A Structural and Energetic Model for the Slow-onset Inhibition

A Dissertation Presented

by

Cheng-Tsung Lai

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

May 2014

Stony Brook University

The Graduate School

Cheng-Tsung Lai

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Carlos L. Simmerling – Dissertation Advisor
Professor, Department of Chemistry

Peter J. Tonge – Dissertation Advisor
Professor, Department of Chemistry

David F. Green – Chairperson of Defense
Associate Professor, Department of Applied Mathematics and Statistics

Robert C. Rizzo – Committee Member
Associate Professor, Department of Applied Mathematics and Statistics

Dale G. Drueckhammer – Outside Member
Professor, Department of Chemistry

This dissertation is accepted by the Graduate School

Charles Taber
Dean of the Graduate School

Abstract of the Dissertation

A Structural and Energetic Model for the Slow-onset Inhibition

by

Cheng-Tsung Lai

Doctor of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2014

Slow-onset inhibitors are of particular interest in drug discovery programs as the slow dissociation of the inhibitor from the target-inhibitor complex prolongs target occupancy and improves *in vivo* efficacy. While slow-onset inhibition is observed in many enzymes and the kinetic equations used to describe slow-onset inhibition were derived more than 26 years ago, the structural basis for slow-onset inhibition is still not generally well understood, hindering prediction and control of slow-onset binding kinetics. An enzyme system with known experimental kinetics and structural data for multiple classic and slow-onset inhibition complexes would be an ideal model system for the study of structure basis of slow-onset inhibition. InhA, the *Mycobacterium tuberculosis* enoyl-ACP reductase, is a validated target for the development of tuberculosis chemotherapeutics and an excellent model system for the study of slow-onset inhibition. Inhibition of InhA follows a two-step binding mechanism in which formation of the initial enzyme-inhibitor (EI) complex is followed by a slow conformational change that leads to the final enzyme-inhibitor complex (EI*). In this work, based on analyses of

all available crystal structures, we found that the active-site conformation of InhA can be characterized as open or closed. Unrestrained molecular dynamics simulation started from the open conformation demonstrated that the active-site helix-6&7 region was moving toward to a semi-closed conformation in both rapid reversible and slow-onset inhibitors bound complexes. On the contrary, when simulation started from the closed conformation, the active-site helix-6&7 region maintained in a relative stable conformation in the slow-onset inhibitor bound complexes. We hypothesized that the open and closed conformations represent the initial EI and final EI* complexes, respectively, and conformational change from open to closed is the slow structural isomerization step. By using partial nudged elastic band (PNEB) and umbrella sampling (US) simulations, we were able to obtain a continued energy landscape along the open-closed conformational change path. For the rapid reversible inhibition complexes, the energy landscapes exhibit two types of energy profile, either preferring the open state, or having little preference and low energy barrier between the open and closed states. On the other hand, the slow-onset inhibition complexes have a relative stable closed state with a more significant energy barrier between the open and closed states.

In order to modulate the life-time (residence time) of the enzyme-inhibitor complex, it is important to understand the interactions that modulate this induced-fit mechanism and, specifically, to determine the structure of the transition state that lies on the reaction coordinate between the open and closed states. Structural analyses identified several active-site residues that regulate the free energy barrier in the open-closed path. Replacement of these key residues with amino acids possessing smaller side chains results in a decrease of energy barrier. The energy barrier can be restored by rational reintroduction of a steric clash at the transition state. This was accomplished in two different ways: through mutation of a residue at a different position to a

larger side chain, or through modification of the ligand with a bulky substituent. These loss and regain of function studies validate these key residues controlling conformational change, and will provide a platform for the design of inhibitors with longer residence time and better *in vivo* potency.

Although the free energy profile derived from the PNEB/US approach provides a way to distinguish rapid reversible and slow-onset inhibitors, PNEB/US approach is computationally expensive and time-consuming. It is unlikely to use this approach to examine every inhibitor, thus, a rapid screening approach – using docking to rapidly examine inhibitors was designed. The dock score had same trend as the US results. The combined dock/PNEB/US protocol could provide an approach for slow-onset inhibitor screening.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Full form of the abbreviation
AM1-BCC	semi-empirical (AM1) with bond charge correction (BCC)
DSSP	Database of secondary structure assignments
E. coli	*Escherichia coli*
FAS	Fatty acid synthesis
FEP	Free energy perturbation
ftuFabI	*Francisella tularensis* FabI
GAFF	General AMBER force field
GPU	Graphics processing unit
HIV	Human immunodeficiency virus
HTS	High-throughput screening
INH	Isoniazid
LIE	linear interaction energy
MIC	Minimum inhibitory concentration
MD	Molecular dynamics
MDR-TB	Multidrug-resistant TB
MM	Molecular mechanics
MM/PBSA	Molecular mechanics/Poisson Bolzmann surface area
MM/GBSA	Molecular mechanics/Generalized Born surface area
MTB	*Mycobacterium tuberculosis*
NADH/NAD$^+$	Nicotinamide adenine dinucleotide

NADPH/NADP$^+$	Nicotinamide adenine dinucleotide phosphate
NEB	Nudged elastic band
PCA	Principal component analysis
PDB	Protein data bank
PMF	Potential of mean force
PNEB	Partial nudged elastic band
RESP	Restrained electrostatic potential
RMSD	Root mean square deviation
RMSF	Root mean square fluctuation
SASA	Solvent accessible surface area
TB	Tuberculosis
TI	Thermodynamic integration
TIP3P	Transferable intermolecular potential 3 points
US	Umbrella sampling
VDW	Van der Waals
WHAM	Weighted histogram analysis method
WHO	World Health Organization

Acknowledgments

It has been a very fulfilling six years studying for my Ph.D. in Stony Brook University under the mentoring of Dr. Carlos Simmerling and Dr. Peter Tonge. I started my Ph.D. journey in Tonge lab as a rotation student, followed by next two rotations in Rizzo lab and Simmerling lab. After the rotations, I joined both Simmerling lab and Tonge lab working on the collaborating InhA project. Dr. Peter Tonge has given me plenty of room and freedom to explore my research interests and supported me when I needed. Dr. Carlos Simmerling brought me into the world of molecular dynamics simulation and has been extremely patient in hearing my research results and ideas during our discussions, and provided me valuable guidance to pursue the hypothesis. Both Dr. Simmerling and Dr. Tonge have made great influence in my life, professionally and personally.

I am very fortunate to have Dr. David Green, Dr. Robert Rizzo, and Dr. Dale Drueckhammer on my thesis committee. They have provided valuable inputs into my research and have been very supportive for me. I would like to thank Dr. David Green, the chairperson of my dissertation committee, for his constructive criticism and valuable suggestions. I am grateful for Dr. Robert Rizzo, for taking me as a rotation student in his lab. I have gained lots of knowledge on computational chemistry from Dr. Rizzo's courses and during the rotation. I would especially thank Dr. Dale Drueckhammer, my outside member, for his kind support.

I would also like to thank all the past and present members of the Simmerling lab and the Tonge lab. Without their support, I wouldn't be able to complete my Ph.D. study. It is my great pleasure to know and to work with them. They have been with me through the ups and downs and I am thankful for their companies as well as their feedbacks. I would also like to thank

people from Green lab and Rizzo lab. I enjoyed our brainstorming discussions and thank you for feeding me new ideas from different viewpoints.

At the end, my greatest thank goes to my family, especially my parents, for their unconditionally love and sacrifice. I also want to thank my lovely girlfriend, Ming, who shares joy and sorrow with me along the road. She always brings positive thoughts when I am frustrated. Without their support, I wouldn't finish my Ph.D. study.

List of Publications

1. **Lai, C.-T.**, Li, H.-J, Garcia-Diaz, M., Tonge, P. J., Simmerling, C. "Rational modulation of the induced-fit conformational change for slow-onset inhibition in *M. tuberculosis* enoyl-ACP reductase InhA" (In preparation)

2. Li, H.-J.*, **Lai, C.-T.***, Pan, P., Yu, W., Liu, N., Garcia-Diaz, M., Simmerling, C., Tonge, P. J. (2014) "A structural and energetic model for the slow-onset inhibition of InhA, an enzyme drug target from Mycobacterium tuberculosis" ACS Chemical Biology 9, 986-993. (*equal contribution)

3. Pan, P., Knudson, S., Bommineni, G. R., Li, H.-J., **Lai, C.-T.**, Liu, N., Garcia-Diaz, M. Simmerling, C., Patil, S. S., Slayden, R. A., Tonge, P. J. (2014) "Time-dependent diaryl ether inhibitors of InhA, structure-activity relationship studies of enzyme inhibition, antibacterial activity, and in vivo efficacy" ChemMedChem 9, 776-791.

4. Brust, R., Haigney, A., Lukacs, A., Gil, A., Hossain, S., Addison, K., **Lai, C.-T.**, Towrie, M., Greetham, G., Clark, I., Illarionov, B., Bacher, A., Kim, R.-R., Fischer, M., Simmerling, C., Meech, S., Tonge, P. J. (2014) Ultrafast dynamics of blsA, a photoreceptor from the pathogenic bacterium Acinetobacter baumannii, Journal of Physical Chemistry Letters 5, 220-224.

5. Kapilashrami, K., Bommineni, G. R., Machutta, C. A., Kim, P., **Lai, C.-T.**, Simmerling, C., Picart, F., and Tonge, P. J. (2013) Thiolactomycin-based β-ketoacyl-AcpM synthase A (KasA) inhibitors: fragment-based inhibitor discovery using transient one-dimensional nuclear overhauser effect NMR spectroscopy, Journal of Biological Chemistry 288, 6045-6052.

6. Li, X., Liu, N., Zhang, H., Knudson, S. E., Li, H.-J., **Lai, C.-T.**, Simmerling, C., Slayden, R. A., and Tonge, P. J. (2011) CoA adducts of 4-oxo-4-phenylbut-2-enoates: inhibitors of MenB from the M. tuberculosis menaquinone biosynthesis pathway, ACS Medicinal Chemistry Letters 2, 818-823.

Chapter One

Introduction

1.1 Background of Tuberculosis

Tuberculosis (TB) is a contagious disease caused by the bacillus *Mycobacterium tuberculosis* (MTB) that mainly targets the lungs, but also attacks any part of the body, such as the brain, the kidneys, or the spine. TB spreads from person to person through the air. According to the World Health Organization (WHO) 2012 report, there were 8.7 million new infected cases and 1.4 million people died from TB around the world in 2011. Geographically, Asia is the most affected area, accounting for 59% of all cases. Other affected areas are Africa (26%), Eastern Mediterranean region (7.7%), the European region (4.3%), and the region of the Americas (3%) (*1*).

TB can be characterized as latent TB infection or TB disease. People become TB infected by breathing in TB bacteria. TB bacteria can live in the human body without causing illness because the human's immune system can fight the TB bacteria. In this situation, this is called latent TB infection. People with latent TB infection do not feel sick, do not have TB symptom, and will not spread TB bacteria to others. However, TB bacteria will become active if the infected person has a weak or compromised immune system, and in this situation it is called TB disease. The symptoms of TB disease include chest pain, coughing up blood, fatigue, low appetite, weight loss, chills, fever, and night sweating. Because the latent TB can become active TB in an immunocompromised body, a 20-fold increased rate of activating latent TB is found in the HIV infected people (*2, 3*).

Current treatment of TB consists of a six month regimen of four first-line drugs: **Isoniazid**, **Rifampicin**, **Ethambutol** and **Pyrazinamide** (**Figure 1.1**) (*4*). **Isoniazid** and **Rifampicin** are the two most powerful TB drugs targeting enoyl-ACP reductase and DNA-dependent RNA polymerase, respectively. For multidrug-resistant TB (MDR-TB), defined as resistance to both **Isoniazid** and **Rifampicin**, the treatment requires other more expensive and more toxic second-line drugs, and the treatment time is at least 24 months. WHO estimated the cost to stop TB between 2013 and 2015 will be $8 billion per year, including $5 billion for the drug-susceptible TB, $2 billion for the MDR-TB and $1 billion for the TB/HIV interventions (*1*). Therefore, a cheaper and more efficient way to treat TB is always needed.
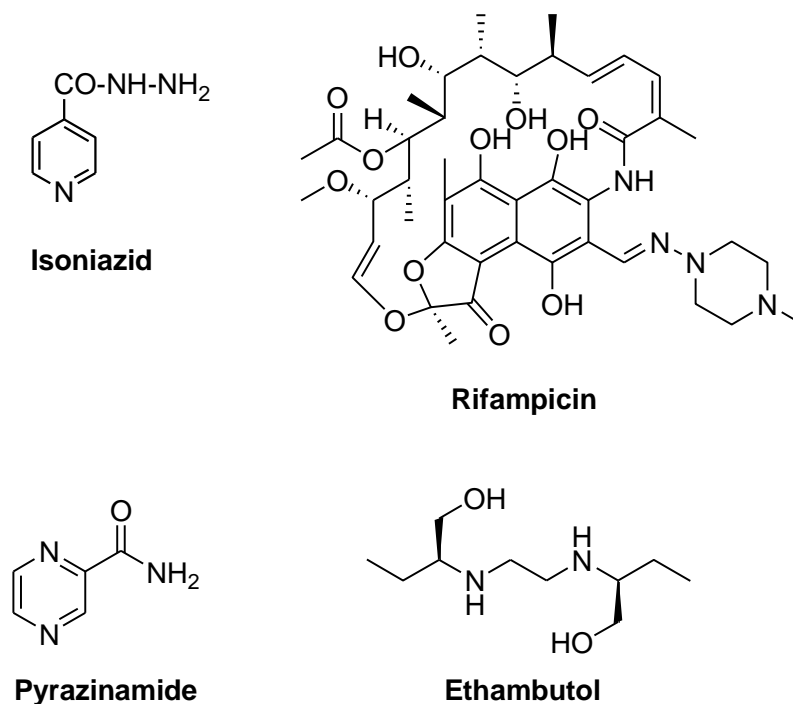
Figure 1.1. The first-line anti-TB drug. Isonizaid, **Rifampicin**, **Pyrazinamide**, and **Ethambutol**.

1.2 New TB drugs underdevelopment

There is an urgent need for new TB drugs. TB drug-resistant strains were being isolated clinically one year after the use of the first-line drugs **Isoniazid** and **Rifampicin** in early 1950s (*5*). Because the MDR-TB requires a 24-month treatment regimen and the second-line drugs are associated with significant side-effects, new drugs to shorten and simplify TB treatment have always been needed. **Figure 1.2** summarizes new TB drugs that are currently underdevelopment: several enzymes and leads have been identified, such as enoyl-ACP reductase (InhA), oxidoreductase (DprE), and DNA gyrase (GyrB). There are four drugs in phase III stage, including **Delamanid** (OPC-67683), **Gatifloxacin**, **Moxifloxacin** and **Rifapentine**. Among these drugs, **Delamanid** is known to target the mycolic acid biosynthesis but to what enzyme it targets remains unknown (*6*). **Gatifloxacin** and **Moxifloxacin** are known to target DNA gyrase (*7*). **Rifapentine**, similar to **Rifampicin**, targets RNA polymerase.

Lead optimization	Preclinical development	Phase I	Phase II	Phase III
DprE inhibitors	CPZEN-45	BTZ-043	AZD-5847	Delamanid
GyrB inhibitors	DC-159a	TBA-354	TMC-207	Gatifloxacin
InhA inhibitors	Q-201		Linezolid	Moxifloxacin
LeuRS inhibitors	SQ-609		PA-824	Rifapentine
MGyrX1 inhibitors	SQ-641		SQ-109	
Mycobacterial Gyrase inhibitors			PNU-100480	
Translocase-1 inhibitors				

Figure 1.2. The development pipeline of new TB drugs. (reproduced from WHO 2012 report (*1*).)

1.3 Enoyl-ACP reductase (InhA), the target of isoniazid

Isoniazid, the most effective first-line TB drug**,** destroys the TB cell wall integrity by blocking the biosynthesis of mycolic acid (*8*). Mycolic acid, a long chain fatty acid (C50~C60), is a common component of extra cellular matrix in gram-positive bacteria (**Figure 1.3**). Bacteria synthesize mycolic acid via the fatty acid synthesis II (FAS II) pathway, whereas humans use FAS I pathway to generate short chain fatty acid. Another significant difference between FAS I and II pathway is that FAS I system uses a multifunctional enzyme for all catalytic process, while multiple unifunctional enzymes are found in the FAS II system of bacteria. Hence the FAS II pathway is a promising target for developing anti-bacteria drugs. The elongation process of FAS II consists of four steps: condensation, reduction, dehydration and a second reduction (**Figure 1.4**). The second reduction step in this elongation process is catalyzed by enoyl-acyl carrier protein reductase, which is the target of **Isoniazid**. The enoyl-acyl carrier protein reductase is generally known as FabI (fatty acid biosynthesis protein I), and is named as InhA in MTB because it is coded by the inhA gene (*9*).
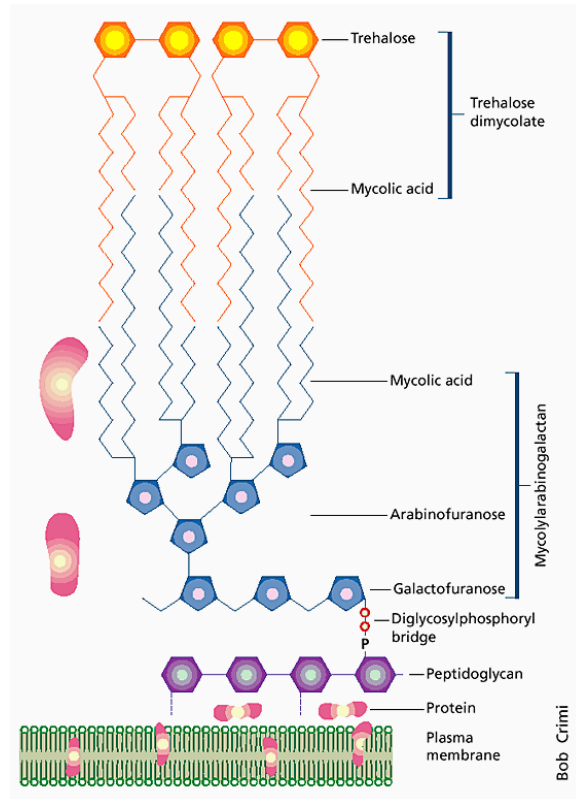
Figure 1.3. Schematic representation of the mycobacterial cell wall. Mycolic acid is an essential component of the mycobacterial cell wall. Figure taken from (*8*).
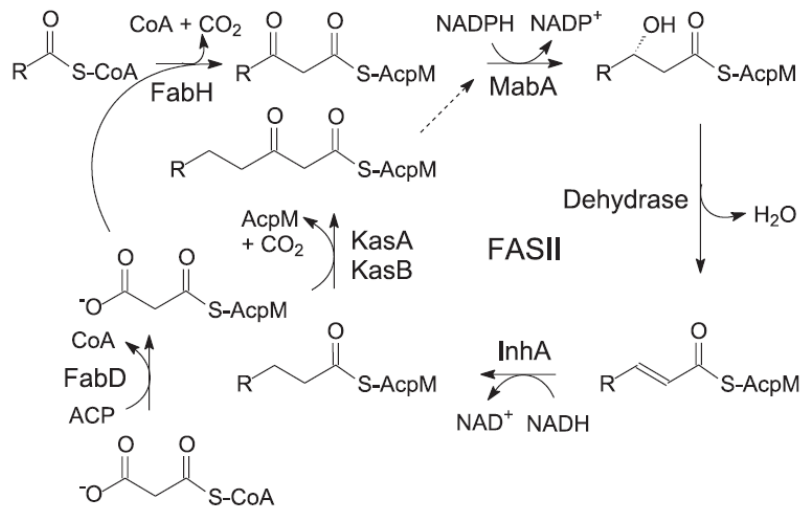
Figure 1.4. FAS II biosynthesis pathway. Long chain fatty acids are synthesized by discrete and unifunctional enzymes. The last step of each elongation cycle is catalyzed by enoyl-ACP reductase which converts enoyl-ACP to acyl-ACP.

Isoniazid is a pro-drug that requires activation by KatG, a catalase/peroxidase enzyme found in MTB. Upon activation, Isoniazid and the cofactor NAD(H) form a covalent INH-NAD adduct. The INH-NAD adduct inhibits InhA by competing with the NADH for the active-site driven by better binding affinity to InhA (**Figure 1.5**). **Figure 1.6** shows the superposition of INH-NAD adduct and the substrate analogue bound InhA, illustrating how the INH-NAD adduct blocks InhA enzymatic function: INH-NAD occupies the position of residue Y158, making residue Y158 move to the substrate binding pocket and therefore prevents substrate binding to the active-site.

Isoniazid   **NAD⁺**   **INH-NAD adduct**
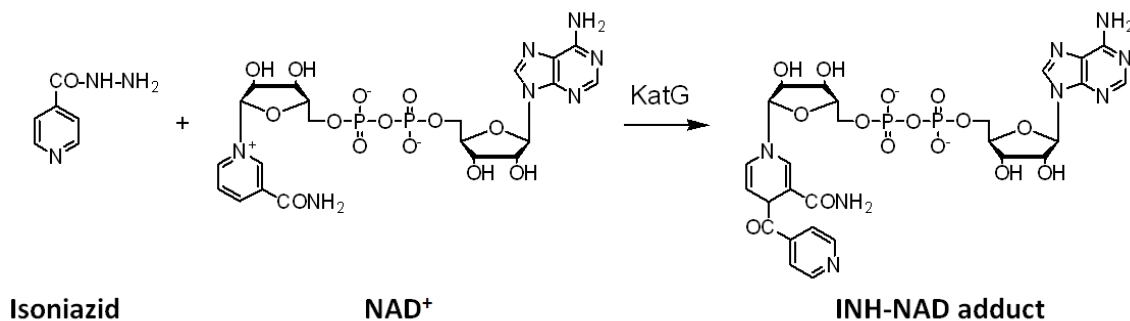
Figure 1.5. Activation of prodrug isoniazid. Isoniazid and cofactor NAD$^+$ form INH-NAD adduct. The INH-NAD adduct has similar structure to the cofactor NADH.
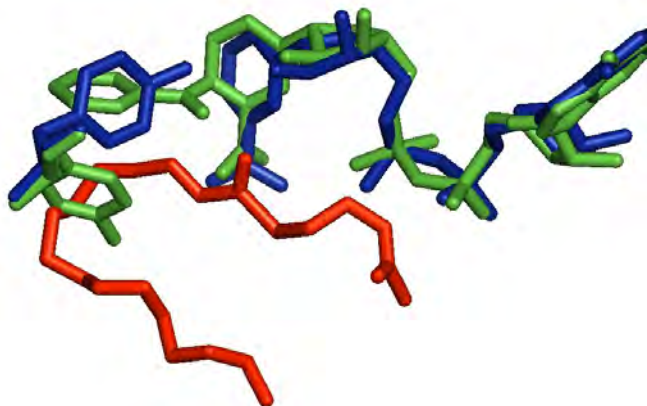
Figure 1.6. Superimposed structure of INH-NAD and substrate analog bound InhA. The INH-NAD adduct (green) occupies the original position of residue Y158 (blue) and pushes Y158 to the substrate binding pocket, preventing substrate binding to the active-site. The substrate analogue is shown in red.

The INH-NAD adduct is a slow and tight-binding inhibitor of InhA (*10*). The binding of INH-NAD to InhA involves a two-step enzyme inhibition mechanism. The initial step is a fast and weak binding process (K_i = 16 nM) followed by a slow conversion to the final complex with an overall K_i = 0.75 nM. Because KatG mutants have been reported to be linked with some clinical drug-resistant TB strains, compounds that do not require activation while directly inhibiting InhA could be promising to overcome the resistance.

Triclosan (**Figure 1.7**), unlike **Isoniazid**, inhibits FabI directly without the requirement of activation by another enzyme. **Triclosan** is a widely used antimicrobial agent that found in many consumer products, such as soaps, toothpastes, mouthwashes, deodorants, shaving creams and cleaning supplies. **Triclosan** is a slow-onset inhibitor of *E. coli* FabI*,* but is a weak binding (K_i = 0.2 μM) and rapid reversible inhibitor of MTB InhA (*11, 12*). Using structure-based drug design approach, 2-(o-Tolyloxy)5-hexylphenol (**PT70**, **Figure 1.7**) was designed and shown to

be a slow-onset and tight binding inhibitor of InhA with an overall K_i value of 22 pM and a residence time of 24 min (*13*).
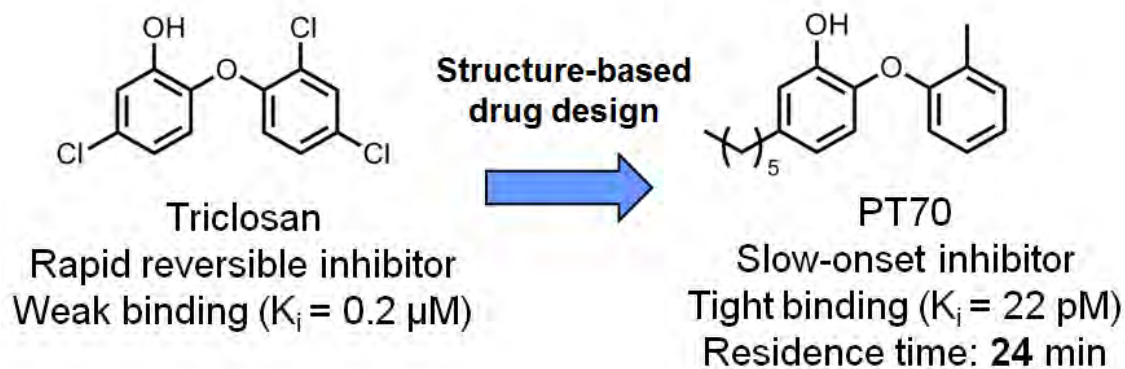
Figure 1.7. Structures of Triclosan and 2-(o-Tolyloxy)5-hexylphenol (PT70). Triclosan is a rapid reversible and weak inhibitor of InhA. Using structure-based drug design approach, **PT70** was found to be a tight binding and slow-onset inhibitor of InhA.

1.4 The two-step binding mechanism and the residence time

Enzyme-inhibitor complex formation often occurs through two mechanisms. The first mechanism describes common enzyme-inhibitor binding events, which only have a simple, one-step association/dissociation process (**Scheme 1.1A**). The second mechanism involves an enzyme isomerization step after initial inhibitor binding, leading to a tighter, reversible enzyme-inhibitor complex (**Scheme 1.1B**). The latter one is also known as induced-fit mechanism. From kinetics experiments, it is known that this isomerization process is much slower than the initial inhibitor binding step for most enzyme systems.

(A)

$$E+I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EI$$

(B)

$$E+I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EI \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EI^*$$
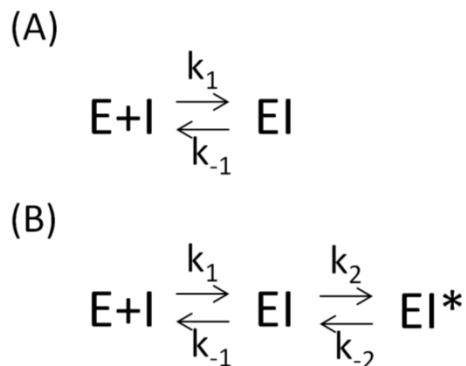
Scheme 1.1. One- (A) and two-step (B) enzyme-inhibitor binding mechanisms

Inhibitor binding kinetics has emerged as an important parameter in lead optimization (*14-18*). Traditionally, a strong binder is thought to have better efficacy than a weak binder (*19*). However, many results have argued that *in vivo* efficacy does not always strongly correlate with binding affinity (*20*). One of the reasons is that the environment where drug-target encounters is different: the *in vitro* system is a closed environment, whereas the *in vivo* system is an open environment. In a closed system, the enzyme and inhibitor concentrations remain constant. In this situation, thermodynamic equilibrium constants such as K_d, K_i or IC_{50} can be measured accurately. On the other hand, in an open system, the enzyme and inhibitor diffuse between cellular compartments over time with changing concentrations, resulting in a dynamic and non-equilibrium process. The thermodynamic equilibrium measurements (ie. equilibrium constant K_i) determined *in vitro* are no longer appropriate to represent or predict *in vivo* efficacy. Hence, an equilibrium independent parameter would be a better measurement for enzyme-inhibitor interaction, such as, the dissociation rate constant. The dissociation rate constant is independent with respect to inhibitor and target concentrations, making an ideal parameter to predict inhibitor efficacy *in vivo*. Moreover, because an inhibitor has its function only when being occupied by the target, the time of an inhibitor resides in the target provides a direct indicator of inhibitor

9

efficacy. Ideally, an inhibitor with a small dissociation rate constant will dissociate slower and reside in the target longer than an inhibitor with a large dissociation rate constant. Because the dissociation rate constant has a unit of s^{-1} or $minute^{-1}$, the dissociation rate constant then can directly represent the life-time of target-inhibitor complex.

Residence time, the reciprocal of the dissociation rate constant, represents the life-time of an enzyme-inhibitor complex (*14, 15*). For the one-step kinetic mechanism, the residence time is defined as the reciprocal of dissociation constant ($1/k_{-1}$). In a two-step kinetic mechanism, the k_{off} includes both forward and reverse steps in enzyme isomerization and the dissociation of the initial EI complex ($k_{off} = k_{-1}k_{-2}/(k_{-1}+k_{2}+k_{-2})$). A hypothetical model suggests that for a series of compounds with the same K_i, a long half-life drug can maintain a higher drug concentration and target occupancy than a short half-life drug (**Figure 1.8**) (*18*). Moreover, the observations from many current drugs demonstrate that they have long residence times on their targets, suggesting drug-target residence time is an important component of *in vivo* drug activity (*18*). In particular, the study of *Francisella tularensis* (ftuFabI) showed that residence time has better correlation with *in vivo* efficacy than does the binding affinity (*21*).
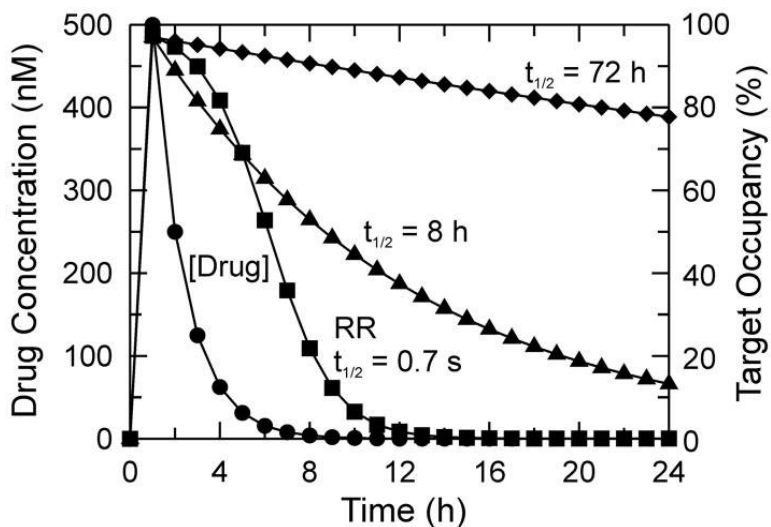
Figure 1.8. A hypothetical model of drug concentration versus time (*18*)**.** The long residence time inhibitor has a higher target occupancy than short residence time inhibitor. The drug is assumed to have the same K_i (14 nM) and to reach a maximum concentration (Cmax = 500 nM) at the target site 1 h after dosing and have an elimination half-life of 1 h. For rapid reversible (RR) inhibitor, it is assumed that drug and target are at equilibrium, so that free drug can rebind to the target. For slow-onset inhibitor, it is assumed that drug does not rebind to the target. The drug concentration at time t is calculated as $D(t) = C_{max} * 2^{(-t_{1/2}/1)}$. Figure taken from (*18*).

1.5 X-ray crystal structures of InhA

The biological assembly of InhA is a homotetramer; each monomer contains 269 amino acids (**Figure 1.9**). As shown in **Figure 1.9B**, InhA has a sandwich-like architecture: β-strands in the middle are surrounded by two layers of α-helices. According to CATH classification, InhA belongs to the 3-layer (αβα) sandwich architecture and the Rossmann fold (*22*). The monomer of InhA contains 14 helices and 11 sheets based on the DSSP analysis (**Figure 1.10**) (*23*).
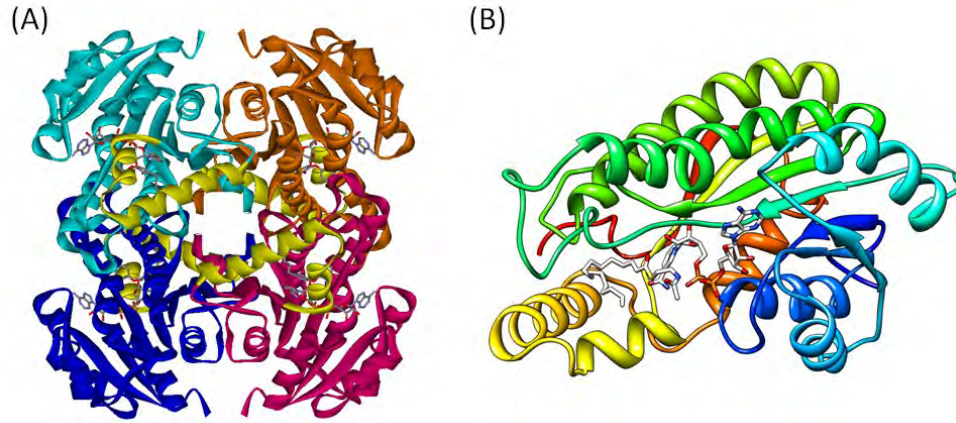
Figure 1.9. The X-ray crystal structure of InhA. (A) Biological assembly and (B) monomer. (PDB code: 1BVR)
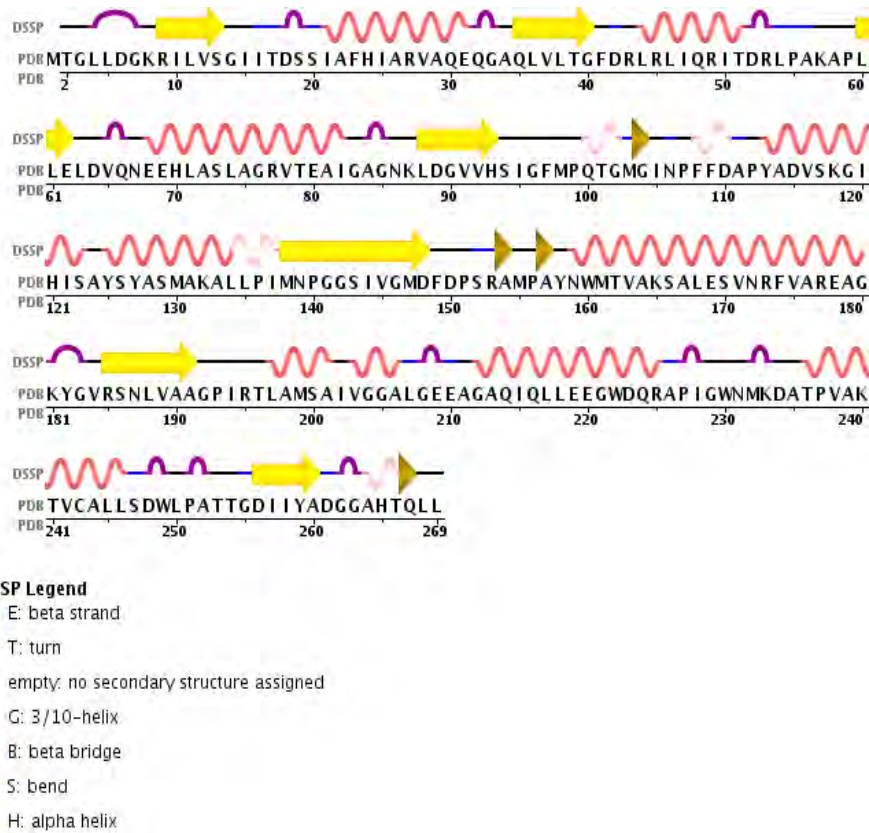
Figure 1.10. Secondary structure of InhA. InhA contains 14 helices and 11 sheets (PDB code: 2X23). The secondary structure was determined by DSSP approach (*23*).

The active-site helix-6 (residue 198 to 206) has been hypothesized to be correlated with slow-onset inhibition. From crystal structure observations, InhA has a disordered helix-6 in the presence of rapid reversible inhibitors, such as **triclosan, PT3** and **PT5**. Conversely, this helix-6 is ordered when InhA is bound with a slow-onset inhibitor (**PT70**) (**Figure 1.11**) (*13*). Similar observations are also noted in *E.coli* FabI. FabI has an ordered active-site helix with the slow-onset inhibitor **triclosan** (*24*). These observations lead to the hypothesis that slow-onset inhibition is coupled to ordering of active-site loop (*13*). The detailed mechanism causing slow-onset inhibition as well as how to control slow-onset inhibition, however, has yet to be completely characterized.
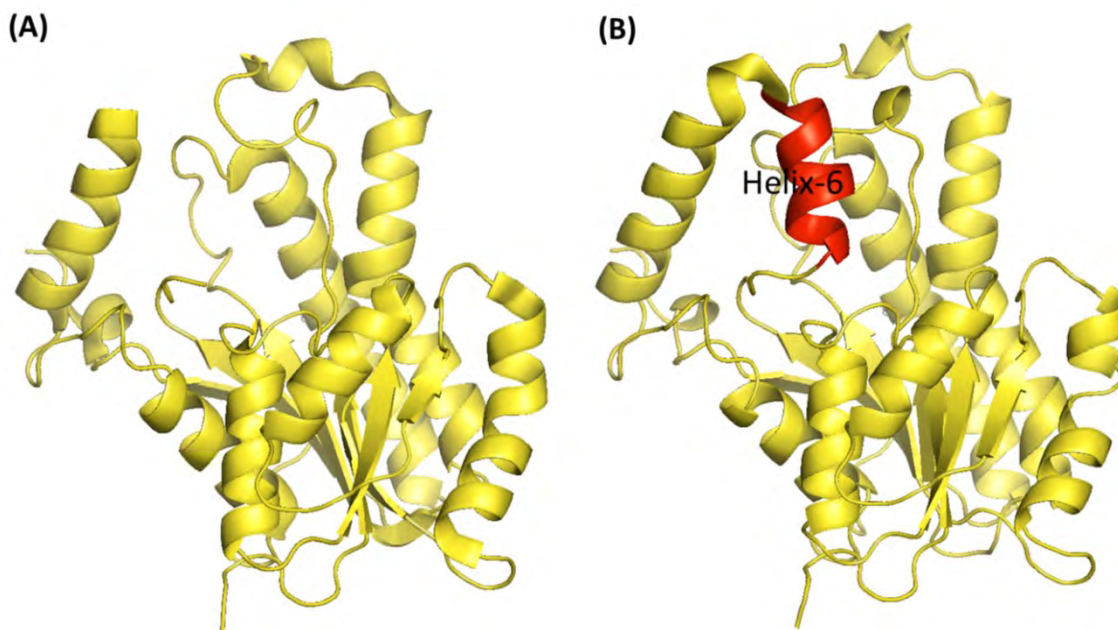
Figure 1.11. Ordering of the helix-6. (A) InhA bound with a rapid reversible inhibitor (B) InhA bound with a slow-onset inhibitor. The slow-onset inhibitor bound InhA has an ordered helix-6.

Because the kinetics results imply slow-onset inhibition of InhA involves a slow conformation change step and the crystal structures only provide static picture, there is a knowledge gap between the resulted kinetic and structural data. Moreover, rational modulation of the residence time of an enzyme-inhibitor complex remains very challenging, since it requires insight into the detailed structural mechanism of isomerization step, and, specifically, the structure of the transition state; these are typically not directly accessible to experimental characterization. We used MD simulation to bridge the gap between the kinetic and structural studies. Because the conformational change is slow, a standard MD simulation approach will not be able to sample this slow structural isomerization step. In this study, we used several advanced MD approaches in addition to the standard MD.

1.6 Molecular dynamics simulation

1.6.1 Background

Molecular dynamics simulation is a computational approach that calculates the movements of atoms and molecules based on the theory of molecular mechanics. The motion of atoms and molecules are determined by numerically solving the Newtonian equations of motion. The forces driving atoms to move are defined by molecular mechanics force fields, which are mathematical description of physical parameters. The common terms of molecular mechanics force fields are bond, angle, dihedral angle and non-bonding interactions (**Figure 1.12**). Unlike quantum mechanics, the force fields method ignores the electronic motions and only calculates the energy as a function of the nuclear positions; therefore the force fields are empirical. Since

the motion of atoms and molecules are based on the force field, an accurate force field is indeed important for molecular dynamics simulation.
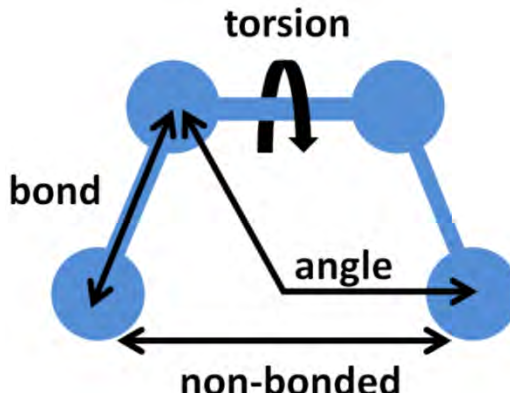
$$U = \sum_{all\ bonds} \frac{1}{2} K_b (b - b_0)^2$$

$$+ \sum_{all\ angles} \frac{1}{2} K_\theta (\theta - \theta_0)^2$$

$$+ \sum_{all\ torsions} K_\phi [1 - \cos(n\phi)]$$

$$+ \sum_{i,j\ nonbonded} \varepsilon_{ij} \left[\left(\frac{R_{ij}}{r_{ij}} \right)^{12} - 2 \left(\frac{R_{ij}}{r_{ij}} \right)^6 \right]$$

$$+ \sum_{i,j\ nonbonded} \frac{q_i q_j}{4 \pi \varepsilon_0 \varepsilon r_{ij}}$$

Figure 1.12. General molecular mechanics force field terms (*25*).

Currently, the most commonly used force field in AMBER is ff99SB (*26*), which has been cited more than 1600 times (until Jan 2014) since 2006. The common way to obtain empirical force field is to fit the force field parameters to experimental or quantum mechanics data. Thus, the accuracy of force fields is always dependent on the quality of experimental results and the available computational resource used to train the force field parameters at that time. For example, the most well-known force field associated with AMBER simulation package during 1990s was the 1994 Cornell force field 94 (ff94) (*27*), which was widely used during that time but less common used today due to the development of more accurate force fields such ff99SB. Because of the limitation of computational resource at that time, the backbone parameters were fit only to a small number of glycine and alanine dipeptide from quantum

mechanical results. The limitation of this force field has been reported to be overstabilization of α-helix secondary structure (*28, 29*). A subsequent modification of ff94, ff96 attempted to reproduce the energy difference between extended and constrained α-helical energies, but it overestimates β-strand propensity (*30*). With more efficient computing power, ff99 refit the backbone force field by including alanine tetrapeptides along with alanine dipeptides (*31*). Although ff99 tried to refit the force field using tetrapeptides to correct the ff94 problem, it introduced incorrect conformation preference for glycine and other amino acids. A further modification ff99SB (*26*) improved the φ, ψ-dihedral terms in the ff99 energy function by fitting both alanine and glycine tetrapeptides from quantum mechanical results, thus achieving a better balance of secondary structure propensities for both glycine and alanine.

1.6.2   Free energy calculation of protein-ligand interactions

It has been more than two decades since people first used computational approaches to calculate and predict binding free energies for protein-ligand complexes. In this section, I briefly summarize some common approaches for free energy calculation for protein-ligand interaction.

1.6.2.1 Endpoint methods

There are two commonly used binding free energy calculation approaches that fall into the endpoint method category: (1) linear interaction energy (LIE) (*32*); (2) molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) and molecular mechanics/generalized born surface area (MM/GBSA) (*33, 34*).

In the LIE approach, the binding free energy is estimated as the energy difference between bound and free states. The basic assumption is that the free energy of binding is a linear response on the polar energies (controlled with the parameter β) and nonpolar energies (controlled with the parameter α) as shown in equation 1.1

$$\Delta G = \alpha \left(U_{vdw,bound} - U_{vdw,free} \right) + \beta \left(U_{elec,bound} - U_{elec,free} \right) \qquad \text{Eq. 1.1}$$

The coefficient $\alpha = 0.161$ was initially derived from training to experimental binding affinity data of four endothiapepsin inhibitors with a theoretically derived β value (β = 0.5) (32). One of the limitations of the LIE approach is that different protein systems may require different sets of coefficients α and β. Thus, for the new system of interest, experimental binding affinity data may be required to train the coefficients before further predictions can be made. But, it is quite surprising that the initial values of α and β also work on other enzyme systems, such as HIV protease (35, 36), trypsin (37) and glucose/galactose binding protein (38). Further parameterization approaches, such as including the surface area term, have been incorporated into LIE calculation to improve accuracy of prediction (39-41).

MM/PBSA and MM/GBSA are two widely used endpoint free energy calculation methods. MM/PBSA and MM/GBSA are post-processing approaches – the free energy is calculated based on existing MD trajectories. Equation 1.2 shows that the binding free energy of ligand (ΔG_{bind}) can be described as the free energy of complex minus the free energies of unbound protein ($G_{unbound\ protein}$) and free ligand ($G_{free\ ligand}$).

$$\Delta G_{bind} = G_{complex} - \left(G_{unbound\ protein} + G_{free\ ligand} \right) \qquad \text{Eq. 1.2}$$

The binding free energy (ΔG_{bind}) consists of three terms: molecular mechanical energy (ΔE_{MM}), entropy ($T\Delta S_{solute}$) and solvation energy (ΔG_{solv}) as shown in equation 1.3

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S_{\text{solute}}$$ Eq. 1.3

where the ΔE_{MM} represents the change in molecular mechanics potential energy upon binding. $T\Delta S_{\text{solute}}$ represents the entropic contribution at temperature T in Kelvin. The configuration entropy can be estimated from normal mode analysis (*42, 43*). The solvation energy is the free energy change to transfer a molecule from vacuum to solvent and can be further divided into polar and nonpolar parts as in equation 1.4

$$\Delta G_{\text{solv}} = \Delta G_{\text{PB}} \, (\text{or} \, \Delta G_{\text{GB}}) + \Delta G_{\text{np}}$$ Eq. 1.4

The polar part (ΔG_{PB} or ΔG_{GB}) is calculated by continuum solvent PB or GB methods.

The nonpolar part (ΔG_{np}) is often estimated to be proportional to the solvent accessible surface area (SASA) as in equation 1.5

$$\Delta G_{\text{np}} = \gamma \, \text{SASA} + \beta$$ Eq. 1.5

where γ is 0.00542 kcal/mol·Å^2 and β is 0.92 kcal/mol from the fits of linear alkanes (*34*).

Recent studies have shown the accuracy of MM/PB(GB)SA approach relies on the length of MD simulation, entropy calculation, solute dielectric constant, and force field (*44, 45*). In short, the result of binding free energy has a high correlation with the length of MD simulation, but longer simulations do not necessarily lead to a better result. The conformational entropy has large fluctuations in the whole trajectory; hence a larger number of frames are required to achieve a converged value. The free energy results are very sensitive to the choice of dielectric constant, thus one should choose appropriate value based on the protein/ligand binding interface. The ff99 (*31*) force field achieved the best prediction with MM/GBSA calculations, while the ff99SB (*26*) force field did the best prediction with MM/PBSA calculations. Ligand partial

charges using the RESP (*46*) charges model gave the best performance compared to AM1-BCC (*47*) and ESP (*48*) charge models. Although the MM/PBSA approach has been used widely in many drug design studies, this method has its limitations. Structured water molecules are often not accounted for in this calculation. Explicitly considering structured water in the calculation may improve the accuracy of prediction (*49*). Another common limitation is that the implicit solvation PB(GB)SA model cannot accurately calculate solvation energy for deeply buried residues or ligands (*50*). Nevertheless, the entropy term seems to be the major source of uncertainty of MM/PB(GB)SA calculations (*44*). The overall free energy shall converge faster if the single-trajectory approach is used instead of the three-trajectory approach (*51*). For similar compounds, the entropy term may be neglected to reduce uncertainty, as the values are roughly equal (*51*).

1.6.2.2 Free energy perturbation or pathway methods

In free energy perturbation (FEP) theory, the Hamiltonian of the target system (H_1) is represented as the sum of the reference Hamiltonian (H_0) and the perturbation term (ΔH) as shown in equation 1.6.

$$H_1(x, p_x) = H_0(x, p_x) + \Delta H(x, p_x) \qquad\qquad \text{Eq. 1.6}$$

where x is the 3N Cartesian coordinates and their conjugate momenta p_x.

The difference of free energy (ΔA) between the target and the reference system can be written in terms of the ratio of the partition function Q_1 and Q_0.

$$\Delta A = -\beta^{-1}\ln\frac{Q_1}{Q_0} \qquad\qquad \text{Eq. 1.7}$$

where β is $1/k_B T$ and $Q = \frac{1}{h^{3N}N!}\iint \exp(-\beta H)\,dx\,dp_x$ $\qquad\qquad$ Eq. 1.8

When substituting the partition function with the expression of the Hamiltonian (**Eq. 1.8**), the equation 1.7 become

$$\Delta A = -\beta^{-1}\ln\langle\exp[-\beta\Delta H(x, p_x)]\rangle_0 \qquad\qquad \text{Eq. 1.9}$$

where $<\ldots>_0$ denotes the ensemble average of the reference state. Because the integration of the kinetic term in the Hamiltonian will be cancelled out, equation 1.9 can be written as

$$\Delta A = -\beta^{-1}\ln\langle\exp(-\beta\Delta U)\rangle_0 \qquad\qquad \text{Eq. 1.10}$$

where $\varDelta U$ is the difference of potential energy between the target and the reference systems.

One thing to be concerned is that the target and the reference systems should be very similar; otherwise the free energy will fluctuate and not converge. To solve this problem, the overall free energy can be calculated by breaking the change into N small steps using perturbations linked by a pathway along the target and reference states as shown in equation 1.11. This is why this approach is also called a pathway approach. Because the intermediates in equation 1.11 may have no physical meaning, this approach is also known as an alchemical approach.

$$\Delta A = \sum_i^{N-1}\Delta A_{i,i+1} \qquad\qquad \text{Eq. 1.11}$$

The FEP approach has been employed widely and one application is the thermodynamics integration (TI) method (*52*). Theoretically, the energy difference between two states can be calculated by numerically integrating the derivative with respect to the distance (i.e. parameter λ) along the path between target and reference states.

$$\Delta A = A(\lambda = 1) - A(\lambda = 0) = \int_0^1\langle\partial U/\partial\lambda\rangle_\lambda\, d\lambda \qquad\qquad \text{Eq. 1.12}$$

$$U(x, \lambda) = \lambda U_1(x) + (1 - \lambda)U_0(x) \qquad\qquad \text{Eq. 1.13}$$

where $U(\lambda)$ is the λ-coupled potential function that corresponds to U(reference) for $\lambda=0$ and U(target) for $\lambda=1$. The integration is carried out over the average of the λ derivative of the coupled potential function at given λ values.

The difference between FEP and TI approaches is shown in Figure 1.11. Both FEP and TI apply a parameter λ to describe the transformation between reference ($\lambda=0$) and target ($\lambda=1$) state. The free energy difference from the FEP approach is the difference between initial and final states (**Figure 1.13A**), while the free energy difference from TI approach is the area under the line since y-axis is partial derivative of the energy (**Figure 1.13B**).
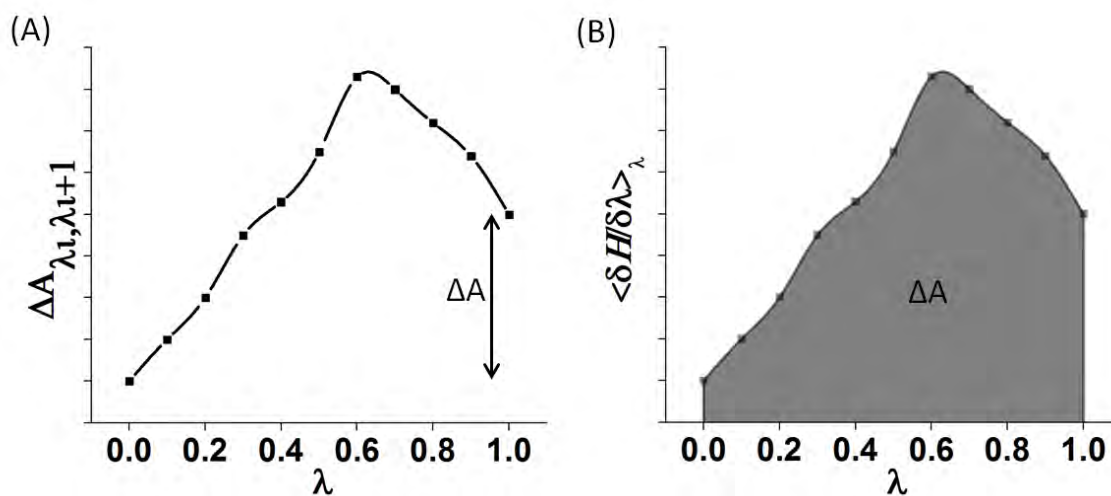
Figure 1.13. Free energy differences by (A) perturbation and (B) thermodynamics integration.

Depending on how the thermodynamic cycle is designed, TI can be used to calculate the relative binding free energy between two ligands with the same receptor or the relative binding free energy between wild-type and mutant receptors with the same inhibitor. For example, in a thermodynamic cycle, **a** and **b** represent the binding free energies for compound **1** and **2**,

respectively. **c** and **d** are the transformation energies for compound **1** into compound **2** in the complex and the compound-alone environment, respectively (**Figure 1.14A**). The binding free energy can be measured through experimental approaches, but is very difficult through simulation approaches. On the other hand, the transformation energy can be measured through a numerical approach, but is very difficult to measure in experimental approaches. Because **a**, **b**, **c**, and **d** are within a thermodynamic cycle, the relative binding free energy (**b** minus **a**) equals to the relative transformation energy (**c** minus **d**). The relative binding free energy between wild-type and mutant with same inhibitor can also be calculated in a similar way (**Figure 1.14B**).
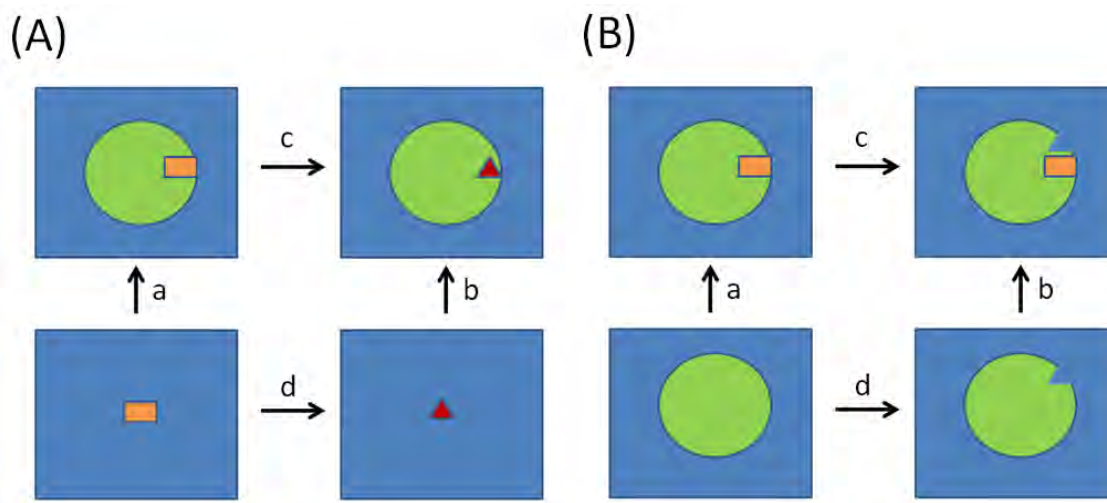
Figure 1.14. Thermodynamic cycle of (A) Two different ligands and (B) Wild-type and mutant receptors.

TI calculation requires complicated setup and extensive simulation time compared to LIE and MM/PBSA approaches. As mentioned above, TI applies a parameter λ to describe the transformation between reference ($\lambda=0$) and target ($\lambda=1$) state. For each λ state, in order to get a smooth transformation that limits the errors in the numerical integration, we have to calculate the

22

perturbations of electrostatic and van der Waals energies separately. A standard procedure of TI calculation for each λ state requires three steps: (1) removing the partial charges on the target atoms; (2) soft-core vdw transformation; and (3) adding the partial charges on the target atoms. A one-step transformation approach, in which electrostatic and van der Waals forces are simultaneously modified, has been implemented into AMBER and reproduced very similar results with those from the three-step transformation approach (*53*). But the one-step transformation approach may require longer time to get a converged free energy result (*53*). A most recent implementation in AMBER has made the TI calculation optimized to run with the pmemd program and gives 2.5 times greater computational efficiency than the sander program (*54*).

1.6.2.3 The unbinding pathway method

The ligand unbinding approach is another way to calculate binding free energy. As binding free energy is the energy difference between the ligand-bound and ligand-free states, one can get the binding free energy by gradually removing the ligand from the active-site and calculating the potential of mean force along the path (**Figure 1.15**). The ligand dissociation path can be obtained from steered MD simulation, and the free energy along this pathway can be calculated via the non-equilibrium Jarzynski relationship (*55-57*) or equilibrium umbrella sampling approaches (*58*). Several studies have shown this approach calculates binding free energies comparable to experimental results (*59-65*). The advantage of this approach is not only provides relative free energy between ligand-bound and ligand-free states, but also the energy profile along the unbinding path(s). The energy profile along the path is very valuable for the study of binding kinetics because all energy barriers in the ligand unbinding path will contribute

23

to the net dissociation rate of the inhibitor (*66*). The disadvantage of this approach is that there might be multiple dissociation paths, thus multiple or long timescale simulations may be required to obtain a converged free energy. Also, the translational entropy is excluded from this approach.
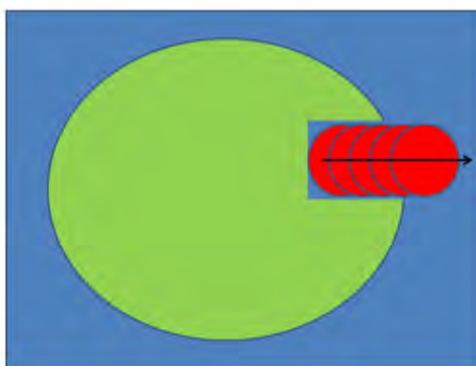
Figure 1.15. A cartoon representation of ligand unbinding path.

1.6.2.4 Common limitations of free energy calculation

Accurately predicting the binding free energy is always a challenging problem for computational simulations. For all of the above approaches, accurate force fields are especially critical. Unlike quantum mechanics, molecular mechanics ignores the electronic motion and the force fields method is empirical. Thus, there is no perfect or correct force field and there is always room to improve molecular dynamics simulation accuracy. For example, new force fields terms like polarizable force fields have been developed to overcome the limitation of fixed charges models (*67-75*). Sampling and convergence problems are also critical for MD simulations, and several enhanced sampling approaches have been developed to overcome the problems, such as the replica exchange approach (*76*), adaptively biased MD (*77*), and accelerated MD (*78*), to name a few. Moreover, the implementation of GPU (graphics processing

unit) in MD simulation makes long timescale simulation much easier than before that definitely helps limit the convergence issue (*79-81*). Overall, one should bear in mind the limitation of the free energy calculations when applying these approaches in calculating binding free energies.

1.6.3 Studying of InhA using MD simulations

Several studies of InhA using MD simulation have obtained good agreement with experimental results. For example, the drug-resistance problem is a popular topic in InhA research (*82-85*). Clinical mutants of I16T, I21V, I47T, S94A, and I95P have been shown to decrease the binding affinity of the cofactor in experimental studies (*86*). The MD simulation results demonstrate that the NADH pyrophosphate moiety in I21V and I16T mutants undergoes considerable conformational changes to reduce its binding affinity to the receptor (*82*). In addition, other researchers have used MD simulation to study the binding affinity of the tautomeric forms of INH-NAD adduct (*87*). The tautomeric forms of the INH-NAD adduct caught the community's attention because the chemically related INH-NADP adduct forms a 4R conformation in MTB dihydrofolate reductase but a 4S conformation of INH-NAD adduct is found in MTB InhA (*88*). The binding affinity calculation suggested that the 4S form of INH-NAD adduct in InhA has stronger binding affinity than the 4R form, thus representing the predominant active form of the INH-NAD adduct in the crystal structure (*87*). In addition to InhA, our previous MD simulations and binding free energy calculations in the *E. coli* FabI system also achieved assuring agreement with the experimental results and indicated the shape of the inhibitor is very important for the overall binding free energies in this enzyme system (*89, 90*). Moreover, another study of *Plasmodium falciparum* FabI also demonstrates the extreme

value of MD simulation in assisting the drug discovery targeting the enoyl-ACP reductase enzymes (*91*).

1.6.4   Advanced MD simulation approaches used in this dissertation

1.6.4.1 Partial nudged elastic band (PNEB)

Nudged elastic band (NEB) is a MD simulation approach to find low-energy-path structures between two fixed end-point structures (*92*). The structures between the end-points structures can be treated as multiple beads connected to the previous and next beads with a string, and these beads are spaced evenly between the end-point structures. NEB was originally derived from the plain elastic band method (*93*), however, because the plain elastic band method has the tendency to cut corners in the energy landscape, NEB truncates the spring forces in directions perpendicular to the tangent of the path to prevent corner cutting. The force on atoms can be divided into perpendicular ($\perp$) and parallel force ($\parallel$), such as

$$F = F^{\perp} + F^{\parallel} \qquad\qquad \text{Eq. 1.14}$$

$$F^{\perp} = -\nabla V(P_i) + (\nabla V(P_i) \cdot \tau)\tau \qquad\qquad \text{Eq. 1.15}$$

$$F^{\parallel} = [k_{i+1}(P_{i+1} - P_i) - k_i(P_i - P_{i-1}) \cdot \tau]\tau \qquad\qquad \text{Eq. 1.16}$$

where P_i is the 3N dimensional position vector of image i, k_i is the spring constant between image i-1 and image i, V is the potential described by the force field, and τ is the 3N dimensional tangent unit vector that describes the path.

With an additional energy minimization step (i.e. simulated annealing (*94*)), the NEB results can be used to represent the minimum-energy path (*92*). Partial NEB (PNEB) is a recently modified NEB that allows users applying spring force on a specific region. With this implementation, simulation of large system in explicit solvent has become applicable (*95*).

1.6.4.2 Steered MD

Steered MD applies an external force on the target atoms or molecules, and drives changes in the coordinate that allows us to explore biological processes on the timescale where MD simulation is accessible. For example, ligand unbinding process and large conformational changes in molecules require long timescale simulation. By applying external forces on the ligand to pull it from the active-site, we can study the unbinding process in a smaller timescale (*59-64*).

1.6.4.3 Umbrella sampling and potential of mean force

Umbrella sampling is a MD simulation approach used to obtain free energy along a reaction coordinate (*58*). By applying a bias potential to sample system's configurations along a particular reaction coordinate, such as angles, distance, …etc., both rare configurations (high energy) and dominant configurations (low energy) are adequately sampled. The unbiased free energy, also called the potential of mean force (PMF), can be calculated by post-processing a series of umbrella sampling simulations with the weighted histogram analysis method (WHAM) (*96-98*). The limitation of this approach is that a well-defined reaction coordinate(s) is required

for umbrella sampling; otherwise the calculated PMF energy values will underestimate the true free energy. However, because protein motions involve too many atoms (degrees of freedom) to assign individual reaction coordinates, one is unlikely to find a perfect reaction coordinate. Therefore, we should treat umbrella sampling results as guides of quality and trends rather than quantity and absolute values.

1.7 Overview of this dissertation

Since 2004, slow-onset inhibition of InhA has been characterized by enzyme kinetics assays. Although X-ray crystal structures provide us with an understanding of how the inhibitors interact in the active-site, the information provides only a static picture that cannot explain the mechanism of slow-onset inhibition which is a dynamic process involving large conformational changes. Because the goal of InhA lead optimization is to design long residence time inhibitors, a better understanding of the conformational change mechanism is the key component of the research project. The goal of my research is to understand the detailed interactions between the inhibitor and the receptor at the atomic level as well as to explore the possible structural mechanism of slow-onset inhibition. A good structural model for slow-onset inhibition requires a multidimensional energy landscape that defines the relative population of different conformational states and the free energy barriers between them. Since the atomistic structural transition is very difficult to measure with experimental approaches, my research applies advanced MD simulations to study the slow-onset inhibition mechanism at the atomic scales.

This dissertation is arranged as follows: In chapter two, I propose the induced-fit conformational change mechanism for the slow-onset inhibition. In chapter three, I show the

detailed interaction between inhibitors and InhA. In chapter four, I show the free energy landscape along the open-closed conformational change path and demonstrate how to change the key interactions to control the transition state and the free energy landscape. In chapter five, I show the virtual screening approaches to screen new InhA inhibitors.

Chapter Two

The induced-fit conformational changes of *M. tuberculosis* InhA: Insights from molecular dynamics simulation and principal component analysis

2.1 Introduction

Protein dynamics plays an important role in biological function. Protein will undergo conformational changes after binding of substrate/ligand, post translational modification, or in different solvent environments (*99-104*). The magnitude of conformational changes can be as large as involving entire protein domains or as small as some sidechain movement. For example, adenyl kinase (ADK), a phosphotransferase enzyme in cellular energy homeostasis, undergoes a large conformational rearrangement of the ATP-binding and AMP-binding domains from an open state to a closed state (*105*). The backbone RMSD between two different conformations of ADK is as high as 7Å. This "open-to-closed" conformational change is also found in different enzyme systems, such as HIV protease (*106*), 1-deoxy-d-ylulose-5-phosphate reductoisomerase (*107*), glutamine-binding protein (*108*) and formate dehydrogenase (*109*).

The conformational change of protein after binding of ligands has often been explained by two different models: induced-fit conformational change and conformational selection (also called population shift) (**Figure 2.1**). The induced-fit model can be traced back to 1958, when Koshland proposed this model that the ligand-free protein (apo) structure changes its conformation after binding of ligand, leading to the distinct ligand-bound conformation (holo) (*110*). This model is supported by a large amount of structures from different protein systems showing apo and holo structures have different conformations. The conformational selection

model, a recently growing model based on the results of NMR and MD simulations, holds a different view to the induced-fit model (*111*). This model suggests that multiple protein conformations preexist but with different probabilities (i.e. conformations of open and closed exist at the same time, while the open conformation is the dominant state) and the dominant population shifts to the other one after ligand is bound. These two models with different viewpoints are able to explain most ligand binding mechanisms. For some ambiguous cases, it seems that either model can explain the ligand binding mechanism (*112*). A general rule has been suggested that the induced-fit model better explains those strong binding protein-ligand complexes while the conformational selection model fits better for weak binding complexes (*113*). For enzymes with lid-gate active-sites, the induced-fit mechanism seems to be more appropriate because the protein can only permit the ligand with an open active-site (*114*). Because induced-fit and conformation selection model are two extreme cases, a mixed-model has been proposed to explain semi-closed/open conformation of the maltose binding protein (*112*). In the mixed-model, the first step is a population shift triggered by ligand binding into a semi-closed/open state, followed by an induced-fit step to the closed state. The first step involves a large domain motion while the second step is only local conformational changes.
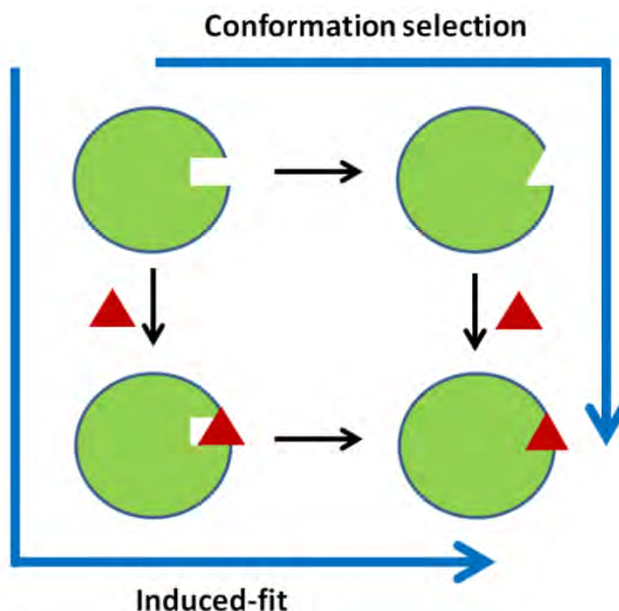
Figure 2.1. Induced-fit and conformational selection binding mechanism (*104*).

The substrate-binding loop (helix-6) of InhA behaves differently in the rapid reversible and slow-onset inhibitors bound complexes. The active-site loop of InhA has been hypothesized to control the process of slow-onset inhibition (*13*). With reversible inhibitors bound to the active-site, such as **Triclosan** (PDB ID: 2B35), **PT3** (PDB ID: 2B36), and **PT5** (PDB ID: 2B37) (inhibitor structures in **Figure 2.2**), the crystal structure has a weak electron density map (disordered) in the helix-6 region (**Figures 2.3**) (*12*). When InhA is bound with **PT70**, a slow-onset inhibitor of InhA, helix-6 is ordered (**Figure 2.4**) (*13*). Similar phenomena are also observed in *E. coli* FabI, the homologue of InhA, where the substrate-binding loop is ordered in the presence of **triclosan** (PDB ID:1QSG), a slow-onset inhibitor of *E. coli* FabI **(Figure 2.4)** (*24*). Based on the observations of MTB InhA and *E. coli* FabI, the dis/ordering of the substrate-binding loop was hypothesized to be correlated with slow-onset inhibition of InhA, stating InhA shall have an ordered substrate-binding loop if a slow-onset inhibitor is bound to the enzyme

32

(*13*). However, a recently solved InhA structure co-crystallized with the rapid reversible inhibitor **PT155** (PDB ID: 4OXK, 4OXN) exhibits an ordered substrate-binding loop (**Figure 2.4**). One possible reason is the crystal packing effects; however, after examining the structures, it seems the crystal packing effect is not the reason. Therefore, the structure-based mechanism of slow-onset inhibition is more complicated than the concept of ordering of the substrate-binding loop.
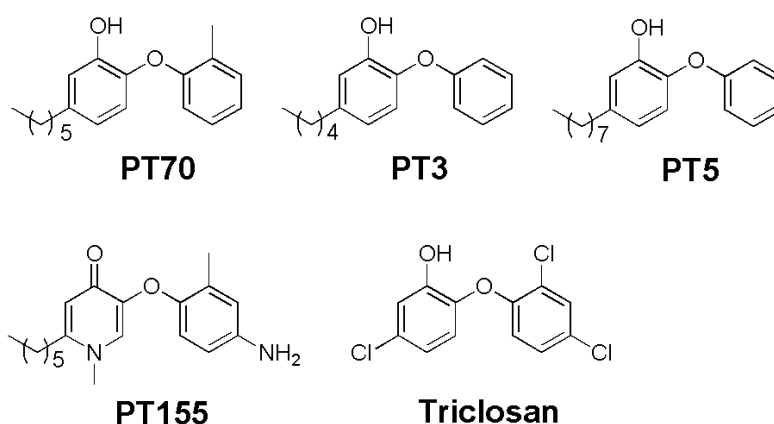
Figure 2.2. Structures of PT70, PT3, PT5, PT155, and triclosan.
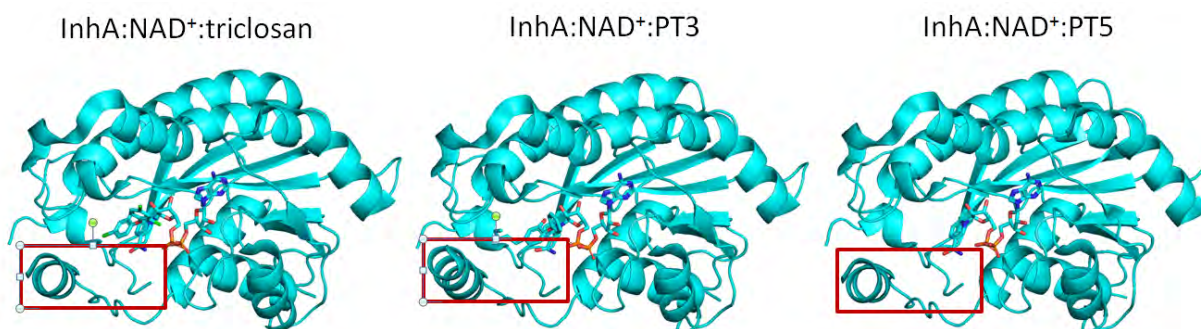
Figure 2.3. The triclosan, PT3, and PT5 bound InhA structures. The substrate-binding helix-6 (red box) is missing (disordered) in the presence of rapid reversible inhibitors **triclosan**, **PT3** and **PT5**. PDB ID for **triclosan**, **PT3** and **PT5** bound InhA are 2B35, 2B36, 2B37, respectively.
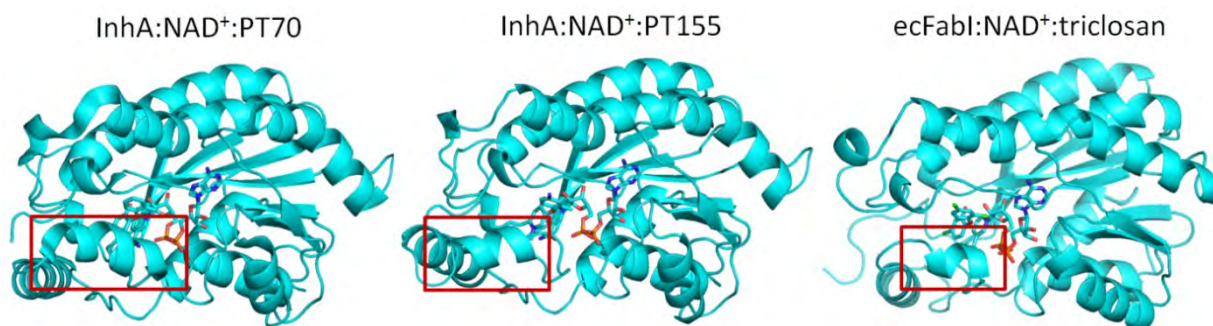
Figure 2.4. The PT70 and PT155 bound InhA, and triclosan bound ecFabI. The ordered helix-6 was only found previously in the slow-onset inhibitors bound InhA. Recently, InhA bound with **PT155**, a rapid reversible inhibitor of InhA, also has an ordered helix-6.

To get an idea of what kind of structural motions are correlated with the slow-onset inhibition of InhA, a systematic characterization of the active-site conformation and a dynamic view of the protein structure are required. While crystal structures provide valuable information regarding how the inhibitor binds to receptor at the atomic level, the results of x-ray crystal structures are static and neglect dynamics. A static picture, however, is insufficient to provide a complete picture of the kinetic properties of slow-onset inhibition, which involves a slow conformational change step as suggested by enzyme kinetics. Moreover, the non-physiological conditions used for the crystal structures such as high salt and low temperature as well as crystal packing effects may mislead the interpretation of kinetics data. To better understand the structural differences among all InhA structures as well as to obtain a dynamic view into how InhA interacts with the inhibitors, all InhA crystal structures were compared first, followed by a series of standard (unrestrained) MD simulations and structural comparisons. After analyzing all the available crystal structures of InhA, we found the ordered active-site loop can be further characterized as "closed" or "open". The MD simulation results demonstrated the helix-6 and 7 were very flexible when the simulations started from the open state, whether InhA was bound with slow-onset or with rapid reversible inhibitors. On the other hand, when the MD simulations

34

started from the closed state, the helix-6 and 7 were more stable with slow-onset inhibitor than with rapid reversible inhibitors. Structural comparison and principal component analysis of the trajectories suggested that the open conformation of InhA tended to close its active-site, regardless of whether it was bound with rapid reversible or slow-onset inhibitors. Conversely, the closed active-site tended to keep the active-site closed when slow-onset inhibitor was bound, whereas rapid reversible inhibitor bound InhA tended to open the active-site. These results implied that the open and closed states likely represent the initial EI and final EI* states as described in the **Scheme 1.1,** and the induced-fit conformational changes from open to closed states are likely to be the structure basis mechanism of slow-onset inhibition in the InhA system.

2.2 Simulation Details

2.2.1   Initial structures

The tetramer InhA:NAD$^+$:**PT70** and InhA:NAD$^+$:**C16-NAC** structures (PDB ID: 2X23 (*13*) and 1BVR (*11*), respectively) were used to build the starting structures of the closed and open states of InhA. The inhibitor poses of **PT70**, **PT3**, **PT5** and **triclosan** were taken from the crystal structures (PDB ID: 2X23, 2B36, 2B37, and 2B35, respectively). We assumed the inhibitors occupy in the same position, so the inhibitors share the same poses in the closed and open states.

2.2.2 Setup of Molecular Dynamics Simulation

AMBER ff99SB (*26*) and GAFF (*115*) force field parameters were assigned to the protein and inhibitor, respectively. The partial atomic charges of each inhibitor were derived using RESP approach (*46, 116*). The RESP charges were computed using Gaussian98 (*117*) with the HF/6-31G* basis set. The force field parameters of the cofactor NAD^+ were taken from other studies (*118, 119*). Each complex was solvated in a truncated octahedral TIP3P (*120*) water box with a minimum distance of 14 Å between the water box edge and solute, resulting in ~ 80,000 atoms in total. The MD simulations were run at 300 K, constant 1 atm pressure, and periodic boundary conditions. Twenty nanosecond (ns) MD simulations were performed for each compound. The first 1 ns was the equilibration procedure, followed by a 19 ns production run. The first step of equilibration was 10,000 steps of steepest descent minimization with 100 kcal mol^{-1} $Å^{-2}$ restraints on all atoms except water molecular and hydrogen atoms. The second step was heating the system from 100 to 300 K at constant volume over 100 ps with 100 kcal mol^{-1} $Å^{-2}$ restraints on non-water and non-hydrogen atoms, followed by 100 ps with the same restraints at constant 300 K temperature and 1 atm pressure. The third step was 250 ps MD with a restraint weight of 10 kcal mol^{-1} $Å^{-2}$ on the non-water and non-hydrogen atoms at constant 300 K temperature and 1 atm pressure. The following steps only restrained the backbone atoms and gradually reduced the restraints weight from 10 to 0.1 kcal mol^{-1} $Å^{-2}$ at constant 300 K and 1 atm. This was carried out by 100 ps with 10 kcal mol^{-1} $Å^{-2}$ restraint, followed by 100 ps with 1 kcal mol^{-1} $Å^{-2}$ restraint, and 100 ps with 0.1 kcal mol^{-1} $Å^{-2}$ restraint. The last step of equilibration was 250 ps of unrestrained MD.

2.2.3 Principal component analysis (PCA)

Principal component analysis is an orthogonal linear transformation method that converts a set of observations into a set of linear variables called principal components (*121*). Since the molecular motion involves too many atoms and the pattern of motion is hard to be determined by "viewing" the trajectory, the advantage of PCA for trajectory analysis is that it reveals the dominant motion. In this study, 19 ns production trajectories were analyzed using ptraj and the lowest frequency motion (first mode) of each system was shown in this study.

2.2.4 Inhibitor unbinding calculation

Steered MD simulation approach was used to push out the inhibitor to get the initial unbinding path. The unbinding reaction coordinate was assumed to be the distance between the inhibitor and receptor. This was done by gradually pushing out the inhibitor using 3 kcal mol^{-1} Å^{-1} force on the center of mass of backbone atoms of residues 150, 155, 190, 193, 218, 222, 225, 261, 264, 265 and A ring of inhibitor. After the initial path was formed, umbrella sampling and WHAM (weight histogram analysis method) were used to calculate the free energy profile along the path. For the umbrella sampling, 83 windows (with interval of 0.5 Å) were performed. 5 ns umbrella sampling per window was performed and the first 1 ns of the trajectories were excluded in the free energy calculation.

2.3 Results and Discussion

2.3.1 Closed, open, and disordered conformations

After examining all 46 published PDB structures (until Dec, 2013), we found the active-site can be characterized as closed, open and disordered based on the positions of helix-6 and 7 (**Figure 2.5**). **Table 2.1** lists the information of all InhA crystal structures. The major conformation is the open form, as found in the binary complex (InhA:NAD$^+$), substrate-analogue bound complex (InhA:NAD$^+$:**C16-NAC**), and most of the inhibitor-bound ternary complexes (InhA:NAD$^+$:inhibitor). The closed conformation is found in the diphenyl ether inhibitor bound structures, such as **PT70**, **PT10, and PT91**. The major difference between the open and closed conformations is the position of helix-6. In the open conformation, helix-6 moves away from strand-4 (residue 95 to 98) and the distance between helix-6 and strand-4 is ~ 10 Å. In the closed conformation, the active-site entrance is covered by helix-6, resulting in a "closed" and "locked" active-site pocket. The backbone RMSD of helix-6 and 7 between the open and closed conformation is 4.4 Å. The disordered helix-6 is found in the binary complex and some inhibitor-bound ternary complexes.
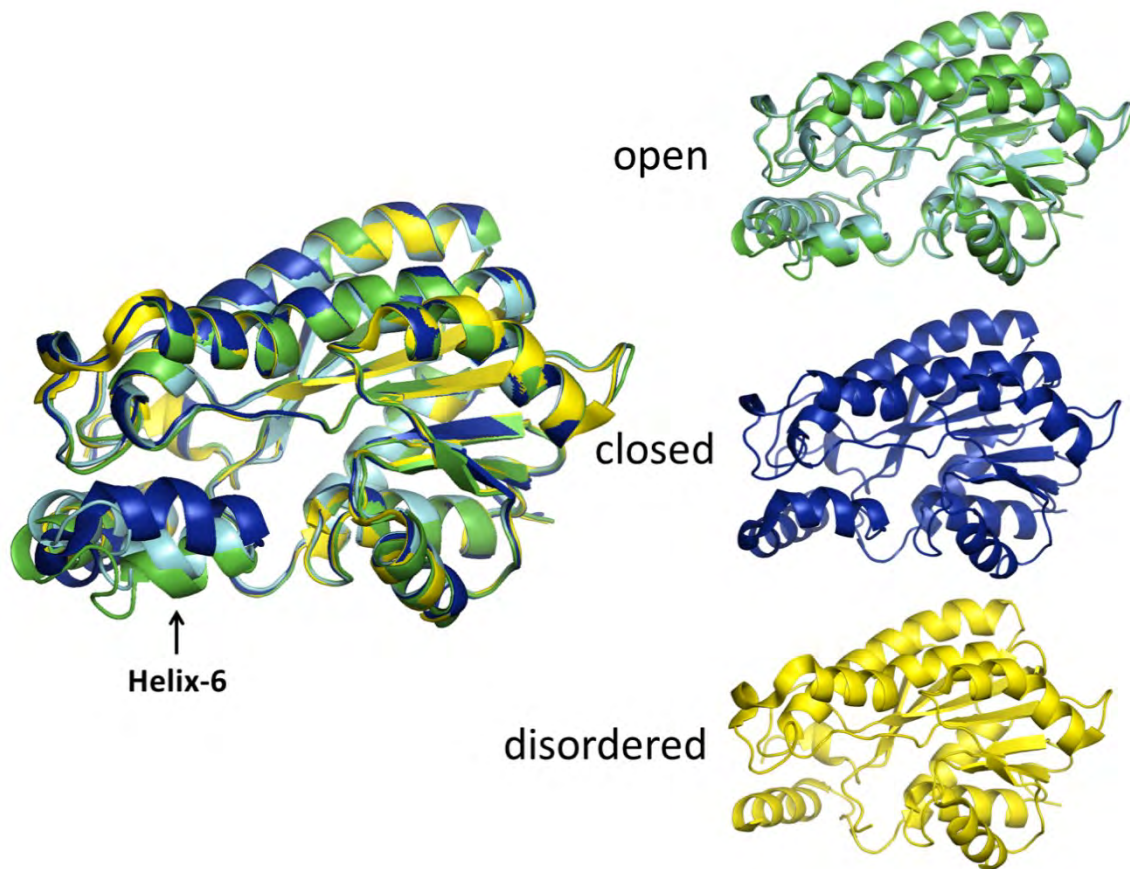
Figure 2.5. Superposing different active-site conformations. The active-site can be characterized as open, closed, or disordered. Representative structures are PDB ID 2AQ8 (cyan), 1BVR (green), 2X23 (blue), and 2B35 (yellow).

Table 2.1. Summary of active-site conformation

PDB ID	Description	Resolution (Å)	Inhibitor	Active-site form	Reference
4OXY	complex with NAD and inhibitor	2.35	PT10	closed	(*122*)
4OYR	complex with NAD and inhibitor	2.30	PT91	closed	(*122*)
4OXN	complex with NAD and inhibitor	2.29	PT155	open and closed	(*122*)
4OXK	complex with NAD and inhibitor	1.84	PT155	open and closed	(*122*)
4BGE	S94A mutant in complex with pyridomycin	2.25	pyridomycin	open	(*123*)
4BGI	S94A mutant in complex with NAD and OH-141	2.09	3-hydroxy-N-[(2R,5R,6S,9S,10S,11R)-10-hydroxy- 5,11-dimethyl-3,7,12-trioxo-2-(propan-2-yl)- 9-(pyridin-3-ylmethyl)-1,4-dioxa-8-azacyclododecan- 6-yl]pyridine-2-carboxamide	open and disordered	(*123*)
4BII	Complex with pyridomycin	1.95	pyridomycin	open	(*123*)
4DQU	D148G mutant in complex with NAD	2.45	no inhibitor	open	(*124*)
4DRE	complex with NAD	2.40	no inhibitor	open	(*124*)
4DTI	S94A mutant in complex with NAD	1.90	no inhibitor	open	(*124*)
3OEW	complex with NAD	2.20	(4S)-2-methyl-2,4-pentanediol	open	(*125*)
3OEY	T266E mutant in complex with NAD and inhibitor	2.00	(4S)-2-methyl-2,4-pentanediol	open	(*125*)
3OF2	T266D mutant in complex with NAD and inhibitor	2.00	(4S)-2-methyl-2,4-pentanediol	open	(*125*)
2X22	complex with NAD and inhibitor	2.10	5-hexyl-2-(2-methylphenoxy)phenol (PT70)	closed	(*13*)

2X23	complex with NAD and inhibitor	1.81	5-hexyl-2-(2-methylphenoxy)phenol (PT70)	closed	(*13*)
3FNE	complex with NAD and inhibitor	1.98	2-(2,4-dichlorophenoxy)-5-(pyridin-2-ylmethyl)phenol	closed, open, and disordered	(*126*)
3FNF	complex with NAD and inhibitor	2.30	5-benzyl-2-(2,4-dichlorophenoxy)phenol	closed, open, and disordered	(*126*)
3FNG	complex with NAD and inhibitor	1.97	5-(cyclohexa-1,5-dien-1-ylmethyl)-2-(2,4-dichlorophenoxy)phenol	open	(*126*)
3FNH	complex with NAD and inhibitor	2.80	2-(2,4-dichlorophenoxy)-5-(2-phenylethyl)phenol	open	(*126*)
2PR2	complex with INH-NADP	2.50	INH-NADP	open	(*127*)
2NSD	complex with NAD and inhibitor	1.90	N-(4-methylbenzoyl)-4-benzylpiperidine	open	(*128*)
2IDZ	complex with INH-NAD	2.00	INH-NAD adduct	open	(*129*)
2IE0	I21V mutant in complex with INH-NAD	2.20	INH-NAD adduct	open	(*129*)
2IEB	S94A mutant in complex with INH-NAD	2.20	INH-NAD adduct	open	(*129*)
2IED	S94A mutant	2.14	no inhibitor	open	(*129*)
2H9I	complex with ETH-NAD	2.20	ETH-NAD adduct	open	(*128*)
2NTJ	complex with PTH-NAD	2.50	PTH-NAD adduct	open and closed	(*128*)
2H7I	complex with NAD and inhibitor	1.62	(3S)-1-cyclohexyl-5-oxo-n-phenylpyrrolidine- 3-carboxamide	open	(*130*)
2H7L	complex with NAD and inhibitor	1.73	(3S)-n-(3-bromophenyl)-1-cyclohexyl-5-oxopyrrolidine- 3-carboxamide	open	(*130*)
2H7M	complex with NAD and inhibitor	1.62	(3S)-1-cyclohexyl-n-(3,5-dichlorophenyl)-5-oxopyrrolidine-3-	open	(*130*)

41

			carboxamide		
2H7N	complex with NAD and inhibitor	1.90	(3S)-n-(5-chloro-2-methylphenyl)-1-cyclohexyl- 5-oxopyrrolidine-3-carboxamide	open	(*130*)
2H7P	complex with NAD and inhibitor	1.86	(3S)-n-(3-chloro-2-methylphenyl)-1-cyclohexyl- 5-oxopyrrolidine-3-carboxamide	open	(*130*)
2NV6	S94A mutant in complex with INH-NAD	1.90	INH-NAD adduct	open	(*131*)
2AQ8	complex with NAD	1.92	no inhibitor	open	(*132*)
2AQH	I21V mutant in complex with NAD	2.01	no inhibitor	open	(*132*)
2AQI	I47T mutant in complex with NAD	2.20	no inhibitor	open	(*132*)
2AQK	S94A mutant in complex with NAD	2.30	no inhibitor	open	(*132*)
2B35	complex with NAD and inhibitor	2.30	Triclosan	disordered	(*12*)
2B36	complex with NAD and inhibitor	2.80	5-pentyl-2-phenoxyphenol (PT3)	disordered	(*12*)
2B37	complex with NAD and inhibitor	2.60	5-octyl-2-phenoxyphenol (PT5)	disordered	(*12*)
1P44	complex with NAD and inhibitor	2.70	5-{[4-(9h-fluoren-9-yl)piperazin-1-yl]carbonyl}- 1h-indole	open	(*133*)
1P45	complex with NAD and inhibitor	2.60	Triclosan	open and closed	(*133*)
1BVR	complex with NAD and C16-fatty-acyl-substrate	2.80	trans-2-hexadecenoyl-(n-acetyl-cysteamine)-thioester	open	(*11*)
1ZID	complex with INH-NAD	2.70	INH-NAD adduct	open	(*134*)
1ENY	complex with NAD	2.20	no inhibitor	open	(*135*)
1ENZ	complex with NAD	2.70	no inhibitor	open	(*135*)

Because the active-site conformation can be characterized as open and closed, it would be interesting to know what kind of conformational change will happen to the active-site helix-6 in the presence of the inhibitors, especially in the presence of rapid reversible inhibitors; since the crystal structures do not have clear structure in this region. We speculated that the slow-onset inhibitors bound InhA should remain a stable closed active-site conformation while the rapid reversible inhibitors bound InhA would have flexible helix-6 when the simulations started from the closed conformation. On the contrary, because the open conformation is likely to be the initial conformation which inhibitors encounter, it would be interesting to know what kind of conformational change will happen to the InhA upon binding of the slow-onset and rapid reversible inhibitors. We speculated that all slow-onset and rapid reversible inhibitors bound InhA would move toward to a closed active-site when the simulations started from the open conformation.

2.3.2   MD simulation started from the closed conformation

To get a better idea why InhA with **PT70** has an ordered active-site loop while InhA with **triclosan**, **PT3** or **PT5** has a disordered active-site loop, a series of MD simulations were performed with these inhibitors. Because the crystal structures already show that InhA has an ordered helix-6 with slow-onset inhibitor **PT70** and this helix becomes disordered with rapid reversible inhibitors **triclosan**, **PT3,** and **PT5**, we hypothesized the active-site loop region (helix-6 and 7) would remain in a stable state in the **PT70** bound complex when the simulation started from the closed conformation. On the other hand, this active-site loop region would become flexible with the rapid reversible inhibitors **triclosan**, **PT3,** and **PT5**.

Because **PT70** bound InhA has a closed conformation, it is expected that the simulation also have a stable closed conformation. To know whether the structure deviates from the initial crystal structure, we measured the whole protein backbone RMSD (root mean square deviation). As expected, the helix-6 and 7 region was stable with the slow-onset inhibitor **PT70**. During the 20 ns simulation, the protein had an average backbone RMSD ~ 1.0 Å compared to the crystal structure, indicating the protein structure was stable (**Figure 2.6A**). Upon a closer examination of the substrate-binding loop region, the average backbone RMSD of the helix-6 and 7 region was ~ 1.7 Å compared to the crystal structure (**Figure 2.6B**). The low RMSD values suggested the protein does not undergo large conformational change. In order to know the flexibility of protein structure, RMSF (root mean square fluctuation) referenced to the average structure was also calculated. As shown in **Figure 2.6C**, the RMSF of helix-6 was below 1 Å, suggesting helix-6 was stable. The MD result, as expected, is consistent with the observations made from crystal structure, demonstrating that InhA with **PT70** has a stable and ordered helix-6.
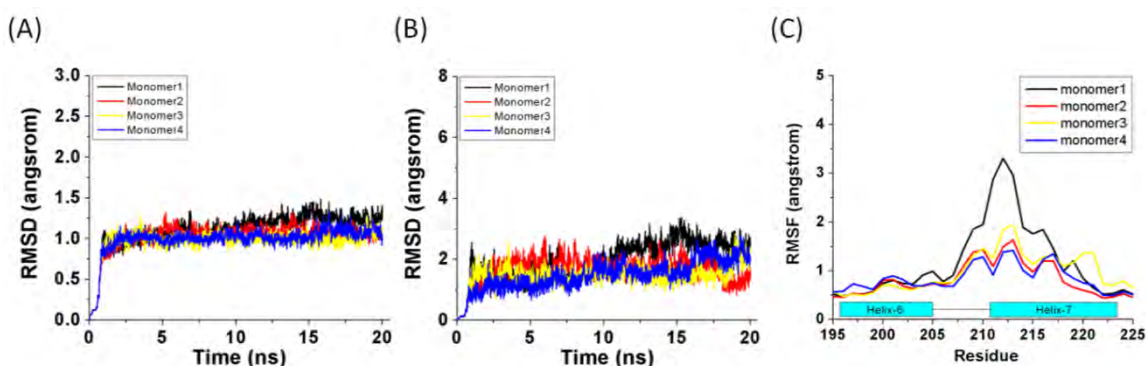
Figure 2.6. RMSD and RMSF of the InhA:NAD$^+$:PT70 complex. (A) Backbone RMSD of the whole protein. (B) Backbone RMSD of the Helix-6&7 region. (C) Backbone RMSF of the helix-6&7 region. RMSD was referenced to the closed crystal structure. RMSF was referenced to the average structure.

On the contrary, the MD simulation results showed that the helix-6 region became flexible in the presence of rapid reversible inhibitors. From previous enzyme kinetics and crystallography results, we know **triclosan**, **PT3**, and **PT5** are rapid reversible inhibitors of InhA and the active-site helix-6 is disordered. To know whether MD simulations also demonstrate a flexible helix-6 conformation, we simulated the **triclosan**, **PT3**, and **PT5** bound complexes started from the closed conformation and measured the RMSD and RMSF. In the 20 ns MD simulation result of InhA:NAD$^+$:**triclosan** complex, the protein was more flexible than the above InhA:NAD$^+$:**PT70** complex, with a higher backbone RMSD value between 0.8 to 1.7 Å. It is noticed that some monomers had larger RMSD (**Figure 2.7A**). The highest RMSD of helix-6 and 7 region reaches ~ 5.8 Å (**Figure 2.7B**). The RMSF plot also suggested that helix-6 was more flexible in the **triclosan** bound complex **(Figure 2.7C)** than in the **PT70** bound complex (**Figures 2.6C**). In addition to **triclosan**, the other two rapid reversible inhibitors **PT3** and **PT5** bound complexes also had more flexible helix-6 and 7 (**Figure 2.8**). Structural analyses of the initial and last frames of trajectories illustrated that helix-6 moved to a semi-closed conformation (**Figure 2.9**).
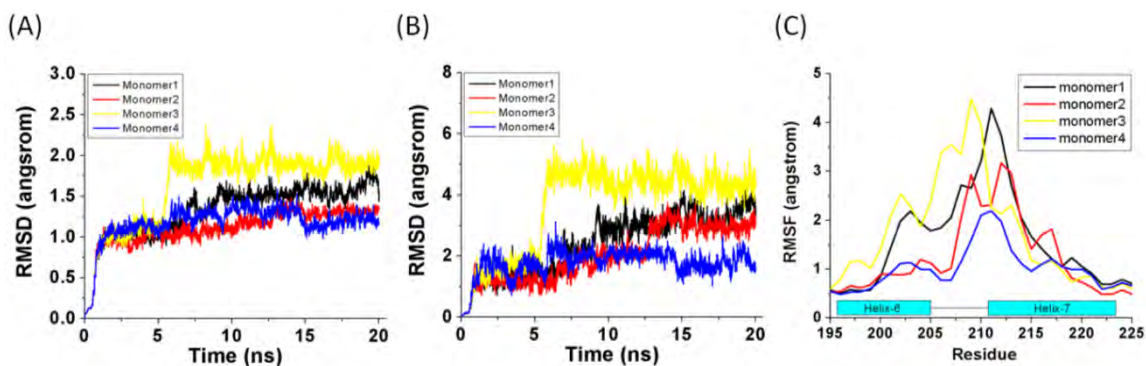
Figure 2.7. RMSD and RMSF of the InhA:NAD$^+$:triclosan complex. (A) Backbone RMSD of the whole protein. (B) Backbone RMSD of the helix-6&7 region. (C) Backbone RMSF of the helix-6&7 region. RMSD was referenced to the closed crystal structure. RMSF was referenced to the average structure.
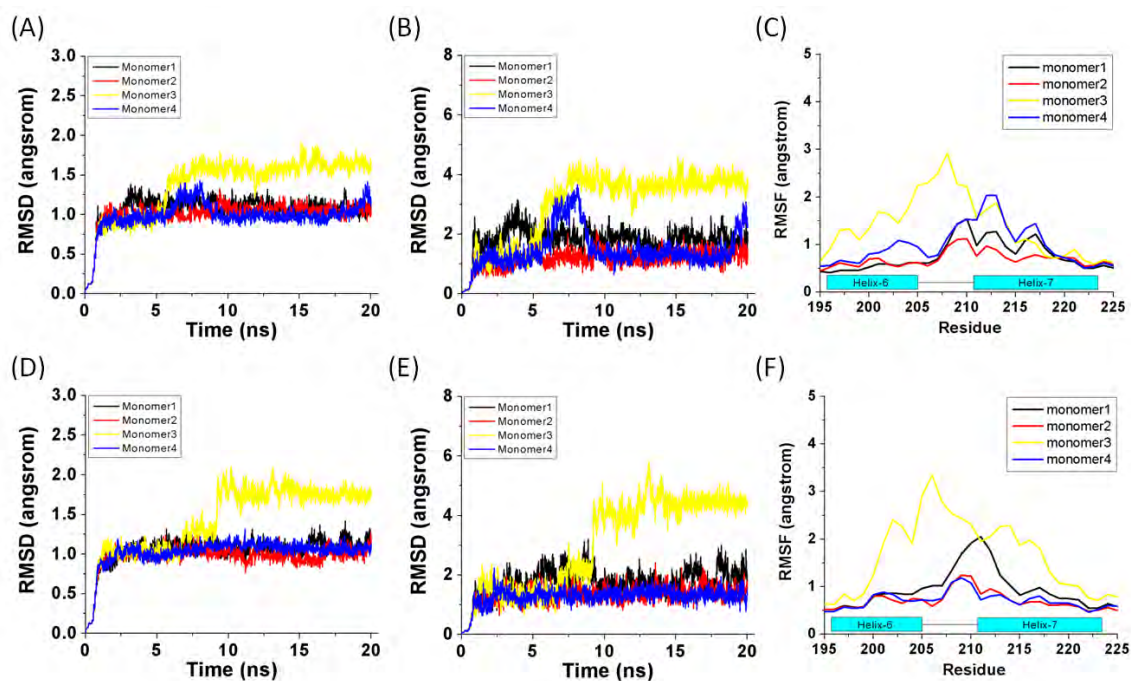
Figure 2.8. RMSD and RMSF of the InhA:NAD$^+$:PT3 (A~C) and InhA:NAD$^+$:PT5 (D~F) complexes. (A) and (D) Backbone RMSD of the whole protein. (B) and (E) Backbone RMSD of the helix-6&7 region. (C) and (F) Backbone RMSF of the helix-6&7 region. RMSD was referenced to the closed crystal structure. RMSF was referenced to the average structure.

46
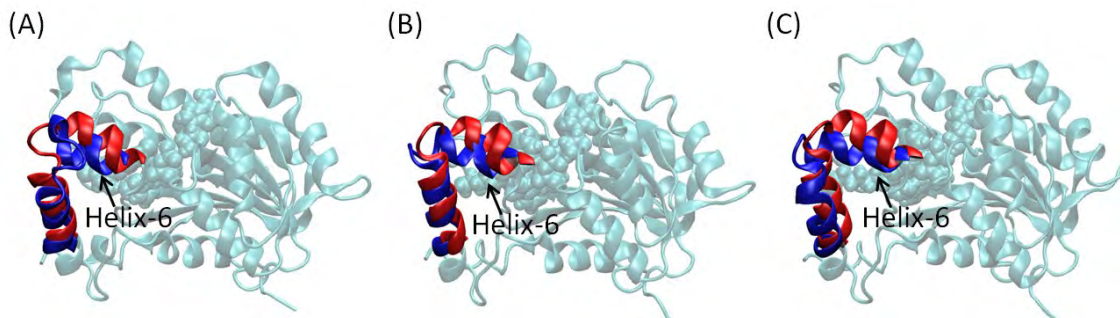
Figure 2.9. The last snapshots of monomer-3 from MD simulations. (A) InhA:NAD$^+$:**triclosan**. (B) InhA:NAD$^+$:**PT3**. (C) InhA:NAD$^+$:**PT5**. The active-site helix-6 moved toward to a more open conformation. First and last snapshopts are colored in red and blue, respectively.

2.3.3   MD simulation started from the open conformation

The open conformation of InhA is observed in the apo, binary, substrate bound, and most inhibitor bound structures (**Table 2.1**). As shown in **Figure 2.10**, the active-site entrance is covered and blocked by helix-6 in the closed conformation; it is unlikely that the inhibitor binds to the closed conformation during the early stage of the enzyme-inhibitor association step. Instead, it is more likely the protein follows the induced-fit mechanism in which the inhibitor bind to the open conformation of InhA, which has a wider active-site entrance to allow an inhibitor to dock in, followed by a conformational change, leading to the closed state. Because the open conformation is likely to be the initial conformation which inhibitors encounter, it would be interesting to know what kind of conformational change will happen to the InhA upon binding of the slow-onset and rapid reversible inhibitors. To know how the open helix-6 and 7 region acts with the inhibitors, we performed same simulations but used the substrate bound InhA as the starting structure. We hypothesized the InhA with rapid reversible inhibitors would

47

have a flexible helix-6 as shown in the crystal structure. For the slow-onset inhibitor **PT70** bound InhA, we hypothesized the helix-6 region would move toward to a more closed conformation.
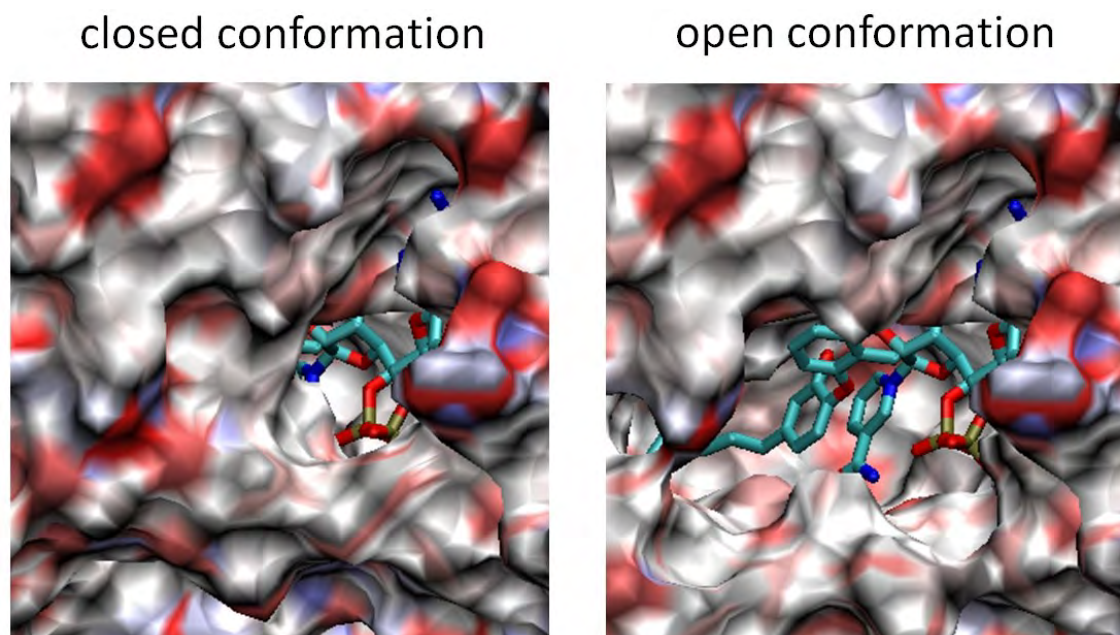
Figure 2.10. Surface representation of the closed and open active-site conformations. The substrate entrance is covered and blocked in the closed conformation (left).

For the slow-onset complex InhA:NAD^+:**PT70**, the average backbone RMSD value (referenced to the open conformation) was ~ 1.3 Å and is slightly higher than the above results started from the closed conformation (**Figures 2.11A and 2.6A**). Not only was RMSD value higher, the RMSF value was also higher than the closed state (~ 3 Å v.s. 1 Å). The high RMSD and RMSF values of the helix-6 and 7 regions suggested helix-6 is flexible (**Figures 2.11B, and 2.6B**). Taking monomer-1 as an example, the superimposed structures of the first and last frames of the trajectory demonstrate that helix-6 moved toward to a more closed conformation (**Figure 2.12**).
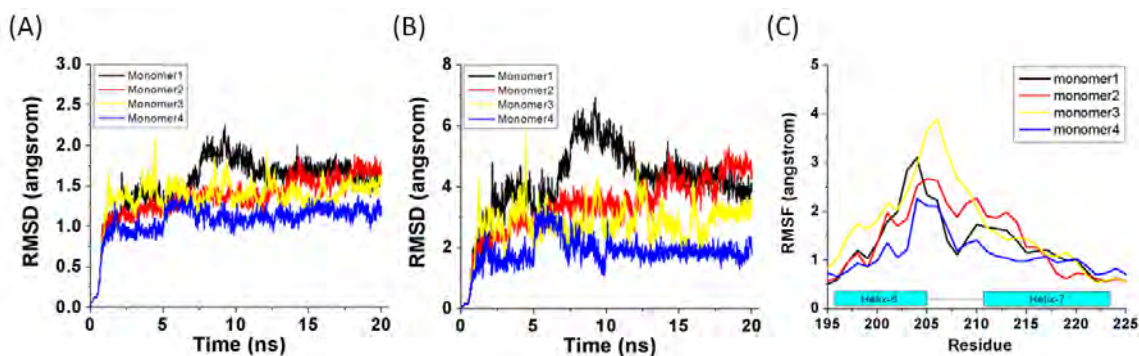
Figure 2.11. RMSD and RMSF of the InhA:NAD^{+}:PT70 complex. (A) Backbone RMSD of the whole protein. (B) Backbone RMSD of the helix-6&7 region. (C) Backbone RMSF of the helix-6&7 region. RMSD was referenced to the open crystal structure. RMSF was referenced to the average structure.
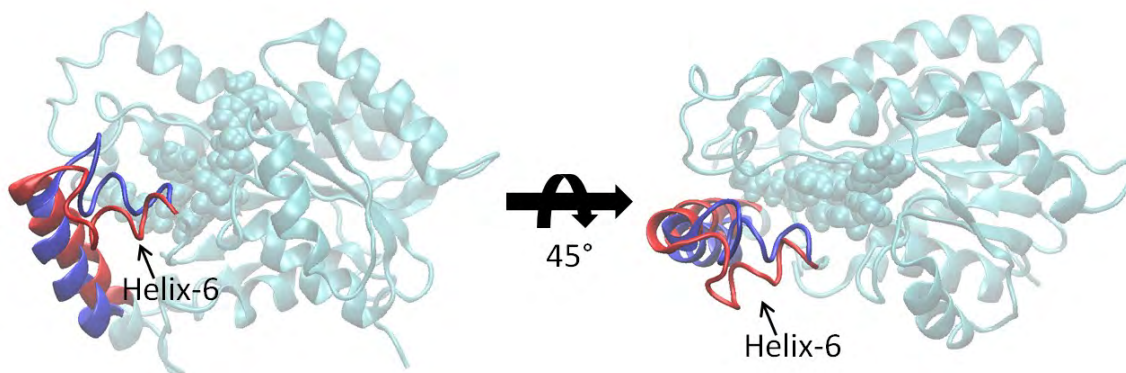
Figure 2.12. The first and last frames of monomer-1 from the InhA:NAD^{+}:PT70 complexes. First and last frames are colored in red and blue, respectively. The helix-6 moves toward to a more closed conformation.

InhA with rapid reversible inhibitors started from the open conformation also has a flexible helix-6. For **triclosan** bound InhA, the backbone RMSD fluctuates at the early stage of simulation, suggesting helix-6 and 7 moved away from the open conformation (**Figures 2.13A, 2.13B**). The high RMSF value also indicates the helix-6 and 7 region is flexible (**Figure 2.13C**).

49

Similar to the InhA:NAD[+]:**PT70** complex, structural comparison of the initial and final snapshots indicated that the helix-6 moved toward to a more closed conformation (**Figure 2.14**). The other two rapid reversible inhibitor complexes InhA:NAD[+]:**PT3** and InhA:NAD[+]:**PT5** also have flexible helix-6 supported by their high RMSD and RMSF values (**Figure 2.15**).
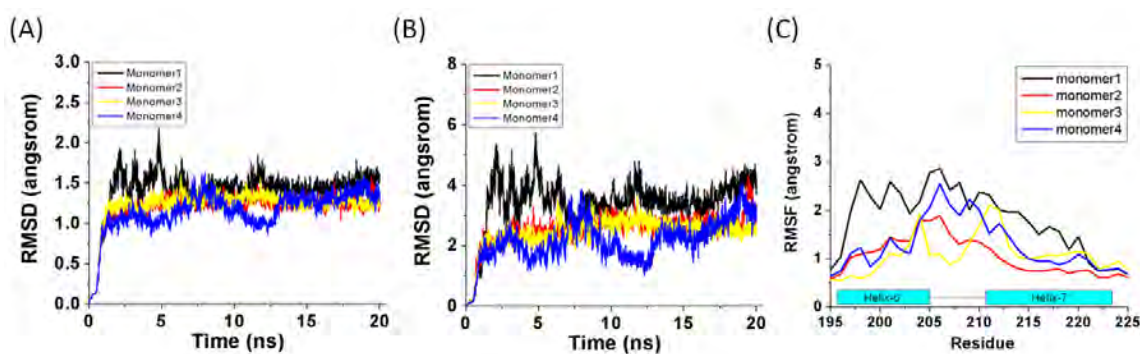
Figure 2.13. RMSD and RMSF of the InhA:NAD[+]:triclosan complex. (A) Backbone RMSD of the whole protein. (B) Backbone RMSD of the helix-6&7 region. (C) Backbone RMSF of the helix-6&7 region. RMSD was referenced to the open crystal structure. RMSF was referenced to the average structure.
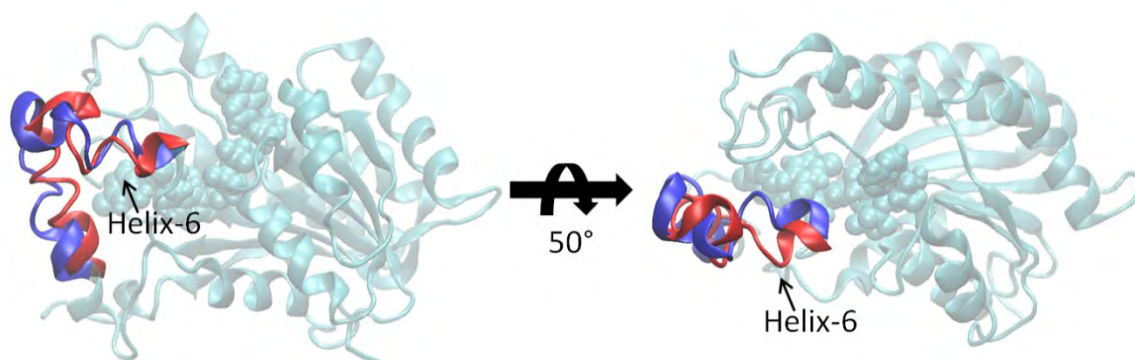
Figure 2.14. The first and last frames of monomer-1 from the InhA:NAD[+]:triclosan complexes. First and last frames are colored in red and blue, respectively. The helix-6 moves toward to a more closed conformation.
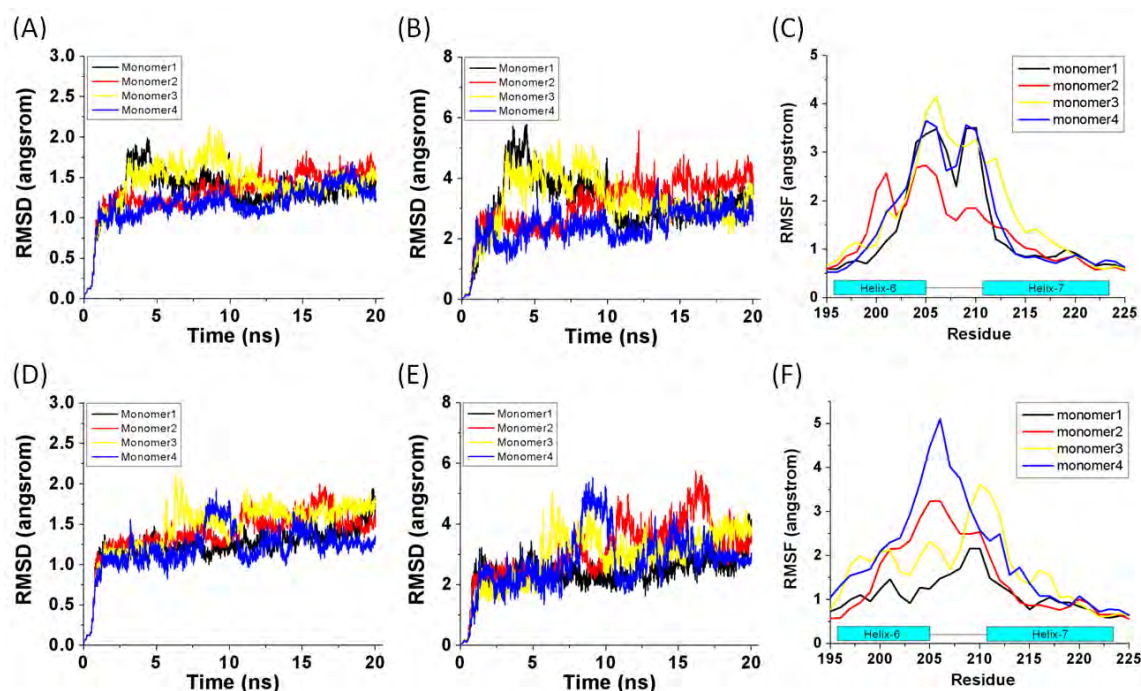
Figure 2.15. RMSD and RMSF of the InhA:NAD⁺:PT3 (A~C) and InhA:NAD⁺:PT5 (D~F) complexes. (A) and (D) Backbone RMSD of the whole protein. (B) and (E) Backbone RMSD of the helix-6&7 region. (C) and (F) Backbone RMSF of the helix-6&7 region. RMSD was referenced to the open crystal structure. RMSF was referenced to the average structure.

2.3.4 Principal components analysis (PCA)

The 20 ns regular (unrestrained) MD simulations revealed the dynamic properties of the helix-6 and 7 region. It is interesting to know the motion of helix-6&7 and quantify the kind of motion that may contribute to the slow-onset inhibition. Although comparing the simulation trajectories, as those shown in **Figures 2.9, 2.12** and **2.14**, already implied the active-site has an open-to-closed or closed-to-open motion, however, it would be better to have a quantitative method to measure this motion. Here, we used principle components analysis (PCA) method to decompose the major motion. PCA is basically an eigenvector-based multivariate analysis: it helps us to find out the major eigenvector (with respect to the largest eigenvalue) of each atom in

the trajectories. For the MD simulations started from the closed conformation, we know the helix-6 had a larger motion in the rapid reversible inhibitors (**triclosan**, **PT3**, and **PT5**) bound complexes. The PCA results demonstrated that the helix-6 region has an open-closed motion, however this motion is not an "ideal" open-closed motion (**Figures 2.16** and **Figure 2.17**). An "ideal" open-closed motion is a direct path between the crystal structures open and closed conformations (**Figure 2.17A**). The motions from the PCA seem to perpendicular to the ideal open-closed motion. One of the possible explanations is that we only performed single simulation for each inhibitor bound complexes, and this would bias the PCA results. Multiple runs will improve the quality of the results. If we assume multiple different runs have these similar results, a possible explanation is that the ideal open-closed path is not an energy favorable path; the helix-6 need to detour to find a low energy path. Examining the residues on the helices 6 and 7, we found there are several hydrophobic and bulky sidechain residues, particularly the sidechain of residues M199, V203, and I215 are packing to each other, forming steric clashes (**Figure 2.17B**). It is possible the residues bump to each other during the active-site open-closed motion, so that the helix-6 has to move to another direction to avoid the steric clashes.
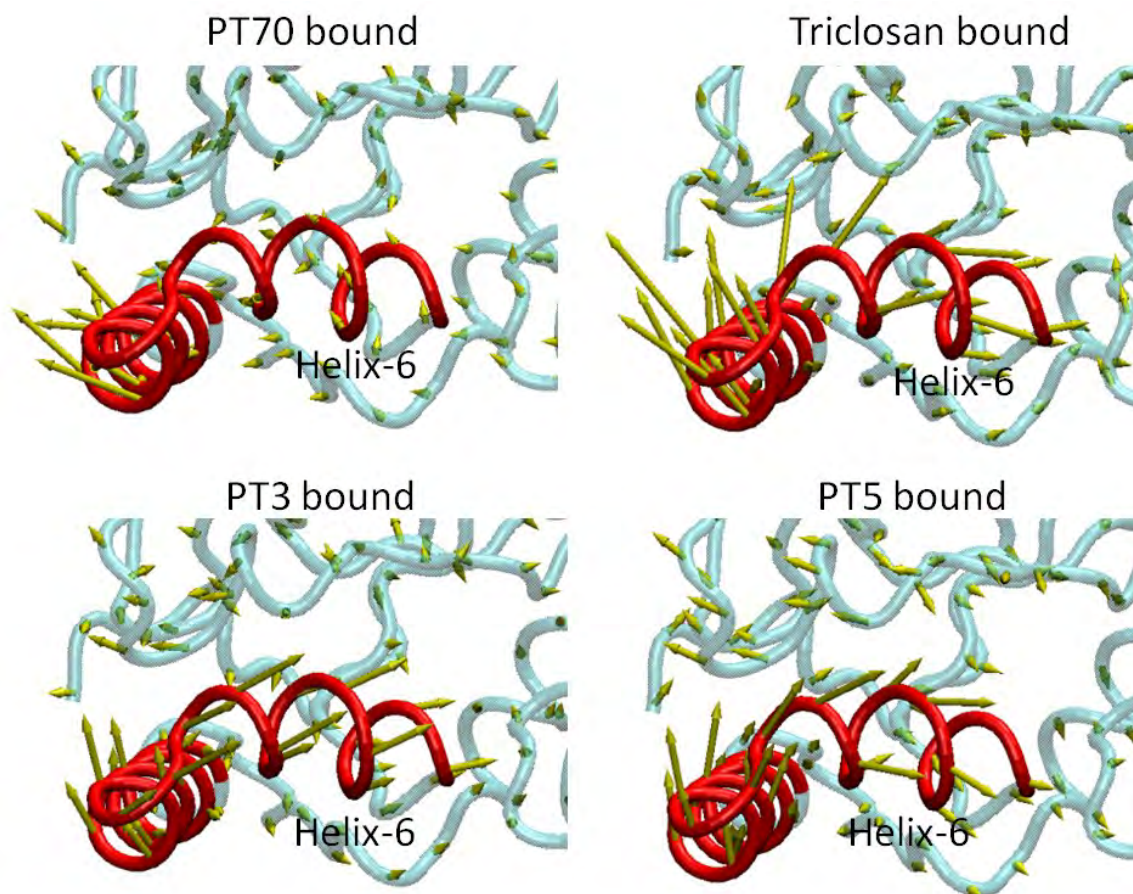
Figure 2.16. The PCA motion vectors calculated based on the simulations started from the closed conformation. Rapid reversible inhibitors bound complexes seem to have a closed-to-open motion. The arrows only show one direction from the starting structure. Figures were generated using the VMD plugin NMWiz (*136*).
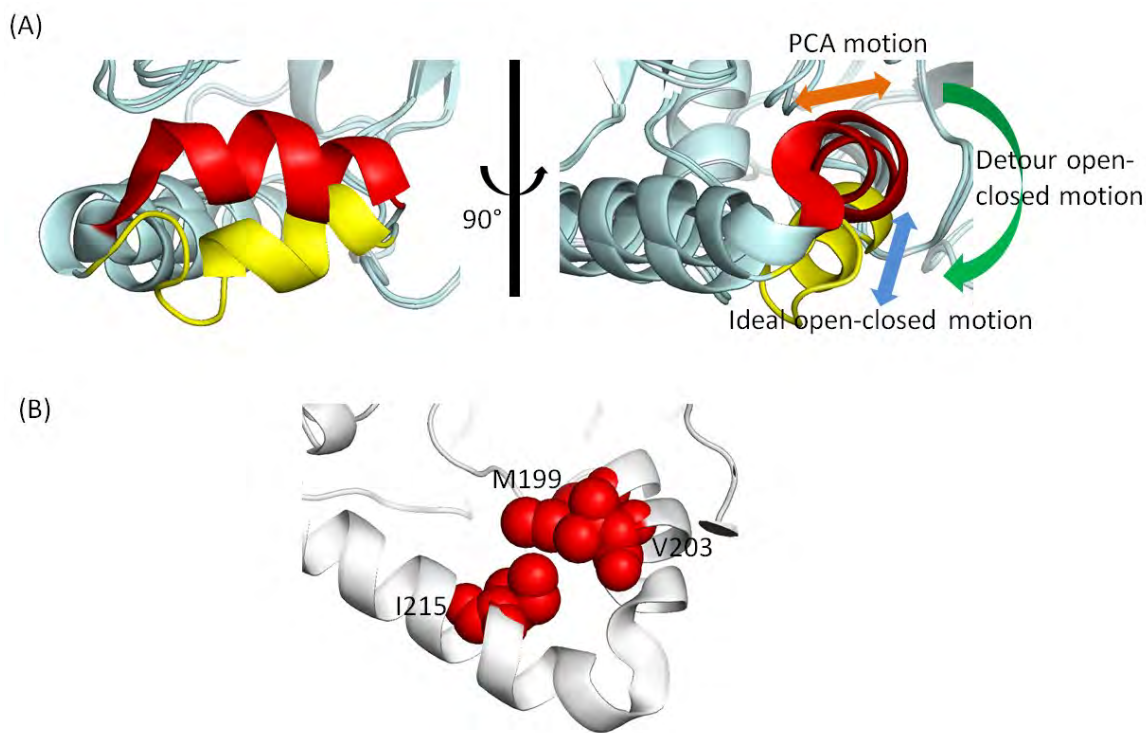
Figure 2.17. The (A) The PCA, ideal, and detour open-closed motions and (B) hydrophobic residues on helices 6 and 7.

For the simulations started from the open conformation, all inhibitor bound complexes had a flexible helix-6. As shown in **Figures 2.12, 2.14**, both rapid reversible inhibitors (**triclosan**, **PT3** and **PT5**) and slow-onset inhibitor (**PT70**) bound complexes have a semi-open active-site conformation after 20 ns simulation. From the PCA calculation, the active-site helix-6 seems to also have a detour open-closed motion: the PCA vectors are not parallel the open-closed path (**Figure 2.18**).
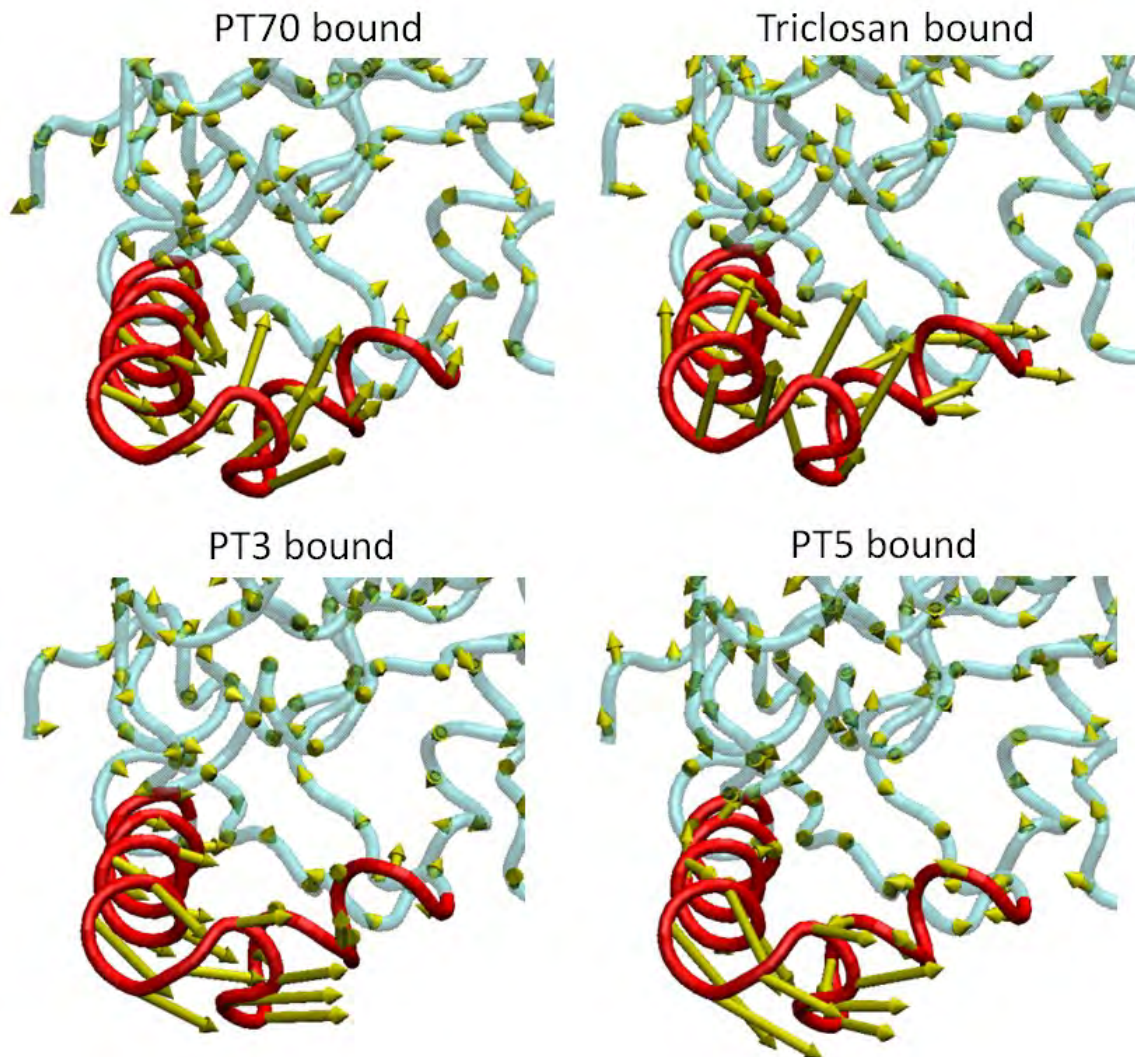
Figure 2.18. The PCA motion vectors calculated based on the simulations started from the closed conformation. Both slow-onset and rapid reversible inhibitors bound complex seem to have an open-to-closed motion. The arrows only show one direction from the starting structure. Figures were generated using the VMD plugin NMWiz (*136*).

2.3.5   Inhibitor unbinding calculation

Because the active-site entrance is blocked in the closed conformation (**Figure 2.10**), an inhibitor is unlikely to bind the closed conformation at the initial enzyme-inhibitor association step. Instead of this, an inhibitor is more likely to bind the open conformation, followed by a

conformational change, leading to the closed conformation. The above MD simulations imply the active-site has an open-closed motion and the open-closed motion is likely to be the structural isomerization step in the slow-onset inhibition mechanism. However, it is also possible the initial inhibitor binding step is the rate-limiting step where experiments measure; because the inhibitor is competing the active-site pocket with the substrate, the slow-onset inhibition property could happen if an inhibitor moves slowly from the active-site gate to the pocket. If this is the case, the inhibitor binding step should encounter energy barriers. To understand the inhibitor association process is a challenging task with current MD simulation techniques. Although a long timescale MD simulation (μs and ms) may be able to capture the event of an inhibitor finding the binding site (*137*), it is still a computationally expensive task. Alternatively, the inhibitor unbinding calculation is a more feasible approach and could provide information about the reverse binding step. There have been several studies that applied steered MD to generate the ligand unbinding process, and the calculated binding free energies have very good agreements with the experimental results (*59-65*). The unbinding calculation provides not only the free energy difference (ΔG) between the bound and unbound states, but also the energy profile along the unbinding path. The energy profile can help us to discover energy barrier during the inhibitor unbinding step.

To understand whether the inhibitor dissociating process has bumpy barrier, we applied the steered MD for the **PT70** bound complex. Because the open conformation of InhA is likely to be the EI state in the two-step inhibitor binding model (**Scheme 1.1B**), we simulated the inhibitor unbinding process started from the open conformation. We used steered MD method to gradually push off the **PT70** from the active-site based by increasing the distance between the center of mass of yellow atoms (receptor) and the ligand (**Figure 2.19**), followed by the umbrella

sampling/WHAM methods to obtain the free energy along this unbinding path. We initially performed 1 ns umbrella sampling and found the free energies were not converged, and then we extended the simulation to 4 ns.
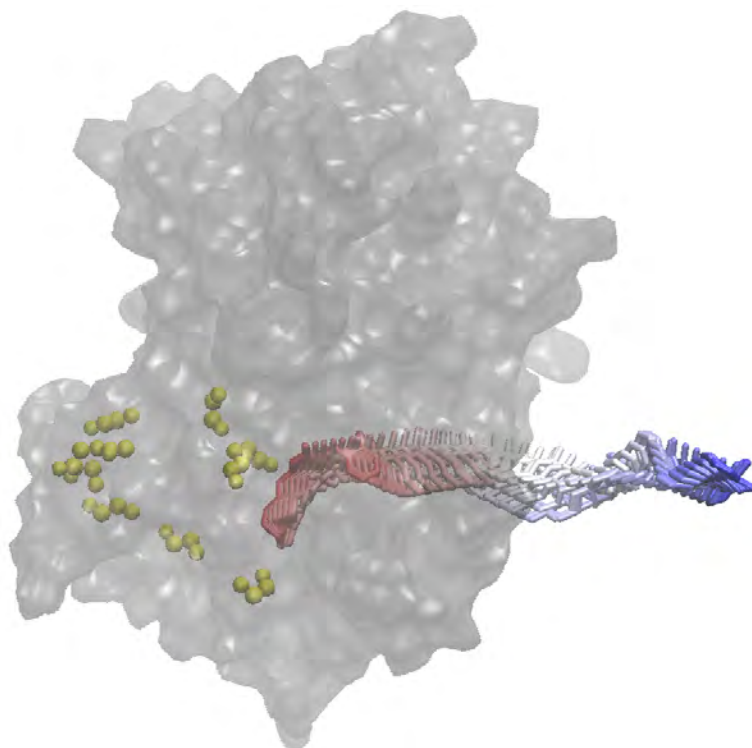
Figure 2.19. The unbinding path of PT70. The inhibitor was gradually pushed out the active-site by increasing the distance between inhibitor and some receptor atoms (yellow).

The free energy difference between the lowest point (bound state) and the plateau state (unbound state) is ~ 9 kcal/mol (**Figure 2.20, 4 ns result**) and this value is close to the experimental result (-11.05 kcal/mol; based on $K_i = 7.8$ nM (*13*)). The 2 kcal/mol difference may be due to the translational entropy which was not accounted for in this calculation. A review paper has shown the translational and rotational entropy for small ligands binding to proteins is estimated to be 4 - 5 kcal/mol (*138*).
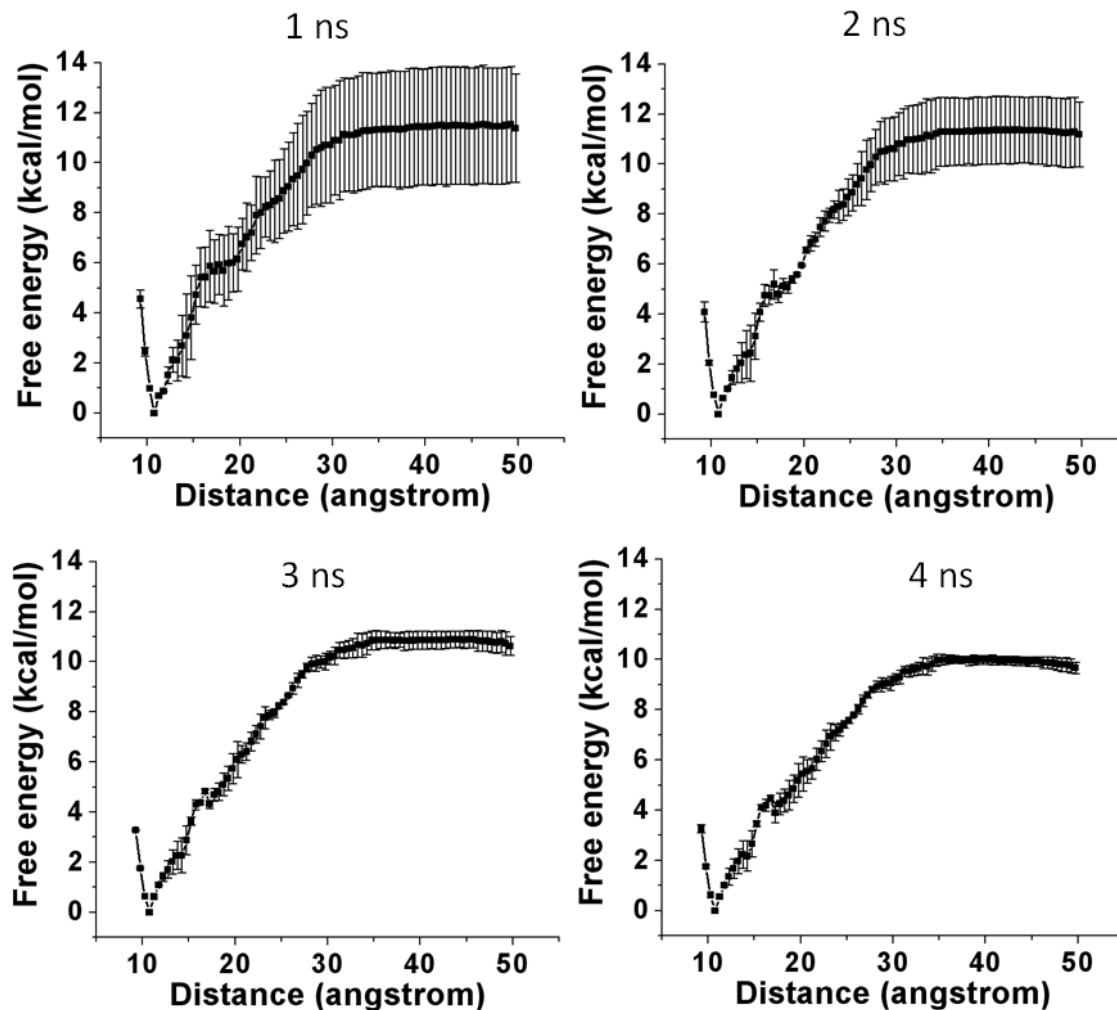
Figure 2.20. The energy profile of PT70 dissociation from the active-site. There is no significant bump barrier during the inhibitor unbinding path. The errorbar is the difference between two independent runs.

There is no significant bumpy barrier along the binding path. Another observation from the unbinding simulation is that the free energy increased gradually during the inhibitor dissociation process, reaching a plateau state when the inhibitor is completely out of the active-site. It is noted that the energy of the binding path is smooth without having bumpy barriers. This result seems to suggest that the EI to E+I process is not the key step causing slow-onset inhibition for **PT70**.

As mentioned previously, the inhibitor unbinding may require long timescale calculations to get a converged result. In our case, the first 1 ns and 2 ns results are not converged, but seems to converge at 4 ns (**Figure 2.20**). Because the inhibitor unbinding process is likely to have multiple paths, particularly a radial pattern, using one reaction coordinate is not a good choice; inhibitor can move to different spots in the other orthogonal space. Therefore, a long timescale simulation is commonly used to limit the sampling problem. Introduction another orthogonal reaction coordinate(s) could be able to limit the converged problem as well as to improve the quality of energy profile. Since we only use one reaction coordinate, in the future it would be better to introduce more precise reaction coordinate to get a more accurate inhibitor unbinding energy profile.

2.3.6   The hypothesis of open-to-closed structural changes for slow-onset inhibition

Based on the results of X-ray crystal structures and MD simulations, we proposed the open-to-closed conformational changes mechanism for the slow-onset inhibition of InhA. The summary of crystallographic and MD simulations results are listed below.

1. The InhA:NAD$^+$ complex and the substrate bound (InhA:NAD$^+$:**C16-NAC**) complex have open active-site conformations, while the slow-onset inhibitor bound InhA has a closed active-site conformation. Because the active-site entrance is blocked in the closed active-site, it is unlikely that the inhibitor initially binds to the closed InhA. Instead, it is more likely the inhibitor binds to the open conformation, which is the substrate bound conformation. These observations imply the open conformation is the initial EI state.

2. All rapid reversible and slow-onset inhibitor bound InhA complexes started from the open conformation led to a semi-closed active-site conformation after 20 ns MD simulations. From the analyses of the MD simulation trajectories, the helix-6 and 7 region has an open-to-closed motion, suggesting the enzyme tends to close its active-site in the presence of inhibitor.

3. When simulation started from closed conformation, InhA maintains a stable and closed active-site when bound with slow-onset inhibitor, but has flexible helix-6 with rapid reversible inhibitors bound. From the crystal structure observation, the active-site helix-6 is ordered with **PT70**, but disordered with rapid reversible inhibitors **triclosan**, **PT3** and **PT5**. MD simulations demonstrated the **PT70** bound InhA has a relatively more stable helix-6 conformation than rapid reversible bound InhA. This suggests that the closed conformation is likely to be the final EI* state and only the slow-onset inhibitor can stably reside in this type of active-site.

4. The inhibitor unbinding calculation showed that there is no significant bump barrier during the ligand binding process. This seems to rule out the initial inhibitor binding step is the rate-limiting step where kinetics studies measured.

Based on all the above preliminary MD simulation results, we proposed the open and closed conformations represent the initial EI state and final EI* state, respectively and the induced-fit conformational change from the open to closed states is highly like to explain the two-step slow-onset binding mechanism (**Figure 2.21**). Because the MD simulations we show in this chapter are preliminary observations from only four inhibitors bound complexes, the summary we draw here may not true for other inhibitors. However, since our purpose here is to get an idea of what kind of motions may correlate with the slow-onset inhibition, we found the

open-closed conformational change is highly possible to be the isomerization step. A more detailed study of the open-closed conformational change is presented in chapter four. In chapter four, we applied pathway sampling approach along the open-closed path for different inhibitors bound complexes and obtained the energy profiles during the active-site closing step.

$$E+I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} EI \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} EI^*$$
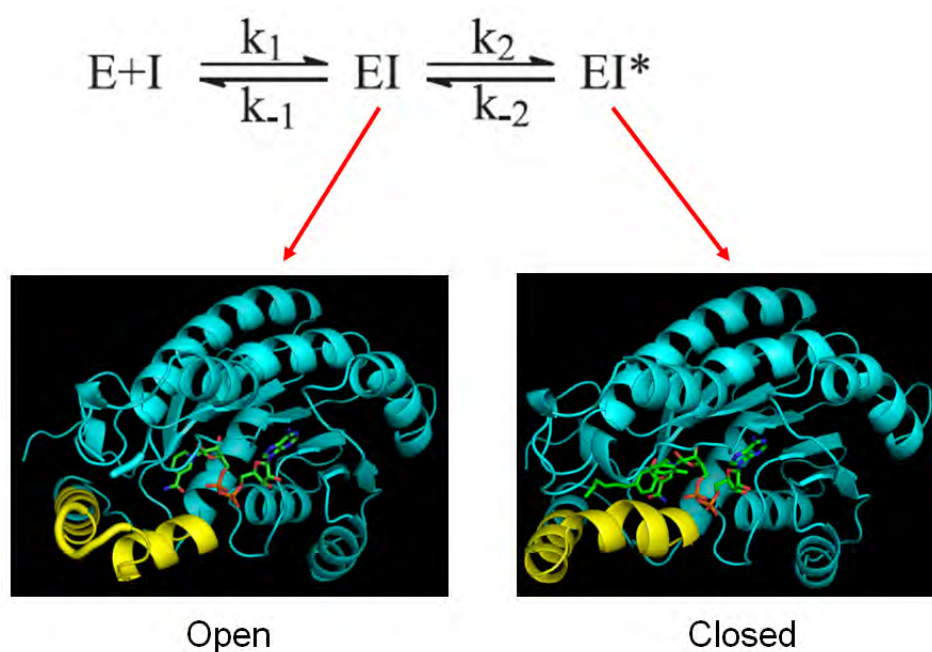
Open   Closed

Figure 2.21. The open-closed conformational changes model for the slow-onset inhibition of InhA. In this model, open and closed conformations represent EI and EI* states, respectively.

2.4 Summary

To understand why certain inhibitors have ordered helix-6 while others do not, a systematic characterization of all InhA structures was done first, followed by a series of MD simulations. After analyzing all the available InhA structures, we found the ordered active-site loop can be further characterized as "closed" or "open". The MD simulation started from the

61

open conformation had a flexible helix-6 in both slow-onset and rapid reversible inhibitors bound InhA complexes. Conversely, when the MD simulations were started from the closed conformation, the helix-6 region remained stable only with slow-onset inhibitors. Further principal component analysis of the trajectories indicated that the open active-site has a motion of closing active-site. On the contrary, the closed active-site tends to keep the active-site closed only when the slow-onset inhibitor is bound. The crystallographic and MD studies demonstrate that the closed conformation of InhA is the stable state where slow-onset inhibitors reside, implying the closed conformation is the EI* state in the two-step binding model. On the other hand, the open conformation is likely to be the EI state where an inhibitor initially binds to. Based on these results, we proposed the open-to-closed conformational changes mechanism for the slow-onset inhibition of InhA, where the open and closed states represent EI and EI* states, respectively. More detailed study of the open-closed conformational change is shown in chapter four.

Chapter Three

How robustly does the MM/PBSA approach reproduce the binding free energy in the InhA system?

3.1 Introduction

Computation plays an important role in assisting lead optimization by predicting the interaction between the inhibitor and the target. As mentioned in chapter one, there are several different computational approaches people commonly use in calculating protein-ligand binding free energy. One can decide which approach based on the accuracy and computationally efficiency. For example, docking approach provides a fast but less accurate result, while the free energy perturbation (FEP) and thermodynamics integration (TI) method give most accurate results but are computationally expensive. The molecular mechanics/Poisson-Boltzmann surface area (MM/PBSA) approach is a post-processing approach that requires less computing power than the FEP and TI methods but the prediction is less reliable than FEP and TI method (*139*). Although the MM/PBSA is less accurate than the FEP and TI methods, the computationally efficient and acceptable prediction make it be widely used in the studies of protein-ligand and protein-protein interactions (*44*). A further advantage of MM/PBSA approach is that the interaction energy can be decomposed into per-residue level, providing additional insights into the energetic of the investigated system.

To know whether the MD simulation can robustly predict the binding free energy and further use these results to better understand the mechanism of slow-onset inhibition in the InhA system, 10 inhibitors (5 rapid reversible and 5 slow-onset inhibitors) with known experimental

binding affinity ranging from μM to pM were selected for the MM/PBSA simulations. Since the active-site conformation of InhA is characterized as closed or open, we performed the MD simulations for both closed and conformations.

The binding free energy, calculated from the MM/PBSA approach, had a good correlation with the experimental results, suggesting the MD simulation reproduces the trend of inhibitor binding in the InhA system. The calculated binding free energies are in consistent with the crystal structure observations, demonstrating that the slow-onset inhibitors favor the closed conformation while the rapid reversible inhibitors do not have a strong preference for the closed or open conformation. Further energy decomposition analyses show that the van der Wall (vdW) interaction is the most important interaction between inhibitor and enzyme. Per-residue energy decompositions indicate that residues G96, F149, Y158, M161 and V203 contribute most to the inhibitor binding.
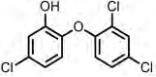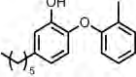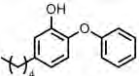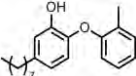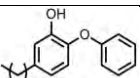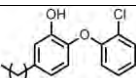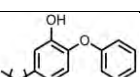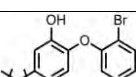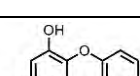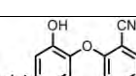
3.2 Simulation details

3.2.1   Initial structures

The InhA:NAD$^+$:**PT70** and InhA:NAD$^+$:**C16-NAC** tetramer structures (PDB ID: 2X23 (*13*) and 1BVR (*11*), respectively) were used to build the starting structures of closed and open states. For those inhibitors that do not have crystal structures, the initial poses of inhibitor were obtained from molecular docking with DOCK 6.3 package (*140*). The sphere and grid files were generated from SPHGEN and GRID modules of DOCK 6.3, respectively. The grid file for the grid-based energy scoring had a space of 0.3 Å. Default parameters were used in the flexible

docking (*140, 141*). We assumed that the analogues occupy similar position as **PT70**, thus results with a lowest RMSD value in the diphenyl ether moiety were chosen as the initial structures.

Table 3.1. List of inhibitors studied in this chapter.

Name	Structure	Slow-onset inhibitor	K_i (nM)		Name	Structure	Slow-onset inhibitor	K_i^* (nM)
Triclosan		No	220 ± 20		**PT70**		Yes	0.04 ± 0.01
PT3		No	11.8 ± 4.5		**PT82**		Yes	0.01 ± 0.00
PT4		No	9.4 ± 0.5		**PT91**		Yes	0.96 ± 0.13
PT5		No	1.1 ± 0.2		**PT92**		Yes	0.20 ± 0.05
PT52		No	5000		**PT119**		Yes	2.14 ± 0.35

Experimental K_i values were taken from (*13, 142*).

3.2.2   Setup of Molecular Dynamics Simulation

AMBER ff99SB (*26*) and GAFF (*115*) force field parameters were assigned to the protein and inhibitor, respectively. The partial atomic charges of each inhibitor were derived from the RESP approach (*46, 116*). The RESP charges were computed using Gaussian98 (*117*) with the HF/6-31G* basis set. The force field of cofactor NAD^+ was taken from other studies (*118, 119*). Each complex was solvated in a truncated octahedral TIP3P (*120*) water box with a minimum distance of 14 Å between the water box edge and solute, resulting in ~ 80,000 atoms in

total. The MD simulations were run at 300 K under constant 1 atm pressure and periodic boundary conditions. 20 ns MD simulation was performed for each compound. The first 1 ns was the equilibration procedure, followed by a 19 ns unrestrained production run. The first step of equilibration was 10,000 steps of steepest descent minimization with 100 kcal mol^{-1} Å^{-2} restraints on all atoms except water molecular and hydrogen atoms. The second step was heating the system from 100 to 300 K at constant volume over 100 ps with 100 kcal mol^{-1} Å^{-2} restraints on non-water and non-hydrogen atoms, followed by 100 ps with the same restraints at constant 300 K temperature and 1 atm pressure. The third step was 250 ps MD with restraint weight of 10 kcal mol^{-1} Å^{-2} on the non-water and non-hydrogen atoms at constant 300 K temperature and 1 atm pressure. The following steps only restrained the backbone atoms and gradually reduced the restraints weight from 10 to 0.1 kcal mol^{-1} Å^{-2} at constant 300 K and 1 atm. This was carried out by 100 ps with 10 kcal mol^{-1} Å^{-2} restraint, followed by 100 ps with 1 kcal mol^{-1} Å^{-2} restraint, and 100 ps with 0.1 kcal mol^{-1} Å^{-2} restraint. The last step of equilibration was 250 ps of unrestrained MD.

3.2.3   Molecular Mechanics/Poisson-Boltzmann surface (MM/PBSA)

The binding free energy was calculated from the 19 ns production simulation trajectories (1900 frames) using the mmpbsa.py program (*34, 143*). To reduce the noise and cancel the errors, single trajectory approach was used (*44*). The dielectric constant of 78.5 and 1 were assigned to exterior and interior, respectively. The conformational entropy was estimated using normal mode analysis (*144, 145*). Because the normal mode analysis is computationally expensive, to reduce the computing time we skipped every 10 frames. Thus, 190 frames were calculated.

66

3.2.4   Energy decomposition

The energy decomposition was calculated using MM-GBSA approach. The non-bonded energy terms (vdW and electrostatic) between inhibitor and receptor were decomposed into per-residue level. This was done by using the SANDER program where idecomp flag was set to 2.

3.3 Results and Discussion

To better understand whether the MD simulation can robustly predict the binding affinity in the InhA system, five rapid reversible and five slow-onset inhibitors with known experimental binding affinity were selected. We performed 20 ns MD simulations for each inhibitor bound to the biological assembly tetramer closed and open conformations, and used MM/PBSA approach to calculate the binding free energies. We speculated that, for the slow-onset inhibitors, since the slow-onset inhibitors bound crystal structures have a closed active-site conformation, the calculated binding free energies in the closed conformation would have more favorable values than those obtained in the open conformation. Moreover, as mentioned in chapter two, we speculated the closed state is likely to be the EI* state, so we hypothesized that the experimental binding affinities would have better correlation with those calculated from the closed state than those from the open state. For the rapid reversible inhibitors, because the crystal structures (PDB ID: 2B35, 2B36, 2B37) have missing electron density in the helix-6 region, we speculated that the rapid reversible inhibitors will not have a strong preference for the closed and open conformations.

3.3.1 MD simulation stability started from the closed conformation

To estimate whether the MD simulations were in a stable stage, structural properties were monitored during the course of the trajectories. Because the slow-onset inhibitor **PT70** bound crystal structure has a closed conformation, we speculated that the other slow-onset inhibitors bound complexes started from the closed conformation would also maintain a stable closed conformation. On the contrary, since the rapid reversible inhibitors bound complexes have a disordered helix-6, the rapid reversible inhibitors bound complexes would be more flexible.

As expected, in the presence of slow-onset inhibitors, the receptor was stable during the 20 ns course of MD simulation. From the measurement of backbone RMSD of the whole protein, it showed that the RMSD was always below 1.3 Å (**Figure 3.1**). Taking **PT70** bound complex as an example, all monomers had RMSD values ~1.0 Å during the 20 ns simulation, suggesting the receptor maintained the initial closed conformation (**Figure 3.1A**). As for other slow-onset complexes, the RMSDs were slightly higher than the InhA:NAD$^+$:**PT70** complex, but the RMSD values were always below 1.3 Å. The slightly higher RMSDs could be due to the starting structure is taken from the **PT70** bound complex, so that other inhibitor bound complexes required some structural relaxations. Overall, the RMSD measurements demonstrate that the receptor remained in a stable state in the presence of slow-onset inhibitors.
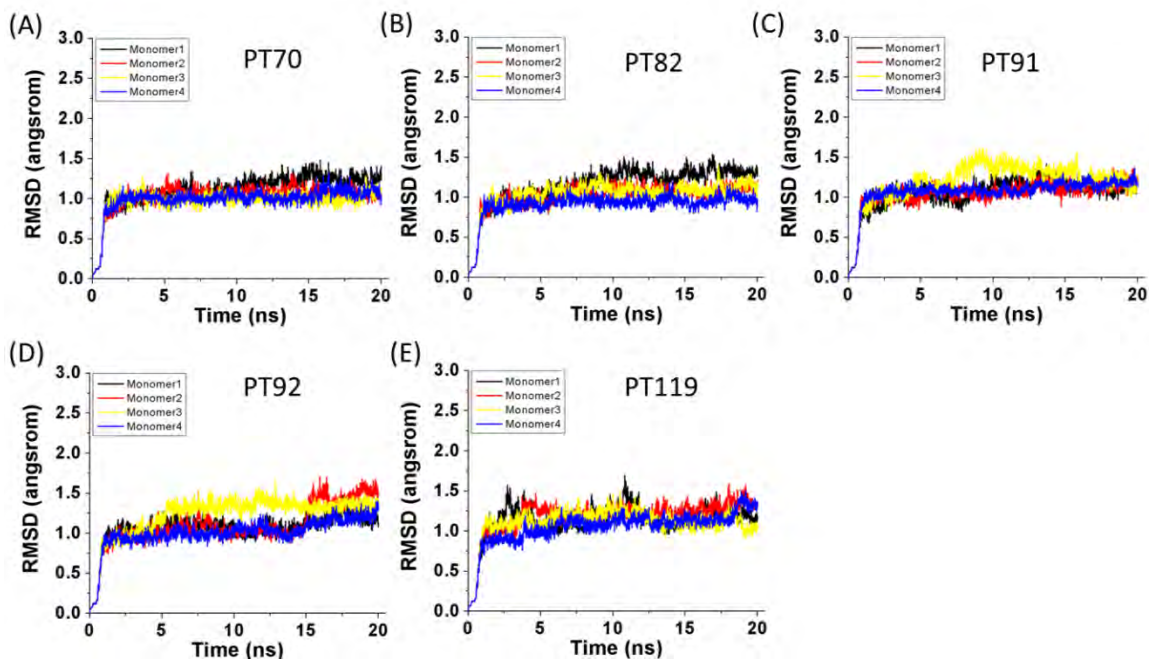
Figure 3.1. InhA backbone RMSD bound with different slow-onset inhibitors. (A) **PT70**. (B) **PT82**. (C) **PT91**. (D) **PT92**. (E) **PT119**. These simulations were started from the closed conformation. RMSD was referenced to the initial closed conformation.

For the rapid reversible inhibition complexes, because the crystal structures have a disordered helix-6, we speculated that the helix-6 would not maintain a stable closed conformation. Based on the RMSD measurements, some monomers had higher RMSD values (**Figure 3.2**). For example, the RMSD of monomer-3 went up to 2.4 Å in the InhA:NAD$^+$:**triclosan** complex (**Figure 3.2A**). Similar results were found in the **PT3**, **PT5**, and **PT52** bound complexes, showing some monomers deviated from the starting closed conformation (**Figures 3.2B ~ 3.2E**). Because these were only 20 ns MD simulations, it is possible that other monomers would also behave the same if the simulations run longer.
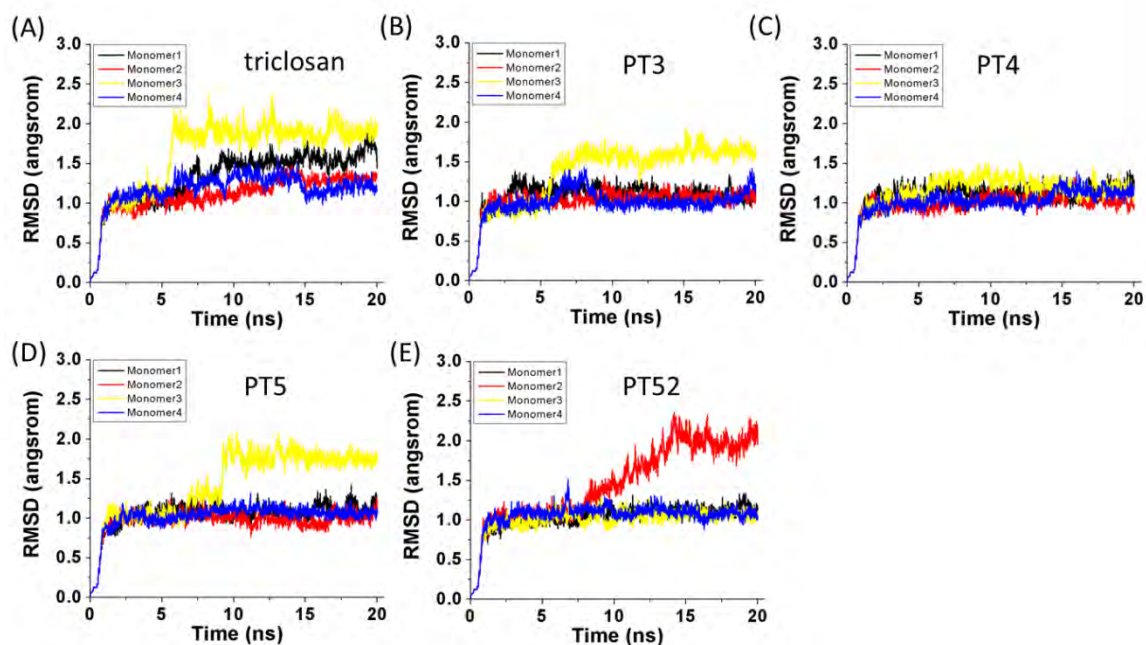
Figure 3.2. InhA backbone RMSD bound with different rapid reversible inhibitors. (A) **triclosan**. (B) **PT3**. (C) **PT4**. (D) **PT5**. (E) **PT52**. These simulations were started from the closed conformation. RMSD was referenced to the initial closed conformation.

3.3.2 MD simulation stability started from the open conformation

To gauge whether the receptor were in a stable stage, RMSD were monitored during the course of the trajectories. Because the slow-onset inhibitor bound InhA have a closed active-site conformation, we speculated that the active-site will move away from the open conformation when the simulations started from the open conformation. For the rapid reversible inhibitors, since the crystal structures have a flexible helix-6, we speculated the RMSD values would fluctuate during the course of the trajectories.

The MD simulation results showed, as expected, that the receptor deviated from the starting open conformation. For the slow-onset inhibitor bound complexes, the average RMSD was ~ 1.7 Å (**Figure 3.3**). Not like the results started from the closed conformation, the RMSD

70

fluctuates between 1.0 ~ 2.5 Å during the 20 ns simulation. Similar flexible phenomena were found in the rapid reversible inhibitor bound complexes, exhibiting a higher RMSD values than those results started from the closed conformation (**Figure 3.4**).

In summary of these measurements of protein backbone RMSD, InhA had a relative higher RMSD values when the MD simulations started from the open conformation regardless with slow-onset or rapid reversible inhibitors. On the contrary, when simulations started from the closed conformation, receptor was more stable in the presence of slow-onset inhibitors.
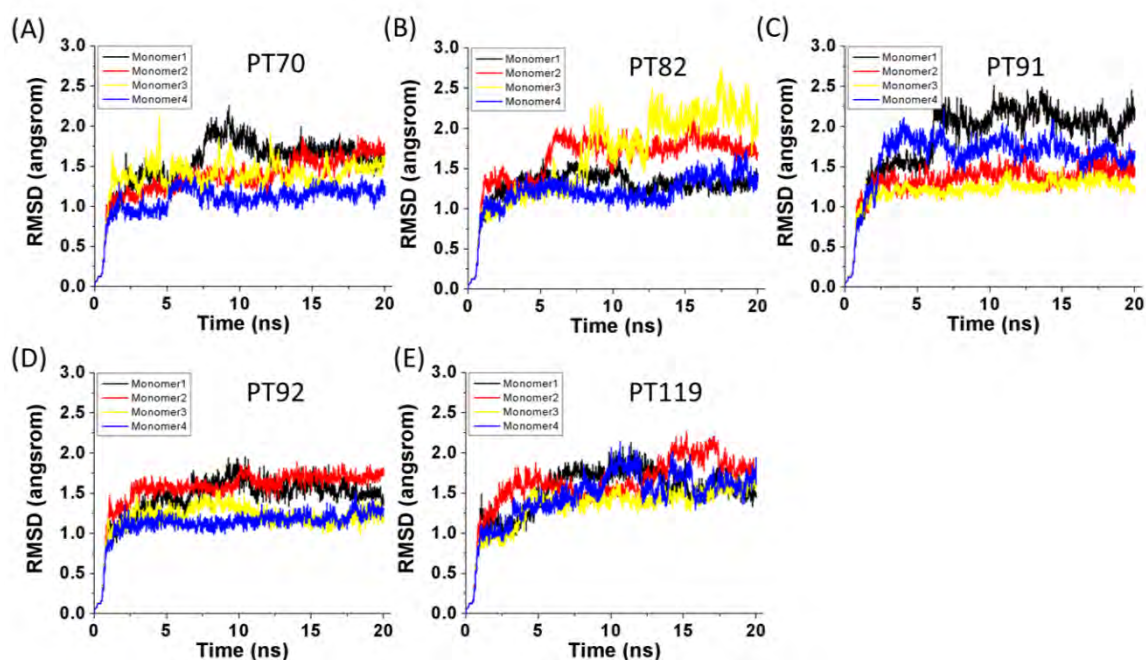
Figure 3.3. InhA backbone RMSD bound with different slow-onset inhibitors. (A) **PT70**. (B) **PT82**. (C) **PT91**. (D) **PT92**. (E) **PT119**. These simulations were started from the open conformation. These simulations were started from the open conformation. RMSD was referenced to the initial open conformation.
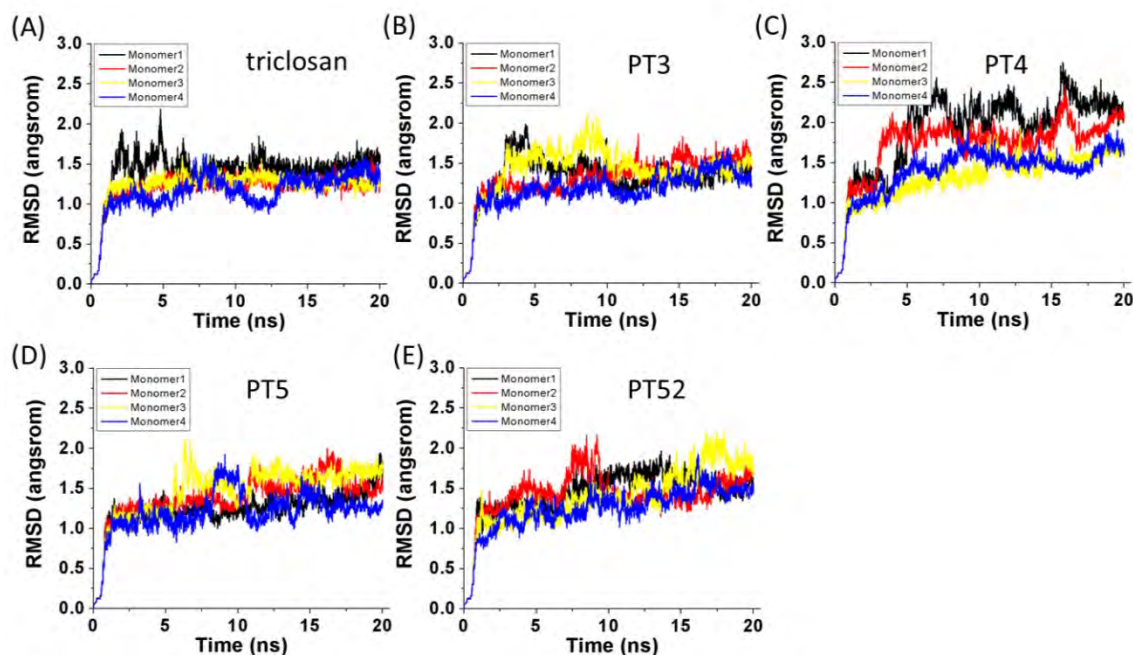
Figure 3.4. InhA backbone RMSD bound with different rapid reversible inhibitors. (A) **triclosan**. (B) **PT3**. (C) **PT4**. (D) **PT5**. (E) **PT52**. These simulations were started from the open conformation. These simulations were started from the open conformation. RMSD was referenced to the initial open conformation.

3.3.3 MM/PBSA binding free energy calculated from the closed and open states

To know whether MD simulation can robustly predict the binding affinity in the InhA system as well as to get a better understanding of what kind of non-bonding interaction (van der Waals or electrostatics) controls the inhibitor binding, we used MM/PBSA approach to calculate the binding free energy. Because the MM/PBSA method is a post-processing calculation approach, the quality of the binding free energy value depends on the resulted trajectories. From the above RMSD measurements, it has already shown some receptor-ligand complexes were very flexible. It is possible that the simulation length is not enough. Studies have shown that longer MD simulation is not always necessary to achieve better prediction for the MM/PBSA approach; particularly a 5 ns MD simulation is enough (*44, 45*). Moreover, the effective

sampling is mostly coupled with the accuracy of force field not the length of MD simulation. Therefore, when the single trajectory protocol is used, long timescale MD simulations do not necessarily lead to accurate prediction (*44, 45*). For our study, we used single trajectory protocol; we think these 20 ns simulations are sufficient to qualitatively answer the question we asked here. More reliable free energy calculation approaches, like umbrella sampling and thermodynamic integration calculations are presented in chapter four.

We first examined the reliability of these calculated binding free energy values. A common way to judge the calculated binding free energy is to compare with the experimental binding free energy in terms of comparing the correlation. Because the closed state is likely to be the EI* state (as mentioned in chapter two), we speculated the experimental binding free energy values would have better correlation with those calculated from the closed state than those from the open state.

As expected, the free energies calculated from the closed conformation had a good correlation with experimental results. **Table 3.2** lists the van der Waals, electrostatics, solvation energy, entropy, and overall free energy calculated from the MM/PBSA approach. By plotting the experimental binding free energy versus the overall calculated binding free energies, we found the correlation coefficient (R^2) are promising (**Figure 3.5**). The values of correlation coefficient (R^2) are 0.96 and 0.81 for rapid reversible and slow-onset inhibitors, respectively. When combining all ten inhibitors results, the overall R^2 value is 0.96 (**Figure 3.5C**), suggesting the MM/PBSA approach reproduces a good trend of the binding affinity for the InhA system.

One thing to note is that the slope deviates from one. Other studies had showed the effects of entropy calculation could drive the slope away from one (*44, 146, 147*). In addition,

the fixed partial charge model also can cause the slope deviating from one (*148*). The polarizable

force fields may reduce the imperfection of fixed charge models (*67-69*).
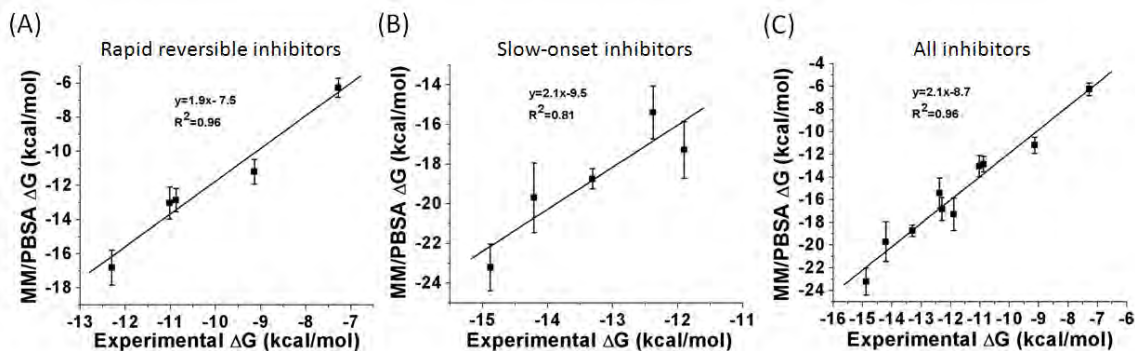
Figure 3.5. The correlation between the experimental binding free energies and the MM/PBSA binding free energies calculated from the closed InhA. (A) Rapid reversible inhibitors. (B) Slow-onset inhibitors. (C) All inhibitors. Detailed data are listed in **table 3.2**.

For those free energies calculated from the open conformation, only rapid reversible

inhibitors had a good correlation with experimental results. The correlation coefficients for rapid

reversible and slow-onset inhibitors are 0.89 and 0.54, respectively (**Figure 3.6**). For the rapid

reversible inhibitors, the calculated binding free energies from the closed and open states both

have good correlations with experiment results ($R^2 = 0.96$ and 0.89, respectively). However, for

the slow-onset inhibitors, the correlation coefficient drops from 0.81 (closed state) to 0.54 (open

state). These results match to our expectation that the open conformation is unlikely to be the

final EI* complex for the slow-onset inhibitors, so that the binding free energies calculated from

the open state have a weak correlation with the experimental values.
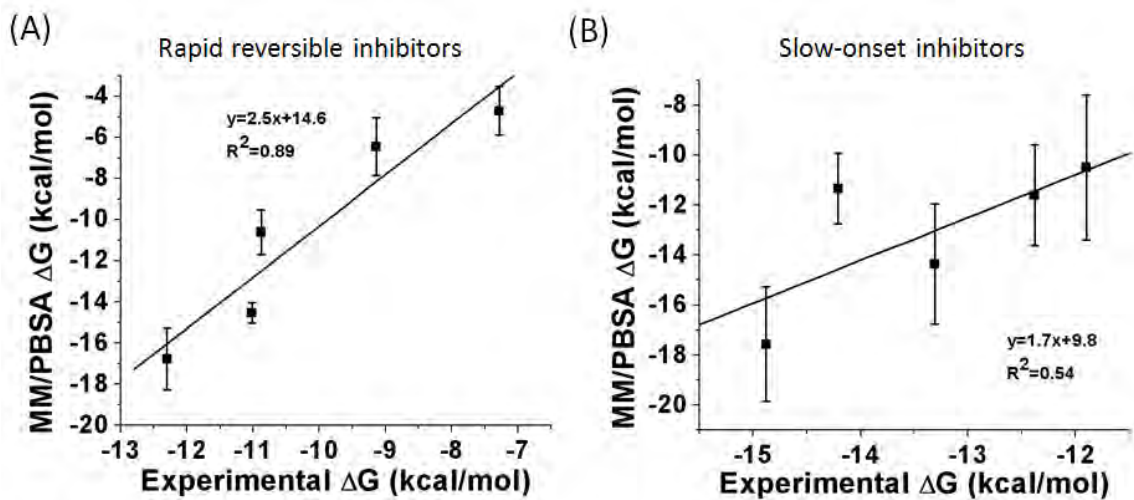
Figure 3.6. The correlation between the experimental binding free energies and the MM/PBSA binding free energy calculated from the open InhA. (A) Rapid reversible inhibitors. (B) Slow-onset inhibitors. Detailed values are shown in **table 3.3**.

Table 3.2. MM/PBSA binding free energies calculated from the closed conformation.

	ΔG_{explt}	ΔE_{vdw} A	ΔE_{col} B	ΔG_{sol} C	$\Delta G_{entropy}$ D	$\Delta E_{vdw} + \Delta E_{col} + \Delta G_{sol} + \Delta G_{entropy}$ A+B+C+D
Triclosan	-9.14 ± 0.05	-39.81 ± 0.27	-9.44 ± 0.84	22.32 ± 0.49	15.75 ± 0.53	**-11.19 ± 0.72**
PT3	-10.88 ± 0.19	-44.95 ± 0.20	-14.02 ± 0.51	27.57 ± 0.29	18.55 ± 1.02	**-12.85 ± 0.69**
PT4	-11.02 ± 0.03	-46.89 ± 0.66	-12.48 ± 1.11	26.34 ± 0.93	20.00 ± 1.00	**-13.03 ± 0.94**
PT5	-12.30 ± 0.10	-51.46 ± 0.47	-12.76 ± 0.62	26.80 ± 0.59	20.61 ± 0.36	**-16.81 ± 1.02**
PT52	-7.28	-34.77 ± 0.31	-9.79 ± 0.87	21.64 ± 1.10	16.66 ± 0.43	**-6.26 ± 0.56**
						R^2=0.96
PT70	-14.21 ± 0.06	-50.90 ± 0.30	-13.97 ± 0.73	27.79 ± 0.88	19.45 ± 0.59	**-19.69 ± 1.75**
PT82	-14.88 ± 0.08	-54.98 ± 0.46	-16.54 ± 2.65	29.17 ± 1.13	19.15 ± 0.37	**-23.20 ± 1.18**
PT91	-12.38 ± 0.08	-50.41 ± 0.53	-9.73 ± 0.91	25.28 ± 0.48	19.47 ± 1.06	**-15.39 ± 1.33**
PT92	-13.31 ± 0.13	-51.42 ± 0.24	-11.64 ± 0.32	25.32 ± 0.21	18.99 ± 0.49	**-18.74 ± 0.51**
PT119	-11.90 ± 0.09	-49.00 ± 0.72	-14.76 ± 2.77	29.95 ± 1.80	16.53 ± 0.51	**-17.28 ± 1.43**
						R^2=0.81

Experimental K_i values are taken from ref (*13, 142*).

The energy value was the mean value from four monomers. Statistical error was the standard error of mean from four monomers. It was estimated by calculating the standard deviation of four monomers and divided by $\sqrt{4}$. Although we performed the simulation in the biological tetramer assembly, we assumed four monomers work independently and there is no cooperativity and allostery in this system. Therefore, we treated one tetramer run as a four individual runs.

Table 3.3. MM/PBSA binding free energies calculated from the open conformation.

	ΔG_{explt}	ΔE_{vdw}	ΔE_{col}	ΔG_{sol}	$\Delta G_{entropy}$	$\Delta E_{vdw} + \Delta E_{col} + \Delta G_{sol} + \Delta G_{entropy}$
		A	B	C	D	A+B+C+D
Triclosan	-9.14 ± 0.05	-36.92 ± 0.68	-9.58 ± 0.78	22.51 ± 1.34	17.54 ± 1.36	**-6.45 ± 1.41**
PT3	-10.88 ± 0.19	-44.27 ± 0.34	-13.19 ± 1.04	28.10 ± 1.33	18.75 ± 0.95	**-10.61 ± 1.08**
PT4	-11.02 ± 0.03	-46.76 ± 0.73	-13.86 ± 0.23	27.86 ± 0.26	18.23 ± 0.85	**-14.53 ± 0.50**
PT5	-12.30 ± 0.10	-48.94 ± 1.56	-12.13 ± 1.60	25.76 ± 1.48	18.55 ± 1.19	**-16.76 ± 1.50**
PT52	-7.28	-30.99 ± 1.53	-6.26 ± 1.75	16.95 ± 2.57	15.59 ± 0.76	**-4.71 ± 1.18**
						R^2=0.89
PT70	-14.21 ± 0.06	-47.12 ± 1.14	-11.75 ± 0.91	26.60 ± 1.57	20.93 ± 1.04	**-11.34 ± 1.41**
PT82	-14.88 ± 0.08	-51.80 ± 1.45	-11.84 ± 1.18	25.53 ± 1.03	20.55 ± 0.55	**-17.56 ± 2.30**
PT91	-12.38 ± 0.08	-47.26 ± 1.09	-8.79 ± 1.33	24.01 ± 1.25	20.45 ± 0.81	**-11.60 ± 2.01**
PT92	-13.31 ± 0.13	-47.80 ± 1.32	-11.63 ± 2.40	26.03 ± 1.08	19.05 ± 1.14	**-14.35 ± 2.40**
PT119	-11.90 ± 0.09	-44.67 ± 1.95	-16.22 ± 5.31	33.18 ± 3.71	17.22 ± 1.65	**-10.49 ± 2.90**
						R^2=0.54

Experimental K_i values were taken from ref (*13, 142*).

The energy value was the mean value from four monomers. Statistical error was the standard error of mean from four monomers. It was estimated by calculating the standard deviation of four monomers and divided by $\sqrt{4}$. Although we performed the simulation in the biological tetramer assembly, we assumed four monomers work independently and there is no cooperativity and allostery in this system. Therefore, we treated one tetramer run as a four individual runs.

3.3.4 Free energy difference between the closed and open states

As mentioned in chapter two, because we speculated that the open and closed states are likely to be the EI and EI* states respectively, the slow-onset inhibitors shall have better binding energies in the closed state than in the open state. For the rapid reversible inhibitors, since the **triclosan**, **PT3**, and **PT5** bound ternary crystal structures have missing electron density in the helix-6 region, these inhibitors would not show a strong preference for the closed or open active-site conformation. It would be interesting to know whether the calculated binding free energies are consistent with the crystallographic observations. To know this, we compared the energy difference between the closed and open states.

For the slow-onset inhibitors, as expected, the binding free energies of the closed state have lower (favorable) values than those from the open states. By subtracting the MM/PBSA of closed and open states, we found that the slow-onset inhibitors were 3 ~ 8 kcal/mol more favorable in the closed state (**Table 3.4**). On the contrary, the rapid reversible inhibitors only modestly favored the closed state by 0 ~ 2 kcal/mol (**Table 3.4**). The MD simulation results are consistent with the crystal structure observations showing that the slow-onset inhibitors prefer the closed conformation while the rapid reversible inhibitors do not have strong preference for the closed or the open conformation. Moreover, for the rapid reversible inhibitors, because the binding free energy values at the closed and open states are similar, it explains why the R^2 values are good for the both closed and open states. As for the slow-onset inhibitors, because the crystal structures have a closed conformation, it is reasonable that the calculated binding free energies from the open state did not have a good correlation with the experimental binding free energies.

Table 3.4. Free energy differences between the closed and open conformation.

	$\Delta\Delta E_{vdw}$ A	$\Delta\Delta E_{col}$ B	$\Delta\Delta G_{sol}$ C	$\Delta\Delta G_{entropy}$ D	$\Delta\Delta E_{vdw}+\Delta\Delta E_{col}+\Delta\Delta G_{sol}+\Delta\Delta G_{entropy}$ A+B+C+D
Triclosan	-2.89 ± 0.68	0.14 ± 0.78	-0.19 ± 1.34	1.79 ± 1.36	**-4.73 ± 1.41**
PT3	-0.68 ± 0.34	-0.83 ± 1.04	-0.53 ± 1.33	0.20 ± 0.95	**-2.23 ± 1.08**
PT4	-0.13 ± 0.73	1.38 ± 0.66	-1.52 ± 1.11	-1.77 ± 0.93	**1.50 ± 0.94**
PT5	-2.52 ± 1.56	-0.63 ± 1.60	1.04 ± 1.48	-2.06 ± 1.19	**-0.06 ± 1.50**
PT52	-3.78 ± 1.53	-3.53 ± 1.75	4.69 ± 2.57	-1.07 ± 1.10	**-1.56 ± 1.18**
PT70	-3.78 ± 1.14	-2.22 ± 0.91	1.19 ± 1.57	1.48 ± 1.04	**-8.34 ± 1.75**
PT82	-3.18 ± 1.45	-4.7 ± 1.18	3.64 ± 2.65	1.40 ± 1.13	**-5.64 ± 2.30**
PT91	-3.15 ± 1.09	-0.94 ± 1.33	1.27 ± 1.25	0.98 ± 0.81	**-3.80 ± 2.01**
PT92	-3.62 ± 1.32	-0.01 ± 2.40	-0.71 ± 1.08	0.06 ± 1.14	**-4.39 ± 2.40**
PT119	-4.33 ± 1.95	1.46 ± 5.31	-3.23 ± 3.71	0.69 ± 1.80	**-6.78 ± 2.90**

These values were the energy difference between **tables 3.2 and 3.3**.

3.3.5 Comparisons of experimental binding free energy with the calculated VDW, electrostatic, entropy and solvation energies.

Previous study has shown that the VDW energy term controls the overall inhibitor binding in the *E. coli* FabI system (*90*). Because InhA is homolog of FabI, it would be interesting to know whether the VDW interaction is also important for the inhibitor binding in the InhA system as this information can assistant future lead optimization. In order to know what kind of energy terms control the inhibitor binding in the InhA enzyme, we compared the experimental binding free energies with the different MM/PBSA energy terms (VDW, electrostatic, entropy, and solvation).

As shown in **Figure 3.7**, the VDW energy term had the strongest correlation ($R^2 = 0.94$)

with the experimental binding free energies, suggesting that the inhibitor's shape is important for

inhibitor binding in the InhA enzyme. This result is similar to the previous *E. coli* FabI study

indicating the VDW energy term controls the overall inhibitor binding (*90*). Because we only

simulated 10 inhibitors bound complexes; larger and more diverse inhibitor sets may be

considered in order to further validate and test the conclusion we have here.
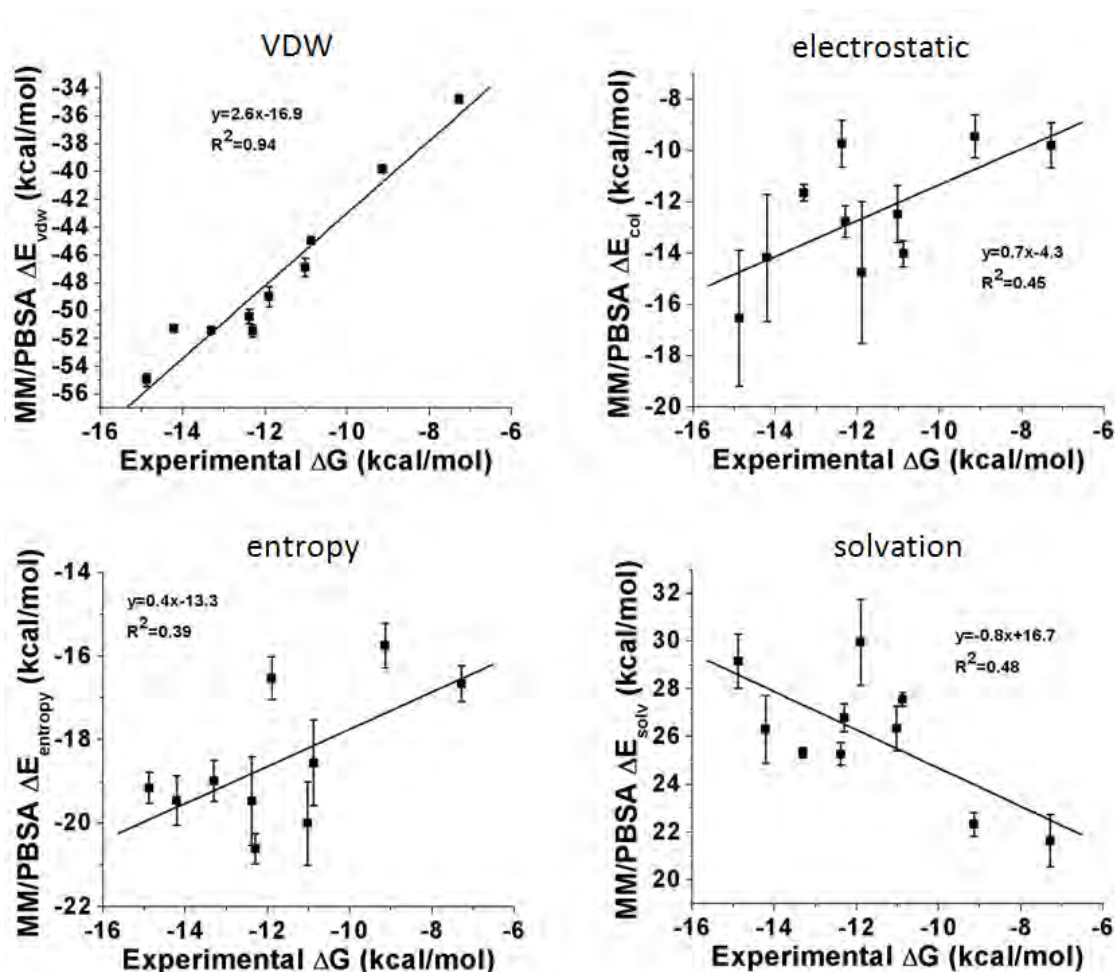
Figure 3.7. The correlation between the experimental binding free energies and the MM/PBSA VDW, electrostatic, entropy and solvation energies. The experimental binding free energies have the strongest correlation with the MM/PBSA VDW energies.

3.3.6 Per-residue Energy decomposition

Based on the observation from X-ray crystal structure, residues F149 ,Y158, A198, I202, M199, and V203 have been suggested to be the key residues for the inhibitor binding (*13*). Although we can estimate which residues are important in terms of measuring their distance between inhibitor, however, this approach is less quantitative. To get quantitative interaction energies, we calculated the non-bonded interactions between the inhibitor and each receptor residue.

Figure 3.8A shows the non-bonded interaction energies (van der Waals + electrostatic) between inhibitor **PT70** and receptor in the closed conformation. From this per-residue energy decomposition, residues G96, F149, Y158, M161, M199, and V203 are suggested to be key residues for receptor-ligand binding. Based on the energy decomposition results and the relative position of these key residues (**Figure 3.8B**), it seems like residues G96 backbone has a potential to form an H-bond with the inhibitor. For future inhibitor design, addition of some functional group at the *meta* or *para* position of the B-ring may create a H-bond with the residue G96.
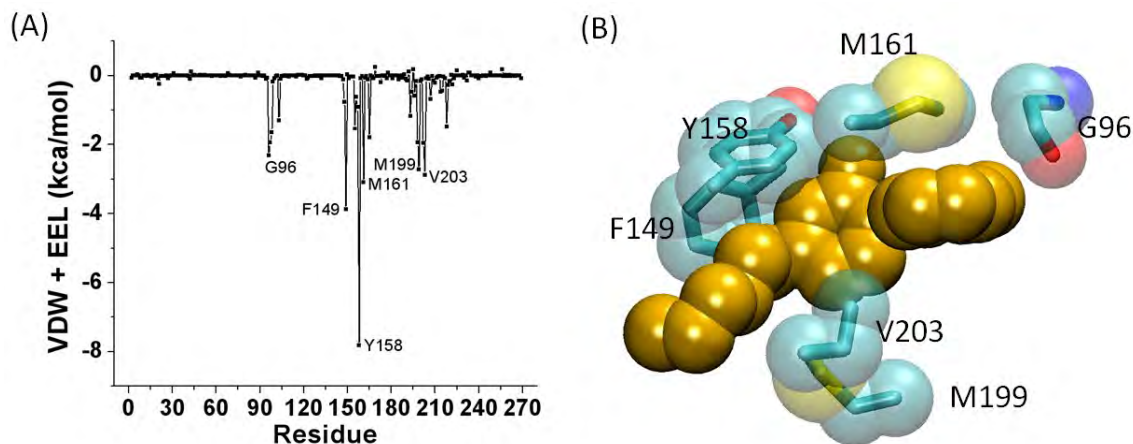
Figure 3.8. Per-residue energy decomposition of PT70. Residue G96, F149, Y158, M161, M199, and V203 contribute most to the inhibitor binding. (A) Energy decomposition (B) The relative position of those residues from (A).

3.3.7 What is missing from the free energy calculation in terms of explaining slow-onset inhibition?

From the MM/PBSA calculation, we found the slow-onset inhibitors prefer closed conformation while the rapid reversible inhibitors do not have strong preference for the closed or open conformation and these results are in consistent with the crystallographic observations. We demonstrated that the MM/PBSA approach reproduced a good trend of the binding affinity in the InhA system. However, these calculated free energies are from the ground states, and are insufficient to address the slow-onset inhibition mechanism. As we mentioned in chapter two, we speculated that the closed and open are highly possible to be the EI and EI* states, respectively, and the conformational change from the open to the closed state is the isomerization step. To test this model, we have to validate whether there are energy barriers between the closed and open states and an energy profile along the open-closed path will be required. We present the open-

82

closed path energy calculations in a separately chapter (see chapter four). In that chapter, we applied pathway sampling approach and more reliable free energy calculation approaches, such as umbrella sampling and thermodynamic integration to better understand the transition state barrier. Although the MM/PBSA results cannot explain the conformational change step, the ability of MM/PBSA reproduces the binding affinity seems promising. For future InhA inhibitors design, the MM/PBSA approach can be used to predict the inhibitors' binding affinity before synthesis of compound.

3.4 Summary

To know whether MD simulation can robustly reproduce the binding affinity in the InhA system as well as to get a better understanding of what kind of non-bonding interaction (van der Waals or electrostatics) controls the inhibitor binding, we performed MD simulations for 10 inhibitors bound complexes (with both closed and open conformations) and used the MM/PBSA approach to calculate the binding free energy.

The calculated binding free energies from the MM/PBSA approach had a good correlation (R^2 = 0.96) with the experimental values, indicating MD simulation reproduces the trend of binding affinity in the InhA enzyme. Detailed comparisons between closed and open states free energies demonstrate that the slow-onset inhibitors had more favorable binding free energy at the closed state than the open state by 2 ~ 8 kcal/mol, while rapid reversible inhibitors did not have strong preference for the closed or open conformations. By comparing the different energy terms, we found that the inhibitor's shape is crucial for binding in the InhA enzyme. Per-

residue energy decomposition shows that residues F149, Y158, M161, M199 and V203 play important roles for the inhibitor binding.

Chapter Four

A structural and energetic model for the slow-onset inhibition of the InhA

4.1 Introduction

Drug-target residence time has become an important parameter in lead optimization (*14-18*). Traditionally, a strong binder is thought to have better efficacy than a weak binder does (*19*). However, many examples suggest that the *in vivo* efficacy does not always linearly correlate with binding affinity (*20*). Instead of binding affinity, binding kinetics has caught drug discovery community's attention and an increasing number of groups have emphasized both structure-kinetics relationship and structure-activity relationship (*149, 150*). Residence time, the reciprocal of the dissociation rate constant, represents the life-time of an enzyme-inhibitor complex (*15*). The observations of many current drugs have long residence times on their targets suggest that drug-target residence time is an important component of *in vivo* drug activity (*18*). Especially, the study of the InhA homologue in *Francisella tularensis* (ftuFabI) demonstrated that inhibitor's residence time had better correlation with *in vivo* efficacy than the binding affinity did (*21*).

There are few studies that used free energy landscape model to predict the inhibitor association/dissociation kinetics. A recently published paper used free energy landscape theory to predict the binding kinetics of Huperzine A, an drug of Alzheimer's disease, and the predicted binding and activation energies deviated less than 1 kcal/mol to the experimental values (*66*). Another opportunity of using free energy profile to explain slow kinetics is found in the study of comparing charged and uncharged ligands in the Q_A site of photosynthetic bacterial reaction centers (*151*).

In principle, there are two possible ways to increase the inhibitor residence time. The inhibition of InhA with the INH-NAD adduct occurs through a two-step binding mechanism in which the initial rapid formation of a weakly bound enzyme-inhibitor complex (EI) is followed by a slow conversion step, leading to a higher affinity complex (EI*). The first approach to increase the residence time is to stabilize the final EI* complex and the second approach is to destabilize the transition state between EI and EI* states. For InhA lead optimization, **PT70** is a tight binding inhibitor of InhA with a K_i value in pM level that would be very challenging to increase the residence time by solely considering the binding affinity of EI* state. Alternatively, increase of the free energy barrier in the isomerization step could be a more feasible approach. Thus, a detailed understanding of the interactions that modulate the barrier to this isomerization step would thus facilitate lead optimization and rational design in binding kinetics.

In this chapter, the goal is to obtain the free energy landscape along the open-closed path and to understand the detailed conformational changes during the transition, further to obtain a model for the transition state that would provide insight into rational control of the energy barrier. Since these states are inherently difficult to characterize experimentally, we used a variant of the nudged elastic band (NEB) (*92*) simulation method, partial NEB (PNEB) (*95*) to generate a set of structural snapshots representing the conformational pathway between the crystallographic open and closed states. We used free energy calculations to evaluate the energy profile for loop closing, and to understand the impact on both the binding affinity and the energy barrier when the protein or ligand was modified. The free energy profiles demonstrated that rapid reversible inhibitor complex exhibited two types of potential of mean force (PMF), either with preference for the open state, or with little preference and low energy barrier between the open and closed states. On the other hand, slow-onset inhibition complexes had a relative stable closed state with

a more significant energy barrier between the open and closed states. These energy profiles explain why rapid reversible inhibitors bound complex can have ordered open (like **PT155** bound InhA) or disordered (like **triclosan**, **PT3** and **PT5** bound InhA) helix-6, while slow-onset inhibitors bound complex (like **PT70** bound InhA) have only ordered closed helix-6 from the x-ray crystallography structures.

Structural analyses of the transition path revealed specific active-site residues that create steric hindrance at the transition state. We hypothesized that these steric clashes control the slow-onset inhibition of the InhA enzyme, and changes of side chain or inhibitor size in this region could modulate slow-onset inhibition. Further loss and regain of energy barrier studies demonstrate that the free energy barrier at the transition state can be controlled by mutations and designed inhibitors. MD simulation showed that removal of the steric clashes by replacement of amino acids with those possessing smaller side chains resulted in reduction of the energy barrier. Furthermore, the energy barrier was regained in the mutants when using either of the two following approaches. The steric clash lost when a side chain is mutated to a smaller group can be regained in a double mutant with a nearby, compensating change to a larger side chain. Likewise, energy barrier lost in the single mutant can be regained when using inhibitors designed with bulky groups intended to restore the steric clash. These loss and regain of function studies validate our mechanistic hypothesis for the key interactions controlling active-site open-closed conformational change. Moreover, we identified the potential for the same mechanistic approach in FabI enzymes from other important pathogens.

4.2 Method

4.2.1   Molecular dynamics simulation setup

For the InhA system, the initial pose of inhibitor **PT70** was taken from the first monomer of the InhA:NAD$^+$:**PT70** ternary complex (PDB ID: 2X23 (*13*)) while the initial poses of **PT162** and **PT163** (**Figure 4.1**) were generated by the DOCK 6.3 suite of docking software (*140*). The procedure for preparing the binding-site for docking has been described previously (*141*). Flexible ligand docking and default parameters were used. We assumed that the diphenyl ether of **PT162** and **PT163** occupy similar position as **PT70** does, thus the docking poses with lowest RMSD value in the diphenyl ether moiety was chosen as the initial structures. For the saFabI system, the initial structure of **PT4** was taken from the first monomer of the saFabI:NAD$^+$:**PT4** ternary complex (PDB ID: 4BNH (*152*)) while the initial structures of **PT1** and **PT166** (**Figure 4.1**) were generated by the DOCK 6.3 suite of docking software with same flexible docking protocol (*140*).

For the InhA system, the initial closed and open structures were taken from the first monomers of the InhA:NAD$^+$:**PT70** ternary complex (PDB ID: 2X23 (*13*)) and the InhA:NAD$^+$:**C16-NAC** complex (PDB ID: 1BVR (*11*)), respectively. For the saFabI system, the initial closed structures were taken from the first monomers of the saFabI:NAD$^+$:**PT4** ternary complex (PDB ID: 4BNH (*152*)). AMBER ff99SB (*26*) and GAFF (*115*) force field were assigned to the protein and inhibitor, respectively. The force field parameters of the cofactor NAD$^+$ and NADP$^+$ were obtained from previous studies (*118, 119*). The partial atomic charges of inhibitors were computed using Gaussian98 (*153*) with the HF/6-31G* basis set, followed by RESP fitting (*46, 116*). The mutants were generated using Swiss-Pdb viewer (*154*). Each

complex was solvated in a truncated octahedral TIP3P (*120*) water box with a minimum distance of 8 Å between the water box edge and solute, resulting in ~ 23,000 atoms in total. SHAKE (*155*) was used to constrain bond lengths involving hydrogen. The particle mesh Ewald method (*156*) was used for calculating electrostatic energy with an 8Å nonbonded cutoff. Each system was equilibrated initially by high restraints on non-water atoms, and gradually reduced and removed the positional restraints on the backbone atoms using the following procedures. The first step was 10,000 steps of steepest descent minimization with 100 kcal mol^{-1} Å^{-2} restraints on all atoms except water molecular and hydrogen atoms. The second step was heating the system from 100 to 300 K at constant volume over 100 ps with 100 kcal mol^{-1} Å^{-2} restraints on non-water and non-hydrogen atoms, followed by 100 ps with the same restraints at constant 300 K temperature and 1 atm pressure. The third step was 250 ps MD with restraint weight of 10 kcal mol^{-1} Å^{-2} on the non-water and non-hydrogen atoms at constant 300 K temperature and 1 atm pressure. The following steps only restrained the backbone atoms and gradually reduced the restraints weight from 10 to 0.1 kcal mol^{-1} Å^{-2} at constant 300 K and 1 atm. This was carried out by 100 ps with 10 kcal mol^{-1} Å^{-2} restraint, followed by 100 ps with 1 kcal mol^{-1} Å^{-2} restraint, and 100 ps with 0.1 kcal mol^{-1} Å^{-2} restraint.

Figure 4.1. Inhibitors studied in this chapter.

4.2.2   Open conformation of saFabI

The open conformation of saFabI was generated from steered MD simulation. This was done by gradually increasing the distance between center of mass (COM) of residue 194 to 198 backbone atoms and COM of residue 92 to 96 and 111 to 120 backbone atoms to 20 Å with a 10 kcal/mol Å^{-2} restraint over 500 ps MD simulation.

4.2.3 PNEB simulation

The partial nudged elastic band (PNEB) (*95*) simulation approach was used in this study to generate structure models along the low-energy pathway between the open and closed conformations. The equilibrated open and closed structures were used as the two end-point structures and 30 windows (including end-points) were used in the simulation. The NVT ensemble was used with PNEB. A spring force was applied to the backbone atoms of α-helices 6 and 7 (for InhA: residues 196 to 223; for saFabI residues 189 to 215) in the subsequent steps. In the first 40 ps, the system was equilibrated at 300 K with a Langevin collision frequency of 50 ps^{-1} and spring force of 20 kcal mol^{-1} $Å^{-2}$. The next step was 100 ps equilibration at 300 K with a 20 ps^{-1} Langevin collision frequency and a spring force of 75 kcal mol^{-1} $Å^{-2}$. After the conformations were generated along the open-to-closed path, simulated annealing approach was used to optimize the local energy minimized path. This was done by heating the system from 300 to 375 K gradually over 175 ps, and subsequently cooling back to 300 K gradually over 175 ps with a 20 ps^{-1} Langevin collision frequency and a spring force of 75 kcal mol^{-1} $Å^{-2}$. A 20 kcal mol^{-1} $Å^{-2}$ Cartesian restraint was applied to the backbone atoms from residues 2 to 195 and 225 to 268 (for saFabI: residues 1 to 188 and 216 to 254) to maintain structure stability during the heating process. After this simulated annealing, a 600 ps production run was performed with the same Langevin collision frequency and spring force at 300 K. In order to maintain structure stability at the tetramer interface, a 10 kcal mol^{-1} $Å^{-2}$ Cartesian restraint was applied to the backbone atoms from residues 2 to 195, and 225 to 268 (for saFabI: residues 1 to 188 and 216 to 254) during the production run.

4.2.4   Umbrella sampling

Energy landscapes were obtained by umbrella sampling (US). Two reaction coordinates (step and shear torsions) were used to describe the motion of α-helices 6 and 7 (**Figure 4.2**). US windows were spaced at 3º increments in both step and shear torsions. An extra 6º buffer region was added around the grid region sampled in the PNEB simulation, resulting in a total of ~192 grid points. Initial structures at each grid point were selected from the PNEB production trajectories with dihedral values (step/shear torsions) closest to the respective grid point. 500 ps MD simulation in the NVT ensemble at 300 K with a Langevin collision frequency of 75.0 ps^{-1} was performed for each grid point. The Weighted Histogram Analysis Method (WHAM) (*98*) approach and analysis program (*96*) was then used to obtain the potential of mean force (PMF) from the umbrella sampling results. Convergence of the free energy calculations was tested by extending the umbrella sampling of InhA:NAD$^+$, InhA:NAD$^+$:**PT155**, and InhA:NAD$^+$:**PT70** complex for another 500 ps. There was no significant change in the free energy from 1st 500 ps, 2nd 500 ps, and 1 ns runs. We assumed other complexes also have similar results, thus 500 ps umbrella sampling were run for other complexes (**Figure 4.3**).
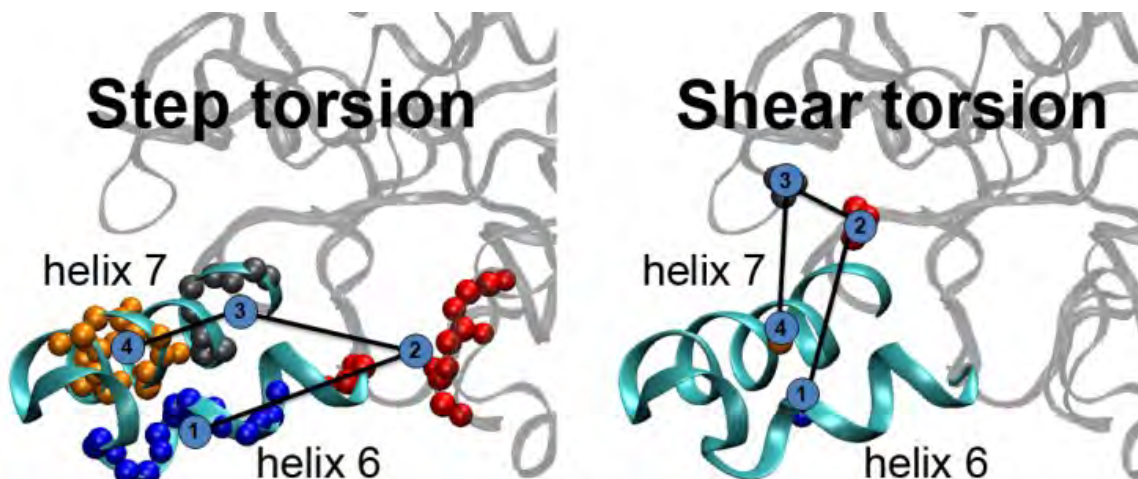
Figure 4.2. Step and shear torsion angles. In the step torsion, point 1 and 4 are the center of mass (COM) of backbone atoms from residue 200 to 205 and residue 211 to 216, respectively. Point 2 is the COM of backbone atoms from residue 19 to 21 and 196. Point 3 is the COM of backbone atoms from residue 219 to 222. In the shear torsion, points 1 and 4 are the beta-carbon of residue 203 and 215, respectively. Points 2 and 3 are the COM of backbone atoms from residue 98 and 158, respectively. For the saFabI system, similar groups of atoms were used. In the step torsion, point 1 and 4 are the COM of backbone atoms from residue 195 to 196 and residue 202 to 206, respectively. Point 2 is the COM of backbone atoms from residue 16 to 18 and 193. Point 3 is the COM of backbone atoms from residue 216 to 219. In the shear torsion, points 1 and 4 are the beta-carbon of residue 195 and 202, respectively. Points 2 and 3 are the COM of backbone atoms from residue 93 to 95 and 157 to 161, respectively.
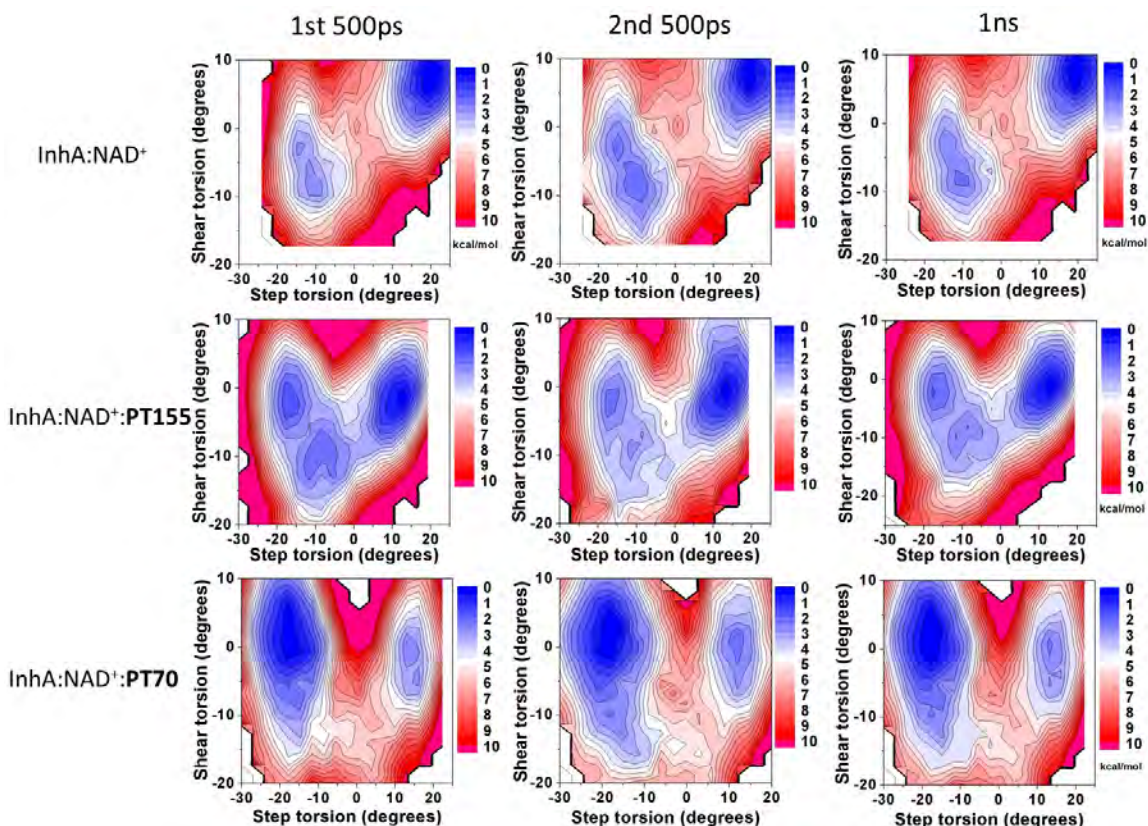
Figure 4.3. Convergence test of PMF. There is no significant change in the free energy from 1st 500 ps, 2nd 500 ps and 1 ns runs.

4.2.5 Thermodynamic integration calculation

Thermodynamic integration (TI) calculation was used to calculate the relative binding free energy ($\Delta\Delta G$) between two states. Blow is the scheme showing the thermodynamic cycle of wild-type and mutant enzymes with a same inhibitor and their associated free energies.
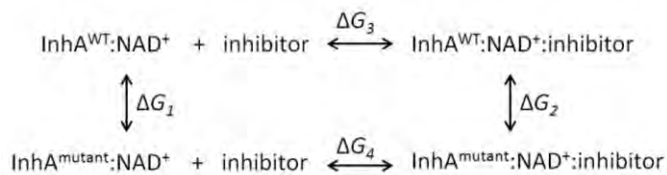
where $\Delta\Delta G = \Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3$

ΔG_3 and ΔG_4 represent the binding free energies of an inhibitor to the wild-type and mutant InhA, while ΔG_1 and ΔG_2 represent the energies of alchemical transformation in the enzyme along and complex environment. The TI method calculates the alchemical transformation based on the following equation:

$$\Delta G = \int_0^1 \langle \frac{\partial V(\lambda)}{\partial \lambda} \rangle_\lambda \, d\lambda$$

where λ represents the coupling parameter describing state between initial and final states, and $V(\lambda)$ is the λ-coupled potential function. The $\partial V/\partial \lambda$ value can be obtained from the Sander output files. To get a smooth transformation from the initial state to the final state, three steps approach were used in this study (53). First step was removing the partial charges on the target atoms. Second step was soft-core vdW transformation. Third step was adding partial charges to the target atoms. The free energy changes associated with each of the steps can be estimated from numerically integrating the $\partial V/\partial \lambda$ values from multiple simulations at a fixed λ value based on the following equation.

$$\Delta G = \sum_{i=1}^{n} W_i \langle \frac{\partial V}{\partial \lambda} \rangle_{\lambda_i}$$

Based on the AMBER tutorial, 19 λ points (0.05 to 0.95 with interval 0.05) were used in this study (157). The 0.05 weight value (W_i) was used for each λ point. For the endpoint ($\lambda=0$ and 1), trapezoid rule was used to linearly extrapolate the ∂V value from the closest two values and their associated weights were 0.025.

Each system was equilibrated initially with 5,000 steps of steepest descent minimization with 100 kcal mol^{-1} Å^{-2} restraints on all atoms except water molecular and hydrogen atoms, followed by heating the system from 100 to 300 K at constant volume over 100 ps, and 100 ps

MD at constant 300 K and 1 atm with same restraints. Next step was 250 ps MD with same temperature and pressure but changed the restraints to 10 kcal mol^{-1} Å^{-2}. Finally, the restraints were only applied on backbone atoms and decreased gradually. This was carried out by 100 ps with 10 kcal mol^{-1} Å^{-2} restraints, followed by 100 ps with 1 kcal mol^{-1} Å^{-2} restraints, and 100 ps with 0.1 kcal mol^{-1} Å^{-2} restraints. 1 ns production run was performed in NPT ensemble using Langevin thermostat with a collision frequency of 2.0 ps^{-1}. The $\partial V/\partial \lambda$ statistics values collected from the production run were used to numerically estimate the $\Delta\Delta G$.

4.3 Results and Discussion

In this chapter, we aimed to obtain the free energy landscape along the open-closed path and to understand the detailed conformational changes during the conformational isomerization, further to obtain the transition state structure that would provide insight into rational control of the induced-fit conformational change. Because these transition state structures are inherently difficult to characterize experimentally, we used a variant of the nudged elastic band (NEB) (*92*) simulation method, partial NEB (PNEB) (*95*) to generate a set of continued structures representing the conformational pathway between the crystallographic open and closed states. We used free energy calculation (umbrella sampling) to evaluate the energy landscape along the open-closed path, and to understand the impact on both the binding affinity and the energy barrier when the protein or ligand was modified.

4.3.1 The continued structures along the open-closed path

Because we need the energy profile along the open-closed path to evaluate the transition state barrier, first we have to get continued structures between the crystallographic open and closed conformations, followed by the free energy calculations. To obtain continuous structures along the open-closed path, partial nudged elastic band (PNEB) simulation approach was used. During the transition from open to closed, helix 6 moved up and helix 7 moved down simultaneously, but helix-6 had a larger motion than helix-7 (**Figure 4.4**). In order to quantify the different conformations along this transitional path, two reaction coordinates (torsion angles) were defined (**Figure 4.2**). The first reaction coordinate was step torsion, which characterized the up-down motion of helix-6 and helix-7. From structural analyses, some potential steric interactions were observed between side chain residues from helix-6 and helix-7. Hence, the second reaction coordinate, shear torsion, was introduced to describe the motion between side chain atoms of V203 and I215. Based on these two reaction coordinates, the closed (PDB ID: 2X23) and open (PDB ID: 1BVR and 2AQ8) crystal structures are clustered into specific ranges, suggesting these reaction coordinates are good descriptors for quantifying the open-closed motion. The open conformation locates in the region has step torsion 15 ~ 20 degrees, and shear torsion -4 ~ 5 degrees; whereas the closed conformation has step torsion in -19 degrees and shear torsion in 3 degrees. (**Figure 4.5**).
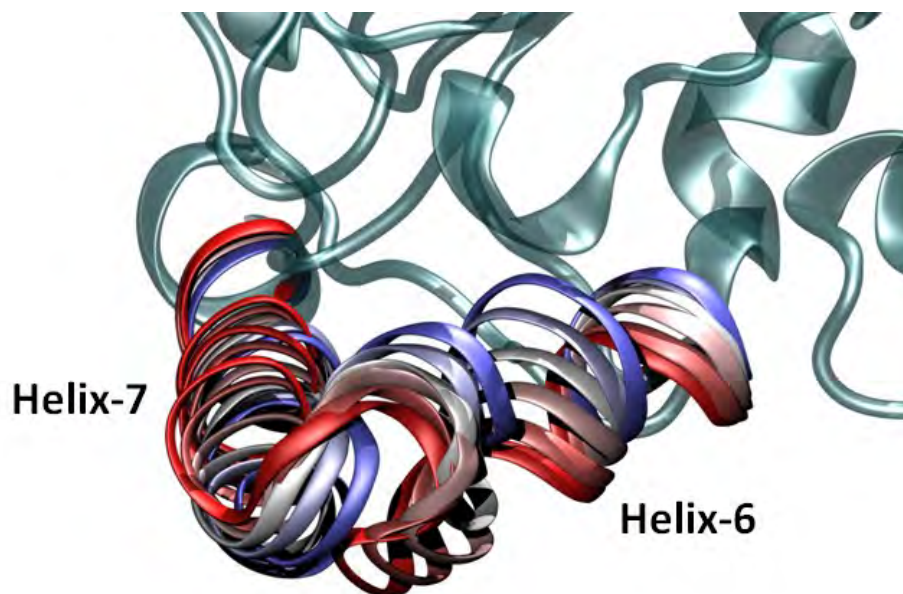
Figure 4.4. The continued structure between open and closed InhA. The blue color represents closed state while red color represents open state.
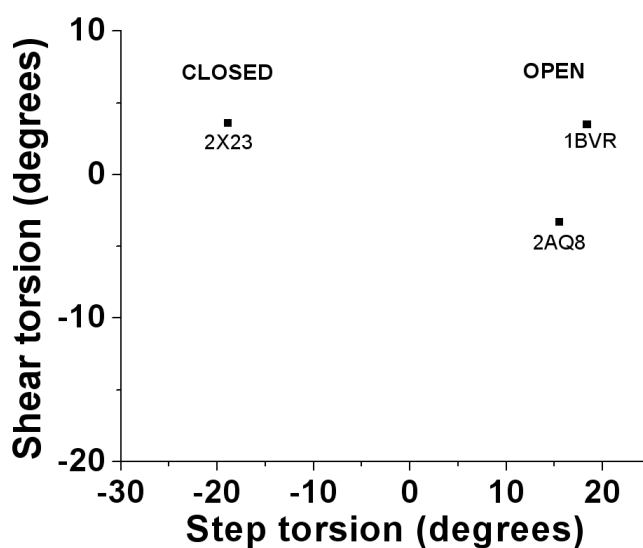
Figure 4.5. Mapping the closed and open crystal structures into the reaction coordinates of step and shear torsion angles. Based on measurements of step and shear torsions, the closed (PDB ID: 2X23) and open (PDB ID: 1BVR and 2AQ8) crystal structures are clustered into specific ranges, suggesting these reaction coordinates are good descriptors for quantifying the open-closed motion.

4.3.2 Free energy profile of InhA:NAD$^+$ binary complex

To better understand why certain inhibitor bound complexes have open, closed, and disordered helix 6, we use umbrella sampling based on the resulted PNEB structures to obtain a 2D free energy profile along the open-closed path. We first calculated the binary complex (without inhibitor) and repeated the calculations for the inhibitor bound complexes. Based on the crystal structure observation of the binary complex (InhA:NAD$^+$), InhA has an open active-site conformation (*132*). We speculated that the energy profile would have a very stable energy value at the open state.

As expected, the energy profile of InhA:NAD$^+$ complex derived from the umbrella sampling had a good agreement with the crystal structure observation, showing that the global minimum is in the open conformation (**Figure 4.6**). In this PMF, two local energy minima are located in the open and closed positions, and an energy barrier is in between the open and closed states. The global minimum, which locates in the open state, is ~3 kcal/mol more favorable over the closed state. In addition, a ~6 kcal/mol energy barrier prevents the helix-6 moving from the open state to closed state.
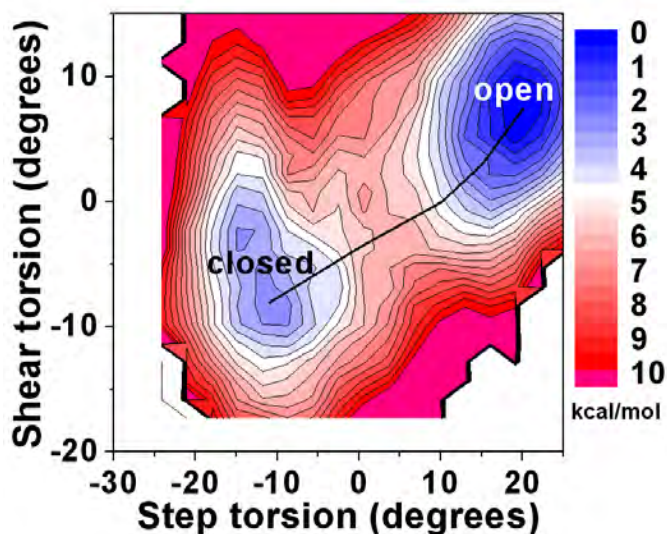
Figure 4.6. PMF of InhA:NAD+ binary complex. The global minimum is in the open state, in agreement with crystal structure observation (PDB ID: 1BVR). An energy barrier is between open and closed states.

4.3.3   Free energy profiles of rapid reversible inhibition complexes

To better understand why certain inhibitor bound complexes have open, closed, and disordered helix 6, we performed free energy calculation along the open-closed path for different inhibitor bound complexes. In chapter three we showed the binding free energy calculations at the closed and open conformation only provide the ground state information and are insufficient to explain the slow-onset inhibition mechanism which involved a structural isomerization step. Although **triclosan**, **PT3** and **PT155** all are rapid reversible inhibitors of InhA, **triclosan** and **PT3** bound InhAs have a disordered helix-6 but **PT155** bound InhA has an ordered and open helix-6 conformation based on the crystal structure observations. The **PT155** bound InhA crystal structure is exception to the hypothesis of "slow-onset inhibition is coupled to ordering of active-site" in previous study (*13*). It is very interesting to know why InhA with **PT155** has an ordered active-site. One of the explanations is the crystal packing causes an ordered active-site loop.

However, after examining the crystal structure, crystal packing does not seem to be the reason. It would be interesting to know what kind of mechanism resulted in an ordered/disordered active-site loop. In chapter 2, we showed the hypothesis of active-site open-to-closed conformation change might correlate with the slow-onset inhibition. If this model is correct, we expect a rapid reversible complex could exhibit two kinds of energy profiles. The first one is a flat energy landscape, with no energy barrier and no strong preference of open and closed states. The other one is that the complex favors open state (**Figure 4.7**). From the enzyme kinetics view, both of these scenarios would exhibit one step binding kinetics. We hypothesized that **triclosan** and **PT3** bound InhA would have a flat energy barrier since the crystal structures have a disordered active-site, while **PT155** bound InhA would have a similar energy profile as the InhA:NAD$^+$ complex, exhibiting a global minimum in the open state and an energy barrier between the open and close states. To test this hypothesis, we ran the PNEB/US simulations for the **triclosan**, **PT3** and **PT155** bound complexes.
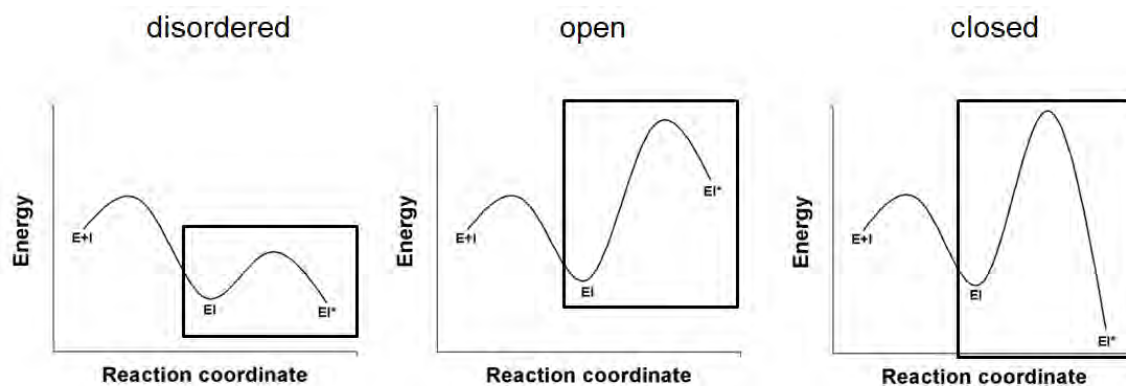
Figure 4.7. The possible energy profiles for disordered open, and closed active-site loop. For the rapid reversible inhibition complexes, if the crystal structure has a disordered active-site loop, the EI to EI* state would have similar free energy values. In addition, if the crystal structure has an open active-site loop, the EI state would be the global minimum. For a slow-onset inhibition complex, the global minimum would always in the EI* state and an energy barrier in between the two states.

As expected, InhA with rapid reversible inhibitor favors open state or has low energy barrier between open and closed states. Similar to the InhA:NAD$^+$ binary complex, the **PT155** bound ternary complex (denoted as InhA:NAD$^+$:**PT155**) favors open conformation (**Figure 4.8A**). The free energy barriers of open-to-closed and closed-to-open were 3.40 ± 0.09 and 2.15 ± 0.27 kcal/mol, respectively. We further measured the step torsion and shear torsion of the **PT155** bound crystal structure (PDB ID: 4OXK) and found that the crystal structure is nearly in the global minimum region of the PMF (step and shear torsion are 9.2 and 1.5 degrees, respectively). The free energy landscape rationalizes the **PT155** bound crystal structure, demonstrating even a rapid reversible can also have an ordered helix-6.

Unlike the **PT155** bound complex, the InhA:NAD$^+$:**PT3** complex had equal free energy in the open and closed states (**Figure 4.8B**). The free energy barriers were 1.97 ± 0.16 and 2.57 ± 0.03 kcal/mol in the open-to-closed and closed-to-open directions, respectively. Because of the flat energy landscape, the active-site helices are expected to move randomly between the open

and closed conformation, resulting an indistinct electron density in the helix-6 region. Again, the PMF not only consistent with the crystal structure observation but also tell why the electron density of helix-6 is blurred in the **PT3** bound crystal structure (PDB ID: 2B36). Similar result was also found in InhA:NAD$^+$:**triclosan** complex, showing a flat free energy profile and explains why the crystal structure has a disordered helix-6 (**Figure 4.8C**).
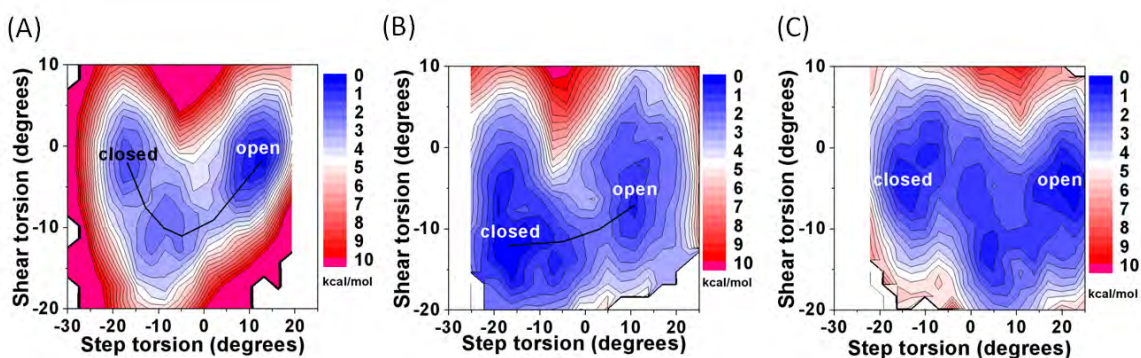
Figure 4.8. PMF of rapid reversible complexes. (A) InhA:NAD$^+$:**PT155** (B) InhA:NAD$^+$:**PT3** and (C) InhA:NAD$^+$:**triclosan**. **PT155** bound InhA favors open conformation, while **PT3** and **triclosan** bound InhA have flat energy barrier.

4.3.4 Free energy profiles of slow-onset inhibition complexes

The energy profiles of rapid reversible inhibitors bound complexes explained the rapid reversible can have an open or disordered helix-6 conformation. These results match to our expectation when treating the open and closed active-site as the EI and EI* states (**Figure 4.7**). If our model is correct, we speculated the slow-onset inhibition complexes will also match to our expectation in **Figure 4.7**; the closed state is the global minimum in the energy profile and an energy barrier between the open and closed states. **PT70** and **PT92** are slow-onset inhibitors of InhA that have two-step inhibition kinetic property (*13, 142*), we simulated these two inhibitors

bound complexes and expected that the PMFs will exhibit a very stable closed state and an energy barrier in between the closed and open states.

The free energy profiles, as expected, when InhA bound with these two inhibitors, the closed state had a lower free energy then the open state and an energy barrier in between them (**Figures 4.9**). For the **PT70** bound complex, the free energy barrier from open to closed was 3.26 ± 0.55 kcal/mol, whereas that of closed to open was 6.70 ± 0.41 kcal/mol. For the **PT92** bound complex, the free energy barrier from open to closed and closed to open were 1.63 ± 0.61 and 6.81 ± 0.80 kcal/mol, respectively. These PMF profiles support the two step slow-onset inhibition and induced-fit models which the enzyme inhibitor initially form an EI complex (open), followed by an isomerization step, leading to a tightly binding EI* complex (closed). Moreover, the much deeper well in the closed conformation gives helix-6 a much better chance to stay structurally ordered in the crystal structure, which is consistent with the crystal structure observation that active-site helices are ordered in the **PT70** and **PT92** bound structures.
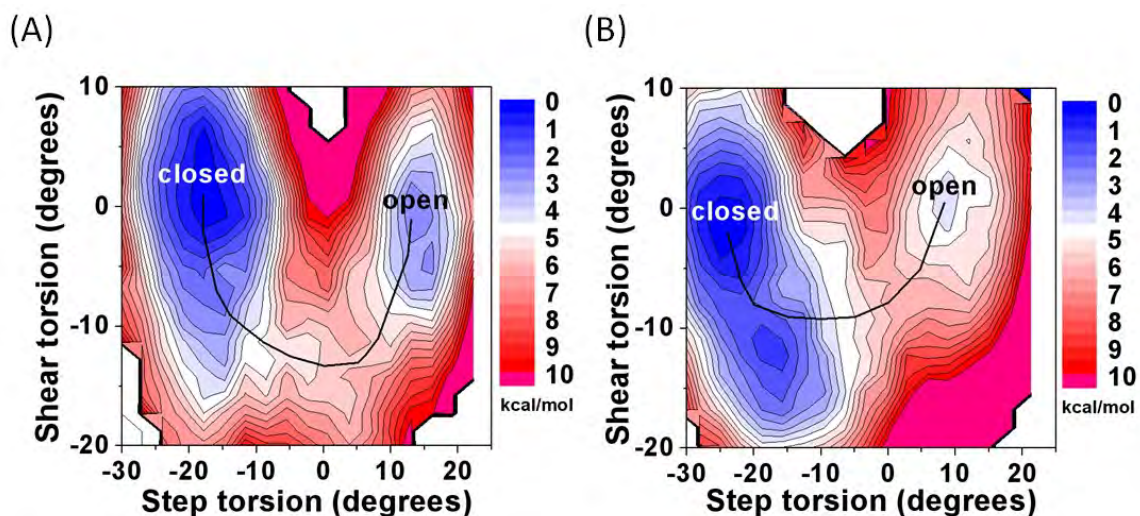
Figure 4.9. PMF of slow-onset inhibition complexes. (A) InhA:NAD$^+$:**PT70** and (B) InhA:NAD$^+$:**PT92**.

4.3.5 Key interactions during the open-closed conformational change

In the above energy profiles, the rapid reversible complexes exhibit two types of PMF, either preferring the open state, or having little preference and low energy barrier between the open and closed states. On the other hand, slow-onset inhibition complexes have a relatively stable closed state with a more significant energy barrier between the open and closed states. The InhA^{WT}:NAD$^+$:**PT70** complex shows an energy barrier at the transition state from 5 to 10 degrees of step torsion, and -10 to -15 degrees of shear torsion (**Figure 4.9A**). Although comparison of the PMFs can suggest whether a particular enzyme-inhibitor complex may exhibit slow-onset inhibition, the PMF itself provides little information regarding which residues and/or what kinds of interactions contribute to the energy barrier. Importantly, the value of the MD simulations is that they allow us to extract structures near the PMF transition state and gain

insights into the atomic-level interactions that give rise to the slow-onset behavior. This transition state model can then be used to propose subsequent rational control of binding kinetics.

Structural analysis of the transition state structure from the **PT70** bound PNEB results revealed that the side chains of active-site residues M199, V203, A211, and I215 encounter a steric clash near the transition state (**Figure 4.10**). We speculated that a slow-onset inhibitor, such as **PT70**, crowds the helices 6 and 7 regions, particularly constrains residues M199, V203A, A211, and I215 during active-site loop closing step. To measure the progress of this reorganization, we used the relative position of residues V203 and I215 to distinguish between open and closed states. By placing and viewing the protein structure as shown in **Figure 4.10**, residue V203 is below residue I215 in the open state, and vice versa in the closed state. This motion, during which residue I215 passes residues near M199 and V203, and residue V203 passes residues A211 and I215, is similar to opening a purse clasp that requires extra force to move past the steric clash. Conversely, a smaller inhibitor, such as **triclosan** and **PT3**, cannot constrain these residues during the loop closing step; therefore, these sidechains bypass the energy unfavorable steric clash.
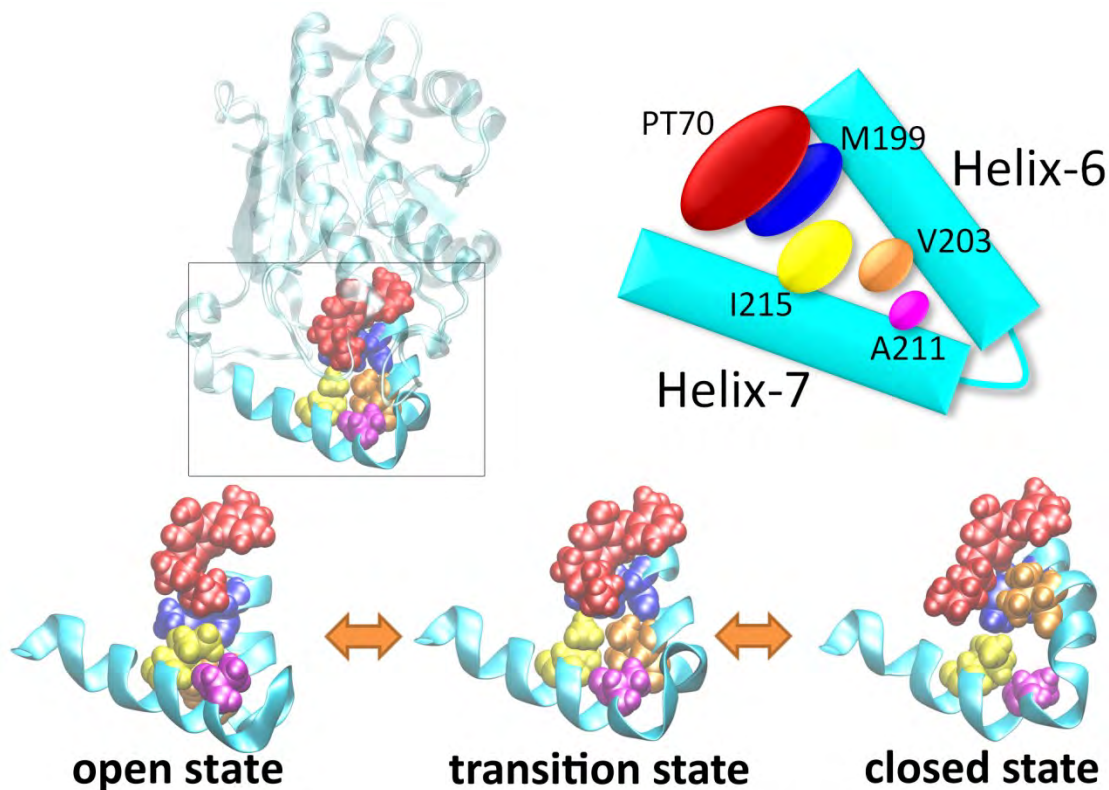
Figure 4.10. Key residues involved in the transition between open and closed states. Residues M199, V203, and I215 form a "coin purse clasp" conformation and encounter an energetically unfavorable conformation at the transition state.

To further examine the steric clashes among these residues, we measured inter-residue backbone distances during the transition. We observed that residues V203 and I215 follow a non-linear path during loop opening, initially becoming close, but then needing to move to longer distances before they become close again (**Figure 4.11**). This suggests that the side chains approach to form vdW contact, but the direct path to closed has a steric clash, and as a result the low-energy path takes a detour by separating the loop region prior to complete opening. We speculated that these steric clashes between resides V203 and I215 control the conformational changes between open and closed states for the slow-onset inhibition complex InhA:NAD$^+$:**PT70**.
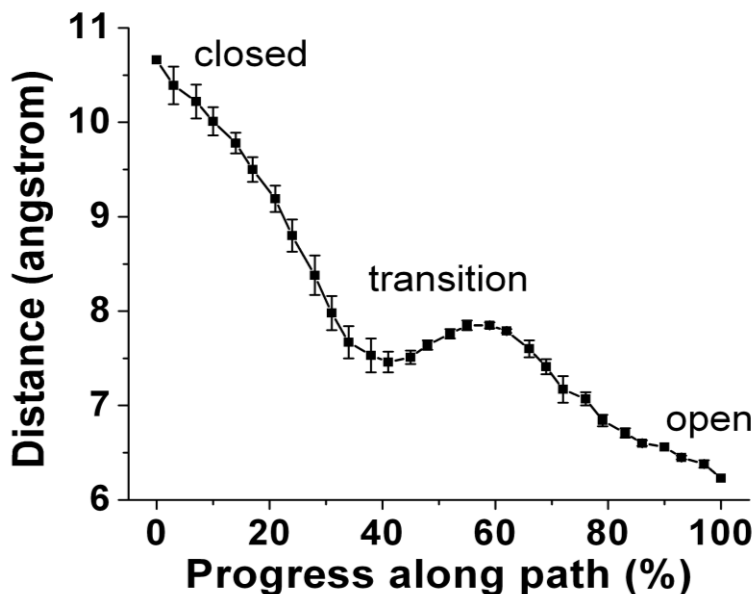
Figure 4.11. Backbone distance betweenV203 and I215. These helices come closer during loop opening, but a steric clash in the transition state causes the path to deviate from linear behavior near the midpoint.

4.3.6 Removal of steric hindrance results in loss of transition state barrier

To confirm our hypothesis that steric clash between residues V203 and I215 controls conformational changes, we performed *in silico* mutagenesis studies. If our hypothesis about the steric clashes between residues M199, V203, and I215 is true, then decreasing the steric clashes by replacing residues V203 and I215 with a smaller side chain, such as alanine, would decrease the energy barrier at the transition state.

We repeated our wild-type simulation protocol for I215A (denoted as InhA^{I215A}:NAD$^+$:**PT70**) and V203A (denoted as InhA^{V203A}:NAD$^+$:**PT70**) mutants bound to **PT70**. In **Figure 4.12**, we show the backbone distance between residues at position 203 and 215 for these mutants. In both cases, a marked difference is seen as compared to the wild type (**Figure**

4.11); when either of the side chains is replaced by alanine, the "hump" in the profile associated with the steric clash disappears.
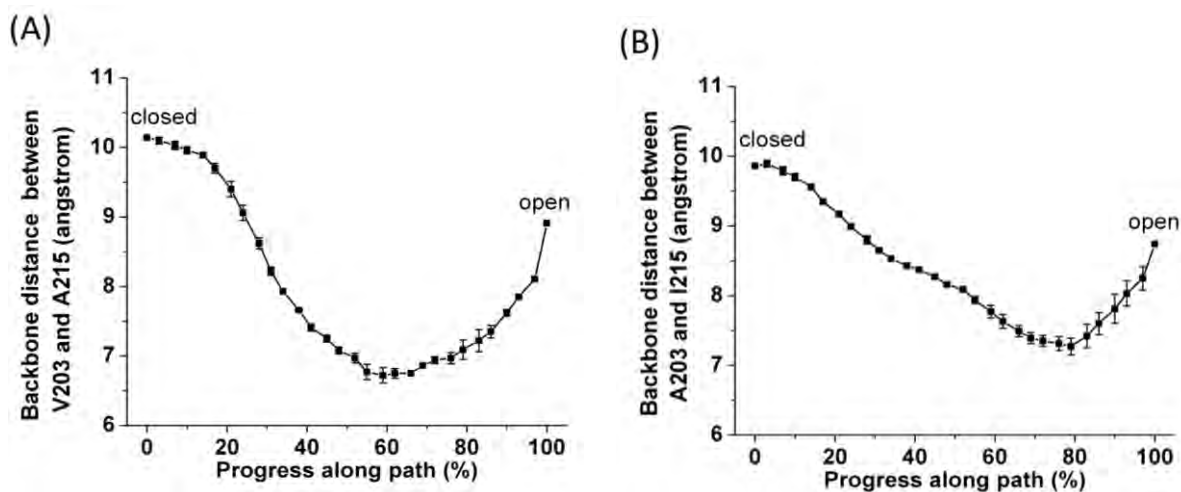
Figure 4.12. Backbone distance between residues 203 and 215. (A) InhA^{I215A}:NAD$^+$:**PT70** complex (B) InhA^{V203A}:NAD$^+$:**PT70** complex. Unlike the wild type, there is no "hump" in the transition state**.**

If this model is correct, we expect to see a reduction in the free energy barrier as a result of this loss of the steric conflict in the transition state for I215A and V203A mutants. As shown in **Figure 4.13A**, the energy landscape of the InhA^{I215A}:NAD$^+$:**PT70** complex is indeed qualitatively flatter compared to the InhA^{WT}:NAD$^+$:**PT70** complex (**Figure 4.9A**); only a 0.34 ± 0.20 kcal/mol and 2.73 ± 0.22 kcal/mol energy barrier remains for open to closed and closed to open directions, respectively (**Table 4.1**). Although the energy barrier becomes much smaller, the complex still prefers the closed state. For the V203A mutant, the energy barrier from open to closed decreased to 1.31 ± 0.95 kcal/mol and the overall energy landscape is flatter compared to

the InhA^{WT}:NAD$^+$:**PT70** complex (**Figures 4.13B and 4.9A**). The energy barrier from closed to

open drops from 6.70 ± 0.41 kcal/mol to 2.52 ± 0.48 kcal/mol (**Table 4.1**).
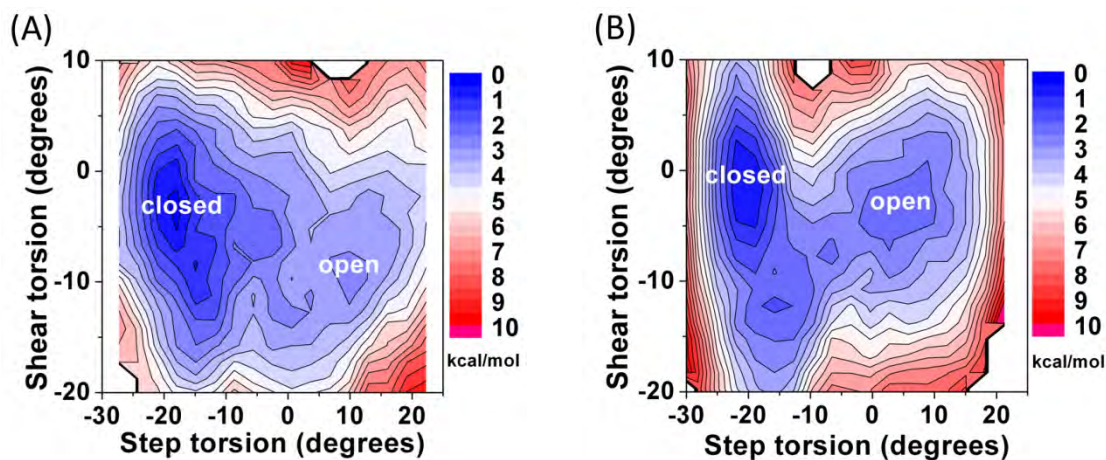
Figure 4.13. PMFs of (A) InhA^{I215A}:NAD$^+$:PT70 and (B) InhA^{V203A}:NAD$^+$:PT70. Both mutants have significantly flatter free energy landscapes than obtained for InhA^{WT}:NAD$^+$:**PT70**.

Table 4.1: Summary of energy barriers.

Complex	Energy barrier[a] (kcal mol^{-1})	
	open to closed	closed to open
InhA^{WT}:NAD$^+$	6.29 ± 0.22	3.32 ± 0.51
InhA^{WT}:NAD$^+$:**PT155**	3.06 ± 0.02	1.29 ± 0.09
InhA^{WT}:NAD$^+$:**PT3**	1.97 ± 0.16	2.57 ± 0.03
InhA^{WT}:NAD$^+$:**Triclosan**	1.04 ± 0.45	1.77 ± 0.51
InhA^{WT}:NAD$^+$:**PT70**	3.26 ± 0.55	6.70 ± 0.41
InhA^{WT}:NAD$^+$:**PT92**	1.63 ± 0.61	6.81 ± 0.80
InhA^{V203A}:NAD$^+$:**PT70**	1.31 ± 0.95	2.52 ± 0.48
InhA^{I215A}:NAD$^+$:**PT70**	0.34 ± 0.20	2.73 ± 0.22
InhA$^{V203F/I215A}$:NAD$^+$:**PT70**	0.29 ± 0.24	7.16 ± 0.24

110

InhA$^{V203L/I215A}$:NAD$^+$:**PT70**	5.11 ± 0.09	3.31 ± 0.25
InhA$^{V203A/I215F}$:NAD$^+$:**PT70**	1.43 ± 0.51	6.32 ± 0.09
InhA$^{V203A/I215M}$:NAD$^+$:**PT70**	0.69 ± 0.48	4.26 ± 0.19
InhA$^{V203A/I215Q}$:NAD$^+$:**PT70**	2.30 ± 0.25	4.80 ± 0.36
InhA^{V203A}:NAD$^+$:**PT162**	1.41 ± 0.77	2.73 ± 0.17
InhA^{I215A}:NAD$^+$:**PT163**	2.24 ± 0.50	6.20 ± 0.03

[a]These energy values were taken from the PMFs by subtracting the transition state and the ground state energies. Statistical error was the difference between two independent runs.

The PMFs give the relative free energy during loop closing. However, changes to the apparent energy barrier could arise either from decreasing the transition state energy, or increasing the ground state energy. Additional calculations are required to determine which of these are responsible for the changes observed in the figures shown. In order to calibrate the two PMFs to the same energy scale and allow direct comparison, we performed thermodynamic integration (TI) to calculate relative binding free energy ($\Delta\Delta G$) between wild-type and mutation InhA, both in the closed state. This allows us to calculate the offset in the loop closing PMFs that were obtained from umbrella sampling. Because the accuracy of TI simulation is sensitive to the length of simulation, we initially performed 500 ps simulation then extended to 1 ns to check the convergence. Convergence of the TI calculations was tested by comparing the 100 ps to 1 ns results. There was no significant change in the free energy after 600 ps (**Figure 4.14**).
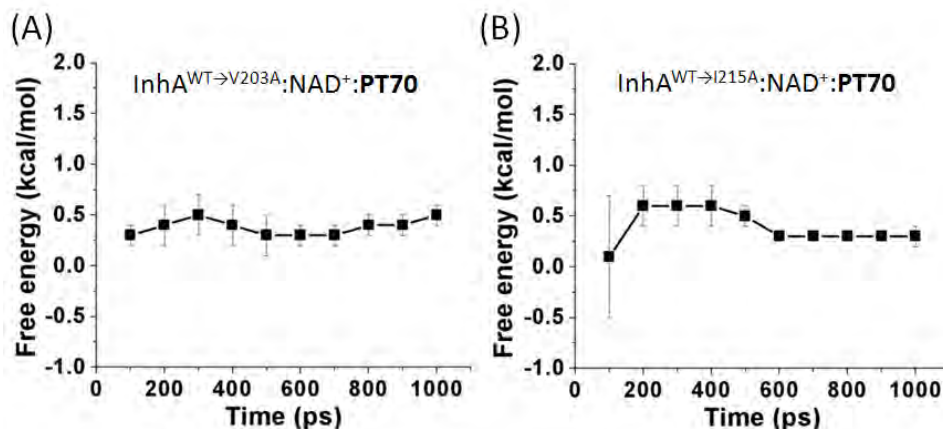
Figure 4.14. Convergence of thermodynamics integration calculation. (A) InhA$^{WT \rightarrow V203A}$:NAD$^+$:PT70 (B) InhA$^{WT \rightarrow I215A}$:NAD$^+$:PT70. Error bar is the difference between two independent runs. There is no significant change in the free energy after 600 ps.

The TI results show that the relative binding free energies of $\Delta\Delta G_{WT \rightarrow I215A}$ and $\Delta\Delta G_{WT \rightarrow V203A}$ increase 0.32 ± 0.07 and 0.52 ± 0.12 kcal/mol, respectively, indicating these mutations modestly destabilize the binding affinity of **PT70** in the closed state (**Table 4.2**). Because the I215 sidechain atoms only interact with the end of the **PT70** alkyl tail, the small change of binding free energy is reasonable for the I215A mutant. The V203 sidechain atoms interact with the diphenyl ether moiety of inhibitor, and mutation has a somewhat larger effect, weakening binding of **PT70** by 0.52 kcal/mol.

Table 4.2. Thermodynamic integration results.

	Step 1 (removing partial charges)[a]	Step 2 (soft-core transformation)[a]	Step 3 (adding partial charges)[a]	Overall (Step1+Step2+Step3)[a]
$\Delta\Delta G_{WT \rightarrow V203A}$ with PT70	0.70 ± 0.01	-0.89 ± 0.16	0.71 ± 0.04	**0.52 ± 0.12**
$\Delta\Delta G_{WT \rightarrow I215A}$ with PT70	0.03 ± 0.02	0.06 ± 0.04	0.24 ± 0.05	**0.32 ± 0.07**

[a]1 ns TI results.

Combining the umbrella sampling and TI results, the data suggest that the I215A and V203A mutants change mainly the transition state. By measuring the average step and shear torsions of the PNEB trajectories and obtaining the corresponding free energy values from the 2D PMF, we were able to get a 1D PMF along the open-closed path. **Figure 4.15** shows the corresponding spots in the 2D PMF where the free energy values were extracted for 1D PMF.
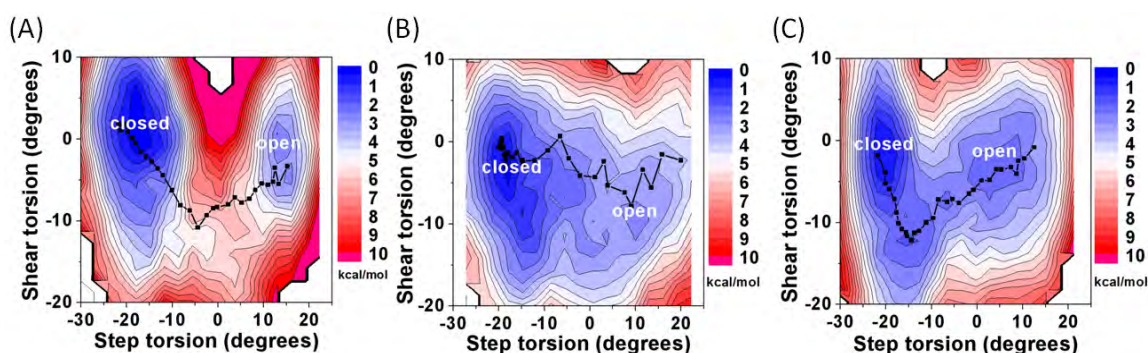
Figure 4.15. The corresponding spots where free energy values were extracted for 1D PMF.
(A) Wild type (B) I215A mutant (C) V203A mutant.

We calibrated the offset of mutants in closed state from the results of TI (wild-type closed state was set to 0) and plot a 1D PMF to map the relative free energy between wild-type and mutants. 1D PMF shows that the I215A and V203A mutations stabilize ~3.0 kcal/mol at the transition state and destabilize ~0.5 kcal/mol at the ground state (**Figure 4.16**). Because the I215A and V203A mutants predominantly decrease the energy barrier at the transition state rather than change of the ground state, this reinforces our hypothesis that the steric clashes between V203 and I215 control the closed to open conformational changes.
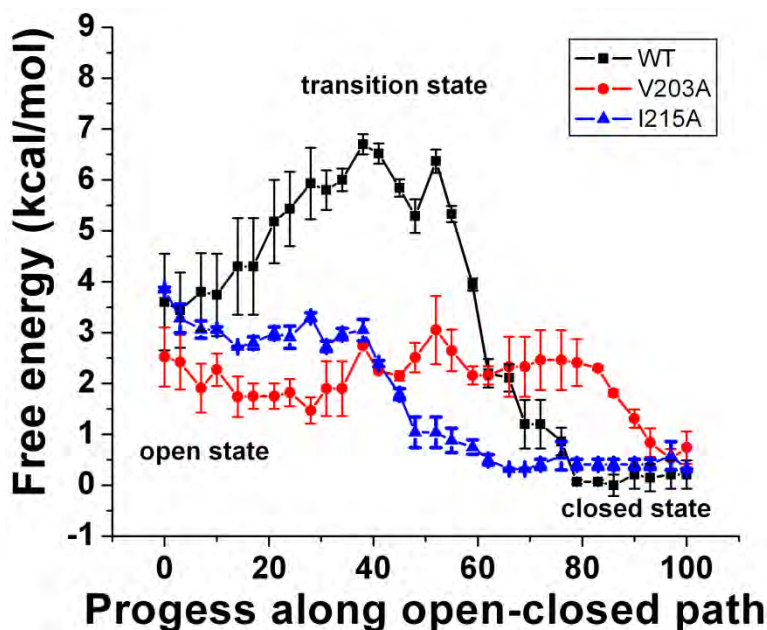
Figure 4.16. 1D PMF of wild-type, I215A, and V203A InhA with PT70. Mutations predominantly stabilize the transition state. Statistic error was the difference between two independent runs.

4.3.7   Regain of energy barrier

We hypothesized the steric clashes of side chain atoms between residues V203 and I215 control the conformational changes of InhA. We confirmed that replacement of residues V203 and I215 with alanine decreased the steric hindrances, and resulted in decrease of energy barrier. If this hypothesis is correct, the reduced energy barrier in the mutants should be recoverable by restoring steric clashes at the transition state. We approached this goal in two ways. First, we hypothesized that if making one side chain of the pair smaller abrogated the energy barrier, increasing the size of the partner side chain might restore the clash in a double mutant. Secondly, we hypothesized that bulky substituents on the alkyl tail of the inhibitor would be able to reach the region where the steric clash occurs in the wild-type InhA and compensate for the smaller

114

side chain in the mutant. If this model is correct, we expect inhibitors modified in this manner to exhibit energy barrier for the mutant InhA.

We explored several double mutants to recover the energy barrier. For the V203A mutant, we changed I215 to the larger Phe (InhA$^{V203A/I215F}$), Met (InhA$^{V203A/I215M}$), and Gln (InhA$^{V203A/I215Q}$). The PMFs of InhA$^{V203A/I215F}$:NAD$^+$:**PT70** and InhA$^{V203A/I215M}$:NAD$^+$:**PT70** complexes show a <1 kcal/mol increase in the energy barrier from open to closed (**Figures 4.17A, 4.17B** and **Table 4.1**). The InhA$^{V203A/I215Q}$:NAD$^+$:**PT70** complex recovers the highest energy barrier among the double mutants and has 2.30 ± 0.25 kcal/mol energy barrier from open to closed (**Figure 4.17C** and **Table 4.1**). However, it is still smaller compared to the wild type. The possible reason is that the wild type clash pair has two β-branched side chains and is difficult to avoid the rotamer changes.

For the I215A mutant, we also changed its clash partner V203 to Phe (InhA$^{V203F/I215A}$). However, the PMF in this case shows no significant energy barrier from open to closed (**Figure 4.17D** and **Table 4.1**). In addition, we changed V203 to Leu (InhA$^{V203L/I215A}$) for I215A mutant, but the PMF shows that InhA$^{V203L/I215A}$:NAD$^+$:**PT70** complex favors the open rather than closed state (**Figure 4.17E** and **Table 4.1**). The results of double mutants suggest that the energy barrier at transition state can be partially recovered if new steric clash is reintroduced during the loop closing step.
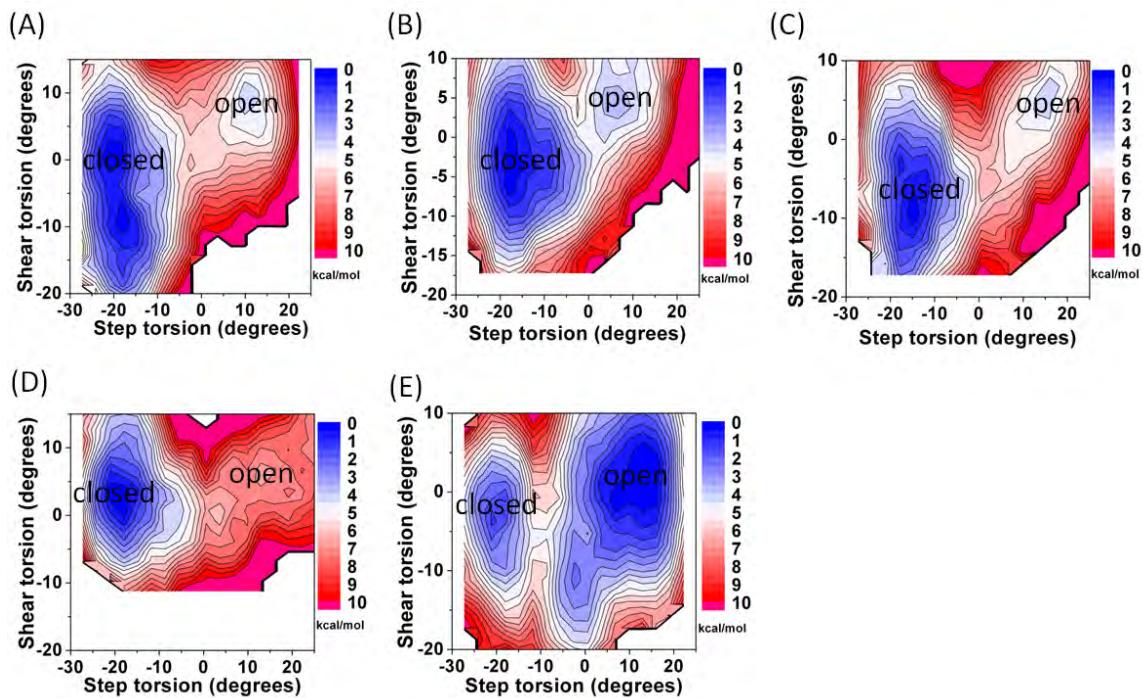
Figure 4.17. PMF of double mutants. (A) InhA$^{V203A/I215F}$:NAD$^+$:**PT70** (B) InhA$^{V203A/I215M}$:NAD$^+$:**PT70** (C) InhA$^{V203A/I215Q}$:NAD$^+$:**PT70** (D) InhA$^{V203F/I215A}$:NAD$^+$:**PT70** (E) InhA$^{V203L/I215A}$:NAD$^+$:**PT70**.

Although we attempted to recover the energy barrier at transition state by double-mutant designs, the open to closed energy barrier is never as high as the wild type (**Figure 4.9A and Table 4.1**). Nevertheless, in most drug discovery programs it is unrealistic to change the protein sequence to make inhibitors perform better. We therefore hypothesized that we could apply the same rationale that worked for the side chains, but applied to the ligand in order to recover a steric clash at the transition state. We speculated that a compound fit in the closed state and create steric clashes with residue V203/A215 would be an ideal compound to recover energy barrier. After examining the structure and location of the inhibitor in the transition state, we focused on the alkyl tail of **PT70** where an appropriately bulky group could crowd the binding loop during closing and compensate for the smaller side chain in the mutant enzyme. We

116

selected the inhibitor 5-((4-phenyl-1H-1,2,3-triazol-1-yl)methyl)-2-(o-tolyloxy)phenol (denoted **PT163**), intended to recover transition state barrier for I215A. In **Figure 4.18**, we show **PT163** docked to closed InhA^{I215A}; as expected, the phenyl group on the tail region fits into the empty space at A215 near V202 and L218. We expected that the phenyl moiety could play a role in producing a steric conflict with V203 in the I215A mutant, analogous to I215 in the wild type. We repeated the computational protocol described above using the I215A mutant with PT163 (InhA^{I215A}:NAD$^+$:**PT163**). As expected, the barrier between closed and open forms of the I215A mutant is much higher with **PT163** than for **PT70** (**Figure 4.19A and Table 4.1**). Further evidence that the energy barrier arises from the designed effect is shown in **Figure 4.19B**, where we show that the profile for the distance between V203 and A215 is comparable to **PT70** with the wild type InhA that a "hump" appears at the transition state.
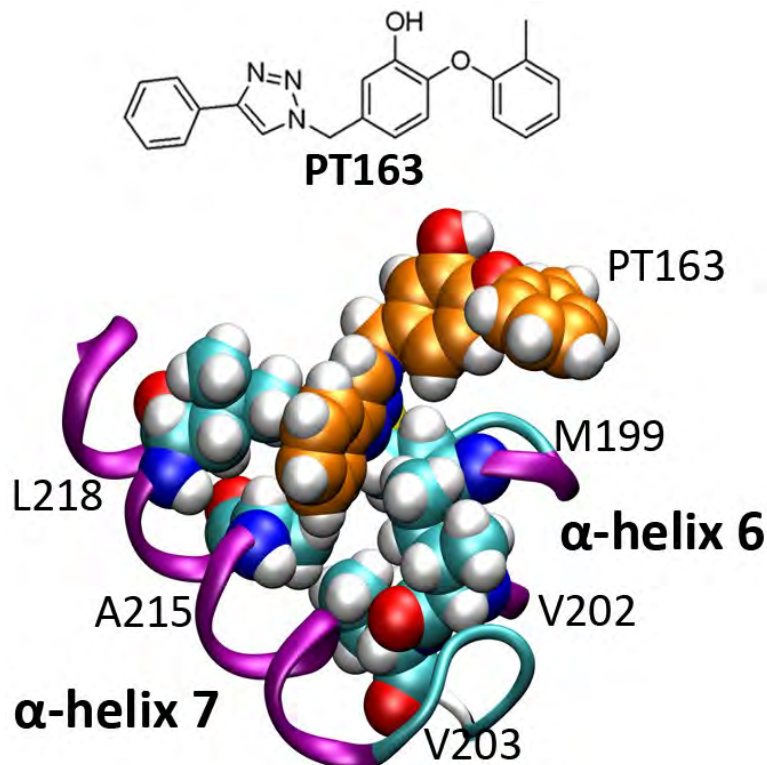
Figure 4.18. Docked model of PT163 (orange) showing interactions with the binding loop of InhA[I215A].
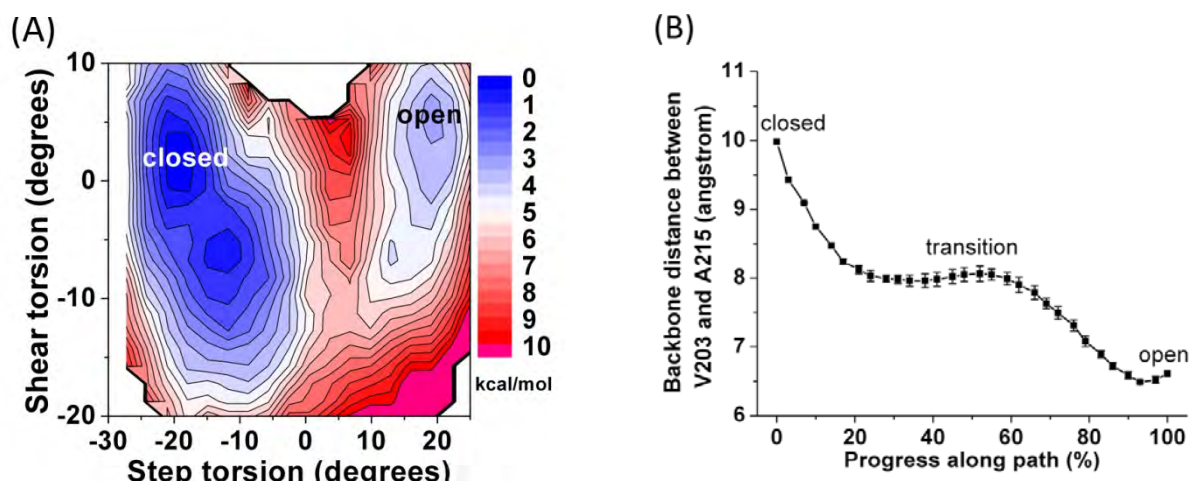
Figure 4.19. Simulation results for loop closing in the InhA[I215A]:NAD[+]:PT163 complex. (A) PMF profile (B) Backbone distance between residues V203 and A215. Both plots are similar to those obtained with **PT70** with wild type InhA, suggesting that the modification in **PT163** restores the energy barrier that was lost in the I215A mutant with **PT70**.

118

We speculated the bulky tail of **PT163** could also change the binding at the ground state since this larger inhibitor can make more extensive contact with the surrounding residues. To examine how **PT163** changes the binding at the ground state, we calculated the relative binding free energy between **PT163** and **PT70** in I215A mutant using TI approach. The TI result shows that the overall binding free energy of PT163 weakens by 5.58 ± 0.76 kcal/mol compared to **PT70** at the closed state (**Figure 4.20A** and **Table 4.3**).
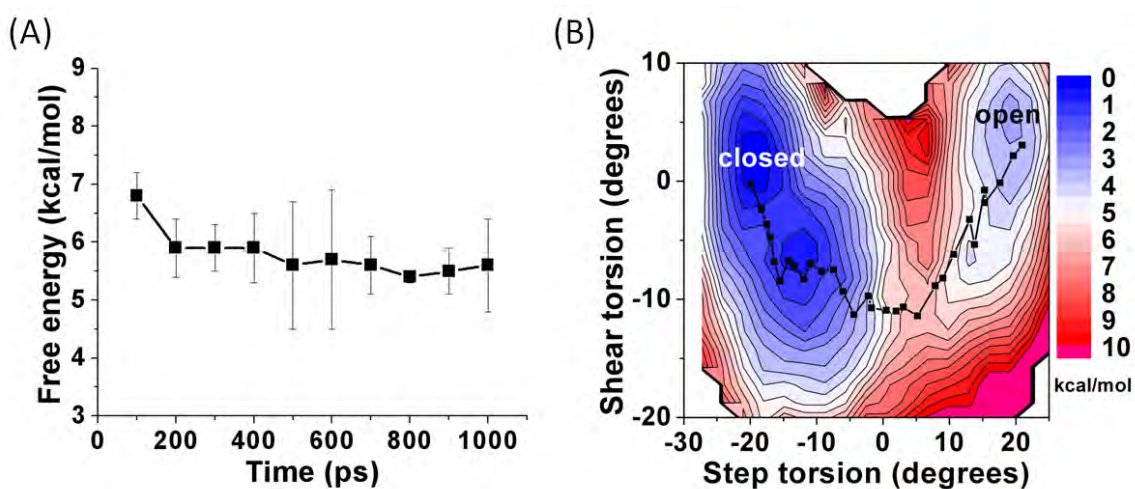
Figure 4.20. (A) 100 ps to 1 ns InhA:NAD$^+$:PT70→PT163 TI results and (B) The corresponded spots in the InhA^{I215A}:NAD$^+$:PT163 where the free energy values were extracted for 1D PMF.

Table 4.3. Thermodynamic integration results.

	Step 1 (removing partial charges)[a]	Step 2 (soft-core transformation)[a]	Step 3 (adding partial charges)[a]	Overall (Step1+Step2+Step3)[a]
$\Delta\Delta G_{PT70\rightarrow PT163}$ in I215A mutant	1.45 ± 0.06	-0.83 ± 0.46	4.96 ± 0.15	**5.58 ± 0.76**

[a] 1 ns TI results.

119

Combining the umbrella sampling and TI results suggests that the **PT163** indeed increases the energy barrier at the transition state in the I215A mutant. The 1D PMF shows that the PT163 destabilizes ~8.6 and ~5.6 kcal/mol at the transition and ground states, respectively (**Figure 4.21**), confirming the design goal; the bulky group indeed has a larger (3.0 kcal/mol) impact on the energy barrier than on binding affinity. Again, these results reinforce our confidence that a steric clash between the two helices of the active site loop is the source of the barrier controlling conformational changes between open and closed states, and that this mechanism can be explained productively through future, more detailed rational ligand design efforts.
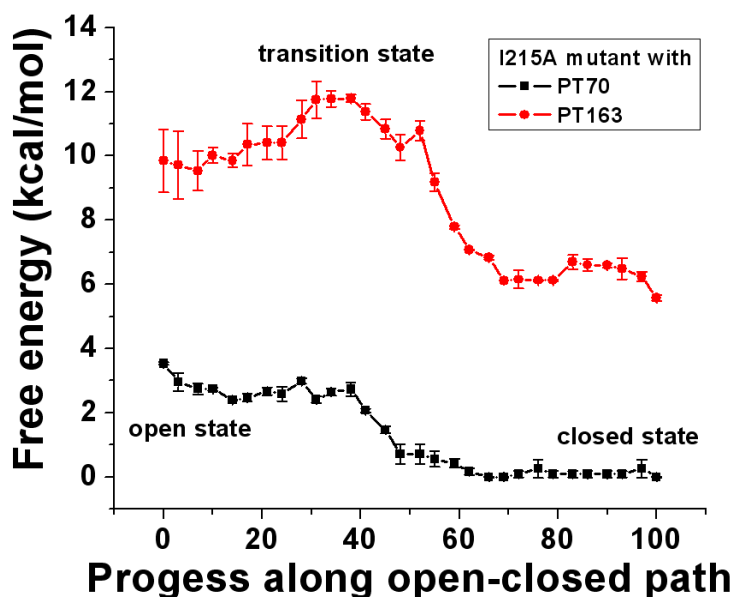
Figure 4.21. 1D-PMF of I215A with PT70 and PT163. By comparing to **PT70**, **PT163** destabilizes the transition state to a significantly larger amount than the ground states. Statistic error was the difference between two independent runs.

We attempted to dock several designed inhibitors to the V203A mutant to position a bulky group near the A203 side chain. However, A203 is ~ 8 Å away from the tip of alkyl tail of **PT70** and it was difficult to design a functional group to precisely fill the space. As an alternate approach, we selected 5-phenethyl-2-(o-tolyloxy)phenol (**PT162** and **Figure 4.22**) and speculated that constraining residues M199 and I215 with the phenethyl group could be an alternative way to create a steric clash during the transition. We repeated the simulation protocol for the V203A mutant with PT162 (InhA^{V203A}:NAD$^+$:**PT162**). The PMF was consistent with our expectation of a less-flat energy landscape; the energy barrier from open to closed (1.41 ± 0.77 kcal/mol) is modestly higher than the InhA^{V203A}:NAD$^+$:**PT70** complex (**Figure 4.23** and **Table 4.1**). However, because the closed and open states have similar energy and the barrier between them is relative small, the complex is unlikely to exhibit slow-onset inhibition behavior.
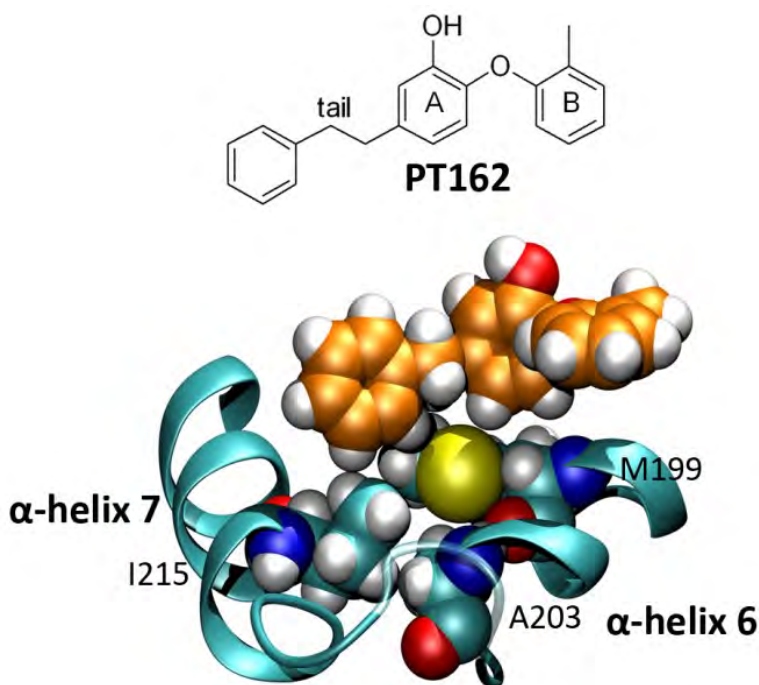
Figure 4.22. Docked model of PT162.
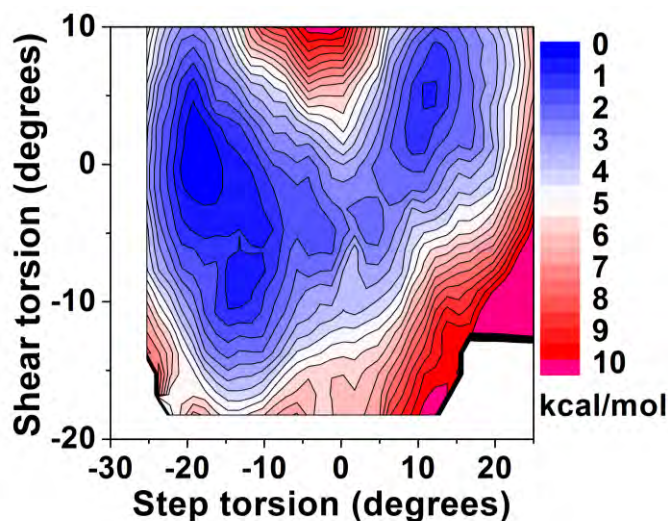
Figure 4.23. The PMF of InhA$^{\text{V203A}}$:NAD$^+$:PT162 complex. The PMF has a modest incensement of energy barrier in the closed to open path.

4.3.8   Transferability of the model to other pathogens

We predict a similar potential for steric hindrance, and thus opportunity for slow-onset inhibition, in other FabI enzymes. We superimposed InhA with other FabIs (*E. coli (24)*, *Francisella tularensis (158)*, *Staphylococcus aureus (159)*, *Bacillus subtilis (160)*, *Helicobacter pylori (161)*, and *Bacillus anthracis (162)*) and found conserved residues Ala and Phe corresponding to the residues V203 and I215 of InhA (**Figure 4.24**). Similar to InhA, these FabI enzymes have closed and ordered α-helix 6 to represent the closed state. However, helix-6 appears disordered in the open state. This is likely because these FabIs have shorter active-site helix 6, and when the active site opens, helix-6 loses stabilizing contacts. Because the Ala and Phe are conserved, we speculate that steric clashes between residues F203 and A197 (using *E. coli* residue numbering) might produce slow-onset inhibition in these FabIs.
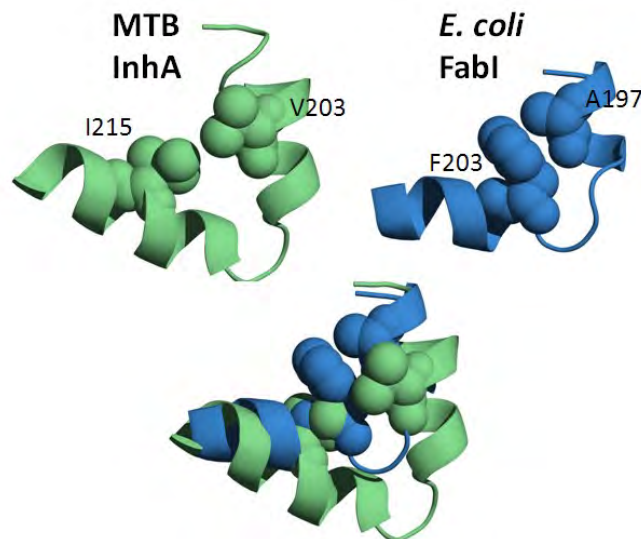
Figure 4.24. The superimposed structure of MTB InhA and *E. coli* FabI.

4.3.9   Energy profile of *Staphylococcus aureus* FabI

Staphylococcus aureus is a dangerous pathogen to immunocompromised patients in hospital (*163, 164*). *S. aureus* infection becomes a serious problem since the methicillin-resistant strains were found (*165*). There is an emerging need for new drugs that target *S. aureus.* FAS II pathway is a promising target since humans do not have this long chain fatty acid synthesis pathway. Among the FAS II enzymes, FabI (homologue of InhA) enzyme has been most studied in terms of drug discovery. Currently, there are three saFabI inhibitors in clinical trials (*166-168*).

Slow-onset inhibition property is also found in *Staphylococcus aureus* FabI (saFabI). Previous studies indicated some diphenyl either compounds are slow-onset inhibitors of saFabI (*152, 169, 170*). The active-site helix-6 dis/ordered scenario is also observed in the saFabI. The apo crystal structure has a disordered active-site helix-6, suggesting this helix-6 is highly flexible in the absence of ligand. On the other hand, when saFabI is bound with a slow-onset inhibitor

PT4, the helix-6 region forms an ordered helical conformation. Because the open-closed conformational change in the InhA is hypothesized to be correlated with the slow-onset inhibition, it would be interesting to know whether the saFabI enzyme also has similar structural mechanism of slow-onset inhibition.

Because the NEB simulation setup requires two endpoints structure and there is no available open conformation of saFabI, we modeled the open active-site conformation from the closed conformation. We used steered MD simulation approach to push helix-6 away from strand-4 to generate an open active-site. As mentioned in **4.3.8**, FabI has a short helix-6, it is expected this helix will become distorted when the active-site is open. The steered MD result, as expected, helix-6 had a distorted helical conformation (**Figure 4.25**).
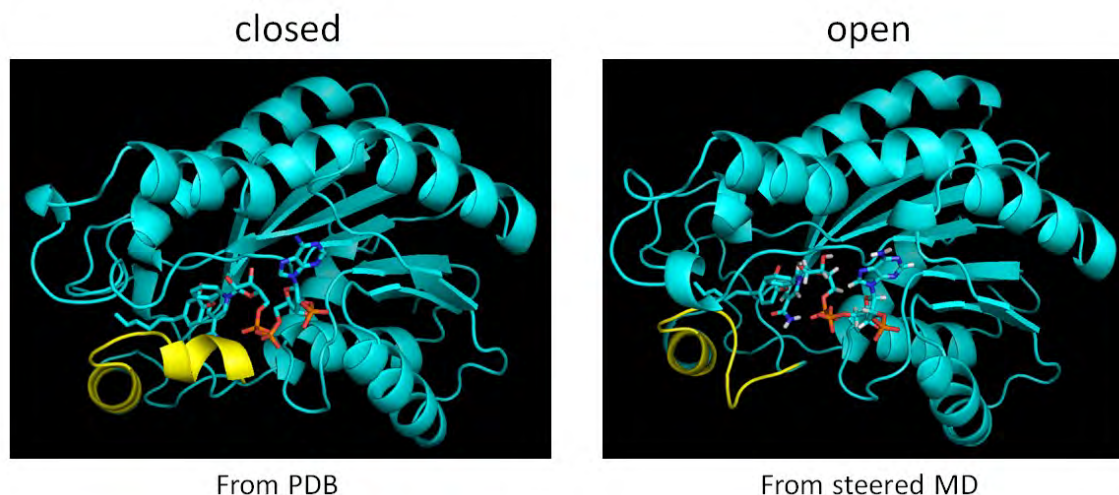
Figure 4.25. closed and open conformation of saFabI. The open active-site has a distorted helix-6.

We adapted the InhA PNEB/US protocol for the saFabI simulations. The PNEB approach was used to generate low energy path structures between the open and closed states followed by

umbrella sampling to get the free energy profiles. The closed conformation was taken from the crystal structure (PDB ID: 4BNH) while the open conformation was taken from the above steered MD result. From the enzyme kinetics study, **PT4** and **PT1** are slow-onset inhibitors of saFabI with a residence time of 460 and 80 minutes, respectively (*152*). If the open-closed conformational change is the rate limited step where experiments measured, and the reaction coordinates used in the umbrella sampling properly captures this motion, we expect these complexes will have a slow-onset inhibition PMF – the closed state is the global minimum and an energy barrier in between closed and open states.

The PMFs of saFabI:NADP$^+$:**PT4** and saFabI:NADP$^+$:**PT1** complexes had a slow-onset inhibition energy landscape. The PMF of saFabI:NADP$^+$:**PT4** complex demonstrates that the closed state is the global minimum and an energy barrier between closed and open state (**Figure 4.26A**). The energy barriers from closed to open and open to closed are ~4 kcal/mol and ~2 kcal/mol, respectively. These values are smaller than the InhA:NAD$^+$:**PT70** complex, which are 6.70 kcal/mol and 3.26 kcal/mol, respectively. To know whether other slow-onset inhibition complexes also have similar energy profile, we performed same simulation on another slow-onset inhibitor **PT1**. From the enzyme kinetics study, **PT1** has a residence time of 80 minutes that is 1/6 time shorter than **PT4** (*152*). Because **PT1** has smaller residence time than **PT4**, we expect this complex would have smaller energy barrier in the PMF. Similar to saFabI:NADP$^+$:**PT4** complex, the closed conformation is the global minimum in the PMF of saFabI:NADP$^+$:**PT1** complex, and an energy barrier separates the closed and open states (**Figure 4.26B**). The energy barriers from closed to open and open to closed are ~3.5 kcal/mol and ~1.5 kcal/mol, respectively. Although the difference of energy barrier between **PT4** and **PT1** complex

is only 0.5 kcal/mol, it seems the PMFs have a positive trend: longer residence time inhibitor has higher energy barrier.
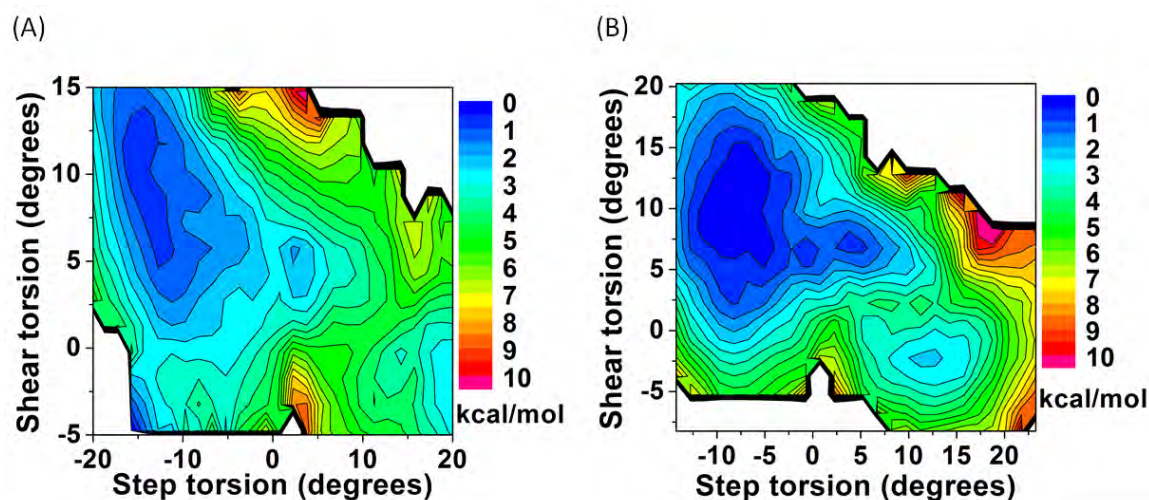
Figure 4.26. The PMFs of (A) saFabI:NADP$^+$:PT4 and (B) saFabI:NADP$^+$:PT1 complexes.

The small difference of energy barrier between saFabI:NADP$^+$:**PT4** and saFabI:NADP$^+$:**PT1** complexes might be due to reaction coordinate problem. It is possible the reaction coordinate adapted from InhA could not capture the open-closed motion in the saFabI system. Another possibility is that the open-closed conformation change is not the rate limited step where kinetics experiment measured. If the later one is the case, the rapid reversible complex would also have similar energy profile as the slow-onset inhibition complex does. To examine whether this is the case as well as to provide a control set, a simulation of rapid reversible complex would be required.

To know the energy profile of a rapid reversible complex, inhibitor **PT166** was selected. According to enzyme kinetics result (unpublished result by Andrew Chang), **PT166** is a rapid

126

reversible inhibitor of saFabI. As the PMF shown in **Figure 4.27**, the closed state is still the global minimum, suggesting this complex still favors the closed state. The energy barrier from closed to open is ~2.5 kcal/mol while the open to closed path is ~1 kcal/mol. There is a ~1 kcal/mol energy barrier difference between the slow-onset saFabI:NADP$^+$:**PT4** and rapid reversible saFabI:NADP$^+$:**PT166** complexes in the open-to-closed path. By comparing all saFabI PMFs, it seems the energy profiles have a good trend between slow-onset and rapid reversible complexes. However, because the ~1 kcal/mol small energy difference between slow-onset and rapid reversible complexes, we cannot conclude whether open-closed conformational change is correlated with slow-onset inhibition in the saFabI enzyme. For future study of saFabI, better reaction coordinates will be required.
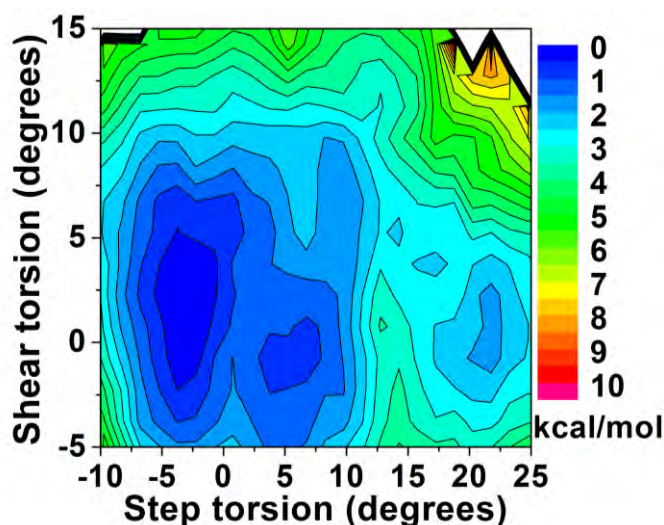
Figure 4.27. PMF of saFabI:NADP$^+$:PT166 complex.

4.3.10 Perspective and limitations for designing new InhA inhibitors

For future InhA lead optimization, especially for tight binding inhibitors, it is very challenging to increase the residence time by solely considering improving the binding affinity. Increase of the free energy barrier by introducing unfavorable interactions along the active-site loop closing step may be a more feasible approach. Although one can apply the protocol we present here to screen candidate inhibitors for significant energy barriers prior to *in vitro/vivo* experiments, one should also consider the limitations of this approach.

The first limitation is that umbrella sampling requires well-defined reaction coordinate(s). It is always challenging to define complex dynamics using a small number of descriptors, since molecular motion can be inherently high dimensional. If the defined reaction coordinate(s) do not accurately describe/capture the motion that corresponds to the energy barrier, the umbrella sampling results would underestimate the true free energy barrier (*171*). However, if considering the balance between accuracy and computing time of umbrella sampling, it seems less than three reaction coordinates is commonly used currently (*172*). Furthermore, the energy barriers obtained from the PMF do not directly correspond to rate constants, and more complex approaches would be needed to more accurately predict kinetics (*173, 174*). Thus, we use our umbrella sampling results as a qualitative guide.

A second limitation in using a transition state model to predict changes to residence times is that the closing pathway, and hence the transition state may change with different inhibitors. If considering the roughness of energy landscape in a higher dimension, there will be multiple paths for the active-site loop closing. Raising the barrier on a given path may lead to the loop

closing via a different path, potentially needing reevaluation of the reaction coordinate descriptors.

Third, it is always very challenging to increase energy barrier at the transition state without altering the ground state for the active-site loop closing. Careful structural visualization should be performed to understand whether such effects may occur; as described above, binding affinity calculations can be used to separate these two effects (ground state vs. transition state), but these calculations can be computationally intensive.

4.4 Summary

In summary, we show here that computationally mapping the pathway of loop closing allows us to obtain models for the transition state. The free energy profiles demonstrate that rapid reversible inhibitor complexes exhibit two types of PMF, either preferring the open state, or having little preference and low energy barrier between the open and closed states. On the other hand, slow-onset inhibition complexes have a relative stable closed state with a more significant energy barrier between the open and closed states. These PMF profiles support the induced-fit models which the enzyme inhibitor initial form in an open state, followed by an isomerization step, leading to a more energetic favorable closed state. These open and closed states are matched well to the two-step slow-onset binding mechanism, where open and closed conformations represent the EI and EI* states, respectively.

Analyses of these structures provide insight into the specific interactions that modulate the energy barrier. We hypothesized that a steric clash between side chains on 2 helices that pass closely during loop closing are responsible for the binding kinetics. The model suggests that

129

replacement of residues V203 and I215 with amino acids possessing smaller side chains would remove the steric clash. Computational mutagenesis studies of the InhA^{V203A} and InhA^{I215A} mutants substantiated the expected decreases in the free energy barrier.

To confirm that the pair interaction is important, we determined that increase in size of one clash partner can compensate somewhat for reduced size in the other; the double mutant V203A/I215Q shows increased computational barrier heights as compared to the single V203A mutant. Further evidence of the validity of this model is provided by rationally redesign inhibitor **PT163** to restore the energy barrier as the designed inhibitor compensates the reduced clash partner V203/I215A and recreates steric clashes in the loop closing path in the mutant InhA.

The ability to rationally remove energy barrier through a protein mutation, partially recover it with a double mutant, and then selectively restore it in the single mutants with a ligand modification strongly supports the model of the steric clash mechanism that a crowded active-site environment created by inhibitor, particularly restraints on residues M199, V203A, A211 and I215, controls the loop closing step. We conclude that steric clashes between the active site loop helices control the conformational changes between open and closed states of InhA and this information can be used for future inhibitor design to increase the residence time of receptor-inhibitor complex, such as using portion of inhibitor to constraint the steric clashes. Examination of homologous FabI sequences in other pathogens reveals similar interactions between the corresponding loops, suggesting that this may be a general approach for developing slow-onset inhibitors of this critical pathway.

Chapter Five

Virtual screening of new slow-onset InhA inhibitors

5.1 Introduction

Slow-onset inhibitors are of particular interests in drug discovery programs as the slow dissociation of the inhibitor from the target-inhibitor complex prolongs target occupancy and improves *in vivo* efficacy. To understand the role of residence time in modulating inhibitor activity in the InhA system, inhibitors of different residence time ranging from minutes to hours will be required. However, this is limited by the small numbers of known InhA inhibitors. In this chapter, we tried to use computational approaches to screen new slow-onset inhibitors of InhA. As mentioned in chapter four, a slow-onset inhibition complex of InhA would prefer closed state and have a high energy barrier between the closed and open states. It would be valuable if we use the resulting PNEB and umbrella sampling (US) protocols to find out some inhibitors with high transition state barrier. In the first part of this chapter, we used the PNEB/US approaches to virtually screen slow-onset inhibitors of InhA. We simulated the **PT163** analogues complexes and some known MIC (minimum inhibitory concentration) compounds.

In the second part of this chapter, we tried to develop a docking protocol to rapidly screen the slow-onset inhibitor. The free energy profile derived from the PNEB/US approach can be used to distinguish rapid reversible and slow-onset inhibitors; however, the PNEB/US approach is computationally expensive and time-consuming. It is unlikely to use this approach to examine every inhibitor we are interested in, thus, a rapid screening approach is needed to facilitate the inhibitor screening process. We thought that the docking calculation has the potential to rapidly

examine the inhibitor's binding free energies at the closed, open, and transition states. To know whether the docking calculation can reproduce the same trend of the binding free energies at the closed, open, and transition states, we compared the dock scores of five slow-onset and five rapid reversible inhibitors calculated by DOCK (*140*) program. The dock scores had the same trend as the umbrella sampling results: the closed state had lowest score among three states, and transition state had higher energy barriers for the slow-onset inhibitors than the rapid reversible inhibitors did. The application of docking score could provide an opportunity to pre-examine inhibitor binding affinity before time-consuming PNEB/US simulations.

5.2 Simulation detail

5.2.1   Initial structures

The InhA:NAD$^+$:**PT70** and InhA:NAD$^+$:**C16-NAC** structures (PDB ID: 2X23 (*13*) and 1BVR (*11*), respectively) were used to build the starting structures of closed and open states. For those inhibitors that do not have crystal structures, the initial poses were obtained from molecular docking with DOCK 6.3 package (*140*). The sphere and grid files were generated from SPHGEN and GRID modules of DOCK 6.3, respectively. The grid file for the grid-based energy scoring had a space of 0.3 Å. Default parameters were used in the flexible docking. We assumed the analogues occupy similar position as **PT70** does, thus the result with a lowest RMSD value in the diphenyl ether moiety was chosen as the initial pose.

5.2.2 Molecular dynamics simulation setup

AMBER ff99SB (*26*) and GAFF (*115*) force field were assigned to the protein and inhibitor, respectively. The force field of cofactor NAD^+ was taken from other studies (*118, 119*). The partial atomic charges of inhibitors were computed using Gaussian98 (*153*) with the HF/6-31G* basis set, followed by RESP fitting (*46, 116*). Each complex was solvated in a truncated octahedral TIP3P (*120*) water box with a minimum distance of 8 Å between the water box edge and solute, resulting in ~ 23,000 atoms in total. SHAKE (*155*) was used to constrain bond lengths involving hydrogen. The particle mesh Ewald method (*156*) was used for calculating electrostatic energy with an 8 Å nonbonded cutoff. Each system was equilibrated initially by high restraints on non-water atoms, and gradually reduced and removed the positional restraints on the backbone atoms using the following procedures. The first step was 10,000 steps of steepest descent minimization with 100 kcal mol^{-1} $Å^{-2}$ restraints on all atoms except water molecules and hydrogen atoms. The second step was heating the system from 100 to 300 K at constant volume over 100 ps with 100 kcal mol^{-1} $Å^{-2}$ restraints on non-water and non-hydrogen atoms, followed by 100 ps with the same restraints at constant 300 K temperature and 1 atm pressure. The third step was 250 ps MD with restraint weight of 10 kcal mol^{-1} $Å^{-2}$ on the non-water and non-hydrogen atoms at constant 300 K temperature and 1 atm pressure. The following steps only restrained the backbone atoms and gradually reduced the restraints weight from 10 to 0.1 kcal mol^{-1} $Å^{-2}$ at constant 300 K and 1 atm. This was carried out by 100 ps with 10 kcal mol^{-1} $Å^{-2}$ restraint, followed by 100 ps with 1 kcal mol^{-1} $Å^{-2}$ restraint, and 100 ps with 0.1 kcal mol^{-1} $Å^{-2}$ restraint.

5.2.3 PNEB simulation

The partial nudged elastic band (PNEB) (*95*) simulation approach was used in this study to generate structure models along the low-energy pathway between the open and closed conformations. The equilibrated open and closed structures were treated as the two end-point structures and 30 windows (including end-points) were used in the simulation. A spring force was applied to the backbone atoms of α-helices 6 and 7 (residues 196 to 223) in the subsequent steps. In the first 40 ps, the system was equilibrated at 300 K with a Langevin collision frequency of 50 ps^{-1} and spring force of 20 kcal mol^{-1} $Å^{-2}$. The next step was 100 ps equilibration at 300 K with a 20 ps^{-1} Langevin collision frequency and a spring force of 75 kcal mol^{-1} $Å^{-2}$. After the conformations were generated along the open-to-closed path, simulated annealing was used to optimize the local energy minimized path. This was done by heating the system from 300 to 375 K gradually over 175 ps, and subsequently cooling back to 300 K gradually over 175 ps with a 20 ps^{-1} Langevin collision frequency and a spring force of 75 kcal mol^{-1} $Å^{-2}$. A 20 kcal mol^{-1} $Å^{-2}$ Cartesian restraint was applied to the backbone atoms from residues 2 to 195 and 225 to 268 to prevent protein unfolding during the heating process. After this simulated annealing, a 600 ps production run was performed with the same Langevin collision frequency and spring force at 300 K. To prevent large conformation changes on regions outside helices 6 and7, a 10 kcal mol^{-1} $Å^{-2}$ Cartesian restraint was applied to the backbone atoms from residues 2 to 195, and 225 to 268 during the production run.

5.2.4 Umbrella sampling

Energy landscapes were obtained by umbrella sampling (US). Two reaction coordinates (step and shear torsions) were used to describe the motion of helices 6 and 7. US windows were spaced at 3º increments in both step and shear torsions. An extra 6º buffer region was added around the grid region sampled in the PNEB simulation, resulting in a total of 192 grid points. Initial structures at each grid point were selected from the PNEB production trajectories with dihedral values (step/shear torsions) closest to the respective grid point. 500 ps MD simulation in the NVT ensemble at 300 K with a Langevin collision frequency of 75.0 ps^{-1} was performed for each grid point. The Weighted Histogram Analysis Method (WHAM) (*98*) approach and analysis program (*96*) was then used to obtain the potential of mean force (PMF) from the umbrella sampling results.

5.2.5 Docking calculation

Two sets of closed, open, and transition state structures were generated (6 receptor structures). The first set of receptor structures were from the **triclosan** bound PNEB snapshots. Another set of receptor structures were from the **PT70** bound PNEB snapshots. The receptor structure of each state was the representative structure based on the clustering analyses. Docking calculation was performed with DOCK 6.3 package (*140*). The sphere and grid files were generated from SPHGEN and GRID modules of DOCK 6.3, respectively. The grid file for the grid-based energy scoring had a space of 0.3 Å. Default parameters were used in the flexible

docking. We assumed the inhibitors occupy similar position as **PT70** does, thus result with a lowest RMSD value in the diphenyl ether moiety was chosen as the correct pose.

5.3 Results and Discussion

Slow-onset inhibitors are of particular interest in drug discovery programs as the slow dissociation of the inhibitor from the target-inhibitor complex prolongs target occupancy and improves *in vivo* efficacy. In order to validate the role of residence time in modulating inhibitor activity in the InhA system, inhibitors with different residence time are needed. There are only a few known slow-onset inhibitors of InhA. Computational simulation could assist the inhibitor screening process by calculating the inhibitor's energy profile. As shown in Chapter four, a slow-onset inhibition complex of InhA shall prefer closed state and have an energy barrier between the closed and open states. The resulting PNEB and umbrella sampling (US) protocols have the potential to screen slow-onset inhibitors of InhA. We first tried to screen some **PT163** analogues (**Table 5.1**), followed by some known MIC (minimum inhibitory concentration) compounds.

Table 5.1. PT163 analogues.

Code	Structure	Code	Structure
PT163		SB94	
SB92		SB101	
SB93		SB102	

5.3.1 Energy profiles of PT163 analogues

In chapter 4, we showed that **PT163** recovered the transition state barrier for the I215A mutant by introducing a bulky functional group which mimics the sidechain of residue I215 to create steric clashes with residue V203. For the wild-type InhA, **PT163** is a good starting structure for designing inhibitor with higher transition state barrier. In principle, an appropriate size of bulky tail could create a higher energy barrier. To know how the bulky tail tunes the transition state barrier, we simulated different **PT163** analogues (**Table 5.1**) bound complexes.

We first simulated the **PT163** bound with wild-type InhA. The PMF of InhA^{WT}:NAD$^+$:**PT163** complex exhibited a slow-onset inhibition profile – a preference for the closed state with an energy barrier at the transition state (**Figure 5.1**). The energy barriers in the open-to-closed and closed-to-open paths are ~2.2 kcal/mol and ~6.7 kcal/mol, respectively. The energy barrier from closed to open is similar to the InhA^{WT}:NAD$^+$:**PT70** complex, but the open-to-closed path is ~1 kcal/mol less than the InhA^{WT}:NAD$^+$:**PT70** complex (**Figure 4.9A and Table 4.1**).
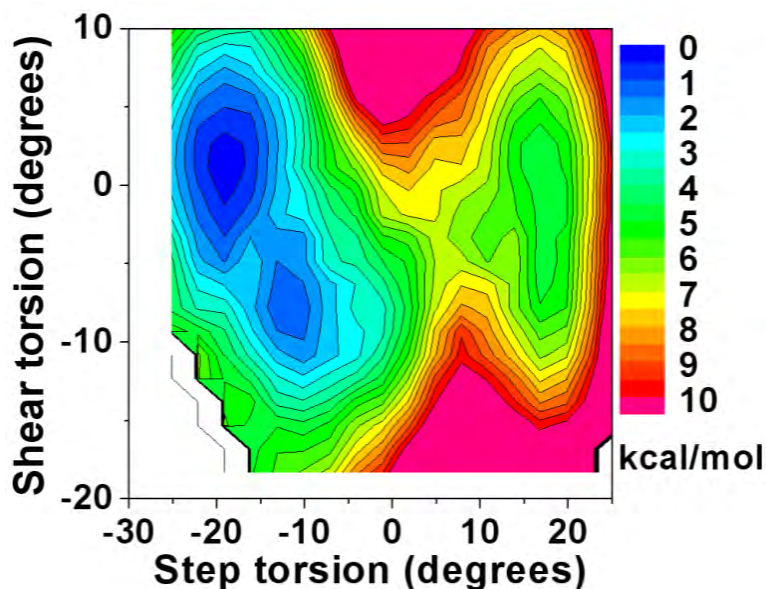
Figure 5.1. PMF of InhA^{WT}:NAD$^+$:PT163 complex. The PMF exhibits a slow-onset inhibition profile.

To understand the change of relative binding affinity between inhibitors **PT163** and **PT70** as well as to calibrate the offset of ground state between these two inhibitors, thermodynamic integration (TI) calculation was performed in the closed state. Surprisingly, the closed state destabilized 4.9 ± 0.2 kcal/mol (**Table 5.2**). According to the equation of $\Delta G = -RT$ $\ln(K_i^1/K_i^2)$, this ~5 kcal/mol decrease of binding affinity will increase K_i by ~1000 folds from pM level to nM level. Although InhA^{WT}:NAD$^+$:**PT163** complex exhibits a slow-onset inhibition energy profile, the decrease of binding affinity at the closed state may pose **PT163** a less potent inhibitor of InhA.

Table 5.2. Thermodynamic integration result.

	Step 1 (removing partial charges)[a]	Step 2 (soft-core transformation)[a]	Step 3 (adding partial charges)[a]	Overall (Step1+Step2+Step3)[a]
$\Delta\Delta G_{PT70\rightarrow PT163}$ in wild-type InhA	1.51 ± 0.15	-1.48 ± 0.56	4.90 ± 0.28	**4.93 ± 0.37**

[a]1 ns TI results. Statistical error was the difference between two independent runs.

In order to know how the bulky tail of inhibitor tunes the energy barrier, a series of modifications of the tail were tested (**Table 5.1**). We first tested three smaller size compounds (**SB92**, **SB93**, and **SB94**). Surprisingly, the global minimum shifted from the closed state to the open state in the **SB92** and **SB93** ternary complexes, suggesting these complexes changed mainly the preference for the closed/open conformation rather than the transition state barrier (**Figures 5.2A, 5.2B**). The **SB94** bound complex had equal preference for the closed and open states (**Figure 5.2C**). On the contrary, when the size of tail increased (**SB101** and **SB102**), the open state became less favorable (**Figures 5.2D, 5.2E**). As a result, the open-to-closed barrier became very small or disappeared.
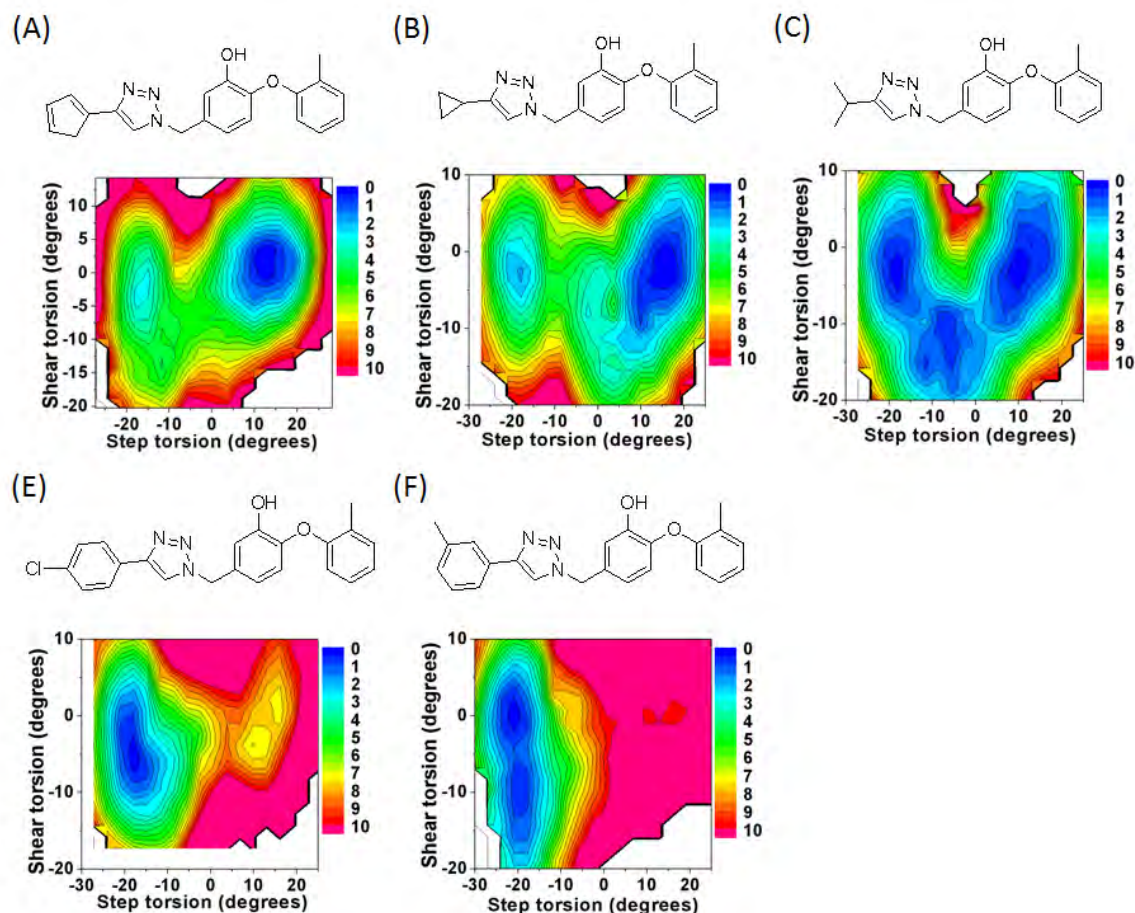
Figure 5.2. PMF of PT163 analogues.

5.3.2 Energy profiles of known minimum inhibitory concentration (MIC) compounds

Since there is an urgent need for the discovery of new anti-TB agents as well as to facilitate TB drug research, several groups have shared their *in vivo* high-throughput screening results to public (*175, 176*). PubChem, a cheminformatics database, stores many biological activities information for small molecules (*177, 178*). There are ~ 10,000 compounds in PubChem that have inhibition activity against TB, however, for most of these inhibitors the targeted enzymes are unknown. In this section, some **PT70**-like known MIC compounds were

selected and tested by the PNEB/US simulations to see whether they exhibit a slow-onset energy profile (**Figure 5.3**). The rationale is that if the compound exhibits a slow-onset inhibition energy profile, there is a high possibility that this compound is a slow-onset inhibitor of InhA.
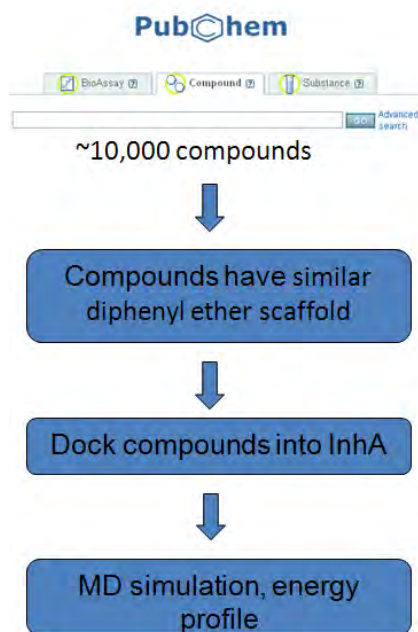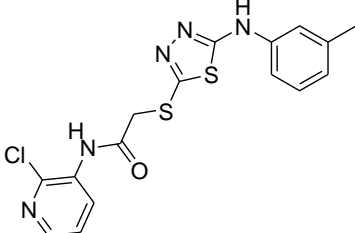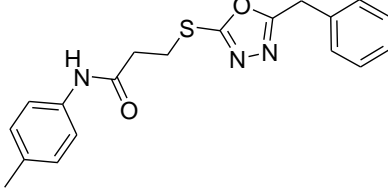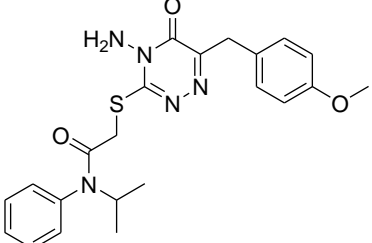
Figure 5.3. Procedure of mining known MIC compounds.

Table 5.3 lists the compounds we selected from the PubChem database. These compounds have MIC_{90} (minimum inhibitory concentration) in μM. We performed PNEB and US simulations for these compounds. The energy profiles demonstrated that inhibitors **CID4851382** and **CID18580849** had high energy barriers in the closed to open path, but low energy barriers in the open to closed path (**Figures 5.4A, 5.4B**), while inhibitor **CID992163** had a low energy barrier between open and closed states (**Figure 5.4C**). Because these PMFs did not exhibit a slow-onset inhibitor energy profile, these compounds are unlikely to be slow-onset inhibitors of InhA.

Table 5.3. Known MIC compounds.

CID No.	Structure	MIC$_{90}$
CID4851382		3 μg/mL (8 μM)
CID992163		10 μg/mL (29 μM)
CID18580849		32 μg/mL (74 μM)
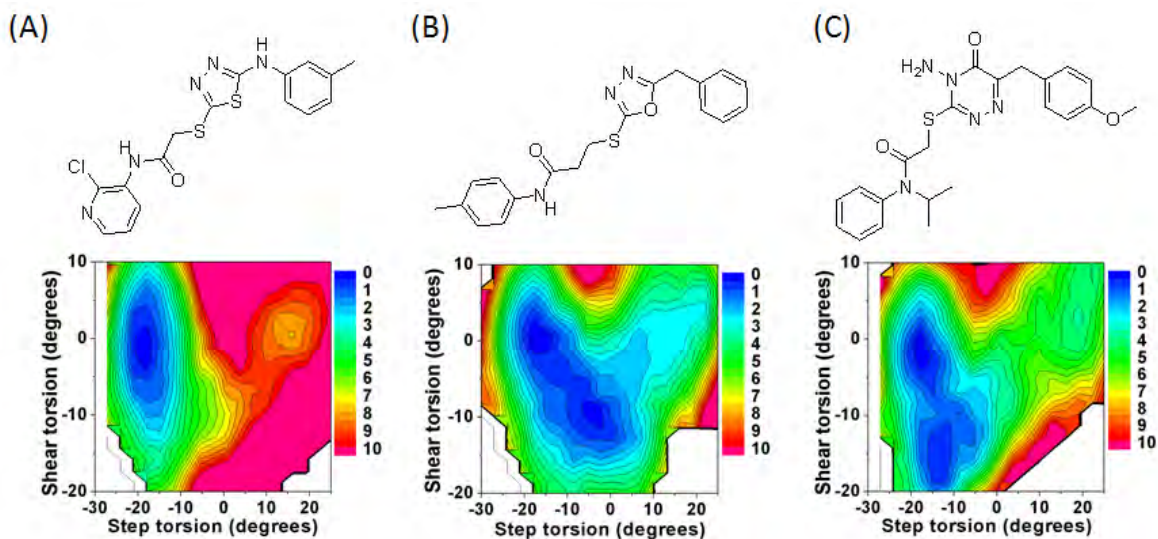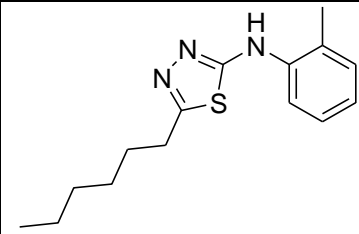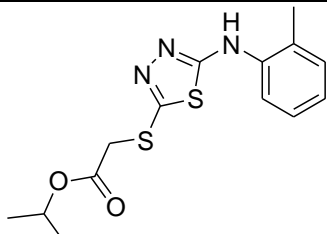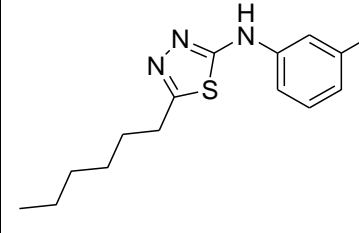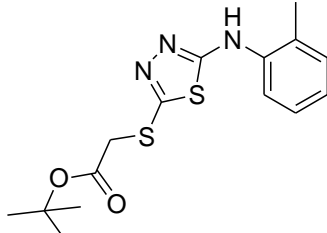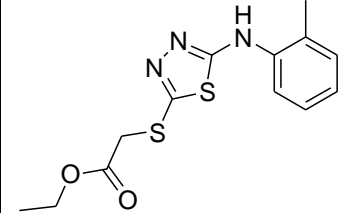
Figure 5.4. PMFs of known MIC compounds. (A) **CID4851382**. (B) **CID992163**. (C) **CID18580849**.

One of the possibilities explains the above compounds not having good PMF could be that the tail region is too bulky, so we tested other compounds with simpler tails to see whether they exhibit a better PMF (**Table 5.4**). Unfortunately, none of these compounds had a good PMF as **PT70** (**Figure 5.5**).

Table 5.4. CID4851382 analogues.

Code	Structure	Code	Structure
PT471		CID7753966	
SB109		CID4963957	
CID4879834			
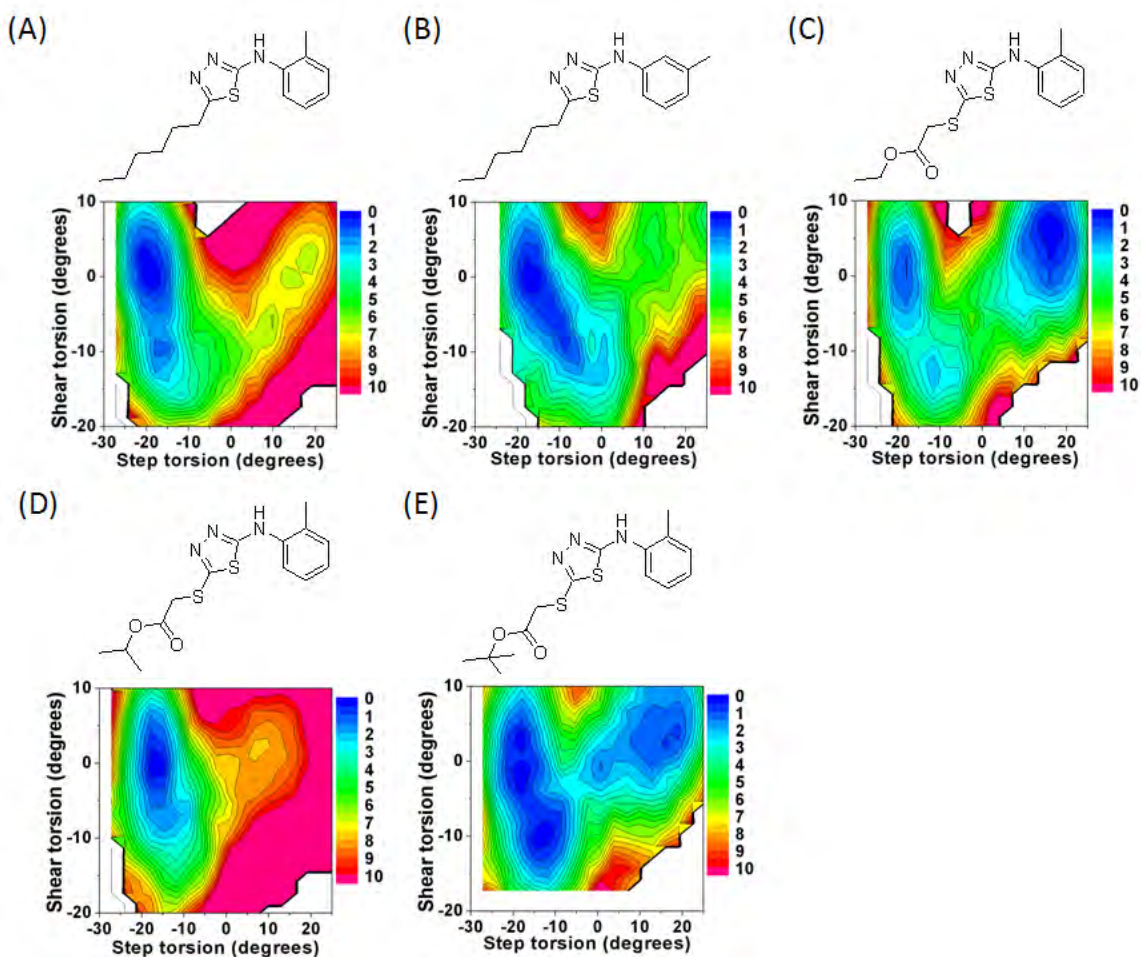
Figure 5.5. PMF of CID4851382 analogues. (A) PT471. (B) **SB109**. (C) **CID4879834**. (D) **CID7753966**. (E) **CID4963957**.

5.3.3 What is missing in terms of inhibitor screening using the PNEB/US approach

The binding free energy calculation of the ground state is required to calibrate the 2D-PMF. As mentioned in chapters two and four, a slow-onset inhibitor of InhA would favor the closed state and have an energy barrier between the closed and open states. A "good" slow-onset inhibitor would have a strong binding affinity at the closed state and high energy barrier between the closed and open states. From the PNEB/US calculations, it only tells the relative energies between the closed, open, and transition states. It is possible that an inhibitor has a slow-onset

inhibition 2D-PMF but the binding affinity of the ground state is weak; in this case, since this inhibitor does not have a good specificity to the target, it is not likely to be a good inhibitor. We can use the thermodynamic integration (TI) calculation to calibrate the offset of the ground state as we did in chapter four. Since the TI calculation is computationally expensive, we did not perform this calculation for each inhibitor during the virtual screening process here. For future virtual inhibitor screening, calculating the binding affinity of the closed state (ie. MM/PBSA method) prior to the PNEB/US calculations will help to eliminate weak binding inhibitors.

For the transition state barrier, we should also consider the limitation of umbrella sampling. As mentioned in chapter four, if the reaction coordinates used during the umbrella sampling cannot accurately capture the motion, the free energy barrier value will be underestimated. Because some of the inhibitors shown in this chapter have different scaffold than the **PT70**, the current reaction coordinate may not be able to capture the rare event during the active-site loop closing. In this case, the umbrella sampling may not deliver the true energy barrier.

5.3.4   Examination of docking protocol

PNEB/US approach can be used to distinguish rapid reversible and slow-onset inhibitors by comparing the energy profiles, but this approach is computationally expensive and time-consuming. It is unlikely to use this approach to examine every inhibitor we are interested. Therefore, a rapid screening approach is needed to facilitate the inhibitor screening process. Docking approach provides a rapid calculation, but this approach is less accurate than the umbrella sampling. However, it will save huge time if we use docking approach to examine

inhibitor before PNEB/US simulation. If the dock scores of the closed/transition/open states are highly correlative to the PMF results, then we can use docking approach to pre-select inhibitors. Here, we aimed to test whether the dock score can reproduce similar trend of binding free energies at the closed, open, and transition states as seen in the PMF.

Because docking approach is highly sensitive to the receptor structure of choice, we tried two different sets of closed/transition/open structures from different inhibitor bound PNEB snapshots. One set of receptor structures were from the **triclosan** bound PNEB snapshots and the other one were from the **PT70** bound PNEB snapshots. We docked a series of rapid reversible and slow-onset inhibitors in these six receptor structures and compared their dock scores.

5.3.3.1 Use of PT70 bound PNEB structures for docking calculation

We first tested the slow-onset inhibitor **PT70** bound PNEB receptor structures. We docked five slow-onset and five rapid reversible inhibitors to these receptor structures and compared their dock scores. For slow-onset inhibitors, the dock scores had a similar trend as seen in the PMF. The dock scores calculated at the closed state are between -67 ~ -70 kcal/mol, while the dock scores at the open state are between -61 ~ -65 kcal/mol (**Figure 5.7A**). The transition states had the highest dock scores (~-55 kcal/mol) among these three states. Overall, these dock scores had a same trend as the PMF, showing the closed state has a lowest free energy value and the transition state has the highest free energy among the three states.

For the rapid reversible inhibitors, the dock scores calculated in the closed state are between -45 ~ -48 kcal/mol, while in the open conformation are between -39~ -45 kcal/mol (**Figure 5.6A**). The transition state had dock scores between -35~-42 kcal/mol. This is similar to

the slow-onset inhibitors, demonstrating that the closed state has lowest grid score, and the transition state has the highest grid score.

To see whether the slow-onset inhibitors have higher transition state barrier than the rapid reversible inhibitors do, we compared the energy difference at the transition state and the ground states. From the PMF observation, the slow-onset inhibition complexes have a relative higher energy barrier compared to that of rapid reversible complexes. By comparing the difference between the transition and ground states of these dock scores, the ΔGridscore (transition state minus closed state (TS-closed) and transition state minus open state (TS-open)) demonstrated that the transition state barriers of the slow-onset inhibition complexes are larger than the rapid reversible inhibition complexes (**Figure 5.6B**). For the slow-onset inhibitors, the average ΔGridscore of TS-open and TS-closed are ~15 kcal/mol and ~9 kcal/mol, respectively. These values are ~4 kcal/mol larger than those of rapid reversible inhibitors, which are 10 kcal/mol and 5 kcal/mol, respectively. These ΔGridscore have a similar trend as seen in the PMFs: the rapid reversible inhibition complexes have smaller energy barriers than the slow-onset inhibition complexes.
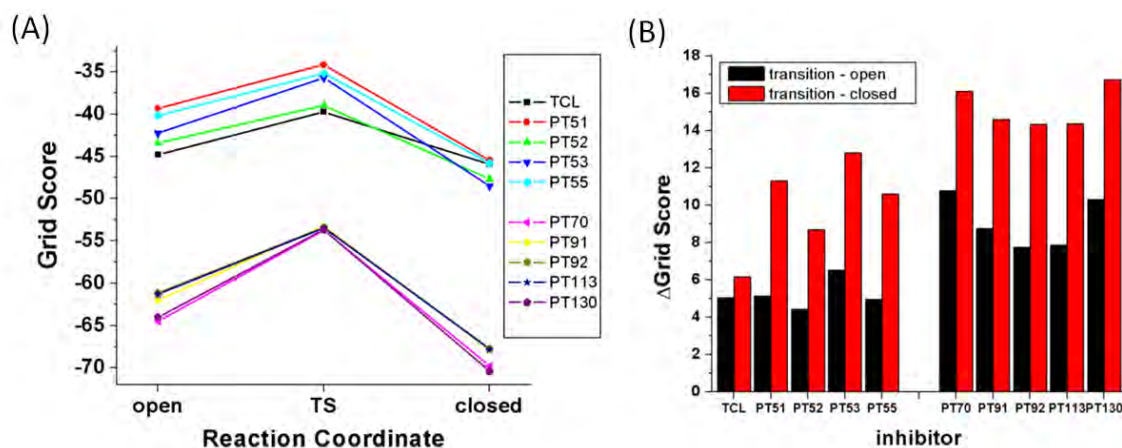
Figure 5.6. Grid score calculated from the PT70 bound PNEB open, transition state, and closed structures. (A) Grid score. (B) ΔGrid Score.

5.3.3.2 Use of triclosan bound PNEB structures for docking calculation

From the above results, it seems the dock scores reproduce the free energy pattern of the PMF; however, it may bias by the selected receptor structures. Next, we tried the rapid reversible inhibitor **triclosan** bound PNEB receptor structures. We docked the same five slow-onset and five rapid reversible inhibitors to the receptor structures and compared their dock scores.

For the slow-onset inhibitors, the dock scores calculated in the closed and open states are between -65 ~ -67 kcal/mol and -54 ~ -57 kcal/mol, respectively. The dock scores at the transition state are -37 ~ -48 kcal/mol and are higher than those calculated in the closed and open structures. These results also had a same trend as the PMF, demonstrating the closed state is energetic favorable than the open state (**Figure 5.7A**). For the rapid reversible inhibitors, the closed state dock scores are between -44 ~ -47 kcal/mol, while the open state dock scores are between -35 ~ -40 kcal/mol. The dock scores of transition state structure are between -35 ~ -38 kcal/mol.

Similar to the above results, the transition state barriers for the slow-onset inhibitors are larger than the rapid reversible inhibitors. For the slow-onset inhibitors, the average ΔGridscore of TS-open and TS-closed were 21 kcal/mol and 11 kcal/mol, respectively (**Figure 5.7B**). These values are 8~11 kcal/mol larger than those values of the rapid reversible inhibitors, which are 10 kcal/mol and 2 kcal/mol, respectively.
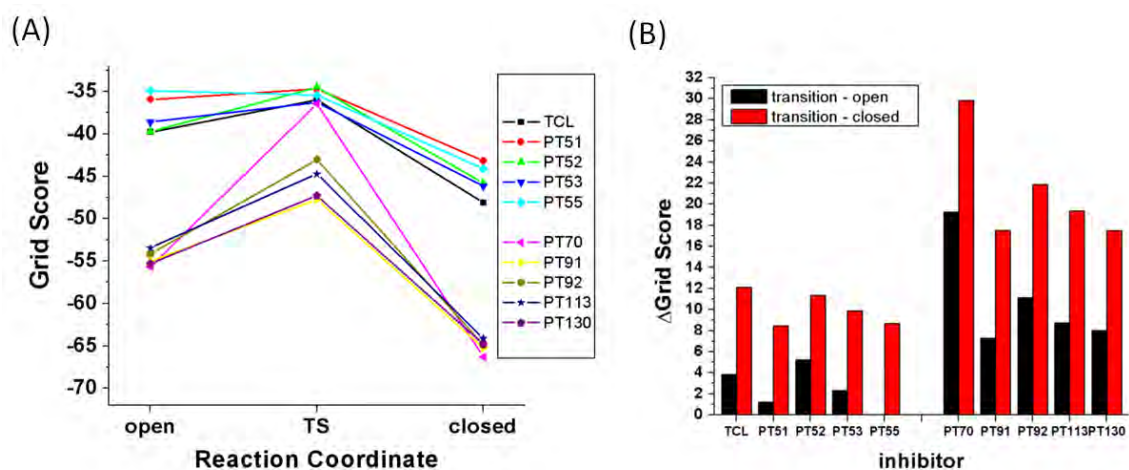
Figure 5.7. Grid score calculated from the triclosan bound PNEB open, transition state, and closed structures. (A) Grid score. (B) ΔGrid Score.

The energy barriers calculated from the **triclosan** bound PNEB structures are higher than those from the **PT70** bound PNEB structures. The ground state dock scores calculated using the **triclosan** bound PNEB structures roughly shifted ~3 kcal/mol compared to those calculated using **PT70** bound PNEB structures (**Figures 5.7A, 5.8A**). Moreover, the slow-onset inhibitors had higher energy barriers in the **triclosan** bound PNEB structures. For the **PT70** bound PNEB structures, the average energy barrier from the closed state to the transition state is ~15 kcal/mol. This energy barrier went up to ~21 kcal/mol in the **triclosan** bound PNEB structures. It is

149

possible that the **PT70** bound active-site pockets and the resulted PNEB closed/transition/open states structures were "preset" for the **PT70**-like inhibitors (particularly the slow-onset inhibitors). Therefore, slow-onset inhibitors had more favorable dock scores. On the contrary, the active-site pockets from the **triclosan** bound PNEB structures are smaller, so that the slow-onset inhibitors might bump to the receptor, resulting in an unfavorable dock scores.

In contrast, for the rapid reversible inhibitors, the energy barriers from the closed state to the transition state are very similar (both were ~10 kcal/mol) in both **PT70** bound PNEB and **triclosan** bound PNEB structures. Because **PT70** has larger structure than **triclosan**, the **PT70** bound active-site pocket is expected to be larger than the **triclosan** bound one. The larger **PT70** bound PNEB active-site pockets can always tolerate the smaller **triclosan**-like inhibitors, so the triclosan analogues had similar energy barrier in both the **PT70** bound PNEB and **triclosan** bound PNEB structures.

Overall, the dock scores have the same trend as the umbrella sampling results suggested. For future InhA inhibitor screening, we could incorporate the docking calculation to estimate their free energies at the closed, open, and transition states before the computationally expensive PNEB and umbrella sampling calculations. Although here we only tested two sets of PNEB structures, it is always better to introduce more sets of closed/open/transition states structures for the dock calculation to limit the bias problem.

5.4 Summary

To validate the role of residence time in modulating inhibitor activity in the InhA system, inhibitors with a range of residence time from short to long are needed. Since only a few slow-

onset inhibitor of InhA are available, we tried to screen possible slow-onset inhibitors of InhA with computational approaches. We first used the PNEB/US approaches to virtually screen some **PT163** analogues complexes and some known MIC (minimum inhibitory concentration) compounds. However, none of them exhibits a good slow-onset inhibition PMF.

Since PNEB and umbrella sampling calculations is computationally expensive and time-consuming. We thought that the docking calculation has the potential to rapidly examine the inhibitor's binding free energy at the closed, open, and transition states. We found the dock scores had the same trend as the umbrella sampling results: the closed state had lowest score among three states, and transition state had higher energy barriers for the slow-onset inhibitors than the rapid reversible inhibitors did. Since we only tested ten inhibitors and six receptor structures, for future virtual screening process, introducing more sets of closed/open/transition states structures for dock calculation will be necessary to limit the bias problem. The use of docking score could provide an avenue to pre-examine inhibitor binding affinity before time-consuming PNEB/US simulations.

Appendix 1: Inhibitor partial charges

Unit: kcal/mol

Triclosan partial charges
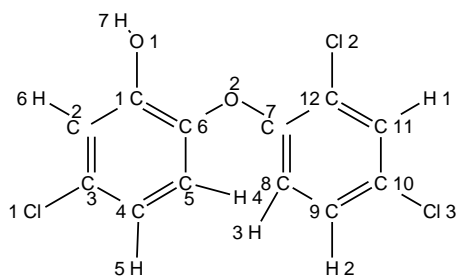
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.245554	C2	-0.200726	C3	-0.029640	C4	-0.087045
C5	-0.224964	C6	0.201578	C7	0.300778	C8	-0.256649
C9	-0.055501	C10	-0.050668	C11	-0.044691	C12	-0.031694
O1	-0.554889	O2	-0.297574	Cl1	-0.092772	Cl2	-0.095189
Cl3	-0.106579	H1	0.145569	H2	0.150212	H3	0.158601
H4	0.182727	H5	0.144719	H6	0.154300	H7	0.444541

PT3 partial charges
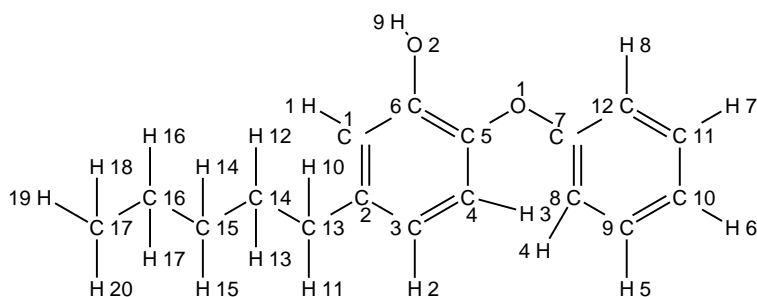
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	-0.223315	C2	-0.007234	C3	-0.229547	C4	-0.211969
C5	0.148221	C6	0.285367	C7	0.222865	C8	-0.135579
C9	-0.197672	C10	-0.123410	C11	-0.197672	C12	-0.135579
C13	-0.056773	C14	0.002988	C15	-0.002068	C16	0.026176
C17	-0.035302	O1	-0.352217	O2	-0.609758	H1	0.180658
H2	0.155157	H3	0.183429	H4	0.137848	H5	0.161508
H6	0.136727	H7	0.161508	H8	0.137848	H9	0.452110
H10	0.041085	H11	0.041085	H12	0.010396	H13	0.010396
H14	0.009321	H15	0.009321	H16	-0.008443	H17	-0.008443
H18	0.006989	H19	0.006989	H20	0.006989		

PT4 partial charges
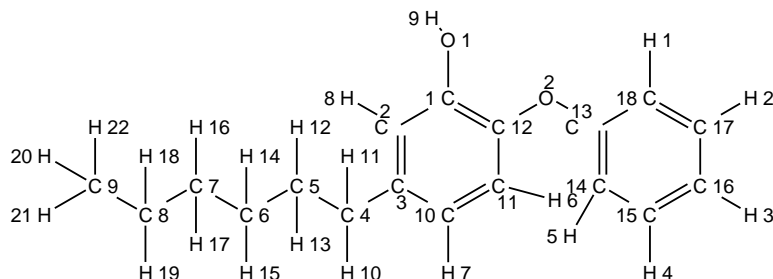
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.268524	C2	-0.203761	C3	-0.022553	C4	-0.103453
C5	0.013812	C6	0.017742	C7	-0.025854	C8	0.049305
C9	-0.044019	C10	-0.194991	C11	-0.239239	C12	0.167956
C13	0.224557	C14	-0.133611	C15	-0.208147	C16	-0.103685
C17	-0.208147	C18	-0.133611	O1	-0.608157	O2	-0.357438
H1	0.138634	H2	0.162281	H3	0.131165	H4	0.16228
H5	0.138634	H6	0.188624	H7	0.144716	H8	0.177924
H9	0.451250	H10	0.056699	H11	0.056699	H12	0.009159
H13	0.009159	H14	0.005925	H15	0.005925	H16	0.002696
H17	0.002696	H18	-0.010689	H19	-0.010689	H20	0.007227
H21	0.007227	H22	0.007227				

PT5 partial charges
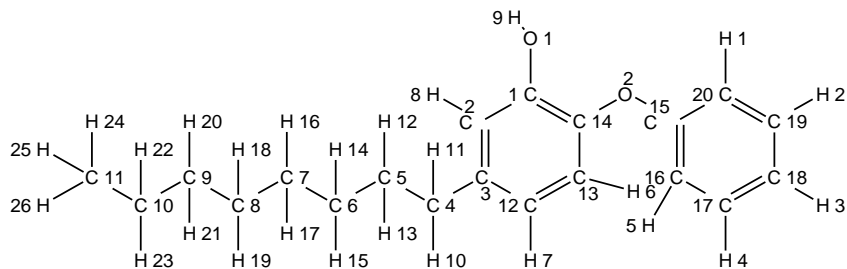
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.273801	C2	-0.292990	C3	0.016460	C4	-0.080397
C5	0.018651	C6	0.006889	C7	-0.019985	C8	-0.013415
C9	0.023785	C10	0.022808	C11	-0.064279	C12	-0.167562
C13	-0.222108	C14	0.138705	C15	0.172845	C16	-0.112445
C17	-0.193929	C18	-0.137043	C19	-0.193929	C20	-0.112445
O1	-0.557785	O2	-0.267307	H1	0.127506	H2	0.154839
H3	0.134906	H4	0.154839	H5	0.127506	H6	0.176353
H7	0.143121	H8	0.151063	H9	0.431158	H10	0.046319
H11	0.046319	H12	0.013168	H13	0.013168	H14	-0.003041
H15	-0.003041	H16	0.008237	H17	0.008237	H18	-0.004934
H19	-0.004934	H20	0.003729	H21	0.003729	H22	0.000503
H23	0.000503	H24	0.010808	H25	0.010808	H26	0.010808

PT52 partial charges

Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.233624	C2	-0.188355	C3	-0.016323	C4	-0.075891
C5	-0.243896	C6	0.178935	C7	0.162655	C8	-0.111744
C9	-0.178260	C10	-0.152766	C11	-0.178260	C12	-0.111744
O1	-0.544971	O2	-0.257369	Cl 1	-0.108545	H1	0.120878
H2	0.154633	H3	0.140046	H4	0.154633	H5	0.120878
H6	0.182598	H7	0.137230	H8	0.145982	H9	0.436032

PT70 partial charges

Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	-0.305877	C2	-0.005783	C3	-0.207809	C4	-0.182149
C5	0.112182	C6	0.269957	C7	0.123130	C8	-0.141905
C9	-0.194331	C10	-0.164129	C11	-0.234181	C12	0.111613
C13	-0.192668	C14	-0.013634	C15	0.024577	C16	-0.010196
C17	-0.005967	C18	0.065218	C19	-0.045568	O1	-0.227720
O2	-0.536475	H1	0.166880	H2	0.154695	H3	0.166490
H4	0.133435	H5	0.156808	H6	0.145586	H7	0.161364
H8	0.062701	H9	0.062701	H10	0.062701	H11	0.419315
H12	0.031061	H13	0.031061	H14	0.010100	H15	0.010100
H16	-0.000394	H17	-0.000394	H18	-0.002013	H19	-0.002013
H20	-0.017212	H21	-0.017212	H22	0.008652	H23	0.008652
H24	0.008652						

154

PT82 partial charges

Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.240795	C2	-0.157575	C3	-0.014628	C4	-0.046227
C5	-0.017992	C6	0.012018	C7	-0.008337	C8	-0.032087
C9	0.000606	C10	0.025916	C11	-0.067850	C12	-0.248655
C13	-0.157993	C14	0.047381	C15	0.180087	C16	-0.225056
C17	-0.154507	C18	-0.161377	C19	-0.235725	C20	0.106562
C21	-0.204103	O1	-0.597671	O2	-0.248518	H1	0.064271
H2	0.064271	H3	0.064271	H4	0.162463	H5	0.146445
H6	0.152121	H7	0.172945	H8	0.174921	H9	0.155076
H10	0.166210	H11	0.447658	H12	0.040396	H13	0.040396
H14	0.006656	H15	0.006656	H16	0.013450	H17	0.013450
H18	0.006018	H19	0.006018	H20	0.005784	H21	0.005784
H22	0.005276	H23	0.005276	H24	0.000536	H25	0.000536
H26	0.012683	H27	0.012683	H28	0.012683		

PT91 partial atomic charges
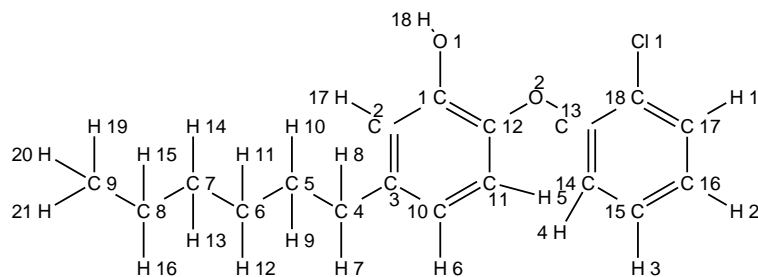
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.263299	C2	-0.318851	C3	-0.004596	C4	-0.038783
C5	0.020501	C6	-0.009081	C7	-0.010164	C8	0.054551
C9	-0.042670	C10	-0.199367	C11	-0.200547	C12	0.184665
C13	0.292876	C14	-0.191376	C15	-0.166094	C16	-0.190899
C17	-0.086354	C18	-0.022094	O1	-0.548482	O2	-0.294294
Cl1	-0.119188	H1	0.140273	H2	0.151520	H3	0.163964
H4	0.137692	H5	0.168835	H6	0.157041	H7	0.041508
H8	0.041508	H9	0.012043	H10	0.012043	H11	0.000754
H12	0.000754	H13	0.001131	H14	0.001131	H15	-0.014078
H16	-0.014078	H17	0.170982	H18	0.427894	H19	0.008678
H20	0.008678	H21	0.008678				

PT92 partial atomic charges
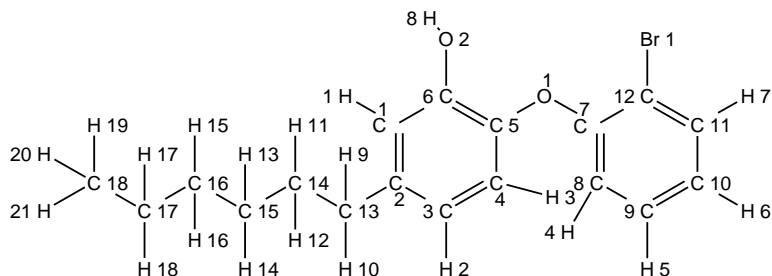
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	-0.272048	C2	-0.021360	C3	-0.191209	C4	-0.179507
C5	0.121524	C6	0.239883	C7	0.272739	C8	-0.152979
C9	-0.184872	C10	-0.155948	C11	-0.072072	C12	-0.107639
C13	-0.036432	C14	0.028997	C15	-0.004680	C16	-0.010591
C17	0.056060	C18	-0.043943	O1	-0.216836	O2	-0.554134
Br1	-0.091016	H1	0.157112	H2	0.153929	H3	0.157876
H4	0.132010	H5	0.163768	H6	0.140505	H7	0.138119
H8	0.436221	H9	0.040829	H10	0.040829	H11	0.009003
H12	0.009003	H13	-0.000839	H14	-0.000839	H15	0.000284
H16	0.000284	H17	-0.014184	H18	-0.014184	H19	0.008780
H20	0.008780	H21	0.008780				

PT119 partial atomic charges
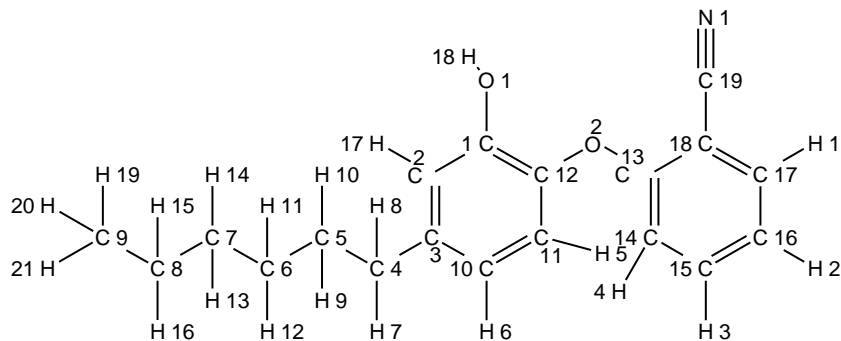
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	0.238774	C2	-0.269914	C3	0.020846	C4	-0.103698
C5	0.007662	C6	0.000630	C7	-0.021632	C8	0.055081
C9	-0.034178	C10	-0.198218	C11	-0.220266	C12	0.197522
C13	0.234682	C14	-0.218646	C15	-0.119037	C16	-0.171974
C17	-0.172551	C18	0.018429	C19	0.320513	O1	-0.555479
O2	-0.257935	N1	-0.468208	H1	0.182316	H2	0.148230
H3	0.158097	H4	0.159798	H5	0.175534	H6	0.149310
H7	0.060149	H8	0.060149	H9	0.012045	H10	0.012045
H11	0.007322	H12	0.007322	H13	0.003016	H14	0.003016
H15	-0.013893	H16	-0.013893	H17	0.155038	H18	0.433582
H19	0.006138	H20	0.006138	H21	0.006138		

PT155 partial atomic charges
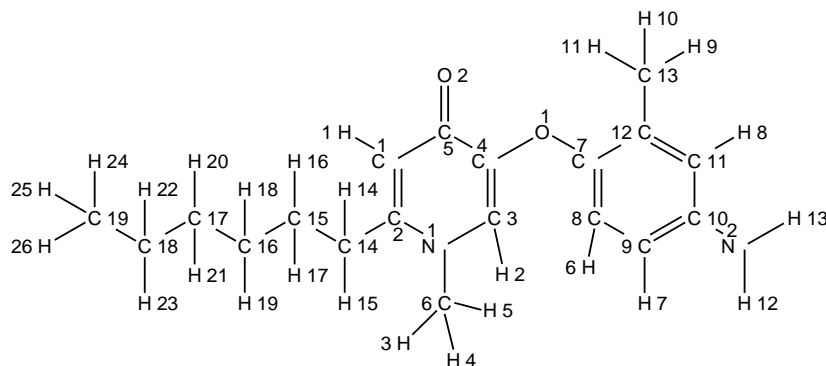
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	-0.438965	C2	0.053087	C3	-0.207546	C4	0.026918
C5	0.628641	C6	-0.133660	C7	0.171512	C8	-0.260329
C9	-0.206477	C10	0.290422	C11	-0.303745	C12	0.055573
C13	-0.125502	C14	-0.009664	C15	-0.004296	C16	-0.020813
C17	-0.002104	C18	0.032576	C19	-0.136394	N1	0.019303
N2	-0.908259	O1	-0.284312	O2	-0.601764	H1	0.184029
H2	0.230152	H3	0.083804	H4	0.083804	H5	0.083804
H6	0.208299	H7	0.157432	H8	0.172183	H9	0.045560
H10	0.045560	H11	0.045560	H12	0.378067	H13	0.378067
H14	0.048581	H15	0.048581	H16	0.015768	H17	0.015768
H18	0.002572	H19	0.002572	H20	0.006981	H21	0.006981
H22	0.014048	H23	0.014048	H24	0.031192	H25	0.031192
H26	0.031192						

PT162 partial charges
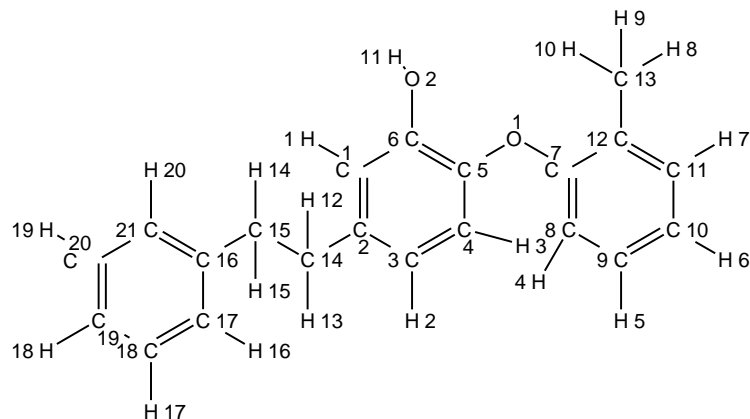
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	-0.241594	C2	-0.016797	C3	-0.187413	C4	-0.222738
C5	0.183975	C6	0.229272	C7	0.191402	C8	-0.172167
C9	-0.194219	C10	-0.176389	C11	-0.206289	C12	0.092179
C13	-0.180069	C14	0.008306	C15	-0.028610	C16	-0.000242
C17	-0.136522	C18	-0.17531	C19	-0.103650	C20	-0.175317
C21	-0.136522	O1	-0.269172	O2	-0.558044	H1	0.152014
H2	0.147673	H3	0.166923	H4	0.139061	H5	0.158347
H6	0.144587	H7	0.148949	H8	0.058580	H9	0.058580
H10	0.058580	H11	0.431523	H12	0.027243	H13	0.027243
H14	0.037373	H15	0.037373	H16	0.135539	H17	0.141943
H18	0.126925	H19	0.141943	H20	0.135539		

PT163 partial charges
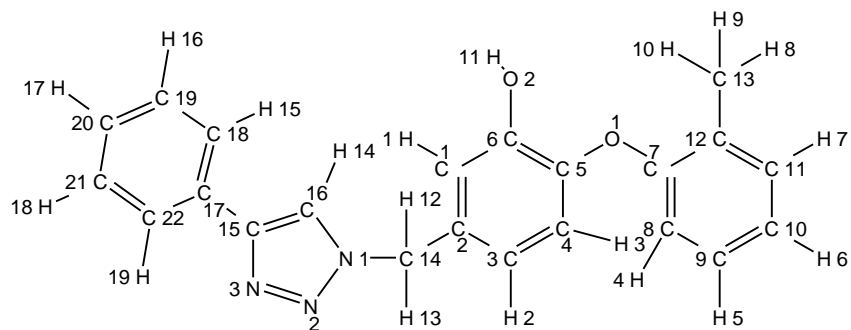
Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges	Atom name	Partial charges
C1	-0.254089	C2	0.010275	C3	-0.232043	C4	-0.186847
C5	0.065180	C6	0.293665	C7	0.168604	C8	-0.180610
C9	-0.187078	C10	-0.155166	C11	-0.215591	C12	0.090623
C13	-0.200415	C14	-0.024151	C15	0.386267	C16	-0.440644
C17	-0.115933	C18	-0.076475	C19	-0.146492	C20	-0.143830
C21	-0.146492	C22	-0.076475	N1	0.407390	N2	-0.356124
N3	-0.247600	O1	-0.239238	O2	-0.584944	H1	0.186235
H2	0.187292	H3	0.183086	H4	0.154768	H5	0.161491
H6	0.150046	H7	0.157123	H8	0.066447	H9	0.066447
H10	0.066447	H11	0.428407	H12	0.074335	H13	0.074335
H14	0.217408	H15	0.102391	H16	0.136819	H17	0.135947
H18	0.136819	H19	0.102391				

References

1.   WHO. (2012) WHO Report, Switzerland.

2.   Getahun, H., Gunneberg, C., Granich, R., and Nunn, P. (2010) HIV Infection—Associated Tuberculosis: The Epidemiology and the Response, *Clinical Infectious Diseases 50*, S201-S207.

3.   Pawlowski, A., Jansson, M., Sköld, M., Rottenberg, M. E., and Källenius, G. (2012) Tuberculosis and HIV Co-Infection, *PLoS Pathog 8*, e1002464.

4.   Bass, J. B., Jr., Farer, L. S., Hopewell, P. C., O'Brien, R., Jacobs, R. F., Ruben, F., Snider, D. E., Jr., and Thornton, G. (1994) Treatment of tuberculosis and tuberculosis infection in adults and children. American Thoracic Society and The Centers for Disease Control and Prevention, *Am J Respir Crit Care Med 149*, 1359-1374.

5.   Kiepiela, P., Bishop, K. S., Smith, A. N., Roux, L., and York, D. F. (2000) Genomic mutations in the katG, inhA and aphC genes are useful for the prediction of isoniazid resistance in Mycobacterium tuberculosis isolates from Kwazulu Natal, South Africa, *Tuber Lung Dis 80*, 47-56.

6.   Matsumoto, M., Hashizume, H., Tomishige, T., Kawasaki, M., Tsubouchi, H., Sasaki, H., Shimokawa, Y., and Komatsu, M. (2006) OPC-67683, a Nitro-Dihydro-Imidazooxazole Derivative with Promising Action against Tuberculosis In Vitro and In Mice, *PLoS Med 3*, e466.

7.   Sriram, D., Aubry, A., Yogeeswari, P., and Fisher, L. M. (2006) Gatifloxacin derivatives: Synthesis, antimycobacterial activities, and inhibition of Mycobacterium tuberculosis DNA gyrase, *Bioorganic & Medicinal Chemistry Letters 16*, 2982-2985.

8.   Tonge, P. J. (2000) Another brick in the wall, *Nat Struct Biol 7*, 94-96.

9.   Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., Wilson, T., Collins, D., de Lisle, G., and Jacobs, W. R., Jr. (1994) inhA, a gene encoding a target for isoniazid and ethionamide in Mycobacterium tuberculosis, *Science 263*, 227-230.

10.   Rawat, R., Whitty, A., and Tonge, P. J. (2003) The isoniazid-NAD adduct is a slow, tight-binding inhibitor of InhA, the Mycobacterium tuberculosis enoyl reductase: adduct affinity and drug resistance, *Proc Natl Acad Sci U S A 100*, 13881-13886.

11.   Rozwarski, D. A., Vilchèze, C., Sugantino, M., Bittman, R., and Sacchettini, J. C. (1999) Crystal Structure of the Mycobacterium tuberculosis Enoyl-ACP Reductase, InhA, in Complex with NAD+ and a C16 Fatty Acyl Substrate, *Journal of Biological Chemistry 274*, 15582-15589.

12.   Sullivan, T. J., Truglio, J. J., Boyne, M. E., Novichenok, P., Zhang, X., Stratton, C. F., Li, H.-J., Kaur, T., Amin, A., Johnson, F., Slayden, R. A., Kisker, C., and Tonge, P. J. (2006) High Affinity InhA Inhibitors with Activity against Drug-Resistant Strains of Mycobacterium tuberculosis, *ACS Chemical Biology 1*, 43-53.

13.   Luckner, S. R., Liu, N., am Ende, C. W., Tonge, P. J., and Kisker, C. (2010) A slow, tight binding inhibitor of InhA, the enoyl-acyl carrier protein reductase from Mycobacterium tuberculosis, *J Biol Chem 285*, 14330-14337.

14.   Copeland, R. A., Pompliano, D. L., and Meek, T. D. (2006) Drug-target residence time and its implications for lead optimization, *Nat Rev Drug Discov 5*, 730-739.

15.   Tummino, P. J., and Copeland, R. A. (2008) Residence time of receptor-ligand complexes and its effect on biological function, *Biochemistry 47*, 5481-5492.

16.   Swinney, D. C. (2009) The role of binding kinetics in therapeutically useful drug action, *Current opinion in drug discovery & development 12*, 31-39.

17.   Swinney, D. (2008) Applications of Binding Kinetics to Drug Discovery, *Pharm Med 22*, 23-34.

18.   Lu, H., and Tonge, P. J. (2010) Drug-target residence time: critical information for lead optimization, *Current Opinion in Chemical Biology 14*, 467-474.

19.   Zhang, R., and Monsma, F. (2010) Binding kinetics and mechanism of action: toward the discovery and development of better and best in class drugs, *Expert Opinion on Drug Discovery 5*, 1023-1029.

20.   Swinney, D. C. (2004) Biochemical mechanisms of drug action: what does it take for success?, *Nat Rev Drug Discov 3*, 801-808.

21.   Lu, H., England, K., am Ende, C., Truglio, J. J., Luckner, S., Reddy, B. G., Marlenee, N. L., Knudson, S. E., Knudson, D. L., Bowen, R. A., Kisker, C., Slayden, R. A., and Tonge, P. J. (2009) Slow-onset inhibition of the FabI enoyl reductase from francisella tularensis: residence time and in vivo activity, *ACS Chem Biol 4*, 221-231.

22.   Sillitoe, I., Cuff, A. L., Dessailly, B. H., Dawson, N. L., Furnham, N., Lee, D., Lees, J. G., Lewis, T. E., Studer, R. A., Rentzsch, R., Yeats, C., Thornton, J. M., and Orengo, C. A. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures, *Nucleic Acids Research 41*, D490-D498.

23.   Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features, *Biopolymers 22*, 2577-2637.

161

24.   Stewart, M. J., Parikh, S., Xiao, G., Tonge, P. J., and Kisker, C. (1999) Structural basis and mechanism of enoyl reductase inhibition by triclosan, *Journal of Molecular Biology 290*, 859-865.

25.   Ponder, J. W., and Case, D. A. (2003) Force Fields for Protein Simulations, In *Advances in Protein Chemistry* (Valerie, D., Ed.), pp 27-85, Academic Press.

26.   Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins 65*, 712-725.

27.   Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., and Kollman, P. A. (1995) A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *Journal of the American Chemical Society 117*, 5179-5197.

28.   Okur, A., Strockbine, B., Hornak, V., and Simmerling, C. (2003) Using PC clusters to evaluate the transferability of molecular mechanics force fields for proteins, *Journal of Computational Chemistry 24*, 21-31.

29.   García, A. E., and Sanbonmatsu, K. Y. (2002) α-Helical stabilization by side chain shielding of backbone hydrogen bonds, *Proceedings of the National Academy of Sciences 99*, 2782-2787.

30.   Kollman, P., Dixon, R., Cornell, W., Fox, T., Chipot, C., and Pohorille, A. (1997) The development/application of a 'minimalist' organic/biochemical molecular mechanic force field using a combination of ab initio calculations and experimental data, In

Computer Simulation of Biomolecular Systems (Gunsteren, W., Weiner, P., and Wilkinson, A., Eds.), pp 83-96, Springer Netherlands.

31.   Wang, J., Cieplak, P., and Kollman, P. A. (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules?, *Journal of Computational Chemistry 21*, 1049-1074.

32.   Åqvist, J., Medina, C., and Samuelsson, J.-E. (1994) A new method for predicting binding affinity in computer-aided drug design, *Protein Engineering 7*, 385-391.

33.   Srinivasan, J., Cheatham, T. E., Cieplak, P., Kollman, P. A., and Case, D. A. (1998) Continuum Solvent Studies of the Stability of DNA, RNA, and Phosphoramidate−DNA Helices, *Journal of the American Chemical Society 120*, 9401-9409.

34.   Kollman, P. A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D. A., and Cheatham, T. E., 3rd. (2000) Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models, *Acc Chem Res 33*, 889-897.

35.   Hansson, T., and Åqvist, J. (1995) Estimation of binding free energies for HIV proteinase inhibitors by molecular dynamics simulations, *Protein Engineering 8*, 1137-1144.

36.   Hultén, J., Bonham, N. M., Nillroth, U., Hansson, T., Zuccarello, G., Bouzide, A., Åqvist, J., Classon, B., Danielson, U. H., Karlén, A., Kvarnström, I., Samuelsson, B., and Hallberg, A. (1997) Cyclic HIV-1 Protease Inhibitors Derived from Mannitol: Synthesis, Inhibitory Potencies, and Computational Predictions of Binding Affinities, *Journal of Medicinal Chemistry 40*, 885-897.

37.   ÅQvist, J. (1996) Calculation of absolute binding free energies for charged ligands and effects of long-range electrostatic interactions, *Journal of Computational Chemistry 17*, 1587-1597.

38.   Johan, and Mowbray, S. L. (1995) Sugar Recognition by a Glucose/Galactose Receptor: Evaluation of Binding Energetics From Molecular Dynamics Simulations, *Journal of Biological Chemistry 270*, 9978-9981.

39.   Wang, W., Wang, J., and Kollman, P. A. (1999) What determines the van der Waals coefficient β in the LIE (linear interaction energy) method to estimate binding free energies using molecular dynamics simulations?, *Proteins: Structure, Function, and Bioinformatics 34*, 395-402.

40.   Valiente, P. A., Gil L, A., Batista, P. R., Caffarena, E. R., Pons, T., and Pascutti, P. G. (2010) New parameterization approaches of the LIE method to improve free energy calculations of PlmII-Inhibitors complexes, *Journal of Computational Chemistry 31*, 2723-2734.

41.   Jones-Hertzog, D. K., and Jorgensen, W. L. (1997) Binding Affinities for Sulfonamide Inhibitors with Human Thrombin Using Monte Carlo Simulations with a Linear Response Method, *Journal of Medicinal Chemistry 40*, 1539-1549.

42.   Ma, J., and Karplus, M. (1997) Ligand-induced conformational changes in ras p21: a normal mode and energy minimization analysis, *Journal of Molecular Biology 274*, 114-131.

43.   Karplus, M., and Kushick, J. N. (1981) Method for estimating the configurational entropy of macromolecules, *Macromolecules 14*, 325-332.

44. Hou, T., Wang, J., Li, Y., and Wang, W. (2010) Assessing the Performance of the MM/PBSA and MM/GBSA Methods. 1. The Accuracy of Binding Free Energy Calculations Based on Molecular Dynamics Simulations, *Journal of Chemical Information and Modeling 51*, 69-82.

45. Xu, L., Sun, H., Li, Y., Wang, J., and Hou, T. (2013) Assessing the Performance of MM/PBSA and MM/GBSA Methods. 3. The Impact of Force Fields and Ligand Charge Models, *The Journal of Physical Chemistry B 117*, 8408-8421.

46. Bayly, C. I., Cieplak, P., Cornell, W., and Kollman, P. A. (1993) A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model, *The Journal of Physical Chemistry 97*, 10269-10280.

47. Jakalian, A., Jack, D. B., and Bayly, C. I. (2002) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: II. Parameterization and validation, *Journal of Computational Chemistry 23*, 1623-1641.

48. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., and Weiner, P. (1984) A new force field for molecular mechanical simulation of nucleic acids and proteins, *Journal of the American Chemical Society 106*, 765-784.

49. Nurisso, A., Blanchard, B., Audfray, A., Rydner, L., Oscarson, S., Varrot, A., and Imberty, A. (2010) Role of Water Molecules in Structure and Energetics of Pseudomonas aeruginosa Lectin I Interacting with Disaccharides, *Journal of Biological Chemistry 285*, 20316-20327.

50. Singh, N., and Warshel, A. (2010) Absolute binding free energy calculations: On the accuracy of computational scoring of protein–ligand interactions, *Proteins: Structure, Function, and Bioinformatics 78*, 1705-1723.

51. Homeyer, N., and Gohlke, H. (2012) Free Energy Calculations by the Molecular Mechanics Poisson–Boltzmann Surface Area Method, *Molecular Informatics 31*, 114-122.

52. Kollman, P. (1993) Free energy calculations: Applications to chemical and biochemical phenomena, *Chemical Reviews 93*, 2395-2417.

53. Steinbrecher, T., Joung, I., and Case, D. A. (2011) Soft-core potentials in thermodynamic integration: Comparing one- and two-step transformations, *Journal of Computational Chemistry 32*, 3253-3263.

54. Kaus, J. W., Pierce, L. T., Walker, R. C., and McCammon, J. A. (2013) Improving the Efficiency of Free Energy Calculations in the Amber Molecular Dynamics Package, *Journal of Chemical Theory and Computation 9*, 4131-4139.

55. Jarzynski, C. (1997) Nonequilibrium Equality for Free Energy Differences, *Physical Review Letters 78*, 2690-2693.

56. Hummer, G., and Szabo, A. (2001) Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proceedings of the National Academy of Sciences 98*, 3658-3661.

57. Hummer, G., and Szabo, A. (2003) Kinetics from Nonequilibrium Single-Molecule Pulling Experiments, *Biophysical Journal 85*, 5-15.

58. Torrie, G. M., and Valleau, J. P. (1977) Nonphysical sampling distributions in Monte Carlo free-energy estimation: Umbrella sampling, *Journal of Computational Physics 23*, 187-199.

59. de Ruiter, A., and Oostenbrink, C. (2011) Free energy calculations of protein–ligand interactions, *Current Opinion in Chemical Biology 15*, 547-552.

60. Huang, D., and Caflisch, A. (2011) The Free Energy Landscape of Small Molecule Unbinding, *PLoS Comput Biol 7*, e1002002.

61. Buch, I., Sadiq, S. K., and De Fabritiis, G. (2011) Optimized Potential of Mean Force Calculations for Standard Binding Free Energies, *Journal of Chemical Theory and Computation 7*, 1765-1772.

62. Vashisth, H., and Abrams, C. F. (2008) Ligand Escape Pathways and (Un)Binding Free Energy Calculations for the Hexameric Insulin-Phenol Complex, *Biophysical Journal 95*, 4193-4204.

63. Woo, H.-J., and Roux, B. (2005) Calculation of absolute protein–ligand binding free energy from computer simulations, *Proceedings of the National Academy of Sciences of the United States of America 102*, 6825-6830.

64. Nicolini, P., Frezzato, D., Gellini, C., Bizzarri, M., and Chelli, R. (2013) Toward quantitative estimates of binding affinities for protein–ligand systems involving large inhibitor compounds: A steered molecular dynamics simulation route, *Journal of Computational Chemistry 34*, 1561-1576.

65. Allen, W. J., and Bevan, D. R. (2011) Steered Molecular Dynamics Simulations Reveal Important Mechanisms in Reversible Monoamine Oxidase B Inhibition, *Biochemistry 50*, 6441-6454.

66. Bai, F., Xu, Y., Chen, J., Liu, Q., Gu, J., Wang, X., Ma, J., Li, H., Onuchic, J. N., and Jiang, H. (2013) Free energy landscape for the binding process of Huperzine A to acetylcholinesterase, *Proceedings of the National Academy of Sciences 110*, 4273-4278.

67.   Ponder, J. W., Wu, C., Ren, P., Pande, V. S., Chodera, J. D., Schnieders, M. J., Haque, I.,

        Mobley, D. L., Lambrecht, D. S., DiStasio, R. A., Head-Gordon, M., Clark, G. N. I., Johnson,

        M. E., and Head-Gordon, T. (2010) Current Status of the AMOEBA Polarizable Force Field,

        The Journal of Physical Chemistry B 114, 2549-2564.

68.   Baker, C. M., Anisimov, V. M., and MacKerell, A. D. (2010) Development of CHARMM

        Polarizable Force Field for Nucleic Acid Bases Based on the Classical Drude Oscillator

        Model, *The Journal of Physical Chemistry B 115*, 580-596.

69.   Baker, C. M., Lopes, P. E. M., Zhu, X., Roux, B., and MacKerell, A. D. (2010) Accurate

        Calculation of Hydration Free Energies using Pair-Specific Lennard-Jones Parameters in

        the CHARMM Drude Polarizable Force Field, *Journal of Chemical Theory and

        Computation 6*, 1181-1198.

70.   Patel, S., and Brooks, C. L. (2004) CHARMM fluctuating charge force field for proteins: I

        parameterization and application to bulk organic liquid simulations, *Journal of

        Computational Chemistry 25*, 1-16.

71.   Patel, S., Mackerell, A. D., and Brooks, C. L. (2004) CHARMM fluctuating charge force

        field for proteins: II Protein/solvent properties from molecular dynamics simulations

        using a nonadditive electrostatic model, *Journal of Computational Chemistry 25*, 1504-

        1514.

72.   Lamoureux, G., and Roux, B. (2003) Modeling induced polarization with classical Drude

        oscillators: Theory and molecular dynamics simulation algorithm, *The Journal of

        Chemical Physics 119*, 3025-3039.

73. Lamoureux, G., Harder, E., Vorobyov, I. V., Roux, B., and MacKerell Jr, A. D. (2006) A polarizable model of water for molecular dynamics simulations of biomolecules, *Chemical Physics Letters 418*, 245-249.

74. Wang, Z.-X., Zhang, W., Wu, C., Lei, H., Cieplak, P., and Duan, Y. (2006) Strike a balance: Optimization of backbone torsion parameters of AMBER polarizable force field for simulations of proteins and peptides, *Journal of Computational Chemistry 27*, 781-790.

75. Cieplak, P., Caldwell, J., and Kollman, P. (2001) Molecular mechanical models for organic and biological systems going beyond the atom centered two body additive approximation: aqueous solution free energies of methanol and N-methyl acetamide, nucleic acid base, and amide hydrogen bonding and chloroform/water partition coefficients of the nucleic acid bases, *Journal of Computational Chemistry 22*, 1048-1057.

76. Roitberg, A. E., Okur, A., and Simmerling, C. (2007) Coupling of Replica Exchange Simulations to a Non-Boltzmann Structure Reservoir, *The Journal of Physical Chemistry B 111*, 2415-2418.

77. Babin, V., Roland, C., and Sagui, C. (2008) Adaptively biased molecular dynamics for free energy calculations, *The Journal of Chemical Physics 128*, -.

78. Hamelberg, D., Mongan, J., and McCammon, J. A. (2004) Accelerated molecular dynamics: A promising and efficient simulation method for biomolecules, *The Journal of Chemical Physics 120*, 11919-11929.

79. Le Grand, S., Götz, A. W., and Walker, R. C. (2013) SPFP: Speed without compromise—A mixed precision model for GPU accelerated molecular dynamics simulations, *Computer Physics Communications 184*, 374-380.

80.   Götz, A. W., Williamson, M. J., Xu, D., Poole, D., Le Grand, S., and Walker, R. C. (2012) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born, *Journal of Chemical Theory and Computation 8*, 1542-1555.

81.   Salomon-Ferrer, R., Götz, A. W., Poole, D., Le Grand, S., and Walker, R. C. (2013) Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald, *Journal of Chemical Theory and Computation 9*, 3878-3888.

82.   Schroeder, E. K., Basso, L. A., Santos, D. S., and de Souza, O. N. (2005) Molecular dynamics simulation studies of the wild-type, I21V, and I16T mutants of isoniazid-resistant Mycobacterium tuberculosis enoyl reductase (InhA) in complex with NADH: toward the understanding of NADH-InhA different affinities, *Biophys J 89*, 876-884.

83.   Wahab, H. A., Choong, Y. S., Ibrahim, P., Sadikun, A., and Scior, T. (2009) Elucidating isoniazid resistance using molecular modeling, *J Chem Inf Model 49*, 97-107.

84.   Cohen, E., Machado, K., Cohen, M., and Norberto de Souza, O. (2011) Effect of the explicit flexibility of the InhA enzyme from Mycobacterium tuberculosis in molecular docking simulations, *BMC Genomics 12*, S7.

85.   Costa, A. P., Pauli, I., Dorn, M., Schroeder, E., Zhan, C.-G., and Norberto de Souza, O. (2012) Conformational changes in 2-trans-enoyl-ACP (CoA) reductase (InhA) from M. tuberculosis induced by an inorganic complex: a molecular dynamics simulation study, *J Mol Model 18*, 1779-1790.

86.   Basso, L. A., Zheng, R., Musser, J. M., Jacobs, W. R., and Blanchard, J. S. (1998) Mechanisms of Isoniazid Resistance in Mycobacterium tuberculosis: Enzymatic

Characterization of Enoyl Reductase Mutants Identified in Isoniazid-Resistant Clinical Isolates, *Journal of Infectious Diseases 178*, 769-775.

87.   Stigliani, J. L., Arnaud, P., Delaine, T., Bernardes-Genisson, V., Meunier, B., and Bernadou, J. (2008) Binding of the tautomeric forms of isoniazid-NAD adducts to the active site of the Mycobacterium tuberculosis enoyl-ACP reductase (InhA): a theoretical approach, *J Mol Graph Model 27*, 536-545.

88.   Argyrou, A., Vetting, M. W., Aladegbami, B., and Blanchard, J. S. (2006) Mycobacterium tuberculosis dihydrofolate reductase is a target for isoniazid, *Nat Struct Mol Biol 13*, 408-413.

89.   Rafi, S., Novichenok, P., Kolappan, S., Zhang, X., Stratton, C. F., Rawat, R., Kisker, C., Simmerling, C., and Tonge, P. J. (2006) Structure of acyl carrier protein bound to FabI, the FASII enoyl reductase from Escherichia coli, *J Biol Chem 281*, 39285-39293.

90.   Rafi, S. B., Cui, G., Song, K., Cheng, X., Tonge, P. J., and Simmerling, C. (2006) Insight through molecular mechanics Poisson-Boltzmann surface area calculations into the binding affinity of triclosan and three analogues for FabI, the E. coli enoyl reductase, *J Med Chem 49*, 4574-4580.

91.   Lindert, S., and McCammon, J. A. (2012) Dynamics of Plasmodium falciparum enoyl-ACP reductase and implications on drug discovery, *Protein Science 21*, 1734-1745.

92.   Mathews, D. H., and Case, D. A. (2006) Nudged Elastic Band Calculation of Minimal Energy Paths for the Conformational Change of a GG Non-canonical Pair, *Journal of Molecular Biology 357*, 1683-1693.

93.   Elber, R., and Karplus, M. (1987) A method for determining reaction paths in large molecules: Application to myoglobin, *Chemical Physics Letters 139*, 375-380.

94.   Kirkpatrick, S., Gelatt, C. D., and Vecchi, M. P. (1983) Optimization by Simulated Annealing, *Science 220*, 671-680.

95.   Bergonzo, C., Campbell, A. J., Walker, R. C., and Simmerling, C. (2009) A Partial Nudged Elastic Band Implementation for Use with Large or Explicitly Solvated Systems, *Int J Quantum Chem 109*, 3781.

96.   Grossfield, A. (2007) An implementation of WHAM: the Weighted Histogram Analysis Method.

97.   Roux, B. (1995) The calculation of the potential of mean force using computer simulations, *Computer Physics Communications 91*, 275-282.

98.   Kumar, S., Rosenberg, J. M., Bouzida, D., Swendsen, R. H., and Kollman, P. A. (1995) Multidimensional free-energy calculations using the weighted histogram analysis method, *Journal of Computational Chemistry 16*, 1339-1350.

99.   Sharff, A. J., Rodseth, L. E., Spurlino, J. C., and Quiocho, F. A. (1992) Crystallographic evidence of a large ligand-induced hinge-twist motion between the two domains of the maltodextrin binding protein involved in active transport and chemotaxis, *Biochemistry 31*, 10657-10663.

100. Qasba, P. K., Ramakrishnan, B., and Boeggeman, E. (2005) Substrate-induced conformational changes in glycosyltransferases, *Trends in Biochemical Sciences 30*, 53-62.

101. Trodler, P., Schmid, R., and Pleiss, J. (2009) Modeling of solvent-dependent conformational transitions in Burkholderia cepacia lipase, *BMC Structural Biology 9*, 38.

102. Groban, E. S., Narayanan, A., and Jacobson, M. P. (2006) Conformational Changes in Protein Loops and Helices Induced by Post-Translational Phosphorylation, *PLoS Comput Biol 2*, e32.

103. Hyeon, C., Jennings, P. A., Adams, J. A., and Onuchic, J. N. (2009) Ligand-induced global transitions in the catalytic domain of protein kinase A, *Proceedings of the National Academy of Sciences 106*, 3023-3028.

104. Copeland, R. A. (2011) Conformational adaptation in drug–target interactions and residence time, *Future Medicinal Chemistry 3*, 1491-1501.

105. Beckstein, O., Denning, E. J., Perilla, J. R., and Woolf, T. B. (2009) Zipping and Unzipping of Adenylate Kinase: Atomistic Insights into the Ensemble of Open↔Closed Transitions, *Journal of Molecular Biology 394*, 160-176.

106. Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, *Proceedings of the National Academy of Sciences of the United States of America 103*, 915-920.

107. Henriksson, L. M., Unge, T., Carlsson, J., Åqvist, J., Mowbray, S. L., and Jones, T. A. (2007) Structures of Mycobacterium tuberculosis 1-Deoxy-D-xylulose-5-phosphate Reductoisomerase Provide New Insights into Catalysis, *Journal of Biological Chemistry 282*, 19905-19916.

108.   Bermejo, G. A., Strub, M.-P., Ho, C., and Tjandra, N. (2010) Ligand-Free Open–Closed Transitions of Periplasmic Binding Proteins: The Case of Glutamine-Binding Protein, *Biochemistry 49*, 1893-1902.

109.   Shabalin, I. G., Filippova, E. V., Polyakov, K. M., Sadykhov, E. G., Safonova, T. N., Tikhonova, T. V., Tishkov, V. I., and Popov, V. O. (2009) Structures of the apo and holo forms of formate dehydrogenase from the bacterium Moraxella sp. C-1: towards understanding the mechanism of the closure of the interdomain cleft, *Acta Crystallographica Section D 65*, 1315-1325.

110.   Koshland, D. E. (1958) Application of a Theory of Enzyme Specificity to Protein Synthesis, *Proceedings of the National Academy of Sciences 44*, 98-104.

111.   Boehr, D. D., Nussinov, R., and Wright, P. E. (2009) The role of dynamic conformational ensembles in biomolecular recognition, *Nat Chem Biol 5*, 789-796.

112.   Bucher, D., Grant, B. J., and McCammon, J. A. (2011) Induced Fit or Conformational Selection? The Role of the Semi-closed State in the Maltose Binding Protein, *Biochemistry 50*, 10530-10539.

113.   Okazaki, K.-I., and Takada, S. (2008) Dynamic energy landscape view of coupled binding and protein conformational change: Induced-fit versus population-shift mechanisms, *Proceedings of the National Academy of Sciences 105*, 11182-11187.

114.   Sullivan, S. M., and Holyoak, T. (2008) Enzymes with lid-gated active sites must operate by an induced fit mechanism instead of conformational selection, *Proceedings of the National Academy of Sciences 105*, 13829-13834.

115. Wang, J., Wolf, R. M., Caldwell, J. W., Kollman, P. A., and Case, D. A. (2004) Development and testing of a general amber force field, *Journal of Computational Chemistry 25*, 1157-1174.

116. Cornell, W. D., Cieplak, P., Bayly, C. I., and Kollmann, P. A. (1993) Application of RESP charges to calculate conformational energies, hydrogen bond energies, and free energies of solvation, *Journal of the American Chemical Society 115*, 9620-9631.

117. Torrie, G. M., and Valleau, J. P. (1974) Monte Carlo free energy estimates using non-Boltzmann sampling: Application to the sub-critical Lennard-Jones fluid, *Chemical Physics Letters 28*, 578-581.

118. Pavelites, J. J., Gao, J., Bash, P. A., and Mackerell, A. D. (1997) A molecular mechanics force field for NAD+ NADH, and the pyrophosphate groups of nucleotides, *Journal of Computational Chemistry 18*, 221-239.

119. Walker, R. C., de Souza, M. M., Mercer, I. P., Gould, I. R., and Klug, D. R. (2002) Large and Fast Relaxations inside a Protein: Calculation and Measurement of Reorganization Energies in Alcohol Dehydrogenase, *The Journal of Physical Chemistry B 106*, 11658-11665.

120. Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983) Comparison of simple potential functions for simulating liquid water, *The Journal of Chemical Physics 79*, 926-935.

121. Ringner, M. (2008) What is principal component analysis?, *Nat Biotech 26*, 303-304.

122. Li, H.-J., Lai, C.-T., Pan, P., Yu, W., Liu, N., Bommineni, G. R., Garcia-Diaz, M., Simmerling, C., and Tonge, P. J. (2014) A structural and energetic model for the slow-onset inhibition

175

of the Mycobacterium tuberculosis enoyl-ACP reductase InhA, *ACS Chemical Biology 9*, 986-993.

123. Hartkoorn, R. C., Pojer, F., Read, J. A., Gingell, H., Neres, J., Horlacher, O. P., Altmann, K.-H., and Cole, S. T. (2013) Pyridomycin bridges the NADH- and substrate-binding pockets of the enoyl reductase InhA, *Nat Chem Biol advance online publication*.

124. Hartkoorn, R. C., Sala, C., Neres, J., Pojer, F., Magnet, S., Mukherjee, R., Uplekar, S., Boy-Röttger, S., Altmann, K.-H., and Cole, S. T. (2012) Towards a new tuberculosis drug: pyridomycin – nature's isoniazid, *EMBO Molecular Medicine 4*, 1032-1042.

125. Molle, V., Gulten, G., Vilchèze, C., Veyron-Churlet, R., Zanella-Cléon, I., Sacchettini, J. C., Jacobs Jr, W. R., and Kremer, L. (2010) Phosphorylation of InhA inhibits mycolic acid biosynthesis and growth of Mycobacterium tuberculosis, *Molecular Microbiology 78*, 1591-1605.

126. Freundlich, J. S., Wang, F., Vilchèze, C., Gulten, G., Langley, R., Schiehser, G. A., Jacobus, D. P., Jacobs, W. R., and Sacchettini, J. C. (2009) Triclosan Derivatives: Towards Potent Inhibitors of Drug-Sensitive and Drug-Resistant Mycobacterium tuberculosis, *ChemMedChem 4*, 241-248.

127. Argyrou, A., Vetting, M. W., and Blanchard, J. S. (2007) New Insight into the Mechanism of Action of and Resistance to Isoniazid: Interaction of Mycobacterium tuberculosis enoyl-ACP Reductase with INH-NADP, *Journal of the American Chemical Society 129*, 9582-9583.

128. Wang, F., Langley, R., Gulten, G., Dover, L. G., Besra, G. S., Jacobs, W. R., and Sacchettini, J. C. (2007) Mechanism of thioamide drug action against tuberculosis and leprosy, *The Journal of Experimental Medicine 204*, 73-78.

129. Dias, M. V. B., Vasconcelos, I. B., Prado, A. M. X., Fadel, V., Basso, L. A., de Azevedo Jr, W. F., and Santos, D. S. (2007) Crystallographic studies on the binding of isonicotinyl-NAD adduct to wild-type and isoniazid resistant 2-trans-enoyl-ACP (CoA) reductase from Mycobacterium tuberculosis, *Journal of Structural Biology 159*, 369-380.

130. He, X., Alian, A., Stroud, R., and Ortiz de Montellano, P. R. (2006) Pyrrolidine Carboxamides as a Novel Class of Inhibitors of Enoyl Acyl Carrier Protein Reductase from Mycobacterium tuberculosis, *Journal of Medicinal Chemistry 49*, 6308-6323.

131. Vilcheze, C., Wang, F., Arai, M., Hazbon, M. H., Colangeli, R., Kremer, L., Weisbrod, T. R., Alland, D., Sacchettini, J. C., and Jacobs, W. R. (2006) Transfer of a point mutation in Mycobacterium tuberculosis inhA resolves the target of isoniazid, *Nat Med 12*, 1027-1029.

132. Oliveira, J. S., Pereira, J. H., Canduri, F., Rodrigues, N. C., de Souza, O. N., de Azevedo Jr, W. F., Basso, L. A., and Santos, D. S. (2006) Crystallographic and Pre-steady-state Kinetics Studies on Binding of NADH to Wild-type and Isoniazid-resistant Enoyl-ACP(CoA) Reductase Enzymes from Mycobacterium tuberculosis, *Journal of Molecular Biology 359*, 646-666.

133. Kuo, M. R., Morbidoni, H. R., Alland, D., Sneddon, S. F., Gourlie, B. B., Staveski, M. M., Leonard, M., Gregory, J. S., Janjigian, A. D., Yee, C., Musser, J. M., Kreiswirth, B., Iwamoto, H., Perozzo, R., Jacobs, W. R., Sacchettini, J. C., and Fidock, D. A. (2003)

Targeting Tuberculosis and Malaria through Inhibition of Enoyl Reductase: COMPOUND ACTIVITY AND STRUCTURAL DATA, *Journal of Biological Chemistry 278*, 20851-20859.

134. Rozwarski, D. A., Grant, G. A., Barton, D. H. R., Jacobs, W. R., and Sacchettini, J. C. (1998) Modification of the NADH of the Isoniazid Target (InhA) from Mycobacterium tuberculosis, *Science 279*, 98-102.

135. Dessen, A., Quemard, A., Blanchard, J., Jacobs, W., and Sacchettini, J. (1995) Crystal structure and function of the isoniazid target of Mycobacterium tuberculosis, *Science 267*, 1638-1641.

136. Bakan, A., Meireles, L. M., and Bahar, I. (2011) ProDy: Protein Dynamics Inferred from Theory and Experiments, *Bioinformatics 27*, 1575-1577.

137. Shan, Y., Kim, E. T., Eastwood, M. P., Dror, R. O., Seeliger, M. A., and Shaw, D. E. (2011) How Does a Drug Molecule Find Its Target Binding Site?, *Journal of the American Chemical Society 133*, 9181-9183.

138. Murray, C., and Verdonk, M. (2002) The consequences of translational and rotational entropy lost by small molecules on binding to proteins, *J Comput Aided Mol Des 16*, 741-753.

139. Gilson, M. K., and Zhou, H.-X. (2007) Calculation of Protein-Ligand Binding Affinities*, *Annual Review of Biophysics and Biomolecular Structure 36*, 21-42.

140. Lang, P. T., Brozell, S. R., Mukherjee, S., Pettersen, E. F., Meng, E. C., Thomas, V., Rizzo, R. C., Case, D. A., James, T. L., and Kuntz, I. D. (2009) DOCK 6: combining techniques to model RNA-small molecule complexes, *RNA 15*, 1219-1230.

141.   Mukherjee, S., Balius, T. E., and Rizzo, R. C. (2010) Docking Validation Resources: Protein Family and Ligand Flexibility Experiments, *Journal of Chemical Information and Modeling 50*, 1986-2000.

142.   Pan, P., Knudson, S. E., Bommineni, G. R., Li, H.-J., Lai, C.-T., Liu, N., Garcia-Diaz, M., Simmerling, C., Patil, S. S., Slayden, R. A., and Tonge, P. J. (2014) Time-Dependent Diaryl Ether Inhibitors of InhA: Structure–Activity Relationship Studies of Enzyme Inhibition, Antibacterial Activity, and in vivo Efficacy, *ChemMedChem 9*, 776-791.

143.   Miller, B. R., McGee, T. D., Swails, J. M., Homeyer, N., Gohlke, H., and Roitberg, A. E. (2012) MMPBSA.py: An Efficient Program for End-State Free Energy Calculations, *Journal of Chemical Theory and Computation 8*, 3314-3321.

144.   Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1996) Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium, *The Journal of Physical Chemistry 100*, 19824-19839.

145.   Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995) Pairwise solute descreening of solute charges from a dielectric medium, *Chemical Physics Letters 246*, 122-129.

146.   Chang, C.-E. A., Chen, W., and Gilson, M. K. (2007) Ligand configurational entropy and protein binding, *Proceedings of the National Academy of Sciences 104*, 1534-1539.

147.   Chen, W., Chang, C.-E., and Gilson, M. K. (2004) Calculation of Cyclodextrin Binding Affinities: Energy, Entropy, and Implications for Drug Design, *Biophysical Journal 87*, 3035-3049.

148. Rocklin, G. J., Boyce, S. E., Fischer, M., Fish, I., Mobley, D. L., Shoichet, B. K., and Dill, K. A. (2013) Blind Prediction of Charged Ligand Binding Affinities in a Model Binding Site, *Journal of Molecular Biology 425*, 4569-4583

149. Schneider, E. V., Böttcher, J., Huber, R., Maskos, K., and Neumann, L. (2013) Structure–kinetic relationship study of CDK8/CycC specific compounds, *Proceedings of the National Academy of Sciences 110*, 8081-8086.

150. Vilums, M., Zweemer, A. J. M., Yu, Z., de Vries, H., Hillger, J. M., Wapenaar, H., Bollen, I. A. E., Barmare, F., Gross, R., Clemens, J., Krenitsky, P., Brussee, J., Stamos, D., Saunders, J., Heitman, L. H., and Ijzerman, A. P. (2013) Structure–Kinetic Relationships—An Overlooked Parameter in Hit-to-Lead Optimization: A Case of Cyclopentylamines as Chemokine Receptor 2 Antagonists, *Journal of Medicinal Chemistry 56*, 7706–7714.

151. Madeo, J., Mihajlovic, M., Lazaridis, T., and Gunner, M. R. (2011) Slow Dissociation of a Charged Ligand: Analysis of the Primary Quinone QA Site of Photosynthetic Bacterial Reaction Centers, *Journal of the American Chemical Society 133*, 17375-17385.

152. Chang, A., Schiebel, J., Yu, W., Bommineni, G. R., Pan, P., Baxter, M. V., Khanna, A., Sotriffer, C. A., Kisker, C., and Tonge, P. J. (2013) Rational Optimization of Drug-Target Residence Time: Insights from Inhibitor Binding to the Staphylococcus aureus FabI Enzyme–Product Complex, *Biochemistry 52*, 4217-4228.

153. Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Zakrzewski, V. G., Montgomery, J. A., Jr., R. E. S., Burant, J. C., Dapprich, S., Millam, J. M., Daniels, A. D., Kudin, K. N., Strain, M. C., Farkas, O., Tomasi, J., Barone, V., Cossi, M., Cammi, R., Mennucci, B., Pomelli, C., Adamo, C., Clifford, S., Ochterski, J., Petersson, G.

A., Ayala, P. Y., Cui, Q., Morokuma, K., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Cioslowski, J., Ortiz, J. V., Baboul, A. G., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Gomperts, R., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Gonzalez, C., Challacombe, M., Gill, P. M. W., Johnson, B. G., Chen, W., Wong, M. W., Andres, J. L., Head-Gordon, M., Replogle, E. S., and Pople, J. A. (1998) Gaussian 98 (Revision A.5), Gaussian, Inc., Pittsburgh PA.

154.   Guex, N., and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-Pdb Viewer: An environment for comparative protein modeling, *Electrophoresis 18*, 2714-2723.

155.   Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes, *Journal of Computational Physics 23*, 327-341.

156.   Cheatham, T. E., III, Miller, J. L., Fox, T., Darden, T. A., and Kollman, P. A. (1995) Molecular Dynamics Simulations on Solvated Biomolecular Systems: The Particle Mesh Ewald Method Leads to Stable Trajectories of DNA, RNA, and Proteins, *Journal of the American Chemical Society 117*, 4193-4194.

157.   Steinbrecher, T. (2007) AMBER TUTORIAL A9: Thermodynamic Integration using soft core potentials.

158.   Mehboob, S., Truong, K., Santarsiero, B. D., and Johnson, M. E. (2010) Structure of the Francisella tularensis enoyl-acyl carrier protein reductase (FabI) in complex with NAD+ and triclosan, *Acta Crystallographica Section F 66*, 1436-1440.

159. Priyadarshi, A., Kim, E. E., and Hwang, K. Y. (2010) Structural insights into Staphylococcus aureus enoyl-ACP reductase (FabI), in complex with NADP and triclosan, *Proteins: Structure, Function, and Bioinformatics 78*, 480-486.

160. Kim, K.-H., Ha, B. H., Kim, S. J., Hong, S. K., Hwang, K. Y., and Kim, E. E. (2011) Crystal Structures of Enoyl-ACP Reductases I (FabI) and III (FabL) from B. subtilis, *Journal of Molecular Biology 406*, 403-415.

161. Lee, H. H., Moon, J., and Suh, S. W. (2007) Crystal structure of the Helicobacter pylori enoyl-acyl carrier protein reductase in complex with hydroxydiphenyl ether compounds, triclosan and diclosan, *Proteins: Structure, Function, and Bioinformatics 69*, 691-694.

162. Tipparaju, S. K., Mulhearn, D. C., Klein, G. M., Chen, Y., Tapadar, S., Bishop, M. H., Yang, S., Chen, J., Ghassemi, M., Santarsiero, B. D., Cook, J. L., Johlfs, M., Mesecar, A. D., Johnson, M. E., and Kozikowski, A. P. (2008) Design and Synthesis of Aryl Ether Inhibitors of the Bacillus Anthracis Enoyl-ACP Reductase, *ChemMedChem 3*, 1250-1268.

163. Archer, G. L. (1998) Staphylococcus aureus: A Well-Armed Pathogen, *Clinical Infectious Diseases 26*, 1179-1181.

164. Kluytmans, J., van Belkum, A., and Verbrugh, H. (1997) Nasal carriage of Staphylococcus aureus: epidemiology, underlying mechanisms, and associated risks, *Clinical Microbiology Reviews 10*, 505-520.

165. Oliveira, D. C., Tomasz, A., and de Lencastre, H. (2002) Secrets of success of a human pathogen: molecular evolution of pandemic clones of meticillin-resistant Staphylococcus aureus, *The Lancet Infectious Diseases 2*, 180-189.

166. Park, H. S., Yoon, Y. M., Jung, S. J., Kim, C. M., Kim, J. M., and Kwak, J.-H. (2007) Antistaphylococcal activities of CG400549, a new bacterial enoyl-acyl carrier protein reductase (FabI) inhibitor, *Journal of Antimicrobial Chemotherapy 60*, 568-574.

167. Escaich, S., Prouvensier, L., Saccomani, M., Durant, L., Oxoby, M., Gerusz, V., Moreau, F., Vongsouthi, V., Maher, K., Morrissey, I., and Soulama-Mouze, C. (2011) The MUT056399 Inhibitor of FabI Is a New Antistaphylococcal Compound, *Antimicrobial Agents and Chemotherapy 55*, 4692-4697.

168. Kaplan, N., Albert, M., Awrey, D., Bardouniotis, E., Berman, J., Clarke, T., Dorsey, M., Hafkin, B., Ramnauth, J., Romanov, V., Schmid, M. B., Thalakada, R., Yethon, J., and Pauls, H. W. (2012) Mode of Action, In Vitro Activity, and In Vivo Efficacy of AFN-1252, a Selective Antistaphylococcal FabI Inhibitor, *Antimicrobial Agents and Chemotherapy 56*, 5865-5874.

169. Xu, H., Sullivan, T. J., Sekiguchi, J.-i., Kirikae, T., Ojima, I., Stratton, C. F., Mao, W., Rock, F. L., Alley, M. R. K., Johnson, F., Walker, S. G., and Tonge, P. J. (2008) Mechanism and Inhibition of saFabI, the Enoyl Reductase from Staphylococcus aureus†, *Biochemistry 47*, 4228-4236.

170. Schiebel, J., Chang, A., Lu, H., Baxter, Michael V., Tonge, Peter J., and Kisker, C. (2012) Staphylococcus aureus FabI: Inhibition, Substrate Recognition, and Potential Implications for In Vivo Essentiality, *Structure (London, England : 1993) 20*, 802-813.

171. Zheng, L., Chen, M., and Yang, W. (2009) Simultaneous escaping of explicit and hidden free energy barriers: Application of the orthogonal space random walk strategy in

generalized ensemble based conformational sampling, *The Journal of Chemical Physics 130*, 234105.

172.   Wojtas-Niziurski, W., Meng, Y., Roux, B., and Bernèche, S. (2013) Self-Learning Adaptive Umbrella Sampling Method for the Determination of Free Energy Landscapes in Multiple Dimensions, *Journal of Chemical Theory and Computation 9*, 1885-1895.

173.   Elber, R. (2007) A Milestoning Study of the Kinetics of an Allosteric Transition: Atomically Detailed Simulations of Deoxy Scapharca Hemoglobin, *Biophysical Journal 92*, L85-L87.

174.   Faradjian, A. K., and Elber, R. (2004) Computing time scales from reaction coordinates by milestoning, *The Journal of Chemical Physics 120*, 10880-10889.

175.   Ananthan, S., Faaleolea, E. R., Goldman, R. C., Hobrath, J. V., Kwong, C. D., Laughon, B. E., Maddry, J. A., Mehta, A., Rasmussen, L., Reynolds, R. C., Secrist Iii, J. A., Shindo, N., Showe, D. N., Sosa, M. I., Suling, W. J., and White, E. L. (2009) High-throughput screening for inhibitors of Mycobacterium tuberculosis H37Rv, *Tuberculosis 89*, 334-353.

176.   Maddry, J. A., Ananthan, S., Goldman, R. C., Hobrath, J. V., Kwong, C. D., Maddox, C., Rasmussen, L., Reynolds, R. C., Secrist Iii, J. A., Sosa, M. I., White, E. L., and Zhang, W. (2009) Antituberculosis activity of the molecular libraries screening center network library, *Tuberculosis 89*, 354-363.

177.   Bolton, E. E., Wang, Y., Thiessen, P. A., and Bryant, S. H. (2008) Chapter 12 PubChem: Integrated Platform of Small Molecules and Biological Activities, In *Annual Reports in Computational Chemistry* (Ralph, A. W., and David, C. S., Eds.), pp 217-241, Elsevier.

178.   Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., Zhou, Z., Han, L., Karapetyan, K., Dracheva, S., Shoemaker, B. A., Bolton, E., Gindulyte, A., and Bryant, S. H. (2012) PubChem's BioAssay Database, *Nucleic Acids Research 40*, D400-D412.