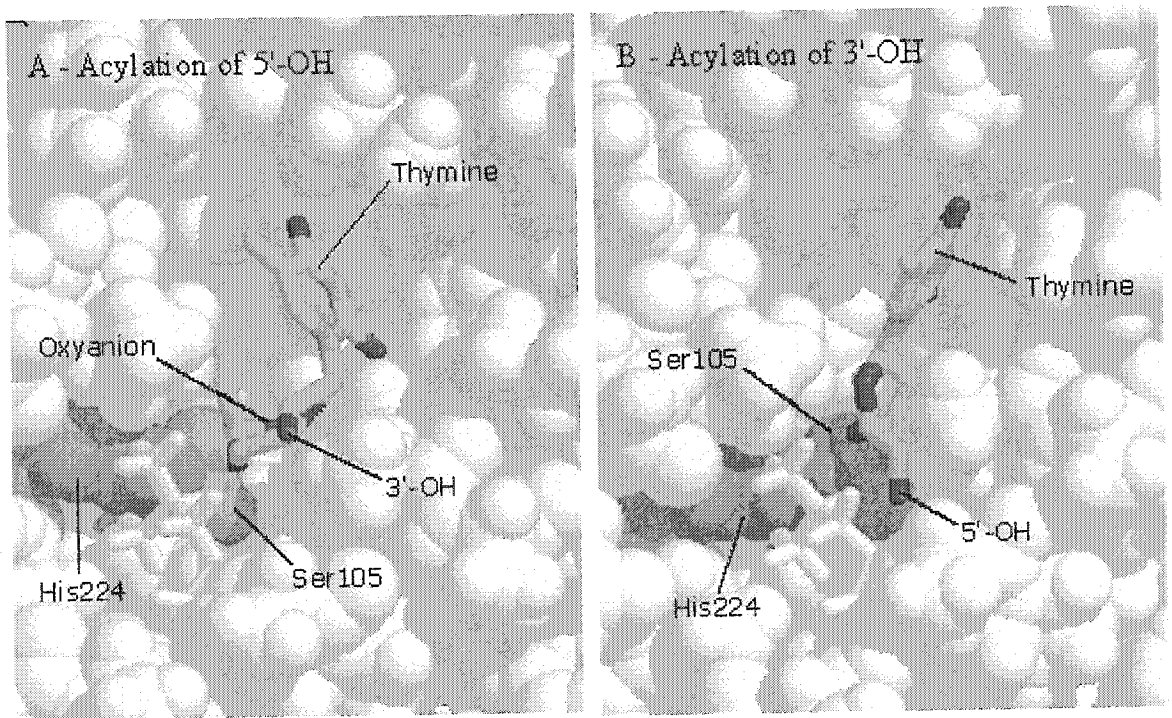


Figure 4: Best models of phosphonates in CAL-B. The best conformations of 5' and 3' transtion states in CAL-B are shown above. The tetrahedral intermediate for the 5' acylation reaction (A) situates the thymine ring on the right side of the large hydrophobic pocket, making it the more productive conformation due to better enzyme substrate interaction. The tetrahedral intermediate for the 3' acylation reaction (B) situates the thymine ring outside the hydrophobic pocket. Again, the binding thymine ring appears to be the key factor in the explanation for the selectivity of the enzyme.

From the models above, binding of the thymine ring appears to be a key determinant of the selectivity of the enzyme. The fast reacting alcohols in both enzymes have the base ring bound to a hydrophobic region of the active site. Comparing 5' and 3' acylation in CAL-B (Figure 3b) shows favorable binding of the thymine ring in the tetrahedral intermediate of 5' acylation. This binding of the thymine ring appears to account for the regio-preference of CAL-B for the 5' position. PCL shows favourable binding of the thymine ring in the tetrahedral intermediate of the 3' acylation reaction and favors acylation at the more hindered 3' position. Hence, the unusual regioselectivity of PCL can be explained by the favorable binding of the thymine ring during acylation at the 3' position.

Acyl chain preferences

The propyl group, used in the models described previously, is quite flexible, and adapts well to the structure of the large hydrophobic pocket. The nonyl group, which wasn't modeled, would be considered ideal for such a reaction, as lipases naturally catalyze the hydrolysis of long chain lipids. Effects of the substrate acyl groups were investigated using benzoate and phenylacetate models of PCL-3'-A (Table 4). In both CAL-B and PCL the large hydrophobic pocket is narrow and restrictive near the catalytic serine. The phenyl and benzyl acyl chains were found to have unfavorable steric interactions which distorted the enzyme due to the rigid structure of these aromatics and the convoluted structure of the large hydrophobic pocket in which they are situated. Though the phenyl ring is restricted in the pocket and afforded no structures without steric problems, the benzyl ring can be situated un-bound and above the pocket without steric problems.

Table 4: Molecular modeling results of benzoate and phenylacetate PCL-3'-A models.[a]

Acyl Chain (R)	*Key H-Bonds*	Binding	Steric Interaction
phenyl	3 of 5	in large pocket	present w/ lipase [b]
benzyl	5 of 5	unbound	no interaction

[a] Models were assembled using the PCL-3'-A structure for PCL acylation reaction at the 3' position of thymidine. [b] Due to steric interactions a planar model of the phenyl ring inside the large pocket was never obtained and the phenyl ring shifts away from the pocket with continued minimization. The steric interaction also distorts the enzyme structure.

Discussion

The unusual regioselectivity for the more hindered 3' alcohol of PCL in it's acylation of nucleosides may be attributed to the presence of an alternate large binding pocket below the catalytic triad of the enzyme. The thymine ring of the 3' tetrahedral intermediate, for the reaction in this study, binds in this alternate pocket indicating that this intermediate binds preferably over the 5' intermediate having an unbound thymine ring. In CAL-B, there exists no such "alternate" binding region, and correspondingly, there are poor binding models for the 3' tetrahedral intermediate. The best binding model for CAL-B, having the thymine ring bound in the large pocket, is that of the 5' acylation tetrahedral intermediate, which is also the preferred acylation site. From these results the binding of the thymine ring appears to be critical in determining the regioselectivity of the two lipases, and the altermate pocket of PCL appears to be responsible for it's unusual regioselectivity.

Added support for this binding model is given by the structure of an enzyme similar to PCL, PPL (lipase from porcine pancreas)[9]. This lipase, like PCL, also contains an alternate binding region below the catalytic triad of the enzyme. PPL catalyzes nucleoside acylation in accordance with the above binding models, favoring 3'

acylation.[10] Again the major structural difference between PPL and CAL-B is the presence of this alternate pocket, as is the case with PCL.

Furthermore, with the binding of the base ring being a primary determinant of the regioselectivity of these lipases, one would expect ribose acylation reactions to proceed with little to no regioselectivity. This is the case for reaction ribose acylation reactions catalyzed by both PCL and CAL-B.[11] From the models obtained, the ribose ring seems quite free to bind in multiple conformations without a base ring. The thymine ring during our conformational search was the cause of most steric problems.

To continue the investigation of this binding model site directed mutagenesis maybe used to explore the importance of the alternate hydrophobic pocket of PCL. As well, investigating the acylation of other substrates with an α-trans orientations (cyclopropanes, dienes) may provide further evidence of our binding model.

Experimental Section

General

The program Insight II, version 95.0, was used for viewing the structures. The geometric optimizations were performed using Discover, version 2.9.7 (*Accelrys*, San Diego CA, USA), using the AMBER[12] force field. The distance dependent dielectric constant was set to 4.0 to mimic the electrostatic shielding of the solvent and the 1-4 van der Waals interactions were scaled to 50%. The crystal structures 3lip[13] (PCL) and 1lbs[14] (CAL-B) were obtained from the Protein Data Bank (www.rcsb.org/pdb/). The CAL-B structure contains a phosphonate while the PCL active site is empty in the crystal structure.

The crystal structures of both enzymes were relaxed before addition of the substrate to these experimental structures. The PDB files 3lip and 1lbs were opened with

all atoms of the crystal structure, including water molecules and structural Ca^{2+} ions. Hydrogen atoms were added to the structures, they were then tested for partial charge balance and corrected for atom types within the AMBER force field. The pH of the catalytic histidine is adjusted to pH 4 to protonate the residue and the structure was then relaxed. Geometry optimizations were carried out in two steps. First, the steepest descent algorithm corrects major high energy structural problems and gets the molecule to a local minima with a RMS deviation of about 0.05 Å mol^{-1}. Second, conjugate gradient algorithm gives a faster and more precise minimization near the local minima, obtaining a RMS value of 0.005 Å mol^{-1} or lower. The entire structure was relaxed by a systematic manner that avoids large structural changes. First only the water and cations were optimized after which the side chains of amino acids were released and optimized as well. The entire complex of enzyme and solvent molecules was then released for the final optimization. For all further geometry optimizations in this study the solvent molecules and structural Ca^{2+} ions were never fixed. This was to ensure that water molecules were mobile and that their position did not hinder the nucleoside or the enzyme.

Geometry Optimization of Phosphonate Core

A phosphonate core of the tetrahedral intermediate (Figure 1a), with butyl and ethyl substituents, was assembled using the phosphonate present in the crystal structure. For PCL the phosphonate core was built on to the catalytic serine. For CAL-B the phosphonate was inverted to and atoms were deleted inorder to obtain our phosphonate core. The partial charges on the phosphonate atoms were set to those of the tetrahedral

intermediate (carbon, not phosphorus), the values of which were obtained by semi-empirical calculation performed using Chem3D using the AM1 parameters. The following partial charges were assigned to the phosphonate: oxyanion = -0.74, phosphorus = 0.58, alcohol oxygen = -0.51, serine oxygen = -0.50 and serine hydroxyl carbon = 0.22.

For the optimization of the phosphonate core a systematic approach of releasing different motifs was adopted. Initially, the phosphonate was allowed to adjust to the enzyme active site by keeping the entire enzyme fixed during the geometry optimization. The side chains of the enzyme were then released and allowed to adjust and finally the entire complex is allowed to adjust to the presence of the phosphonate core. This approach avoids drastic changes in the lipase structure caused by non optimal conformations of the phosphonate. A hydrogen bond calculation is performed after each minimization in order to assure the presence of the critical hydrogen bonds around the catalytic site. To identify hydrogen bonds, we searched for a donor atom to acceptor atom distance of less than 3 Å and a donor atom - hydrogen - acceptor atom angle of 120° or greater.

Phosphonates including thymidinyl group

After obtaining a catalytically productive phosphonate core, the remainder of the nucleoside was and different conformations were produced by manual adjustment of dihedral angles identified in Figure 4. Geometry optimizations were again performed using the same approach used to optimize the phosphonate core with systematic releasing of the substrate, amino acid side chains and the entire enzyme complex. Again,

minimizations are carried out with both the steepest descent and conjugate gradients until an RMS value of at least 0.005 Å mol^{-1} is obtained.

The conformational search involving reaction at the 3'-hydroxyl involved 120^{O} rotations of two dihedral angles identified in figure 4. With these angle adjustments all possible staggered conformations for both dihedral angles were then optimized. For structures of the 5'-hydroxyl reaction, however, one extra dihedral angle is available for adjustment with the primary alcohol. Due to the long and flexible structure of this intermediate small adjustments of the dihedral angles cause large changes in the orientation of the thymine moiety. As well, interaction of substrate conformations with the convoluted structure of the enzyme produces more local minima than a nucleoside in free space. Therefore, manual adjustments of 120^{O} do not prove to be sufficient to find all possible local minima. For the 5'-hydroxyl intermediates a more extensive searching process was required. Adjustments of approximately 10^{O}-20^{O} were performed for both dihedral angles adjacent to the primary alcohol. Structures with obvious steric problems were ignored. Additional manual adjustments were done to orient substrate moieties into visible pockets.

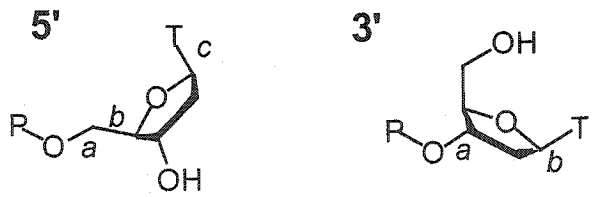


Figure 5. Manipulated dihedral angles for 5' and 3' structures.
Dihedral angles corresponding to all the bonds labeled above were adjusted during manual conformational searching. For most angles, adjustments of 120 degrees were made of each dihedral angle to produce possible conformations. Angle *a* of the 5' structure was manipulated using smaller adjustments to account for large changes in substrate conformations caused by these adjustments. In addition, different conformations of the ribose ring are explored.

Acyl Chain Effects

The models above correspond to the reaction of a butanoyl acyl group (R = propyl in Equation 1). The propyl chain was oriented to fit into the large hydrophobic pocket with a conformation resembling acyl groups from phoshonate crystal structures. Two other acyl groups were modeled for the PCL catalyzed reaction at the 3'-hydroxyl: benzoate (R = Ph) and phenyl acetate ($R = CH_2Ph$). The two acyl groups were added to the secondary alcohol model in PCL, PCL-3'-A and results are summarized in Table 4. The phenyl moiety only has one dihedral angle available for adjustment during manual searching. Adjusting this dihedral angle, which simply involves rotation of the planar phenyl ring, exhibited little difference between structures. Modeling of the benzyl moiety primarily involved adjustment of the dihedral angle adjacent to the phosphorus atom, the other dihedral angle having little effect. The nonyl structure was not modeled, due to the large number of dihedral angles and possible structures, and was considered ideal for lipase hydrolysis reactions.

References

1 Vince, R., Hua, M. *J. Med. Chem.*, **1990**; 33, 17-21.

2 Gotor, V., Morís, F. *J. Org. Chem.*, **1992**, 57, 2490-2492.

3 Weissfloch, A. N. E., Kazlauskas, R. J. *J. Org. Chem.*, **1995**, 60, 6959-6969.

4 Nozaki, K., Uemura, A., Yamashita, I., Yasumoto, M. *Tetrahedron Lett.*, **1990**, 31, 7327-7328.

5 Gotor, V., Moris, F. *Synthesis*, **1992**, 7, 626-628. : Reaction performed in pyridine at 60° and 250rpm for 48-72h. 1.5g or lipase are used for 15mL reaction.

[6] Moris, F., Gotor, V. *J. Org. Chem.*, **1993**, *58*, 653-660. : Reaction performed in THF at 0° and 250 rpm for 8 h. 0.4 g of lipase are used for 15 mL reaction.

[7] Ferrero, M., Gotor, V. *Monatshefter für Chemie*, **2000**, *131*, 585-616.

[8] Kazlauskas, R. J. *Tibtech*, **1994**, 12, 464-472.

[9] van den Berg, B., Tessari, M., de Haas, G. H., Verheij, H. M., Boelens, R., Kaptein, R. *EMBO J.*, **1995**, 14, 4123-4131.

[10] Hermoso, J., Pignol, D., Kerfelec, B., Crenon, I., Chapus, C., Fontecilla-Camps, J. C. *J. Biol. Chem.*, **1996**, *271*, 18007-18016.

[11] Bornscheuer, U. T., Kazlauskas, R. J. *Hydrolases in Organic Synthesis*, Wiley-VCH, Weinheim, **1999.**

[12] Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Weiner, P. *J. Am. Chem. Soc.*, **1984**, *106*, 765-784.

[13] Schrag, J. D., Li, Y., Cygler, M., Lang, D., Burgdorf, T., Hecht, H. J., Schmid, R., Schomburg, D., Rydel, T. J., Oliver, J. D., Strickland, L. C., Dunaway, C. M., Larson, S. B., Day, J., McPherson, A. *Structure*, **1997**, *5*, 187-202.

[14] Uppenberg, J., Oehrner, N., Norin, M., Hult, K., Kleywegt, G. J., Patkar, S., Waagen, V., Anthonsen, T., Jones, T. A. *Biochemistry*, **1995**, *34*, 16838-16851.

Chapter 3: Molecular Modeling and Screening of PFE Mutants

Abstract

An esterase from *Pseudomonas fluorescens* (PFE) has been used to selectively hydrolyze methyl-3-bromo-2-methylpropionate. Using a rational approach four residues were subjected to random mutations and libraries of mutants were screened for improved mutants. Improved mutants were found in Phe199, Val122 and Trp29 libraries. Molecular modeling suggests hydrogen bonding interactions may be an explanation for the improvement in enantioselectivity of the Val122Ser mutant. A second substrate, methyl-2-chloro propionate, was used to screen the library and a Val122Ala mutant was identified to increase in enantioselectivity from 2.6 (wild-type) to 14.

Introduction

Lipases and esterases are natural hydrolases that are interesting for synthetic chemistry. Both have a similar catalytic triad containing a catalytic serine and also contain an oxyanion hole. While lipases generally demonstrate selectivity for the alcohol moiety of esters, esterases are more often selective for the acid moiety of esters.

Esterases are stable in organic solvent and are, therefore, a very useful tool for synthetic chemistry.[1] Their synthetic interest is primarily for asymmetric reactions that cannot be performed chemically with high enantioselectivity. These esterases can also be modified to increase the selectivity through techniques such as directed evolution.

An esterase from *Pseudomonas fluorescens* (PFE) is currently being studied for its enantioselectivity towards a variety of small esters. From preliminary screening results the wild-type enzyme was found to be selective for (*S*)-methyl-3-bromo-2-

methylpropionate (MBMP). PFE was subsequently mutated in order to improve the selectivity, using a rational approach. Although the structure of PFE is not currently available, a homology model has been derived from the structure of non-heme haloperoxidases, which have 46-51% sequence homology with PFE.[2]

By modeling the substrate in the active site of the enzyme 4 amino acids around the active site were selected for mutation. Four libraries were produced of mutants at each site. The first improved mutant was found in the Trp29 mutant library. The mutation was a Trp-to-Leu substitution that improved the enantioselectivity of the enzyme from 12 to 58. Further screening identified numerous improved mutants in the Val122 library. The mutant library selectivity towards both MBMP and methyl-2-chloro propionate is investigated in this study.

Results

Origin of Improved Enantioselectivity of Val-122-Ser Mutant

The best mutant among the Val122 library for MBMP hydrolysis was found to be a Ser mutation (E = 61). Using the homology model of PFE we modeled the substrate and the mutation in order to identify what interactions increase the enantioselectivity of the enzyme. A model of a phosphonate analog of the substrate was obtained with catalytic hydrogen bonds present.

The active site of PFE is very constricted, having many large hydrophobic residues. Conformations of the substrate were primarily affected by rotation of the dihedral angle between the carbonyl carbon and the chiral carbon of the substrate. Conformational searching provided one ideal binding model. The model orients the C3 bromine carbon between Phe199, Phe144 and Ser122 residues, Figure 2. Additionally,

the bromine atom is in hydrogen bonding distance with the hydroxyl group and the Ser mutation. The model does not account for the increase in enantioselectivity observed with the Trp29Leu mutant.

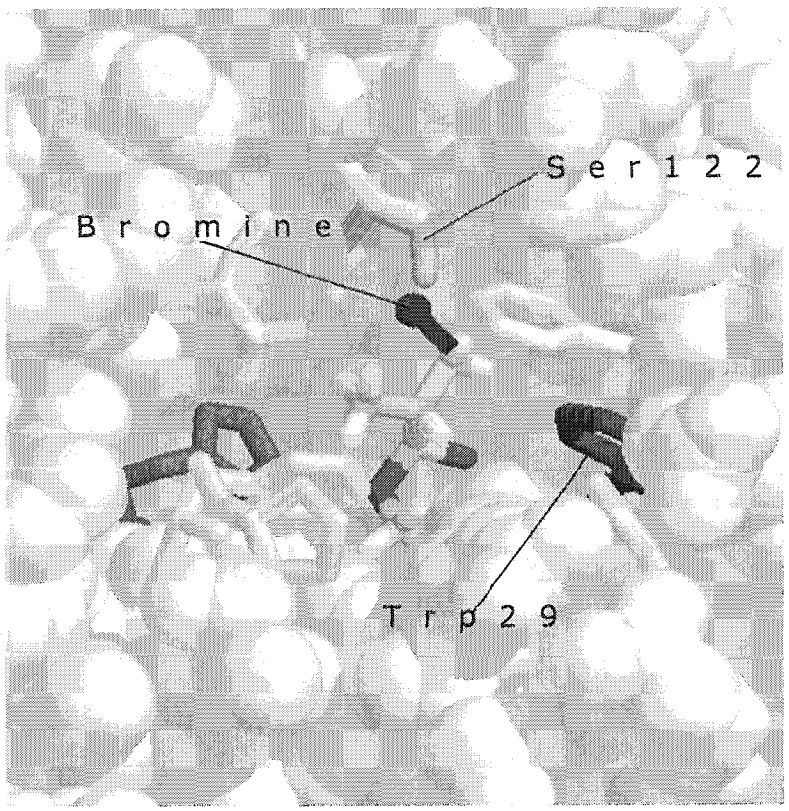


Figure 2: Proposed binding model of MBMP tetrahedral intermediate in PFE. The active site of PFE is very crowded due to large hydrophobic residues. White stick representation of these hydrophobic residues is shown above to allow the substrate to be viewed. The catalytic histidine is in red.

Screening of Val122 Library with methyl-2-chloro propionate

Due to the crowded nature of the PFE active site the enzyme would be expected to best catalyze reaction with small molecules. Methyl-2-chloro propionate was used to screen the Val122 library to identify mutants, which could hydrolyze the substrate

enantioselectively. Quick E was used to screen the compound and identify selective mutants.

The wild-type enzyme had low enantio-selectivity (E = 2.6) for (*R*)-methyl-2-chloro propionate. Upon screening the library many improved mutants were found. The best Quick E value obtained was 14.1. A list of the best mutants is given in Table 1. The best mutants were sequenced to determine the amino acid mutation. Ala was the mutation of the most enantio-selective enzyme while Leu was the mutation of two other improved mutants.

Mutant	Quick E	Mutation
V122-37	14.1	Ala
V122-20	13.0	Leu
V122-11	10.8	Leu
V122-6[a]	10.6	---
V122-48[a]	10.6	---
V122-78[a]	9.4	---
V122-25[a]	6.5	---
V122-80[b]	2.6	Val (wt)

Table 1: Quick E and amino acid mutations from V122 mutant library. Above are the most enantio-selective mutants from the V122 library. [a] Mutants that were not sequenced. [b] Wild-type enzyme.

Discussion

Rational mutation of PFE proved to be successful, providing various mutants with improved enantio-selectivity. The Val122Ser mutation and the Trp29Leu mutations displayed large increases in enantio-selectivity. The origin of these improvements is not certain, but modeling has provided some ideas.

The Val122Ser mutation likely introduces a hydrogen bond with the bromine of the substrate. The size of the Val and Ser amino acid side chains are approximately the same, hence the addition of a polar alcohol group appears to be the major difference between the wild-type and the mutant. A Thr residue would normally be considered to have the same effect, but due to the stereochemistry of the side chain the hydroxyl group cannot be oriented towards the substrate.

The results from the screening of methyl-2-chloro propionate are difficult to explain. The greatest improvement occurs from a Val to Ala mutation, which decreases in size. The mutation would provide more room for binding but the size of the methyl and chloro groups around the chiral center are similar. From Val122Ser mutant model, a polar interaction between the Ser alcohol and the substrate chlorine would have been expected to increase the enantio-selectivity. The improved selectivity is very surprising, however, considering the similarity and size of the chloro and methyl groups.

The Trp29Leu mutation is also difficult to explain. The residue is too far for either polar or steric interactions with the bromine of the substrate in the binding model presented here. A second binding model may account for this improvement in selectivity as the C3 carbon within 4 Å of the Trp side chain.

Experimental Section

Molecular Modeling of MBMP in PFE Val122Ser Mutant

The program Insight II, version 95.0, was used for viewing the structures. The geometric optimizations were performed using Discover, version 2.9.7 (*Accelrys*, San Diego CA, USA), using the CFF91[3] force field. A homology model from previous work was used to represent the structure of PFE.[2]

All optimizations were performed step-wise. The homology model of PFE was first optimized without any substrate present, allowing the amino acid side chains for be optimized first followed by the entire enzyme. A structure was obtained with hydrogen bonding within the catalytic triad of Ser95-His252-Asp223. A phosphonate core with methyl and methoxy substituents was attached to Ser95. This structure was optimized in order to obtain a catalytic structure with all key hydrogen bonds. Again a step-wise approach is used, optimizing first the substrate followed by amino acid side chains and then the entire complex. The entire MBMP substrate was constructed and structures were optimized after manual conformational searching.

Production of Mutant Enzymes

A library was prepared of all active V122 mutants in a 96-well microplate. An overnight culture of this bacterial library was prepared for enzyme production. 96-well blocks with 2 mL wells were used for production of mutant enzymes. To each well 1 mL of LB broth (100 μg/mL) and 10 μL of the overnight cultures was added and incubated at 37 °C and 350 rpm for 3 h. Following incubation 50 μL of rhamnose (4% w/v) was added to induce protein expression as the PFE containing plasmid has a rhamnose induced expression vector. The cultures were incubated for 6 h and then centrifuged (2700 x g at 4 °C) for 10 min. The supernatant was discarded and the cell pellets were resuspended in 400 μL of 5 mM BES buffer (pH 7.2) containing 0.4 mg/mL of lysozyme. The suspensions were incubated for 45 min followed by an overnight freeze (-20 °C) thaw cycle. The lysed cells were incubated with 10 μg/mL of Rnase A and 1.7 μg/mL of DNase I for 15 min and centrifuged (2700 x g at 4 °C) for 15 min. The supernatant was used directly as the esterase source for Quick E.

Screening of V122 library with methyl-2-chloro propionate

The Quick E^4 screening method was used to determine the enantioselectivity of the reactions. The supernatant from protein production in the 96-well blocks was used directly a source of enzyme. 10 μL of the enzyme solution was added to each well of a 96-well microplate. Substrate solutions for both enantiomers were prepare in 5 mM BES buffer (pH 7.2, 0.33 mM Triton X-100) containing acetonitrile, *p*-nitrophenol and resorufin acetate. 90 μL of this solution was added to wells of the microplate giving a 100 μL reaction volume with 7% v/v acetonitrile, 0.11 mM resorufin acetate and 1.0 mM of both enantiomers of the substrate. Substrate and resorufin acetate hydrolysis was monitored at 404 nm and 574 nm respectively in a microplate reader.

Reference

1 Bornscheuer, U. T., Kazlauskas, R. J. *Hydrolases in Organic Synthesis*, Wiley-VCH, Weinheim, **1999**.

2 Horsman, G. P. *Master of Science Thesis* McGill University, Department of Chemistry, **2001**.

3 Maple, J.R., Hwang, M.-J., Jalkanen, K.J., Stockfisch, T.P., Hagler, A.T., *J. Comp. Chem.*, **1998**, 19, 430-458.

4 Janes, L. E., Kazlauskas, R. J. *J. Org. Chem.*, **1997**, 62, 4560-4561.

Chapter 4: Modeling of β-lactam Hydrolysis by CAL-B

Abstract

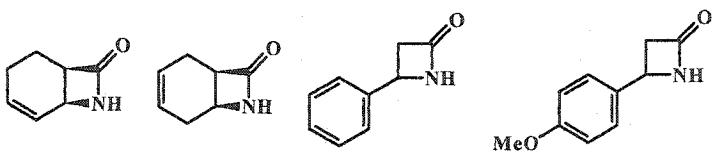
β-Lactams have been used as precursors for enantiopure β amino acid synthesis. A selective ring opening reaction is catalyzed by a lipase from *Candida antarctica* B (CAL-B). The reaction results show dependence on temperature, the solvent used and the presence of a secondary alcohol. The reaction is also highly enantioselective. The enantioselectivity of the reaction is explained by molecular modeling. Unfavorable steric interactions with the (S)-enantiomer of the substrate seem to determine the enantioselectivity. There is also a proposed model of an alcohol bridge between the catalytic His224 and the amide of the lactam ring. This model allows all possible hydrogen bonds to contribute to the stabalization of the tetrahedral intermediate.

Introduction

β-Lactams have been of interest and used as antibiotics for many years, penicillin being a common example. Preparation of these inhibitors as pure enantiomers has been achieved via lipase catalyzed reactions.[1] More recently β-lactams have been of interest as synthetic precursors for the synthesis of β-amino acids.[2] Consequently, enantioselective ring opening of β-lactam rings is now being investigated and various methods have been developed. Enantioselective catalysis of activated lactams has been obtained by enzymatic methods.[3]

The research presented here focuses on obtaining enantiopure β-amino acids directly from unactivated β-lactams using *Candida antarctica* B lipase (CAL-B).as a

enzymatic catalyst. A variety of β-lactams were used, however, compound 1 in Figure 1 was the focus of the research.



(±)-1 (±)-2 (±)-3 (±)-4

Figure 1: β-Lactams tested with CAL-B for selective ring opening. Compound 1 is the focus of this study.

Results

CAL-B was used as the catalyst for the ring opening reactions. For compound 1 CAL-B was selective for the (*R*)-enantiomer. The reaction rate is quite slow in all cases, however, specific modifications to experimental conditions cause an increased rate and enantio-selectivity of the reaction. The reaction rate was higher with *iso*-propyl ether than with toluene. The reaction rate is also ideal at 60 °C. The reaction was slow in water and the enantioselectivity was poor. Among the alcohols used as nucleophiles in the ring opening reaction, 2-octanol was the best, having good enantioselectivity (E = 129) and reaction rate. The results obtained by Eniko Forro (Institute of Pharmaceutical Chemistry, University of Szeged, Hungary) are summarized in Table 1.

Table 1. Novozym 435-catalyzed ring opening of (±)-1 (0.05 M) in water and in the presence of different alcohols (R'OH) in organic media (1:15, v/v) at different temperatures

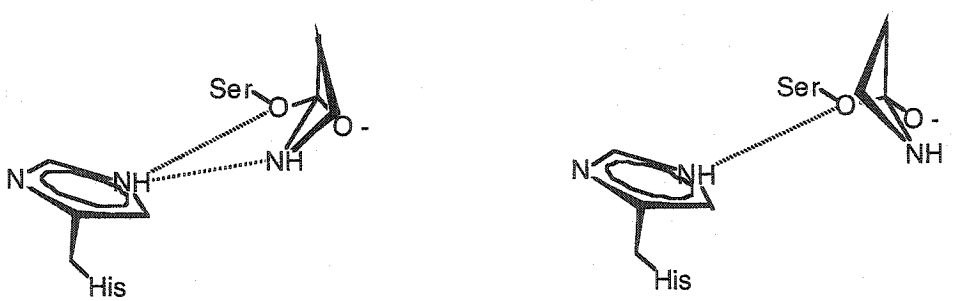
Row	Temp. (°C)	Novozym 435 (mg mL^{-1})	Solvent : R'OH	Time (h)	Conv. (%)	ee_s[a] (%)	ee_p (%)	*E*
1	60	20	H_2O	72	36[b]	32	57	5
2	60	20	Toluene:CH_3CH_2OH	44	2	2	>95[c]	>40
3	60	20	Toluene:Cl_3CCH_2OH	44	16	18	>95[c]	>47
4	60	20	Toluene:$CH_3(CH_2)_6OH$	44	14	16	>95[c]	>46
5	60	20	Toluene:$CH_3(CH_2)_{11}OH$	44	31	43	>95[c]	>60
6	60	20	Toluene:C_6H_5OH	44	**No reaction**			
7	60	20	Toluene:$(CH_3)_2CHOH$	48	3	3	>95[c]	>40
8	60	20	Toluene:$(ClCH_2)_2CHOH$	44	5	1	19[c]	~1
9	60	20	Toluene:$(BrCH_2)(CH_3CH_2)CHOH$	44	36	*Rac*	*Rac*	
10	60	20	Toluene:$CH_3(CH_2)_5CH(CH_3)OH$	44	34	48	>95[c]	>63
11	60	20	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$	44	36	53	>95[c]	>66
12	60	30	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$	43 66	39 47	61 84	>95[c] >95[c]	>73 >104
13	60	30	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$ + Et_3N	43 68	39 46	71 81	>95[c] >95[c]	>73 >98
14	60	50	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$	43	48	89	>95[c]	>117
15	60	75	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$	44	49	92	>95[c]	>129
16	60	10	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$	43	15	17	>95[c]	>46
17	70	10	*i*Pr_2O:$CH_3(CH_2)_5CH(CH_3)OH$	40	28	36	>95[c]	>56
18	70	10	*i*Pr_2O:(+)-$CH_3(CH_2)_5CH(CH_3)OH$	40	28	37	>95[c]	>56
19	70	10	*i*Pr_2O:(-)-$CH_3(CH_2)_5CH(CH_3)OH$	40	28	37	>95[c]	>56
20	60	20	Toluene:$C_6H_5CH(CH_3)OH$	46	20	24	>95[c]	>49
21	60	20	Toluene:$(CH_3)_3COH$	46	14	16	>95[c]	>45
22	60	30	*i*Pr_2O:[CH_3CH_2OH: $CH_3(CH_2)_5CH(CH_3)OH$ (1:1)]	48	14	8	~49[c]	~3

[a]According to GC. [b]Determined by NMR. [c]Calculated by using an internal standard (*n*-decane).

The ideal conditions for selective β-lactam ring opening are at 60 °C in *iso*-propyl ether with 2-octanol as the nucleophile. Molecular modeling was used to explain these specific conditions of catalysis.

Modeling was first used to explain the enantioselectivity of the reaction. Few possible conformations of the first tetrahedral intermediate of the substrate are available to explore. The most significant conformational change is the pucker of the lactam ring,

Figure 2. There are two possible puckered conformations of the ring. Both puckered conformations were investigated with both enantiomers.



Conformation 1 **Conformation 2**

Figure 2: Puckered conformations of lactam ring. The lactam ring has two puckerd conformations. Conformation 1 orients the lactam amine towards His224 allowing a hydrogen bond while conformation 2 orients the amine away from His224.

In both of the puckered conformations, the (*S*)-enantiomer has steric contact with Ile189. A model for the tetrahedral intermediate of the (*S*)-enantiomer was not obtained without obvious steric problems. The (*R*)-enantiomer provides a less hindered model of the tetrahedral intermediate for both puckered conformations. Ideally, having the amine of the lactam ring oriented towards the catalytic histidine (His224) would be ideal, Figure 2. However, having this conformation, the hydrogen bond between the oxyanion and Gln106 is lost, which may account for the slow reaction rate. A conformation having a hydrogen bond with the Glu106-oxyanion does not allow hydrogen bonding between His224 and the lactam amine.

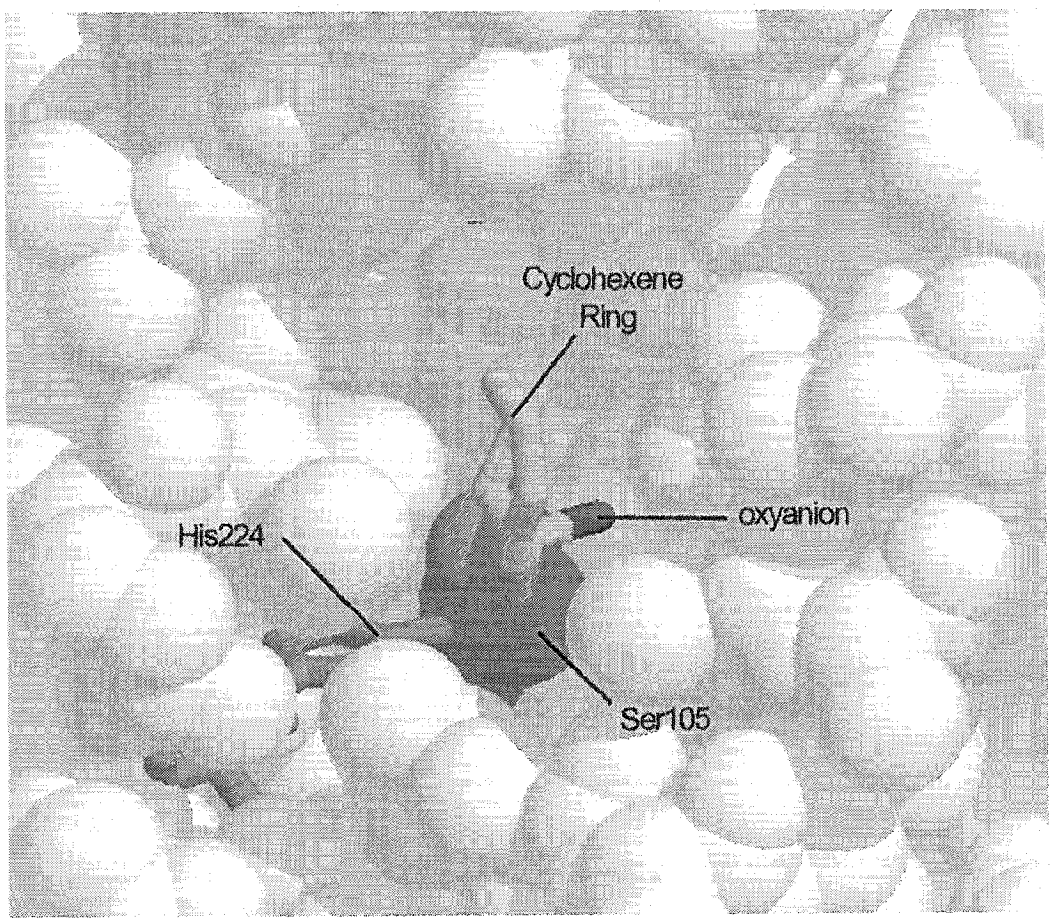


Figure 2: Conformation 1 of the (*R*)-enantiomer of the β-lactam tetrahedral intermediate in CAL-B. This conformation of the lactam ring allows for hydrogen bonding between His224 and the lactam amine. The oxyanion is unable to have a hydrogen bond with Gln106. This conformation may explain the slow rate of the reaction because of the missing hydrogen bond with the oxyanion. A structure with the Gln-106-oxyanion having a hydrogen bonding interaction with the lactam amide would be more catalytically productive.

The second puckered conformation of the (*R*)-enantiomer, Figure 3, also does not allow for a His224 and lactam amine hydrogen bond. This conformation, however, "fits" well inside the active site and has no steric problems. The (*R*)-enantiomer in both puckered conformations produces the best binding model and steric problems of the (*S*)-enantiomer appear to determine the enantioselectivity of the CAL-B in the ring opening reaction. However, the lack of a catalytic hydrogen bond in the best models of the (*R*)-enantiomer indicates the models are not catalytically viable.

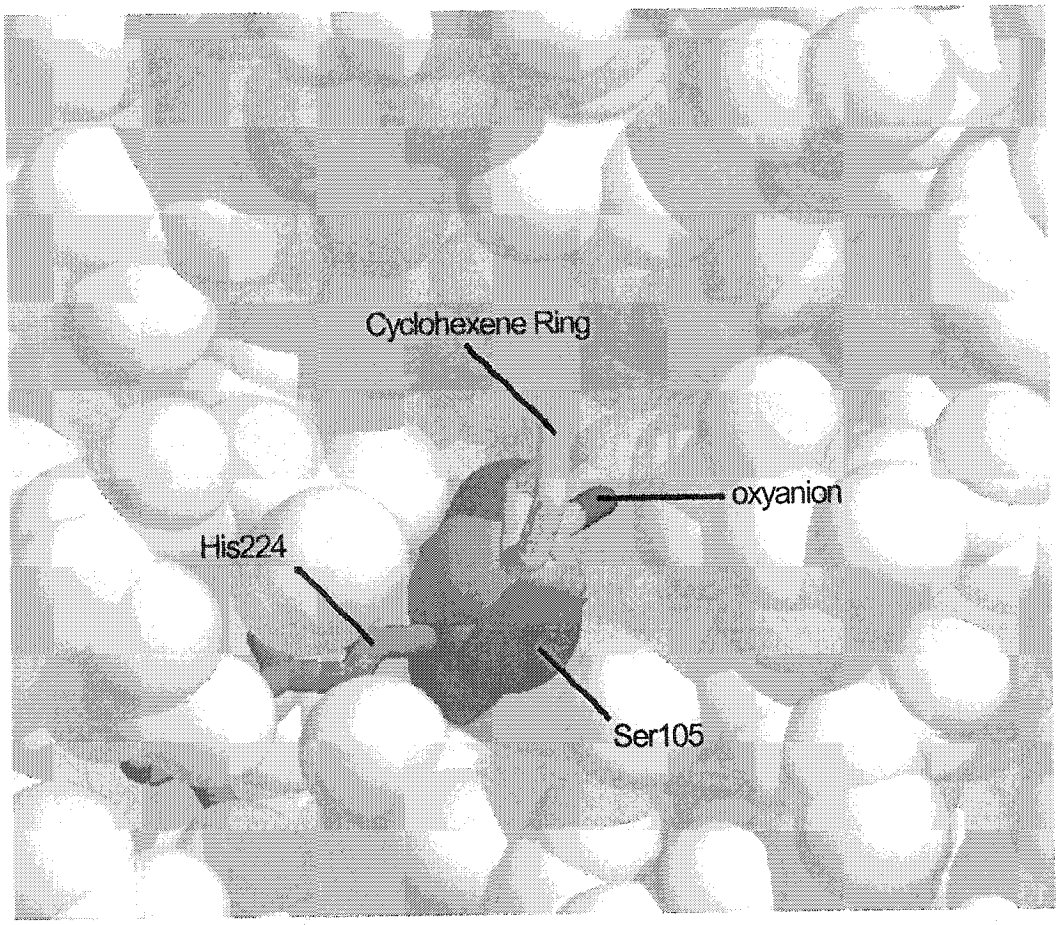


Figure 3: Conformation 2 of the (*R*)-enantiomer of the β-lactam tetrahedral intermediate in CAL-B. This conformation of the lactam ring does not allow for hydrogen bonding between His224 and the lactam amine. The oxyanion is stabilized by all three hydrogen bonding structures (amide and hydroxyl of Thr40 and amide of Gln106) in the CAL-B oxyanion hole.

In order to address this problem the role of the nucleophilic alcohol in these models of the first tetrahedral intermediate is considered. An alcohol bridge between His224 and the lactam amine is proposed as a possible catalytic structure for models missing this hydrogen bond, Figure 4. This alcohol molecule would facilitate the transfer of protons between His224 and the lactam amine. The possibility of having the proposed bridge and hydrogen bonding network was confirmed by the modeling of a water molecule.

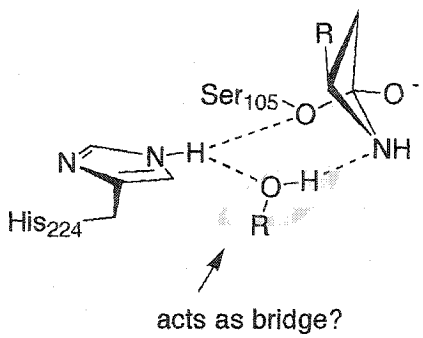


Figure 4: Proposed alcohol bridge for lactam ring opening.

The importance and possibility of having a long chain secondary alcohol in this catalytic model was explored. The (*R*)-enantiomer of 2-butanol was first modeled. The model shows the alcohol binds well within the active site having the C1 carbon in the medium pocket of CAL-B. From this model the remaining atoms of 2-octanol were attached. Conformations of the alkyl chain display it binding above Ile189 and the hexene ring of the β-lactam, Figure 5. Because both enantiomers of 2-octanol catalyze the reaction with similar results (Row 18 and 19 of Table 1) the (*S*)-enantiomer was also constructed from the bridging water molecule. The model suggests the alkyl chain would bind next to Ile285 but an extensive conformational search was not performed.

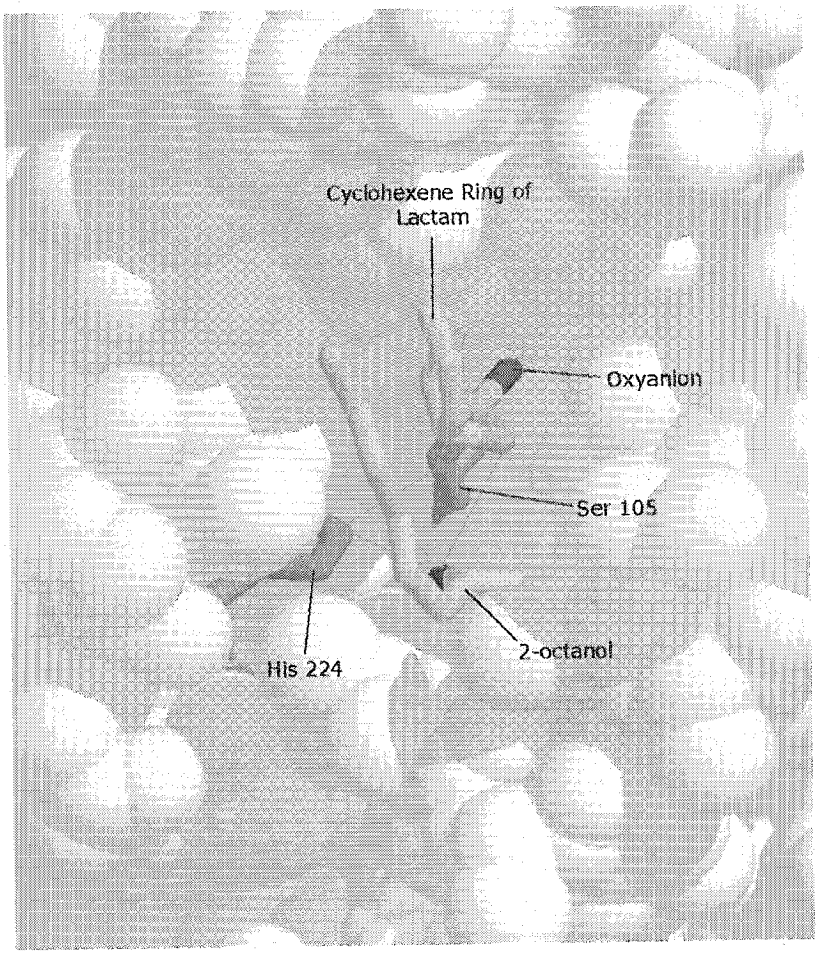


Figure 5: Binding model of 2-octanol and β-lactam tetrahedral intermediate. The alcohol bridges His224 and the lactam amine. The C1 carbon of 2-octanol binds in the medium pocket of CAL-B. The alkyl chain of 2-octanol is oriented next to the cyclohexene ring of the β-lactam and Ile189.

Discussion

The models obtained for the β-lactam ring opening reaction by CAL-B provide an explanation for the enantioselectivity of the reaction. The slow reacting (*S*)-enantiomer of the substrate has obvious unfavored steric interaction with Ile189 in both binding conformations. These unfavorable steric interactions determine the selectivity of the enzyme. The favored (*R*)-enantiomer binds well and without steric problems, in both

puckered conformations of the lactam ring. However, the conformations of he favored enantiomer in both cases do not allow for a hydrogen bond between His224 and the lactam amine. A single conformation of the (*R*)-enantiomer allows for the formation of this hydrogen bond, however, it lacks the Gln106-oxyanion hydrogen bond. This may be the best conformation and may explain the slow reaction rates. However, if the His224 and lactam amine hydrogen bond can be bridged in the other conformations, they would be more catalytically productive.

This research has proposed a alcohol bridging model of the catalytic intermediate. The alcohol facilitates the transfer of a proton between His224 and the lactam amine. This model also positions the nucleophilic alcohol in position to hydrolyze the substrate. The alcohol binding is quite good with the C1 carbon of the alcohol fitting into the medium pocket of CAL-B. This model is a possible catalytic structure, and experimental evidence still needs to be provided to confirm such a model.

Reactions with 2-hexanol were used to test the importance of the chain length of the alcohol. No significant changes were observed. With regards to the proposed binding model, having the alkyl chain oriented above Ile189 and the lactam substrate, the 2-hexanol reaction is understandably similar to that of 2-octanol. The alkyl chain does not bind in any pocket of the enzyme, and therefore, binding interactions and the size of this moiety would not cause significant changes in activity and selectivity.

The temperature dependance of the reaction has not been explored in modeling studies. This reaction has very specific condtions to obtain ideal selectivity and reaction rates. Although need for all the conditions cannot be rationalized, the selectivity appears to be dependent on poor steric interactions. The alcohol bridging model is a catalytic

structure proposed to explain the models obtained for the favored (*R*)-enantiomer, but must be confirmed.

Experimental Section

General

The program Insight II, version 95.0, was used for viewing the structures. The geometric optimizations were performed using Discover, version 2.9.7 (*Accelrys*, San Diego CA, USA), using the CFF91[4] force field. The crystal structure 1lbs[5] was obtained from the Protein Data Bank (www.rcsb.org/pdb/). The structure was obtained with a phosphonate attached to the catalytic serine.

A relaxed enzyme structure having a phosphonate core with key hydrogen bonds and partial charges present was available from previous studies. This structure was used to build the β-lactam tetrahedral intermediate.

Geometry optimizations were carried out in two steps. First, the steepest descent algorithm corrects major high energy structural problems and gets the molecule to a local minima with a RMS deviation of about 0.05 Å mol^{-1}. Second, conjugate gradient algorithm gives a faster and more precise minimization near the local minima, obtaining a RMS value of 0.005 Å mol^{-1} or lower. For all geometry optimizations in this study the solvent molecules and structural Ca^{2+} ions were never fixed.

β-Lactam Modeling

The β-lactam was built in both puckered conformations. Conformations were changed by breaking bonds between lactam ring atoms followed by dihedral bond rotation. The geometry was optimized step-wise, first allowing the substrate to be optimized, followed by the amino acid side chains.

2-Octanol Modeling

A water molecule was place in the active site between His224 and the lactam amine. The single molecule was optimized until hydrogen bonds were present. The C1 to C4 atoms of the molecule were attached. Manual conformational searching was used. Geometry optimization was performed in a the step-wise manner described above. C5 to C8 atoms were attached to the best binding conformation of 2-butanol.

References

1 Nagai, H., Shiozawa, T., Achiwa, K., Terao, Y. *Chem. Pharm. Bull.*, **1993**, *41*, 1933

2 Adam, W., Groer, P., Humpf, H-U., Saha-Möller, C.R. *J. Org. Chem.*, **2000**, *65*, 4919

3 Achilles, K., Schirmeister, T., Otto, H-H. *Arch. Pharm. Pharm. Med. Chem.*, **2000**, *333*, 243.

4 Wilson, E. B., Decius, J. C., Cross, P. C. *Molecular Vibrations*, Dover, New York, **1980**.

5 Uppenberg, J., Oehrner, N., Norin, M., Hult, K., Kleywegt, G. J., Patkar, S., Waagen, V., Anthonsen, T., Jones, T. A. *Biochemistry*, **1995**, *34*, 16838-16851.

Chapter 5: Acetyl Selective Hydrolysis by ThermoGen Esterase E018b

Abstract

Because thermophile enzymes can now be easily produced in the laboratory, they have become candidates for selective catalysts for chemical reactions. A ThermoGen esterase library has been screened for acyl chain selectivity. Though most of the esterases in the library favor hexanoyl or octanoyl chains, E018b shows high selectivity for acetyl groups. 7% acetonitrile has been found to be the ideal solvent condition and high selectivity is demonstrated under these conditions. The selectivity is comparable to a well known highly acetyl selective enzyme, acetyl esterase from orange peel (AE).

Introduction

Ester hydrolysis, though a common chemical process, is difficult to perform selectively with chemical methods. Often harsh conditions, such as acid hydrolysis, may cause problems with other groups of the molecule. Hence, a method to selectively hydrolyze different acyl groups would prove to be quite useful. One such approach would be to use enzyme catalyzed reactions. Both esterases and lipases have been demonstrated to be effect catalysts of ester hydrolysis. These enzymatic reactions have been performed with differnent selectivities (enantio-selectivity, regio-selectivity)

A mutant ThermoGen[1] esterase, E018b, has been shown to have high acetyl selectivity in previous studies[2], Figure 1. This previous work has also indicated that the selectivity of E018b is comparable to other known acetyl selective enzymes such as orange peel acetylesterase (AE) and acetylcholine esterase (AcE). However, some

complication were observed with the hydrolysis of mono-hexonate mono-acetate bisphenol A. The selectivity appeared to reverse with different organic solvents and the selective reaction was quite slow (work by Dave Dietrich at McGill University).

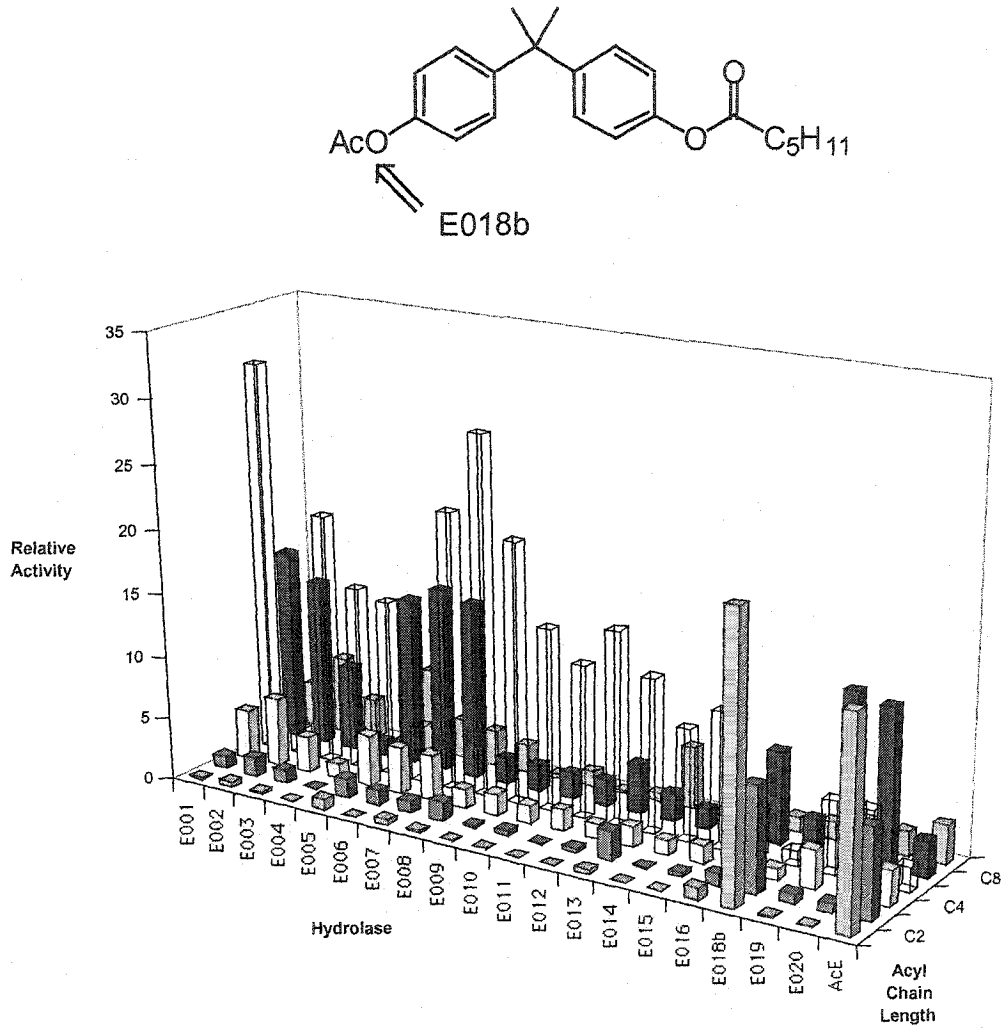
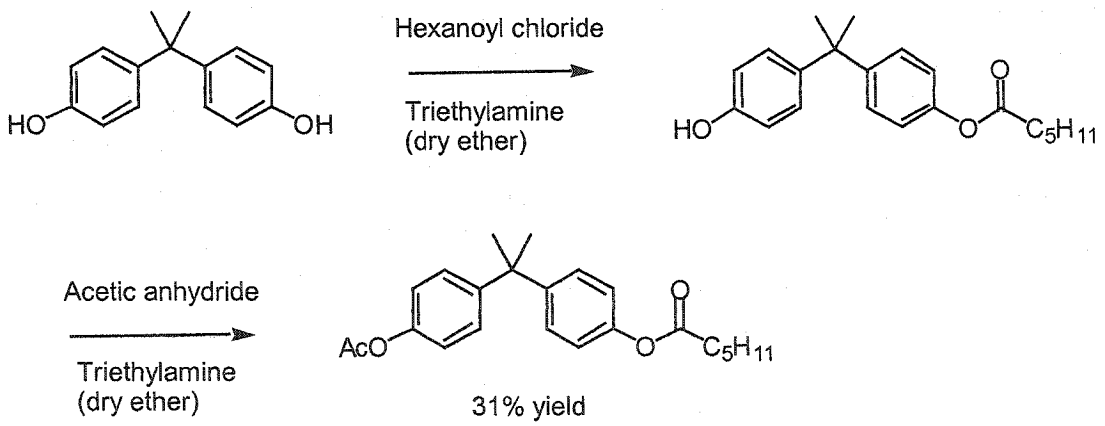


Figure 1: Acyl chain selectivity of E018b and ThermoGen esterase library. E018b is selective for acetyl esters. The diagram above shows that other ThermoGen esterases do not display the same selectivity but are more selective for larger acyl chains. AcE, a well known acetyl selective esterase, and E018b have similar selectivity. Figure from the thesis of Neil Somers[2]

The following research explores ideal conditions to obtain a highly acetyl selective reaction providing high yields with E018b. Different temperatures, organic solvents, solvent concentrations, buffers, buffer concentrations and enzyme concentrations are tested. Reactions are tested using mono-hexonate mono-acetate bisphenol A. The selectivity and yields are compared with that of other esterases, AE and AcE. As well, acetyl selectivity with other substrates is tested.

Results

The synthesis of mono-hexanoate mono-acetate bisphenol A was performed with a 31% yield starting from bisphenol A. Mono-hexanoate bisphenol A as a standard was prepared with a 32% yield from bisphenol. Additionally, mono-acetate bisphenol A was prepared from bisphenol A with a 34% yield.



To analyze the results of the enzymatic reactions, gas chromatography could not be used because well defined peaks could not be obtained for all the products and starting material. With HPLC, good separation was obtained using 95:5 hexane/isopropanol solution.

BES buffer was used for initial reactions, however, no product was seen for any of the reactions. The buffer was then changed to a 100 mM phosphate buffer. The reaction with 7% acetonitrile solution proceeded at a low rate, similar to previous results. However, using higher concentrations of acetonitrile (20 and 50% solutions) yielded little to no product after one week.

Previous results had indicated that performing the reaction in toluene reversed the selectivity of the enzyme (favoring hexanoate hydrolysis). However, the results from this study show that none of the reactions with toluene, hexanes and tetrahydrofuran proceeded. Using a 50% ether solution, very small amounts of product were seen by TLC after 1 week. The reaction in toluene previously was demonstrated to be faster than that in 7% acetonitrile, however, no product was seen from the latest results. Further tests, with HPLC analysis, confirmed that the reaction does not proceed in a 50% toluene solution.

The 7% acetonitrile reaction at room temperature provides yields of 17-20% after 96h with no significant increase in yield with after as much as one week. After using resorufin acetate to test the enzyme stability in 7% acetonitrile it was found that the E018b was denatured after 96h. The yield of the reaction was increased by increasing the temperature to 50 °C and using about 7-10 mg of the esterase. The selectivity of the enzyme for all reactions in 7% acetonitrile remained very good, as only the acetyl hydrolyzed product is observed.

Although the acetonitrile solutions are not biphasic like the toluene solutions, they are still a suspension. The substrate does not dissolve well in a 7% solution. Decreasing the substrate concentration still produced a suspension and no increase in yields was

observed. Triton X was tested as a solvating agent, but no reaction was observed. Testing the enzyme with resorufin acetate after 96h, the enzyme from the triton X solution enzyme remained active.

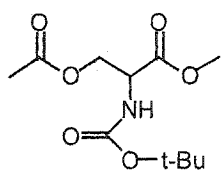
Two other esterases, AcE and AE were tested as a comparison, Table 1. With mono-hexonate mono-acetate bisphenol A in 7% acetonitrile solution, AcE did not react (<3% yield). AE selectivity hydrolyzed the acetate but with lower yields (12%) than E018b.

Table 1: Acetyl hydrolysis of mono-hexanoate mono-acetate bisphenol A by E018b.

Enzyme	Temperature	mg of enzyme	% yield
E018b	Room temp.	1-2	17-21
E018b	50°	7-10	29
AcE	Room temp.	~2	<3
AE	Room temp.	~1.5	12

Only acetyl hydrolysis is observed. Reaction time + 96h
Solvent: 7% acetonitrile, 100mM phosphate buffer.

For the synthesis of O-acetyl N-Boc serine methyl ester, the L enantiomer was obtained with a 71% yield. Much of the D enantiomer is lost, hence no yield was available for the product.. Enzymatic reaction with these substrates produced only one product, suggesting that only the acetyl group is hydrolyzed, which is consistent with previous results. The reaction



was acidified before extraction to assure the acid was not in its ionic form, allowing for it's extraction. The yield of the product was ~38% after 48h. An improved yield would be expected as the enzyme is active for 96h, and can be replenished.

Discussion

ThermoGen esterase E018b serves as a very good catalyst for the selective removal of acetyl ester groups. In all the reactions in this study only ester hydrolysis was observed. In order to have synthetic use, the ideal conditions for the reaction need to be identified.

The esterase is more active at high temperatures, demonstrated by the increase in the reaction rate at 50 °C. The ideal solvent for the reaction is 7% acetonitrile in phosphate buffer. There is no reversal of selectivity with the use of toluene. In solvents such as toluene and with the exception of ether, which creates a two phase system, the esterase did not show any activity.

In acetonitrile, however, the enzyme was denatured after 96h. Because of this inactivation only moderate yields were obtained. The yield of the reaction was increased by increasing the enzyme concentration. Another option is to replenish the enzyme every 96 hours to obtain higher yields.

Finally, substrate flexibility was demonstrated by catalysis of both the O-acetyl N-Boc serine methyl ester and mono-hexanoate mono-acetate bisphenol A.

Experimental Section

Synthesis of mono-hexanoate mono-acetate bisphenol A

Bisphenol A (0.05 mol, 11.4 g), hexanoyl chloride (0.05 mol, 6.5 mL) and triethyl amine (0.075 mol, 5.5 mL) were combined in 50 mL of absolute ether. The triethylamine was added slowly to the reaction mixture and the reaction was kept under inert conditions (N_2). The reaction time was 5 h and progress was monitored by TLC. The reaction mixture was extracted with two washes of water, 1N HCl, and saturated NaOH solutions, then dried over sodium sulphate. The product was isolated via silica column using a 20:1 hexanes/ethyl acetate solution as an eluent. The yield of the product was 3.1 g (38%). The product was confirmed by ^{1}H-NMR: 1.6 (s, 6H, CH_3 of propane), 2.4 (t, 2H, α-CH_2 of hexanoyl), 0.8-1.5 (m, 9H, $(CH_2)_3CH_3$ of hexanoyl), 6.6-7.2 (4d, 8H, aromatic CH). Using the mono-hexonate product, acetylation was carried out in 40 mL of absolute ether combining mono-hexonate bisphenol A (0.0095 mol, 3.1 g), acetic anhydride (0.08 mol, 7.4 mL) and triethylamine (0.02 mol, 1.4 mL) which was added slowly. The reaction time was 3.5h and progress was monitored via TLC. The reaction proceded to completion. The reaction mixture was washed with water, 1N HCl and saturated NaOH, then dried over sodium sulphate. The yield of the product was 2.1 g (82%). Purity of the product was confirmed by ^{1}H-NMR: 1.6 (s, 6H, CH_3 of propane), 2.4 (t, 2H, α-CH_2 of hexanoyl), 0.8-1.5 (m, 9H, $(CH_2)_3CH_3$ of hexanoyl), 6.7-7.2 (4d, 8H, aromatic CH), 2.3 (s, 3H, CH_3 of acetyl).

Hydrolysis w/ Themogen E018b

Enzymatic reactions were performed in 100mM phosphate buffer (pH 7.2). Substrate concentrations were generally between 0.0020- 0.00050 M and an enzyme concentration between 1-2 mg/ml. Single organic solvents were used at concentrations of: 7% acetonitrile, 50% toluene, 50% tetrahydrofuran, 50% ether, 50% hexanes and 50% DMSO. Reactions times were 96h and aliquots were taken at 24h and 48h. The reaction mixture was extracted with an equivalent volume of ether and then washed with water, HCl and NaOH and dried over sodium sulphate. For the reactions >96h, 1ml of enzyme solution was added with 75μL of acetonitrile every 96h.

HPLC analysis

The reaction products were extracted with ether, concentrated and dissolved in a 95:5 hexanes/isopropanol solution. Chromatography was performed using a Chiralcel OD column with a 1.0 flow rate for 20 min and the mobile phase was 95:5 hexane/isopropanol. RTs for the starting material and products were found to be: 5 min for mono-acetate mono-hexonate bisphenol A, 8.5 min for mono-hexonate bisphenol A, 12 min for mono-acetate bisphenol A and 16 min. for bisphenol A.

Synthesis of O-acetyl N-BOC serine methyl ester

In 20ml of methylene chloride each, *N*-BOC-L-serine methyl ester (0.0012 mol, 251 mg) and *N*-BOC-D-serine methyl ester (0.0018 mol, 395 mg) were combined with acetic anhydride (0.042 mol) and triethylamine (0.015 mol). The reaction was carried out under inert conditions (N_2) for 6h. The resulting reaction mixtures were concentrated and dissolved in absolute ether, followed by washes with water, 1N HCl and saturated NaOH and dried over sodium sulphate. Purity was confirmed by NMR.

E018b hydrolysis of O-acetyl N-BOC serine methyl ester

Enzymatic reactions were performed in 50mM BES buffer (pH 7.2). 100 μL of the substrate were used and an enzyme concentration between 0.5-1.0 mg/mL. The solvent contained 7% acetonitrile and the reaction time was 40h. The reaction mixture was extracted twice with an equivalent volume of ether and then washed with water, 1N HCl and saturated NaOH, then dried over sodium sulphate. Gas chromatography of the reaction produce gave only one product peak at RT 24.7 for D-serine and 25.0 for L-serine. The RT of *N*-BOC-L-serine methyl ester was determined to be 24.2.

References

1 ThermoGen Inc., Chicago Technology Park 2201 West Cambell Park Drive, Chicago, Illinois, 60612.

2 Somers, N. A. *Master of Science Thesis*, McGill University, Department of Chemistry, **1999**.