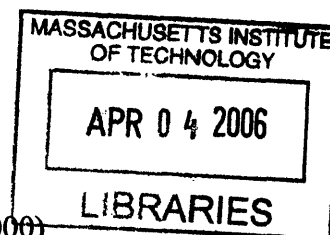


**Structural Analysis of Hydroxypropylphosphonic Acid Epoxidase:  
A Fosfomycin Biosynthetic Enzyme**

by

Luke J. Higgins

B.S., Biochemistry and Molecular Biology (2000)  
University of Maryland



**ARCHIVES**

Submitted to the Department of Chemistry  
in Partial Fulfillment of the Requirements for the Degree of  
Doctor of Philosophy in Biological Chemistry

at the

Massachusetts Institute of Technology

[February 2006]  
August 2005

© 2005 Massachusetts Institute of Technology  
All rights reserved

Signature of Author .....

Department of Chemistry  
August 8, 2005

Certified by .....

Dr. Catherine L. Drennan  
Associate Professor of Chemistry  
Thesis Supervisor

Accepted by .....

Dr. Robert W. Field  
Chairman, Departmental Committee on Graduate Studies

This doctoral thesis has been examined by Committee of the Department of Chemistry as follows:

.....  
Dr. Stephen J. Lippard  
Professor of Chemistry  
Committee Chair

.....  
Dr. Catherine L. Drennan  
Associate Professor of Chemistry  
Thesis Supervisor

.....  
Dr. JoAnne Stubbe  
Professor of Chemistry

## ACKNOWLEDGEMENTS

I would like to thank my thesis advisor Dr. Catherine Drennan. I thank Cathy for the flexibility that she gave me as a graduate student and for supporting my professional aspiration to become a physician-scientist.

Graduate school can be a grueling process for the graduate student but also to family and friends. My family deserves my sincere gratitude for all their support especially my Grandmother, my Dad, Beth, my Great Aunt Marguerite, Great Uncle Frank, Great Aunt Nancy, Aunt Rosie, Uncle Peter, Uncle Jim, Aunt Janie; cousins Frank, Mag (“Twig”), Tom, Lisa, Nancy, and Andy; the kids, Patrick (“the poet”), Kaitlyn, Nicole, Amanda, Tracy, Ryan, Molly, Colin (my Godson), Logan, Haily, Aden, Casey, Megan, Kyle, and Erin. Thank you all!

Coming to MIT allowed me to meet the love of my life, and I thank you Colleen for your undying love and support, especially during the last six months of thesis writing and seemingly endless revisions. Meeting you has made this all worth it! Thanks darlin’!

Finally, there are a lot of people at MIT that deserve my thanks especially Eric Schreiter, Michael Sintchak, Yvain Nicolet, and Mohammad Seyedsayamdost. You guys allowed me to bounce ideas off you, bug you about crystallography, and you all listened to me vent when MIT politics were getting me down. And to the rest of the Drennan Laboratory, past and present, you guys are the best! Thanks to you all!

## TABLE OF CONTENTS:

<b>ACKNOWLEDGEMENTS</b>	<b>3</b>
<b>ABSTRACT</b>	<b>9</b>
<b>Chapter I: Introduction</b>	<b>10</b>
I.A. Fosfomycin: a clinically useful antibiotic	10
I.B. Biochemical insight into the reactions catalyzed by HppE: a unique mononuclear iron enzyme	12
<i>I.B.1. Fosfomycin biosynthesis</i>	12
<i>I.B.2. Substrate binding</i>	13
<i>I.B.3. Mechanism of HppE catalysis</i>	14
<i>I.B.4. Regiospecificity and substrate specificity in HppE catalysis</i>	15
I.C. Comparing and contrasting mononuclear iron chemistry within the cupin superfamily	16
<i>I.C.1. <math>\alpha</math>-KG dependent cupin mononuclear iron enzymes:</i> <i>The TauD paradigm</i>	18
<i>I.C.2. <math>\alpha</math>-KG independent mononuclear iron enzymes:</i> <i>IPNS and ACC oxidase</i>	20
<i>I.C.3. Deacetoxycephalosporin synthase (DAOCS):</i> <i>“Booby trapped” mechanism</i>	21
<i>I.C.4. Clavamate synthase (CAS): A ligand rearrangement         mechanism</i>	23
I.D. Thesis Goals	24
I.E. Figures	25

I.F. References	45
<b>Chapter II: Structural Insight into the Biosynthesis of the Antibiotic</b>	<b>52</b>
<b>Fosfomycin by a Mononuclear Iron Enzyme</b>	
Summary	52
II.A. Introduction	53
II.B. Methods	55
<i>II.B.1. Site-directed mutagenesis, protein purification, assays</i>	55
<i>II.B.2. Crystallization and data collection</i>	56
<i>II.B.3. Structure Determinations</i>	58
II.C. Results	61
<i>II.C.1. Overall structure of HPP epoxidase</i>	61
<i>II.C.2. The HppE active site</i>	62
<i>II.C.3. Substrate binding and conformational changes</i>	64
II.D. Discussion	66
<i>II.D.1. Mechanism of HppE</i>	66
<i>II.D.2. Structural and mechanistic comparisons</i>	70
II.E. Conclusion	71
II.F. Acknowledgements	72
II.G. Figures and Tables	73
II.H. References	90
<b>Chapter III: Hydroxypropylphosphonic Acid Epoxidase: Structural</b>	<b>95</b>
<b>Insights into the Mechanism of Regioselectivity</b>	
Summary	95

III.A. Introduction	96
III.B. Methods	98
<i>III.B.1. Protein purification and crystallization</i>	98
<i>III.B.2. Structure determinations</i>	99
III.C. Results	100
<i>III.C.1. R-HPP-Co(II)-HppE: Three binding modes</i>	100
<i>III.C.2. Comparison of bidentate R- and S-HPP binding</i>	102
<i>III.C.3. Cantilever hairpin conformational change</i>	103
<i>III.C.4. Conserved water binding site</i>	104
III.D. Discussion	105
<i>III.D.1. Mechanism of regioselectivity</i>	105
<i>III.D.2. Substrate induced-fit conformational change</i>	107
<i>III.D.3. Mechanistic implications of conserved water binding site</i>	108
III.E. Conclusion	109
III.F. Figures and Tables	110
III.G. References	128
<b>Chapter IV: Hydroxypropylphosphonic Acid Epoxidase: An Enzyme</b>	<b>130</b>
<b>with Broad Substrate Specificity</b>	
Summary	130
IV.A. Introduction	131
IV.B. Methods	133
<i>IV.B.1. Protein purification and crystallization</i>	133
<i>IV.B.2. Structure determination and refinement</i>	134

IV.C. Results	135
IV.D. Discussion	137
IV.E. Conclusion	140
IV.F. Figures and Tables	141
IV.G. References	152
<b>Chapter V: Hydroxypropylphosphonic Acid Epoxidase: A Cupin</b>	<b>154</b>
<b>Enzyme Like No Other?</b>	
Summary	154
V.A. Introduction	155
V.B. Methods	157
V.C. Discussion	158
<i>V.C.1. HppE: A structural dicupin</i>	158
<i>V.C.2. Cupin Motifs</i>	160
<i>V.C.3. Conformational changes and active site protection:</i>	161
<i>TauD, IPNS, HppE</i>	
<i>V.C.4. The RXS motif and <math>\alpha</math>-KG dependence</i>	163
<i>V.C.5. The other HppE from Pseudomonus syringae</i>	164
V.D. Figures and Tables	165
V.E. References	180
<b>VI. Appendix: Ribonucleoside triphosphate reductase and thioredoxin</b>	<b>185</b>
VI.A. Cloning, expression and purification of thioredoxin from	185
<i>Lactobacillus leichmannii</i>	
<i>VI.A.1. Isolation of genomic DNA from Lactobacillus leichmannii</i>	185

<i>VI.A.2. Cloning thioredoxin gene from Lactobacillus leichmannii genomic DNA</i>	186
<i>VI.A.3. PCR of Tr in TOPO vector and ligation into a protein expression vector</i>	187
<i>VI.A.4. Ligation of the thioredoxin gene into pET 23a</i>	188
<i>VI.A.5. Purification of C31S-Tr from Lactobacillus leichmannii</i>	189
VI.B. Purification of C731S-RTPR	190
VI.C. RTPR-Tr coupling experiments	191
<i>VI.C.1. Overview</i>	191
<i>VI.C.2. Overall coupling strategy</i>	191
<i>VI.C.3. Coupling procedure</i>	192
VI.D. RTPR in complex with cob(II)alamin, 5'-dA, substrate and/or Effector	193
<i>VI.D.1. C731,736-RTPR protein purification</i>	193
<i>VI.D.2. C731,736-RTPR crystallization</i>	194
<i>VI.D.3. Preliminary data analysis</i>	196
VI. Figures and Tables	198
VI.F. References	214
<b>Curriculum Vitae</b>	<b>215</b>



# Structural Analysis of Hydroxypropylphosphonic Acid Epoxidase:

## A Fosfomicin Biosynthetic Enzyme

by

Luke J. Higgins

Submitted to the Department of Chemistry  
on August 8, 2005 in Partial Fulfillment of the  
Requirements for the Doctor of Philosophy in  
Chemistry

### ABSTRACT

An X-ray crystallographic study of the fosfomicin biosynthetic enzyme hydroxypropylphosphonic acid epoxidase (HppE) from *Streptomyces wedmorensis* is presented. Structural analysis of this cupin mononuclear iron enzyme in complex with its *in vivo* substrate, and comparison with apo- and holo-enzyme structures, provides insight into the mechanism of fosfomicin biosynthesis. *In vivo*, the enzyme catalyzes the biosynthesis of the epoxide fosfomicin from its substrate, *S*-2-hydroxypropylphosphonic acid (*S*-HPP). *In vitro*, HppE is able to catalyze the conversion of the substrate enantiomer, *R*-2-hydroxypropylphosphonic acid (*R*-HPP), to a single ketone product in a stereospecific manner. X-ray crystal structures of HppE in complex with the *R*-HPP substrate suggest a mechanism for the regiospecific reactions catalyzed by the enzyme. This regiospecificity of HppE catalysis is also observed when substrates *R*- and *S*-2-phenylethylphosphonic acid (HPEP) are used, where a methyl group has been substituted by a larger phenyl substituent. X-ray crystal structures in complex with *S*-HPEP suggest a conserved mechanism of regioselectivity and a mechanism of ligand rearrangement upon co-substrate (i.e. dioxygen) binding. Finally, HppE is compared to enzymes and proteins within the cupin superfamily, defining a new structural subfamily of cupin mononuclear iron enzymes.

Thesis Supervisor: Dr. Catherine L. Drennan  
Title: Associate Professor of Chemistry

## Chapter I: Introduction

### I.A. Fosfomycin: a clinically useful antibiotic

Fosfomycin, or (1*R*,2*S*)-1,2-epoxypropylphosphonic acid, is a broad spectrum antibiotic used clinically for the treatment of lower urinary tract infections<sup>1-4</sup> and has recently shown drug synergy against methicillin-resistant *Staphylococcus aureus* (MRSA) obtained from patients suffering from third-degree burns<sup>5</sup>. Furthermore, several vancomycin resistant, fosfomycin susceptible clones of MRSA were recently identified<sup>6</sup>. This broadening of MRSA resistance to existing antibiotic therapy underscores the importance of understanding the biosynthesis of fosfomycin.

From a clinical perspective, fosfomycin is an ideal drug in that it has high bioavailability and low drug toxicity. When administered orally, fosfomycin is not susceptible to degradation via the first-pass effect, and when administered intravenously, the drug is chemically intact after renal excretion. The relative lack of toxicity is due to the unique structure of fosfomycin, in that the drug is very specific for its *in vivo* drug target<sup>7</sup>.

Fosfomycin inhibits bacterial cell wall biosynthesis via the formation of a covalent adduct with UDP-N-acetylglucosamine enolpyruvyltransferase (MurA) (Figure I.1.)<sup>8</sup>. MurA catalyzes the first committed step in the biosynthesis of the peptidoglycan cell wall<sup>9</sup>. Pathogenic bacteria that are susceptible to fosfomycin inhibition express a MurA enzyme that utilizes an active site cysteine<sup>10</sup>. This cysteine facilitates the tautomerization of phosphoenolpyruvate (PEP) and the subsequent dephosphorylation of substrate (Figure I.2.) by playing a role as a general acid in the protonation of the C3 of PEP<sup>11</sup>. The active site cysteine is also thought to form a reversible covalent adduct with

the PEP tautomer but this role is not essential for catalysis<sup>10,11</sup>. Upon exposure to fosfomycin, however, the epoxide reacts with this active site cysteine forming an irreversible covalent adduct, precluding substrate binding and catalysis as evidenced by the fact that cysteine to aspartic acid mutation prevents the formation of a covalent adduct without abolishing activity<sup>10,11</sup>.

Pathogens have evolved several mechanisms of resistance to fosfomycin including: (1) active site modification of MurA, (2) fosfomycin-modifying enzymes, and (3) MurA overexpression<sup>9,12-18</sup>. Active site modification arises from an evolutionary pressure exerted by electrophilic agents like fosfomycin. The effect of this modification is to reduce the nucleophilicity of the catalytic acid while maintaining its ability to facilitate PEP tautomerization. The pathogen *Mycobacterium tuberculosis*, for example, has evolved a MurA that utilizes aspartic acid as the catalytic acid instead of cysteine<sup>11</sup>. A second strategy of bacterial fosfomycin resistance serves to preemptively disarm the fosfomycin warhead (i.e. the epoxide). Four different fosfomycin-modifying enzymes have been characterized: FosA<sup>15</sup>, FosB<sup>13</sup>, FosC<sup>16</sup>, and FosX<sup>19</sup>. These enzymes catalyze modification of fosfomycin by glutathione, cysteine, ATP, or via hydrolysis, respectively<sup>13,15,16,19</sup>. For FosA, crystallographic and biochemical analysis reveal a ligand triad coordinating an active site manganese ion<sup>20</sup>. Fosfomycin has been shown to bind the manganese metal in a bidentate fashion via the oxirane oxygen and the phosphonate substituent (Figure I.3.)<sup>20</sup>. A nearby binding site for glutathione serves to orient and stabilize the nucleophilic thiolate, facilitating glutathione conjugation<sup>20</sup>.

## **I.B. Biochemical insight into the reactions catalyzed by HppE: a unique mononuclear iron enzyme**

### *I.B.1. Fosfomycin biosynthesis catalyzed by HppE*

Fosfomycin is one of a few known natural products containing a carbon-phosphorus (C-P) bond<sup>21</sup>, other examples include bialaphos, plumbemycin, fosfazinomycin, fosmidomycin, and phosalacine. The fosfomycin biosynthetic gene cluster in *Streptomyces wedmorensis* has been cloned and reveals that the fosfomycin biosynthetic pathway is similar to other C-P bond containing natural products with respect to the first two steps in the pathway, the conversion of PEP to phosphopyruvate (PnPy) and the subsequent decarboxylation of PnPy to phosphonoacetaldehyde (PnAA) (Figure I.4.)<sup>21</sup>. After this, the pathways diverge, and PnAA is methylated by PnAA methylase in what is predicted by sequence and cofactor requirements to be a novel methylation<sup>22</sup> reaction. The product of this methylation, (*S*)-2-hydroxypropylphosphonic acid (*S*-HPP), is converted into the epoxide fosfomycin by (*S*)-2-hydroxypropylphosphonic acid epoxidase (HppE) via an oxidative cyclization reaction<sup>23</sup>.

HppE is thought to be a member of the cupin superfamily of proteins based on sequence analysis. A subset of cupin enzymes bind metals using a conserved triad of ligands, His<sub>2</sub>(Asp/Glu). HppE is a representative of this class of cupin proteins and requires iron for catalysis. Cupin mononuclear iron proteins like HppE catalyze a diverse set of reactions including those involved in DNA repair and antibiotic biosynthesis. Apo-HppE is purified as a homotetramer (monomer MW=21,200 Da) and is reconstituted *in vitro* with iron ammonium sulfate to afford an active enzyme<sup>23,24</sup>. *In vitro* activity requires dioxygen and an electron source (e.g. FMN and NADH)<sup>23</sup>. The turnover rate

using FMN and NADH for *S*-HPP is  $149 \pm 2$  nmol/mg·min<sup>23</sup>. The enzyme is inactive using NADH alone<sup>25</sup>. *In vivo* active site reduction is thought to require a reductase; this protein has yet to be identified<sup>23,24,26,27</sup>. However, an *in vitro* surrogate to this reductase is an enzyme known as E3, an NADH-dependent-[2Fe2S] containing flavoprotein reductase from the 3,6-dideoxyhexose biosynthetic pathway in *Yersinia pseudotuberculosis*. The turnover rate for fosfomycin synthesis using *S*-HPP as substrate and the E3 reductase is 20-fold higher than that when FMN and NADH are used as an electron source (turnover rate for *S*-HPP =  $2.2 \mu\text{mol/mg}\cdot\text{min}$ )<sup>23</sup>. These observations are consistent with two single electron transfers to the HppE active site during catalysis. Although the identity of the HppE reductase is unknown, a yeast two hybrid system revealed that only one protein, orfD, in the fosfomycin biosynthetic gene cluster in *Streptomyces wedmorensis* interacts specifically with HppE. Unfortunately, orfD has not yet been shown capable of reducing the HppE active site.

### *I.B.2. Substrate binding*

Like cupin mononuclear iron enzymes 1-aminocyclopropane-1-carboxylate oxidase (ACC oxidase)<sup>28</sup> and isopenicillin N synthase (IPNS)<sup>29</sup>, HppE is an  $\alpha$ -KG independent mononuclear iron enzyme<sup>23</sup>. *S*-HPP oxidation and an *in vivo* reductase provide the requisite reducing equivalents. Analogous to  $\alpha$ -KG which binds to the active site iron of  $\alpha$ -KG-dependent enzymes like deacetoxycephalosporin synthase (DAOCS)<sup>30</sup>, taurine  $\alpha$ -KG dioxygenase (TauD)<sup>31</sup>, and clavamate synthase (CAS)<sup>32</sup>, *S*-HPP has been shown to bind near or directly to the iron center by EPR<sup>23</sup>. EPR analysis of Fe(III)-HppE is consistent with two ligand environments and substrate binding to ferric HppE results in

a perturbation of the EPR spectra<sup>23</sup>. *S*-HPP binding influences both ferric HppE metal environments and is therefore consistent with at least two substrate binding modes. The EPR signal coalesces into a homogeneous spectrum, consistent with a single binding mode, upon addition of nitric oxide to ferrous HppE in the presence of *S*-HPP; nitric oxide is used as a surrogate of dioxygen<sup>23</sup>.

### *I.B.3. Mechanism of HppE catalysis*

Several mechanisms have been proposed for the biosynthesis of fosfomicin catalyzed by HppE<sup>23,27</sup>. All mechanistic proposals employ a Fe-oxo species as an oxidant which facilitates hydrogen atom abstraction and subsequent substrate radical formation (Figure I.5.). When Fe(II)-HppE is incubated with *S*-HPP under aerobic conditions for 1 hr a radical forms ( $g=2.0$ ). This radical produces a narrow linewidth ( $< 40$  G) that is consistent with an organic radical rather than a protein radical (Feng Yan Thesis, 2005)<sup>33</sup>. Currently, one can only speculate that the EPR data are the result of a substrate radical loosely coupled to the iron center. Experiments using substrate analogs and radical clocks also support a radical mechanism although results are not definitive<sup>25</sup>.

Further indirect evidence for a radical mechanism is an off-pathway (i.e. distinct from the fosfomicin biosynthetic pathway) monooxygenase reaction catalyzed by HppE<sup>27</sup> (Figure I.6.). When HppE is reconstituted with ferrous ammonium sulfate and subsequently treated with 10 equivalents of ascorbate (a reductant) under aerobic conditions, the enzyme catalyzes a self-hydroxylation of Y105 to produce DOPA-Y105<sup>27</sup>. The oxygen atom inserted into Y105 is derived from molecular oxygen<sup>27</sup>.

Analogous to TauD<sup>34</sup>, a ferryl-oxo species could be the reactive iron-oxo intermediate that facilitates *S*-HPP turnover by HppE. Using P450-like chemistry<sup>35</sup>, hydrogen atom abstraction at the C1 position of substrate would yield a transient substrate radical intermediate (Figure I.5.). Oxygen rebound could afford a diol intermediate and subsequently the epoxide fosfomycin. In order to probe this mechanistic possibility, [2-<sup>18</sup>O]-*S*-HPP and [<sup>18</sup>O]-dioxygen were separately used as substrates. With respect to the former, <sup>18</sup>O was found in the epoxide product, and in the latter case, the isotope label was not found in the epoxide product<sup>23</sup>. This experiment shows that the HppE mechanism is not the result of oxygen rebound, nor is HppE a monooxygenase. When on-pathway, hydrogen peroxide is not a product of HppE catalysis, nor can hydrogen peroxide be used as a substrate. This lack of a peroxide shunt implies that a ferrous peroxide intermediate is not a long lived species during catalysis. In short, the reaction catalyzed by HppE appears to be the result of an epoxidation mechanism without precedent in biology.

#### *I.B.4. Regiospecificity and substrate specificity in HppE catalysis*

HppE recognizes both enantiomers of substrate, catalyzing two distinct reactions with regiospecificity. Specifically, *S*-HPP is converted to the epoxide fosfomycin while *R*-HPP, the C2 enantiomer, is exclusively converted to the ketone 2-oxo-propylphosphonate (turnover rate using an FMN and NADH reducing system =  $162 \pm 7$  nmol/min·mg)<sup>23</sup>. Experiments with C2-substituted difluoroanalogs (discussed in chapter 3) are consistent with a regiospecific mechanism where hydrogen abstraction occurs at the C1 position when the *S*-stereoisomer is used whereas hydrogen atom abstraction at

the C2 position occurs with respect to *R*-HPP catalysis (Figure I.7.)<sup>26</sup>. Previous feeding experiments using fosfomycin producing microbe *S. fradiae* are consistent with regiospecific hydrogen atom abstraction of the pro-*R* hydrogen<sup>36</sup>; in that, C2 deuterium labeling of the pro-*R* hydrogen results in isotope washout (i.e. not observed in fosfomycin) while labeling the pro-*S* hydrogen results in isotope retention (i.e. observed in fosfomycin).

HppE also catalyzes the conversion of (*S*)-2-hydroxy-2-phenylethylphosphonic acid (*S*-HPEP) to (1*R*,2*S*)-2-phenyl-1,2-epoxyethylphosphonic acid, and (*R*)-2-hydroxy-2-phenylethylphosphonic acid (*R*-HPEP) to 2-oxo-2-phenylethylphosphonic acid<sup>25</sup> (Figure I.8. and chapter 4). Paradoxically, complete regiospecificity is maintained despite such broad substrate specificity (i.e. phenyl to methyl substitution)<sup>25</sup>. Experiments analogous to those described for *S*-HPP and *R*-HPP substrates have not been done on HPEP substrates, so little is known about the mechanism of regiospecificity and substrate specificity when these analogs are used.

### **I.C. Comparing and contrasting mononuclear iron chemistry within the cupin superfamily**

The cupin superfamily of proteins has been called the “most functionally diverse superfamily” in biology<sup>37,38</sup> (see chapter 5). An important subfamily are the mononuclear iron enzymes which catalyze a diverse array of reactions and perform a variety of functions within the biological milieu including superoxide scavenging, DNA repair, and antibiotic biosynthesis<sup>39</sup>. In general, these enzymes house an active site monoanionic triad of ligands, His<sub>2</sub>(Asp/Glu), that coordinate the iron<sup>40</sup>. Despite the broad spectrum of



reactions catalyzed there are certain tasks that each enzyme must accomplish: (1) fine-tuning of the active site to bind substrate, co-substrate and/or cofactor, (2) activating the iron center for dioxygen binding and subsequent oxidation of Fe<sup>II</sup> to Fe<sup>III</sup>, (3) shielding of high energy iron-oxo intermediates from solvent, and (4) reduction of dioxygen to form a substrate hydroxyl group or water<sup>41,42</sup>.

The strategy employed for iron activation by these enzymes is quite similar. The “resting state” of the enzyme to which dioxygen binds weakly is a six coordinate ferrous center with the monoanionic triad contributing three ligands and waters occupying the remaining exchangeable sites. To activate iron, substrate or co-substrate binds directly to iron via electron donating substituents thereby lowering the Fe<sup>III/II</sup> potential and favoring iron oxidation by dioxygen. This modulation of the redox potential allows for formation of a ferric superoxide intermediate (Fe<sup>III</sup>-O<sub>2</sub><sup>•-</sup>)<sup>43</sup>.

Mononuclear iron enzymes are either alpha-ketoglutarate ( $\alpha$ -KG) dependent (e.g. TauD<sup>31</sup> and DAOCS<sup>44</sup>) or independent (e.g. IPNS<sup>45</sup>, ACC oxidase<sup>28</sup>, and HppE<sup>24</sup>). Those that are dependent on  $\alpha$ -KG bind the cofactor in a bidentate fashion to the active site iron via the 2-oxo and C1 carboxylate oxygen atoms<sup>39</sup>; this five-coordinate, square pyramidal interaction leaves an open coordination site for subsequent dioxygen binding. Furthermore, coordination of the negatively charged  $\alpha$ -KG molecule in effect lowers the reduction potential of the ferrous center, facilitating iron oxidation by dioxygen.

Like HppE, IPNS<sup>29,46</sup> and ACC oxidase are  $\alpha$ -KG independent<sup>47</sup>, and both enzymes bind substrate via direct coordination to the iron center and both substrates are negatively charged<sup>29,48,49</sup>. Therefore, similar to  $\alpha$ -KG dependent enzymes, substrate binding to the iron center of  $\alpha$ -KG independent enzymes facilitates dioxygen binding via

electron donation, lowering the reduction potential of the ferrous center. In addition, substrate binding can result in the displacement of water, opening a coordination on Fe for dioxygen binding.

In order to provide a context for understanding fosfomycin biosynthesis catalyzed by HppE, the specific strategies employed with respect to iron activation and subsequent stabilization of high energy iron-oxo intermediates by the mononuclear iron enzymes TauD, DAOCS, IPNS, and ACC oxidase will be discussed below. Finally CAS is presented as an example of an enzyme, like HppE, that is able to recognize multiple substrates and catalyzes three distinct reactions.

#### *I.C.1. $\alpha$ -KG dependent cupin mononuclear iron enzymes: The TauD paradigm*

The HppE substrate, *S*-HPP, has been proposed to serve similar catalytic functions as the cofactor  $\alpha$ -KG in other mononuclear iron enzymes including TauD. Also, substrate binding to HppE has been proposed to facilitate iron oxidation and the subsequent formation of a transient ferryl species.  $\alpha$ -KG dependent monooxygenases and dioxygenases, like TauD, are thought to couple  $\alpha$ -KG oxidation to substrate oxidation via a conserved mechanism<sup>39</sup>. Although several  $\alpha$ -KG dependent enzymes have been studied in detail, TauD remains the only enzyme where a ferryl intermediate has been directly observed<sup>50</sup>, and therefore it serves as an ideal paradigm to describe the consensus mechanism for this subfamily of enzymes. Structural analysis of TauD, like HppE (chapter 2), illustrate how substrate binding triggers a conformational change that may serve to protect the ferryl species during catalysis.

TauD catalyzes the oxidation of taurine to sulfite and aminoacetaldehyde under conditions depleted of inorganic sulfur (Figure I.9.)<sup>31,51</sup>. Substrate oxidation, in this case hydroxylation, is coupled to  $\alpha$ -KG oxidative decarboxylation. The octahedral ferrous center binds  $\alpha$ -KG displacing two water molecules maintaining the six coordinate ligand geometry<sup>52,53</sup>. Substrate binding provides the impetus for converting the six-coordinate metal center to a five-coordinate center through the loss of the remaining water ligand<sup>53</sup>. This open coordination site is coincident with the dioxygen binding site. Interestingly, the  $K_d$  for dioxygen binding is significantly reduced for the substrate bound quaternary complex compared to the ternary complex (i.e.  $\alpha$ -KG alone). Furthermore, oxidative decarboxylation of  $\alpha$ -KG in the absence of substrate is much slower than in its presence<sup>31,34,50</sup>. It is therefore clear that quaternary complex formation facilitates dioxygen binding and catalytic initiation.

The next steps in the TauD catalytic cycle allow for oxygen insertion into  $\alpha$ -KG and subsequently into substrate. An Fe(III) bound superoxide species can attack the C2 carbonyl carbon of  $\alpha$ -KG forming a putative, cyclic intermediate (Figure I.10.). Subsequent decomposition of this transient species could result in the oxidative decarboxylation of  $\alpha$ -KG concomitant with formation of a ferryl-oxo (Fe(IV)=O) intermediate. TauD is the only enzyme where data has been presented supporting the identity of this intermediate. Mössbauer experiments on TauD are consistent with either a high spin Fe(IV)-oxo or Fe(III)-oxo species<sup>34</sup>. To differentiate between these, a single electron reduction of the high valent intermediate formed, following rapid freeze quench was performed, and the resulting Mössbauer parameters were consistent with an Fe(III)-oxo species<sup>34,50</sup>. Thus the initial species was indeed a ferryl-oxo intermediate. The

resulting substrate radical is oxidized via oxygen insertion, or rebound, using the second oxygen atom derived from molecular oxygen (Figure I.10.).

### *I.C.2. $\alpha$ -KG independent mononuclear iron enzymes: IPNS and ACC oxidase*

HppE catalyzes fosfomycin biosynthesis in an  $\alpha$ -KG independent fashion. Both IPNS and ACC oxidase are also  $\alpha$ -KG independent enzymes. In contrast to HppE, IPNS catalyzes a four electron oxidation of substrate whereas ACC oxidase, like HppE, catalyzes a two electron oxidation of substrate. Structural analysis of IPNS, like HppE (chapter 2), illustrate how substrate binding triggers a conformational change that serves to protect the ferryl species during catalysis.

The final step in penicillin biosynthesis requires the cupin mononuclear iron enzyme isopenicillin N synthase (IPNS)<sup>54,55</sup>. IPNS converts adipoylcysteinyvaline (ACV) to penicillin N (Figure I.11.) via two consecutive oxidative cyclization reactions. IPNS is structurally similar to  $\alpha$ -KG dependent mononuclear iron enzymes and the reason IPNS does not require  $\alpha$ -KG can be easily rationalized<sup>29</sup>. Most importantly, the conversion of ACV to penicillin N is a four electron oxidation, sufficient to completely reduce molecular oxygen to water. All  $\alpha$ -KG dependent mononuclear iron enzymes couple two electron substrate oxidation to the two electron oxidative decarboxylation of  $\alpha$ -KG<sup>56</sup>. As described above,  $\alpha$ -KG not only serves as a reducing agent but also to activate ferrous center for dioxygen binding<sup>42</sup>. IPNS is not designed to bind  $\alpha$ -KG and therefore has evolved a different mechanism of iron activation. As shown in Figure I.12., the cysteinyl sulfur atom of ACV directly coordinates to the metal center, resulting in the loss of a single water molecule and the displacement of Gln330<sup>29</sup>. Like TauD and other  $\alpha$ -KG

dependent enzymes, substrate binding results in a coordination change of six to five about the iron center<sup>29</sup>. With respect to IPNS, the direct coordination of the substrate thiolate is proposed to activate the iron center through electron donation facilitating dioxygen binding. ACV thus serves to two functions: provide reducing equivalents and facilitate dioxygen binding, eliminating the need for  $\alpha$ -KG.

ACC oxidase is another cupin mononuclear iron enzyme that is  $\alpha$ -KG independent<sup>28</sup>. This enzyme catalyzes an interesting decomposition reaction, evolving ethylene, carbon dioxide, and cyanide (Figure I.13.). ENDOR (electron nuclear double resonance) analysis of substrate binding to ACC oxidase show that 1-aminocyclopropane-1-carboxylate binds in a bidentate fashion via the amino and carboxylate substituents *en route* to ethylene biosynthesis. This bidentate binding, like  $\alpha$ -KG, serves to activate the metal center for dioxygen binding through electron donation<sup>49</sup>. In contrast to IPNS, substrate oxidation is only a two electron process thus requiring two additional electrons to afford complete reduction of dioxygen to water. ACC oxidase uses ascorbate as a cofactor to facilitate electron transfer to the metal center<sup>47,57</sup>. Currently no structural data are available to provide insight into how substrate binds to the ACC oxidase active site, nor has a binding site for ascorbate been defined.

### *I.C.3. Deacetoxycephalosporin synthase (DAOCS): "Booby trapped" mechanism*

Like HPP (chapter 2), penicillin N, the DAOCS substrate, binds directly to the iron during catalysis. In contrast to HPP, penicillin N binding does not induce a conformational change. Rather, DAOCS has been proposed to trigger ferryl formation upon direct coordination of substrate with iron via a ping-pong mechanism. An argument

is presented in chapter 2 in an attempt to rationalize why HppE does not employ such a mechanism.

DAOCS is an  $\alpha$ -KG dependent mononuclear iron enzyme that catalyzes the synthesis of all cephalosporin antibiotics using various penicillin derivatives (e.g. penicillin N) as substrates<sup>44</sup> (Figure I.14.). It utilizes a unique strategy for protection of high energy iron-oxo intermediates compared to TauD and IPNS<sup>30</sup>. The DAOCS mechanism of catalytic initiation is thought to be identical to that of TauD in that  $\alpha$ -KG activates the iron center for dioxygen binding and a ferric superoxide intermediate attacks the C2 carbonyl of  $\alpha$ -KG facilitating decarboxylation. However, the mechanism of DAOCS diverges here and a cyclic peroxo intermediate is thought to be a precursor to a planar peroxo form, which is estimated to be lower in energy when compared to the ferryl-oxo species<sup>30</sup>. Based on kinetic, computational, and crystallographic evidence, substrate binding to the DAOCS active site, at the same site where  $\alpha$ -KG binds, forces the decomposition of this planar peroxo intermediate, loss of succinate and carbon dioxide, and formation of a more reactive ferryl-oxo species<sup>30</sup>. Exchange of the negatively charged succinate product with a neutral thioester is thought to increase the reactivity of this ferryl-oxo species facilitating hydrogen atom abstraction of the proximal methyl moiety on the penicillin substrate (Figure I.15.). In summary, DACOS is designed to stabilize a peroxo intermediate, i.e. protect the reactive species until substrate is present<sup>30</sup>. This strategy allows for a ping-pong mechanism without damage from highly reactive Fe-oxo species in the absence of substrate.

#### *I.C.4. Clavaminic Synthase (CAS): A ligand rearrangement mechanism*

HppE has been shown to catalyze different reactions using different substrates and substrate analogs (chapters 3 and 4). Similarly, CAS is a mononuclear iron enzyme that catalyzes three distinct reactions in the biosynthetic pathway of clavulanic acid and is therefore fine-tuned to bind three unique substrates (Figure I.16.)<sup>58</sup>. In an  $\alpha$ -KG dependent fashion, the first reaction is a hydroxylation of an unactivated methylene carbon, a process that likely to involve an oxygen rebound similar to that proposed in TauD catalysis and P450 enzymes<sup>31,35,51</sup>. The second reaction catalyzed by CAS is an oxidative ring cyclization that occurs subsequent to hydrogen atom abstraction by a putative ferryl-oxo species. This reaction is quite similar to the epoxidation reaction catalyzed by HppE. Finally, CAS facilitates the hydrogenation of a third substrate.

In addition to being able to catalyze three reactions, CAS is proposed to have a unique mechanism of catalytic initiation relative to other cupin mononuclear iron enzymes. Similar to the TauD paradigm,  $\alpha$ -KG binds in a bidentate fashion to the metal center and is activated for dioxygen binding in the presence of deoxyguanidinoproclavaminic acid (substrate 1). Rather than binding to the open coordination site, dioxygen surrogate nitric oxide binds at the site coincident with the carboxylate ligand from  $\alpha$ -KG requiring a ligand rearrangement mechanism where this carboxylate ligand moves to the open coordination site (Figure I.17.)<sup>32</sup>. Paradoxically, substrate is not positioned in an appropriate orientation, relative to this new dioxygen binding site, to facilitate hydrogen atom abstraction at the  $\beta$ -methylene. To address this paradox, a rearrangement of the ferryl intermediate has been proposed, although little evidence is available to support this mechanism<sup>32</sup>.

### *I.D. Thesis Goals*

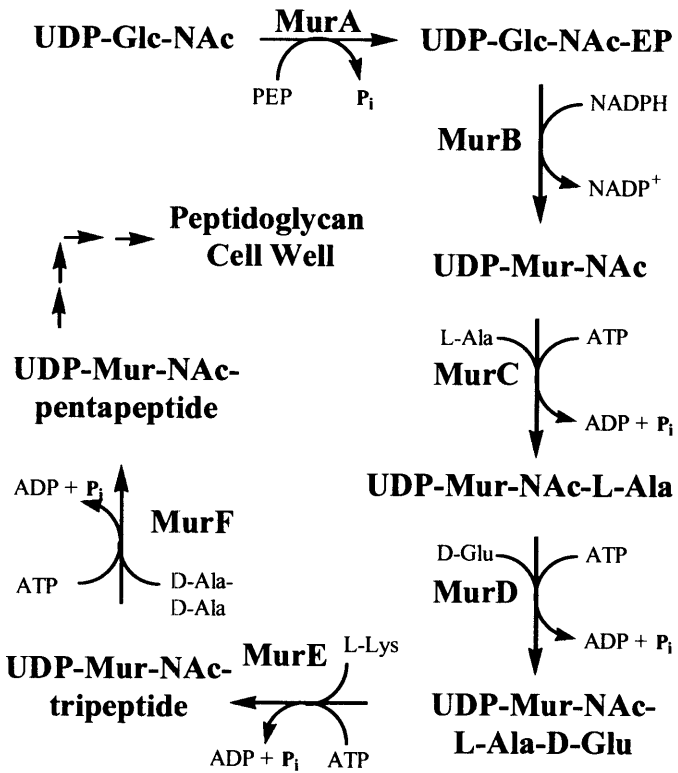
In this thesis, X-ray crystallography is used to answer the following questions about HppE: (1) how HppE is able to recognize and respond to its substrate, *S*-HPP, (2) why *S*-HPP is able to prime iron for dioxygen binding in the absence of  $\alpha$ -KG, (3) how the unique epoxidation reaction catalyzed by HppE may occur, (4) how HppE is able to catalyze two distinct reactions, recognizing both enantiomers of substrate (*R*- and *S*-HPP), (5) how HppE is able to bind larger hydrophobic substrates while maintaining complete regioselectivity, and (6) how HppE compares to other enzymes within the cupin superfamily?



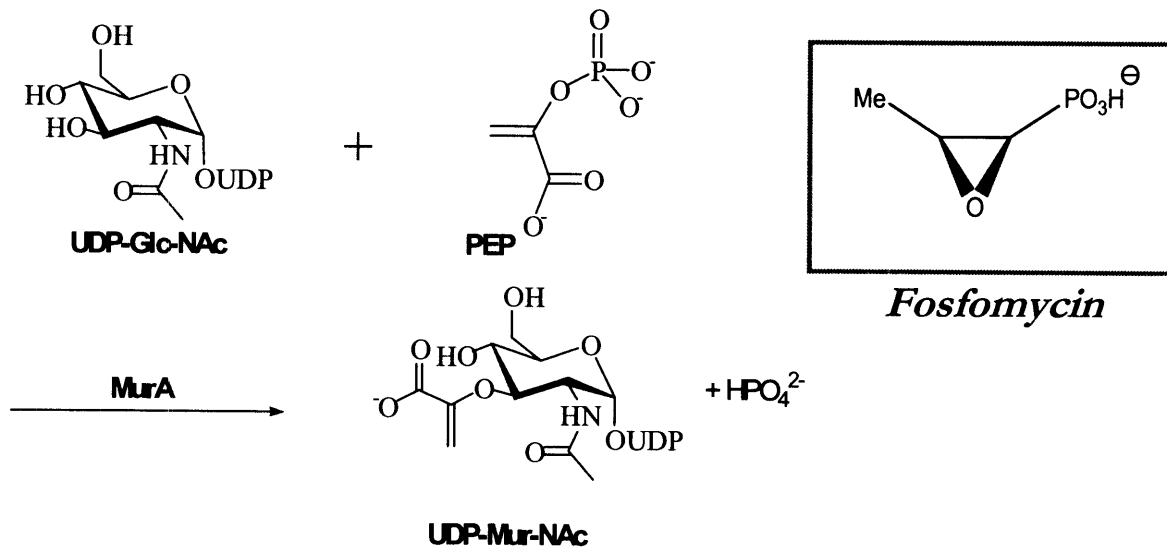
## **I.E. Figures**

**Figure I.1. Bacterial cell wall biosynthesis. (a)** MurA (enolpyruvyl-UDP-GlcNAc synthase) catalyzes the transfer of the enolpyruvyl moiety of phosphoenolpyruvate (PEP) to UDP-*N*-acetylglucosamine (UDP-Glc-Nac) resulting in the formation of enolpyruvyl-UDP-*N*-acetylglucosamine (UDP-GlcNAc-EP) and phosphate (Pi). MurB catalyzes the reduction of UDP-GlcNAc-EP to UDP-*N*-acetylmuramic acid (UDP-Mur-NAc). Enzymes MurC-MurF catalyze the conjugation of a pentapeptide to UDP-Mur-NAc forming UDP-Mur-NAc-pentapeptide in an ATP dependent fashion. UDP-Mur-NAc-pentapeptide is a vital constituent of the bacterial cell wall. **(b)** The MurA reaction is shown with the MurA inhibitor, fosfomicin, shown in the yellow inset.

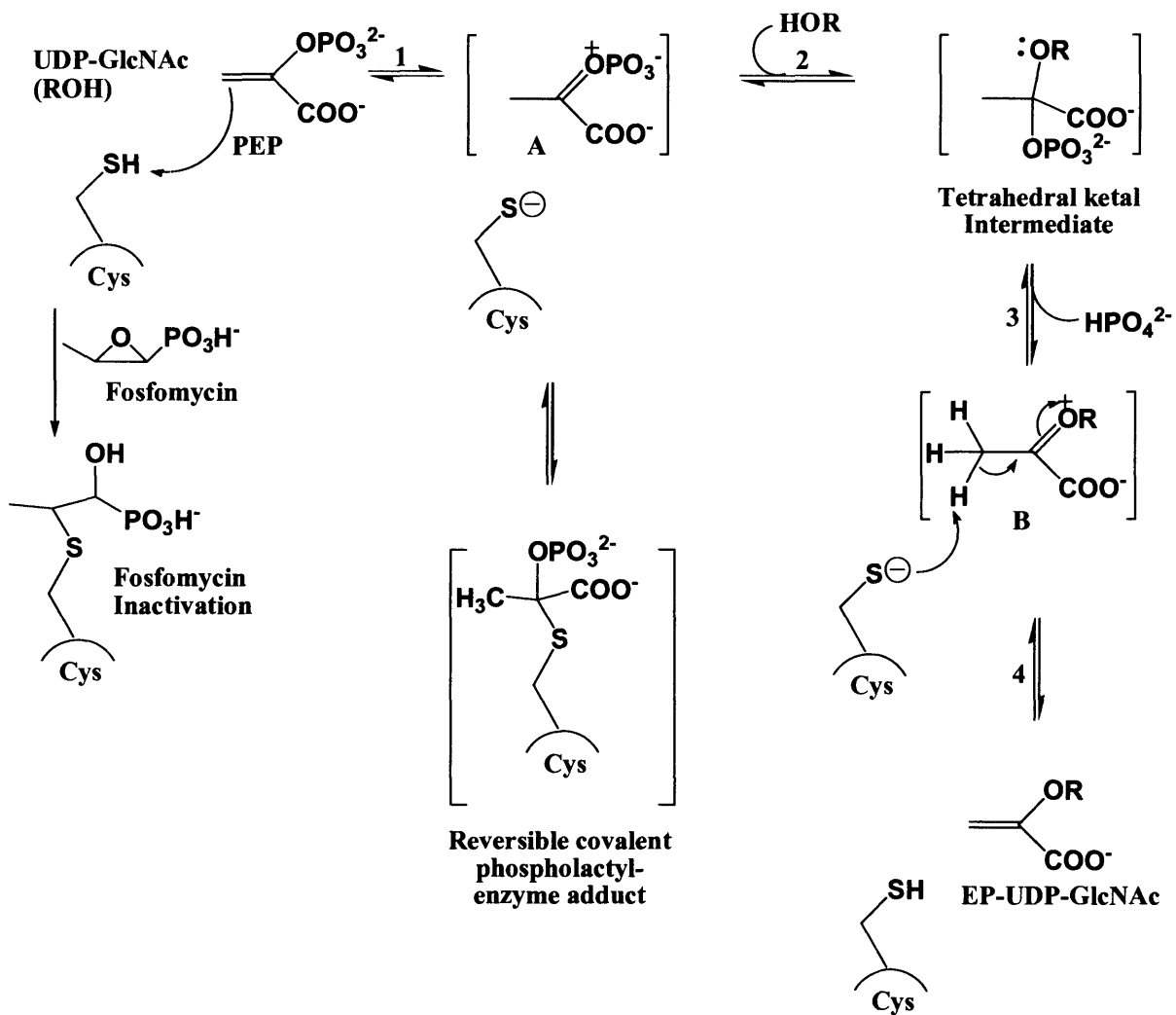
(a)



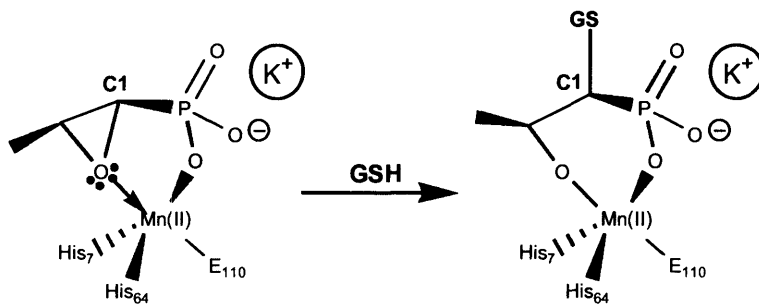
(b)



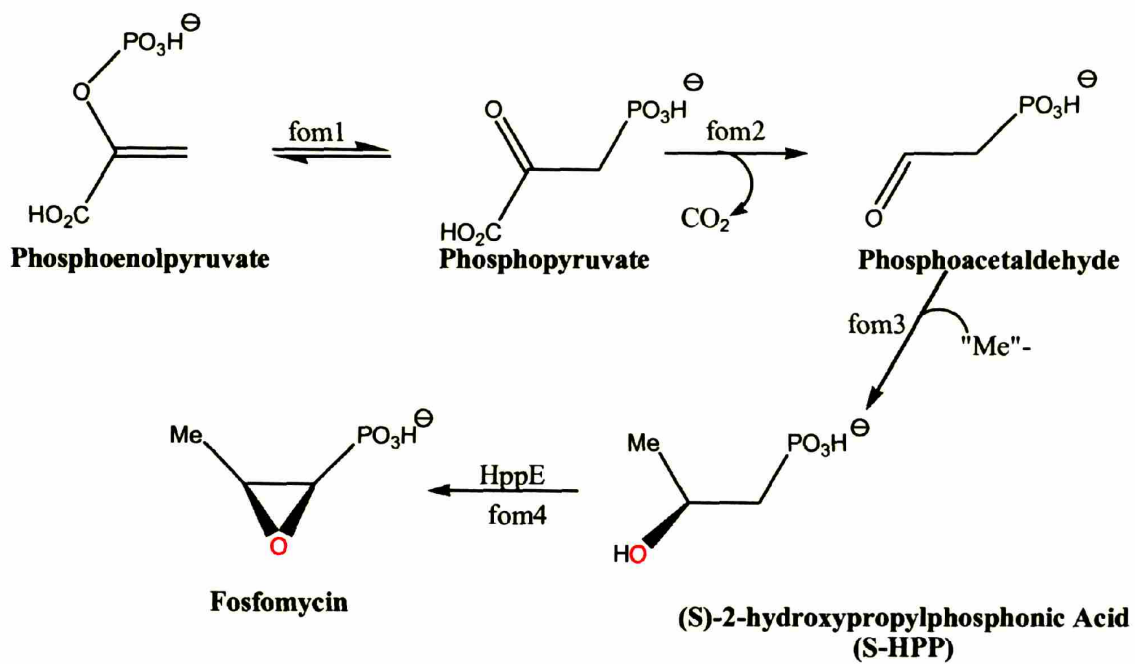
**Figure I.2. Fosfomicin mechanism of MurA inhibition<sup>11</sup>.** MurA facilitates the tautomerization of PEP (1) via the active site Cys. The C3 hydroxyl of UDP-GlcNac (ROH) attacks the electrophilic center in **A**, resulting in dephosphorylation (2,3). Tautomerization of **B**, again facilitated by Cys, yields EP-UDP-GlcNac (4). A reversible covalent phospholactyl-enzyme adduct, not required for catalysis, forms subsequent to Cys attack of **A**. In contrast, an irreversible fosfomicin adduct forms subsequent to Cys attack of the epoxide.



**Figure I.3. FosA catalyzed inactivation of fosfomycin<sup>20</sup>.** Fosfomycin binds in a bidentate fashion to the Mn(II) center of FosA via the epoxide and phosphonate oxygens. The phosphonic acid moiety is further stabilized by a potassium cation. Fosfomycin binding to FosA increases the electrophilicity at C1 and facilitates the conjugation of glutathione (GSH).

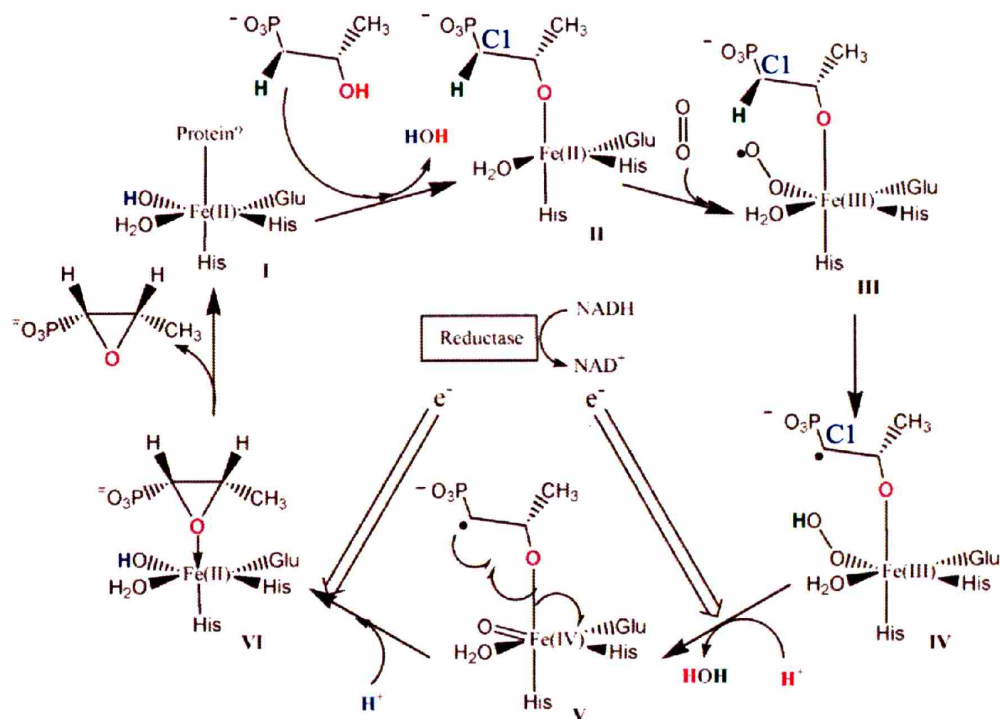


**Figure I.4. Biosynthesis of fosfomycin in *Streptomyces wedmorensis*.**

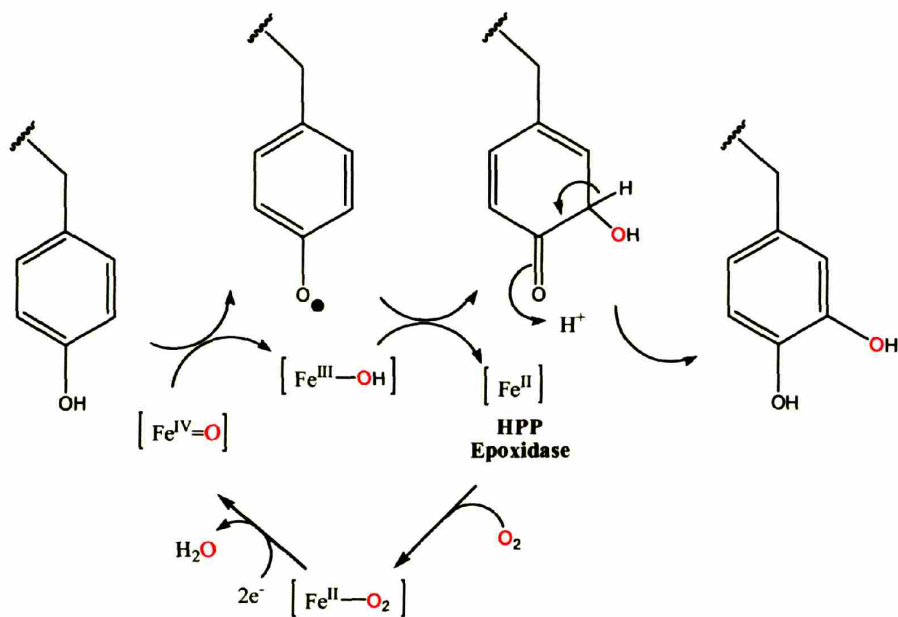


**Figure I.5. One Possible Mechanism of HppE Catalysis<sup>23</sup>.**

The hydrogen atom being abstracted is shown in green. The substrate hydroxyl oxygen and hydrogen are shown in purple and red, respectively. Protons added during catalysis are shown in blue and pink. The C1 position of substrate is shown in blue.

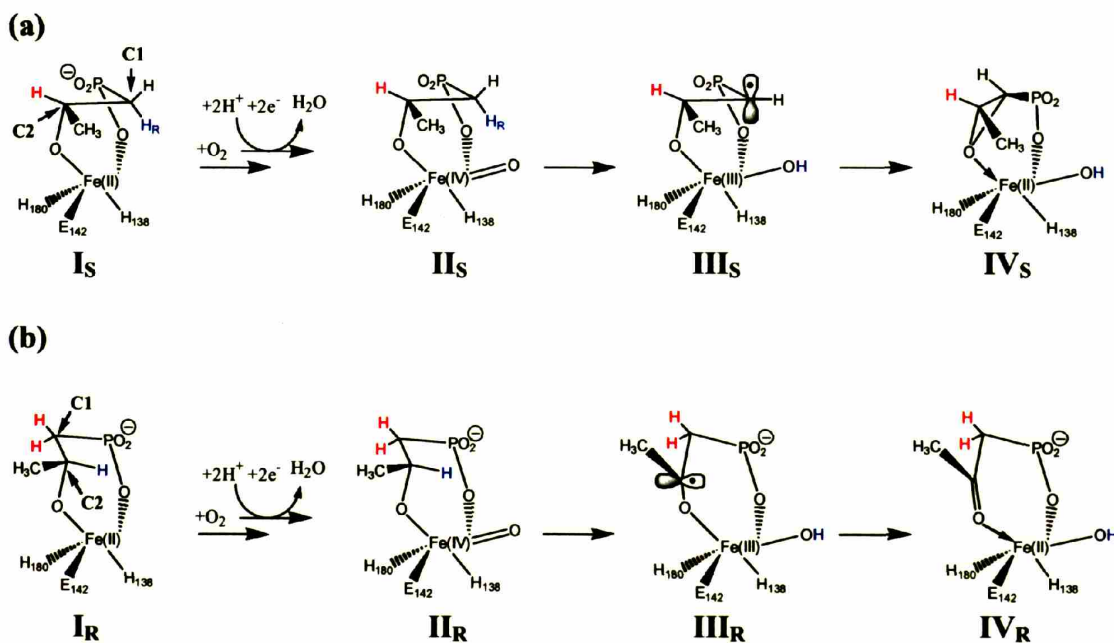


**Figure I.6. HppE self-hydroxylation of tyrosine<sup>27</sup>.** Oxygen atoms derived from dioxygen are shown in red.



**Figure I.7. Potential mechanisms of regioselective H-atom Abstraction.**

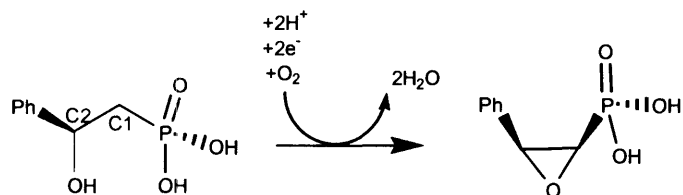
**(a)** A scheme illustrating the conversion of *S*-HPP ( $I_S$ ) to fosfomycin ( $IV_S$ ). Triad ligands are labeled throughout. The putative ferryl-oxo intermediate ( $II_S$ ) and the C1-centered radical intermediate ( $III_S$ ) are shown. Based on the structure data presented herein, the pro-R hydrogen atom at the C2-position *S*-HPP is accessible to H-atom abstraction (blue) while the hydrogen atom at the C1-position is inaccessible to H-atom abstraction (red). **(b)** A scheme illustrating the conversion of *R*-HPP ( $I_R$ ) to 2-oxo-propylphosphonic acid ( $IV_R$ ). The putative ferryl-oxo intermediate ( $II_R$ ) and the C2-centered radical intermediate ( $III_R$ ) are shown. The pro-R hydrogen atom at the C2-position *R*-HPP is inaccessible to H-atom abstraction (red) while the hydrogen atom at the C1-position is accessible to H-atom abstraction (blue).



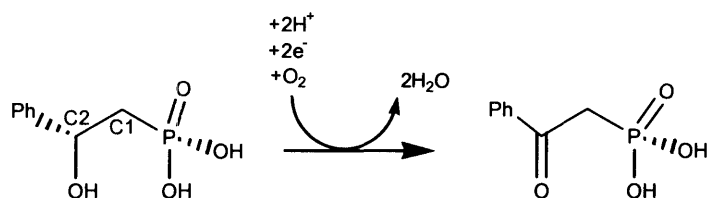


**Figure I.8. HppE catalysis using HPEP substrates.** HppE converts **(a)** (*S*)-2-hydroxy-2-phenylethylphosphonic acid to (1*R*,2*S*)-2-phenyl-1,2-epoxyethylphosphonic acid and **(b)** (*R*)-2-hydroxy-2-phenylethylphosphonic acid to 2-oxo-2-phenylethylphosphonic acid. The C1 and C2 positions of each substrate are labeled (blue).

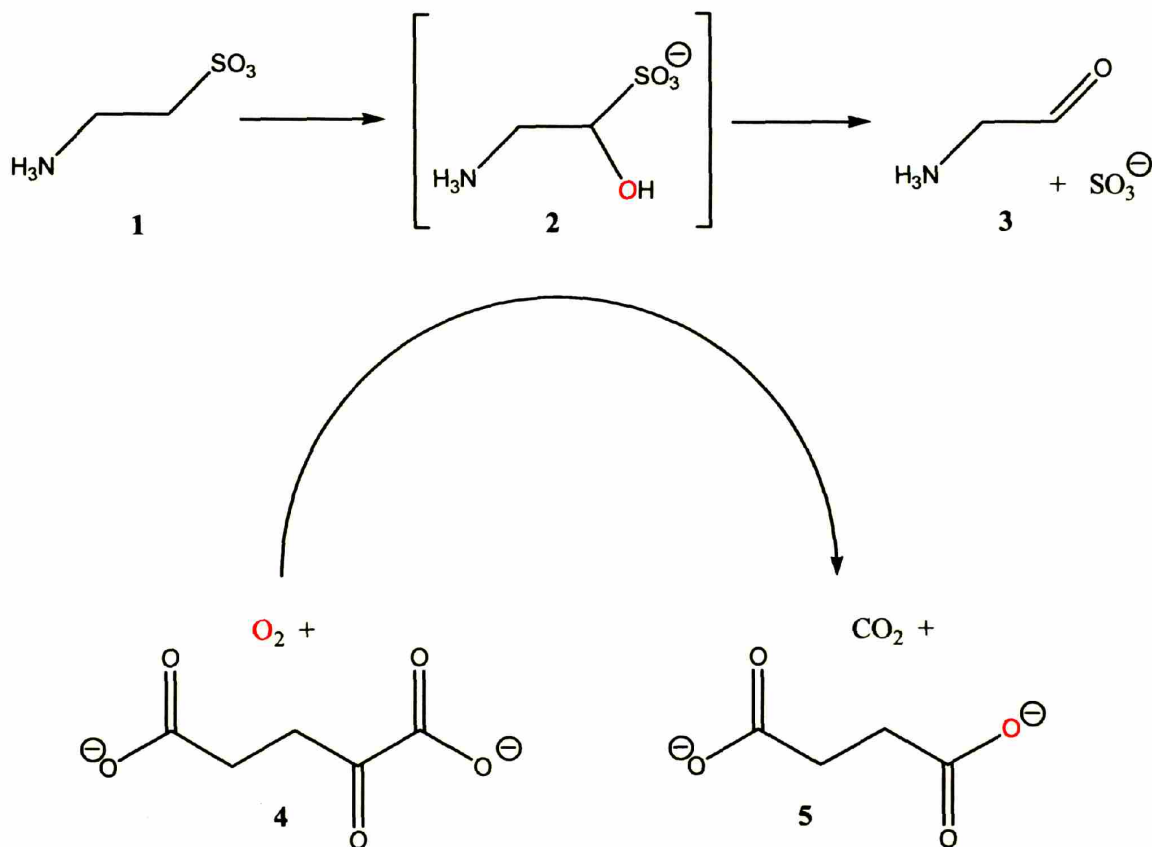
**(a)**



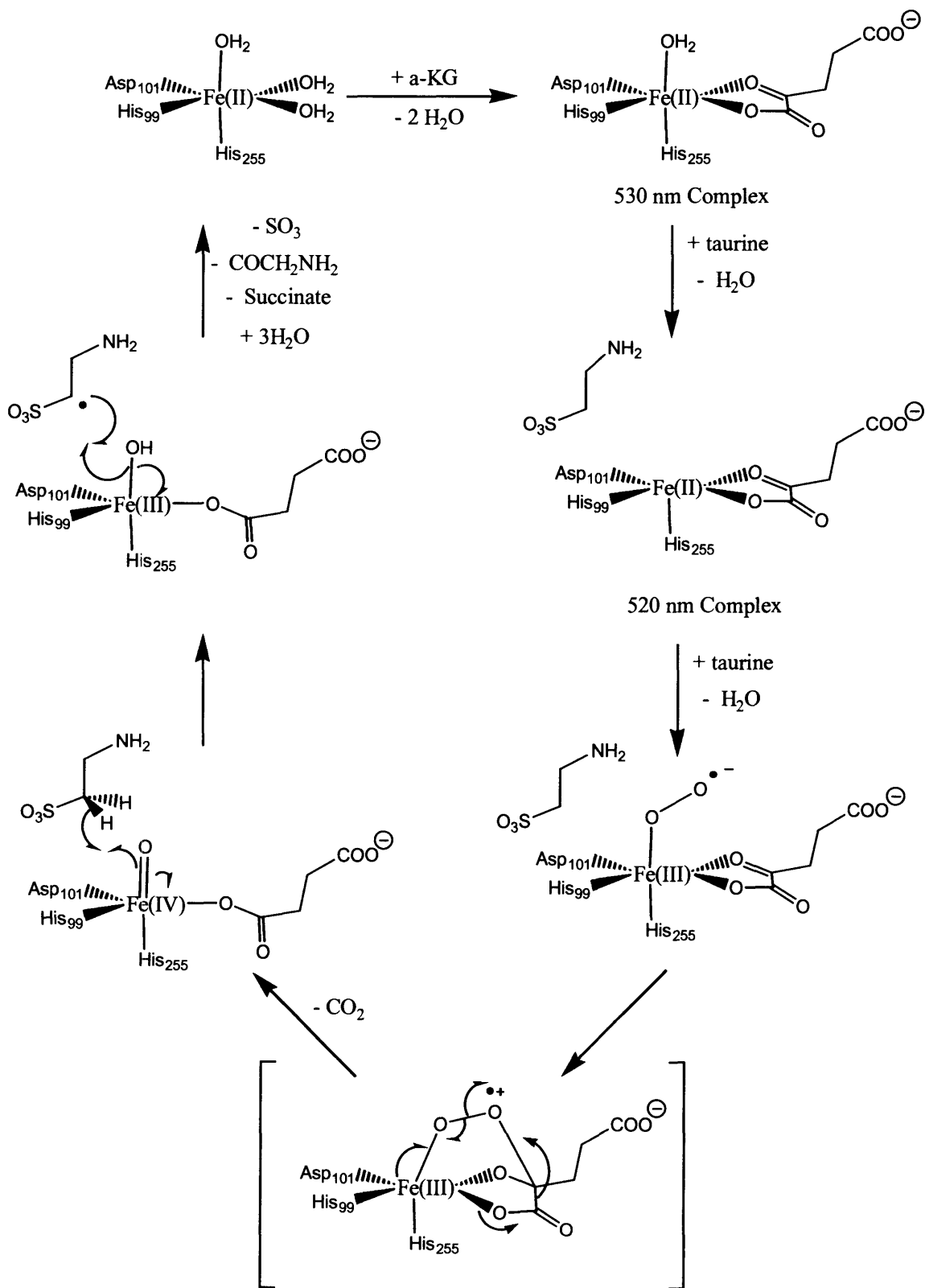
**(b)**



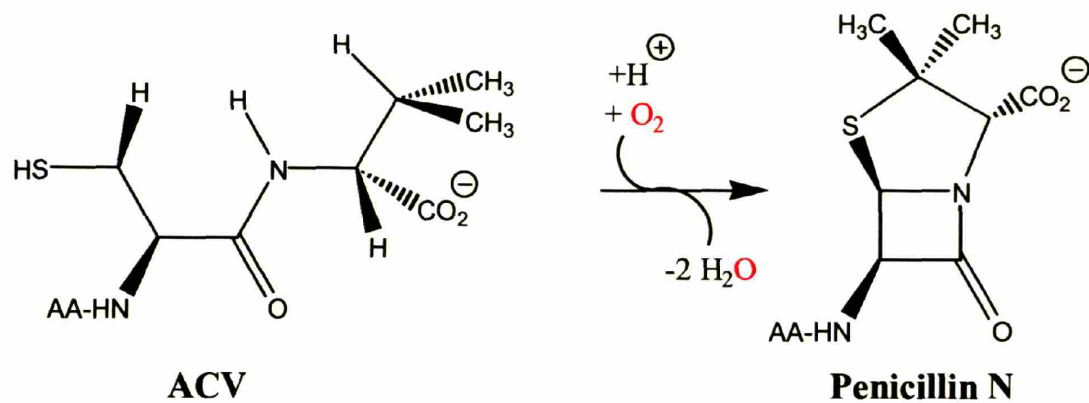
**Figure I.9. The reaction catalyzed by TauD<sup>31</sup>.** Substrate oxidation is coupled to  $\alpha$ -KG decarboxylation as illustrated in the context of taurine dioxygenase. Oxygen atoms derived from molecular oxygen are colored red. (1: Taurine; 2: 2-amino-hydroxyethylsulfonic acid; 3: aminoacetaldehyde; 4:  $\alpha$ -KG; 5: succinate).



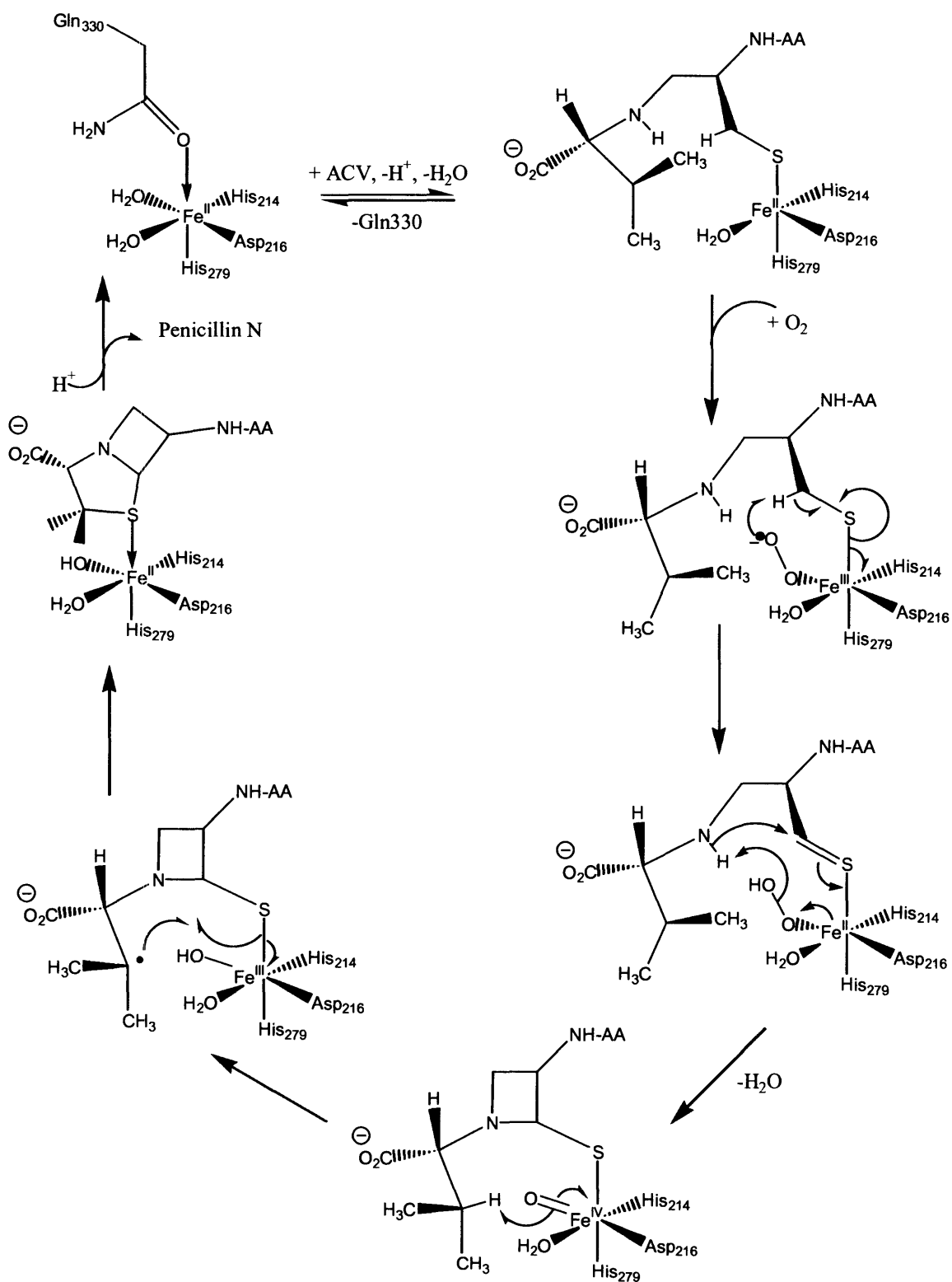
**Figure I.10. Proposed mechanism of taurine hydroxylation by TauD<sup>34,50,51</sup>.**



**Figure I.11. The reaction catalyzed by IPNS<sup>45</sup>.** Oxygen atoms derived from molecular oxygen are colored red. Both oxygen atoms are lost as water.



**Figure I.12. Proposed mechanism of IPNS catalysis<sup>29,48</sup>.**



**Figure I.13. Reaction catalyzed by ACCO<sup>28,47</sup>.**

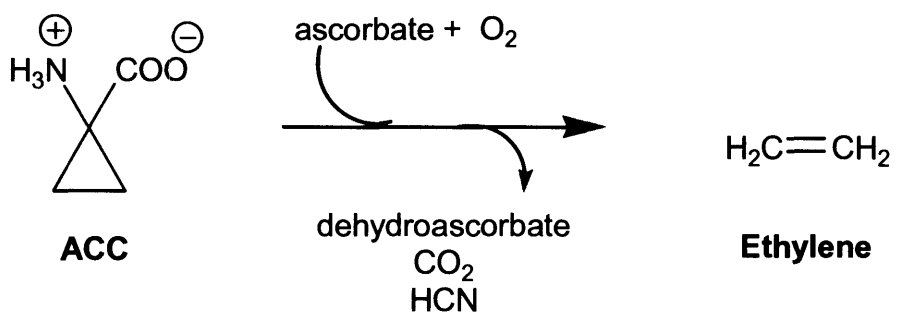
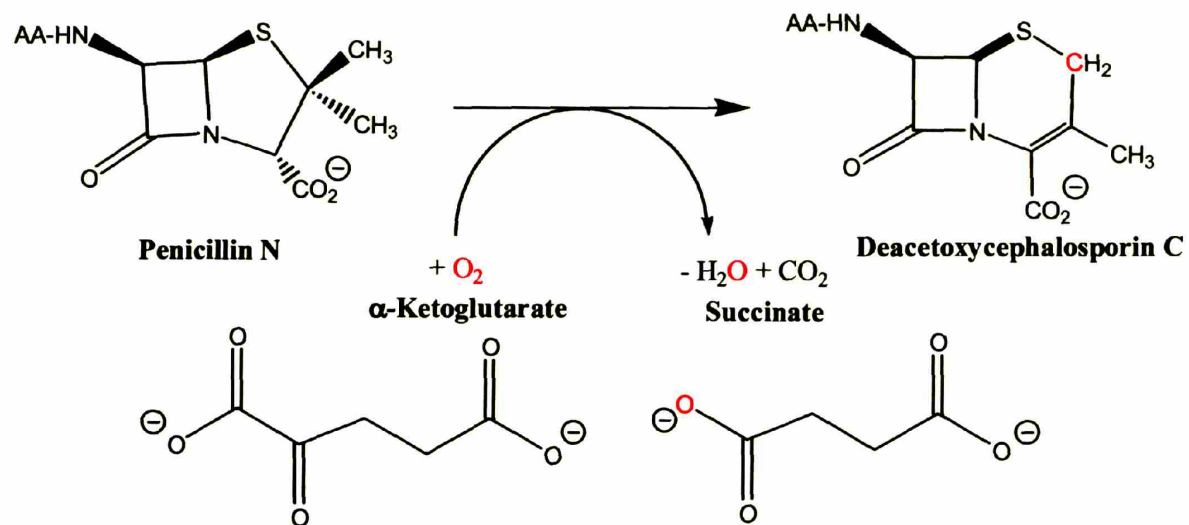
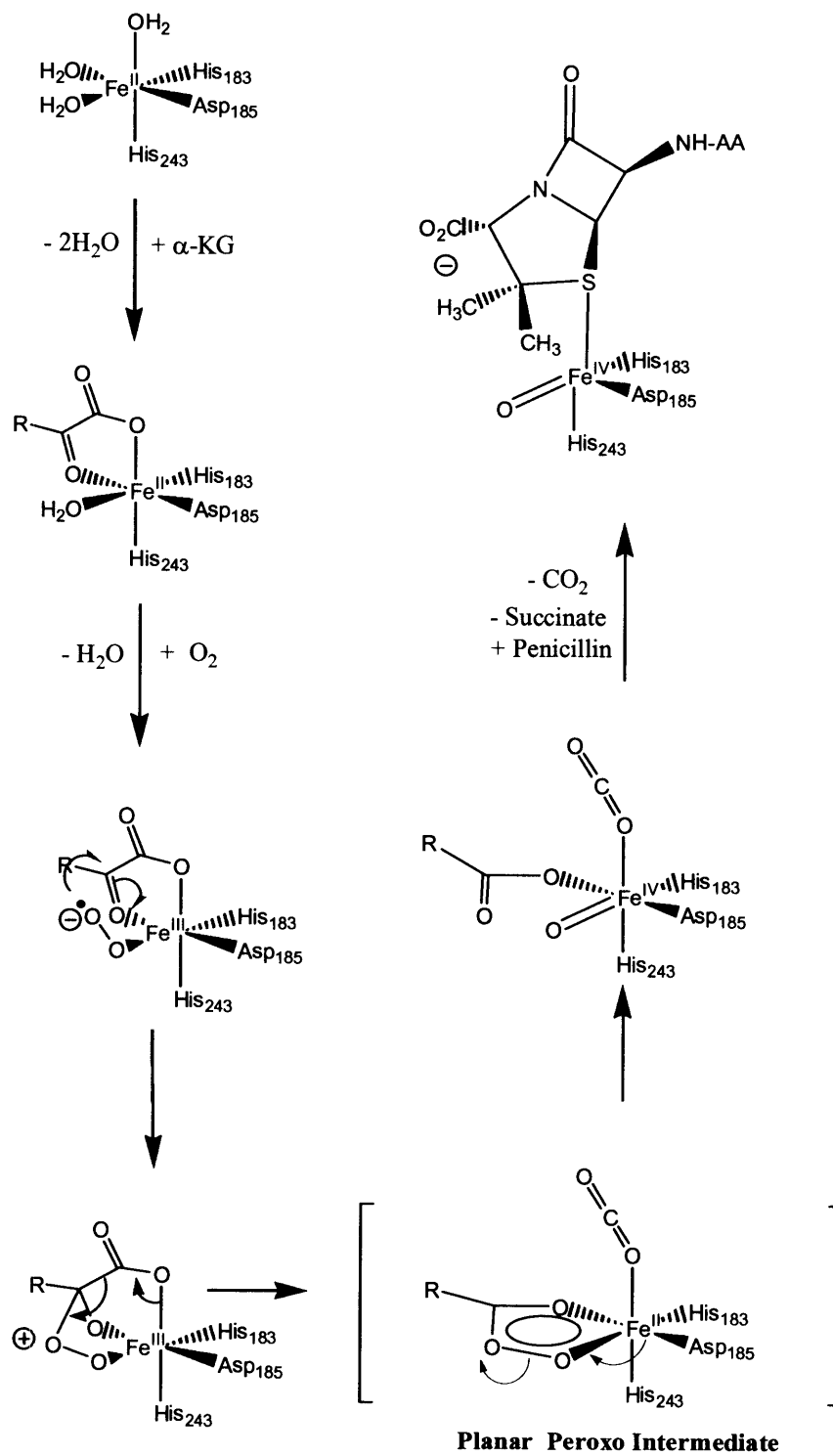


Figure I.14. Reaction catalyzed by DAOCS<sup>44</sup>.



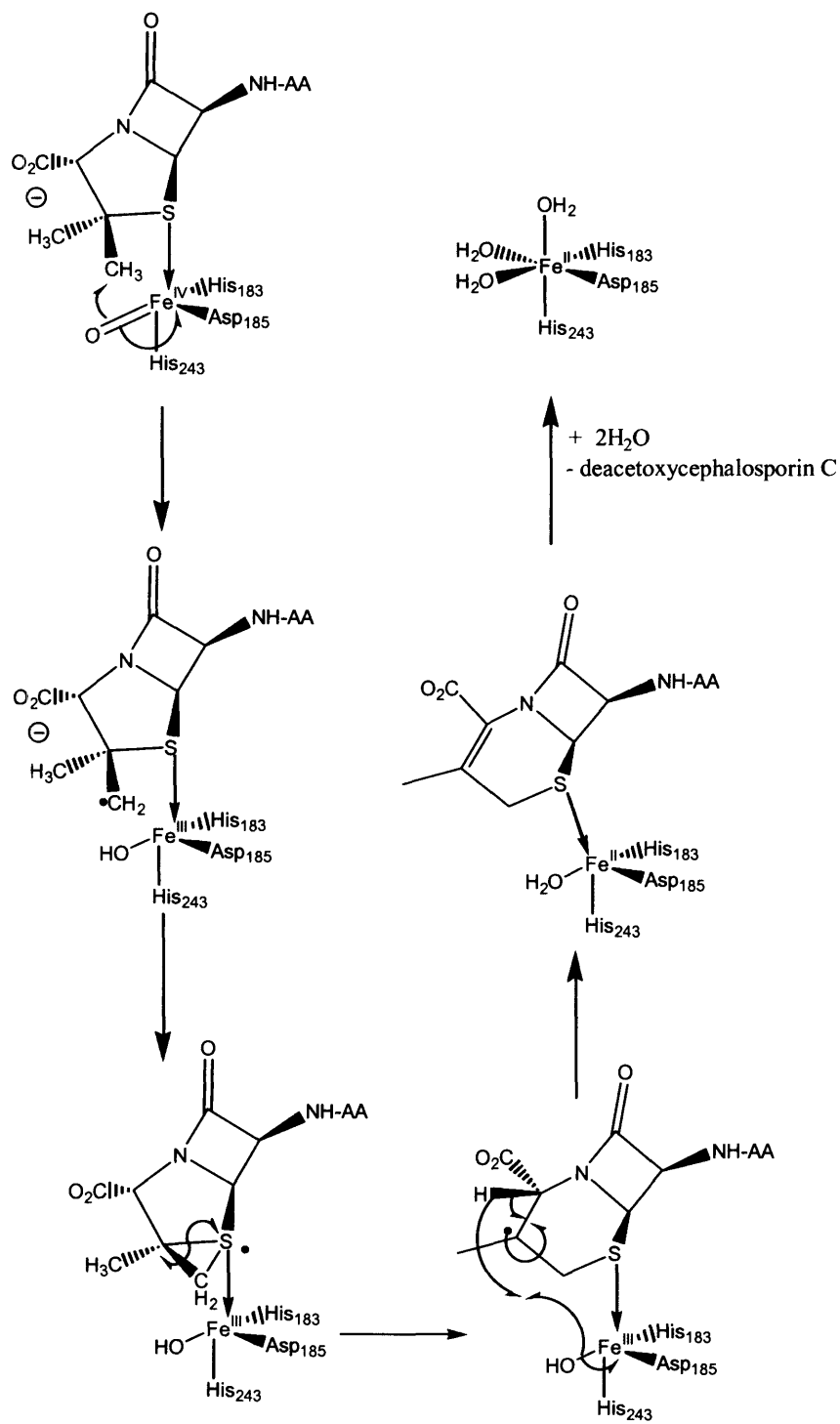
**Figure I.15. Proposed mechanism of DAOCS catalysis<sup>30</sup>.**

**(a) Substrate binding triggers “booby-trapped” Fe-oxo species**

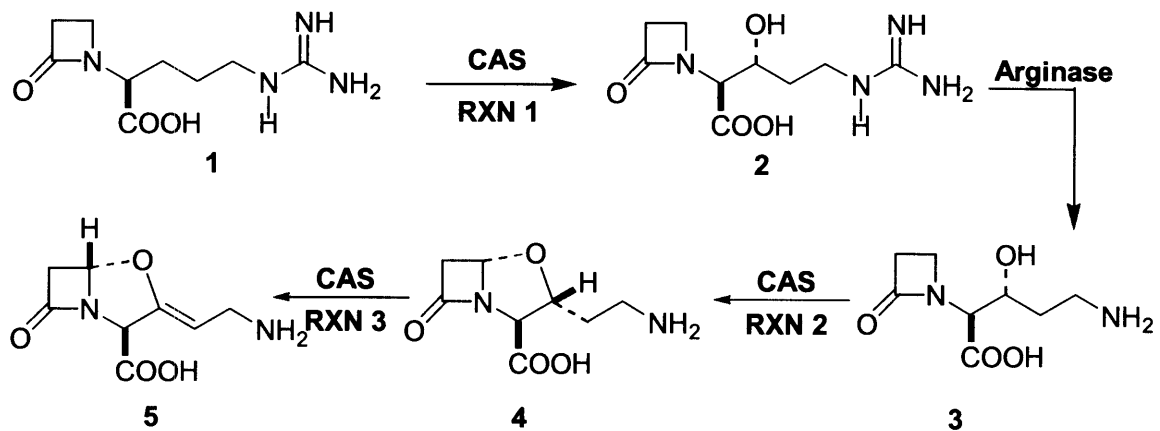




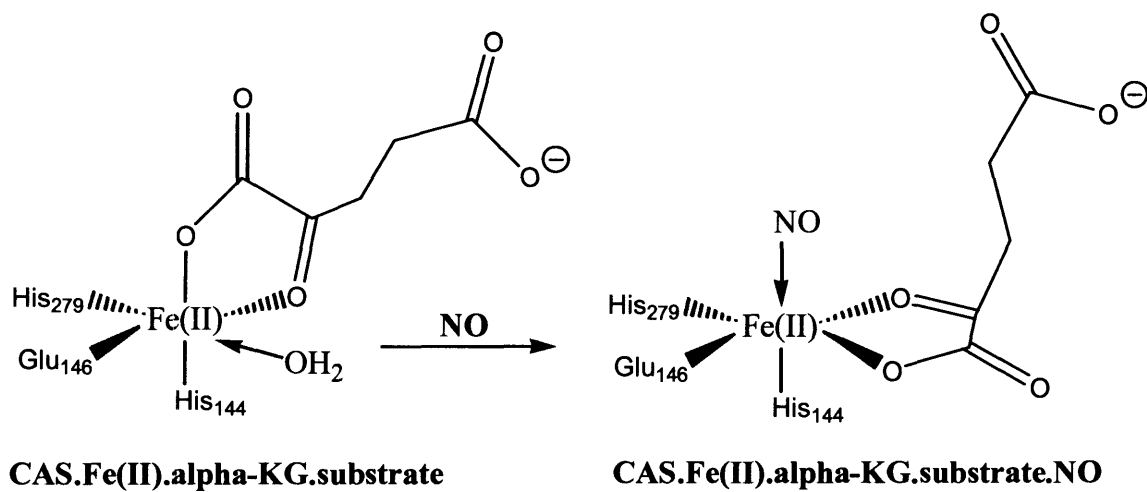
**(b) hydrogen atom abstraction and cephalosporin synthesis**



**Figure I.16. Reactions catalyzed by CAS<sup>58</sup>. 1, deoxyguanidinoproclavaminic acid; 2, guanidinoproclavaminic acid; 3, proclavaminic acid; 4.**



**Figure I.17. Ligand rearrangement mechanism of CAS upon NO binding<sup>32</sup>.**



## I.F. References

1. Anderson, R. U. Management of lower urinary tract infections and cystitis. *Urol Clin North Am* **26**, 729-735, viii (1999).
2. Hamilton, E. M. Management of upper and lower urinary tract infections in pregnant women. *J Am Acad Nurse Pract* **8**, 599-663 (1996).
3. Bailey, R. R. Management of lower urinary tract infections. *Drugs* **45 Suppl 3**, 139-144 (1993).
4. Nicolle, L. E. The optimal management of lower urinary tract infections. *Infection* **18 Suppl 2**, S50-52 (1990).
5. Nakazawa, H., Kikuchi, Y., Honda, T., Isago, T. & Nozaki, M. Enhancement of antimicrobial effects of various antibiotics against methicillin-resistant *Staphylococcus aureus* (MRSA) by combination with fosfomycin. *J Infect Chemother* **9**, 304-309 (2003).
6. Cassone, M., Campanile, F., Pantosti, A., Venditti, M. & Stefani, S. Identification of a variant "Rome clone" of methicillin-resistant *Staphylococcus aureus* with decreased susceptibility to vancomycin, responsible for an outbreak in an intensive care unit. *Microb Drug Resist* **10**, 43-49 (2004).
7. Schito, G. C. Why fosfomycin trometamol as first line therapy for uncomplicated UTI? *Int J Antimicrob Agents* **22 Suppl 2**, 79-83 (2003).
8. Kahan, F. M., Kahan, J. S., Cassidy, P. J. & Kropp, H. The mechanism of action of fosfomycin (phosphonomycin). *Ann N Y Acad Sci* **235**, 364-386 (1974).
9. El Zoeiby, A., Sanschagrín, F. & Levesque, R. C. Structure and function of the Mur enzymes: development of novel inhibitors. *Mol Microbiol* **47**, 1-12 (2003).

10. Marquardt, J. L., Brown, ED, Lane, WS, Haley, TM, Ichikawa, Y, Wong, CH & Walsh, CT. Kinetics, stoichiometry, and identification of the reactive thiolate in the inactivation of UDP-GlcNAc enolpyruvyl transferase by the antibiotic fosfomycin. *Biochemistry* **33**, 10646-10651 (1994).
11. Kim, D. H., Lees, WJ, Kempell, KE, Lane, WS, Duncan, K & Walsh, CT. Characterization of a Cys115 to Asp substitution in the *Escherichia coli* cell wall biosynthetic enzyme UDP-GlcNAc enolpyruvyl transferase (MurA) that confers resistance to inactivation by the antibiotic fosfomycin. *Biochemistry* **35**, 4923-4928 (1996).
12. Armstrong, R. N. Mechanistic imperatives for the evolution of glutathione transferases. *Curr Opin Chem Biol* **2**, 618-623 (1998).
13. Cao, M., Bernat, B. A., Wang, Z., Armstrong, R. N. & Helmann, J. D. FosB, a cysteine-dependent fosfomycin resistance protein under the control of sigma(W), an extracytoplasmic-function sigma factor in *Bacillus subtilis*. *J Bacteriol* **183**, 2380-2383 (2001).
14. Bernat, B. A. & Armstrong, R. N. Elementary steps in the acquisition of Mn<sup>2+</sup> by the fosfomycin resistance protein (FosA). *Biochemistry* **40**, 12712-12718 (2001).
15. Bernat, B. A., Laughlin, L. T. & Armstrong, R. N. Fosfomycin resistance protein (FosA) is a manganese metalloglutathione transferase related to glyoxalase I and the extradiol dioxygenases. *Biochemistry* **36**, 3050-3055 (1997).
16. Garcia, P., Arca, P. & Evaristo Suarez, J. Product of fosC, a gene from *Pseudomonas syringae*, mediates fosfomycin resistance by using ATP as cosubstrate. *Antimicrob Agents Chemother* **39**, 1569-1573 (1995).

17. Arca, P., Hardisson, C. & Suarez, J. E. Purification of a glutathione S-transferase that mediates fosfomycin resistance in bacteria. *Antimicrob Agents Chemother* **34**, 844-848 (1990).
18. Villar, C. J., Hardisson, C. & Suarez, J. E. Cloning and molecular epidemiology of plasmid-determined fosfomycin resistance. *Antimicrob Agents Chemother* **29**, 309-314 (1986).
19. Beharry, Z. & Palzkill, T. Functional analysis of active site residues of the fosfomycin resistance enzyme FosA from *Pseudomonas aeruginosa*. *J Biol Chem* **280**, 17786-17791 (2005).
20. Rife, C. L., Pharris, R. E., Newcomer, M. E. & Armstrong, R. N. Crystal structure of a genomically encoded fosfomycin resistance protein (FosA) at 1.19 Å resolution by MAD phasing off the L-III edge of Tl(+). *J Am Chem Soc* **124**, 11001-11003 (2002).
21. Seto, H. & Kuzuyama, T. Bioactive natural products with carbon-phosphorus bonds and their biosynthesis. *Nat Prod Rep* **16**, 589-596 (1999).
22. Hidaka, T., Goda, M, Kuzuyama, T, Takei, N, Hidaka, M & Seto, H. Cloning and nucleotide sequence of fosfomycin biosynthetic genes of *Streptomyces wedmorensis*. *Mol Gen Genet* **249**, 274-280 (1995).
23. Liu, P., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Biochemical and spectroscopic studies on (S)-2-hydroxypropylphosphonic acid epoxidase: a novel mononuclear non-heme iron enzyme. *Biochemistry* **42**, 11577-11586 (2003).

24. Zhao, Z., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Mechanistic studies of HPP epoxidase: configuration of the substrate governs its enzymatic fate. *Angew Chem Int Ed Engl* **41**, 4529-4532 (2002).
25. Liu, P. in *Department of Chemistry 229* (University of Minnesota, 2001).
26. Liu, P., Murakami, K, Seki, T, He, X, Yeung, SM, Kuzuyama, T, Seto, H & Liu, H. Protein purification and function assignment of the epoxidase catalyzing the formation of fosfomycin. *J Am Chem Soc* **123**, 4619-4620 (2001).
27. Liu, P, Mehn, MP, Yan, F, Zhao, Z, Que, L Jr & Liu, HW. Oxygenase activity in the self-hydroxylation of (s)-2-hydroxypropylphosphonic acid epoxidase involved in fosfomycin biosynthesis. *J Am Chem Soc* **126**, 10306-10312 (2004).
28. Dong, J. G., Fernandez-Maculet, J. C. & Yang, S. F. Purification and characterization of 1-aminocyclopropane-1-carboxylate oxidase from apple fruit. *Proc Natl Acad Sci U S A* **89**, 9789-9793 (1992).
29. Roach, P. L., Clifton, IJ, Hensgens, CM, Shibata, N, Schofield, CJ, Hajdu, J & Baldwin, JE. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830 (1997).
30. Valegard, K., Terwisscha, van Scheltinga, AC, Dubus, A, Ranghino, G, Oster, LM, Hajdu, J & Andersson, I. The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. *Nat Struct Mol Biol* **11**, 95-101 (2004).
31. Eichhorn, E., van der Ploeg, J. R., Kertesz, M. A. & Leisinger, T. Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J Biol Chem* **272**, 23031-23036 (1997).

32. Zhang, Z., Ren, J, Harlos, K, McKinnon, CH, Clifton, IJ & Schofield, CJ. Crystal structure of a clavamate synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett* **517**, 7-12 (2002).
33. Yan, F. in *Chemistry 222* (University of Texas, Austin, 2005).
34. Price, J. C., Barr, E. W., Tirupati, B., Bollinger, J. M., Jr. & Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **42**, 7497-7508 (2003).
35. Guengerich, F. P. Cytochrome P450 oxidations in the generation of reactive electrophiles: epoxidation and related reactions. *Arch Biochem Biophys* **409**, 59-71 (2003).
36. Woschek, A., Wuggenig, F., Peti, W. & Hammerschmidt, F. On the transformation of (S)-2-hydroxypropylphosphonic acid into fosfomicin in *Streptomyces fradiae*--a unique method of epoxide ring formation. *Chembiochem* **3**, 829-835 (2002).
37. Dunwell, J. M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**, 7-17 (2004).
38. Hanauske-Abel, H. M. & Gunzler, V. A stereochemical concept for the catalytic mechanism of prolylhydroxylase: applicability to classification and design of inhibitors. *J Theor Biol* **94**, 421-455 (1982).



39. Schofield, C. J. & Zhang, Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr Opin Struct Biol* **9**, 722-731 (1999).
40. Hegg, E. L. & Que, L., Jr. The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur J Biochem* **250**, 625-629 (1997).
41. Costas, M., Mehn, M. P., Jensen, M. P. & Que, L., Jr. Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem Rev* **104**, 939-86 (2004).
42. Mehn, M. P., Fujisawa, K., Hegg, E. L. & Que, L., Jr. Oxygen activation by nonheme iron(II) complexes: alpha-keto carboxylate versus carboxylate. *J Am Chem Soc* **125**, 7828-7842 (2003).
43. Bassan, A., Borowski, T. & Siegbahn, P. E. Quantum chemical studies of dioxygen activation by mononuclear non-heme iron enzymes with the 2-His-1-carboxylate facial triad. *Dalton Trans*, 3153-3162 (2004).
44. Kovacevic, S., Weigel, B. J., Tobin, M. B., Ingolia, T. D. & Miller, J. R. Cloning, characterization, and expression in *Escherichia coli* of the *Streptomyces clavuligerus* gene encoding deacetoxycephalosporin C synthetase. *J Bacteriol* **171**, 754-760 (1989).
45. Palissa, H., von Dohren, H., Kleinkauf, H., Ting, H. H. & Baldwin, J. E. Beta-lactam biosynthesis in a gram-negative eubacterium: purification and characterization of isopenicillin N synthase from *Flavobacterium sp.* strain SC 12.154. *J Bacteriol* **171**, 5720-5728 (1989).

46. Sami, M., Brown, T. J., Roach, P. L., Schofield, C. J. & Baldwin, J. E. Glutamine-330 is not essential for activity in isopenicillin N synthase from *Aspergillus nidulans*. *FEBS Lett* **405**, 191-194 (1997).
47. Zhang, Z., Barlow, J. N., Baldwin, J. E. & Schofield, C. J. Metal-catalyzed oxidation and mutagenesis studies on the iron(II) binding site of 1-aminocyclopropane-1-carboxylate oxidase. *Biochemistry* **36**, 15999-6007 (1997).
48. Roach, P. L., Clifton, IJ, Fulop, V, Harlos, K, Barton, GJ, Hajdu, J, Andersson, I, Schofield, CJ & Baldwin, JE. Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* **375**, 700-704 (1995).
49. Tierney, D. L., Rocklin, A. M., Lipscomb, J. D., Que, L., Jr. & Hoffman, B. M. ENDOR studies of the ligation and structure of the non-heme iron site in ACC oxidase. *J Am Chem Soc* **127**, 7005-7013 (2005).
50. Price, J. C., Barr, E. W., Glass, T. E., Krebs, C. & Bollinger, J. M., Jr. Evidence for hydrogen abstraction from C1 of taurine by the high-spin Fe(IV) intermediate detected during oxygen activation by taurine:alpha-ketoglutarate dioxygenase (TauD). *J Am Chem Soc* **125**, 13008-13009 (2003).
51. Ryle, M., J., Liu, A, Muthukumaran, RB, Ho, RY, Koehntop, KD, McCracken, J, Que, L Jr & Hausinger, RP. O<sub>2</sub>- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42**, 1854-1862 (2003).
52. Elkins, J. M., Ryle, MJ, Clifton, IJ, Dunning, Hotopp, JC, Lloyd, JS, Burzlaff, NI, Baldwin, JE, Hausinger, RP & Roach, PL. X-ray crystal structure of *Escherichia*

- coli* taurine/alpha-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **41**, 5185-5192 (2002).
53. O'Brien, J. R., Schuller, D. J., Yang, V. S., Dillard, B. D. & Lanzilotta, W. N. Substrate-induced conformational changes in *Escherichia coli* taurine/alpha-ketoglutarate dioxygenase and insight into the oligomeric structure. *Biochemistry* **42**, 5547-5554 (2003).
54. Dubus, A., Sami, M, Brown, TJ, Schofield, CJ, Baldwin, JE & Frere, JM. Studies of isopenicillin N synthase enzymatic properties using a continuous spectrophotometric assay. *FEBS Lett* **485**, 142-146 (2000).
55. Blackburn, J. M., Sutherland, J. D. & Baldwin, J. E. A heuristic approach to the analysis of enzymic catalysis: reaction of delta-(L-alpha-aminoadipoyl)-L-cysteinyl-D-alpha-aminobutyrate and delta-(L-alpha-aminoadipoyl)-L-cysteinyl-D-allylglycine catalyzed by isopenicillin N synthase isozymes. *Biochemistry* **34**, 7548-7562 (1995).
56. Hausinger, R. P. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit Rev Biochem Mol Biol* **39**, 21-68 (2004).
57. Zhang, Z., Ren, J. S., Clifton, I. J. & Schofield, C. J. Crystal structure and mechanistic implications of 1-aminocyclopropane-1-carboxylic acid oxidase--the ethylene-forming enzyme. *Chem Biol* **11**, 1383-1394 (2004).
58. Townsend, C. A. New reactions in clavulanic acid biosynthesis. *Curr Opin Chem Biol* **6**, 583-589 (2002).

## **Chapter II: Structural Insight into the Biosynthesis of the Antibiotic Fosfomycin by a Mononuclear Iron Enzyme\***

### **Summary**

The biosynthetic pathway of the clinically important antibiotic fosfomycin uses enzymes that catalyze reactions without precedent in biology. Among these is hydroxypropylphosphonic acid epoxidase, which represents a new subfamily of non-heme mononuclear iron enzymes. We have determined six X-ray structures of this enzyme: the apoenzyme at 2.0 Å resolution; a native Fe(II)-bound form at 2.4 Å resolution; a tris-(hydroxymethyl)aminomethane-Co(II)-enzyme complex structure at 1.8 Å resolution; a substrate-Co(II)-enzyme complex structure at 2.5 Å resolution; and two substrate-Fe(II)-enzyme complexes at 2.1 and 2.3 Å resolution, respectively. These structural data suggest how this enzyme is able to recognize and respond to its substrate with a conformational change that protects the radical-based intermediates formed during catalysis. Comparisons with other family members suggest why substrate binding is able to prime iron for dioxygen binding in the absence of  $\alpha$ -ketoglutarate (a co-substrate required by many mononuclear iron enzymes), and how the unique epoxidation reaction of hydroxypropylphosphonic acid epoxidase may occur.

\* A version of this chapter has been published: Higgins, L. J., Yan, F., Liu, P., Liu, H.W. & Drennan, C. L. Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature* (2005) (Online at [www.nature.com](http://www.nature.com)).

## II.A. Introduction

Iron-dependent (*S*)-2-hydroxypropylphosphonic acid epoxidase (HppE) catalyzes the final step in the biosynthesis of fosfomycin<sup>1</sup>. Fosfomycin, or (1*R*,2*S*)-1,2-epoxypropylphosphonic acid, is an antibiotic used in the treatment of lower urinary tract infections<sup>2</sup>, and is effective against methicillin-resistant<sup>3</sup> and vancomycin-resistant<sup>4</sup> strains of *Staphylococcus aureus*. In addition to the attention it receives as an antibiotic, fosfomycin is one of a few known natural products containing a carbon-phosphorus (C-P) bond. The biosynthesis of this unusual molecule in *Streptomyces wedmorensis* mirrors that of other C-P bond containing natural products with respect to the first two steps, the conversion of phosphoenolpyruvate (PEP) to phosphonopyruvate (PnPy) and the subsequent decarboxylation of PnPy to phosphonoacetaldehyde (PnAA)<sup>1</sup> (Figure II.1). After this, the pathways diverge, and PnAA is methylated by PnAA methylase in what is predicted by sequence and cofactor requirements to be a novel methylation reaction<sup>5</sup>. The product of this methylation, (*S*)-2-hydroxypropylphosphonic acid (*S*-HPP), is converted into the epoxide fosfomycin by HppE via an oxidative cyclization reaction with retention of the substrate hydroxyl oxygen atom<sup>6,7</sup>. This latter transformation may be unique in biology<sup>7</sup>, since epoxidation reactions, such as those catalyzed by the P-450 cytochrome family and non-heme iron dependent monooxygenases, typically involve the conversion of substrate to an epoxide with oxygen atom insertion from dioxygen<sup>8,9</sup>.

Based on sequence analysis, HppE is a member of the cupin superfamily, characterized by an antiparallel  $\beta$ -barrel ('*cupa*' is the Latin term for small barrel)<sup>10</sup>. Non-heme mononuclear iron-dependent proteins, such as HppE, are an interesting subset of this superfamily. These enzymes use a facial triad with two histidine ligands and one

aspartic acid or glutamic acid, His<sub>2</sub>(Glu/Asp), to catalyze a variety of different reactions, including DNA repair and antibiotic biosynthesis. An interesting question is how enzymes with the same structure, metal, and ligand triad can catalyze a diverse range of reactions.

With respect to the HppE enzyme mechanism, biochemical experiments show that the reaction is initiated by a stereospecific and regiospecific hydrogen atom abstraction to yield a substrate radical intermediate en route to epoxide formation<sup>11</sup>. A similar ring closure mechanism has been proposed in the biosynthesis of penicillin by IPNS<sup>12</sup>. Specifically, the bicyclic ring system of penicillin is derived from the four-electron oxidation of  $\delta$ -(L- $\alpha$ -aminoadipoyl)-L-cysteinyl-D-valine (ACV). HppE and IPNS differ from most characterized cupin mononuclear iron enzymes in that these enzymes do not catalyze oxygen insertion into substrate or co-substrate. Also, both HppE and IPNS are independent of cofactors/co-substrates such as  $\alpha$ -ketoglutarate ( $\alpha$ -KG), tetrahydropterin, ascorbate, and iron-sulfur clusters, utilized by other mononuclear non-heme iron proteins as a source of reducing equivalents<sup>7</sup>. The reason for this cofactor independence is different, however. Unlike fosfomicin biosynthesis, the conversion of ACV to penicillin N is a four-electron oxidation coupled to a four-electron reduction of dioxygen to water. Therefore, IPNS does not require an external electron source. In contrast, HppE requires two electrons transferred from an endogenous reductase, in addition to the two electrons supplied by substrate, to completely reduce dioxygen to water. The putative HppE reductase remains to be identified.

Herein we present the first crystal structures of HppE, a representative of a new subfamily of mononuclear iron enzymes, and discuss how this enzyme is specifically

tuned to facilitate the unusual radical-mediated epoxidation reaction found in the pathway of fosfomicin biosynthesis.

## **II.B. Methods**

### *II.B.1. Site-directed mutagenesis, protein purification, assays*

Site-directed mutagenesis of the HppE gene (*fom4*) was carried out in Dr. Ben Liu's laboratory using the previously reported procedure<sup>16</sup>. The oligonucleotide primers used for the Lys23Ala mutagenesis were pK23AFY\_1/pK23AFY\_2 (pK23A\_1: 5'-CGGCGCGAGCAGGTCGCGATGGACCACGCCGCC-3'; pK23AFY\_2: 5'-GGCGGCGTGGTCCATCGCGACCTGCTCGCGCCG-3'), and for Glu142Ala mutagenesis were pE142AFY\_1/pE142AFY\_2 (pE142AFY\_1: 5'-GCCACGCCGCAACGCGTTCCTCTTCGTGCTCG-3'; pE142AFY\_2: 5'-CGAGCACGAAGAGGAACGCGTTGCCGGCGTGCC-3'). In preparation of selenomethionine-labeled HppE (SeMet-HppE) performed in Dr. Ben Liu's laboratory, the plasmid pPL001<sup>1</sup> that contains *fom4* was used to transform methionine-auxotrophic *E. coli* strain B834(DE3) (Novagen, Madison, WI). The transformed cells were grown in LeMaster medium supplemented with 25mg/L-selenomethionine. Expression and purification of native-HppE, mutant HppE, and SeMet-HppE were performed in Dr. Ben Liu's laboratory as described previously<sup>1</sup>. Wild type and mutant HppE were reconstituted with Fe(II)(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub> as described previously<sup>1,7</sup>. Following removal of excess metal with a G-10 column, these proteins were tested for their epoxidase activity using the previously established <sup>31</sup>P-NMR assay method<sup>7</sup>.

### *II.B.2. Crystallization and data collection*

Apo-HppE crystals were grown at room temperature using the hanging drop vapor diffusion method, whereas Fe-HppE and Tris-Co(II)-HppE were grown using the sitting drop vapor diffusion method. Apo-HppE crystals were grown from protein solution (30 mg/mL (1.4 M) HppE, 20 mM Tris-hydrochloride pH 8.0) mixed in a 1:1 ratio (2  $\mu$ L each) with precipitant solution (1.9 M sodium malonate, pH 7.0) and equilibrated over 0.5 mL of precipitant solution. Colorless square bipyramidal crystals of apo-HppE grew in 24 hrs. Tris-Co(II)-HppE crystals were grown from protein solution (30 mg/ml (1.4 M) in 20 mM Tris-HCl pH 8.5) mixed in a 1:1 ratio with precipitant solution (2.0 M ammonium sulfate, 0.1 M Tris-HCl pH 8.5), and CoCl<sub>2</sub> as an additive (0.3  $\mu$ L at 100mM). Tris-Co(II)-HppE crystals grew in 24 hrs and were slightly pink relative to background.

*S*-HPP-Co(II)-HppE crystals were obtained from Tris-Co(II)-HppE crystals, by first transferring crystals from Tris to HEPES buffer to displace any Tris molecules bound to the Co(II), and then by soaking crystals in a solution containing substrate in a 12:1 enzyme to substrate ratio. In particular, Tris-Co(II)-HppE crystals were transferred to 4.7  $\mu$ L of the following solution: 2.5 M ammonium sulfate, 400 mM sodium chloride, 100 mM HEPES pH 7.5 for 10 min. (*S*)-2-Hydroxypropylphosphonic acid (0.3  $\mu$ L at 2.3 M) was added to the drop and the resulting solution was placed over a reservoir of soaking solution for an additional 14 hrs. As is the case with the Tris-Co(II)-HppE crystals, Co(II) atoms appear to be present at full occupancy as judged by comparison of B-factors for metal atoms compared to neighboring protein atoms.

*S*-HPP-Fe(II)-HppE hexagonal crystals (form 1) were also prepared from Tris-Co(II)-HppE hexagonal crystals. Under anaerobic conditions, these crystals were soaked



for four hrs in a storage solution (4.5  $\mu\text{L}$ ) containing metal chelator  $\text{Na}_2\text{EDTA}$  (2 M ammonium sulfate, 100 mM HEPES pH 7.5, 100 mM  $\text{Na}_2\text{EDTA}$ ) to remove bound cobalt ions. Crystals are then washed iteratively into the storage solution without  $\text{Na}_2\text{EDTA}$  (2 M ammonium sulfate, 100 mM HEPES pH 7.5) a total of four times in 4.5  $\mu\text{L}$  aliquots. Crystals were then transferred to a new storage solution drop (4.5  $\mu\text{L}$ ) and 100mM iron sulfate (1.0  $\mu\text{L}$ ) was subsequently added. After 4 hr, sodium (*S*)-2-hydroxypropylphosphonate was added to the drop (0.3  $\mu\text{L}$  at 2.3 M) and soaked for an additional 14 hrs. Soaking was performed over a well-solution equivalent to the storage solution. This protocol led to ~80% occupancy of Fe(II) in the crystal as judged by B-factor comparisons. The identity of iron as the metal bound in the active site following addition of iron sulfate was confirmed by an anomalous scattering experiment performed at 1.7482 Å. It is interesting to note that co-crystallization with substrate yields apo-HppE crystals, probably due to substrate-metal chelation.

Fe(II)-HppE crystals were grown in a Coy chamber under an argon/hydrogen gas atmosphere from protein solution (30 mg/ml in 20 mM Tris-HCl pH 8.5) mixed in a 1:1 ratio with precipitant solution (2.0 M ammonium sulfate, 0.1 M Tris-HCl pH 8.5), and  $\text{FeSO}_4$  as an additive (0.3  $\mu\text{L}$ ; 100mM). Crystals grew after 24 hrs. To ensure anaerobiosis, all solutions were purged with argon gas before they were transferred into the Coy chamber, and a solution of methyl viologen and sodium hydrosulfite (1.5:1.0 molar ratio) was used as a colorimetric indicator for the presence of dioxygen in the chamber. To prepare *S*-HPP-Fe(II)-HppE tetragonal crystals (form 2), Fe(II)-HppE tetragonal crystals were transferred to 4.7  $\mu\text{L}$  of the following solution: 2.5 M ammonium sulfate, 400 mM sodium chloride, 100 mM HEPES pH 7.5 for 10 min. Sodium (*S*)-2-

Hydroxypropylphosphonate acid (0.3  $\mu$ L at 2.3 M) was added to the drop and the resulting solution was placed over a reservoir of soaking solution for an additional 14 hrs. This procedure leads to full occupancy of Fe(II) at the active site as judged by the comparison of B-factors of iron compared to neighboring protein atoms.

With the exception of the Fe(II)-HppE, all crystals were transferred to a cryoprotectant solution (30% xylitol, 2.0M ammonium sulfate, and 100mM Tris-HCl pH 8.0) for 30 s and subsequently cooled under a nitrogen stream at 100K. In order to cryocool the Fe(II)-HppE and *S*-HPP-Fe(II)-HppE crystals, the above cryoprotectant was purged with argon gas and transferred to the Coy chamber. Crystals were transferred to the cryoprotectant for 30 s and plunged into liquid nitrogen while under an argon/hydrogen atmosphere. Data were collected at the synchrotron light sources listed in Table II.1. and Table II.2., and were subsequently integrated and scaled in DENZO and SCALEPACK, respectively<sup>13</sup>.

### *II.B.3. Structure determinations*

Experimental phases were obtained using data from the Tris-Co(II)-SeMet-HppE crystals. A single dataset, collected at the Se absorption peak wavelength (0.9791Å), was used to solve the structure (Table II.1.). Nine Se atoms per asymmetric unit were located and refined in CNS<sup>14</sup>. These sites were used to calculate phases in CNS with a mean figure of merit of 0.548 to 2.5 Å resolution. Following density modification by solvent flipping in CNS, RESOLVE<sup>14,15</sup> was used to build residues 7-73 of one  $\alpha$ -domain and 103-198 of one  $\beta$ -domain. The remaining residues were built manually into the density modified experimental map using XFIT (Duncan E. McRee, The Scripps Research Institute), and

this initial model was refined as a rigid body in CNS<sup>14</sup>. Iterative rounds of model building in XFIT and refinement in CNS followed at 2.5 Å resolution. The CNS protocol included simulated annealing against a MLF target and *B*-factor refinement; non-crystallographic symmetry restraints and sigma cutoff were not used. Topology files for Tris were obtained using HIC-UP<sup>15</sup>. The final model was refined against a 1.8 Å resolution native Tris-Co(II)-HppE dataset (Table II.1.). This model contains residues 5-87, 102-158, 160-198 in molecule A and 6-87, 102-198 molecule B (198 residues per molecule). There is no electron density for residues 88-101.

The refined Tris-Co(II)-HppE structure, including all modeled protein atoms, was used as a molecular replacement model to solve both the Fe(II)-HppE and the apo-HppE structure. The former was solved using AMORE<sup>16</sup> (correlation coefficient of 0.58; R-factor of 0.48) and the latter was solved using EPMR<sup>17</sup> (correlation coefficient of 0.54; R-factor of 0.46), in both cases using data from 15-4.0 Å resolution. Rigid body refinement was subsequently performed in CNS. The models were adjusted manually in XFIT and refined in CNS, using the same protocol as described above, at the highest available resolution (Table II.1.). The final Fe(II)-HppE model contains 6-159, 161-198 in molecule A and 6-198 in molecule B, and 6-96, 100-198 in molecule C. The final apo-HppE model contains residues 6-95, 101-136, 140-157, 162-198 in molecule A and 6-97, 101-136, 141-158, 162-198 in molecule B (198 residues per molecule). The electron density was not interpretable for the remaining residues. The apo-HppE data were highly anisotropic, and therefore not of as good a quality as the other data sets that were used to refine HppE structures. As a result, the apo-HppE structure, with R-factors of 27.8% and

32.4% ( $R_{\text{free}}$ ), is not as accurate as the other structures, but is useful in terms of rough comparison with the metal-bound structures.

As *S*-HPP-Co(II)-HppE, form-1 *S*-HPP-Fe(II)-HppE, and Tris-Co(II)-HppE crystals are isomorphous (Table II.1-2.), the *S*-HPP-Co(II)-HppE, and *S*-HPP-Fe(II)-HppE structures were solved by rigid body refinement in CNS using only protein atoms from the refined Tris-Co(II)-HppE structure as an initial model. The  $9\sigma$  electron density peaks in the Co(II) structure and  $6\sigma$  peaks in the Fe(II) structure readily identified the location of the phosphonic moiety of the substrate, and allowed for the placement of the substrate in each model, respectively. Topology and parameter files for the substrate were obtained using HIC-UP<sup>15</sup>. Manual adjustment of the model was done using XFIT and refinement was performed in CNS as described above. The final *S*-HPP-Co(II)-HppE model contains residues 6-90, 101-187, 190-198 in molecule A and 6-198 in molecule B (198 residues per molecule); there is no electron density for remaining residues of molecule A. The final *S*-HPP-Fe(II)-HppE hexagonal model contains residues 7-87, 102-198 in molecule A and 7-87, 103-156, 160-198 in molecule B. The electron density for omitted residues was not sufficient for model building.

Since the form-2-*S*-HPP-Fe(II)-HppE and Fe(II)-HppE tetragonal crystals are isomorphous (Table II.2.), the *S*-HPP-Fe(II)-HppE tetragonal structure was solved by rigid body refinement using only protein atoms from the refined Fe(II)-HppE structure as an initial model. The substrate orientation was defined by  $10\sigma$  electron density peaks corresponding to the phosphorus as in the hexagonal substrate structures. Refinement was carried out as described above. The final *S*-HPP-Fe(II)-HppE tetragonal model contains residues 7-198 in molecule A and 6-198 molecule B and C.

Simulated annealing composite omit maps were used to validate all HppE structures. For the six structures presented, all residues are in the allowed regions of the Ramachandran plot as calculated by PROCHECK<sup>18</sup>: Fe(II)-HppE: 87.0% in the most favored regions, 11.7% in additionally allowed regions, and 1.3% in generously allowed regions; apo-HppE: 90.9% in the most favored regions and 9.1% in additionally allowed regions, and 0.3% in generously allowed regions; Tris-Co(II)-HppE: 91.2% in the most favored regions, 9.5% in additionally allowed regions, and 0.3% in generously allowed regions; S-HPP-Co(II)-HppE: 89.7% in the most favored regions, 9.6% in additionally allowed regions, and 0.6% in generously allowed regions; S-HPP-Fe(II)-HppE form 1: 90.7% in the most favored regions, 8.9% in additionally allowed regions, and 0.3% in generously allowed regions; S-HPP-Fe(II)-HppE form 2: 88.6% in the most favored regions, 10.8% in additionally allowed regions, and 0.6% in generously allowed regions; S-HPP-Co(II)-HppE: 89.7% in the most favored regions, 9.6% in additionally allowed regions, and 0.6% in generously allowed regions.

## **II.C. Results**

### *II.C.1. Overall structure of HPP epoxidase*

The three-dimensional structure of HppE (Table II.1-2.) confirms its identity as a member of the cupin superfamily, possessing the requisite  $\beta$ -barrel fold in which antiparallel  $\beta$ -stands are wrapped around a barrel core in a jelly roll variant of a Greek Key or  $\beta$ -sandwich motif (Figure II.2.a). As expected, the epoxidase is a physiological homotetramer with one iron per monomer<sup>1</sup>. Each monomer consists of two domains: the  $\alpha$ -domain, which is all  $\alpha$ -helical, and a  $\beta$ -domain, which consists of all anti-parallel  $\beta$ -

strands in a jellyroll  $\beta$ -barrel motif (Figure II.2.a.). As predicted by sequence alignments<sup>19</sup>, the facial triad, His138, Glu142, and His180, is housed within the  $\beta$ -barrel, defining the HppE active site. The spacing of the first two triad members (His138 & Glu142) is different in HppE and germin oxalate oxidase (HxxxE)<sup>20</sup> relative to the majority of other proteins within the superfamily (HxD/E), an important consideration for bioinformatics studies. The smaller  $\alpha$ -domain has only 47 residues, consisting of five helices, and represents a unique helical domain fold (Figure II.2.a.), according to a Dali server data base search. An extended thirty-two-residue linker separates the  $\alpha$ - and  $\beta$ -domains of the HppE such that the closest distance between the two domains is approximately 15 Å (Figure II.2.a.). This separation translates to a striking quaternary structure that is unique in the cupin superfamily (Figure II.2.b. and Figure II.2.c.)<sup>21</sup>.

### *II.C.2. The HppE active site*

Structural comparison of apo-HppE with Fe(II)-HppE shows that regions around the metal binding site in the  $\beta$ -domain, which are disordered in the apo-structure, become ordered upon metal binding. The ordering of strand 5, and the loop between strands 5 and 6, bring triad ligand His138 into proximity of the iron (Figure II.3.a. and Figure II.3.b.). In one of three molecules in the asymmetric unit of the Fe(II)-HppE structure, two waters occupy exchangeable sites yielding a five-coordinate metal center (Figure II.3.b. and Figure II.4.b.). In the other two molecules, the electron density is less well defined but is consistent with direct coordination of three water molecules to Fe(II) (data not shown).

In the mononuclear iron proteins, residues responsible for substrate specificity, regioselectivity, and stereoselectivity are found above the ligand triad on the other end of the  $\beta$ -barrel<sup>22</sup>. Based upon the overall HppE structure, residues on the “tunable” face of the  $\beta$ -domain that are potentially significant for catalysis include: Tyr105, Tyr103, Tyr102, and Arg97. In HppE, both  $\beta$  and  $\alpha$  domains contribute to the active site. Helices 1 and 2 of the  $\alpha$ -domain comprise one side of the active site (Figure II.5.b.), Lys23 from the  $\alpha$ -domain is positioned near the metal center (Figure II.3.). Interestingly, HppE oxidizes Tyr105 to dihydroxyphenylalanine (DOPA) via a self-hydroxylation reaction<sup>23</sup>, correctly predicting that Tyr105 lies near the iron site (Figure II.3.b.). In a Tyr105Phe mutant protein, Tyr103 is oxidized to DOPA<sup>23</sup>, consistent with Tyr103 being close to the iron site, but not as close as Tyr105, since it is only modified in the mutant protein (Figure II.3.b.). Although HppE is not a monooxygenase with respect to the biosynthesis of fosfomicin, it does catalyze these self-hydroxylation reactions in the absence of substrate, a phenomenon observed in other family members such as taurine/ $\alpha$ -KG dioxygenase (TauD)<sup>24</sup>. Although self-hydroxylation is off-pathway, it is interesting in terms of the reactivity and the plasticity of the HppE active site. In the structure of Fe(II)-HppE, Tyr105 is within 8.7 Å and Tyr103 is within 10.8 Å, indicating that regions of the active site must be conformationally flexible to allow these tyrosine residues to approach the iron and hydroxylation can occur. Results presented below confirm this conformational flexibility.

### *II.C.3. Substrate Binding and Conformational changes*

To investigate substrate binding without turnover, Co(II)-HppE was used as a catalytically inert model system<sup>7</sup>, and anaerobic crystals of Fe(II)-HppE were obtained in two different crystal forms (form-1, P6<sub>5</sub>22 and form-2, P4<sub>2</sub>2<sub>1</sub>2) and soaked with substrate (Table II.1-2.). The presence of a moderately heavy atom in the substrate (i.e. phosphorus) gave rise to 6-9 $\sigma$  omit electron density peaks indicating the position of the phosphonic acid moiety and the orientation of the substrate in these structures. Surprisingly, these structures reveal two modes of substrate binding, suggesting a two-step binding process prior to catalysis. In one of the two molecules present in the asymmetric unit of Co(II)-HppE and form-1 Fe(II)-HppE, substrate binds in a monodentate fashion with an oxygen atom of the phosphonic acid moiety coordinating the metal and displacing one bound water (Figure II.3.c,f. and Figure II.4.d-e.). In the other molecule of Co(II)-HppE and form-1 Fe(II)-HppE, as well as in form-2 Fe(II)-HppE, S-HPP forms a bidentate interaction with the metal center via the 2-hydroxyl oxygen and an oxygen atom from the phosphonic acid moiety, substituting for water ligands in the free form, and leaving a third coordination site open for subsequent dioxygen binding (Figure II.3.d,g. and Figure II.4.a,f. and Figure II.5.a.). In both Fe(II) and Co(II) crystal forms, bidentate binding of S-HPP induces a conformational change in the  $\beta$ -barrel, while substrate bound in a monodentate fashion physically precludes this conformational change from occurring. The conformational change involves  $\beta$ -strands 2 and 3 (Figure II.2.), which form a  $\beta$ -hairpin substructure and act as a cantilever responding to the positioning of substrate in the bidentate orientation by moving in and covering one side of the active site (Figure II.5.b.). Movements near the hinge, or



fulcrum, of the hairpin are subtle while those near the end are more appreciable (Figure II.5.b.). Specifically, Tyr105 moves toward the metal center by 1.8 Å, positioning its hydroxyl group within hydrogen bonding distance of a phosphonate oxygen of *S*-HPP (2.7 Å), while Tyr102 moves 8.2 Å toward the metal, occupying approximately the same position as Tyr103 in the Fe(II)-HppE structure (Figure II.3.b. and Figure II.3.d,g.). In addition, Arg97, on the N-terminus of strand 2, swings toward the metal center through a distance of approximately 7.4 Å (Figure II.5.b.). There is little change in other proximal residues of the β-barrel. Relative to the cantilever fulcrum defined by residues Ile92 and Asn106, the β-hairpin swings through an angle of 29 degrees, decreasing the solvent exposed surface area of the β-barrel. In addition to the linear motion toward the active site, the right-handed twist (Figure II.5.b.) of the hairpin increases upon substrate binding, conserving the hydrogen bonding interactions of the β-sheet. Alignment of the two substrate binding modes reveals that the monodentate binding position of *S*-HPP sterically precludes hairpin motion toward the metal center, in that it occupies the position that Arg97 holds in the closed conformation (Figure II.5.a.).

The movement of the cantilever does not entirely seal off the active site. Dioxygen can still access the iron, and the negatively charged phosphonic acid moiety of *S*-HPP is still exposed to solvent (Figure II.5.c.). In contrast, the hydrophobic portion of *S*-HPP, where hydrogen atom abstraction occurs, is completely buried in a small hydrophobic cavity. The conformational change also creates an extensive hydrogen bonding network between the substrate, bound water molecules, and the enzyme (Figure II.4.a,f.). Mutation of two of the residues involved in the hydrogen bonding network, Tyr105Phe and Lys23Ala, leads to complete loss of activity as determined by a <sup>31</sup>P-NMR

assay, under conditions where wild-type HppE has activity and ligand triad mutant Glu142Ala does not. The movement of the cantilever also positions three Tyr residues (102, 103, 105) in a row, leading from the active site to the surface of the protein (Figure II.3.d,g.). This result suggests that these Tyr residues could play a role in electron transfer between the putative reductase and the iron in the active site.

The structure of *tris*-(hydroxymethyl) aminomethane (Tris) in complex with Co(II)-HppE provides insight into substrate specificity and active site closure (Figure II.3.e. and Figure II.4.c.). The Tris molecule is best fit to the electron density when modeled as binding to the Co(II) metal center via two hydroxymethyl oxygen atoms and the nitrogen atom of the aminomethane moiety (Figure II.3.e. and Figure II.4.c.). This coordination geometry places the remaining hydroxymethyl substituent within hydrogen bonding distance to Lys23 (2.9 Å). With Tris in the active site, one strand of the hairpin cantilever is in the substrate-free position whereas the other strand is disordered. Given that the physiological substrate of HppE is smaller than Tris, is charged instead of neutral, and has additional hydrogen bond acceptors, it is not surprising that the hairpin cantilever does not close in as it does when bidentate substrate is bound. Thus, the cantilever hairpin is only positioned for catalysis when a molecule of the correct size is bound in the appropriate conformation.

## **II.D. Discussion**

### *II.D.1. Mechanism of HppE*

In the past ten years there has been a series of structural studies of mononuclear iron proteins, including the work on enzymes involved in antibiotic biosynthesis  $\alpha$ -KG

independent IPNS<sup>12,25</sup> and  $\alpha$ -KG dependent Deacetoxycephalosporin-C synthase (DAOCS)<sup>26</sup>. The structures of HppE described here, along with the previous biochemical studies<sup>1,7,11,23</sup>, allow us to compare HppE with these other enzymes to probe the fundamental properties of this enzyme subfamily. There are a series of mechanistic challenges that mononuclear iron proteins in the cupin family must face: fine-tuning the protein structure to bind substrate, co-substrate, and/or cofactor; activation of iron for dioxygen binding; acquisition of reducing equivalents; protection of radical intermediates during catalysis. With this structural information, we can consider how HppE accomplishes the above, and why it differs from so many other family members in not using  $\alpha$ -KG as a co-substrate.

The mechanism of substrate binding in HppE has been investigated using biochemical and biophysical techniques<sup>7</sup>. As shown by electron paramagnetic resonance (EPR) spectroscopy, the metal environment is heterogeneous when substrate binds to Fe(III)-HppE. This heterogeneity coalesces into a single homogeneous species in the EPR spectrum of Fe(II)-HppE bound with substrate and nitric oxide, a catalytically inert surrogate for dioxygen<sup>7</sup>. These EPR data indicate that substrate binds directly to iron, and are consistent with multiple substrate binding modes. In the crystal structures, we find that substrate binds directly to the metal, and that there are two binding modes, monodentate and bidentate. Since there are no lattice contacts to explain the multiple substrate conformations and since the two conformations are found with both Co(II) and Fe(II) reconstituted enzymes, we propose that both binding modes are relevant, and that *S*-HPP binds in a two-step process. In this proposed two-step process, *S*-HPP first binds in a monodentate fashion via an oxygen atom of the phosphonic acid substituent such that

a water molecule is displaced. Monodentate binding is followed by rotation about the C(1)-P-bond and the C(1)-C(2) bond of the substrate, allowing for bidentate coordination of *S*-HPP to the metal center concomitant with the displacement of remaining water molecules (Figure II.5.b. and Figure II.6.). With the substrate in a bidentate configuration, the charge and spatial complementarity are such that they afford the 8 Å shift of the cantilever hairpin, covering the hydrophobic portion of the substrate bound in the active site.

Following substrate binding, the next steps in the mechanism are iron activation and dioxygen binding. Here, the negative charge of substrate could serve to enhance the affinity of dioxygen for the iron center by favoring higher oxidation states, a common mechanistic theme for this family of proteins<sup>9,26-29</sup>. It is interesting that HppE has evolved to recognize this unusual C-P bond containing natural product, and to utilize its physical properties to activate iron for dioxygen binding. The active site iron is activated for dioxygen binding with substrate configured in either monodentate or bidentate modes. In either conformation of the cantilever, dioxygen can access the iron through a very small channel created at the interface of the  $\alpha$  and  $\beta$  domains (Figure II.5.c.). Interestingly, this small channel leads to the only open coordination site on the metal in the bidentate *S*-HPP complex. This coordination site is where we have modeled dioxygen (Figure II.5.d.).

After formation of a ternary complex with substrate and dioxygen, C-1 hydrogen atom abstraction by an activated Fe-oxo species would yield a transient substrate radical intermediate that undergoes cyclization to yield fosfomycin<sup>11</sup>. There are several plausible mechanisms for active site reduction and epoxide ring formation, including hydrogen

atom abstraction via an Fe(IV)-oxo intermediate (mechanism A) and hydrogen atom abstraction by an Fe(III)-hydroperoxide intermediate (mechanism B) (Figure II.6.). The four electron oxidation of ACV catalyzed by IPNS employs both direct hydrogen atom abstraction by an Fe-superoxo species, a variation of mechanism B, and substrate radical formation via an Fe(IV)-oxo species, as in mechanism A<sup>12,30</sup>. The hydrogen atom abstraction is both stereo- and regiospecific<sup>11</sup>, and the dioxygen-binding model in Figure II.4.d. can explain the observed stereo- and regiospecificity. An oxygen species bound in the open coordination site could only access the relevant hydrogen on C-1, the pro-*R* hydrogen. The C-2 hydrogen points away from the oxygen-binding site, and is protected from abstraction by the C-2 methyl substituent.

If the model in Figure II.4.d. is correct, then Lys23 from the neighboring molecule should play a role in catalysis since it is positioned directly above the proposed dioxygen binding site. In fact, replacement of Lys23 with Ala results in a protein with no detectable activity, supporting this model. The exact role that Lys23 is playing is not clear. In both mechanisms A and B, the negatively charged Fe-peroxide intermediate would need to be stabilized and subsequently protonated. Lys23 could be involved in both processes.

Like other mononuclear iron enzymes, HppE and IPNS must shield reactive intermediates from solvent. Both enzymes have evolved different strategies to meet this need. In IPNS, ACV binding facilitates ligand exchange, triggering the extension of the C-terminal helix by five residues which seals the  $\beta$ -barrel<sup>12</sup>. In HppE, the cantilever hairpin moves toward the metal center concomitant with bidentate substrate binding, serving to partly occlude the  $\beta$ -barrel (Figure II.5.b.). Comparison of Figures II.3.c-g.

shows that the structure of HppE is finely tuned to respond to the correct substrate bound in an appropriate orientation (bidentate binding) before adopting this closed catalytic conformation.

### *II.D.2. Structural and mechanistic comparisons*

Structural comparisons of HppE, IPNS, and DAOCS, suggest why HppE is an  $\alpha$ -KG independent enzyme in contrast to the majority of cupin mononuclear iron proteins. First, one must consider that  $\alpha$ -KG can serve two functions: it can serve as a source of two electrons for dioxygen reduction, and it can activate iron for dioxygen binding. With respect to the latter function, the bidentate binding of  $\alpha$ -KG to iron via its negatively charged C-1 carboxylate and the 2-oxo substituent, primes the iron for dioxygen binding, and subsequent iron oxidation<sup>31</sup>. This function is not necessary in HppE since the substrate is negatively charged and can independently activate iron for dioxygen binding. Similarly, IPNS, another  $\alpha$ -KG independent enzyme within this subfamily, facilitates dioxygen binding via coordination of a negatively charged thiolate substituent of substrate (ACV)<sup>12</sup>. A second consideration is that bidentate substrate binding in HppE fills two of three open coordination sites on iron, preventing simultaneous bidentate binding of  $\alpha$ -KG. Most characterized  $\alpha$ -KG dependent iron enzymes simultaneously bind  $\alpha$ -KG and substrate in their catalytic positions, so that a highly reactive Fe-oxygen species is not formed until substrate is available<sup>9</sup>. A notable exception, DAOCS, binds  $\alpha$ -KG and substrate penicillin to the same coordination site on the iron<sup>26</sup>. In this case, DAOCS has evolved a mechanism by which the Fe-oxygen species formed by reaction with  $\alpha$ -KG is "booby-trapped" (i.e. reactivity is contained) until the trap is sprung by

substrate binding<sup>26</sup>. While HppE could have evolved a similar protection mechanism to allow ping-pong binding of  $\alpha$ -KG and substrate to the iron, this enzyme does not require  $\alpha$ -KG for activation, and thus a better solution appears to be the use of the putative HppE reductase. While all cupin mononuclear iron proteins have to activate iron for dioxygen binding and obtain reducing equivalents for the conversion of dioxygen to water, the strategies employed are different. When coupled to substrate oxidation,  $\alpha$ -KG can fulfill both requirements. In IPNS, substrate alone can fulfill both roles. In HppE, however, substrate facilitates dioxygen binding, but cannot provide all of the electrons, and thus this enzyme requires an endogenous reductase. In the latter respect, HppE deviates from other cupin mononuclear iron enzymes.

## **II.E. Conclusion**

This chapter presents the first structures of HppE, revealing a new subfamily of cofactor-independent mononuclear iron proteins. Structural analysis provides insight into the mechanism of substrate binding and catalytic initiation *en route* to the biosynthesis of fosfomycin. The growing threat of antibiotic resistant microorganisms accentuates the importance of understanding the chemical mechanisms governing antibiotic biosynthesis. This knowledge can be applied to protein engineering of mononuclear iron scaffolds and the biosynthesis of unique antibiotic analogs.

## **I.F. Acknowledgements**

Feng Yang performed all the protein purification, mutagenesis experiments, and enzyme assays. Pinghua Liu developed the purification protocol. Dr. Ben Liu is a collaborator on this project.



## II.G. Figures and Tables

**Table II.1. Data and Refinement Statistics for Co and Apo HppE Structures**

	<b>SeMet- Tris-Co(II)<sup>b</sup></b>	<b>Native- Tris-Co(II)</b>	<b>S-HPP-Co(II)</b>	<b>Apo</b>
<b>Data Collection</b>				
Space Group	P6 <sub>5</sub> 22	P6 <sub>5</sub> 22	P6 <sub>5</sub> 22	P4 <sub>3</sub> 2 <sub>1</sub> 2
Cell Dimensions				
<i>a</i> , <i>b</i> (Å)	86.63	86.40	86.23	89.10
<i>c</i> (Å)	220.78	220.25	220.22	163.62
Beamline	NSLS-X26C	APS-8BM	APS-8BM	APS-8BM
Wavelength (Å)	0.9791	0.9791	0.9791	0.9791
Resolution (Å)	50-2.5	50-1.80	50-2.30	50-2.0
R <sub>sym</sub> (%) <sup>a,b</sup>	10.7 (25.2)	4.8 (36.6)	8.6 (36.0)	7.8 (29.3)
I/σ(I) <sup>a</sup>	66.9 (9.1)	35.0 (3.9)	26.4 (6.4)	19.3 (3.8)
Total Observations	577887	312589	381588	211784
Unique Reflections	17792	45155	22149	40511
Completeness (%) <sup>a</sup>	99.8 (98.8)	98.2 (99.9)	99.3 (100.0)	90.0 (91.7)
Redundancy	32.5	6.9	17.2	5.2
<b>Refinement</b>				
Resolution (Å)		22-1.80	43-2.30	49.9-2.0
No. reflections		43377	21796	39099
R <sub>work</sub> <sup>d</sup> (R <sub>free</sub> ) <sup>e</sup> (%)		22.5 (24.2)	23.8 (28.1)	27.8 (32.4)
No. Molecules per asymmetric unit <sup>c</sup>		2	2	2
No. atoms				
Protein		2620	2803	2709
Iron		0	0	0
Cobalt		2	2	0
S-HPP		0	16	0
Sulfate		0	0	0
Tris		16	0	0
Water		295	73	177
Average B-Factors (Å <sup>2</sup> )				
Protein atoms		35.7	48.8	37.7
ligands		43.8	56.3	-
metal		33.7	43.4	-
water		47.5	48.4	43.6
R.m.s deviations				
Bond lengths (Å)		0.01	0.007	0.006
Bond angles (°)		1.7	1.5	1.4

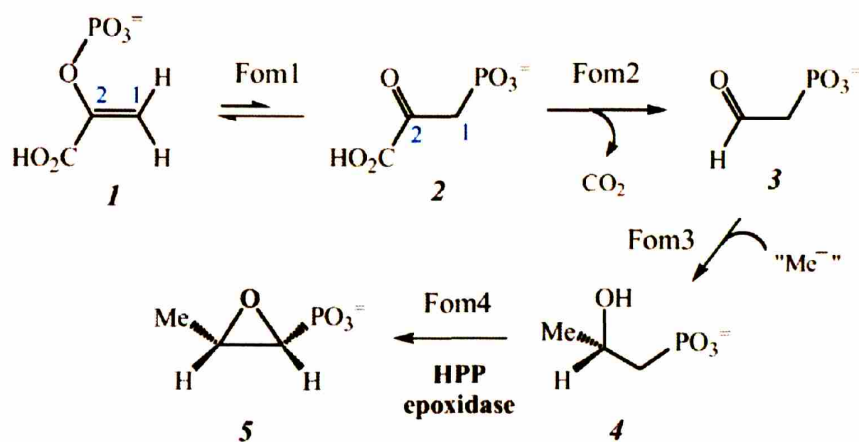
**Table II.2. Data and Refinement Statistics for Fe HppE Structures**

<b>Data Collection</b>	<b>Fe(II)</b>	<b>S-HPP-Fe(II) Form 1</b>	<b>S-HPP-Fe(II) Form 2</b>
Space Group	P4 <sub>2</sub> 2 <sub>1</sub> 2	P6 <sub>5</sub> 22	P4 <sub>2</sub> 2 <sub>1</sub> 2
Unit Cell Dimensions			
<i>a</i> , <i>b</i> (Å)	111.87	86.56	111.65
<i>c</i> (Å)	153.16	220.36	152.07
Beamline	APS-8BM	APS-8BM	ALS 5.0.2
Wavelength (Å)	0.9791	1.7384	1.1271
Resolution Range (Å)	30-2.4	30-2.1	30-2.3
R <sub>sym</sub> (%) <sup>a,b</sup>	8.2 (35.8)	8.3 (36.2)	7.0 (21.8)
I/σ(I) <sup>a</sup>	16.7 (3.0)	19.1 (3.2)	20.8 (6.8)
Total Observations	186672	198326	265410
Unique Reflections	36510	28639	42988
Completeness (%) <sup>a</sup>	94.6 (94.9)	97.4 (93.1)	99.2 (99.3)
Redundancy	5.1	6.9	6.2
<b>Refinement</b>			
Resolution Range (Å)	15-2.4	29-2.1	30-2.3
No. reflections	34838	28115	41747
R <sub>work</sub> <sup>d</sup> (R <sub>free</sub> ) <sup>e</sup> (%)	24.9 (28.7)	24.4 (26.6)	22.5 (25.2)
No. Molecules per asymmetric unit	3	2	3
No. atoms			
Protein	4294	2590	4341
Iron	3	2	3
Cobalt	0	0	0
S-HPP	0	16	24
Sulfate	15	0	0
Water	127	145	146
Average B-Factors (Å <sup>2</sup> )			
Protein atoms	41.4	34.3	36.3
ligands	-	63.4	32.8
metals	42.5	46.2	30.0
water	39.8	38.3	36.7
R.m.s deviations			
Bond lengths (Å)	0.009	0.009	0.010
Bond angles (°)	1.5	1.6	1.6

<sup>a</sup>Values in parentheses are the highest-resolution shell. <sup>b</sup> $R_{sym} = \sum_{hkl} |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \langle I(hkl) \rangle$ , where  $I_i(hkl)$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I(hkl) \rangle$  is the mean of the intensity for the miller index (hkl). <sup>c</sup>ASU, asymmetric unit. <sup>d</sup> $R_{factor} = \sum_{hkl} |F_o(hkl) - F_c(hkl)| / \sum_{hkl} |F_o(hkl)|$ . <sup>e</sup> $R_{free} = R_{factor}$  for a test set of reflections (5%). <sup>f</sup>Statistics are for data set in which Friedel pairs are not merged.

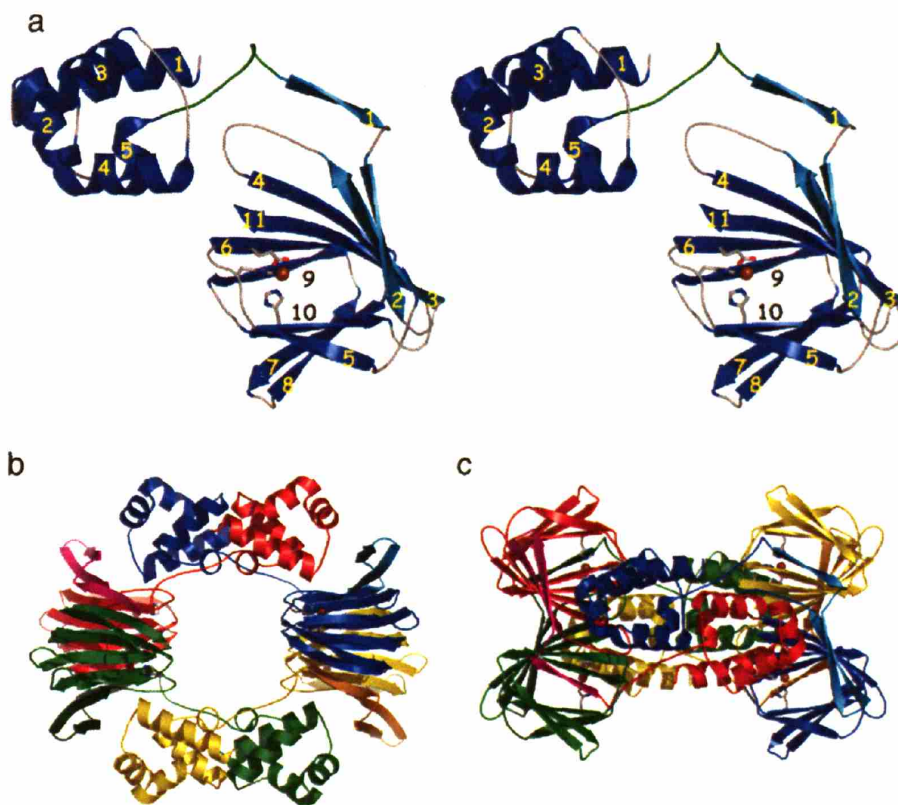
**Figure II.1. Fosfomycin biosynthetic pathway.**

The fosfomycin biosynthetic pathway requires phosphoenolpyruvate mutase (Fom1), phosphonopyruvate decarboxylase (Fom2), phosphonoacetaldehyde methyltransferase (Fom3) and HppE (Fom4). 1, phosphoenolpyruvate (PEP); 2, phosphonopyruvate (PnPy); 3, phosphonoacetaldehyde (PnAA); 4, (*S*)-2-hydroxypropylphosphonic acid (*S*-HPP); 5, fosfomycin. C1 and C2 positions are shown in blue.



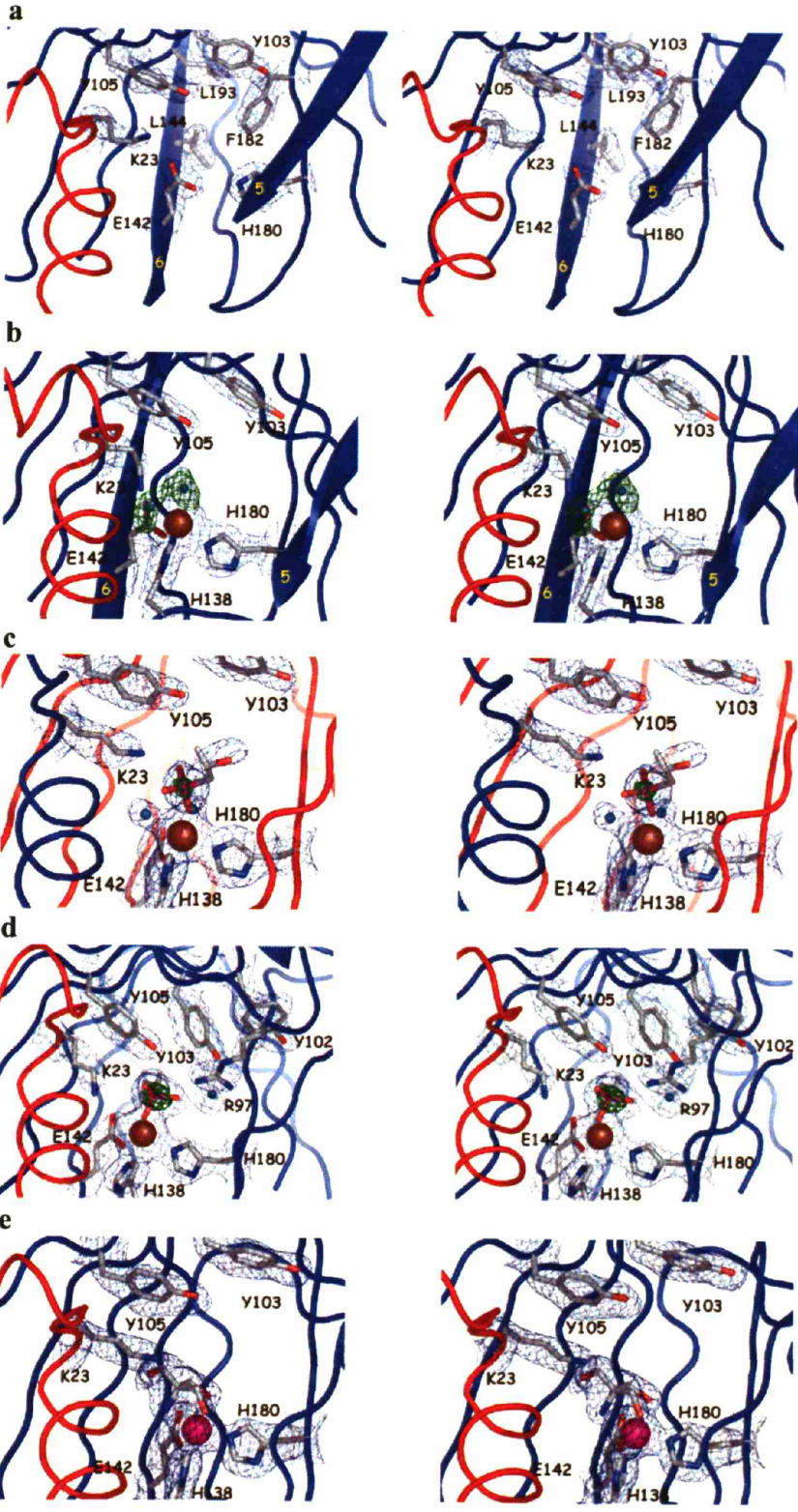
**Figure II.2. The overall structure of HppE.**

(a) An HppE monomer consists of an  $\alpha$ -domain (blue), an interdomain linker (green) to a single  $\beta$ -strand 1 (cyan) and a  $\beta$ -domain (blue). This stereoview highlights the cantilever hairpin ( $\beta$ -strands 2 and 3) in cyan, facial triad ligands (Glu 142, His 138 and His 180) in ball-and-stick, and iron as a brown sphere. Helices and strands are numbered separately and sequentially with respect to the primary structure. (b) HppE tetramer, coloured by molecule, is shown down one of the three two-fold axes of symmetry. The cantilever hairpin is coloured cyan in the blue molecule, magenta in the red molecule, dark yellow in the yellow molecule, and dark green in the green molecule. (c) HppE tetramer oriented along a second two-fold axis of symmetry, orthogonal to that in b. The cantilever hairpins are coloured as in b. Figs 2–4 were made in PyMol<sup>32</sup>.

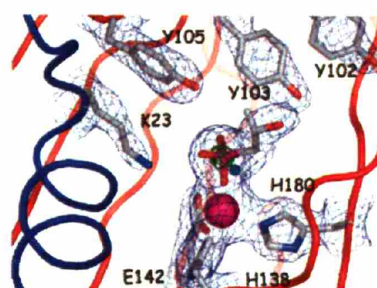
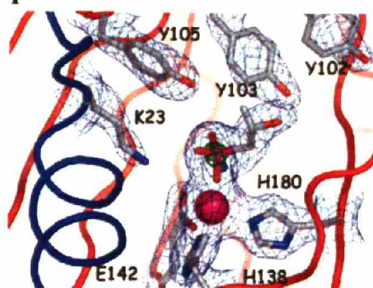


### Figure II.3. The HppE Active Site.

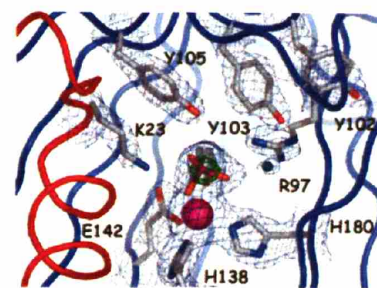
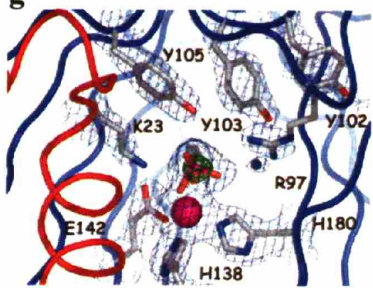
Stereoviews of HppE active sites displayed with  $2Fo-Fc$  maps contoured from 1.0-1.5 $\sigma$ . **(a)** Apo-HppE active site with strands 5 and 6 (yellow labels),  $\alpha$ -domain (red),  $\beta$ -domain (blue), and carbon (gray), oxygen (red), nitrogen (blue). **(b)** Fe(II)-HppE active site.  $Fo-Fc$  electron density peaks contoured at 3 $\sigma$  (green) indicate the position of bound water molecules (cyan) to Fe(II) (rust). **(c)** Monodentate substrate binding in form-1 *S*-HPP-Fe(II)-HppE active site colored as above. An omit map contoured at 6 $\sigma$  shows an electron density peak (green) coincident with the phosphorus atom. **(d)** Bidentate substrate binding in a form-2 *S*-HPP-Fe(II)-HppE active site with domains and atoms colored as above. **(e)** The Tris-Co(II)-HppE active site colored as above with cobalt (magenta sphere). **(f)** Monodentate substrate binding in an *S*-HPP-Co(II)-HppE active site with  $\alpha$ -domain (blue),  $\beta$ -domain (red), and carbon (gray), oxygen (red), nitrogen (blue), and cobalt in magenta sphere and phosphorus in magenta ball-and-stick. An omit map contoured at 9 $\sigma$  shows an electron density peak (green) coincident with the phosphorus atom here. The HppE  $\alpha$ -domain is colored in blue and the  $\beta$ -domain in red to emphasize that this is a different active site in the asymmetric unit than the one depicted in (g). **(g)** Bidentate substrate binding in an *S*-HPP-Co(II)-HppE active site with  $\alpha$ -domain (red),  $\beta$ -domain (blue), and carbon (gray), oxygen (red), nitrogen (blue), and cobalt in magenta sphere and phosphorus in magenta ball-and-stick.



f



g



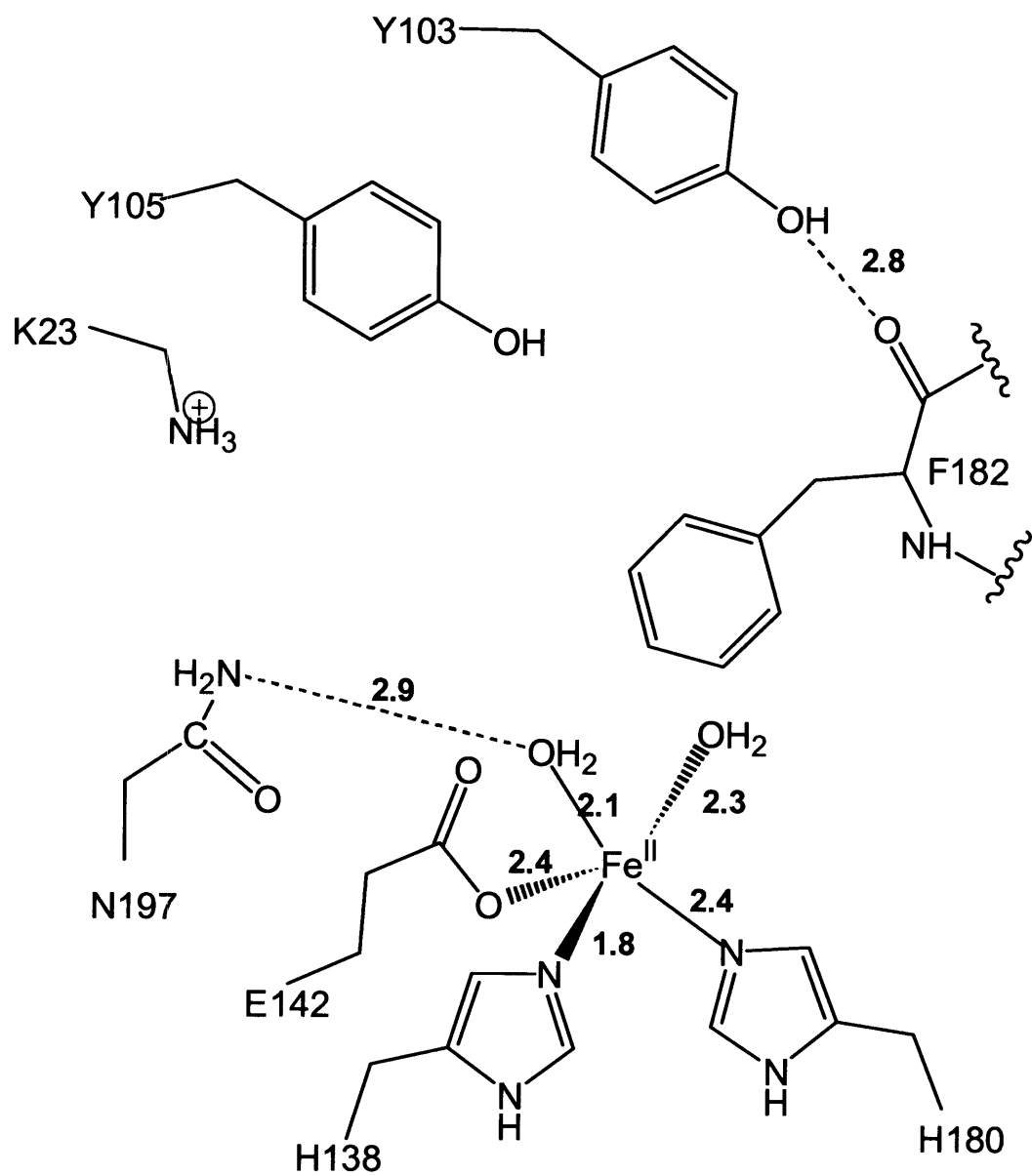
**Figure II.4. Schematic diagrams of protein, metal, and ligand interactions.**

**(a)** Residues within the substrate hydrophobic cavity are labelled (semicircles) and include Phe 182. Hydrogen-bonding contacts (dotted lines) are shown with distances indicated in Å. For simplicity, the interaction of the Arg 97 guanidinium with the phosphonic acid moiety (~ 3.4–3.6 Å) is not shown. Water shown as a grey sphere. **(b)** Fe(II)-HppE **(c)** Tris-Co(II)-HppE **(d)** *S*-HPP-Fe(II)-HppE Form 1 **(e)** *S*-HPP-Co(II)-HppE monodentate coordination **(f)** *S*-HPP-Co(II)-HppE bidentate coordination.

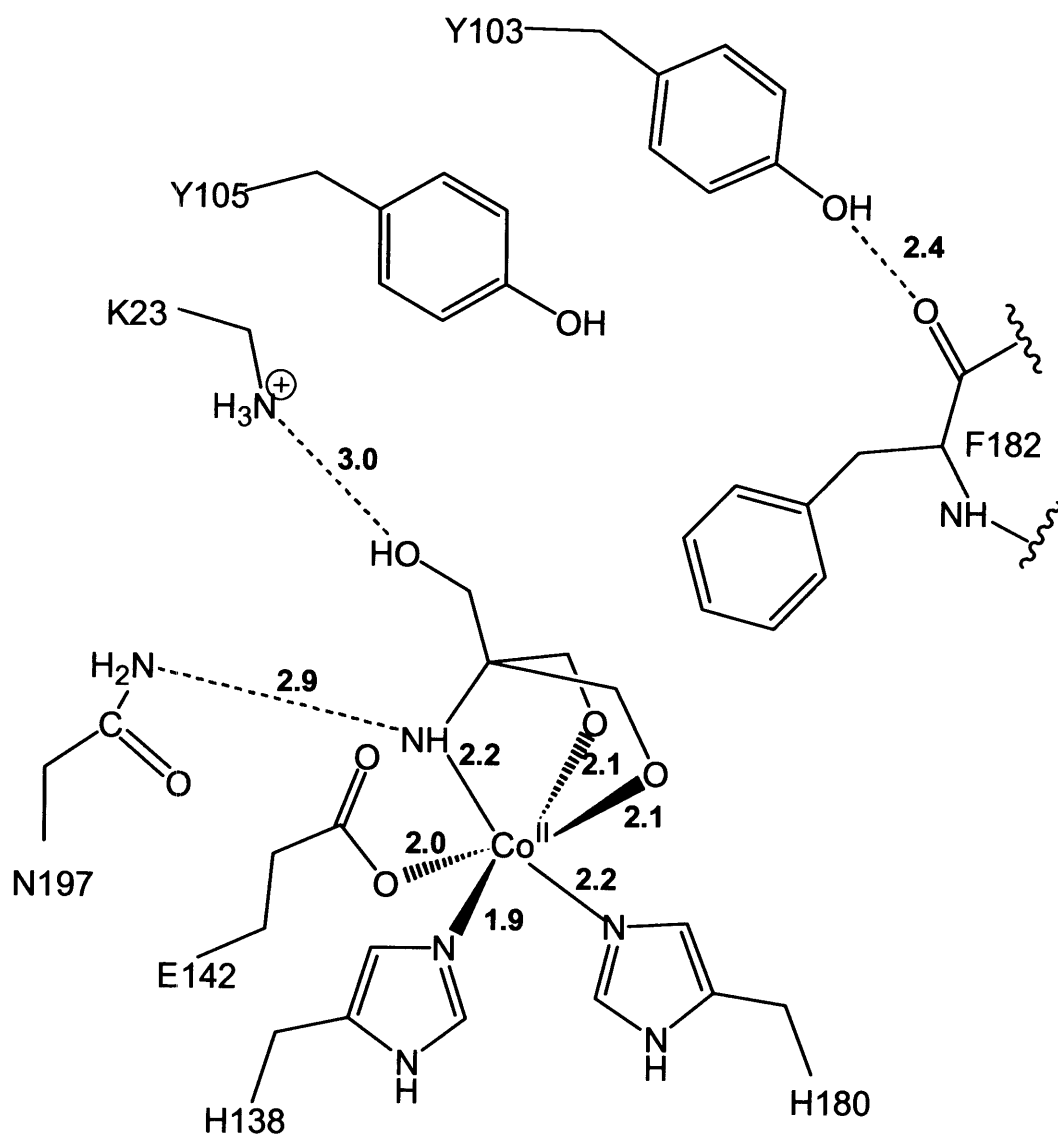




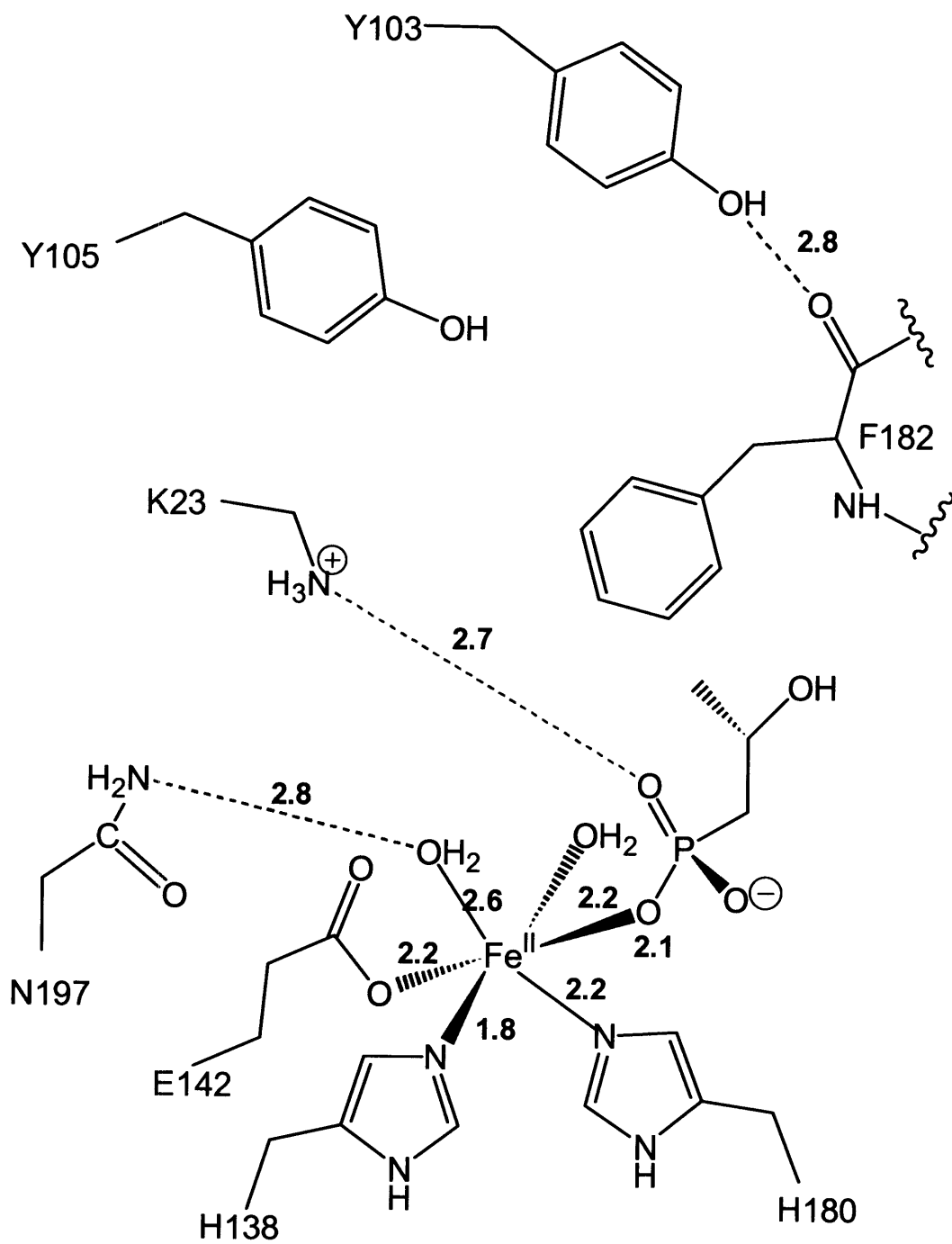
(b)



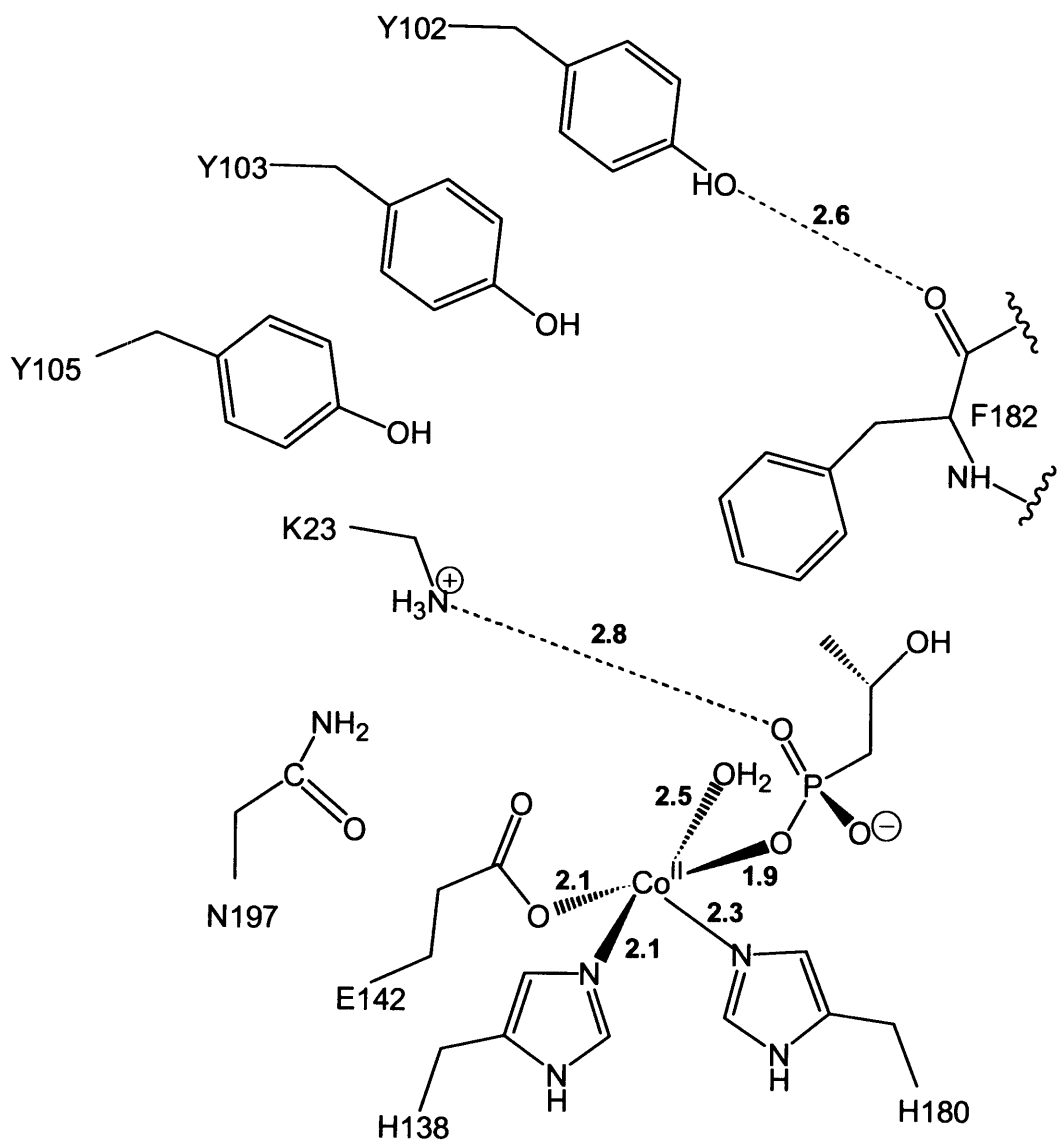
(c)



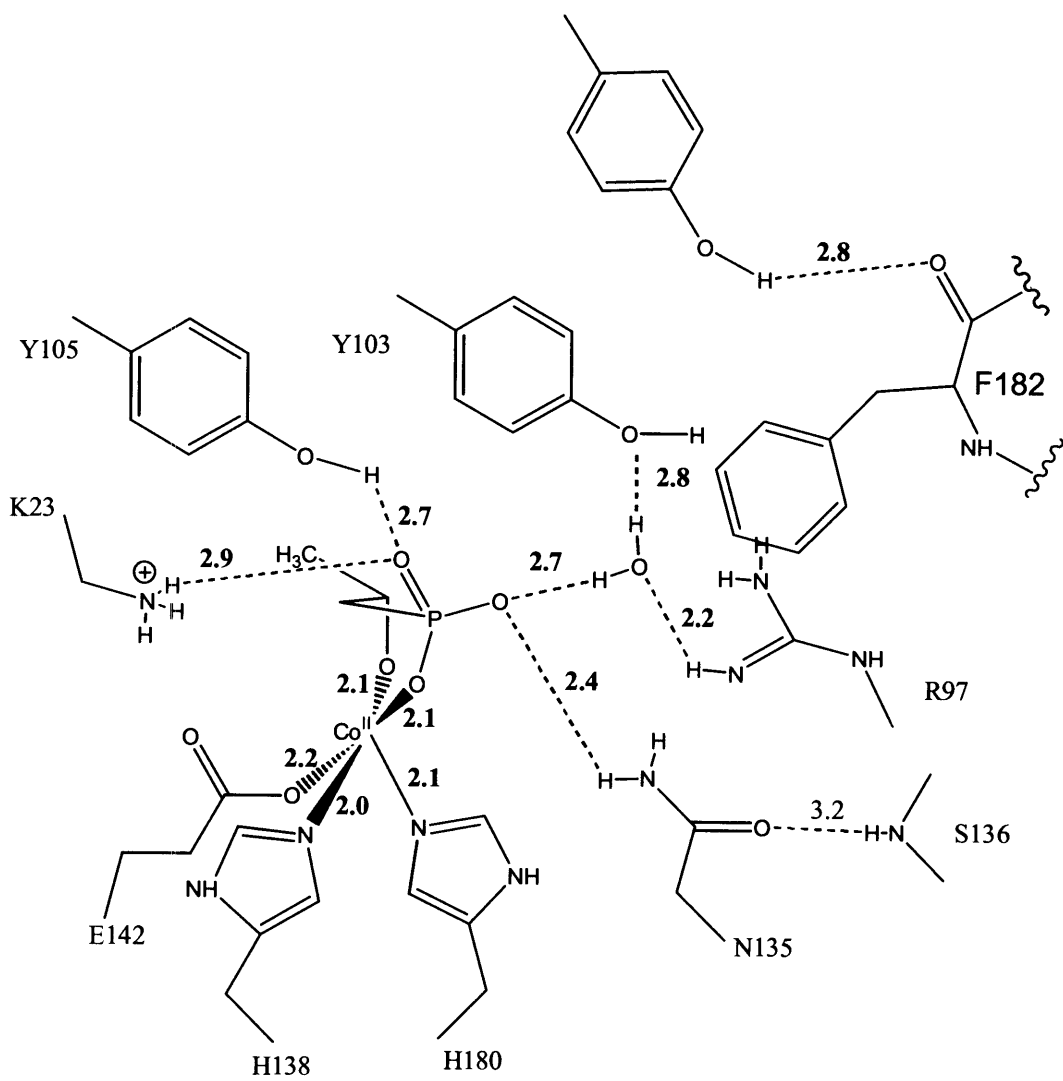
(d)



(e)

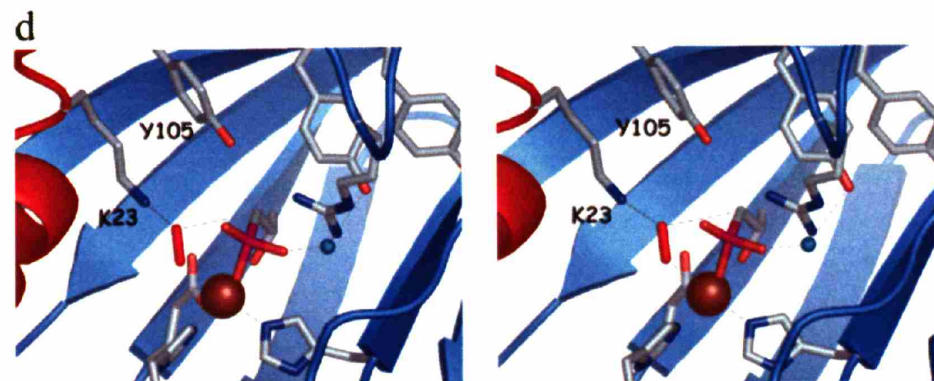
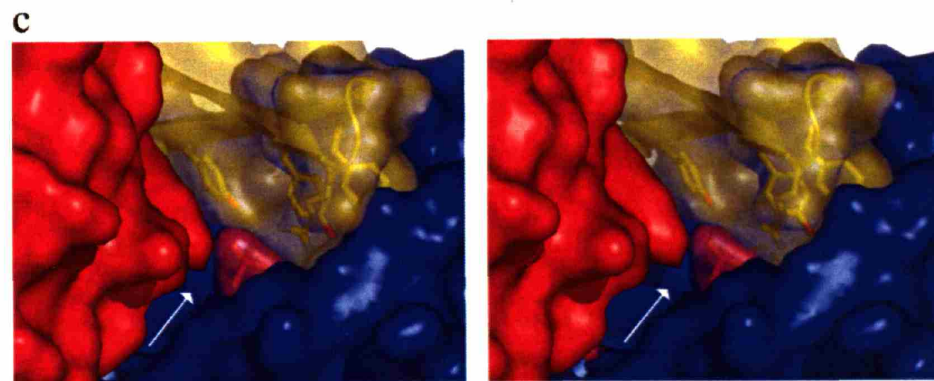
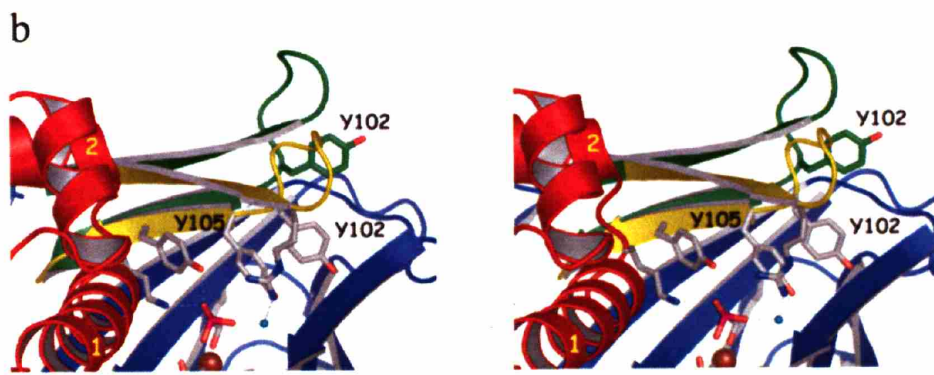
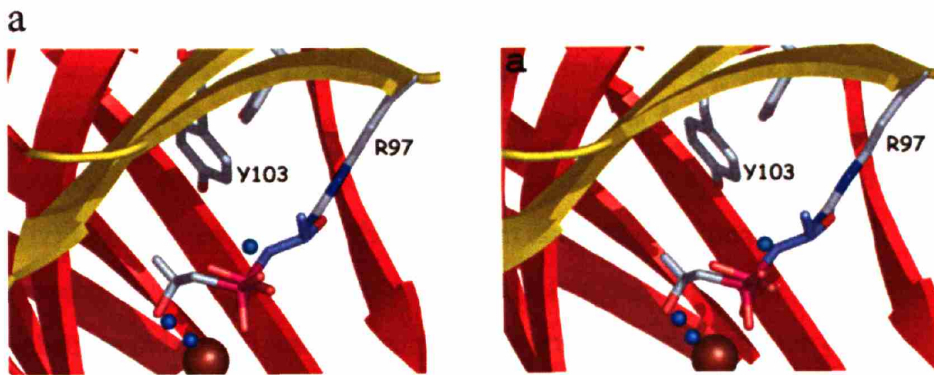


(f)



**Figure II.5. Structural insights into catalysis.**

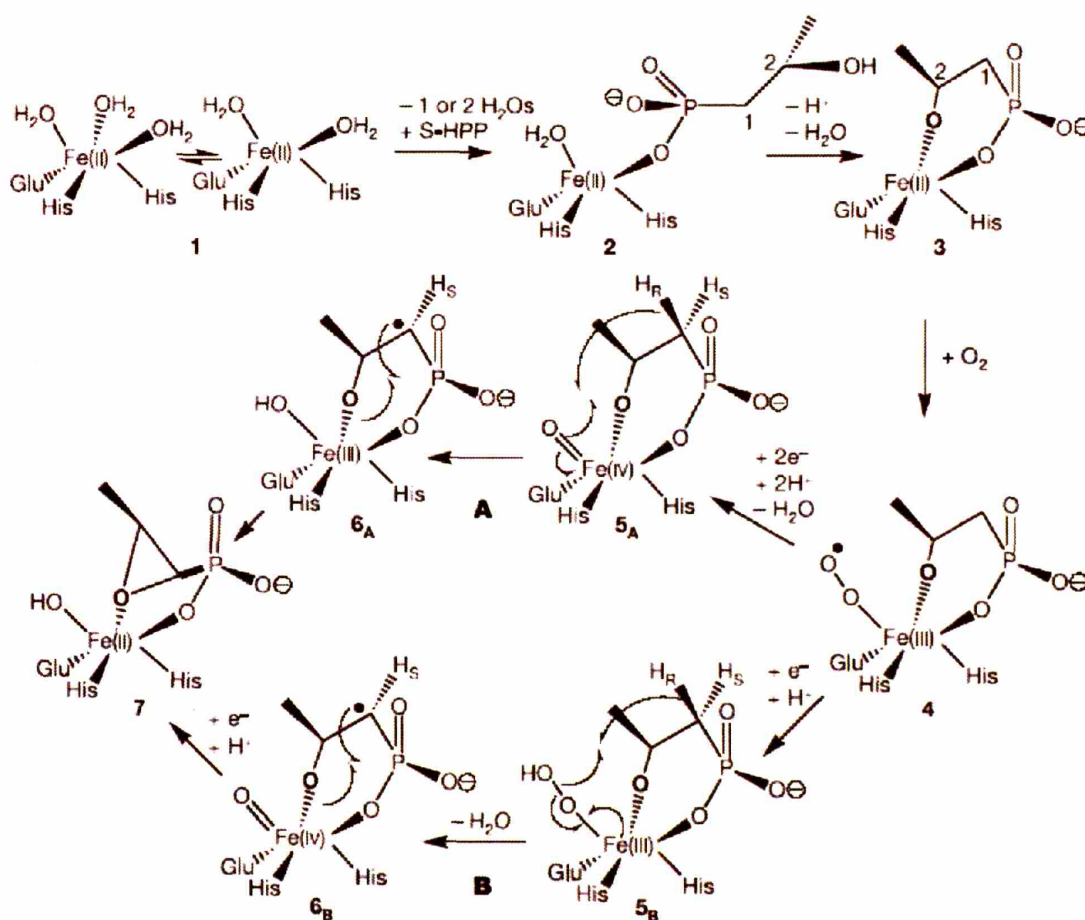
Structural insights into catalysis (stereoviews). **(a)** Superposition of monodentate and bidentate substrate binding modes in Fe(II)-HppE with the cantilever hairpin (yellow) is in the closed conformation.  $\beta$ -domain is in red, carbon (gray), oxygen (red), nitrogen (blue), with water molecules associated with the monodentate structure (blue spheres), and bidentate structure (cyan spheres). **(b)** A  $\beta$ -hairpin changes conformation between substrate-free (ribbons and residues in green) and substrate-bound (ribbons in yellow, residues in grey) states of Fe(II)-HppE. Both positions of Tyr102 are labeled. Ile92 and Asn106, which represent the start and end of the hairpin, respectively, are positioned under helix 2, and are not visible in this orientation. **(c)** Van-der-Waals surface representation after hairpin closure with  $\alpha$ -domain (red),  $\beta$ -domain (blue), cantilever hairpin (yellow), and substrate (magenta). White arrow indicates a proposed channel for dioxygen. **(d)** Model for a dioxygen bound structure. Dashed lines illustrate potential ligand-metal and dioxygen-sidechain interactions, as well as a possible path for hydrogen atom abstraction from an iron-hydroperoxo intermediate (mechanism B in Figure II.6.).





**Figure II.6. Possible mechanisms for substrate radical formation.**

In a two-step binding model, the substrate binds iron first as a monodentate (2) and then a bidentate ligand (3), displacing 2–3 bound waters. Substrate binding activates the iron for dioxygen binding (4). The mechanisms diverge at intermediate (4). In mechanism A, an Fe(IV)-oxo species abstracts the hydrogen atom (5<sub>A</sub>). In mechanism B, the hydrogen atom is abstracted by an iron-hydroperoxo intermediate (5<sub>B</sub>). For both mechanisms, the formation of a substrate radical (6) leads to ring closure and product formation (7).



## II.H. References

1. Liu, P., Murakami, K, Seki T, He, X, Yeung, SM, Kuzuyama, T, Seto, H & Liu, H. Protein purification and functional assignment of the epoxidase catalyzing the formation of fosfomycin. *J. Am. Chem. Soc.* **123**, 4619-4620 (2001).
2. Lobel, B. Short term therapy for uncomplicated urinary tract infection today. Clinical outcome upholds the theories. *Int. J. Antimicrob. Agents* **22**, 85-87 (2003).
3. Nakazawa, H., Kikuchi, Y., Honda, T., Isago, T. & Nozaki, M. Enhancement of antimicrobial effects of various antibiotics against methicillin-resistant *Staphylococcus aureus* (MRSA) by combination with fosfomycin. *J. Infect. Chemother.* **9**, 304-309 (2003).
4. Cassone, M., Campanile, F., Pantosti, A., Venditti, M. & Stefani, S. Identification of a variant "Rome clone" of methicillin-resistant *Staphylococcus aureus* with decreased susceptibility to vancomycin, responsible for an outbreak in an intensive care unit. *Microb. Drug Resist.* **10**, 43-49 (2004).
5. Hidaka, T., Goda, M, Kuzuyama, T, Takei, N, Hidaka, M & Seto, H. Cloning and nucleotide sequence of fosfomycin biosynthetic genes of *Streptomyces wedmorensis*. *Mol. Gen. Genet.* **249**, 274-280 (1995).
6. Hammerschmidt, F. Biosynthesis of Natural-Products with a P-C Bond. 8. On the Origin of the Oxirane Oxygen Atom of Fosfomycin in *Streptomyces fradiae*. *J. Chem. Soc. Perkin Trans. 1* **8**, 1993-1996 (1991).
7. Liu, P., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Biochemical and spectroscopic studies on (S)-2-hydroxypropylphosphonic acid epoxidase: a novel mononuclear non-heme iron enzyme. *Biochemistry* **42**, 11577-11586 (2003).

8. Guengerich, F.P. Cytochrome P450 oxidations in the generation of reactive electrophiles: epoxidation and related reactions. *Arch. Biochem. Biophys.* **409**, 59-71 (2003).
9. Costas, M., Mehn, M.P., Jensen, M.P. & Que, L., Jr. Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem. Rev.* **104**, 939-986 (2004).
10. Dunwell, J.M. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol. Genet. Eng. Rev.* **15**, 1-32 (1998).
11. Zhao, Z., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Mechanistic studies of HPP epoxidase: configuration of the substrate governs its enzymatic fate. *Angew. Chem. Int. Ed. Engl.* **41**, 4529-4532 (2002).
12. Roach, P.L., Clifton, IJ, Hensgens, CM, Shibata, N, Schofield, CJ, Hajdu, J & Baldwin, JE. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830 (1997).
13. Otwinowski Z, M. W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* **276**, 307-326 (1997).
14. Brunger, A. T., Adams, PD, Clore, GM, DeLano, WL, Gros, P, Grosse-Kunstleve, RW, Jiang, JS, Kuszewski, J, Nilges, M, Pannu, NS, Read, RJ, Rice, LM, Simonson, T & Warren, GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 ( Pt 5)**, 905-921 (1998).

15. Terwilliger, T.C. Automated main-chain model building by template matching and iterative fragment extension. *Acta Crystallogr. D* **59**, 38-44 (2003).
16. Navaza, J. AMoRe: an Automated Package for Molecular Replacement. *Acta Cryst. A* **50**, 164-182 (1994).
17. 29. Kissinger, C.R., Gehlhaar, D.K. & Fogel, D.B. Rapid automated molecular replacement by evolutionary search. *Acta Crystallogr. D* **55**, 484-491 (1999).
18. Laskowski, R., MacArthur, M., Moss, D. & Thornton, J. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291 (1993).
19. Dunwell, J.M., Culham, A., Carter, C.E., Sosa-Aguirre, C.R. & Goodenough, P.W. Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.* **26**, 740-746 (2001).
20. Woo, E.J., Dunwell, J.M., Goodenough, P.W., Marvier, A.C. & Pickersgill, R.W. Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. *Nat Struct Biol* **7**, 1036-1040 (2000).
21. Dunwell, J.M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**, 7-17 (2004).
22. Hausinger, R.P. FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes. *Crit. Rev. Biochem. Mol. Biol.* **39**, 21-68 (2004).
23. Liu, P., Mehn, MP, Yan, F, Zhao, Z, Que, L Jr & Liu, HW. Oxygenase activity in the self-hydroxylation of (S)-2-hydroxypropylphosphonic acid epoxidase involved in fosfomycin biosynthesis. *J. Am. Chem. Soc.* **126**, 10306-10312 (2004).

24. Ryle, M.J., Liu, A, Muthukumaran, RB, Ho, RY, Koehntop, KD, McCracken, J, Que, L Jr & Hausinger, RP. O<sub>2</sub>- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42**, 1854-1862 (2003).
25. Roach, P.L., Schofield, C.J., Baldwin, J.E., Clifton, I.J. & Hajdu, J. Crystallization and preliminary X-ray diffraction studies on recombinant isopenicillin N synthase from *Aspergillus nidulans*. *Protein Sci.* **4**, 1007-1009 (1995).
26. Valegård, K., Terwisscha, van Scheltinga, AC, Dubus, A, Ranghino, G, Oster, LM, Hajdu, J & Andersson, I. The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. *Nat. Struct. Mol. Biol.* **11**, 95-101 (2004).
27. Price, J.C., Barr, E.W., Tirupati, B., Bollinger, J.M., Jr. & Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **42**, 7497-7508 (2003).
28. Zhou, J., Kelly WL, Bachmann BO, Gunsior M, Townsend CA & Solomon EI. Spectroscopic studies of substrate interactions with clavaminic acid synthase 2, a multifunctional alpha-KG-dependent non-heme iron enzyme: correlation with mechanisms and reactivities. *J. Am. Chem. Soc.* **123**, 7388-7398 (2001).
29. Solomon, E.I., Brunold, TC, Davis, MI, Kemsley, JN, Lee, SK, Lehnert, N, Neese, F, Skulan, AJ, Yang, YS & Zhou, J. Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* **100**, 235-350 (2000).

30. Roach, P.L., Clifton, IJ, Hensgens, CM, Shibata, N, Long, AJ, Strange, RW, Hasnain, SS, Schofield, CJ, Baldwin, JE, Hajdu, J. Anaerobic crystallisation of an isopenicillin N synthase.Fe(II).substrate complex demonstrated by X-ray studies. *Eur. J. Biochem.* **242**, 736-740 (1996).
31. Hegg, E.L. & Que, L., Jr. The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur. J. Biochem.* **250**, 625-629 (1997).
32. DeLano, W.L. The PyMOL Molecular Graphics System. (DeLano Scientific, San Carlos, CA). (2002).

## Chapter III: Hydroxypropylphosphonic Acid Epoxidase: Structural Insights into the Mechanism of Regioselectivity

### Summary

Hydroxypropylphosphonic acid epoxidase is a unique mononuclear iron enzyme that catalyzes the oxidative epoxidation of (*S*)-2-hydroxypropyl phosphonic acid in the biosynthesis of the antibiotic fosfomycin in *Streptomyces wedmorensis*. In addition, the enzyme is able to convert the *R*-enantiomer of substrate to 2-oxo-propylphosphonic acid with regiospecificity. To probe this regiospecific mechanism of HppE, we determined two structures of Co(II)-HppE in complex with the *R*-stereoisomer of substrate at 2.1 Å resolution. These structures along with the previously determined structure of *S*-HPP-Co(II)-HppE elegantly illustrate how HppE is able to catalyze two distinct reactions, recognizing both enantiomers of substrate. Specifically, both enantiomers of substrate bind such that a single hydrogen atom is accessible for hydrogen atom abstraction. Three modes of *R*-stereoisomer binding are presented. In both bidentate binding modes, the C2 hydrogen alone is accessible for abstraction. With respect to the *S*-stereoisomer bidentate binding mode, only the C1 pro-*R* hydrogen is accessible. In addition, a water binding site near the putative dioxygen coordination site is observed both in the *R*- and *S*-stereoisomer complex structures and may serve to stabilize reactive iron-oxo species and/or facilitate O-O bond homolysis or heterolysis in the HppE catalytic mechanism.

### III.A. Introduction

Hydroxypropylphosphonic acid epoxidase (HppE) catalyzes the conversion of (*S*)-2-hydroxypropylphosphonic acid (*S*-HPP) to the antibiotic fosfomycin in *Streptomyces wedmorensis* (Figure III.1.a.)<sup>1</sup>. HppE is a mononuclear iron enzyme within the cupin superfamily and thus has a triad of ligands that coordinate the active site metal within a conserved  $\beta$ -barrel motif, the cupin-fold<sup>1-3</sup> (Figure III.2.a-b.). This ligand triad, also called the facial triad [His<sub>2</sub>(Glu/Asp)], lies at one end of the cupin-fold leaving three exchangeable coordination sites within the barrel core for substrate, co-substrate, and/or dioxygen binding. The catalytic diversity of cupin mononuclear iron enzymes derives from the tunable face of the cupin scaffold above the facial triad anchor. Residues on the opposite face of the facial triad are evolutionarily designed to accommodate a particular substrate and/or co-substrate within this invariant scaffold. HppE is specifically “tuned” to recognize its substrate, *S*-HPP, and has been shown to undergo an induced-fit conformational change upon bidentate binding (Figure III.2.c.) (chapter 2)<sup>4</sup>.

Irrespective of the reaction catalyzed, all mononuclear iron enzymes must provide a mechanism to: (1) activate the iron center to facilitate dioxygen binding, (2) protect high energy intermediates throughout the catalytic cycle, and (3) completely reduce dioxygen<sup>5</sup>. These enzymes require an electron donating substrate or co-substrate to lower the Fe<sup>II/III</sup> redox potential, favoring oxidation by dioxygen<sup>5</sup>. For HppE this is accomplished by substrate binding to the ferrous center. *S*-HPP binds in two modes, in a monodentate fashion via the C1 phosphonate substituent alone and in a bidentate fashion via both the C1 phosphonate and the C2 hydroxyl moiety (chapter 2)<sup>4</sup>. Both bidentate and monodentate binding modes are, in principle, able to facilitate dioxygen binding and



subsequent formation of a superoxide intermediate ( $\text{Fe-O}_2^{\bullet-}$ ). In contrast to the monodentate binding mode, bidentate substrate binding also triggers a conformational change in a  $\beta$ -hairpin, or cantilever hairpin, directly above the facial triad (Figure III.2.c.)<sup>4</sup>. This induced fit conformational change serves to reduce solvent accessibility, in effect protecting high energy iron-oxo intermediates that arise during catalysis. With respect to reduction of dioxygen, substrate also plays a crucial role, providing two reducing equivalents. The other two electrons are believed to come from a reductase *in vivo*<sup>2</sup>.

HppE recognizes both *S*-HPP and its enantiomer *R*-HPP and proves to be an efficient catalyst with respect to both reactions (Figure III.1.a-b.)<sup>6</sup>. Interestingly, *R*-HPP is not converted to an epoxide by HppE. Rather this stereoisomer exclusively yields the ketone 2-oxo-propylphosphonic acid (Figure III.1.b.)<sup>6</sup>. To probe the regioselective mechanism of HppE catalysis, both enantiomers of substrate were synthesized incorporating fluorines in place of C1 methylene hydrogens (*S*-FHPP and *R*-FHPP)<sup>6</sup>. *S*-FHPP was a strong binding inhibitor of HppE activity while *R*-FHPP was an active substrate, converted in a kinetically competent fashion to 1-difluoro-2-oxopropylphosphonic acid (Figure III.1.c-d.)<sup>6</sup>. This result suggests that a C1 hydrogen atom is abstracted from *S*-HPP substrate and a C2 hydrogen atom is abstracted from *R*-HPP. A radical centered at C1 would then yield the epoxide product whereas the C2 radical yields the ketone<sup>6</sup>. While the results from these fluorinated analogs provide an explanation for why a different product is formed from each isomer, it was unclear as to how a different hydrogen atom would be abstracted in each case. Do *R*- and *S*-stereoisomers bind the iron active site differently? If so, what is the cause of the

differential binding? If they bind in the same fashion, can differences in bond strengths explain the observed regioselectivity? To address these questions, we have solved two structures at 2.1 Å resolution of Co(II)-HppE with the *R*-stereoisomer of substrate bound. We find that *R*-HPP binds in a manner similar to the *S*-HPP binding modes and differ most strikingly with respect to the relative orientation of the C1, C2, and C3 positions. Furthermore, the structures illustrate that hydrogen atom accessibility to the iron-oxo species, is sufficient to explain HppE's regioselectivity.

### **III.B. Methods**

#### *III.B.1. Protein purification and crystallization*

Apo-HppE was expressed and purified by Feng Yan at the University of Texas, Austin as described previously<sup>1</sup>. In order to show the protein is active, an aliquot of apo-protein was reconstituted by Feng Yan at the University of Texas Austin with Fe(II)(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>·6H<sub>2</sub>O as described previously<sup>1,2</sup>. Following removal of excess metal with a G-10 column, this sample was tested for epoxidase activity using an established <sup>31</sup>P-NMR assay<sup>2</sup>. The remaining apo-protein solution was exchanged into 20 mM Tris-hydrochloride pH 8.0 and concentrated to 30 mg/ml.

Cocrystallizations of Apo-HppE with Fe(II) and substrate led to apo-HppE crystals, and soaking Fe(II)-HppE crystals with *R*-HPP was unsuccessful. Therefore, crystals of *R*-HPP in complex with Co(II)-HppE were obtained by soaking substrate into Tris-Co(II)-HppE, a hexagonal crystal form. These hexagonal crystals were grown from apo-HppE by mixing 2.0 μL of protein solution with 2.0 μL of precipitant solution (2.0 M ammonium sulfate, 0.1 M Tris-HCl pH 8.5). CoCl<sub>2</sub> was then added to the 4.0 μL drop

(0.3  $\mu\text{L}$  at 100mM). Tris-Co(II)-HppE crystals grew in 24 hrs. Soak 1 *R*-HPP-Co(II)-HppE crystals were obtained from Tris-Co(II)-HppE crystals by first transferring crystals from Tris buffer to HEPES buffer to displace any Tris molecules coordinating Co(II) via mass action. The removal of Tris was verified through a structure determination (data not shown). Crystals were then soaked in a solution containing *R*-HPP substrate. Specifically, Tris-Co(II)-HppE crystals were transferred to 4.0  $\mu\text{L}$  of the following soaking solution: 2.5 M ammonium sulfate, 400 mM sodium chloride, 100 mM HEPES pH 7.5 for 10 min. (*R*)-2-Hydroxypropylphosphonic acid (1.0  $\mu\text{L}$  at 2.5 M) was added to the drop and the resulting solution was placed over a reservoir of soaking solution for an additional 14 hrs.

Similarly, to obtain soak 2 *R*-HPP-Co(II)-HppE bidentate structure, Tris-Co(II)-HppE crystals were transferred to 2.0  $\mu\text{L}$  of HEPES soaking solution for 10 min. Sodium (*R*)-2-Hydroxypropylphosphonate (2.0  $\mu\text{L}$  at 2.5 M) was subsequently added to the drop. The crystals were placed over a reservoir of soaking solution for 14 hrs. Both *R*-HPP-HppE complex crystals were transferred to a cryo-protectant solution (30% xylitol, 2.0M ammonium sulfate, and 100mM Tris-HCl pH 8.0) for 30 s and subsequently cooled under a liquid nitrogen stream at 100K. Data were collected at National Synchrotron Light Sources, and were subsequently integrated and scaled in DENZO and SCALEPACK, respectively<sup>7</sup> (Table III.1.).

### *III.B.2. Structure determinations*

Since Tris-Co(II)-HppE crystals ( $P6_322$ ;  $a, b = 86.40 \text{ \AA}$ ,  $c = 220.25 \text{ \AA}$ ), *R*-HPP-Co(II)-HppE (soak 1) crystals, and *R*-HPP-Co(II)-HppE (soak 2) crystals are isomorphous (Table III.1.), both *R*-HPP-Co(II)-HppE structures were solved by rigid

body refinement in CNS<sup>8</sup> using only protein atoms from the refined Tris-Co(II)-HppE structure as an initial model (PDB code 1ZZC). The CNS protocol included simulated annealing against a maximum likelihood function target, positional refinement, and *B*-factor refinement; non-crystallographic symmetry restraints and sigma cutoff were not used. Manual adjustment of the model was done using XFIT<sup>9</sup> and iterative refinement was performed in CNS. The electron difference density peaks in both the *R*-HPP-Co(II)-HppE structures ( $8\sigma$  in soak 1 and  $9\sigma$  in soak 2) readily identify the location of the phosphonic moieties of substrate, and allowed for the placement of *R*-HPP substrate in each model. Topology and parameter files for *R*-HPP were obtained using HIC-Up<sup>10</sup>. The final model for *R*-HPP-Co(II)-HppE soak 1 contains residues 6-89 and 103-198 in molecule A, and 7-87 and 102-198 in molecule B (198 residues per molecule). The final model *R*-HPP-Co(II)-HppE soak 2 contains residues 6-89 and 103-198 in molecule A and 7-87 and 102-198 in molecule B.

Simulated annealing composite omit maps were used to validate both *R*-HPP-Co(II)-HppE structures. For both structures, all residues are in the allowed regions of the Ramachandran plot as calculated by PROCHECK<sup>11</sup>: *R*-HPP-Co(II)-HppE soak 1: 90.2% in the most favored regions, 9.8% in additionally allowed regions, and 0% in generously allowed regions; *R*-HPP-Co(II)-HppE soak 2: 90.6% in the most favored regions, 8.7% in additionally allowed regions, and 0.7% in generously allowed regions.

### III.C. Results

#### III.C.1. *R*-HPP-Co(II)-HppE: Three binding modes

To provide insight into the regiospecific mechanism of hydrogen atom abstraction catalyzed by HppE toward the synthesis of fosfomicin and 2-oxo-propylphosphonic acid,

two crystal structures of Co(II)-HppE in complex with (*R*)-2-hydroxypropyl phosphonic acid (*R*-HPP) were solved to 2.1 Å resolution. Co(II)-HppE is catalytically inert under aerobic conditions, and it was previously shown that X-ray structures of the Co(II)-HppE substrate complex are virtually identical to the Fe(II)-HppE complex structures (Chapter 2)<sup>4</sup>.

The two *R*-HPP crystal structures reveal three modes of substrate binding. In monodentate-1 binding mode (soak 2), *R*-HPP binds in a monodentate fashion via the phosphonate substituent alone (Figure III.3.a.; Figure III.4.a.) similar to that observed in the monodentate *S*-HPP complex crystal structures (chapter 2). In a second binding mode *R*-HPP also coordinates in a monodentate fashion but via the C2 hydroxyl moiety alone (soak 1) (Figure III.3.b-c.; Figure III.4.b.). In a third binding mode, *R*-HPP is coordinated to the metal center in a bidentate fashion (soak 2) via both the C2 hydroxyl and C1 phosphonate substituents (Figure III.3.d-e.; Figure III.4.c.), similar to that of the bidentate *S*-HPP binding mode (Figure III.3.f-g.; Figure III.4.d.). Importantly, binding modes 2 and 3 are similar. The major difference is the orientation of the phosphonate moiety in the HppE active site and the resulting Co(II)-phosphonate distances (2.2 versus 2.6 Å). In addition to the fact that 2.6 Å is long to be considered a bond, the electron density is discontinuous between Co(II) and phosphonate oxygen in the monodentate-2 structure whereas the bidentate structure shows continuous density (Figure III.3.b-c versus Figure III.3.d-e.).

In all three *R*-HPP binding modes, the positioning of the phosphorus atom of substrate was defined by a high  $\sigma$  peak (10  $\sigma$ , 8  $\sigma$ , and 9  $\sigma$  in binding modes monodentate-1, monodentate-2, and bidentate, respectively) (data not shown). The

phosphonate substituent is stabilized by charge-charge interactions and/or hydrogen bonding interactions with Y105 and K23. In both binding monodentate-2 and bidentate binding modes, the phosphonate is further stabilized by Y103 hydrogen bonding through a bridging water molecule (Figure III.4.b-c). This water molecule is in a position nearly coincident with that observed in the *S*-HPP:HppE structures (Fe(II) and Co(II)) (Figure III.3.b-g).

### III.C.2 Comparison of bidentate *R*- and *S*-HPP binding

As demonstrated previously, the HppE active site is “tuned” to recognize an amphipathic substrate molecule of appropriate size (chapter 2)<sup>4</sup>. The HppE hydrophobic pocket, including F182, L193, and L144, accommodates the C2 methyl group of *S*-HPP<sup>4</sup>. Interestingly, the position of the C2 methyl moiety of the *R*-HPP molecule (bidentate binding mode) is within 1 Å of the C2 methyl moiety of *S*-HPP (Figure III.5.a-c). Therefore, the same hydrophobic binding pocket serves to accommodate the *R*-enantiomer of substrate.

Although hydrogen atoms are not discernable in the electron density, the orientation of sp<sup>3</sup> carbon atoms (i.e. tetrahedral carbon atoms) can be accurately modeled within the electron density. Therefore, the positions of the hydrogen atoms covalently bound to these carbon atoms can be inferred. In the bidentate *S*-HPP-Fe(II)-HppE structure, dioxygen is proposed to bind *trans* to triad ligand H180 (chapter 2)<sup>4</sup>. The C1, pro-*R* hydrogen of *S*-HPP points toward this dioxygen binding site while the C2 hydrogen is oriented in the opposite direction (Figure III.5.c.; Figure III.5.d.II.). In both *R*-HPP monodentate-2 binding mode and bidentate binding mode structures, the C2

methylene hydrogens point away while the C1 hydrogen points toward the dioxygen binding site (Figure III.5.c.; Figure III.5.d.I.). Thus, despite the similar bidentate binding modes, the orientation of C1, C2 carbons and their hydrogen atoms are quite different.

### *III.C.3. Cantilever hairpin conformational change*

HppE uses an induced-fit mechanism to protect high-energy Fe-oxo species formed during catalysis<sup>4</sup>. A  $\beta$ -hairpin called the cantilever hairpin responds to the physiological substrate *S*-HPP when bound to the enzyme in the correct orientation by moving in and sealing off the active site<sup>4</sup>. Comparison of HppE structures with alternative substrate *R*-HPP bound in three different binding modes shows that this enantiomer of HPP is not capable of fully ordering and closing the cantilever over the active site (Table III.2. ; Figure III.6.b.). However, monodentate-2 and bidentate modes of *R*-HPP binding do elicit a partial response to the hairpin. While the hairpin is partially disordered in both structures, electron density shows that the residues of the strands are, for the most part, in the closed position (Table III.2. ; Figure III.6.b.). The electron density of these strands is of better quality in the bidentate *R*-HPP structure than in the monodentate-2 structure (data not shown), indicating that bidentate binding of *R*-HPP elicits a more complete conformational response. Hairpin residues Y103 and Y105 are in the same position in all *R*-HPP structures as they are in the fully closed *S*-HPP structure (Figure III.6.a.). Y102, on the other hand, is disordered in all monodentate-substrate structures, but is ordered and superimposable in structures of either enantiomer displaying the bidentate binding mode (Figure III.6.a.). R97 is the only hairpin residue that plays a role in *S*-HPP binding that is not ordered in the *R*-HPP bidentate structures (Figure III.6.a.). Thus, *R*-HPP when

bound bidentate to iron achieves much of the induced fit response of the physiological substrate, but only the natural substrate can completely close and order the hairpin residues over the active site.

#### *III.C.4. Conserved water binding site*

In both *R*-HPP monodenate-2 binding mode and bidentate binding mode structures, a water molecule binds in a position *trans* to facial triad ligand H180 but does not coordinate the metal. Interestingly, a water molecule was observed in a nearly identical position in one of three active sites within the asymmetric unit of the tetragonal Fe(II)-HppE structure in complex with *S*-HPP (chapter 2) (Figure III.3.f; Figure III.4.d.). In this (*S*)-HPP-Fe(II)-HppE structure, the water molecule is within hydrogen bonding distance to both the  $\gamma$ -amino and  $\gamma$ -carbonyl oxygen atom of Asn 197 (2.9 Å and 3.1 Å, respectively) and one of the  $\gamma$ -carboxylate oxygens of Glu 142 (2.7 Å). The binding site is also 3.8 Å away from the  $\epsilon$ -amino substituent of Lys 23 and 4.2 Å from one of the phosphonic acid oxygen atoms of (*S*)-HPP. Similarly, the analogous water molecule in both *R*-HPP complex structures is within hydrogen bonding distance to the  $\gamma$ -amino substituent of Asn 197, one of the  $\gamma$ -carboxylate oxygens of Glu 142, another water molecule, and one of the phosphonic acid oxygen atoms of *R*-HPP (Figure III.4.b-c.). The position of this water binding site is near the putative dioxygen coordination site, *trans* to facial triad ligand H180; the only open coordination site on iron when substrate is bound in a bidentate fashion.



### III.D. Discussion

#### III.D.1. Mechanism of regioselectivity

Experiments using fluorinated analogs suggest that the C1 hydrogen of the *S*-enantiomer of substrate is abstracted to produce a C1-centered radical intermediate *en route* to the epoxide fosfomycin, subsequent to oxidative ring closure<sup>6,12</sup>. In the synthesis of 2-oxo-propylphosphonic acid, the C2 hydrogen of the *R*-enantiomer of substrate is thought to be abstracted to produce a C2-centered radical intermediate, resulting in the formation of a ketone<sup>6</sup>. Our goal with these structural analyses is to address the mechanism of this selectivity. We find that both stereoisomers bind the active site metal in a bidentate fashion via the phosphonate and hydroxyl oxygen atoms. This substrate conformation orients the phosphonate substituent toward solvent while the C2 methyl substituent is positioned into a hydrophobic pocket. Notably, the position of the C2 methyl substituents in *R*-HPP and *S*-HPP bidentate binding modes within the hydrophobic pocket is similar (Figure III.5.). In the *R*-HPP and *S*-HPP bidentate structures, the metal center is five-coordinate with a distorted square pyramidal geometry, leaving a single open coordination site *trans* to triad ligand H180 for dioxygen binding (Figure III.3.d-g.). Subsequent to dioxygen binding, active site reduction is proposed to facilitate the formation of a transient, high energy iron-oxo species *trans* to H180<sup>2,13</sup>. This species would be expected to have sufficient oxidizing power to facilitate hydrogen atom abstraction of unactivated C-H bonds<sup>5</sup>. The structures of *R*-HPP in complex with Co(II)-HppE show three modes of substrate binding that are quite similar to those observed in the *S*-HPP-Co(II)-HppE structures. However, the orientation of the C1 and C2 carbon atoms, and thus their hydrogen atoms, relative to the putative dioxygen binding site are

different, providing a structural rather than chemical rationale for the observed regioselectivity in HppE catalysis. Structures of *S*-HPP in complex with HppE indicate that the C1 pro-R hydrogen of *S*-HPP would be the only hydrogen accessible for hydrogen atom abstraction by the iron-oxo intermediate (Figure III.5.d.II.). The C1 methyl of substrate points toward the iron-oxo coordination site while the C2 methylene is oriented in the opposite direction. In contrast, both *R*-HPP monodentate-2 and bidentate binding modes position the C2 methyl substituent, the  $\beta$ -hydrogen atom, and the C1 methylene hydrogens in the opposite orientations (Figure III.5.c-d.). From this analysis, it appears that the regioselectivity observed in HppE catalysis derives from the position and accessibility of the C1 and C2 hydrogens relative to that of the iron-oxo species (Figure III.7.).

When comparing bidentate structures of both substrate stereoisomers, the different orientations of C1 and C2 carbons appears to be due to conformational restrictions on the substrates. When bound in a bidentate fashion with the C2 hydroxyl and phosphonate oxygens anchored to the iron, newman projections illustrate that both enantiomers are in a staggered conformation about the C2-C1 and C1-P bonds (Figure III.8.). This staggered conformation minimizes the internal steric clash within the substrate molecules. Since the HppE hydrophobic pocket does not force the C2 methyl substituent of *R*-HPP in a particular orientation within the barrel, one could envision an alternative model of the *R*-isomer where the phosphorus, C1, and C2 atoms of the *R*-stereoisomer superpose with that of the *S*-HPP bidentate structure (Figure III.5.d.III.). Although no deleterious protein-substrate interactions would occur, this hypothetical orientation would position the C2 methyl close to the phosphonate moiety ( $\sim 2.8$  Å

compared to 5.2 Å observed in the X-ray structure). Based on Van der Waals radii for a phosphonate substituent and methyl substituent, a phosphonate oxygen and methyl carbon distance below ~3.2 Å would cause such a steric clash<sup>14</sup>. If the hypothetical structure (Figure III.5.d.III.) were observed, then it would be likely that C-H bond energies would dictate the product formed. Given the structural data, no such choice is available. Thus it does not appear to be the enzyme, but the substrate in its bidentate binding mode that dictates the regioselectivity.

### *III.D.2. Substrate induced-fit conformational change*

Previously, we have proposed the *S*-enantiomer of substrate (*S*-HPP) binds in a two-step binding process involving monodentate and bidentate binding modes (chapter 2). Following this model, substrate initially binds monodentate via coordination of the phosphonic acid moiety of substrate with the active site metal. Substrate subsequently binds in a bidentate fashion triggering an induced fit conformational change within the cantilever hairpin of the  $\beta$ -domain (Figure III.2.c.).

The *R*-HPP-Co(II)-HppE structure reveals a similar mechanism of catalytic initiation toward the synthesis of 2-oxo-propylphosphonic acid. *R*-HPP monodentate-1 binding mode is similar to that of the monodentate binding mode observed in the *S*-HPP-HppE complex structures; a binding mode stabilized by K23 and Y105 in both enantiomer structures. With respect to catalytic initiation, *R*-HPP subsequently binds in a bidentate fashion (bidentate binding mode) which is stabilized by hydrophobic interactions between the C2 methyl moiety of *R*-HPP and the HppE hydrophobic pocket as well as hydrogen bonding interactions with the *R*-HPP phosphonate substituent. In contrast to the

*S*-HPP-Co(II)-HppE bidentate complex structure, strand 1 of the cantilever hairpin remains disordered in the *R*-HPP bidentate binding mode and there is not interpretable density for Arg 97 within hairpin strand 1. The partial movement of the cantilever hairpin (i.e. strand 2 only) illustrates how HppE is “tuned” to bind the physiologically relevant enantiomer of substrate via an induced-fit conformational change (Figure III.6.). Although HppE can bind *R*-HPP in a similar fashion to that of *S*-HPP and catalyze its conversion to a unique product, sufficient structural differences exist to prevent complete closure of the HppE active site.

### *III.D.3. Mechanistic implications of conserved water binding site*

Several mechanisms have been proposed for the biosynthesis of fosfomycin (chapter 2); all require a high energy iron-oxo species to facilitate hydrogen atom abstraction from substrate resulting in the formation of a transient substrate radical intermediate<sup>2,4,6,13</sup>. Regardless of the precise mechanism, the two electron reduction of dioxygen and concomitant protonation (two protons) results in the loss of a single water molecule. It is therefore interesting that a water binding site is observed directly above the putative dioxygen binding site *trans* to H180 and at the end of the putative dioxygen channel in the tetragonal *S*-HPP-Fe(II) crystal structure. A water molecule in a nearly identical position is observed in both the *R*-HPP monodentate-2 and bidentate binding mode structures (Figure III.3.b-f. and Figure III.4.b-d.). In all three structures, the water molecule is stabilized via charge-dipole and hydrogen bonding interactions with K23 and N197, respectively. From a mechanistic perspective, it is plausible that these two residues facilitate O-O bond cleavage by providing a binding site for the leaving water molecule.

If one models superoxide binding to iron at distance  $\sim 2.1$  Å (e.g. NO-Fe(II) bond distance in IPNS:ACV:NO structure is 2.1 Å)<sup>15</sup>, the distal oxygen is coincident with the position of this water molecule. Therefore, K23 and N197 could also serve to stabilize the distal oxygen prior to O-O bond cleavage. In support of this hypothesis, it has been shown that K23A mutant protein is inactive<sup>4</sup>.

### III.E.1. Conclusion

This work describes how HppE is able to catalyze two distinct reactions, recognizing both enantiomers of substrate, with regiospecificity. The structures described support a regioselective mechanism of hydrogen atom abstraction that depends on the orientations of the C1 and C2 hydrogen atoms relative to the dioxygen binding site (i.e. iron-oxo coordination site). It is likely that internal sterics lead to the different orientations of C1 and C2 hydrogen atoms, and consequently the regioselectivity. These structures also show that a  $\beta$ -hairpin acts as a finely-tuned cantilever, only closing completely to a fully ordered state when the physiological substrate (*S*-HPP) is bound in the correct orientation. Bidentate *R*-HPP binding elicits a similar response, but one strand is not ordered. Monodenate binding fails to achieve the induced-fit conformational change to the same degree. Finally, a water binding site near the putative dioxygen coordination site may be important for facilitating O-O bond cleavage during catalysis by providing a binding site for the leaving water or this site may represent how dioxygen is stabilized when bound to iron.

### III.F. Figures and Tables

**Table III.1. Refinement and Data Statistics**

	Soak 1 R-HPP- Co(II)	Soak 2 R-HPP- Co(II)
<b>Data Collection</b>		
Space Group	P6 <sub>5</sub> 22	P6 <sub>5</sub> 22
Unit Cell Dimensions		
a (Å) = b	86.3	86.3
c (Å)	218.9	219.1
Beamline	ALS 5.0.2	ALS 5.0.2
Wavelength (Å)	1.1271	1.1271
Resolution Range (Å)	30-2.1	30-2.1
Observations	182423	392445
Unique Reflections	26816	28960
Redundancy	6.8	13.6
Completeness (%) <sup>a</sup>	92.2 (92.5)	99.9 (100.0)
I/σ(I) <sup>a</sup>	21.1 (4.5)	40.2 (5.1)
Rsym (%) <sup>a,b</sup>	5.4 (27.6)	6.9 (32.8)
<b>Refinement</b>		
R-factor <sup>d</sup> (Rfree <sup>e</sup> ) (%)	23.0 (26.6)	23.0 (25.7)
Resolution Range (Å)	30.0-2.1	30.0-2.1
No. Reflections	26122	28369
No. Molecules per asu <sup>c</sup>	2	2
Number of Atoms:		
Protein	2662	2682
Cobalt	2	2
R-HPP	16	16
Water	145	158
RMS Deviations:		
Bond Lengths(Å)	0.0095	0.011
Bond Angles(°)	1.57	1.71
Average B-Factors (Å <sup>2</sup> )		
Protein atoms	37.4	36.9
R-HPP	70.3	62.7
Cobalt	42.4	41.0
Water	42.5	42.0

<sup>a</sup>Values in parentheses are the highest-resolution shell. <sup>b</sup> $R_{sym} = \sum_{hkl} |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \langle I(hkl) \rangle$ , where  $I_i(hkl)$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I(hkl) \rangle$  is the mean of the intensity for the miller index (hkl). <sup>c</sup>ASU, asymmetric unit.

<sup>d</sup> $R_{factor} = \sum_{hkl} ||F_o(hkl)| - |F_c(hkl)|| / \sum_{hkl} |F_o(hkl)|$ . <sup>e</sup> $R_{free} = R_{factor}$  for a test set of reflections (5%).

<sup>f</sup>Statistics are for data set in which Friedel pairs are not merged.

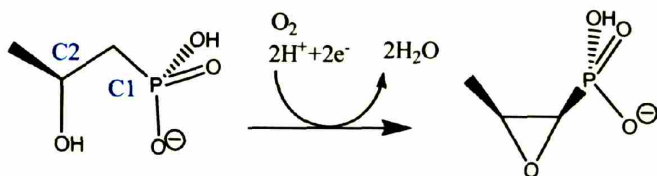
**Table III.2. Relative positions of cantilever hairpin strands 1 and 2.** “Open” refers to the cantilever hairpin conformation in the Fe(II)-HppE structure active site which is widely accessible. “Closed” refers to the cantilever hairpin conformation in the *S*-HPP-Fe(II) or *S*-HPP-Co(II)-HppE bidentate structures where top part of active site is sealed. “closed + disordered” \* = electron density suggests a closed position but modeling of strand is not possible (an example of this is shown in Figure III.6.b.).

	<b>strand 1</b>	<b>strand 2</b>
<b>R-HPP-Co(II)</b>	-	-
mono-1	open	open
mono-2	closed + disordered*	closed + ordered from Y103
bidentate	closed + disordered*	closed + ordered from Y102
<b>S-HPP-Co(II)<sup>4</sup></b>	-	-
mono	open	closed + ordered from Y103
bidentate	closed + all ordered	closed and all ordered
<b>Tris-Co(II)<sup>4</sup></b>	open + disordered	open + ordered from Y103

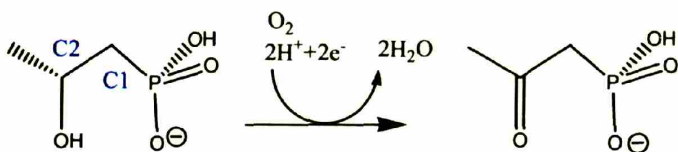
**Figure III.1. Overall reactions catalyzed by HppE.**

Overall reaction catalyzed by HppE using the (a) *S*-HPP as substrate<sup>1</sup> and (b) *R*-HPP as substrate<sup>6</sup>. (c) No reaction is observed using (*S*)-1,1-difluoro-2-hydroxypropylphosphonic acid (*S*-FHPP)<sup>6</sup>. (d) HppE catalyzes the formation of 1,1-difluoro-2-oxo-propylphosphonic acid using (*R*)-1,1-difluoro-2-hydroxypropylphosphonic acid (*R*-FHPP) as substrate<sup>6</sup>. The C1 and C2 positions are labeled (blue).

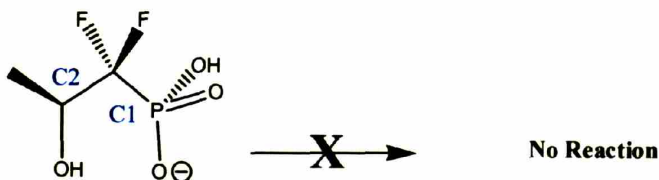
(a)



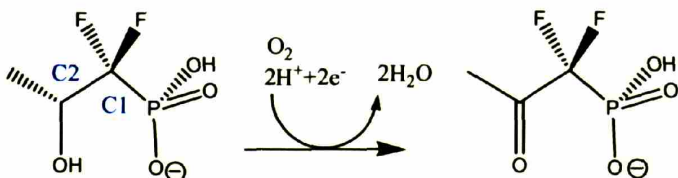
(b)



(c)



(d)



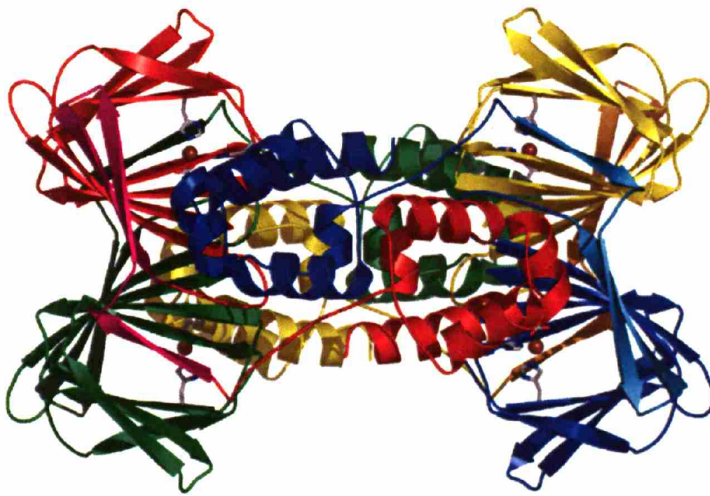


**Figure III.2. HppE structure.** (a) The HppE monomer is shown with the  $\alpha$ - and  $\beta$ -domains in blue, strand 1 and the cantilever hairpin in cyan, and the inter-domain linker in green. Triad ligands are shown in ball-in-stick: nitrogen (blue), oxygen (red), and carbon (grey). (b) The HppE tetramer is shown colored by molecule: red, yellow, blue, and green. Strand 1 and the cantilever hairpin is colored as in (a) for the blue molecule, magenta for the red molecule, dark yellow for the yellow molecule, and dark green for the green molecule. (c) The HppE active site is shown in space-fill representation: molecule 1 is blue ( $\alpha$ -domain), molecule 2 is yellow ( $\beta$ -domain), the cantilever hairpin is red. The hairpin is shown in the open conformation, using the Fe(II)-HppE structure, and the closed conformation, using the *S*-HPP-Fe(II)-HppE structure.

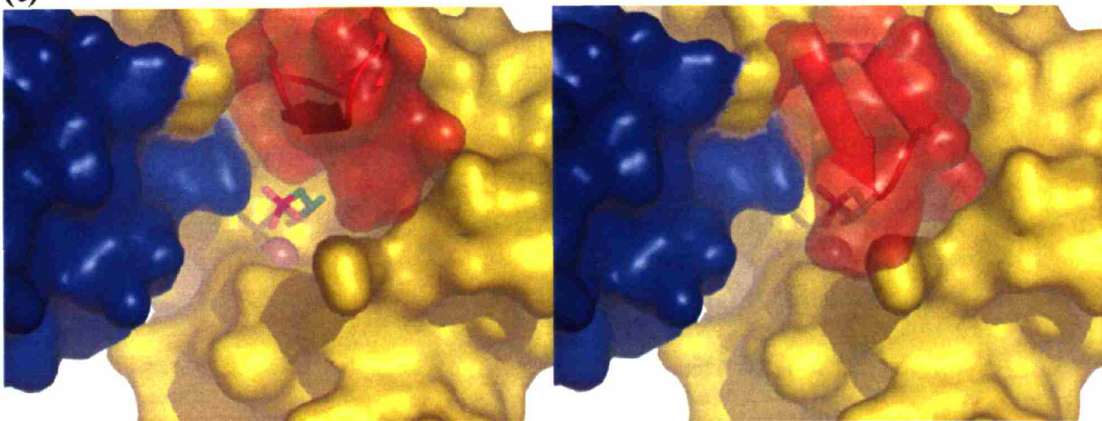
(a)



(b)

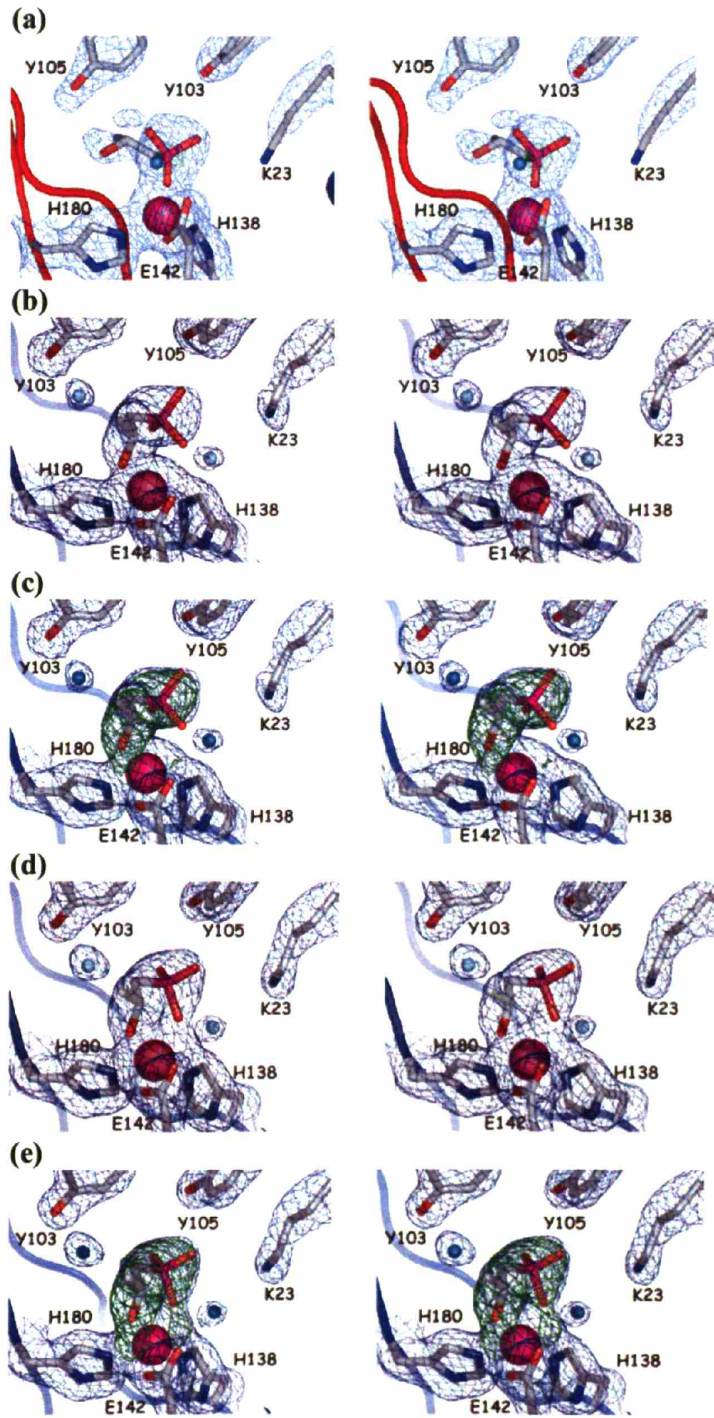


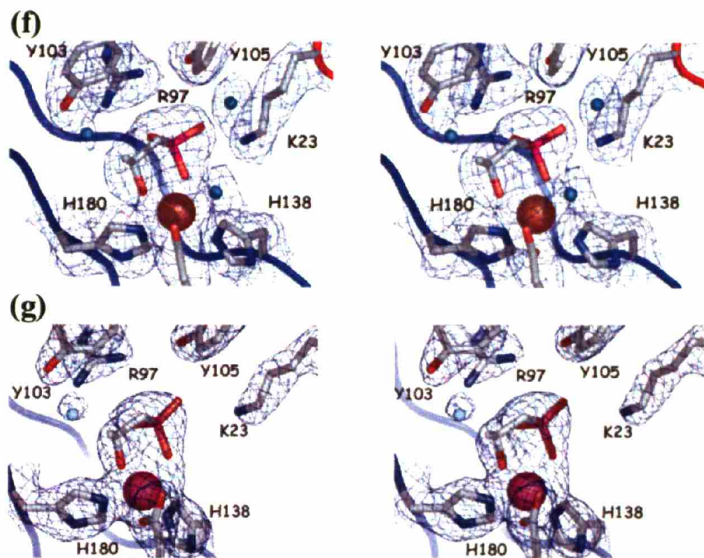
(c)



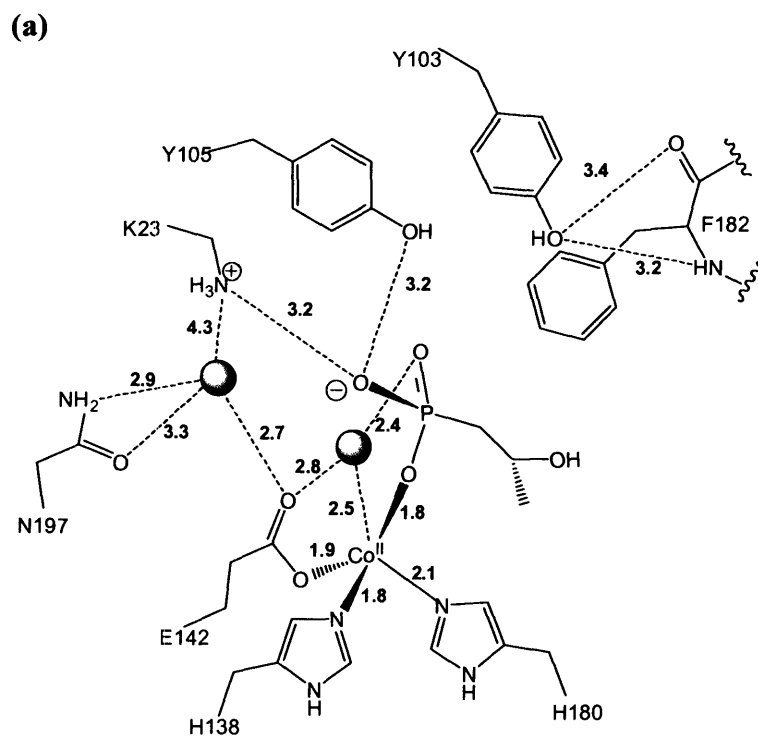
**Figure III.3. HppE active sites displayed with  $2Fo-Fc$  maps contoured at  $1.0 \sigma$ .**

Stereoviews of (a) *R*-HPP monodenate-1 binding mode structure with  $\beta$ -strands shown as red ribbons, substrate and residues shown as ball-and-stick (carbon (gray), oxygen (red), nitrogen (blue), phosphorus (magenta)), water molecules as spheres (cyan), and the Co(II) metal as a sphere (magenta). (b,c) *R*-HPP monodenate-2 binding mode with  $\beta$ -strands shown as blue ribbons and substrate is shown as in (a), with an omit-map of substrate is contoured at  $3.0 \sigma$  ((c), green). (d,e) *R*-HPP bidentate binding mode structure is shown as in (b). An omit-map of substrate is contoured at  $3.0 \sigma$  ((e), green). (f) *S*-HPP-Fe(II)-HppE molecule C structure with  $\beta$ -strands shown as blue ribbons, the loop containing Lys 23 is shown in red, residues and water molecules are shown as in (a-e), and the Fe(II) metal is shown as a sphere (brown)<sup>4</sup>. (g) *S*-HPP-Co(II)-HppE structure showing the bidentate binding mode shown as in (f). The Co(II) metal is shown as a sphere (magenta)<sup>4</sup>.

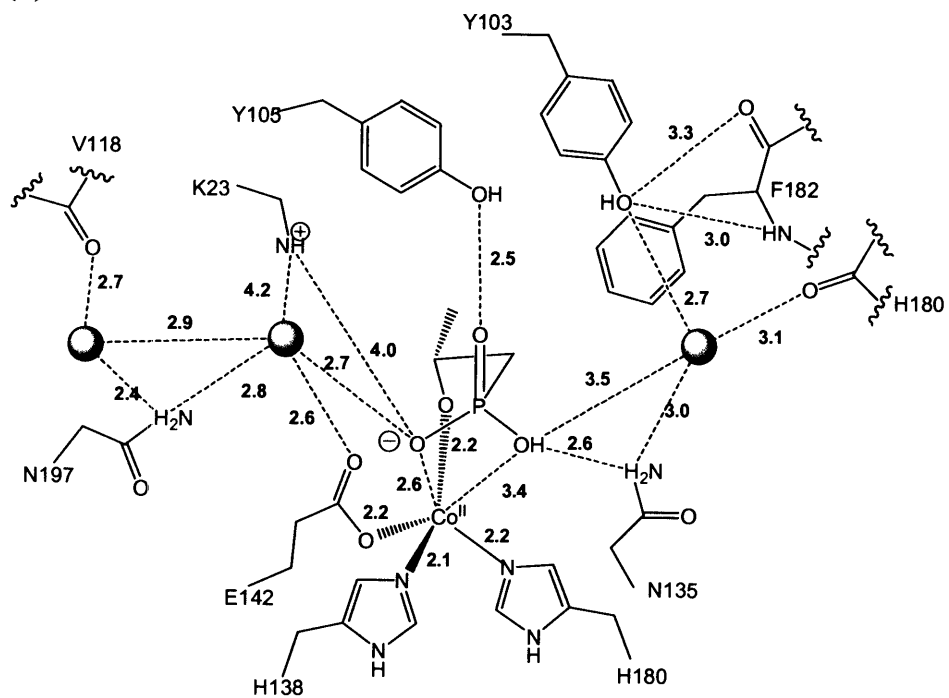




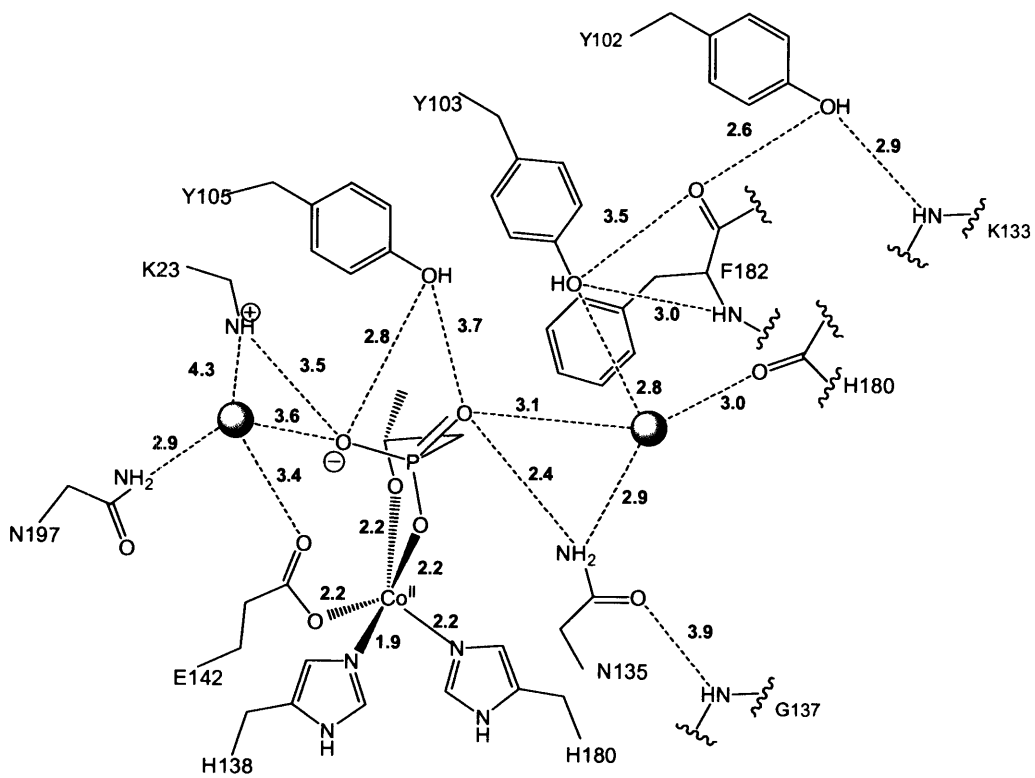
**Figure III.4. Two dimensional active site contact maps.** Distances (Å) shown with dotted lines and grey spheres are water molecules (a) *R*-HPP monodenate-1 binding mode; (b) *R*-HPP monodenate-2 binding mode; (c) *R*-HPP bidentate binding mode; (d) *S*-HPP-Fe(II)-HppE molecule C.



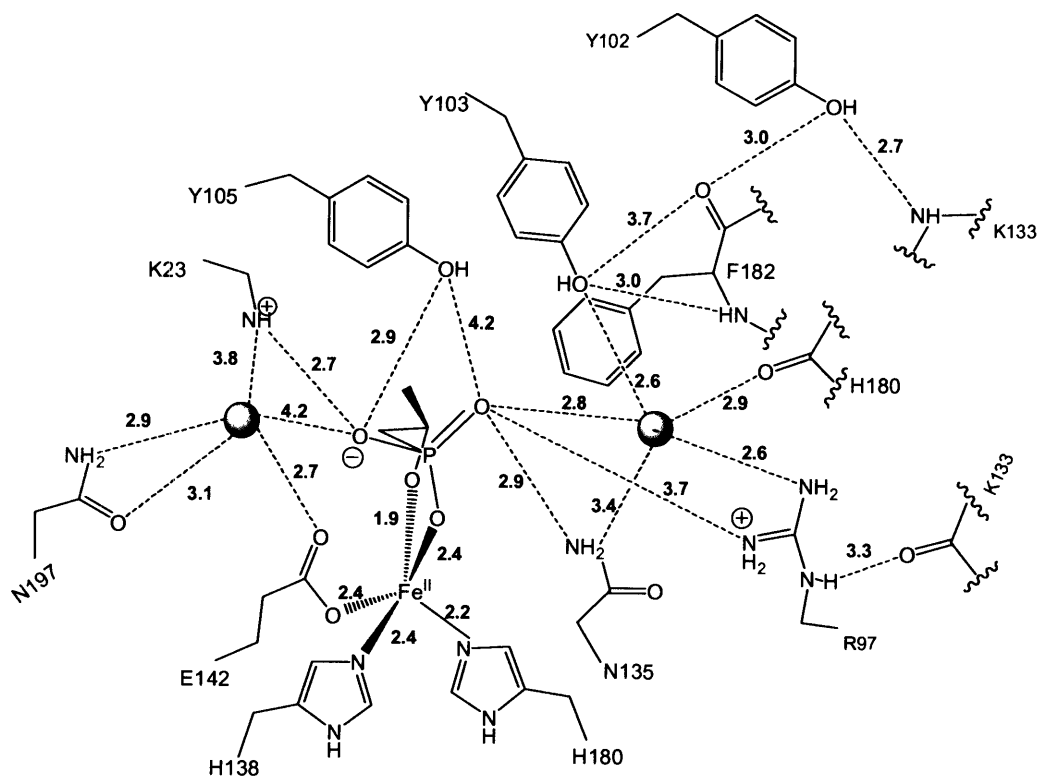
(b)



(c)



(d)





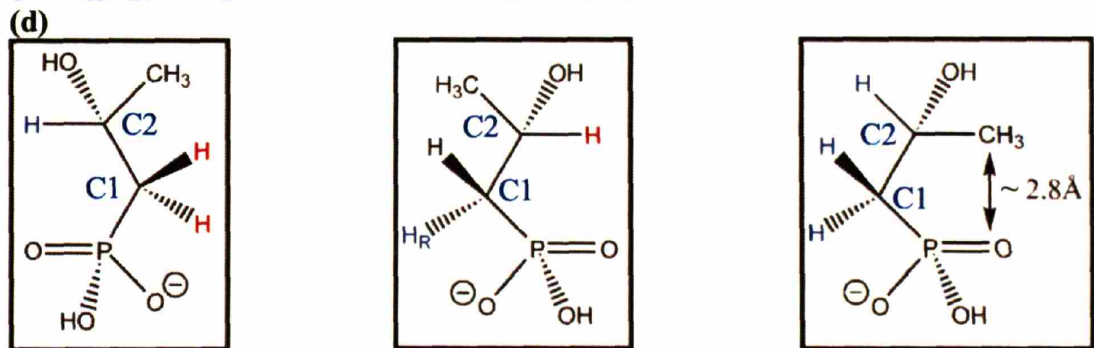
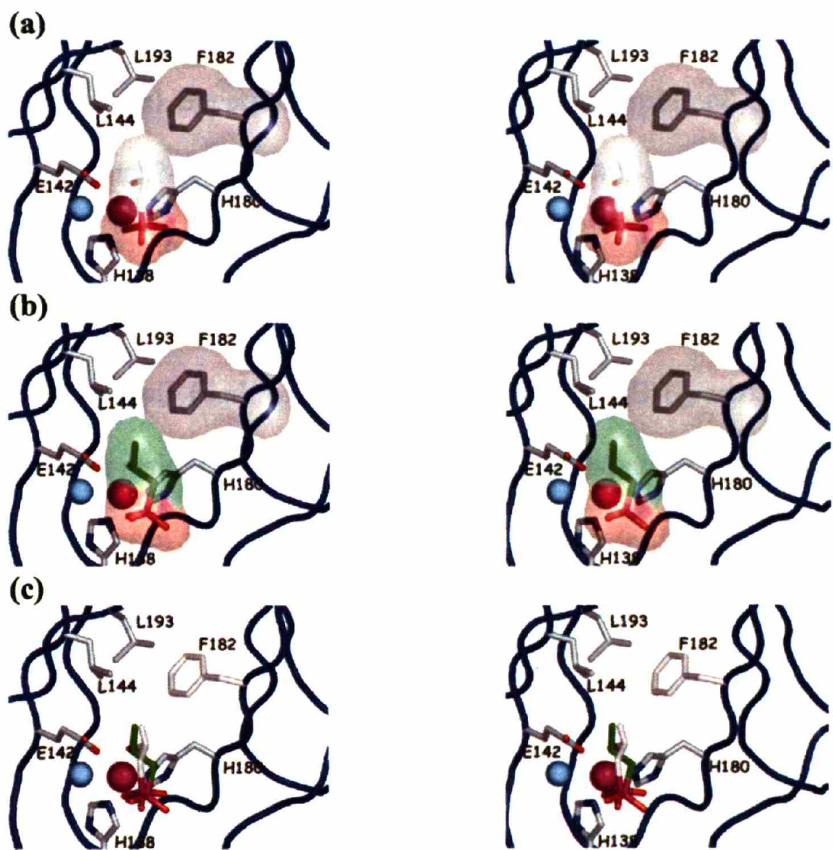
**Figure III.5. Structural insights into regioselectivity of HppE.**

(a) *S*-HPP-Co(II) bidentate binding mode structure (in stereo) with  $\beta$ -strands shown as blue ribbons, substrate and residues shown as ball-and-stick (carbon (gray), oxygen (red), nitrogen (blue), phosphorus (magenta)), and the Co(II) metal as a sphere (magenta). Van der Waals surfaces are shown around the *S*-HPP molecule and F182. A water molecule near the putative dioxygen binding site is shown as a sphere (cyan).

(b) *R*-HPP-Co(II) bidentate binding mode structure (in stereo) with  $\beta$ -strands and residues shown as in (a). Substrate is shown in ball-in-stick representation, colored as in (a) except for carbon (green). Van der Waals surfaces are shown around the *R*-HPP molecule and F182. A water molecule is shown as (a).

(c) *S*-HPP-Co(II) and *R*-HPP-Co(II) bidentate binding mode structures aligned (in stereo) with  $\beta$ -strands shown as blue ribbons, *S*-HPP and residues shown as ball-and-stick (carbon (gray), oxygen (red), nitrogen (blue), phosphorus (magenta)), and the Co(II) metal as a sphere (magenta). The *R*-HPP molecule carbon atoms are shown in green. A water molecule is shown as (a).

(d) Stereoisomer substrate conformations. (I) *R*-HPP shown in the conformation observed in the bidentate binding mode structures. (II) *S*-HPP shown in the conformation observed in monodenate-2 binding mode and 3 structures. (III) *R*-HPP shown in a hypothetical conformation (not observed) aligned by the phosphorus, C1-carbon, and C2-carbon atoms of *R*-HPP with those of *S*-HPP in the bidentate binding mode. The predicted distance between the methyl and phosphonate oxygen atoms is shown. In (a)-(c), blue hydrogens are accessible for hydrogen atom abstraction and red hydrogens are inaccessible.



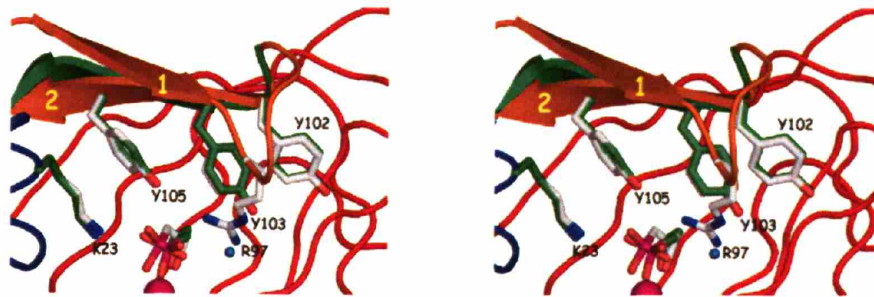
I R-HPP

II S-HPP

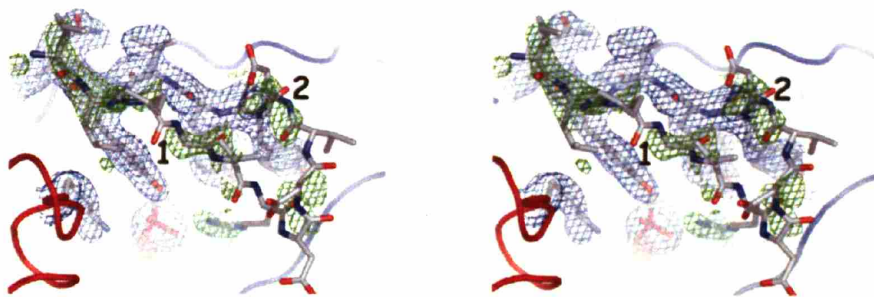
III Hypothetical R-HPP

**Figure III.6. Stereoviews of cantilever hairpin induced-fit conformational change.** (a) Superposition of the cantilever hairpin and substrate for *S*-HPP-Co(II)-HppE and *R*-HPP-Co(II)-HppE structures. The *R*-HPP and *S*-HPP substrate molecules are shown in ball-and-stick: carbon (grey, *S*-HPP; green, *R*-HPP), oxygen (red), and phosphorus (magenta). The cantilever hairpin from each structure is shown: *S*-HPP-Co(II)-HppE (dark orange); *R*-HPP-Co(II)-HppE (green). Y105, Y103, Y102, R97 (hairpin residues) and K23 are shown in ball-in-stick: carbon (grey for *S*-HPP-Co(II)-HppE; green for *R*-HPP-Co(II)-HppE), oxygen (red), and nitrogen (blue). The *S*-HPP-Co(II)-HppE cobalt ion (magenta) and water (cyan) molecule are shown as spheres. The *S*-HPP-Co(II)-HppE is shown with:  $\alpha$ -domain (blue),  $\beta$ -domain (red), and the *R*-HPP-Co(II)-HppE structure is not shown. (b) The cantilever hairpin in the *S*-HPP-Co(II)-HppE structure shown in ball-in-stick representation. A 2Fo-Fc map (blue) and a Fo-Fc map (positive density, green) contoured at 1.0  $\sigma$  and 3.0  $\sigma$ , respectively, calculated from the *R*-HPP bidentate structure is shown over the *S*-HPP bidentate model. The positive difference density (green) is coincident with the position of residues 101-106, strand 1 of the hairpin in the *S*-HPP bidentate model. This is an example of “closed and disordered” (see Table III.2.) where electron density shows residues in a closed position but density is not good enough for strand 1 to be modeled. Strand 1 and strand 2 of the cantilever hairpin are shown (black numbers).

(a)

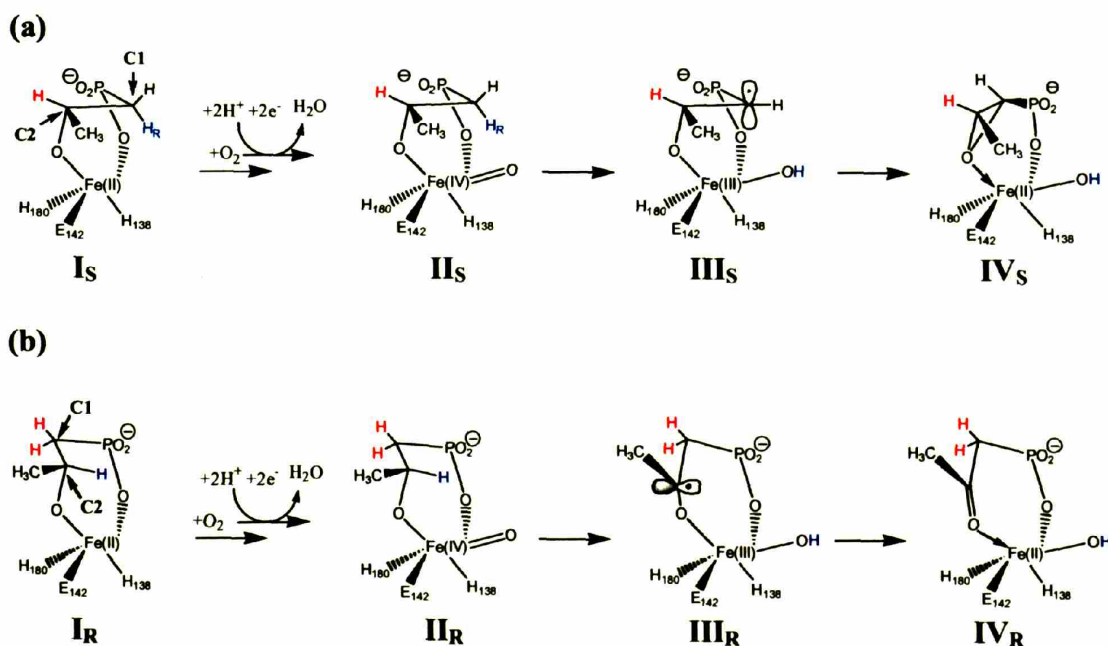


(b)



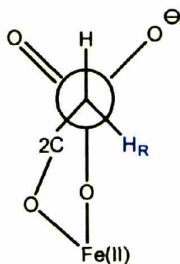
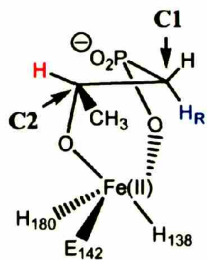
**Figure III.7. Potential mechanisms of regioselective H-atom abstraction.**

(a) A scheme illustrating the conversion of *S*-HPP ( $I_S$ ) to fosfomicin ( $IV_S$ ). Triad ligands are labeled throughout. The putative ferryl-oxo intermediate ( $II_S$ ) and the C1-centered radical intermediate ( $III_S$ ) are shown. Based on the structure data presented herein, the pro-R hydrogen atom at the C2-position *S*-HPP is accessible to H-atom abstraction (blue) while the hydrogen atom at the C1-position is inaccessible to H-atom abstraction (red). (b) A scheme illustrating the conversion of *R*-HPP ( $I_R$ ) to 2-oxo-propylphosphonic acid ( $IV_R$ ). The putative ferryl-oxo intermediate ( $II_R$ ) and the C2-centered radical intermediate ( $III_R$ ) are shown. The pro-R hydrogen atom at the C2-position *R*-HPP is inaccessible to H-atom abstraction (red) while the hydrogen atom at the C1-position is accessible to H-atom abstraction (blue).

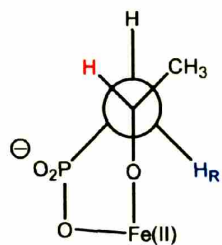


**Figure III.8. Newman projections of *R*- and *S*-HPP anchored in a bidentate fashion to iron. (a) *S*-HPP and (b) *R*-HPP substrates shown looking down the C1-P bond (top) and C2-C1 bond (bottom) illustrating bidentate substrate binding in a staggered conformation. Hydrogen atoms accessible to hydrogen atom abstraction are shown in blue. Hydrogen atoms inaccessible to hydrogen atom abstraction are shown in red (See Figure III.7.).**

(a)

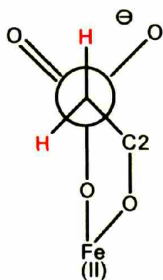
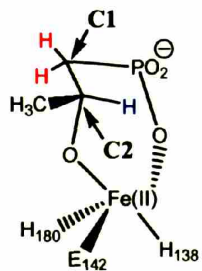


**S-HPP: Down C1-P bond**

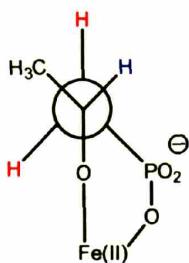


**S-HPP: Down C2-C1 bond**

(b)



**R-HPP: Down C1-P bond**



**R-HPP: Down C2-C1 bond**

### III.G. References

1. Liu, P., Murakami, K, Seki, T, He, X, Yeung, SM, Kuzuyama, T, Seto, H & Liu, H. Protein purification and function assignment of the epoxidase catalyzing the formation of fosfomycin. *J Am Chem Soc* **123**, 4619-4620 (2001).
2. Liu, P., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Biochemical and spectroscopic studies on (S)-2-hydroxypropylphosphonic acid epoxidase: a novel mononuclear non-heme iron enzyme. *Biochemistry* **42**, 11577-11586 (2003).
3. Dunwell, J. M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**, 7-17 (2004).
4. Higgins, L. J., Yan, F., Liu, P., Liu, H. W. & Drennan, C. L. Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature* (2005). ([www.nature.com](http://www.nature.com) doi: 10.1038/nature03924)
5. Costas, M., Mehn, M. P., Jensen, M. P. & Que, L., Jr. Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem Rev* **104**, 939-986 (2004).
6. Zhao, Z., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Mechanistic studies of HPP epoxidase: configuration of the substrate governs its enzymatic fate. *Angew Chem Int Ed Engl* **41**, 4529-4532 (2002).
7. Otwinowski Z, M. W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods Enzymol* **276**, 307-326 (1997).
8. Brunger, A. T., Adams, PD, Clore, GM, DeLano, WL, Gros, P, Grosse-Kunstleve, RW, Jiang, JS, Kuszewski, J, Nilges, M, Pannu, NS, Read, RJ, Rice, LM, Simonson, T & Warren, GL. Crystallography & NMR system: A new software



- suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 ( Pt 5)**, 905-921 (1998).
9. McRee, D. E. XtalView/Xfit--A versatile program for manipulating atomic coordinates and electron density. *J Struct Biol* **125**, 156-165 (1999).
  10. Kleywegt, G. Hetero-compound Information Centre - Uppsala (HIC-Up). (2005).
  11. Laskowski RA, M. M., Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291 (1993).
  12. Hammerschmidt, F. Biosynthesis of Natural Products with a P-C Bond. 7. Synthesis of [1,1-2H<sub>2</sub>]-, [2,2-2H<sub>2</sub>]-, (R) and (S)-[1-2H<sub>1</sub>](2-Hydroxyethyl)phosphonic acid and (R,S)-[1-2H<sub>1</sub>](1,2-Dihydroxyethyl)phosphonic acid and incorporation studies into fosfomycin in *Streptomyces fradia*. *J Org Chem* **56**, 2364-2370 (1991).
  13. Liu, P., Mehn, MP, Yan, F, Zhao, Z, Que, L Jr & Liu, HW. Oxygenase activity in the self-hydroxylation of (s)-2-hydroxypropylphosphonic acid epoxidase involved in fosfomycin biosynthesis. *J Am Chem Soc* **126**, 10306-10312 (2004).
  14. Bondi, A. van der Waals Volumes and Radii. *Journal of Physical Chemistry* **68**, 441-451 (1964).
  15. Roach, P. L. Clifton, IJ, Hensgens, CM, Shibata, N, Schofield, CJ, Hajdu, J & Baldwin, JE. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830 (1997).

## Chapter IV: Hydroxypropylphosphonic Acid Epoxidase: An Enzyme with Broad Substrate Specificity

### Summary

Hydroxypropylphosphonic acid epoxidase (HppE) is a unique mononuclear iron enzyme within the cupin superfamily of proteins that catalyzes the biosynthesis of the antibiotic fosfomicin in *Streptomyces wedmorensis* from *S*-hydroxypropylphosphonic (*S*-HPP). The enzyme catalyzes the conversion of the *R*-HPP enantiomer to 2-oxo-propylphosphonic acid. Further demonstrating the broad substrate specificity of this enzyme, HppE catalyzes the conversion of (*S*)-2-hydroxy-2-phenylethylphosphonic acid (*S*-HPEP) to (1*R*,2*S*)-2-phenyl-1,2-epoxyethylphosphonic acid, and (*R*)-2-hydroxy-2-phenylethylphosphonic acid to 2-oxo-2-phenylethylphosphonic acid. This work presents the first substrate analog structure of HppE: *S*-HPEP in complex with Co(II)-HppE at 2.2 Å resolution. This structure provides insight into how HppE is able to bind larger hydrophobic substrates while maintaining complete regioselectivity.

#### IV.A. Introduction

The last enzyme in the fosfomycin biosynthetic pathway in *Streptomyces wedmorensis* is hydroxypropylphosphonic acid epoxidase (HppE)<sup>1</sup>. HppE is a mononuclear iron enzyme within the cupin superfamily of proteins that is structurally unique when compared to all other known mononuclear iron enzymes<sup>2</sup>. The enzyme is a tetramer<sup>1</sup>, each monomer consists of two domains, an  $\alpha$ -domain and a  $\beta$ -domain (chapter 2)<sup>2</sup>. All mononuclear iron enzymes within the cupin superfamily house a triad of ligands that coordinates the iron center within a  $\beta$ -barrel fold<sup>3</sup>. This ligand triad, His<sub>2</sub>(Glu/Asp), binds the iron in a trigonal pyramidal geometry. As a consequence of this ligand geometry, three coordination sites remain available for substrate and/or co-substrate binding<sup>3-5</sup>. In the absence of substrate, HppE is either six coordinate or five coordinate with water molecules occupying the vacant coordination sites (chapter 2)<sup>2</sup>. *En route* to fosfomycin biosynthesis (Figure IV.1.a.), HppE is specifically designed to bind (*S*)-2-hydroxypropylphosphonic acid, *S*-HPP, in a bidentate fashion via the 2-hydroxyl and phosphonic acid oxygens (chapter 2)<sup>2</sup>. Bidentate binding of *S*-HPP to cobaltous HppE and ferrous HppE results in an induced fit conformational change in two strands of the  $\beta$ -barrel, known as the cantilever hairpin<sup>2</sup>. These two strands serve to seal the active site and protect intermediates formed during catalysis. With this bidentate substrate conformation, the negative charge of the phosphonate moiety of substrate is stabilized via charge-charge interactions and hydrogen bonding with R97 and K23 while the hydrophobic portion of substrate is buried in a hydrophobic pocket (chapter 2)<sup>2</sup>.

HppE recognizes the C2 enantiomer of substrate, *R*-HPP. Rather than catalyzing the formation of an enantiomeric epoxide, *R*-HPP is stereospecifically converted to 2-

oxo-propylphosphonic acid (Figure IV.1.b.)<sup>6</sup>. In contrast to *S*-HPP, bidentate *R*-HPP binding does not yield a complete conformational change in the cantilever hairpin, movement only occurs in one strand of the hairpin (chapter 3). This observation underscores the importance of the appropriate size, charge distribution, and substrate orientation to hairpin motion and active site closure.

In both substrate enantiomer Co(II)-HppE complex structures, bidentate binding leaves a sixth coordination site open for dioxygen binding<sup>2</sup> (chapter 2 and chapter 3). From the orientation of this binding site it is possible to explain the observed one-to-one regioselective mechanism of hydrogen atom abstraction. Presumably, hydrogen atom abstraction at the C1 position results in epoxide formation, while hydrogen atom abstraction at the C2 position results in ketone product formation. The bidentate binding modes of the substrate enantiomers allows for only a single accessible site for hydrogen atom abstraction, the C1 position for the *S*-enantiomer and the C2 position for the *R*-enantiomer (chapter 3).

Further illustrating the catalytic flexibility of the HppE active site, the enzyme is able to facilitate the oxidation of various substrate analogs. Surprisingly, enzymatic activity is observed when the C2 methyl substituent of either substrate enantiomer is changed to a phenyl moiety, although the rate of catalysis is attenuated 10-fold<sup>7</sup>. It seems intuitive that broad substrate specificity would not correlate with high regio- or stereoselectivity. However, HppE binds (*S*)-2-hydroxypropylphosphonic acid (*S*-HPEP) and exclusively converts this substrate analog into (1*R*,2*S*)-2-phenyl-1,2-epoxyethylphosphonic acid (Figure IV.1.c.). Analogous to *R*-HPP, the enzyme also

recognizes (*R*)-2-hydroxypropylphosphonic acid (*R*-HPEP), exclusively converting it into 2-oxo-2-phenylethanoic acid (Figure IV.1.d.)<sup>7</sup>.

The observation that HppE is able to catalyze the conversion of *S*-HPEP to an epoxide and *R*-HPEP to a ketone with regiospecificity raises several questions: (1) how does a hydrophobic pocket designed to bind HPP accommodate a phenyl substituent at the C2 position rather than methyl group, and (2) what is the mechanism of regioselective hydrogen atom abstraction? To address these questions, the structure of Co(II)-HppE in complex with the substrate analog, *S*-HPEP, was solved to 2.2 Å resolution.

## IV.B. Methods

### IV.B.1. Protein purification and crystallization

As in chapter 2 and chapter 3, protein purification was performed in the laboratory of Dr. Ben Liu by Dr. Feng Yan. The *fom4* gene encoding HppE was previously cloned into a pET24b vector (PL001)<sup>1,8</sup>. BL21(DE3) cells were transformed with PL001 and subsequently grown in LB medium at 37°C. Cultures were induced at an O.D. at 600 nm of 0.6 to a final concentration of 0.1 mM IPTG. Cells were broken via sonication and purified using (1) a 40-60% ammonium sulfate fractionation, (2) gravity DEAE-Sepharose anion exchange chromatography, and (3) FPLC Mono Q anion exchange chromatography as described previously<sup>1</sup>. The protein solution was dialyzed into 20mM Tris-hydrochloride pH 8.0 and concentrated to 30 mg/ml.

Hexagonal Tris-HppE crystals were grown using the sitting drop vapor diffusion method (chapters 2 and 3). The 30 mg/ml protein solution (2μL) was mixed in a 1:1 ratio with 2 μL of HppE precipitant solution, 2.0 M ammonium sulfate, 0.1 M Tris-HCl

pH 8.5. Cobaltous chloride (0.3  $\mu\text{L}$  at 100mM) was added to the sitting drop. The solution was equilibrated over 500  $\mu\text{L}$  of HppE precipitant solution for 24 hrs. Hexagonal crystals grew in approximately 24 hrs. Several crystals were then transferred to 5  $\mu\text{L}$  of 2.5 M ammonium sulfate, 400 mM NaCl, 100 mM HEPES pH 7.5 (HppE soaking solution). This protocol is used in order to displace the Tris molecules from the HppE active site via mass action as previously described<sup>2</sup>. Crystals were allowed to equilibrate in a hanging drop over 500  $\mu\text{L}$  of well solution for 4 hrs; both the drop and the well contained the HEPES soaking solution. Crystals were subsequently transferred to a fresh drop (4  $\mu\text{L}$ ) of HppE soaking solution. *S*-HPEP (1 $\mu\text{L}$  at 3 M) was immediately added to this drop without mixing (final volume of 5  $\mu\text{L}$ ). The drop was then placed over 500  $\mu\text{L}$  of HppE soaking solution for approximately 1 hr. Crystals were then transferred to a cryo-protectant (30% xylitol, 2.0M ammonium sulfate, and 100mM Tris-HCl pH 8.0), soaked for less than 10 s, and subsequently cooled with gaseous nitrogen at 100 K. Data were collected at the National Synchrotron Light Source (NSLS), and data were subsequently integrated and scaled in DENZO and SCALEPACK, respectively<sup>9</sup> (Table IV.1.).

#### *IV.B.2. Structure determination and refinement*

Since *S*-HPEP crystals are derived from the Tris-HppE hexagonal crystals described in chapter 2<sup>2</sup>, the Tris-HppE model (PDB Entry 1ZZC) including only protein atoms was used for rigid body refinement in CNS<sup>10</sup> against the *S*-HPEP dataset. The first round of refinement allowed for rigid body movement of both the HppE  $\alpha$ - and  $\beta$ -domains (residues 8-72 and residues 103-198) as well as the interdomain region (residues

103-198) (3 groups). This refinement was performed sequentially at 4 Å and at 2.2 Å resolution. The next round of refinement included all secondary structural elements within the asymmetric unit as rigid bodies (51 groups). This latter refinement was performed at 2.2 Å resolution. Iterative model building in XFIT and subsequent refinement in CNS resulted in the final *S*-HPEP complex structure (Table IV.1.). Topology and parameter files for the *S*-HPEP molecule were obtained using HIC-Up<sup>11</sup>. Positional and *B*-factor refinement were used throughout the refinement. Finally, the model was refined via simulated annealing against a MLF (maximum likelihood function) target. A simulated annealing composite omit map was used to validate the final model. Non-crystallographic symmetry restraints and sigma cutoff were not used. The final *S*-HPEP structure contains residues 6-87;102-131;136-198 in molecule A and 6-87;102-158;159-198 in molecule B (198 residues per molecule). Ramachandran plot as calculated by PROCHECK<sup>12</sup>: 90.3% in the most favored regions, 9.7% in additionally allowed regions, and 0% in generously allowed regions.

#### IV.C. Results

Previous structures (chapter 2)<sup>2</sup> have shown that Co(II)-HppE is a good model system for Fe(II)-HppE and is stable under aerobic conditions thus we have used the Co-enzyme for structural studies of *S*-HPEP binding. The structure of *S*-HPEP complex of cobaltous HppE (Figure IV.2.) reveals a new binding mode for substrate distinct from that of both *R*-HPP and *S*-HPP complex structures, described previously (chapters 2 and 3). Like the HPP enantiomer structures, *S*-HPEP binds the metal center (i.e. cobalt) in a bidentate fashion, via an oxygen atom from the phosphonate moiety and a second oxygen

from the C2 hydroxyl substituent. The difference in binding modes is in the substrate orientation. The proposed dioxygen binding site *trans* to H180 in the fosfomycin biosynthetic mechanism (chapter 2)<sup>2</sup> is occupied by the phosphonate moiety of *S*-HPEP, and the phosphonate binding site in the bidentate *S*-HPP structures, *trans* to E142, is occupied by a water molecule. As a result, the metal center is six coordinate whereas all *S*-HPP and *R*-HPP bidentate structures are five coordinate. The phosphonate is stabilized via H-bonding to the side-chain amide of N197, the  $\epsilon$ -amino group of K23, and the backbone carbonyl of E20 through a bridging water molecule (Figure IV.3.). The  $\epsilon$ -amino group of K23 also serves to stabilize the phosphonate via charge-charge interactions. Notably, the hydroxyl substituent of tyrosine 105, which hydrogen bonds to HPP phosphonates, is not proximal to the HPEP phosphonate oxygens (5.5 Å). The orientation of the phosphonate precludes any interaction with the guanidinium of R97 as observed in HPP structures and may be important for strand 1 movement in the cantilever hairpin.

. The water coordinating cobalt is stabilized by an extended hydrogen bonding network via a water molecule and including side-chain amide of N135, another water molecule, the backbone nitrogen of F182, the carbonyl of H180, and the hydroxyl substituent of Y103 of the cantilever hairpin. Also in contrast to the majority of *S*-HPP and *R*-HPP structures, *S*-HPEP does not display a monodentate binding mode. Instead the *S*-HPEP in the second molecule in the asymmetric unit is disordered.



#### IV.D. Discussion

The structure of *S*-HPEP Co(II)-HppE complex illustrates how the HppE active site, designed to recognize *S*-HPP and catalyze fosfomycin synthesis, accommodates a phenyl substituent at the C2 position in place of a methyl group. The unique binding mode of *S*-HPEP relative to both HPP substrate enantiomers suggests a ligand rearrangement mechanism prior to HppE catalysis.

Like the *R*-HPP binding modes and unlike the natural substrate (chapter 3), *S*-HPEP binding does not result in an induced fit conformational change in which the cantilever hairpin closes to seal the active site (Figure IV.4.a.). Substrate analog binding does, however, induce a cooperative conformational change of hydrophobic pocket residues similar to that observed in both the *R*-HPP and *S*-HPP complex crystal structures (Figure IV.5.). The C2 phenyl substituent *S*-HPEP packs directly against L144 and the phenyl moiety of F182 (Figure IV.6.). Movement of the latter induces a cooperative displacement of M153 and M172 in turn. Above the buried C2 phenyl of *S*-HPEP are Y105 and Y103 (Figure IV.6.) which provide a steric constraint to cantilever motion similar to that observed in the Tris-Co(II)-HppE structure (chapter 2), which like HPEP is larger than substrate, and physically prevents residues of the cantilever hairpin from moving in toward the Fe, and sealing off the active site<sup>2</sup>. Finally since the C2 phenyl must bind within a hydrophobic pocket designed to accommodate the smaller C2 methyl substituent of *S*-HPP, L193 moves toward the core of the  $\beta$ -barrel relative to the *S*-HPP structure (Figure IV.5.c.) to create a slightly larger binding pocket.

The different orientation of *S*-HPEP relative to the triad ligands (chapters 2 and 3), could be a result of a rearrangement that is necessary to minimize C2 phenyl substituent

steric clashing (Figure IV.7.). If *S*-HPEP were to bind in a manner more similar to the *S*-HPP bidentate binding mode (i.e. phosphonate *trans* to E142), significant steric clashes would occur between the HppE hydrophobic core (L193 and L144) and the C2 phenyl substituent of *S*-HPEP (Figure IV.7.b.).

In the observed *S*-HPEP binding mode dioxygen would bind *trans* to E142, displacing the water. Dioxygen binding in this position, however, can not explain the regioselectivity of product formation. If one were to predict based on a modeling exercise alone, the reaction pathway (i.e. whether hydrogen atom abstraction occurs at the C1 or C2 position) for the observed conformation of *S*-HPEP, the C2 position seems most attractive (Figure IV.8.). From a thermodynamic perspective, one could argue that the ketone product (**VIIB**) would be favored when compared to the epoxide (**VIIA**)<sup>13</sup>. Also from a kinetic perspective, the radical intermediate in pathway **B** should be lower in energy than the radical intermediate in pathway **A**, therefore the transition state energy en route to **VIA** would be higher than that of **VIB**; that is, assuming a late transition state. Paradoxically, **VIIA** is the only observed product, suggesting this orientation of *S*-HPEP is not catalytically relevant.

Thus we propose that a ligand rearrangement is required prior to catalytic initiation (Figure IV.9.). Specifically, dioxygen binding could trigger a change in the phosphonate coordination site to a position *trans* to E142. Although *S*-HPEP is sterically prohibited from binding in a way directly analogous to that of *S*-HPP bidentate binding mode (Figure IV.7.b.), *S*-HPEP binding can be modeled such that the C2 phenyl substituent is in a position that lies somewhere between the *S*-HPP binding mode (Figure IV.7.b.) and the observed *S*-HPEP binding mode (Figure IV.7.a.). Although this

hypothetical binding mode (Figure IV.7.c.) is probably higher in energy than the observed binding mode, it positions the C2 substituent snugly within the HppE hydrophobic pocket allowing Y105 (strand 2 cantilever hairpin residue) to move closer to the active site and to hydrogen bond with the phosphonate moiety as it does in the *S*-HPP structures (Figure IV.7.d.). This hypothetical *S*-HPEP binding mode is consistent with the mechanism proposed for fosfomycin biosynthesis<sup>2</sup> and the observed product obtained from *in vitro* *S*-HPEP turnover (Figure IV.9.).

A similar mechanism of ligand rearrangement has been proposed for the biosynthesis of clavulanic acid catalyzed by another mononuclear iron enzyme, clavamate synthase (CAS)<sup>14</sup>. CAS is an  $\alpha$ -KG-dependent enzyme having relatively broad substrate specificity, catalyzing three non-sequential reactions in the biosynthetic pathway<sup>15</sup>. In the absence and presence of CAS substrate,  $\alpha$ -KG binds in a bidentate fashion to the metal center via the 2-oxo and C1 carboxylate oxygens such that the carboxylate ligand is *trans* to the cupin motif 1 histidine ligand, equivalent to H138 in HppE. Upon NO binding, a dioxygen surrogate, the  $\alpha$ -KG carboxylate ligand shifts to the coordination site *trans* to the motif 2 histidine ligand (H180 in HppE) while the position of the 2-oxo ligand remains unchanged<sup>14</sup>.

In addition, the ligand rearrangement mechanism (Figure IV.9.) is consistent with the observed regioselectivity in the conversion of *R*-HPEP to 2-oxo-2-phenylethylphosphonic acid (Figure IV.1.d.)<sup>7</sup>. Since *R*-HPP bidentate binding mimics that of *S*-HPP with respect to the orientation of the C2 substituent (i.e. methyl), it is plausible that the C2 phenyl of *R*-HPEP would bind in a similar orientation within the hydrophobic pocket to that of *S*-HPEP. Upon dioxygen binding, phosphonate ligand rearrangement

from a position *trans* to H180 to a position *trans* to E142 would allow for hydrogen atom abstraction at the C1 position alone, resulting in the exclusive formation of the observed ketone product.

#### **IV.E. Conclusion**

Although tuned for binding *S*-HPP en route to fosfomycin synthesis, the plasticity of the HppE active site allows for several substrate analogs to bind and turnover to yield a variety of products. Key features common to all HppE substrate-enzyme interactions include bidentate substrate binding via the hydroxyl and phosphonate oxygens and the use of a pliable hydrophobic pocket to accommodate the alkyl substituents of substrate. The broad specificity of the HppE active site provides the potential of using this enzyme to tailor make epoxide products and molecules containing the unusual carbon-phosphorus bond.

## IV.F. Figures and Tables

**Table IV.1. Data and Refinement Statistics**

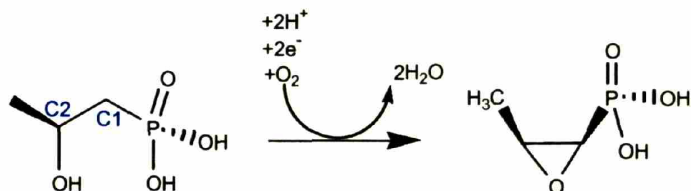
	<b>S-HPEP-Co(II)</b>
<b>Data Collection</b>	
Space Group	P6 <sub>5</sub> 22
Unit Cell Dimensions	
a (Å) = b	86.31
c (Å)	220.1
Beamline	SSRL 9-1
Wavelength (Å)	0.9797
Resolution Range (Å)	50-2.2
Observations	165934
Unique Reflections	25324
Redundancy	6.6
Completeness (%) <sup>a</sup>	99.2 (99.8)
I/σ(I) <sup>a</sup>	14.0 (6.1)
R <sub>sym</sub> (%) <sup>a,b</sup>	8.8 (31.1)
<b>Refinement</b>	
R-factor <sup>d</sup> (R <sub>free</sub> <sup>e</sup> ) (%)	24.1 (21.4)
Resolution Range (Å)	50-2.2
No. Molecules per asu <sup>c</sup>	2
Number of Atoms:	
Protein	2636
Cobalt	2
S-HPEP	26
Water	232
RMS Deviations:	
Bond Lengths(Å)	0.0084
Bond Angles(°)	1.5
<b>Average B-Factors (Å<sup>2</sup>)</b>	
Protein atoms	43.1
S-HPEP	31.3

<sup>a</sup>Values in parentheses are the highest-resolution shell. <sup>b</sup> $R_{sym} = \sum_{hkl} |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \langle I(hkl) \rangle$ , where  $I_i(hkl)$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I(hkl) \rangle$  is the mean of the intensity for the miller index (hkl). <sup>c</sup>ASU, asymmetric unit.

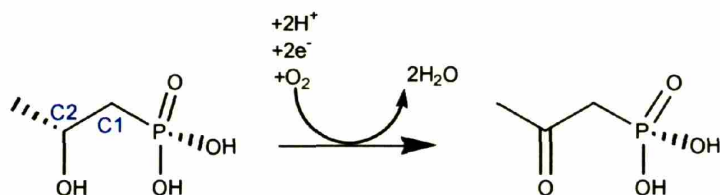
<sup>d</sup> $R_{factor} = \sum_{hkl} ||F_o(hkl)| - |F_c(hkl)|| / \sum_{hkl} |F_o(hkl)|$ . <sup>e</sup> $R_{free} = R_{factor}$  for a test set of reflections (5%).

**Figure IV.1. Reactions catalyzed by HppE.** HppE converts **(a)** (*S*)-2-hydroxypropylphosphonic acid to fosfomycin, **(b)** (*R*)-2-hydroxypropylphosphonic acid to 2-oxo-propylphosphonic acid, **(c)** (*S*)-2-hydroxy-2-phenylethylphosphonic acid to (1*R*,2*S*)-2-phenyl-1,2-epoxyethylphosphonic acid, and **(d)** (*R*)-2-hydroxy-2-phenylethylphosphonic acid to 2-oxo-2-phenylethylphosphonic acid. The C1 and C2 positions of each substrate are labeled (blue).

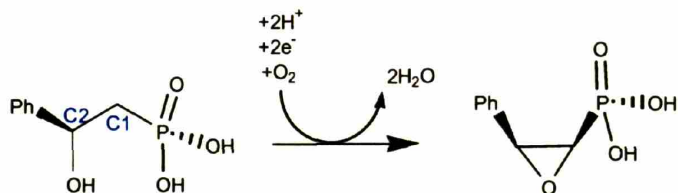
**(a)**



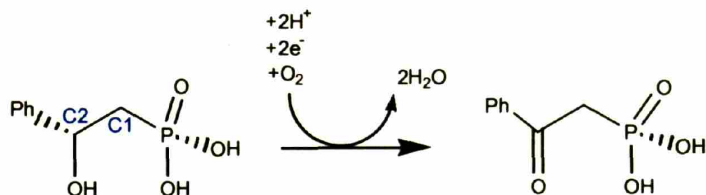
**(b)**



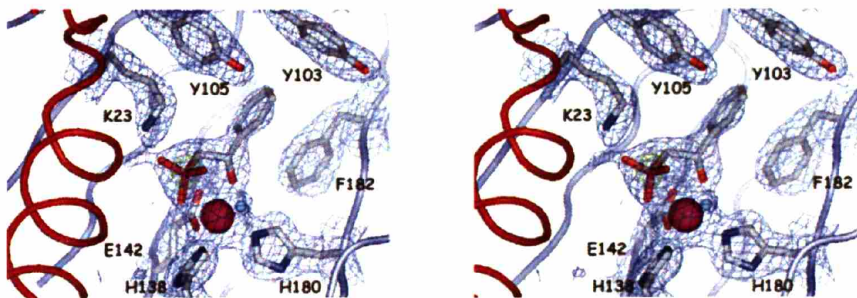
**(c)**



**(d)**



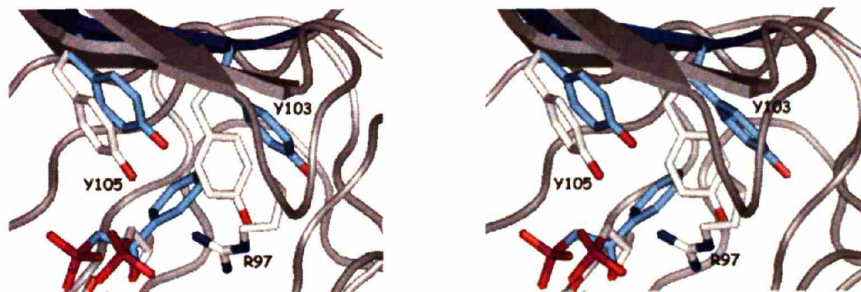
**Figure IV.2. The HppE active site displayed with 2Fo-Fc maps contoured at 1.0 $\sigma$ .** The *S*-HPEP:Co(II)-HppE structure is shown with  $\beta$ -strands shown as slate tubes ( $\beta$ -domain) and the  $\alpha$ -domain is shown as red tubes. *S*-HPEP, F182, facial triad residues (E142, H138, and H180), and cantilever hairpin residues Y103 and Y105 are shown in ball-and-stick (carbon (grey), oxygen (red), nitrogen (blue), phosphorus (magenta)), a water molecule as a sphere (marine), and the Co(II) metal also as a sphere (magenta).







**Figure IV.4. Alignment of cantilever hairpin in *S*-HPP:Co(II)-HppE and *S*-HPEP:Co(II)-HppE structures.** Select residues and substrate, or substrate analog, are shown in ball-in-stick representation (oxygen (red), nitrogen (blue), and phosphorus (magenta)); carbon in *S*-HPP:Co(II)-HppE (light grey) carbon in *S*-HPEP:Co(II)-HppE (cyan). The cantilever hairpin is shown with arrows. (grey; *S*-HPP:Co(II)-HppE and blue; *S*-HPEP:Co(II)-HppE.)

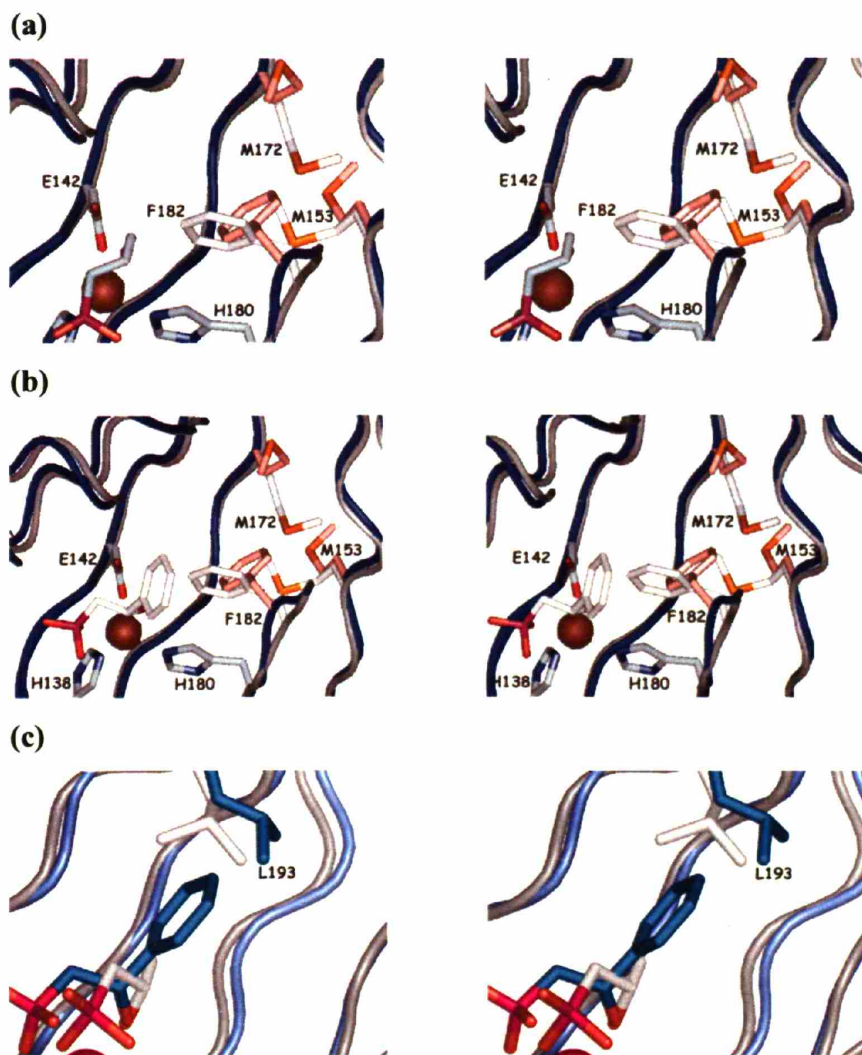


**Figure IV.5. Cooperative movement of hydrophobic pocket residues upon substrate binding.**

**(a)** Alignment of Fe(II)-HppE (grey loops) and *S*-HPP:Fe(II)-HppE (blue loops) structures. F182, M153, and M172 are shown in ball-in-stick representation (carbon is light grey in Fe(II)-HppE, salmon in *S*-HPP:Fe(II)-HppE), and sulfur is orange. Triad residues (E142, H138, and H180) and *S*-HPP are also shown in ball-in-stick representation (carbon (grey), oxygen (red), nitrogen (blue), and phosphorus (magenta)).

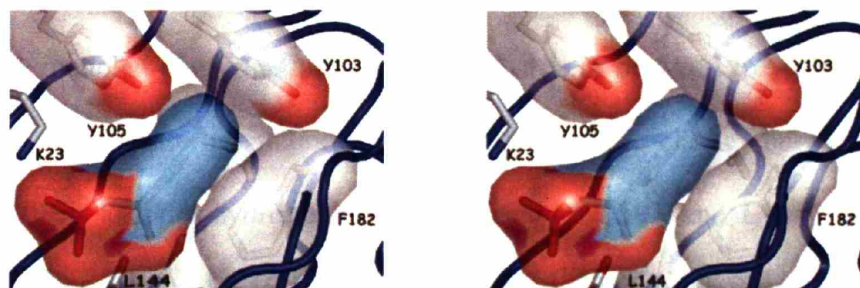
**(b)** Alignment of Fe(II)-HppE (grey loops) and *S*-HPEP:Co(II)-HppE (blue loops) structures. Colored as in (a).

**(c)** Alignment of *S*-HPP:Co(II)-HppE and *S*-HPEP:Co(II)-HppE structures. L193, *S*-HPP, and *S*-HPEP are shown in ball-in-stick representation (oxygen (red), nitrogen (blue), phosphorus (magenta), carbon in *S*-HPP (light grey), and carbon in *S*-HPEP (cyan)). Both *S*-HPEP-Co(II)-HppE and *S*-HPP:Co(II)-HppE structures are shown (purple and grey loops, respectively).

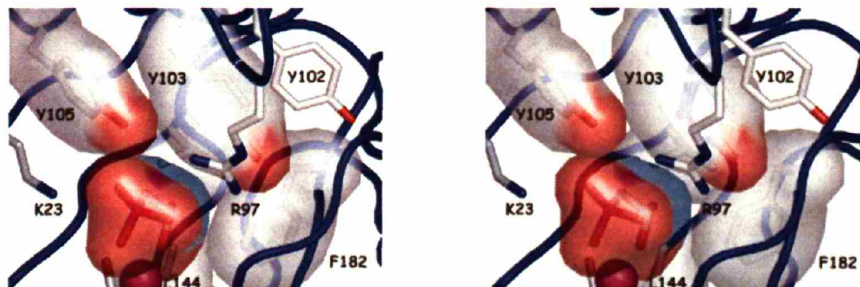


**Figure IV.6. Van der Waals surface representations *S*-HPP and *S*-HPEP bidentate binding modes.** Van der Waals surface representation of the (a) *S*-HPEP:Co(II)-HppE complex and the (b) *S*-HPP:Co(II)-HppE complex. Residues R97, Y102, Y103, and Y105 are from the cantilever hairpin. The protein backbone is shown as blue loops. The hydrophobic portion of substrate, or substrate analog, is shown in cyan; the hydrophilic portion of substrate, or substrate analog, is shown in red. L193 (top) is positioned directly behind the C2-phenyl substituent of *S*-HPEP and is not visible in this orientation. All residues are shown in ball-in-stick representation (carbon (grey), oxygen (red), and nitrogen (blue)). Both (a) and (b) are in approximately the same orientations.

(a)

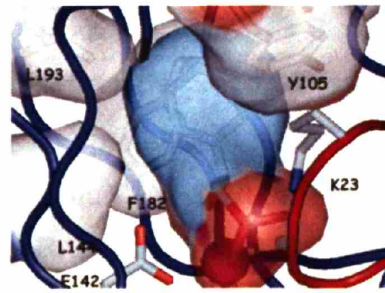
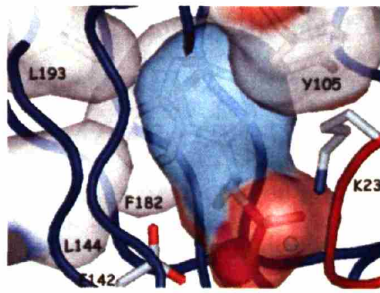


(b)

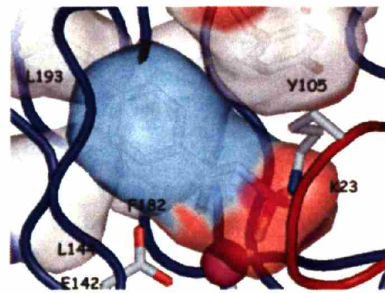
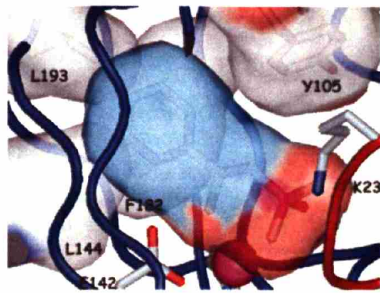


**Figure IV.7. Observed and hypothetical *S*-HPEP binding modes within the HppE hydrophobic pocket.** (a) The conformation of *S*-HPEP observed in the X-ray crystal structure with Van der Waal surfaces. The protein backbone is shown as blue loops. The hydrophobic portion of *S*-HPEP is shown in cyan; the hydrophilic portion of substrate, or substrate analog, is shown in red. All residues are shown in ball-in-stick representation (carbon (grey), oxygen (red), and nitrogen (blue)). (a)-(d) are in the same orientations. (b) A modeled conformation of *S*-HPEP where the phosphonate, C1 atom, and C2 substituents are aligned as *S*-HPP bound in the *S*-HPP-Co(II)-HppE crystal structure. (c) A modeled conformation of *S*-HPEP that is in between the conformation in (a) and (b). This substrate orientation should minimize steric clashes between the C2 phenyl of *S*-HPEP and the hydrophobic pocket residues L193 and L144 (d) The same modeled conformation of *S*-HPEP as in (c) shown against the enzyme with the cantilever hairpin in the closed position, as it is in the *S*-HPP-Co(II)-HppE structure.

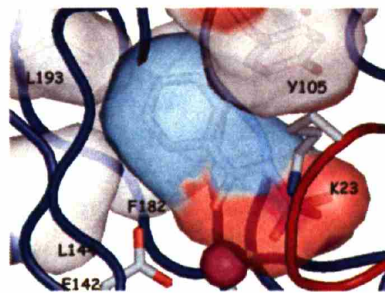
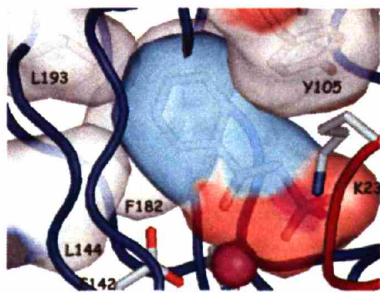
(a)



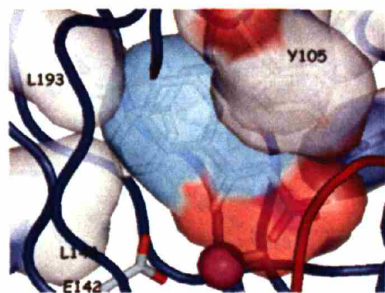
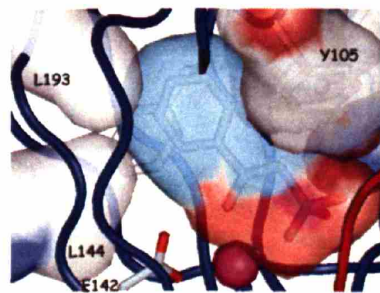
(b)



(c)

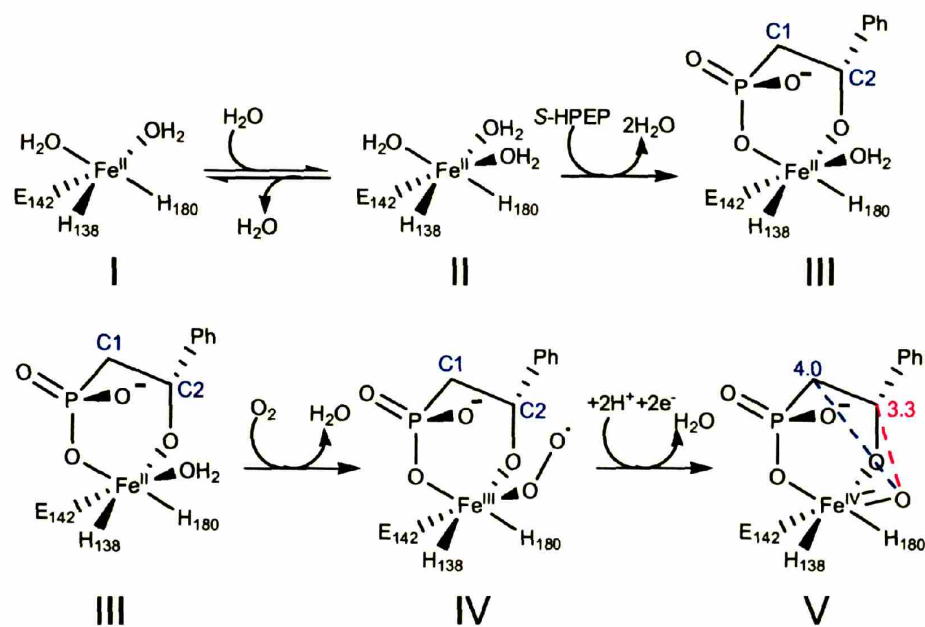


(d)

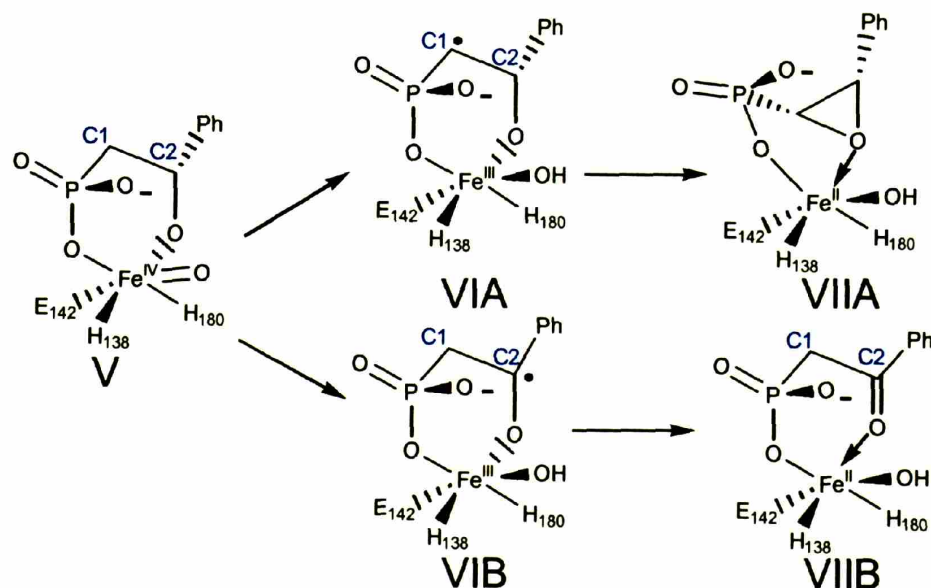


**Figure IV.8. S-HPEP observed binding mode is inconsistent with the known *in vitro* reaction. (a) Formation of the putative ferryl-oxo intermediate with dioxygen binding *trans* to triad residue E142. Distances are estimated from the proposed dioxygen binding site *trans* to E142 to the C1 (blue) and C2 (magenta) positions of S-HPEP are shown **V**. (b) Predicted results of hydrogen atom abstraction from C1 and C2: pathway **A**, hydrogen atom abstraction at the C1-position, and pathway **B**, hydrogen atom abstraction from the C2-position. C1 and C2 positions of S-HPEP are labeled (blue).**

(a)

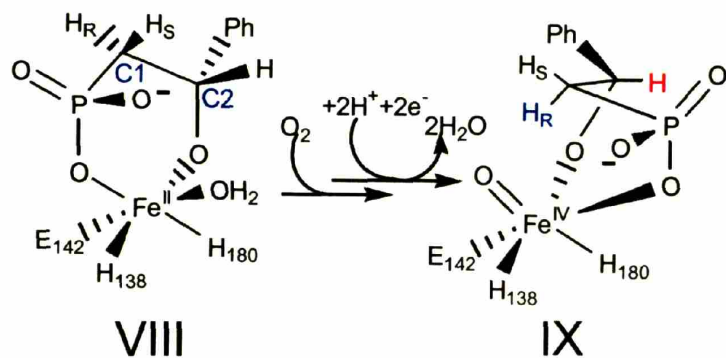


(b)



**Figure IV.9. Ligand rearrangement mechanism.**

The C2 pro-R hydrogen atom would be accessible to hydrogen atom abstraction (blue) and the C2 hydrogen atom (red) would be inaccessible in this hypothetical *S*-HPEP binding mode, IX. C1 and C2 positions of *S*-HPEP are labeled (blue).



#### IV.G. References

1. Liu, P., Murakami, K, Seki, T, He, X, Yeung, SM, Kuzuyama, T, Seto, H & Liu, H. Protein purification and function assignment of the epoxidase catalyzing the formation of fosfomycin. *J Am Chem Soc* **123**, 4619-4620 (2001).
2. Higgins, L. J., Yan, F., Liu, P., Liu, H. W. & Drennan, C. L. Structural insight into antibiotic fosfomycin biosynthesis by a mononuclear iron enzyme. *Nature* (2005).
3. Dunwell, J. M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**, 7-17 (2004).
4. Dunwell, J.M. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol Genet Eng Rev* **15**, 1-32 (1998).
5. Dunwell, J.M., Culham, A., Carter, C.E., Sosa-Aguirre, C.R. & Goodenough, P.W. Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* **26**, 740-746 (2001).
6. Zhao, Z., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Mechanistic studies of HPP epoxidase: configuration of the substrate governs its enzymatic fate. *Angew Chem Int Ed Engl* **41**, 4529-4532 (2002).
7. Liu, P. in *Department of Chemistry* 229 (University of Minnesota, 2001).
8. Liu, P., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Biochemical and spectroscopic studies on (S)-2-hydroxypropylphosphonic acid epoxidase: a novel mononuclear non-heme iron enzyme. *Biochemistry* **42**, 11577-11586 (2003).



9. Otwinowski Z, M. W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326 (1997).
10. Brunger, A. T., Adams, PD, Clore, GM, DeLano, WL, Gros, P, Grosse-Kunstleve, RW, Jiang, JS, Kuszewski, J, Nilges, M, Pannu, NS, Read, RJ, Rice, LM, Simonson, T & Warren, GL. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D Biol Crystallogr* **54 ( Pt 5)**, 905-921 (1998).
11. Kleywegt, G. Hetero-compound Information Centre - Uppsala (HIC-Up). (2005).
12. Laskowski RA, M. M., Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Cryst.* **26**, 283-291 (1993).
13. Sundberg, F.A.C.a.R.J. Advanced Organic Chemistry Third Edition: Part A. Vol. 3 651-719 (Plenum Press, New York, 1990).
14. Zhang, Z., Ren, J, Harlos, K, McKinnon, CH, Clifton, IJ & Schofield, CJ. Crystal structure of a clavaminic synthase-Fe(II)-2-oxoglutarate-substrate-NO complex: evidence for metal centered rearrangements. *FEBS Lett* **517**, 7-12 (2002).
15. Townsend, C.A. New reactions in clavulanic acid biosynthesis. *Curr Opin Chem Biol* **6**, 583-589 (2002).

## **Chapter V: Hydroxypropylphosphonic Acid Epoxidase: A Cupin Enzyme Like No Other?**

### **Summary**

Crystallographic analysis of hydroxypropylphosphonic acid epoxidase (HppE) from *Streptomyces wedmorensis* revealed a unique mononuclear iron enzyme defining a new sub-family of cupin enzymes (Chapter 2). Here we provide a comparative structural analysis of HppE relative to other cupin enzymes. We find that HppE is structurally more similar to dicupin proteins like conavalin and germin oxalate oxidase than monocupin mononuclear iron enzymes like taurine 2-oxoglutarate dioxygenase and isopenicillin N synthase, and may provide an evolutionary link between these two major classes of cupin proteins.

## V.A. Introduction

HppE was previously determined to be a member of the cupin protein superfamily<sup>1</sup>, which are defined by two primary sequence motifs, cupin motif 1 and 2 (Table V.1.) separated by an intermotif region (IMR) that is both variable in sequence and length. From sequence alignments alone it is difficult to unambiguously assign a given protein to this superfamily due to a substantial variability in primary structure. As shown in Table V.2., a large subset of cupin proteins require a metal for catalysis. For enzymes that do bind an active site metal, like HppE, sequence alignments are more useful for confirming the identity of a putative cupin protein. Specifically, cupin motif 1 houses two to three of the protein ligands that coordinate the active site metal (G(X)<sub>5</sub>HXHX(X)<sub>1-4</sub>(E/D)(X)<sub>6</sub>G) while cupin motif-2 (G(X)<sub>5</sub>PXG(X)<sub>2</sub>H(X)<sub>3</sub>N) provides a third or fourth protein ligand. Cupin motif 1 in HppE consists of facial triad ligands H138 and E142, and cupin motif 2 includes triad ligand H180 (Table V.1. and Figure V.1.). With respect to tertiary structure, all cupin proteins have a conserved antiparallel  $\beta$ -barrel motif that is a Greek key variant of the jelly-roll fold. The ligand triad is housed within this cupin scaffold, at one end of the  $\beta$ -barrel<sup>2-5</sup>.

Cupin proteins have a diverse array of biological functions such as DNA repair, antibiotic biosynthesis, and nutrient storage in plants. X-ray crystallographic analysis of cupin proteins, including cupin type phosphoglucose isomerase (PGI), canavalin, germin oxalate oxidase (GOO), 1-aminocyclopropane-1-carboxylic acid oxidase (ACCO), deacetoxycephalosporin C synthase (DAOCS), taurine:alpha-ketoglutarate dioxygenase (TauD), isopenicillin N synthase (IPNS), has provided a wealth of information about the cupin superfamily. PGI is a metal dependent (Fe(II) or Ni(II)) enzyme that facilitates the

isomerization of glucose-6-phosphate and fructose-6-phosphate in glycolysis and gluconeogenesis. Canavalin, GOO, and ACCO are important in plant biochemistry. Canavalin is a plant seed storage protein with no known catalytic function. GOO is a Mn(II) dependent enzyme that catalyzes oxalate oxidation to yield carbon dioxide. ACCO catalyzes the biosynthesis of the plant signaling molecule ethylene and is a mononuclear iron enzyme. Mononuclear iron enzymes are an important subfamily of cupin proteins and include DAOCS, TauD, and IPNS. DAOCS catalyzes the biosynthesis of cephalosporin antibiotics, and TauD catalyzes the oxidation of taurine to sulfite and aminoacetaldehyde providing a sulfur source for bacteria. Both DAOCS and TauD are dependent of  $\alpha$ -ketoglutarate ( $\alpha$ -KG) as a co-substrate, a criterion that defines a large subset of cupin mononuclear iron enzymes. In contrast, IPNS is an  $\alpha$ -KG independent cupin mononuclear iron enzyme, an example of a smaller subset of cupin enzymes. Cupin mononuclear iron enzymes require a high energy (i.e. high reduction potential) iron-oxo intermediate to facilitate catalysis. Substrate binding triggers a conformational change in both IPNS and TauD that serves to close the active site, reducing solvent accessibility, thus protecting the high energy intermediates that arise during catalysis.

There are two main classes of cupin proteins: monocupins and dicupins<sup>3</sup>. A majority of characterized cupin proteins are monocupins, in that a single polypeptide chain consists of a single cupin domain (i.e. two cupin motifs + IMR). Dicupins, such as conavalin, have two cupin domains within a single polypeptide which is thought to be the result of gene duplication (Figure V.2.). Dicupins consist of a conserved  $\beta$ -dimer governed primarily by hydrophobic interactions between juxtaposed cupin domains. Interestingly, the quaternary structures of several monocupin proteins are similar to

dicupin proteins. Specifically, the  $\beta$ -dimer of cupin domains seen in dicupins (e.g. canavalin; Figure V.2.) is also observed in monocupin proteins like wheat protein GOO, once referred to as the archetypical cupin. The dimer interface of the former is an intramolecular interaction and in the latter an intermolecular interaction.

From a structural perspective it is less useful to characterize cupin proteins as monocupins (one cupin motif per polypeptide chain) or dicupins (two cupin motifs per polypeptide chain), and more appropriate to classify them based on the presence or absence of a  $\beta$ -dimer. Herein these two structural subclasses will be referred to as structural dicupins and monocupins, respectively. As shown in Table V.2., the latter class of cupins includes acireductone dioxygenase (ARD), clavamate synthase (CAS), deacetoxycephalosporin synthase (DAOCS), and isopenicillin N synthase (IPNS) and the former includes canavalin, GOO, homogentisate dioxygenase (HGO), and quercetin dioxygenase (QCD). In this chapter, the HppE structures will be compared and contrasted to other cupin proteins and non-heme mononuclear iron proteins within the superfamily.

## V.B. Methods

All structural alignments were done using LSQMAN<sup>6</sup> via the brute force algorithm. For alignments of the  $\beta$ -domain of HppE, the tetragonal Fe(II)-HppE model was used (residues 96-198, molecule C). For  $\beta$ -dimer alignments, the tetragonal Fe(II)-HppE model was used (residues 96-198, molecules A and B). The following cupin protein models were used (PDB codes are in parentheses): ARD (1M4O)<sup>7</sup>, CAS (1DRT)<sup>8</sup>, DAOCS (1DCS)<sup>9</sup>, TauD (1GQW)<sup>10</sup>, canavalin (1CAV)<sup>11</sup>, GOO protein (1FI2)<sup>12</sup>, HGO (1EY2)<sup>13</sup>, QCD (1H1M)<sup>14</sup>, and PGI (1J3R). RMSD values (Table V.3.a-b.) were

calculated in LSQMAN. The ACCO structure was omitted from these alignments since no PDB file is available although a structure has been published<sup>15</sup>. The primary sequence alignment of HppE from *Streptomyces wedmorensis* with HppE from *Pseudomonas syringae* was obtained using ClustalW<sup>16</sup>.

## V.C. Discussion

### V.C.1. HppE: A structural dicupin

The fosfomicin biosynthetic enzyme hydroxypropylphosphonic acid epoxidase (HppE) is a mononuclear iron enzyme with a unique monomeric architecture and quaternary structure<sup>1,17</sup>. The HppE monomer consists of two domains, an  $\alpha$ -domain and a  $\beta$ -domain (Figure V.3.a.). The former is a unique tertiary structure and consists of five helices (all  $\alpha$ -helical)<sup>18</sup>. The latter consists of eleven  $\beta$ -strands arranged in a jelly-roll variant of a  $\beta$ -barrel. The domains are separated by a 32-residue linker. The HppE monomeric architecture translates to a unique quaternary structure. Specifically, strands 4,6,9, and 11 form one sheet within the  $\beta$ -barrel, and stack directly against a symmetry equivalent sheet in a juxtaposed  $\beta$ -barrel forming the HppE  $\beta$ -dimer (Figure V.3.b.). This apposition of  $\beta$ -domains is governed mainly by hydrophobic interactions. In addition, strand 1 of one molecule crosses the  $\beta$ -dimer interface to complete the sheet of the C2 symmetry related molecule. This intertwining of strands further stabilizes the HppE  $\beta$ -dimer (Figure V.3.b.). In order to “build” the HppE physiological tetramer, two HppE  $\beta$ -dimers (two HppE monomers) associate to form an  $\alpha$ -dimer interface (four HppE monomers). This quaternary arrangement orients the  $\alpha$ -domain of one HppE monomer

above the barrel of a second HppE monomer, forming a unique domain interface (Figure V.3.c-d) (Chapter 2).

Comparative analyses of the  $\beta$ -domains and  $\beta$ -dimers within cupin proteins reveal striking similarities with that of HppE. Similar to canavalin<sup>11</sup>, GOO<sup>12</sup>, and PGI<sup>19</sup>, the HppE  $\beta$ -domain contains only  $\beta$ -strands with loops connecting each secondary structural element (Figure V.1.). Both HppE and GOO have non- $\beta$ -sheet regions, but these regions are separated from the  $\beta$ -domain by a variable linker (Figure IV.4.a.). This architecture is in contrast to many mononuclear iron enzymes that maintain the same basic jelly-roll topology but have helices and other structural units that lie between the  $\beta$ -strands of the barrel (Figure V.5.)<sup>3</sup>. HppE shares other features with dicupin proteins including its  $\beta$ -dimer (i.e. the apposition of two HppE  $\beta$ -domains) and the fact that the  $\beta$ -dimers are intertwined such that strand 1 of the  $\beta$ -domain from one molecule completes the sheet of a juxtaposed molecule at the dimer interface (Figure V.3.b and Figure V.4.b.). The modularization of secondary structural elements into domains as found in HppE and GOO allows for the creation of interesting and unique quaternary structures (Figure V.6. and Figure V.3.c-d.). For example, the intertwining of the HppE monomeric units creates a multi-layer structure with four active sites at the interface of the  $\alpha$ - and  $\beta$ -domains. In contrast, the GOO monomeric architecture results in a planar structure with a more open active site relative to that of HppE (Figure V.6.). As discussed in detail in Chapter 2, the orientation of the HppE  $\alpha$ -helical domain relative to the HppE  $\beta$ -domain has catalytic significance in that the  $\alpha$ -domain serves to partially occlude the HppE active site, a common theme with respect to protection of high energy iron-oxo intermediates arising during the HppE catalytic cycle (chapter 2), a function not needed in GOO.

### *V.C.2. Cupin Motifs*

HppE differs from most other mononuclear iron enzymes that have been studied crystallographically (e.g. DAOCS<sup>9,20</sup>, IPNS<sup>21</sup>, TauD<sup>10</sup>, and ACCO<sup>15</sup>) with respect to the position of facial triad ligands within the primary, secondary, and tertiary structure. The primary sequence defining motif 1 triad ligands in HppE is <sup>138</sup>HXXXE<sup>142</sup>, in contrast to a majority of mononuclear iron enzymes which have a HX(E/D) sequence<sup>22</sup>. In the HppE crystal structures, E142 lies on the N-terminal end of strand 6 and His 138 on a loop between strands 5 and 6 (Figure V.1.b.). While in mononuclear iron enzymes DAOCS, IPNS, TauD, and ACCO, triad residues in cupin motif 1 lie on a loop between strands 5 and 6 (HppE numbering) (Figure V.7.a.). Interestingly, His 138 of HppE occupies the same space within the cupin fold as the Asp or Glu triad residue of all structurally characterized monocupin enzymes and Asn 135 of HppE is coincident in space with the motif 1 triad His of monocupin enzymes (Figure V.7.a.). In contrast, the position of HppE triad ligands within the cupin fold is identical to that observed in the dicupin subclass of mononuclear iron proteins that bind a metal, regardless of the metals identity (Figure V.7.b.). Based on this comparative analysis, it is tempting to speculate that HppE evolved from a dicupin progenitor. One can envision an evolutionary progression from a conavalin-like progenitor, to a GOO-like enzyme, to HppE. Conavalin is a plant storage protein that is devoid of any known catalytic function and does not bind a metal (Figure 2). GOO catalyzes the oxidation of oxalate to two equivalents of carbon dioxide, a reaction that is dependent upon Mn<sup>II</sup>. Interestingly, the monomeric and quaternary architecture of the GOO protein and canavalin are quite similar (Figure V.2. and Figure V.6.). Nature designed the GOO protein to bind manganese, converting an otherwise inert



protein into a viable bioinorganic catalyst. Similar to PGI, HppE is both designed to bind a required metal cofactor (i.e. iron) and is “tuned” to recognize an organic substrate.

PGI is a dicupin mononuclear iron enzyme, the only dicupin enzyme with the exception of HppE that is iron dependent. It has a unique motif 1 ligand spacing for a cupin protein, H(X)<sub>6</sub>E, although in three dimensions these ligands are coincident with the position of HppE motif 1 triad ligands, HXXXE (Figure V.7.c.). Like some other cupin proteins, PGI cupin motif 1 has an additional histidine ligand available for metal coordination, HXH(X)<sub>6</sub>E. The position of this ligand is coincident with that of N135 of HppE and the motif 1 E/D ligand in monocupin enzymes, HX(D/E) (e.g. TauD) (Figure V.7.). Thus four ligands of PGI represent the composite of HppE and TauD ligands.

### *V.C.3. Conformational changes and active site protection: TauD, IPNS, HppE*

Like the HppE story, studies of TauD are a good example of how biochemical studies coupled with structural analysis can synergistically improve our understanding of enzyme mechanisms<sup>10,23-27</sup>. TauD catalyzes the hydroxylation of taurine in an  $\alpha$ -KG dependent fashion. It has been shown that the  $K_D$  of  $\alpha$ -KG decreases considerably in the presence of excess taurine (by a factor of ~60) and the oxidative decarboxylation of  $\alpha$ -KG occurs at a much greater rate when substrate is present<sup>27</sup>. Therefore, TauD is specifically “tuned” to couple  $\alpha$ -KG oxidative decarboxylation and taurine hydroxylation. X-ray crystal structures of apo-TauD, an Fe(II): $\alpha$ -KG:TauD ternary complex, and a taurine:Fe(II): $\alpha$ -KG:TauD quaternary complex provide insight into the TauD mechanism<sup>10,24</sup>. Specifically, the apo-TauD structure and the ternary complex structures are close to identical. In the ternary complex structure,  $\alpha$ -KG binds in a bidentate fashion

to the ferrous center, as in all known  $\alpha$ -KG dependent mononuclear iron enzymes, and the distal carboxylate, relative to the metal center, is stabilized via charge-charge interactions with the guanidinium moiety of R266<sup>24</sup>. This arginine is conserved in many  $\alpha$ -KG dependent mononuclear iron enzymes within the so-called RXS motif<sup>22</sup>. Furthermore, the TauD ternary complex structure (i.e.  $\alpha$ -KG alone) has a relatively open active site and is fairly solvent accessible. Interestingly, when substrate binds to the ternary complex or to apo-TauD, a conformational change occurs in a helical hairpin positioned above the cupin barrel, and in a loop positioned along the axis of the barrel which serves to occlude the TauD active site, reducing solvent accessibility<sup>24</sup> (Figure V.8.a.). The local metal environment is perturbed favoring a change in coordination from six to five via the loss of a water molecule<sup>24</sup>. Diffusion of dioxygen into the TauD active site is still possible in this closed down conformation<sup>24</sup> and upon binding to the open sixth coordination site catalysis is initiated.

With respect to penicillin biosynthesis catalyzed by IPNS, substrate binding also serves as an enzyme conformational trigger. In the IPNS crystal structure, a C-terminal helix extends from the final strand of the cupin  $\beta$ -barrel (strand 11 by HppE numbering) and serves to partially occlude the IPNS active site (Figure V.8.b.). A C-terminal loop extends from this helix and is anchored to the metal center via direct coordination of Q330 to iron in a monodenate fashion. Substrate binding, via direct coordination of  $\delta$ -(1- $\alpha$ -aminoadipoyl)-l-cysteinyl-D-valine (ACV) to iron, triggers the release of Q330. After the release of Q330, the C-terminal helix is able to extend and further occlude the IPNS active site from solvent<sup>21,28,29</sup>. Subsequent binding of dioxygen is then required for catalytic initiation.

As discussed previously (chapter 2), in HppE, a hairpin comprised of strands 2 and 3 (residues 92-106) acts as a cantilever which occludes the active site, upon bidentate substrate (*S*-HPP) binding (Figure V.8.c.). Appropriate substrate size, charge, and orientation triggers this hairpin motion, sweeping through an angle of approximately 23°. This active site occlusion occurs at the same face of the cupin  $\beta$ -barrel as that of IPNS and TauD.

#### *V.C.4. The RXS motif and $\alpha$ -KG dependence*

Notably, HppE, ACCO, and IPNS are the only  $\alpha$ -KG independent cupin mononuclear iron enzymes whose crystal structures have been solved<sup>15,21</sup>. Unlike HppE, one could argue that both ACCO and IPNS derive from  $\alpha$ -KG dependent mononuclear iron enzymes. Although neither enzyme is “tuned” to bind the  $\alpha$ -KG molecule, both retain the conserved arginine residue within the RXS motif. IPNS could have divergently evolved utilizing this RXS motif to bind the distal carboxylate of its substrate ACV. Although structures do not exist of ACCO in complex with substrate or co-substrate, a conserved arginine residue found in its sequence could serve a similar role in substrate or co-substrate recognition. HppE does not have this RXS motif. The only arginine in the active site, R97 in strand 1 of the cantilever hairpin, is in a different position within the cupin fold. Similarly, GOO<sup>12</sup>, canavalin<sup>11</sup>, and PGI<sup>19</sup> are also devoid of the RXS motif. This provides further evidence that HppE is not derived from a monocupin progenitor and is more likely a modification of a dicupin protein.

#### V.C.5. The other HppE from *Pseudomonus psyringae*

There are only two known organisms that catalyze the biosynthesis of fosfomicin from *S*-HPP, *Pseudomonus psyringae* and *Streptomyces wedmorensis*. A *Pseudomonus psyringae* homologue of HppE<sup>30</sup> has been demonstrated *in vitro* to be analogous to that of HppE from *Streptomyces wedmorensis* (unpublished results; Professor Ben Liu Laboratory). The sequence homology between the two HppE isoforms is greater than 90% with a sequence identity of 35.5%. Comparing the two sequences (Figure V.9.), the facial triad is conserved consisting of motif 1 ligands, H138/E142\*, and motif 2 ligand, H180. Residues important for substrate binding and cantilever hairpin motion in HppE from *Streptomyces wedmorensis* are conserved including Y105, Y103, Y102, and R97 (chapter 2). Also F182, important for defining the HppE hydrophobic pocket, K23 and N197, important for stabilization of the water binding site *trans* to H180, and N135, proximal to the phosphonate substituent of substrate enantiomers *R*- and *S*-HPP in bidentate binding modes, are conserved. Given the high sequence homology and conservation of catalytically important residues, the results presented throughout this work may be extrapolated to the *Pseudomonus syringae* HppE enzyme.

\* Numbering follows the sequence from *Streptomyces wedmorensis* HppE

## V.D. Figures and Tables

**Table V.1. Alignments of cupin motifs 1 and 2.** The definition of cupin motifs 1 and 2 is shown with metal ligands in blue. Inconsistencies with the cupin motif definition are shown in red. The number of residues in between cupin motifs is shown (IMR). HppE from *Streptomyces wedmorensis* and *Pseudomonas psyringae* are listed as HppE S.w. and HppE P.s., respectively. Protein acronyms are defined in Table V.2.

	<b>Cupin Motif 1</b>	<b>IMR</b>	<b>Cupin Motif 2</b>
<b>Definition</b>	$G(X)_5HXH(X)_{1-4}(E/D)(X)_6G$	-	$G(X)_5PXG(X)_2H(X)_3N$
<b>HppE S.w.</b>	D (X) <sub>5</sub> SXH (X) <sub>3</sub> E (X) <sub>6</sub> G	19	G (X) <sub>5</sub> EXH (X) <sub>2</sub> H (X) <sub>3</sub> A
<b>HppE P.s.</b>	Q (X) <sub>5</sub> GXH (X) <sub>3</sub> E (X) <sub>6</sub> G	20	G (X) <sub>5</sub> LXN (X) <sub>2</sub> H (X) <sub>3</sub> N
<b>GOO</b>	G (X) <sub>5</sub> HXH (X) <sub>4</sub> E (X) <sub>6</sub> G	23	G (X) <sub>5</sub> PXG (X) <sub>2</sub> H (X) <sub>3</sub> N
<b>PGI</b>	F (X) <sub>5</sub> HXH (X) <sub>6</sub> E (X) <sub>6</sub> G	21	G (X) <sub>5</sub> PXY (X) <sub>2</sub> H (X) <sub>3</sub> N
<b>ARD</b>	E (X) <sub>5</sub> HXH (X) <sub>3</sub> E (X) <sub>6</sub> G	19	N (X) <sub>5</sub> PXH (X) <sub>2</sub> H (X) <sub>3</sub> M
<b>QCD</b>	A (X) <sub>5</sub> HXH (X) <sub>4</sub> E (X) <sub>6</sub> G	20	G (X) <sub>5</sub> PXN (X) <sub>2</sub> H (X) <sub>3</sub> I
<b>DAOCS</b>	E (X) <sub>5</sub> AXHXD (X) <sub>6</sub> I	40	L (X) <sub>5</sub> VXA (X) <sub>2</sub> H (X) <sub>3</sub> A
<b>TAUD</b>	N (X) <sub>5</sub> NXHXD (X) <sub>6</sub> P	135	D (X) <sub>5</sub> NXV (X) <sub>2</sub> H (X) <sub>2</sub> N
<b>CAS</b>	T (X) <sub>5</sub> EXHXE (X) <sub>6</sub> Q	113	G (X) <sub>6</sub> NXR (X) <sub>2</sub> H (X) <sub>3</sub> P

**Table V.2. Cupin structures solved and metal dependence.** Structural monocupins and structural dicupins are defined in the text.

Protein	Abbreviations	PDB code	Structural		Enzyme	Metal
			Monocupin	Dicupin		
Hydroxypropylphosphonic acid epoxidase	HppE	1ZZC		x	x	Fe
Acireductone dioxygenase	ARD	1M4O	x		x	Ni
Anthocyanidin synthase	ANS	1GP4	x		x	Fe
Auxin Binding Protein	-	1LR5		x		Zn
Canavalin	-	2CAV		x		None
Clavamate Synthase	CAS	1DS1	x		x	Fe
Deacetoxycephalosporin Synthase	DAOCS	1DCS	x		x	Fe
dTDP-dehydrorhamnose epimerase	-	1DZR			x	None
	-	1EP0		x	x	None
GAB protein	-	1JR7	x		x	Fe
GOO	GOO	1FI2		x	x	Mn
Homogentisate dioxygenase	HGO	1EYB		x	x	Fe
Isopenicillin N Synthase	IPNS	1BK0	x		x	Fe
Oxalate decarboxylase	-	1J58		x	x	Mn
Phaseolin	-	2PHL				None
Phosphoglucose isomerase	PGI	1PLZ		x	x	Fe?
Phosphomannose isomerase	PMI	1PMI	x		x	Zn
Progycinin	-	1FXZ				None
Proline-3-hydroxylase	-	1E5S			x	Fe
Quercetin dioxygenase	QCD	1GQG		x	x	Cu
Taurine-alpha-ketoglutarate dioxygenase	TauD	1OS7	x		x	Fe
Transcription factor	-	2ARC	x			None
Unknown protein	-	1LKN	x		?	?

**Table V.3. Comparison of the HppE  $\beta$ -domain and HppE  $\beta$ -dimer with cupin proteins. Root mean squared deviations are shown (Table V.2. for abbreviations).**

**(a)**

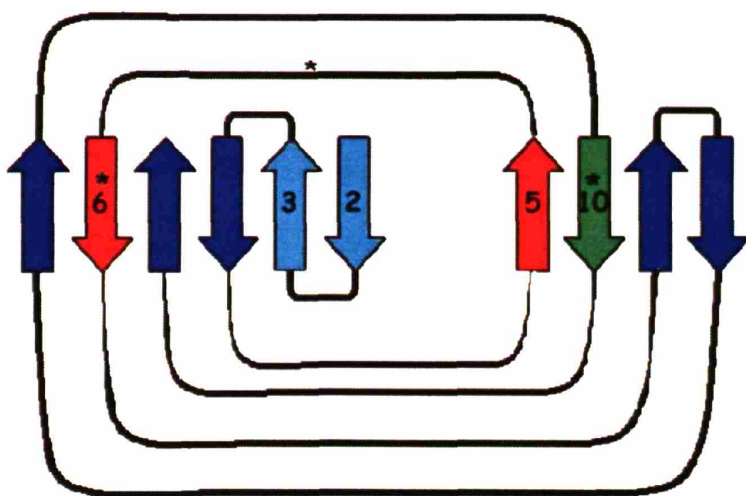
	<b>CAS</b>	<b>DAOCS</b>	<b>IPNS</b>	<b>TauD</b>
<b><math>\beta</math>-domain</b>	1.4	1.8	2	1.4

**(b)**

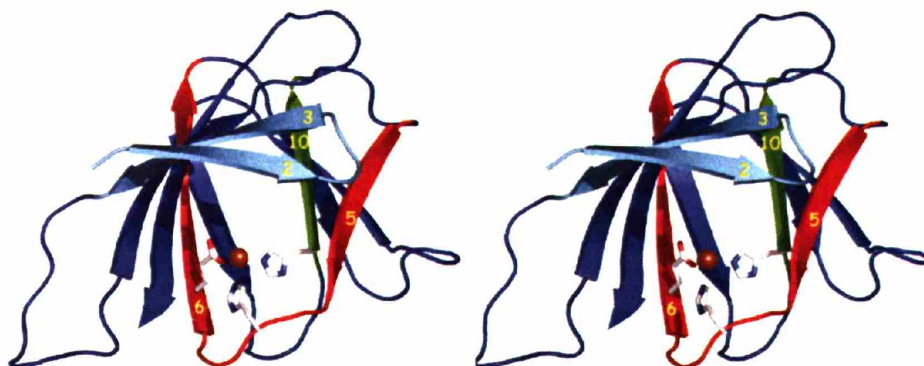
	<b>Canavalin</b>	<b>GOO</b>	<b>PGI</b>
<b><math>\beta</math>-dimer</b>	1.4	1.5	1.7

**Figure V.1. HppE  $\beta$ -domain and cupin motifs 1 and 2.** (a) Topology diagram of the HppE  $\beta$ -domain. Cupin motif 1 consists of  $\beta$ -strands 5 and 6 (red), cupin motif 2 consists of  $\beta$ -strand 10 (green), and the cantilever hairpin consists of  $\beta$ -strands 2 and 3 (cyan). The relative position of facial triad ligands is shown with black asterisks. (b) The HppE  $\beta$ -domain shown in ribbon representation (stereo) colored as in (a). Facial triad ligands are shown in ball-in-stick representation: carbon (grey), oxygen (red), and nitrogen (blue). Facial triad ligands include: E142 (strand 6), H138 (5/6 inter-strand loop), and H180 (strand 10).

(a)

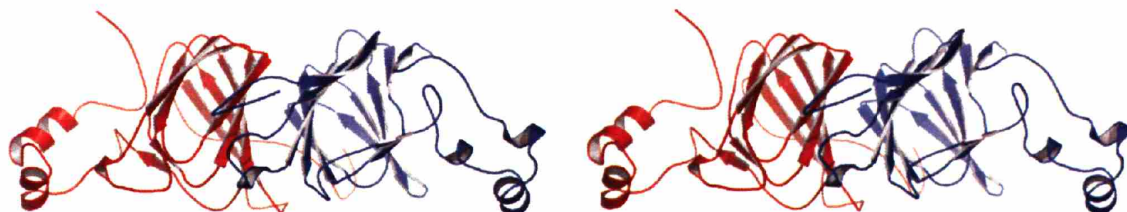


(b)



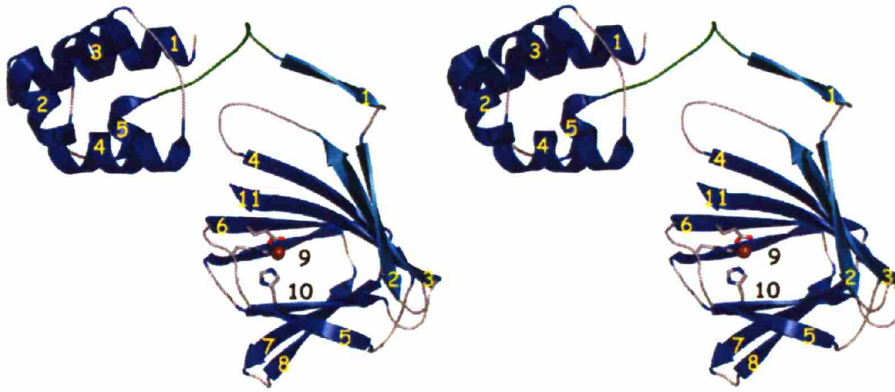


**Figure V.2. Canavalin dicupin structure.** A single polypeptide chain of the dicupin conavalin is shown in ribbon representation colored by domain (red and blue).



**Figure V.3. Tertiary and quaternary structure of HppE.** (a) The HppE monomer is shown in ribbon representation (stereo). The  $\alpha$ -domain (all  $\alpha$ -helical) and  $\beta$ -domain (all  $\beta$ -strands) is colored blue, numbered separately and sequentially by secondary structural element. Cross-interface strand 1 and the cantilever hairpin (strands 2 and 3) are colored cyan. The inter-domain linker is shown as a green loop. Facial triad ligands (E142, H138, and H180) are shown in ball and stick representation: carbon (grey), oxygen (red), and nitrogen (blue). Iron is represented as a brown sphere. (b) The HppE  $\beta$ -dimer is shown in ribbon representation (stereo). Strands are colored blue and the cross-interface strand (strand 1) is colored yellow. (c) The HppE tetramer is shown down a C2 symmetry axis in ribbon representation, colored by molecule: molecule 1 (blue, cantilever hairpin/strand 1 in cyan), molecule 2 (red, cantilever hairpin/strand 1 in red), molecule 3 (yellow, cantilever hairpin/strand 1 in dark yellow), and molecule 4 (green, cantilever hairpin/strand 1 in dark green). (d) The HppE tetramer in an orientation perpendicular to and colored as in (c).

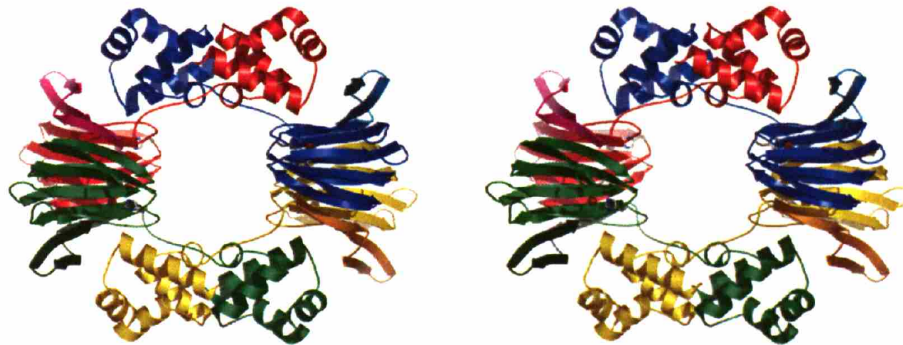
(a)



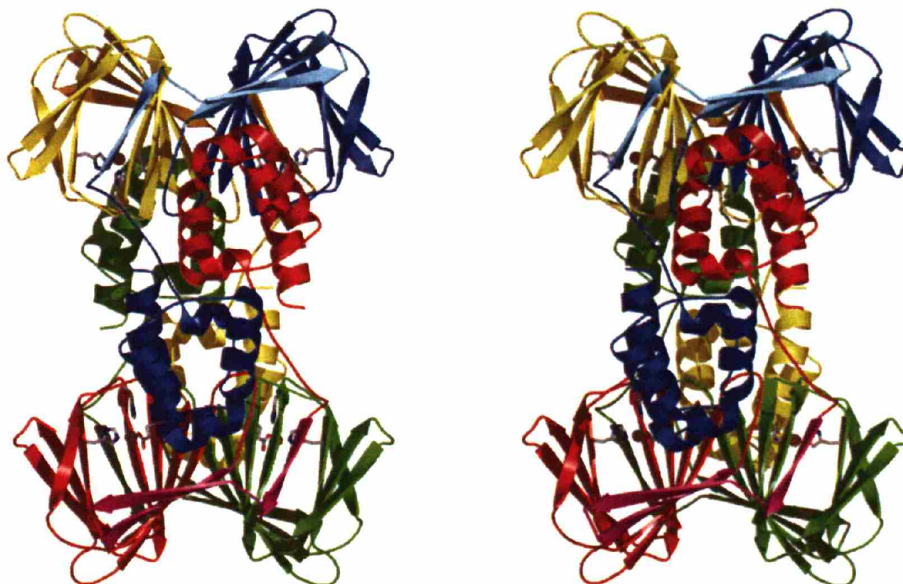
(b)



(c)

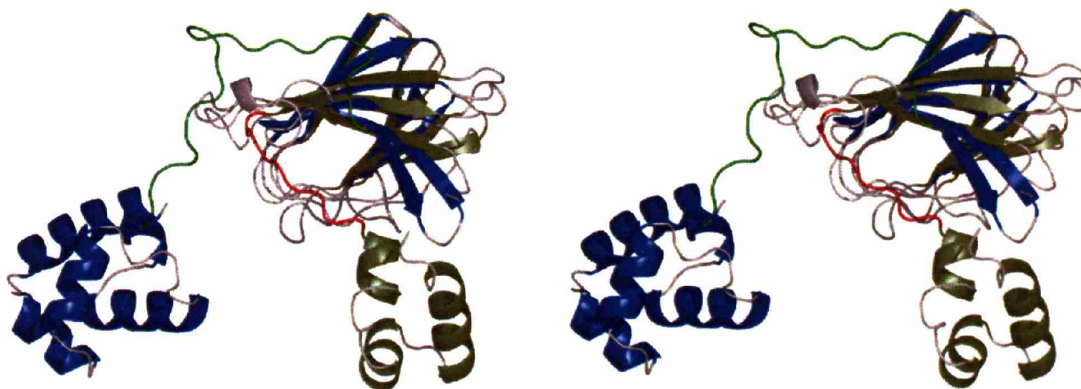


(d)

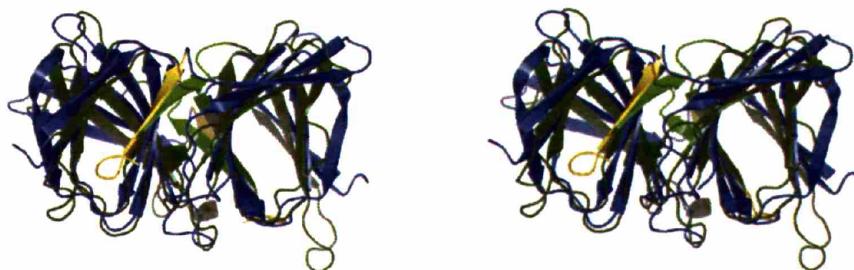


**Figure V.4. Dicupin structural alignment.** (a) The HppE (blue) and GOO (olive) monomers are aligned (Table V.3. RMSD value) shown in ribbon representation (stereo). The HppE and GOO inter-domain linkers are shown in green and red, respectively. (b) Structural alignment of the HppE  $\beta$ -dimer (blue) and GOO dimer (olive) are shown in ribbon representation. The cross-interface strands are shown in yellow and green, respectively.

(a)

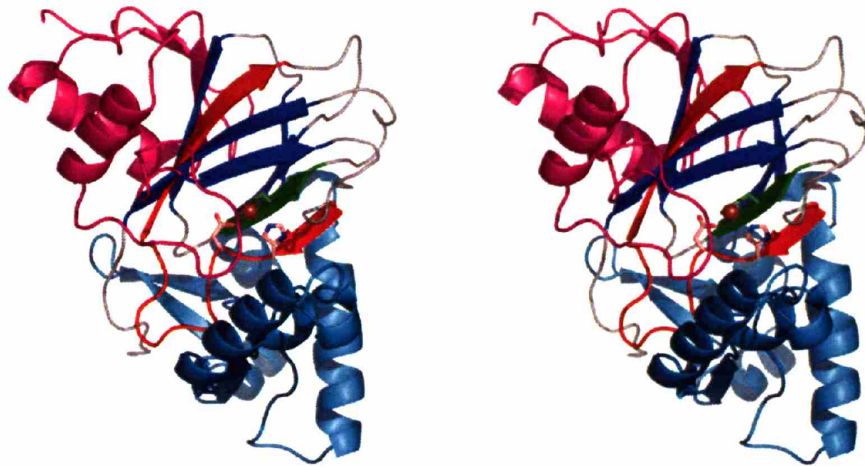


(b)

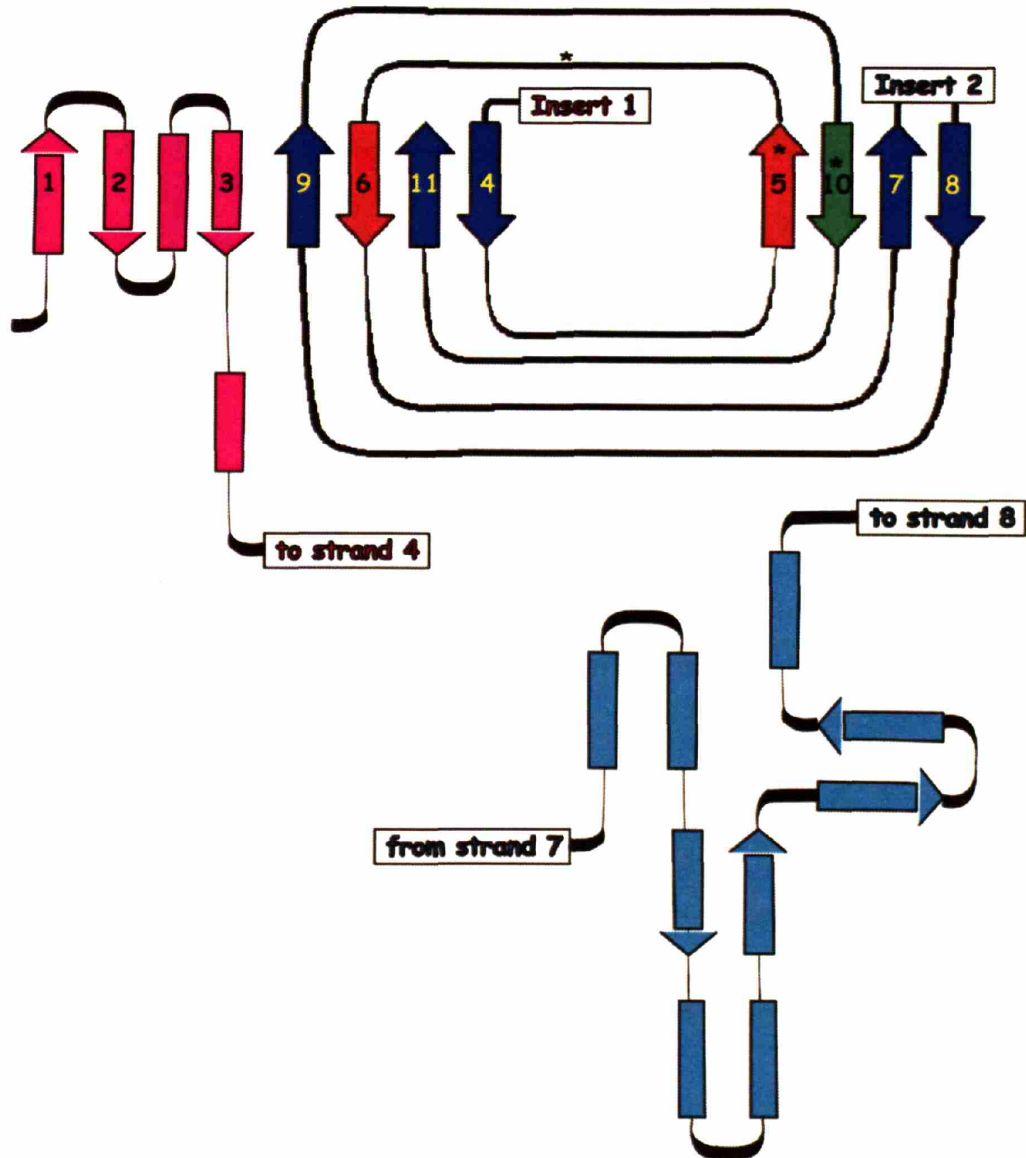


**Figure V.5. The monocupin TauD.** (a) The TauD monomer is shown in ribbon representation with cupin strands 5 and 6 in red, cupin strand 10 in green, and remaining strands defining the jelly-roll fold in blue (strands 4,7-9, and 11). Insert 1 is shown in magenta and insert 2 is shown in cyan (b) Topology diagram of the TauD colored as in (a). Strands are numbered as in Figure V.3.a. Asterisks indicate the positions of the facial triad ligands.

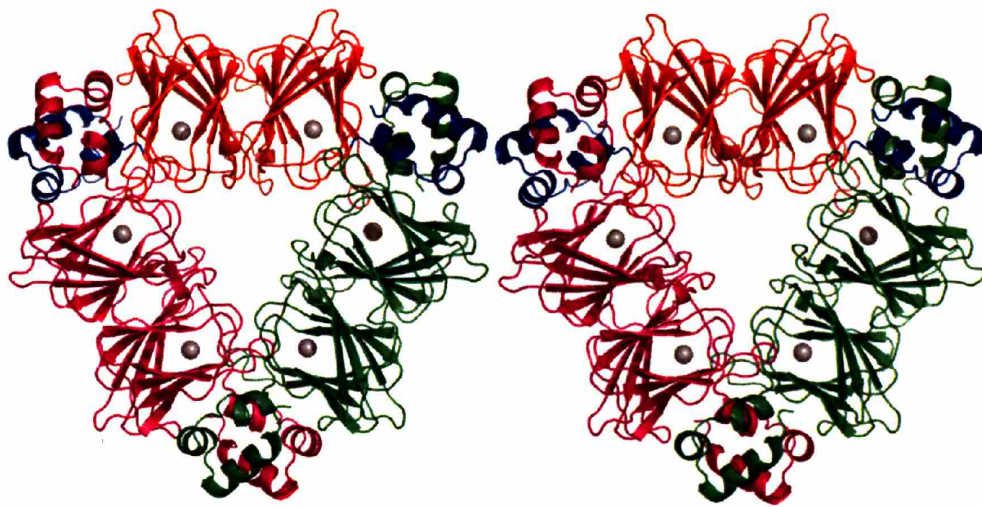
(a)



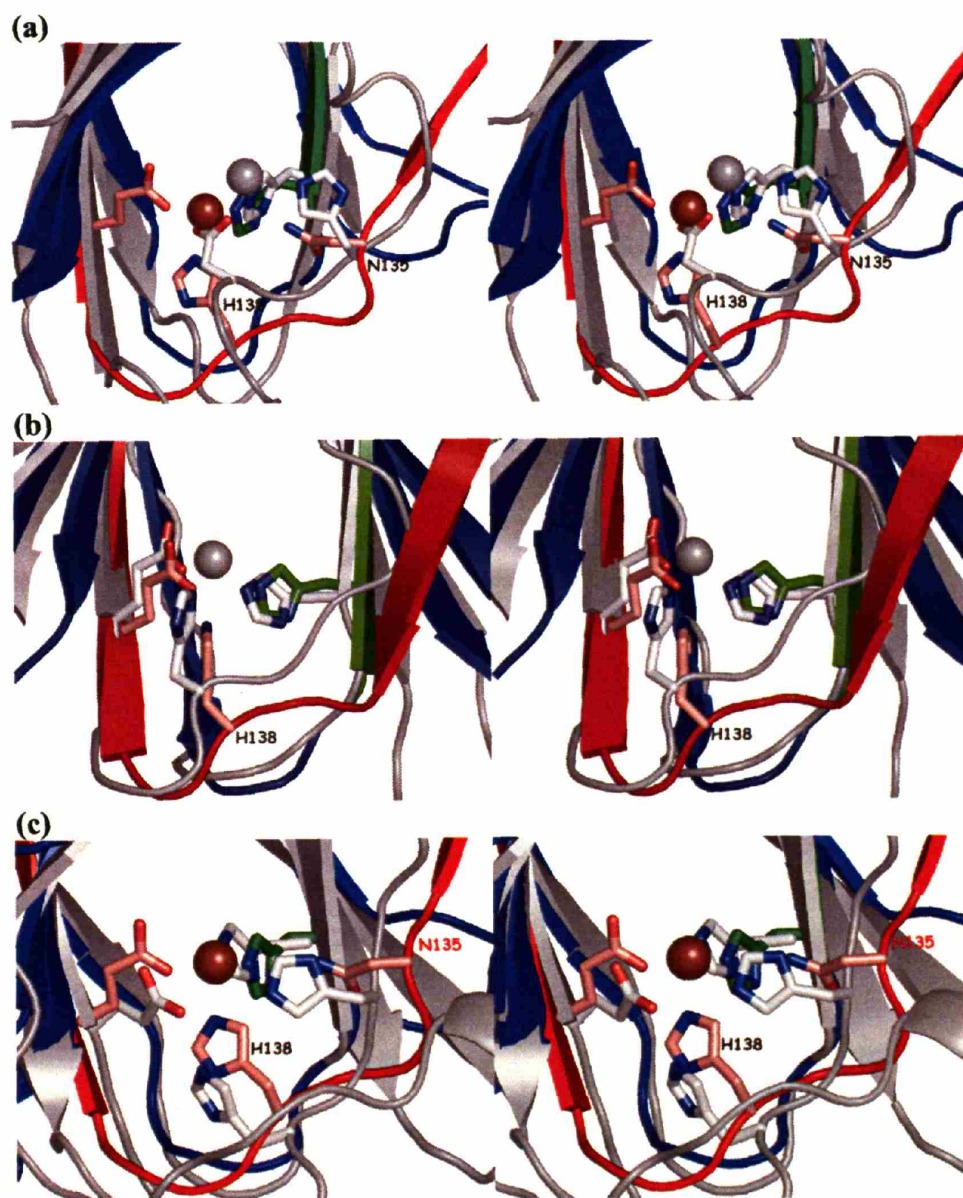
(b)



**Figure V.6. Quaternary structure of GOO.** The GOO hexamer is shown in ribbon representation (stereo), colored by GOO dicupin dimer: dimer 1 (red  $\beta$ -strands and blue  $\alpha$ -helices), dimer 2 (magenta), dimer 3 (green). Mn(II) ions (grey spheres) define the GOO active site.



**Figure V.7. Cupin proteins and the facial triad.** Structural alignments of the HppE  $\beta$ -dimer colored as in Figure V.3.b. with (a) TauD, (b) GOO, and (c) PGI in grey. HppE facial triad ligands are shown in ball-in-stick representation: carbon (motif 1, red; motif 2, green), nitrogen (blue), oxygen (red). The HppE active site iron is shown as a brown sphere. Metal ligands for the aligned proteins structure are shown in ball-in-stick representation: carbon (grey), nitrogen (blue), and oxygen (red).

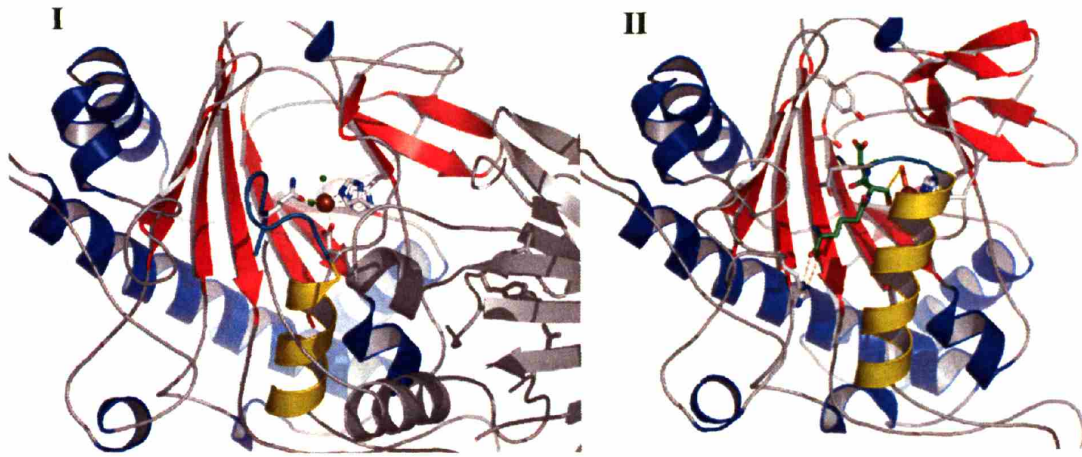




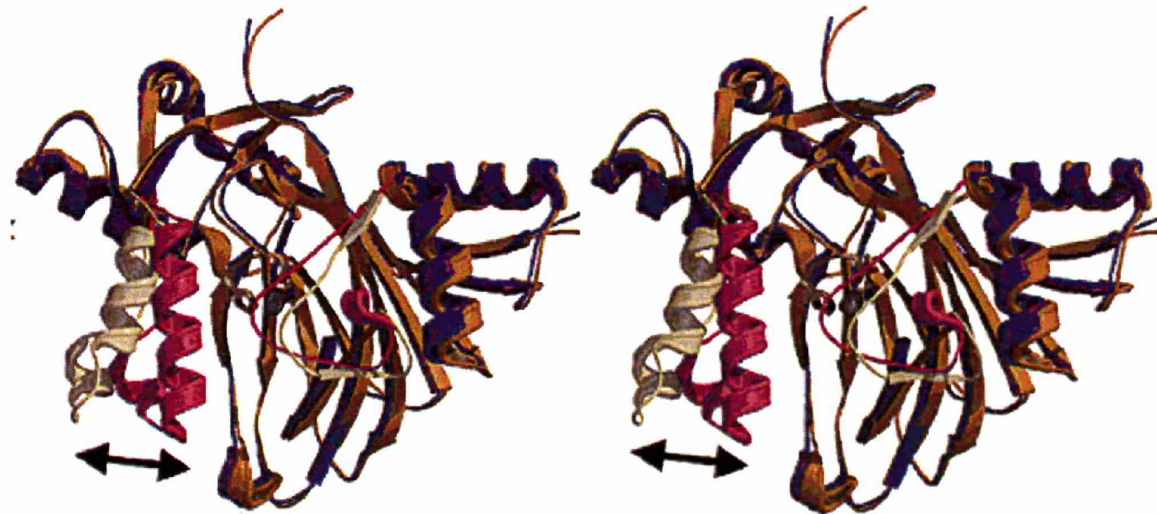
**Figure V.8. Conformational Changes Upon Substrate Binding.**

**(a)** (I) Fe(II)-IPNS structure in the absence of substrate shown in ribbons with strands (red) and helices (blue). A C-terminal helix (yellow) is shown with a C-terminal loop (cyan). (II) The ACV-Fe(II)-IPNS structure, substrate bound, shown as in I. The ACV substrate is shown in ball-in-stick (green)<sup>29</sup>. **(b)** (in stereo) Fe(II)-TauD structure in the absence of substrate shown in ribbons (dark orange) aligned with the Fe(II)-TauD structure in complex with substrate (dark purple). A helical loop is colored beige (no substrate) and purple (with substrate) and undergoes a conformational change upon substrate binding<sup>24</sup>. **(c)** Alignment of Fe(II)-HppE cantilever hairpin (green) with the *S*-HPP bidentate bound Fe(II)-HppE hairpin (yellow). Remaining strands are colored blue and helices are colored red.

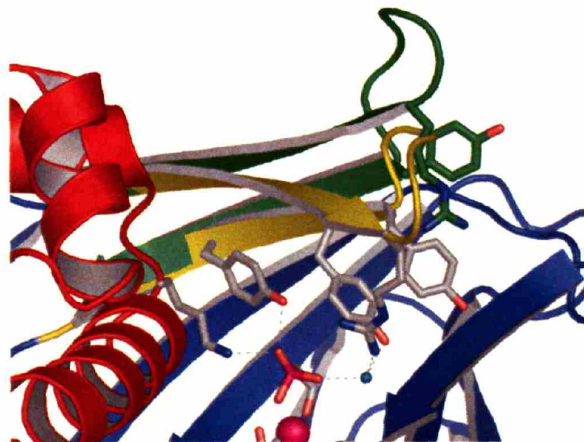
(a)



(b)



(c)



**Figure V.9. HppE sequence alignment.** Primary sequence alignment of HppE from *Streptomyces wedmorensis* (*S.w.*) and *Pseudomonas syringae* (*P.s.*). Facial triad residues (H138, E142, and H180) are colored blue. K23, cantilever hairpin residues (R97, Y102, Y103, Y105), N135, F182, and N197 are colored orange, green, light green, magenta, and purple, respectively.

**S.w.** MSNTKTASTGFAELLKDRREQV**K**MDHAALASLLGETPETVAAWENGEGGE  
**P.s.** MD-VRTLAVGKAHLEALLATR-**K**MTLEHLQDVR---HDATQVYFDGLEHL

**S.w.** LTLTQLGRIAHVLGTSIGALTPPAGNDLDDGVIIQMPDERPILKGV**R**DNV  
**P.s.** QNVAQYLAIPLSEFFVGQ-TQS---DLDDGVKIARRNGGFKREEI**R**GGV

**S.w.** D**Y**Y**V**Y**N**CLVVRTKRAPSLVPLVVDVLTDPDDAKF**N**SG**H**AG**N**E**F**LFVLEGE  
**P.s.** H**Y**Y**T**Y**E**HLVTTNQDPGLMALRLDLHSDDEQPLRL**N**GG**H**GS**R**E**I**VYVTRGA

**S.w.** IHMKW-GDKENPKEALLPTGASMFVEEHV**P**H**A**F**T**AAKGTGSAKLI**A**V**N**F-  
**P.s.** VRVRWVGDNDELKEDVLNEGDSIFILPNV**P**H**S**F**T**NHVGGAKSEII**A**I**N**Y**G**

## V.E. References

1. Liu, P., Murakami, K, Seki, T, He, X, Yeung, SM, Kuzuyama, T, Seto, H & Liu, H. Protein purification and function assignment of the epoxidase catalyzing the formation of fosfomycin. *J Am Chem Soc* **123**, 4619-4620 (2001).
2. Khuri, S., Bakker, F.T. & Dunwell, J.M. Phylogeny, function, and evolution of the cupins, a structurally conserved, functionally diverse superfamily of proteins. *Mol Biol Evol* **18**, 593-605 (2001).
3. Dunwell, J.M., Purvis, A. & Khuri, S. Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**, 7-17 (2004).
4. Dunwell, J.M., Culham, A., Carter, C.E., Sosa-Aguirre, C.R. & Goodenough, P.W. Evolution of functional diversity in the cupin superfamily. *Trends Biochem Sci* **26**, 740-746 (2001).
5. Dunwell, J.M. Cupins: a new superfamily of functionally diverse proteins that include germins and plant storage proteins. *Biotechnol Genet Eng Rev* **15**, 1-32 (1998).
6. Kleywegt, G.J. LSQMAN. (Uppsala, SWEDEN, 1992).
7. Pochapsky, T.C., Pochapsky, SS, Ju, T, Mo, H, Al-Mjeni, F & Maroney, MJ. Modeling and experiment yields the structure of acireductone dioxygenase from *Klebsiella pneumoniae*. *Nat Struct Biol* **9**, 966-972 (2002).
8. Zhang, Z., Ren, J, Stammers, DK, Baldwin, JE, Harlos, K & Schofield, CJ. Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase. *Nat Struct Biol* **7**, 127-133 (2000).

9. Valegard, K., Terwisscha, van Scheltinga, AC, Dubus, A, Ranghino, G, Oster, LM, Hajdu, J & Andersson, I. Structure of a cephalosporin synthase. *Nature* **394**, 805-809 (1998).
10. Elkins, J.M., Ryle, MJ, Clifton, IJ, Dunning, Hotopp, JC, Lloyd, JS, Burzlaff, NI, Baldwin, JE, Hausinger, RP & Roach, PL. X-ray crystal structure of *Escherichia coli* taurine/alpha-ketoglutarate dioxygenase complexed to ferrous iron and substrates. *Biochemistry* **41**, 5185-5192 (2002).
11. Ko, T.P., Ng, J.D. & McPherson, A. The three-dimensional structure of canavalin from jack bean (*Canavalia ensiformis*). *Plant Physiol* **101**, 729-744 (1993).
12. Woo, E.J., Dunwell, J.M., Goodenough, P.W., Marvier, A.C. & Pickersgill, R.W. Germin is a manganese containing homohexamer with oxalate oxidase and superoxide dismutase activities. *Nat Struct Biol* **7**, 1036-1040 (2000).
13. Titus, G.P., Mueller HA, Burgner J, Rodriguez De Cordoba S, Penalva MA, Timm DE. Crystal structure of human homogentisate dioxygenase. *Nat Struct Biol* **7**, 542-546 (2000).
14. Steiner, R.A., Kalk, K.H. & Dijkstra, B.W. Anaerobic enzyme-substrate structures provide insight into the reaction mechanism of the copper-dependent quercetin 2,3-dioxygenase. *Proc Natl Acad Sci U S A* **99**, 16625-16630 (2002).
15. Zhang, Z., Ren, J.S., Clifton, I.J. & Schofield, C.J. Crystal structure and mechanistic implications of 1-aminocyclopropane-1-carboxylic acid oxidase--the ethylene-forming enzyme. *Chem Biol* **11**, 1383-1394 (2004).

16. Higgins, D. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Vol. 2005 (1997).
17. Liu, P., Liu, A, Yan, F, Wolfe, MD, Lipscomb, JD & Liu, HW. Biochemical and spectroscopic studies on (S)-2-hydroxypropylphosphonic acid epoxidase: a novel mononuclear non-heme iron enzyme. *Biochemistry* **42**, 11577-11586 (2003).
18. Higgins, L.J., Yan, F., Liu, P., Liu, H.W. & Drennan, C.L. Structural insight into antibiotic fosfomicin biosynthesis by a mononuclear iron enzyme. *Nature* (2005).
19. Hansen, T., Schlichting, B., Felgendreher, M. & Schonheit, P. Cupin-type phosphoglucose isomerases (Cupin-PGIs) constitute a novel metal-dependent PGI family representing a convergent line of PGI evolution. *J Bacteriol* **187**, 1621-1631 (2005).
20. Valegard, K., Terwisscha, van Scheltinga, AC, Dubus, A, Ranghino, G, Oster, LM, Hajdu, J & Andersson, I. The structural basis of cephalosporin formation in a mononuclear ferrous enzyme. *Nat Struct Mol Biol* **11**, 95-101 (2004).
21. Roach, P.L., Clifton, IJ, Fulop, V, Harlos, K, Barton, GJ, Hajdu, J, Andersson, I, Schofield, CJ & Baldwin, JE. Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. *Nature* **375**, 700-704 (1995).
22. Schofield, C.J. & Zhang, Z. Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. *Curr Opin Struct Biol* **9**, 722-731 (1999).

23. Eichhorn, E., van der Ploeg, J.R., Kertesz, M.A. & Leisinger, T. Characterization of alpha-ketoglutarate-dependent taurine dioxygenase from *Escherichia coli*. *J Biol Chem* **272**, 23031-23036 (1997).
24. O'Brien, J.R., Schuller, D.J., Yang, V.S., Dillard, B.D. & Lanzilotta, W.N. Substrate-induced conformational changes in *Escherichia coli* taurine/alpha-ketoglutarate dioxygenase and insight into the oligomeric structure. *Biochemistry* **42**, 5547-5554 (2003).
25. Price, J.C., Barr, E.W., Tirupati, B., Bollinger, J.M., Jr. & Krebs, C. The first direct characterization of a high-valent iron intermediate in the reaction of an alpha-ketoglutarate-dependent dioxygenase: a high-spin FeIV complex in taurine/alpha-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. *Biochemistry* **42**, 7497-7508 (2003).
26. Price, J.C., Barr, E.W., Glass, T.E., Krebs, C. & Bollinger, J.M., Jr. Evidence for hydrogen abstraction from C1 of taurine by the high-spin Fe(IV) intermediate detected during oxygen activation by taurine:alpha-ketoglutarate dioxygenase (TauD). *J Am Chem Soc* **125**, 13008-13009 (2003).
27. Ryle, M.J., Liu, A, Muthukumar, RB, Ho, RY, Koehntop, KD, McCracken, J, Que, L Jr & Hausinger, RP. O<sub>2</sub>- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase. *Biochemistry* **42**, 1854-1862 (2003).
28. Burzlaff, N.I., Rutledge, PJ, Clifton, IJ, Hensgens, CM, Pickford, M, Adlington, RM, Roach, PL, Baldwin, JE. The reaction cycle of isopenicillin N synthase observed by X-ray diffraction. *Nature* **401**, 721-724 (1999).

29. Roach, P.L., Clifton, IJ, Hensgens, CM, Shibata, N, Schofield, CJ, Hajdu, J & Baldwin, JE. Structure of isopenicillin N synthase complexed with substrate and the mechanism of penicillin formation. *Nature* **387**, 827-830 (1997).
30. Kuzuyama, T., Seki, T., Kobayashi, S., Hidaka, T. & Seto, H. Cloning and expression in *Escherichia coli* of 2-hydroxypropylphosphonic acid epoxidase from the fosfomycin-producing organism, *Pseudomonas syringae* PB-5123. *Biosci Biotechnol Biochem* **63**, 2222-2224 (1999).



## **VI. Appendix**

### **Ribonucleoside triphosphate reductase and thioredoxin from *Lactobacillus leichmannii* \***

\* This work was done in collaboration with Mohammad Seyedsayamdost in the laboratory of JoAnne Stubbe. The author and Mohammad contributed equally to this work.

#### **VI.A. Cloning, expression, and purification of thioredoxin from *Lactobacillus leichmannii***

##### *VI.A.1. Isolation of genomic DNA from *Lactobacillus leichmannii**

Cell stocks (500  $\mu$ L) of *Lactobacillus leichmannii* were purchased from ATCC, transferred to MRS (Becton Dickinson) growth medium and frozen as instructed by the manufacturer. The cell stock was thawed on ice for approximately 20 min, grown at 37° C for 90 min in MRS medium. A 250  $\mu$ L aliquot of cells was used to inoculate 5 mL of autoclaved MRS medium. The cells were grown at 37°C for 2 hrs and spun down in 1.5 ml aliquots, freezing the cell pellets. Two of these aliquots were used for genomic DNA isolation.

Each 1.5 ml aliquot was resuspended in 400  $\mu$ L of 50 mM Tris, 10 mM EDTA, 10 mM NaCl, pH 7.5. Lysozyme (~2 mg) was subsequently added to the cell suspension. After incubating the sample at 37°C for approximately 30 min, 20  $\mu$ L of 30% SDS and 10  $\mu$ L of proteinase K (10 mg/ml) was added to the lysozyme digest and incubated at 37°C for 30 min. After cell lysis, protein precipitation was facilitated by an ammonium sulfate cut (400  $\mu$ L of 4 M ammonium acetate) and the resulting solution was centrifuged at 5000 rpm for 15 min. A phenol:chloroform DNA extraction was performed using the

supernatant of the ammonium sulfate cut. Specifically, a 1:1 buffered phenol:chloroform solution was added to the aqueous DNA sample and centrifuged at 11000 rpm for 15 min to yield optimal separation. This step was repeated twice followed by a chloroform extraction. DNA was precipitated from the aqueous layer using an equal volume of isopropyl alcohol. After approximately 15 min the DNA (white ppt.) was spun down at 11000 rpm for 10 min, the supernatant was discarded, and the resulting pellet was resuspended in 400  $\mu$ L of 100 mM sodium acetate at pH 6.0. The pellet was sequentially washed with ethanol ( $\sim$ 800  $\mu$ L) and 70% ethanol (1 mL), centrifuged at 11000 rpm for 10 min, and was allowed to dry for 5 min and 1 hr, respectively. After the 70% ethanol wash, the pellet was resuspended in sterile ddH<sub>2</sub>O. Each 1.5 mL aliquot produced two DNA samples, gDNA A and gDNA B. An extraction step was performed with gDNA B replacing the 70% ethanol wash. Here, a 1mL aliquot of chilled isopropyl alcohol was added to the ethanol solution in order to precipitate gDNA B. The two DNA samples (A and B) were combined and analyzed by agarose gel electrophoresis (Figure VI.1.).

To remove RNA that precipitated with the DNA, an 18  $\mu$ g aliquot of genomic DNA was subjected to DNase-free RNase digestion (100 ng) at 37°C for 30 min. The digestion was analyzed by agarose gel electrophoresis (Figure VI.2.).

#### *VI.A.2. Cloning thioredoxin gene from Lactobacillus leichmannii genomic DNA*

To amplify the thioredoxin (Tr) gene using *Lactobacillus leichmannii* genomic DNA, the following primers were designed based on the thioredoxin protein sequence (Hogenkamp unpublished results):

Forward Primer: ATG GTG GAT GAA ATT ACC GAT GCG ACC TTT GAA G

Reverse Primer: TCA ATC GGT ATA CTG ATC CAG AAT TTC GTT CAG T

PCR reactions were performed using both Taq and Pfu polymerase (Table VI.1. and Table VI.2.) via the following cycling protocol:

PCR Conditions:

- 1- 94°C, 4 min
- 2- 54°C, 1 min
- 3- 72°C, 1 min
- 4- 94°C, 1 min
- 5- Repeat Step 2-4, 29 times
- 6- 54°C, 1 min
- 7- 72°C, 10 min

An agarose gel was run to analyze the PCR results (Figure VI.3.).

Bands from reactions 3 and 4 (Figure VI.3.) were cut out of the gel, the DNA was extracted using a standard protocol (Qiagen gel extraction kit), and the product was blunt ligated into a pCR 2.1 TOPO cloning vector and subsequently transformed into *E. coli* competent cells as described by the manufacturer (Invitrogen). Two colonies were obtained from each reaction (3 and 4), grown using LB medium as recommended in the TOPO cloning vector protocol, and a DNA miniprep (Qiagen) was performed on both samples. Each DNA sample was sequentially sequenced at the MIT Biopolymers Laboratory. Both samples gave identical DNA sequence (Figure VI.4.) results, and when translated, the protein sequence matched the Hogenkamp thioredoxin sequence.

#### *VI.A.3. PCR of Tr in TOPO vector and ligation into a protein expression vector*

pCR 2.1 TOPO vector with thioredoxin (Tr) blunt ligated (pTr-clone) was used in the following subcloning procedure. The objective was to obtain a PCR product with NdeI and BlnI flanking restriction sites. To accomplish this objective, the following primers were designed (restriction sites are shown in red):

Forward Primer (NdeI): GCGCGCA-TATGGTGGATGAAATTACCGATGCGACC

Reverse Primer (BlnI): CACCGCT-AGCTCAATCGGTATACTGATCCAGAATTTTCG

Using pTr-clone template DNA, PCR reactions were performed as indicated in Table

VI.3. The following PCR conditions were used:

1- 95°C, 5 min

2- 55°C 1 min

3- 68°C, 1 min

4-95°C, 1 min

5- repeat steps 3 and 4 30 times

6- 55°C, 1 min

7- 68°C, 10 min

8- 4°C, store

The results of these reactions are shown in Figure VI.5.

#### *VI.A.4. Ligation of the thioredoxin gene into pET 23a*

The pET 23a vector (novagen) (3 µg) was digested using 20 U of NdeI and BlnI (NEB enzymes) in NEB buffer 4 with a total volume of 30 µL. After a 2 hr digestion at 37°C, 10 U of alkaline phosphatase from calf intestine was added and incubated for 1 hr at 37°C. After isolation by agarose gel electrophoresis, a Qiagen gel extraction kit was used as indicated by the manufacturer in order to purify the digested pET vector. PCR reaction 7 (Table VI.3. and Figure VI.5.) was used for ligation into a pET 23a vector. The PCR product was digested as in Table VI.4. using a standard protocol (NEB). The digested product was purified using a Qiagen gel extraction kit.

A ligation reaction was prepared using the following (final volume 20 µL): 2 µL of 10X T4-DNA ligase buffer, 400 U of T4 DNA ligase, 80 ng of Tr-gene insert, and 138 ng of digested pET 23a. After an overnight ligation at 4°C, the plasmids were amplified via transformation into XL1-Blue competent cells. Cells were grown and plasmid was

isolated using a Qiagen miniprep kit (standard protocols were followed). The pET23a-Tr plasmid was sequenced and the DNA sequence was verified.

The pET23a-Tr plasmid was used to transform BL21-DE3 pLys competent cells as described by the manufacturer in order to generate sufficient protein for future experiments.

#### *VI.A.5. Purification of C31S-Tr from Lactobacillus leichmannii*

The pET23a-Tr plasmid was used to obtain C31S-Tr using the Quickchange kit from Stratagene. Primers were designed (serine to cysteine mutation site shown in red below) and the plasmid (pTrC31S) was obtained as described in the Quickchange protocols. The DNA sequence of pTrC31S was confirmed. The plasmid was transformed into BL21-DE3 pLys cells for optimal protein expression.

Forward Primer:

5'- GG GCA ACC TGG TGC GGT CCT **AGC** CGG ATG CAG TCA C - 3'

Reverse Primer:

5'- GAC TGG TGA CTG CAT CCG **GCT** AGG ACC GCA CCA GGT - 3'

A single colony of BL21-DE3 cells with pTrC31S was used to inoculate a small culture (5 mL of LB medium with 100  $\mu\text{g}/\text{mL}$  of ampicillin). A 1 mL aliquot of a saturated cell culture was used to inoculate 2 L of LB with ampicillin. Cells were grown to an  $\text{OD}_{600}$  of 0.8 at 37°C and were then induced with IPTG (1 mM). After 3 hrs, the cells were centrifuged for 15 min at 5000 x g. This growth gave 6.3 g of cell paste.

The cell pellet (6.3 g) was resuspended in 60 mL of buffer (50 mM Tris, 3 mM EDTA, pH 7.4, 6.3 mg PMSF, 13.3 mg TLCK). The cells were lysed using a French press at 14000 psi (Two passes were required). The cell debris was separated from the cell lysate via centrifugation at 14000 x g for 30 min. Streptomycin sulfate (4.8 %) in buffer A (50 mM KPi, 3 mM EDTA, pH 7.3) was used to remove DNA from the lysate. The streptomycin sulfate was added over 15 min and the precipitated DNA was removed via centrifugation (14000 x g for 30 min). The supernatant was diluted to a final volume of 160 mL using buffer A. A DE-52 anion exchange column was then equilibrated in buffer A using 50 mL of resin and the 160 mL sample was loaded onto this column. A linear gradient (600x600 mL) of buffer A to buffer B (200 mM KPi, 3 mM EDTA, pH 7.3) was then used for elution (flow rate ~ 1 mL/min). Fractions 88-95 were pooled using SDS-PAGE (not shown) and concentrated to 20 mL, see Figure VI.6. for elution profile.

A 7.5 mL aliquot from the DE-52 concentrated elution was then loaded onto a G-75 size exclusion column equilibrated in buffer A. The column was run at a flow rate of ~ 0.3 mL/min and fractions 29-40 were pooled based on SDS-PAGE (not shown). The G-75 elution profile is shown in Figure VI.7. The fractions were concentrated to 195  $\mu$ M. A purification gel and table are shown in Figures VI.8. and VI.9., respectively.

## **VI.B. Purification of C731S-RTPR**

The C731S-RTPR mutant glycerol stock was obtained from the Stubbe laboratory. C731S-RTPR was purified in the Stubbe laboratory by Mohammad Seyedsayamdost. The protocol requires an ammonium sulfate cut (45%) followed by a DE-52 anion exchange column and finally an anion exchange H-Q column.

## VI.C. RTPR-Tr coupling experiments

### VI.C.1. Overview

Previous work on *Escherichia coli* thioredoxin suggests that one active site cysteine is more nucleophilic than the other, specifically C32<sup>1,2</sup>. The analogous residue in thioredoxin from *Lactobacillus leichmannii* is C28. It is reasonable to hypothesize that this cysteine attacks the oxidized C-terminus of ribonucleotide triphosphate reductase from *Lactobacillus leichmannii* (RTPR), either C731 or C736. This section describes attempts at coupling C731S-RTPR to C31S-Tr from *Lactobacillus leichmannii* which would trap a species that exists only transiently *in vivo*.

### VI.C.2. Overall coupling strategy

The goal of this work is to couple RTPR to thioredoxin, trapping a heterodisulfide interaction between the nucleophilic cysteine of thioredoxin and the electrophilic cysteine of the RTPR C-terminus. Since thioredoxin has two cysteines and RTPR has ten, a C31S-thioredoxin mutant protein was activated for RTPR coupling using a small molecule disulfide reagent (5-5'-dithiobis-(2-nitrobenzoic acid); DTNB (Sigma)). This protein (Tr-TNB) was mixed with C731S-RTPR at a ratio of 20:1 (Tr to RTPR) in an attempt to couple the two proteins together.

### VI.C.3. Coupling procedure

The concentration of C31S-Tr was 195  $\mu\text{M}$  and that of C731S-RTPR was 133  $\mu\text{M}$ , both in 50mM potassium phosphate, pH 7.6. C31S-Tr was prereduced by incubating with DTT such that 0.375  $\mu\text{moles}$  of DTT in 50mM potassium phosphate, pH 7.6 was added

12 nmoles of thioredoxin mutant protein (final volume = 66.5  $\mu$ L). This sample was allowed to equilibrate for approximately 30 min on ice. DTT was removed from the C31S-Tr sample using a prepacked micro Bio-spin P-6 column (BioRad). Specifically, the column was exchanged four times using 500  $\mu$ L of 50 mM potassium phosphate, pH 7.6 iteratively as described in the micro Bio-spin column instruction manual (BioRad). The C31S-Tr sample was loaded onto the potassium phosphate buffer equilibrated column and the flow-through containing the reduced C31S-Tr sample was retained as described in the BioRad literature. A second column was prepared using 50mM potassium phosphate, pH 7.6 and the flow-through was transferred to this column to ensure that all the DTT was removed from the sample.

The eluent of the second column was used for TNB (5-thio-2-nitrobenzoic acid) labeling. An aliquot of 30 nmoles of DTNB in 100 mM sodium acetate, pH 5.1 was added to 11.4 nmoles of C31S-Tr reduced and desalted protein (final volume = 70  $\mu$ L). The reaction was incubated at 4°C for 30 min. The reaction solution was transferred to a micro Bio-spin column that was equilibrated in 50 mM potassium phosphate, pH 7.6 as described in the Bio-Rad literature to remove TNB (yellow) product from the Tr-TNB coupled protein. The eluent was stored on ice until needed.

A aliquot of C731S-RTPR (4 nmoles) was reduced with 500 nmoles of DTT in 50 mM potassium phosphate, pH 7.6) and incubated on ice for 30 min. As described for the C31S-Tr reduction, the DTT was removed in two iterations (i.e. sequentially) using two micro Bio-spin columns equilibrated in 50 mM potassium phosphate, pH 7.6. The eluent from the second column was stored on ice until needed.



In order to couple C731S-RTPR to C31S-Tr, the reduced C731S-RTPR was mixed with C31S-Tr coupled to TNB (C31S-Tr-TNB). An aliquot of C731S-RTPR is mixed with an aliquot of C31S-Tr-TNB (1:20) and the reaction was incubated at 4°C for 30 min. A polyacrylamide gel (10% T (polyacrylamide concentration)) was loaded as described in Figure VI.10. The left-side of the gel is a reducing SDS-PAGE gel and the right side is a non-reducing (i.e. no  $\beta$ -mercaptoethanol (BME)) SDS-PAGE gel.

Figure VI.10. is consistent with a heterodisulfide linkage between C731S-RTPR and C31S-Tr; however the coupling experiment did not yield a homogeneous species. In order to obtain a homogeneous sample for crystallization of the RTPR:Tr complex, a mutant of RTPR devoid of all cysteines except C731 or C736 could be made. If this RTPR mutant protein can be purified to homogeneity, then the aforementioned coupling reaction should be more efficient and selective.

#### **VI.D. RTPR in complex with cob(II)alamin, adenine 5'-deoxynucleoside, substrate and/or effector.**

##### *VI.D.1. C731,736S-RTPR protein purification*

The C731/736S-RTPR mutant protein was purified by Mohammad Seyedsayamdost in the Stubbe laboratory following the protocol outlined in appendix section B, Purification of C731S-RTPR. The mutant protein sample was dialyzed into a new buffer (RTPR storage buffer: 100 mM sodium citrate pH 5.6, 1.6M ammonium sulfate, 8% PEG 8000, 20% glycerol) and concentrated to 20 mg/ml.

#### *VI.D.2. C731,736S-RTPR crystallization*

Crystals were grown using the C731,736S-RTPR mutant protein sample (Appendix E1) at 20 mg/ml in the RTPR storage buffer. The precipitant solution used for crystallization was made as follows: (1) 5 mL of 1 M sodium citrate, pH 5.6, 23 mL of 3.5 M ammonium sulfate, 10 mL of glycerol (neat), and 4 mL of ddH<sub>2</sub>O were mixed to homogeneity via vortexing in a 50 mL conical vial, (2) the solution was filtered through a 0.2  $\mu$ m syringe filter, (3) 8 mL of 50% PEG 8K was added to this solution and vortexed on high to give a white, translucent suspension, (4) the heterogeneous suspension was allowed to settle for ~ 24 hrs or until a clear phase separation was visible, (5) the bottom layer of this mixture was used as the RTPR precipitant solution. Importantly, this protocol can be varied by replacing the addition of 8 mL of 50% PEG 8K with 8 mL of ddH<sub>2</sub>O. After filtering the solution, PEG solid can be added (4 g). It is important to grind the PEG solid with a mortar and pestle in order to obtain a fine white powder before addition. The solution is vortexed to a heterogeneous suspension, removing all large particles, and subsequently treated as described above.

For the purposes of crystallization described here, the protocol using 50% PEG 8K (aq) was used to obtain the RTPR precipitant solution. It is important to note that variability in crystallization results is due to the variability in PEG and glycerol concentrations. PEG is saturated in the RTPR precipitant solution but the concentration of PEG in solution varies with concentration of glycerol which is difficult to accurately quantitate. For this reason, three different drop ratios (protein:precipitant) were used in the crystallization trials, 1:1, 1:1.5, and 1.5:1. The hanging drop vapor diffusion method was used with Libro plates (Hampton Research) such that the well solution was 500  $\mu$ L

with 3 drops used per coverslip (siliconized). Drops consisted of the following protein to precipitant volumes: 2  $\mu\text{L}$  + 2  $\mu\text{L}$ , 2  $\mu\text{L}$  + 3  $\mu\text{L}$ , and 3  $\mu\text{L}$  + 2  $\mu\text{L}$ . Crystals grew in  $\sim$  36 hrs with approximate dimensions, 200 x 100 x 100  $\mu\text{m}$ . This crystal form was solved previously by Michael Sintchak in the Drennan Laboratory<sup>3</sup>.

For soaking experiments with cob(II)alamin, crystals were grown anaerobically in a Coy chamber under an argon atmosphere. In an effort to define the cob(II)alamin and adenine 5'-deoxynucleoside (5'dA) binding sites in RTPR, the following combinations were used: (1) cob(II)alamin, 5'-dA, and dGTP (3x soak, includes post carbon-cobalt homolysis products and effector) and (2) cob(II)alamin, 5'-dA, ATP, and dGTP (3x soak, includes post carbon-cobalt homolysis products, substrate, and effector).

The soaking solution used in experiment 1 was 2.3 M ammonium sulfate, 12.2% PEG 8K, 18% glycerol, 12.2 mM DTT, 367 mM HEPES, pH 7.6. dGTP and 5'dA were dissolved in this soaking solution to final concentrations of 137 mM and 15 mM, respectively. The 5'dA solution (2.4  $\mu\text{moles}$ ), dGTP solution (4.5  $\mu\text{moles}$ ), and soaking solution were mixed (250  $\mu\text{L}$  total volume) and degassed by gently bubbling with argon for 30 min. In the anaerobic Coy chamber, 6.6  $\mu\text{L}$  of cob(II)alamin (synthesized by Dr. John Robblee in the Laboratory of JoAnne Stubbe) at 11 mM was added to 30  $\mu\text{L}$  of the dGTP:5'dA solution. The solution was allowed to equilibrate for 4 hrs by diffusion; a uniform brown color resulted. The ratio dGTP:cob(II)alamin:5'dA is 1.9:1.1:1 at final concentrations of 14.8 mM, 9 mM, and 8 mM, respectively. Crystals of C731,736S-RTPR were transferred to the 36.6  $\mu\text{L}$  drop and allowed to equilibrate for 16.5 hrs over a 500  $\mu\text{L}$  well of soaking solution (The concentration of RTPR is unknown since this is a

soaking experiment.). The crystals were then cooled using liquid nitrogen for data collection as described below.

The soaking solution used in experiment 2 was identical to that used in soaking experiment 1. ATP was dissolved in this soaking solution to yield a final concentration of 342 mM. The 5'dA solution (2.4  $\mu$ moles), dGTP solution (4.5  $\mu$ moles), ATP solution (9.2  $\mu$ moles), and soaking solution were mixed (250  $\mu$ L total volume) and degassed by gently bubbling with argon for 30 min. Cob(II)alamin was added, as in experiment 1, to a 30  $\mu$ L drop of 5'dA:dGTP:ATP solution and allowed to equilibrate for 4 hrs. The ratio ATP:dGTP:cob(II)alamin:5'dA is 3.8:1.9:1.1:1 at final concentrations of 30.3 mM, 14.8 mM, 9mM, and 8mM, respectively. Crystals of C731,736S-RTPR were transferred to the 36.6  $\mu$ L drop and allowed to equilibrate for 16.5 hrs over a 500  $\mu$ L well of soaking solution.

After soaking crystals were transferred to the following cryoprotectant solution (36.6  $\mu$ L) with substrate, effector, 5'dA, and cob(II)alamin: 1.9 M ammonium sulfate, 10% PEG 8K, 20% glycerol, 300 mM HEPES, pH 7.6. Crystals were cooled in liquid nitrogen in the anaerobic chamber for less than 30 s. Crystals were stored for data collection.

### *VI.D.3. Preliminary data analysis*

Data sets of the 3x soak crystals (Table VI.5.) and the 4x soak crystals (data not shown) were collected using a rotating copper anode source (1.542 Å), 200 images collected at 0.75° oscillations. Data processing in DENZO and SCALEPACK<sup>4</sup> and rigid body refinement in CNS using the apo-RTPR model (PDB code 1L1L) revealed that

neither substrate nor co-substrate was present in the active site or in the effector site. In the 3x soak (Figure VI.11.), the cob(II)alamin was disordered relative to the adenylnpenthylcobalamin:RTPR structure (unpublished results from Michael Sintchak in the Drennan Laboratory). The cob(II)alamin in the 4xsoak was disordered such that modeling was not possible (i.e. the structure was apo-RTPR). Future crystallographic analysis will be required to obtain an informative cob(II)alamin:5'dA structure in complex with substrates and/or effectors.

## VI.E. Figures and Tables

**Table VI.1. PCR reactions for Tr gene amplification (Reactions 1-4)**

Component	Reaction 1	Reaction 2	Reaction 3	Reaction 4
Genomic DNA	100 ng	200 ng	100 ng	200 ng
F-Primer	100 nM	100 nM	400 nM	400 nM
R-Primer	100 nM	100 nM	400 nM	400 nM
DNTP	250 uM	250 uM	250 uM	250 uM
Polymerase	2.5 U Taq	2.5 U Taq	2.5 U Pfu	2.5 U Pfu
10 x Buffer	5 µl	5 µl	5 µl	5 µl
MgSO <sub>4</sub>	--	--	4 mM	4 mM
Water	41.5 µl	41 µl	33 µl	32.5 µl
Total	50 µl	50 µl	50 µl	50 µl

**Table VI.2. PCR reactions for Tr gene amplification (Reactions 5-8)**

<u>Component</u>	<u>Reaction 5</u>	<u>Reaction 6</u>	<u>Reaction 7</u>	<u>Reaction 8</u>
Genomic DNA	62 ng	124 ng	62 ng	124 ng
F-Primer	100 nM	100 nM	400 nM	400 nM
R-Primer	100 nM	100 nM	400 nM	400 nM
dNTP	250 uM	250 uM	250 uM	250 uM
Polymerase	2.5 U Taq	2.5 U Taq	2.5 U Pfu	2.5 U Pfu
10 x Buffer	5 µl	5 µl	5 µl	5 µl
MgSO <sub>4</sub>	--	--	4 mM	4 mM
Water	41 µl	40 µl	33 µl	32 µl
Total	50 µl	50 µl	50 µl	50 µl

**Table VI.3. Reactions set up to incorporate restriction sites into the PCR product**

Component	Reax 1/2	Reax 3/4	Reax 5/6	Reax 7/8
10 x Buffer	5 $\mu$ l	5 $\mu$ l	5 $\mu$ l	5 $\mu$ l
Primer 1	1.6 (125 ng)	1.6 (125 ng)	1.6 (125 ng)	1.6 (125 ng)
Primer 2	1.6 (125 ng)	1.6 (125 ng)	1.6 (125 ng)	1.6 (125 ng)
dNTP (20mM)	2 $\mu$ l	2 $\mu$ l	2 $\mu$ l	2 $\mu$ l
Template	25 ng	25 ng	50 ng	50 ng
MgSO <sub>4</sub> (0.1M)	--	2 $\mu$ l	--	2 $\mu$ l
Water	37.8 $\mu$ l	35.8 $\mu$ l	36.8 $\mu$ l	34.8 $\mu$ l
Pfu	2.5 Units	2.5 Units	2.5 Units	2.5 Units



**Table VI.4. Double digestion of PCR reaction 7 in Table VI.2. and Figure VI.5.**

	Volume ( $\mu$ l)	Amount
PCR Product	10	2 ug
10 x NEB 4 Buffer	4	--
Bip I	2	20 U
Nde I	1	20 U
ddH <sub>2</sub> O	23	--
Total	40	--

**Table VI.5. Data table for the RTPR:cobalamin structure.**

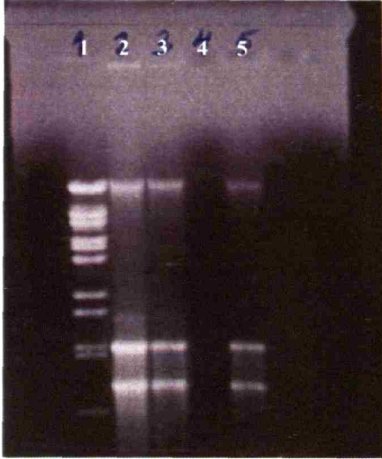
	<b>3xSoak</b>
<b>Data Collection</b>	
Space Group	P2 <sub>1</sub>
Unit Cell Dimensions	
a (Å)	120.6
b (Å)	114.4
c (Å)	121
Wavelength (Å)	1.54
Resolution Range (Å)	30-2.8
Observations	157008
Unique Reflections	67803
Redundancy	2.3
Completeness (%) <sup>a</sup>	89.0 (84.7)
I/σ(I) <sup>a</sup>	11.1 (2.2)
Rsym (%) <sup>a,b</sup>	6.9 (34.2)
<b>Refinement</b>	
R-factor <sup>d</sup> (Rfree <sup>e</sup> ) (%)	27.2 (31.8)
Resolution Range (Å)	50-2.2
No. Molecules per asu <sup>c</sup>	4
Number of Atoms:	
Protein	22352

<sup>a</sup>Values in parentheses are the highest-resolution shell. <sup>b</sup> $R_{sym} = \sum_{hkl} |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \langle I(hkl) \rangle$ , where  $I_i(hkl)$  is the  $i^{\text{th}}$  measured diffraction intensity and  $\langle I(hkl) \rangle$  is the mean of the intensity for the miller index (hkl). <sup>c</sup>ASU, asymmetric unit.

<sup>d</sup> $R_{factor} = \sum_{hkl} ||F_o(hkl)| - |F_c(hkl)|| / \sum_{hkl} |F_o(hkl)|$ . <sup>e</sup> $R_{free} = R_{factor}$  for a test set of reflections (5%).

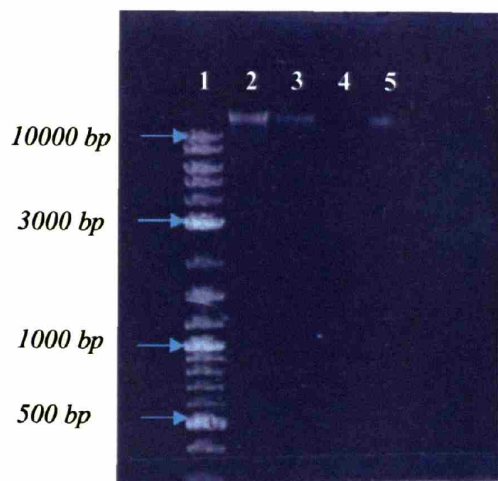
**Figure VI.1. Agarose gel of isolated genomic DNA from *Lactobacillus leichmannii*.**

Lane 1: 1kb DNA Ladder (Promega); Lane 2: 7.2  $\mu\text{g}$  of genomic DNA; Lane 3: 3.1  $\mu\text{g}$  of genomic DNA; Lane 4: no sample; Lane 5: 3.1  $\mu\text{g}$  of genomic DNA.



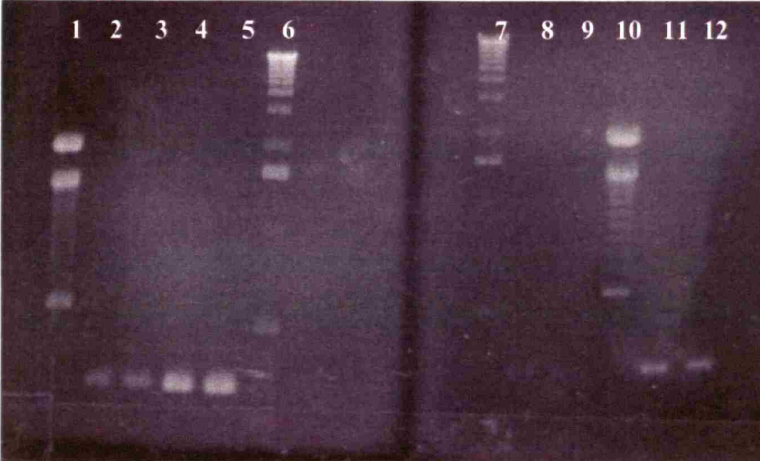
**Figure VI.2. Agarose gel of genomic DNA RNase digest**

Lane 1: 2 log DNA Ladder (NEB); Lane 2: 7.4  $\mu\text{g}$  of genomic DNA; Lane 3: 2.2  $\mu\text{g}$  of DNA; Lane 4: no sample; Lane 5: 3.7  $\mu\text{g}$ .



**Figure VI.3. PCR of the Tr gene.**

Lane 1: 100 kb DNA ladder; Lanes 2-5: reactions 1-4 (Table VI.1.); Lanes 6-7: 1000 kb DNA ladder; Lanes 8-9: reactions 5-6 (Table VI.2.); Lane 10: 100 kb DNA ladder; Lanes 11-12: reactions 7-8 (Table VI.2.).



**Figure VI.4. Sequence of cloned *L. leichmannii* TR gene**

ATG GTG GAT GAA ATT ACC GAT GCG ACC TTT GAA GAA GAA ACC  
AAG GAT GGC TTT GTT CTG GTT GAC TTT TGG GCA ACC TGG TGC  
GGT CCT TGC CGG ATG CAG TCA CCA GTC ATC GAA GCT TTG GCT  
GAA GAA AGA CAA GAC GTG CAC TTC ACT AAG ATG GAC GTT GAC  
GAC AAC CCG GAA ACG GCC AAG AAC TAT GGC ATC ATG GCC ATC  
CCA ACC CTC TTG ATC AAG AAG GAC GGG CAA GTG GTT GAC CGC  
TTG ACT GGC TAT ACT CCA AAG GAA AAG CTG AAC GAA ATT CTG  
GAT CAG TAT ACC GAT TGA

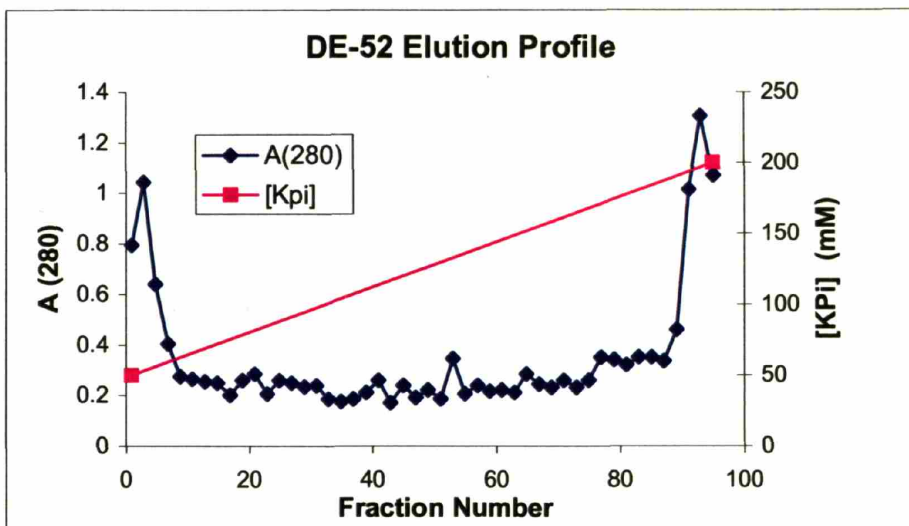
**Figure VI.5. Subcloning PCR reactions.**

Lanes 1-3: 5  $\mu$ L of reactions 1-3; Lane 4: DNA 100 bp Ladder; Lanes 5-7: 5  $\mu$ L of reactions 4-6; Lane 8: 100 bp DNA ladder; Lanes 9-10: 5  $\mu$ L of reactions 7-8 (Table VI.3.). (expected #bp = 312)



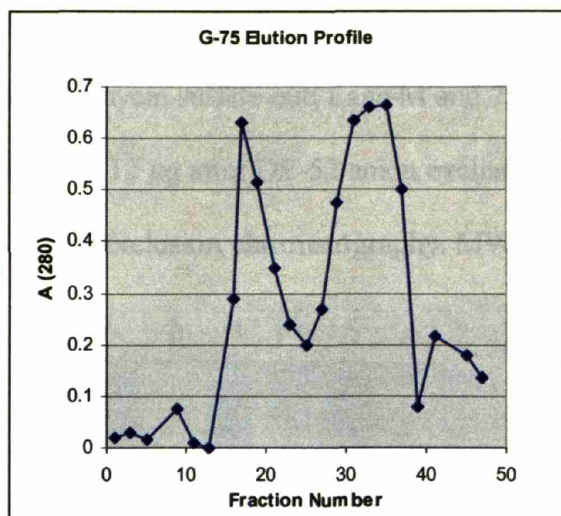
**Figure VI.6. Elution profile from DE-52 anion exchange chromatography.**

The Kpi line (pink) corresponds to the salt gradient.



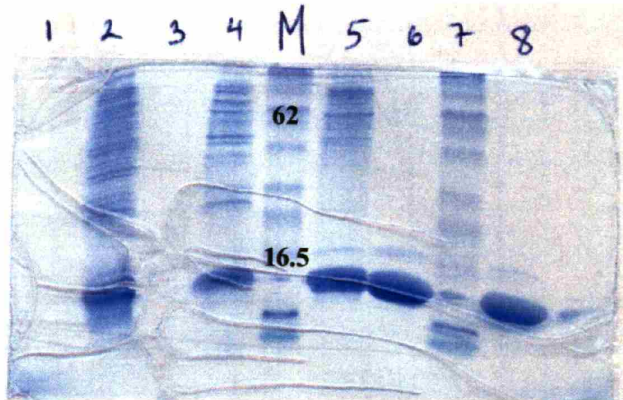


**Figure VI.7. Elution profile from G-75 size exclusion chromatography.**



**Figure VI.8. Final C31S-Tr purification gel**

Lane 1: no sample; Lane 2: 12  $\mu\text{g}$  crude extract; Lane 3: no sample; Lane 4: 12  $\mu\text{g}$  after streptomycin sulfate cut; Lane M and 7: NEB Broad Range Protein Marker protein ladder; Lane 5: 12  $\mu\text{g}$  after DE-52 anion exchange chromatography; Lane 6 and 8: 18  $\mu\text{g}$  after G-75 size exclusion chromatography. MW Tr = 13 kDa.

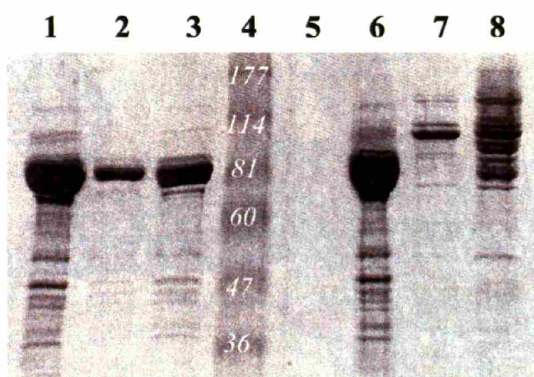


**Figure VI.9. Purification table**

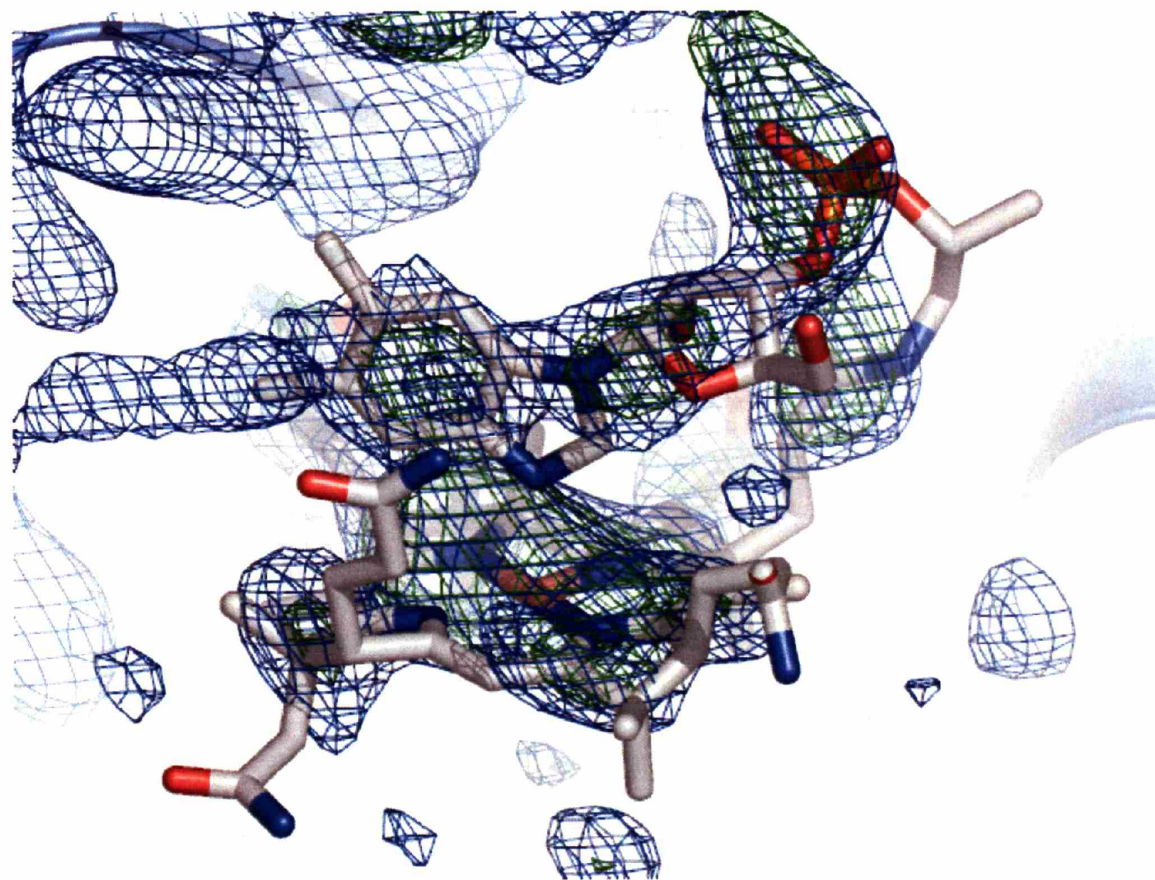
<u>Sample</u>	<u>Conc. (mg/ml)</u>	<u>Volume (ml)</u>	<u>Amount (mg)</u>
Crude Extract	2.4	80	192
Strp. Sulfate	2.2	100	217
DE-52 conc.	2.3	20	47
Final	1.2	24	29

**Figure VI.10. C731S-RTPR:C31S-Tr reducing and non-reducing polyacrylamide.**

Lanes 1-3 were boiled in both SDS and BME whereas lanes 5-8 were boiled in SDS-alone (i.e. no BME added) [Bio-Rad SDS-PAGE loading buffer]. Lanes 1,2 and lanes 6,7 were loaded using 5  $\mu$ L aliquots. Lanes 3 and 8 were run using 10  $\mu$ L aliquots. Lanes 1,6: C731S-RTPR alone; Lane 2,3,7,8: C31S-Tr:C731S-RTPR coupling reaction; Lane 4: protein ladder (Invitrogen Pre-stained ladder), numbers (white) in kDa; Lane 5: no sample. Note: Tr alone is not visible on the gel.



**Figure VI.11. Electron density near the cobalamin binding site of RTPR.** A 2FoFc map (blue) is contoured at  $1\sigma$  and a FoFc map (green) is contoured at  $3\sigma$ . The cob(II)almin molecule is shown in ball-in-stick representation. The RTPR protein proximal to the cobalamin binding site is shown as a cartoon (slate).



## VI.F. References.

1. Chivers, P. T. & Raines, R. T. General acid/base catalysis in the active site of *Escherichia coli* thioredoxin. *Biochemistry* **36**, 15810-15816 (1997).
2. Chivers, P. T., Prehoda, KE, Volkman, BF, Kim, BM, Markley, JL & Raines, RT. Microscopic pKa values of *Escherichia coli* thioredoxin. *Biochemistry* **36**, 14985-14991 (1997).
3. Sintchak, M. D., Arjara, G., Kellogg, B. A., Stubbe, J. & Drennan, C. L. The crystal structure of class II ribonucleotide reductase reveals how an allosterically regulated monomer mimics a dimer. *Nat Struct Biol* **9**, 293-300 (2002).
4. Otwinowski Z, M. W. Processing of X-ray Diffraction Data Collected in Oscillation Mode. *Methods in Enzymology* **276**, 307-326 (1997).

# Curriculum Vitae

Luke James Higgins

## EDUCATION

- 2000-2005:** PhD in Biological Chemistry; Massachusetts Institute of Technology (MIT)  
Thesis Title: *Structural Analysis of Hydroxypropylphosphonic Acid Epoxidase: A Fosfomycin Biosynthetic Enzyme*  
Graduate Advisor: Dr. Catherine L. Drennan, PhD
- 1996-2000:** B.S. in Biochemistry and Molecular Biology; University of Maryland

## AWARDS/HONORS

*Phi Beta Kappa, UMD*  
*Golden Key National Honor Society, UMD*  
*American Institute of Chemists: Excellence in Biochemistry, UMD*  
*Departmental Award for Academic Excellence in Biochemistry, UMD*  
*UMD Provost Undergraduate Research Award*  
*NIH/CEHS Research Assistantship: Edgerton Center, MIT*  
*Department of Chemistry Fellowship, MIT*

## Teaching Experience:

- Chemistry Tutor (2 years), UMD Chemistry Tutorial Center  
Teaching Assistantships, MIT
- 5.12 Organic Chemistry
  - 5.310 Introductory Chemistry Laboratory
  - 5.78 Practical Protein Crystallography

## MEETINGS

ACS 224<sup>th</sup> National Meeting  
Poster Presentation: *The Elusive C-terminus of Ribonucleotide Reductase*

Cobalamin Gordon Conference:  
Poster Presentation: *Catalytic Activation of Class II Ribonucleotide Reductase from *Lactobacillus leichmannii**

## PUBLICATIONS

Luke J. Higgins, Feng Yan, Pinghua Liu, Hung-wen Liu & Catherine L. Drennan  
*Structural Insight into Antibiotic Fosfomycin Biosynthesis by a Mononuclear Iron Enzyme*; Nature. 2005 Jul 13; [doi: 10.1038/nature03924].



Room 14-0551  
77 Massachusetts Avenue  
Cambridge, MA 02139  
Ph: 617.253.5668 Fax: 617.253.1690  
Email: docs@mit.edu  
<http://libraries.mit.edu/docs>

## **DISCLAIMER OF QUALITY**

Due to the condition of the original material, there are unavoidable flaws in this reproduction. We have made every effort possible to provide you with the best copy available. If you are dissatisfied with this product and find it unusable, please contact Document Services as soon as possible.

Thank you.

**Some pages in the original document contain color pictures or graphics that will not scan or reproduce well.**