Stony Brook University

Applying Computer Aided Drug Design To Target Multi-Drug Resistant Tuberculosis

A Dissertation Presented

by

Salma Banu Rafi

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 2008

Stony Brook University

The Graduate School

Salma Banu Rafi

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree,

hereby recommend acceptance of this dissertation.

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

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

Dale Deutsch, Ph.D. – Chairperson of Defense
Professor, Department of Biochemistry and Cell Biology

John Shanklin, Ph.D.
Senior Biochemist, Brookhaven National Laboratory
Adjunct Professor, Department of Biochemistry and Cell Biology,
Stony Brook University

Wendy Cornell, Ph.D.
Director, Molecular Systems Modeling Group, Merck & Co. Inc.

This dissertation is accepted by the Graduate School

Lawrence Martin

Dean of the Graduate School

Abstract of the Dissertation

Applying Computer Aided Drug Design To Target Multi-Drug Resistant Tuberculosis

by

Salma Banu Rafi

Degree of Philosophy

in

Biochemistry and Structural Biology

Stony Brook University

2008

One person dies every 15 seconds because of tuberculosis. Isoniazid (INH) is the frontline drug used to combat this disease caused by *Mycobacterium tuberculosis*. Isoniazid's proposed target is InhA, an enoyl reductase enzyme in fatty acid synthesis–II pathway. To inhibit InhA, isoniazid requires activation by catalase-peroxidase (KatG). Mutations in the *katG* gene are the single largest determinant of isoniazid resistance, thus complicating the TB epidemic by the rising tide of multi-drug- resistant strains. Our research focuses on designing new drugs that do not require activation.

Designing new inhibitors is a challenge due to a lack of detailed structural information about the interaction of the enzyme with its substrate and the acyl

carrier protein that delivers it. The same problem is encountered in InhA homolog *E.coli* enzyme, FabI. Firstly, we employed X-ray crystallography. Although the experiment revealed the relative orientation of the enzymes, a dynamic interface resulted in poor electron density for the substrate and side-chains in the protein:protein interface. Starting with this data, we successfully employed molecular modeling and MD simulations to predict the structure of the productive FabI-ACP complex. The resulting model is in excellent agreement with kinetic studies on wild-type and mutant FabIs. Importantly, the binding mode of the substrate differs from that of current inhibitors. Secondly, we complemented this structural study by applying the MM-PB(GB)SA to the binding by FabI of a series of triclosan analogs that span 450,000 orders of magnitude in affinity. The same compounds were screened experimentally. We obtained a 98% correlation between calculated and measured relative binding affinities.

Similar MM-GBSA calculations performed with InhA (*M.tuberculosis* enzyme) reveal that the binding energy is correlated to the extent of binding loop ordering, with loop ordering accompanied by stronger affinity. This is a critical observation since the FabI binding loop becomes ordered by triclosan binding, while that of InhA does not. We have performed simulations to test this hypothesis and determine how the inhibitor can be modified to optimize interaction with the binding loop and obtain a strong binding affinity to InhA, as is seen for FabI. Obtaining such lead compounds is a critical step toward effective and affordable treatment of drug-resistant TB.

This work is dedicated to my

mother **Jahan Ara**

and

bhaiya **Aslam**

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Abbreviation	Full Form of the Abbreviation
5PP	5-pentyl-2-phenoxyphenol
8PP	5-octyl-2-phenoxyphenol
ACP	Acyl Carrier Protein
AIDS	Acquired Immuno Deficiency Syndrome
CoA	Coenzyme A
COM	Center of mass
CPP	5-chloro-2-phenoxyphenol
CPU	Central Processing Unit
DOTS	Directly Observed Treatment, Short course
E. coli	Escherichia coli
EGB	Generalized Born energy
ENR	Enoyl ACP reductase
FAS	Fatty Acid Synthesis
FPP	5-fluoro-2-phenoxyphenol
GB	Generalized Born
HF	Hartree-Fock
HIV	Human Immunodeficiency Virus
INH	Isoniazid
M. smegmatis	Mycobacterium smegmatis
M. tuberculosis	Mycobacterium tuberculosis

MD	Molecular dynamics
MDR-TB	Multi Drug Resistant- Tuberculosis
MM	Molecular mechanics
MMGBSA	Molecular mechanics- Generalized Born Surface Area
MMPBSA	Molecular mechanics- Poisson Bolzmann Surface Area
MP2	Møller-Plesset perturbation theory of the second order
NADH/NAD^+	Nicotinamide Adenine Dinucleotide
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NMA	Normal mode analysis
NMR	Nuclear magnetic resonance
PB	Poisson Boltzmann
PCA	Principal component analysis
PDB	Protein data bank
PME	Particle Mesh Ewald
PP	2-phenoxyphenol
RESP	Restrained electrostatic potential
RMSD	Root mean square deviation
SASA	Solvent accessible surface area
TB	Tuberculosis
TCS	Triclosan
TIP3P	Transferable intermolecular potential 3 points
vdw	van der Waals
WHO	World Health Organization

XDR-TB	Extensively Drug Resistant – Tuberculosis

Publications

Salma B. Rafi, Guanglei Cui, Kun Song, Xiaolin Cheng, Peter J. Tonge, and Carlos Simmerling, *"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*" *Journal of Medicinal Chemistry*, 2006, 49, 4574-4580

Salma Rafi, Polina Novichenok, Subramaniapillai Kolappan, Xujie Zhang, Christopher F. Stratton, Richa Rawat, Caroline Kisker, Carlos Simmerling, and Peter J. Tonge "*Structure of Acyl Carrier Protein Bound to FabI, the FASII Enoyl Reductase from Escherichia coli" Journal of Biological Chemistry,* Dec 22 2006, Volume 281, Number 51, 39285-39293

Chapter 1 Introduction

1.1 Tuberculosis

1.1.1 Background and Significance

Tuberculosis (TB), an illness that mainly affects the respiratory system, is one of the world's most pernicious diseases and is caused by *Mycobacterium tuberculosis (M. tuberculosis)*. TB most commonly attacks the lungs (as pulmonary TB) but can also affect the central nervous system, the lymphatic system, the circulatory system, the genitourinary system, bones, joints and even the skin. *M. tuberculosis,* an intracellular pathogen survives and replicates within macrophages, the cells of the immune system that normally kill bacteria as part of the immune response to infection (*1, 2*). This triggers the release of toxic substances that the macrophage would normally utilize to kill the foreign organism, injuring of the surrounding tissues and forming cavitary lesions known as tubercles (*3*). Most infected individuals fail to progress to full-blown disease because the TB bacilli are 'walled off' by the immune system inside this tissue nodule. This stage is known as the latent form of TB and could, at a later stage, reactivate, for example in immuno-compromised individuals. If left untreated TB can be fatal.

Over one-third of the world's population has been exposed to the TB bacterium, and new infections occur at a rate of one per second (*4*). In 2004,

mortality and morbidity statistics included 14.6 million chronic active TB cases, 8.9 million new cases, and 1.6 million deaths, mostly in developing countries (*4*). There were an estimated 9.2 million new cases of TB in 2006 (139 per 100,000 population). This is an increase from 9.1 million cases in 2005, due to population growth. India, China, Indonesia, South Africa and Nigeria rank first to fifth respectively in terms of absolute numbers of cases. The African Region has the highest incidence rate per capita (363 per 100,000 population) (*5*). Without control, an estimated 35 million will die from TB by 2020. Due to its deadly impact, the World Health Organization declared TB a global health emergency in 1993, and the Stop TB Partnership developed a Global Plan to Stop TB that aims to save 14 million lives between 2006 and 2015 (*6*).

In addition, a rising number of people in the developed world are contracting TB because their immune systems are compromised by immunosuppressive drugs, substance abuse, or HIV/AIDS (*7*). The rise in HIV infections and the neglect of TB control programs have enabled a resurgence of TB. The emergence of drug-resistant strains has also contributed to this new epidemic with, from 2000 to 2004, 20% of TB cases being resistant to standard treatments and 2% resistant to second-line drugs (*8*). TB incidence varies widely, even in neighboring countries, apparently because of differences in health care systems. The resistant strains are created by bad medical practice such as over-prescribing antibiotics and mainly due to the long course of treatment. This provides the selection mechanism needed for the development of resistance in

those mycobacteria that are not destroyed by the drugs. The increased length of chemotherapy and the need for intensive initial drug therapy can be directly attributed to the slow doubling time of the tubercle bacilli of approximately 24 hours.

The first TB vaccine is Bacille Calmette-Guerin (or BCG). It is also the only available TB vaccine. BCG has very limited impact on the global TB burden. In young children, BCG vaccination prevents development of severe and fatal forms of TB, but it is not very effective in reducing the greater number of infectious cases in adults (*9*). An effective control strategy called DOTS (Directly Observed Treatment, Short course) has been introduced by WHO in response to the global TB epidemic. The main features of DOTS involve government commitment to TB control, a standardized treatment regimen with directly observed treatment for at least the first two months, a sufficient supply of drugs, case detection by sputum smear microscopy and a standardized recording and reporting system that allows assessment of treatment results. Although DOTS is highly effective, its implementation has been slow with overall low coverage. Also, this strategy is cumbersome and labor intensive.

The goal of the research is to design and develop new inhibitors to combat TB. The experimental data available does not give an atomistic view of the dynamic nature of the enzymes and their interactions. Dynamic information of the enzyme interactions with the inhibitor and enzyme interactions with the substrate

can help to accelerate the drug design efforts. Hence I employ molecular modeling and molecular dynamic simulations to gain insight into the enzyme-inhibitor and enzyme-substrate interactions.

1.1.2 TB Chemotherapy

The chemotherapy of infectious diseases, using sulfonamide and penicillins, had been underway in the late 1930's, but these molecules were ineffective against *M. tuberculosis*. In 1943, in test animals, streptomycin, purified from *Streptomyces griseus*, combined maximal inhibition of *M. tuberculosis* with relatively low toxicity. On November 20, 1944, the antibiotic was administered for the first time to a critically ill TB patient. The new drug had side effects - especially on the inner ear - but the fact remained, *M. tuberculosis* was no longer a bacteriological exception, it could be assailed and beaten into retreat within the human body.

A rapid succession of anti-TB drugs appeared in the following years. These were important because with streptomycin monotherapy, resistant mutants began to appear within a few months, endangering the success of antibiotic therapy. However, it was soon demonstrated that this problem could be overcome with the combination of two or three drugs.

Following streptomycin, *p*-aminosalicylic acid (1949), isoniazid (INH) (1952), pyrazinamide (1954), cycloserine (1955), ethambutol (1962) and rifampin (rifampicin; 1963) were introduced as anti-TB agents. Aminoglycosides such as capreomycin, viomycin, kanamycin and amikacin, and the newer quinolones (e.g. ofloxacin and ciprofloxacin) are only used in drug resistance situations. Combinations of a *B*-lactam antibiotic with a *B*-lactamase inhibitor enhance treatment effectiveness, but the newer drugs, including the macrolides, have not received much clinical testing.

Two properties of anti-TB drugs are important: antibacterial activity, (highest in INH, rifampin, streptomycin) and their capacity to inhibit the development of resistance (the most effective drugs being INH, rifampin, ethambutol).

1.1.3 MDR-TB, XDR-TB: need for new TB chemotherapeutics

Multidrug-resistant tuberculosis (MDR-TB) is TB due to organisms which show high-level resistance to both INH and rifampicin, with or without resistance to other anti-TB drugs. The molecular basis of resistance to INH and rifampicin (and some other drugs) is now largely understood (Table 1). Resistance to INH is due to mutations at one of two main sites, in either the *katG* or *inhA* genes (*10, 11*). Resistance to rifampicin is nearly always due to point mutations in the *rpo*

gene in the beta subunit of DNA-dependent RNA polymerase (*12*). These mutations are not directly connected, and so separate mutations are required for organisms to change from a drug-susceptible isolate to MDR-TB.

Table 1: Genetic sites for drug resistance in TB

Drug	Target	Gene
INH	Catalase-peroxidase enzyme	*katG*
INH–ethionamide	Mycolic acid synthesis	*inhA*
Rifampicin	RNA polymerase	*rpoB*
Streptomycin	Ribosomal S12 protein	*rpsL*
	16S rRNA	*rrs*
Quinolones	DNA gyrase	*gyrA*

MDR-TB is much more difficult to treat than sensitive TB, as it requires administration of more expensive, second-line antibiotics for up to two years. Consequently, there is a compelling need for the development of new TB therapeutics that are effective against both sensitive and drug-resistant strains of *M. tuberculosis*.

1.1.4 Mycobacterial Cell Wall

The secret of survival of the mycobacteria inside the host is their cell wall. Mycobacteria are surrounded by a thick waxy unusual cell wall, which ensures its survival inside human macrophages – cells that kill the invading bacteria. Mycobaterial cell wall consists of covalently linked complex of peptidoglycan, arabinogalactan and mycolic acids (*13, 14*). A key component of this is mycolic acids. These are very long chain (C60-C90) fatty acids that are esterified to an arabinogalactan polymer or to trehalose (Figure 1).

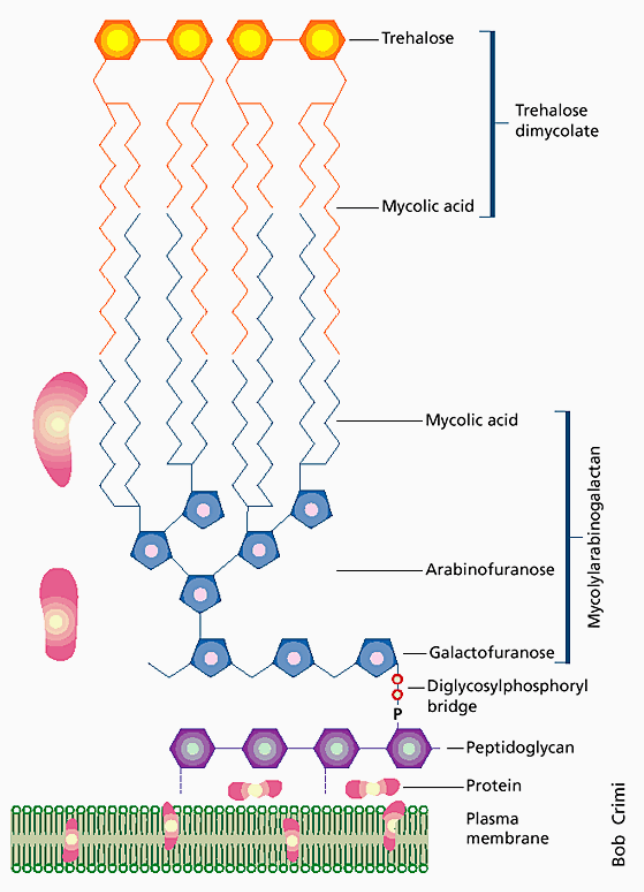


Figure 1: Schematic representation of Mycobacterial Cell Wall. Figure reproduced from (*15*)

Evidence to date strongly suggests that normal mycolic acid metabolism crucial to the FAS-II (Fatty Acid Synthase) system lies in the early stage of mycolic acid biosynthesis in mycobacteria (*16, 17*)

1.1.5 Fatty Acid Synthesis

Fatty acid synthesis (FAS) system is vital to all living organisms. Prior to 1958, it was assumed that FAS is accomplished by the reversal of the fatty acid beta-oxidation cycle for the degradation of fatty acids occurring in mitochondria (*18*). The two pathways were realized to be distinct upon the establishment that bicarbonate is required for fatty acid biosynthesis. Wakil and Brady demonstrated that bicarbonate is required for the synthesis of malonyl-coA which contributes the two carbon units for the growing fatty acid. It was later shown that the FAS enzymes were located in the cytosol rather that the beta-oxidation system in the mitochondria (*19, 20*). Using *E. coli* and yeast systems, all the enzymatic steps of the FAS pathway were identified by 1960 (*20, 21*). Following this, a great deal of progress was made on the enzymatic characterization of fatty acid biosynthesis system from *E. coli*. A key discovery was the identification of the acyl carrier protein (ACP) which served as the substrate carrier for the reaction intermediates covalently attached to this protein (*22*).

There are two principal classes of fatty acid synthases. FAS Type I systems utilize a single large, multifunctional polypeptide and are common to both mammals, birds and fungi (although the structural arrangement of fungal and mammalian synthases differ). FAS Type II, or bacterial systems and plants, use discrete, monofunctional enzymes which are used iteratively to elongate and reduce the fatty acid chain (*23*). Mammalian FAS-I is a 550 Da homodimer with multiple structural and functional domains. This is also known as the “associated” system. Here, all the enzymatic activities and the acyl chain carrier, ACP, reside on one single multifunctional polypeptide chain. The major product of FAS-I is palmitic acid (C16) which is released from the complex by thioesterase (*24, 25*). Fatty acids are the main constituents of fat, membrane lipids and substrates for protein acylation in humans. Diet regulates the expression of FAS and its activity (*26*).

FAS-II, where all the seven individual enzyme activities and the ACP exist as separate proteins encoded by unique genes, is aptly called “dissociated” system (*27*). The biosysnthesis of fatty acids is vital to the survival of the bacteria as it is the first step in the formation of membrane lipids (*28*). Regardless of the type of FAS system, the actual reactions are identical in sequence. Repeated steps of condensation, reduction, dehydration and another reduction results in the formation of long chain fatty acids. Mycobacteria are unusual in that they possess both the FAS-I and FAS-II systems. This includes *M. tuberculosis,* the first prokaryote shown to have the FAS-I system (*29*). In these organisms, the

products of FAS-I are fatty acyl CoA substrates (24-26 carbon atoms in length) that are fed into the FAS-II system, which further elongates the FAS-I primers to about 56 carbon atoms (*23*). ACP acts as the substrate carrier for the FAS-II cycle.

The *E. coli* FAS-II pathway is the most extensively studied system of all the FAS-II pathways known and serves as a paradigm for understanding all FAS-II systems (*30*). A conserved serine residue in ACP carries all the intermediates in this pathway through a 4'-phosphopantetheine moiety. The acyl chain elongates by two carbon atoms every time it passes through the FAS cycle. Acetyl-coA acts as the primer unit and malonyl-CoA as the elongator (*31*). The major products of this pathway are C16-C18 straight chain fatty acids (*32*). The basic steps of FAS-II are common to all bacteria and the genes for individual proteins are generally identified by homology to the corresponding *E. coli* enzymes (*30*). The first enzyme in the FAS pathway is acetyl-CoA carboxylase, which is a multi subunit enzyme encoded by four different genes *accABCD* (**Figure 2**). The product of this reaction is malonyl CoA which is transferred to the ACP by malonyl-CoA:ACP transacylase, FabD. The fatty acid elongation is initiated by the condensation of malonyl-ACP and acetyl-CoA to yield beta-ketobutyryl-ACP and CO_2. The reaction is catalyzed by beta-ketoacyl-ACP synthase III (FabH). The next four reactions are repeated until fatty acids of the appropriate chain length are synthesized. First, beta-ketoacyl-ACP reductase (FabG) catalyzes the NADPH-dependant reduction of beta-ketobutyryl-ACP to

beta-hydroxyacyl-ACP. Next, either of the two dehydratases, FabA or FabZ, catalyzes the removal of water to produce trans-2-enoyl-ACP. The final step of the elongation cycle is catalyzed by the enoyl-ACP reductase (FabI, FabK or FabL) which catalyze the reduction of the enoyl-ACP to yield saturated acyl-ACP. Subsequent cycles of this pathway are initiated by the condensation of the respective acyl-ACP with malonyl-ACP by the beta-ketoacyl-ACP synthase FabB or FabF. FabI was demonstrated to be the only enoyl-ACP reductase in *E. coli* (*33*). However, two isoforms of this enzyme (FabK, FabL and FabV) were recently identified in other organisms through genetic experiments (*34, 35*).

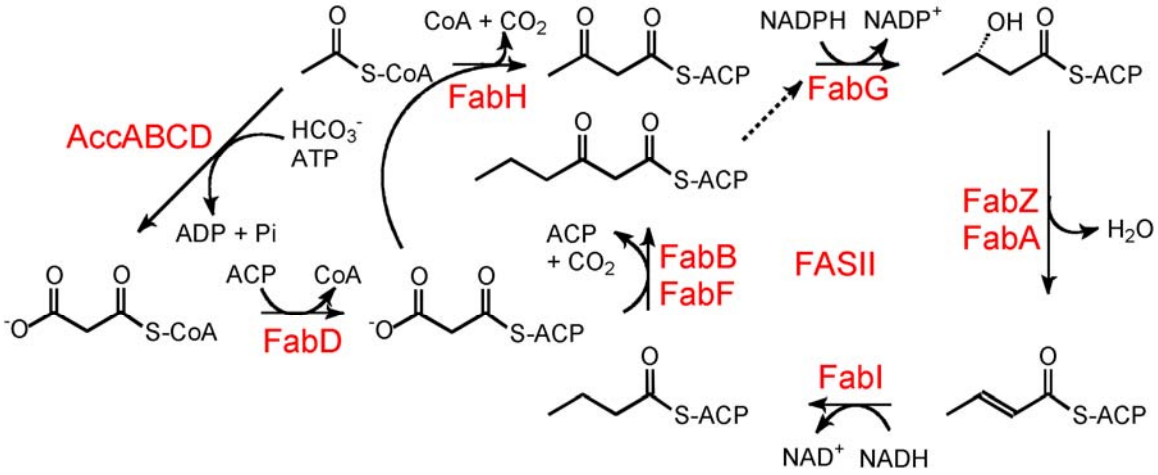


Figure 2: Bacterial Fatty Acid Synthesis (FAS-II)

Bacterial FAS system has attracted a great deal of interest to itself as it has been repeatedly validated to be an excellent target for therapeutic intervention for two reasons. Firstly, the lack of similarity between the mammalian and bacterial FAS system, and the differences in the active site organization,

make this pathway a good target for type selective intervention (*36*). Secondly, this pathway is vital for the survival of several pathogens, such as *E. coli, M. tuberculosis, M. smegmatis* (*37, 38*). Many of the currently available and widely used anti-bacterial agents target the FAS-II pathway, especially the condensing enzymes and the enoyl reductases.

Bacterial enoyl ACP reductase (ENR) catalyzes the NADH-dependant reduction of the trans double bound between the C2 and C3 of the fatty acyl substrate (Figure 3).

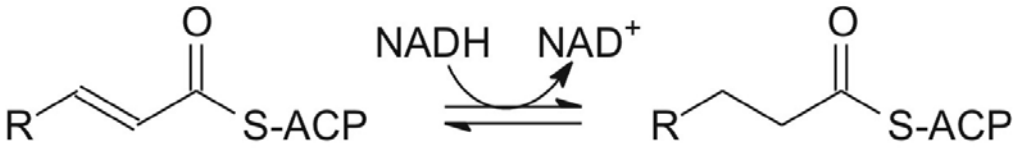


Figure 3: Reaction catalyzed by Enoyl-ACP reductase

This reaction is essential and is the last step of the FAS system. FabI was shown to play a determinant role in completing cycles of fatty acid biosynthesis pathway in *E. coli*. Extracts of temperature sensitive FabI mutants, fabI (T_s), were unable to synthesize saturated fatty acyl ACPs and addition of FabI protein restored acyl-ACP production. Since the equilibrium of the dehydration reaction lies much in favor of beta-hydroxyacyl-ACP, FabI was demonstrated to “pull” the fatty acids through the cycle (*33*).

1.1.5.1 Inhibitors of Enoyl-ACP Reductase Enzymes

There are three types of enoyl-ACP reductase inhibitors: diazaborine, INH and triclosan (TCS). All three inhibitors bind the ENRs by forming a tight complex with the cofactor NADH. Diazaborines and TCS bind the the ENR•NAD^+ form of the enzyme and interact with the 2'-hydroxyl group of the co-factor. INH occupies a slightly different position in the active site of *M. tuberculosis* enoyl reductase InhA and forms a covalent bond with the carbon at 4' position of the co-factor which is believed to participate in the transfer of an hydride ion during the catalysis. Diazaborines and INH form covalent bonds with the cofactor, but TCS does not.

1.1.5.1.1 Diazaborine

Diazaborines represent a group of antimicrobial agents what contain a heterocyclic 1,2-diazine ring and boron as a third heteroatom. The antibacterial activity of diazaborines is almost exclusively confined to gram-negative bacteria and *M. tuberculosis* (*39*).

1.1.5.1.2 Isoniazid

Isoniazid is also called isonicotinyl hydrazid or INH (**Figure 4**). INH is a first-line anti-TB medication used in the prevention and treatment of tuberculosis. It was discovered in 1952 by Roche, when for the first time, a cure for TB was considered reasonable.

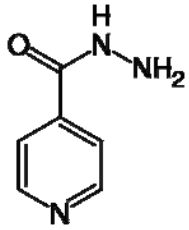


Figure 4: Isoniazid

The mode of action and the cellular target remained unknown for many years. The early stages of INH-resistant strains of *M. tuberculosis* showed a correlation between resistance and catalase-peroxidase activity, both encoded by a gene, *katG* (*10, 40*). This means that the product of the *katG* gene plays an important role in INH inhibiting the target enzyme, which suggests that intracellular activation may be necessary. Indeed, INH is a prodrug and must be activated by catalase-peroxidase enzyme katG to form isonicotinic acyl anion or radical (*41*). These forms then react with a NADH radical or anion to form isonicotinic acyl-NADH complex (Figure 5). This complex binds tightly to keto-enoyl reductase known as InhA and prevents access of the natural substrate. This mechanism inhibits the synthesis of mycolic acid in the mycobacterial cell wall.

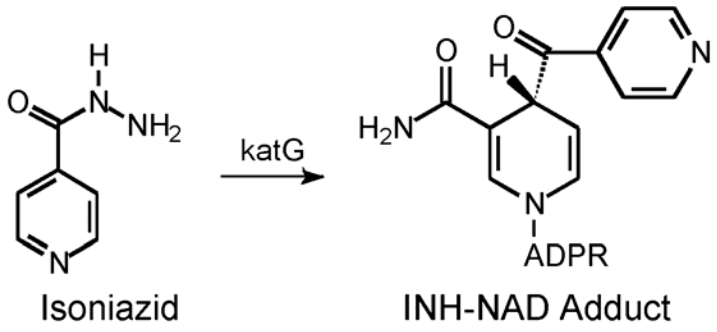


Figure 5: Activation of Isoniazid by katG

In *M. tuberculosis* besides enoyl-ACP reductase (InhA), another enzyme, beta-ketoacyl-ACP synthase (KasA) has also been proposed as the target of INH (*42*). Both these enzymes are part of the FAS-II system. Through genetic experiments it was shown in 1994 that INH specifically targets the *inhA* gene (*43*). Also, the product of this gene was shown to be NADH dependant enoyl-ACP reductase with long chain substrate specificity (*44*). Currently 25% of *M. tuberculosis* strains resistant to INH are due to direct mutations with the *inhA* gene (*45*). The most common mutation is S95A of InhA which is a similar position as G93 in FabI. Other mutations lead to decreased affinity of NADH or up-regulated expression of wild type enzyme (*44, 46*) Biochemical studies in the Tonge laboratory have shown that INH-NAD adduct is a slow, tight binding competitive inhibitor of InhA with an overall dissociation constant of K_i = 0.75 nM (*42*). The crystal structure of the binary complex of InhA:INH-NAD adduct has been solved in 1998 by Sacchetini and coworkers (*47*) (Figure 6, Figure 7). The isonicotinic-acyl group is in the place of the 4S hydrogen of NADH which is involved in hydride transfer to the enoyl-ACP substrate.

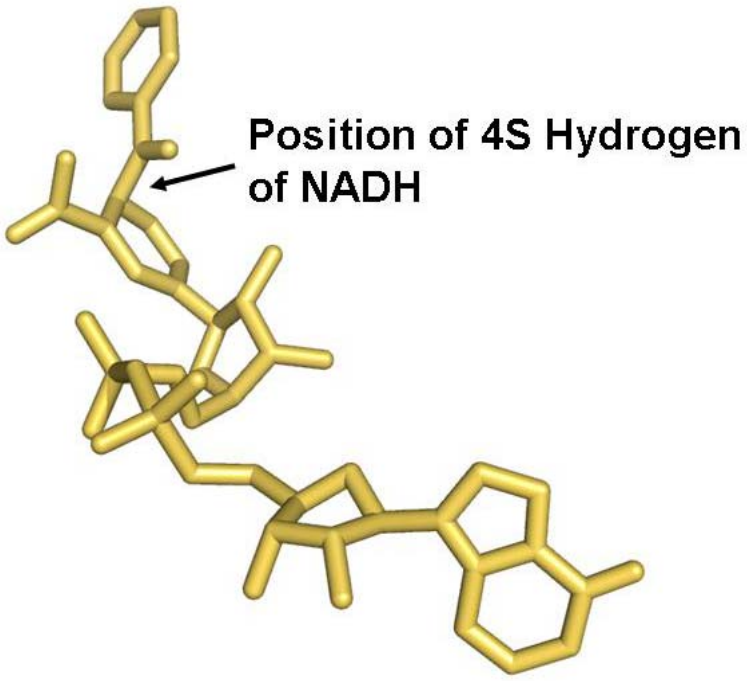


Figure 6: INH-NAD adduct in the active site of InhA enzyme

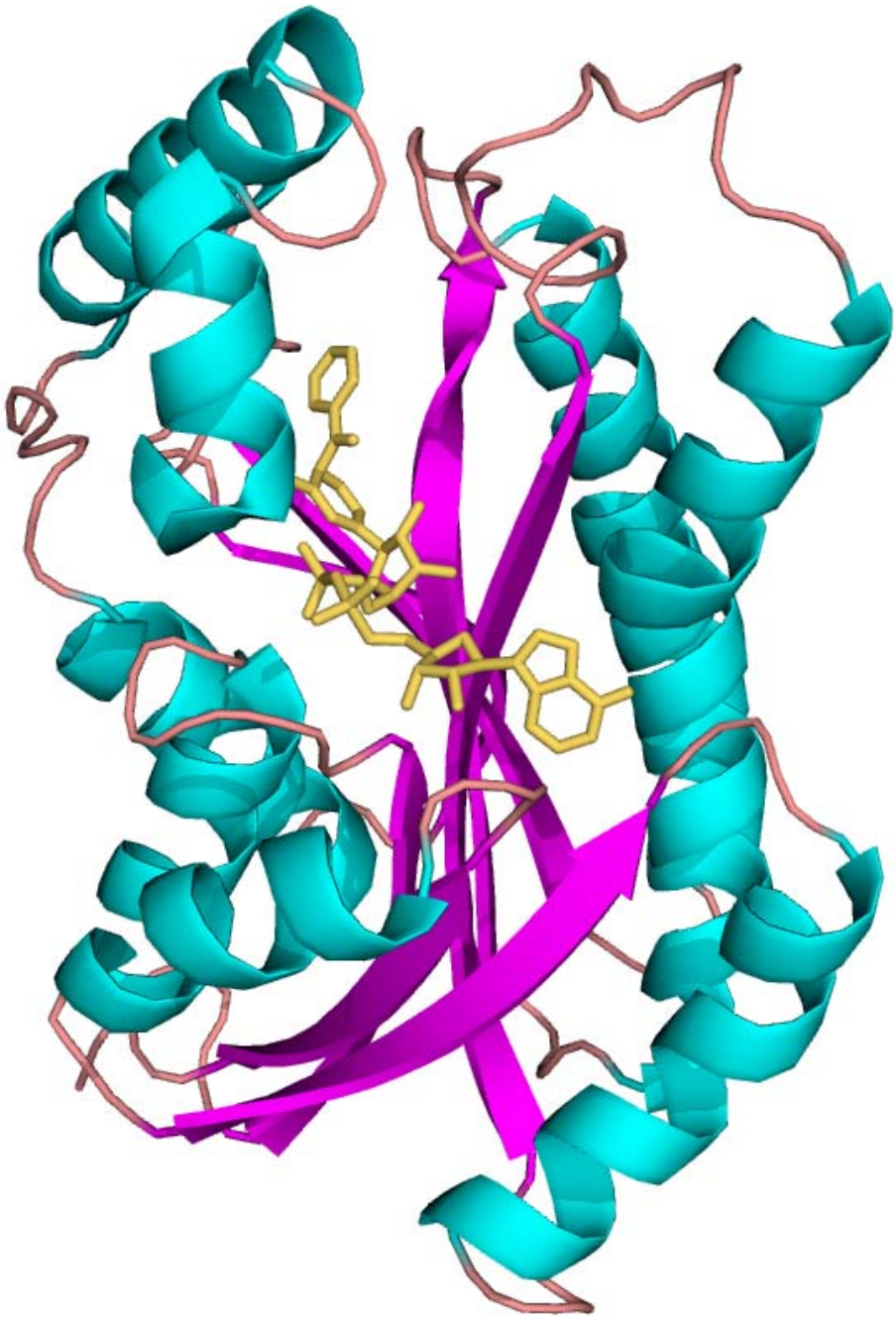


Figure 7: Crystal structure of InhA with INH-NAD adduct bound in the active site (PDB ID: 1ZID) (*47*). The protein is shown in cyan and purple, the INH-NAD adduct is shown in yellow.

1.1.5.1.3 Triclosan

TCS (2,4,4'-trichloro-2'-hydroxydiphenyl ether) is a broad spectrum antibacterial agent that was originally patented as a herbicide (Figure 8) (*30*). Because of its favorable safety profile, for about 30 years TCS is being used in soaps (0.15-0.30%), deodorants, toothpastes, shaving creams, mouth washes, and cleaning supplies and is infused in an increasing number of consumer products, such as kitchen utensils, toys, bedding, socks, and trash bags, sometimes as the proprietary Microban treatment. It has been shown to be effective in reducing and controlling bacterial contamination on the hands and on treated products (*48, 49*).

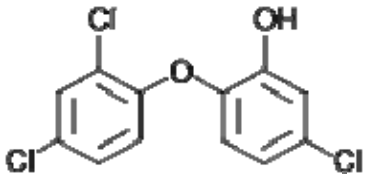


Figure 8: Triclosan

Untill recently it was assumed that TCS, being a small hydrophobic molecule, caused a general and non-specific membrane disruption via diffusion thus making the bacterial envelope more porous and preventing the uptake of nutrients. Therefore, the aquisition of cellular resistance would be unlikely (*50*). However, in 1998, it was demostrated through genetic experiments that TCS specifically targets fatty acid biosynthesis. Levy and coworkers isolated TCS-resistant *E. coli* and mapped the resistance to the *fabI* gene (*51*). Also, TCS

resistant G93V mutant of FabI was isolated and the purified protein was resistant to inhibition by TCS whereas the wild type protein was not (*52*). It was also shown that another mechanism of resistance was upregulation of *fabI* expression. It was noticed that inhibition of *E. coli* fabI with TCS increased over time. Through a detailed kinetic analysis, TCS was shown to be a slow, tight-binding, picomolar inhibitor of FabI, binding to NAD^+ form of the co-factor (*53, 54*).

The resistance to G93V mutant to TCS inhibition was also explained by the biochemical experiments. It was shown that G93V mutation reduces affinity of the enzyme for TCS by 9000 fold with little effect on catalysis. Though TCS bound to the FabI•NAD^+ complex, it did not exhibit slow onset inhibition. The G93V mutation could obstruct TCS binding site (*55*). The mutation G93S leads to diazaborine resistance (*56*), and mutation of the homologous residue in InhA (S94A) results in decreased sensitivity of InhA to INH (*44, 46*). In addition, although the S94A InhA mutation also gives rise to TCS resistance in *Mycobacterium smegmatis (57)*, it does not appear to change the affinity of TCS for InhA (*58*). The mutations at this position have little effect on the kinetic parameters for substrate reduction and the inhibitor binding is severely compromised. Thus, this residue is a general determinant of bacterial resistance to FabI inhibitors. Clearly then, novel FabI inhibitors must be designed to circumvent mutation at G93 (FabI) or S94 (InhA). Numerous crystal structures of

FabI in complex with NAD^+ and TCS were determined by several laboratories (*53, 54, 59-63*).

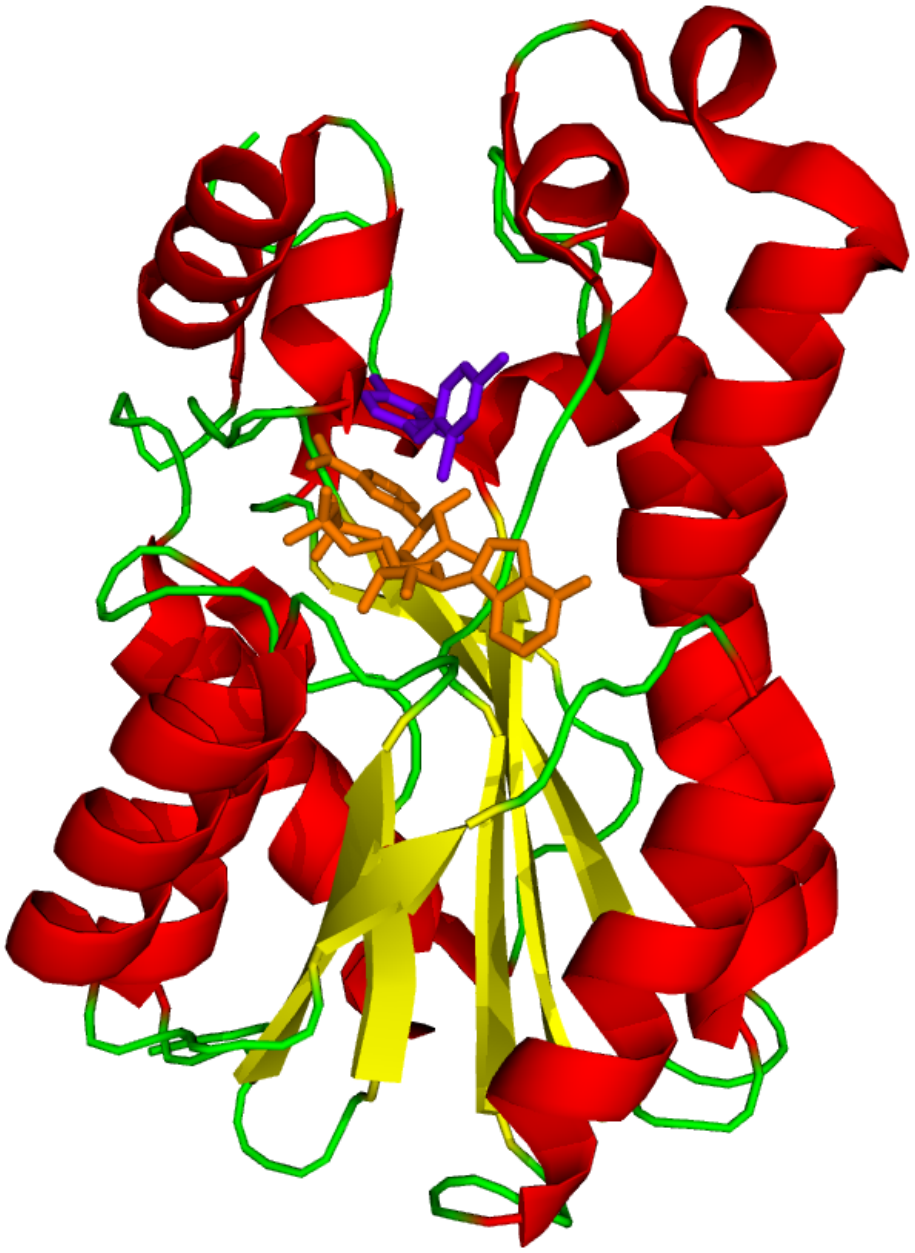


Figure 9: Crystal structure of FabI with TCS and NAD^+ bound in the active site (PDB ID: 1QSG). The protein is shown in red, yellow and green, TCS is shown in violet and NAD^+ is shown in orange

These crystal structures reveal that the tricloan binding in the active site of FabI is very simiar to the binding of diazaborines (Figure 9). TCS is positioned in close proximity to the cofactor with hydroxychlorophenyl ring coplanar to the nicotinamide ring of NAD^+. This allows for π-π stacking interactions which are also observed in all of the diazaborine structures. The inhibitor is bound in the hydrophobic pocket of FabI formed by amino acid residues G93, F94, A95, Y146, Y156, M159, K163, I200 and F203. The hydroxyl group of TCS hydrogen bonds to the Y156 OH and the 2' hydroxyl group of the cofactor (Figure 10).

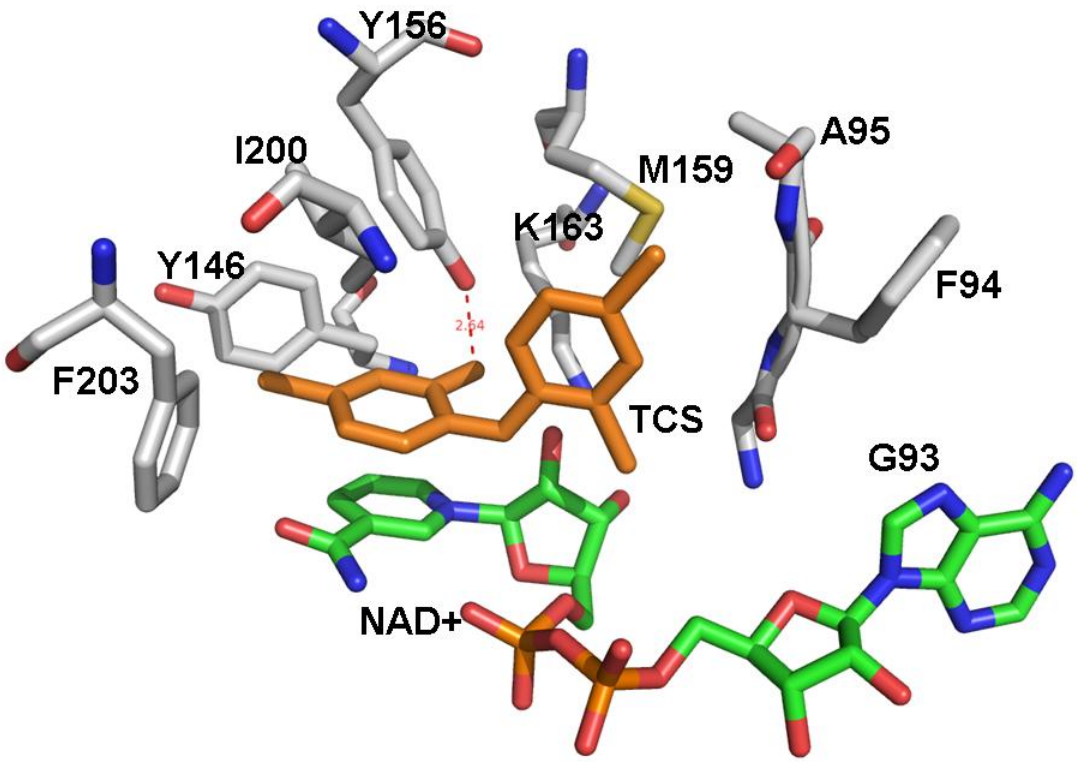


Figure 10: Active site of FabI with TCS and NAD^+ bound in it (1QSG.pdb)

A key difference in the crystal structure of ternary complex of FabI•NAD^+•TCS comparing to the structure of the binary complex of FabI and NAD^+ is that the substrate binding loop comprising of residues 196-205 is ordered whereas these residues are disordered in the binary complex. TCS is a slow onset inhibitor. The proposed mechanism of FabI inhibition by TCS in given in Figure 11. TCS binds exclusively to the FabI•NAD^+ form of the enzyme. This can be understood by a 2 step inhibition mechanism in which the initial rapid binding of the inhibitor to the enzyme is followed by a second slow step that results in the final enzyme-inhibitor complex. We believe that the slow step in inhibition is coupled to the conformational change in the protein, namely, the substrate binding loop ordering.

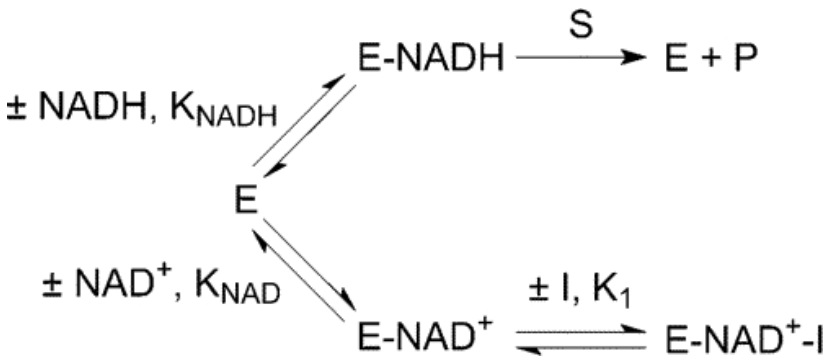


Figure 11: Proposed Mechanism of inhibition of FabI by TCS (55)

The two step inhibition mechanism can be represented by **Scheme 1**. Here, k_2 represents the slow step of the inhibition.

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

Scheme 1: Two step inhibition of FabI

This reaction can be represented in terms of free energy and the reaction coordinate as shown in Figure 12. In order for the enzyme to reach the EI* state, it has to overcome a huge energy barrier ($‡_2$). We hypothesize that this is coupled to the conformational change in the protein upon inhibitor binding like the substrate binding loop ordering.

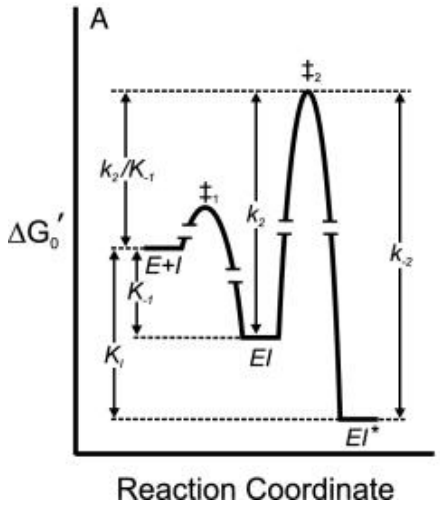


Figure 12: Free energy profile for the interaction of TCS with FabI. The free energy of the transition state for the conversion of EI to EI* ($‡_2$) is determined from k_2.

It was first recognized by Heath and coworkers that the mutations in *fabI*, responsible for resistance to TCS were analogous to the mutations in *inhA,* imparting resistance to INH in *M. tuberculosis (36, 52).* Further kinetic studies have shown that TCS also inhibits the InhA ezyme from *M. tuberculosis,* binding

to the enzyme with submicromolar affinity. Importantly, it has been demonstrated that INH-resistant mutants were sensitive to TCS inhibition (*58*). The crystal structure of the ternary complex of InhA•NAD^+•TCS has been recently solved by Kuo *et. al.* and also by our collaborator Dr. Kisker (*64*).

TCS binds to InhA in a similar way as it binds with FabI. The *M. tuberculosis* enoyl reductase (InhA) is 36% identical and 67 % similar to the *E. coli* (FabI) (*58*). However, there is a 30,000 fold difference in binding of TCS to FabI versus InhA with K_i values of 7 pM and 0.22 μM, respectively (*50, 58*). It is to be noted that TCS is a slow onset inhibitor of FabI, binding preferentially to E•NAD^+ form of the enzyme, but is a classical, uncompetitive inhibitor of InhA binding to the E•NAD^+ form as well (*55, 58*).

1.1.6 Acyl Carrier Protein (ACP)

1.1.6.1 Background and significance

The acyl carrier protein (ACP) is an important component in fatty acid biosynthesis, lipid A formation, membrane-derived oligosaccharide, lipoic acid and polyketide biosynthesis with the growing chain bound during synthesis as a thiol ester at the distal thiol of a 4'-phophopantethiene moiety (*65, 66*). ACP, thus has the remarkable capability to recognize very diverse proteins. Whereas ACP is the substrate carrier in the FAS system, beta-oxidation of fatty acid utilizes coenzyme A as the substrate carrier. FAS pathway is a synthetic pathway,

whereas beta-oxidation of fatty acids is a degradative pathway. Utilization of 2 different forms for substrate carriers viz., ACP and CoA, permit the organism to distinguish between the synthetic and degradative cycles in fatty acid metabolism (*67*).

It was first demonstrated in the fatty acid biosynthesis system that a protein can act as a coenzyme (*21, 68, 69*). All the enzymatic components of the FAS system were identified by the year 1961. Lynen and coworkers attempted to isolate the free intermediates in the synthetic process catalyzed by a purified fatty acid synthesizing complex derived from yeast. Their repeated failures led them to propose that FAS intermediates are protein bound (*70*). Isolation of pure ACP and the chemical and enzymatic synthesis of ACP thioester substrates have allowed the elucidation of reaction sequence which explains FAS.

The protein is expressed in the inactive *apo* form which is modified with the 4'-phosphopantetheine moiety. This is post-translationally attached to a conserved serine residue on the ACP by the action of holo-acyl carrier protein synthase (ACPS), a phosphopantetheinyl transferase in a magnesium dependent reaction (*71-73*). Although majority of experiments have been carried out with *E. coli* ACP, a protein that functions like ACP has been identified in every biological system which catalyzes the de novo synthesis of fatty acids (*74*). The ACPs are small negatively-charged α-helical bundle proteins with a high degree of structural and amino acid similarity. ACP is located in the cytoplasm of the cell

and is considered to be one of the most abundant proteins in *E. coli* (~60,000 molecules per cell) (*75*). The structures of a number of ACPs have been solved using various NMR and crystallography techniques.

1.1.6.2 Role of ACP in fatty acid biosynthesis

A great deal of interest in ACP was generated due to its central importance in FAS pathway, which has been repeatedly been validated as an important target for antibiotic development. ACP is responsible for acyl group activation and it is a necessary cofactor in all FAS reactions, including initiation, elongation, and transfer to the membrane bilayer (*18, 32, 76, 77*). A schematic view of the ACP molecule with an acyl phosphopantetheine chain attached to S36 is given in Figure 13. The 4'-phosphopantethiene moiety has two crucial functions: provides a flexible chain that can reach into the active sites of ACP-recognizing enzymes and serves as a point of attachment of the acyl groups (*78*).

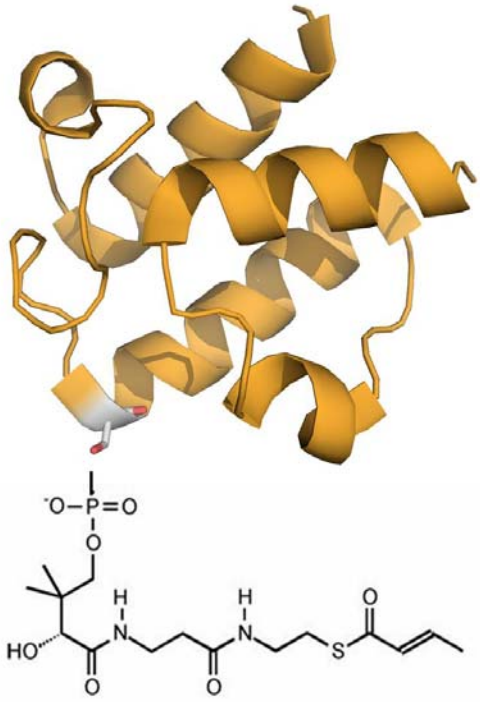


Figure 13: ACP molecule with the acyl-phosphopantetheine chain attached to S36

Protein-protein interactions are of central importance in a variety of biological processes like cellular structure, immune response, signal transduction and apoptosis to just name a few (*79*) and hence there has been a growing interest in understanding and manipulating them. ACP has been identified to be vital for the survival of *E. coli* (*80*). Yet, the details of the molecular recognition of ACP by the FAS enzymes are poorly understood since no crystal structures of ACP in complex with any FAS enzymes are available. There is a great deal of genetic, biochemical, and crystallographic evidence which points towards fatty acid biosynthesis enzymes being important antibacterial targets. Hence, identifying the specific interactions which govern ACP recognition by the FAS

enzymes provides an exciting new area for drug development. Designing small peptide antagonists of ACP.protein interfaces, would result in failure to deliver the substrate into the active site of enzyme and consequently, synthesize fatty acids and eventually in bacterial growth inhibition.

1.2 Molecular Dynamics Simulations

Structure based drug design and structural biology widely use molecular dynamic (MD) simulations (*81-83*). MD simulations have several advantages when compared with conventional experimental methods. First, they exhibit biological events occurring on very short time scale. Unlike many experimental structural biology methods, which only generate time-averaged results, the timescale of MD simulations can be infinitely small, usually between femtoseconds to picoseconds. Second, MD simulation provides an approach to directly observe the biological events at atomic resolution. The time-sequence trajectories generated by MD simulation exhibit the atomic motions with the aid of graphic software such as MOIL-view (*84*), pymol (*85*) and VMD(*86*). Third, MD simulations can be used to explain biological events not only qualitatively, but also quantitatively. A relatively new concept, “iso-structure does not necessarily imply iso-energy” (*87*), is now generally accepted in the structural biology community. This means that the systems with same or similar structures can be studied using the computational approaches to evaluate the energy differences.

The goal of my research is to use MD simulations to simulate the biological events, explain the relevant experimental results from a microscopic view, and understand the mechanisms of enzymes from both structure and energy points of view.

1.2.1 Basics, history and perspectives

Although MD simulations have been widely used as an essential tool for biological and other systems, its fundamental principle is rather simple. The particles in the systems are defined as spheres. Their movements obey Newton's three laws of motion. In the simplest hard-ball model, before the collision, particles move with constant velocity. After the collision, the change of the velocity is directly proportional to the force acting on the particle and inversely proportional to the mass of the particle. During the collision, to every force applied there is an equal force applied in the opposite direction. In real applications, more smooth functional forms are usually used to replace this hard-ball model.

MD simulations were first introduced by Alder and Wainwright in 1957, and applied on hard sphere model to study phase transitions(*88*). The first MD simulation of a realistic system was done by Rahman and Stillinger in their simulation of liquid water in 1974(*89*). The first protein simulations were carried

out in 1977 by MacCommon et. al on the bovine pancreatic trypsin inhibitor (BPTI) (*90*) . Although the simulation was short (9.7 ps) and in vacuum, the results paved the way for understanding protein motion as essential in function.

Since first applied to biomolecular systems 30 years ago, MD simulations have been widely used in studies of biomolecules. One common pattern in applying MD in biological systems is using structural experimental methods, such as NMR or x-ray crystallography to capture static snapshots, or kinetic experimental methods to obtain macroscopic data. Then MD simulations are used to mimic the motions of proteins from one conformation to the other or reproduce the macroscopic data to explain the results using microscopic structures. It is always important to compare results of MD simulations with relevant experimental data. Numerous MD simulations have been applied to complex biological problems such as protein folding, ion channel, motor protein function, and enzyme catalysis etc. Two recently reviews have summarized most recent applications of MD simulations.(*81, 82*).

Despite countless efforts to improve the efficiency and accuracy of MD simulations, three challenges remain for conducting a successful MD simulation: force field, searching and sampling. The accuracy of the force field is essential for the accuracy of the simulation. Typically, a force field is defined by an equation and a parameter set accommodating the equation. The parameter set of a force field is usually generated by fitting on the quantum calculation results

or experimental results. The performance of a force field can be biased towards the native conformation or the most stable conformation. A well known example is that most commonly used force field, amber ff94, which has a strong bias favoring α-helix conformation (*91*). The Generalized Born (GB) model, an implicit way to simulate solvent effects, has also been demonstrated that the GB model has the tendency to over-stabilize salt bridges (*92*). Therefore, careful evaluation of the force field on the subject system before production simulation is essential.

In spite of the increasing computing power, searching for alternative stable conformations besides the initial structure or native structure remains a challenging problem. It has long been realized that many biomolecules could have more than one stable conformation, which means besides the native conformation, there are intermediate conformations which are also very important. Many conformational changes are correlated with enzyme functions. However, these stable conformations are separated by energy barriers. Overcoming these energy barriers and searching for alternative conformations is a major challenge in studying protein function by computational methods. Simulated annealing (*93*), softcore potential (*94*), targeted MD and other methods have been used to conquer this problem.

Sampling is a step further than searching. We can use a simple 2-state-model system as an example. Starting from one state, once the simulation reaches the other state, it can be called a successful searching. Sampling is for

generating the Boltzmann populations of the two states. The relative populations of the two states are determined by their free energies, and at equilibrium, they are constant. A successful sampling needs “correct” potential energy for both states and multiple transitions between the two states to guarantee that the sampling is enough.

The constant increase in computing power is drawing a more and more promising perspective for biomolecular MD simulations. The increase of computing power consists of two components. One is the evolution of the computer hardware. Over the past few decades, hardware computing power has increased by a factor of 10 about every 5 years. The other improvement is from the creation of more efficient computing algorithms. Table 2 below shows the timings of a standard Amber benchmark over a decade (*95*). The simulated system is dihydrofolate reductase (159 residues) in TIP3P water (23,558 atoms). The table shows the speed on a single CPU (Intel Xeon x86_64, 3.4 GHz). The increase in speed is solely due to the improvement of the algorithm.

Table 2: AMBER benchmarks

Code	**Release date**	**speed, ps/day**
Amber 4.1	June, 1995	103
Amber 5	November, 1997	104
Amber 6	December, 1999	121
Amber 7	March, 2002	135
Amber 8	March, 2004	179
Amber 9	March, 2006	249
Amber 10	April, 2008	314

Along with the rapid growth of computing power, great efforts have been made to improve the accuracy of the current force field and to design more sophisticated force fields. For example, in MD simulations, atoms are considered as the elementary particles and no electrons are included, so, in order to describe the enzyme reaction we will need hybrid quantum-classical (Quantum mechanics/Molecular mechanics) modeling. The individual behavior of the subatomic particles that make up all forms of matter, can often be satisfactorily described using quantum mechanics (QM). In molecular mechanics (MM), equations are used to calculate the interations and the corresponding energies. These equations follow the classical laws of physics and the atoms are represented as particles (nuclei) without considering the electrons. The atoms are connected via bonds, this means that it is not possible to simulate reactions in molecular mechanics. Bonds cannot be formed or broken. QM/MM approach is a molecular simulation method that combines the strength of both QM (accuracy) and MM (speed) calculations.

1.2.2 Force field

A force field is the mathematical description of physical interactions within a system. The force field used in classical MD specifies how an atom interacts with the rest of the system. Simply, a force field is an equation with a set of parameters. It uses the nuclear positions of the atoms to calculate the potential

energy of the system, usually with some additional terms to improve the accuracy of the output.

The Amber force field is one of the most commonly used force fields, and has been used through out my study (*96*). There are four components in the Amber force field: bond energy, angle energy, torsion energy, and non-bonded interaction energy. Two different force fields can have different parameters or functional forms.

Error! Objects cannot be created from editing field codes.

Error! Objects cannot be created from editing field codes.

1.1

Equation 1.1 is the general form of the Amber force field (*96*). The first term is the bond term which models the interactions between pairs of bonded atoms. The second term is the summation of all the valence bond angles bending. These two terms are modeled by the harmonic potential, k_i and k'_i are the force constants, and l_i-$l_{i,0}$, Θ_i-$\Theta_{i,0}$ are the deviations from their respective equilibrium values. The third term is the torsional potential that models how the energy changes as a bond rotates. The fourth term models the non-bonded interactions which are usually modeled by using the Coulomb potential term for electrostatic interactions and a Lennard-Jones potential for van der Waals interactions. Usually the largest difference between two different force fields

comes from the philosophies for optimization of the non-bonded parameters. An illustration of these four terms is shown in Figure 14.

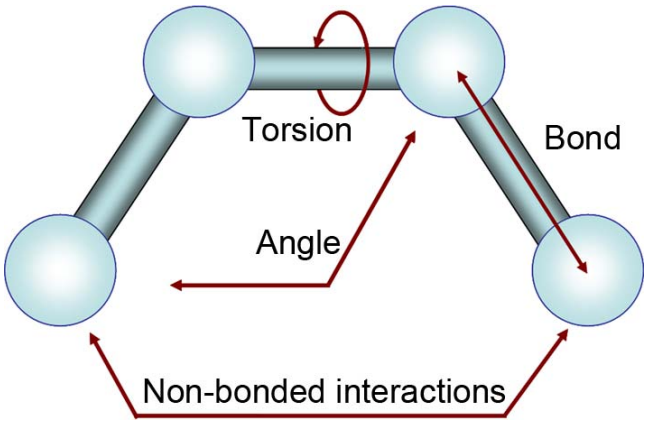


Figure 14: Four types of interactions in amber force field

One of the most important qualities of force field is that its functional form and parameters must be transferable. This means that the same set of parameters can be used to model a series of related molecules. However, careful evaluation of the accuracy of the force field on the subject before further analyzing simulation results is always critical.

1.2.3 Solvation effects

In most cases we are interested in the properties of molecules in solution, usually in aqueous solution. Therefore, it is essential to calculate not only the interactions between the atoms of biomolecules, but also the solvent effects on

these interactions. There are two different approaches to include the solvent effects in molecular dynamics. One is the explicit solvent model, the other one is the continuum solvent model, also known as the implicit solvent model.

In the explicit solvent model the solvent molecules are included in the calculation. The simplest and most commonly used model is the TIP3P water model (*97*), in which the water molecules are rigid and only non-bonded interactions involve water molecules. The electrostatic interaction is calculated based on Coulomb's law and the dispersion and repulsion forces are calculated using the Lennard-Jones potential. The charges are located at the centers of three atoms (that is what "3" stands for). There are several other water models, such as TIPS (*98*), SPC (*99*) which uses a similar strategy. In complex water models, there are more charge sites on dummy atoms to represent a more realistic electron distribution around the oxygen atom, such as TIP4P(*97*), TIP4Pew(*100*), TIPS2 (*101*), TIP5P (*102*). There are also studies for using flexible water models (*103*) and including polarization effect (*104*).

In explicit solvent models we need to calculate the interaction between the particles and every solvent molecule individually, therefore it is a very "computationally expensive" job, and the cost increases with the number of interaction sites in the water model. Algorithms such as Particle Mesh Ewald method (*105*), particle-particle/particle-mesh (P3M) method (*106*) are developed to accelerate the simulations.

Even with advanced algorithms, one obvious drawback of the explicit solvent model is the large system size due to numerous solvent molecules. An alternative approach is using the continuum solvent models, such as the Generalized Born (GB) model. This model represents the solvent implicitly as a continuum with the dielectric properties of water, and also includes the charge screening effects of salt. This lowers the calculation expense in two aspects: first, it removes the calculation of the interactions and motions involving solvent molecules; second, the absence of solvent friction accelerates the dynamics of the solute. However, the computing cost of the GB model scales up with the size of the system. For large systems, the computing cost may be greater than using the explicit solvent model.

Both explicit solvent and implicit solvent models have certain strengths and weaknesses. Calculations using explicit solvent generally yield more accurate results. In some cases, such as simulations involving water bridges, explicit water molecules are essential for the calculation. However, systems using explicit solvent have many more atoms (in most case mostly water molecules), demanding additional computing resources. Due to the fraction force from the water molecules, the dynamics of the solute is also slower in the explicit solvent. This may be useful for understanding the real time scale of a biological event. However, for some studies where rate is not a factor, but the results of the motion are important, the implicit solvent model will be more efficient. Another

important advantage of implicit solvent is that it calculates solvent free energy, instead of energy. In the explicit solvent model, water atoms are considered to be part of the system, and solvent-solute interactions are included in the non-bonded calculations. There is no independent solvation term. In implicit solvent, the simulated system only includes solute. The molecular mechanic interactions are calculated first in a vacuum. Then an additional term, the solvation free energy, is calculated and included in the energy of the system. The implicit solvent mimics the average effects of the water molecules, therefore it generates solvation free energies. It is essential for methods such as Molecular Mechanics-Poisson Boltzmann (Generalized Born) Surface Area (MM-PB(GB)SA) approach which needs an explicit solvation free energy term in the calculation. Just a note, MM-PB(GB)SA is a post-processing analysis, which means it is still recommended to generate the structures using explicit solvent first, then using implicit solvent model to analyze the solvation free energy on existing structures.

1.3 Advanced MD simulation methods

As mentioned earlier, due to the rugged potential surface of biomolecules and the limited computing power, advanced sampling methods based on MD simulations are commonly needed to enhance sampling. In this section I have introduced the binding free energy calculation method which I have used in my research.

1.3.1 MM-PB(GB)SA method

MM-PB(GB)SA method is another free energy calculation method (*107*). Unlike other free energy calculation methods like Free Energy Perturbation (FEP) or Thermodynamic Integration (TI) method, this is a post-processing method, which means it calculates the free energy based on existing structure snapshots. In MM-PB(GB)SA calculation, the free energy of the system is divided into three parts: molecular mechanic energy (MM), polar solvation free energy (PB or GB, depending on the method), and non-polar solvation free energy (SA). The MM energy usually consists of electrostatic energy and van der Waals energy of the system in vacuum. In some applications the bond energies, angle energies, and dihedral angle energies are also included in MM energy. The entropy of the system can be included by using normal mode analysis or other methods.

In this method to estimate the free energy of a complex system, one carries out a molecular dynamics simulation, typically in a periodic box with water and counterions, and correct representation of long-range electrostatic effects, saving a set of representative structures. Then one post-processes these structures, removes any solvent and counterion molecules, and calculates the free energy, *G*, according to the following equation:

$$\Delta G_{bind} = \Delta E_{MM} + \Delta G_{PBSA} - TS_{MM} \qquad 1.3$$

where G is the calculated average free energy, and E_{MM} is the average molecular mechanical energy,

$$\Delta E_{MM} = \Delta E_{bond} + \Delta E_{angle} + \Delta E_{tors} + \Delta E_{vdw} + \Delta E_{elec} \qquad 1.4$$

where these correspond to the bond, angle, torsion, van der Waals, and electrostatic terms in the molecular mechanical force field, evaluated with no nonbonded cutoff. G_{PBSA} is the solvation free energy calculated with a numerical solution of the Poisson-Boltzmann equation and an estimate of the nonpolar free energy with a simple surface area term, and $-TS_{MM}$ is the solute entropy, which can be estimated by quasi harmonic analysis of the trajectory or by using normal-mode analysis (*108*). The solvent entropy is included in the solvation free energy term. The free energy due to ionic strength effects can be added with a continuum approach, as described below for nucleic acids (*108*) .

The ability to accurately calculate ΔG, the average free energy you can calculate $\underline{\Delta}$G for a given macromolecular system in various different conformations or structures, adds a very important methodology to our computational arsenal. This has been possible before with free energy perturbation but only for small systems and very limited conformational or topological changes. By using a continuum model, it is implicitly integrating out all

the solvent coordinates and simplifying the problem. Also, by calculating the absolute free energy directly with eq 1.4 between the two "end points" instead of calculating the relative free energy along a mapping coordinate, we are avoiding computations on less interesting intermediate states. Nonetheless, it is clear that applications of eq 1.3 will have intrinsically much larger errors than free energy perturbation/thermodynamic integration calculations. What is surprising is that, despite these larger uncertainties, we can often calculate ΔG in respectable agreement define with experiment (*107*).

1.4 Overview of My Research

1.4.1 Predicting How Structural Modifications Will Impact binding to FabI: Insight Through Molecular Mechanics Poisson Boltzmann Surface Area Calculations

There is a continuous need to develop new drugs to keep pace with emerging drug resistance in clinically important pathogens. The drug discovery process will be greatly aided by inexpensive yet reliable computational methods that predict the binding affinities of ligands for drug targets. In the present work we have used the Molecular Mechanics Poisson Boltzmann Surface Area (MM-PBSA) method to calculate the affinity of a series of TCS analogs for the *E. coli* enoyl reductase FabI (ecFabI). TCS binds with very high affinity to the

FabI•NAD^+ complex, with a K_i value of 7 pM, and the TCS analogues span a 450,000 range of binding affinities. Significantly, a high value of correlation (r^2 = 0.96) is observed between the calculated binding energies and those determined experimentally. Further examination of the components contributing to the free energy of binding and their correlation to the experimental values reveals that the van der Waals energies are the most correlated component of the total affinity (r^2 = 0.74), indicating that the shape of the inhibitor is very important in defining the binding energies for this system. The validation of MM-PBSA for the *E. coli* FabI system serves as a platform for inhibitor design efforts focused on the homologous enzyme in *M. tuberculosis.*

1.4.2 Elucidating the Structure of Acyl Carrier Protein Bound to FabI, the FASII Enoyl Reductase from Escherichia coli using computational modeling and X-ray crystallography

Acyl carrier proteins play a central role in metabolism by transporting substrates in a wide variety of pathways including the biosynthesis of fatty acids and polyketides. However, despite their importance, there is a paucity of direct structural information concerning the interaction of ACPs with enzymes in these pathways. Here we report the structure of an acyl-ACP substrate bound to the *Escherichia coli* fatty acid biosynthesis enoyl reductase enzyme (FabI), based on a combination of X-ray crystallography and molecular dynamics simulation. The structural data are in agreement with kinetic studies on wild-type and mutant

FabIs, and reveal that the complex is primarily stabilized by interactions between acidic residues in the ACP helix α2 and a patch of basic residues adjacent to the FabI substrate-binding loop. Unexpectedly, the acyl-pantetheine thioester carbonyl is not hydrogen-bonded to Y156, a conserved component of the short chain alcohol dehydrogenase/reductase superfamily active site triad. FabI is a proven target for drug discovery and the present structure provides insight into the molecular determinants that regulate the interaction of ACPs with target proteins. Also discussed is whether ACP binds in a similar fashion to the other enoyl reductase enzymes. Keen observation of different crystal structures and crystal contacts led us to the identification of a patch of basic residues in InhA that may play an important role in interaction with the AcpM, the molecule that brings the substrate to the active site of the InhA. This is very important as the residues participating in this key interaction are different than those participating in the similar interaction of ecFabI with ACP, the protein that brings in the substrate in ecFabI. This study has been tested by mutagenesis and fluorescent experiments by Xujie Zhang of the Tonge laboratory.

1.4.3 Understanding the substrate binding loop ordering mechanism in the M.tuberculosis enoyl reductase InhA

A study of the available crystal structures of the enoyl reductases reveals that the substrate binding loop region is disordered in many. This elusive loop is comprised of residues 197-205 in ecFabI and residues 196-211 in *M. tuberculosis* enoyl reductase enzyme, InhA. The “ordering” of this loop could be coupled with the slow onset inhibition of the enzyme. This is evident from TCS binding to ecFabI, diazaborines binding to ecFabI, INH binding to *M. tuberculosis* InhA and PT70 binding to *M. tuberculosis* InhA. All these inhibitors bring about slow onset inhibition of the enzyme. According to our hypothesis, the slow onset step should be related to the substrate binding loop getting ordered. In Chapter 4, the studies on the cause and effects of loop ordering have been discussed. It is also seen from the MD simulations that the conformation of the NAD^+ has an effect on the inhibition of the enzyme. In order to understand this behavior and to see if the ordering is the cause or the effect of high affinity inhibitor binding to the enzyme, the crystal packing effects in some of the crystal structures was investigated. Here, the crystal lattices of the InhA enzyme bound with INH-NAD adduct and PT70 were produced, and the interactions studied. In the crystal lattice of INH-NAD adduct bound to InhA, it is seen that the Glu210 of one monomer is at hydrogen bonding distance to R45 of the neighboring monomer. The crystal packing could have an effect on the loop conformation.

Chapter 2 Predicting How Structural Modifications Will Impact binding to FabI: Insight Through Molecular Mechanics Poisson Boltzmann Surface Area Calculations

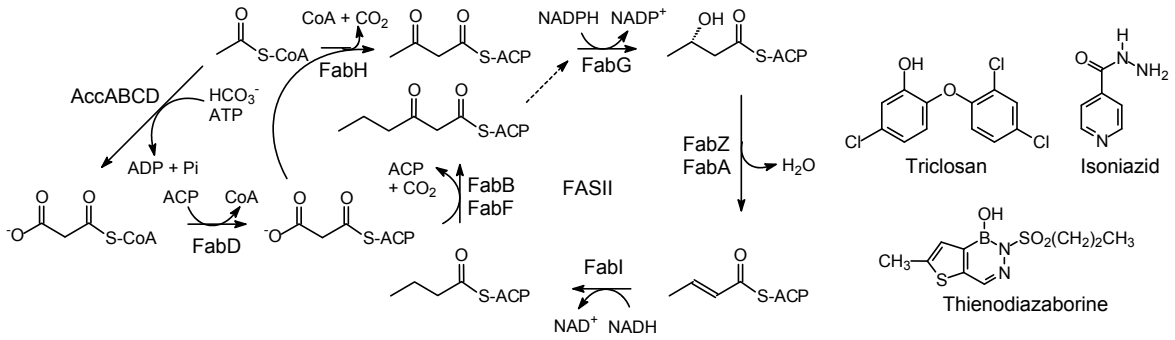
This work has been published and text has been taken from "*Insight through Molecular Mechanics Poisson-Boltzmann Surface Area Calculations into the Binding Affinity of Triclosan and Three Analogues for FabI, the E. coli Enoyl Reductase*" by Salma B. Rafi, Guanglei Cui, Kun Song, Xiaolin Cheng, Peter J. Tonge, and Carlos Simmerling, *Journal of Medicinal Chemistry*, 2006, 49, 4574-4580

2.1 Introduction

There is a compelling need for the development of new therapeutics, especially against unexploited drug targets, that are effective against drug-resistant strains of medically important pathogens. A promising target is the fatty acid synthase (FAS) pathway in bacteria. Fatty acid biosynthesis is a fundamental and vital component of cellular metabolism, and provides the building blocks for the formation of the bacterial cell wall. The bacterial fatty acid synthase system (FAS-II) is very different from that of yeast or the animal fatty

acid synthase system (FAS-I) with respect to the structural organization. The enzymes involved in FAS-II are monofunctional discrete proteins while FAS-I contains all required enzymes in a single polypeptide chain, resulting in a multifunctional unit (*109-111*). Because of its vital role, and the organizational differences with its mammalian counterpart, the FAS-II system has become an attractive target for drug discovery efforts (*112*).

The *E. coli* FAS-II pathway has been extensively studied (*109*) (Scheme 2). Successive rounds of elongation involve the addition of two carbons to the growing fatty acid from malonyl CoA followed by sequential reduction, dehydration and reduction steps. Both the condensing enzymes and terminal reductases have previously been the focus of inhibitor design efforts. Thiolactomycin (*113*) and cerulenein inhibit the condensing enzymes (*114*), while a variety of inhibitors have been developed that target the enoyl reductase FabI (*115*). The diazaborine class of compounds are an early example of FabI inhibitors (*116*), while the anti-TB drug isonaizid (INH) (*43, 117*) inhibits the FabI homologue in *M. tuberculosis* (InhA). More recently it has been shown that the TCS inhibits the FabI enzymes from a variety of organisms (*51*). Initial SAR studies on the interaction of TCS with FabI utilized analogues such as 2,2'-dihydroxy diphenyl ether, 2,2'-dihydroxy diphenyl thioether and 2-hydroxy-3-phenoxybenzaldehyde. These have been studied as potent FabI inhibitors (*52*).



Scheme 2: The Type II Fatty Acid Biosynthesis Pathway and FabI Inhibitors.

TCS is effective against many pathogenic organisms, such as *P. falciparum* (*118*) and *S. aureus (119),* via inhibition of the FabI enzyme. While TCS also inhibits InhA, the enoyl reductase from *M. tuberculosis* (*120*), the K_i for the inhibition is only 0.2 μM. Thus TCS is a promising lead compound for the development of potent InhA inhibitors, and a goal of the present studies is to develop methods that will be useful for rational modification of TCS in order to improve binding affinity for InhA.

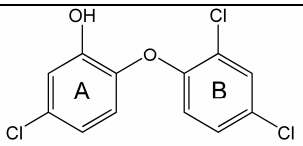
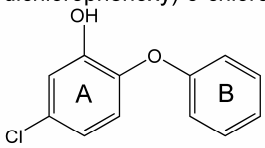
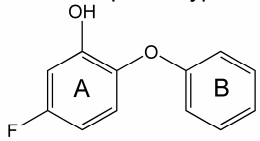
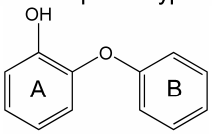
Computation can play an important role in lead compound optimization by predicting structural changes that will improve the affinity of inhibitors for their target enzymes. Several approaches exist, such as free energy perturbation (FEP), thermodynamic integration (TI), Molecular Mechanics-Poisson Boltzmann Surface Area (MM-PBSA) (*108*) and Molecular Mechanics-Generalized Born Surface Area (MM-GBSA) (*121*). The more recent MM-PBSA approach has been well studied (*122, 123*) and used with success on many protein-ligand systems, such as avidin (*124*), HIV reverse transcriptase (*125*), neuraminidase (*126*),

cathepsin D (*127*), Sem-5 (*128*), growth factor receptor binding protein2 (*129*) and matrix metalloproteases (*130*). MM-PBSA is advantageous in that it can be used for many types of intermolecular complexes and that it is “universal” and does not require fitting of additional parameters; however, it is rather time consuming (although much less so than FEP or TI) and in some cases fails to accurately rank ligands. This may arise from approximations inherent in MM-PBSA, which will be discussed in more detail below. In particular, explicit water molecules are employed during molecular dynamics simulations but subsequently replaced with a continuum water model for evaluation of binding affinities.

In order to validate the use of the MM-PBSA method for the development of improved InhA inhibitors, we report its use in calculating relative binding affinities to FabI for a series of TCS analogs. Experimental binding affinities for this TCS analog series have been published previously where it has been shown that small structural changes have a dramatic effect on affinity for FabI(*131*). While removal of the TCS B ring chlorine atoms have only a small effect on binding to FabI, removal of the A ring chlorine results in a 450,000-fold decrease in affinity (Table 3). In addition, replacement of the A ring chlorine with methyl or fluoro substituents also have a dramatic effect on binding. Importantly, we show that the MM-PBSA approach results in excellent reproduction of the experimental data, with a correlation coefficient of 0.98 between computational and experimental relative affinities. Further analysis suggests that van der Waals

energies are largely responsible for the variation in FabI affinity among the inhibitors. This in turn could be because of the changes in the co-factor conformation and we note that the internal hydrogen bond in NAD^+ observed in the FabI/NAD^+/TCS complex is mediated by a water molecule in the other FabI/NAD^+/inhibitor systems.

Table 3: Inhibition Data[a] for TCS and Three Diphenyl Ether TCS Analogs Binding to FabI

Name	Chemical structure	K_1 with FabI
TCS	 2-(2,4-dichlorophenoxy)-5-chlorophenol	7.0 pM
CPP	 5-chloro-2-phenoxyphenol	1.1 pM
FPP	 5-fluoro-2-phenoxyphenol	1.5nM
PP	 2-phenoxyphenol	0.5 μM

[a] K_1 is the dissociation constant of the inhibitor from the enzyme-NAD^+ product complex. Data is taken from (*55*).

2.2 Methods

2.2.1 System Preparation

The crystal structure of the FabI/NAD^+/TCS complex (1QSG)(*132*) was used to build the starting structures for the ternary complex with each of the TCS analogs. Although FabI is a tetramer in solution, a fragment with a single binding site was generated in order to reduce computational cost. The fragment included all residues within 20Å of TCS in the complex, including 246 residues from this monomer and 115 amino acid residues from neighboring monomers. In order to prevent the fragments from drifting away during the MD simulations, in the fragment system, all atoms beyond 15Å from TCS were weakly restrained with a 0.5 kcal/ (mol•$Å^2$) force constant. Three additional systems were built in a similar fashion using the FabI/NAD^+/TCS structure as the template, with the assumption that TCS and the 3 analogs are similar in shape and hence their binding mode to FabI will be similar to that of TCS. Thus, for each system, the TCS was replaced by the analog (CPP, PP or FPP) using the same coordinates for common atoms. Missing atoms were built using the LEAP module of Amber.

Each of the systems was then solvated in a truncated octahedral box of TIP3P water(*97*) with an 8Å buffer between the solute and box edge, resulting in a system with 7262 water molecules and ~27,000 atoms in total. Counterions were not used in any calculations. For each system, a total of 500ps of MD simulation at 300K, with a constant pressure of 1 atm., periodic boundary

conditions and particle mesh Ewald(*105, 133*) treatment of electrostatics were performed with a time step of 1fs. Snapshots were saved every 1ps yielding a total of 500 frames. The first 50 ps of data were regarded as equilibration and not used in the binding affinity analysis.

2.2.2 Force field parameters

Standard Amber ff99 forcefield parameters(*134*) were assigned to the protein. TCS and the analogs were parameterized as follows. The initial geometry of TCS was obtained from the crystal structure of the FabI/NAD^+/TCS complex (PDB ID: 1QSG (*132*)) and the TCS analogs were created through manual modification of the TCS structure. Each structure was optimized using Gaussian 98 (*135*) at HF/6-31G* and partial atomic charges (Tables S1-S4) were derived with standard RESP methodology (*136-138*). In order to obtain the torsion angle parameters for the diphenyl ether linkage (Table S5), a potential energy scan (PES) was performed on these two angles with 36 degree intervals, and each of the resulting 100 geometries was optimized at RHF/6-31G* followed by calculation of single point energies using MP2/6-31G*. These methods and basis set were chosen to be consistent with the procedure used in development of ff99. The QM and MM energy differences were calculated for each conformer, and the torsion angle Fourier series parameters were obtained using multivariate

least square fitting. These parameters were developed on TCS and used for all analogs.

2.2.3 MM-PBSA calculations

MM-PBSA was used to calculate the relative binding free energies of TCS and its analogs to the FabI fragment. The details of this method have been presented elsewhere(*123*). Briefly, the binding affinity for a protein-ligand complex corresponds to the free energy of association in solution as shown in equation 2.1:

$$\Delta G_{bind} = G_{complex} - (G_{unbound\ protein} + G_{free\ ligand}) \quad 2.1$$

while the relative affinities for two ligands can be calculated using equation 2.2:

$$\Delta\Delta G_{bind(1\rightarrow 2)} = \Delta G_{bind(2)} - \Delta G_{bind(1)} \quad 2.2$$

In MM-PBSA, the binding affinity in equation 2.1 is typically calculated using equation 2.3:

$$\Delta G_{\text{bind}} = \Delta E_{\text{MM}} + \Delta G_{\text{solv}} - T\Delta S_{\text{solute}} \quad 2.3$$

where ΔE_{MM} represents the change in molecular mechanics potential energy upon formation of the complex, calculated using all bonded and non-bonded interactions. Solvation free energy, G_{solv}, is composed of the electrostatic component (G_{PB}) and a non-polar component (G_{np}):

$$\Delta G_{solv} = \Delta G_{PB} + \Delta G_{np} \qquad 2.4$$

G_{PB} was calculated using the DelPhi program (*139*) with PARSE radii (*140*). The cubic lattice had a grid spacing of 0.5Å, dielectric constants of 1 and 80 were used for interior and exterior respectively and 1000 linear iterations were performed. The hydrophobic contribution to the solvation free energy, $G_{np,}$ was calculated using the solvent accessible surface area (SASA) (*140*) from the MSMS program (*141*), where γ = 0.00542 kcal/(mol Å^2) and β = 0.92 kcal/mol with a solvent probe radius of 1.4 Å.

$$\Delta G_{np} = \gamma SASA + \beta$$

$T\Delta S_{solute}$ represents the entropic contribution to binding affinity at temperature *T.* The 4 ligands used in these calculations are TCS and three structurally very similar analogs. For a series of compounds with similar structures and binding modes, the entropy contribution can be omitted if one is only interested in relative binding affinities (*125, 142*). Since this calculation converges slowly and can have large uncertainties, we omitted the entropic contribution to $\Delta\Delta$G. We assume that in calculating the relative affinities, it gets

cancelled out. The calculated error bars are standard errors (SE). SE=standard deviation /√ N where N is the number of trajectory snapshots used in the calculations.

2.3 Results and Discussion

Despite spanning a range of 10^6 in dissociation constants, the diphenyl ether FabI inhibitors in Table 3 are structurally similar. All have a hydroxyl group on the A ring. TCS has a chlorine meta to the hydroxyl group on the A ring and two chlorine atoms on the B ring. Replacement of the B ring chlorines with hydrogens results in CPP, which, experimentally, is a 7 fold better inhibitor of FabI than TCS (*131*) (Table 3). In contrast, replacement of the A ring chlorine in CPP with fluorine (FPP) results in a dramatic change in binding affinity, and experimentally FPP binds 1,300 fold less tightly to FabI than CPP. Similarly, replacement of the fluorine with a hydrogen to give PP results in another large decrease in binding affinity and, overall, PP binds 450,000 fold less tightly to FabI than CPP.

An important validation of the MM-PBSA approach to future ligand design efforts is to reproduce the experimental changes in binding affinity that occurs as the TCS skeleton is modified. In particular, the method should reproduce the high sensitivity to the substituent at the meta position on the A ring and the relative insensitivity to removal of the chlorine atoms on the B ring.

2.3.1 Molecular Dynamics Simulations

During the course of the MD simulations, the protein fragment was quite stable (<1Å RMSD, data not shown). This and all other structural comparisons and RMSD values are reported with respect to the crystal structure of the FabI/NAD^+/TCS ternary complex (PDB ID: 1QSG (*132*)). The low RMSD likely reflects the use of positional restraints on atoms farther than 15Å from the inhibitor. Neither the ligands nor the NAD^+ were restrained in any of the simulations. The RMSD of the inhibitors remained low (<2Å, data not shown) and did not vary significantly between the different analogs. Interestingly, the RMSD values for the NAD^+ in the ternary complex did vary between the systems; the RMSD is in the range of 0.5 – 1.0 Å with TCS, but is notably higher (1 – 2 Å) for the other analogs (Figure 15). Since the RMSD is low for the TCS inhibitor that was used to obtain the crystal structure, it seems unlikely that the higher RMSD values that we observe for the other analogs arise solely from the simulation protocol or inaccuracies in the potential function. Instead, it is possible that the conformation of the NAD^+ indeed differs when the B ring chlorine atoms are removed. This may have implications for the binding affinity of the ligands and is discussed in more detail below.

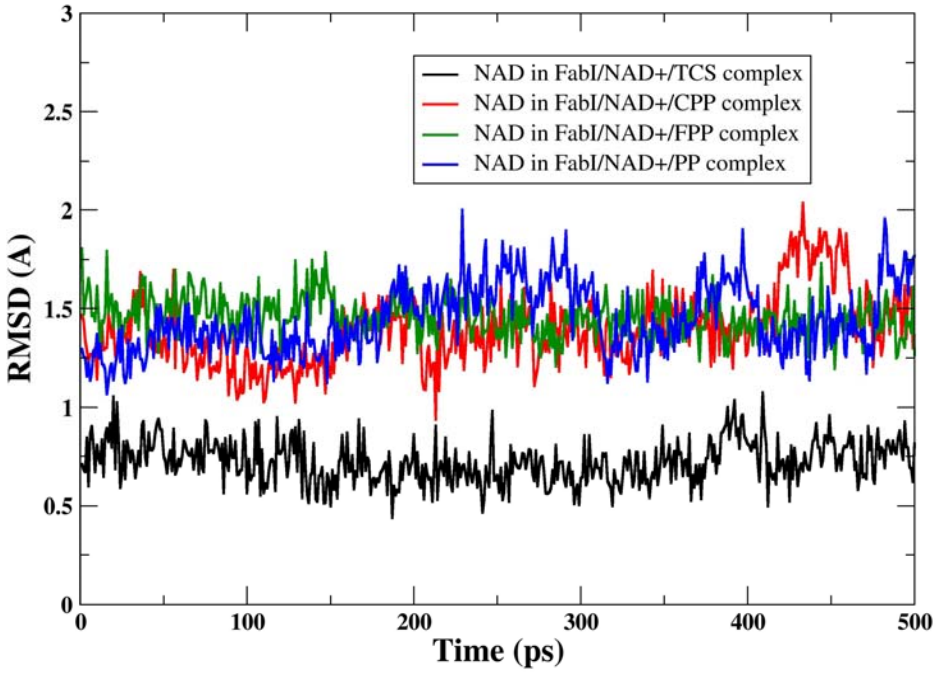


Figure 15: RMSD vs Time plot for NAD^+ in the FabI/NAD^+/inhibitor simulations. RMSD values were calculated between the position of NAD^+ during the simulation and the position of NAD^+ in the FabI/NAD^+/TCS crystal structure. For the FabI/NAD+/TCS system, the RMSD of NAD^+ (black) is notably lower (<1 Å) than those for the FabI/NAD^+/CPP (red), FabI/NAD^+/FPP(green) and FabI/NAD^+/PP(blue) systems.

In order to obtain an atomic-detail view of the changes that are reflected in these RMSD values, cluster analysis was performed on the 4 trajectories. In Figure 16, the representative structures for the inhibitor, NAD^+ and selected active site residues are compared to the experimentally determined structure of the FabI/NAD^+/TCS ternary complex. Consistent with the low RMSD values shown above, the structure of the complex obtained with TCS reproduces the crystallographic data (Figure 16), with an RMSD of ~0.75 Å for NAD^+ and ~1Å

for TCS, again suggesting that the simulations can adequately reproduce the experimental structure data for this inhibitor. In contrast, the complexes with the other 3 inhibitors show more significant differences from the TCS complex (Figure 16 B, C, D) with ~1.5Å RMSD for NAD^+ and ~2 Å for TCS. The close match between the simulated and experimental structures obtained with TCS and the difference between the simulated structures for TCS and the other analogs again suggests that the B ring chlorines play a role in the details of the NAD^+ conformation.

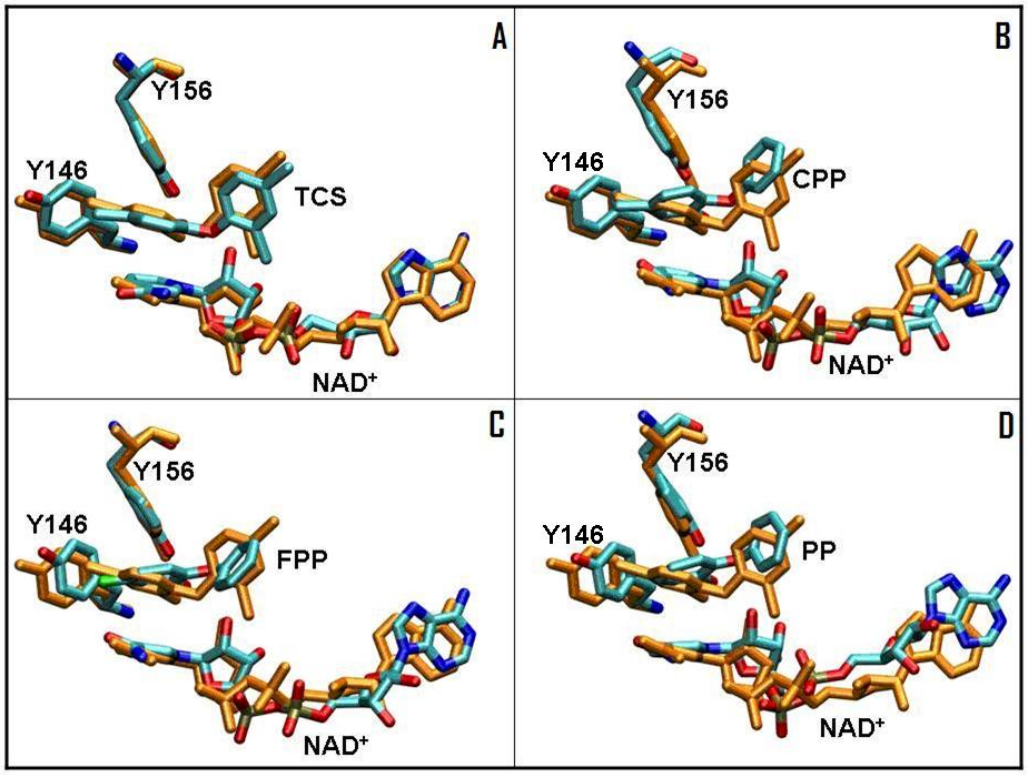


Figure 16: Comparison of the Theoretical and Experimental Active Site Structures for the Enzyme-Inhibitor Systems. The active sites of the four FabI/NAD^+/inhibitor systems from simulations (cyan/blue/red). A: FabI/NAD^+/TCS; B: FabI/NAD^+/CPP C: FabI/NAD^+/FPP; D: FabI/NAD^+/PP. For comparison, the crystal structure of the FabI/NAD^+/TCS complex is shown in orange. For clarity, only heavy atoms are shown for the simulated systems.

A comparison of the crystal structure of FabI/NAD^+/TCS complex to representative structures obtained by cluster analysis of the four simulations in explicit solvent reveals that the intramolecular hydrogen bond between the phosphate and the NH of the carboxamide group in the NAD^+ of the FabI/NAD^+/TCS system (Figure 17) is replaced by one bridging water molecule in the CPP, FPP and PP systems (Figure 17B, C, D).

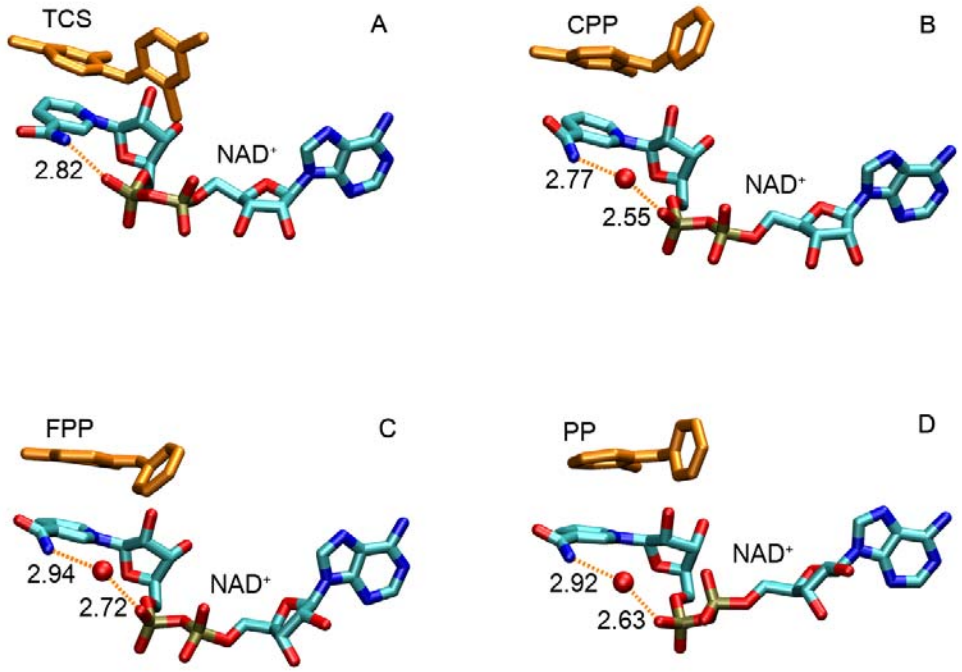


Figure 17: The Bridging Water Molecule in the NAD^+ of the CPP, FPP and PP Complexes. Calculated structures showing the presence of a bridging water molecule (red atom) between the NAD^+ carboxamide and phosphate groups in the CPP, FPP and PP complexes. The inhibitors are shown in orange. A: FabI/NAD^+/TCS. B: FabI/NAD^+/CPP. C: FabI/NAD^+/FPP. D: FabI/NAD^+/PP. Consistent with the crystal structure, an intramolecular hydrogen bond is observed between the NAD^+ carboxamide and phosphate groups in the TCS complex. When the inhibitor B ring chlorine atoms are removed (B, C and D), the intramolecular hydrogen bond is replaced by a water molecule.

2.3.2 Comparison of calculated and experimental binding affinities

The stability of the structures in the simulations and the close reproduction of the experimental binding mode for TCS suggest that it is reasonable to use the simulation data for further analysis. The binding free energies calculated using the MM-PBSA approach are compared with the experimental free energies of binding (*131*) (absolute affinities, Table 4 ; relative affinities for all inhibitor pairs, Table 5). The MM-PBSA data reasonably reproduce the experimental absolute binding free energies, with an overall correlation coefficient of 0.85. The error for TCS is <2 kcal/mol, and is somewhat larger (3 – 4 kcal/mol) for the other inhibitors (Table 5). These differences in absolute affinities are quite reasonable given the approximations that were made in the simulations; the entropic component was neglected and conformational change in the ligand and/or receptor upon binding were also neglected due to our use of a single trajectory for the bound and free states. It is likely that these effects are similar for each inhibitor and thus will largely cancel in the relative affinities. It is also possible that the removal of all of the water in the MM-PBSA calculation resulted in the increased error for CPP, FPP and PP, each of which has a structured water molecule forming a hydrogen bonding bridge in the NAD^+ cofactor (Figure 17).

Table 4: Individual Energy Components for the Calculated Absolute Binding Free Energies.

System	ΔE_{vdw} (N=450) A	ΔE_{coul} (N=450) B	ΔG_{polar} (N=450) C	$\Delta G_{non\text{-}polar}$ (N=450) D	$\Delta E_{electro}$ = B+C+D	$\Delta G_{mm\text{-}pbsa}$ A+ B+C+D	ΔG_{expt} = RT ln(K_1)
TCS	-37.28 ± 0.11	-11.70 ± 0.10	39.15 ± 0.15	-4.36 ± 0.01	23.09 ± 0.15	-14.19 ± 0.14	-15.46
CPP	-31.91 ± 0.12	-11.43 ± 0.09	35.11 ± 0.16	-3.98 ± 0.01	19.70 ± 0.15	-12.20 ± 0.13	-16.56
FPP	-29.30 ± 0.11	-9.73 ± 0.12	35.97 ± 0.30	-3.97 ± 0.01	22.27 ± 0.28	-7.03 ± 0.25	-11.76
PP	-27.10 ± 0.15	-10.92 ± 0.11	35.98 ± 0.37	-3.74 ± 0.01	21.32 ± 0.38	-5.78 ± 0.28	-8.73
	r^2 = 0.63	*r^2 = 0.31*	*r^2 = 0.05*	*r^2 = 0.50*	*r^2 = 0.03*	*r^2 = 0.85*	

The calculated absolute affinities indicate decreasing affinity for FabI in the order TCS, CPP, FPP and PP. This rank order matches the experimental trends with the exception that the order of TCS and CPP are reversed, with CPP binding more tightly than TCS in the experimental data. This may again reflect error in MM-PBSA energies introduced by removal of the structured water in CPP that was not present in the TCS complex. However, the simulation is in agreement with experiment in that the effect arising from removal of both B ring chlorines (TCS→CPP) is much smaller than the effects arising from removal of the A ring chlorine among the other inhibitors. For example, CPP and TCS differ by 2 kcal/mol in simulations and 1.1 kcal/mol in experiment, while CPP and PP differ by 6.4 kcal/mol in simulation and 7.7 kcal/mol in experiment.

Table 4 also shows the individual energy components contributing to the total calculated absolute binding free energies. It is interesting to note that the correlation with the total ΔG_{expt} is nearly zero for the ΔG_{polar} term, indicating that

differences in desolvation of the inhibitor and enzyme upon binding are not directly responsible for the large variation in binding affinities of these analogs. Likewise, the correlation with the Coulomb electrostatic energies is quite poor (r^2=0.3). Although the correlation with the $\Delta G_{nonpolar}$ (SASA) term is better (r^2=0.5), this term varies by only a few tenths of kcal/mol between the inhibitors.

Adding together the Coulomb and solvation terms does not improve the correlation (r^2=0.03), although this sum is large and positive for all of the inhibitors (20 to 25 kcal/mol), suggesting that the favorable electrostatic interactions between inhibitor and enzyme (-10 to -15 kcal/mol) are insufficient to completely overcome the large desolvation penalties (35 to 40 kcal/mol).

The van der Waals component shows the highest correlation with experimental affinities (r^2=0.6), consistent with previous MM-PBSA studies that also showed high correlation between van der Waals components and experimental binding affinities (*128, 143*). In spite of the low level of correlation for the Coulomb/solvation sum, adding it to the van der Waals term improves the quality of the fit, with an overall correlation coefficient of 0.85 for the total binding free energies. Overall, this suggests that the affinities of these inhibitors for FabI are dominated by shape complementarity, but the total affinity arises from a more complex interplay between all of these components.

These effects are more apparent when one examines the relative binding affinities between all six pairs of inhibitors (Figure 18). As expected, the correlation in the relative affinities is much improved, likely due to substantial cancellation of the errors arising from neglect of entropy and conformational changes in the enzyme in the absolute affinities. As shown in Figure 18 and Table 5, the relative binding affinities of the analogs show an impressive correlation coefficient of 0.98 between calculation and experiment. In addition, the slope of the best-fit line (obtained through linear regression) is 1.01, indicating that the data is not only well correlated but also that the differences in affinities are reproduced nearly quantitatively by the simulations.

As noted above, the sign of $\Delta\Delta G$ is incorrect for the TCS-CPP pair; TCS appears to bind 2 to 3 kcal/mol too strongly in all of the simulations as compared to the analogs without the B ring chlorine atoms. A possible explanation for this discrepancy is the removal of the structured water molecule that forms the hydrogen bond bridge in the NAD^+ of the CPP, FPP and PP complexes (Figure 17). It is difficult to define the position of this water molecule in the unbound state as this water was not present in the TCS bound FabI fragment system, and was thus not included in the calculations.

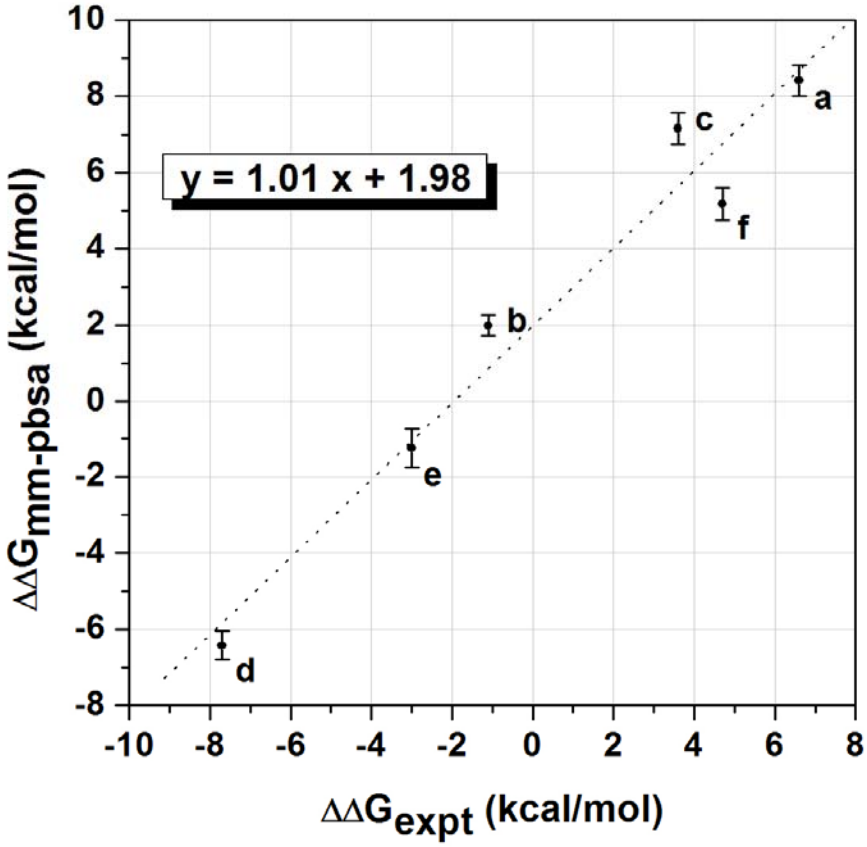


Figure 18: Correlation Between the Relative Experimental and Theoretical Binding Free Energies. A high correlation and nearly linear slope is observed between relative binding free energies obtained experimentally and the binding energies calculated using the MM-PBSA approach. a: PP-TCS, b: CPP-TCS, c: FPP-TCS, d: CPP-PP, e: FPP-PP, f: FPP-CPP

Table 5: Individual Energy Components for the Calculated Relative Binding Free Energies.

System	$\Delta\Delta E_{vdw}$ (N=450) A	$\Delta\Delta E_{coul}$ (N=450) B	$\Delta\Delta G_{polar}$ (N=450) C	$\Delta\Delta G_{non\text{-}polar}$ (N=450) D	$\Delta\Delta E_{electro}$ = B+C+D	$\Delta\Delta G_{mm\text{-}pbsa}$ = A+ B+C+D	$\Delta\Delta G_{expt}$ = RT $\ln(K_1)$
a: PP-TCS	10.17 ± 0.17	0.78 ± 0.15	-3.16 ± 0.34	0.62 ± 0.01	-1.76 ± 0.37	8.41 ± 0.41	6.6
b: CPP-TCS	5.36 ± 0.15	0.27 ± 0.13	-4.03 ± 0.18	0.36 ± 0.01	-3.39 ± 0.19	1.99 ± 0.27	-1.1
c: FPP-TCS	7.98 ± 0.16	1.97 ± 0.15	-3.18 ± 0.36	0.39 ± 0.01	-0.82 ± 0.35	7.16 ± 0.42	3.6
d: CPP-PP	-4.80 ± 0.14	-0.51 ± 0.15	-0.87 ± 0.32	-0.25 ± 0.01	-1.63 ± 0.36	-6.42 ± 0.38	-7.7
e: FPP-PP	-2.20 ± 0.17	1.19 ± 0.17	-0.01 ± 0.44	-0.23 ± 0.01	0.95 ± 0.47	-1.25 ± 0.50	-3
f: FPP-CPP	2.60 ± 0.14	1.70 ± 0.15	0.86 ± 0.38	0.01 ± 0.01	2.56 ± 0.34	5.17 ± 0.43	4.7
	$r^2 = 0.74$	$r^2 = 0.49$	$r^2 = 0.047$	$r^2 = 0.56$	$r^2 = 0.051$	$r^2 = 0.96$	

Based on the high correlation of the relative binding affinities with the experimental data, we further investigated which energy components were most correlated with the experimental relative affinities. All components contributing to each data point in **Figure 18** are shown in **Table 5**. Similar to the absolute affinities, the van der Waals energies are well correlated ($r^2 = 0.74$) with the experimental values (which, of course, reflect all interactions). In addition, the van der Waals terms are the largest component of the *relative* affinity for each of the pairs. This indicates that the shape of the inhibitor plays an important role in defining the relative affinities of these inhibitors for FabI, much as the large magnitude of the van der Waals term in the total affinities suggested that this term was also the most important for defining the absolute affinity of each of the inhibitors. One might imagine that the interactions that give rise to strong binding

need not be the same as those that discriminate among the inhibitors, but in the present case these arise from the same type of interaction. The $\Delta G_{nonpolar}$ term contributes less than 1 kcal/mol to each of the relative affinities, consistent with the small differences in solvent accessible surface area among the inhibitors. The sum of Coulomb and solvation terms shows little correlation, also consistent with our observations for the total affinities, but once again the addition of this term to the van der Waals component improves the quality of the fit (with r^2 increasing from 0.74 to 0.96 upon inclusion of electrostatic and desolvation effects).

Due to the importance of the van der Waals interactions not only in the absolute affinities but also in defining the relative binding of the analogs, we performed a further decomposition of this component into terms involving each amino acid in the enzyme. Our goal was to gain insight into specific active site residues that influence the relative affinities, with potential application in the design of improved inhibitors that optimize these key interactions. We considered separately the effects of changing substituents on the A ring (PP vs. CPP) and B ring (TCS vs. CPP). The residues contributing significantly to the difference in binding are shown in Figure 19 with energy values in Table 6.

With respect to the difference in binding of CPP and TCS, van der Waals interactions with the NAD^+ cofactor contribute ~3 of the total 5.3 kcal/mol. This is reasonable considering that the chlorine at the ortho position on the B ring is in

direct contact with the NAD^+ phosphates in both the crystal structure and the simulations of the FabI/NAD^+/TCS ternary complex. This contact may also play a role in stabilizing the observed position of the NAD^+ in the binding pocket, since all simulations with inhibitors that lacked B ring chlorines showed a change in the NAD^+ conformation (**Figure 15**). The other residues contributing significantly (greater than 0.5 kcal/mol) to the van der Waals binding energy difference for B ring chlorines (TCS-CPP) are Gly93, Phe94, Ala196 and Ala197, residues that are all close to the B ring. Important residues for removal of the A ring chlorine are Tyr146, Pro191, Ala196, Ala197 and Ile200. Ala196 and Ala197 are in the substrate binding loop. Residues Gly93, Phe94, Tyr146, Pro191 are a part of the active site pocket, with Tyr146 and Pro191 in close proximity to the A ring (Figure 19).

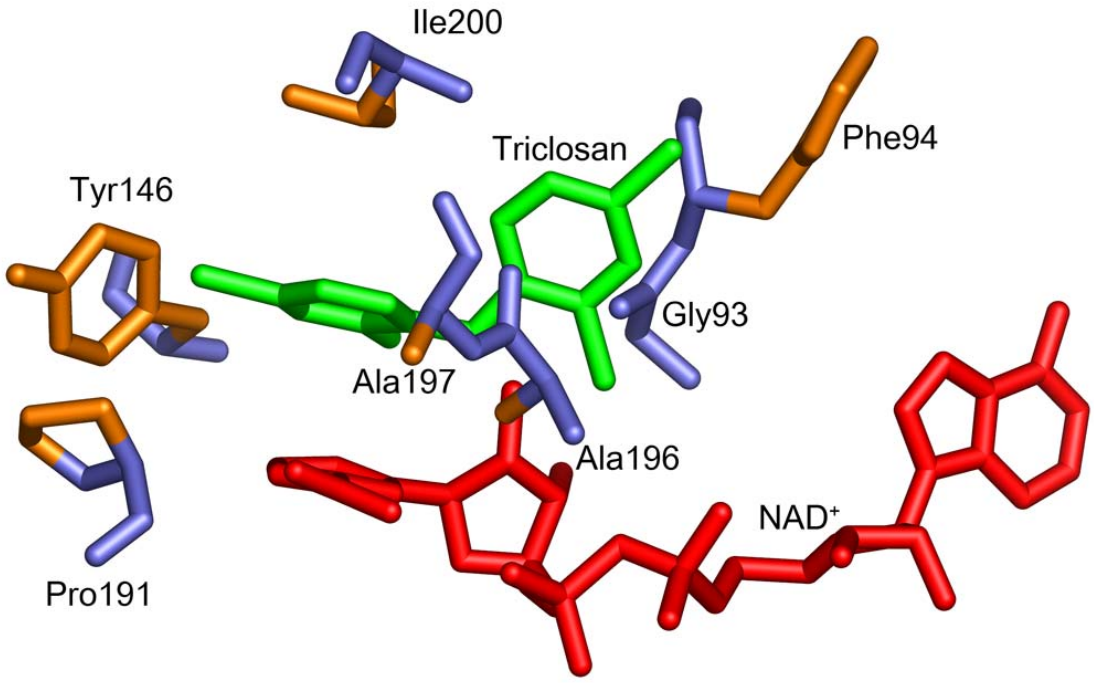


Figure 19: The Active Site Cavity of FabI Showing the Position of TCS, NAD^+ and Selected Active Site Residues. NAD^+, red; TCS, green; backbone, blue; side chains, orange.

Table 6: Residues in FabI that are Sensitive to Removal of the TCS Chlorine Atoms. Residues that contribute more than 0.5 kcal/mol in van der Waals energies for removal of A ring (TCS vs. CPP) and B ring (CPP vs. PP) chlorine atoms. Residues are shown in Figure 19.

$\Delta\Delta E_{vdw}$(TCS-CPP) (kcal/mol)		**$\Delta\Delta E_{vdw}$(CPP-PP) (kcal/mol)**	
Gly 93	-0.81	Tyr 146	-0.61
Phe 94	-0.81	Pro 191	-0.61
Ala 196	-1.21	Ala 196	0.85
Ala 197	1.01	Ala 197	-1.74
NAD+	-2.97	Ile 200	-0.52

2.4 Conclusions

We applied the computationally inexpensive MM-PBSA method to calculate the relative binding affinities of a series of inhibitors to FabI, the *E. coli* enoyl reductase enzyme. The rank ordering of the calculated ligand affinities was correct with the exception of TCS and CPP, the pair that represents the smallest difference in affinity in both simulation and experiment. TCS was found to bind somewhat too strongly relative to all of the other inhibitors (as compared to experiment), possibly due to a difference in water structuring between TCS and the other inhibitors that we observed in simulations. Incorporation of these effects into an implicit solvent model such as that used with MM-PBSA is non-trivial. This water molecule was not present in the FabI•Nad^+•TCS crystal structure and our MD simulations derived FabI•Nad^+•TCS structure. Hence, it is difficult to define the position of this water molecule in the unbound state of the protein.

Overall, the binding free energy data obtained from MM-PBSA were in excellent quantitative agreement with experimental K_1 data, with a correlation coefficient of 0.98 and a slope of 1.01. Particularly notable is that the calculations also reproduce the high sensitivity to removal of the chlorine atom on the TCS A ring as well as the relative insensitivity to removal of the B ring chlorines. This high level of agreement validates the data and suggests that it is reasonable to

further examine the specific interactions and energy components that influence binding. Energy decomposition analysis was performed to study the contribution of different interaction types towards the relative affinities. This showed that the van der Waals energies were most highly correlated to experimental data for both the absolute and relative affinities. In addition, the Coulomb electrostatic interaction between protein and inhibitor was observed to be insufficient to overcome the unfavorable desolvation energy. Adding together the Coulomb and solvation terms does not improve the correlation (r^2=0.03), although this sum is large and positive for all of the inhibitors (20 to 25 kcal/mol), suggesting a possible route to further optimization of affinity. Together, these imply that the shape of the ligand plays an important role in determining its binding affinity. This information will be of critical importance for designing more potent inhibitors of the homologous enzyme in *M. tuberculosis.*

Supporting Information: The force field parameters for TCS and analogs are given in Table 7, Table 8, Table 9, Table 10 and Table 11

Table 7: TCS residue AMBER Atom Types and Partial Charges

Atom Name	AMBER atom type	Partial charge
C1	CA	0.0924
C2	CA	0.2056
C3	CA	-0.0559
C4	CA	-0.0729
C5	CA	-0.1186
C6	CA	-0.1412
O1	OS	-0.2648
O2	OH	-0.5688
C7	CA	0.1204
C8	CA	0.0087
C9	CA	-0.0300
C10	CA	-0.0436
C11	CA	-0.1176
C12	CA	-0.0956
CL2	Cl	-0.0883
CL3	Cl	-0.0900
CL1	Cl	-0.0908
H4	HA	0.1559
H5	HA	0.1548
H6	HA	0.1629
H7	HO	0.4354
H8	HA	0.1419
H9	HA	0.1708
H10	HA	0.1292

Table 8: CPP residue AMBER Atom Types and Partial Charges

Atom Name	AMBER atom type	Partial charge
C1	CA	0.1022
C2	CA	0.1845
C3	CA	-0.0690
C4	CA	-0.0485
C5	CA	-0.1189
C6	CA	-0.1638
O1	OS	-0.2753
O2	OH	-0.5549
C7	CA	0.1152
C8	CA	-0.0629
C9	CA	-0.2176
C10	CA	-0.0982
C11	CA	-0.2174
C12	CA	-0.0533
H1	HA	0.1090
H2	HA	0.1302
CL	Cl	-0.1034
H4	HA	0.1573
H5	HA	0.1513
H6	HA	0.1653
H7	HO	0.4196
H8	HA	0.1698
H9	HA	0.1698
H10	HA	0.1090

Table 9: FPP residue AMBER Atom Types and Partial Charges

Atom Name	AMBER atom type	Partial charge
C1	CA	0.1082
C2	CA	0.1783
C3	CA	-0.1160
C4	CA	0.1204
C5	CA	-0.1712
C6	CA	-0.2049
O1	OS	-0.2748
O2	OH	-0.5531
C7	CA	0.1079
C8	CA	-0.0571
C9	CA	-0.2220
C10	CA	-0.0822
C11	CA	-0.2219
C12	CA	-0.0490
H1	HA	0.1065
H2	HA	0.1235
F	F	-0.1622
H4	HA	0.1617
H5	HA	0.1641
H6	HA	0.1808
H7	HO	0.4206
H8	HA	0.1679
H9	HA	0.1679
H10	HA	0.1065

Table 10: PP residue AMBER Atom Types and Partial Charges

Atom Name	AMBER atom type	Partial charge
C1	CA	0.0804
C2	CA	0.2802
C3	CA	-0.1723
C4	CA	-0.1897
C5	CA	-0.2006
C6	CA	-0.1505
O1	OS	-0.2720
O2	OH	-0.5753
C7	CA	0.1271
C8	CA	-0.0904
C9	CA	-0.2016
C10	CA	-0.1077
C11	CA	-0.2201
C12	CA	-0.0549
H1	HA	0.1185
H2	HA	0.1337
H3	HA	0.1619
H4	HA	0.1580
H5	HA	0.1570
H6	HA	0.1619
H7	HO	0.4236
H8	HA	0.1571
H9	HA	0.1571
H10	HA	0.1185

Table 11: Amber dihedral angle parameters for the biphenyl ether linkage. Un and Vn are the amplitudes (force constant) for two torsions of n (n=1,2,3 in this case) periodicity respectively. The equi values are the corresponding phase angles

U: C8-C7-O-C1, V: C2-C1-O-C7

	amplitude	periodicity	phase
U1	0.7329	1	-121.8378
U2	1.0994	2	164.7281
U3	0.3439	3	57.2180
V1	5.6737	1	-110.9929
V2	4.6134	2	178.7944
V3	1.6328	3	-43.2034

Chapter 3 Elucidating the Structure of Acyl Carrier Protein Bound to FabI, the FASII Enoyl Reductase from Escherichia Coli using computational modeling and X-ray crystallography

This work has been published and most of the text has been taken from the "*Structure of Acyl Carrier Protein Bound to FabI, the FASII Enoyl Reductase from Escherichia coli"* by Salma Rafi, Polina Novichenok, Subramaniapillai Kolappan, Xujie Zhang, Christopher F. Stratton, Richa Rawat, Caroline Kisker, Carlos Simmerling, and Peter J. Tonge; *Journal of Biological Chemistry, Dec 22, 2006, Volume 281, Number 51*

3.1 Introduction

Acyl carrier proteins (ACPs) play an essential role in a diverse array of metabolic pathways including the biosynthesis of fatty acids (*144, 145*), polyketides (*146*), membrane-derived oligosaccharides (*147*), lipopolysaccharides (*148, 149*) and phospholipids (*150*). In each case the growing substrate is attached via a thioester to the ACP phosphopantetheine group. ACPs must therefore be able to recognize and interact, in an acyl group-

dependent manner, with a wide variety of enzymes. In eukaryotic type I fatty acid synthesis (FASI) and in polyketide biosynthesis, the ACP occurs as part of a larger polypeptide that is also associated with other catalytic activities. In contrast, in bacterial type II fatty acid biosynthesis (FASII), each of the enzyme activities as well as the ACP are encoded by separate polypeptide chains (*145*). ACPs that function in FASII-mediated biosynthesis are small, highly soluble, acidic proteins that vary in molecular weight from 7.5 kDa (*E. coli*) to 13 kDa (*M. tuberculosis*) (*144, 151-154*).

Despite the central role that ACPs play in metabolism, structural details of their interaction with target proteins are sparse. While the structures of ACPs from a variety of different species have been determined by X-ray crystallography (*75*) and NMR spectroscopy (see for example (*155, 156*)), only one structure has been determined of ACP in complex with another protein, the holo ACP synthase (AcpS) (*157*), and no structural information is available for the interaction between ACP and enzymes of the fatty acid biosynthesis pathway. AcpS attaches the phosphopantetheine to the ACP serine and thus, while valuable, the complex of AcpS and ACP differs fundamentally from other ACP-protein complexes and does not provide insight into the delivery of substrate by ACP.

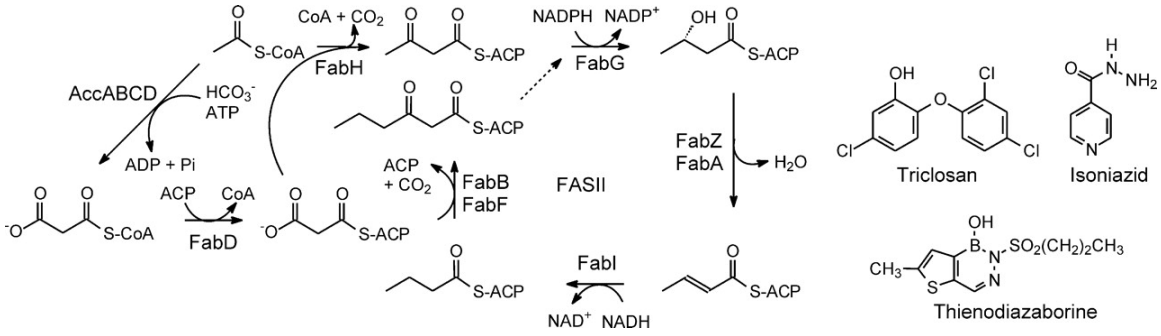
The NMR studies reveal that ACPs are highly flexible, a structural feature that is thought to be important for transporting the growing acyl chain and for protein recognition. From our MD simulations it can be seen that the residue

rmsd for the ACP molecule is higher (~3 Å) (data not shown). In addition, based on X-ray crystallographic analysis of butyryl-ACP, it is thought that ACP adopts two major conformations, optimized for either delivering the acyl chain to the target enzyme or for transporting the growing acyl chain between enzymes (*75*). The flexibility of ACP, coupled with the relatively low affinity of ACP for target proteins, has hindered crystallographic efforts to obtain direct structural data on the interaction of ACP with any of the target proteins (*158, 159*).

Using docking and site directed mutagenesis, Rock and coworkers have rigorously analyzed the interaction of ACP with two FASII enzymes. In agreement with the X-ray structure of ACP bound to AcpS (*157*), these studies confirmed the importance of the ACP recognition helix (helix α2) in FASII protein binding and identified a patch of basic residues on the *E. coli* β-ketoacyl-ACP synthase (FabH) and β-ketoacyl-ACP reductase (FabG) enzymes responsible for interaction with the ACP helix α2 (*158, 159*).

The efforts in my research are focused on the FASII enoyl reductase enzyme FabI, a target for antibacterial diazaborine compounds (*116*) and TCS (*51, 120, 160-162*), while the anti-TB drug INH inhibits InhA, the FabI homolog from *M. tuberculosis (117, 163, 164)* (**Scheme 3**). Modeling studies, based on the information provided by the interaction of ACP with beta-ketoacyl-ACP synthase III, FabH and beta-ketoacyl-ACP reductase, FabG (*158, 159*), suggest that ACP should interact with a cluster of basic residues adjacent to the FabI

substrate binding loop. This loop is disordered in binary FabI-cofactor complexes, but becomes ordered in the ternary FabI:NAD^+:TCS complex. The ordered loop provides two entries into the active site, termed the major and minor portals. Sacchettini and coworkers have determined the structure of a C16 substrate bound to InhA and have discussed the role of the major and minor portals in substrate recognition by the FabI enzymes (*165*). The orientation of the C16 substrate in InhA suggests that substrates enter the FabI active site through the major portal.



Scheme 3: The type II FAS pathway and FabI inhibitors

In the current study we have determined the structure of ACP bound to the *E. coli* FabI enzyme. X-ray crystallography was performed by collaborator Dr. Kisker. X-ray crystallographic data obtained from a complex between FabI and dodecenoyl-ACP revealed most of the main chain electron density for both FabI and ACP. However, the observed relative orientation of ACP and FabI leaves the ACP S36 residue too far from the active site to deliver the substrate through the major portal.

Since some aspects of the structure were not resolved by the crystallographic data, we employed computational methods to model the missing details and importantly to ascertain whether the ACP could deliver substrate into the FabI active site in the observed complex. Through MD simulations, we generated a model for a productive complex between ACP and FabI. The details of the interaction between FabI and ACP in the resulting model are supported by mutagenesis studies, and provide the first detailed description of ACP recognition by a FASII enzyme. Intriguingly, the structural data indicate that the substrate enters the FabI active site through the minor portal and furthermore suggest that the substrate thioester carbonyl group does not form a hydrogen bond with Y156, a conserved active site residue.

3.2 Materials and Methods

The preparation of substrates and enzymes and the kinetic assays were performed by the Tonge laboratory personnel. X-ray crystallography was performed by Kisker lab personnel. Crystallographic statistics are given in Table 12. The coordinates and structure factors have been submitted to the RCSB (PDB ID 2FHS).

Table 12: Data collection and refinement statistics: [a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i - \langle I \rangle|/\Sigma_{hkl}\Sigma_i \langle I \rangle$, where I_i is the ith measurement and $\langle I \rangle$ is the weighted mean of all measurements of I. $\langle I/\sigma I \rangle$ indicates the average of the intensity divided by its average standard deviation. Numbers in parentheses refer to the respective highest resolution data shell in each data set. $R_{cryst} = \Sigma||F_o| - |F_c||/\Sigma|F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes. R_{free} same as R_{cryst} for 5% of the data randomly omitted from the refinement. Ramachandran statistics indicate the fraction of residues in the most favored, additionally allowed, generously allowed, and disallowed regions of the Ramachandran diagram, as defined by the program PROCHECK (41)

Data collection	
Space group	$P6_522$
Cell dimensions	
a, b, c (Å)	127.7, 127.7, 206.7
Resolution (Å)	50.0-2.7
R_{merge} **(%)**[a]	8.1 (71.2)
Mean $\langle I/\sigma I \rangle$	25.4 (2.7)
Completeness (%)	97.1 (98.4)
Multiplicity	10.1 (7.3)
Refinement	
Resolution range (Å)	30.0-2.7
Unique reflections	25,906
Ramachandran statistics	87.1/11.7/1.2/0.0
R_{cryst}	0.226
R_{free}	0.263
Number of nonhydrogen atoms (protein/solvent)	3947/79
r.m.s. deviations from ideal values	
Bond length (Å)	0.010
Bond angle (°)	1.235

3.2.1 Molecular Dynamics Simulations

Computational modeling and molecular dynamics simulations were performed with the Amber suite of programs (*166*). The missing ACP atoms were added using the butyryl-ACP crystal structure (pdb code 1L0I) (*75*) by overlapping the two structures, and replacing the partial coordinate set with the coordinates from 1LOI. Similarly, the structure of TCS bound to FabI in the presence of NAD^+ (pdb code 1QSG) (*161*) was used to place the NADH cofactor. A comparison between the TCS bound FabI crystal structure (PDB ID: 1QSG) and the current crystal structure was made. It was seen that the current crystal structure was missing the density for some of the side chains and the corresponding residues were all termed “alanines” due to the lack of the side chains. The correct amino acids were assigned using the TCS bound FabI crystal structure (PDB ID: 1QSG). Missing FabI side chains and all hydrogens were added using Xleap The coordinates for the phosphopantetheine moiety were taken from the crystal structure of holo-acyl carrier protein-synthase in complex with holo-acyl carrier protein (1F80) (*157*). Maestro Molecular Modeling software was used to build the acyl chain for the phosphopantetheine moiety (PP). Force field parameters were the ff99 set for proteins (*167*) and published parameters (*168, 169*) for NADH. The Amber antechamber module and GAFF force field (*170*) with am1bcc charges (*171*) were used to generate the parameters for PP and the attached acyl chain. In order to present the correct face of the crotonyl double bond to the NADH, the substrate must be bound in an *s-trans*

conformation such that the *si* face of Cβ is oriented toward the NADH pro4(*S*) proton so that hydride transfer will generate the expected 3(*S*)-enoyl product (*172*). Since spontaneous isomerization of this double bond is unlikely to occur during relatively short MD simulations, we modeled the crotonyl in the *s-trans* conformation.

Initial minimization of the FabI:NADH:ACP ternary complex was performed in a stepwise fashion by restraining the backbone Cα atoms and allowing the side chains to move, with each step consisting of 1000 cycles. The restraints were gradually removed in each step (force constants from 10, 7, 4, 1 and 0 kcal mol^{-1} $Å^{-2}$). Equilibration dynamics was performed on the minimized structure, with a constant temperature of 300 K maintained by coupling to a thermostat using the Langevin algorithm with the collision frequency set to 1 ps^{-1}. This reduced viscosity has been shown to facilitate rapid structural rearrangement (*173*). During dynamics, restraints on the backbone atoms were gradually released in a stepwise fashion in 25ps increments (force constants 10, 7, 4 to 1 kcal/mol/$Å^2$). Further equilibration with no restraints was performed for 500ps at 300K. The time step was 1fs. During all simulations, all possible nonbonded interactions were evaluated at each time step (i.e. no cutoff was employed). Solvation effects were incorporated using the Generalized Born model (igb=1) as implemented in Amber (*174-176*). RMSD values were calculated using the initial model structure (prior to equilibration) as a reference.

The production phase consisted of 1 ns simulations at 300 K with parameters as described above. Simulations were fully unrestrained with the exception of a distance restraint between the C3 atom of the substrate and C4 of the NADH nicotinamide ring, with a force constant of 10 kcal mol^{-1} $Å^{-2}$. The initial separation was ~18 Å, and the final distance was 4.6 Å. We modified Amber to apply the restraint force only to the substrate C3 in order to avoid pulling NADH out of the binding pocket. After the substrate was drawn into the active site, a 100 ps simulation was performed to equilibrate the system with the substrate in the active site. Finally, this restraint was released and an additional 100 ps of fully unrestrained simulation was performed.

The sensitivity of the results to the force field parameters was studied by repeating the procedure using the ff99SB protein force field (*177*) and a newer variant of the GB solvation model (igb=5) (*178, 179*). We recently showed that this combination of force field and solvent model was able to accurately reproduce experimental data for large conformational changes in HIV-1 protease that occur upon addition or removal of an inhibitor (*180, 181*). After the substrate was drawn into the FabI active site, the system was simulated for an additional 2500 ps.

3.3 Results

3.3.1 X-ray Crystallography

FabI forms a 110.9 kDa tetramer with 222 symmetry in which each monomer is formed by a central β-sheet that contains seven β-strands sandwiched by eight helices. Two FabI monomers (A and C) within the tetramer form a complex with ACP, resulting in a stoichiometry of 2:1 for the FabI:ACP interaction (Figure 20). The ACP-bound FabI monomers and the uncomplexed FabI monomers are oriented in such a way that they form a nearly continuous β-sheet across the dimer interface.

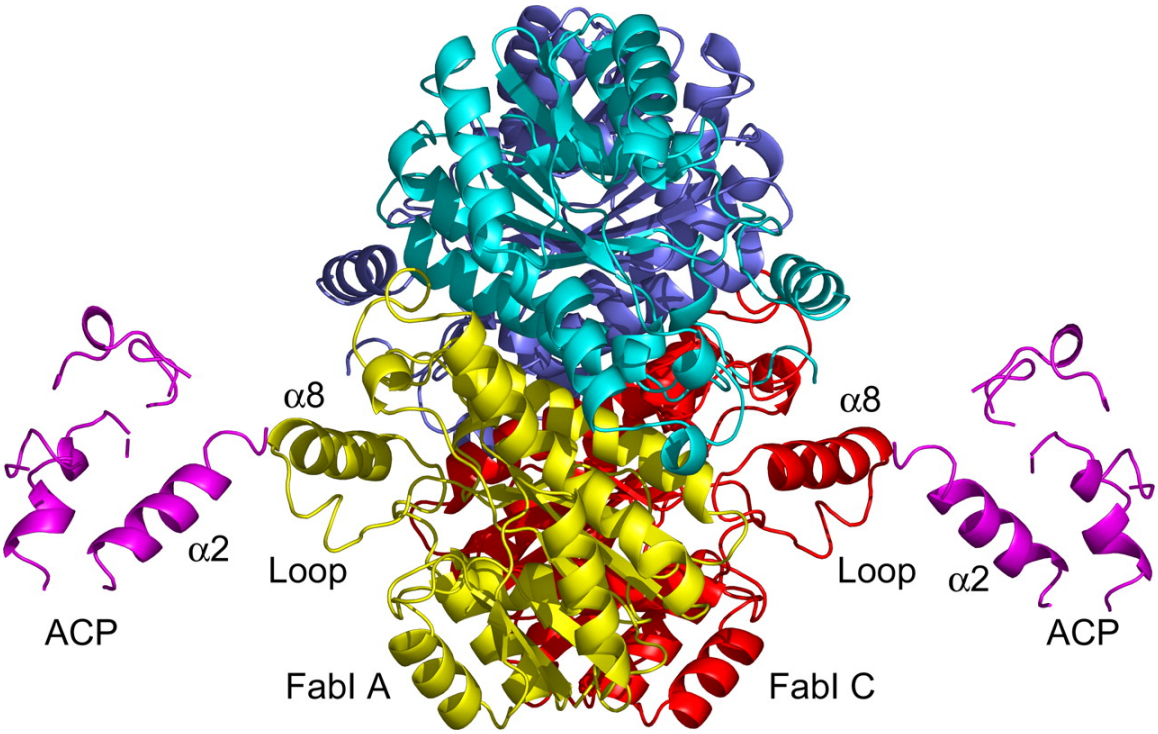


Figure 20: X-ray crystallographic structure of the FabI·ACP complex. X-ray structure of the FabI·ACP complex. Two ACP molecules (*pink*) are bound to the FabI tetramer and interact with FabI monomers labeled *A* (*yellow*) and *C* (*red*). The principal interaction interface is between ACP helix α2 and FabI helix α8. The FabI helix α8 is adjacent to the FabI substrate binding loop (*Loop*). The figure was made with pymol (*85*)

Superposition of the ACP-bound FabI monomers with FabI from the TCS-bound structure (1QSG.pdb) (*161*) results in an RMS deviation of 0.58 Å for 252 Cα atoms for residues 2-194 and 200-259. Thus, the overall structure of FabI does not change upon binding to ACP. However, the substrate binding loop (residues 191-200) undergoes a major conformational change upon complexation with ACP, while FabI helix α8 (residues 201-213) is shifted towards helix α2 of ACP (*161*) (Figure 21). This could possiblAlthough the density is missing for most of the side chains in ACP, the distance between the main chain

atoms of these two helices indicates that helix α8 of FabI interacts with helix α2 of ACP with a main chain separation of about 8.5 Å, over a helix-helix interface length of roughly 11 Å. The substrate binding loop in the TCS-bound structure, which is connected to helix α8, forms a lid on top of the TCS moiety and the nicotinamide ring, thereby shielding them from the solvent. In contrast, this loop adopts an open lid conformation in the FabI-ACP structure, likely due to interaction with ACP, although the details of this interaction cannot be elucidated from the crystallographic data due to missing side chain density. Finally, the observation that only two ACP molecules are observed in the complex with the FabI tetramer could be due to crystal packing. We attempted to dock ACP molecules to the corresponding positions on the other monomers and obtained steric conflicts between the existing ACP molecules, suggesting that the lack of 1:1 stoichiometry is biologically relevant. One FabI monomer – ACP pair was used for this. On the current crystal structure, the FabI monomer was overlapped on the FabI protein and saved the coordinates for the ACP. This was followed by overlapping the FabI on the other monomer seen in the crystal structure and the coordinates for ACP were saved. This gave the corresponding position of ACP to the other FabI monomer.

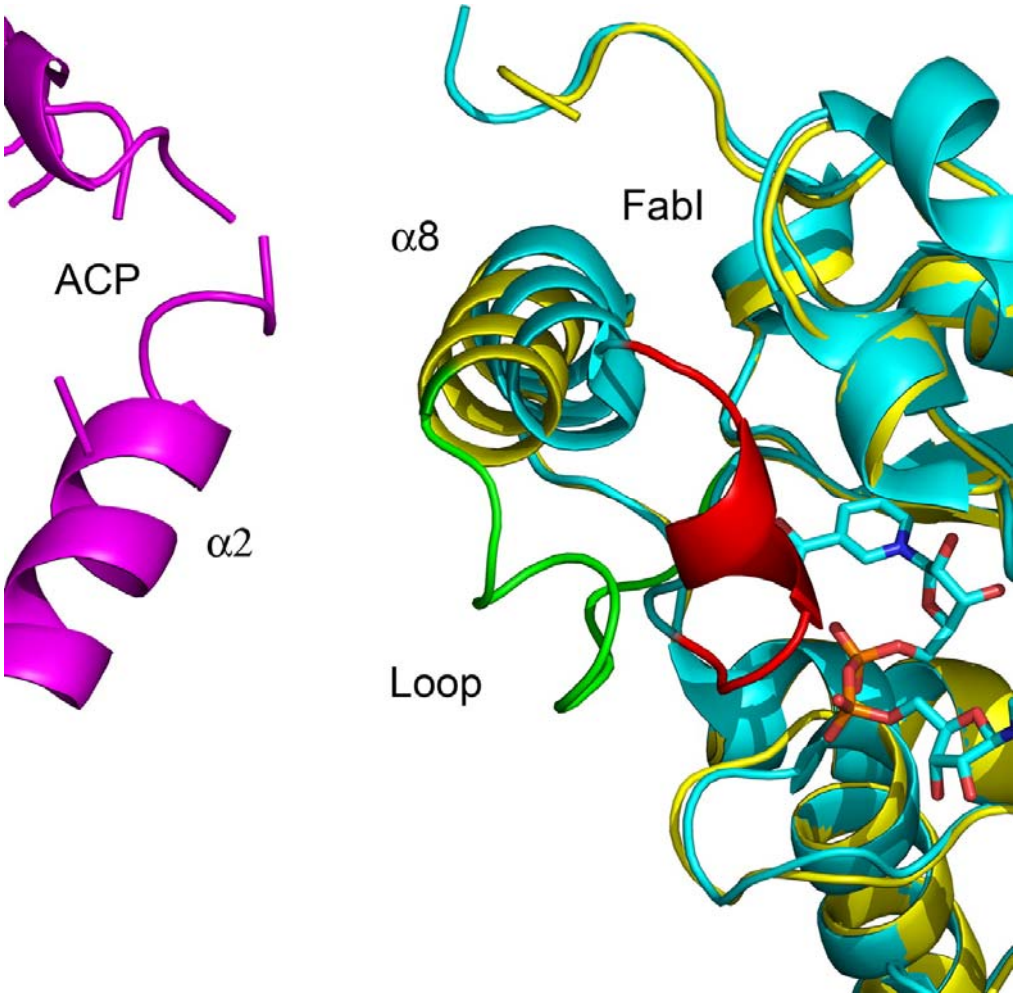


Figure 21: The substrate binding loop in the FabI·ACP and FabI·TCS complexes. Superposition of FabI bound to TCS (*cyan*) and bound to ACP (*yellow*). The FabI substrate binding loop is colored *red* in the FabI·TCS structure and *green* in the FabI·ACP structure. For clarity NAD^+ is shown but TCS is not. The figure was made with pymol (*85*)

3.3.2 Model Building and Molecular Dynamics

Since the side chains for both molecules in the region of the FabI-ACP interface were not resolved by the crystallographic data, we began our computational study by building a model that included these coordinates (see

Methods for details), followed by a series of energy minimization and molecular dynamics (MD) simulations to relax the resulting structure of the complex. The proteins were stable, with the RMSD of FabI and ACP individually remaining below 3 Å. The FabI system was unrestrained in these MD simulations unlike as done in the FabI-inhibitor studies. This is because, here, the interaction of FabI with another protein, namely, ACP is studied. The relatively small contact interface between the proteins permitted a somewhat larger degree of relative motion of the two proteins, with the overall RMSD of the system reaching a plateau of ~3.5 Å.

Structural analysis of the resulting model indicates that the complex is predominantly stabilized through hydrogen bonding interactions between basic residues of FabI in helix α8 and acidic residues of ACP in helix α2, consistent with the structural and modeling studies on the interaction of ACP with AcpS, FabH and FabG (*157-159*). However, while the X-ray crystallographic data and computational modeling studies provide information on the nature of the interaction between FabI and ACP, the absence of electron density for the ACP pantetheine hindered our ability to predict how the ACP delivers the substrate into the FabI active site. Consequently, we set out to build a model of a productive FabI-ACP complex using the computationally refined crystallographic model as a starting point and attaching a crotonyl-phosphopantetheine group to the side chain of S36 in ACP. Importantly, these simulations also help determine

whether the relative positions of FabI and ACP observed in the crystal structure permit delivery of substrate to the active site.

During MD simulations of this structure, the substrate remained outside the FabI active site. This is expected since diffusion of the substrate into the protein is likely an inaccessible event during standard MD simulations on the nanosecond timescale. We thus obtained a model for the productive complex by placing a distance restraint between the crotonyl C3 carbon and the NADH C4 carbon, reducing the target value during a 1 ns MD simulation to draw the substrate into the active site. No other restraints were employed to enforce any particular binding mode. Subsequently, the restraint between the substrate and FabI active site was released and a further completely unrestrained simulation was performed for 100 ps.

Similar to the behavior without substrate, the individual proteins remained relatively stable during these simulations, with backbone RMSD less than or equal to 3.0 Å as compared to the initial model. We observed a greater movement of the ACP molecule with respect to FabI, which allowed S36 to be oriented towards the active site cavity. In order to gain more specific insight into the structural changes that occurred during substrate entry, we performed a superposition of the initial and final structures of the entire protein and then separately calculated the RMSD values for each residue without refitting (**Figure 22**). These show that the RMSD for the majority of FabI was < 2.0 Å, whereas

the active site loop residues (95-115 and 190-210) showed a much larger deviation of > 5 Å. The larger fluctuations for these same loop regions even in the absence of substrate suggest that their flexibility may accommodate entry of substrate into the active site. The RMSD values for ACP are distributed more evenly across the sequence (3.5 Å); this is not unexpected since the initial model was based on the crystal structure of ACP in the absence of FabI with the butyryl moiety accommodated in an internal cavity of the ACP.

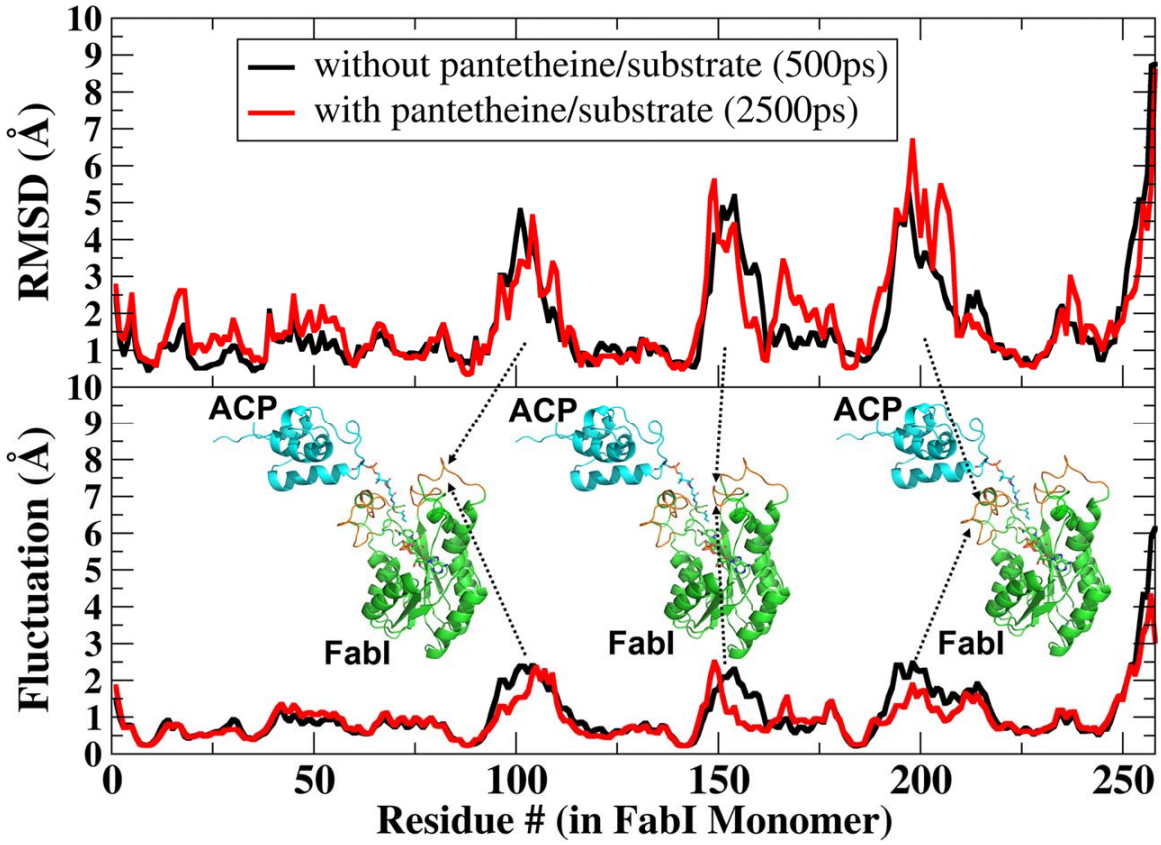


Figure 22: r.m.s. deviation values and fluctuations for each FabI residue during MD simulations of the complex. Average r.m.s. deviation values and positional fluctuations for each residue in FabI during simulations of the complex without substrate (*black lines*) and after drawing the substrate into the active site (*red lines*). Most regions of the protein have deviations of only 1-2 Å from the crystal structure, whereas several loop regions (indicated by the *arrows*) show larger r.m.s. deviation values that arise from increased, as indicated by higher

fluctuations. These loops show similar behavior in the presence and absence of the substrate.

Interactions in the Productive FabI-ACP Complex

The structure of the final FabI-ACP structure is shown in **Figure 23**. ACP interacts with FabI helix α8 and delivers the substrate to the active site between helix α8 and a loop comprised of FabI residues 152-156. Analysis of this structure reveals several important interactions at the FabI-ACP interface and also between the phosphopantetheine and the FabI protein (**Figure 24**). Residues K201, R204, and K205 from helix α8 of FabI contact residues D35, D38, E41, and E48 of ACP helix α2, while FabI K201 also interacts with Q14 in ACP. The side chain amino group of FabI K205, which interacts with D35 in ACP helix α2, is hydrogen bonded to the phosphopantetheine phosphate (O7). In addition, the backbone carbonyl of FabI K205 forms a hydrogen bond to the pantetheine hydroxyl group (O10) while Hε2 of H209 is hydrogen bonded to the pantetheine 4" amide carbonyl oxygen. Finally, the pantetheine 2" amide nitrogen forms a hydrogen bond to the backbone carbonyl of D202 (**Figure 24**). These interactions between the ACP pantetheine and FabI undoubtedly play an important role in stabilizing the FabI-ACP complex and in positioning the substrate within the active site. In the final structure, the distance between the crotonyl C3 and the NADH pro4(*S*) proton is 3 Å. Analysis of the structure also

reveals that the crotonyl thioester carbonyl is located 4 Å from the Y146 hydroxyl group, suggesting that Y146 may form a hydrogen bond to the thioester during substrate reduction (Figure 24).

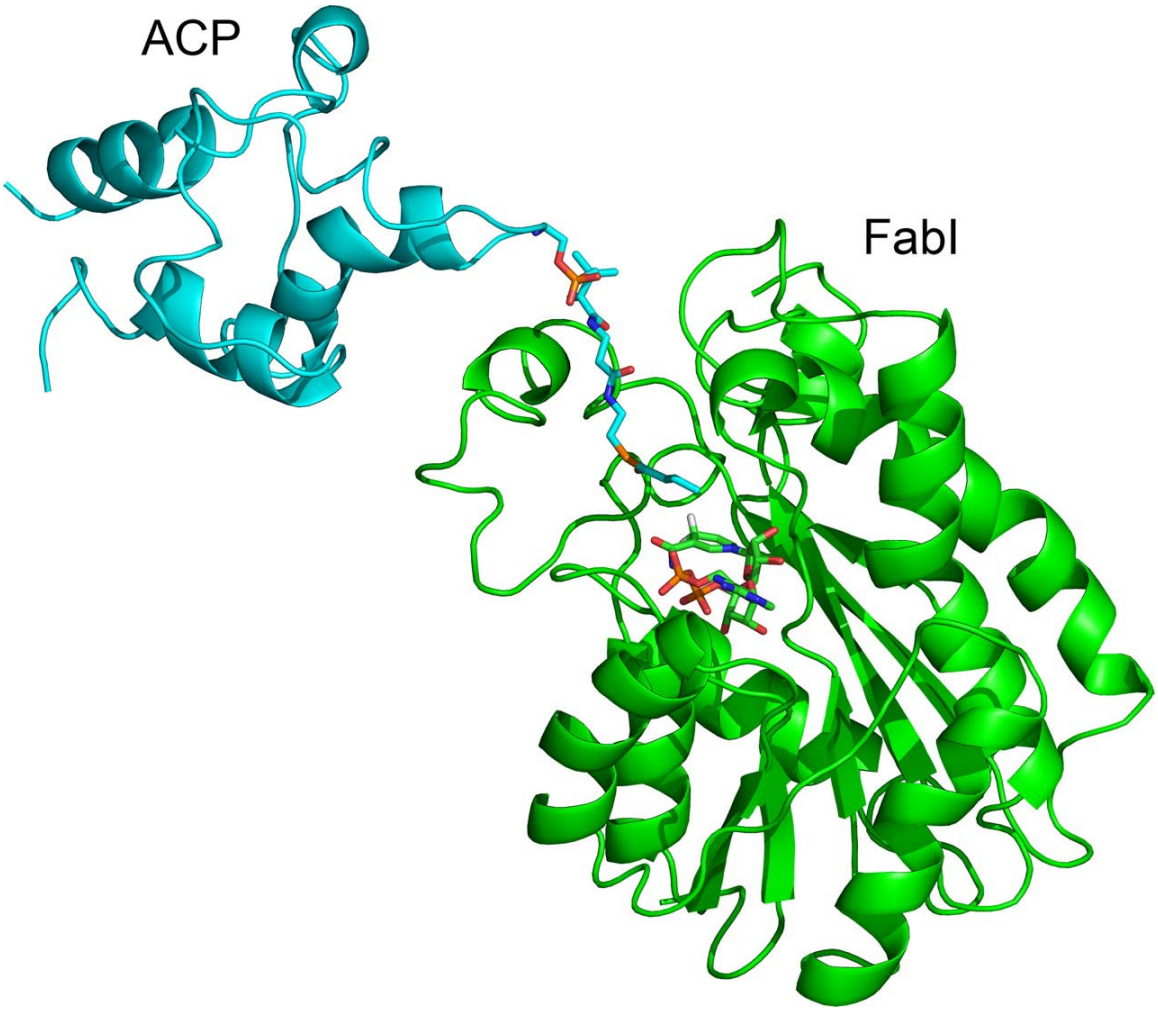


Figure 23: The structure of ACP bound to FabI following MD simulations. Final structure of the FabI·ACP complex. FabI is colored *green* and ACP is colored *cyan*. The figure was made with pymol (*85*)

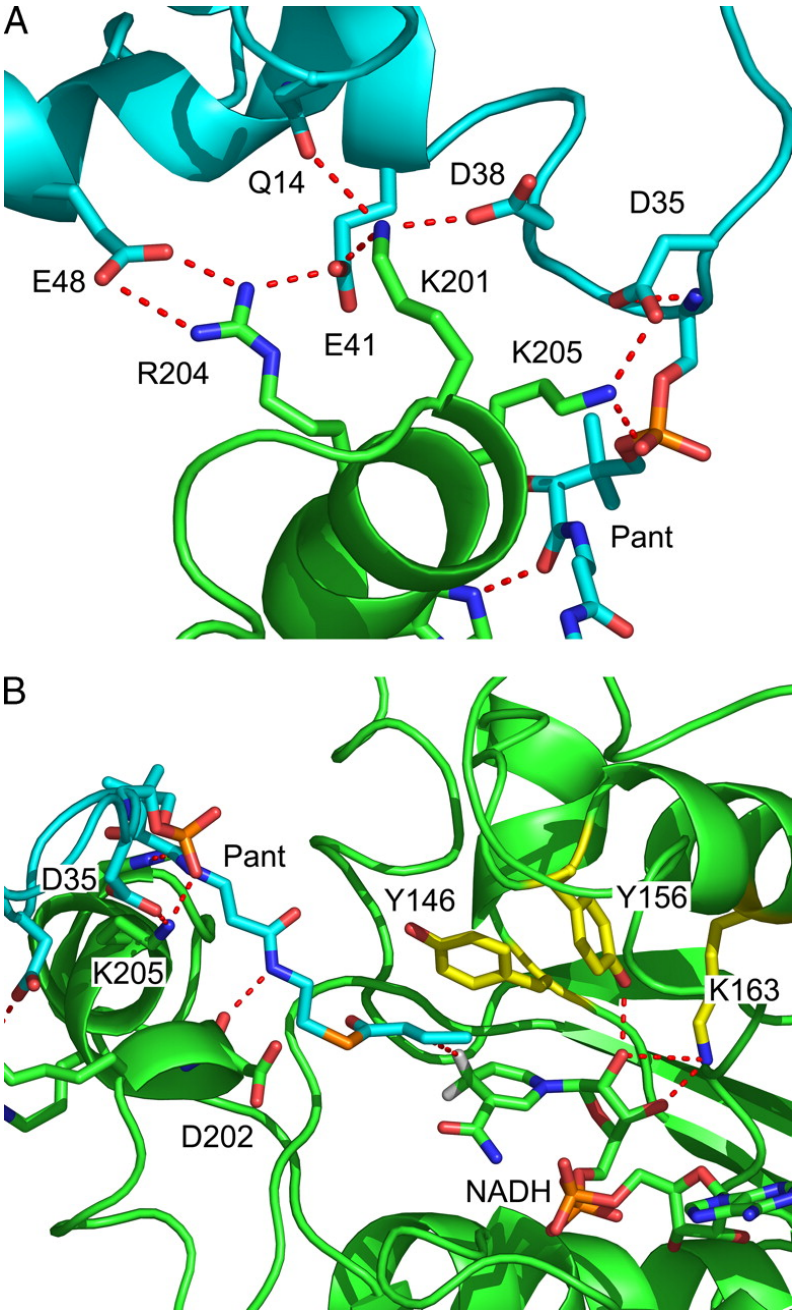


Figure 24: Interactions between FabI and ACP. *A*, interactions between ACP (*cyan*) and FabI (*green*) at the helix α2 (ACP)-helix α8 (FabI) interface. *B*, interactions between crotonyl-pantetheine and FabI. The pantetheine (*cyan*) is hydrogen bonded to residues in FabI helix α8 (*green*). FabI residues in the conserved active site triad (Tyr^{146}, Tyr^{156}, and Lys^{163}) are colored *yellow*. The crotonyl group of the substrate (*cyan*) is bound in the *s-trans* conformation and the crotonyl carbonyl group is oriented toward Tyr^{146} (*yellow*). The C-3 carbon of the crotonyl group is 3 Å from the NADH pro4(*S*) proton (*white*). In addition, the NADH ribose (*cyan*) is hydrogen bonded to Tyr^{156} and Lys^{163}. The figure was made with pymol (*85*)

In order to investigate the influence of the computational protocol on the results, we repeated the process using a different protein force field and solvent model (see Methods). We recently employed this combination to successfully simulate ligand-induced conformational changes in HIV-1 protease (*180, 181*). Importantly, we also demonstrated that this particular implicit solvent model accurately reproduces the stability of salt bridges between protein side chains as compared to simulations in explicit solvent (*182*). Simulations with these parameters gave very similar results to those described above. The crotonyl thioester carbonyl is located even closer (2.7 Å) to the Y146 hydroxyl group, providing further evidence that Y146 may form a hydrogen bond to the thioester during substrate reduction.

3.3.3 Kinetic Analysis of Wild-type and Mutant FabI Enzymes

The structural studies described above identify several key interactions in the FabI-ACP complex that have been evaluated by replacing the basic FabI residues with acidic groups and Ala. The data in **Table 13** demonstrates that replacement of K201, R204 and K205 with Ala has little or no effect on the kinetic parameters for reduction of DD-CoA whereas k_{cat}/K_m for reduction of DD-ACP is reduced 5 (K201, K205) to 50 (R204) fold. In addition, replacement of R204 and K205 with Glu causes a further reduction in k_{cat}/K_m for reduction of DD-ACP

without affecting k_{cat}/K_m for the DD-CoA substrate. Similar to the Ala mutants, substitution of Glu for R204 has a larger impact on substrate reduction (250-fold) compared to K205 (14-fold). Finally, replacement of K201 with Glu resulted in an enzyme with little activity toward either substrate and we were unable to determine accurate kinetic parameters for this mutant.

Table 13: Kinetic parameters for wild-type and mutant FabIs

	DD-CoA				DD-ACP			
	k_{cat}	K_m	k_{cat}/K_m	Ratio[a]	k_{cat}	K_m	k_{cat}/K_m	Ratio[a]
	min^{-1}	μM	$\mu M^{-1}\ min^{-1}$		min^{-1}	μM	$\mu M^{-1}\ min^{-1}$	
Wild-type	800 ± 40	24 ± 3	33 ± 6	1	900 ± 16	3.3 ± 0.2	270 ± 20	1
K201A[b]	733 ± 20	22 ± 1	33 ± 3	1	582 ± 12	13 ± 1	45 ± 4	0.2
R204A	1470 ± 53	18 ± 2	82 ± 12	2.5	1550 ± 90	291 ± 31	5 ± 0.8	0.02
R204E	1300 ± 70	16 ± 2	81 ± 14	2.5	94 ± 4	90 ± 9	1.0 ± 0.1	0.004
K205A	716 ± 38	24 ± 5	30 ± 7	0.9	860 ± 30	14 ± 2	61 ± 11	0.2
K205E	440 ± 8	21 ± 1	21 ± 1	0.6	1030 ± 20	48 ± 2	22 ± 1	0.07
Y156F	1160 ± 35	19 ± 1	61 ± 5	1.8	1140 ± 50	4.6 ± 0.5	250 ± 40	0.9
Y146F	24 ± 1	10 ± 1	2.4 ± 0.3	0.07	14 ± 1	2.4 ± 0.2	6 ± 1	0.02

[a] Ratio of k_{cat}/K_m values for wild-type and mutant FabIs
[b] The K201E mutant had very low activity with both DD-CoA and DD-ACP substrates

We have also examined the importance of Y156 and Y146, the two active site Tyr residues, in substrate reduction. Replacement of Y156 with Phe has no effect on substrate reduction in agreement with previous studies on both FabI (*162*) and InhA (*120, 183*), which questioned the importance of Y156 (Y158 in

InhA) in catalysis (*184*). In contrast, mutagenesis of Y146 has a larger impact on catalysis, with k_{cat} and k_{cat}/K_m for DD-ACP decreasing by around 50-fold compared to wild-type FabI. A similar decrease in kinetic parameters for Y146F was also observed for the DD-CoA substrate.

3.4 Discussion

ACPs are small, acidic proteins that fulfill an essential role in metabolism through their interactions with a diverse array of target enzymes. However, despite their central role, detailed structural information on the acyl-group specific recognition of ACPs by their target proteins has remained elusive, presumably partly as a result of the conformational flexibility of the ACP molecule. Here we report the first structural data for the direct interaction of an acyl-ACP substrate with a target enzyme, the FASII FabI enzyme, based on a combination of X-ray crystallography and computational modeling. FabI is the enoyl reductase in the bacterial FASII pathway and a target for antibacterial drug discovery. Thus, not only does the structure of the FabI:ACP complex provide general insight into how target proteins recognize and bind to acyl-ACPs, the present structure also provides a foundation for the development of novel FabI inhibitors that antagonize the interaction of FabI with its natural substrate.

In agreement with previous predictions, several acidic residues in and close to the ACP helix α2 (D35, D38 and E41) form stable electrostatic interactions with three basic amino acids (K201, R204 and K205) located adjacent to the FabI substrate binding loop. Importantly, replacement of K201, R204 or K205 by Ala or Glu results in significant decreases in k_{cat}/K_m for reduction of DD-ACP, caused both by a decrease in k_{cat} and an increase in K_m, without affecting the kinetic parameters for reduction of the corresponding CoA substrate. The effect on DD-ACP reduction was larger for the Glu substitutions, with R204 demonstrating the most sensitivity to mutation. Thus, these residues are involved in specific interactions with the protein portion of DD-ACP, as was shown in similar experiments with FabH and FabG (*158, 159*). The FabI-ACP interactions position S36, the ACP residue that carries the phosphopantetheine, above an opening into the active site formed by the substrate binding loop, helix α8 comprised of residues 192-206 and a mobile loop comprised of residues 152-156. These two loops move apart to allow the phosphopantetheine to deliver the substrate to the active site through the minor portal (*165*). While the C3 carbon of the enoyl substrate is at a distance of approximately 3 Å from the pro4(*S*) NADH proton, the positioning of the substrate into the active site is unexpected given previous structural data on inhibitors bound to FabI and the hexadecenoyl-N-acetylcysteamine (C16-NAC) substrate (*165*) bound to InhA. Below we discuss the orientation of the substrate with respect to the catalytic triad in the FabI active site.

FabI is a member of the short chain alcohol dehydrogenase/reductase family. This superfamily is characterized by a conserved triad of active site residues. In FabI the triad is comprised of Y146, Y156 and K163, while in InhA, the *M. tuberculosis* enoyl reductase, the triad is F149, Y158 and K165. Mechanistic information on the role of these residues in catalysis has been provided by site-directed mutagenesis coupled with structural data primarily arising from enzyme-inhibitor rather than enzyme-substrate complexes. Inhibitors that have been structurally characterized in complex with these enzymes include compounds such as the diazaborines (*116*) and INH (*117*) which modify the NAD(H) cofactor, and those, such as TCS, that bind noncovalently to the enzyme-cofactor complex (*160, 161*).

TCS, which binds with picomolar affinity to the *E. coli* FabI (*162, 185, 186*), has been proposed to bind to the enzyme as a substrate analog (*59*). This hypothesis has gained support from the structure of hexadecenoyl-N-acetylcysteamine (C16-NAC) bound in a stable ternary complex to InhA in the presence of NAD^+ (*165*). Y156 of FabI (Y158 in InhA) forms a hydrogen bond to the phenol of TCS and the carbonyl oxygen of the C16 substrate. Mutagenesis clearly supports the importance of Y156 (Y158) in TCS binding, and replacement of this residue by Phe increases the K_i for TCS inhibition by 400 (160)-fold (*120, 162*). However, the impact of mutating Y156 (Y158) with respect to its role in catalysis is much less pronounced. In FabI, the kinetic parameters are unaffected by replacement of Y156 with Phe, while the corresponding mutation in InhA

(Y158F) has a slightly stronger effect on catalysis, reducing k_{cat} 24-fold compared to wild-type. Intriguingly, however, the Y158S InhA mutant displays wild-type activity (*183*). Thus, Y156 does not play a significant role in substrate reduction and, in agreement with these data, the substrate carbonyl group in the present FabI-ACP structure is pointing away from Y156 (Figure 24).

The function of the second aromatic residue in the enoyl reductase triad (Y146/F149) is also not clear, and Rozwarski et al. have argued that F149 in InhA is involved in directing the NADH to deliver a hydride to the correct position on the substrate and/or in modulating the interaction of the bound substrate with a channel of water molecules that leads away from the active site (*165*). In the simulation model, the FabI Y146 hydroxyl is located within hydrogen bonding distance from the substrate carbonyl, suggesting that Y146 rather than Y156 provides electrophilic assistance during substrate reduction. To probe the role of this residue in catalysis, the Y146F mutant was characterized and shown to catalyze substrate reduction with a k_{cat}/K_m value 14 (DD-CoA) to 50 (DD-ACP)-fold lower than wild-type FabI. These data suggest that the Y146 hydroxyl is directly involved in catalysis, consistent with the hydrogen bonding interaction revealed by the MD simulations. We note, however, that the homologue of Y146 in InhA is a Phe (F149), and so is unable to provide a hydrogen bond to the substrate. Indeed, Raman studies on deuterated NADD cofactor bound to InhA suggest that F149 most likely is involved in correctly positioning the cofactor for

hydride transfer (unpublished results). Thus, one possibility is that Y146 (FabI) and F149 (InhA) play different roles in catalysis.

Returning to previous structural studies, Sacchettini and coworkers have discussed two entry points for substrates into the active sites of enoyl reductases termed the major and minor portals (*165*). Based on the structures of inhibitors bound to FabI and InhA and that of the InhA:C16-NAC complex, the fatty acid substrate would have been expected to enter the FabI active site through the major portal. Instead, the present structure indicates that binding of ACP to FabI delivers the acyl-pantetheine to the active site between loops comprised of residues 192-206 and 152-156 such that the fatty acid enters the active site through, or adjacent to, the minor portal. This structure is consistent with our mutagenesis data and also the ACP-FabI binding interface proposed by Rock and coworkers (*158, 159*). Indeed, interaction of ACP with the basic patch of residues on FabI leaves the ACP S36 residue too far from the active site to deliver the substrate through the major portal, thus predicating an alternative entry point for the substrate. Interestingly, in the recently determined structure of 5-octyl-2-phenoxyphenol bound to InhA, the analogous channel in InhA is occupied by the alkyl chain of the inhibitor which binds to the enzyme with a K_i value of 1 nM (2B37.pdb) (*187*). Taken together, these data indicate that substrate binding occurs in the opposite orientation to that expected from studies using inhibitors or the truncated C16-NAC substrate. In the case of InhA, we do not as yet have a structure of the relevant complex with ACP and it is possible

that ACP delivers the substrate to this enzyme through the major portal, as suggested by the C16-NAC structure. Alternatively, without the ACP to locate it, the C16-NAC molecule could bind to InhA in a mode that more closely approximates an inhibitor rather than a substrate. The two resulting possibilities then are that the C16-NAC could be reduced by the enzyme in this alternative binding mode or that the C16-NAC is bound nonproductively to the enzyme in the complex used for the structural studies. While we are unable to differentiate between these possibilities at the present time, we note that the C16-NAC molecule is bound to the enzyme in the presence of NAD^+ rather than NADH, as observed for the diphenyl ether inhibitors. Future studies on the interaction of ACP with InhA will shed light on these possibilities. It thus leads us to the question of whether all FabIs interact with ACP in the same way.

3.5 Investigation of the interaction of AcpM with InhA in order to deliver the substrate into the active site of InhA

The interaction of ACP with FabI, enoyl ACP reductase from *E. coli* has been studied using X-ray crystallography, molecular modeling and biochemical techniques. Based on the ecFabI-ACP module, one would expect the AcpM with the substrate to interact with the basic residues in the substrate binding loop region of InhA and deliver the substrate through the minor portal of InhA. The position corresponding to the basic residues in the substrate binding loop of FabI

is occupied by V193, Q206 and L207 in InhA. The interaction of these residues with the acidic residues of AcpM may not be favorable, and indeed Xujie Zhang demonstrated using site-directed mutagenesis that replacement of these residues, had little impact on the interaction of AcpM with InhA. Subsequent analysis resulted in the identification of two alternative patches of basic amino acids, R45, R49, R53 and R195, R225 and R233, the former being close to the InhA major portal (Figure 25). Modification of R45,49,53 resulted in a dramatic reduction in k_{cat}/K_m for reduction of C12-AcpM but not of C12-CoA, indicating that these residues play a critical role in the interaction of AcpM with InhA, possibly via the AcpM helix α2. In addition, mutagenesis of R195, R225 and R233 also resulted in a reduction in k_{cat}/K_m for reduction of C12-AcpM, albeit to a smaller extent than replacement of residues inthe R45 patch. These data resulted in an alternative model for the interaction of AcpM with InhA, in which the substrate is delivered through the major portal. AcpM differs from ACP in the addition of 38 residues at the C-terminus, and we propose that this C-terminal tail interacts with R195, R225 and R233 in InhA. In support of this model, mutagensis of R45, R49 and R53 also reduces the k_{cat}/K_m for reduction of C12-ACP, indicating that both ACP and AcpM interact with this basic patch in InhA. This divergence in substrate delivery between FabI and InhA may be related to the chain length specificities of the two enzymes, with InhA capable of reducing enoyl-AcpMs up to C56 in length.


ecolifabI ------MGFLSGKRILVTGVASKLSIAYGIAQAMHREGAELAFTYQND---KLKGRVEEF 51
photolistFabI ------MGFMTGKRILITGVASKLSIAYGVAKAMHDQGAELAFTYQND---KLKPRVEEF 51
LegiolistFabI MGGDTIVGFLTGKKALIVGLASNRSIAYGIAKAFHNQGAELAFTYQNE---KLQSRVEAM 57
pyloriFabI ------MGFLKGKKGLIVGVANNKSIAYGIAQSCFNQGATLAFTYLNE---SLEKRVRPI 51
aureusFabI ------MLNLENKTYVIMGIANKRSIAFGVAKVLDQLGAKLVFTYRKE---RSRKELEKL 51
subtilisFabI -----MNFSLEGRNIVVMGVANKRSIAWGIARSLHEAGARLIFTYAGE---RLEKSVHEL 52
cyanolistFabI MTTKISMLNLTGKNALVTGIANNRSIAWGIAQQLHAAGANLGITYLPDERGKFEKKVSEL 60
mtbInhA -----MTGLLDGKRILVSGIITDSSIAFHIARVAQEQGAQLVLTGFDR-----IRLIQRL 50
bovisInhA -----MTGLLDGKRILVSGIITDSSIAFHIARVAQEQGAQLVLTGFDR-----IRLIQRL 50
marimumInhA -----MAGLLEGKRILVSGIITDSSIAFHIARVAQEQGAQLVLTGFDR-----MRLIQRL 50
burulistInhA -----MAGLLEGKRILVSGIITDSSIAFHIARVAQEQGAQLVLTGFDR-----MRLIQRL 50
lepraeInhA -----MAGLLEGKRILVSGIITDSSIAFHIAKVAQEAGAQLVLTGFDR-----IRLIQRL 50
ecolifabI AAQLG--SDIVLQCDVAEDASIDTMFAELGKVWP---KFDGFVHSIGFAPGDQLDGDYVN 106
photolistFabI AASLN--STIVLPCDVAEDESIEVLFTELSKIWP---KFDGFVHSIGFAPADQLDGNYVS 106
LegiolistFabI ASEFN--STLVFPCDVASDEEIKAVFDNLRNHWD---KLDILVHSVAYAPADQISGDFVE 112
pyloriFabI AQELN--SPYVYELDVSKEEHFKSLYNNIKQDLG---SLDFIVHSVAFAPKEALEGSLLE 106
aureusFabI LEQLNQPEAHLYQIDVQSDEEVINGFEQIGKDVG---NIDGVYHSIAFANMEDLRGRFSE 108
subtilisFabI AGTLDRNDSIILPCDVTNDAEIETCFASIKEQVG---VIHGIAHCIAFANKEELVGEYLN 109
cyanolistFabI VEPLN--PSLFLPCNVQNDEQIQSTFDTIRDKWG---RLDILIHCLAFANRDDLTGDFSQ 115
mtbInhA TIRLP-AKAPLLELDVQNEEHLASLAGRVTEAIGAGNKLDGVVHSIGFMPQTGMGINPFF 109
bovisInhA TIRLP-AKAPLLELDVQNEEHLASLAGRVTEAIGAGNKLDGVVHSIGFMPQTGMGINPFF 109
marimumInhA VIRLP-QKAPLIELDVQNEEHLASLAGRVTEVIGEGNNLDGVVHSIGFMPQSGMGINPFF 109
burulistInhA VIRLP-QKAPLIELDVQNEEHLASLAGRVTEVIGEGNNLDGVVHSIGFMPQSGMGINPFF 109
lepraeInhA AIRLP-DKAPLIELDVQNEEHLATLAERVTAEIGEGNKLDGVVHSIGFMPQTGMGTNQFF 109
ecolifabI AVTREGFKIAHDISSYSFVAMAKACR-SMLNPGSALLTLSYLGAERAIPNYNVMGLAKAS 165
photolistFabI SVTREGFRIAHDISSYSFVAMAKTCR-EMLNPNSALLTLTYLGAERSIPNYNVMGLAKAS 165
LegiolistFabI CANREGFRIAHDISAYSLIGLSQAALPMMLDTQGSILTLSYYGAEKAVPNYNVMGVAKAS 172
pyloriFabI TS-KSAFNTAMEISVYSLIELTNTLK-PLLNNGASVLTLSYLGSTKYMAHYNVMGLAKAA 164
aureusFabI TS-REGFLLAQDISSYSLTIVAHEAK-KLMPEGGSIVATTYLGGEFAVQNYNVMGVAKAS 166
subtilisFabI TN-RDGFLLAHNISSYSLTAVVKAAR-PMMTEGGSIVTLTYLGGELVMPNYNVMGVAKAS 167
cyanolistFabI TS-RAGFATALDISTFSLVQLSGAAK-PLMTEGGSIITLSYLGGVRAVPNYNVMGVAKAG 173
mtbInhA DAPYADVSKGIHISAYSYASMAKALL-PIMNPGGSIVGMDFD-PSRAMPAYNWMTVAKSA 167
bovisInhA DAPYADVSKGIHISAYSYASMAKALL-PIMNPGGSIVGMDFD-PSRAMPAYNWMTVAKSA 167
marimumInhA DAPYEDVSKGIHISAYSYASLAKALL-PIMNPGGSIVGMDFD-PTRAMPAYNWMTVAKSA 167
burulistInhA DAPYEDVSKGIHISAYSYASLAKALL-PIMNPGGSIVGMDFD-PTRAMPAYNWMTVAKSA 167
lepraeInhA DAPYEDVSKGIHISTYSYASLAKALL-LIMNSGGSIVGMDFD-PTRAMPAYNWMTVAKSA 167
ecolifabI LEANVRYMANAMGPEGVRVNAISAGPIRTLAASGIKD----------IRKMLAHCEAVTP 215
photolistFabI LEANVRYMANAMGAEGIRVNGISAGPIRTLAASGIKD----------IRKMLAHCESVNP 215
LegiolistFabI LEASVRYLAASLGSRGLRINAISAGPIKTLAAAGIKD----------IRKIHAAYANITP 222
pyloriFabI LESAVRYLAVDLGKHNIRVNALSAGPIRTLASSGIAD----------IRMILKWNEINAP 214
aureusFabI LEANVKYLALDLGPDNIRVNAISAGPIRTLSAKGVGG----------INTILKEIEERAP 216
subtilisFabI LDASVKYLAADLGKENIRVNSISAGPIRTLSAKGISD----------FNSILKDIEERAP 217
cyanolistFabI LEASVRYLASELGSQNIRVNAISAGPIRTLASSAVGG----------ILDMIHHVEQVAP 223
mtbInhA LESVNRFVAREAGKYGVRSNLVAAGPIRTLAMSAIVGGALGEEAGAQIQLLEEGWDQRAP 227
bovisInhA LESVNRFVAREAGKYGVRSNLVAAGPIRTLAMSAIVGGALGEEAGAQIQLLEEGWDQRAP 227
marimumInhA LESVNRFVAREAGKYGVRSNLVAAGPIRTLAMSAIVGGALGEEAGAQIQLLEDGWDQRAP 227
burulistInhA LESVNRFVAREAGKYGVRSNLVAAGPIRTLAMSAIVGGALGEEAGAQIQLLEDGWDQRAP 227
lepraeInhA LESVNRFVAREAGKYGVRSNLVAAGPIRTLAMSAIVGGAFGEEAGAQMQLLEEGWDQRAP 227

Figure 25: The Multiple Sequence Alignment of FabI sequences from 12 different species. In red is shown the acidic residues in the substrate binding loop region, in blue is shown the basic residues in the substrate binding loop region and in green are shown a different patch of basic residues identified to be involved in AcpM interaction

Based on these findings, we propose a model for the interaction of AcpM with InhA. AcpM is constituted of 115 amino acids and is different from ACP from *E. coli* (77 amino acids) in having an extra tail (composed of all the additional amino acids). To S41 of AcpM is attached the acylated phospho-pantetheine. The acidic residues (D40, D46) can interact with residues R45, R49 and R53 of InhA and the long tail can interact with residues R195, R225 and R233. These interactions can help position the AcpM in such a way that the substrate can be delivered into the active site of InhA through the major portal (Figure 26).

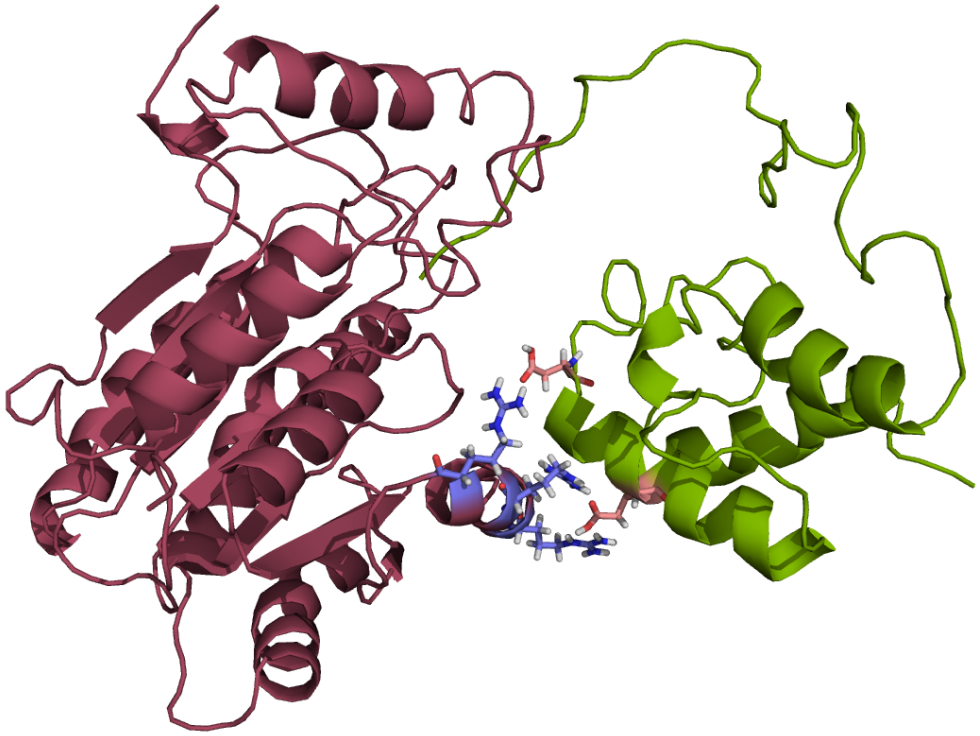


Figure 26: Model of InhA interaction with AcpM. InhA is shown in Raspberry. Residues R45, R49 and R53 are shown in blue. ACP is shown in green and the residues R40 and R46 are shown in red.

3.6 Summary

We have used a combination of X-ray crystallography and molecular dynamics to provide the first structural insight into the detailed interaction of ACP with a target enzyme. The structural data are substantiated by mutagenesis, and reveal that interactions between ACP and FabI are largely electrostatic in nature. Of key importance are interactions between acidic residues in ACP helix α2 and basic residues in FabI helix α8. The ACP phosphopantetheine delivers the enoyl substrate to the active site between helix α8 and a mobile loop comprised of residues 152-156. In agreement with site-directed mutagenesis, the conserved

active site Tyr Y156 is not directly involved in interactions with the substrate. While some of the specific interactions between FabI and ACP differ somewhat between simulations performed using different protein force fields and solvent models, they provide the consistent view that ACP is able to successfully deliver the substrate into the FabI active site through the minor portal and from the position observed in the crystal structure of the complex. We have also identified another patch of basic residues (R45, R49 and R53) in the InhA enzyme that may interact with the acidic amino acids of AcpM. In this case, the AcpM may deliver the substrate through the major portal. Knowledge of the structural determinants that control the interaction of ACP with target enzymes is of critical importance for the design of inhibitors against these enzymes and for fully understanding the multiple roles of ACP in metabolic pathways.

Chapter 4 Understanding the substrate binding loop ordering mechanism in the *M.tuberculosis* enoyl reductase InhA

4.1 InhA as a drug target

InhA is the enoyl reductase enzyme from *Tuberculosis* bacterium. It is a part of FAS-II pathway. INH, the frontline drug used in the treatment of TB infection, targets FAS-II. At least 2 enzymes, KasA and InhA of the FAS-II pathway have been proposed as targets for INH. KasA is one of the three keto-acyl synthases and InhA is an enoyl reductase that catalyzes the NADH dependent reduction of long chain *trans*-2-enoyl-ACPs. As mutations were observed in the *inhA* gene in INH-resistant clinical isolates, Jacob and coworkers first identified InhA, as a target of INH. Also, they identified a point mutant (S94A) that conferred resistance to INH and ethionamide in *M. tuberculosis* and *M. bovis* (*188-191*) Subsequently, Blanchard, Sacchettini and coworkers demonstrated that InhA was inhibited *in vitro* by a covalent adduct formed between activated INH and the nicotinamide head group of NAD (*192, 193*). Mutant InhA was shown to have decreased affinity of the purified enzyme for NADH (*192-195*). In addition, Vilcheze et al. used a temperature sensitive mutation in the *inhA* gene to show the phenotypic response to InhA inactivation

in *M. smegmatis* was identical to that caused by treatment with INH (*196*). Thus InhA was validated to be a definite target for drug discovery. The second putative INH target is KasA. Barry and coworkers elicited the direct role of KasA in INH action who observed the formation of a covalent 80 kDa complex consisting of KasA-AcpM (the TB acyl carrier protein) and INH upon treating *M. tuberculosis* with INH. KasA mutants have been identified in the INH-resistant clinical isolates.(*197*) Research thus far has not been able to definitively determine the ultimate target for INH.(*196, 198-201*) But it is quite possible for INH to have multiple targets in the mycobacterium. In my research InhA would be used as the target for the drug design efforts.

4.2 Diphenyl ethers as the lead compounds

Using structure-based drug design, members of Tonge laboratory have developed a series of alkyl diphenyl ethers that are inhibitors of InhA. The most potent compound has a *K*i´ value of 1 nM for InhA and MIC99 values of 2–3 μg mL-1 (6–10 μM) for both drug-sensitive and drug-resistant strains of *M. tuberculosis* (*202*). Overexpression of InhA in *M. tuberculosis* results in a 9–12-fold increase in MIC99. This is consistent with the belief that these compounds target InhA within the cell. In addition, transcriptional response studies have been performed and these reveal that the alkyl diphenyl ethers fail to upregulate a putative efflux pump and aromatic dioxygenase. These are the detoxification

mechanisms that are triggered by the lead compound TCS. These diphenyl ether-based InhA inhibitors do not require activation by the mycobacterial KatG enzyme (*202*), thus bypassing the normal mechanism of resistance to the front line drug INH and thus accounting for their activity against INH-resistant strains of *M. tuberculosis*.

SAR studies from several sources have thrown light into the interaction of TCS derivatives with enzyme (*34, 50, 52, 203, 204*). Although TCS is a picomolar inhibitor of ecFabI, this compound inhibits InhA with a K_i' value of only 0.2 μM (*54, 55, 58*). InhA and ecFabI are structural and functional homologs (1.6 Å rmsd in α-carbon position) and are 36% identical in sequence (*23*). Consequently, our goal is to understand the 30 000-fold difference in the affinity of TCS for ecFabI and InhA, if it were coupled to conformational change in the protein and to exploit this information to modify TCS and thereby design high-affinity InhA inhibitors.

4.3 The substrate binding loop of InhA, FabI

There is an available 1.75 Å X-ray structure of TCS bound to ecFabI (Figure 27a (*63*). To add to this information, the structure of TCS bound to InhA was determined (*202*) at 2.3 Å resolution. Consistent with the uncompetitive inhibition of InhA by TCS (*58*), the inhibitor binds to InhA in the presence of NAD^+ (Figure 27b). In this structure it is seen that the A ring of TCS stacks with the

nicotinamide ring of NAD^+ and forms a hydrogen bond to Tyr158 and the 2´-hydroxy group of NAD^+ (Figure 27b). This is a highly conserved hydrogen-bonding pattern observed in complexes of TCS with other enoyl reductase enzymes (Figure 27b, black dotted lines) (*53, 63, 64*). The other interactions between TCS and InhA are mostly hydrophobic, and the B ring is positioned orthogonally to the A ring with one of the chlorine atoms pointing toward the NAD^+, while the other is solvent-exposed.

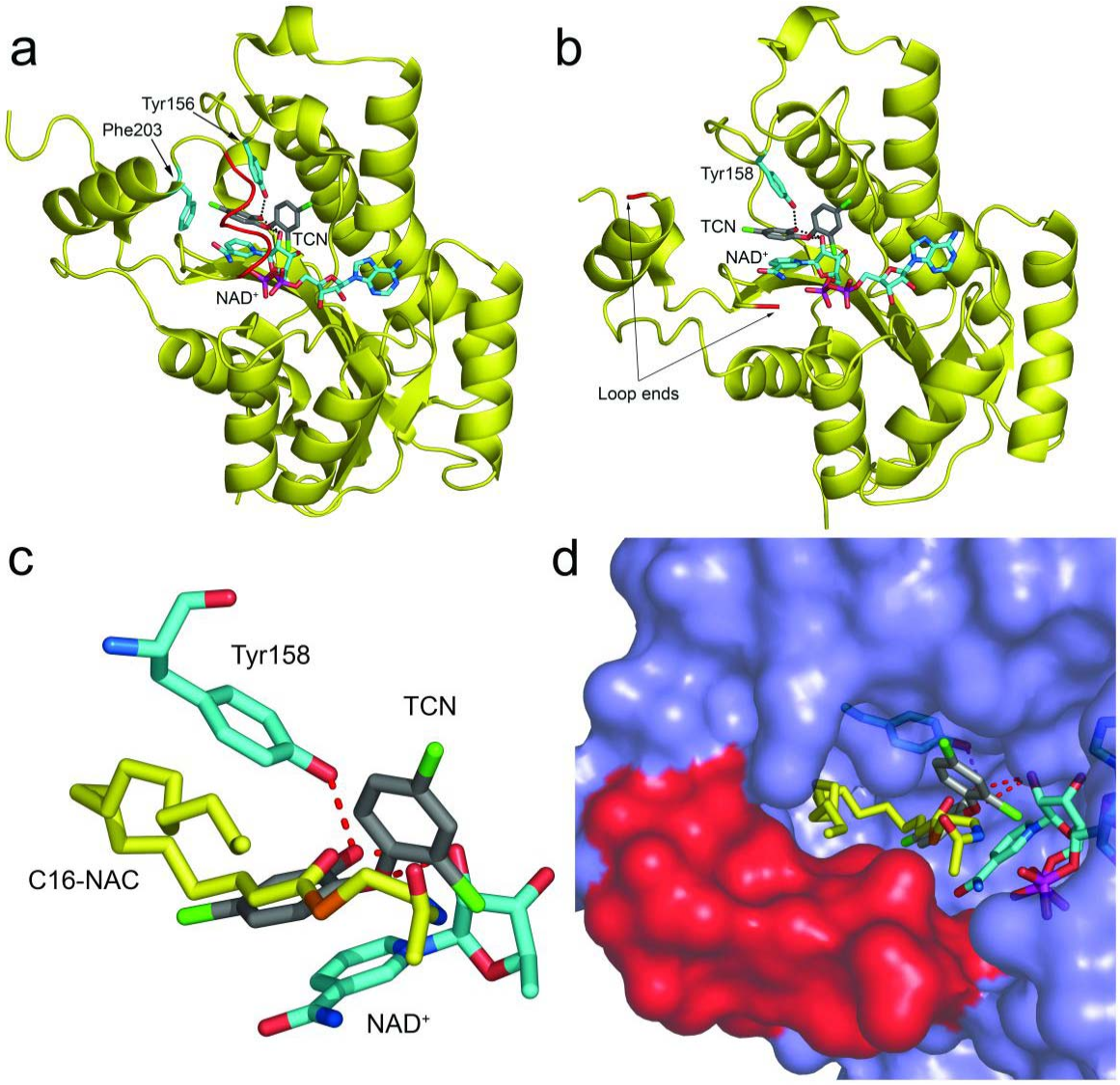


Figure 27: Structures of TCS bound to the *E. coli* FabI and InhA. Structures of TCS (TCN; gray molecule) in complex with NAD^+ (blue molecule) and a) the *E. coli* FabI (1QSG.pdb) and b) InhA. In the FabI-TCS complex, the active site loop is ordered (colored red) and residue Phe203 (blue) is shown adjacent to the loop, while in the InhA-TCS complex, the loop is disordered (loop ends colored red). In panels c and d, the InhA-TCS structure has been overlayed with the structure of the C16-*N*-acetylcytseamine substrate (C16-NAC; yellow) bound to InhA and NAD+ (1BVR.pdb;(*23*)). In the C16-NAC complex, the active site loop (red surface) is ordered. The figures reproduced from (*202*).

The structures of TCS bound to ecFabI and InhA are very similar. However, there is an important difference in the conformation of the residues comprising the

substrate binding loop region that covers the active site of each enzyme (*63*). This loop (residues 195–200; Figure 27a, red loop) becomes ordered when TCS binds to ecFabI (*63*) or when FabI is complexed with a diazaborine inhibitor (*44*). As TCS brings about slow onset inhibition of ecFabI (*54, 55*), a plausible explanation for the slow step in TCS binding is the ordering of this loop. In contrast, the substrate-binding loop comprised of residues 197–211, is "disordered" in the InhA–TCS complex (Figure 27b, loop ends colored red), and TCS does not bring about slow onset inhibition of InhA (*58*). These data suggest that ordering of the FabI (InhA) active site loop is coupled to high-affinity enzyme inhibition. In support of this hypothesis, we note that the InhA loop is ordered in the structure of the INH–NAD adduct bound to InhA (*47*) and that the INH–NAD adduct brings about the slow onset inhibition of InhA with a K_i value of 0.75 nM (*42*).

The longer substrate-binding loop of InhA as compared to the ecFabI (*23, 47*), led to the speculation that larger substituents on the diphenyl ether A ring might result in additional contacts between TCS and residues comprising the InhA substrate-binding loop. The superposition of the structures of InhA in complex with TCS (Figure 27c, gray molecule) and hexadecenoyl-*N*-acetylcysteamine (C16-NAC; Figure 27c, yellow molecule; PDB code 1BVR) (*23*) further strengthened this hypothesis. It is to be noted that in the InhA–C16-NAC complex (Figure 27d, red surface), the substrate binding loop is ordered and the superposition of the two structures revealed that the thioester carbonyl group and

the first three carbons of the acyl chain of the C16-substrate superimpose accurately with half of the TCS A ring and the chlorine atom of ring A (Figure 27c, yellow and gray molecules, respectively). This led to the modification of TCS by removing the chlorine atoms and adding an acyl chain to the A ring carbon. This resulted in alkyl diphenyl ethers with varying length of alkyl chain (**Figure 28**). Also, by adding a methyl group at the ortho position on the B ring of the diphenyl ether, and a 6 carbon chain in the A ring resulted in another compound called PT70 (**Figure 29**), also in the diphenyl ether series.

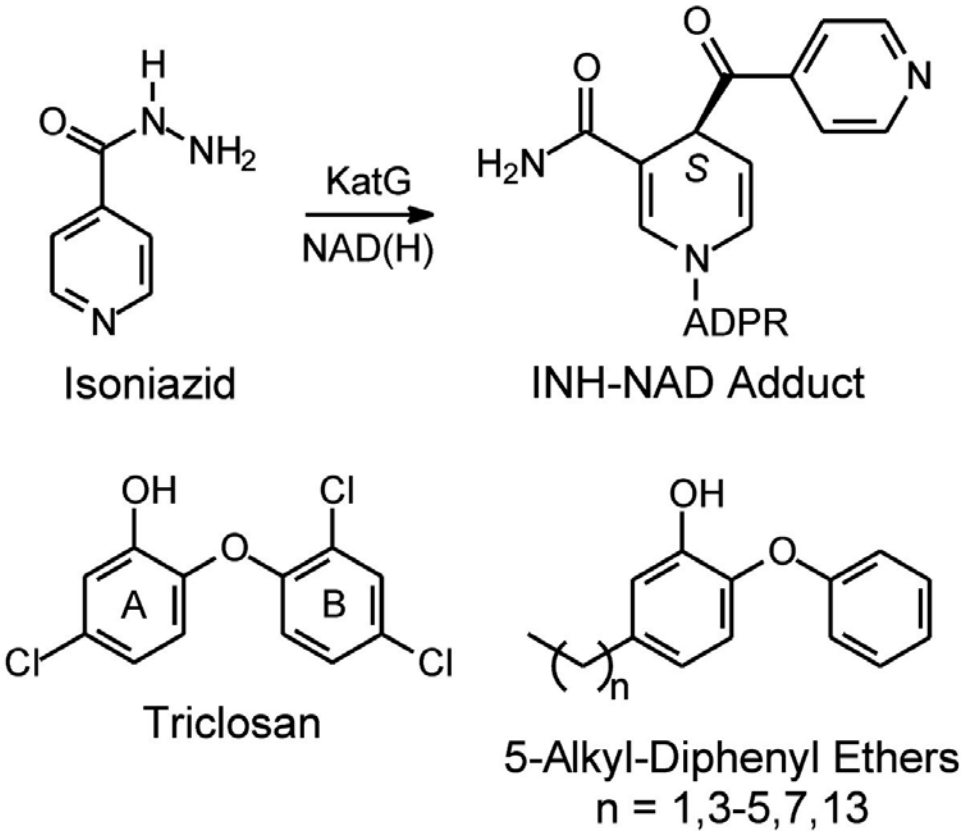


Figure 28: Enoyl reductase inhibitors. Figure reproduced from (*202*)

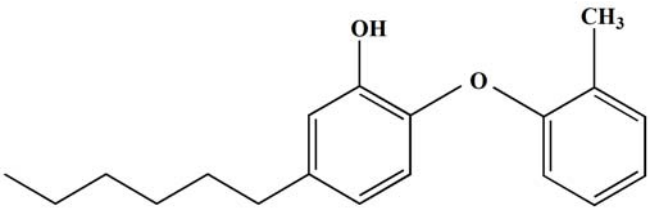


Figure 29: Inhibitor PT70

In order to understand the interaction of the alkyl diphenyl ethers with InhA, the structures of the enzyme complexes with the pentyl (5PP) and octyl (8PP) derivatives were solved to 2.8 and 2.6 Å resolutions, respectively. The rings of these inhibitors bind in a similar orientation as the TCS rings (Figure 31, gray molecules). Also, a similar hydrogen-bonding pattern is observed between inhibitor hydroxyl substituents and Tyr158 as well as the 2´-hydroxyl group of NAD^+ (Figure 31, black dotted lines). The 5PP alkyl chain is in a similar position in the active site as the alkyl chain of the C16 substrate. The pentyl chain forms predominantly hydrophobic interactions with residues Phe149, Met155, Tyr158, and Leu218. Superposition of the 8PP inhibitor structure with the C16 structure shows that the first four carbons of the octyl chain superimpose well with the C16 fatty-acyl chain. Unlike the C16 substrate, which bends and forms a U-shape within the active site, the octyl chain adopts a linear conformation and burrows into the protein, forming hydrophobic contacts with Phe149, Met155, Ala157, Pro156, Ile215, and Leu218. To accommodate the linear conformation, Leu218 is shifted by an average of 2 Å. The structure of PT70 bound to InhA was also solved at 1.8 Å resolution (unpublished data from the Kisker lab) and has an

ordered substrate binding loop conformation. Unlike the other diphenyl ether inhibitors binding to InhA, PT70 brings about slow onset inhibition.

A qualitative study of all the available crystal structures of InhA in the Protein Data Bank was performed (**Table 14**). The superposition of all these 27 structures on each other with the protein backbone fit on residues 5 to 185 reveals a very interesting overlap of the substrate binding loop region (**Figure 30**). 18 of these structures are from crystals with space group P 6 222. All these adopt the more closed loop conformation (**Figure 30**B). All the structures that adopt the comparatively open substrate binding loop conformation are not from the crystals with P 6 222 space group (**Figure 30**A). It is worth investigating if the crystal packing arrangements have anything to do with the loop conformations.

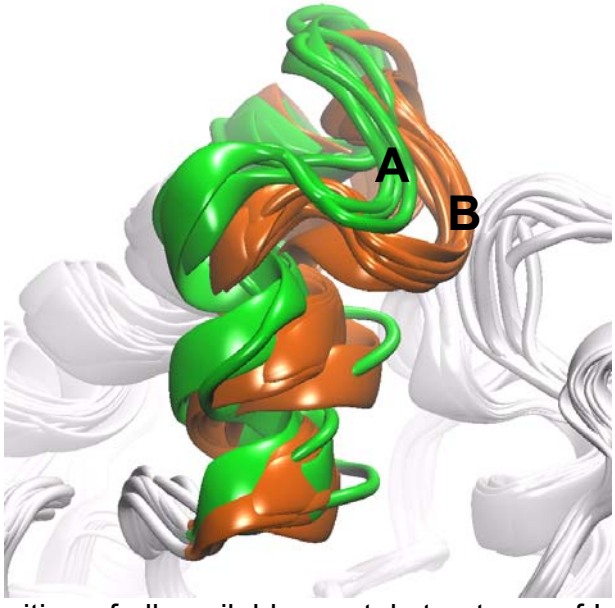


Figure 30: Superposition of all available crystal structures of InhA. Residues 197-215 are shown in color. All the structures from crystals with P 6 222 space group have a more closed loop conformation (B). None of the available crystal structures with a more “open” loop conformation (A) are belong to the P 6 222 space group.

Table 14: A tabulation of all available crystal structures of InhA

PDB ID	Structure Descriptor	Resolu-tion (Å)	Space Group
1BVR	*M. tuberculosis* Enoyl-ACP Reductase (INHA) in complex with NAD^+ AND C16-fatty-acyl-substrate	2.8	C 1 2 1
1E1F	Crystal Structure Of A Monocot (Maize Zmglu1) Beta-Glucosidase In Complex With P-Nitrophenyl-Beta-D-Thioglucoside	2.6	P 1 21 1
1ENY	Crystal Structure And Function Of The Isoniazid Target Of *M. Tuberculosis*	2.2	P 62 2 2
1ENZ	Crystal Structure And Function Of The Isoniazid Target Of *M. Tuberculosis*	2.7	P 62 2 2
1P44	Targeting tuberculosis and malaria through inhibition of enoyl reductase: compound activity and structural data	2.7	C 1 2 1
1P45	Targeting tuberculosis and malaria through inhibition of enoyl reductase: compound activity and structural data	2.6	I 21 21 21
1ZID	Long Fatty Acid Chain Enoyl-Acp Reductase (Inha) In Complex With An Isonicotinic-Acyl-Nadh Inhibitor	2.7	P 62 2 2
2AQ8	Crystal structure of wild-type of Enoyl-ACP(CoA) reductase from *M. tuberculosis* in complex with NADH.	1.92	P 62 2 2
2AQH	Crystal structure of Isoniazid-resistant I21V Enoyl-ACP(CoA) reductase mutant enzyme from *M. tuberculosis* in complex with NADH	2.01	P 62 2 2
2AQI	Crystal structure of Isoniazid-resistant I47T Enoyl-ACP(CoA) reductase mutant enzyme from *M. tuberculosis* in complex with NADH	2.2	P 62 2 2
2AQK	Crystal structure of Isoniazid-resistant S94A Enoyl-ACP(CoA) reductase mutant enzyme from *M. tuberculosis* in complex with NADH	2.3	P 62 2 2
2B35	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) inhibited by triclosan	2.3	C 1 2 1
2B36	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) inhibited by 5-pentyl-2-phenoxyphenol	2.8	C 1 2 1
2B37	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) inhibited by 5-octyl-2-phenoxyphenol	2.6	C 2 2 21
2H7I	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) complexed with 1-cyclohexyl-5-oxo-N-phenylpyrrolidine-3-carboxamide	1.62	P 62 2 2
2H7L	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) complexed with N-(3-bromophenyl)-1-cyclohexyl-5-oxopyrrolidine-3-carboxamide	1.73	P 62 2 2
2H7M	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) complexed with 1-cyclohexyl-N-(3,5-dichlorophenyl)-5-oxopyrrolidine-3-carboxamide	1.62	P 62 2 2
2H7N	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) complexed with N-(5-chloro-2-methylphenyl)-1-cyclohexyl-5-oxopyrrolidine-3-carboxamide	1.9	P 62 2 2
2H7P	Crystal structure of *M. tuberculosis* enoyl reductase (InhA) complexed with N-(3-chloro-2-methylphenyl)-1-cyclohexyl-5-oxopyrrolidine-3-carboxamide	1.86	P 62 2 2
2H9I	*M. tuberculosis* InhA bound with ETH-NAD adduct	2.2	P 62 2 2
2IDZ	Crystal structure of wild type Enoyl-ACP(CoA) reductase from *M. tuberculosis* in complex with NADH-INH	2	P 62 2 2
2IE0	Crystal Structure of Isoniazid-resistant I21V Enoyl-ACP(COA) Reductase Mutant Enzyme From *M. tuberculosis* in Complex with NADH-INH	2.2	P 62 2 2
2IEB	Crystal Structure of Isoniazid-resistant S94A ENOYL-ACP(COA) Reductase Mutant Enzyme from *M. tuberculosis* in Complex with NADH-INH	2.2	P 62 2 2
2IED	Crystal Structure Of Isoniazid-Resistant S94A Enoyl-Acp(Coa) Reductase Mutant Enzyme From *M. Tuberculosis* Uncomplexed	2.14	P 1
2NSD	Enoyl acyl carrier protein reductase InhA in complex with N-(4-methylbenzoyl)-4-benzylpiperidine	1.9	I 21 21 21
2NTJ	*M. tuberculosis* InhA bound with PTH-NAD adduct	2.5	I 21 21 21
2NTV	*M. leprae* InhA bound with PTH-NAD adduct	2.1	I 21 21 21
2NV6	*M. tuberculosis* InhA (S94A) bound with INH-NAD adduct	1.9	P 62 2 2
2PR2	Structure of *M. tuberculosis* enoyl-ACP reductase with bound INH-NADP.	2.5	P 62 2 2

4.4 Hypothesis

Similar to the InhA–TCS complex, the majority of the substrate-binding loop is disordered in both the pentyl and octyl ligand structures (**Figure 31**, loop ends colored red). These ligands (TCS, pentyl- and octyl- derivatives of TCS) are uncompetitive inhibitors of InhA. The crystal structure of InhA with PT70 (unpublished data) has an ordered loop conformation. It is to be noted that PT70 is slow onset inhibitor of InhA. This makes a compelling hypothesis that the ordering of the substrate binding loop can serve as the molecular explanation for the slow onset inhibition brought about in this enzyme. We note that the pentyl and octyl chains interact only with the very last residues of the loop, and further elaboration of the inhibitors is presumably required to cause loop ordering. We predict that compounds with this property will be slow, tight-binding InhA inhibitors with sub-nanomolar affinities for the enzyme. There is also the possibility that the crystal packing may affect the substrate binding loop conformation in InhA.

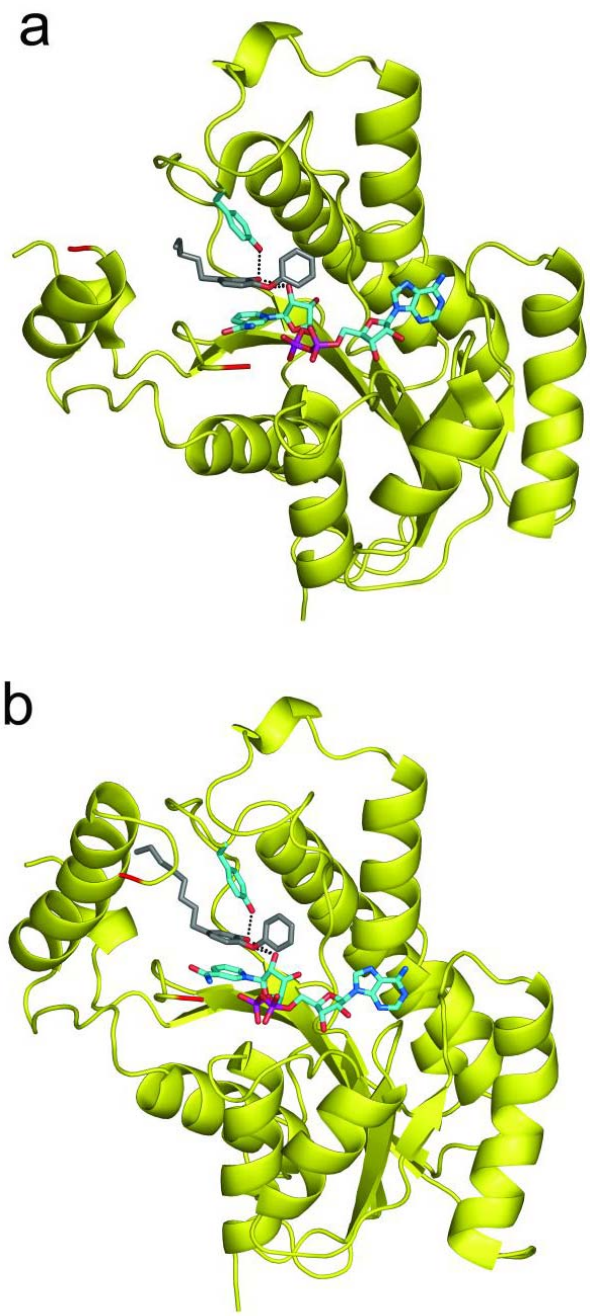


Figure 31: Structures of InhA in complex with the alkyl diphenyl ethers 5PP and 8PP. InhA in complex with NAD^+ (blue molecule) and a) 5-pentyl-2-phenoxyphenol (5PP, gray) and b) 5-octyl-2-phenoxyphenol (8PP, gray). The color scheme is the same as that used in Figure 1, panel b. In both subunits, the active site loops are disordered and the loop ends are colored red. Also shown are hydrogen bonds (black dots) between the inhibitor (gray) and Tyr158 (blue molecule) as well as the 2´-hydroxyl group of NAD^+ (blue molecule). The figures are reproduced from (*202*)

As the X-ray crystallography structures are time averaged structures, and the substrate binding loop ordering phenomenon requires understanding at the dynamic level, I have employed MD simulations. The use of molecular modeling and MD simulations will complement the X-ray crystallographic structures and will throw light on the dynamics of loop ordering and inhibitor binding to InhA. The results of this study will also help in predicting modifications to the inhibitors that will bind better to InhA.

4.5 MD simulations performed and the effect in Substrate binding loop ordering

In the table below (**Table 15**), the various MD simulations performed with different loop conformations are summarized. They are explained in detail in the following sections.

Table 15: Summary of the different MD simulations performed on the InhA system with different loop conformations

S. No.	InhA System	Inhibitor	Loop conformation (residues 196 – 217)	Duration of MD	Flexibility of the Loop during the simulation
1	Tetramer	TCS	From INH-NAD adduct bound InhA (1ZID.pdb)	5ns	Flexible all during the simulation
2	Tetramer	5PP	From INH-NAD adduct bound InhA (1ZID.pdb)	5ns	Flexible
3	Tetramer	5PP	From INH-NAD adduct bound InhA (1ZID.pdb)	62ns	Adopts partly ordered conformation
4	Tetramer	PT70	From PT70 bound InhA (unpublished data)	Run1: 6ns Run2: 6ns	Not Flexible Not Flexible
5	Tetramer	Triclosan	From PT70 bound InhA (unpublished data)	5ns	Flexible Flexible
6	Tetramer	5PP	From PT70 bound InhA (unpublished data)	8ns	Flexible Flexible
7	Tetramer	PT70->5PP	From PT70 bound InhA (unpublished data)	Run1: 5ns Run2: 5ns	Not Flexible Not Flexible
8	Fragment	5PP	Partly ordered loop conformation (from equilibrated structure of simulation #3)	5ns	Not flexible (stays in the partly ordered conformation)
9	Fragment	5OH	Partly ordered loop conformation (from equilibrated structure of simulation #3)	8ns	Not flexible (stays in the partly ordered conformation)

In order to understand how the binding free energy of different inhibitors binding to InhA is affected by the different loop conformations, MM-GBSA calculations were performed.

The crystal structures of both TCS bound InhA (pdb ID: 2B35) and 5PP bound InhA (pdb ID: 2B36) (*202*) are available. These structures are however not complete as the substrate binding loop region is missing in the electron density map. The missing number of residues from each monomer is different. Initially, in

order to complete the model, we used the then available complete X-ray structure of INH bound InhA (PDB ID: 1ZID) (*47*). Later, the availability of PT70 bound InhA structure, seemed to be a better template for the loop model as PT70 is a alkyl diphenyl ether that brings about slow onset inhibition of InhA. In our case, parametrizing an alkyl diphenyl ether was easier as we already have the parameters for TCS, another diphenyl ether. A comparison of the structures of PT70 bound InhA and INH-NAD adduct bound InhA shows both structures overlap very well with an rmsd of <0.5 Å over the protein backbone of residues 5-180. The only difference is that the loop with PT70 is more closed than that seen in INH-NAD adduct bound InhA (**Figure 32**). Hence we have 2 different loop conformations in InhA with slow onset inhibitors bound in the active site. Hence, in order to test the effect of the initial loop conformation on its dynamics and to understand the substrate binding loop ordering mechanism, I have performed MD simulations of InhA bound to TCS and 5PP with both the different loop conformations.

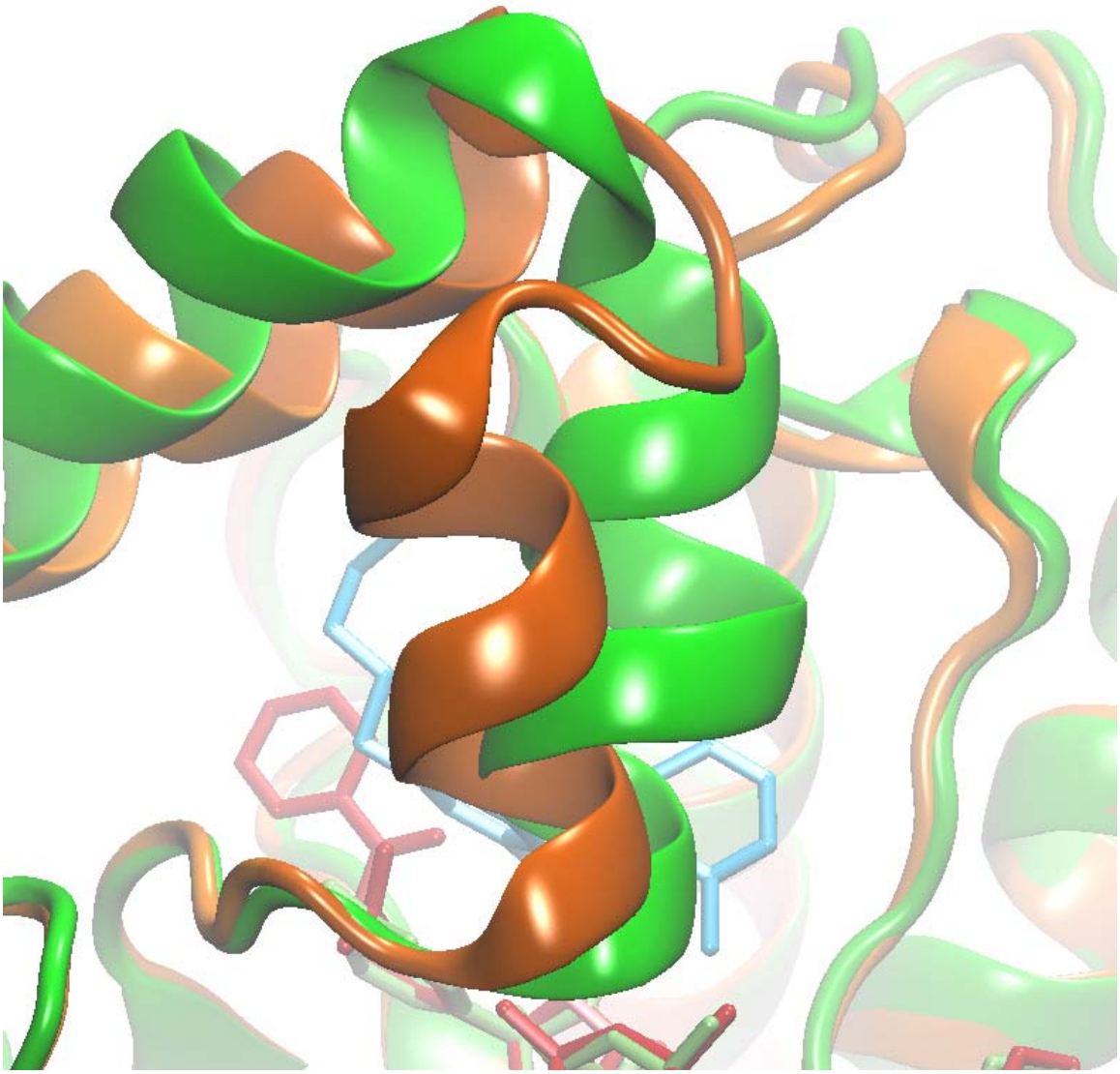


Figure 32: Superposition of PT70 bound InhA structure with INH-NAD adduct bound InhA structure. Orange : INH-NAD adduct bound InhA; Green: PT70 bound InhA. Red: INH-NAD adduct; Cyan: PT70; Lime: NAD^+

4.5.1 Simulations of InhA:TCS:NAD^+ tetramer system and InhA:5PP:NAD^+ tetramer system with INH-NAD adduct bound InhA loop conformation

4.5.1.1 System Setup

The crystal structure of the InhA/NAD^+/TCS complex (2B35) (*132, 202*) was used to build the starting structures for the ternary complex with the TCS. InhA is a tetramer in solution. Hence, a tetramer of InhA was used in our simulations. The substrate binding loop region was disordered in the X-ray crystal structure. The slow, tight binding inhibitor INH when bound to InhA, the loop is ordered in the X-ray crystal structure (unpublished data from the Kisker laboratory). This substrate binding loop (residues 196-216) conformation was used to build the missing substrate binding loops of the monomers. The InhA coordinates with the missing susbtrate binding loop residues were overlapped and the protein residues (5-185) were fit with the PT70 bound InhA co-ordinates. Both set of co-ordinates were saved. The missing residue coordinates were taken from those of PT70 bound InhA structure. The residue D147 was protonated in each of the InhA monomers. Missing atoms were built using the LEAP module of Amber.

Each of the systems was then solvated in a box of TIP3P water(*97*) with an 8Å buffer between the solute and box edge, resulting in a system with 19087

water molecules and ~73700 atoms in total. Counterions were not used in any calculations. In a similar manner, a system with InhA bound to 5PP was modeled with the loop conformation grafted from INH bound InhA X-ray structure. 5PP bound InhA tetramer system was solvated in a box of 19364 water molecules, making the system size to ~74600 atoms. For each system, a total of ~5ns of MD simulation at 300K, with a constant pressure of 1 atm., periodic boundary conditions and particle mesh Ewald(*105, 133*) treatment of electrostatics were performed with a time step of 1fs. Snapshots were saved every 1ps. The first 3ns of data were regarded as equilibration and not used in the binding affinity analysis.

4.5.1.2 MM-GBSA calculations

MM-GBSA was used to calculate the relative binding free energies of TCS and 5PP to the InhA. The details of this method have been presented earlier in Chapter 1. Briefly, the binding affinity for a protein-ligand complex corresponds to the free energy of association in solution as shown in equation below:

$$\Delta G_{bind} = G_{complex} - (G_{unbound\ protein} + G_{free\ ligand})$$

while the relative affinities for two ligands can be calculated using equation below:

$$\Delta\Delta G_{bind(1\rightarrow 2)} = \Delta G_{bind(2)} - \Delta G_{bind(1)}$$

In MM-GBSA, the binding affinity in the above equations is typically calculated using equation below:

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

where ΔE_{MM} represents the change in molecular mechanics potential energy upon formation of the complex, calculated using all bonded and non-bonded interactions. Solvation free energy, G_{solv}, is composed of the electrostatic component (G_{PB}) and a non-polar component (G_{np}):

$$\Delta G_{solv} = \Delta G_{GB} + \Delta G_{np}$$

G_{GB} was calculated using the sander module. The hydrophobic contribution to the solvation free energy, $G_{np,}$ was calculated using the solvent accessible surface area (SASA) (*140*) also from the sander module by using the parameter gbsa = 2.

$T\Delta S_{solute}$ represents the entropic contribution to binding affinity at temperature *T.* The ligands used in these calculations are TCS and 5PP. For a series of compounds with similar structures and binding modes, the entropy contribution can be omitted if one is only interested in relative binding affinities (*125, 142*). This may not be true for the alkyl chain of 5PP as it has multiple rotatable bonds, not present in TCS. Since this calculation converges slowly and can have large uncertainties, we omitted the entropic contribution to $\Delta\Delta$G.

4.5.1.3 Results

From the simulations of the InhA:TCS:NAD^+ tetramer system, it was seen that the backbone rmsd of the all the residues except for the ones from the loop region is < 1.5 Å (data not shown). The loop region residues move away from the conformation that they were built in. This is not uncommon as the loop conformation was taken from the INH bound InhA structure (pdb ID:1ZID) (*47*).

The fluctuation plot shows that the loop region remains flexible during most part of the simulation (Figure 33). The purpose of this plot is to find out which residue fluctuates most during the MD simulation.

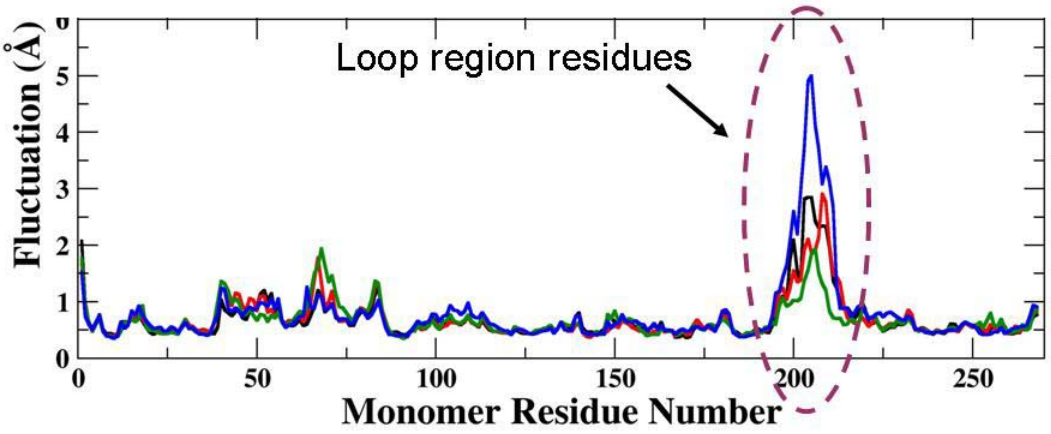


Figure 33: Fluctuation Plot: The loop region of the monomers of TCS bound InhA shows highest flexibility. The loop region was originally grafted from the INH-NAD adduct bound InhA crystal structure. Black, red, green and blue respective stand for fluctuations in monomers 1, 2, 3 and 4.

The MM-GBSA method was used to calculate the binding energy of TCS to InhA. The tetramer trajectory was split into the individual monomers and the energies calculated. The following results were obtained (Table 16). The average energy is -33.4 kcal/mol.

Table 16: Binding energies of TCS to InhA calculated using MM-GBSA approach

Monomer System	InhA/TCS/NAD$^+$ $E_{MM}+E_{GB}+E_{SA}$ (kcal/mol)
Monomer1	-34.2
Monomer2	-32.6
Monomer3	-36.0
Monomer4	-30.7

Similar approach was followed with the 5PP bound InhA simulations. In these simulations too, the backbone rmsd of all the residues remained below 1.5Å, but the substrate binding loop remained flexible, as seen in the fluctuations plot (Figure 34).

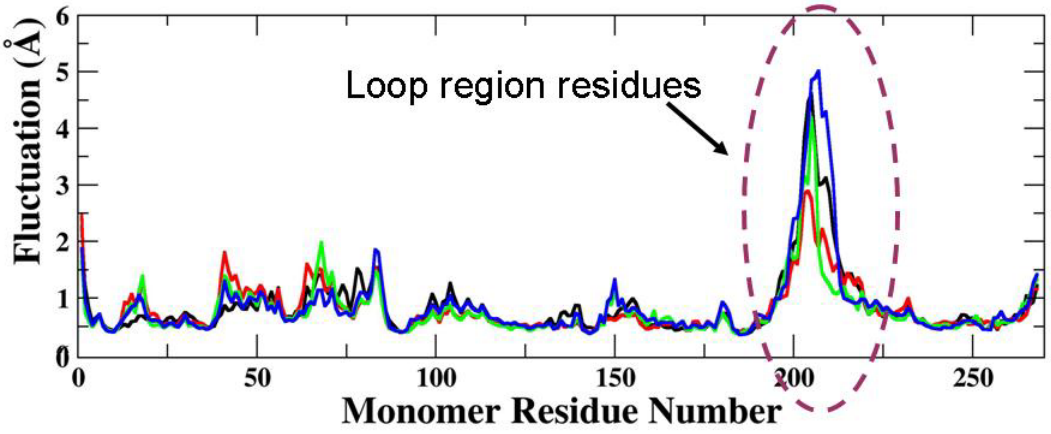


Figure 34: Fluctuation Plot: The loop region of the monomers of 5PP bound InhA shows highest flexibility. The loop region was originally grafted from the INH-NAD adduct bound InhA crystal structure. Black, red, green and blue respective stand for fluctuations in monomers 1, 2, 3 and 4.

The MM-GBSA method was used to calculate the binding energy of 5PP to InhA. The tetramer trajectory was split into the individual monomers and the energies calculated. The following results were obtained (Table 17). The average energy is -37.4 kcal/mol.

Table 17: Binding Energy of 5PP to InhA using MM-GBSA approach

Monomer System	InhA/5PP/NAD$^+$ $E_{MM}+E_{GB}+E_{SA}$ (kcal/mol)
Monomer1	-42.0
Monomer2	-43.3
Monomer3	-31.4
Monomer4	-33.1

From the above results, it is clear that the MM-GBSA approach is able to reproduce the correct trend in binding affinities, with inhibitor 5PP binding better than TCS. It is interesting to note that 2 of the 4 monomers in the 5PP bound InhA simulations (monomer 1 and monomer 2) have significantly lower energies viz., -42.0 kcal/mol and -43.3 kcal/mol as compared to the other two (-31.4 kcal/mol and -33 kcal/mol). This leads us to the question that when all the 4 protein monomers have the same sequence and structure, then why does the binding energies to each of the monomers differ? This led to the investigation of the loop conformations in the 5PP bound InhA. Surprisingly, the monomers with the worst energies (highest) had either most “ordered” or most “disordered” conformations as compared to the initial conformation (ie as seen in 1ZID.pdb). In **Figure 35** are shown the substrate binding loops of the 4 monomers and that in 1ZID.pdb (orange). The susbtrate binding loop monomers 3 (iceblue color) maintains its initial built in conformation and that monomer 4 (dark blue) is

completely disordered. The binding free energy calculations reveal that these 2 monomers have the highest binding free energy viz., -31.4 kcal/mol and -33 kcal/mol respectively. The substrate binding loop of monomers 1 (pink) and monomer 2 (red) do not stay in their initially modeled conformation (1ZID.pdb loop conformation – orange (*47*)), and are not completely disordered as seen in monomer 4 (dark blue). They adopt a partially ordered conformation. The binding free energy calculations show that these have best (lowest) free energies (-42.0 kcal/mol and -43.3 kcal/mol respectively). Thus, clearly, there is a correlation between the substrate binding loop ordering and the binding free energy of the inhibitor.

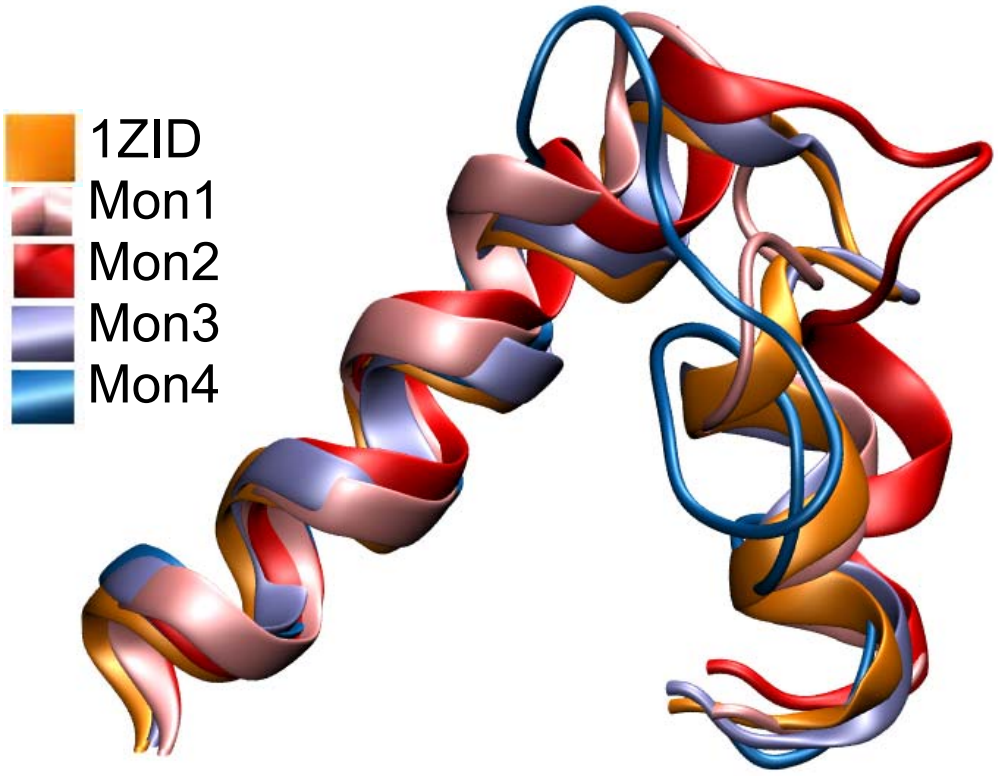


Figure 35: Superposition of the substrate binding loops of the different monomers of the 5PP bound InhA tetramer on the 1ZID (*47*) substrate binding loop conformation

The study involving substrate binding loop conformation will not be relevant if the inhibitor is not taken into account. The position of the 5PP was carefully studied in the active site of InhA. It was seen that the inhibitor position had changed from the position they were modeled in, or from their conformation in the crystal structure (Figure 36). The left panel of Figure 36 shows the 5PP in the active site of different monomers of InhA. Clearly, they are not in the same conformation in all the monomers.

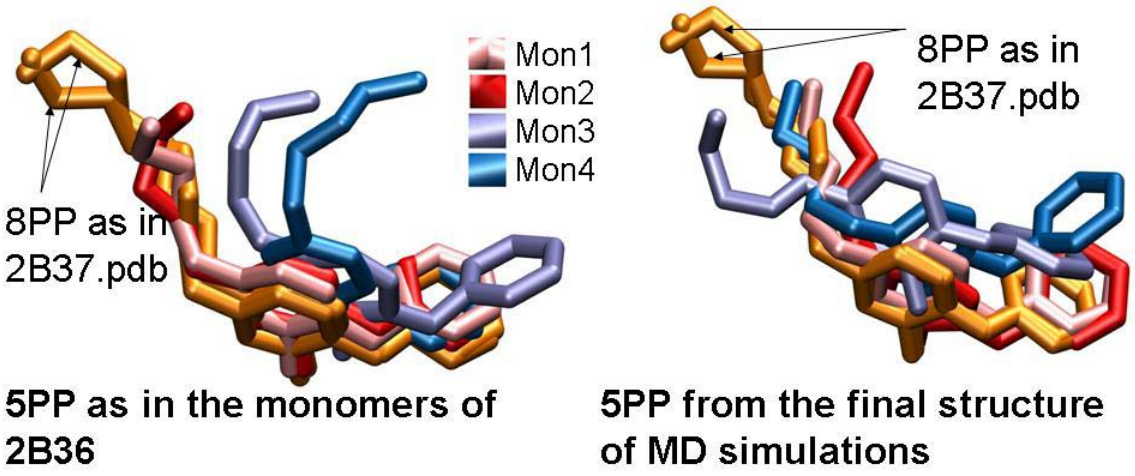


Figure 36: Left: Comparison of the conformation of inhibitor 5PP to 8PP (orange) in the crystal structures 2B36 and 2B37. Overlap produced by fitting protein backbone residues 5 to 185 of each monomer. Right: Comparison of the conformation of inhibitor 5PP from the MD simulations to 8PP (orange) in the crystal structure 2B37. Overlap produced by fitting protein backbone residues 5 to 185 of each monomer.

Though the substrate binding loop conformation for all the 4 monomers was alike initially (taken from 1ZID.pdb), the inhibitor conformation in their respective active sites was not similar (**Figure 36** left panel). The interaction of the inhibitor with the loop will differ based on its initial conformation. As seen in

Figure 36 (right panel), 5PP, during the course of the MD simulation, do change from their initial conformation. This difference may have very well led to the differences in the conformation of their respective substrate binding loops. Thus, the different susbtrate binding loop conformations or 5PP conformations are not an artifact of MD simulations. To understand the loop dynamics, an addition 62 ns of explicit solvent MD simulation was performed on the 5PP bound InhA tetramer system. The conformation of the loop was analyzed using 2D-RMSD plots (Figure 37). The data for this was generated using MOIL-VIEW software. This type of analysis helps us to see if a particular conformation is revisited during the course of the MD simulations. In Figure 37, lighter areas away from the diagonal indicate that the conformations are revisited. In monomer 1 the initial loop conformation is not revisited during the initial part of the simulation and only to very little extent in the very end of the simulation. From the 2D-RMSD plot of monomer 2, we see that though the initial conformation is changed, the new conformation stays stable during most part of the MD simulation. We see that monomer 3 loop conformation is the most stable in the sense that its conformation in the beginning of the MD simulations is revisited repeatedly during the entire MD simulation. Upon analysis, as seen in **Figure 35**, the substrate binding loop of monomer 3 (iceblue color) keeps the initially modeled loop conformation. The initial conformation of monomer 4's loop is completely lost and is not revisited during the course of the MD simulation. This "ordered" conformation of the loop in monomer 3 comes with a "displaced" inhibitor 5PP. However, monomer 2 loop conformation is different from the initial conformation,

but is stable and also has the inhibitor in place. This monomer also had the most favorable binding free energy to InhA. Thus, monomer 2 substrate binding loop conformation was studied in more detail.

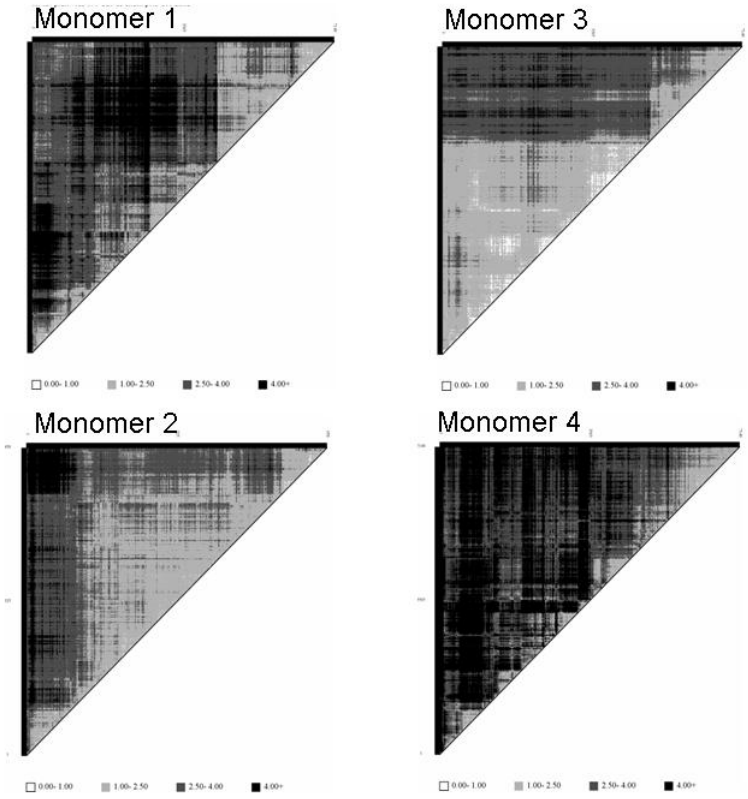


Figure 37: 2D-RMSD Plots: (Fit to protein) Loop RMSD

The residues 195-201 and 207-223 of the substrate binding loop of 1ZID.pdb adopt a α-helical conformation. The residues 196-201 and residues 217-224 of the substrate binding loop of monomer 2 adopt a helical conformation. This can be considered as an intermediate conformation with between the one observed in INH bound InhA (PDB ID:1ZID) (*47*). and the completely disordered one as seen in monomer 4.

The active site of monomer 2 at a random snapshot in the production phase run of MD simulations, showed an interesting arrangement of a water molecule (Figure 38). There was a water molecule in the active site which was forming hydrogen bonds with I193 and with T195.

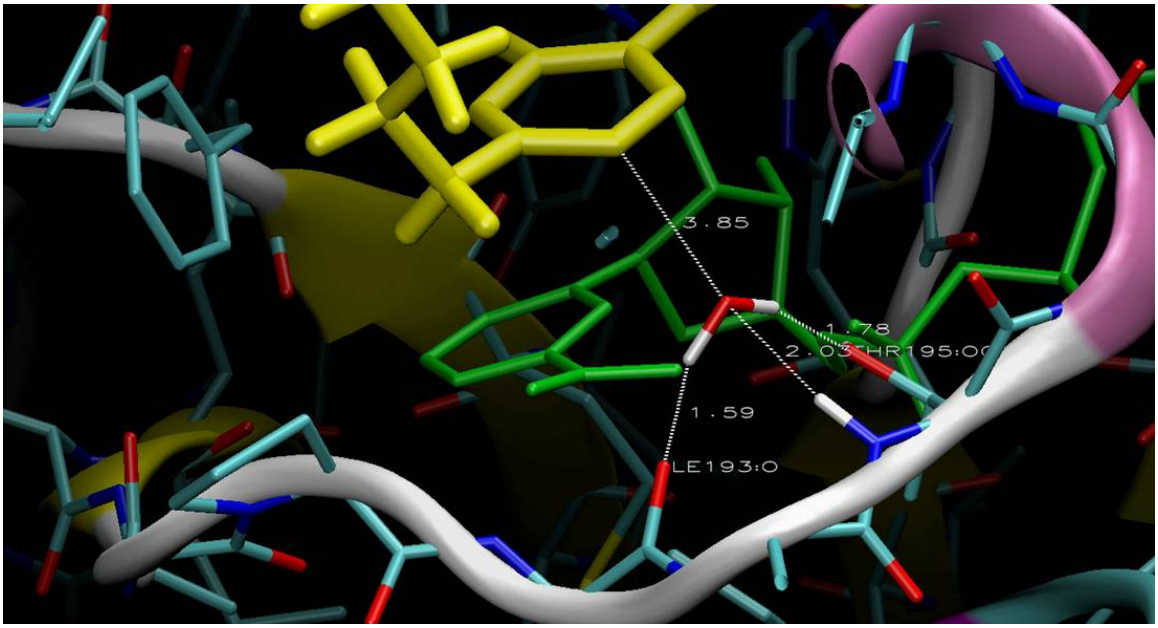


Figure 38: The active site of monomer 2 in InhA:5PP:NAD^+ after MD simulation showing a water molecule forming hydrogen bonding contacts with the substrate binding loop residues. In yellow, is the inhibitor 5PP, in green is the co-factor NAD^+ (a production phase MD snapsot)

This is a very interesting observation. This can be potentially exploited in our drug design efforts. The inhibitor could be modified with a hydrogen-bond forming group. This can replace the hydrogen bonds formed by this water molecule and can increase the solubility of the inhibitor (Discussed later in this Chapter).

4.5.2 Simulations of PT70 bound to InhA

PT70 is the only alkyl diphenyl ether exhibiting slow tight binding inhibition to the *M. tuberculosis* enoyl reductase InhA. Thus, to test our hypothesis, it can prove as an excellent control to see if the loop remains stable in the "ordered" conformation during the MD simulations. Also, this can throw light on the dynamics and contacts involved in slow onset inhibition. In order to investigate if the loop movement observed before is a random event or is it structured to do so, I performed MD simulations on the PT70 bound InhA tetramer system.

4.5.2.1 System Setup

This is our control system as it is taken directly from the X-ray structure solved by collaborator Dr. Kisker (unpublished data). Added hydrogens and solvated the system with 8 Å TIP3P water box resulting in a total of ~78600 atoms. The parameters for PT70 were derived in the same way as for TCS. Counterions were not used in any calculations. A total of ~6 ns of MD simulation at 300K, with a constant pressure of 1 atm., periodic boundary conditions and particle mesh Ewald (*105, 133*) treatment of electrostatics was performed with a time step of 1fs. Snapshots were saved every 1ps yielding a total of 6000 frames.

The first 3000 ps of data were regarded as equilibration and not used in the binding affinity analysis.

4.5.2.2 Results

In order to see if our results are reproducible, two MD runs were performed with all the same conditions and starting structures. When fluctuation of the residues was plotted against time, it was seen that the fluctuation for the loop region was very low (Figure 39).

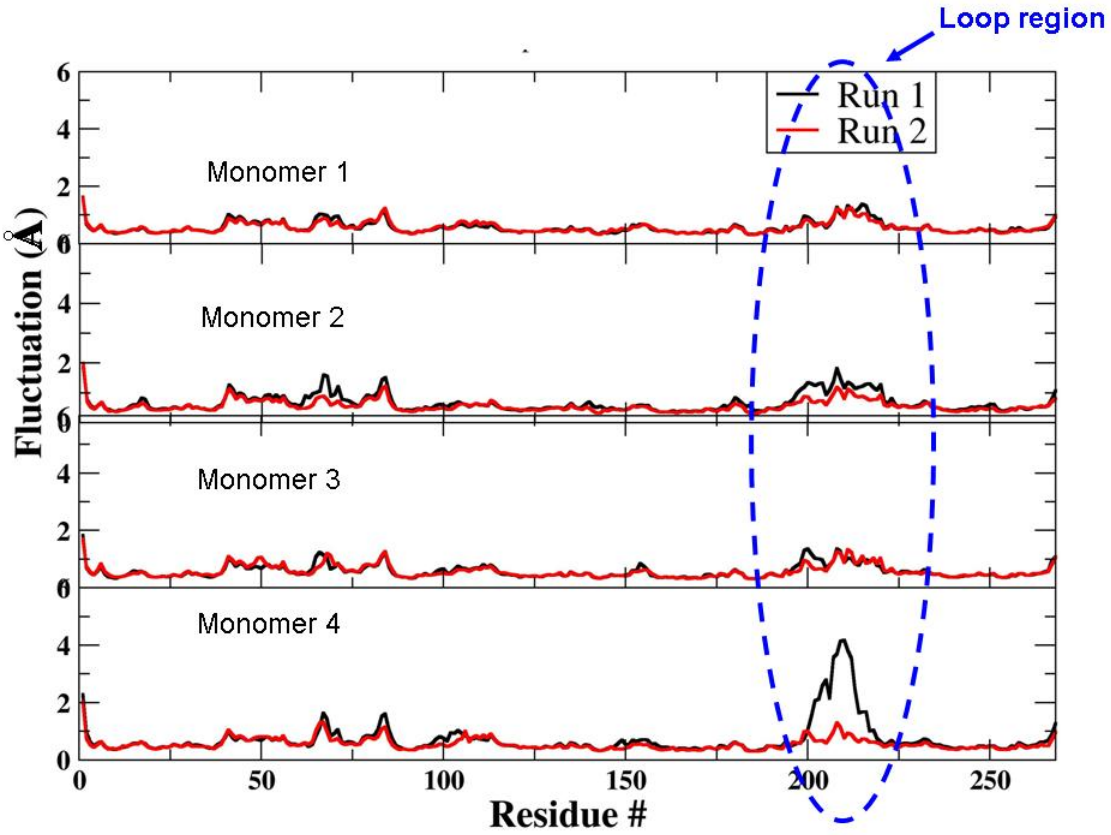


Figure 39: Fluctuation vs. Residue # of the 4 monomers of InhA:PT70:NAD^+ system: This plot represents the fluctuations of the residues during the course of the MD simulation.

Compared to the results of the TCS and 5PP bound InhA simulations, it can be seen that the loop region flexibility is <1 Å when a slow, tight binding inhibitor like PT70 is bound to InhA. This is inline with our hypothesis that ordering of the substrate binding loop region is correlated to the slow, tight binding of the inhibitor to the enzyme.

The last 1ns of this simulation data was used to calculate the binding free energy of PT70 binding to InhA using the MM-GBSA approach. It was calculated to be ~-43.6 kcal/mol (Table 18). Though this number may not be accurate, the results are in agreement with the PT70 binding better than 5PP binding better than TCS.

Table 18: Table showing the binding free energies of different monomers of InhA:PT70:NAD^+ system calculated using MM-GBSA approach

System	ΔMME	ΔVDW	ΔEEL	ΔEGB	ΔEEL+ΔEGB	ΔSA	ΔG
Mon 1	-67.8±3.2	-49.6±2.5	-18.2±2.7	25.4±1.6	7.1±2.1	-4.1±0.1	-46.5±2.9
Mon 2	-65.1±3.3	-46.9±2.5	-18.1±2.8	27.1±2.2	9.0±2.2	-4.2±0.1	-42.2±2.6
Mon 3	-63.0±3.1	-45.2±2.5	-17.8±2.2	27.3±1.7	9.6±1.9	-4.2±0.1	-39.9±3.3
Mon 4	-58.5±3.2	-41.2±2.3	-17.3±2.5	14.1±1.6	-3.2±1.7	-1.1±0.1	-45.6±2.8

4.5.3 Simulations on the InhA:TCS:NAD^+ tetramer system and InhA:5PP:NAD^+ tetramer system with PT70 bound InhA loop conformation

4.5.3.1 System Preparation

The crystal structure of the InhA/NAD^+/TCS complex (2B35) (*132, 202*) was used to build the starting structures for the ternary complex with the TCS. InhA is a tetramer in solution. Hence, a tetramer of InhA was used in our simulations. The substrate binding loop region was disordered in the X-ray crystal structure. The slow, tight binding inhibitor PT70 when bound to InhA, the loop is ordered in the X-ray crystal structure (unpublished data). This substrate binding loop (residues 196-216) conformation was used to build the missing substrate binding loops of the monomers. Missing atoms were built using the LEAP module of Amber. Standard Amber ff99SB forcefield parameters were assigned to the protein. TCS and 5PP were parameterized as follows. The initial geometry of TCS was obtained from the crystal structure of the InhA/NAD^+/TCS complex (PDB ID: 2B35 (*132, 202*)); The TCS parameters generated for the simulations of FabI/triclsoan/NAD^+ system were used in these simulations too. 5PP parameters were generated as described for TCS in chapter 2.

Each of the systems was then solvated in a box of TIP3P water(*97*) with an 8Å buffer between the solute and box edge, resulting in a system with 19087 water molecules and ~73700 atoms in total. Counterions were not used in any calculations.

In a similar manner, a system with InhA bound to 5PP was modeled with the loop conformation grafted from PT70 bound InhA X-ray structure. 5PP bound InhA tetramer system was solvated in a box of 19364 water molecules, making the system size to ~74600 atoms. For each system, a total of ~5ns of MD simulation at 300K, with a constant pressure of 1 atm., periodic boundary conditions and particle mesh Ewald (*105, 133*) treatment of electrostatics were performed with a time step of 1fs. Snapshots were saved every 1ps. The first 3ns of data were regarded as equilibration and not used in the binding affinity analysis.

4.5.3.2 Results

The simulations with InhA bound to TCS shows that during the entire duration of the simulation, the protein remained stable with a backbone rmsd of <1.5 Å for the residues not comprising the loop (Figure 40).

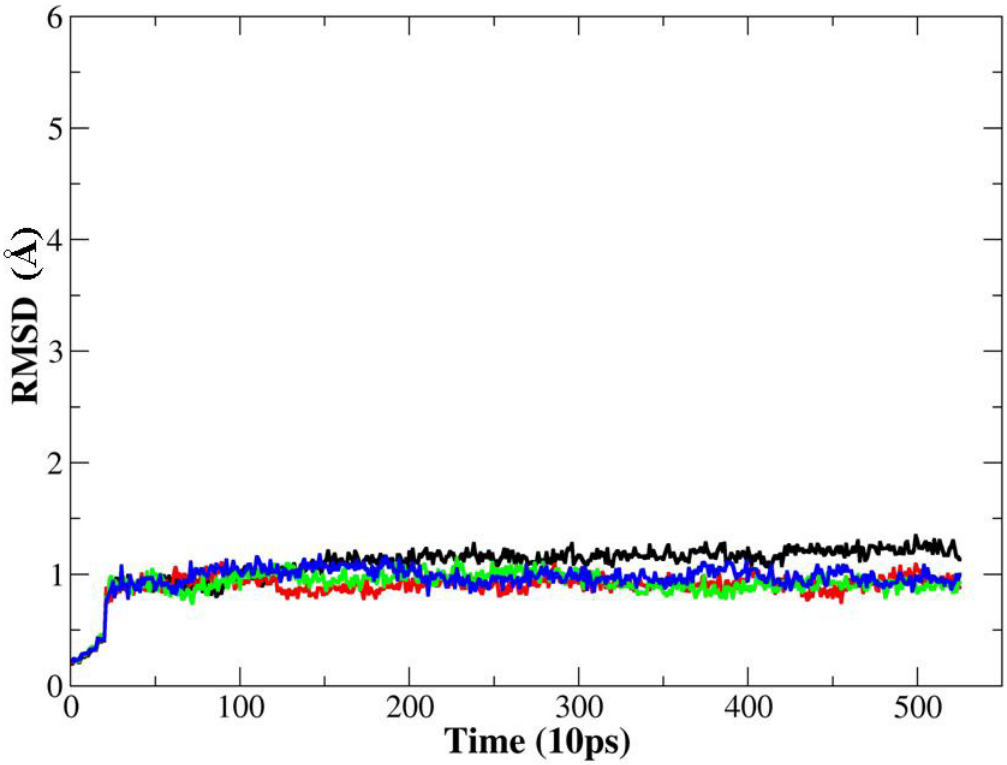


Figure 40: Rmsd vs Time Plot for the 4 monomers of the InhA:TCS:NAD^+ system MD simulation with InhA:PT70:NAD^+ structure loop conformation. The RMSD is backbone rmsd of the all the residues except the loop region residues

The X-ray structure of TCS bound to InhA (*202*) has the substrate binding loop region residues (residues 195-211) to be disordered. This could possibly be due to flexibility of the loop region. Consistent with this observation, the loop region of InhA:TCS:NAD^+ system, showed high fluctuations compared to the other regions of the protein as evident from the fluctuation plot (Figure 41).

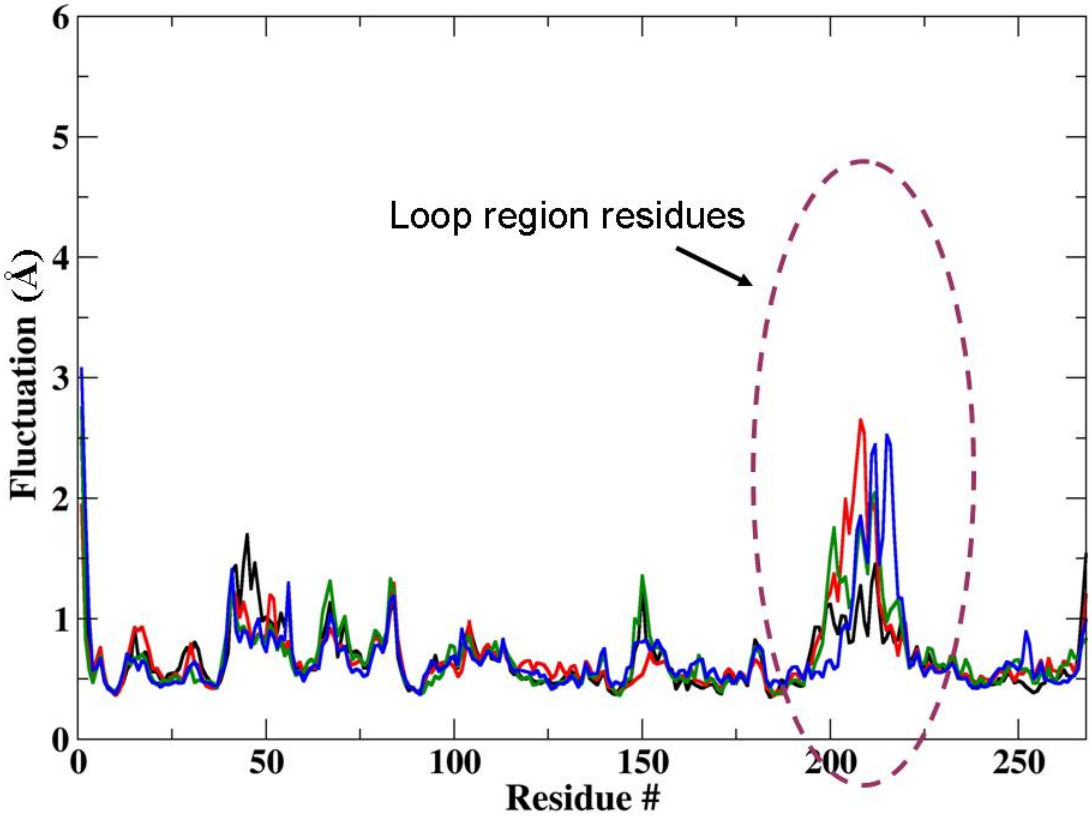


Figure 41: Fluctuation Plot for the 4 monomers of InhA:TCS:NAD^+ system with InhA:PT70:NAD^+ structure loop conformation: Fluctuation vs Residue # showing that the residues comprising the loop region fluctuate more during the course of the MD simulations

Similar to the observation in the TCS bound InhA simulations, the simulations with 5PP bound to InhA showed that the RMSD of the non-loop residues remained <1.5 Å (data not shown) and the fluctuations plot (**Figure 42**) showed that the loop region was very flexible.

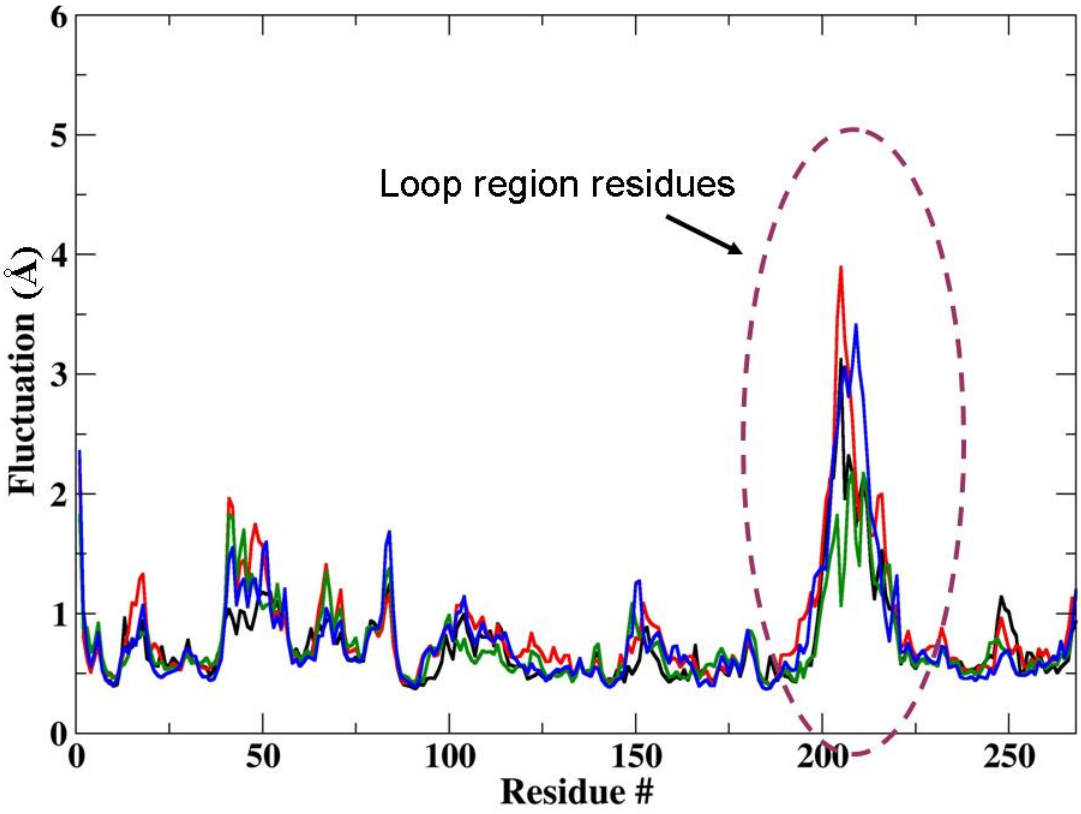


Figure 42: Fluctuation Plot for the 4 monomers of InhA:5PP:NAD^+ system with InhA:PT70:NAD^+ structure loop conformation: Fluctuation vs Residue # showing that the residues comprising the loop region fluctuate more during the course of the MD simulations

These results consistent with my previous simulations of the InhA system with bound TCS and 5PP but with the loop conformation as seen in INH bound InhA.

4.5.4 Simulations with PT70->5PP bound InhA

When a PT70 is bound to InhA, the loop remains ordered. Also, if 5PP is bound to InhA the substrate binding loop is disordered. Does replacing PT70 with 5PP “disorder” the loop? This would provide great insight into the mechanism of substrate binding loop ordering. In order to answer this question, I set up simulations on the PT70 bound InhA X-ray structure, except that I modified the inhibitor to be 5PP.

4.5.4.1 System Setup

The coordinates for all the residues constituting this system came from the PT70 bound InhA X-ray crystal structure, except that the inhibitor PT70 was manually modified into 5PP by deleting one carbon from the alkyl chain on the A ring and deleting the methyl group on the B ring of the diphenyl ether. Missing atoms were built using the LEAP module of Amber. Standard Amber ff99SB forcefield parameters were assigned to the protein. The 5PP parameters generated for the previous simulations of InhA/5PP/NAD^+ system were used in these simulations too.

The system was then solvated in a box of TIP3P water (*97*) with an 8Å buffer between the solute and box edge, resulting in ~75000 atoms in total. Counterions were not used in any calculations. For each system, a total of ~5ns of MD simulation at 300K, with a constant pressure of 1 atm., periodic boundary conditions and particle mesh Ewald (*105, 133*) treatment of electrostatics were performed with a time step of 1fs. Snapshots were saved every 10 ps. The first 3ns of data were regarded as equilibration.

4.5.4.2 Results

The simulations with InhA bound to PT70->5PP shows that during the entire duration of the simulation, the protein remained stable with a backbone rmsd of <1.5 Å for the residues not comprising the loop (data not shown). Unexpectedly, the substrate binding loop remained ordered. If indeed loop ordering is related to slow onset inhibition, then replacement of PT70 with an analog (5PP) that is only a rapid reversible inhibitor of the enzyme should have caused the loop to become disordered. One explanation for the observation that the loop remained ordered following introduction of 5PP may be that loop ordering and disordering is a kinetically slow process and the MD simulations were simply not long enough to encompass the loop disordering process. (**Figure 43**).

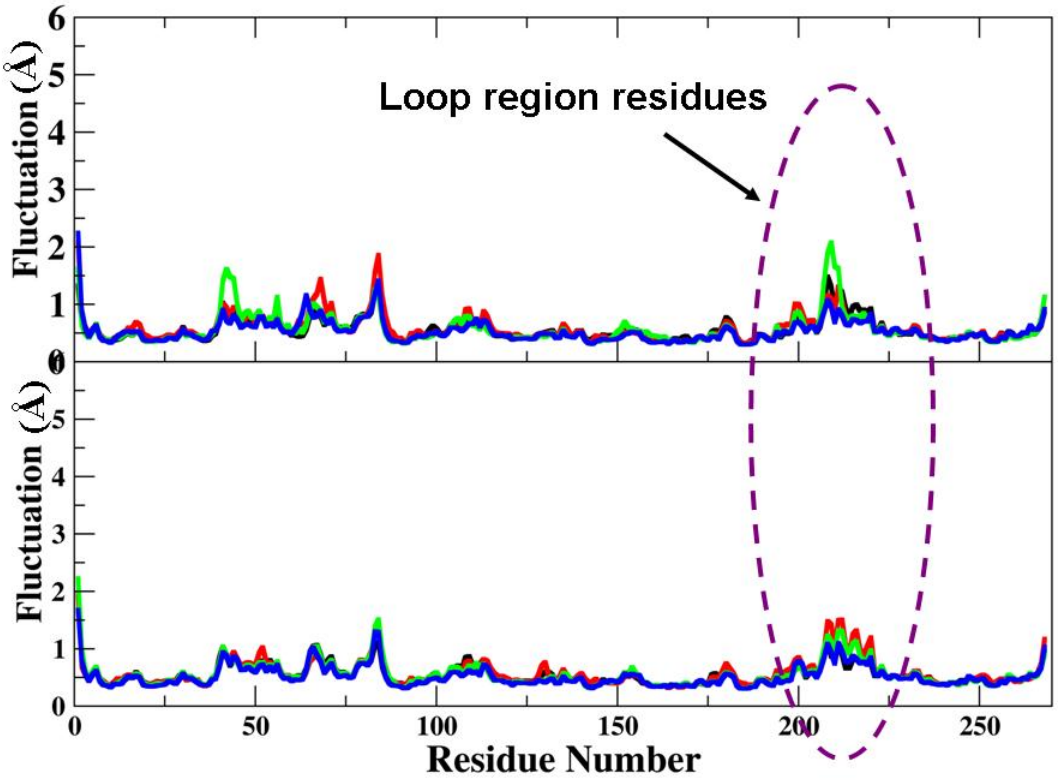


Figure 43: Residue Fluctuation Plot for the 2 separate runs of InhA tetramer system with 5PP (modified PT70) and NAD^+ bound in the active site of each monomer

If the inhibitor directs the loop ordering, then if 5PP is bound in the active site which is a uncompetitive inhibitor of InhA, the loop is expected to be flexible. A reasonable explanation for this observation could be that as the starting structure is at a global minima and the 5PP binding is not able to displace the loop from its conformation. Further analysis on the structure is required in order to completely understand this observation.

4.6 Effect of crystal packing contacts on the substrate binding loop ordering in M.tuberculosis enoyl reductase enzyme InhA

The crystal structure of the enoyl reductases, ecFabI and InhA are interesting in that the substrate binding loop density is not seen until an inhibitor that brings about slow onset inhibition is bound to the enzyme. The substrate binding loop conformation in the crystal structures (PDB IDs: 1ZID (*47*), 2B35, 2B36 (*202*), PT70 bound InhA (unpublished data)) can also be influenced by the crystal contacts (*205*). In the 5PP bound InhA structure (PDB ID: 2B36), it is seen that the different monomers in the same asymmetric unit have different number of substrate binding loop residues disordered. It is difficult to explain this until the crystal packing effects are taken into consideration. In this section, the crystal contacts and its influence on the inhA enzyme structure are explored.

The crystal structure of INH bound to InhA has a well defined substrate binding loop conformation. When the crystal lattice was reproduced using the Pymol software, it was seen that the substrate binding loop is in close contact with that of the neighboring InhA molecules (Figure 44). The acidic residues in the substrate binding loop are at a distance favorable for forming salt bridges with those in the substrate binding loop region of the neighboring molecule.

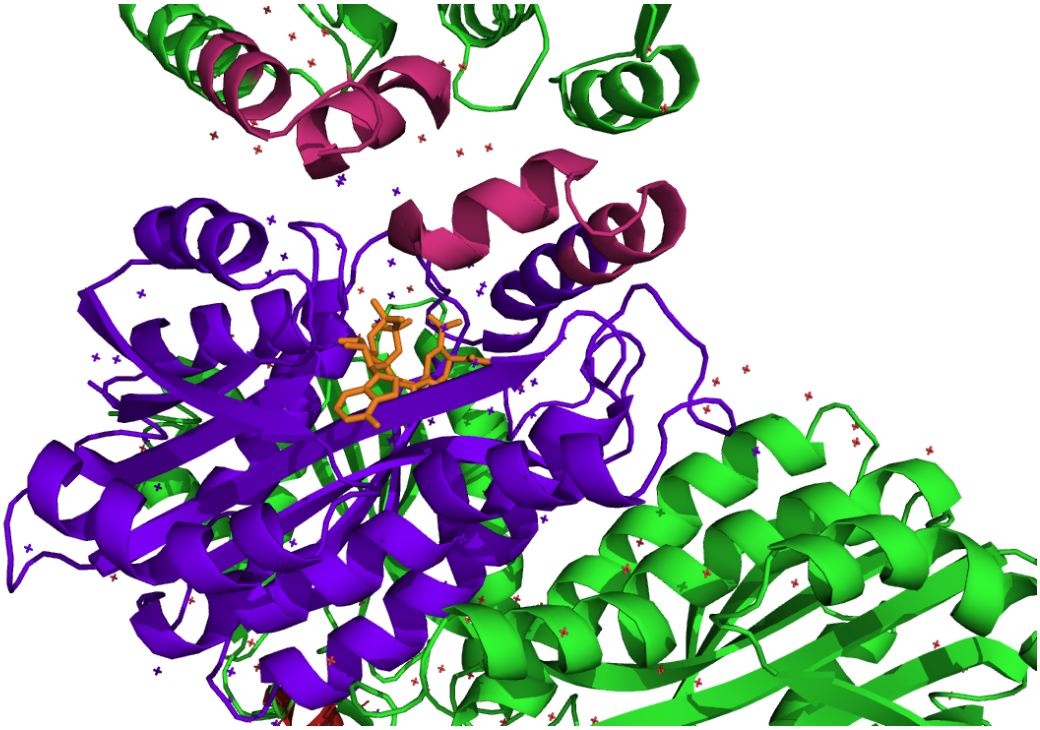


Figure 44: Crystal lattice for 1ZID.pdb using Pymol software. InhA in the asymmetric unit shown in violet. The neighboring InhA are shown in green. The substrate binding loop is shown in pink.

Close observation of the interface of the 2 molecules of InhA reveals that E210 is at salt bridge distance (3.07 Å) to R45 of the neighboring InhA molecule (Figure 45). This observation is important as it may be possible that the substrate binding loop ordering is not an effect of inhibitor binding, but just an artifact of crystal packing effects. In order to understand the effect of crystal packing on substrate binding loop ordering mechanism, I employed MD simulations.

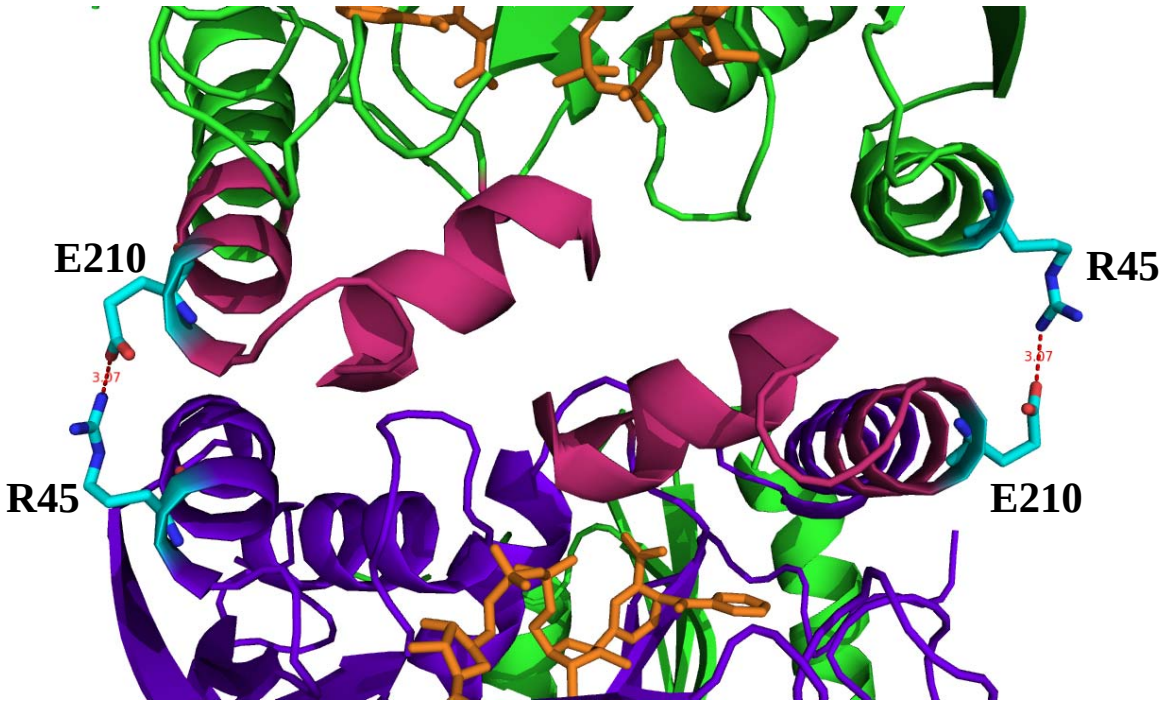


Figure 45: InhA molecules in crystal lattice produced from 1ZID.pdb. InhA from 1ZID.pdb is shown in violet, neighboring InhA molecules are shown in green. The substrate binding loop residues 195-215 are shown in pink. INH-NAD adduct is shown in orange.

4.6.1 System Preparation:

MD Simulations were set up with 2 monomers of InhA, packing against each other, as seen in the crystal lattice of 1ZID.pdb. The difference in the models was that in the active site, the INH-NAD adduct was replaced by the inhibitor 5PP and cofactor NAD^+. Counterions were not used in any calculations. After the minimization and equilibration, a short MD of 500 ps was performed at 300K, with a constant pressure of 1 atm., periodic boundary conditions and particle mesh Ewald (*105, 133*) treatment of electrostatics were performed with a time step of 1fs. Snapshots were saved every 1ps yielding a total of 500 frames.

Also, another simulation was simultaneously run on a similar system, but with just NAD^+ in the active site (no inhibitor) under same conditions.

4.6.2 Results:

It is seen from the simulations that in the system with 5PP in the active site, the loop continues to stay in the “ordered” conformation. The inhibitor 5PP is seen to move out of the active site and NAD^+ changes conformation. Under the same conditions, in the second system, the loop continues to stay in the “ordered” conformation. Unlike the previous system, here the NAD^+ has not changed the conformation drastically (Figure 46).

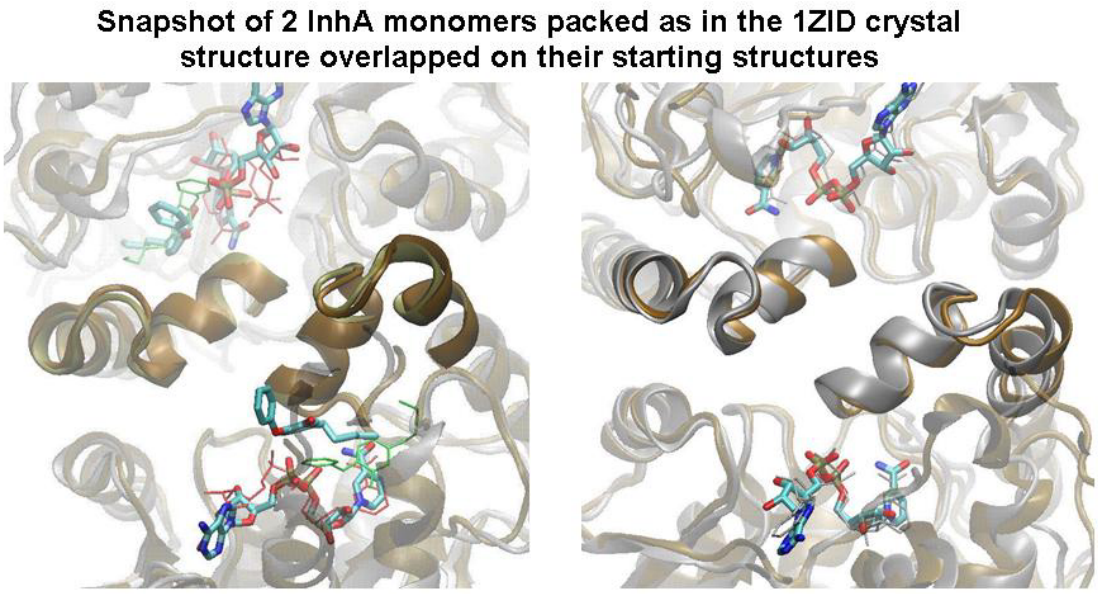


Figure 46: Overlap of 2 snapshots (tan) of the simulations of InhA with crystal contacts on their starting structures (silver). Left: system with 5PP and NAD^+ in the active site. Right: System with just NAD^+ in the active site.

4.6.3 Discussion:

The substrate binding loop stays in “ordered” conformation during the simulation even though the inhibitor 5PP is not a slow, tight binding inhibitor. This is explained by the fact that there is no space for the substrate binding loop to move or get “disordered” into. It may be possible that even with INH-NAD adduct bound to the active site of InhA, the substrate binding loop may be ordered due to crystal contacts and not due to the interactions of the INH-NAD adduct. However, simulations with PT70 bound to InhA demonstrated that the loop remained ordered even in the absence of the intermolecular contacts found in the

protein crystals. Thus, the ordered substrate binding loop as seen in the INH-NAD adduct bound InhA or with PT70 bound InhA X-ray structure, may be not the effect of crystal packing, but may cause the crystal to pack in that particular way.

4.7 Predicting modifications to the InhA inhibitor

From the analysis of simulations performed on the InhA:5PP:NAD^+ tetramer system, it was seen that the monomer 2 (with the lowest binding free energy) has a water molecule in the active site cavity forming hydrogen bonds to the substrate binding loop residues (**Figure 38**). In order to introduce an hydrogen-bond forming group, 5PP is modified into 5OH with a CH_2OH group at the para position to the hydroxyl group on the A-ring of the diphenyl ether (Figure 47). The parameters for this inhibitor were generated in a similar way as those for TCS and 5PP. The modifications in this inhibitor will enhance its interactions with the substrate binding loop and help it to stay “ordered”, thereby binding better with the enzyme InhA.

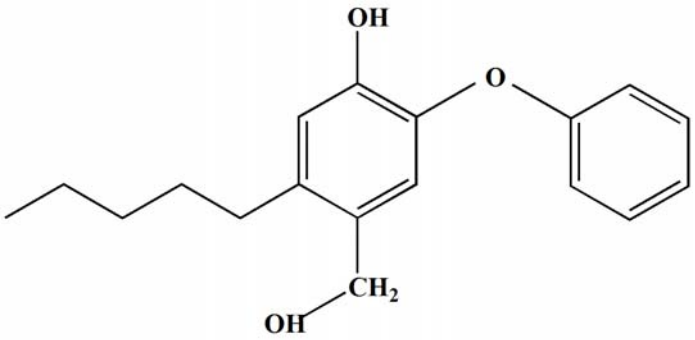


Figure 47: Modified InhA inhibitor, 5OH

4.7.1 Generation of InhA fragment

In order to reduce the total size of the inhA tetramer system, a downsized system was generated. As monomer 2 of the 5PP bound InhA tetramer system gave the most favorable energies and a reasonable substrate binding loop conformation, it was used for the InhA fragment generation. This system includes all the residues within 15 Å of the center of mass of the inhibitor and the co-factor NAD^+. The residues farther than 15 Å of the inhibitor and NAD^+, but part of the substrate binding loop were included in the fragment. Everything beyond 11 Å from the center of mass of the inhibitor and the cofactor were loosely restrained with a force constant of 0.5 kcal/mol/$Å^2$, except the residues constituting the substrate binding loop (Figure 48). Two such systems were generated, viz., one with 5PP in the active site and the other with the 5OH. Each of these systems includes 213 residues as opposed to 1080 residues in the tetramer system.

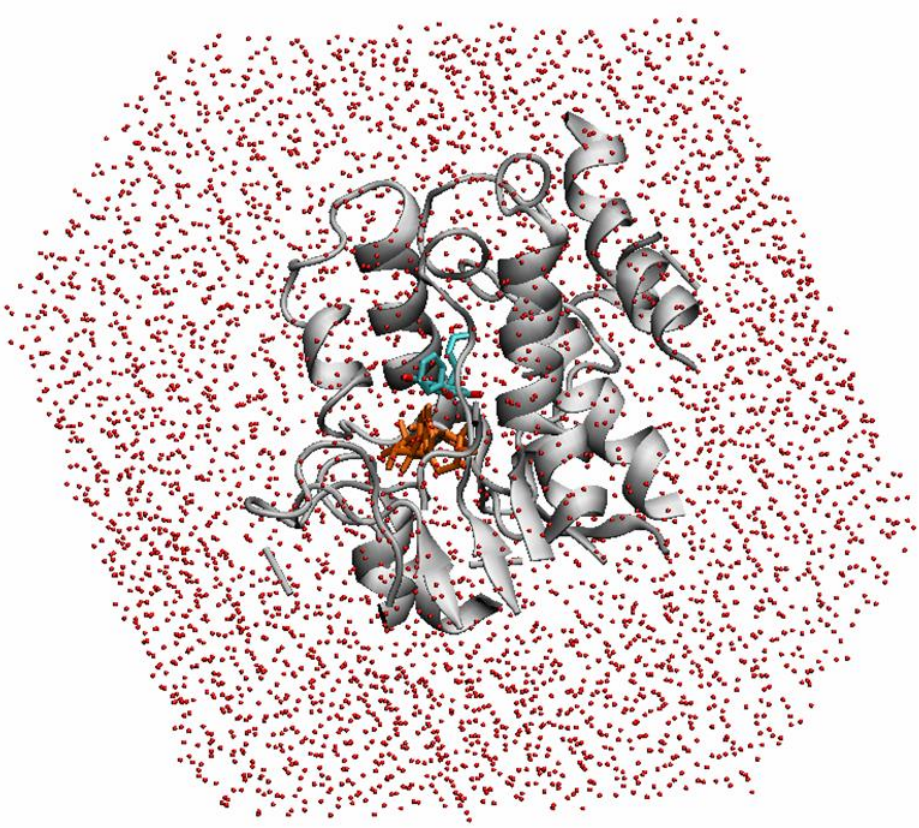


Figure 48: InhA fragment in explicit solvent

4.7.2 Results from Simulations with InhA:5PP:NAD^+ fragment system and InhA:5OH:NAD^+ fragment system

For each of the two systems, the minimization and equilibration was followed by production phase MD. A total of 8ns of explicit solvent MD simulations were performed. The initial 4ns of data were not considered for post processing. It was seen from the simulations that the overall backbone RMSD of the entire backbone was very low. This is not surprising as the protein beyond 11 Å was

restrained. The interesting part is that the loop region residues show low fluctuations (**Figure 49**). The fluctuations of the loop region for the InhA:5OH:NAD^+ system is slightly higher than for the InhA:5PP:NAD^+ system.

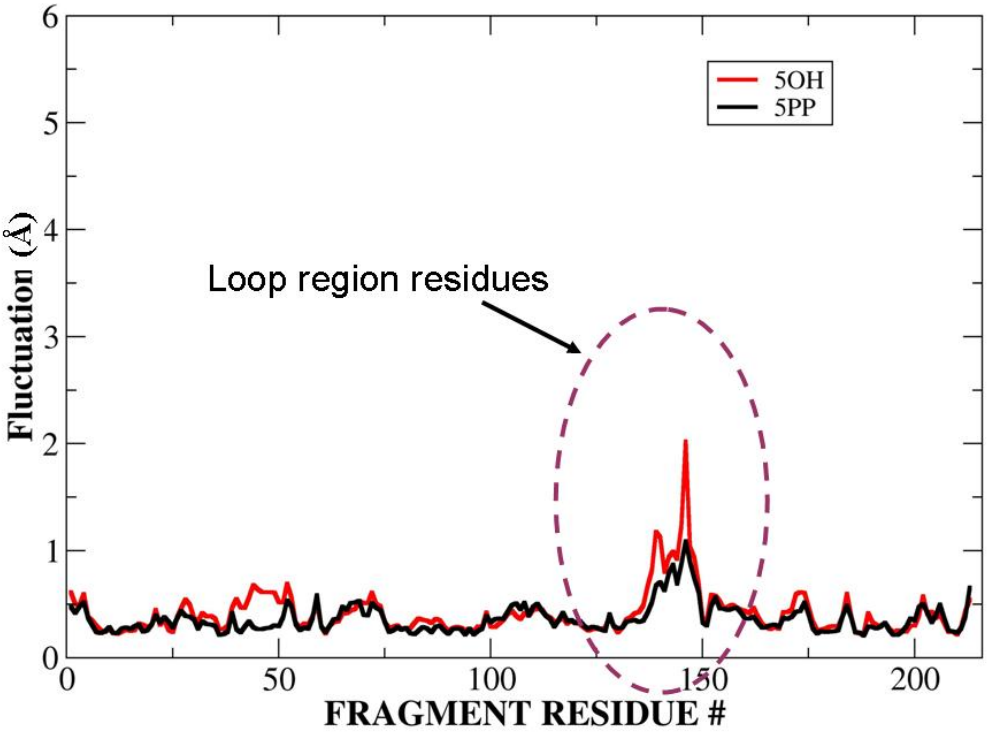


Figure 49: Fluctuation Plot for the InhA fragment system: The InhA:5OH:NAD^+ system shows a higher substrate binding loop fluctuations than the InhA:5PP:NAD^+ system.

This is unexpected as 5OH is designed to interact better with the substrate binding loop of the InhA enzyme. The MM-GBSA approach was implemented to calculate the binding free energy of the 2 inhibitors, 5PP and 5OH to the InhA enzyme.

Table 19: The binding free energies of InhA:5OH:NAD^+ and InhA:5PP:NAD^+ fragment systems calculated using MM-GBSA approach. All the energies are given in kcal/mol units

Fragment InhA System	ΔMME	ΔVDW	ΔEEL	ΔEGB	ΔEEL+ ΔEGB	ΔSA	ΔG	ΔΔG
5OH	-53.0	-42.9	-10.0	23.4	13.4	-3.9	-33.5	
5PP	-61.4	-43.8	-17.6	24.0	6.4	-3.8	-41.1	-7.6

The results of this are shown in the below table (**Table 19**). The simulations with the 5PP fragment system serve as a control to be compared to our InhA:Inhibitor tetramer system simulations. The binding free energy calculated for the 5PP binding to the monomer 2 of the InhA tetramer system is in close agreement to the binding energy calculated for the same in the InhA fragment system. From the simulations, it is seen that 5PP binds ~8 kcal/mol better than 5OH to InhA. As seen in table 6, most of the difference in binding arises from the electrostatics and EGB energies. One would expect these energies to be favorable as the inhibitor was modified to introduce hydrogen bonds. In order to understand the exhibited energetics of the system, cluster analysis was performed on both the fragment systems. I created a stripped down trajectory from the tetramer InhA trajectory (explicit solvent) with only 50 solvent molecules around 5PP bound to the monomer 2 using ptraj program. Then, I clustered this trajectory using Moil-View software by fitting the protein to residues 113 to 127 backbone CA and residues 159 to 172 backbone CA. The rmsd is

calculated on the residues 190 to 220 backbone C_α (substrate binding loop region residues). The cutoff was set to 1.5 Å. the best representative structure with the cluster with 269 members was saved. Similar clustering is also performed on the 5OH bound InhA fragment trajectory and the representative structure is saved.

The cluster analysis of the InhA fragment with 5OH bound showed that the introduced OH group misplaces the water molecule in the active site and does not replace the hydrogen bonds formed by it, unlike predicted (Figure 50).

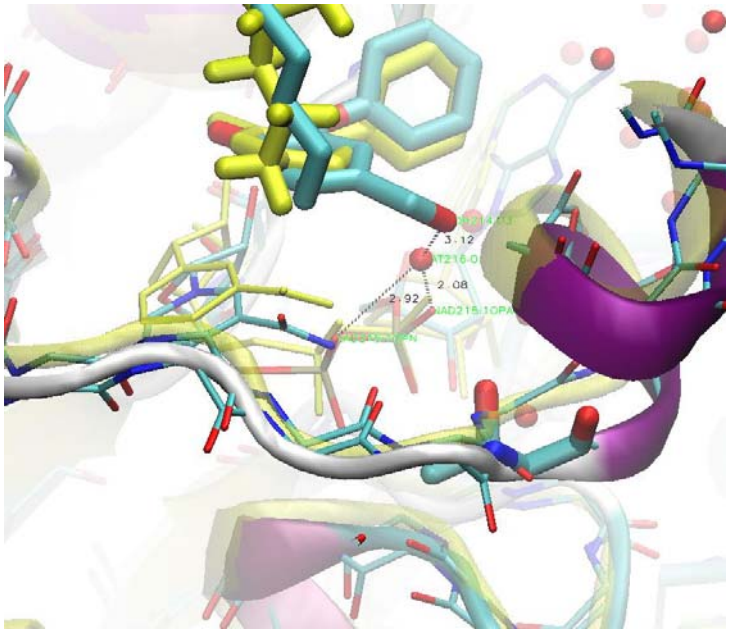


Figure 50: Overlap of InhA:5OH:NAD^+ fragment system (name type coloring) and monomer 2 of InhA:5PP:NAD^+ tetramer system (yellow). The waters from the latter are shown as red spheres.

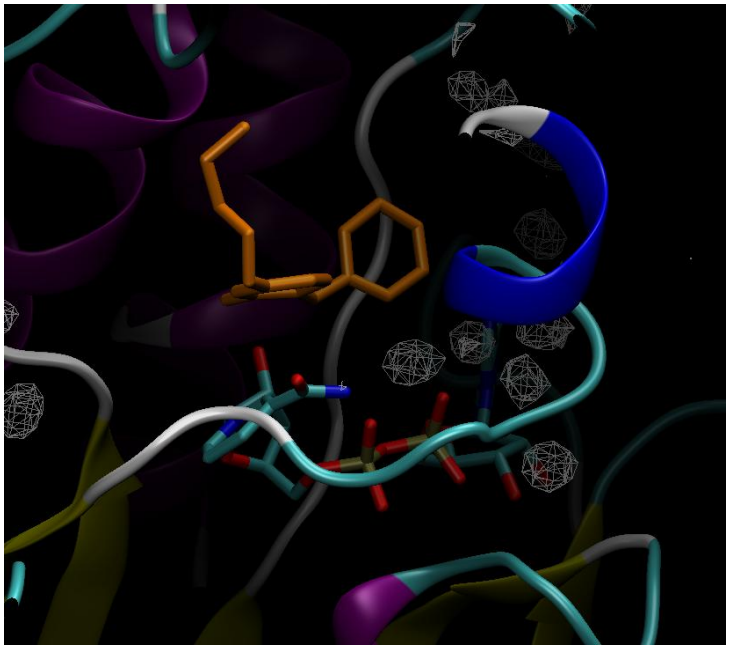


Figure 51: Water density grid (grey mesh) for the InhA:5PP:NAD^+ tetramer system. 5PP is shown in orange and the protein in name type coloring.

Further analysis on the 5PP bound InhA tetramer system using the water density grid calculations revealed that there is a structured water molecule sharing hydrogen bonds with the 2 phosphates of the co-factor NAD^+ (**Figure 51**). This water molecule is displaced resulting in poor electrostatics of the fragment system with 5OH bound to InhA.

4.8 Conclusions

The crystal structures of uncompetitive inhibitors bound to InhA have the substrate binding loop region disordered. PT70, also alkyl diphenyl ether, brings about slow onset inhibition of InhA. In order to investigate if substrate binding

loop ordering is the molecular explanation of slow onset inhibition, MD simulations on InhA bound with different inhibitors were performed. Consistent with the experimental observation, the substrate binding loop stays flexible during the course of the simulation if uncompetitive inhibitors like TCS or 5PP are bound to InhA irrespective of the starting loop conformation. MD simulations with PT70 in the active site of InhA strengthen our hypothesis that substrate binding loop ordering is coupled to slow onset inhibition of the enzyme, as the substrate binding loop stays ordered during the course of the simulation. Also, relative binding free energy calculations using MM-GBSA approach were performed to understand the how the loop ordering is coupled to binding free energy. TCS, 5PP and PT70 were correctly ranked with PT70 binding better than 5PP binding better than TCS. It was observed in the case of 5PP bound InhA that if the loop stays “ordered” as modeled (1ZID.pdb loop conformation), the inhibitor is pushed out of the active site and *vice versa*. In either case the energy of the system is comparatively higher than when the loop is partially ordered. Additional MD simulations on the InhA system with PT70 modified into 5PP, also have the substrate binding loop conformation to be ordered, unlike expected. The PT70 bound InhA loop conformation could be at a global minimum. In order to go to another conformation, it is possible that it may have to cross a large energy barrier. This is a reasonable explanation for an ordered substrate binding loop conformation when 5PP is in the active site of InhA when the starting loop conformation is as seen in PT70 bound InhA X-ray structure. Further analysis of the structural trajectories is necessary to draw a conclusive statement about loop

ordering. In the TCS and 5PP bound InhA structures (2B35.pdb, 2B36.pdb), surprisingly, the monomers in the same asymmetric unit have different number of amino acids missing from substrate binding loop region. To investigate the effect of crystal packing on the substrate binding loop conformation, further analysis and MD simulations were performed. The crystal lattice of INH-NAD adduct bound InhA structure (1ZID.pdb), revealed that the E210 in the substrate binding loop region is at salt-bridge distance from R45 of the neighboring InhA monomer. This kind of interactions may also lead to ordered substrate binding loop conformation. Two short MD simulations of 500 ps each, of 2 monomers at the crystal interface were performed. Instead of the INH-NAD adduct in the active site, NAD^+ and 5PP were modeled in one of the systems. Even though an uncompetitive inhibitor such as 5PP is bound in the active site, the substrate binding loop stays ordered. Sterically, the loop has no space to get disordered. Consequently, the inhibitor moves away from the active site. On the other system with NAD^+ in the active site (and no inhibitor) of InhA, MD simulation was performed and the NAD^+ stayed in the built in conformation. Crystal packing effects can still play a critical role in determining the susbtrate binding loop conformation.

One of the goals of this research is to predict modifications to TCS such that its interactions with the substrate binding loop are enhanced and it binds with greater affinity to InhA. Based on the simulations of the 5PP bound InhA system (with INH-NAD adduct bound InhA loop conformation), there was a water

molecule in the active site of monomer 2 that made multiple hydrogen bonding contacts with residues near the substrate binding loop region. Based on this observation, the 5PP inhibitor was modified to inhibitor 5OH by adding a CH_2OH group at the ortho- position to the alkyl chain. The relative binding free energy of 5OH binding to InhA was calculated using MM-GBSA calculations. 5PP ranked better than 5OH by 8 kcal/mol. The reason was because a conserved water molecule was displaced instead of being replaced by the CH_2OH group. If the contacts of the inhibitor with the loop are elaborated, for example by replacing the structured water molecule in the active site cavity, the inhibitor will presumably cause loop ordering. We predict that compounds with this property will be slow, tight binding InhA inhibitors with sub-nano molar affinities for the enzyme.

References:

1. Unanue, E. R., and Allen, P. M. (1987) The basis for the immunoregulatory role of macrophages and other accessory cells, *Science 236*, 551-557.

2. Takemura, R., and Werb, Z. (1984) Secretory products of macrophages and their physiological functions, *The American journal of physiology 246*, C1-9.

3. Stengel, A., Fox, H. (1915) *A text book of pathology*, W. B. Sanders Company.

4. (WHO)., W. H. O. (March 2006, Retrieved on 6 October 2006.) Tuberculosis Fact sheet N°104 - Global and regional incidence, (WHO), W. H. O., Ed.).

5. (WHO), W. H. O. (2008) WHO report 2008

   Global tuberculosis control - surveillance, planning, financing, Key Points, ((WHO), W. H. O., Ed.).

6. (WHO)., W. H. O. (Retrieved on 3 October 2006.) Stop TB Partnership. , *World Health Organization (WHO).* .

7. American Thoracic Society, C. D. C., Infectious Diseases Society of America. (2003) Treatment of tuberculosis., Division of Tuberculosis Elimination, Centers for Disease Control and Prevention (CDC).

8. Wright, A., MPH, Stop TB Dept, WHO;Bai G , PhD, Barrera L , MS, Boulahbal F, PhD, N Martín-Casabona, MD, PhD, C Gilpin, PhD, F Drobniewski, MD, PhD, M Havelková, MD, PhD, R Lepe, PhD, R Lumb, MAppSc, B Metchock, DrPH, F Portaels, PhD, M Rodrigues, PhD, S Rüsch-Gerdes, PhD, A Van Deun, MD, V Vincent, PhD, WHO/International Union Against Tuberculosis and Lung Disease Network of Supranational Reference Laboratories. V Leimane, MD, V Riekstina, MD, PhD, G Skenders, MD, State Agency for Tuberculosis and Lung Diseases, Riga, Latvia. T Holtz, MD, R Pratt, K Laserson, ScD, C Wells, MD, P Cegielski, MD, Div of Tuberculosis Elimination, National Center for HIV, STD, and TB Prevention; NS Shah, MD, EIS Officer, CDC. (2006) Emergence of Mycobacterium tuberculosis with extensive resistance to second-line drugs—worldwide, 2000–2004.

9. Bannon, M. J. (1999) BCG and tuberculosis, *Archives of disease in childhood 80*, 80-83.

10. Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. (1992) The Catalase Peroxidase Gene and Isoniazid Resistance of Mycobacterium-Tuberculosis, *Nature 358*, 591-593.

11. Piatek, A. S., Telenti, A., Murray, M. R., El-Hajj, H., Jacobs, W. R., Jr., Kramer, F. R., and Alland, D. (2000) Genotypic analysis of Mycobacterium tuberculosis in two distinct populations using molecular beacons: implications for rapid susceptibility testing, *Antimicrobial agents and chemotherapy 44*, 103-110.

12. Telenti, A., Imboden, P., Marchesi, F., Lowrie, D., Cole, S., Colston, M. J., Matter, L., Schopfer, K., and Bodmer, T. (1993) Detection of rifampicin-resistance mutations in Mycobacterium tuberculosis, *Lancet 341*, 647-650.

13. Winder, F. G., and Collins, P. B. (1970) Inhibition by isoniazid of synthesis of mycolic acids in Mycobacterium tuberculosis, *Journal of general microbiology 63*, 41-48.

14. Wang, L., and Takayama, K. (1972) Relationship between the uptake of isoniazid and its action on in vivo mycolic acid synthesis in Mycobacterium tuberculosis, *Antimicrobial agents and chemotherapy 2*, 438-441.

15. Tonge, P. J. (2000) Another brick in the wall, *Nature structural biology 7*, 94-96.

16. Takayama, K., Qureshi, N., Raetz, C. R. H., Ribi, E., Peterson, J., Cantrell, J. L., Pearson, F. C., Wiggins, J., and Johnson, A. G. (1984) Influence of Fine-Structure of Lipid-a on Limulus Amebocyte Lysate Clotting and Toxic Activities, *Infection and Immunity 45*, 350-355.

17. Takayama, K., Qureshi, N., Ribi, E., and Cantrell, J. L. (1984) Separation and Characterization of Toxic and Nontoxic Forms of Lipid-A, *Reviews of Infectious Diseases 6*, 439-443.

18. Majerus, P. W., and Vagelos, P. R. (1967) Fatty acid biosynthesis and the role of the acyl carrier protein, *Adv Lipid Res 5*, 1-33.

19. Gibson, D. M., Titchener, E. B., and Wakil, S. J. (1958) Studies on the mechanism of fatty acid synthesis. V. Bicarbonate requirement for the synthesis of long-chain fatty acids, *Biochimica et biophysica acta 30*, 376-383.

20. Lynen, F. (1961) Biosynthesis of saturated fatty acids, *Federation proceedings 20*, 941-951.

21. Wakil, S. J., Pugh, E. L., and Sauer, F. (1964) The Mechanism of Fatty Acid Synthesis, *Proc Natl Acad Sci U S A 52*, 106-114.

22. Smith, S., Witkowski, A., and Joshi, A. K. (2003) Structural and functional organization of the animal fatty acid synthase, *Progress in lipid research 42*, 289-317.

23. Rozwarski, D. A., Vilcheze, 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, *The Journal of biological chemistry 274*, 15582-15589.

24. Tai, M. H., Chirala, S. S., and Wakil, S. J. (1993) Roles of Ser101, Asp236, and His237 in catalysis of thioesterase II and of the C-terminal region of the enzyme in its interaction with fatty acid synthase, *Proc Natl Acad Sci U S A 90*, 1852-1856.

25. Smith, S. (1994) The animal fatty acid synthase: one gene, one polypeptide, seven enzymes, *Faseb J 8*, 1248-1259.

26. Brink, J., Ludtke, S. J., Kong, Y., Wakil, S. J., Ma, J., and Chiu, W. (2004) Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis, *Structure 12*, 185-191.

27. Tsukamoto, Y., Wong, H., Mattick, J. S., and Wakil, S. J. (1983) The architecture of the animal fatty acid synthetase complex. IV. Mapping of active centers and model for the mechanism of action, *The Journal of biological chemistry 258*, 15312-15322.

28. Marrakchi, H., Zhang, Y. M., and Rock, C. O. (2002) Mechanistic diversity and regulation of Type II fatty acid synthesis, *Biochemical Society transactions 30*, 1050-1055.

29. Vilcheze, C., Morbidoni, H. R., Weisbrod, T. R., Iwamoto, H., Kuo, M., Sacchettini, J. C., and Jacobs, W. R., Jr. (2000) Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis, *J Bacteriol 182*, 4059-4067.

30. Heath, R. J., White, S. W., and Rock, C. O. (2001) Lipid biosynthesis as a target for antibacterial agents, *Progress in lipid research 40*, 467-497.

31. Kater, M. M., Koningstein, G. M., Nijkamp, H. J., and Stuitje, A. R. (1994) The use of a hybrid genetic system to study the functional relationship between prokaryotic and plant multi-enzyme fatty acid synthetase complexes, *Plant molecular biology 25*, 771-790.

32. Magnuson, K., Jackowski, S., Rock, C. O., and Cronan, J. E., Jr. (1993) Regulation of fatty acid biosynthesis in Escherichia coli, *Microbiol Rev 57*, 522-542.

33. Heath, R. J., and Rock, C. O. (1995) Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of fatty acid elongation in Escherichia coli, *The Journal of biological chemistry 270*, 26538-26542.

34. Heath, R. J., Li, J., Roland, G. E., and Rock, C. O. (2000) Inhibition of the Staphylococcus aureus NADPH-dependent enoyl-acyl carrier protein reductase by triclosan and hexachlorophene, *The Journal of biological chemistry 275*, 4654-4659.

35. Heath, R. J., and Rock, C. O. (2000) A triclosan-resistant bacterial enzyme, *Nature 406*, 145-146.

36. Campbell, J. W., and Cronan, J. E., Jr. (2001) Bacterial fatty acid biosynthesis: targets for antibacterial drug discovery, *Annual review of microbiology 55*, 305-332.

37. Barry, C. E., Lee, R. E., Mdluli, K., Sampson, A. E., Schroeder, B. G., Slayden, R. A., and Yuan, Y. (1998) Mycolic acids: Structure, biosynthesis and physiological functions, *Prog Lipid Res 37*, 143-179.

38. Slayden, R. A., and Barry, C. E., 3rd. (2000) The genetics and biochemistry of isoniazid resistance in mycobacterium tuberculosis, *Microbes and infection / Institut Pasteur 2*, 659-669.

39. Davis, M. C., Franzblau, S. G., and Martin, A. R. (1998) Syntheses and evaluation of benzodiazaborine compounds against M. tuberculosis H37Rv in vitro, *Bioorganic & medicinal chemistry letters 8*, 843-846.

40. Middlebrook, G., Cohn, M. L., and Schaefer, W. B. (1954) Studies on isoniazid and tubercle bacilli. III. The isolation, drug-susceptibility, and catalase-testing of tubercle bacilli from isoniazid-treated patients, *American review of tuberculosis 70*, 852-872.

41. Zhang, Y., Heym, B., Allen, B., Young, D., and Cole, S. (1992) The catalase-peroxidase gene and isoniazid resistance of Mycobacterium tuberculosis, *Nature 358*, 591-593.

42. 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.

43. 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.

44. Quemard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R., Jr., and Blanchard, J. S. (1995) Enzymatic characterization of the target for isoniazid in Mycobacterium tuberculosis, *Biochemistry 34*, 8235-8241.

45. Musser, J. M. (1995) Antimicrobial agent resistance in mycobacteria: molecular genetic insights, *Clinical microbiology reviews 8*, 496-514.

46. Basso, L. A., Zheng, R., Musser, J. M., Jacobs, W. R., Jr., and Blanchard, J. S. (1998) Mechanisms of isoniazid resistance in Mycobacterium tuberculosis: enzymatic characterization of enoyl reductase mutants identified in isoniazid-resistant clinical isolates, *The Journal of infectious diseases 178*, 769-775.

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

48. Bhargava, H. N., and Leonard, P. A. (1996) Triclosan: applications and safety, *American journal of infection control 24*, 209-218.

49. Jones, R. D., Jampani, H. B., Newman, J. L., and Lee, A. S. (2000) Triclosan: a review of effectiveness and safety in health care settings, *American journal of infection control 28*, 184-196.

50. Sivaraman, S., Sullivan, T. J., Johnson, F., Novichenok, P., Cui, G., Simmerling, C., and Tonge, P. J. (2004) Inhibition of the bacterial enoyl reductase FabI by triclosan: a structure-reactivity analysis of FabI inhibition by triclosan analogues, *J Med Chem 47*, 509-518.

51. McMurry, L. M., Oethinger, M., and Levy, S. B. (1998) Triclosan targets lipid synthesis, *Nature 394*, 531-532.

52. Heath, R. J., Yu, Y. T., Shapiro, M. A., Olson, E., and Rock, C. O. (1998) Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis, *J Biol Chem 273*, 30316-30320.

53. Heath, R. J., Rubin, J. R., Holland, D. R., Zhang, E., Snow, M. E., and Rock, C. O. (1999) Mechanism of triclosan inhibition of bacterial fatty acid synthesis, *The Journal of biological chemistry 274*, 11110-11114.

54. Ward, W. H., Holdgate, G. A., Rowsell, S., McLean, E. G., Pauptit, R. A., Clayton, E., Nichols, W. W., Colls, J. G., Minshull, C. A., Jude, D. A., Mistry, A., Timms, D., Camble, R., Hales, N. J., Britton, C. J., and Taylor, I. W. (1999) Kinetic and structural characteristics of the inhibition of enoyl (acyl carrier protein) reductase by triclosan, *Biochemistry 38*, 12514-12525.

55. Sivaraman, S., Zwahlen, J., Bell, A. F., Hedstrom, L., and Tonge, P. J. (2003) Structure-activity studies of the inhibition of FabI, the enoyl reductase from Escherichia coli, by triclosan: kinetic analysis of mutant FabIs, *Biochemistry 42*, 4406-4413.

56. Turnowsky, F., Fuchs, K., Jeschek, C., and Hogenauer, G. (1989) Envm Genes of Salmonella-Typhimurium and Escherichia-Coli, *Journal of Bacteriology 171*, 6555-6565.

57. McMurry, L. M., McDermott, P. F., and Levy, S. B. (1999) Genetic evidence that InhA of Mycobacterium smegmatis is a target for triclosan, *Antimicrobial agents and chemotherapy 43*, 711-713.

58. Parikh, S. L., Xiao, G., and Tonge, P. J. (2000) Inhibition of InhA, the enoyl reductase from Mycobacterium tuberculosis, by triclosan and isoniazid, *Biochemistry 39*, 7645-7650.

59. Levy, C. W., Roujeinikova, A., Sedelnikova, S., Baker, P. J., Stuitje, A. R., Slabas, A. R., Rice, D. W., and Rafferty, J. B. (1999) Molecular basis of triclosan activity, *Nature 398*, 383-384.

60. Qiu, X., Janson, C. A., Court, R. I., Smyth, M. G., Payne, D. J., and Abdel-Meguid, S. S. (1999) Molecular basis for triclosan activity involves a flipping loop in the active site, *Protein Sci 8*, 2529-2532.

61. Roujeinikova, A., Levy, C. W., Rowsell, S., Sedelnikova, S., Baker, P. J., Minshull, C. A., Mistry, A., Colls, J. G., Camble, R., Stuitje, A. R., Slabas, A. R., Rafferty, J. B., Pauptit, R. A., Viner, R., and Rice, D. W. (1999) Crystallographic analysis of triclosan bound to enoyl reductase, *J Mol Biol 294*, 527-535.

62. Roujeinikova, A., Sedelnikova, S., de Boer, G. J., Stuitje, A. R., Slabas, A. R., Rafferty, J. B., and Rice, D. W. (1999) Inhibitor binding studies on enoyl reductase reveal conformational changes related to substrate recognition, *J Biol Chem 274*, 30811-30817.

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

64. 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., Jr., Sacchettini, J. C., and Fidock, D. A. (2003) Targeting tuberculosis and malaria through inhibition of Enoyl reductase: compound activity and structural data, *J Biol Chem 278*, 20851-20859.

65. Keating, M. M., Gong, H., and Byers, D. M. (2002) Identification of a key residue in the conformational stability of acyl carrier protein, *Biochim Biophys Acta 1601*, 208-214.

66. Worsham, L. M., Earls, L., Jolly, C., Langston, K. G., Trent, M. S., and Ernst-Fonberg, M. L. (2003) Amino acid residues of Escherichia coli acyl carrier protein involved in heterologous protein interactions, *Biochemistry 42*, 167-176.

67. Elovson, J., and Vagelos, P. R. (1968) Acyl carrier protein. X. Acyl carrier protein synthetase, *J Biol Chem 243*, 3603-3611.

68. Majerus, P. W., Alberts, A. W., and Vagelos, P. R. (1964) The Acyl Carrier Protein of Fatty Acid Synthesis: Purification, Physical Properties, and Substrate Binding Site, *Proc Natl Acad Sci U S A 51*, 1231-1238.

69. Wakil, S. J. (1989) Fatty acid synthase, a proficient multifunctional enzyme, *Biochemistry 28*, 4523-4530.

70. Vagelos, P. R., Majerus, P. W., Alberts, A. W., Larrabee, A. R., and Ailhaud, G. P. (1966) Structure and function of the acyl carrier protein, *Fed Proc 25*, 1485-1494.

71. Flugel, R. S., Hwangbo, Y., Lambalot, R. H., Cronan, J. E., Jr., and Walsh, C. T. (2000) Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in Escherichia coli, *J Biol Chem 275*, 959-968.

72. Gehring, A. M., Lambalot, R. H., Vogel, K. W., Drueckhammer, D. G., and Walsh, C. T. (1997) Ability of Streptomyces spp. acyl carrier proteins and coenzyme A analogs to serve as substrates in vitro for E. coli holo-ACP synthase, *Chem Biol 4*, 17-24.

73. Lambalot, R. H., Gehring, A. M., Flugel, R. S., Zuber, P., LaCelle, M., Marahiel, M. A., Reid, R., Khosla, C., and Walsh, C. T. (1996) A new enzyme superfamily - the phosphopantetheinyl transferases, *Chem Biol 3*, 923-936.

74. Dixon, M. (1979) in *Enzymes, 3rd Edition*, pp 155-199.

75. Roujeinikova, A., Baldock, C., Simon, W. J., Gilroy, J., Baker, P. J., Stuitje, A. R., Rice, D. W., Slabas, A. R., and Rafferty, J. B. (2002) X-ray crystallographic studies on butyryl-ACP reveal flexibility of the structure around a putative acyl chain binding site, *Structure 10*, 825-835.

76. Wakil, S. J., Stoops, J. K., and Joshi, V. C. (1983) Fatty acid synthesis and its regulation, *Annu Rev Biochem 52*, 537-579.

77. Prescott, D. J., and Vagelos, P. R. (1972) Acyl carrier protein, *Adv Enzymol Relat Areas Mol Biol 36*, 269-311.

78. Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III, *J Biol Chem 276*, 8231-8238.

79. Peczuh, M. W., and Hamilton, A. D. (2000) Peptide and protein recognition by designed molecules, *Chem Rev 100*, 2479-2494.

80. Xu, G. Y., Tam, A., Lin, L., Hixon, J., Fritz, C. C., and Powers, R. (2001) Solution structure of B. subtilis acyl carrier protein, *Structure 9*, 277-287.

81. Karplus, M., and McCammon, J. A. (2002) Molecular dynamics simulations of biomolecules, *Nature structural biology 9*, 646-652.

82. van Gunsteren, W. F., Bakowies, D., Baron, R., Chandrasekhar, I., Christen, M., Daura, X., Gee, P., Geerke, D. P., Glattli, A., Hunenberger, P. H., Kastenholz, M. A., Ostenbrink, C., Schenk, M., Trzesniak, D., van der Vegt, N. F. A., and Yu, H. B. (2006) Biomolecular modeling: Goals, problems, perspectives, *Angewandte Chemie-International Edition 45*, 4064-4092.

83. Jorgensen, W. L. (2004) The many roles of computation in drug discovery, *Science 303*, 1813-1818.

84. Simmerling, C., Elber, R. and Zhang, J. (1995) *MOIL-View - A Program for Visualization of Structure and Dynamics of Biomolecules and STO- A Program for Computing Stochastic Paths, in Modelling of Biomolecular Structure and Mechanisms,* , Kluwer, Netherlands

85. DeLano, W. L. (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos, CA, USA.

86. Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: Visual molecular dynamics, *Journal of Molecular Graphics 14*, 33-&.

87. Plum, G. E., Grollman, A. P., Johnson, F., and Breslauer, K. J. (1995) Influence of the oxidatively damaged adduct 8-oxodeoxyguanosine on the conformation, energetics, and thermodynamic stability of a DNA duplex, *Biochemistry 34*, 16148-16160.

88. Alder, B. J., and Wainwright, T. E. (1957) Phase Transition for a Hard Sphere System, *J Chem Phys 27*, 1208-1209.

89. Stilling.Fh, and Rahman, A. (1974) Improved Simulation of Liquid Water by Molecular-Dynamics, *J Chem Phys 60*, 1545-1557.

90. Mccammon, J. A., Gelin, B. R., and Karplus, M. (1977) Dynamics of Folded Proteins, *Nature 267*, 585-590.

91. 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.

92. Geney, R., Layten, M., Gomperts, R., Hornak, V., and Simmerling, C. (2006) Investigation of salt bridge stability in a generalized born solvent model, *Journal of Chemical Theory and Computation 2*, 115-127.

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

94. Hornak, V., and Simmerling, C. (2004) Development of softcore potential functions for overcoming steric barriers in molecular dynamics simulations, *Journal of Molecular Graphics & Modelling 22*, 405-413.

95. Case, D. A., Cheatham, T., Darden, T., Merz, K., Simmerling, C., Luo, R., Wang, J. (2008) amber.scripps.edu.

96. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *Journal of Computational Chemistry 26*, 1668-1688.

97. 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, *Journal of Chemical Physics 79*, 926-935.

98. Jorgensen, W. L. (1981) Quantum and Statistical Mechanical Studies of Liquids .10. Transferable Intermolecular Potential Functions for Water, Alcohols, and Ethers - Application to Liquid Water, *Journal of the American Chemical Society 103*, 335-340.

99. H.J.C. Berendsen, J. P. M. P., W.F. van Gunsteren, and J. Hermans. (1981) In Intermolecular Forces, (B. Pullman (Reidel, D., Ed.), p p. 331.

100. Horn, H. W., Swope, W. C., Pitera, J. W., Madura, J. D., Dick, T. J., Hura, G. L., and Head-Gordon, T. (2004) Development of an improved four-site water model for biomolecular simulations: TIP4P-Ew, *J Chem Phys 120*, 9665-9678.

101. Jorgensen, W. L. (1982) Quantum and Statistical Mechanical Studies of Liquids .24. Revised Tips for Simulations of Liquid Water and Aqueous-Solutions, *J Chem Phys 77*, 4156-4163.

102. Mahoney, M. W., and Jorgensen, W. L. (2000) A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions, *J Chem Phys 112*, 8910-8922.

103. Levitt, M., Hirshberg, M., Sharon, R., Laidig, K. E., and Daggett, V. (1997) Calibration and testing of a water model for simulation of the molecular dynamics of proteins and nucleic acids in solution, *J Phys Chem B 101*, 5051-5061.

104. Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987) The Missing Term in Effective Pair Potentials, *J Phys Chem-Us 91*, 6269-6271.

105. Darden, T., York, D., and Pedersen, L. (1993) Particle Mesh Ewald - an N.Log(N) Method for Ewald Sums in Large Systems, *Journal of Chemical Physics 98*, 10089-10092.

106. Hockney, R. W., and Eastwood, J. W. (1988) *Computer simulation using particles*, Special student ed., A. Hilger, Bristol [England] ; Philadelphia.

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

108. 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.

109. Cronon, J. E., Jr., Rock, C.O. (1996) *Biosynthesis of Membrane Lipids in Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*, American Society of Microbiology, Washington, D.C.

110. Rock, C. O., and Jackowski, S. (2002) Forty years of bacterial fatty acid synthesis, *Biochemical and Biophysical Research Communications 292*, 1155-1166.

111. Rock, C. O., and Cronan, J. E. (1996) Escherichia coli as a model for the regulation of dissociable (type II) fatty acid biosynthesis, *Biochim Biophys Acta 1302*, 1-16.

112. Zhang, Y. M., Lu, Y. J., and Rock, C. O. (2004) The reductase steps of the type II fatty acid synthase as antimicrobial targets, *Lipids 39*, 1055-1060.

113. Hayashi, T., Yamamoto, O., Sasaki, H., Okazaki, H., and Kawaguchi, A. (1984) Inhibition of fatty acid synthesis by the antibiotic thiolactomycin, *J Antibiot (Tokyo) 37*, 1456-1461.

114. Price, A. C., Choi, K. H., Heath, R. J., Li, Z., White, S. W., and Rock, C. O. (2001) Inhibition of beta-ketoacyl-acyl carrier protein synthases by thiolactomycin and cerulenin. Structure and mechanism, *J. Biol. Chem. 276*, 6551-6559.

115. Moir, D. T. (2005) Identification of inhibitors of bacterial enoyl-acyl carrier protein reductase, *Curr. Drug Targets Infect. Disord. 5*, 297-305.

116. Baldock, C., Rafferty, J. B., Sedelnikova, S. E., Baker, P. J., Stuitje, A. R., Slabas, A. R., Hawkes, T. R., and Rice, D. W. (1996) A mechanism of drug action revealed by structural studies of enoyl reductase, *Science 274*, 2107-2110.

117. Rozwarski, D. A., Grant, G. A., Barton, D. H. R., Jacobs WR, J. r., and Sacchettini, J. C. (1998) Modification of the NADH of the isoniazid target (InhA) from *Mycobacterium tuberculosis*, *Science 279*, 98-102.

118. Surolia, N., and Surolia, A. (2001) Triclosan offers protection against blood stages of malaria by inhibiting enoyl-ACP reductase of Plasmodium falciparum (vol 7, pg 167, 2000), *Nature Medicine 7*, 636-636.

119. Heath, R. J., Li, J., Roland, G. E., and Rock, C. O. (2000) Inhibition of the Staphylococcus aureus NADPH-dependent enoyl- acyl carrier protein reductase by triclosan and hexachlorophene, *J. Biol. Chem. 275*, 4654-4659.

120. Parikh, S. L., Xiao, G., and Tonge, P. J. (2000) Inhibition of InhA, the enoyl-reductase from *Mycobacterium tuberculosis,* by triclosan and isoniazid, *Biochemistry 39*, 7645-7650.

121. Gohlke, H., Kiel, C., and Case, D. A. (2003) Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras-RalGDS complexes, *Journal of Molecular Biology 330*, 891-913.

122. Massova, I., and Kollman, P. A. (2000) Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding, *Perspectives in Drug Discovery and Design 18*, 113-135.

123. 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.

124. Kuhn, B., and Kollman, P. A. (2000) Binding of a diverse set of ligands to avidin and streptavidin: An accurate quantitative prediction of their relative

affinities by a combination of molecular mechanics and continuum solvent models, *Journal of Medicinal Chemistry 43*, 3786-3791.

125. Wang, J. M., Morin, P., Wang, W., and Kollman, P. A. (2001) Use of MM-PBSA in reproducing the binding free energies to HIV-1 RT of TIBO derivatives and predicting the binding mode to HIV-1 RT of efavirenz by docking and MM-PBSA, *Journal of the American Chemical Society 123*, 5221-5230.

126. Masukawa, K. M., Kollman, P. A., and Kuntz, I. D. (2003) Investigation of neuraminidase-substrate recognition using molecular dynamics and free energy calculations, *Journal of Medicinal Chemistry 46*, 5628-5637.

127. Huo, S., Wang, J., Cieplak, P., Kollman, P. A., and Kuntz, I. D. (2002) Molecular dynamics and free energy analyses of cathepsin D-inhibitor interactions: insight into structure-based ligand design, *J Med Chem 45*, 1412-1419.

128. Wang, W., Lim, W. A., Jakalian, A., Wang, J., Wang, J. M., Luo, R., Bayly, C. T., and Kollman, P. A. (2001) An analysis of the interactions between the Sem-5 SH3 domain and its ligands using molecular dynamics, free energy calculations, and sequence analysis, *Journal of the American Chemical Society 123*, 3986-3994.

129. Suenaga, A., Hatakeyama, M., Ichikawa, M., Yu, X. M., Futatsugi, N., Narumi, T., Fukui, K., Terada, T., Taiji, M., Shirouzu, M., Yokoyama, S., and Konagaya, A. (2003) Molecular dynamics, free energy, and SPR analyses of the interactions between the SH2 domain of grb2 and ErbB phosphotyrosyl peptides, *Biochemistry 42*, 5195-5200.

130. Donini, O. A. T., and Kollman, P. A. (2000) Calculation and prediction of binding free energies for the matrix metalloproteinases, *Journal of Medicinal Chemistry 43*, 4180-4188.

131. Sivaraman, S., Sullivan, T. J., Johnson, F., Novichenok, P., Cui, G. L., Simmerling, C., and Tonge, P. J. (2004) Inhibition of the bacterial enoyl reductase FabI by triclosan: A structure-reactivity analysis of FabI inhibition by triclosan analogues, *Journal of Medicinal Chemistry 47*, 509-518.

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

133. Petersen, H. G. (1995) Accuracy and Efficiency of the Particle Mesh Ewald Method, *Journal of Chemical Physics 103*, 3668-3679.

134. Wang, J. M., 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.

135. Frisch, M. J. T., G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Zakrzewski,, V. G.; Montgomery, J., J. A.; Stratmann, R. E.; Burant, J. C.; Dapprich, S.; Millam, J. M.; Daniels, A., D.; Kudin, K. N. S., M. C.; Farkas, O.; Tomasi, J.; Barone, V.; Cossi, M.; Cammi, R.; Mennucci, B.;, Pomelli, C. A., C.; Clifford, S.; Ochterski, J.; Petersson, G. A.; Ayala, P. Y.; Cui, Q.; Morokuma, K.;, Malick, D. K. R., A. D.; Raghavachari, K.; Foresman, J. B.; Cioslowski, J.; Ortiz, J. V.; Stefanov, B., B.; Liu, G. L., A.; Piskorz, P.; Komaromi, I.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Keith, T.;, Al-Laham, M. A. P., C. Y.; Nanayakkara, A.; Gonzalez, C.; Challacombe, M.; Gill, P. M. W.; Johnson,, and B.; Chen, W. W., M. W.; Andres, J. L.; Gonzalez, C.; Head-Gordon, M.; Replogle, E. S.; Pople, J. A. *Gaussian 98, Revision A.5, Gaussian, Inc., Pittsburgh PA, 1998.*

136. Cornell, W. D., Cieplak, P., Bayly, C. I., and Kollman, 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.

137. Bayly, C. I., Cieplak, P., Cornell, W. D., and Kollman, P. A. (1993) A Well-Behaved Electrostatic Potential Based Method Using Charge Restraints for Deriving Atomic Charges - the Resp Model, *Journal of Physical Chemistry 97*, 10269-10280.

138. Cieplak, P., Cornell, W. D., Bayly, C., and Kollman, P. A. (1995) Application of the Multimolecule and Multiconformational Resp Methodology to Biopolymers - Charge Derivation for DNA, Rna, and Proteins, *Journal of Computational Chemistry 16*, 1357-1377.

139. Honig, B., and Nicholls, A. (1995) Classical Electrostatics in Biology and Chemistry, *Science 268*, 1144-1149.

140. Sitkoff, D., Sharp, K. A., and Honig, B. (1994) Accurate Calculation of Hydration Free-Energies Using Macroscopic Solvent Models, *Journal of Physical Chemistry 98*, 1978-1988.

141. Sanner, M. F., Olson, A. J., and Spehner, J. C. (1996) Reduced surface: An efficient way to compute molecular surfaces, *Biopolymers 38*, 305-320.

142. Zhuang, S. Z., J.; Jiang, Y.; Mao, X.; Zhang, B.; Liu, H.; Yu, Q. (2005) Some insights into the stereochemistry of inhibition of macrophage migration inhibitory factor with 2-fluoro-p-hydroxycinnamate and its

analogues from molecular dynamics simulations, *Journal of Medicinal Chemistry 48*, 7208-7214.

143. Rizzo, R. C., Toba, S., and Kuntz, I. D. (2004) A molecular basis for the selectivity of thiadiazole urea inhibitors with stromelysin-1 and gelatinase-A from generalized born molecular dynamics simulations, *Journal of Medicinal Chemistry 47*, 3065-3074.

144. Rock, C. O., and Cronan, J. E. (1981) Acyl carrier protein from *Escherichia coli*, *Methods Enzymol. 71* 341-351.

145. Rock, C. O., and Jackowski, S. (2002) Forty years of bacterial fatty acid synthesis, *Biochem. Biophys. Res. Commun. 292*, 1155-1166.

146. Cane, D. E., and Walsh, C. T. (1999) The parallel and convergent universes of polyketide synthases and nonribosomal peptide synthetases, *Chem. Biol. 6*, R319-325.

147. Tang, L., Weissborn, A. C., and Kennedy, E. P. (1997) Domains of Escherichia coli acyl carrier protein important for membrane-derived-oligosaccharide biosynthesis, *J. Bacteriol. 179*, 3697-3705.

148. Brozek, K. A., and Raetz, C. R. (1990) Biosynthesis of lipid A in Escherichia coli. Acyl carrier protein-dependent incorporation of laurate and myristate, *J. Biol. Chem. 265*, 15410-15417.

149. Sweet, C. R., Williams, A. H., Karbarz, M. J., Werts, C., Kalb, S. R., Cotter, R. J., and Raetz, C. R. (2004) Enzymatic synthesis of lipid A molecules with four amide-linked acyl chains. LpxA acyltransferases selective for an analog of UDP-N-acetylglucosamine in which an amine replaces the 3"-hydroxyl group, *J. Biol. Chem. 279*, 25411-25419.

150. Cronan, J. E. (2003) Bacterial membrane lipids: where do we stand?, *Annu. Rev. Microbiol. 57*, 203-224.

151. Vagelos, P. R., Majerus, P. W., Alberts, A. W., Larrabee, A. R., and Ailhaud, G. P. (1966) Structure and function of the acyl carrier protein, *Fed. Proc. 25*, 1485-1494.

152. Majerus, P. W., and Vagelos, P. R. (1967) Fatty acid biosynthesis and the role of the acyl carrier protein, *Adv. Lipid Res. 5*, 1-33.

153. Magnuson, K., Jackowski, S., Rock, C. O., and Cronan, J. E., Jr. (1993) Regulation of fatty acid biosynthesis in Escherichia coli, *Microbiol. Rev. 57*, 522-542.

154. Heath, R. J., White, S. W., and Rock, C. O. (2001) Lipid biosynthesis as a target for antibacterial agents, *Prog. Lipid Res. 40*, 467-497.

155. Holak, T. A., Nilges, M., Prestegard, J. H., Gronenborn, A. M., and Clore, G. M. (1988) Three-dimensional structure of acyl carrier protein in solution determined by nuclear magnetic resonance and the combined use of dynamical simulated annealing and distance geometry, *Eur. J. Biochem. 175*, 9-15.

156. Wong, H. C., Liu, G., Zhang, Y. M., Rock, C. O., and Zheng, J. (2002) The solution structure of acyl carrier protein from *Mycobacterium tuberculosis*, *J. Biol. Chem. 277*, 15874-15880.

157. Parris, K. D., Lin, L., Tam, A., Mathew, R., Hixon, J., Stahl, M., Fritz, C. C., Seehra, J., and Somers, W. S. (2000) Crystal structures of substrate binding to Bacillus subtilis holo-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites, *Structure Fold. Des. 8*, 883-895.

158. Zhang, Y. M., Rao, M. S., Heath, R. J., Price, A. C., Olson, A. J., Rock, C. O., and White, S. W. (2001) Identification and analysis of the acyl carrier protein (ACP) docking site on beta-ketoacyl-ACP synthase III, *J. Biol. Chem. 276*, 8231-8238.

159. Zhang, Y. M., Wu, B., Zheng, J., and Rock, C. O. (2003) Key residues responsible for acyl carrier protein and beta-ketoacyl-acyl carrier protein reductase (FabG) interaction, *J. Biol. Chem. 278*, 52935-52943.

160. Heath, R. J., Rubin, J. R., Holland, D. R., Zhang, E., Snow, M. E., and Rock, C. O. (1999) Mechanism of triclosan inhibition of bacterial fatty acid synthesis, *J. Biol. Chem. 274*, 11110-11114.

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

162. Sivaraman, S., Zwahlen, J., Bell, A. F., Hedstrom, L., and Tonge, P. J. (2003) Structure-Activity Studies of the Inhibition of FabI, the Enoyl Reductase from Escherichia coli, by Triclosan: Kinetic Analysis of Mutant FabIs, *Biochemistry 42*, 4406-4413.

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

164. Quemard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R., Jr., and Blanchard, J. S. (1995) Enzymatic Characterization of the Target for Isoniazid in Mycobacterium tuberculosis, *Biochemistry 34*, 8235-8241.

165. Rozwarski, D. A., Vilcheze, 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, *J. Biol. Chem. 274*, 15582-15589.

166. Case, D. A., Cheatham, T. E., Darden, T., Gohlke, H., Luo, R., Merz, K. M., Onufriev, A., Simmerling, C., Wang, B., and Woods, R. J. (2005) The Amber biomolecular simulation programs, *J. Comput. Chem. 26*, 1668-1688.

167. Wang, J. M., 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?, *J. Comput. Chem. 21*, 1049-1074.

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

169. 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, *J. Phys. Chem. B 106*, 11658-11665.

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

171. Jakalian, A., Bush, B. L., Jack, D. B., and Bayly, C. I. (2000) Fast, efficient generation of high-quality atomic Charges. AM1-BCC model: I. Method, *J. Comput. Chem. 21*, 132-146.

172. Fillgrove, K. L., and Anderson, V. E. (2000) Orientation of coenzyme A substrates, nicotinamide and active site functional groups in (di)enoyl-coenzyme A reductases, *Biochemistry 39*, 7001-7011.

173. Zagrovic, B., and Pande, V. (2003) Solvent viscosity dependence of the folding rate of a small protein: Distributed computing study, *J. Am. Chem. Soc. 24*, 1432-1436.

174. Still, W. C., Tempczyk, A., Hawley, R. C., and Hendrickson, T. (1990) Semianalytical Treatment Of Solvation For Molecular Mechanics And Dynamics, *J. Am. Chem. Soc. 112*, 6127-6129.

175. Hawkins, G. D., Cramer, C. J., and Truhlar, D. G. (1995) Pairwise Solute Descreening Of Solute Charges From A Dielectric Medium, *Chem. Phys. Lett. 246*, 122-129.

176. 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, *J. Phys. Chem. 100*, 19824-19839.

177. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A. and Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters, *Proteins*, In Press.

178. Feig, M., Onufriev, A., Lee, M. S., Im, W., Case, D. A., and Brooks, C. L. (2004) Performance comparison of generalized born and Poisson methods in the calculation of electrostatic solvation energies for protein structures, *J. Comput. Chem. 25*, 265-284.

179. Onufriev, A., Bashford, D., and Case, D. A. (2004) Exploring protein native states and large-scale conformational changes with a modified generalized born model, *Proteins 55*, 383-394.

180. Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 protease flaps spontaneously close to the correct structure in simulations following manual placement of an inhibitor into the open state, *J. Am. Chem. Soc. 128*, 2812-2813.

181. Hornak, V., Okur, A., Rizzo, R. C., and Simmerling, C. (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations, *Proc. Natl. Acad. Sci. U.S.A. 103*, 915-920.

182. Geney, R., Layten, M., Gomperts, R., Hornak, V., and Simmerling, C. (2006) Investigation of salt bridge stability in a generalized born solvent model, *J. Chem. Theory Comput. 2*, 115-127.

183. Parikh, S., Moynihan, D. P., Xiao, G., and Tonge, P. J. (1999) Roles of tyrosine 158 and lysine 165 in the catalytic mechanism of InhA, the enoyl-ACP reductase from *Mycobacterium tuberculosis*, *Biochemistry 38*, 13623-13634.

184. Fillgrove, K. L., and Anderson, V. E. (2001) The mechanism of dienoyl-CoA reduction by 2,4-dienoyl-CoA reductase is stepwise: Observation of a dienolate intermediate, *Biochemistry 40*, 12412-12421.

185. Ward, W. H. J., Holdgate, G. A., Rowsell, S., McLean, E. G., Pauptit, R. A., Clayton, E., Nichols, W. W., Colls, J. G., Minshull, C. A., Jude, D. A., Mistry, A., Timms, D., Camble, R., Hales, N. J., Britton, C. J., and Taylor, I. W. F. (1999) Kinetic and structural characteristics of the inhibition of enoyl (acyl carrier protein) reductase by triclosan, *Biochemistry 38*, 12514-12525.

186. Sivaraman, S., Sullivan, T. J., Johnson, F., Novichenok, P., Cui, G., Simmerling, C., and Tonge, P. J. (2004) Inhibition of the Bacterial Enoyl Reductase FabI by Triclosan: A Structure-Reactivity Analysis of FabI Inhibition by Triclosan Analogs, *J. Med. Chem. 47*, 509-518.

187. 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 Chem. Biol. 1*, 43-53.

188. Ramaswamy, S. V., Reich, R., Dou, S. J., Jasperse, L., Pan, X., Wanger, A., Quitugua, T., and Graviss, E. A. (2003) Single nucleotide polymorphisms in genes associated with isoniazid resistance in Mycobacterium tuberculosis, *Antimicrobial Agents and Chemotherapy 47*, 1241-1250.

189. Banerjee, A., Dubnau, E., Quemard, A., Balasubramanian, V., Um, K. S., Wilson, T., Collins, D., Delisle, G., and Jacobs, W. R. (1994) Inha, a Gene Encoding a Target for Isoniazid and Ethionamide in Mycobacterium-Tuberculosis, *Science 263*, 227-230.

190. Johnsson, K., and Schultz, P. G. (1994) Mechanistic Studies of the Oxidation of Isoniazid by the Catalase Peroxidase from Mycobacterium-Tuberculosis, *Journal of the American Chemical Society 116*, 7425-7426.

191. Johnsson, K., King, D. S., and Schultz, P. G. (1995) Studies on the Mechanism of Action of Isoniazid and Ethionamide in the Chemotherapy of Tuberculosis, *Journal of the American Chemical Society 117*, 5009-5010.

192. Quemard, A., Sacchettini, J. C., Dessen, A., Vilcheze, C., Bittman, R., Jacobs, W. R., and Blanchard, J. S. (1995) Enzymatic Characterization of the Target for Isoniazid in Mycobacterium-Tuberculosis, *Biochemistry 34*, 8235-8241.

193. Quemard, A., Dessen, A., Sugantino, M., Jacobs, W. R., Sacchettini, J. C., and Blanchard, J. S. (1996) Binding of catalase-peroxidase-activated isoniazid to wild-type and mutant Mycobacterium tuberculosis enoyl-ACP reductases, *Journal of the American Chemical Society 118*, 1561-1562.

194. Dessen, A., Quemard, A., Blanchard, J. S., Jacobs, W. R., and Sacchettini, J. C. (1995) Crystal-Structure and Function of the Isoniazid Target of Mycobacterium-Tuberculosis, *Science 267*, 1638-1641.

195. Basso, L. A., Zheng, R. J., 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.

196. Vilcheze, C., Morbidoni, H. R., Weisbrod, T. R., Iwamoto, H., Kuo, M., Sacchettini, J. C., and Jacobs, W. R. (2000) Inactivation of the inhA-encoded fatty acid synthase II (FASII) enoyl-acyl carrier protein reductase induces accumulation of the FASI end products and cell lysis of Mycobacterium smegmatis, *Journal of Bacteriology 182*, 4059-4067.

197. Mdluli, K., Slayden, R. A., Zhu, Y. Q., Ramaswamy, S., Pan, X., Mead, D., Crane, D. D., Musser, J. M., and Barry, C. E. (1998) Inhibition of a Mycobacterium tuberculosis beta-ketoacyl ACP synthase by isoniazid, *Science 280*, 1607-1610.

198. Kremer, L., Dover, L. G., Morbidoni, H. R., Vilcheze, C., Maughan, W. N., Baulard, A., Tu, S. C., Honore, N., Deretic, V., Sacchettini, J. C., Locht, C., Jacobs, W. R., and Besra, G. S. (2003) Inhibition of InhA activity, but not KasA activity, induces formation of a KasA-containing complex in mycobacteria, *Journal of Biological Chemistry 278*, 20547-20554.

199. Slayden, R. A., and Barry, C. E. (2002) The role of KasA and KasB in the biosynthesis of meromycolic acids and isoniazid resistance in Mycobacterium tuberculosis, *Tuberculosis 82*, 149-160.

200. Slayden, R. A., Lee, R. E., and Barry, C. E. (2000) Isoniazid affects multiple components of the type II fatty acid synthase system of Mycobacterium tuberculosis, *Molecular Microbiology 38*, 514-525.

201. Slayden, R. A., and Barry, C. E. (2000) The genetics and biochemistry of isoniazid resistance in Mycobacterium tuberculosis, *Microbes and Infection 2*, 659-669.

202. 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.

203. Perozzo, R., Kuo, M., Sidhu, A. S., Valiyaveettil, J. T., Bittman, R., Jacobs, W. R., Jr., Fidock, D. A., and Sacchettini, J. C. (2002) Structural elucidation of the specificity of the antibacterial agent triclosan for malarial enoyl acyl carrier protein reductase, *J Biol Chem 277*, 13106-13114.

204. Muralidharan, J., Suguna, K., Surolia, A., and Surolia, N. (2003) Exploring the interaction energies for the binding of hydroxydiphenyl ethers to enoyl-acyl carrier protein reductases, *J Biomol Struct Dyn 20*, 589-594.

205. Layten, M., Hornak, V., and Simmerling, C. (2006) The open structure of a multi-drug-resistant HIV-1 protease is stabilized by crystal packing contacts, *J Am Chem Soc 128*, 13360-13361.