Stony Brook University

The official electronic file of this thesis or dissertation is maintained by the University Libraries on behalf of The Graduate School at Stony Brook University.

Structural Mechanisms of Catalysis and Inhibition of Enzymatic Drug Targets

Involved in Fatty Acids and Menaquinone Biosyntheses of *Mycobacterium tuberculosis*

A Dissertation Presented

by

Huei-Jiun Li

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Doctor of Philosophy

in

Chemistry

Stony Brook University

May 2012

Stony Brook University

The Graduate School

Huei-Jiun Li

We, the dissertation committee for the above candidate for the

Doctor of Philosophy degree, hereby recommend

acceptance of this dissertation.

Peter J. Tonge – Dissertation Advisor
Professor of Chemistry Department

Francis Johnson - Chairperson of Defense
Professor of Chemistry Department

Frank W. Fowler – Committee Member of Defense
Professor of Chemistry Department

Markus Seeliger – External Committee Member of Defense
Assistant Professor of Pharmacological Sciences Department, Stony Brook University

This dissertation is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

ii

Abstract of the Dissertation

Structural Mechanisms of Catalysis and Inhibition of Enzymatic Drug Targets

Involved in Fatty Acids and Menaquinone Biosyntheses of *Mycobacterium tuberculosis*

by

Huei-Jiun Li

Doctor of Philosophy

in

Chemistry

Stony Brook University

2012

Slow-onset inhibition is observed for many successful drugs on the market presumably since this leads to a better pharmacodynamic profile. However, the structural basis for slow-onset inhibition is largely unexplored and there is no rational rule for developing such kinetic properties. In our efforts to combat *M. tuberculosis* through the inhibition of its FAS-II enoyl-ACP reductase InhA, we have discovered a diphenyl ether scaffold that exhibits slow-onset inhibition, and used this as a model system to uncover the nature of the slow step that occurs on the timescale of minutes and longer. Previously it was found that slow-onset inhibition correlated to the ordering of helix-6 on InhA. In the present work, from the conformational space covered by a large number of crystal structures, separate conformational states of InhA are identified and the energy barrier revealed by two-dimensional energy profiles. The size of the barrier and relative stabilities of the two major conformational states rationalize observations in kinetic experiments and crystal structures, supporting that slow-onset inhibitors initially bind with an InhA conformational state competent in binding the natural substrate, the long chain fatty acyl-carrier protein, and the complex subsequently undergoes isomerization to an otherwise naturally unfavorable conformer of InhA. Analysis of crystal structures along the binding reaction coordinate reveals that the helix-6 ordering event is the consequence of a large-scale local refolding process involving at least 30 residues induced by the interactions with the inhibitor.

MenB, the 1,4-dihydroxy-2-naphthoyl-CoA synthase from the bacterial menaquinone biosynthesis pathway, catalyzes an intramolecular Claisen condensation (Dieckmann reaction) in which the electrophile is an unactivated carboxylic acid. Mechanistic studies on this crotonase family member have been hindered by partial active site disorder in existing MenB X-ray structures. In the

current work the 2.0 Å structure of O-succinylbenzoyl-aminoCoA (OSB-NCoA) bound to the MenB from *Escherichia coli* provides important insight into the catalytic mechanism by revealing the position of all active site residues. This has been accomplished by the use of a stable analogue of the O-succinylbenzoyl-CoA (OSB-CoA) substrate in which the CoA thiol has been replaced by an amine. The resulting OSB-NCoA is stable and the X-ray structure of this molecule bound to MenB reveals the structure of the enzyme-substrate complex poised for carbon-carbon bond formation. The structural data support a mechanism in which two conserved active site Tyr residues, Y97 and Y258, participate directly in the intramolecular transfer of the substrate α-proton to the benzylic carboxylate of the substrate, leading to protonation of the electrophile and formation of the required carbanion. Y97 and Y258 are also ideally positioned to function as the second oxyanion hole required for stabilization of the tetrahedral intermediate formed during carbon-carbon bond formation. In contrast, D163, which is structurally homologous to the acid-base catalyst E144 in crotonase, is not directly involved in carbanion formation and may instead play a structural role by stabilizing the loop that carries Y97. When similar studies were performed on the MenB from *Mycobacterium tuberculosis*, a twisted hexamer was unexpectedly observed, demonstrating the flexibility of the interfacial loops that are involved in the generation of the novel tertiary and quaternary structures found in the crotonase superfamily. This work reinforces the utility of using a stable substrate analogue as a mechanistic probe in which only one atom has been altered leading to a decrease in α-proton acidity.

The designated *trpE* gene (*Rv1609*) from *Mycobacterium tuberculosis* was expressed in *E. coli* and characterized. While TrpE displays normal NH_4^+ and Mg^{2+} dependence and inhibition by tryptophan, as expected for an anthranilate synthase, its ability to produce anthranilate is affected by contaminating enzymes that display chorismate mutase and prephenate dehydratase activities. Introduction of the *trpG* (*Rv0013*) gene product into the reaction mixture preferably partitions chorismate into the direction of anthranilate production. Kinetic analysis demonstrates that the improved efficiency of anthranilate production is achieved by greater affinity of TrpE for chorismate and also an increase in k_{cat} for the reaction. Intriguingly, there is no evidence for a tight complex between TrpE and TrpG, unlike other characterized anthranilate synthases that exhibit cooperativity. Moreover, while AS activity was optimized, the reaction intermediate ADIC was found to accumulate although other anthranilate synthases are not known to accumulate ADIC. The study reveals the importance of complex formation for AS catalysis and again raises the long standing question of how anthranilate synthases catalyze pyruvate elimination.

Table of Contents

List of Figures

xi

List of Tables

List of Abbreviations

ACP	Acyl carrier protein
AcpM	Mycobacterial acyl carrier protein
ADC	4-amino-4-deoxychorismate
ADIC	2-amino-2-deoxyisochorismate
ADCS	ADC synthase
ADICS	ADIC synthase
AS	Anthranilate synthase
ASU	Asymmetric unit
CoA	Coenzyme A
CM	Chorismate mutase
CS	Crotonase superfamily
DCI	Dienoyl-CoA isomerase
DHNA-CoA	1,4-dihydroxy-2-naphthoyl-CoA
DOTS	Directly observed treatment-short course
ecFabI	FabI from *Escherichia coli*
ECH	Enoyl-CoA hydratase
ECI	Δ^3- Δ^2-enoyl-CoA isomerase
EMB	Ethambutol
ETH	Ethionamide
FabI	Enoyl-ACP reductase
FAS-II	Type II fatty acid synthase
ftuFabI	FabI from *Francisella tularensis*
HCHL	Hydroxycinnamoyl-CoA hydratase-lyase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
ICS	isochorismate synthase
INH	Isoniazid
InhA	Enoyl-ACP reductase from *Mycobacterium tuberculosis*
MDR	Multi-drug resistant
MenB	1,4-dihydroxy-2-naphthoyl-CoA synthase
MenF	Menaquinone-specific isochorismate synthase
MMCD	Methylmalonyl CoA decarboxylase
MPD	2-methyl-2,4-pentandiol

mtMenB	MenB from *Mycobacterium tuberculosis*
M. tuberculosis	*Mycobacterium tuberculosis*
NADH	Nicotinamide adenine dinucleotide, reduced form
NAD$^+$	Nicotinamide adenine dinucleotide, oxidized form
PAS	*para*-aminosalicylic acid
PDB	Protein data bank
PDT	Prephenate dehydratase
PTH	Prothionamide
RIF	Rifampicin
RMSD	Root mean square deviation
SBL	Substrate-binding loop
SS	Salicylate synthase
TB	Tuberculosis

Chapter 1 Targeting fatty acid and menaquinone biosynthesis pathways in *Mycobacterium tuberculosis*

1.1 Tuberculosis and treatments

Tuberculosis is a global health threat

Mycobacterium tuberculosis is the causative agent of tuberculosis (TB), the disease that most commonly affects the lungs and is transmitted from person to person via droplets from the throat and lungs of people with the active respiratory disease. More than two billion people, or one third of the world's total population, are infected with the TB bacillus [1], which may develop into two types of TB. The healthy immune system of the infected people protects them from developing any symptoms or becoming infectious (latent TB). Others with active TB in the lungs can show symptoms such as coughing with sputum and blood, chest pains, weakness, weight loss, fever and night sweats, and they also become infectious (active TB) [2]. TB caused 1.45 million deaths globally in 2010, and 8.8 million incident cases occurred in Asia (59%), Africa (26%), the Eastern Mediterranean region (7%), the European region (5%), and the region of the Americas (3%) [3,4]. TB remains one of the top ten causes of death worldwide [5].

Antibiotic treatment of drug-susceptible TB

TB is an ancient disease [6-8], but the modern treatments only began with the discovery of the first effective TB drug, streptomycin [9], in 1944, followed by others such as para-aminosalicylic acid (PAS), isoniazid (INH), pyrazinamide (PZA), ethionamide (ETH), prothionamide (PTH), ethambutol (EMB), and rifampicin (RIF) in the 1950s to 60s (**Figure 1.1**) [10-14]. The current standard therapy introduced in the mid-1990s, the directly observed treatment—short course (DOTS), is a two-phase, six-month course of antibiotic treatment. The first two months treated with the four drugs, INH, RIF, PZA and EMB, are followed by the continued treatment of INH and RIF for another four months [15]. This regimen recommended by World Health Organization for new cases of drug-susceptible-TB has cure rates of around 90% in HIV-negative patients.
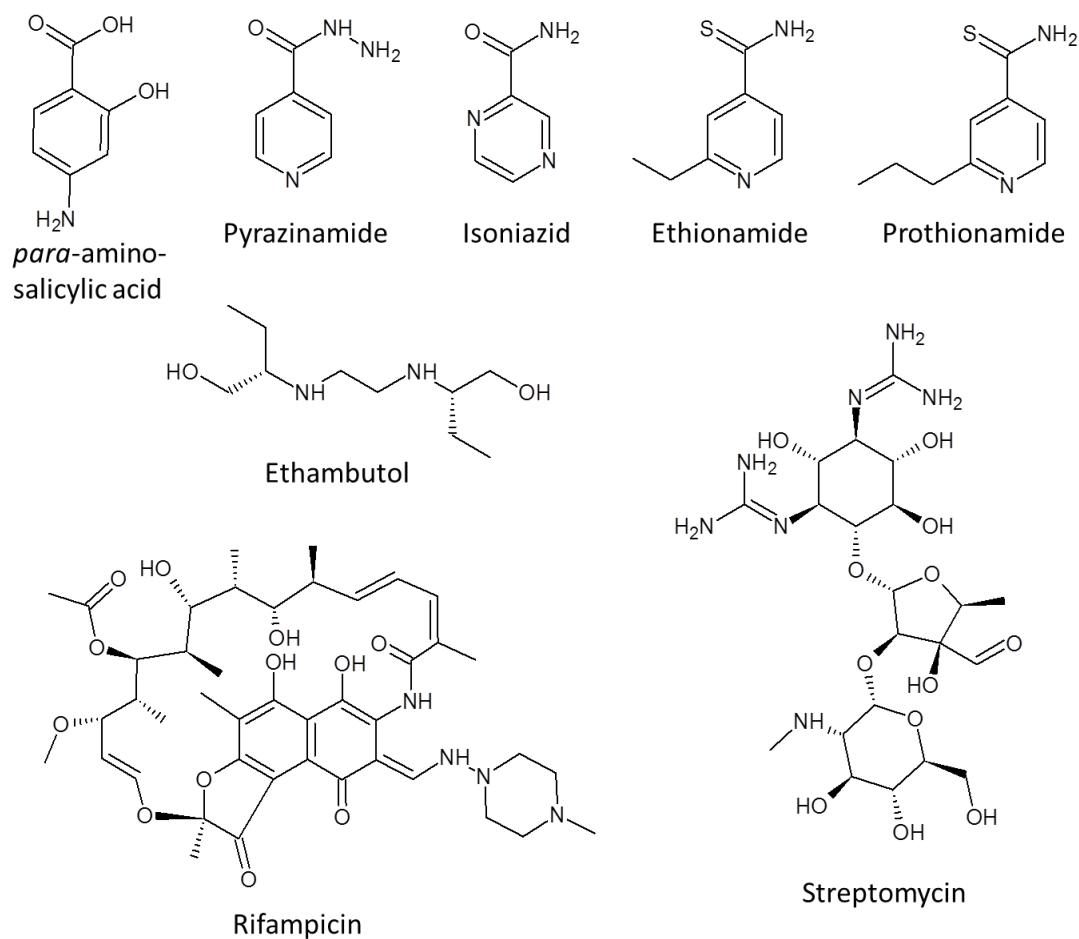
Figure 1.1 Structures of current anti-TB drugs

Treatment of drug-resistant TB

However, multi-drug resistant TB (MDR-TB) that is resistant to at least two front-line drugs, INH and RIF, has emerged due to the length of treatment, difficulty in patient compliance, improper diagnosis and improper use of drugs, which lowers the success rate of DOTS. It is estimated that 15% of previously treated patients develop MDR at the global level [16,17]. The treatment becomes even longer, requiring at least 20 months, and involves second-line drugs that produce multiple side effects. Drugs used for MDR-TB treatment are grouped according to efficacy, experience of use and drug class (**Table 1.1**) [15,18]. The cure rates are lower, usually in the range 60–75% [3].

Table 1.1 Groups of drugs used to treat MDR-TB

Group	Drugs (abbreviations)
Group 1: First-line oral agents	• pyrazinamide (Z) • ethambutol (E) • rifabutin (Rfb)
Group 2: Injectable agents	• kanamycin (Km) • amikacin (Am) • capreomycin (Cm) • streptomycin (S)
Group 3: Fluoroquinolones	• levofloxacin (Lfx) • moxifloxacin (Mfx) • ofloxacin (Ofx)
Group 4: Oral bacteriostatic second-line agents	• para-aminosalicylic acid (PAS) • cycloserine (Cs) • terizidone (Trd) • ethionamide (Eto) • protionamide (Pto)
Group 5: Agents with unclear role in treatment of drug resistant-TB	• clofazimine (Cfz) • linezolid (Lzd) • amoxicillin/clavulanate (Amx/Clv) • thioacetazone (Thz) • imipenem/cilastatin (Ipm/Cln) • high-dose isoniazid (high-dose H) • clarithromycin (Clr)

The need for new drugs

The development of new diagnostics, new drugs and vaccines are still needed to achieve the goal of TB elimination [19]. New drugs are required to address current concerns including, the length and complexity of treatment, the efficacy and tolerability of treatment for MDR-TB and co-infection and co-treatment with HIV. In addition, new drugs could also help to treat latent TB infection in people without active TB disease [20]. The current preventive therapy usually consists of 6–9 months of treatment using INH [3]. Within this context, our research will focus on the discovery stage along the pipeline for developing new anti-TB drugs (**Figure 1.2**).
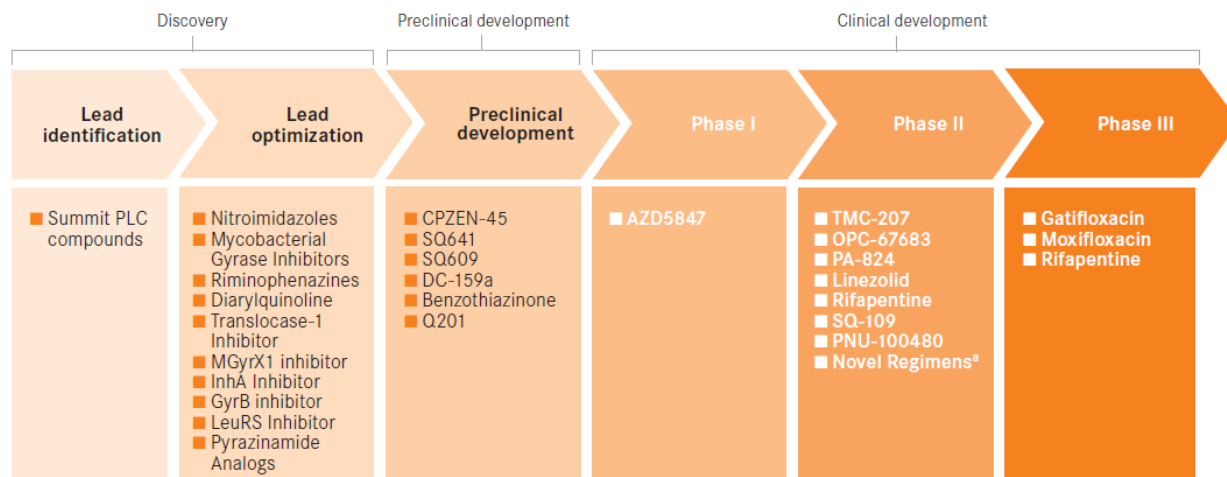
Figure 1.2 The status of the pipeline for new anti-TB drugs in July 2011

Note: This table only includes projects that have identified a promising molecule (known as a "lead" compound).

a: The first clinical trial (NC001) of a novel TB drug regimen testing the three-drug combination of PA-824, moxifloxacin and pyrazinamide was initiated in November 2010.

Source: WHO Global tuberculosis control 2011[3]

1.2 Targeting mycobacterial cell wall biosynthesis

Cell wall biosynthesis as a drug target

Among the TB drugs in use, INH, ETH, PTH, EMB and cycloserine target cell wall biosynthesis [21-34], RIF and fluoroquinolones target RNA and DNA synthesis [35,36], and streptomycin inhibits protein synthesis [37-39]. In more recent drug development efforts, enzymes involved in mycobacterial cell wall biosynthesis are still frequent targets [40-43]. The cell wall of mycobacteria has a multi-layered and extremely hydrophobic structure that prevents penetration of many hydrophilic antibiotics (**Figure 1.3**). There are two major components of lipopolysaccharides, the arabinogalactan (AG) esterified with mycolic acids (mAG), and lipoarabinomannans (LAM), both essential for mycobacterial survival. The biosynthesis of mycolic acids, the very long C60-C90 fatty acids, is the target of INH.
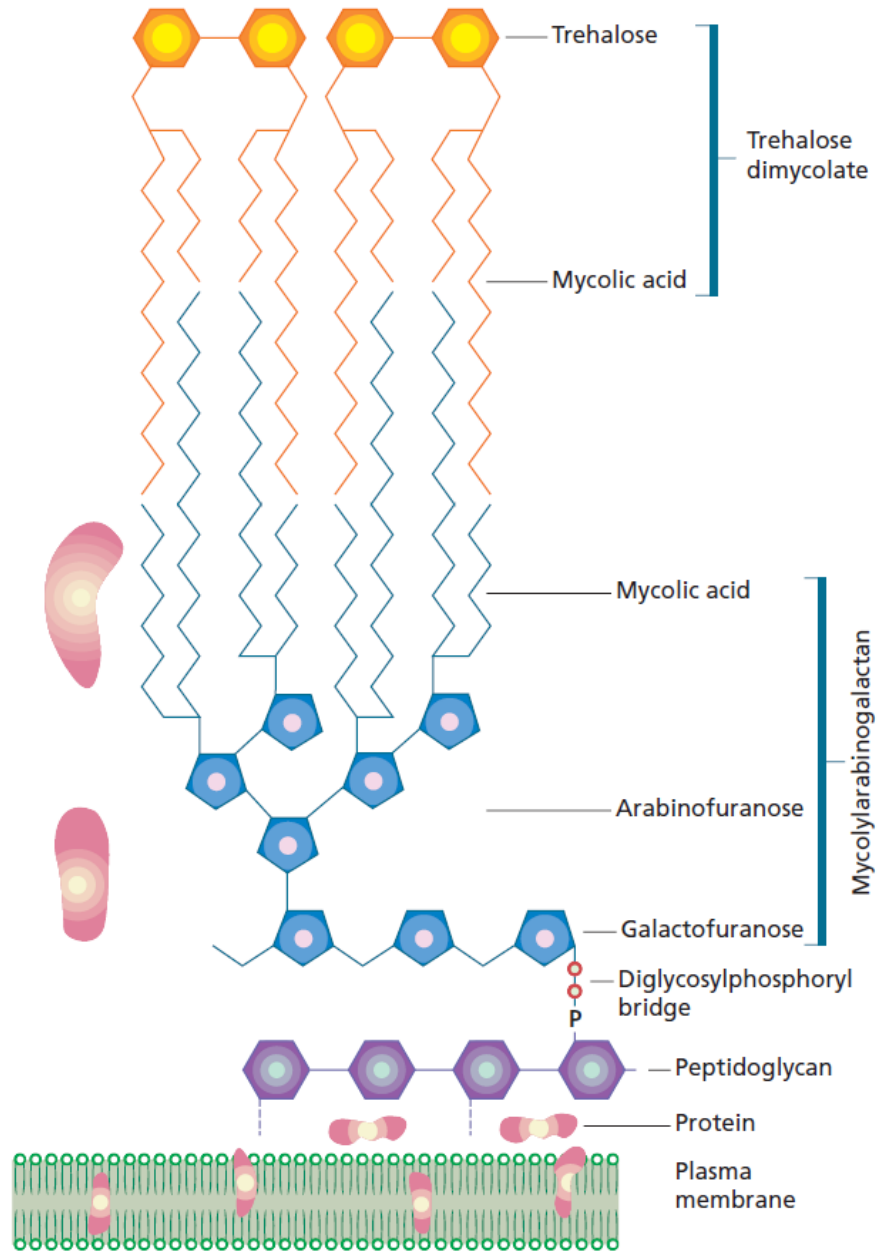
Figure 1.3 Composition of the mycobacterial cell wall
Chain length is not to scale. Figure taken from Ref [40,44].

INH targets mycolic acid and fatty acid biosynthesis

INH is a pro drug that is activated by the mycobacterial catalase-peroxidase KatG and forms an adduct with NADH that inhibits InhA, the 2-*trans*-enoyl-ACP reductase in the type II fatty acid synthase system (FAS-II) of mycobacteria [24,45]. The FAS-II system is distinguished from FAS-I because FAS-II is composed of separate polypeptides each possessing individual activities (**Figure 1.4**) whereas FAS-I is a single polypeptide with multiple enzyme activities. While in general FAS-I is found in mammals and FAS-II found in plants and bacteria [46-51], mycobacteria possess both FAS-I and FAS-II. FAS-I likely makes fatty acids up to C26, which are then extended up to C56 by FAS-II [52-54]. These fatty acids then serve as the precursors of mycolic acids, the C54-C63 β-hydroxy fatty acids with a C22-C24 α-alkyl side chain [55-57]. Hence the inhibition of mycolic acid biosynthesis by INH results from the inhibition of fatty acid biosynthesis. The FAS-II system, in general, has been demonstrated to be an attractive target for novel antibacterial agents [58-62].
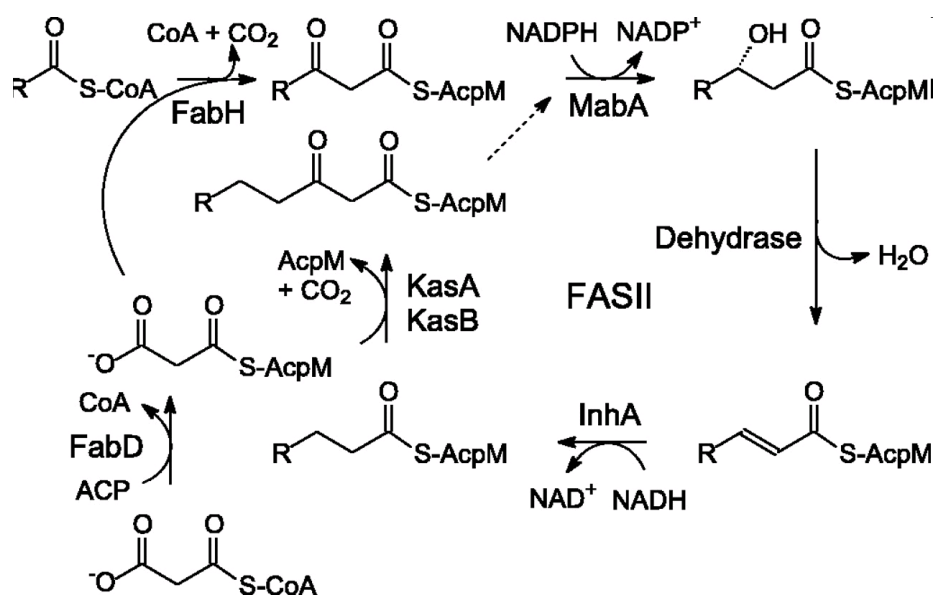
Figure 1.4 The mycobacterial FAS-II pathway

FAS-II enoyl-ACP reductase, InhA

The target of INH, InhA, catalyzes the NADH-dependent reduction of the *trans*-2 double bond on fatty acyl substrates [63]. InhA has been shown to prefer chain lengths of C12 to C24 [63] and up to C30 has been observed [64]. Activities for chain lengths as long as C56 and as short as C4 have also been suggested [65,66]. Although InhA is active with coenzyme A-based substrates *in vitro*, the *in vivo* fatty acyl substrates are carried by the mycobacterial acyl carrier protein (AcpM) [57,67]. The homologue of InhA in the FAS-II system of other bacteria, FabI, is also found to

be the target of known antibiotics diazaborines and triclosan, making this enzyme a target of intensive interest in the FAS-II pathway [40,62,68-70].

1.3 Mycobacterial respiration

Respiration in persistent *M. tuberculosis*

Drugs targeting cell wall biosynthesis are primarily active against the growing bacilli [71]. However, there are other bacterial populations with diverse metabolic status in lesions that are not killed by INH. Some are dormant and not killed by any current TB drugs [14,72-76]. Hence there have been attentions in discovering inhibitors that target the respiratory pathway (**Figure 1.5**) thought to be essential for the non-replicating and persistent bacilli [77-86].
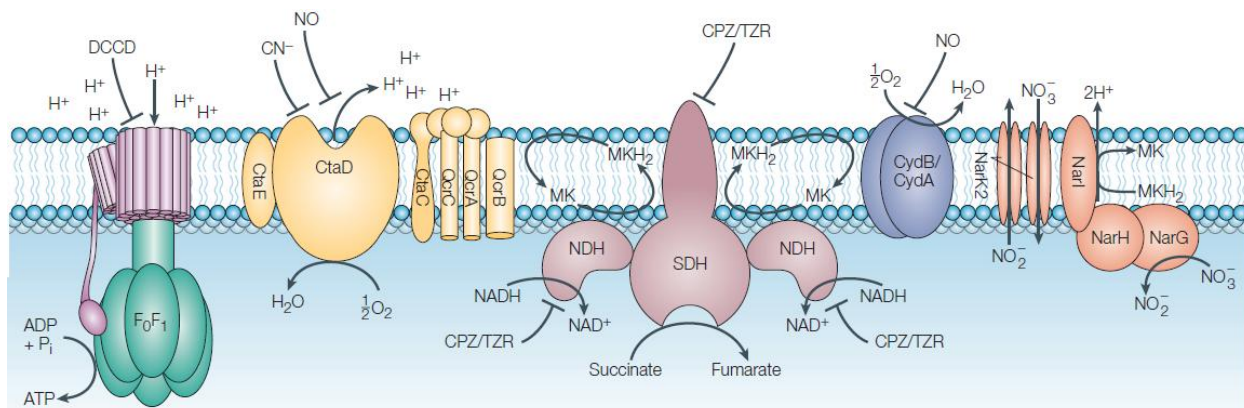
Figure 1.5   Components of the electron-transport chain leading to ATP synthesis in the respiration of *Mycobacterium tuberculosis* under limiting oxygen conditions

Some respiratory inhibitors that have been used to characterize genetic control of respiratory metabolism are shown according to their known site of action. CN⁻, cyanide; NO, nitric oxide; CPZ, chlorpromazine; DCCD, dicyclohexylcarbodiimide; TZR, thioridazine.
NDH: type II non-proton pumping NADH dehydrogenases; SDH: succinate dehydrogenase; NarK2: nitrite-efflux system; NarGHJI: nitrate reductase; CydAB: cytochrome bd-type menaquinol oxidase; MK: menaquinone [76,87-92]. Figure taken from Ref [76].

Targeting menaquinone biosynthesis

M. tuberculosis uses menaquinone, a polyisoprenylated naphthoquinone (**Figure 1.6**), as a lipid-soluble redox cofactor in its electron transport chain (**Figure 1.5**). Quinones are central to the electron transport chain for all kingdoms of life, but the specific quinone utilized depends

on species [93]. Menaquinone is employed in most Gram positive and some Gram negative bacteria [94,95]. In plants and cyanobacteria, a closely-related phylloquinone (**Figure 1.6**), serves as an electron carrier in photosysthesis [96]. Ubiquinone, a benzoquinone, is used in eukaryotes and some prokaryotes [97]. The biosyntheses of menaquinone and phylloquinone are similar while that of ubiquinone is not homologous [98-103]. Although humans utilize menaquinone (vitamin K2) and phylloquinone (vitamin K_1) for the γ-carboxylation of glutamate residues [104,105], mammalian cells are unable to undertake the de novo synthesis of menaquinone, and thus bacterial enzymes in the menaquinone biosynthesis pathway are potential targets for antibacterial drug discovery [106-109]. Menaquinone is believed to be the only quinone in *M. tuberculosis*, and the biosynthes pathway is therefore a promising drug target [110-117].

The menaquinone biosynthesis pathway

The structure of menaquinone consists of a "head group" with polar substituents and a non-polar isoprenoid tail [99]. The distribution of chain lengths and the degrees of saturation in the isoprenoid tail in different bacterial species vary [94]. Menaquinone-9 with the second isoprene unit saturated is the most abundant quinone in mycobacteria [118]. Sulfur-containing menaquinones are also found in species including *M. tuberculosis* [119,120].

The classical biosynthesis pathway of menaquinone from chorismate was originally elucidated in *Escherichia coli*, *Bacillus subtilis* and *Mycobacterium phlei* [121,122] and recently refined by Jiang et al. [123,124] (**Figure 1.6**). After branching from chorismate which is synthesized by the shikimate pathway, a succinyl group is attached by the action of MenD, and the first aromatic ring of menaquinone furnished by MenC, the o-succinylbenzoate synthase. Subsequently, the second ring is built up from the succinyl group by a ring-closure reaction catalyzed by MenB through the activation of the succinyl carboxylate catalyzed by MenE, and the restoration of the free carboxylate in the following step by a thioesterase that is yet to be identified. The isoprenyl group and methyl group are then attached in the last steps of the pathway. Although an alternative pathway through a futalosine intermediate has been discovered in *Streptomyces coelicolor* [125], *M. tuberculosis* contains only homologous genes in the classical pathway including *menA*, *menB*, *menC*, *menD*, *menE* and *menG* [126].
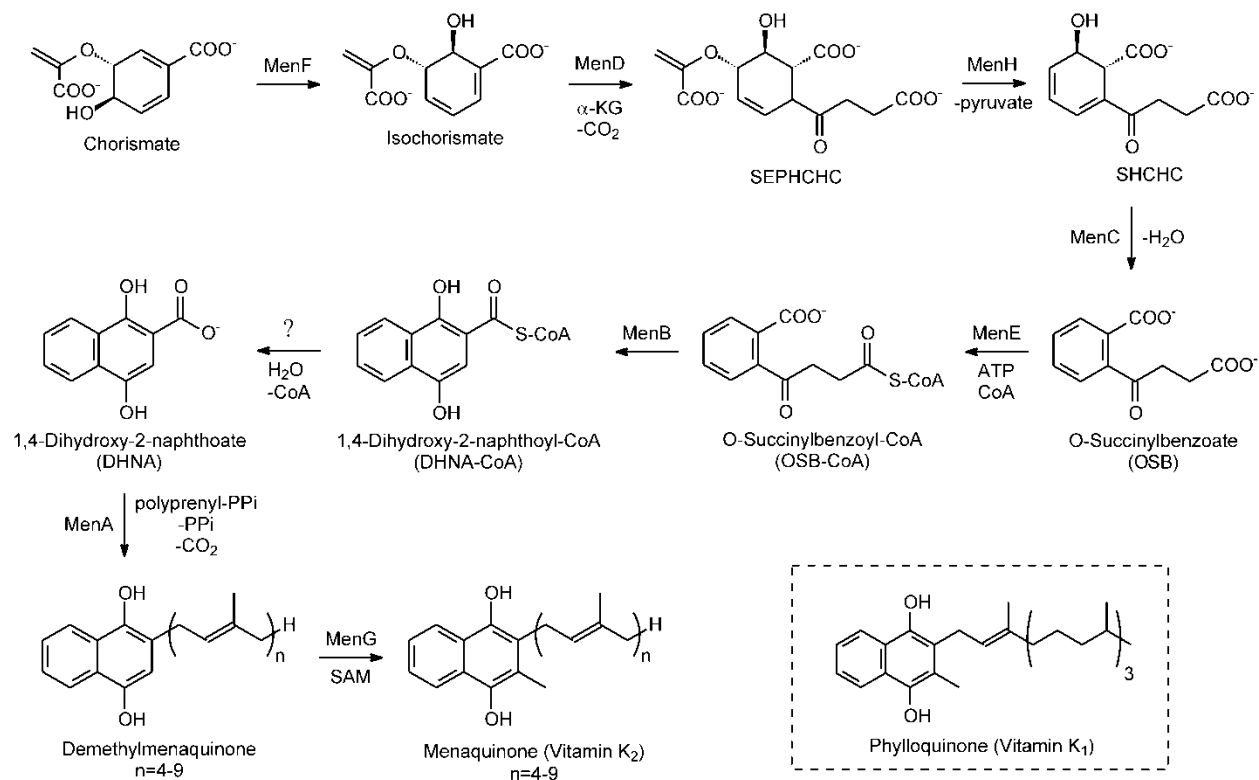
Figure 1.6 The menaquinone biosynthesis pathway and the structure of phylloquinone

1.4 Chorismate-utilizing enzymes

Menaquinone is one of several aromatic metabolites derived from chorismate (**Figure 1.7**), the product of shikimate pathway (**Figure 1.8**) [127-129]. Thus, in a cell five or more enzymes compete for chorismate. The shikimate pathway and chorismate-utilizing terminal pathways are attractive anti-TB and antimicrobial drug targets due to their central importance and their absence in humans [130-132]. The structural and mechanistic similarity among chorismate-utilizing enzymes provides the possibility for one drug to simultaneously inhibit multiple pathways and to increase the barrier for the bacteria to develop drug resistance [133-135].

9
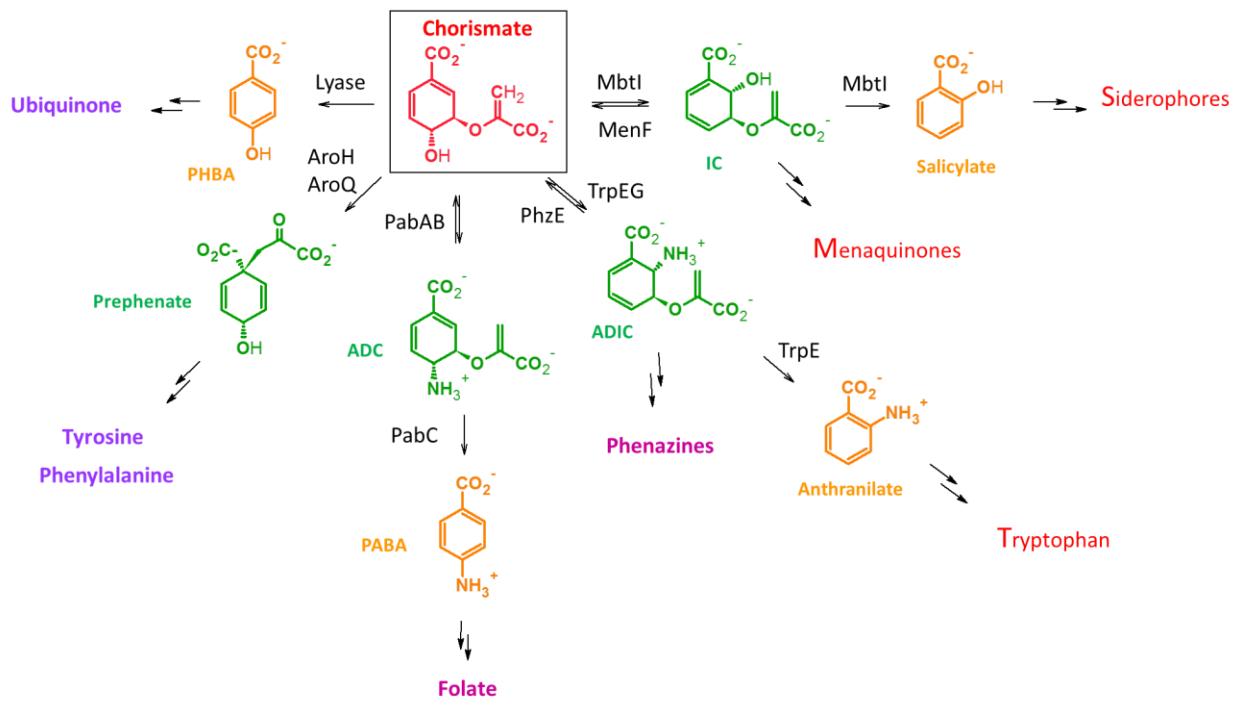
Figure 1.7 Metabolites and biosynthesis pathways derived from chorismate
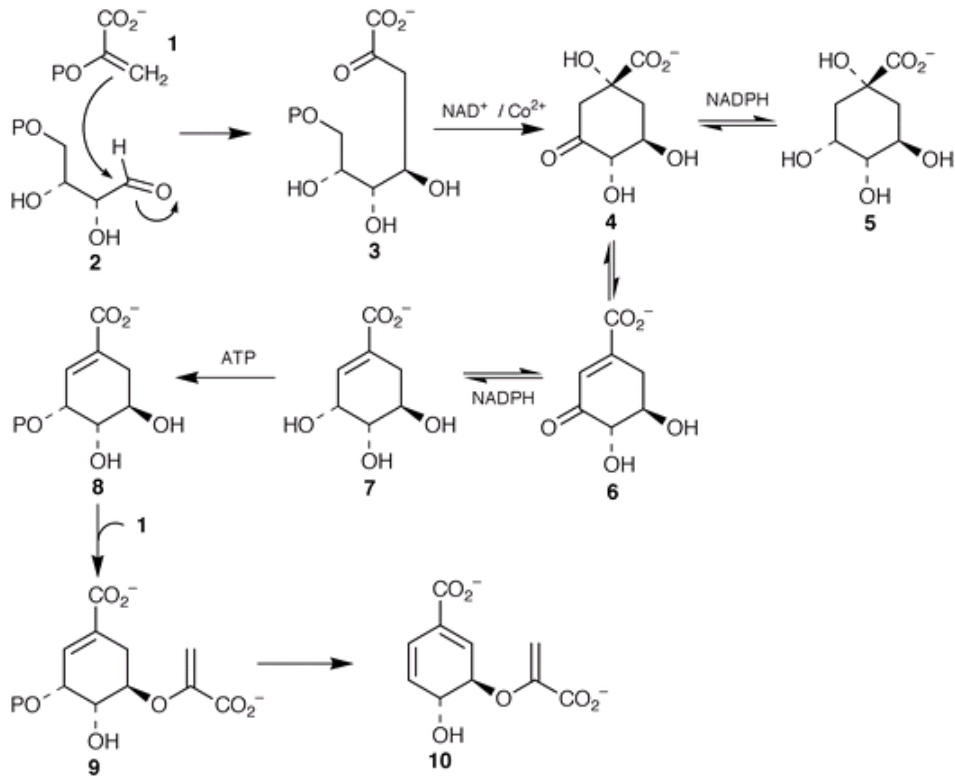
Figure 1.8 The shikimate pathway leading to the biosynthesis of chorismate (10)
Source: Ref [136]

The menaquinone-specific isochorismate synthase, MenF, identified in *E. coli* (ecMenF) is not annotated in mycobacterial genomes including that of *M. avium* [137-139]. The BLAST search for *M. tuberculosis* genes homologous to ecMenF returns four open reading frames, *entC*, *mbtI*, *trpE* and *pabB* with the highest sequence similarities (**Table 1.2**) [126]. MbtI exhibits isochorismate synthase (ICS) activity but further converts isochorismate to salicylate, and the gene is part of the highly regulated mycobactin operon that produces the essential iron chelator [140,141]. No ICS activity of the annotated EntC can be detected, and it is not known whether the annotated TrpE and PabB possess ICS activity.

Table 1.2 The open reading frames in *M. tuberculosis* with homology to MenF of *E. coli*

ORF	Annotation	Length (a.a.)	% Identity	% Similarity
Rv3215	EntC, unknown function	372	30 (79/262)	44 (117/262)
Rv2386c	MbtI, mycobactin biosynthesis	450	28 (77/267)	48 (123/267)
Rv1609	TrpE, tryptophan biosynthesis	516	23 (72/302)	43 (130/302)
Rv1005	PabB, folate biosynthesis	458	31 (60/192)	46 (90/192)

1.5 Research overview

In this thesis work, we have explored the possibilities of targeting these biosynthesis pathways in *M. tuberculosis*. Chapter 2 is focused on the FAS-II enoyl-ACP reductase, InhA, a well-validated drug target. A class of inhibitors, the diphenyl ethers, have been developed as an alternative, non-covalent solution to the covalent inhibition mechanism of action employed by INH. A number of diphenyl ethers have been found to display slow-onset inhibition kinetics, a property that can improve a drug's pharmacokinetic and pharmacodynamic profile *in vivo*. However, there is not yet a strong foundation of structural understanding that provides a clear guidance for further rational design of such kinetic property. We study the structural basis of slow-onset inhibition of InhA by diphenyl ethers and hope the principles can be applied to other inhibition systems.

MenB is the 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) synthase in the menaquinone biosynthesis pathway. Our group has demonstrated the activity of MenB inhibitors against replicating and nonreplicating *M. tuberculosis* [78,79]. However, unlike the InhA system where a wealth of structural information is available on substrate recognition and catalysis that has guided the inhibitor optimization efforts, structure-based design of MenB inhibitors is

hampered due to the lack of an intact active site in existing X-ray strucutres. In chapter 3, we overcome this difficulty and study the catalytic mechanism of MenB in more detail.

The identity of the isochorismate synthase for menaquinone production in *M. tuberculosis* is unclear. In chapter 4, we ask whether the two genes, *Rv1609* and *Rv1005*, annotated as *trpE* and *pabB*, respectively, could fulfill this function. The characterization then leads to further functional investigations that could stimulate the development of potential inhibitors targeting members in the enzyme family that all bind chorismate but catalyze distinct chemical transformations.

Chapter 2 Structural mechanism of slow-onset InhA inhibition by the diphenyl ethers

2.1 Introduction

Inhibition of FabI by diphenyl ethers — affinity and kinetics

The development of diphenyl ether inhibitors of InhA in the Tonge lab originated from the discovery of ecFabI inhibition by triclosan (**Figure 2.1**) [142–144], which binds preferentially to the E·NAD$^+$ form of the enzyme with picomolar affinity (**Table 2.1**). With respect to NADH, the inhibition kinetics appears competitive. The onset of inhibition is slow and the apparent association and dissociation rate constants (k_{on} and k_{off}) are 2.6×10^7 M^{-1}s^{-1} and 1.9×10^{-4} s^{-1} (dissociation $t_{1/2}$ 61 min), respectively [144–147]. This slow, tight, and specific inhibition of ecFabI accounts for the effective antibacterial activity of triclosan, which has been widely used in daily consumer products since the 1960s [148].
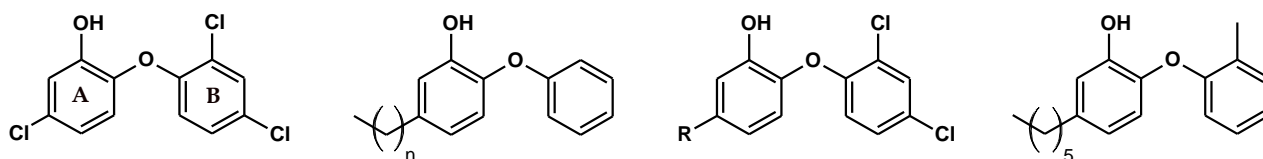
Figure 2.1 Structures of triclosan, 5-alkyl diphenyl ethers and PT70 (left to right)

FabIs from *Plasmodium falciparum*, *Staphylococcus aureus*, *Francisella tularensis* and *Burkholderia pseudomallei* are 24–64% identical, 49–80% similar, respectively, to ecFabI in sequence (calculated by BLAST 2) and triclosan has also been shown to inhibit these FabIs with nanomolar to picomolar affinity and slow-onset kinetics [149–155] (**Table 2.1**). In contrast, InhA from *M. tuberculosis* is 33% identical and 51% similar to ecFabI, but slow-onset inhibition by triclosan is not observed and the inhibition constant is a moderate value of 0.2 µM [156], i.e. four orders of magnitude worse in affinity compared to ecFabI. The 5-alkyl diphenyl ethers developed in the Tonge lab improve the binding affinity up to 1.1 nM (**Figure 2.1**, n=7) [157] while those developed by Freundlich et al. (**Figure 2.1**, B ring chlorines retained) improve the IC$_{50}$ up to 50 fold over that of triclosan [158]. Attaching a methyl group on the B ring (**Figure**

2.1, PT70) results in slow-onset inhibition, further improving the affinity to 22 pM. The dissociation rate constant for the inhibited complex is 7.2×10^{-4} s^{-1} (t$_{1/2}$ 17 min) [159]. Hence similar level of slow and tight inhibition as triclosan on ecFabI has been achieved by PT70 with respect to InhA.

Table 2.1 Comparison of FabI sequence from various organisms and kinetic properties of inhibition by triclosan

Identity and similarity are relative to ecFabI.

	% Identity	% Similarity	Slow-onset	k_{on} (M^{-1}s^{-1}) k_{off} (s^{-1})	K$_i$ (nM)
Plasmodium falciparum	24	49	Yes [149,150]	1.3×10^4 4.2×10^{-4} [151]	0.096 [150] K$_d$= 32 [151]
Mycobacterium tuberculosis	33	51	No [156]		220 [156,157]
	(PT70)		Yes	3.3×10^7 (calc.) 7.2×10^{-4}	0.022 [159]
Staphylococcus aureus	44	65	Yes [152,153]		4.9 [153]
Bacillus anthracis	47	67	ND		IC$_{50}$=600 [160]
Bacillus subtilis	52	69	Yes [161]		IC$_{50}$=16000 [161] (ecFabI: 2000)
Francisella tularensis	58	77	Yes [154]	8.2×10^6 (calc.) 4.2×10^{-4}	0.051 [154]
Burkholderia pseudomallei	64	80	Yes [155]	2.1×10^5 (calc.) 3.3×10^{-4}	1.6 [155]
Haemophilus influenzae	75	85	Yes [162]	1.9×10^4 2.9×10^{-2}	IC$_{50}$=100 [162] (no preincubation) K$_d$= 1500
Escherichia coli	100	100	Yes [144-147]	2.6×10^7 [144] 1.9×10^{-4} [144]	0.022 [144] 0.023 [146] 0.007 [147]

Slow-binding inhibition is frequently encountered in enzyme kinetics experiments, and such kinetic behavior is often exploited for probing enzymatic catalysis, or is desired as the basis for potent, selective, or effective *in vivo* chemotherapeutics [163,164]. The term "slow-binding inhibition" was first introduced in 1982 to identify the time-dependent feature of some inhibition kinetic observations in contrast to classical reversible inhibition where the assumption of rapid equilibrium between enzyme and enzyme-inhibitor complex is made in the mathematical treatment [165-167]. On the timescale of a common enzyme assay, the inhibition does not apparently reach its full potency instantaneously but instead takes seconds, minutes or longer to establish a steady state. Hence it is also referred to as slow-onset or time-dependent inhibition. However, development of more powerful methods has made it possible to monitor the transient phase of enzymatic reactions in fractions of a second, and it becomes relative what should be regarded as slow inhibition. Szedlacsek defined slow-binding inhibition by comparing with the timescale of enzyme-catalyzed substrate conversion. In this view, slow inhibition has to occur within the timescale of measurement of substrate conversion in order to observe the time dependency [167,168]. In practice, inhibitors with a broad range of association and dissociation rate constants (k_{on} and k_{off}) have been treated as slow-binding inhibitors. k_{off} is generally smaller than 1 s^{-1} and can be as small as 10^{-8} s^{-1}. k_{on} can be smaller than 1 M^{-1}s^{-1} but can also be larger than 10^7 M^{-1}s^{-1} which is almost diffusion controlled. As a result, a broad range of possible dissociation constants for a slow-binding inhibitor from sub-millimolar to sub-picomolar has been observed [165]. In other words, inhibitors traditionally considered slow-binding possess a k_{off} that is generally considered slow, but confusingly might possess a k_{on} that is considered fairly rapid. From this aspect, slow-binding inhibition can be defined as the enzyme-inhibitor binding process for which k_{off} is less than 1 s^{-1}. Time dependency of enzyme inhibitor association can still be observed in the case of a large k_{on} when the inhibitor concentration is sufficiently low (as the experimental condition employed for a high-affinity inhibition assay), given that the apparent first-order association rate constant is $k_{on}[I]$ [169].

Structural basis of slow-onset inhibition

Slow-onset inhibition kinetics is characterized by the curvature of the reaction progress curve before there is significant substrate depletion. During the same period, the reaction rates remain linear in the presence of classical inhibitors [165,167,169,170]. The slow-onset kinetics of InhA inhibition by PT70 is described by a two-step, induced-fit model in which k_2 and k_{-2} are small compared to k_1 and k_{-1} (**Figure 2.2**, mechanism a) [159], suggesting a slow

conformational change of the enzyme complex induced by the inhibitor following rapid initial binding [171–175]. This kinetic mechanism may be distinguished experimentally from the single-step mechanism [176–178] and the two-step, conformation-selection mechanism [179] (**Figure 2.2**, mechanism b and c) by the effect of inhibitor concentration on the initial velocity or the rate of approach to the steady state [167,176,179,180]. In other words, kinetic analysis of PT70 inhibition indicates there is a distinguishable intermediate enzyme-inhibitor species from the final complex, and the slow isomerization step occurs after rather than before inhibitor binding.

(a) $\quad E{\cdot}NAD^+ + I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E{\cdot}NAD^+{\cdot}I \underset{k_{-2}}{\overset{k_2}{\rightleftharpoons}} E^*{\cdot}NAD^+{\cdot}I$

(b) $\quad E{\cdot}NAD^+ + I \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E^*{\cdot}NAD^+{\cdot}I$

(c) $\quad E{\cdot}NAD^+ \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} E^*{\cdot}NAD^+ \underset{k_{-2}}{\overset{I \quad k_2}{\rightleftharpoons}} E^*{\cdot}NAD^+{\cdot}I$
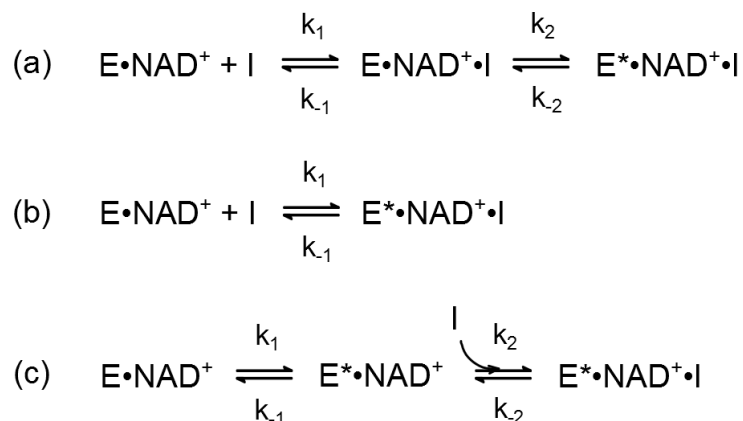
Figure 2.2 Kinetic mechanism of slow-onset InhA inhibition
Kinetic mechanism of slow-onset InhA inhibition by PT70, (a), and two other contrasting mechanisms for slow-onset inhibition, (b) and (c). In (a) and (c) the second and first step, respectively, is rate-limiting while in (b) the single step is slow.

Since the two-step, induced-fit mechanism is suggested to be the most common among the three kinetic models developed for slow-onset inhibition [175,180] and a conformational change is implied in the slow step, the structural studies of slow-onset inhibition naturally aim to identify such a conformational change. However, the large-amplitude conformational change is often not observed in the structures before and after inhibitor binding [163,181] and there seems still no developed general or systematic approach to undertake the task of elucidating the structural mechanism of slow-onset inhibition. Therefore it is difficult to infer from the studies of other slow-onset inhibition systems what must occur in the slow step of InhA inhibition by PT70.

Among the few attempts to generalize the structural observations in the transition from the initial to the final complex, Garvey suggests that the most general observation involves small protein conformational adjustments which could also be accompanied by covalent adduct formation with the inhibitor, or ionization of the inhibitor. Translocation of the inhibitor within the protein could also constitute the slow step [163]. Copeland suggests that these structural transitions tend to lead to greater occlusion of the binding pocket from bulk solvent due to predominant hydrophobic interactions [175]. Such occlusion often leads to ordering of structural elements of the protein over the inhibitor binding site. Both suggest that similar structural elements of the protein are often involved in recognition in both the initial and final complexes, but the interactions tend to be strengthened, or additional recognition elements will become engaged, in the final complex.

One example of large-amplitude conformational changes is found with the inhibitors of the epidermal growth factor receptor (EGFR). The slower inhibitor GW572016 with a residence time of 300 min binds to the inactive-like form of EGFR while the more rapid inhibitor with residence time < 10 min binds to the active-like form of EGFR. The change in the shape of ATP binding site and conformation of the A loop going from the active to inactive form of EGFR was speculated to be responsible for the difference in off rate [182].

The time-dependent inhibition of p38 MAP kinase by BIRB 796 was attributed to the conformational change in the Asp-Phe-Gly (DFG) motif. BIRB 796 has a k_{on} of 0.0051 nM^{-1} $\cdot min^{-1}$ and k_{off} of 0.00050 min^{-1}. Another diaryl urea inhibitor with lower affinity has a k_{on} of 0.0072 $nM^{-1} \cdot min^{-1}$ and k_{off} of 8.4 min^{-1}. Both bind to a "DFG-out" conformation of p38 MAP kinase while all other known Ser/Thr kinase structures bear the "DFG-in" conformation. It was suggested that infrequent sampling of the "DFG-out" conformation was the cause of time-dependent inhibition [183].

Oseltamivir carboxylate is a slow-binding inhibitor of influenza A/WS/33 neuraminidase with apparent k_{on} and k_{off} values of 0.035 $nM \cdot min^{-1}$ and 0.034 min^{-1}, respectively. Slow inhibition was abolished in the H274Y mutant but retained in the H274G, H274S, H274N, H274Q mutants, leading to the conclusion that the side chain volume of residue 274 would determine the binding kinetics of oseltamivir carboxylate. Based on the conformational change of Glu-276 observed in the active site of influenza A N9 neuraminidase, it was postulated that smaller side chains at residue 274 permit the conformational change of the neighboring Glu-276 to lead to slow inhibition [184]. The conformational change in this case is relatively subtle.

Structural investigation of slow-onset inhibition of FabI

The hint for the structural basis of the slow step of InhA inhibition by PT70 comes again from the structural studies of FabI inhibition by triclosan. Similar to FabI structures in complex with triclosan (**Table 2.2**), PT70 forms a ternary complex with InhA and NAD^+ in the crystal structure [159], consistent with the uncompetitive inhibition kinetics with respect to NAD^+. In all these structures, the A ring of the inhibitor stacks with the nicotinamide ring of the cofactor while the plane of the B ring is perpendicular to that of the A ring. The phenolic hydroxyl group on the A ring is hydrogen bonded to the phenolic hydroxyl group of a conserved active site Tyr and the 2'-hydroxyl group of the nicotinamide ribose, which in turn is hydrogen bonded to the bridging oxygen of the inhibitor and a conserved active site Lys (**Figure 2.3**, top). The signature active site Tyr, Lys and nicotinamide ribose involved in these interactions have been proposed to stabilize or protonate the enolate intermediate during catalysis [63,185-189], and the A ring of triclosan, especially if the phenol is ionized (pK_a=7.9), may mimic the reaction intermediate or the transition state [144,187,190-192].

Table 2.2 Crystal structures of FabI in complex with triclosan

PDB	Enzyme	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]	Ref
1QG6	ecFabI	P2$_1$	74.2	80.3	82.8	90.0	104.5	90.0	4	1.90	[144]
1QSG	ecFabI	P1	73.7	82.1	84.2	89.5	87.4	77.8	8	1.75	[187]
1C14	ecFabI	P6$_1$22	79.0	79.0	328.8	90.0	90.0	120.0	2	2.00	[190]
1D8A	ecFabI	P6$_1$22	80.6	80.6	327.6	90.0	90.0	120.0	2	2.20	[191]
N/A	ecFabI	P6$_1$22	79.7	79.7	327.7	90.0	90.0	120.0	2	1.80	[145]
3OIF	bsFabI	P2$_1$2$_1$2$_1$	64.1	83.7	203.6	90.0	90.0	90.0	4	2.60	[193]
2QIO	baFabI	P2$_1$2$_1$2$_1$	72.4	89.0	186.1	90.0	90.0	90.0	4	2.44	[160]
1D7O	bnFabI	P4$_2$2$_1$2	70.5	70.5	118.0	90.0	90.0	90.0	1	1.90	[192]
3NRC	ftuFabI	P2$_1$2$_1$2	123.1	85.1	51.8	90.0	90.0	90.0	2	2.10	[194]
2PD3	hpFabI	P2$_1$	73.4	94.9	75.4	90.0	106.2	90.0	4	2.50	[195]
1P45	InhA	I2$_1$2$_1$2$_1$	94.8	104.1	189.8	90.0	90.0	90.0	2	2.60	[196]
2B35	InhA	C2	100.0	81.8	188.7	90.0	95.7	90.0	6	2.30	[157]
3F4B	pbFabI	P2$_1$	64.3	121.2	87.8	90.0	109.0	90.0	4	2.49	U.P.
1NHG	pfFabI	P4$_3$2$_1$2	133.1	133.1	84.2	90.0	90.0	90.0	2	2.43	[197]
1UH5	pfFabI	P4$_3$2$_1$2	133.0	133.0	83.7	90.0	90.0	90.0	2	2.20	[198]
2O2Y	pfFabI	P2$_1$	88.2	82.4	94.8	90.0	90.8	90.0	4	2.20	[199]
3GR6	saFabI	P2$_1$2$_1$2$_1$	77.6	118.4	149.8	90.0	90.0	90.0	4	2.28	[200]
2WYW	ttFabI	P2$_1$	57.7	127.6	66.5	90.0	108.2	90.0	4	1.90	U.P.
2O2S	tgFabI	P3$_2$21	78.1	78.1	188.0	90.0	90.0	120.0	2	2.60	[199]

bs: *Bacillus subtilis*; ba; *Bacillus anthracis*; bn: *Brassica napus*; ftu: *Francisella tularensis*; hp: *Helicobacter pylori*; InhA: *Mycobacterium tuberculosis*; pb: *Plasmodium berghei*; pf: *Plasmodium falciparum*; sa: *Staphylococcus aureus*; tt: *Thermus thermophiles*; tg: *Toxoplasma gondii*

1. N: Number of subunits in the asymmetric unit
2. R: resolution
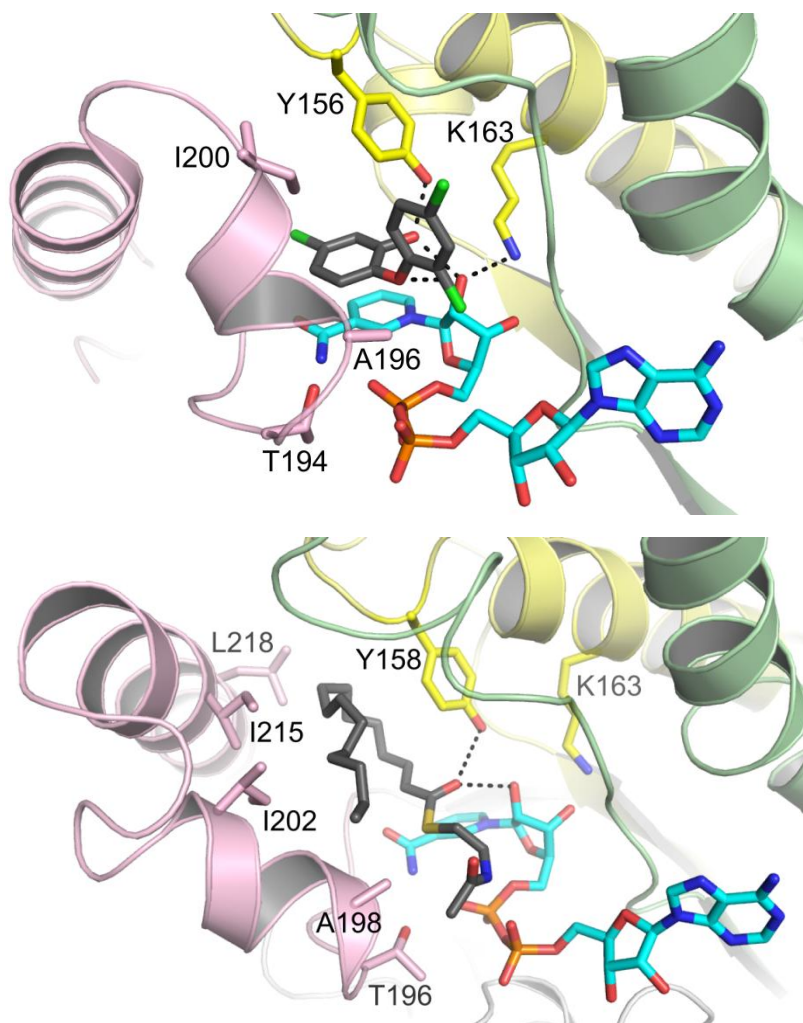
Figure 2.3 Triclosan and C16-NAC bound in ecFabI and InhA

Top: ecFabI bound with NAD$^+$ and triclosan (PDB 1QSG). Bottom: InhA bound with NAD$^+$ and substrate analogue, C16-NAC (PDB 1BVR), with the substrate-binding loop shown in pink

For ecFabI, a loop region (the pink loop carrying T194, A196 and I200 in **Figure 2.3**) that is disordered in the binary complex bound with the cofactor (PDB 1DFI [186,201]), becomes ordered in the ternary crystal structures obtained from co-crystallization with triclosan, and establishes extensive van der Waals contacts with the inhibitor [144,187,190,191]. The same loop has also been found to bind the hydrophobic acyl chain of the substrate [188] (**Figure 2.3**, bottom). Like ecFabI, this "substrate-binding loop" is ordered and makes close contacts with the inhibitor when the slow-onset PT70 is bound to InhA [159] (**Figure 2.4**, left). In contrast, it is disordered in the ternary complex structures with triclosan, PT03 or PT05 (Figure 2., n = 4

20

and 7) for which slow-onset inhibition is not observed (**Figure 2.4**, right) [157]. These observations point to the possibility that the slow inhibition kinetics of either InhA or FabI can be attributed to the slow ordering of the substrate-binding loop (SBL).
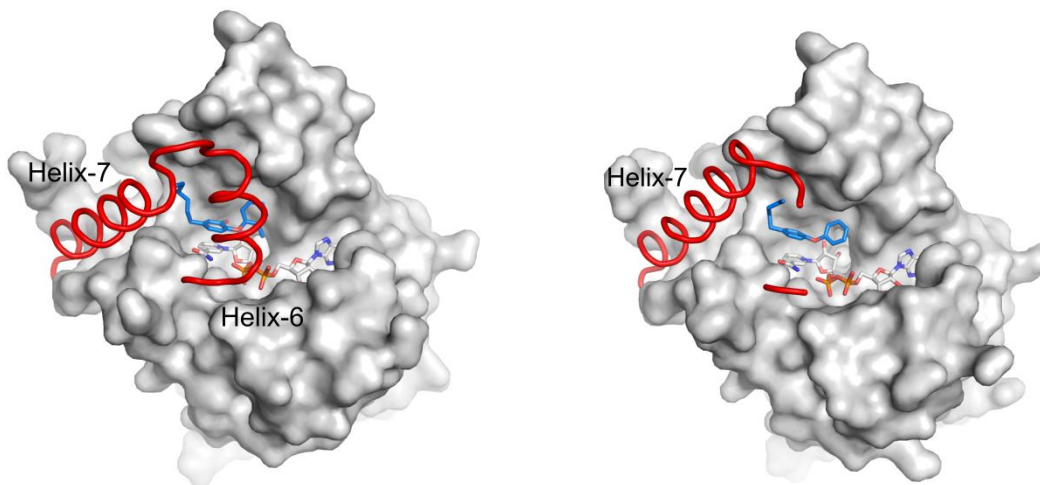
Figure 2.4 SBL ordering of InhA

Left: ordering of SBL helix-6 (residues 195-210) in InhA upon binding of slow-onset PT70 (PDB code 2X23). Right: helix-6 is disordered with a classical inhibitor, PT03 (PDB code 2B36).

The structural data from ftuFabI and saFabI have also correlated SBL ordering with slow-onset inhibition (**Table 2.3**). An ordered SBL is also observed in the structure of InhA in complex with INH-NAD [24,202] known to display slow-onset inhibition [45]. However, the InhA SBL is ordered in the crystal structure even in the absence of the inhibitor (**Table 2.3**). This ordered SBL conformation is different from that observed in the PT70 complex, with helix-6 much farther away from the active site and the cofactor (**Figure 2.5**). A similar large-amplitude conformational change of the SBL has been observed for ecFabI (**Figure 2.5**) and it is speculated that the inhibition potency of triclosan has to do with its ability to stabilize the flipping SBL [187,190]. The variation of SBL conformation for FabIs from different species has been correlated to the difference in binding affinity for triclosan [198].

Table 2.3 Correlation of slow-onset inhibition with SBL ordering in the crystal structures

PDB	Enzyme	Complex	Inhibitor	Slow-onset	N[1]	SBL	Ref
1ENY					1		203
2AQ8	InhA	binary			1	ordered	204
3OEW					1		205
1ZID	InhA	binary	INH–NAD	✔	1	ordered	24
2IDZ					1		202
2B35			triclosan		6		157
2B36	InhA	ternary	PT03		6	disordered	
2B37			PT05		6		
2X22	InhA	ternary	PT70	✔	2	ordered	159
2X23					4		
1DFI	ecFabI	binary			4	disordered	201
1QG6					4		144
1QSG	ecFabI	ternary	triclosan	✔	8	ordered	187
1C14					2		190
1D8A					2		191
2JJY	ftuFabI	apo, binary			4	disordered	154
3NRC	ftuFabI	ternary	triclosan	✔	2	ordered	194
3GNS	saFabI	apo			1	disordered	200
3GNT					2	disordered	
3GR6	saFabI	ternary	triclosan	✔	4	ordered	200

1. Number of subunits in the asymmetric unit
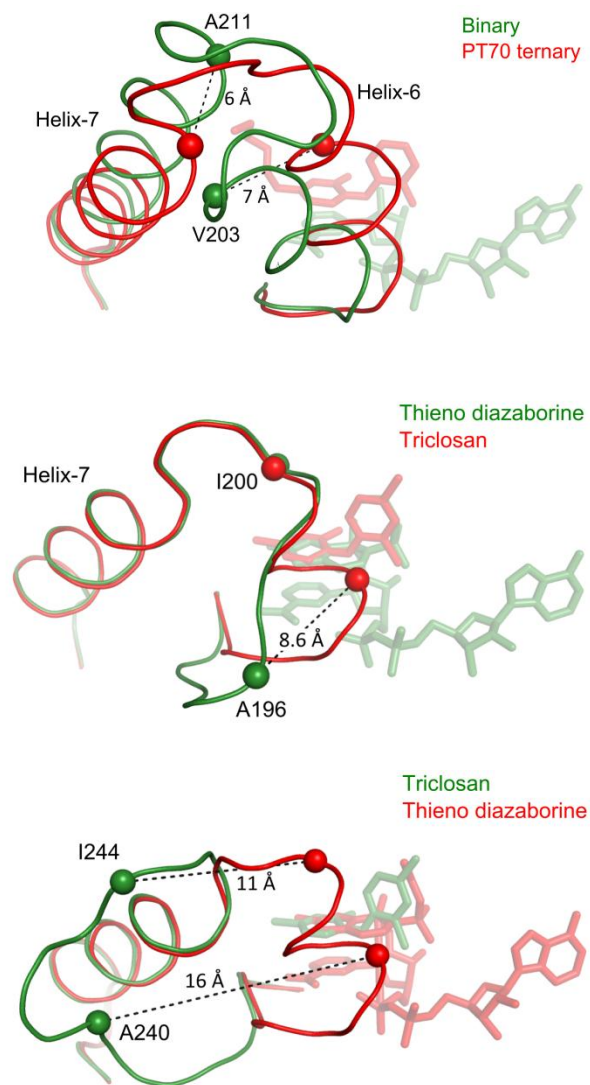
Figure 2.5 Large-amplitude conformational change of the SBL in InhA and FabI.
(a) Overlay of InhA binary structure in complex with NAD in green (PDB code 2AQ8) and ternary structure in complex with NAD$^+$ and PT70 in red (PDB 2X23). (b) ecFabI in complex with thieno diazaborine in green (PDB 1DFH) and with triclsoan in red (PDB 1QSG). (c) bnFabI in complex with triclosan in green (PDB 1D7O) and the A138G mutant in complex with thieno diazaborine in red (PDB 1CWU).

Timescale of protein motions

Slow-onset inhibition results in a slow-dissociating complex with a half-life comparable to the timescale of drug elimination from the blood after dosing, thus increasing the occupancy on the target for a prolonged period of time [180]. Rational design of a drug with this kinetic property requires an understanding of the structural mechanism of the slow step of binding. Presumably, the anticipated structural transition needs to overcome an energy barrier that fits to the timescale of slow-onset inhibition. However, there seems again no specific rule to predict which types of structural transitions must occur on this slow timescale (**Figure 2.6**).
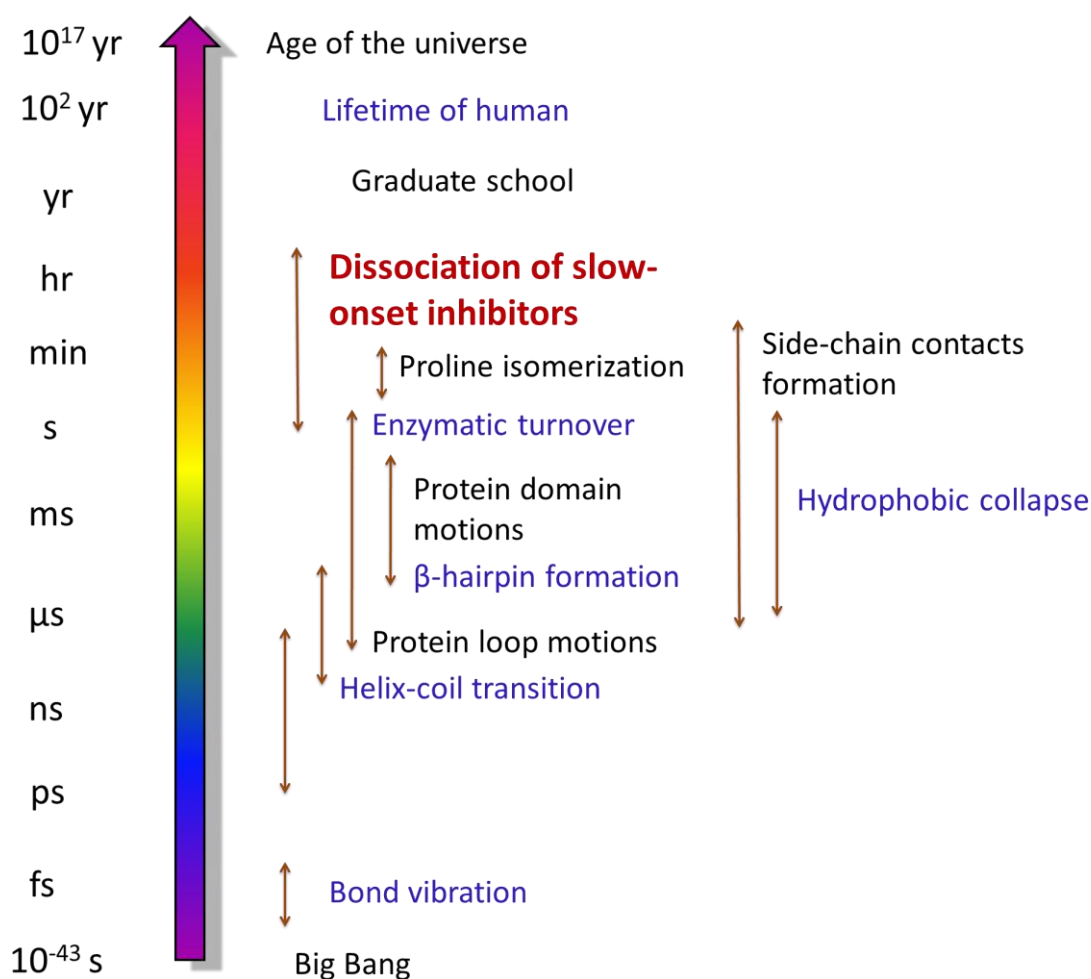
Figure 2.6 Timescale of protein motions

It is intriguing how a small difference on the B ring structure between PT03, PT05 and PT70 leads to the difference in SBL ordering in their complex crystal structures. The nature of the ordering process and why it occurs on a slow timescale is also a mystery. In this project we first speculated that the disorder of SBL in the crystal structure might be a consequence of fluctuations between different conformational states, and that it might be possible to characterize these conformations using ternary complexes of diphenyl ethers with different kinetic properties. We set out to determine the crystal structures of selected ternary complexes of wild type and mutant InhA. The hypothesis of SBL ordering by slow-onset inhibition will also be tested by comparing SBL density in the complex structures of slow-onset and classical inhibitors.

2.2 Methods

InhA preparation

InhA was overexpressed His-tagged at the N- terminus and purified by Ni affinity and gel filtration chromatography equilibrated with 30 mM PIPES, 150 mM NaCl and mM EDTA at pH 6.8. The protein was stored in the same buffer at −80 °C until thawed for crystallization setup.

ftuFabI preparation

ftuFabI was prepared by Yang Lu according to the procedure described in Ref [154]. C-terminal His tag was retained. Stored in 30 mM PIPES, 150 mM NaCl, 1.0 mM EDTA, pH 8.0 at −80 °C until thawed for crystallization setup.

Crystallization of the binary complex InhA:NAD

9.8 mg/mL of wild-type InhA was incubated with 2 mM of NAD^+ in 4% DMSO. The crystal was obtained by mixing 2 uL of the incubated solution with 2 uL of precipitant solution containing 100 mM N-(2-acetamido)iminodiacetic acid (ADA) pH 6.8, 150 mM ammonium acetate and 15.5 % PEG 4000, and setting up for hanging drop, vapor diffusion crystallization against 500 uL of precipitant solution at room temperature. Prior to freezing in liquid nitrogen, the crystal was soaked in the solution containing 100 mM ADA pH 6.8, 200 mM ammonium acetate, 18% PEG 4000, 4% DMSO, 150 mM NaCl and 2 mM NAD+ and transferred to the final solution containing 100 mM ADA pH 6.8, 150 mM ammonium acetate, 18% PEG

4000, 25% DMSO, 120 mM NaCl and 2 mM NAD+ in a stepwise fashion.

Crystallization of the ternary complex InhA:NAD:PT92

5 mg/mL of InhA was incubated with 2 mM NAD$^+$ and 800 μM PT92 in 8% DMSO for two hours at room temperature before mixing with equal volume of precipitant solution containing 100mM Bis-tris pH 6.4, 200 mM NaCl, 14% PEG 3350 and 4% DMSO and setting up for hanging drop, vapor diffusion crystallization. The crystals were left growing for ten days before soaking in the solution containing 100 mM Bis-tris pH 6.4, 310 mM NaCl, 16% PEG 3350, 25% DMSO, 2mM NAD$^+$ and 800 μM PT92 and freezing in liquid nitrogen. The diffraction data was collected at beamline X29 at NSLS.

Crystallization of the ternary complex ftuFabI:NAD:PT04

75C3: 0.5 μL protein solution + 0.5 μL reservoir solution. The protein solution contains 5.8 mg/mL ftuFabI, 2 mM NAD$^+$ and 2 mM PT04. The reservoir solution contains 100 mM Bis-Tris pH 5.8 and PEG MME 550. The cryo solution contains 100 mM Bis-Tris pH 5.8, 40% PEG MME 550, 75 mM NaCl, 1 mM NAD$^+$, and 2 mM PT04. Solutions for crystal 70F8 are the same except there is 50 mM CaCl$_2$ in the reservoir and cryo solutions.

Crystallization of the ternary complex InhA:NAD:PT155

Reservoir solution: 100mM HEPES pH 8.0 and 32% Jeffamine ED-2001 pH 7.0, hanging drop. Two crystal forms were fished from the same drop without further cryo protection and flash frozen in liquid nitrogen. Datasets 126D5_1 and 126D5_2 were collected at X25 and X12C, respectively.

Crystallization of other ternary complexes was performed similarly using conditions obtained from screening. The screening solutions were supplied by Hampton Research.

General procedure of structure determination

Crystals were cryo-protected and frozen in liquid nitrogen, and subsequently mounted in the center of synchrotron X-ray beam and a gaseous stream from liquid nitrogen. The diffraction images were collected while rotating the crystals over 180° along the axis perpendicular to the direction of the X-ray beam in most cases with varied size of oscillation angles, exposure time, detector distance and beam size for each dataset under the control of CBASS. The diffraction data were indexed, integrated and scaled using HKL2000 [206] and structures solved using MolRep [207] or Phaser incorporated in the CCP4 suite [208] and an InhA

model from the protein data bank. Ligands were built into the difference map calculated from the initial solutions. Structure parameters for ligands were generated using eLBOW [209] or libcheck in Coot and manually modified when necessary. Coordinates, ADP, occupancy and TLS refinement, simulated annealing, and water picking were performed using the CNS suite [210,211] and more recently Phenix [212]. Real-space refinement including the modeling of ligands, water, mutation, occupancies, rotamers, alternative conformers, was performed in Coot [213].

Refinement status

Crystal ID	Complex	R_{work}	R_{free}
21B3	InhA:NAD	0.173	0.242
48A3	A198S InhA:NAD:PT70	0.154	0.181
49A3	A198S InhA:NAD:PT10	0.214	0.268
70F8	ftuFabI:NAD:PT04	0.158	0.259
75C3	ftuFabI:NAD:PT04	0.142	0.191
88B3	InhA:NAD:PT115	0.149	0.183
90A3	InhA:NAD:PT119	0.155	0.187
124D3	InhA:NAD:PT130	0.231	0.278
126D5_1	InhA:NAD:PT155	0.149	0.170
126D5_2	InhA:NAD:PT155	0.153	0.190
236G8	V203A InhA:NAD	0.185	0.267
239G7_2	I215A InhA:NAD:PT70	0.206	0.263
249B3_5	V203A InhA:NAD:PT70	0.160	0.218
256B1_2	I215A InhA:NAD:PT70	0.216	0.258
261A1_2	V203A InhA:NAD:PT162	0.168	0.200
257D5_2	V203A InhA:NAD:PT162	0.204	0.267
280F11	V203A InhA:NAD:PT163	0.186	0.229
281G11	I215A InhA:NAD:PT163	0.181	0.246

Structural analyses

To quantify the difference in the SBL conformation in difference crystal structures, C_α atoms of residues 193 to 226 comprising helix-6 and helix-7 are superimposed by least-square fit and the RMSD values calculated by Coot. Parameters of the secondary structure are extracted by PDBsum. Values of the solvent accessible surface area are generated by PISA [214].

Two-dimensional plots of crystal structures

In the different InhA complex crystal structures, the conformations of helix-6 and helix-7 vary relative to the rest of the protein. To simplify structural comparison, the crystal structures are represented by only two coordinates that describe the movement of helix-6 and helix-7 relative to each other and to the rest of the protein. The first coordinate is defined as the dihedral angle in degrees measured from the C_α atoms of residues 233, 203, 211 and 231. Residue 233 and 231 are not on the SBL and are relatively stable in their positions, and therefore the dihedral angle describes the change in the relative positions of residues 203 and 211, and is sensitive to the distance from helix-6 and helix7 to the ligand binding site (**Figure 2.3.4**). The second coordinate is defined as the angle in degree measured from the C_α atoms of residue 203, 149, 215. Residue 149 is not on the SBL and is relatively stable, and the angle is sensitive to the rotation of helix-6 and helix-7 along their helical axes.

2.3 Results

The InhA:NAD binary complex structure

The reported binary complex structures of InhA bound to NAD(H) (PDB code 1ENY, 2AQ8, 3OEW) in the protein data bank are all in a hexagonal crystal form in which crystal packing occurs exclusively via the SBL region (**Table 2.4**). It is not surprising that an ordered SBL is observed in these structures. In a different crystallization condition, orthorhombic crystals can be obtained (**Table 2.4**) with occasional observation of the same hexagonal crystals. In the orthorhombic crystal structure there are six subunits in the asymmetric unit, and the SBL is not always involved in close crystal contacts. Disorder of SBL is observed when crystal contacts are absent, and the extent of disorder is apparently dependent on the crystal packing environment. Hence the number of residues in the SBL whose positions can be determined unambiguously varies. One of the subunits is packed closely against the SBL, and the entire SBL can be traced with certainty. Its conformation is found to be similar to that observed in the hexagonal crystal, both with an open helix-6. The C_α atoms of their residues 193 to 226 align with an RMSD value of 0.58 Å (**Table 2.5**).

Table 2.4 Unit cell parameters of wild type InhA crystal structures in complex with NAD(H)

ID	Space Group	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]
21B3	C222$_1$	81.87	99.80	375.73	90.00	90.00	90.00	6	2.3
1ENY[3]	P6$_2$22	100.14	100.14	140.45	90.0	90.0	120.0	1	2.2
2AQ8[3]	P6$_2$22	96.02	96.02	138.16	90.0	90.0	120.0	1	1.9
3OEW[3]	P6$_2$22	98.55	98.55	139.71	90.0	90.0	120.0	1	2.2

1. N: Number of subunits in the asymmetric unit
2. R: resolution
3. From the PDB

Table 2.5 RMSD values of the binary complex and PT70 ternary complex

RMSD of C_α atoms on helix-6 and helix-7, or residue 193 to 226, by least-square alignment of the same region in subunit 3 of the orthorhombic binary complex structure (Crystal ID 21B3) with other InhA structures

RMSD(Å)	1ENY	2AQ8	3OEW	2X23-1	2X23-2	2X23-3	2X23-4	2X22-1	2X23-2
21B3-3	0.52	0.58	0.52	2.71	2.93	2.77	2.97	3.10	3.04

InhA ternary complexes with NAD$^+$ and PT70, PT91, PT92, and PT10
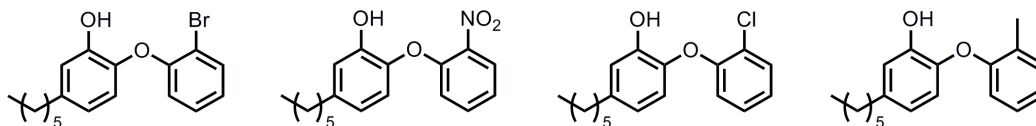
Figure 2.7 Structures of PT92, PT10, PT91 and PT70 (left to right)

Crystal form and mobility of SBL

PT91, PT92, and PT10 are analogues of PT70 (**Figure 2.7**) that also display slow-onset inhibition kinetics with InhA (unpublished data from Pan Pan). As shown in Error! Reference source not found., all of these inhibitors including PT70 can be co-crystallized with InhA and NAD$^+$ in the same crystal form belonging to space group $P2_12_12_1$, which is different from the two crystal forms that have been reported for the PT70 complex in the PDB. There are four subunits in the asymmetric unit, two of which have their SBL extensively involved in crystal packing, and the density for the SBL can be clearly observed. The other two subunits do not engage much of their SBL in crystal packing, and display noticeably higher B factors or disorder from residue 205 to 220. As demonstrated by subunit 4 in the PT92 ternary complex structure, there is only little discontinuous density for residues 207 to 220 (**Figure 2.8**). This is similar to the observation of B factors in the PT70 complex structure in the PDB (2X23) in which two of the four subunits less involved in crystal packing through the SBL display higher B factor in a similar region (subunit 2 and 4 in **Figure 2.9**). The "hot spot" centers on helix-7. The other PT70 complex structure from the PDB, 2X22, exhibits the same trend.

Table 2.6 Unit cell parameters of wild type InhA crystal structures in complex with NAD$^+$ and diphenyl ether inhibitors

ID	Inhibitor	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]
33B2	PT92	P2$_1$2$_1$2$_1$	72.85	90.49	161.83	90.00	90.00	90.00	4	1.6
28D3_3	PT10	P2$_1$2$_1$2$_1$	88.83	91.08	148.91	90.00	90.00	90.00	4	2.4
29C5_2	PT91	P2$_1$2$_1$2$_1$	74.84	90.65	164.44	90.00	90.00	90.00	4	2.3
24C6	PT70	P2$_1$2$_1$2$_1$	74.34	91.15	164.06	90.00	90.00	90.00	4	2.8
2X22[3]	PT70	C222$_1$	89.81	157.51	91.23	90.00	90.00	90.00	2	2.1
2X23[3]	PT70	P2$_1$	88.48	90.27	89.56	90.00	118.76	90.00	4	1.8

1. N: Number of subunits in the asymmetric unit
2. R: resolution
3. From the PDB
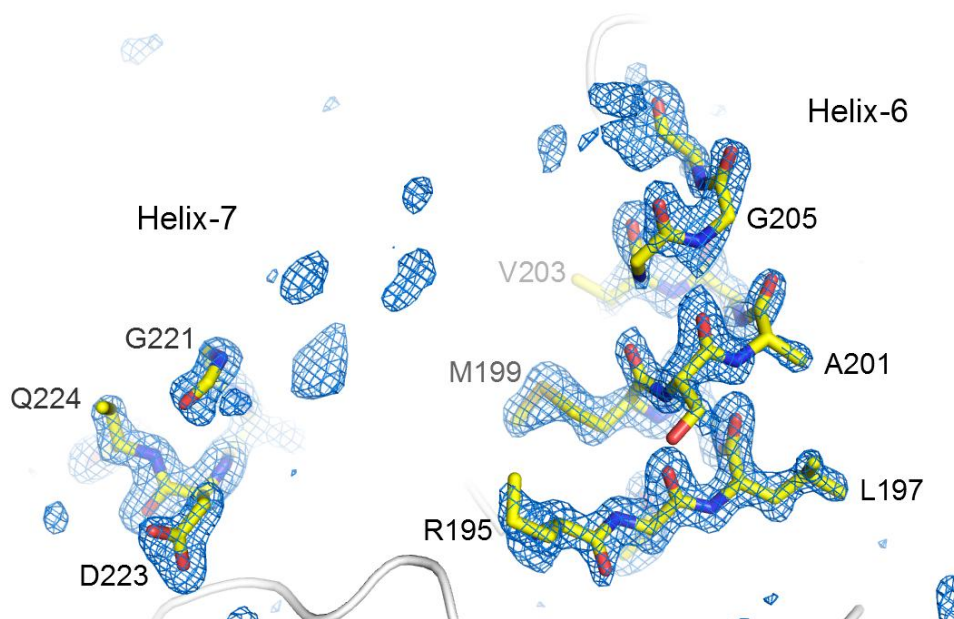
Figure 2.8 SBL density in the PT92 ternary complex

The simulated annealing Fo−Fc map contoured at 3 σ is generated by omitting residues 195 to 225 in subunit 4 from the model
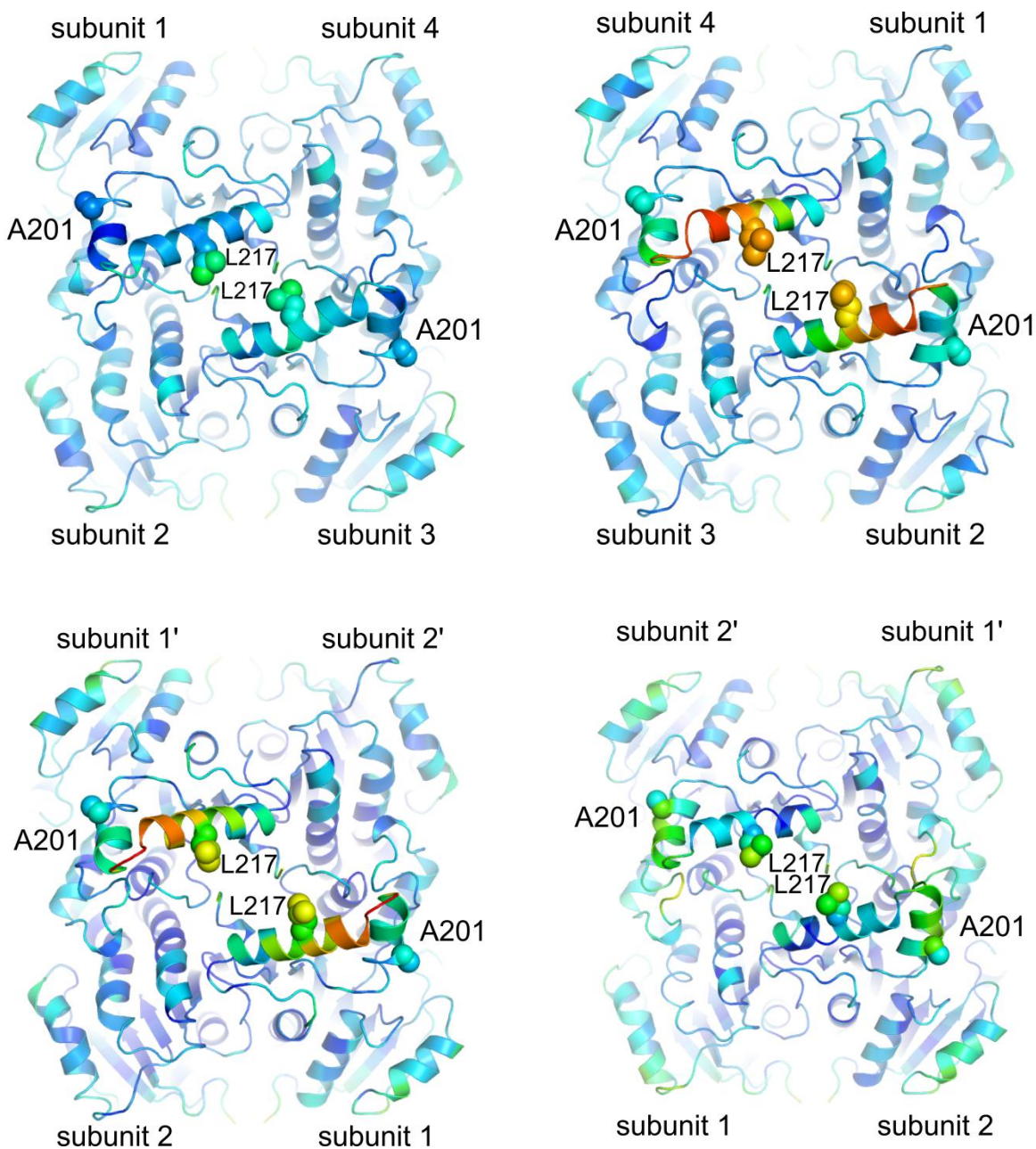
Figure 2.9 Region of high mobility in the PT70 ternary complex

Crystal structures of InhA in complex with PT70 (PDB 2X23, top; 2X22, bottom) colored by B factor with a spectrum from red to blue showing highest to lowest B factors. The two opposite views each shows two SBLs from two of the four subunits in the asymmetric unit. The side chains of A201 and L217 on the SBL are shown by spheres.

The overall consistency in conformation within the three PT70 complex structures in **Table 2.6** is high, but helix-7 appears to fluctuate significantly compared with the rest of the protein (**Figure 2.10**). The RMSD values obtained from aligning residue 193 to 226 comprising helix-6 and helix-7 range from 0.25 to 1.23 Å (**Table 2.7**). For comparison, the RMSD values obtained by aligning helix-5, corresponding to residue 158 to 183 on the opposite side of the ligand binding pocket, is much lower, ranging from 0.062 to 0.23 Å (**Table 2.10**). If the alignment is further decomposed into the individual helices, it can be seen that the RMSD values for helix-7 (**Table 2.9**) are clearly higher than those for helix-6 (**Table 2.8**) except three alignments involving subunit 2 of PDB structure 2X22, as revealed by the cyan subunit in **Figure 2.10**. The greater conformational variability of helix-7 is consistent with its higher B factors and higher degree of disorder when these inhibitors are bound.
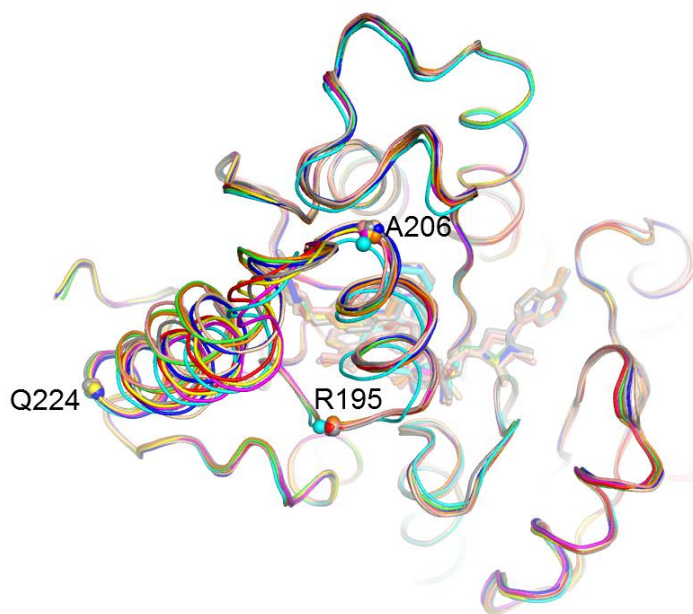
Figure 2.10 Variability of helix-7 in the PT70 complex
Overlay of all the subunits in the PT70 ternary complex structure from three different crystals in **Table 2.6**.

Table 2.7 RMSD values of helix-6 and helix-7 in the PT70 complex

RMSD of C_α atoms on helix-6 and helix-7, or residue 193 to 226, by least-square alignment of the same region in all the subunits from the three PT70 complex structures

RMSD(Å)	2X23-1	2X23-2	2X23-3	2X23-4	2X22-1	2X22-2	24C6-1	24C6-2
2X23-1	0							
2X23-2	0.88	0						
2X23-3	0.62	0.76	0					
2X23-4	0.91	0.25	0.74	0				
2X22-1	1.11	0.33	0.92	0.40	0			
2X22-2	1.23	0.62	0.98	0.64	0.58	0		
24C6-1	0.26	0.83	0.60	0.86	1.05	1.15	0	
24C6-2	0.81	0.87	0.35	0.84	1.00	1.00	0.75	0

Table 2.8 RMSD values of helix-6 in the PT70 complex

RMSD of C_α atoms on helix-6, or residue 193 to 206, by least-square alignment of the same region in all the subunits from the three PT70 complex structures

RMSD(Å)	2X23-1	2X23-2	2X23-3	2X23-4	2X22-1	2X22-2	24C6-1	24C6-2
2X23-1	0							
2X23-2	0.215	0						
2X23-3	0.285	0.296	0					
2X23-4	0.170	0.117	0.276	0				
2X22-1	0.207	0.099	0.282	0.135	0			
2X22-2	0.671	0.523	0.652	0.593	0.529	0		
24C6-1	0.154	0.196	0.258	0.172	0.213	0.625	0	
24C6-2	0.317	0.316	0.176	0.296	0.331	0.594	0.288	0

Table 2.9 RMSD values of helix-7 in the PT70 complex

RMSD of C_α atoms on helix-7, or residue 206 to 226, by least-square alignment of the same region in all the subunits from the three PT70 complex structures

RMSD(Å)	2X23-1	2X23-2	2X23-3	2X23-4	2X22-1	2X22-2	24C6-1	24C6-2
2X23-1	0							
2X23-2	0.835	0						
2X23-3	0.521	0.831	0					
2X23-4	0.812	0.272	0.794	0				
2X22-1	1.048	0.332	1.010	0.471	0			
2X22-2	0.945	0.364	0.915	0.383	0.374	0		
24C6-1	0.262	0.777	0.555	0.765	0.973	0.880	0	
24C6-2	0.717	1.000	0.382	0.953	1.150	1.026	0.708	0

Table 2.10 RMSD values of helix-5 in the PT70 complex

RMSD of C_α atoms on helix-5, or residue 158 to 183, by least-square alignment of the same region in all the subunits from the three PT70 complex structures

RMSD(Å)	2X23-1	2X23-2	2X23-3	2X23-4	2X22-1	2X22-2	24C6-1	24C6-2
2X23-1	0							
2X23-2	0.084	0						
2X23-3	0.066	0.090	0					
2X23-4	0.077	0.062	0.086	0				
2X22-1	0.105	0.091	0.121	0.094	0			
2X22-2	0.093	0.106	0.093	0.109	0.097	0		
24C6-1	0.199	0.184	0.187	0.196	0.196	0.185	0	
24C6-2	0.183	0.168	0.184	0.170	0.165	0.198	0.234	0

Despite the greater conformational variability of helix-7 in the PT70 ternary complex structures, the RMSD values are small ($< 1.3\,\text{Å}$) compared to those measured by aligning the ternary and binary complexes ($> 2.7\,\text{Å}$,**Table 2.5**). As shown in **Figure 2.5**, there is large-amplitude conformational difference on both helix-6 and helix-7 between the ternary and binary structures. Many residues (195 to 219) on these two helices interact with one another and since helix-6 and helix-7 do not move in the same direction, many of these residues (202 to 215) dramatically change their positions relative to one another. In contrast, the relative positions of residues in different subunits within the PT70 ternary complex remain similar. To demonstrate this difference in the level of conformational change, the structures are plotted in the two dimensions defined by the two coordinates shown in **Figure 2.11**. In **Figure 2.12**, it is clear that the difference within the ternary complex or within the binary complex is small compared to the difference between the ternary and binary structures.
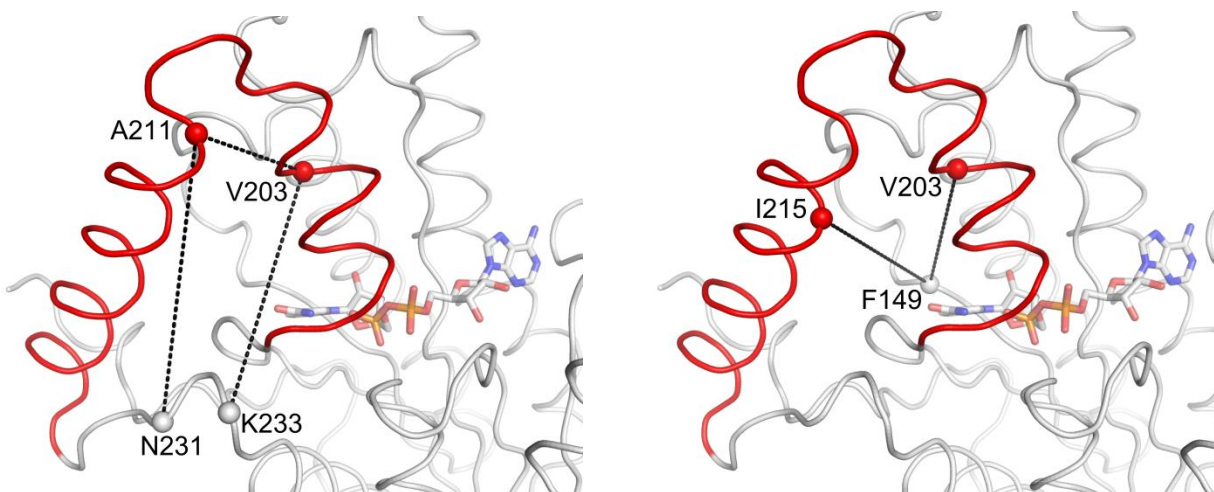
Figure 2.11 The two coordinates used to make the 2D plot of the crystal structures
$\delta 1$ is the dihedral angle defined by the C_{α} atoms of residue 233, 203, 211 and 231. $\delta 2$ is the angle formed by the C_{α} atoms of residue 203, 149 and 215.
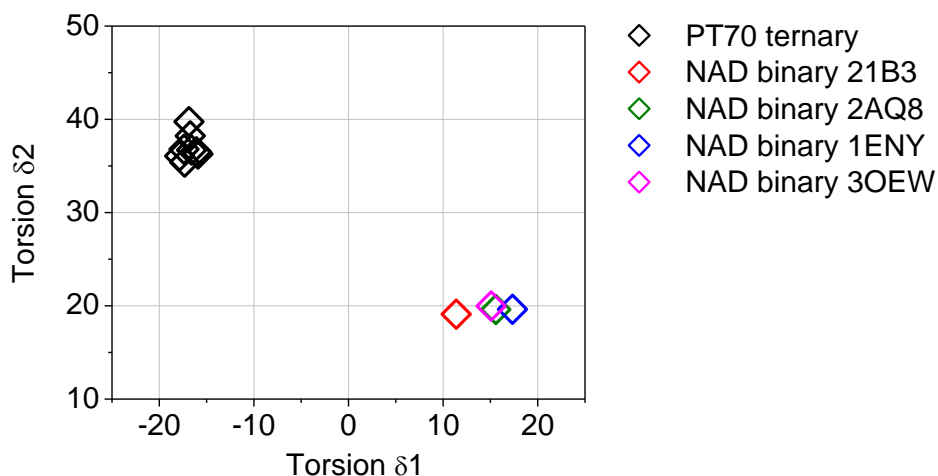
Figure 2.12 Crystal structures of the binary complex and PT70 ternary complex represented on the 2D plot

There are a total of 10 InhA monomers from the three PT70 ternary complex structures whereas the binary complex strucutres contain only one InhA monomer in the asymmetric unit. $\delta 1$ and $\delta 2$ are measured in degrees.

Variation of SBL conformation bound with PT70 analogues

The ordered conformations of the SBL bound with PT70, PT91, PT92 and PT10 are compared in **Table 2.11**. Among the three PT70 analogue complexes, PT91 gives the most similar SBL conformation with that of PT70. For PT92 and PT10, the alterations in substituent on the B ring (**Figure 2.13**) have more noticeable impacts on the SBL conformations. The SBL in subunit 2 of the PT92 complex clearly adopts a conformation distinct from any observed with PT70 (**Figure 2.14**, top) while that in subunit 1 remains similar. Helix-6 in subunit 1 of the PT10 complex is also clearly displaced when aligned with the PT70 complex (**Figure 2.14**, bottom). These observations are reflected in the RMSD values. While the SBL with PT91 generally aligns with PT70 within 1 Å of RMSD, subunit 2 of the PT92 complex and subunit 1 of the PT10 complex give RMSD values greater than 2.8 Å and 1 Å, respectively (**Table 2.11**). The two-coordinate structural plot also demonstrates this difference (**Figure 2.15**).
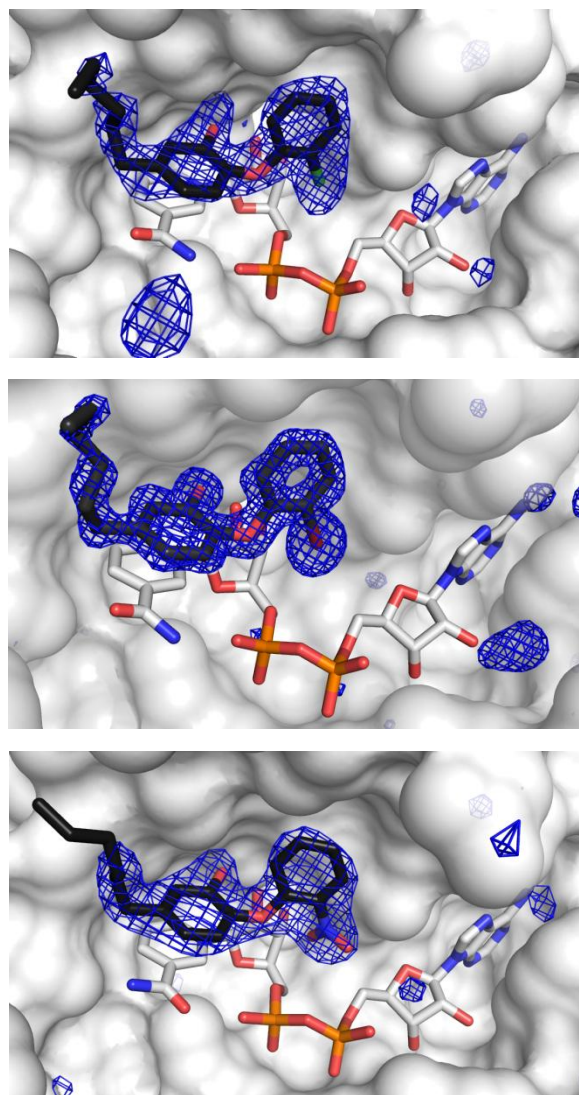
Figure 2.13 Maps of PT91, PT92 and PT10

Simulated annealing Fo-Fc maps calculated by omitting PT91 (top), PT92 (middle), and PT10 (bottom) from their respective structures. Shown are subunit 2 for PT91, subunit 1 for PT92 and subunit 1 for PT10.

Table 2.11 RMSD values of the SBL in the ternary complexes of PT70, PT91, PT92 and PT10

RMSD of C_α atoms on helix-6 and helix-7, or residue 193 to 226, by least-square alignment of the same region

RMSD(Å)	PT91-1	PT91-2	PT92-1	PT92-2	PT10-1	PT10-2
2X23-1	0.69	0.15	1.13	2.82	1.38	1.15
2X23-2	0.76	0.83	0.48	3.07	1.09	0.58
2X23-3	0.18	0.60	1.01	2.94	1.17	0.96
2X23-4	0.74	0.87	0.50	3.11	1.02	0.61
2X22-1	0.91	1.07	0.29	3.27	1.09	0.53
2X22-2	0.92	1.18	0.71	3.28	1.20	0.40
PT91-1	0					
PT91-2	0.66	0				
PT92-1	1.00	1.09	0			
PT92-2	2.97	2.82	3.27	0		
PT10-1	1.17	1.34	1.10	3.04	0	
PT10-2	0.92	1.10	0.60	3.11	1.18	0
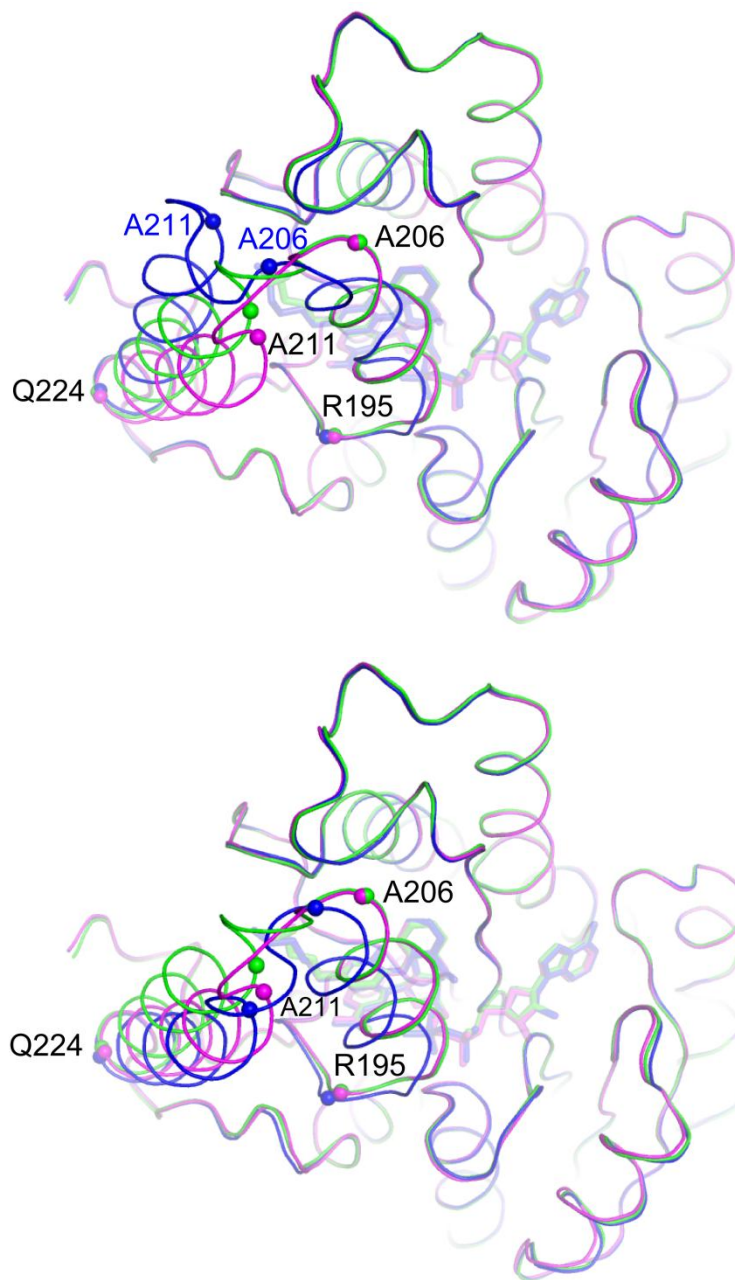
Figure 2.14 Variation of SBL conformation in the ternary complexes of PT92 and PT10
Overlay of subunit 2 from PT92 complex structure (top, blue) and of subunit 1 from PT10 complex structure (bottom, blue) with PT70 complex structures 2X23-1(green) and 2X22-1 (magenta)
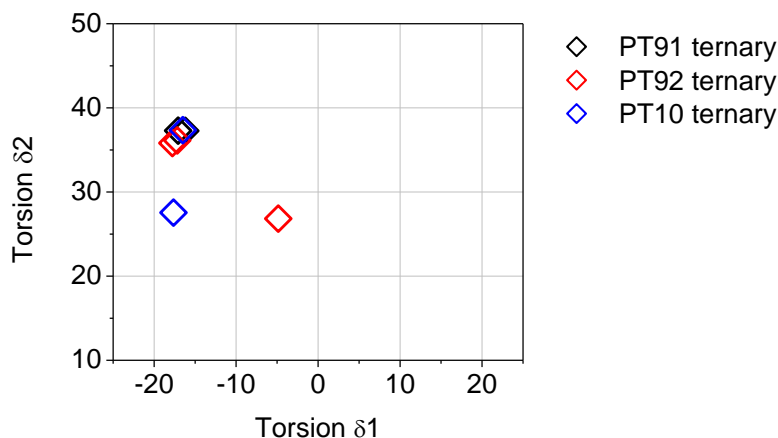
Figure 2.15 Structures of PT91, PT92 and PT10 in the 2D plot

$\delta 1$ and $\delta 2$ are measured in degrees.

Impact of B ring structure on SBL conformation

As shown in **Figure 2.16**, the B ring of PT70 establishes extensive interactions with helix-6, strand-4, helix-5, and NAD$^+$. These interactions are critical for PT70's improved affinity for InhA [159]. In subunit 2 of the PT92 complex structure where large deviations from the PT70 structure are observed (**Figure 2.14**, top), the helix-6 contacts shift away from the B ring of the inhibitor by 0.5 Å, which also removes the interactions between helix-6 and strand-4, exposing the B ring to bulk solvent. This appears to result from the bulkier bromo group that interacts unfavorably with both helix-6 and strand-4 (**Figure 2.16**, middle). Similarly, in subunit 1 of the PT10 complex structure, the bulkier nitro group on the B ring of PT10 compared with PT70 pushes helix-6 away from strand-4 and exposes the B ring to the solvent (**Figure 2.16**, bottom). However, the conformation of helix-6 is not the same as that observed for subunit 2 in the PT92 complex. The B ring interacts with helix-6 through A201 in the case of PT92, but through I202 in the case of PT10. This difference between PT92 and PT10 appears to also lead to the significant change in the unit cell dimensions despite the fact that both complex structures are in the same space group (**Table 2.6**).

41
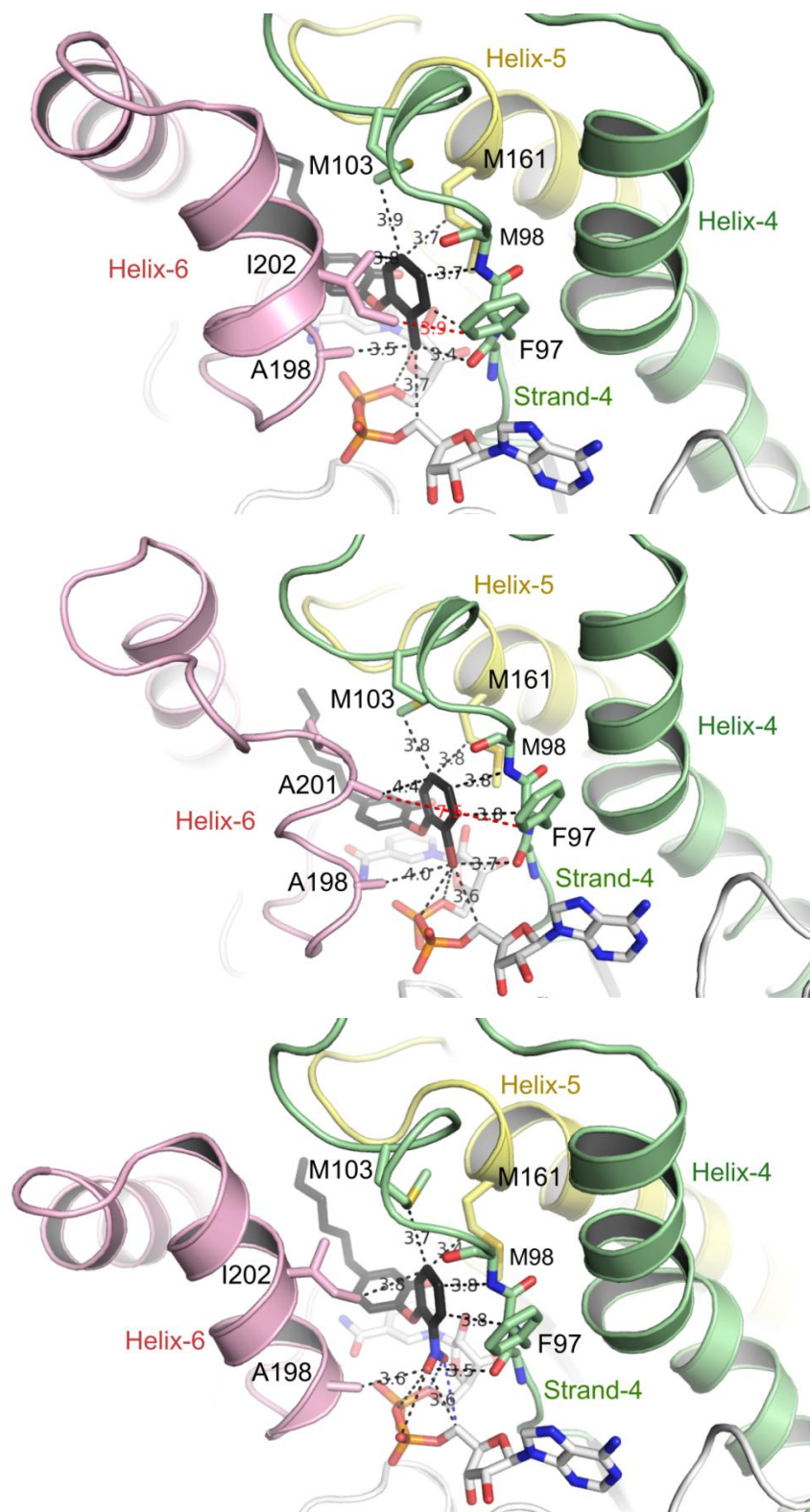
Figure 2.16 Interactions of the inhibitor B rings with InhA.
Top: subunit 1 in the PT70 complex structure 2X23; middle: subunit 2 in the PT92 complex structure; bottom: subunit 1 from the PT10 complex structure.

The InhA A198S mutant ternary complexes with NAD$^+$ and PT70 or PT10

Given the sensitivity of the SBL conformation to the change in the B ring ortho substitution on the inhibitor, we further explored the impact of mutating A198 which is directly involved in the interaction between helix-6 and B ring ortho substituent. The A198S mutant is active, and the k_{cat} and k_{cat}/K_m values decrease by 17 and 4.2 fold, respectively (unpublished data from Dr. Nina Liu). The kinetic progress curves show that PT70 loses the slow-onset characteristics with the mutant while PT10 remains slow-onset. The inhibition constant for the mutant by PT70 is 2860-fold larger than that for the wild type (Dr. Nina Liu). Crystallization of the mutant in similar conditions as those used for the wild type tends to yield clusters that cannot be used in the diffraction experiment. Hence a new condition was explored, and high concentrations of acetate or formate below pH 6 were found to yield single crystals that gave good diffraction (**Table 2.12**). Above pH 6, the crystals become clusters of needles.

Table 2.12 Unit cell parameters of the InhA A198S mutant crystal structures in complex with NAD$^+$ and PT70 or PT10

ID	Inhibitor	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]
48A3	PT70	I4$_1$	90.48	90.48	181.64	90.00	90.00	90.00	2	1.9
49A3	PT10	I4$_1$	90.96	90.96	182.71	90.00	90.00	90.00	2	1.9

1. N: Number of subunits in the asymmetric unit

2. R: resolution

The structures show that S198 interacts with the inhibitor B ring ortho substituent and hydrogen-bonds to the NAD$^+$ phosphate. However the slightly bulkier side chain on residue 198 might create some hindrance with the B ring substituent since the distance is smaller than 3.3 Å. Residues 203 to 208 appear less ordered as indicated by poorer quality of density and higher B factors. The SBL conformation for the A198S mutant bound to PT70 appears to be intermediate between the wild type PT70 ternary structure and the binary structure, and is most similar to the conformation observed for subunit 2 of the PT92 complex.

InhA ternary complexes with NAD$^+$ and PT115 or PT119

The residence time determined from progress curve analysis is 263 min and 80 min for PT115 and PT119, respectively, significantly longer than that of PT70. The K_i for PT119 is 2.1 nM (unpublished data from Pan Pan).
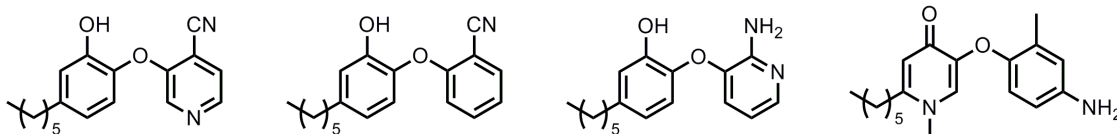
Figure 2.17 Structures of PT115, PT119, PT130 and PT155

Crystallization

The same high salt conditions used for crystallizing the A198S mutant also yielded single crystals with similar diffraction quality (**Table 2.13**). This set of conditions cannot be used to crystallize the ternary complexes with PT04 or PT07 (PT04 with 14-carbon acyl chain), or the NAD$^+$ binary complex. Crystals of similar appearance can be obtained for the ternary complex with PT70, but the crystals do not diffract. These observations suggest that binding with these different inhibitors changes the structure of InhA in some fundamental way.

Table 2.13 Unit cell parameters of the InhA crystal structures in complex with NAD$^+$ and PT115 or PT119

ID	Inhibitor	Space Group	a(Å)	b(Å)	c(Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]
88B3	PT115	I4$_1$22	90.47	90.47	182.42	90.00	90.00	90.00	1	1.9
90A3	PT119	I4$_1$22	90.62	90.62	182.63	90.00	90.00	90.00	1	1.8

Crystal form and SBL conformation

The two structures in **Table 2.13** are in space group I4$_1$22 containing one subunit in the asymmetric unit (ASU), and can be considered to be in the same crystal form as the A198S mutant structure, which is in space group I4$_1$ and has two subunits in ASU. The difference lies in whether the two subunits in the I4$_1$ structure are treated exactly the same, which will reduce the structure to the higher-symmetry space group, I4$_1$22. As observed in the A198S mutant structures, residues 203 to 208 appear less ordered as indicated by poorer quality of

density and higher B factors. This common feature in the $I4_1(22)$ crystal form is related to the pattern of crystal packing which leaves more space around residues 205 to 208 compared to other parts of the SBL. The PT115 complex has a SBL conformation almost identical to that observed in the PT70 complex with the A198S mutant. A RMSD of 0.12 Å is obtained by matching the C_α atoms from residue 193 to 226. Residual density reveals that some buffer component is ordered around the SBL region. Acetate molecules are likely bound between helix-6 and strand-4, and between helix-6 and helix-7.

The InhA ternary complex with NAD^+ and PT130

The structure of PT130 is shown in **Figure 2.17**. PT130 also displays longer residence time (189 min) on InhA (unpublished data from Pan Pan).

The $I4_1(22)$ crystal form and SBL conformation

The similar SBL conformational observed in the $I4_1(22)$ crystal form raises the concern that the SBL conformation is determined by crystal packing rather than the inhibitor structure. The presence of the acetate might also influence the SBL conformation. The PT130 complex structure is found to be in the same crystal form despite the absence of acetate (**Table 2.14**). In addition, the same crystal form is obtained for the PT115 complex in the high formate conditions. Two structures from the PDB in this crystal form (3FNG, 3FNH) were crystallized in the same HEPES/MPD buffer condition from which the $P6_222$ hexagonal binary complex structure was obtained. Together these observations suggest the inhibitors still play a role in determining the SBL conformation and the crystal form.

The PT130 ternary complex structure shows similar regions of higher B factors and degrees of disorder compared to other structures in the same crystal form. Nevertheless, the SBL conformation is different from the others and resembles that observed for the PT70 complex (**Figure 2.18**). This apparently reflects the fact that the B ring substituent is smaller than that of PT115.

Table 2.14 Unit cell parameters of the InhA crystal structure in complex with NAD$^+$ and PT130

ID	Inhibitor	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]
124D3	PT130	I4$_1$	90.67	90.67	185.41	90.00	90.00	90.00	2	2.0
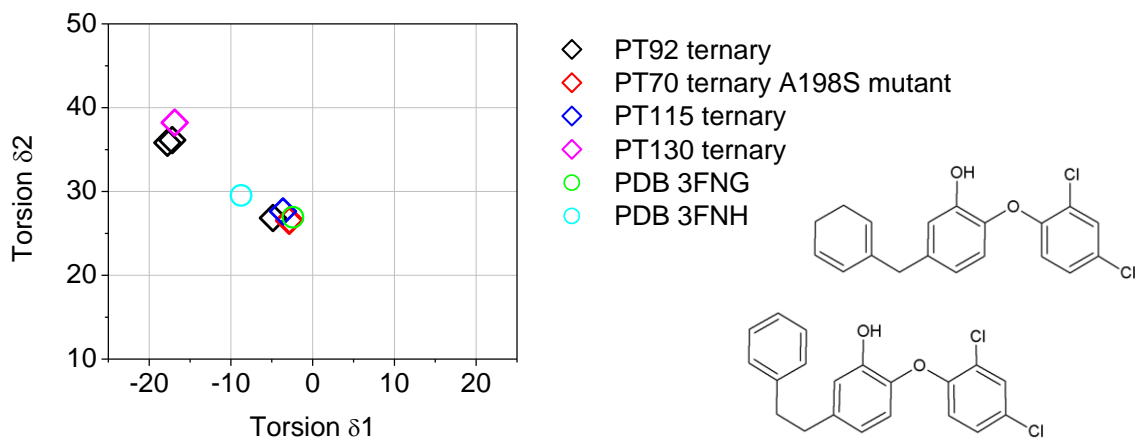
Figure 2.18 Structures in the I4$_1$(22) crystal form compared with the PT92 complex.
Structures of inhibitors in the PDB structures 3FNG (top) and 3FNH (bottom) are shown on the right. δ1 and δ2 are measured in degrees.

The ftuFabI ternary complex with NAD^+ and PT04

PT04 is a slow-onset inhibitor of ftuFabI [154]. The K_i is 2.7 nM, and the residence time is 143 min. The binary complex structure of ftuFabI bound to NAD^+ (PDB code 2JJY) has a disordered SBL. We attempt to co-crystallize PT04 with ftuFabI and NAD^+ to see what effect PT04 has on the SBL and if the observation is common to the InhA system.

Crystallization

Two crystallization conditions exist for PDB structures 2JJY and 3NRC, which are the NAD^+ binary complex and the triclosan ternary complex structure, respectively. The binary crystals grow from 200 mM magnesium acetate and 20% PEG 3350 in two days. The triclosan ternary complex crystals grow from 2M $(NH_4)_2SO_4$ and 100 mM phosphate-citrate at pH 4.2 in four days. Crystallization was tested using the magnesium acetate condition in the absence or presence of PT04 but no crystal was reproduced in the hanging drops.

The crystals of ftuFabI binary complex bound to NAD^+ grow from sitting drops containing 100 mM Bis-Tris pH 6.5, 50 mM $CaCl_2$ and 30% PEG MME 550 and form sticks and clusters. Needles grow from sitting drops containing 100 mM HEPES pH 7.5, 25% PEG 3350, and 200 mM of either $(NH_4)_2SO_4$, Li_2SO_4, or $MgCl_2$. The crystals from PEG MME 550 cannot be reproduced in hanging drops. The crystals from $(NH_4)_2SO_4$ cannot be reproduced in hanging drops or in the presence of PT04.

PT04 ternary complex crystals also grow from the sitting drop containing 100 mM Bis-Tris pH 6.5, 50 mM $CaCl_2$ and 30% PEG MME 550, but again cannot be reproduced in either sitting drops in a different container, or in hanging drops. The crystals only reproduce in sitting drops in the same container and diffract to 2.7 Å (**Table 2.15**, 70F8). The crystals reproduce in 20 min after setup at Bis-Tris pH 6.5 and 7.0, and in 4 days in the absence of PT04. The crystals of PT04 ternary complex grow at Bis-Tris pH 5.8 in the absence of $CaCl_2$, and diffract to 2.0 Å (**Table 2.15**, 75C3).

Table 2.15 Unit cell parameters of the ftuFabI crystal structures in complex with NAD^+ and PT04

ID	Inhibitor	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N^1	$R (Å)^2$
70F8	PT04	$P2_1$	105.99	71.36	126.77	90.00	100.55	90.00	8	2.7
75C3	PT04	$P2_1$	73.97	70.92	89.35	90.00	100.44	90.00	4	2.0

SBL ordering

In the PT04 ternary complex structure of ftuFabI (dataset 75C3), density for the SBL in subunit 1 and 4 is clear. Helix-6 of subunit 1 packs against the C-terminal His-tag from a symmetry-related subunit. SBL in subunit 4 is ordered without close crystal contacts. SBL in subunit 2 is not involved in crystal contacts, and has poor density between helix-6 and helix-7 (residue 199 to 202). Nor is the SBL in subunit 3 involved in crystal contacts, and there is no clear density for part of helix-6 and helix-7 (residue 198 to 206). Density for the C-terminal His-tag can be observed.

The C-terminus and tetrameric interactions

The His-tag folds along the surface of the ftuFabI tetramer, in contrast to the position found in the ftuFabI:NAD$^+$ binary complex structure (PDB code 2JJY) which folds perpendicular to the surface of the tetramer and enters the binding site of both substrates of a symmetry-related subunit. In either case, the C-terminus has a chance to interact with helix-7 of a different subunit within the tetramer, noticeably through Y209. In InhA there is no tag at the C-terminus, but the tail also has a chance to interact with helix-7 of a different subunit within the tetramer. The binding of inhibitor sometimes removes these interactions and is expected to have some impact on the tetrameric interactions.

SBL conformation

Helix-6 is slightly more open when PT04 is bound in ftuFabI compared to the triclosan structure, 3NRC. In subunit 3 of the PT04 ternary complex structure of ftuFabI, residual density is found in an even more open position. This suggests that PT04 has some propensity to bind in a different mode with a more open SBL conformation. In the binary complex structure 2JJY, the His-tag found in the active site implies that the SBL adopts an open conformation, which could be characterized by further experiments or modeling. In the ternary complex of either PT04 or triclosan, helix-6 which interacts with and covers over the inhibitor occupies the space expected for the fatty acyl chain as in the case of InhA (discussed later).

The InhA ternary complex with NAD^+ and PT155

The structure of PT155 is shown in **Figure 2.17**. PT155 is a diphenyl ether analogue with the A ring replaced by a 4−pyridone. However, due to resonance the A ring is expected to display partial aromatic characteristics. The NMR spectrum supports a hybrid structure (unpublished data from Pan Pan). The inhibition kinetics of InhA by PT155 does not exhibit the slow−onset characteristics.

Crystal form

The crystals for the PT155 complex are obtained from the same HEPES/Jeffamine condition that crystallizes the PT130 complex, but different crystal forms (**Table 2.16**) and SBL conformations are observed. The inhibitor therefore seems to dictate the crystal form and the SBL conformation.

Table 2.16 Unit cell parameters of the InhA crystal structures in complex with NAD^+ and PT155

ID	Inhibitor	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N^1	R (Å)^2
126D5_1	PT155	$P2_12_12_1$	89.27	97.44	182.58	90.00	90.00	90.00	4	1.8
126D5_2	PT155	$I2_12_12_1$	88.95	97.15	187.01	90.00	90.00	90.00	2	2.3

SBL density and conformation

There are four subunits in the ASU in the $P2_12_12_1$ structure. Density for the SBL in subunit 1, 3 and 4 is clear. Residues 204 to 210 in subunit 2 display disorder or higher B factors. There are two subunits in the ASU in the $I2_12_12_1$ structure. Residues 199 to 210 in subunit 1 display disorder or higher B factor while the density for the SBL in subunit 2 is clear.

Interestingly, a wide range of the SBL conformations are observed for the PT155 complex, spanning the conformational space between the binary complex and the PT70 ternary complex (**Figure 2.19**). These different SBL conformations provide a variety of binding modes for a single inhibitor, and map out a possible binding path for its analogue, the diphenyl ether. The two subunits of the PT155 complex that reside at the two ends of this path closely resemble the binary complex and the PT70 ternary complex, respectively, whereas the

subunits with intermediate structures closely resemble the ternary complexes of other diphenyl ethers, which is also revealed by the RMSD values (**Table 2.17**).

Table 2.17 RMSD values in the SBL of the PT155 ternary complex
RMSD of C_α atoms on helix-6 and helix-7, or residue 193 to 226, by least-square alignment of the same region

RMSD (Å)	PT155I-1	PT155P-2	PT155P-1	PT155P-3	PT155P-4	PT155I-2
Binary[1]	2.36	3.17	2.90	1.35	1.41	0.32
PT92 ternary[2]	1.80	2.76	1.32	2.73	2.77	2.90
PT92 ternary[3]	1.07	1.85	3.30	3.18	3.24	3.07
PT10 ternary[4]	1.09	1.31	3.08	3.20	3.23	2.93
PT70 ternary[5]	1.14	1.80	3.14	3.02	3.06	2.91

1: PDB structure 2AQ8; 2: PT92 complex subunit 2; 3: PT92 complex subunit 1; 4: PT10 complex subunit 1; 5: PDB structure 2X23 subunit 4
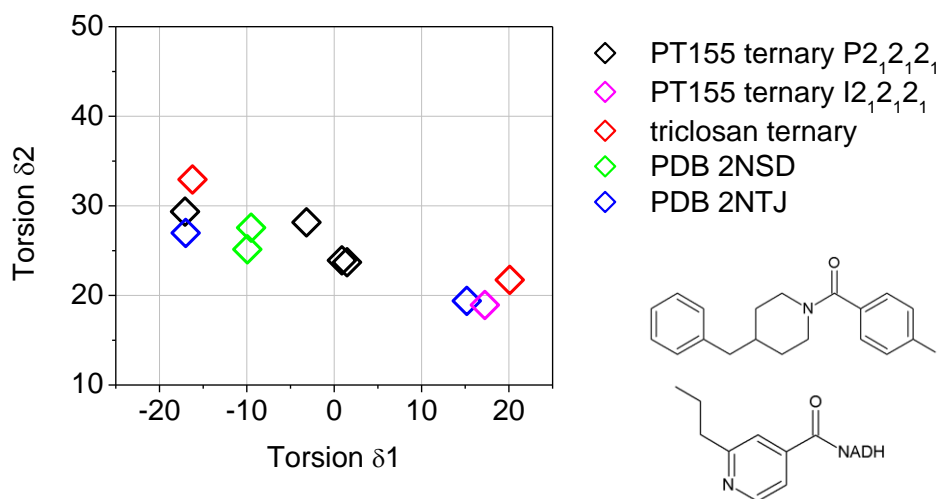
Figure 2.19 Structures of the PT155 complex and other reported InhA ternary complexes in the I2$_1$2$_1$2$_1$ crystal form in the 2D plot.
Structures of inhibitors in the PDB structures 2NSD and 2NTJ are shown on the right.

The substrate-like binding mode

In the subunit of the PT155 complex which resembles the binary complex, the C6 aliphatic tail of PT155 swings toward the B ring compared with the orientation found in other diphenyl ether complex structures so that the tail can reach the position occupied by C10 of

the substrate (**Figure 2.20**). This makes it possible for the tail to interact better with helix-6 as well as strand-4 like the substrate does. In addition, part of the Jeffamine molecule from the crystallization buffer becomes ordered and fills the gap between helix-6 and strand-4, thus covering the ligand binding pocket (**Figure 2.21**). Together these interactions hold the SBL in the position close to that found in the binary complex and in the C16-NAC ternary complex (**Figure 2.22**).
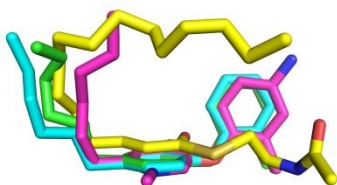
Figure 2.20 Comparison of C6 tail orientation

Magenta: PT155 in $I2_12_12_1$, subunit 2; Green: PT70, 2X23 subunit 1; cyan: PT92, subunit 2; yellow: C16-NAC complex PDB 1BVR, subunit 1
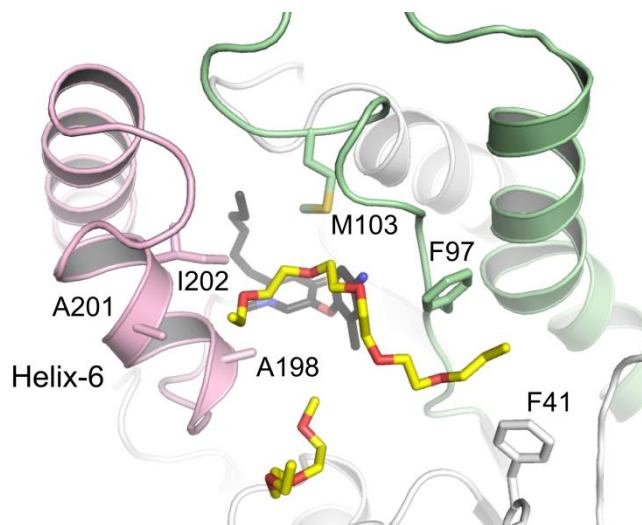
Figure 2.21 The PT155 complex bound with ordered Jeffamine
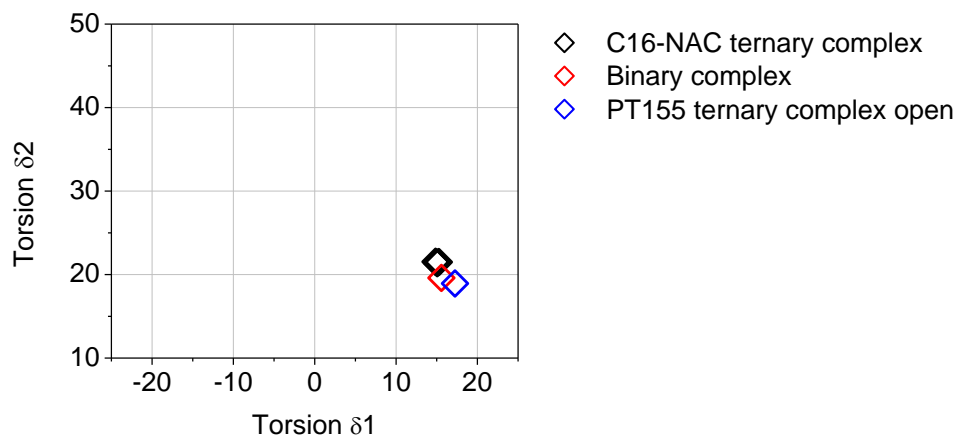
Figure 2.22 PT155 complex in 2D plot in comparison with the binary complex and C16-NAC ternary complex

Subunit 2 in the PT155 ternary complex from the $I2_12_12_1$ crystal, the binary complex (PDB 2AQ8) and the C16−NAC ternary complex (PDB 1BVR)
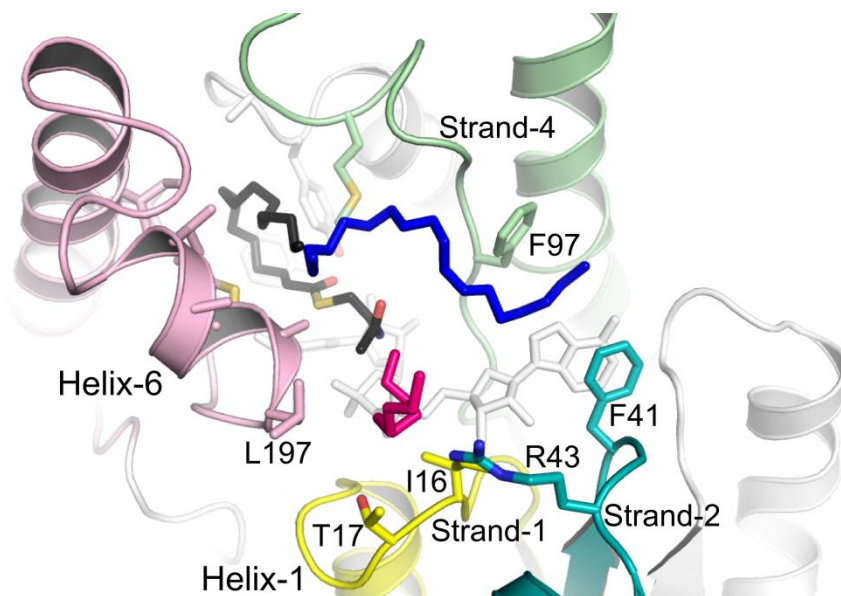
Figure 2.23 Putative binding mode of the long chain fatty acyl substrate and the phosphopantetheine arm that delivers the substrate

In blue and heavy pink are Jeffamine molecules from the PT155 complex structure superimposed on the C16−NAC complex structure 1BVR.

Binding model of long chain fatty acyl substrate

The structural overlay of the bound Jeffamine molecules and the C16–NAC complex reveals how a long chain fatty acyl substrate > C30 may bind to InhA (**Figure 2.23**). After passing the hydrophobic crevice formed between helix–6 and strand–4, the long chain further exploits another hydrophobic binding site formed by strand–4, strand–2 and the adenosine portion of NAD. On the other hand, the phosphopantetheine arm that links the fatty acyl chain to the acyl carrier protein passes through a binding site between helix–6, strand–1 and strand–2. R43 at the C–terminal end of strand–2 is ideally position to interact with the phosphate on the phosphopantetheine arm. In ecFabI, there is a lysine at this position. Since arginine or lysine is a common strategy to recognize phosphate in the ligand, this binding mode is likely common for InhA and FabI. Thus in the presence of the natural substrate with the long fatty acyl chain and the phosphopantetheine group, the active site is securely enclosed despite helix–6 is open and separate from strand–4, ensuring that the catalysis takes place without perturbation of the bulk solvent. In support of this binding mode of the natural substrate, mutagenesis studies have suggested that the mycobacterial acyl carrier protein (AcpM) interacts with helix–2 of InhA (Dr. Xujie Zhang), which will bring AcpM to a proper location that matches with the putative phosphopantetheine binding site identified here.

Slow-onset inhibition correlates with the outlying conformational space

Since the structures of apo, binary and ternary complex with substrate of InhA are available, the range of the SBL conformational space that is utilized in the catalytic cycle can be depicted (**Figure 2.24**). The space bound with the diphenyl ethers overlaps with this area, but many structures clearly lie outside of this area. Significantly, this outlying area is heavily occupied by slow-onset inhibitors, suggesting that an energy barrier exists at the boundary that prevents InhA from adopting the outlying conformation in the absence of these diphenyl ethers.
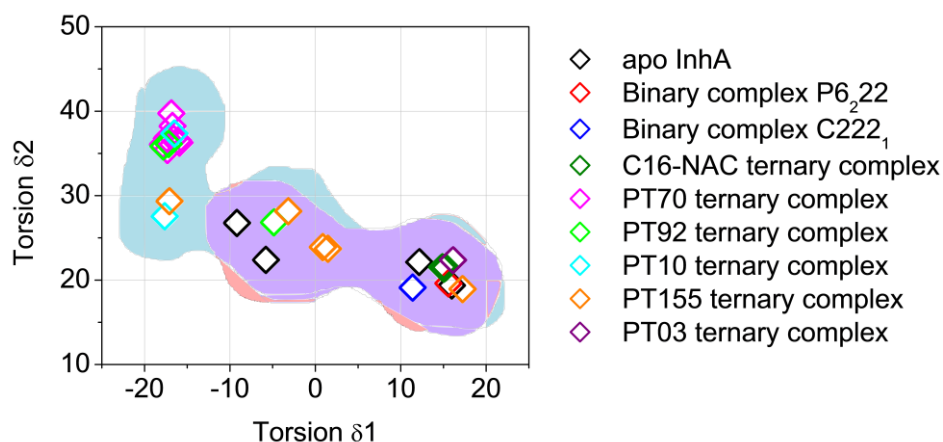
Figure 2.24 Comparison of structures in the absence and presence of diphenyl ethers
The range of space occupied in the catalytic cycle is colored red. The range of space occupied by diphenyl ether-bound complexes is colored blue. The overlapped region is colored purple.

Slow-onset inhibitors induce self-inhibition of InhA

In these outlier structures, helix-6 approaches strand-4 and its hydrophobic residues occupy the space that would normally be occupied by the natural substrate (**Figure 2.25**). Specifically, A201, I202 and V203 occupy the space meant for the long fatty chain while L197 displaces the phosphopantetheine group when PT70 is bound. Thus, the potency of PT70 arises not only from its competition with the thioester moiety and the front portion of the fatty acyl chain of the substrate, but also from the combined effect with helix-6 that competes with the rear portion of the fatty acyl chain as well as the phosphopantetheine arm. In other words, helix-6 itself becomes an inhibitor and binding of PT70 has an impact beyond the protein-inhibitor interface [215]. Presumably, for normal functions of the enzyme such an event should be prevented under normal conditions, i.e., in the absence of the diphenyl ether, by an energy barrier.
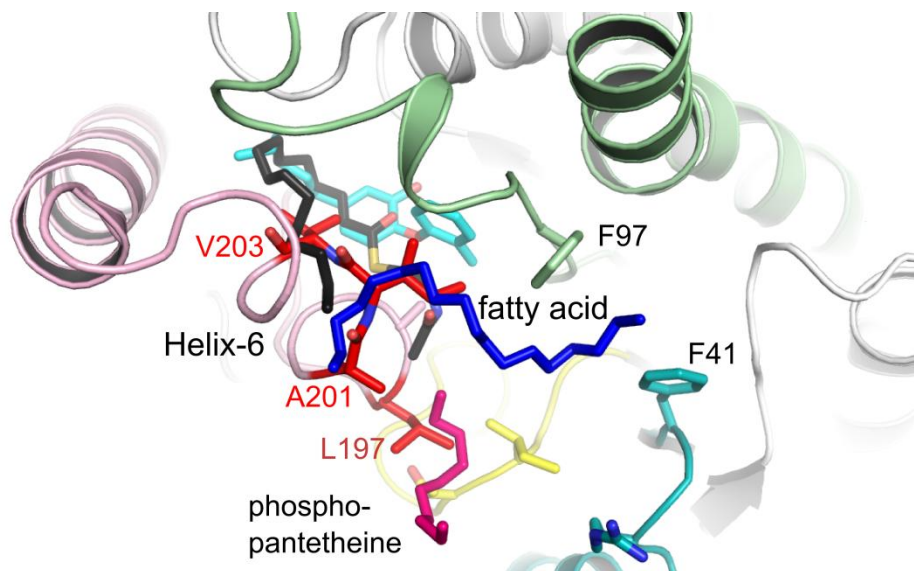
Figure 2.25 Helix-6 competes with the binding site of the natural substrate
Shown is the structure of PT70 complex superimposed with C16-NAC, and Jeffamine molecules in the PT155 complex mimicking the long fatty acyl chain and the phosphopantetheine group of the substrate.

Binding of short hydrophobic chain and chain length selection

Furthermore, the outlier structures suggest that a shorter fatty acyl chain (e.g. ~C10) may increase its binding affinity by stretching into the opposite direction and bind in the hydrophobic space surrounded by helix-7, strand-5, and the C-terminus of a different subunit in the tetramer (**Figure 2.26**). However, if InhA has to cross an energy barrier to adopt this binding mode, the reaction, even if it does take place, cannot compete with that on the long chain substrate which binds readily to InhA [63]. The existence of such an energy barrier can therefore provide a means for InhA to kinetically select against a shorter acyl substrate [188]. In this regard, the diphenyl ethers with a short C6 tail can be considered to mimic a short acyl substrate, and slow-onset inhibition and substrate selection in InhA catalysis can be attributed to the same cause.
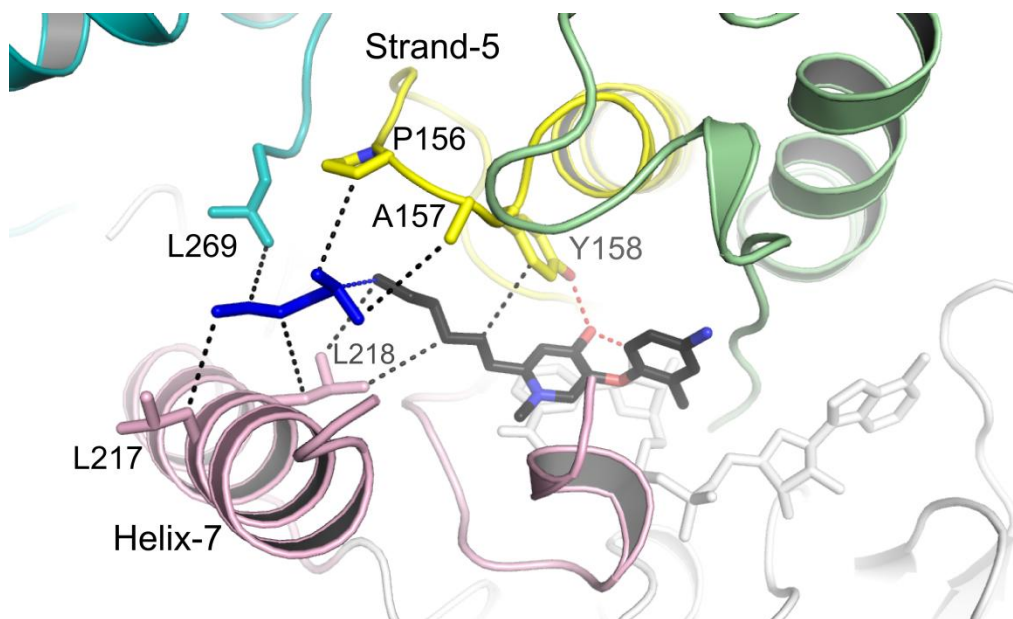
Figure 2.26 Possible binding mode for a short hydrophobic chain.
Shown is one subunit of the PT155 complex residing in the outlying area of the 2D plot. The location of the hydrophobic chain is indicated by the tail of PT155 in black and the buffer molecule in blue.

The slow step of inhibition

Based on the above observations, it may be hypothesized that the slow step in the two-step, induced-fit kinetic mechanism for slow-onset diphenyl ether inhibition (**Figure 2.2**) is the barrier-crossing process over the boundary between the purple and blue area in the 2D plot (**Figure 2.24**). The initial $E \cdot NAD^+ \cdot I$ complex will then most likely have a structure (or a range of structures) with a $\delta 1$ value of $-10°\sim20°$ along the highlighted path while the final $(E \cdot NAD^+ \cdot I)^*$ complex has a structure with $\delta 1$ and $\delta 2$ values in the range of $-15°\sim -20°$ and $25°\sim 40°$, respectively. In this process, helix-6 shifts toward strand-4 and docks into the binding site for the natural long chain fatty acyl substrate while helix-7 is pulled away from strand-5 which opens some room for a short chain substrate (**Figure 2.27**).
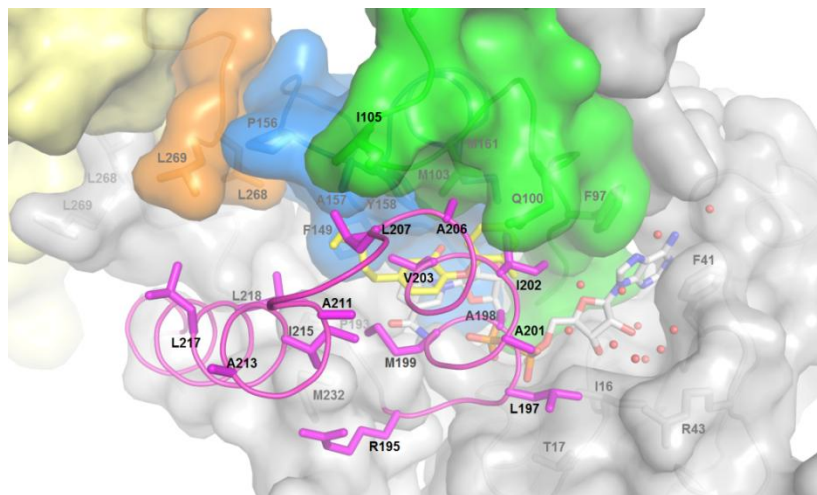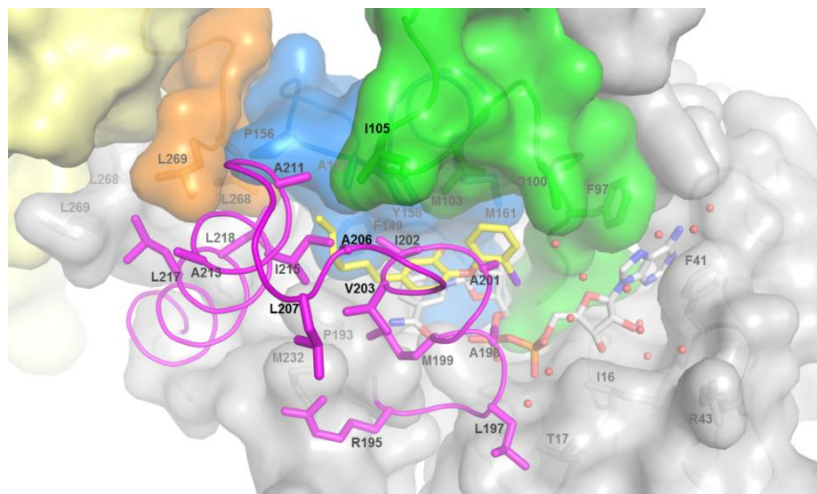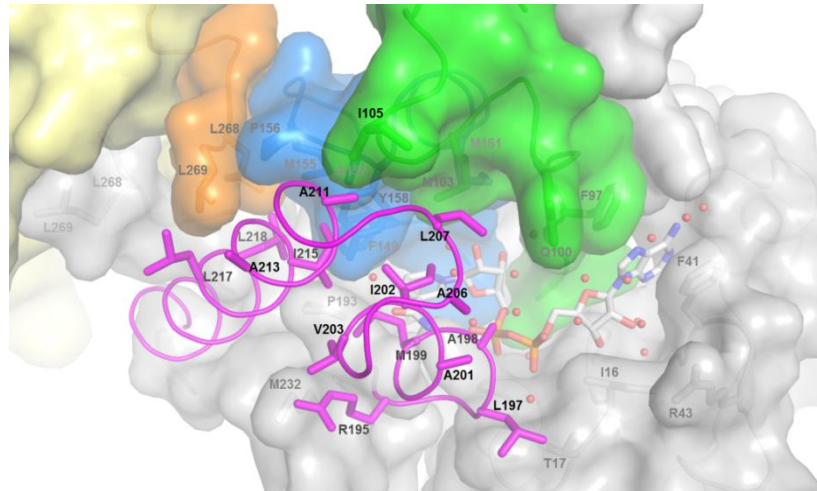
Figure 2.27 Structural transition of InhA to the final inhibited complex

Top: binary complex PDB 2AQ8; middle: PT92 ternary complex subunit 2; bottom: PT70 ternary complex PDB 2X23 subunit 2. SBL is in magenta ribbon, strand-4 in green, strand-5 in blue, and the C-terminus from the separate yellow subunit in the tetramer is in orange. The solvent molecules in these structures also imply the locations prepared for binding the natural long chain fatty acyl substrate.

57

Characterization of the barrier

Given this range of structures, a time-independent molecular dynamics simulation method, partial nudged elastic band (PNEB) [216], can be used to generate a series of intermediate structures that constitute a lowest-energy path between the two ends for the desired complex, which allows us to further characterize the barrier. The simulated path for the PT70 ternary complex suggests that the passing of the side chains of V203 on helix-6 and I215 on helix-7 with respect to each other during the transition to the final complex could cause severe steric hindrance (Eric Cheng Tsung Lai) (**Figure 2.28**). The free energy profile along the path is then calculated using the umbrella sampling methodology [217-220] at the two-dimensional level based on two coordinates describing the structural change (**Figure 2.29**).
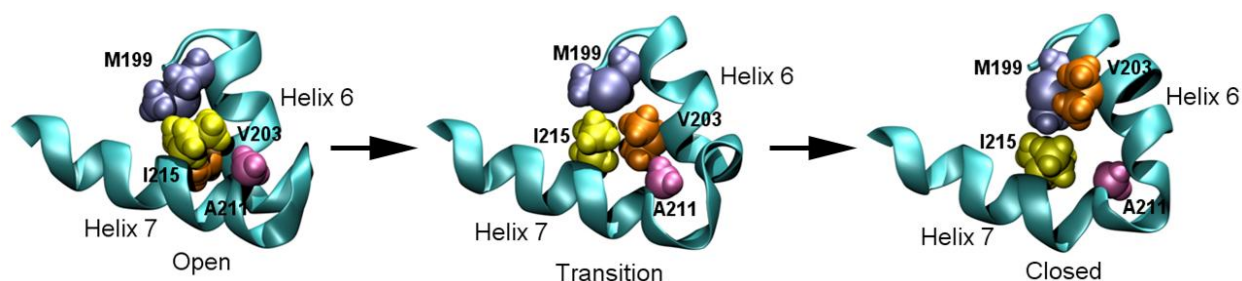
Figure 2.28 Change of relative positions of side chains on helix-6 and helix-7 along the path (Eric Cheng Tsung Lai)
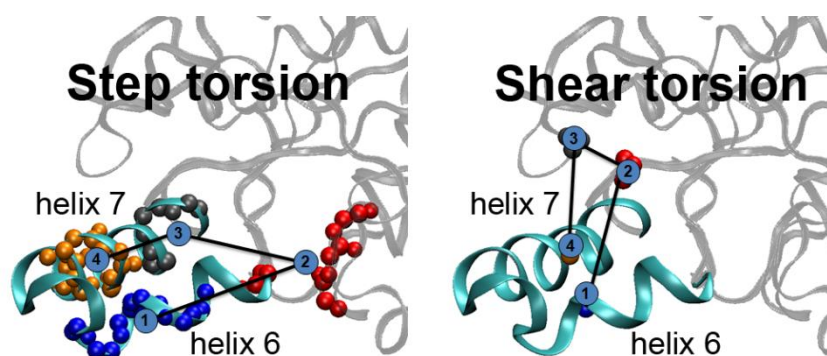
Figure 2.29 Definition of the two coordinates used for umbrella sampling simulation
Left: the first coordinate is the step torsion defined by point 1-4. Point 1 represents the center of mass of backbone atoms from residue 200-205. Point 2 represents the center of mass of backbone atoms from residue 19-21 and 196. Point 3 represents the center of mass of backbone atoms from residue 219-222. Point 4 represents the center of mass of backbone atoms from residue 211-216. Right: the second coordinate is the shear torsion defined by point 1-4. Point 1 represents C_β of residue 203. Point 2 represents the center of mass of backbone atoms from residue 98. Point 3 represents the center of mass of backbone atoms from residue 158. Point 4 represents C_β of residue 215 (Eric Cheng Tsung Lai).

58

Free energy profiles in the presence and absence of diphenyl ethers

The 2D free energy profiles have been calculated for the binary complex and the ternary complexes of PT70, PT92, PT155 and PT03 (Eric Cheng Tsung Lai) (**Figure 2.30**). A barrier on the scale of 5 kcal/mol can be clearly identified for the binary complex and the PT70, PT92 ternary complexes. This size of barrier is consistent with the scale of the rate constants $(k\sim10^{-4} \text{ s}^{-1})$ we observe in the slow step of inhibition.

Consistency with crystal structures

The energy profiles are highly consistent with the population of crystal structures on the 2D plot (**Figure 2.24**). As expected, the slow-onset PT70 prefers the conformation on the opposite side of the barrier in contrast to the complex without the inhibitor. PT92 shares this feature with PT70 although some structural deviation from the PT70 complex is expected. On the contrary, the energy profile for the PT155 complex is similar to that of the binary complex in that it prefers the conformation on the same side of the barrier. However, some population shift toward the other side is observed as seen in the crystal structures (**Figure 2.19**). The energy profile in the presence of PT03 becomes much shallower and there is no preference for either side of the barrier, which leads to the lack of density seen in the crystal structure [157] due to either dynamic or static disorder [221]. In fact, the un-modeled electron density in the map of the PT03 complex structure can be interpreted as low occupancy of multiple conformations (**Figure 2.31**), which is consistent with its energy profile.
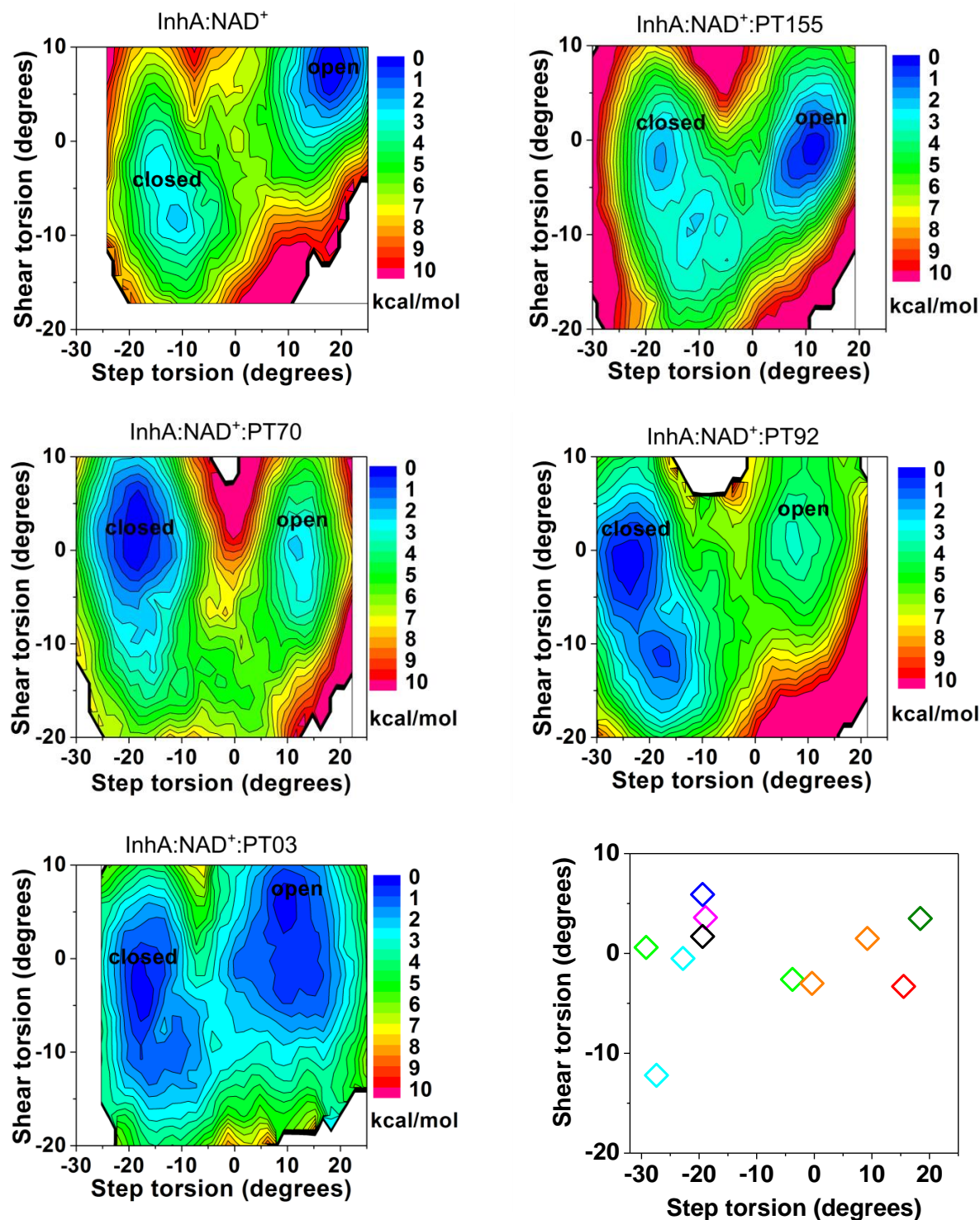
Figure 2.30 2D free energy profiles of the binary complex and the ternary complexes of PT70, PT92, PT155 and PT03

The last panel shows the crystal structures located by the same coordinates defined in Figure 2.29. Selected subunits are shown for the following complexes: ◊ Binary complex 2AQ8 ◊ C16−NAC ternary complex 1BVR ◊ PT70 ternary complex 2X23 ◊ PT10 ternary complex ◊ PT91 ternary complex ◊ PT92 ternary complex ◊ PT130 ternary complex ◊ PT155 ternary complex (Eric Cheng Tsung Lai)

Selected subunits from the crystal structures are plotted on the same coordinates in order to directly compare the energy profiles and the distribution of crystal structures (**Figure 2.30**). The resulting plot follows the same trend as the 2D plot in the previous sections (**Figure 2.24**). Crystal structures of the PT70 and PT92 complexes are observed at the positions that match well with the minima in the energy profiles. However, one subunit from each of the PT92 and PT155 complexes appears to sit on the barrier around step torsion zero. These two subunits are the same as the ones around ($\delta 1$, $\delta 2$) = (−3, 28) in **Figure 2.24**. For both of them, crystal contacts might have played a more important role in stabilizing these high-energy conformations. Nevertheless, these structures provide important insight into the intermediates for the enzyme to complete the major conformational change, which will be discussed later.
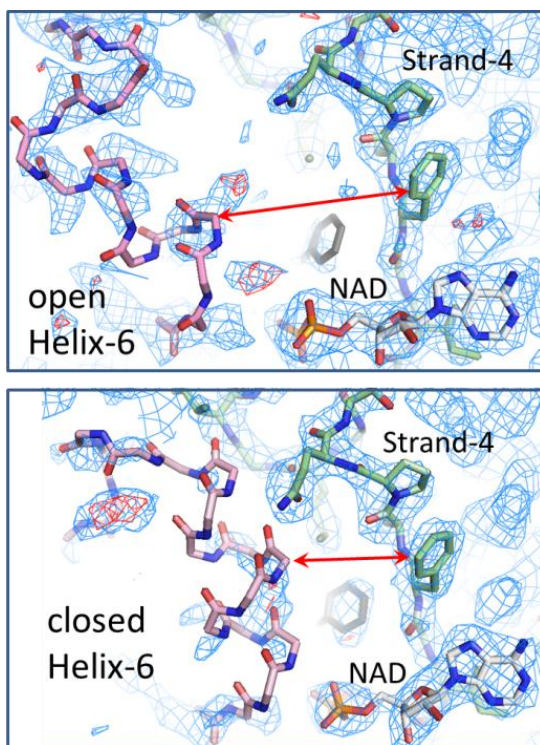
Figure 2.31 Residual SBL density in the PT03 complex structure

Density maps from two subunits in the PDB structure 2B36 are shown in blue mesh. NAD^+, PT03 and strand-4 are shown in stick. Helix-6 is disordered, and models of open and closed conformations are superimposed.

Consistency with kinetic observations

The above results indicate that slow-onset inhibition kinetics is only observed (PT70, PT92) when the driving force to cross the barrier is sufficient. The minima behind the barrier are 2-3 kcal/mol more stable than the minima before the barrier for the PT70 and PT92 complexes. This is not the case for PT155, and slow-onset behavior is not observed. For PT03, slow-onset is not observed possibly also because of the lack of the strong driving force. The close stabilities on both sides of the barrier makes $E·NAD^+·I$ and $(E·NAD^+·I)^*$ kinetically indistinguishable [169]. Secondly, the reduced size of the barrier for both directions could have made the rate-determining step so fast that it is no longer observed on the timescale of the enzyme assay (minutes).

Two-step vs one-step slow-onset kinetics

The energy profiles for PT70 and PT92 also show that the sufficient stability before crossing the barrier is the basis of the kinetic observation of the two-step binding. If the initial $E·NAD^+·I$ is high-lying then the initial binding will be weak and the slow-onset process will appear as one-step.

Induced-fit vs conformation-selection

The energy profiles help us see the reason why PT70 binding apparently proceeds through an induced-fit (**Figure 2.32**, route 2) mechanism instead of conformation-selection (**Figure 2.32**, route 1). The sufficient stability of the initial PT70 complex as well as the overlapping of its conformation with that of the preferred binary complex is essential for the observation of the induced-fit route. Consider the situation where the initial ternary complex is much more unstable and the energy profile of the binary complex has a much smaller barrier, the conformation-selection route will be expected to predominate.
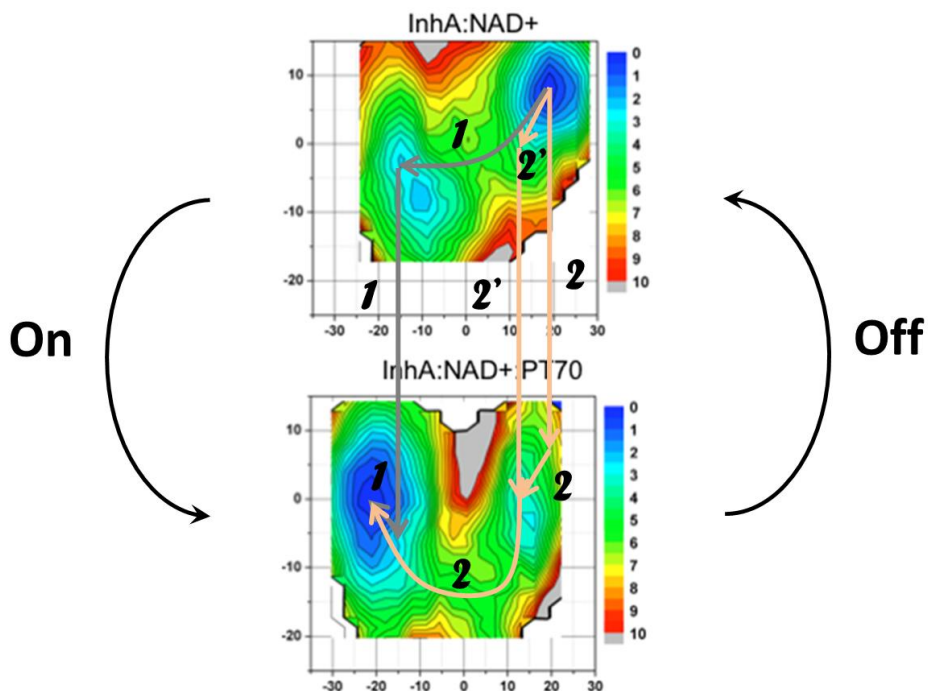
Figure 2.32 The on-route of PT70 binding

Route 1 (grey): conformation-selection, or binding after the rate-determining conformational change step. Route 2 (salmon): induced-fit, or binding before the rate-determining conformational change step.

Contributing factors of the barrier

In this binding reaction, E•NAD⁺ + I <=> E•NAD⁺•I, the intrinsic binding energy [215] for any conformer, $\Delta\Delta G'(\delta 1, \delta 2)$, in the 2D plot is given by,

$$\Delta\Delta G'(\boldsymbol{\delta 1}, \boldsymbol{\delta 2}) = \Delta G'_{E \cdot NAD^+ \cdot I}(\boldsymbol{\delta 1}, \boldsymbol{\delta 2}) - \Delta G'_{E \cdot NAD^+}(\boldsymbol{\delta 1}, \boldsymbol{\delta 2}) - \Delta G'_I(\boldsymbol{\delta 1'}, \boldsymbol{\delta 2'})$$

In other words, the energy profile of the ternary complex is the superposition of energy profiles from three components [222],

$$\Delta G'_{E \cdot NAD^+ \cdot I}(\delta 1, \delta 2) = \Delta\Delta G'(\delta 1, \delta 2) + \Delta G'_{E \cdot NAD^+}(\delta 1, \delta 2) + \Delta G'_I(\delta 1', \delta 2')$$

,where δ1', δ2' can be considered as the "complementary coordinates" of δ1 and δ2 [223].

If the structure of I does not change much when there is a conformational change of the ternary complex from (δ1, δ2) to (δ3, δ4), then this term $\Delta G_I'(\delta1', \delta2')$ is negligible when we consider the change in $\Delta G_{E \cdot NAD^+ \cdot I}'(\delta1, \delta2)$.

Comparison with the energy profile of the binary complex reveals that the height of the forward side of the barrier for the PT70 ternary complex appears to be dominated by the energy change from the binary complex while the height of the reverse side is dominated by the change in the binding energy. In other words, in the uphill stage of the barrier-crossing process, the ternary complex is primarily trying to overcome an unfavorable structural change of the enzyme whereas in the downhill stage the stability is primarily achieved by optimization of binding.

Reconstruction of the path of conformational change

To identify the unique features of the protein conformational change that distinguish it from others such as loop motions and domain motions that could happen on the timescale of μ s to ms [181,224,225], we need to further compare the structures of the intermediates along the coordinates of the conformational change to those at the starting and end points. The position of the local minimum in the energy profile of the binary complex suggests that the protein structure before inhibitor binding in solution might be even better represented by the crystal structure of the C16-NAC ternary complex than that of the binary complex, 2AQ8 (**Figure 2.30**). Nevertheless, it has been noted that there is no major structural reorganization and no hydrogen bonds are broken upon binding of C16-NAC despite that the SBL shifts outward by 4 Åto accommodate the substrate [188]. The small energy difference between the two conformations suggests that they can interconvert in solution much faster than the slow-onset inhibition timescale. Therefore, the crystal structure 2AQ8 is still a good representation of the binary complex.

The energy profiles of PT70, PT92 and PT155 all display a local minimum for the initial ternary complex around torsion angles (10, 0), a slight shift from the minimum for the binary complex around (18, 7) (**Figure 2.30**). This position overlaps with the conformational space occupied by subunit 3 and 4 in the crystal structure of the PT155 complex in $P2_12_12_1$, which

reveals the adjustments the enzyme has made from a binary complex to establish initial binding with the diphenyl ether. The torsion angles of (10, 0) are equivalent to ($\delta 1$, $\delta 2$) = (1, 24). Subunit 2 in the crystal structure of the PT155 ternary complex is essentially the same as the binary complex structure 2AQ8 (**Table 2.17**), and may represent the very initial complex before the enzyme has made the first round of adjustments towards torsion angles of (10, 0).

Table 2.18 RMSD values for structures in the high δ 1region

RMSD (Å)	Binary[1]	Binary[2]	C16 ternary[3]	PT155I–2[4]	PT155P–3[5]
Binary[1]	0				
Binary[2]	0.58	0			
C16 ternary[3]	0.76	0.64	0		
PT155I–2[4]	0.32	0.76	0.88	0	
PT155P–3[5]	1.35	1.18	1.20	1.48	0

1 PDB structure 2AQ8
2 Binary structure 21B3 subunit 3
3 PDB structure 1BVR
4 PT155 ternary complex structure in $I2_12_12_1$, 126D5_2 subunit 2
5 PT155 ternary complex structure in $P2_12_12_1$, 126D5_1 subunit 3

The crystal structure of subunit 1 of the PT155 ternary complex in $P2_12_12_1$ and subunit 2 of the PT92 ternary complex are both located in the vicinity of the barrier. The SBL from the two subunits align with an RMSD value of 1.32 Å (**Table 2.17**) and the latter structure is arbitrarily chosen for the following comparison.

After crossing the barrier, the local energy minimum for the binary complex lies around torsion angles (−11, −8), and overlaps to varied extent with the minimum for each of the ternary complexes (**Figure 2.30**). The protein structure of subunit 1 in the ternary complex of PT10 represents this conformational state that has passed the major barrier but further

optimization of the stability of the ternary complex might still be available for certain inhibitors such as PT70. Finally, subunit 2 from the PT70 ternary complex structure 2X23 is chosen to represent the structure in the ternary complex that is further optimized primarily by moving toward higher shear torsion angles.

Barrier-crossing coupled with refolding

Change in secondary structure

With this chosen set of representative structures that connect the whole barrier-crossing process through smaller steps, we can characterize the barrier with more details. The protein component alone, $\Delta G'_{E \cdot NAD^+}(\delta 1, \delta 2)$, is first considered here. It has been suggested that breaking and making large numbers of hydrogen bonds can provide a barrier for interconversion of protein conformers [215]. The secondary structure of helix-6 and helix-7 reveals that the SBL, helix-6 in particular, cannot shift as an intact helix along the barrier-crossing path on the conformational space spanning δ1 values of 20 to −20 (**Figure 2.33**). In the binary complex, helix-6 maintains ideal α-helical properties, but this structure is disrupted by traveling down the reaction coordinate (**Table 2.19**). The number of residues per turn in the α-helix increases from the ideal value of 3.6 to 4.8 for the structure on the front edge of the barrier, and the helix is distorted as indicated by the angle of deviation from the ideal value of 0°. The SBL finds a new conformation with comparable helicity to the native conformation after passing the barrier and these numbers drop. However, within this new SBL conformation, the stability of the ternary complex can be further optimized by a slight twist mainly within helix-6. Hence there is a slight increase of these numbers again at the end point of the reaction coordinate.
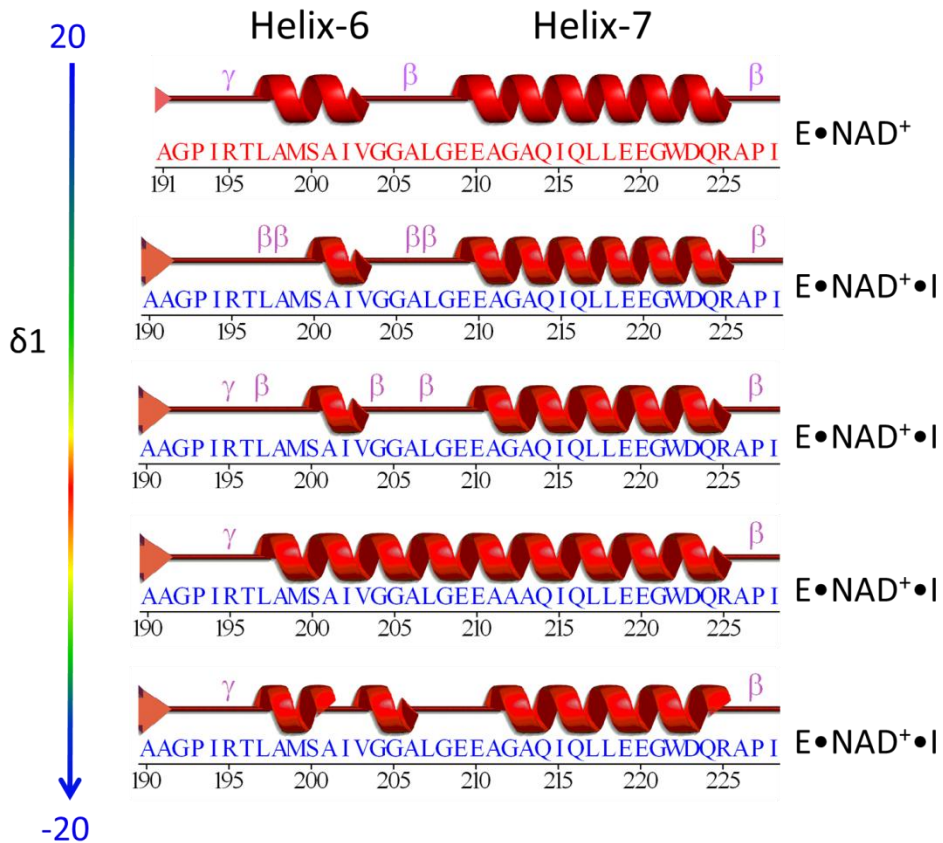
Figure 2.33 Change in the secondary structure across δ1=16 to -18

δ1 is defined as in **Figure 2.11**. The top structure is represented by the binary complex InhA:NAD$^+$ (PDB structure 2AQ8), the second by the PT155 ternary complex in P2$_1$2$_1$2$_1$, subunit 3, the third by the PT92 ternary complex, subunit 2, the fourth by the PT10 ternary complex, subunit 1 and the last by the PT70 ternary complex, 2X23, subunit 2.

Table 2.19 Secondary structure parameters for helix-6 across δ1=16 to -18
The first binary complex is represented by the PDB structure 2AQ8, the second complex by the PT155 complex structure in $P2_12_12_1$, subunit 3, the third by the PT92 complex structure, subunit 2, the fourth by the PT10 ternary complex, subunit 1, and the last by the PT70 complex structure, 2X23, subunit 2.

Complex	$\delta 1$	Start	End	Type	No. residues	Residues per turn	Deviation from ideal (°)
E·NAD$^+$	16	L197	V203	H	7	3.62	7.9
E·NAD$^+$·I	1	S200	V203	H	4	3.88	38.6
E·NAD$^+$·I	-5	S200	V203	H	4	4.81	32.2
E·NAD$^+$·I	-18	L197	L207	H	11	3.72	12.3
E·NAD$^+$·I	-17	L197	A201	H	5	3.78	24.5
		V203	A206	H	4	3.60	39.6

Hydrogen bonding patterns in the main chain

The change of hydrogen bonding pairing within the main chain again reflects this disruption (**Figure 2.34, Table 2.20**). The trend in the loss and gain of energy by breaking and formation of main chain hydrogen bonds (**Figure 2.35**) matches the shape of the energy profile of the binary complex especially prior to reaching the barrier. Although the inhibitor might not be binding to the lowest-energy conformer of the binary complex, the model structures along with the energy profiles of the ternary complexes still suggest that it takes at least the energy of 8 hydrogen bonds for the ternary complex to undergo this conformational change. This energy-costly process is driven by the interactions with the diphenyl ether. In contrast, the conformational change caused by binding C16-NAC does not require such a large-scale rearrangement of hydrogen bonding [188].

$\delta 1=16$

$\delta 1=1$

δ 1=−5
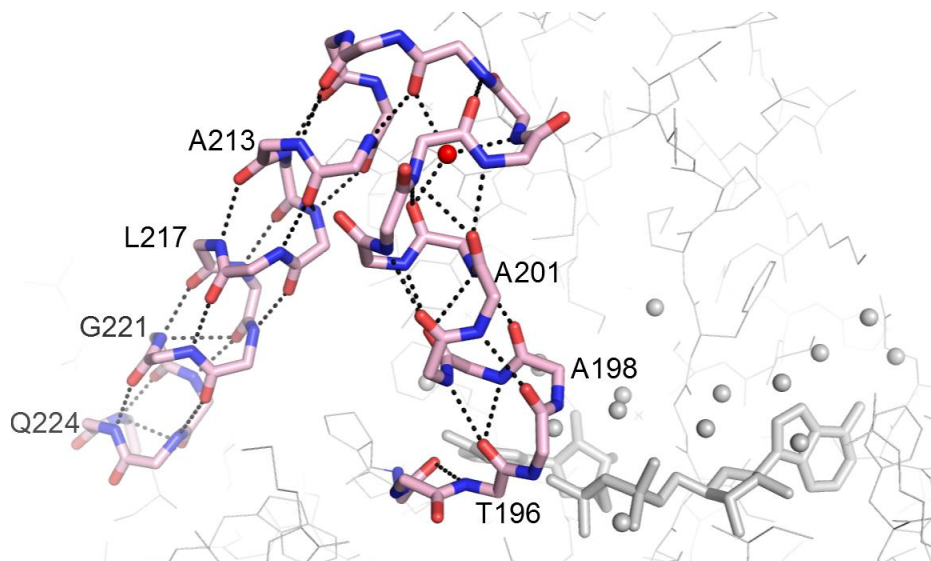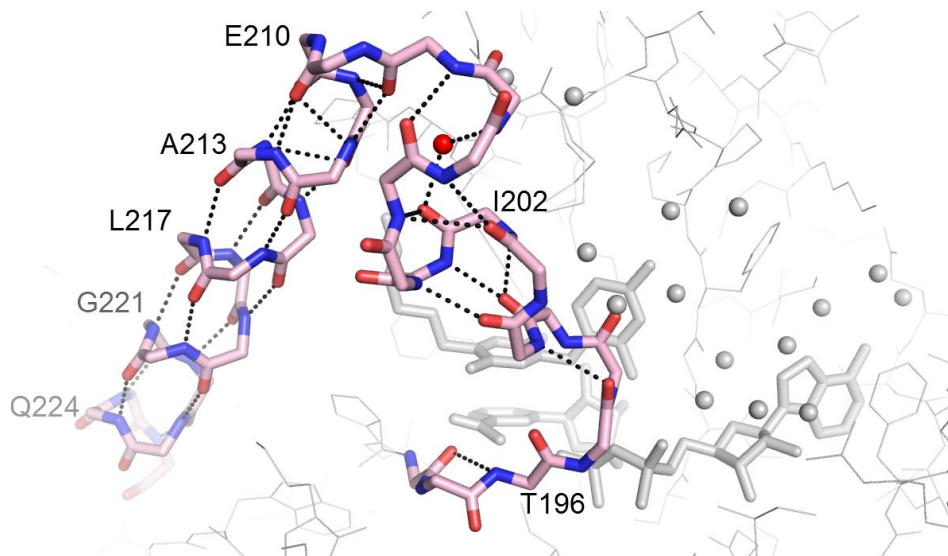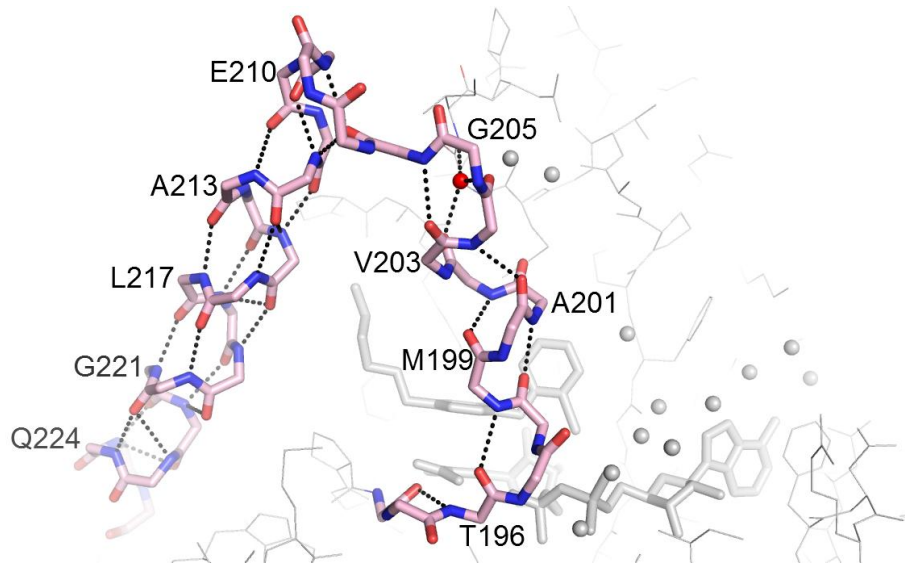
δ 1=−18
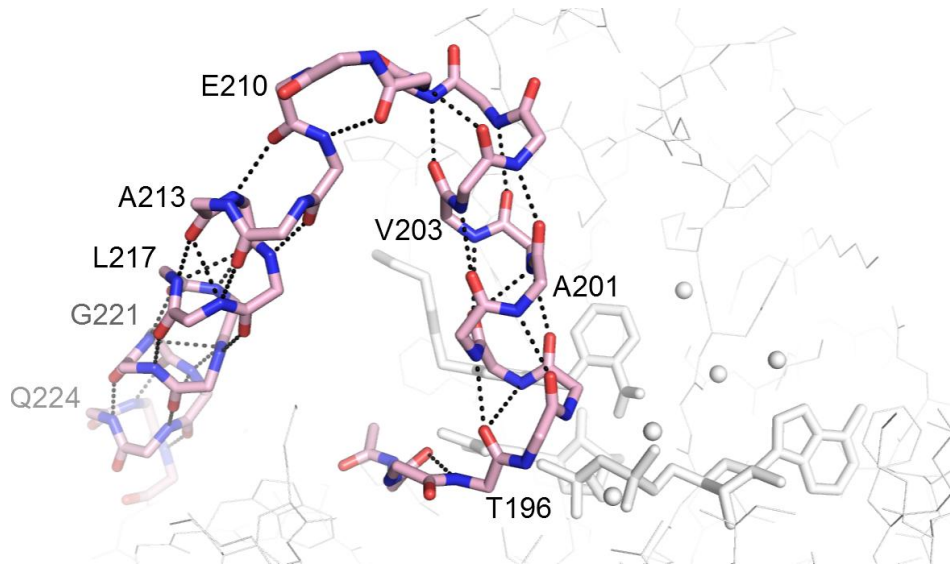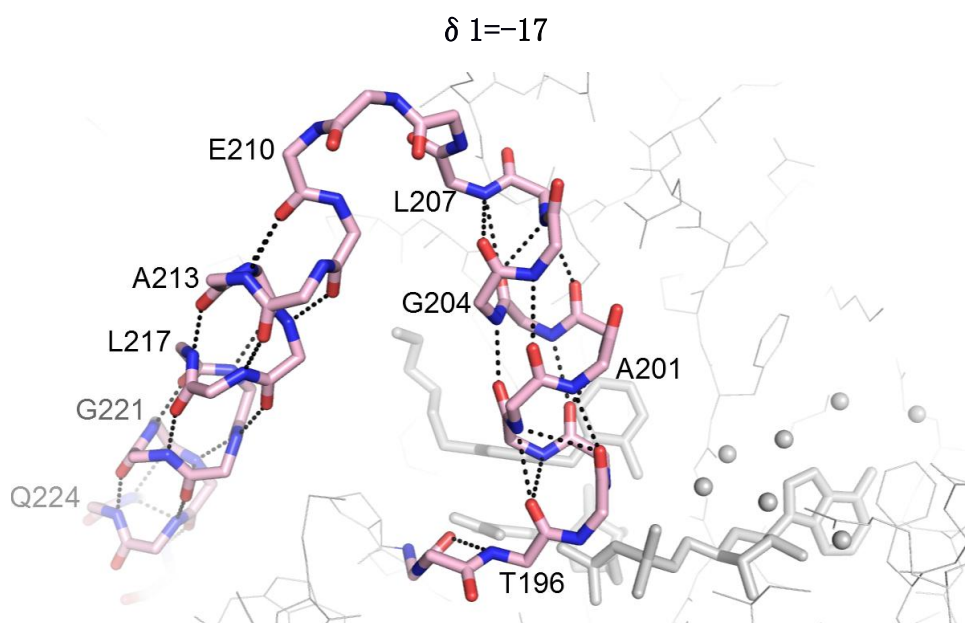
Figure 2.34 Change in the hydrogen bonding pattern for secondary structure formation across δ1=16 to -18

Hydrogen bonds within the main chain are represented by dash. Side chains are not shown for clarity. The first structure is represented by the binary complex PDB structure 2AQ8, the second by the PT155 ternary complex in $P2_12_12_1$, subunit 3, the third by the PT92 ternary complex, subunit 2, the fourth by the PT10 ternary complex, subunit 1, and the last by the PT70 ternary complex, 2X23, subunit 2. Water molecules at ligand binding sites are shown in spheres.

Table 2.20 Change of hydrogen bonding patterns across δ1=16 to -18

The first binary complex is represented by the PDB structure 2AQ8, the second complex by the PT155 complex structure in P2$_1$2$_1$2$_1$, subunit 3, the third by the PT92 complex structure, subunit 2, the fourth by the PT10 ternary complex, subunit 1, and the last by the PT70 complex structure, 2X23, subunit 2.

δ 1=16		δ 1=1		δ 1=-5		δ 1=-18		δ 1=-17	
218	214	218	214	218	214	218	214	218	214
217	213	217	213	217	213	217	213	217	213
216	212	216	212	216	212	216	212	216	212
215	211	215	211	215	211	215	211	215	211
214	210	214	210	214	210	214	210	214	210
213	209	213	209	213	209	213	209	213	209
212	208	212	208	212	208	212	208	212	208
211	207	211	207	211	207	211	207	211	207
210	206	210	206	210	206	210	206	210	206
209	205	209	205	209	205	209	205	209	205
208	204	208	204	208	204	208	204	208	204
207	203	207	203	207	203	207	203	207	203
206	202	206	202	206	202	206	202	206	202
205	201	205	201	205	201	205	201	205	201
204	200	204	200	204	200	204	200	204	200
203	199	203	199	203	199	203	199	203	199
202	198	202	198	202	198	202	198	202	198
201	197	201	197	201	197	201	197	201	197
200	196	200	196	200	196	200	196	200	196
199	195	199	195	199	195	199	195	199	195
198	194	198	194	198	194	198	194	198	194
197		197		197		197		197	
196		196		196		196		196	

--·--·--·-- Via water
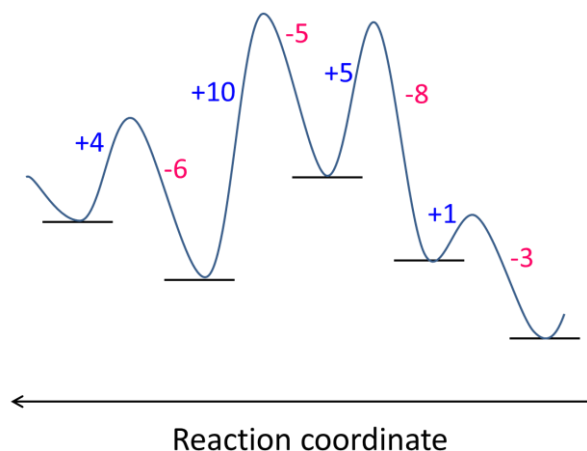
Figure 2.35 Numbers of main chain hydrogen bonds broken and formed in steps crossing the barrier

Side chains, neighboring structural elements and other interactions

Concomitant with the change in the secondary structure of helix-6 and helix-7 is the change in their backbone and side chain interactions with each other and with several nearby turning elements between strand-4 and helix-4, strand-5 and helix-5, strand-1 and helix-1, helix-7 and helix-8, and the C-terminus of the neighboring subunit (**Figure 2.27**). This constitutes a local but large-scale refolding event. Besides the energy involved in main chain hydrogen bonding discussed above, other energy components involving these different interactions of backbone and side chains may also contribute to the barrier. For instance, a salt bridge formed between the side chains of R195 and E219 from helix-6 and helix-7, respectively, apparently important in maintaining the native conformation of the SBL appears to be broken along the reaction coordinate. In addition, hydrophobic interactions are crucial in guiding protein folding [226-228], and the size of barrier for folding may be correlated to factors such as the length of the sequence and the surface area involved [229-231]. Since the burial of surface area can provide a clue as to how well the structure is folded, the solvent accessible surface area for the different conformers along the reaction coordinate is examined (**Table 2.21**). Indeed, the protein conformation adopted by the binary complex has a significantly smaller exposed area (11923 Å) compared to the conformations adopted by the ternary complexes (12247–12501 Å), indicating that the protein sacrifices its well-folded structure to interact with the inhibitor.

Table 2.21 Solvent accessible surface area and surface area buried by the ligands (Å) in each of the conformations along the reaction coordinate

Ternary 0 is represented by the PT155 complex structure in $I2_12_12_1$, subunit 2, ternary 1 by the PT155 complex structure in $P2_12_12_1$, subunit 3, ternary 2 by the PT92 complex structure, subunit 2, ternary 3 by the PT10 ternary complex structure, subunit 1, and ternary 4 by the PT70 complex structure, 2X23, subunit 2.

	Binary	Ternary 0	Ternary 1	Ternary 2	Ternary 3	Ternary 4
Total protein ASA	11923	12246.9	12501.3	12346.3	12448.4	12490.2
BSA by inhibitor	0	223.5 (1.8%)	286.6 (2.3%)	243.5 (2.0%)	293.9 (2.4%)	231.2 (1.9%)
BSA by NAD	404.6 (3.4%)	382.5 (3.1%)	409.3 (3.3%)	407.8 (3.3%)	416.9 (3.3%)	386.7 (3.1%)
Total inhibitor ASA		558.4	564.8	543.4	549.1	533.3
BSA by protein		401.4 (71.9%)	439.5 (77.8%)	414.9 (76.4%)	403.4 (73.5%)	412.6 (77.4%)
BSA by NAD		126.1 (22.6%)	125.3 (22.2)	128.5 (23.6%)	143.3 (26.1%)	120.6 (22.6%)

ASA: accessible surface area
BSA: buried surface area

Table 2.22 Solvent accessible surface area for selected residues and that buried by the inhibitor (Å) in each of the conformations along the reaction coordinate

The same structures from Table 2.21 are used.

	Binary	Ternary 0	Ternary 1	Ternary 2	Ternary 3	Ternary 4
I16	68.07	59.53	43.17	69.02	70.71	37.95
T17	27.89	35.13	32.62	27.80	32.79	26.35
I95	29.18	32.60	33.79	32.91	34.63	34.65
		0	0	0	0.78	1.26

G96	48.75	46.33	40.23	45.41	47.11	37.93
		20.15	16.00	18.83	21.76	14.83
F97	85.40	82.84	89.35	80.16	77.35	66.87
		12.52	11.87	6.74	5.86	1.34
M98	15.36	12.68	15.19	16.86	19.19	2.38
		6.96	11.72	8.88	7.66	1.40
M103	21.82	30.78	10.61	8.66	17.22	1.84
		21.95	6.67	5.34	4.97	1.84
F149	24.37	31.96	30.11	28.29	38.05	32.66
		19.29	18.02	17.64	24.66	21.03
M155	36.05	44.24	42.56	43.30	65.90	65.93
		1.23	2.48	2.42	12.15	7.08
P156	57.72	53.94	66.15	54.54	122.41	120.84
		0	5.46	4.75	9.91	4.17
A157	4.49	9.53	4.77	5.89	40.88	18.86
		3.66	0.84	0.50	12.29	10.88
Y158	21.43	33.66	19.42	17.81	42.41	28.09
		31.70	17.24	14.75	39.07	25.13
M161	21.35	17.07	20.85	19.83	18.62	17.54
		12.77	15.78	14.41	13.95	13.24
K165	19.44	15.20	15.87	14.13	14.55	13.65
		4.39	4.54	4.31	4.39	4.14
P193	5.19	8.04	19.74	9.88	10.01	5.86
		4.22	14.48	6.59	6.29	3.29
I194	12.89	16.95	28.42	12.85	10.98	10.44
		6.78	15.51	3.61	1.00	2.67
R195	63.68	77.78	83.77	120.90	88.32	87.32

T196	31.96	31.89	35.18	21.08	14.42	21.24
		8.86	8.83	3.33	1.70	4.06
L197	123.96	129.81	84.40	130.01	106.95	64.49
A198	65.63	64.59	92.74	66.52	62.65	52.95
		12.18	45.49	36.25	24.66	26.02
M199	19.64	23.96	53.23	36.40	35.00	22.61
		20.48	49.60	16.48	25.04	15.75
A201	46.24	43.24	47.99	59.73	69.09	56.20
		0	0	16.22	0	0
I202	20.08	30.78	12.51	35.54	85.08	36.82
		17.80	0	34.82	43.02	9.88
V203	23.79	30.10	26.92	53.35	29.73	30.03
		0	7.20	0	6.19	30.03
A206	62.65	68.40	84.99	0	16.09	4.52
L207	24.02	18.97	67.12	155.22	71.00	30.48
		0.67	0	0	0	2.34
A211	0	0.67	0.17	20.30	6.19	13.65
I215	2.01	14.02	19.91	12.70	13.88	17.42
		14.02	19.23	9.53	0	3.68
L217	121.48	122.85	135.02	115.20	98.78	108.02
(X)	16.43	14.57	20.19	28.26	0	18.33
(Net)	105.05	108.28	114.83	86.94	98.78	89.69
L218	26.87	33.52	45.36	45.28	72.33	60.13
(I)		3.86	14.72	14.05	28.56	27.18
(X)	26.75	23.98	26.04	24.88	0	0
(Net)	0.12	9.54	19.32	20.4	28.56	27.18
M232	47.28	41.82	31.70	15.84	0.17	0.12

L268	162.15	161.47	164.72	170.89	177.64	157.14
(X)	122.34	134.85	141.08	146.18	134.68	107.27
(Net)	39.81	26.62	23.64	24.71	42.96	49.87
L269	197.18	188.88	173.06	182.19	179.85	195.32
(X)	130.83	107.67	116.46	130.57	44.72	66.29
(Net)	66.35	81.21	56.60	51.62	135.13	129.03

(X): BSA by the neighboring subunit

To see the change of the exposure of each residue surrounding the inhibitor binding site and how this might contribute to the energy barrier, the accessible surface area for each residue in the different conformers is plotted against the reaction coordinate (**Table 2.22, Figure 2.36**). The plots show that some residues around the binding site, such as T17, I95 and M161, do not change their exposure significantly along the reaction coordinate. Some display a trend of decreasing their surface exposure, such as F97, A206 and M232. Some display a trend of increase, such as M155, P156, A157, A201, I215 and L218. Residues R195, V203, L207 and A211 all display a spike in the high-energy intermediate 2 (point 3), suggesting that they are the critical contributors to the barrier in the energy profile of the binary complex, which is also carried over into the profile of the ternary complex. While R195 is a common surface residue and the exposure of V203 provides the opportunity to interact with the inhibitor, the exposure of L207 is clearly an unfavorable component.

It is also found that the C-terminal residues L269 and L268 from the neighboring subunit lose significant interactions with helix-7 after the conformational change (**Figure 2.36**, last panel), which extends the impact of inhibition beyond a single subunit. The proximity of two SBLs from two adjacent subunits in the InhA tetramer implies that there will be changes in inter-subunit interactions during the conformational change induced by inhibition (**Figure 2.37**). This factor has not been incorporated into the calculation of energy profiles since the simulations are performed with a monomer, but might also contribute to the actual energy barrier.
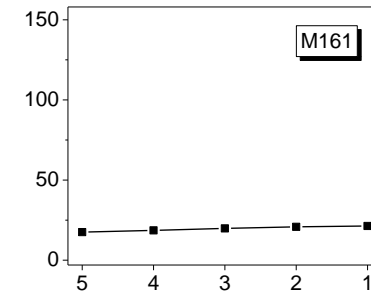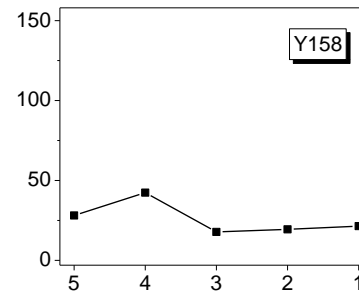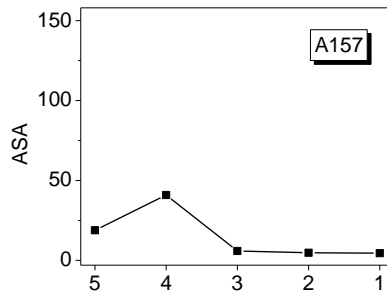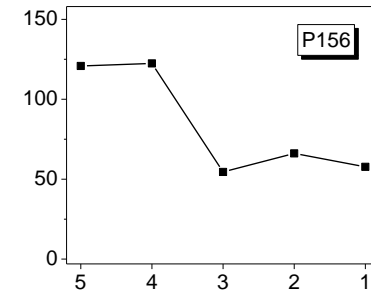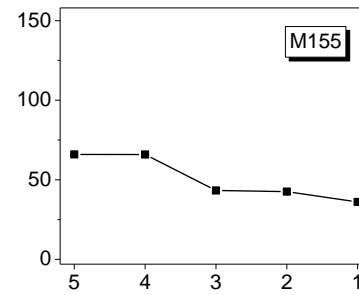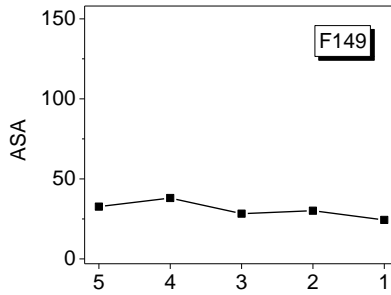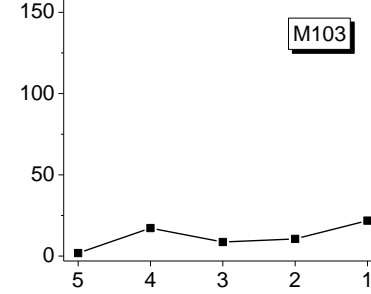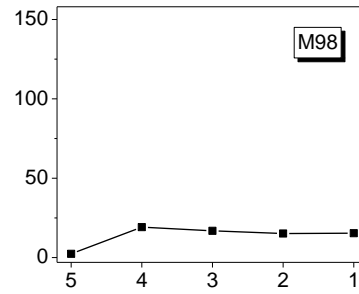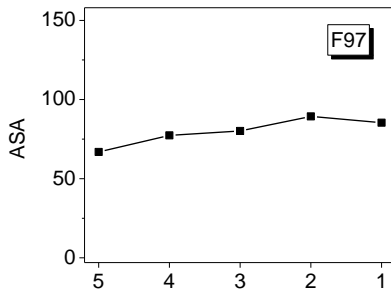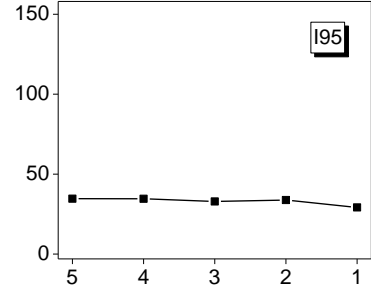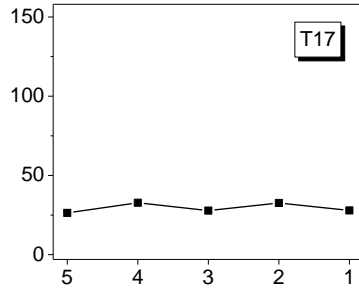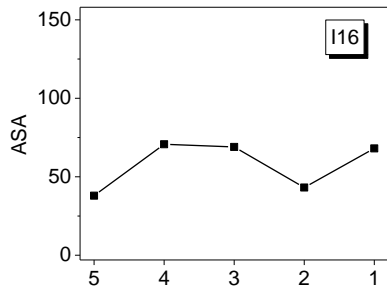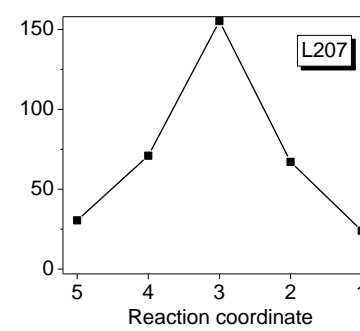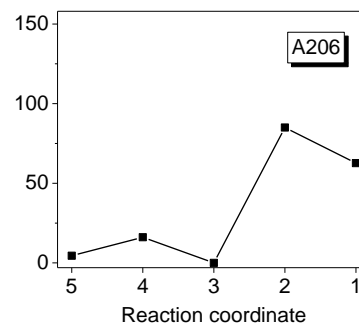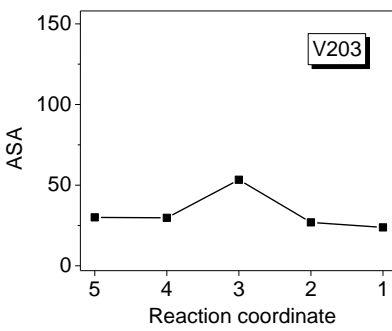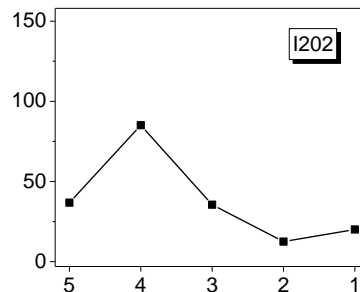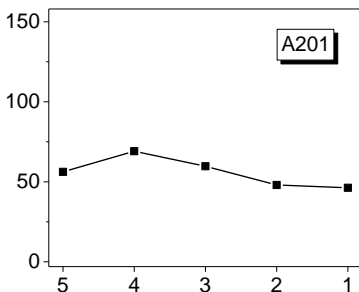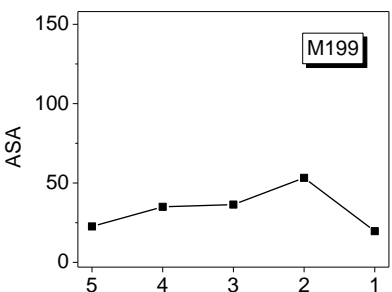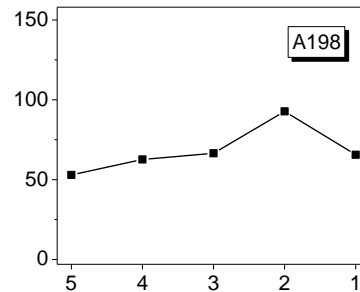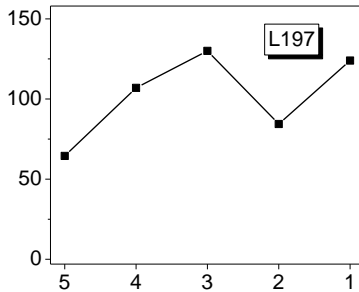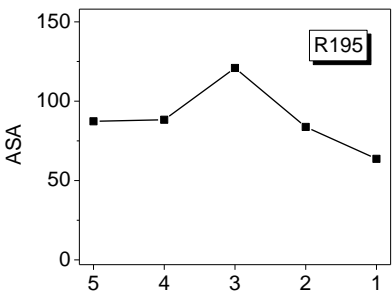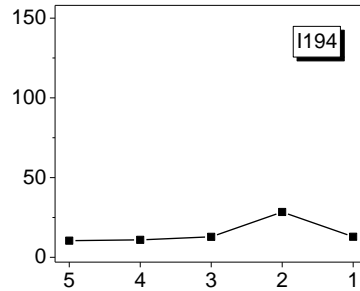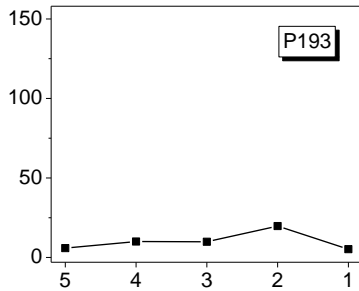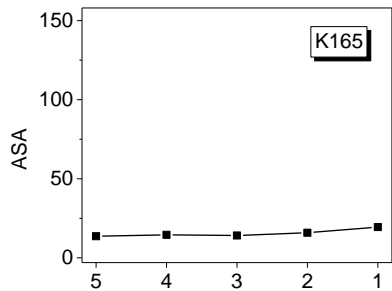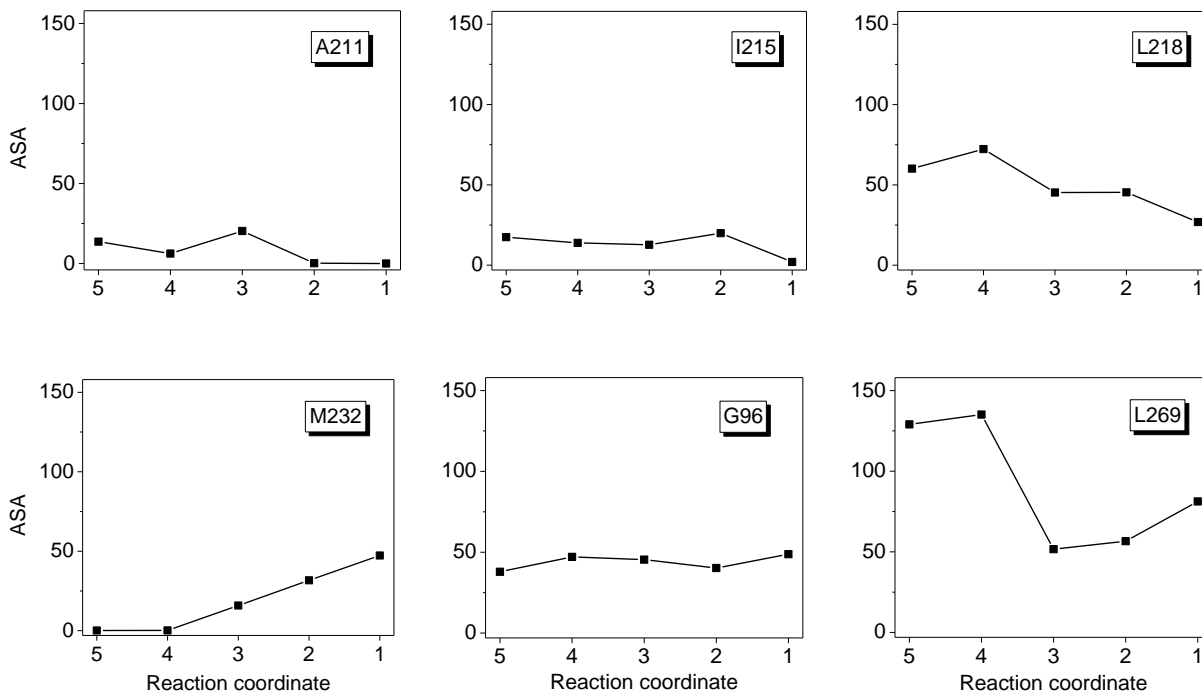
Figure 2.36 Change of accessible surface area (ASA in Å) for selected residues along the barrier-crossing coordinate

The value of ASA is obtained for protein alone as if there were no ligand or other subunits. Point 1: binary complex represented by the PDB structure 2AQ8; point 2: PT155 complex structure in $P2_12_12_1$, subunit 3; point 3: PT92 complex structure, subunit 2; point 4: PT10 ternary complex, subunit 1; point 5: PT70 complex structure, 2X23, subunit 2. For L269 (last panel) the burial by a separate subunit is also considered (**Table 2.22**) and incorporated into the plot.
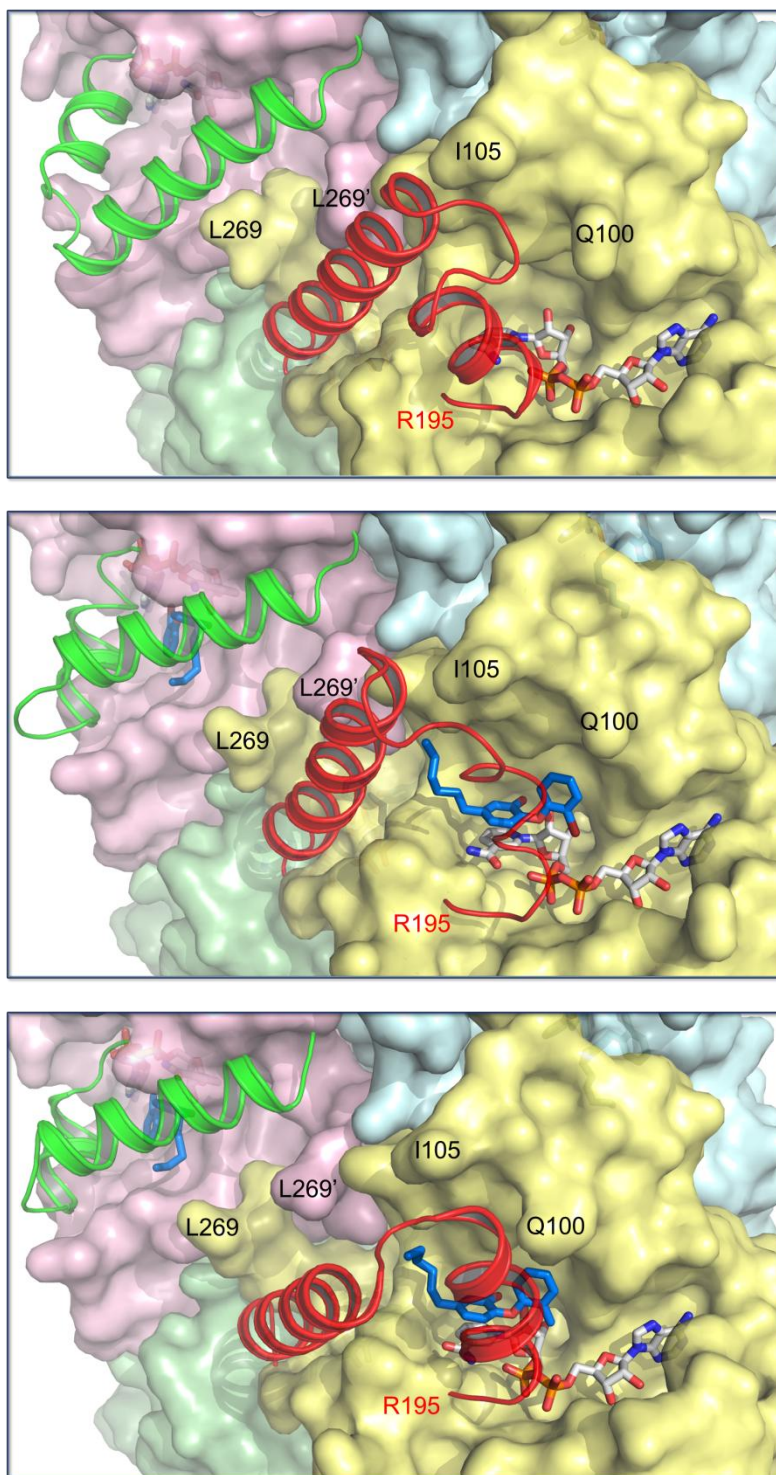
Figure 2.37 Change of cross-subunit interactions of the SBL in the InhA tetramer
Top: binary complex structure 2AQ8. Middle: ternary complex of PT92, subunit 2 in yellow. Bottom: ternary complex of PT70, subunit 2 in yellow.The tetramer is colored by subunit. The SBL of the yellow and pink subunit is in red and green, respectively.

The primary difference between the energy profile of the PT70 ternary complex and the binary complex is the altered stability on the two sides of the barrier. To see if this term, $\Delta\Delta G^{'}(\delta 1, \delta 2)$, the intrinsic binding energy, may contribute to the stabilization of the ternary complex after crossing the barrier, the amount of the surface area buried by the inhibitor is examined. Although the number is not expected to be an accurate measure of the binding energy, it is an estimation of the amount of contacts made between the protein and the inhibitor. The buried surface area of the inhibitor by the protein and NAD together is nearly 100% for all the ternary complexes along the reaction coordinate except in Ternary 0 that is buried by 94.5% (**Table 2.21**), indicating that significant interactions between the enzyme and the inhibitor have been established even before crossing the barrier. This is consistent with the great affinity already achieved by the initial complex as indicated by an initial K_i of 1–10 nM from the kinetic analysis of PT70 and PT92 inhibition (Pan Pan). However, it is not obvious from the total buried surface area of the inhibitor how the intrinsic binding energy might stabilize the ternary complex after crossing the barrier. The buried surface area of individual residues is also examined, and the plots suggest that the most important contributor of stabilization going from point 4 to point 5 is V203 (**Figure 2.38**).
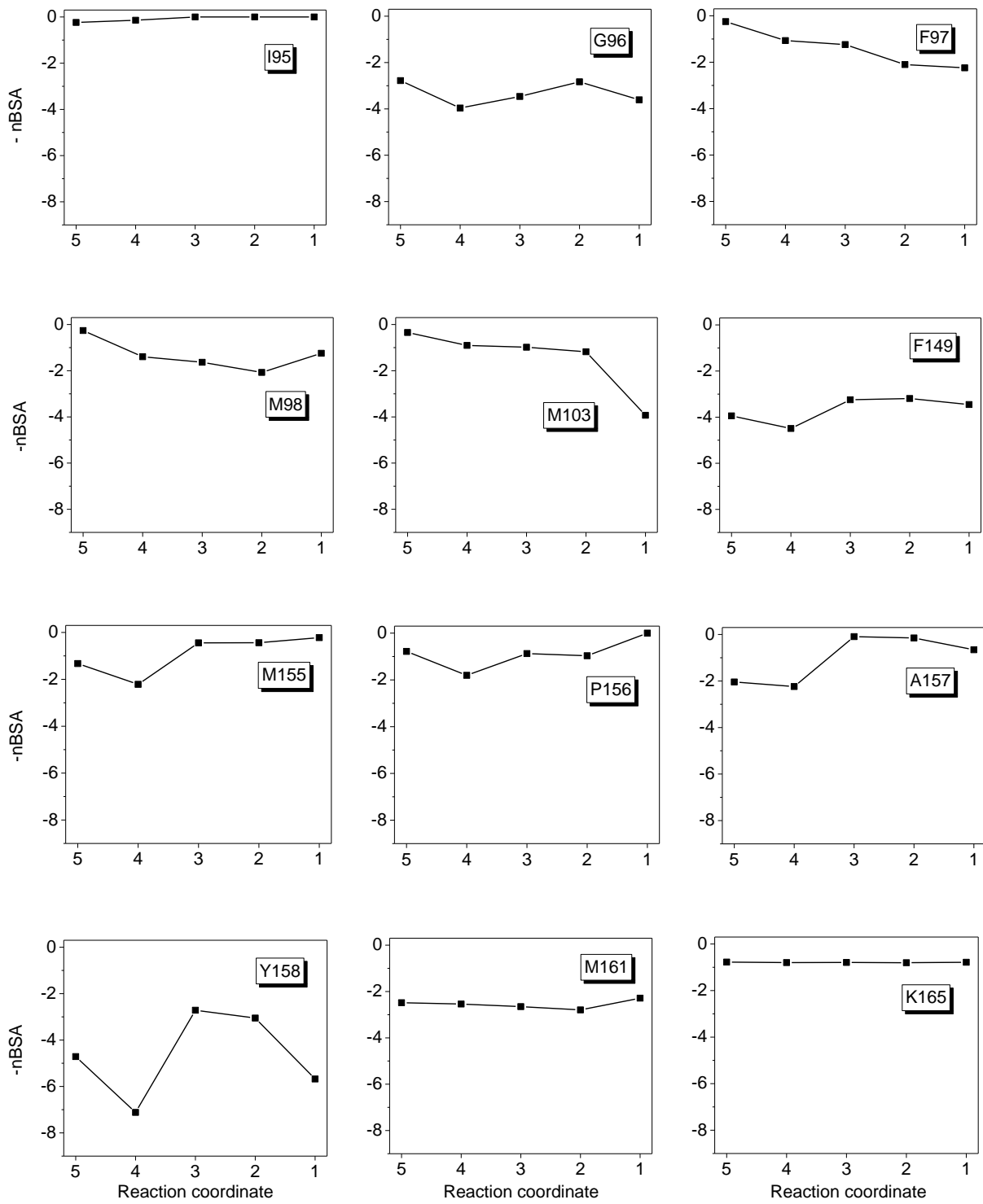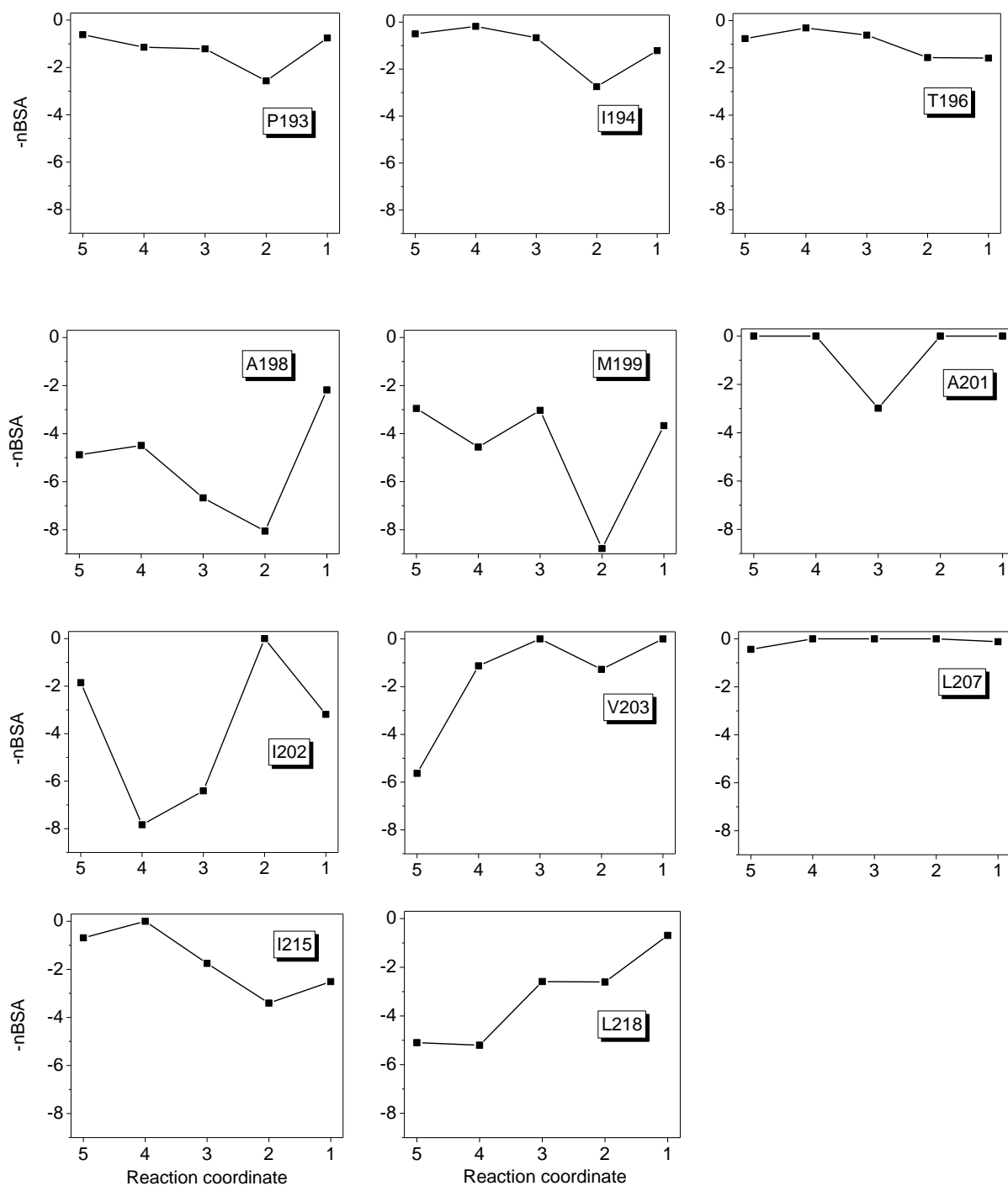
Figure 2.38   Normalized buried surface area for selected residues along the barrier-crossing coordinate

The accessible surface area of the residue buried by the inhibitor is divided by the total accessible surface area of the respective inhibitor, and presented as a negative percentage (%). Point 1: PT155 complex structure in $I2_12_12_1$, subunit 2; point 2: PT155 complex structure in $P2_12_12_1$, subunit 3; point 3: PT92 complex structure, subunit 2; point 4: PT10 ternary complex, subunit 1; point 5: PT70 complex structure, 2X23, subunit 2.

84

Residues such as F149 and M161 make good contacts with the inhibitor throughout the conformational change and there is not much change in the buried surface area (**Figure 2.38**). The interactions with these residues are similar in the intermediate and final complexes (**Figure 2.39**). As the conformational change proceeds, the inhibitor gains better contacts with V203 and L218, but loses contacts with I215. The evolution of the interactions between the enzyme and the diphenyl ether suggests that the initial conformation of the enzyme is optimal for making contacts with the A ring and the tail of the inhibitor, but not as good for binding the B ring. The conformational change of the enzyme makes it possible for the ternary complex to optimize the interactions with the diphenyl core of the inhibitor. In addition, the more "relaxed" helix-7 as a result of the conformational change could fit better with the entropic demands of the tail attached to the A ring. This means that the stabilization of the ternary complex after the barrier might be attributed to the third term, $\Delta G_I' \left(\delta 1', \delta 2' \right)$, the stability of the inhibitor conformation by itself. The structural transition to the final complex observed here fits into the picture of generalization made by Garvey and Copeland [163,175].

The conformational change brings the enzyme to an unnatural conformational state because V203 whose physiological function is not to bind the ligand is now involved in ligand binding while the opposite is true for I215. As suggested by the PNEB result, the passing of V203 and I215 with respect to each other signals the transition to this unnatural conformational state. In summary, the barrier for the binary complex to convert between the open and closed states can be composed of the energy required for disruption of the secondary structure including breaking several hydrogen bonds and the exposure or unpacking of hydrophobic side chains. The size of the barrier for the ternary complex is dictated by that of the binary complex as well as free energy of the inhibitor and contributions from interactions with the inhibitor.
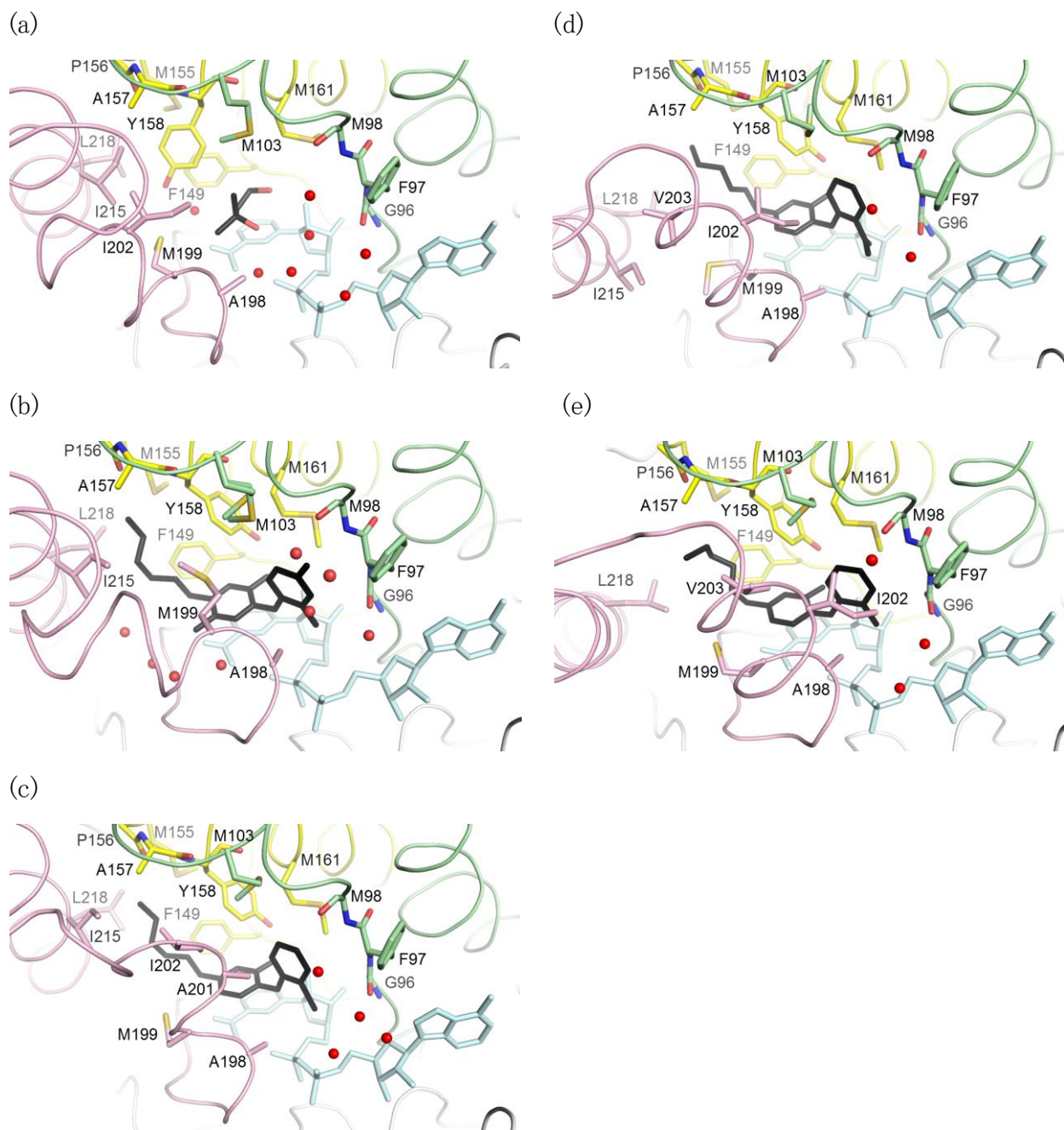
Figure 2.39 The evolution of interacting residues from the initial to the final complex
(a) PDB structure 2AQ8 with re-modeled 2-methyl-2,4-pentanediol (MPD) in the active site; (b) the initial complex represented by subunit 3 of the PT155 ternary complex structure; (c) the intermediate complex represented by subunit 2 of the PT92 ternary complex structure; (d) the intermediate complex represented by subunit 1 of the PT10 ternary complex structure; (e) the final complex represented by subunit 2 of the PT70 ternary complex structure 2X23.

Effects of V203A and I215A mutations

The side chains of V203 and I215 swap their roles in binding the inhibitor during the conformational change and are speculated to cause steric hindrance. The effect of mutating these two residues to smaller side chains is explored. Both the energy profiles of the PT70 ternary complexes with the V203A and I215A mutants display a reduced barrier size in both forward and reverse directions (**Figure 2.40**) (Eric Cheng Tsung Lai). The V203A and I215A mutants have similar catalytic parameters to those of the wild type, but lose slow-onset kinetics with PT70, and K_i increases from 22 pM for the wild type to 67 nM and 102 nM for the V203A and I215A mutants, respectively (**Table 2.23**) (Dr. Nina Liu).

V203A:NAD$^+$:PT70   I215A:NAD$^+$:PT70
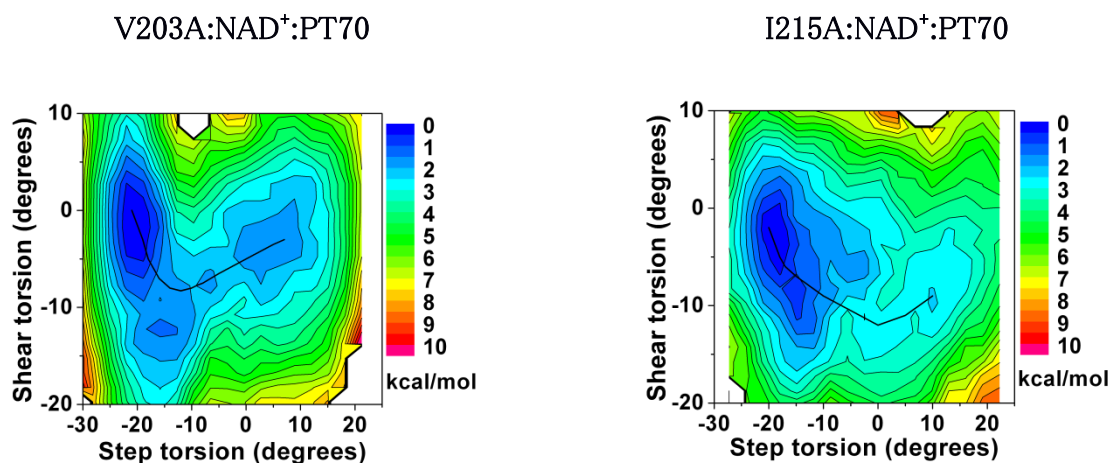
Figure 2.40 Energy profiles of the PT70 ternary complexes with V203A and I215A mutants

Table 2.23 Kinetic properties for catalysis and inhibition by PT70 for wild-type InhA and the V203A and I215A mutants

InhA	k_{cat} (min^{-1})	K_m (μM)	k_{cat}/K_m (min$^{-1}\cdot\mu$M^{-1})	Inhibition kinetics	Ki (nM)
Wild type	278±26	27±7	10±3	slow-onset	0.022±0.001
V203A	78±7	13±3	6±1	classical	67±7
I215A	109±5	13±2	8.4±1	classical	102±12

The crystal structure of the PT70 ternary complex with the V203A mutant contains four subunits in the ASU and only one has continuous density for the entire SBL (**Figure 2.41**). The position of helix-6 and helix-7 has also deviated from that observed in the wild type structure and appears in the higher δ1 conformational space in the 2D plot (**Figure 2.42**). The structure with the I215A mutant in space group $P6_222$ is also not observed in the same conformational space as wild type anymore. The I215A mutant structure in complex with PT70 in $C222_1$ has poor density for PT70. Nevertheless, Y158 adopts a rotamer that is consistent with an interaction with the phenolic group on the A ring. Hence the lack of density is most likely due to disorder rather than the absence of PT70 in the structure. Weak density for the SBL is observed in two of the six subunits in ASU. Although the subunits with better density appear in the high δ1 conformational space in the 2D plot, the subunits with weaker density display low occupancy of the SBL in both low and high δ1 conformational space. This is similar to the observation in the crystal structure of the PT03 ternary complex with the wild type enzyme. The binary complex structure of the V203A mutant appears in a more similar position on the 2D plot to the I215A mutant in complex with PT70 than to the wild type binary complex. The mutation has apparently reduced the stability of the binary complex in its native conformation.

Table 2.24 Unit cell parameters of the mutant InhA crystal structures in complex with NAD$^+$ and PT70

ID	Mutation	Space Group	a(Å)	b(Å)	c (Å)	α (°)	β (°)	γ (°)	N[1]	R (Å)[2]
249B3	V203A	C2	126.10	91.90	102.85	90.00	106.45	90.00	4	2.3
236G8	V203A[1]	$P6_222$	97.34	97.34	260.39	90.00	90.00	120.00	2	2.9
239G7	I215A	$P6_222$	99.79	99.79	139.79	90.00	90.00	120.00	1	2.5
256B1	I215A	$C222_1$	82.32	100.23	379.10	90.00	90.00	90.00	6	2.0

1. Binary complex
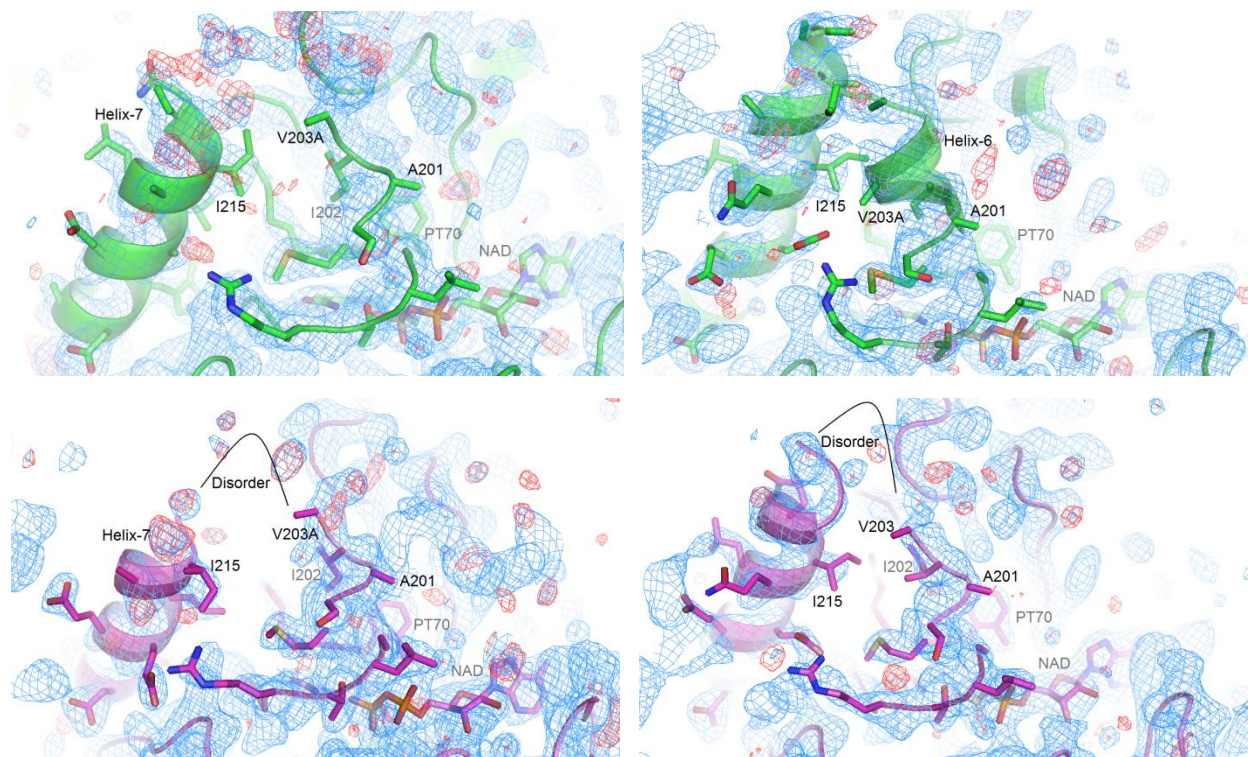
Figure 2.41 Density for the SBL in the ternary complex structure of the V203A mutant bound to PT70
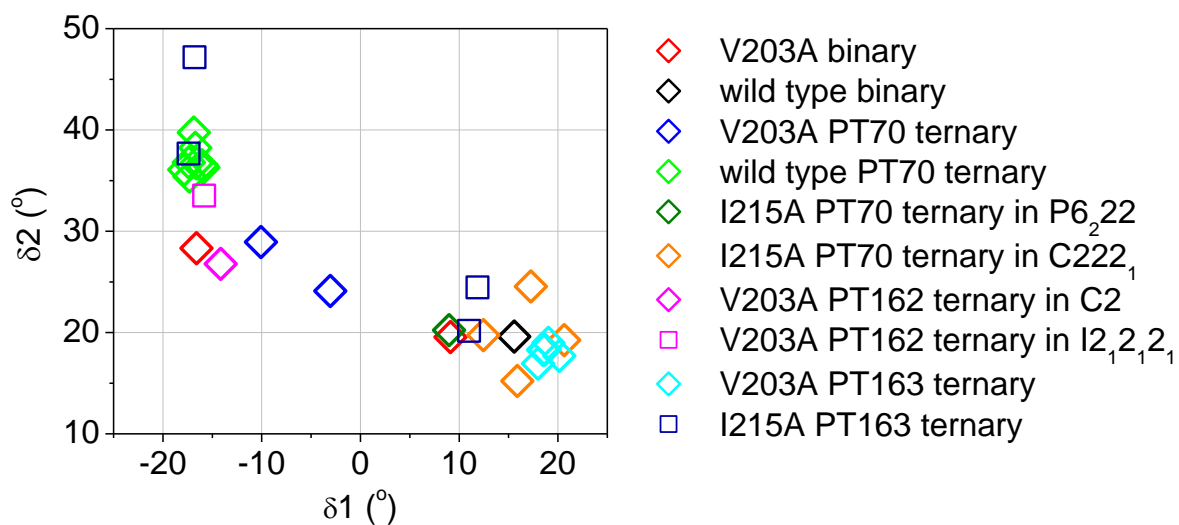
Figure 2.42 Crystal structures of the V203A and I215A mutants in the 2D plot

PT162 and PT163 have a bulkier side chain attached to the A ring than that of PT70 and are expected to impact the interactions with residues 203 and 215. The energy profiles of PT162 ternary complex with the V203A mutant and the PT163 ternary complex with the I215A mutant show the recovery of the size of the barrier compared to the respective PT70 complexes (**Figure 2.43**) (Eric Cheng Tsung Lai). Slow-onset inhibition is also observed for the V203A mutant by PT162 and for the I215A mutant by PT163 (**Table 2.25**) (Pan Pan).

V203A:NAD:PT162   I215A:NAD:PT163
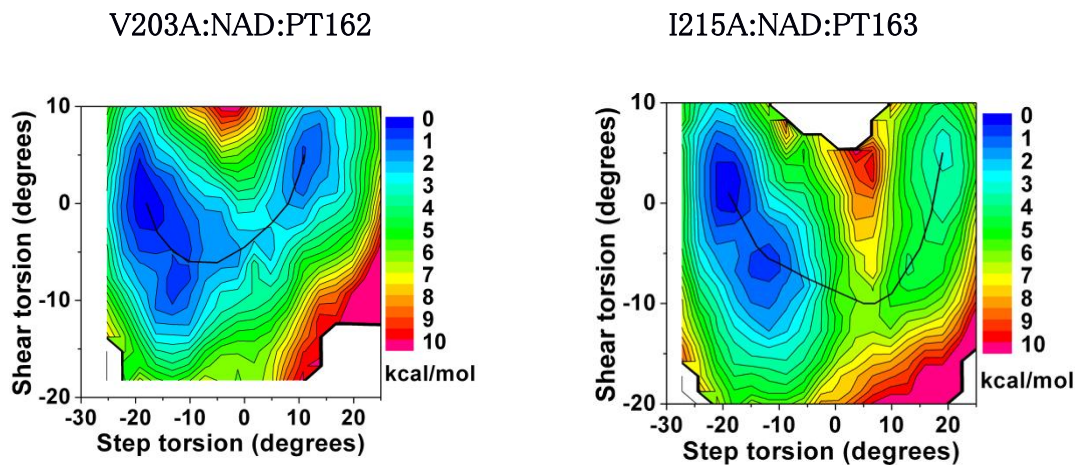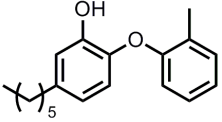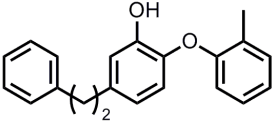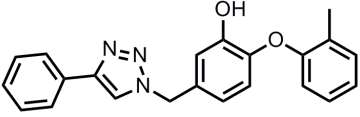
Figure 2.43 Energy profiles of the ternary complexes of PT162 bound to the V203A mutant and PT163 bound to the I215A mutant

Table 2.25 Kinetic properties of inhibition of the V203A and I215A mutants by PT70, PT162 and PT163

Compound		wt InhA		V203A		I215A	
		IC$_{50}$ (nM)[†]	T$_R$ (min)	IC$_{50}$ (nM)[†]	T$_R$ (min)	IC$_{50}$ (nM)[‡]	T$_R$ (min)
	PT070	50.7±4	30.3	n.d.	R	n.d.	R
	PT162	360±78	R	140±28	13.6	n.d.	R
	PT163	400±48	R	n.d.	n.d.	151.9±10.9	70

[†][E] = 100 nM; [‡][E] = 40 nM; T$_R$: residence time; R: not slow-onset

Table 2.26 Unit cell parameters of the V203A InhA crystal structures in complex with NAD$^+$ and inhibitors

ID	Inhibitor	Space Group	a(Å)	b(Å)	c(Å)	α (°)	β (°)	γ (°)	N[1]	R(Å)[2]
261A1	PT162	C2	118.28	67.90	77.58	90.00	122.70	90.00	2	1.8
257D5	PT162	I2$_1$2$_1$2$_1$	90.65	97.90	184.70	90.00	90.00	90.00	2	3.6
280F11	PT163	C222$_1$	82.00	100.33	376.93	90.00	90.00	120.00	6	2.0
281G11	PT163[1]	P2$_1$2$_1$2$_1$	92.01	97.56	184.71	90.00	90.00	90.00	4	3.4

1. I215A mutant

The PT162 ternary complex with the V203A mutant in space group C2 contains two subunits in the ASU (Table 2.26). The SBL density from residue 204 to 214 in subunit 2 is poor. The PT162 ternary complex with the V203A mutant in space group I2$_1$2$_1$2$_1$ contains two subunits in the ASU. The structure is low−resolution but the C$_\alpha$ positions of the SBL can be

traced. However, the inhibitor is only found in subunit 1. The positions of both structures shift back toward the low $\delta 1$ conformational space in the 2D plot (**Figure 2.42**), consistent with the recovery of kinetic properties and the energy profiles. The PT163 ternary complex structure with the I215A mutant contains four subunits in the ASU. The C_α positions of the SBL are traceable despite the low resolution. Two of the four subunits show the recovery of the SBL to low $\delta 1$ conformations.

The PT163 ternary complex structure with the V203A mutant contains 6 subunits in ASU (**Table 2.26**). This is the same crystal form as for the PT70 ternary complex with the I215A mutant, 256B1 (**Table 2.24**). Four out of the six subunits contain clear density for the diphenyl core. The density for the SBL is clear except in subunit 5 where poor density is observed for residues 203 to 218, showing low occupancy of the SBL in the low $\delta 1$ conformational space. Other subunits contain a SBL in the high $\delta 1$ region.

The altered energy profiles of the PT70 ternary complex with mutant enzymes (**Figure 2.40**) may be understood using the energy components discussed earlier. Assuming that the mutations change only the two terms, $\Delta G'_{E \cdot NAD^+}(\delta 1, \delta 2)$ and $\Delta\Delta G'(\delta 1, \delta 2)$, the smaller side chains reduce the van der Waals contacts and the possible steric hindrance in the energy barrier of the wild type enzyme. Since I215 contributes more surface area than V203 for the folded structure in the binary complex (**Figure 2.36**), the I215A mutation is expected to have a larger impact on the stability of the binary complex than V203A, leading to greater reduction of the barrier size in the forward direction in the energy profile. In addition, I215 contributes to binding earlier and V203 later in the reaction coordinate (**Figure 2.38**), and so the V203A mutation destabilizes the final complex more while I215A destabilizes the initial complex more. These factors lead to the reduction of the barrier size in both cases but some difference in the relative stability of the initial and final complexes. The change in $\Delta\Delta G'(\delta 1, \delta 2)$ is in agreement with the increase in K_i.

The recovered barrier from the mutant energy profiles with PT162 and PT163 may be attributed to the change in $\Delta\Delta G'(\delta 1, \delta 2)$. Compared with the corresponding PT70 ternary complex, the energy profile of the PT162 complex recovers the barrier primarily on the reverse side (**Figure 2.43**), which may be rationalized by the recovery of the van der Waals contacts of residue 203 with the bulkier tail. PT163 has an even bulkier side chain which can increase the intrinsic binding energy for both the final and initial complexes compared to the

PT70 counterpart, but at least not by the same amount in the high-energy intermediates because there is too little space for the bulky group in these conformers and steric hindrance causes a decrease in binding energy.

2.4 Discussion

Correlation of SBL ordering and slow-onset inhibition

Generalization is difficult due the artifact of crystal packing and the ambiguity sometimes encountered in the reaction progress curve analysis. Although ordering of helix-6 in the ternary complex structure of PT70 can be attributed to stabilization by the inhibitor rather than by crystal packing [159], it is not always straightforward to make a conclusion from the crystal structural data how much ordering of the SBL is induced by the inhibitor. The SBL is a region longer than 20 amino acids and the local quality of density is very dependent on the crystal packing environment, which varies in different crystal forms and for different subunits in the same crystal form. The ordering is also more of a matter of degree, and not that of all-or-nothing. However, the crystal structural data is usually only presented in the form of coordinates and the information about mobility is obscured. The refinement parameter that is most relevant to the degree of ordering is the B factor. When an atom does not stay consistently in the same position throughout the crystal lattice, its density in the map becomes weaker and its position cannot be determined with high certainty, which is reflected by a higher B factor. The high B factors obtained in protein crystal structural data do not distinguish thermal fluctuations from "real" collective motions. When the density becomes too weak that it is impossible to model the position of the atom, its coordinates are left out and it is described as being disordered. Hence the disorder for a certain region in the crystal structure simply means the lack of density, and not the lack of secondary structure (i.e. unstructured) as in the language of protein folding. The disordered region might still have a secondary structure that changes positions over the crystal lattice (static disorder) or over time (dynamic disorder) [221].

The SBL in InhA is a surface region vulnerable to disorder in the crystal structure, and the flexibility of this region is most likely functionally relevant. Even smaller-amplitude motions can lead to diminished density. The disorder seen in the binary complex structure, 21B3, and that in the PT03 ternary complex structure, 2B36, are not of the same nature. The

93

residual density in these data suggests that the disorder in the binary complex is due to smaller-amplitude motions, e.g. conformational transitions of the SBL between ($\delta 1$, $\delta 1$) values of (11, 19) and (15, 22). In contrast, the disorder in the PT03 ternary complex is due to larger-amplitude motions, e.g. conformational transitions between (11, 19) and (-17, 35).

Similarly, the attempt to make correlations with slow-onset kinetics may be complicated by the ambiguity in drawing a conclusion from the linearity in the reaction progress curves. In the case of the two-step, slow-onset mechanism, the curvature in the progress curve is observed only when k_{-2} is smaller than k_2, and both are of sufficiently small magnitude, presumably $<< 1$ s^{-1}. It is likely that the curvature is not detectable when k_2 and k_{-2} become comparable in magnitude, or when k_2 and k_{-2} increase so much that the steady-state is reached within a few seconds.

The good level of density of the SBL observed in the crystal structure of INH-NAD adduct complex cannot rule out the contribution from crystal packing stabilization. The 21 crystal structures in PDB in the same crystal form (**Table 2.28**) all show a similar B factor profile with an increase around residues 200 to 210. So this region will likely be more mobile in solution. In contrast, the good level of density of the SBL observed in the crystal structure of the C16-NAC complex may be attributed to the extensive contacts made between the acyl chain of C16-NAC and helix-6. Even so, the B factor profile still displays an increase around residues 195 to 205.

Despite these complications, it is clear that helix-6 is intrinsically more mobile than the rest of the protein, and slow-onset diphenyl ethers indeed reduce the mobility of helix-6 because of the well-established interactions between the diphenyl core and residues on helix-6, as well as the enhanced interactions between helix-6 and strand-4. However, this is achieved at the expense of the interactions between helix-7 and strand-5 and the C-terminus of the neighboring subunit, leading to the increase of mobility of helix-7 (**Figure 2.10**). Hence the observed ordering of helix-6 by PT70 is the consequence of the conformational change induced by the inhibitor.

InhA conformational states and enzyme function

It has become apparent that helix-6 and helix-7 most likely possess defined although variable secondary structures even when their densities are not well observed in the crystal structure. This conformational variability leads to separate conformational states of InhA as revealed in **Figure 2.24.** The clustering of these different conformers observed experimentally suggests that there exists at least one major barrier on the InhA energy landscape, and we have identified one such barrier from computation (**Figure 2.30**). In other words, we have identified two major conformational substates of InhA, termed helix-6-open and helix-6-closed, corresponding to high and low $\delta 1$ values, respectively. We know that only the former will be competent in binding the physiological substrate, the long fatty acyl chain attached to the acyl carrier protein. This is analogous to the three conformational substates (CS) observed in the tier 0 energy landscape of myoglobin, each of which binds CO with different rates [232,233]. The size of the energy barriers between these CS is on the scale of several kT, and protein fluctuations between these CS result in tier-0 dynamics on the timescale of μs and slower at physiological temperature [224]. Given the timescale of slow-onset inhibition, the protein motions relevant to the slow step should be categorized into tier-0 dynamics or even beyond.

While it appears in the case of InhA inhibition by diphenyl ethers that inhibitor binding stabilizes an InhA conformer that is incompetent in binding its natural substrate, it has been similarly proposed that conversions to the inactive form of epidermal growth factor receptor (EGFR) and infrequent-sampled conformation of p38 MAP kinase are responsible for slow kinetics of inhibition [182,183]. Slow-onset inhibition by clofarabine diphosphate (CIFDP) changes the quaternary structure of human ribonucleotide reductase (hRNR) into an inactive form [234]. In contrast, it is suggested that the final closed conformation of PDF induced by slow inhibition by actinonin is the catalytically competent form given that actinonin is a transition state analogue [235].

The characterization of these different CS of InhA from a large number of crystal structures which can subsequently be linked to the energetic information from computation is facilitated by the reduction of the multiple dimensions of structural data to lower dimensions, in our case, two. The atomic coordinates in the crystal structures are compressed into two chosen coordinates that can reveal the conformational change in question. This approach has been used in a number of studies, including the inhibition of acetylcholinesterase (AChE) by

the neurotoxin, fasciculin-2 (FAS2) [222], the domain closure of adenylate kinase (AdK) [236], and the binding pathway of slow-onset actinonin along the conformational states of peptide deformylase (PDF) [235].

Structural transitions on the slow timescale

While molecular motions on the fast timescale correspond to local flexibility, e.g. bond vibration and amino acids side chain rotation (**Figure 2.6**), it is expected that those on the slow timescale are larger-amplitude collective motions [224]. However, it is unclear what structural factors distinguish some collective motions faster than ms from those slower than seconds. The generalization of the large-amplitude collective motions still seems inconclusive since many studies attribute slow kinetics to structural changes such as small adjustments of protein conformations, water structure at the active site, or changes on the inhibitors such as covalent bond formation to the protein, hydration/dehydration of a carbonyl, ionization, or inhibitor conformational change or movement [163,165,237]. Some are attributed to changes in the oligomeric state, or many small local changes on a larger scale [238]. Protein folding studies suggest that prolyl peptidyl cis-trans isomerization, hydrophobic collapse or side chain contact formation may occur on the timescale of seconds and longer [225,239]. Prolyl isomerization has to overcome a large energy barrier of 20-22 kcal/mol, exhibiting half times of 10-100 s at 25°C because of rotation about a partial double bond [240].

We have constructed a pathway leading InhA from its native predominate conformation to a less favorable one stabilized by the slow-onset diphenyl ethers. Representative conformers along the pathway are available including the higher-energy intermediate, allowing us to examine the change in energy in more molecular details. Some structural analyses have suggested disruption of secondary structure and steric hindrance of side chains contribute to the energy barrier. Further computation will be required to distinguish the contribution of van der Waals interactions from others such as polar interactions, internal energy and entropy. By dissection of the energy of the ternary complex into three components — protein, inhibitor, and their interactions, it is also possible to analyze the contributions to the barrier in a more systematic and quantitative manner, and establish the structure-activity relationship between inhibitor structure and inhibition kinetics that will provide valuable information for medicinal chemists.

Relationship between inhibitor structure and inhibition kinetics

Although more analysis is required to elucidate the molecular basis of the energy barrier, our general understanding of the conformational change has made it possible to do more rational design of inhibition kinetics. The data indicates that all diphenyl ethers with moderate size of substituents to the core rings can shift the equilibrium of the two major conformational states toward the helix-6-closed form to varied extent. Given the structures of the initial complex, high-energy intermediate and final complex, we can assess the relative stabilities of the ternary complexes of the designed inhibitor and predict whether slow-onset inhibition will be observed. Slower kinetics may be achieved by destabilization of the high-energy intermediate relative to the final complex or even the initial complex. Compared to traditional rational design, we take into consideration not only the final conformational state but also a few other intermediate states. This more dynamic view is important for more accurate prediction of the kinetics and affinity of inhibition.

The increase of energy barrier for the I215A mutant by replacing PT70 with PT163 supports that there is limited binding space for the tail of the diphenyl ether in the high-energy intermediate, which can be destabilized by increasing the volume of the tail. Reducing the volume of the tail to C5 as in PT03 appears to stabilize the high-energy intermediate although both the final and initial complexes might also be destabilized by removal of the B ring substituent which reduces van der Waals contacts and increases the entropic penalty of binding (**Figure 2.30**). The final complex can also be destabilized when the B ring substituent is too big as in PT115. However, it is unclear how the high-energy intermediate is destabilized to a greater extent that leads to a slower off rate than the PT70 complex.

Induced-fit and conformation-selection

There are often debates on whether the conformational change associated with ligand binding should proceed through the induced-fit or the conformation-selection mechanism [223,241-246]. The experimental support for either mechanism relies on the observation of the intermediates, either the ligand bound complex in the conformational state of the unbound receptor or the pre-existing receptor conformer in the absence of the ligand. Both intermediates can be identified in our 2D energy profiles of binary and ternary complexes, and

so it is a matter of relative stabilities of these two intermediates that determines which one might be detected experimentally. Our structural model either from crystal structure or from simulation indicates that there is sufficient stability of the initial complex that rationalizes the observation of the induced-fit mechanism from kinetic experiments. However, it is certainly possible that in the microsteps, conformation-selection could come into play (**Figure 2.32**). The energy profiles provide strong validation of a structural model for conformational change through the comparison with kinetic observation.

Applicability to other systems

Two other classes of slow-onset InhA inhibitors are known, the INH-NAD adduct, and the inorganic complex, pentacyano(isoniazid)ferrate(II) (PIF) [247]. The analogue of INH-NAD adduct, ethionamide-NAD (ETH-NAD), is bound to two conformers in the crystal structure (PDB 2NTJ) [248] similar to the initial and final complexes of diphenyl ethers (**Figure 2.19**). It is very likely that these inhibitors all induce a similar conformational change in the slow step.

The methods we have developed in this study are generally applicable to other FabIs. Although the SBL of InhA is longer than that in most other homologues, some critical features are conserved including hydrophobic residues in the positions analogous toV203 and I215 in the final complex structure, and residues analogous to R195 and E219 that can potentially form a salt bridge in the initial complex. For ecFabI, an ACP-bound model with an open helix-6 could represent the initial complex structure (2FHS). The final complex structure bound with triclosan (1QSG) displays a broken salt bridge similar to the final complex of InhA, and an intermediate structure represented by the complex of diazaborine inhibitor (1DFH) exhibits a weakened salt bridge. However it is still inconclusive how much the energy involved in this salt bridge contributes to the overall energy barrier. The carboxylate and guanidinium side chains of these two residues are often mobile in the crystal structures, and tend to form artificial salt bridges with symmetry-related subunits.

The function-related slow-onset inhibition mechanism in which the native conformational state of the enzyme forms the initial complex with the inhibitor which then isomerizes to an outlying conformational state in tier-0 or beyond is an appealing mechanism that might lead to a general strategy for *de novo* design of slow-onset kinetics. Provided the substrate-bound

conformation of one enzyme is known, a flexible substrate binding element can be modeled [249–251] to generate a distinct conformer slightly less stable, which can then be stabilized by the designed inhibitor fit into the newly-shaped binding pocket. The energy profiles can then be calculated to predict whether the conformational change can be achieved when the inhibitor is bound.

2.5 Conclusion

Through the observations in several ternary complex crystal structures of diphenyl ethers, it is found that the SBL conformation is sensitive to subtle changes in inhibitor structure and protein mutation, ranging from the helix-6-open form competent in binding the natural substrate to the helix-6-closed form favored by slow-onset diphenyl ethers. The local quality of density depends on both the interactions with the inhibitor and crystal packing. Slow-onset diphenyl ethers achieve optimization of binding through interactions primarily with helix-6 and stabilize the helix-6-closed conformational state higher in energy than the helix-6-open state in the absence of the inhibitor. The energy barrier between the two conformational states is revealed by 2D energy profiles from simulations. The size of the barrier and relative stabilities of the two conformational states rationalize observations in kinetic experiments and crystal structures, and the relevance of this conformational change in the slow step of inhibition is further substantiated.

The distribution of crystal structures along the reaction coordinate of the conformational change enables construction of a pathway composed of initial and final complexes as well as higher-energy intermediates. Structural analyses of these complexes suggest that disruption of the native secondary structure of SBL and steric hindrance caused by side chains may contribute to the energy barrier, and a systematic approach is proposed to further explore the molecular basis of the source of the barrier in terms of energy components and establish the relationship between inhibitor structure and inhibition kinetics. The methods developed in this study may find general applicability to other slow-onset inhibition systems. Knowledge of the reaction pathway and a linkage between structure, function and energy are necessary to understand kinetic observations. The local refolding process will also appeal to people studying the mechanism and rate of protein folding.

Appendix
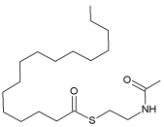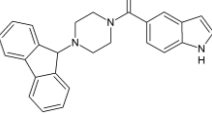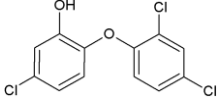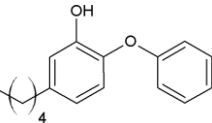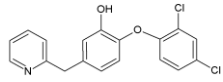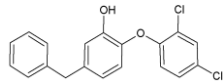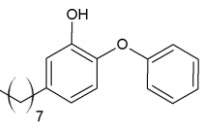
Secondary structure of PT115 ternary complex:
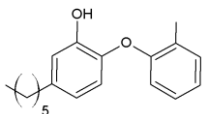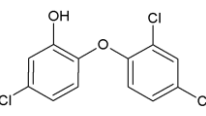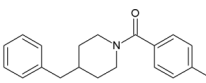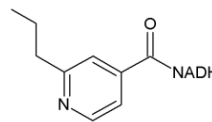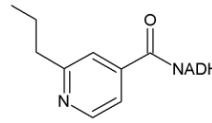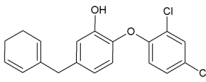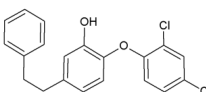
Table 2.27   Secondary structure parameters for SBL in the crystal structure of PT115 ternary complex 88B3
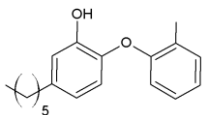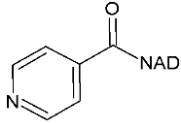
Complex	δ 1	Start	End	Type	No. residues	Residues per turn	Deviation from ideal (°)
E·NAD$^+$·I	-4	L197	V205	H	9	3.92	20.7
		E210	R225	H	16	3.77	15.9
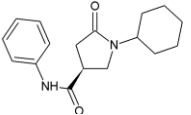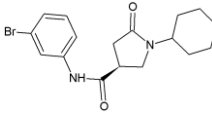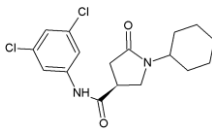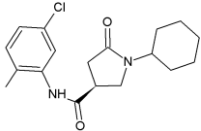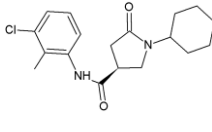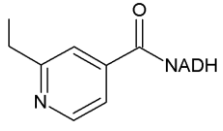
InhA crystal structures in PDB:
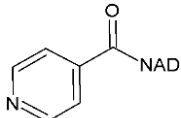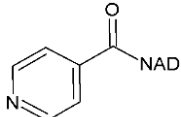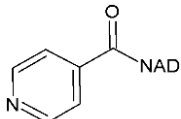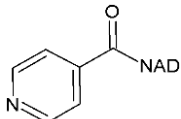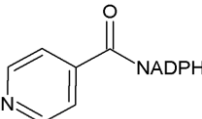
Table 2.28 Crystal structures of InhA in PDB

PDB ID		Ligands		Space Group	a (Å)	b (Å)	c (Å)	α (°)	β (°)	γ (°)	Resolution (Å)
1BVR	wt		NAD	C 2	101.08	83.45	192.81	90.0	95.24	90.0	2.8
1P44	wt		NAD	C 2	100.62	83.22	192.97	90.0	94.95	90.0	2.7
2B35	wt		NAD	C 2	99.96	81.83	188.66	90.0	95.69	90.0	2.3
2B36	wt		NAD	C 2	99.96	81.83	188.66	90.0	95.69	90.0	2.8
3FNE	wt		NAD	C 2	125.59	92.34	103.02	90.0	106.4	90.0	1.98
3FNF	wt		NAD	C 2	125.65	92.38	102.41	90.0	106.51	90.0	2.3
2B37	wt		NAD	C 2 2 2$_1$	82.34	100.29	379.27	90.0	90.0	90.0	2.6

PDB ID	Variant	Structure	Ligand	Space Group	a	b	c	α	β	γ	Resolution
2X22	wt		NAD	C 2 2 2_1	89.81	157.51	91.23	90.0	90.0	90.0	2.1
1P45	wt		NAD	I 2_1 2_1 2_1	94.82	104.12	189.75	90.0	90.0	90.0	2.6
2NSD	wt		NAD	I 2_1 2_1 2_1	91.26	91.33	184.26	90.0	90.0	90.0	1.9
2NTJ	D2A			I 2_1 2_1 2_1	91.59	100.52	186.5	90.0	90.0	90.0	2.5
2NTV	M. leprae			I 2_1 2_1 2_1	91.23	100.02	186.62	90.0	90.0	90.0	2.1
3FNG	wt		NAD	I 4_1 2 2	90.05	90.05	183.89	90.0	90.0	90.0	1.97
3FNH	wt		NAD	I 4_1 2 2	90.02	90.02	183.14	90.0	90.0	90.0	2.8
2IED	S94A	--	--	P 1	54.69	63.52	65.18	97.21	85.81	102.87	2.14

102

2X23	wt		NAD	P 2$_1$	88.48	90.27	89.56	90.0	118.76	90.0	1.81
1ENY	wt	--	NAD	P 6$_2$ 2 2	100.14	100.14	140.45	90.0	90.0	120.0	2.2
1ENZ	S94A	--	NAD	P 6$_2$ 2 2	100.14	100.14	139.95	90.0	90.0	120.0	2.7
1ZID	T2A			P 6$_2$ 2 2	100.53	100.53	138.96	90.0	90.0	120.0	2.7
2AQ8	wt	--	NAD	P 6$_2$ 2 2	96.02	96.02	138.16	90.0	90.0	120.0	1.92
2AQH	I21V	--	NAD	P 6$_2$ 2 2	95.95	95.95	138.24	90.0	90.0	120.0	2.01
2AQI	I47T	--	NAD	P 6$_2$ 2 2	95.44	95.44	138.71	90.0	90.0	120.0	2.2
2AQK	S94A	--	NAD	P 6$_2$ 2 2	96.25	96.25	138.05	90.0	90.0	120.0	2.3

3OEW	wt	--	NADH	P 6$_2$ 2 2	98.55	98.55	139.71	90.0	90.0	120.0	2.2
3OEY	T266E	--	NADH	P 6$_2$ 2 2	98.15	98.15	139.44	90.0	90.0	120.0	2.0
3OF2	T266D	--	NADH	P 6$_2$ 2 2	98.37	98.37	139.89	90.0	90.0	120.0	2.0
2H7I	wt		NAD	P 6$_2$ 2 2	97.54	97.54	140.52	90.0	90.0	120.0	1.62
2H7L	wt		NAD	P 6$_2$ 2 2	97.55	97.55	140.35	90.0	90.0	120.0	1.73
2H7M	wt		NAD	P 6$_2$ 2 2	97.79	97.79	140.66	90.0	90.0	120.0	1.62
2H7N	wt		NAD	P 6$_2$ 2 2	97.06	97.06	140.42	90.0	90.0	120.0	1.9
2H7P	wt		NAD	P 6$_2$ 2 2	97.78	97.78	141.12	90.0	90.0	120.0	1.86
2H9I	wt			P 6$_2$ 2 2	97.93	97.93	140.42	90.0	90.0	120.0	2.2

104

2IDZ	wt		P 6$_2$ 2 2	97.09	97.09	136.94	90.0	90.0	120.0	2.0
2IE0	I21V		P 6$_2$ 2 2	96.59	96.59	136.25	90.0	90.0	120.0	2.2
2IEB	S94A		P 6$_2$ 2 2	96.83	96.83	136.42	90.0	90.0	120.0	2.2
2NV6	D2A, S94A		P 6$_2$ 2 2	98.28	98.28	139.24	90.0	90.0	120.0	1.9
2PR2	wt		P 6$_2$ 2 2	99.13	99.13	139.27	90.0	90.0	120.0	2.5

Chapter 3 * Mechanism of catalysis by 1,4-dihydroxy-2-naphthoyl-CoA synthase

* This chapter is based on the work published in *Biochemistry 50*, 9532-9544 (2011).

3.1 Introduction

MenB catalyzed reaction

A key step in the biosynthesis of menaquinone is the reaction catalyzed by the 1,4-dihydroxynaphthoyl-CoA synthase (MenB) in which the second naphthoquinone aromatic ring is formed through an intramolecular Claisen (or Dieckmann) condensation involving the succinyl side chain of O-succinylbenzoate (OSB) (**Figure 1.6**) [109,252,253]. The previously proposed MenB reaction mechanism can be dissected into a sequence of four steps: (1) Enolate formation by proton abstraction, (2) C-C bond formation by carbanion attack, (3) ketone formation by elimination of water, and (4) aromatization by keto-enol tautomerism (**Figure 3.1**) [112].
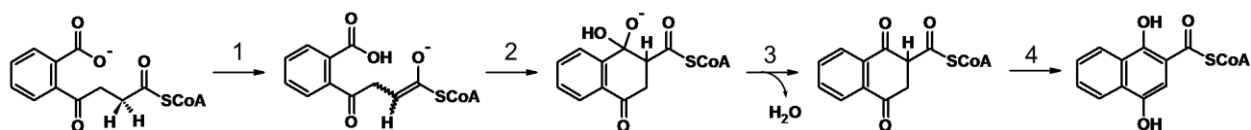
Figure 3.1 Dieckmann condensation catalyzed by MenB

The discovery of MenB occurred when cell-free extract preparations of *Mycobacterium phlei*, *E. coli*, *Bacillus subtilis* and *Micrococcus luteus* were found to convert OSB-CoA to DHNA (1,4-dihydroxy-2-naphthoate synthase activity) [254-257]. It had been unclear whether the hydrolysis of DHNA-CoA to form DHNA was carried out by MenB or simply uncatalyzed [96]. A separate thioesterase was proposed to fulfill this function in *E. coli* [258]. Later, characterization of the purified MenB from *Mycobacterium tuberculosis* verified that CoA was retained in the MenB reaction product [112], and the DHNA-CoA thioesterase in the phylloquinone biosynthesis of cyanobacteria was also identified and characterized [96], supporting that the actual MenB reaction is the conversion of OSB-CoA to DHNA-CoA.

106

The MenB-catalyzed ring closure reaction is analogous to a Dieckmann reaction which requires activation of the two carboxyl groups [259]. However, the enzymatic conversion from OSB to DHNA proceeds via activation of only the nucleophilic carboxyl group, which is considered to be unusual [260-262]. In Claisen condensations, such as that catalyzed by the β-ketoacyl-ACP synthases, normally both nucleophile and electrophile are activated through the formation of thioesters [263,264]. Although it was once thought that the aromatic carboxylate was activated when the MenE product was first characterized [257,260,261,265], the aliphatic rather than aromatic thioester was later shown to be the competent MenB substrate, and the discrepancy was attributed to the ability of MenE to produce both the aliphatic and aromatic thioesters and the relative instability of the former [266,267].

MenB is a member of the crotonase superfamily

MenB is a member of the crotonase superfamily that uses a common fold to catalyze a diverse set of reactions (**Figure 3.2**, **Figure 3.3**). Enoyl-CoA hydratase (crotonase) is the first member in the superfamily for which the crystal structure was determined [268]. The crotonase fold contains repeated ββα units that assemble into two approximately perpendicular β-sheets surrounded by α-helices [269]. Members of the crotonase superfamily (CS) share low sequence similarity and catalyze reactions including hydration, isomerization, decarboxylation, dehalogenation, C-C bond formation and cleavage, oxidation, hydrolysis, and fragmentation of β-diketones. The substrates of the majority of these enzymes are CoA thioesters (**Figure 3.4**). A common theme in the catalysis of these reactions is the stabilization of a CoA thioester enolate oxyanion intermediate by an oxyanion hole formed by two protein backbone NH groups in the active site [264,270-273].
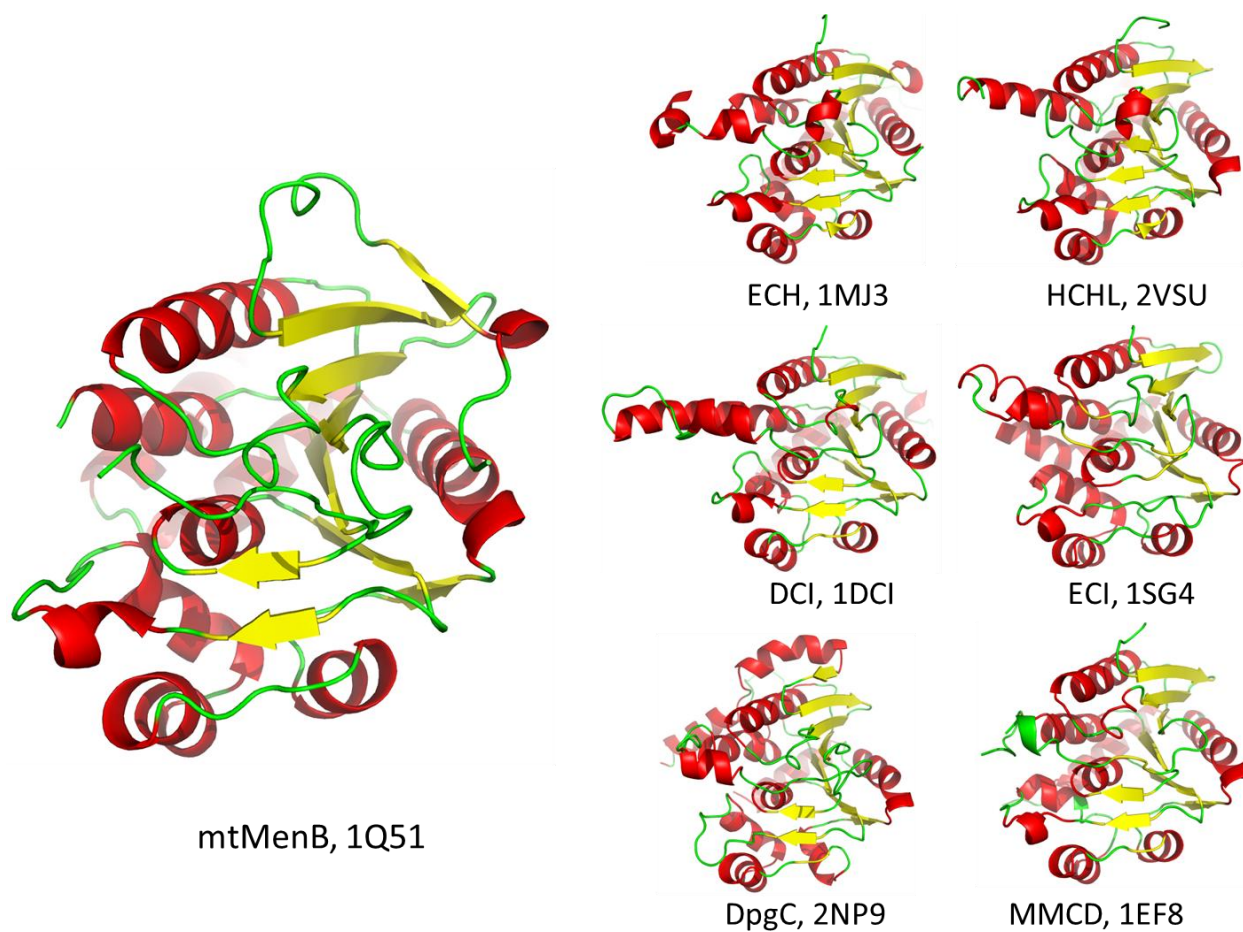
Figure 3.2 The common fold found in the crotonase superfamily
The crystal structures from PDB are colored by secondary structure.
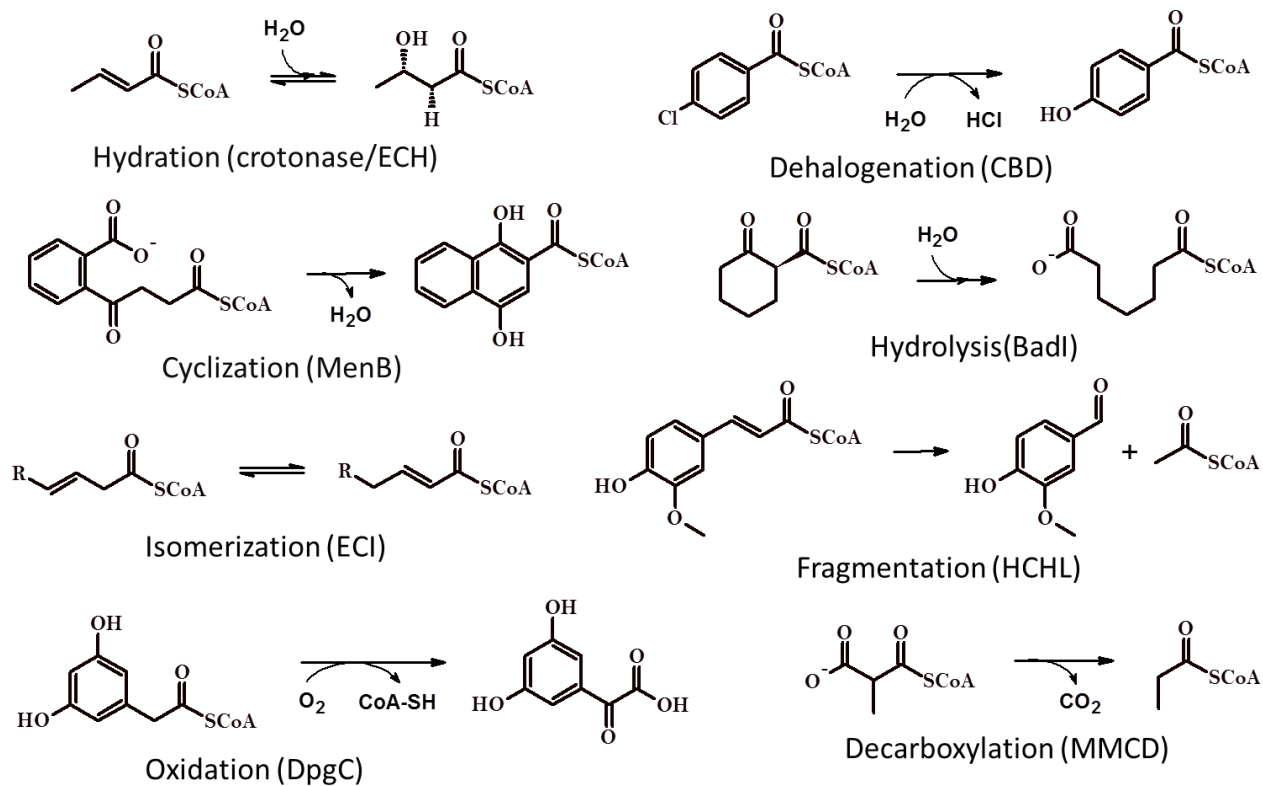
Figure 3.3 Reactions catalyzed by members of the crotonase superfamily

OAH1

```
HCHL:2VSU|E     25 NRPEKRNAMSPTLNREMIDVLETLEQDPAAGVLVLTG---------AGEAWTAGMDLKEYF 76
ECH:1MJ3|A      23 NRPKALNALCNGLIEELNQALETFEEDPAVGAIVLTG--------GEKAFAAGADIKEMQ 74
mtMenB:1Q51|A   53 NRPEVRNAFRPHTVDELYRVLDHARMSPDVGVVLLTGNGPSPK--DGGWAFCSGGDQRIRG 111
DCI:1DCI|A      20 NRPEKRNAMNRAFWRELVECFQKISKDSDCRAVVVSG---------AGKMFTSGIDLMDMA 71
hmECI:1SG4|A    21 KNP-PVNSLSLEFLTELVISLEKLENDKSFRGVILTSD--------RPGVFSAGLDLTEMC 72
DpgC:2NP9|A    183 CRDDRLNAEDGQQVDDMETAVDLALLDPGVRVGLLRGGVMSHPRYRGKRVFSAGINLKYLS 243
MMCD:1EF8|A     20 NYGRKLNALSKVFIDDLMQALSDLNR-PEIRCIILRAP------SGSKVFSAGHDIHELP 72
bmFabM          22 DRPDKLNSFTRDMHRELQSALDDVQAANAR-ALVLTG-------AGRGFCAGQDLADLD 72
yECI:1PJH|C     25 INPDNLNALEGEDYIYLGELLELADRNRDVYFTIIQS---------SGRFFSSGADFKGIA 76
rpBadI          20 NRPDKMNAFRGTTCDELIKALYKAGYDKDVGAIVLAG--------AGDRAFCTGGDQSTHD 72
CBD:1NZY|A      19 KLPRHRNALSVKAMQEVTDALNRAEEDDSVGAVMITG---------AEDAFCAGFYLREIP 70
```

```
HCHL:2VSU|E    110 IAMVNGWCPGGGFAPLVA 127
ECH:1MJ3|A     101 IAAVNGYALGGGCELAMM 118
mtMenB:1Q51|A  151 ICLVNGWAAGGGHSLHVV 168
DCI:1DCI|A     111 IAAIHGGCIGGGVDLISA 128
hmECI:1SG4|A   101 VSAINGACPAGGCLVALT 118
DpgC:2NP9|A    288 VAAVDGFAIGGGAQLLLV 305
MMCD:1EF8|A    100 ISMVEGSVWGGAFEMIMS 117
bmFabM         104 IAAVNGTAAGAGANLAFA 121
yECI:1PJH|C    116 ICCLNGPAIGLSAALVAL 133
rpBadI         100 IARVQGYAIGGGNVLATI 117
CBD:1NZY|A     104 LAAINGVAAGGGLGISLA 121
```

```
                256 QGMKQFLDDKSIKPGLQAYK275
                244 EGMSAFVEKR--KANFKDH-260
                294 EGRDAFLQKR--PPDWSPFP311
                256 KSVQAAMEKK--DSKSITFS273
                246 KSLQMYLERLKEEKG-----260
                429 DKVGRFGGRP---PA-----440
                245 EGMNAFLEKR--KPNFVGH-261
                247 EGVRAFIEKR--APRFEGR-263
                266 KRFRQLGSKQR-KHRL----280
                242 EGVKALQEKR--KPEFRKYI259
                247 PCLTRFLDGHRADRPQVELP266
```

OAH2

Figure 3.4 Residues for CoA recognition in the crotonase superfamily

Many CS members assemble into trimers, which further dimerize into hexamers in some cases. "Domain swapping" of the C-terminal α-helical domains contributes to the differences in their oligomeric assembly [112,274-276]. Three types of folding behavior of the C-terminal domain have been identified. In one type, the C-terminal domain folds back and covers the active site of the same monomer. In the second, the C-terminal domain covers the active site of the next monomer within the same trimer. Finally, the type exemplified by MenB has a C-terminus crossing the trimer-trimer interface and covering the active site of the opposing monomer within the other trimer (**Figure 3.5**).
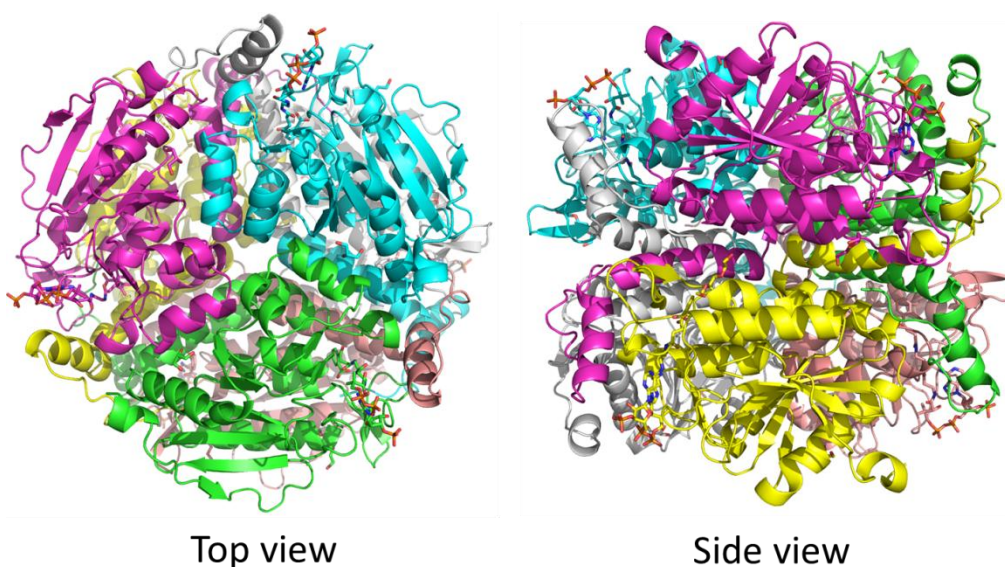
Top view   Side view

Figure 3.5 Hexameric assembly of mtMenB.
The hexamer colored by subunit is a dimer of trimers. The trimer is revealed by the top view and dimer revealed by the side view in which the trimer-trimer interface is perpendicular to the plane of paper. The smaller C-terminal domain crosses this interface. The ligands bound at the interfaces are shown in stick.

Catalytic mechanism

The conserved oxyanion hole in the superfamily is also found in MenB where it plays a critical role by acidifying the OSB-CoA α-protons and promoting carbanion formation (**Figure 3.6**) [109]. However the identity of the base that abstracts the α-proton has still not been fully resolved. The mechanism originally proposed by our group involved an intramolecular proton transfer from C2 to the OSB carboxylate resulting in carbanion formation and also leading to

protonation of the acid, thus making it a better electrophile [109]. In contrast, more recently it has been suggested that an active site Asp or a bicarbonate cofactor functions as the base [277]. In addition, by analogy with enzymes such as the β-ketoacyl-ACP synthases [264], a second oxyanion hole must be present in the active site to stabilize the tetrahedral oxyanion intermediate formed by carbon-carbon bond formation. However the identity of this second oxyanion hole, which is not normally a feature of crotonase superfamily members, is an open question.
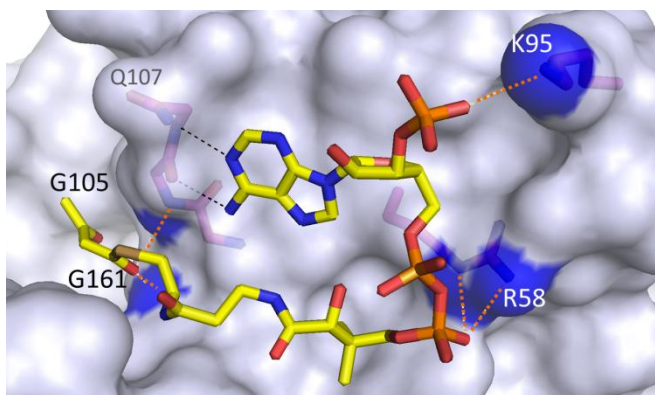
Figure 3.6 The oxyanion hole in mtMenB

The enolate oxyanion in the mtMenB reaction can be stabilized by backbone NH groups of G105 and G161, a conserved feature in the crotonase superfamily.

Our ability to fully elucidate the mechanism of MenB and address the questions raised above has been hindered by the lack of structural data in which the MenB active site is intact. Currently structures exist of the MenB enzymes from *Mycobacterium tuberculosis* (mtMenB, 1Q51, 1RJN; [109,278]), *Staphylococcus aureus* (saMenB, 2UZF; [279]), *Salmonella typhimurium* (stMenB 3H02; [280]) and *Geobacillus kaustophilus* (gkMenB, 2IEX; [281]). However in each case a portion of the MenB active site is disordered. This region displays variability in sequence even within the MenB family and also little conservation in structure throughout the crotonase superfamily, which has hindered our attempts to use sequence homology to predict the function of amino acids in this part of the active site. Although several of these enzymes were crystallized in the presence of acyl-CoA ligands, including the product of the reaction, only in the case of acetoacetyl-CoA can the acyl group be visualized. We hypothesized that a stable analogue of the substrate that retained all enzyme-substrate interactions would make an ideal structural probe and consequently we used a chemoenzymatic approach to synthesize OSB-aminoCoA (OSB-NCoA) in which the thioester sulfur is replaced by a nitrogen. The resulting decrease in α-proton acidity resulting from conversion of the CoA thioester into an amide reduces the stability of the carbanion sufficiently so that the Claisen condensation reaction is prevented. Subsequent structural studies revealed that the OSB-NCoA successfully traps the

previously disordered active site region in a well-defined conformation, providing valuable information on substrate recognition and the catalytic mechanism of this intriguing reaction.

3.2 Methods

Preparation of OSB-NCoA and analogues

OSB-aminoCoA (OSB-NCoA) was synthesized chemoenzymatically by Dr. Xiaokai Li according to the route shown in **Figure 3.7**. This method is based on the strategy employed for the synthesis of acyl-CoA analogues initially described by Drueckhammer [282,283] and employed by us for the synthesis of crotonyl-oxyCoA [284] in which the OSB-aminopantetheine is initially synthesized chemically and then converted to the final product using enzymes from the CoA biosynthetic pathway. OSB-CoA methyl ester and OCPB-CoA were synthesized by Dr. Huaning Zhang.
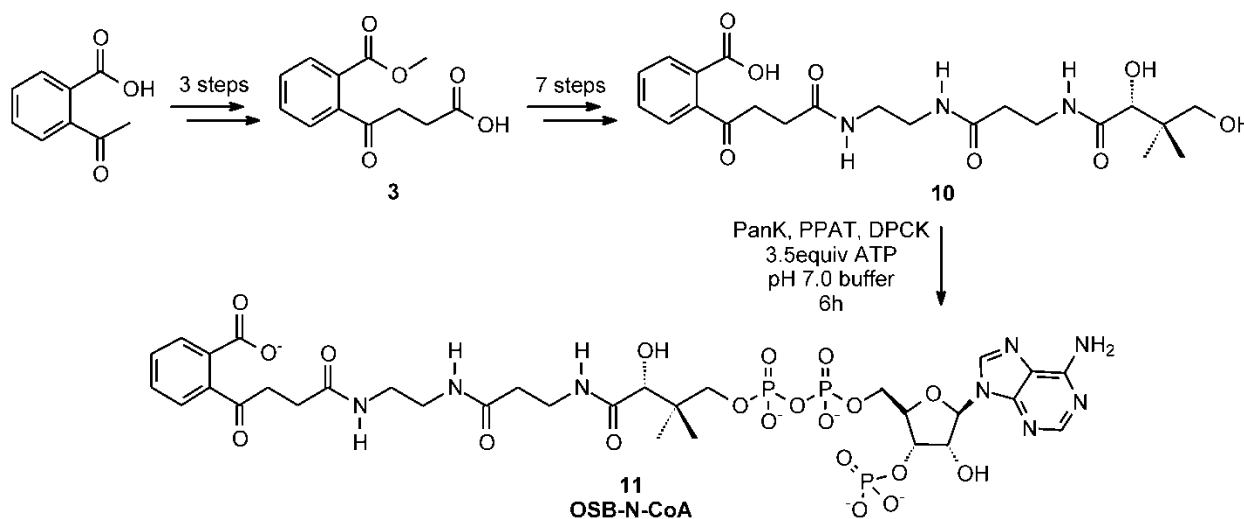
Figure 3.7 Chemoenzymatic synthesis of OSB-NCoA

Preparation of wild-type and mutant MenB enzymes

The MenB enzyme from *Mycobacterium tuberculosis* (mtMenB) was expressed and purified as described previously [109]. The *E. coli menb* gene *b2262* (858 bp) was cloned into the pET-15b plasmid (Novagen) and placed in frame with an N-terminal His-tag sequence by Dr. Yuguo Feng. Protein expression was performed by Dr. Nina Liu using BL21 (DE3) *E. coli* cells essentially as described for mtMenB. A 1 l culture of cells was induced with 1 mM isopropyl-β-D-

thiogalactopyranoside (IPTG) at an OD_{600} of 0.8 and harvested by centrifugation after shaking for 12 h at 25°C. The cell pellet was then resuspended in 30 ml of His binding buffer (5 mM imidazole, 0.5 M NaCl, 20 mM Tris HCl, pH 7.9) and lysed by 3 passages through a French Press cell (1000 psi). Cell debris was removed by centrifugation at 33,000 rpm for 60 min at 4 °C. ecMenB was purified using His-tag affinity chromatography. The supernatant was loaded onto a His-bind column (1.5 cm x 15 cm) containing 4 ml of His-bind resin (Novagen) that had been charged with 9 ml of charge buffer (Ni^{2+}). The column was washed with 60 ml of His-binding buffer and 30 ml of wash buffer (60 mM imidazole, 0.5 M NaCl, 20 mM Tris HCl, pH 7.9). Subsequently, the protein was eluted using a gradient of 20 ml elution buffer (60-500 mM imidazole, 0.5 M NaCl, 20 mM Tris HCl, pH 7.9). Fractions containing ecMenB were collected and the imidazole removed using a Sephadex G-25 chromatography column (1.5 cm x 55 cm) with a buffer containing 20 mM NaH_2PO_4, 0.1 M NaCl at pH 7.0 as eluent. The protein was found unstable with the His-tag, which was removed by treating the protein with biotinylated thrombin overnight at RT. After the biotinylated thrombin was captured with streptavidin agarose, the protein was concentrated using a Centricon-30 (Amicon) concentrator and stored at -80 ºC. The concentration of ecMenB was determined by measuring the absorption at 280 nm and using an extinction coefficient of 36,040 $M^{-1}cm^{-1}$ calculated from the primary sequence. The purified protein was >95% pure on SDS-PAGE which gave an apparent MW of ~30 kDa.

Quikchange mutagenesis was used to prepare the D185E, D185G and S190A mutants of mtMenB and the Y97F and G156D mutants of ecMenB by Dr. Yuguo Feng, Huaning Zhang, Nina Liu and Xiaokai Li. These mutant proteins were expressed, purified and stored as described for the wild-type enzymes.

Crystallization, data collection and structure determination

apo-ecMenB and OSB-NCoA:ecMenB

Apo-ecMenB was crystallized using the sitting drop vapor diffusion technique. 0.5 µl of 350 µM protein solution was mixed with 0.5 µl of reservoir solution (200 mM sodium malonate at pH 7.0 and 20% PEG 3350, Hampton Research) and equilibrated against 75 µl of reservoir solution. Plates of 0.3 mm length in the longest dimension appeared in one day. The crystals were cryo-protected using a solution that contained 200 mM sodium malonate at pH 7.0, 21% PEG 3350, 25% glycerol and 20 mM NaH_2PO_4 at pH 7 and cryo-cooled in liquid nitrogen.

ecMenB in complex with OSB-NCoA was crystallized similarly using a sitting drop comprising 0.5 µl of 350 µM protein solution containing 1.7 mM of OSB-NCoA and 0.5 µl of reservoir solution (100 mM Bis-tris at pH 6.5, 200 mM NaCl and 25% PEG 3350, Hampton Research). Small bipyramids appeared in one day, and a long rod (0.6 mm in the long dimension) with a hollow core appeared in two days (**Figure 3.8**). The bipyramids did not

improve and were not used for the diffraction experiment. A portion of the long rod was cryo-protected in a solution containing 100 mM Bis-Tris at pH 6.5, 200 mM NaCl, 26% PEG 3350, 25% glycerol, 20 mM NaH_2PO_4 at pH 7 and 200 µM OSB-NCoA, and cryo-cooled in liquid nitrogen.

Figure 3.8 Crystal of ecMenB in complex with OSB-NCoA

The long rod used for data collection appears on the background of small bipyrimidal crystals.

Data were collected at beamline X29 at the National Synchrotron Light Source (NSLS) in Brookhaven National Laboratory, indexed, integrated and scaled using HKL2000 [285]. The binary structure was solved by MolRep [207] using the structure of MenB from *Salmonella typhimurium* (stMenB) as a search model (PDB code 3H02). Clear densities for OSB-NCoA and the missing residues 89-103 in the search model were revealed and built into the model. Residues in ecMenB inconsistent with the stMenB structure were fixed. Restraints for OSB-NCoA were generated by eLBOW [286] using the appropriate SMILES string [287] (see below), and coordinates were refined in Phenix [288]. Manual model building was performed in Coot [289]. The density adjacent to G156 and the succinyl group of OSB-NCoA could not be simply accounted for by water. Modeling of a chloride ion present in the crystallization buffer eliminated the majority of the difference density. Several cycles of refinement followed by manual model building reduced both the working and free R factors significantly to below 20%. Densities for part of a PEG 3350 molecule (essentially three units of ethylene glycol) and glycerol from the cryo solution were found on the surface and were included in the structure. Protein geometry and all-atom contacts including added hydrogens were analyzed by MolProbity [290]. Some Asn, Gln and His residues were subsequently flipped following this analysis. Although MolProbity suggested clashes between the OSB C2 and the oxygen attached to C7 in some monomers, the density map suggests that these atoms are indeed brought into close proximity.

The structural parameter file (.cif) for OSB-NCoA was generated by eLBOW using the following SMILES string [291]:

CC(C)(COP(=O)([O-])OP(=O)([O-])OC[C@H]3O[C@@H](n2cnc1c(N)ncnc12)[C@H](O)[C@@H]3OP(=O)([O-])[O-])[C@@H](O)C(=O)NCCC(=O)NCCNC(=O)CCC(=O)c4ccccc4C(=O)[O-]

115
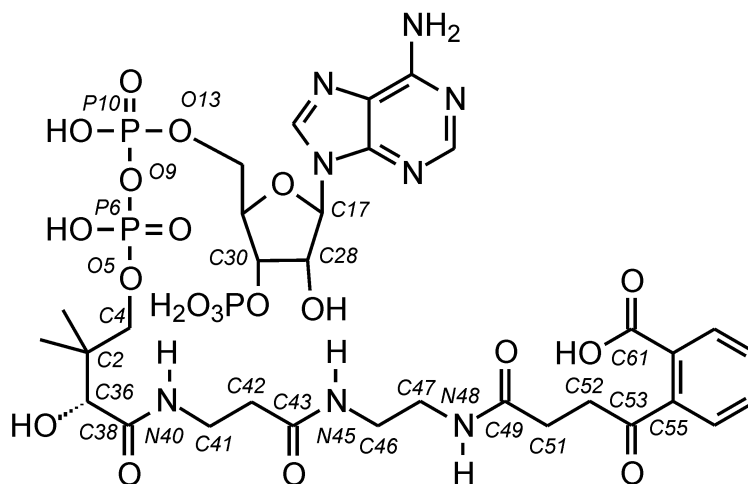
Figure 3.9 Atom numbering in the ligand, OSB-NCoA, used in the refinement

The resulting atom numbers used in the refinement of the OSB-NCoA protein-ligand structures are given in **Figure 3.9**. The generated cif file was edited to improve chemical plausibility. The bond type of the free OSB carboxylate (*C61*) was changed to 'deloc' with a C-O bond length of 1.25Å. All the C-C, C-N and C-O single bond lengths on the main-chain between *O5* and *C55* were allowed less restraint by re-assigning a value_dist_esd of 0.05 from the default value of 0.02. Bond angles centering on *C2, C4, C36, C41, C42, C46, C47, C51* and *C52* as well as the three angles *O9-P6-O5*, *O13-P10-O9*, and *C30-C28-C17* were relaxed by re-assigning a value_angle_esd of 9.0 from the default value of 3.0. These modifications improved the geometry around C36, which otherwise gave a short C36-C38 bond length of < 1.50 Å. The planarity in the file was also tested to see whether the C4 carbonyl plane (defined by *C52, C53, O54 and C55*) co-planar with the benzene ring would improve the fit. However, modeling either way gave different errors. The restraint was not used for the final refinement and the C4 carbonyl remained out of the plane formed by the aromatic ring. The planarity of the aromatic carboxylate (*C61, O62, O63*), C4 carbonyl (*C52, C53, O54, C55*) and the three amide bonds was maintained by restraining the dist_esd value of the constituent atoms on the plane.

mtMenB

A solution of 280 μM mtMenB was incubated with 3.7 mM OSB-NCoA for 1 h on ice. One μl of this protein solution was then mixed with 1 μl of reservoir solution containing 100 mM HEPES pH 8.0, 300 mM NaCl and 19% PEG 6000 and set up for hanging drop, vapor diffusion crystallization against 500 μl of reservoir solution at RT. The crystal grew into an irregular rod, 0.2 mm in the long dimension, and was cryo-cooled in liquid nitrogen on day 6 after stepwise transfer into a cryo solution containing 100 mM HEPES pH 8.0, 20 mM NaH_2PO_4 pH 7.0, 1 mM OSB-NCoA, 100 mM NaCl, 20% PEG 6000, and 22% glycerol. In a separate setup, 280 μM mtMenB was incubated with 1 mM OSB-NCoA for 40 min at RT after which 1 μl of this solution was mixed with 1 μl of reservoir solution containing 100 mM Tris pH 8.0, 200 mM Li_2SO_4 and 19% PEG 3350 and set up for hanging drop, vapor diffusion crystallization against 500 μl of reservoir solution at RT. Crystals of various shapes were obtained and cryo-cooled in liquid nitrogen on day 14 after soaking in a cryo solution containing 100 mM Tris pH 8.0, 200 mM Li_2SO_4, 23% PEG 3350, 22% glycerol, 20 mM NaH_2PO_4 pH 7.0 and 1 mM OSB-NCoA.
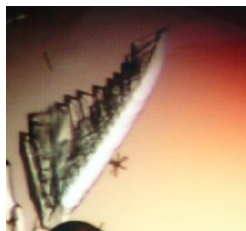
Figure 3.10 Crystal of mtMenB from the altered hexamer

Diffraction data were collected at beamline X29 at the NSLS, indexed, integrated and scaled using HKL2000. Structure solutions were found by MolRep or Phaser [292] molecular replacement using the mtMenB structure 1Q51 as the search model [109]. The same hexagonal crystal form was observed from both crystallization conditions, but an additional rhombohedral crystal form was obtained from the second condition. The rhombohedral structure was solved with two monomers in the asymmetric unit, but the second monomer was found by molecular replacement only after removing the C-terminal domain from the 1Q51 search model.

Enzyme kinetics

Steady-state kinetic experiments were performed by Dr. Yuguo Feng, Huaning Zhang, Nina Liu and Xiaokai Li and Ramona Franciska Gajda. The kinetic parameters for ecMenB and mtMenB were determined in 20 mM Na_2HPO_4, 150 mM NaCl, 1 mM $MgCl_2$, pH 7.0 buffer at 25 °C using a coupled assay [109]. The formation of DHNA-CoA at 392 nm was monitored using a CARY-300 spectrophotometer, and reactions were initiated by addition of MenB to solutions containing OSB, ATP (120 μM), CoA (120 μM) and ecMenE (4 μM). The final assay volume was 500 μl. Kinetic parameters were calculated by fitting the initial velocity data to the Michaelis-Menten equation and using an extinction coefficient of 4, 000 $M^{-1}cm^{-1}$ for DHNA-CoA at 392 nm.

3.3 Results and Discussion

Statistics from diffraction data processing, model refinement and validation are summarized in **Table 3.1**.

Table 3.1 Data collection and refinement statistics

	ecMenB:OSB-NCoA	apo-ecMenB	mtMenB:OSB-NCoA	apo-mtMenB
PDB ID	3T88	3T89	3T8A	3T8B
Data Collection				
Space group	$P2_12_12$	$P2_12_12_1$	$P6_122$	R3
Unit cell dimensions a, b, c (Å)	140.49, 141.79, 89.12	76.51, 134.00, 153.36	87.13, 87.13, 414.84	132.23, 132.23, 71.14
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Redundancy	6.9 (6.0)	6.8 (6.5)	15.2 (9.5)	5.7 (5.6)
Completeness (%)	99.8 (99.1)	98.2 (95.8)	99.6 (95.2)	100.0 (100.0)
No. unique reflections	120517	113570	45905	55893
I/σI	18.5 (2.4)	17.1 (2.6)	26.1 (3.4)	21.2 (2.2)
R_{merge}	0.163 (0.674)	0.101 (0.537)	0.113 (0.542)	0.083 (0.681)
Refinement				
Resolution range (Å)	44.8-2.00	42.9-1.95	46.6-2.25	37.0-1.65
No. atoms	14295	12315	6066	3703
Protein	13004	11770	5883	3417
Water	811	449	107	280
Ligands/ions	480	96	76	6
Average *B* factors	22.4	27.1	37.9	21.9
Protein	21.9	27.0	37.7	21.2
Waters	26.6	29.3	34.3	31.2
Ligands/ions	28.4	31.6	62.2	21.7
R_{work}/R_{free}	0.155/0.195	0.186/0.219	0.194/0.237	0.141/0.187
RMSD from ideal values in				
bond length (Å)	0.008	0.005	0.007	0.004
bond angle (°)	1.2	0.84	0.95	0.83

Values for the highest resolution shell are given in parentheses.

Active site disorder

MenB catalyzes an intramolecular Claisen condensation reaction in which the electrophile is an unactivated carboxylic acid. Apart from the oxyanion hole that stabilizes the carbanion (enolate) and which is a common feature of crotonase superfamily enzymes, insight into other aspects of the mechanism including α-proton abstraction, carbon-carbon bond formation and stabilization of the resulting tetrahedral oxyanion requires specific information on how the acyl portion of the substrate interacts with active site residues. Based on previous structures of mtMenB in complex with acetoacetyl-CoA and DHNA-CoA, L134, I136 and L137 have been suggested to make hydrophobic contacts with the aromatic ring of OSB while D185, A192, S190 and Y287 are potential proton donors and acceptors involved in the MenB reaction [109,278]. However the exact role that each residue plays in the reaction is difficult to determine given the active site disorder, which is normally observed in MenB X-ray structures and which, for mtMenB, includes residues 108-125 (the A-loop) (**Figure 3.11**). Although the acyl group for acetoacetyl-CoA can be observed, this is a poor mimic of the natural substrate. In addition, the acyl group of the product, DHNA-CoA, cannot be observed most likely due to disorder. Nevertheless, we noticed that in the saMenB structure with acetoacetyl-CoA, only 8 residues (residues 79-86) were missing from the A-loop [279] presumably due to its shorter sequence in this nonconserved region. We therefore speculated that an acyl-CoA ligand that better mimicked the native substrate combined with a MenB enzyme with a shorter loop might increase the chance of revealing an intact active site together with details of the enzyme-substrate interactions.
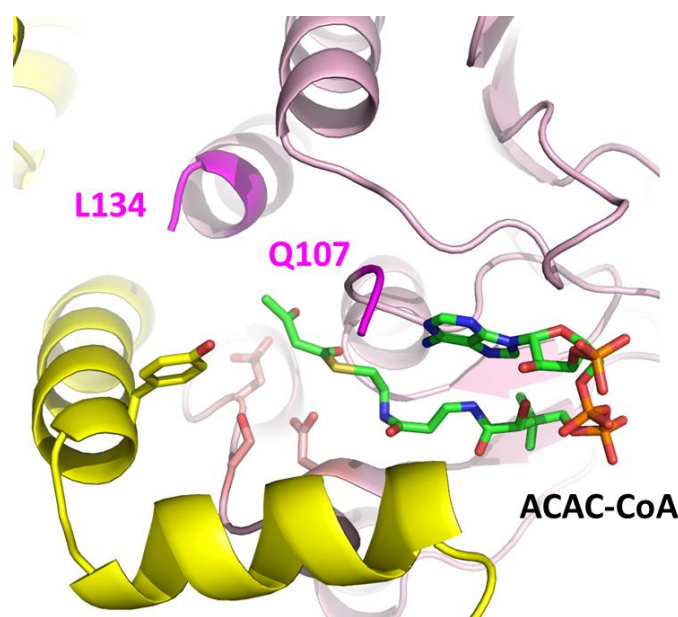
Figure 3.11 Disorder of the A-loop in mtMenB

The OSB-NCoA substrate analogue

We speculated that replacing the thioester in OSB-CoA with an amide would reduce α-proton acidity sufficiently to prevent formation of the carbanion required for carbon-carbon bond formation. This approach has previously been utilized in structural studies of succinyltransferase [293], while an acyl-CoA amide analogue has also been used to probe the mechanism of substrate recognition and catalysis by the crotonase superfamily member dihydroxyphenylglyoxylate synthase (DpgC) [294-296]. In addition the amide analogue will be more stable to hydrolysis than the thioester counterpart, and so Dr. Xiaokai Li prepared OSB-NCoA using the chemoenzymatic route in **Figure 3.7**. As anticipated, no reaction could be detected when OSB-NCoA was incubated with ecMenB or mtMenB even at high (100 µM) enzyme concentrations. We thus used OSB-NCoA as a ligand for co-crystallization with ecMenB and mtMenB.

Structure of the ecMenB:OSB-NCoA complex

While apo-ecMenB crystallizes in the same space group ($P2_12_12_1$) as apo-stMenB (3H02), co-crystallization of ecMenB with OSB-NCoA resulted in a different crystal form with space group $P2_12_12$. Both structures were refined to a similar resolution of 2 Å with R factors below 0.2 (**Table 3.1**). The asymmetric unit contains a hexamer in both cases with the same tertiary and quaternary structure observed for other MenB enzymes. The crotonase fold and hexameric assembly are shown in **Figure 3.12**. The three C-terminal helices (H14-H16, the C-loop) fold across the trimer-trimer interface and contribute to part of the active site in the opposing trimer, a characteristic of all known MenB hexamers [109,278,279] (**Figure 3.12**, **Figure 3.13**).
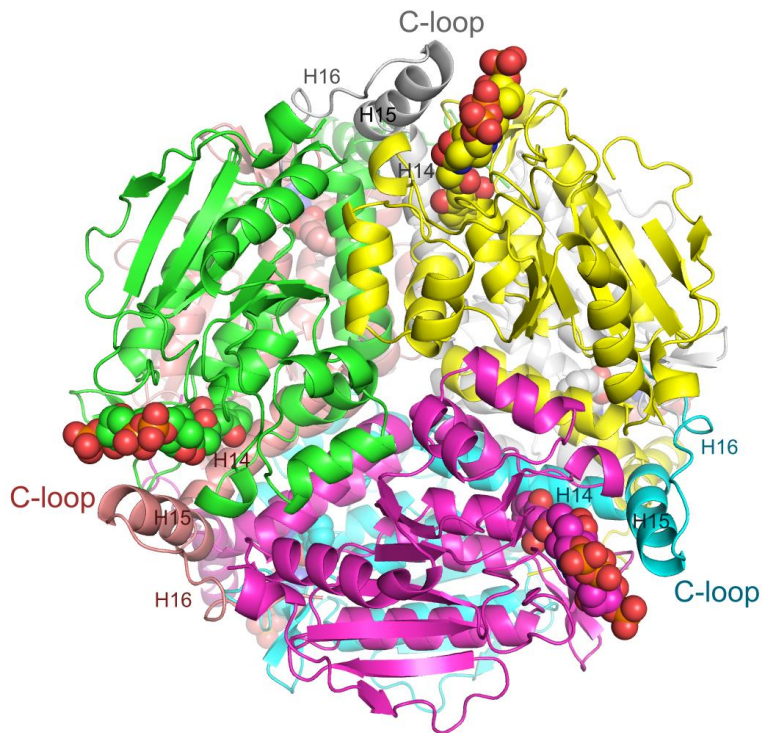
Figure 3.12 The x-ray structure of the ecMenB:OSB-NCoA complex

Hexameric assembly of ecMenB in complex with OSB-NCoA viewed along the three-fold axis of trimers. The protein is colored by chain and OSB-NCoA is shown in spheres. The C-loop comprising H14-H16 from the back trimer folds across the trimer-trimer interface and covers the active site of the front trimer.
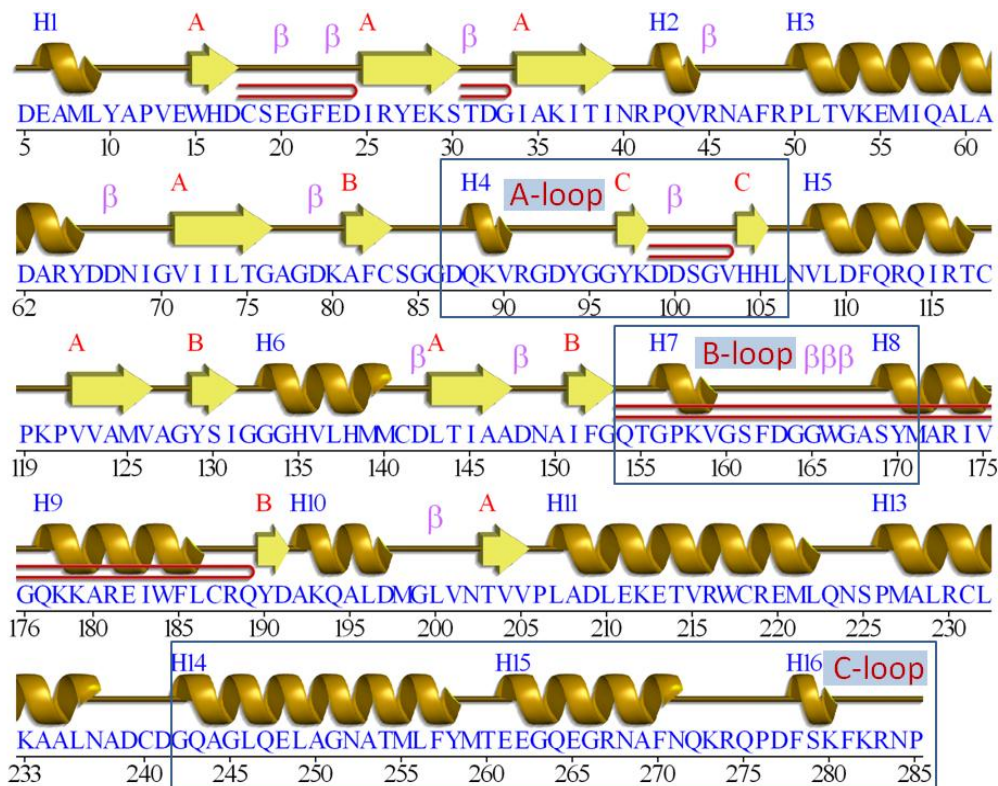
Figure 3.13 Assignment of secondary structures in the ecMenB structure in complex with OSB-NCoA.

The diagram was generated by PDBsum [297]. The locations of the A, B and C-loops are indicated by boxes.

Conformation of the A-loop

In the structure of OSB-NCoA bound to ecMenB, the ligand can be observed in all six active sites of the hexamer. Significantly, there is no active site disorder and thus the entire active site of MenB is revealed for the first time (**Figure 3.14**). The A-loop, comprising Q88 to L106 in ecMenB, is locked in the same conformation for all six monomers, forming an additional α coil followed by a loop leading to a β turn and a short β hairpin (**Figure 3.15**). This extensive β-like secondary structure is apparently only trapped by substrate binding since secondary structure prediction in the absence of ligand only infers a disordered loop or α-helical structure [298]. The A-loop not only covers the active site, but also packs against the C-terminal helices from two neighboring monomers (yellow and cyan monomers, **Figure 3.15**), contributing to the integrity of the MenB hexameric assembly. This includes the C-terminus of a monomer from the opposing trimer which, together with the A-loop, forms part of the active site (C-loop, yellow monomer, **Figure 3.15**) (**Figure 3.16**). In contrast, the A-loop remains disordered for apo-

123

ecMenB as observed in other known MenB structures, and part of the C-loop is also less ordered or displays higher B factors. Placing the hexamer for the enzyme-ligand complex in the lattice of the apo structure generates clashes between the A-loop and an adjacent apo hexamer, indicating incompatibility of this A-loop conformation in the apo crystal form. Apparently the crystal form is altered for the binary complex due to the change in A-loop location upon binding of OSB-NCoA.
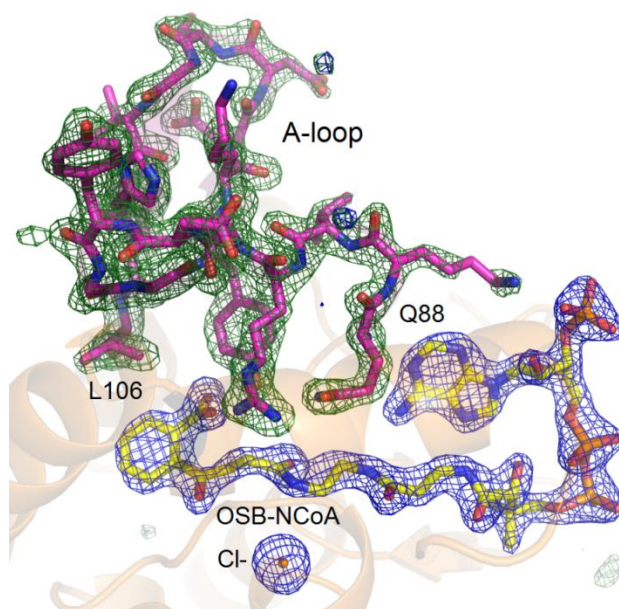
Figure 3.14 Ordered A-loop and OSB-NCoA

Simulated annealing fo-fc omit maps showing density of OSB-NCoA, the bound anion, and the ordered active site loop region between Q88 and L106 (A-loop). The green map is generated by omitting Q88-L106 from chain C in the model, and the blue map generated by omitting OSB-NCoA and the chloride ion from chain C. The mesh level is contoured at 4.5 σ.
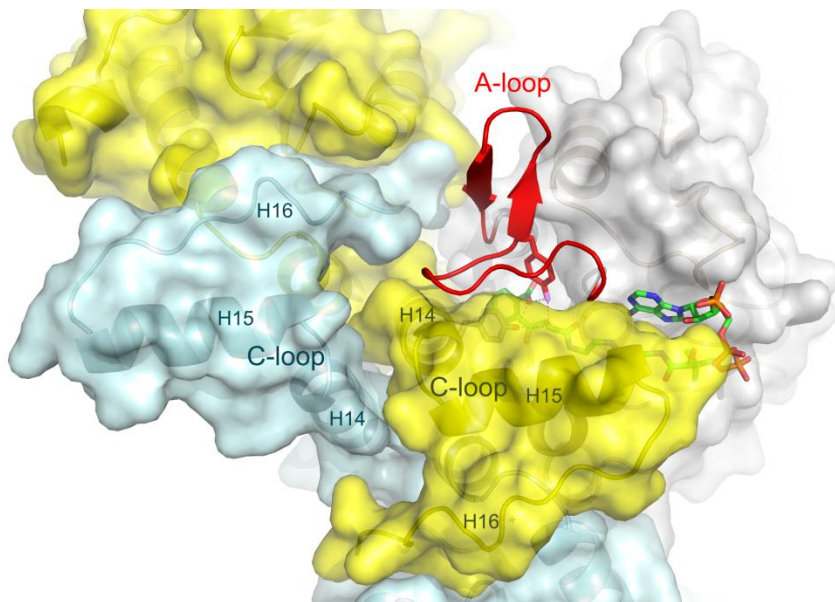
Figure 3.15 Secondary structure of the A-loop

The three monomers in grey, yellow and cyan from **Figure 3.12** are displayed. The A-loop from the grey monomer covers the active site and makes contacts with the C-terminal helices from two neighboring monomers. OSB-NCoA is shown in sticks.
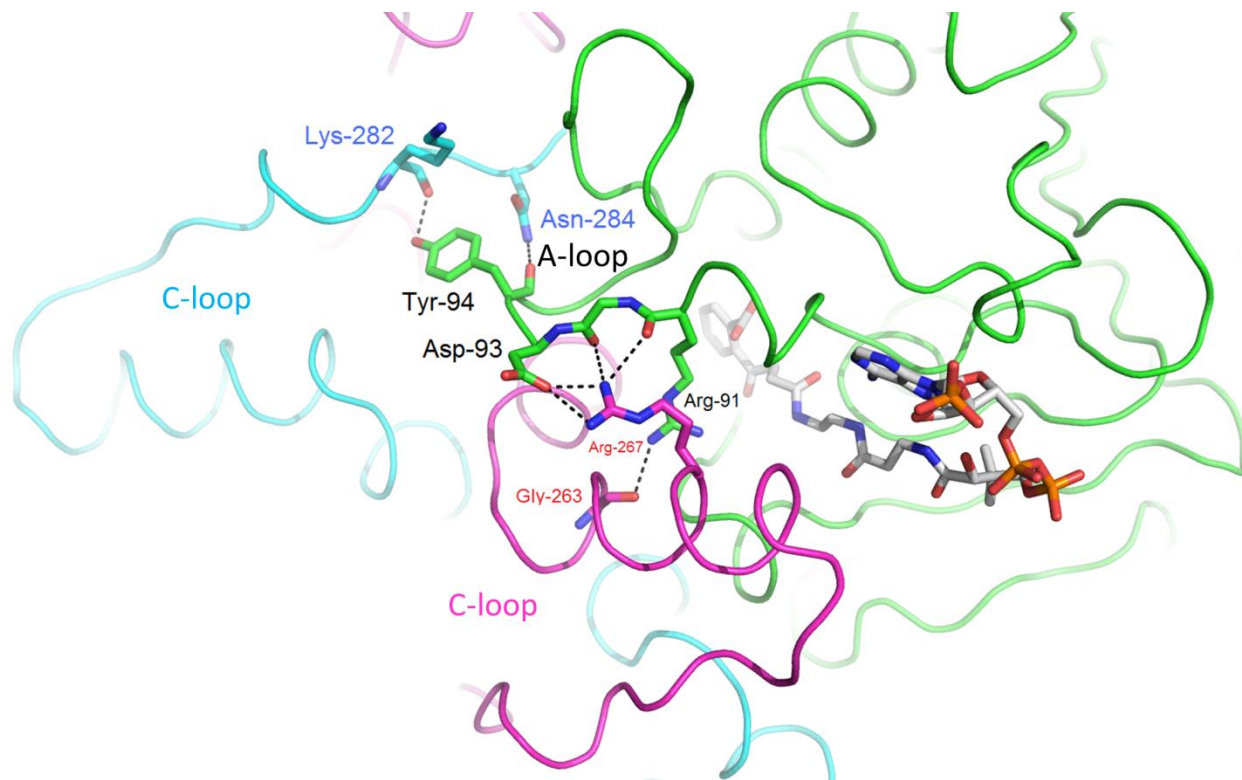
Figure 3.16 Interactions of the A-loop with neighboring subunits

125

The flexible A-loop becomes ordered upon binding OSB-NCoA and interacts with the C-loop from two other subunits.

C-loop conformational change upon binding OSB-NCoA

The C-loop is another region that has a tendency to show higher mobility in MenB. In mtMenB, the C-loop after Gly-271 has been shown to display higher B factors than the rest of the structure, or even become disordered after Ala-286 [112,278]. The two ecMenB structures presented here also display higher B factors in the C-terminal helix after Glu-261. The disorder is worse in the apo structure, with many residues losing their molecular features in the density map. Residues after Glu-261 in Chain C are too disordered to be modeled completely. Nevertheless, it is clear that the C-loop has shifted toward the substrate binding pocket upon binding of OSB-NCoA, most significantly at the C-terminal end of helix 15 (around Asn-271); the largest C_α displacement can be greater than 3 Å (**Figure 3.17**). Similar displacement up to 1.1 Å was observed for mtMenB upon binding of acetoacetyl CoA [112]. The movement of the C-loop in ecMenB is stabilized by the additional salt bridge between Lys-273, Lys-89 and the ribose phosphate of OSB-NCoA, and hydrogen bonding interactions between the hydroxyl of ribose, helix 4, and the C-terminal end of helix 15 via mainchain of Phe-270, sidechain of Asn-271 and ordered water molecules, bringing helix 15 closer to helix 4 and the CoA pocket (**Figure 3.18**). In the apo ecMenB structure, these residues tend to become disordered. Although these residues are conserved or similar in mtMenB (Phe-299, Leu-300, Lys-302 and Arg-108), Lys-302 is 4.5 Å from the ribose phosphate of acetoacetyl-CoA (compared to 2.5 Å in ecMenB), and Arg-108 and water molecules are not ordered, making many of these ecMenB interactions apparently absent in mtMenB bound with acetoacetyl-CoA. mtMenB appears to position a lysine at an additional spot to assist in CoA binding— an extended β turn between β strand A4 and strand B1 has allowed the inserted Lys-95 to reach the proximity of ribose phosphate of CoA from the other side (**Figure 3.19**). These results show that there is slight variability in binding the CoA portion of substrate between mtMenB and ecMenB, and the substrate is apparently shielded better when the correct substrate is bound, which might be the strategy employed by MenB for substrate selection.
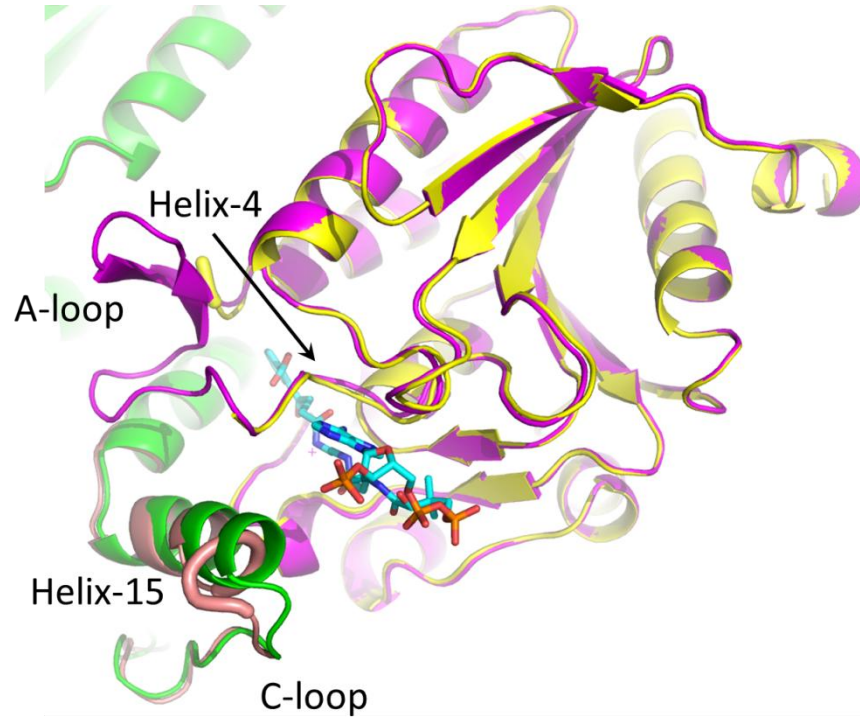
Figure 3.17 Conformational change upon binding OSB-NCoA

ecMenB monomer and the C-loop of the opposing monomer covering the active site. Apo (in yellow and salmon) and complex with OSB-NCoA (in magenta and green) are superimposed. Upon binding OSB-NCoA (shown in sticks), the A-loop (Gln-88 to Leu-106) becomes ordered, and the C-loop (especially around Asn-271) shifts toward the CoA binding pocket.
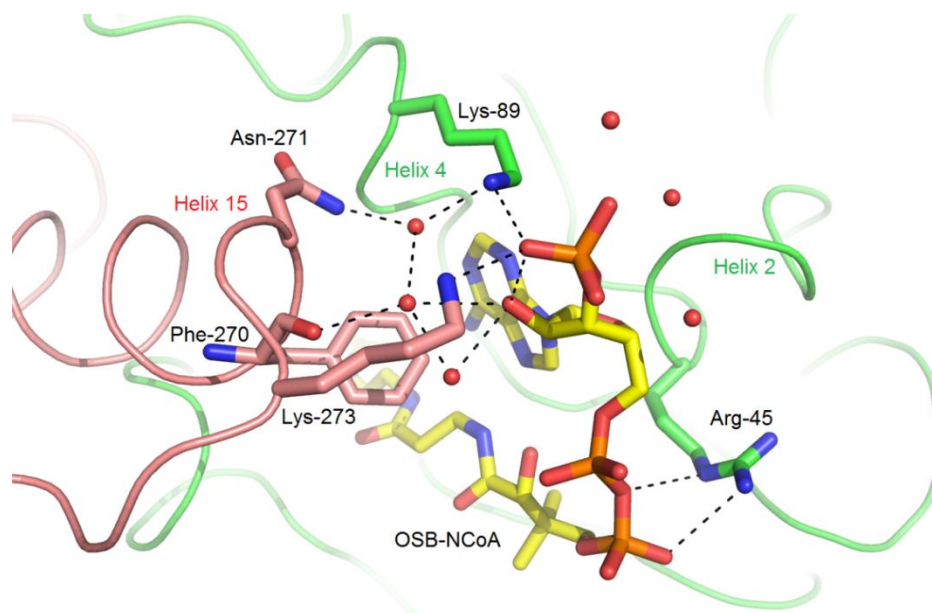
Figure 3.18 CoA binding in the ecMenB:OSB-NCoA complex

ecMenB bound with OSB-NCoA, showing specific interactions that bring helix 15 in the C-loop closer to helix 4 and the CoA binding pocket. Lys-273 and Lys-89 contribute to the salt bridge while Phe-270 and Asn-271 are hydrogen-bonded via ordered water molecules. These CoA-binding residues along with Arg-45 tend to become disordered in the apo ecMenB structure.
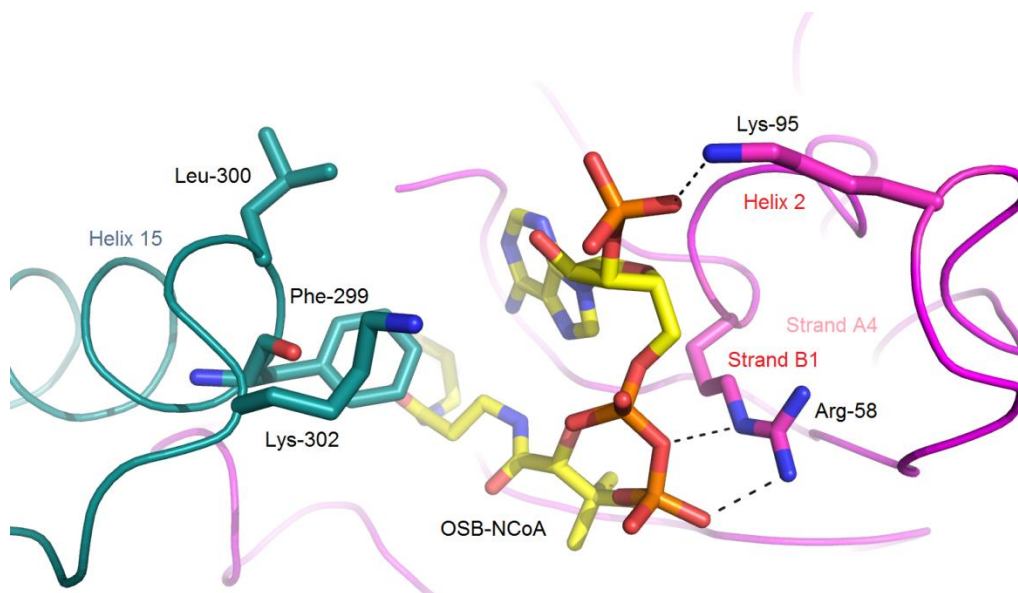
Figure 3.19 mtMenB bound with acetoacetyl-CoA

Phe-299, Lys-302, and Arg-58 are equivalent to Phe-270, Lys-273 and Arg-45 in ecMenB. In this case Lys-302 is placed too far from the ribose phosphate of OSB-NCoA to interact productively. The loss may be compensated by interacting with Lys-95, which is inserted into the extended beta turn between strand A4 and B1. Lys-89 in ecMenB is replaced with Arg-108 in mtMenB, which is part of the disordered loop in the structure.

Binding mode of the substrate

OSB-NCoA adopts the same conformation in each of the 6 active sites. The CoA portion of the ligand is bound in the U-shaped conformation that is seen in the structures of other crotonase family members [109,275,294,299-304]. In addition, the OSB portion is bound in such a way that the C7 carboxylate and C4 carbonyl are clearly out of the plane of the benzene ring, which is also the case when OSB is bound in OSB synthase from *Amycolatopsis* sp. (1SJB; [305]) and OSB-CoA synthase from *Thermobifida fusca* (2QVH; [306]). In the MenB reaction carbon-carbon bond formation occurs between the OSB C2 and C7 atoms which are within 3.5 Å of each other in the OSB-NCoA structure. In contrast, the distance between the C2 and C7 atoms is > 5.3 Å in the OSB and OSB-CoA synthases where C-C bond formation does not occur (1FHV, 1SJB and 2QVH [305,307]). The OSB-NCoA benzoate is surrounded by Y97, F48, L106, V108, L109, Q112, A163 and F162 from the same monomer, Y258 and T254 from the opposing monomer, and an array of eight ordered water molecules. The side chains of L106, V108, L109 and T254 provide favorable hydrophobic interactions with the benzene ring, while the hydroxyl groups from Y97, Y258, and an ordered water molecule provide hydrogen bonding interactions to the carboxylate (**Figure 3.20**). Two additional water molecules extend the hydrogen bonding network to the backbone carbonyl of F48. Other ordered water molecules in the binding pocket mediate a second hydrogen bonding network that involves the backbone carbonyl groups of F162 and G133, and the side chains of T254, D163 and Q112. In addition, S161, an ordered water and a chloride ion are located adjacent to the OSB succinyl moiety. The chloride ion is modeled instead of a water molecule because the electron density is larger than that expected for water. In addition, the negatively charged chloride also interacts favorably with the side chain of Q154 and the backbone NH group of T155. The water molecule is adjacent to the OSB C2 and makes polar interactions with the chloride ion, the oxygen atom of the C4 carbonyl on OSB and the side chain of S161. Finally, the OSB-NCoA amide carbonyl oxygen is hydrogen bonded in the oxyanion hole formed by the backbone NH groups of G86 and G133 (**Figure 3.20**), a feature that is characteristic of the crotonase superfamily. These interactions orient the OSB moiety into the conformation required for carbon-carbon bond formation, and also prevent the OSB carboxylate from attacking the OSB C-4 carbonyl, which is the first step in the uncatalyzed decomposition of OSB-CoA to spirodilactone.
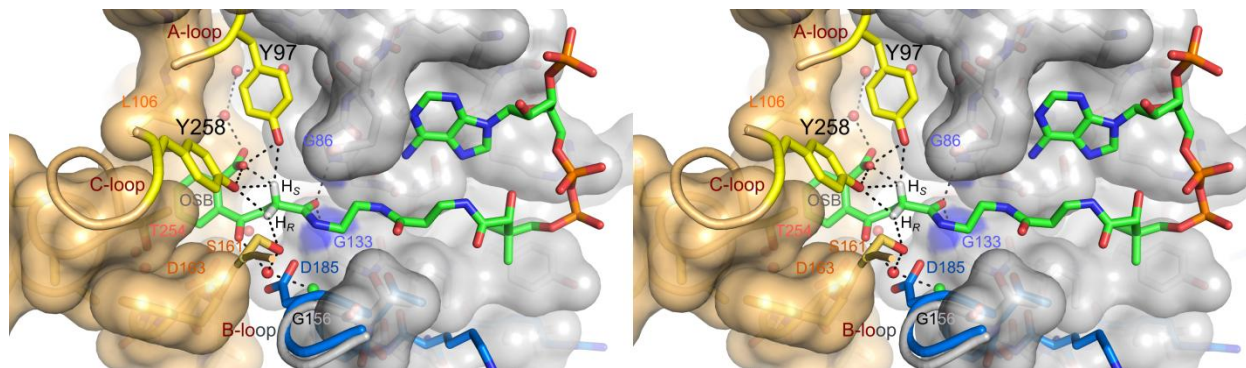
Figure 3.20 Substrate recognition in the active site

The stereo figure shows the interactions between the OSB group and active site residues. The two hydrogen atoms have been added to the C2 carbon (white). Waters are shown as red spheres and the chloride ion as green sphere. The surface of the oxyanion hole is colored blue. The blue loop from mtMenB (PDB 1Q51) is superimposed and the side chain of D185 from mtMenB is displayed. Note that D185 is glycine in ecMenB.

Intramolecular proton transfer

Compared with the structure of OSB determined by energy minimization [308], the bound OSB in the MenB active site adopts a reactive conformation in which the C7 carboxylate is positioned significantly closer toward C2 (**Figure 3.21**). In the unbound structure the oxygen atom on the C7 carboxylate is more than 3.3 Å away from C2 whereas in the bound OSB this distance is between 2.7 and 2.9 Å. Although this is considered an unfavorably close approach of the two atoms by MolProbity, the density map supports the observed interaction which, for the natural substrate, will facilitate the transfer of the pro-2S proton from C2 to C7. This proton transfer is likely facilitated by two nearby Tyr residues, Y97 and Y258 (**Figure 3.20**), which are expected to have pKa values intermediate between the α-proton and the OSB-carboxylate. Instead of direct transfer from C2 to C7, it is possible that Y97 acts as a proton shuttle given its *syn* position with respect to the carboxylate [309] and its orientation with respect to the C2-H*s* bond. In addition, the hydrogen bonding pattern in which Y97 and Y258 donate two in-plane hydrogen bonds to one oxygen atom of the C7 carboxyl group and an ordered water donates one out-of-plane hydrogen bond to the second oxygen suggests that the C7 carboxylate is deprotonated [310]. Therefore the C7 carboxylate remains capable of accepting a proton in this binding environment. Thus, intramolecular proton transfer from C2 to C7 not only leads to carbanion formation but also protonates the carboxylate thus making it a better electrophile. In this regard, we note that the involvement of a substrate carboxylate as an acid/base catalyst has also been suggested for 3(S)-methylglutaconyl-CoA hydratase, another member of the crotonase superfamily [311].
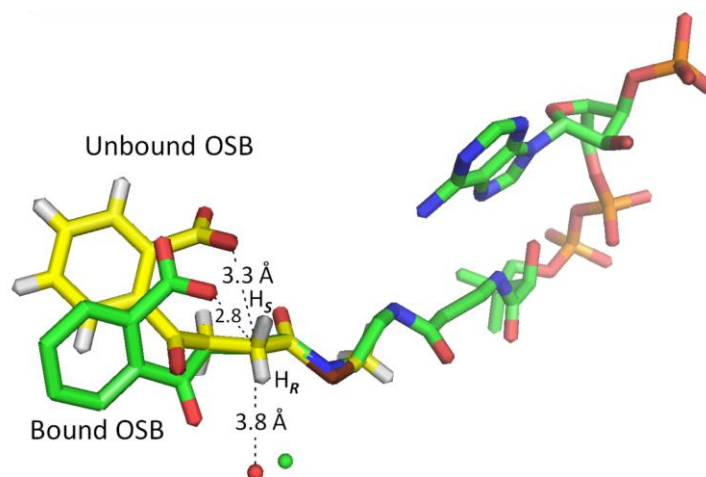
Figure 3.21 Comparison of bound and unbound OSB conformations
Overlay of energy-minimized unbound OSB (yellow) and OSB bound to ecMenB (green), showing the reactive conformation in which the distance between the C7 carboxylate oxygen and C2 carbon has decreased.

The tetrahedral oxyanion hole

By analogy to the Claisen condensation reactions catalyzed by the β-ketoacyl-ACP synthases in fatty acid biosynthesis, a second oxyanion hole is required to stabilize the tetrahedral oxyanion that results from carbanion attack on the protonated OSB carboxyl group [263,264]. The two Tyr residues, Y97 and Y258, are ideally positioned to fulfill this function. Sequence alignment demonstrates that these residues are conserved in the MenB family despite significant variability in the A-loop that carries Y97 (**Figure 3.23**). The two Tyr residues are also conserved in BadI, a crotonase superfamily member that catalyzes a retro-Dieckmann condensation presumably via an analogous tetrahedral oxyanion intermediate (**Figure 3.22**) [312]. The alignment with Y97 is not straightforward without the present structure given that additional Tyr residues are present in the A-loop of other MenB enzymes and also because of the active site disorder that characterizes other MenB structures. Once Y97 is aligned properly, the conservation of G96 is also revealed. G96 is adjacent to Leu-106 in space, another conserved residue making direct hydrophobic contacts with the aromatic ring of OSB. Replacement of Y97 with a phenylalanine in ecMenB leads to an inactive enzyme (**Table 3.2**) that preserves the overall structure of the enzyme (**Figure 3.24**), and previously we demonstrated that the homologue of the second Tyr in mtMenB was essential (Y287) [109]. These observations strongly suggest that the two Tyr residues play a critical role in the MenB reaction.

A similar strategy is employed by hydroxycinnamoyl-CoA hydratase–lyase (HCHL), another crotonase superfamily member that uses two Tyr residues as a molecular "pincer" to recognize the substrate, initiate deprotonation and exert strain on the substrate. The two Tyr residues in HCHL are located on analogous loops to those found in MenB (**Figure 3.23**), one of which is also only observed when the analogous A-loop in HCHL becomes ordered [304].

Figure 3.22 The retro-Dieckmann, ring-opening reaction catalyzed by BadI

A-loop

```
M.tuberculosis   99  WAFCSGCDQRIRGR-SGYQYASG----------DTAD-------TVDVAR----------AGRLHILEVQRLIRFM
C.glutamicum     93  WAFCSGCDQRIRGR-SGYQYATE---------HARDDATADVSTVDIARTKVE------GGRLHILEVQRLIRTM
C.michiganensis  88  WAFCSGCDQRIRGR-DGYKYAEG---------ETAEG-------VDPAR---------AGRLHILEVQRLIRFM
S.oneidensis     86  YSFSSGCDQRIRGK-DGYKYEGA---------EEGKA-------DLARM---------GRLHILEVQRLIRFM
S.aurantiaca     99  WAFCSGCDQRIRGK-DGYKYEG----------EEGKA-------DPARL---------GRLHILEVQRLIRFL
H.marismortui   120  WAFCAGGDQSIRGE-SGYEYSESPRDSDTASGEQSDP-------RETSDGVEDDAPENVGRLHILEVQRQIRHI
S.ruber          84  WAFSSGCDQQIRGQ-DGYQYED----------EAGDP-------DATRP---------GRLHILEVQRLIRFM
G.forsetii       88  WAFCSGCDQKARGH-QGYVGEDG----------------------------------YHRLNILEVQRLIRFM
T.elongatus      74  YAFCAGGDQSIRGE-AGYLDEQG----------------------------------VARLNVLDLQRQIRSL
B.anthracis      68  RAFCSGCDQKVRGH-GGYVGDDQ----------------------------------IPRLNVLDLQRLIRAI
S.aureus         69  LAFCSGCDQKKRGH-GGYVGEDQ----------------------------------IPRLNVLDLQRLIRII
E.coli           80  KAFCSGCDQKVRGDYGGYKDDSG----------------------------------VHHLNVLDFQRQIRTC
S.typhimurium    80  KAFCAGGDQKVRGDYGGYQDDSG----------------------------------VHHLNVLDFQRQIRTC
C.reinhardtii    71  LAFCSGCDQSVRGK-GGYVGEDG----------------------------------IARLNVLDLQVQIRRL
rpBadI           60  RAFCTGGDQSTHDG--NYDGRGT----------------------------------VGLPMEELHTAIRDV
ECH:1MJ3         92  KAFAAGADIKEMQNRTFQDCYSG----------------------------------KFLSHWDHITRI
HCHL:2VSU        64  EAWTAGMDLKEYFREVDAGPEILQEK-------------------------------IRREASQWQWKLLRMY
MMCD:1EF8        60  KVFSAGHDIHELPS-GGRDPLSY---------------------------------DDPLRQITRMIQKF
```

B-loop   C-loop

```
M.tuberculosis  157  WAAGCGHSLHVVCDLTLASREYARFKQTDADVGSFDGGYGSAYLARQV   272  LVGQQLFAG--EATRLAYM-TD
C.glutamicum    162  WAAGCGHSLHVVCDLTIASRQEARFKQTDADVGSFDGGYGSAYLARQV   277  LMGQQVFAG--EATRLAYM-TD
C.michiganensis 146  WAAGCGHSLHVVCDLTIASAEHGRFKQTDADVGSFDGGYGSAYFARQV   261  MVGQQVFAG--EATRLAYG-TD
S.oneidensis    143  WAVGCGHSLHVVCDLTLASKEHAIFKQTDPDVASFDSGYGSAYLAKMI   258  MVGQQLFAG--EATRLAYA-SA
S.aurantiaca    155  WAVGCGHSLHVVCDMTLASREHAVFKQTDPDVASFDSGYGSALLARQI   270  LVGQQLFAG--EATRLAYG-TE
H.marismortui   196  WAVGCGHSLHVVCDLTLASEEHAKFLQTDPDVGSFDGGFGSAYLARQV   311  MVGQQVFAG--EATRLAYM-TD
S.ruber         140  WAVGCGHSLHVVCDLTLASTEHATFKQTDPDVASFDGSALLAQQV      255  LVGQQVFAG--EATRLAYM-TD
G.forsetii      136  WAVGCGHSLHVVCDLTLASKEHAIFKQTDADVTSFDAGYGSAYLAKMV   251  MVGQQVFAG--EVTRLAYM-TD
T.elongatus     122  YAIGCGHVLHLVCDLTIAA-ENAIFGQTGPKVGSFDAGFGASYLARIV   236  QAGLQELAG--NATLLYYL-TQ
B.anthracis     116  YAIGCGHVLHIVCDLTIAA-DNAVFGQTGPKVGSFDGGYGAGYLARMV   230  LAGIQQLAG--DATLLYYT-TD
S.aureus        117  YAVGCGHVLNVVCDLTIAA-DNAIFGQTGPKVGSFDAGYGSGYLARIV   231  LAGLQQMAG--DATLLYYT-TD
E.coli          129  YSIGCGHVLHMMCDLTIAA-DNAIFGQTGPKVGSFDGGWGASYMARIV   243  QAGLQELAG--NATMLFYM-TE
S.typhimurium   129  YSIGCGHVLHMMCDLTIAA-ENAIFGQTGPKVGSFDGGWGASYMARIV   243  QAGLQELAG--NATMLFYM-TE
C.reinhardtii   119  YAVGCGHILHMVCDLTIAA-DNAIFGQTGPKVGSFDAGYGSTHMARLV   233  QAGIQELGG--NATMLFYH-SE
rpBadI          106  YAIGCGNVLATICDLTICS-EKAIFGQVGPKMGSVDPGYGTAFLARVV   220  QAGIAGMGM--YALKLYYD-TD
ECH:1MJ3        137  YALGCGCELAMMCDIIYAG-EKAQFGQPEILLGTIPGAGGTQRLTRAV   251  TLTEGNKLE--KKLFYSTFATD
HCHL:2VSU       116  WCFGGFAPLVACDLAICA-DEATFGLSEINYGIPPGNLVSKAMADTV    230  TWEQNEDYLYAKLDQSRLLDTE
MMCD:1EF8       106  SVWGCAFEMIMSSDLIIAA-STSTFSMTPVNLGVPYNLVGIHNLTRDA   220  HTMNSDEFERIQGMRRAVYDSE
```

Figure 3.23 Sequence alignment of MenB enzymes and other crotonase superfamily members

MenB sequences were selected from a wide range of organisms including γ-proteobacteria, δ-proteobacteria, firmicutes, actinobacteria, bacteroidetes, cyanobacteria, plants and archaea. The alignment was performed using ClustalW and the figure generated using Jalview [313,314]. Alignment of HCHL in the A-loop region was edited according to the alignment of the crystal structures. The enolate oxyanion hole is shown in magenta, the tetrahedral oxyanion hole in purple, OSB-binding residues in green, D185 of mtMenB in blue, G156 of ecMenB in orange, known or possible catalytic residues in other crotonase superfamily members are shown in red [304,311,315].

133

Figure 3.24 Overall structure of Y97F ecMenB

The preliminary structure of ecMenB Y97F mutant has been overlaid with that of wild-type ecMenB. The RMSD by aligning all Cα atoms in one monomer is 0.33 Å. The R_{work} and R_{free} values for the mutant structure are 0.153 and 0.203, respectively. The overall fold is maintained but there is no detectable activity, suggesting that the substrate is binding in the active site in some non-productive manner. The A-loop is disordered. If the substrate could bind exactly the same way as observed in the wild-type structure, then we should expect to see at least some low-level activity even though the Y97 hydroxyl has been deleted.

Structure of the mtMenB:OSB-NCoA complex

We also attempted to co-crystallize OSB-NCoA with mtMenB which has an A-loop that is 9 residues longer than in ecMenB. The resulting structure is in space group P6$_1$22 and contains three MenB monomers in the asymmetric unit (**Table 3.1**). The same hexameric assembly observed in existing MenB structures can be generated by a symmetry operation, and the RMSD by aligning the Cα atoms in the three monomers with the equivalent monomers in the hexamer of mtMenB (1Q51) is 0.376 Å. However, in contrast to ecMenB, the A-loop is still disordered in all three monomers. Density for OSB-NCoA can be observed but is weak in chain C possibly due to disorder or low occupancy. In chain B, the density for the CoA portion is improved and there is more residual density for the acyl portion, but the conformation for OSB cannot be determined unambiguously (**Figure 3.25**). There is little evidence for the presence of OSB-NCoA in chain A. Instead, the CoA binding site is occupied by the N-terminus of a

monomer from a neighboring MenB hexamer in the crystal lattice that prevents the ligand from binding (**Figure 3.26**), and hence full occupancy cannot be achieved as in the ecMenB structure. Although the longer A-loop might be responsible for the disorder in the mtMenB structure, it is also plausible that the A-loop only adopts a productive conformation when all six active sites are occupied with the substrate. Such a mechanism might exist to ensure that the MenB hexamer utilizes all its active sites efficiently in catalysis.
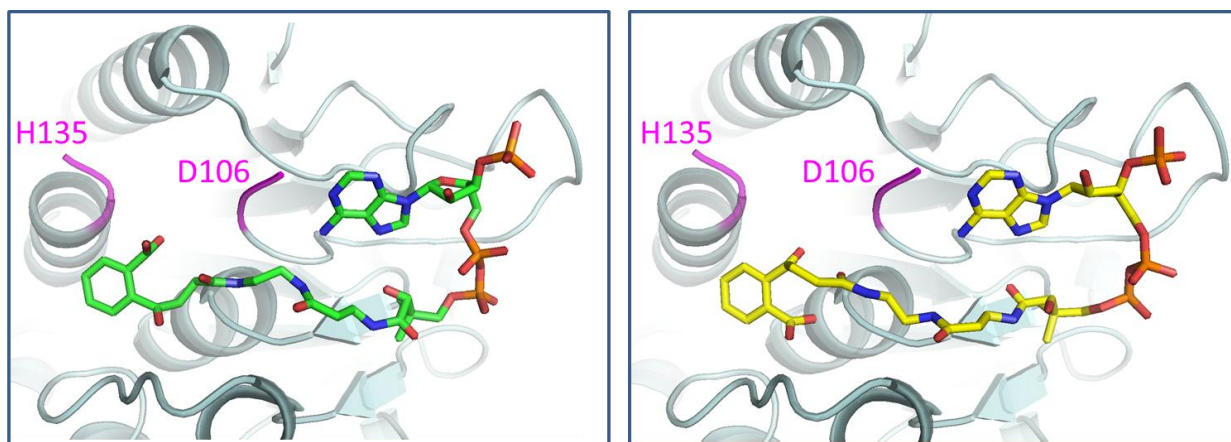
Figure 3.25 Disorder in binding OSB in the mtMenB structure
The acyl terminal of OSB-NCoA might bind in the active site in more than one way that leads to disorder
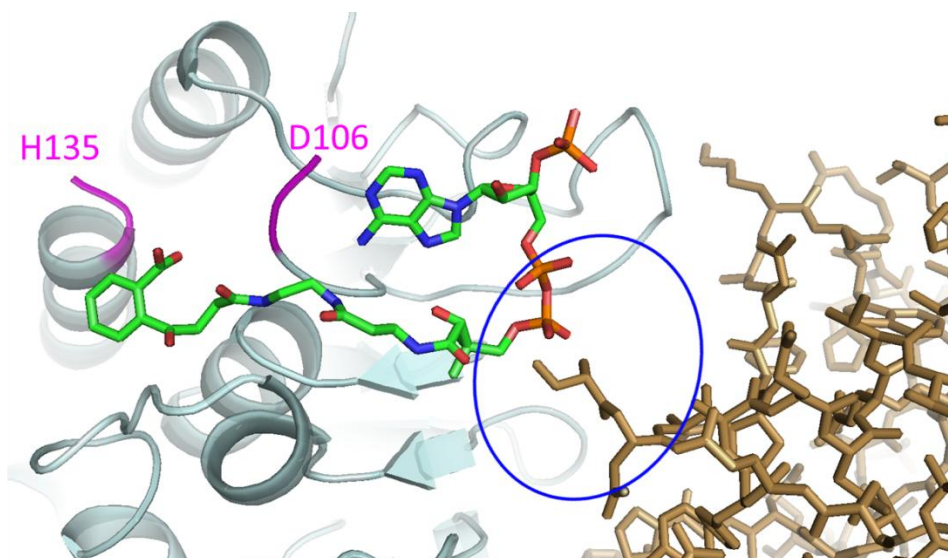
Figure 3.26 Interference of CoA binding site by the neighboring mtMenB hexamer in the crystal
The ligand is superimposed to the binding site that is interrupted by a neighboring hexamer.

135

Novel mtMenB hexameric assembly

In order to improve the occupancy of OSB-NCoA in mtMenB we attempted to identify a different crystal form in which the N-terminal tail from one monomer did not block the CoA binding site in an adjacent hexamer. We subsequently found a rhombohedral crystal (**Figure 3.10**) with high diffraction quality, which belongs to space group R3 and contains two monomers in the asymmetric unit. The RMSDs between each monomer and that in 1Q51 are 0.627 Å and 0.364 Å over 211 and 197 Cα atoms, respectively. However, upon symmetry operation the asymmetric unit generates a "staggered" hexamer in striking contrast to the eclipsed configuration in all other known MenB structures (**Figure 3.12**). In this new crystal form, the two trimers in the hexamer are rotated 60° with respect to each other along their three-fold axis (**Figure 3.27**). Consistent with this change in quaternary structure, there is local disruption to the tertiary structure. While the A-loop is disordered, there is also significant disorder in the C-terminal region. In chain A, the C-loop corresponding to residues 273-314 is disordered, and the six preceding residues fold in a different direction resulting in a shift of 16 Å for the Cα of L272 compared to its position in 1Q51, suggesting a large-scale dislocation of the C-loop. In addition, the B loop, which is comprised of residues 183-199 and also forms part of the active site, displays a different conformation in chain A so that D187 travels 9 Å from the protein surface into the oxyanion hole (**Figure 3.28**). The flexibility of these loops is further reflected in chain B where the B-loop, the C-loop and the preceding helix (residues 255-314) give little electron density. As a result, the normal active site seen in other known structures is severely disrupted, and OSB-NCoA is not found in the structure.
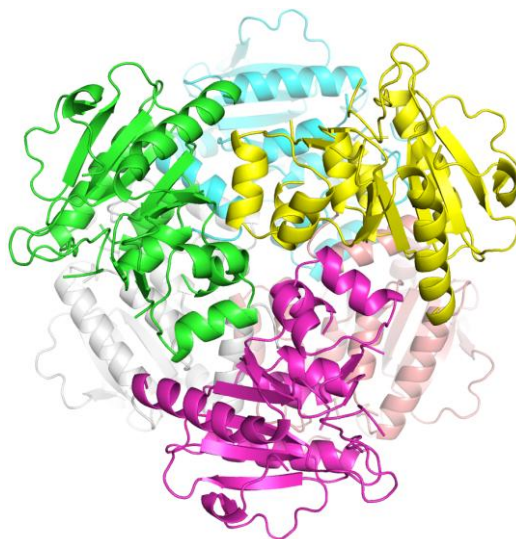
Figure 3.27 The altered mtMenB hexameric assembly
The "staggered" MenB hexameric assembly in contrast to the eclipsed assembly (**Figure 3.12**) normally observed in MenB enzymes.
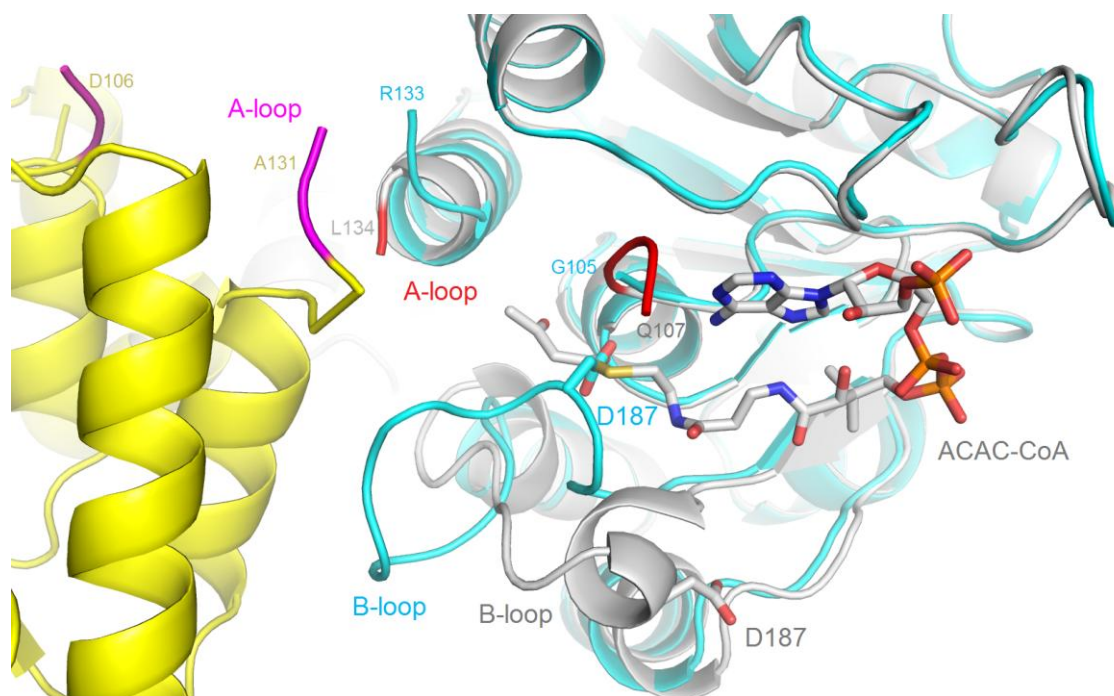
Figure 3.28 Change in the active site and subunit interactions in the altered hexamer
New contacts at the trimer-trimer interface and the change in conformation of the B-loop that
leads to occlusion of the oxyanion hole. The cyan and yellow monomers from **Figure 3.27** are
displayed. The ends of the disordered A-loop are labeled. Residues 106-110 in the cyan
monomer although ordered are not shown for clarity. The structure of mtMenB in complex with
acetoacetyl-CoA shown in grey is superimposed with the ends of the disordered A-loop shown in
red. The A-loop from yellow monomer comes into close proximity to the A- and B-loop from
the cyan monomer. D187 from the B-loop binds in the oxyanion hole, preventing binding of the
acyl-CoA ligand. The C-loop does not cover the same binding site as in **Figure 3.15**, **Figure
3.20** or **Figure 3.29.**

The disorder of the A-loop is a common observation among crotonase superfamily
members including enoyl-CoA hydratase (ECH) [302,316,317], methylmalonyl CoA decarboxylase
(MMCD) [275], Δ^3- Δ^2-enoyl-CoA isomerase (ECI) [318], and hydroxycinnamoyl-CoA hydratase-lyase
(HCHL) [304], while in MenB the flexibility of the B-loop and C-loop has also been observed [109,278].
These loops together form the interface between the two trimers in the hexamer that buries
more than 3450 Å, or 24% of the total surface area, and which is likely the reason for the
stability of the mtMenB structure [278]. However, the novel hexameric structure reported here
demonstrates that the flexibility of these loops can actually lead to a trimer-trimer rotation
which establishes new contacts between two A-loops and between the B-loop and A-loop from
the two trimers (**Figure 3.28**). While the functional relevance of this rotation is unclear at
present, the observation is consistent with the versatility of these loops in the controlling

function within the superfamily [319]. These loops comprise the active site (**Figure 3.29**), and share little similarity amongst superfamily members. The C-loop is known to be involved in domain swapping that results in variation of the hexameric assembly [109,274,275]. However, variation of the hexameric assembly within the same enzyme is less well known. Two forms of hexamers have been observed for the yeast peroxisomal Δ^3- Δ^2-enoyl-CoA isomerase (Eci1p) where it was speculated that the hexamer might dissociate into trimers and associate with a different partner for entry into peroxisomes [318]. While lower oligomeric states have been reported for saMenB [279], there is little evidence for the presence of trimeric mtMenB. In solution mtMenB is found exclusively as a hexamer [278], and in the present rhombohedral crystal form the mtMenB monomers clearly form discrete hexameric units. However, only 10% of the total surface area is buried upon trimer-trimer association in contrast to the 24% that is buried in the normal MenB hexamer. Thus the staggered hexamer is likely less stable than the hexameric structure usually adopted by MenB.
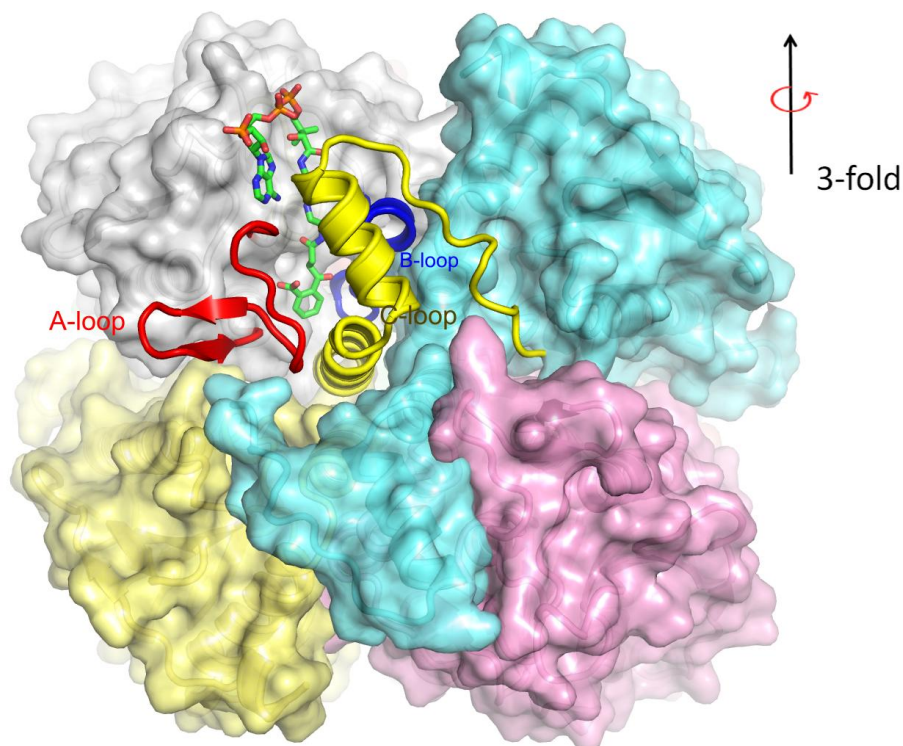
Figure 3.29   Positions of the variable A-loop, B-loop and C-loop in the crotonase superfamily
The three loops highlighted using the ecMenB structure correspond to the flexible regions found in MenB and the highly variable regions in the crotonase superfamily. The hexamer is viewed from the side and one set of A-, B- and C-loops is displayed in red, deep blue and yellow.

Active sites of mtMenB and ecMenB

Although the A-loop and acyl portion of the ligand are not observed in mtMenB, conservation of active site residues identified from the ecMenB structure, including those involved in OSB binding and the two oxyanion holes (**Figure 3.23**), indicates that mtMenB is capable of using the same catalytic mechanism as ecMenB. The non-conserved regions, including the 9 extra residues in the A-loop of mtMenB, are expected to be outside the active site while catalysis takes place. One exception, though, is D185 of mtMenB which is located next to the OSB succinyl group. This residue is replaced by G156 in ecMenB, and a water/chloride ion replaces the D185 side chain (**Figure 3.20**). Throughout the MenB family, this residue is either an Asp or a Gly (**Figure 3.23**). Although G156 or the water molecule does not appear to play a major role during α-deprotonation or stabilization of the two oxyanion intermediates based on our ecMenB structure, a glutamate or water in a similar position has been proposed to perform acid/base catalysis in a number of crotonase superfamily members including ECH (**Figure 3.30**) [315], dienoyl-CoA isomerase (DCI) (**Figure 3.31**) [320], ECI (**Figure 3.32**, **Figure 3.33**) [321], HCHL, [304] and DpgC [295]. The bicarbonate dependence in the MenB activities of *E. coli*, *S. aureus* and *B. subtilis* and the presence of bicarbonate in the crystal structure of stMenB also has led to a proposal that D185 is used for α-deprotonation in mtMenB and that a bicarbonate cofactor performs the same function when the Asp is replaced by a Gly as in ecMenB, saMenB and bsMenB [277]. Our structural analysis of ecMenB, however, does not support this role for bicarbonate, and bicarbonate is also not observed in the structures of saMenB and gkMenB [279,281]. The altered function of D185 in mtMenB compared to other members in the crotonase superfamily might be a result of an intrinsically higher pKa for the Cα proton in its substrate, OSB-CoA (**Figure 3.34**).
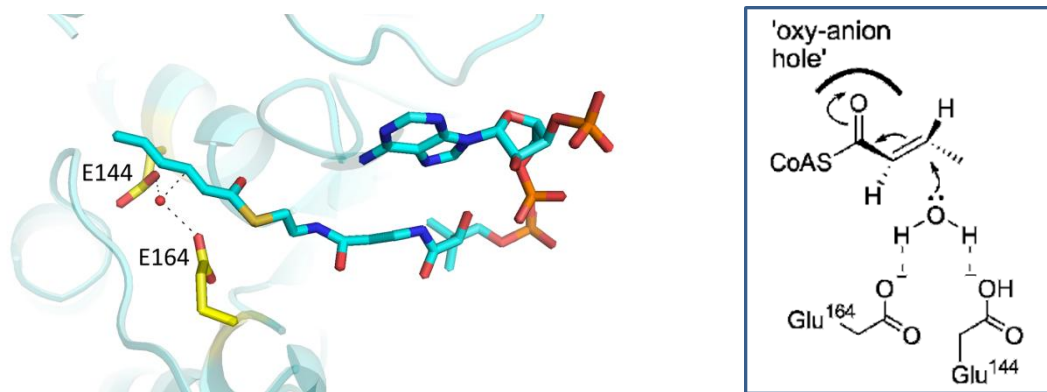
Figure 3.30 Mechanism of the ECH-catalysed reaction
PDB structure 1MJ3.
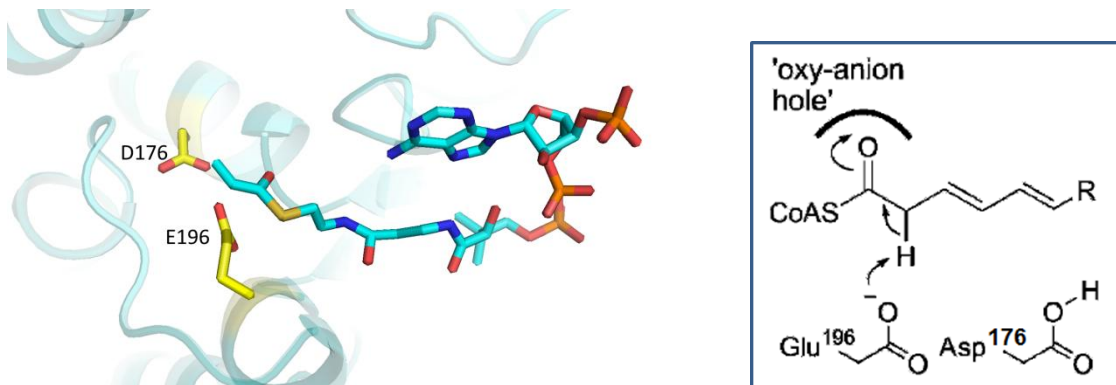
Figure 3.31 Mechanism of the DCI-catalyzed reaction
PDB structure 1DCI with a superimposed ligand.
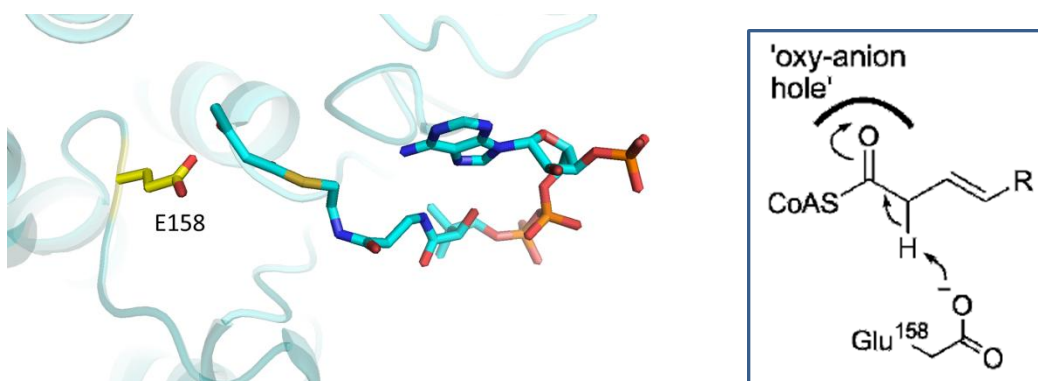
Figure 3.32 Mechnism of the yeast ECI-catalyzed reaction
PDB structure 1PJH with a superimposed ligand.
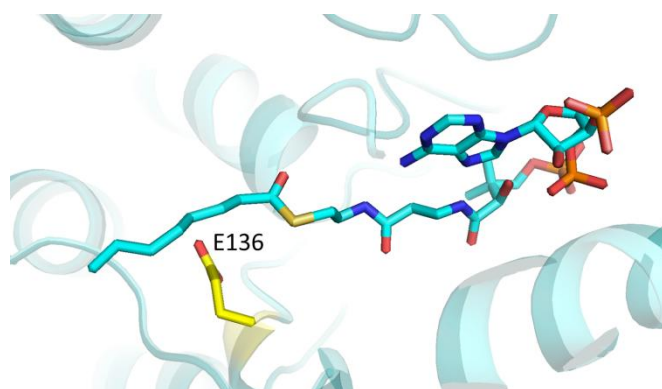
Figure 3.33 Mechanism of the human ECI-catalyzed reaction
PDB structure 1SG4.
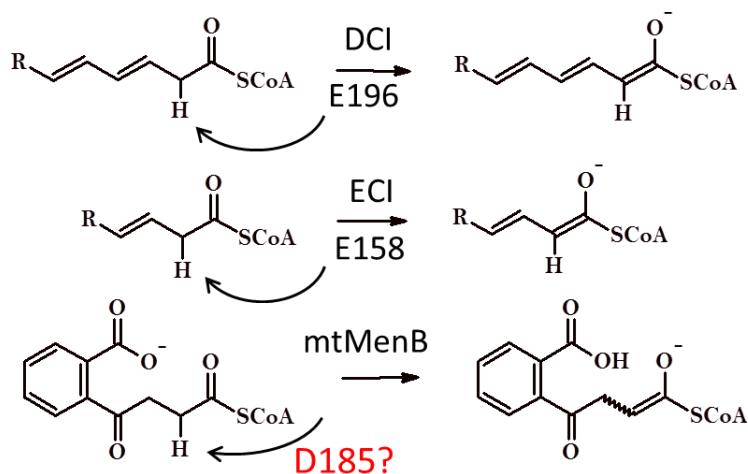
Figure 3.34 Comparison of the Cα pK$_a$ of the DCI, ECI and MenB substrates

Nevertheless, both the ordered water molecule in ecMenB and D185 in mtMenB are positioned so that they could aid in the recognition of the OSB C4 carbonyl and assist in the final tautomerization step that leads to DHNA-CoA. D185 in mtMenB clearly plays a critical role in the reaction since k$_{cat}$ for the D185E and D185N mtMenB mutants is reduced 200 and 2000-fold, respectively (**Table 3.2**). In addition, neither the D185G mtMenB mutant nor the G156D ecMenB mutant has detectable activity, indicating that the roles of these residues in the two enzymes cannot be simply reversed. Detailed comparison of the two active sites reveals that this result can be attributed to the disruption of the respective hydrogen bonding network optimized for each enzyme when the mutation is introduced (**Figure 3.35** and **Figure 3.36**). The network formed by water molecules W1-W5 is conserved while that formed by W6-W8 is not conserved.
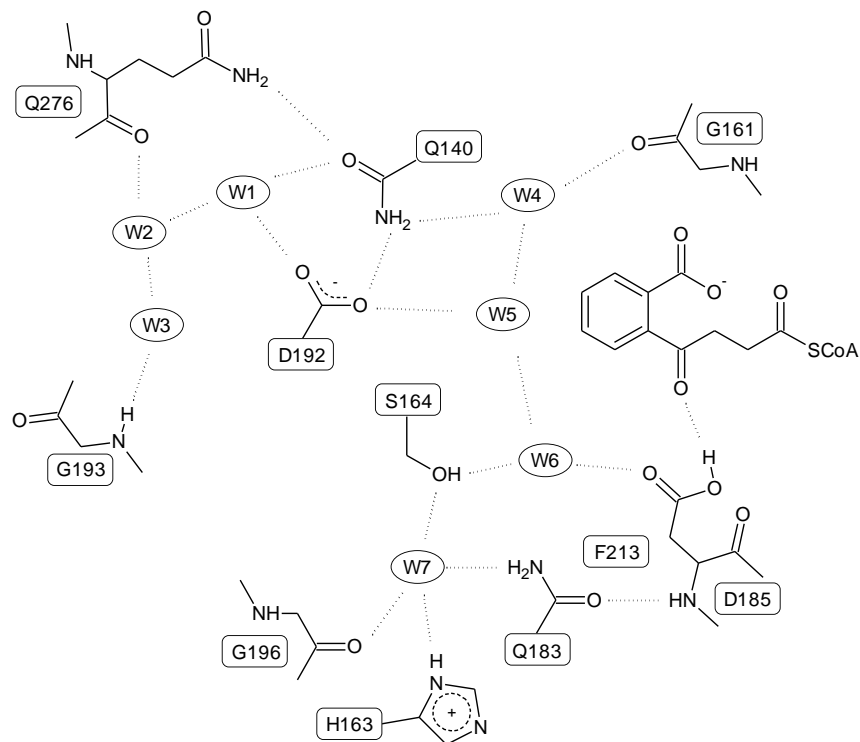
Figure 3.35 Network of interactions involving D185 in mtMenB
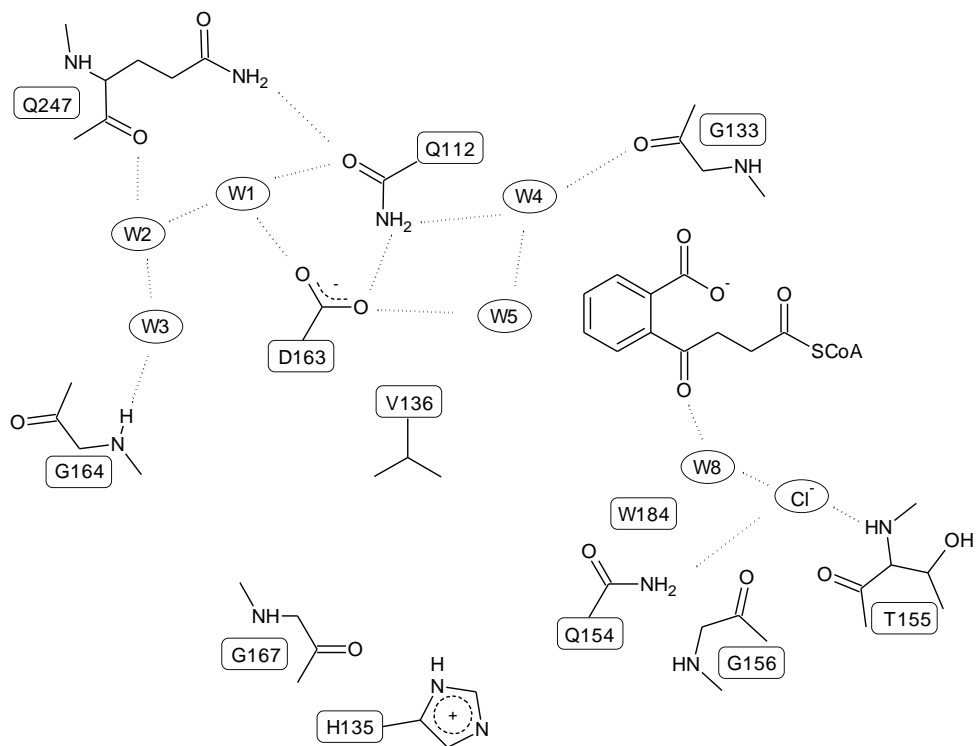
Figure 3.36 Network of interactions involving W8 and Cl⁻ in ecMenB

The anionic binding site in ecMenB

Overlay of ecMenB (green, cyan, magenta), gkMenB (purple) and mtMenB (yellow) shows different ligands in the anionic binding site occupied by D185 of mtMenB (**Figure 3.37**, top). Water (W8), chloride (green), malonate (cyan) and formate (magenta) in crystallization buffers can be found in different ecMenB structures. Two water molecules (grey crosses) are modeled in gkMenB that coincide with W8 and chloride in one of the ecMenB structures. The same spots are occupied by two of the oxygen atoms from malonate. When it is not possible to occupy the two spots simultaneously, the chloride site is preferred, as formate binding reveals. Bicarbonate from the stMenB structure also occupies the chloride site (**Figure 3.37**, bottom). These different ligands mark the two spots most consistently occupied by water or ions in ecMenB, gkMenB as well as in saMenB and stMenB and of most physiological relevance. In mtMenB, one oxygen atom on the side chain of D185 aligns close to the W8 site.
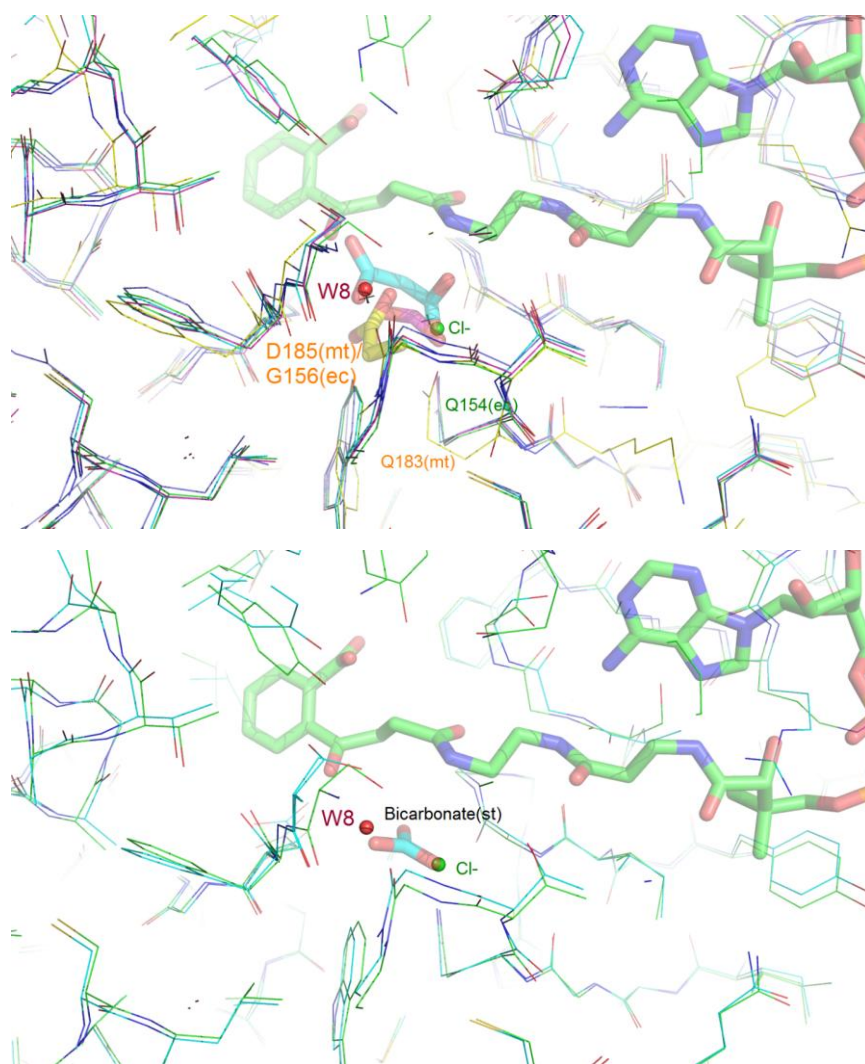
Figure 3.37 The anionic binding site in ecMenB

Substrate analogues as mechanistic probes

The importance of the OSB carboxylate in α-proton abstraction is substantiated by the observation that the methyl ester analogue of OSB-CoA is inactive (**Figure 3.38**, **Table 3.2**). If the carbanion can be generated without intramolecular proton transfer to the OSB carboxylate, the OSB-CoA methyl ester is expected to be a good substrate for MenB since methanol and water have similar leaving group abilities. In addition, the K_d for the OSB-CoA methyl ester (11.5 ± 1.2 µM; data not shown) is similar to the K_m value for OSB-CoA (22.4 ± 2.1 µM), suggesting that the loss of activity for the methyl ester is not simply due to an inability to bind to MenB. In contrast, OCPB-CoA which lacks the C4 carbonyl is still able to undergo ring closure, albeit with lower efficiency than OSB-CoA, indicating that MenB retains the ability to abstract the substrate α-proton when the OSB carboxylate is not modified (**Table 3.2**). These results support our original mechanism for mtMenB in which the OSB carboxylate abstracts the C2 α-proton.
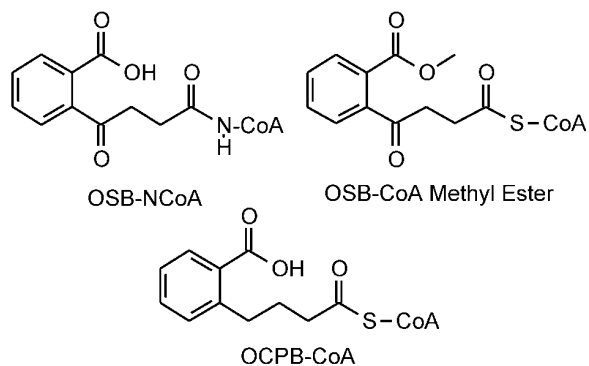
Figure 3.38 Mechanistic probes for the MenB reaction

Table 3.2 Kinetic Parameters for ecMenB and mtMenB

Enzyme	$k_{cat} \cdot 10^2 \, (s^{-1})$	$K_m \cdot 10^6 \, (M)$	$k_{cat}/K_m \cdot 10^{-4} \, (s^{-1} \cdot M^{-1})$
mtMenB			
wild-type	46.2 ± 1.5	22.4 ± 2.1	2.1 ± 0.2
D185E	0.23 ± 0.02	4.8 ± 0.3	0.048 ± 0.005
D185N	0.022 ± 0.002	3.1± 1.1	0.007 ± 0.003
S190A	0.17 ± 0.02	40.2 ± 4.5	0.004 ± 0.001
Y287F[a],		No activity	
OCPB-CoA	0.025 ± 0.003	106 ± 31	0.00024± 0.00007
OSB-CoA		No activity	
ecMenB			
wild-type	6.2 ± 0.2	25.9 ± 3.3	0.24 ± 0.03
Y97F		No activity	
G156D[b,c]		No activity	

[a]Up to 5 µM enzyme [109].
[b]Up to 3 µM enzyme.
[c]Jiang et al. [277] reported similar results.

The roles of D163 and S161

ECH contains two active site glutamates, E144 and E164 (**Figure 3.30**) [315]. While E164 is the structural homologue of D185/G156, a second Asp is present in the ecMenB (D163) that is conserved amongst MenB enzymes (D192 in mtMenB) and that is in a similar location in the active site as E144 in ECH (**Figure 3.39**, **Figure 3.30**). While we have previously shown that the D192N mtMenB mutant was inactive [109], our structural data indicates that D192/D163 points away from the substrate and is involved in a conserved hydrogen bond network so that it cannot directly participate in acid/base catalysis (**Figure 3.35**, **Figure 3.36**, **Figure 3.40**). A similar example is found for MMCD, where E113 is hydrogen-bonded to an arginine and points away from the substrate, preventing it from direct involvement in catalysis [275]. Here D192 is hydrogen-bonded to Q140 at the C-terminal end of the A-loop and is apparently important in maintaining the shape of the substrate binding pocket.
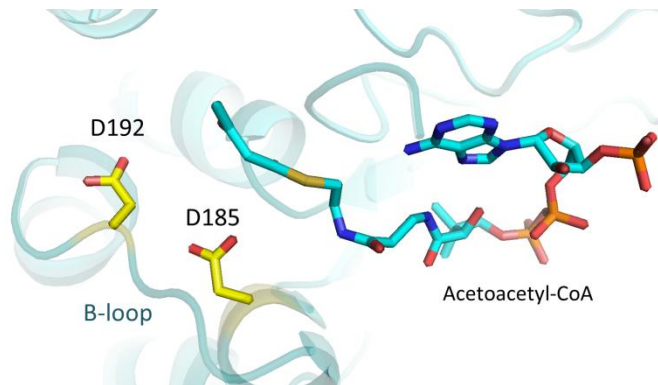
Figure 3.39 Active site Glu residues in mtMenB
PDB structure 1Q51.
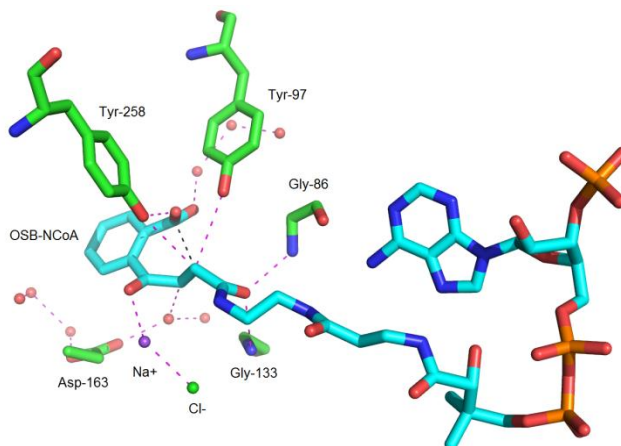
Figure 3.40 The conserved binding network involving D163 and water in ecMenB

Finally, previous studies also demonstrated that S190 in mtMenB was important but not essential for catalysis (**Table 3.2**). The structures of mtMenB and ecMenB reveal that the side chain of S190 (S161 in ecMenB) can interact with Y287, D185, the sulfur atom of the thioester moiety and the oxygen atom of the C4 carbonyl group. Its location in the active site suggests that this residue is involved in the tautomerization step.

Mechanism of the MenB-catalyzed reaction

The mechanism of the MenB-catalyzed reaction based on these new observations can be described as follows (**Figure 3.41**). Two active site Tyr residues from the A-loop and C-loop play a central role in orientating the OSB C7 carboxylate group so that it can abstract the pro-2S proton in either ecMenB or mtMenB. The resulting enolate oxyanion is stabilized in the oxyanion hole formed by two backbone amides, and the C7 tetrahedral intermediate generated

146

by C-C formation is stabilized by the second oxyanion hole formed by the two Tyr residues. The intramolecular proton transfer leads to protonation of the C7 carboxylate, which increases the electrophilicity of this group while the surrounding hydrogen bond network assists in elimination of water from the tetrahedral oxyanion. The two active site Asp residues in mtMenB or the Asp and water in ecMenB together with the active site serine are involved in correctly positioning the substrate for the reaction and are also likely involved in substrate tautomerization.
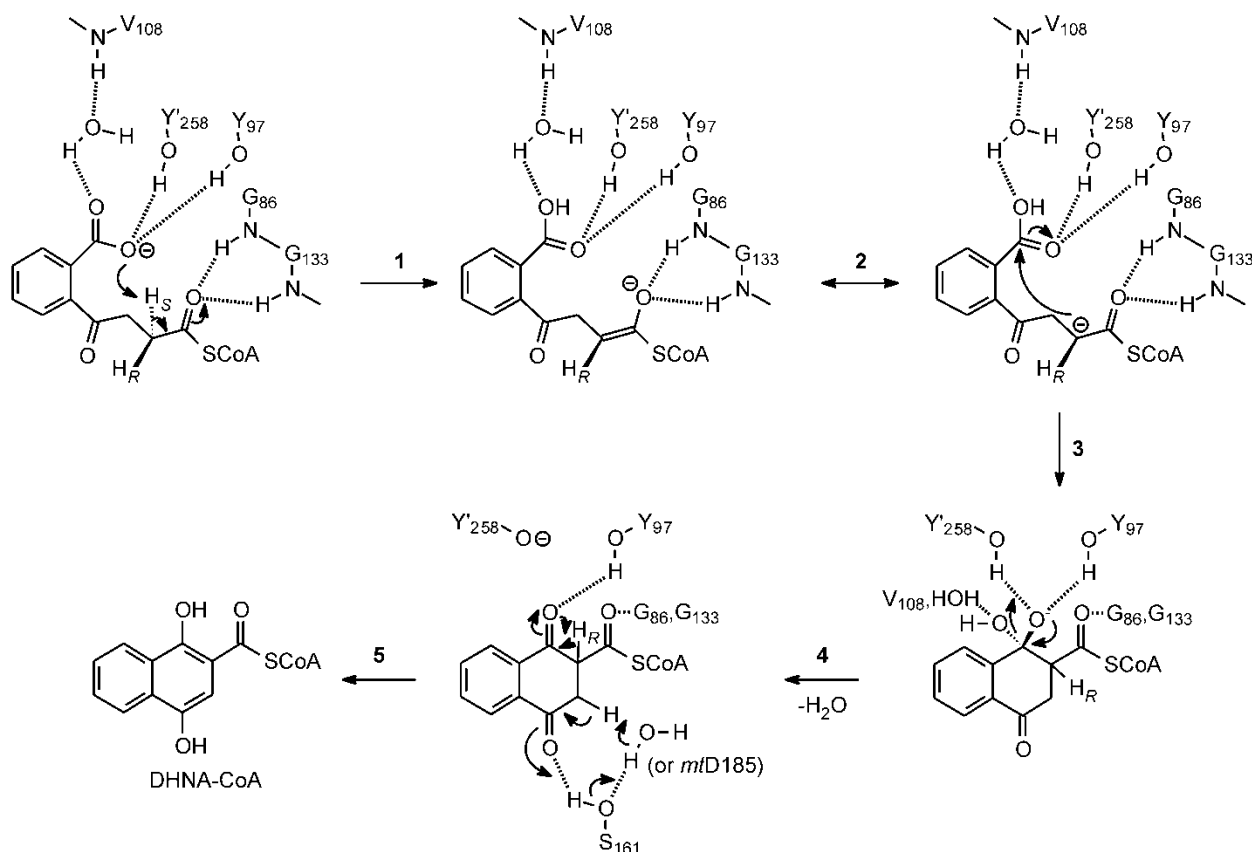
Figure 3.41 Proposed mechanism of the MenB catalyzed reaction
The OSB carboxylate is positioned to abstract the pro-2S proton, assisted by Y258 and Y97. Proton transfer results in a carbanion/enolate that is stabilized by the classic crotonase oxyanion hole formed by the backbone NH groups of G86 and G133. Subsequent attack of the carbanion on the protonated OSB carboxylic acid leads to a tetrahedral intermediate that is stabilized by hydrogen bonds to Y258 and Y97, which we propose form the second oxyanion hole required for Claisen condensations. This creates a stereo center at the α-carbon and it is not clear at present whether the stereochemistry here is the same as that observed in BadI [322]. Elimination of water yields a β-ketothioester which formally is the keto tautomer of the DHNA-CoA product. Subsequent tautomerization to the enol results in a thermodynamically favorable aromatization of the ring. In the Figure we show Y258 acting as the general acid, however adjacent water molecules and other nearby residues are also potential candidates especially if ring formation

(step 3) leads to a rearrangement of the active site. A network of hydrogen bonds that interact with the OSB-NCoA keto group is a candidate for assisting in the keto-enol tautomerism and include S161, ordered water molecule(s) and D163. In mycobacterial MenB enzymes the side chain of D185 (mtMenB numbering) is close to S161 and could also be involved in tautomerization.

Although the X-ray structure of ecMenB in complex with OSB-NCoA strongly suggests that the substrate pro-2S proton is transferred to C7, Igbavboa et al [323] previously reported that MenB catalyzed the exchange of the pro-2R OSB-CoA proton with solvent more rapidly than the pro-2S proton. Although we cannot rationalize this discrepancy, it is possible that other enzymes could be present in the cell extract used for the labeling experiments that might catalyze the exchange of the pro-2R proton with solvent, while the instability of OSB-CoA, which decomposes readily to spirodilactone in solution (**Figure 3.42**), could also complicate measurements of stereospecificity. In addition, it is also plausible that the interaction between the OSB carboxylate and the pro-2S proton reduces the propensity of this proton to exchange with solvent. Interestingly, however, the pro-2S stereochemistry agrees with that determined for BadI, which catalyzes a retro-Dieckmann, ring-opening reaction, that is essentially the reverse of the MenB reaction [322]. BadI is also a member of the crotonase superfamily and sequence analysis reveals that many active site residues including Y97 and Y258 are conserved between the two enzymes. If MenB and BadI share a common mechanism, then the pro-2S proton in the BadI product could be derived from the Tyr residues or from the carboxylic acid of the substrate which is generated following ring opening (**Figure 3.43**).

OSB-CoA   **Spirodilactone**

Figure 3.42 Thermal decomposition of OSB-CoA into OSB spirodilactone

Figure 3.43 Proposed mechanism of the retro-Dieckmann reaction catalyzed by BadI

3.4 Conclusion

We have successfully trapped the ecMenB active site in a catalytically relevant conformation using OSB-NCoA, a stable substrate analogue. This structure reveals the positions of all the catalytic residues for the first time, and shows the substrate poised for proton abstraction and carbon-carbon bond formation. Coupled with site-directed mutagenesis and studies with additional substrate analogues, the mechanistic studies reveal how this crotonase superfamily member has adapted to catalyze an intramolecular Claisen condensation reaction. Common or similar strategies are likely employed by other members that catalyze Claisen-like reactions, including BadI, 6-oxocamphor hydrolase and *Anabaena* β-diketone hydrolase [324,325]. The serendipitous finding of a novel MenB hexameric assembly also highlights the tight connection between evolution, catalysis and oligomerization within the crotonase scaffold.

Appendix

XK1

XK2
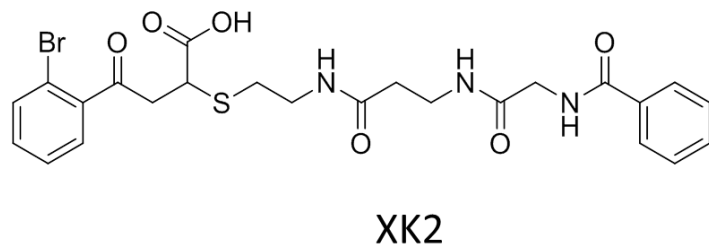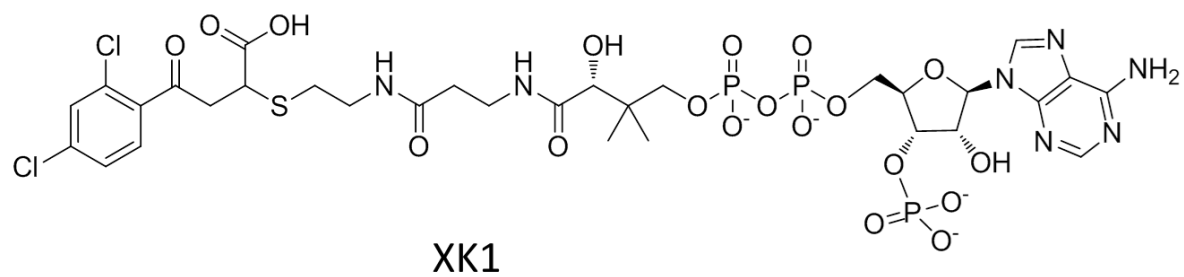
Figure 3.44 Structures of MenB inhibitors used for co-crystallization

Table 3.3 Summary of observed MenB crystal froms

ID	Protein	Ligand	Space Group	a	b	c	α	β	γ	Resolution
JT-soak3	*mt*MenB	OSB-NCoA	P $6_1$22	87.46	87.46	416.90	90.00	90.00	120.00	2.3 Å
JT-B6_2	*mt*MenB	OSB-NCoA	P$2_12_12_1$	87.04	139.17	157.98	90.00	90.00	90.00	2.8 Å
130D3_1	*mt*MenB	OSB-NCoA	P $6_1$22	87.13	87.13	414.84	90.00	90.00	120.00	2.3 Å
136A4_1	*mt*MenB	XK1	R 3	141	141	290	90	90	120	~4 Å
155B2_3	*mt*MenB	(XK1)	P 4_3	80.01	80.01	249.81	90.00	90.00	90.00	2.00 Å
214E10_2	*mt*MenB	-	R 3	131.85	131.85	71.11	90.00	90.00	120.00	1.95 Å
217B2_1	*mt*MenB	-	P222							2.4 Å
217B1_2	*mt*MenB	-	R3							1.8 Å
217C2_3	*mt*MenB	-	P$2_12_12_1$							2.4 Å
218B1_1	*mt*MenB	(OSB-NCoA)	R 3	132.23	132.23	71.14	90.00	90.00	120.00	1.7 Å
218B1_2	*mt*MenB									1.8 Å
218B1_5	*mt*MenB	OSB-NCoA	P $6_1$22	86.63	86.63	417.61	90.00	90.00	120.00	2.0 Å
218B1_7	*mt*MenB	OSB-NCoA	P $6_1$22							
218B3_2	*mt*MenB	OSB-NCoA	P$2_12_12_1$							2.6 Å
218B3_5	*mt*MenB	OSB-NCoA	C 2							3 Å
218A1_2	*mt*MenB	OSB-NCoA	P 2	90	140	142				
218A2_1	*mt*MenB	OSB-NCoA	P222/P3							Bad
218C3_2	*mt*MenB	OSB-NCoA	P222							
225D9_1	*mt*MenB	(XK2)	R 3	136.31	136.31	137.89	90.00	90.00	120.00	3.0 Å
225E9_1	*mt*MenB	(XK2)	P $6_1$22	86.51	86.51	417.58	90.00	90.00	120.00	1.9 Å

225E10_1	*mt*MenB	(XK2)	C 2	149.79	86.31	139.84	90.00	101.81	90.00	2.5 Å
225E10_3	*mt*MenB	(XK2)	R 3	132.28	132.28	71.24	90.00	90.00	120.00	1.8 Å
225E10_4	*mt*MenB	(XK2)	P$2_12_12_1$	85.52	140.13	146.46	90.00	90.00	90.00	3.2 Å
225E10_5	*mt*MenB	(XK2)	P $6_1$22	85.79	85.79	417.36	90.00	90.00	120.00	3.3 Å
225H10_3	*mt*MenB	(XK2)	R 3	136.02	136.02	138.29	90.00	90.00	120.00	2.7 Å
194G11_1	*ec*MenB	-	P$2_12_12_1$	76.51	134.00	153.36	90.00	90.00	90.00	2.0 Å
195G9_1	*ec*MenB	OSB-NCoA	P$2_12_1$2	140.49	141.79	89.12	90.00	90.00	90.00	2.0 Å
196D12_1	*ec*MenB	XK1	P 1	75.57	80.57	80.56	61.08	68.63	83.66	2.0 Å
240A12_2	*ec*MenB Y97F	CoA, succinate	P$2_12_12_1$	73.86	133.03	153.30	90.00	90.00	90.00	2.0 Å
242F8_1	*ec*MenB Y97F	(3-BPA+CoA)	P$2_12_12_1$	75.20	134.28	153.88	90.00	90.00	90.00	2.3 Å
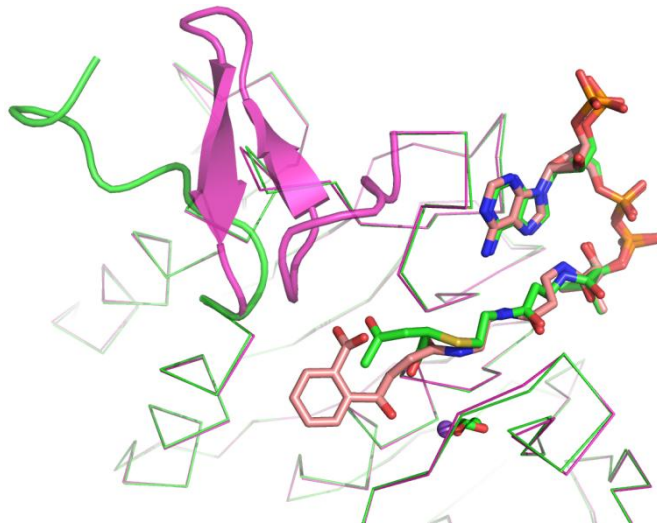
Figure 3.45 Conformational change of ecMenB upon inhibition
The A-loop opens when XK1 is bound. Green: ecMenB in complex with XK1; magenta: ecMenB in complex with OSB-NCoA.

Chapter 4 Characterization of the putative anthranilate synthase in *Mycobacterium tuberculosis*

4.1 Introduction

Anthranilate synthase

Anthranilate synthases (ASs) catalyze the conversion of chorismate into anthranilate which then leads to the biosynthesis of tryptophan in microorganisms, fungi and plants (**Figure 4.1**) [326]. Two nonidentical subunits are found in anthranilate synthases: AS-I, usually called TrpE in bacteria, catalyzes anthranilate formation using ammonia (AS-NH$_3$ activity), and contains a tryptophan binding site for feedback inhibition. AS-II, or TrpG in bacteria, is a glutamine amidotransferase that supplies the NH$_3$ from glutamine *in vivo* for the TrpE active site (AS-Gln activity) (**Figure 4.2**). The two subunits are usually found associated in the form of α$_2$β$_2$ heterotetramer or αβ heterodimer (**Figure 4.3**) [327-329].

The reaction catalyzed by AS-I is composed of two steps (**Figure 4.4**). The first is a reversible S$_N$2" 1,5-substitution by amine addition at C2 of chorismate and elimination of the hydroxyl group at C4 to give 2-amino-2-deoxyioschorismate (ADIC). The subsequent irreversible step of pyruvate elimination yields anthranilate (ADIC lyase activity). Evidence of the intermediate comes from the conversion of synthetic ADIC to anthranilate by the enzyme and its isolation from the reaction catalyzed by the H398M mutant of AS from *Salmonella typhimurium* [330-332]. Normally, ADIC remains bound to the wild-type enzyme and does not accumulate [333].
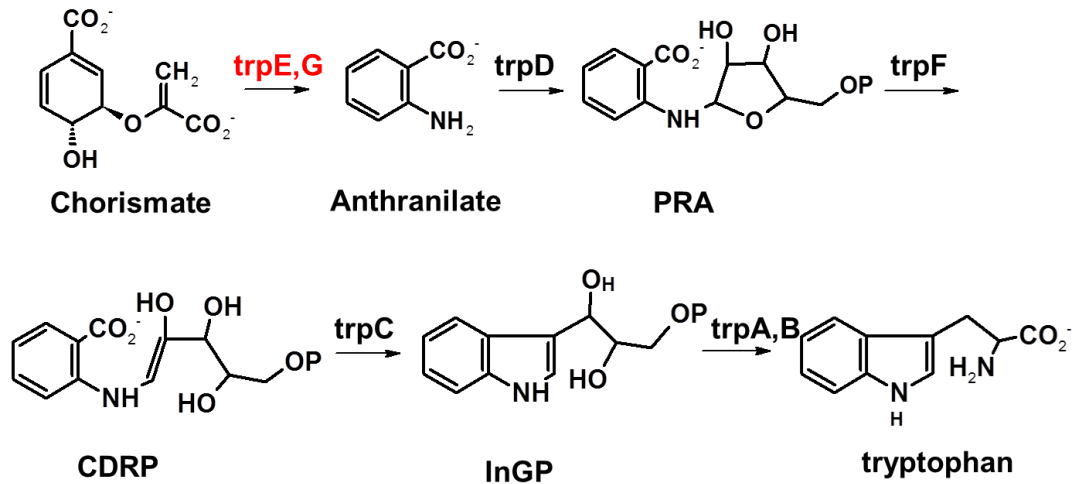
Figure 4.1 Biosynthesis of tryptophan from chorisimate
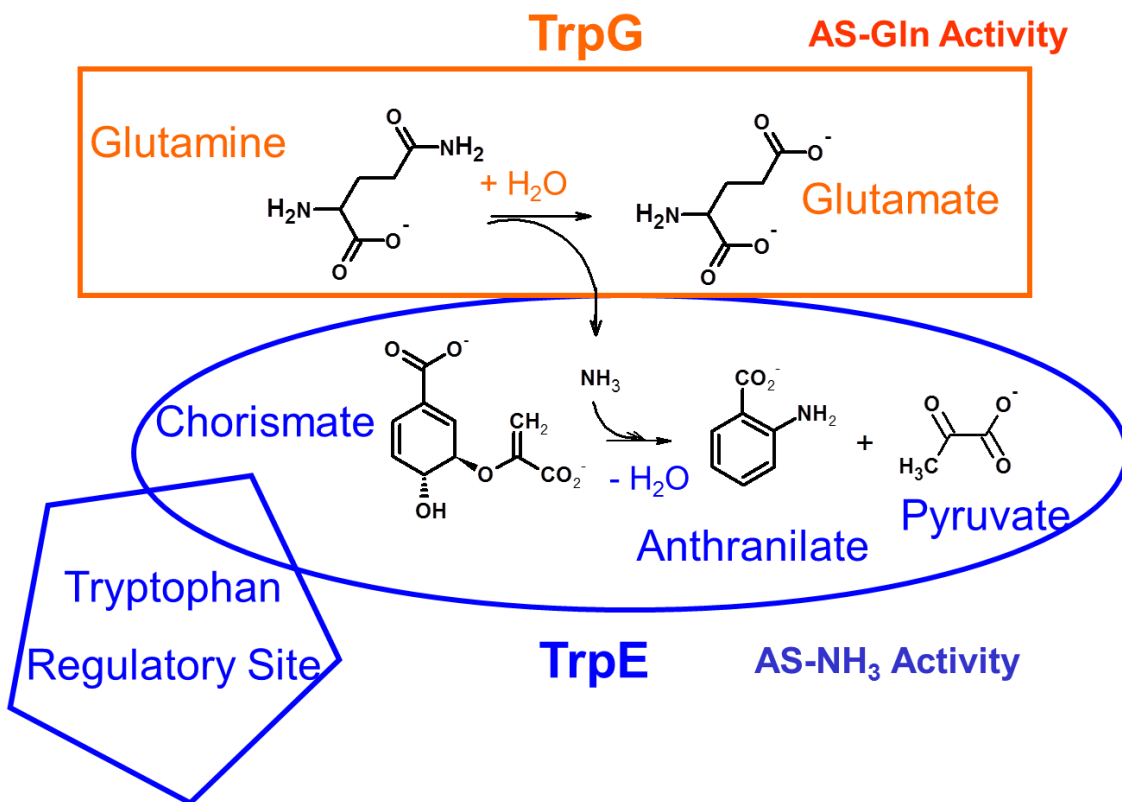
Figure 4.2 Reactions catalyzed by subunits of the anthranilate synthase
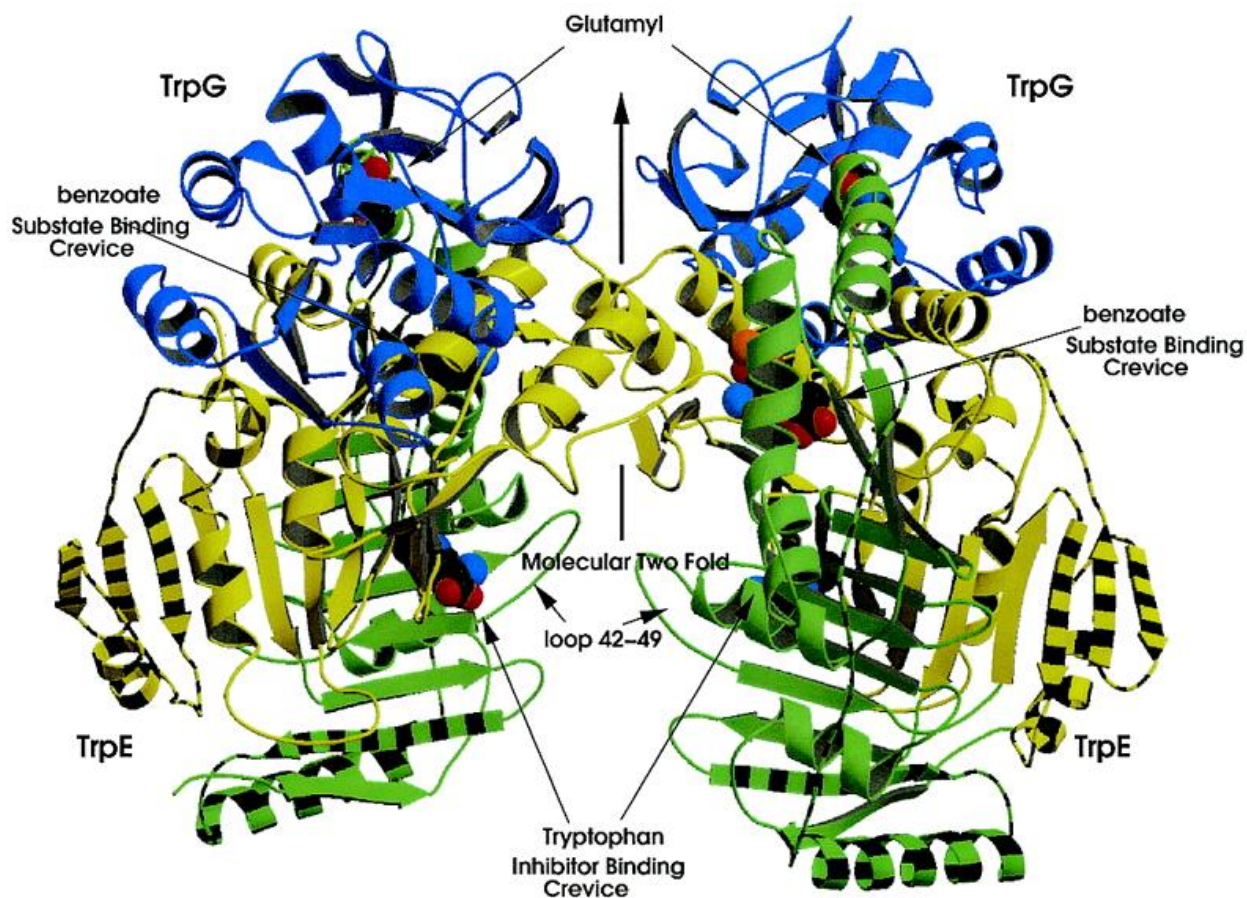
Figure 4.3 Structure of anthranilate synthase from *Serratia marcescens*
It is a heterotetramer [327].
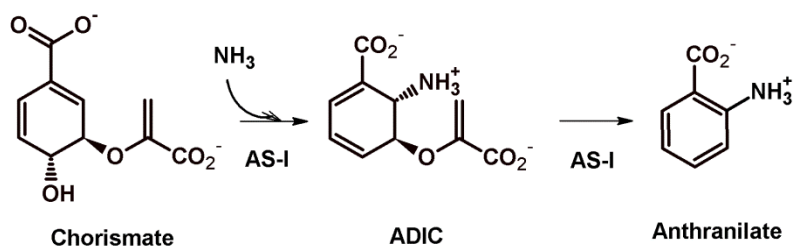
Figure 4.4 Steps catalyzed by AS-I

Homologues in the chorismate-utilizing enzyme family

The hydroxylation analogue of this AS-I catalyzed amination reaction is performed by the salicylate synthase (SS). Close relatives in the chorismate-binding enzyme family include

isochorismate synthase (ICS), 2-amino-2-deoxyisochorismate synthase (ADICS) and 4-amino-4-deoxychorismate synthase (ADCS), which catalyzes hydroxylation and amination without eliminating pyruvate. In the ADCS reaction, the nucleophile that attacks C2 of chorismate is the є-amino group of a lysine residue. The covalent intermediate is then converted to ADC by the reverse S_N2'' addition of ammonia to C4 (**Figure 4.5**) [131,134,334]. Like ASs, the ammonia required for ACDS is generated from glutamine by a glutamine amidotransferase *in vivo*. These enzymes share a similar catalytic mechanism and high sequence homology. Thus it is an interesting mechanistic subject how these enzymes use a common active site template to achieve specificity for each reaction.
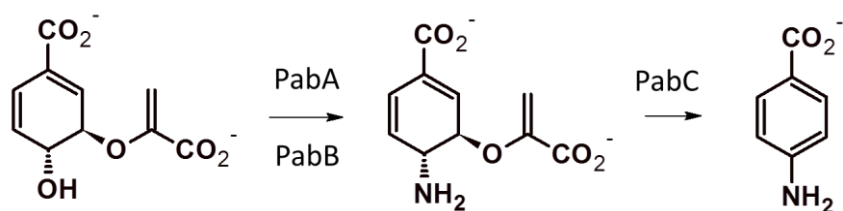
Figure 4.5 Synthesis of PABA following ADCS (PabB)

In *E. coli*, TrpG and PabA are separate glutamine amidotransferases dedicated to the biosyntheses of tryptophan and PABA/folate, respectively. However, the *trpG* and *pabA* in *M. tuberculosis* is designated to the same open reading frame, Rv1005c [126], which is not clustered to other *trp* or *pab* genes (**Table 4.1**, **Figure 4.39**), and how it is involved in these two pathways is an open question.

Table 4.1 Locations of putative genes for PABA synthesis in M. tuberculosis

ORF	Annotation	Location (kb)
Rv0013	pabA (trpG)	15
Rv1005c	pabB	1124
Rv0812	pabC	906

Unlike the stable oligomeric structures of ASs, the association of ecPabB with PabA is thought to occur only transiently, and PabB is found to be monomeric in the crystal structure

(**Figure 4.6**) [335-339]. In plants, PabA and PabB are fused and are two domains of a single protein [340]. The hydroxylation enzymes MbtI, Irp9, MenF and EntC that do not require the amidotransferase are also found to be monomers [132,140,341-345].
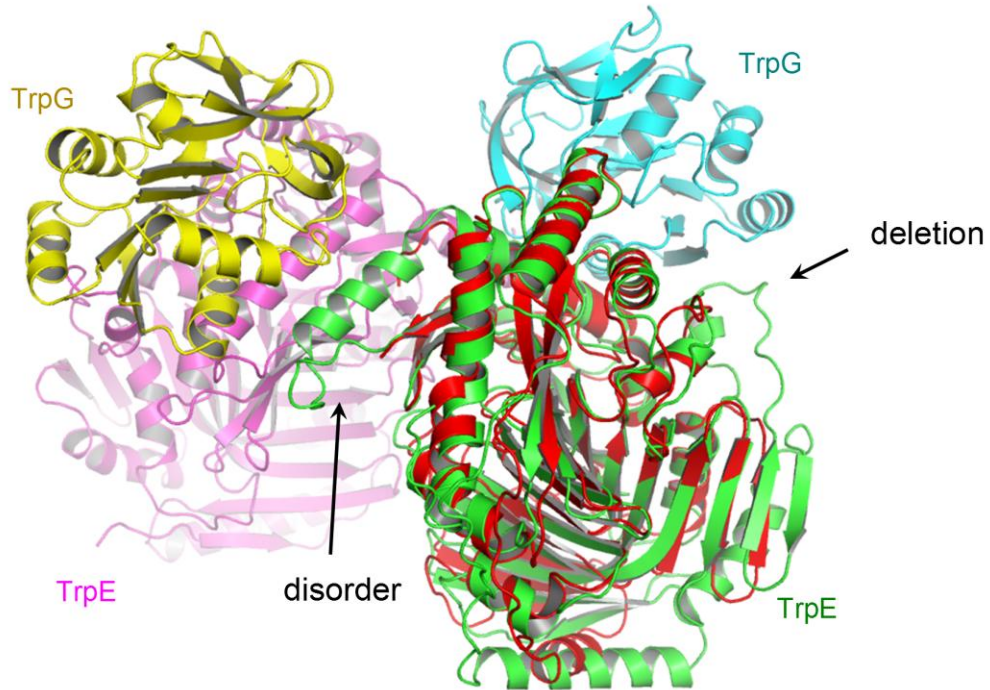
Figure 4.6 Superposition of ecPabB structure to the heterotetramer of smAS

ecPabB is shown in red. Arrows indicate regions involved in oligomerization of smAS that are disordered or deleted in ecPabB.

4.2 Methods

Cloning, expression and purification of annotated *trpE* and *trpG* (*pabA*)

Rv1609, the predicted *trpE* from *M. tuberculosis*, was cloned from H37Rv genomic DNA into vector pET15b using the forward primer 5'- GGAATTCCATATGCACGCCGACCTCGCAGCCA -3' and the reverse primer 5'- CCGCTCGAGTTAGCAGCCACTGCGGTTC -3', between the two restriction sites, NdeI and XhoI. *Rv0013*, the predicted *trpG* or *pabA* from *M. tuberculosis*, was cloned into vector pET23b using the forward primer 5'- GGAATTCCATATGCGGATCCTGGTCGTTGA -3' and the reverse primer 5'- TTTCCTTTTGCGGCCGCTGAGGTTCGGCCAGT -3', between the two restriction sites NdeI and NotI. Such design leaves an N-terminal and a C-terminal His-tag on TrpE and TrpG, respectively. *E. coli* cells BL21(DE3) were transformed with the plasmids containing *trpE* and *trpG* individually,

158

grown to OD larger than 0.5 and induced for expression with IPTG at room temperature overnight. Cell pellets were harvested, resuspended, and TrpE and TrpG each purified by Ni affinity chromatography using Ni-NTA resin. The eluate from Ni column was subjected to gel filtration chromatography on Superdex 200 equilibrated with 50 mM Tris at pH 7.5.

Assay of AS activity

The standard reaction mixture contained 9 mM $MgCl_2$, 110 mM NH_4Cl, 48 µM chorismate in 50 mM triethanolamine buffer at pH 8.5, and the reaction was initiated by the addition of TrpE. UV-vis absorption of the reaction mixture was scanned from 240-400nm at fixed time intervals by Cary 300 spectrophotometer to follow the progress of the reaction. The absorption at time zero was subtracted to reveal only the difference from the initial mixture.

Kinetics of AS activity

The reaction mixture contained NH_4Cl 50 mM, $MgCl_2$ 10 mM, chorismate and TrpE at specified concentrations in 50 mM Tris. When TrpG is included, glutamine at 2.4 mM substitutes for NH_4Cl. AS activity is followed by change in UV-vis absorption at 314 nm.

The lactate dehydrogenase (LDH) coupled assay

TrpE 1.4 µM, Mg^{2+} 10 mM, NH_4^+ 50 mM, NADH 320 µM, LDH 1 µL in Tris buffer 50 mM, pH 8.5. The reaction was followed by the absortion at 360 nm.
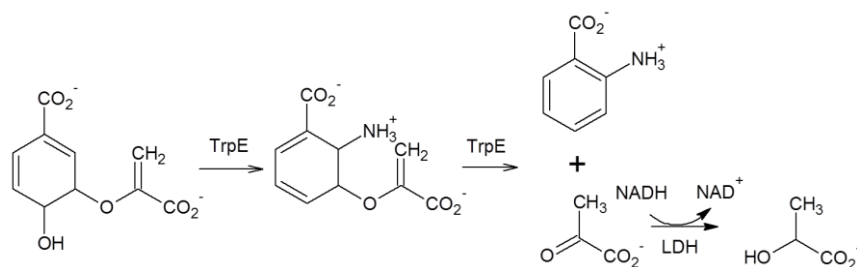
Figure 4.7 LDH coupled TrpE assay

Isothermal calorimetry titration (ITC)

The ITC experiment was performed on a MicroCal VP-ITC unit with a cell volume of 1.445 ml. 170 µM TrpE in 100 mM potassium phosphate pH 7.0, 0.1 mM EDTA and 1.0 mM DTT was titrated with 5 mM tryptophan in the same buffer with 4 µL steps at 25 °C. All solutions were degassed under vacuum prior to use. Nonlinear least-squares curve fitting was performed using Origin v 5.0.

Reactions followed by NMR spectroscopy

^{1}H NMR spectra were collected on a Varian INOVA 600 MHz spectrometer at 25 °C. The PRESAT gradient spin echo sequence was used to suppress the signals from water. All the reaction components are mixed except the enzymes, which are added last to initiate the reaction.

Complex formation followed by gel filtration

Gel filtration chromatograms were obtained by running TrpE or pre-mixed TrpEG through Superdex 200 columns Hiload 16/60 or 10/300 GL operated by the AKTA FPLC system.

Complex formation followed by analytical ultracentrifugation (AUC)

TrpE sample stock with A(280 nm) 0.86 was mixed with the TrpG sample stock with A(280 nm) 0.99 at a ratio of 3/1 (v/v). The control cell contains only the protein buffer. The program started from RPM 3000 followed by 4 stages with increasing rotation speed from RPM 6000, 8500, 10000 to 13000 each for 22 hours at 4°C.

Cloning and expression of annotated *pabB*

Rv1005c, the annotated *pabB*, was cloned from H37Rv genomic DNA into expression vectors pET23b, pET28a and pETDuet using the primers in **Table 4.2**. The expression of the pETDuet plasmid is designed to attach an N-terminal His-tag to TrpG (PabA) and no tag on PabB.

Table 4.2 Vectors and primers for cloning pabB

Vectors	Restriction	Primers
pET23b	NdeI	F: 5'-GGAATTCCATATGAACTTAGCGTGGGAGCTT-3'
	NotI	R: 5'-TTTTCCTTTTGCGGCCGCACTTTGCTGGCTAA-3'
Deletion mutagenesis		F: 5'-CAAAGTGCGGCCGCTCGAGCACCACCAC-3'
		R: 5'-GTGGTGGTGCTCGAGCGGCCGCACTTTG-3'
pET28a	NdeI	F: 5'-GGAATTCCATATGAACTTAGCGTGGGAGCTT-3'
	NotI	R: 5'-ATAGTTTAGCGGCCGCCTACCGCACTTTGCTGG-3'
	NdeI	F: 5'-GGAATTCCATATGAACTTAGCGTGGGAGCTT-3'
	NheI	R: 5'-CTAGCTAGCGCACTTTGCTGGCTAACCGA-3'
pETDuet (MCS1: pabA) (MCS2: pabB)	EcoRI	F: 5'- GGAATTCATGCGGATCCTGGTCGTTGACAA -3'
	NotI	R: 5'-ATAGTTTAGCGGCCGCTCACGCTGAGGTTCGGCCA-3'
	NdeI	F: 5'-GGAATTCCATATGAACTTAGCGTGGGAGCTT-3'
	BglII	R: 5'-GAAGATCTCTACCGCACTTTGCTGGCTAA -3'

Expression of *pabB*

Expression by the *E. coli* pLysS cells was induced by IPTG at 22˚C and 37˚C on the scale of 50 mL and 500 mL of LB cultures.

Co-expression of pabA and pabB

E. coli pLysS cells bearing both the plasmids trpG-pET23b and pabB-pET28a were doubly-selected by ampicillin and kanamycin. 500 mL of cell culture was grown to OD larger than 0.7 and induced by 0.2 mM IPTG for expression at 16 ˚C for 8 hours.

Expression of *pabB* by the Arctic *E. coli* strain

pabB cloned into pET28a, pET23b and pETDuet was expressed alone and co-expressed with *trpG* by the Arctic *E. coli* strain. In each case a 50 mL culture was induced by 0.1 mM IPTG at 11.5 °C for 20 hours.

4.3 Results

TrpE purification

TrpE expresses well, solubility is good, and the yield is greater than 5 mg of purified protein from 1L of LB media. TrpE elutes from the G200 gel filtration column as three major peaks (A1, A11, B11, low to high elution volume), all of which run as the same band on protein gel (58 kDa, **Figure 4.8**) and possess anthranilate synthase activity although with slightly different kinetics (**Figure 4.9**). It is speculated the higher MW fractions accumulate the reaction intermediate, ADIC.
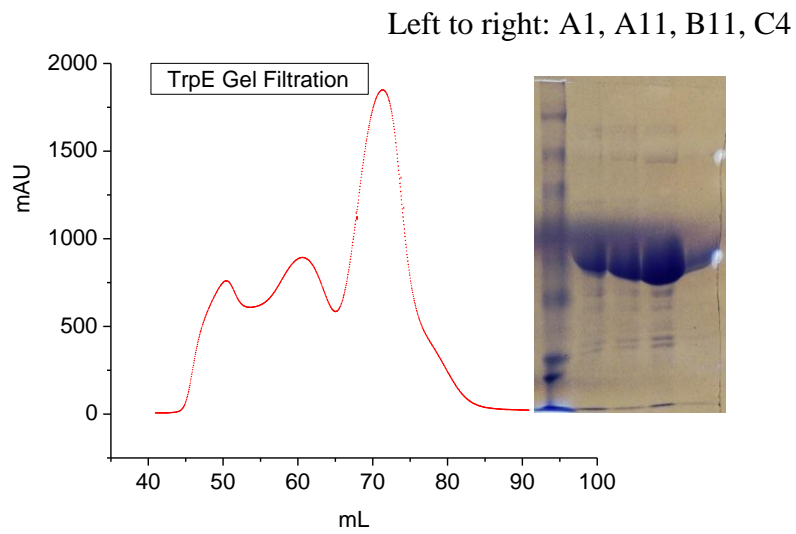
Figure 4.8 TrpE gel filtration chromatogram and SDS PAGE for each major peak

162
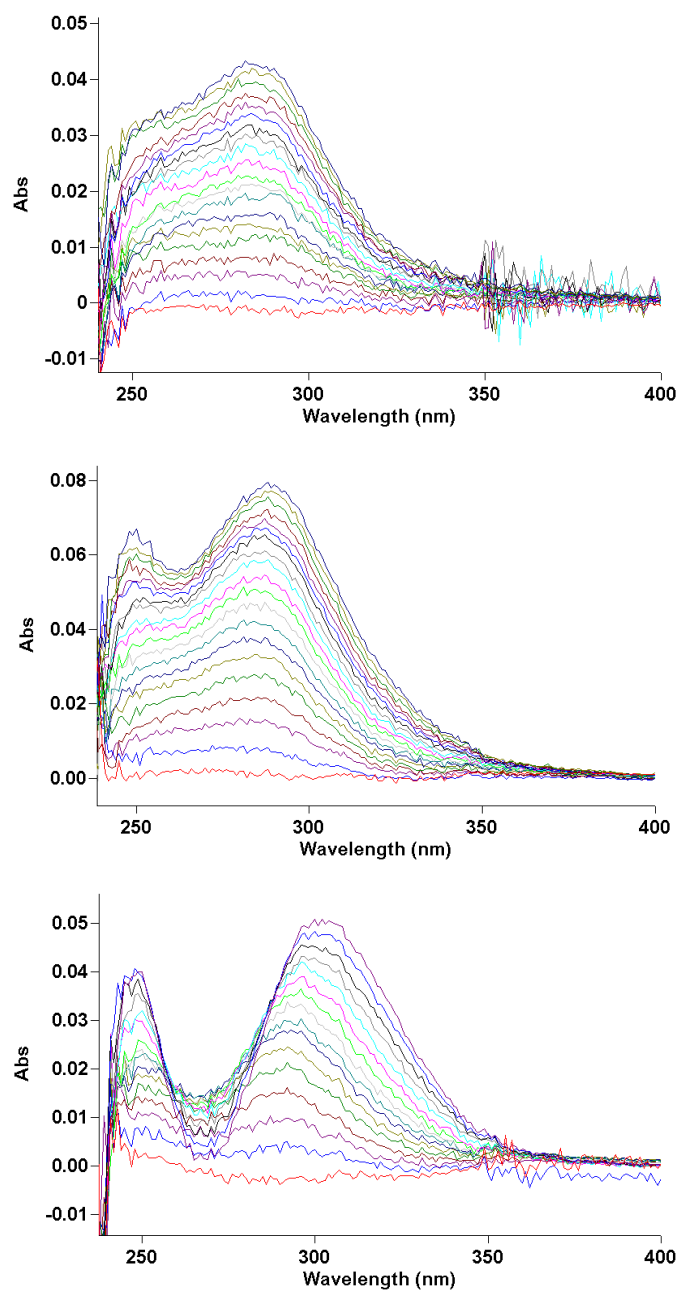
Figure 4.9 Kinetics of TrpE activity from fractions of different MW.
Top: fraction A1; middle: fraction A11; bottom: fraction B11. Reactions span 40 min.

Characterization by MALDI

TrpE can be identified from the database following digestion (**Figure 4.10**).
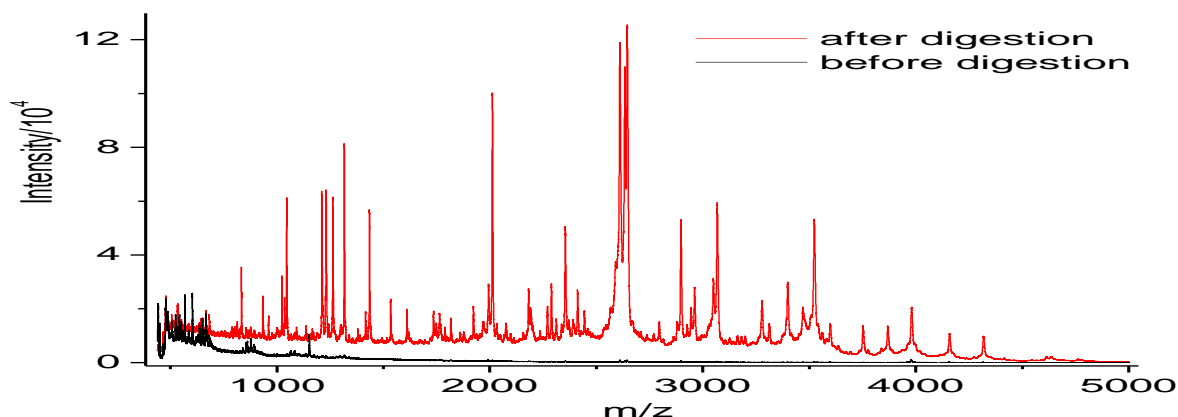
Species	protein	match %	score
MYCLE	UPF0051 protein ML0593	24%	91082
MYCBO	TrpE	32% (20/61 peaks)	82160
MYCTU	TrpE	32% (20/61 peaks)	82160

Figure 4.10 Identification of TrpE by MALDI

TrpE behavior as an anthranilate synthase

The TrpE subunit of AS is capable of taking ammonium salts such as sulfate and chloride as the nitrogen donor, and is known to be Mg^{2+}-dependent and subject to feedback inhibition by tryptophan [346,347]. The purified mtTrpE was tested for these characteristics. The UV-vis spectra were taken at fixed time intervals after the reaction was initiated, and the trace from time zero was subtracted so that only the difference from time zero will be observed. The anthranilic acid standard absorbs at 314 nm and chorismate absorbs at 274 nm. In the presence of Mg^{2+} and NH_4^+, reduction of the 274 nm peak and growth of the 314 nm peak can be observed, indicating the conversion from chorismate into anthranilate. There is no anthranilate formation in the absence of either Mg^{2+} or NH_4^+; nor is anthranilate formed in the presence of tryptophan. These observations are characteristic of an anthranilate synthase. The same reaction carried out at different pH values (**Figure 4.11**, panel 1 and 6 at pH 8.5 and 7.4, respectively) suggests different kinetics in anthranilate formation. One rather unexpected observation is that chorismate is still disappearing while anthranilate cannot be detected (panel 2, 3, 4). Such

164

activity is only absent when no enzyme is added (panel 5). Therefore this activity is enzymatic and not due to un-catalyzed chorismate transformation. The isothermal calorimetry curve shows that tryptophan binds to TrpE with a binding constant of 3.2 µM, consistent with the effective inhibition of TrpE by 8 µM of tryptophan (**Figure 4.12**).
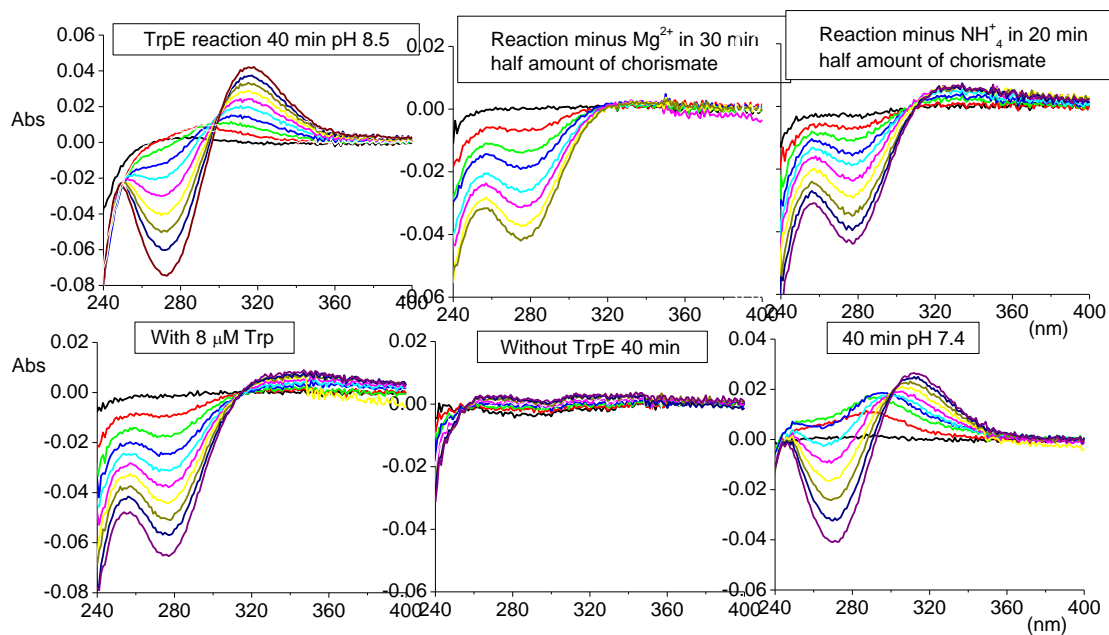
Figure 4.11 Anthranilate formation in the reactions catalyzed by mtTrpE.

Anthranilate synthase activities.of the reaction mixture (a) with everything except the enzyme; (b) with everything except $MgCl_2$; (c) with everything except NH_4Cl; (d) with everything; (e) with everything plus 8 µM tryptophan; (f) with everything but the buffer was adjusted from pH 8.5 to pH 7.4.
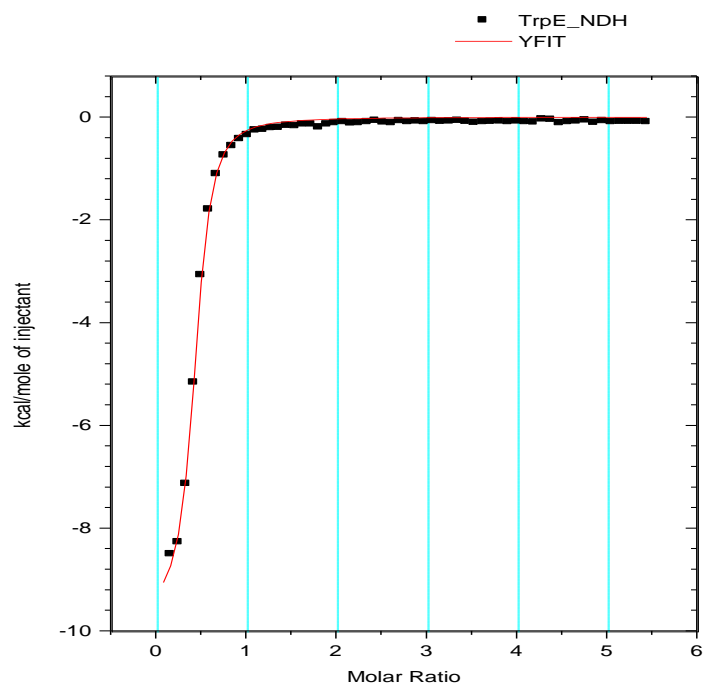
Figure 4.12 ITC Titration curve of TrpE by tryptophan.
The isothermal curve shows that tryptophan binds to TrpE, and binding is saturated at a molar ratio of 1. Data fitting gives a dissociation constant of 3.2 μM.

The reaction products

The TrpE reaction mixtures were separated by HPLC (**Figure 4.13**). At pH 8.5, more anthranilate and less ADIC were produced compared with pH 7.5. PAS does not inhibit anthranilate formation at 100 µM.
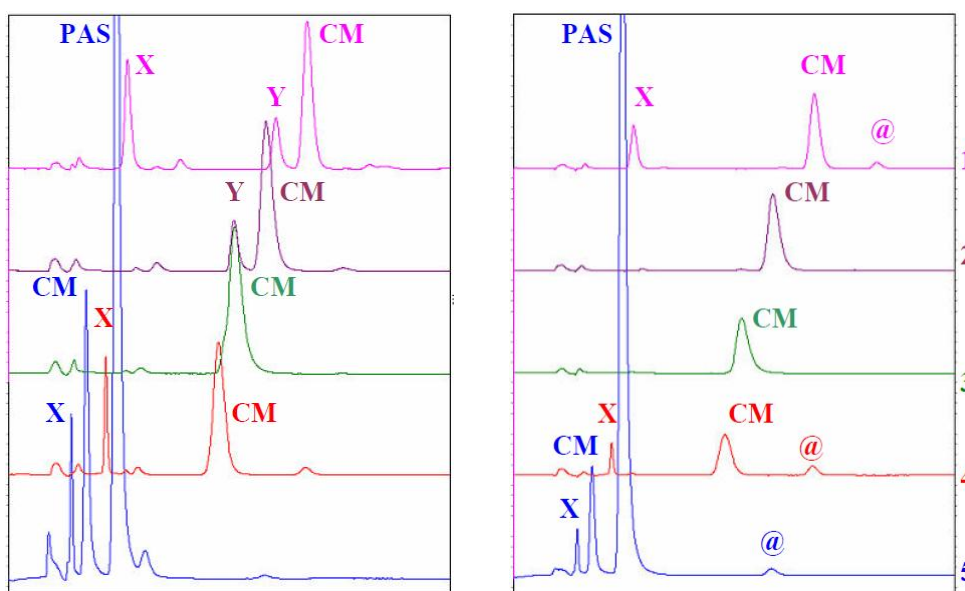
Figure 4.13 HPLC separations of the TrpE reaction mixtures

Left: 280 nm; right 310 nm. 1 µM TrpE, 150±15 µM chorismate, 10 mM Mg^{2+}, 50mM NH_4^+, 30 min, 25°C, in Tris 50 mM; samples were prepared in 1.3% acetic acid before injection, 1 mL/min, 5% acetic acid isocratic.

Trace 1: reaction at pH 7.5
Trace 2: control with no NH_4^+, pH 8.5
Trace 3: control with no enzyme, pH 8.5
Trace 4: reaction at pH 8.5
Trace 5: reaction in the presence of 100 M PAS, pH 8.5

@—Anthranilate
CM—chorismate
PAS-- p-aminosalicylate
X—probably aminodeoxyisochorismate (ADIC)
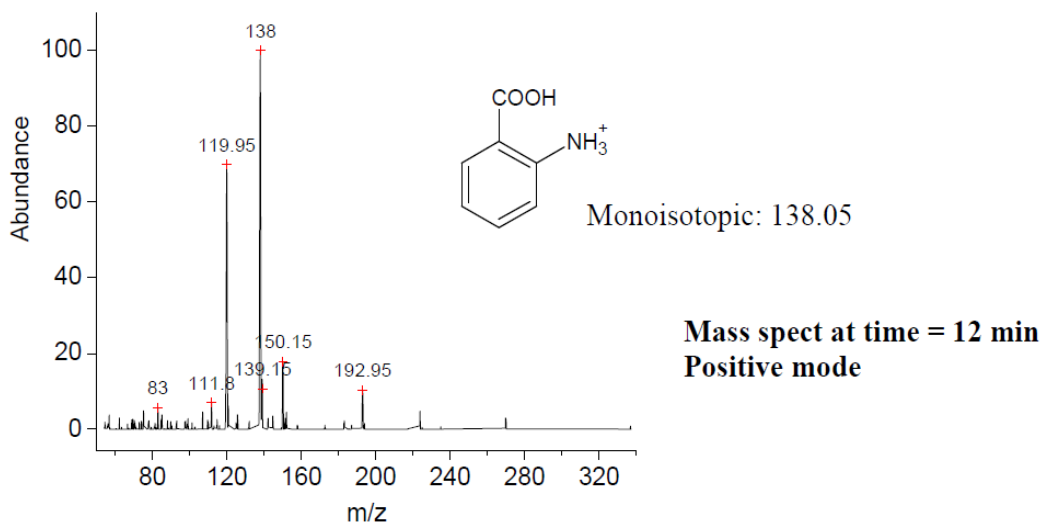
Figure 4.14 Mass spectrum of the TrpE reaction products
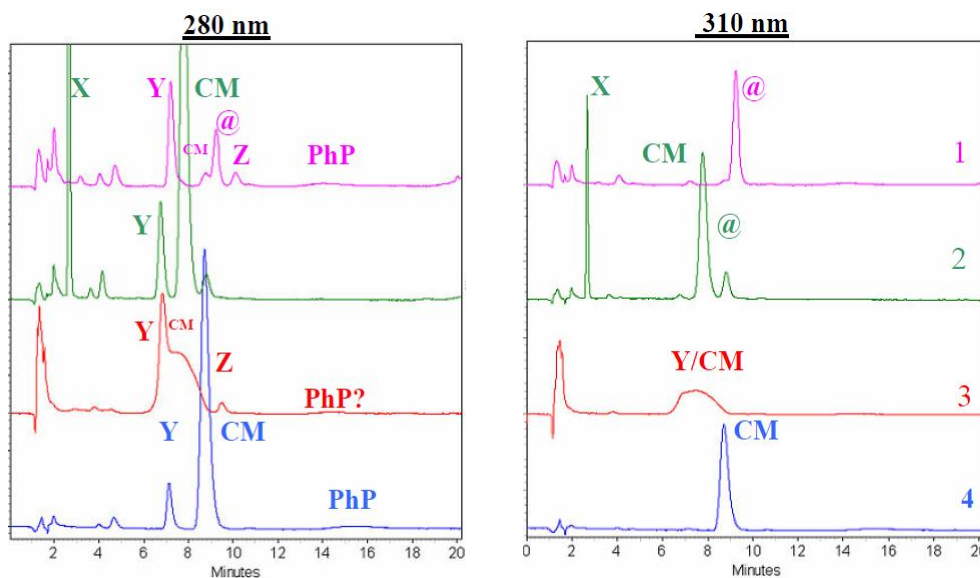
Figure 4.15 HPLC separations of longer TrpE reaction mixtures

Trace 1 (magenta): reaction at pH 8.5 for 4 hours
Trace 2 (green): reaction at pH 7.5 for 0.5 hour
Trace 3 (red): reaction without Mg^{2+} at pH 8.5 for 80 min
Trace 4 (blue): standard chorismate (CM) + phenylpyruvate (PhP)

After 4 hours, the anthranilate peak predominates and the chorismate peak is very small. Peak X (ADIC) is also small, indicating the reaction is near competion. The separation for the Mg^{2+}-deficient reaction is worse. The phenylpyruvate peak is broadened and hard to see.

Characterization of reaction products by ^{1}H NMR spectroscopy

The product of the second activity observed in UV-vis experiments is apparently UV-insensitive. When the reaction is monitored by NMR, it is revealed that the second activity converts chorismate into phenylpyruvate via prephenate (**Figure 4.17**) [348], which accompanies anthranilate synthesis in the presence of Mg^{2+} (**Figure 4.16b**), but is the exclusive event when Mg^{2+} is absent (**Figure 4.16c**). Chorismate is marginally stable and slowly undergoes rearrangement into prephenate in the absence of enzyme (**Figure 4.16a**). Comparison with the uncatalyzed reaction shows all of the anthranilate synthase (AS), chorismate mutase (CM) and prephenate dehydratase (PDT) activities observed in **Figure 4.16b** and **Figure 4.16c** are enzymatic.

Small intermediate peaks during anthranilate formation in the early stage of the reaction can be observed, presumably from the intermediate ADIC (**Figure 4.16b**, 25 min). Lower concentrations of enzyme slow down all three activities, especially the PDT activity (**Figure 4.16b**, blue trace), and anthrnilate and prephenate appear to be the final products after one day.

Observations from UV-vis and NMR taken together, the TrpE preparation catalyzes chorismate transformation into two reaction paths, one gives anthranilate via ADIC that requires Mg^{2+}, the other gives phenylpyruvate via prephenate that does not require Mg^{2+} (**Figure 4.17**).
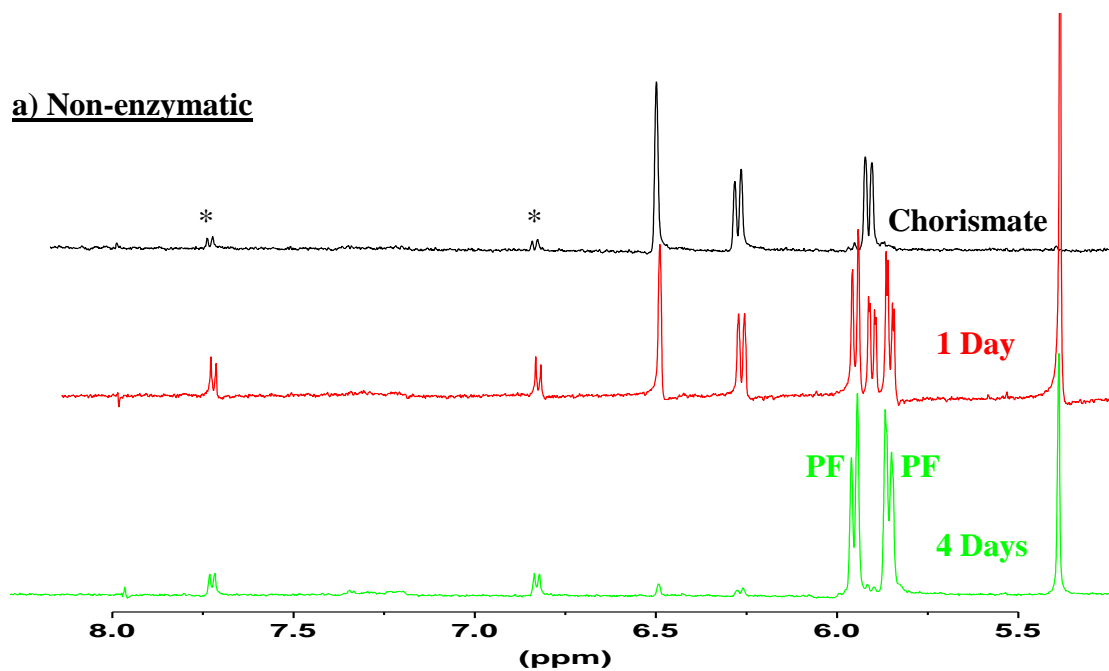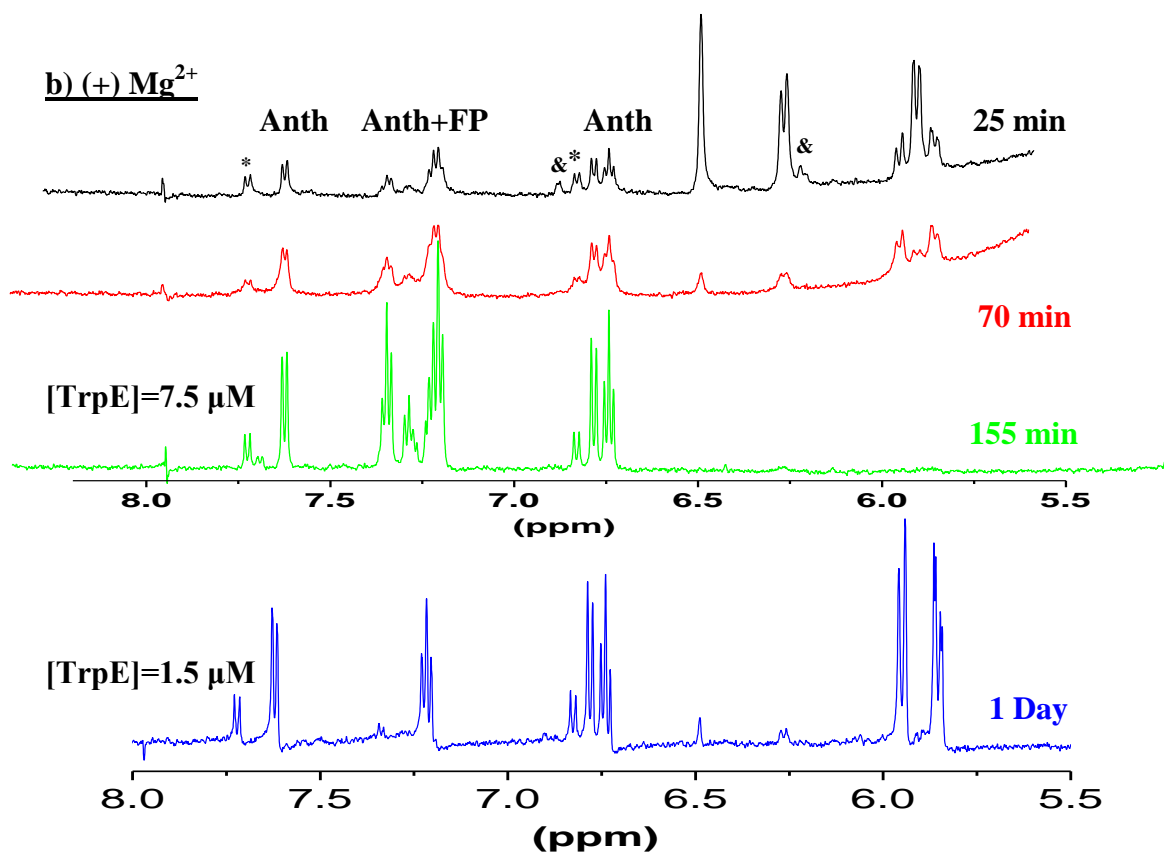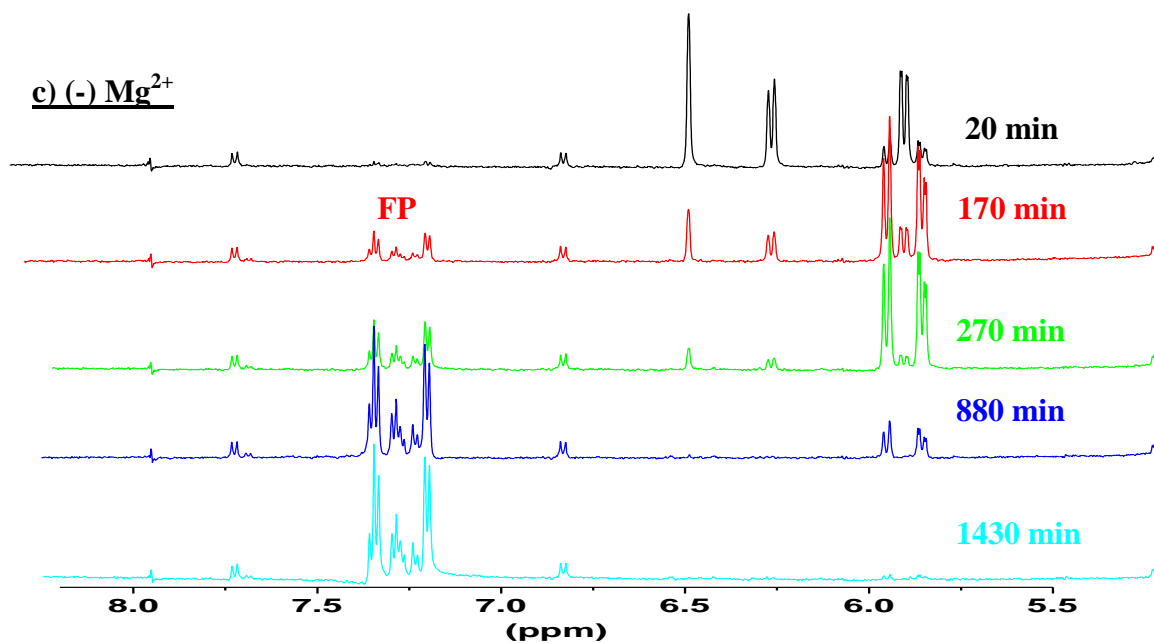
Figure 4.16 TrpE Reactions monitored by ^{1}H NMR
Chorismate 1.3 mM, TrpE 7.5 µM if present, Mg^{2+} 10 mM if present, NH_4^+ 50 mM in Tris buffer 20 mM, pH 8.3. Anth: anthranilate; PF: prephenate; FP: phenylpyruvate. (a) Control experiment without any enzyme over four days. (b) TrpE reactions with Mg^{2+} over 155 min; blue spectrum is a separate experiment at a lower concentration of TrpE, 1.5 µM, over one day. (c) TrpE reaction without Mg^{2+} over one day.
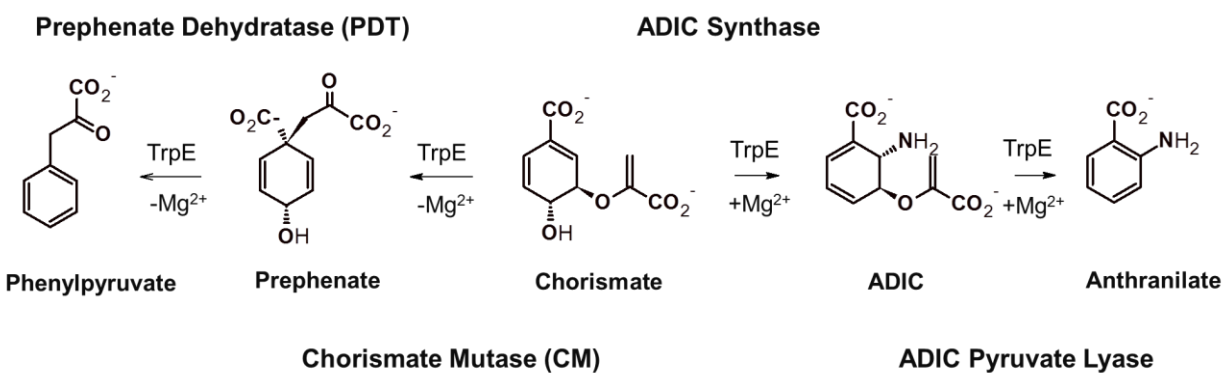
Figure 4.17 Activities of the TrpE preparation

To examine the possibility that the observed CM and PDT activities are from *E. coli* contaminating proteins, the effect of phenylalanine on those activities are tested. The *E. coli* P-protein, comprising both CM and PDT that lead to the biosynthesis of phenylalanine and tyrosine, is inhibited by phenylalanine [349]. It is shown that the rate of chorismate consumption in the absence of Mg^{2+} is not affected by phenylalanine (**Figure 4.18**). Thus it is not the normal *E. coli* P-protein that gives rise to the side activities.

(-) Phenylalanine   **(+) Phenylalanine**
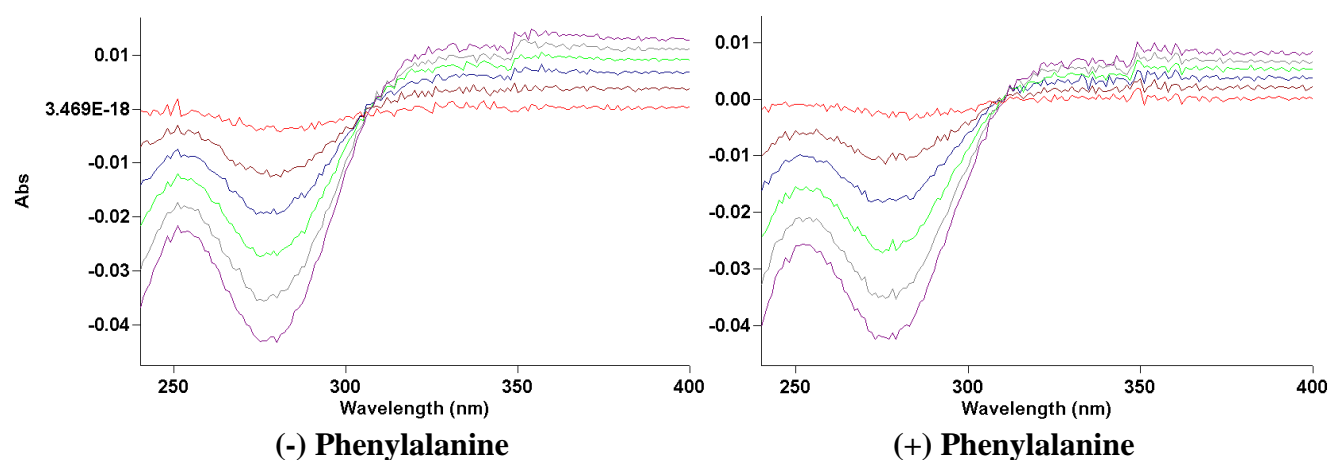
Figure 4.18 Effect on CM and PDT activities by phenylalanine
Chorismate 150 μM, TrpE 1.2 μM, no Mg^{2+}, NH_4^+ 50 mM in Tris buffer 50 mM, pH 8.5. Phenylalanine 25 μM for the (+) experiment. Time interval is 6 min. (a) without phenylalanine (b) with phenylalanine

Table 4.3 Final product distribution at different pH
Chorismate 150 μM, TrpE 1.2 μM, no Mg^{2+}, NH_4^+ 50 mM in Tris buffer 50 mM, pH 8.5. Phenylalanine 25 μM for the (+) experiment. Time interval is 6 min.

Buffers	%Anth	%FP
pH 7.1, (+)Mg^{2+}*	81	19
pH 9.0, (+)Mg^{2+}	60	40
pH 8.3, (+)Mg^{2+}	55	45
pH 7.1, (+)Mg^{2+}	14	86
pH 8.3, (-)Mg^{2+}	0	100

* In the presence of TrpG

pH dependence of TrpE activities

The ratio of the products formed at the end of the reactions performed at different pH values are summarized from the NMR data (**Table 4.3**). The product distribution manifests the

outcome of the kinetic race between the two reaction paths. Basic conditions favors the TrpE AS activity in the range investigated.

TrpG effectively activates anthranilate synthesis

The slow kinetics and promiscuity seen for mtTrpE reaction prompt us to ask whether TrpG is necessary for maintaining better efficiency of anthranilate production in *M. tuberculosis*. *Rv0013*, predicted to encode the amidotransferase for both anthranilate synthase and ADC synthase in *M. tuberculosis*, was cloned, expressed and purified. The yield of TrpG is greatly limited by its poor solubility, and the majority of the expressed TrpG is found in the insoluble fraction (26 kDa) (**Figure 4.19**), and the recovered soluble TrpG also tends to form irreversible high MW aggregates, and unlike TrpE, gradually loses activity at 4°C in a few days. The loss of activity is slower if TrpE is present (**Figure 4.20**).
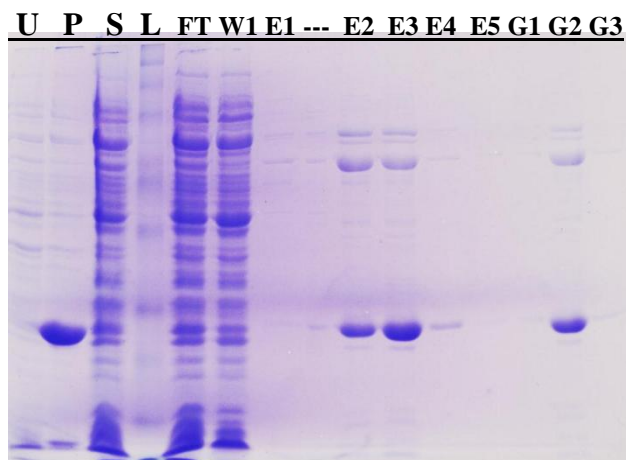
Figure 4.19 SDS PAGE from TrpG purification by His Bind and G25 resin
"E" lanes are eluates from the His Bind column and "G" lanes are from the G25 gel filtration column. The gel contains 12% acrylamide.

Nevertheless, TrpG proves to be the indispensible component of *M. tuberculosis* anthranilate synthase. With glutamine as the nitrogen donor, anthranilate is produced much more efficiently by TrpE and TrpG working together (AS-Gln activity, **Figure 4.20**, Day 1). As a result, the reaction is complete within minutes as compared to hours of time required for TrpE-only reaction to use up all the chorismate provided (**Figure 4.11**, reactions were not complete). Such efficiency is observed at pH 7.5, and with a starting glutamine concentration of around 2 mM, compared to higher pH and high concentration of ammonium salt required to enhance AS activity of TrpE-only reaction. The well-defined isosbestic points (**Figure 4.20**) also suggest cleaner conversion of chorismate compared to the TrpE-only reaction (**Figure 4.11**). As

expected, chorismate is better partitioned into the AS pathway and lower amounts of phenylpyruvate are observed at the end of the TrpEG reaction (

Figure 4.21, **Table 4.3**). These observations clearly show the advantages of TrpEG reaction over TrpE-only reaction.

TrpEG can still utilize ammonium chloride for amination (AS-NH_3 activity), but glutamine is a more efficient nitrogen donor than the ammonium salt (data not shown). In summary, the AS reaction efficiency is in the following order: TrpEG-Gln > TrpEG-NH_3 >> TrpE-NH_3.

Day 1: TrpE-TrpG

Day 4: TrpE-TrpG
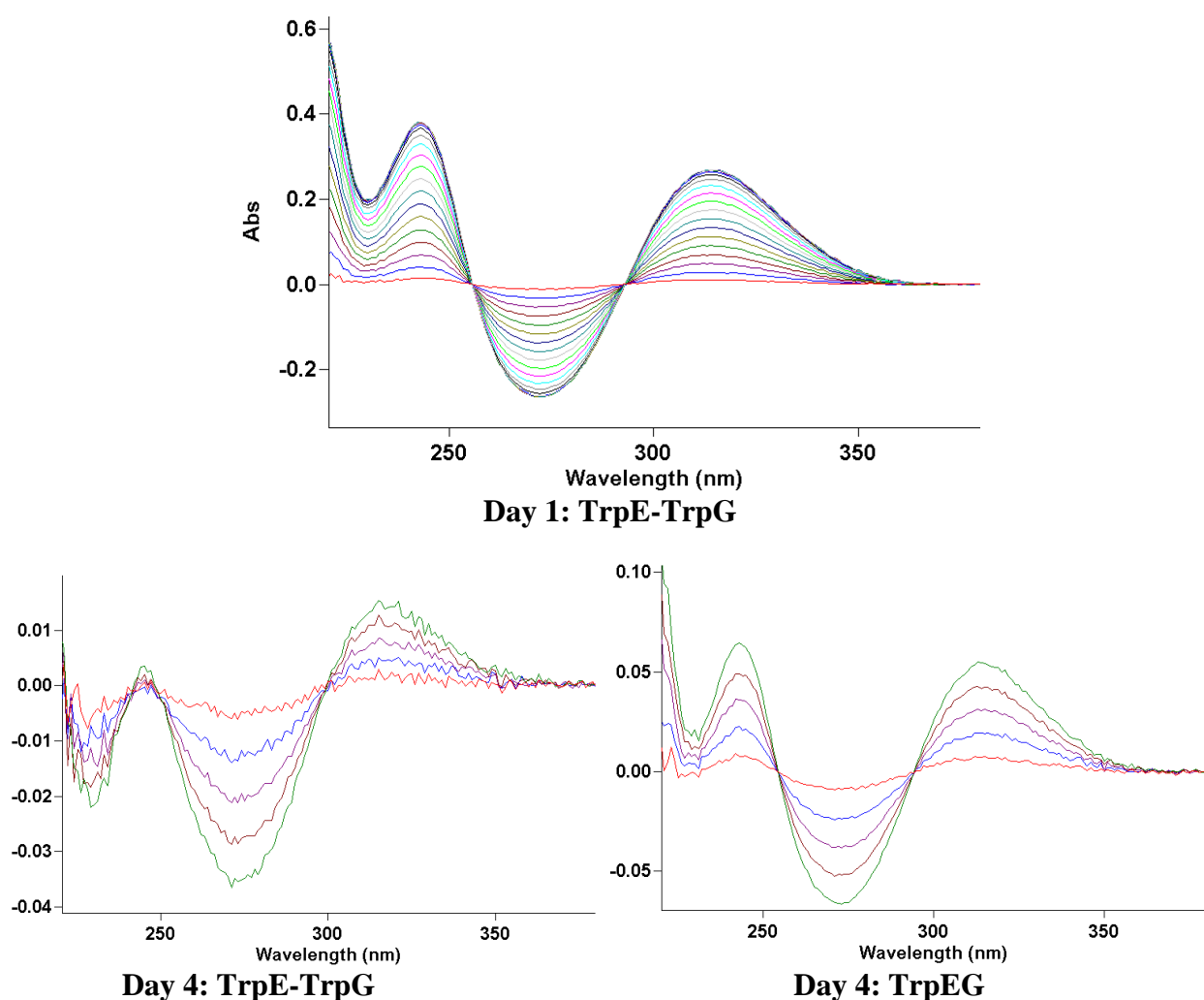
Day 4: TrpEG

Figure 4.20 The effect of TrpG on the TrpE activity

TrpG stabilization by TrpE. AS Gln-dependent activity with fresh TrpG (Day 1), and aged TrpG (Day 4). TrpE-TrpG were mixed on Day 4 while TrpEG were mixed on Day 1. Tris 50 mM pH 7.5, Mg2+10 mM, glutamine < 2.4 mM, [chorismate] ~ 175 μM, TrpE 0.7 μM, TrpG 1 μM, scans every 1 min

174
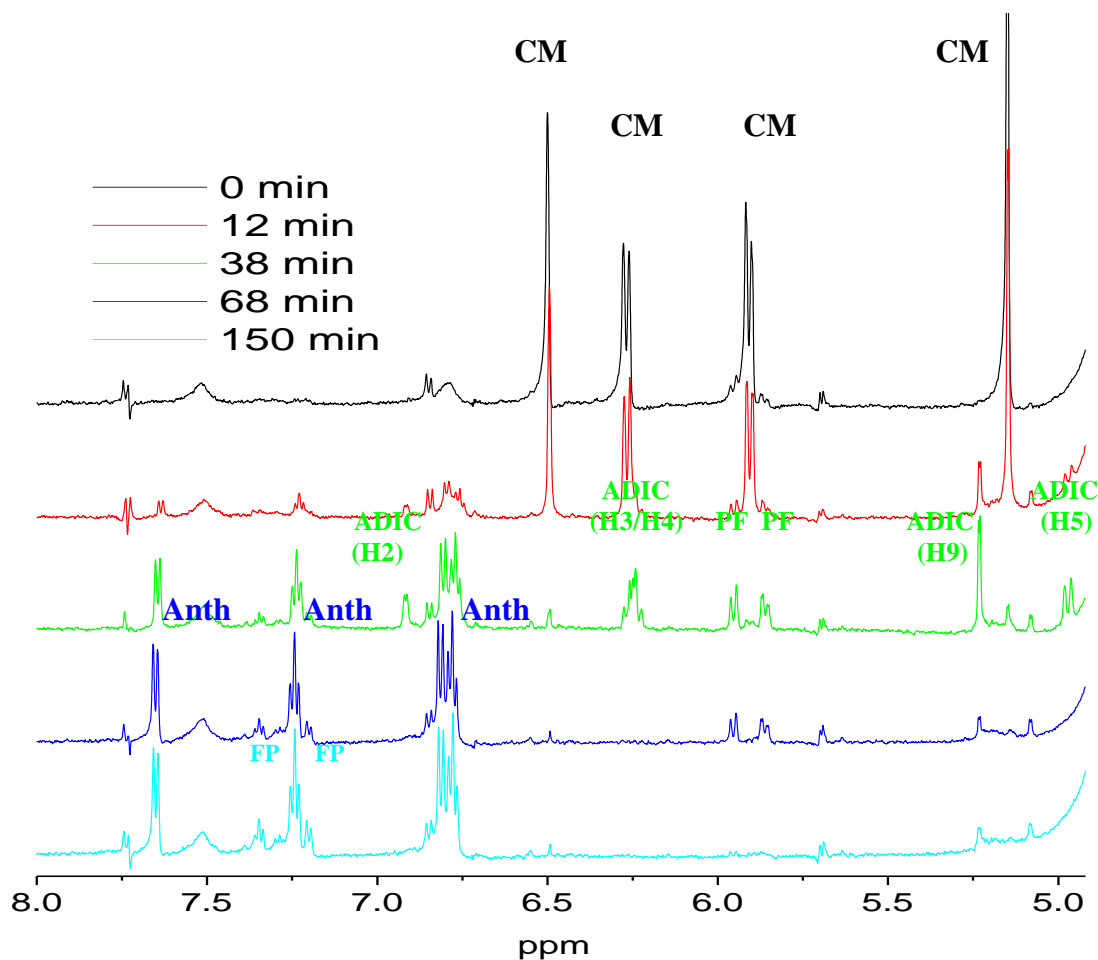
Figure 4.21 TrpEG reaction monitored by ^{1}H NMR

Chorismate 890 µM, TrpE 3.9 µM, TrpG 5.1 µM, Mg^{2+} 10 mM, glutamine < 5.1 mM, Tris 50 mM, pH 7.1

The improved AS efficiency is also reflected in the kinetic parameters. The difference in chorismate affinity for TrpE with and without TrpG is obvious when saturation experiments are compared (**Figure 4.22**, **Figure 4.23**). In the absence of TrpG, it takes milli-molars of chorismate to saturate TrpE. The error at high substrate concentration is large, and the K_m for chorismate is in the range of 600 µM. Using extinction coefficients at 314 nm of 303 and 3060 cm^{-1}M^{-1} for chorismate and anthranilate, respectively, and a difference extinction coefficient of 2757 cm^{-1} M^{-1}, the k_{cat} value for TrpE-only reaction is estimated to be 3 min^{-1}. In contrast, saturation in the presence of TrpG is much easier. The K_m and k_{cat} values are estimated to be 25 µM and 11 min^{-1}, respectively. When [TrpG]/[TrpE] is further optimized, k_{cat} can be increased to 36 min^{-1} (**Figure**

175

4.33). Thus, TrpG enhances AS efficiency by improving both K_m and k_{cat}, which strongly suggests physical interactions between TrpE and TrpG.

Enzyme Kinetics Data

Parameter	Value	Std. Error
Vmax	0.0133	0.0033
Km	0.5667	0.4219
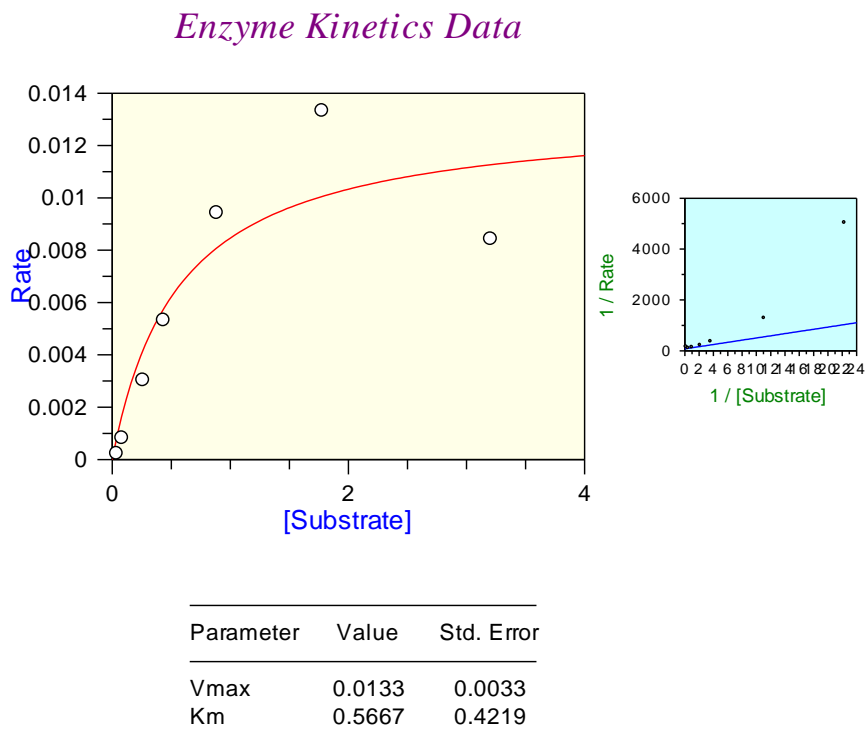
Figure 4.22 Saturation of TrpE by chorismate
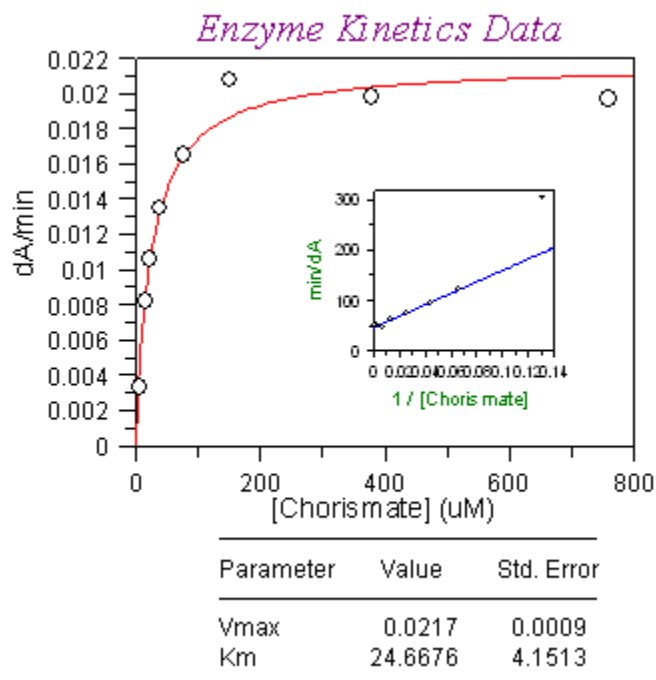
Figure 4.23 Saturation of TrpEG by chorismate

The improved kinetic properties are certain despite that there are issues in the accuracy of kinetics in the TrpE assays (**Figure 4.24**).
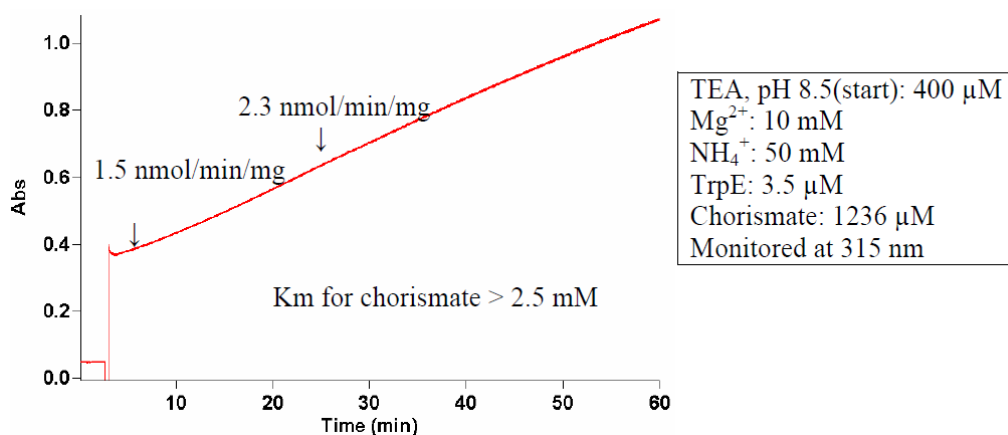
Figure 4.24 Sigmoidal initial velocity in the TrpE reaction
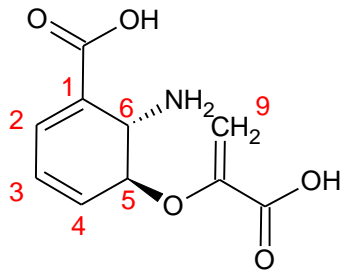
Figure 4.25 Structure of ADIC

Accumulation of ADIC during the course of the reaction

The peaks from the intermediate grow even more prominent in the TrpEG reaction than in the TrpE-only reaction (

Figure 4.21, **Figure 4.16**) and allow the assignment to ADIC (**Table 4.4**, **Figure 4.25**). Nevertheless, it takes around 40 min for [ADIC] to reach a maximum before disappearing into anthranilate. If the rate of anthranilate formation is proportional to [ADIC], the observation will explain the little lag in the beginning of the progress curve for anthranilate formation observed in the UV-vis experiment.

Table 4.4 ADIC chemical shifts from different preparations

δ		H2	H3/H4	H9a	H5	H9b	H6
*	ADIC(rxn)	6.92	6.25 (m)	5.23	4.97		
	D_2O	(d, J = 4.2 Hz)		(d, J = 2.4 Hz)	(d, J = 11.2)		
a[350]	ADIC	7.37	6.48 (m)	5.60	5.08	4.90(m)	4.51
	CD_3OD	(d, J = 4.5 Hz)		(d, J = 3.0 Hz)	(d, J = 2.7)		(d, J = 5.2)
b[351]	ADIC free	7.05	6.28 (m)	5.34 (s)	4.99	4.59 (s)	4.45
	CD_3OD	(br, s)			(d, J = 12.3)		(d, J = 12.3)
	ADIC·TFA	7.37	6.50 (m)	5.60	5.08	4.90 (t)	4.50
	CD_3OD	(d, J = 5.1 Hz)		(d, J = 3.0 Hz)	(d, J = 3.0)		(d, J = 4.8)
c[352]	ADIC·TFA	7.37	6.47 (m)	5.63	5.07	DOH	4.53
	CD_3OD	(d, J = 5 Hz)		(d, J = 3 Hz)	(d, J = 3)		(d, J = 6)

*current work

The large coupling constant at H5 suggests that most of the ADIC adopts the conformation where both the enolpyruvyl and amino groups are equatorial [351] (although the absolute configuration drawn in this paper does not seem to be right). However this does not reflect the enzyme-bound conformation since ADIC is present in greater stoichiometry than the enzyme.

Improved purity of TrpE and TrpG that reduces side activities

Promiscuous CM or PDT activity has been found associated with several chorismate or isochorismate utilizing enzymes including the salicylate synthase MbtI [140] and Irp9 [353], isochorismate pyruvate lyase PchB [354], and a mutant anthranilate synthase [353], and this might be extended to mtAS. However, the mtTrpG preparation was found to display similar CM activity as well, which implies the same impurity from E. coli. Extensive wash during Ni affinity chromatography was then applied to both TrpE and TrpG purification, and it is seen indeed that the side activity can be reduced by a great extent (**Figure 4.26**). The oligomeric state of TrpE is not affected by the impurities. Because of the lower yield of TrpG, it is less easy to get rid of some of the impurity bands, and so an additional ion exchange step was applied to TrpG purification. The relatively higher pI of TrpG separates it from other proteins (**Figure 4.27**). Fraction F9 was tested for CM activity and unable to account for the observed CM activity (data not shown).
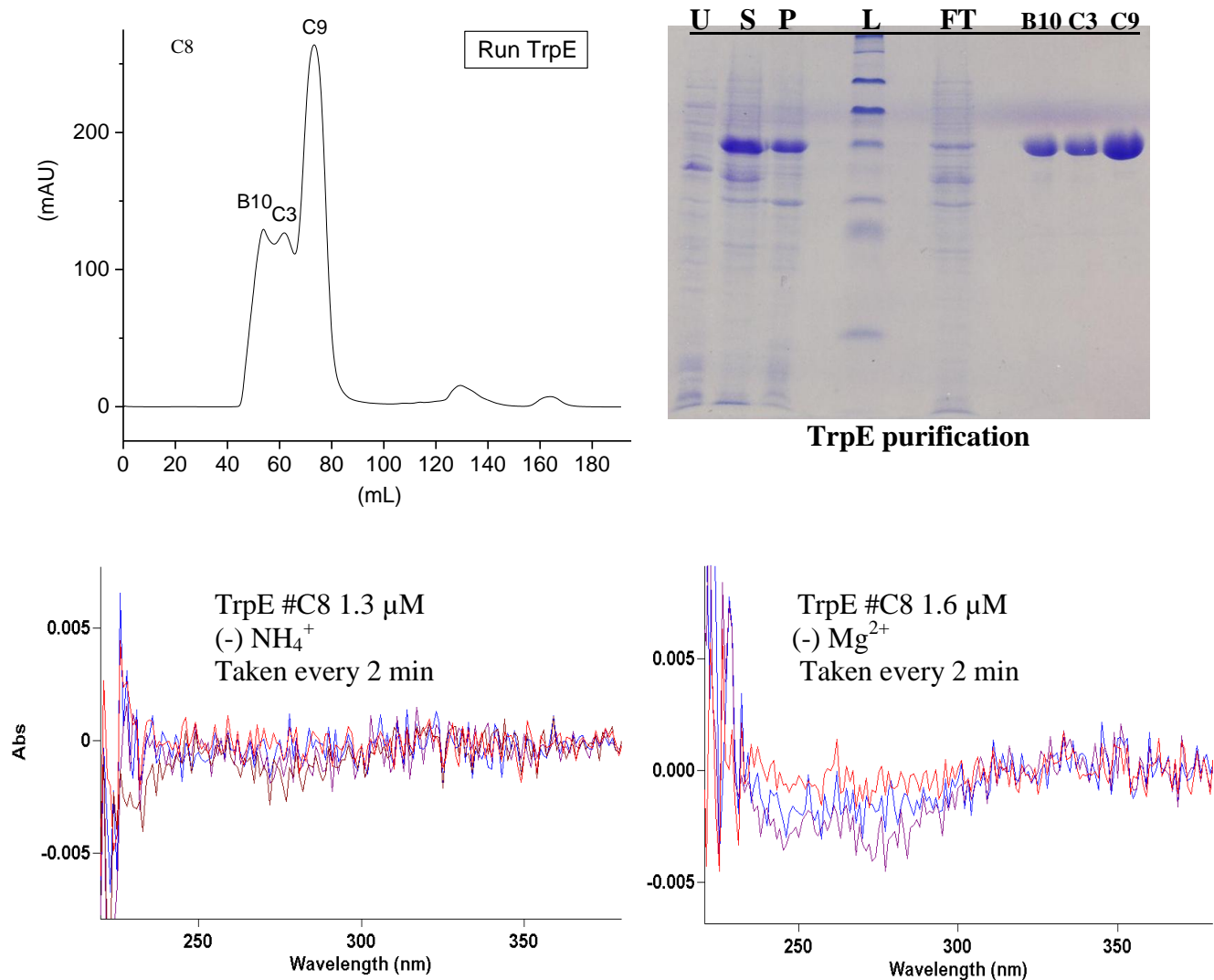
Figure 4.26 Purification of TrpE with extensive wash and test of side activities
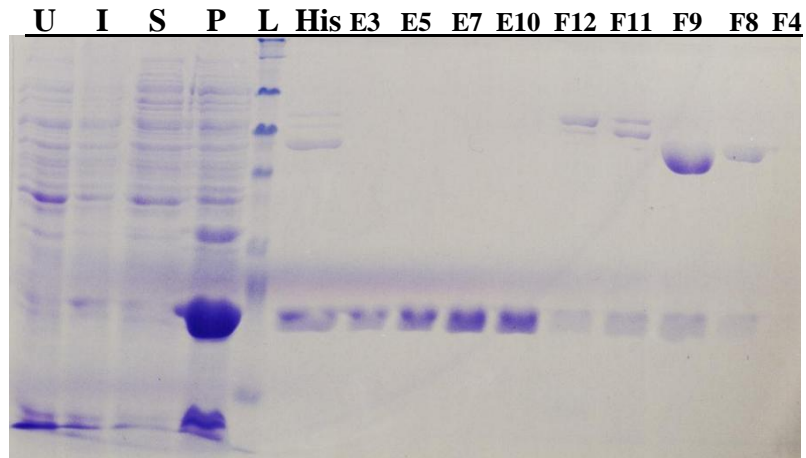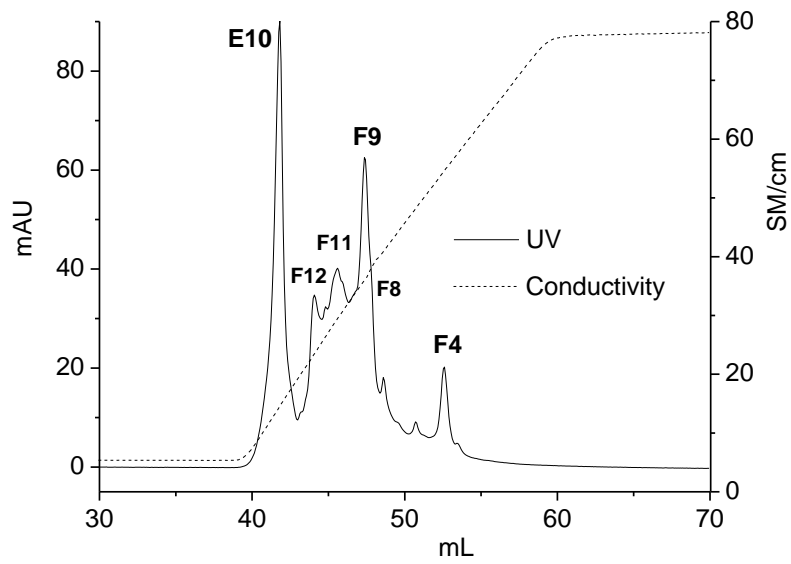
Figure 4.27 TrpG purification by anion exchange column Mono Q 5/50 GL.
Elution after several injections.

As expected, prephenate or phenylpyruvate formation can barely compete with AS activity (**Figure 4.29**). In addition, higher [TrpG]/[TrpE] was applied for optimization in this reaction since it was found that 1:1 was not the best available ratio for AS activity (**Figure 4.28**). In fact, ADIC synthase as well as AS synthase activity speeds up so much that ADIC appears to accumulate. The variation of each reaction component is summarized in **Figure 4.30**. Phenylpyruvate formation only becomes noticeable when chorismate is almost depleted. The accumulated ADIC seems reluctant to convert into anthranilate. The rate of chorismate consumption does not translate directly into the rate of anthranilate formation; regardless, the rate of anthranilate formation remains nearly constant until chorismate is consumed.
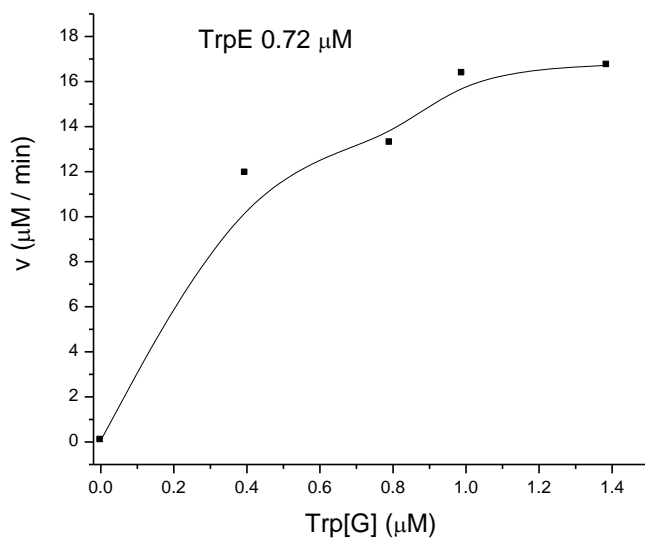
Figure 4.28 TrpE saturation by TrpG.
TrpE fixed at 720 nM while varying TrpG concentration to optimize AS activity. Mg^{2+} 10 mM, glutamine 2.4 mM, chorismate 198 µM, Tris 50 mM pH 7.5
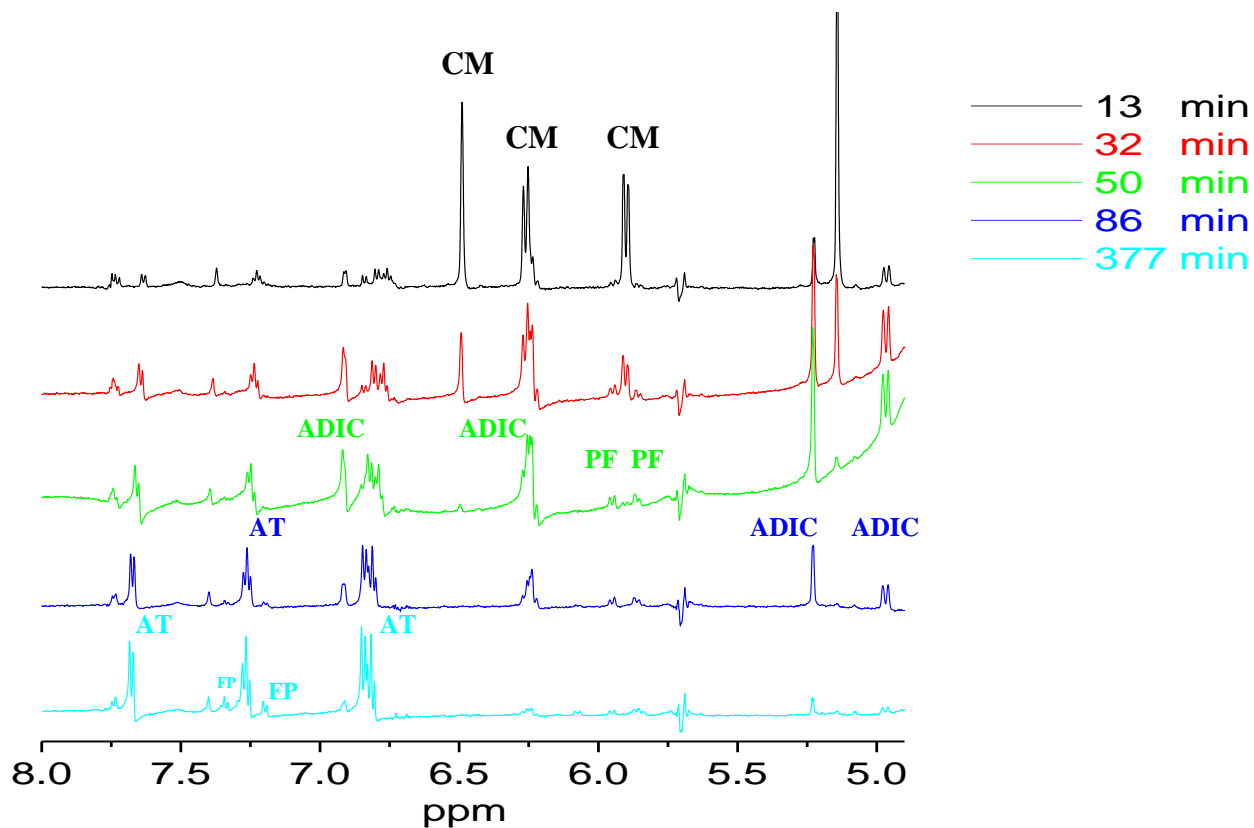
Figure 4.29 TrpEG reaction monitored by ^{1}H NMR
Tris 50 mM in D_2O, pH 7.1, Mg^{2+} 10 mM, glutamine 2.4 mM, chorismate 1.1 mM, TrpE 1.0 µM, TrpG 2.7 µM. TrpE and TrpG were pre-mixed and added to the mixture to initiate the reaction.
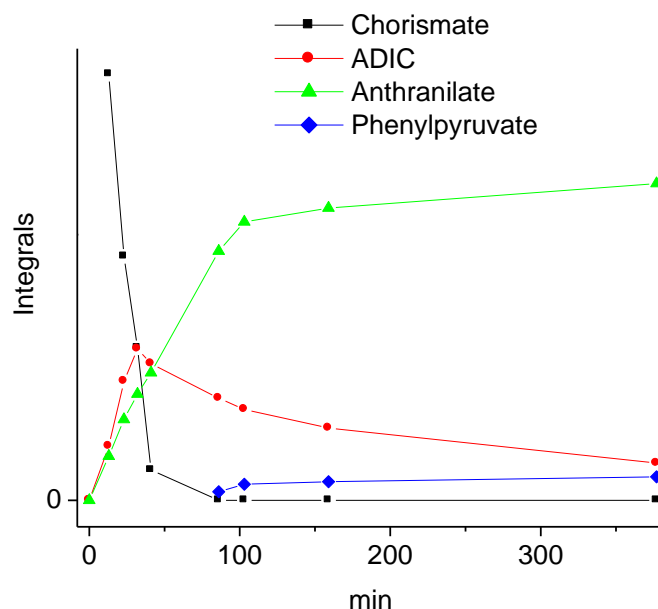
Figure 4.30 Variation in the amount of components with time in the course of the reaction

Oligomeric states of the subunits and the absence of complex

Most characterized anthranilate synthases are heterodimers or heterotetramers formed by TrpE and TrpG subunits [355]. All three of the known structures of anthranilate synthases are tetrameric [327-329]. Given the much better kinetics of TrpEG reaction, one would also expect mtAS to form a complex. Therefore gel filtration experiments were carried out to determine the oligomeric states of mtb AS.

TrpE elutes from the gel filtration column 10/300 GL as three separate peaks (**Figure 4.31a, b**) as in the large scale gel filtration experiments (**Figure 4.8**, **Figure 4.26**). At 200 mM salt concentration, the monomeric peak becomes more populated compared to that in the 50 mM salt buffer. TrpG tends to form very high MW, irreversible aggregates (**Figure 4.31c**). The monomeric TrpE and TrpG fractions are mixed and injected again in 200 mM salt buffer. One higher MW peak is observed in addition to the individual monomeric peaks (**Figure 4.31d**); however, the monomeric TrpE appears to generate the same peak (**Figure 4.31e**) and therefore the peak is not attributed to TrpEG complex. At 50 mM salt where complex formation might be more favorable, the TrpEG mix still comes off as simple addition of monomeric TrpG (**Figure 4.31f**) and the re-equilibrated TrpE (**Figure 4.31a**). The expected complex cannot be detected under these conditions.
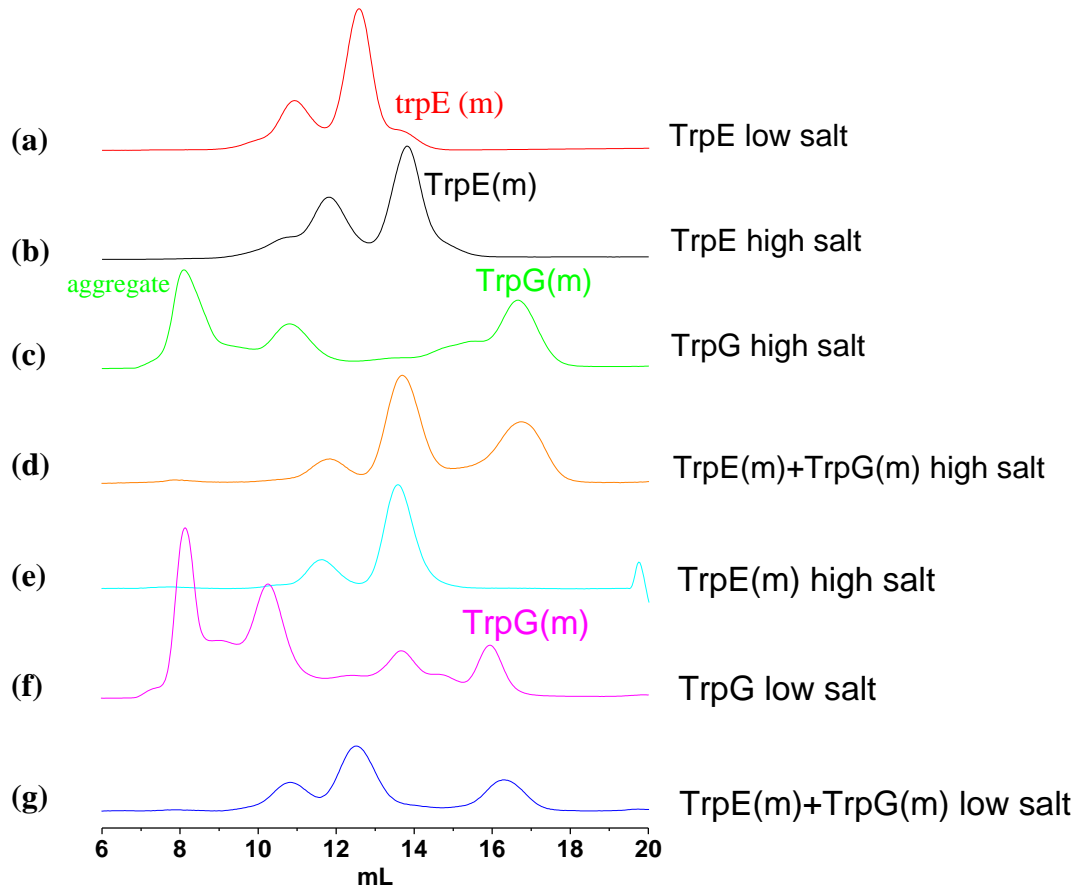
Figure 4.31 Gel filtration for determining the oligomeric status of TrpEG.

Low salt buffer: 50 mM Tris, pH 7.5, high salt buffer: 20 mM HEPES, 200 mM KCl, pH 8, Superdex 200 10/300 GL, flow 0.5 mL/min

The data from the AUC experiments were fit using a hetero-association model (**Figure 4.32**). The theoretical molecular weight of the heterotetrameric TrpEG complex is 164 kDa and a fitting result of 178 kDa was obtained. However, the fit is not good and MW ranging from 110-230 kDa can be obtained by fitting data at different speed and selecting different data points.
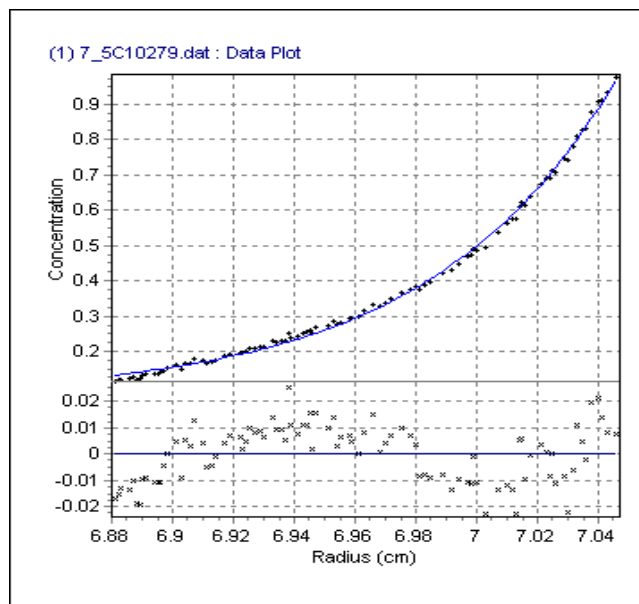
Figure 4.32 The fit curve from AUC experiment

Tryptophan-induced cooperativity in chorismate binding

Among the heterotetrameric anthranilate synthases, the ones from *Salmonella typhimurium* and *Serratia marcescens* display cooperativity [356,357] while that of *Sulfolobus solfataricus* does not [358]. Coincidently, the oligomeric assembly in AS-st and AS-sm are through the TrpE subunit while that in AS-ss is through the TrpG subunit, according to the crystal structures [327-329]. The cooperativity might give a hint into the oligomeric structure of mtb AS.

Cooperativity in chorismate binding is signified by the sigmoidal shape in the plot of velocity versus chorismate concentration as well as the upward curvature in the double reciprocal plot (**Figure 4.33**). Inhibition is most likely competitive with respect to chorismate because of the constant V_{max} value. The Hill coefficient *n* is close to 1 at zero tryptophan concentration (Trp=0) but > 1 in the presence of tryptophan, indicative of positive cooperativity. The double reciprocal data at Trp=0 is fit to the linear L-B equation, and k_{cat} of 36 min^{-1} and K_m of 15 μM are obtained. Compared with the previous data from **Figure 4.23**, the further saturated TrpE by TrpG acquires even lower K_m and larger k_{cat}, supporting the benefit from subunit interactions.

186
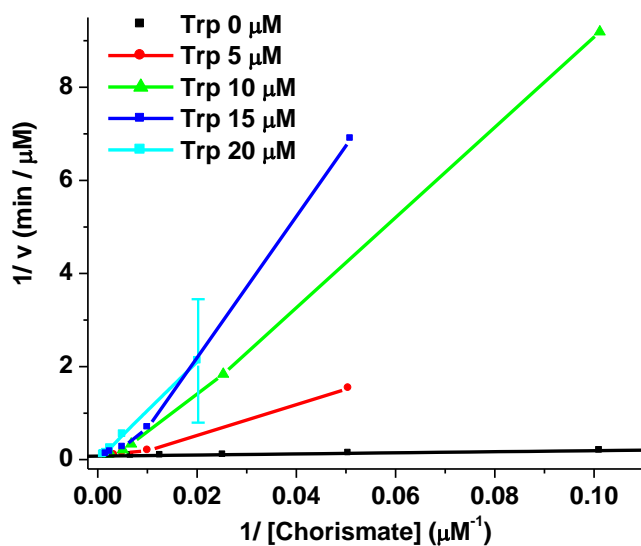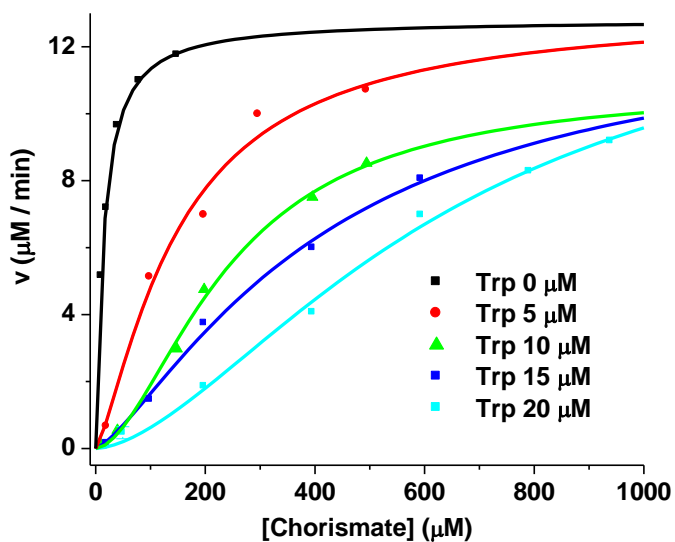
Figure 4.33 Cooperativity revealed by AS activity at different concentration of tryptophan.
TrpE 360 nM, TrpG 990 nM, Mg2+ 10 mM, glutamine 2.4 mM, Tris 50 mM pH 7.5. The v vs. [S] plot is fit to the Hill equation, $v_i = V_{max} [S]^n / (K^n + [S]^n)$

Table 3. Parameters obtained from tryptophan inhibition

[Trp] (μM)	V_{max} (μM/min)	K (μM)	n	Chi^2	R^2
0	12.8 ± 0.5	14.8 ± 1.2	1.07 ± 0.12	0.041	0.997
5	13.1 ± 3.0	151 ± 64	1.33 ± 0.48	0.594	0.982
10	10.9 ± 1.6	242 ± 49	1.76 ± 0.35	0.060	0.997
15	12.9 ± 4.2	418 ± 216	1.35 ± 0.31	0.088	0.996
20	14.2 ± 3.4	641 ± 202	1.65 ± 0.32	0.103	0.995

pabB expression and purification

Co-expression with *trpG* (*pabA*)

    pabB and *trpG* can both express, but the majority of the protein is found in the pellet (**Figure 4.34**). PabB also degrades into a band ~33 kDa. The yield of PabB from the His Bind resin is much lower than TrpG. Expression by the Arctic *E. coli* strain did not improve the solubility (**Figure 4.35**). Expression with pETDuet is not good.
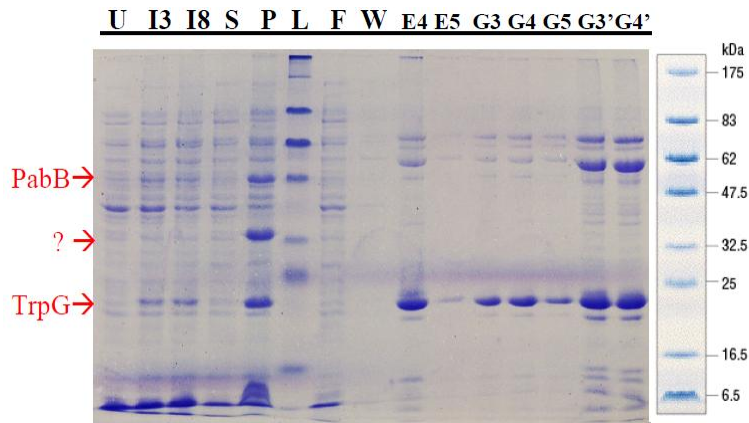
Figure 4.34 Co-expression of pabA and pabB.
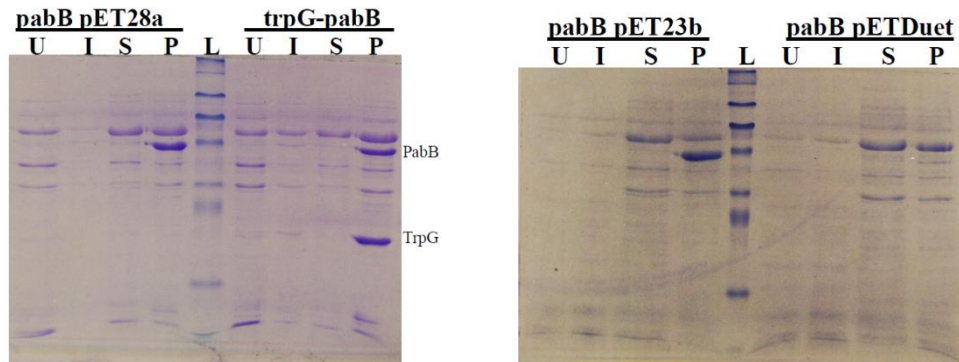
Figure 4.35 pabB expression by the Arctic strain

4.4 Discussion

Selectivity for amination over hydroxylation

Isochorismate synthase (ICS) and salicylate synthase (SS) catalyze the same chorismate transformations as ADIC synthase and anthranilate synthase (AS), respectively, except the nucleophilic substrate is water instead of ammonia. ICS, SS and AS are structurally homologous and the active sites are very similar [140,341], and they share a common mechanism of nucleophilic addition at C2 of chorismate [134]. Nevertheless, they are selective for the individual nucleophile and therefore specific for the desired hydroxylation or amination reaction. A lysine residue in the active site plays a crucial role in the selection for the water nucleophile [132,140], and the lysine appears to be conserved in the hydroxylation enzymes while a glutamine is seen in the corresponding position for amination enzymes (**Figure 4.42**). In spite of the specificity observed in general, there seems to be some space for promiscuity: ICS has been shown to display 2.5% ADIC synthase activity in the presence of ammonium salt [133], and the K193Q mutant of Irp9 displays AS activity; on the other hand, the Q to K mutant in TrpE does not render AS capable of doing hydroxylation [353]. In the characterization of mtAS products, no isochorismate or salicylate is observed, consistent with the notion that the active site lysine is required for hydroxylation to occur.

CM and PDT activities

CM and PDT activities comparable to the AS activity of mtAS are observed in the TrpE-only reaction. Chorismate is known to undergo nonenzymatic Claisen rearrangement into prephenate followed by decarboxylation-dehydration into phenylpyruvate [359]. Enzymatically, the two chemical steps are also involved in the biosynthesis of phenylalanine. The control experiments without mtTrpE show that uncatalyzed CM and PDT activities are indeed present, but the enzymatic CM and PDT activities are also observed. The question is whether such activities derive from TrpE active site or from impurity E. coli enzymes.

Subsequent assays with purer enzymes and with TrpG have removed most of the CM and PDT activities, which suggests the activities derive from E. coli impurities. **Table 4.5** - **Table 4.9** compare the kinetic parameters for CM and AS along with other chorismate enzymes from different organisms. CMs or PDTs are generally more efficient than AS as characterized. The k_{cat}/K_m values of *E. coli* CM or PDT can be up to three orders of magnitude larger than that of mtTrpE. Therefore trace amount of *E. coli* CM or PDT in the mtAS preparation can easily interfere with the AS assays. The biggest impurity band from TrpG preparation cannot account for the observed CM activity. This is possible since the efficient *E. coli* CM might well be invisible on the gel but still give rise to detectable CM activity. When total enzyme concentration is five

times lower in the reaction in **Figure 4.16**, PDT activity cannot be detected probably because the concentration of impurity PDT is too low to be active.

However, that the CM/PDT activity cannot be inhibited by phenylalanine leaves question for the real identity of the impurity protein. In the (-)Mg^{2+} reaction of **Figure 4.16c**, PDT activity does not occur with the same efficiency as in the (+)Mg^{2+} reaction of **Figure 4.16b**, and appears to be dependent on the AS activity; in other words, the two activities might have occurred within the same active site. Promiscuous CM activity has been found in PchB, MbtI and Irp9 that possess isochorismate pyruvate lyase (IPL) activity, and the promiscuity has been rationalized by the common sigmatropic process involved in the two activities [140,360]. Since the ADIC pyruvate lyase is analogous to IPL activity, it is also possible for AS to possess the promiscuous CM activity. PDT activity has also been reported for mutant Irp9 and AS-sm although not mechanistically correlated [353]. The PDT activity observed for AS could still be the promiscuity in the same active site and not simply a result of contamination.

Competing off the side activities by AS complex formation

While the observed CM/PDT activity for mtb TrpE cannot all be attributed to either contamination or promiscuity, it is clear that TrpG helps TrpE enhance the AS activity and compete off the CM/PDT activity, and this might reflect an *in vivo* state where chorismate needs to partition more into the tryptophan pathway than into other chorismate branching pathways. Dissociation of TrpG from TrpE results in the loss of chorismate binding affinity by more than ten-fold, and the remaining bound chorismate can be further eradicated by excess amount of tryptophan. Our experiment suggests that it will be extremely difficult for mtTrpE to produce significant amount of anthranilate *in vivo* in the absence of TrpG because the TrpE-only reaction requires high pH and high ammonium salt for a reasonable rate. Suppression of anthranilate formation can be achieved by disrupting the interactions between TrpE and TrpG. Intriguingly, mt-trpG is predicted to serve as pabA involved in the PABA biosynthesis as well (**Figure 4.36**), unlike the TrpG in *E. coli* and many other organisms that are faithful to tryptophan biosynthesis. This means mtTrpE has to compete with PabB for the amidotransferase. The absence of a tight complex between mtTrpE and TrpG could be the result of the versatility of TrpG. There exist diverse ways of regulation for tryptophan production among different species [361,362], and for *M. tuberculosis*, association and dissociation of the AS and PABA synthase complexes could also play a role.
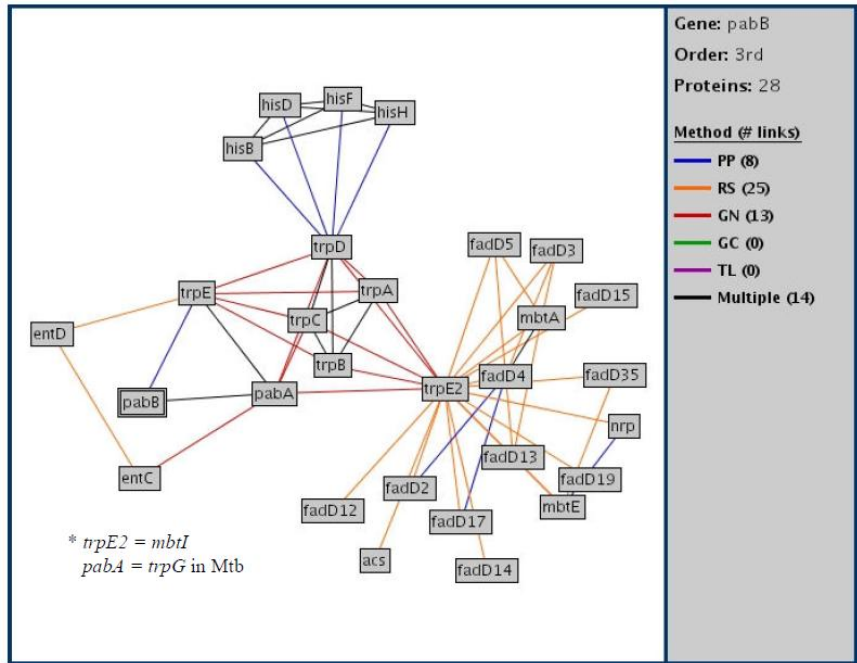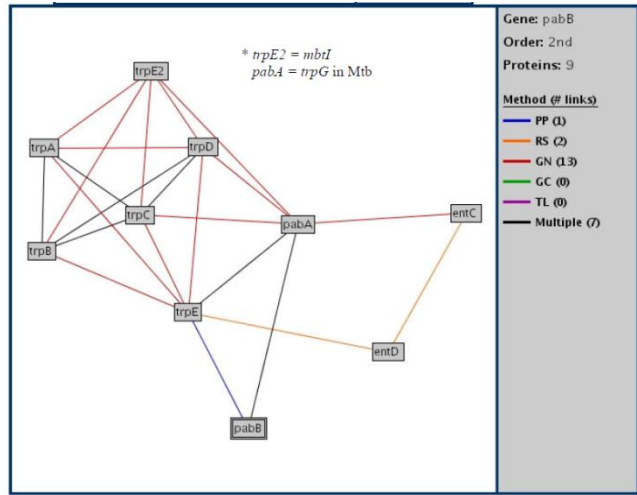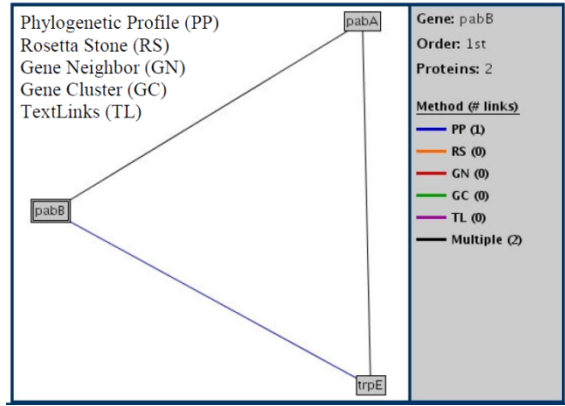
Figure 4.36 Functional linkage to *pabA* (*trpG*) suggested by ProLink.

AS "turned on" by complex formation — a common mechanism?

It is widely accepted in the literature that the TrpE subunit by itself can still perform anthranilate synthesis in the presence of ammonium salt. TrpG appears to only provide an alternative way of making anthranilate by utilizing glutamine as a source of nitrogen. Although better AS activity by TrpEG compared to TrpE alone has been occasionally mentioned, and subunit communication has been shown to mediate feedback inhibition by tryptophan, the importance of complex formation for the AS function has not been paid much attention. In our experiment with mtAS, the complex becomes much more important because of the poor AS activity of TrpE alone. It can be seen that for *E. coli* and *S. typhimurium*, although there is a trend that TrpE-alone activity is not as good as TrpEG, the efficiency of the TrpE-alone reaction is still ten-fold higher than that of mtTrpE (**Table 4.6**, **Table 4.7**), and that is why the absence of the complex does not seem to be vital in the former two cases. Paradoxically, *E. coli* and *S. typhimurium* have tightly controlled trp operons with TrpEG expressed stoichiometrically, while *M. tuberculosis* who seems to be desperate for TrpEG complex lacks such an organization.

Although such activation upon complex formation was not generalized, the trend can be observed from **Table 4.5**-**Table 4.9** and from the following examples. It was reported that specific activity of AS from *S. typhimurium* increased from 462 to 5000 nmole/min/mg upon complexation with the subunit containing the amidotransferase [363]. The *E. coli* AS has a higher apparent K_m for TrpE-alone than for the complex (**Table 4.6**). TrpE from *S. marcescens* was reported to have about 25 % of the specific activity obtained with the TrpEG complex as well as a poorer K_m [364]. The hysteresis (lag before steady state in kinetics) observed for AS of *Bacillus subtilis* was attributed to the re-association of the TrpEG complex [365], implying that the complex enhances AS activity. More recently, the α subunit of rice AS (analogous to TrpE) was reported to require the β subunit (analogous to TrpG) for maximum enzymatic activity although K_m for the complex was not smaller in this case [366].

Glutaminase activity of TrpG "turned on" by complex formation

The benefit of complex formation appears to be better known for promoting the glutaminase activity of TrpG. The binding sequence for AS was determined to be chorismate first followed by glutamine, and so the glutaminase activity is poor before the complex is assembled [367]. In fact, it was suggested generally for enzymes involving glutamine amidotransferases that glutaminase as well as the synthase activity are optimum only when the two subunits are coupled. The partnership between mtTrpE and TrpG is obvious given the fact that TrpE is activated by TrpG and TrpG is stabilized by TrpE. The benefit is mutual, and one

would expect that complex formation exerts conformational changes on both subunits, signaling substrate binding and catalysis on both sides.

Currently all the available AS crystal structures are heterotetrameric complexes. A structure of autonomous TrpE is not available for studying the structural change upon complexation. The structure of a close relative, PabB, from *E. coli*, is available without PabA, but the PabAB complex structure is not available because of the absence of a tight complex. mtTrpE in this aspect might resemble *E. coli* PabB than TrpE, and might be able to provide structural information for a different type of AS. However the first crystal screening yielded only precipitation; the oligomers of different sizes observed in TrpE solution might have hindered crystallization. Nevertheless, TrpE is very stable at 4°C and further screening with different conditions will be worth trying. A crystallographic study of PLP synthase complex and the individual subunits demonstrates the effect of subunit interactions [368].

Sequence alignment of mtTrpE and smTrpE based on secondary structure prediction reveals the same region of deletion as in ecPabB from that involved in subunit interactions with TrpG in smTrpE (residues 91-96, mtTrpE, or residues 65-109, smTrpE, **Figure 4.37, Figure 4.6**), suggesting one reason why mtTrpE or ecPabB does not form as tight a complex with its amidotransferase as smTrpE.

```
mtTrpE    1 MHADLAATTSREDFRLLAAEHRVVPVTRKVLADSETPLSAYRKLA  45
smTrpE    1 ------------------------LKVQASY-RGDPTTLFHQLC  19

mtTrpE   46 ANRPGTFLLESAENGRSWSRWSFIGAGAPTALTVREGQAVWLGAV  90
smTrpE   20 GARPATLLLESAEINDKQNLQSLLVIDSALRITALGHTVSVQALT  64

mtTrpE   91 PKD----------------------------------------APT  96
smTrpE   65 ANGPALLPLLDEALPPEVRNQARPNGRELTFPAIDAVQDEDARLR 109

mtTrpE   97 GGDPLRALQVTLELLATADRQSEPGLPPLSGGMVGFFAYDMVRRL 141
smTrpE  110 SLSVFDALRTILTL------VDSPADEREAVMLGGLFAYDLVAGF 148

mtTrpE  142 ERLPERAVDDLCLPDMLLLLATDVAAVDHHEGTITLIANAVNWNG 186
smTrpE  149 ENLPALRQ-DQRCPDFCFYLAETLLVLDHQRGSARLQASVFSEQA 192

mtTrpE  187 TDERVDWAYDDAVARLDVMTAALGQPLPSTVATFSRPEPRHRAQR 231
smTrpE  193 SE------AQRLQHRLEQLQAELQQPPQP-IPHQKLENMQLSCNQ 230

mtTrpE  232 TVEEYGAIVEYLVDQIAAGEAFQVVPSQRFEMDTDVDPIDVYRIL 276
smTrpE  231 SDEEYGAVVSELQEAIRQGEIFQVVPSRRFSLPCP-APLGPYQTL 274
```

Figure 4.37 Deletion of sequence for subunit interactions in mtTrpE

193

Table 4.5 Chorismate enzymes from *M. tuberculosis*

	Name	k_{cat} (min^{-1})	K_m (µM)	Effectors	Ref
Isochorismate synthase (IC)	MbtI	3.1	34		140
Isochorismate pyruvate lyase (IPL)	MbtI	2.1	2.6		140
Anthranilate synthase (AS)	TrpEG	36	15	Trp	Huei
	TrpE	3.2	570	Trp	Huei
p-Aminodeoxychorismate synthase (ADCS)	PabB	Uncharacterized			
Chorismate mutase (CM) Rv1885c	*AroQ	3600	500		369
				Trp	370
		1560	1200	Phe, Tyr, Trp -activators, low conc. -inhibitors, high conc.	371
		3000	180	Phe, Tyr, Trp < 10%	372
Chorismate mutase (CM) Rv0948c		14	5000		371
Prephenate dehydratase (PDT)	PheA	4125[#]	500	Phe, Tyr, Trp activators	373
Chorismate lyase (CL)					

* Only V_{max} was reported, based on MW 33000

Table 4.6 Chorismate enzymes from *E. coli*

		k_cat (min⁻¹)	K_m (μM)	Effectors	Ref
Isochorismate synthase (IC)	MenF	213 ± 5	192 ± 7		[132]
	EntC				
Anthranilate synthase (AS)	TrpEG		5.5		[374]
	TrpE	146	30		[374,375]
p-Aminodeoxychorismate synthase (ADCS)	PabAB-Gln	31.8	13 ± 1		[134]
	PabB	1.92	58 ± 8		[134]
Chorismate mutase (CM)	AroQ	3840	390		[376]
P-protein N-terminal 1-113		2482	300	Tyr	[377]
P-protein, chorismate		3300	64		[378]
Prephenate dehydratase (PDT), P-protein, prephenate	PheA	2160	44		[378,379]
Chorismate Lyase (CL)					

Table 4.7 Chorismate enzymes from *Salmonella typhimurium*

		k_{cat} (min^{-1})	K_m (μM)	Ref
Anthranilate synthase (AS)	TrpEG-Gln	720	2.3	[356]
	TrpEG-NH$_3$	558	4.0	[332]
	TrpE	11	110	[346]
Chorismate mutase (CM)	*AroQ	534	142	

Table 4.8 Chorismate enzymes from *Serratia marcescens*

		k_{cat} (min^{-1})	K_m (μM)	Ref
Anthranilate synthase (AS)	TrpEG-Gln		3.6	[357]
	TrpEG-NH$_3$		15	[357]
	TrpE		~ 100	[364]

Table 4.9 Chorismate enzymes from *Bacilus subtilis*

		k_{cat} (min^{-1})	K_m (μM)	Ref
Anthranilate synthase (AS)	TrpEG-Gln			
	TrpEG-NH$_3$			
	TrpE			
Chorismate mutase (CM)	AroQ	2460	74± 8	[376]

Oligomeric states and cooperativity

While cooperativity is observed for AS complexes from *S. typhimirium* and *S. marcescens*, it is not observed for TrpE alone, and is thought not possible for the heterodimeric AS [358]. PabAB is not known to display cooperativity. It is thus unique for mtAS to give only loose complex like PabAB but display cooperativity like those tight tetrameric anthranilate synthases. The mtAS complex might be a loose heterotetramer. In addition, the multiple oligomeric states for TrpE are not seen for other TrpEs. Since the increase in [TrpE] does not give a linear increase in AS activity, the oligomeric forms of TrpE might be inactive. It is not clear whether *M. tuberculosis* will adopt a dormant state of TrpE.

(TrpG)"(Inactive)

k^{-1} ~ a few days at 4C

TrpG(Inactive)

$+ NH_4^+$ - Pyruvate

+ Chorismate $Mg^{2+}2+$ + Chorismate   - Anthranilate

TrpE $Mg^{2+}2+$ cat. cycle

+ Glutamine   TrpG

(NH_4^+) catalytic cycle (TrpE.TrpG)$_n$ ⇌ TrpE ⇌ (TrpE) ⇌ (TrpE)

(Very active)   (Active)   (Inactive)   (Inactive)

- Pyruvate

- Anthranilate   + Trp   + Trp

- Glutamate

(TrpE.TrpG)$_n$(Trp)$_m$ TrpE.Trp

(Inactive)   (Inactive)
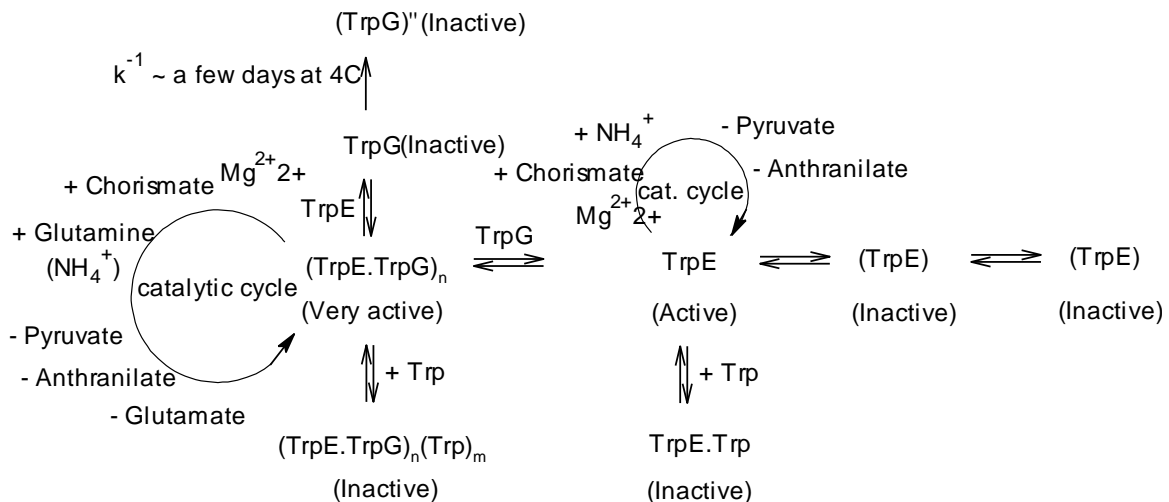
Figure 4.38 Summary of *M. tuberculosis* anthranilate synthase activities

mtAS is also different from those of *S. typhimurium* and *S. marcescens* in that not all the trp genes are clustered in the operon (**Figure 4.39**). *trpG* in mtb is 1800 kb upstream of trpE. Unlinked trpE/G is also observed in *Bacillus subtilis*, *Pseudomonas putida* and *Acinetobacter calcoacericus*. Interestingly, they are reported to give heterodimeric AS complexes, and the three genera are also known to have a single glutamine amidotransferase for both AS and PABS [326]. The gene organization, the commitment in multiple biosynthetic pathways and the oligomeric states of the enzymes might actually be correlated.
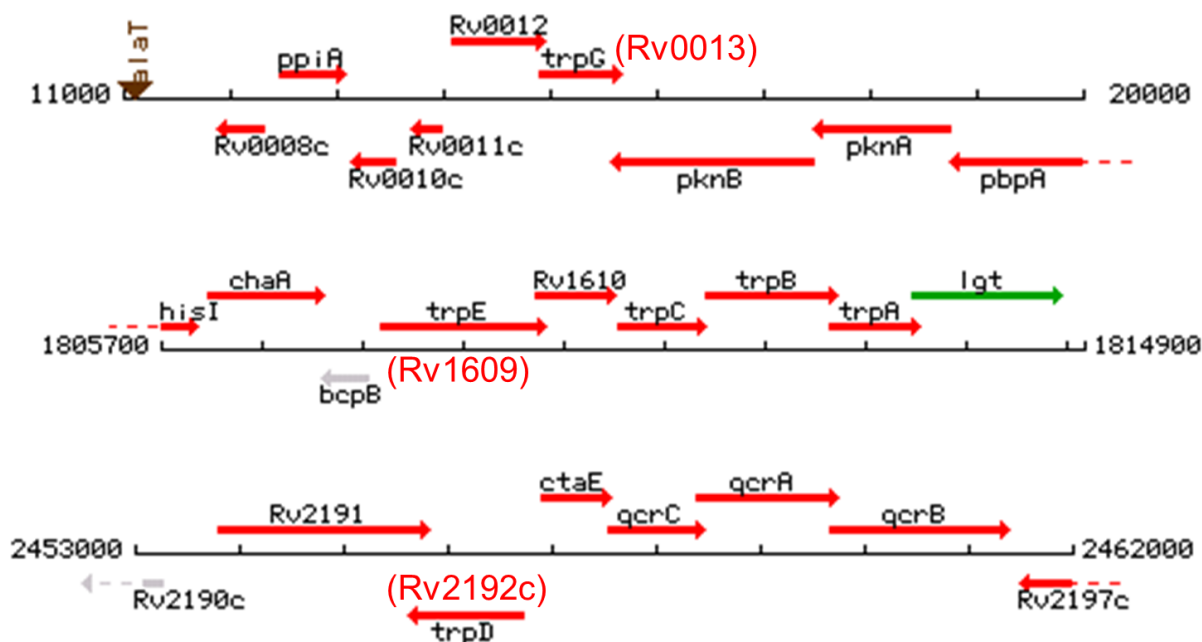
Figure 4.39 The *trp* genes in *M. tuberculosis*

Mechanism of pyruvate elimination

To isolate the intermediate from enzymatic reaction, and to investigate how anthranilate synthases cleave the enolpyruvyl group from ADIC, researchers have attempted to mutate residues within the active site, but have not been able to convert an AS into a pure ADIC synthase although ADIC was successfully isolated from the mutant enzyme reaction [332]. There are only two natural ADIC synthases reported, SgcD which is involved in the synthesis of the benzoxazolinate moiety of C-1027, an antitumor antibiotic from streptomyces [350], and PhzE, involved in the biosynthesis of phenazines which are redox-active secondary metabolites produced by bacteria as virulence factors, antibiotics or respiratory pigments [380-382]. The sequence comparison between the AS and ADIC does not reveal residues unique for pyruvate elimination (**Figure 4.40**).

The mechanism of pyruvate elimination from isochorismate in PchB and MbtI is thought to be a concerted sigmatropic process, similar to that of CM and CPL, and the decomposition process in solution [140,359,360]. In principle, pyruvate elimination from ADIC will be analogous to this mechanism [383]. However, it has been puzzling how some enzymes eliminate the enolpyruvyl group while others retain it, but all of them share a conserved active site. That MbtI eliminates pyruvate at pH 8 and retains pyruvate at pH 7 suggests that ionization state of the

enzyme dictates the fate of the enolpyruvyl group. In the mtTrpE reaction, such pH dependence is not obvious; at pH 7.1 compared to pH 8.5, a slower AS activity was observed, but ADIC was not accumulated. The TrpEG reaction at pH 7.1 accelerates AS activity but at the same time accumulates ADIC, which is not observed for other characterized AS [333]. The accumulation of ADIC persists long after chorismate is depleted, indicating the pyruvate lyase activity is somehow inactivated or inhibited in this last phase of the reaction. It could be due to the dissociation of TrpEG complex, or the inhibition from the accumulated AS reaction products, anthranilate, pyruvate, or even the degradation product of ADIC, the amino analogue of isoprephenate [384]. While it is obvious that TrpG improves ADIC synthase activity, it will be interesting to know whether subunit interactions also mediate pyruvate elimination. Further structural investigation into differences between pyruvate-retaining and pyruvate-eliminating enzymes as well as between amidotransferase-dependent and amidotransferase-independent enzymes might shed more light into the modulation of these catalytic activities.

Sequence alignment and functional implications

```
              10         20         30         40         50         60         70         80         90
TrpE-st/1-520   MQTPKPTLELLTCDAAYRENPTALFHQVCGDRPATLLLESADIDSKDDLKSLLLVDSALRITALGDTVTIQALSD-NGASLLPLLDTALP----
TrpE-sm/1-519   MNT-KPQLTLLKVQASYRGDPTTLFHQLCGARPATLLLESAEINDKQNLQSLLVIDSALPITALGHTVSVQALTA-NGPALLPVLDEALP----
TrpE-Mtb/1-516  ----MHADLAATTSRED----FRLLAAEHRVVPVTRKVLADS----ETPLSAYRKLAANRPGTFLLESAENG-RSWSRWSFIGAGAP----
TrpE-ss/1-421   ------------------------------MEVHPISEFASPFEVFKCIERDFKVAGLLESIGGP-QYKARYSVIAWSTN----
EntC-ec/1-391   ------------------------------MDTSLAEEVQQTMATLAPN-RFFFMSPYRSFTTSGCFA-RFDEPAVNGDSPDSPFQQ
EntC-Mtb/1-372  ------------------------------VSAHVATLHPEPPFALCGPRGTLIARGVRT-RYCD--------------
PabB-ec/1-453   ----MKTLSPAVITLLWRQDAAEFYFSRLSHLPWAMLLHSGYADHPYSRFDIVVAEPICTLTTFGKETVVSESEK-RTTTT-----
PabB-Mtb/1-458  -------MNLAWELSTRTKSPRSHLRCENPQFCQARTVRIDRLGDLGGAPAVLRAVGRATSRLDLPPPAALTGEWFGALAVIAPSVS--
MenF-ec/1-431   -------------MQSLTTALENLLRHLSQEIPATPGIR-|-VIDIPFPLKDAFDALSWLASQQTYPQFYWQQRNGDEEAVVLGAITRFTS----
MbtI-Mtb/1-450  -----MSELSVATGAVSTASSSIPMPAGVNPADLAAELAAVVTESVDEDYLLYECDGQWVLAAGVQAMVELDSDELRVIRDGVTRRQQWS----
```

```
        100        110        120        130        140        150        160        170        180        190
AGVENDVLPAGRVLRFPPVSPLLDEDARLCSLSVFDAFRLLQGVVNIPTQEREAMFFG---GLFAYDLVAGFEALPHLEAGNNC-PDYCFYLAETLMVIDHQKKS
PEVRNQARPNGRELTFPAIDAVQDEDARLRSLSVFDALRTLLTLVDSPADEREAVMLG---GLFAYDLVAGFENLPAVRQDQRC-PDFCFYLAETLLVLDHQRGS
TALT---VREGQAVWLGAVP-KDAPTGGDPLRALQVTLELLATADRQSEPGLPFLSGGMVGFFAYDMVRRLERLPERAVDDLCLPDMLLLLATDVAAVDHHEGT
GYLK---------------------IHDDPVNILNGYLKDLKLADIPG-----LFKGGMIGYISYDAVRFWEKIRDLKPAAEDWPYAEFFTPDNIIIYDHNEGK
---------------------KLAALFADAKAQGIKNPVMVGAIPFDPRQPSSLYIPESWQSFSRQEKQASARRFTRSQSLNVVERQAIPEQTTFE
---------------------VRAAQAALRSGTAPILLGALPFDVSRPAALMVPDGVLRARKLPDWPTG-----PLPKVRVAAALPPPADYL
--------------------DDPLQVLQQVLDRADIR-PTHNEDLPFQGGALGLFGYDLGRRFESLPEIAEQDIVLPDMAVGIYDWALIVDHQRHT
--------------------IQPVSGDDVFSGPPGTGGPDATGAVGGGWVLSYPDAGADGRPHRIPEAAGGWTDCVLRRDRDGQWWYESLSG
--------------------LDQAQRFLRQHPEHADLRIWGLNAFDPSQGNLLLPRLEWRRCGGKATLRLTLFSESSLQHDAIQAKEFIATLVSI
--------------------GRPGAALGEAVDRLLLETDQAFGWVAFEFGVHRYGLQQRLAPHTPLARVFSPRTRIMVSEKEIRLFDAGIRHREAID
```

```
200        210        220        230        240        250        260        270        280        290        300
TRIQASLFTASDREK-----QRLNARLAYLSQQLTQPAP-PLPVTVPVPDMRCECNQSDDAFGAVVRQLQKAIRAGEIFQVVPSRRFSLPCPS---PLAAYYVL
ARLQASVFSEQASEA-----QRLQHRLEQLQAELQQPPQ-PIPHQKLENMQLSCNQSDEEYGAVVSELQEAIRQGEIFQVVPSRRFSLPCPA---PLGPYQTL
ITLIANAVNWNGTDERVDWAYDDAVARLDVMTAALGQPLPSTVATFSRPEPRHRAQRTVEEYGAIVEYLVDQIAAGEAFQVVPSQRFEMDTDV--DPIDVYRIL
VYVNADLSSVGGCGDIG--------EFKVSFYDESLN-----------------KNSYERIVSESLEYIRSGYIFQVVLSRFYRYIFSG--DPLRIYYNL
QMVARAAALTATPQ----------------------------VDKVVLSRLIDITTDAAIDGVLLER---------------L
TRIGRARDLLAAFDG-----------------------------PLHVVLARAVQLTADAPLDARVLLRR---------------L
VSLLSHNDVNARRAWLESQQFSPQEDFTLTSDWQSN---------------MTREQYGEKFRQVQEYLHSGDCYQVNLAQRFHATYSG--DEWQAFLQL
APIADWLASALATTRAS----VARPAPACRIDWEP---------------ADRAAHRDGVLACLEAIGAGEVYACVCTQFAGTVTG--SPLDFFIDG
KPLPGLHLTTTREQHWP----------------------DKTGWTQLIELATKTIAEGELDKVVLARATDLHFASPVNAAAMMAAS
RLLATGVREVPQSRSVDVS------------------DDPSGFRRRVAVAVDEIAAGRYHKVILSRCVEVPFAID--FPLTYRLG
```

```
310        320        330        340        350        360        370        380        390        400
KKSNPSPYMFFMQDN-------DFTLFGASPESSLKYDAASRQIEIYPIAGTRPRGRRADGTLDRDLDSRIELDMRTDHHKELSEHLMLVDLARNDLARICTPGS
KDNNPSPYMFFMQDD-------DFTLFGASPESALKYDAGNRQIEIYPIAGTRPRGRRADGSLDLDLDSRIELEMRTDHHKELAEHLMLVDLARNDLARICQAGS
RVTNPSPYMYLLQVPNSDGA-VDFSIVGSSPEALVTVHEG--WATTHPIAGTRWRGRTDD-----EDVLLEKELLADDKERAEHLMLVDLGRNDLGRVCTPGT
RRINPSPYMFYLKFD-------EKYLIGSSPELLFRVQDN--IVETYPIAGTRPRGADQE-----EDLKLELELMNSEKDKAEHLMLVDLARNDLGKVCVPGT
IAQNPVSYNFHVPLADGG----VLLGASPELLVARSGN--RFSSIPLAGSAPRQPDEV-----LDREAGNRLLASEKDRHEHELVTQAMKEVLRERSS--E
VVADPTAYGYLVDLTSAGNDDTGAALVGASPELLVARSGN--RVMCKPFAGSAPRAADPK-----LDAANAAALASSAKNRHEHQLVVDTMRVALEPLCE--D
NQANRAPFSAFLRLEQG------AILSLSPERFILCDNS--EIQTRPIKGTLPRLPDPQ-----EDSKQAVKLANSAKDRAENLMIVDLMRNDIGRVAVAGS
FGRTAPSRSAFVAGPWG------AVASLSPELFLRRRGS--VTSSPIKGTLPDAPPS-----------ALRASAKEVAENIMIVDLVRNDLGRVAVTGT
RRLNLNCYHFYMAFDGEN-----AFLGSSPERLWRRRDK--ALRTEALAGTVANNPDDK-----QAQQLGEWLMADDKNQRENMLVVEDICQRLQADTQTLD
RRHNTPVRSFLLQLGGIR------ALGYSPELVTAVRADG-VVITEPLAGTRALGRGPA------IDRLARDDLESNSKEIVEHAISVRSSLEEITDIAEPGS
```

```
410        420        430        440        450        460        470        480        490        500
RYVADLTKVDRYSYVMHLVSRVVG-ELRHDLDALHAYRACMNMGTLSGAPKVRAMQLIADAEGQRRGSYGGAVGYFTAHGDLDTCIVIRSALVENGI--
RYVADLTKVDRYSFVMHLVSRVVG-TLRADLDVLHAYQACMNMGTLSGAPKVRAMQLIAALRSTRRGSYGGRVGYFTAHRHLDTCIVIRSAYVEDGH--
VRVEDYSHIERYSHVMHLVSTVTG-KLGAGDTLDAVTACFPAGTVSGAPKPMAMNIIETLEEYKRGPYAGAVGFISADGNAEFAIAIRTALMRNGT--
VKVPELMYVEKYSHVQHIVSKVIG-TLKKKYNALNVLSATFPAGTVSGAPKPMAMNIIETLEEYKRGPYAGAVGFISADGNAEFAIAIRTAFLNKEL--
LHVPSSPQLITTPTLWHLATPFEG-KANSQENALTLACLLHPTPALSGFPHQAATQVIAELEPFDRELFGGIVGWCDSEGNGEWVVTIRCAKLRENQ--
LTIPAQPQLNRTAAVWHLCTAITGRLRNISTTAIDLALHPTPAVGGVPTKAATELIAELE-GDRGFVAGAVWGCDGRDGDGHWVVSIRCAQLSADRRA
VKVPELFVVEPFPAVHHLVSTITA-QLPEQLHASDLLRAAFPGGSITGAPKVRAMEIIDELEPQRRNAWCGSIGYLSFCGNMDTSITIRTLTAINGQ--
VTVPELLVVRPAPGVWHLVSTVSA-RVPLEEPMSALLDAAFPPASVTGTPKLRARQLISQWERYRRGIYCGTVGLASPVAGCELNVAIRTVEFDTAG-N
VLPPQVLRLRKVQHLRRCIWTSLN-----KADDVICLHQLQPTAAVAGLPRDLARQFIARHEPFTREWYAGSAGYLSLQ-QSEFCVSLRSAKISGNV--
AAVIDFMTVRERGSVQHLGSTIRA-RLDPSSDRMAALEALFPAVTASGIPKAAGVEAIFRLDECPRGLYSGAVVMLSADGGLDAALTLRAAYQVGGR--
```

```
510        520        530        540        550
ATVQAGAGIVLDSVPQSEADETRNKARAVLRAIATAHHAQETF------
RTVQAGAGVVQDSIRRREADETRNKARAVLRAIATAHHAKEVF------
AYVQAGGGVVADSNGSYEYNEARNKARAVLNAIAAAETLAAPGANRSGC
LRIHAGAGIVYDSNPESEYFETEHKLKALKTAIGVR-------------
VRLFAGAGIVPASSPLGEWRETGVKLSTMLNVFGLH-------------
ALAHAGGGIVAESDPDDELEETTTKFATILTALGVG-------------
IFCSAGGGIVADSQEEAEYQETFDKVNRILKQLEK--------------
AVLGVGGGITADSDPDAEWAECLHKAAPIVGLPAATRTTPARLASKVR-
VRLYAGAGIVRGSDPEQEWQEIDNKAAGLRTLLQME-------------
TWLRAGAGIIEESEPEREFEETCEKLSTLTPYLVARQ------------
```

Figure 4.40 Sequence alignment of TrpE, EntC, PabB, MenF and MbtI

Unique for MenF, EntC and MbtI:

ecMenF K190 which activates water for hydroxylation on chorismate C2, is Q for all amination enzymes, K for all hydroxylation enzymes although this particular alignment did not align K very well for EntC.

![green bar]

Unique for PabB:

ecPabB K274 which attacks chorismate C2, is A for all other homologs.

![magenta outline box]

Trp binding site:

stTrpE fragment P291-Y-292-M293 (**Figure 4.41**) is conserved for all TrpE. Y455 and D463 are conserved in all TrpE except ssTrpE (which lacks cooperativity) and also conserved in ecPabB (which binds Trp).

![green outline box]

This site is occupied by W55 in MenF-ec and by Y48 in MbtI-Mtb.
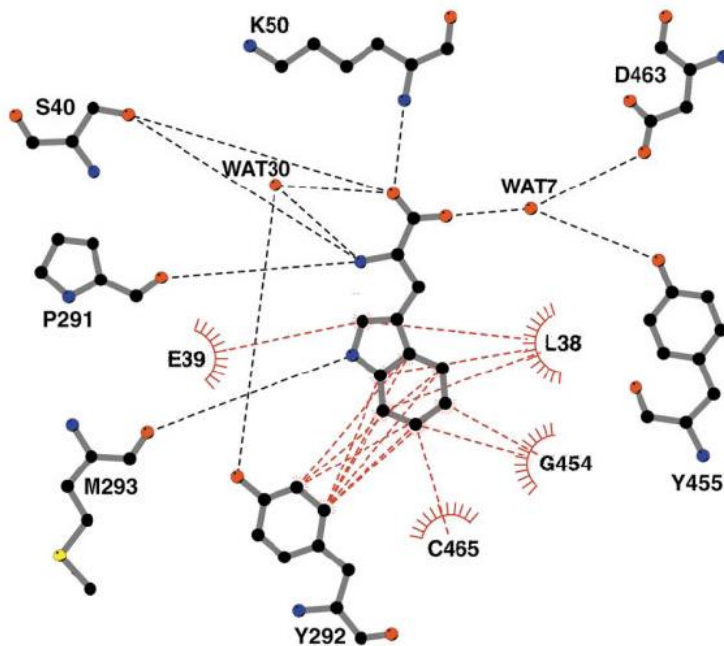
Figure 4.41 Tryptophan binding site in stTrpE

Unique for pyruvate elimination (**Figure 4.42**) [132]?
ecMenF A344 is also A in EntC, and T for all pyruvate eliminating enzymes, S for PabB, T for PhzE (ADC synthase).
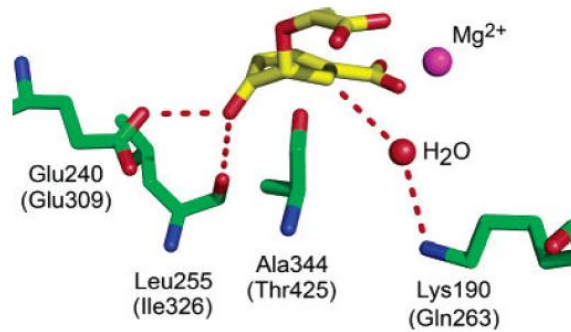
Figure 4.42 Active site of ecMenF and smTrpE

Active site residues (**Figure 4.43**)
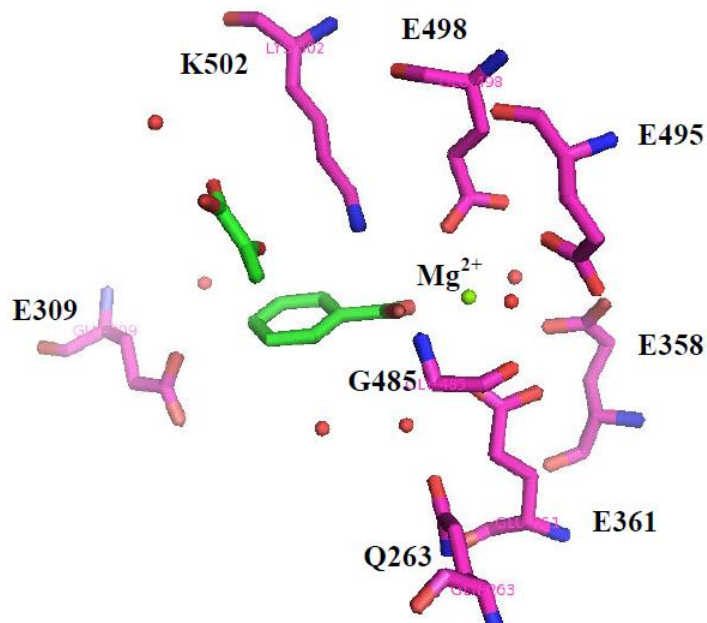
Figure 4.43 smTrpE active site

The PabB reaction requires the formation of a covalent intermediate between K274 and C2 of chorismate, which eliminates the 4-OH [131,134,334]. Isochorismate synthase or anthranilate synthase does not involve this covalent intermediate and an alanine replaces the lysine in their sequences. The annotated mtPabB has a lysine at this position suggesting the same function as ecPabB. Moreover, for the hydroxylation enzymes MenF, EntC and MbtI, another lysine residue corresponding to the K190 of *E. coli* MenF is required to activate water for nucleophilic attack on C2 of chorismate [132,140]. All amination enzymes including mtPabB replace this lysine with a glutamine, making it very unlikely for mtPabB to function as a hydroxylation enzyme.

4.5 Conclusion

mtTrpE displays slow AS-NH$_3$ activity along with detectable CM and PDT activities that could be due to contamination or intrinsic promiscuity. Addition of TrpG improves AS reaction efficiency and specificity implying that subunit interactions are important for catalytic function. Gel filtration did not reveal a complex between TrpE and TrpG, but the cooperativity in chorismate binding suggests the presence of a heterotetramer. Unlike the most well-characterized anthranilate synthases from *S. typhimurium* and *S. marcescens*, the reaction intermediate ADIC can be released into solution in the TrpEG reaction. The different oligomeric states and reaction kinetics observed here from that of *S. typhimurium* and *S. marcescens* reflect the difference in gene organization and suggest a different mechanism of regulation among the chorismate pathways.

References

(1) WHO 2010.

(2) Wouk, H. *Tuberculosis (Health Alert)*; Benchmark Books: New York, 2009.

(3) WHO *WHO report 2011* **2011**.

(4) WHO; World Health Organization: 2010.

(5) WHO; World Health Organization: 2011.

(6) Finer, K. R.; Alcamo, I. E. *Tuberculosis (deadly diseases & epidemics)*; Chelsea House Publishers: New York, 2003.

(7) Madkour, M. M. *Tuberculosis*; 1st ed.; Springer, 2004.

(8) Roberts, C.; Buikstra, J. *The bioarchaeology of tuberculosis: a global view on a reemerging disease*; 1st ed.; University Press of Florida, 2008.

(9) Schatz, A.; Bugie, E.; Waksman, S. A. *Proceedings of the Society for Experimental Biology and Medicine* **1944**, *55*, 66.

(10) Lehmann, J. *Lancet* **1946**, *1*, 15.

(11)   Bernheim, F. *Science (New York, N.Y.)* **1940**, *92*, 204.

(12)   Chorine, V. *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences* **1945**, *220*, 150.

(13)   Fox, H. H. *Science* **1952**, *116*, 129.

(14)   Ying, Z. In *Annual Review of Pharmacology and Toxicology* 2005; Vol. 45, p 529.

(15)   WHO *Treatment of tuberculosis guidelines*; 4th ed.; World Health Organization, 2010.

(16)   WHO *Anti-tuberculosis drug resistance in the world: fourth global report*, World Health Organization, 2008.

(17)   Dye, C. In *Antibiotic resistance: from genes to global prevalence, the biomedical & life sciences collection*; Gillespie, S., Ed.; Henry Stewart Talks Ltd, London: 2009.

(18)   WHO *Guidelines for the programmatic management of drug-resistant tuberculosis: emergency update 2008*, World Health Organization, 2008.

(19)   Abu-Raddad, L. J.; Sabatelli, L.; Achterberg, J. T.; Sugimoto, J. D.; Longini, I. M., Jr.; Dye, C.; Halloran, M. E. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 13980.

(20)   Koul, A.; Arnoult, E.; Lounis, N.; Guillemont, J.; Andries, K. *Nature* **2011**, *469*, 483.

(21)   Winder, F. G.; Collins, P. B. *Journal of general microbiology* **1970**, *63*, 41.

(22)   Zhang, Y.; Heym, B.; Allen, B.; Young, D.; Cole, S. *Nature* **1992**, *358*, 591.

(23)   Banerjee, A.; Dubnau, E.; Quemard, A.; Balasubramanian, V.; Um, K. S.; Wilson, T.; Collins, D.; Delisle, G.; Jacobs, W. R. *Science* **1994**, *263*, 227.

(24)   Rozwarski, D. A.; Grant, G. A.; Barton, D. H. R.; Jacobs, W. R.; Sacchettini, J. C. *Science* **1998**, *279*, 98.

(25)   Timmins, G. S.; Deretic, V. *Molecular Microbiology* **2006**, *62*, 1220.

(26)   Cade, C. E.; Dlouhy, A. C.; Medzihradszky, K. F.; Salas-Castillo, S. P.; Ghiladi, R. A. *Protein Science* **2010**, *19*, 458.

(27)   DeBarber, A. E.; Mdluli, K.; Bosman, M.; Bekker, L. G.; Barry, C. E. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 9677.

(28)   Baulard, A. R.; Betts, J. C.; Engohang-Ndong, J.; Quan, S.; McAdam, R. A.; Brennan, P. J.; Locht, C.; Besra, G. S. *Journal of Biological Chemistry* **2000**, *275*, 28326.

(29)   Vannelli, T. A.; Dykman, A.; de Montellano, P. R. O. *Journal of Biological Chemistry* **2002**, *277*, 12824.

(30)   Takayama, K.; Kilburn, J. O. *Antimicrobial Agents and Chemotherapy* **1989**, *33*, 1493.

(31)   Wolucka, B. A.; McNeil, M. R.; Dehoffmann, E.; Chojnacki, T.; Brennan, P. J. *Journal of Biological Chemistry* **1994**, *269*, 23328.

(32)   Mikusova, K.; Slayden, R. A.; Besra, G. S.; Brennan, P. J. *Antimicrobial Agents and Chemotherapy* **1995**, *39*, 2484.

(33)   Telenti, A.; Philipp, W. J.; Sreevatsan, S.; Bernasconi, C.; Stockbauer, K. E.; Wieles, B.; Musser, J. M.; Jacobs, W. R. *Nature Medicine* **1997**, *3*, 567.

(34)   Feng, Z. Y.; Barletta, R. G. *Antimicrobial Agents and Chemotherapy* **2003**, *47*, 283.

(35)   Williams, D. L.; Spring, L.; Collins, L.; Miller, L. P.; Heifets, L. B.; Gangadharam, P. R. J.; Gillis, T. P. *Antimicrobial Agents and Chemotherapy* **1998**, *42*, 1853.

(36)   Andersson, M. I.; MacGowan, A. P. *Journal of Antimicrobial Chemotherapy* **2003**, *51*, 1.

(37)   Anand, N.; Davis, B. D. *Nature* **1960**, *185*, 22.

(38)   Spotts, C. R.; Stanier, R. Y. *Nature* **1961**, *192*, 633.

(39)   Davies, J.; Gorini, L.; Davis, B. D. *Molecular Pharmacology* **1965**, *1*, 93.

(40)   Tonge, P. J.; Kisker, C.; Slayden, R. A. *Current Topics In Medicinal Chemistry* **2007**, *7*, 489.

(41)   Umesiri, F. E.; Sanki, A. K.; Boucau, J.; Ronning, D. R.; Sucheck, S. J. *Medicinal Research Reviews* **2010**, *30*, 290.

(42)   Rivers, E. C.; Mancera, R. L. *Current Medicinal Chemistry* **2008**, *15*, 1956.

(43)   Makarov, V.; Manina, G.; Mikusova, K.; Moellmann, U.; Ryabova, O.; Saint-Joanis, B.; Dhar, N.; Pasca, M. R.; Buroni, S.; Lucarelli, A. P.; Milano, A.; De Rossi, E.; Belanova, M.; Bobovska, A.; Dianiskova, P.; Kordulakova, J.; Sala, C.; Fullam, E.; Schneider, P.; McKinney, J. D.; Brodin, P.; Christophe, T.; Waddell, S.; Butcher, P.; Albrethsen, J.; Rosenkrands, I.; Brosch, R.; Nandi, V.; Bharath, S.; Gaonkar, S.; Shandil, R. K.; Balasubramanian, V.; Balganesh, T.; Tyagi, S.; Grosset, J.; Riccardi, G.; Cole, S. T. *Science* **2009**, *324*, 801.

(44)   Tonge, P. J. *Nature Structural Biology* **2000**, *7*, 94.

(45)   Rawat, R.; Whitty, A.; Tonge, P. J. *PNAS* **2003**, *100*, 13881.

(46)   Kolattukudy, P. E.; Poulose, A. J.; Buckner, J. S. *Methods Enzymol.* **1981**, *71 Pt C*, 103.

(47)   Wakil, S. J.; Stoops, J. K.; Joshi, V. C. *Annual Review of Biochemistry* **1983**, *52*, 537.

(48)   Fulco, A. J. *Prog. Lipid Res.* **1983**, *22*, 133.

(49)   Smith, S. *FASEB J.* **1994**, *8*, 1248.

(50)   Kater, M. M.; Koningstein, G. M.; Nijkamp, H. J. J.; Stuitje, A. R. *Plant Mol. Biol.* **1994**, *25*, 771.

(51)   Rock, C. O.; Cronan, J. E. *Biochimica Et Biophysica Acta-Lipids and Lipid Metabolism* **1996**, *1302*, 1.

(52)   Bloch, K. *Advances in Enzymology and Related Areas of Molecular Biology* **1977**, *45*, 1.

(53)   Kikuchi, S.; Rainwater, D. L.; Kolattukudy, P. E. *Archives of Biochemistry and Biophysics* **1992**, *295*, 318.

(54)   Qureshi, N.; Sathyamoorthy, N.; Takayama, K. *Journal of Bacteriology* **1984**, *157*, 46.

(55)   Brennan, P. J.; Nikaido, H. *Annual Review of Biochemistry* **1995**, *64*, 29.

(56)   Lee, R. E.; Brennan, P. J.; Besra, G. S. *Tuberculosis* **1996**, *215*, 1.

(57)   Takayama, K.; Wang, C.; Besra, G. S. *Clinical Microbiology Reviews* **2005**, *18*, 81.

(58)   Heath, R. J.; White, S. W.; Rock, C. O. *Progress in Lipid Research* **2001**, *40*, 467.

(59)   Campbell, J. W.; Cronan, J. E. *Annual Review of Microbiology* **2001**, *55*, 305.

(60)   Heath, R. J.; White, S. W.; Rock, C. O. *Applied Microbiology and Biotechnology* **2002**, *58*, 695.

(61)   Wright, H. T.; Reynolds, K. A. *Current Opinion in Microbiology* **2007**, *10*, 447.

(62)   Lu, H.; Tonge, P. J. *Accounts of Chemical Research* **2008**, *41*, 11.

(63)   Quemard, A.; Sacchettini, J. C.; Dessen, A.; Vilcheze, C.; Bittman, R.; Jacobs, W. R.; Blanchard, J. S. *Biochemistry* **1995**, *34*, 8235.

(64)   Marrakchi, H.; Laneelle, G.; Quemard, A. *Microbiology-Uk* **2000**, *146*, 289.

(65)   Qureshi, N.; Sathyamoorthy, N.; Takayama, K. *Journal of Bacteriology* **1984**, *157*, 46.

(66)   Gurvitz, A.; Hiltunen, J. K.; Kastaniotis, A. J. *Appl. Environ. Microbiol.* **2008**, *74*, 5078.

(67)   Mdluli, K.; Slayden, R. A.; Zhu, Y. Q.; Ramaswamy, S.; Pan, X.; Mead, D.; Crane, D. D.; Musser, J. M.; Barry, C. E. *Science* **1998**, *280*, 1607.

(68)   Ling, L. L.; Xian, J.; Ali, S.; Geng, B. L.; Fan, J.; Mills, D. M.; Arvanites, A. C.; Orgueira, H.; Ashwell, M. A.; Carmel, G.; Xiang, Y. B.; Moir, D. T. *Antimicrobial Agents and Chemotherapy* **2004**, *48*, 1541.

(69)   Moir, D. T. *Current drug targets. Infectious disorders* **2005**, *5*, 297.

(70)   Pan, P.; Tonge, P. J. *Curr Top Med Chem.* **2012**.

(71)   Mitchison, D. A. *Frontiers in Bioscience* **2004**, *9*, 1059.

(72)   Mitchison, D. A. *Chest* **1979**, *76*, 771.

(73)   Zhang, Y. *Frontiers in Bioscience* **2004**, *9*, 1136.

(74)   Gideon, H. P.; Flynn, J. L. *Immunologic Research* **2011**, *50*, 202.

(75)   Wayne, L. G. *Eur. J. Clin. Microbiol. Infect. Dis.* **1994**, *13*, 908.

(76)   Boshoff, H. I. M.; Barry, C. E. *Nature Reviews Microbiology* **2005**, *3*, 70.

(77)   Weinstein, E. A.; Yano, T.; Li, L. S.; Avarbock, D.; Avarbock, A.; Helm, D.; McColm, A. A.; Duncan, K.; Lonsdale, J. T.; Rubin, H. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 4548.

(78)   Li, X.; Liu, N.; Zhang, H.; Knudson, S. E.; Slayden, R. A.; Tonge, P. J. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6306.

(79)   Li, X.; Liu, N.; Zhang, H.; Knudson, S. E.; Li, H.-J.; Lai, C.-T.; Simmerling, C.; Slayden, R. A.; Tonge, P. J. *ACS Medicinal Chemistry Letters* **2011**, *2*, 818.

(80)   Andries, K.; Verhasselt, P.; Guillemont, J.; Gohlmann, H. W. H.; Neefs, J. M.; Winkler, H.; Van Gestel, J.; Timmerman, P.; Zhu, M.; Lee, E.; Williams, P.; de Chaffoy, D.; Huitric, E.; Hoffner, S.; Cambau, E.; Truffot-Pernot, C.; Lounis, N.; Jarlier, V. *Science* **2005**, *307*, 223.

(81)   Koul, A.; Dendouga, N.; Vergauwen, K.; Molenberghs, B.; Vranckx, L.; Willebrords, R.; Ristic, Z.; Lill, H.; Dorange, I.; Guillemont, J.; Bald, D.; Andries, K. *Nature Chemical Biology* **2007**, *3*, 323.

(82)   Sohaskey, C. *Recent patents on anti-infective drug discovery* **2011**, *6*, 139.

(83)   Piccaro, G.; Filippini, P.; Giannoni, F.; Scipione, L.; Tortorella, S.; De Vita, D.; Mellini, P.; Fattorini, L. *Journal of Chemotherapy* **2011**, *23*, 175.

(84)   Rao, S. P. S.; Alonso, S.; Rand, L.; Dick, T.; Pethe, K. *Proceedings of the National Academy of Sciences of the United States of America* **2008**, *105*, 11945.

(85)   Bald, D.; Koul, A. *Fems Microbiology Letters* **2010**, *308*, 1.

(86)   Gengenbacher, M.; Rao, S. P. S.; Pethe, K.; Dick, T. *Microbiology-Sgm* **2010**, *156*, 81.

(87)   Sohaskey, C. D.; Wayne, L. G. *Journal of Bacteriology* **2003**, *185*, 7247.

(88)   Boshoff, H. I. M.; Myers, T. G.; Copp, B. R.; McNeil, M. R.; Wilson, M. A.; Barry, C. E. *Journal of Biological Chemistry* **2004**, *279*, 40174.

(89)   Sherman, D. R.; Voskuil, M.; Schnappinger, D.; Liao, R. L.; Harrell, M. I.; Schoolnik, G. K. *Proceedings of the National Academy of Sciences of the United States of America* **2001**, *98*, 7534.

(90)   Voskuil, M. I.; Schnappinger, D.; Visconti, K. C.; Harrell, M. I.; Dolganov, G. M.; Sherman, D. R.; Schoolnik, G. K. *Journal of Experimental Medicine* **2003**, *198*, 705.

(91)   Nathan, C.; Shiloh, M. U. *Proceedings of the National Academy of Sciences of the United States of America* **2000**, *97*, 8841.

(92)   Kana, B. D.; Weinstein, E. A.; Avarbock, D.; Dawes, S. S.; Rubin, H.; Mizrahi, V. *Journal of Bacteriology* **2001**, *183*, 7076.

(93)   Kurosu, M.; Begari, E. *Molecules* **2010**, *15*, 1531.

(94)   Bishop, D. H. L.; Pandya, K. P.; King, H. K. *Biochem. J.* **1962**, *83*, 606.

(95)   Collins, M. D.; Goodfellow, M.; Minnikin, D. E.; Alderson, G. *J. Appl. Bacteriol.* **1985**, *58*, 77.

(96)   Widhalm, J. R.; van Oostende, C.; Furt, F.; Basset, G. J. C. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 5599.

(97)   Lester, R. L.; Crane, F. L. *J. Biol. Chem.* **1959**, *234*, 2169.

(98)   Johnson, T. W.; Shen, G. Z.; Zybailov, B.; Kolling, D.; Reategui, R.; Beauparlant, S.; Vassiliev, I. R.; Bryant, D. A.; Jones, A. D.; Golbeck, J. H.; Chitnis, P. R. *J. Biol. Chem.* **2000**, *275*, 8523.

(99)   Johnson, T. W.; Zybailov, B.; Jones, A. D.; Bittl, R.; Zech, S.; Stehlik, D.; Golbeck, J. H.; Chitnis, P. R. *J. Biol. Chem.* **2001**, *276*, 39512.

(100)   Lefebvre-Legendre, L.; Rappaport, F.; Finazzi, G.; Ceol, M.; Grivet, C.; Hopfgartner, G.; Rochaix, J. D. *J. Biol. Chem.* **2007**, *282*, 13250.

(101)   Gross, J.; Meurer, J.; Bhattacharya, D. *Bmc Evolutionary Biology* **2008**, *8*.

(102)   Gross, J.; Cho, W. K.; Lezhneva, L.; Falk, J.; Krupinska, K.; Shinozaki, K.; Seki, M.; Herrmann, R. G.; Meurer, J. *J. Biol. Chem.* **2006**, *281*, 17189.

(103)   Gross, J., Ludwig Maximilians Universität München, 2006.

(104)   Dowd, P.; Ham, S. W.; Naganathan, S.; Hershline, R. *Annu. Rev. Nutr.* **1995**, *15*, 419.

(105)   Olson, R. E. *Annu. Rev. Nutr.* **1984**, *4*, 281.

(106) Kurosu, M.; Narayanasamy, P.; Biswas, K.; Dhiman, R.; Crick, D. C. *J. Med. Chem.* **2007**, *50*, 3973.

(107) Li, X.; Liu, N.; Zhang, H.; Knudson, S. E.; Slayden, R. A.; Tonge, P. J. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6306.

(108) Lu, X.; Zhang, H.; Tonge, P. J.; Tan, D. S. *Bioorg. Med. Chem. Lett.* **2008**.

(109) Truglio, J. J.; Theis, K.; Feng, Y.; Gajda, R.; Machutta, C.; Tonge, P. J.; Kisker, C. *J. Biol. Chem.* **2003**, *278*, 42352.

(110) Noll, H. *J. Biol. Chem.* **1958**, *232*, 919.

(111) Collins, M. D.; Pirouz, T.; Goodfellow, M.; Minnikin, D. E. *J. Gen. Microbiol.* **1977**, *100*, 221.

(112) Truglio, J. J.; Theis, K.; Feng, Y. G.; Gajda, R.; Machutta, C.; Tonge, P. J.; Kisker, C. *J. Biol. Chem.* **2003**, *278*, 42352.

(113) Dhiman, R. K.; Mahapatra, S.; Slayden, R. A.; Boyne, M. E.; Lenaerts, A.; Hinshaw, J. C.; Angala, S. K.; Chatterjee, D.; Biswas, K.; Narayanasamy, P.; Kurosu, M.; Crick, D. C. *Mol. Microbiol.* **2009**, *72*, 85.

(114) Lu, X.; Zhang, H.; Tonge, P. J.; Tan, D. S. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5963.

(115) Lu, X.; Zhou, R.; Sharma, I.; Li, X.; Kumar, G.; Swaminathan, S.; Tonge, P. J.; Tan, D. S. *ChemBioChem* **2012**, *13*, 129.

(116) Kurosu, M.; Crick, D. C. *Medicinal Chemistry* **2009**, *5*, 197.

(117) Sassetti, C. M.; Boyd, D. H.; Rubin, E. J. *Molecular Microbiology* **2003**, *48*, 77.

(118) Collins, M. D.; Jones, D. *Microbiol. Rev.* **1981**, *45*, 316.

(119) Holsclaw, C. M.; Sogi, K. M.; Gilmore, S. A.; Schelle, M. W.; Leavell, M. D.; Bertozzi, C. R.; Leary, J. A. *ACS chemical biology* **2008**, *3*, 619.

(120) Ishii, M.; Kawasumi, T.; Igarashi, Y.; Kodama, T.; Minoda, Y. *J. Bacteriol.* **1987**, *169*, 2380.

(121) Bentley, R.; Meganathan, R. *Microbiol. Rev.* **1982**, *46*, 241.

(122) Meganathan, R. *Vitam. Horm.* **2001**, *61*, 173.

(123) Jiang, M.; Cao, Y.; Guo, Z. F.; Chen, M.; Chen, X.; Guo, Z. *Biochemistry* **2007**, *46*, 10979.

(124) Jiang, M.; Chen, X.; Guo, Z. F.; Cao, Y.; Chen, M.; Guo, Z. *Biochemistry* **2008**, *47*, 3426.

(125) Hiratsuka, T.; Furihata, K.; Ishikawa, J.; Yamashita, H.; Itoh, N.; Seto, H.; Dairi, T. *Science* **2008**, *321*, 1670.

(126) Cole, S. T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S. V.; Eiglmeier, K.; Gas, S.; Barry, C. E.; Tekaia, F.; Badcock, K.; Basham, D.; Brown, D.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Hornby, T.; Jagels, K.; Krogh, A.; McLean, J.; Moule, S.; Murphy, L.; Oliver, K.; Osborne, J.; Quail, M. A.; Rajandream, M. A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, J.; Squares, R.; Squares, S.; Sulston, J. E.; Taylor, K.; Whitehead, S.; Barrell, B. G. *Nature* **1998**, *393*, 537.

(127) Knaggs, A. R. *Natural Product Reports* **2000**, *DOI: 10.1039/a902464f*.

(128) Gosset, G. *Current Opinion in Biotechnology* **2009**, *20*, 651.

(129) Walsh, C. T.; Haynes, S. W.; Ames, B. D. *Natural Product Reports* **2012**, *29*, 37.

(130) Ducati, R. G.; Basso, L. A.; Santos, D. S. *Current Drug Targets* **2007**, *8*, 423.

(131) Kerbarh, O.; Bulloch, E. M. M.; Payne, R. J.; Sahr, T.; Rébeillé, F.; Abell, C.; Biochem. Soc. Trans.: 2005; Vol. 33, p 763.

(132) Kolappan, S.; Zwahlen, J.; Zhou, R.; Truglio, J. J.; Tonge, P. J.; Kisker, C. *Biochemistry* **2007**, *46*, 946.

(133) Kozlowski, M. C.; Tom, N. J.; Seto, C. T.; Sefler, A. M.; Bartlett, P. A.; JACS: 1995; Vol. 117, p 2128.

(134) He, Z.; StigersLavoie, K. D.; Bartlett, P. A.; Toney, M. D. *JACS* **2004**, *126*, 2378.

(135)   Ziebart, K. T.; Dixon, S. M.; Avila, B.; El-Badri, M. H.; Guggenheim, K. G.; Kurth, M. J.; Toney, M. D. *Journal of Medicinal Chemistry* **2010**, *53*, 3718.

(136)   Knaggs, A. R. *Natural Product Reports* **2003**, *20*, 119.

(137)   Daruwala, R.; Bhattacharyya, D. K.; Kwon, O.; Meganathan, R.; J. Bacteriology: 1997; Vol. 179, p 3133.

(138)   Daruwala, R.; Kwon, O.; Meganathan, R.; Hudspeth, M. E. S. *Fems Microbiology Letters* **1996**, *140*, 159.

(139)   Li, L. L.; Bannantine, J. P.; Zhang, Q.; Amonsin, A.; May, B. J.; Alt, D.; Banerji, N.; Kanjilal, S.; Kapur, V. *Proceedings of the National Academy of Sciences of the United States of America* **2005**, *102*, 12344.

(140)   Zwahlen, J.; Kolappan, S.; Zhou, R.; Kisker, C.; Tonge, P. J. *Biochemistry* **2007**, *46*, 954.

(141)   Quadri, L. E. N.; Sello, J.; Keating, T. A.; Weinreb, P. H.; Walsh, C. T. *Chemistry & Biology* **1998**, *5*, 631.

(142)   McMurry, L. M.; Oethinger, M.; Levy, S. B. *Nature* **1998**, *394*, 531.

(143)   Heath, R. J.; Yu, Y. T.; Shapiro, M. A.; Olson, E.; Rock, C. O. *J. Biol. Chem.* **1998**, *273*, 30316.

(144)   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.; Taylor, I. W. F. *Biochemistry* **1999**, *38*, 12514.

(145)   Heath, R. J.; Rubin, J. R.; Holland, D. R.; Zhang, E. L.; Snow, M. E.; Rock, C. O. *J. Biol. Chem.* **1999**, *274*, 11110.

(146)   Sivaraman, S.; Zwahlen, J.; Bell, A. F.; Hedstrom, L.; Tonge, P. J. *Biochemistry* **2003**, *42*, 4406.

(147)   Sivaraman, S.; Sullivan, T. J.; Johnson, F.; Novichenok, P.; Cui, G.; Simmerling, C.; Tonge, P. J. *Journal of Medicinal Chemistry* **2004**, *47*, 509.

(148)   Jones, R. D.; Jampani, H. B.; Newman, J. L.; Lee, A. S. *Am. J. Infect. Control* **2000**, *28*, 184.

(149)   Kapoor, M.; Dar, M. J.; Surolia, A.; Surolia, N. *Biochemical and Biophysical Research Communications* **2001**, *289*, 832.

(150)   Kapoor, M.; Reddy, C. C.; Krishnasastry, M. V.; Surolia, N.; Surolia, A. *Biochemical Journal* **2004**, *381*, 719.

(151)   Kapoor, M.; Mukhi, P. L. S.; Surolia, N.; Suguna, K.; Surolia, A. *Biochemical Journal* **2004**, *381*, 725.

(152)   Heath, R. J.; Li, J.; Roland, G. E.; Rock, C. O. *J. Biol. Chem.* **2000**, *275*, 4654.

(153)   Xu, H.; Sullivan, T. J.; Sekiguchi, J.-i.; Kirikae, T.; Ojima, I.; Stratton, C. F.; Mao, W.; Rock, F. L.; Alley, M. R. K.; Johnson, F.; Walker, S. G.; Tonge, P. J. *Biochemistry* **2008**, *47*, 4228.

(154)   Lu, H.; England, K.; Ende, C. A.; Truglio, J. J.; Luckner, S.; Reddy, B. G.; Marlenee, N. L.; Knudson, S. E.; Knudson, D. L.; Bowen, R. A.; Kisker, C.; Slayden, R. A.; Tonge, P. J. *ACS Chemical Biology* **2009**, *4*, 221.

(155)   Liu, N.; Cummings, J. E.; England, K.; Slayden, R. A.; Tonge, P. J. *J. Antimicrob. Chemother.* **2011**, *66*, 564.

(156)   Parikh, S. L.; Xiao, G.; Tonge, P. J. *Biochemistry* **2000**, *39*, 7645.

(157)   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.; Tonge, P. J. *ACS Chemical Biology* **2006**, *1*, 43.

(158)   Freundlich, J. S.; Wang, F.; Vilcheze, C.; Gulten, G.; Langley, R.; Schiehser, G. A.; Jacobus, D. R.; Jacobs, W. R.; Sacchettini, J. C. *Chemmedchem* **2009**, *4*, 241.

(159)   Luckner, S. R.; Liu, N.; am Ende, C. W.; Tonge, P. J.; Kisker, C. *J. Biol. Chem.* **2010**, *285*, 14330.

(160)	Tipparaju, S. K.; Mulhearn, D. C.; Klein, G. M.; Chen, Y.; Tapadar, S.; Bishop, M. H.; Yang, S.; Chen, J.; Ghassemi, M.; Santarsiero, B. D.; Cook, J. L.; Johlfs, M.; Mesecar, A. D.; Johnson, M. E.; Kozikowski, A. P. *Chemmedchem* **2008**, *3*, 1250.

(161)	Heath, R. J.; Su, N.; Murphy, C. K.; Rock, C. O. *J. Biol. Chem.* **2000**, *275*, 40128.

(162)	Marcinkeviciene, J.; Jiang, W.; Kopcho, L. M.; Locke, G.; Luo, Y.; Copeland, R. A. *Archives of Biochemistry and Biophysics* **2001**, *390*, 101.

(163)	Garvey, E. P. *Current Chemical Biology* **2010**, *4*, 64.

(164)	Hiratake, J. *The Chemical Record* **2005**, *5*, 209.

(165)	Schloss, J. V. *Accounts of Chemical Research* **1988**, *21*, 348.

(166)	Morrison, J. F. *Trends in Biochemical Sciences* **1982**, *7*, 102.

(167)	Szedlacsek, S. E.; Duggleby, R. G. *Enzyme kinetics and mechanism, Pt D* **1995**, *249*, 144.

(168)	Smith, H. J., Simons, Claire *Enzymes and Their Inhibitors: Drug Development (Enzyme Inhibitors)*; CRC Press, 2004.

(169)	Morrison, J. F.; Walsh, C. T. *Advances in enzymology and related areas of molecular biology* **1988**, *61*, 201.

(170)	Williams, J. W.; Morrison, J. F. In *Methods Enzymol.*; Daniel, L. P., Ed.; Academic Press: 1979; Vol. Volume 63, p 437.

(171)	Williams, J. W.; Morrison, J. F.; Duggleby, R. G. *Biochemistry* **1979**, *18*, 2567.

(172)	Frieden, C.; Kurz, L. C.; Gilbert, H. R. *Biochemistry* **1980**, *19*, 5303.

(173)	Schloss, J. V.; Porter, D. J. T.; Bright, H. J.; Cleland, W. W. *Biochemistry* **1980**, *19*, 2358.

(174)	Baici, A.; Gygermarazzi, M. *Eur. J. Biochem.* **1982**, *129*, 33.

(175)	Copeland, R. A. *Future medicinal chemistry* **2011**, *3*, 1491.

(176)	Cha, S. *Biochemical Pharmacology* **1975**, *24*, 2177.

(177)	Cha, S. *Biochemical Pharmacology* **1976**, *25*, 1561.

(178)	Cha, S. *Biochemical Pharmacology* **1976**, *25*, 2695.

(179)	Duggleby, R. G.; Attwood, P. V.; Wallace, J. C.; Keech, D. B. *Biochemistry* **1982**, *21*, 3364.

(180)	Tummino, P. J.; Copeland, R. A. *Biochemistry* **2008**, *47*, 5481.

(181)	Schramm, V. L. *Annual Review of Biochemistry* **2011**, *80*, 703.

(182)	Wood, E. R.; Truesdale, A. T.; McDonald, O. B.; Yuan, D.; Hassell, A.; Dickerson, S. H.; Ellis, B.; Pennisi, C.; Horne, E.; Lackey, K.; Alligood, K. J.; Rusnak, D. W.; Gilmer, T. M.; Shewchuk, L. *Cancer Research* **2004**, *64*, 6652.

(183)	Pargellis, C.; Tong, L.; Churchill, L.; Cirillo, P. F.; Gilmore, T.; Graham, A. G.; Grob, P. M.; Hickey, E. R.; Moss, N.; Pav, S.; Regan, J. *Nat Struct Mol Biol* **2002**, *9*, 268.

(184)	Wang, M. Z.; Tai, C. Y.; Mendel, D. B. *Antimicrob. Agents Chemother.* **2002**, *46*, 3809.

(185)	Rafferty, J. B.; Simon, J. W.; Baldock, C.; Artymiuk, P. J.; Baker, P. J.; Stuitje, A. R.; Slabas, A. R.; Rice, D. W. *Structure* **1995**, *3*, 927.

(186)	Baldock, C.; Rafferty, J. B.; Stuitje, A. R.; Slabas, A. R.; Rice, D. W. *Journal of Molecular Biology* **1998**, *284*, 1529.

(187)	Stewart, M. J.; Parikh, S.; Xiao, G. P.; Tonge, P. J.; Kisker, C. *Journal of Molecular Biology* **1999**, *290*, 859.

(188)	Rozwarski, D. A.; Vilcheze, C.; Sugantino, M.; Bittman, R.; Sacchettini, J. C. *J. Biol. Chem.* **1999**, *274*, 15582.

(189)	Roujeinikova, A.; Sedelnikova, S.; de Boer, G. J.; Stuitje, A. R.; Slabas, A. R.; Rafferty, J. B.; Rice, D. W. *J. Biol. Chem.* **1999**, *274*, 30811.

(190)	Qiu, X.; Abdel-Meguid, S. S.; Janson, C. A.; Court, R. I.; Smyth, M. G.; Payne, D. J. *Protein Science* **1999**, *8*, 2529.

(191)	Levy, C. W.; Roujeinikova, A.; Sedelnikova, S.; Baker, P. J.; Stuitje, A. R.; Slabas, A. R.; Rice, D. W.; Rafferty, J. B. *Nature* **1999**, *398*, 383.

(192) 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.; Rice, D. W. *Journal of Molecular Biology* **1999**, *294*, 527.

(193) Kim, K.-H.; Ha, B. H.; Kim, S. J.; Hong, S. K.; Hwang, K. Y.; Kim, E. E. *Journal of Molecular Biology* **2011**, *406*, 403.

(194) Mehboob, S.; Truong, K.; Santarsiero, B. D.; Johnson, M. E. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **2010**, *66*, 1436.

(195) Lee, H. H.; Moon, J.; Suh, S. W. *Proteins-Structure Function and Bioinformatics* **2007**, *69*, 691.

(196) 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.; Fidock, D. A. *JBC* **2003**, *278*, 20851.

(197) Perozzo, R.; Kuo, M.; Sidhu, A. b. S.; Valiyaveettil, J. T.; Bittman, R.; Jacobs, W. R.; Fidock, D. A.; Sacchettini, J. C. *J. Biol. Chem.* **2002**, *277*, 13106.

(198) Pidugu, L. S.; Kapoor, M.; Surolia, N.; Surolia, A.; Suguna, K. *Journal of Molecular Biology* **2004**, *343*, 147.

(199) Muench, S. P.; Prigge, S. T.; McLeod, R.; Rafferty, J. B.; Kirisits, M. J.; Roberts, C. W.; Mui, E. J.; Rice, D. W. *Acta Crystallographica Section D-Biological Crystallography* **2007**, *63*, 328.

(200) Priyadarshi, A.; Kim, E. E.; Hwang, K. Y. *Proteins-Structure Function and Bioinformatics* **2010**, *78*, 480.

(201) Baldock, C.; Rafferty, J. B.; Sedelnikova, S. E.; Baker, P. J.; Stuitje, A. R.; Slabas, A. R.; Hawkes, T. R.; Rice, D. W. *Science* **1996**, *274*, 2107.

(202) Dias, M. V. B.; Vasconcelos, I. B.; Prado, A. M. X.; Fadel, V.; Basso, L. A.; De Azevedo, W. F.; Santos, D. S. *Journal of Structural Biology* **2007**, *159*, 369.

(203) Dessen, A.; Quemard, A.; Blanchard, J. S.; Jr, W. R. J.; Sacchettini, J. C. *Science* **1995**, *267*, 1638.

(204) Oliveira, J. S.; Pereira, J. H.; Canduri, F.; Rodrigues, N. C.; de Souza, O. N.; de Azevedo, W. F.; Basso, L. A.; Santos, D. S. *Journal of Molecular Biology* **2006**, *359*, 646.

(205) Molle, V.; Gulten, G.; Vilcheze, C.; Veyron-Churlet, R.; Zanella-Cleon, I.; Sacchettini, J. C.; Jacobs, W. R., Jr.; Kremer, L. *Molecular Microbiology* **2010**, *78*, 1591.

(206) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276: Macromolecular Crystallography, part A*, 307.

(207) Vagin, A.; Teplyakov, A. *J. Appl. Crystallogr.* **1997**, *30*, 1022.

(208) Bailey, S. *Acta Crystallographica Section D-Biological Crystallography* **1994**, *50*, 760.

(209) Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2009**, *65*, 1074.

(210) Brunger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallographica Section D-Biological Crystallography* **1998**, *54*, 905.

(211) Brunger, A. T. *Nature Protocols* **2007**, *2*, 2728.

(212) Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2010**, *66*, 213.

(213) Emsley, P.; Cowtan, K. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2004**, *60*, 2126.

(214) Krissinel, E.; Henrick, K. *Journal of Molecular Biology* **2007**, *372*, 774.

(215) Teague, S. J. *Nat Rev Drug Discov* **2003**, *2*, 527.

(216)   Bergonzo, C.; Campbell, A. J.; Walker, R. C.; Simmerling, C. *International Journal of Quantum Chemistry* **2009**, *109*, 3781.

(217)   Roux, B. *Computer Physics Communications* **1995**, *91*, 275.

(218)   Grossfield, A. *WHAM: the weighted histogram analysis method version 2.0.4, http://membrane.urmc.rochester.edu/content/wham*, 2007.

(219)   Torrie, G. M.; Valleau, J. P. *Journal of Computational Physics* **1977**, *23*, 187.

(220)   Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A. *Journal of Computational Chemistry* **1995**, *16*, 1339.

(221)   Petsko, G. A.; Ringe, D. *Annual Review of Biophysics and Bioengineering* **1984**, *13*, 331.

(222)   Bui, J. M.; McCammon, J. A. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 15451.

(223)   Okazaki, K.-i.; Takada, S. *PNAS* **2008**, *105*, 11182.

(224)   Henzler-Wildman, K.; Kern, D. *Nature* **2007**, *450*, 964.

(225)   Nölting, B. *Protein folding kinetics: biophysical methods*; Springer: Berlin, 1999.

(226)   Baldwin, R. L. *Science* **2002**, *295*, 1657.

(227)   Dyson, H. J.; Wright, P. E.; Scheraga, H. A. *Proceedings of the National Academy of Sciences of the United States of America* **2006**, *103*, 13057.

(228)   Lesk, A. M.; Rose, G. D. *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences* **1981**, *78*, 4304.

(229)   Portman, J. J. *Current Opinion in Structural Biology* **2010**, *20*, 11.

(230)   De Sancho, D.; Doshi, U.; Munoz, V. *Journal of the American Chemical Society* **2009**, *131*, 2074.

(231)   De Sancho, D.; Munoz, V. *Physical Chemistry Chemical Physics* **2011**, *13*, 17030.

(232)   Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G. *Science* **1991**, *254*, 1598.

(233)   Ansari, A.; Berendzen, J.; Braunstein, D.; Cowen, B. R.; Frauenfelder, H.; Hong, M. K.; Iben, I. E. T.; Johnson, J. B.; Ormos, P.; Sauke, T. B.; Scholl, R.; Schulte, A.; Steinbach, P. J.; Vittitow, J.; Young, R. D. *Biophys. Chem.* **1987**, *26*, 337.

(234)   Aye, Y.; Stubbe, J. *Proceedings of the National Academy of Sciences of the United States of America* **2011**, *108*, 9815.

(235)   Fieulaine, S.; Boularot, A.; Artaud, I.; Desmadril, M.; Dardel, F.; Meinnel, T.; Giglione, C. *Plos Biology* **2011**, *9*.

(236)   Beckstein, O.; Denning, E. J.; Perilla, J. R.; Woolf, T. B. *Journal of Molecular Biology* **2009**, *394*, 160.

(237)   Walker, M. C.; Kurumbail, R. G.; Kiefer, J. R.; Moreland, K. T.; Koboldt, C. M.; Isakson, P. C.; Seibert, K.; Gierse, J. K. *Biochemical Journal* **2001**, *357*, 709.

(238)   Frieden, C. *Annual Review of Biochemistry* **1979**, *48*, 471.

(239)   Dobson, C. M. In *Mechanisms of protein folding*; 2 ed.; Pain, R. H., Ed.; Oxford University Press: New York, 2000, p 1.

(240)   Balbach, J.; Schmid, F. X. In *Mechanisms of protein folding*; 2 ed.; Pain, R. H., Ed.; Oxford University Press: New York, 2000, p 212.

(241)   Hammes, G. G.; Chang, Y. C.; Oas, T. G. *Proceedings of the National Academy of Sciences of the United States of America* **2009**, *106*, 13737.

(242)   Zhou, H. X. *Biophysical Journal* **2010**, *98*, L15.

(243)   Sullivan, S. M.; Holyoak, T. *Proceedings of the National Academy of Sciences* **2008**, *105*, 13829.

(244)   Csermely, P.; Palotai, R.; Nussinov, R. *Trends in Biochemical Sciences* **2010**, *35*, 539.

(245)   Bucher, D.; Grant, B. J.; McCammon, J. A. *Biochemistry* **2011**, *50*, 10530.

211

(246)   Weikl, T. R.; von Deuster, C. *Proteins: Structure, Function, and Bioinformatics* **2009**, *75*, 104.

(247)   Oliveira, J. S.; de Sousa, E. H. S.; de Souza, O. N.; Moreira, l. S.; Santos, D. S.; Basso, L. A. *Current Pharmaceutical Design* **2006**, *12*, 2409.

(248)   Wang, F.; Langley, R.; Gulten, G.; Dover, L. G.; Besra, G. S.; Jacobs, W. R.; Sacchettini, J. C. *Journal of Experimental Medicine* **2007**, *204*, 73.

(249)   Durrant, J. D.; McCammon, J. A. *Current Opinion in Pharmacology* **2010**, *10*, 770.

(250)   Schames, J. R.; Henchman, R. H.; Siegel, J. S.; Sotriffer, C. A.; Ni, H. H.; McCammon, J. A. *Journal of Medicinal Chemistry* **2004**, *47*, 1879.

(251)   Craig, I. R.; Pfleger, C.; Gohlke, H.; Essex, J. W.; Spiegel, K. *Journal of Chemical Information and Modeling* **2011**, *51*, 2666.

(252)   Bryant, R. W., Jr.; Bentley, R. *Biochemistry* **1976**, *15*, 4792.

(253)   Meganathan, R.; Bentley, R. *J. Bacteriol.* **1979**, *140*, 92.

(254)   Meganathan, R.; Bentley, R. *J. Bacteriol.* **1979**, *140*, 92.

(255)   Shaw, D. J.; Guest, J. R.; Meganathan, R.; Bentley, R. *J. Bacteriol.* **1982**, *152*, 1132.

(256)   Meganathan, R.; Bentley, R.; Taber, H. *J. Bacteriol.* **1981**, *145*, 328.

(257)   Meganathan, R.; Folger, T.; Bentley, R. *Biochemistry* **1980**, *19*, 785.

(258)   Meganathan, R. *FEMS Microbiol. Lett.* **2001**, *203*, 131.

(259)   Bryant, R. W.; Bentley, R. *Biochemistry* **1976**, *15*, 4792.

(260)   Heide, L.; Arendt, S.; Leistner, E. *J. Biol. Chem.* **1982**, *257*, 7396.

(261)   Kolkmann, R.; Knauel, G.; Arendt, S.; Leistner, E. *FEBS Lett.* **1982**, *137*, 53.

(262)   Heide, L.; Leistner, E. *FEBS Lett.* **1981**, *128*, 201.

(263)   Heath, R. J.; Rock, C. O. *Nat. Prod. Rep.* **2002**, *19*, 581.

(264)   Zhang, H.; Machutta, C. A.; Tonge, P. J. In *Comprehensive Natural Products Chemistry II Chemistry and Biology*; Mander, L., Lui, H.-W., Eds. 2010; Vol. 8, p 231.

(265)   Bentley, R.; Meganathan, R. *Microbiol. Rev.* **1982**, *46*, 241.

(266)   Kolkmann, R.; Leistner, E. *Tetrahedron Lett.* **1985**, *26*, 1703.

(267)   Dewick, P. M. *Nat. Prod. Rep.* **1986**, *3*, 565.

(268)   Engel, C. K.; Mathieu, M.; Zeelen, J. P.; Hiltunen, J. K.; Wierenga, R. K. *EMBO J.* **1996**, *15*, 5135.

(269)   Hamed, R. B.; Batchelar, E. T.; Clifton, I. J.; Schofield, C. J. *Cell. Mol. Life Sci.* **2008**, *65*, 2507.

(270)   Gerlt, J. A.; Babbitt, P. C. *Annu. Rev. Biochem.* **2001**, *70*, 209.

(271)   Holden, H. M.; Benning, M. M.; Haller, T.; Gerlt, J. A. *Acc. Chem. Res.* **2001**, *34*, 145.

(272)   Hamed, R. B.; Batchelar, E. T.; Clifton, I. J.; Schofield, C. J. *Cell. Mol. Life Sci.* **2008**, *65*, 2507.

(273)   Bell, A. F.; Wu, J.; Feng, Y.; Tonge, P. J. *Biochemistry* **2001**, *40*, 1725.

(274)   Bennett, M. J.; Schlunegger, M. P.; Eisenberg, D. *Protein Sci.* **1995**, *4*, 2455.

(275)   Benning, M. M.; Haller, T.; Gerlt, J. A.; Holden, H. M. *Biochemistry* **2000**, *39*, 4630.

(276)   Hubbard, P. A.; Yu, W. F.; Schulz, H.; Kim, J. J. P. *Protein Sci.* **2005**, *14*, 1545.

(277)   Jiang, M.; Chen, M.; Guo, Z. F.; Guo, Z. *J. Biol. Chem.* **2010**, *285*, 30159.

(278)   Johnston, J. M.; Arcus, V. L.; Baker, E. N. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **2005**, *61*, 1199.

(279)   Ulaganathan, V.; Agacan, M. F.; Buetow, L.; Tulloch, L. B.; Hunter, W. N. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **2007**, *63*, 908.

(280)   Minasov, G.; Wawrzak, Z.; Skarina, T.; Onopriyenko, O.; Peterson, S. N.; Savchenko, A.; Anderson, W. F. *Center for Structural Genomics of Infectious Diseases (CSGID)* **2009**.

(281)   Jeyakanthan, J.; Kanaujia, S. P.; Vasuki Ranjani, C.; Sekar, K.; BaBa, S.; Ebihara, A.; Kuramitsu, S.; Shinkai, A.; Shiro, Y.; Yokoyama, S. *RIKEN Structural Genomics/Proteomics Initiative (RSGI)* **2006**.

(282)   Martin, D. P.; Drueckhammer, D. G. *J.Am.Chem.Soc.* **1992**, *114*, 7287.

(283)   Mishra, P. K.; Drueckhammer, D. G. *Chem. Rev.* **2000**, *100*, 3283.

(284)   Dai, M.; Feng, Y.; Tonge, P. J. *J. Am. Chem. Soc.* **2001**, *123*, 506.

(285)   Otwinowski, Z.; Minor, W. *Meth. Enzymol.* **1997**, *276*, 307.

(286)   Moriarty, N. W.; Grosse-Kunstleve, R. W.; Adams, P. D. *Acta Crystallogr. D Biol. Crystallogr.* **2009**, *65*, 1074.

(287)   Weininger, D. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 237.

(288)   Adams, P. D.; Afonine, P. V.; Bunkoczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 213.

(289)   Emsley, P.; Cowtan, K. *Acta Crystallogr. D Biol. Crystallogr.* **2004**, *60*, 2126.

(290)   Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr. D Biol. Crystallogr.* **2010**, *66*, 12.

(291)   Weininger, D. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 237.

(292)   McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658.

(293)   Beaman, T. W.; Vogel, K. W.; Drueckhammer, D. G.; Blanchard, J. S.; Roderick, S. L. *Protein Sci.* **2002**, *11*, 974.

(294)   Widboom, P. F.; Fielding, E. N.; Liu, Y.; Bruner, S. D. *Nature* **2007**, *447*, 342.

(295)   Fielding, E. N.; Widboom, P. F.; Bruner, S. D. *Biochemistry* **2007**, *46*, 13994.

(296)   Liu, Y.; Bruner, S. D. *ChemBioChem* **2007**, *8*, 617.

(297)   Laskowski, R. A.; Hutchinson, E. G.; Michie, A. D.; Wallace, A. C.; Jones, M. L.; Thornton, J. M. *Trends Biochem. Sci.* **1997**, *22*, 488.

(298)   Kelley, L. A.; Sternberg, M. J. E. *Nat. Protocols* **2009**, *4*, 363.

(299)   Benning, M. M.; Taylor, K. L.; Liu, R. Q.; Yang, G.; Xiang, H.; Wesenberg, G.; DunawayMariano, D.; Holden, H. M. *Biochemistry* **1996**, *35*, 8103.

(300)   Engel, C. K.; Mathieu, M.; Zeelen, J. P.; Hiltunen, J. K.; Wierenga, R. K. *EMBO J.* **1996**, *15*, 5135.

(301)   Wu, W. J.; Anderson, V. E.; Raleigh, D. P.; Tonge, P. J. *Biochemistry* **1997**, *36*, 2211.

(302)   Bell, A. F.; Feng, Y. G.; Hofstein, H. A.; Parikh, S.; Wu, J. Q.; Rudolph, M. J.; Kisker, C.; Whitty, A.; Tonge, P. J. *Chem. Biol.* **2002**, *9*, 1247.

(303)   Hall, P. R.; Wang, Y. F.; Rivera-Hainaj, R. E.; Zheng, X. J.; Pustai-Carey, M.; Carey, P. R.; Yee, V. C. *EMBO J.* **2003**, *22*, 2334.

(304)   Bennett, J. P.; Bertin, L.; Moulton, B.; Fairlamb, I. J. S.; Brzozowski, A. M.; Walton, N. J.; Grogan, G. *Biochem. J.* **2008**, *414*, 281.

(305)   Thoden, J. B.; Ringia, E. A. T.; Garrett, J. B.; Gerlt, J. A.; Holden, H. M.; Rayment, I. *Biochemistry* **2004**, *43*, 5716.

(306)   Sugadev, R.; Burley, S. K.; Swaminathan, S. *New York SGX Research Center for Structural Genomics (NYSGXRC)* **2007**.

(307)   Thompson, T. B.; Garrett, J. B.; Taylor, E. A.; Meganathan, R.; Gerlt, J. A.; Rayment, I. *Biochemistry* **2000**, *39*, 10662.

(308)   Schuttelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallographica Section D-Biological Crystallography* **2004**, *60*, 1355.

(309)   Gandour, R. D. *Bioorg. Chem.* **1981**, *10*, 169.

(310)   Ramanadham, M.; Jakkal, V. S.; Chidambaram, R. *FEBS Lett.* **1993**, *323*, 203.

(311)   Wong, B. J.; Gerlt, J. A. *Biochemistry* **2004**, *43*, 4646.

(312)   Pelletier, D. A.; Harwood, C. S. *J. Bacteriol.* **1998**, *180*, 2330.

(313)   Clamp, M.; Cuff, J.; Searle, S. M.; Barton, G. J. *Bioinformatics* **2004**, *20*, 426.

(314)   Waterhouse, A. M.; Procter, J. B.; Martin, D. M. A.; Clamp, M.; Barton, G. J. *Bioinformatics* **2009**, *25*, 1189.

(315)   Hofstein, H. A.; Feng, Y.; Anderson, V. E.; Tonge, P. J. *Biochemistry* **1999**, *38*, 9508.

(316)   Bahnson, B. J.; Anderson, V. E.; Petsko, G. A. *Biochemistry* **2002**, *41*, 2621.

(317)   Engel, C. K.; Kiema, T. R.; Hiltunen, J. K.; Wierenga, R. K. *J. Mol. Biol.* **1998**, *275*, 847.

(318)   Mursula, A. M.; Hiltunen, J. K.; Wierenga, R. K. *FEBS Lett.* **2004**, *557*, 81.

(319)   Xiang, H.; Luo, L.; Taylor, K. L.; Dunaway-Mariano, D. *Biochemistry* **1999**, *38*, 7638.

(320)   Modis, Y.; Filppula, S. A.; Novikov, D. K.; Norledge, B.; Hiltunen, J. K.; Wierenga, R. K. *Structure* **1998**, *6*, 957.

(321)   Muller-Newen, G.; Stoffel, W. *Biochemistry* **1993**, *32*, 11405.

(322)   Eberhard, E. D.; Gerlt, J. A. *J. Am. Chem. Soc.* **2004**, *126*, 7188.

(323)   Igbavboa, U.; Leistner, E. *Eur. J. Biochem.* **1990**, *192*, 441.

(324)   Leonard, P. M.; Grogan, G. *J. Biol. Chem.* **2004**, *279*, 31312.

(325)   Bennett, J. P.; Whittingham, J. L.; Brzozowski, A. M.; Leonard, P. M.; Grogan, G. *Biochemistry* **2007**, *46*, 137.

(326)   Crawford, I. P. *Annual Review of Microbiology* **1989**, *43*, 567.

(327)   Spraggon, G.; Kim, C.; Nguyen-Huu, X.; Yee, M.-C.; Yanofsky, C.; Mills, S. E. *PNAS* **2001**, *98*, 6021.

(328)   Morollo, A. A.; Eck, M. J. *Nat Struct Mol Biol* **2001**, *8*, 243.

(329)   Knochel, T.; Ivens, A.; Hester, G.; Gonzalez, A.; Bauerle, R.; Wilmanns, M.; Kirschner, K.; Jansonius, J. N.; PNAS: 1999; Vol. 96, p 9479.

(330)   Teng, C. Y. P.; Ganem, B. *Journal of the American Chemical Society* **1984**, *106*, 2463.

(331)   Policastro, P. P.; Au, K. G.; Walsh, C. T.; Berchtold, G. A. *Journal of the American Chemical Society* **1984**, *106*, 2443.

(332)   Morollo, A. A.; Bauerle, R. In PNAS: 1993; Vol. 90, p 9983.

(333)   Walsh, C. T.; Liu, J.; Rusnak, F.; Sakaitani, M. *Chemical Reviews* **1990**, *90*, 1105.

(334)   He, Z.; Toney, M. D. *Biochemical Journal* **2006**, *45*, 5019.

(335)   Parsons, J. F.; Jensen, P. Y.; Pachikara, A. S.; Howard, A. J.; Eisenstein, E.; Ladner, J. E.; Biochemistry: 2002; Vol. 41, p 2198.

(336)   Roux, B.; Walsh, C. T. *Biochemistry* **1992**, *31*, 6904.

(337)   Rayl, E. A.; Green, J. M.; Nichols, B. P. *Biochimica Et Biophysica Acta-Protein Structure and Molecular Enzymology* **1996**, *1295*, 81.

(338)   Viswanathan, V. K.; Green, J. M.; Nichols, B. P. *Journal of Bacteriology* **1995**, *177*, 5918.

(339)   Ye, Q. Z.; Liu, J.; Walsh, C. T. *Proceedings of the National Academy of Sciences of the United States of America* **1990**, *87*, 9391.

(340)   Basset, G. J. C.; Quinlivan, E. P.; Ravanel, S.; Rebeille, F.; Nichols, B. P.; Shinozaki, K.; Seki, M.; Adams-Phillips, L. C.; Giovannoni, J. J.; Gregory, J. F.; Hanson, A. D. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, *101*, 1496.

(341)   Harrison, A. J.; Yu, M.; Gardenborg, T.; Middleditch, M.; Ramsay, R. J.; Baker, E. N.; Lott, J. S. *Journal of Bacteriology* **2006**, *188*, 6081.

(342)   Manos-Turvey, A.; Bulloch, E. M. M.; Rutledge, P. J.; Baker, E. N.; Lott, J. S.; Payne, R. J. *Chemmedchem* **2010**, *5*, 1067.

(343)   Kerbarh, O.; Chirgadze, D. Y.; Blundell, T. L.; Abell, C. *Journal of Molecular Biology* **2006**, *357*, 524.

(344)   Parsons, J. F.; Shi, K. M.; Ladner, J. E. *Acta Crystallographica Section D-Biological Crystallography* **2008**, *64*, 607.

(345)   Sridharan, S.; Howard, N.; Kerbarh, O.; Blaszczyk, M.; Abell, C.; Blundell, T. L. *Journal of Molecular Biology* **2010**, *397*, 290.

(346)   Zalkin, H.; Kling, D. *Biochemistry* **1968**, *7*, 3566.

(347)   Patel, N.; Holmes, W. M.; Kane, J. F. *Journal of Bacteriology* **1974**, *119*, 220.

(348)   DeClue, M. S.; Baldridge, K. K.; Kast, P.; Hilvert, D. *Journal of the American Chemical Society* **2006**, *128*, 2043.

(349)   Dopheide, T. A. A.; Crewther, P.; Davidson, B. E. *JBC* **1972**, *247*, 4447.

(350)   Van Lanen, S. G.; Lin, S.; Shen, B. *PNAS* **2008**, *105*, 494.

(351)   Anthony A. Morollo, J. M. G. F., and Ronald Bauerle *JACS* **1993**, *115*, 816.

(352)   Peter P. Policastro, K. G. A., Christopher T. Walsh, and Glenn A. Berchtold *JACS* **1984**, *106*, 2443.

(353)   Olivier Kerbarh; Alessio Ciulli; Dimitri Y. Chirgadze; Tom L. Blundell; Abell, C. *ChemBioChem* **2007**, *8*, 622.

(354)   Gaille, C.; Kast, P.; Haas, D. *JBC* **2002**, *277*, 21768.

(355)   Romero, R. M.; Roberts, M. F.; Phillipson, J. D. *Phytochemistry* **1995**, *39*, 263.

(356)   Caligiuri, M. G.; Bauerle, R. *JBC* **1991**, *266*, 8328.

(357)   Zalkin, H.; Hwang, L. H. *JBC* **1971**, *246*, 6899.

(358)   Tutino, M. L.; Tosco, A.; Marino, G.; Sannia, G. *Biochemical and Biophysical Research Communications* **1997**, *230*, 306.

(359)   Wright, S. K.; DeClue, M. S.; Mandal, A.; Lee, L.; Wiest, O.; Cleland, W. W.; Hilvert, D. *JACS* **2005**, *127*, 12957.

(360)   Kunzler, D. E.; Sasso, S.; Gamper, M.; Hilvert, D.; Kast, P. *JBC* **2005**, *280*, 32827.

(361)   Xie, G.; Keyhani, N. O.; Bonner, C. A.; Jensen, R. A. *Microbiol. Mol. Biol. Rev.* **2003**, *67*, 303.

(362)   Merino, E.; Jensen, R. A.; Yanofsky, C. *Current Opinion in Microbiology* **2008**, *11*, 78.

(363)   Nagano, H.; Zalkin, H. *JBC* **1970**, *245*, 3097.

(364)   Tso, J. Y.; Zalkin, H. *JBC* **1981**, *256*, 9901.

(365)   Kane, J. F.; Holmes, W. M.; Smiley, K. L., Jr.; Jensen, R. A. *Journal of Bacteriology* **1973**, *113*, 224.

(366)   Kanno, T.; Koji, K.; Yasuko, I.-K.; Kyo, W.; Yuzuru, T. *Plant Molecular Biology* **2004**, *V54*, 11.

(367)   Massière, F.; Badet-Denisot, M. A. *Cellular and Molecular Life Sciences (CMLS)* **1998**, *54*, 205.

(368)   Strohmeier, M.; Raschle, T.; Mazurkiewicz, J.; Rippe, K.; Sinning, I.; Fitzpatrick, T. B.; Tews, I. *PNAS* **2006**, *103*, 19284.

(369)   Kim, S.-K.; Reddy, S. K.; Nelson, B. C.; Vasquez, G. B.; Davis, A.; Howard, A. J.; Patterson, S.; Gilliland, G. L.; Ladner, J. E.; Reddy, P. T. *Journal of Bacteriology* **2006**, *188*, 8638.

(370)   Qamra, R.; Prakash, P.; Aruna, B.; Hasnain, S. E.; Mande, S. C. *Biochemistry* **2006**, *45*, 6997.

(371)   Prakash, P.; Aruna, B.; Sardesai, A. A.; Hasnain, S. E. *JBC* **2005**, *280*, 19641.

(372)   Sasso, S.; Ramakrishnan, C.; Gamper, M.; Hilvert, D.; Kast, P. *FEBS Journal* **2005**, *272*, 375.

(373)   Prakash, P.; Pathak, N.; Hasnain, S. E. *JBC* **2005**, *280*, 20666.

(374)   Ito, J., Yanofsky, C. *Journal of Bacteriology* **1969**, *97*, 734.

(375)   Ito, J.; Cox, E. C.; Yanofsky, C. *Journal of Bacteriology* **1969**, *97*, 725.

(376)   Mandal, A.; Hilvert, D. *JACS* **2003**, *125*, 5598.

(377)   Sheng, Z.; Kongsaeree, P.; Clardy, J.; Wilson, D. B.; Ganem, B. *Bioorganic & Medicinal Chemistry* **1996**, *4*, 1015.

(378)   Turnbull, J.; Cleland, W. W.; Morrison, J. F. *Biochemistry* **1990**, *29*, 10245.

(379)   Zhang, S.; Pohnert, G.; Kongsaeree, P.; Wilson, D. B.; Clardy, J.; Ganem, B. *JBC* **1998**, *273*, 6248.

(380)   Mavrodi, D. V.; Blankenfeldt, W.; Thomashow, L. S. In *Annual Review of Phytopathology* 2006; Vol. 44, p 417.

(381)   Mentel, M.; Ahuja, E. G.; Mavrodi, D. V.; Breinbauer, R.; Thomashow, L. S.; Blankenfeldt, W. *Chembiochem* **2009**, *10*, 2295.

(382)   Li, Q.-A.; Mavrodi, D. V.; Thomashow, L. S.; Roessle, M.; Blankenfeldt, W. *Journal of Biological Chemistry* **2011**, *286*, 18213.

(383)   DeClue, M. S.; Baldridge, K. K.; Kunzler, D. E.; Kast, P.; Hilvert, D. *JACS* **2005**, *127*, 15002.

(384)   Ganem, C.-Y. P. T. a. B. *JACS* **1984**, *106*, 2463.