STRUCTURAL STUDIES OF ENZYMES INVOLVED IN ANTIBIOTIC RESISTANCE AND PHOSPHONATE DEGRADATION

BY

VINAYAK AGARWAL

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Biophysics and Computational Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Satish K. Nair, Chair, Director of Research
Professor Susan A. Martinis
Professor Wilfred A. van der Donk
Professor Huimin Zhao

Abstract

Biological chemistry revolves around the abiogenetic interplay of high energy labile chemical linkages which shuttle energy and chemical potential, and stable and degradation resistant bonds that make up the biological polymers which offer structural and functional support. Insights into secondary metabolism in microorganisms have shed light on fascinating arrays of genetically encoded small molecule natural products which have utilized the chemical frameworks of transient high energy compounds, and replaced key labile bonds with those borrowed from stable structural molecules. Compounds such as these can be designed as mimics of enzymatic transition states, which thus inhibit the respective enzymes by binding to them, but are not turned over to products. One such example is the Trojan horse tRNA synthetase inhibitor antibiotic microcin C7, which mimics the aminoacyl adenylate intermediate of the tRNA synthetase enzyme by replacing the high energy phosphodiester bond of the enzyme intermediate with a stable phosphoramine-amide linkage. Second examples are the phosphonate molecules which replace the high energy phosphoryl and carboxyl ester moieties of various enzyme intermediates by stable phosphonate linkages. These compounds in addition to possessing bioinhibitory activities are also thought to be store houses of carbon and phosphorous in the microbial metabolome as they cannot be easily degraded. The body of work presented in this dissertation delves into molecular mechanisms of enzymatic degradation of the stable phosphoramine-amide bond of microcin C7, as well as the hydrolysis of the phosphonate bond in a ubiquitous phosphonate molecule- phosphonoacetate. Both these studies, presented in the form of two chapters, involve the atomic resolution structure determination of the enzyme catalysts during various stages of their catalytic cycles, and in complex with different substrate molecules which delineate their specificity and selectivity. All structural investigations have been followed

by extensive biochemical validation of the structural results. These results have allowed us to postulate the reaction mechanisms for these catalysts. Our proposals reveal how enzymatic architectures have been borrowed from primary metabolism, and the enzyme active sites slightly tweaked to achieve remarkable and unprecedented chemical transformations by these enzymes of microbial secondary metabolism. Our studies also have implications for the design of inhibitors of these enzymes, as similar mechanisms may be used to avert the bioinhibitory action of various antibiotics by drug resistant pathogenic microorganisms.

Acknowledgements

I would like to express my gratitude to all those who have given me the support and the encouragement in the process of completing my doctoral degree in the Center for Biophysics and Computational Biology in the University of Illinois at Urbana-Champaign. My research work contains their indispensable contributions.

I am deeply grateful to my mentor, Professor Satish K. Nair, who has been a constant source of encouragement and wisdom during the course of my graduate studies in his laboratory. He generously provided me with his time and attention, and the necessary resources to conduct my research. Under his tutelage, I have been able to learn and develop as a scientist. I will continue to look up to Dr. Nair for scientific inputs, encouragement as well as inspiration during the course of my future endeavors.

I would also like to thank Dr. Houjin Zhang and Brian Bae, members of the Nair laboratory and who have since moved on to their respective scientific careers. Their tireless efforts to inculcate in me the basic qualities requisite of a graduate student during the early phases of my research are much appreciated.

I have been fortunate to be able to work amongst, and learn from fellow students, post-doctoral associates and scientists at the Mining Microbial Genomic Theme, at the Institute for Genomic Biology, University of Illinois at Urbana-Champaign. I would specially like to thank Professor Wilfred van der Donk, Professor William W. Metcalf, Dr. Huimin Zhao, and Dr. Svetlana A. Borisova for their scientific assistance, theoretical inputs and constant encouragement. I would like to thank Professor Konstantin Severinov at the Rutgers, State

University of New Jersey, and members of his laboratory for their scientific assistance and collaboration.

My most sincere thanks to Cindy Dodds, for guiding me through the graduate school requirements and procedures. I appreciate the hard work and professionalism of all the staff members of the School of Chemical Sciences and the university in taking care of the needs of the students and scientists.

My family has always stood by me during my travails through the duration of my graduate studies. I owe them a debt of gratitude. A special mention and thanks to Neha Garg, for her help, support and grounding for the last five years, and for a lifetime ahead.

Table of contents

Chapter 1: Introduction[1,2]

Interactions of prokaryotic microorganisms with human beings and plants in the biosphere is widespread and significant, with the number of bacterial cells within the human body exceeding that of the human eukaryotic cells themselves. This interaction is designed to be mutually beneficial to both the human or plant host and prokaryotes. However, certain microorganisms have been known to cause disease in human beings and plants, hence called pathogenic microorganisms. The mechanisms by which these pathogenic organisms cause disease are extremely diverse and affect nearly all types of tissues. With the advancement of medical science, humans have developed two frontline defense mechanisms to fight infections and diseases caused by microbes. The first line of defense are the vaccines, which are designed to tune the human host's adaptive immune response to recognize and eliminate the pathogenic and harmful microbes, while selectively preserving the beneficial microbe population within the host. The second is the discovery and development of antibiotics, which are small organic

---

[1]Parts of this chapter has been adapted with permission from:

1. Agarwal V, Tikhonov A, Metlitskaya A, Severinov K, Nair SK. Proc Natl Acad Sci U S A. 2012 Mar 20;109(12):4425-30. Referred to as 'Agarwal et. al., 2012' in this chapter. Reproduced with the permission of the publisher and is available from http://www.pnas.org/ and using DOI: 10.1073/pnas.1114224109.
2. Agarwal V, Metlitskaya A, Severinov K, Nair SK. J Biol Chem. 2011 Jun 17;286(24):21295-303. Referred to as 'Agarwal et. al., 2011b' in this chapter. Reproduced with the permission of the publisher and is available from http://www.jbc.org/ and using DOI: 10.1074/jbc.M111.226282.
3. Agarwal V, Nair SK. Med Chem Comm. 2012. Referred to as 'Agarwal and Nair, 2012' in this chapter. Reproduced with the permission of the publisher and is available from http://pubs.rsc.org/en/journals/journalissues/md and using the DOI: 10.1039/c2md20032e.
4. Agarwal V, Borisova SA, Metcalf WW, van der Donk WA, Nair SK. Chem Biol. 2011 Oct 28;18(10)1230-40. Referred to as 'Agarwal et. al., 2011a' in this chapter. Reproduced with the permission of the publisher and is available from http://www.cell.com/ and using DOI: 10.1016/j.chembiol.2011.07.019.

[2]Parts of this chapter are in preparation for publication:

1. Agarwal V, Chen JH, Chekan JR, Borisova SA, Metcalf WW, van der Donk WA, Nair SK. Structural and functional characterization of phosphonoacetaldehyde dehydrogenase enzyme provides missing link for phosphonoacetate degradation, 2012, in preparation for publication.

molecules, designed to target and kill the harmful microorganisms within the human body. Antibiotics are derived from naturally occurring molecules, which are used within the biosphere to render selective advantage during interactions of different living species for competition for nutrients and survival advantage. Serendipitous discovery and exploitation of these molecules by human beings dates back in ancient history (Forrest, 1982a, b; Lindblad, 2008) with numerous reports such as the use of the bark from cinchona tree, which contains the antibiotic quinine, used to treat malaria. Antibiotics generally have well defined mechanisms of action against the pathogenic microbes, and target specific biological processes within these microbes to either retard their growth (bacteriostatic antibiotics), or kill the microbe cell (bacteriocidal antibiotics).

The first report of an isolated antibiotic molecule used to effectively treat infectious diseases is the development and use of penicillin, discovered by Alexander Fleming in 1928 from the fungus *Penicillium notatum.* Penicillin was used to treat diseases such as syphilis and infections caused by staphylococci and streptococci. The term 'antibiotic' was coined by eminent microbiologist Selman Waksman in 1942 (Waksman, 1947), and is derived from *antibiosis*, which translates to 'against life'. Hailed as a wonder drug, penicillin marked the beginning of the antibiotic age, in which antibiotics have been developed to treat almost all major human non-viral infectious diseases. Thousands of antibiotic molecules have been discovered and applied to fight bacterial, fungal, protozoan and yeast infections. Synthetic derivatives of naturally occurring antibiotic molecules have also been generated with more desirable properties. These molecules have marked a decreased incidence of outbreaks of infectious diseases and have contributed to increased life expectancy in the world.

However, repeated challenge of the pathogenic microorganisms against these antibiotics, as well as natural evolution of their genetic information leads to the microorganisms developing

resistance against these antibiotics, by which they are able to evade the action of the antibiotic molecules (Allen et al., 2010; Davies and Davies, 2010; Moellering, 1998). This process has led to the development of microbial resistance against nearly all available antibiotic molecules, and has precipitated to an immense medical and financial challenge. The arsenal available to fight infections is becoming increasingly limited and the rate of development of new molecules has not been able to keep pace with the emergence of antibiotic resistance. Several factors contribute to the incidence of antibiotic resistance among susceptible microorganisms. Primary among these is the widespread and indiscriminate overuse of antibiotics in the clinical environment, which exposes pathogenic microorganisms to an ever increasing concentration of different antibiotic molecules, which in turn increases the selection pressure to develop resistance against these antibiotic molecules (Goossens et al., 2005; Tacconelli et al., 2008). Secondly, the use of antibiotics in agriculture and livestock rearing also leads to increased exposure of pathogenic bacteria to antibiotic molecules. Many of the antibiotics present in agriculture and livestock feeds are also used for treating humans, and hence development of resistant bacterial strains can be transferred from one setting to the other very easily (Angulo et al., 2004; Mathew et al., 2007; Singer et al., 2003). Microbes are also proficient in genetic recombination and the transfer of genes from an antibiotic resistant microbe to a non-resistant microbe, which leads to spread of resistance among microbial species. Antibiotic resistance genes are usually borne on highly mobile genetic vehicles called plasmids, which are easily transmitted among microbes (Koch, 1981).

Microorganisms develop resistance to antibiotics by myriad mechanisms (Tenover, 2006). Efflux pumps in the microbial cell walls and membranes actively transport the antibiotic molecule out of the cell and hence decrease its concentration within the cell. Consequently ever

increasing doses of antibiotics are required against such microbes (Li and Nikaido, 2009; Lin et al., 2009; Van Bambeke et al., 2000). A primary example is the *tetA* gene which affects the efflux of tetracycline out of the susceptible cells (Van Bambeke et al., 2003). The antibiotic target within the target cell can be altered and modulated, so that the antibiotic is no longer able to bind and engage the target. Random mutations within the antibiotic target genes may lead to the emergence and subsequent selection of microbe species which harbor a desensitized antimicrobial target. Modulation of the D-alanine-D-alanine dipeptide within the bacterial cell wall to D-alanine-D-lactate mediates resistance against vancomycin (French, 1998). Presence of additional antibiotic resistant enzyme targets within the susceptible cells mediates resistance against the topical antibiotic mupirocin (Patel et al., 2009). Antibiotics can also be chemically modified by specific enzymes within the target pathogenic microbe, which renders the antibiotic ineffective (Wright, 2005). Chemical modifications include, but are not limited to hydrolysis, acetylation, adenylation and phosphorylation of the antibiotic molecules. Hydrolysis of antibiotics is of particular concern against β-lactam antibiotics, such as the parent compound and derivatives of penicillin (Bradford, 2001). Also the aminoglycoside antibiotics, such as kanamycin and neomycin are acetylated by dedicated acetyltransferase and phosphotransferase enzymes (Smith and Baker, 2002). It is thus a continuous endeavor on part of the academic research as well as the pharmaceutical industry to search for antibiotics which have new and as yet unexplored mode of actions. Also new chemical classes of drug molecules are pursued to mitigate the effects of resistance against the current drug chemical entities.

An underappreciated approach towards combating antibiotic resistance is inhibiting the processes which mediate resistance against the antibiotic. A shining example of application of this methodology is the development of clavulanate, which is an irreversible covalent inhibitor of

the β-lactamase enzyme which hydrolyzes the β-lactam antibiotics (Drawz and Bonomo, 2010). When used in conjunction with the β-lactam antibiotic meropenem, clavulanate was shown to be active against multi drug resistant (MDR) and extremely drug resistant (XDR) strains of tuberculosis causing *Mycobacterium* strains (Hugonnet et al., 2009). Clavulanate in conjunction with amoxicillin is clinically used under the trade name Augmentin. Similarly, β-lactam antibiotic ampicillin is clinically used in conjunction with the β-lactamase inhibitor sulbactam under the trade name Unasyn. Additional efforts are underway to develop inhibitors of the antibiotic efflux pumps, which would increase the efficacy of antibiotics such as tetracycline (Van Bambeke et al., 2000; Van Bambeke et al., 2003). Characterization of the aminoglycoside acetyltransferase enzymes and efforts towards the development of their mechanism based inhibitors (Chen et al., 2011; Magalhaes et al., 2008; Green et. al., 2012) is directed at improving the efficacy of these aminoglycoside antibiotics commonly used in the treatment regimen against tuberculosis.

In this rich background of efforts to reinvigorate the efficacy of antibiotics by combating their inactivating processes, we have described the microbial mechanisms which can be employed to inactivate nucleoside antibiotics such as the Trojan horse aspartyl tRNA synthetase inhibitor antibiotic microcin C7 (McC7). The chemical mechanisms used to inactivate McC7 are distinct from previously used mechanisms employed by resistant microbes against aminoacyl tRNA synthetase inhibitor antibiotics (Hurdle et al., 2005). We describe the crystal structure of biochemical characterization of two enzymes- MccF and MccE which respectively hydrolyze and acetylate McC7 to render it ineffective (Agarwal et al., 2011b; Agarwal et al., 2012). Our investigations shed light on how similar enzymes already present in the microbial metabolome can be used to provide resistance against other nucleoside antibiotics. Our efforts also provide

the structural groundwork for the design of inhibitors against these enzymes. The structures of the two enzymes demonstrate how enzymes of the microbial primary metabolism have been borrowed by processes of secondary metabolism. Though the core tertiary structures of MccF and MccE resemble their counterparts from primary metabolism, subtle changes have been incorporated within the active sites to recognize and chemically modify McC7. A structure based explanation of enzymatic specificity, and biochemical validation of the structural findings is the underlying theme in this body of work.

An emerging class of antibiotics is the phosphonate molecules (Metcalf and van der Donk, 2009). Phosphonate molecules mimic diverse phosphorylated enzyme intermediates, but have replaced the labile oxygen-phosphorus (O-P) bonds of these enzyme intermediates by stable carbon-phosphorus (C-P) bonds, while preserving both the tetrahedral geometry of the phosphorus atom and the charge on the phosphoryl oxygen atoms. Their applicability to medicine is exemplified by the anti-malaria phosphonate antibiotic FR-900098 and its derivatives (Uh et al., 2011), antifungal rhizocticin (Kugler et al., 1990), antibacterial plumbemycin and the broad spectrum antibiotics fosfomycin (Falagas et al., 2008) and dehydrophos (Circello et al., 2011). Besides antibiotics, phosphonate molecules find extensive use in agriculture and industry as well. Phosphonate molecules are also essential phosphorus, carbon and nitrogen sources. In view of their utilization as growth sources by microbes, it is not surprising that microbes have evolved myriad mechanisms to hydrolyze the C-P bonds in phosphonates to release phosphate and the corresponding organic molecules (White and Metcalf, 2007). These hydrolysis mechanisms offer unprecedented insights into varied reaction mechanisms of microbial secondary metabolism.

A ubiquitous phosphonate molecule in the microbial metabolome is phosphonoacetate (PnA). Though the biogenic origin of PnA had not been established until last year, the widespread prevalence of phosphonoacetate hydrolase enzyme in prokaryotes hinted towards dedicated PnA biosynthetic machinery being present as well. Indeed, the biosynthesis of PnA was subsequently established from the most widespread phosphonate molecule- 2-aminoethyl phosphonate (Borisova et al., 2011). In this work, we have structurally and biochemically characterized the PnA biosynthetic enzyme PhnY. PhnY oxidizes phosphonoacetaldehyde (produced by the transamination of 2-AEP) to generate PnA. Our structural and kinetic data establishes that PhnY borrows the mechanism and substrate recognition motifs from the essential primary metabolic enzyme – glyceraldehyde 3-phosphate dehydrogenase. Slight structural perturbations in the active site have rendered the enzyme specific towards phosphonate molecules, albeit at a kinetic cost. We have also determined the structure of, and established a mechanistic proposal for PnA hydrolase enzyme PhnA (Agarwal et al., 2011a). The active site of PhnA has been adapted from the ubiquitous alkaline phosphatase enzymes, though the stabilization of the reaction intermediate required a structural reorganization of the bi-metallo active site to accommodate the phosphonate molecule. Our findings for the PhnA enzyme further build upon the findings from a recent report of the structure of the PhnA enzyme (Kim et al., 2011), albeit with a different mechanistic proposal for reaction intermediate stabilization. This body of work is also motivated by structural explanation for substrate specificity within the enzyme active site, and biochemical validation of the structural findings.

As has been mentioned previously, and will be delineated in the subsequent chapters, we have relied heavily on high resolution structure determination by the technique of X-ray crystallography. All enzymes investigated have been co-crystallized in the presence of their

respective substrates and cofactors, and inhibitors where applicable. We have further employed myriad biochemical assays and site specific mutagenesis to probe the role of enzyme active site amino acids in catalysis and in generating substrate specificity.

Through our findings we have attempted to build an appreciation for the immense untapped potential for the discovery of enzymatic chemistry involved in small molecule natural product biosynthesis, modification and degradation. Further sequencing of microbial genomes to discover cryptic small molecule gene clusters, together with metagenomic approaches to discover novel small molecule natural products within the microbial metabolome will further enrich this field.

References

Agarwal, V., Borisova, S.A., Metcalf, W.W., van der Donk, W.A., and Nair, S.K. (2011a). Structural and mechanistic insights into C-P bond hydrolysis by phosphonoacetate hydrolase. Chem Biol *18*, 1230-1240.

Agarwal, V., Metlytskaya, A., Severinov, K.V., and Nair, S.K. (2011b). Structural basis for Microcin C7 inactivation by the MccE acetyltransferase. J Biol Chem.

Agarwal, V., Tikhonov, A., Metlitskaya, A., Severinov, K., and Nair, S.K. (2012). Structure and function of a serine carboxypeptidase adapted for degradation of the protein synthesis antibiotic microcin C7. Proc Natl Acad Sci U S A *109*, 4425-4430.

Allen, H.K., Donato, J., Wang, H.H., Cloud-Hansen, K.A., Davies, J., and Handelsman, J. (2010). Call of the wild: antibiotic resistance genes in natural environments. Nat Rev Microbiol *8*, 251-259.

Angulo, F.J., Nargund, V.N., and Chiller, T.C. (2004). Evidence of an association between use of anti-microbial agents in food animals and anti-microbial resistance among bacteria isolated from humans and the human health consequences of such resistance. J Vet Med B Infect Dis Vet Public Health *51*, 374-379.

Borisova, S.A., Christman, H.D., Metcalf, M.E., Zulkepli, N.A., Zhang, J.K., van der Donk, W.A., and Metcalf, W.W. (2011). Genetic and biochemical characterization of a pathway for the degradation of 2-aminoethylphosphonate in Sinorhizobium meliloti 1021. J Biol Chem *286*, 22283-22290.

Bradford, P.A. (2001). Extended-spectrum beta-lactamases in the 21st century: characterization, epidemiology, and detection of this important resistance threat. Clin Microbiol Rev *14*, 933-951, table of contents.

Chen, W., Biswas, T., Porter, V.R., Tsodikov, O.V., and Garneau-Tsodikova, S. (2011). Unusual regioversatility of acetyltransferase Eis, a cause of drug resistance in XDR-TB. Proc Natl Acad Sci U S A *108*, 9804-9808.

Circello, B.T., Miller, C.G., Lee, J.H., van der Donk, W.A., and Metcalf, W.W. (2011). The antibiotic dehydrophos is converted to a toxic pyruvate analog by peptide bond cleavage in Salmonella enterica. Antimicrob Agents Chemother *55*, 3357-3362.

Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. Microbiol Mol Biol Rev *74*, 417-433.

Diaz-Sanchez, A.G., Gonzalez-Segura, L., Rudino-Pinera, E., Lira-Rocha, A., Torres-Larios, A., and Munoz-Clares, R.A. (2011). Novel NADPH-cysteine covalent adduct found in the active site of an aldehyde dehydrogenase. Biochem J.

Drawz, S.M., and Bonomo, R.A. (2010). Three decades of beta-lactamase inhibitors. Clin Microbiol Rev *23*, 160-201.

Falagas, M.E., Kanellopoulou, M.D., Karageorgopoulos, D.E., Dimopoulos, G., Rafailidis, P.I., Skarmoutsou, N.D., and Papafrangas, E.A. (2008). Antimicrobial susceptibility of multidrug-resistant Gram negative bacteria to fosfomycin. Eur J Clin Microbiol Infect Dis *27*, 439-443.

Forrest, R.D. (1982a). Development of wound therapy from the Dark Ages to the present. J R Soc Med *75*, 268-273.

Forrest, R.D. (1982b). Early history of wound treatment. J R Soc Med *75*, 198-205.

French, G.L. (1998). Enterococci and vancomycin resistance. Clin Infect Dis *27 Suppl 1*, S75-83.

Goossens, H., Ferech, M., Vander Stichele, R., and Elseviers, M. (2005). Outpatient antibiotic use in Europe and association with resistance: a cross-national database study. Lancet *365*, 579-587.

Hugonnet, J.E., Tremblay, L.W., Boshoff, H.I., Barry, C.E., 3rd, and Blanchard, J.S. (2009). Meropenem-clavulanate is effective against extensively drug-resistant Mycobacterium tuberculosis. Science *323*, 1215-1218.

Hurdle, J.G., O'Neill, A.J., and Chopra, I. (2005). Prospects for aminoacyl-tRNA synthetase inhibitors as new antimicrobial agents. Antimicrob Agents Chemother *49*, 4821-4833.

Kim, A., Benning, M.M., OkLee, S., Quinn, J., Martin, B.M., Holden, H.M., and Dunaway-Mariano, D. (2011). Divergence of chemical function in the alkaline phosphatase superfamily: structure and mechanism of the P-C bond cleaving enzyme phosphonoacetate hydrolase. Biochemistry *50*, 3481-3494.

Koch, A.L. (1981). Evolution of antibiotic resistance gene function. Microbiol Rev *45*, 355-378.

Kugler, M., Loeffler, W., Rapp, C., Kern, A., and Jung, G. (1990). Rhizocticin A, an antifungal phosphono-oligopeptide of Bacillus subtilis ATCC 6633: biological properties. Arch Microbiol *153*, 276-281.

Li, X.Z., and Nikaido, H. (2009). Efflux-mediated drug resistance in bacteria: an update. Drugs *69*, 1555-1623.

Lin, L., Ling, B.D., and Li, X.Z. (2009). Distribution of the multidrug efflux pump genes, adeABC, adeDE and adeIJK, and class 1 integron genes in multiple-antimicrobial-resistant clinical isolates of Acinetobacter baumannii-Acinetobacter calcoaceticus complex. Int J Antimicrob Agents *33*, 27-32.

Lindblad, W.J. (2008). Considerations for determining if a natural product is an effective wound-healing agent. Int J Low Extrem Wounds *7*, 75-81.

Magalhaes, M.L., Vetting, M.W., Gao, F., Freiburger, L., Auclair, K., and Blanchard, J.S. (2008). Kinetic and structural analysis of bisubstrate inhibition of the Salmonella enterica aminoglycoside 6'-N-acetyltransferase. Biochemistry *47*, 579-584.

Mathew, A.G., Cissell, R., and Liamthong, S. (2007). Antibiotic resistance in bacteria associated with food animals: a United States perspective of livestock production. Foodborne Pathog Dis *4*, 115-133.

Metcalf, W.W., and van der Donk, W.A. (2009). Biosynthesis of phosphonic and phosphinic acid natural products. Annu Rev Biochem *78*, 65-94.

Moellering, R.C., Jr. (1998). Antibiotic resistance: lessons for the future. Clin Infect Dis *27 Suppl 1*, S135-140; discussion S141-132.

Patel, J.B., Gorwitz, R.J., and Jernigan, J.A. (2009). Mupirocin resistance. Clin Infect Dis *49*, 935-941.

Singer, R.S., Finch, R., Wegener, H.C., Bywater, R., Walters, J., and Lipsitch, M. (2003). Antibiotic resistance--the interplay between antibiotic use in animals and human beings. Lancet Infect Dis *3*, 47-51.

Smith, C.A., and Baker, E.N. (2002). Aminoglycoside antibiotic resistance by enzymatic deactivation. Curr Drug Targets Infect Disord *2*, 143-160.

Tacconelli, E., De Angelis, G., Cataldo, M.A., Pozzi, E., and Cauda, R. (2008). Does antibiotic exposure increase the risk of methicillin-resistant Staphylococcus aureus (MRSA) isolation? A systematic review and meta-analysis. J Antimicrob Chemother *61*, 26-38.

Tenover, F.C. (2006). Mechanisms of antimicrobial resistance in bacteria. Am J Infect Control *34*, S3-10; discussion S64-73.

Uh, E., Jackson, E.R., San Jose, G., Maddox, M., Lee, R.E., Boshoff, H.I., and Dowd, C.S. (2011). Antibacterial and antitubercular activity of fosmidomycin, FR900098, and their lipophilic analogs. Bioorg Med Chem Lett *21*, 6973-6976.

Van Bambeke, F., Balzi, E., and Tulkens, P.M. (2000). Antibiotic efflux pumps. Biochem Pharmacol *60*, 457-470.

Van Bambeke, F., Glupczynski, Y., Plesiat, P., Pechere, J.C., and Tulkens, P.M. (2003). Antibiotic efflux pumps in prokaryotic cells: occurrence, impact on resistance and strategies for the future of antimicrobial therapy. J Antimicrob Chemother *51*, 1055-1065.

Waksman, S.A. (1947). What is an antibiotic or an antibiotic substance? Mycologia *39*, 565-569.

White, A.K., and Metcalf, W.W. (2007). Microbial metabolism of reduced phosphorus compounds. Annu Rev Microbiol *61*, 379-400.

Wright, G.D. (2005). Bacterial resistance to antibiotics: enzymatic degradation and modification. Adv Drug Deliv Rev *57*, 1451-1470.

Chapter 2: Structural and biochemical characterization of antibiotic microcin C7 inactivating enzymes[1]

Abstract

Several classes of naturally occurring antibiotics target the aminoacyl tRNA synthetase (aaRS) enzymes. The aspartyl tRNA synthetase inhibitor microcin C7 (McC7) consists of an aspartyl-adenylate conjugated to a hexapeptide carrier that facilitates active import into bacterial cells through an oligopeptide transport system. Subsequent proteolytic processing releases the toxic compound inside the target cell. As the processing proteases are broad spectrum non-specific enzymes, producing strains of McC7 must protect themselves against auto-toxicity that may result from premature processing within the producer cell itself. The *mccF* and *mccE* genes confer resistance against McC7 by hydrolyzing the amide bond that connects the peptide and nucleotide moieties of McC7, and acetylating the backbone primary amine of the antibiotic respectively. We have determined the wild type apo crystal structure of MccF and of catalytically deficient mutant form of the enzyme in complex with the processed form of McC7 and various other substrate analogs. The MccF structure is similar to that of dipeptide LD-carboxypeptidase, but with an additional loop proximal to the active site that serves as the primary determinant for recognition of adenylated substrates. Wild-type MccF only hydrolyzes

---

[1]Parts of this chapter have been adapted with permission from:

1. Agarwal V, Tikhonov A, Metlitskaya A, Severinov K, Nair SK. Proc Natl Acad Sci U S A. 2012 Mar 20;109(12):4425-30. Referred to as 'Agarwal et. al., 2012' in this chapter. Reproduced with the permission of the publisher and is available from http://www.pnas.org/ and using DOI: 10.1073/pnas.1114224109.
2. Agarwal V, Metlitskaya A, Severinov K, Nair SK. J Biol Chem. 2011 Jun 17;286(24):21295-303. Referred to as 'Agarwal et. al., 2011' in this chapter. Reproduced with the permission of the publisher and is available from http://www.jbc.org/ and using DOI: 10.1074/jbc.M111.226282.
3. Agarwal V, Nair SK. Med Chem Comm. 2012. Referred to as 'Agarwal and Nair, 2012' in this chapter. Reproduced with the permission of the publisher and is available fromhttp://pubs.rsc.org/en/journals/journalissues/md and using the DOI: 10.1039/c2md20032e.

the aminoacyl-nucleotide conjugates that contain anionic side chains. We show that substitutions of two active site MccF residues result in a specificity switch towards hydrophobic aminoacyl-nucleotide substrates. We also demonstrate that the enzyme is promiscuous for the acceptance of anionic side chains conjugated to other purine nucleotides. These results suggest how MccF-like enzymes may be used to avert various toxic aminoacyl-nucleotides. We also present characterization data for the acetylation mechanism of MccE through biochemical and structural biological studies of the acetyltransferase domain of MccE. We establish the activity of MccE against processed McC7 by kinetic analysis and show that enzyme activity is inhibited by adenosine monophosphate (AMP). The co-crystal structure of the MccE-acetyl CoA binary complex reveals a structural similarity with the Gcn5-acetyltransferase superfamily but with a novel aromatic platform located adjacent to the thioester. Ternary complex structures of MccE-acetyl CoA with AMP, and MccE-CoA with the substrate mimics aspartyl sulfamoyl adenosine (DSA) and glutamyl sulfamoyl adenosine (ESA) show a mode of substrate recognition similar to those found in cap binding proteins and nucleotidyltransferases. The ternary complex of MccE-acetyl CoA with processed McC7 reveals a product structure with the acetyl group located on the α-amino moiety. These studies also demonstrate the biochemical potential of mechanisms for the resistance against future antibiotics which may be developed as derivatives of McC7 or other naturally occurring aminoacyl nucleotide conjugates. They also serve to lay a platform for the development of inhibitors of these enzymes to circumvent bacterial resistant against McC7 and other antibiotics targeting aaRS enzymes.

Introduction

Aminoacyl tRNA synthetase enzymes

Aminoacyl-tRNA synthetases (aaRSs) are essential components of the ribosomal protein synthesis machinery. AaRSs catalyze the attachment of the correct amino acid to its cognate transfer RNA (tRNA), which is then utilized by the ribosome to elongate a peptide chain (Ibba and Soll, 2000; Ling et al., 2009). The aaRS reaction can be divided into two half reactions. The first acylation reaction involves the formation of an enzyme bound aminoacyl-adenylate (Figure 2.1), which is then resolved into aminoacylated-tRNA through nucleophilic attack of the 2' or 3'-hydoxyl of the terminal adenine ribose of the tRNA in the second reaction (Ibba and Soll, 2000). Inhibition of either one of these two enzymatic steps disrupts tRNA charging, which, in turn, stalls elongation of growing polypeptide chains (Tao and Schimmel, 2000). In addition, several aaRS enzymes also harbor proofreading catalytic activities, in which misacylation of tRNA with a non-cognate amino acid can be corrected.

Several features highlight the suitability of these enzymes for drug discovery: a) the divergence between prokaryotic and eukaryotic aaRSs can be exploited for the development of drugs that only target the bacterial enzymes (Raczniak et al., 2001); b) each aaRS is phylogenetically conserved, and antimicrobials that target a particular bacterial aaRS have the potential to inhibit homologous enzymes from a range of other bacteria (Raczniak et al., 2001); and c) most bacterial species contain twenty distinct aaRSs, each representing an independent target for drug discovery (Hurdle et al., 2005; Pohlmann and Brotz-Oesterhelt, 2004). Consequently, research efforts over the last two decades have focused on the identification and characterization of these natural product aaRS inhibitors.

AaRSs as targets for antibiotic development

Several naturally occurring small molecule antimicrobials target the aaRSs, validating these enzymes as drug targets (Kim et al., 2003; M. Ibba, 2005; Robert Chênevert, 2003; Vondenhoff and Van Aerschot, 2011). One class of such natural product aaRS inhibitors consist of the polyketides such as mupirocin (targets IleRS) (Thomas et al., 2010), reveromycin (targets IleRS) (Fremlin et al., 2011), borrelidin (targetsThrRS) (Otoguro et al., 2003), and granaticin (targets LeuRS) (Hurdle et al., 2005). Mupirocin is currently the world's leading topical antibiotic for the treatment of methicillin resistant *Staphylococcus aureus* (MRSA) infections (Boyce, 2001), and borrelidin was shown to be effective against *Plasmodium*, the causative agent for malaria (Otoguro et al., 2003), besides its anti-angiogenesis properties. Successes have also been achieved with synthetic aaRS inhibitors, such as benzoxaborole compound AN-2690 that inhibits LeuRS (Baker et al., 2006; Rock et al., 2007). AN-2690 is currently under clinical trials for the treatment of onychomycosis (Alley et al., 2007), which is a fungal infection targeting the fingers and toenails.

A second class of aaRS inhibitors consist of non-hydrolyzable analogs of the aminoacyl-adenylate intermediate that is generated during the first aaRS half reaction (Tao and Schimmel, 2000). While such inhibitors can be readily generated by synthetic chemistry methods, they are ineffective as antibiotics because these compounds can also inhibit human aaRSs (Heacock et al., 1996) and most lack *in vivo* activity as they do not readily penetrate bacterial cells (Kim et al., 2003; Schimmel et al., 1998). Nature has overcome this limitation by using a "Trojan horse" strategy for the deployment of aminoacyl-adenylate analogs into target bacteria. In essence, the active portion of the drug (i.e. the aminoacyl-adenylate) is conjugated with a specialized module that facilitates transport into bacterial cells. Once inside the cell, the Trojan horse module is

removed to yield the active compound, which can then exert antibiotic activity. Examples of naturally occurring aminoacyl-adenylate analog-based antibiotics that employ this Trojan horse strategy include the LeuRS inhibitor agrocin 84 (Figure 2.2; linked through a phosphoramidate bond to a D-glucofuranosyloxyphosphoryl moiety) (Reader et al., 2005), the SerRS inhibitor albomycin (Figure 2.2; containing a covalent linkage to a hydroxymate siderophore) (Stefanska et al., 2000) and the AspRS inhibitor McC7 (Figure 2.2) (Metlitskaya et al., 2006).

Mechanism of resistance against aaRS inhibitors

The overarching chemical scheme for conferring resistance against aaRS inhibitors in producing strains is the presence of an additional aaRS enzyme (Ataide and Ibba, 2006; Hurdle et al., 2005). Mutations within the non-resistant housekeeping aaRS can also mediate immunity against the action of the antibiotic, though these mutations have usually been associated with fitness costs to the cell. Examples of biochemically characterized resistant aaRSs are encoded within the biosynthetic clusters of mupirocin, borrelidin, and agrocin84. In addition to these compounds, resistant aaRS enzymes have also been found for the TrpRS inhibitor indolmycin (Kitabatake et al., 2002; Vecchione and Sello, 2009) the trojan horse SerRS inhibitor albomycin (Zeng et al., 2009) and high throughput screening derived MetRS inhibitors SB-362916 and SB-441513 (Gentry et al., 2003) among other examples. Genome sequencing and phylogenetic analysis studies suggest that these antibiotic resistant aaRS alleles are widespread in the microbiota, (Brown et al., 2003) which raises the possibility that the efficacy of the aaRS inhibitors as antibiotics may already be compromised. Most naturally occurring inhibitors target aaRS enzymes that adenylate and transfer hydrophobic or amphipathic amino acids. Inhibitors

that target aaRSs that utilize polar amino acids have not been found in the natural microbial metabolome, and to the best of our knowledge, small molecule high throughput screening strategies have not identified lead compounds against these aaRS enzymes (Payne et al., 2007). McC7 represents the only known compound that targets an aaRS enzyme for a charged hydrophilic amino acid transfer.

Structure and antimicrobial activity of McC7

McC7 was first isolated from *Escherichia coli* H22 in 1984 (Garcia-Bustos et al., 1984; Garcia-Bustos et al., 1985) and subsequent structure determination revealed that the drug consists of a heptapeptide (Met-Arg-Thr-Gly-Asn-Ala-Asp) with a formyl group at the methionine amino terminus and an adenylate linked at the main chain carboxyl of the terminal aspartyl residue through a phosphoramidate bond (Guijarro et al., 1995) (Figure 2.2). In addition to containing the carboxy-terminal aspartyl acyl-AMP anhydride mimic, McC7 is also esterified at the *N*-acyl phosphoramidate group via a 3-aminopropyl appendage. Functional studies demonstrate that aminopropylation contributes significantly to the potency of McC7, as half-inhibition of AspRS with either a non-hydrolyzable aspartyl sulfamoyl-adenylate (DSA) or McC7 lacking the aminopropyl group requires a 10-fold higher concentrations than with mature McC7 (Metlitskaya et al., 2009). McC7 in its intact form is not bioactive.

Bioactivity for McC7 is generated by the cleavage of the first six residues of the McC7 heptapeptide to generate processed McC7 (**2**) (Figure 2.7). Processed McC7 consists only of the terminal aspartyl residue connected to the adenine nucleotide by a phosphoramide bond, and the aminopropyl appendage upon the phospho- oxygen atom. Processed McC7 thus mimics the

aspartyl adenylate formed by the aspartyl tRNA synthetase (AspRS) enzyme (**1**), and can thus bind to the AspRS active site to inhibit the charging of $tRNA^{Asp}$ (Metlitskaya et al., 2006). Modeling studies suggest that the aminopropyl group likely binds to a pocket adjacent to the active site of its target AspRS (Metlitskaya et al., 2009). Processed McC7 is bioactive against several members of the *Enterobacteriaceae* family (Garcia-Bustos et al., 1985). The producer strain of McC7 – *E. coli* H22, was shown to inhibit the growth of the human pathogen *Shigella flexneri* and subsequently results in the clearance of the pathogen from the gut of gnotobiotic mice (Cursino et al., 2006).

Biosynthesis of McC7 and the Trojan horse mechanism of action

The McC7 biosynthetic cluster from *E. coli* H22 consists of six genes consisting of ORFs that encode for both production and autoimmunity (Gonzalez-Pastor et al., 1995). Subsequently, function of several of these genes has been experimentally verified, and the biochemical route for McC7 maturation has also been established (Figure 2.3) (Kazakov et al., 2008; Metlitskaya et al., 2009; Novikova et al., 2010; Novikova et al., 2007; Severinov et al., 2007; Tikhonov et al., 2010). The *mccA* gene codes for a seven-residue, ribosomally-synthesized peptide that represents one of the smallest structural gene identified (Gonzalez-Pastor et al., 1994). The MccA heptapeptide is carboxy-terminally conjugated to AMP by MccB to yield an unprocessed McC7 precursor (i.e. the seven-residue peptide with the carboxy-terminal adenylate). Biochemical reconstitution of the reaction *in vitro* demonstrates that two stoichiometric equivalents of ATP are consumed and the reaction proceeds through a cyclic succinimide intermediate (Roush et al., 2008). The reaction results in the conversion of carboxy-terminal

asparagine of unmodified MccA into an aspartate. Other genes in the cluster include *mccC*, which shares sequence identity with members of the major facilitator superfamily, and is likely the efflux pump for McC7 export. The *mccD* gene product contains signature motifs of an S-adenosyl-methionine dependent enzyme, and the *mccE* gene product contains an amino terminal decarboxylase domain and a carboxy-terminal acetyltransferase domain. The combined actions of MccD and MccE were shown to be responsible for the installation of the propylamine ester on the nucleotide phosphate group (Metlitskaya et al., 2009) though the timing and function of these two enzymes remain to be determined.

Following biosynthesis, the McC7 heptapeptide (adenylated at the carboxy terminus and esterified by propylamine at the nucleotide phosphate) is exported out of the producer cell, and taken up by sensitive bacterial strains by virtue of recognition of the heptapeptide by specific active transporters (YejABEF in the case of *E. coli*) (Novikova et al., 2007). Bacteria that lack this transporter have reduced virulence as this operon is necessary for competitive fitness (Eswarappa et al., 2008; Qimron et al., 2004; Yang et al., 1996). Once inside the target cell, McC7 is processed by a series of enzymatic steps. First, the formyl group is removed from the amino-terminal methionine by peptide deformylase. Deformylated McC7 is then progressively digested in the N to C direction by any one of three cellular aminopeptidases (PepA, PepB and PepN in *E. coli*) (Kazakov et al., 2008). Upon hydrolysis of the peptide bond connecting the sixth and seventh residues of the peptide, the toxic processed McC7 is released inside the target cell (Metlitskaya et al., 2009). It is only the processed form of McC7 that can inhibit AspRS. The first six amino acids of the peptide mediate the transport of the antibiotic inside the susceptible cells and their removal is necessary for antibiotic action. Hence McC7 is classified as a 'Trojan horse' antibiotic.

Mechanisms for self-immunity against McC7

McC7 can be processed to its bioactive form within the producer cell, as the PepA-N proteases are present within *E. coli* H22 as well. This necessitates the presence of self-immunity mechanisms within the producer cell to protect against the action of the antibiotic. What is notable is that McC7 resistance is not mediated by a resistant AspRS enzyme, but rather through the different mechanism of chemical modifications of the antibiotic (Agarwal et al., 2011; Agarwal et al., 2012; Novikova et al., 2010; Tikhonov et al., 2010). Chemical modification mediated resistance against antibiotics, such as hydrolysis of polyketides and β-lactams, and acetylation and phosphorylation of aminoglycosides, has been documented before. However, this strategy is to date not found as a resistance mechanism against aaRS inhibitors and is therefore unique to McC7.

Self-resistance against McC7 is mediated by the MccF enzyme, which has been experimentally shown to cleave the peptide bond between the main chain carboxylate of the aspartyl residue of processed McC7 and the nucleotide phosphoramine (Figure 2.3) (Tikhonov et al., 2010). The carboxy terminal domain of MccE ($MccE^{CTD}$) also provides self-resistance for the producer cells against the action of processed McC7 by acetylating the antibiotic (Figure 2.3) (Novikova et al., 2010). Notably, MccE and MccF provide the only known examples of chemical modifications of an aaRS-targeting antibiotic to mediate resistance (Agarwal and Nair, 2012). In all other aaRS inhibitor antibiotics characterized, the principle mechanism for resistance has been the presence of an antibiotic resistant aaRS enzyme encoded by the biosynthetic cluster. It should also be noted that a McC7 resistant AspRS enzyme has not been isolated as yet.

Our studies described herein have structurally and biochemically characterized both immunity conferring enzymes for McC7- MccF and MccE^{CTD}. Atomic resolution three dimensional structures for these catalysts have been determined using X-ray crystallography, and the structural findings have been corroborated with extensive biochemical analysis. These findings have been used to explain the substrate engagement and mechanism for MccF and MccE^{CTD}. Structural and biochemical basis for the broad substrate scope for both enzymes has been established, with a view of development of resistance against other aaRS inhibitor molecules which can compromise their antibiotic efficacy.

Experimental procedures

Cloning, protein expression and purification

The construction of plasmid DNA bearing the wild type and Ser118Ala mutant MccF enzymes has been described previously (Tikhonov et al., 2010). These plamids were generously provided to us by the Severinov laboratory (Rutgers, The State University of New Jersey). All mutants were constructed using standard site directed mutagenesis protocols. All plasmids were sequeced (ACGT Inc., Wheeling- IL) to verify the site of mutagensis.

The gene encoding the C-terminal acetyl transferase domain of MccE, encoded by amino acids 405 to 589 was amplified using the polymerase chain reaction with forward primer 5'-AAGCAGCCGCATATGCGGAAGTATGATGTATCTCTTACGCC bearing an engineered NdeI restriction site (underlined) and the reverse primer 5'-CTAGCTCGAGTTAACCAATTACTTTCGAATAAATATTTTGATC with an engineered XhoI restriction site (underlined). The PCR product was purified, and following restriction digest, was ligated into pET28a(+) plasmid (Novagen) to yield an amino-terminal hexa-histidine tag and a thrombin digestion site preceding the inserted gene. The plasmid was sequenced to confirm the integrity of the coding sequence.

MccF wild type and all MccF mutant enzymes were purified under identical conditions. All plasmids were transformed into competent *E. coli* Rosetta2(DE3) strains for recombinant protein expression. A single colony of the transformed cells was grown overnight in Luria-Bertani (LB) medium supplemented with 100 µg/mL ampicillin and 25 µg/mL chloramphenicol. This was subsequently used to innoculate 6 L of LB medium similarly supplemented with antibiotics, and grown at 37° C till the O.D. at 600 nm reached 0.6. Isopropyl β-D-

thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM and the culture was shifted to 18° C and grown for an additional 16 hours. All subsequent steps were performed at 4° C. The bacterial cells were harvested by centrifugation and resuspended in 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10% glycerol buffer and lysed by multiple passes through an Avestin C5 homogenizer. The cell lysate was clarified by centrifugation and the supernatant was loaded onto a 5 mL bed volume His-Trap column (GE Biosciences) equilibrated with lysis buffer. Subsequent to loading, the column was extensively washed with 20 mM Tris-HCl (pH 8.0), 1 M NaCl, 30 mM imidazole buffer and eluted by a continous linear gradient to 250 mM imidazole concentration in the wash buffer. Protein purity was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and greater than 95% pure fractions were applied to Superdex-75 16/60 size exclusion chromatography column (GE Biosciences). The protein was isochratically eluted in 20 mM HEPES-Na (pH 7.5), 100 mM KCl buffer and concentrated to 0.6 mM final concentration before being flash frozen in liquid nitrogen and storage at -80° C in small aliquouts. For the purification of the MccF Asn220Leu/Lys247Leu mutant enzyme, all buffers were supplemented with 10% glycerol.

Selenomethione incorporated MccF wild type enzyme was prepared by previously established protocols (Van Duyne et al., 1993). Culture was grown under identical conditions in M9 minimal medium supplemented with antibiotics and 5 g/L D-glucose. When the O.D. reached 0.4, 100 mg/L L-leucine, 100 mg/L L-lysine, 100 mg/L L-valine, 50 mg/L L-threonine, 50 mg/L L-phenylalanine and 50 mg/L L-isoleucine amino acids were added to the culture medium. After 20 minutes of further growth at 37° C, 50 mg/L L-selenomethione amino acid was added to the culture medium. When the O.D. reached 0.6, culture was induced with IPTG and grown for an

additional 16 hours at 18° C. The selenomethionine derivitized protein was purified identically to the wild type and mutant forms of MccF.

The plasmid bearing the $MccE^{CTD}$ was similarly transformed into competent Rosetta2(DE3) cells and recombinant protein expression carried out under identical conditions. The media was supplemented with 50 µg/mL kanamycin and 25 µg/mL chloramphenicol. The recombinant protein bearing the hexahistidine tag was isolated by Ni-NTA chromatography using the same protocol as that for the MccF enzyme. Greater than 95% pure fractions thus obtained were extensively dialyzed against 20 mM Tris-HCl (pH 8.0), 500 mM NaCl with the addition of 1 unit/mg of thrombin enzyme to proteolytically remove the hexahistidine tag. SDS-PAGE was used to verify the complete digestion of the hexahistidine tag. The protein was concentrated and applied to Superdex-75 16/60 size exclusion chromatography column. The protein was isochratically eluted in 20 mM HEPES-Na (pH 7.5), 100 mM KCl buffer and concentrated to 0.6 mM final concentration before being flash frozen in liquid nitrogen and storage at -80° C in small aliquouts. All mutants of $MccE^{CTD}$ were purified in an identical manner.

Protein crystallization, X-ray diffraction data collection and structure determination

Crystals of the apo MccF wild type enzyme and all other variants were grown using the hanging drop vapor diffusion technique. Commercially available sparse matrix crystallization screens were used to identify initial crystallization conditions. Three crystallization conditions could be reproduced- (i) 12% PEG 20000, 0.1 M MES-Na (pH 6.5) (ii) 10% PEG 8000, 8% ethylene glycol, 0.1 M HEPES-Na (pH 7.5), and (iii) 0.1 M Tris-HCl (pH 8.5) 8% PEG 8000.

Typically, 1 μL of the protein (9 mg/mL in 20 mM HEPES-Na (pH 7.5), 100 mM KCl) was added to 1 μL of the crystallization condition and equilibrated against the same solution at 9^{o} C. Selenomethionine-incorporated MccF was crystallized under similar conditions. Complex with processed McC7, aspartyl-sulfamoyl-adenosine (DSA), glutamoyl-sulfamoyl-adenosine (ESA), glutamoyl-sulfamoyl-inosine (ESI) and glutamoyl-sulfamoyl-guanosine (ESG) were generated by soaking apo crystals with 5 mM final concentration of substrate compounds added to the crystallization condition for 12 hours to 2 days. Crystals of the MccF-Trp186Phe in complex with AMP were generated by incubating the protein with 50 mM AMP and subsequent crystallization using the previous conditions. Crystals were vitrified by soaking in crystallization condition supplemented with 30% ethylene glycol prior to direct immersion in liquid nitrogen. All X-ray diffraction data was collected at the Life Sciences Collaborative Access Team (Sector-21), Argonne National Laboratory, Lemont- IL.

Initial crystallographic phases were determined by single wavelength anomalous diffraction utilizing anomalous scattering from the five selenium-substituted methionine residues per monomer of MccF. A four-fold redundant data set was collected from crystals of selenomethionine-incorporated MccF at the selenium absorption edge, to a limiting resolution of 1.45 Å (overall R_{merge} = 7.2, I/σ (I) = 13.7 in the highest resolution shell) utilizing a Mar 300 CCD detector. Data were indexed and scaled using the HKL2000 package (Otwinowski et al., 2003). Selenium sites were identified and refined for phase calculation using HySS (Grosse-Kunstleve and Adams, 2003), yielding an initial figure of merit of 0.439 (acentric/centric = 0.454/0.137) to 1.45 Å resolution. The resultant electron density map was of exceptional quality and permitted most of the main-chain and 50% of side-chain residues to be automatically built using the phenix-autobuild procedure (Terwilliger et al., 2008). The remainder of the model was

fitted using Coot (Emsley and Cowtan, 2004) and further improved by rounds of refinement with REFMAC5 (Murshudov et al., 1997) and manual building. Subsequent rounds of model building and crystallographic refinement utilized data from a native crystal of MccF grown under similar conditions that diffracted to 1.30 Å resolution (overall R_{merge} = 6.5, I/σ (I) = 6.7 in the highest resolution shell). Cross-validation, using 5% of the data for the calculation of the free R factor, was utilized throughout the model-building process in order to monitor building bias (Kleywegt and Brunger, 1996).

The crystal structures of MccF mutants Ser118Ala, Asn220Ala/Lys247Ala/Ser118Ala and Trp186Phe in complex with AMP were determined to resolutions of 1.5 Å, 1.7 Å and 1.5 Å, respectively, and MccF-Ser118Ala co-crystal structures with DSA, ESA, ESI, ESG and processed McC7 to resolutions of 1.5 Å, 1.3 Å, 1.2Å, 1.23 Å and 1.3 Å respectively, were determined by molecular replacement using the coordinates of wild type apo MccF as a search probe. Each of the structures was refined and validated using the procedures detailed above. For each of the structures, the stereochemistry of the model was monitored throughout the course of refinement using PROCHECK (Laskowski et al., 1996).

Initial crystallization conditions for $MccE^{CTD}$ were obtained by sparse-matrix sampling method using commercial screens. Crystals of $MccE^{CTD}$ were grown using the hanging vapor drop diffusion method. For crystallization of the apo protein, 1 µL of protein at 13.5 mg/mL concentration, incubated with 2 mM acetyl coenzyme A (acetyl CoA), or coenzyme A (CoA), for 2 hours on ice, was mixed with 1 µL of precipitant solution (50 mM magnesium chloride, 50 mM HEPES-Na (pH 7.0), 25% PEG550MME) and equilibrated over a well containing the precipitant solution at 9° C. Co-crystallization for substrate complexes involved further incubation of the protein with 5 mM of substrates for 1 hour on ice. Crystals grew within 3 days

and were briefly soaked in precipitant solution with PEG concentration increased to 30% prior to flash-cooling in liquid nitrogen. Ternary complex co-crystal structures with AMP, processed McC7, DSA, and ESA were grown under similar conditions and manipulated in the same manner. Co-crystals with AMP required incubation with 50 mM substrate prior to crystallization.

The structure of $MccE^{CTD}$-acetyl CoA complex was determined by molecular replacement (McCoy, 2007; McCoy et al., 2007; Qimron et al.) using the coordinates of a putative acetyltransferase (PDB entry 1NSL; 35% sequence identity), whose structure was determined by the Midwest Consortium for Structural Genomics (Brunzelle et al., 2004). Use of the full-length structure as a search probe failed to yield a successful MR solution. Noting the conservation in primary sequence was limited over the amino-terminal seventy residues, further efforts at MR utilized a model from which the amino terminal 70 residues were removed. One such model, in which all non-conserved residues were changed to serine, resulted in two MR solutions without any packing clashes that were consistent with the self-rotation function data, with minimal clashes. Following rigid body refinement of the initial MR solution, phases from the resultant model were further improved by two-fold density averaging and the use of Prime-and-Switch maps (Adams et al., 2002; Terwilliger et al., 2008), allowing for a significant portion of the model to be manually rebuilt using XtalView (McRee, 1999). Further manual fitting was interspersed with automated rebuilding using ARP/wARP (Perrakis et al., 1997) and rounds of refinement using REFMAC5 (Murshudov et al., 1997; Murshudov et al., 1999). Cross-validation, using 5% of the data for the calculation of the free R factor (Brunger, 1992), was utilized throughout model building process in order to monitor building bias. Although clear density for acetyl CoA could be observed prior to crystallographic refinement, the ligand was manually built into the model only after the free R factor dropped below 30%. The

stereochemistry of the models was routinely monitored throughout the course of refinement using PROCHECK (Laskowski et al., 1996).

The crystal structures of the MccE^{CTD}-acetyl CoA ternary complex with AMP, McC7, DSA and ESA were all determined by the molecular replacement method using the final refined coordinates of the MccE^{CTD}-acetyl CoA complex as a search probe. Multiple rounds of manual model building were interspersed with refinement using REFMAC5 to complete structure refinement. Cross-validation used 5% of the data in the calculation of the free R factor. Crystal parameters, data collection parameters and refinement statistics for each of the structures are summarized in Table 2.5.

End point assays to monitor hydrolysis of synthetic sulfamoyl nucleotide substrate analogs by wild type MccF enzyme

Purified MccF enzyme was diluted to a final concentration of 0.2 µM in 20 mM HEPES-Na (pH 7.5) buffer and incubated with 200 µM of substrates for 60 minutes in 75 µL assay volume at room temperature. The assays were quenched with equal volume of 4% formic acid and heating at 80° C for 10 minutes. The reactions were centrifuged and 100 µL volume was injected to a analytical C18 column (Vydac; 5 µm particle size, 4.6 × 250 mm) using an Agilent 1260 high performance liquid chromatography (HPLC) system. Chromatography was performed using 2% acetonitrile as Buffer A and 80% acetonitrile as Buffer B. The elution gradient profile was as follows: wash with 5 mL of buffer A, elute with continous linear gradient to buffer B across 10 mL volume, wash with 5 mL Buffer B, reduce acetonitrile concentration to 2% using a

linear continous gradient across 4 mL volume, requilibrate with 6 mL Buffer A. All reactions were performed in triplicate. Peak fractions were manually collected and flash frozen in liquid nitrogen. The fractions were then lyophilized to dryness and resuspended in water. Electrospary ionization mass spectroscopy (ESI-MS) data was collected using a Waters Quattro II mass spectrometer equipped with a electrospray ion source at the Mass Spectrometry Laboratory at the School of Chemical Sciences, University of Illinois at Urbana Champaign. Control reactions lacking the enzyme were performed in identical conditions and similarly analyzed.

Continuous coupled assay to determine the kinetic constants for the hydrolysis of ESA, ESI and ESG by wild type and mutant MccF enzymes

Kinetics for cleavage of ESA, ESI and ESG by MccF was monitored by coupling the glutamate production to reduced nicotinamide adenine dinucleotide (NADH) generation by bovine glutamate dehydrogenase (bGDH) (Sigma G2626-50MG), which was monitored by continuous increase in absorbance at 340 nm wavelength using a Cary4000 UV-VIS spectrophotometer. The assays were performed at room temperature in 20 mM HEPES-Na (pH7.5), 100 mM KCl buffer. NAD^+ and bGDH concentrations were kept fixed at 2 mM and 1.7 µM respectively in the assays. Linear correlation between the initial rate of reaction and MccF enzyme concentration, at 500 µM substrate concentrations confirmed that the coupling enzymatic reaction was not rate limiting. MccF wild type enzyme and Trp186Phe, Arg246Ala, Asn220Ala and Lys247Ala mutants were assayed at 150 nM enzyme concentration. Ser118Ala, Trp186Ala and Asn220Ala/Lys247Ala mutants were assayed at 750 nM enzyme concentration. Baseline absorbance was observed for one minute before the reactions were initiated by the

addition of enzyme. All data points were collected in triplicate and kinetic parameters determined by fitting the data to the Michaelis-Menten equation.

Discontinuous assay to determine the kinetic constants for the hydrolysis of DSA by wild type and mutant MccF enzymes

Kinetics for the cleavage of DSA by MccF was determined by discontinuous HPLC separation of the reactant- DSA and product- sulfamoyl adenosine (SA) on an analytical scale C18 column (Vydac, 5μm particle size, 4.6 × 250 mm), monitored at 254 nm wavelength. MccF wild type, Asn220Ala and Lys247Ala mutant enzymes were assayed at 50 nM enzyme concentration. 75 μL enzymatic reactions were quenched by the addition of equal volumes of 4% formic acid and heating at 80º C for two minutes. 100 μL of the quenched reaction was injected to the C18 column and chromatography performed using 2% acetonitrile as buffer A and 80% acetonitrile as buffer B on an Agilent 1260 HPLC system. The elution gradient profile was as described previously. Different amounts of DSA without the addition of the enzyme were diluted into the reaction buffer, quenched and heated as above were injected, and a standard curve of the moles of reactant versus the area under the elution peak was determined. As the extinction coefficient of DSA and SA can be assumed to be the same at 254 nm wavelength, this standard curve was used to determine the number of moles of product formed in the subsequent enzymatic reaction time points. Measurements of the enzymatic rates were within the linear range of substrate turnover with respect to time, which was confirmed by the linear correlation between the measured initial rates and enzyme concentration across four different enzyme concentrations for each substrate concentration used. All data points were collected in triplicate and kinetic

parameters determined by fitting the data to the Michaelis-Menten equation. Steady state kinetic parameters for DSA hydrolysis by the Asn220Ala and Lys247Ala mutant enzymes were determined in an identical manner.

Discontinuous assay to monitor the hydrolysis of FSA by MccF mutant enzyme

Hydrolysis of phenylalanine-sulfamoyl-adenosine (FSA) by MccF-Asn220Leu/Lys247Leu was monitored in a manner identical to the kinetics determination of DSA hydrolysis by wild type MccF. 135 μM FSA was incubated with 450 nM of MccF-Asn220Leu/Lys247Leu enzyme at room temperature in 20 mM HEPES-Na (pH 7.5) buffer. Enzyme dilutions were made in 20 mM HEPES-Na (pH 7.5) 100 mM KCl buffer. Reactions were initiated by the addition of enzyme. Aliquots from the reaction mixture were taken at different time points, quenched and substrate and products were separated by HPLC. Substrate and product peak elution fractions were collected, lyophilized and dissolved in water and analyzed by ESI-MS to confirm the identity of the substrate and product chemical species. Calibration curve between number of moles of substrate and area under the substrate peak was generated and used to calculate the amount of substrate consumed at different time points.

Isothermal titration calorimetry to determine disassociation constant for binding of AMP to MccE^{CTD}

Measurements were made at 25° C using a VP-ITC calorimeter (MicroCal, Inc, Northampton, MA). Protein after size exclusion chromatography was extensively dialyzed

against 50 mM potassium phosphate buffer (pH 7.0) and concentrated to a final concentration of 0.2 mM, followed by incubation with a final concentration of 0.5 mM acetyl CoA for 1 hour on ice. A stock of (AMP) at a concentration of 50 mM was prepared in the spent dialysis buffer. The protein sample in a 1.4495 mL reaction cell was injected with 25 successive 10 µL aliquots of (AMP) at a spacing of 300 seconds, with spent dialysis buffer in the reference cell. Based on the crystallographic results, the number of binding sites for (AMP) was presumed to be one. Data was fitted by nonlinear regression using a single site binding model (MicroCal Origin), and thermodynamic parameters were calculated using the Gibbs free energy equation ($\Delta G = \Delta H - T\Delta S$), and the relation $-RT\ln K_a = \Delta G$.

Continuous assay to determine MccE^{CTD} kinetic parameters

Reaction rates were measured spectrophotometrically by determining the rate of increase in absorbance at 412 nm due to the reaction between the free sulfhydryl group of CoA, generated by the enzymatic transfer of the acetyl group from acetyl CoA to the substrate molecule, and 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB). The assays were performed at room temperature in a final volume of 100 µL and monitored continuously using a CARY UV-VIS spectrophotometer. Each reaction consisted of 2.5 mM final concentration of DTNB and 1.5 mM acetyl CoA. Five different substrate concentrations were used and the assays were initiated by the addition of the enzyme to a final concentration of 100 pM and the increase in absorbance was monitored for 5 minutes. Rates for each substrate concentration were measured in triplicate.

Results and discussion

MccF overall three dimensional structure and description of the catalytic triad

MccF wild type structure in the apo state was determined by single wavelength anomalous dispersion to a resolution of 1.2 Å and refined to R/R_{free} of 0.184/0.201. The relevant data collection and refinement statistics are presented in Table 2.1. Barring the hexa-histidine tag, six residues at the N-terminus and three residues at the C-terminus, all amino acids have been described in the final model. The model also incorporates several molecules of ethylene glycol which bear their origin to the crystallization condition and cryo-protectant. The crystal asymmetric unit consists of two monomers, and represents a single biological homodimer unit as determined by analytical size exclusion chromatography. A total surface area of 3681 $Å^2$ is occluded between the two monomer chains. Multiple hydrogen bonds exist at the dimer interface between basic arginine side chains and acidic carboxylate and hydroxyl side chains. One of these interactions is the hydrogen bond between the side chain of Arg254 of one monomer and Thr223 of the other monomer. Mutation of Arg254 to alanine resulted in insoluble protein, which is suggestive of disruption of dimer packing due to loss of these hydrogen bonding interactions.

Each MccF monomer can be divided into three distinct domains (Figure 2.4A). The amino terminal domain comprises of residues 7-164 and is made up of four parallel β sheets flanked by α helices on either side of β sheet. It is followed by a rather unusually long single loop made up of residues 165-208 which is henceforth referred to as the 'catalytic loop'. The carboxy terminal domain is made of residues 209-341 and has a six membered mixed β sheet flanked by three α helices on one side of the sheet. The catalytic loops are positioned in between the amino and carboxy terminal domains. Dimerization is mediated by helices and loops from

both the amino and carboxy terminal domains and the catalytic loops are positioned opposite to the dimerization surface. Two α helices from the carboxy terminal domains contact each other at the dimer interface and essentially provide all the substrate recognition residues for coordination of the substrate acidic side chain (*vide infra*) (Figure 2.4B). The MccF dimer forms a solvent exposed concave cavity which houses both the monomer active sites which point towards each other.

Numerous contacts exist between residues of the catalytic loop and the side chain and main chain atoms of the amino and carboxy termini domains. These interactions serve to position the catalytic loop correctly. Strong ion pair exists between the side chains of Glu182 and Arg193, which lends conformational rigidity to this extended loop region. The carboxy and amino terminal domain arrangements, and the secondary structure elements are conserved between MccF and the metal independent LD-carboxypeptidase (PDB ID: 1ZRS) (Korza and Bochtler, 2005), with the notable absence of the catalytic loop in the LD-carboxypeptidase crystal structure. Sequence alignment of MccF with LD-carboxypeptidases identifies greatest sequence divergence in the region encoding for the catalytic loop, and thus can be used as a differentiation marker to identify LD-carboxypeptidases mis-annotated as MccF like proteases in sequence databases (Figure 2.5). Analysis of MccF homologs' catalytic loop sequences reveals that hydrophobic residues corresponding to Trp186 are conserved, suggesting that these proteins should also interact with adenylated targets. Curiously, many MccF homologs are encoded by genes that are not part of recognizable McC7-like biosynthetic operons. A recent crystal structure report of a MccF homolog from *Bacillus anthracis*, which is does not produce McC7 or any other known small molecule natural product of similar structure, has reported a nearly identical

arrangement of the catalytic loop tryptophan residue (Nocek et al., 2012). The intracellular function of these proteins remains to be determined.

Serine proteases have been shown to catalyze peptide and ester bond hydrolysis by an extensively studied 'charge relay' mechanism which comprises of an invariant serine nucleophile, a general acid-base histidine and a carboxylate side chain from an aspartate or glutamate residue with the histidine side chain in hydrogen bonding distance to both serine and carboxylate side chains (Hedstrom, 2002). Crystal structure of MccF identifies the catalytic triad residues to be Ser118, His311 and Glu243, and is thus consistent with the earlier report of Ser118Ala mutant to be catalytically inactive (Tikhonov et al., 2010). The Glu243 side chain is buried in a polar environment with one of the oxygen atoms of the carboxylate side chain being hydrogen bonded to the Nδ1 of His311 side chain and also to a well ordered water molecule. The second oxygen atom is hydrogen bonded to Tyr137 side chain hydroxyl and side chain amide of Asn218. Nε2 atom of His311 side chain is appropriately placed at a distance of 2.8 Å from the Ser118 side chain hydroxyl to act as the general acid-base catalyst. Rather unexpectedly, alternate conformations can be assigned to the Ser118 side chain in one of the monomer active sites. Coordination with His311 is possible in only one of these conformations and the serine hydroxyl is hydrogen bonded to two well ordered water molecules in the other conformation. Two ethylene glycol molecules are bound on the other face of the catalytic triad in the other monomer.

We next determined the 1.5 Å resolution crystal structure of the MccF-Ser118Ala mutant in the apo form to refinement statistics of R/R_{free} 0.163/0.190 (Table 2.1). The alignment of the wild type structure to the Ser118Ala mutant structure unequivocally establishes that the loss in catalytic activity of the Ser118Ala mutant enzyme is due to the loss of the catalytic serine side

chain hydroxyl nucleophile, as all catalytic site and neighboring residue side chains align very well with each other and no major overall structural changes are apparent. Hence we hypothesize that the co-crystal complexes of the MccF-Ser118Ala enzyme with different substrate molecules are descriptive of the productive state of substrate binding by MccF.

Crystal structure of MccF-Ser118Ala enzyme in complex with the physiological substrate-processed McC7

In order to delineate the interactions between processed McC7 and MccF enzyme, crystals of the catalytically inactive MccF-Ser118Ala mutant were grown and soaked with processed McC7. The difference electron density maps displayed unambiguous density for processed McC7 in both subunits of the enzyme, and were used to model in the processed McC7 molecule to obtain a model at 1.3 Å resolution which has been refined to R/R_{free} of 0.174/0.186 (Table 2.1) (Figure 2.6). Analogous to extensive literature on serine proteases, the interactions of the processed McC7 molecule are herein described with reference to the nucleotide binding site as the P1 site, and the aspartate side chain of processed McC7 as the P1' interaction site.

The interactions at the P1 site are primarily modulated by the π-stacking interactions of the adenine ring of processed McC7 and coordination of the 2' and 3' ribose hydroxyls (Figure 2.6). The adenine ring stacks against the indole side chain of the residue Trp186. The Trp186 residue is harbored on the 'catalytic loop' (Figure 2.4), and is thus absent from the LD-carboxypeptidase active site (Korza and Bochtler, 2005) (Figure 2.5). The absence of a co-crystal structure of the LD-carboxypeptidase LcdA with its physiological substrate precludes the comparison of the substrate binding sites of the two enzymes, though the absence of the catalytic

loop provides a possible explanation for the reported inability of LcdA to detoxify McC7 and its analogs (Tikhonov et al., 2010).

The 2’ and 3’ hydroxyl groups of the ribose sugar are respectively positioned 2.9 Å and 2.7 Å away from the side chain carboxylate oxygen atoms of Glu277 and engage in hydrogen bonding interactions. The 2’ hydroxyl is also positioned 2.7 Å away from guanidine side chain of Arg246, which in turn is also hydrogen bonded to Ser183 side chain hydroxyl. Interaction between carboxy terminus domain residue Arg246 and catalytic loop residue Ser183 could play a role in maintaining the proper conformation of the catalytic loop, and thus the positioning of the Trp186 side chain for stacking with the adenine group of McC7. This is also exemplified by an extensive set of hydrogen bonding and salt bridge interactions formed by the polar residues in the primary sequence vicinity of Trp186, which lend rigidity to the conformation of the catalytic loop.

Interaction with the enzyme backbone atoms stabilize the oxyanion reaction intermediate, with the carbonyl oxygen of the phosphoramidate linkage positioned 3.05 Å and 2.8 Å away from the backbone amide nitrogen atoms of Asp119 and Gly92 respectively. One of the phosphate oxygen atom is positioned 2.95 Å away from the backbone amide nitrogen of Gly91. Unexpectedly, the propylamine appendage to the phosphate is fully extended and pointing outwards from the active site. The terminal primary amine does not find any interaction partners, and neither is the propyl chain buttressed by hydrophobic side chain residues. Similar lack of interactions for the propylamine appendage were observed in the crystal structure of $MccE^{CTD}$ with processed McC7 (PDB ID: 3R9G) (Agarwal et al., 2011). However the conformation of processed McC7 is very different in the two structures. While the aspartate side chain and

propylamine appendage are essentially parallel for MccE^{CTD}, they are perpendicular to each other for MccF (*vide infra*).

The aspartate side chain of processed McC7 is coordinated at the P1' position of the MccF active site. One of the carboxylate oxygen atoms is bound to the side chain of Lys247 and is positioned 3.1 Å away from it (Figure 2.6). The other carboxylate oxygen is 2.9 Å away from the side chain amide nitrogen of Asn220. This oxygen atom is also tightly coordinated by the guanidine side chain of an alternate conformation of Arg254 residue from the other MccF monomer (shown in grey in Figure 2.6). The distance between the oxygen atom and the guanidine nitrogen atoms is 3.05 Å and 2.9 Å. In the other conformation of Arg254 side chain, it mediates dimerization between the MccF monomers by coordinating to the Thr223 side chain hydroxyl as observed in the MccF wild type apo enzyme structure. The guanidine nitrogen atoms of Arg254 are also 3.2 Å and 3.25 Å away from the side chains of Asn220 and Thr221 respectively, of the other MccF monomer. These alternate conformations of Arg254 side chain and synergistic substrate binding interactions described here are observed for both MccF monomers. Side chains of Thr221 and Asn218 also point inwards towards the highly polar tightly coordinated substrate binding site. It should be noted that Asn218 also modulates the correct positioning of the catalytic triad residue Glu243.

MccF can also hydrolyze and inactivate unprocessed McC7, in which the hexapeptide Trojan horse has not been cleaved off (Tikhonov et al., 2010). Our numerous attempts at crystal soaking as well as co-crystallization with intact McC7 failed to yield the substrate in the active site of the enzyme. However, the co-crystal structure of MccF-Ser118Ala in complex with processed McC7 offers rationale for acceptance of McC7 as a substrate by the enzyme. The amino terminal backbone nitrogen atom of processed McC7 is positioned without any specific

contacts with the enzyme, and points outwards from the active site. Peptide elongation upon that nitrogen atom would most probably lead to the Trojan horse carrier hexapeptide being solvent exposed without specific interactions with the MccF enzyme. Also the McC7 amino acid immediately upstream of the aspartate residue is an alanine, which could easily be accommodated in the solvent exposed substrate binding site of MccF without creating steric or conformational clashes with the enzyme.

Structure and kinetics for hydrolysis of synthetic sulfamoyl adenylates of aspartate and glutamate by MccF

Numerous in-vitro assays were unsuccessfully trialed to measure the kinetics for the hydrolysis of processed McC7 by MccF. The hydrolysis products of processed McC7 could not be detected by coupling to an aspartate dehydrogenase enzyme (Yang et al., 2003), nor could be retained reproducibly on HPLC reverse phase or ion exchange columns. Attempts at derivatizing the primary amines, or the free carboxylates of the processed McC7 hydrolysis products for detection by gas chromatography mass spectroscopy (GC-MS) were also unsuccessful. Hence, we shifted our attention to synthetic sulfamoyl adenylates of aspartate (DSA) (**3** in Figure 2.7) and glutamate (ESA) (**4**) to monitor the kinetics of hydrolysis by MccF wild type and mutant enzymes. These synthetic substrates were previously reported to be accepted as substrates by the MccF enzyme, as well as inhibit bacterial growth with an identical mode of action as that for McC7 (Novikova et al., 2010; Tikhonov et al., 2010; Van de Vijver et al., 2009a; Van de Vijver et al., 2009b). Synthetic aminoacyl sulfamoyl adenylates have been demonstrated to inhibit various other cellular primary and secondary metabolic processes which involve adenylated enzyme intermediates and have thus gained attention as antibiotic candidates (Finking et al.,

2003; Gutierrez-Lugo and Bewley, 2011; May et al., 2005; Qiao et al., 2007; Tremblay et al., 2008).

Hydrolysis of both ESA and DSA was found to be tractable by kinetic assays and HPLC analysis (Figure 2.7B). ESI-MS on the HPLC purified reactants and products for the MccF catalyzed hydrolysis of DSA and ESA substrates confirmed the site of hydrolysis (Figure 2.8). The mode of binding for DSA and ESA by the enzyme was determined by crystal structure of MccF-Ser118Ala mutant in presence of both substrates (Figure 2.9) (Table 2.2).

Inability of the MccF-Ser118Ala mutant enzyme to hydrolyze DSA and ESA was confirmed spectrophotometrically and by HPLC analysis. In order to delineate the interactions between ESA and MccF enzyme, crystals of the catalytically inactive Ser118Ala mutant were obtained and soaked with ESA. The difference electron density maps displayed unambiguous density for ESA in both subunits of the enzyme, and were used to model in the ESA molecule to obtain a model at 1.3 Å resolution, which has been refined to R/R_{free} of 0.179/0.197 (Figure 2.9).

As observed in the crystal structure of MccF-Ser118Ala with processed McC7, the interactions at the P1 site for ESA are modulated by the π-stacking interactions of the adenine ring of ESA by Trp186 side chain. Mutation of Trp186 to phenylalanine resulted in a fourfold increase in the K_M for the enzyme (Table 2.3). Crystal structure of the Trp186Phe/Ser118Ala mutant enzyme in the apo state shows that the phenylalanine side chain is essentially parallel to the indole ring for the wild type enzyme and thus can make stacking interaction with the substrate, which is further confirmed by the 1.5 Å resolution cocrystal structure of the MccF-Trp186Phe mutant enzyme in complex with adenosine monophosphate (Figure 2.10). A Trp186Ala mutation resulted in complete loss of observable activity for the enzyme. The π-

stacking of the adenine ring is consistent with the mode of binding of McC7, and its analogs with the McC7 detoxifying enzyme $MccE^{CTD}$, in which the adenine ring stacks between tryptophan and a phenylalanine side chains, mutations to alanine of either one of which lead to complete loss of activity of $MccE^{CTD}$ (Agarwal et al., 2011). The 2' and 3' ribose hydroxyls are coordinated by side chains of Glu277 and Arg246. Mutation of Arg246 to alanine did not alter the k_{cat} of the enzyme, but instead lead to a fivefold increase in the K_M for the substrate.

In contrast to identical interactions at the P1 site for ESA and processed McC7, variability was found at the P1' site for ESA recognition. The glutamate side chain of ESA contacts both Asn220 and Lys247 side chains, but the ESA carboxylate is skewed towards the Lys247 primary amine. In contrast to the aspartate carboxylate of McC7 in which only one of the oxygen atoms contacted Lys247 side chain, both the ESA oxygen atoms are within hydrogen bonding distance of Lys247 now. Though one of the ESA oxygen atoms still contacts the Asn220 side chain amide, this coordination distance is now increased to 3.1 Å. This oxygen atom has also lost contacts with Asn218 and Thr221 side chains. This greater reliance on the side chain of Lys247 is evident from kinetic analysis as well. Mutation of Lys247 to alanine leads to over 30 fold reduction in the catalytic efficiency of the enzyme. Loss in the interaction with Asn220, by mutagenesis of Asn220 to an alanine, can be compensated by interactions of both oxygen atoms to Lys247, which is evident from only 2 fold increase in K_M and no change in k_{cat} for ESA hydrolysis. However mutation of both Asn220 and Lys247 residues to alanine lead to complete loss in detectable catalytic activity for the enzyme. Crystal structure of the MccF-Asn220Ala/Lys247Ala/Ser118Ala mutant enzyme in the apo state (determined to a resolution of 1.7 Å and R/R_{free} 0.173/0.201) shows no rearrangement of the P1 site residues and hence loss in

activity can be attributed solely to the loss in coordination for the glutamate side chain of ESA at the P1' site (Figure 2.11).

The mode of stabilization of the oxyanion transition state formed during hydrolysis of ESA was found identical to that for processed McC7 described previously. The carbonyl oxygen atom of the amide bond is positioned 3.2 Å and 2.8 Å away from the backbone amide nitrogen atoms of Asp119 and Gly92 residues respectively. The other sulfate oxygen atom, which would carry the propylamine arm by analogy to the processed McC7 cocrystal structure, finds no interaction partners with the enzyme and is pointing outward from the active site.

In order to confirm that the nature of aspartate chain coordination is preserved between DSA and processed McC7 we determined the co-crystal structure of MccF-Ser118Ala enzyme in complex with DSA to a resolution of 1.5 Å and to refinement statistics of R/R_{free} of 0.181/0.202 (Figure 2.9A). Briefly, the interactions at the P1 site, of the sulfamoyl oxygen atoms and the oxyanion hole are essentially identical to ESA and processed McC7 co-crystal structures. The DSA aspartate chain is also similarly coordinated as for processed McC7. The two carboxylate oxygen atoms are individually coordinated by Asn220 and Lys247. The oxygen atom coordinated by Asn220 (3.1 Å) is also in hydrogen bonding distance to Asn218 (3.3 Å) and Thr221 (3.1 Å) side chains. Thus the binding of DSA, which is distinct from that of ESA, mimics the binding of processed McC7 in the MccF active site.

As estimated by the ratio of k_{cat}/K_M, cleavage of DSA appears to be about 8 fold more efficient than the hydrolysis of ESA by the enzyme (Table 2.3). This substrate preference is reflected by an order of magnitude higher turnover number for DSA, while the K_M for the two substrates is comparable. This reflection of substrate preference by change of k_{cat} and not K_M is

in concert with the earlier work with serine proteases such as chymotrypsin, elastase and pepsin (Baumann et al., 1970, 1973; Sachdev and Fruton, 1975; Thompson and Blout, 1973), in which synthetic substrates designed with altered binding interactions at the P1' and subsequent downstream binding subsites have lead to non-significant changes in the K_M for these enzymes, while displaying large variations in the k_{cat}. For MccF, the order of magnitude increase in k_{cat} for DSA as compared to ESA can be attributed to the differential mode of stabilization of the acylenzyme transition state (formed after the departure of the adenine sulfonamide) in the active site. As the aspartate chain makes more contacts with the enzyme than the glutamate side chain, the activation energy for the formation of this transition state would be lower, leading to an increase in the rate of hydrolysis of DSA. It should also be noted that the mutations at the P1 site- Trp186Phe and Arg246Ala, had much larger alterations in the K_M values for ESA hydrolysis, and much less significant changes in k_{cat} values.

MccF active site can accommodate other purine nucleobases at the P1 site

In order to explore whether the MccF active site can accept nucleobases other than adenine at the P1 position, we investigated the hydrolysis of ESI (**5** in Figure 2.12A) and ESG (**6**) by wild type MccF enzyme. ESI and ESG are identical to ESA except the identity of the nucleobases. Glutamyl nucleotide conjugates were chosen in place of aspartyl nucleotide conjugates for this study, as their hydrolysis could be monitored by a continuous coupled spectrophotometric assay, as opposed to a discontinuous HPLC based assay for DSA described previously. ESI and ESG were generously synthesized and provided by Gaston Vondenhoff and

Dr. Arthur Van Aerschot (Laboratory of Medicinal Chemistry, Rega Institute for Medical Research, Katholieke Universiteit Leuven, Belgium).

Wild type MccF could indeed catalyze the efficient cleavage of both ESI and ESG, as judged by separation of the reactants and products by HPLC (Figure 2.12B). ESI-MS confirmed that the site of hydrolysis remained as the amide connecting the amino acid and the sulfamoyl-amine (Figure 2.13). We next determined the crystal structure of MccF-Ser118Ala enzyme in complex with ESI and ESG to limiting resolutions of 1.2 Å and 1.23 Å respectively, with refinement statistics of R/R_{free} of 0.159/0.175 and 0.155/0.172 respectively.

As anticipated, MccF active site accommodates the glutamate side chain of ESI and ESG in an identical manner as that for ESA in the P1' site. At the P1 site, the nucleotide ribose 2' and 3' hydroxyls are also identically coordinated by Glu277. However slight variations are found in the π-stacking interactions with the indole side chain of Trp186.

While π-stacking interactions can provide extensive surface area for van der Waals contact, such interactions result in electrostatic repulsion as the delocalized π-electrons result in a concentration of negative electrostatic potential. However, the electrostatic potential of heterocyclic adenine and guanine nucleobases of DNA/RNA differ markedly from that of homocyclic conjugated systems such as benzene. Calculations of the electrostatic potential map of adenine illustrates that the negative charge is concentrated on N1, N3, and N7, while C8 and N9 have significant positive potential (Boehr et al., 2002). Similar studies on the electrostatic potential of tryptophan suggest a relatively greater negative potential on the six-member ring of the indole side chain, than on the five-member heterocycles (Dougherty, 1996). In the MccF-Ser118Ala co-crystal structure with ESA, the side chain of Trp186 and the adenine ring of the

substrate are oriented to maximize favorable electrostatic interactions between the two heterocyclic π-electron systems (Figure 2.14A). The N1 and N3 of the adenine, which have the greatest negative potential, are positioned furthest away from the six-membered ring of the indole side chain (which would bear the greatest negative potential). In addition, the positive potential bearing C8 and N9 positions of the adenine are stacked nearest to the indole six-membered ring. The electronegative N3 is also positioned near the basic guanidine side chain of Arg246. The two ring systems are aligned so as to minimize the electrostatic repulsion between the two heterocyclic π-electron systems.

This juxtapositioning of the inosine nucleobase is nearly identical to that for adenine. However the K_M for the hydrolysis of ESI (91 μM) is higher than that for ESA. This increase in K_M for ESI can be attributed to a decrease in van der Waals contact with Trp186 side chain, as the inosine is not aromatic across the entirety of the fused heterocycles due to the presence of a double bonded carbonyl oxygen at the C6 atom of the inosine ring. This oxygen atom is also markedly electron rich, as revealed by electrostatic potential map of inosine (Schramm, 2007) which can cause electrostatic repulsion with the indole electron rich six member ring (Figure 2.14B). The k_{cat} for the hydrolysis of ESI, which has previously been shown to be dependent on the interactions in the P1' site of the enzyme is nearly identical to that for ESA (Table 2.4).

The guanine nucleobase is also not aromatic across the entirety of the fused heterocycles due to the presence of similar double bonded carbonyl oxygen at the C6 atom of the guanine ring. Hence the finding that the K_M for the hydrolysis of ESG (33 μM) is lower than that for ESA was unexpected. However the cocrystal structure of MccF with ESG provides a rationale for this kinetic observation. The primary amine at the C2 carbon atom of the guanine ring is hydrogen bonded to the side chain hydroxyl of Ser183. This hydrogen bonding also causes a slight rotation

of the guanine ring, as compared to the adenine ring of ESA, so that it makes more extensive van der Waals contact with the six member ring of the Trp186 indole (Figure 2.14C). Consistent with our previous hypothesis, the k_{cat} for ESG hydrolysis was found to be nearly the same as that for ESA hydrolysis (Table 2.4).

These results demonstrate that the MccF active site can indeed accept different purine nucleobases at the P1 site and accommodate different length of carboxylate side chains at the P1' site, albeit with slight structural modifications and kinetic proficiency alterations.

Modulation of substrate specificity of MccF at the P1' position

As demonstrated earlier (Tikhonov et al., 2010), and structurally and kinetically characterized in the previous sections, MccF is specific for the presence of acidic side chain coordination to the adenine nucleotide. The substrate recognition elements for the adenine nucleotide and the acidic side chain are separate and not coupled to distant structural features by enzyme conformational changes. Also, independent changes at the P1 or the P1' positions of the enzyme lead to compromise of catalytic activity. Consistent with the previous engineering efforts of serine proteases (Leis and Cameron, 1994; Perona and Craik, 1995), we hypothesized that by changing amino acid side chains at the P1' position, we could engineer the enzyme to accept sulfamoyl adenylates of hydrophobic amino acids as substrates. Contacts to side chain carboxylate of McC7, DSA and ESA are provided by the side chains of Asn218, Asn220, Thr221, Lys247 and Arg254. Asn218 was not mutated as it provides contacts to position the catalytic triad residue Glu243. Mutation to alanine of Arg254 had lead to insoluble protein. Thr221 side chain hydroxyl was not in coordination for the ESA side chain, and is also 3.6 Å

away from the backbone amide nitrogen of Asp119. Hence by a process of elimination, residues chosen for mutagenesis were Asn220 and Lys247. We created four double mutant combinations in which both these residues were mutated to either leucine or phenylalanine and tested their activity against the hydrolysis of FSA. All mutants had compromised stability and solubility. The proteins were particularly poorly behaved for the three mutants who had either of the side chains mutated to phenylalanine, with the Asn220Phe/Lys247Phe being completely insoluble after cell lysis. The Asn220Leu/Lys247Leu mutant enzyme could be purified and gel filtration chromatography confirmed that it existed as a dimer in solution, and hence used for activity analysis against FSA. Time resolved HPLC separation of substrates and products showed that Asn220Leu/Lys247Leu enzyme could indeed catalyze FSA cleavage, albeit the reaction was much slower than that of wild type enzyme with physiological substrates like DSA and ESA (Figure 2.15A,B). ESI-MS analysis of the reactant and product peaks confirmed that the site of cleavage was at the sulfamoyl amide bond (Figure 2.15C,D). The rate of cleavage of FSA by Asn220Ala/Lys247Ala enzyme is negligible under similar experimental conditions, which demonstrates that the acceptance of hydrophobic sulfamoyl adenosine as a substrate is due to the introduction of hydrophobic side chains in the P1' recognition site. However the poor stability of the mutant enzyme and slow catalytic rate represents an enzymatic system in which all requirements for substrate recognition and catalysis have not been currently satisfied.

Altered reaction mechanism for the hydrolysis of McC7 by MccF

The co-crystal structure of MccF in the presence of its physiological substrate-processed McC7 also poses a mechanistic challenge regarding the protonation of the α-amine of the

departing carboxy terminal peptide after the cleavage of the scissile peptide bond. The pK_a of the catalytic triad histidine side chain has been determined to be close to 7.0 (Bachovchin et al., 1981) and can thus protonate the departing α-NH_2 of the peptide, which would have a pK_a of greater than 9.0. However, in the case a phosphoramide leaving group, such as the adenosine phosphoramide of McC7, the pK_a of the α-NH_2 would be close to 4.5 (Gamcsik et al., 1993), and thus not possible to be protonated by the catalytic histidine side chain (His311 for MccF). However the lone pair of electrons on the α-amine of adenosine phosphoramine is in conjugation with the phosphate oxygen atom, which in turn is within hydrogen binding distance with the backbone amide nitrogen of Gly91 (Figure 2.16). Hence we propose that the α-amine, in this case, overcomes the requirement of protonation by the catalytic histidine side chain residue, as the developing negative charge on the α-amine nitrogen can then be delocalized by resonance conjugation to the significantly more electronegative phosphate oxygen atom, which is also in hydrogen bonding distance to a backbone amide nitrogen atom (Figure 2.16). This interaction constitutes the 'second oxyanion hole' for the stabilization of an altered transition state in the catalytic cycle of MccF.

Overall three dimensional structure of MccE^{CTD}

MccE is a bi-domain protein with the C-terminal domain (CTD) mediating acetylation of the backbone primary amine of the aminoacyl nucleotide conjugates, independent of the N-terminal domain. As the full length MccE protein could not be crystallized despite our exhaustive crystallization screening efforts, we focused our attention to the MccE^{CTD}. Limited proteolysis was used to determine the domain boundaries for the CTD, but this approach was

unsuccessful due to recalcitrance of MccE to proteolytic digestion. Hence we cloned numerous construct of the MccE^{CTD} based on sequence similarity with other known GCN5 related N-acetyltransferase (GNAT) enzymes. Primary among these were the *E. coli* RimI and RimJ enzymes, which have been structurally and kinetically characterized for the acetylation of ribosomal proteins (Vetting et al., 2008; Vetting et al., 2005a). Various MccE^{CTD} constructs were purified and screened for crystallization. Only the construct encompassing the residues Asp409 to Ile589 could be crystallized reproducibly and diffraction data collected. Crystals could only be obtained in the presence of the physiological cofactor- acetyl CoA. No crystals could be obtained in the presence of either CoA, or other CoA molecules such as succinyl CoA and acetoacetyl CoA.

The three-dimensional structure of MccE^{CTD} was solved by the molecular replacement method using the coordinates of the putative ribosomal *N*-acetyltransferase YdaF from *B. subtilis* (PDB entry 1NSL; 30% sequence identity) (Brunzelle et al., 2004). A successful molecular replacement solution could only be identified using a truncated search model from which the amino terminal 70 residues were removed. Extensive manual and automated rebuilding led to a complete model of the protein chain, encompassing residues Asp409 through Ile589. The addition of solvent molecules and acetyl CoA at positions with suitably well-defined electron density features resulted in the final model with a free R factor of 0.227 and good Ramachandran statistics. All data collection and refinement data are listed in Table 2.5.

The overall structure of MccE^{CTD} consists of a GNAT superfamily fold consisting of seven antiparallel β strands flanked by four α helices (Dyda et al., 2000; Vetting et al., 2005b) (Figure 2.17A). The structure is similar to that of uncharacterized acetyltransferase YdaF from *B. subtilis* (PDB code: 1NSL; RMSD of 1.6 Å over 176 aligned Cα atoms) and *Salmonella*

typhimurium RimL (PDB code 1S7L; RMSD of 1.2 Å over 174 aligned Cα atoms) (Brunzelle et al., 2004; Vetting et al., 2005c). The structure of the polypeptide can be roughly divided into two domains: an amino-terminal domain composed of two large α-helices separated by a short 3_{10} helix and a carboxy-terminal α/β fold that corresponds to the acetyl CoA binding site. The most significant deviations from other GNAT family members are along the amino-terminal 70 residues that define the ligand-binding site of $MccE^{CTD}$. Unlike *Salmonella*, RimL that is shown to be a dimer both in solution and in the crystal (Vetting et al., 2005c), and *B. subtilis* YdaF that shares the same dimerization interface in the crystal (Brunzelle et al., 2004), $MccE^{CTD}$ is a monomer in the crystal as well as in solution, as determined by the elution profile on size exclusion chromatography with calibrated standards (Figure 2.18).

Within the $MccE^{CTD}$ co-crystal structure, clear electron density corresponding to acetyl CoA can be observed at a position similar to that observed in other structures of GNAT family members (Figure 2.17B). The adenosine group is located at the surface of the molecule and engages in crystal packing contacts with the adenine ring of a symmetry-related molecule of $MccE^{CTD}$. The pantetheine and pyrophosphate moieties engage in hydrogen bonding and van der Walls interactions with the polypeptide, in a manner similar to that observed in the co-crystal structure of RimL (Vetting et al., 2005c). Electron density is evident for the entire ligand, and the acetyl group is situated at the base of the active site pocket, where the carbonyl group hydrogen bonds with the backbone amide of Tyr510 (O-N distance of 3.1 Å) and is flanked by residues Ser553 and Glu572. These residues are implicated to act as general acid/base catalysts for the acetyltransferase reaction and substitution of either of these residues for alanine was shown to compromise the catalytic activity and biological function of MccE (Novikova et al., 2010).

Interactions with the physiological substrate – processed McC7

Processed McC7 was generously provided by the Severinov laboratory (Rutgers, The State University of New Jersey) and used to obtain a 1.35 Å resolution co-crystal structure with $MccE^{CTD}$-acetyl CoA. In the structure, clear electron density can be observed for the entire substrate molecule, including the aspartyl side chain and the propylamine decoration on the phosphoramidate oxygen. Unexpectedly, the structure also shows continuous electron density corresponding to the attachment of an acetyl group on the α-amino nitrogen of processed McC7 (Figure 2.19). Density for the substrate molecule is equally clear for both molecules in the crystallographic asymmetric unit, confirming that the structure presents a view of the acetylated processed McC7 product complex.

The binding pocket for the substrate is composed of a number of hydrophobic residues. The substrate molecule is oriented perpendicular to the central β strands of the GNAT fold, resulting in the deposition of the tail of the CoA donor near the α-amino group of the peptide portion of the acceptor substrate. The CoA donor and the substrate are located on opposite sides of polypeptide and acetyl transfer occurs through a cavity in the center of $MccE^{CTD}$.

In the co-crystal structure, the polypeptide makes minimal contacts with the amino acid aspartyl side chain of the substrate. Instead, the most significant determinant for substrate specificity is the adenine ring of the substrate. The ring is sandwiched between two aromatic residues, Trp453 and Phe466, which are located within the two large helices in the amino terminus of $MccE^{CTD}$, resulting in a strong π-stacking interaction with the aromatic adenine ring of the substrate (Figure 2.19). The two aromatic residues are part of a hydrophobic cluster that

also includes Ile440, Met451, Val493, and Trp511, all of which engage in van der Waals contact with the adenine ring of the substrate.

Aside from this sandwich π-stacking interaction, there are minimal contacts between the polypeptide and the substrate molecule. With respect to the aminoacyl portion of the substrate, the main chain carbonyl oxygen is within hydrogen bonding distance (2.6 Å) from Nδ2 of Asn497, while for the side chain, the closest interacting residue is Lys545, in which Nζ is 3.7 Å away from Oδ2 of processed McC7. The ribose oxygen atoms are engaged in hydrogen bonding interactions with Ser495 (O2'- Oγ distance of 3.3 Å) and the backbone carbonyl of Tyr510 (O3'-O distance of 3.3 Å). One of the phosphoramine oxygen atoms is within hydrogen bonding distance (2.9 Å) from the indole nitrogen of Trp453. Other than the π-stacking interaction with Trp453 and Phe466, there are minimal contacts with any of the polar atoms of the adenine ring. Both N6 and N7 are solvent exposed, while N3 is 3.1 Å away from the indole nitrogen of Trp511.

Although the 3-aminopropyl group of McC7 is directed outwards to bulk solvent and does not make contacts with the polypeptide, electron density corresponding to this modification can clearly be observed in the co-crystal structure. The acetyl group is situated in a hydrophobic cleft defined by Ile508, Tyr510, and Ile544, where the acetyl oxygen is located within hydrogen bonding distance (3.0 Å) from the backbone carbonyl of Ile508.

The mode of purine base engagement by π-stacking interactions between two aromatic amino acid side chains (Trp453 and Phe466 of $MccE^{CTD}$) is reminiscent of recognition mechanisms used by eukaryotic and viral protein that recognize the 7-methylguanosine cap at the 5' end of host messenger RNA (mRNA) (Fechter and Brownlee, 2005), and by bacterial and

viral nucleotidyltransferases, such as DNA/RNA ligases and RNA capping enzymes (Shuman and Lima, 2004). With respect to mRNA cap binding proteins, despite the fact that each contain structurally distinct folds, they all share a hydrophobic binding pocket where the methylated guanosine is sandwiched between two aromatic residues, one of which is usually a tryptophan or tyrosine (Guilligay et al., 2008; Hodel et al., 1996; Marcotrigiano et al., 1997; Mazza et al., 2002). Preference for the methylated over non-methylated guanosine is dictated by a stronger interaction between the delocalized positive charge on the 7-methylguanosine and the π-electrons of the aromatic residues. In the nucleotidyltransferase superfamily, the purine nucleotide substrate is similarly sandwiched between a conserved aromatic residue and a conserved aliphatic residue, and additional contacts with either a conserved glutamate (phage/bacterial DNA ligase), threonine (viral DNA ligase), or lysine (RNA capping enzymes) residue.

In order to determine whether $MccE^{CTD}$ could indeed recognize and bind purine ribonucleotides independent of the identity of the aminoacyl conjugates, we used isothermal titration calorimetry to monitor binding between AMP and $MccE^{CTD}$. The binding isotherm for the interaction was fitted by nonlinear regression using a simple bimolecular interaction model. Two concentrations of AMP were used to generate binding isotherms and independently fitted. The data is consistent with a 1:1 interaction between the ligand and protein with a dissociation constant of about 1.03 mM (Figure 2.20A) (Table 2.6). We also determined the 1.3 Å resolution co-crystal structure of $MccE^{CTD}$-acetyl CoA in complex with AMP. The cocrystal structure shows clear electron density for both AMP and for the intact acetyl group of acetyl CoA (Figure 2.20B). The location of the AMP is nearly identical to that occupied by the adenosine moiety in the co-crystal structure with processed McC7. The remainder of the substrate binding cleft is

occupied by a number of solvent molecules. The polypeptide engages AMP with a similar set of interactions as those observed in the co-crystal structure with McC7. Additionally, the side chain of Asn497 is now oriented towards the AMP, where Oδ1 is within hydrogen bonding distance (3.2 Å) from the O2' of the ribose.

An indirect evidence for the requirement of the structural recognition of the adenine ring of McC7 for its acetylation by $MccE^{CTD}$ is presented by the crystal structure of the *E. coli* RimI and RimL enzymes. Though $MccE^{CTD}$ and RimI are highly homologous to each other at the primary sequence level, RimI has been shown to be unable to acetylate and provide resistance against McC7 (Severinov laboratory, unpublished results). In order to facilitate the structural comparison between the active sites of $MccE^{CTD}$ and RimI, we determined the crystal structure of RimI in the presence of acetyl CoA to a limiting resolution of 1.8 Å (Figure 2.21). The overall structures, and the mode of binding of acetyl CoA for $MccE^{CTD}$ and RimI are remarkably similar, and a similar pair of parallel planar side chains is evident in the crystal structure of RimI which could sandwich the adenine ring of McC7 by π-stacking interactions. However the adenine binding pocket of RimI is occluded by the presence of the side chains and Glu19 and Asn68 side chains. This steric occlusion of the adenine binding pocket by Glu19 and Asn68 side chains likely contributes to the inability of RimI to acetylate processed McC7.

In contrast, the *E. coli* RimL enzyme has demonstrated to acetylate processed McC7. We determined the crystal structure of the RimL enzyme to a limiting resolution of 1.8 Å. Though the crystal structure proved recalcitrant to crystallographic refinement owing to twinning or other pathologies associated with the X-ray diffraction data, an incomplete model demonstrated the presence of a pair of parallel planar amino acid side chains which are appropriately positioned for π-stacking interactions with the adenine ring of McC7. This putative adenine binding site is

not occluded by other amino acid side chains, and likely presents a structural explanation for the acetylation of McC7 by RimL.

Substrate promiscuity of MccE^{CTD} explained by cocrystal structures with DSA and ESA

MccE^{CTD} has been demonstrated to acetylate a wide variety of amino nucleotide conjugates in addition to the physiological substrate- processed McC7. It was shown to acetylate DSA, ESA and the leucyl-sulfamoyl-adenosine (LSA) (Novikova et al., 2010). In order to probe the determinants of substrate promiscuity for the enzyme, we obtained crystal structure of the MccE^{CTD} in presence of acetyl-CoA and DSA and ESA. Crystallization was also attempted in the presence of LSA, valinyl-sulfamoyl-adenosine (VSA) and prolyl-sulfamoyl-adenosine (PSA), but diffraction quality crystals could not be obtained in the presence of these substrates and acetyl CoA.

MccE^{CTD} engages DSA and ESA through contacts similar to those observed in co-crystal structure with processed McC7 (Figure 2.22). The primary determinant for substrate engagement is the π-stacking of the adenine ring. No contacts could be observed for either the aspartyl or the glutamyl amino acid side chains. In each of the two structures, no electron density can be observed for the acetyl group of the donor, suggesting that hydrolysis of the thioester has occurred *in situ*. In addition, the side chain of Cys546 is rotated towards the active site pocket where it forms a covalent disulfide linkage with CoA. Although these structures suggest that MccE^{CTD} proceeds through an enzyme-acetyl intermediate, this is unlikely as mutational analysis of the equivalent residue in RimL suggest that this cysteine does not play a role in catalysis. A similar cysteyl-acetyl linkage is observed in the RimL co-crystal structure (Vetting et al., 2005a).

The lack of specific contacts with substrate also explains the ability of MccE^{CTD} to detoxify the structurally distinct seryl-tRNA synthetase inhibitor albomycin (Novikova et al., 2010) and establish the structural premise for the observed substrate promiscuity for the enzyme. Processed albomycin is composed of an N^4-carbamyl-5-methyl-4-imino-cytidine conjugated to a thioxylofuranosyl pyrimidine (Stefanska et al., 2000), and the nucleoside moiety may be similarly accommodated into the MccE^{CTD} hydrophobic substrate-binding pocket. The rate and specificity for the modification of processed albomycin by MccE^{CTD} is not yet clear as these studies were conducted as end-point assays using whole cell extracts (Novikova et al., 2010).

Kinetic analysis for the acetylation of ESA by MccE^{CTD}

To further characterize the acetylation activity of the MccE^{CTD} enzyme for aminoacyl nucleotide conjugated substrates, we carried out kinetic analysis of MccE^{CTD} by reacting 5,5'-dithiobis(2-nitrobenzoic acid) (Ellman's reagent) with the free sulfhydryl of the product CoA that is formed upon acetylation of the substrate. Using ESA as a substrate, the spectrophotometric studies yield a K_M value of 55 ± 7 μM and k_{cat} value of 23.7 ± 2.2 min^{-1}. As neither the side chain of the sulfamoyl adenylate or the propylamine of McC7 makes any specific contacts with MccE^{CTD}, the kinetic constants for DSA and processed McC7 are likely very similar to those experimentally determined for ESA. The k_{cat} value obtained for MccE^{CTD} is in close agreement with that of the *S. typhimurium* RimL acetyltransferase for a ribosomal peptide substrate (Vetting et al., 2005a), though the K_M value obtained for MccE^{CTD} is nearly 50 times higher than that for RimL. However the k_{cat} values for the N-acetylation of aminoglycoside antibiotics by the respective resistance causing acetyltransferase enzymes is nearly two orders of

magnitude higher than that for MccE^{CTD} (Hamano et al., 2004; Norris et al., 2010). Curiously, these N-acetyltransferases also display remarkable substrate promiscuity.

Model for the synergistic activity for full length MccE

Full length MccE enzyme is a bi-domain protein. Putative enzymatic role can be assigned to the MccE N-terminal domain (MccE^{NTD}), in addition to the characterized McC7 acetylating activity of the MccE^{CTD}. Together with MccD, full length MccE was shown to contribute towards the attachment of the propylamine appendage upon the phosphoramide oxygen atom (Metlitskaya et al., 2009). Sequence homology identifies MccD to be a S-adenosyl-methionine (SAM) dependent methyl transferase, while MccE^{NTD} bears homology to pyridoxal-5'-phosphate (PLP) dependent decarboxylases. The in-vitro activity for either MccD or MccE enzymes has not been reconstituted as yet. The chemical structure of McC7 or processed McC7 does not bear a methyl group on any nucleophillic nitrogen or oxygen atom. Hence the postulate that MccD would transfer a methyl group from SAM, without the recruitment of radical chemistry, is highly unlikely. We propose that MccD instead transfers a 2-ammoniobutanoate moiety from SAM to one of the phosphoramide oxygen atoms (Figure 2.23). The reaction mechanism would involve nucleophilic attack on the Cγ atom of the methionine side chain of SAM, rather than at the Cε methyl group. This enzymatic reaction has been reported before in the biosynthesis of polyamines in plants and nicotianamine biosynthesis in archaea (Dreyfus et al., 2009; Roje, 2006). The transfer of 2-ammoniobutanoate would convert SAM to methyl-thio-adenosine (MTA). The 2-ammoniobutanoate could then be decarboxylated by MccE^{NTD} to generate the propylamine appendage. This hypothesis is lent further credence by the sequence homology of

MccE^{NTD} with Type III ornithine and diaminopimelate decarboxylase enzymes (Gokulan et al., 2003).

The putative transfer of 2-ammoniobutanoate from SAM to the phosphoramide oxygen atom could be succeeded by the removal of the McC7 Trojan horse hexapeptide within the producer cell by the action of the PepA-N proteases. In the event that this proteolysis event precedes the decarboxylation by MccE^{NTD}, MccE^{CTD} could act concomitantly with MccE^{NTD} to acetylate the now free aspartate main chain primary amine to prevent the synthesis of a toxic species within the producer cell. A concomitant action on one substrate molecule by two different domains of the MccE enzyme cannot be ruled out by a steric argument. The propylamine appendage of processed McC7 points outwards from the active site of MccE^{CTD} and does not make any contacts with the enzyme. Hence it is sterically 'available' to be accepted within the MccE^{NTD} active site. As the propylamine appendage of the McC7 makes it a more potent AspRS inhibitor (Metlitskaya et al., 2009), it is conceivable that the McC7 biosynthetic machinery has been evolved to position a self-immunity determinant to act synergistically with an antibiotic maturation event.

MccE^{NTD} and MccE^{CTD} can also share some McC7 recognition structural motifs. Our current hypothetical model for the full length MccE enzyme is based upon the recognition of the adenine moiety by π-stacking interactions in MccE^{CTD}, while the aspartate side chain (which also points outwards from the MccE^{CTD} active site) and the carboxylated propylamine appendage are recognized within the MccE^{NTD}, as shown in Figure 2.23. This hypothesis is currently under investigation.

Future directions

Several questions remain with respect to the natural biological role of the MccF enzyme. The ubiquitous presence of the MccF gene in the microbial metabolome across diverse ecological niches makes us speculate that MccF might have a conserved biological function associated with growth and fitness of these microbes. Several other small molecule stable aminoacyl nucleotide conjugates are produced as natural products by bacteria. Primary examples with demonstrated biological activities are agrocin 84 (Reader et al., 2005), albomycin (Pramanik et al., 2007), ascamycin (Osada and Isono, 1985; Ubukata et al., 1985), pacidamycin (Winn et al., 2010), capuramycin (Reddy et al., 2008) and polyoxins (Zhang and Miller, 1999), among others. Our results have demonstrated that mutations within the *E. coli* MccF active site can expand upon the substrate scope of the enzyme. The enzyme is also intrinsically pre-organized to recognize different nucleotides. Hence it can be speculated that homologs of MccF might offer resistance against the bioinhibitory activities of several more natural products, other than McC7. As inhibitors of antibiotic degrading enzymes can be valuable adjuvants to the dosing regimen of an antibiotic, as exemplified by the development of clavulanic acid adjuvant in broad spectrum antibiotic amoxicillin (Drawz and Bonomo, 2010; Hugonnet et al., 2009), inhibitors of MccF might allow the aminoacyl nucleotide natural products to explore and even more diverse target spectrum. Our structural and kinetic results have provided a framework which would make the development of inhibitors of MccF easier. This process would also be aided by previous extensive literature on the development of synthetic inhibitors for serine protease enzymes.

Our inability to reconstitute the enzymatic activity for the MccD and MccE enzymes continues to rankle and confound us. We have unsuccessfully trialed a variety of experimental

conditions to install either the aminopropyl or the 2-ammoniobutanoate moieties on the purified MccB reaction product. Future efforts will be directed at abolishment of the MccE activity within the producer cell, and then isolating a McC7 biosynthetic pathways intermediate compound. This could serve as the physiological substrate for the MccE enzyme in-vitro. This abolishment of activity for the MccE enzyme would be carried out by site directed mutagenesis within the $MccE^{NTD}$, and not by knocking out the entire MccE enzyme, as the activity of the $MccE^{CTD}$ is associated with fitness costs for the producer cell (Novikova et al., 2010). As mentioned before, efforts to crystallize full length MccE enzyme have not borne fruit as of now. Presumably, the presence of the physiological substrate of the enzyme will serve to restrict the conformational flexibility of the bi-domain protein, and make the enzyme more amenable to crystallization.

Several antibiotics have been demonstrated to biologically active only when associated with peptidyl Trojan horse carriers. Examples of these peptidyl-Trojan horse dependent antibiotics are rhizocticins (Kugler et al., 1990), plumbemycins (Fredenhagen et al., 1995) and dehydrophos (Circello et al., 2011), among others. It has also been demonstrated that just by changing the nature of the peptidyl Trojan horse carrier, the antifungal rhizocticins can be converted to antibacterial plumbemycins, as they share a common bioinhibitory chemical entity. Hence it is conceivable that the microbial target spectrum for McC7 can be expanded by altering the amino acids in the Trojan horse hexapeptide. The first steps towards this direction have been reported in literature (Van de Vijver et al., 2009b; Vondenhoff et al., 2011). This process would also be aided by the structural and biochemical characterization of the McC7 import machinery. Import of McC7 into susceptible relies on the YejABEF ATP-binding cassette (ABC) transporter (Eswarappa et al., 2008; Novikova et al., 2007), with the YejA protein acting as the extracellular

McC7 recognition determinant. We have been able to generate crystals of *E. coli* YejA in presence of MccA and collect preliminary native diffraction data sets. Elucidation of the YejA-MccA complex structure would allow for the rational redesign of the antibiotic with decreased minimum inhibitory concentration (MIC) values and better pharmacokinetic profiles. Semi-synthetic alteration in the identity of the Trojan horse carrier to analogous dipeptides (rhizocticins and plumbemycins), sugar molecules (agrocin 84) or siderophores (albomycin) can also expand upon the target spectrum of McC7.

Tables

Table 2.1 Data collection and refinement statistics for MccF wild type enzyme, Ser118Ala and Ser118Ala in complex with processed McC7.

	MccF (native)	Ser118Ala	MccF-McC7
Data collection			
Cell dimensions			
a, b,c (Å), β(°)	54.4,85.9,73.1, 101.2	54.5,85.8,73.2, 101.2	54.3,85.7,72.9, 101.3
Resolution (Å)[1]	50-1.2 (1.24-1.2)	50-1.5 (1.55-1.5)	50-1.3 (1.35-1.3)
R_{sym} (%)	6.8 (35.2)	6.9 (12.1)	5.0 (28.2)
I/σ(I)	38.7 (3.9)	46.3 (19.2)	29.8 (4.5)
Completeness(%)	98.9 (93.8)	99.7 (99.8)	97.2 (93.8)
Redundancy	7.3 (5.8)	4.6 (4.6)	4.4 (4.1)
Refinement			
Resolution (Å)	25-1.2	25-1.5	25-1.3
No. reflections	192,327	99,890	148,363
R_{work} / R_{free}[2]	18.3/20.1	17.7/19.8	17.4/18.6
Number of atoms			
Protein	5,313	5,188	5,277
Ligand	-	-	70
Water	961	981	978

1. Highest resolution shell is shown in parenthesis.
2. R-factor = $\Sigma(|F_{obs}|-k|F_{calc}|)/\Sigma |F_{obs}|$and R-free is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table 2.2 Data collection and refinement statistics for MccF Ser118Ala enzyme in complex with DSA, ESA, ESI and ESG.

	MccF-ESA	MccF-DSA	MccF-ESI	MccF-ESI
Data collection				
Cell dimensions				
a, b, c (Å) β(°)	54.5, 85.9, 73.1 101.2	54.4, 85.7, 72.9 101.3	54.5, 86.1, 73.2 101.2	54.4, 85.7, 73.0 101.2
Resolution (Å)[1]	50-1.35 (1.4-1.35)	50-1.5 (1.55-1.5)	40–1.20 (1.22–1.2)	50–1.23 (1.25–1.23)
R_{sym} (%)	4.5 (14.5)	5.8 (24.4)	6.2 (62.6)	9.0 (41.5)
I/σ(I)	39.6 (9.5)	27.9 (5.4)	27.5 (2.0)	26.6 (3.3)
Completeness(%)	98.2 (96.2)	95.6 (85.0)	98.5 (96.2)	100.0 (100.0)
Redundancy	3.8 (3.6)	6.0 (5.3)	5.8 (5.3)	6.2 (6.0)
Refinement				
Resolution (Å)	25-1.3	25-1.5	25-1.20	25-1.23
No. reflections	149,510	94,666	193215	180541
R_{work} / R_{free}[2]	17.9/19.7	17.9/20.2	0.159 / 0.175	0.155 / 0.172
Number of atoms				
Protein	5,261	5,261	5343	5335
Ligand	64	62	64	66
Water	957	864	617	589

1. Highest resolution shell is shown in parenthesis.
2. R-factor = $\Sigma(|F_{obs}|-k|F_{calc}|)/\Sigma |F_{obs}|$and R-free is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table 2.3 Steady state kinetic parameters for DSA and ESA hydrolysis by MccF wild type and mutant enzymes.

	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($M^{-1} s^{-1}$)	Relative k_{cat}/K_M [1]
		DSA		
Wild type	74.3 ± 6.1	417.7 ± 33.3	9.4E4	1.00
Asn220Ala	186.3 ± 20.2	133.4 ± 12.2	1.2E4	0.13
Lys247Ala	133.9 ± 15.1	196.3 ± 20.1	2.4E4	0.26
Ser118Ala	n.d. [2]	n.d.		
		ESA		
Wild type	51.6 ± 4.9	35.0 ± 4.1	1.1E4	1.00
Trp186Phe	218.8 ± 19.4	23.3 ± 4.0	1.8E3	0.16
Arg246Ala	154.3 ± 16.3	33.7 ± 2.5	3.6E3	0.32
Asn220Ala	105.9 ± 9.5	37.2 ± 4.9	5.9E3	0.52
Lys247Ala	238.9 ± 25.2	5.0 ± 0.3	3.5E2	0.03
Trp186Ala	n.d.	n.d.		
Asn220Ala/Lys247Ala	n.d.	n.d.		
Ser118Ala	n.d.	n.d.		

[1] k_{cat}/K_M of mutant enzyme relative to wild type enzyme
[2] n.d.: no detectable activity precluded kinetic parameter determination

Table 2.4 Steady state kinetic parameters for ESA, ESI and ESG hydrolysis by MccF wild type enzyme.

	K_M (μM)	k_{cat} (min^{-1})	k_{cat}/K_M ($M^{-1} s^{-1}$)
ESA	57.6 ± 4.2	36.7 ± 3.1	1.06E4
ESI	90.9 ± 10.5	38.1 ± 4.2	0.70E4
ESG	33.4 ± 2.2	34.9 ± 3.1	1.74E4

* Determined in an independent experiment from Table 2.4.

Table 2.5 Data collection, phasing and refinement statistics for MccE^{CTD} enzyme in the apo form, and in complex with processed McC7, DSA, ESA and AMP.

	Acetyl CoA	DSA	ESA	AMP	McC7
Data collection					
Space group	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$	$P2_12_12$
Cell dimensions					
a, b, c (Å)	78.9, 94.7, 53.9	78.9, 94.8, 53.3	78.9, 94.8, 53.4	78.5, 94.9, 53.3	77.9, 95.1, 53.2
Resolution (Å)[1]	50-1.6 (1.66-1.6)	50-1.25 (1.29-1.25)	50-1.2 (1.24-1.2)	50-1.3 (1.35-1.3)	50-1.35 (1.4-1.35)
R_{sym} (%)	6.1 (36.1)	6.3 (42.7)	6.5 (44.7)	6.5 (23.5)	6.6 (49.7)
I / σ(I)	28.7 (4.8)	46.9 (2.9)	51.7 (3.3)	40.5 (11.2)	45.7 (2.6)
Completeness(%)	98.8 (96.9)	99.5 (96.6)	99.3 (96.0)	99.8 (100.0)	99.2 (95.2)
Redundancy	7.2 (7.1)	9.1 (5.4)	8.6 (6.9)	12.1 (11.8)	6.5 (4.4)
Refinement					
Resolution (Å)	25.0-1.6	25.0-1.25	25.0-1.2	25.0-1.3	25.0-1.35
No. reflections	50,128	105,359	118,379	93,420	82,500
R_{work} / R_{free}[2]	20.3/22.8	18.2/20.2	18.5/19.9	17.6/19.7	18.7/21.6
Number of atoms					
Protein	2800	2864	2836	2787	2787
(Ac)CoA	102	96	96	102	96
Substrate	-	62	64	46	76
Water	419	610	665	696	517

1. Highest resolution shell is shown in parenthesis.
2. R-factor = $\Sigma(|F_{obs}|-k|F_{calc}|)/\Sigma |F_{obs}|$and R-free is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table 2.6 Thermodynamic parameters obtained by fitting the isothermal titration calorimetry binding isotherms to a bimolecular interaction model for binding of AMP to $MccE^{CTD}$. Isotherms were obtained for two concentrations of AMP and similarly analyzed.

	50mM AMP	**20mM AMP**
N	1.08	1.10
K^*	963 ± 97 M^{-1}	948 ± 23 M^{-1}
ΔH	-1.4E4 ± 475.5 cal/mol	-1.5E4 ± 174.2 cal/mol
ΔS	-34.4 cal/mol.deg	-37.8 cal/mol.deg
ΔG	-4.1E3 cal/mol	-4.1E3 cal/mol

* This value refers to association constant

Figures

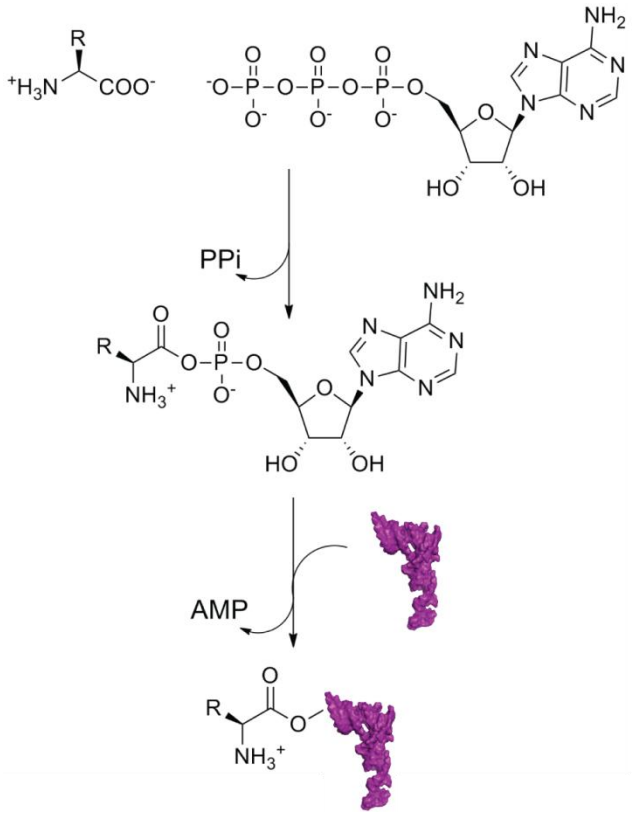

Figure 2.1 Two-step mechanism of aaRS enzymes. In the first step, the cognate amino acid is adenylated, which is then transferred to the ribose hydroxyl of the tRNA molecule in the second step. One molecule of ATP is consumed in this process. The tRNA molecule is shown in surface representation in purple.

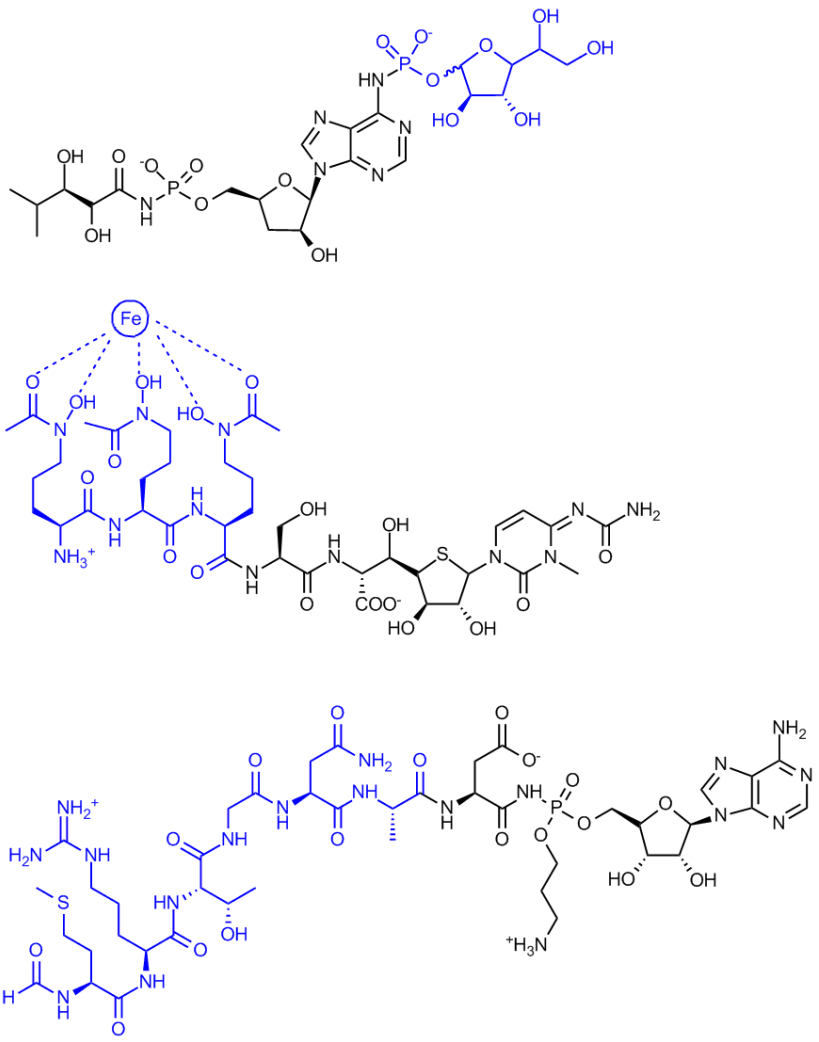

Figure 2.2 Structures of trojan horse aminoacyl-nucleotide conjugate natural product aaRS inhibitors- agrocin 84 (top), albomycin (middle) and microcin C7 (bottom). The Trojan horse moieties which need to be removed prior to inhibitory action of the molecule are shown in blue. The tri-hydroxamate siderophore moiety of albomycin is shown coordinating a ferrous ion.

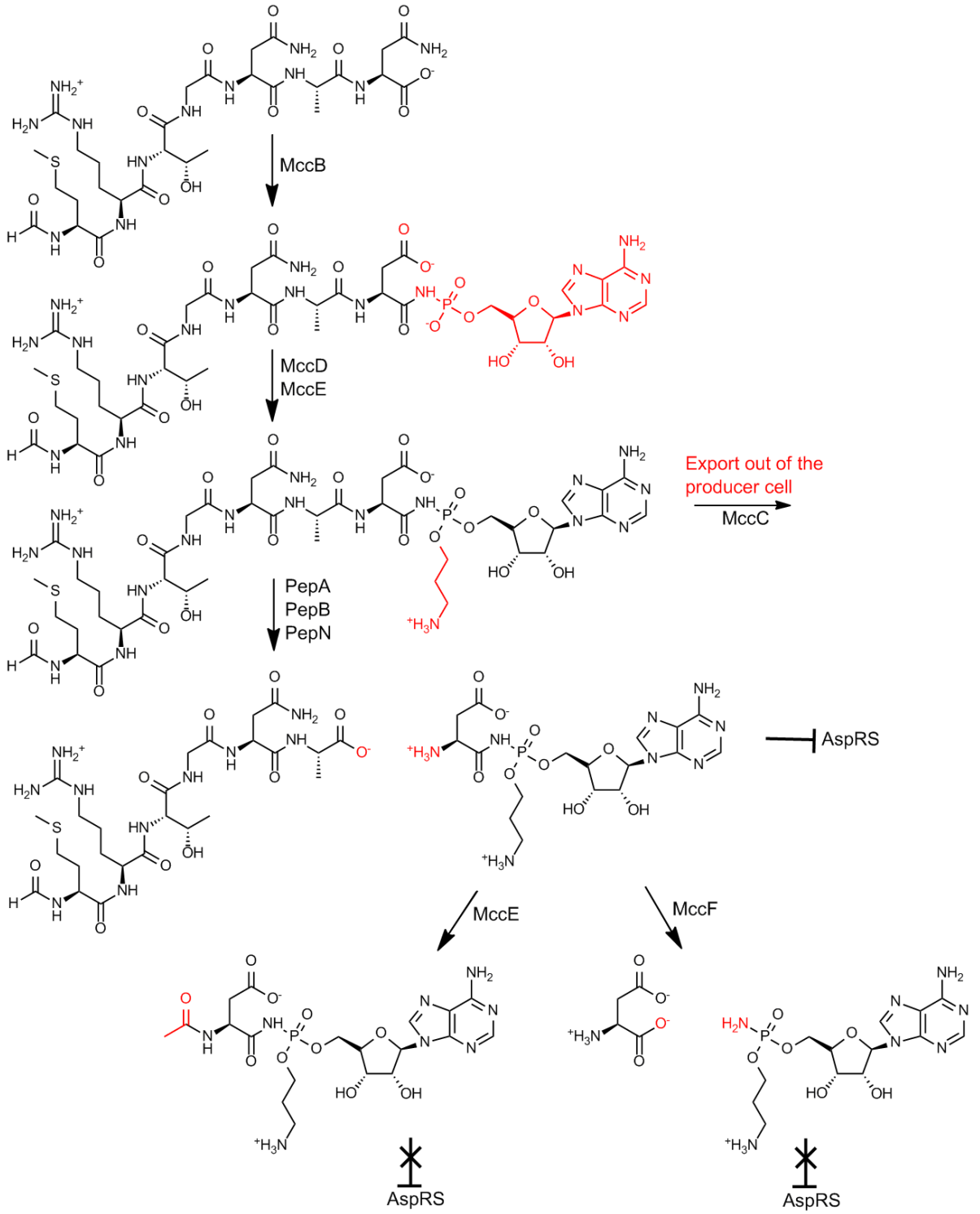

Figure 2.3 Biosynthetic pathway for the production and self-immunity for microcin C7 antibiotic. The chemical transformations at each enzymatic step are highlighted in red.

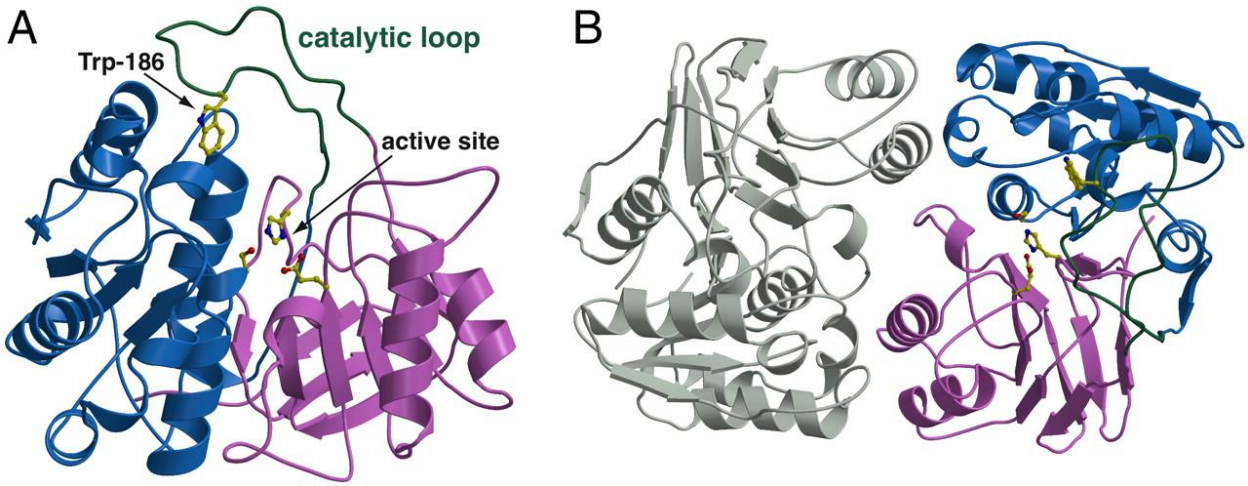

Figure 2.4 Crystal structure of wild type MccF in the apo state. (A) MccF monomer structure showing the three domain architecture. The amino terminal domain is shown in blue, the catalytic loop in green and the carboxy terminal domain in pink. The active site residues are shown in stick ball representation with the carbon atoms colored yellow. **(B)** Homodimer organization of MccF. One dimer is colored as before. The other dimer is colored grey.

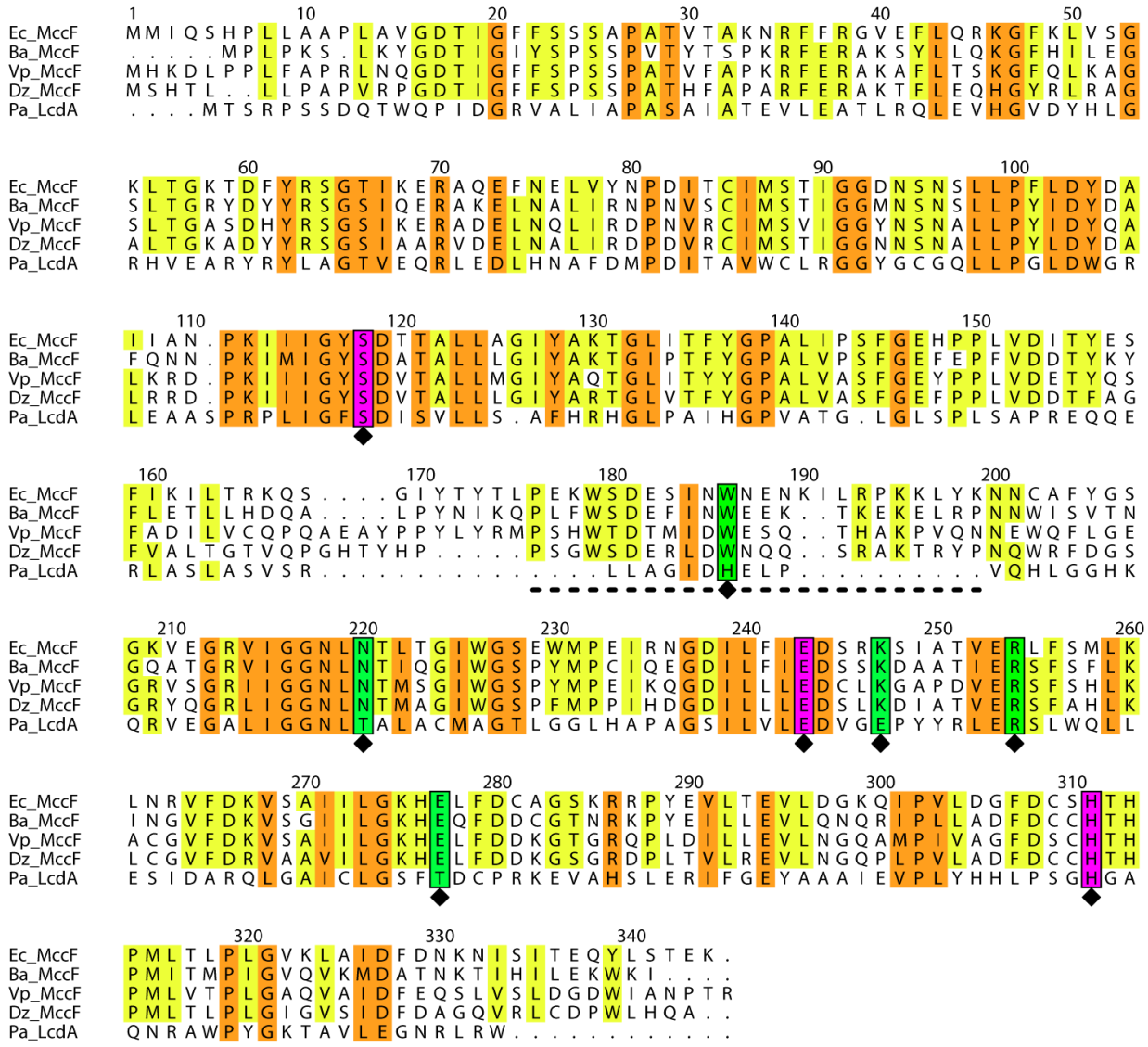

Figure 2.5 Sequence alignment of *E. coli* MccF (Ec_MccF) with homologs from human pathogens *B. anthracis* (Ba_MccF), *Vibrio parahaemolyticus* (Vp_MccF), plant pathogen *Dickeya zeae* (Dz_MccF) and LD-carboxypeptidase LcdA from *P. aureginosa* (Pa_MccF). Note the absence of the catalytic loop from Pa_LcdA (marked by black dashed line under the sequence). The catalytic triad residues (highlighted in pink boxes) are conserved among all sequences, but McC7 recognition residues (highlighted in green boxes) are absent from LcdA, with the exception of Arg254 which also mediates enzyme dimerization, and is hence conserved in LcdA.

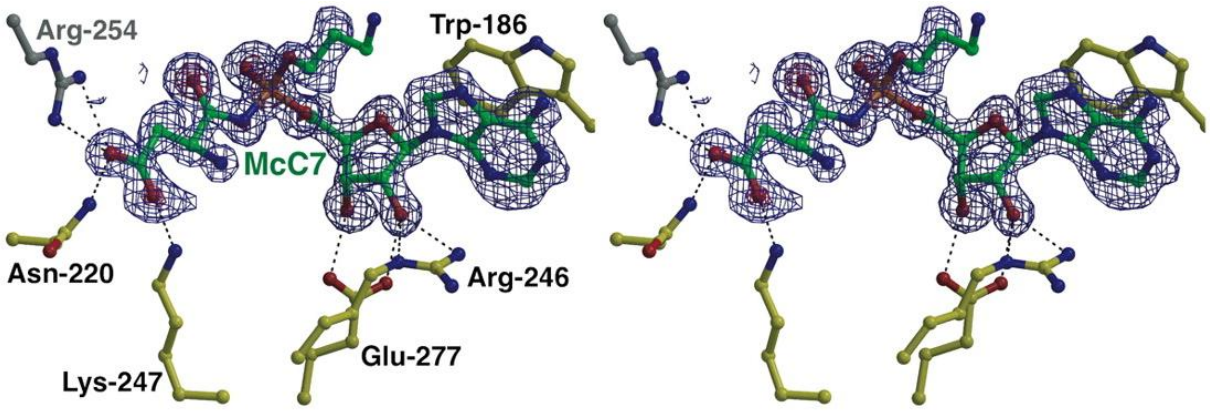

Figure 2.6 Stereo view of processed McC7 bound to the active site of MccF. Stereo view showing the active site features of the MccF-S118A in complex with processed McC7. The MccF carbon atoms are shown in yellow ball-and-stick and the McC7 carbon atoms are colored in green. Superimposed is a difference Fourier electron density map (contoured at 2.7σ over background in blue) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of McC7 deleted prior to one round of refinement.

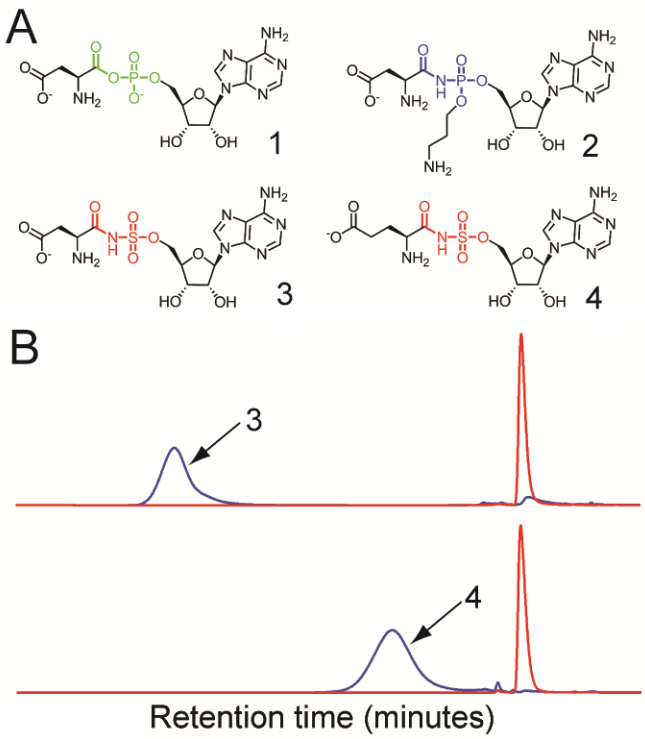

Figure 2.7 Hydrolysis of DSA and ESA by MccF. (A) Chemical structures of AspRS enzyme reaction intermediate- aspartyl adenylate (**1**), processed McC7 (**2**), DSA (**3**) and ESA (**4**). The phosphoester bond in **1** is shown in green, the phosphoramine-amide in **2** is shown in blue, and the sulfamoyl-amide bond in **3** and **4** is shown in red. **(B)** HPLC separation of the reactants DSA and ESA in blue, and the product SA in red. The substrate peaks are marked by arrows.

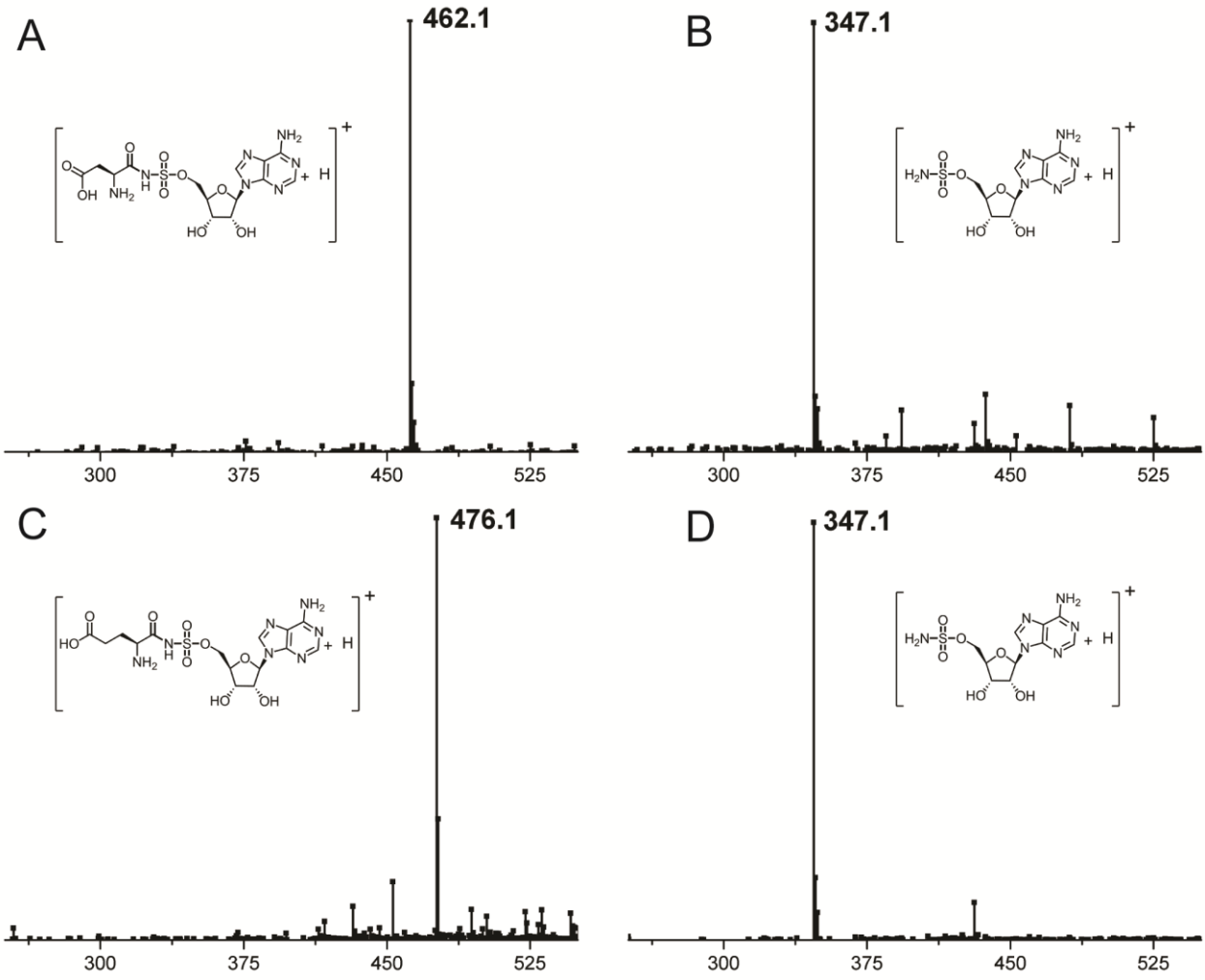

Figure 2.8 ESI-MS spectrum for the reactants and products of hydrolysis reactions of DSA and ESA by wild type MccF enzyme. (A) and **(C) S**pectrum for the substrates DSA and ESA respectively. **(B)** and **(D)** product SA from the hydrolysis reaction of DSA and ESA respectively. The major peaks are labeled with their respective m/z values. Inset shows the chemical species corresponding to the peak mass.

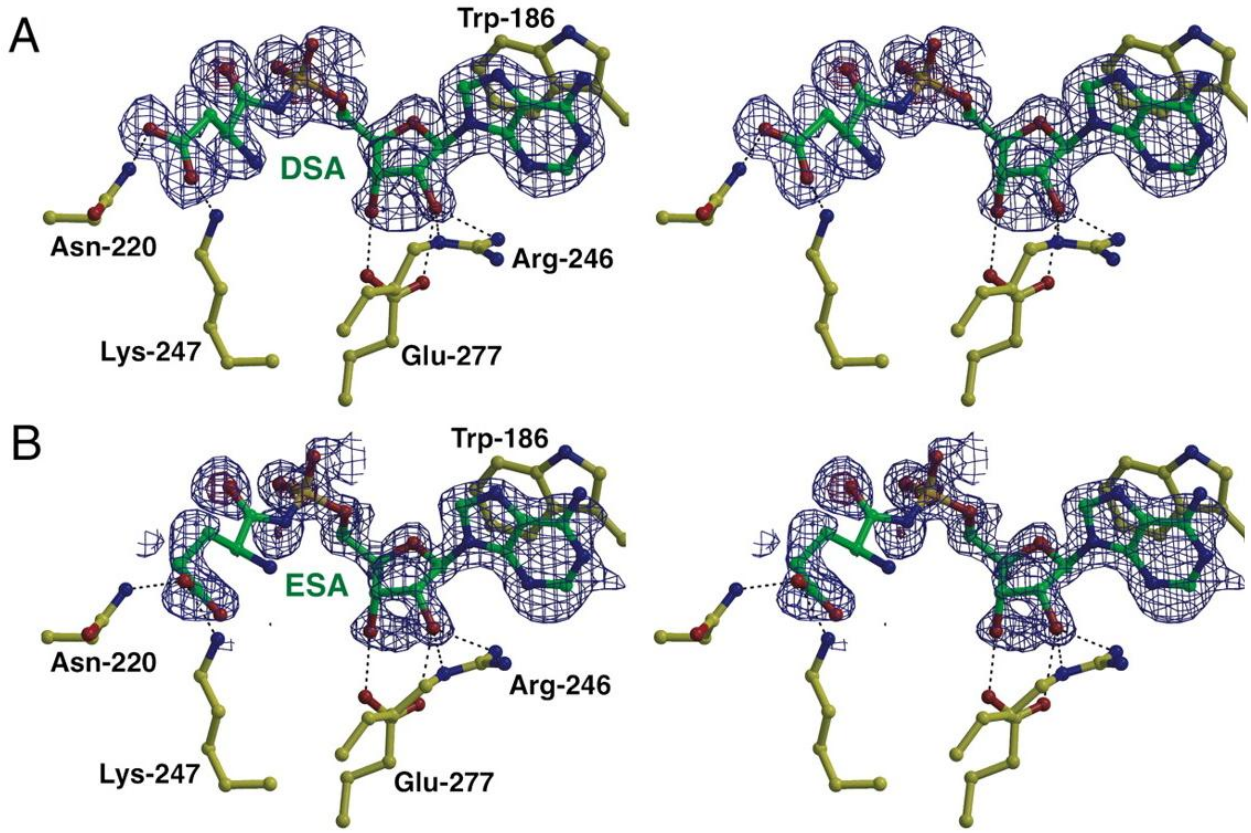

Figure 2.9 Stereo views of DSA and ESA bound to the active site of MccF. Stereo view showing the active site features of the MccF-S118A in complex with **(A)** DSA and **(B)** ESA. The MccF carbon atoms are shown in yellow ball-and-stick and the substrate carbon atoms are colored in green. Superimposed is a difference Fourier electron density map (contoured at 2.7σ over background in blue) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of DSA and ESA deleted prior to one round of refinement.

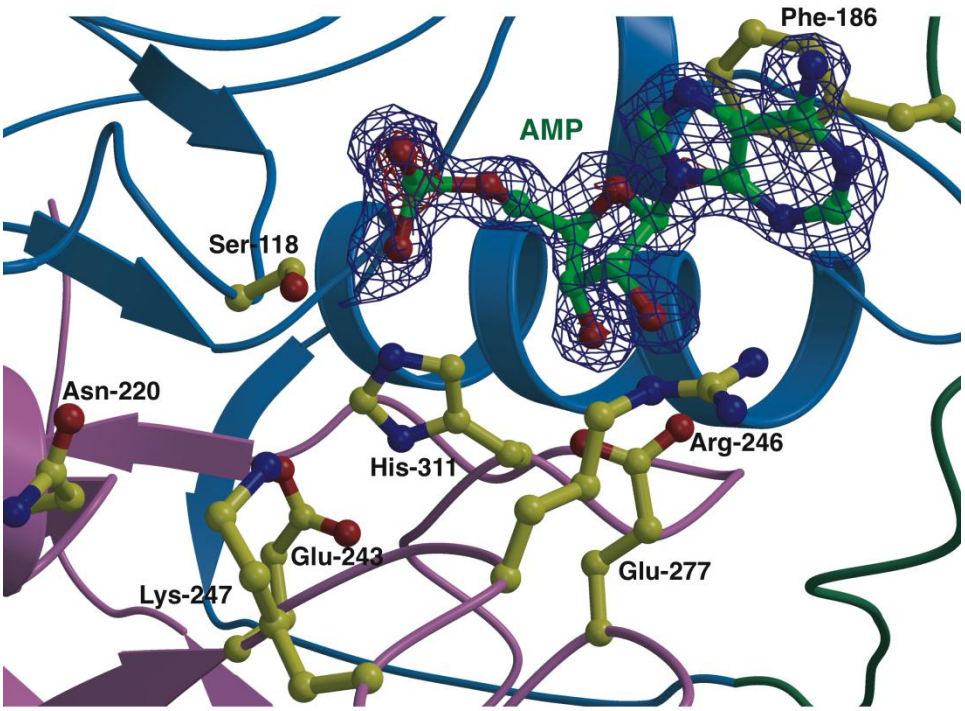

Figure 2.10 Active site view of the co-crystal structure of Trp186Phe MccF with AMP. Enzyme side chain residue carbon atoms are colored in yellow, AMP carbon atoms are colored in green. Superimposed is a difference Fourier electron density map (contoured at 2.7σ over background in blue) calculated with coefficients $|F_{obs}|$ - $|F_{calc}|$ and phases from the final refined model with the coordinates of AMP deleted prior to one round of refinement. MccF N-terminal domain is colored blue, C-terminal domain is colored pink and the catalytic loop is colored green.

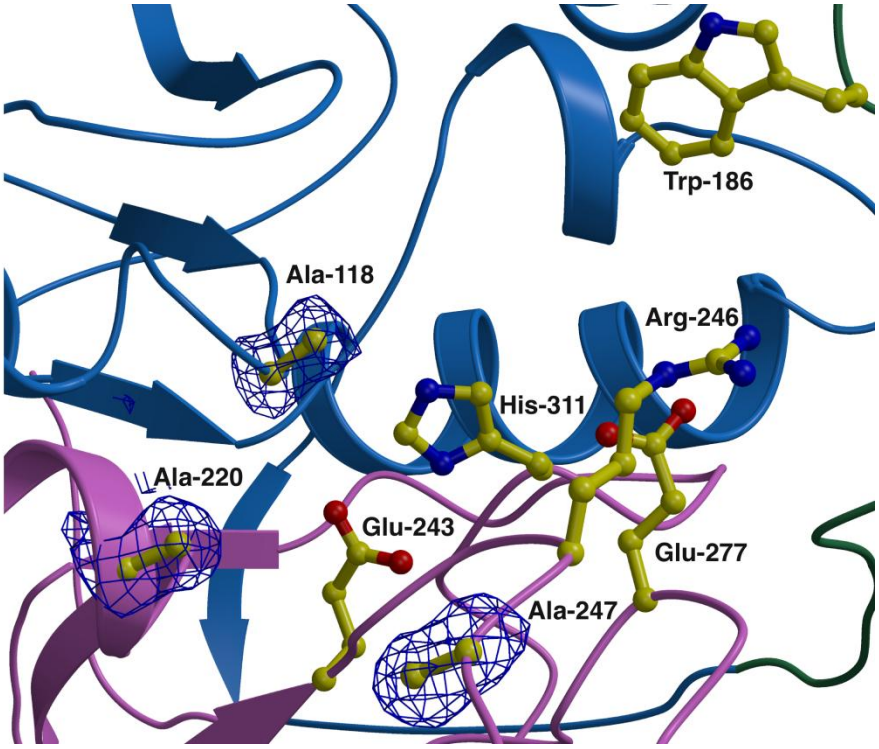

Figure 2.11 Active site view of Ser118Ala/Asn220Ala/Lys247Ala MccF. Enzyme side chain residue carbon atoms are colored in yellow. Superimposed is a difference Fourier electron density map (contoured at 2.7σ over background in blue) calculated with coefficients $|F_{obs}|$ - $|F_{calc}|$ and phases from the final refined model with the coordinates of the side chains of Ala-118, Ala-220, and Ala-247 deleted prior to one round of refinement. MccF N-terminal domain is colored blue, C-terminal domain is colored pink and the catalytic loop is colored green.

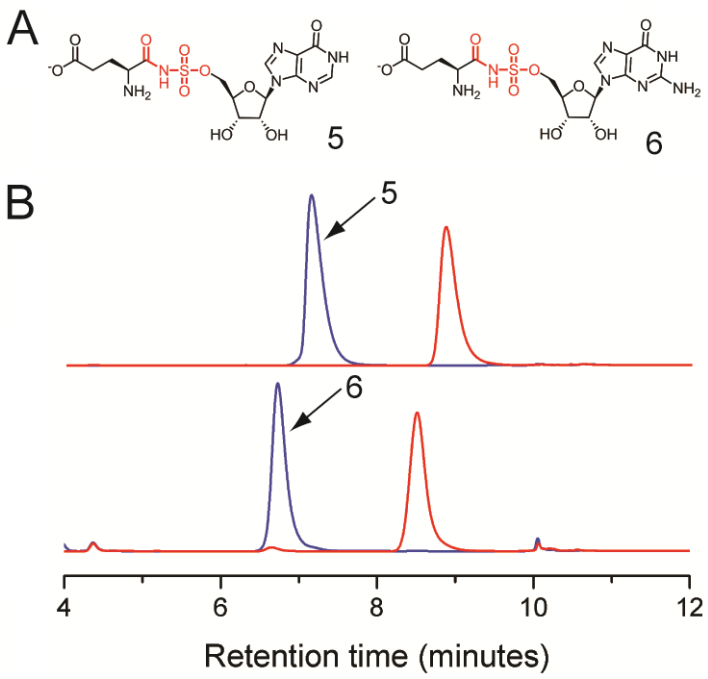

Figure 2.12 Hydrolysis of ESI and ESG by MccF. (A) Chemical structures of ESI (**5**) and ESG (**6**). The sulfamoyl-amide bond in **5** and **6** is shown in red. **(B)** HPLC separation of the reactants ESI and ESG in blue, and the products SI and SG in red. The substrate peaks are marked by arrows.

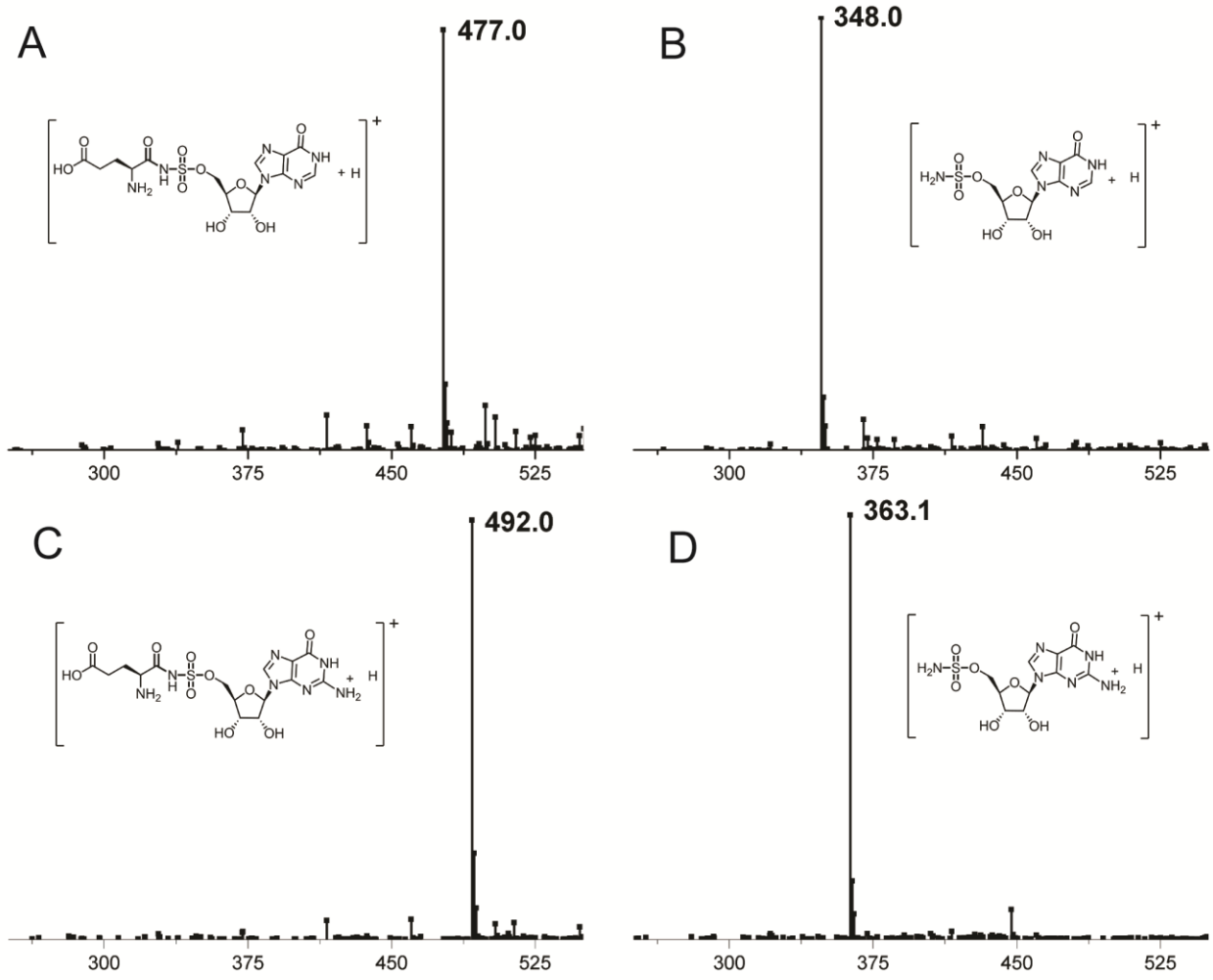

Figure 2.13 ESI-MS spectrum for the reactants and products of hydrolysis reactions of ESI and ESG by wild type MccF enzyme. (A) and **(C) S**pectrum for the substrates ESI and ESG respectively. **(B)** and **(D)** products SI and SG from the hydrolysis reaction of ESI and ESG respectively. The major peaks are labeled with their respective m/z values. Inset shows the chemical species corresponding to the peak mass.

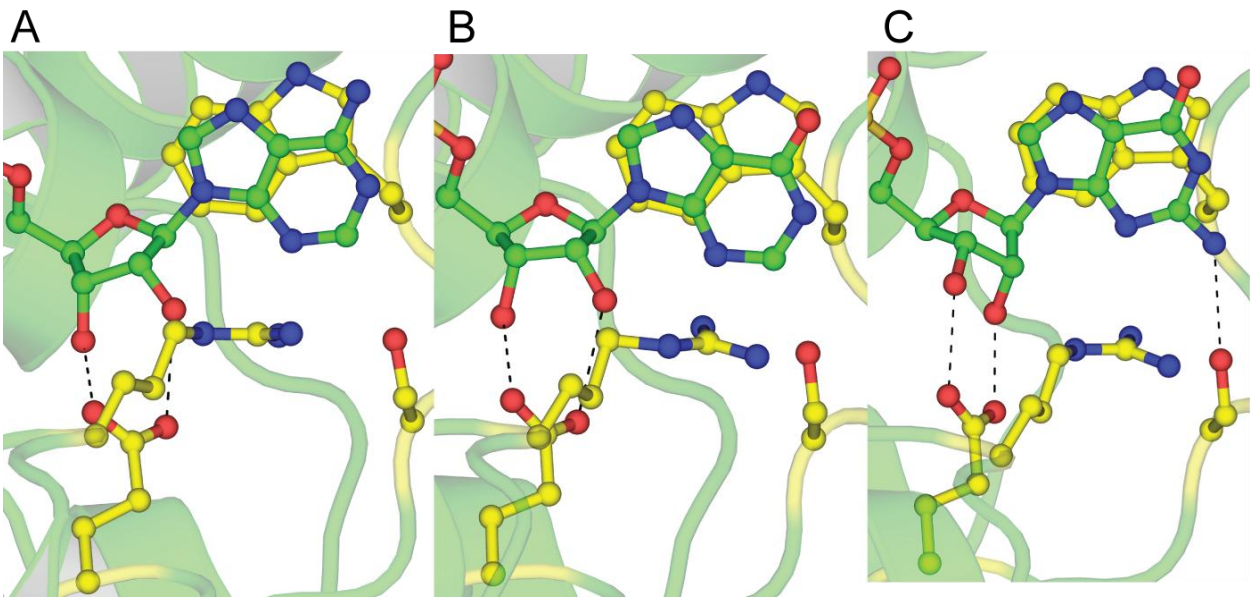

Figure 2.14 Ligand binding at the P1 site for MccF-Ser118Ala in complex with (A) ESA, (B) ESI and (C) ESG. Ligands are shown in stick-ball representation with the carbon atoms colored green. Side chain atoms of MccF residues Ser183, Trp186, Arg246 and Glu277 residues are shown in stick-ball representation with carbon atoms colored yellow. Hydrogen bonds are shown as black dashes.

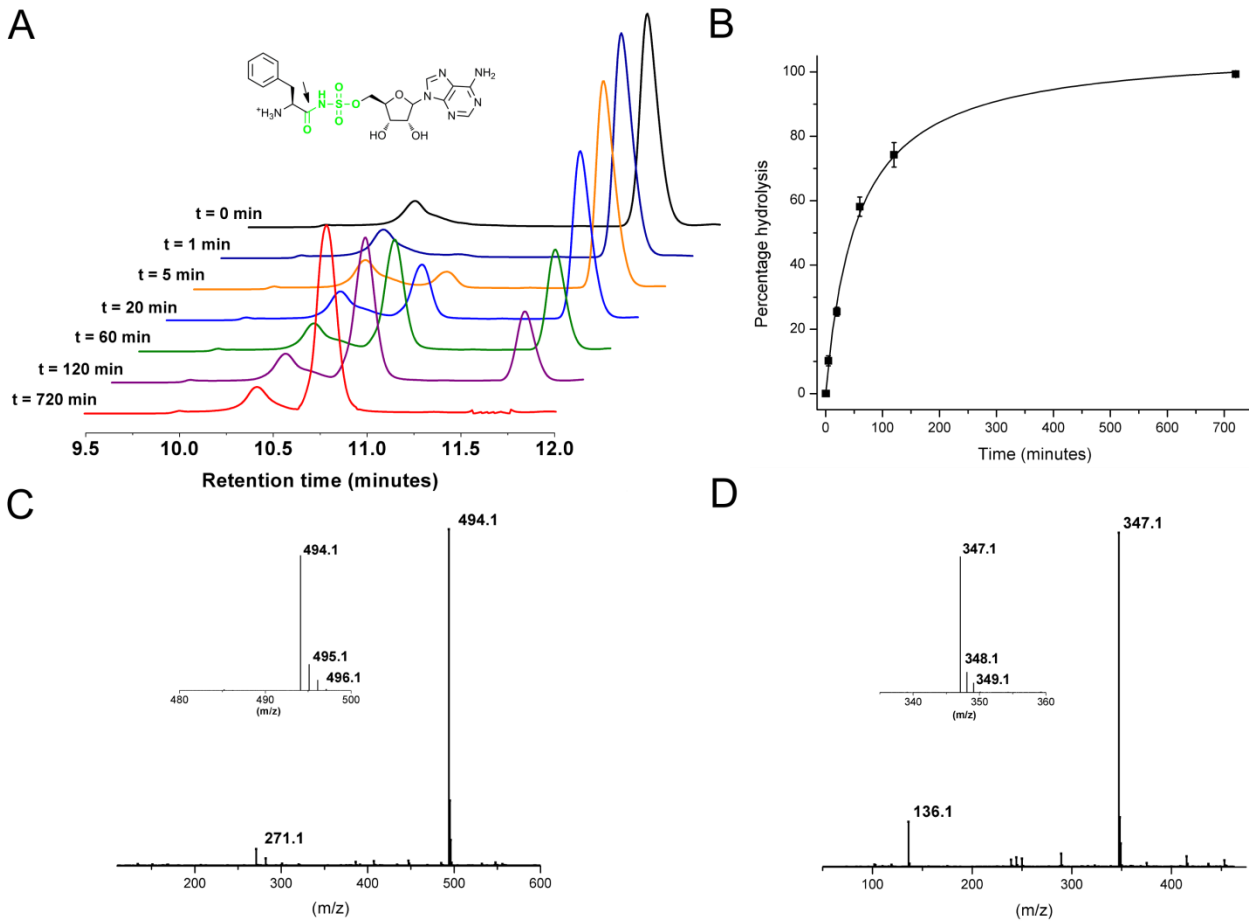

Figure 2.15 Cleavage of FSA by MccF Asn220Leu/Lys247Leu mutant enzyme. (A) Time resolved HPLC separation of the substrate (FSA) and the product (SA). The inset shows the chemical structure of FSA with the site of cleavage marked by an arrow. **(B)** Reaction progression, determined by the rate of appearance of the product, fit to a sigmoidal curve. **(C)** ESI-MS spectrum of the substrate FSA with the monoisotopic masses shown for the major peak shown in inset. **(D)** ESI-MS spectrum of the product SA with the monoisotopic masses shown for the major peak shown in inset.

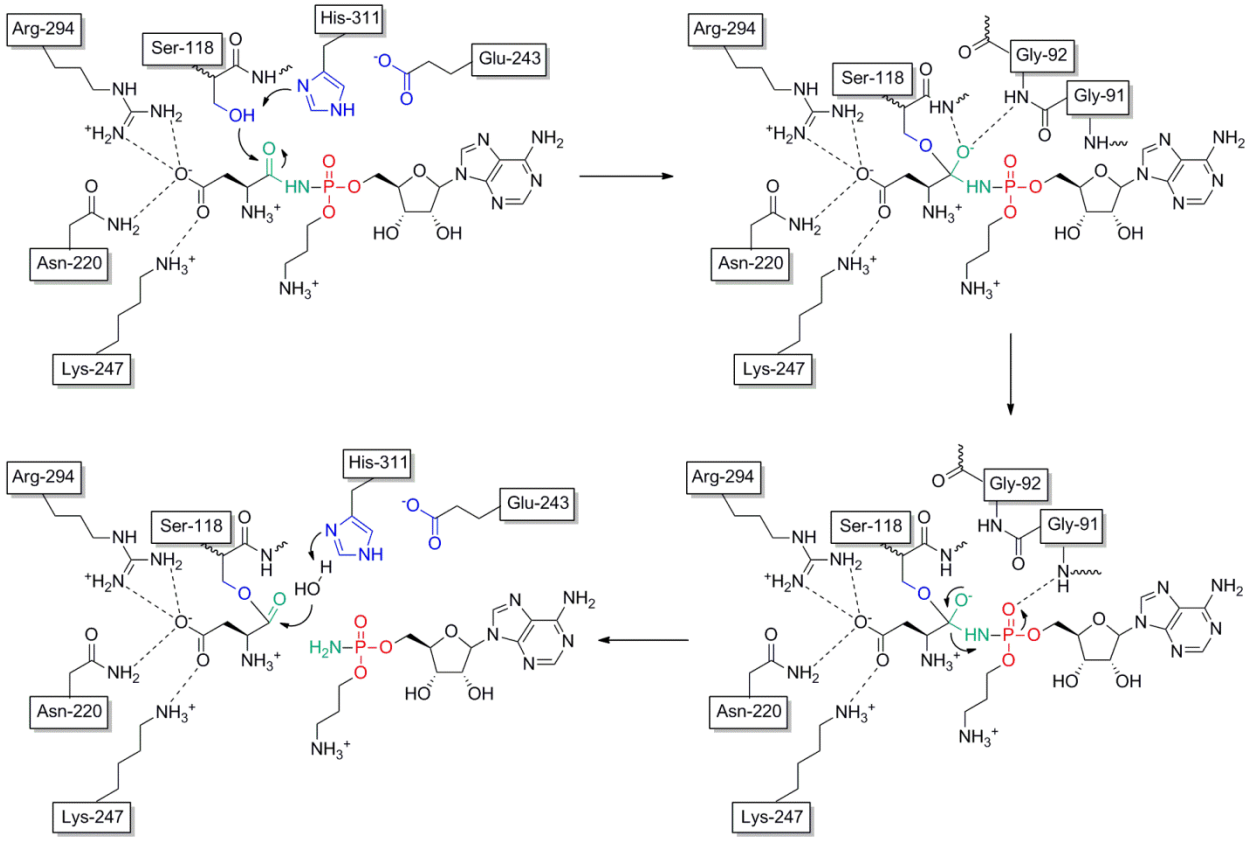

Figure 2.16 Catalytic cycle for the hydrolysis of McC7 by MccF. The catalytic triad residues are shown in blue. The phosphor- moiety of the phosphoramide is shown in red, and the scissile amide bond is shown in green. Putative hydrogen bond interactions are shown as dashes.

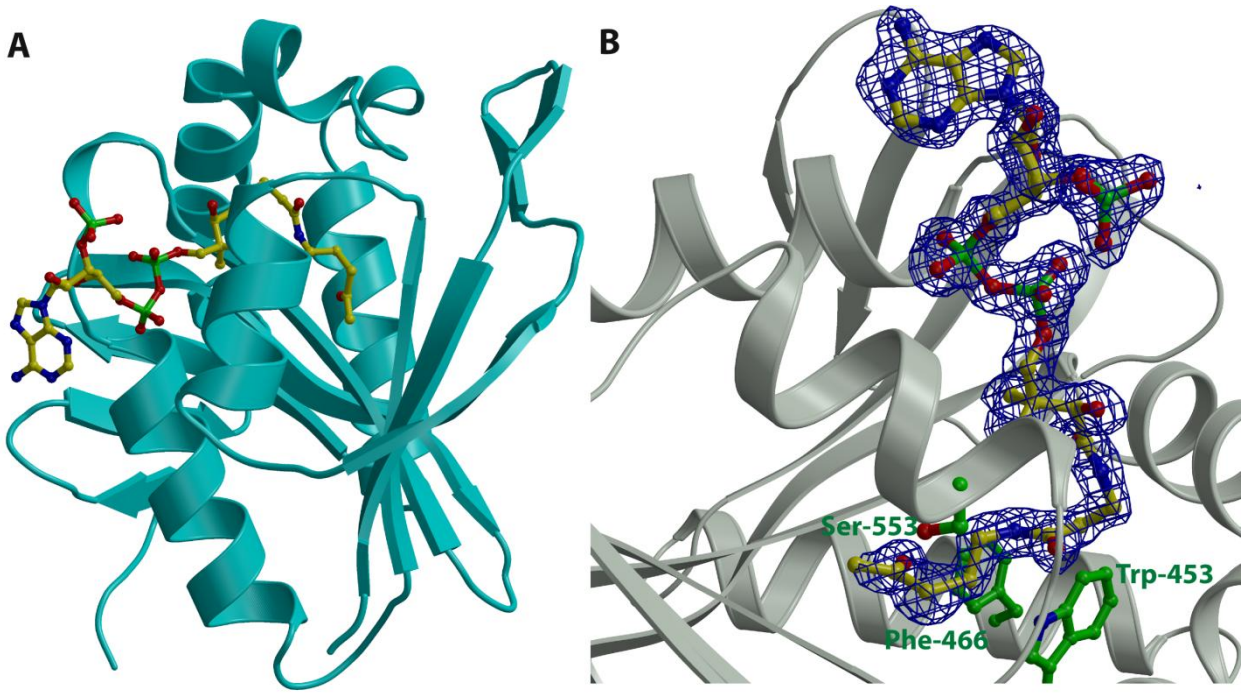

Figure 2.17 (A) Ribbon diagram showing the overall structure of $MccE^{CTD}$ (in cyan) in complex with acetyl CoA (yellow ball-and-stick). **(B)** Close-up view of the acetyl CoA binding site. Superimposed is a difference Fourier electron density map (in blue), calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of acetyl CoA deleted prior to one round of refinement. The map is contoured at 3σ over background.

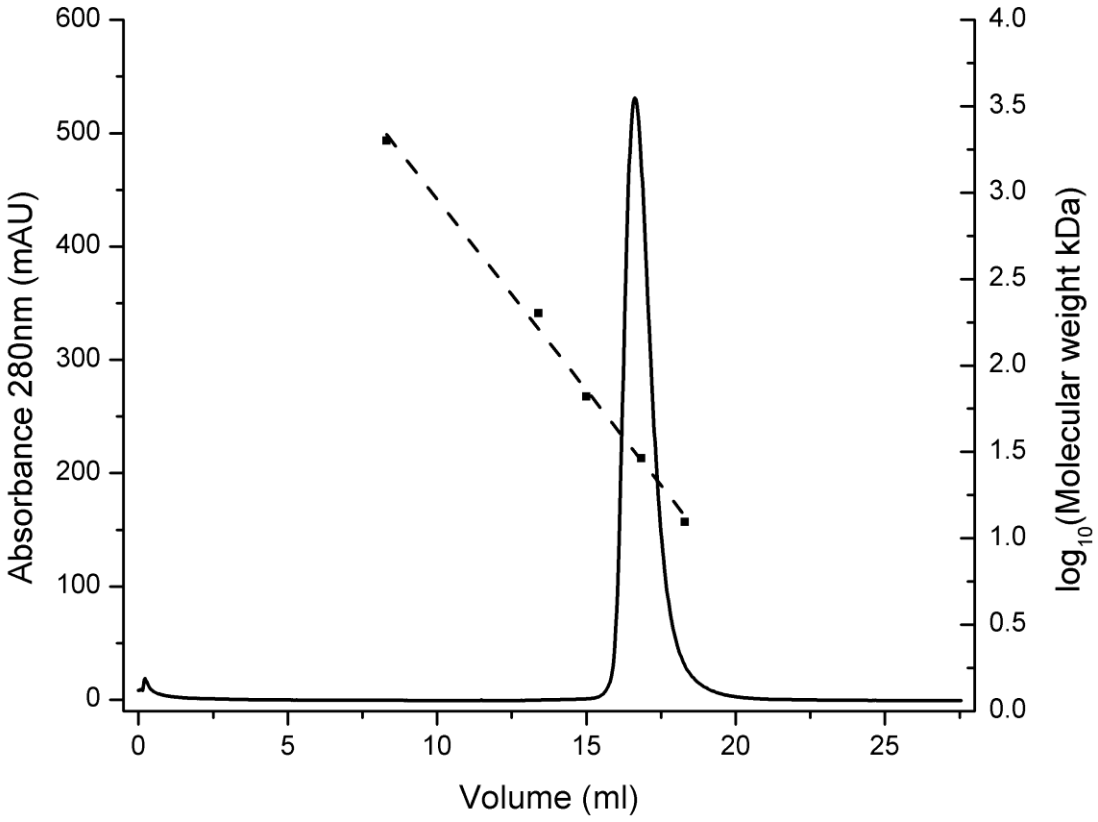

Figure 2.18 Size exclusion chromatography elution profile for MccE^{CTD}. MccE^{CTD} elution profile is shown as a solid line with linear fit (dashed line) of peak elution volumes for molecular weight standard proteins shown as square dots. Chromatography was performed using Superdex 200 HR 10/30 analytical gel filtration column (GE Lifesciences) in buffer 20 mM HEPES-Na (pH 7.5), 100 mM KCl, at a flow rate of 0.5 mL/min. The molecular weight standards were Blue Dextran (~2,000,000 Da), beta-amylase (~200,000 Da), Albumin (~66,200 Da), carbonic anhydrase (~29,000 Da) and cytochrome c (~12,400 Da) and were obtained from Sigma.

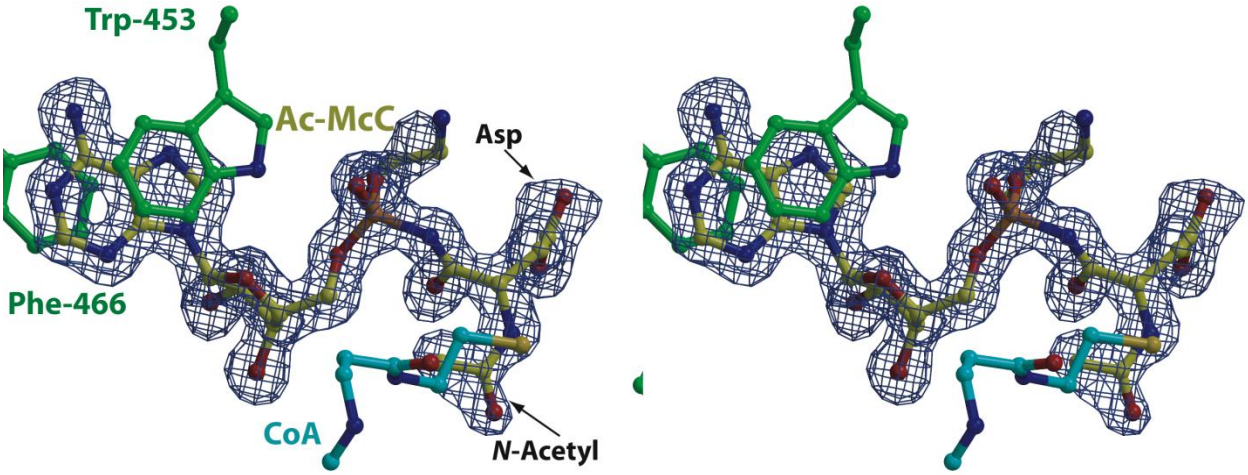

Figure 2.19 Stereoview of the active site of MccE^{CTD} in complex with CoA and acetylated McC7. Superimposed is a difference Fourier electron density map (contoured at 3σ over background and shown in blue) calculated with coefficients $|F_{obs}|$ - $|F_{calc}|$ and phases from the final refined model with the coordinates of substrate deleted prior to one round of refinement. The adenine stacking residues Trp453 and Phe466 are labeled.

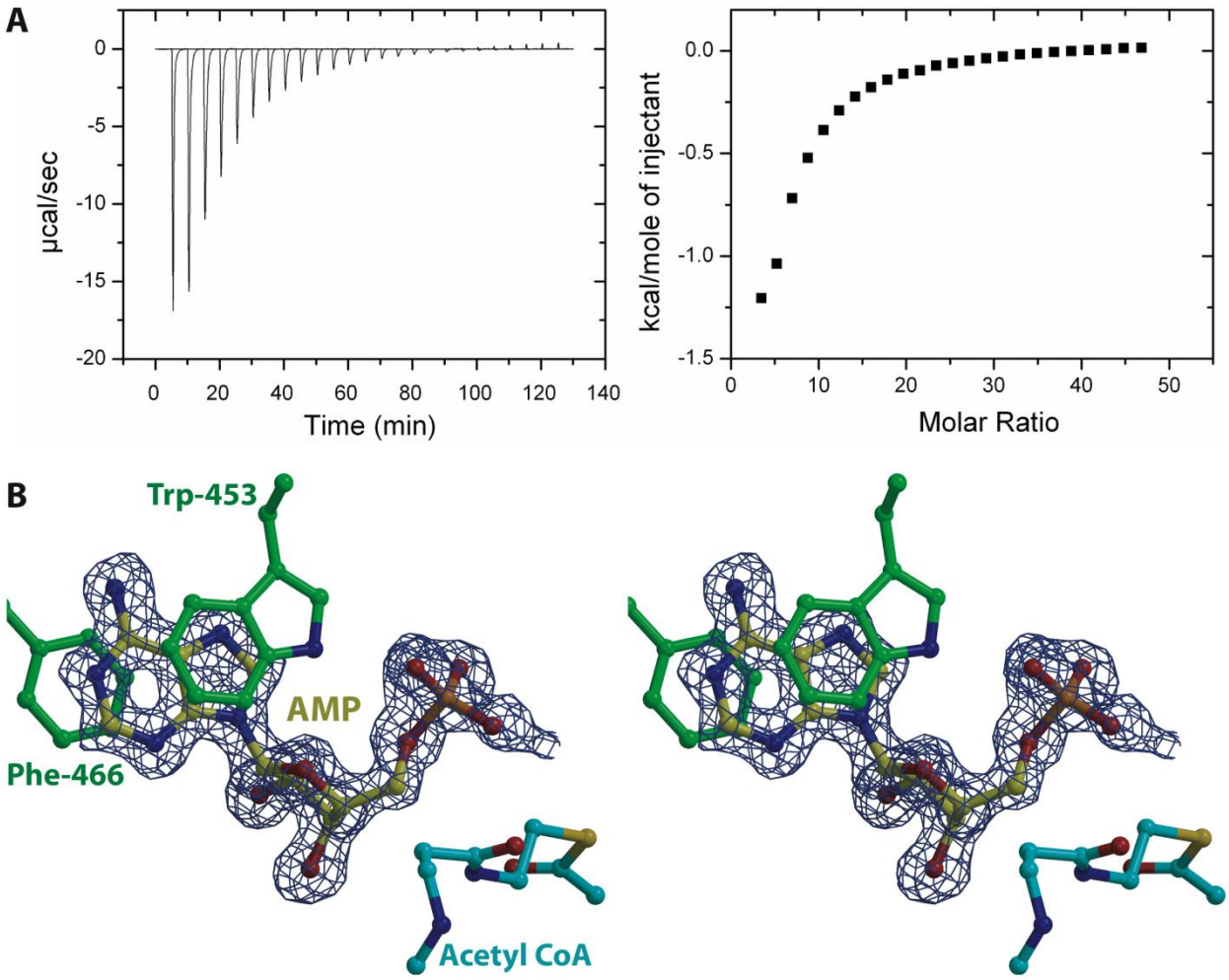

Figure 2.20 ITC and crystallographic demonstration of AMP binding to MccE^{CTD}. (A) Raw data and corresponding binding isotherm for the interaction of MccE^{CTD} with AMP. **(B)** Stereoview of the active site of MccE^{CTD} in complex with acetyl CoA and AMP. Substrate and protein residues are colored and the superimposed difference Fourier electron density map (contoured at 3σ over background and shown in blue) is calculated as in Figure 2.18.

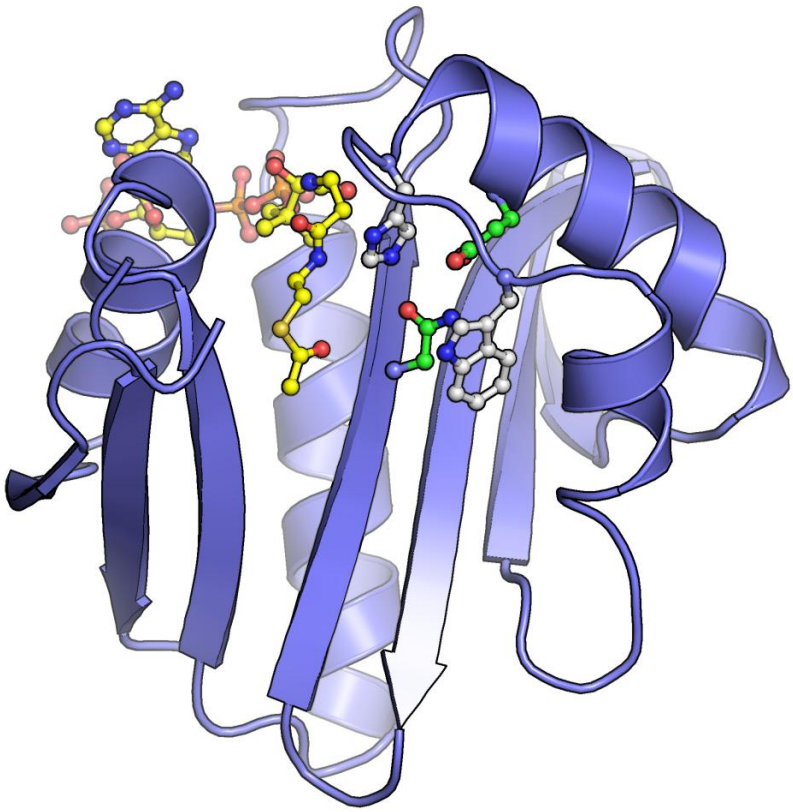

Figure 2.21 Structure of RimI in complex with acetyl CoA. RimI is shown in carton representation with acetyl CoA shown in yellow ball-and-stick. The putative adenine π-stacking amino acid side chains are shown in ball-and-stick with the carbon atoms colored grey. The Glu19 and Asn68 side chains which occlude the adenine binding site are shown in ball-and-stick with the carbon atoms colored green.

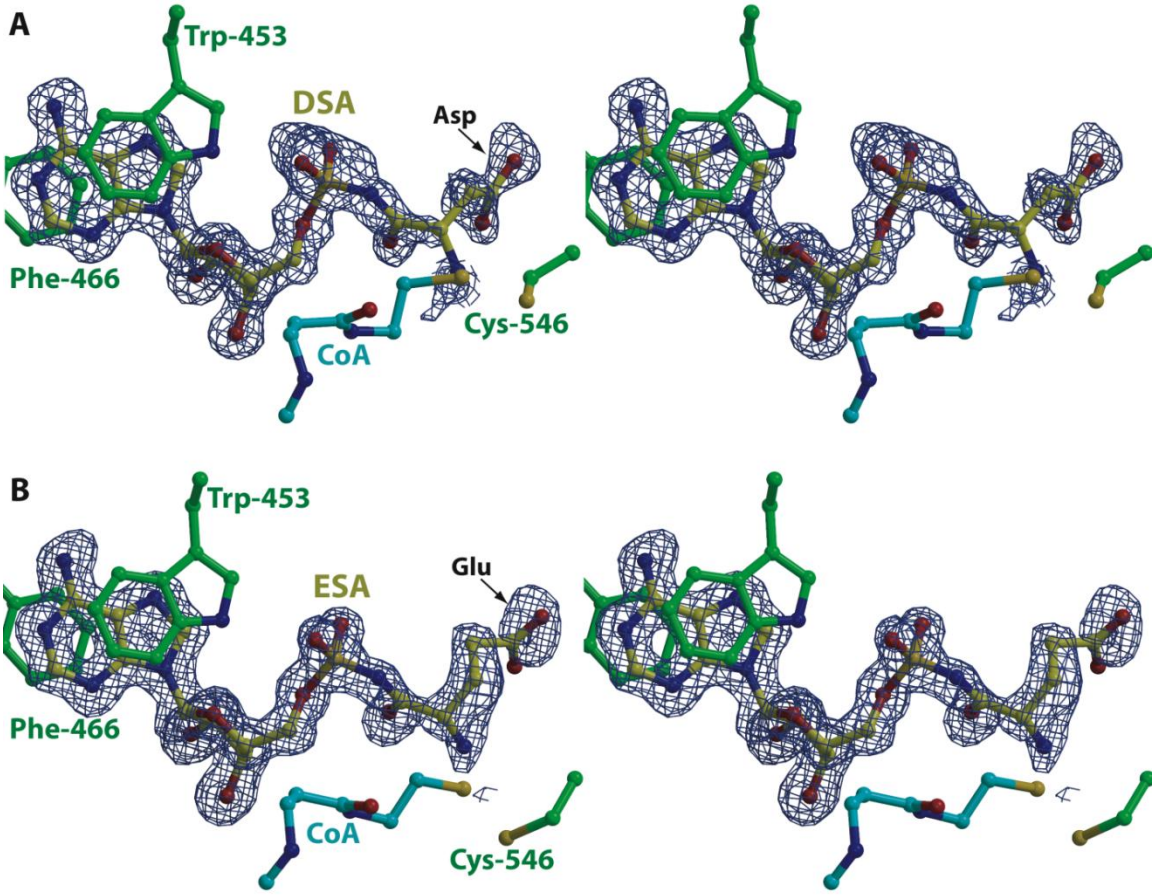

Figure 2.22 Active site of MccE^{CTD} in complex with DSA and ESA. (A) Stereoview of the active site of MccE^{CTD} in complex with CoA and DSA. The DSA substrate is colored in yellow, the CoA molecule is colored in cyan, polypeptide residues colored in green include Trp-453 and Phe-466 that form a π-stacked sandwich with the adenine ring of the substrate and Cys-546 that is disulfide bonded to the CoA thiol. Superimposed is a difference Fourier electron density map (contoured at 3σ over background and shown in blue) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of DSA deleted prior to one round of refinement. **(B)** Stereoview of the active site of MccE^{CTD} in complex with CoA and ESA. Substrate and protein residues are colored as above and the superimposed difference Fourier electron density map (contoured at 3σ over background and shown in blue) is calculated as above.

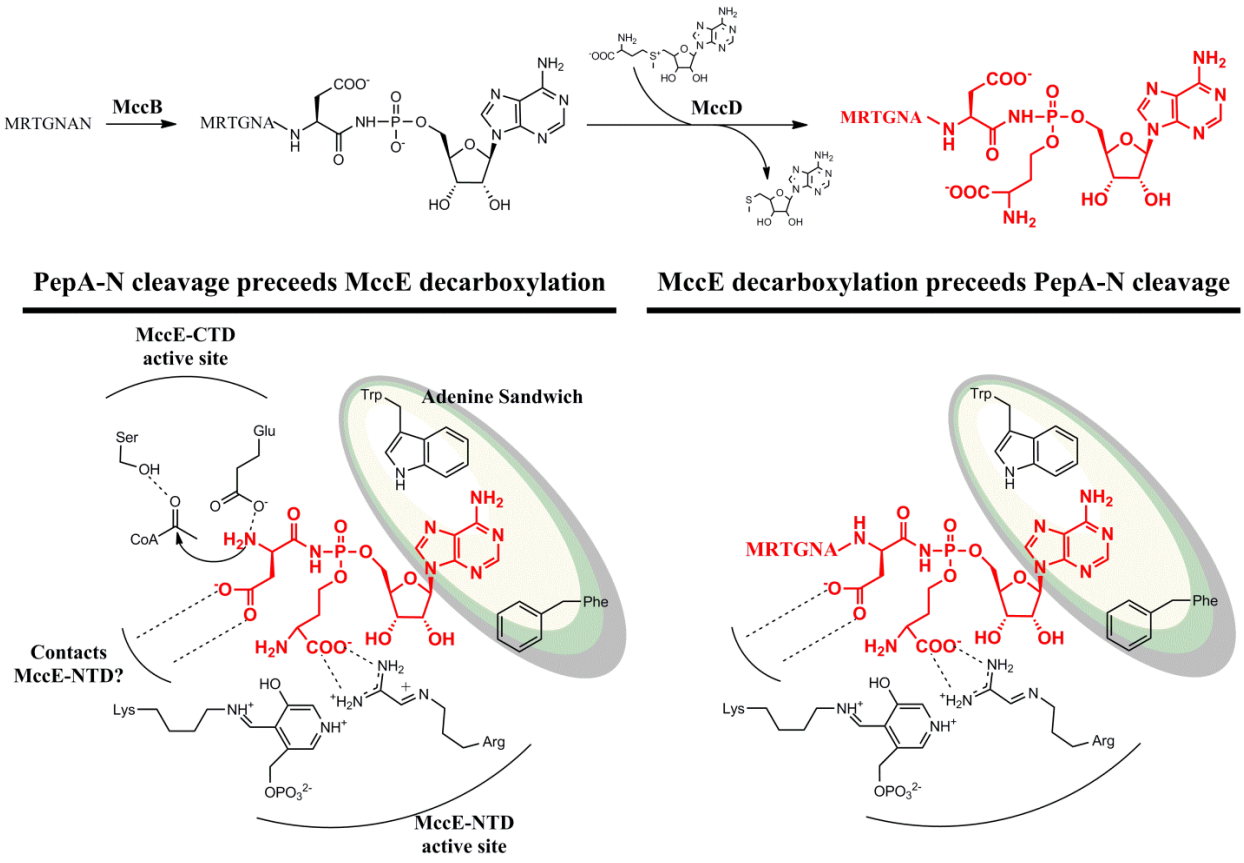

Figure 2.23 Proposed hypothetical model for the concomitant action of the MccE bi-domain protein. We postulate that MccD transfers a 2-ammoniobutanoate moiety from S-adenosyl-methionine to the phosphoramide oxygen atom, and releases methyl-thio-adenosine (top). The 2-aamoniobutanoate moiety is then decarboxylated by MccE^{NTD} to generate McC7 (bottom right). In the event that the non-specific PepA-N cleavage occurs prior to decarboxylation by MccE^{NTD}, subsequent decarboxylation by MccE^{NTD} would generate the toxic processed McC7 within the producer cell. In such a scenario, MccE^{CTD} would concomitantly acetylate the aspartyl primary amine to generate a non-toxic species within the producer cell (bottom right).

References

Adams, P.D., Grosse-Kunstleve, R.W., Hung, L.W., Ioerger, T.R., McCoy, A.J., Moriarty, N.W., Read, R.J., Sacchettini, J.C., Sauter, N.K., and Terwilliger, T.C. (2002). PHENIX: building new software for automated crystallographic structure determination. Acta Crystallogr D Biol Crystallogr *58*, 1948-1954.

Agarwal, V., Metlytskaya, A., Severinov, K.V., and Nair, S.K. (2011). Structural basis for Microcin C7 inactivation by the MccE acetyltransferase. J Biol Chem.

Agarwal, V., and Nair, S.K. (2012). Aminoacyl tRNA synthetases as targets for antibiotic development. Med Chem Comm.

Agarwal, V., Tikhonov, A., Metlitskaya, A., Severinov, K., and Nair, S.K. (2012). Structure and function of a serine carboxypeptidase adapted for degradation of the protein synthesis antibiotic microcin C7. Proc Natl Acad Sci U S A *109*, 4425-4430.

Alley, M.R., Baker, S.J., Beutner, K.R., and Plattner, J. (2007). Recent progress on the topical therapy of onychomycosis. Expert Opin Investig Drugs *16*, 157-167.

Ataide, S.F., and Ibba, M. (2006). Small molecules: big players in the evolution of protein synthesis. ACS Chem Biol *1*, 285-297.

Bachovchin, W.W., Kaiser, R., Richards, J.H., and Roberts, J.D. (1981). Catalytic mechanism of serine proteases: reexamination of the pH dependence of the histidyl 1J13C2-H coupling constant in the catalytic triad of alpha-lytic protease. Proc Natl Acad Sci U S A *78*, 7323-7326.

Baker, S.J., Zhang, Y.K., Akama, T., Lau, A., Zhou, H., Hernandez, V., Mao, W., Alley, M.R., Sanders, V., and Plattner, J.J. (2006). Discovery of a new boron-containing antifungal

agent, 5-fluoro-1,3-dihydro-1-hydroxy-2,1- benzoxaborole (AN2690), for the potential treatment of onychomycosis. J Med Chem *49*, 4447-4450.

Baumann, W.K., Bizzozero, S.A., and Dutler, H. (1970). Specificity of alpha-chymotrypsin. Dipeptide substrates. FEBS Lett *8*, 257-260.

Baumann, W.K., Bizzozero, S.A., and Dutler, H. (1973). Kinetic investigation of the alpha-chymotrypsin-catalyzed hydrolysis of peptide substrates. The relationship between peptide-structure N-terminal to the cleaved bond and reactivity. Eur J Biochem *39*, 381-391.

Boehr, D.D., Farley, A.R., Wright, G.D., and Cox, J.R. (2002). Analysis of the pi-pi stacking interactions between the aminoglycoside antibiotic kinase APH(3')-IIIa and its nucleotide ligands. Chem Biol *9*, 1209-1217.

Boyce, J.M. (2001). MRSA patients: proven methods to treat colonization and infection. J Hosp Infect *48 Suppl A*, S9-14.

Brown, J.R., Gentry, D., Becker, J.A., Ingraham, K., Holmes, D.J., and Stanhope, M.J. (2003). Horizontal transfer of drug-resistant aminoacyl-transfer-RNA synthetases of anthrax and Gram-positive pathogens. EMBO Rep *4*, 692-698.

Brunger, A.T. (1992). Free R value: a novel statistical quantity for assessing the accuracy of crystal structures. Nature *355*, 472-475.

Brunzelle, J.S., Wu, R., Korolev, S.V., Collart, F.R., Joachimiak, A., and Anderson, W.F. (2004). Crystal structure of Bacillus subtilis YdaF protein: a putative ribosomal N-acetyltransferase. Proteins *57*, 850-853.

Circello, B.T., Miller, C.G., Lee, J.H., van der Donk, W.A., and Metcalf, W.W. (2011). The antibiotic dehydrophos is converted to a toxic pyruvate analog by peptide bond cleavage in Salmonella enterica. Antimicrob Agents Chemother *55*, 3357-3362.

Cursino, L., Smajs, D., Smarda, J., Nardi, R.M., Nicoli, J.R., Chartone-Souza, E., and Nascimento, A.M. (2006). Exoproducts of the Escherichia coli strain H22 inhibiting some enteric pathogens both in vitro and in vivo. J Appl Microbiol *100*, 821-829.

Dougherty, D.A. (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. Science *271*, 163-168.

Drawz, S.M., and Bonomo, R.A. (2010). Three decades of beta-lactamase inhibitors. Clin Microbiol Rev *23*, 160-201.

Dreyfus, C., Lemaire, D., Mari, S., Pignol, D., and Arnoux, P. (2009). Crystallographic snapshots of iterative substrate translocations during nicotianamine synthesis in Archaea. Proc Natl Acad Sci U S A *106*, 16180-16184.

Dyda, F., Klein, D.C., and Hickman, A.B. (2000). GCN5-related N-acetyltransferases: a structural overview. Annu Rev Biophys Biomol Struct *29*, 81-103.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Eswarappa, S.M., Panguluri, K.K., Hensel, M., and Chakravortty, D. (2008). The yejABEF operon of Salmonella confers resistance to antimicrobial peptides and contributes to its virulence. Microbiology *154*, 666-678.

Fechter, P., and Brownlee, G.G. (2005). Recognition of mRNA cap structures by viral and cellular proteins. J Gen Virol *86*, 1239-1249.

Finking, R., Neumuller, A., Solsbacher, J., Konz, D., Kretzschmar, G., Schweitzer, M., Krumm, T., and Marahiel, M.A. (2003). Aminoacyl adenylate substrate analogues for the inhibition of adenylation domains of nonribosomal peptide synthetases. Chembiochem *4*, 903-906.

Fredenhagen, A., Angst, C., and Peter, H.H. (1995). Digestion of rhizocticins to (Z)-L-2-amino-5-phosphono-3-pentenoic acid: revision of the absolute configuration of plumbemycins A and B. J Antibiot (Tokyo) *48*, 1043-1045.

Fremlin, L., Farrugia, M., Piggott, A.M., Khalil, Z., Lacey, E., and Capon, R.J. (2011). Reveromycins revealed: new polyketide spiroketals from Australian marine-derived and terrestrial Streptomyces spp. A case of natural products vs. artifacts. Org Biomol Chem *9*, 1201-1211.

Gamcsik, M.P., Ludeman, S.M., Shulman-Roskes, E.M., McLennan, I.J., Colvin, M.E., and Colvin, O.M. (1993). Protonation of phosphoramide mustard and other phosphoramides. J Med Chem *36*, 3636-3645.

Garcia-Bustos, J.F., Pezzi, N., and Asensio, C. (1984). Microcin 7: purification and properties. Biochem Biophys Res Commun *119*, 779-785.

Garcia-Bustos, J.F., Pezzi, N., and Mendez, E. (1985). Structure and mode of action of microcin 7, an antibacterial peptide produced by Escherichia coli. Antimicrob Agents Chemother *27*, 791-797.

Gentry, D.R., Ingraham, K.A., Stanhope, M.J., Rittenhouse, S., Jarvest, R.L., O'Hanlon, P.J., Brown, J.R., and Holmes, D.J. (2003). Variable sensitivity to bacterial methionyl-tRNA synthetase inhibitors reveals subpopulations of Streptococcus pneumoniae with two distinct methionyl-tRNA synthetase genes. Antimicrob Agents Chemother *47*, 1784-1789.

Gokulan, K., Rupp, B., Pavelka, M.S., Jr., Jacobs, W.R., Jr., and Sacchettini, J.C. (2003). Crystal structure of Mycobacterium tuberculosis diaminopimelate decarboxylase, an essential enzyme in bacterial lysine biosynthesis. J Biol Chem *278*, 18588-18596.

Gonzalez-Pastor, J.E., San Millan, J.L., Castilla, M.A., and Moreno, F. (1995). Structure and organization of plasmid genes required to produce the translation inhibitor microcin C7. J Bacteriol *177*, 7131-7140.

Gonzalez-Pastor, J.E., San Millan, J.L., and Moreno, F. (1994). The smallest known gene. Nature *369*, 281.

Grosse-Kunstleve, R.W., and Adams, P.D. (2003). Substructure search procedures for macromolecular structures. Acta Crystallogr D Biol Crystallogr *59*, 1966-1973.

Guijarro, J.I., Gonzalez-Pastor, J.E., Baleux, F., San Millan, J.L., Castilla, M.A., Rico, M., Moreno, F., and Delepierre, M. (1995). Chemical structure and translation inhibition studies of the antibiotic microcin C7. J Biol Chem *270*, 23520-23532.

Guilligay, D., Tarendeau, F., Resa-Infante, P., Coloma, R., Crepin, T., Sehr, P., Lewis, J., Ruigrok, R.W., Ortin, J., Hart, D.J., *et al.* (2008). The structural basis for cap binding by influenza virus polymerase subunit PB2. Nat Struct Mol Biol *15*, 500-506.

Gutierrez-Lugo, M.T., and Bewley, C.A. (2011). Susceptibility and mode of binding of the Mycobacterium tuberculosis cysteinyl transferase mycothiol ligase to tRNA synthetase inhibitors. Bioorg Med Chem Lett *21*, 2480-2483.

Hamano, Y., Hoshino, Y., Nakamori, S., and Takagi, H. (2004). Overexpression and characterization of an aminoglycoside 6'-N-acetyltransferase with broad specificity from an epsilon-poly-L-lysine producer, Streptomyces albulus IFO14147. J Biochem *136*, 517-524.

Heacock, D., Forsyth, C.J., Shiba, K., and Musier-Forsyth, K. (1996). Synthesis and aminoacyl-tRNA synthetase inhibitory of prolyl adenylate analogs. Bioorg Chem *24*, 273-289.

Hedstrom, L. (2002). Serine protease mechanism and specificity. Chem Rev *102*, 4501-4524.

Hodel, A.E., Gershon, P.D., Shi, X., and Quiocho, F.A. (1996). The 1.85 A structure of vaccinia protein VP39: a bifunctional enzyme that participates in the modification of both mRNA ends. Cell *85*, 247-256.

Hugonnet, J.E., Tremblay, L.W., Boshoff, H.I., Barry, C.E., 3rd, and Blanchard, J.S. (2009). Meropenem-clavulanate is effective against extensively drug-resistant Mycobacterium tuberculosis. Science *323*, 1215-1218.

Hurdle, J.G., O'Neill, A.J., and Chopra, I. (2005). Prospects for aminoacyl-tRNA synthetase inhibitors as new antimicrobial agents. Antimicrob Agents Chemother *49*, 4821-4833.

Ibba, M., and Soll, D. (2000). Aminoacyl-tRNA synthesis. Annu Rev Biochem *69*, 617-650.

Kazakov, T., Vondenhoff, G.H., Datsenko, K.A., Novikova, M., Metlitskaya, A., Wanner, B.L., and Severinov, K. (2008). Escherichia coli peptidase A, B, or N can process translation inhibitor microcin C. J Bacteriol *190*, 2607-2610.

Kim, S., Lee, S.W., Choi, E.C., and Choi, S.Y. (2003). Aminoacyl-tRNA synthetases and their inhibitors as a novel family of antibiotics. Appl Microbiol Biotechnol *61*, 278-288.

Kitabatake, M., Ali, K., Demain, A., Sakamoto, K., Yokoyama, S., and Soll, D. (2002). Indolmycin resistance of Streptomyces coelicolor A3(2) by induced expression of one of its two tryptophanyl-tRNA synthetases. J Biol Chem *277*, 23882-23887.

Kleywegt, G.J., and Brunger, A.T. (1996). Checking your imagination: applications of the free R value. Structure *4*, 897-904.

Korza, H.J., and Bochtler, M. (2005). Pseudomonas aeruginosa LD-carboxypeptidase, a serine peptidase with a Ser-His-Glu triad and a nucleophilic elbow. J Biol Chem *280*, 40802-40812.

Kugler, M., Loeffler, W., Rapp, C., Kern, A., and Jung, G. (1990). Rhizocticin A, an antifungal phosphono-oligopeptide of Bacillus subtilis ATCC 6633: biological properties. Arch Microbiol *153*, 276-281.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR *8*, 477-486.

Leis, J.P., and Cameron, C.E. (1994). Engineering proteases with altered specificity. Curr Opin Biotechnol *5*, 403-408.

Ling, J., Reynolds, N., and Ibba, M. (2009). Aminoacyl-tRNA synthesis and translational quality control. Annu Rev Microbiol *63*, 61-78.

M. Ibba, M.F., S. Cusack (2005). (Georgetown,TX, Landes Bioscience).

Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1997). Cocrystal structure of the messenger RNA 5' cap-binding protein (eIF4E) bound to 7-methyl-GDP. Cell *89*, 951-961.

May, J.J., Finking, R., Wiegeshoff, F., Weber, T.T., Bandur, N., Koert, U., and Marahiel, M.A. (2005). Inhibition of the D-alanine:D-alanyl carrier protein ligase from Bacillus subtilis increases the bacterium's susceptibility to antibiotics that target the cell wall. FEBS J *272*, 2993-3003.

Mazza, C., Segref, A., Mattaj, I.W., and Cusack, S. (2002). Large-scale induced fit recognition of an m(7)GpppG cap analogue by the human nuclear cap-binding complex. Embo J *21*, 5548-5557.

McCoy, A.J. (2007). Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr D Biol Crystallogr *63*, 32-41.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J Appl Crystallogr *40*, 658-674.

McRee, D.E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. J Struct Biol *125*, 156-165.

Metlitskaya, A., Kazakov, T., Kommer, A., Pavlova, O., Praetorius-Ibba, M., Ibba, M., Krasheninnikov, I., Kolb, V., Khmel, I., and Severinov, K. (2006). Aspartyl-tRNA synthetase is the target of peptide nucleotide antibiotic Microcin C. J Biol Chem *281*, 18033-18042.

Metlitskaya, A., Kazakov, T., Vondenhoff, G.H., Novikova, M., Shashkov, A., Zatsepin, T., Semenova, E., Zaitseva, N., Ramensky, V., Van Aerschot, A., *et al.* (2009). Maturation of the translation inhibitor microcin C. J Bacteriol *191*, 2380-2387.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr D Biol Crystallogr *55*, 247-255.

Nocek, B., Tikhonov, A., Babnigg, G., Gu, M., Zhou, M., Makarova, K.S., Vondenhoff, G., Aerschot, A.V., Kwon, K., Anderson, W.F., *et al.* (2012). Structural and Functional Characterization of Microcin C Resistance Peptidase MccF from Bacillus anthracis. J Mol Biol.

Norris, A.L., Ozen, C., and Serpersu, E.H. (2010). Thermodynamics and kinetics of association of antibiotics with the aminoglycoside acetyltransferase (3)-IIIb, a resistance-causing enzyme. Biochemistry *49*, 4027-4035.

Novikova, M., Kazakov, T., Vondenhoff, G.H., Semenova, E., Rozenski, J., Metlytskaya, A., Zukher, I., Tikhonov, A., Van Aerschot, A., and Severinov, K. (2010). MccE provides

resistance to protein synthesis inhibitor microcin C by acetylating the processed form of the antibiotic. J Biol Chem *285*, 12662-12669.

Novikova, M., Metlitskaya, A., Datsenko, K., Kazakov, T., Kazakov, A., Wanner, B., and Severinov, K. (2007). The Escherichia coli Yej transporter is required for the uptake of translation inhibitor microcin C. J Bacteriol *189*, 8361-8365.

Osada, H., and Isono, K. (1985). Mechanism of action and selective toxicity of ascamycin, a nucleoside antibiotic. Antimicrob Agents Chemother *27*, 230-233.

Otoguro, K., Ui, H., Ishiyama, A., Kobayashi, M., Togashi, H., Takahashi, Y., Masuma, R., Tanaka, H., Tomoda, H., Yamada, H., *et al.* (2003). In vitro and in vivo antimalarial activities of a non-glycosidic 18-membered macrolide antibiotic, borrelidin, against drug-resistant strains of Plasmodia. J Antibiot (Tokyo) *56*, 727-729.

Otwinowski, Z., Borek, D., Majewski, W., and Minor, W. (2003). Multiparametric scaling of diffraction intensities. Acta Crystallogr A *59*, 228-234.

Payne, D.J., Gwynn, M.N., Holmes, D.J., and Pompliano, D.L. (2007). Drugs for bad bugs: confronting the challenges of antibacterial discovery. Nat Rev Drug Discov *6*, 29-40.

Perona, J.J., and Craik, C.S. (1995). Structural basis of substrate specificity in the serine proteases. Protein Sci *4*, 337-360.

Perrakis, A., Sixma, T.K., Wilson, K.S., and Lamzin, V.S. (1997). wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. Acta Crystallogr D Biol Crystallogr *53*, 448-455.

Pohlmann, J., and Brotz-Oesterhelt, H. (2004). New aminoacyl-tRNA synthetase inhibitors as antibacterial agents. Curr Drug Targets Infect Disord *4*, 261-272.

Pramanik, A., Stroeher, U.H., Krejci, J., Standish, A.J., Bohn, E., Paton, J.C., Autenrieth, I.B., and Braun, V. (2007). Albomycin is an effective antibiotic, as exemplified with Yersinia enterocolitica and Streptococcus pneumoniae. Int J Med Microbiol *297*, 459-469.

Qiao, C., Gupte, A., Boshoff, H.I., Wilson, D.J., Bennett, E.M., Somu, R.V., Barry, C.E., 3rd, and Aldrich, C.C. (2007). 5'-O-[(N-acyl)sulfamoyl]adenosines as antitubercular agents that inhibit MbtA: an adenylation enzyme required for siderophore biosynthesis of the mycobactins. J Med Chem *50*, 6080-6094.

Qimron, U., Madar, N., Mittrucker, H.W., Zilka, A., Yosef, I., Bloushtain, N., Kaufmann, S.H., Rosenshine, I., Apte, R.N., and Porgador, A. (2004). Identification of Salmonella typhimurium genes responsible for interference with peptide presentation on MHC class I molecules: Deltayej Salmonella mutants induce superior CD8+ T-cell responses. Cell Microbiol *6*, 1057-1070.

Raczniak, G., Ibba, M., and Soll, D. (2001). Genomics-based identification of targets in pathogenic bacteria for potential therapeutic and diagnostic use. Toxicology *160*, 181-189.

Reader, J.S., Ordoukhanian, P.T., Kim, J.G., de Crecy-Lagard, V., Hwang, I., Farrand, S., and Schimmel, P. (2005). Major biocontrol of plant tumors targets tRNA synthetase. Science *309*, 1533.

Reddy, V.M., Einck, L., and Nacy, C.A. (2008). In vitro antimycobacterial activities of capuramycin analogues. Antimicrob Agents Chemother *52*, 719-721.

Robert Chênevert, S.B., Jacques Lap (2003). Inhibitors of Aminoacyl-tRNA Synthetases as Antibiotics and Tools for Structural and Mechanistic Studies.

Rock, F.L., Mao, W., Yaremchuk, A., Tukalo, M., Crepin, T., Zhou, H., Zhang, Y.K., Hernandez, V., Akama, T., Baker, S.J., *et al.* (2007). An antifungal agent inhibits an aminoacyl-tRNA synthetase by trapping tRNA in the editing site. Science *316*, 1759-1761.

Roje, S. (2006). S-Adenosyl-L-methionine: beyond the universal methyl group donor. Phytochemistry *67*, 1686-1698.

Roush, R.F., Nolan, E.M., Lohr, F., and Walsh, C.T. (2008). Maturation of an Escherichia coli ribosomal peptide antibiotic by ATP-consuming N-P bond formation in microcin C7. J Am Chem Soc *130*, 3603-3609.

Sachdev, G.P., and Fruton, J.S. (1975). Kinetics of action of pepsin on fluorescent peptide substrates. Proc Natl Acad Sci U S A *72*, 3424-3427.

Schimmel, P., Tao, J., and Hill, J. (1998). Aminoacyl tRNA synthetases as targets for new anti-infectives. Faseb J *12*, 1599-1609.

Schramm, V.L. (2007). Enzymatic transition state theory and transition state analogue design. J Biol Chem *282*, 28297-28300.

Severinov, K., Semenova, E., Kazakov, A., Kazakov, T., and Gelfand, M.S. (2007). Low-molecular-weight post-translationally modified microcins. Mol Microbiol *65*, 1380-1394.

Shuman, S., and Lima, C.D. (2004). The polynucleotide ligase and RNA capping enzyme superfamily of covalent nucleotidyltransferases. Curr Opin Struct Biol *14*, 757-764.

Stefanska, A.L., Fulston, M., Houge-Frydrych, C.S., Jones, J.J., and Warr, S.R. (2000). A potent seryl tRNA synthetase inhibitor SB-217452 isolated from a Streptomyces species. J Antibiot (Tokyo) *53*, 1346-1353.

Tao, J., and Schimmel, P. (2000). Inhibitors of aminoacyl-tRNA synthetases as novel anti-infectives. Expert Opin Investig Drugs *9*, 1767-1775.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. Acta Crystallogr D Biol Crystallogr *64*, 61-69.

Thomas, C.M., Hothersall, J., Willis, C.L., and Simpson, T.J. (2010). Resistance to and synthesis of the antibiotic mupirocin. Nat Rev Microbiol *8*, 281-289.

Thompson, R.C., and Blout, E.R. (1973). Dependence of the kinetic parameters for elastase-catalyzed amide hydrolysis on the length of peptide substrates. Biochemistry *12*, 57-65.

Tikhonov, A., Kazakov, T., Semenova, E., Serebryakova, M., Vondenhoff, G., Van Aerschot, A., Reader, J.S., Govorun, V.M., and Severinov, K. (2010). The mechanism of microcin C resistance provided by the MccF peptidase. J Biol Chem *285*, 37944-37952.

Tremblay, L.W., Fan, F., Vetting, M.W., and Blanchard, J.S. (2008). The 1.6 A crystal structure of Mycobacterium smegmatis MshC: the penultimate enzyme in the mycothiol biosynthetic pathway. Biochemistry *47*, 13326-13335.

Ubukata, M., Osada, H., and Isono, K. (1985). Synthesis and biological activity of nucleoside antibiotics, ascamycin and its amino acid analogs. Nucleic Acids Symp Ser, 81-83.

Van de Vijver, P., Vondenhoff, G.H., Denivelle, S., Rozenski, J., Verhaegen, J., Van Aerschot, A., and Herdewijn, P. (2009a). Antibacterial 5'-O-(N-dipeptidyl)-sulfamoyladenosines. Bioorg Med Chem *17*, 260-269.

Van de Vijver, P., Vondenhoff, G.H., Kazakov, T.S., Semenova, E., Kuznedelov, K., Metlitskaya, A., Van Aerschot, A., and Severinov, K. (2009b). Synthetic microcin C analogs targeting different aminoacyl-tRNA synthetases. J Bacteriol *191*, 6273-6280.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol *229*, 105-124.

Vecchione, J.J., and Sello, J.K. (2009). A novel tryptophanyl-tRNA synthetase gene confers high-level resistance to indolmycin. Antimicrob Agents Chemother *53*, 3972-3980.

Vetting, M.W., Bareich, D.C., Yu, M., and Blanchard, J.S. (2008). Crystal structure of RimI from Salmonella typhimurium LT2, the GNAT responsible for N(alpha)-acetylation of ribosomal protein S18. Protein Sci *17*, 1781-1790.

Vetting, M.W., de Carvalho, L.P., Roderick, S.L., and Blanchard, J.S. (2005a). A novel dimeric structure of the RimL Nalpha-acetyltransferase from Salmonella typhimurium. J Biol Chem *280*, 22108-22114.

Vetting, M.W., de Carvalho, L.P., Yu, M., Hegde, S.S., Magnet, S., Roderick, S.L., and Blanchard, J.S. (2005b). Structure and functions of the GNAT superfamily of acetyltransferases. Arch Biochem Biophys *433*, 212-226.

Vetting, M.W., deCarvalho, L.P., Roderick, S.L., and Blanchard, J.S. (2005c). A novel dimeric structure of the RimL Nalpha-acetyltransferase from Salmonella typhimurium. J Biol Chem *280*, 22108-22114.

Vondenhoff, G.H., Blanchaert, B., Geboers, S., Kazakov, T., Datsenko, K.A., Wanner, B.L., Rozenski, J., Severinov, K., and Van Aerschot, A. (2011). Characterization of peptide chain length and constituency requirements for YejABEF-mediated uptake of microcin C analogues. J Bacteriol *193*, 3618-3623.

Vondenhoff, G.H., and Van Aerschot, A. (2011). Aminoacyl-tRNA synthetase inhibitors as potential antibiotics. Eur J Med Chem *46*, 5227-5236.

Winn, M., Goss, R.J., Kimura, K., and Bugg, T.D. (2010). Antimicrobial nucleoside antibiotics targeting cell wall assembly: recent advances in structure-function studies and nucleoside biosynthesis. Nat Prod Rep *27*, 279-304.

Yang, C.H., Azad, H.R., and Cooksey, D.A. (1996). A chromosomal locus required for copper resistance, competitive fitness, and cytochrome c biogenesis in Pseudomonas fluorescens. Proc Natl Acad Sci U S A *93*, 7315-7320.

Yang, Z., Savchenko, A., Yakunin, A., Zhang, R., Edwards, A., Arrowsmith, C., and Tong, L. (2003). Aspartate dehydrogenase, a novel enzyme identified from structural and functional studies of TM1643. J Biol Chem *278*, 8804-8808.

Zeng, Y., Roy, H., Patil, P.B., Ibba, M., and Chen, S. (2009). Characterization of two seryl-tRNA synthetases in albomycin-producing Streptomyces sp. strain ATCC 700974. Antimicrob Agents Chemother *53*, 4619-4627.

Zhang, D., and Miller, M.J. (1999). Polyoxins and nikkomycins: progress in synthetic and biological studies. Curr Pharm Des *5*, 73-99.

Chapter 3: Structural and biochemical characterization of enzymes of a novel phosphonate degradation pathway provide missing links for the biosynthesis and degradation of phosphonoacetate[1,2]

Abstract

Phosphonates are ubiquitous natural organic compounds characterized by the presence of a carbon-phosphorus (C-P) bond that is resistant to enzymatic and chemical degradation commonly afforded upon oxygen-phosphorus (O-P) bond-containing phosphate esters. Due to the stability of the phosphonate bond, natural and synthetic phosphonates have commercial application as antibiotics, herbicides and detergents. Several microbial species have evolved dedicated enzymatic machineries for the hydrolytic degradation of the phosphonate molecules for their utilization as carbon and phosphorus sources. Previous investigations have revealed that phosphonoacetaldehyde (PnAA) and phosphonoacetate (PnA) can be enzymatically hydrolyzed to generate phosphate and either acetaldehyde or acetate, respectively. A recently discovered PnAA oxidase gene in *Sinorhizobium meliloti* 1021 links these two hydrolytic processes and provides a previously unknown catabolic mechanism for PnA production in the microbial metabolome. Central to this study was the discovery of a previously unknown PnAA dehydrogenase enzyme, in addition to a PnA specific metal dependent hydrolase. Here, we carry

---

[1]Parts of this chapter have been reproduced from:

Agarwal V, Borisova SA, Metcalf WW, van der Donk WA, Nair SK. Chem Biol. 2011 Oct 28;18(10)1230-40. Referred to as 'Agarwal et.al., 2011' in this chapter. Reproduced with the permission of the publisher and is available from http://www.cell.com/ and using DOI: 10.1016/j.chembiol.2011.07.019.

[2]Parts of this chapter are in preparation for publication:

Agarwal V, Chen JH, Chekan J, Borisova SA, Metcalf WW, van der Donk WA, Nair SK. Structural and functional characterization of phosphonoacetaldehyde dehydrogenase enzyme provides missing link for phosphonoacetate degradation, 2012, in preparation for publication.

out full biochemical and structural biological characterization of this PnAA dehydrogenase enzyme - PhnY. We present crystal structures of PhnY in the apo state, as well as complexes with substrate, cofactor and product. Kinetic characterization of structure-based active site mutants demonstrates how a highly conserved enzyme active site has been modified by Nature to generate new metabolic activities. We also present detailed biochemical and structural characterization of the PnA hydrolase enzyme- PhnA that provides insights into the mechanism of C-P bond cleavage. The 1.35 Å resolution crystal structure reveals a catalytic core similar to those of alkaline phosphatases (AP) and nucleotide pyrophosphatases (NPP) but with notable differences, such as a longer metal-metal distance. Detailed structure-guided analysis of active site residues and four additional cocrystal structures with PnA substrate, acetate, phosphonoformate (PnF) inhibitor, and a covalently bound transition state mimic provide insight into active site features that may facilitate cleavage of the C-P bond. These studies expand upon the array of reactions that can be catalyzed by enzymes of the AP superfamily.

Introduction

Natural occurrence and application of phosphonate molecules

Phosphonates are organic molecules characterized by the presence of C-P bond in place of the O-P bond found in phosphate esters. Consequently, the oxidation state of the phosphorus atom in phosphonate is +3, as compared to +5 found in phosphate and phosphate esters. However, the tetrahedral geometry around the phosphorus atom is preserved in phosphonates, and, in combination with their remarkable recalcitrance towards chemical and enzymatic hydrolysis, renders phosphonates as competitive inhibitors of biochemical processes that involve phosphoryl group transfer reactions (Metcalf and van der Donk, 2009). Nature has utilized this chemical mimicry in the development of several compounds which have potent antibiotic, herbicidal, and other bio-inhibitory activities. One class of such compounds includes the rhizocticins and plumbemycins that utilize the unnatural amino acid- (*Z*)-L-2-amino-5-phosphono-3-pentenoic acid to competitively inhibit the β-γ replacement of phosphohomoserine catalyzed by threonine synthase (Borisova et al., 2010). The antimalarial phosphonate molecule FR900098 acts as a competitive inhibitor of 1-deoxy-D-xylulose 5-phosphate reductoisomerase and stalls isoprenoid production in *Plasmodium falciparum* (Zingle et al., 2010). Phosphonate molecules, such as glyphosate find extensive application in agriculture as an herbicide. Glyphosate inhibits the essential enzyme- 5-enolpyruvateshikimate-3-phosphate synthase (EPSPS) (Steinrucken and Amrhein, 1984). X-ray crystallographic studies of EPSPS in complex with glyphosate show that glyphosate functions by occupying the binding site of the phosphoenolpyruvate, mimicking an intermediate state of the ternary enzyme substrate complex (Schonbrunn et al., 2001). Chemical structures of some naturally occurring phosphonate molecules and the enzyme reaction intermediates which they mimic are shown in Figure 3.1.

Several synthetic phosphonate molecules find applications in detergents, fuels, water purification, and in pulp and paper manufacture, among other uses.

Phosphonate compounds also occur in nature in the form of phosphonylated biological macromolecules such as lipids, sugars and glycoproteins. Their presence in the biosphere is exemplified by estimates that up to 30% of the available phosphorus in the oceans exist in the form of phosphonic acids (Clark et al., 1999). In light of the fact that phosphorus is often the limiting growth nutrient in the natural environment, it is not surprising that microbes have evolved metabolic pathways for the hydrolysis of phosphonate compounds for their subsequent utilization as sources for phosphorus, and carbon.

Enzymatic routes for biodegradation of phosphonate molecules

Until recently, only four distinct pathways for the degradation of phosphonate compounds had been characterized (Figure 3.2A). The most widespread and substrate promiscuous pathway among these is the C-P lyase pathway, which utilizes a membrane associated multi-enzyme complex to catalyze phosphonate bond hydrolysis using [Fe-S]-based radical chemistry (Kamat et al., 2011; Shames et al., 1987; Wackett et al., 1987). All other phosphonate degradation pathways absolutely require a substrate with a β-carbonyl moiety adjacent to the phosphonate bond. PnAA (**1** in Figure 3.2), produced by transamination of the most abundant naturally occurring phosphonate molecule- 2-aminoethylphosphonate (2-AEP), is hydrolyzed by the enzyme- phosphonatase through a bicovalent catalytic mechanism involving a Schiff-base intermediate (Morais et al., 2000). Phosphonopyruvate (PnPy), formed either by the enzymatic isomerization of phosphoenolpyruvate (Seidel et al., 1988), or by transamination of

phosphonoalanine (Ternan et al., 1998), is hydrolyzed by the enzyme- phosphonopyruvate hydrolase, using a mechanism similar to the isomerization of PnPy, differing only in the nucleophile which attacks the metaphosphate intermediate (Chen et al., 2006). Lastly, enzymatic hydrolysis of PnA (**2** in Figure 3.2) to phosphate and acetate has been demonstrated. Phosphate starvation-independent PnA hydrolysis activity was observed in the crude extracts of *Pseudomonas fluorescens* 23F (McMullan et al., 1992). The enzyme was subsequently purified from the native producer, and shown *in vitro* to catalyze the zinc-dependent hydrolysis of PnA (McGrath et al., 1995). Subsequently, the corresponding gene (designated *phnA*) was cloned and the gene-product was heterologously expressed in bacteria and yeast (Kulakova et al., 1997). PnA hydrolysis activity has since been demonstrated in *Streptomyces* spp. (Panas et al., 2006), *Penicillium* spp. (Forlani et al., 2006), marine bacteria, and coral holobionts (Thomas et al., 2010), establishing the presence of PhnA-mediated PnA hydrolysis in the microbial metabolome. It should be noted that the biogenic origins of PnA had not be established until this stage, and have been elucidated very recently only (Borisova et al., 2011).

Apart from the highly promiscuous C-P lyase pathway, the C-P bond-cleaving enzymes in the other three pathways require the presence of a β-carbonyl moiety adjacent to the phosphorus atom. However, the utilization of this β-carbonyl moiety by the three enzymatic mechanisms is vastly different. The PnAA hydrolyzing enzyme uses this group to generate a Schiff-base intermediate that activates the phosphonate group for attack by an active site nucleophile and provides resonance stabilization for the developing negative charge on carbon (Morais et al., 2000). PnPy hydrolase utilizes the β-carbonyl group for stabilization of the pyruvate leaving group by delocalizing the developing negative charge on the methylene group to the carbonyl oxygen (Chen et al., 2006; Kulakova et al., 2003; Ternan et al., 2000). The

mechanism employed by PnA hydrolase suggested by the co-crystal structures presented previously and in this study (Agarwal et al., 2011; Kim et al., 2011), also requires charge delocalization onto the β-carbonyl.

Biosynthesis and degradation of PnA in *Sinorhizobioium meliloti* 1021

The enzymatic basis of conversion of 2-AEP into PnA and its subsequent hydrolysis to phosphate in the soil dwelling bacterium *S. meliloti* 1021, has recently been characterized (Borisova et al., 2011) (Figure 3.2B). The corresponding gene cluster consists largely of previously described enzymatic activities, including those for transamination of 2-AEP to PnAA (catalyzed by the ORF *phnW)* and for hydrolysis of PnA into inorganic phosphate (catalyzed by the ORF *phnA)*. However, this cluster also includes a novel nicotinamide adenine dinucleotide (NAD^+) dependent aldehyde dehydrogenase (ALDH) (*phnY*) whose gene product was shown to carry out the oxidization PnAA to PnA. PhnY thus provides a biogenic route for the production PnA and establishes a fifth pathway for microbial phosphonate degradation. Further bioinformatic analysis demonstrates that similar 2-AEP degradation pathways are present across many different bacterial species and PnAA oxidation might be widespread across diverse microbial populations (Villarreal-Chiu et al., 2012). Indeed, an $NADP^+$-dependent PnAA oxidation activity was subsequently reported in the cell-extracts of a marine Roseobacter-*Roseovarius nubinhibens* ISM (Cooley et al., 2011).

Primary sequence analysis reveals PhnA to be a member of the alkaline phosphatase (AP) superfamily that can catalyze a range of phosphoryl and sulfuryl transfer reactions (Coleman, 1992; Galperin et al., 1998). APs, and the related NPPs, catalyze the hydrolysis of phosphate

monoesters and diesters, respectively. The active site of these enzymes consists of a bi-metallo core and contains a catalytically required Ser or Thr that forms a covalent phosphoenzyme intermediate (Holtz et al., 1999). *S. meliloti* 1021 PhnA was shown to catalyze the hydrolysis of PnA to yield phosphate and acetate (Borisova et al., 2011).

Here, we present detailed biochemical and crystallographic characterization of the PnAA dehydrogenase PhnY from *S. meliloti* 1021 in the apo form, and of cocrystal complexes with the physiological substrate- PnAA, with and without the presence of the cofactor NAD^+ in the active site. To facilitate the description of how the PhnY active site is primed for 'phosphonated'-aldehyde dehydrogenation, in comparison to phosphorylated-aldehyde substrates of primary metabolism, we have also determined the crystal structure of wild type PhnY in complex with glyceraldehyde-3-phosphate (G3P) (**3** in Figure 3.2) and NAD^+. This is also the first report of a non-phosphorylating ALDH enzyme in complex with the reaction product- PnA, which allows for a comparison between the substrate and product binding interactions within the enzyme active site. We present kinetic characterization of structure guided site-specific mutants. Our structural and biochemical data, in the context of previous studies on the reaction mechanism of ALDHs, allow for the mechanistic proposal of the catalytic cycle for the oxidation of PnAA to PnA by PhnY.

In order to elucidate the mechanistic basis for C-P (rather than O-P) bond hydrolysis within the AP bi-metallo active site, we present detailed kinetic and structural studies of PhnA from *S. meliloti* 1021. Crystal structures of the wild-type PhnA enzyme confirm structural similarity to APs but with notable differences that likely direct activity toward the cleavage of a C-P bond. Kinetic analyses of metal-substituted PhnA and corresponding anomalous scattering diffraction experiments on zinc- and manganese-containing PhnA establish the in vitro metal

preference. Cocrystal structures with substrate PnA, product acetate ion, inhibitor PnF, and the transition state mimic vanadate, together with kinetic and structural analysis of site-specific active site variants, suggest a plausible mechanism for PnA hydrolysis by PhnA. Our results expand upon a very recent study of a PnA hydrolase from *P. fluorescens* 23F (Kim et al., 2011).

Experimental procedures

Cloning, protein expression and purification

The construction of plasmid DNA bearing the wild type and PhnY and PhnA enzymes has been described previously (Borisova et al., 2011). These plamids were generously provided to us by the laboratories of Dr. Wilfred A. van der Donk and Dr. William W. Metcalf (University of Illinois at Urbana Champaign). All mutants were constructed using standard site directed mutagenesis protocols. All plasmids were sequeced (ACGT Inc., Wheeling- IL) to verify the site of mutagensis.

PhnY wild type and all PhnY mutant enzymes were purified under identical conditions. All plasmids were transformed into competent *E. coli* Rosetta2(DE3) strains for recombinant protein expression. A single colony of the transformed cells was grown overnight in Luria-Bertani (LB) medium supplemented with 50 µg/mL kanamycin and 25 µg/mL chloramphenicol. This was subsequently used to innoculate 6 L of LB medium similarly supplemented with antibiotics, and grown at 37° C till the O.D. at 600 nm reached 0.6. Isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 0.5 mM and the culture was shifted to 18° C and grown for an additional 16 hours. All subsequent steps were performed at 4° C. The bacterial cells were harvested by centrifugation and resuspended in 20 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10% glycerol buffer and lysed by multiple passes through an Avestin C5 homogenizer. The cell lysate was clarified by centrifugation and the supernatant was loaded onto a 5 mL bed volume His-Trap column (GE Biosciences) equilibrated with lysis buffer. Subsequent to loading, the column was extensively washed with 20 mM Tris-HCl (pH 8.0), 1 M NaCl, 30 mM imidazole buffer and eluted by a continous linear gradient to 250 mM imidazole

concentration in the wash buffer. Greater than 95% pure fractions thus obtained were extensively dialyzed against 20 mM Tris-HCl (pH 8.0), 500 mM NaCl with the addition of 1 unit/mg of thrombin enzyme to proteolytically remove the hexahistidine tag. Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was used to verify the complete digestion of the hexahistidine tag. The protein was then concentrated by spin filtration and were applied to Superdex-75 16/60 size exclusion chromatography column (GE Biosciences). The protein was isochratically eluted in 20 mM HEPES-Na (pH 7.5), 300 mM KCl buffer and concentrated to 0.2 mM final concentration before being flash frozen in liquid nitrogen and storage at -80° C in small aliquouts.

Selenomethione incorporated PhnY wild type enzyme was prepared by previously established protocols (Van Duyne et al., 1993). Culture was grown under identical conditions in M9 minimal medium supplemented with antibiotics and 5 g/L D-glucose. When the O.D. reached 0.4, 100 mg/L L-leucine, 100 mg/L L-lysine, 100 mg/L L-valine, 50 mg/L L-threonine, 50 mg/L L-phenylalanine and 50 mg/L L-isoleucine amino acids were added to the culture medium. After 20 minutes of further growth at 37° C, 50 mg/L L-selenomethione amino acid was added to the culture medium. When the O.D. reached 0.6, culture was induced with IPTG and grown for an additional 16 hours at 18° C. The selenomethionine derivitized protein was purified identically to the wild type and mutant forms of PhnY.

The plasmid bearing the wild type PhnA was similarly transformed into competent Rosetta2(DE3) cells and recombinant protein expression carried out under identical conditions. The media was supplemented with 50 µg/mL kanamycin and 25 µg/mL chloramphenicol. No exogenous metal ions were supplemented in the growth media.The recombinant protein bearing the hexahistidine tag was isolated by Ni-NTA chromatography using the same protocol as that

for the PhnY enzyme. Greater than 95% pure fractions thus obtained were extensively dialyzed against 20 mM Tris-HCl (pH 8.9), 50 mM NaCl with the addition of 1 unit/mg of thrombin enzyme to proteolytically remove the hexahistidine tag. SDS-PAGE was used to verify the complete digestion of the hexahistidine tag. Protein was then loaded onto a Q-sepharose ion exchange column (GE Biosciences) equilibrated in the dialysis buffer and eluted by a linear continuous gradient to dialysis buffer supplemented with 1 M NaCl. Concentrated elution fractions were directly applied to Superdex-75 16/60 size exclusion chromatography column. The protein was isochratically eluted in 20 mM HEPES-Na (pH 7.5), 100 mM KCl buffer and concentrated to 1 mM final concentration before being flash frozen in liquid nitrogen and storage at -80° C in small aliquouts. The purified recombinant protein was a monomer in solution as determined using analytical size exclusion chromatography. PhnA protein used in the activity assays did not have the N-terminal hexahistidine tag removed.

Protein crystallization, X-ray data collection and structure determination

The crystals of PhnY were grown at 15° C using hanging drop vapor-diffusion method by mixing 1 μL protein sample at 8 mg/mL concentration, with equal volume of reservoir mother liquor. Apo-form crystals were obtained in mother liquor containing 22-25% PEG3350, 0.1 M sodium cacodylate buffer (pH 6.5), 0.2 M sodium phosphate dibasic and 5-10% glycerol within two days and grew to their maximum size in one week. For crystallization in presence of cofactor NAD^+, substrate PnAA, product PnA, and substrate analog G3P, the salt in mother liquor needed to be substituted from sodium phosphate. Several salts at various concentrations were trialed for this purpose. 0.2 M ammonium chloride was found to be the best candidate, as

judged by crystal morphology, and subsequently used for growing all crystals. Crystals were cryoprotected by brief soaking in mother liquor supplemented with 25% glycerol and vitrified in liquid nitrogen. The crystals appeared as plates which cracked instantly on exposure to the cryoprotectant solution. A number of different cryoprotectants were thus trialed and glycerol was found to be the best candidate, as judged by X-ray diffraction patterns and spot splitting.

A fourfold redundant data set was collected from a single crystal of *S. meliloti* 1021 PhnY to a limiting resolution of 2.1 Å (overall R_{merge} = 0.072, I/σ(I) = 2.9 in the highest resolution shell) utilizing a Mar 300 CCD detector (LS-CAT, Sector 21 ID-D, Advanced Photon Source, Argonne, IL). Data were indexed and scaled using the HKL-2000 package (Otwinowski et al., 2003). The structure of PhnY was solved by SAD phasing using procedures described previously. Most of the main chain could be built using ARP/wARP (Perrakis et al., 1997). Cross-validation used 5% of the data in the calculation of the free R factor (Kleywegt and Brunger, 1996). The remainder of the model was fitted using XtalView and further improved by rounds of refinement with REFMAC (Murshudov et al., 1997; Murshudov et al., 1999) interspersed with rounds of manual building using XtalView (McRee, 1999).

The cocrystal structures of PhnY in complex with PnAA, NAD^+-PnAA, NAD^+-PnA, and NAD^+-G3P, were determined to resolutions of 2.1 Å, 2.1 Å, 2.1 Å and 2.24 Å, respectively, by molecular replacement using the coordinates of apo PhnY as a search probe. Cross-validation was routinely used throughout the course of model building and refinement using 5% of the data in the calculation of the free R factor. For each of the structures, the stereochemistry of the model was monitored throughout the course of refinement using PROCHECK (Laskowski et al., 1996).

Apo wild type PhnA protein at a final concentration of 20 mg/mL in 20 mM HEPES-Na (pH 7.5), 100 mM KCl buffer was used for sparse matrix crystallization screening trials, using the hanging drop vapor diffusion technique. Diffraction quality crystals were obtained in two mother liquor conditions at 9º C, 25% PEG 3350, 0.2 M sodium chloride, 0.1 M Tris-HCl (pH 8.5) and 25% PEG 3350, 0.2 M ammonium acetate, 0.1 M HEPES-Na (pH 7.5). Crystals typically took three days to grow and were briefly soaked in mother liquor supplemented with 15% glycerol prior to vitrification in liquid nitrogen. PhnA mutants were crystallized according to the protocols for wild-type enzyme. The PhnA mutant Thr68Ala was incubated with 2 mM zinc chloride and 10 mM PnA for 2 hours on ice prior to sparse screening for crystallization. Co-crystals of PhnA Thr68Ala mutant in complex with PnA were obtained in the crystallization condition 20% PEG 3350, 0.2 M ammonium chloride and briefly soaked in cryoprotectant solution of mother liquor supplemented with 20% glycerol, 5 mM zinc chloride and 50 mM PnA prior to vitrification in liquid nitrogen. Co-crystals of the complex with PnF were obtained by soaking apo protein crystals in 10 mM of PnF in the mother liquor for 12 hours. Metal ions were soaked by supplementing the mother liquor with 5 mM of zinc chloride, 5 mM manganese chloride, or 5 mM of iron(II) ammonium sulfate for 3 hours. To generate the covalently bound vanadate complex, apo protein was incubated on ice for 10 minutes with 2 mM freshly boiled sodium orthovanadate solution and then crystallized in the manner described above.

A ten-fold redundant data set was collected from crystals of *S. meliloti* 1021 PhnA to a limiting resolution of 1.35 Å (overall R_{merge} = 0.068, I/σ(I) = 1.8 in the highest resolution shell) utilizing a Mar 300 CCD detector (LS-CAT, Sector 21 ID-D, Advanced Photon Source, Argonne, IL). The structure of PhnA was solved by single wavelength anomalous diffraction utilizing anomalous scattering from a mercury derivative (six-fold redundancy with R_{merge} = 0.084, I/σ(I)

= 3.2 in the highest resolution shell). Data were indexed and scaled using the HKL-2000 package (Otwinowski et al., 2003). Mercury sites were identified using HySS and the heavy atom substructure was imported to SHARP for maximum likelihood refinement and phase calculation, yielding an initial figure of merit of 0.497 to 1.9 Ǻ resolution. Solvent flattening using DM further improved the quality of the initial map and most of the main chain could be built using ARP/wARP (Perrakis et al., 1997). Cross-validation used 5% of the data in the calculation of the free R factor (Kleywegt and Brunger, 1996). The remainder of the model was fitted using XtalView and further improved by rounds of refinement with REFMAC (Murshudov et al., 1997; Murshudov et al., 1999) interspersed with rounds of manual building using XtalView (McRee, 1999).

The co-crystal structures of PhnA with PnA, acetate, PnF, and vanadate were determined, to resolutions of 2.1 Å, 2.0 Å, 1.6 Å, and 1.8 Å, respectively, by molecular replacement using the coordinates of native PhnA as a search probe. Each of the structures was refined and validated using the procedures detailed above. Cross-validation was routinely used throughout the course of model building and refinement using 5% of the data in the calculation of the free R factor. For each of the structures, the stereochemistry of the model was monitored throughout the course of refinement using PROCHECK (Laskowski et al., 1996). Relevant data collection and refinement statistics are provided in Table 3.2.

PhnY enzyme kinetics

The kinetics of conversion of PnAA to PnA by wild type PhnY enzyme and site specific mutants of PhnY were determined by monitoring the concomitant and equimolar reduction of the

NAD^+ by PhnY using procedures described previously (Borisova et al., 2011). The Asn158Ala, Arg290Ala and Glu385Ala mutant enzymes were assayed at 15 μM final enzyme concentration and 1 mM final PnAA concentration. The low level activity precluded determination of K_M values. k_{cat} values were verified by a linear correlation between the rate of change of absorbance and enzyme concentration used in the assay for each mutant enzyme. The Cys291Ala and Glu254Ala mutant enzymes were assayed at various substrate and enzyme concentrations but no activity could be observed. Activity for G3P was monitored at 2 mM final concentration of substrate using three different enzyme concentrations (7 μM, 15 μM and 20 μM final enzyme concentrations). The low level activity precluded determination of K_M values. The Glu254Asp mutant protein was quite unstable during purification processes and could not be concentrated above ~ 20 μM concentration. Hence comparison assays for G3P dehydrogenation between the wild type and Glu254Asp enzymes were conducted at 10 μM enzyme concentration.

PhnA enzyme kinetics

The generous help of Dr. Svetlana A. Borisova in conducting the kinetics experiments for PhnA is gratefully acknowledged. The formation of inorganic phosphate from PnA by the action of N-terminally hexahistidine tagged PhnA (PhnA-N-His) was detected by a discontinuous assay using a Malachite Green phosphate assay kit (BioAssay Systems, Hayward, CA). Assay mixtures (500 μL total volume) containing 50 mM HEPES-Na (pH 7.5) and 0.42 μM PhnA-N-His were pre-incubated at 30° C for 8 minutes and the reaction was initiated by the addition of PnA stock solutions to final concentrations of 0-400 μM. Aliquots of the reaction mixture (80 μL) were taken out every 20 seconds over a period of 2 minutes, quenched by the addition to 20 μL of

Malachite Green reagent prepared as per the manufacturer's instructions, and incubated at room temperature for 30 minutes for color development. The assays were done in duplicate. The absorbance at 620 nm was plotted against the reaction time and the rate of A_{620} increase was converted to the rate of phosphate formation using a linear calibration curve prepared with known concentrations of inorganic phosphate standard (0-40 µM). The initial rates of phosphate formation were fit to the Michaelis-Menten equation ($V_0=([S]*V_{max})/([S]+K_M)$) using the IGOR Pro 6.1 software package (WaveMetrics, Portland, OR) in order to determine steady state kinetic parameters of PhnA-N-His.

To evaluate the divalent metal dependence of PhnA-N-His, 10 µM $ZnCl_2$, $MgCl_2 \times 6H_2O$, $MnCl_2 \times 4H_2O$, $CoCl_2 \times 6H_2O$, or $(NH_4)_2Fe(SO_4)_2 \times 6H_2O$ were added to the assay mixture prior to the pre-incubation period and assays were performed as described above. The concentrations of PhnA-N-His were adjusted (0.07-0.48 µM) to allow for the detection of product formation within the linear range of the assay. Assays in the presence of oxygen-sensitive Fe(II) were set up in an anaerobic chamber obtained from Coy Laboratory Products, Inc. (Grass Lake, MI) under an atmosphere of N_2 and H_2 (95%/5%). The aliquots of buffer-enzyme solution containing Fe(II) were subsequently brought outside the glove box in tightly capped eppendorf tubes, followed by pre-incubation and reaction initiation with PnA as described above.

Metal-free PhnA-N-His was prepared by treatment of the protein with 8.3 mM EDTA sodium salt at 4° C for 3 hours with gentle agitation followed by size exclusion chromatography using a PD-10 desalting column (GE Healthcare, Piscataway, NY) eluted with 50 mM HEPES-Na (pH 7.5), 0.2 M NaCl, 10% glycerol. The resulting apo PhnA-N-His was tested in the activity assay as described above with and without addition of Zn^{2+}, Mn^{2+}, or Fe^{2+}. When combinations of two divalent metal cations were studied (at 10 µM each) initial rates of product formation

were measured at 200 μM PnA. Enzymatic activity of PhnA mutants purified as above but not treated with EDTA (at 10 μM, except for the Thr68Ala mutant at 5 μM) was measured as described above in the presence of 20 μM Mn^{2+} and 0.2 or 1.0 mM PnA. The formation of P_i was monitored over a period of 0.5-1 hours.

Results and discussion

Overall crystal structure of wild type PhnY in apo state

The structure of PhnY from *S. meliloti* 1021 has been determined in the wild type apo state to 2.1 Å resolution. Crystallographic phases were determined by single wavelength anomalous diffraction data collected on crystals of selenomethionine labeled protein, followed iterative rounds of manual building and automated refinement to generate a final model with R/R_{free} = 0.199/0.246. The crystallographic asymmetric unit consists of eight protein chains comprising four sets of biologically relevant homodimers. The arrangement of these protein chains as sets of homodimers is consistent with the behavior of PhnY as a dimer in solution, as judged by its elution profile using size exclusion chromatography (Figure 3.3). Members of the aldehyde dehydrogenase (ALDH) superfamily have been previously found to exist as dimers as well as tetramers in solution.

Each PhnY monomer has a tertiary structure that can be divided into three distinct domains (Figure 3.4). Residues 1-123, 147-256 and 461-475 together constitute the 'cofactor-binding domain' and adopt a classical α/β Rossmann fold. Residues 257-460 comprise of the 'catalytic domain' with a α/β fold comprised of a central seven membered β sheet flanked by α helices on either side. Residues 124-146, together with residues 476-486 constitute the three-membered antiparallel β sheet 'oligomerization' domain that provides the principal contacts for dimerization. The secondary structure loops at the interface of the cofactor-binding domain and the catalytic domain of PhnY provide nearly all the residues implicated in substrate binding and catalysis (*vide infra*). The active site is located at the solvent occluded interface between the cofactor-binding and catalytic domains, with the catalytic nucleophile- Cys291 positioned on a

loop that connects the third and fourth β sheet of the catalytic domain. Another distinctive long loop region in the vicinity of the active site at the domain interface spans residues 440-467 and connects the penultimate β strand of the catalytic domain to a short α helix of the cofactor binding domain, and finally culminates in a β strand in the oligomerization domain.

ALDH superfamily of enzymes can be divided into three subfamilies based on the reaction catalyzed. These are the phosphorylating ALDHs, the non-phosphorylating ALDHs, and the coenzyme A dependent non-phosphorylating ALDHs. The structure of PhnY is similar to that of members of non-phosphorylating ALDH superfamily. A DALI search against the Protein Data Bank identifies the strongest structural homology to G3P dehydrogenase (GAPDH) from *Streptococcus mutans* (Cobessi et al., 1999) (PDB Code 2EUH, Z-score=54.3 , RMSD=1.6 Å over 468 aligned Cα atoms), aminoaldehyde dehydrogenase from *Pisum sativum* (Tylichova et al., 2010) (PDB Code 3IWK, Z-score=53.1 , RMSD=1.9 Å over 468 aligned Cα atoms), human succinic semialdehyde dehydrogenase (Kim et al., 2009) (PDB Code 2W8R, Z-score=52.6, RMSD=1.8 Å over 464 residues), and betaine dehydrogenase from *P. aeruginosa* (Diaz-Sanchez et al., 2011b) (PDB Code 2WOX, Z-score=52.3, RMSD=1.8 Å over 463 residues).

Electron density corresponding to a phosphate ion (presumably originating from the crystallization mother liquor) can be observed at the active site of PhnY (Figure 3.5). The phosphate ion is present near Cys291, where the three of the four phosphate oxygen atoms are coordinated by hydrogen bonds with the side chains of Asn158 (Nδ2 - O1 distance of 3.1 Å), His159 (Nε2 - O2 distance of 2.8 Å), Arg290 (Nη2 - O2 distance of 3.3 Å) and Arg447 (Nη2 - O3 distance of 3.3 Å). A primary sequence alignment of ALDHs that catalyze dehydrogenation reaction for a phosphorylated substrates and ALDHs which do not require phosphorylated substrates identifies the strict conservation of this basic, phosphate binding pocket formed by

three Arg and one His/Tyr residues in the former (Figure 3.6). The basic phosphate binding pocket is found conserved in the *S. meliloti* 1021 PhnY, Roseobacter PhnY and the *S. mutans* GAPDH, while being absent from the succinic semialdehyde and betaine dehydrogenases. Hence a primary sequence alignment of a non-phosphorylating ALDH enzyme can used to predict whether the enzyme substrate would be phosphorylated or non-phosphorylated, by the presence or absence of the basic phosphate binding pocket.

Cocrystal structure with physiological phosphonate substrate PnAA

Freshly prepared PnAA was synthesized as described previously (Borisova et al., 2011), and generously provided by Dr. Svetlana A. Borisova (Mining Microbial Genomes theme, Institute for Genomic Biology, University of Illinois at Urbana Champaign). Enzymatically synthesized PnAA was used at 10 mM concentration without further purification. Crystals of PhnY with covalently bound PnAA were obtained by incubating the wild type enzyme with 5 mM PnAA, and subsequent crystallization using solutions devoid of phosphate salts. A complete dataset was collected from a single crystal and the structure determined using the apo PhnY structure as a search model for molecular replacement. The resultant 2.1 Å resolution cocrystal structure is nearly identical to that of the apo structure with no secondary or tertiary structural rearrangements apparent. The $|F_{obs}| - |F_{calc}|$ difference electron density map shows unambiguous density for all atoms of PnAA in the active site of all eight protein monomers in the crystallographic asymmetric unit (Figure 3.7A). PnAA is bound at the interface of the cofactor-binding and catalytic domains. The side chain thiol of Cys291 undergoes a 120^{o} rotation, relative to the apo structure, which correctly positions this residue for interactions with the substrate. The

Cys291-Sγ atom in this conformation is 3.0 Å and 3.3 Å away from the backbone amide nitrogen atoms of Cys291 and Thr292 respectively. Similar interactions, in related members of the ALDH superfamily, have been implicated in modulating the nucleophilicity of the Cys side chain for nucleophilic attack.

Two lines of evidence suggest that Cys291 engages PnAA through a covalent linkage: first, continuous electron density in the $|F_{obs}| - |F_{calc}|$ map is consistent with the substrate aldehyde carbon atom positioned 1.65 Å away from the Cys291-Sγ atom, indicative of covalent bond between these two atoms. Second, the electron density at the aldehyde carbon atom suggested that the stereochemistry is not planar sp^2, as would be expected for an aldehyde moiety. Instead, the bond angle between the three atoms- (i) PnAA-aldehyde O (ii) PnAA-Cβ (iii) PnAA-Cα atoms is 111°, and the bond angle the aldehyde (i) PnAA-aldehyde O (ii) PnAA-Cβ (iii) Cys291-Sγ atoms is 107.7° (phosphonate carbon atom of PnAA is denoted as Cα, the aldehyde bearing carbon atom of PnAA is denoted as Cβ) (Figure 3.7B). These bond angles are indicative of a tetrahedral sp^3 stereochemistry at the PnAA-Cβ atom, which can be only be generated by the nucleophilic addition of the Cys291 side chain thiol (Cys291-Sγ) on the PnAA-Cβ atom to generate a tetrahedral thiohemiacetal reaction intermediate. Thus the hydrogen atom, which would occupy the fourth position in the tetrahedral geometry around the PnAA-Cβ atom, would point opposite to the plane defined by (i) PnAA-Cα (ii) PnAA-aldehyde O (iii) Cys291-Sγ atoms. Consistent with the proposed reaction mechanism for ALDH enzymes, this geometry would revert back to planar sp^2at the PnAA-Cβ atom once the hydride is transferred to the nicotinamide ring of NAD^+ cofactor to generate a thiolate ester of the substrate with the Cys291-Sγ atom (Figure 3.10), but as crystals were grown without the cofactor, this structure represents the substrate in the form of a catalytically relevant covalent tetrahedral thiohemiacetal intermediate.

The mechanistic scheme for the reaction catalyzed by PhnY is discussed in greater detail in later sections.

The covalent bond between the Cys291-Sγ atom and the PnAA-Cβ atom converts the PnAA aldehyde moiety into a hydroxyl group; with the hydroxyl oxygen positioned 3.1 Å from the amide side chain of Asn158 (Figure 3.7A). As next nearest interaction partner for this hydroxyl is His159 (positioned 4.0 Å away), Asn158 provides the principal interaction for the stabilization of the tetrahedral thiohemiacetal intermediate hydroxyl. An additional stabilizing interaction is provided by the backbone amide nitrogen atom of Cys291 residue. Thus the side chain of Asn158 and main chain amide nitrogen atom of Cys291 constitutes the 'oxyanion hole' for stabilization of the thiohemiacetal intermediate hydroxyl.

Interactions with the oxygen atoms of the PnAA substrate phosphonate moiety are provided by Arg108 (Nη1 – O1 distance of 2.9 Å) (Nη2 – O1 distance of 3.0 Å), Arg290 (Nη2 - O2 distance of 3.1 Å) (Nε – O3 distance of 2.8 Å) and Arg447 (Nη2 – O1 distance of 2.8 Å) (Figure 3.13A). Additional interactions include a hydrogen bond with His159 (Nε2 – O3 distance of 2.6 Å) and a water-mediated hydrogen bond with the backbone carbonyl oxygen atom of Gly445 (not shown in figure 3.7).

Kinetics of PnAA oxidation by wild type and mutant PhnY enzymes

The structural model of the PhnY in covalent complex with PnAA described in the previous section has been verified by extensive mutational and kinetic analysis. Cys291, Asn158, His159, Arg108, Arg290 and Arg447, which contribute to substrate engagement, were targeted for mutagenesis to alanine, and the change in kinetic properties of the mutant enzymes was

determined by monitoring the increase in absorbance at 340 nm due to production of reduced nicotinamide adenine dinucleotide (NADH) during the course of oxidation of PnAA to PnA (Borisova et al., 2011). For the wild type enzyme, the Michaelis-Menten steady state kinetics parameters calculated from at least three independent initial rate measurements for each different substrate concentration yielded $k_{cat} = 2.1 \pm 0.4$ sec^{-1} and $K_M = 3.2 \pm 0.2$ μM for PnAA which are in good agreement with the already published values (Borisova et al., 2011) (Table 3.1).

No activity could be observed when the catalytically requisite residue Cys291 was replaced by Ala. The PhnY-His159Ala mutation led a modest decrease in k_{cat} to $(16.33 \pm 1.36) \times 10^{-2}$ sec^{-1} and an increase in K_M value to 136.61 ± 12.39 μM with respect to PnAA. However, the Asn158Ala mutation resulted in a k_{cat} of $(4.25 \pm 0.44) \times 10^{-3}$ sec^{-1} that is reduced by nearly 500-fold relative to the wild-type. K_M values for the mutant enzyme could not be determined as the initial rates could not be reproducibly measured at lower substrate concentrations. A linear correlation of this k_{cat} value with enzyme concentration used in the assay verified that the measured absorbance increase was at a fully saturating substrate concentration. Replacement of Arg108 by Ala led to 80 fold decrease in k_{cat} to $(25.36 \pm 0.94) \times 10^{-3}$ sec^{-1} and a 25 fold increase in K_M to 75.08 ± 10.38 μM with respect to PnAA. The Arg290Ala mutation results in a drastic decrease in k_{cat} to $(4.75 \pm 0.77) \times 10^{-3}$ sec^{-1}. For this particular mutant the K_M value for could not be determined. The Arg447Ala mutation reduced the k_{cat} value to $(35.13 \pm 4.88) \times 10^{-3}$ sec^{-1} and drastically increased the K_M to 392.93 ± 27.82 μM. As all of the above residues directly contact the PnAA substrate, the K_M values for mutations at these sites have been determined with respect to PnAA.

As noted before, binding of PnAA in a productive conformation in the PhnY active site requires a 120° rotation of the catalytic Cys-Sγ atom. This rotation is affected upon substrate

binding. Hence it can be inferred that the PhnY active site is not pre-organized in a catalytically proficient conformation. The required catalytically proficient conformation is achieved by the covalent binding of the PnAA-Cβ to Cys291 side chain, and the subsequent rotation of the Cys291 side chain, which in turn is driven by the favorable enthalpy of binding of PnAA phosphonate oxygen atoms to the basic binding pocket. This enthalpic contribution of the phosphate binding pocket is reflected in the reduction of k_{cat} values by mutagenesis, concomitant with increase in K_M values (Table 3.1). Another structural finding which supports this hypothesis is the positioning of the side chain of Arg108 residue in the PhnY apo crystal structure in presence of P_i (Figure 3.5) and in the presence of PnAA (Figure 3.7). In the presence of P_i, the Arg108 side chain is too far to coordinate to the oxygen atoms of P_i and is hence moved away from the active site. This is analogous to an Arg108Ala mutation with respect to the active site. However binding of PnAA organizes the Arg108 side chain to a substrate binding productive conformation, where it is absolutely required for efficient catalysis, as the Arg108Ala mutant is nearly 2000 fold catalytically less proficient than the wild type enzyme (as judged by the ratio of k_{cat}/K_M values).

Structure of the ternary complex with NAD^+ and PnAA offers insight into the deacylation state of the enzyme

As ALDHs follow ordered sequential kinetics for substrate oxidation (Cobessi et al., 2000; D'Ambrosio et al., 2006; Marchal and Branlant, 1999; Marchal et al., 2000), structural data on a cofactor bound complex can only be achieved in a ternary complex. In order to determine the enzyme active site architecture by the concomitant binding of the substrate and the cofactor

molecules, we attempted to co-crystallize the enzyme with fivefold molar excess of the substrate as compared to the cofactor. Once crystals were obtained, they were soaked with crystallization solution supplemented with 5 mM of NAD^+ and harvested by cryoprotection and vitrification in liquid nitrogen at different time points, prior to data collection. We envisaged that this approach should potentially yield an architecture which would describe the tetrahedral geometry around the aldehyde carbon prior to hydride transfer and in which the cofactor is in the productive hydride transfer conformation. Alternately, it should yield the planar geometry around the aldehyde carbon subsequent to the hydride transfer with the cofactor is positioned in the conformation which would support deacylation of the transition state by nucleophilic attack by a water molecule. The two scenarios can also offer an insight into the rate limiting step of the ALDH catalytic cycle (discussed in detail in a later section). Visualization of the first scenario will suggest that the hydride transfer step is rate limiting, while visualization of the second scenario might suggest that the deacylation step is rate limiting.

However, this experimental approach, despite our exhaustive efforts in diffraction data collection after different time points after the crystal soaking, was unsuccessful in describing either of the two functional states for the enzyme. The electron density maps obtained for the cofactor after all time points demonstrated a lack of electron density for the nicotinamide ring of the cofactor (Figure 3.8), and hence were not used for the description of the cofactor. Similar results were obtained when PhnY was crystallized in complex with NAD^+ alone. These observations have been reported in literature before, and their functional implications with respect to the proposed reaction mechanism are discussed in greater detail in a later section.

However the stereochemistry around the Cβ atom of PnAA has shifted to a planar sp^2 hybridization, as judged by the bond angles between the three atoms- (i) PnAA-aldehyde O (ii)

PnAA-Cβ (iii) PnAA-Cα atoms (120.9°), and the bond angle between (i) PnAA-aldehyde O (ii) PnAA-Cβ (iii) Cys291-Sγ (118.4°) (Figure 3.8B). This shift in stereochemistry is indicative of the post-hydride state of the enzyme. This is also supported by the cofactor nicotinamide ring not being positioned in the productive hydride transfer conformation. The most unexpected result, however, was the finding of the side chain positioning of the catalytic base residue- Glu254 (*vide infra*). The glutamate side chain carboxylate points inwards towards the active site, but enzyme deacylation has not yet occurred. This leads us to postulate that deacylation, rather than hydride transfer is the rate limiting step for the ALDH reaction mechanism.

Structure of the ternary complex with NAD^+ and PnA

A complete description for all atoms of the NAD^+ cofactor could only be obtained when wild type PhnY enzyme was co-crystallized with the cofactor and the reaction product- PnA. The 2.1 Å resolution cocrystal structure of PhnY in complex with NAD^+ and PnA shows clear and continuous density for the cofactor in all copies of the monomer peptides in the crystallographic asymmetric unit (Figure 3.9). In this structure, NAD^+ is bound such that the nicotinamide ring is situated at the interface of the cofactor-binding and catalytic domains. Interactions stabilizing this conformation of the cofactor are described starting from the adenine ring, and proceeding to the nicotinamide ring.

The adenine ring of NAD^+ is positioned in a hydrophobic pocket defined by Val238, Leu241 and Ile242 on one face, and Pro214 at the other face. The 2’ and 3’ hydroxyls of the adenosine ribose sugar are both stabilized by interactions with Lys181 side chain amine. The 2’ hydroxyl makes additional contacts with Glu184, while the 3’ hydroxyl contacts the backbone

amide of Thr183 and the backbone carbonyl of Thr155. The carboxyl side chain of Glu385 provides the only interaction for the nicotinamide ribose 2' and 3' hydroxyl groups. The nicotinamide ring is nearly parallel to the Cys291 side chain, which is situated at the *re*- face of the nicotinamide ring. Although the distance between the Cys291-Sγ atom and the nicotinamide C4 atom is only 2.3 Å, the two atoms are unlikely to be covalently linked. There is no continuous electron density between the Cys291-Sγ and the nicotinamide C4 atoms, unlike the case of *P. aeruginosa* betaine ALDH that does indeed form a covalent complex between these atoms (Diaz-Sanchez et al., 2011a, b). A hydrophobic binding pocket composed of Pro156, Met163 and Thr233 is located at the *si*- face of the nicotinamide ring. A comparative decrease in conformational ordering of the cofactor, starting from the adenine to the nicotinamide terminus of NAD^+, is evident in the quality of the difference electron density maps (Figure 3.9).

The ternary complex structure with NAD^+ shows Glu385 side chain provides the only interaction with the nicotinamide ribose 2' and 3' hydroxyl groups. Consequently, the Glu385Ala mutation results in a 300 fold reduction in k_{cat} value to $(6.91 \pm 0.87) \times 10^{-3}$ sec^{-1}. In order to elucidate the molecular basis for this reduction in catalytic activity, we solved the 2.3 Å resolution cocrystal structure of PhnY Glu385Ala mutant enzyme in complex with NAD^+. The structure shows that cofactor binding occurs in a non-productive manner that would hinder hydride transfer and likely accounts for the lowered activity of this mutant (data not shown). We could not determine the K_M values for NAD^+ or PnAA for this mutant due to the drastic reduction in activity, but, based on the cocrystal structure, we speculate that the K_M values with respect to PnAA would not be altered to same degree as that for NAD^+.

The PhnY-NAD^+-PnA cocrystal structure also allows for the description of the structural changes at the active site that accompany the oxidation of the substrate PnAA to product PnA.

The cocrystal structure is consistent with a non-covalent product bound complex, as evidenced by the fact that the electron density between the carboxylate carbon atom of PnA and the Cys291-Sγ is discontinuous, and that the distance between the two atoms in the refined structure is 3.7 Å (Figure 3.9). Relative to the PnAA complex structure, the position of the PnAA-Cα atom and the adjoining Cβ-carboxylate has shifted, such that only one of the carboxylate oxygen atoms now co-ordinates with the Asn158 side chain amide, while no interactions exist for the other oxygen atom. In addition, the tetrahedral angle at the PnAA-Cα carbon is reduced by about 5^{o}. These features distinguish the PnA product structure from that of the covalent complex with the PnAA intermediate.

Conformational flexibility of the cofactor is a prerequisite for the proposed reaction mechanism for PhnY

In the proposed reaction mechanism for ALDHs, the reduced cofactor is not released from the enzyme prior to deacylation, but rather 'flips-out' from the enzyme active site to adopt alternate conformations (Cobessi et al., 2000; D'Ambrosio et al., 2006). These alternate conformations of the cofactor have been extensively investigated in literature, and at least three different conformations have been structurally described with the cofactor occupying varying positions that may reflect different spatial positions along the reaction trajectory. A conserved feature of all these cofactor conformations is the relatively higher thermal (B) factors for the nicotinamide ring and the preceding ribose sugar atoms, and lower B values for the adenine ring and the succeeding ribose sugar. This is also clearly reflected in the electron density maps obtained for the cofactor in various different studies, with a lack of electron density for the

nicotinamide ring of the cofactor in some examples. A major determinant for whether the nicotinamide ring of the cofactor could be positioned into the enzyme active site is the conformation of the side chain of the catalytic Cys291. In the PhnY apo crystal structure, the Cys291-Sγ atom is in a position which would sterically clash at the *re* face of nicotinamide ring π-electron system. Upon binding to the substrate PnAA, the Cys291 side chain thiol undergoes a 120° rotation along the Cys291-Cβ atom. The outward rotation of the cysteine side chain makes sufficient room for the nicotinamide side chain to be positioned in a productive conformation for catalysis, thus providing the first structural premise for the ordered-sequential mechanism. A superposition of the PhnY-PnAA cocrystal structure with that of PhnY-PnA-NAD^+ cocrystal structure reveals that the active site geometry defined by the two structures is apt for the transfer of the hydride from the tetrahedral thiohemiacetal carbon (described by the PnAA cocrystal structure) to the C4 carbon atom from the *re* face of the cofactor nicotinamide ring (described by the PnA-NAD^+ cocrystal structure) (Figure 3.10). The thiohemiacetal intermediate is arranged such that the hydride to be transferred to the cofactor would point towards the *re* face of the nicotinamide ring. The change in the pucker of the nicotinamide ring subsequent to hydride transfer may result in steric clashes with these closely spaced atoms, and is presumed to be responsible for the cofactor adopting the 'flipped-out' conformation reported in literature, and as observed in the Glu385Ala-NAD^+ cocrystal structure.

The second premise for the ordered-sequential mechanism is provided by the position of the side chain of Glu254. Based on the structural data and sequence alignments (Figure 3.6), Glu254 was presumed to the general base that deprotonates a water molecule for nucleophilic attack to initiate the deacylation half reaction. The PhnY-Glu254Ala mutant enzyme was found to be catalytically inactive. The manner in which the cofactor binds in the hydride transfer

conformation described above occludes the Glu254 side chain from pointing towards the active site by steric clash between the Glu254 side chain and the *si* face of the nicotinamide ring of the cofactor. Hence, before the cofactor nicotinamide ring departs the active site subsequent to the hydride transfer, the second deacylation half reaction cannot occur. The fact that the cofactor binding for hydride transfer precludes deacylation of the enzyme provides a rationale for the observed flexibility of the nicotinamide ring. It should be noted that the Glu254 is markedly flexible, and does not have any strong stabilizing contacts in its occluded state. This conformational flexibility of the catalytic base has previously been observed for other ALDHs (Bains et al., 2011). We also observe Glu254 positioned in this occluded state in the substrate and product cocrystal structures, while it is relatively disordered in several of the apo structure monomers. Hence the driving force for the conformational change of the Glu254 side chain to bring it to a catalytically efficient state is not immediately clear.

Based on our structural and kinetic data, a reaction mechanism scheme for PhnY can be proposed (Figure 3.11). Pursuant to this scheme, binding of PnAA active site causes the rotation of the side chain of Cys291. This removes the steric occlusion for the nicotinamide ring of the cofactor caused by the Cys291-Sγ atom in the apo state of the enzyme. Thus the cofactor can now be positioned so that the hydride of the thiohemiacetal intermediate points towards to the *re* face of the nicotinamide ring, and is aptly positioned to be transferred to the C4 atom of the nicotinamide ring. Subsequent to hydride transfer, the thiohemiacetal intermediate converts to a planar thioester. This structural change, in conjunction with the loss in planarity of the nicotinamide ring subsequent to hydride transfer causes the cofactor to leave the active site of the enzyme. The departure of the cofactor allows the catalytic base- Glu254 to be correctly positioned so that it can deprotonate a water molecule and position it correctly for nucleophilic

addition at the thioester intermediate. This addition generates a thiohemiketal intermediate which is resolved by the departure of the Cys291 side chain to yield PnA. Differences in the mode of binding of PnAA and PnA within the PhnY active site cause the departure of the product and the enzyme regenerates for the next catalytic cycle.

Cocrystal structure with phosphorylated aldehyde substrate G3P and cofactor NAD^+

In order to determine the catalytic repertoire of the PhnY enzyme, we explored dehydrogenation activity of PhnY for non-physiological phosphorylated substrates. A hypothetical phosphorylated mimic of PnAA would be unstable in aqueous buffers due to the presence of a phosphate anhydride, with demonstrated increase in its labile nature at neutral to alkaline pH (Lipmann, JBC, 1944). Hence, G3P (**3** in Figure 3.2) was chosen as a candidate substrate mimic. PhnY could indeed accept G3P as a substrate, as judged by the spectrophotometric assay (Figure 3.12), albeit with a much reduced k_{cat} value of 0.039 ± 0.002 sec^{-1}, which is nearly 1500 fold slower than G3P oxidation by *S. mutans* GAPDH (Cobessi et al., 2000) (Table 3.1). This demonstrates that the active site of PhnY is adapted for the oxidation of PnAA, as compared to the GAPDH active site which is adapted for G3P acceptance. The postulated structural basis of this substrate preference is delineated in this section. The low level of activity, coupled with the high enzyme concentration needed for detection, precluded the determination of K_M, as initial turnover rates could not be measured reproducibly at lower substrate concentrations.

The crystal structure of the PhnY enzyme in presence of G3P was determined by co-crystallizing the enzyme with 10 mM G3P and 1 mM NAD^+. Racemic mixture of G3P was

obtained from Sigma-Aldrich and directly used without further purification. Though the crystallization conditions remained the same, diffraction quality co-crystals of the enzyme could not be obtained with G3P alone, as with the physiological substrate PnAA described previously. Electron density could not be observed in the 2.1 Å co-crystal structure obtained for the entirety of the cofactor. Electron density for the cofactor is missing for the nicotinamide ring and the ribose sugar connected to it. As discussed previously, this lack of electron density has previously been observed for other members of the ALDH family (Cobessi et al., 2000; D'Ambrosio et al., 2006; Langendorf et al., 2010).

Unambiguous robust electron density for all atoms of the G3P substrate molecule can be observed in the active site (Figure 3.13). No major secondary or tertiary structural rearrangements are required for the accommodation of G3P, as compared to the binding of PnAA in the active site. The phosphate group bearing carbon atom of G3P is denoted as Cα, followed by the Cβ bearing the hydroxyl, and the Cγ bearing the aldehyde group. The aldehyde carbon of G3P is covalently bonded to the side chain sulfur atom of Cys291-Sγ atom and is positioned 1.7 Å away. The stereochemistry around the G3P-Cγ atom, as determined by the bond angles, points towards a tetrahedral geometry. Though the complex with G3P has been obtained in the presence of the cofactor, this tetrahedral geometry indicates that the hydride transfer to the cofactor has not yet occurred. This probably is a consequence of 10 fold higher concentration of G3P than that of NAD^+ used during crystallization. The aldehyde oxygen atom, which is thus modeled as a hydroxyl is 2.7 Å away from the side chain amide of Asn158. No interactions can be observed for the Cβ-hydroxyl, and the bridging oxygen atom of the phosphate moiety to the Cα methylene carbon atom. The phosphate moiety of G3P is positioned analogously to the phosphonate group of PnAA, by hydrogen bond interactions with the basic triad formed by

Arg108 (Nη1 – O1 distance of 3.0 Å) (Nη2 – O1 distance of 3.3 Å), Arg290 (Nη2 - O2 distance of 2.9 Å) and Arg447 (Nη2 – O2 distance of 3.1 Å) (Figure 3.14).

Elasticity of the phosphate binding pocket; structural basis for substrate specificity in ALDHs

A comparison of the structure of PhnY-PnAA, with that of GAPDH-G3P-NADPH (PDB Code: 2ESD), provides the framework to probe the structural requirements necessary to differentiate between a phosphonate and phosphate ester within the ALDH active site. It should be noted that the cofactor bound to GAPDH in PDB- 2ESD is not in a productive hydride transfer conformation and is hence assumed not to provide any constraints upon substrate binding in the active site. A superposition of the PhnY-PnAA (Figure 3.14A) structure with that of the GAPDH-G3P-NADPH (Figure 3.14B) complex shows that the two substrates, PnAA and G3P, differ by the distance between the aldehyde carbon, on which catalysis occurs, and the phosphonate/phosphate head group that provides the major binding interactions. The G3P substrate in GAPDH is bound in a distinctly planar sp^2 conformation in all four monomers, and the stereochemistry at this atom was established based on control experiments that demonstrated that reduction of the cofactor had occurred within the crystals (D'Ambrosio et al., 2006).

While the two structures can be aligned with high conservation of secondary structural elements, a view of the phosphonate/phosphate binding pocket reveals that the selectivity for a particular substrate is built into the active site primarily by the relative positions of the three Arg residues (Figure 3.14). When the two catalytic site Cys (PhnY-Cys291 and GAPDH-Cys284) are superimposed, the aldehyde carbon and oxygen atoms of the substrate also align, along with the

side chain of PhnY-Asn158 and GAPDH-Asn154. However, due to a longer distance between the aldehyde carbon and the phosphate head group in G3P, GAPDH-Arg103 and GAPDH-Arg283 are significantly pushed away from the active site as compared to the corresponding residues- PhnY-Arg108 and PhnY-Arg290 respectively. GAPDH-Arg437 (analogous to PhnY-Arg447) is also distorted away from the PhnY phosphate binding site. A rationale can also be derived for the mutation of the highly conserved GAPDH-Tyr155 to a histidine (PhnY-His159). His159 provides an additional contact to one of the phosphonate oxygen atoms and a bulkier tyrosine side chain could not be accommodated at this position due to steric clash with the closely spaced phosphonate group in the PhnY active site. Thus, we hypothesize that the phosphate binding pocket is fairly elastic and this freedom to adopt slightly different conformations on a highly conserved secondary and tertiary structural core provides for the modulation of substrate preference for ALDH enzymes.

This argument hypothesizing the conformational elasticity of the phosphate binding pocket is lent further credence by two observations of the present study. Firstly, aligning the PhnY cocrystal structures in the presence of inorganic phosphate and substrate PnAA demonstrates a similar elasticity of the phosphate binding pocket (Figure 3.5 and Figure 3.7). The inorganic phosphate deviates significantly from the expected binding site delineated by the phosphonate group in the PnAA cocrystal structure. The phosphate is moved towards the catalytic Cys291 residue and makes contacts with Asn158 side chain as well. This lateral shift of the phosphate ion makes the Arg108 residue side chain to lose contacts with the phosphate and it can then be positioned away from the active site by stabilizing interactions with Asp111. Secondly, a comparison of the PhnY-PnAA structure with that of the PhnY-NAD^+-PnA structure (Figure 3.7 and Figure 3.9) demonstrates a movement of the phosphonate group of PnA in the

opposite direction; towards the phosphate binding pocket. As the PnA-Cβ cannot be covalently bonded to Cys291-Sγ, PnA is laterally shifted away from the catalytic site and towards the phosphonate binding site. This lateral shift in the PnA phosphonate oxygen atoms is accommodated by a corresponding shift in the Arg108 side chain. Rather surprisingly, of the four residues which comprise of the phosphonate binding pocket for PhnY, only Arg108 is placed on an α-helix, while His159, Arg290 and Arg447 are all placed on loops at the interface of the catalytic domain and the cofactor binding domain. Hence the modulation of the phosphate binding pocket elasticity primarily by the movement of the Arg108 residue was rather unexpected.

A comparison of the GAPDH and PhnY active sites with their respective substrates bound also provides a rationale for the requirement of a phosphate ester substrate (G3P) by GAPDH, while PhnY accepts a phosphonate substrate (PnAA). The bridging phosphate ester oxygen atom in G3P is coordinated by the side chain guanidine group of GAPDH-Arg437 residue. One of the Nη atoms of Arg437 is placed 2.6 Å away from this oxygen atom and thus forms a strong hydrogen bonding interaction (highlighted in red in Figure 3.14B). This interaction is absent in the PhnY-PnAA complex active site. The corresponding Arg residue for PhnY- Arg447 is positioned away from the bridging PnAA-Cα atom and the Nη atom of Arg447 side chain contacts one of the phosphonate oxygen atoms at a distance of 2.8 Å (Figure 3.14A). An additional interaction for G3P phosphate ester oxygen atom is also provided by the side chain of GAPDH-Tyr155 residue, while the analogously placed PhnY-His159 residue only hydrogen bonds to the phosphonate oxygen atoms of PnAA. Hence it can be postulated that the requirement for the acceptance of a phosphate ester substrate for an ALDH enzyme led to additional interaction of the enzyme active site residue side chains with the phosphate bridging

ester oxygen atom (highlighted in red in Figure 3.14B).These interactions have consequently been lost for a phosphonate molecule, thus providing an example of substrate directed evolution in the ALDH family.

Structurally restrained diversification of the substrate portfolio for ALDHs

It should be noted that although the substrates of PhnY and GAPDH resemble each other, PhnY is more than an order of magnitude slower than the reported kinetic values for *S. mutans* GAPDH (Marchal et al., 2000). This kinetic observation can be explained on the basis of the vicinity of the oxygen atom of the thiohemiacetal intermediate to other molecules of the substrate. For PhnY, the thiohemiacetal transition state Cβ-hydroxyl is positioned only 2.9 Å away from one of the phosphonate oxygen atoms (Figure 3.7). On the other hand, for GAPDH the corresponding hydroxyl has no such destabilizing steric/electrostatic repulsions from the G3P phosphate oxygen atoms. In addition, the G3P substrate has taken advantage of the free rotation possible around the G3P-Cβ carbon to position the 2' hydroxyl away from the oxygen atom of the thiohemiacetal intermediate. A similar arrangement is hypothetically possible for the PhnY substrate as well, as free rotation is possible at the PnAA-Cα atom. In such a hypothetical, more energetically favorable scenario for substrate binding in the PhnY active site, the phosphorus atom would rather be positioned where one of the hydrogen atoms is for the PnAA-Cα atom, and would thus be positioned away from the Cβ aldehyde oxygen atom. However, such a positioning of the phosphonate group would make the phosphonate oxygen atoms lose all contacts with PhnY-His159 and PhnY-Arg290 side chains. However, by analogy to the active site for GAPDH,

and by the kinetic data obtained in this study for the His159Ala and Arg290Ala mutations, such a hypothetical enzyme-substrate complex would not be catalytically proficient.

Another argument in support of the above discussion is provided by the comparison between G3P binding in the PhnY and GAPDH active sites. In GAPDH active site, G3P is fully extended with the aldehyde oxygen atom being more than 7.0 Å away from the nearest phosphate oxygen atom. However, for G3P binding in the PhnY active site, the tetrahedral angle between the G3P Cγ-Cβ-Cα atoms has been reduced to only 96.5 degrees (Figure 3.14C). Hence, we would like to propose that PhnY represents a classical example of how diverse catalytic activities have been built upon a conserved structural core in nature. This divergence of activity, concomitant with the preservation of structure at a kinetic cost for catalysis, has not been previously demonstrated in the ALDH superfamily.

Overall structure of apo wild type PhnA enzyme

The PhnA protein from *S. meliloti* 1021 was heterologously produced in *E. coli* with an amino-terminal hexahistidine tag. For crystallization, the tag was removed via thrombin protease cleavage. Initial crystallographic phases for the wild-type PhnA were determined by single isomorphous replacement using diffraction data collected on native crystals soaked with mercury salts, and the structure has been subsequently refined against data to 1.35 Å resolution. Relevant data collection and refinement statistics are provided in Table 3.2. The overall structure of PhnA is shown in Figure 3.15.

The structure of PhnA consists of two distinct domains, a core domain that is highly homologous within members of the AP superfamily, and a divergent capping domain that has

only previously been observed in the structure of NPP from *Xanthomonas axonopodis pv. citri* (Zalatan et al., 2006) and in the crystal structure of PnA hydrolase from *P. fluorescens* 23F (Kim et al., 2011). The core domain comprises residues Met4-Gly85, Asp104-Met253, and Ser376-Ala416 and consists of a seven-membered β sheet flanked by eight α helices on either side. The capping domain has a typical α/β/α fold and is composed of residues Lys254-Arg375. Of particular note is the β-loop-β appendage comprising of residues Ile86-Asn103, which extends from the helices of the core domain and contacts the central β sheet of the capping domain (Figure 3.15). This appendage is also present in the crystal structure of NPP (Zalatan et al., 2006), but absent in the structures of Aps hydrolyzing phosphate monoesters. The PhnA structures described here are more complete and of higher resolution than the previously reported structure of *P. fluorescens* 23F PnA hydrolase (PDB ID: 1EI6) (Kim et al., 2011), in which several residues proximal to the active site were disordered and consequently were not modeled. Of note the current structure presents complete modeling of all residues of the β-loop-β appendage and all surface loops of the core domain, along with alternate conformations of several residues, primary among which is the M2 metal binding residue, Asp29 (*vide infra*).

Active site metal ion

Members of the AP superfamily are characterized by the presence of two requisite metal ions, and the nature of the metal-coordinating protein ligands is highly conserved (Coleman, 1992). Likewise, PhnA contains two metal ions (referred to as M1 and M2, consistent with nomenclature for this superfamily) with M1 coordinated by His215, His377, and Asp211 and M2 coordinated by Asp29, Asp250, and His251 (Figures 3.16 and Figure 3.18). As in other members of the AP superfamily, metal M2 likely activates the catalytic Thr68 for nucleophilic attack at the

phosphorus atom (Ghosh et al., 1986; Kim et al., 2011; Schwartz and Lipmann, 1961) and metal M1 likely stabilizes the leaving group generated after hydrolysis and activates a water molecule for nucleophilic displacement at the phosphorus atom during hydrolysis of the phosphorylated enzyme intermediate (Coleman, 1992).

The bound metal ions in PhnA (as purified from heterologous over expression in *E. coli*) are zinc ions; their identity and location were determined using anomalous diffraction data collected at the zinc absorption edge (Figure 3.16A). Interestingly, the distance between the two metal ions was 4.6 Å. This distance is significantly greater than the distance between the metal ions in *E. coli* AP (4.26-4.28 Å; PDB ID: 1ED9) (Figure 3.16B) and NPP (4.26-4.36 Å; PDB ID: 2GSN) (Figure 3.16C), and to the best of our knowledge is greater than the distance reported for any other AP superfamily members (Stec et al., 2000; Zalatan et al., 2006).

In alkaline phosphatases, a third metal (typically magnesium) is located in the active site and is postulated to be critical for deprotonation of the serine side chain, and subsequent generation of the alkoxide nucleophile (Kim and Wyckoff, 1991). This third metal is absent in PhnA, as has also been observed in the structure of NPP (Zalatan et al., 2006). Absence of the third metal ion in NPP has led to re-evaluation of its role in phosphoryl group transfer reactions mediated by APs. More recent studies comparing the reactivities of monoester and diester substrates suggest that the magnesium ion stabilizes a non-bridging oxygen atom in the transition state for phosphate monoester hydrolysis (Zalatan et al., 2008). In the PhnA structure, side chains of residues Cys27, Asp29, Asn72, Tyr206, and Thr208 form a hydrogen bonding network replacing the region corresponding to the third metal site in alkaline phosphatases. The mechanistic implications of the absence of this third metal in the structure of PhnA are not immediately clear.

Metal ion specificity

APs can utilize a vast array of divalent metal ions in the catalytic M1 and M2 positions, with the metal ion specificity subject to change by alterations in the active site amino acid side chains (Wojciechowski and Kantrowitz, 2002). To probe the specific metal ion preference for PhnA, we monitored the catalytic efficiency of the enzyme that was overexpressed and purified from *E. coli*, and after its incubation with Mg^{2+}, Zn^{2+}, Co^{2+}, Fe^{2+}, and Mn^{2+} (Table 3.3). As previously reported (Borisova et al., 2011), the enzyme purified from *E. coli* efficiently catalyzed the hydrolysis of PnA without addition of metal ions. When PhnA was incubated with Zn^{2+} prior to the activity assay, only a small improvement in k_{cat} was observed along with an increase in K_M. However, when PhnA was pre-incubated with Mn^{2+} or Fe^{2+}, two to three fold improvement in k_{cat} was observed with only small variations in the K_M values (Table 3.3). Activity assays in the presence of Mg^{2+} or Co^{2+} resulted in decreased k_{cat} values with a substantially lower K_M value of PhnA reconstituted with Co^{2+} compensating for the lower turnover number of this enzyme.

Enzyme purified from *E. coli* was treated with the metal ion chelator ethylenediaminetetraaceticacid (EDTA) to remove the metal ions already bound to the enzyme. The resulting apo-PhnA was devoid of enzymatic activity. However, the activity was fully reconstituted by brief incubation of PhnA with different divalent metal ions prior to the activity assays (Table 3.3). Thus, addition of Zn^{2+} produced a form of PhnA with the kinetic parameters nearly identical to those of the as-isolated enzyme. The greatest increase in the enzymatic activity was observed upon reconstitution of EDTA-treated PhnA with Mn^{2+} resulting in a k_{cat}/K_M value four-fold higher than that found for as-isolated PhnA. When combinations of equal concentrations of either Fe^{2+} and Zn^{2+} or Mn^{2+} and Zn^{2+} were tested for apo-PhnA reconstitution, the activity of the enzyme did not exceed that of Zn-reconstituted PhnA suggesting that it is

present predominantly as the Zn-bound form in either preparation and that PhnA has the greatest affinity for Zn^{2+}.

In order to probe whether differences in activity were a consequence of structural changes between the zinc and manganese substituted enzymes, we collected anomalous diffraction data on Mn^{2+}-substituted PhnA. Double difference Fourier maps were calculated which showed unambiguous positive density for metal ions in the M1 and M2 positions only at the manganese absorption edge assuring that both sites of the recombinant PhnA are occupied by manganese under the conditions used. The structure of Mn^{2+}-PhnA is virtually identical to that of Zn^{2+}-PhnA including the metal-metal distance (4.6 Å), suggesting that the structural results detailed below are valid for enzyme containing either metal. It should be noted that the values of kinetic constants observed for any of the metal-bound forms of PhnA (Table 3.3) are within the physiologically relevant range. Thus, it is likely that the identity of the PhnA-bound metal in vivo is strongly influenced by the intracellular levels of metal ions. It is likely that the recombinant isolation of PhnA as a zinc-bound form is reflective not only of its metal ion selectivity but also of the intracellular Zn^{2+}, Mn^{2+}, and Fe^{2+} ion levels, estimated in *E. coli* at approximately 100 μM for zinc and iron and 10 μM for manganese (Outten and O'Halloran, 2001). The intracellular levels of metal ions in *S. meliloti* 1021 have not been reported to date.

Model for a transition state structure

Hydrolysis of phosphate esters by APs takes place by an in-line double displacement mechanism in which the alkoxide ion, generated on a serine or threonine side chain, first attacks the tetracoordinated phosphorus atom. This displacement of a metal-stabilized leaving group by

the enzyme nucleophile occurs via a trigonal bipyramidal transition state, leading to the formation of a covalent phosphoseryl (or phosphothreonyl) intermediate (Coleman, 1992). Hydrolysis of this intermediate involves attack of a metal-bound hydroxide onto the phosphorus atom displacing the Ser/Thr via another trigonal bipyramidal transition state. The covalent adduct of orthovanadate ion with the catalytic serine or threonine side chain mimics the trigonal bipyramidal geometry of the transition state with the nucleophile and the leaving group occupying the axial positions; three of the oxygen atoms attached to the vanadium atom occupy the equatorial positions. Complexes with vanadate have been reported for both AP and NPP and have led to insights into the stabilization of the reaction transition state (Holtz et al., 1999; Zalatan et al., 2006). These structural data have facilitated the identification of stabilizing interactions for each of the equatorial oxygen atoms. For both enzymes, one oxygen atom is directed in between the metal ions and is stabilized by electrostatic interactions. Two other equatorial oxygen atoms in *E. coli* AP are stabilized by interactions with the guanidinium group of Arg166. For NPP, in which phosphate diester hydrolysis results in two, instead of three, non-bridging equatorial oxygen atoms, Asn111 (in *X. axonopodis*) interacts with one of the oxygen atoms to stabilize the transition state.

In order to delineate the interactions that occur in the transition state for phosphonate C-P bond hydrolysis by PhnA, we have determined the co-crystal structure of a covalent complex with orthovanadate at 1.65 Å resolution (Figure 3.17A and Figure 3.20). As observed for NPP (Zalatan et al., 2006) and AP (Holtz et al., 1999), the trigonal bipyramidal structure of the orthovanadate adduct is slightly distorted, with the O-V-O bond angle for the axial oxygen atoms being 168^{o} in PhnA compared to 157^{o} for NPP and 170^{o} for AP. But unlike AP and NPP, the three equatorial oxygen atoms are not in a plane with the vanadate atom but tilted towards the

M2 ion (Figure 3.20). The side chain hydroxyl of Thr68 is placed in one of the axial positions and the oxygen atom at the other axial position interacts with the zinc ion in the M1 site, which corresponds to the leaving group stabilization by the metal ion in a phosphate ester hydrolysis reaction. One of the equatorial oxygen atoms is situated in between the zinc atoms at a distance of 1.8 Å to M2 and a longer distance of 2.6 Å to M1, while the second equatorial oxygen is within hydrogen bonding distance (2.7 Å) of the amide bond nitrogen of Thr68. However, unlike NPP, where an additional hydrogen bond from the side chain of Asn111 further stabilizes the equatorial oxygen, no appropriately placed amino acid side chains are within hydrogen bonding distance. A water molecule coordinated by the side chain of Asn89 is placed 3.5 Å away from this equatorial oxygen and this water molecule additionally interacts with the backbone nitrogen of Asn69 (Figure 3.20). When Asn89 was replaced with valine the catalytic activity of the PhnA-Asn89Val mutant was reduced approximately 10^4-fold as compared to the wild-type PhnA (Table 3.4). Crystal structure of the Asn89Val mutant determined at 1.8 Å resolution revealed that both metal ions are preserved in the mutant structure and no major rearrangements of the amino acid side chains around the active site could be observed. However, the arrangement of the water molecules in the active site is different from that of the wild-type enzyme (data not shown) supporting the contribution of the water-mediated contact to Asn89 side chain to transition state stabilization. In AP, the third equatorial oxygen atom is hydrogen bonded to Arg166; in NPP, this oxygen carries a substituent in the phosphodiester substrate and does not have any interaction partners in the NPP crystal structure. In PhnA, this third oxygen makes water-mediated contacts in the PhnA-vanadate structure (Figure 3.20). A water molecule, hydrogen bonded to the carboxylate side chain of Asp29 is positioned 3.2 Å away.

Crystal structure of PhnA with acetate bound

Diffraction data from crystals of PhnA soaked with PnA reveal the appearance of electron density consistent with an acetate ion bound to the M1 metal ion (Figures 3.17B and 3.19). Two lines of evidence suggest that the resultant electron density for acetate derives from turnover of the PnA substrate by PhnA: first, diffraction data from crystals grown under identical conditions but without the addition of PnA do not show density for acetate; second, no contamination of acetate could be detected in samples of PnA used for co-crystallization as determined by proton NMR spectroscopy.

In the 2.0 Å resolution PhnA-acetate co-crystal structure, both oxygen atoms of the acetate ion are coordinating to the M1 metal ion at a distance of 2.1 Å (Figure 3.19). Coordination of the two oxygen atoms of the carboxylate moiety of acetate by the M1 metal ion places the methyl group of acetate 2.4 Å from the M1 metal ion and 3.9 Å from the M2 metal ion. Although not observed in the structure, the spacing between the acetate and M2 metal ion is sufficient to accommodate a phosphate group covalently bound to Thr68. One water molecule is positioned 2.85 Å away from the methyl group of the acetate and is stabilized by hydrogen bond interactions with the backbone carbonyl of Ile287 and backbone amide nitrogen of Asn69.

Co-crystal structure of PhnA-Thr68Ala with PnA and postulated reaction mechanism

Attempts at co-crystallization of wild-type PhnA enzyme with substrate, or soaking of apo wild-type enzyme crystals with PnA were unsuccessful as across all concentrations of substrate employed, catalytic turnover of the substrate lead to incorporation of the acetate ion in the active site of the enzyme and no density of the phosphate moiety could be observed. Hence

co-crystallization of PnA with Thr68Cys, Thr68Ala, and Thr68Gly mutants (exhibiting approximately 10^5, 10^3, and 10^5-fold decreases in activity, respectively, Table 3.4) was attempted. Of these, crystals could not be obtained for the Thr68Gly mutant enzyme, while crystals of the Thr68Cys enzyme led to incorporation of malate ion from the crystallization condition into the active site (data not shown). Crystals could also be obtained for the Thr68Ala mutant and showed robust density for PnA in the active site (Figures 3.17C and 3.21). The identity of the ligand is affirmed by the presence of strong electron density at a position corresponding to the electron dense phosphorus atom of the substrate. The phosphonate group of the substrate molecule is bound in a manner consistent with the transition state model predicted by the vanadate covalent crystal structure. The phosphorus atom is 3.4 Å away from the side chain methyl group of alanine at position 68 (corresponding to the position of the Thr nucleophile). One oxygen atom of the phosphonate of PnA points towards the M2 metal ion and is positioned at a distance of 2.0 Å. A second oxygen atom is hydrogen bonded to the backbone amide of Ala68, 2.7 Å from the amide nitrogen. A water mediated hydrogen bond exists between Asp29 and the third oxygen atom (Figure 3.21). Another equidistant water molecule is also hydrogen bonded to the side chain of Asn69. Superposition of the apo wild-type PhnA structure to the PhnA-Thr68Ala-PnA structure positions the catalytic threonine side chain hydroxyl 1.8 Å from the phosphorus atom and the theoretical bond angle between the hydroxyl oxygen, phosphorus atom and methylene group of PnA is 163^o, close to the 168^o seen for the vanadate transition state mimic. For comparison, the bond length between the vanadate atom and the catalytic threonine hydroxyl oxygen in the covalent complex is 2.0 Å. Thus the position of the phosphonate group in the Thr68Ala mutant is slightly skewed towards the threonine hydroxyl as compared to the vanadate transition state model.

The acetate moiety of the substrate is bound in a different manner to the Thr68Ala mutant compared to the acetate-bound state of the wild-type enzyme described above. While the acetate ion binds to the M1 metal ion in a bidentate manner in wild-type PhnA, the substrate acetate moiety binds to M1 of the Thr68Ala mutant in a monodentate fashion with one of the oxygen atoms at a distance of 2.5 Å from M1 and the second oxygen pointing away from the metal, and hydrogen bonded to a solvent molecule. This water in turn is hydrogen bonded to the backbone carbonyl of Ile287 (Figure 3.21). The distance of the methylene group of PnA to the M1 metal ion is 2.2 Å, close to the observed distance of 2.4 Å between the M1 ion and the methyl group of bound acetate; the distance to M2 is 4.4 Å. As a consequence of binding of the substrate to both metals the carbon-carbon-phosphorus bond angle is quite small (102^{o}) with the carbon-phosphorus bond oriented in a way that upon its cleavage would allow electron delocalization into the π system of the metal-bound carboxylate. Hence, it appears that the bound conformation of PnA to the Thr68Ala mutant resembles the productive complex for catalysis.

Asymmetric stabilization of the transition state by slight alterations in the active site of PhnA

The suggested mode of monodentate stabilization of an enolate intermediate by a metal ion is reminiscent of mandelate racemase enzyme. Mechanistic and crystallographic studies of this enzyme have demonstrated that an enolate intermediate of mandelate is stabilized by a divalent magnesium ion that binds only one of the oxygen atoms of the intermediate. This oxygen also interacts with a Lys. The second oxygen atom of the enolate of mandelate is coordinated by a strictly conserved glutamate side chain, and the enolate anion is further

stabilized by resonance with an aromatic ring (Kallarakal et al., 1995). In PhnA, only a water molecule is hydrogen bonded to the second oxygen of the carboxylate.

From our data, the stabilization of developing negative charge on the leaving group by the M1 metal ion appears to be a conserved feature of the alkaline phosphatase superfamily members, albeit the manner by which the active site geometry is used differs considerably from phosphate ester hydrolysis. Compared to AP and NPP, PhnA must activate a much weaker, carbon-based leaving group instead of an oxygen-based leaving group. As a measure of leaving group ability, the pK_a of the conjugate acid of the acetate enolate is likely to be >30 (Richard and Amyes, 2001) compared to a typical alcohol pK_a of ~15-18). In addition to the different leaving group ability, the charge distribution on the leaving group is also significantly different. During phosphate ester hydrolysis, significant charge build-up occurs on the bridging oxygen of the leaving group and this charge is stabilized by the M1 ion (Zalatan and Herschlag, 2006). For phosphonate hydrolysis, the charge is delocalized such that most of the charge likely resides on the oxygen atoms of the acetate enolate (Figure 3.22). This charge would be situated further from the oxygen of the threonine nucleophile coordinated to the M2 ion. Thus, a longer metal-metal distance appears to be required to simultaneously stabilize the developing charge on the leaving group with the M1 ion and activate the threonine nucleophile with the M2 ion. The involvement of Thr68 as a nucleophile is supported by a recent study on PhnA from *P. fluorescens* 23F (Kim et al., 2011) that reported loss of activity when the corresponding residue was mutated and that demonstrated labeling of the threonine hydroxyl with a phosphate group when incubated with γ–^{32}P-ATP, a slow substrate for PhnA. Thus, only the relatively long Zn-Zn distance allows the PnA substrate to simultaneously bind to both metals in a conformation that supports formation of an acetate enolate intermediate.

How PhnA can stabilize the much weaker leaving group with essentially the same active site architecture as that of AP and NPP to achieve effective catalysis is at present not clear. Unlike the extensive studies of enzymatic and non-enzymatic hydrolysis of phosphate monoesters and phosphate diesters, which have provided a detailed picture of transition state structures, very little is known about the transition state structure for C-P bond cleavage. Although at present it is not clear which features of the active site geometry of PhnA facilitate efficient catalysis beyond electrostatic stabilization of the leaving group, a plausible explanation can be offered for the previously reported poor activity for phosphate ester hydrolysis (Kim et al., 2011). AP has been shown to utilize strong electrostatic stabilization of the negative charge of the non-bridging oxygen atoms by the binuclear Zn^{2+} cluster (Nikolic-Hughes et al., 2005). On the basis of structures of AP with vanadate bound (PDB Code: 1B8J) (Holtz et al., 1999), one of the non-bridging oxygen atoms is believed to interact with both metals in the transition state (Figure 3.23). On the other hand, the vanadate structure of PhnA is decidedly non-symmetric with the corresponding equatorial oxygen, interacting much more strongly with M2 than with M1.

The formation of a stabilized enolate dianion was also recently proposed for PnA hydrolysis by *P. fluorescens* 23F PhnA based on structural and mechanistic studies (Kim et al., 2011). In that study, two lysine residues (Lys126 and Lys128) were implicated in furnishing potential stabilization of the enolate intermediate. These residues are conserved in *S. meliloti* 1021 PhnA (Lys130 and Lys132), however, they do not appear to be involved in any binding interactions with the active site ligands in any of the structures reported here. The active sites of *S. meliloti* 1021 and *P. fluorescens* 23F enzymes are very similar with the Zn-Zn distance slightly shorter in the latter (4.47 Å). However, a few differences exist between the two enzymes,

most notably the recombinant enzyme from *S. meliloti* 1021, as isolated from *E. coli*, is monomeric, whereas the enzyme from *P. fluorescens* 23F is a dimer (Kim et al., 2011). Detailed mechanistic studies will be required to distinguish between the two strategies of enolate stabilization proposed based on the structural information.

Crystal structure of inhibitor PnF-bound PhnA

Kinetic analyses in previous studies (Kim et al., 2011; McGrath et al., 1995) have identified PnF as an inhibitor of PnA hydrolase. Quinn and coworkers hypothesized that the inhibition of PnA hydrolase from *P. fluorescens* 23F could be due to the similar structures of PnA and PnF (McGrath et al., 1995). In order to delineate the inhibitory mechanism of PnF, we determined the crystal structure of PhnA in the presence of PnF to a resolution of 1.6 Å (Figure 3.17D). While no rearrangement of the metal binding side chains was observed, the distance between the zinc ions increased to 4.7 Å.

Compared to the vanadate binding site in the covalent complex described previously, the phosphonate moiety of PnF is skewed towards the M1 zinc ion. One oxygen atom of the phosphonate is stabilized by electrostatic interactions with the M1 ion positioned 1.85 Å away (Figure 3.17D). A water-mediated hydrogen bonding interaction also exists between a second oxygen atom of the phosphonate moiety and the side chain of Asp29. Similarly, the third oxygen atom interacts with the side chain of Asn89 through a water-mediated hydrogen bond. However, the phosphorus atom is rather far (3.4 Å) from the side chain of the catalytic Thr68 residue for nucleophilic attack of the alkoxide ion on this atom. The angle defined by the oxygen atom of the Thr68 side chain and the C-P bond is 71^{o}, suggesting that inhibitor binding does not mimic

substrate binding but rather that it binds in a non-productive manner. The carboxylate group is positioned towards the sites of stabilization of the vanadate equatorial oxygen atoms. One of the carboxylate oxygen atoms is 3.0 Å away from the backbone amide of Thr68 and is involved in a hydrogen bonding interaction, and the other oxygen atom is stabilized by a water molecule, which in turn is hydrogen-bonded to the side chain of Asn89 and backbone amide of Asn69. Hence, based upon the predicted reaction transition state, as delineated by the PnA co-crystal structure and vanadate covalent complex structure, inhibition of enzymatic activity by PnF does not seem to be dependent upon an analogous binding mode of the substrate and inhibitor molecules, but rather upon the tight but non-productive binding interactions of this molecule in the active site of the enzyme. In the crystal structure of PnA with AP (Holtz et al., 2000), which has a shorter Zn-Zn distance, PnA is bound in the active site in a conformation analogous to the PnF binding observed in this study for PhnA, hence acting as a competitive inhibitor of AP rather than a substrate.

Future directions

The productive hydride transfer complex described for PhnY in this study is a hypothetic model generated by the superpositioning of the PhnY-PnAA and PhnY-NAD^+-PnA crystal structures. Visualization of this state of the enzyme experimentally has been elusive, despite our exhaustive efforts in this regard. An attractive strategy would be mutation of the catalytic Cys291 residue to alanine, and then probe the crystal structure of the enzyme in the presence of PnAA and NAD^+. Though such a PhnY-Cys291Ala mutant would undoubtedly loose the covalent contact with the substrate, it would also alleviate the steric hindrance for the proper positioning of the cofactor nicotinamide ring in the active site. The molecular biological groundwork towards the realization of this aim has been laid, with the construction of the mutant plasmid, and purification of the PhnY-Cys291Ala enzyme. Crystallization of the ternary complex has not been achieved to date, as crystals could not be reproduced in the original condition. We aim to exhaustively screen crystallization conditions in aim to generate diffraction quality crystals of the ternary complex, and determine the high resolution co-crystal structure.

Several questions remain unanswered with respect to the involvement of the active site residues in the vicinity of the catalytic threonine residue of PhnA. Analogous to the model of transition state stabilization in AP and NPP, PhnA-Asn89 should ideally be suitably placed to make contacts with the phosphonate equatorial oxygen atoms. However, our present structural data identifies it to be too distant to be in direct contact with the oxygen atoms. The exact reasoning of why the phosphonate hydrolase active site diverges in this respect as compared to AP and NPP is not presently clear.

In order to address this issue, we aimed to generate crystal structures of mutant of PhnA in which the Asn89 residue has been replaced with a variety of hydrophobic, basic and acidic amino residues. An array of PhnA- Asn89Ala, Asn89Lys, Asn89Arg, Asn89Glu, Asn89Asp and Asn89Gln mutants was generated, and recombinant proteins purified and crystallized as for the wild type and Thr89Ala enzyme. We seek to determine the high resolution crystal structure of these mutants, and in conjunction with the kinetic characterization, aim to elucidate the exact catalytic role of these residues. A previous study delineating the crystal structure of *P. aeruginosa* PhnA enzyme also hinted at alteration in the kinetic properties of the enzyme by the mutation of two histidine residues which lie on the β-appendage. These histidine residues are also conserved in PhnA, though the present crystal structures do not shed light on their mode of involvement in catalysis. We also aim generate crystal structures of these mutants and probe in molecular detail the alterations in the PhnA active site due to their mutagenesis.

Tables

Table 3.1 Steady state kinetic parameters for PhnY wild type and mutant enzymes.

	K_M (μM)	k_{cat} (sec^{-1})
PnAA as substrate		
Wild type	3.2 ± 0.2	2.1 ± 0.4
Cys291Ala	n.d. [1]	n.d.
Glu254Ala	n.d.	n.d.
Asn158Ala	n.d.	$(4.3 \pm 0.4) \times 10^{-3}$
Arg108Ala	75.1 ± 10.4	$(25.7 \pm 0.9) \times 10^{-3}$
His159Ala	136.6 ± 12.4	$(16.3 \pm 1.4) \times 10^{-2}$
Arg290Ala	n.d.	$(4.8 \pm 0.8) \times 10^{-3}$
Arg447Ala	392.9 ± 27.8	$(35.1 \pm 4.9) \times 10^{-3}$
Glu385Ala	n.d.	$(6.9 \pm 0.9) \times 10^{-3}$
G3P as substrate		
Wild type	n.d.	$(3.9 \pm 0.2) \times 10^{-2}$
Cys291Ala	n.d.	n.d.

[1] n.d.: Not determined. Low levels of activity precluded kinetic parameter determination.

Table 3.2 PhnA Data collection, phasing and refinement statistics.

	PhnA (native)	PhnA-Vanadate	PhnA-PnA	PhnA-Acetate	PhnA-PnF
Data collection					
Space group	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$	$P4_32_12$
Cell dimensions					
a, b, c (Å)	111.8, 111.8, 72.8	111.1, 111.1, 72.4	111.4, 111.4, 72.8	111.4, 111.4, 72.9	111.6, 111.6, 72.5
Resolution (Å)[1]	50-1.35 (1.4-1.35)	50-1.8 (1.83-1.8)	40-2.0 (2.07-2.0)	50-2.0 (2.03-2.0)	50-1.6 (1.66-1.6)
R_{sym} (%)	6.8 (60.8)	6.5 (69.3)	7.5 (47.4)	7.1 (16.9)	8.0 (38.0)
I / σ(I)	36.8 (1.7)	36.2 (2.2)	32.1 (2.5)	47.5 (5.5)	33.0 (3.9)
Completeness(%)	98.3 (84.6)	98.0 (91.6)	99.4 (94.3)	100.0 (99.9)	99.3 (96.4)
Redundancy	10.4 (4.0)	10.9 (7.8)	11.2 (5.5)	11.3 (10.9)	11.3 (9.0)
Refinement					
Resolution (Å)	25.0-1.35	25.0-1.8	25.0-2.0	25.0-2.0	25.0-1.6
No. reflections	94,012	39,623	29,711	29,946	57,337
R_{work} / R_{free}[2]	19.7/20.9	20.5/24.1	18.9/23.4	17.6/21.9	19.3/21.6
Number of atoms					
Protein	3204	3194	3192	3199	3200
Metal	2	2	2	2	2
Ligand	-	5	8	4	7
Water	583	351	280	379	565

[1] Highest resolution shell is shown in parenthesis.

[2] R-factor = $\Sigma(|F_{obs}|-|F_{calc}|)/\Sigma |F_{obs}|$ and R-free is the R value for a test set of reflections consisting of a random 5% of the diffraction data not used in refinement.

Table 3.3 Kinetic parameters for PnA hydrolysis by PhnA in the presence of different metal ions (10 μM).

PhnA-N-His	k_{cat}, s^{-1}	K_M, μM	k_{cat}/K_M, $M^{-1}s^{-1}$	k_{cat}/K_M, rel.
As isolated[a]	0.9 ± 0.03	22.1 ± 2.3	4.1×10^4	1
$+Zn^{2+}$[a]	1.1 ± 0.02	36.8 ± 1.8	2.9×10^4	0.7
$+Fe^{2+}$	3.6 ± 0.1	43.0 ± 4.7	8.4×10^4	2.0
$+Mn^{2+}$	2.6 ± 0.1	33.7 ± 2.8	7.7×10^4	1.9
$+Co^{2+}$	0.6 ± 0.02	11.1 ± 1.6	5.1×10^5	1.2
$+Mg^{2+}$	0.2 ± 0.01	33.9 ± 4.8	0.5×10^4	0.1
Apo PhnA	No activity detected			
$+Mn^{2+}$	5.5 ± 0.2	33.7 ± 3.0	16.2×10^4	4.0
$+Fe^{2+}$	3.6 ± 0.2	44.1 ± 7.7	8.0×10^4	2.0
$+Zn^{2+}$	1.2 ± 0.03	30.3 ± 3.0	3.8×10^4	0.9

[a]Reported in (Borisova, et al., 2011)

Table 3.4 Initial rates for PnA hydrolysis by wild-type PhnA and by PhnA mutants.

PhnA-N-His variant	V/[E], s^{-1} 0.2 mM PnA[a]	V/[E], s^{-1} 1 mM PnA[a]
wild-type	3.8 ± 0.4	3.4 ± 0.5
Thr68Ala	$(2.4 \pm 0.2) \times 10^{-3}$	$(1.9 \pm 0.2) \times 10^{-3}$
Thr68Cys	$(5.3 \pm 0.7) \times 10^{-5}$	$(2.5 \pm 0.5) \times 10^{-5}$
Thr68Gly	$(1.7 \pm 0.4) \times 10^{-5}$	$(2.9 \pm 0.4) \times 10^{-5}$
Asn89Val	$(2.9 \pm 0.2) \times 10^{-4}$	$(3.3 \pm 0.2) \times 10^{-4}$

[a] The activity of these mutants was too low to obtain accurate kinetic parameters. Because the activities at two different substrate concentrations are very similar (with the exception perhaps of Thr68Gly), the values shown likely represent $k_{cat.}$

Figures

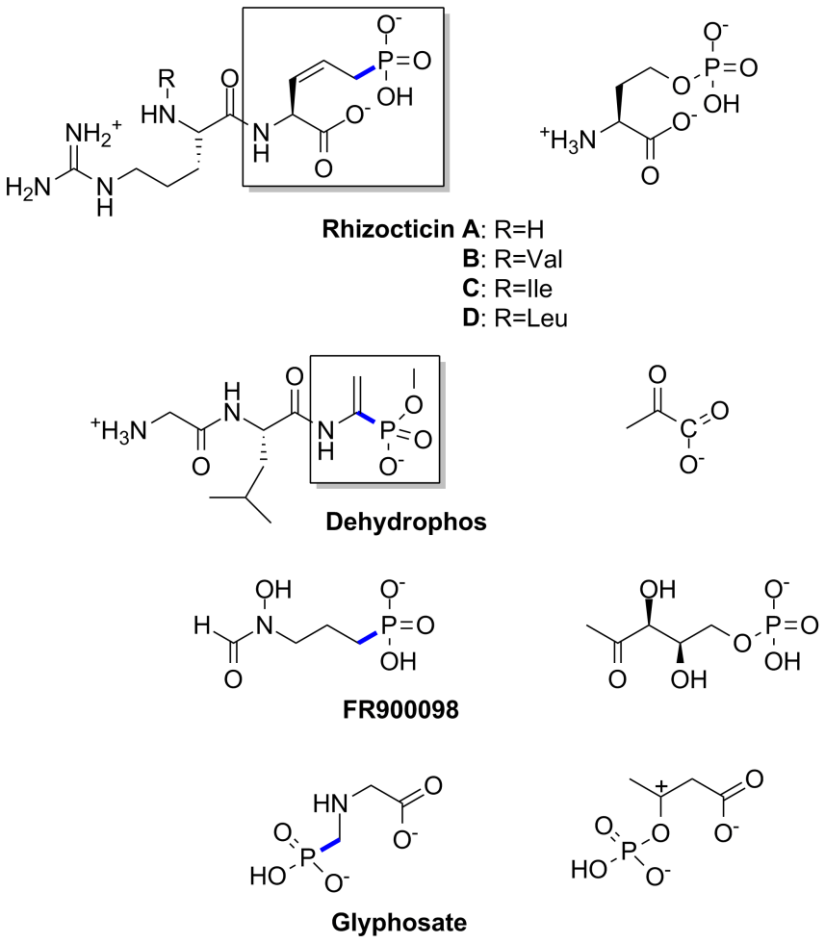

Figure 3.1 Mimicry of phosphate and carboxylate groups of enzyme reaction intermediates by phosphonate compounds. Hydrolysis of the rhizocticin (left) tri-peptide generates APPA (boxed) which mimics the intermediate for the threonine synthase enzyme (right). Hydrolysis of dehydrophos (left) and subsequent hydrolytic rearrangement of the 1-aminovinyl-phosphonate *O*-methyl ester (boxed) generates a structural analog of pyruvate (right) which inhibits pyruvate dehydrogenase and pyruvate oxidase. FR900098 (left) is an analog of 1-deoxy-d-xylulose 5-phosphate (right), which is a substrate for the isoprenoid biosynthetic enzyme deoxyxylulose reducto-isomerase. Widely used herbicide molecule glyphosate inhibits the EPSP synthase enzyme by mimicking the phosphoenol pyruvate oxonium ion (right) transition state of the enzyme. Phosphonate bonds are shown in blue. Bioactivity for these molecules is owed to the stability of the C-P bond, as compared to the labile nature of the O-P bond.

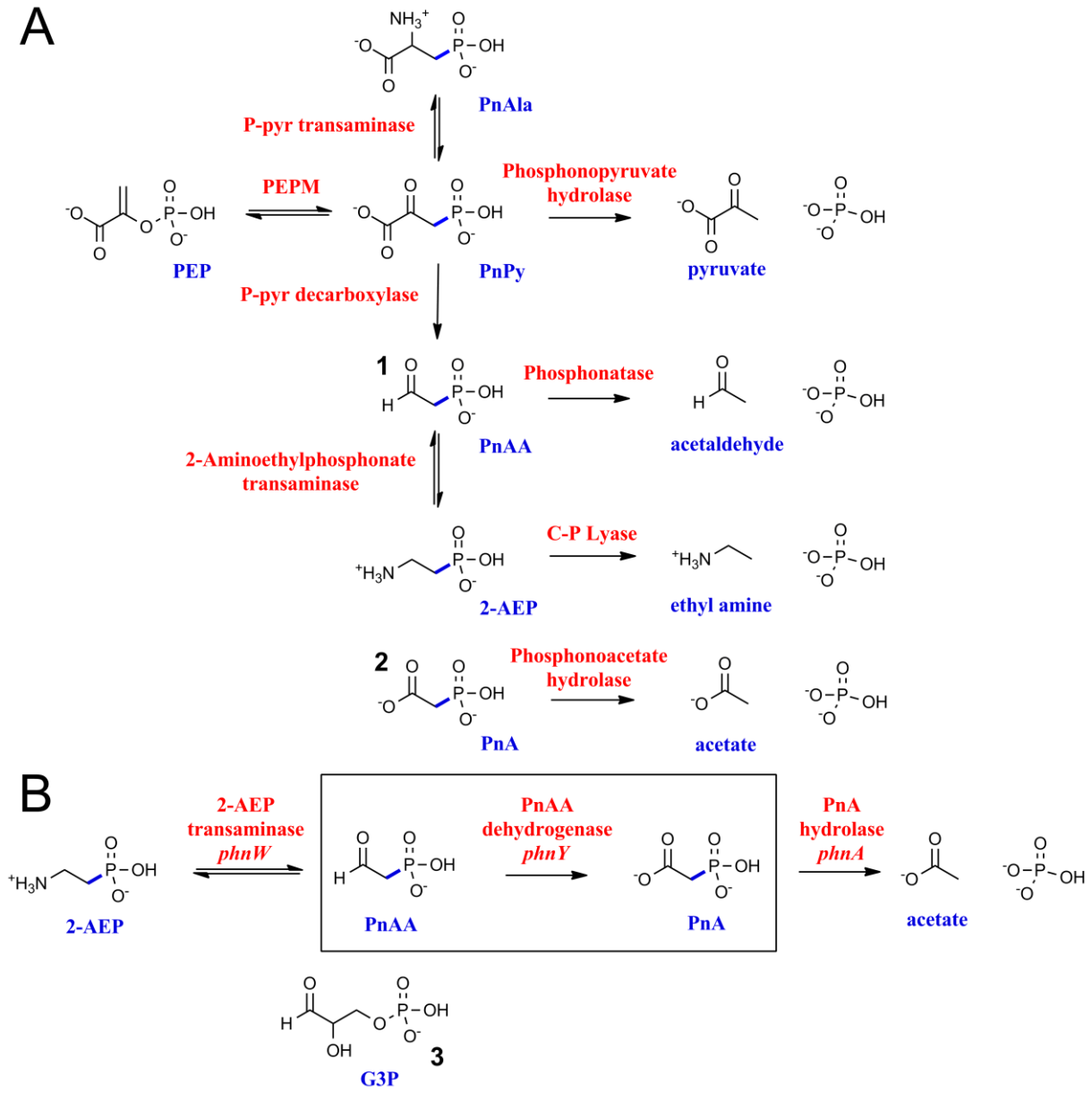

Figure 3.2 Biochemical pathways for microbial degradation of phosphonate molecules in nature. Phosphonate bonds are shown in blue. Enzymes are labeled in red and the compounds are labeled in blue. Reaction catalyzed by PhnY is boxed. **(A)** Phosphonate degradation pathways characterized previous to Borisova et. al., 2011. Note that only one reaction for the highly substrate promiscuous C-P lyase pathway is shown. **(B)** 2-AEP is converted to PnAA (**1**) by PhnW, which is oxidized to PnA (**2**) by PhnY and subsequently hydrolyzed by PhnA in a novel phosphonate degradation pathway in *S. meliloti* 1021. PhnY can also oxidize G3P (**3**) (*vide infra*).

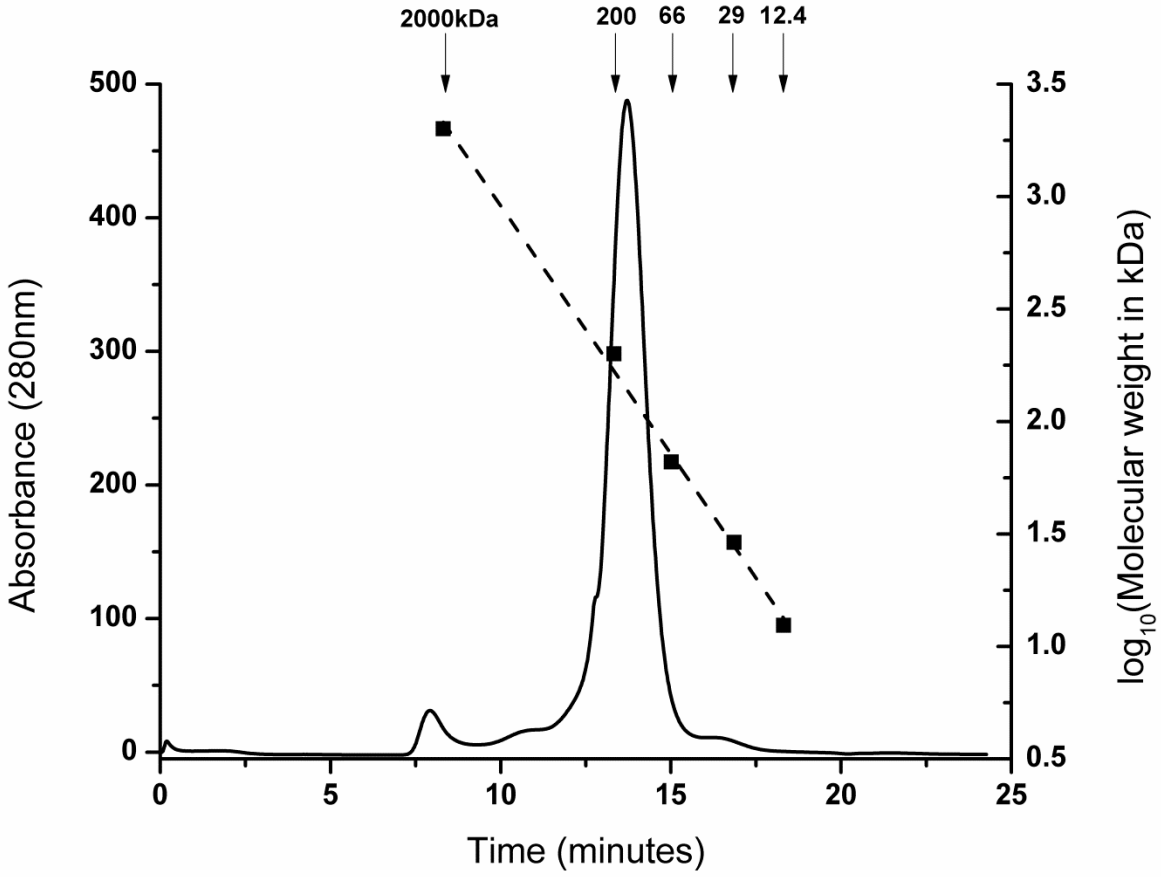

Figure 3.3 Size exclusion chromatography elution profile for PhnY. PhnY elution profile is shown as a solid line with linear fit (dashed line) of peak elution volumes for molecular weight standard proteins shown as square dots. Chromatography was performed using Superdex 200 HR 10/30 analytical gel filtration column (GE Lifesciences) in buffer 20 mM HEPES-Na (pH 7.5), 300 mM KCl, at a flow rate of 0.5 mL/min. The molecular weight standards were Blue Dextran (~2,000,000 Da), beta-amylase (~200,000 Da), Albumin (~66,200 Da), carbonic anhydrase (~29,000 Da) and cytochrome c (~12,400 Da) and were obtained from Sigma.

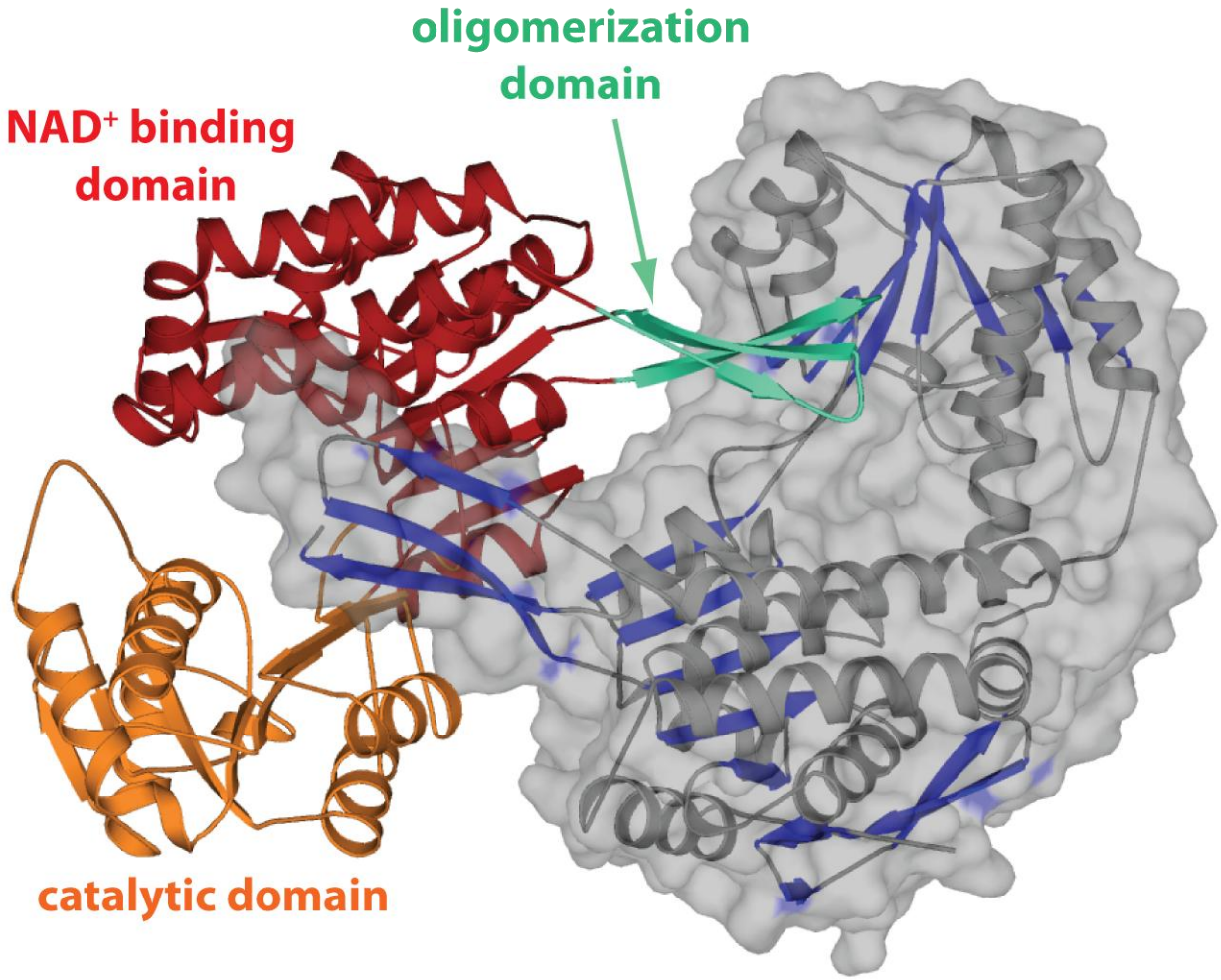

Figure 3.4 Overall structure of the PhnY dimer. The cofactor binding domain of one monomer is shown in red, the catalytic domain in orange, and the dimerization domain in cyan. The other monomer is shown in cartoon representation with a partially transparent surface superimposed. The β-sheets of the second monomer are colored blue.

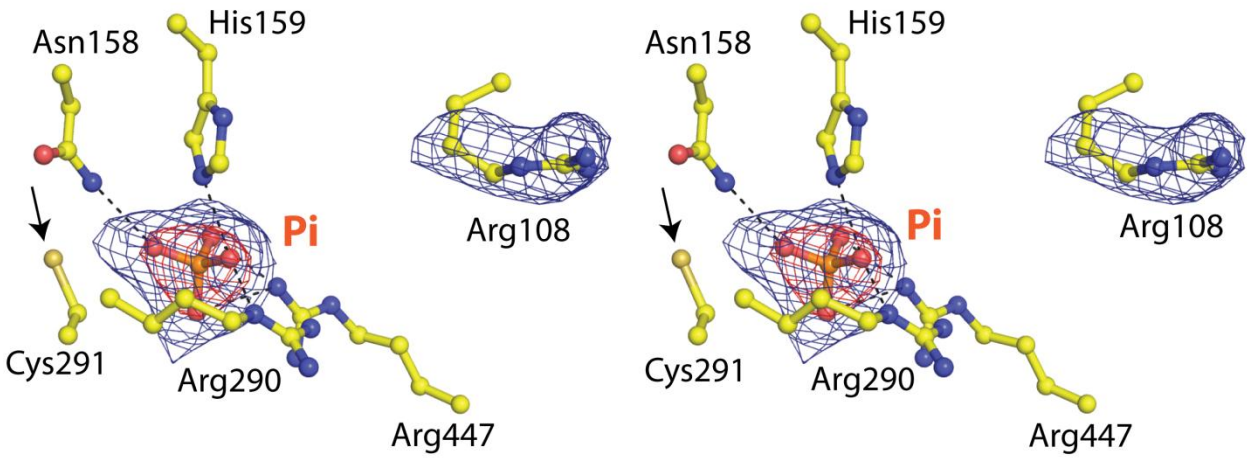

Figure 3.5 Stereoview of active site of PhnY in complex with inorganic phosphate (Pi). Pi and amino acid side chains are shown in stick-ball representation with the carbon atoms colored in yellow. The sulfur atom of Cys291 side chain is denoted by arrow. Superimposed is a difference Fourier electron density map (contoured at 3.0 σ over background and shown in blue and 9.0σ over background and shown in red) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of P_i and side chain of Arg108 deleted prior to one round of refinement. Hydrogen bonds are shown as black dashes.

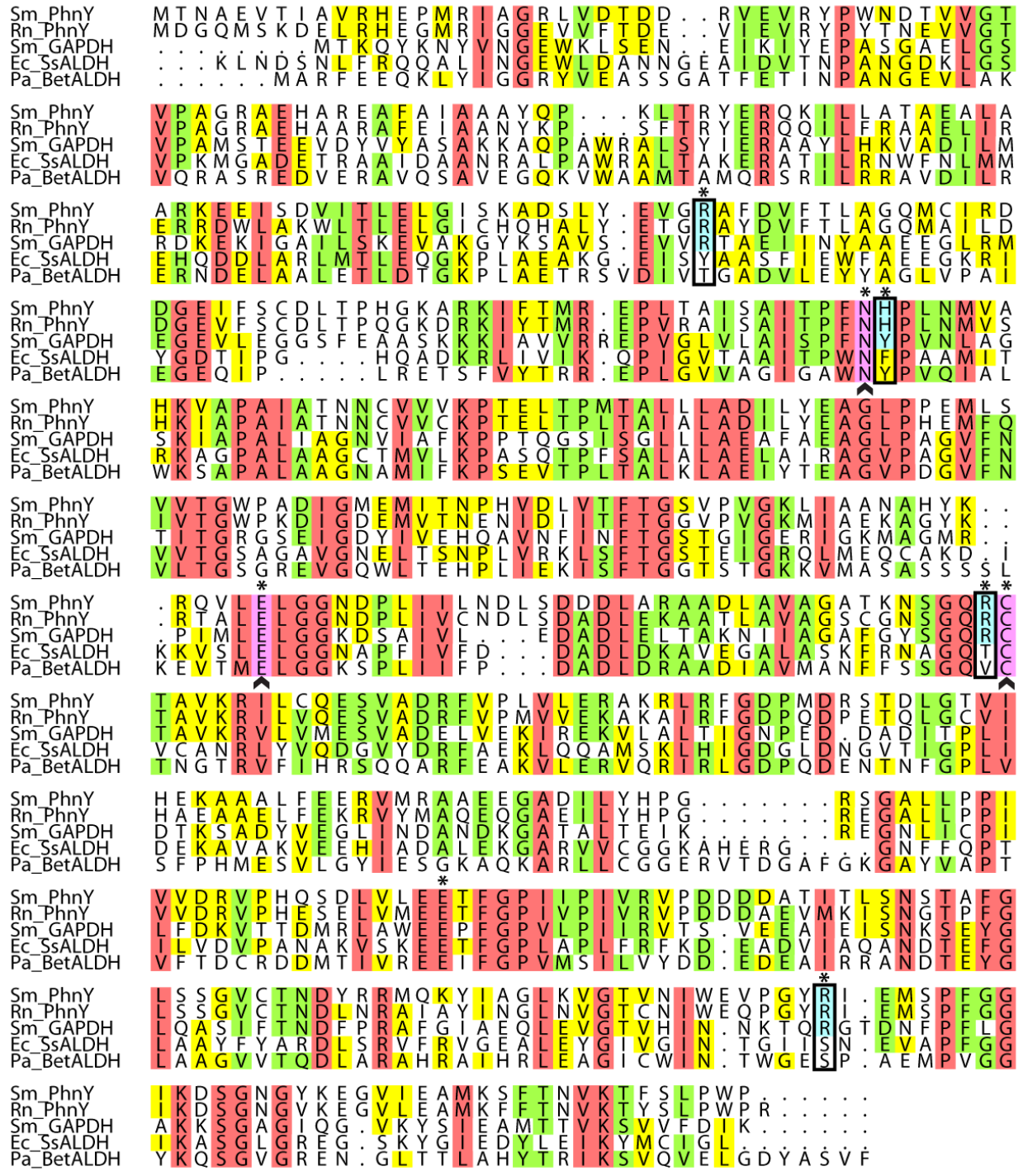

Figure 3.6 Primary sequence alignment between experimentally verified ALDH members. Sm_PhnY: *S. meliloti* PhnY, Rn_PhnY: *R. nubinhibens* PhnY, Sm_GAPDH: *S. mutans* GAPDH, Ec_SsALDH: *E. coli* succinic semialdehyde dehydrogenase, Pa_BetALDH: *P.aeruginosa* betaine ALDH. The absolutely conserved catalytic nucleophile cysteine, catalytic

(Figure 3.6 contd.)

base glutamate and the oxyanion stabilizing asparagine residues are highlighted in pink and marked by (^) under the sequences. The phosphate/phosphonate binding pocket forming arginine, histidine and tyrosine residues are highlighted in blue and boxed. These residues are not conserved in Ec_SsALDH and Pa_BetALDH. All residues of interest are highlighted by (*) above the sequences.

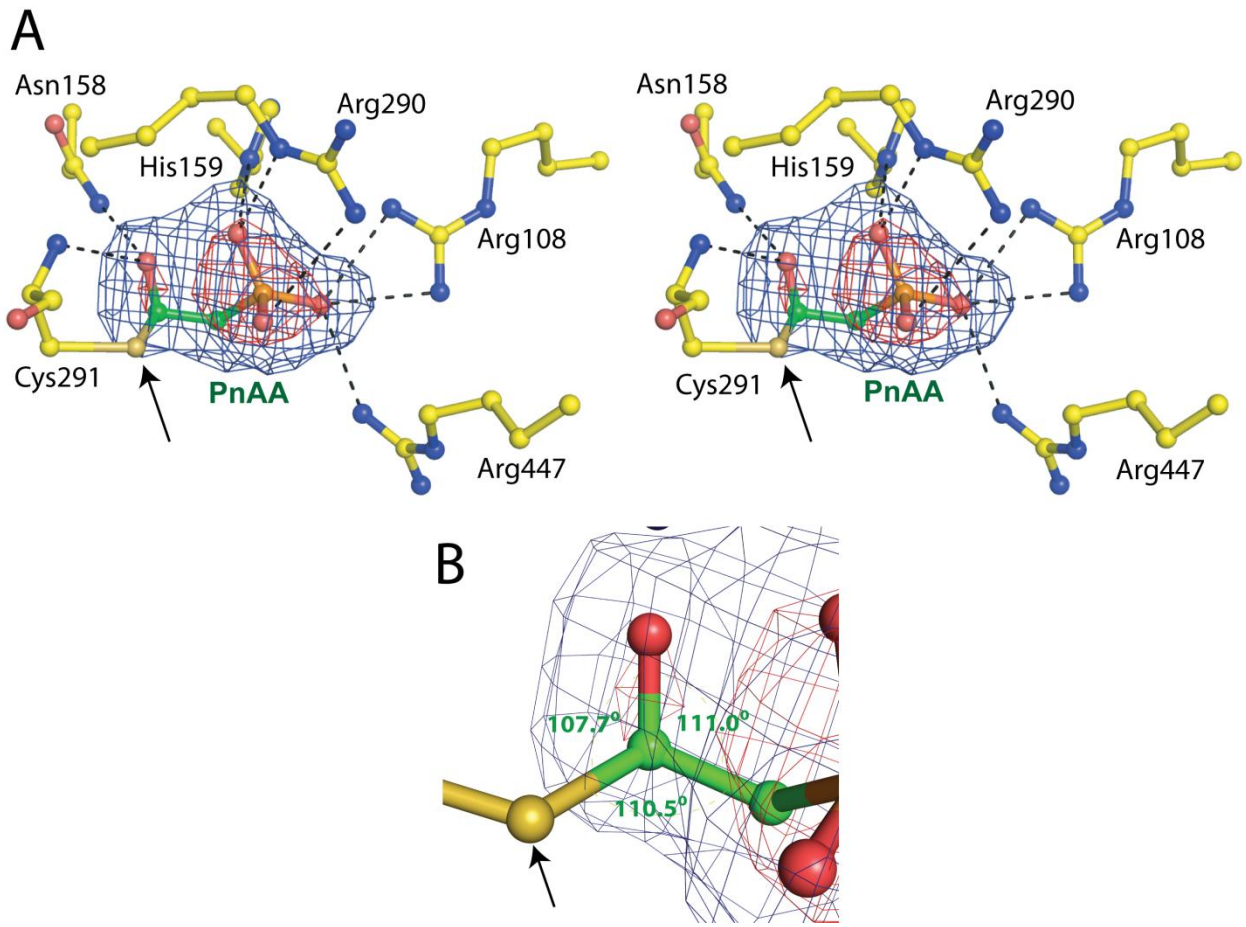

Figure 3.7 Active site of PhnY in covalent complex with PnAA. The PnAA substrate and amino acid side chains are shown in stick-ball representation with the carbon atoms colored in green and yellow respectively. The sulfur atom of Cys291 side chain is denoted by arrow. Superimposed is a difference Fourier electron density map (contoured at 3.0σ over background and shown in blue and 9.0σ over background and shown in red) calculated with coefficients $|F_{obs}|$ - $|F_{calc}|$ and phases from the final refined model with the coordinates of PnAA deleted prior to one round of refinement. Hydrogen bonds are shown as black dashes. **(A)**Stereoview of the PhnY active site, with **(B)** angles around the PnAA-Cβ atom labeled. Note that the angles around the PnAA-Cβ atom denote a tetrahedral geometry which can be achieved only prior to hydride transfer the cofactor NAD^+. Though not immediately apparent, the PnAA-Cβ atom in (B) is below the plane of the paper.

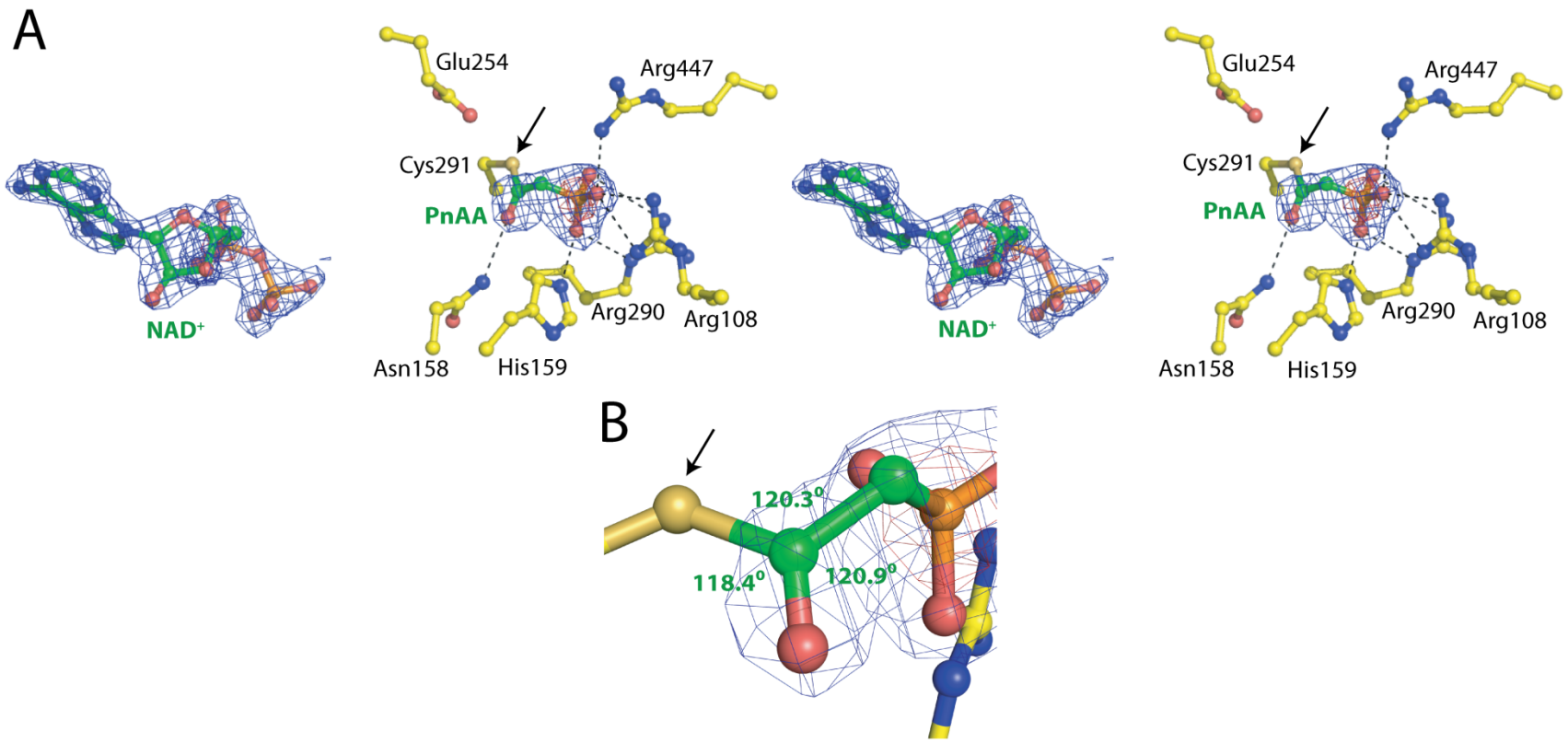

Figure 3.8 PhnY in covalent complex with substrate PnAA and truncated cofactor NAD$^+$. The PnAA substrate, NAD$^+$ cofactor and amino acid side chains are shown in stick-ball representation and colored as before. The sulfur atom of Cys291 side chain is denoted by arrow. Superimposed is a difference Fourier electron density map (contoured at 3.0σ over background and shown in blue and 9.0σ over background and shown in red) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of PnAA and NAD$^+$ deleted prior to one round of refinement. Hydrogen bonds are shown as black dashes.**(A)**Stereoview of the PhnY active site, with **(B)** angles around the PnAA-Cβ atom labeled. Note that the angles around the PnAA-Cβ atom denote a planar geometry which can be achieved only after hydride transfer the cofactor NAD$^+$.

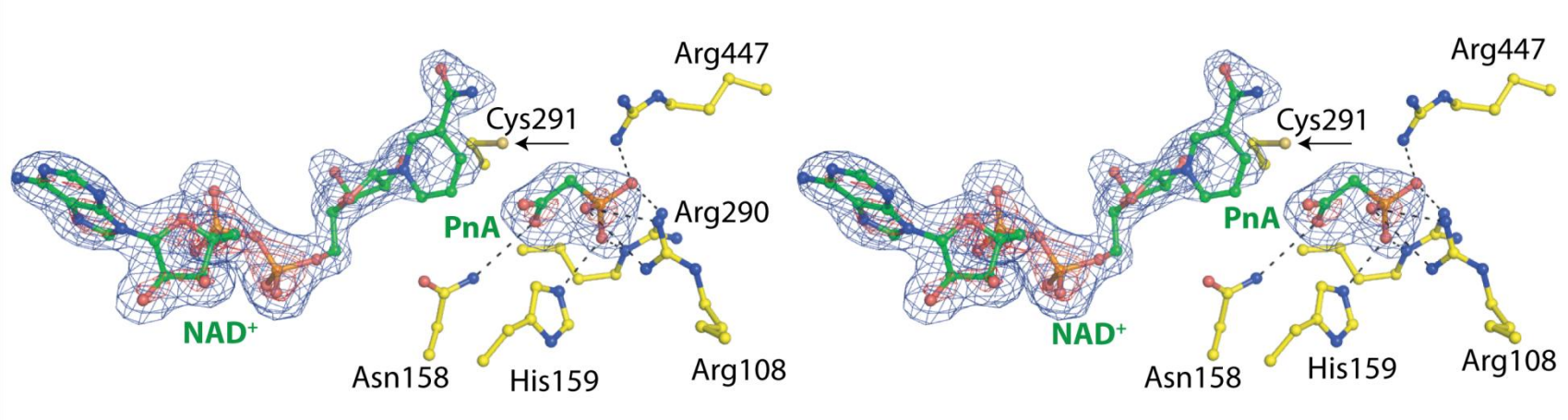

Figure 3.9 Stereoview of active site of PhnY in non-covalent complex with product PnA and cofactor NAD^+. The PnAproduct, NAD^+ cofactor and amino acid side chains are shown in stick-ball representation and colored as before. The sulfur atom of Cys291 side chain is denoted by arrow. Superimposed is a difference Fourier electron density map (contoured at 3.0σ over background and shown in blue and 9.0σ over background and shown in red) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of PnA and NAD^+ deleted prior to one round of refinement. Hydrogen bonds are shown as black dashes.

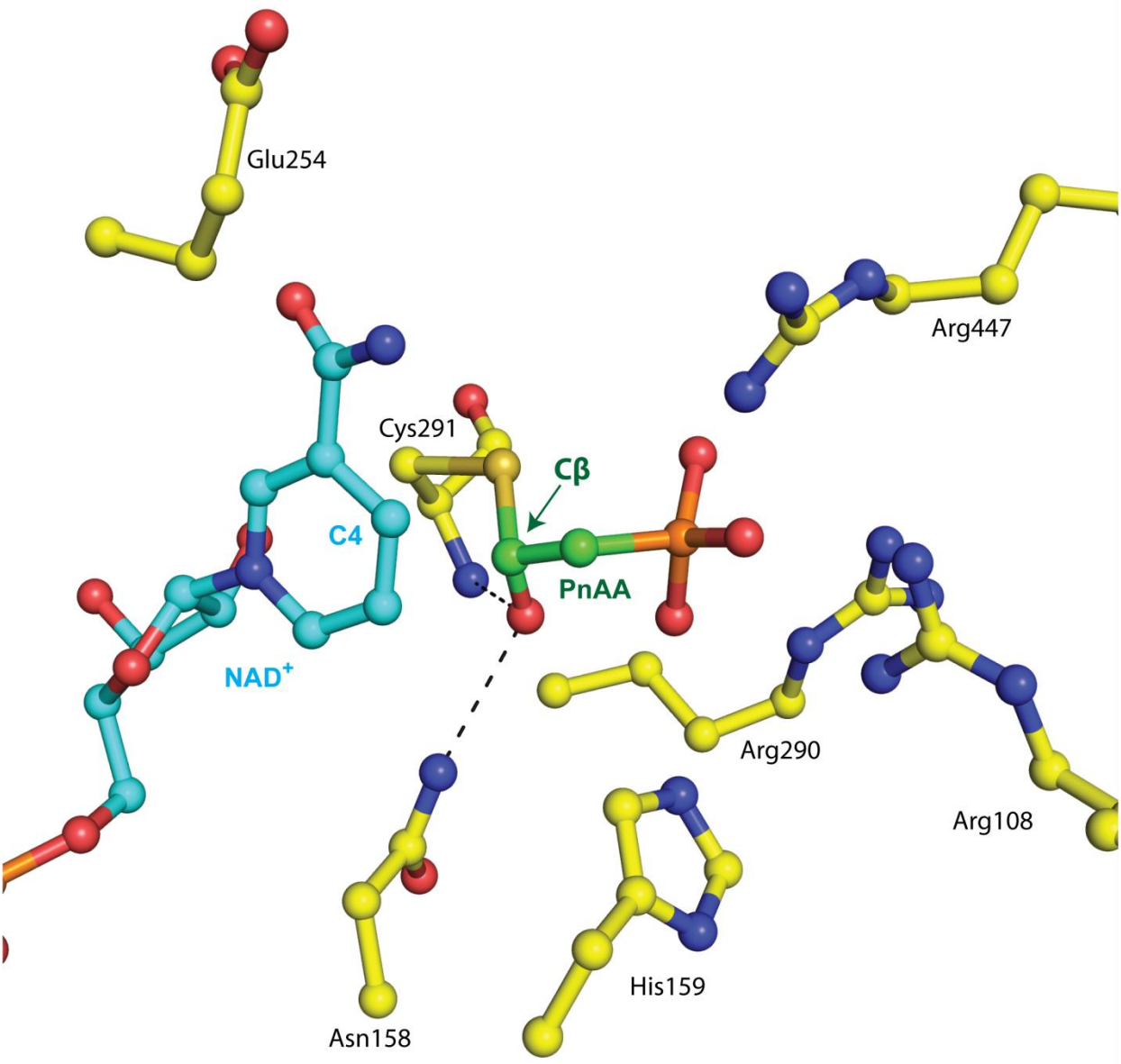

Figure 3.10 Model of the PhnY active site for hydride transfer from PnAA to NAD$^+$. The model is derived from the superposition of the PhnY crystal structure in the presence of PnAA and the PhnY crystal structure in the presence of PnA and NAD$^+$. Oxyanion stabilizing interactions for the PnAA Cβ hydroxyl are shown as black dashes. Note the relative positioning of the PnAA Cβ atom, and the C4 atom of the nicotinamide ring. Also the catalytic base- Glu254 is sterically occluded by the cofactor. Rotation of the Cys291 side chain thiol by 120° away from the active site would clash with the nicotinamide ring.

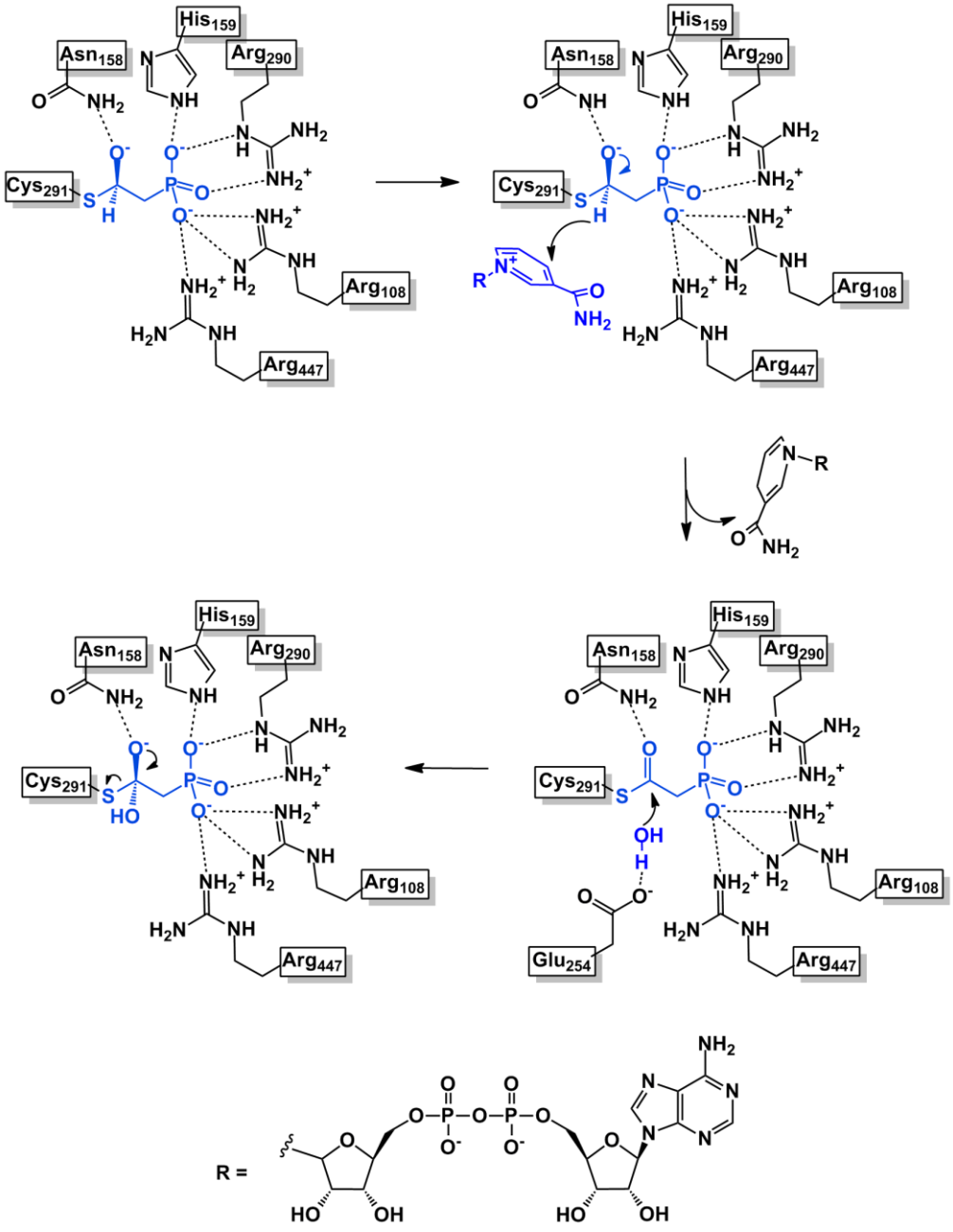

Figure 3.11 Proposed reaction mechanism for PhnY. Side chain thiol of Cys291 attacks at the aldehyde carbon atom of PnAA to generate a thiohemiacetal intermediate which is stabilized by hydrogen bonding with Asn158 side chain. A hydride is then transferred from the intermediate to the C4 atom of the nicotinamide rind of the cofactor. The cofactor departs from the active site allowing the catalytic base, Glu254 to correctly position and deprotonate a water molecule for attack at the Cys291 bound sp^2 hybridized carbon atom. The intermediate thus generated resolves itself by departure of the Cys291 side chain and PnAA is released.

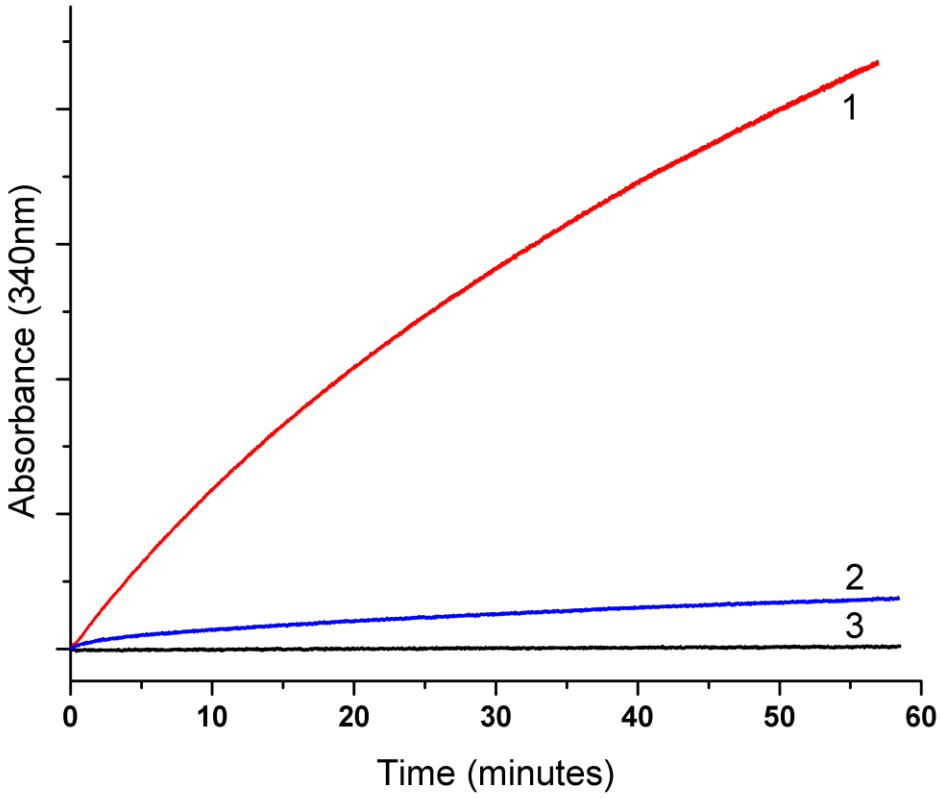

Figure 3.12 Relative activity of wild type PhnY to oxidize PnAA and G3P to PnA and 3-phosphoglycerate respectively. Activity is measured by absorbance increase due to production of NADH from NAD+ during the course of the reaction. Curve 1 (red) corresponds to PnAA as substrate, curve 2 (blue) to G3P as substrate and curve 3 (black) is a negative control without enzyme with PnAA as substrate.

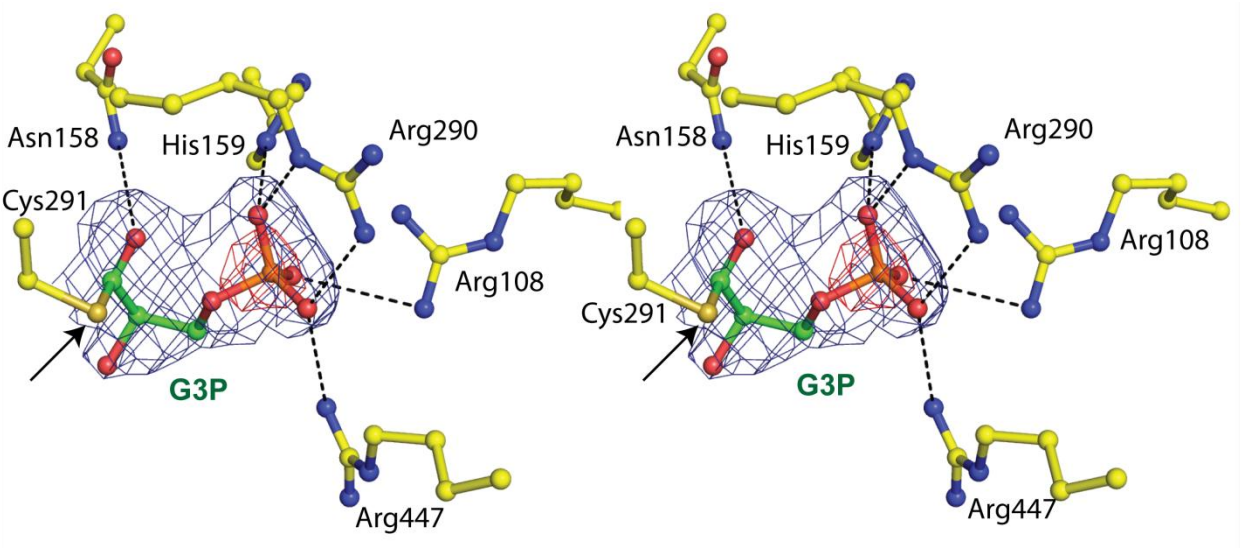

Figure 3.13 Stereoview of active site of PhnY in covalent complex with G3P. The G3P substrate and amino acid side chains are shown in stick-ball representation with the carbon atoms colored in green and yellow respectively. The sulfur atom of Cys291 side chain is denoted by arrow. Superimposed is a difference Fourier electron density map (contoured at 3.0σ over background and shown in blue and 9.5σ over background and shown in red) calculated with coefficients $|F_{obs}|$ - $|F_{calc}|$ and phases from the final refined model with the coordinates of G3P deleted prior to one round of refinement. Hydrogen bonds are shown as black dashes.

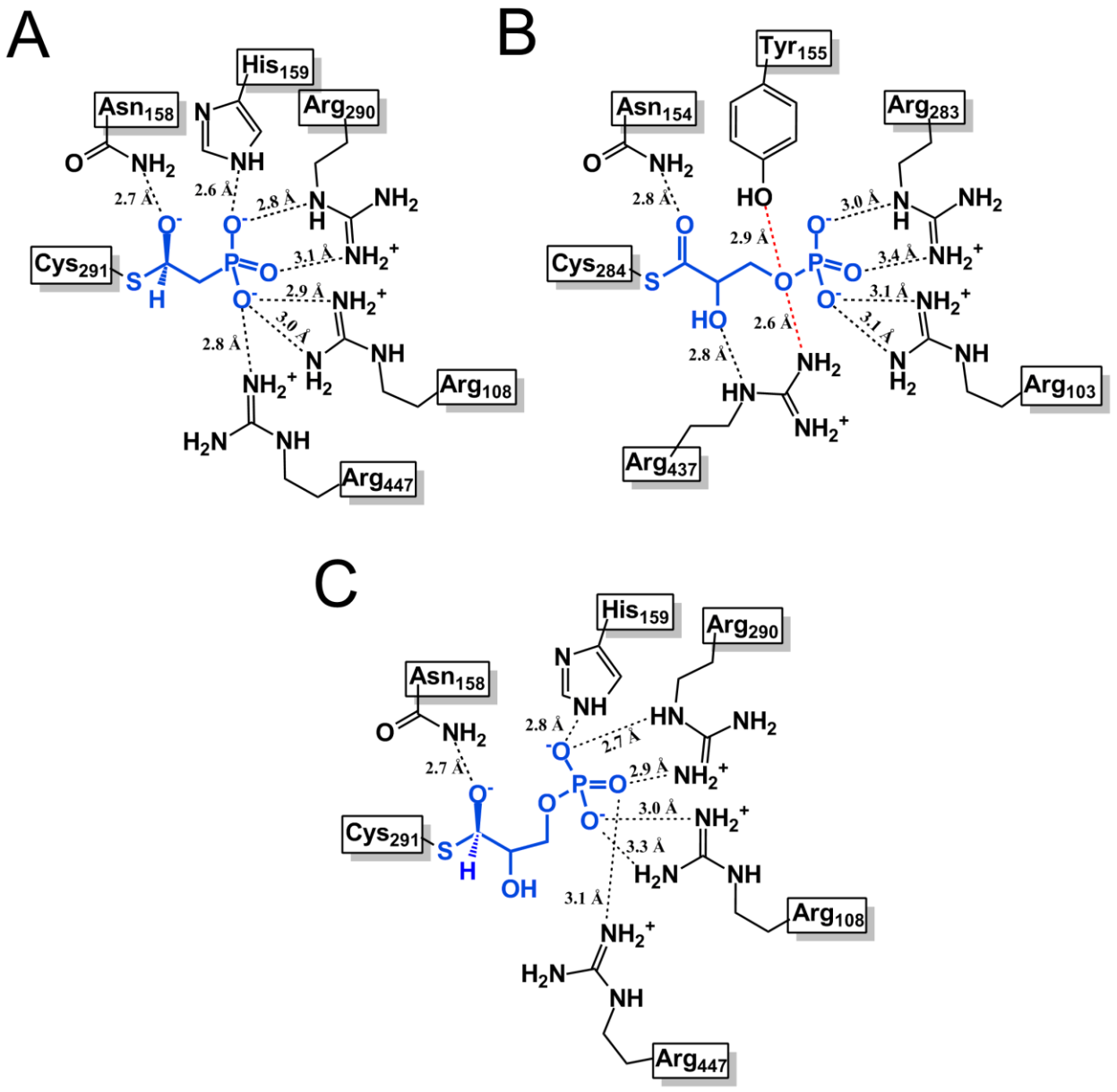

Figure 3.14 Active site arrangements for binding of (A) PnAA in PhnY active site, (B) G3P in GAPDH active site, and (C) G3P in PhnY active site. The substrate molecules are shown in blue. Putative hydrogen bonding interactions are shown as dashed lines. The distances between hydrogen bonded atoms are labeled. Note the extended nature of G3P in GAPDH active site, and the contorted nature of G3P in PhnY active site.

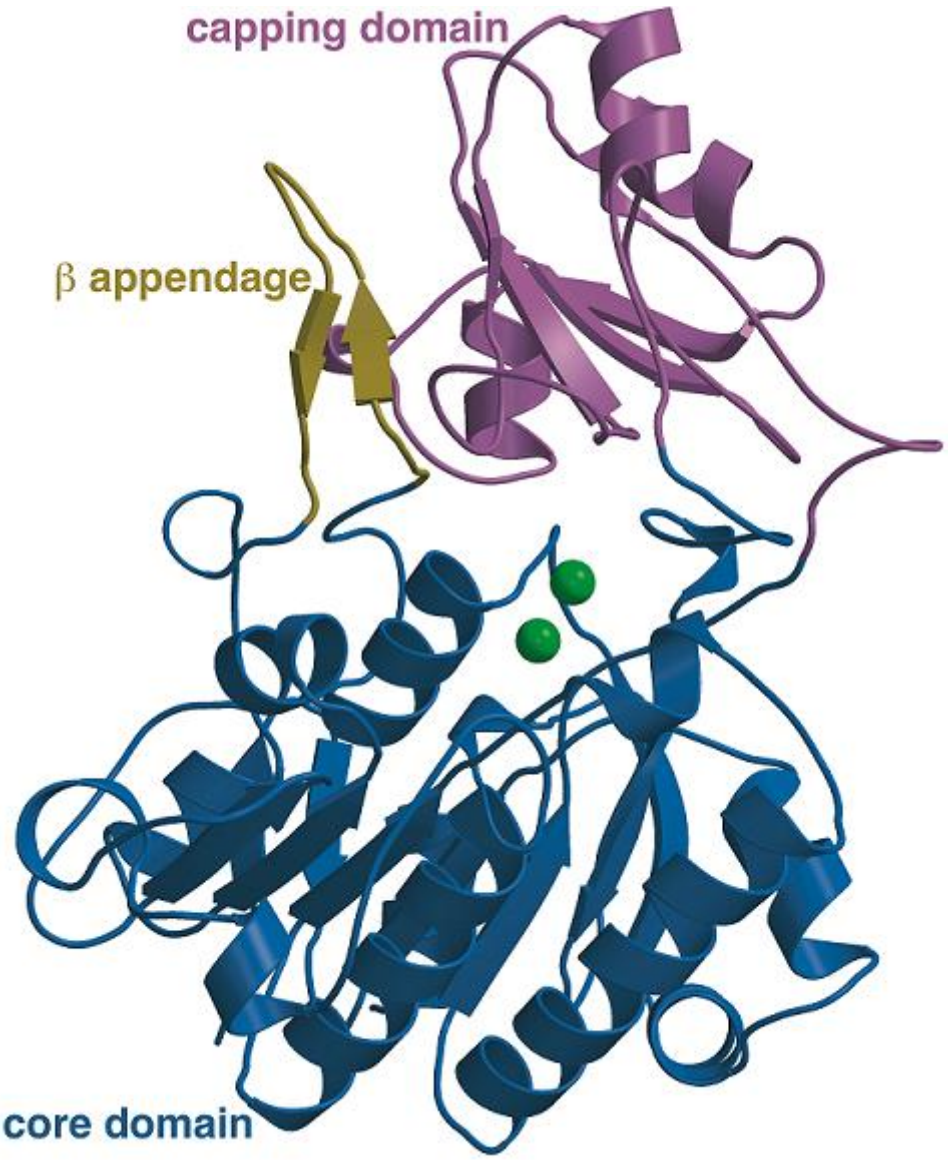

Figure 3.15 Crystal structure of *S. meliloti* PhnA. The core domain is shown in blue, the capping domain in pink and the beta appendage in gold. The two zinc ions bound at the interface of the core and the capping domains are shown as green spheres.

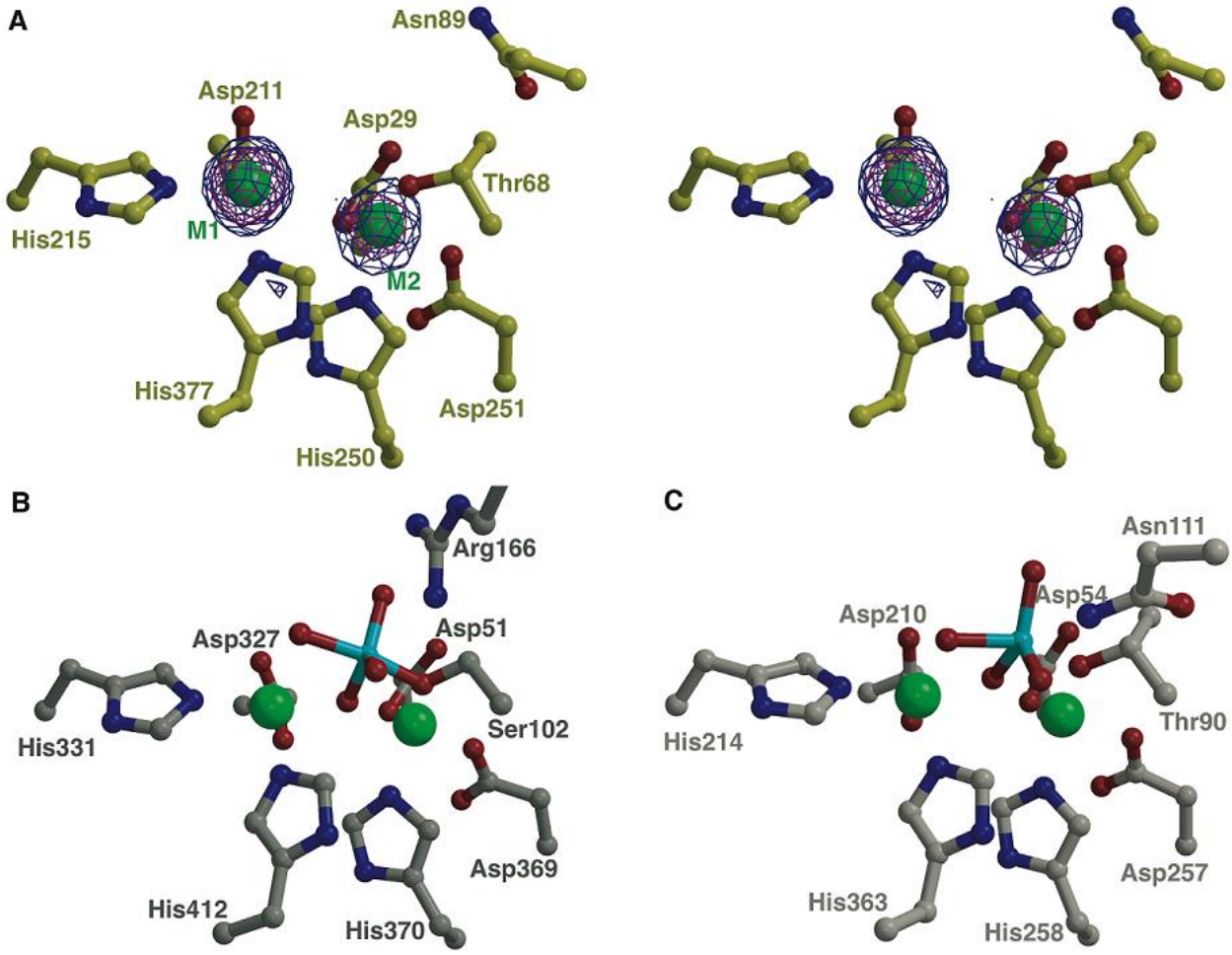

Figure 3.16 Active site view of PhnA as compared to AP and NPP.(A) Stereo view of PhnA active site showing a Bijvoet difference map, calculated using Fourier coefficients |F(+)| and |F(-)| from data collected at the zinc absorption edge and phases from the final refined model without any metal ions. The map is contoured at 3σ (in blue) and 10σ (in red).The coordinates from the final refined model are superimposed and the two metal ions are shown as green spheres. The active site view **(B)** of *E. coli* AP (PDB Code: 1B8J) and **(C)** of *X. axonopodis* NPP (PDB Code: 2GSO), both in complex with the inhibitor vanadate (colored in cyan) are shown for comparison. Water molecules are omitted for clarity.

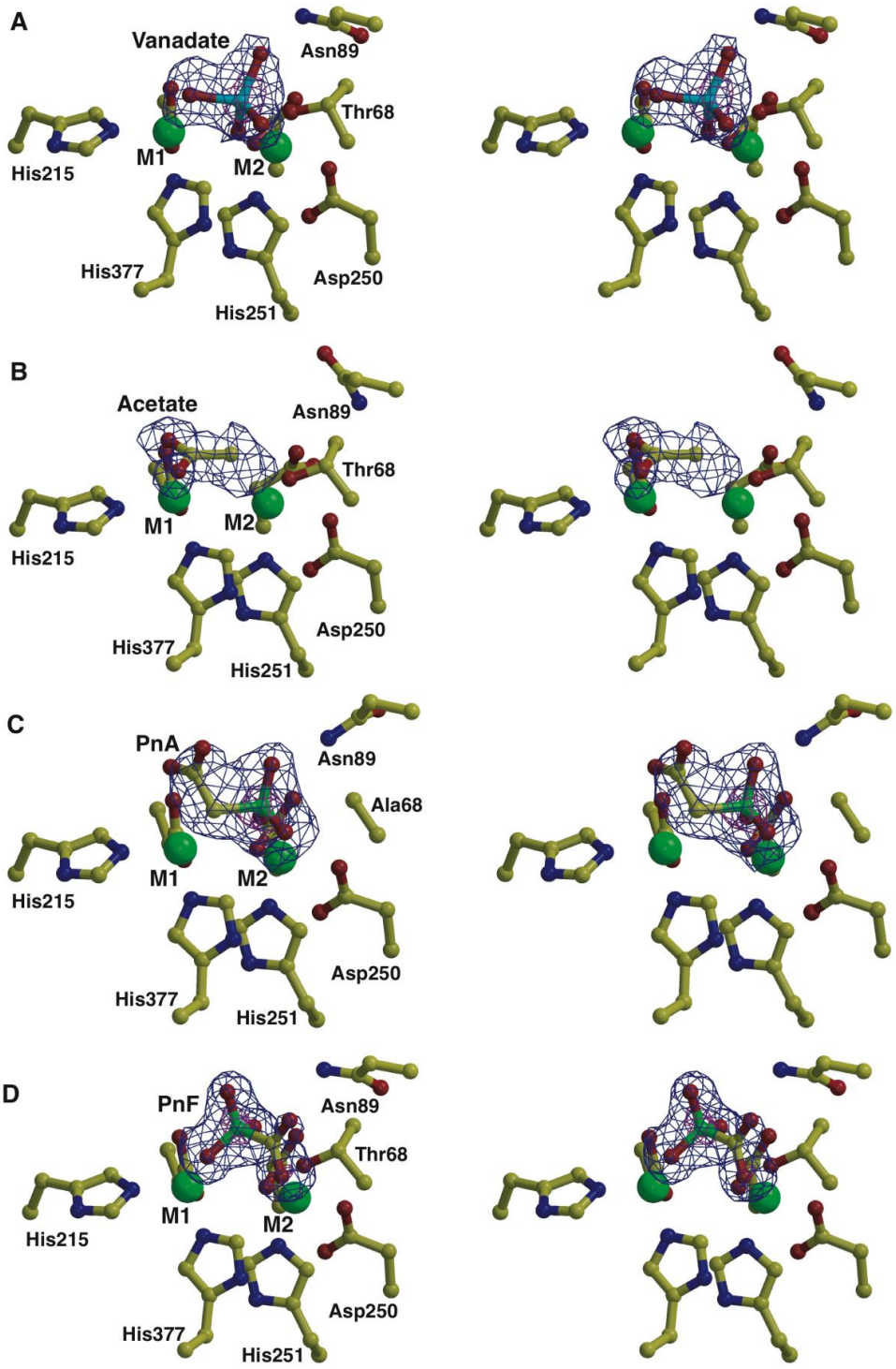

Figure 3.17 Stereo views of PhnA active site with various ligands bound.(A) Stereo view showing the active site features of the PhnA-vanadate structure. The vanadate is colored in cyan, polypeptide residues are shown in yellow and the active site metals are shown as green spheres. Superimposed is a difference Fourier electron density map (contoured at 3σ over background in

(Figure 3.17 contd.)

blue and 8σ over background in pink) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of vanadate deleted prior to one round of refinement. **(B)** Stereo view of the active site features of the PhnA-acetate complex. Residues are colored as above and superimposed is a difference Fourier electron density map (contoured at 3σ over background in blue) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of acetate deleted prior to one round of refinement. **(C)** Stereo view of the active site features of the PhnA-Thr68Ala-PnA complex. Residues are colored as above and superimposed is a difference Fourier electron density map (contoured at 3.3σ over background in blue and 10σ over background in pink) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of PnA deleted prior to one round of refinement. **(D)** Stereo view of the active site features of the PhnA-PnF complex. Residues are colored as above and superimposed is a difference Fourier electron density map (contoured at 3σ over background in blue and 8σ over background in pink) calculated with coefficients $|F_{obs}| - |F_{calc}|$ and phases from the final refined model with the coordinates of PnF deleted prior to one round of refinement. In all panels, water molecules are omitted for clarity.

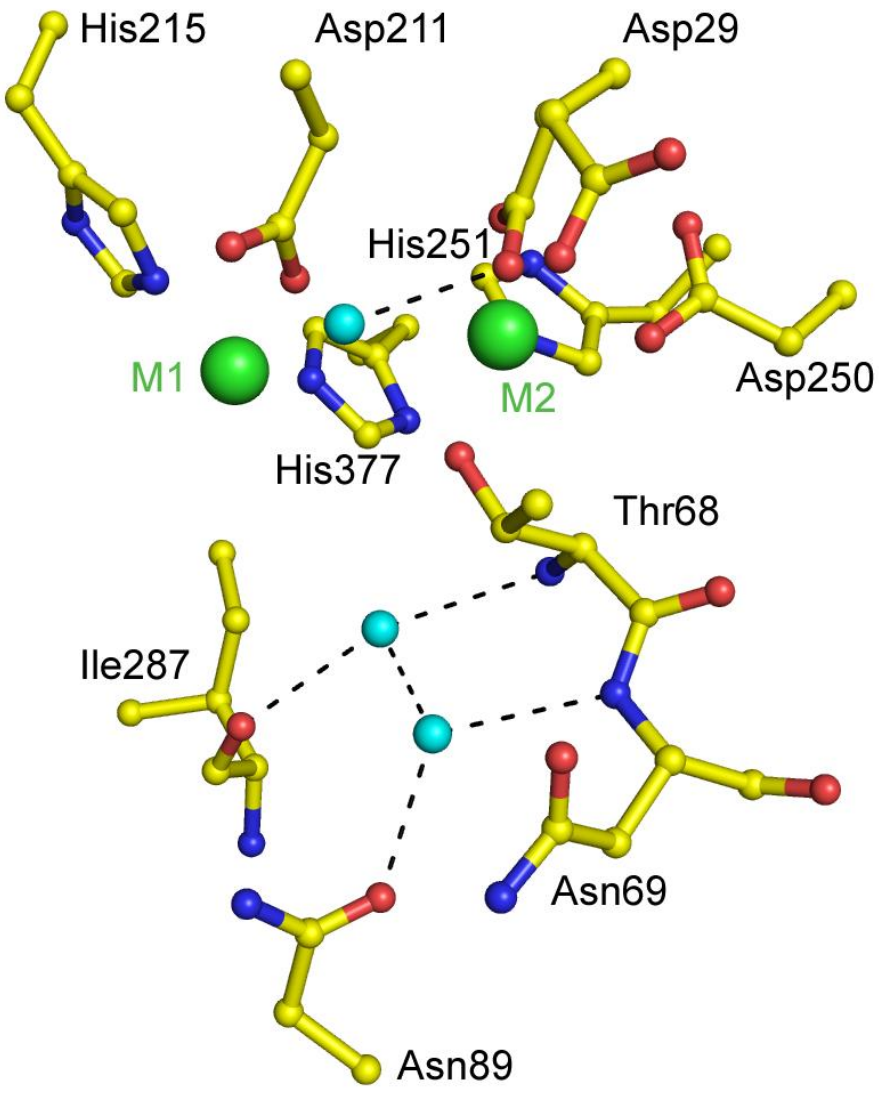

Figure 3.18 Active site view of the structure of PhnA. The water molecules are shown as cyan spheres. Metal ions are shown as green spheres. Enzyme carbon atoms are shown in yellow, oxygen in red and nitrogen in blue. Note that two different conformations of Asp29 are shown.

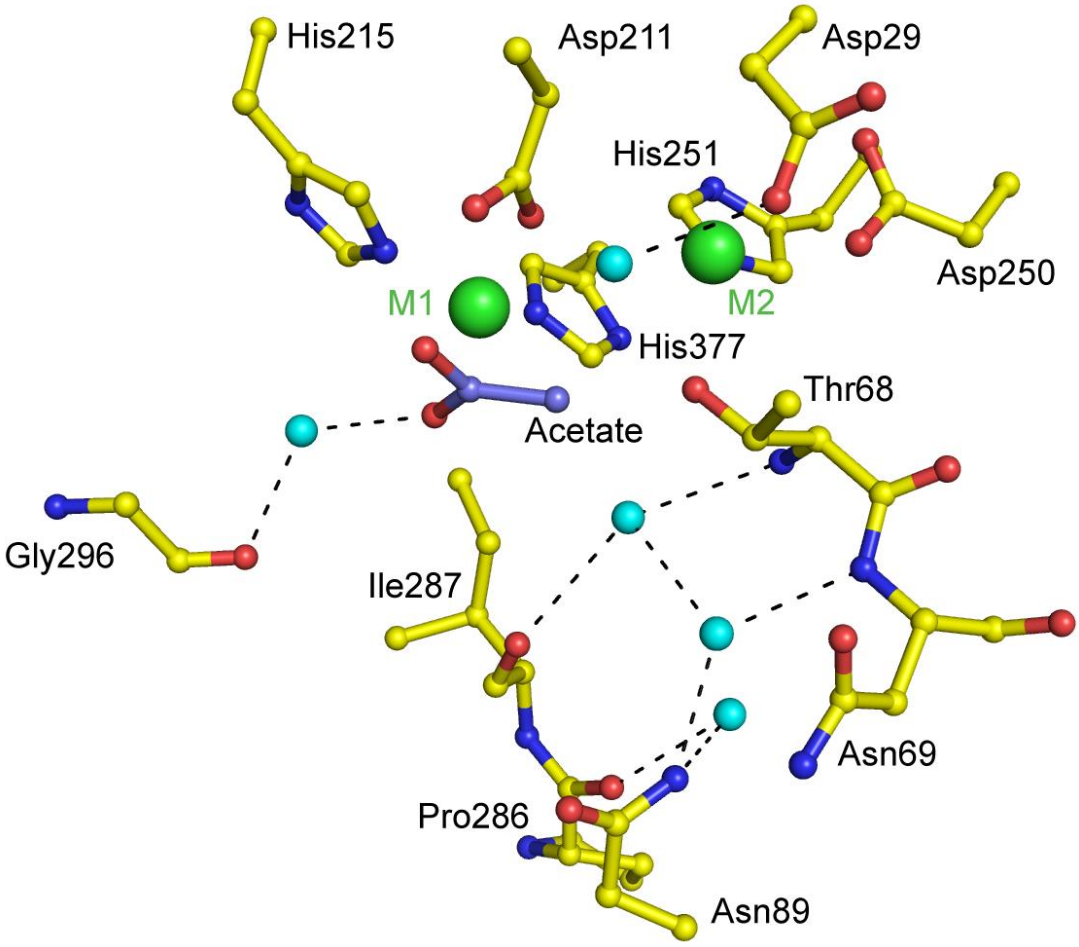

Figure 3.19 Active site view of the structure of PhnA with acetate bound. The water molecules are shown as cyan spheres. Metal ions are shown as green spheres. Enzyme carbon atoms are shown in yellow, oxygen in red and nitrogen in blue. Acetate carbon atoms are shown in blue.

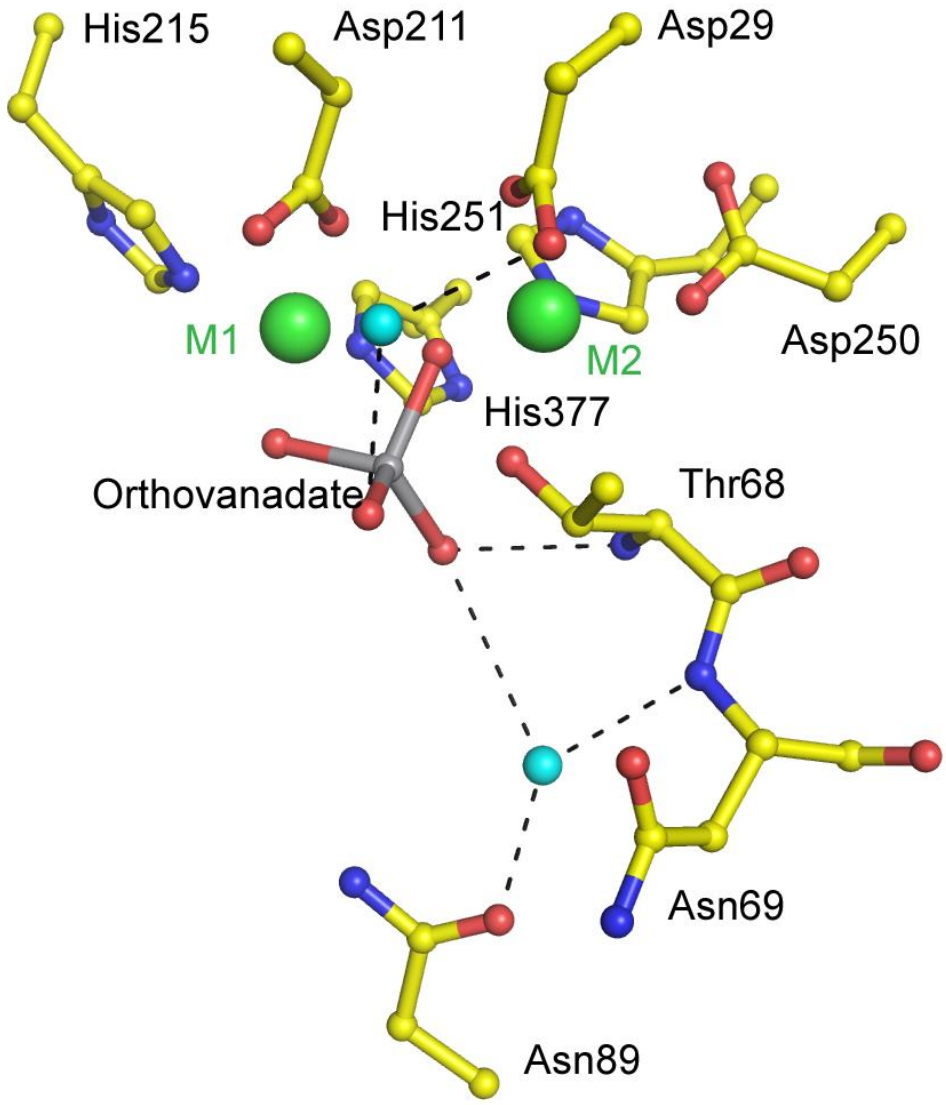

Figure 3.20 Active site view of the structure of PhnA with vanadate bound. The water molecules are shown as cyan spheres. Metal ions are shown as green spheres. Enzyme carbon atoms are shown in yellow, oxygen in red and nitrogen in blue. Vanadium atom is shown in grey. Note that Asn89 is too distant to make direct contacts with the equatorial oxygen atoms of the transition state mimic, as found in the crystal structures of AP and NPP. Rather Asn89 makes water mediated contacts with the transition state mimic.

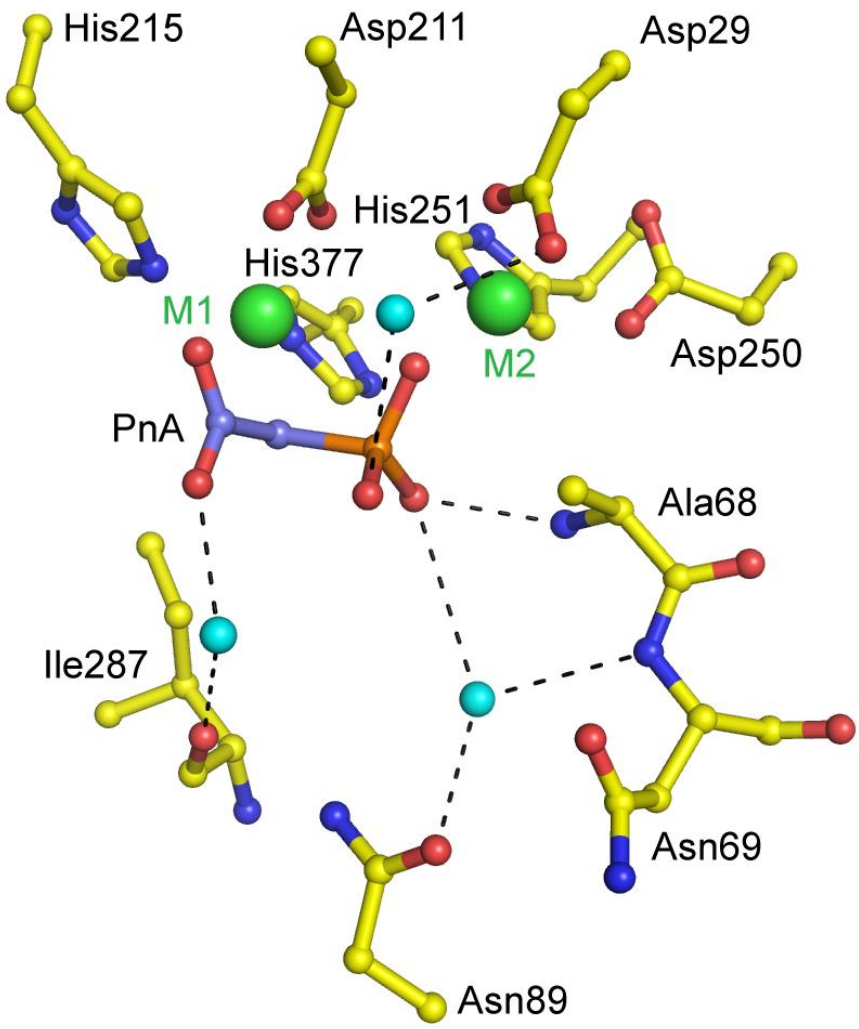

Figure 3.21 Active site view of the structure of PhnA-Thr68Ala with PnA bound. The water molecules are shown as cyan spheres. Metal ions are shown as green spheres. Enzyme carbon atoms are shown in yellow, oxygen in red and nitrogen in blue. Phosphorus atom is shown in orange.

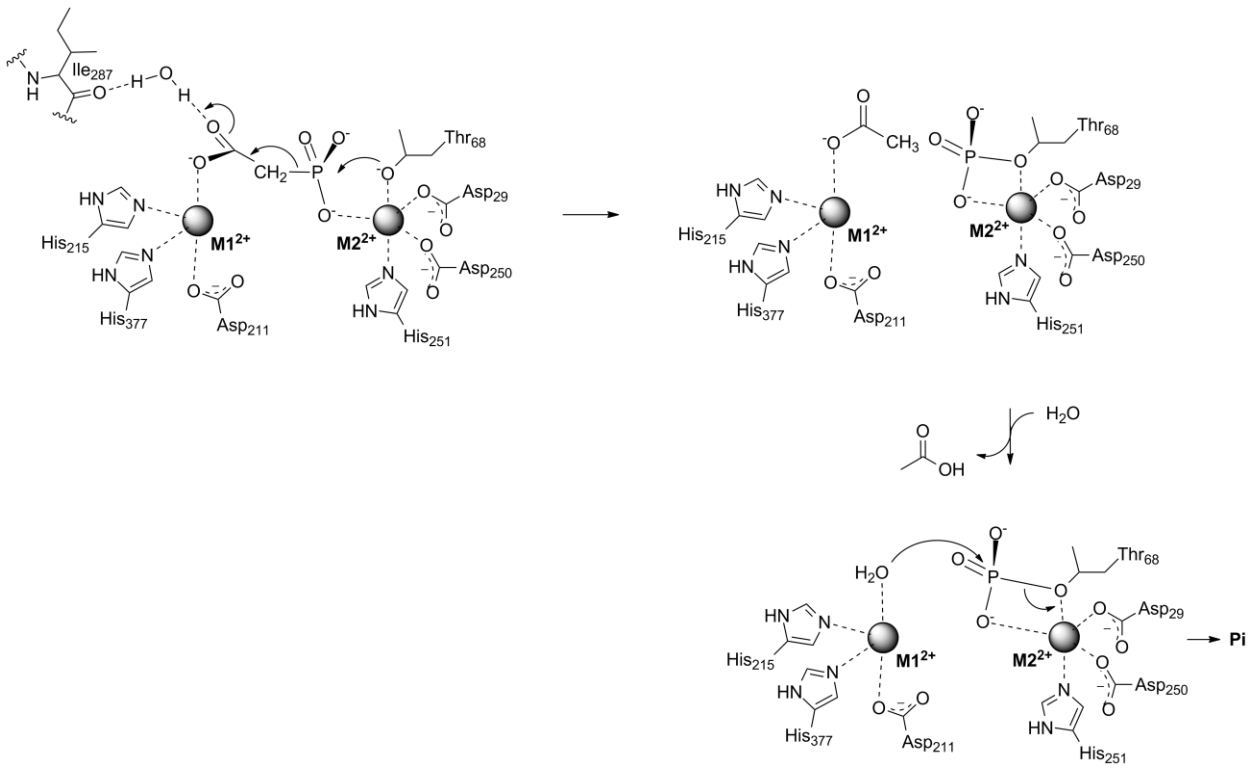

Figure 3.22 Putative mechanism for PnA hydrolysis by PhnA. Note the monodentate binding of the acetate moiety of PnA by the M1 metal ion.

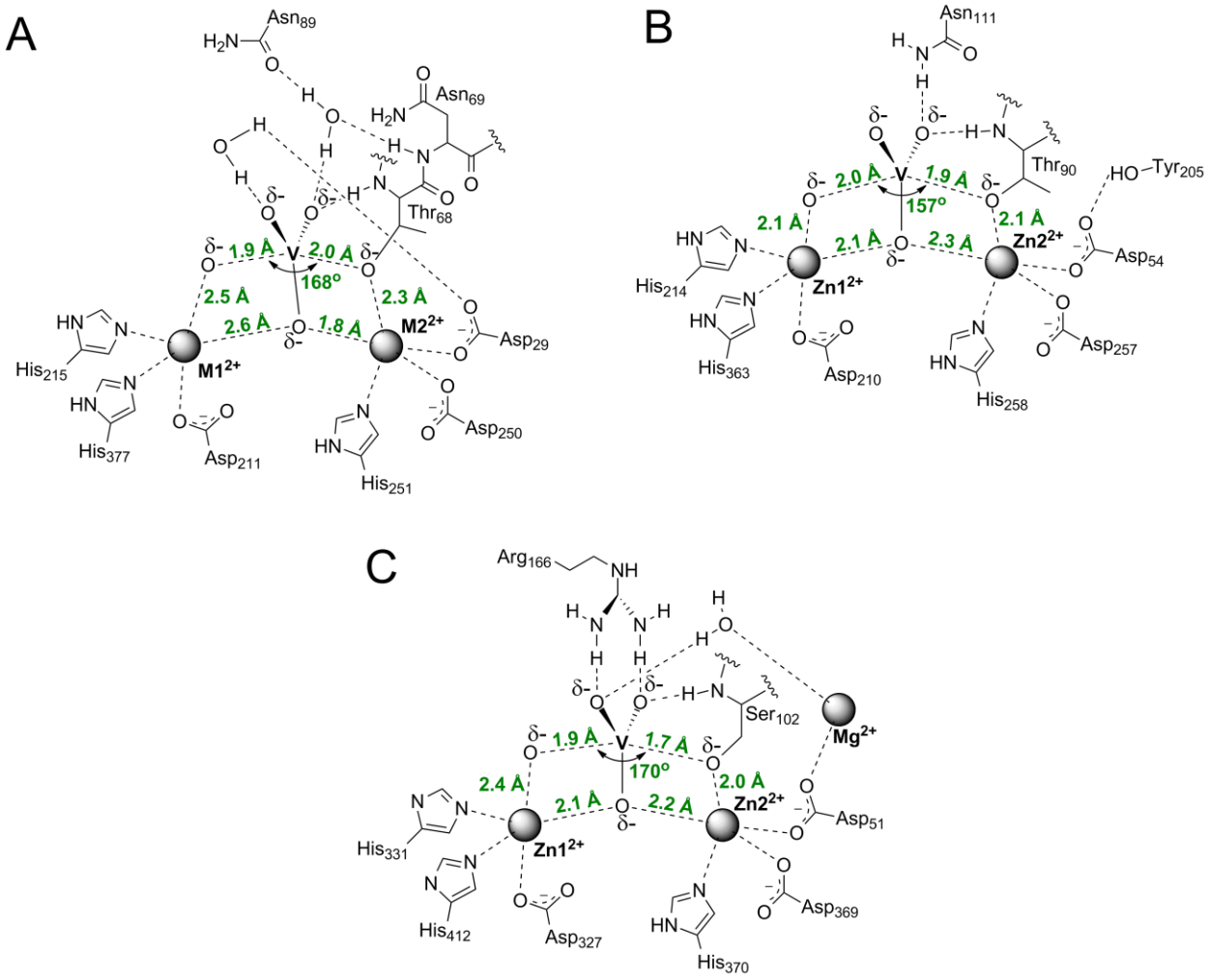

Figure 3.23 Scheme of (A) *S. meliloti* PhnA active site structure with vanadate bound used as transition state model as compared to (B) NPP (PDB Code: 2GSO) and (C) AP (PDB Code: 1B8J). Note the asymmetric nature of the equatorial oxygen atom co-ordination between the active site metal ions in the PhnA active site, as opposed to the symmetric nature of the same vanadate equatorial oxygen atom binding in the NPP and AP active site. Also note that the AP active site possess an additional Mg^{+2} ion, which is absent from PhnA and NPP active sites.

References

Agarwal, V., Borisova, S.A., Metcalf, W.W., van der Donk, W.A., and Nair, S.K. (2011). Structural and mechanistic insights into C-P bond hydrolysis by phosphonoacetate hydrolase. Chem Biol *18*, 1230-1240.

Bains, J., Leon, R., Temke, K.G., and Boulanger, M.J. (2011). Elucidating the reaction mechanism of the benzoate oxidation pathway encoded aldehyde dehydrogenase from Burkholderia xenovorans LB400. Protein Sci *20*, 1048-1059.

Borisova, S.A., Christman, H.D., Metcalf, M.E., Zulkepli, N.A., Zhang, J.K., van der Donk, W.A., and Metcalf, W.W. (2011a). Genetic and biochemical characterization of a pathway for the degradation of 2-aminoethylphosphonate in Sinorhizobium meliloti 1021. J Biol Chem *286*, 22283-22290.

Borisova, S.A., Circello, B.T., Zhang, J.K., van der Donk, W.A., and Metcalf, W.W. (2010). Biosynthesis of rhizocticins, antifungal phosphonate oligopeptides produced by Bacillus subtilis ATCC6633. Chem Biol *17*, 28-37.

Chen, C.C., Han, Y., Niu, W., Kulakova, A.N., Howard, A., Quinn, J.P., Dunaway-Mariano, D., and Herzberg, O. (2006). Structure and kinetics of phosphonopyruvate hydrolase from Variovorax sp. Pal2: new insight into the divergence of catalysis within the PEP mutase/isocitrate lyase superfamily. Biochemistry *45*, 11491-11504.

Clark, L.L., Ingall, E.D., and Benner, R. (1999). Marine organic phosphorus cycling: Novel insights from nuclear magnetic resonance. American Journal of Science *299*, 724-737.

Cobessi, D., Tete-Favier, F., Marchal, S., Azza, S., Branlant, G., and Aubry, A. (1999). Apo and holo crystal structures of an NADP-dependent aldehyde dehydrogenase from Streptococcus mutans. J Mol Biol *290*, 161-173.

Cobessi, D., Tete-Favier, F., Marchal, S., Branlant, G., and Aubry, A. (2000). Structural and biochemical investigations of the catalytic mechanism of an NADP-dependent aldehyde dehydrogenase from Streptococcus mutans. J Mol Biol *300*, 141-152.

Coleman, J.E. (1992). Structure and mechanism of alkaline phosphatase. Annu Rev Biophys Biomol Struct *21*, 441-483.

Cooley, N.A., Kulakova, A.N., Villarreal-Chiu, J.F., Gilbert, J.A., McGrath, J.W., and Quinn, J.P. (2011). Phosphonoacetate biosynthesis: in vitro detection of a novel NADP(+)-dependent phosphonoacetaldehyde-oxidizing activity in cell-extracts of the marine Roseovarius nubinhibens ISM. Mikrobiologiia *80*, 329-334.

D'Ambrosio, K., Pailot, A., Talfournier, F., Didierjean, C., Benedetti, E., Aubry, A., Branlant, G., and Corbier, C. (2006). The first crystal structure of a thioacylenzyme intermediate in the ALDH family: new coenzyme conformation and relevance to catalysis. Biochemistry *45*, 2978-2986.

Diaz-Sanchez, A.G., Gonzalez-Segura, L., Rudino-Pinera, E., Lira-Rocha, A., Torres-Larios, A., and Munoz-Clares, R.A. (2011a). Novel NADPH-cysteine covalent adduct found in the active site of an aldehyde dehydrogenase. Biochem J.

Diaz-Sanchez, A.G., Gonzalez-Segura, L., Rudino-Pinera, E., Lira-Rocha, A., Torres-Larios, A., and Munoz-Clares, R.A. (2011b). Novel NADPH-cysteine covalent adduct found in the active site of an aldehyde dehydrogenase. Biochem J *439*, 443-452.

Forlani, G., Klimek-Ochab, M., Jaworski, J., Lejczak, B., and Picco, A.M. (2006). Phosphonoacetic acid utilization by fungal isolates: occurrence and properties of a phosphonoacetate hydrolase in some penicillia. Mycol Res *110*, 1455-1463.

Galperin, M.Y., Bairoch, A., and Koonin, E.V. (1998). A superfamily of metalloenzymes unifies phosphopentomutase and cofactor-independent phosphoglycerate mutase with alkaline phosphatases and sulfatases. Protein Sci *7*, 1829-1835.

Ghosh, S.S., Bock, S.C., Rokita, S.E., and Kaiser, E.T. (1986). Modification of the active site of alkaline phosphatase by site-directed mutagenesis. Science *231*, 145-148.

Holtz, K.M., Stec, B., and Kantrowitz, E.R. (1999). A model of the transition state in the alkaline phosphatase reaction. J Biol Chem *274*, 8351-8354.

Holtz, K.M., Stec, B., Myers, J.K., Antonelli, S.M., Widlanski, T.S., and Kantrowitz, E.R. (2000). Alternate modes of binding in two crystal structures of alkaline phosphatase-inhibitor complexes. Protein Sci *9*, 907-915.

Kallarakal, A.T., Mitra, B., Kozarich, J.W., Gerlt, J.A., Clifton, J.G., Petsko, G.A., and Kenyon, G.L. (1995). Mechanism of the reaction catalyzed by mandelate racemase: structure and mechanistic properties of the K166R mutant. Biochemistry *34*, 2788-2797.

Kamat, S.S., Williams, H.J., and Raushel, F.M. (2011). Intermediates in the transformation of phosphonates to phosphate by bacteria. Nature.

Kim, A., Benning, M.M., Oklee, S., Quinn, J., Martin, B.M., Holden, H.M., and Dunaway-Mariano, D. (2011). Divergence of chemical function in the alkaline phosphatase superfamily: structure and mechanism of the p-C bond cleaving enzyme phosphonoacetate hydrolase. Biochemistry *50*, 3481-3494.

Kim, E.E., and Wyckoff, H.W. (1991). Reaction mechanism of alkaline phosphatase based on crystal structures. Two-metal ion catalysis. J Mol Biol *218*, 449-464.

Kim, Y.G., Lee, S., Kwon, O.S., Park, S.Y., Lee, S.J., Park, B.J., and Kim, K.J. (2009). Redox-switch modulation of human SSADH by dynamic catalytic loop. EMBO J *28*, 959-968.

Kleywegt, G.J., and Brunger, A.T. (1996). Checking your imagination: applications of the free R value. Structure *4*, 897-904.

Kulakova, A.N., Kulakov, L.A., and Quinn, J.P. (1997). Cloning of the phosphonoacetate hydrolase gene from Pseudomonas fluorescens 23F encoding a new type of carbon-phosphorus bond cleaving enzyme and its expression in Escherichia coli and Pseudomonas putida. Gene *195*, 49-53.

Kulakova, A.N., Wisdom, G.B., Kulakov, L.A., and Quinn, J.P. (2003). The purification and characterization of phosphonopyruvate hydrolase, a novel carbon-phosphorus bond cleavage enzyme from Variovorax sp Pal2. J Biol Chem *278*, 23426-23431.

Langendorf, C.G., Key, T.L., Fenalti, G., Kan, W.T., Buckle, A.M., Caradoc-Davies, T., Tuck, K.L., Law, R.H., and Whisstock, J.C. (2010). The X-ray crystal structure of Escherichia coli succinic semialdehyde dehydrogenase; structural insights into NADP+/enzyme interactions. PLoS One *5*, e9280.

Laskowski, R.A., Rullmannn, J.A., MacArthur, M.W., Kaptein, R., and Thornton, J.M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR *8*, 477-486.

Marchal, S., and Branlant, G. (1999). Evidence for the chemical activation of essential cys-302 upon cofactor binding to nonphosphorylating glyceraldehyde 3-phosphate dehydrogenase from Streptococcus mutans. Biochemistry *38*, 12950-12958.

Marchal, S., Rahuel-Clermont, S., and Branlant, G. (2000). Role of glutamate-268 in the catalytic mechanism of nonphosphorylating glyceraldehyde-3-phosphate dehydrogenase from Streptococcus mutans. Biochemistry *39*, 3327-3335.

McGrath, J.W., Wisdom, G.B., McMullan, G., Larkin, M.J., and Quinn, J.P. (1995). The purification and properties of phosphonoacetate hydrolase, a novel carbon-phosphorus bond-cleavage enzyme from Pseudomonas fluorescens 23F. Eur J Biochem *234*, 225-230.

McMullan, G., Harrington, F., and Quinn, J.P. (1992). Metabolism of phosphonoacetate as the sole carbon and phosphorus source by an environmental bacterial isolate. Appl Environ Microbiol *58*, 1364-1366.

McRee, D.E. (1999). XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. J Struct Biol *125*, 156-165.

Metcalf, W.W., and van der Donk, W.A. (2009). Biosynthesis of phosphonic and phosphinic acid natural products. Annu Rev Biochem *78*, 65-94.

Morais, M.C., Zhang, W., Baker, A.S., Zhang, G., Dunaway-Mariano, D., and Allen, K.N. (2000). The crystal structure of bacillus cereus phosphonoacetaldehyde hydrolase: insight into catalysis of phosphorus bond cleavage and catalytic diversification within the HAD enzyme superfamily. Biochemistry *39*, 10385-10396.

Murshudov, G.N., Vagin, A.A., and Dodson, E.J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr *53*, 240-255.

Murshudov, G.N., Vagin, A.A., Lebedev, A., Wilson, K.S., and Dodson, E.J. (1999). Efficient anisotropic refinement of macromolecular structures using FFT. Acta Crystallogr D Biol Crystallogr *55*, 247-255.

Nikolic-Hughes, I., O'Brien P, J., and Herschlag, D. (2005). Alkaline phosphatase catalysis is ultrasensitive to charge sequestered between the active site zinc ions. J Am Chem Soc *127*, 9314-9315.

Otwinowski, Z., Borek, D., Majewski, W., and Minor, W. (2003). Multiparametric scaling of diffraction intensities. Acta Crystallogr A *59*, 228-234.

Outten, C.E., and O'Halloran, T.V. (2001). Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. Science *292*, 2488-2492.

Panas, P., Ternan, N.G., Dooley, J.S., and McMullan, G. (2006). Detection of phosphonoacetate degradation and phnA genes in soil bacteria from distinct geographical origins suggest its possible biogenic origin. Environ Microbiol *8*, 939-945.

Perrakis, A., Sixma, T.K., Wilson, K.S., and Lamzin, V.S. (1997). wARP: improvement and extension of crystallographic phases by weighted averaging of multiple-refined dummy atomic models. Acta Crystallogr D Biol Crystallogr *53*, 448-455.

Richard, J.P., and Amyes, T.L. (2001). Proton transfer at carbon. Curr Opin Chem Biol *5*, 626-633.

Schonbrunn, E., Eschenburg, S., Shuttleworth, W.A., Schloss, J.V., Amrhein, N., Evans, J.N., and Kabsch, W. (2001). Interaction of the herbicide glyphosate with its target enzyme 5-enolpyruvylshikimate 3-phosphate synthase in atomic detail. Proc Natl Acad Sci U S A *98*, 1376-1380.

Schwartz, J.H., and Lipmann, F. (1961). Phosphate incorporation into alkaline phosphatase of E. coli. Proc Natl Acad Sci U S A *47*, 1996-2005.

Seidel, H.M., Freeman, S., Seto, H., and Knowles, J.R. (1988). Phosphonate biosynthesis: isolation of the enzyme responsible for the formation of a carbon-phosphorus bond. Nature *335*, 457-458.

Shames, S.L., Wackett, L.P., Labarge, M.S., Kuczkowski, R.L., and Walsh, C.T. (1987). Fragmentative and Stereochemical Isomerization Probes for Homolytic Carbon to Phosphorus

Bond Scission Catalyzed by Bacterial Carbon Phosphorus Lyase. Bioorganic Chemistry *15*, 366-373.

Stec, B., Holtz, K.M., and Kantrowitz, E.R. (2000). A revised mechanism for the alkaline phosphatase reaction involving three metal ions. J Mol Biol *299*, 1303-1311.

Steinrucken, H.C., and Amrhein, N. (1984). 5-Enolpyruvylshikimate-3-phosphate synthase of Klebsiella pneumoniae 2. Inhibition by glyphosate [N-(phosphonomethyl)glycine]. Eur J Biochem *143*, 351-357.

Ternan, N.G., Hamilton, J.T., and Quinn, J.P. (2000). Initial in vitro characterisation of phosphonopyruvate hydrolase, a novel phosphate starvation-independent, carbon-phosphorus bond cleavage enzyme in Burkholderia cepacia Pal6. Arch Microbiol *173*, 35-41.

Ternan, N.G., McGrath, J.W., and Quinn, J.P. (1998). Phosphoenolpyruvate phosphomutase activity in an L-phosphonoalanine-mineralizing strain of burkholderia cepacia. Appl Environ Microbiol *64*, 2291-2294.

Thomas, S., Burdett, H., Temperton, B., Wick, R., Snelling, D., McGrath, J.W., Quinn, J.P., Munn, C., and Gilbert, J.A. (2010). Evidence for phosphonate usage in the coral holobiont. ISME J *4*, 459-461.

Tylichova, M., Kopecny, D., Morera, S., Briozzo, P., Lenobel, R., Snegaroff, J., and Sebela, M. (2010). Structural and functional characterization of plant aminoaldehyde dehydrogenase from Pisum sativum with a broad specificity for natural and synthetic aminoaldehydes. J Mol Biol *396*, 870-882.

Van Duyne, G.D., Standaert, R.F., Karplus, P.A., Schreiber, S.L., and Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. J Mol Biol *229*, 105-124.

Villarreal-Chiu, J.F., Quinn, J.P., and McGrath, J.W. (2012). The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. Front Microbiol *3*, 19.

Wackett, L.P., Shames, S.L., Venditti, C.P., and Walsh, C.T. (1987). Bacterial carbon-phosphorus lyase: products, rates, and regulation of phosphonic and phosphinic acid metabolism. J Bacteriol *169*, 710-717.

Wojciechowski, C.L., and Kantrowitz, E.R. (2002). Altering of the metal specificity of Escherichia coli alkaline phosphatase. J Biol Chem *277*, 50476-50481.

Zalatan, J.G., Fenn, T.D., Brunger, A.T., and Herschlag, D. (2006). Structural and functional comparisons of nucleotide pyrophosphatase/phosphodiesterase and alkaline phosphatase: implications for mechanism and evolution. Biochemistry *45*, 9788-9803.

Zalatan, J.G., Fenn, T.D., and Herschlag, D. (2008). Comparative enzymology in the alkaline phosphatase superfamily to determine the catalytic role of an active-site metal ion. J Mol Biol *384*, 1174-1189.

Zalatan, J.G., and Herschlag, D. (2006). Alkaline phosphatase mono- and diesterase reactions: comparative transition state analysis. J Am Chem Soc *128*, 1293-1303.

Zingle, C., Kuntz, L., Tritsch, D., Grosdemange-Billiard, C., and Rohmer, M. (2010). Isoprenoid biosynthesis via the methylerythritol phosphate pathway: structural variations around phosphonate anchor and spacer of fosmidomycin, a potent inhibitor of deoxyxylulose phosphate reductoisomerase. J Org Chem *75*, 3203-3207.