2013

COMPUTATIONAL MODELING, DESIGN, AND CHARACTERIZATION OF COCAINE-METABOLIZING ENZYMES FOR ANTI-COCAINE MEDICATION

Lei Fang
University of Kentucky, fralium@hotmail.com

Recommended Citation
Fang, Lei, "COMPUTATIONAL MODELING, DESIGN, AND CHARACTERIZATION OF COCAINE-METABOLIZING ENZYMES FOR ANTI-COCAINE MEDICATION" (2013). Theses and Dissertations--Pharmacy. 39.
https://uknowledge.uky.edu/pharmacy_etds/39

This Doctoral Dissertation is brought to you for free and open access by the College of Pharmacy at UKnowledge. It has been accepted for inclusion in Theses and Dissertations--Pharmacy by an authorized administrator of UKnowledge. For more information, please contact UKnowledge@lsv.uky.edu.
STUDENT AGREEMENT:

I represent that my thesis or dissertation and abstract are my original work. Proper attribution has been given to all outside sources. I understand that I am solely responsible for obtaining any needed copyright permissions. I have obtained and attached hereto needed written permission statements(s) from the owner(s) of each third-party copyrighted matter to be included in my work, allowing electronic distribution (if such use is not permitted by the fair use doctrine).

I hereby grant to The University of Kentucky and its agents the non-exclusive license to archive and make accessible my work in whole or in part in all forms of media, now or hereafter known. I agree that the document mentioned above may be made available immediately for worldwide access unless a preapproved embargo applies.

I retain all other ownership rights to the copyright of my work. I also retain the right to use in future works (such as articles or books) all or part of my work. I understand that I am free to register the copyright to my work.

REVIEW, APPROVAL AND ACCEPTANCE

The document mentioned above has been reviewed and accepted by the student’s advisor, on behalf of the advisory committee, and by the Director of Graduate Studies (DGS), on behalf of the program; we verify that this is the final, approved version of the student’s dissertation including all changes required by the advisory committee. The undersigned agree to abide by the statements above.

Lei Fang, Student
Dr. Chang-Guo Zhan, Major Professor
Dr. Jim Pauly, Director of Graduate Studies
A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Pharmacy at the University of Kentucky

By
Lei Fang
Lexington, Kentucky

Advisor: Dr. Chang-Guo Zhan,
Professor of Pharmaceutical Sciences
Lexington, Kentucky
2013

Copyright © Lei Fang 2013
ABSTRACT OF DISSERTATION

COMPUTATIONAL MODELING, DESIGN, AND CHARACTERIZATION OF COCAINE-METABOLIZING ENZYMES FOR ANTI-COCAIN MEDICATION

Cocaine is a widely abused and addictive drug, resulting in serious medical and social problems in modern society. Currently, there is no FDA-approved medication specific for cocaine abuse treatment. The disastrous medical and social consequences of cocaine abuse have made the development of an anti-cocaine medication a high priority. However, despite decades of efforts, traditional pharmacodynamic approach has failed to yield a truly useful small-molecule drug due to the difficulties inherent in blocking a blocker like cocaine without affecting the normal functions of the transporters or receptors. An alternative approach, *i.e.* pharmacokinetic approach, is to interfere with the delivery of cocaine to its receptors/transporters and/or accelerate its metabolism in the body. It would be an ideal anti-cocaine medication to accelerate cocaine metabolism producing biologically inactive metabolites.

Two natural enzymes may catalyze hydrolysis of cocaine: human butyrylcholinesterase (BChE) and bacterial cocaine esterase (CocE). However, the wild-type enzymes are not suitable as anti-cocaine therapeutics, due to the low catalytic activity, thermostability, or short biological half-life. In this investigation, we performed integrated computational-experimental studies to rationally design and discover mutants of these enzymes in order to improve the catalytic activity, thermostability, and/or biological half-life. To rationally design desirable mutants of the enzymes, we have successfully developed computational models, including those for BChE gating, glycosylated BChE structure, BChE-substrate complex structures, BChE dimer/tetramer structures, CocE monomer/dimer structures, and CocE-substrate complex structures. Development of the computational models enabled us to rationally design new amino-acid mutations that may improve the catalytic activity, thermostability, and/or prolonged biological half-life. The computational design was followed by wet experimental tests, including both *in vitro* and *in vivo* experiments, leading to discovery of new enzyme forms with not only a high catalytic efficiency against cocaine, but also an improved thermostability and/or prolonged biological half-life. The identified new mutants of BChE and CocE are expected to be valuable candidates for development of a more efficient enzyme therapy for cocaine abuse. The encouraging outcomes of the present study also suggest that the structure-and-mechanism-based design and integrated computational-experimental approach is promising for rational drug design and discovery.
KEYWORDS: Protein drug design, Human butyrylcholinesterase, Bacterial cocaine esterase, Mutant, Anti-Cocaine

Lei Fang

____________________
Student’s signature

07/31/2013

____________________
Date
COMPUTATIONAL MODELING, DESIGN, AND CHARACTERIZATION OF COCAINE-METABOLIZING ENZYMES FOR ANTI-COCaine MEDICATION

DISsERTATION

by

Lei Fang

Chang-Guo Zhan
____________________
Director of Dissertation
Jim Pauly
____________________
Director of Graduate Study
07/31/2013
____________________
Date
ACKNOWLEDGEMENTS

I would like to express my sincerest thanks to many people for their kind support and help during my graduate studies at the University of Kentucky. First, I am grateful to my advisor Dr. Chang-Guo Zhan who brought me into this field and patiently guided me throughout my graduate studies. Especially, he taught me the idea of critical and logical thinking which will benefit my future career and whole life. I would like to thank my committee members Dr. Peter Spielmann, Dr. Kyung-Bo Kim, and Dr. Steven Van Lanen for their advice, encouragement and help throughout my study. I would also like to thank Dr. David Rodgers for serving as the outside examiner of my thesis defense and for x-ray crystallography analysis. In addition, I thank Dr. Fang Zheng for her kind advice during my thesis research.

Besides, I would like to thank many members of Dr. Zhan’s group, including Dr. Xi Chen, Dr. Yongmei Pan, Dr. Daquan Gao, Dr. Yuan Yao, Dr. Wenchao Yang, Dr. Xiaqin Huang, Dr. Vinod Kasam, Dr. Mohamed Diwan M. AbdulHameed, Dr. Junjun Liu, Dr. Haiting Lu, Dr. Zhenyu Jin, Dr. Dongmei Li, Dr. Gefei Hao, Dr. Beilei Lei, Dr. Bo Yang, Dr. Xinyun Zhao, Dr. Donghui Wei, Dr. Ning-Ning Wei, Dr. Marius J. Vilkas, Dr. Adel Hamza, Dr. Liu Xue, Ms Shurong Hou, Ms. Min Yao, Mr. Xiabin Chen, and Ms. Yan Qiao, for their help, support and friendship in the past years. I would also like to thank former visiting professors in the group, including Professors Jennifer L. Muzyka, Alan C. Goren, and Fei Liao, for their help. I also thank Ms. Catina Rossoll for her kind help throughout my graduate studies. I would like to thank all the people (including Dr. Martin Chow in the Center of Structural Biology) and friends who have supported and helped me during my life in Lexington.

Finally, I would like to thank my father (Xiaomin Fang) and mother (Huijun Chen) for their selfless and endless love, sacrifices and constant support during my whole life. I would also thank my wife Shurong Hou whose love and support are my immeasurable strengths. Finally, I would like to specially thank my son Kevin who has been inspiring me to work hard for my career development.
TABLE OF CONTENTS

ACKNOWLEDGEMENTS ................................................................. iii
TABLE OF CONTENTS .................................................................... iv
LIST OF FIGURES ........................................................................ vi
LIST OF TABLES ........................................................................... viii

Chapter One: Cocaine Abuse and Anti-Cocaine Medication ...................... 1
  1.1 Cocaine and its abuse ............................................................. 1
  1.2 Mechanism for the action of cocaine ....................................... 3
  1.3 Treatment of cocaine abuse ................................................... 4
  1.4 Human butyrylcholinesterase as a possible therapeutic enzyme for treatment of cocaine abuse .................................................. 7
  1.5 Bacterial cocaine esterase as a possible therapeutic enzyme for treatment of cocaine abuse .................................................. 8
  1.6 Further development of the therapeutic enzymes ....................... 9

Chapter Two: Active-Site Gating and Substrate Specificity of Human Butyrylcholinesterase and Acetylcholinesterase .......................... 12
  2.1 Overview of the active-site gating ........................................... 12
  2.2 Computational methods used to understand the active-site gating .......................................................... 14
  2.3 Stability of the MD trajectories .................................................. 16
  2.4 Active-site doors of AChE and BChE monomers ......................... 18
  2.5 Active-site doors of AChE and BChE tetramers .......................... 27
  2.6 Conclusion ............................................................................. 35

Chapter Three: Substrate Selectivity of High-Activity Mutants of Human Butyrylcholinesterase ...................................................... 36
  3.1 Desirable substrate selectivity of the human BChE mutants .......... 36
  3.2 Computational methods used to assess the substrate selectivity ......... 39
  3.3 Insights from the computational modeling ................................... 39
  3.4 Experient Validation .................................................................. 45
  3.5 Conclusion ............................................................................. 46

Chapter Four: Model of Glycosylated Human Butyrylcholinesterase .......... 47
  4.1 Understanding the detailed structure of a glycosylated protein ........ 47
  4.2 Computational methods used to study the structure of glycosylated BChE 48
  4.3 Structural features of the glycosylated BChE ................................. 49
  4.4 Solvent Accessible Surface Area (SASA) ...................................... 52
  4.5 Conclusion ............................................................................. 60

Chapter Five: Computational Design of Amino-Acid Mutations to Prolong the Biological Half-life of a Highly Efficient Cocaine Hydrolase Engineered from Human Butyrylcholinesterase ............................................. 61
LISTS OF FIGURES

Figure 1.1 Molecular structure of (-)-cocaine free base ................................................. 2
Figure 1.2 Structure of human butyrylcholinesterase in ribbon model .............................. 7
Figure 1.3 Structure of cocaine esterase (dimer) in ribbon model ................................. 9
Figure 2.1 The continuous Connolly surface of a protein under different probe radii.... 16
Figure 2.2 Time-dependent RMSD values for the atomic positions of the protein backbone in the MD-simulated enzyme structures. (A) AChE monomer, and (B) BChE monomer ............................... 17
Figure 2.3 Time-dependent RMSD values for the atomic positions of the protein backbone in the MD-simulated enzyme structures. (A) AChE tetramer, and (B) BChE tetramer ........................................ 18
Figure 2.4 The main-door radii of the monomers: (A, top) AChE; (B, bottom) BChE... 19
Figure 2.5 (A, upper) The eight residues with aromatic or indole side chains in the gorge of AChE. (B, middle) The eight residues in the gorge of BChE. (C, bottom) Superimposing of the gating residues in AChE and BChE...... 22
Figure 2.6 The overall and detailed view of different entrances to the active site of AChE.................................................................................................................. 25
Figure 2.7 The overall and detailed view of different entrances to the active site of BChE.................................................................................................................. 26
Figure 2.8 The overall view of AChE tetramer and BChE tetramer ......................... 28
Figure 2.9 (A) Acetylcholine (ACh) can enter the active site of AChE/BChE monomer or AChE/BChE tetramer subunits A and C when the main door is open. (B) to (D) The path to active site of AChE/BChE tetramer subunit B or D (left part, labeled in multiple colors) is blocked by subunit A or C (right part, labeled in green). ................................................................................................. 30
Figure 2.10 (A) – (D) Time-dependence of the key inter-residue distances between the center of mass of representative residues in subunit B and that of the main blocking residues in complementary subunit A in the MD-simulated BChE tetramer. (E) A snapshot of the MD-simulated BChE tetramer...................... 34
Figure 3.1 The energy-minimized ES and TS1 structures for (-)-cocaine interacting with wild-type human BChE and E12-7................................. 39
Figure 3.2 The energy-minimized ES and TS1 structures for ACh interacting with wild-type human BChE and E12-7 ......................................................... 40
Figure 3.3 The energy-minimized ES structures for ACh interacting with E14-3, E12-4, and E13-1 ................................................................. 41
Figure 3.4 The energy-minimized ES and TS1 structures for ATC interacting with wild-type human BChE and E12-7 ..................................................... 42
Figure 3.5 The energy-minimized ES and TS1 structures for BTC interacting with wild-type human BChE and E12-7 ..................................................... 43
Figure 3.6 The energy-minimized ES and TS1 structures for (+)-cocaine interacting with wild-type human BChE and E12-7 ........................................ 44
Figure 4.1 The composition of the glycan in human BChE ............................................ 47
Figure 4.2 Time-dependent RMSD values for the atomic positions of the protein backbone and glycan backbone in the MD-simulated unglycosylated and
glycosylated human BChE structures................................................................. 48
Figure 4.3 Overall structure of the glycosylated human BChE model derived from the
MD simulation.......................................................... 49
Figure 4.4 Representative PEGylation reactions.................................................... 50
Figure 4.5 Locations of representative lysine residues.............................................. 57
Figure 5.1 The time-dependent Ca–Ca distances between key residues from different
subunits (a and b) in the MD-simulated BChE dimer structure......................... 66
Figure 5.2 The overall and detailed view of modeled BChE dimer structure.............. 68
Figure 5.3 Nondenaturing gels (8%) stained for the BChE activity of (A) E12-7, (B)
E364-532, (C) E377-516, and (D) E535................................................................. 69
Figure 5.4 Kinetic data for the hydrolysis of (-)-cocaine catalyzed by BChE mutants:
(A) E12-7; (B) E364-532; (C) E377-516; and (D) E535................................. 70
Figure 5.5 Time-dependent concentrations of the active enzyme in rat plasma after
the i.v. injection of the enzyme at a dose of 0.15 mg/kg: (A) E12-7, (B)
E364-532, (C) E377-516, and (D) E535................................................................. 73
Figure 6.1 The time-dependent Ca–Ca distances between L196 and I301 during the
whole MD simulation in CocE dimer................................................................. 79
Figure 6.2 The overall and detailed view of modeled CocE dimer structure.............. 81
Figure 6.3 Plots of measured reaction rates (with error bars) versus the substrate
concentration for (−)-cocaine hydrolysis catalyzed by CocE mutants.............. 82
Figure 6.4 Thermostability profile of E196-301....................................................... 84
Figure 6.5 Structural stability of E172-173 and two cross-subunits disulfide bonds
containing mutant E196-301.......................... 85
Figure 6.6 Backbone alignments between the crystal structures of E172-173 and E196-
301.......................................................... 88
Figure 6.7 Time-dependence of important H–O distances (related to hydrogen bonds)
from MD-simulated E172-173 and E196-301 structures................................. 89
Figure 6.8 in vivo protection of E196-301 and PEGylated E196-301 against cocaine-
induced lethality in mice................................................................. 91
LIST OF TABLES

Table 2.1 The average and maximum radii of different doors in monomers. .................. 20
Table 2.2 The average and maximum radii of main doors in tetramers ....................... 29
Table 2.3 The inter-residue distances ($R_{AB}$, in Å) between the center of mass of
representative residues in subunit B and that of the main blocking residues in
complementary subunit A of AChE and BChE tetramers. ............................. 31
Table 2.4 The calculated average vdW radii (Å) of the residues involved in the inter-
residue distance calculations for residues in subunits A and B of the AChE
and BChE tetramers. ............................................................................ 32
Table 4.1 The relative SASA values of the Nε atoms of lysine residues on the surface of
human BChE calculated by using different probe radii................................... 55
Table 4.2 The Percent of Remain (PÔR) of SASA of the Nε atoms of lysine residues on
BChE surface calculated using different probe radii ................................. 57
Table 5.1 The Cα-Cα distances between key residues in various dimer/tetramer BChE
structures obtained from computational modeling/simulation. ...................... 68
Table 5.2 Kinetic parameters determined for (-)-cocaine hydrolysis catalyzed by the
BChE mutants. ..................................................................................... 75
Table 6.1 The Cα-Cα distances between key residues in CocE dimer obtained from
computational simulation ........................................................................ 91
Table 6.2 Kinetic parameters determined for (-)-cocaine hydrolysis catalyzed by
CocE mutants. ..................................................................................... 94
Table 6.3 Summary of the MD-simulated key distances (in Å) between the hydrogen
atoms of key residue and the carbonyl oxygen of (-)-cocaine benzoyl
ester in the transition-state structures of CocE. ........................................ 98
Chapter One: Cocaine Abuse and Anti-Cocaine Medication

Cocaine is a widely abused and addictive drug which blocks dopamine reuptake in the central nervous system (CNS). Currently, there is no FDA-approved medication specific for cocaine abuse treatment. The disastrous medical and social consequences of cocaine abuse have made the development of an anti-cocaine medication a high priority. However, despite decades of efforts, traditional pharmacodynamic approach has failed to yield a truly useful small-molecule drug due to the difficulties inherent in blocking a blocker like cocaine without affecting the normal functions of the receptors or transporters (such as dopamine transporter). An alternative approach, i.e. pharmacokinetic approach, is to interfere with the delivery of cocaine to its receptors and/or accelerate its metabolism in the body. It would be an ideal anti-cocaine medication to accelerate cocaine metabolism producing biologically inactive metabolites. Two natural enzymes, i.e. human butyrylcholinesterase (BChE) and bacterial cocaine esterase (CocE), can catalyze hydrolysis of cocaine into non-toxic metabolites, demonstrating their potential in anti-cocaine medication.

1.1 Cocaine and its abuse

Cocaine is a natural alkaloid found in the coca plant (mainly from Erythroxylum coca and Erythroxylum novogranatense) originally from Andean ridge in South America, and normally extracted from plant leaves.\(^1\)\(^,\)\(^2\) The ancient Inca Indians grew coca for both medical purpose and religious rituals for a long history due to its ability to boost energy, relieve fatigue, and lessen hunger.\(^1\)\(^,\)\(^3\) As the discovery of the new world in sixteenth century, coca plant is known by the outside world. However, the content of cocaine in coca leaves decreases significantly during storage,\(^4\) it was not until 1855 that this alkaloid was isolated by a German chemist Friedrich Gaedcke with the name of “erythroxyline”.\(^5\) Albert Niemann further improved the purification process of this alkaloid in 1859 and named it with the current name “cocaine”.\(^2\)\(^,\)\(^6\)
Before the isolation of cocaine, chewing coca leaves is the main form of consuming cocaine. In fact, due to the low gastrointestinal absorption of cocaine and the relatively low content of cocaine in coca leaves (about 1% of the weight of the leaves), chewing coca leaves does not cause the toxicity associated with purified cocaine.\(^3\) Currently, there are two forms of cocaine, including salt and base. Cocaine hydrochloride is the most common salt and is the main form in the street until the late 1970s. It is a water-soluble powder, normally used nasally or mixed with water and injected intravenously.\(^7\) The high melting temperature (195 °C) of the salt makes it not suitable to be smoked because it will decompose before vaporization. Cocaine base is the alkaloid form of cocaine including freebase and “crack”, mainly emerging in the street after 1970s. Crack cocaine is in fact a low purity form of freebase cocaine, and is named from the sound of cocaine crystals popping during preparation. Freebase and crack cocaine are insoluble in water, making it difficult for injection. However, the relatively low melting temperature (98°C) makes it easily be smoked.

![Molecular structure of (-)-cocaine free base.](image)

Cocaine has limited medical use as a local anesthetic (in dentistry and ophthalmology). More importantly, it is a CNS stimulant. Consuming cocaine in any aforementioned forms, either by nasal insufflation, intravenous injection, or smoking, can lead to the feeling of euphoria or “high”. Depending on the route of administration and the dosage, cocaine can take effect in from less than one minute to around 15 minutes and last for 15 minutes to one hour.\(^8\) Besides the feeling of euphoria, effects of cocaine also include feeling of energetic, talkative, increased alertness, reduced appetite, and sensation of light, sound and touch.\(^9\) Long-term use of cocaine could cause serious physical
problems, including cardiovascular failures, respiratory difficulties, and gastrointestinal complications, and mental problems, including depression, agitation, nervous, and tired but unable to sleep. Cocaine overdose could be more serious and even lethal, due to the possible respiratory failure, stroke, cerebral hemorrhage, or heart-failure after seizures induced by cocaine’s toxicity.\textsuperscript{10,11}

As the addictive and toxic properties of cocaine have been known, the production, distribution, and sale of cocaine products are restricted in most countries.\textsuperscript{12} However, cocaine abuse is still a widespread problem throughout society.\textsuperscript{12}

\section*{1.2 Mechanism for the action of cocaine}

Cocaine is a powerful CNS stimulant, causing its reinforcing and toxic effects by binding with various monoaminergic neurotransmitter systems, such as dopaminergic,\textsuperscript{13,14} serotonergic,\textsuperscript{15} and norepinephrinergic\textsuperscript{16} systems. The most important target of cocaine is the dopamine transporter (DAT).\textsuperscript{17} During the normal dopaminergic signaling cycle, neurotransmitter dopamine (DA) is released into synaptic cleft from the transmitting neuron and binds with dopamine receptors on the receiving neuron to conduct the signaling. Dopamine will be then recycled \textit{via} DAT located on the transmitting neuron to end the signaling cycle. However, cocaine entering into the brain would interfere with this signaling cycle. It will bind with the DAT to form a complex that can block the binding of dopamine.\textsuperscript{17} This can result in accumulation of dopamine in the synaptic cleft and lead to the prolonged and enhanced effect of dopaminergic signaling, which causes the euphoria feeling as reported by cocaine users.

Moreover, one-time use of cocaine could cause the change of the dopaminergic signaling, \textit{i.e.} the down-regulation of dopamine receptors and up-regulation of dopamine transporter for at least a month.\textsuperscript{18} So, cocaine user would not achieve the same “high” as they felt at the first time they used the drug. As normalization of the dopaminergic function is usually an extremely slow process, there is less dopamine in the synapse for signaling in the brain of a cocaine user. Gradually, a tolerance to the drug will be developed, leading to the drug seeking or carving. So, it is necessary for therapeutic treatment of cocaine addiction to first (directly or indirectly) block the stimulant effects
of cocaine. Without the stimulant effects of cocaine, there would be a better chance to bring the function of brain's communication system back to normal.\textsuperscript{19}

Besides affecting the neurological system, cocaine also has a severe impact to the cardiovascular system. Increased heart rate, blood pressure, and vasoconstriction are caused by cocaine binding with noradrenergic transporters to result in elevated noradrenergic signaling.\textsuperscript{20} Cocaine also binds to sodium ion channel, which are responsible for the local anesthesia and may contribute to cardiac arrhythmias.\textsuperscript{21}

1.3 Treatment of cocaine abuse

Cocaine is now widely abused all over the world with an estimated annual cocaine users ranging from 13.3 million to 19.7 million in 2010, especially in North American which is the world's largest consumer of cocaine.\textsuperscript{12} Cocaine is now the second most popular illegal recreational drug in the United States behind marijuana and the most co-abused drug along with alcohol.\textsuperscript{20} However, it accounts for nearly half of all the emergency department (ED) visits involving the use of illicit drugs in 2008.\textsuperscript{20} Cocaine abuse has resulted in serious medical and social problems in modern society.\textsuperscript{22,23} Currently, there is no anti-cocaine medication approved by the U.S. Food and Drug Administration (FDA). So cocaine abuse has made the development of an effective pharmacological treatment a high priority. Both pharmacodynamics and pharmacokinetic method have been attempted to search for possible candidates of anti-cocaine therapeutics.\textsuperscript{22,24}

1.3.1 Pharmacodynamics method

Traditional pharmacodynamic method was used to find small-molecule agents that are either to block cocaine binding with DAT or alter the effects of cocaine by acting at sites other than DAT.\textsuperscript{22,25} For example, vanoxerine (GBR-12909), a selective DAT inhibitor with a high affinity, was shown to compete with cocaine at the binding of DAT as a promising candidate. However, due to its cardiovascular side effects, the US National Institute on Drug Abuse (NIDA) discontinued its clinical trials (from 2003-2008).\textsuperscript{25} Tetrahydropalmatine (THP), a non-selective DA receptor antagonist, was shown
efficacy in reducing craving and lowered opiate withdrawal symptoms in heroin addicts. It is now in Phase I human clinical trial for treatment of cocaine addiction. Atomoxetine, a norepinephrine re-uptake inhibitor finished its Phase II clinical trial in 2012 as a potential therapeutic compound of cocaine abuse in patients with ADHD. Beside these agents, more small-molecule based agents are under different phases of clinical trials. However, though tried for a long time, due to the complex interrelations of neuronal circuits, it is very challenging to find a safe and specific therapeutic agent against cocaine through a pharmacodynamics-based method.

### 1.3.2 Pharmacokinetic method

Altering the pharmacokinetics of cocaine as treatment of cocaine abuse is gaining more and more attention in recent years. Pharmacokinetic method aims to interfere with the delivery of cocaine to its sites of action and/or accelerate its metabolism in the body through protein therapeutics including immunological or enzyme strategy.

Immunological strategy aims at searching for protein biologics to sequester cocaine (via binding with an antibody). Passive immunization with monoclonal antibodies was developed and shown to reduce cocaine self-administration in rats. The most effective active immunization vaccine to date is TA-CD, a Succinyl Norcocaine-Cholera Toxin B Vaccine. According to the clinical trials, reduced cocaine use was observed in subjects with a reasonably high (≥43 μg/mL) IgG anti-cocaine antibody levels, which is only attained by 38% of the vaccinated subjects. Report also showed that antibody response of most subjects is 6 weeks later, and cannot last for longer than 12 weeks. The major disadvantage of the antibody strategy is that each antigen-binding site of the antibody can bind with only one cocaine molecule. Catalytic antibody might be another direction, which can not only sequester cocaine but can accelerate cocaine metabolism. The most efficient anti-cocaine catalytic antibody reported so far is 15A10. It can block the reinforcing effects of cocaine self-administration in rat models and reduced cocaine-induced seizures and deaths in a dose-dependent manner. However, the catalytic efficiency of the catalytic antibody is very low (\(K_M = 220 \mu M\) and \(k_{cat} = 2.3 \text{ min}^{-1}\)). Immunological strategy has shown some potential in anti-cocaine medication,
especially for cocaine addiction patients which have relatively low plasma cocaine concentration (1-5 µM).\textsuperscript{19} However, it still has limitation in treating cocaine overdose patients with plasma cocaine concentration around 3-200 µM.\textsuperscript{19} In addition, improved vaccines and boosters are necessary to adequately block the effects of cocaine.\textsuperscript{31}

Compared with immunological strategy, developing a highly efficient cocaine hydrolase can be a more promising direction.\textsuperscript{19,20,22,28,34} The desired therapeutic enzymes can hydrolyze cocaine in the serum with the high efficiency before it reaches its sites of action to avoid the cocaine toxicity, and can free itself for further binding with and hydrolyzing another cocaine molecule. Before searching for qualified candidates, it is necessary to understand the metabolic pathways of cocaine in human. There are two ester groups in cocaine, benzoyl ester and methyl ester. Benzoyl ester hydrolysis generates ecgonine methyl ester (EME), whereas methyl ester hydrolysis yields benzoylecgonine (BE).\textsuperscript{23} In human, it is estimated that hydrolysis accounts for about 95\% of cocaine metabolism in human.\textsuperscript{22,35} Human butyrylcholinesterase (BChE) is the principal cocaine hydrolase in plasma which is capable of hydrolyzing cocaine at the benzoyl ester,\textsuperscript{36} while two liver carboxylesterases, known as hCE-1 nd hCE-2, are responsible for hydrolyzing cocaine at the methyl ester and benzoyl ester, respectively.\textsuperscript{22,37} The remaining 5\% is deactivated through oxidation by the liver microsomal cytochrome P450 system, producing norcocaine.\textsuperscript{22,23} Among the metabolites (EME, BE, and norcocaine), EME appears to be the least pharmacologically active and may even cause vasodilation,\textsuperscript{38} whereas both BE and norcocaine appear to cause vasoconstriction and lower the seizure threshold, similar to cocaine itself.\textsuperscript{22,38-40} Moreover, norcocaine is hepatotoxic and a local anesthetic.\textsuperscript{22,41,42} Clearly, amplifying the metabolic pathway through hydrolysis at the cocaine benzoyl ester is most suitable for rapid cocaine clearance in the body. Currently, human BChE and bacterial cocaine esterase (CocE) are the two most interesting candidate enzymes.\textsuperscript{20,28} These enzymes will be discussed further in the following sections.
1.4 Human butyrylcholinesterase as a possible therapeutic enzyme for treatment of cocaine abuse

Human butyrylcholinesterase (BChE) is a plasma enzyme widely distributed in the body, and the only substrate in the body is acetylcholine. The biological function of BChE is not completely clear yet; studies have suggested that it may be involved in the regulation of cell proliferation and differentiation during early neuronal development, and neurodegenerative disorders such as Alzheimer's disease. On the other hand, human BChE has a long history of clinical use in treating patients incapable of hydrolyzing succinylcholine and mivacurium, and no adverse effects have been observed.

Figure 1.2 Structure of human butyrylcholinesterase (PDB ID: 1P0M) in ribbon model.

Many studies have demonstrated that administration of an exogenous enzyme (human BChE purified from donated human blood) substantially decreases cocaine’s half-life. Beside, BChE has potential advantages over active immunization since its administration would immediately enhance cocaine metabolism and would not require an immune response to be effective which is important for patients with cocaine overdose in ED. Combining all these reasons, enhancement of cocaine metabolism by administration of BChE is considered as a promising pharmacokinetic approach for the treatment of cocaine abuse. However, the catalytic activity of this plasma enzyme against naturally occurring (-)-cocaine ($K_M = 4.5 \mu M$ and $k_{cat} = 4.1 \text{ min}^{-1}$) is three orders of
magnitude lower than that against the biologically inactive (+)-cocaine enantiomer ($K_M = 6400 \ \mu M$ and $k_{cat} = 4.7 \ \text{min}^{-1}$).\textsuperscript{19,48} (+)-Cocaine can be cleared from plasma in seconds and prior to partitioning into the CNS, whereas (-)-cocaine has a plasma half-life of approximately 47 min or longer (for an intravenous dose of cocaine 0.2 mg/kg), long enough for the manifestation of the CNS effects, which peak in minutes.\textsuperscript{49} Thus, positron emission tomography (PET) mapping of the binding of (-)-cocaine and (+)-cocaine in baboon CNS showed marked uptake corresponding to (-)-cocaine at the striatum along with other areas of low uptake, whereas no CNS uptake corresponding to (+)-cocaine was observed. (+)-cocaine was hydrolyzed by BChE so rapidly that it never reached the CNS for PET visualization.\textsuperscript{50} Furthermore, the actual half-life of (-)-cocaine in plasma is dependent on the dose of cocaine received.\textsuperscript{19} When the enzyme is saturated, the (-)-cocaine hydrolysis speed has already reached its maximum and will no longer change by increasing dose of (-)-cocaine. Thus, the higher the dose, the longer the actual half-life of (-)-cocaine in plasma in the case of cocaine overdose.\textsuperscript{19,34}

1.5 Bacterial cocaine esterase as a possible therapeutic enzyme for treatment of cocaine abuse

Besides BChE, bacterial cocaine esterase (CocE) found in Rhodococcus spp. strain MB1 from the rhizosphere soil of coca plant can also hydrolyze the benzoyl ester of (-)-cocaine ($K_M = 640 \ \text{nM}$ and $k_{cat} = 468 \ \text{min}^{-1}$).\textsuperscript{51} Therefore, the catalytic efficiency ($k_{cat}/K_M$) of native CocE is approximately 800-fold higher than that of native BChE against (-)-cocaine. Studies have shown that CocE can help prevent acute cocaine toxicity from a lethal dose of cocaine.\textsuperscript{52} CocE has a relatively short half-life \textit{in vivo}. In rat, CocE has a biological half-life of approximately 10 min.\textsuperscript{53} In addition, being a bacterial protein, CocE may elicit a robust immune response.\textsuperscript{34}
1.6 Further development of the therapeutic enzymes

Regarding the use of these two enzymes in anti-cocaine therapeutic treatment development, both BChE and CocE have their own advantages and disadvantages. BChE from human source can be well tolerated in humans, but the low catalytic efficiency of native BChE against naturally occurring (−)-cocaine must be improved. Compared with native BChE, the catalytic efficiency of native CocE against (−)-cocaine is approximately 800-fold higher. The disadvantages of using native CocE are related to its bacterial origin and thermoinstability. In addition, both BChE (recombinant form of the protein) and CocE have short biological half-lives in animal models. But the biological half-lives could be prolonged through protein engineering. Hence, both BChE and CocE could be engineered to become valuable anti-cocaine therapeutic agents and could complement each other.

In recent years, our research group has successfully designed a series of BChE mutants with a considerably improved catalytic efficiency against (−)-cocaine based on the understanding of catalytic mechanism of the enzymatic cocaine hydrolysis using an integrated computational-experimental approach. One of the BChE mutants designed in our group, i.e. A199S/S287G/A328W/Y332G BChE (denoted as E14-3), fused with human serum albumin (denoted as AlbuBChE) is under human clinical trial by Teva Pharmaceutical Industries Ltd; the Phase II trial has been completed, and the Phase III trial is being planned. When AlbuBChE was administered to human subjects by
intramuscular injection, no significant increase in blood pressure was observed, nor was lethargy reported. The findings indicate that intramuscular administration of AlbuBChE to humans at the specified dosages (50 to 300 mg per subject) did not result in any unacceptable side effects and that the specified dosages will allow for the successful treatment of biological effects of cocaine exposure. Human subjects administrated AlbuBChE did not report a significant increase in cocaine craving. In contrast, the desire to use cocaine again and the “overall drug liking” are significantly decreased following AlbuBChE dosing.19

Another promising BChE mutant designed in our lab, i.e. A199S/F227A/S287G/A328W/Y332G BChE (denoted as E12-7) which has been known as the most active cocaine hydrolase reported, has a ~2000-fold improved catalytic efficiency against (-)-cocaine compared to the wild-type BChE, a ~2-fold improved catalytic efficiency against (-)-cocaine compared to E14-3.61 The in vivo data in a mice model demonstrated that the minimum effective dose of the enzyme to protect mice from cocaine-induced lethality (180 mg/kg) is 0.01 mg per mouse (with a weight of ~30 grams).61

To improve the performance of CocE, our lab also designed a series of CocE mutants. One of the designed CocE mutants, i.e. the T172R/G173Q CocE, has a ~30-fold improved thermostability (in terms of the in vitro half-life at 37°C) compared to the wild-type CocE and maintains its catalytic activity against cocaine.62 The T172R/G173Q CocE is currently in Phase II clinical trial.

In summary, the rationally designed BChE and CocE mutants have shown their values in development of anti-cocaine therapeutic agents. This thesis is a brief summary of our recently accomplished integrated computational-experimental studies that aim to design and discover new mutants of these enzymes for the purpose of further improving their in vitro and in vivo profiles concerning the catalytic activity against (-)-cocaine, thermostability, and biological half-life. To rationally design desirable mutants of the enzymes, we have successfully developed computational models, including those for the BChE gating structure, glycosylated BChE structure, BChE-substrate complex structures, BChE dimer/tetramer structures, CocE monomer/dimer structures, and CocE-substrate
complex structures. Development of the computational models enabled us to rationally design new amino-acid mutations that may improve the catalytic activity, thermostability, and/or biological half-life. The computational design was followed by wet experimental tests, including both *in vitro* and *in vivo* experiments, leading to discovery of new enzyme forms with not only a high catalytic efficiency against cocaine, but also an improved thermostability and prolonged biological half-life. In addition, it has been demonstrated that a PEGylated CocE mutant can be used to fully protect mice from a lethal dose of cocaine (180 mg/kg, LD$_{100}$) for at least three days. The identified new mutants of BChE and CocE are expected to be valuable candidates for development of a more efficient enzyme therapy for cocaine abuse. In addition, the encouraging outcomes of the present study suggest that the structure-and-mechanism-based design and integrated computational-experimental approach is reliable for rational protein drug design and discovery.
Chapter Two: Active-Site Gating and Substrate Specificity of Human
Butyrylcholinesterase and Acetylcholinesterase

To design BChE mutants with improved catalytic efficiency against naturally occurring, biologically active (-)-cocaine, it is necessary to understand the detailed structure and catalytic mechanism of the enzyme. The general catalytic mechanism of BChE hydrolyzing (-)-cocaine has been described in previous reports from our group.\(^\text{61,63,64}\) In this chapter, the active-site gating mechanism will be described as it is very important in understanding how substrates can enter the active site and how the hydrolysis products can leave the active site of the enzyme.\(^\text{65}\) Further, human BChE and human acetylcholinesterase (AChE) are highly homologous proteins with distinct substrate preferences. To gain more structural information about BChE, the active-site gating and substrate specificity of monomers and tetramers of BChE and AChE were also compared after performing molecular dynamics (MD) simulations on the enzymes solvated in water.

### 2.1 Overview of the active-site gating

Butyrylcholinesterase (BChE) and acetylcholinesterase (AChE) are two cholinesterases distinguished by their substrate preferences.\(^\text{66}\) AChE is considered to be specific to its natural substrate acetylcholine (ACh).\(^\text{67}\) By rapidly catalyzing the hydrolysis of ACh, this enzyme is responsible for the termination of cholinergic neurotransmission at neuronal and neuromuscular synapses.\(^\text{68-70}\) AChE was found to exist in multiple forms.\(^\text{71,72}\) So far, X-ray crystal structures of the catalytic domains of different species have been identified. These include human AChE (hAChE),\(^\text{73}\) mouse AChE (mAChE),\(^\text{74}\) and Torpedo California AChE (TcAChE).\(^\text{75}\) Previous studies have shown that the catalytic domains of AChE in various species overlap with each other well.\(^\text{73}\) Identical dimers have been observed among different species of AChE with two opposite-positioned subunits associating with each other to form a four helix bundle,\(^\text{73,76}\) causing the active sites of the two subunits to face in opposite directions. Recently, the model of
an AChE tetramer complexed with collagen-tail protein (ColQ) was built by docking of the tetramerization domain of an AChE tetramer on the catalytic domains of the *Electrophorus electricus* AChE (*Ee*AChE) tetramer.\textsuperscript{72} Kinetic data have shown that AChE is one of the most efficient enzymes, with its catalytic efficiency almost reaching the diffusion limit.\textsuperscript{69} However, crystal structures of AChE indicate that the active site of this enzyme is buried at the bottom of a deep and narrow gorge lined with a few aromatic residues.\textsuperscript{73-75} Many studies have been reported to search for the cause for the high catalytic efficiency.\textsuperscript{69,77-87} The most likely cause is that there may be some alternative doors in addition to the main door that lead to the active site. These alternative doors can be channels for small molecules such as water and ions, so that the entering of the reactants and the leaving of products can be much smoother. So far, three alternative doors have been proposed, including the back door, the side door, and the acyl loop door. The back door was first proposed by Sussman *et al.*\textsuperscript{78} in 1993 and supported by a short MD simulation (119 ps) done by McCammon *et al.*\textsuperscript{79} in 1994. Later, it was also supported by longer MD simulations.\textsuperscript{81-84} This back door is basically formed by four residues, including Trp86, Gly448, Tyr449, and Ile451 (*h*AChE residue numbering). The gating mechanism is described as the temporary side chain (indole ring) movement of Trp86. Later in 1997, a side door was found by McCammon *et al.*\textsuperscript{69,77} using the same method. This side door is mainly formed by six residues, including Asp74, Thr75, Leu76, Thr83, Glu84, and Asn87. The gating mechanism of this door was also mainly due to the fluctuations of the side chains of these residues. Based on previous studies, this side door was only found in the AChE complex, either with small molecule or with protein. As explained in one of the studies done by McCammon *et al.*,\textsuperscript{81} the substrate here may help to induce the structural change of the residues that form the side door. In 1998, another alternative door was proposed by Brooks *et al.*\textsuperscript{86} The method they used is different from the method used by McCammon *et al.*\textsuperscript{79} In the paper, they studied two MD trajectories by using the umbrella sampling method.\textsuperscript{88} They found this door formed by residues Arg247, Phe297, and Trp236 could be a better exit for the acetic acid and acetate ion releasing from AChE. No further information (such as the size and open frequency of this door) has been released on this door in AChE. But a study on BChE in 2005 found evidence for
the existence of an analogous door and investigated the size of the door, although the size of the door is very small, with the average radius of $0.8 \pm 0.2$ Å during the whole simulation time.\textsuperscript{85}

Compared with AChE, the natural function of BChE is unclear;\textsuperscript{23} however, it has been shown that this enzyme is the principal cocaine hydrolase in human plasma. Thus, this enzyme could be a promising option to solve the problems of cocaine abuse as we mentioned in Chapter One. High-activity mutants of this enzyme have become promising drug candidates for cocaine abuse, considering their powerful effects of protecting rodents from cocaine toxicity. The first crystal structure of the catalytic domain of human BChE (hBChE) monomer was identified by Nachon \textit{et al}.\textsuperscript{66,47} Studies have shown that the predominant form of native BChE in plasma is the tetramer.\textsuperscript{55,89} A model of hBChE tetramer complexed with a proline-rich peptide was built by our group recently, based on homology modeling of the AChE tetramer followed by MD simulation of the explicit water solvated system.\textsuperscript{90} In comparison with the high selectivity of AChE, BChE was once called “nonspecific cholinesterase” due to its wide range of substrates, including acetylcholine, butyrylcholine, succinylcholine, organophosphates, and cocaine.\textsuperscript{67} In the study described in this chapter, we compared the main door and alternative doors of the monomers (AChE and BChE) as well as the main door of the tetramers (AChE and BChE) by using MD simulations.\textsuperscript{65} The gating residues of BChE and AChE are found to be responsible for their distinct substrate specificities. The simulations of the tetramers of AChE and BChE indicate that although there are some structural differences, the two active sites of both enzymes facing to the dimer-dimer interface may not function due to their restricted accessibility to substrates.

2.2 Computational methods used to understand the active-site gating

2.2.1 Molecular dynamics simulations of BChE and AChE

Our computational modeling and simulations started from available X-ray crystal structures. The X-ray crystal structures of the hAChE and hBChE monomers were from the Protein Data Bank (PDB ID: 1B41\textsuperscript{91} and 1P0M\textsuperscript{92}, respectively). The structure of ColQ of [AChE]$_4$-ColQ (PDB ID: 1VZJ\textsuperscript{93}) was truncated to the proline-rich attachment
domain (PRAD) consisting of the N-terminal 17 residues to improve the simulation efficiency. Thus, the AChE tetramer (PDB ID: 1C2O\textsuperscript{94}) complexed with PRAD sequence is denoted as [AChE]\textsubscript{4}-PRAD in this chapter. The initial structure of hBChE tetramer with proline-rich peptide CCLMPPPPLFPFFF (denoted as [BChE]\textsubscript{4}-PRAD) was built by our group based on the homology modeling of [AChE]\textsubscript{4}-ColQ.\textsuperscript{90}

The general procedure for carrying out the MD simulations is similar to that used in section 2.2 of chapter 2. Briefly, all energy minimizations and MD simulations were carried out by using the AMBER 9 program package. The Amber force field (ff03) was used to establish the force field of the protein. A counter ion (Na\textsuperscript{+}) was used to neutralize the system and, then, the system was immersed in an orthorhombic box of TIP3P water molecules with a minimum solute-wall distance of 10 Å.\textsuperscript{95} The whole system was carefully equilibrated and fully energy-minimized. After that, the system was gradually heated in the NPT ensemble from 10 K to 300 K over 60 ps. Then, MD simulation was performed for 10 ns under the normally adopted temperature (300 K). During the MD simulation, Particle Mesh Ewald (PME) method was employed to deal with the long-range electrostatic interactions.\textsuperscript{96} The SHAKE procedure was employed to constrain the lengths of all covalent bonds involving hydrogen atoms,\textsuperscript{97} with a time step of 2.0 fs. The atomic coordinates were saved every 1 ps for subsequent sampling and analysis.

2.2.2 Evaluation of the door size

To measure the radii of different doors of AChE and BChE, the MSMS program\textsuperscript{98} was employed here in a way similar to how the method has been used elsewhere.\textsuperscript{69,79,80} Essentially, the MSMS program generates a continuous Connolly surface\textsuperscript{99} of the protein by using a probe as shown in Figure 2.1. The shape of each door is approximately treated as a circle, using radius to describe its size.\textsuperscript{69,79,80} If the radius of the door is bigger than the radius of the current probe, then this probe can enter the door and generate the surface of the gorge which will also be part of the continuous Connolly surface of the protein, as in Figure 2.1 (A); otherwise, the probe cannot enter the gorge through the door and no Connolly surface will be generated, as in Figure 2.1 (B). Thus, we can use probes with different radii to detect the radii of different doors including the main door and the
alternative ones. To precisely measure the radius, a series of artificial dummy atoms were used to block the doors other than the one currently being measured. Here we chose several residues as the probing targets, including Gly122, Glu202, and Ser203 (hAChE residue numbering). If any of these residues appear in the Connolly surface, that suggests they have a direct contact with the probe; otherwise their Connolly surface area value would be zero.

Figure 2.1 The continuous Connolly surface of a protein under different probe radii. (A) Probe with smaller radius can enter the cave. (B) Probe with a larger radius cannot enter the cave.

2.3 Stability of the MD trajectories

To explore dynamic stability of the models, the time-dependent root-mean-square deviation (RMSD) of the atomic positions of the key atoms from their initial positions were evaluated here. Depicted in Figures 2.2 and 2.3 are the RMSD values of the protein backbone of BChE/AChE monomer and tetramer, respectively.
Figure 2.2 Time-dependent RMSD values for the atomic positions of the protein backbone in the MD-simulated enzyme structures. (A) AChE monomer, and (B) BChE monomer.

As illustrated in Figures 2.2 and 2.3, all the trajectories start to enter the relatively stable production phase after 2 ns of MD simulation. The RMSD values are all between the ranges of 1.0 Å to 2.5 Å during the entire simulation time. Based on the RMSD values of these trajectories, snapshots after the 2-ns MD simulation can be extracted for later analysis. Considering the large size of AChE tetramer and BChE tetramer, only the catalytic domain of each subunit (without tetramerization domain) was analyzed here. The stable MD trajectories depicted in Figure 2.3 suggest that the tetramer structures of the enzymes were stabilized during the simulations.
2.4 Active-site doors of AChE and BChE monomers

To search for all the possible doors of each enzyme, the structures extracted from the MD simulation were analyzed. According to the results, there are some similarities and some differences compared with previous studies. In the study presented in this chapter, all the aforementioned doors (main door, side door, back door, and acy loop back
door) were found to exist, although the size of the acyl loop back door was relatively smaller. In previous work, two species of AChE (TcAChE and mAChE) were studied. The hAChE had not previously been examined. Regarding BChE, hBChE is the only species that has been studied. But considering that many X-ray crystal structures have been determined and different initial structures may cause different results, another X-ray crystal structure with the PDB ID of 1P0M, which is used by our group for molecular mechanics (MM) and hybrid quantum mechanics/molecular mechanics (QM/MM) calculations, was considered. For the monomers, we tried to calculate all the possible doors; for the tetraromers, we focused on the main doors.

Figure 2.4 depicts the main entrance radius values of AChE monomer and BChE monomer during the MD simulations. The average radius of the main door of AChE monomer are 1.26 Å with the maximum radius of 2.30 Å as shown in Table 2.1. During the last 4 ns of the MD simulations, the average radius was 1.51 Å. The entrance radius seldom reached the maximum value. This result suggested that during the 10-ns MD simulation, the main door of AChE did not open spontaneously to allow even small substrates, such as acetylcholine (ACh), to enter the active site. This result was similar to that reported by McCammon et al. in which the average radius was 1.52 Å with the maximum radius of 2.40 Å. As they concluded, this maximum radius may not be big enough for even small substrates to enter the active site easily, whereas it was also not likely to hinder its binding (within the active site) significantly. It has the “dynamic selectivity” that can prevent the main door from binding much bigger substrates such as cocaine. Figure 2.4 (B) shows the main entrance radius values of the BChE monomer. The range of the radius was from 1.20 Å to 4.00 Å, which was larger than that of the AChE. The average entrance radius of the main door was 2.75 Å with the maximum radius of 4.00 Å. Compared with the “dynamic selectivity” of AChE, BChE can bind not only small substrates such as ACh, but also large substrates such as cocaine. These data were also close to the results calculated by Suárez et al. (average radius of 3.1 Å and maximum radius of 3.8 Å).
Figure 2.4 The main-door radii of the monomers: (A, top) AChE; (B, bottom) BChE.

Table 2.1 The average and maximum radii of different doors in monomers.

<table>
<thead>
<tr>
<th>Door</th>
<th>Average Radius (Å)</th>
<th>Maximum Radius (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AChE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>main</td>
<td>1.26 ± 0.2</td>
<td>2.30</td>
</tr>
<tr>
<td>back</td>
<td>&lt;1.20</td>
<td>1.80</td>
</tr>
<tr>
<td>side</td>
<td>&lt;1.20</td>
<td>2.10</td>
</tr>
<tr>
<td>acyl loop</td>
<td>&lt;1.20</td>
<td>1.50</td>
</tr>
<tr>
<td>Monomer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>main</td>
<td>2.75 ± 0.3</td>
<td>4.00</td>
</tr>
<tr>
<td>back</td>
<td>1.63 ± 0.2</td>
<td>2.30</td>
</tr>
<tr>
<td>side</td>
<td>1.69 ± 0.2</td>
<td>2.30</td>
</tr>
<tr>
<td>acyl loop</td>
<td>&lt;1.20</td>
<td>1.70</td>
</tr>
</tbody>
</table>
The difference between the two enzymes can be explained by the different structures around the main entrance area. In AChE, according to the MD simulation, there are eight aromatic residues associated with the main door. Three of them are located on the entrance of the active site (Tyr72, Trp286, and Tyr341); five of them are located almost the same depth inside the gorge (Tyr124, Phe295, Phe297, Tyr337, and Phe338), as shown in Figure 2.5 (A). In BChE, however, six of these residues are changed to smaller ones, including Tyr72 to Asn68, Tyr124 to Gln119, Trp286 to Ala277, Phe295 to Leu286, Phe297 to Val288, and Tyr337 to Ala328, with both sets of residues shown in Figure 2.5 (B) and (C). These smaller residues may significantly enlarge the entrance radius, which makes it possible for substrates as large as cocaine to enter the active site.
Figure 2.5 (A, upper) The eight residues with aromatic or indole side chains in the gorge of AChE. (B, middle) The eight residues in the gorge of BChE. (C, bottom) Superimposing of the gating residues in AChE and BChE.
For the monomers, we also analyzed the alternative doors suggested in previous studies. In this study, all these alternative doors were found and shown in Figures 2.6 and 2.7. The average radii of these three doors in AChE are all below 1.20 Å, but with different maximum radii and different fractions of opening snapshots (values larger than 1.40 Å will be considered as open, the approximate radius of a water molecule). A small fraction (139 out of 8000) of the sampling snapshots showed the opening of the back door. This finding is consistent with a previous study in which 78 out of 10000 snapshots showed the back-door opening event. The flip of the indole ring of Trp 82 is believed to be responsible for this opening. In the current study, the side door has a longer opening time; 3463 out of 8000 sampling snapshots showed the open state. The fluctuations of the side chains of residues that form the side door are responsible for the opening.

After the acyl loop back door was first suggested in 1998 by Brooks et al., no further studies had been reported to examine this door in AChE before our analysis of the door in the study described in this chapter. According to the analysis, 77 out of the 8000 snapshots were associated with the opening of acyl loop back door with the maximum value of 1.50 Å. This was the first time that the existence and the size of the acyl loop back door of AChE were observed and detected by performing MD simulation and the Connolly Surface analysis, although the observed opening snapshots were few and the size of the door was small.

In BChE, the characteristics of these three alternative doors have some similarities but also some differences compared with those of AChE. The calculated average radius of the back door is 1.63 Å with the maximum radius of 2.30 Å. The average radius of the side door is 1.69 Å with the maximum radius of 2.30 Å. The calculated values of the radii are systematically larger than the corresponding values calculated previously by Suárez et al., who reported an average radius of 0.8 Å with a maximum radius of 1.2 Å for the back door and an average radius of 1.1 Å with a maximum radius of 2.2 Å for the side door. One of the possible reasons might be associated with the different versions of the Amber force field. The ff03 was used in this MD simulation with Amber 9 program, whereas the specific Amber force field version used for their MD simulation with Amber 7 was not specified.
The acyl loop back door in BChE was also observed in this study. As in AChE, the calculated average radius for the acyl loop back door is smaller than 1.20 Å. The snapshots with the door opening (defined as the radius being larger than 1.4 Å) are even fewer in BChE. 77 snapshots (out of 8000) show the opening event in AChE, whereas only 9 snapshots (out of 8000) showed the opening event in BChE. In the MD simulation reported previously by Suárez et al., no opening event was observed. All of the computational results suggest that the acyl loop back door in BChE may be insignificant for its catalytic function. Overall, for both AChE and BChE, the acyl loop back door is not as important as the other doors.

To summarize the results calculated for the monomers, there are some similarities but also some differences between the AChE and BChE monomers. BChE has a larger main entrance than AChE, which seems reasonable as BChE can bind substrates as large as cocaine. Two alternative doors (side door and back door) in BChE are larger in size than in AChE. The acyl loop back doors in both enzymes are too small to be significant channels leading to the active sites. Based on these results, it is believed that the back door and the side door are more likely and more important alternative doors than the acyl loop back door, as the former ones have larger radii and can exist for longer time. As discussed in a previous study, there is still no direct experimental evidence to prove the function of these alternative doors. However, multiple simulations done by different groups using different forms of enzymes and different force fields have all predicted the existence of these doors (see Figures 2.6 and 2.7 for these alternative doors).
Figure 2.6 The overall and detailed view of different entrances to the active site of AChE. The blue grid means the Connolly surface. The red arrows point from outside of the protein Connolly surface to the inside of active site in (B), (C), (D), and (E). The residues that form the entrances are showed in different color: main door in orange, side door in cyan, back door in yellow, acyl loop door in green, the active site in purple. (A) All doors,
the red arrows point the location of different doors. (B) The main door. (C) The side door. (D) The back door. (E) The acyl loop door.

Figure 2.7 The overall and detailed view of different entrances to the active site of BChE. The blue grid means the Connolly Surface. The red arrows point from outside of the protein Connolly surface to the inside of active site in (B), (C), (D), and (E). The residues
that form the entrances are showed in different color: main door in orange, side door in cyan, back door in yellow, acyl loop door in green, the active site in purple. (A) All doors, the red arrows point the location of different doors. (B) The main door. (C) The side door. (D) The back door. (E) The acyl loop door.

2.5 Active-site doors of AChE and BChE tetramers

Since AChE can exist in the forms of dimer and tetramer and the dominant form of native hBChE in plasma is the tetramer, MD simulations were performed to simulate both [AChE]₄-PRAD and [BChE]₄-PRAD. For the tetramer systems, we mainly focused on the main doors. The overall structures of [AChE]₄-PRAD and [BChE]₄-PRAD were similar, as shown in Figure 2.8. The parts inside the red squares as labeled in Figure 2.8 (A) and (B) indicated the tetramerization domains, while the others were the catalytic domains of the two enzymes. The four subunits of each enzyme were labeled in four different colors. The tetramers were dimers of the dimers with each dimer oriented vertically. As mentioned in a previous paper,⁹⁰ the two dimers were oriented in an anti-parallel pattern, so that the subunits of the tetramers are diagonally equivalent. As shown in Figure 2.8 (A) and (B), the two active sites labeled with triangles were fully exposed to the solvent; the other two active sites labeled with circles faced to the interface between two dimers (labeled with rectangles).
The data summarized in Table 2.2 indicate that for the AChE tetramer, the average radii of the main door of each subunit are 1.67 Å, 0.85 Å, 1.65 Å, and 0.80 Å, with the maximum radii of 2.40 Å, 1.30 Å, 2.30 Å, and 1.20 Å, respectively. In the BChE tetramer, these values are 2.07 Å, 1.24 Å, 2.30 Å, and 1.63 Å, with the maximum radii of 3.70 Å, 2.60 Å, 3.40 Å, and 2.80 Å, respectively. In both enzymes, two active sites facing the solvent have a better chance of binding the substrates compared with the two active sites facing to the dimer-dimer interface, which can be seen in Figures 2.8 and 2.9. Figure 2.9 (A) shows that if the main doors of the enzyme (subunits A and C) that facing to the
solvent are opened wide enough, then substrates such as ACh could enter it by following the direction indicated by the white arrow. However, for the subunits B and D (the left part, labeled in multiple colors in Figure 2.9 (B) to (D)) of these two tetramers, the probability of finding a path that can lead to the active site in subunit B/D without being blocked by subunit A/C would be very low, if not impossible. The conformational flexibility of ACh was considered in this analysis. Illustrated in Figure 2.9 (B) to (D) are representative scenarios in which ACh entering the active site of subunit B/D (left part, labeled in multiple colors) is blocked by subunit A/C (the right part, labeled in green).

Table 2.2 The average and maximum radii of main doors in tetramers

<table>
<thead>
<tr>
<th>Door</th>
<th>Average Radius (Å)</th>
<th>Maximum Radius (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>main (subunit A)</td>
<td>1.67 ± 0.3</td>
<td>2.40</td>
</tr>
<tr>
<td>AChE Tetramer main (subunit B)</td>
<td>0.87 ± 0.1</td>
<td>1.30</td>
</tr>
<tr>
<td>main (subunit C)</td>
<td>1.65 ± 0.2</td>
<td>2.30</td>
</tr>
<tr>
<td>main (subunit D)</td>
<td>0.80 ± 0.1</td>
<td>&lt;1.20</td>
</tr>
<tr>
<td>main (subunit A)</td>
<td>2.07 ± 0.3</td>
<td>3.70</td>
</tr>
<tr>
<td>BChE Tetramer main (subunit B)</td>
<td>1.24 ± 0.1</td>
<td>2.60</td>
</tr>
<tr>
<td>main (subunit C)</td>
<td>2.30 ± 0.2</td>
<td>3.40</td>
</tr>
<tr>
<td>main (subunit D)</td>
<td>1.63 ± 0.2</td>
<td>2.80</td>
</tr>
</tbody>
</table>
Figure 2.9 (A) Acetylcholine (ACh) can enter the active site of AChE/BChE monomer or AChE/BChE tetramer subunits A and C when the main door is open. The white arrow shows the entering direction. (B) to (D) The path to active site of AChE/BChE tetramer subunit B or D (left part, labeled in multiple colors) is blocked by subunit A or C (right part, labeled in green). The white crosses mean the clash between ACh and the subunits. Subunits A and D form one dimer, whereas subunits B and C form the other dimer, in the tetramer. The figure was drawn based on the simulated BChE tetramer structures.

Further, in order to identify the specific residues in subunit A/C of AChE/BChE tetramer that block ACh entering the active site of the complementary subunit B/D, the distances between the key residues on the main door of subunit B/D and the nearby residues of the complementary subunit A/C in the simulated tetramer structures of AChE and BChE were examined carefully. For convenience of discussion, here it was focused on the inter-residue distances between three representative residues on the entrance of the main door in subunit B and residues in complementary subunit A. The three representative residues of AChE are Tyr72, Trp286, and Tyr341, whereas the corresponding residues of BChE are Asn68, Ala277, and Tyr332. The inter-residue
distances between subunits C and D are similar to the corresponding inter-residue distances between subunits A and B. For each snapshot of the MD-simulated tetramer structure, we evaluated the distance between the center of mass of each representative residue in subunit B and that of each residue in complementary subunit A. According to the calculated distances, for AChE, the residues in subunit A that are the closest to Tyr 72, Trp 286, and Tyr 341 in subunit B are Asp320, Phe321, Gln322, and Gln420. The corresponding residues in subunit A of BChE are Leu309, Gly310, Gln311, and Trp412. The calculated inter-residue distances involving these residues of BChE are depicted in Figure 2.10; the plots for the corresponding inter-residue distances in AChE (data not shown) are similar to the results obtained for BChE. The inter-residue distances calculated for the snapshots between 2000 and 10000 ps are used to determine the average inter-residue distance for each pair of residues. The determined average inter-residue distances are summarized in Table 2.3.

Table 2.3 The inter-residue distances ($R_{AB}$, in Å) between the center of mass of representative residues in subunit B and that of the main blocking residues in complementary subunit A of AChE and BChE tetramers.

<table>
<thead>
<tr>
<th>Distance$^a$ in the AChE tetramer</th>
<th>Distance$^a$ in the BChE tetramer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tyr72B</td>
<td>Trp286B</td>
</tr>
<tr>
<td>Asp320A</td>
<td>17.50 (7.02)</td>
</tr>
<tr>
<td>Phe321A</td>
<td>20.67 (8.98)</td>
</tr>
<tr>
<td>Gln322A</td>
<td>15.21 (4.21)</td>
</tr>
<tr>
<td>Gln420A</td>
<td>18.03 (7.30)</td>
</tr>
</tbody>
</table>

$^a$ The distance ($R_{AB}$) between the center of mass of a residue in subunit B and a residue in complementary subunit A. Given in the parentheses are the corresponding $\Delta R_{AB}$ values ($\Delta R_{AB} = R_{AB} - R_A - R_B$, where $R_{AB}$ refers to the average inter-residue distance between residues A and B, and $R_A$ and $R_B$ are the average vdW radii of residues A and B, respectively). The $R_A$ and $R_B$ values used to determine the $\Delta R_{AB}$ values are listed in Table 2.4.
Table 2.4 The calculated average vdW radii (Å) of the residues involved in the inter-residue distance calculations for residues in subunits A and B of the AChE and BChE tetramers.

<table>
<thead>
<tr>
<th>AChE</th>
<th>vdW radius$^a$</th>
<th>BChE</th>
<th>vdW radius$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp320A</td>
<td>4.54</td>
<td>Leu309A</td>
<td>4.80</td>
</tr>
<tr>
<td>Phe321A</td>
<td>5.75</td>
<td>Gly310A</td>
<td>3.98</td>
</tr>
<tr>
<td>Gln322A</td>
<td>5.06</td>
<td>Gln311A</td>
<td>5.01</td>
</tr>
<tr>
<td>Gln420A</td>
<td>4.79</td>
<td>Trp412A</td>
<td>5.90</td>
</tr>
<tr>
<td>Tyr72B</td>
<td>5.94</td>
<td>Asn68B</td>
<td>4.51</td>
</tr>
<tr>
<td>Trp286B</td>
<td>5.94</td>
<td>Ala277B</td>
<td>4.03</td>
</tr>
<tr>
<td>Tyr341B</td>
<td>5.96</td>
<td>Tyr332B</td>
<td>5.79</td>
</tr>
</tbody>
</table>

$^a$ The vdW radius of a residue is defined as the longest distance between the center of mass of the residue and the vdW surface of the residue. The vdW radius calculations were based on the MD trajectories between 2000 and 10000 ps.

The van der Waals (vdW) radius of each residue involved in the inter-residue distance calculations were calculated here, and the average vdW radius for the snapshots between 2000 and 10000 ps were determined. The determined average vdW radii are summarized in Table 2.4. Based on the calculated average inter-residue distances and average vdW radii, another type of distance (which may be called effective inter-residue distance) was evaluated, denoted by $\Delta R_{AB}$:

$$\Delta R_{AB} = R_{AB} - R_A - R_B$$

in which $R_{AB}$ refers to the average inter-residue distance between residues A and B, whereas $R_A$ and $R_B$ are the average vdW radii of residues A and B, respectively. The $\Delta R_{AB}$ values, that are also summarized in Table 2.3 (the values in parentheses), are a better indicator concerning how close the two residues are to each other.

As seen from the $\Delta R_{AB}$ values summarized in Table 2.3, the smallest $\Delta R_{AB}$ value is ~2.9 Å between Ala277 in subunit B and Gln311 in subunit A for the BChE tetramer (see Figure 2.10 (E)). The other two residues (Leu309 and Gly310) in subunit A are also close to Ala277 in subunit B. So, the entrance of the main door in subunit B of the BChE tetramer is blocked mainly by Leu309, Gly310, and Gln311 residues in complementary subunit A, while Trp412 had only a minor role in the blocking.
For the AChE tetramer, the smallest $\Delta R_{AB}$ value is ~2.2 Å between Trp286 in subunit B and Asp320 in subunit A. The other three residues (Phe321, Gln322, and Gln420) in subunit A are also close to Trp286 etc. in subunit B. So, the entrance of the main door in subunit B of the AChE tetramer is blocked mainly by Asp320, Phe321, Gln322, and Gln420 residues in complementary subunit A.

To summarize the results obtained from the comparison between the AChE and BChE tetramers, although there are some differences in the size and shape of the active site cavities of the two tetramers, they share some similarity. Two active sites that face the solvent can be the functional active sites, while the other two are not. Considering the door size and the restricted route a substrate could go through to get close to the active site gorge, the possibility of the substrates entering the two interface-facing active sites of both enzymes could be significantly decreased. This observation is also supported by previous experimental studies showing that a possible AChE tetramer with molecular weight of 250 kDa had two effective active sites per molecule.\textsuperscript{100}
Figure 2.10 (A) – (D) Time-dependence of the key inter-residue distances between the center of mass of representative residues in subunit B and that of the main blocking residues in complementary subunit A in the MD-simulated BChE tetramer. (E) A snapshot of the MD-simulated BChE tetramer, showing the relative positions of the residues involved in the inter-residue distance calculations.
2.6 Conclusion

The active sites and gating mechanisms of AChE and BChE are compared after performing MD simulations on the water-solvated systems of the monomers and tetramers of hBChE and AChE. It has been demonstrated that the different gating mechanisms due to the conformational dynamics of gating residues of AChE and BChE are responsible for their different substrate specificities. The simulations of the tetramers of AChE and BChE indicate that although there are some structural differences, both enzymes could have two dysfunctional active sites due to their special locations.
Chapter Three: Substrate Selectivity of High-Activity Mutants of Human Butyrylcholinesterase

An ideal, therapeutically valuable mutant of human BChE should have not only a high catalytic activity against (-)-cocaine and long biological half-life in human body, but also certain selectivity for (-)-cocaine over neurotransmitter acetylcholine (ACh) such that one would not expect systemic administration of the BChE mutant to interrupt cholinergic transmission. In this study, enzyme-substrate complex models were built and simulated to account for the mutation-caused changes of the catalytic activities of BChE against both (-)-cocaine and ACh. In addition, it is also interesting to know the catalytic activities of the designed BChE mutants against other known substrates, including acetylthiocholine (ATC), butyrylthiocholine (BTC), and (+)-cocaine, in comparison with the corresponding catalytic activity against (-)-cocaine.

3.1 Desirable substrate selectivity of the human BChE mutants

As well known, computational design of high-activity mutants of an enzyme is extremely challenging, particularly when the chemical reaction process is rate-determining for the enzymatic reaction. To computationally design a mutant enzyme with an improved catalytic activity for a given substrate, one needs to design possible amino-acid mutations that can accelerate the rate-determining step of the catalytic reaction process while other steps of the reaction are not slowed down by the mutations. The fundamental reaction pathway for BChE-catalyzed hydrolysis of (-)-cocaine was uncovered in our former studies by extensive molecular dynamics (MD) simulations and reaction-coordinate calculations using quantum mechanics (QM) and hybrid quantum mechanics/molecular mechanics (QM/MM). The computational studies revealed that the rate-determining step of (-)-cocaine hydrolysis catalyzed by the A328W/Y332A and A328W/Y332G mutants of BChE is the first step of the chemical reaction process. Therefore, starting from the A328W/Y332A or A328W/Y332G mutant, rational design of BChE mutants against (-)-cocaine has been focused on
decreasing the energy barrier for the first reaction step (depicted in Scheme 3.1) without significantly affecting the other steps. Our group has developed unique computational strategies and protocols based on the virtual screening of rate-determining transition states of the enzymatic reaction to design enzyme mutants with improved catalytic activity.\textsuperscript{58-60,104-107} The computational design was then followed by \textit{in vitro} experiments, including site-directed mutagenesis, protein expression, and enzyme activity assays, and \textit{in vivo} experiments. The integrated computational-experimental studies have led to discovery of a series of BChE mutants with a significantly improved catalytic efficiency against (-)-cocaine.\textsuperscript{58-60,104-108}

The study described in this chapter was focused on the substrate selectivity of our discovered high-activity mutants of human BChE. Computational modeling was carried out to analyze the catalytic activities of the promising BChE mutants against other known substrates, including acetylcholine (ACh), acetylthiocholine (ATC), butyrylthiocholine (BTC), and (+)-cocaine, in comparison with the corresponding catalytic activity against (-)-cocaine. In particular, ACh is the only known natural substrate of BChE in the body. The catalytic activities of wild-type BChE against ACh, ATC, BTC, and (+)-cocaine are all much higher than that against (-)-cocaine.\textsuperscript{109-111} So, it is interesting to know whether the same amino-acid mutations designed to considerably increase the catalytic activity of BChE against (-)-cocaine also considerably increase the catalytic activities of BChE against other substrates. According to the simulation results, the computationally designed mutations only considerably improve the catalytic efficiency of human BChE against (-)-cocaine, without significantly improving the catalytic activities against other substrates, which were proven by the obtained kinetic data (to be published elsewhere, not shown in this thesis).
Scheme 3.1 The first reaction step of (-)-cocaine hydrolysis catalyzed by a BChE mutant including the A199S mutation.
3.2 Computational methods used to assess the substrate selectivity

Various substrates interacting with human BChE and its mutants were modeled for their enzyme-substrate binding complexes (denoted as ES’s) and transition states for the initial reaction step (denoted as TS1’s) by using the same modeling strategy and approach that we used to study (-)-cocaine interacting with the enzymes. The strategy of performing an energy minimization or molecular dynamics (MD) simulation on a transition state structure of an enzymatic reaction using a classical force field (molecular mechanics) has been described in section 2.2 of chapter 2. During the energy minimization or MD simulation on the TS1 structure, the lengths of transition bonds (i.e. the covalent bonds that are being broken or formed gradually during the initial reaction step) were restrained while all other geometric parameters were allowed to move. The transition bond lengths used in the modeling of the TS1 structures for each pair of enzyme and substrate were based on our previously reported molecular modeling and reaction-coordinate calculations on BChE-catalyzed hydrolysis of (-)-cocaine or ACh or ATC. The computational procedures for modeling a TS1 structure were the same as those for modeling the corresponding ES structure, except for keeping the transition bond lengths restrained during the energy minimization or MD simulation on the TS1 structure.

The initial structures of BChE and the mutants used in the molecular modeling were prepared on the basis of our previous MD simulation on the enzyme-substrate complex for wild-type BChE binding with (-)-cocaine. Our previous MD simulations on the enzyme-substrate complexes started from the X-ray crystal structure deposited in the Protein Data Bank (PDB ID: 1P0M). The procedure for carrying out the MD simulations on the enzyme-substrate interactions in water is the same as that described in section 2.2 of chapter 2.

3.3 Insights from the computational modeling

Molecular modeling enabled us to understand how human BChE and its mutants interact with ACh, ATC, BTC, (+)-cocaine, and (-)-cocaine in the ES and TS1 structures. According to the modeling, for (-)-cocaine interacting with wild-type BChE, there is only
one hydrogen bond (H-bond) between the carbonyl oxygen of (-)-cocaine and the oxyanion hole (G116, G117, and A199) in the ES structure, and there are two H-bonds in the TS1 structure, as seen in Figure 3.1 (A) and (B). With the A199S/F227A/S287G/A328W/Y332G mutant (E12-7), there are two H-bonds between the carbonyl oxygen of (-)-cocaine and the oxyanion hole (G116, G117, and S199) in the ES structure and three H-bonds in the TS1 structure, as seen in Figure 3.1 (C) and (D). The extra H-bond in both the ES and TS1 structures is with the hydroxyl group of S199 after the A199S mutation. The modeled ES and TS1 structures suggest that this mutant should have a significantly higher catalytic activity against (-)-cocaine compared to the wild-type BChE.

Figure 3.1 The energy-minimized ES and TS1 structures for (-)-cocaine interacting with wild-type human BChE and its A199S/F227A/S287G/A328W/Y332G mutant (E12-7).
For substrate ACh interacting with the wild-type BChE or anyone other mutants, there are always two H-bonds between the carbonyl oxygen of the substrate and the oxyanion hole (G116, G117, and A/S199) in the ES structure and three H-bonds in the TS1 structure. Depicted in Figures 3.2 are the modeled ES and TS1 structures with wild-type BChE and a representative mutant E12-7. Depicted in Figures 3.3 are the modeled ES structures with A199S/S287G/A328W/Y332G mutant (noted as E14-3), A199S/F227A/S287G/A328W/E441D mutant (noted as E12-4), and A199S/F227A/S287G/A328W/Y332G/E441D mutant (noted as E13-1).

Figure 3.2 The energy-minimized ES and TS1 structures for ACh interacting with wild-type human BChE and its A199S/F227A/S287G/A328W/Y332G mutant (E12-7).
Figure 3.3 The energy-minimized ES structures for ACh interacting with E14-3, E12-4, and E13-1.

For other substrates (including ATC, BTC, and (+)-cocaine) interacting with the wild-type BChE or E12-7, there are always two H-bonds between the carbonyl oxygen of the substrate and the oxyanion hole (G116, G117, and A/S199) in the ES structure and three H-bonds in the TS1 structure. Depicted in Figures 3.4 to 3.6 are the modeled ES and TS1 structures with wild-type BChE and E12-7. The hydroxyl group of S199 in the mutant does not form an extra H-bond in the ES or TS1 structure for the mutant interacting with any substrate other than (-)-cocaine, suggesting that E12-7 should not have dramatically improved catalytic activities against ATC, BTC, and (+)-cocaine compared to the wild-type BChE.
Figure 3.4 The energy-minimized ES and TS1 structures for ATC interacting with wild-type human BChE and its A199S/F227A/S287G/A328W/Y332G mutant (E12-7).
Figure 3.5 The energy-minimized ES and TS1 structures for BTC interacting with wild-type human BChE and its A199S/F227A/S287G/A328W/Y332G mutant (E12-7).
3.4 Experimental validation

The computational insights were validated by experimental kinetic analysis (see the kinetic parameters summarized in Table 3.1). As seen in Table 3.1, the BChE mutant E12-7 has a 2000-fold improved catalytic efficiency ($k_{cat}/K_M$) against (-)-cocaine. The same mutations do not dramatically improve the catalytic efficiencies of human BChE against the other substrates.
Table 3.1 Kinetic parameters determined for substrate hydrolyses catalyzed by wild-type human BChE and E12-7. The *in vitro* activity assays were performed by Shurong Hou, Liu Xue, and Wenchao Yang in our lab.\(^{116}\)

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Enzyme</th>
<th>(K_M) (µM)</th>
<th>(k_{cat}) (min(^{-1}))</th>
<th>(k_{cat}/K_M) (M(^{-1})min(^{-1}))</th>
<th>RCE(^a)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-)-cocaine</td>
<td>wild-type BChE</td>
<td>4.5</td>
<td>4.1</td>
<td>(9.1 \times 10^5)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E12-7</td>
<td>3.1</td>
<td>5,700</td>
<td>(1.8 \times 10^9)</td>
<td>2,020</td>
</tr>
<tr>
<td>(+)-cocaine</td>
<td>wild-type BChE</td>
<td>4.7</td>
<td>6,420</td>
<td>(1.4 \times 10^9)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E12-7</td>
<td>4.6</td>
<td>8,990</td>
<td>(2.0 \times 10^9)</td>
<td>1.43</td>
</tr>
<tr>
<td>ACh</td>
<td>wild-type BChE</td>
<td>148</td>
<td>61,200</td>
<td>(4.1 \times 10^8)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E12-7</td>
<td>37</td>
<td>11,900</td>
<td>(3.2 \times 10^8)</td>
<td>0.78</td>
</tr>
<tr>
<td>ATC</td>
<td>wild-type BChE</td>
<td>33</td>
<td>20,200</td>
<td>(6.1 \times 10^8)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E12-7</td>
<td>20</td>
<td>14,800</td>
<td>(7.2 \times 10^8)</td>
<td>1.19</td>
</tr>
<tr>
<td>BTC</td>
<td>wild-type BChE</td>
<td>17</td>
<td>29,500</td>
<td>(1.7 \times 10^9)</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>E12-7</td>
<td>13</td>
<td>28,000</td>
<td>(2.2 \times 10^9)</td>
<td>1.24</td>
</tr>
</tbody>
</table>

3.5 Conclusion

In this study, enzyme-substrate complex models were built and simulated to account for the mutation-caused changes of the catalytic activities of BChE and its mutants against both (-)-cocaine and other substrates (including acetylcholine, acetylthiocholine, butyrylthiocholine, and (+)-cocaine). According to the simulation results, the computationally designed mutations only considerably improve the catalytic efficiency of human BChE against (-)-cocaine, without significantly improving the catalytic activities against other substrates. The computational insights have been confirmed by the experimental kinetic analysis.\(^{116}\)

Copyright © Lei Fang 2013
Chapter Four: Model of Glycosylated Human Butyrylcholinesterase

In order to effectively suppress cocaine reward in the brain for a long period of time after an exogenous cocaine hydrolase administration, the therapeutic enzyme should have not only a high catalytic efficiency against cocaine, but also a sufficiently long circulation time. It has been known that PEGylation modification of a therapeutic protein can prolong the biological half-life of the protein without affecting its biological function. However, the asparagine-linked glycans on the surface of glycosylated BChE may interfere with the PEGylation modification. In the study described in this chapter, we built a three-dimensional (3D) model of glycosylated human BChE to investigate the influence of glycans on the PEGylation modification.

4.1 Understanding the detailed structure of a glycosylated protein

To effectively suppress cocaine reward in the brain for a long period of time after an exogenous cocaine hydrolase administration, the therapeutic enzyme should have not only a high catalytic efficiency against (-)-cocaine, but also a sufficiently long circulation time for cocaine addiction treatment. Unfortunately, studies have shown that recombinant BChE expressed in various expressing systems all have a short biological half-life in animal models ranging from minutes to hours. Success in PEGylation modification of peptides and proteins would be an ideal option to improve the biological half-life. However, as the potential PEGylation sites are always located on the protein surface, some post-translational modification, such as glycosylation, may interfere with the PEGylation. So, the nine asparagine-linked glycans on the surface of BChE could probably hinder the PEGylation modification. It is essentially important for rational design of protein engineering, particularly the PEGylation, to know the three-dimensional (3D) structure of the glycosylated human BChE.

On the other hand, it is extremely challenging to obtain a single crystal of a glycosylated protein and determine the crystal structure, particularly for human BChE. In order to prepare a single crystal of human BChE, one had to make certain mutations on
the glycosylation sites.\textsuperscript{92} Hence, available X-ray crystal structures of human BChE or its mutant do not include all structural information about the glycans of the native protein.\textsuperscript{120}

In the study described in this chapter, a 3D model of the glycosylated human BChE starting from the available X-ray crystal structure of the unglycosylated BChE was modeled and simulated. According to the simulated results, glycans did not change the overall stability of the BChE structure, but could increase the flexibility of some local structures. The Solvent Accessible Surface Area (SASA) analysis revealed that some lysine residues have a significant decrease in SASA due to direct or indirect effects of their surrounding glycans. The modeling and simulation provide valuable insights concerning how the glycans affect the BChE structure and how the glycans influence the PEGylation modification of BChE. These insights will be useful in guiding future rational design of BChE engineering.

4.2 Computational methods used to study the structure of glycosylated BChE
4.2.1 Construction of the glycosylated BChE model

Our computational modeling the glycosylated human BChE structure started from the unglycosylated BChE structure discussed in chapters 2 and 3. The starting (unglycosylated) BChE structure was taken from the X-ray crystal structure deposited in Protein Data Bank at 2.0 Å resolution (PDB ID: 1P0M).\textsuperscript{92} A few modifications were made to fix this structure. First, all small molecules and the carbohydrate chains were removed, except the crystal waters. Second, the missing residues (Asp2, Asp3, Asp378, Asp379, and Gln455) were added manually by TLEAP module in Amber 9 package.\textsuperscript{121} Further, the mutated residues Gln17, Gln455, Gln481, and Gln486 were changed back to Asn17, Asn455, Asn481, and Asn486, respectively. Finally, the protonation states of histidine residues were modified after manually inspecting their local environments. His126, His214, His372, and His438 were protonated on their Nδ atoms, whereas His77, His207, and His423 were protonated on their Nε atoms. This initial structure was then optimized using the SANDER module of the Amber 9 package\textsuperscript{121,122} for constructing the glycosylated BChE model.
The initial model of the glycosylated human BChE was constructed based on the above optimized BChE structure using the Glycam Biomolecule Builder. Nine biantennary glycans are located on the surface of BChE. Each glycan is covalently linked to the Nδ atom of asparagine (Asn17, Asn57, Asn106, Asn241, Asn256, Asn341, Asn451, Asn481, and Asn486). Each of the simulated glycans was composed of 11 carbohydrate residues (see Figure 4.1), according to the experimental results based on serum human BChE from Kolarich et al. For convenience, the nine glycans are named in accordance with their linkage residue number, i.e. Glycan17, Glycan57, Glycan106, Glycan241, Glycan256, Glycan341, Glycan451, Glycan481, and Glycan486.

\[
\text{Neu5Ac2} \rightarrow 6\text{Galβ1} \rightarrow 4\text{GlcNacβ1} \rightarrow 2\text{Manα1} \rightarrow 6\text{Manβ1} \rightarrow 4\text{GlcNacβ1} \rightarrow 4\text{GlcNac}
\]

Figure 4.1 The composition of the glycan in human BChE.

4.2.2 Molecular dynamic simulations

The procedure for carrying out the molecular dynamic (MD) simulations is similar to that used section 2.2 of chapter 2. Briefly, all energy minimizations and MD simulations were carried out by using the AMBER 9 program package. The Amber force field (ff03) was used to establish the potentials of protein, and Glycam06 force field was used to establish the potentials of glycans.

The initial model of the glycosylated human BChE was built based on the fully energy-minimized BChE structure and, then, was further carefully equilibrated via MD simulation and energy-minimized again at the end. The same procedure used to simulate the unglycosylated human BChE structure was employed to simulate the glycosylated human BChE structure for 10 ns.

4.3 Structural features of the glycosylated BChE

Through molecular modeling and MD simulations, a dynamically stable 3D model of the glycosylated human BChE in comparison with that of the unglycosylated human BChE was obtained. To explore dynamic stability of the models, the time-dependent root-mean-square deviation (RMSD) of the atomic positions of the key atoms
from their initial positions was evaluated. Depicted in Figure 4.2 are the RMSD values of the protein backbone and the heavy atoms of glycans. As seen in Figure 4.2 (A), the RMSD values for the protein backbone in both models (i.e. the unglycosylated and glycosylated human BChE structures) reached the equilibrium after 1 ns during the MD simulations and kept stable in the remaining time of the simulations. The average RMSD value of the backbone of the unglycosylated BChE during the stable stage was 1.26 Å which is close to that of the glycosylated BChE, 1.24 Å, suggesting that the dynamic stability of BChE at room temperature was not changed significantly by the glycans.

Figure 4.2 Time-dependent RMSD values for the atomic positions of the protein backbone and glycan backbone in the MD-simulated unglycosylated and glycosylated human BChE structures. (A) BChE (red) and Glycosylated BChE (black); (B) Glycan17 (black), Glycan57 (red), Glycan106 (green), Glycan241 (blue), and Glycan256 (pink); (C) Glycan341 (black), Glycan455 (red), Glycan481 (green), and Glycan486 (blue).

As seen in Figures 4.2 (B) and (C), after 1 ns of the MD simulation, the RMSD values for the heavy atoms of the glycans all reached the dynamically equilibrium stage, with the RMSD values fluctuating between ~2 and ~6 Å. The average RMSD values of the heavy atoms of Glycan17, Glycan57, Glycan106, Glycan241, Glycan256, Glycan341, Glycan455, Glycan481, and Glycan486 were 3.37 Å, 4.65 Å, 4.38 Å, 3.52 Å, 3.55 Å, 4.09 Å, 3.79 Å, 4.33 Å, and 4.32 Å, respectively.

Depicted in Figure 4.3 is the MD-simulated structure (the last snapshot) of the glycosylated BChE. Nine glycans are located on the surface of the enzyme. If α-helix where residue Ser198 is located can be viewed as a virtual central axis of the protein,
there are more glycans (Glycan17, Glycan57, Glycan106, Glycan241, Glycan455, Glycan481, and Glycan486) on the left side of the protein than the right side (Glycan241, Glycan256, and Glycan341). In addition, glycans are all far away from the active-site entrance of the enzyme. Hence, the glycosylation should have no significant influence on the catalytic function of human BChE.

Figure 4.3 Overall structure of the glycosylated human BChE model derived from the MD simulation. The orange ribbons represent BChE, and the blue sticks refer to the nine glycans. The cyan sticks represent the nine linkage asparagine residues. The red ribbon refers to the α-helix where Ser198 is located.

It is interesting to know the potential hydrogen bonding between the glycans and BChE itself. Hence, hydrogen bond analysis was performed using PTRAJ module in AMBER 9 package based on the sampling snapshots (once every 20 ps after 1 ns) from the MD simulation. The hydrogen bond analysis was based on a simple criterion that the distance between the hydrogen and the hydrogen-bond acceptor (oxygen or nitrogen) was shorter than 3.0 Å. Hydrogen bonds were counted according to their overall appearance time during the simulation. For example, if a hydrogen bond existed in 50% of the
collected snapshots, it was counted as half (0.5) of a hydrogen bond. According to the results, there were only 1.7 hydrogen bonds in average formed between each glycan and the BChE protein during the MD simulation, suggesting that the hydrogen bonding between the glycans and the BChE protein is moderate.

4.4 Solvent Accessible Surface Area (SASA)

It has been known that PEGylation can significantly prolong the biological half-life of a therapeutic protein without affecting its biological functions. BChE can be PEGylated through lysine linkage chemistry, as there are lysine residues on the BChE surface. However, asparagine-linked glycans on the surface of glycosylated BChE may interfere with the PEGylation modification of the enzyme. In the study described in this chapter, we tried to evaluate the influence of the steric hindrance of glycans on the potential PEGylation reaction sites, specifically the Nε atoms of lysine residues. To achieve this, the SASA values of the Nε atoms of all the 31 lysine residues on the enzyme surface were calculated and analyzed by using the MSMS program for both the unglycosylated and glycosylated BChE models. Three probes with different radii were chosen to represent the approximate sizes of different functional groups of different PEGylation reagents. A probe with a radius of 1.4 Å represents a small functional group, such as carbonyl group. A probe with a radius of 4 Å represents a medium functional group, such as epoxide group. A probe with a radius of 8 Å represents a large functional group, such as benzotriazole group (see Figure 4.4).

**Figure 4.4 Representative PEGylation reactions.**

A: $\text{PEG} - \text{O} - \text{H} + \text{H}_2\text{N} - \text{R} \rightarrow \text{PEG} - \text{O} - \text{NHR}$

B: $\text{PEG} - \text{O} - \text{C} - \text{CH}_2 + \text{H}_2\text{N} - \text{R} \rightarrow \text{PEG} - \text{O} - \text{OH} - \text{NHR}$

C: $\text{PEG} - \text{O} - \text{N} - \text{N} + \text{H}_2\text{N} - \text{R} \rightarrow \text{PEG} - \text{O} - \text{C} - \text{NH}_2\text{R}$
The relative SASA values of the Nε atoms of lysine residues calculated using different probe radii in the unglycosylated and glycosylated BChE models are listed in Table 4.1. In the unglycosylated BChE, the relative SASA values of the Nε atoms calculated using the radii of 1.4, 4, and 8 Å range from 4.79% to 45.99%, 0.01% to 25.83%, and 0.0% to 16.97% with the average values of 30.12%, 11.34%, and 5.77%, respectively. In the glycosylated BChE model, the relative SASA values of the Nε atoms of lysine residues calculated using probe radii of 1.4, 4, and 8 Å range from 6.17% to 42.06%, 0.01% to 20.23%, and 0.0% to 10.67% with the average values of 25.98%, 7.84%, and 3.06%, respectively. In general, the relative SASA values of the Nε atoms of lysine residues on the enzyme surface decrease as the probe radius increases in both the unglycosylated and glycosylated BChE models, demonstrating that a larger probe radius or a larger PEGylation reagent has less chance to contact with the reaction sites (the Nε atoms of lysine residues) and, thus, is expected to have a lower PEGylation reaction rate.
Table 4.1 The relative SASA values of the Nε atoms of lysine residues on the surface of human BChE calculated by using different probe radii.

<table>
<thead>
<tr>
<th>Residue</th>
<th>SASA in unglyco-BChE&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SASA in glyco-BChE&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>R = 1.4 Å</td>
<td>R = 4 Å</td>
</tr>
<tr>
<td>Lys9</td>
<td>43.70%</td>
<td>21.05%</td>
</tr>
<tr>
<td>Lys12</td>
<td>45.99%</td>
<td>19.92%</td>
</tr>
<tr>
<td>Lys44</td>
<td>29.58%</td>
<td>7.93%</td>
</tr>
<tr>
<td>Lys45</td>
<td>27.86%</td>
<td>7.23%</td>
</tr>
<tr>
<td>Lys51</td>
<td>41.39%</td>
<td>16.52%</td>
</tr>
<tr>
<td>Lys60</td>
<td>35.19%</td>
<td>15.91%</td>
</tr>
<tr>
<td>Lys103</td>
<td>25.57%</td>
<td>4.40%</td>
</tr>
<tr>
<td>Lys105</td>
<td>42.93%</td>
<td>25.82%</td>
</tr>
<tr>
<td>Lys131</td>
<td>20.87%</td>
<td>2.53%</td>
</tr>
<tr>
<td>Lys180</td>
<td>25.71%</td>
<td>5.40%</td>
</tr>
<tr>
<td>Lys190</td>
<td>42.47%</td>
<td>22.43%</td>
</tr>
<tr>
<td>Lys248</td>
<td>39.73%</td>
<td>19.92%</td>
</tr>
<tr>
<td>Lys262</td>
<td>35.38%</td>
<td>16.16%</td>
</tr>
<tr>
<td>Lys267</td>
<td>29.02%</td>
<td>11.45%</td>
</tr>
<tr>
<td>Lys313</td>
<td>6.75%</td>
<td>0.64%</td>
</tr>
<tr>
<td>Lys314</td>
<td>35.78%</td>
<td>17.37%</td>
</tr>
<tr>
<td>Lys323</td>
<td>23.53%</td>
<td>6.55%</td>
</tr>
<tr>
<td>Lys339</td>
<td>32.99%</td>
<td>8.78%</td>
</tr>
<tr>
<td>Lys348</td>
<td>38.05%</td>
<td>19.40%</td>
</tr>
<tr>
<td>Lys355</td>
<td>31.37%</td>
<td>10.70%</td>
</tr>
<tr>
<td>Lys366</td>
<td>4.79%</td>
<td>0.31%</td>
</tr>
<tr>
<td>Lys407</td>
<td>20.69%</td>
<td>0.61%</td>
</tr>
<tr>
<td>Lys408</td>
<td>10.88%</td>
<td>0.60%</td>
</tr>
<tr>
<td>Lys427</td>
<td>10.35%</td>
<td>0.01%</td>
</tr>
<tr>
<td>Lys458</td>
<td>39.02%</td>
<td>18.53%</td>
</tr>
<tr>
<td>Lys469</td>
<td>14.93%</td>
<td>1.57%</td>
</tr>
<tr>
<td>Lys476</td>
<td>28.91%</td>
<td>2.46%</td>
</tr>
<tr>
<td>Lys494</td>
<td>41.90%</td>
<td>18.84%</td>
</tr>
<tr>
<td>Lys499</td>
<td>37.96%</td>
<td>17.04%</td>
</tr>
<tr>
<td>Lys513</td>
<td>31.80%</td>
<td>13.12%</td>
</tr>
<tr>
<td>Lys528</td>
<td>38.61%</td>
<td>18.47%</td>
</tr>
</tbody>
</table>

<sup>a</sup>SASA of the Nε atom of lysine residues in the unglycosylated BChE.

<sup>b</sup>SASA of the Nε atom of lysine residues in the glycosylated BChE.

<sup>c</sup>R refers to the probe radius (1.4, 4, or 8 Å).

<sup>d</sup>The relative SASA value refers to the calculated exposed percent of the surface area of the Nε atom. The total surface area of the Nε atom (with a van der Waals radius of 1.55 Å) is $4\pi(R+1.55)^2$ in which R is the probe radius.
To further analyze the steric hindrance of the glycosylation on the surface lysine residues, those lysine residues that showed a significant decrease in SASA (when the Percent of the Remain (POR) of SASA is less than 80%) are selected for further discussion, and those residues are listed in Table 4.2. When the probe radius was 1.4 Å, nine lysine residues were found to have a large decrease in SASA: Lys44, Lys45, Lys190, Lys248, Lys314, Lys 339, Lys355, Lys458, and Lys476. This group of residues, denoted as Group I for convenience, has the POR values ranging from 19.67% to 78.57%. When the probe radius was increased to 4 Å, four more lysine residues (Lys105, Lys348, Lys494, and Lys499) were added to the Group I, resulting in Group II (Lys44, Lys45, Lys105, Lys190, Lys248, Lys314, Lys 339, Lys355, Lys348, Lys458, Lys476, Lys494, and Lys499), with the POR values ranging from 1.31% to 76.13%. When the probe radius was further increased to 8 Å, 15 lysine residues (i.e. the all the lysine residues listed in Table 4.2, denoted as Group III for convenience) showed a large decrease in SASA with the POR values ranging from 0.22% to 73.35%. Clearly, as the probe radius increases, the number of lysine residues with larger POR values increases, suggesting that the effects of the glycans on the PEGylation modification reactions are more significant when a larger reagent is used.
Table 4.2 The Percent of Remain (POR) of SASA of the Nε atoms of lysine residues on BChE surface calculated using different probe radii

<table>
<thead>
<tr>
<th>Residue</th>
<th>R = 1.4 Å\textsuperscript{a}</th>
<th>R = 4 Å</th>
<th>R = 8 Å</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lys44</td>
<td>69.78%\textsuperscript{b}</td>
<td>46.78%</td>
<td>17.70%</td>
</tr>
<tr>
<td>Lys45</td>
<td>70.17%</td>
<td>35.55%</td>
<td>12.00%</td>
</tr>
<tr>
<td>Lys60</td>
<td>99.37%</td>
<td>88.81%</td>
<td>73.35%</td>
</tr>
<tr>
<td>Lys105</td>
<td>98.39%</td>
<td>78.35%</td>
<td>62.88%</td>
</tr>
<tr>
<td>Lys190</td>
<td>78.57%</td>
<td>43.69%</td>
<td>17.13%</td>
</tr>
<tr>
<td>Lys248</td>
<td>66.70%</td>
<td>29.77%</td>
<td>12.98%</td>
</tr>
<tr>
<td>Lys262</td>
<td>95.05%</td>
<td>86.63%</td>
<td>71.33%</td>
</tr>
<tr>
<td>Lys314</td>
<td>32.20%</td>
<td>11.28%</td>
<td>2.19%</td>
</tr>
<tr>
<td>Lys339</td>
<td>73.60%</td>
<td>59.11%</td>
<td>29.21%</td>
</tr>
<tr>
<td>Lys348</td>
<td>88.94%</td>
<td>76.13%</td>
<td>54.30%</td>
</tr>
<tr>
<td>Lys355</td>
<td>19.67%</td>
<td>1.31%</td>
<td>0.22%</td>
</tr>
<tr>
<td>Lys458</td>
<td>68.84%</td>
<td>22.40%</td>
<td>1.89%</td>
</tr>
<tr>
<td>Lys476</td>
<td>59.18%</td>
<td>21.54%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Lys494</td>
<td>88.76%</td>
<td>75.85%</td>
<td>51.53%</td>
</tr>
<tr>
<td>Lys499</td>
<td>98.29%</td>
<td>72.01%</td>
<td>14.30%</td>
</tr>
</tbody>
</table>

\textsuperscript{a}R represents the probe radius (1.4, 4, or 8 Å).

\textsuperscript{b}The POR of SASA of the Nε atom of lysine residues in the glycosylated BChE compared to SASA of the Nε atom of lysine residues in the unglycosylated BChE. POR = SASA of the Nε in glycol-BChE / SASA of the Nε in (glycosylated) BChE.

To understand the relation between the lysine residues with the large decrease in SASA and their locations on the enzyme surface, we searched for the lysine residues that are located within a certain distance (4 or 8 Å) from each glycan. Specifically, the distance between a lysine residue and a glycan is defined as the internuclear distance between the Nε atom of the lysine and a glycan atom closest to the Nε atom. Based on the collected snapshots of the MD-simulated structure, 12 lysine residues were found to locate within 4 Å from a glycan: Lys44, Lys45, Lys60, Lys105, Lys190, Lys248, Lys313, Lys314, Lys348, Lys355, Lys458, and Lys476 (denoted as Group IV). 17 lysine residues were found to locate within 8 Å from a glycan: Lys44, Lys45, Lys60, Lys103, Lys105, Lys190, Lys248, Lys261, Lys313, Lys314, Lys348, Lys355, Lys366, Lys427, Lys458, Lys469, and Lys476 (denoted as Group V). Compared to Group IV, Group V, has five more residues (Lys103, Lys261, Lys366, Lys427, and Lys469). However, according to
the data in Table 4.1, Lys103, Lys366, Lys427, and Lys469 all have very small SASA values calculated using all of the three probe radii in both the unglycosylated and glycosylated BChE models. This means that, even without glycans on enzyme surface, theses lysine residues are hardly available to contact with the PEGylation reagents. So, ignoring these four lysine residues, Groups IV and V are almost the same, with Lys262 being the only additional lysine residue in Group V.

In comparison between Groups I and IV, it was noted that eight of the nine residues in Group I were included in Group IV, suggesting that glycans on BChE surface indeed mainly decrease the SASA of the Nε atoms of most lysine residues close to the glycans. However, Lys339 in Group I was not included in Group IV. After inspecting the location of Lys339, it was noted that Lys339 is located in an area surrounded by loop residues (Gln71-Glu80, Tyr332-Asn341, and Glu421-Glu432). The glycan closest to Lys339 is Glycan341, which is more than 10 Å away (see Figure 4.5 (A)). The large decrease in SASA of the Nε atom of Lys339 may be due to the amino-acid residues located in these highly flexible loops which may more significantly block the contact between Lys339 and the probe in the glycosylated BChE. In fact, the RMSD values of the backbone atoms of these loops (Gln71-Glu80, Tyr332-Asn341, and Glu421-Glu432) in the MD-simulated glycosylated BChE were 0.84 Å, 1.27 Å, and 0.71 Å, respectively, larger than the corresponding RMSD values of 0.77 Å, 1.19 Å, and 0.55 Å in the MD-simulated unglycosylated BChE. The glycans might not significantly change the overall stability of the BChE protein as mentioned above, but they may increase the flexibility of local protein structures. From this point of view, Lys399 is indirectly influenced by the glycosylation. In Group I, Lys314 and Lys355 are affected most by the glycosylation with the POR values being 32.20% and 19.67%, respectively. As seen in Figure 4.5 (B) and (C), Lys314 is very close to some carbohydrate residues of Glycan106, and Lys355 is almost fully blocked by several carbohydrate residues at the end of Glycan241.

Of the 13 lysine residues in Group II, 10 residues were included in Group IV. Besides the aforementioned Lys339, residues Lys94 and Lys499 were also not included in Group IV. These two residues were located in the same loop (Asn481-Lys499) and they were very close to each other. Similar to Lys399, the large decrease in SASA of the
Ne atom of Lys494 and Lys499 was likely due to the reason that the amino-acid residues located in the highly flexible loop block the contact between Lys494/Lys499 and the probe. The RMSD value of the backbone atoms of the loop where these two residues are located in the MD-simulated glycosylated BChE was 1.22 Å, larger than that (0.90 Å) in the MD-simulated unglycosylated BChE. The larger RMSD in the MD-simulated glycosylated BChE was likely caused by Glycan481 and Glycan486 that are all connected to the same loop as Lys494 and Lys499 (see Figure 4.5 (D) for the detail). These computational data are consistent with previously reported experimental observation that Asn481 and Asn486 must be mutated to glutamine residues to avoid the glycosylation on these two sites and, thus, to obtain a single crystal for the X-ray diffraction. The probe radius also plays an important role here. According to the data in Table 4.2, a larger probe radius is associated with the smaller POR values, indicating that larger reagents would amplify the influence of the glycosylation. The POR value of Lys314 decreased from 32.20% associated with a probe radius of 1.4 Å to 11.28% associated with a probe radius of 4 Å.
Figure 4.5 Locations of representative lysine residues. The orange ribbons represent BChE. The blue sticks represent the glycans. The cyan sticks refer to the linkage residue asparagine. The green sticks represent the surrounding residues. The red ribbon refers to the α helix where the Ser198 located. The pink sticks represent the lysine residues: (A) Lys339; (B) Lys314; (C) Lys355; (D) Lys494 and Lys499.

Compared to the fact that Groups I and II only included some of the lysine residues in Table 4.2, Group III included all of the 15 lysine residues in Table 4.2, which further reflects the aforementioned observation that the steric hindrance of the glycosylation on the lysine residues is more significant when a larger probe is used. In comparison between Groups III and IV, 11 of the 14 lysine residues listed in Group III are also listed in Group IV; the remaining four residues are Lys262, Lys339, Lys494, and
Lys499. Of these four exceptions, three (i.e. Lys339, Lys494, and Lys499) have been discussed above, and the only newly added one is Lys262. Notably, Lys262 is included in Group V, suggesting this residue is still significantly affected by the glycosylation when a larger PEGylation agent is used.

4.5 Conclusion

In the study described in this chapter, a 3D structural model of the glycosylated human BChE has been developed to examine the effects of the glycosylation on the possible PEGylation modification reactions. The calculated Solvent Accessible Surface Area (SASA) values revealed that most surface lysine residues located within 8 Å, or more precisely 4 Å, of glycans in the glycosylated BChE model have a significant decrease in SASA compared to the corresponding SASA values in the unglycosylated BChE, demonstrating the direct influence of the glycosylation. Three lysine residues (Lys399, Lys494, and Lys499) that are not very close to the glycans still have a significant decrease in SASA due to the indirect influence of the glycosylation.
Chapter Five: Computational Design of Amino-Acid Mutations to Prolong the Biological Half-life of a Highly Efficient Cocaine Hydrolase Engineered from Human Butyrylcholinesterase

Our recently designed and discovered cocaine hydrolase, particularly E12-7 engineered from human butyrylcholinesterase (BChE), has the promise of becoming a valuable cocaine abuse treatment. An ideal anti-cocaine therapeutic enzyme should have not only a high catalytic efficiency against cocaine, but also a sufficiently long biological half-life. However, recombinant human BChE and the known BChE mutants have a much shorter biological half-life compared to the native human BChE. The present study aimed to extend the biological half-life of the cocaine hydrolase without changing its high catalytic activity against cocaine by using an integrated computational-experimental approach. The strategy was to computationally design possible amino-acid mutations that can introduce cross-subunit disulfide bond(s) and, thus, change the distribution of the oligomeric forms and extend the biological half-life.

5.1 Overview of our strategy to prolong the biological half-life

In order to effectively suppress cocaine reward for a long period of time after administration of an exogenous CocH, the therapeutic enzyme (CocH) should have not only a high catalytic efficiency against cocaine, but also a sufficiently long circulation time (biological half-life). Native human BChE has a biological half-life of ~24 hours in mice and ~7 to 12 days in humans. However, recombinant forms of wild-type human BChE and the known BChE mutants have a much shorter biological half-life compared to the native human BChE. The difference between the native and recombinant human BChE proteins in biological half-life is associated with the difference in the distribution of the oligomeric forms and the post-translational modification. Native BChE consists of more than 95% of tetramer, whereas predominant forms of recombinant BChE are monomer and dimer. In addition, the native BChE is fully glycosylated with whole
nine N-linked oligosaccharides, whereas recombinant BChE is either not fully glycosylated or close to native BChE.\textsuperscript{119,133-135}

The present study aimed to extend the biological half-life of E12-7, \textit{i.e.} the A199S/F227A/S287G/A328W/Y332G mutant of human BChE, without changing its catalytic activity against cocaine. It is very popular for extending the biological half-life of a protein to chemically modify the protein surface with polyethylene glycol (PEG). However, it is well-known that PEG can elicit an immune response in humans.\textsuperscript{136-138} For this reason, we are reluctant to PEGylate the human protein which is accommodated perfectly in humans. We prefer to extend the biological half-life of E12-7 through rationally designed further mutations on its amino-acid residues that are not available for intermolecular interactions with any other proteins in the body; such type of amino-acid mutations is not expected to produce immune response.

On the other hand, it is a grand challenge to design amino-acid mutations that can extend the biological half-life of a protein, although computational design has been used successfully to identify thermostable mutants of proteins. This is because the biological half-life of a protein is determined by many factors, in addition to the thermostability. For a protein to have a desirably long biological half-life, the protein must be thermostable enough at the body temperature (37°C) to have an \textit{in vitro} half-life at 37°C which is at least not shorter than the desirable biological half-life. So, when a protein has a short biological half-life because the protein is thermally unstable, one may computationally design a thermostable mutant of the protein to extend the biological half-life.\textsuperscript{62} However, for most proteins, their short biological half-lives are due to factors other than the thermostability. Specifically for the therapeutically interesting protein concerned in the study described in this chapter, BChE (wild-type or the mutant) is very thermostable at 37°C. But the recombinant form of the thermostable protein is quickly eliminated from the body. So, further improving the protein thermostability is not expected to extend its biological half-life. One must account for other factors affecting the biological half-life.

In the study described in this chapter, computational modeling of E12-7 was performed to design further amino-acid mutations that can favorably change the distribution of the oligomeric forms and, thus, extend the biological half-life. In particular,
the computational design was focused on possible mutations that may introduce disulfide bond(s) between different subunits in the multimeric forms (dimer and tetramer) and, thus, stabilize the multimeric forms and decrease the concentration of the monomer. This computational design strategy was based on the observation that the multimeric forms of BChE have a longer biological half-life compared to the monomer. The computational design was followed by experimental studies in vitro and in vivo, leading to discovery of novel BChE mutants that not only retain the high catalytic activity of E12-7 against cocaine, but also have a significantly prolonged biological half-life, demonstrating that it is possible to computationally design amino-acid mutations on therapeutic proteins with a prolonged biological half-life without changing the biological functions.

5.2 Computational methods used to performed for the mutant design

The initial BChE structure modeled in the study described in this chapter came from our previously reported computational study on the BChE tetramer. Based on the tetramer structure, the tetramer consists of two equivalent dimers, and the computational design was focused on the amino acid residues on the inter-subunit interface within a dimer. The procedure for carrying out the molecular dynamics (MD) simulation on the dimer in this study was similar as that used in section 2.2 of chapter 2. All molecular mechanics optimization and MD simulation were carried out by using the AMBER 9 program package. The MD simulation was performed under normal temperature (300 K) for 50 ns.

5.3 Methods for in vitro studies

5.3.1 Materials used for in vitro studies

Cloned Pfu DNA polymerase and DpnI endonuclease were obtained from Stratagene (La Jolla, CA). All oligonucleotides were synthesized by the Eurofins (Huntsville, AL). The QIAprep Spin Plasmid Miniprep Kit was obtained from QIAGEN (Valencia, CA). Chinese hamster ovary (CHO) cells and culture medium were from life technologies (Grand Island, NY). The transfection kit was from Mirus Bio LLC (Madison, WI). Anti-BChE (mouse monoclonal antibody) was purchased from AntibodyShop
(Gentofte, Denmark), and goat anti-mouse IgG HRP conjugate was from Zymed Laboratories (South San Francisco, CA). \[^{3}H\](−)-Cocaine (50 Ci/mmol) was purchased from PerkinElmer Life and Analytical Sciences (Waltham, MA). QFF ion exchanger was from GE Healthcare.

5.3.2 Site-directed mutagenesis

Site-directed mutagenesis of human BChE cDNA was performed by using the QuikChange method. Further mutations required to produce a new BChE mutant cDNA were generated from the cDNA corresponding to the A199S/F227A/S287G/A328W/Y332G mutant of human BChE in a pRc/CMV expression plasmid. Using plasmid DNA as template and primers with specific base-pair alterations, mutations were made by polymerase chain reaction with Pfu DNA polymerase for replication fidelity. The PCR product was treated with DpnI endonuclease to digest the parental DNA template. The digested product was transformed into *Escherichia coli*, amplified, and purified. The DNA sequences of the mutants were confirmed by DNA sequencing.

5.3.3 Protein expression

The A199S/F227A/S287G/A328W/Y332G BChE and newly designed BChE mutants were expressed in CHO-S cells in free-style CHO expression medium. Cells were first grown to a density of \(\sim 1.0 \times 10^6\) in 1L shake flask and transfected using TransIT-PRO Transfection Kit. Cells were incubated at 37°C in a CO\(_2\) incubator for 6 days. The culture medium was then harvested for the BChE activity assays.

5.3.4 Protein purification

Purification of each enzyme mutant in medium was achieved by using an ion exchange chromatography. In brief, 1L of medium with a BChE mutant was diluted with the same volume of 20 mM Tris-HCl, pH 7.4. Pre-Equilibrated QFF anion exchanger was added to the diluted medium in 1% of its volume and incubated at 4°C with shaking (100 rpm) overnight. More than 95% enzyme activity was found to bind to the resin after the
incubation. The suspension was then packed in a column (5 × 50 cm), and the medium was allowed to flow through rapidly with the aid of suction (50–100 ml/min). The QFF resin was repacked again in a washing buffer after all of the medium was excluded. After washing the column with 20 mM Tris-HCl, pH 7, the enzyme was eluted by 20 mM Tris-HCL, pH 7.0, plus 0.3 M NaCl. The eluate was concentrated and changed to phosphate-buffered saline by Millipore centrifugal filter device. Finally, the purified enzyme was stored at 4°C or −80°C. The entire purification process was carried out in cold room at 4°C.

5.3.5 Gels stained for enzyme activity

8% non-denaturing polyacrylamide gels were run at constant current of 8 mA at 4 °C, overnight. The gels were stained for BChE activity with 1 mM butyrylthiocholine iodide as substrate in the Karnovsky and Roots staining procedure.143

5.3.6 Enzyme activity assays

To measure (−)-cocaine and benzoic acid, the product of (−)-cocaine hydrolysis catalyzed by BChE, sensitive radiometric assays were used based on toluene extraction of [3H](−)-cocaine labeled on its benzene ring.130 In brief, to initiate the enzymatic reaction, 100 nCi of [3H](−)-cocaine was mixed with culture medium. The enzymatic reactions proceeded at room temperature (25°C) with varying concentrations of (−)-cocaine. The reactions were stopped by adding 200 μl of 0.1 M HCl, which neutralized the liberated benzoic acid whereas ensuring a positive charge on the residual (−)-cocaine. [3H]Benzoic acid was extracted by 1 ml of toluene and measured by scintillation counting. Finally, the measured (−)-cocaine concentration-dependent radiometric data were analyzed by using the standard Michaelis-Menten kinetics so that the catalytic parameters were determined along with the use of a well-established standard enzyme-linked immunosorbent assay protocol.104
5.4 Methods for \textit{in vivo} studies

Sprague-Dawley (Male or female) rats (200-250 g) were obtained from Harlan (Indianapolis, IN) and were housed in groups of 2 to 4 per cage. All rats were allowed ad libitum access to food and water, and were maintained on a 12-h light-dark cycle with lights on at 8 AM in a room kept at a temperature of 21-22°C. Each rat was used only once. Experiments were performed in the same colony room in accordance with the Guide for the Care and Use of Laboratory Animals as adopted and promulgated by the National Institutes of Health. The experimental protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Kentucky.

Purified BChE mutants (0.15 mg/kg) were injected intravenously (\textit{i.v.}, via tail vein) into rats. After the enzyme injection, around 50-75 μl blood was collected into a capillary tube from the saphenous veins at different time points (2 min, 15 min, 30 min, 1 hr, 2 hr, 3 hr, 5 hr, 8 hr, 12 hr, 24 hr, 48 hr, and 72 hr). Serum isolated by centrifugation (5000 g, 15 min) from the blood sample will be used for analysis using the \textit{in vitro} enzyme activity assays mentioned above. The time-dependent concentrations of the enzyme in plasma were fitted to a double-exponential equation described by Kronman.\textsuperscript{144}

5.5 BChE mutant design: Insights from molecular modeling

The goal of the study described in this chapter was to design new BChE mutants that had a prolonged biological half-life without a significant change in the high catalytic activity of E12-7 against (-)-cocaine. The new computational design strategy relies on the MD-simulated dynamic structure of the BChE dimer and the idea that the BChE monomers can be covalently bonded to form covalent dimers by introducing intermolecular disulfide bond(s). So, we have carried out a long MD simulation (50 ns) on the BChE dimer structure in order to obtain a dynamically stable BChE dimer structure. To search for appropriate mutational sites to introduce the cross-subunit disulfide bond(s), a self-developed script was used to scan the key internuclear distances between the \textit{Ca} atoms of the residues on the dimer interface from the snapshots extracted from MD trajectory. Essentially, each pair of residues from different subunits was evaluated computationally for the simulated \textit{Ca–Ca} distance. If the simulated \textit{Ca–Ca}
distance was within 7 Å, the pair of the residues would be checked manually for further evaluation of the detailed interactions. The most hopeful pairs of residues may be mutated to cystines for introducing possible cross-subunit disulfide bond(s).

Table 5.1 The Cα-Cα distances between key residues in various dimer/tetramer BChE structures obtained from computational modeling/simulation.

<table>
<thead>
<tr>
<th>BChE structure</th>
<th>Cα-Cα distance (Å)</th>
<th>F364a</th>
<th>F364b</th>
<th>V377a</th>
<th>V377b</th>
<th>N535a</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>M532b</td>
<td>M532a</td>
<td>A516b</td>
<td>A516a</td>
<td>N535b</td>
</tr>
<tr>
<td>BChE model A&lt;sup&gt;a&lt;/sup&gt;</td>
<td>9.02</td>
<td>9.31</td>
<td>5.83</td>
<td>6.20</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>BChE model B&lt;sup&gt;b&lt;/sup&gt;</td>
<td>9.98</td>
<td>10.29</td>
<td>8.15</td>
<td>8.82</td>
<td>9.07</td>
<td></td>
</tr>
<tr>
<td>MD-simulated BChE structure (50 ns MD)&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Maximum</td>
<td>10.71</td>
<td>8.08</td>
<td>15.70</td>
<td>10.47</td>
<td>10.84</td>
</tr>
<tr>
<td></td>
<td>Minimum</td>
<td>5.33</td>
<td>4.42</td>
<td>4.71</td>
<td>5.56</td>
<td>4.53</td>
</tr>
<tr>
<td></td>
<td>Average</td>
<td>7.93</td>
<td>5.61</td>
<td>7.73</td>
<td>6.89</td>
<td>6.83</td>
</tr>
</tbody>
</table>

<sup>a</sup>A simple homology model of BChE tetramer obtained from the homology modeling using the X-ray crystal structure of Electrophorus electricus AChE (EeAChE) tetramer (PDB ID: 1C2O)<sup>94</sup> as a template. Because the simple homology modeling was based on the sequence alignment only (without any energy minimization or MD simulation on the backbone), the Cα-Cα distances between F364 and M532 in the BChE tetramer are simply the corresponding Cα-Cα distances between L373 and A542 in the EeAChE tetramer, and the Cα-Cα distances between V377 and A516 in the BChE tetramer are simply the corresponding Cα-Cα distances between L386 and A526 in the EeAChE tetramer. The sequence alignment indicates that N535 in BChE corresponds to T545 in EeAChE which is not available missing in the X-ray crystal structure.

<sup>b</sup>BChE tetramer model obtained from the combined use of the X-ray crystal structures of the EeAChE tetramer (PDB ID: 1C2O) and the BChE monomer (PDB ID: 1P0M)<sup>92</sup>; each subunit of the EeAChE tetramer was superimposed with the BChE monomer<sup>139</sup> in order to obtain the BChE tetramer (with the backbone frozen during the energy minimization and MD simulation).
Fully relaxed MD simulation of the BChE dimer structure starting from the BChE model B. The maximum, minimum, and average values of the Cα-Cα distances refer to the stable MD trajectory (10 to 50 ns).

Figure 5.1 The time-dependent Cα–Cα distances between key residues from different subunits (a and b) in the MD-simulated BChE dimer structure: (A) F364 and M532; (B) V377 and A516; and (C) N535.

As seen in Table 5.1, the detailed analysis of the MD trajectory predicted that mutations F364C/M532C, V377C/A516C, and N535C may introduce the desirable cross-subunit disulfide bond(s). Depicted in Figure 5.1 are the simulated time-dependent Cα–Cα distances for these pairs of the residues. Depicted in Figure 5.2 (A) is the MD-simulated BChE dimer structure consisting of subunits a and b. Figure 5.2 (B)–(G) shows the details of the MD-simulated BChE dimer structure concerning these important pairs of residues. The interface is mainly composed of 4 α helices (A516a – M532a, E363a – Y373a, A516b – M532b, and E363a – Y373b) and several loop parts. F364 is located on one α helix (E363a – Y373a), M532 is located on the end of one α helix (A516b – M532b) as shown in Figure 5.2 (B) and (C). V377 is located on the loop part of the interface. The location of A516 is similar to that of M532. Both A516 and M532 are located on the same α-helix (A516 – M532), with A516 being on the head and M532 being on the end of the α-helix.

According to the locations of these residues, it is possible to introduce a pair of cross-subunit disulfide bonds (C364a-C532b and C532a-C364b) through the F364C/M532C mutations on E12-7. The computationally designed new mutant, i.e. the
A199S/F227A/S287G/A328W/Y332G/F364C/M532C mutant (denoted as E364-532 for convenience), may have a pair of cross-subunit disulfide bonds: one between C364 of subunit a (C364a) and C532 of subunit b (C532b), and the other between C532 of subunit a (C532a) and C364 of subunit b (C364b). Similarly, the V377C/A516C mutations may also introduce a pair of cross-subunit disulfide bonds (C377a-C516b and C516a-C377b), and the designed new mutant, i.e. the A199S/F227A/S287G/A328W/Y332G/V377C/A516C mutant, is denoted as E377-516 for convenience. In addition, we noted in the simulated E12-7 structure that N535 was right in the middle of the interface, so that N535a and N535b from the two subunits are close to each other. For this reason, a single mutation on N535 (i.e. the N535C mutation) may introduce a single cross-subunit disulfide bond (C535a-C535b). The designed new mutant, i.e. the A199S/F227A/S287G/A328W/Y332G/N535C mutant, is denoted as E535 for convenience. It is also interesting to note that residues V377, M532, and N535 are either on the loop or next to the loop, implying that the positions of these residues have some flexibility which may help to form the desirable cross-subunit disulfide bonds.

As noted above, the computational prediction was based on the MD-simulated dynamically stable BChE dimer structure. For comparison, we also examined two static models of the BChE tetramer structures. One was a simple model of BChE tetramer (BChE model A in Table 5.1) obtained from the homology modeling using the X-ray crystal structure of Electrophorus electricus AChE (EeAChE) tetramer (PDB ID: 1C2O) as a template. The other one was a BChE tetramer model (BChE model B in Table 5.1) obtained from the combined use of the X-ray crystal structures of the EeAChE tetramer (PDB ID: 1C2O) and the BChE monomer (PDB ID: 1P0M). As seen in Table 5.1, based on BChE model A, we could only predict the possible cross-subunit disulfide bonds that can be introduced by the V377C/A516C mutations. As seen in Table 5.1, based on BChE model B, we could not predict any of the aforementioned cross-subunit disulfide bonds.
Figure 5.2 (A) The modeled BChE dimer structure (subunits a and b). (B) Key residues F364 (a/b) and M532 (a/b) on the dimer interface. (C) Cross-subunit disulfide bonds formed on the interface of E364-532 by C364a-C532b and C532a-C364b. (D) Key residues V377 (a/b) and A516 (a/b) in the dimer interface. (E) Cross-subunit disulfide bonds formed on the interface of E377-516 by C377a-C516b and C516a-C377b. (F) Key residues N535 (a/b) on the dimer interface. (G) Cross-subunit disulfide bond formed on the interface of E535 by C535a-C535b. Different subunits are shown in different colors: subunit a in green and subunit b in orange. The active-site residues are shown in pink. Key residues are colored in blue and red using the ball-and-stick model. E364-532 refers to A199S/F227A/S287G/A328W/Y332G/F364C/M532C BChE.

5.6 Oligomeric forms of the proteins

Based on the computational insights, we carried out in vitro experimental tests, including site-directed mutagenesis, protein expression and purification, on the E12-7, E364-532, E377-516, and E535 under the same experimental conditions. Depicted in Figure 5.3 are nondenaturing gels (8%) stained for the BChE activity. As shown in Figure
5.3, E12-7 exists in a mixture of the monomer (~70%), dimer (~10%), and tetramer (~20%), whereas E364-532, E377-516, and E535 only exist in the dimer (~80%) and tetramer (~20%) without the monomer at all. The data shown in Figure 5.3 indicate that the desirable cross-subunit disulfide bonds were likely formed in the designed enzymes E364-532, E377-516, and E535, and that formation of the cross-subunit disulfide bond(s) can help to completely eliminate the monomer and increase the dimer concentration.

![Figure 5.3 Nondenaturing gels (8%) stained for the BChE activity of (A) E12-7, (B) E364-532, (C) E377-516, and (D) E535.](image)

Figure 5.3 Nondenaturing gels (8%) stained for the BChE activity of (A) E12-7, (B) E364-532, (C) E377-516, and (D) E535. The gels were run with the constant current of 8 mA at 4°C overnight. The gels were stained for the BChE activity with butyrylthiocholine iodide as substrate at room temperature until the protein bands with the enzymatic activity were clearly identified. E364-532 refers to A199S/F227A/S287G/A328W/Y332G/F364C/M532C BChE, E377-516 refers to A199S/F227A/S287G/A328W/Y332G/V377C/A516C BChE, and E535 refers to A199S/F227A/S287G/A328W/Y332G/N535C BChE.

5.7 Catalytic activity against (−)-cocaine

Knowing the effects of the computationally designed mutations on the distribution of the oligomeric forms of the enzyme, we further carried out the in vitro enzyme activity assays to determine the effects of the same mutations on the catalytic activity of the enzyme against (−)-cocaine. To minimize the possible systematic experimental errors of
In vitro kinetic analysis, we expressed the enzymes and performed kinetic studies on E12-7 and the new mutants (E364-532, E377-516 and E535) at the same time under the same experimental conditions, and compared the catalytic activity of the new mutants to that of E12-7 against (-)-cocaine. Depicted in Figure 5.4 are the kinetic data, and summarized in Table 5.2 are the determined kinetic parameters. As summarized in Table 5.2, the catalytic parameters of the computationally designed new BChE mutants are very close to or essentially the same as corresponding catalytic parameters of E12-7 against (-)-cocaine. In terms of the $K_M$ values, only E377-516 has a slightly larger $K_M$ value (3.8 µM) compared to that (3.1 µM) of E12-7. Concerning the $k_{cat}$ values, E535 and E12-7 have a very similar $k_{cat}$ value (5700 min$^{-1}$ for E12-7 and 5770 min$^{-1}$ for E535) with a negligible difference (70 min$^{-1}$ which might be within the possible experimental errors). The $k_{cat}$ values of E364-532 and E377-516 are slightly lower than that of E12-7. So, we may say that the computationally designed mutations have remarkably changed the distribution of the oligomeric forms of the enzyme without significantly affecting the high catalytic activity of E12-7 against (-)-cocaine.
Figure 5.4 Kinetic data for the hydrolysis of (-)-cocaine catalyzed by BChE mutants: (A) E12-7; (B) E364-532; (C) E377-516; and (D) E535. The reaction rates were determined by using a sensitive radiometric assays based on toluene extraction of [3H](-)-cocaine labeled on its benzene ring. E364-532 refers to A199S/F227A/S287G/A328W/Y332G/F364C/M532C BChE, E377-516 refers to A199S/F227A/S287G/A328W/Y332G/V377C/A516C BChE, and E535 refers to A199S/F227A/S287G/A328W/Y332G/N535C BChE.
Table 5.2 Kinetic parameters determined for (-)-cocaine hydrolysis catalyzed by the BChE mutants.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_M$ (µM)</th>
<th>$k_{cat}$ (min$^{-1}$)</th>
<th>$k_{cat}/K_M$ (M$^{-1}$ min$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E12-7$^a$</td>
<td>3.1</td>
<td>5700</td>
<td>$1.8 \times 10^9$</td>
</tr>
<tr>
<td>E364-532</td>
<td>3.1</td>
<td>4650</td>
<td>$1.5 \times 10^9$</td>
</tr>
<tr>
<td>E377-516</td>
<td>3.8</td>
<td>4980</td>
<td>$1.3 \times 10^9$</td>
</tr>
<tr>
<td>E535</td>
<td>3.1</td>
<td>5770</td>
<td>$1.9 \times 10^9$</td>
</tr>
</tbody>
</table>


5.8 Biological half-life

Based on the encouraging in vitro data discussed above, we further tested E12-7, E364-532, E377-516, and E535 in vivo for their pharmacokinetic in rats. First of all, five rats (n=5) were tested for E12-7 to determine its biological half-life as the standard reference. For comparison, the other three enzymes (E364-532, E377-516, and E535) were also tested for their biological half-lives in rats. Depicted in Figure 5.5 were the time-dependent average concentrations of the enzymes in plasma after the i.v. injection of the enzymes. It had been well-known that measured time-dependent concentrations of the active enzyme $[E]$ in plasma follows a double exponential equation, i.e. 

$$[E] = A \exp(-\alpha_1 t) + B \exp(-\alpha_2 t)$$

which accounts for both the enzyme distribution process (the fast phase, associated with $\alpha_1$) and elimination process (the slow phase, associated with $\alpha_2$). The elimination half-life was also known as the biological half-life. Analysis of the time courses depicted in Figure 5.5 revealed the biological half-lives of the enzymes. The determined biological half-life of E12-7 was 7.3 hr in rats. With the cross-subunit disulfide bond(s), the computationally designed new mutants all had a significantly prolonged biological half-life which was 13.2, 13.1, and 13.3 hr for E364-532, E377-516, and E535, respectively. So, the computationally designed new enzymes (E364-532,
E377-516, and E535) indeed had a significantly prolonged biological half-life in rats without significantly changing the high catalytic activity of E12-7 against (-)-cocaine.

Figure 5.4 Time-dependent concentrations of the active enzyme in rat plasma after the i.v. injection of the enzyme at a dose of 0.15 mg/kg: (A) E12-7, (B) E364-532, (C) E377-516, and (D) E535. E12-7 refers to A199S/F227A/S287G/A328W/Y332G BChE, E364-532 refers to A199S/F227A/S287G/A328W/Y332G/F364C/M532C BChE, E377-516 refers to A199S/F227A/S287G/A328W/Y332G/V377C/A516C BChE, and E535 refers to A199S/F227A/S287G/A328W/Y332G/N535C BChE.

The in vivo data suggest that the new mutants, particularly E535, could be more valuable than E12-7 in the future further development of a cocaine abuse treatment. It has been known that, for a given therapeutically useful enzyme, the biological half-life in humans is usually much longer than that in rats. As noted above, the biological half-life of the native human BChE is ~24 hr in mice and ~7 to 12 days in humans. Assuming that the biological half-life of a BChE mutant in humans is at least 7-fold longer than that in
mice and rats, then we may theoretically predict the biological half-lives of the BChE mutants in humans from their biological half-lives in rats. Thus, E12-7 (with a biological half-life of ~7 hr in rats) is expected to have a biological half-life of at least two days in humans, and the newly designed enzymes (E364-532, E377-516, and E535, all with a biological half-life of ~13 hr in rats) are expected to have a biological half-life of about four days or longer in humans.

5.9 Conclusion

The MD simulation on the dimer structure of a promising cocaine hydrolase, \textit{i.e.} E12-7 engineered from human BChE, led us to predict that the F364C/M532C, V377C/A516C, and N535C mutations can produce cross-subunit disulfide bond(s) in the enzyme and, thus, stabilize the dimer structure. Based on the MD simulation, three new mutants of human BChE (E364-532, E377-516, and E535) were predicted to have a more stable dimer structure with the desirable cross-subunit disulfide bond(s) and, therefore, a different distribution of the oligomeric forms and a prolonged biological half-life. Following the computational prediction, \textit{in vitro} and \textit{in vivo} experimental studies have demonstrated that the computationally designed new BChE mutants, \textit{i.e.} E364-532, E377-516, and E535, indeed have a remarkably different distribution of the oligomeric forms (dimer and tetramer) without monomer at all and a significantly prolonged biological half-life in rats. The computationally designed new mutations (F364C/M532C, V377C/A516C, and N535C) have successfully extended the biological half-life of E12-7 in rats from ~7 hrs. to ~13 hrs., without significantly changing the high catalytic activity of E12-7 against (-)-cocaine. The newly designed and discovered BChE mutants, particularly E535, could be more valuable than E12-7 in the future development of a novel enzyme therapy for cocaine abuse. The encouraging outcomes of the present study suggest that the structure-and-mechanism-based computational design and integrated computational-experimental approach are promising for rational protein design, discovery, and development. The general computational protein design strategy and approach may also be valuable for engineering other proteins.

Copyright @ Lei Fang 2013
Chapter Six: Design, Preparation, and Characterization of Cocaine Esterase against Cocaine Toxicity with Improved Activity and Thermostability

Bacterial cocaine esterase (CocE) was known as the most efficient natural enzyme for metabolizing the naturally occurring (-)-cocaine, suggesting that it might be a promising candidate in treating cocaine overdose and addiction. However, a major obstacle of wild-type CocE is the thermostability and short biological half-life. The present study aimed to extend the thermostability and biological half-life of CocE by introducing cross-subunit disulfide bonds and PEGylation modification.

6.1 Overview of the CocE engineering

Bacterial cocaine esterase (CocE) was recognized as the most efficient natural enzyme for metabolizing the naturally occurring (-)-cocaine. Studies have shown that CocE can help prevent extreme cocaine toxicity and even from the lethal effects of cocaine. However, a major obstacle to the clinical application of CocE is the thermostability of wild-type CocE with a half-life of only a few minutes at physiological temperature (37°C). It is highly desirable to develop thermostable mutants of CocE for therapeutic treatment of cocaine toxicity.

To improve enzyme thermostability, two strategies are normally used with the purpose of improving the weak interactions inside the enzymes by either through non-covalent forces, such as hydrogen bond, or through covalent linkage, such as disulfide bond. Besides improving the stability of the enzyme, it is also important to maintain the activity of the enzyme. Based on a commonly accepted “stability-function theory”, there is a balance between protein stability and protein function, especially for enzyme. Enzymes fold from amino acid sequences to their tertiary structures to form the hydrophobic interiors and hydrophilic exteriors, in other words, from unstable amino acid sequences to stable structures. Structure flexibility of enzyme is important for facilitating its binding and reaction with a substrate. Thus, improving enzyme stability may possibly lead to decrease/loss of the enzyme activity. So, it is challenging to improve the
thermostability of an enzyme without decreasing its catalytic activity,\textsuperscript{147,148} and it is extremely rare to simultaneously improve both the thermostability and catalytic activity.\textsuperscript{149} However, one of the CocE mutants, \textit{i.e.} the T172R/G173Q CocE, designed in our lab has a ~30-fold improved thermostability (in terms of the \textit{in vitro} half-life at 37°C) compared to the wild-type CocE and maintains its catalytic activity against cocaine,\textsuperscript{62} giving us the hope to design new CocE mutants with improved thermostability without interfering with its activity against cocaine.

T172R/G173Q CocE (denoted as E172-173 for convenience) was designed through introducing favorable non-covalent force including a hydrogen bond between domains I and II.\textsuperscript{62} In this study, we aimed to further extend the thermostability of E172-173 through covalent bonding, \textit{i.e.} by introducing disulfide bonds. Our design strategy relies on the MD-simulated structure of the CocE dimer and the idea that the dimer structure can be stabilized by introducing disulfide bonds between the two subunits of the dimer.\textsuperscript{54} The computational design was followed by \textit{in vitro} experimental studies, leading to the discovery of a new CocE mutant that has a significantly improved thermostability and an improved catalytic efficiency against cocaine as well. The new mutant was modified further \textit{via} PEGylation. The PEGylated CocE mutant was used to fully protect mice from a lethal dose of cocaine (180 mg/kg, LD\textsubscript{100}) for at least three days, indicating that it might be a more promising candidate for development of novel anti-cocaine therapeutics. In addition, the integrated computational-experimental studies may also be valuable for other protein drug design, discovery, and development efforts that aim to treat other drugs of abuse or metabolic diseases.

\textbf{6.2 Computational methods used for the mutant design}

The starting CocE dimer (E172-173) structure was taken from the X-ray crystal structure (deposited at Protein Data Bank) at 2.0 Å resolution (PDB ID: 3I2F\textsuperscript{54}). The general procedure for carrying out the molecular dynamic (MD) simulations was similar to that used in section 2.2 of chapter 2. Briefly, all molecular mechanics optimization and MD simulations were carried out by using the AMBER 9 program package. A 50-ns MD simulation was performed under the normally adopted temperature (300 K).
6.3 Methods for *in vitro* studies

6.3.1 Site-directed mutagenesis

Similar to the procedure described in Chapter 5, for the study described in this chapter, point mutations were also generated using QuikChange method. Further mutations required to produce a new CocE mutant cDNA were generated from the cDNA corresponding to the E172-173 in the pET-22b (+) bacterial expression vector. All mutants were sequenced in both directions over the entire coding region. Using plasmid DNA as template and primers with specific base-pair alterations, mutations were made by polymerase chain reaction with Pfu DNA polymerase for replication fidelity. The PCR product was treated with DpnI endonuclease to digest the parental DNA template. The digested product was transformed into *Escherichia coli*, amplified, and purified. The DNA sequences of the mutants were confirmed by DNA sequencing.

6.3.2 Protein expression and purification

The CocE mutants were expressed in *Escherichia coli* BL-21 (DE3) cells grown at 37 °C. Protein expression was induced with 1 mM isopropyl-β-thiogalactopyranoside (Sigma Aldrich) for ~15 h at 18°C. Cells were pelleted, resuspended in 50 mM Tris-HCl, pH 8.0, 150 mM NaCl buffer with protease inhibitor cocktail (Sigma) and lysed using a French press (Thermo Fisher Scientific). The 6His-tagged enzymes were then enriched using HisPur cobalt resin (Thermo Fisher Scientific). The eluted fractions were concentrated by using an Amicon Ultra-50K centrifuge (Millipore, Billerica, MA). The enzyme concentrations were determined using a CB-Protein Assay kit (from CALBIOCHEM) with bovine serum albumin as a standard.

6.3.3 Enzyme activity assays

The procedure of enzyme activity assays is the same as that described in Chapter 5. To test the thermostability of the mutants, enzymes were diluted to 200 μg/ml, stored in sealed glass tubes and incubated at 37°C. One tube will be taken out of the incubation cabinet at various time points (0, 2, 9, 12, 31, and 100 days) and assayed for activity.
against (-)-cocaine as mentioned above. The percentage of remaining activity was plotted against time to assess the half-life of each enzyme.

6.3.4 Melting temperature measurements

Circular Dichroism (CD) data were collected on a J-810 spectropolarimeter (JASCO, Easton, MD). UV wavelength of 280 nm was used to run with temperature increasing from 10°C to 60°C in 1°C increments, and a 1-cm pathlength water-jacketed cylindrical quartz cell.

6.3.5 PEGylation

Purified enzyme was conjugated with maleimide-linked branched PEG with molecular weight of 40 KDa (JenKem Technology) overnight in PBS buffer, pH 7.4 at the PEG to enzyme ratio of 20.

6.4 Methods for in vivo studies

Similar to the description of rats in Chapter 5, for the study described in this chapter, male CD-1 mice (25–30 g) were purchased from Harlan (Indianapolis, IN) and were housed in groups of four mice per cage. All mice were allowed ad libitum access to food and water and were maintained on a 12 h light–dark cycle with lights on at 6:30 a.m. in a room kept at a temperature of 21–22°C. Experiments were performed in accordance with the Guide for the Care and Use of Laboratory Animals as adopted and promulgated by the National Institutes of Health. The experimental protocols were approved by the Institutional Animal Care and Use Committee (IACUC) at the University of Kentucky.

The purified enzyme was administrated intravenously (i.v., via tail vein) and (-)-Cocaine HCl (obtained from National Institute on Drug Abuse, Bethesda, MD) was administered intraperitoneally, at a volume of ~0.2 mL/mouse. Cocaine-induced toxicity was characterized by the occurrence of lethality. Lethality was defined as cessation of observed movement and respiration. A single dose (30 mg/kg) of E196-301 or PEGylated E196-301 was administered intravenously 1 min before the first intraperitoneal administration of 180 mg/kg cocaine (n=5). The mice were challenged daily with 180
mg/kg cocaine until no mouse survived. Following cocaine administration, mice were immediately placed individually for observation. The presence or absence of lethality was recorded for 60 min following cocaine administration.

6.5 CocE mutant design: Insights from molecular modeling

The goal of the study described in this chapter was to design CocE mutants that have an improved thermostability. Our new computational design strategy relied on the MD-simulated structure of the CocE dimer and the idea that the dimer structure can be stabilized by introducing disulfide bonds between the two subunits of the dimer. So, we carried out a long MD simulation (50 ns) on the CocE dimer structure in order to obtain a dynamically stable dimer structure. To search for appropriate mutational sites to introduce the cross-subunit disulfide bonds, a self-developed script was used to scan the key internuclear distances between the Cα atoms of the residues on the dimer interface from the snapshots extracted from MD trajectory as described in chapter 5.

Figure 6.1 The time-dependent Cα–Cα distances between L196 and I301 during the whole MD simulation in CocE dimer. Here a and b indicated after the residue numbers refer to subunits a and b, respectively.
Table 6.1 The Cα-Cα distances between key residues in CocE dimer obtained from computational simulation.

<table>
<thead>
<tr>
<th>CocE structure</th>
<th>Cα–Cα distances (Å)</th>
<th>L196a – I301b</th>
<th>I301a – L196b</th>
</tr>
</thead>
<tbody>
<tr>
<td>CocE crystal structure&lt;sup&gt;a&lt;/sup&gt;</td>
<td>7.29</td>
<td>7.20</td>
<td></td>
</tr>
<tr>
<td>MD-simulated CoCE structure (50 ns MD)&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Maximum 10.85</td>
<td>7.19</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Minimum 5.16</td>
<td>4.51</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Average 7.15</td>
<td>5.44</td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>T172R/G173Q CocE crystal structure, PDB ID: 3I2F.

<sup>b</sup>Fully relaxed MD simulation of the CocE dimer structure starting from the T172R/G173Q CocE crystal structure. The maximum, minimum, and average values of the Cα-Cα distances refer to the stable MD trajectory (10 to 50 ns).

As seen in Table 6.1, the detailed analysis of the MD trajectory predicted that extra mutations L196C/I301C may introduce the desirable cross-subunit disulfide bonds between the subunits of E172-173. Depicted in Figure 6.1 are the simulated time-dependent Cα–Cα distances for the pair of the residue. Depicted in Figure 6.2(A) is the MD-simulated CocE dimer structure consisting of subunits a and b. Figure 6.2(B) is the detail of the MD-simulated CocE dimer structure concerning this important pair of residues. The interface between the two subunits is mainly composed of some residues from all the domains (I, II, and III) in the form of α-helixs, β-sheets, and loops. L196 is located on one α-helix (E184-N197) in domain II, I301 is located on one loop in domain I.
According to the locations of these residues, it is highly possible to introduce two pairs of cross-subunit disulfide bonds (C196a-C301b and C301a-C196b) through the L196C/I301C/ mutations on E172-173. The computationally designed new mutant, i.e. the T172R/G173Q/L196C/I301C mutant (denoted as E196-301 for convenience), may have a pair of cross-subunit disulfide bonds: one between C196 of subunit a (C196a) and C301 of subunit b (C301b), and the other between C301 of subunit a (C301a) and C196 of subunit b (C196b). It is also interesting to note that residue I301 is on the loop, implying that the loop flexibility may help to form the desirable cross-subunit disulfide bonds.

Also as noted above, our computational prediction was based on the MD-simulated dynamically stable CocE dimer structure. For comparison, we also examined the static crystal structure of the CocE dimer. As seen in Table 6.1, based on the crystal structure, we could not predict any of the aforementioned cross-subunit disulfide bonds.

6.6 Thermostability assessment

Based on the computational insights, we carried out wet experimental tests, including site-directed mutagenesis, protein expression, purification, and enzyme activity
assays on the E172-173 and E196-301. To minimize the possible systematic experimental errors of the kinetic data, we expressed the enzymes and performed the kinetic studies on E172-173 and E196-301 under the same experimental conditions and compared their catalytic efficiency against (-)-cocaine. Michaelis-Menten kinetics of the enzymatic hydrolysis of (-)-cocaine was determined by performing the sensitive radiometric assays using [³H](-)-cocaine (labeled on its benzene ring) with varying concentrations of the substrate. Depicted in Figure 6.3 are the measured kinetic data, and summarized in Table 6.2 are the determined kinetic parameters.

Figure 6.3 Plots of measured reaction rates (with error bars) versus the substrate concentration for (-)-cocaine hydrolysis catalyzed by CocE mutants. (A) E196-301, 25°C, (B) E196-301, 37°C, (C) E172-173, 25°C, (D) E172-173, 37°C. E172-173 refers to T172R/G173Q CocE mutation, and E196-301 refers to T172R/G173Q/L196C/I301C CocE mutation.
Table 6.2 Kinetic parameters determined for (−)-cocaine hydrolysis catalyzed by CocE mutants.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>K_M (µM)</th>
<th>k_cat (min⁻¹)</th>
<th>k_cat/K_M (M⁻¹min⁻¹)</th>
<th>T (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>E196-301</td>
<td>2.4</td>
<td>1950</td>
<td>8.2×10⁸</td>
<td>25</td>
</tr>
<tr>
<td>E172-173</td>
<td>4.5</td>
<td>1400</td>
<td>3.13×10⁸</td>
<td></td>
</tr>
<tr>
<td>E196-301</td>
<td>1.5</td>
<td>3450</td>
<td>2.30×10⁷</td>
<td>37</td>
</tr>
<tr>
<td>E172-173</td>
<td>2.9</td>
<td>2600</td>
<td>9.15×10⁻⁸</td>
<td></td>
</tr>
</tbody>
</table>

*E172-173 refers to T172R/G173Q CocE mutation, and E196-301 refers to T172R/G173Q/L196C/I301C CocE mutation.

After determination of the kinetic parameters of E196-301, the purified protein was assessed for thermostability. As seen in Figure 6.4, the enzyme showed a relatively fast decrease of the activity (around 20%) in the first few days. Though the disulfide bonds were expected to form spontaneously (no extra oxidation reagent was employed here to facilitate the formation), a small percent of the new mutant (E196-301) molecules did not form the desirable cross-subunit disulfide bonds. Those E196-301 molecules without the desirable cross-subunit disulfide bonds could lose the activity more rapidly. The *in vitro* half-life of the new mutant without cross-subunit disulfide bonds is expected to the same as that of E172-173 which is ~6 hr (at 37°C).⁶² However, starting from the ninth day, the further enzyme activity in the subsequent 90 days was only about 15%, demonstrating its high thermostability. Overall, the enzyme still retained more than 60% of enzymatic activity after incubation at 37°C for 100 days, indicating that its *in vitro* half-life at 37°C is longer than 100 days.
Figure 6.4 Thermostability profile of E196-301. Percentage of the remaining enzyme activity was plotted against days (0, 2, 9, 12, 31, and 100 days) of incubation at 37 °C. E196-301 refers to T172R/G173Q/L196C/I301C CocE.

The circular dichroism measurements of thermal denaturation were performed to detect the apparent denaturation temperature ($T_m$) of E172-173 and E196-301. According to the data in Figure 6.5, the $T_m$ value of E172-173 and E196-301 are ~47°C and ~49°C, thus E196-301 has a ~2°C elevated $T_m$ compared with E172-173.
Figure 6.5 Date from circular dichroism measurements of E172-173 and E196-301. E172-173 refers to the T172R/G173Q CocE, and E196-301 refers to the T172R/G173Q/L196C/I301C CocE.

6.7 Confirmation of new cross-subunit disulfide bond(s) between the subunits

The above experimental results suggest the formation of the designed cross-subunit disulfide bonds in the CocE dimer. Further, in collaboration with Dr. David Rodgers, we tried to obtain the crystal structure of E196-301 that can be used to directly confirm the formation of the cross-subunit disulfide bonds between L196 and I301 (the detailed experimental procedure will be described elsewhere). The obtained X-ray crystal structure of E196-301 is depicted in Figure 6.6 in comparison with that of E172-173.
Figure 6.6 shows the alignment between E172-173 and E196-301 with a RMSD value of 0.299 Å, indicating the high similarity of the structures. However, the alignment result also shows a slight shift of two α-helices (H2 and H3 in domain II) in E196-301 compared with that in E172-173 as indicated by the arrows in Figure 6.6. The shift created a slightly enlarged active-site cavity in E196-301, which could possibly facilitate the catalytic reaction. A slightly larger active site could accommodate such a large substrate (cocaine) better and, thus, make the enzyme more active against cocaine; the similar mechanism was noted for the catalytic activity of certain BChE mutants against cocaine.\textsuperscript{105}
Figure 6.7 Time-dependence of important H–O distances (related to hydrogen bonds) from the MD-simulated E172-173 and E196-301 structures. Y44HH–CocO represents the distance between the hydroxyl hydrogen (HH) of the Y44 side chain and the oxygen atom (CocO) of carbonyl oxygen of (−)-cocaine benzoylester. Y118H–CocO refers to the distance between hydrogen (H) of the Y118 main chain and the oxygen atom (CocO) of carbonyl oxygen of (−)-cocaine benzoylester. E172-173 refers to T172R/G173Q CocE, and E196-301 refers to T172R/G173Q/L196C/I301C CocE.

Table 6.3 Summary of the MD-simulated key distances (in Å) between the hydrogen atoms of key residue and the carbonyl oxygen of (−)-cocaine benzoylester in the transition-state structures of CocE.

<table>
<thead>
<tr>
<th>Hydrogen bond</th>
<th>Distances</th>
<th>Maximum (Å)</th>
<th>Minimum (Å)</th>
<th>Average (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>E172-173a</td>
<td>3.02</td>
<td>1.44</td>
<td>1.87</td>
</tr>
<tr>
<td></td>
<td>E196-301a</td>
<td>3.17</td>
<td>1.44</td>
<td>1.80</td>
</tr>
<tr>
<td>Y44HH-CocO</td>
<td>E172-173a</td>
<td>3.14</td>
<td>1.61</td>
<td>2.16</td>
</tr>
<tr>
<td></td>
<td>E196-301a</td>
<td>2.89</td>
<td>1.61</td>
<td>2.13</td>
</tr>
</tbody>
</table>

aE172-173 refers to T172R/G173Q CocE, and E196-301 refers to T172R/G173Q/L196C/I301C CocE.

To further examine the possible mechanism concerning the improved catalytic activity, MD simulations were performed on the transition-state structures (TS1’s) of
E172-173 and E196-301. Depicted in Figure 6.7 are the simulated time-dependent H–O distances (related to the hydrogen bonds) in E172-173 and E196-301 during the MD simulations for 50 ns. The detailed analysis of the key H–O distances between enzyme residues and cocaine is summarized in Table 6.34. In E172-173, the H–O distance of CocO with Y44HH is 3.02 Å in maximum, 1.44 Å in minimum, and 1.87 Å in average, while the H–O distance with Y114 is 3.14 Å in maximum, 1.61 Å in minimum, and 2.16 Å in average. In E196-301, the H–O distance between Y44HH and CocO is 3.17 Å in maximum, 1.44 Å in minimum, and 1.80 Å in average, whereas the H–O distance between Y114H and CocO is 2.89 Å in maximum, 1.61 Å in minimum, and 2.13 Å in average. According to these H–O distances, the lengths of the two hydrogen bonds in E196-301 are shorter than the corresponding ones in E172-173, suggesting that cocaine has stronger hydrogen bonding with E196-301 compared to that with E172-173. The enhanced hydrogen bonding helps to stabilize the transition-state structure during the catalytic reaction and, thus, lower the energy barrier, which explains the improved catalytic activity of the new mutant.

6.8 in vivo protection against cocaine-induced lethality in mice

After confirming the high thermostability of the E196-301, the PEGylation modification was conducted to further engineer E196-301. According to the activity assays, the PEGylated E196-301 maintained its activity against cocaine as that of E196-301. To test the ability of E196-301 and PEGylated E196-301 in protecting mice from a lethal dose of cocaine, E196-301 or PEGylated E16-301 was administered intravenously (at a single dose of 30 mg/kg) 1 min before the first intraperitoneal administration of 180 mg/kg cocaine (n=5). Depicted in Figure 6.8 are the data for the in vivo protection of unPEGylated E196-301 and PEGylated E196-301 against cocaine-induced lethality in mice. As seen in Figure 6.8, unPEGylated E196-301 protected the mice from death after the first injection of 180 mg/kg cocaine, but lost efficacy at second injection of 180 mg/kg cocaine 24 hours later. PEGylated E196-301 performed much better. It fully protected all the mice (n=5) for at least 72 hr from the acute toxicity of a lethal dose of cocaine (180 mg/kg, LD_{100}): no mice died during the cocaine challenge at 72 hr, three
mice died during the cocaine challenge at 96 hr, and the remaining two mice died during the final cocaine challenge at 120 hr (see Figure 6.8). According to the data depicted in Figure 6.8, PEGylated E196-301 can protect the three mice (60%) with the protection time \( t_p \) being between 72 hr and 96 hr: 72 hr < \( t_p \) < 96 hr, or we have \( t_p = 84 \pm 12 \) hr. For the remaining two mice (40%), 96 hr < \( t_p \) < 120 hr, or we have \( t_p = 108 \pm 12 \) hr according to the data in Figure 6.8. Overall, PEGylated E196-301 can protect the mice with an average protection time \( \langle t_p \rangle \) of ~94 hr, *i.e.* we have \( \langle t_p \rangle = \sim 94 \) hr.

![Graph showing survival percentage vs. time (hour) for E196-301 and PEGylated E196-301.](image)

**Figure 6.8** *in vivo* protection of E196-301 (black squares) and PEGylated E196-301 (red triangles) against cocaine-induced lethality in mice. A single dose (30 mg/kg) of E196-301 or PEGylated E16-301 was administered intravenously 1 min before the first intraperitoneal administration of 180 mg/kg cocaine (n=5). The mice were challenged daily with 180 mg/kg cocaine *via* intraperitoneal administration until no mouse survived. E196-301 refers to T172R/G173Q/L196C/I301C CocE.

### 6.9 Conclusion

In the study described in this chapter, molecular dynamics simulation on the dimer structure of a promising cocaine–metabolizing enzyme, *i.e.* E172-173 engineered
from wild-type CocE, led us to predict that the extra L196C/I301C mutations on E172-173 can produce cross-subunit disulfide bonds in the enzyme and, thus, stabilize the dimer structure. Following the computational prediction, our \textit{in vitro} experimental studies have demonstrated that the computationally designed new CocE mutant (T172R/G173Q/L196C/I301C), \textit{i.e.} E196-301, indeed has significantly improved thermostability and catalytic activity against cocaine. \textit{in vivo} tests showed that the PEGylated E196-301 can fully protect mice from a lethal dose of cocaine (180 mg/kg, LD$_{100}$) for at least three days, whereas the unPEGylated E196-301 can only protect mice less than 24 Hours. In summary, the newly designed and discovered CocE mutant E196-301 could be more valuable than E172-173 in the future development of a novel enzyme therapy for cocaine abuse. The encouraging outcomes of the study described in this chapter suggest that the structure-and-mechanism-based computational design and integrated computational-experimental approach are promising for rational protein drug design, discovery, and development. The general computational protein design strategy and approach may also be valuable for engineering other proteins.
Chapter Seven: Concluding Remarks and Future Plan

Human butyrylcholinesterase (BChE) and bacterial cocaine esterase (CocE) have shown their potential in treating cocaine abuse. However, the wild-type enzymes themselves are not suitable for use as anti-cocaine therapeutics, due to the low catalytic activity, thermoinstability, and/or short biological half-life. In this investigation, we performed integrated computational-experimental studies to rationally design and discover mutants of these enzymes in order to improve the catalytic activity, thermostability, and/or biological half-life. To rationally design desirable mutants of the enzymes, we have successfully developed computational models, including those for the BChE gating, glycosylated BChE structure, BChE-substrate complex structure, BChE dimer/tetramer structures, CocE monomer/dimer structures, and CocE-substrate complex structures. Development of the computational models enabled us to rationally design new amino-acid mutations that may improve catalytic activity, thermostability, and/or biological half-life. The computational design was followed by wet experimental tests, including both in vitro and in vivo experiments, leading to encouraging outcomes.

7.1 Summary of the major conclusions obtained from this investigation

1) The active sites and gating mechanisms of AChE and BChE have been compared by performing MD simulations on the water-solvated systems of the monomers and tetramers of hBChE and AChE. It has been demonstrated that different gating mechanisms due to the conformational dynamics of gating residues of AChE and BChE are responsible for their difference in the substrate specificity. The MD simulations of the tetramer structures of AChE and BChE have revealed that although there are some structural differences, both enzymes have a common feature that they could have two dysfunctional active sites due to the special locations of the active-site gate.

2) According to the BChE-substrate complex models, the computationally designed mutations only considerably improve the catalytic efficiency of human BChE against (-)-cocaine, without significantly improving the catalytic activities against other
substrates (including acetylcholine, acetylthiocholine, butyrylthiocholine, and (+)-
cocaine).

3) A 3D structural model of the glycosylated human BChE has been developed to
examine the possible effects of the glycosylation on the possible PEGylation
modification reactions, revealing the direct and indirect influence of the glycosylation.

4) The computationally designed new mutations (F364C/M532C, V377C/A516C,
and N535C) have successfully extended the biological half-life of
A199S/F227A/S287G/A328W/Y332G BChE in rats from ~7 hr to ~13 hr, without
significantly changing the high catalytic activity of
A199S/F227A/S287G/A328W/Y332G BChE against (-)-cocaine.

5) The computationally designed new mutations (L196C/I301C) have
successfully extended the in vitro half-life of T172R/G173Q CocE from ~6 hrs. to at least
100 days, and significantly improved the catalytic activity against cocaine. In addition, it
has been demonstrated that the PEGylated T172R/G173Q/L196C/I301C CocE can be
used to fully protect mice from a lethal dose of cocaine (180 mg/kg, LD_{100}) for at least
three days.

7.2 Future plan concerning rational design of therapeutic enzymes for cocaine abuse

Future rational design of human BChE mutants should be focused on improving
the biological half-life while maintaining the high catalytic activity against (-)-cocaine
and selectivity for (-)-cocaine over neurotransmitter acetylcholine.

For CocE-based protein drug design, future rational design should be focused on
improving the biological half-life and reducing the immune response. In addition,
improving the catalytic efficiency of the thermostable CocE mutants against (-)-cocaine
may be another direction, and the substrate selectivity of the thermostable CocE mutants
for (-)-cocaine over neurotransmitter acetylcholine should also be accounted for in the
computational design.

Copyright © Lei Fang 2013
REFERENCE

(2) Stolberg, V. B. Journal of Ethnicity in Substance Abuse 2011, 10, 126.
(5) Gaedcke, F. Archiv der Pharmazie 1855, 132, 141.
(6) Niemann, A. Archiv der Pharmazie 1860, 153, 129.
(25) Preti, A. Addiction Biology 2007, 12, 133.
(26) Xi, Z.-X. In Addictions - From Pathophysiology to Treatment; Belin, D., Ed.; InTech: 2012, p 311.


VITA
Lei Fang

Education
Aug. 2008 – Present, Ph.D. candidate in Pharmaceutical Sciences, University of Kentucky, Lexington, KY.
Sep. 2005 – Jun. 2008, Master in Physical Chemistry (Computational Chemistry), Graduate University of Chinese Academy of Sciences (GUCAS), Beijing, China
Sep. 2000 – Jun. 2004, Bachelor in Polymer Science and Engineering, Minor Degree in Computer Science and Application, Sichuan University (SCU), Chengdu, China.

Publications


for publication.


**Meeting Abstracts**


Additional Information
2008. 9 – 2009.5, Teaching Assistant at College of Pharmacy, UK
2009. 6 – Present, Research Assistant at College of pharmacy, UK
2010.9 – 2012.5, Historian of Association of American Pharmaceutical Scientists Student Chapter of UK
2009.9 – Present, American Association of Pharmaceutical Scientists (AAPS) member