

**STRUCTURAL AND MECHANISTIC INSIGHTS FROM HIGH
RESOLUTION CRYSTAL STRUCTURES OF THE TOLUENE-4-
MONOOXYGENASE CATALYTIC EFFECTOR PROTEIN,
NAD(P)H OXIDASE AND CHOLINE OXIDASE**

A Thesis
Presented to
The Academic Faculty

by

George T. Lountos

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry & Biochemistry

Georgia Institute of Technology
December 2005

**STRUCTURAL AND MECHANISTIC INSIGHTS FROM HIGH
RESOLUTION CRYSTAL STRUCTURES OF THE TOLUENE-4-
MONOOXYGENASE CATALYTIC EFFECTOR PROTEIN,
NAD(P)H OXIDASE AND CHOLINE OXIDASE**

Approved by:

Dr. Allen M. Orville, Advisor
School of Chemistry & Biochemistry
Georgia Institute of Technology

Dr. Loren Dean Williams
School of Chemistry & Biochemistry
Georgia Institute of Technology

Dr. Christoph J. Fahrni
School of Chemistry & Biochemistry
Georgia Institute of Technology

Dr. Andreas S. Bommarius
School of Chemical & Biomolecular
Engineering
Georgia Institute of Technology

Dr. Giovanni Gadda
Departments of Chemistry & Biology
Georgia State University

Date Approved: 11/11/2005

This thesis is dedicated to my mother, father, and brother for all of their support.

Also, this thesis is dedicated to the memory of my grandparents, George & Vassiliki Lountos, and George & Konstantina Panakis, and Uncle Gus Lountos

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my research advisor, Dr. Allen Orville for his guidance and support over these past five years. Dr. Orville has taught me a lot and his guidance was essential for the success of these projects.

I would also like to thank the members of my committee, Dr. Loren Williams, Dr. Andreas Bommarius, Dr. Christoph Fahrni, and Dr. Giovanni Gadda for serving on my committee and for providing valuable criticism, discussions, advice, and suggestions. I have had the opportunity to work closely with Dr. Bommarius and Dr. Gadda on collaborative projects and I have learned a lot of new concepts as a result.

I would like to thank Dr. Brian Fox of the University of Wisconsin-Madison and his research group for providing purified T4moD used for crystallization and for performing the biochemical characterization of the T4moD mutants. Dr. Fox also worked closely with us on publishing the structures of T4moD and provided many insights and suggestions. Dr. Bettina Riebel, Dr. Andreas Bommarius, and William Wellborne provided purified NAD(P)H oxidase for the crystallographic studies of the enzyme along with many discussions. Mahmoud Ghanem of Georgia State University was very helpful in obtaining the MALDI-TOF spectra in the NAD(P)H project. I would also like to thank Fan Fan and Dr. Giovanni Gadda of Georgia State University for providing purified choline oxidase. Fan was very helpful in the initial crystallization experiments of the enzyme and thus is an “honorary crystallographer.” Dr. Rajeev Prabhakar of Emory University is currently working on theoretical studies of the FAD coordinates in choline oxidase and thus has provided helpful insights.

I would also like to express my gratitude to Jamie Summerour who served as Instructor for the analytical chemistry labs that I served as a TA for many years. She was a pleasure to work with and was very helpful in allowing me flexibility in my teaching schedule when I needed to be away from class for research, especially when I was away on synchrotron data collection trips during the semester.

I would like to also thank the staff at SER-CAT who were a wonderful resource. Almost all of the data in my thesis was collected at the SER-CAT facilities which are truly first-class. Dr. Zhongmin Jin was very helpful in the mail-in crystallography program and I would also like to thank Prof. Andy Howard who spent a lot of time teaching us new concepts in data collection and data processing. Marie Graham was also very helpful in scheduling our visits and providing administrative support.

Many thanks are in order also to the members of Dr. Williams' group. Dr. Shelly Howerton, Dr. Kris Woods, Dr. Tinoush Moulaei, Dr. Seiji Komeda, Tatsuya Maehigashi, and Derrick Watkins were a constant source of support or advice.

The members of my research group Linda Manning and Akanksha Nagpal also deserve special recognition. I have learned a lot from interacting with them and will always cherish their friendship and memories.

Finally, I would like to thank my parents and brother for providing a loving and supporting family for which I am deeply grateful to have.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF TABLES	x
LIST OF FIGURES	xi
LIST OF SYMBOLS AND ABBREVIATIONS	xiv
SUMMARY	xv
 <u>CHAPTER</u>	
1 Crystallization and Preliminary Analysis of Native and N-Terminal Truncated Isoforms of Toluene-4-Monooxygenase Catalytic Effector Protein	1
Abstract	1
Introduction	2
Materials and Methods	3
Protein Expression and Purification	3
Crystallization and X-ray Data Collection	4
Results and Discussion	11
Conclusions	12
References	13
 2 Crystal Structures and Functional Studies of T4moD, the Toluene-4-monooxygenase Catalytic Effector Protein	 15
Abstract	15
Introduction	16
Materials and Methods	19
Construction of T4moD Isoforms	19

Preparation and Characterization of T4moD	19
Crystallization and X-ray Diffraction Data Collection	20
Crystal Structure Determination	20
Results	22
Crystal Structure Determination	22
Overall Structure Description	24
N-Terminal Region	32
Catalytic Activity of N-Terminal Deletions of T4moD	32
Discussion	33
Comparison of T4moD Crystal Structures and NMR Structures	33
Environment of Asn34	36
Environment of Glu64	39
N-Terminal Region	42
Role of N-Terminal Residues in Effector Proteins	42
Fidelity Between Catalytic Effector Proteins	44
Conclusion	44
References	46
3 Crystallization and Preliminary Analysis of a Water-forming NADH Oxidase from <i>Lactobacillus sanfranciscensis</i>	50
Abstract	50
Introduction	50
Materials and Methods	53
Protein Expression and Purification	53
Crystallization and X-ray Data Collection	54
Results and Discussion	55
Conclusion	58

References	59
4 X-ray Crystal Structure of NAD(P)H Oxidase From <i>Lactobacillus sanfranciscensis</i> at 1.8 Å Resolution	63
Abstract	63
Introduction	64
Materials and Methods	69
Preparation, Crystallization, & Data Collection	69
Crystal Structure Determination	70
Mass Spectrometry	74
Results	74
Overall Structure Description	74
Binding Mode of High Affinity Ligand	78
Active Site Architecture	82
Discussion	84
Structural Homologs of <i>L.san</i> -Nox2	84
Comparison of <i>L.san</i> -Nox2 and NADH Peroxidase Crystal Structures	87
Comparison of <i>L.san</i> -Nox2 and Human Glutathione Reductase	91
Substrate Delivery	94
Orientation of Tyr159	96
O ₂ Access and Reactivity	98
Mechanistic Insights	99
Conclusion	103
References	104
5 Structural and Mechanistic Insights from the Crystal Structure of Choline Oxidase	110

Abstract	110
Introduction	111
Materials and Methods	124
Enzyme Preparation, Purification, Crystallization, & X-ray Data Collection	124
Crystal Structure Determination	127
Results	130
Crystal Structure Determination	130
Overall Structure Determination	134
The Flavin Binding Site	137
Substrate Binding Site	145
Discussion	152
Identification of Putative Substrate Binding Site	152
The FAD in Choline Oxidase	157
Correlation of the Structure to the Proposed Mechanism	164
Conclusion	169
References	171
VITA	179

LIST OF TABLES

	Page
Table 1.1: Data Collection Statistics for SeMet-T4moD	7
Table 1.2: Data Collection Statistics for T4moD	8
Table 1.3: Phasing Statistics for SeMet-T4moD	10
Table 2.1: Data Collection & Model Refinement Statistics for T4moD	23
Table 2.2: Root Mean Squared differences between the T4moD and X-ray structures and Related NMR structures	25
Table 2.3: Hinge Angle Calculated Between β -sheets of Effector Proteins	28
Table 2.4: Compariosn of Catalytic Properties and Regiospecificity of Natural and N-Terminal Deleted Isoforms of Toluene-4-Monooxygenase Effector Protein	43
Table 3.1: Data Collection Statistics for NADH oxidase	57
Table 4.1: Data Collection Statistics for <i>L.san</i> -Nox2	72
Table 4.2: Crystallographic Refinement Statistics for <i>L.san</i> -Nox2	73
Table 4.3: Structural Alignments of <i>L.san</i> -Nox2 and Homologs	85
Table 5.1: X-ray Diffraction Data Collection Statistics for Choline Oxidase	132
Table 5.2: Refinement Statistics for Choline Oxidase	133
Table 5.3: Selected Bond Lengths and Angles in FAD Isoalloxazine Rings	143

LIST OF FIGURES

	Page
Figure 1.1: The toluene-4-monooxygenase enzyme complex	3
Figure 1.2: Crystals of Δ N10-T4moD	6
Figure 2.1: Comparison of the X-ray crystal structures of the T4moD variants	25
Figure 2.2: The RMS differences and B-factor profiles for the main chain atoms in each residue in chain A & B of T4moD	26
Figure 2.3: Primary Sequence and Secondary Structure Alignment for T4moD and other effector proteins	28
Figure 2.4: A divergent stereoview of the environment around the hydrophobic cavity of T4moD	29
Figure 2.5: The electrostatic potential mapped onto the solvent exposed surface of Δ N4-T4moD	31
Figure 2.6: Divergent stereoview of the superimposed X-ray and NMR structures of T4moD	34
Figure 2.7: Differences in hydrogen-bonding patterns observed in the crystal and NMR structures of T4moD surrounding residue Asn34	38
Figure 2.8: Differences in hydrogen-bonding patterns observed in the crystal and NMR structures of T4moD surrounding several charged residues	41
Figure 3.1: Reactions catalyzed by lactic acid bacteria and other facultative or strict anaerobic bacteria.	52
Figure 3.2: Crystals of <i>L.sanfranciscensis</i> NADH oxidase	56
Figure 4.1: The overall reactions catalyzed by <i>L.san</i> -Nox2	67
Figure 4.2: Some <i>in vivo</i> forms of cysteine and its common modifications observed in antioxidant enzymes or proteins.	76
Figure 4.3: The X-ray crystal structure of NAD(P)H oxidase (<i>L.san</i> -Nox2) refined to 1.8 Å resolution	76
Figure 4.4: Overlay of the FAD-binding and the NAD(P)H binding domains of <i>L.san</i> -Nox2	77

Figure 4.5: UV-Vis spectrum of the supernatant from heat denatured <i>L.san</i> -Nox2	79
Figure 4.6: Negative-ion mode MALDI-TOF spectrum of the supernatant of denatured <i>L.san</i> -Nox2	80
Figure 4.7: Electron density maps for the bound ADP molecule	81
Figure 4.8: A divergent stereoview of the binding mode of ADP within the active site of <i>L.san</i> -Nox2	82
Figure 4.9: Structural analysis of the redox state of Cys42	83
Figure 4.10: Stereoview of the active site residues surrounding the <i>si</i> -face of the FAD and Cys42-SOH.	84
Figure 4.11: Primary sequence and secondary structure alignment of <i>L.san</i> -Nox2 and related homologs	86
Figure 4.12: A divergent stereoview of the structural overlays of the C _α backbone	89
Figure 4.13: Structural overlay of the <i>L.san</i> -Nox2 and NADH peroxidase active sites	91
Figure 4.14: Surface representations of the monomers of <i>L.san</i> -Nox2 and hGr	93
Figure 4.15: Surface representations of <i>L.san</i> -Nox2	96
Figure 4.16: Structural overlays of the substrate-binding pocket in <i>L.san</i> -Nox2 and NADH peroxidase	97
Figure 4.17: Hypothetical models of the FAD-C4a:H ₂ O ₂ complexes	99
Figure 4.18: Structure-based reaction mechanism for <i>L.san</i> -Nox2	102
Figure 5.1: A structure-based hypothesis for reactivity in flavoenzymes	113
Figure 5.2: The reaction catalyzed by choline oxidase	116
Figure 5.3: The asynchronous hydride transfer mechanism for choline oxidase with the alkoxide intermediate stabilized by His466	119
Figure 5.4: The three proposed mechanisms for the oxidation of unpolarized alcohols by members of the GMC family	121
Figure 5.5: The reaction mechanism for the oxidation of choline to glycine betaine by choline oxidase based on mechanistic and structural studies	123
Figure 5.6: A view of crystals of choline oxidase under polarized light	126

Figure 5.7: The three-dimensional structure of choline oxidase refined to 1.86 Å resolution	136
Figure 5.8: A view of the $2mF_o-DF_c$ and mF_o-DF_c electron-density maps for the FAD cofactor and surrounding residues during the course of structure refinement	138
Figure 5.9: Approximately orthogonal views of the final 1.86 Å $2mF_o-DF_c$ electron density maps for the FAD isoalloxazine ring	141
Figure 5.10: Approximately orthogonal views of the 1.86 Å electron density maps for the FAD isoalloxazine ring in choline oxidase	142
Figure 5.11: A divergent stereoview of the environment surrounding the FAD isoalloxazine ring in choline oxidase	145
Figure 5.12: A view of the monomeric structure of choline oxidase highlighting the loop region that is comprised of residues 64-95.	146
Figure 5.13: A representation of the well-defined cavity located in the substrate-binding domain of choline oxidase	148
Figure 5.14: A view of the $2mF_o-DF_c$ electron density for the bound DMSO	149
Figure 5.15: A view of the molecular surface of choline oxidase dimer with the solvent channel that protrudes into the active site	151
Figure 5.16: A representation of the solvent channel that protrudes toward the cavity in the substrate-binding domain.	152
Figure 5.17: A divergent stereoview of the active site of choline oxidase with manually docked choline	155
Figure 5.18: The proposed reaction mechanism involved in the generation of the FAD-C4a adduct in the crystal of choline oxidase	161
Figure 5.19: Gas-phase DFT calculations using the FAD C4a-O ₂ ⁻ model with a two-electron reduced flavin	162
Figure 5.20: Gas-phase DFT calculations using the FAD C4a-OH model with a two-electron reduced flavin	163
Figure 5.21: A comparison of the enzyme-bound FAD C4a-O ₂ ⁻ adduct from the X-ray structure and the enzyme-bound FAD C4a-OH model obtained after DFT calculations.	164
Figure 5.22: A proposed structural role for His351 in the asynchronous hydride transfer mechanism	168

LIST OF SYMBOLS AND ABBREVIATIONS

Å	Angstroms
K	Kelvin
CHCA	α -cyano-4-hydroxycinnamic acid
ESI	Electrospray Ionization
hGR	Human Glutathione Reducatase
<i>L.san</i> -Nox2	NAD(P)H Oxidase
MAD	Multiple Anomalous Dispersion
MALDI-TOF	Matrix Assisted Laser Desorption Ionization-Time of Flight
NPX	NADH peroxidase
PCR	Polymerase Chain Reaction
SeMet	Selenomethionine
T4MO	Toluene-4-monooxygenase
T4moD	Toluene-4-monooxygenase Catalytic Effector

SUMMARY

The following thesis describes three bodies of work that use X-ray crystallography to obtain structural insights into the structure-function relationship of three different enzymes. The three enzymes that were studied in this thesis include the toluene-4-monooxygenase catalytic effector protein from *Pseudomonas mendocina*, NAD(P)H oxidase from *Lactobacillus sanfranciscensis*, and choline oxidase from *Arthrobacter globiformis*.

The first two chapters of this thesis focus on the determination of the crystal structure of the toluene-4-monooxygenase catalytic effector protein from the toluene-4-monooxygenase, a four component protein complex that catalyzes the regiospecific, NADH-dependent hydroxylation of toluene to yield *p*-cresol. High-resolution crystal structures (up to 1.7 Å resolution) were obtained of the wild-type enzyme and two engineered variants missing either four or ten amino acids from the N-terminal end of the enzyme. The goal of the project was to obtain structural information to complement biochemical studies that were initiated to study the role of the N-terminal sequence in effector protein catalysis. Results from the crystallographic studies indicate that there are significant structural differences between the X-ray crystal structure of T4moD and the structure previously determined by NMR spectroscopy. Additionally, deletion of the N-terminal residues of the enzyme does not appear to result in any significant structural deviations relative to the wild-type enzyme and biochemical studies indicate that the N-terminal residues are not essential for catalysis. The high-resolution X-ray structures presented here have helped to define potential differences in electrostatic surfaces that may govern the feasibility of protein-protein interactions. Moreover, the present

structures reveal a single, well-defined cavity of size suitable for toluene binding near a region of the protein surface that has substantial different electrostatic properties among the effector protein family members. These results give new information and raise important new questions about these small, enigmatic components of the diiron hydroxylase enzyme complexes.

Chapters three and four present the first crystal structure of a water-forming NAD(P)H oxidase. The enzyme is of considerable interest as it catalyzes the oxidation of two equivalents of NAD(P)H and reduces one equivalent of oxygen to yield two equivalents of water, without releasing hydrogen peroxide after the reduction of the first equivalent of NAD(P)H. The 1.8 Å resolution structure reveals the presence of a redox active cysteine residue in the active site that exists as a sulfenic acid and plays an important mechanistic role by reducing hydrogen peroxide to water, thus preventing the release of hydrogen peroxide from the active site. Additionally, a tightly bound ADP cofactor was discovered in the enzyme structure which occupies the binding pocket of the NAD(P)H binding domain of the enzyme and is hypothesized to play an important role in influencing the dual substrate specificity exhibited by NAD(P)H oxidase.

Finally, chapter five presents the crystal structure of choline oxidase refined to a resolution of 1.86 Å. Choline oxidase catalyzes the four-electron oxidation of choline to glycine betaine via two sequential, FAD-dependent reactions. The study of choline oxidase is of considerable interest for medical and biotechnological applications, since intracellular accumulation of glycine betaine allows normal function under conditions of hyperosmotic and temperature stress in pathogenic bacteria. Consequently, the study of choline oxidase has potential for the development of therapeutic agents that inhibit

glycine betaine biosynthesis and render pathogenic bacteria susceptible to either conventional treatments or the innate immune system, and for the engineering of drought and temperature resistance in economically relevant crops. The crystal structure reveals the presence a well-defined cavity within the active site that allows for the identification of the putative binding site for choline and residues involved in substrate-binding and catalysis. Additionally, the structure reveals that the FAD adopts a highly unusual conformation and contains a C4a-adduct which is proposed to be either an FAD C4a-O₂⁻ or FAD-C4a-OH complex. These results provide the first direct observation of an oxygen reaction intermediate in any flavoenzyme oxidase by X-ray crystallography. Additionally, the structural data obtained from this research allow for new insights into the catalytic mechanism of choline oxidase.

CHAPTER 1

**CRYSTALLIZATION AND PRELIMINARY ANALYSIS OF NATIVE
AND N-TERMINAL TRUNCATED ISOFORMS OF TOLUENE-4-
MONOOXYGENASE CATALYTIC EFFECTOR PROTEIN**

Abstract

¹Single crystals have been obtained of the toluene-4-monooxygenase catalytic effector protein, the SeMet-enriched protein and a truncated isoform missing ten amino acids from the N-terminus. Complete X-ray diffraction data sets have been collected and analyzed to 2.0, 3.0, and 1.96 Å resolution for the native, SeMet and truncated isoform crystals, respectively. The native and SeMet proteins crystallized in space group $P6_122$ (unit cell parameters $a = b = 86.41 \pm 0.15$, $c = 143.90 \pm 0.27$ Å), whereas the truncated isoform crystallized in space group $P2_13$ ($a = b = c = 86.70 \pm 0.47$ Å). Matthews coefficient calculations suggest either two or three molecules per asymmetric unit in the $P6_122$ space group and two molecules per asymmetric unit in the $P2_13$ space group. Experimental phases from MAD analysis of the SeMet isoform and molecular replacement of the truncated isoform confirm the presence of two molecules per asymmetric unit in each case. These crystallographic results are the first available for the evolutionary related but functionally diversified catalytic effector proteins from the multicomponent diiron monooxygenase family.

¹ Adapted from Orville, A.M., Studts, J.M., Lountos, G.T., Mitchell, K.H., and Fox, B.G. (2005) Crystallization and preliminary analysis of native and N-terminal truncated isoforms of toluene-4-monooxygenase catalytic effector protein. *Acta Cryst. D59*, 572-575.

Introduction

Toluene 4-monooxygenase (T4MO) from *Pseudomonas menodocina* KR1 is a multicomponent enzyme complex (1, 2) that exhibits remarkably high regioselectivity for the NADH- and O₂-dependent hydroxylation of toluene to form *para*-cresol (Figure 1.1). The T4MO enzyme complex consists of an NADH oxidoreductase (T4moF, 33 kDa), a Rieske-type ferredoxin (T4moC, 12.5 kDa), a catalytic effector protein (T4moD, 11.6 kDa) and a diiron-containing hydroxylase (T4moH, 212 kDa) with a ($\alpha\beta\gamma$)₂ quaternary structure. T4MO is a member of an evolutionarily related family of oxygenases that includes four subgroups distinguished by their specificity for different natural substrates (3, 4). Methane monooxygenase is the best characterized member of this family and crystal structures are available for hydroxylase components from *Methylococcus capsulatus* Bath and *Methylosinus trichosporium* OB3b (5, 6). Previous single-turnover and peroxide shunt results have shown that the hydroxylase diiron center is the unique site of O₂ reactivity (7). However, protein-protein interactions involving the catalytic effector protein cause changes in the spectroscopic features of the hydroxylase diiron center (8), changes in the lifetimes of reactive intermediates (9) and changes in the product distributions observed from substrates capable of yielding more than one product (10, 11). These results underscore the fundamental importance of effector protein-hydroxylase interactions in diiron-enzyme catalysis.

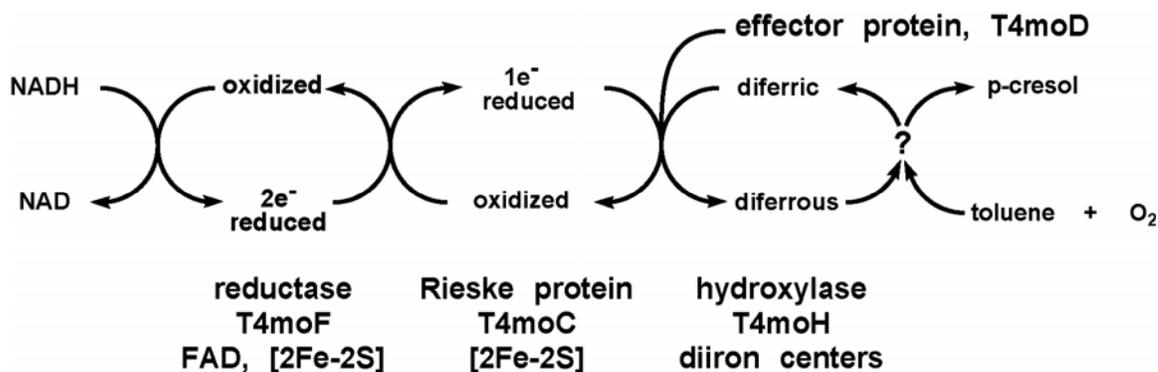


Figure 1.1. The toluene 4-monoxygenase enzyme complex. The catalytic effector protein (T4moD, 11.6 kDa) is the subject of this work. The natural enzyme complex produces >95% yield of *p*-cresol from NADH, O₂, and toluene (11).

Presently, NMR structures of effector proteins are available for four of the 24 members of the family: T4MO (T4moD) (3), phenol hydroxylase (DmpM) (12), and methane monooxygenase (MmoB) from *Methylosinus trichosporium* OB3b (13) and *Methylococcus capsulatus* Bath (14). These proteins represent three of the four functional classes identified for this enzyme family (3). As a group, these proteins exhibit similar secondary-structure topology but significant variability in tertiary structure that may be related to functional divergence and/or an incomplete set of NMR distance restraints. In this report, we describe the crystallization and preliminary analysis of crystals of native T4moD, a selenomethionine-enriched form (SeMet-T4moD) and an isoform lacking ten residues from the N-terminus (Δ N10-T4moD). The data are of sufficient quality to permit crystal structural determinations for both T4moD and (Δ N10-T4moD), which are in progress.

Material and Methods

Protein Expression and Purification

T4moD was expressed in *Escherichia coli* BL21 (DE3) and purified as previously reported (15). SeMet-T4moD was expressed in the Met auxotroph *E. coli* B834 (DE3) in

a minimal medium augmented with selenomethionine. ESI-MS analysis indicated that the N-terminal selenomethionine was removed during expression and that ~95% incorporation was obtained at the two remaining methionine residues. The Δ N10–T4moD isoform was constructed by PCR using pJDP01 as the template (16). The following oligonucleotides were used: DtruncF (5'-gcttaCATATGaataacgttgaccgattaccg-3') and NtermR (5'-caaggggtatgctagtattgctcagcgg-3'). Capitals indicate the *Nde*I site used for cloning, which also placed the start codon two residues before β -strand 1, the first identifiable secondary-structure element in the NMR structure (3).

Crystallization and X-ray data collection

The initial crystallization studies were performed using screening kits from Hampton Research and were optimized by screening additives and cryoconditions. All crystals were grown using the hanging-drop vapor-diffusion technique with standard 24-well Linbro plates from ~2 μ L of protein (typically in 25 mM MOPS pH 6.75, 7% v/v glycerol, 0.15 M NaCl) mixed with 2 μ L reservoir solution. T4moD crystallized at 298 K with 975 mM sodium/potassium phosphate pH 4.7, 400 mM NaCl, 50 mM succinate pH 5.5 at a protein concentration of approximately 40 mg/mL. The approximately 0.2 x 0.2 x 0.2 mm crystals grew reproducibly within approximately one week under aerobic conditions. Crystals of SeMet-T4moD (approximately 0.15 x 0.15 x 0.15 mm) were grown in an anaerobic chamber (Coy Labs, maintained with a 95% N₂ and 5% H₂ atmosphere) under nearly identical crystallization conditions supplemented with 1 mM dithiothreitol. Single crystals of T4moD were transferred to mother liquor augmented with 30% glycerol as the cryoprotectant immediately prior to flash-freezing in a 100 K cold stream. For Δ N10–T4moD, the reservoir solution contained 2.0 M ammonium sulfate and 5% (v/v) 2-propanol and the drop (10 mg/mL protein) was augmented with 1 μ L of 7.5% (v/v) 1,2,3-heptanetriol. Crystals appeared after approximately 3 days of equilibration at 277 K and reached maximum dimensions of approximately 0.2 x 0.2 x

0.2 mm within one week (Figure 1.2). The Δ N10–T4moD crystals were transferred to Paratone-N (Hampton Research), the excess mother liquor was pulled away and the crystals were flash-frozen by rapid submersion in liquid N₂.

All X-ray diffraction data were collected from crystals held at approximately 100 K. The T4moD data set was collected with 0.7433 Å X-rays at the Advanced Photon Source (APS), Argonne National Laboratory. The four-wavelength MAD data set was collected at beamline 1-5 at the Stanford Synchrotron Radiation Source. An ADSC Quantum 4 detector was used to collect two 16° wedges of unique data and their inverse-beam wedges at the wavelengths indicated in Table 1.1. Each image was exposed for 45 seconds with a 1° oscillation. Diffraction data from Δ N10–T4moD were collected during the commissioning of beamline 22-ID by the South East Regional Collaborative Access Team (SER-CAT) at APS using a MAR CCD 165 detector (Table 1.2). Each image was collected with a 1 second exposure time and a 1° oscillation range. The T4moD data set was processed with DENZO and SCALEPACK (17). The SeMet-T4moD and Δ N10–T4moD data set was processed with MOSFLM and SCALA from the CCP4 suite of programs (18).

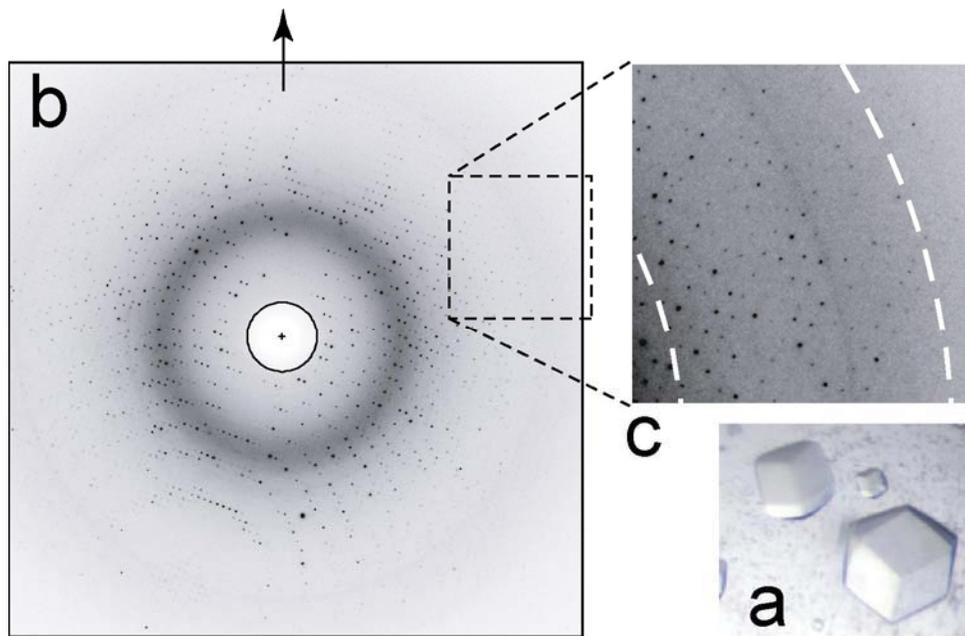


Figure 1.2. A) Crystals of Δ N10-T4moD measuring approximately 0.3 mm along the diagonal. B) the X-ray diffraction pattern obtained with 1 second exposure and 1° oscillation along the vertical axis (vertical line). The data were collected at SER-CAT beamline 22-ID with a MAR CCD 165 area detector and a 103 mm crystal to film distance. C) An expanded and contrast-adjusted view of the diffraction pattern between 2.0 and 1.5 Å resolution (inner and outer dashed arcs), respectively.

Table 1.1 Data Collection Statistics for Se-Met T4moD

Crystal	SeMet-T4moD ^{a,b}			
	inflection (f')	peak (f')	remote 1	remote 2
Wavelength (Å)	0.979880	0.979571	0.925256	1.068830
Resolution range (Å)	33 - 3.0	33 - 3.0	33 - 3.0	37 - 3.0
High res. shell (Å)	3.08 - 3.0	3.08 - 3.0	3.08 - 3.0	3.08 - 3.0
Space group	<i>P</i> 6 ₁ 22			
Total reflections	44128	44088	44703	39510
Unique reflections	6781	6791	6800	6293
Multiplicity (overall) ^c	6.5 (5.9)	6.5 (6.0)	6.6 (6.8)	6.3 (5.3)
Completeness (%, overall)	98.9 (98.9)	98.9 (91.3)	99.2 (99.2)	99.0 (91.0)
Anomalous comp. (%, overall)	98.1	98.0	98.6	88.3
R _{sym} (%, overall) ^d	7.1 (27.5)	7.1 (29.0)	7.2 (32.7)	7.5 (33.0)
I/σ(I) (overall) ^e	4.8 (2.7)	4.8 (2.6)	4.7 (2.3)	5.0 (2.5)

^a Unit cell dimensions for *P*6₁22 crystals are $a = b = 86.41 \pm 0.15$ Å, $c = 143.9 \pm 0.27$ Å; $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$. ^b Data collected at beamline 1-5 of SSRL with an ADSC Quantum 4 detector. ^c Values in parenthesis are for the highest resolution shell of data ^d R_{sym} (I) gives the average agreement between the independently measured intensities such as $\sum_h \sum_i |I_i - I| / \sum_h \sum_i I$, where I is the mean intensity of the *i* observations of reflection *h*. ^e I/σ(I) is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

Table 1.2 Data Collection Statistics for T4moD^a crystals

Crystal	T4moD ^b	Δ N10-T4moD ^c	Δ N10-T4moD ^d	Δ N10-T4moD ^d
Wavelength (Å)	0.7433	1.5418	1.000	1.000
Resolution range (Å)	37 - 2.05	23 - 2.2	16 - 1.96	17 - 2.00
High res. shell (Å)	2.11 - 2.05	2.32 - 2.20	2.07-1.96	2.11-2.00
Space group	<i>P</i> 6 ₁ 22	<i>P</i> 2 ₁ 3	<i>P</i> 2 ₁ 3	<i>P</i> 2 ₁ 3
Total reflections	34207	235752	339875	333434
Unique reflections	19009	11161	15889	15236
Multiplicity (overall) ^e	1.8 (1.2)	21.1 (21.1)	21.4 (21.0)	21.9 (21.4)
Completeness (% , overall)	96.5 (78.2)	99.9 (99.9)	99.8 (99.8)	99.8 (99.8)
R _{sym} (% , overall) ^f	6.6 (29.4)	10.9 (30.0)	6.4 (31.7)	5.8 (34.3)
I/ σ (I) (overall) ^g	6 (2.5)	6.2 (2.4)	9.6 (2.3)	10.1 (2.2)

^a Unit cell dimensions for *P*6₁22 crystals are $a = b = 86.41 \pm 0.15$ Å, $c = 143.9 \pm 0.27$ Å; $\alpha = \beta = 90^\circ$, $\gamma = 120^\circ$ and for *P*2₁3 crystals $a = b = c = 86.7 \pm 0.47$ Å, $\alpha = \beta = \gamma = 90^\circ$. ^b Data collected at beamline 1-5 of SSRL with an ADSC Quantum 4 detector. ^c Data collected with rotating Cu anode and an MSC Raxis IV++ detector. ^d Data collected at beamline 22-ID at APS with a marCCD 165 detector. ^e Values in parenthesis are for the highest resolution shell of data ^eR_{sym}(I) gives the average agreement between the independently measured intensities such as $\sum_h \sum_i |I_i - I| / \sum_h \sum_i I$, where I is the mean intensity of the i observations of reflection h. ^g I/ σ (I) is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

The four-wavelength MAD data set of SeMet-T4moD was further scaled and analyzed with FHSCALE and SCALEIT from the CCP4 suite of programs (18). The Se atoms were located with SOLVE (19) using data from all four wavelengths and X-ray cross-sectional estimates of anomalous scattering factors at each wavelength as determined with CROSSEC (18). The Se sites were refined with SHARP (20) to produce the experimental phases currently in use for model building and refinement. The resulting phasing statistics from SHARP are presented in Table 1.3.

Table 1.3: Phasing Statistics for SeMet-T4moD^a				
Data set	inflection (f')	peak (f'')	remote 1	remote 2
Wavelength (Å)	0.979880	0.979571	0.925256	1.068830
Resolution range (Å)	33 - 3.0	33 - 3.0	33 - 3.0	37 - 3.0
R _{cullis} centric (iso) ^b	0.35	0.34	---	0.78
R _{cullis} acentric (iso/ano)	0.39 / 0.93	0.39 / 0.78	--- / 0.83	0.74 / 0.97
R _{kraut} centric (iso/ano) ^c	0.06 / 0.12	0.07 / 0.13	0.08 / 0.13	0.09 / 0.14
R _{kraut} acentric (iso/ano)	0.02 / 0.04	0.02 / 0.04	0.03 / 0.04	0.03 / 0.04
Phasing power, centric (iso) ^d	4.11	4.00	0	0.12
Phasing power, acentric (iso/ano)	4.43 / 1.04	4.11 / 1.86	0 / 1.65	0.12 / 0.86
Figure of merit ^e	0.521 (1593 centric refs.); 0.519 (5181 acentric refs.)			
	0.896 (after solvent flattening, 6815 refs.)			
<p>^a Compiled from SHARP (de La Fortelle & Bricogne, 1997).</p> <p>^b $R_{\text{cullis}} = \langle \text{phase-integrated lack of closure} \rangle / \langle F_{\text{ph}} - F_{\text{p}} \rangle$, where F_{ph} is the structure factor obtained from the inflection, peak or remote 2 data sets and F_{p} is obtained from the 0.925256 Å data set. The isomorphous and anomalous differences are designated as iso and ano, respectively.</p> <p>^c $R_{\text{kraut}} = \Sigma \varepsilon_{\text{iso}} / \Sigma \Delta_{\text{iso}}$ for isomorphous differences and $\Sigma \varepsilon_{\text{ano}} / \Sigma \Delta_{\text{Bijvoet}}$ for anomalous differences, where ε_{iso} and ε_{ano} are the isomorphous and anomalous phase-integrated lack of closure respectively, and Δ_{iso} is the isomorphous difference and Δ_{Bijvoet} is the Bijvoet difference.</p> <p>^d Phasing power = $\langle [F_{\text{h(calc)}} / \text{phase-integrated lack of closure}] \rangle$, where $F_{\text{h(calc)}}$ is the calculated structure factor for the three Se atoms.</p> <p>^e Figure of merit statistics describe a confidence level for the calculated centroid phases.</p>				

Results and Discussion

The *tmoD* gene (NCBI accession No. M65106) encodes a 103-residue 11,618 Da protein that contains no Cys or Trp residues (21). There are three Met residues (1, 74, and 103) in T4moD, but the N-terminal Met is efficiently removed during expression in *E. coli* BL21 (DE3). The calculated instability index of 49.7 suggests that T4moD may be “unstable” (22). However, purified T4moD is stable and a high-resolution NMR structure has recently been completed (3). Moreover, previous structural and catalytic studies suggest a potential functional role for the amino-terminal region of the catalytic effector protein families (9, 23). To initiate comparative functional studies of the T4MO complex, a truncated isoform of T4moD was constructed, purified, and crystallized.

The diffraction data from native T4moD and SeMet-T4moD crystals are consistent with space groups $P6_122$ or $P6_522$ ($a = b = 86.41 \text{ \AA} \pm 0.15$, $c = 143.9 \text{ \AA} \pm 0.27$). Matthews coefficient and solvent-content analysis (24) suggest either two ($V_M = 3.5 \text{ \AA}^3 \text{ Da}^{-1}$, 64.8% solvent) or three ($V_M = 2.3 \text{ \AA}^3 \text{ Da}^{-1}$, 47.2% solvent) molecules per asymmetric unit. Although SOLVE did not produce a reasonable result from analysis of the $P6_522$ space group data, three Se sites were identified in the asymmetric unit for the $P6_122$ space group. The overall estimated figure of merit reported from SOLVE was 0.34, with a Z score of 15.22 and peak intensities at the Se sites of 39.5, 39.1, and 26.8 σ . The presence of three well defined Se atoms is consistent with V_M calculations assuming a trimer in the asymmetric unit given that, for example, Met74 is well ordered and that Met103 is disordered in the crystal lattice. However, an alternative postulate of a dimer in the asymmetric unit is tenable by assuming that the two monomers are structurally inequivalent and that one of the C-terminal SeMet residues is disordered. The Se sites were refined with SHARP and the resulting experimental phases and electron density maps clearly show that the asymmetric unit contains two T4moD molecules. The overall folds of the monomers are similar to the NMR structure (3). However, there are at

least two regions where the structure of each monomer appears to differ significantly from the other and from the NMR structure, which includes the amino-terminal regions.

Crystals of the Δ N10-T4moD isoform are colorless non-birefringent cubes (Figure 1.2) and the diffraction data is consistent with the cubic space group $P2_13$ (unit-cell parameters $a = b = c = 86.75 \text{ \AA}$). Despite the different space group, the unit-cell edges of the cubic form are nearly identical to the two short cell edges of the hexagonal crystal form. A typical data set from Δ N10-T4moD consists of 339,875 observations of 15,889 unique reflections in the resolution range 15.83-1.96 \AA . The overall completeness of 99.8% (99.8% in the 2.07-1.96 \AA highest resolution shell) and overall R_{sym} of 6.4% (31.7% in the highest resolution shell) are both excellent. The data are both strong as indicated by the overall $I/\sigma(I)$ of 9.6. Assuming two Δ N10-T4moD monomers in the asymmetric unit, the calculated V_M of the crystals is $2.64 \text{ \AA}^3 \text{ Da}^{-1}$, corresponding to a solvent content of 53%. Molecular replacement is currently under way using the partially refined structure of the appropriately truncated native T4moD model and confirms the Matthews coefficient estimates for this isoform.

Conclusions

High quality crystals of native T4moD, a SeMet enriched isoform, and an isoform with 10 amino acids removed from the N-terminus have been obtained. High resolution X-ray diffraction data have been collected from each isoform which is of sufficient quality to enable the determination of the three-dimensional structure of each isoform.

REFERENCES

1. Whited, G. M., and Gibson, D. T. (1991) *J Bacteriol* 173, 3010-6.
2. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Munck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) *Biochemistry* 35, 9106-19.
3. Hemmi, H., Studts, J. M., Chae, Y. K., Song, J., Markley, J. L., and Fox, B. G. (2001) *Biochemistry* 40, 3512-24.
4. Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Muller, J., and Lippard, S. J. (2001) *Angew. Chem. Int. Ed. Engl.* 40, 2782-2807.
5. Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) *Nature* 366, 537-43.
6. Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D., and Ohlendorf, D. H. (1997) *Protein Sci.* 6, 556-68.
7. Andersson, K. K., Froland, W.A., Lee, S.-K., Lipscomb, J.D. (1991) *New J. Chem.* 15, 411-415.
8. Fox, B. G., Liu, Y., Dege, J. E., and Lipscomb, J. D. (1991) *J Biol Chem* 266, 540-50.
9. Chang, S. L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (2001) *Biochemistry* 40, 9539-51.
10. Froland, W. A., Andersson, K. K., Lee, S. K., Liu, Y., and Lipscomb, J. D. (1992) *J. Biol. Chem.* 267, 17588-97.
11. Mitchell, K. H., Studts, J. M., and Fox, B. G. (2002) *Biochemistry* 41, 3176-88.
12. Qian, H., Edlund, U., Powlowski, J., Shingler, V., and Sethson, I. (1997) *Biochemistry* 36, 495-504.
13. Chang, S. L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (1999) *Biochemistry* 38, 5799-812.
14. Walters, K. J., Gassner, G. T., Lippard, S. J., and Wagner, G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 7877-82.
15. Studts, J. M., and Fox, B. G. (1999) *Protein Expr. Purif.* 16, 109-19.

16. Xia, B., Pikus, J. D., Xia, W., McClay, K., Steffan, R. J., Chae, Y. K., Westler, W. M., Markley, J. L., and Fox, B. G. (1999) *Biochemistry* 38, 727-39.
17. Otwinowski, Z., and Minor, W. (1997) *Methods in Enzymology* 276, 307-326.
18. Collaborative Computational Project, N. (1994) *Acta Cryst. D50*, 760-763.
19. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Cryst. D55*, 849-61.
20. La Fortelle, E. d., and Bricogne, G. (1997) *Methods in Enzymology* 276, 472-494.
21. Yen, K. M., Karl, M. R., Blatt, L. M., Simon, M. J., Winter, R. B., Fausset, P. R., Lu, H. S., Harcourt, A. A., and Chen, K. K. (1991) *J Bacteriol* 173, 5315-27.
22. Guruprasad, K., Reddy, B. V., and Pandit, M. W. (1990) *Protein Eng* 4, 155-61.
23. Brandstetter, H., Whittington, D. A., Lippard, S. J., and Frederick, C. A. (1999) *Chem. Biol.* 6, 441-9.
24. Matthews, B. W. (1968) *J Mol Biol* 33, 491-7.

CHAPTER 2

CRYSTAL STRUCTURES AND FUNCTIONAL STUDIES OF T4MOD, THE TOLUENE 4-MONOOXYGENASE CATALYTIC EFFECTOR PROTEIN²

Abstract

Toluene 4-monooxygenase (T4MO) is a four component complex that catalyzes the regiospecific, NADH-dependent hydroxylation of toluene to yield *p*-cresol. The catalytic effector (T4moD) of this complex is a 102-residue protein devoid of metals or organic cofactors. It forms a complex with the diiron hydroxylase component (T4moH) that influences both the kinetics and regiospecificity of catalysis. Here, we report crystal structures for native T4moD and two engineered variants with either four (Δ N4-) or 10 (Δ N10-) residues removed from the N-terminal at 2.1, 1.7, and 1.9 Å resolution, respectively. The crystal structures have C-alpha root-mean-squared differences of less than 0.8 Å for the central core consisting of residues 11-98, showing that alterations of the N-terminal have little influence on the folded core of the protein. The central core has the same fold topology as observed in the NMR structures of T4moD, the methane monooxygenase effector protein (MmoB) from two methanotrophs, and the phenol hydroxylase effector protein (DmpM). However, the root-mean-squared differences between comparable C-alpha positions in the X-ray structures and NMR structure vary from approximately 1.8 Å to greater than 6 Å. The X-ray structures exhibit an estimated overall coordinate error from 0.095 (0.094) Å based on the R-value (R-free) for the

² The following chapter is adapted from Lountos, G.T., Mitchell, K.H., Studts, J.M., Fox, B.G., & Orville, A.M. (2005) Crystal Structures and Functional Studies of T4moD, the Toluene 4-Monooxygenase Catalytic Effector Protein, *Biochemistry* 44, 7131-7142.

highest resolution Δ N4-T4moD structure to 0.211 (0.196) Å for the native T4moD structure. Catalytic studies of the Δ N4-, Δ N7-, and Δ N10- variants of T4moD show statistically insignificant changes in k_{cat} , K_{M} , $k_{\text{cat}}/K_{\text{M}}$, and K_{I} relative to the native protein. Moreover, there was no significant change in the regiospecificity of toluene oxidation with any of the T4moD variants. The relative insensitivity to changes in the N-terminal region distinguishes T4moD from the MmoB homologues, which each require the ~33 residue N-terminal region for catalytic activity.

Introduction

T4MO is a member of the diiron monooxygenase family of enzymes (1). These evolutionarily related enzyme complexes are classified as either aliphatic or aromatic hydroxylases (2). MMO, a member of the aliphatic monooxygenase class, is the most thoroughly studied diiron monooxygenase and serves as the structural and functional paradigm for the family (3, 4). The aromatic hydrocarbon hydroxylases can be grouped into two subclasses based on operon structure, sequence analysis, and the primary substrate that is oxidized (2). Subclass I contains the toluene-2-monooxygenase/phenol hydroxylase enzymes (T2MO/phenol), while subclass II contains the toluene 4-monooxygenase/benzene monooxygenase enzymes (T4MO/benzene).

All members of the diiron monooxygenase family are multi-component enzyme complexes that consist of either a one- or two-protein electron-transfer chain, a catalytic effector protein that contains no metal ions or organic cofactors, and a terminal hydroxylase (2, 5-7). Specific interactions between the hydroxylase and reductase/ferredoxin components are required for electron transfer during catalysis (3, 5, 7). X-ray crystal structures of terminal hydroxylases MmoH from *Methylococcus capsulatus* (Bath) and *Methylosinus trichosporium* Ob3B and the toluene/*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 have been solved (8-10). Furthermore,

NMR structures of the Rieske ferredoxin T4moC³, the ferredoxin domain, and the flavin domain of *M. capsulatus* (Bath) oxidoreductase have also been solved (11-13).

The effector protein has also been shown to serve an essential role in catalysis through formation of a protein-protein complex with the hydroxylase (14-20). NMR structures of effector proteins from each catalytic effector subclass have been solved (21-24). Although these structures exhibit approximately the same topology of secondary structural elements, there are significant differences in the 3D structures, which may arise in part from the different numbers of NOE restraints used to calculate the solution structures (21-24). The present crystal structures, determined to 2.1 Å resolution and better, allow comparison with the NMR structure of T4moD and other effector proteins.

In each of the previously determined NMR structures, a portion of the N-terminal region was disordered. For the MMO subclass, the disordered region is ~33 residues long. Deletion of the N-terminal end results in the complete loss of catalytic effector activity (25, 26). More specifically, His33, which is five residues before the start of β -strand 1 in the folded core of the protein, has a controlling influence on the rate of formation and decay of transient catalytic intermediates (27). Intriguingly, mutagenesis of His33 gave no apparent effect on steady-state catalysis, which instead appears to be governed by rate-limiting product release (27). Mutagenesis also changed the product specificity (25, 26). As His33 is conserved in all sequenced examples of MmoB, the effector proteins are thought to control internal steps of the multi-step catalytic cycle through formation of specific protein-protein interactions (16, 18, 19, 25-31).

Sequence alignments and structural analysis have shown that there is substantial variation in the length of the disordered N-terminal regions of effector proteins from the various enzyme subclasses (21-26, 32). For example, the N-terminal disordered region of

³ The crystal structure of T4moC has also been determined to 1.48 Å resolution (PDB code 1VM9).

the effector proteins from the aromatic ring monooxygenases are only ~3 (subclass I, T2MO/phenol) or ~13 (subclass II, T4MO/benzene) residues long, respectively. Furthermore, sequence alignments revealed that there was no sequence conservation among these N-terminal sequences and that only four of the effector proteins from subclass II contained a His residue anywhere in the N-terminal sequence. Consequently, the role or requirement of these shorter N-terminal sequences in effector protein catalysis has not yet been established. In the T4moD solution structure, the disordered region consists of the sequence Ser1-Asn12⁴.

Here, we report X-ray crystal structures and functional analysis of natural T4moD and truncated variants with 4,7, and 10 residues deleted from the N-terminal of the native protein. These latter variants are designated Δ N4-, Δ N7-, and Δ N10-T4moD, respectively. The crystal structures of the natural enzyme and the Δ N4- and Δ N10- variants were refined to resolutions between 2.1 and 1.7 Å in two space groups. Collectively, these structures exhibit low root-mean-squared differences, indicating that the protein core is well-ordered, independent of the presence of the N-terminal region, and is not influenced by the different crystal lattice packing environments present in the two space groups. Comparison of the T4moD X-ray crystal structures with the NMR structure reveals substantial agreement with the topology of the fold. However, the high resolution X-ray structures allow for a reassessment of certain structural features proposed from our previous NMR work on T4moD, including hydrogen-bonding interactions, the positions of several side chains, and the nature of an internal cavity. Progressive removal of the N-terminal residues from T4moD is also shown to have only a modest influence on the steady-state kinetic properties and no significant effect on the

⁴ The residue numbering used in this work is based on counting after removal of the N-terminal Met residue, giving Ser1 as the first residue in the mature protein obtained from bacterial expression.

regiospecificity of product formation in the reconstituted T4MO complex. Thus, specific interactions of residues in the N-terminal end are apparently not required for aromatic hydroxylation in the T4moH, in contrast to the essential role for the N-terminal end in the MmoB homologues during methane hydroxylation.

Materials and Methods

Construction of T4moD Isoforms

The Δ N4-, Δ N7-, and Δ N10-T4moD isoforms were created by removal of the indicated number of amino acid residues from the N-terminal region. These isoforms were constructed by PCR using the expression vector pJDP01 (1, 33, 34) as the template, Vent DNA polymerase (New England Biolabs, Beverly, MA), and a Perkin-Elmer Model 9600 thermocycler (Perkin-Elmer, Foster City, CA). The following oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA) and used as forward primers to create the deletions: Δ N4-(5'-agcacaCATATGgatcaggctttacataac-3'); Δ N7-(5'-tggetgatCATATGttacataacaataacg-3'); Δ N10-(5'-gctttaCATATGaataacgttgaccgattatccg-3'). Capital letters indicate the position of an *NdeI* restriction site used to place the start codon at the correct location relative to the ribosome-binding site in pJDP01. The reverse primer was 5'-caaggggttatgctagtattgctcagcgg-3'. To generate PCR products, a 50 μ L reaction was subjected to 30 cycles of denaturation at 94° C for 45 seconds, anneal at 60° C for 30 seconds, and extension at 72° C for 45 seconds. The PCR products were gel-purified and digested with *NdeI* and *BamHI* and cloned into similarly digested pET3a (Novagen, Madison, WI). The final constructs were sequenced using the Big Dye sequencing kit (Perkin-Elmer) at the University of Wisconsin Biotechnology Center.

Preparation and Characterization of T4moD

The T4moD variants and all other T4MO proteins were overexpressed and purified as previously reported (1, 35, 36). Yields of purified T4moD were ~10-15 mg/L of culture medium. The SeMet-labeled T4moD was produced by fed-batch fermentation

in expression host *Escherichia coli* B834(DE3) as previously described (35).

Electrospray ionization mass spectrometry was performed at the University of Wisconsin Biotechnology Center using a Perkin-Elmer Sciex API 365 triple quadrupole mass spectrometer. Steady-state kinetic assays and determinations of product distributions were as previously reported (22, 34). Velocity data for *p*-cresol formation versus effector protein concentration were fit to $v = k_{\text{cat}}[S]/(K_M + [S] + [S]^2/K_I)$ using the NonLinearRegress routine of Mathematica (v.4.0.1.0, Wolfram Research, Inc. Champaign, IL), where k_{cat} is the maximal reaction velocity, $[S]$ is the concentration of effector protein, K_M is the apparent Michaelis constant for formation of the activating complex of effector protein, and K_I is the apparent equilibrium constant for formation of the inhibiting complex of effector protein. The k_{cat} -values are reported as turnover numbers relative to diiron center in the ($\alpha\beta\gamma$) protomer of T4moH.

Crystallization and X-ray Diffraction Data Collection

Single crystals were obtained for the natural, the SeMet-enriched and the Δ N10-variants as previously described (37). Crystals of the Δ N4-variant were obtained from the same conditions used for the Δ N10-variant (37). The X-ray diffraction data were collected from cryo-preserved crystals at approximately 100 K and processed as previously reported with the diffraction data for the Δ N4-variant collected at the SER-CAT facility of the Advanced Photon Source (APS), Argonne National Laboratory (37).

Crystal Structure Determination

The structure of native T4moD was solved with MAD phasing using a four-wavelength data set collected from a SeMet-T4moD crystal (37). Two molecules of T4moD were present in the asymmetric unit. Three Se atoms were located with SOLVE (38) using data between 33 and 3.0 Å resolution and refined with SHARP (39) as previously reported (37). The experimental phases between 33 and 3.0 Å resolution were appended to the 2.09 Å resolution native T4moD data set. Phase extension and solvent-

flattening to 2.09 Å was performed with the CCP4 suite of programs (40). Model building was performed with the program O (41, 42) and refinements were with CNS (43) and/or REFMAC5 (44, 45). Cross-validated $2mF_o-DF_c$ and mF_o-DF_c maps (46) were used to evaluate the model and correct errors. Water molecules were located and refined in the final stages of refinement with ARP/Waters (47) and REFMAC5.

The Δ N4- and Δ N10-T4moD variants crystallized in space group $P2_13$ with unit cell dimensions of $a = b = c = 86.8$ Å. Two protein molecules were present in the asymmetric unit. The structures for the Δ N4- and Δ N10-T4moD isoforms were solved by molecular replacement in CNS with the native T4moD structure as the search model after appropriate truncation of the N-terminal residues, removal of solvent molecules, and assignment of the B -factors to 20 Å². The final structure for Δ N4-T4moD had an R-factor of 0.15 and an R-free of 0.19 with data between 37 and 1.71 Å resolution. The refined Δ N10-T4moD structure had an R-factor of 0.19 and an R-free of 0.25 with data between 16 and 1.96 Å resolution.

Validation of the refined models and Ramachandran analysis were done using SFCHECK (48) and PROCHECK (48-50). Solvent exposed surface areas were calculated with a 1.4 Å probe radius with DMS implemented in the MidasPlus package from Computer Graphics Laboratory, University of California, San Francisco, CA (supported by NIH P41 RR-01081) (51), CNS, CCP4, or Swiss-PDB Viewer (v.3.7b2) (52). Secondary structure assignments were made with KSDSSP (53). The rms difference between models was calculated with CCP4 or Swiss-PDBViewer. Structure figures were prepared using Swiss-PDBViewer and PovRay (v.3.5) or PyMOL (DeLano Scientific LLC, Castro City, CA).

Results

Crystal Structure Determination

The X-ray data collection and refinement statistics are shown in Table 2.1. Although an NMR structure for T4moD was previously determined (22), various search models derived from the solution structure ensemble did not yield a molecular replacement solution for the native crystal structure. Consequently, the SeMet-enriched T4moD was crystallized and used to determine the experimental phases via MAD data analysis techniques. Phasing statistics for SeMet-T4moD are reported in Orville et al. (2003) (37). The 3.0 Å resolution electron density map resulting from refining three Se atoms in space group $P6_122$ was interpretable and used to build the majority of the initial model. The phase-extended, solvent-flattened map at 2.1 Å resolution for native T4moD was then used to complete the model building. The model for native T4moD was then refined against the 2.1 Å resolution data obtained from the native protein. The final refined model for native T4moD has an R-factor of 0.24 and R-free of 0.28 with data between 37 and 2.1 Å resolution.

Table 2.1. Data Collection & Model refinement statistics

T4moD Variant	Native-	ΔN4-	ΔN10-
Data Collection Statistics			
Resolution range (Å)	37-2.05	37-1.71	16-1.96
Highest Resolution Shell (Å)	(2.11-2.05)	(1.80-1.71)	(2.07-1.96)
Unique reflections	19009	23407	15889
Completeness (%) ^a	96.5 (78.2)	99.99 (99.9)	99.8 (99.8)
R _{sym} (%) ^{a,b}	6.6 (29.4)	12.5 (35.1)	6.4 (31.7)
I/ σ (I) ^c	6 (2.5)	3.4 (1.9)	9.6 (2.3)
Refinement Statistics			
Resolution range (Å)	37-2.1	37-1.71	16-1.96
No. of reflections	17059	21007	14262
R-factor	0.243	0.154	0.192
R-free ^d	0.289	0.185	0.249
No. of non-H protein atoms	1579	1473	1404
No. of water molecules	58	238	167
No. of 1,2,3-heptanetriol molecules	0	0	1
Mean B, protein atoms (Å ²)	50.7	14.5	25.0
Mean B, water molecules (Å ²)	49.6	33.1	36.8
Mean B, 1,2,3-heptanetriol molecule (Å ²)	-	-	31.8
RMS deviations from ideal			
Bond lengths (Å)	0.018	0.016	0.019
Bond angles (°)	1.6	1.6	1.7

^a Numbers in parentheses are for the highest resolution shell

^b The average agreement between the independently measured intensities

^c The root-mean-squared value of the intensity measurements divided by their estimated standard deviation

^d Calculated with 10% of the data.

Each of the truncated T4moD variants crystallized in space group $P2_13$ and molecular replacement was used to solve these structures using a search model derived from the native crystal structure. The crystals of Δ N4-T4moD diffracted to the highest resolution (1.7 Å) and the resulting atomic model was refined to a final R-factor of 0.15 and R-free of 0.19. Ramachandran analysis of all the crystal structures showed that over 90% of the residues were located in the most favored region. In the natural T4moD, Arg45 is in the disallowed region in both chain A and B. The residue is also in the

disallowed region for chain A or B of the Δ N4-T4moD and Δ N10-T4moD structures, respectively. The high-resolution structure of Δ N4-T4moD suggests an alternate conformation for Tyr47 in the B chain. At this position, the $2mF_o-DF_c$ electron density maps are consistent with the two most common conformers found in the O library. The X-ray structure of the Δ N10-T4moD variant contains a single 1,2,3-heptanetriol molecule, an additive in the crystallization solution, located at the interface of the two molecules in the asymmetric unit. Each of the heptanetriol hydroxyl groups hydrogen-bond with either a solvent molecule or the carbonyl oxygen atoms of residues Leu80-A, Leu80-B, or Glu77-B.

Overall Structure Description

The three crystal structures of T4moD are similar, especially for the core region (residues 11-98) as illustrated in Figure 2.1. Indeed, the rms difference between C_α atoms in the core is less than 0.8 Å (Table 2.2) despite several differences in secondary structure assignments (discussed below). Figure 2.2B-F shows the rms differences and B -values for the individual residue positions. The secondary structure assignments for each structure are illustrated in Figure 2.3. The structures differ most significantly at the N-terminal region for natural T4moD. In chain A, this region adopts an α -helical structure, designated α N, whereas in the B-chain and for both chains in the Δ N4-T4moD, this region is largely disordered. The other regions exhibiting slight differences are located between the β 1 strand and the α 1 helix and the hairpin turn between strands β 2 and β 3. Each of these regions also corresponds to larger thermal factors, suggesting some inherent structural variability to the fold.

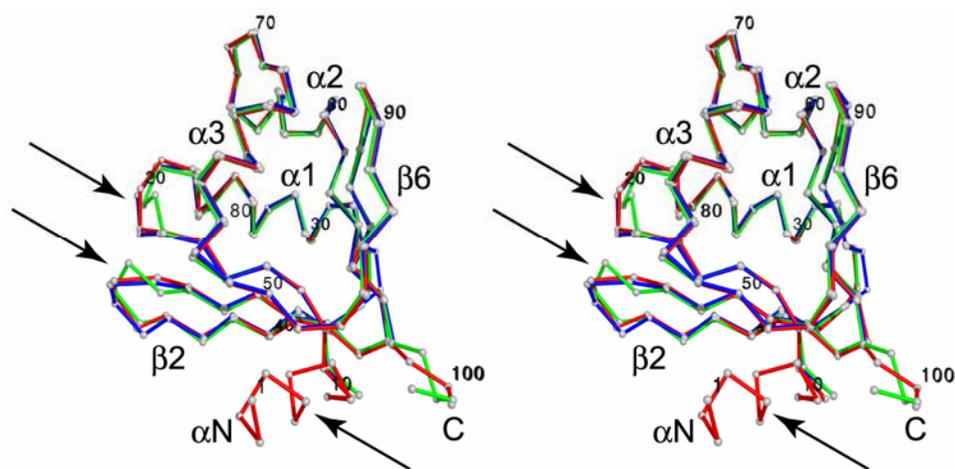


Figure 2.1. Comparison of the X-ray crystal structures of the T4moD variants. A divergent stereo-overlay of the C_{α} backbone trace for natural-(red), $\Delta N4$ - (green), and $\Delta N10$ -T4moD (blue), with arrows highlighting the regions that differ slightly in the three structures.

Table 2: Root Mean Squared Differences between the T4moD X-ray Structures and Related NMR Structures^a

	$\Delta N4$ -A	$\Delta N10$ -A	T4moD NMR	MMO (OB3b)	MMO (Bath)	DmpM
native-A ^b	0.59	0.45	1.88	1.74	1.88	6.96
$\Delta N4$ -A ^c		0.66	1.95	1.77	1.91	6.90
$\Delta N10$ -A ^c			1.91	1.68	2.10	6.86

^aSuperposition of C_{α} atoms using SwissPdb-Viewer and the iterative fit function.

^bresidues 5-100 in either chain A or chain B. ^cResidues 11-98. ^dAlignment with the best representative NMR structure [PDB 1G10, (22)]. ^eThirty-four C_{α} atoms between residues 36-126 in the best representative structure from PDB code 2MOB_10 (21). ^fFifty-nine C_{α} atoms between residues 36-138 in the best representative structure from PDB code 1CKV_11 (24). ^gFifty-four C_{α} atoms between residues 31-84 in the best representative structure from PDB code 1HQI_3 (23). The best representative structure was selected from analysis of the ensemble NMR structure using OLDERADO (54).

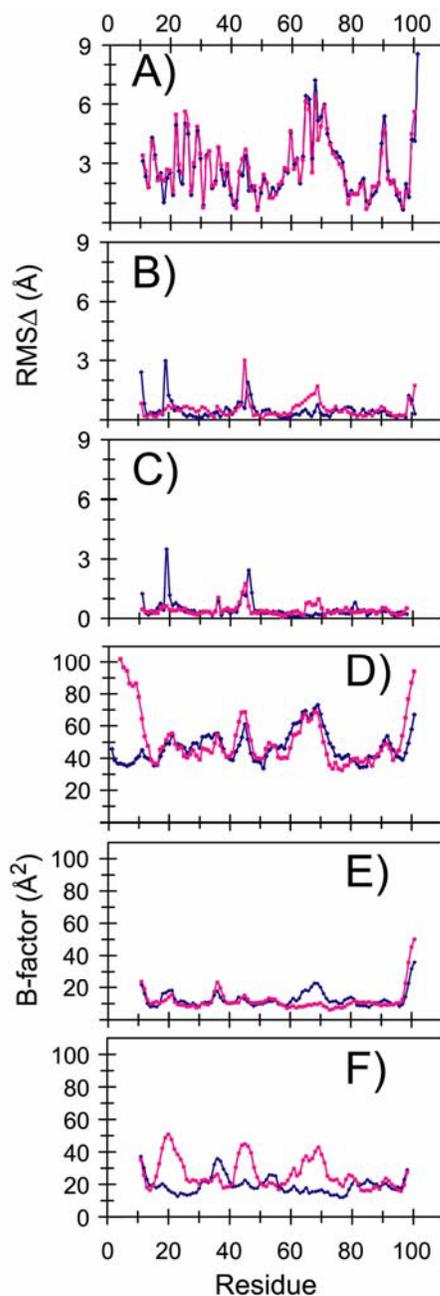


Figure 2.2: The rms differences and B -factors for the main-chain atoms in each residue in chain A (blue) and chain B (magenta). The rms differences between A) $\Delta N4$ -T4moD and the average NMR structure of native T4moD from 1G10 (22), B) $\Delta N4$ -T4moD and the X-ray structure of native T4moD, and C) $\Delta N4$ -T4moD and $\Delta N10$ -T4moD. The average B -factor values for all backbone atoms in each residue for each chain in D) native T4moD, E) $\Delta N4$ -T4moD, and F) $\Delta N10$ -T4moD.

From one perspective, the shape of T4moD is roughly triangular with approximately 24 Å x 20 Å x 23 Å edges when measured at C_α atoms of residues 46, 54, and 60. From the orthogonal perspective, the molecule is roughly rectangular with a width and length of approximately 20 Å by 30 Å when measured at C_α atoms of residues 31-89 and 54-69, respectively. Two sides of the triangular shape are comprised of antiparallel β-sheets, and the backbone trace crosses the corner connecting them three times. A Gly residue is present at each place the backbone chain crosses the vertex of the triangle (Gly36, Gly54, and Gly85) in T4moD. Figure 2.3 also shows a structure-edited sequence alignment of the effector proteins, and the consensus alignment shows that each of these positions is conserved to a differing extent across the entire family. For example, according to the more extensive alignment shown in Hemmi et al. (22), Gly36 is present in all eight of the sequences most closely related to T4moD but not in MMOB and DmpM. In addition, Gly54 is present in 17 of the 24 sequences aligned for the superfamily, while Gly85 is present in all sequences of the superfamily (22). Thus, these Gly residues appear to provide the conformational flexibility required for the backbone to radically alter direction. The angles between the β-sheets range from 74° to 85° in the X-ray structures and from 86° to 113° in the NMR structures (Table 2.3).

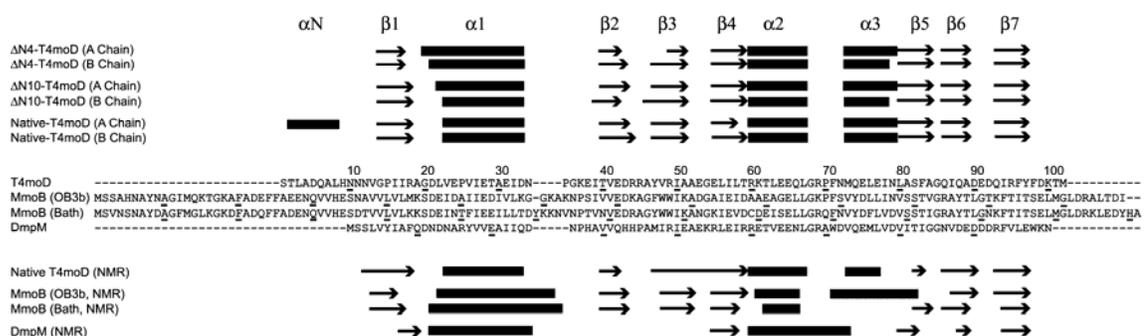


Figure 2.3. Primary sequence and secondary structure alignment for T4moD and other effector proteins. The amino acid sequence alignment was from CLUSTALW using default weights (T4moD gi:45479224; MmoB OB3b, gi:44616 (21); MmoB (Bath), gi:127207 (24); DmpM, gi:118693 (23)). Every 10th residue is underlined with the numbering fixed to that of the native T4moD after post-translational removal of the N-terminal Met. The arrows and rectangles indicate the β -strands and α -helices as assigned from KSDSSP. The secondary structure elements from the T4moD X-ray structures are shown on the top of the figure for both A- and B-chains in the asymmetric units. The secondary structure elements assigned from the NMR structures of T4moD and the other effector proteins are shown on the bottom of the figure (21-24).

Table 2.3. Hinge Angle Calculated Between β -Sheets of Effector Proteins^a

T4moD Variant	X-ray ^b		NMR			
	Chain A	Chain B	T4moD ^c	MMOB (OB3b) ^d	MMOB (Bath) ^e	DmpM ^f
Native-	78°	82°	86°	96°	88°	113°
Δ N4-	74°	85°				
Δ N10-	81°	83°				

^aThe angle calculated was between the β -sheets within the indicated structure. ^bThis work, using the C $_{\alpha}$ of residues 46-52-91 of the given chain. ^cCalculated from 1G10 (22) using residues 46-52-91. ^dCalculated from 2MOB-10 using residues 76-80-119 (21). ^eCalculated from 1CKV_11 using residues 77-81-120 (24). ^fCalculated from 1HQI_3 using residues 36-42-81 (23). The best representative structure was selected from analysis of the ensemble NMR structure using OLDERADO (54).

The interior of T4moD is packed exclusively with hydrophobic residues. Although one small cavity is detected in each chain (Figure 2.4), the electron density maps do not reveal any ordered solvent molecules within the cavity even at high resolution. The volume of the cavity is somewhat variable depending on the chain and structure, but is typically $\sim 50 \text{ \AA}^3$, which is approximately comparable to the volume required to accommodate a toluene molecule (55-59). The borders of the cavity are comprised of Leu22, Pro25, Val26, Leu63, Leu67, Phe71, and Leu76. A few hydrophobic residues are also on the exterior surface of the protein. However, either charged or polar residues typically flank the exterior hydrophobic residues. For example, Glu55, Thr59, Gln86, and Arg95 surround the external residue Ile57. Thus, they provide hydrophobic interactions with Ile57 and polar interactions with solvent molecules. In addition, Ile78, Phe83, and Ile87 are on the exterior surface. These residues are located near the two-fold interface between the two chains in the asymmetric unit. Consequently, their exposure to solvent is limited in the crystal lattice.

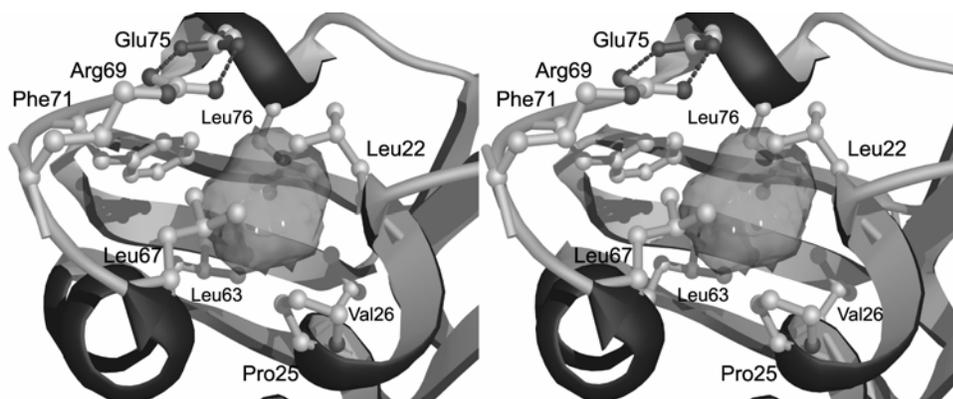


Figure 2.4: A divergent stereoimage of the environment around the hydrophobic cavity illustrated with chain A of the $\Delta N4$ -T4moD crystal structure.

All of the charged residues in T4moD are located on the exterior of the protein, and their distribution is not uniform. Consequently, there are regions with distinctive positive or negative electrostatic surface potential (Figure 2.5). For example, $\alpha 1$ is an amphiphatic helix with Asp21, Glu24, Glu28, Glu31, and Asp33 all on the exterior surface, which yields a strong negative electrostatic surface patch (red in Figure 2.5). The other side of the helix is comprised of Val23, Val26, Ile27, and Ala30, and these residues all project into the hydrophobic core of the protein. In contrast to the negatively charged patch, the region around the hairpin turn between $\beta 2$ and $\beta 3$ strands contains four arginine residues (Arg18, Arg44, Arg45, Arg49), which yields a distinct positive electrostatic patch (blue in Figure 2.5).

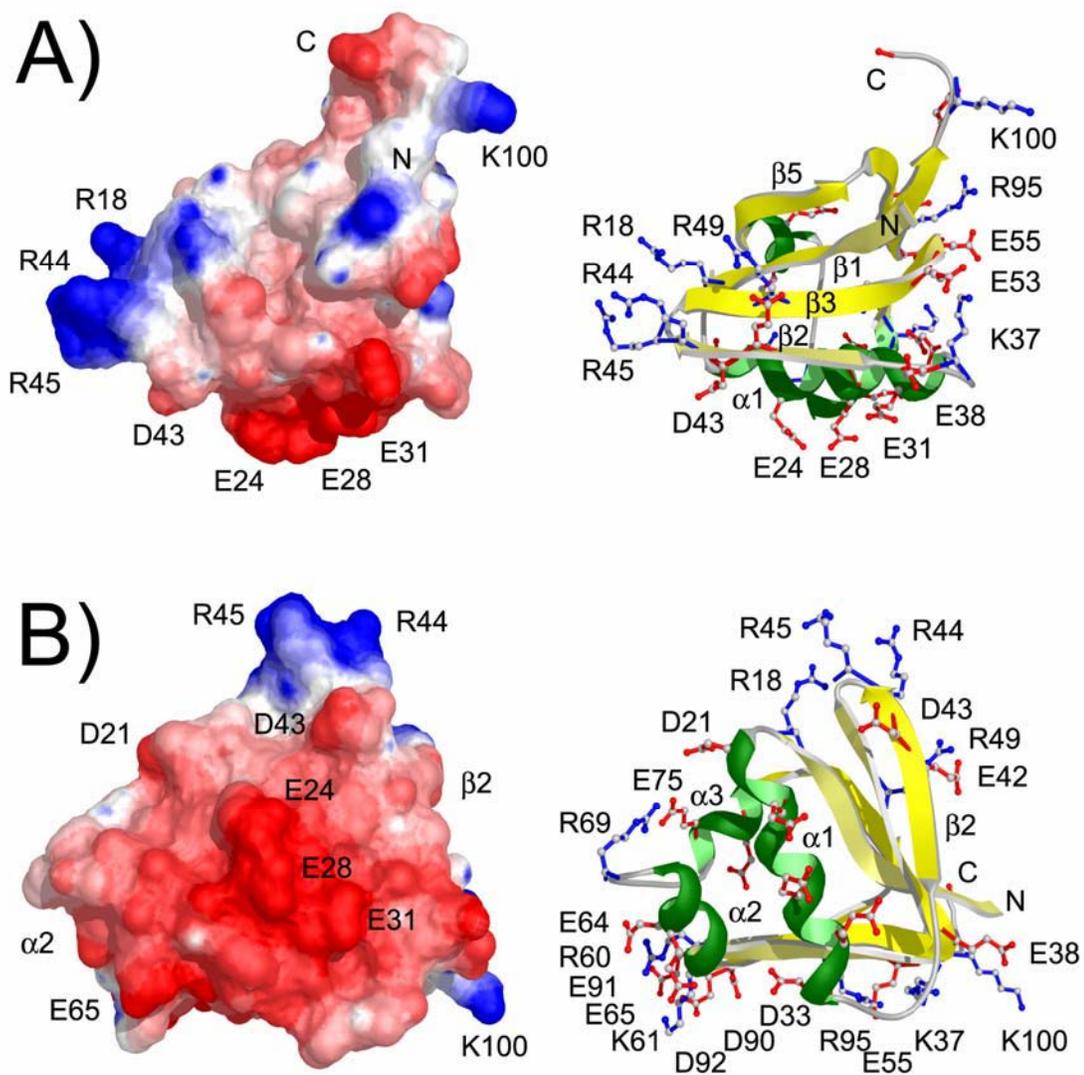


Figure 2.5. The electrostatic potential mapped onto the solvent exposed surface of $\Delta N4$ -T4moD. A) The solvent exposed surface determined with a 1.4 Å probe radius and colored according to the calculated Coulombic electrostatic potentials (-6, -1, and 1.75 are red, white, and blue, respectively). The ribbon drawing of the backbone trace shown on the right is in the identical orientation and scale. The $\alpha 1$ helix is on the bottom, approximately in the plane of the page with its N-terminal end on the left. B) Another orientation of the molecule in which the $\alpha 1$ helix is in the front.

N-Terminal Region

The N-terminal regions in the two chains of the native T4moD structure differ. In addition, there is no electron density visible for Δ N4-T4moD prior to Asn11 in both chains. The electron density for the native T4moD suggests that residues Ser1-Ala7 in chain A adopt an α -helical configuration. The electron density is weaker for the analogous residues in chain B, which suggests the presence of a majority-disordered fraction and a minority-ordered fraction for these residues. The *B*-factors for these regions also differ and are consistent with decreased order for the N-terminal region of the B-chain. (Figure 2.2D). Analysis of the crystal lattice packing reveals that these regions are located in different environments. In chain A, the N-terminal residues are located where four chains pack together. Consequently, only approximately 20% of the surface area for the N-terminal α -helix is exposed to bulk solvent. In contrast, nearly the entire N-terminal region in chain B is exposed to solvent because it projects into a large, solvent-filled space in the crystal lattice. The electron density maps for the two chains in Δ N4-T4moD also suggest that the N-terminal regions are disordered, and analysis of lattice packing revealed that these residues are exposed to the bulk solvent between molecules in the crystal lattice.

Catalytic Activity of N-Terminal Deletions of T4moD

Table 2.4 shows the steady-state catalytic parameters and regioselectivity for toluene oxidation measured from the various N-terminal truncations of T4moD. These characterizations were made according to previously described experimental procedures (33, 34) and included control reactions with the natural isoform. The results show that the disordered N-terminal region of T4moD is not essential for catalysis by the T4MO complex. Thus a 10-residue deletion of the N-terminal gave statistically insignificant changes in k_{cat} , apparent K_M , and k_{cat}/K_M . Likewise, the apparent K_I for inhibition of the toluene hydroxylation reaction was also unaffected by removal of the N-terminal from T4moD. Furthermore, the regioselectivity was nearly unchanged, with the percentage of

p-cresol slightly decreased from 96.2% in native T4moD to 94.5% in Δ N4-T4moD.

These observations also suggest that no single residue in the disordered region of T4moD is likely to be responsible for the observed modest changes in catalytic parameters.

Our previous studies on the reconstitution of T4MO with effector proteins from the subclass I (T2MO/phenol) and subclass II (T4MO/benzene) monooxygenases (22) showed that TbuV from subclass II was able to effectively complement catalytic activity, while S1 from subclass I was not. To further define the reactions of heterologous effector proteins with T4MO, the natural MmoB effector protein and the catalytically inactive Δ N29-MmoB were also tested as part of this work. As with the S1 effector, neither of the MmoB variants were able to complement the T4MO complex in the standard T4MO in vitro assay (1).

Discussion

Comparison of the T4moD Crystal Structures and NMR Structures

Figure 2.6 shows a ribbon trace of the average NMR structure of native T4moD superimposed on the Δ N4-T4moD X-ray structure using C $_{\alpha}$ atoms from residues 11-98. The average rms difference from the overlay is 1.95 Å, and the individual rms differences are plotted with respect to residue position in Figure 2.2A. For Δ N4-T4moD, the structure refined from the higher resolution data gave 93.7% of the residues in the most favored region of the Ramachandran plot. In comparison, only 65.6% of the residues in the NMR structure were in the most favored region. The refined X-ray structures yield an estimated overall coordinate error of 0.095 Å for Δ N4-T4moD, 0.180 Å for Δ N10-T4moD, and 0.211 Å for native T4moD. The correlation coefficients were 0.961 (Δ N4-T4moD), 0.943 (Δ N10-T4moD), and 0.945 (native T4moD). For comparison, the core region of the NMR structure (Asn12-Phe98) exhibited rms differences to the averaged structure of 0.71 Å for the backbone atoms and 1.24 Å for all non-hydrogen atoms (22). The crystal structures of the natural and variants of T4moD are more similar to each other than they are to the

average NMR structure (Figures 2.1 and 2.2 and Table 2.2). Indeed, the rms differences between C_{α} atoms in the T4moD crystal structures are less than 0.8 Å. The rms differences between the C_{α} atoms in the crystal structures and the average T4moD NMR structure are approximately 2.0 Å (Table 2.2), and increase to ~6 Å when specific secondary structure elements are considered (e.g. the $\alpha 2$ helix in T4moD) or increase beyond 6 Å when the T4moD x-ray structure is compared with the DmpM NMR structure. The high resolution X-ray data also more accurately describes the positions of Arg69-Leu80, since these residues were not well defined in the NMR structure (22). As illustrated in Figure 2.2A, this region exhibits the highest rms differences between the X-ray and NMR structure.

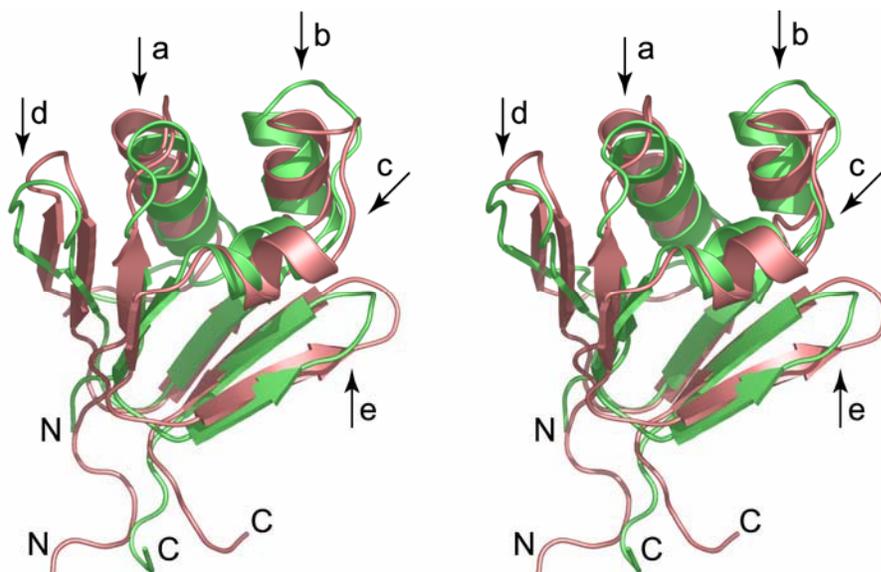


Figure 2.6: Divergent stereoview of the superimposed X-ray and NMR [1G10 (22)] structures of T4moD using residues 11-98. The chain A from the crystal structure in $\Delta N4$ -T4moD is shown in green, and the average NMR structure is red. The positions where the largest differences in the C_{α} traces occur are indicated with arrows and described in the text.

Several notable differences between the T4moD NMR and the X-ray structures are readily apparent (Figures 2.6 and 2.2A) and will be discussed with respect to the Δ N4-T4moD variant. The three α -helices in the X-ray and NMR structures exhibit significant differences. For example, the α 1 helix consists of three turns in the NMR structure, but of four turns in the X-ray structures. Although the α 1 helix terminates at analogous locations (Asp33), the helix starts at Gly19 in chain A of the Δ N4-T4moD and at Val23 in the NMR structure (Figure 2.6a). Consequently, this region yields rms differences of ~ 5 Å. The axes of the α 2 helices intersect with an angle of approximately 45° in the two structures (Figure 6b). Although the α 1 and α 2 helical axes are nearly parallel in the NMR structure, they are related by $\sim 45^\circ$ in the X-ray structures. The connecting loop between α 2 and α 3 helices and the orientation of the α 3 helix are also different in the structures (Figure 2.6c). Consequently, the region between residues 60 and 80 yields the highest rms differences between the NMR and X-ray structures (see also below). Finally, the hairpin turns connecting the β 2- β 3 and β 6- β 7 strands also diverge significantly (Figure 2.6d and 2.6e) and yield peaks in the rms difference plot (Figure 2.2A).

Each T4moD crystal structure exhibits a well-formed hydrophobic core with all charged residues on the exterior surface of the protein. Thus, the crystal structures appear more compact than the NMR structures (Figure 2.6). Consistent with these differences, the molecular surface area for the core residues (11-98) of the crystal structure (Δ N4-T4moD, 4170 \AA^2) is smaller than the average NMR structure (5753 \AA^2). Moreover, the crystal structures have a continuous solvent accessible surface area with only small depressions on the surface, while analysis of the surface area of the NMR structure reveals a large, tubular invagination that extends through the entire protein. Because of this feature, the enclosed solvent excluded volume for the NMR structure (9580 \AA^3) is smaller than that of the crystal structures (Δ N4-T4moD, 10450 \AA^3).

The hinge angle between the two β -sheets is listed in Table 2.3 and has been suggested to be a structural feature that differentiates the effector proteins (22). Three conserved glycine residues are located at the vertex of the hinge in T4moD (Gly36, Gly54, and Gly85), and extensive hydrogen bonding across the two β -sheets apparently stabilizes each particular angular relationship between the β -sheets. Thus, differences in hydrogen bonding between the α 1, α 2, and α 3 helices and the two β -sheets apparently stabilize the different hinge angles among the family members. It is also worth considering that the hydrophobic cavity is located between these three α -helices and toluene binding (if it does) would potentially influence this hinge angle.

Environment of Asn34

The sequence Asn34-Pro35-Gly36 is highly conserved across the catalytic effector protein superfamily (for example, see Figure 7 in Hemmi et al. (22)). Despite this sequence conservation, the X-ray and NMR structures of T4moD do not produce reasonable overlays when only the C_{α} , backbone atoms, or all atoms of these three residues are superimposed. Although the overlay between selected residues is reasonable, the rest of the structures are completely out of register. Moreover, when the α 1 helix and these three conserved residues are used, this does not yield satisfactory superposition. Thus, significant structural differences are present in this region of the structures. During the analysis of the T4moD NMR structure, the Asn34 side-chain amide protons were assigned to resonances that were 3σ outliers to the average chemical shifts assigned to amide groups in other proteins (http://www.bmrb.wisc.edu/data_access/outlier_selection_grid.html). These protons were also unusually resistant to exchange with solvent, despite being located near the surface of the protein (22), implying a unique chemical environment for the side chain. Figure 2.7A shows the $2mF_o-DF_c$ electron density for this region of the protein and the quality of the corresponding atomic model. The X-ray structures show that Asn34 and Pro35 cap the C-terminal end of the α 1 helix

and that no solvent molecules are hydrogen-bonded to Asn34^N. In the Δ N4-T4moD crystal structure, nearly every atom of Asn34 that can participate in hydrogen bonding does so. For example, the hydrogen atoms associated with Asn34^{N δ 2} are positioned to donate hydrogen bonds to the carbonyl oxygen atoms of Ala30 and Ile57. Asn34^{O δ 1} is poised to accept hydrogen bonds from Ile57^N and a solvent atom. In addition, Asn34^N donates a hydrogen bond to Ala30^O whereas Asn34^O accepts a hydrogen bond from Lys37^N.

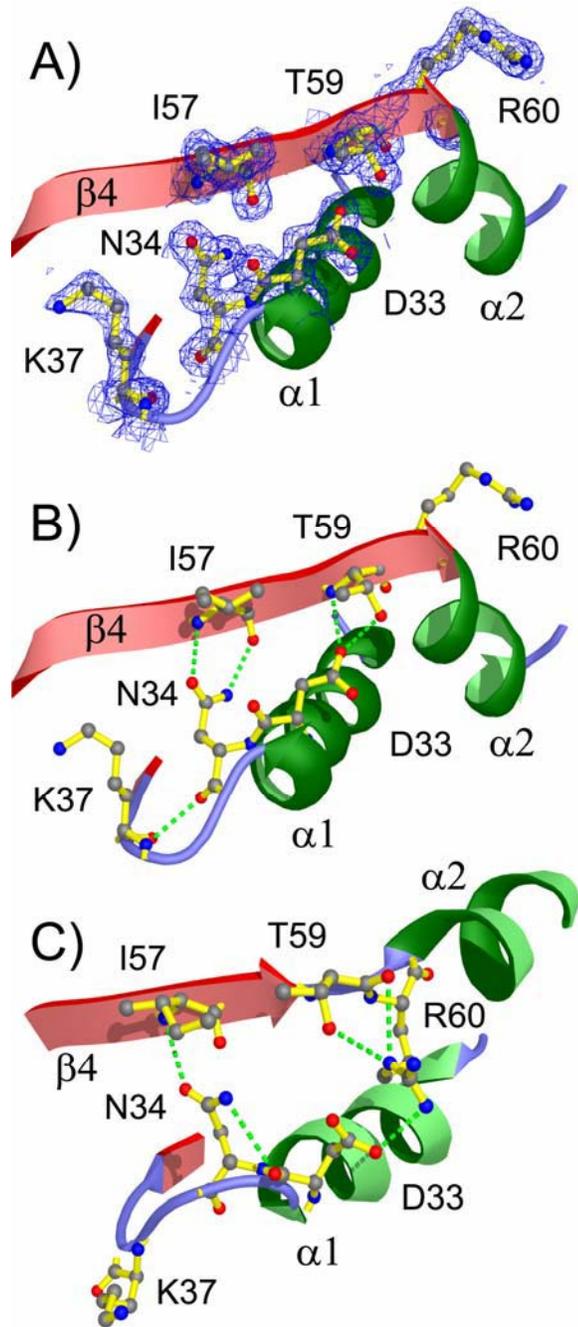


Figure 2.7: Differences in hydrogen bonding patterns observed in the crystal and NMR [1G10 (22)] structures of T4moD surrounding residue Asn34. A) The $2mF_o - DF_c$ electron density (1σ , 1.7 \AA resolution) superimposed on the refined $\Delta N4$ -T4moD model. B) X-ray structure of $\Delta N4$ -T4moD with hydrogen bonding indicated. Note that the position of Arg60 and $\alpha 2$ is different in the X-ray structure relative to the NMR structure. C) NMR structure (22). Note that the hydrogen bonding and electrostatic interaction between Arg60 and Asp33 and Thr59 are not observed in the crystal structures.

The hydrogen-bonding pattern assigned from the NMR structure of T4moD is considerably different, beginning where Asn34^N provides a hydrogen bond to Glu31^O and Asn34^O accepts a hydrogen bond from Gly36^N (see Figure 2.7C). The differences between the X-ray and NMR structures increase as the hydrogen-bonding patterns propagate further from Asn34. They maximize at Arg60, where the NMR structures indicate that Arg60 forms an ionic interaction with Asp33 (see Figure 2.7C). In contrast, the X-ray structure shows an ionic interaction between Arg60 and Glu64 (see below and Figure 2.8). Thus, some of the largest differences between the structures, also indicated in the rms difference plots (Figure 2.2A), are in the regions near Asp33 and Arg60. Two consequences of the differences in hydrogen bonding detected by the X-ray structures are that the $\alpha 1$ helix is longer and truly amphiphatic. Indeed, all five negatively charged residues are on the exterior surface, and all the hydrophobic residues are buried inside the hydrophobic core of the structure.

Environment of Glu64

The crystal structures reveal ionic interactions between Glu64 and Arg60 and between Glu75 and Arg69 (see Figure 2.8A,B). An interaction between Glu75 and Arg69 partially sequesters the hydrophobic cavity discussed above from bulk solvent and thus resembles an ionic gate for the cavity. The NMR structure of T4moD refined to a consensus family with the side chain of Glu64 projecting into the hydrophobic interior of the protein (Figure 2.8C), and the lack of a basic residue capable of forming a neutralizing charge pair or other potential hydrogen-bonding interactions was noted (22). Moreover, Arg60, Arg69, and Glu75 were not assigned to charge-pairing interactions. The refined NMR structures accommodated these differences through a slightly larger hinge angle between the two β -sheets (Table 2.3) as well as allowing the large solvent accessible channel described above. In the alignment of the NMR and X-ray structures, the C $_{\alpha}$ of Glu64 in the NMR structure nearly superimposes with the C $_{\alpha}$ of Leu63 in the crystal structures. Consequently, the Glu64 side chain of the NMR structure nearly

superimposes with the small hydrophobic cavity in the crystal structures (compare parts B and C of Figure 2.8).

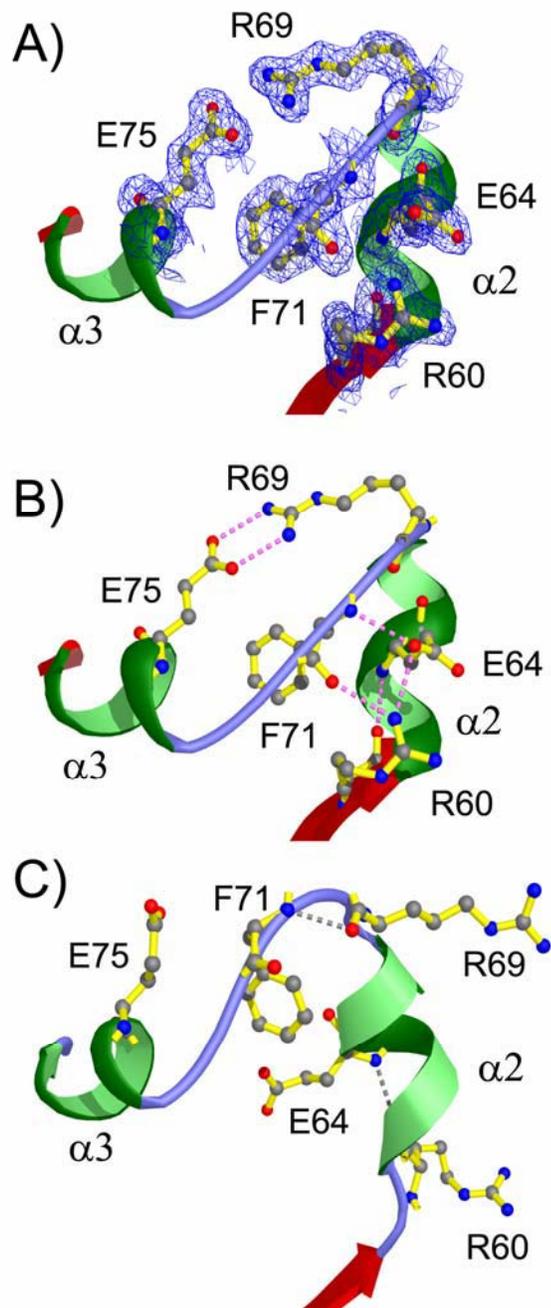


Figure 2.8: Differences in hydrogen-bonding patterns observed in the crystal and NMR [1G10 (22)] structures of T4moD surrounding several charged residues. A) The $2mF_o - DF_c$ electron density (1σ , 1.7 \AA resolution) superimposed on the refined $\Delta N4$ -T4moD model. B) X-ray structure of $\Delta N4$ -T4moD showing that Arg60 and Glu64 form a charge pair, and Phe71 occupies an interior position. C) NMR structure (22) that placed Glu64 in an interior cavity. Interactions between Arg60 and Glu64 and Arg69 and Glu75 observed in the crystal structures were not predicted by the NMR structures.

N-Terminal Region

The T4moD NMR structure includes a disordered region comprise of the first 10 residues, followed by a well-determined structure for residues 11-98 (22). Similarly, disordered and ordered regions are also observed in the crystal structures and appear to correlate with exposure to bulk solvent. Indeed, of the six chains in the crystal structures reported here, only the N-terminal end of chain A in the native T4moD projects into a crowded, lattice-packed environment, and only it exhibits a well-defined secondary structure. Thus, the α -helical structure in the N-terminal region of native T4moD appears to be dependent upon the environment or macromolecular crowding. Although the propensity for secondary structure to be sequence- and context-dependent has been documented, especially within T4 lysozyme (60-62), lattice interaction or macromolecular crowding that induce secondary structure is more difficult to asses (63, 64).

Role of N-Terminal Residues in Effector Proteins

The N-terminal disordered region of the MmoB effector protein is required for catalysis. This portion of the protein plays a role in the formation of the catalytic intermediates used by MMO for the oxidation of methane (19, 26, 28), the most difficult hydrocarbon to oxidize. The methane oxidation requires a highly reactive intermediate, possibly a diferryl species (3, 5, 20, 65-67). In contrast, other diiron enzymes have evolved to catalyze aromatic ring hydroxylations, an energetically less demanding reaction.

Deletion of the N-terminal region from T4moD had only modest influence on the steady-state kinetic properties and the regiospecificity of product formation of the reconstituted T4MO complex (Table 2.4). Furthermore, catalytic complementation of the T4MO complex with an effector protein from each subclass has now shown that only TbuV, an effector protein from an enzyme complex that is in the same subclass as T4MO, was able to give catalytic activity (22, 68) . Specifically, reconstitution studies

performed in this work with MmoB and Δ N29-MmoB, which likely have structures closely similar to T4moD (Figure 2.3 and Table 2.2), showed no catalytic complementation. These results indicate that, even as the three-dimensional structures of the effector proteins are relatively conserved, highly specific residue-dependent contacts will likely be required for function (Figure 2.5).

Table 2.4. Comparison of the Catalytic Properties and Regiospecificity of the Natural and N-Terminal Deleted Isoforms of Toluene 4-Monooxygenase Effector Protein.

T4moD variant ^c	Kinetic parameters ^a				Percent of products ^b			
	k_{cat} (s ⁻¹)	K_{M} (μM)	$k_{\text{cat}}/K_{\text{M}}$ ($\mu\text{M s}^{-1}$)	K_{I} (μM)	<i>p</i> -cresol	<i>m</i> -cresol	<i>o</i> -cresol	benzyl alcohol
Native-	3.3(0.1)	7.0(0.7)	0.47(0.1)	98(10)	96.2(0.2)	1.5(0.1)	0.9(0.0)	1.4(0.1)
Δ N4-	3.4(0.2)	2.3(0.3)	1.47(0.6)	106(17)	95.9(0.2)	1.6(0.1)	1.0(0.1)	1.5(0.2)
Δ N7-	2.8(0.1)	5.3(0.5)	0.53(0.2)	110(11)	95.2(0.2)	1.8(0.1)	1.2(0.1)	1.8(0.2)
Δ N10-	3.1(0.2)	3.2(0.5)	0.97(0.4)	116(19)	94.5(0.2)	1.8(0.1)	1.4(0.1)	2.2(0.2)

^aApparent kinetic parameters determined with each T4moD variant treated as the variable substrate and all other components of the reconstituted enzyme complex present in optimal amounts and NADH, toluene, and O₂ present in saturating concentrations. All kinetic experiments were performed in triplicate with at least seven different concentrations of T4moD. The standard error is shown in parenthesis. The k_{cat} values are reported relative to the $\alpha\beta\gamma$ protomer concentration of the T4moH component.

^bPercentage of product distribution observed from toluene oxidation. The standard error is shown in parentheses. ^cN-terminal sequence of the mature native protein obtained after in vivo post-translational removal of the N-terminal Met residue encoded by the expression plasmid, consisting of STLADQALHNNN- β 1. β 1 refers to the start of β -strand 1, the first identifiable element of secondary structure in T4moD. Δ N4-, deletion of 4 residues from mature N-terminal to give DQALHNNN- β 1; Δ N7-, deletion of 7 residues from the mature N-terminal to give LHNNN- β 1, Δ N10-, deletion of 10 residues from the mature N-terminal to give NN- β 1.

Fidelity Between Catalytic Effector Proteins

The distribution of charged residues on the surface of the $\alpha 1$ helices is different among the effector proteins (Figures 2.3 and 2.5). For example, TbuV, which cross-reacts with T4moH, has similarly charged residues as the T4moD. In contrast, MmoB, which does not cross-react with T4moH, has several residues that are either opposite in charge, polar, or even hydrophobic in places that correspond to Asp or Glu residues. Moreover, a positive electrostatic patch observed in T4moD (Figure 2.5, blue) is derived from Arg18, Arg44, Arg45, and Arg49. Amino acid sequence alignments suggest that similarly charged residues would be located in analogous positions in TbuV. In contrast, the MmoB proteins do not have Arg or Lys residues that correspond with these four arginine residues. Thus, we hypothesize that the $\alpha 1$ helix and electrostatic interactions may be important contributors to specific complex formation between T4moD and T4moH. Changes in these specific interactions may also be responsible for the lack of catalytic complementation of T4moH by the MmoB components. In T4moD, the $\alpha 1$ helix is flanked on one side by the $\beta 2$ strand and on the other by the $\alpha 2$ helix (Figure 2.5). It is interesting to note in the analysis of the crystal packing that an exposed $\beta 2$ strand forms an antiparallel β -sheet across the asymmetric unit interface in both space groups. If the $\alpha 1$ helix and the $\beta 2$ strand are critical features for complex formation, their interaction with T4moH may align the hydrophobic cavity in T4moD and a substrate access tunnel proposed from the recent crystal structure for the hydroxylase component of the toluene/*o*-xylene monooxygenase from *Pseudomonas stutzeri* OX1 (9).

Conclusion

The effector proteins have no cofactors or metals yet are able to exert exquisite control over the reactivity of the multiprotein diiron hydroxylase complex. The combination of X-ray and NMR structures now available indicates that the effector

proteins have a highly homologous three-dimensional structure, but there are profound differences in ability of the structurally conserved proteins to provide heterologous complementation of catalytic activity. The high resolution X-ray structures presented here have helped to define the potential differences in electrostatic surfaces that may govern the feasibility of protein-protein interactions. Moreover, the present structures reveal a single, well-defined cavity of size suitable for toluene binding near a region of the protein surface that has substantial different electrostatic properties among the effector protein family members. These results give new information and raise important new questions about these small, enigmatic components of the diiron hydroxylase enzyme complexes.

REFERENCES

1. Pikus, J. D., Studts, J. M., Achim, C., Kauffmann, K. E., Munck, E., Steffan, R. J., McClay, K., and Fox, B. G. (1996) *Biochemistry* 35, 9106-19.
2. Leahy, J. G., Batchelor, P. J., and Morcomb, S. M. (2003) *FEMS Microbiol. Rev.* 27, 449-79.
3. Baik, M. H., Newcomb, M., Friesner, R. A., and Lippard, S. J. (2003) *Chem. Rev.* 103, 2385-419.
4. Wallar, B. J., and Lipscomb, J. D. (1996) *Chem. Rev.* 96, 2625-2658.
5. Lipscomb, J. D. (1994) *Annu. Rev. Microbiol.* 48, 371-99.
6. Colby, J., Dalton, H., and Whittenbury, R. (1979) *Annu. Rev. Microbiol.* 33, 481-517.
7. Merkx, M., Kopp, D. A., Sazinsky, M. H., Blazyk, J. L., Muller, J., and Lippard, S. J. (2001) *Angew. Chem. Int. Ed. Engl.* 40, 2782-2807.
8. Elango, N., Radhakrishnan, R., Froland, W. A., Wallar, B. J., Earhart, C. A., Lipscomb, J. D., and Ohlendorf, D. H. (1997) *Protein Sci.* 6, 556-68.
9. Sazinsky, M. H., Bard, J., Di Donato, A., and Lippard, S. J. (2004) *J. Biol. Chem.* 279, 30600-10.
10. Rosenzweig, A. C., Frederick, C. A., Lippard, S. J., and Nordlund, P. (1993) *Nature* 366, 537-43.
11. Skjeldal, L., Peterson, F. C., Doreleijers, J. F., Moe, L. A., Pikus, J. D., Westler, W. M., Markley, J. L., Volkman, B. F., and Fox, B. G. (2004) *J. Biol. Inorg. Chem. in press.*
12. Muller, J., Lugovskoy, A. A., Wagner, G., and Lippard, S. J. (2002) *Biochemistry* 41, 42-51.
13. Chatwood, L. L., Muller, J., Gross, J. D., Wagner, G., and Lippard, S. J. (2004) *Biochemistry* 43, 11983-91.
14. Fox, B. G., Liu, Y., Dege, J. E., and Lipscomb, J. D. (1991) *J Biol Chem* 266, 540-50.
15. Froland, W. A., Andersson, K. K., Lee, S. K., Liu, Y., and Lipscomb, J. D. (1992) *J. Biol. Chem.* 267, 17588-97.
16. Green, J., and Dalton, H. (1985) *J. Biol. Chem.* 260, 15795-801.

17. Liu, Y., Nesheim, J. C., Lee, S. K., and Lipscomb, J. D. (1995) *J. Biol. Chem.* 270, 24662-5.
18. Lloyd, J. S., Bhambra, A., Murrell, J. C., and Dalton, H. (1997) *Eur. J. Biochem.* 248, 72-9.
19. MacArthur, R., Sazinsky, M. H., Kuhne, H., Whittington, D. A., Lippard, S. J., and Brudvig, G. W. (2002) *J. Am. Chem. Soc.* 124, 13392-3.
20. Shu, L., Nesheim, J. C., Kauffmann, K., Munck, E., Lipscomb, J. D., and Que, L., Jr. (1997) *Science* 275, 515-8.
21. Chang, S. L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (1999) *Biochemistry* 38, 5799-812.
22. Hemmi, H., Studts, J. M., Chae, Y. K., Song, J., Markley, J. L., and Fox, B. G. (2001) *Biochemistry* 40, 3512-24.
23. Qian, H., Edlund, U., Powlowski, J., Shingler, V., and Sethson, I. (1997) *Biochemistry* 36, 495-504.
24. Walters, K. J., Gassner, G. T., Lippard, S. J., and Wagner, G. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 7877-82.
25. Brandstetter, H., Whittington, D. A., Lippard, S. J., and Frederick, C. A. (1999) *Chem. Biol.* 6, 441-9.
26. Chang, S. L., Wallar, B. J., Lipscomb, J. D., and Mayo, K. H. (2001) *Biochemistry* 40, 9539-51.
27. Wallar, B. J., and Lipscomb, J. D. (2001) *Biochemistry* 40, 2220-33.
28. Brazeau, B. J., and Lipscomb, J. D. (2003) *Biochemistry* 42, 5618-31.
29. Brazeau, B. J., Wallar, B. J., and Lipscomb, J. D. (2003) *Biochem. Biophys. Res. Commun.* 312, 143-8.
30. Callaghan, A. J., Smith, T. J., Slade, S. E., and Dalton, H. (2002) *Eur. J. Biochem.* 269, 1835-43.
31. Kopp, D. A., Berg, E. A., Costello, C. E., and Lippard, S. J. (2003) *J. Biol. Chem.* 278, 20939-45.
32. Coufal, D. E., Blazyk, J. L., Whittington, D. A., Wu, W. W., Rosenzweig, A. C., and Lippard, S. J. (2000) *Eur. J. Biochem.* 267, 2174-85.

33. Pikus, J. D., Mitchell, K. H., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (2000) *Biochemistry* 39, 791-9.
34. Pikus, J. D., Studts, J. M., McClay, K., Steffan, R. J., and Fox, B. G. (1997) *Biochemistry* 36, 9283-9.
35. Studts, J. M., and Fox, B. G. (1999) *Protein Expr. Purif.* 16, 109-19.
36. Studts, J. M., Mitchell, K. H., Pikus, J. D., McClay, K., Steffan, R. J., and Fox, B. G. (2000) *Protein Expr. Purif.* 20, 58-65.
37. Orville, A. M., Studts, J. M., Lountos, G. T., Mitchell, K. H., and Fox, B. G. (2003) *Acta Cryst. D59*, 572-5.
38. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Cryst. D55*, 849-61.
39. de la Fortelle, E., and Bricogne, G. (1997) *Methods Enzymol.* 276, 472-494.
40. Collaborative Computational Project, N. (1994) *Acta Cryst. D50*, 760-763.
41. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* 277, 173-208.
42. Kleywegt, G. J., and Jones, T. A. (1997) *Methods Enzymol.* 277, 208-230.
43. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Cryst. D54*, 905-21.
44. Murshudov, G. N., Vagin, A. A., Lebedev, A., Wilson, K. S., and Dodson, E. J. (1999) *Acta Cryst. D55*, 247-55.
45. Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998) *Acta Cryst. D54*, 1285-94.
46. Read, R. J. (1997) *Methods Enzymol.* 277, 110-128.
47. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr D Biol Crystallogr* 49, 129-47.
48. Vaguine, A. A., Richelle, J., and Wodak, S. J. (1999) *Acta Cryst. D55*, 191-205.
49. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) *Proteins* 12, 345-64.
50. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) *J Mol Biol* 231, 1049-67.

51. Ferrin, T. E., Huang, C. C., Jarvis, L. E., and Langridge, R. (1988) *Journal of Molecular Graphics* 6, 2-12.
52. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* 18, 2714-23.
53. Kabsch, W., and Sander, C. (1983) *Biopolymers* 22, 2577-637.
54. Kelley, L. A., and Sutcliffe, M. J. (1997) *Protein Sci.* 6, 2628-30.
55. Xu, J., Baase, W. A., Quillin, M. L., Baldwin, E. P., and Matthews, B. W. (2001) *Protein Sci.* 10, 1067-78.
56. Moe, L. A., Hu, Z., Deng, D., Austin, R. N., Groves, J. T., and Fox, B. G. (2004) *Biochemistry* 43, 15688-701.
57. Eriksson, A. E., Baase, W. A., Wozniak, J. A., and Matthews, B. W. (1992) *Nature* 355, 371-3.
58. Morton, A., Baase, W. A., and Matthews, B. W. (1995) *Biochemistry* 34, 8564-75.
59. Morton, A., and Matthews, B. W. (1995) *Biochemistry* 34, 8576-88.
60. Sagermann, M., Baase, W. A., and Matthews, B. W. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 6078-83.
61. Sagermann, M., Baase, W. A., Mooers, B. H., Gay, L., and Matthews, B. W. (2004) *Biochemistry* 43, 1296-301.
62. Sagermann, M., and Matthews, B. W. (2002) *J. Mol. Biol.* 316, 931-40.
63. Juers, D. H., and Matthews, B. W. (2001) *J. Mol. Biol.* 311, 851-62.
64. Zhang, X. J., Wozniak, J. A., and Matthews, B. W. (1995) *J. Mol. Biol.* 250, 527-52.
65. Gherman, B. F., Baik, M. H., Lippard, S. J., and Friesner, R. A. (2004) *J. Am. Chem. Soc.* 126, 2978-90.
66. Brazeau, B. J., Austin, R. N., Tarr, C., Groves, J. T., and Lipscomb, J. D. (2001) *J. Am. Chem. Soc.* 123, 11831-7.
67. Brazeau, B. J., and Lipscomb, J. D. (2000) *Biochemistry* 39, 13503-15.
68. Mitchell, K. H., Studts, J. M., and Fox, B. G. (2002) *Biochemistry* 41, 3176-88.

CHAPTER 3

**CRYSTALLIZATION AND PRELIMINARY ANALYSIS OF A
WATER-FORMING NADH OXIDASE FROM *LACTOBACILLUS*
*SANFRANCISCENSIS***

Abstract

⁵Single crystals have been obtained of NADH oxidase (Nox), a flavoenzyme cloned from *Lactobacillus sanfranciscensis*. The enzyme catalyzes the oxidation of two equivalents of NAD(P)H and reduces one equivalent of oxygen to yield two equivalents of water, without releasing hydrogen peroxide after the reduction of the first equivalent of NAD(P)H. The enzyme crystallizes in space group $P2_12_12_1$ with unit cell parameters $a = 59.6$, $b = 92.6$, $c = 163.5$ Å. The crystals diffract to 1.85 Å resolution using synchrotron radiation. Matthews coefficient calculations suggest the presence of two molecules per asymmetric unit ($V_M = 2.3$ Å³ Da⁻¹, 45.5% solvent content), which has been confirmed by the molecular replacement solution using a search model derived from NADH peroxidase (PDB code 1F8W).

Introduction

Lactobacillus sanfranciscensis is used in the production of sourdough bread and is thus an important member of the family of lactic acid bacteria. It is an aerotolerant anaerobe and an obligatory heterofermentative microorganism that obtains most of its energy from the fermentation of maltose (*I*). Genomic sequencing and analysis of several

⁵ The following chapter is an adaptation from Lountos, G.T., Riebel, B.R., Wellborne, W.B, Bommarius, A.S., and Orville, A.M. (2004) Crystallization and Preliminary Analysis of a Water-forming NADH oxidase from *Lactobacillus sanfranciscensis*, *Acta Cryst D*40, 2044-2047.

lactic acid bacteria suggests that most do not synthesize hemes or cytochromes (2-6). Consequently, they do not use an electron-transport chain, O₂, and oxidative phosphorylation for energy metabolism, but rather satisfy all of their energy requirements with glycolysis. Fermentation helps to maintain an optimized intracellular NAD⁺/NADH ratio that is essential for efficient glycolysis. It also yields lactic acid, which acidifies the media and reduces competition from other microorganisms (7, 8). However, *L. sanfranciscensis* tolerates O₂ in order to maintain a symbiotic relationship with other aerobic microbes, most notably the yeast found in sourdough breads, and for the fermentation of other cereals.

Oxidative stress in many lactic acid bacteria and other facultative and strict anaerobic bacteria is managed in part by the expression of one or more flavin-dependent NAD(P)H oxidase(s) (2, 4, 7-11). The enzymes catalyze reactions of the type illustrated in Figure 3.1 and thus can be further classified as (i) H₂O₂-producing NADH oxidases (PrxR and some Nox), (ii) peroxidases that reduce H₂O₂ to H₂O and (iii) water-producing NAD(P)H oxidases (Nox). In addition to the FAD cofactor (except for Prx and other peroxiredoxins which do not contain flavins), each enzyme includes one or more redox-active cysteine residue(s). This catalytically essential residue alternates between the thiol/thiolate and sulfenic acid states (Nox and Npx), the thiol/thiolate and the disulfide state (PrxR and Prx) or the disulfide and sulfenic acid state (Prx), for recent reviews see (12-18). The NAD(P)H oxidases exhibit a strong preference for O₂ as the electron acceptor (19-22). Moreover, they only rarely release H₂O₂ as the product, in contrast to most flavin-dependent oxidases (21, 23-26). For example, during aerobic NADH turnover by *L. sanfranciscensis* Nox, less than 0.5% of the reducing equivalents can be detected as H₂O₂ suggesting that hydrogen peroxide is not released from the active site (27, 28). In contrast, the Npx and Prx enzymes do not react with O₂, but rather use H₂O₂ or alkylperoxides as the electron acceptor (14, 29-32). Most Nox and Npx enzymes

exhibit a strong preference for NADH, whereas Nox from *L. sanfranciscensis* will use either reduced nicotinamide adenine dinucleotide.



Figure 3.1. Reactions catalyzed by lactic acid bacteria and other facultative or strict anaerobic bacteria.

All the enzymes may facilitate regeneration of oxidized pyridine nucleotides for glycolysis and help protect the organism against oxidative stress. For example, the NADH oxidase from *Streptococcus pyogenes* and *L. delbrueckii* have been shown to contribute significantly to aerobic metabolism under conditions of high O₂ stress (33, 34). Although no crystal structures have been reported for an H₂O-producing NADH oxidase, several have been characterized and described in the literature (35-39). BLASTP (v.2.2.8) searches with the protein sequence of *L. sanfranciscensis* Nox return other water-producing enzymes with sequence identities ranging from 59% for the enzyme from *L.plantarum* WCFS1 (NP_786664.1) to 39% for the crystallographically defined Npx from *Enterococcus faecalis* (32, 40). In each case, the most highly conserved portions from the sequence comparisons are associated with either the redox-active cysteine (Cys41 in Npx), the FAD-binding or the NAD(P)H-binding regions. The Nox from *L. sanfranciscensis* (ATCC 27651) is a homodimer comprised of approximately 50 kDa subunits, similar to other NADH oxidases. In contrast, Npx is a homotetramer with D2

symmetry relating the four subunits. In order to describe the structural basis of the substrate preference and lack of H₂O₂ generation by *L. sanfranciscensis* Nox, we have initiated a crystallographic analysis of the enzyme. Here, we report the crystallization conditions and initial X-ray diffraction analysis of the H₂O-forming NADH oxidase from *L. sanfranciscensis*.

Materials and Methods

Protein expression and purification

The cloning, expression and preparation of *L. sanfranciscensis* Nox was essentially as previously reported, with slight modifications (27, 28). Briefly, starter cultures of *Escherichia coli* strain JM101 containing the *nox2* gene from *L. sanfranciscensis* in the pKK223-3 vector were grown in 5 mL LB media augmented with 100 µg/mL ampicillin (LB/amp) at 310 K to an OD₆₀₀ of 1.0. 1 L cultures of LB/amp medium were inoculated with the starter cultures and grown at 303 K in baffled 2.8 L Fernbach shake flasks with 200 rev/min shaking. When the culture reached an OD₆₀₀ of 0.7, protein expression was induced by the addition of 1.0 mM IPTG for 3 hours. Additional ampicillin was added to 200 µg/mL at the induction time and 1.5 hours later. Cultures were harvested by centrifugation at 277 K and the resulting cell pellet was frozen at 193 K.

The protein was purified by a two-step procedure at 277 K unless otherwise noted. Approximately 31g of frozen cells were thawed and suspended in 30 mL 100 mM 1-methylpiperazine buffer pH 5.0 plus 1 mM EDTA, 5 mM DTT and 20 mM spermine. The cell slurry was sonicated at approximately 277-283 K for six 2 minute pulses and centrifuged at 16,000 rev/min. Acid precipitation of a range of impurities was accomplished by dialyzing the cell-free extract overnight at 303 K with three changes of 1.5 L 20 mM 1-methylpiperazine pH 5.0 plus 1 mM EDTA and 5 mM DTT. After centrifugation, the enzyme was loaded onto a Hiprep 16/10 Source 30Q column (Amersham Pharmacia, Piscataway, NJ, USA) and washed with ten column volumes of

20 mM 1-methylpiperazine pH 5.0 plus 5 mM DTT. The enzyme was then eluted with a linear gradient from 0 to 1 M NaCl in the same buffer at a flow rate of 5 mL/min. SDS-PAGE analysis suggested that the protein was >95% pure.

L. sanfranciscensis Nox activity was typically assayed with 0.2 mM NADH at 303 K in 0.1 M triethanolamine (TEA) pH 7.5. The reaction was followed at 340 nm and the activity was calculated using a $6220 \text{ M}^{-1} \text{ cm}^{-1}$ extinction coefficient for NADH. Protein concentration was determined by the Bradford method (Pierce Chemical, Rockford, IL, USA). The specific activity of *L. sanfranciscensis* Nox was typically 221 units/mg.

Crystallization and X-ray data collection

Initial crystallization conditions for NADH oxidase were determined using sparse-matrix screens from Hampton Research (Laguna Niguel, CA, USA) and Nextal Biotechnologies (Montreal, Quebec, Canada). Typically, 2 μL protein (10 mg/mL in 20 mM 1-methylpiperazine pH 5.0, 0.15 sodium chloride, and 5 mM DTT) was mixed with an equal volume of reservoir solution on a silanized cover slip. Equilibration took place at room temperature by vapor diffusion in 24-well VDX plates (Hampton Research). Crystals appeared after approximately 5 days with reservoir solution containing 100 mM HEPES buffer pH 7.0-7.7, 4-10% (v/v) 2-propanol and 18-24% (w/v) polyethylene glycol 4000. Optimization of the initial conditions involved mixing 2 μL protein solution with 2 μL reservoir solution and 1 μL freshly prepared 100 mM dithiothreitol and screening pH and precipitant concentration. Crystals used for data collection were obtained by transferring seed crystals appearing in drops equilibrated over 100 mM HEPES pH 7.5, 24% (w/v) polyethylene glycol 4000 and 6% (v/v) 2-propanol by streak-seeding with a cat whisker into a pre-equilibrated drop containing 2 μL 10 mg/mL protein solution, 2 μL reservoir solution [100 mM HEPES pH 7.0, 18% (w/v) polyethylene glycol 4000 and 8% (v/v) 2-propanol] and 1 μL 100 mM dithiothreitol. Crystals from the streak seeding

experiment appeared after approximately 4 days and reached maximum dimensions of 0.2 x 0.1 x 0.1 mm within two weeks.

Crystals for X-ray diffraction data collection were harvested with a nylon loop and transferred to mother liquor supplemented with 15% (w/v) polyethylene glycol 400 and allowed to soak for approximately 30 seconds. The crystals were flash frozen by rapid immersion into liquid N₂. All diffraction data were collected from crystals held at approximately 100 K on beamline 22-ID operated by the South East Regional Collaborative Access Team (SER-CAT) at the Advanced Photon Source (APS), Argonne National Laboratory using a MAR CCD 225 detector. Each image was collected with a 1 second exposure time, a 0.5° oscillation range and a 180 mm crystal-to-detector distance. The data were integrated with HKL2000 and merged with SCALEPACK (41). Molecular replacement was carried out with the MOLREP program from the CCP4 suite of programs (42, 43). A search model was prepared using a monomer of NADH peroxidase (PDB code 1F8W, (32)) with all non-identical residues to *L. sanfranciscensis* Nox mutated to alanine and searching for two molecules in the asymmetric unit with data between 46 and 3 Å resolution. Rigid-body refinement and simulated annealing of the MOLREP solution was performed with CNS (44).

Results and Discussion

The first crystals of *L. sanfranciscensis* Nox appeared as a cluster of small yellow rods and plates that were birefringent under polarized light. Refinement of pH and precipitant concentration yielded larger reproducible crystals but frequently failed to yield single crystals. Moreover, the single crystals obtained from these conditions typically only diffracted to approximately 3.5 Å resolution. The addition of dithiothreitol and streak-seeding yielded single crystals that were suitable for data collection at high resolution (Figure 3.2). The statistics for the X-ray diffraction data collection are reported in Table 3.1. The data sets were obtained from 180° of ϕ rotation, are strong [overall $I/\sigma(I) = 30.7$], of good quality ($R_{\text{sym}} = 9.9\%$) and redundant (overall 8.8-fold

multiplicity). The space group is $P2_12_12_1$ and Matthews coefficient calculations (45, 46) suggest the presence of two molecules in the asymmetric unit ($V_M = 2.3 \text{ \AA}^3 \text{ Da}^{-1}$, 45.5% solvent content).

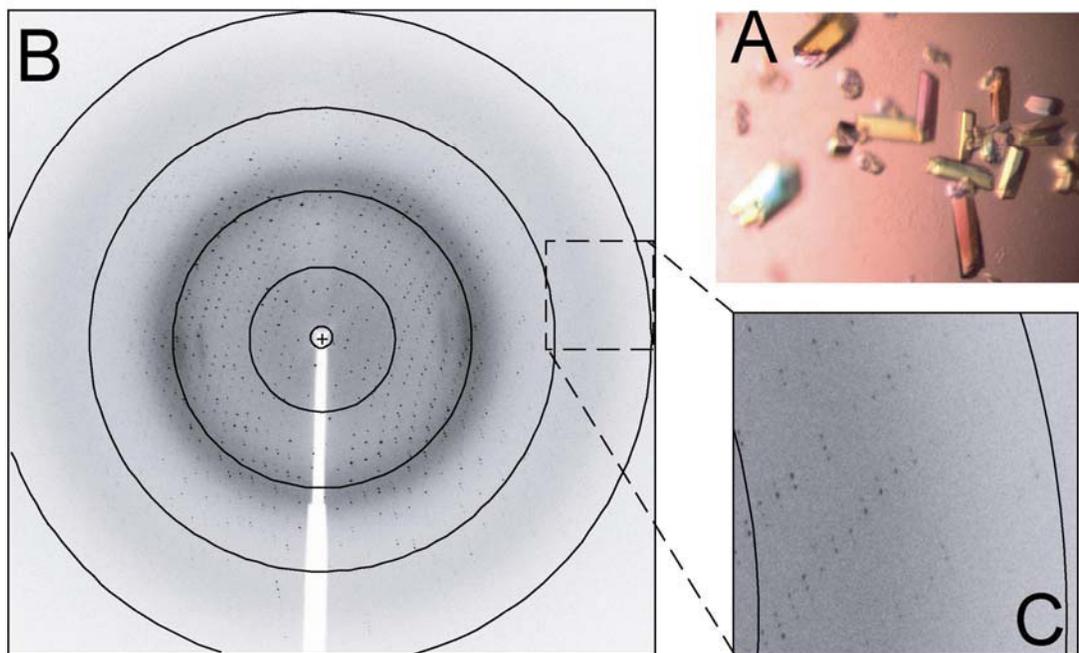


Figure 3.2. A) Crystals of *L. sanfranciscensis* NADH oxidase with dimensions of approximately 0.2 x 0.1 x 0.1 mm photographed under polarized light. B) the X-ray diffraction pattern obtained with 1 second exposure and 0.5° oscillation range about the vertical axis.; arcs indicate 7.1, 3.6, and 1.8 Å resolution. C) An expanded and contrast-adjusted view of the diffraction pattern between 2.4 and 1.8 Å perpendicular to the ϕ rotation axis.

Table 3.1: Data Collection Statistics for NADH Oxidase

X-Ray source	SER-CAT beamline 22-ID, APS
Beamline	22-ID
Wavelength (Å)	0.9997
Detector	MAR CCD 225
Resolution Range (Å)	46.4-1.85
Highest Resolution Shell (Å)	1.92-1.85
Mosaic spread (°)	~0.7
Space group	$P2_12_12_1$
Unit-cell parameters	
a (Å)	59.6
b (Å)	92.6
c (Å)	163.5
Total Reflections	675265
Unique Reflections	76768
Multiplicity	8.8 (5.4)
Completeness (%)	98.1(86.3)
$I/\sigma(I)^*$	30.7 (2.6)
$R_{\text{sym}}(\%)^\dagger$	9.9 (42.1)

Data were collected with a MAR CCD 225 detector. Values for the highest resolution shell of data are given in parentheses. $^\dagger R_{\text{sym}}(I)$ gives the average agreement between the independently measured intensities such as $\sum_h \sum_i |I_i - I| / \sum_h \sum_i I$ where I is the mean intensity of the i observations of reflection h . $^*I/\sigma(I)$ is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

The R303M mutant isoform of NADH peroxidase (Npx) from *E. faecalis* (PDB code 1F8W) has 39% sequence identity to *L. sanfranciscensis* Nox (32). A monomeric molecular search model consisted of 447 residues of the 452 residues in a full-length monomer of *L. sanfranciscensis* Nox. All non-identical residues to Nox were truncated to alanine unless they were glycine residues in Npx. All B factors were set to 20.0 \AA^2 . The top molecular replacement solution obtained from MOLREP yielded a correlation coefficient of 0.4 for two molecules in the asymmetric unit. The top result was approximately two times greater than the next best solution. The solution was subjected to rigid-body refinement and simulated annealing to 3.0 \AA resolution which improved the R-factors to $R_{\text{crist}} = 0.43$ and $R_{\text{free}} = 0.48$. The resulting electron density maps were clearly interpretable. Moreover, greater than 3σ positive difference features were apparent for the active site FAD and the side chains of the residues missing those atoms

in the search model. Refinement of the model to the high-resolution limit is currently in progress.

Conclusion

The structural analysis of *L. sanfranciscensis* Nox will help to establish the structural basis for the nearly stoichiometric production of H₂O₂, the almost complete lack of H₂O₂ detected after reduction of the first equivalent of NAD(P)H and the apparent promiscuity for reduced nicotinamide adenine dinucleotide substrates. Indeed, *L. sanfranciscensis* Nox exhibits nearly identical K_M values for NADH and NADPH (6.7 and 6.1 μ M, respectively), whereas the NADH oxidase from *Borrelia burgdorferi* or *L. brevis* only accept NADH (27, 28, 47). Moreover, comparisons to the other homologs may also reveal features that differentiate the various family members. For example, the structures of NADH peroxidase and the biochemical analysis of the NADH oxidase from *Enterococcus faecalis* reveal that a highly conserved redox-active cysteine residue plays an essential role in the catalytic cycle (48, 49). *L. sanfranciscensis* Nox contains the analogous Cys42 residue that is proposed to cycle between a thiol/thiolate anion and a sulfenic acid (Cys-SOH). This residue is proposed to be largely responsible for altering the enzyme reaction coordinate to yield H₂O rather than H₂O₂ (13, 21, 31, 50, 51). The X-ray diffraction data are of sufficient quality and resolution to support a refined crystal structure of *L. sanfranciscensis* Nox which is in progress.

REFERENCES

1. Gobbetti, M. C., A. (1997) *Food Microbiol* 14, 175-188.
2. Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarne, K., Weissenback, J., Ehrlich, S.D., & Sorokin, A. (2001) *Genome Res.* 11, 731-753.
3. Klaenhammer, T., Altermann, E., Arigoni, F., Bolotin, A., Breidt, F., Broadbent, J., Cano, R., Chaillou, S., Deutscher, J., Gasson, M., van de Guchte, M., Guzzo, J., Hartke, A., Hawkins, T., Hols, P., Hutkins, R., Kleerebezem, M., Kok, J., Kuipers, O., Lubbers, M., Maguin, E., McKay, L., Mills, D., Nauta, A., Overbeek, R., Pel, H., Pridmore, D., Saier, M., van Sinderen, D., Sorokin, A., Steele, J., O'Sullivan, D., de Vos, W., Weimer, B., Zagorec, M., and Siezen, R. (2002) *Antonie Van Leeuwenhoek* 82, 29-58.
4. Kleerebezem, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O. P., Leer, R., Turchini, R., Peters, S. A., Sandbrink, H. M., Fiers, M. W., Stiekema, W., Lankhorst, R. M., Bron, P. A., Hoffer, S. M., Groot, M. N., Kerkhoven, R., de Vries, M., Ursing, B., de Vos, W. M., and Siezen, R. J. (2003) *Proc Natl Acad Sci USA* 100, 1990-5.
5. de Vos, W. M., Bron, P. A., and Kleerebezem, M. (2004) *Curr Opin Biotechnol* 15, 86-93.
6. Siezen, R. J., van Enkevort, F. H., Kleerebezem, M., and Teusink, B. (2004) *Curr Opin Biotechnol* 15, 105-15.
7. De Angelis, M., Bini, L., Pallini, V., Cocconcelli, P. S., and Gobbetti, M. (2001) *Microbiology* 147, 1863-73.
8. De Angelis, M., and Gobbetti, M. (2004) *Proteomics* 4, 106-22.
9. Rabus, R., Ruepp, A., Frickey, T., Rattei, T., Fartmann, B., Stark, M., Bauer, M., Zibat, A., Lombardot, T., Becker, I., Amann, J., Gellner, K., Teeling, H., Leuschner, W. D., Glockner, F. O., Lupas, A. N., Amann, R., and Klenk, H. P. (2004) *Environ Microbiol* 6, 887-902.

10. Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D., Richardson, D. L., Kerlavage, A. R., Graham, D. E., Kyrpides, N. C., Fleischmann, R. D., Quackenbush, J., Lee, N. H., Sutton, G. G., Gill, S., Kirkness, E. F., Dougherty, B. A., McKenney, K., Adams, M. D., Loftus, B., Venter, J. C., and et al. (1997) *Nature* 390, 364-70.
11. Baker, L. M., Raudonikiene, A., Hoffman, P. S., and Poole, L. B. (2001) *J Bacteriol* 183, 1961-73.
12. Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., 3rd, Charrier, V., and Parsonage, D. (1999) *Biochemistry* 38, 15407-16.
13. Claiborne, A., Mallett, T. C., Yeh, J. I., Luba, J., and Parsonage, D. (2001) *Adv Protein Chem* 58, 215-76.
14. Poole, L. B., Reynolds, C. M., Wood, Z. A., Karplus, P. A., Ellis, H. R., and Li Calzi, M. (2000) *Eur J Biochem* 267, 6126-33.
15. Hofmann, B., Hecht, H. J., and Flohe, L. (2002) *Biol Chem* 383, 347-64.
16. Wood, Z. A., Schroder, E., Robin Harris, J., and Poole, L. B. (2003) *Trends Biochem Sci* 28, 32-40.
17. Argyrou, A., and Blanchard, J. S. (2004) *Prog Nucleic Acid Res Mol Biol* 78, 89-142.
18. Takeda, K., Nishiyama, Y., Yoda, K., Watanabe, T., Nimura-Matsune, K., Mura, K., Tokue, C., Katoh, T., Kawasaki, S., and Niimura, Y. (2004) *Biosci Biotechnol Biochem* 68, 20-7.
19. Ahmed, S. A., and Claiborne, A. (1989) *J Biol Chem* 264, 19863-70.
20. Ahmed, S. A., and Claiborne, A. (1989) *J Biol Chem* 264, 19856-63.
21. Mallett, T. C., and Claiborne, A. (1998) *Biochemistry* 37, 8790-802.
22. Mallett, T. C., Parsonage, D., and Claiborne, A. (1999) *Biochemistry* 38, 3000-11.
23. Fraaije, M. W., and Mattevi, A. (2000) *Trends Biochem Sci* 25, 126-32.
24. Massey, V. (2000) *Biochem Soc Trans* 28, 283-96.
25. Ward, D. E., Donnelly, C. J., Mullendore, M. E., van der Oost, J., de Vos, W. M., and Crane, E. J., 3rd. (2001) *Eur J Biochem* 268, 5816-23.

26. Kengen, S. W., van der Oost, J., and de Vos, W. M. (2003) *Eur J Biochem* 270, 2885-94.
27. Riebel, B. R., Gibbs, P. R., Wellborne, W. B., and Bommarius, A. S. (2002) *Adv. Synth. Catal.* 344, 1156-1168.
28. Riebel, B. R., Gibbs, P. R., Wellborne, W. B., and Bommarius, A. S. (2003) *Adv. Synth. Catal.* 345, 707-712.
29. Poole, L. B., and Claiborne, A. (1986) *J Biol Chem* 261, 14525-33.
30. Poole, L. B., and Claiborne, A. (1988) *Biochem Biophys Res Commun* 153, 261-6.
31. Parsonage, D., and Claiborne, A. (1995) *Biochemistry* 34, 435-41.
32. Crane, E. J., 3rd, Yeh, J. I., Luba, J., and Claiborne, A. (2000) *Biochemistry* 39, 10353-64.
33. Gibson, C. M., Mallett, T. C., Claiborne, A., and Caparon, M. G. (2000) *J Bacteriol* 182, 448-55.
34. Marty-Teyssset, C., de la Torre, F., and Garel, J. (2000) *Appl Environ Microbiol* 66, 262-7.
35. Koike, K., Kobayashi, T., Ito, S., and Saitoh, M. (1985) *J Biochem (Tokyo)* 97, 1279-88.
36. Ross, R. P., and Claiborne, A. (1991) *J Mol Biol* 221, 857-71.
37. Peterson, S. N., Hu, P. C., Bott, K. F., and Hutchinson, C. A. (1993) *Journal of Bacteriology* 175, 7918-7930.
38. Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., Fitzgerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S., and Venter, J. C. (1996) *Science* 273, 1058-1073.
39. Matsumoto, J., Higushi, M., Shimada, M., Yamamoto, Y., and Kamio, Y. (1996) *Biosci. Biotech. Biochem.* 60, 39-43.
40. Stehle, T., Ahmed, S. A., Claiborne, A., and Schulz, G. E. (1991) *J Mol Biol* 221, 1325-44.
41. Otwinowski, Z., and Minor, W. (1997) *Methods in Enzymology* 276, 307-326.

42. Vagin, A. A., and Teplyakov, A. (1997) *Journal of Applied Crystallography* 30, 1022.
43. COLLABORATIVE COMPUTATIONAL PROJECT, N. (1994) *Acta. Cryst D50*, 760-763.
44. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* 54 (Pt 5), 905-21.
45. Matthews, B. W. (1968) *J Mol Biol* 33, 491-7.
46. Kantardjieff, K. A., and Rupp, B. (2003) *Protein Sci* 12, 1865-71.
47. Hummel, W., and Riebel, B. (2003) *Biotechnol Lett* 25, 51-4.
48. Ross, R. P., and Claiborne, A. (1992) *J Mol Biol* 227, 658-71.
49. Stehle, T., Claiborne, A., and Schulz, G. E. (1993) *Eur J Biochem* 211, 221-6.
50. Mande, S. S., Parsonage, D., Claiborne, A., and Hol, W. G. (1995) *Biochemistry* 34, 6985-92.
51. Yeh, J. I., Claiborne, A., and Hol, W. G. (1996) *Biochemistry* 35, 9951-7.

CHAPTER 4

X-RAY CRYSTAL STRUCTURE OF NAD(P)H OXIDASE FROM *LACTOBACILLUS SANFRANCISCENSIS* AT 1.8 Å RESOLUTION

Abstract

⁶The flavin-dependent enzyme NAD(P)H oxidase (*L.san-Nox2*), found in *Lactobacillus sanfranciscensis* and in homologous form in many lactic acid bacteria, plays a critical role in managing oxidative stress. *L.san-Nox2* catalyzes the oxidation of two equivalents of NAD(P)H and reduces one equivalent of oxygen to yield two equivalents of water. Remarkably, the enzyme does not release hydrogen peroxide after the oxidation of the first equivalent of NAD(P)H and reaction with O₂. In order to understand the structure-function relationship of the enzyme, its mechanistic role in preventing release of hydrogen peroxide from the active site and its promiscuity for either NAD(P)H substrates, our laboratories have overexpressed, purified, and solved the three-dimensional X-ray structure to 1.8 Å resolution. The structural analysis reveals that the enzyme crystallizes as a dimer with each monomer consisting of a FAD binding domain, a NAD(P)H binding domain, and a dimerization domain. The crystal structure also shows that the highly conserved, active site Cys42 is located on the *si*-face of the FAD and exists as sulfenic acid (Cys-SOH). During the course of refinement, an unexpected ligand was discovered that is bound on the *re*-face of the FAD in the NAD(P)H binding

⁶ The following work is a collaboration between Lountos, G.T., Riebel, B.R., Wellborne, W.B., Bommarius, A.S. and Orville A.M. The author gratefully acknowledges the assistance of Mahmoud Ghanem and Prof. Giovanni Gadda of Georgia State University in obtaining the MALDI-TOF spectra.

domain. The density clearly supports the presence of adenine, ribose, and two phosphate moieties and MALDI-TOF analysis indicates that this molecule is ADP. The bound ADP along with the redox-active Cys42 residue are hypothesized to play a critical role in preventing hydrogen peroxide dissociation from the active site. Additionally, the bound ADP molecule may potentially influence the promiscuity for reduced nicotinamide adenine dinucleotides observed in *L.san*-Nox2. Based on crystallographic and biochemical data, we propose insights into the mechanism of *L.san*-Nox2. Furthermore, we present a comparison with the crystal structures of NADH peroxidase and glutathione reductase to determine similarities and differences between the homologs.

Introduction

Oxidative stress is potentially lethal to all life forms. Therefore, ubiquitous and diverse strategies have evolved in nature to sense, respond to, avoid or eliminate reactive oxygen and nitrogen species. Oxidative stress in many lactic acid bacteria, other facultative and strict anaerobic bacteria is managed, in part, by the expression of one or more flavin-dependent NAD(P)H oxidase(s) (1-8). *Lactobacillus sanfranciscensis* is an aerotolerant anaerobe and an obligatory heterofermentive microorganism (9). Genomic sequencing and analysis of several lactic acid bacteria suggest that most do not synthesize hemes or cytochromes (2, 7, 10-12). Consequently, they are unable to use an electron transport chain or oxidative phosphorylation and thus must satisfy all of their energy requirements with glycolysis. The NAD(P)H oxidases play an important functional role in regenerating oxidized pyridine nucleotides for glycolysis and help protect the organisms from oxidative stress. Studies have shown that the NADH oxidases from *Archaeoglobus fulgidus*, *Streptococcus pyogenes*, *Streptococcus mutans*, and *Lactobacillus*

delbrueckii contribute significantly to their aerobic metabolism under conditions of oxidative stress (13-16).

Sequence analysis of these enzymes reveals a highly conserved cysteine residue analogous to Cys42 in the *L.san*-Nox2. Typical sequence identities range from 39% for the crystallographically defined NADH peroxidase, 23% for the putative NADH peroxidase from *Streptococcus pyogenes* MGAS315, and 59% for the NADH oxidase from *Lactobacillus plantarum* WCFS1 (17-19). In each case, the most highly conserved regions include either the redox-active cysteine (Cys42 in *L.san*-Nox2) or the FAD and NAD(P)H binding domains. Although, to our knowledge, no crystal structure has been reported for a H₂O-producing NAD(P)H oxidase, sequence analysis suggests that these enzymes constitute a distinct class of FAD-dependent oxidoreductases. The crystal structure of *L.san*-Nox2 reveals that the enzyme shares close structural homology to the well-characterized family of pyridine nucleotide disulfide reductases (20) which include glutathione reductase (21), NADH peroxidase (18), thioredoxin reductases (22), dihydrolipoamide dehydrogenases (23), trypanothione reductases (24), and LpDA from *Mycobacterium tuberculosis* (25). Many of these types of enzymes are also important antioxidant proteins in some human pathogens (3, 4, 26-28).

Our laboratories have isolated and cloned the gene, overexpressed the protein, purified, characterized and solved the crystal structure for *L.san*-Nox2, the NAD(P)H oxidase from *Lactobacillus sanfranciscensis*. *Lactobacillus sanfranciscensis* alleviates oxidative stress and regenerates NAD(P)⁺ for glycolysis via this FAD-dependent NAD(P)H oxidase. The enzyme oxidizes NADH and NADPH at a V_{max}-ratio of 3:1, but with very similar K_M values of 6.1 and 6.7 μM, respectively. However, *L.san*-Nox2 does

not release H₂O₂ during turnover, but rather forms two H₂O molecules from O₂ and two equivalents of NAD(P)H. Indeed, we have demonstrated that less than 0.5% of the reducing equivalents can be detected as H₂O₂. (29, 30) Therefore, hydrogen peroxide must not be released from the active site during turnover. Sequence analysis and correlation with related enzymes suggest that the enzyme also uses a redox active cysteine residue (Cys42) that alternates between the thiol/thiolate and the sulfenic acid states during turnover (31). Thus, the first equivalent of NAD(P)H does yield H₂O₂, which then reacts with Cys42 to form the Cys42-SOH intermediate. The second equivalent of NAD(P)H is then used to reduce the sulfenic acid intermediate to the thiolate, which releases the second H₂O molecule. The essential role of the sulfenic acid intermediate is supported by analysis of mutants in homologous enzymes where the homologous residue was replaced with a serine or other amino acids. These mutants produced H₂O₂ instead of water during turnover (32-35). Two features of the reaction scheme (Figure 4.1) warrant further comment. First, the *L.san*-Nox2 evolved a mechanism to facilitate the generation of H₂O₂ which is very common in flavoprotein oxidases (36, 37), but then to prevent it from escaping from the active site. This is particularly important for *L.san*-Nox2 as the facultative anaerobe lacks a heme-based catalase to protect it from H₂O₂ (2, 7, 10-12). Second, the *L.san*-Nox2 also evolved a mechanism to enable O₂ reactivity at step 2 but not at step 5. Aberrant H₂O₂ produced at step 5 would likely inactivate the enzyme via generation of a Cys-SO₃H intermediate and create additional oxidative stress. Thus, *L.san*-Nox2 has evolved an efficient mechanism to protect the organism from oxidative stress.

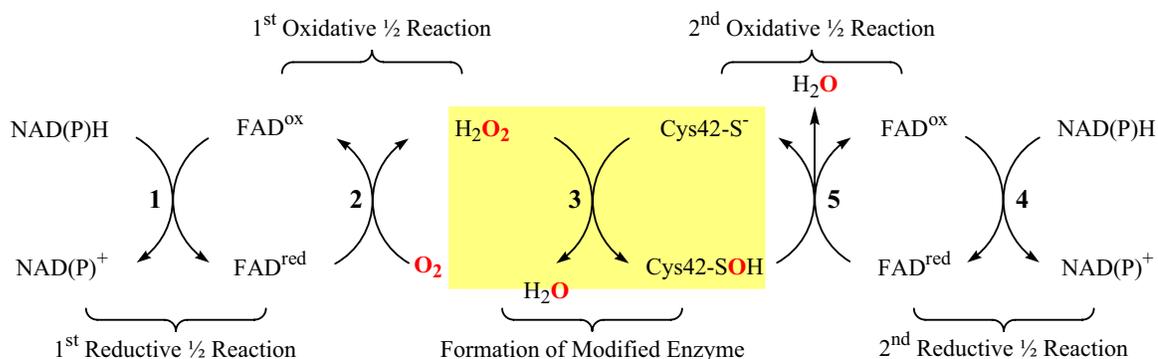


Figure 4.1. The overall reactions catalyzed by *L.san*-Nox2. The H_2O_2 generated in the active site does not dissociate from the enzyme, but rather reacts with the active site Cys42 (Step 3 within the shaded box).

The role of cysteine-sulfenic acids in biology and oxidative stress is currently garnering attention (31, 38). Oxidative stress is often caused by an imbalance of redox-active species in the cell of which reactive oxygen species are the principal molecular propagators and include hydrogen peroxide, superoxide, hydroxyl radical, or even molecular O_2 . Additionally, reactive nitrogen species such as nitric oxide, nitrite, or peroxynitrite are also very important contributors to oxidative stress. Recently, it has become clear that many of these reactive oxygen species are important constituents of cell signaling pathways and are also employed in fighting pathogenic infections (4, 39-41). Oxidative stress is implicated in the pathology of a wide range of human diseases such as atherosclerosis, hypertension, stroke, cancer, and neurodegenerative diseases such as Parkinson's disease (42-46). In response, Nature has evolved strategies to combat oxidative stress which include deployment of thiol-containing or seleno-containing proteins and metabolites that serve to scavenge the potentially reactive oxygen and nitrogen species (Figure 4.2). However, "second-generation" reactive species such as sulfenic or sulfinic acids ($\text{R-SO}_2\text{H}$), disulfide-S-oxides, or thienyl radicals are also

potentially deleterious to the cell as they are not effective mediators of reactive oxygen/nitrogen species (3, 26). Cysteine sulfenic acid is unstable in solution and can be further oxidized to the sulfonic acid (R-SO₃H) under aerobic conditions (31, 38, 47). Moreover, the oxidation of protein-based cysteine residues to the sulfenic, sulfinic and sulfonic acids has received significant attention in recent years and is quite prevalent in nature including humans (27). For example, the oxidation of essential cysteine residues in enzymes and proteins include the peroxiredoxins, the Parkinson's disease protein DJ-1, protein tyrosine phosphatase 1B, carbonic anhydrase III, the guanine nucleotide binding protein H-ras, and human serum albumin (40, 48-52). Often the reactive cysteine residue(s) are critical for the antioxidant defense of the cell. Reactive oxygen and nitrogen species initiate these modifications, which also include cysteine-S-nitrosylation and impact remarkable diverse biochemical systems and cell signaling pathways (41, 53-58). It is hypothesized that the immediate environment around the particular cysteine/cysteine sulfenic acid residue influences the susceptibility to further oxidation, but structural contributions to this process are not well understood. Recent studies have shown that some structural contributions may include intramolecular hydrogen bonding, the absence of other vicinal protein thiols that could promote the formation of disulfide bonds, limited solvent accessibility, and association with apolar elements of the protein (27, 31, 38, 47, 59).

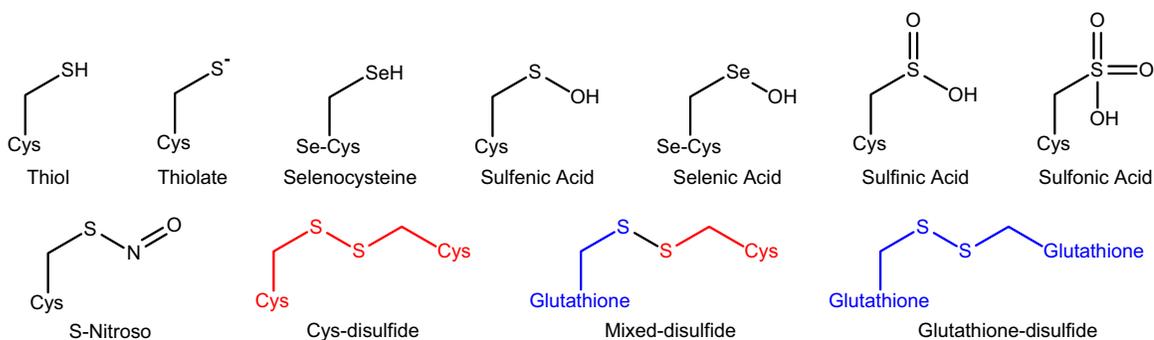


Figure 4.2. Some *in vivo* forms of cysteine and its common modification observed in antioxidant enzymes or proteins.

Here we report the crystal structure of native NAD(P)H oxidase from *Lactobacillus sanfranciscensis* (which will be referred to as *L.san-Nox2*) refined to a resolution of 1.8 Å. The high resolution structure confirms the presence of the active-site cysteine-sulfenic acid which was found to exist in two alternate conformations. Surprisingly, a tightly bound ligand which we have modeled as ADP was found on the *re*-face of the active site FAD although no exogenous ADP was added to the protein during the course of purification or crystallization. The entire ADP molecule fits well in the electron density. The determination and analysis of the three-dimensional structure of *L.san-Nox2* is the first for any water-forming member of this class (60). It completes the structural characterization for the representative members in each subclass and provides novel insights toward establishing the structure-function relationship of the enzyme.

Materials and Methods

Preparation, Crystallization, & Data Collection

L.san-Nox2 was cloned, expressed, purified, and crystallized as previously reported (60). X-ray diffraction data were collected from cryoprotected crystals held at approximately 100 K at beamline 22-ID of the SER-CAT facility of the Advanced

Photon Source (APS), Argonne National Laboratory and processed as previously reported (60). Briefly, the crystals crystallized in space group $P2_12_12_1$, diffracted up to 1.8 Å resolution and contain two molecules per asymmetric unit. An additional data set was collected from another crystal of *L.san*-Nox2 grown in identical conditions in order to confirm the presence of a bound ADP in this structure as well. The crystal diffracted to a resolution of 2.1 Å with very similar unit cell dimensions. The structure description of *L.san*-Nox2 will be based on the highest resolution (1.8 Å) structure.

Crystal Structure Determination

The structure of *L.san*-Nox2 was solved by molecular replacement using NADH peroxidase (PDB code; 1F8W) as a search model and the program MOLREP as previously described (60). The molecular replacement solution was initially subjected to rigid body refinement and simulated annealing with CNS (61) to a resolution of 3.0 Å resolution. Manual model rebuilding was performed with the program O (62, 63) and the model was refined against the 1.8 Å resolution data using maximum likelihood refinement in the program REFMAC5 (64) from the CCP4 suite of programs (65). Progress of the refinement was monitored by R_{free} (66, 67), which was calculated using 5% of the reflections, and cross-validated, σ_A weighted $2mF_o - DF_c$ and $mF_o - DF_c$ maps to evaluate the model and correct errors (68). Water molecules were located and refined in the final stages of refinement with ARP/Waters (69) and REFMAC5. The 2.1 Å resolution crystal structure was determined by subjecting the 1.8 Å structure to rigid body refinement against the 2.1 Å data followed by maximum likelihood refinement with REFMAC5. The ADP, FAD, and solvent molecules were removed from the model prior to the refinement and the Cys42 residue was modeled as alanine to prevent model bias.

The quality of the stereochemical parameters of the refined model were evaluated with PROCHECK (70, 71) and indicate good stereochemical properties.

The X-ray data collection and refinement statistics for *L.san*-Nox2 are shown in Tables 4.1 and 4.2. Initial maps obtained after rigid-body refinement and simulated annealing resulted in interpretable electron density with greater than 3σ positive difference features clearly visible for the active site FAD and the side chains for many of the truncated residues. In order to minimize model bias, Cys42 was modeled as alanine until the majority of the structure was traced correctly into the maps. An ADP molecule was fit into the extra density on the *re*-face of the flavin which consisted of greater than 6.5σ positive difference features and was refined. Simulated-annealing OMIT maps were also calculated in CNS by omitting ADP, FAD, and Cys42 from the map calculations in order to check for model bias (61). The final model for *L.san*-Nox2 refined to an R-factor of 0.17 and R-free of 0.22 for data between 46.4 and 1.8 Å (77355 reflections). The 2.1 Å structure was refined to an R-factor of 0.20 and R-free of 0.26 for data between 50 and 2.1 Å (49877 reflections). Ramachandran analysis of the highest resolution structure showed that 90.3% of the residues were located in the most favored region, 9.3% in additionally allowed region, 0.4% in generously allowed regions, and none in the disallowed region. The estimated overall coordinate error for the model based on the R-factor was 0.137 (0.132 for R-free) and the correlation coefficient of the maps was 0.965.

Table 4.1. Data Collection Statistics for <i>L.san-Nox2</i>		
	Crystal 1	Crystal 2
Space group	<i>P2₁2₁2₁</i>	<i>P2₁2₁2₁</i>
Unit Cell Dimensions (Å)	<i>a</i> = 59.6 <i>b</i> = 92.6 <i>c</i> = 163.5	<i>a</i> = 59.2 <i>b</i> = 92.5 <i>c</i> = 163.2
Resolution Range (Å) ^a	46.4-1.80 (1.85-1.80)	50-2.10 (2.15-2.10)
Total Reflections	693981	344645
Unique Reflections	81507	52622
Completeness (%)	96.1 (73.8)	98.7 (95.2)
Multiplicity	8.5 (4.1)	6.5 (4.7)
<i>I</i> / σ (<i>I</i>) ^b	29.0 (1.9)	20.0 (2.2)
<i>R</i> _{sym} (%) ^c	10.2 (48.9)	9.9 (48.8)

^aNumbers in parentheses are for the highest resolution shell. ^b The average agreement between the independently measured intensities. ^cThe root-mean squared value of the intensity measurements divided by their estimated standard deviation

Table 4.2 Crystallographic Refinement Statistics for <i>L.san</i>-Nox2		
	Crystal 1	Crystal 2
Resolution Range (Å)	46.4-1.80	50-2.10
No. of Reflections	77370	49878
R-factor	0.178	0.199
R-free ^a	0.223	0.254
No. of non-H protein atoms	6980	6980
No. of FAD molecules	2	2
No. of ADP molecules	2	2
No. of water molecules	708	225
Mean <i>B</i> , protein atoms (Å ²)	27.1	34.6
Mean <i>B</i> , FAD (2) (Å ²)	22.7	27.9
Mean <i>B</i> , ADP (2) (Å ²)	48.0	49.7
Mean <i>B</i> , water atoms (Å ²)	37.2	225
RMS deviations from ideal		
Bond lengths (Å)	0.015	0.015
Bond angles (°)	1.542	1.546
Estimated Coordinate error	0.136	0.257
Correlation coefficient	0.965	0.951

^a Calculated with 5% of the data.

Solvent exposed surface areas were calculated with a 1.4 Å probe radius with Swiss-PDB viewer (v3.7b2) (72) or VEGA (<http://www.ddl.unimi.it>) (73). Library files for the modified Cys42 residue and cofactors were made using the Dundee PRODRG server (74). Secondary structure assignments were made using KSDSSP (75). Structural homologs in the Protein Data Bank were found using MSDfold (76). The rms differences between models was calculated with SSM (<http://www.ebi.ac.uk/msd-serv/ssm>) or Swiss-PDB viewer (v3.7b2). Structure figures were prepared using Swiss-PDBViewer (v3.7b2) and PovRay (v.3.5) or Pymol (DeLano Scientific LLC, Castro City, CA).

Mass Spectrometry

The supernatant containing released ligands from heat denatured *L.san*-Nox2 was analyzed using MALDI-TOF mass spectrometry. Samples were prepared by gel filtration using a Sephadex G-25 column (PD-10) (Amersham-Pharmacia Biotech) and then heat denatured by boiling the enzyme for 1.5 hours. MALDI-TOF spectra of the supernatant were collected using an ABI Voyager DE-pro mass spectrophotometer at the Georgia State University (Atlanta, GA) Mass Spectrometry facility.

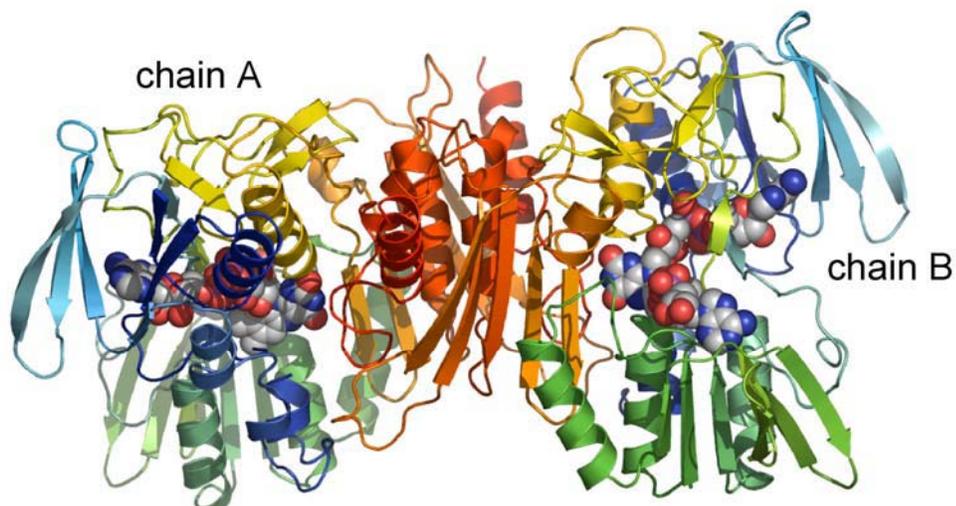
Results

Overall Structure Description

The determination of the crystal structure of *L.san*-Nox2 reveals structural homology with the pyridine nucleotide disulfide reductases family of enzymes (Table 4.3). The asymmetric unit contains one holoenzyme which is composed of two identical subunits related by two-fold symmetry (Figure 4.3). Indeed, structural overlays between the two monomers reveal rms differences of approximately 0.4 Å even though non-crystallographic symmetry restraints were not applied during refinement. Each subunit is roughly divided into three major domains; an N-terminal FAD-binding domain (residues

1-120), a middle NAD(P)H binding domain (residues 150-250), and a C-terminal dimerization domain (residues 325-451). Approximately 6000 Å² is buried at the dimer interface which corresponds to 30% of the surface area for each monomer. The C-terminal residue 452 in chain A and residues 450-452 in chain B were not visible in the electron density maps and therefore were not modeled. The FAD and NAD(P)H binding domains both adopt a Rossman fold topology (Figure 4.4) and the structure of the individual domains are very similar when compared (rms difference of 1.09 Å over 66 common C_α atoms). In addition to non-covalently bound FAD, the enzyme also contains a bound ligand molecule in the NAD(P)H binding domain which we have modeled as ADP.

A)



B)

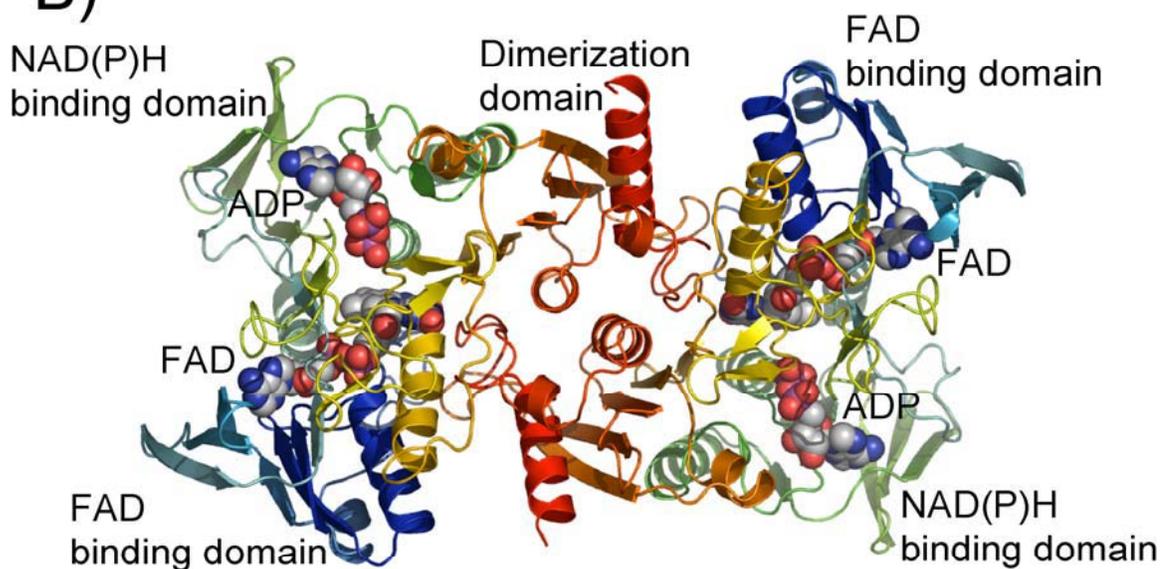


Figure 4.3. The X-ray crystal structure of NAD(P)H oxidase (*L.san*-Nox2) refined to 1.8 Å resolution illustrating A) the view perpendicular to the 2-fold axis and B) parallel to it are related by a 90° rotation about the horizontal axis. The FAD and ADP are shown as CPK atoms with the C, N, and O atoms colored in gray, blue and red, respectively.

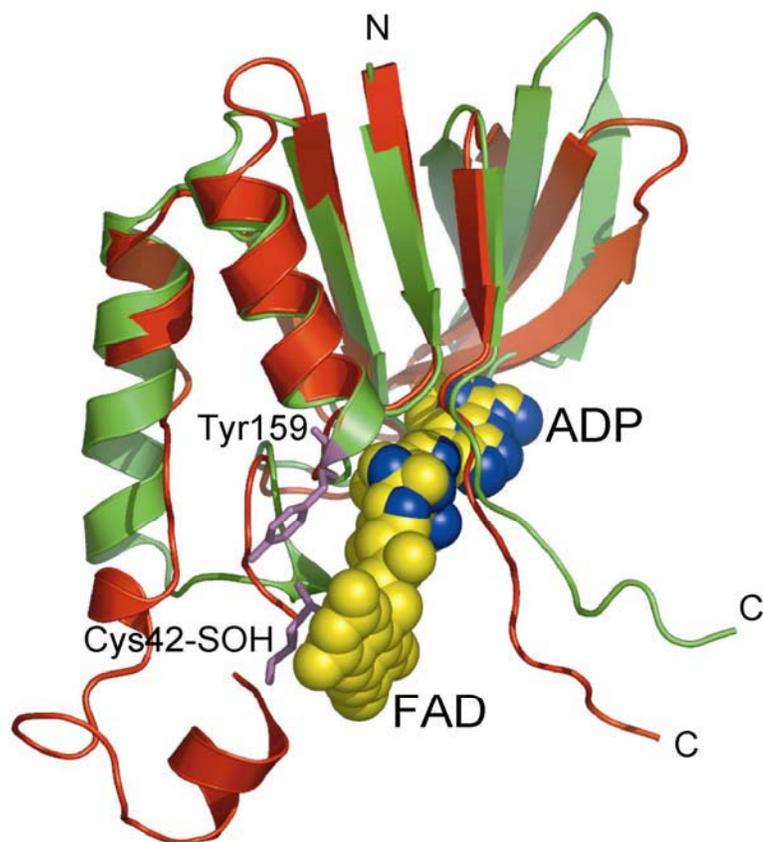


Figure 4.4. Overlay of the FAD-binding (green, residues 1-120) and the NAD(P)H-binding (red, residues 150-250) domains of *L.san-Nox2*. The FAD (yellow) and ADP (blue) cofactors are illustrated as CPK atoms. Both domains adopt a Rossman fold topology and exhibit very closely related structures. The Cys42-SOH and Tyr159 residues are shown in sticks (magenta).

Binding Mode of High Affinity Ligand

During the course of refinement, strong positive difference features (greater than 6.5σ) were present on the *re*-face of the FAD that indicated the presence of a bound ligand. The electron density clearly supports the presence of at least an adenine, ribose, and two phosphate moieties. The extra density is located on the *re*-face of the FAD in the NAD(P)H binding domain. Crystal structures from the closely related NADH peroxidase in complex with NADH show that this region is the binding site for NADH (77). The

supernatant (Figure 4.5) containing released ligands from the heat denatured enzyme was analyzed by MALDI-TOF mass spectrometry (Figure 4.6). The negative ion mode analysis reveals two sharp m/z ratios of 784 and 425.9 that correlate with the molecular weights of FAD and ADP, respectively. Simulated-annealing OMIT maps further support the fit of the ADP to the density (Figure 4.7). The ligand is nestled in the NAD(P)H binding domain in an extended conformation towards the FAD *re*-face (Figure 4.8). There is no evidence for covalent attachment of the ligand to the enzyme. The ligand also binds near the GXGXXG motif in which the carboxylate side chain of Asp179 hydrogen bonds with the 2' hydroxyl group of the ribose portion of ADP. The binding of ADP is stabilized by hydrogen bonding of the two well-defined phosphate groups with Ile160^N, Tyr159^N, and Tyr188^{OH}. Additional stabilizing hydrogen bonds occur between the ribose carbonyl groups and His181^{Nε2}.

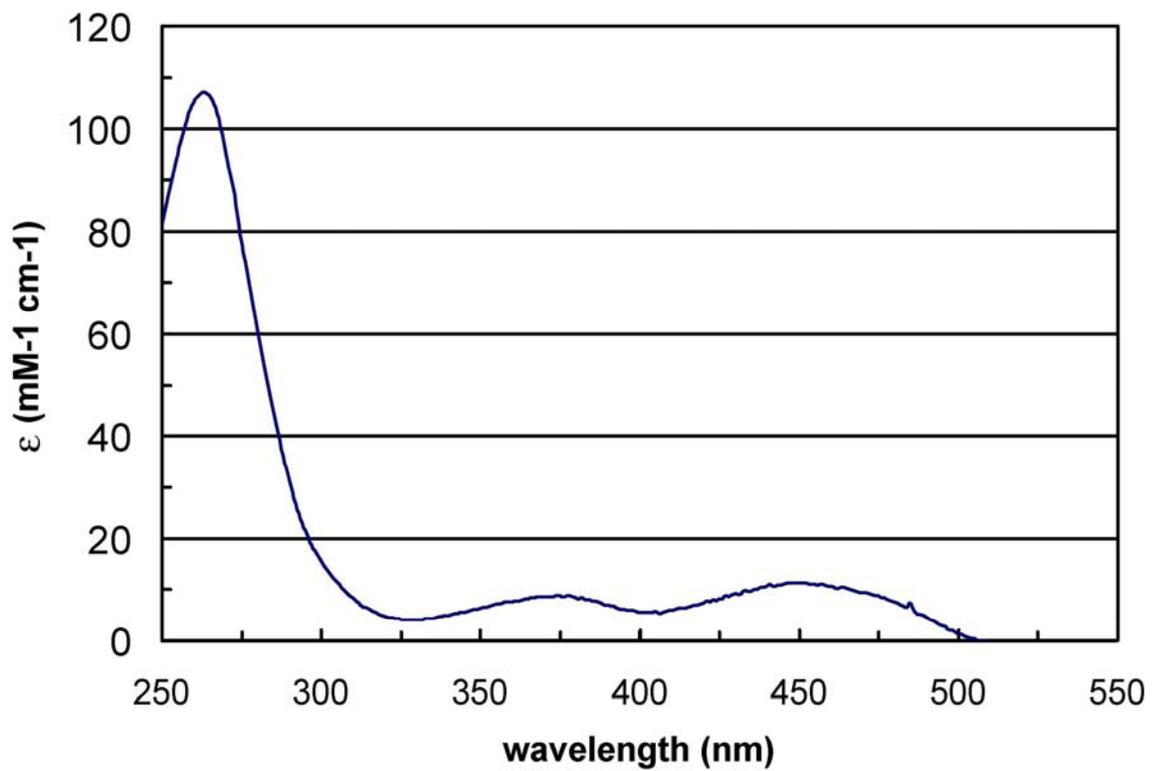


Figure 4.5 UV-Vis Spectrum of the supernatant from heat denatured *L.san*-Nox2.

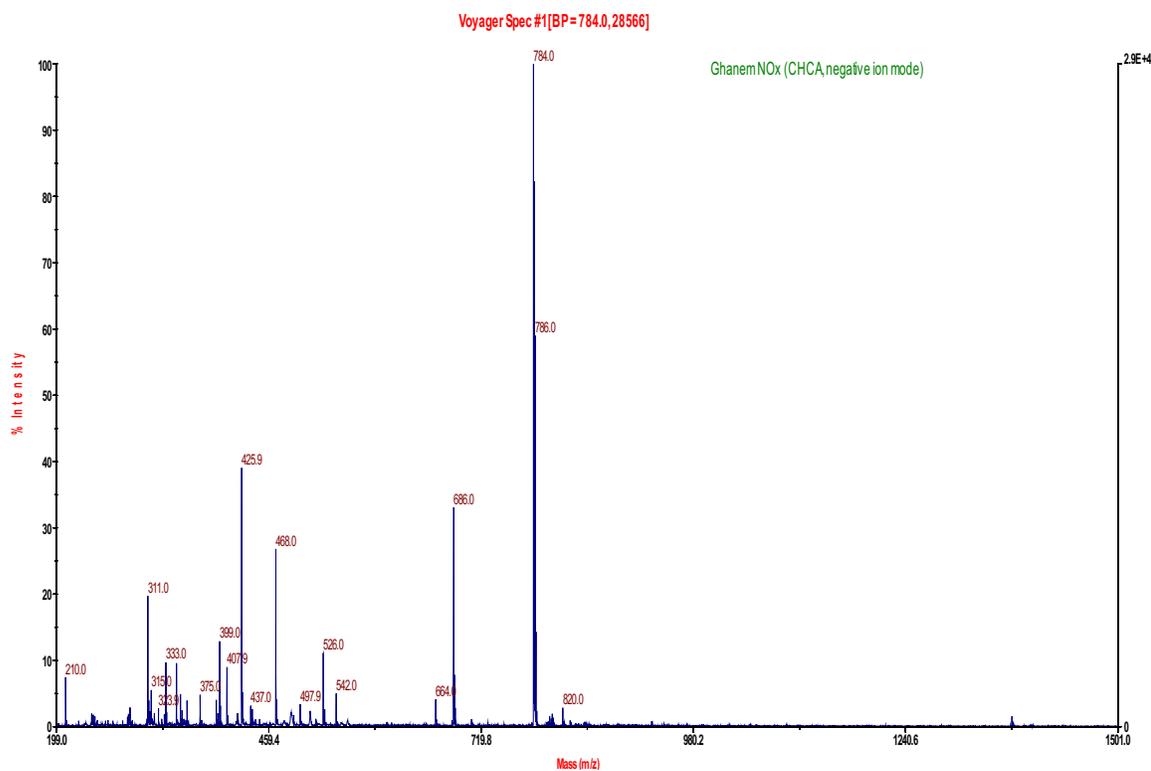


Figure 4.6. Negative-ion mode MALDI-TOF spectrum of the supernatant of denatured *L.san-Nox2*. The negative ion peaks with m/z ratios of 784 and 425.9 correlate with the molecular weights of FAD (785.7) and ADP (427), respectively. An ADP molecule was fit into the mF_o-DF_c difference electron density and is supported by the MALDI-TOF peak at 425.9. The peaks at 289, 311, 333, 351, 375, and 395 correspond to the molecular weights of the CHCA matrix.

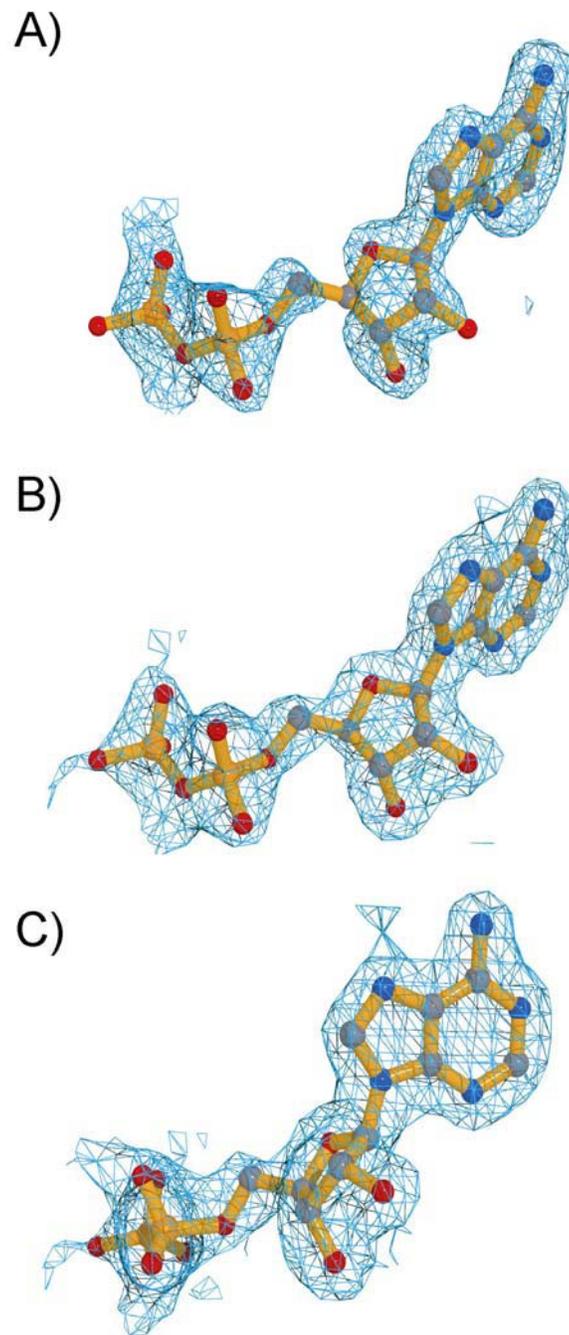


Figure 4.7. Electron density maps for the bound ADP molecule. A) The fit of the ADP molecule to the $2mF_o-DF_c$ (contoured at 1σ , 1.8 Å resolution) simulated annealing OMIT map obtained after omitting ADP and FAD from the map calculations B-C) The fit of the ADP molecule to the final $2mF_o-DF_c$ maps (contoured at 1σ , 1.8 Å resolution)

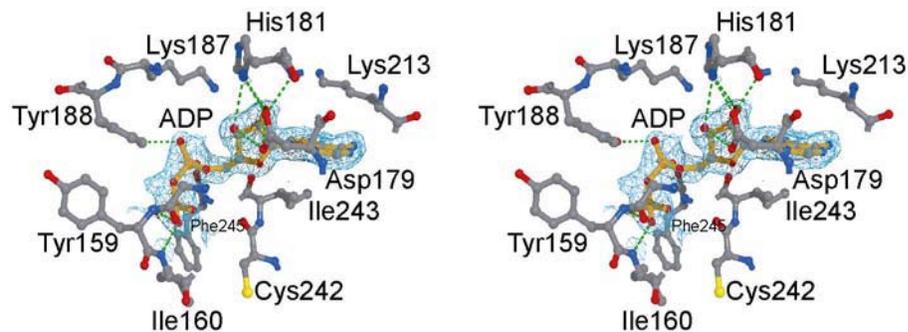


Figure 4.8. A divergent stereoview of the binding mode of ADP within the active site of *L.san*-Nox2. The hydrogen bonding interactions are represented by green dashed lines. The ADP molecule is superimposed upon the final $2mF_o-DF_c$ electron density maps (contoured at 1σ , 1.8 Å resolution). The bonds in the ADP molecule are colored in orange and the C, N, O, and S atoms of the amino acid residues are colored in gray, blue, red, and yellow, respectively.

Active Site Architecture

The electron density for the entire portion of the non-covalently bound FAD cofactor is unambiguous. A portion of the FAD in chain A is shielded from the exterior solvent by the second chain of the dimer via the backbone of Pro424b, Phe235b, and Met236b. In this contact, Phe423b is in close proximity to make hydrogen bonds via the backbone carbonyl with FAD N3 and O4 atoms. The electron density maps further indicate that the active site FAD is bent slightly with an approximate angle of 168.4° along C6-N5-C4 atoms and 169.6° along the C9-N10-N1 atoms of the isoalloxazine ring. The redox-active Cys42 is located on the *si*-face of the FAD and appears to be oxidized to the sulfenic acid (Cys42-SOH) that adopts two alternate conformations (Figure 4.9). Each conformation is stabilized by hydrogen bonding with His10^{Nε2} or the O2' of the FAD ribitol moiety. A number of refinement models were analyzed for this residue including Cys42-SH, Cys42-SO₂H, and Cys42-SO₃H. None of these trials produced models devoid of mF_o-DF_c difference features and often had *B*-values for the S and/or O atoms that were much different than their bonded atoms within the residue. A refinement model comprised of Cys42-SOH in two 0.5 occupancy conformations produced a model

with very well defined $2mF_o-DF_c$ electron density and no difference features. Moreover, the B -values with the residue varied smoothly from atom to atom. Additional trials with occupancies of the two conformations that ranged in 10% increments from 1:0 for each conformation were analyzed but these did not improve the models or maps. Slightly different orientations of the His10 residue are noticed that may depend on the orientation of the Cys42-SOH group. The Cys42-SOH residue further is in close proximity to the FAD-C4a (3.04 Å); thus it partially blocks access to the FAD-C4a atom from the *si*-face. In addition to the close proximity to His10, analysis of the protein environment around Cys42-SOH indicates the residue is sequestered by mostly nonpolar residues (Leu40, Gly43 from chain A and Phe435, Pro427 from chain B). A polar Ser41 residue lies adjacently to the Cys42 position. There are no vicinal thiol residues within close proximity to the Cys42-SOH.

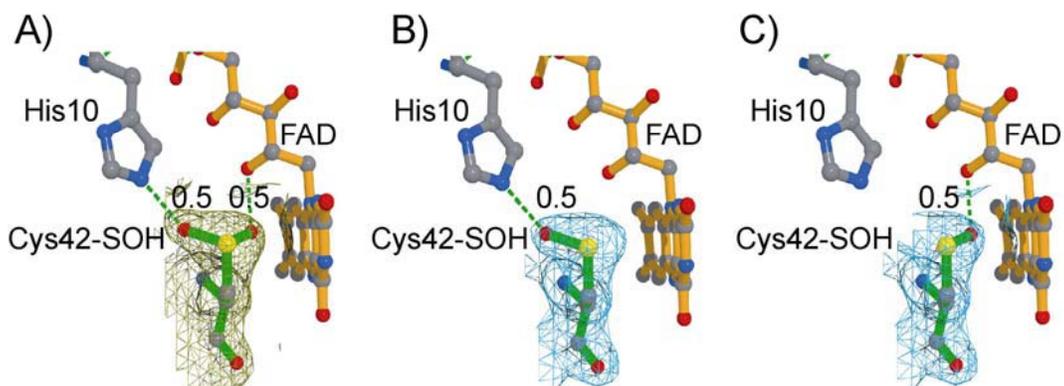


Figure 4.9. Structural analysis of the redox state of Cys42. A) The mF_o-DF_c positive difference electron density (gold, contoured at 3σ , 1.8 Å resolution) obtained from simulated annealing OMIT maps, with Cys42 omitted from the map calculations, superimposed upon the Cys42-SOH residue which is modeled with the distal oxygen occupying two alternate conformations with half-occupancy. B-C) The Cys42-SOH residue exists in two alternate conformations and is stabilized by hydrogen bonding interactions with either His10 or the FAD. Cys42-SOH also partially blocks access of O_2 to the FAD-C4a atom. The atoms in B-C are superimposed upon the final $2mF_o-DF_c$ electron density (blue, contoured at 1σ , 1.8 Å resolution). The bonds in the Cys42-SOH residue are colored in green and those of the FAD in orange. The C, N, O, and S atoms are colored in gray, blue, red, and yellow, respectively.

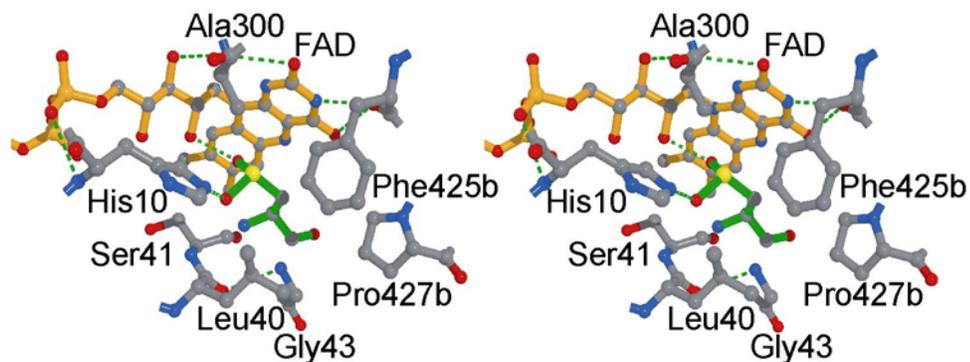


Figure 4.10. Stereoview of the active site residues surrounding the *si*-face of the FAD and Cys42-SOH. The Cys42-SOH residue is colored in green and the C, N, O, and S atoms are colored in gray, blue, red, and yellow, respectively.

Structural Homologs of L.san-Nox2

The structure of *L.san-Nox2* consists of a homodimer related by two-fold symmetry. Structural analysis of the monomer fold and a search of the Protein Data Bank reveal structural homology with the pyridine nucleotide disulfide reductase family of proteins. A detailed analysis of the structural overlays with the closest family members is listed in Table 4.3. Overall, these enzymes share very similar folds even at levels of low sequence identity. Due to the structural homology and the presence of homologous redox active cysteine(s) in the active site, the family members may be evolutionary related (Figure 4.11). The closest structural homolog is NADH peroxidase (pdb code: 1JOA) which exhibits rms deviations of 1.45 Å over 439 common C_{α} atoms when the monomers are superimposed (47).

Table 4.3: Structural Alignments of <i>L.san</i>-Nox2 with Homologs				
Enzyme	Q-Score^a	N_{align}^b	RMSD^c	% sequence identity
NADH Peroxidase ^d	0.78	439	1.45	39
Putidaredoxin Reductase ^e	0.49	368	2.01	21
Ferredoxin Reductase ^f	0.48	356	2.07	19
NADH oxidase/ Nitrite Reductase ^g	0.43	324	2.25	24
Dihydrolipoamide dehydrogenase ^h	0.34	368	2.93	20
Trypanothione Reductase ⁱ	0.31	348	2.64	17
Glutathione Reductase ^j	0.28	329	2.83	18
LPDA(Rv3303C) ^k	0.27	343	3.24	20

Structural Analysis of *L.san*-Nox2 with relevant homologs. Structures were superimposed using the Secondary Structure Matching server. ^a The Q score is defined as $N_{align} * N_{align} / (1+(RMSD/R_o)^2) * N_{res1} * N_{res2}$, where N_{align} is the length of alignment and N_{res1} and N_{res2} are the number of total residues in the *L.san*-Nox2 and target structure. ^cThe root mean squared difference ^b Number of C_α atoms used in alignment ^dAlignment with PDB code 1JOA, ^e1Q1W, ^f1F3P, ^g1XHC, ^h1LVL, ⁱ1AOG, ^j1GRA, and ^k1XDI. In all cases, chain A was used for structural overlaps.

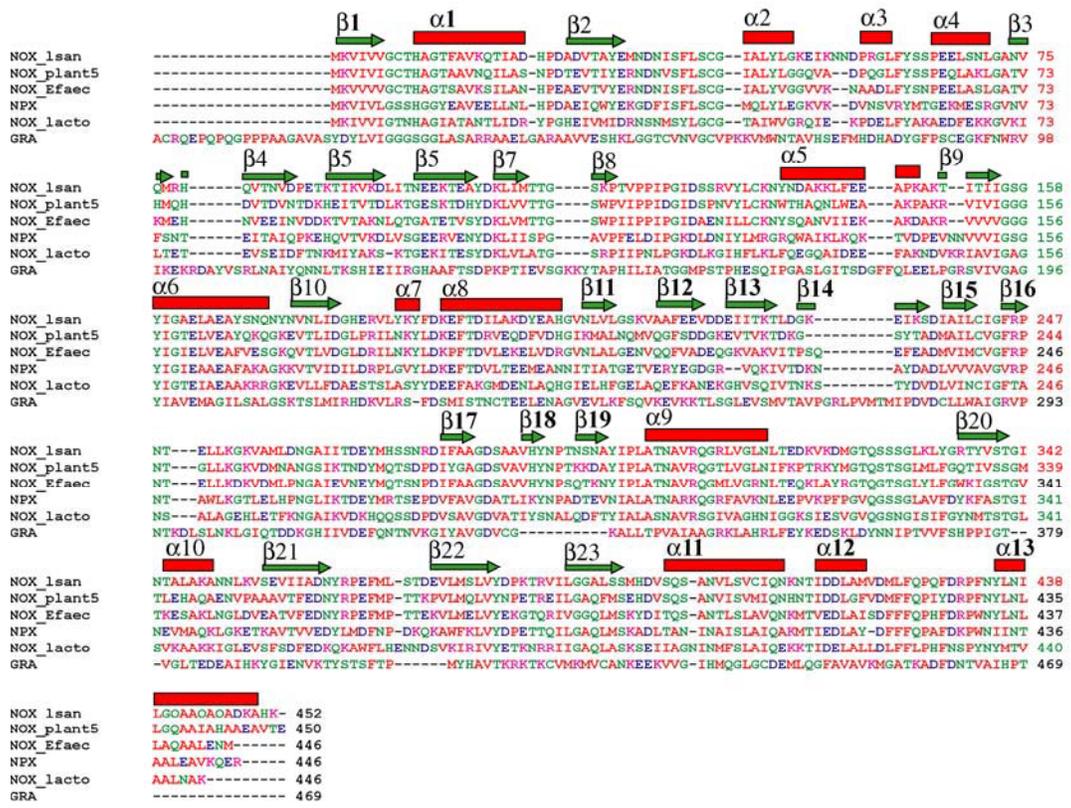


Figure 4.11 Primary sequence and secondary structure alignment of *L.san*-Nox2 and related homologs. The secondary structure elements for *L.san*-Nox2 are illustrated above the sequence for *L.san*-Nox2. The green arrows represent β-strands and the red rectangle represents α-helices. The amino acid sequence alignment was from CLUSTALW using default weights. The sequences (from top to bottom) used in the alignment are NAD(P)H oxidase (*Lactobacillus sanfranciscensis*, gi: 11862874), NADH oxidase (*Lactobacillus plantarum*, gi:28379772), NADH oxidase (*Enterococcus faecalis*, gi:29375784), NADH peroxidase (gi:29375784), NADH oxidase (*Lactococcus lactis*, gi:15672373), and human glutathione reductase (gi:18655540).

Comparison of L.san-Nox2 and NADH Peroxidase Crystal Structures

A comparison of various structurally defined members of the pyridine nucleotide disulfide reductase family indicate that the closest structural homology and sequence identity are found between *L.san*-Nox2 and NADH peroxidase. The monomers of *L.san*-Nox2 and NADH peroxidase superimpose with an rms difference of 1.4 Å over 439 common C_α atoms (Figure 4.12A). A striking difference between the two homologs, however, is that the biological unit of NADH peroxidase is a homotetramer with *D*₂ symmetry relating the four 46 kDa subunits. The quaternary structure contains an extensive interface between two monomers and only a weak association between the dimer of dimers (18, 47). In contrast, *L.san*-Nox2 is a homodimer comprised of two 50 kDa subunits. The NADH peroxidase tetramer is stabilized by only a few weak contacts. Analysis of the structure indicates that there are no salt bridges between the dimer of dimers. The main stabilizing interactions consist of hydrogen bonding and van der Waals contacts. Examination of interchain contacts indicates that the main contacts involve Lys53, Val55, Asn56, Ile138, Lys141, Gln142, Val145, Asp146, Pro147, Lys170, Ala171, Gly172, and Thr300. Sequence comparisons to *L.san*-Nox2 indicates that only Lys53, Pro147, and Thr300 are conserved in *L.san*-Nox2. Charge differences are also observed between the two structures in which Asn56, Ile138, Gln142, Asp146, and Lys170 of NPX are found to be Asp, Lys, Glu, Ala, Asn residues respectively in *L.san*-Nox2. Furthermore, the hydrophobic Val55 in NADH peroxidase is found to be Asn, a polar residue, in *L.san*-Nox2. Consequently, these differences in amino acid identity may contribute to the observation that *L.san*-Nox2 is a dimer. Moreover, crystal packing analysis in *L.san*-Nox2 does not yield a similar quaternary structure. When the *L.san*-Nox2 dimer is superimposed with the NADH peroxidase dimer, the rms differences

increase to approximately 1.9 Å (for 956 common C_α atoms). Analysis of the three common domains suggests additional structural differences. The largest structural deviations are found in the FAD binding domain (residues 1-119). Upon monomer superposition, all 119 C_α atoms overlay with rms deviation of ~4.8 Å but increases to ~6.2 Å when compared to the alignment of dimer structures. Indeed, the greatest differences are found between residues 53 and 120 in which rms differences range between 4 and 8 Å. In comparison, when comparing monomer vs. dimer structures with respect to the NAD(P)H binding domain, there is only an increase of 0.5 Å rms difference. The largest deviation in this domain occurs within residues 121-148 where rms differences range from ~3.4-8.0 Å. Thus, it appears that significant differences are found in the orientation of the secondary structure elements in the FAD binding domain that contribute to significant differences between the two structures which may favor a tetramer structure in NADH peroxidase and a dimer structure in *L.san*-Nox2. The most striking difference between the two homologs, however, is the presence of a bound ADP molecule in the NAD(P)H binding domain of *L.san*-Nox2 as isolated and crystallized. Native NADH peroxidase does not contain any bound ligands within the NAD(P)H binding domain as isolated.

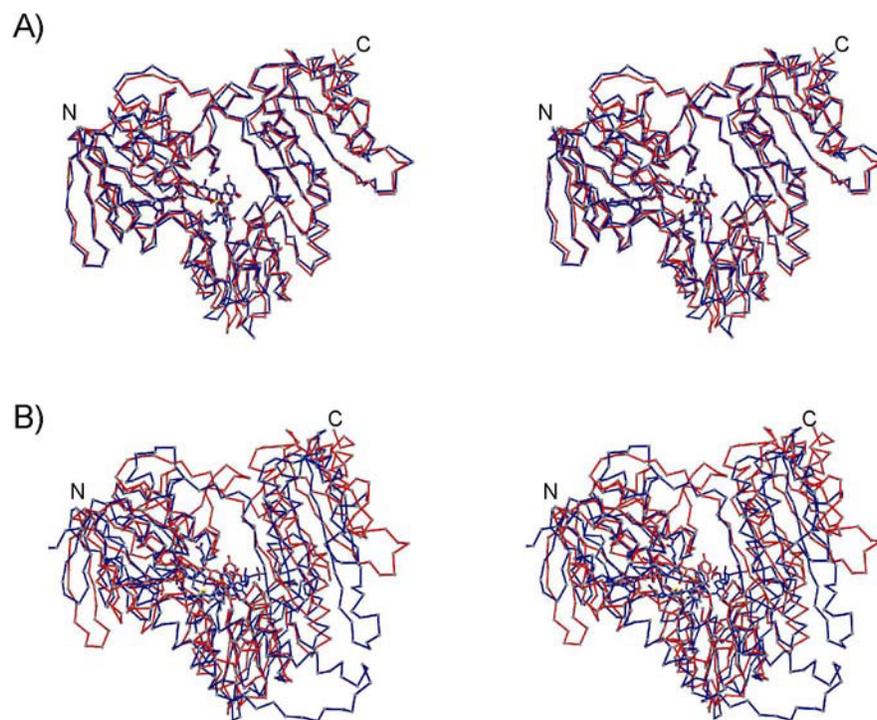


Figure 4.12. A divergent stereoview of the structural overlays of the C_{α} backbone trace of A) *L.san*-Nox2 (red) and NADH peroxidase (blue, PDB code: 1JOA) and B) *L.san*-Nox2 (red) and human glutathione reductase (blue, PDB code: 1GRA).

A comparison of the active site environment between NADH peroxidase (with native sulfenic acid) and *L.san*-Nox2 further reveals similarities and differences (Figure 4.13). Residues His10 and Cys42 are conserved among the two structures although the positions of the residues vary between the two structures. The positions of the C_{α} atom for His10 differ by 1.2 Å and 0.7 Å for Cys42. Furthermore, His10 is observed in two alternate conformations in *L.san*-Nox2. There is also a significant shift of 1.0 Å in the location of the sulfur atoms. The hydrogen-bonding environment also differs with respect to the two residues. In the NADH peroxidase structure, the proximal oxygen of Cys42-SOH hydrogen bonds with the FAD-N5 atom (3.2 Å) and is within 3.4 Å of the His10^{Ne2}

atom, whereas in the *L.san*-Nox2 structure, the alternate conformations of Cys42-SOH allow the oxygen to hydrogen bond with either the FAD O2' atom (2.5 Å) in one orientation and to the His10^{Nε2} atom (3.2 Å) in the other orientation. Consequently, the location of the proximal oxygen atoms of Cys42-SOH differs in the two structures. As His10 is conserved among many of the NAD(P)H oxidase and NADH peroxidase sequences, it has been proposed that His10 functions as an essential acid-base catalyst in NADH peroxidase (18). However, crystal structures from the NADH peroxidase C42S mutant and wild-type NADH peroxidase reveal a hydrogen bond between the Arg303 guanidinium moiety and the His10 imidazole. Based on these observations, it has been proposed that His10 remains unprotonated throughout the catalytic cycle (34). Studies based on the NADH peroxidase H10Q and H10A mutants suggest that His10 in NADH peroxidase is not essential for catalytic activity, but does function in part to stabilize the Cys42-SOH redox center within the active site environment (78). The mechanistic role of His10 in *L.san*-Nox2 awaits further biochemical study. Additional differences exist among residues lining the solvent access channel to the *si*-face of the FAD. In NADH peroxidase, there exists an ionic interaction between Glu14 and Asp303 and furthermore, the Arg303^{N1} atom hydrogen bonds with His10^{Nε1}. These interactions are non-existent in *L.san*-Nox2 due to the fact that these residues are nonpolar (Phe14 and Val304, respectively).

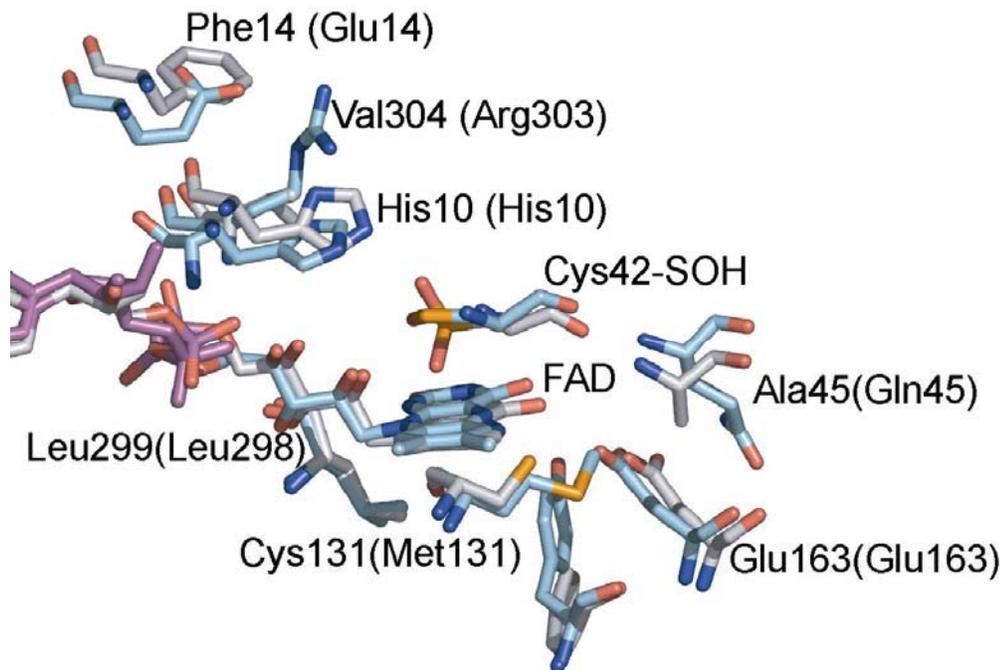


Figure 4.13. Structural overaly of *L.san*-Nox2 and NADH peroxidase active sites. The carbon atoms of *L.san*-Nox2 are colored in gray and those of NADH peroxidase are colored in cyan. The residue labels in parenthesis correspond to the NADH peroxidase residues. The N, O, S, and P atoms are colored in blue, red, yellow, yellow, and magenta, respectively.

Comparison of L.san-Nox2 and human Glutathione reductase

The monomeric structures of *L.san*-Nox2 and human glutathione reductase (hGR) overlay with rms differences of approximately 2.8 Å over 329 common C_α atoms (Figure 4.12b). The overall fold is very similar and both proteins exhibit roughly the same topology of secondary structural elements and domains (21). Analysis of the molecular surface of the *L.san*-Nox2 and hGR reveal that both share homology in the active site which includes the redox-active cysteine residues (Figure 4.14). The *L.san*-Nox2 Cys42 residue is analogous to the hGR Cys63 residue and the nearby Cys58 residue in hGR allows for the formation of a redox-active disulfide bond (79). The C_α positions of the

aligned Cys42 in *L.san*-Nox2 and Cys63 differ by 2.0 Å. Results from our laboratories indicate that the total turnover number for *L.san*-Nox2 is greatly increased in the presence of thiol-reducing agents, especially DTT (29, 30). Catalysis of *L.san*-Nox2 with NADH and O₂ is “turnover-limited”. For example, extended enzyme reactions at low enzyme concentrations (typically in the range of 2-20 nM) do not convert all of available NADH. The lack of conversion depends on the substrate-to-enzyme ratio and the presence or absence of thiol reagents such as DTT. In the absence of DTT, the total turnover number is approximately 5000 ± 1500. In contrast, the total turnover number in the presence of 1 mM DTT is more than 120,000 which is 24 times greater (29, 30). In comparison, the hydrogen peroxide producing NADH oxidase from *Lacotococcus lactis* exhibits a total turnover number of greater than 82,000 (80). Thus, DTT has a dramatic effect on *L.san*-Nox2 catalysis, but its precise role remains to be determined. As shown in Figure 4.14, the close homology of Cys42 in *L.san*-Nox2 and Cys63 in hGr suggests a putative binding site for exogenous thiols similar to the glutathione binding site in hGR. In the hGR-gluthione complex structure, hydrogen bonds between the glutathione substrate and the side chains of Ser30, Arg37, Tyr114, and Arg347 stabilize the complex (79). Comparison with *L.san*-Nox2 indicates that these residues correspond with His10, Lys17, His79, Arg308. Thus, the only strictly conserved residue corresponds to Arg308 and the other residues are conservatively substituted. These residues also surround the solvent channel that protrudes toward the *si*-face of the FAD and Cys42 allowing for a possible port of access for reducing agents. Previous crystallographic studies have shown that in NADH peroxidase, the Cys42 residue is susceptible to overoxidation to the sulfonic acid (Cys42-SO₃H) upon prolonged oxygen exposure (47). Thus, our observations involving

the use of exogenous thiols (DTT) suggests that they may play an important role in preventing the premature overoxidation of Cys42 in *L.san*-Nox2 which may involve binding of thiol-reducing agents near Cys42. We hypothesize that these agents may react with the Cys42-SOH resulting in the reduction of Cys42-SOH to the thiolate anion. It is also worthy to note that DTT was an essential component of the crystallization conditions that led to high quality crystals (60).

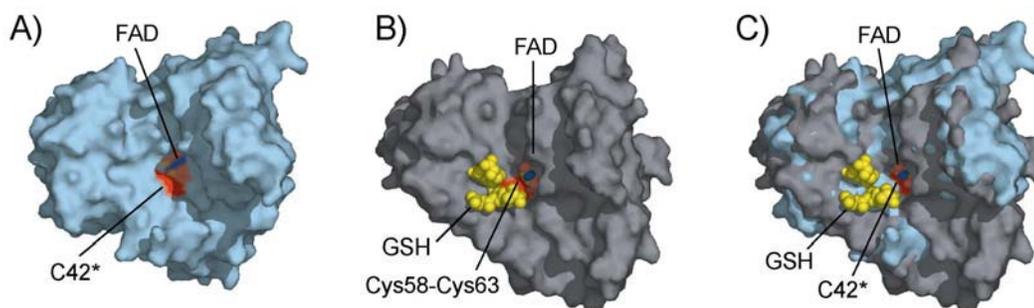


Figure 4.14. Surface representations of the monomer of A) *L.san*-Nox2 (cyan) and B) human glutathione reductase (gray) bound to two glutathione substrates (CPK atoms, yellow) (PDB code: 1GRA) compared to C) an overlay of the two indicate similar binding surfaces with respect to exogenous thiols.

The structural overlays for the monomeric structures of *L.san*-Nox2 and hGr are more similar to each other, however, than when the two biological units are superimposed. Upon superimposing the dimer structures, it is clearly evident that there are significant differences between the two structures. Using the Swiss-PDB viewer iterative fit function, the best structural overlay resulted in a rms difference of 1.8 Å over only 212 common C_α atoms. In comparison, the same routine gave an rms difference of 1.7 Å over only 204 common C_α atoms upon monomer overlays. Analysis of the two dimeric structures reveals that although the molecules in one half of the dimer superimpose well, the second subunits in the dimer are completely out of register with respect to each other

(rms difference of 34.8 Å for residues 18-450). Thus, there is a significant shift in the orientation of the second molecule in the dimer of hGr when compared to the dimer structure of *L.san*-Nox2 that more than likely is due to significant differences in interactions at the dimer interface. Analysis of the monomer overlays indicates the greatest structural differences occur in the dimerization domains of the two homologs.

Substrate Delivery

Previous work has shown that *L.san*-Nox2 exhibits dual substrate specificity for NAD(P)H (29, 30). Indeed, both substrates have similar K_M values for NADH and NADPH (6.7 and 6.1 μM, respectively). There are two channels that start from the solvent-exposed surface of the protein and extend to the active site; one to the *re*-face and the other to the *si*-face of the flavin. The channel beginning from the solvent-exposed surface of the enzyme and traversing to the *si*-face of the FAD is lined primarily with hydrophobic residues. In NADH peroxidase, the homologous channel is lined with more polar residues and is proposed to be the channel in which H₂O₂ is delivered to the active site (18). Another channel traverses from the solvent directly to the *re*-face of the FAD and forms a cavity which is bordered by Tyr159, Tyr188, Ile160, Glu163, Ser328, Gly329, Leu330, and the phosphate moieties of the ADP (Figure 4.15). If the observed ligand remains tightly bound to the enzyme, it can be envisioned that NAD(P)H substrate may be delivered via this channel. The channel is appropriately positioned to allow for an NAD(P)H molecule to enter the active site and position the nicotinamide ring beneath the isoalloxazine ring. Thus, we hypothesize that the dual substrate promiscuity could arise from the ability of the solvent channel to accept either NAD(P)H molecules. Additionally, there are two channels that are adjacent to the FAD *re*-face. The

second channel traverses from the *re*-face of the FAD to the *si*-face and is partially bordered by Cys42-SOH. Specificity for NAD(P)H is due, in part, to the identity of the amino acid located 19-21 residues downstream of the last glycine in the signature motif (GXGXXG). The motif is involved in recognizing the ADP portion of pyridine nucleotides (81-84). In general, NADP⁺ specificity usually involves an arginine at this position whose side chain interacts with the 2'-phosphate of NADP⁺ by ionic and hydrogen bonding interactions. In contrast, NAD⁺ specificity involves an acidic amino acid at this position in which the carboxylate side chain makes hydrogen bonding interactions with the hydroxyl groups of the ribose containing the adenine ring. In *L.san*-Nox2, Asp179 occupies this position and makes hydrogen bonding contacts with the ribose hydroxyl of the bound ADP. Thus, with ADP occupying the binding site and an available entrance tunnel to the *re*-side of the FAD, the bound ADP molecule may play an influential role in allowing *L.san*-Nox2 to accept either NAD(P)H substrate. Further biochemical experiments are needed to define the precise role of ADP within *L.san*-Nox2. Additional steric contributions from residues surrounding the cavities in the active site may also contribute to preventing escape of hydrogen peroxide from the active site prior to its reaction with the redox-active Cys42.

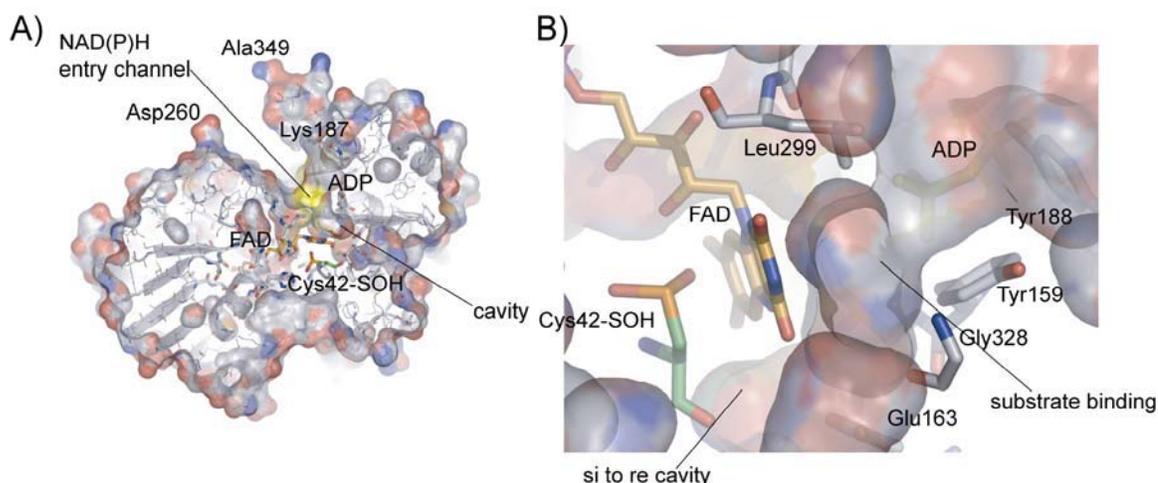


Figure 4.15. Surface representations of *L.san*-Nox2 reveals a probable A) entrance channel for NAD(P)H molecules towards the *re*-face of the active site FAD. B) The channel extends deep into the active site and forms a “cavity” in front of the *re*-face of the FAD isoalloxazine ring which would allow for the positioning of the nicotinamide ring of the entering NAD(P)H molecule next to the isoalloxazine ring. The availability of the entrance channel and the bound ADP in the Rossmann fold domain are postulated to contribute to the observed promiscuity for NAD(P)H substrates exhibited by *L.san*-Nox2. Furthermore, an additional channel between the *re*-face of the FAD and the *si*-face exists which is hypothesized to allow H₂O₂ migration towards the redox-active cysteine resulting in its reduction to H₂O after nucleophilic attack by the thiolate form of Cys42.

Orientation of Tyr159

Comparison of the sequence alignments of *L.san*-Nox2, NADH peroxidase, and hGR reveal that Tyr159 in *L.san*-Nox2 is highly conserved among the proteins. In the various structures from these homologs, the equivalent Tyr159 residues occupy one of two orientations; either the “in” orientation which partially occludes the FAD, or the “out” conformation which permits ligand binding at the *re*-face of the FAD. The crystal structures of NADH peroxidase in complex with NADH, and hGR in complex with NADP⁺, and *L.san*-Nox2 indicate that Tyr159 in *L.san*-Nox2 occupies the “out” position in order to accommodate the bound ligand (77, 79). The hydrogen bonding of the

hydroxyl moiety of the equivalent Tyr159 residue is important for stabilizing each orientation of the residue in the homologous structures. In *L.san*-Nox2, the “out” conformation is held in place by a hydrogen bond between Tyr159^{OH} and Ser^{O γ} . Thus, conformational changes and the orientation of the Tyr159 appear to play an important role towards the affinity for the substrate. A comparison of the aligned structures of NADH peroxidase in complex with NADH and *L.san*-Nox2 (Figure 4.16) indicate that the ADP molecule occupies a similar position with respect to the ADP portion of the NADH molecule (77).

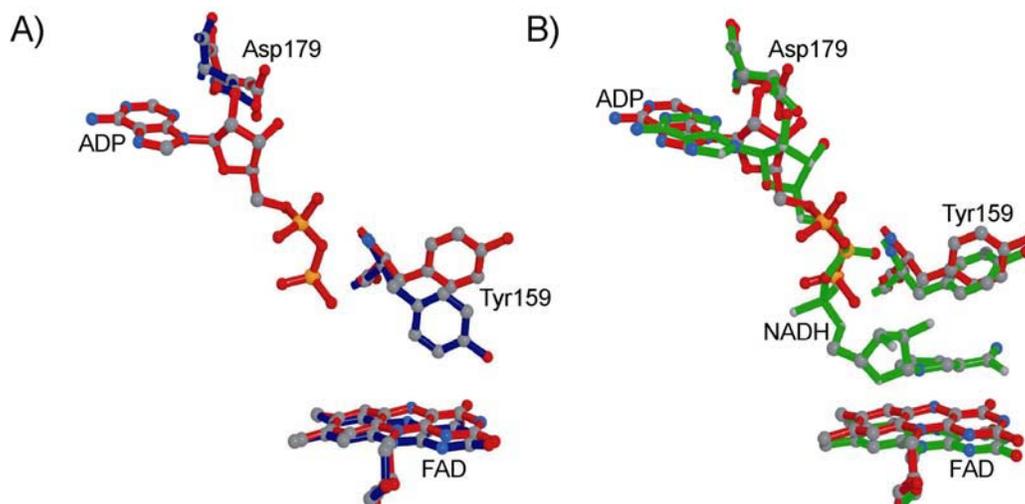


Figure 4.16. Structural overlays of the substrate-binding pocket in *L.san*-Nox2 and NADH peroxidase. The figure illustrates the location of the bound ADP molecule on the *re*-face of the FAD in *L.san*-Nox2. A) The superimposed structures of *L.san*-Nox2 (red) and unliganded NADH peroxidase (blue). B) The superimposed structures of *L.san*-Nox2 (red) and NADH peroxidase in complex with NADH (green). The conformation of the Tyr159 residue is hypothesized to influence ligand binding. In the homologous NADH peroxidase structure, the Tyr159 residue undergoes a conformational change upon ligand binding in the NAD(P)H binding domain (77, 79).

O₂ Access and Reactivity

Gas-phase theoretical studies indicate that the isoalloxazine ring has different conformations associated with the oxidized, reduced, and FAD C4a-H₂O₂ intermediate states (85-87). These redox-dependent alterations will likely influence reactivity with O₂ and reactivity with the Cys42-SOH intermediate. Previous work on NADH oxidase from *Enterococcus faecalis* 10C1, which shares 59% sequence identity with *L.san*-Nox2, has shown evidence for a C4a-peroxyflavin intermediate in studies of the C42S mutant isoform of NADH oxidase (33, 38). Mallett et al. proposed a direct reaction of Cys42-S⁻ with the peroxyflavin which would yield an immediate product as FAD-C4a-hydroxide. In this model, Mallett et. al predict that bound NAD⁺ at the *re* face does not dissociate prior to the O₂ reaction, and the O₂ reaction must be directed at the flavin *si*-face. This would generate the C4a-peroxy flavin with the distal oxygen projected into a position optimal for transfer to Cys42-S⁻ to regenerate the oxidized Cys42-SOH center. However, there is no crystal structure available to provide any structural insights. In order to predict the structural basis for reactivity with O₂ using our crystal structure, we have modeled in the C4a-peroxy complex using MOPAC simulations in CHEM3D (Cambridgesoft Corporation, Cambridge, MA). In the model, the reduced flavin can have the C4a-peroxy at either the *re* or *si* face (Figure 4.17). However, analysis of the active site architecture of *L.san*-Nox2 shows that the protein backbone packs tightly under the *si*-face of the FAD and the Cys42-SOH residue is in van der Waals contact with the FAD leaving no space for a C4a-peroxy adduct due to steric crowding. However, the alternative model shows that a C4a-peroxy adduct on the *re*-face is more likely due to available space. Furthermore, if a C4a-peroxy adduct does indeed form on the *re*-face of

the FAD, there is a short channel that traverses from the *re* to the *si*-face. This would allow for the product hydrogen peroxide to traverse to the *si* face and react with the Cys42-S⁻ residue. Thus, our structure predicts that due to steric crowding of the FAD-C4a position on the *si*-face, it is sterically more favorable for a potential C4a-peroxy adduct to project from *re*-face of the flavin. In order for a dioxygen adduct to form on the *si*-face of the FAD, a conformational change involving Cys42 would be necessary.

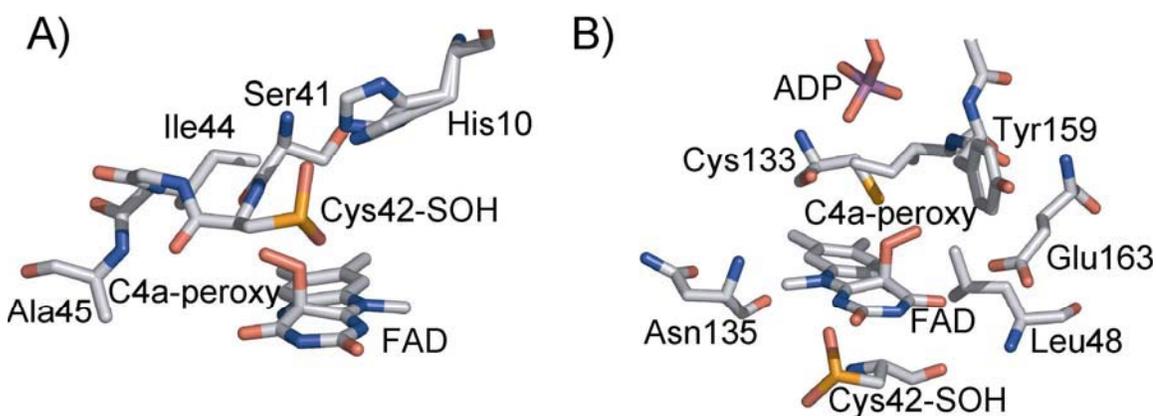


Figure 4.17. Hypothetical models of the FAD-C4a:H₂O₂ complexes modeled with MOPAC simulations in CHEM3D using either a A) *si*-face complex or a B) *re*-face complex. Each model was overlaid with the isoalloxazine ring of *L.san*-Nox2. Note in panel A, that the C4a-H₂O₂ would result in a steric clash between the sulfur of Cys42 and the dioxygen adduct, whereas in panel B there are no steric clashes with the dioxygen adduct. The Cys42-SOH and His10 residues are modeled with both alternate conformations. The C, N, O, S, and P atoms are colored in gray, blue, red, orange, and magenta, respectively.

Mechanistic Insights

The high resolution crystal structure of *L.san*-Nox2 provides structural information for better understanding the reaction mechanism. The oxidation state of the redox-active Cys42 residue provides critical information for rationalizing the lack of

H₂O₂ release from the active site of *L.san*-Nox2 (Figure 4.18). Cys42 is proposed to alternate between the thiol/thiolate (E) and the sulfenic acid states (E*) during turnover. Analysis of the two C4a-peroxy flavin models suggests that the enzyme may react with O₂ at the FAD *re*-face to yield a C4a-peroxy intermediate similar to that observed in the *Enterococcus faecalis* NADH oxidase C42S mutant. The first equivalent of NAD(P)H during the first oxidative half reaction may enter the active site via a channel traversing from the solvent toward the *re*-face of the FAD, reduce the FAD and yield H₂O₂ after reaction of the reduced FAD with O₂. The existence of a nearby channel from the *re*-face traversing to the *si*-face suggests that upon break down of the C4a-peroxy intermediate, dissociated H₂O₂ is delivered to Cys42 allowing for Cys42-S⁻ to react with H₂O₂ via a nucleophilic attack to yield Cys42-SOH intermediate which is stabilized via hydrogen bonds to His10 and the FAD. The bound ADP would thus also potentially serve the role of steering H₂O₂ toward Cys42-S⁻ by blocking the escape of H₂O₂ from the *re*-face. The second equivalent of NAD(P)H is then used to deliver the second set of electrons to the FAD. Since the FADH⁻ is in van der Waals contact with Cys42-SOH (E*), it transfers two electrons to Cys42-SOH and reduces the sulfenic acid moiety to the thiolate (E) and releases the second H₂O molecule. The importance of Cys42 in preventing H₂O₂ escape has been demonstrated by site-directed mutagenesis in NADH peroxidase in which Cys42 was mutated to Ala or Ser and also in the *Enterococcus faecalis* NADH oxidase C42S mutant. These mutants generated H₂O₂ as opposed to H₂O (33-35).

Our crystal structure is most consistent with the E* form indicated in the reaction mechanism. Based on the crystal structure, we propose that the Cys42-SOH residue has two possibilities for conversion back to the thiol/thiolate state of the resting enzyme (E).

First, the NAD(P)H-dependent process is similar to the first half of the reaction cycle. However, since one conformation of the Cys42-SOH is in van der Waals contact with the reduced FAD, the reduction of the Cys42-SOH is predicted to be much faster than aberrant reaction with O₂. Second, the DTT-dependent process is analogous to that of the non-flavin human peroxidase which also passes through a cysteine sulfenic acid intermediate (48, 88). It is also consistent with a large, DTT-dependent increase in total turnover number observed in the catalytic process. Thus, *L.san*-Nox2 has two relatively independent mechanisms to eliminate O₂ and neither pathway releases H₂O₂. For example, during oxidative stress and conditions with low availability of NAD(P)H, the thiol-dependent process provides protection. In contrast, if thiol based reagents are scarce during oxidative stress, but the NAD(P)H level is abundant, then the *L.san*-Nox2 still affords protection. Thus, to our knowledge, our laboratories are the first to describe these two types of structural contributions toward the stoichiometric conversion of one O₂ molecule into two H₂O molecules in any flavoprotein.

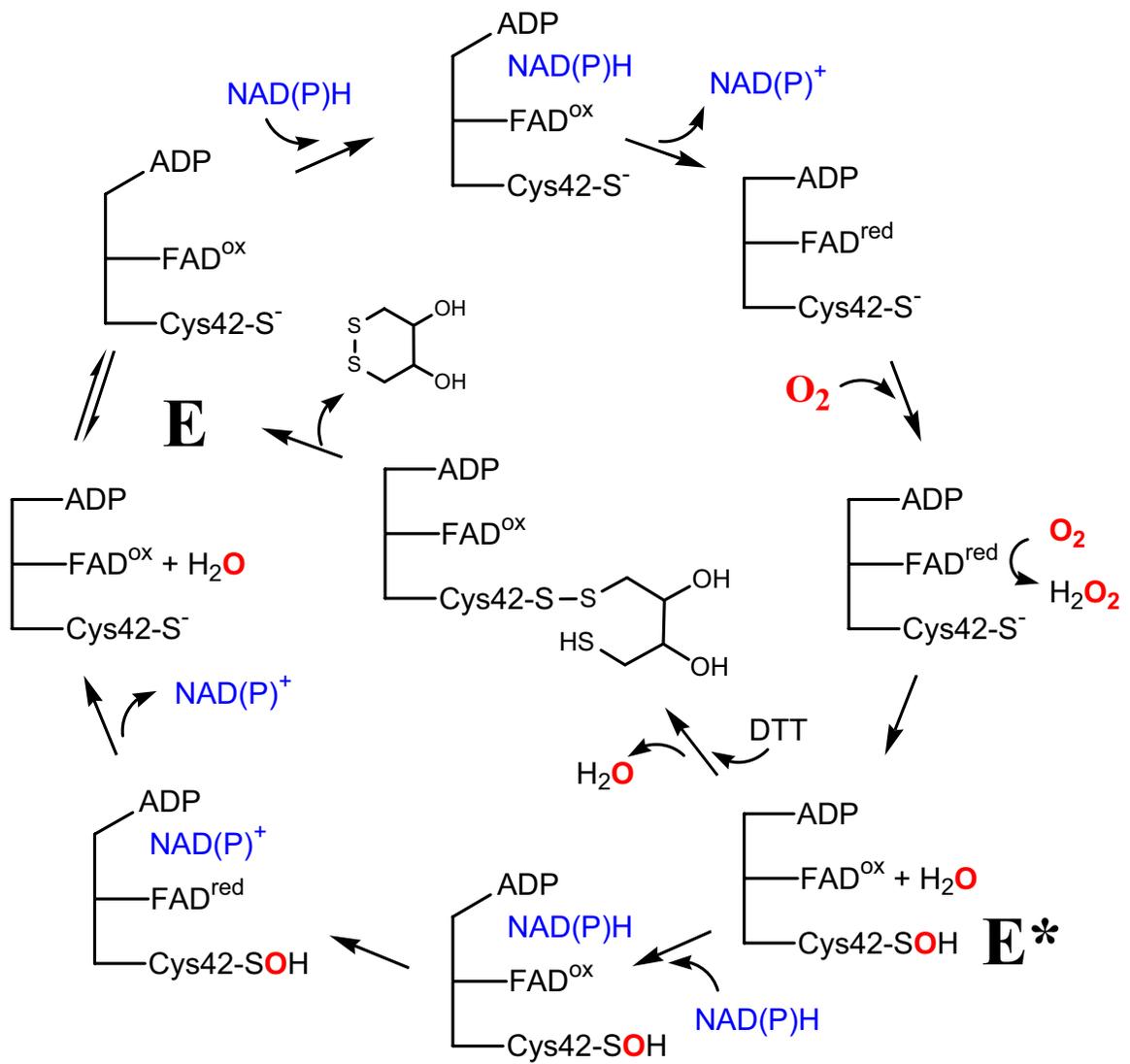


Figure 4.18. Structure-based reaction mechanism for *L.san*-Nox2

Conclusion

Our high-resolution crystal structure of *L.san*-Nox2 provides the first structure of a water-forming NADH oxidase and allows for analysis of the structural basis for the conversion of O₂ into two molecules of H₂O by the FAD-dependent enzyme. The determination of the X-ray structure confirms that *L.san*-Nox2 is a structural homolog of the pyridine nucleotide disulfide reductase family of enzymes and most closely resembles the structure of NADH peroxidase. The structure provides a detailed look at the active site of the enzyme and confirms the presence of Cys42-SOH that is proposed to play a critical role in the conversion of H₂O₂ into H₂O via a thiolate intermediate. Furthermore, the discovery of a high affinity molecule which copurifies and cocrystallizes with the enzyme raises important new questions on the role of ADP in the enzyme mechanism. Further biochemical studies are needed to fully understand the role of the bound ligand. The structure also provides further insights into the reaction mechanism involving the alternating states of the Cys42-SOH and thiolate/thiol forms of Cys42 that may involve reaction with reducing agents such as DTT. Finally, the structure provides a basis for future site-directed mutagenesis experiments to probe the mechanistic role of certain active site residues. The structure of *L.san*-Nox2 will also further provide critical structural information in its development as a potential biocatalyst for the regeneration of oxidized pyridine nucleotides (29, 30).

References

1. De Angelis, M., and Gobbetti, M. (2004) *Proteomics* 4, 106-22.
2. Bolotin, A., Wincker, P., Mauger, S., Jaillon, O., Malarne, K., Weissenbach, J., Ehrlich, S. D., and Sorokin, A. (2001) *Genome Res* 11, 731-53.
3. Poole, L. B. (2005) *Arch Biochem Biophys* 433, 240-54.
4. Baker, L. M., Raudonikiene, A., Hoffman, P. S., and Poole, L. B. (2001) *J Bacteriol* 183, 1961-73.
5. Klenk, H. P., Clayton, R. A., Tomb, J. F., White, O., Nelson, K. E., Ketchum, K. A., Dodson, R. J., Gwinn, M., Hickey, E. K., Peterson, J. D., Richardson, D. L., Kerlavage, A. R., Graham, D. E., Kyrpides, N. C., Fleischmann, R. D., Quackenbush, J., Lee, N. H., Sutton, G. G., Gill, S., Kirkness, E. F., Dougherty, B. A., McKenney, K., Adams, M. D., Loftus, B., Venter, J. C., and et al. (1997) *Nature* 390, 364-70.
6. De Angelis, M., Bini, L., Pallini, V., Cocconcelli, P. S., and Gobbetti, M. (2001) *Microbiology* 147, 1863-73.
7. Kleerebezem, M., Boekhorst, J., van Kranenburg, R., Molenaar, D., Kuipers, O. P., Leer, R., Turchini, R., Peters, S. A., Sandbrink, H. M., Fiers, M. W., Stiekema, W., Lankhorst, R. M., Bron, P. A., Hoffer, S. M., Groot, M. N., Kerkhoven, R., de Vries, M., Ursing, B., de Vos, W. M., and Siezen, R. J. (2003) *Proc Natl Acad Sci U S A* 100, 1990-5.
8. Rabus, R., Ruepp, A., Frickey, T., Rattei, T., Fartmann, B., Stark, M., Bauer, M., Zibat, A., Lombardot, T., Becker, I., Amann, J., Gellner, K., Teeling, H., Leuschner, W. D., Glockner, F. O., Lupas, A. N., Amann, R., and Klenk, H. P. (2004) *Environ Microbiol* 6, 887-902.
9. Gobbetti, M., and Corsetti, A. (1997) *Food Microbiology* 14, 175-188.
10. Siezen, R. J., van Enckevort, F. H., Kleerebezem, M., and Teusink, B. (2004) *Curr Opin Biotechnol* 15, 105-15.
11. de Vos, W. M., Bron, P. A., and Kleerebezem, M. (2004) *Curr Opin Biotechnol* 15, 86-93.

12. Klaenhammer, T., Altermann, E., Arigoni, F., Bolotin, A., Breidt, F., Broadbent, J., Cano, R., Chaillou, S., Deutscher, J., Gasson, M., van de Guchte, M., Guzzo, J., Hartke, A., Hawkins, T., Hols, P., Hutkins, R., Kleerebezem, M., Kok, J., Kuipers, O., Lubbers, M., Maguin, E., McKay, L., Mills, D., Nauta, A., Overbeek, R., Pel, H., Pridmore, D., Saier, M., van Sinderen, D., Sorokin, A., Steele, J., O'Sullivan, D., de Vos, W., Weimer, B., Zagorec, M., and Siezen, R. (2002) *Antonie Van Leeuwenhoek* 82, 29-58.
13. Marty-Teyssset, C., de la Torre, F., and Garel, J. (2000) *Appl Environ Microbiol* 66, 262-7.
14. Kengen, S. W., van der Oost, J., and de Vos, W. M. (2003) *Eur J Biochem* 270, 2885-94.
15. Higuchi, M., Yamamoto, Y., Poole, L. B., Shimada, M., Sato, Y., Takahashi, N., and Kamio, Y. (1999) *J Bacteriol* 181, 5940-7.
16. Gibson, C. M., Mallett, T. C., Claiborne, A., and Caparon, M. G. (2000) *J Bacteriol* 182, 448-55.
17. Crane, E. J., 3rd, Yeh, J. I., Luba, J., and Claiborne, A. (2000) *Biochemistry* 39, 10353-64.
18. Stehle, T., Ahmed, S. A., Claiborne, A., and Schulz, G. E. (1991) *J Mol Biol* 221, 1325-44.
19. Beres, S. B., Sylva, G. L., Barbian, K. D., Lei, B., Hoff, J. S., Mammarella, N. D., Liu, M. Y., Smoot, J. C., Porcella, S. F., Parkins, L. D., Campbell, D. S., Smith, T. M., McCormick, J. K., Leung, D. Y., Schlievert, P. M., and Musser, J. M. (2002) *Proc Natl Acad Sci U S A* 99, 10078-83.
20. Kuriyan, J., Krishna, T. S., Wong, L., Guenther, B., Pahler, A., Williams, C. H., Jr., and Model, P. (1991) *Nature* 352, 172-4.
21. Karplus, P. A., and Schulz, G. E. (1987) *J Mol Biol* 195, 701-29.
22. Waksman, G., Krishna, T. S., Williams, C. H., Jr., and Kuriyan, J. (1994) *J Mol Biol* 236, 800-16.
23. Mattevi, A., Obmolova, G., Kalk, K. H., van Berkel, W. J., and Hol, W. G. (1993) *J Mol Biol* 230, 1200-15.
24. Zhang, Y., Bond, C. S., Bailey, S., Cunningham, M. L., Fairlamb, A. H., and Hunter, W. N. (1996) *Protein Sci* 5, 52-61.

25. Argyrou, A., Vetting, M. W., and Blanchard, J. S. (2004) *J Biol Chem* 279, 52694-702.
26. Jacob, C., Giles, G. I., Giles, N. M., and Sies, H. (2003) *Angew Chem Int Ed Engl* 42, 4742-58.
27. Poole, L. B., Karplus, P. A., and Claiborne, A. (2004) *Annu Rev Pharmacol Toxicol* 44, 325-47.
28. Aslund, F., and Beckwith, J. (1999) *Cell* 96, 751-3.
29. Riebel, B. R., Gibbs, P. R., Wellborne, W. B., and Bommarius, A. S. (2002) *Adv. Synth. Catal.* 344, 1156-1168.
30. Riebel, B. R., Gibbs, P. R., Wellborne, W. B., and Bommarius, A. S. (2003) *Adv. Synth. Catal.* 345, 707-712.
31. Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., 3rd, Charrier, V., and Parsonage, D. (1999) *Biochemistry* 38, 15407-16.
32. Mallett, T. C., Parsonage, D., and Claiborne, A. (1999) *Biochemistry* 38, 3000-11.
33. Mallett, T. C., and Claiborne, A. (1998) *Biochemistry* 37, 8790-802.
34. Mande, S. S., Parsonage, D., Claiborne, A., and Hol, W. G. (1995) *Biochemistry* 34, 6985-92.
35. Parsonage, D., and Claiborne, A. (1995) *Biochemistry* 34, 435-41.
36. Fraaije, M. W., and Mattevi, A. (2000) *Trends Biochem Sci* 25, 126-32.
37. Massey, V. (2000) *Biochem Soc Trans* 28, 283-96.
38. Claiborne, A., Mallett, T. C., Yeh, J. I., Luba, J., and Parsonage, D. (2001) *Adv Protein Chem* 58, 215-76.
39. Saurin, A. T., Neubert, H., Brennan, J. P., and Eaton, P. (2004) *Proc Natl Acad Sci U S A* 101, 17982-7.
40. Carballal, S., Radi, R., Kirk, M. C., Barnes, S., Freeman, B. A., and Alvarez, B. (2003) *Biochemistry* 42, 9906-14.
41. Rhee, K. Y., Erdjument-Bromage, H., Tempst, P., and Nathan, C. F. (2005) *Proc Natl Acad Sci U S A* 102, 467-72.
42. Shishehbor, M. H., and Hazen, S. L. (2004) *Curr Atheroscler Rep* 6, 243-50.

43. Virdis, A., Neves, M. F., Amiri, F., Touyz, R. M., and Schiffrin, E. L. (2004) *J Hypertens* 22, 535-42.
44. Floyd, R. A., and Hensley, K. (2002) *Neurobiol Aging* 23, 795-807.
45. Klaunig, J. E., and Kamendulis, L. M. (2004) *Annu Rev Pharmacol Toxicol* 44, 239-67.
46. Qureshi, G. A., Baig, S., Sarwar, M., and Parvez, S. H. (2004) *Neurotoxicology* 25, 121-38.
47. Yeh, J. I., Claiborne, A., and Hol, W. G. (1996) *Biochemistry* 35, 9951-7.
48. Woo, H. A., Kang, S. W., Kim, H. K., Yang, K. S., Chae, H. Z., and Rhee, S. G. (2003) *J Biol Chem* 278, 47361-4.
49. Canet-Aviles, R. M., Wilson, M. A., Miller, D. W., Ahmad, R., McLendon, C., Bandyopadhyay, S., Baptista, M. J., Ringe, D., Petsko, G. A., and Cookson, M. R. (2004) *Proc Natl Acad Sci U S A* 101, 9103-8.
50. Zhang, Z. Y. (2002) *Annu Rev Pharmacol Toxicol* 42, 209-34.
51. Mallis, R. J., Hamann, M. J., Zhao, W., Zhang, T., Hendrich, S., and Thomas, J. A. (2002) *Biol Chem* 383, 649-62.
52. Mallis, R. J., Buss, J. E., and Thomas, J. A. (2001) *Biochem J* 355, 145-53.
53. Lin, T. K., Hughes, G., Muratovska, A., Blaikie, F. H., Brookes, P. S., Darley-Usmar, V., Smith, R. A., and Murphy, M. P. (2002) *J Biol Chem* 277, 17048-56.
54. Mustelin, T., Alonso, A., Bottini, N., Huynh, H., Rahmouni, S., Nika, K., Louisdit-Sully, C., Tautz, L., Togo, S. H., Bruckner, S., Mena-Duran, A. V., and al-Khoury, A. M. (2004) *Mol Immunol* 41, 687-700.
55. Alonso, A., Sasin, J., Bottini, N., Friedberg, I., Osterman, A., Godzik, A., Hunter, T., Dixon, J., and Mustelin, T. (2004) *Cell* 117, 699-711.
56. Park, S. U., Ferrer, J. V., Javitch, J. A., and Kuhn, D. M. (2002) *J Neurosci* 22, 4399-405.
57. Liu, Y., and Gutterman, D. D. (2002) *Clin Exp Pharmacol Physiol* 29, 305-11.
58. Stadtman, E. R. (2001) *Ann N Y Acad Sci* 928, 22-38.
59. Miller, H., Mande, S. S., Parsonage, D., Sarfaty, S. H., Hol, W. G., and Claiborne, A. (1995) *Biochemistry* 34, 5180-90.

60. Lountos, G. T., Riebel, B. R., Wellborn, W. B., Bommarius, A. S., and Orville, A. M. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2044-7.
61. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* 54 (Pt 5), 905-21.
62. Kleywegt, G. J., and Jones, T. A. (1997) *Methods Enzymol.* 277, 208-230.
63. Jones, T. A., and Kjeldgaard, M. (1997) *Methods Enzymol.* 277, 173-208.
64. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr D Biol Crystallogr* 53, 240-55.
65. Collaborative Computational Project, N. (1994) *Acta. Cryst D*50, 760-763.
66. Tickle, I. J., Laskowski, R. A., and Moss, D. S. (1998) *Acta Crystallogr D Biol Crystallogr* 54 (Pt 4), 547-57.
67. Tickle, I. J., Laskowski, R. A., and Moss, D. S. (2000) *Acta Crystallogr D Biol Crystallogr* 56 (Pt 4), 442-50.
68. Read, R. J. (1997) *Methods Enzymol.* 277, 110-128.
69. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr D Biol Crystallogr* 49, 129-47.
70. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) *J Mol Biol* 231, 1049-67.
71. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) *Proteins* 12, 345-64.
72. Guex, N., and Peitsch, M. C. (1997) *Electrophoresis* 18, 2714-23.
73. Pedretti, A., Villa, L., and Vistoli, G. (2004) *J.C.A.M.D* 18, 167-173.
74. Schuttelkopf, A. W., and van Aalten, D. M. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 1355-63.
75. Kabsch, W., and Sander, C. (1983) *Biopolymers* 22, 2577-637.
76. Krissinel, E., and Henrick, K. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2256-68.

77. Stehle, T., Claiborne, A., and Schulz, G. E. (1993) *Eur J Biochem* 211, 221-6.
78. Crane, E. J., 3rd, Parsonage, D., and Claiborne, A. (1996) *Biochemistry* 35, 2380-7.
79. Karplus, P. A., and Schulz, G. E. (1989) *J Mol Biol* 210, 163-80.
80. Jiang, R., and Bommarius, A. S. (2004) *Tetrahedron: Asymmetry* 15, 2939-2944.
81. Lesk, A. M. (1995) *Curr Opin Struct Biol* 5, 775-83.
82. Duax, W. L., Pletnev, V., Addlagatta, A., Bruenn, J., and Weeks, C. M. (2003) *Proteins* 53, 931-43.
83. Carugo, O., and Argos, P. (1997) *Proteins* 28, 10-28.
84. Wierenga, R. K., Terpstra, P., and Hol, W. G. (1986) *J Mol Biol* 187, 101-7.
85. Guo, F., Chang, B. H., and Rizzo, C. J. (2002) *Bioorg Med Chem Lett* 12, 151-4.
86. Rizzo, C. J. (2001) *Antioxid Redox Signal* 3, 737-46.
87. Reibenspies, J. H., Guo, F., and Rizzo, C. J. (2000) *Org Lett* 2, 903-6.
88. Woo, H. A., Chae, H. Z., Hwang, S. C., Yang, K. S., Kang, S. W., Kim, K., and Rhee, S. G. (2003) *Science* 300, 653-6.

CHAPTER 5

STRUCTURAL AND MECHANISTIC INSIGHTS FROM THE CRYSTAL STRUCTURE OF CHOLINE OXIDASE

Abstract

⁷The accumulation of glycine betaine (*N,N,N*-trimethylglycine) in many human pathogens is an essential factor in their stress response toward hyperosmotic environments. The metabolite is most commonly imported or enzymatically generated *in situ*. Choline oxidase from *Arthrobacter globiformis* catalyzes the flavin-dependent, four-electron oxidation of choline to glycine-betaine, with betaine aldehyde as a two-electron intermediate. In the two oxidative half-reactions, two molecules of O₂ are converted into two H₂O₂ molecules. We have solved the X-ray crystal structure of choline oxidase at 1.86 Å resolution using synchrotron radiation. The overall structure of monomeric choline oxidase folds into substrate and flavin-binding domains and is structurally similar to other members of the glucose-methanol-choline oxidase (GMC) family. The crystal structure further reveals a covalent linkage between the His99^{Nε2} and FAD^{C8M} atoms. Moreover, the electron density maps for the FAD also reveal an unusually distorted isoalloxazine ring system with an approximately 120° angle between the pyrimidine and the dimethylbenzene rings. The C4a atom is *sp*³ hybridized, suggesting the presence of a covalent adduct and is supported by the electron density. The C10a atom also appears to

⁷ The following chapter is based on work in collaboration between Lountos, G.T., Fan, F., Gadda, G., and Orville, A.M. Fan Fan provided purified choline oxidase and was involved in initial crystallization experiments. The author also acknowledges the efforts of Megan O'Neill during the initial crystallization trials of choline oxidase. Mahmoud Ghanem and Osbourne Quayle determined the O₂ to H₂O stoichiometry in choline oxidase. Dr. Zhongmin Jin of SER-CAT is also acknowledged for assistance with mail-in crystallography data collection. Dr. Rajeev Prabhakar of Emory University performed the DFT calculations.

be partially sp^3 hybridized. An atomic model fit to the electron density is consistent with an FAD C4a-O₂⁻ or FAD C4a-OH complex. We propose that the anionic complex is generated *in situ* via photo-reduction from the synchrotron X-ray irradiation, followed by O₂ binding. The complex does not release H₂O₂ because the cryogenic conditions do not establish the appropriate proton inventory on the surrounding residues. The unusual feature of the FAD was also confirmed with another data set collected from a different crystal that diffracted to 2.69 Å structure. This is the first direct observation of an oxygen reaction intermediate in any flavoenzyme oxidase by X-ray crystallography. Additionally, the high-resolution crystal structure reveals a cavity in the substrate-binding domain that is sealed off from the exterior of the protein. A model for the choline substrate can be positioned into the cavity and reveals important structural information that allows for the identification of the putative binding site for choline and identification of residues involved in the catalytic mechanism. The availability of the crystal structure of choline oxidase in combination with emerging biochemical and mechanistic data provide many new insights into the structure-function relationship of the enzyme.

Introduction

Flavins and flavoproteins were first discovered in the 1930s (1-3). Over the past seven decades chemists and biochemists have continued to characterize their remarkable diversity (4). It is now recognized that about 4% of microbial proteins are thought to be associated with either an FMN or FAD cofactor. Indeed, in the annotated genome of *Escherichia coli*, there are 205 different predicted flavoproteins. Furthermore, in the

genome of *Homo sapiens* there are more than 360 predicted flavoproteins⁸. A major reason for the ubiquity of flavin-dependent enzymes lies in the chemical versatility of the flavin isoalloxazine ring (4). This heterotricyclic organic cofactor (Figure 5.1) is ideally suited for oxidative or reductive reactions involving one or two electron transfer to and from other redox-active centers as well as reactivity with molecular oxygen (4-6). A manifestation of flavin versatility is the unique ability of flavin-dependent proteins to catalyze a wide range of biochemical reactions ranging from aerobic and anaerobic metabolism, light emission, photosynthesis, DNA repair, plant phototropism, the activation of dioxygen for oxidation and hydroxylation reactions, and regulation of biological clocks (4, 5, 7-18). From a chemical standpoint, the isoalloxazine ring can act as an electrophile, by accepting a hydride equivalent, or a nucleophile, by forming covalent adducts with either protein residues or reaction intermediates at the C4a, N5, C6, and/or C8M positions (see Figure 5.1 for nomenclature) (4).

⁸ G.A Reid, paper presented at the 14th International Symposium on Flavins and Flavoproteins, Cambridge, UK 2002

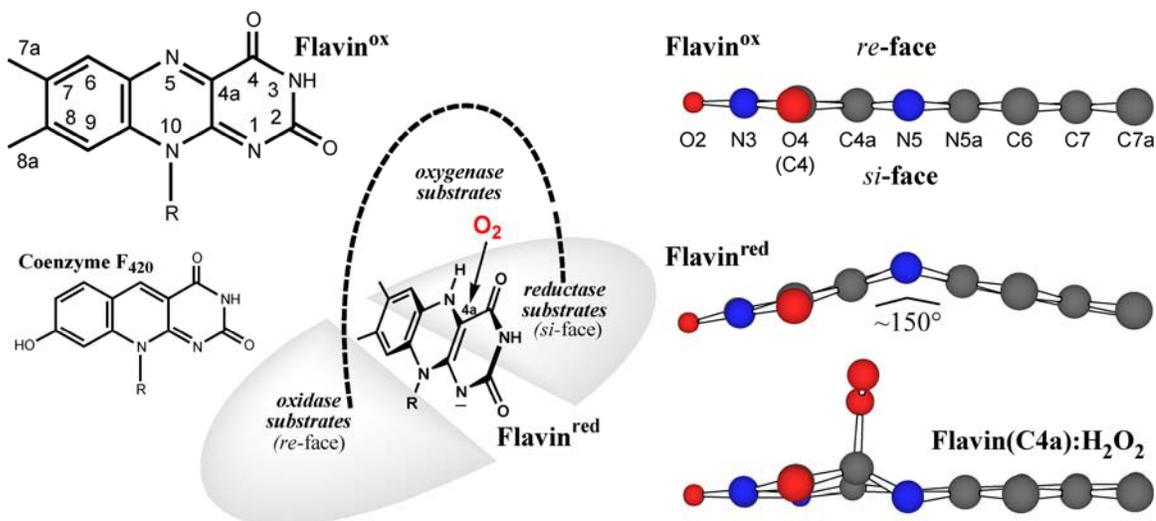


Figure 5.1. A structure-based hypothesis for reactivity in flavoenzymes.

As described above, many biological transformations of substrates are either catalyzed by, or otherwise involve flavoenzymes. FAD and FMN are the most common cofactors found in enzymes (4, 7). Consequently, they facilitate very diverse catalytic activities (19). A myriad of fundamentally similar interactions influence each reaction coordinate, yet, yield remarkable catalytic diversity. Moreover, selective pressure and evolution have enhanced “desirable” reactions, while simultaneously decreasing “deleterious” reactions. Although more than 600 flavoprotein structures have been deposited into the Protein Data Bank, much debate remains as to how the active site microenvironment influences a given reaction coordinate. However, far less is known about how different reactions can occur.

The reactions of a particular flavoprotein derive from a complex combination of interactions. These include the flavin cofactor itself, protein active site residues, substrate

molecule(s) and inclusion or exclusion of solvent molecules and/or ions. A few well-studied flavoenzymes support proposals for several of these features. Structural analysis of several dehydrogenases, oxidases and reductases reveal that substrates often bind parallel to the isoalloxazine ring, but often on opposite faces (Figure 5.1) (7). The dehydrogenases and oxidases frequently bind substrates adjacent to the *re*-face, whereas reductases often bind substrates parallel to the *si*-face. In contrast, *p*-hydroxybenzoate 3-hydroxylase and many other oxygenases, bind substrates near the top edge of the reduced flavin (19-25). Thus, the three-dimensional binding relationship between the substrate(s) and the cofactor profoundly influences the reaction. The reduction potential of the flavin is thought to be influenced by the presence of either a positive or negative charge near the redox active N5-N1 pair (26-31). For example, a positive charge near N1 appears to increase the reduction potential and may also stabilize the anionic state at N1 in the reduced cofactor. Furthermore, hydrogen bonding to N5 is thought to polarize the LUMO and HOMO orbital of the N5 atom of oxidized flavins. Theoretical studies suggest that the hydrogen bond donor approximately in the plane of the isoalloxazine ring will orient the LUMO approximately perpendicular to the plane (32-35). Consequently, the electrophilicity at this position is increased and hydride transfer is enhanced, provided that the substrates are also suitably oriented and parallel to the plane of the isoalloxazine ring. Furthermore, gas-phase theoretical studies and crystallographic studies of free flavins and flavin analogs have shown that the two electron reduced isoalloxazine ring is bent by up to 30° along the N5-N10 axis (36-38). Such a “butterfly” conformational change associated with redox perturbation should significantly influence catalysis in flavoproteins. Reaction with molecular O₂ at the C4a or other acceptor molecules at

either the C4a or N5 atoms can occur from either the *si* or *re*-face. However, there are also structurally defined examples where oxidized flavoproteins have significantly bent isoalloxazine rings and other examples where reduced flavoproteins contain almost planar cofactors (22, 23, 35, 39-42). The protein environment is hypothesized to play a significant role in stabilizing these higher energy states. It is thus clearly evident that the rules which yield a given reaction coordinate remain to be defined in most flavoproteins.

A fundamental reaction of reduced flavoenzymes is with molecular oxygen, which can be orders of magnitude faster or slower than the analogous reactions of flavins in solution (6). The outcome of the reaction also varies greatly, which has been used to differentiate flavoenzymes into different classes (43). The FAD-dependent hydroxylases cleave the dioxygen O-O bond and incorporate one oxygen atom into the organic product, whereas the other oxygen atom is released as water. In contrast, ubiquitous flavin-dependent oxidases use O₂ as an electron acceptor and yield H₂O₂. Despite the different outcomes of their oxidative half-reactions, these two classes of flavoenzymes transfer two reducing equivalents from the reduced flavin to O₂ (6). In the monooxygenases, exemplified by *p*-hydroxybenzoate hydroxylase, luciferase, and microsomal flavin-containing monooxygenase, a transient C4a-hydroxy-flavin species has been established by rapid kinetic methods and its structure characterized by NMR spectroscopy (44-49). In the oxidases, with the exception of rapid kinetic studies on a mutant form of NADH oxidase (50) and pulse radiolysis experiments with glucose oxidase (51) showing that a flavin hydroperoxide species can be obtained from the neutral flavin radical and superoxide anion, a flavin C4a-hydroperoxide has not been detected. This has prompted some to propose that in oxidases, flavin reoxidation proceeds via an outer-sphere electron

transfer process without the involvement of a C4a-hydroperoxide intermediate (6, 52). Regardless of the mechanism, there are no structurally defined C4a-hydroperoxide adducts, despite over 600 flavoprotein structures deposited in the Protein Data Bank to date (53).

Choline oxidase (E.C. 1.1.3.17) catalyzes the four-electron oxidation of choline to glycine betaine (N,N,N-trimethylglycine) via two sequential, FAD-dependent reactions in which betaine aldehyde is formed as an obligatory enzyme-bound intermediate (Figure 5.2) (54, 55). The FAD cofactor is covalently linked via the C8M position to a histidine residue resulting in a midpoint reduction potential in choline oxidase of ~130 mV, which is the highest determined for a flavoenzyme (56-59). In both oxidation reactions catalyzed by choline oxidase, H₂O₂ is formed in the flavin-dependent reduction of O₂ (54). Molecular oxygen acts as the primary electron acceptor in the reaction. The study of choline oxidase is also of considerable interest for medical and biotechnological applications, since intracellular accumulation of glycine betaine allows normal cell function under conditions of hyperosmotic and temperature stress in pathogenic bacteria (60-67) and transgenic plants (68-75). Consequently, the study of choline oxidase has potential for the development of therapeutic agents that inhibit glycine betaine biosynthesis and render pathogenic bacteria susceptible to either conventional treatments or the innate immune system, and for the engineering of drought and temperature resistance in economically relevant crops.

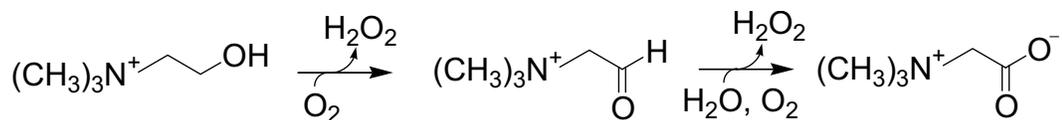


Figure 5.2. The reaction catalyzed by choline oxidase.

The oxidation of alcohols to aldehydes is catalyzed by a number of flavin-dependent enzymes including choline oxidase (54, 76, 77), choline dehydrogenase (78), glucose oxidase (79, 80), cholesterol oxidase (81), and cellobiose dehydrogenase (82). All these enzymes utilize FAD as a cofactor for catalysis and have been grouped into the glucose-methanol-choline (GMC) oxidoreductase enzyme superfamily (83). Although GMC enzymes exhibit low sequence similarity in the substrate-binding domain, the crystal structures of glucose oxidase (84, 85), cholesterol oxidase (86-88), and the flavin domain of cellobiose dehydrogenase (89) show that they all share a highly conserved catalytic site. This suggests that they may exhibit a similar activation mechanism for the oxidation of substrates. However, the nature and identity of both the catalytic base that abstracts the hydroxyl proton from the substrate and the residues that provide the necessary stabilization of the alkoxide intermediate in catalysis has not been fully determined. Mechanistic studies have been conducted on cholesterol oxidase (90, 91), cellobiose dehydrogenase (92), and glucose oxidase (17, 93, 94). A study of these enzymes has suggested that a highly conserved histidine residue found in the GMC family may act as the catalytic base that participates in the oxidation of the alcohol substrate. In choline oxidase, this residue corresponds to His466 (58). Sub-atomic resolution crystal structures from unliganded cholesterol oxidase refined up to 0.92 Å

resolution, however, indicate that the corresponding His447 residue is protonated at the Nε2 position, which points at the substrate hydroxyl group (88). Studies on the His466Ala mutant isoform of choline oxidase by Ghanem and Gadda have provided mechanistic insights into the catalytic role of this residue in choline oxidase (59). In this study, His466 was mutated to alanine by site-directed mutagenesis and the biochemical, spectroscopic, and mechanistic properties of the enzyme were evaluated. Indeed, the mutant exhibited k_{cat} and k_{cat}/K_m values with choline as the substrate that were 60- and 1000-fold lower than the values for the native enzyme. The k_{cat}/K_m value for molecular oxygen, however, was unaffected by the mutation (59). This result suggested that the residue is involved in the oxidation of choline, but not in the reduction of molecular oxygen. Furthermore, the lack of involvement of His466 in the oxidative half-reaction also agrees well with previous kinetic data on choline oxidase as a function of pH which illustrated that no ionizable groups with pK_a values of 6-10 are required for the oxidation of the enzyme-bound reduced flavin in catalysis (95). Additional data support the notion that His466 is not likely the active site base with a pK_a of 7.5 that abstracts the hydroxyl proton of the substrate in the reductive half-reaction in which choline is oxidized to betaine aldehyde. Evidence for this notion comes from the pH dependence of the imidazole effect with the His466Ala mutant which showed that imidazolium is the catalytically relevant species involved in the partial rescue of activity of the mutant isoform. Indeed, it was demonstrated that enzymatic activity could be rescued with the addition of exogenous imidazolium, but not imidazole which is consistent with the notion that the His466 residue is indeed protonated (59). These data raise questions of the notion that the His466 residue is the active site base involved in the oxidation of choline because

if this were the case, maximal rescue of enzymatic activity of the mutant enzyme should have been observed with increasing pH. Thus, it has been proposed that His466 contributes to catalysis by modulating the electrophilicity of the FAD and the polarity of the active site. In addition, it is suggested that the residue is involved in the stabilization of the negative charge of the alkoxide intermediate that is formed in the oxidation of choline to betaine aldehyde (Figure 5.3). Evidence for stabilization of the alkoxide intermediate comes from the observation that there is a complete loss of the ability of the His466Ala mutant to form an N5-flavin adduct with sulfite, which is restored in the presence of exogenous imidazole (59).

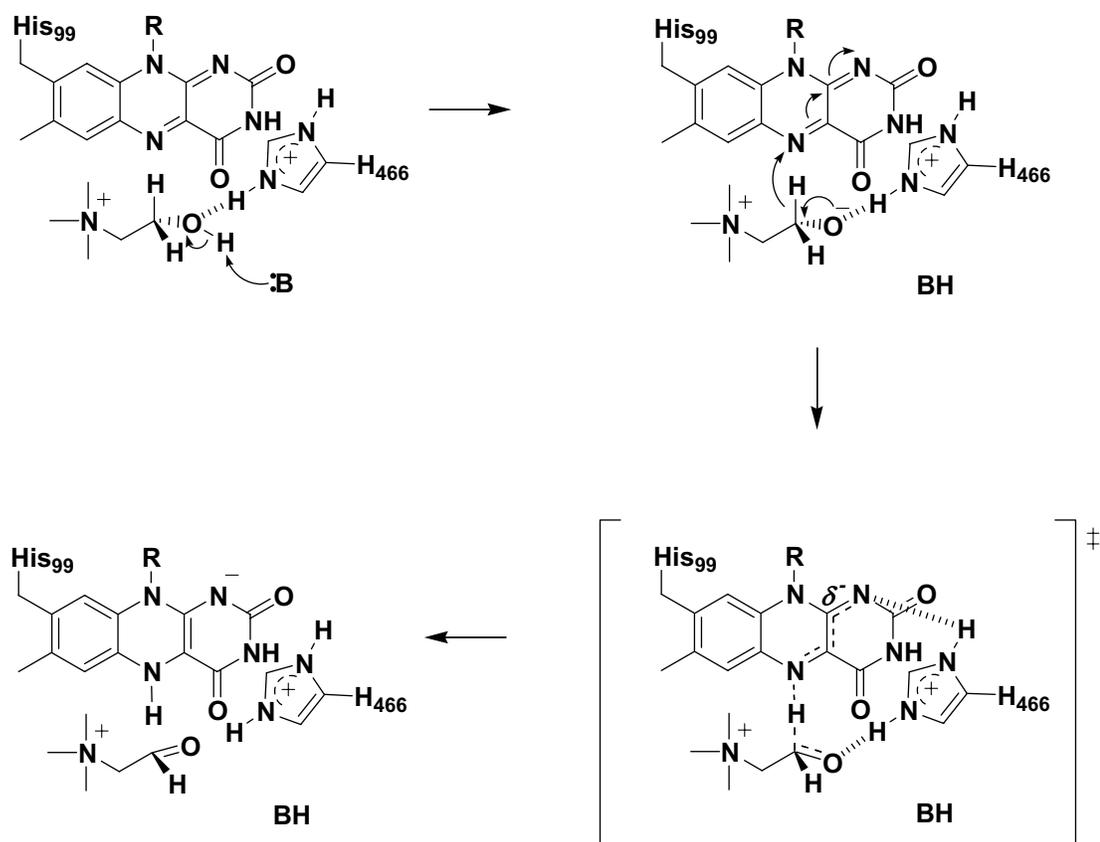


Figure 5.3. The asynchronous hydride transfer mechanism for choline oxidase with the alkoxide intermediate stabilized by His466. Adapted from (59)

Several possible mechanisms for the activation of unpolarized alcohols have been proposed for the members of the GMC family of enzymes. X-ray crystal structures of glucose oxidase (85, 93), cholesterol oxidase (87, 88), and the flavin domain of cellobiose dehydrogenase (89, 96), support the notion that the catalytic mechanism involves the removal of the hydroxyl proton by an active site base and concomitant transfer of a hydride from the substrate α -carbon to the flavin cofactor (Figure 5.4, path a). Mechanistic studies using kinetic isotope effects to probe the relative timing of OH and CH bond cleavages have been hindered in glucose oxidase (97) and cholesterol oxidase (91, 98) due to the occurrence of steps that are slower than those that directly involved in the reductive or oxidative reactions. An asynchronous hydride transfer mechanism has been suggested based on studies on methanol oxidase (99) using isotope effects and substrate analogs in which OH bond cleavage occurs to a great extent before the CH bond cleavage (Figure 5.4, path b). Alternatively, it has been suggested from crystallographic studies on cholesterol oxidase (87) and cellobiose dehydrogenase (96) that a single electron is transferred to the flavin concomitantly with abstraction of the hydroxyl proton which is then followed by the transfer of the α -hydrogen to the flavin (Figure 5.4, path c). Biochemical studies, however, have not proved successful in providing evidence for such a mechanism involving the formation of the two-radical species. Additionally, a carbanion mechanism has been proposed for the oxidation of polarized alcohols catalyzed by flavocytochrome b_2 (100, 101) in which catalysis is initiated by abstraction of the substrate α -proton by an active site base to form a carbanion, followed by the formation of a covalent N5-flavin adduct that would subsequently decay to yield the aldehyde product and reduced flavin. However, this

mechanism is least likely for the oxidation of unpolarized alcohols due to energetic factors related to the stabilization of the negative charge (102).

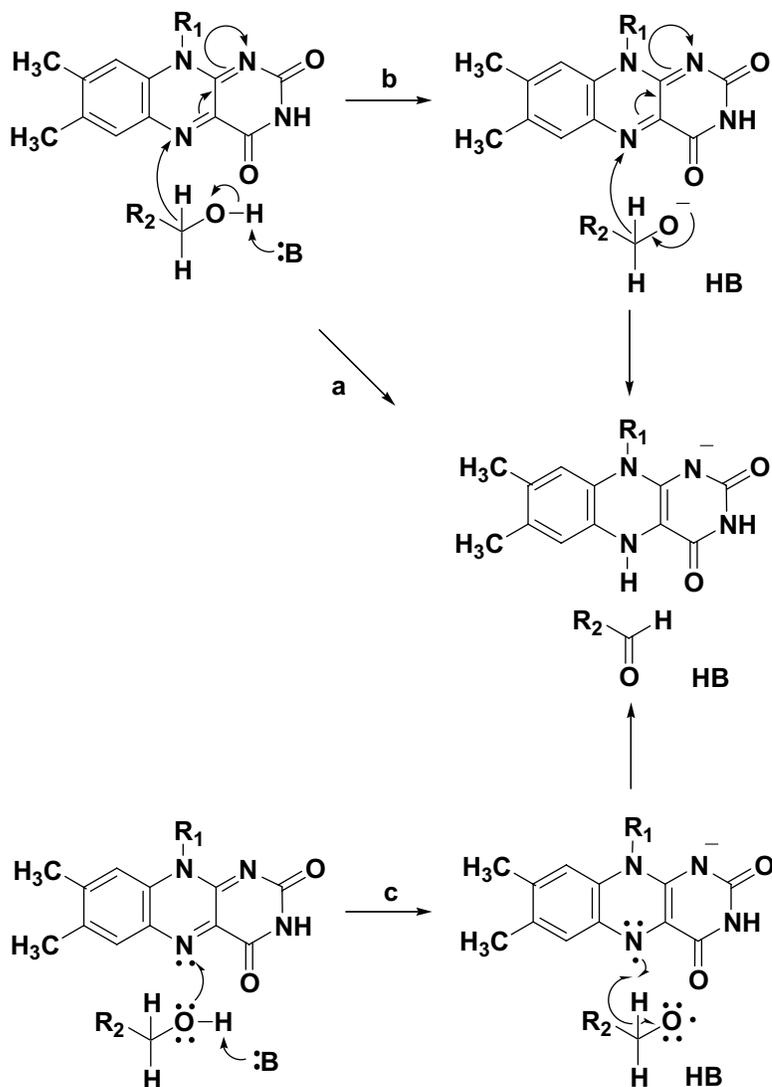


Figure 5.4. The three proposed mechanisms for the oxidation of unpolarized alcohols by members of the GMC family include a) a concerted hydride transfer mechanism, b) an asynchronous hydride transfer mechanism, and c) an oxygen radical mechanism. Adapted from (77).

In a study of choline oxidase using primary deuterium and solvent kinetic isotope effects, insights into the mechanism for substrate oxidation by choline oxidase using both

steady-state kinetics and rapid kinetic techniques have been obtained by Fan and Gadda (77). Kinetic isotope effects with isotopically substituted substrate and solvent have been used to obtain insights into the mechanism of alcohol oxidation in the reaction catalyzed by choline oxidase. The reaction involves the removal of both the hydroxyl proton and the hydrogen bound to the α -carbon of the choline substrate in either a concerted or stepwise fashion, depending on the relevant timing of CH and OH bond cleavage. In this study, the oxidation of choline catalyzed by choline oxidase was found to occur through the formation of an alkoxide species resulting from the removal of the substrate hydroxyl proton occurring before hydride transfer to the flavin. Evidence for this mechanism, in which the choline alkoxide proton is not in flight in the transition state for CH bond cleavage, comes from the substrate and solvent deuterium kinetic isotope effects determined on the rate of reduction of the flavin by choline, showing a substrate isotope effect of ~ 9 and a solvent isotope effect of unity (77). The data rule out a concerted hydride transfer mechanism in which both the proton and the hydride are concomitantly in flight in the transition state. Thus the emerging biochemical data suggest that choline oxidase follows an asynchronous hydride transfer mechanism (77).

A detailed picture of the catalytic mechanism of choline oxidase has emerged from pH and kinetic isotope effect studies (55, 77, 103), as well as mechanistic studies with substrate and product analogues (104), site-directed mutants (59), and X-ray crystallography (Figure 5.5). During enzymatic turnover with choline, a catalytic base with pK_a of 7.5 activates choline by abstracting the hydroxyl proton from the alcohol substrate (95). Stabilization of the transient choline-alkoxide species is provided by electrostatic interactions with the positively charged imidazolium of His466 (59) which is

located at $\sim 3.3 \text{ \AA}$ from the FAD N1 atom. A hydride is then transferred from the α -carbon of the activated alcohol to the FAD N5 atom, resulting in the reduction of the flavin. Thus, the first reductive half-reaction during turnover with choline provides a proton and a hydride to the enzyme and the flavin, respectively (77). In the subsequent oxidative half reaction, two electrons are transferred from the reduced flavin to O_2 . The final delivery of two protons from the catalytic base, shown as His351, and the FAD N5 atom yields oxidized FAD and H_2O_2 , which readily dissociates from the enzyme active site. Enzymatic turnover is then completed with a second oxidation reaction in which the enzyme-bound aldehyde is oxidized to a glycine betaine, although this reaction has not been mechanistically characterized yet.

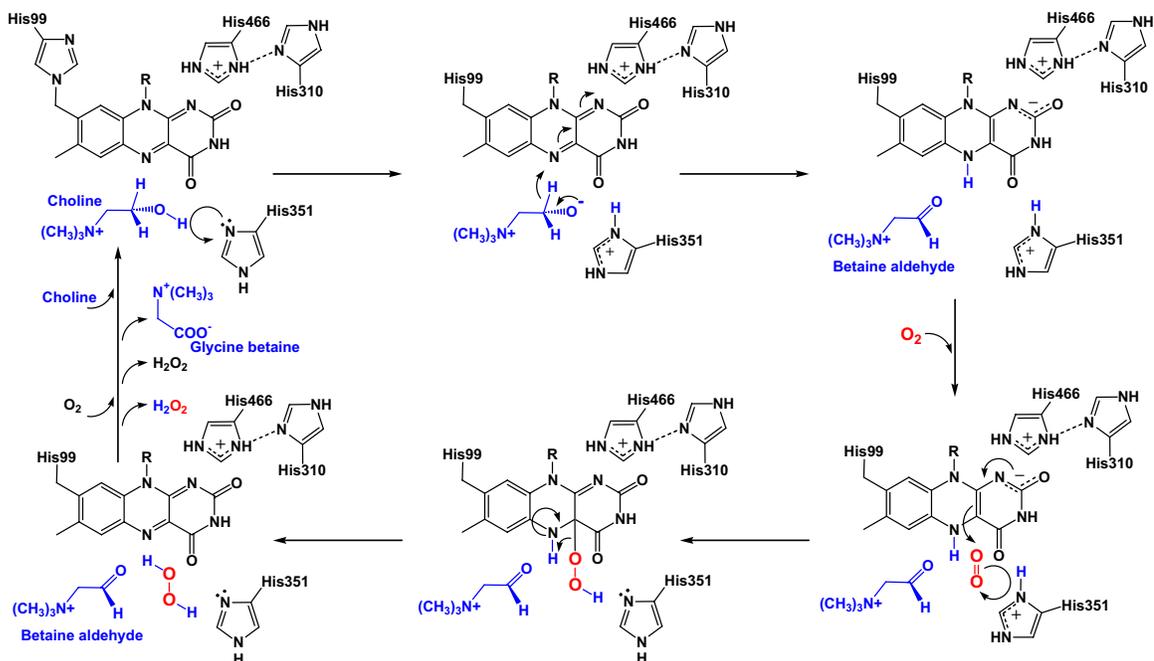


Figure 5.5. The reaction mechanism for the oxidation of choline to glycine betaine by choline oxidase based on mechanistic and structural studies.

Here, we report the crystal structure of choline oxidase refined at resolutions of 1.86 Å and 2.69 Å from two independent crystals and data sets. Since nearly identical structures were obtained from both data sets, we will focus the structure description on the highest-resolution structure. In preparation for the X-ray diffraction data collection, the crystals of oxidized enzyme in aerobic mother liquor were flash cooled by plunging them into liquid N₂. The crystals were yellow after freezing which is consistent with the oxidized FAD within the enzyme. The diffraction data were collected under cryogenic conditions using synchrotron X-rays. However, the structure reveals a surprising feature of the active site FAD in which it adopts a highly unusual conformation with respect to the isoalloxazine ring and appears to contain a C4a-adduct species. The determination of the crystal structure of choline oxidase provides detailed structural information that can be correlated with emerging biochemical and mechanistic studies of the enzyme. In addition, it provides important information in the design of new site-directed mutants to further probe the chemical mechanism and provides novel insights into the identity of specific residues that may be important constituents in the catalytic mechanism.

Materials and Methods

Enzyme Preparation, Purification, Crystallization, & X-ray Data Collection

Choline oxidase from *Arthrobacter globiformis* was cloned, expressed, and purified in high purity and yield as previously described (58) Fully oxidized enzyme was obtained as previously described (104). Crystals of choline oxidase were grown by the hanging drop vapor diffusion method with standard VDX plates from Hampton Research (Aliso Viejo, CA). Initial crystallization conditions for choline oxidase were obtained

from commercially available sparse-matrix screening kits from Hampton Research and Nextal Biotechnologies (Montreal, Quebec). Typically, 2 μ L of protein solution (4.9 mg/mL) was mixed with 2 μ L of crystallization solution on a silanized cover slip and sealed over the reservoir solution. Screening experiments were conducted by incubating the plates at 4°, 15°, and 23° C for each type of experiment. Initial conditions for choline oxidase crystals were obtained from 0.1M Tris-HCl, pH 8.5 and 1.5 M ammonium sulfate at 23° C which yielded a shower of thin, yellow needles (Figure 5.6, panel A).

Optimization of the conditions involved screening pH , precipitant concentration, and temperature in addition to different types of buffers. Additional optimization trials involved the screening of several small molecule additives from the Hampton Research Additive Screen. From the additive screens, dimethylsulfoxide (DMSO) was selected as a supplemental additive to the crystallization solution that contributed to improved crystals. In order to obtain optimal crystals used in data collection, 2 μ L of choline oxidase (4.9 mg/mL) was mixed with 2 μ L of reservoir solution containing 0.1 M Bis Tris Propane, pH 8.5, 1.2 M ammonium sulfate, and 10% v/v dimethylsulfoxide on a silanized cover slip and sealed over the reservoir containing 1 mL of crystallization solution. The tray was incubated at 23° C for approximately two weeks. Yellow rod-like crystals were obtained within 4 days of incubation and grew to maximal dimensions of 0.2 x 0.05 x 0.05 mm within two weeks (Figure 5.6, panel B).

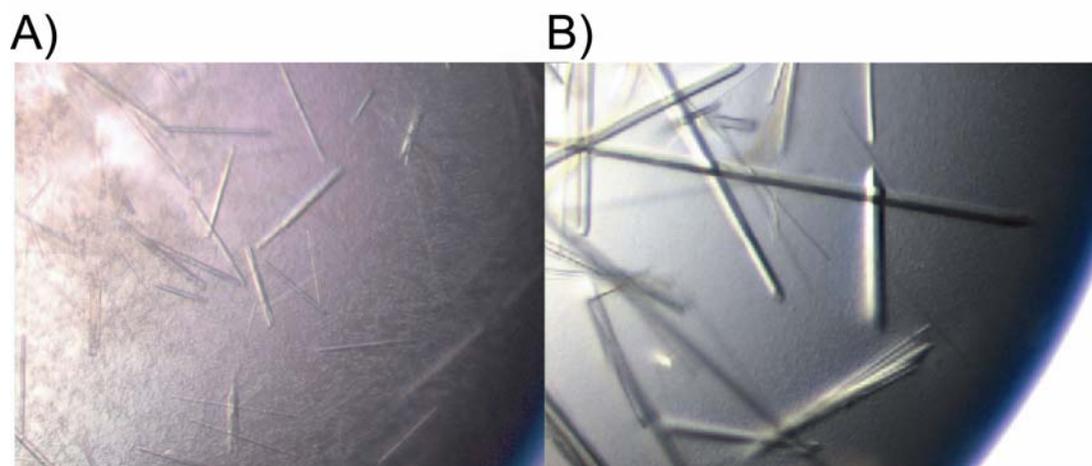


Figure 5.6. A view of crystals of choline oxidase under polarized light. Initial crystals of choline oxidase appeared as A) a shower of small needles and microcrystalline clusters. Refinement and optimization of the initial conditions resulted in B) rod like crystals which grew to a maximum dimensions of 0.2 x 0.05 x 0.05 mm which were of sufficient quality to allow for collection of X-ray diffraction data.

A single crystal for data collection at cryogenic conditions was prepared by transferring a crystal from the mother liquor into a 2 μ L solution of 3.4 M sodium malonate, pH 7.0 (105) and allowed to soak for two minutes at 23° C prior to capturing the crystal with a nylon loop. The crystal was flash-frozen by quick submersion into liquid nitrogen. Two independent data sets were collected from crystals held at approximately 100 K at the SER-CAT facilities at the Advanced Photon Source, Argonne National Laboratory. The highest resolution data set (1.86 Å) was collected at beamline 22-ID using a wavelength of 1.0 Å and a MARCCD 300 detector. The crystal to detector distance was set at 250 mm and data were collected using fine slicing with a 0.2° oscillation angle and two seconds exposure time. The second data set (2.69 Å) was collected at beamline 22-BM using a wavelength of 1.0 Å and a MARCCD 225 detector. The crystal to detector distance was set to 350 mm and data were collected with an

oscillation angle of 0.5° and a 60 second exposure time. The diffraction images were processed with HKL2000 and scaled with SCALEPACK (106).

Crystal Structure Determination

The crystal structure of choline oxidase was solved using molecular replacement. Initial phases for model refinement were obtained by molecular replacement using the coordinates of the 1.9 Å resolution structure of glucose oxidase (PDB code 1CF3) as a search model (93). Glucose oxidase exhibits 26% sequence identity with choline oxidase. The search model was modified by analyzing the sequence alignment between choline oxidase and glucose oxidase. Unaligned portions of the sequence in glucose oxidase were deleted from the model and all non-identical residues to choline oxidase were mutated to alanine. All *B*-factors for the model were set at 20.0 \AA^2 . Furthermore, all solvent, ligand, and FAD molecules were removed from the search model. Molecular replacement was performed using the program MOLREP (107) from the CCP4 suite of programs (108). Cross rotation and translation searches were performed with data from 15.0-3.5 Å resolution and searching for two molecules in the asymmetric unit. Since at this point, the two possible space groups were either $P4_32_12$ or $P4_12_12$, molecular replacement was performed in both space groups. The best solution from MOLREP in each space group was then subjected to rigid body refinement with data from 50.0-1.86 Å resolution using CNS (109). Space group $P4_32_12$ gave interpretable electron density maps and was determined to be the correct space group at this point. The starting model at this point had an R-factor of 0.48 and R-free of 0.54. Multiple rounds of manual model rebuilding and positional and isotropic *B*-factor refinement were carried out with O (110) and REFMAC5 (111), respectively. Progress of the refinement was monitored by R_{free} , which

was calculated using 5% of the reflections, and cross-validated, σ_A weighted $2mF_o-DF_c$ and mF_o-DF_c maps to evaluate the model and correct errors (112). The electron density maps obtained at this point were then vastly improved by using the PRIME & SWITCH phasing routine (113) from RESOLVE (114). In order to carry out the map improvements with PRIME & SWITCH, phases calculated from the model were obtained using SIGMAA in CCP4 and were input as the starting phases into PRIME & SWITCH. The new maps obtained from this routine resulted in electron density maps that were used to correct errors in the backbone trace and to locate the positions of the amino acid side chains.

The location of the FAD was identified after several rounds of refinement and model building and was placed into the electron density. Library files containing the topology and parameter files for the FAD were prepared using the Dundee PRODRG server (115). Initially, planar restraints were enforced on the FAD isoalloxazine ring throughout several rounds of refinement. When the overall R-factor decreased below 0.30, water molecules were located with ARP/Waters (116), refined with REFMAC5, and manually inspected. A single DMSO molecule, an additive in the crystallization solution, was fit into electron density near the N5 position of the FAD isoalloxazine ring that indicated the presence of a small molecule and refined. For comparison, a sulfate molecule (a possible candidate from the crystallization solution) was also fit in this position and refined, but DMSO was deemed to fit the density better. The electron density for the FAD molecule clearly indicated a significant bend in the isoalloxazine ring at this point in the refinements. The library file for the FAD was adjusted with respect to planar restraints and atom hybridization as suggested by the electron density

features. The isoalloxazine ring was then manually adjusted to properly fit the pyrimidine ring into the electron density and planar restraints were removed with respect to the pyrimidine ring and sp^3 hybridization was conferred upon the C4a and C10a atoms. After refinement, the pyrimidine ring fit well into the electron density, but greater than 4.5 σ difference features were observed in the mF_o-DF_c electron density maps near the C4a atom. At this point, three models were refined with different FAD models. In one model, the sp^3 C4a atom was bonded to a single oxygen atom (FAD-C4a-OH) and the model was refined after appropriate adjustment of the library files. Another model was analyzed in which the sp^3 C4a atom was bonded to a dioxygen adduct (FAD-C4a-O₂⁻) and then refined. A final model consisted of a water molecule placed in the positive difference density peak and refined. The FAD-C4a-O₂⁻ model was selected as the best model that fit the electron density and was used for the completion of the model refinements.

Occupancy refinements of the distal oxygen atom of the bound O₂ were also performed and an occupancy of 0.5 was selected for this atom. To check for model bias, simulated annealing OMIT maps (109) were prepared in CNS as follows. The FAD and DMSO atoms were removed from the model and all residues within a 3.5 Å radius of His99 were selected for exclusion from the OMIT map calculation. Simulated annealing was carried out with a starting temperature of 1000 K and steps of 50 K with data from 50.0-1.86 Å. The $2mF_o-DF_c$ and mF_o-DF_c simulated annealing OMIT maps were calculated and analyzed with respect to the model. The final refinement of the model was then completed with REFMAC5. Structure validation of model geometry was carried out with PROCHECK (117, 118).

Solvent exposed surface areas were calculated with a 1.4 Å probe radius with Swiss-PDB viewer (v3.7b2) or VEGA (<http://www.ddl.unimi.it>) . Secondary structure assignments were made using KSDSSP (119). Structural homologs in the Protein Data Bank were found using MSDFold . The rms differences between models was calculated with SSM (<http://www.ebi.ac.uk/msd-serv/ssm>) or Swiss-PDB viewer (v3.7b2). Structure figures were prepared using Swiss-PDBViewer (v3.7b2) and PovRay (v.3.5) or Pymol (DeLano Scientific LLC, Castro City, CA).

Results

Crystal Structure Determination

The X-ray data collection and refinement statistics are listed in Tables 5.1 and 5.2. High quality X-ray data were obtained from two crystals of choline oxidase and the best crystals diffracted to a resolution of 1.86 Å using synchrotron radiation. Crystals that were screened on the home source did not exhibit any diffraction, likely due to the small dimensions of the crystals, but diffracted to high resolution at the synchrotron source. The diffraction data from the highest resolution structure are consistent with either space group $P4_32_12$ or $P4_12_12$ with unit cell dimensions of $a = b = 84.4$, $c = 343.5$ Å. Matthew's coefficient (120) and solvent-content calculations suggested two molecules per asymmetric unit ($V_M = 2.5 \text{ \AA}^3 \text{ Da}^{-1}$, 50% solvent content). The data set consists of 381317 observations of 97546 unique reflections in the resolution range 50.0-1.86 Å. The overall completeness of 92.3% (63.2% in the highest resolution shell) and overall R_{sym} of 7.2% (31.5 % in the highest resolution shell) indicate the data are of good quality. A protein-protein BLAST search indicated that the protein with the highest % sequence identity that had an available crystal structure was glucose oxidase (PDB code: 1CF3)

(93). Although both proteins exhibited rather low % sequence identities when compared to each other (26%), molecular replacement was nevertheless successful in obtaining starting phases for the refinements. Molecular replacement using both possible space groups resulted in interpretable electron density only for the $P4_32_12$ space group. The model was refined against the 1.86 Å resolution data which resulted in a final R-factor of 0.161 and R-free of 0.202. The overall coordinate error of the model is 0.128 based on the R-factor and the correlation coefficient is 0.967. Ramachandran analysis of the crystal structures showed that 89.4% of the residues are in the most favored region, 10% in the additionally allowed regions, 0.6% in the generously allowed regions, and 0.1% in the disallowed regions. Ala230 in the A chain was flagged as being in the disallowed region but the electron density maps indicate this residue fits well in the observed density.

Table 5.1 X-ray Diffraction Data Collection Statistics		
Crystal ID	Crystal 1	Crystal 2
X-ray Source	SER-CAT	SER-CAT
Beamline	22-ID	22-ID
Detector	MARCCD 335	MARCCD 225
X-ray Wavelength (Å)	1.0	1.0
Resolution Range (Å) ^a	50-1.86 (1.91-1.86)	50-2.69 (2.76-2.69)
Space group	<i>P4₃2₁2</i>	<i>P4₃2₁2</i>
Unit Cell Dimensions (Å)		
<i>a</i> = <i>b</i> =	84.4	84.4
<i>c</i> =	343.5	343.7
Total Reflections	381317	196964
Unique Reflections	97546	33722
Multiplicity	3.9 (2.8)	5.8 (2.4)
Completeness (%)	92.3 (63.2)	94.2 (69.7)
R_{sym} (%) ^b	7.2 (31.5)	10.8 (27.1)
$I/(\sigma)I^c$	17.7 (2.3)	15.6 (2.8)

^a Values for the highest resolution shell of data are given in parentheses. ^b $R_{\text{sym}}(I)$ gives the average agreement between the independently measured intensities such as $\sum_h \sum_i |I_i - I| / \sum_h \sum_i I$, where I is the mean intensity of the i observations of reflection h . ^c $I/(\sigma)I$ is the root-mean-square value of the intensity measurements divided by their estimated standard deviation.

Table 5.2 Refinement Statistics for Choline Oxidase		
Crystal ID	Crystal 1	Crystal 2
Resolution Range (Å)	50-1.86	50-2.69
No. of Reflections	92403	31943
No. of protein atoms	8171	8168
No. of water molecules	960	152
R-factor	0.161	0.157
R-free ^a	0.202	0.219
Average <i>B</i> -factors (Å ²)		
Protein	22.5	25.5
Water	32.6	17.5
FAD (2)	17.4	20.4
O ₂ ⁻ (2)	24.8	37.2
DMSO (2)	40.0	44.1
Rms deviations for ideal		
Bond length (Å)	0.014	0.018
Bond angles (°)	1.5	1.7
Correlation coefficient	0.967	0.954
Estimated coordinate error	0.128	0.219

^a Calculated with 5% of the data.

Overall Structure Description

Overall, the X-ray structure reveals that choline oxidase crystallizes as a homodimer with approximate dimensions of 88 Å x 70 Å x 46 Å. The choline oxidase monomer contains 546 residues, however the last 19 residues in the C-terminal portion of the monomer were not visible in the electron density maps and thus were not included in the model. The lack of electron density for these residues suggests that the C-terminal portion of choline oxidase is disordered. The overall fold of choline oxidase resembles other members of the GMC family (83, 85, 87, 93) and adopts a two-domain topology similar to the bacterial flavoenzyme *p*-hydroxybenzoate hydroxylase or PHBH-fold (7). A search of the Protein Data Bank reveals that the closest structural homolog is glucose oxidase. When the two monomeric structures are superimposed, they overlay with a rms difference of 1.75 Å over 477 common C_α atoms. The two monomers of choline oxidase interact with each other in an anti-parallel fashion. The buried surface area at the interface of the two monomers is approximately 2429.6 Å² which corresponds to approximately 13% of the total surface area of each monomer. The two monomers in the asymmetric unit are essentially structurally equivalent with rms deviation of only 0.23 Å over 527 common C_α atoms when superimposed despite that non-crystallographic symmetry restraints were not applied during refinement. The monomeric structure consists of two domains, an FAD-binding domain and a substrate binding domain. The FAD-binding domain is formed primarily by residues 1-159, 201-311, and 464-527. The domain consists of a six-stranded parallel β-sheet that is flanked on one side by a three-stranded anti-parallel β-sheet and further surrounded by eight α-helices. The substrate binding domain is formed primarily by residues 160-200 and 312-463. The topology of

the substrate binding domain consists of a distorted six-stranded anti-parallel β -sheet which forms the bottom of the choline oxidase active site and is flanked on the other side by three α -helices which protrude into the bulk solvent. The dimerization contacts between the two monomers are rather limited. Indeed, examination of the dimer contacts reveals that the primary dimer contacts are located on the two edges of the dimer interface. The primary contacts involve interactions between charged residues Arg255-Glu370 (3.2 Å), Asp72-Lys398 (4.1 Å), Asp250-Glu53 (4.3 Å), Arg255-Glu370 (3.2 Å), Asp358-Arg396 (2.4 Å), Arg363-Asp394 (4.2 Å), and Arg363-Asp397 (2.9 Å). The middle portion of the dimer interface contains very few close contacts between subunits.

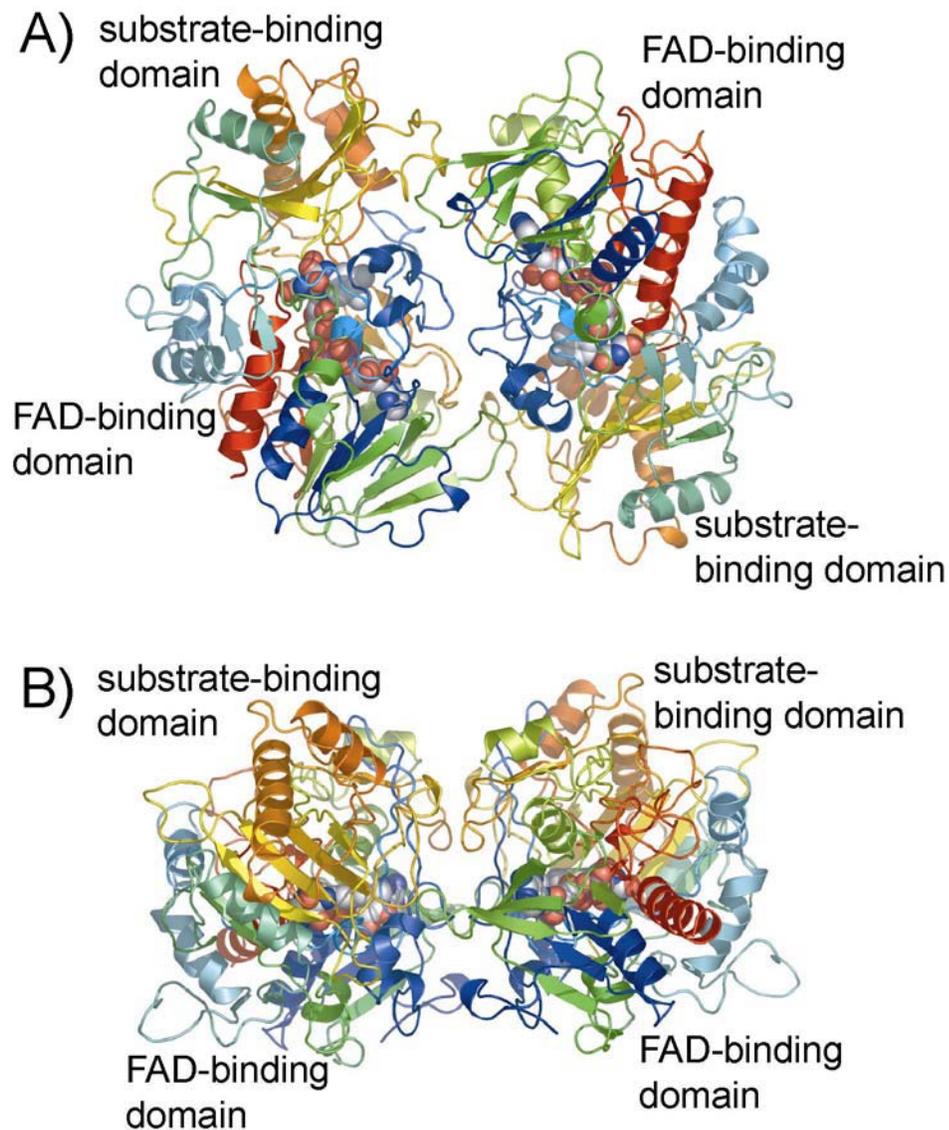


Figure 5.7. The three-dimensional structure of choline oxidase refined to 1.86 Å resolution. A) Choline oxidase is a homodimer with a molecular weight of 120 kDa. The enzyme consists of a FAD-binding and substrate-binding domain. The FAD is shown in CPK atoms with the C, N, and O atoms colored in gray, blue, and red, respectively. B) Another view of choline oxidase after a 90° rotation about the horizontal axis with respect to the orientation in A. The C_α ribbon trace is colored with a gradient from the N-terminal (blue) to the C-terminal (red) end.

The Flavin Binding Site

The high-resolution electron density maps obtained allow for a detailed view of the choline oxidase active site. The flavin molecule is very well buried within the protein and no parts of the molecule are visible from the molecular surface of the protein. Solvent-accessible surface area calculations indicate that 21.5 \AA^2 of a total surface area of 1024.5 \AA^2 is solvent exposed for the FAD. The electron density clearly indicates that the FAD is covalently linked to the His99^{N ϵ 2} atom via the FAD^{C8M} atom of the isoalloxazine ring in contrast to previous mass spectrometry studies by other authors that identified His87 as the residue covalently attached to the FAD^{C8M} (56). The most outstanding and surprising feature of the electron density however, is the observation that the flavin isoalloxazine ring is not planar as anticipated for oxidized flavin. However, the distortions also differ significantly from the structures of reduced flavins, which typically exhibit an approximately 150° angle along the N5-N10 axis defined by the dimethylbenzene and pyrimidine rings (7, 121, 122). An examination of the electron density indicates that although the coplanar dimethylbenzene and piperazine rings in choline oxidase are essentially flat, the plane of the pyrimidine ring is at an approximately 120° angle to the former plane. The pyrimidine ring is also significantly distorted from planarity and adopts a “half-boat” configuration. In this orientation, all atoms in the pyrimidine are in the same plane with the exception of C4a, which lies approximately 0.5 \AA above the pyrimidine ring plane. Moreover, during the refinement of both structures a greater than 4σ $mF_o - DF_c$ positive difference peak associated with the C4a atom became apparent. These features all suggest that the C4a atom is sp^3 hybridized; therefore, we concluded that a covalent adduct was present in the crystal.

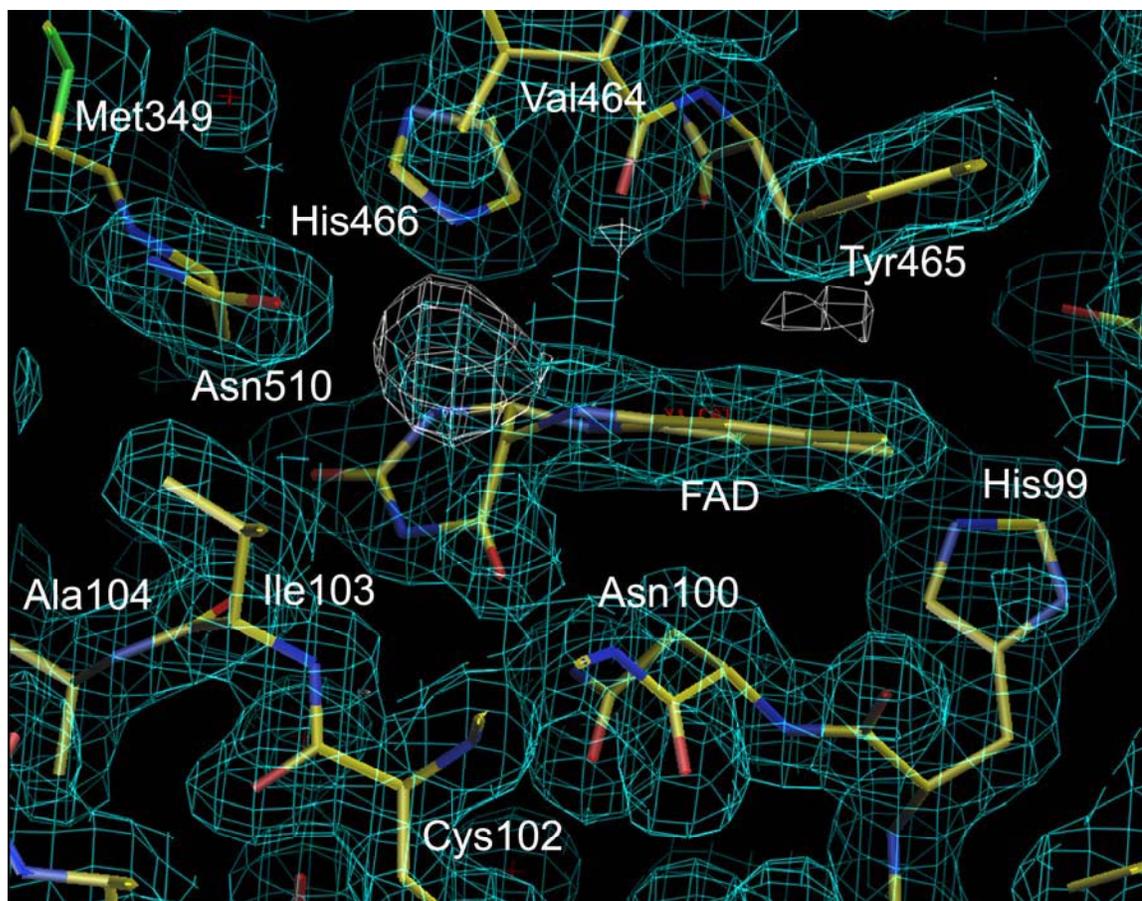


Figure 5.8. A view of the $2mF_o-DF_c$ (blue, contoured at 1σ , 1.86 \AA resolution) and mF_o-DF_c (white, contoured at 3.0σ , 1.86 \AA resolution) electron-density maps for the FAD cofactor and surrounding residues during the course of structure refinement. The C, N, O, and S atoms are colored in yellow, blue, red, and green, respectively. During the refinement process, it was observed that the FAD isoalloxazine ring adopts a distorted structure in which the pyrimidine ring is at an approximate angle of 120° to the coplanar dimethylbenzene and piperazine rings. The mF_o-DF_c electron density maps indicate that a covalent linkage to the C4a atom is present. The electron density maps also indicate that the FAD is covalently linked to the protein via the His99 residue.

The crystallization conditions and materials do not include any reagents known to form a C4a adduct with FAD. On the other hand, the aerobic crystals do contain O_2 , which is proposed to react with reduced flavins at the C4a position of the isoalloxazine ring (6). Therefore, we modeled an O_2 molecule bound to the C4a atom. This model

refined well with REFMAC5 and did not yield any significant difference features (Figure 5.9A). However, occupancy refinements and *B*-factor considerations suggest that the distal oxygen atom of the adduct may be partially disordered due to precession. As a further and independent check, we used simulated annealing OMIT maps as implemented in CNS to refine a model from which the FAD and the atoms associated with C4a were omitted. The resulting mF_o-DF_c and $2mF_o-DF_c$ electron density maps calculated with reflections between 50 and 1.86 Å resolution support the presence of an O₂ species bound to C4a (Figure 5.10). Therefore, we concluded that the FAD in choline oxidase is most likely a C4a adduct. The quality of the refined atomic model of choline oxidase and the fit to the observed electron density reported here is comparable to that of the recently reported oxygen complexes of cytochrome P450cam and naphthalene dioxygenase (123, 124). We also refined several additional models with REFMAC5 to check for additional possibilities. One alternative model included a water molecule centered in the difference peak and unrestrained with respect to the FAD and the C4a atom. Upon convergence of the refinement of this model, the water molecule was only 1.6 Å from the C4a atom and there was continuous electron density between the two atoms (Figure 5.9C). However, since there is no bond present between the C4a and the solvent, it does not satisfy the observed sp^3 hybridization of the flavin C4a atom. We also refined a model consisting of a covalent C4a-OH moiety, which refined to a 1.45 Å bond distance (Figure 5.9B). Although this model fits the electron density well, it is difficult to rationalize because during enzymatic turnover with choline or betaine aldehyde as substrate, a 1:1 stoichiometry of H₂O₂ produced per O₂ consumed is observed. This was established here by determining the effect of catalase on the rate of O₂ consumption with betaine aldehyde

or choline as substrate for choline oxidase at pH 7 and 35° C. In the presence of 5,000 units catalase, the rate of O₂ consumption decreased from 16.8 ± 0.2 s⁻¹ to 8.2 ± 0.5 s⁻¹ with 10 mM choline, consistent with a 1:1 stoichiometry of O₂ consumed per H₂O₂ formed. Similarly, catalase decreased the rate of O₂ consumption from 15.4 ± 0.7 s⁻¹ to 7.4 ± 0.5 s⁻¹ when 5 mM betaine aldehyde was used as a substrate.

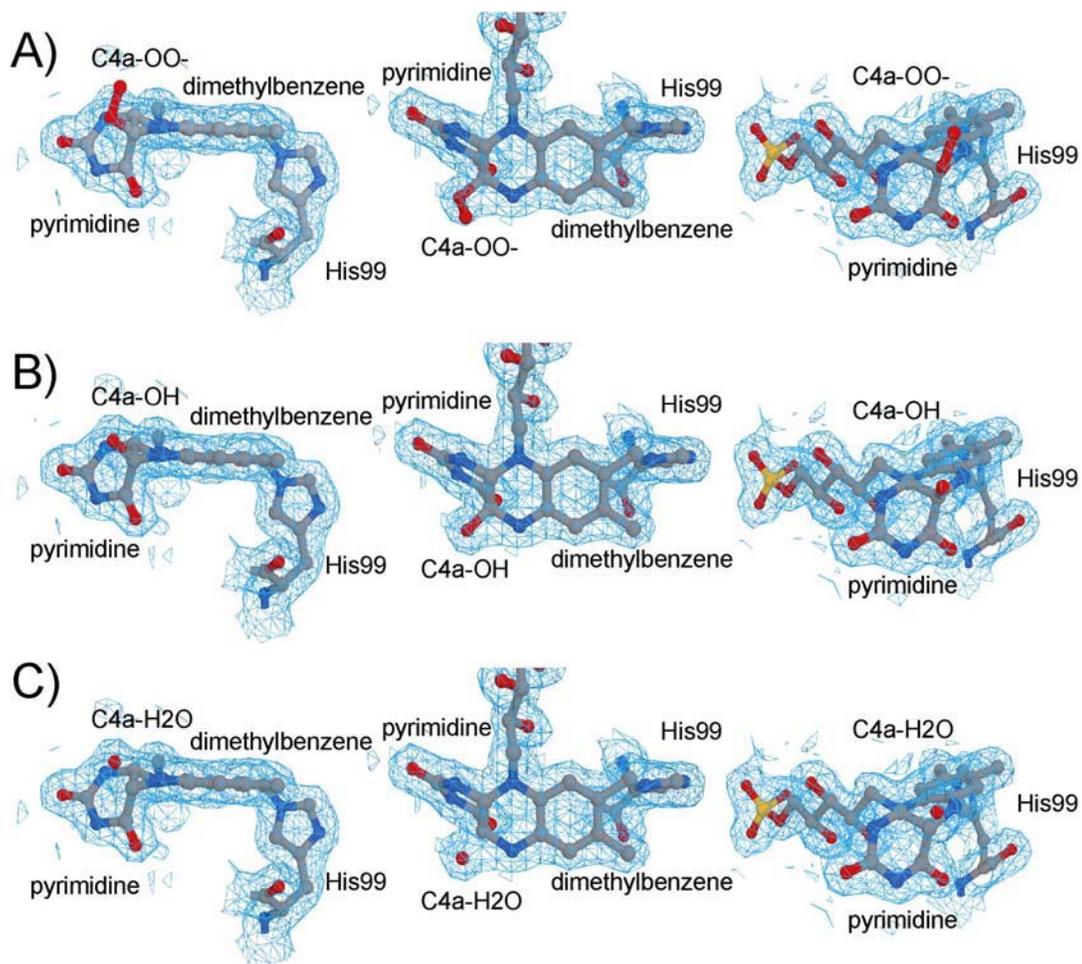


Figure 5.9. Approximately orthogonal views of the final $1.86 \text{ \AA } 2mF_o - DF_c$ electron density maps (contoured at 1σ) for the FAD isoalloxazine ring in choline oxidase and comparison to different models of the FAD-C4a-adduct. The C, N, O, and P atoms are colored in gray, blue, red, and yellow, respectively. In panel A) the FAD-C4a-O₂⁻ adduct is superimposed on the electron density maps B) A representation of the fit of the FAD-C4a-OH model to the density and C) the model with a water molecule placed in the residual density protruding from C4a.

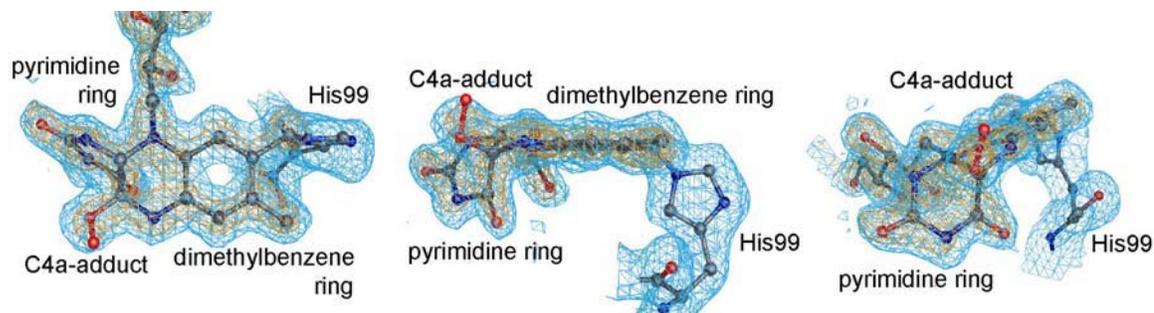


Figure 5.10. Approximately orthogonal views of the 1.86 Å electron density maps for the FAD isoalloxazine ring in choline oxidase. The mF_o-DF_c difference map (contoured at $+3.5 \sigma$) is shown in gold and the $2mF_o-DF_c$ map (contoured at 1σ) is shown in cyan. The maps are simulated-annealing OMIT maps calculated in CNS with the FAD atoms omitted during map calculations. For comparison, the refined atomic model is shown superimposed on the maps with C, N, and O atoms colored in gray, blue, and red, respectively.

In the 1.86 Å resolution structure of choline oxidase, the refined bond length between C4a and the proximal O atom (O_p) converged to a distance of 1.39 Å. The O-O bond distance (O_p-O_d) is approximately 1.40 Å and the C4a-O-O bond angle is 110.4°. Although the estimated coordinate error of the model is approximately 0.1 Å, these parameters agree well with those determined by quantum mechanical calculations for the model FAD-C4a-OOH intermediate in *p*-hydroxybenzoate 3-hydroxylase (25, 125). Thus, the refined model of choline oxidase has reasonable geometry for sp^3 hybridization and the appropriate bond lengths and angles for a C4a-oxygen complex (Table 5.3).

Table 5.3. Selected Bond Lengths and Angles in FAD Isoalloxazine Rings				
Enzyme	CHO ^a	CHO	CHO	<i>p</i> HBA-3H ^b
Model	FAD-C4a-H ₂ O	FAD-C4a-OH	FAD-C4a-OO ⁻	FAD-C4a-OOH
Bond Distance (Å)				
O _p -O _d ^c	N.A. ^d	N.A.	1.38	1.28 - 1.44
C4a-O _p	1.63	1.45	1.40	1.43 - 1.48
C10a-C4a	1.47	1.47	1.46	N.R. ^e
C4-C4a	1.55	1.59	1.55	N.R.
C4a-N5	1.51	1.47	1.48	N.R.
Bond Angle (°)				
C4a-O _p -O _d	N.A.	N.A.	113.7	108.9 - 113.3
C4-C4a-O _p	134.2	123.5	115.4	N.R.
C10a-C4a-O _p	104.1	107.4	108.6	102.2 - 103.2 (108.5 - 109.9)
N5-C4a-C4	123.5	120.8	122.2	N.R.
N1-C10a-C4a	108.2	109.8	108.4	N.R.
C4a				
N5-C4a-O _p	88.9	100.7	105.2	N.R.
C10a-N10-C4a	173.7	173.7	173.6	N.R.
N1				
C4-C4a-C6	125.2	125.2	124.3	N.R.

^a Choline Oxidase (CHO). ^b *p*-hydroxybenzoate 3-hydroxylase. ^cO_p is proximal oxygen and O_d is the distal oxygen. ^d Not applicable. ^e Not reported

Several important interactions are observed between the configuration of the isoalloxazine ring and the protein structure. As shown in Figure 5.11, the pyrimidine ring is involved in a network of hydrogen bonds to the backbone atoms of the protein. This includes hydrogen bonding or electrostatic interactions with the backbone atoms of Asn100, Cys102, Ile103, Pro511, and Asn512. Moreover, if the FAD were planar, then the edge of the pyrimidine ring would be adjacent to the side chains of Asn510 and Ile103. Indeed, the location of Ile103, a hydrophobic residue, adjacent to the hydrophilic pyrimidine ring is an unusual feature of choline oxidase. Nearly all of the backbone atoms between Ile103 and His99, the site of FAD covalent attachment, are involved in contacts with the isoalloxazine ring. Therefore, it appears that the FAD isoalloxazine ring is constrained by the covalent linkage at His99 and the proximity of Ile103 to the pyrimidine ring, which together define the borders of the cavity for the cofactor. Additionally, there are two α -helix dipoles that interact with the flavin in each monomer. The α -helix comprised of residues 23-34 points the N-terminal region directly at the pyrophosphate moiety of the FAD while the N-terminal of the α -helix comprising residues 512-527 points directly at the pyrimidine of the isoalloxazine ring.

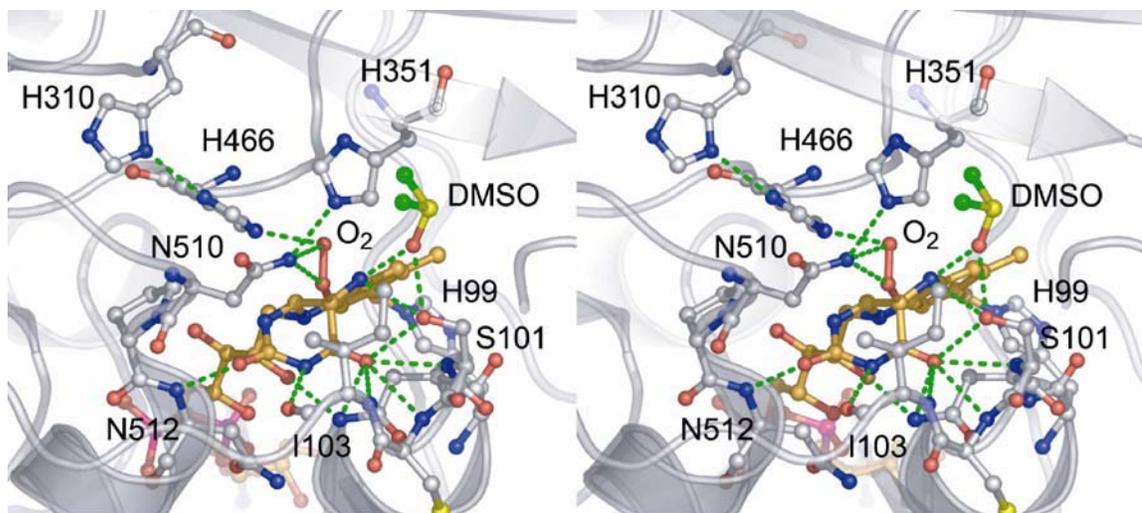


Figure 5.11. A divergent stereo view of the environment surrounding the FAD isoalloxazine ring in choline oxidase. The hydrogen bonding interactions are illustrated by green dashed lines. The bonds for the isoalloxazine ring and DMSO are shown in orange and green, respectively. The C, N, O, and S atoms of the amino acids are colored in gray, blue, red, and yellow, respectively.

Substrate Binding Site

A common feature of the structurally characterized members of the GMC family members is a loop that covers the active site and forms a lid over the putative substrate binding site (89, 126-128). In choline oxidase, this loop region is composed of residues 64-95. Examination of the identity and orientation of the residues that comprise the loop indicate that the loop is amphiphatic in nature. Hydrophobic residues are directed towards the interior of the protein and the more hydrophilic residues are located on the external surface of the molecule. The loop contains eight charged residues all of which protrude away from the interior of the surface of the protein. The charged residues are primarily negatively charged with the exception of three positively charged residues. While the bottom portion of the loop is shielded from bulk solvent by the second monomer of the dimer, residues 74-85 protrude into the bulk solvent. Analysis of the average *B*-factors of

the residues comprising the loop indicate an average B -factor of 20.0 \AA^2 for all residues which suggests that the loop is rigid in the crystal structure. The top portion of the loop that protrudes into the bulk solvent reveals two negatively charged residues, Glu76 and Glu79 whose side chains extend into the solvent.

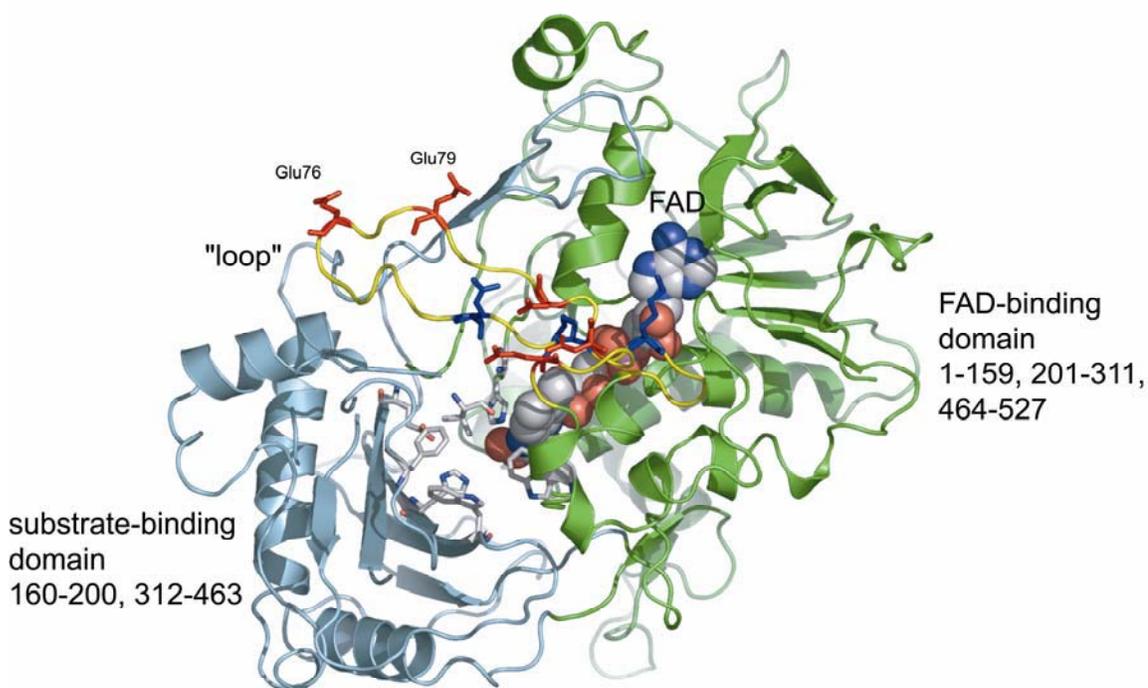


Figure 5.12. A view of the monomeric structure of choline oxidase highlighting the loop region (yellow) that is comprised of residues 64-95. Additionally, the FAD-binding (green) and the substrate-binding (cyan) domains are shown. The structurally defined members of the GMC family all share a loop region that covers the substrate-binding domain and is proposed to act as a lid covering the substrate binding site. In choline oxidase, the crystal structure reveals that this loop is rigid as defined by the average B -factor of 20 \AA^2 and is part of the FAD-binding domain. The loop is amphiphatic in nature with charged residues protruding into the solvent. The negatively charged residues are colored in red and positively charged ones in blue. A portion of the residues comprising the substrate-binding domain are illustrated in stick figures with the C, N, and O atoms colored in gray, blue, and red, respectively. The FAD atom is shown in CPK atoms.

Analysis of the molecular surface of the protein reveals a cavity of an approximate volume of 123 \AA^3 that is located in the substrate binding domain on the *re*-face of the flavin isoalloxazine ring and is sealed off from the exterior of the protein. The cavity is surrounded primarily by hydrophobic residues and a single charged residue that include Trp61, Glu312, Trp331, His351, Phe357, Val464, Tyr465, and His466. Indeed, an analysis of the location of the residues surrounding the cavity reveal that the orientation of the residues form a hydrophobic, aromatic cage around the cavity that is bordered by the flavin isoalloxazine ring on the bottom and is topped by a negatively charged Glu312 which points directly towards the cavity. There are no ordered solvent molecules within the cavity with the exception of a neighboring DMSO molecule that is located within close proximity to the N5 atom of the isoalloxazine ring (2.7 \AA from the DMSO oxygen to N5) and two water molecules that are bound to the Glu312 side chain. Two histidine residues, His466 and His351 border the bottom portion of the cavity and the side chains of both histidines point directly at the cavity. The His466^{N ϵ 2} atom is involved in hydrogen bonding interactions with the flavin isoalloxazine ring N1 (3.1 \AA) while His351^{N ϵ 2} is hydrogen bonded to Asn510^{N δ 2} (3.0 \AA). Another interesting feature of the substrate binding domain that is within close proximity to the cavity is a network of five histidine residues that are within close contact and form a path to the exterior surface of the protein. His466 is interestingly in close contacts with a neighboring His310 residue which in turn is within close proximity to a His381 and His506. His506 and His381 are located close to the solvent exposed molecular surface. The orientation and proximity of these residues to each other suggest that these residues may potentially form a proton relay system to the active site.

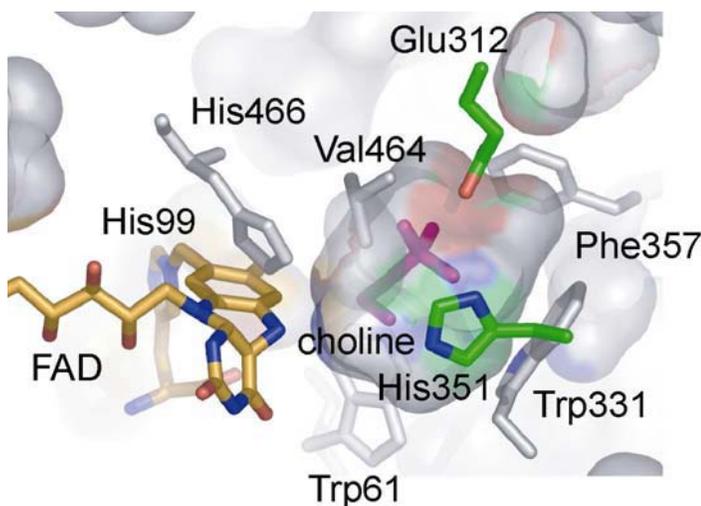


Figure 5.13. A representation of the well-defined cavity located in the substrate-binding domain of choline oxidase. The cavity is surrounded by hydrophobic residues that include Trp61, Trp331, Phe357, and Val464. The cavity is bordered by the FAD isoalloxazine ring on one end and is topped by a negatively charged Glu312 residue. The cavity volume of 123 \AA^3 is of sufficient size to accommodate a choline molecule (shown in magenta).

The electron density maps furthermore indicated a sphere of electron density with greater than 4.5σ positive difference features which was located approximately 4.0 \AA from the N5 atom of the isoalloxazine ring. The shape and size of the electron density clearly indicated that this species was not a water molecule. Thus, it was assumed that the density may correspond to a species from reagents in the crystallization solution. Indeed, the crystallization solution contains both DMSO and ammonium sulfate (which can contribute a sulfate anion). Such solvent molecules are often found to be bound within the solvent channels of proteins thus we examined the possibility that either species could be located at this position. Analysis of the fit of DMSO and sulfate ion into the electron density revealed that DMSO fits better into the density (Figure 5.14). The methyl groups of DMSO point towards Val464 (3.7 \AA) and His351 (3.5 \AA). The oxygen atom of DMSO

is within hydrogen bonding distance with Ser101^{O γ} (2.7 Å), which may contribute to the binding of the molecule. The average *B*-factor of 40.0 Å² for the refined DMSO molecule suggests that this molecule is not tightly bound. We further examined the possibility that DMSO might be an inhibitor of the enzyme as it is bound near the FAD in the substrate-binding domain. However, experiments determined that although DMSO showed competitive inhibition patterns, the inhibition constant was determined to be 462 ± 4 mM (personal communication, Fan, Fan & Gadda, Giovanni).

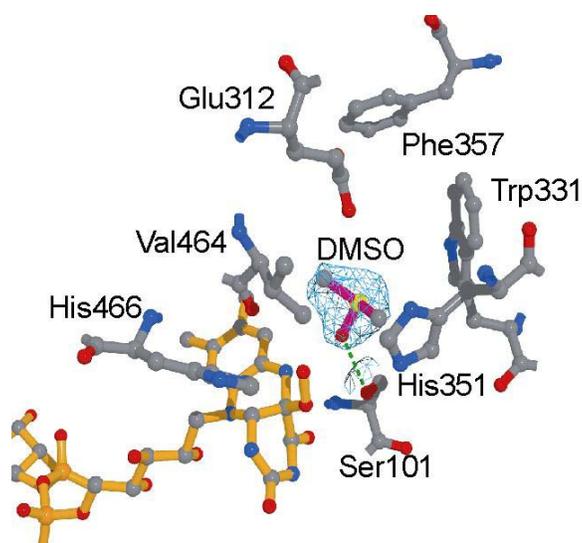


Figure 5.14. A view of the $2mF_o - DF_c$ electron density for the bound DMSO molecule and surrounding residues in the active site of choline oxidase. The oxygen atom of DMSO is within 2.7 Å of the N5 atom of the isoalloxazine ring of the FAD. The DMSO molecule binds at the bottom of the cavity located in the active site.

While it has been proposed that the loop region (residues 64-95) acts as a lid which regulates substrate entry into the substrate binding site of the GMC family members (126, 128), the structure of choline oxidase also reveals a channel which begins from the exterior of the protein and extends toward the defined cavity located in the

active site (Figure 5.15 –5.16). The connectivity of the channel to the cavity is broken by the sulfur of the Met62 side chain and the terminal portion of the channel is bordered by Phe357, Phe252, Pro356, and Trp330. The molecular surface of the protein shows an invagination that protrudes from the bulk solvent into the protein interior towards the active site cavity. Met62 blocks the channel from connecting with the cavity via its sulfur atom and the surface analysis shows a solvent exposed Glu7 residue from the second subunit in the dimer that borders the “entrance” of the channel. The negative charge of this residue and its close location to the entrance of the channel suggests that the negative charge of Glu7 may attract the positive charge of the trimethylammonium group if choline enters the active site through this channel. If choline can enter the active site through this channel, it would require a reorientation of the Met62 side chain to allow for access to the interior cavity.

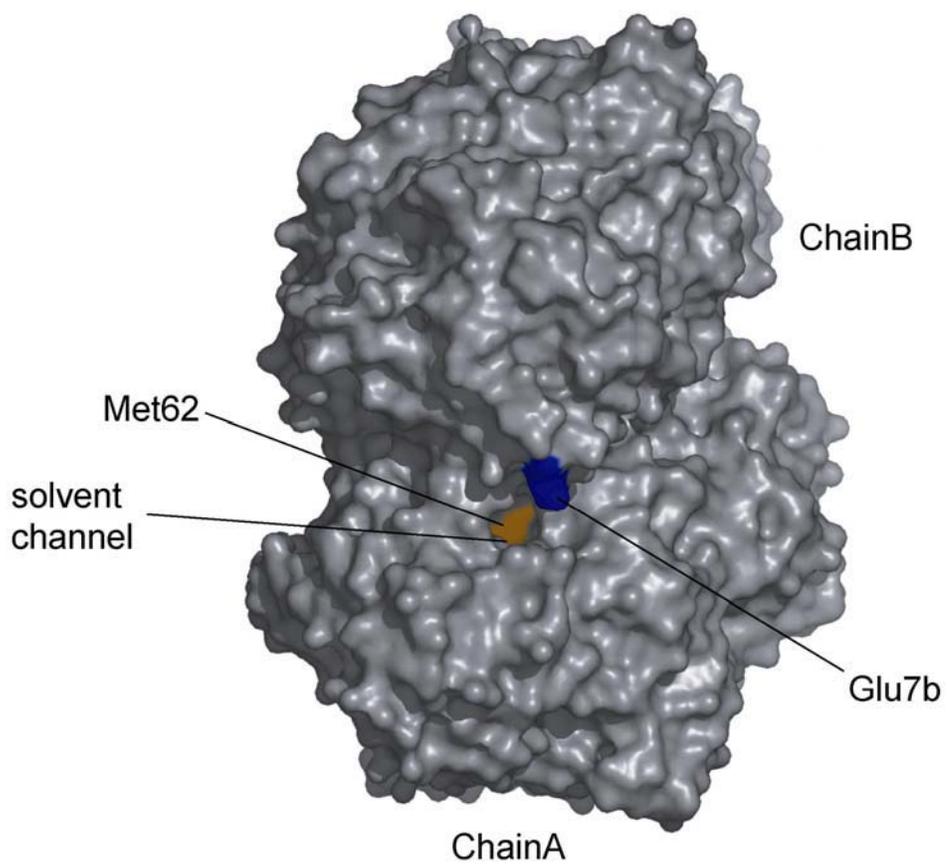


Figure 5.15 A view of the molecular surface of choline oxidase dimer (gray) with the solvent channel that protrudes into the active site highlighted in orange.

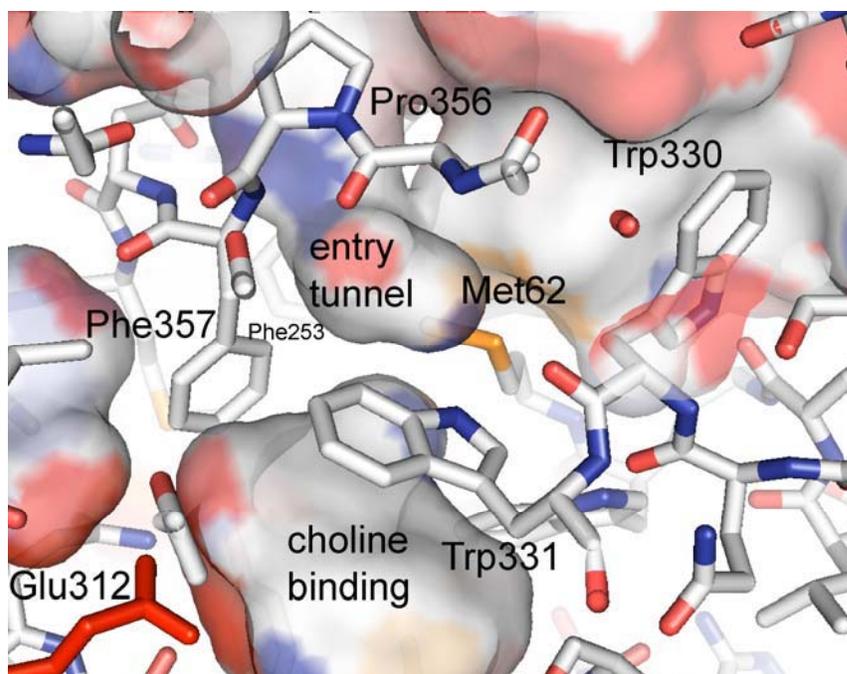


Figure 5.16. A representation of the solvent channel that protrudes toward the cavity in the substrate-binding domain. The channel is blocked from connecting to the cavity by the side chain of Met62. The negatively charged Glu312 residue is colored in red appears to play an important structural role by providing an ionic interaction with the trimethylammonium headgroup of the choline molecule. The C, N, O, and S atoms are colored in gray, blue, red, and orange, respectively.

Discussion

Identification of Putative Substrate Binding Site

The determination of the crystal structure of choline oxidase provides much needed structural information to accompany the many ongoing mechanistic and biochemical studies of the enzyme (55, 59, 77, 103, 104). A major question that needs to be answered to better understand the chemical mechanism of choline oxidase is the identity of amino acid residues involved in substrate binding and catalysis. The high-resolution electron density maps have allowed us to identify the putative binding site for the choline substrate. Previous kinetic studies performed on choline oxidase with choline

analogs indicate that the positively charged trimethylammonium moiety is critically important for the binding of ligands at the active site of choline oxidase and is a major determinant that defines the specificity of the enzyme for the substrate (104). Additionally, kinetic data have indicated that the positively charged amino headgroup of choline plays an important role in the oxidation of the enzyme-bound reduced flavin during catalysis of choline oxidase. Inhibition data have also shown that the binding affinity of ligands is not critically affected by the nature of the side chain attached to the trimethylammonium moiety of the ligand (104). The analysis of the molecular surface of the crystal structure of choline oxidase reveals a cavity on the *re*-face of the flavin isoalloxazine ring with an approximate volume of 123 Å³. The cavity is enclosed mainly by hydrophobic residues that form a hydrophobic cage. Additionally, a negatively charged Glu312 residue protrudes toward the cavity. The volume of the cavity is of sufficient size to accommodate a choline molecule (93 Å³). In order to gain insights between the enzyme and substrate interactions, we took advantage of the location and dimensions of the cavity and manually docked choline into the cavity. Thus, the identification of the cavity and the modeled choline provide structural insights into the putative binding mode of choline substrate and the residues that may be responsible for binding and involved in catalysis.

The docked choline molecule is oriented in a manner in which the positively charged trimethylammonium headgroup is positioned approximately 3.1 Å from the negatively charged side chain residue of Glu312. Structural data based on residues that surround the cavity indicate that the residues that most likely are involved in binding of the choline substrate are Trp61, Trp331, His351, Phe357, Val464, and Glu312. The

active site cavity suggests that the positively charged headgroup of choline points towards the negatively charged Glu312 residue. Glu312 is thus hypothesized to play an important structural role by forming an ionic pair with the positively charged trimethylammonium moiety and thereby properly anchoring the choline molecule in position for efficient catalysis. Recent unpublished data show that the Glu312Asp mutant exhibits lower activity than the wild type enzyme, but the Glu312Ala mutant is completely devoid of activity even when the concentration of choline is 20 times larger than that commonly used to saturate the wild-type enzyme. (personal communication, Quaye, Osbourne & Gadda, Giovanni). Additionally, the trimethylammonium moiety is surrounded by three hydrophobic residues with aromatic side chains (Trp61, Trp331, and Phe357). A hydrophobic Val464 also borders the trimethylammonium moiety. The aromatic residues are hypothesized to further facilitate substrate binding through cation- π interactions between the aromatic side chains and the positive charge of choline. Thus, these interactions are potentially critical in properly positioning choline for catalysis. X-ray crystallographic and site-directed mutagenesis studies on phospholipase C have demonstrated the importance of three residues for substrate specificity in the enzyme's choline binding site (129). These studies demonstrated that Glu4, Tyr56, and Phe66 were important residues in providing a binding site for an inhibitor similar to phosphatidylcholine through electrostatic and cation- π interactions. Additional structural studies on the choline binding site in acetylcholinesterase (130, 131) and CTP:phosphocholine cytidylyltransferase (132) have also demonstrated the importance of cation- π interactions in stabilization of the trimethylammonium moiety.

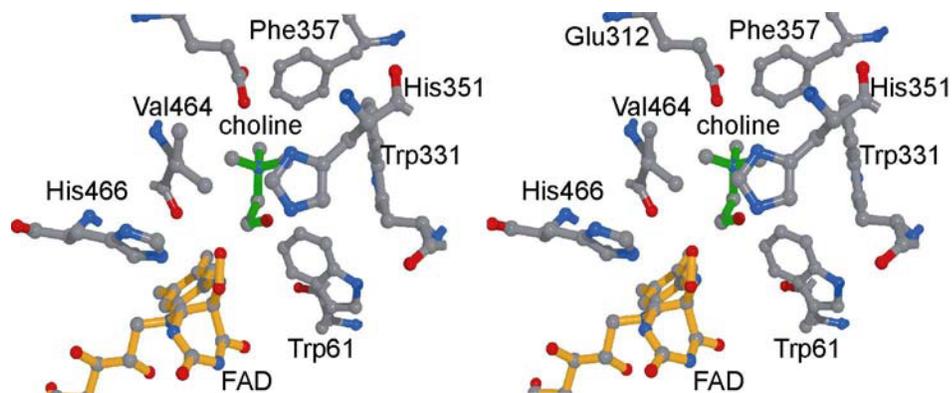


Figure 5.17. A divergent stereoview of the active site of choline oxidase with manually docked choline. The choline molecule was docked into the active site cavity in order to obtain structural insights into the putative residues involved in substrate binding and catalysis. The bonds of the choline molecule are colored in green and the bonds of the FAD molecule are colored in orange. The C, N, and O atoms are colored in gray, red, and blue, respectively.

We further used the location of the bound DMSO molecule to model the position the hydroxyl moiety of the choline molecule in the cavity. Two histidine residues, His466 and His351 are within close proximity to the hydroxyl moiety of the modeled choline substrate. In our model, His351 is in closest proximity to the hydroxyl group (approximately 3.3 Å). We positioned this portion of the molecule towards His351 based on biochemical data from studies on a mutant isoform of choline oxidase in which His466, which is conserved among the GMC family members, was changed into an alanine (59). The conserved residue has been proposed to be the catalytic base in the GMC family. This notion is based on earlier mechanistic studies on cholesterol oxidase, cellobiose dehydrogenase, and glucose oxidase that suggested that a histidine residue

which is fully conserved among the GMC family might act as the specific base that participates in the oxidation of the alcohol substrate. However, the determination of the X-ray crystal structure of unliganded cholesterol oxidase at sub-atomic resolution revealed that the N ϵ 2 position of His447 in cholesterol oxidase is protonated (88, 133). Thus, this finding raises questions on the catalytic role of the conserved histidine residue in the GMC family. The crystal structure of choline oxidase reveals that His351 and His466 are the only residues that are suitably located in the active site of the enzyme with the ability to act as a catalytic base in the oxidation of choline. Biochemical and mechanistic data on a choline oxidase His466Ala mutant suggest that His466 does not participate in the oxidative-half reaction and is likely not the base that abstracts the hydroxyl proton of choline. The enzymatic activity of this mutant is furthermore partially rescued in the presence of exogenous imidazolium but not imidazole (59). Based on the biochemical data, we hypothesize that, in choline oxidase, His351 is most likely the residue that acts as the general base that initiates the oxidation of choline and the general acid that participates in the reduction of O₂. In our model of docked choline, His351 is located approximately 3.3 Å from the choline hydroxyl group. Thus the new structural data that reveal the location of His351 is in close proximity to the choline binding site and the combination of biochemical data from mechanistic studies on the His466Ala mutant provide strong evidence that suggest His351 may be the catalytic base in choline oxidase. The crystallographic data have thus prompted our collaborators to further study the role of His351 in the reaction mechanism which is currently in progress. Additionally, in lieu of the new structural and biochemical data, it is proposed that His466 may play an important role in stabilization of the alkoxide intermediate during the oxidation of choline

and also influence the polarity of the active site. Indeed, studies on the His466Ala mutant have demonstrated that a significant contribution to the electrophilicity of the FAD cofactor is contributed by His466 (59). In the crystal structure of choline oxidase, His466 is involved in a direct interaction with the N1 position of the flavin cofactor via the Nε2 locus and thus may influence the flavin microenvironment via this interaction. The polarity of the active site is essential for efficient proton transfer of the hydroxyl proton from the alcohol substrate to the active site proton acceptor. Data from studies on the His466Ala mutant also suggest that the His466 residue is important for regulating the reactivity of the catalytic base that accepts the hydroxyl proton of the substrate for efficient catalysis (59). Furthermore, the combination of structural and biochemical data available now with respect to identification of the putative choline binding site should provide the groundwork for investigations aimed at the development of specific inhibitors targeted at choline oxidase and glycine betaine synthesis. The development of therapeutic agents that inhibit glycine betaine biosynthesis could potentially render a number of human pathogens more susceptible to conventional treatments.

The FAD in Choline Oxidase

The highly unusual structure of the flavin isoalloxazine ring in choline oxidase raises several questions. The most outstanding question regards the chemical process responsible for the distortion of the flavin isoalloxazine ring and the stability of the structure to allow detection by X-ray crystallography. The electron density maps clearly indicate a significant distortion of the pyrimidine ring from planarity and additionally identify the His99 residue as the residue that covalently links the FAD to the protein. The distortion from planarity is contrary to the flat, planar geometry expected for the

isoalloxazine ring in oxidized FAD but it also significantly differs from the structures of reduced flavins which adopt a bend of approximately 30° along the N5-N10 axis. The presence of a C4a-adduct which we have modeled as a dioxygen atom is also significant in that to our knowledge, this is only the third C4a-adduct crystal structure described for any flavoprotein. For example, *trans*-2-phenylcyclopropylamine forms a FAD-C4a adduct in human monoamine oxidase B (MAO-B), an outer mitochondrial membrane-bound enzyme (39). The other complex was observed as a Cys-S-C4a FMN adduct in the LOV1 domains of the blue-light photoreceptors responsible for phototropism in seed plants. Although the flavin with MAO-B exhibits distortions in the isoalloxazine ring, the distortions observed in the choline oxidase structure are much more significant. Reduced flavin-adducts with covalent ligands attached at the N5 position also affect the distortion of the isoalloxazine ring in MAO-B (39).

We propose that the crystals of choline oxidase formed the FAD-C4a-O₂⁻ *in vitro* in a two-step process (Figure 5.18). The cryogenic data collection conditions and high brilliance synchrotron x-ray beam are known to reduce oxidized flavoproteins and metalloproteins (134, 135). Since the FAD in choline oxidase has a very high redox potential (~130 mV), the thermodynamic driving force is for reduction of the FAD by either one or two electrons (95). The one-electron reduced flavin semiquinone state of choline oxidase can be formed aerobically and is fully stable over prolonged times in the presence of O₂ even after removal of reducing power (58, 95). Consequently, it is likely that the enzyme-bound FAD is fully reduced by two electrons in the X-ray beam. After reduction of the flavin, we propose that O₂ from within the aerobic crystal matrix diffuses to the C4a position of the reduced FAD, in a process similar to that documented at low

temperature crystals for myoglobin-CO photodissociated complexes (136-138). Indeed, geminate recombination and even ligand exchange of diatomic ligands have been demonstrated by spectroscopic methods at temperatures as low as 4 K (139-144). Once the choline oxidase FAD-C4a-O₂⁻ complex is formed, it does not release H₂O₂ because the cryogenic X-ray data collection methods do not establish the appropriate proton inventory on the surrounding residues.

We have also compared the refined coordinates of choline oxidase with Density Functional Theory (B3LYP/6-31G) calculations for a further analysis of the electronic structure of the FAD observed in the crystal structure. Dr. Rajeev Prabhakar of Prof. Keiji Morokumo's group at Emory University is currently performing the calculations. A variety of models are currently being studied by the calculations. As shown in Figure 5.9, the two possibilities for the C4a-adduct indicate a choice between the FAD-C4a-O₂(H) and the FAD-C4a-O(H) models. Calculations were performed with the FAD-C4a-OH model in the gas-phase. Upon optimization, the FAD-C4a-O(H) model with a two electron-reduced flavin refined into a planar structure and indicate an energetic cost of approximately 73 kcal/mol to achieve a bent structure in the flavin (Figure 5.19). Similarly, the FAD-C4a-O₂(H) model with a two electron reduced flavin optimized into a planar structure (-82 kcal/mol) (Figure 5.20). These results clearly indicate the preference of the flavin to adopt a planar conformation in the gas-phase with a very large energetic cost to achieve the bent conformation observed in the crystal structure. However, DFT calculations performed when the protein residues around the flavin are included indicate very similar structures between what is observed in the crystal structure and the structure obtained by DFT calculations. A comparison of the bond angles

between experimental and theoretically calculated values show strikingly similar results. For example, the experimental values for the N5-C10a-N1 bond angle is 115.8° and 108.8° for the N5-C4a-C4 bond angle. In comparison, the calculated values for the same bonds are 121.2° and 113.6°, respectively (Figure 5.21). Preliminary calculations with the FAD C4a-O₂(H) adduct with surrounding protein residues indicate that after DFT calculations, the O-O bond is cleaved resulting in the formation of a C4a-O double bond that results in cleavage of the flavin ring. Thus, DFT calculations currently suggest that the FAD C4a-adduct favors OH as the covalently bound ligand. However, for this step to occur, it would require the presence of a nucleophile to attack the dioxygen adduct in order to form the FAD C4a-OH adduct. The identity of such a nucleophile in the choline oxidase crystal is not known, but there have been reports in the literature in which dimethylsulfide can be hydroxylated to form dimethylsulfoxide (145). Thus, if dimethylsulfide was indeed present, it could have been hydroxylated which would explain the presence of the bound DMSO molecule. However, the DMSO used in the crystallization conditions is of 99.9% purity. Additional calculations with the FAD-C4a-O₂(H) adduct are further being conducted at this moment to further analyze the adduct structure.

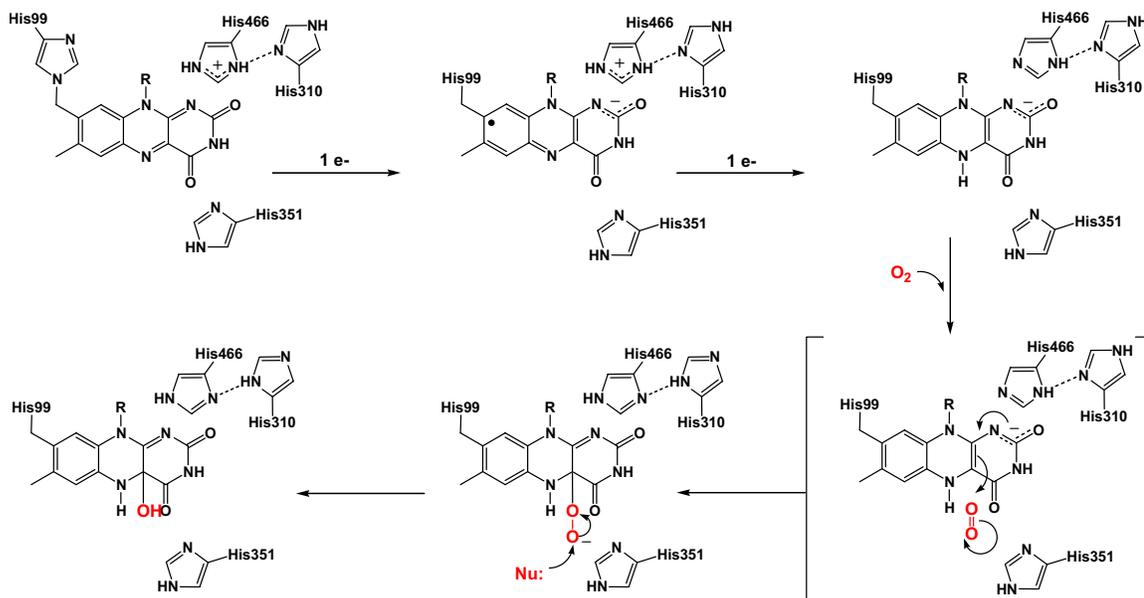


Figure 5.18. The proposed reaction mechanism involved in the generation of the FAD-C4a-adduct in the crystal of choline oxidase initiated by the synchrotron X-rays.

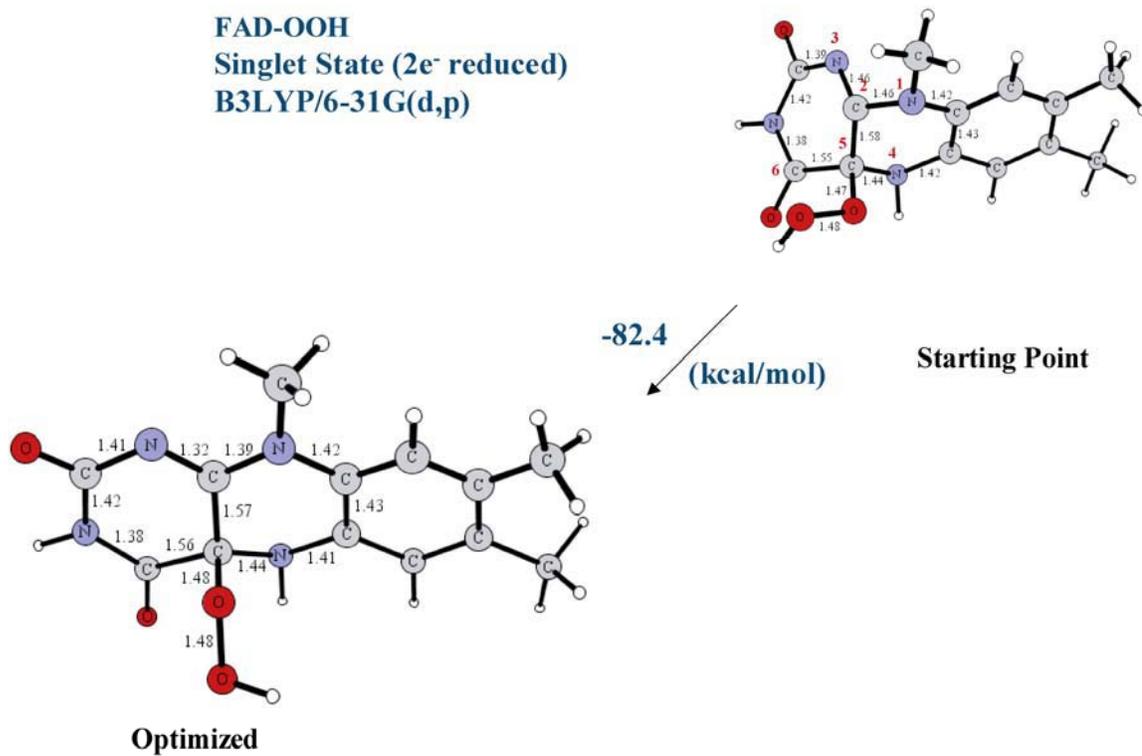


Figure 5.19. Gas-phase DFT calculations (B2LYP/6-31G) using the FAD C4a-O₂⁻ model with a two-electron reduced flavin result in a planar flavin molecule after optimization.

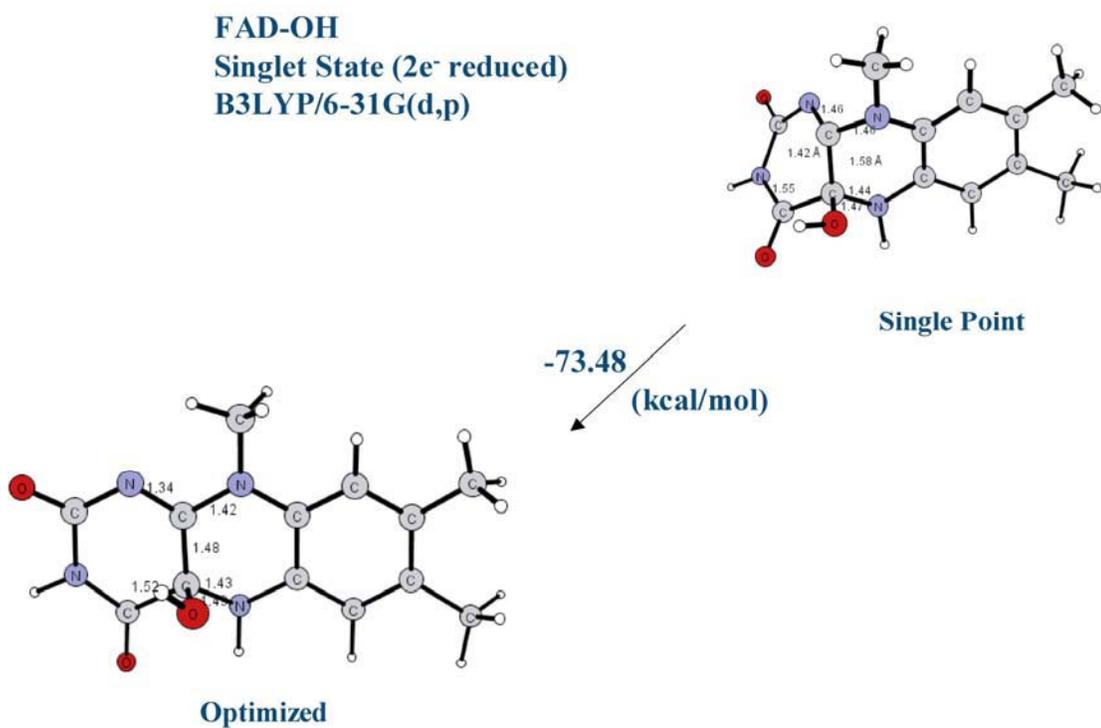


Figure 5.20. Gas-phase DFT calculations (B2LYP/6-31G) using the FAD C4a-OH model with a two-electron reduced flavin result in a planar FAD C4a-OH model after optimization.

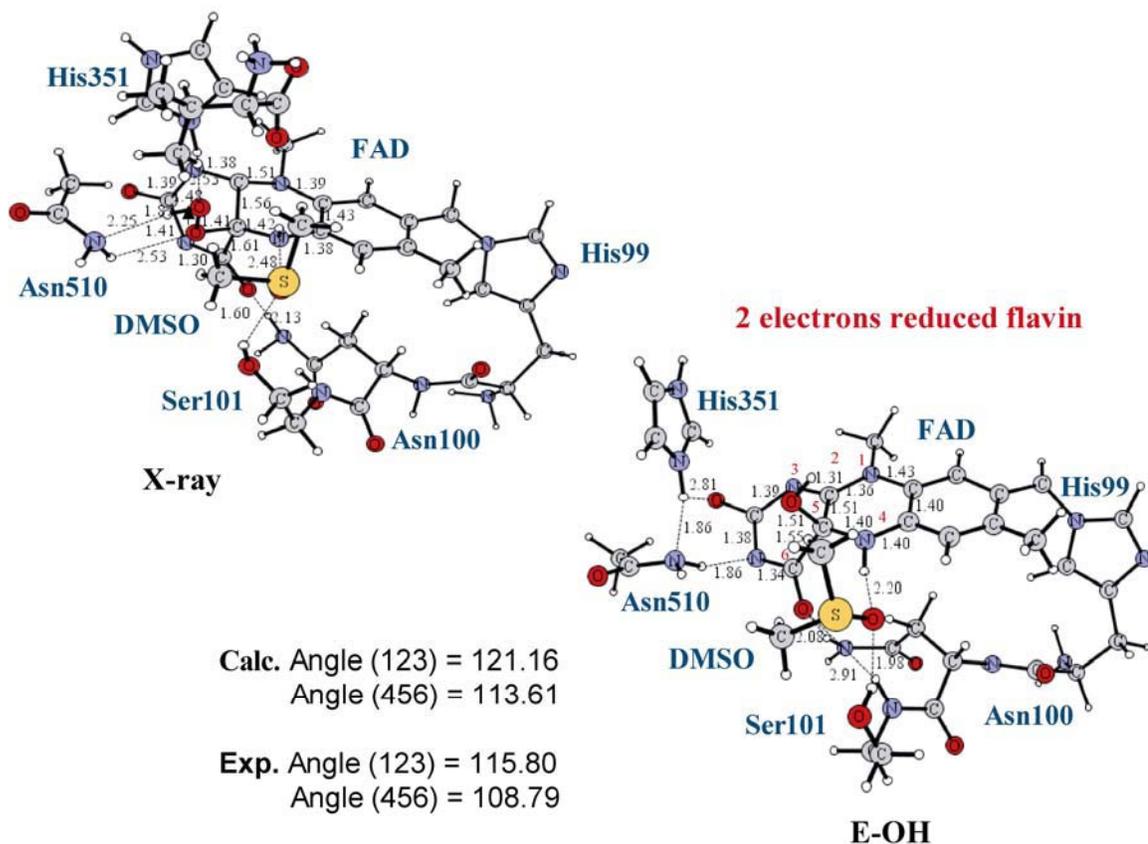


Figure 5.21. A comparison of the enzyme-bound FAD C4a-O₂⁻ adduct from the X-ray structure (left) and the enzyme-bound FAD C4a-OH model (right) obtained after DFT calculations with the protein residues included in the calculations indicate similar bond angles for the pyrimidine ring between experimental and calculated values. Gas-phase DFT calculations indicate that the FAD-adduct optimizes with a planar flavin but when DFT calculations are performed with the protein environment similar geometry to the X-ray structure is obtained.

Correlation of the Structure to the Proposed Reaction Mechanism

The combination of the available crystal structure of choline oxidase, mechanistic studies, and site-directed mutagenesis has now allowed for a more detailed examination of the catalytic mechanism of choline oxidase. As described earlier in the introduction section, the three leading proposals for the activation of unpolarized alcohols by the GMC family of enzymes involve a concerted hydride transfer mechanism, asynchronous hydride transfer mechanism, and an oxygen radical mechanism. Results from mechanistic

investigations on choline oxidase with deuterated substrate and solvent by Fan and Gadda (77) indicate that the chemical steps are fully rate limiting for the overall turnover of choline oxidase. The oxidation of choline by choline oxidase has been shown to occur through the formation of an alkoxide species resulting from the removal of the substrate hydroxyl proton occurring before the hydride transfer to the flavin. Evidence for such a mechanism, in which the choline alkoxide proton is not in flight in the transition state for CH bond cleavage, comes from the substrate and deuterium kinetic isotope effects determined on the rate of reduction of the flavin by choline. The data show a substrate isotope effect of ~ 9 and a solvent isotope effect of unity which rules out a concerted hydride transfer mechanism in which both the proton and the hydride are concomitantly in flight in the transition state (77). The alternative mechanism, in which a single electron is transferred to the flavin concomitantly with the abstraction of the hydroxyl proton before hydrogen transfer to the flavin, has also been determined to be unlikely. Such evidence is supported by observations that such a mechanism would result in the accumulation of flavin radical species during turnover. However, it has been demonstrated that the enzyme-bound flavin is reduced by choline directly to the hydroquinone state without formation of a detectable flavin semiquinone species (58, 95). Furthermore, previous studies have shown that the semiquinone in choline oxidase is catalytically inert, thereby providing indirect evidence against a radical mechanism (95).

Thus, by taking into account data that suggest that substrate oxidation by choline oxidase is via a base-initiated hydride transfer of the substrate to the N5 position of the flavin isoalloxazine ring (77) and the determination that the pK_a value of the catalytic base is 7.5 (95), an emerging chemical mechanism for choline oxidase can be described

taking into account both mechanistic and structural data. The X-ray crystal structure suggests that His351 is the catalytic base rather than His466, which is in agreement with the biochemical and mechanistic study of the His466Ala mutant enzyme. The pH-dependent studies on His466Ala demonstrated the presence of a catalytic base with a shifted pK_a of 9.0, suggesting that His466 affects the polarity of the active site environment and the flavin micro-environment rather than acting as the catalytic base (59). The structural data now available provide further evidence that that His466 residue plays an important role in the stabilization of the alkoxide intermediate during catalysis. It is positioned appropriately to play a stabilizing role of the alkoxide intermediate as indicated by the modeled choline. Furthermore, Asn512 may also play an important role in influencing the redox potential of the FAD and the active site environment as it forms the N-terminal portion of the α -helix dipole that points at the pyrimidine of the isoalloxazine ring. Indeed, it has been demonstrated in cholesterol oxidase that an Asn residue is important in creating an electrostatic environment around the isoalloxazine ring that favors reduction of FAD by unactivated alcohols (146).

The combination of structural and mechanistic studies suggest a chemical mechanism for choline oxidase in which an asynchronous hydride transfer mechanism is most likely for choline oxidase. In the reaction mechanism, we propose that His351 is most likely the active site base that abstracts the hydroxyl proton from the alcohol substrate and to form an alkoxide intermediate which is stabilized by electrostatic interactions with the imidazolium of His466. This step is then followed by a hydride transfer from the α -carbon of the activated alcohol to the N5 atom in the flavin isoalloxazine ring resulting in the formation of reduced flavin and betaine aldehyde.

Indeed, in our structure with modeled choline in the active site, the distance between the α -carbon in choline and the FAD N5 is approximately 4 Å. A combination of electrostatic interactions between the choline headgroup and Glu312 along with cation- π interactions are proposed to be key interactions that properly position the α -carbon of choline to allow for efficient hydride transfer to the flavin N5 position. In the subsequent oxidative half-reaction, two electrons are transferred from the reduced flavin to O₂. The final delivery of two protons from the catalytic base, His351, and the FAD N5 atom yields oxidized FAD and H₂O₂, which readily dissociates from the enzyme active site. Enzymatic turnover is then completed with a second oxidation reaction in which the enzyme bound aldehyde is oxidized to glycine betaine, although this reaction has not been mechanistically characterized yet.

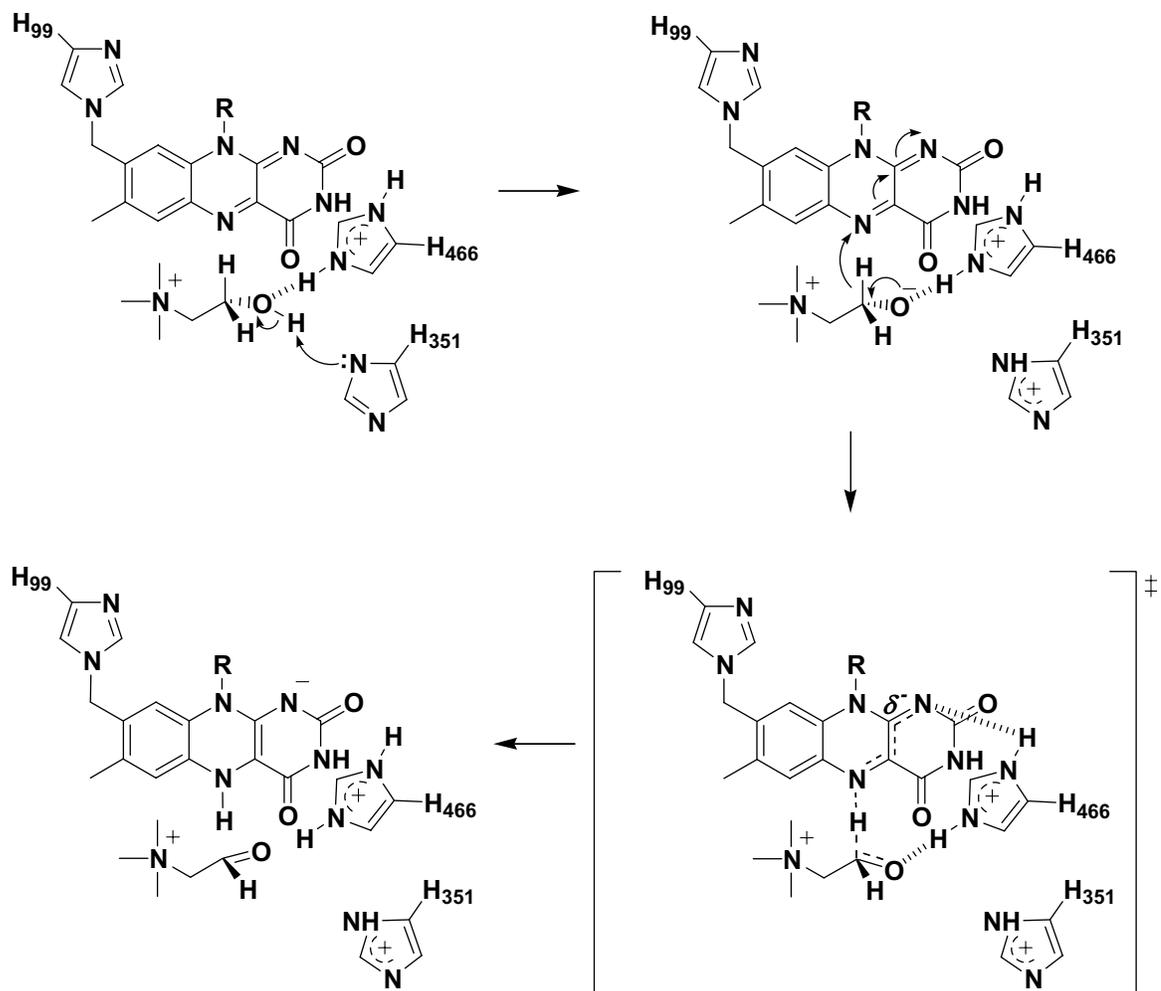


Figure 5.19 . A proposed structural role for His351 in the asynchronous hydride transfer mechanism. His351 is proposed to act as the active site base with a pK_a of 7.5 by abstracting the hydroxyl proton from choline. His466 is proposed to play an important structural role by influencing the electrophilicity of the flavin cofactor and stabilizing the alkoxide intermediate.

Conclusion

The determination of the crystal structure of choline oxidase at 1.86 Å resolution has provided structural information to aid in the study of the catalytic mechanism of the enzyme. While ongoing studies of the chemical mechanism of choline oxidase have proved fruitful in gleaned information regarding the catalytic mechanism, the recently determined crystal structure has allowed for structural insights to complement the many mechanistic studies performed on the enzyme. The structure determination has allowed for the identification of the putative substrate-binding site and the residues that may be directly involved in catalysis. The active site architecture has also revealed that His351 is most likely the active site base that initiates the removal of the hydroxyl proton from the unpolarized alcohol substrate. His466, which is conserved among the GMC family members, appears to play an important role in influencing the polarity of the flavin microenvironment and stabilizing the alkoxide intermediate during catalysis. The x-ray crystal structure indicates that a covalent adduct is present on the C4a position of the flavin which according to DFT calculations is most likely favored to be a FAD C4a-OH adduct. Further theoretical calculations are underway to study the flavin structure observed in the crystal structure. The determination of the structure of choline oxidase thereby provides many significant insights that will guide the further characterization of the overall catalytic mechanism of choline oxidase.

REFERENCES

1. Warburg, O., and Christian, W. (1933) *Biochim. Z.* 266, 377-411.
2. Kuhn, R., Reinemund, K., and Weygard, F. (1934) *Ber.* 67, 1460-1463.
3. Karrer, P., Schopp, K., and Benz, F. (1935) *Helv. Chim. Acta* 18, 426-429.
4. Massey, V. (2000) *Biochem Soc Trans* 28, 283-96.
5. Ghisla, S., and Massey, V. (1989) *Eur J Biochem* 181, 1-17.
6. Massey, V. (1994) *J Biol Chem* 269, 22459-62.
7. Fraaije, M. W., and Mattevi, A. (2000) *Trends Biochem Sci* 25, 126-32.
8. Arakaki, A. K., Ceccarelli, E. A., and Carrillo, N. (1997) *Faseb J* 11, 133-40.
9. Massey, V. (1995) *Faseb J* 9, 473-5.
10. Fisher, A. J., Thompson, T. B., Thoden, J. B., Baldwin, T. O., and Rayment, I. (1996) *J Biol Chem* 271, 21956-68.
11. Francisco, W. A., Abu-Soud, H. M., DelMonte, A. J., Singleton, D. A., Baldwin, T. O., and Raushel, F. M. (1998) *Biochemistry* 37, 2596-606.
12. Macheroux, P., Ghisla, S., and Hastings, J. W. (1993) *Biochemistry* 32, 14183-6.
13. Todo, T., Kim, S. T., Hitomi, K., Otoshi, E., Inui, T., Morioka, H., Kobayashi, H., Ohtsuka, E., Toh, H., and Ikenaga, M. (1997) *Nucleic Acids Res* 25, 764-8.
14. Hitomi, K., Okamoto, K., Daiyasu, H., Miyashita, H., Iwai, S., Toh, H., Ishiura, M., and Todo, T. (2000) *Nucleic Acids Res* 28, 2353-62.
15. Briggs, W. R., and Huala, E. (1999) *Annu Rev Cell Dev Biol* 15, 33-62.
16. Vervoort, J., and Rietjens, I. M. (1996) *Biochem Soc Trans* 24, 127-30.
17. Roth, J. P., and Klinman, J. P. (2003) *Proc Natl Acad Sci U S A* 100, 62-7.
18. Venci, D., Zhao, G., and Jorns, M. S. (2002) *Biochemistry* 41, 15795-802.
19. Palfey, B. A., and Massey, V. (1998) in *Comprehensive Biological Catalysis: A Mechanistic Reference* (Sinnott, M., Ed.) pp 83-154, Academic Press, San Diego.

20. Schreuder, H. A., Prick, P. A., Wierenga, R. K., Vriend, G., Wilson, K. S., Hol, W. G., and Drenth, J. (1989) *J Mol Biol* 208, 679-96.
21. Schreuder, H. A., Mattevi, A., Obmolova, G., Kalk, K. H., Hol, W. G., van der Bolt, F. J., and van Berkel, W. J. (1994) *Biochemistry* 33, 10161-70.
22. Gatti, D. L., Palfey, B. A., Lah, M. S., Entsch, B., Massey, V., Ballou, D. P., and Ludwig, M. L. (1994) *Science* 266, 110-4.
23. Schreuder, H. A., van der Laan, J. M., Swarte, M. B., Kalk, K. H., Hol, W. G., and Drenth, J. (1992) *Proteins* 14, 178-90.
24. Schreuder, H. A., van der Laan, J. M., Hol, W. G., and Drenth, J. (1988) *J Mol Biol* 199, 637-48.
25. Wierenga, R. K., de Jong, R. J., Kalk, K. H., Hol, W. G., and Drenth, J. (1979) *J Mol Biol* 131, 55-73.
26. Ortiz-Maldonado, M., Cole, L. J., Dumas, S. M., Entsch, B., and Ballou, D. P. (2004) *Biochemistry* 43, 1569-79.
27. Moran, G. R., Entsch, B., Palfey, B. A., and Ballou, D. P. (1997) *Biochemistry* 36, 7548-56.
28. Wu, J., Bell, A. F., Luo, L., Stephens, A. W., Stankovich, M. T., and Tonge, P. J. (2003) *Biochemistry* 42, 11846-56.
29. Stankovich, M. T., Sabaj, K. M., and Tonge, P. J. (1999) *Arch Biochem Biophys* 370, 16-21.
30. Johnson, B. D., and Stankovich, M. T. (1993) *Biochemistry* 32, 10779-85.
31. Byron, C. M., Stankovich, M. T., and Husain, M. (1990) *Biochemistry* 29, 3691-700.
32. Bach, R. D., and Dmitrenko, O. (2003) *Journal of Physical Chemistry B* 107, 12851-12861.
33. Ghisla, S., and Thorpe, C. (2004) *Eur J Biochem* 271, 494-508.
34. Bach, R. D., and Dmitrenko, O. (2004) *J Am Chem Soc* 126, 127-42.
35. Satoh, A., Nakajima, Y., Miyahara, I., Hirotsu, K., Tanaka, T., Nishina, Y., Shiga, K., Tamaoki, H., Setoyama, C., and Miura, R. (2003) *J Biochem (Tokyo)* 134, 297-304.

36. Guo, F., Chang, B. H., and Rizzo, C. J. (2002) *Bioorg Med Chem Lett* 12, 151-4.
37. Rizzo, C. J. (2001) *Antioxid Redox Signal* 3, 737-46.
38. Reibenspies, J. H., Guo, F., and Rizzo, C. J. (2000) *Org Lett* 2, 903-6.
39. Binda, C., Li, M., Hubalek, F., Restelli, N., Edmondson, D. E., and Mattevi, A. (2003) *Proc. Natl. Acad. Sci. U. S. A.* 100, 9750-5.
40. Binda, C., Newton-Vinson, P., Hubalek, F., Edmondson, D. E., and, and Mattevi, A. (2002b) *Nat Struct Biol* 9, 22-6.
41. Binda, C., Mattevi, A., and Edmondson, D. E. (2002a) *J Biol Chem* 277, 23973-6.
42. Barna, T. M., Khan, H., Bruce, N. C., Barsukov, I., Scrutton, N. S., and Moody, P. C. (2001) *J Mol Biol* 310, 433-47.
43. Massey, V., and Hemmerich, P. (1980) *Biochem Soc Trans* 8, 246-57.
44. Abu-Soud, H., Clark, A., Francisco, W., Baldwin, T., and Raushel, F. (1993) *J. Biol. Chem.* 268, 7699-7706.
45. Beaty, N. B., and Ballou, D. P. (1981) *J Biol Chem* 256, 4619-25.
46. Jones, K. C., and Ballou, D. P. (1986) *J Biol Chem* 261, 2553-9.
47. Ghisla, S., Hastings, J. W., Favaudon, V., and Lhoste, J.-M. (1978) *PNAS* 75, 5860-5863.
48. Entsch, B., Ballou, D. P., and Massey, V. (1976) *J Biol Chem* 251, 2550-63.
49. Wada, N., Sugimoto, T., Watanabe, H., and Tu, S.-C. (1999) *Photochemistry and Photobiology* 70, 116-122.
50. Mallett, T. C., and Claiborne, A. (1998) *Biochemistry* 37, 8790-802.
51. Massey, V., Schopfer, L. M., and Anderson, R. F. (1988) (King, T. E., Mason, H. S., and Morrison, M., Eds.) pp 147-166, Alan R. Liss, Inc., New York.
52. Roth, J. P., Wincek, R., Nodet, G., Edmondson, D. E., McIntire, W. S., and Klinman, J. P. (2004) *J. Am. Chem. Soc.* 126, 15120-31.
53. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) *Nucleic Acids Res.* 28, 235-42.

54. Ikuta, S., Imamura, S., Misaki, H., and Horiuti, Y. (1977) *J Biochem (Tokyo)* 82, 1741-9.
55. Gadda, G. (2003) *Biochim Biophys Acta* 1646, 112-8.
56. Rand, T., Halkier, T., and Hansen, O. C. (2003) *Biochemistry* 42, 7188-94.
57. Ohta-Fukuyama, M., Miyake, Y., Emi, S., and Yamano, T. (1980) *J Biochem (Tokyo)* 88, 197-203.
58. Fan, F., Ghanem, M., and Gadda, G. (2004) *Arch Biochem Biophys* 421, 149-58.
59. Ghanem, M., and Gadda, G. (2005) *Biochemistry* 44, 893-904.
60. Velasco-Garcia, R., Chacon-Aguilar, V. M., Hervert-Hernandez, D., and Munoz-Clares, R. A. (2003) *Chem Biol Interact* 143-144, 149-58.
61. Ko, R., Smith, L. T., and Smith, G. M. (1994) *J Bacteriol* 176, 426-31.
62. Patchett, R. A., Kelly, A. F., and Kroll, R. G. (1994) *Arch Microbiol* 162, 205-10.
63. Peddie, B. A., Chambers, S. T., and Lever, M. (1996) *J Lab Clin Med* 128, 417-22.
64. Peddie, B. A., Phillips, A., and Chambers, S. T. (1996) *J Antimicrob Chemother* 37, 396-7.
65. Peddie, B. A., Wong-She, J., Randall, K., Lever, M., and Chambers, S. T. (1998) *FEMS Microbiol Lett* 160, 25-30.
66. Peddie, B. A., Lever, M., Randall, K., and Chambers, S. T. (1999) *Antonie Van Leeuwenhoek* 75, 183-9.
67. Sleator, R. D., Wouters, J., Gahan, C. G., Abee, T., and Hill, C. (2001) *Appl Environ Microbiol* 67, 2692-8.
68. Alia, Hayashi, H., Sakamoto, A., and Murata, N. (1998) *Plant J* 16, 155-61.
69. Alia, Hayashi, H., Chen, T. H. H., and Murata, N. (1998) *Plant, Cell and Environment* 21, 232-239.
70. Alia, Kondo, Y., Sakamoto, A., Nonaka, H., Hayashi, H., Saradhi, P. P., Chen, T. H., and Murata, N. (1999) *Plant Mol Biol* 40, 279-88.
71. Hayashi, H., Alia, Mustardy, L., Deshniem, P., Ida, M., and Murata, N. (1997) *Plant J* 12, 133-42.

72. Sakamoto, A., Valverde, R., Alia, Chen, T. H., and Murata, N. (2000) *Plant J* 22, 449-53.
73. Sakamoto, A., and Murata, N. (2000) *J Exp Bot* 51, 81-8.
74. Sakamoto, A., and Murata, N. (2001) *Plant Physiol* 125, 180-8.
75. Sakamoto, A., and Murata, N. (2002) *Plant Cell Environ* 25, 163-171.
76. Ohishi, N., and Yagi, K. (1979) *Biochem Biophys Res Commun* 86, 1084-8.
77. Fan, F., and Gadda, G. (2005) *J Am Chem Soc* 127, 2067-74.
78. Tsuge, H., Nakano, Y., Onishi, H., Futamura, Y., and Ohashi, K. (1980) *Biochim Biophys Acta* 614, 274-84.
79. Weibel, M. K., and Bright, H. J. (1971) *J Biol Chem* 246, 2734-44.
80. Gibson, Q. H., Swoboda, B. E., and Massey, V. (1964) *J Biol Chem* 239, 3927-34.
81. Kamei, T., Takiguchi, Y., Suzuki, H., Matsuzaki, M., and Nakamura, S. (1978) *Chem Pharm Bull (Tokyo)* 26, 2799-2804.
82. Higham, C. W., Gordon-Smith, D., Dempsey, C. E., and Wood, P. M. (1994) *FEBS Lett* 351, 128-32.
83. Cavener, D. R. (1992) *J Mol Biol* 223, 811-4.
84. Wohlfahrt, G., Witt, S., Hendle, J., Schomburg, D., Kalisz, H. M., and Hecht, H. J. (1999) *Acta Crystallogr D Biol Crystallogr* 55, 969-77.
85. Hecht, H. J., Kalisz, H. M., Hendle, J., Schmid, R. D., and Schomburg, D. (1993) *J Mol Biol* 229, 153-72.
86. Vrieling, A., Lloyd, L. F., and Blow, D. M. (1991) *J Mol Biol* 219, 533-54.
87. Li, J., Vrieling, A., Brick, P., and Blow, D. M. (1993) *Biochemistry* 32, 11507-15.
88. Lario, P. I., Sampson, N., and Vrieling, A. (2003) *J Mol Biol* 326, 1635-50.
89. Hallberg, B. M., Henriksson, G., Pettersson, G., and Divne, C. (2002) *J Mol Biol* 315, 421-34.
90. Yin, Y., Liu, P., Anderson, R. G., and Sampson, N. S. (2002) *Arch Biochem Biophys* 402, 235-42.

91. Kass, I. J., and Sampson, N. S. (1998) *Biochemistry* 37, 17990-8000.
92. Rotsaert, F. A., Renganathan, V., and Gold, M. H. (2003) *Biochemistry* 42, 4049-56.
93. Wohlfahrt, G., Witt, S., Hendle, J., Schomburg, D., Kalisz, H. M., and Hecht, H. J. (1999) *Acta Crystallogr. D55*, 969-77.
94. Su, Q., and Klinman, J. P. (1999) *Biochemistry* 38, 8572-81.
95. Ghanem, M., Fan, F., Francis, K., and Gadda, G. (2003) *Biochemistry* 42, 15179-88.
96. Hallberg, B. M., Henriksson, G., Pettersson, G., Vasella, A., and Divne, C. (2003) *J Biol Chem* 278, 7160-6.
97. Bright, H. J., and Gibson, Q. H. (1967) *J Biol Chem* 242, 994-1003.
98. Kass, I. J., and Sampson, N. S. (1998) *Bioorg Med Chem Lett* 8, 2663-8.
99. Menon, V., Hsieh, C.T., Fitzpatrick, P.F. (1995) *Biorg. Chem* 23, 42-53.
100. Urban, P., Lederer, F.J. (1985) *J. Biol. Chem.* 260, 11115-11122.
101. Dubois, J., Chapman, S.K., Mathews, F.S., Reid, G.A., Lederer, F. (1990) 29, 6393-6400.
102. Williams, R. F., Shinkai, S., and Bruice, T. C. (1975) *Proc Natl Acad Sci U S A* 72, 1763-7.
103. Gadda, G. (2003) *Biochim Biophys Acta* 1650, 4-9.
104. Gadda, G., Powell, N. L., and Menon, P. (2004) *Arch Biochem Biophys* 430, 264-73.
105. Holyoak, T., Fenn, T. D., Wilson, M. A., Moulin, A. G., Ringe, D., and Petsko, G. A. (2003) *Acta Crystallogr D Biol Crystallogr* 59, 2356-8.
106. Otwinowski, Z., and Minor, W. (1997) *Methods in Enzymology* 276, 307-326.
107. Vagin, A. A., and Teplyakov, A. (1997) *Journal of Applied Crystallography* 30, 1022.
108. COLLABORATIVE COMPUTATIONAL PROJECT, N. (1994) *Acta. Cryst D50*, 760-763.

109. Brunger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr D Biol Crystallogr* 54 (Pt 5), 905-21.
110. Kleywegt, G. J., and Jones, T. A. (1997) *Methods Enzymol.* 277, 208-230.
111. Pannu, N. S., Murshudov, G. N., Dodson, E. J., and Read, R. J. (1998) *Acta Cryst. D54*, 1285-94.
112. Read, R. J. (1997) *Methods Enzymol.* 277, 110-128.
113. Terwilliger, T. C. (2004) *Acta Crystallogr D Biol Crystallogr* 60, 2144-9.
114. Terwilliger, T. C. (2003) *Methods Enzymol* 374, 22-37.
115. van Aalten, D. M., Bywater, R., Findlay, J. B., Hendlich, M., Hooft, R. W., and Vriend, G. (1996) *J Comput Aided Mol Des* 10, 255-62.
116. Lamzin, V. S., and Wilson, K. S. (1993) *Acta Crystallogr D Biol Crystallogr* 49, 129-47.
117. Laskowski, R. A., Moss, D. S., and Thornton, J. M. (1993) *J Mol Biol* 231, 1049-67.
118. Morris, A. L., MacArthur, M. W., Hutchinson, E. G., and Thornton, J. M. (1992) *Proteins* 12, 345-64.
119. Kabsch, W., and Sander, C. (1983) *Biopolymers* 22, 2577-637.
120. Matthews, B. W. (1968) *J Mol Biol* 33, 491-7.
121. Lennon, B. W., Williams, C. H., Jr., and Ludwig, M. L. (1999) *Protein Sci* 8, 2366-79.
122. Dixon, D. A., Lindner, D. L., Branchaud, B., and Lipscomb, W. N. (1979) *Biochemistry* 18, 5770-5.
123. Schlichting, I., Berendzen, J., Chu, K., Stock, A. M., Maves, S. A., Benson, D. E., Sweet, R. M., Ringe, D., Petsko, G. A., and Sligar, S. G. (2000) *Science* 287, 1615-22.
124. Karlsson, A., Parales, J. V., Parales, R. E., Gibson, D. T., Eklund, H., and Ramaswamy, S. (2003) *Science* 299, 1039-42.

125. Ridder, L., Mulholland, A. J., Rietjens, I. M., and Vervoort, J. (1999) *J Mol Graph Model* 17, 163-75, 214.
126. Hallberg, B. M., Leitner, C., Haltrich, D., and Divne, C. (2004) *J Mol Biol* 341, 781-96.
127. Bannwarth, M., Bastian, S., Heckmann-Pohl, D., Giffhorn, F., and Schulz, G. E. (2004) *Biochemistry* 43, 11683-90.
128. Yue, Q. K., Kass, I. J., Sampson, N. S., and Vrieling, A. (1999) *Biochemistry* 38, 4277-86.
129. Martin, S. F., Follows, B. C., Hergenrother, P. J., and Trotter, B. K. (2000) *Biochemistry* 39, 3410-5.
130. Pomponi, M., Sacchi, S., Colella, A., Patamia, M., and Marta, M. (1998) *Biophys Chem* 72, 239-46.
131. Harel, M., Kryger, G., Rosenberry, T. L., Mallender, W. D., Lewis, T., Fletcher, R. J., Guss, J. M., Silman, I., and Sussman, J. L. (2000) *Protein Sci* 9, 1063-72.
132. Kwak, B. Y., Zhang, Y. M., Yun, M., Heath, R. J., Rock, C. O., Jackowski, S., and Park, H. W. (2002) *J Biol Chem* 277, 4343-50.
133. Vrieling, A., and Sampson, N. (2003) *Curr Opin Struct Biol* 13, 709-15.
134. Carugo, O., Carugo, K. D. (2005) *Trends Biochem Sci* 4, 213.
135. Chance, B., Angiolillo, P., Yang, E. K., and Powers, L. (1980) *FEBS Lett* 112, 178-82.
136. Schlichting, I., Berendzen, J., Phillips, G. N., Jr., and Sweet, R. M. (1994) *Nature* 371, 808-12.
137. Srajer, V., Teng, T., Ursby, T., Pradervand, C., Ren, Z., Adachi, S., Schildkamp, W., Bourgeois, D., Wulff, M., and Moffat, K. (1996) *Science* 274, 1726-9.
138. Teng, T. Y., Srajer, V., and Moffat, K. (1994) *Nat. Struct. Biol.* 1, 701-5.
139. Lim, M., Jackson, T. A., and Anfinsen, P. A. (1997) *Nat. Struct. Biol.* 4, 209-14.
140. Austin, R. H., Beeson, K., Eisenstein, L., Frauenfelder, H., Gunsalus, I. C., and Marshall, V. P. (1973) *Science* 181, 541-3.
141. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H., and Gunsalus, I. C. (1975) *Biochemistry* 14, 5355-73.

142. Orville, A. M., Chen, V. J., Kriauciunas, A., Harpel, M. R., Fox, B. G., Munck, E., and Lipscomb, J. D. (1992) *Biochemistry* 31, 4602-12.
143. Orville, A. M., and Lipscomb, J. D. (1993) *J Biol Chem* 268, 8596-607.
144. Orville, A. M., and Lipscomb, J. D. (1997) *Biochemistry* 36, 14044-55.
145. Canepa, C., Bach, R. D., and Dmitrenko, O. (2002) *J Org Chem* 67, 8653-61.
146. Yin, Y., Sampson, N. S., Vrielink, A., and Lario, P. I. (2001) *Biochemistry* 40, 13779-87.

VITA

GEORGE T. LOUNTOS

George T. Lountos was born to Themistoclis and Spiridoula Lountos in Thomasville, Georgia on February 9, 1978. He attended the Thomasville City Schools and graduated from Thomasville High School in 1996 with honors. He enrolled in the Georgia Institute of Technology in Fall of 1996 and participated in the Georgia Tech Oxford Study Abroad Program during the summer of 1999. In 2000, he received a Bachelor of Science degree in Chemistry with high honor from the Georgia Institute of Technology. As an undergraduate, he was a member of the laboratory of Professor Suzanne Shuker where his research focused on organic synthesis. He was a recipient of the 2000 Merck Index Award awarded by the faculty of the School of Chemistry & Biochemistry. In the Fall of 2000, he joined Professor Allen Orville's laboratory at Georgia Tech and completed his Ph.D. in Chemistry in November 2005. As a Graduate Teaching Assistant, he received the 2003 Outstanding Performance as a GTA Award in the School of Chemistry & Biochemistry.