

ROLES OF AN 'INACTIVE' DOMAIN IN CATALASE-PEROXIDASE CATALYSIS:
MODULATION OF ACTIVE SITE ARCHITECTURE AND FUNCTION BY GENE
DUPLICATION

Except where reference is made to the work of others, the work described in this dissertation is my own or was done in collaboration with my advisory committee. This dissertation does not include proprietary or classified information.

Ruletha D. Baker

Certificate of Approval:

J. Howard Hargis
Professor
Chemistry and Biochemistry

Douglas C. Goodwin, Chair
Associate Professor
Chemistry and Biochemistry

Holly R. Ellis
Associate Professor
Chemistry and Biochemistry

Thomas E. Albrecht-Schmitt
Associate Professor
Chemistry and Biochemistry

Joe F. Pittman
Dean
Graduate School

ROLES OF AN 'INACTIVE' DOMAIN IN CATALASE-PEROXIDASE CATALYSIS:
MODULATION OF ACTIVE SITE ARCHITECTURE AND FUNCTION BY GENE
DUPLICATION

Ruletha Deon Baker

A Dissertation

Submitted to

the Graduate Faculty of

Auburn University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

Auburn, Alabama
December 15, 2006

ROLES OF AN 'INACTIVE' DOMAIN IN CATALASE-PEROXIDASE CATALYSIS:
MODULATION OF ACTIVE SITE ARCHITECTURE AND FUNCTION BY GENE
DUPLICATION

Ruletha Deon Baker

Permission is granted to Auburn University to make copies of this dissertation at its discretion, upon request of individuals or institutions and at their expense.
The author reserves rights to publication.

Signature of Author

Date of Graduation

VITA

Ruletha Deon Baker, daughter of Charles Gaston Jr. and Ruseda Price Baker, was born on March 11, 1973, in Louisville, Alabama. She graduated from George W. Long High School in 1991. In the spring of 1994, she transferred from George C. Wallace Community College with an Associate in Arts degree (cum laude). In the spring of 1997 she graduated from Troy State University in Troy, Alabama, with a Bachelor of Science (cum laude) in Chemistry. She entered the Graduate School, Auburn University, Auburn, Alabama, in August, 2000. She is a member of Phi Lambda Upsilon, National Honorary Chemical Society. She married Dwayne Keith Hartfield on December, 16 2005.

DISSERTATION ABSTRACT

ROLES OF AN 'INACTIVE' DOMAIN IN CATALASE-PEROXIDASE CATALYSIS:
MODULATION OF ACTIVE SITE ARCHITECTURE AND FUNCTION BY GENE
DUPLICATION

Ruletha Deon Baker

Doctor of Philosophy, December 15, 2006
(B.S., Troy State University, 1997)
(A.S.A., George C. Wallace College, 1994)

217 Typed Pages

Directed by Douglas C. Goodwin

Within biological systems iron is a transition metal that allows access to the benefits of molecular oxygen as an oxidant. However, with these benefits come grave consequences if the reactions are not strictly controlled. The most prominent strategy of control and specialization is the protein environment that surrounds iron. Within iron containing proteins, specifically heme proteins, there are four basic levels of structure that impact the iron's function: cofactor structure, protein-supplied ligands, non-ligand active site environment, and protein features that are distant from the active site. This last level remains poorly understood due to a lack of good models to pursue such studies.

Catalase-peroxidases are unique heme proteins because they catalyze peroxide decomposition by two separate mechanisms, catalase and peroxidase, using the same active site. However, were it not for three structural features distant from the active site, catalase-peroxidases would be practically superimposable with monofunctional peroxidases. Given the additional reactivity of catalase-peroxidases it seems probable that these features would be involved in the function and would, therefore, provide excellent models for understanding the roles of features distant from the active site.

The largest of these three features, the C-terminal domain, about 300 amino acids, is situated ≥ 30 Å from the active site. In order to determine its role, we produced a variant of *E. coli* catalase-peroxidase (KatG) lacking its C-terminal domain (KatG^{Nterm}). KatG^{Nterm} was expressed in inclusion bodies and required us to develop a denatured purification procedure, as well as refolding and reconstitution protocols in order to isolate it. A soluble, well-behaved KatG^{Nterm}, was produced; however, it had neither catalase nor peroxidase activity. Far-UV CD and UV-visible spectroscopy confirmed that the protein was correctly folded, and that heme was bound within the active site. Visible absorption and EPR spectra indicated a shift from the predominately high-spin pentacoordinate heme environment to an exclusively hexacoordinate low-spin environment. These results were also observed in the EPR spectrum of KatG^{Nterm}. Peroxyacetic acid and cyanide binding experiments revealed what seem to be an occluded active site. Site-directed substitution of KatG^{Nterm} active site histidine residues with alanine showed that

His106, normally required as a general base, had become a ligand to the heme. Reincorporation of a separately expressed an isolated C-terminal domain, KatG^{Cterm}, resulted in recovery catalase *and* peroxidase activity. Both EPR and UV-visible spectroscopy confirmed the restoration of the overall architecture of the active site. This research demonstrates that the C-terminal domain is essential for activity by maintaining the correct coordination environment for the heme, even though it is some 30 Å away from the active site.

ACKNOWLEDGMENTS

Many people have played a part in the success of the research presented here as well as my graduate career. First, and obvious thanks to Dr. Douglas C. Goodwin for his guidance and support, without which this work would not have been possible. Additionally, thanks for listening and advice during the trying times. Much thanks to the committee members: Dr. Holly R. Ellis, Dr. Thomas E. Albrecht-Schmitt, and Dr. Howard W. Hargis for their time and helpful discussion. Thanks to Dr. Evert C. Duin for his reviews and assistance. Dr. Albrecht-Schmitt, Dr. Tyler Sullens and Thomas Shehee are thanked for their assistance with crystallography. A special thanks to Dr. Holly R. Ellis for so many things but most of all for listening and motivating with a compassionate and enduring spirit. Thanks to Carma O. Cook for all the work with KatG^{Cterm} and with putting “our” domains together. Much thanks to Emily C. Brantley, Kenneth Roberts, Kristen M. Hertwig and Sarah Peelsey for their technical assistance and friendship. To rest of the original “Goodwin Group”, Neil Varnado and Yongjiang Li, I cannot express what your friendship and support has meant, I will miss you both. To my little sister, Charla N. Baker you cannot imagine how I will cherish going to graduate school with you through the thick and the thin it has been a blessing. To my parents I thank you for getting me here, for the support and all the skills you’ve both given me. Lastly, to my husband, Dwayne K. Hartfield, thank you for ever being my rock, lending me your shoulder and for helping me understand the “easy stuff”.

Style manual or journal used:

FASAB Journal

Computer software used:

Word Processor: Microsoft Word (for Windows)

Graphs: Prism (for Macintosh) and Excel (for Windows and Macintosh)

Molecules: ChemDraw

Plasmid diagrams: PlasmidDraw

TABLE OF CONTENTS

LIST OF TABLES.....	xii
LIST OF FIGURES.....	xiii
INTRODUCTION.....	1
LITERATURE REVIEW.....	8
Oxygen Availability, Acquisition, and Reactivity.....	9
Iron cofactors: Structure as a Functional Control.....	17
Heme Protein structure and Function: Strategies for Controlling the Iron. Oxygen Interaction in Biology.....	38
Methods for Determining Structure and Function of Heme Proteins.....	67
METHODS AND MATERIALS.....	108
Materials.....	108
Cloning and Expression of the Heme Receptor Protein (ChuA).....	109
Cloning and Expression of Wild-type KatG.....	114
Construction and Expression of KatG ^{Nterm} and KatG ^{Cterm}	116
Heme Receptor Activity Determination.....	121
Recombination and Characterization of KatG ^{Nterm} and KatG ^{Cterm}	122

RESULTS.....	126
Heme Receptor.....	126
KatG ^{Nterm}	130
Recombination of KatG ^{Nterm} and KatG ^{Cterm}	153
DISCUSSION.....	169
Identifying the Role of the C-terminal Domain.....	171
Restructuring of Active Site.....	172
Future Directions.....	180
Summary.....	182

LIST OF TABLES

1. UV-Visible absorption maxima for some peroxidases.....76
2. MCD values for some hemoproteins.....80
3. Heme absorption maxima spectra for wtKatG, KatG^{Nterm}, LiP, and MnP.....140

LIST OF FIGURES

1. Oxidation of ascorbate by oxygen in solution resulting in ascorbyl radical	11
2. Ground state electron configurations for triplet oxygen ($^3\text{O}_2$) and singlet oxygen ($^1\text{O}_2$).....	13
3. Diagram of the active site ligands of Fe-SOD.....	20
4. Overall reaction scheme for lipoxygenase catalysis.....	23
5. Diagram of ligand in lipoxygenase active site.....	24
6. Proposed reaction mechanism for lipoxygenase.....	25
7. Reaction scheme for methane monooxygenase.....	26
8. Diagram of ligands in methane monooxygenase active site.....	27
9. Diagram of iron-sulfur cluster types.....	28
10. Diagrams of ferridoxin [2Fe-2S] (a) and rubredoxin (b) active site ligands.....	34
11. Reaction scheme for the reversible transformation of citrate to isocitrate catalyzed by aconitase.....	36
12. Diagram of the active site ligands in aconitase.....	37
13. Structures of heme cofactors; heme a, heme b, and heme c.....	38
14. Esterified heme b in peroxidases forming a covalent adduct with the protein, represented by (R).....	41
15. Diagram of active site ligands to heme in cytochrome P450.....	44

16	Reaction mechanism for cytochrome P450.....	45
17.	Reactions converting NO to peroxynitrite and subsequently peroxynitrous acid..	47
18.	Diagram of active site ligands for hemoglobin and myoglobin unbound (left) and O ₂ bound (right).....	49
19.	Diagram of active site ligands for catalases.....	52
20.	Catalase reaction scheme.....	53
21.	Reaction scheme for monofunctional peroxidases.....	55
22.	Diagram of active site ligands for monofunctional peroxidases.....	56
23.	Important active site residues in monofunctional peroxidases and myoglobin....	59
24.	Catalase-peroxidase protein catalase and peroxidase reaction scheme.....	62
25.	Ribbon diagram of catalase-peroxidase KatG.....	64
26.	Distribution effects on the d electrons of iron (III) in (a) no field, (b) weak ligand field, or (c) strong ligand within an octahedral ligand field.....	70
27.	Diagram of absorption bands typical for peroxidases.....	72
28.	Typical absorption spectra for KatG in the Soret (left) and charge transfer regions.....	74
29.	Approximate CD spectra for poly-lysine in three different secondary-structural conformations; (a) α -helix, (b) β -sheet, and (c) random coil.....	78
30.	Approximation of EPR spectra for completely axial signal (a), and completely rhombic signal (b).....	85
31.	Stopped-flow diagram.....	95

32. Ribbon diagram of approximate FhuA interior and cork structures for (a) apo-structure, before heme binding, and (b) holo-structure, with heme bound.....	103
33. Diagram of a model of heme receptor function in the transfer of iron from the exterior environment of the bacterium into the cytoplasm.....	105
34. SDS-PAGE gel of KatG ^{Nterm} solubility.....	131
35. Gel of purified wtKatG and variants.....	133
36. FPLC of KatG ^{Nterm}	134
37. Catalase and peroxidase activities for KatG ^{Nterm} and wtKatG.....	136
38. Far-UV circular dichroism spectra for wtKatG and each of its separately expressed and isolated domains.....	137
39. Absorption spectra for the ferric forms of wtKatG, KatG ^{Nterm} , and free hemin.....	138
40. UV-Visible spectra KatG ^{Nterm} Soret region.....	141
41. UV-Visible spectra of KatG ^{Nterm} charge transfer region.....	142
42. EPR of wtKatG and KatG ^{Nterm}	143
43. Active site histidines of catalase-peroxidase in MnP.....	145
44. UV-Visible absorption spectra for KatG ^{Nterm} and histidine variants.....	146
45. Stopped-flow reaction of KatG ^{Nterm} with peroxyacetic acid (PAA).....	148
46. Cyanide binding spectra for KatG ^{Nterm}	149
47. Time course for cyanide binding to KatG ^{Nterm}	150
48. Effect of cyanide concentration on k_{obs} for cyanide binding to KatG ^{Nterm}	151
49. Effect of cyanide concentration on the amplitude of binding traces for KatG ^{Nterm}	152
50. Changes in active site coordination in MnP and LiP upon loss of Ca ²⁺	154

51. Far UV circular dichroism spectra for wtKatG and each of its separately expressed and isolated domains.....	156
52. Physical and arithmetic mixtures of KatG ^{Nterm} and KatG ^{Cterm}	157
53. Reactivation of catalase activity for KatG ^{Nterm} by addition of KatG ^{Cterm}	158
54. Reactivation of peroxidase activity for KatG ^{Nterm} by addition of KatG ^{Cterm}	159
55. Catalase and peroxidase activities over time for an equimolar mixture of KatG ^{Nterm} and KatG ^{Cterm}	160
56. Effect of H ₂ O ₂ on the catalase activity of an equimolar mixture of KatG ^{Nterm} and KatG ^{Cterm}	161
57. Effect of H ₂ O ₂ on the peroxidase activity of an equimolar mixture of KatG ^{Nterm} and KatG ^{Cterm}	163
58. Effect of separately expressed and isolated KatG ^{Cterm} on the visible absorption spectrum of KatG ^{Nterm}	164
59. Effect of H ₂ O ₂ on a equimolar mixture of KatG ^{Nterm} and KatG ^{Cterm}	166
60. Effect of separately expressed and isolated KatG ^{Cterm} on the EPR spectrum for KatG ^{Nterm}	167
61. Effect of H ₂ O ₂ on the EPR spectrum of the equimolar combination of KatG ^{Nterm} and KatG ^{Cterm}	168
62. Structural representation of the major helices surrounding the active sites of catalase-peroxidase for <i>Burkholderia pseudomallei</i>	174

I. INTRODUCTION

The fact that living organisms are frequently subdivided by virtue of their need or aversion to molecular oxygen (aerobe, facultative anaerobe, etc.) is an indication of its prominence in biology. Indeed, O_2 is a highly versatile oxidant, and as a terminal electron acceptor, it allows access to high yields of ATP from the oxidation of carbon in food sources. Yet, due to its electronic configuration, the direct reaction of O_2 with almost all biological molecules is spin-forbidden. This kinetic barrier can be circumvented by the use of redox active transition metals as catalysts because they are able to occupy multiple oxidation and spin states. It is, then, not unexpected that these transition metals, especially iron, are ubiquitous in biology. Although iron and other transition metals allow access to the benefits of O_2 , as an oxidant the consequences of the uncontrolled interaction between these metals and O_2 are dire. Partially reduced oxygen species, also known as reactive oxygen species, cause damage, often irreparable, to all classes of biological molecules. Not surprisingly, the interaction between iron and O_2 is tightly controlled and specialized at all stages: iron acquisition, transport, and storage; oxygen transport and storage; and any point where the catalytic properties of iron are needed. Not only this, but multiple mechanisms are in place to minimize damage when things go awry.

The most prominent strategy of control and specialization is, in essence, the protein environment that surrounds iron at its points of contact with O₂, its derivatives, and other biologically important molecules. This is evident in all iron-containing proteins, especially those that rely on the iron-porphyrin cofactor known as heme. Heme proteins are integral to nearly all processes that relate to the use of molecular oxygen, not only its benefits, but also its consequences. In each case, it is the environment surrounding the iron that dictates the outcome of the interaction.

There are four basic levels of structure that impact the iron, and therefore, its function. The first is the structure of the cofactor itself, which is relatively limited. At the second level, the protein can supply one or two ligands to the heme iron. The third level encompasses the immediate active site environment of the heme and its ligands, and finally, the fourth level involves long range effects from protein features that are distant from the active site. The first three have been well studied and have provided much insight into the protein structure/function question. However, the fourth level remains poorly understood due to a lack of good models to pursue such studies.

Catalase-peroxidases are unique heme proteins because they catalyze peroxide decomposition by two separate mechanisms, catalase and peroxidase. What is most striking about this is that these enzymes use the same active site for both mechanisms. Given their similar reactions one would expect some similarities between the catalase-peroxidases and the related monofunctional enzymes, monofunctional catalase and

monofunctional peroxidase. In fact, the catalase-peroxidases are dually considered hydroperoxidases or a special class of catalases depending upon the source of classification.

Interestingly, comparison of catalase-peroxidases to monofunctional catalases reveals practically no sequence homology or structural similarity to catalase-peroxidases. Conversely, comparison of catalase-peroxidases to the monofunctional peroxidases reveals a great deal of similarity indeed. The catalase-peroxidases share high sequence homology with the monofunctional peroxidase features, and structurally the two groups would be practically superimposable were it not for three distinct protein features. These include an extension of the loop between the D and E helices, an extension of the loop between the F and G helices (both about 35 amino acids), and a large domain of about 300 amino acids on the C-terminal end of the protein, the C-terminal domain. All three features are conserved across all catalase-peroxidases. Also, all three features are peripheral to the active site. Here it would seem is an ideal model to investigate the fourth level of heme protein structure and function. Catalase-peroxidases, which share a virtually superimposable active site with monofunctional peroxidases, show a distinct functional difference, catalase activity. The structural differences between these two groups of proteins are primarily *outside* of the active site.

Consistent with this hypothesis, The Goodwin Laboratory and others have shown the extended interhelical loops are essential for catalase, but not peroxidase activity. However, of the three features unique to catalase-peroxidases, the C-terminal domain lies the greatest distance from the active site, $\geq 30 \text{ \AA}$. This domain appears to be the result of gene duplication and fusion as its overall fold strongly resembles a monofunctional

peroxidase. Over time the C-terminal domain has lost the ability to bind heme or catalyze any discernable reaction. This would suggest that the domain is a vestigial artifact of the past with no real contribution to catalase-peroxidase function. On the other hand, this domain is a conserved structure in all catalase-peroxidases. Nevertheless, prior to the research described here, the role(s) of the C-terminal domain in catalase-peroxidase structure and catalysis had not been explored and were unknown.

To address this deficiency, we produced a variant of *Escherichia coli* catalase-peroxidase (KatG) lacking its C-terminal domain, KatG^{Nterm}. The protein was expressed at high levels, but it was found exclusively insoluble inclusion bodies, indicating, at the very least that the C-terminal domain was essential for assisting in protein folding during expression. Procedures were developed for purification under denatured conditions, refolding, and reconstitution of KatG^{Nterm}. These procedures produced a soluble and stable protein for evaluation by a variety of spectroscopic and kinetic methods.

KatG^{Nterm} had neither catalase nor peroxidase activity. Circular dichroism (CD) spectra were recorded to determine the overall folded conformation of the protein. These spectra indicated the modification produced no unexpected changes in the secondary structural content of the protein. UV-visible absorption spectra and EPR spectra indicate a hexacoordinate low-spin heme coordination environment. This is in stark contrast to the predominately high-spin heme species which are found in a mixture of pentacoordinate and hexacoordinate states in the wtKatG. Stopped-flow kinetic analysis of KatG^{Nterm} reaction with peroxide (peracetic acid) was used to evaluate compound I formation by the variant. At low peroxide concentrations, no reaction was observed, and at higher peracetic acid concentrations, only the slow bleaching of the heme group was

detected. The binding of CN⁻ to the heme of KatG^{Nterm} was also explored by stopped-flow. Consistent with the peroxide results, binding of CN⁻ was very slow, three orders of magnitude slower than wild-type KatG (wtKatG). These results were consistent with occluded access to the heme iron in KatG^{Nterm} compared to wild-type KatG.

Absorption maxima indicate a strong-field as the sixth ligand to the heme iron resulting in the low-spin state. The EPR spectra for KatG^{Nterm} indicate that the ligand is most likely a histidine. Comparison the monofunctional peroxidases reveals that the strictly conserved distal histidine, which functions as a general base in catalysis, is the likely candidate for the sixth ligand to the heme iron. Proteins lacking the distal histidine, the proximal histidine (the primary ligand for heme in the active site) and both together were produced using site directed mutagenesis. The resulting proteins revealed that the heme is ligated by both of these histidine in KatG^{Nterm} active site. Therefore, the distal histidine appears to be the sixth ligand to the heme iron in KatG^{Nterm}. These data suggest that in the absence of the C-terminal domain structural changes occur which allow the distal histidine to coordinate to the sixth position of the heme iron in the active site resulting in a hexacoordinate low-spin heme environment and a complete loss of both catalase *and* peroxidase activity.

The loss of activity in KatG^{Nterm} was reversed by incubation with an equimolar concentration of separately expressed an isolated C-terminal domain (KatG^{Cterm}). The return of both catalase *and* peroxidase activity was observed. Moreover, this can only be accomplished with a mixture of KatG^{Nterm} and KatG^{Cterm}. Neither protein by itself is active nor is a generic protein able to restore activity in KatG^{Nterm}.

Visible absorption and EPR spectroscopy both show the restoration of an active site architecture that mimics very closely that of wtKatG. Thus, the introduction of KatG^{Cterm} serves not only to draw the distal histidine out of the coordination sphere of the heme iron, but also restores the overall architecture of the active site (*I*). Interestingly, the reactivation process is relatively slow with an apparent first order rate constant of about $4 \times 10^{-5} \text{ s}^{-1}$. The slow reactivation as well as the simultaneous return of both activities indicates that a structural reorganization event external to the active site proceeds alteration of the coordination environment.

This would suggest that the C-terminal domain also may function as a platform to direct the folding of the N-terminal domain for catalysis. Inspection of the known structures of catalase-peroxidases suggest that the *I'* helix may function in just such a role. Very similar phenomena are observed in some monofunctional peroxidases upon the loss of Ca^{2+} . In these proteins, Ca^{2+} serves to stabilize the helices of the active site, particularly the B helix. This prevents coordination of the heme iron by the distal histidine and allows it fill its role as an essential general base in catalysis. Structural analysis of catalase-peroxidases indicates that the C-terminal domain may fill a similar role in the function of these enzymes. Indeed, this points to a general requirement of catalysis in peroxidase and other similar enzymes. That is, reactive groups must be present within the active site, but at the same time, their favorable coordination to the heme must be prevented to maintain a functional active site. The C-terminal domain appears to be a novel strategy to fill this critical role in catalysis.

The research described in this dissertation shows clearly, and for the first time, that the C-terminal domain is essential for catalase-peroxidase functions. From $\geq 30 \text{ \AA}$

away it maintains a very precise active site architecture that includes a correctly poised general base for catalysis, a correct coordination environment for substrate across the heme iron, and the positioning of amino acids for a cross link that is essential for catalase activity. This research shows that catalase-peroxidase, by using a C-terminal domain, solves an important and common obstacle to catalysis by a completely novel mechanism. Furthermore, the use of a C-terminal domain in such a role (as opposed to simple Ca^{2+} coordination) allows its use in other capacities as well. In particular, it supports mechanisms to direct cofactor-independent folding of the active site-bearing domain.

Therefore, in addition to providing key discoveries in the structure and function of catalase-peroxidases that may be applied to important biomedical issues involving these proteins (e.g. antibiotic resistance of *M. tuberculosis* and virulence of other pathogens), this research also identifies how gene duplication and fusion can be used as a strategy to simultaneously fill a broad range of mechanistic requirements. In this report, the structure of the C-terminal domain and the mechanisms by which it supports catalysis hold variable insight for engineering new enzymes to fill novel functions.

II. LITERATURE REVIEW

All organisms have three basic necessities for life which are, a carbon source to provide raw material for the biosynthesis of organic compounds, energy for biosynthesis and transport of said compounds, and lastly, a source of electrons for biosynthesis. Much of the energy currency of the cell is generated by electron transfer ultimately from nutrients sources to a terminal electron acceptor. Of the possible terminal electron acceptors for these processes, oxygen produces the greatest yield of (adenosine triphosphate) ATP followed by nitrate, sulfate and then carbon dioxide. The least productive method is fermentation in which there is no net oxidation of the starting material.

Kinetic rates follow the same trend as energy provided, with oxygen the most rapid, then nitrate, and so on. If an organism uses oxygen as a terminal electron acceptor, its utilization is called aerobic respiration and the organism is called an aerobe. On the contrary if other sources like nitrate are utilized then this is referred to as anerobic respiration and the organisms are called anerobes. Oxygen is fatal to strict anerobes, and though not utilized as a mechanism for ATP production, facultative anerobes must deal

with oxygen in their environment. Therefore, oxygen is a concern for all organisms whether aerobic, strictly anaerobic or facultatively anaerobic. For this reason the availability, acquisition and utilization of oxygen should be addressed.

Oxygen Availability, Acquisition, and Reactivity.

The biospheric oxygen cycle describes the equilibrium which exists between the production and consumption of oxygen. Biospheric oxygen is consumed, for the most part by living organisms through aerobic metabolism, but also by industrial processes and there is some uptake of oxygen due to weathering of fossilized carbon. Conversely oxygen is produced by photosynthesis in plants and net emission from buried atmospheric carbon dioxide (CO₂).

Of course, on the consumption end of the cycle oxygen must first be acquired by the organisms that utilize it. The simplest systems for O₂ acquisition are found in aerobic bacteria. They acquire oxygen from the surrounding environment by diffusion through the plasma membrane. Transport for bacteria is simple because of the membrane surface area of the cell is so large in comparison to its size that the entire cytoplasm has access to oxygen as it diffuses across the membrane. As a rule, the equilibrium of nutrient provision and consumption dictates the growth of an organism. In bacteria, oxygen is the limiting nutrient in that the growth of the size of the cell, and therefore, the membrane will expand until consumption of oxygen exceeds the quantity of oxygen provided by

diffusion. For larger, multicellular organisms, a more complex system is required to solve the diffusion problem.

Insects use an exoskeletal tunnel system which branches toward a specialized vesicle called the tracheole which allows oxygen/carbon dioxide exchange with individual cells. On the other end of the spectrum, in mammalian organisms, inhaled dioxygen passes into the lungs and diffuses into the blood stream via gas exchange between alveoli, air sacs which hold inhaled air, and the complex network of capillaries that surround them. The thin epithelia of the alveoli and capillaries allow erythrocytes to come within $0.7 \mu\text{m}$ of inhaled air, thus allowing for relative efficient oxygen uptake by hemoglobin. Non-mammalian acquire oxygen through gill systems or via absorption through the skin. No matter the method both are similar with respect to the fact that the areas of gas exchange, be it gill, lung or skin, must be rich in vascularization to allow sufficient contact with oxygen sources. The protein centerpiece of these latter systems is hemoglobin, a heme-dependent protein discussed in depth later in this review.

Metabolic Power of Oxygen

Considering the varied methods of obtaining oxygen, no matter the classification, the necessity of oxygen for most organisms seems obvious. The question of why oxygen is so important is more complex. Truly great metabolic power is found in the oxidative potential of oxygen. The reduction of O_2 to H_2O has a potential ($E^{\circ'} = 0.8 \text{ V}$) sufficient to oxidize virtually any biological molecule. The rate of reaction, however, is a

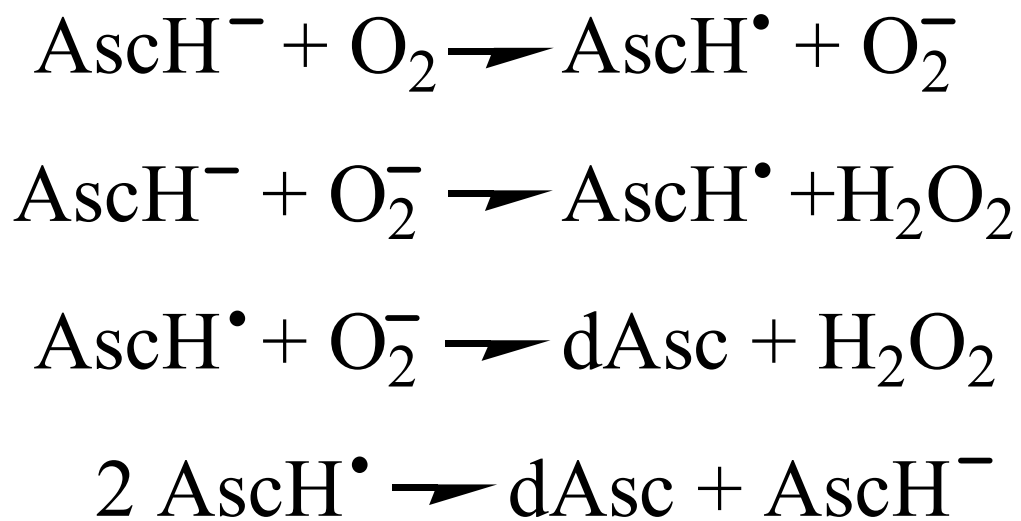


Figure 1. Oxidation of ascorbate by oxygen in solution resulting in ascorbyl radical (2). For these reactions ascorbate anion is represented by AscH⁻, ascorbate radical by AscH[•], and dehydroascorbic acid by dAsc

completely different consideration. The simple oxidation of ascorbate by O_2 in solution gives us a picture of this paradox. Ascorbate, $AscH^-$, at physiological pH (7.4), is an excellent reducing agent, $\Delta E^\circ = -0.06V$ readily undergoing two consecutive, yet reversible, one-electron oxidations to form, first, a relatively unreactive ascorbyl radical ($Asc^{\bullet-}$) [Figure 1] followed by dehydroxyascorbate (dAsc).

Interestingly, when ascorbate is exposed to dioxygen, it is only very slowly oxidized. The relative reduction potentials of ascorbate and O_2 indicate a kinetic rather than thermodynamic barrier. In fact, the key to understanding the reactivity of oxygen lies in the understanding that while the oxidation of biomolecules by O_2 are thermodynamically favorable, kinetically these reactions are extremely slow (3), and the kinetic barrier results from the fact that these reactions are spin-forbidden. Ground- state dioxygen exists in an relatively unreactive, triplet state, 3O_2 . It is considered a free radical. In fact, it is a diradical, as it has two unpaired electrons, and the two unpaired electrons of triplet oxygen have parallel spin ($\uparrow \uparrow$). As a result 3O_2 is unreactive with the vast majority of biomolecules because they are singlet species. Such a kinetic barrier does not exist for singlet oxygen, 1O_2 . Thus, the slow reactions of 3O_2 can be circumvented by converting oxygen to ‘active’ singlet oxygen. Singlet oxygen is produced when sufficient energy is applied to reverse the spin of one of its unpaired electrons, resulting in an antiparallel spin state ($\uparrow \downarrow$) [Figure 2].

In many biological systems the generation of singlet oxygen is not practical as a strategy for activating 3O_2 . Other means must be drawn upon to remove the kinetic barrier. Redox active transition metals, due to their ability to occupy multiple oxidation and spin states, are commonly observed to catalyze otherwise slow oxidations

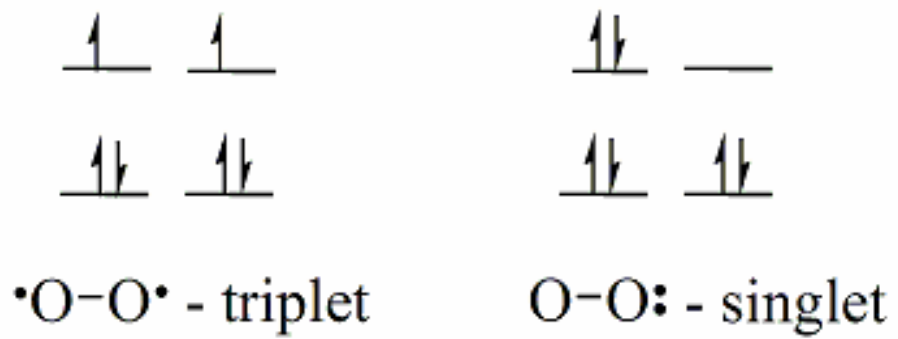
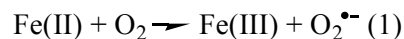
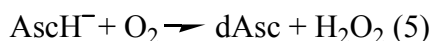


Figure 2. Ground state electron configurations for triplet oxygen ($^3\text{O}_2$) and singlet oxygen ($^1\text{O}_2$)

involving O₂ as an electron acceptor. For example, iron, in its reduced state, is able to transfer an electron to O₂, generating superoxide anion radical (reaction 1).

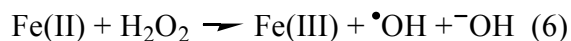


Though more reactive than triplet O₂, superoxide is not a particularly reactive molecule. Nonetheless, under biological conditions superoxide typically reacts with itself to form hydrogen peroxide and triplet oxygen (reaction 2).



Indeed, this very reaction is used to estimate the level of contamination of biological buffers and solutions with redox active transition metals, and is so simple that one can perform the reaction in minutes.

Interestingly, the inclusion of iron with ascorbate and ³O₂ allows for additional reactions. The reduction of H₂O₂ by ferrous iron, for example, produces the highly reactive hydroxyl radical ([•]OH) by the well known Fenton reaction (reaction 6).



Hydroxyl radical (E°= 2.0 V) reacts, often at diffusion controlled rates, with a wide range of compounds, generating hydroxylated, hydrogen abstracted, or otherwise oxidized target compounds (4). Clearly, transition metals have the capacity to activate ³O₂, generating species that are highly reactive and versatile. Of course, this is a necessity for life, but there is a down side. The unscrupulous generation of [•]OH (or species like it) will result in damage (in many cases catastrophic and irreparable) to all classes of biological molecules.

Iron in biological systems

Transition metals, especially iron, are clearly useful in the activation of O₂ to catalyze biologically necessary oxidations. For this reason, iron, in one form or another, is required by virtually all living organisms, and it covers a wide array of functions in proteins and in enzymes (5). It is, however, interesting to consider why iron is preferred over other metals. It is its availability, spin state and redox tunability that make iron attractive for biology. Since iron is the most abundant transition metal in the universe and is the fourth most abundant element in the earth's crust it seems obvious that it would be plentiful in biological systems. However, the acquisition of iron is not a simple process due to the fact that the vast majority of it is unavailable. Therefore, one matter of great importance to life is effective methods to obtain necessary iron in order to support the many systems which require it.

The first problem is that the vast majority of iron in the environment is in the form of insoluble iron oxides. On the front lines of this problem are bacteria which produce and secrete low molecular weight chelators, the siderophores. These molecules have extraordinarily high affinity for iron and thus are able to scavenge it from the extracellular environment. The Fe-siderophore complex is then drawn into the cell via a series of iron transport proteins.

The source of iron for plants comes from absorption of iron compounds through the root systems and eventual storage in the leaves, which accounts for a good portion of ingested animal iron. Among vegetation, darker leafy green vegetables store higher iron concentrations. In animals, the mononuclear iron centered protein, transferrin, is the

major iron transporting protein in the body, transferring iron across cellular membranes through interaction with transferrin receptor. About 0.1% of transferrin bound iron circulates in the plasma as an exchangeable pool. The transferrin protein serves three purposes: it renders iron soluble under physiological conditions, it prevents iron-mediated free radical toxicity, and it facilitates transport into cells.

By far, the vast majority of the body's iron is found in the form of heme (~70%) which is compartmentalized throughout the body. Oxygen transporting proteins, such as hemoglobin (~60%) and myoglobin (~10 %) hold the bulk of heme iron. A small fraction, less than 5%, of heme is found in other proteins devoted to cellular processes such as cytochromes and peroxidases (6). The remainder of the ~4 g of non-heme, ferric iron is found bound to transfer and storage proteins. Approximately 30% of body iron is chelated to ferritin molecules thus rendering it nontoxic. The majority of this stored ferritin-bound iron is found in the liver hepatocytes but is also found in the bone marrow, spleen and the slightest in blood serum (8-10 mg) (7).

A number of pathogenic bacteria are faced with the need to draw iron from a host rather than the external environment. Prospects here are not much improved. Because iron is so difficult to come by, as seen above, and because the consequences of uncontrolled interactions between iron and O₂ (e.g. •OH) are grave, the vast majority of iron in a biological system is sequestered in proteins. Indeed, at physiological pH (7.4), the concentration of free aqueous iron is about 10⁻¹⁸ M.

Consequently, the methods of pathogenic bacteria for acquiring iron capitalize on the iron sources present in their target host, heme in particular. At the center of these pathogenic bacterial iron acquisition systems are the heme receptor proteins. These

proteins are unique to pathogens and are specialized for transferring heme iron into the cell for use as an iron source. This ability is particularly advantageous for bacteria in mammalian hosts where the majority of iron is bound, such as in hemoglobin, and this ability is expected to play a part in their greater virulence than their non-pathogenic counterparts. This is addressed more fully in a future section.

Iron cofactors: Structure as a functional control

Upon acquisition of iron, there remains the problem of tuning its properties, controlling its interactions (especially with O₂ and related species), and directing its reactivity toward desired outcomes. All this is done while simultaneously minimizing the generation and persistence of damaging side products. That this is successfully accomplished is amply demonstrated by the truly broad diversity of biological reactions that rely on iron-dependant proteins.

The control of iron's catalytic capabilities is exerted at three levels of structure. First is the immediate environment of the iron, referred to as the cofactor level. Beyond this, the iron cofactor has an immediate active site context provided by the protein that carries it, and finally, the reactivity of the iron cofactor can be regulated by protein structures peripheral to and even a great distance from the active site.

At the first level, the nature of the cofactor dictates much regarding the coordination environment and oxidation state of iron and, therefore, its reactivity. Iron cofactors can be divided into three broad categories: monofunctional and dinuclear iron

centers, iron-sulfur clusters, and hemes. Iron cofactors are utilized to maximize the catalytic potential of iron be it in reactions directly affecting molecular oxygen or not. Within each category these cofactors can be individually tuned by their protein environments to achieve unique catalytic properties.

Mononuclear and dinuclear iron proteins

Cofactors with mononuclear and dinuclear iron centers cover a range of catalytic processes from aromatic oxidation and cellular iron transport to the post translational modification of proteins. As the name implies, these cofactors contain either one or two iron atoms, respectively. Mononuclear iron centers are either in the Fe (II) or Fe (III) resting state. They appear in important biological processes including antioxidant defense mechanisms and biosynthesis of lipid-based hormones. For example, Fe-SOD and lipoxygenase are both mononuclear iron enzymes. Yet, the reaction carried out by each is quite different, indicating that the protein environment of each cofactor is critical to its prescribed function.

Iron superoxide dismutase

Iron superoxide dismutase (Fe-SOD) contains a single iron atom in its active site which is used to catalyze the redox disproportionation or dismutation of two moles of

superoxide radical ($O_2^{\bullet-}$) to H_2O_2 and molecular oxygen. This serves to very rapidly dispose of $O_2^{\bullet-}$, preventing its toxic effects and helping to prevent formation of the far more destructive $\bullet OH$. This points out the fact that though the iron/ O_2 interaction is tightly controlled in biology, there is still a need for mechanisms to eliminate not only accidentally released reactive oxygen species but also those produced purposefully by some organisms to kill others. Iron and other transition metals are necessary for these processes as well.

The key to Fe-SOD catalysis is in the way it facilitates proton-coupled electron transfer. Fe-SOD activity requires that the reduction potential of the iron is tuned between the reduction potential for $O_2/O_2^{\bullet-}$ and $O_2^{\bullet-} + 2H^+/H_2O_2$. This is intimately tied to the H-bonding environment of the single axial water ligand to the iron center [Figure 3]. The trigonal bipyramidal coordination sphere also includes three strictly conserved histidines (one axial and two equatorial) and a single equatorial aspartate.

The reduction potential of wild-type Fe-SOD is +20mV, due in large part to the interaction of the coordinated solvent with Glutamine 69. Because of the H-bonding network around this residue, Gln69 is constrained to act as an H-bond donor to the coordinated solvent, ultimately stabilizing the ferric state. Upon substitution of Gln69 with His the H-bonding constraints are relaxed such that the His becomes a weak H-bond acceptor and there is a resulting increase in potential to +270 mV. Substitution of Gln69 with Glu introduces a strong H-bond acceptor, and the reduction potential increases dramatically to > +690mV. Indeed, Gln69Glu Fe-SOD is isolated in the ferrous state as

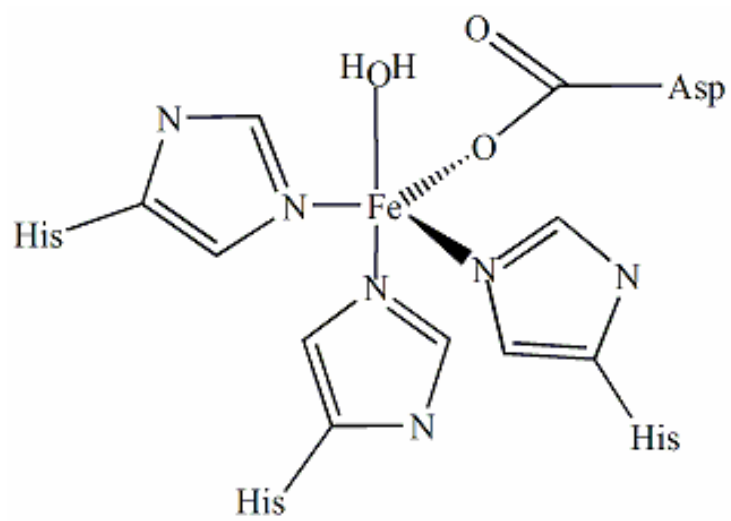


Figure 3. Diagram of the active site ligands of Fe-SOD.

opposed to wt enzyme that is isolated as ferric enzyme. Along with these changes in potential are substantial decreases in enzymatic activity.

Clearly, a critical part of Fe-SOD function is not only the coordination of the iron atom, but also the H-bonded environment around the iron and its ligands. This implicates the overall structure of the Fe-SOD protein as the scaffold that ensures the productive active site conformation, but the specific mechanisms in play and how they may be manipulated are not well understood.

Lipoxygenases

In contrast to Fe-SOD is lipoxygenase. Lipoxygenase (LO) catalyses the regio- and stereospecific hydroperoxidation of *cis*, *cis*-1,4-pentadiene fatty acids to the 1-hydroperoxy-*trans*,*cis*-2,4-diene [Figure 4] (8). This is the first step in the biosynthesis of products for diverse cellular functions such as second messenger molecules (9).

Lipoxygenases are grouped into two families; the plant family which metabolizes linolenic and linoleic acids, and a mammalian family which metabolizes arachidonic acid.

All LOs are two domain monomers composed of a large (692 amino acids) C-terminal domain and a more compact N-terminal domain (146 amino acids) which when separated render the protein completely inactive. The N-terminal domain is an eight-stranded β -barrel domain that is theorized to be important in molecular recognition. The C-terminal domain consists a 20 to 23 α -helix bundle and contains the active site as well

the substrate binding cavity. The larger plant LO, ~94 kDa, has two substrate binding cavities, the first of which has been identified as a dioxygen channel and the other, cavity II, is suggested to be important in the positional specificity of substrates. Mammalian LOs have only one, boot-shaped cavity which is analogous to cavity II in the plant family (10).

The key structural features of LOs can be gleaned from the solved structures for soybean LO-1 and LO-3, the plant family, and rabbit 15-RLO representing the mammalian family. Though the overall sequence homology of all known lipoxygenases is only 25%, there are clear regions of conservation, the active site being one of them (11). All LOs ligate the active site iron with three conserved histidines, a carboxyl group from the C-terminal amino acid residue (usually an isoleucine), and a water molecule, (12, 13). In this respect, the ligation environment is quite similar to Fe-SOD [Figure 5]. Interestingly, the reduction potential of the bound iron appears to be quite high (~ + 600mV), owing to its relatively positive coordination environment. Such a high potential suggests a coordinated H₂O as opposed to OH⁻, which is reasonable based on the hydrophobic environment of the bound water. In this case, the high reduction potential of the iron is an essential part of catalysis as the first and rate limiting step of a two step process, involves the abstraction of the double allylic hydrogen, producing a bisallylic substrate radical which rearranges and the readily reacts with molecular oxygen resulting in the corresponding peroxy radical species [Figure 6] (11, 14). In the final step, the peroxy radical oxidizes the iron center to yield the hydroperoxide product.

The catalytic strategy of lipoxygenase is interesting in that it activates the substrate by converting it to a radical for reaction with ³O₂. This necessitates the high

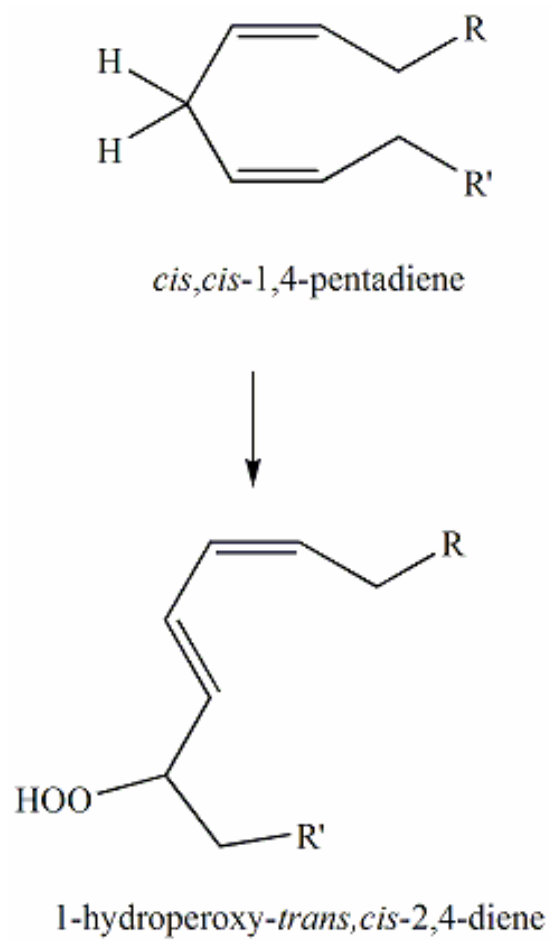


Figure 4. Overall reaction scheme for lipoxygenase catalysis.

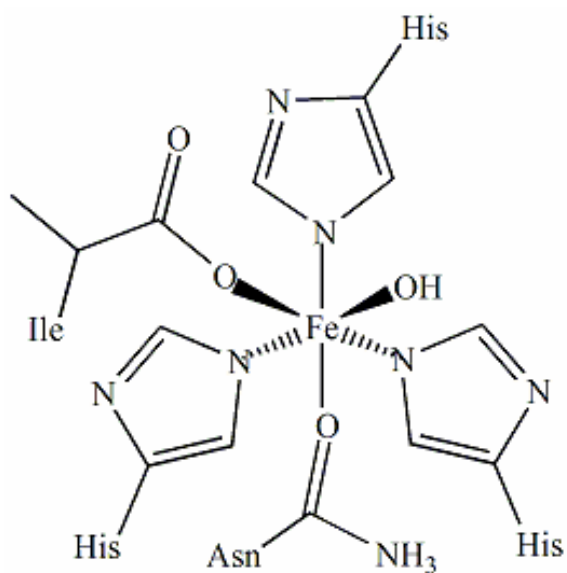


Figure 5. Diagram of ligands in lipoxygenase active site.

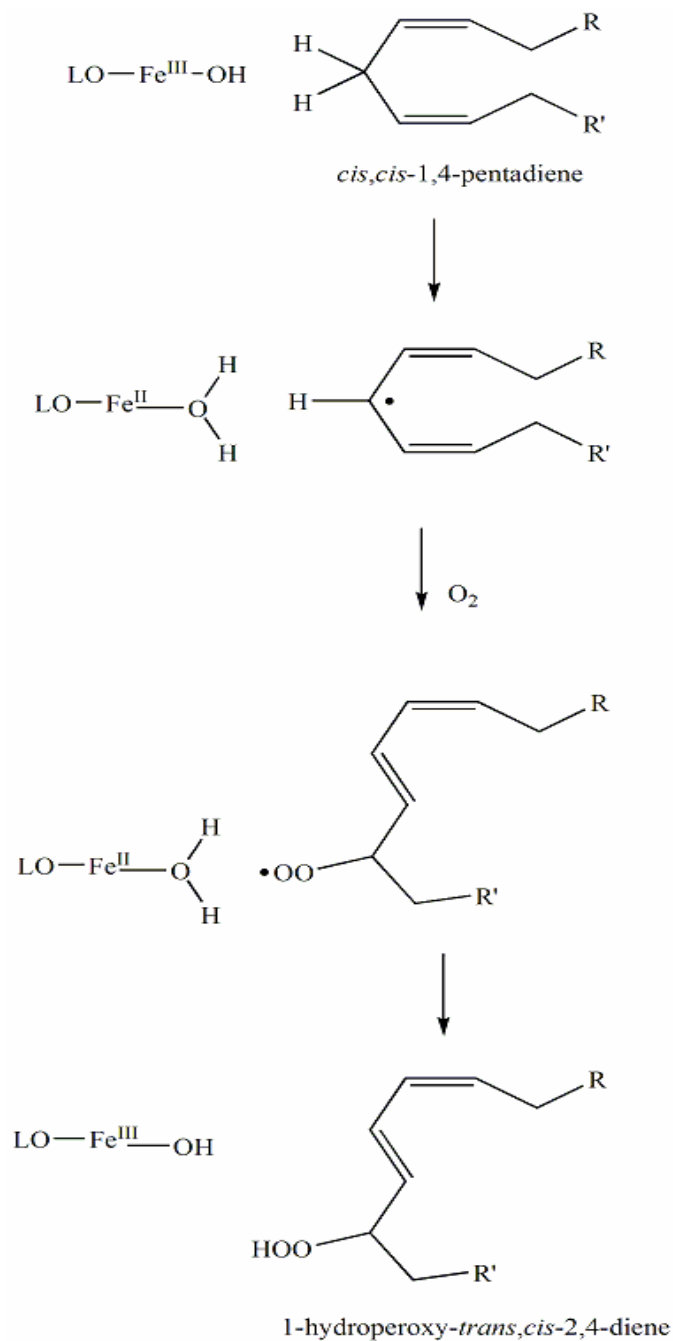


Figure 6. Proposed reaction mechanism for lipoxigenase.

potential for hydrogen atom abstraction from a bisallylic carbon(11). Other iron enzymes activate oxygen, often by way of a ferryl-oxo ($\text{Fe}^{\text{IV}}=\text{O}$) intermediate. In these instances a much more negative coordination environment is expected so that the high valent iron species can be stabilized within the enzyme active site. This coincides with a much lower reduction potential for such proteins.

While the asparagine is included as an axial ligand, it is regarded as a ‘weak ligand’ due to the fact that of the SLO crystal structures the asparagine-iron distances are determined at about 3.1 Å at a resolution of 2.6 Å , 3.2 Å at a resolution of 1.4 Å (15) and ~3 Å at a resolution of 2.6 Å . These distances are generally considered non-bonding which is in agreement with the three groups’ deductions. Also, an exogenous water molecule, at 2.6 Å from the metal center, provides the other axial iron ligand across from the Asn (15). While this interaction would be generally considered questionable above ~ 2 Å resolution, the improved resolution of Minor’s SLO-1 structure supports this proposition (15). Two other interesting characteristics include the observation by Strzypczak-Jankun et al. that based on the structure the exogenous water and Asn ligands could in fact move to bonding distance without disrupting to their environment . The other observation is that Asn is connected to the substrate binding cavity by an extensive hydrogen bonding chain from itself to a leucine residue of cavity II (15).

An interesting feature of lipoxygenase is a connection between external events involved in the reaction and changes in iron coordination. The plant LOs all indicate a possible Asn sixth ligand to the iron. However, the iron–Asn (O) distance and relative orientation varies from one structure to another. Based on DFT calculations, Lehvert and Solomon show that relaxing restrictions on the orientation of Asn 694 allows it to weakly

coordinate the iron (11). The orientation of Asn 694 is restricted in LO by hydrogen bonded interactions that extend to cavity II where the fatty acid substrate binds. It is proposed that substrate binding alters the hydrogen bonding of Asn 694, allowing it to coordinate the iron (15). This is consistent with the observation that substrate binding coincides with a shift to a hexacoordinate iron species in plant LOs.

Methane monooxygenase

Methane monooxygenase, MMO, is a diiron-centered protein included in the large ribonucleotide reductase, RNR, R2 subunit protein superfamily. MMO is found in methanotropic bacteria which use methane for their primary carbon source (16).

Methane monooxygenase uses oxygen to convert methane to methanol and water.



Two varieties of MMO have been identified; the membrane bound particulate monooxygenase, pMMO, the soluble form of monooxygenase, sMMO. Of the two the most characterized is sMMO and, therefore, will be described here. The sMMO system is composed of three proteins; including a dimeric $\alpha_2\beta_2\gamma_2$ subunit hydroxylase, (MMOH), a coupling protein which functions as a regulator enzyme (MMOB), an iron-sulfur flavoprotein reductase (MMOR), which uses two electrons from NADH to recycle MMOH back to its active form prior to catalysis. The active site, with the diiron center, is located in the large α subunit (or R2) of MMOH, where catalysis occurs (17).

In the first step the diferric resting state, MMOH_{ox} , is reduced to diferrous MMOH_{red} by two electrons transferred by MMOR from NADH (16, 18) [Figure 7]. This step includes the coupling protein MMOB which couples the reaction to the electron transfer from NADH to MMOH (19). The diferrous center reacts with O_2 , in the presence of MMOB, resulting three sequential transient states; first an oxygen-activated diferrous- O_2 complex, compound O, next putative Michaelis, bridging peroxo-diferric complex, compound P*, is proposed to form, followed by a diferric hydroperoxo complex, compound P, which accepts two hydrogen ions, releases H_2O and results in the reactive high valence diferryl compound, compound Q (20). In the reaction of MMOH_{red} with O_2 , the presence of MMOB accelerates the rate of reaction by 1000-fold (21). From here the kinetically competent compound Q abstracts a hydrogen from the substrate, forming a hypothetical mixed valent (FeIII/FeIV) radical-substrate adduct, compound R, which supposedly decays rapidly to form a diferric enzyme-product complex, compound T. Compound T decays yielding the methanol product and binds an active site H_2O molecule thereby regenerating MMOH_{ox} (16, 22). At this point the methanol may be converted to formaldehyde or carbon dioxide by other enzymes. In this system MMOH must activate oxygen in order to form a high-valent iron (IV) intermediate needed to hydroxylate the unactivated C-H bond of methane (20).

X-ray crystallographic structures of MMO reveal that the diiron center is buried deep within the core of the protein and is ligated by two histidine and four glutamic acid residues from the surrounding α -helicies (23) [Figure 8]. This highly anionic active site is well suited to provide stability for the high-valent iron intermediate,

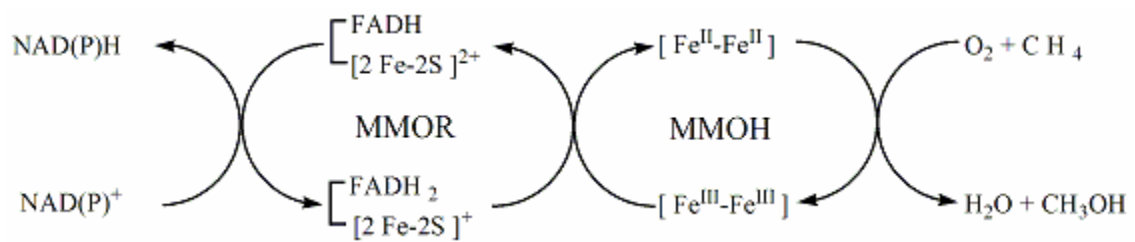


Figure 7. Reaction scheme for methane monooxygenase.

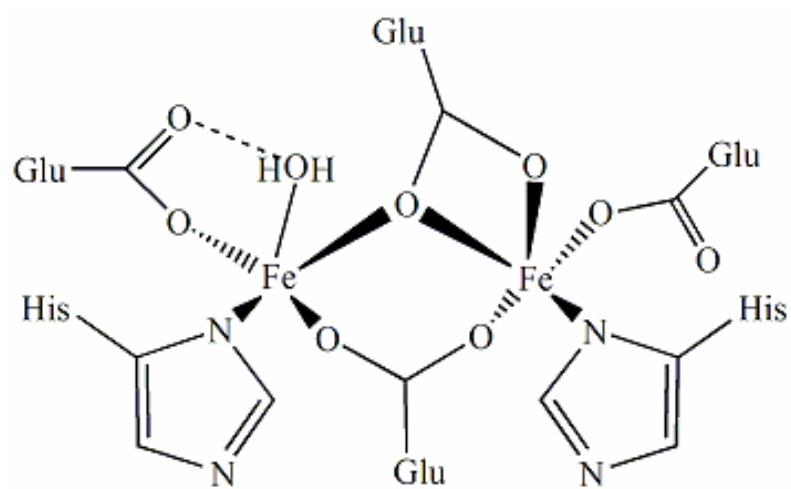


Figure 8. Diagram of ligands in methane monooxygenase active site.

compound Q. On one face of the diiron center is a hydrophobic pocket open to the two irons which is favorable for substrate binding. Comparison of MMO and its structurally similar ribonucleotide reductase R2 subunit family member has revealed mechanistically important features of the active site structure. For instance the flexibility of several of the iron bound glutamates allows for adjustability of their iron coordination environment while also allowing for water/hydroxyl/oxo ligand change, depending upon the environment. These two adaptations allow for great variety in the iron-iron bond distance maximizing both substrate and oxo-species coordination sites. While the hydrophobic face still provides the most probable substrate binding site, the high degree of flexibility is important to allow for controlling the generation of reactive oxygen species (23).

Iron-sulfur proteins

Iron-sulfur clusters are assemblies of iron, inorganic sulfide and cysteine. These cofactors are necessary for numerous functions such as electron transport, catalytic reaction centers (redox and non-redox), regulation, and protein structure stabilization. The iron and sulfur atoms are present in a roughly tetrahedral coordination environment within the proteins that use them. The general denotation for a cluster group is Fe-S found in groupings such as [1Fe-4S], [2Fe-2S], [3Fe-4S], and [4Fe-4S] [Figure 9].

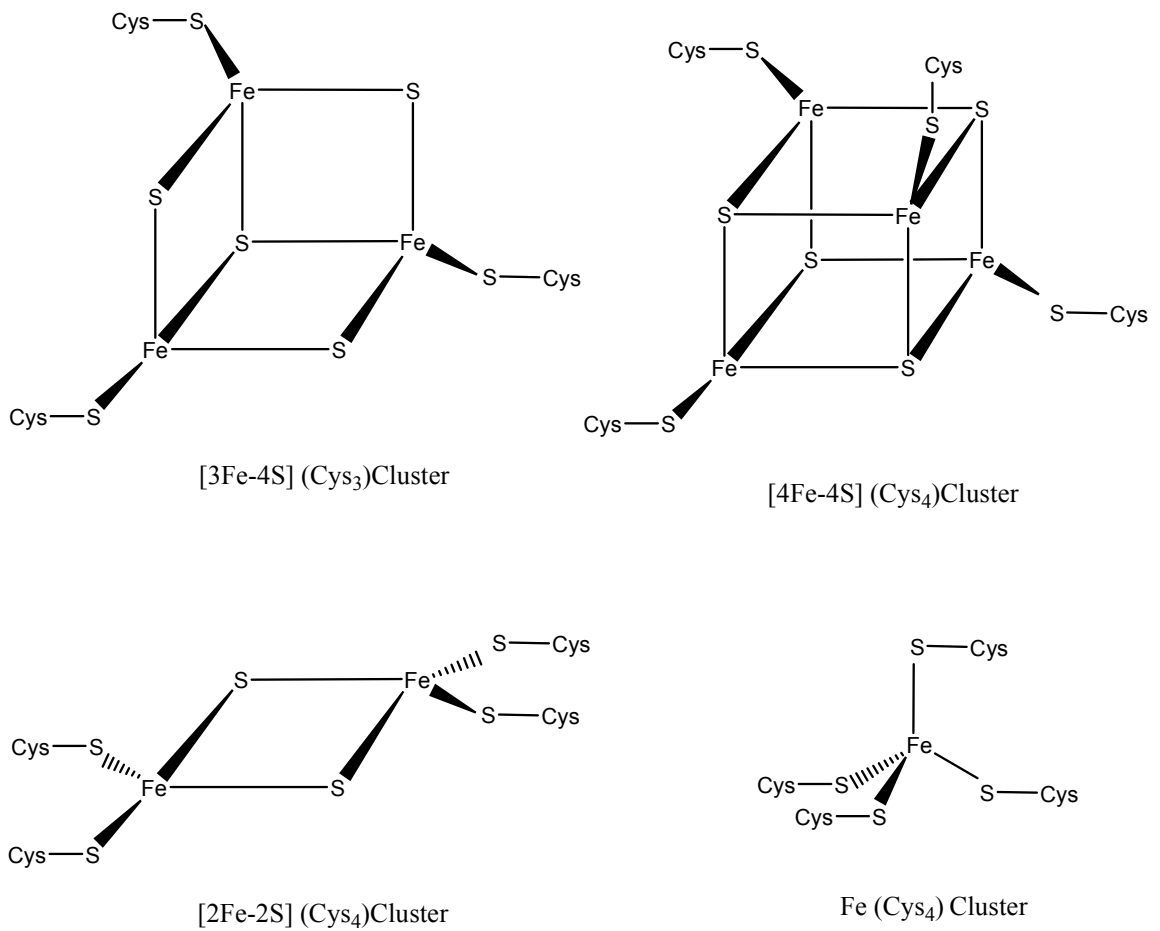


Figure 9. Diagram of iron-sulfur cluster types.

Feridoxin and rubredoxin

Two Fe-S cluster cofactors, ferridoxin and rubredoxin, are small molecular weight electron transport proteins found in green plants and photosynthetic green sulfur metabolizing bacteria. Ferridoxin can have a [2Fe-2S], [3Fe-4S] or [4Fe-4S] center with a general core structure consisting of 1 α -helix and 4 β -strands in a beta(2)-alpha-beta(2) configuration. Ferridoxin is a multifunctional electron carrier which transports, in diverse redox systems, via a [2Fe-2S] center with the sulfur contribution being from acid-labile sulfurs of conserved cysteines [Figure 10]. The iron atoms transfer electrons via a one-electron valence changes from Fe^{3+} to Fe^{2+} .

The protein rubredoxin can act in place of ferridoxin in electron transport. From three dimensional structures it has been determined that the protein is composed of 2 α -helices and 2-3 β -strands (24). The active site of rubredoxin is a Fe-(Cys) S_4 center, or an iron coordinated by four sulfurs contributed by conserved cysteine residues located on two conserved active site loops, which forms an almost regular tetrahedral [Figure 10]. The iron center undergoes electron [Fe(Cys) $_4$] $^{1-}$; [Fe(Cys) $_4$] $^{2-}$ Additional highly conserved residues include a ring of highly acidic amino acids within the active site which are suggested to function in recognition by partnering electron acceptor .

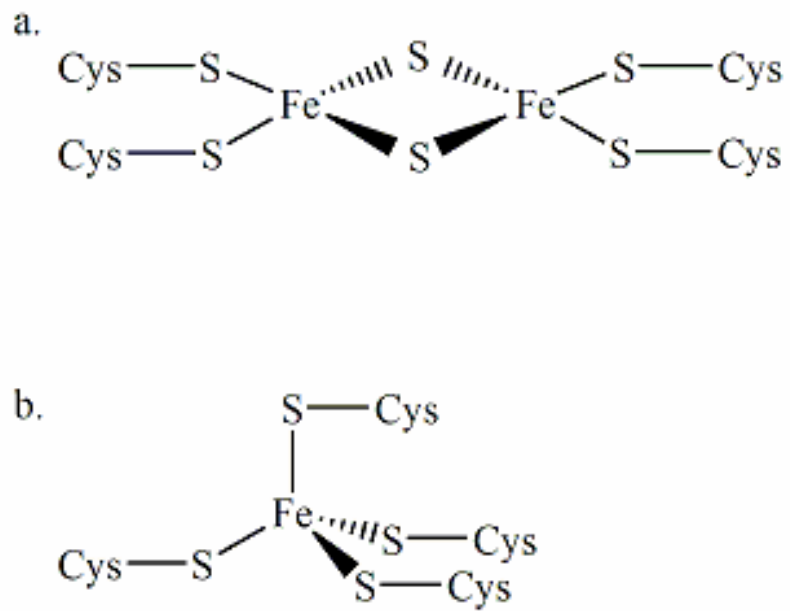


Figure 10. Diagrams of ferridoxin [2Fe-2S] (a) and rubredoxin (b) active site ligands.

Aconitase

In stark contrast is another iron-sulfur protein found in the tricarboxylic acid cycle, aconitase or aconitate hydratase. Aconitase catalyses the reversible, stereospecific isomerization of citrate to isocitrate in a *non-redox* reaction (25) [Figure 11]. Unlike the majority of iron-sulfur proteins the Fe-S cluster of aconitase interacts directly with the substrate. Indeed, instead of the typical coordination of one of the irons by three sulfides and a cysteine thiolate, the thiolate is replaced by a water ligand [Figure 12]. In the first step of the aconitase reaction, citrate binds chelating this iron with the C3 hydroxyl and carboxyl groups. Citrate is dehydrated to produce *cis*-aconitate. *Cis*-aconitate then “flips”, recoordinates, and is hydrated to produce 2R, 3S isocitrate.

The 3D structure of mitochondrial aconitase has been determined and has been shown to contain four α/β domains: the three N-terminal domains have similar folds, each with three $\alpha\beta\alpha$ layers, while the C-terminal domain forms a β -barrel (26). The active form of aconitase contains single cubane like $[\text{Fe}_4\text{S}_4]^{2+}$ cluster, which is converted to an inactive $[\text{Fe}_3\text{S}_4]^+$ form in iron-limiting conditions. Crystal structural data for aconitase reveals three Cys residues have been shown to be ligands of the $[\text{Fe}_4\text{S}_4]$ center (27). In the active state, the labile iron ion of the $[\text{Fe}_4\text{S}_4]$ cluster is not coordinated by Cys but by water molecules. In eukaryotes two isozymes of aconitase are known to exist: one found in the mitochondrial matrix and the other in the cytoplasm. Interestingly, aconitase is a member of the aconitase/ IREBP family which also includes the iron-responsive element binding protein (IREBP). Aconitase, as an IREBP, upregulates genes

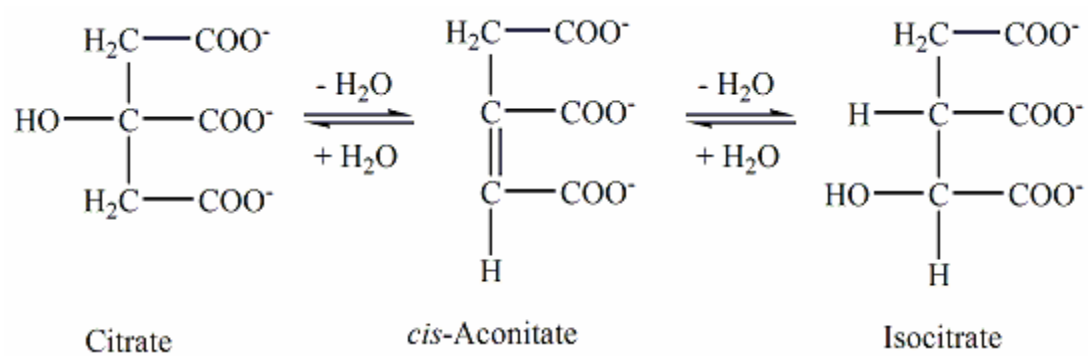


Figure 11. Reaction scheme for the reversible transformation of citrate to isocitrate catalyzed by aconitase.

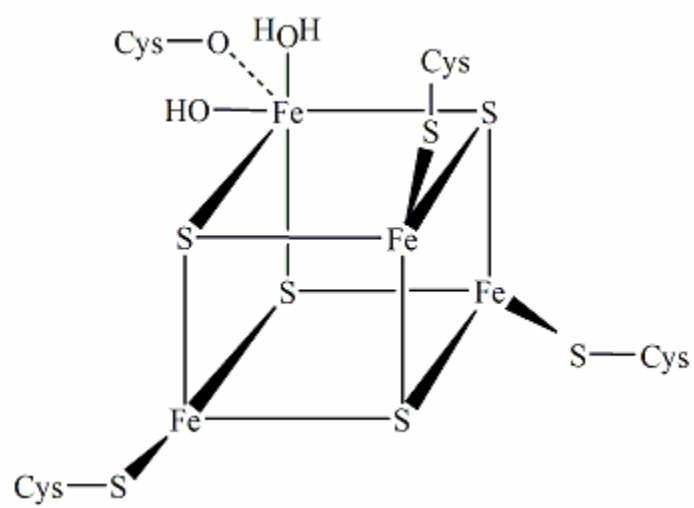


Figure 12. Diagram of the active site ligands of aconitase.

responsible for obtaining more iron when under iron-limiting conditions or in response to iron loss.

Heme Protein Structure and Function: Strategies for controlling the iron oxygen interaction in biology.

Heme cofactors

The third class of iron-containing cofactors is the hemes, which are used by hemoproteins and hemoenzymes. Heme cofactors contain a single iron atom bound to a porphyrin ring system. The porphyrin is the most common ring system of the tetrapyrrole cofactors and is found in three major forms: heme a, b and c [Figure 13]. Heme, in one of these three forms, is widely distributed throughout the biological world both in terms of the organisms that rely on its versatile chemical properties and in terms of the processes that are dependent upon it. Heme is a critical component of O₂ sensing, transport, storage and delivery mechanisms. It is integral to multiple electron transfer processes, including respiratory electron transport and neutrophil oxidative burst. Heme is utilized for the synthesis, sensing, storage, and decomposition of nitric oxide. Heme is required to synthesize all of the prostaglandin hormones and thromboxane. Without heme, plants cannot make the structural polymer lignin. Likewise, the unique ability of so-called white rot fungi to degrade lignin requires heme. Heme is central to

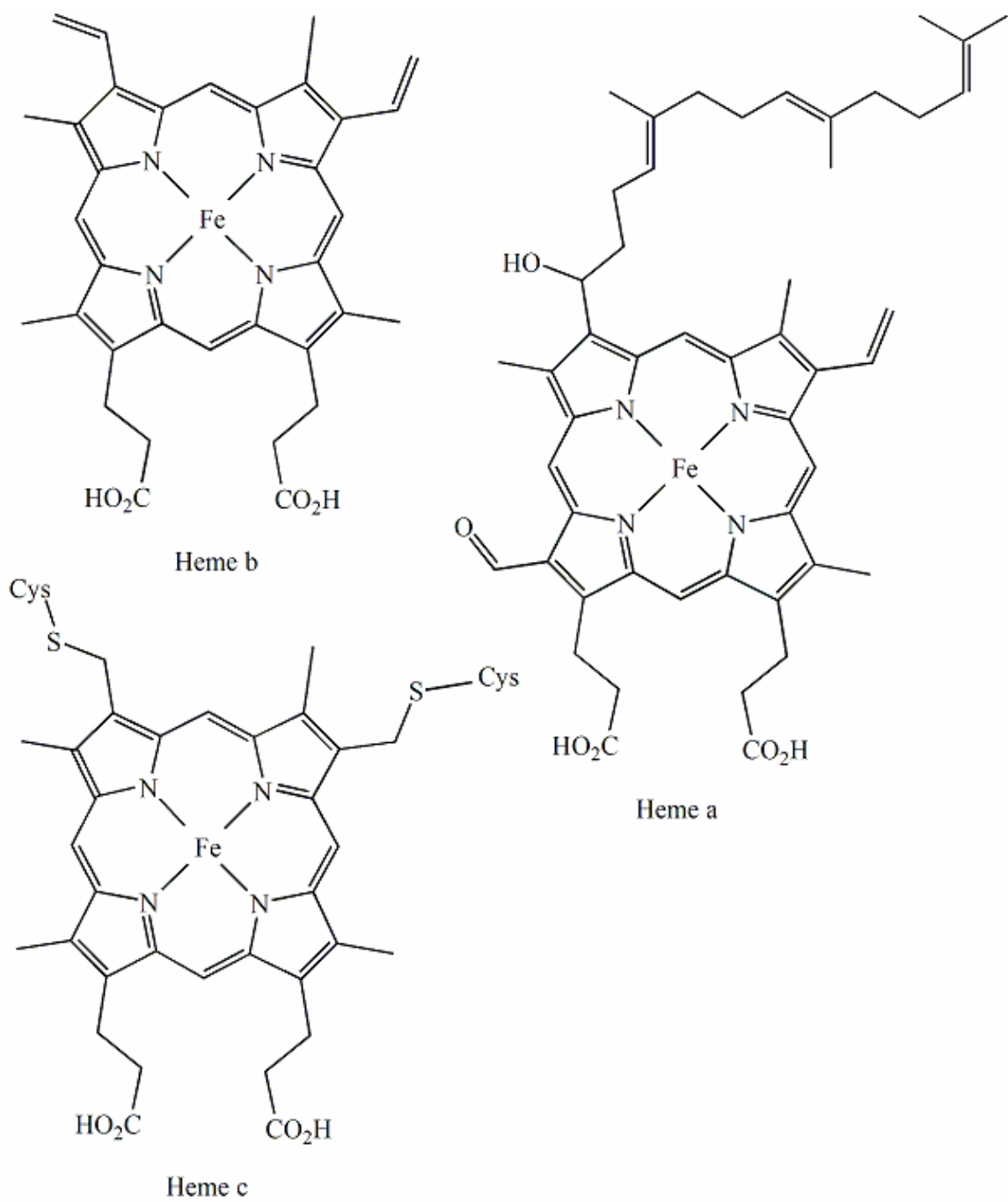


Figure 13. Structures of heme cofactors; heme a, heme b, and heme c.

detoxification and excretion of xenobiotic compounds and to the disposal of reactive oxygen species.

Clearly, heme participates in a variety of biochemical processes. This can be accounted for, to a very limited degree, by the fact that there are different heme cofactor structures. For example, heme is found toward the end of the respiratory electron transport chain because it has a relatively high reduction potential. This is due to the electron withdrawing formyl and isoprenoid side chains that replace a methyl and vinyl group, respectively, in heme b. The covalent attachment of the porphyrin macrocycle to protein as observed in the heme c-bearing proteins and enzymes may impart additional stability. Likewise, it is proposed that covalent adduction of heme b to mammalian peroxidase proteins by esterification may prevent heme modification during halide oxidation [Figure 14].

Nevertheless, it is clear that this small range of variability in cofactor structure can in no way cover the range of function that living organisms depend on heme to deliver. As with the other iron-containing cofactors, it is the protein surrounding the prosthetic group that will dictate the chemical properties of the heme, and therefore, its function in a biological context. Protein structure addresses heme function at three basic levels. First is in the identity of the axial ligands (one or two) supplied to the iron by the protein. Second is the immediate active site environment of the heme and its ligands. Third is the overall structure of the protein, including features that may be a considerable distance from the active site. This can be demonstrated by comparing several examples

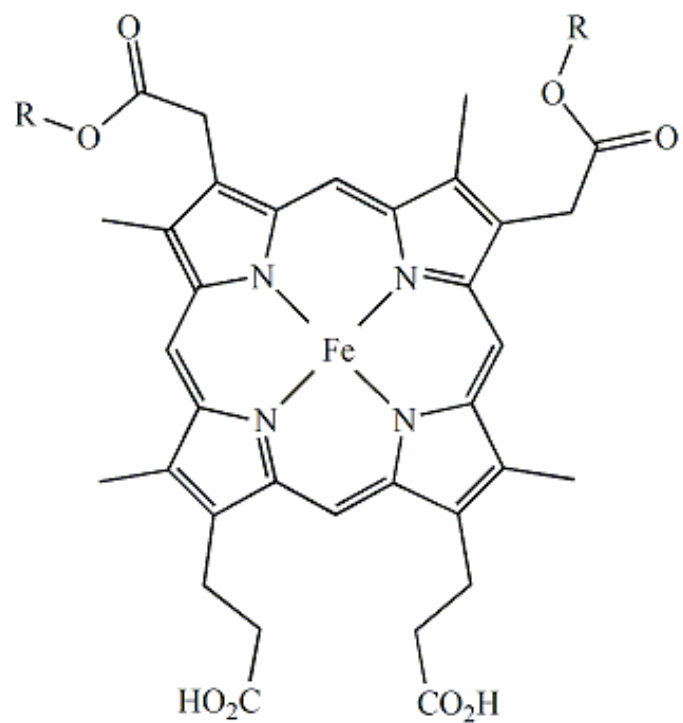


Figure 14. Esterified heme b in peroxidases forming a covalent adduct with the protein, represented by (R).

of dependent proteins with widely varied functional capabilities. It is important to bear in mind that each of these examples uses *the same* prosthetic group, heme b.

Cytochromes P450 and Nitric Oxide Synthases

The cytochromes P450, are important enzymes found in all major classes of organisms. They contain only one heme per polypeptide chain which is buried deeply within the protein. Thus, the P450 controls what substrates may enter the active site and gain access to the highly reactive ferryl-oxo intermediate of the catalytic cycle. As this intermediate is able to hydroxylate, epoxidize, or otherwise insert an oxygen atom into virtually any unactivated organic compound, the stringency of this level of control has great bearing on the biological ramifications of these enzymes. In addition to O₂ and substrate, P450 catalysis requires the input of two electrons from NAD(P)H. This is accompanied by the action of a separate protein, NAD(P)H: cytochrome P450 oxidoreductase. So, what molecules P450 oxygenates depends on the substrates that gain access to the active site, and the timing of their oxygenation depends on the delivery of electrons by P450 reductase.

In humans, P450s are involved in many vital functions. Among these are arachidonate oxygenation to generate biologically active lipid messengers, maturation of steroid hormones, and ω oxidation of fatty acids. Lipophilic xenobiotics (including drugs) are transformed by P450s to more water soluble products. This is to facilitate their excretion in the urine. However, an unfortunate downfall of this latter role of P450s is

that oxygenation of compounds like benzo(a)pyrene activates them to produce potentially mutagenic lesions in DNA and damage other important biological molecules. In bacteria, the P450s provide access to unusual carbon sources for growth by catalyzing the first step in their metabolism, usually hydroxylation. Morpholine, piperidine, and camphor are examples of such carbon sources. Indeed, the most thoroughly studied P450, P450 cam, is expressed at high levels in *Pseudomonas putida* when grown on media containing camphor.

The oxidative power of cytochromes P450 ultimately arises from the activation of O_2 , expressed as a ferryl-oxo/porphyrin radical intermediate, formally the +5 oxidation state. The ability to access this intermediate depends in large part on the ligation environment of the heme. The P450s capitalize on a highly conserved cysteine residue to ligate the heme in the axial position [Figure 15]. This ligation environment gives P450s unique reactivity. The catalytic cycle begins with the enzyme in its ferric state with a bound H_2O or OH^- as a sixth ligand [Figure 16]. The anionic character of the thiolate ligand contributes to the quite low reduction potential of the heme iron ($\sim -300mV$). Binding of substrate displaces the H_2O/OH^- ligand, inducing a shift in spin-state (low spin to high spin) and an increase in potential to $-170mV$. This allows the heme to accept an electron from P450 reductase and be reduced from Fe^{III} to Fe^{II} upon substrate binding. At this point, O_2 binds to the reduced heme iron and a second-electron is delivered, inducing heterolytic cleavage of the bound O_2 . The products of the cleavage are a ferryl/porphyrin radical and water. Water is released from the active site, and the ferryl-oxo radical intermediate initiates O-atom transfer, resulting in insertion of an O-

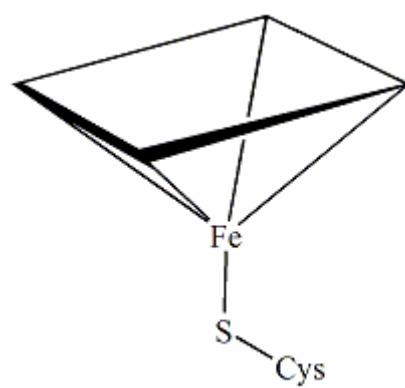


Figure 15. Diagram of active site ligands to heme in cytochrome P450.

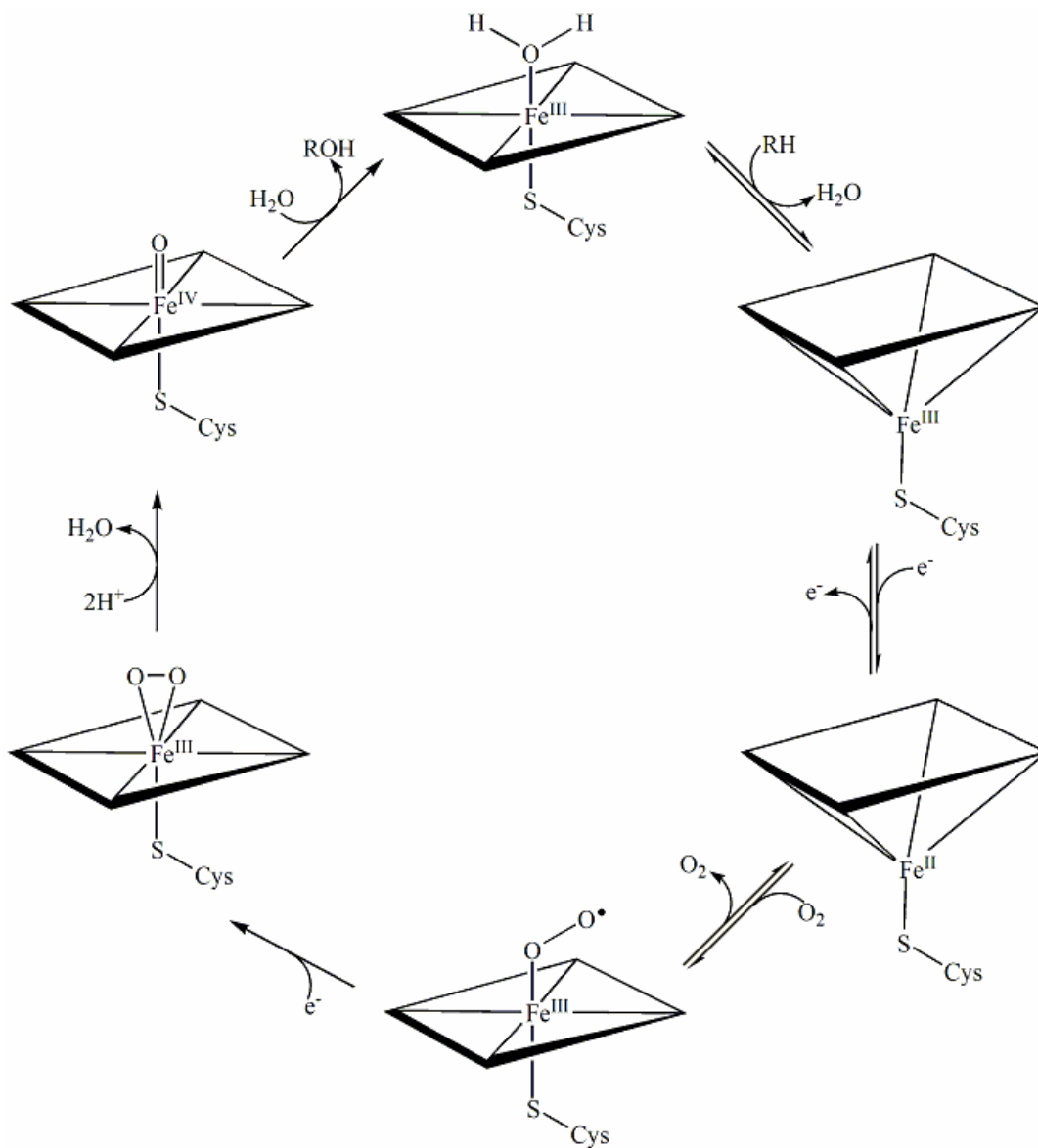


Figure 16. Reaction mechanism for cytochrome P450

atom into a specific C-H bond of the substrate. The key to the process is the ferryl-oxo/porphyrin radical intermediate. In order to stabilize such an electron deficient structure, P450s rely on the electron dense and anionic cysteine thiolate ligand. Selective oxidation of a certain position of the substrate is dictated by binding sites for the hydrocarbon substrates are in the hydrophobic domain of the polypeptide chain in the distal pocket above the heme group.

Nitric oxide synthase is a heme-dependent cofactor enzyme that functions in the production of nitric oxide (NO), a key mediator in which is necessary for smooth muscle contraction, signal transduction. Nitric oxide is a critical messenger in smooth muscle contraction and vasoregulation. It is also generated as part of the immune response of macrophages, leading to production of reactive oxynitrogen species (RONI) as a part of immune response. The RONI are produced when nitric oxide (NO) react with superoxide. Large quantities of nitric oxide are produced by the induction of inducible nitric oxide synthase (iNOS) in response to inflammation due to microbial products or autoimmune reactions. Reactive oxynitrogen intermediates include peroxynitrite ion and peroxynitrous acid [Figure 17]. The coordination environment of the heme of nitric oxide synthase is quite similar to cytochromes P450, supporting highly similar catalytic

Myoglobin and hemoglobin

Hemoglobin, myoglobin, leghemoglobin, and other oxygen storage/transport heme proteins are a stark contrast to the cytochromes P450 and nitric oxide synthases.

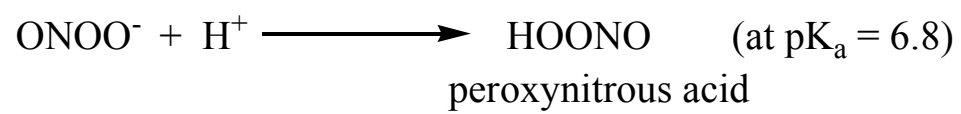
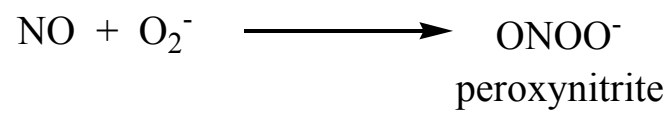


Figure 17. Reactions converting NO to peroxynitrite and subsequently peroxynitrous acid.

These proteins must bind O₂ in a relatively unreactive state and release it unmodified at the appropriate time. Accordingly, the environment of the heme group is quite different. Instead of a cysteine thiolate, the heme is ligated by a histidine imidazole [Figure 18]. The sixth coordination site is left open to accommodate binding of O₂. Importantly, the imidazole ligand is a neutral (not anionic) ligand, contributing to a considerably higher reduction potential of ~ +50mV. This simultaneously favors the ferrous oxidation state to allow for O₂ binding and disfavors higher oxidation states (e.g., Fe^{IV}=O) that would result from oxygen-oxygen bond cleavage. On the distal side, a His serves as an H-bond donor to stabilize the bound O₂ complex, which is suggested to resonate between Fe^{II}-O₂ and Fe^{III}-O₂⁻ states. Here again, the cleavage of the oxygen-oxygen bond is discouraged and oxidative chemistry is prevented.

Both hemoglobin and myoglobin possess the same active site characteristics, and so both proteins bind O₂ in a relatively unreactive state. The effect of O₂ partial pressure on the extent of O₂ binding by these two proteins, however, is markedly different. Myoglobin, a monomeric protein, binds O₂ according to a single equilibrium constant reaching saturation at high concentrations. This gives rise to a rectangular hyperbolic relationship between O₂ partial pressure and the percentage of myoglobin active sites occupied by O₂. Conversely, O₂ binding by hemoglobin (a heterotetramer) is sigmoidal, indicating positive cooperativity. This is due to structural communication of O₂ binding at one active site to the active site of another subunit and constitutes a long-distance or global protein structural influence on function. Binding of O₂ induces a shift in the ferrous iron from a high spin to low spin state. The presence of all six d electrons in the dπ orbitals contributes to the smaller radius for the metal ion, allowing it to move into the

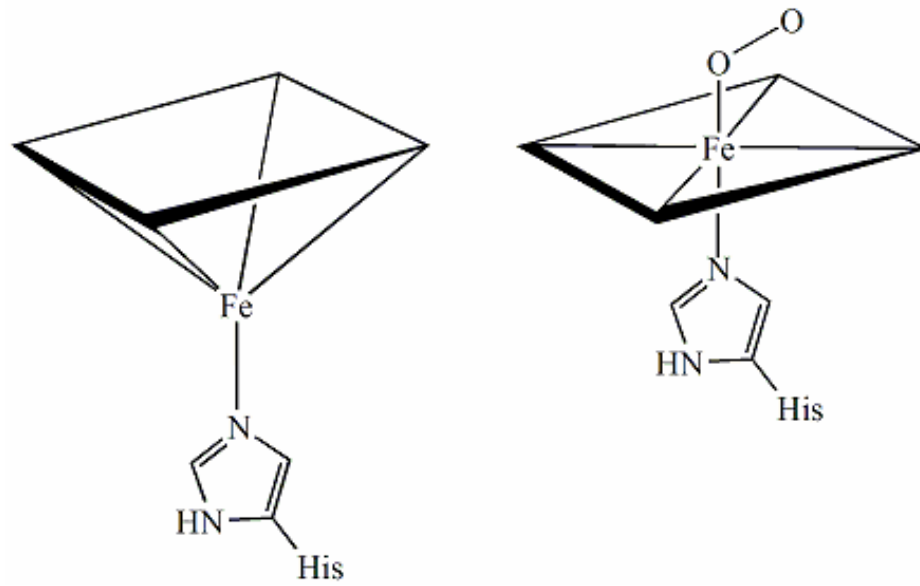


Figure 18. Diagram of active site ligands for hemoglobin and myoglobin unbound (left) and O₂ bound (right).

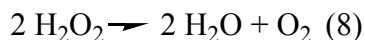
plane of the porphyrin. This 0.6 Å shift of the metal brings with it the histidine ligand, an adjacent valine residue, and the rest of the F-helix. Steric constraints induced by the shift cause an adjustment of the F-helix that is communicated to the intersubunit interface and other active sites of the tetramer. The transition to the so-called “relaxed” or R state facilitates binding of subsequent O₂ molecules.

Monofunctional catalases

Though extensive measures are taken to ensure that oxygen’s activation (and reduction) is carried out in controlled environments, the production of partially reduced oxygen species (i.e., reactive oxygen species) is inevitable in an aerobic environment. Some enzymes like glucose oxidase produce H₂O₂ from the oxidation of their substrates. In some instances misdirected electrons reduce O₂ to O₂^{•-}, which as discussed previously, is disproportionated by SOD or non-enzymatically to generate H₂O₂. The consequences of unmitigated H₂O₂ production are dire, particularly in light of its interaction with reduced transition metals (i.e., generation of •OH). Though important roles for H₂O₂ cell signaling have been identified it is still regarded as a “necessary evil”.

To address this problem, catalysts which rapidly degrade H₂O₂ are distributed widely in nature. Ironically, iron, which lies at the heart of the problem, is a critical part of the solution. Catalases which degrade H₂O₂, producing H₂O and O₂, all rely on a transition metal for activity. Most often that metal is iron and most often the iron is found in a heme prosthetic group. The reaction catalyzed by the heme-dependent

monofunctional catalases is essentially a disproportionation of two H₂O₂ molecules (reaction 8).



The active site of these enzymes is poised to rapidly accomplish the reaction. Indeed, the catalytic abilities of monofunctional catalases are unparalleled. The enzyme accomplishes this reaction at rates approaching 4×10^7 times per second. One striking feature of monofunctional catalases is the fifth ligand to the heme iron. It is the phenolate oxygen of a tyrosine side chain [Figure 19]. As in cytochromes P450, the strong anionic character of the ligand is expected to favor the more oxidized heme species. Indeed, the reduction potential of catalase heme is only known to be below -500 mV. This is consistent with the need to form high valent iron-oxo species in the catalytic cycle [Figure 20]. In this resting state, catalase is in the ferric form. Reaction with one equivalent of H₂O₂ produces the ferryl-oxo/porphyrin radical intermediate known as compound I and one equivalent of H₂O. The second step involves reduction of compound I by a second equivalent of H₂O₂ generating one equivalent of H₂O, one equivalent of O₂ and, recycling the heme back to its ferric form.

Opposite the tyrosine ligand, the distal heme cavity contains a conserved histidine residue whose orientation is unusual in that it is coplanar with the heme. An adjacent serine residue H-bonds to the N^δ of the histidine, thereby increasing its ability to act as an H-bond acceptor. An active site asparagine is also an important, though not entirely essential, residue in catalysis. It is suggested that the His and Asn residues form H-bonds with H₂O₂ in catalase compound I to facilitate its oxidation to O₂.

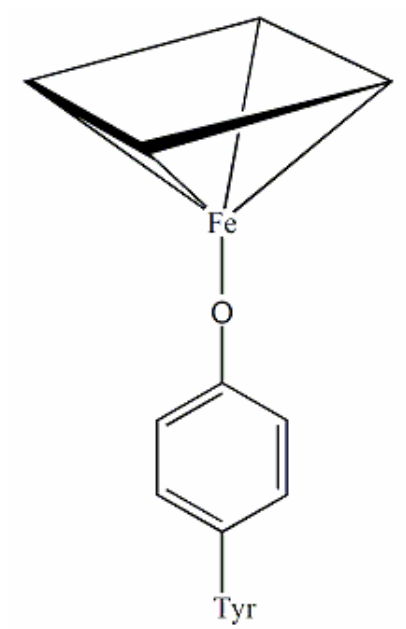


Figure 19. Diagram of active site ligands for catalases.

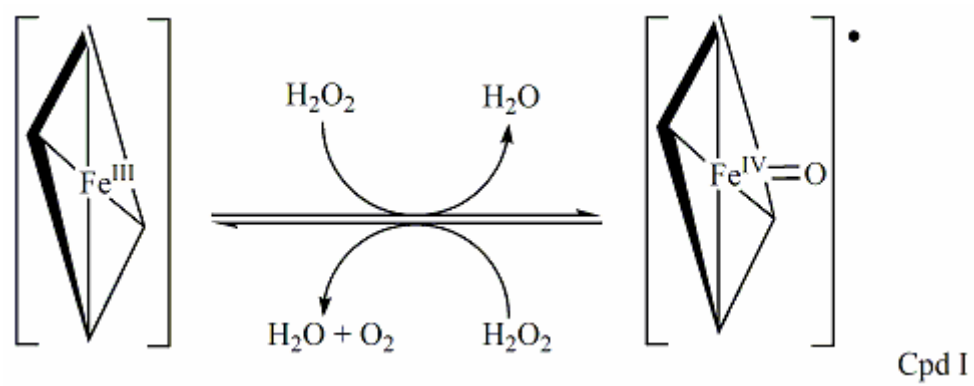
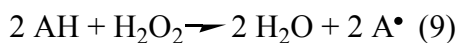


Figure 20. Catalase reaction scheme.

Monofunctional peroxidases

Like the catalases, monofunctional peroxidases also catalyze peroxide decomposition, but they do so (typically) via a three-step process utilizing an exogenous electron donor. The catalytic cycle of peroxidases is reminiscent of the catalase cycle. Indeed, the first step of the catalytic cycle is virtually identical [Figure 21]. In this step H_2O_2 is reduced by two electrons and the enzyme is concomitantly oxidized to the ferryl-oxo porphyrin radical intermediate, compound I. In some peroxidases, like cytochrome c peroxidase, the porphyrin radical is reduced by the protein to generate a ferryl-oxo protein radical intermediate analogous to compound I. In any case, compound I is then reduced, usually in two, sequential, one-electron steps. The first electron reduces the radical, leaving the ferryl-oxo intermediate compound II. Compound II is then reduced by the second electron to generate ferric enzyme, releasing a second equivalent of H_2O .



Peroxidases are found in animals, plants, molds, bacteria and other microorganisms (28). While the vast majority of peroxidases are heme-containing, some have heme derivatives as well as partial iron-porphyrins. The identity of the electron donor that converts compound I to ferric enzyme varies considerably from one peroxidase to another. In fact, most peroxidases are rather promiscuous in that the catalytic cycle can be completed by any one of a number of aromatic compounds. In many cases, the particular function of a peroxidase depends on the identity of the electron donor that is available for oxidation.

For example, some plant peroxidases such as horseradish peroxidase are essential for

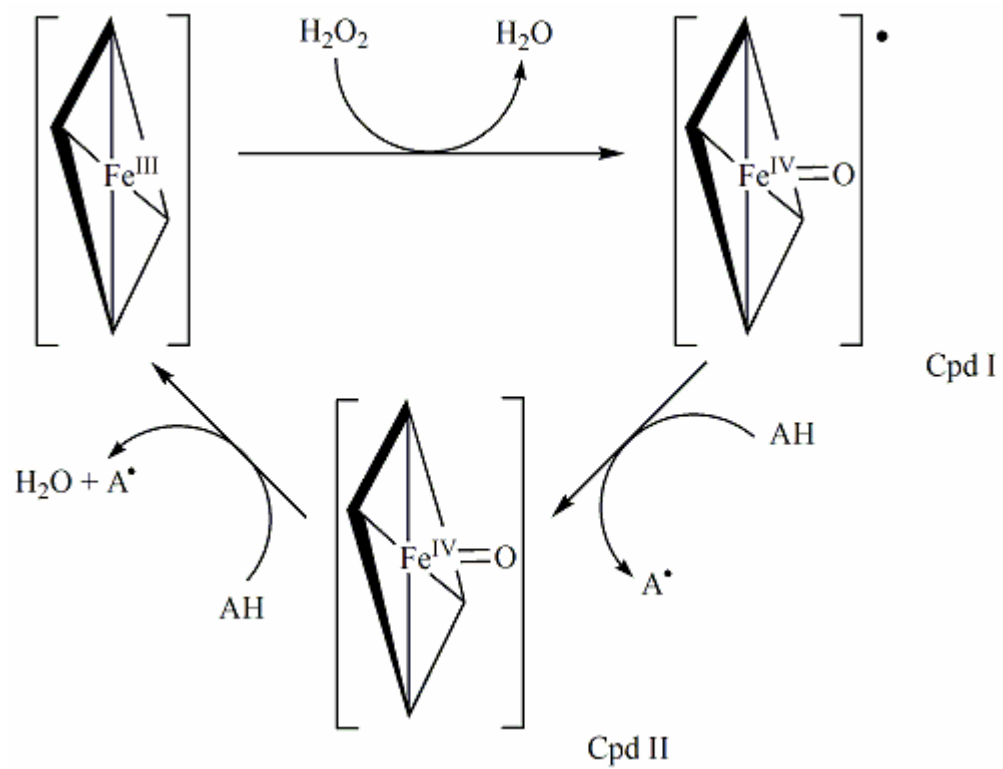


Figure 21. Reaction scheme for monofunctional peroxidases.

lignin biosynthesis. They are produced in an environment that contains the phenylpropanoid precursors of the lignin polymer. Conversely, ascorbate peroxidase, a plant peroxidase found in chloroplasts and root nodules, is needed for the rapid reduction of H_2O_2 to H_2O . The presence of high concentrations of ascorbate in these environments allows for peroxide reduction and only the generation of relatively unreactive and harmless ascorbyl radical. The ability of peroxidase to address several structurally dissimilar electron donors comes from the outer-sphere nature of the electron transfer. These donors associate at or near the heme edge, which is solvent exposed in most peroxidases. This allows for a close approach to the heme and concordant rapid electron transfer. At the same time, there is little in the way of binding stringency, allowing many compounds to fill the role of electron donor in the catalytic cycle.

This is in stark contrast to the monofunctional catalases, which as discussed before, have the heme buried deeply within the protein. The only access to the heme is through a long, narrow, hydrophobic channel. This prevents the close approach of typical peroxidase electron donors to the heme edge, and as a result, electron transfer from these donors to the ferryl-oxo intermediates of catalases is very slow. This is one reason why catalases use H_2O_2 as a donor to complete their catalytic cycles and have very poor peroxidase activity.

In terms of ligands, to the heme iron, the peroxidases are also very distinct from catalases. Most peroxidases are pentacoordinate or hexacoordinate high-spin heme proteins or some mixture of both. The fifth ligand to the heme is almost always the ϵ -nitrogen of a histidine imidazole ring [Figure 22]. The sixth coordination site is often

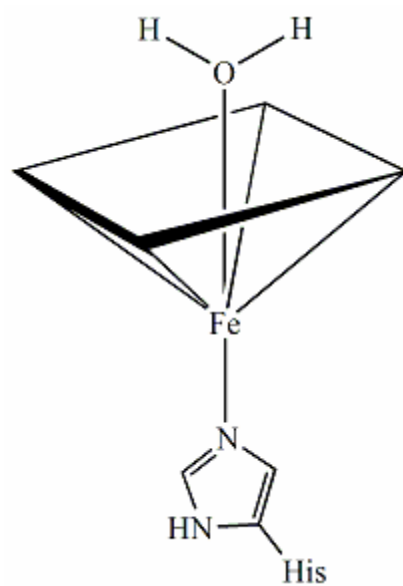


Figure 22. Diagram of active site ligands for monofunctional peroxidases.

unoccupied, or sometimes, an H₂O is weakly coordinated. In this respect, the coordination environment of the heme in peroxidases is much closer to the globins than catalase.

This raises an interesting question. What structural features distinguish peroxidases, which are isolated in their ferric state and catalyze H₂O₂ decomposition through ferryl-oxo intermediates, from the globins, which are isolated as ferrous heme proteins and must be “forced” to react with H₂O₂? The reduction potential of ferric peroxidases (~ -200mV) is clearly distinct from the globins (~ +50mV) (29). In peroxidases, the ferric and higher oxidation states are stabilized by the greater anionic character of the fifth (or proximal) histidine ligand. The greater anionic character is imparted by a strong hydrogen bond with an aspartate that acts as an H-bond acceptor from the δ nitrogen of the proximal histidine. By comparison, the globins supply a much weaker H-bond acceptor, a backbone carbonyl, such that the histidine ligand has a much more neutral character [Figure 23].

The distal cavity of peroxidase is also distinct from the globins. Although the globins and peroxidases both have distal histidine residues, the histidine in globins is ideally positioned to serve as an H-bond donor to the O₂ bound to the iron in a “bent” configuration. The histidine is closer to the plane of the heme and shifted more to its periphery to accomplish this role [Figure 23]. In peroxidases, the distal histidine is required as a general base (ε-nitrogen) for deprotonating the incoming H₂O₂. Consistent with this role, the imidazole ring is further away from the plane of the heme, and it is positioned above the middle of the heme rather than toward its edge. To enhance its basicity, this distal His residue (δ- nitrogen) donates an H-bond to an asparagine or

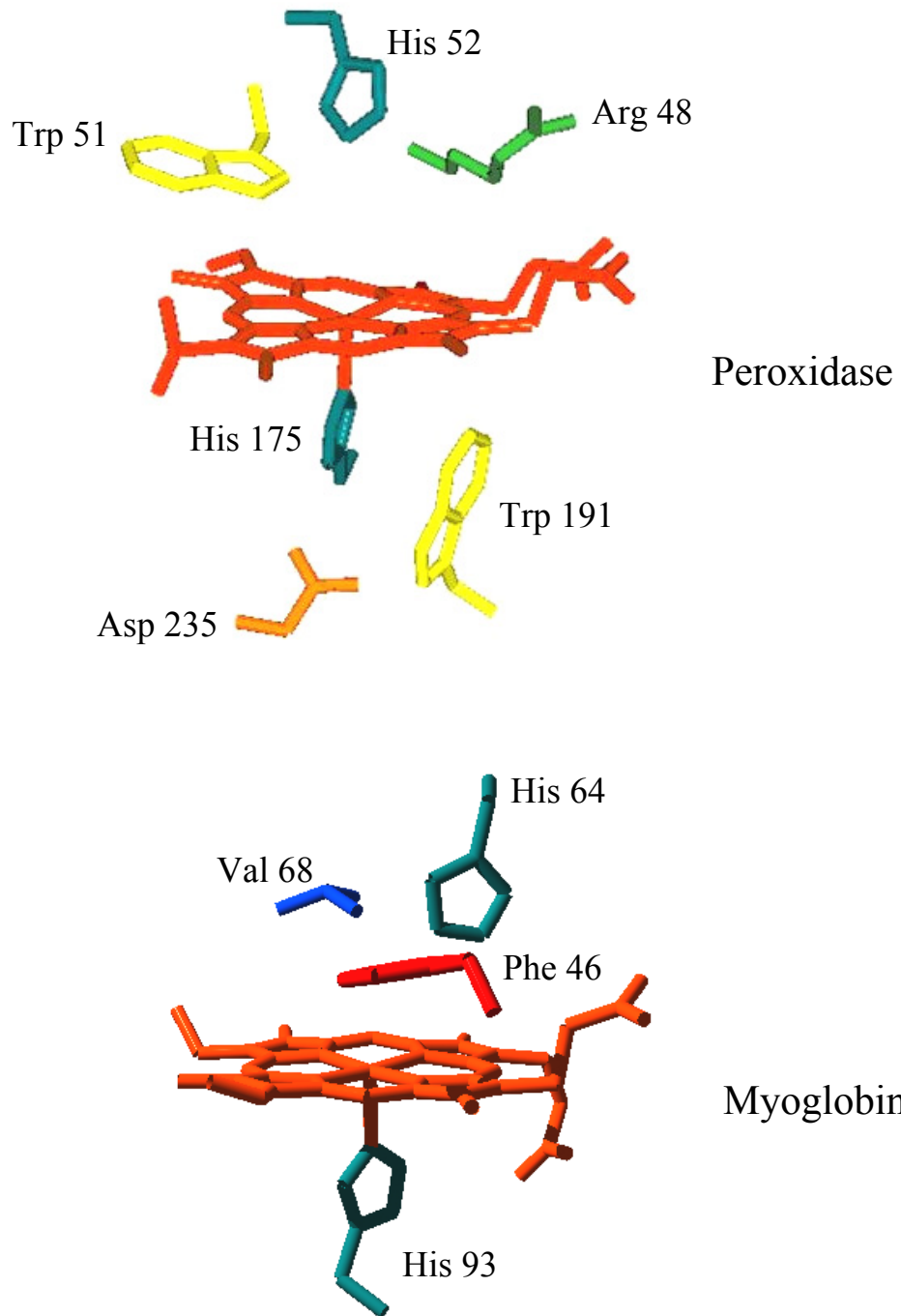


Figure 23. Important active site residues in monofunctional peroxidases and myoglobin
(34)

glutamine carbonyl oxygen. The rest of the distal cavity in globins is occupied by hydrophobic residues to disfavor the binding of inappropriate ligands and favor the nitric oxide oxidation chemistry performed by the $\text{Fe}^{\text{II}}\text{-O}_2$ complex of the globins. Conversely, the peroxidases include an arginine residue in the distal cavity. The positive charge imparted by this residue serves to offset the negative charge that develops on the second oxygen of the H_2O_2 during the formation of compound I.

It is important to briefly address the classification of the peroxidases. The characteristics described above apply, in a general sense, to most peroxidases although there is some variability in the specifics from one group of peroxidases to another. The two main groups of peroxidases are the mammalian, which include lactoperoxidase, myeloperoxidase, and prostaglandin synthase, and the so-called plant peroxidases, which include horseradish peroxidase, lignin peroxidase, and cytochrome *c* peroxidase.

The plant peroxidases are the most relevant to this dissertation, and the structural descriptions given above apply to all known plant peroxidases. The plant peroxidases can be further divided into three classes. Class I includes the intra-organelle enzymes cytochrome peroxidase and ascorbate peroxidase. The fungal secretory peroxidases (e.g., lignin peroxidase [LiP] and manganese peroxidase [MnP]) form Class II and the Class III plant peroxidases include the classical plant secretory enzymes like horseradish peroxidase (HRP) and peanut peroxidase.

Catalase-peroxidases

Even across a limited subset of heme proteins, the preceding discussion amply demonstrates that the protein environment around the heme group profoundly influences its functional capabilities. It should also be apparent that of the four levels by which structure dictates function, an understanding of the first three is bolstered by a wealth of information. It should also be clear that comparatively very little is known about the fourth level. That is, the mechanisms by which protein structures distant from an active site contribute to the functional capabilities of heme in one protein versus another.

For reasons described below, the catalase-peroxidases provide an excellent model to evaluate this poorly understood aspect of protein structure and function. Furthermore, the research described in this dissertation has taken advantage of the catalase-peroxidases to shed light on the mechanisms by which the heme environment can be controlled from a great distance. The benefits to be realized from these similar advances are truly far-reaching.

As the name implies, the catalase-peroxidases are capable of catalyzing at least two reactions, catalase and peroxidase [Figure 24]. In fact, other catalytic abilities have been identified, including peroxynitritase (30), and indole dioxygenase (31). What is particularly striking is that all of these catalytic abilities are accomplished using a single active site. Given the two dominant catalytic abilities, it is instructive to compare catalase-peroxidases to the corresponding monofunctional proteins, catalase and peroxidase.

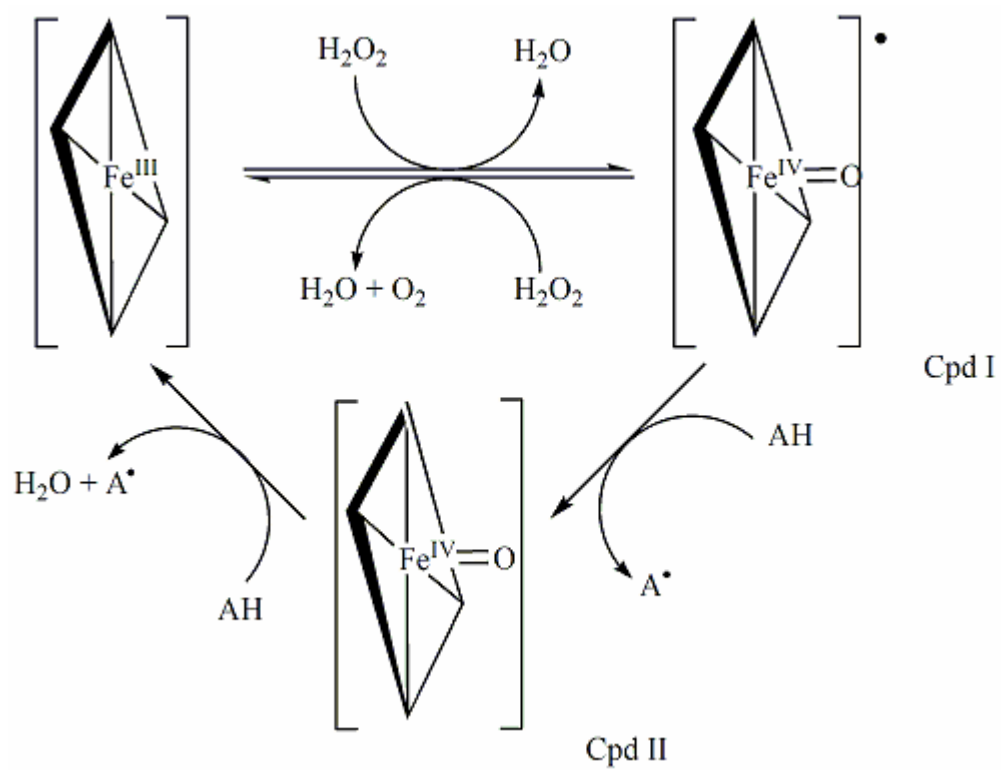


Figure 24. Catalase-oxidase protein catalase and peroxidase reaction scheme.

The first comparison is easy. Catalase-peroxidases bear no sequential or structural resemblance to monofunctional catalases. They both depend on a heme group for catalytic activity, but that is as far as the similarities extend. On the other hand, catalase-peroxidases are highly similar to monofunctional peroxidases. In fact, the active site of the Class I plant peroxidases are virtually superimposable on those of the catalase-peroxidases. For their strong similarity, catalase-peroxidases are considered Class I enzymes. Therefore, the peroxidase activity of the catalase-peroxidases is easy to explain; they have for all intents and purposes, a peroxidase active site. This makes the catalase activity of these enzymes difficult to explain. The rest of the plant peroxidases have little, if any, catalase activity. How is it then that the catalase-peroxidases are able to exhibit robust catalase activity using a peroxidase active site?

In terms of amino acid sequence comparison and structural comparison, it is useful to take a step back from the active site. From this vantage point, three structural features become immediately obvious as present in catalase-peroxidases but absent from monofunctional peroxidases. Two interhelical loops in catalase-peroxidases are extended by ~35 amino acids when compared to the analogous structures in monofunctional peroxidases. The first of these appears between helices D and E and is referred to as the DE insertion or large loop 1 (LL1) [Figure 25]. The second appears between helices F and G and is referred to as the FG insertion or large loop 2 (LL2). The third obvious catalase-peroxidase-unique structure is the 300-amino-acid C-terminal domain. Importantly, *all three* of these structures are peripheral to the active site.

The Goodwin laboratory has shown that the interhelical insertions are both essential for catalase but not peroxidase activity (32). The FG insertion or LL2 likely

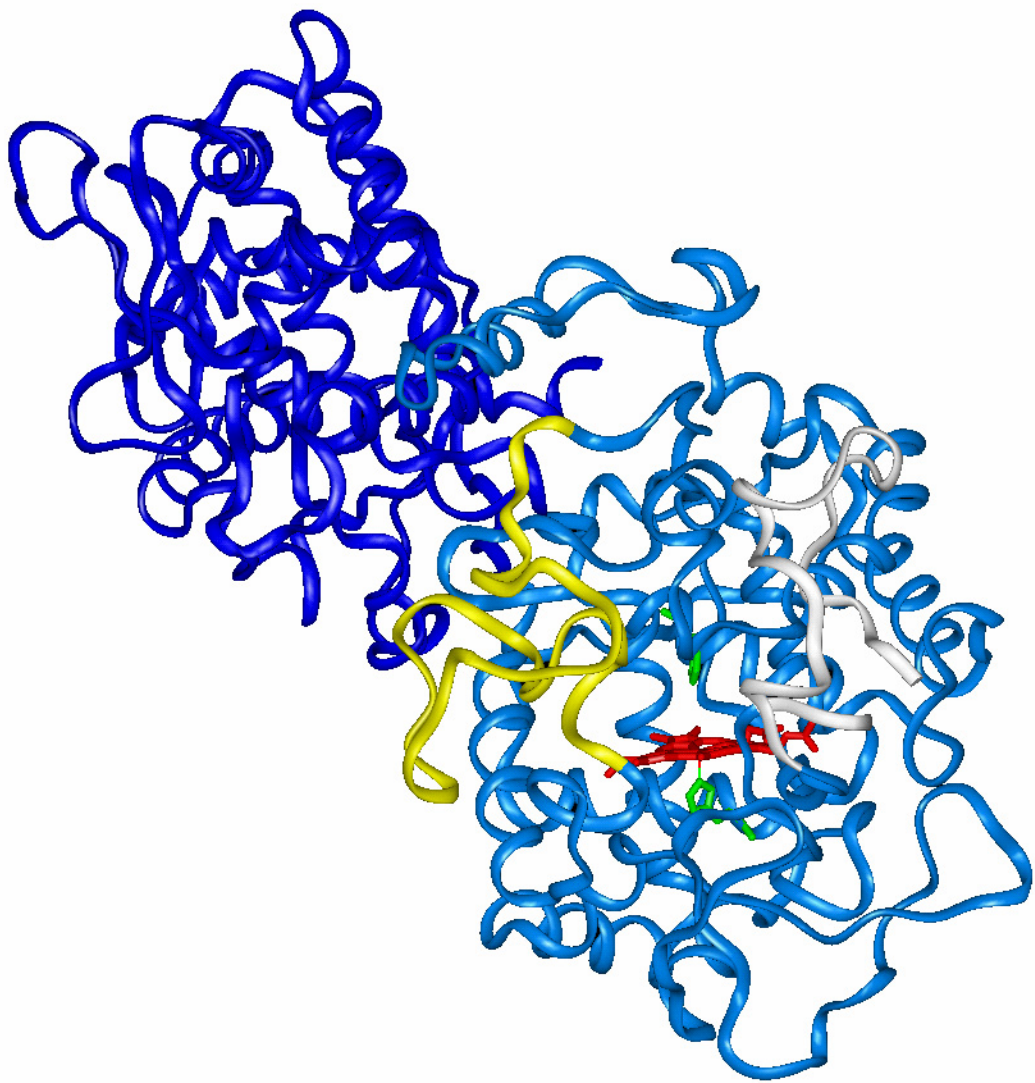


Figure 25. Ribbon diagram of catalase-peroxidase KatG. The DE insertion (white) FG insertion (yellow) and C-terminal domain (darker blue) are highlighted.

catalase-peroxidases (33).

contributes by two mechanisms. First, its N- and C-terminal ends contribute to a significant constriction in access to the heme edge. This would favor passage of H_2O_2 to the active site for oxidation to O_2 , but at the same time limit access to larger electron donors. Consistent with this role, removal of the FG insertion increases the apparent second-order rate constant for reaction with the electron donor ABTS. Second, the FG insertion structure parallels and interacts with an H-bonding network that extends to the distal cavity. Removal of this insertion may alter the H-bonding network of the active site. It has been observed that this network is critical for supporting catalase activity, but peroxidase activity is much less sensitive to its disruption.

The DE insertion or LL1 also forms part of the barrier between the heme edge and the solvent. Its removal also leads to a significant increase in peroxidase activity. The DE insertion also contains a strictly conserved tyrosine residue that is involved in a crosslink between three side chains Trp 105, Tyr 226, and Met 252 (*E. coli* KatG numbering). This covalent adduct has been observed in all of the structures solved for Furthermore, substitution of any one of the members of the adduct with an unreactive side chain (e.g., Tyr226 Phe) disrupts formation of the adduct and eliminates catalase activity. Peroxidase activity, however, is retained and even enhanced. How the covalent adduct contributes to catalase activity is not entirely clear, but it seems to center on the positioning and electronic properties of Trp. The formation of the adduct is dependent upon the presence dependent upon the presence of heme (a redox active porphyrin) and is enhanced by exogenous peroxides. Recently, the Ortiz de Monellano group has proposed and generated considerable evidence for a mechanism for Trp-Tyr-Met adduct formation that relies on high-valent heme intermediates.

The two domain structure of catalases is also distinct from monofunctional peroxidases. Interestingly, both the N-terminal and C-terminal domains bear striking resemblance to monofunctional peroxidases such as CcP (34) and HRP, however, the C-terminal domain does not contain the highly conserved active site residues, and thus, does not have the ability to bind heme or catalyze any reaction. The overall similarity of the C-terminal to the N-terminal and monofunctional peroxidases suggests that the catalase-peroxidase structure is the result of a gene duplication and fusion event (35, 36). The lack of catalytic activity in the C-terminal domain suggests it could be a vestigial “left over” that does not contribute to the function. However, prior to the research described in this dissertation, no work had been done to ascertain its role in catalase-peroxidase catalysis. On the other hand the C-terminal domain is unique to catalase-peroxidases. Of the three catalase-peroxidase-unique structures, it is the greatest distance from the active site ($\geq 30 \text{ \AA}$). Its large structure allows for the capacity to simultaneously participate in several mechanisms to support catalase-peroxidase catalysis. Thus, research directed toward the role(s) of the C-terminal domain has ramifications for understanding the unique versatility of catalase-peroxidases and provides a unique opportunity to investigate a poorly defined dimension of the heme protein structure/function relationship.

Methods for Determining Structure and Function of Heme Proteins

The role(s) of the C-terminal domain in catalase-peroxidases clearly provides a rich model to investigate important questions in heme protein structure and catalysis. It is important to highlight and briefly describe some of the wide array of methods that can be applied to address heme enzymes. While other examples of iron containing systems and molecules exist, the examples presented here effectively demonstrates the necessity of iron in biological systems but more importantly they illustrate the connection between the structure of the surrounding protein on one hand and the function in of the iron cofactor on the other. Given the obvious link between reactivity and protein structure it is therefore essential to understand this relationship.

This poses the question: how does one determine the protein structure? In essence, crystallography provides a snapshot of the static protein and provides valuable insight into the mechanisms by which it functions. However, in most cases the solved crystal structure presents only a starting point for pursuing a true understanding of mechanism and function. The dynamic structural properties of the protein, transient interactions, and active site transformation over time are more thoroughly known through spectroscopic and kinetic techniques. When considering metalloproteins, as with iron cofactors, spectroscopy is critical in the understanding of the coordination, oxidation state and electron spin state of the proteins active site and how these properties change in the proteins or enzymes execution of its function.

The primary spectroscopic techniques used in structural determination are UV-visible absorption, circular dichroism, magnetic circular dichroism, electron

paramagnetic resonance and resonance Raman spectroscopy. With the exception of resonance Raman all of these techniques were utilized in the structural characterization of the catalase-peroxidase proteins presented in this body of work. In conjunction with these with these techniques, kinetic methodology was also employed. For clarity a brief discussion of each is presented focusing on the application of these techniques to hemes and hemoproteins.

UV-visible spectroscopy (UV-vis)

Probably the first and most commonly used spectroscopic technique is UV-visible absorption. Absorption bands in the visible and near UV observed with heme proteins represent either $\pi \rightarrow \pi^*$ transitions, which arise due to excitation of porphyrin π electrons or charge transfer transitions, which occur when the d electrons of iron (III) interact with the π orbitals of the porphyrin. Intra-porphyrin $\pi \rightarrow \pi^*$ transitions can be divided into three different absorptive bands; α bands, β bands and γ (Soret) bands. These absorption bands differ in the energy levels to which they are promoted upon excitation. The ΔE of excitation increases respectively, and thus they are seen at wavelengths of increasing energy. The charge transfer band (CT) transitions occur when porphyrin electrons are excited from their highest-filled energy level into the lowest energy-unfilled d orbital of iron (III). It is evident that the ligand environment of the heme iron affects its optical absorption, both at $\pi \rightarrow \pi^*$ and CT transitions.

To discuss how these interactions affect the transitions one should begin with the simplest case and describe interactions using crystal field theory. In the case of iron (III) alone, Hund's rule applies and the five unpaired valence electrons of iron (III) are evenly distributed among the five degenerate d orbitals [Figure 26a]. In the case where ligands are present, each of the ligands has a lone pair of electrons accessible for coordination with the iron (III) center. Transitions are dependent upon these ligand interactions.

Take, for example, a hexacoordinate (i.e., octahedral) iron center with six ligands of equal bond strength and distance with respect to the iron. A repulsion force, dependent upon the d orbital location, also occurs due to interaction with the ligand lone pair electrons. For instance with ligands interacting with iron from the x, y and z axes, the iron electrons in the $d_{x^2-y^2}$ and d_{z^2} orbitals will be repelled to a greater degree than electrons in the d_{xy} , d_{xz} and d_{yz} orbitals. If the repulsion energy is not too great then Hund's rule still applies because the repulsion strength far offsets the energy difference which would result from an electron dropping into a lower, occupied orbital and thus pairing with the native electron [Figure 26b]. This case is called a weak ligand field environment where a single electron fills each d orbital of iron (III). Here the iron has five unpaired electrons available for bonding. This results in a spin-only magnetic moment and is referred to as a high-spin environment, with respect to the heme iron. In contrast, in a strong ligand field, the d orbitals are more widely separated in energy so that the repulsion energy to pair electrons is lower than the energy required to distribute each of the five iron electrons in individual d orbitals. In a strong ligand field, a spin-orbital magnetic moment from the single unpaired electron of the iron results in a low-spin heme environment [Figure 26c].

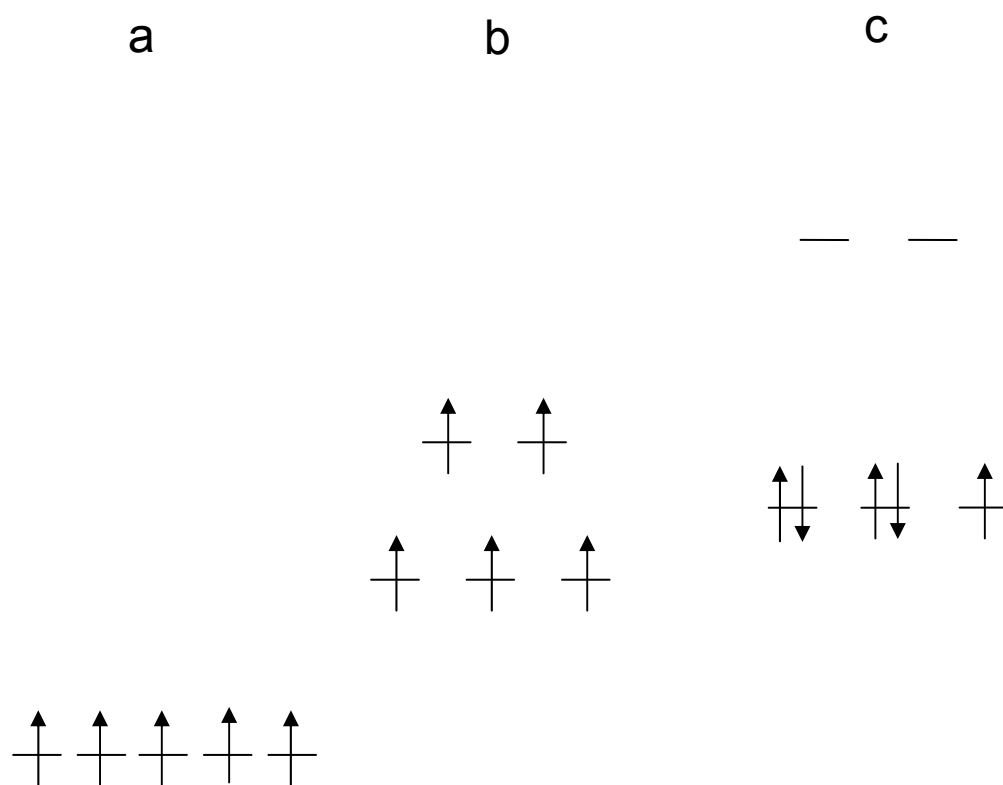


Figure 26. Distribution effects on the d electrons of iron (III) in (a) no field, (b) weak ligand field, or (c) strong ligand within an octahedral ligand field.

The crystal field description is adequate for describing high or low-spin, however, spectral properties are much more complex, requiring an understanding of the d orbital energy levels. Again, the five d orbitals of iron (III) are designated by their position with respect to the xyz coordinate system and include $d_{x^2-y^2}$, d_{z^2} , d_{xz} , d_{yz} and d_{xy} . Their ground state energies are a dependant upon the repulsion energy of the ligands with the electron as they approach the iron. In heme systems, the highest energy orbital is $d_{x^2-y^2}$ due to the fact that the four pyrrole nitrogens of the heme moiety, approach the iron on the x and y axes [Figure 27]. The next lower energy level is the d_{z^2} orbital where the proximal and distal ligands approach the iron (III). The d_{z^2} is followed in lower energy by the d_{xz} and d_{yz} , also known as the $d\pi$ orbitals which are repelled at an energy which is slightly higher than the d_{xy} orbital. As mentioned before there are three $\pi \rightarrow \pi^*$ transitions α (or Q_{00}), β (or Q_{0v}) and γ (also known as Soret or B) which occur within the porphyrin. The three $\pi \rightarrow \pi^*$ transitions are distinguished by an increasing energy corresponding to increasing Greek letter. While the α and β transitions both originate from the same π energy level and both end in the $e_g \pi^*$ level the β transition moves into a higher vibrational state and, therefore, is higher in energy [Figure 27]. The Soret or γ (used interchangeably herein) transitions higher energy is due to the fact that it originates from a lower π energy level, 2_{1u} , but ends at the same $e_g \pi^*$ as the α transition. If the metal ion in the porphyrin has filled d orbitals then the α , β and γ are the only transitions that are observed in the absorption spectrum.

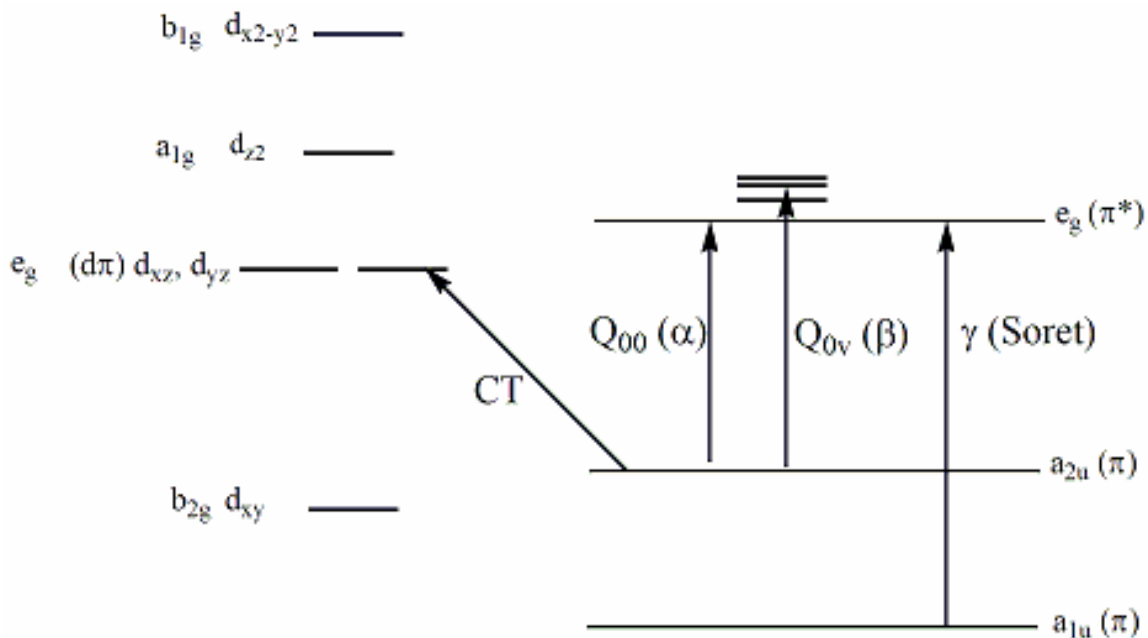


Figure 27. Diagram of absorption bands typical for peroxidases. Energy levels for iron are shown on the left and energy levels for the porphyrin on the right.

If this case is met then the species is called a regular metalloporphyrin. Contrarily, if the metalloporphyrin d orbitals of the metal ion are not completely filled, as with iron, then other transitions are seen where an electron in the highest, filled energy level of the porphyrin is promoted to the lowest, unfilled energy level of the metal ion. Metalloporphyrins which meet this criteria are considered irregular. These transitions are called charge transfer transitions (CT). One of these is illustrated in figure 27 with the 2_{1u} to the e_g transition of the iron (III) $d\pi$ orbital. These transitions usually occur at energy levels lower than that of $\pi \rightarrow \pi^*$ transitions. The charge transfer interactions of the porphyrin and the iron can greatly obscure the α and β transitions of the porphyrin, therefore, it is not atypical to observe the γ and charge transfer transitions alone in the absorption spectra of the iron metalloporphyrin [Figure 28]. These transitions conform to the D_{4h} symmetry model, however if iron is absent from the irregular metalloprotein then the symmetry is reduced to D_{2h} and splitting occurs (37).

This D_{4h} symmetry model is also acceptable to describe the absorption spectra of many peroxidases [Figure 28]. In most native peroxidases the fifth, proximal ligand of the porphyrin is occupied by an active site histidine and the sixth, distal ligand is predominantly unoccupied. There is a small amount of the native protein which has a hexacoordinate low-spin (one unpaired electron) geometry where the sixth ligand is typically occupied by hydroxide as well as a small amount of hexacoordinate high-spin (five unpaired electrons) species. More than one charge-transfer band may be due to

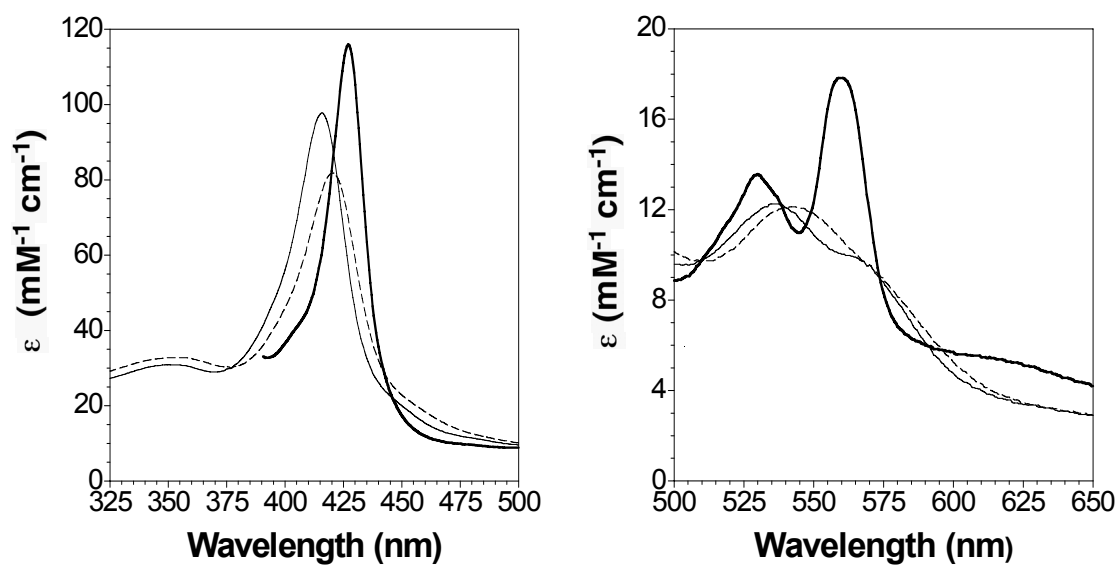


Figure 28. Typical absorption spectra for KatG in the Soret (left) and charge transfer regions.

splitting of the orbital of the iron as a result of distribution between the two orbitals of the d_{eg} energy level [Figure 28]. Also, strong field ligands, such as hydroxide and cyanide, will not have visible charge transfer bands due to the large energy difference in the d orbitals induced by the ligand strength [Table 1]. It is of course important to note that the visible range of this spectroscopy is probably the easiest method for determining protein concentration which is essential for further analysis on a quantified scale.

Circular dichroism (CD)

A deeper understanding of overall protein structure and the protein environment of a chromophore (like heme or flavin) can be determined from circular dichroism spectroscopy, commonly referred to as simply circular dichroism, CD. Circular dichroism passes circularly polarized light, across a wavelength range optimal for optical activation of the chromophore. The chromophore will rotate the plane polarized light, due to refractive indices which shifts the absorption band. Though the plane polarized light, is made up of equally proportioned left and right components of circularly polarized light the refractive indices of the sample cause the left and right components to travel at different velocities. These two components can be recombined, by the instrument resulting in an elliptically polarized beam. From the ellipticity the difference in the refractive indices of the left and right component can be determined, can be determined and comparing these two as a function of wavelength gives the optical rotatory dispersion (ORD) spectrum. A CD spectrum is the result of separating the left and right

Peroxidase	Absorption Band Maxima [nm]				
	Soret (γ)	β	α	CT2 ^a	CT1 ^b
HRP-Native ^c	403	~530	~580	498	643
HRP-CN	423	~539 ^e	~568	- ^d	-
HRP-F	404	560	530(s) ^e	488	612
HRP-OH	416	571	542	-	-
Ferro-HRP	439	557	~585(s)	~510(s)	-
Ferro-HRP-CN	432	536	566	-	-
CcP-Native ^f	408	-	~535(s)	505	645
CcP-CN	424	~570(s)	540	-	-
CcP-F	407	570	535	490	617
Ferro-CcP	438	561	530	-	625
LiP-Native	408	-	-	496	630
Ferro-LiP	436	-	552	-	-

^aCT2, short wavelength charge transfer transition (usually between 490 and 510 nm).

^bCT1, long wavelength charge transfer transition (usually between 600 and 650 nm).

^cSpectral properties for horseradish peroxidase (HRP)(38)

^dAbsorption bands that were not observed or were too weak to make unequivocal assignments are indicated by a dash.

^es denotes a shoulder on a major peak

^fSpectral properties for cytochrome c peroxidase (CcP)(38)

^gSpectral properties for ligninperoxidase (LiP)(39)

Table 1. UV-Visible absorption maxima for some peroxidases.

components by their independent extinction coefficients and comparing the two. The difference in the molar absorptivities is given by $\Delta\varepsilon = \varepsilon_L - \varepsilon_R$ in which $\Delta\varepsilon$ can be either positive or negative values with the units $M^{-1} \text{ cm}^{-1}$. The $\Delta\varepsilon$, when plotted as a function of wavelength, produces the CD spectrum. The results may also be presented in terms of molar ellipticity, θ , which is given by the following equation.

$$\theta = (2.303)(\Delta\varepsilon)(100)(180^\circ)/4\pi \quad (10)$$

Circular dichroism, in the Far-UV range (188-300 nm), provides information on the secondary structural content of proteins of the structural integrity of metalloprotein secondary structure. In the simplest sense the general structure could be determined by comparing protein spectra to established individual spectra of proteins made up of 100% of each; α -helix, β -sheet and random coil [Figure 29], however, the existence of databases which contain CD data on proteins with known structures provides a more accurate comparison. Various programs are available to calculate the percentage of each, α -helix, β -sheet and random coil of the protein of study. These programs provide a general idea of the secondary structure but can be very nearly quantitative when coupled with DNA sequencing.

Magnetic circular dichroism (MCD) is a variation on the traditional CD in which a magnetic field is applied to the sample parallel to the light beam. The magnetic field provokes optical activity in a particular chromophore through its effect on the electronic energy levels of the absorbing species. The MCD technique has been applied to a broad range of heme proteins. Thus a wealth of information allows for comparison of an unknown to other systems. Through MCD heme oxidation and spin states can be

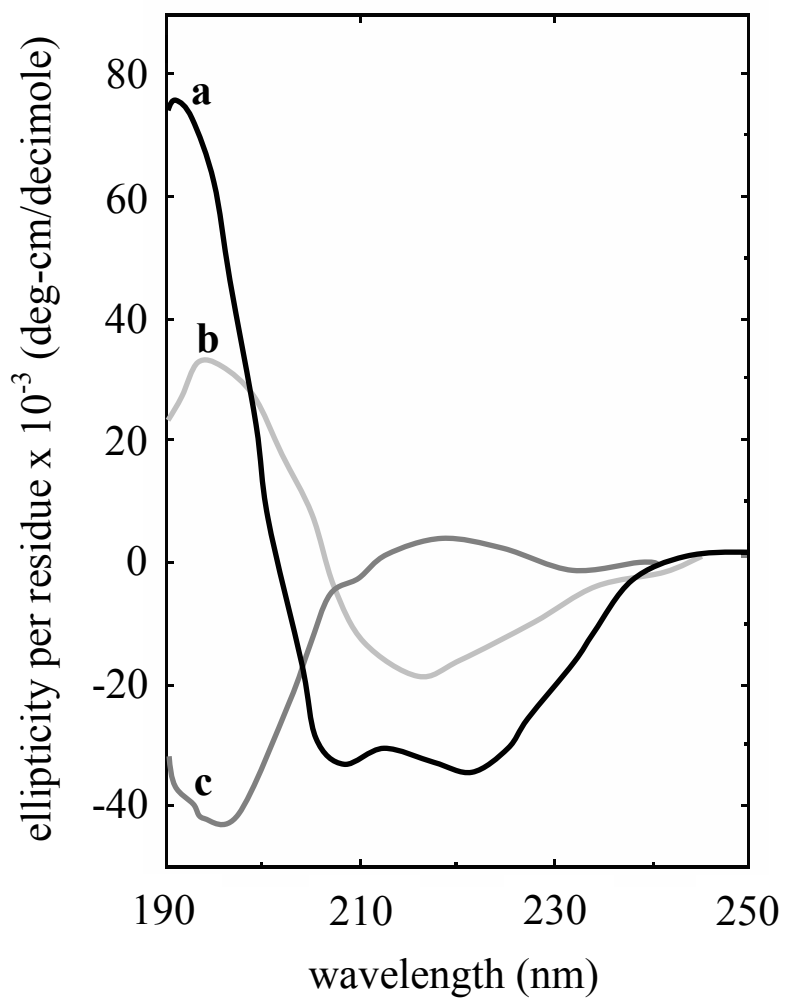


Figure 29. Approximate CD spectra for poly-lysine in three different secondary-structural conformations; (a) α -helix, (b) β -sheet, and (c) random coil.

determined. Moreover, the identity of axial ligands and characteristics of their protein environment can be discerned.

Specific electronic transitions of metal porphyrins, including $\pi\text{-}\pi^*$ and CT, can be focused upon due to Zeeman splitting which occurs when the excited and/or ground states are split in the applied magnetic field, giving rise to a signature pattern of the amount of splitting and the number of effective split levels. Additionally, the number of splits is a non-negative integer given by $2L + 1$, where L is the orbital angular momentum quantum number L of the given atomic level. The transition of electrons from the ground state to excited state of Zeeman sublevels can be induced by circularly polarized light as long as the angular momentum is shifted by ± 1 in the magnetic field.

Ferric iron MCD characteristics provide information about the active site resting ferric iron of many hemoproteins. The MCD spectra of hemoproteins, like UV-Visible spectra, can identify porphyrin $\pi \rightarrow \pi^*$ and porphyrin to metal, $\pi \rightarrow \text{Fe}^{\text{III}}$, charge transitions as well as identifying high- and low-spin ferric states. [Table 2] The low-spin state presents the simpler case of the two. In the low-spin state, the $\pi \rightarrow \pi^*$ transitions are clearly visible in the α - and β -band regions. Charge transfer transitions between π and Fe(III) in the low-spin heme overlap and are only visible in special cases. The low-spin CT transitions which are visible represent transition from the axial ligand to the Fe(III) and changes in this ligand result in substantial wavelength changes (40).

The high-spin ferric state shows more complex MCD spectra. Here the CT transitions from π to Fe(III) are higher in energy than the low-spin charge transfers due to

Heme spin state	Abs. Band Maxima [nm]	Associated transitions
Low-spin Fe(II)	425	$a_g \rightarrow t_{1g}, t_{2g}$
	550	$(\pi \rightarrow \pi^*) e_u$
High-spin Fe(II)	600-1000 ^a	$a_{g1} \rightarrow e_g$
	550	$(\pi \rightarrow \pi^*) e_u$
	671	$a_{1u}(\pi) \rightarrow d_{yz}$
	752	$a_{2u}(\pi) \rightarrow d_{yz}$
	813	$d_{xz} \rightarrow d_{z^2}$
	917	$d_{xz} \rightarrow e_g(\pi^*)$
Low-spin Fe(III)	400	$(\pi \rightarrow \pi^*) e_u$
	450, 550	$a_{1u}, a_{2u}(\pi) \rightarrow e_g(d_{xz}, d_{yz})$
High-spin Fe(III)	350-450, 500-600, 600-650, 700-1000	$(\pi \rightarrow \pi^*) e_u, a_{1u}, a_{2u}(\pi) \rightarrow e_g(d_{xz}, d_{yz})$ ^b

^a Multi-band region.

^b Transition overlap occurs in high-spin state and mixing occurs.

Table 2. MCD absorption maxima for hemoproteins heme states (41).

transfers from π a_{1u} and a_{2u} to two Fe(III) e_g excited states ($d_{xz, yz}$) (42). These stronger CT transitions are, therefore, strongly mixed with the α - and β -bands resulting in four E_u excited states seen at approximately 700-1000, 600-650, 500-600, and 350-450 nm. The stronger influence of the CT transitions in the high-spin spectra provides greater information about the axial ligand specifically with the 700-1000 nm transitions.

Electron paramagnetic resonance spectroscopy (EPR)

Another spectroscopic tool which is especially effective for hemoproteins is electron paramagnetic resonance spectroscopy (EPR), also known as electron spin resonance spectroscopy (ESR), which allows for the study of molecules which contain unpaired electrons. Transition metals, such as iron, typically contain one or more unpaired electrons. Thus, enzymes containing transition metals are well-suited for EPR study. It is important to note that species without unpaired electrons are EPR silent. The magnetic properties of the unpaired electrons allow for interaction with an external applied magnetic field. The sample, in a static applied magnetic field, is excited by an oscillating magnetic field, provided by microwave radiation. The static magnetic field is slowly adjusted until it resonates with the excitation of the sample electrons resulting in EPR absorption.

Several mathematical methods are employed to optimize the EPR spectra and collect information about the electron environment of the chromophore. For instance, the oscillating magnetic field proceeds through a slow sweep of magnetic field strengths in

order to detect resonance and with this slow sweep the instrument noise is increased which can easily overpower the absorption signal. To resolve this, spectra are recorded as first-derivatives given by sampling the absorption curve in increments which provide tangents to the curve. The EPR signal provides information the nature of the unpaired electron and its environment by the intensity, line width, g-values of the signal. From the first-derivative traces the area of the signal can easily be determined by double integration. The area under the signal curve corresponds to the spin concentration.

The line width can be simply described as reciprocals of relaxation times of the affected electrons. In the case of transition metals, this time can be very short resulting in line widths which are too large for the signal to be detected. This problem is resolved by scanning the sample at temperatures approaching 0 K. At such extremely low temperatures, the relaxation is retarded enough to facilitate line narrowing and, therefore, improving signal resolution. Of course, a factor is utilized in order to compensate for the temperature manipulation of the sample. Nicely, the determination of g value, area and intensity are aided by incorporated software programs, WIN-EPR for example.

Splitting of the curve or peak is the result of interactions of the magnetic moments two or more unpaired electrons and is called zero-field splitting. Thus, the zero-field splitting gives information about the number of unpaired electrons and their proximity to each other. Hyperfine splitting is a type of splitting due to interactions between the unpaired electrons and nuclear spin. In these interactions, the electrons involved with the nuclei are restricted to their orbital. In general EPR splitting involves an electron interacting with a single nucleus; therefore, the expected number of lines can be

determined by the following equation.

$$\text{Number of lines} = 2I + 1 \quad (11)$$

Here I represents the spin quantum number of the nucleus which is added to the number of interacting nuclei, in the general case one nucleus. So there may be splitting due to interaction of an unpaired electron with more than one nucleus resulting in peak splitting that is not hyperfine splitting. An example of this condition would be if an electron is interacting with two nuclei resulting in the peak split into a triplet with intensities of 1:2:1 proportionally. If the electron involved in splitting is a π electron then it may interact with many nuclei. If the nuclei are equivalent, then the splitting is considered hyperfine and the number of lines can be predicted using Pascal's triangle. In contrast, if the nuclei are different then the prediction becomes quite complicated.

The g -value of EPR absorption peaks is a factor which specifies the interaction between the spins of the unpaired electrons (radicals) and the applied magnetic field indirectly providing information about the location and environment of the unpaired electrons. With spin one half, $S = \frac{1}{2}$, the g -value accurately describes the spin environment of the radicals, but if the spin is such that $S > \frac{1}{2}$ then g is not a direct measure of the total spin state but rather a subset of the whole. The g -value depends directly on the orientation of the direction magnetic field with paramagnetic species. The relationship is given in the equation below.

$$\nu = g\beta H/h \quad (12)$$

Here h is Planck's constant, β is the Bohr magneton, ν is the frequency of the spectrometer, and H is the applied magnetic field used in detecting a certain signal.

Species with unpaired electrons have double-degenerate states and these energy levels separate in an external magnetic field. The degree of separation depends on the nature of a radical and the magnitude and orientation of the external magnetic field which is proportional to the g-tensor. The g-tensor approximates the coordinate system formed by the x, y and z axes. Values for the g-tensors give the g-values which describe the location and environment of the unpaired electrons. For instance, in the case of a heme protein, if the magnetic field is perpendicular to the plane of the heme porphyrin, then the resulting three g values, from 1 to ten, correspond to g_x , g_y , and g_z . This case applies only to a single crystal sample. Most samples, however, are frozen aqueous or low-temperature glasses of the sample due to the difficulty associated with producing a single crystal suitable for analysis. In these cases a species exhibits g values corresponding to the following equation.

$$g^2_{\text{obs}} = g_x^2 l_x^2 + g_y^2 l_y^2 + g_z^2 l_z^2, \quad (13)$$

Here l_x^2 , l_y^2 and l_z^2 correspond to the direction cosines of the applied magnetic field to the x, y, and z axes on the g-tensor. The g-values are determined by matrix analysis generally by computer software.

There are three types of symmetry associated with the three g-values; rhombic, axial and isotropic. Rhombic symmetry refers to the fact that the symmetry about the heme iron can be no higher than rhombic and occurs when none of the g-values are numerically equal [Figure 30b]. If two of the g-values are equal, then the heme has axial symmetry about the iron [Figure 30a]. Two subsets of the axial symmetry exist. One is where the two equivalent g-values (g_{\perp}) is greater than the unique value (g_{\parallel}). The other

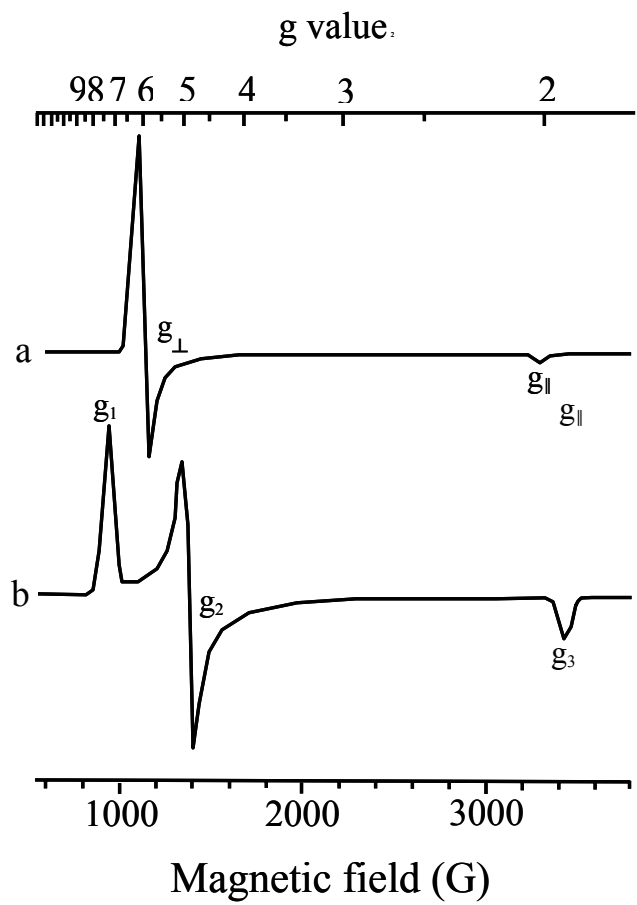


Figure 30. Approximation of EPR spectra for completely axial signal (a), and completely rhombic signal (b).

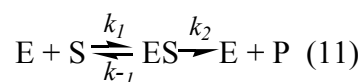
axial case is the reverse, where $g_{\parallel} > g_{\perp}$. Lastly, if all three g-values are numerically equivalent then the symmetry about the heme is isotropic .

Of the spin-states of iron, by far, most EPR studies have observed two spin-states for the iron III; low-spin ($S=1/2$) and high-spin ($S=5/2$). The horseradish peroxidase native indicate a low-spin peroxidase species with values centered at $g=2$ while Compound II is EPR silent (37, 43, 44).

It is also important to consider the low-spin heme environment with respect to crystal field and distortion symmetry. Low-spin hemoproteins have two paired electrons and one unpaired electron in the three axial ground state orbitals which results in unequivalent d_{xz} and d_{yz} thus giving $2/3\Delta$, where Δ is the rhombic contribution is represented by V , which is proportional to the ligand field strength of the axial component along the x and y axes. As the separation of the three ground state orbitals is equal then rhombicity is maximized. Simply, $V/\Delta = 2/3$ which is considered 100% rhombicity. Given that none of the g values are equivalent which is consistent with the conditions for exclusively rhombic symmetry. These three distinct g values result from the possibility that the unpaired electron may reside in any the d_{xz} , d_{yx} and d_{xy} rather than being confined to one specific orbital. This phenomenon is due to spin-orbital interactions.

Kinetic Methods for Determining Function of Heme Proteins

A good understanding of structure must be coupled with an intimate knowledge of function in order to understand structural effects. Essential tool for understanding of enzyme function, catalysis, is provided through kinetic methods. At its simplest, and enzyme mechanism must include a complex between the enzyme and its substrate. This necessitates three reactions, at the very least, to describe conversion of a substrate (S) to product (P) by an enzyme (E) (reaction 11). In its simplest form, catalytic conversion of a single substrate to a single product proceeds by one enzyme-substrate complex (ES).



Of course, an enzyme mechanism can be made more complex by the participation of multiple substrates, generation of multiple products, involvement of multiple intermediates following the formation of ES, and/or allosteric interactions between active sites on separate subunits. Nevertheless, years of evaluation of steady-state behavior of a broad range of enzymes has demonstrated that the model above is useful for determining the broad kinetic parameters governing catalysis by a given enzyme.

Steady state kinetics

Steady state is defined as the point where the sum of the rates of reaction leading to ES is equal to the sum of the rates of reactions leading to the breakdown of ES, in other words

where [ES] doesn't change over time.

$$d[ES]/dt = 0 \quad (12)$$

Here the substrate, interacting with the enzyme, is at a concentration of such excess that its concentration is considered constant. Constant substrate concentration is also ensured by the fact that in this type of kinetic measurements, initial rate or initial velocity (v_0) is considered. Operationally, a rate qualifies as v_0 if it is measured in the region where the substrate is $\geq 90\%$ of its starting value. This gives the nice benefit that the concentration of product is very low, allowing one to ignore the contribution from the back reaction. Since the concentration of the substrate [S] is constant there is not a significant concentration of product [P] being produced so its concentration is considered negligible. Since the concentration of the substrate is constant, it can be assumed that the rate is the initial rate, v_0 . This is called the initial-rate assumption and the initial rate equation, below (equation 13), is derived by making the steady state approximation for [ES] (see equation 12).

$$v_0 = \frac{[E]_0[S] k_{cat}}{K_M + [S]} \quad (13)$$

This equation is named for L. Michaelis and M. L. Menten in honor of their pioneering contribution in developing the first enzyme rate equation. They used the rapid equilibrium assumption to estimate the [ES]. This simply states that the rate constant k_2 is much less than the rate constant k_{-1} . According to this model the initial rate equation is given by equation (14) and the dissociation constant for the ES complex is given by K_S (equation 15).

$$v_o = \frac{k_{cat} [E]_o [S]}{K_S + [S]} \quad (14)$$

$$K_S = \frac{[E][S]}{[ES]} = \frac{k_{-1}}{k_1} \quad (15)$$

Briggs and Haldane expanded the work of Michaelis and Menten by supposing that k_2 maybe equal to k_{-1} or even far greater. In either of these cases, the rapid equilibrium assumption cannot hold, and the steady-state assumption is applied. Here K_S is replaced by K_M , the Michaelis-Menten constant which includes the contribution of k_2 (equation 16).

$$K_M = \frac{k_2 + k_{-1}}{k_1} \quad (16)$$

This equation describes the relationship between velocity and $[S]$. Here $[E]_o$ represents the total enzyme concentration, that is, the sum of the concentrations of all enzyme species. The term K_S represents the true dissociation constant for the ES and is only measurable from steady-state kinetic data under specific limiting cases. On the other hand, the K_M represents the concentration of substrate that produces $\frac{1}{2}$ of the maximum velocity (V_{max}). Thus, this parameter gives a measure of the substrate necessary to achieve maximal catalytic output. K_M is often used as a measure of enzyme affinity for substrate (high-affinity = low K_M and visa-versa), but this is an extrapolation that must be treated with caution because, as the reader will recall, k_2 might be quite large relative to k_{-1} . The term V_{max} may not be immediately obvious in equation 13. In the simple enzyme mechanism of reaction 11, it is represented by the term k_2 . By this simple

mechanism, the rate of reaction can be represented as $k_2[ES]$. As the $[S]$ is increased, the $[ES]$ approaches a limit that is $[E]_0$. Thus, the greatest observable rate must be $k_2[E]_0$, which is assigned as V_{max} . At so-called "saturating" concentrations of S, the rate approaches V_{max} as an asymptote.

More generally, the parameter k_{cat} is used to refer to the apparent first-order rate constant that governs the reactions of all ES and EP complexes. It corresponds to the rate-determining first-order reaction on the forward pathway. Thus, for all enzymes that behave according to Michaelis-Menten kinetics, $k_{cat}[E]_0$ is equal to V_{max} . Thus, $V_{max}/[E]_0$ obtained from laboratory measurements is taken to be k_{cat} . The parameter, k_{cat} , is also known as the turnover number since it represents the number of times the cycle is completed by an active site per unit of time under saturating substrate conditions. Thus, an overall equation to describe enzymes that follow Michaelis-Menten behavior is as follows equation 15.

$$v_o = \frac{k_{cat} [E]_0 [S]}{K_M + [S]} \quad (17)$$

At low concentrations of substrate, v_o , can be described by equation 18, a simplification of equation 17.

$$v_o = \frac{k_{cat} [E]_0 [S]}{K_M} \quad (18)$$

Under these conditions, $[E]_0$ is dominated by the free enzyme, $[E]$, and the equation takes on the form of that describing a second-order reaction between free enzyme, E, and substrate, S. Thus, the ratio of k_{cat}/K_M represents the apparent second-order rate constant.

At low concentrations of S, v_o increases linearly and the slope of that line

corresponds to V_{max}/K_M . Thus, by extension in that same [S] range $v_o/[E]_T$ increases linearly and that slope gives the apparent second-order rate constant k_{cat}/K_M . The steady-state kinetic behavior of catalases and peroxidases can be quite complex. Indeed, monofunctional catalase enzymes do not even follow typical Michaelis-Menten behavior. Fortunately, the catalase activity of catalase-peroxidases does follow Michaelis-Menten kinetics. The apparent K_M ($\sim 3\text{mM}$) falls substantially below H_2O_2 concentrations where enzyme inactivation occurs ($\sim 0.4\text{ M}$), allowing for reasonable estimation of apparent k_{cat} values.

Peroxidases are two-substrate enzymes that follow a modified ping-pong type of mechanism. If kinetic parameters with respect to one substrate are measured in the presence saturating and constant concentrations of the other substrate, reasonable Michaelis-Menten behavior is observed. Under these conditions kinetic parameters can be determined. The difficulty with peroxidase kinetics, particularly as these measurements apply to catalase-peroxidases, is substrate-dependent inhibition with respect to both H_2O_2 and some electron donors. This makes it difficult to identify true saturating conditions with confidence. The solutions to address this issue are to simply compare the maximum observed peroxidase rates of turnover at whatever substrate concentrations they occur. Alternatively, kinetics from one variant to another can be compared on the basis of apparent second-order rate constants.

In any case, the kinetic parameters obtained for any catalytic activity of the catalase-peroxidases should be interpreted with caution. Consequently, descriptions of kinetic behavior in this dissertation will refer to “apparent” k_{cat} and “apparent” K_M values. Along these lines, it is important to point out that all steady-state kinetic parameters were

obtained using non-linear regression methods to avoid the problems with error associated with linear reploting methods like Lineweaver-Burk.

Stopped-flow kinetics

From the preceding discussion, it should be clear that steady-state kinetic analyses are quite useful in outlining the kinetic boundaries of an enzyme-catalyzed reaction. They answer questions like: What is the rate constant for the step that governs turnover? Is the rate limited by diffusion of the substrate to the active site? However, steady-state kinetic methods are also quite limited. For example, they cannot identify the step that limits catalysis, or even the number and character of intermediates that appear between ES complex and EP complex. In fact, as discussed previously, something as simple as the equilibrium constant for the $E + S \rightleftharpoons ES$ reaction cannot necessarily be directed assessed by steady-state kinetic measurements. Rapid kinetic techniques that allow access to presteady-state or transient-state kinetic behavior are integral to more in depth understanding of enzyme mechanism. Stopped flow is one of these rapid kinetic techniques, and it has proven extremely useful in the evaluation of heme proteins and enzymes.

The heme prosthetic group, as discussed previously, is a robust chromophore amenable to UV-Visible, EPR, MCD, and other spectroscopic tools. The most common spectroscopic tool that is combined with stopped-flow is UV-visible absorption . The stopped-flow method allows access to the millisecond time scale ($k_{obs} = \leq 1,000 \text{ s}^{-1}$), and

the intermediates involved in heme enzyme catalysis (e.g. changes in oxidation, spin, and/or ligation states) are often easily tracked by UV-visible absorption. For example, the conversion of ferric peroxidase to Compound I is accompanied by a decrease (~50%) in the molar absorptivity of Soret band. The sequential return of the enzyme to its ferric state in the presence of an electron donor can be monitored by two fold increase in Soret intensity and shift in λ_{max} of ~ 420 nm. This is indicative of Compound II formation. The subsequent return to the ferric state is highlighted by a shift in the Soret band back to λ_{max} of ~405 nm. Each of these steps can be quite rapid with rate constants approaching $10^8 \text{ M}^{-1} \text{ s}^{-1}$.

Heme enzymes with catalase activity use only one substrate (H_2O_2) to complete their catalytic cycles. Usually the first step, Compound I formation, is rate-determining. This makes a stopped-flow experiment using H_2O_2 as the substrate rather uninteresting. In these situations, compound I is formed using alternative hydroperoxides such as *m*-chloroperoxybenzoic acid or peracetic acid. These peroxides support Compound I formation but not Compound I reduction. Thus, rates of Compound I formation with these alternative peroxides can be measured and compared from one catalase to another. Heme proteins are well-known to bind small molecules (CO, NO) and ions (CN^- , F^-) at the heme iron. These binding events are associated with rather dramatic shifts in absorption spectrum, and in many cases, the binding steps mimic steps of the catalytic cycle. For example, cyanide (CN^-) binding by ferric peroxidases resembles in many respects, Compound I formation. It is HCN that is the binding species, and the binding event is accompanied by deprotonation involving an essential general base (45, 46).

Stopped-flow is also a very useful technique to measure rate constants for binding (k_{on}) and dissociation (k_{off}) of these small ligands.

Stopped-flow's advantages include small sample requirements, and the ability to do both single and double mixing reactions. A single-mixing reaction uses only 110 ± 10 μl and sequential-mixing only 180 ± 10 μl , and when considering the low concentration of enzyme needed, usually μM , stopped-flow is considered extremely efficient. The stopped-flow instrument is equipped with drive rams and trigger system which are driven by an inert gas such as N_2 or Ar at a pressure of 8 bar [Figure 31a]. The optical cell is constructed so that one may measure along 2 mm or 10 mm pathlength in order to optimize absorption [Figure 31b]. The spectrophotometer uses a 150 Watt gap shortened Xenon lamp which is coupled with a monochromator for single wavelength scans of absorption.

In addition to rapid single wavelength kinetics measurements the stopped-flow has the capability of recording multi-wavelength spectra of species change during kinetic reactions. This accomplished using a 256 element photodiode array (PDA). Using the same Xenon light source, a wavelength range of approximately 280 to 735 nm can be recorded with a 2.17 nm separation. The transmittance of light through the reacting mixture charges the diode elements and the accumulated charge on the diodes is sequentially measured at a rate of 2.56 ms/scan. The diode charge is digitized by a 12 bit analogue to digital converter (A/D) and is transformed to a measurement of voltage. The voltage measurement is proportional to the integration time, or the time in between a diodes sampling, and also it is proportional to the transmitted output intensity. Each

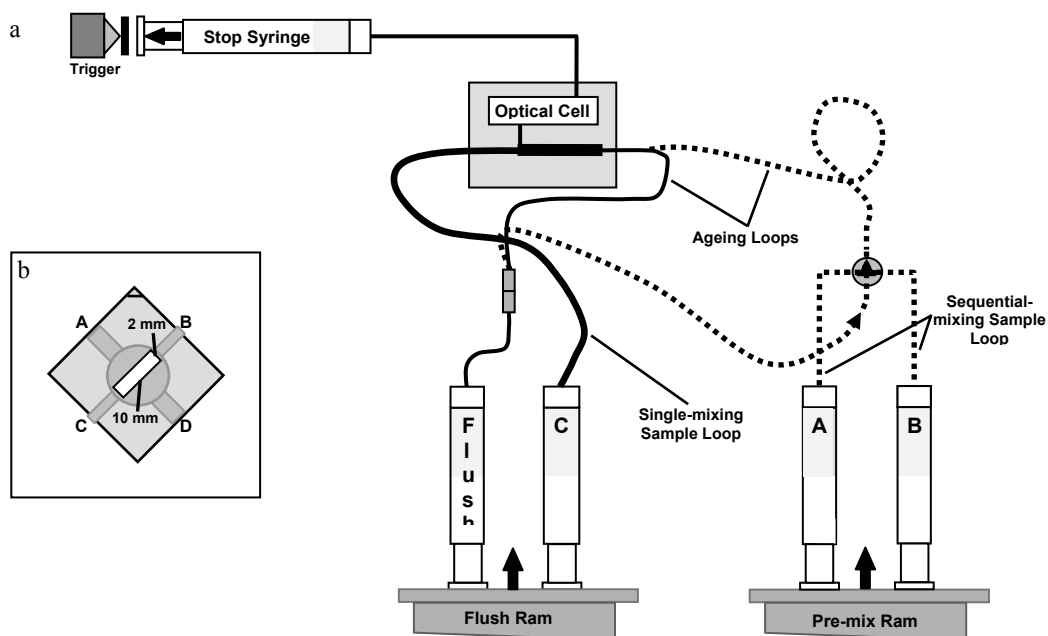


Figure 31. Stopped-flow diagram. (a) Diagram of single-mixing (solid lines) and sequential-mixing (dashed and thick solid line) stopped-flow setup. (b) Optical cell diagram for 2 mm and 10 mm path length options with A-D and B-C representing viewing ports for pathlengths respectively.

diode voltage is compared to a prerecorded 100% transmittance reference output voltage and the absorbance spectrum is calculated by equation 19.

$$\log_{10}(V_{\text{ref}}/V_{\text{sample}}) \quad (19)$$

V. Biomedical Applications of Heme-protein Structure and Function

In a previous section it was demonstrated that catalase-peroxidases offer a novel model to pursue and understand poorly defined aspects of the protein structure and function relationship. The potential benefits and applications to be gleaned from such research are far-reaching. However, it is important to point out that that there are substantial biomedical benefits to be realized from a more complete understanding of catalase-peroxidase structure and function.

Mycobacterium tuberculosis and antibiotic resistance

Worldwide, *M. tuberculosis* is the second greatest killer among infectious diseases, causing approximately two million deaths per year. Also, it is estimated that about one third of the world's population is currently infected with *M. tuberculosis* (TB) with an average of one person per second being newly infected (47). Despite the strides made by the internationally recommended control strategy (DOTS) the greatest threats remaining are AIDS related TB and multi-drug resistant strains (MDR-TB) (47, 48). The World Health Organization (WHO/IUATLD) Global Project on Drug Resistance

Surveillance estimates that greater than 4% of new cases are multi-drug resistant throughout Eastern Europe, Latin America, Africa, and Asia (48).

Isonicotinic acid hydrazide or isoniazid is a frontline anti-tubercular agent. It is easy and cheap to manufacture and has been widely used in the fight against tuberculosis. Indeed a positive TB skin test is enough to warrant an automatic six-month course of isoniazid. Unfortunately, there has been rapid development of resistance to isoniazid. It has been estimated that 70 % of resistant strains carry mutations to the gene for catalase-peroxidase (KatG) that comprise in some way the function of the enzyme. The key to understanding isoniazid resistance comes in the fact that isoniazid is a prodrug which is activated by the catalase-peroxidase enzyme of *M. tuberculosis* (49). Consequently, mutations to the gene for this enzyme figure prominently in *M. tuberculosis*, antibiotic resistance (50).

The mechanism by which isoniazid is activated is not fully known. Similarly, the mechanisms by which many of the commonly observed mutations comprise isoniazid activation are not known. Clearly, a better understanding of catalase-peroxidase structure and function will be beneficial to the problem of isoniazid resistant tuberculosis.

Bacterial virulence

In general, bacterial virulence is a measure of the efficiency of the bacteria to access the host, colonize, alter the host, and move onto a new host (51). In order to accomplish these steps there are mechanisms, called virulence factors, inherent to the

bacteria which contribute to their virulence. Virulence factors can be divided into four categories; offensive mechanisms, nutrient uptake, defensive mechanisms, and regulation. Catalase-peroxidases have been implicated as virulence factors in organisms like *E. coli* 0157:H7, *Yersinia pestis*, and *Legionella pneumophila*. These organisms produce a periplasmic form of the catalase-peroxidase. *E. coli* produces KatP. *L. pneumophila* produces KatA, and *Y. pestis* produces KatY (52-56). These enzymes are associated with virulence, first of all, because they are tied to other factors known to contribute to pathogenicity. For example, KatP is encoded on a large plasmid, p0157, that is carried only by pathogenic strains of *E. coli*. KatY is expressed with a number of virulence factors upon a shift in temperature (26 °C to 37 °C) associated with transfer from the flea to the mammalian host. It is also important to point out that the non-pathogenic relatives of these organisms do not produce periplasmic forms of the catalase-peroxidase.

How could catalase-peroxidases act as virulence factors? As discussed previously, a consequence of daily life in an aerobic environment is the generation of partially reduced or reactive oxygen species. Clearly, H₂O₂ is an important one of these. Enzymes with catalase and/or peroxidase activity are required to safely dispose of H₂O₂ to prevent damage to critical cellular components. Pathogens interact with their hosts in a way that arouses the immune response. A centerpiece of the mammalian defense against bacterial invaders is the generation of copious amounts of reactive oxygen species. The oxidative burst of neutrophils and macrophages rapidly generates high levels of O₂^{•-} which disproportionates to H₂O₂. It has been estimated that H₂O₂ levels reach as high as 100 mM in this biological warfare context. This is only the beginning of

the problem for bacterial pathogens. In neutrophils, H_2O_2 is used by myeloperoxidase to oxidize chloride to hypochlorite anion, the active ingredient in household bleach. In macrophages, the rapid generation of $\cdot\text{NO}$ along with $\text{O}_2^{\cdot-}$ produces peroxynitrite anion through a reaction that proceeds at diffusion-limited rates. Peroxynitrite (ONOO^-) and its conjugate acid, ONOOH , are known to be far more toxic to bacteria than H_2O_2 .

The ability to rapidly degrade H_2O_2 and other peroxides like ONOO^- is likely to afford pathogenic bacteria protection against the host immune response. The periplasmic location of these enzymes is a strategic placement that affords access to the target substrates before they have had the opportunity to damage essential cellular systems (e. g. respiratory electron transport proteins).

The structural basis for efficient peroxide decomposition and how it may be interrupted may hold clues for effective new therapeutic strategies to thwart disease caused by organisms like *Y. pestis* (bubonic plague) and *L. pneumophila* (Legionnaire's disease). In this respect, it is important to recognize that organisms like, *E. coli* 0157:H7, *Burkholderia pseudomallei* and *Y. pestis* are recognized as high-priority threats as agents of bioterrorism.

Virulence and iron acquisition

Iron is as essential to bacteria as it is to other organisms. Bacterial cells must maintain 10^5 to 10^6 iron ions for metabolism, many of which are used for production of

heme as an enzyme cofactor (57). For bacteria the acquisition of iron is a complicated and energy consuming undertaking. Bacteria expend as much as 10% of their total cellular iron and considerable energy to biosynthesis of heme iron (58). When considering the survival of pathogenic bacterial in the host, it is important to recognize that free iron is a rare commodity in the human host with a concentration around 10^{-18} M, much too low to sustain the bacteria (59). The bulk percentage of host iron, about 99.9% is found intracellular in protein complexes such as: myoglobin, hemoglobin, ferritin, and hemosiderin (60). The remaining iron is found extracellularly bound to transferrin and lactoferrin, in secretions, and mucosa (61, 62). Iron is predominately compartmentalized in the host to overcome the low solubility of ferric iron as well as to avoid ferric iron toxicity (62, 63).

To obtain adequate amounts of nutrient iron for survival microorganisms have developed unique strategies to strip it from the host. These strategies of nutrient acquisition involve the isolation, transport, and incorporation of essential iron. For the contribution to the subsistence of bacteria within the host, some nutrient acquisition strategies are considered potential virulence factors. Factors with virulence potential are heme receptor proteins which have been observed in highly virulent pathogenic bacteria but have not been seen in less virulent or non-pathogenic strains of the same bacteria.

Outer membrane heme receptor protein in virulence

Outer membrane heme receptors are integral membrane proteins found in the outer membrane of Gram-negative bacteria. These interact with hemophores and/or heme and other iron compounds to shuttle iron sources into the cytoplasmic space (64). The heme receptor for *E. coli* o157:H7, ChuA, is in a class of very efficient heme receptors because it can bind iron sources directly or can obtain it from interaction with siderophores (65). The variety of *in vivo* iron sources that ChuA can accept includes heme iron, hemoglobin, and ferric citrate (65, 66). The ability of this receptor to transport the entire heme moiety across the outer membrane provides not only an iron source but also a porphyrin source for the bacteria.

The expression of ChuA is regulated transcriptionally by the Fur repressor protein. The Fur repressor requires a divalent heavy-metal ion as an activator, in this case ferrous iron (67). The gene producing the ChuA protein, *chuA*, is nearly identical to the *shuA* gene for the *Shigella dysenteriae* heme receptor ShuA, and both of these are structurally similar to the heme receptor HemR of *Yersinia enterocolitica* with respect to conserved regions, domains and the dependence upon the trans-periplasmic protein TonB for energy transduction (63, 68). For this reason ChuA is considered a HemR-type or TonB-dependant type receptor (5).

There is no resolved crystal structure for ChuA and little is known about heme-receptor structure. Therefore, a suitable model was required to develop a proposed mechanism of transfer. The structure of outer membrane siderophore receptors, which have yielded crystal structures, is expected to be similar to that of the heme receptor (62,

69). For these reasons the structural data for FhuA, the siderophore receptor for *E. coli*, in addition to what is known of HemR is used in this proposed mechanism (68). In both cases, only the regions of conserved residues in common with ChuA form sequence alignments were used to describe the predicted structure and mechanism.

Like FhuA, the heme receptor is most likely a porin-like, monomeric β -barrel composed of antiparallel β -strands connected by short α -helices or loops (63, 68, 70-72) [Figure 32]. The channel formed is sterically occluded from the periplasmic space by a globular, N-proximal domain most commonly referred to as the cork (68, 70). This barrel is much larger than any formed by simple porins, in fact, ChuA is a 69 KDa protein. Like many outer membrane proteins two regions of aromatic residues girdle the exterior of barrel and extend into the lipid bilayer, interacting at the interface between the lipids polar heads and hydrophobic chains (68). The cork is positioned at approximately 45° angle with respect to the membrane creating two cavities; the external cavity which is open to the external environment and the periplasmic cavity which is open to the periplasmic space. The cork is connected to the barrel through extensive hydrogen bonding that decreases slightly upon ligand binding and, therefore, resembling a hinge. The cork is a mixed four stranded β -sheet connected by loops extend above and below the cork into both the external cavity and the periplasmic cavity. Among these loops are three specific loops, apices A, B and C, which are indicated in ligand transfer. Other striking structural features include a band of strictly conserved residues, including Arg, Asn, Asp and Gln, extending the length of the heme receptor. Also, within the external or periplasmic cavity numerous aromatic residues extend from the β -

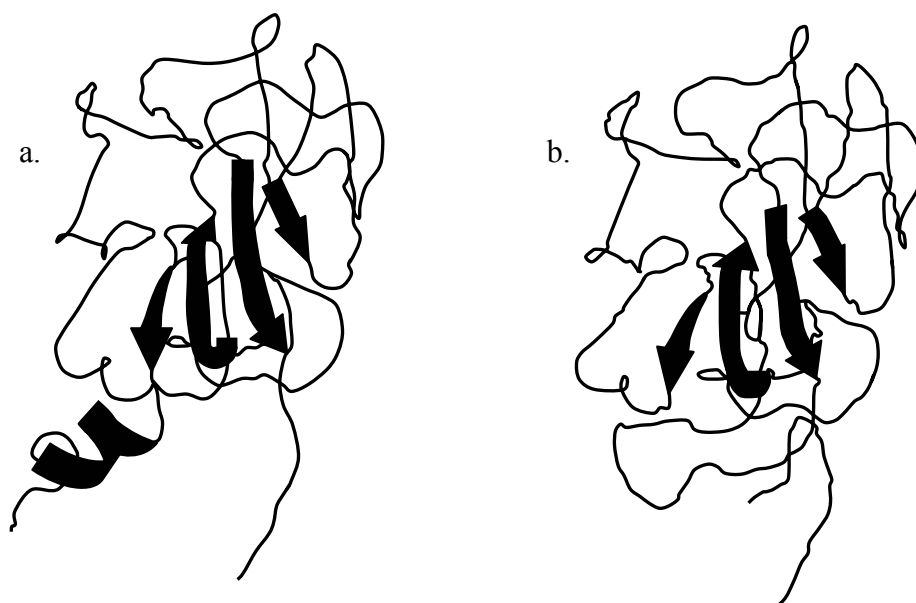


Figure 32. Ribbon diagram of approximate FhuA interior and cork structures for (a) apo-structure, before heme binding, and (b) holo-structure, with heme bound.

strands the length of the barrel's inner walls and from the external surface located loops. Among heme receptors there is an additional conserved domain, the FRAP/NPNL domain found on the barrels external surface. Also, heme receptors have two strictly conserved and functionally important histidine residues. One of these residues is located in the FRAP/NPNL domain of barrels external surface, HisB, while the other histidine, HisC, is located on the cork's external facing side. These residues are essential for acquisition of heme-protein complexes and is necessary for adequate free heme utilization as well (63).

Little is known of the mechanism of heme transport in ChuA and for this reason a simplified model is presented below. This model is based upon structural features of the apo- and holo-FhuA siderophore receptor and HemR heme receptor (63, 68) [Figure 33]. Considering that bacteria are in a ligand-available environment the ligands first interact with aromatic residues found on the surface loops. The high degree of aromaticity in the external cavity provides a suitable ligand environment and thus the ligand is drawn into the external cavity. In heme receptors it is theorized that the invariant HisB residue ligates the heme iron to the barrel in the external cavity. Upon ligand binding the apex B is translated upward toward the ligand forming multiple hydrogen bonds. This induced fit mechanism is also seen with the translocation of all the loops of the cork domain between apex A and the periplasmic cavity. This upward translation disrupts a specific helix in the periplasmic cavity termed the switch helix. The switch helix made up of numerous hydrophobic residues and resides coiled in a complimentary hydrophobic pocket until ligand binding occurs.

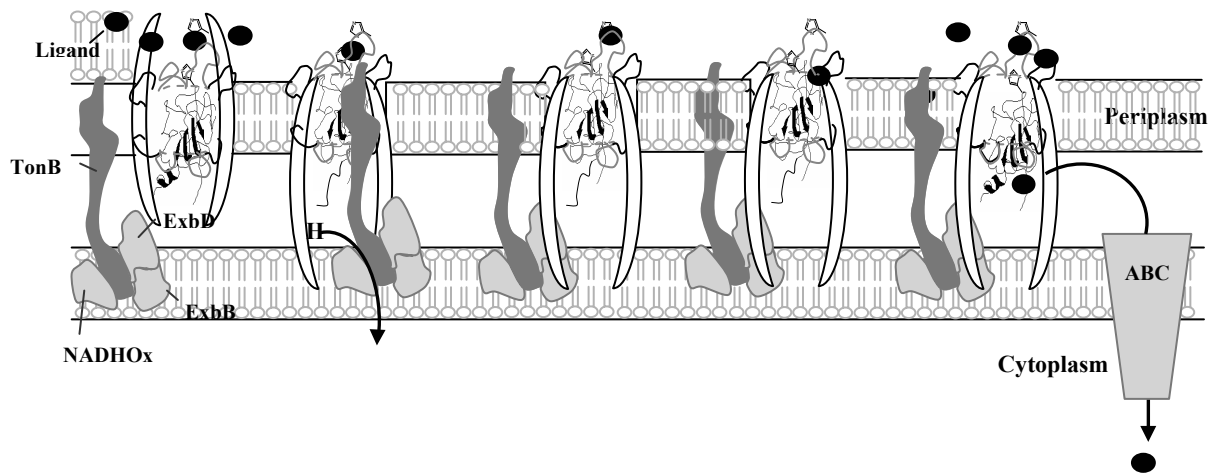


Figure 33. Diagram of a model of heme receptor function in the transfer of iron from the exterior environment of the bacterium into the cytoplasm.

Upon ligand binding a disruption in the interaction between the hydrophobic residues of the switch helix and its complimentary residues cause the helix to be destabilized and as it unwinds into a coil [Figure 33] a region of conserved residues, termed the Ton-box, is moved across the periplasmic channel (68). The Ton-box region, residues 7-11 in FhuA, is strictly conserved and has been shown to form multiple interactions with TonB (68, 73, 74). This helix-to-coil conformational change of the switch helix signals the ligand-bound status of the receptor and the subsequent need for TonB-dependant energy transduction. TonB, with the assistance of ExbB and ExbD proteins, work together to couple cytoplasmic proton motive force with active transport across the outer membrane for nutrients such as iron complexes and vitamin B12 which are too large for simple diffusion (75).

In response to the signal the location of a conserved sequence, the Ton-box, shifts to interact with the receptor. Upon TonB energy transduction allosteric transitions occur which destabilize the ligand binding site thus lowering the affinity for the ligand. In the heme receptor the cork histidine, HisC, is now positioned to accept the ligand due to a now higher-affinity for the ligand than that of HisB. Changes in the affinity between the two histidines are expected to occur due to TonB-receptor interaction which is suggested to mediate a mechanical pulling force applied to the central β -sheet of the cork via electrostatic interactions. With transfer of the ligand to the cork complete it is suggested that TonB disengages, having completed transduction, and the cork apices translate back to their original position. This is essentially a reversal of the first translation but now the cork is loaded with ligand. Upon translation a small water filled channel beneath apex B, previously occluded by the destabilized switch helix, is opened. This putative channel is

large enough to allow the diffusion of the ligand through the periplasmic cavity. This observed transition suggests that the switch helix rewinds into its hydrophobic pocket following energy transduction, which may account for the reverse translation of the cork. The strongly conserved residues that extend the length of the barrel could function as low-affinity binding sites to aid in surface diffusion through the receptor.

This proposed mechanism is consistent with evidence from iron transport systems, however, it is only a model (69). Given the lack of functional and structural knowledge about heme receptors, specifically ChuA, and its likelihood to be a target for novel antibiotics we endeavored to characterize the *E. coli* heme receptor ChuA (76). It is evident that nutrient acquisitions, like iron, are essential to the bacterial survival and virulence.

Therefore, the evidence presented here describes how imperative it is to understand the structure/function roles of proteins. This information may not only be applied to the general understanding of the correlation of the two but also applied to better understand virulence in bacteria, particularly pathogenic bacteria. Strong implications in the understanding of antibiotic resistance and application to the combat of this resistance support this body of work. Additionally, the concept of engineering enzymes to manipulate their function is a new perspective requiring a deeper understanding the connections between structure and function of enzymes.

III. METHODS AND MATERIALS

Materials.

Luria-Bertani (LB) broth, d-aminolevulinic acid (d-ALA), Sephacryl 300 HR, ferrous ammonium sulfate, isopropyl- β -d-galactosidase (IPTG), tris base, tris HCl, sodium chloride, glycerol, ethylenediamine tetraacetic acid (EDTA), tetracycline hydrochloride urea, and Amicon Ultra centrifugal filter units, 5000 MWCO, were purchased from Fisher Scientific (Pittsburgh, PA). 10 DG buffer exchange columns were purchased from BioRad (.). Lithium bromide (99.4 % solution), polyethylene glycol 4000 (PEG) and calcium chloride dihydrate (≥ 99.5 %) were purchased from Fluka (Milwaukee, WI). Cryschem™ crystallography plates, 4 in crystal clear sealing tape and Stock options™ 1M sodium acetate solutions were purchased from Hampton Research (Aliso Viejo, CA). All oligonucleotide primers were produced by Invitrogen (Carlsbad, CA) and all restriction enzymes used as well as calf intestinal phosphatase were purchased from New England Biolabs (Beverly, MA). BSA Protein Assay kit and BSA standard were both products of Pierce (Rockford, IL). Bugbuster and benzonase were purchased from Novagen (Madison, WI). Nickle-nitriloacetic acid resin (Ni-NTA), QIAGEN™ Plasmid Miniprep and Midiprep plasmid purification kit, QIAquick® P

Purification Kit and, Gel Extraction Kit and Factor Xa protease were purchased from Qiagen (Valencia, CA). Peracetic acid, hydrogen peroxide (30%), imidazole, phenylsopharose resin, ampicillin, chloramphenicol, phenylmethylsulfonyl fluoride (PMSF), 2, 2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS), heme and hemin affinity resins, bovine hemin chloride, N-lauroylsarcosine and lysozyme were acquired from Sigma (St. Louis, MO). All *E. coli* strains (XL-1 Blue and BL-21 [DE3] pLysS) as well as *Pfu* polymerase were obtained from Stratagene (La Jolla, CA). All solutions were prepared using either Barnstead (EASYpure II) or Millipore (MilliQ, QPAK II)-purified water (18.2MΩ/cm resistivity).

Methods

Cloning and Expression of the Heme Receptor protein (ChuA).

Cloning of heme receptor gene (chuA).

The gene for the *E. coli* 0157:H7 heme receptor (*chuA*) protein was cloned using the Seamless® Cloning procedure (Stratagene). The *chuA* was amplified from p0157 plasmid isolated from *E. coli* 0157:H7 using primers 0157HRa01 (CCT CAT GGA GCT CTT CTA TGT CAC GTC CCG CAA TTT ACC TCG) and 0157HRa02 (ATC GGG GCT CTT CAG AGC CAT TGA TAA CTC ACG AAA ATT TTT CCG) designed to introduce overhanging sequences for *Nde I* on the N-terminal end and *Xho I* on the C-

terminal end. Additionally, the recipient construct of pET20b(+) was amplified using primers pET20a01 (GCT TGC CTC TTC ACT CGA GCA CCA CCA CCA CCA CCA CTG) and pET20a02 (GCA GCA CTC TTC TCA Tat gta tat ctc ctt ctt aaa gtt aaa c).

The resulting PCR products were combined and digested with Eam1104 I for one hour at 37°C and the digest products were ligated with T4 DNA ligase for one hour at 37°C to produce the final plasmid, pChuA1, was constructed such that ChuA would be expressed with a C-terminal six-histidine tag. The pChuA1 plasmid was into *E. coli* (XL-1 Blue MRF') by the heat-shock procedure for transformation. Transformants were screened based on ampicillin resistance to identify candidate colonies. Plasmids from those candidates were screened for correct pChuA1 construction protein by a diagnostic restriction digest using *XhoI* and *NdeI*. Digest of the correct plasmid would result in the two cuts bands. Positive candidate plasmids were subjected to DNA sequencing analysis (Davis Sequencing Facility, Davis, CA) to ensure that no accidental mutations had occurred during the cloning process. The final plasmid (pChuA1) was transformed, as for KatG clones, into *E. coli* expression host (BL-21 [DE3] pLysS).

Since the expressed ChuA bears a six-histidine tag which needed be ruled out as a competitor with ChuA to bind the heme iron. In order to rule out the possibility that the six-histidine tag would ligate to the heme or hemin affinity resin rather than the actual protein a new plasmid was constructed to eliminate the histidine tag. The following mutagenesis primers were designed to introduce a stop codon (in bold) just prior to the six-histidine tag (underlined): HRm01(+) (GTG AGT TAT CAA TGG **TAA** GAG CAC CAC CAC CAC CAC), and HRm01(-) (GTG GTG GTG GTG GTG CTC **TTA** CCA TTG ATA ACT CAC). The resulting PCR was cloned and screened in accordance with

procedures for the cloning of the wild type in materials and methods. Confirmation by diagnostic restriction digest with *Bgl II* assured that no intended mutations had occurred.

Another approach to further validate heme binding by ChuA, was to utilize the six-histidine tag for purification but to cleave it off after said purification. In order to do this the pChuA1 plasmid was modified by introducing a Factor Xa cleavage oligonucleotide sequence into (77). The sequence introduced is such that Factor Xa protease cleaves following Arg in the following specific sequence: Ile-Glu/Asp-Gly-Arg and provides a unique *Sac I* restriction site. The Factor Xa oligos were prepared such that an overhanging *Xho I* site would be present after annealing the oligo and its complement. After annealing the oligos together they were ligated into the *Xho I* digested pChuA1 vector in accordance with standard ligation procedures (77). Correct ligation was confirmed by the *Sac I* restriction digest and by *Ksp I* and *Sty I* digests whose oligos flank the site of ligation such that the correct size, ~430 bp, can be clearly identified. The resulting plasmid, pChuA2, was transformed, just as before for ChuA.

Expression and purification of ChuA.

Expression of ChuA was carried out in Luria-Bertani broth (2L). Cells were grown to mid-log phase (OD₆₀₀ = 0.5/ml) and expression induced by the addition of IPTG (1 mM). Cultures were grown at 37 °C with constant shaking (250 RPM) for four hours. At four hours post-induction, cells were harvested by centrifugation at 13000 X g

for 15 minutes. The supernatant was discarded and pellets were stored at -80°C until purification.

Expression of ChuA was evaluated by expression analysis assay which utilizes trichloroacetic acid (TCA) to precipitate soluble cellular proteins. In order to accomplish this a volume of cells, sufficient to yield a $0.05 \text{ OD}_{600}/\text{ml}$ reading, was taken from each hourly sample was treated with an equal volume of 10% trichloroacetic acid and centrifuged in a microcentrifuge at maximum rcf for 5 min. The supernatant was discarded and each pellet was washed with 1 ml acetone. The pellets were then dried and resuspended in a minimal amount of SDS-PAGE loading buffer (20 μl), adjusting the pH as appropriate with trizma base. Samples were then separated by SDS-PAGE to allow for visual determination of general protein concentration, size and purity.

Cells harvested from expression were kept on ice during resuspension in 20 mM Tris, 1 mM EDTA, pH 8 and homogenization with a Dounce homogenizer. The chilled sample was French pressed twice at 1000 psi in a small Press cell to lyse the cells (French press, American Instrument Corporation, Santa Ana, CA). All components of the French Press cell and piston were prechilled to avoid protein damage. The lysed cell solution was supplemented with 10 ml of benzonase nuclease (250 U) and 0.001% PMSF and kept on ice. Cell fragments were pelleted at $12000 \times g$ for 20 minutes at 4°C . Due to the insoluble nature of the heme receptor the protein was isolated completely in the insoluble cell fragments of the pellet as verified by SDS-Page analysis of solublized pellet and supernatant fractions.

After centrifugation the pellet was resuspended in 40 ml of TE buffer and frozen for 12 hours at -80°C . The resultant cell pellet solution was thawed on ice and the

protein concentration was determined by standard BCA Protein Assay. The volume of cell solution equivalent to 25 mg of protein was centrifuged at 12000 X g for 10 minutes at 4°C and the remainder of the original cell pellet solution was stored at -80°C. After centrifugation the resulting pellet was resuspended in 5 ml of 0.1M Tris, pH 8. Next, the protein was denatured by the addition of 20 ml of 8 M urea to the cell fragments. Insoluble fragments were pelleted by centrifugation at 12000 X g for 10 minutes at 4°C and the supernatant was applied to 3 ml of equilibrated Ni-NTA resin and incubated overnight at 23°C with constant gentle agitation. The Ni-NTA resin was then collected on a column and washed with ten column volumes of 8 M urea supplemented with 10 mM imidazole. The target protein was then eluted from the column using 8 M urea supplemented with 300 mM imidazole.

The purity of each fraction was estimated by Coomassie-stained SDS-PAGE and the appropriate fractions ($\geq 95\%$ pure) were pooled and dialyzed against Buffer C (20 mM Tris, 200 mM NaCl, pH 8) supplemented with 10% glycerol and 0.1% N-lauroylsarcosine. Dialysis was carried out for 24 hours with multiple changes. The resulting protein concentration was determined by BCA Protein Assay (Pierce), if necessary the sample was concentrated, and stored at -20°C. The $\text{ChuA}^{\Delta\text{HisTag}}$ was expressed by the same procedure as the original heme receptor protein and determined to be in the insoluble pellet fractions. Since it now lacked the six-histidine tag, $\text{ChuA}^{\Delta\text{HisTag}}$ was purified using hemin affinity resin

Cloning and Expression of Wild-type KatG

Cloning of wild-type KatG (wtKatG).

The *katG* gene was amplified from *E. coli* (K-12) genomic DNA using primers ECCPa03 (GTA GAG GGG AGC ACA CAT ATG AGC ACG TCA GAC GAT ATC C) and ECCPa06 (CTG AAC GGG GTC AGA CTC GAG CAG GTC GAA ACG GTC). The resulting PCR product and the pET20b(+) plasmid were digested with *Nde I* and *Xho I*. The appropriate fragments were separated by agarose gel electrophoresis, excised from the gel and extracted with QIAquick Gel Extraction Kit (Qiagen) according to manufacturers' instructions.

The fragments were then ligated according to standard procedures (77) and transformed into *E. coli* (XL-1 Blue) by the heat-shock procedure. Transformants were selected based on ampicillin resistance. The final plasmid (pKatG1) was constructed such that KatG would be expressed with a six-histidine tag. The pKatG1 plasmid was used to transform the *E. coli* expression host (BL-21 (DE3) pLys.S) and transformants were selected on the basis of ampicillin resistance.

Expression of wtKatG.

Expression of wtKatG was carried out in Luria–Bertani broth (2L). Cells were grown to mid-log phase ($OD_{600}/ml = 0.5$) and expression was induced by addition of IPTG (1 mM). At the time of induction, cultures were also supplemented with δ -amino levulinic acid (0.5 mM) and ferrous ammonium sulfate (0.5 mM). Cultures were grown at 37 °C with constant shaking for four hours and growth was monitored hourly at OD_{600} . At 4 h, post-induction, cells were harvested by centrifugation (13,000g), and the cell pellets were stored at -80 °C. Expression of wtKatG was evaluated by expression analysis assay as described for ChuA above.

Purification of wtKatG.

Stored cell pellets were resuspended in Bugbuster reagent supplemented with 0.1 mM phenyl-methylsulfonyl fluoride. Following homogenization, benzonase (250 U) was added, and the mixture was incubated at 23 °C with stirring for 1 h. The cell lysate was then centrifuged at 16,000g. The resulting supernatant was loaded onto a Ni–NTA column by recirculating the solution through the column bed overnight (1 mL/min). The column was then washed with 50 mM phosphate buffer, pH 8.0; 200 mM NaCl supplemented with 2 mM imidazole and KatG was eluted from the Ni-NTA column by a linear gradient ranging in imidazole concentrations from 2 mM to 100 mM in the above buffer. Imidazole was removed by gel filtration. Fractions were screened for purity by SDS-PAGE and by catalase and peroxidase activity to determine the fractions containing

active protein of greater than 90% purity. Fractions meeting these criteria were combined and concentrated by ultrafiltration (30,000 MW cutoff) and stored at -80°C.

Construction and Expression of KatG^{Nterm} and KatG^{Cterm}.

Construction of KatG^{Nterm} Expression Plasmids.

A construct for the expression of KatG lacking its C-terminal domain (KatG^{Nterm}) was produced using primers ECCPm08(+) (GCC GAA AGA AGA TCT GCT CGA GTA AGA TCC GCT GCC) and ECCPm08(-) (GGC AGC GGA TCT TAC TCG AGC AGA TCT TCT TTC GGC) to introduce a Xho I restriction site and a stop codon following codon 432 (Leu) (QuikChange™, Stratagene). The resulting construct was heat-shock transformed into *E. coli* (XL-1 Blue MRF') non-expression host cells, spread onto LB-agar plates supplemented with ampicillin and incubated overnight at 37 °C. Transformants were screened based on ampicillin resistance and plasmid DNA extracted from the candidate colonies was subjected to diagnostic restriction digest using *XhoI* restriction endonuclease.

Digested fragments were separated by agarose gel electrophoresis and the band corresponding to the pET20b plasmid plus the codons of the N-terminal domain was excised, extracted from the gel and religated using T4 DNA ligase. The product was then used to transform *E. coli* (XL-1 Blue) and candidate plasmids were evaluated by diagnostic restriction digests, *XhoI* as above, and DNA sequence analysis. The resulting expression construct (pKatG^{Nterm}) included codons 1 (Met) through 432 (Leu) followed

by codons for Leu, Glu, and a six-histidine tag. This plasmid was used to transform *E. coli* (BL-21 [DE3] pLysS expression host cells.

Further modification of KatG^{Nterm} was accomplished by using site-directed mutagenesis (as above) to replace active site histidines in KatG^{Nterm} with alanines. We used the pKatG^{Nterm} plasmid as a template. Primers ECCPm102(+) (CGT ATG GCC TGG GCC GGC GCG GGG AC) and ECCPm102(-) (GCA TAC CGG ACC CGG CCG CGC CCC TG) were used to produce the H106A substitution. Primers ECCPm101(+) (CTG ATT GCG GGT GGC GCC ACG CTG GGT AAA ACC) and ECCPm101(-) (GAC TAA CGC CCA CCG CGG TGC GAC CCA TTT TGG) were used to produce the H267A substitution. Transformations, restriction digests, and sequence confirmation were carried out as for wtKatG. Plasmids verified to carry the correct mutation were used to transform *E. coli* (BL-21 [DE3] pLysS) to be expressed with a C-terminal six-histidine tag.

In order to render a KatG^{Nterm} mutant in which both histidines were substituted with alanine the primers ECCPm102(+) and ECCPm102(-) with the template DNA of H267A were used to produce KatG^{Nterm}, H106A, H267A. Transformations, restriction digests, and sequence confirmation were carried out as described for wtKatG. The final plasmid was then transformed, as above, into the *E. coli* expression host (BL-21 [DE3] gold pLysS) to be expressed with a C-terminal six-histidine tag.

Construction of the KatG^{Cterm} Expression Plasmid.

Generation of expression constructs for KatG^{Cterm} was accomplished using a deletion mutagenesis procedure adapted from the Seamless cloning procedure by Stratagene. With this new procedure the pKatG1 plasmid was used as a template and primers were designed to: 1) amplify pKatG1 excluding the codons corresponding to the N-terminal domain (Ser 2 – Asp 431), 2) include a restriction site for *Eam 1104 I* (underlined) and 3) yield the necessary *Eam 1104 I* overhanging sequence (bold). We used primers ECCPd301 (GCC GGA CTC TTC G ATG GAA GAT CTG ATC TGG CAA GAT CC) and ECCPd302 (CGT CTG CTC TTC T CAT ATG TAT ATC TCC TTC TTA AAG). Due to the fact that *Eam 1104 I* is a type IIS endonuclease it cuts the target DNA downstream of its target recognition sequence resulting in an overhanging sequence of three nucleotides allowing for religation of the strands such that only the codons of the target deletion are removed. Again, transformations, restriction digests, and sequence confirmation were carried out as for wtKatG. Plasmids verified to carry the correct mutation were used to transform *E. coli* (BL-21 [DE3] pLysS) to be expressed with a C-terminal six-histidine tag.

Expression and Purification of Modified KatG Proteins.

Expression of the KatG variants was carried out as described for wtKatG except that expression cultures were not supplemented with δ -aminolevulinic acid or ferrous ammonium sulfate. The pellets from both domains were resuspended in 200 mM phosphate, pH 7.0, supplemented with lysozyme (0.1 mg/ml) and PMSF (0.1 mM).

Following homogenization, using a Dounce homogenizer, the cells in suspension were lysed by sonication with a Branson 250 sonifier at constant output, 3.5 duty for eight on and off cycles at 45 s intervals each. Following sonication, benzonase (500 units) was added to the lysate and the mixture was stirred at room temperature for 1 h. The cell lysate mixture was then centrifuged at 12000 rpm for 20 min. At this point the supernatant was loaded onto a Ni-NTA column by recirculating through the column bed overnight at 1 ml/min and at 4 °C. The loaded column was then washed with 200 mM phosphate, pH 7. Two sequential washes were performed with 50 mM phosphate buffer, pH 8.0; 200 mM NaCl supplemented with 2 and 20 mM imidazole, respectively. The KatG^{Cterm} was then eluted from the column with the above buffer supplemented with 200 mM imidazole. Fractions containing KatG^{Cterm} were combined and the excess imidazole was removed by dialysis against Buffer A, for 24-36 h with buffer changes every 6 h.

Conversely, the KatG^{Nterm} and site-specific variants thereof were expressed in inclusion bodies, and were therefore, among the insoluble materials of the lysate pellet. Because of this, the post centrifuged lysate pellet was resuspended in 8 M urea, rehomogenized and recentrifuged. After the second centrifugation, 8 M urea equilibrated Ni-NTA resin was added to the supernatant and incubated overnight at 23°C with constant, gentle agitation. The Ni-NTA resin was then collected on a column and washed with ten column volumes of 8 M urea supplemented with 10 mM imidazole. The target protein was then eluted from the column using 8 M urea supplemented with 200 mM imidazole. The purity of each fraction was estimated by Coomassie-stained SDS-PAGE and the appropriate fractions ($\geq 95\%$ pure) were pooled and dialyzed against Buffer A for 24 hours with multiple changes. The resulting protein was then reconstituted, (see

below) and stored at 4°C. Concentration of variants was accomplished using Amicon Ultra centrifugal filter units and any buffer changes necessary for spectroscopy were done using 10 DG columns in accordance with the manufacturers' instructions in both cases.

Heme Solution Preparation and Concentration Determination.

A solution of hemin was prepared to a concentration of 0.5 mg/ml in 100mM KOH solution using bovine hemin chloride (Sigma). Heme concentration was determined by either the pyridine hemichrome assay or by absorption maxima ($\epsilon_{39} = 45 \text{ M}^{-1} \text{ cm}^{-1}$). In the hemichrome assay an initially reference spectra of 1 ml assay mix of potassium ferricyanide (0.3 mM) with 10 μl of prepared heme solution in 35% v/v pyridine, and water is measured from 530–560 nm. Following this 5 mg of sodium hydrosulfite was added and a difference spectrum was collected. The concentration of hemin was determined from the height of the Sorret peak in the reduced/oxidized difference spectrum and $\Delta\epsilon_{557-541} = 20.7 \text{ mM}^{-1} \text{ cm}^{-1}$ for the protoheme pyridine hemichrome.

Reconstitution of wtKatG, KatG^{Nterm} and its variants.

Heme binding or reconstitution was accomplished by incubating the protein at 4°C overnight with an amount of hemin necessary to reconstitute the protein a target

equivalency of hemin (0.75 for wtKatG and 0.95 for KatG^{Nterm} and the variants thereof). Reconstitution converts the apo-protein (heme free) to holo-protein (heme-bound) when incubated with hemin. Spectra for a given KatG protein was recorded from 750-250 nm using a Shimadzu UV1601 spectrophotometer (Columbia, MD) before and after incubation with hemin solution. Heme binding was determined by shifts in spectral lineshapes and molar absorptivities.

Heme Receptor Activity Determination.

Heme receptor Protein Heme Binding determination by affinity chromatography.

The capacity of ChuA to bind heme was also accessed through affinity chromatography heme and hemin coupled resin. Purified protein was incubated with heme resin, at a 1:1 concentration ratio, overnight at 4°C. Likewise, a duplicate sample was incubated with hemin resin, at the same ratio. Individual protein/resin samples were washed three times with five column volumes of Buffer B. The protein/resin was then incubated with five column volumes of 1mM heme solution overnight at 4°C followed by elution. Finally, the protein was eluted by adding one column volume of SDS-PAGE loading buffer and incubating at 95°C for five minutes. Heme affinity binding was confirmed by SDS-PAGE.

Recombination and Characterization of KatG^{Nterm} and KatG^{Cterm}.

Recombination of KatG^{Nterm} and KatG^{Cterm} domains.

Recombination of the KatG^{Nterm} and KatG^{Cterm} domains was accomplished by incubating reconstituted KatG^{Nterm} with a given equivalence of KatG^{Cterm} at 4 °C. For EPR studies (below) 100 equivalents of peroxide substrate was added post recombination and was incubated for 1 h at room temperature prior freezing sample.

UV-Visible Absorption Spectra and Activity Assays for KatG proteins.

Spectra for wtKatG and its variants were recorded using a Shimadzu UV1601 spectrophotometer (Columbia, MD) before and after reconstitution with hemin. All proteins were reconstituted with 0.75 equivalent hemin with the exception of KatG^{Nterm}, which was reconstituted with 0.95 equivalent hemin. Catalase and peroxidase assays were performed for the reconstituted proteins. Catalase activity was determined by evaluating the conversion of hydrogen peroxide to water and molecular oxygen by monitoring the decrease in absorbance of H₂O₂ at 240 nm (39.4 M⁻¹ cm⁻¹). All assays were carried out at 23°C in 100 mM phosphate buffer, pH 7.0. Activities were normalized and compared on the basis of heme content. Peroxidase activity was determined by evaluating the oxidation of ABTS (the electron donor) to its corresponding

ABTS radical at 417 nm (34.7 mM⁻¹ s⁻¹). All assays were carried out in 50 mM acetate buffer, pH 5.0. Like catalase, all activities were normalized and compared on the basis of heme content.

Stopped-flow kinetic studies for KatG^{Nterm} proteins.

Binding of CN⁻ by Kat^{Nterm} was monitored using an SX.18MV Stopped-flow Rapid Reaction Analyzer (Applied Photophysics, Surrey, UK) in single mixing mode. Reactions for the KatG^{Nterm} were set up with one syringe containing 2 μM of protein solution and the second contained either NaCN (20 μM to 2 mM) or peracetic acid (1 μM to 100 μM). Spectra were recorded by diode array. Kinetic constants were determined from single wavelength data recorded at 416 nm. All reactions were carried out at 25°C in 200 mM phosphate buffer, pH 8.0. Data were fit to one site binding curve $y = A e^{-k_{\text{obs}} t}$ and amplitude was fit to the equation for relative amplitude; $\text{Rel. Amp.} = A_{\text{max}} \cdot [\text{CN}^-] / K_D + [\text{CN}^-]$.

Circular dichroism (CD) measurements for KatG proteins.

Far UV CD spectra (188 – 300 nm) were recorded for wtKatG, KatG^{Nterm}, KatG^{Cterm} and combinations of the domain proteins using an AVIV 62DS CD Spectrapolarimeter (Lakewood, NJ) or a Jasco J-810 spectrapolarimeter (Tokyo, Japan).

Preliminary spectra were recorded, with the protein of choice in 10 mM phosphate, pH 7.0 and compared to spectra of protein in Buffer A, for the purpose of accessing buffer interference at higher phosphate concentrations at wavelengths below 200 nm, a protein buffer solution of.

Electron Paramagnetic Resonance (EPR) for KatG and its variants.

EPR spectra were recorded for wtKatG, KatG^{Nterm}, KatG^{Cterm} as well as combinations of the domain proteins using a Bruker EMX spectrometer equipped with an Oxford ESR 900 cryostat and an ITC temperature controller. Spectrometer settings were as follows; temperature, 10K; microwave frequency, 9.38 GHz; microwave power 0.1 mW; modulation amplitude, 10G; modulation frequency, 100 kHz; time constant, 655.36 ms; conversion time, 655.36 ms; and receiver gain, 1×10^5 . Analysis of spectra for baseline subtraction and assignment of g-values was done using WIN-EPR software and Prism was used to reproduce spectra and magnify low-spin regions.

Crystal Screening for KatG^{Nterm}.

Preliminary crystal screens were done by Hauptman-Woodward Medical Research Institute (Buffalo, NY). Candidate salt/buffer crystallography conditions which were identified and set up for optimization using the sitting drop method. Crystal setups were

incubated at a constant 23 °C, in Fisher Brand Isotemp® Model 655D incubator, for 6 weeks and crystal growth was evaluated at time intervals of 24 hr, 1 wk, 2 wk, 3wk and 4 wk during the six weeks. Of the preliminary screen conditions, two well solution salt and buffer conditions were identified for optimization. First a 0.2 M lithium bromide concentration gradient from 5.0 to 6.0 M in a buffer solution of 0.1 M sodium acetate, pH 5.0 and second a 0.025 M calcium chloride dihydrate concentration gradient from 0.025 to 0.15 M supplemented with 20% w/v PEG 4000 also in 0.1 M sodium acetate, pH 5.0. Each of the above conditions was evaluated with salt/buffer only controls which contained no protein, in either the 2 µl drop or the well solution, and a standard crystallography lysozyme control in addition to the KatG^{Nterm} protein crystal setups. A volume of 1µl KatG^{Nterm} protein solution was diluted with an equal volume of the corresponding well solution so as to produce a 2 µl final drop volume and a final concentration of 58.27 µM of KatG^{Nterm}. The lysozyme control insured that conditions were conducive to crystallization based upon the ease at which lysozyme crystallizes. The lysozyme control drop had a total drop volume of 10 µl and a final concentration of 75 mg/ml lysozyme in 0.1 M sodium acetate, pH 4.8. The lysozyme well solution consisted of 1 ml of 10% (w/v) sodium chloride and 25% (v/v) ethylene glycol in 0.1 M sodium acetate buffered at pH 4.8. The sitting drop plates were prepared by first pipetting the correct solution in the well followed by pipetting the drop into concave depression of the sitting drop post. The wells were covered and sealed with crystal clear tape.

III. RESULTS

Heme receptor

The heme receptor protein functions in the transport of heme and other iron sources from the host environment into bacterial cells for use in proteins essential for the bacteria's growth and function. It is not surprising then that the heme receptor is found in many pathogenic bacteria, one of which being *E. coli*. Very little is known about the structure and heme transport mechanism of this integral membrane protein. One of the difficulties with investigating such proteins is that they are naturally found transversing the lipid membrane and are notoriously insoluble and, therefore, very difficult to purify. Considering the essential need for iron, the lack of a way to transport it into the cell would compromise the function of the bacteria, perhaps to the point of killing it. A better understanding of the structure and mechanism of transport could provide information which could be used to develop host defense strategies and possibly anti-bacterial agents.

Purification and heme binding assesment

The purified heme receptor protein was found to be in insoluble inclusion bodies, by SDS-PAGE analysis. This result was not surprising considering that the heme receptor is an integral membrane protein and, therefore, it was found to be expressed in inclusion bodies and was not soluble in typical purification buffer solutions. In order to purify a normally insoluble protein a new denatured purification method (delineated in methods and materials) was developed which utilized urea to denature the insoluble ChuA, brining it into solution and thereby allowing for Ni-NTA purification and subsequent refolding of the protein by dialysis. In order to determine the activity of the protein, it the extent to which the heme receptor protein could bind heme and in what form, protein-bound or free, was evaluated. A chromatography method was established based on affinity of the heme receptor protein to hemin (Fe^{3+} heme) and hemoglobin resins. In order to eliminate non-specific or adventitious protein binding several buffer washes and a hemin wash were designed into the procedure.

The protein flow-through, washes, hemin-solution flow-through and SDS loading buffer eluant were all subjected to SDS-PAGE analysis to identify the presence of the heme receptor. Results revealed a protein band corresponding to that of the ChuA protein eluted in the final SDS-PAGE buffer wash but the protein was not seen in any of the previous washes. While difficult to quantify, the fact that the heme receptor protein was eluted only after addition of the SDS loading buffer but not with the washes, nor with the hemin solution, indicates that the heme receptor was bound very tightly to both the hemin and hemoglobin resins.

Since our expressed and isolated ChuA included a six-histidine tag, for purification, it must be considered that since histidine can ligate metals and that the histidine tag could ligate to the iron in hemin or hemoglobin resin. The later case is highly unlikely since the heme is buried deep within the hemoglobin protein, which provides strong evidence that the actual ChuA protein is binding hemoglobin and not the histidine tag. However, one would expect that this non-specific interaction would be a weaker interaction than that of the heme receptor itself. In fact, the introduction of excess heme would offset the interaction between the histidines and the resin, much like imidazole does in Ni-NTA purification. Since no protein was eluted by the hemin solution wash this indicates that the interaction was, in fact, between the protein and the hemin and hemoglobin affinity resins, indicating that the heme receptor protein was in a active conformation.

Histidine tag deletion

In order to eliminate the possibility that the His-tag was binding to the heme and hemin resins rather than the target protein, a second heme receptor protein, ChuA^{ΔHisTag}, was produced using seamless cloning techniques and expressed. This ChuA^{ΔHisTag} protein was then purified with hemin and hemoglobin affinity resins, thus allowing for concurrent evaluation of heme and hemoglobin binding. The protein was expressed and isolated, to a high degree of purity, therefore, indicating that ChuA^{ΔHisTag} bound hemin and hemoglobin. While the ChuA^{ΔHisTag} protein bound hemin and hemoglobin, as did

ChuA with the His-tag, the amounts of purified protein obtained were lower than that for the Ni-NTA purification method.

Due to the desire for larger quantities of ChuA a system was employed which is routinely used to exploit the high level purification of Ni-NTA but to remove the His-tag post-purification, thereby, eliminating the possibility of interaction of the six-histidine tag with metals. This would allow for higher levels of purified protein to be produced, followed by the removal of the His-tag prior to assessing heme and hemoglobin binding. The specific system utilized here involves the introduction of a Factor Xa cleavage site just prior to the His-tag. The Factor Xa cleavage system (Qiagen) is based upon the introduction of the sequence, Ile-Glu-Gly-Arg, which would be cleaved by the site specific endoprotease Factor Xa protease. The oligonucleotide was introduced via seamless cloning to place the cleavage sequence immediately prior to the His-tag. However, thought restriction digest results indicate that the cleavage sequence was properly incorporated, the cleavage reaction resulted in precipitation of the protein.

Though this was the extent of progress on ChuA characterization the ground work laid here is expected to lead to future progress in understanding heme and iron acquisition by pathogens. Moreover, the methodology developed proved invaluable to the progress on catalase-peroxidase characterization reported in this dissertation.

Catalase-peroxidase

Catalase-peroxidases use only one active site to achieve this catalytic versatility. Interestingly, with few exceptions, active site residues between the catalase-peroxidases and the monofunctional peroxidases are superimposable (33, 36). This suggests that protein structures peripheral to the active site play an important role in the unique catalytic abilities of the catalase-peroxidases. In order to better understand catalase-peroxidase's structure and function roles we investigated the repercussions of removal of the domains from each other and then their recombination.

KatG^{Nterm}

Expression and isolation of KatG^{Nterm}

The C-terminal domain of catalase-peroxidase from *E. coli* (KatG) encompasses roughly 300 amino acids. Elimination of this domain was expected to produce a protein (KatG^{Nterm}) containing 440 amino acids and a molecular weight of 49 kD KatG^{Nterm}. The expression analysis assay for KatG^{Nterm} proved to be successful but as purification began it became evident that the protein was not soluble. SDS-PAGE analysis of the supernatant and urea solubilized pellet fraction revealed that KatG^{Nterm} was found completely in the pellet fraction and thus had been expressed in inclusion bodies [Figure 34]. Since KatG^{Nterm} was expressed exclusively in inclusion bodies it was isolated in the

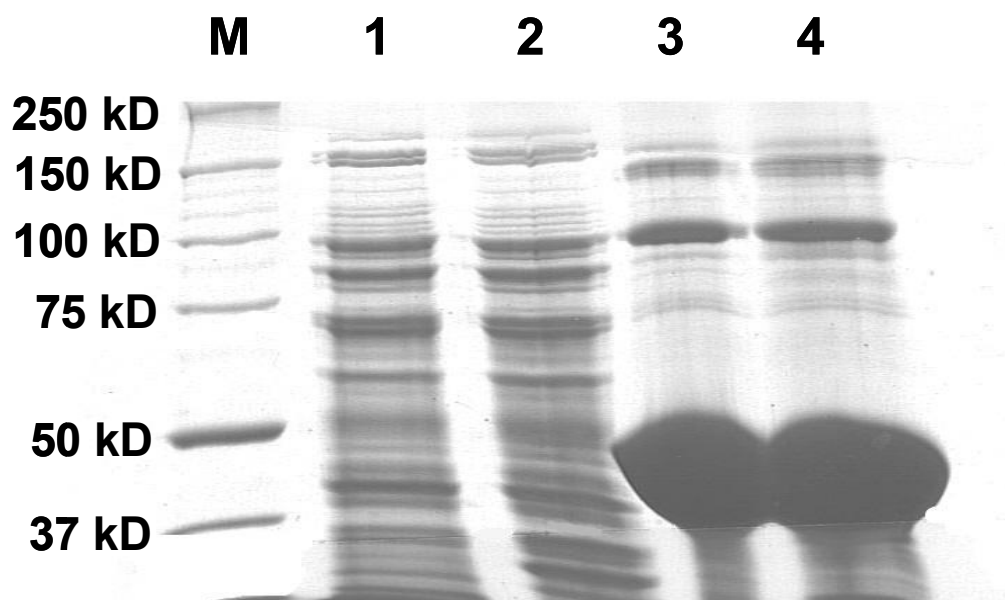


Figure 34. SDS-PAGE gel of KatG^{Nterm} solubility.

presence of 8 M urea via Ni-NTA purification (>95% purity), refolded, and then reconstituted with hemin. A single band corresponding to the expected molecular weight of KatG^{Nterm} was observed by SDS-PAGE, showing that the target protein was successfully expressed and isolated [Figure 35].

Quarternary structure of KatGNterm

The elution profile for native KatG^{Nterm} from a size exclusion matrix (Bio-Select SEC-400, BioRad, Hercules, CA) was dominated by a peak corresponding to a molecular weight of 53.6 kD [Figure 36]. A second peak corresponding to 95.7 kD was observed. These results suggest that the predominant form of KatG^{Nterm} was monomeric with a contribution from a dimeric species (monomer:dimer ~ 4:1). The wtKatG is reported to exist as a tetramer (78).

Evaluation of folded conformation of KatG^{Nterm}

Catalase and peroxidase activity assays revealed that KatG^{Nterm} was completely inactive. Considering that KatG^{Nterm} had no measurable activity and the fact that the protein was purified in a denatured state and then refolded, it was critical to ensure that the protein had been properly refolded and that the active site architecture was intact. Heme absorption spectra recorded for KatG^{Nterm}, though distinct from wtKatG,

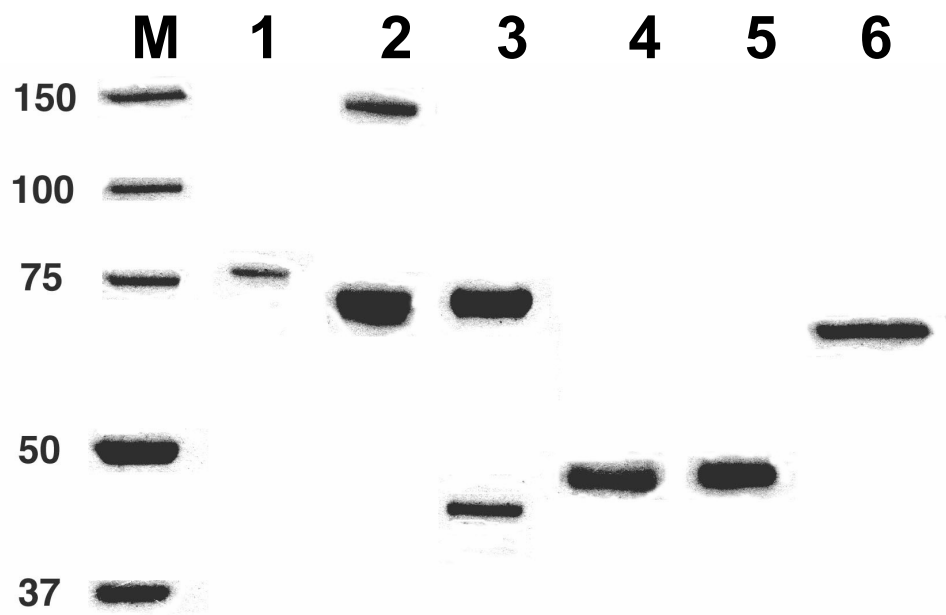


Figure 35. Gel of purified wtKatG and variants. SDS-PAGE of purified wtKatG (lane 1), $\text{KatG}^{\Delta\text{FG}}$ (lane 2), $\text{KatG}^{\Delta\text{DE}}$ (lane 3), $\text{KatG}^{\text{Nterm}}$ (lanes 4 and 5) and $\text{KatG}^{\Delta\text{DE/FG}}$ (lane 6) against a molecular weight marker (M).

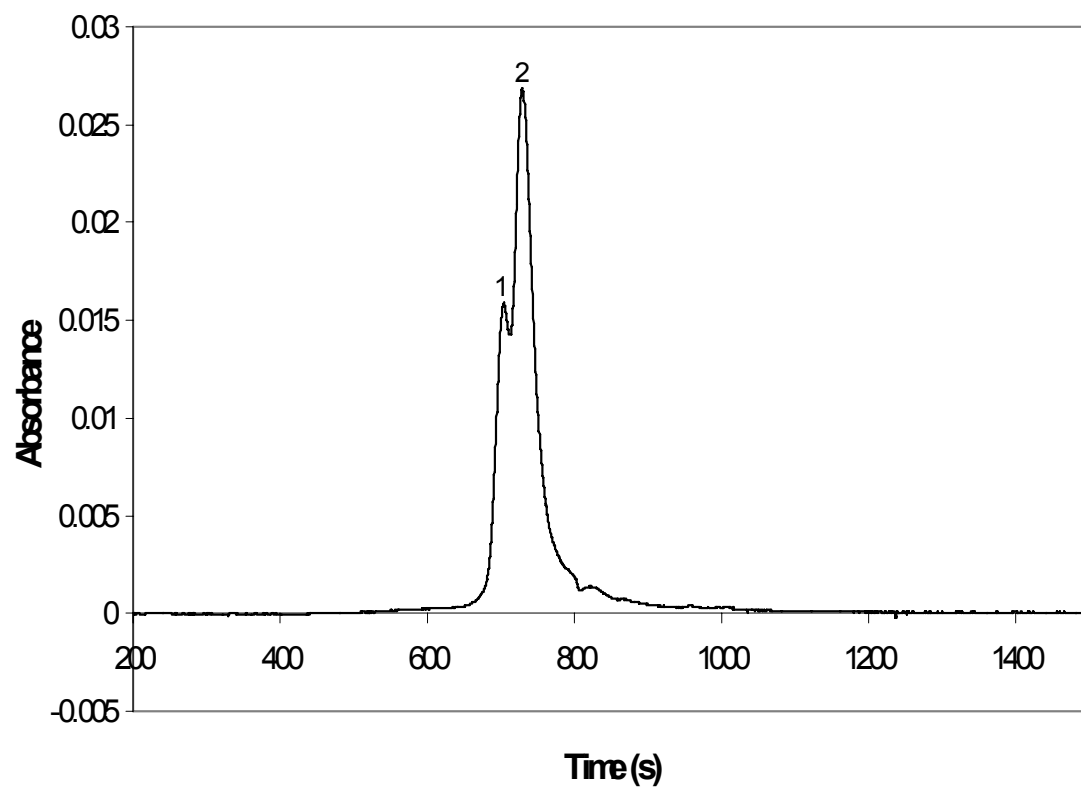


Figure 36. FPLC of KatG^{Nterm}.

were sharp and well defined, contrary to those of hemin either free in solution or adventitiously bound to a misfolded protein [Figure 37]. Additionally, substitution of His 267 with alanine in our KatG^{Nterm} (i.e., H267A KatG^{Nterm}) resulted in a substantial shift in absorption spectrum consistent with a large increase in free hemin. In wtKatG, His 267 is known to be the axial ligand to the heme iron (79). Thus, it appeared that heme bound by KatG^{Nterm} was coordinated within the active site.

Far UV CD spectra showed wtKatG and KatG^{Nterm} to have similar line shapes, but from $[\Theta]_{222}$ values we calculated an 8-10% greater α -helical content for wtKatG than for KatG^{Nterm} [Figure 38]. This is consistent with structural data for catalase-peroxidases showing that the complete monomer (N- and C- terminal domains) has a 6.2% greater α -helical content than the N-terminal domain alone (33). Together, these data strongly suggested that KatG^{Nterm} had an overall folded conformation expected for the catalase-peroxidase N-terminal domain.

3.5 Spectroscopic Properties of KatG^{Nterm}.

Spectra recorded for the ferric KatG^{Nterm} consistently showed a narrow Soret (γ) band with λ_{\max} at 416 nm [Figure 39]. The maxima for the α and β bands were 536 and 568 nm, respectively. Reduction of KatG^{Nterm} with dithionite resulted in a shift in the Soret band to 427 nm and concomitant shifts in the α and β bands to 530 and 560, respectively. These maxima were distinctly different than those of free heme indicating heme was bound, but these maxima are also distinct from wtKatG. Spectra for wtKatG

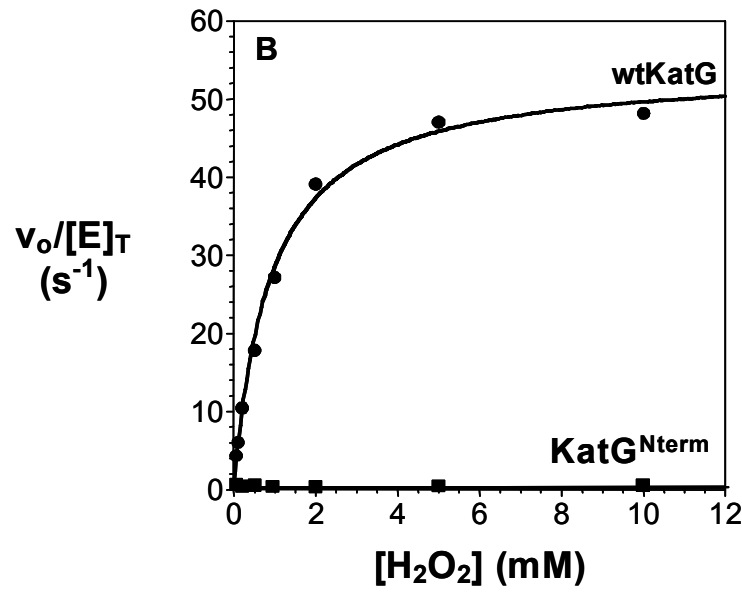
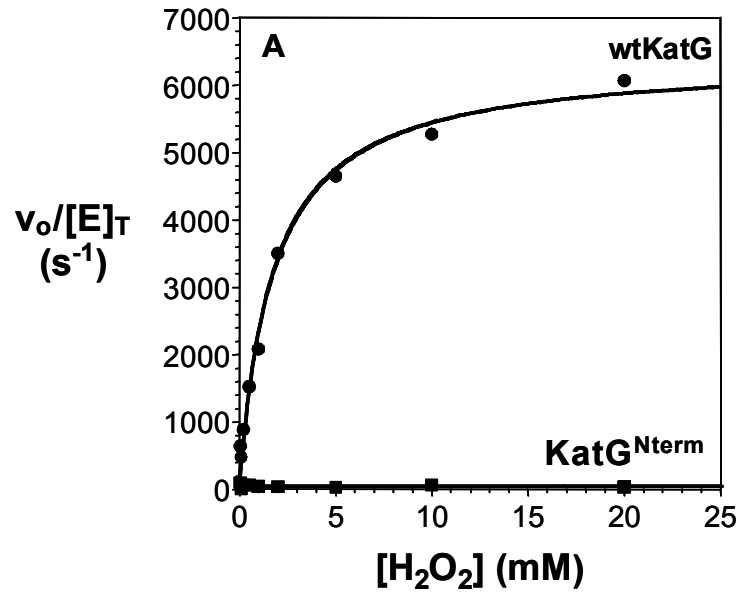


Figure 37. Catalase and peroxidase activities for KatG^{Nterm} and wtKatG. Catalase (A) and peroxidase (B) activity for KatG^{Nterm} as compared to wt KatG.

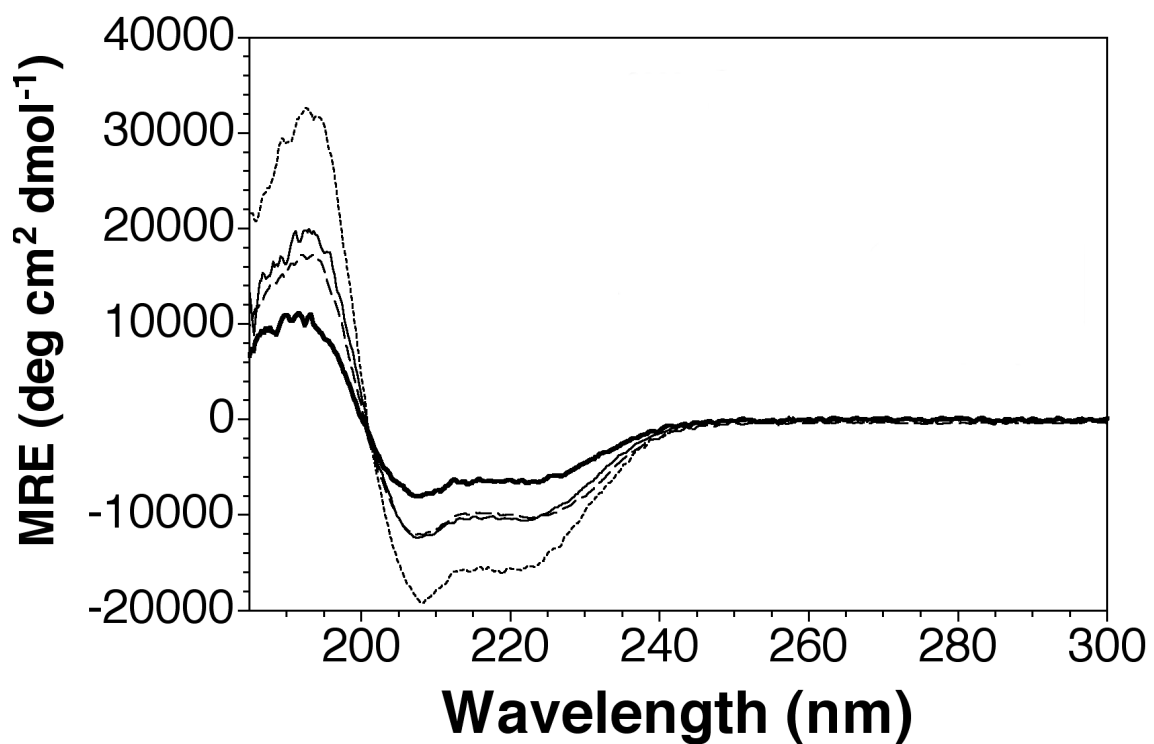


Figure 38. Far -UV circular dichroism spectra for wtKatG and each of its separately expressed and isolated domains. Spectra were recorded for KatG^{Nterm} (a, bold line), KatG^{Cterm} (b, dotted line), wtKatG (c, dashed line), and an equimolar mixture of KatG^{Nterm} and KatG^{Cterm} (d, solid line).

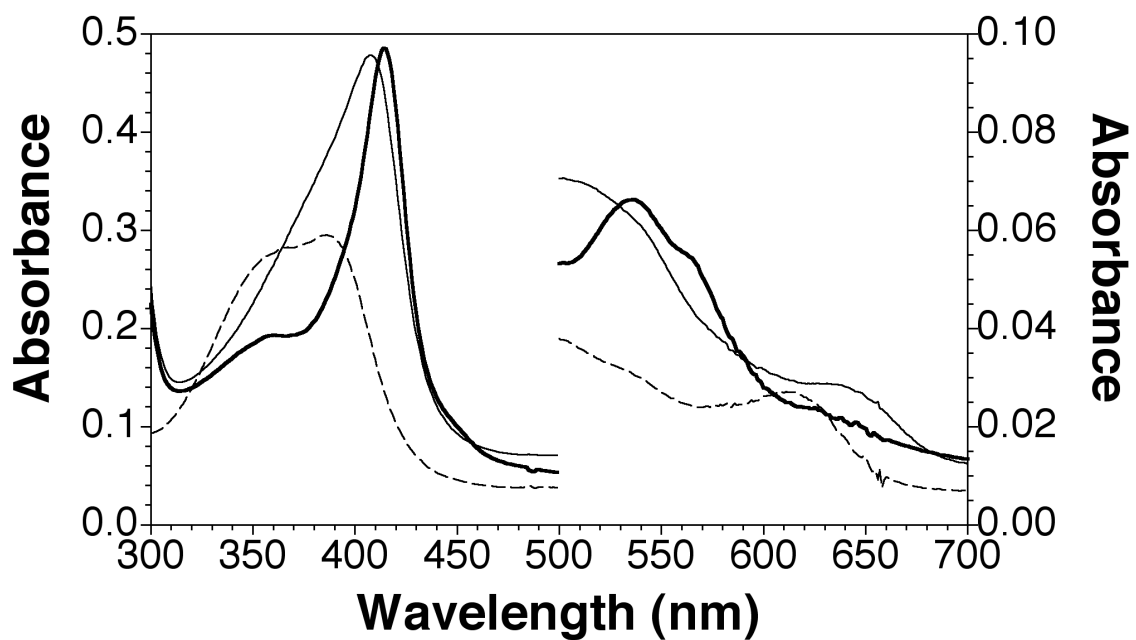


Figure 39. Absorption spectra for the ferric forms of wtKatG, KatG^{Nterm}, and free hemin. Spectra for KatG^{Nterm} (bold line), wtKatG (solid line), and free hemin (dashed line) are shown.

are consistent with a mixture of species dominated by pentacoordinate and hexacoordinate high-spin and quantum mechanically mixed spin species (80).

As compared to ferric wtKatG, the Sorret maximum was blue-shifted 12 nm, and also 21 nm and 31 nm blue-shifts were also observed for the α and β bands (respectively). Also, the molar absorptivity for the α band of reduced KatG^{Nterm} was significantly greater than that of the β band, and these are in stark contrast to those observed for wtKatG [Table 3]. We further examined the spectra for Fe^{III}-CN and dithionite reduced Fe^{II} and found shifts between the wild-type and KatG^{Nterm} which all were consistent with a conversion from primarily pentacoordinate high-spin species, found in wt, to an exclusively hexacoordinate low-spin species seen in KatG^{Nterm} [Figures 40 and 41].

EPR evaluation of KatG^{Nterm}

In order to better understand the change in spin-state and further investigate the coordination of the heme iron, KatG^{Nterm} was evaluated using EPR. The EPR spectrum for ferric wtKatG [Figure 42] was consistent with that of other catalase-peroxidases (80). It is dominated by a heterogeneous mixture of high-spin species as evident from intense signals near $g = 6$ and $g = 2$, both of which are characteristic of a high-spin heme state. Likewise, the predominance of the axial symmetry spectrum indicates high-spin heme ($S = 5/2$). These signals are virtually absent from the EPR spectrum for KatG^{Nterm} but instead, signals at $g=2.9$, 2.3 , 2.1 , and 1.5 are observed, as well as a large increase in the rhombic symmetry. Together these features indicate a hexacoordinate low-spin

Ox. State	Protein	Absorption Band Maxima [nm] (ϵ [$\text{mM}^{-1} \text{cm}^{-1}$]) ^b				
		Soret (γ)	β	α	CT2 ^c	CT1 ^d
Fe ^{III}	wtKatG	408 (121)	- ^e	-	502 (17.5)	629 (9.7)
	KatG ^{Nterm}	416 (98)	536 (12.3)	568 (9.7)		
	LiP ^g	408 (168)	-	-	502 (13)	638 (4.7)
	LiP (Ca ²⁺ -def.)	412 (63)	536 (8.3)	562 (s) ^h	-	-
	LiP (Ca ²⁺ -def.)-	411 (109)	533 (9.4)			
	Im					
	MnP ^f	407 (127)	-	-	504 (8.5)	637 (2.8)
	MnP (Ca ²⁺ -def.)	412 (86)	533 (7.4)	-	-	-
Fe ^{III} -CN	wtKatG	423 (96)	542 (15.6)	-	-	-
	KatG ^{Nterm}	420 (82)	542 (12.1)	-	-	-
	LiP ^g	422 (108)	542 (12)	-		
	LiP (Ca ²⁺ -def.)	420 (77)	541 (9.4)	562 (s)		
	MnP	422 (81)	544 (9.5)	-	-	-
	MnP (Ca ²⁺ -def.)	420 (81)	543 (9.1)	-	-	-
Fe ^{II}	wtKatG	439 (80)	561 (15.1)	581 (10.7)	-	-
	KatG ^{Nterm}	427 (116)	530 (13.6)	560 (17.8)	-	-

^aPhosphate buffer (50 mM, pH 7.0) was used for all spectral measurements.

^b λ_{max} values are in nm, millimolar absorptivities ($\text{mM}^{-1} \text{cm}^{-1}$) are shown in parentheses.

^cCT2, short wavelength charge transfer transition (usually between 490 and 510 nm).

^dCT1, long wavelength charge transfer transition (usually between 600 and 650 nm).

^eAbsorption bands that were not observed or were too weak to make unequivocal assignments are indicated by a dash.

^fSpectral properties of MnP and Ca²⁺-deficient MnP from Sutherland et al. (81).

^gSpectral properties of LiP, Ca²⁺-deficient LiP, and from Nie et al. (81).

^hs denotes a shoulder on a major peak

ⁱSpectral properties LiP (Ca²⁺-def.) with addition of imidazole

Table 3. Heme absorption maxima spectra for wtKatG, KatG^{Nterm}, LiP, and MnP.

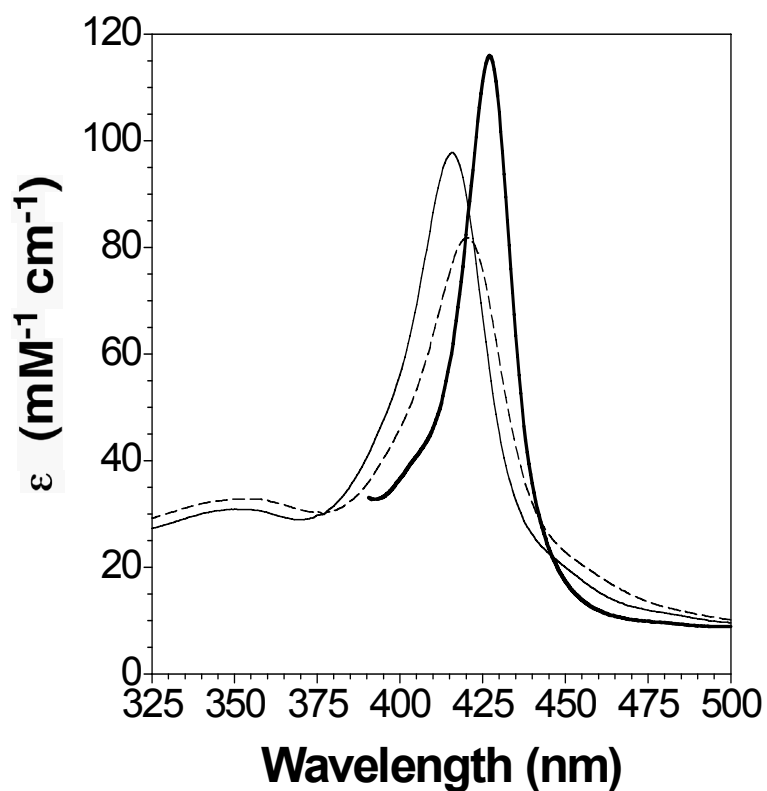


Figure 40. UV-Visible spectra $\text{KatG}^{\text{Nterm}}$ Soret region. Absorption spectra corresponding to the Soret region (300 – 500 nm). Spectra were recorded for ferric (solid line), ferri-cyano (dashed line), and ferrous (bold line) $\text{KatG}^{\text{Nterm}}$. Ferri-cyano $\text{KatG}^{\text{Nterm}}$ was prepared by addition of 2 mM NaCN to ferric $\text{KatG}^{\text{Nterm}}$. Ferrous $\text{KatG}^{\text{Nterm}}$ was obtained by addition of dithionite to ferric $\text{KatG}^{\text{Nterm}}$.

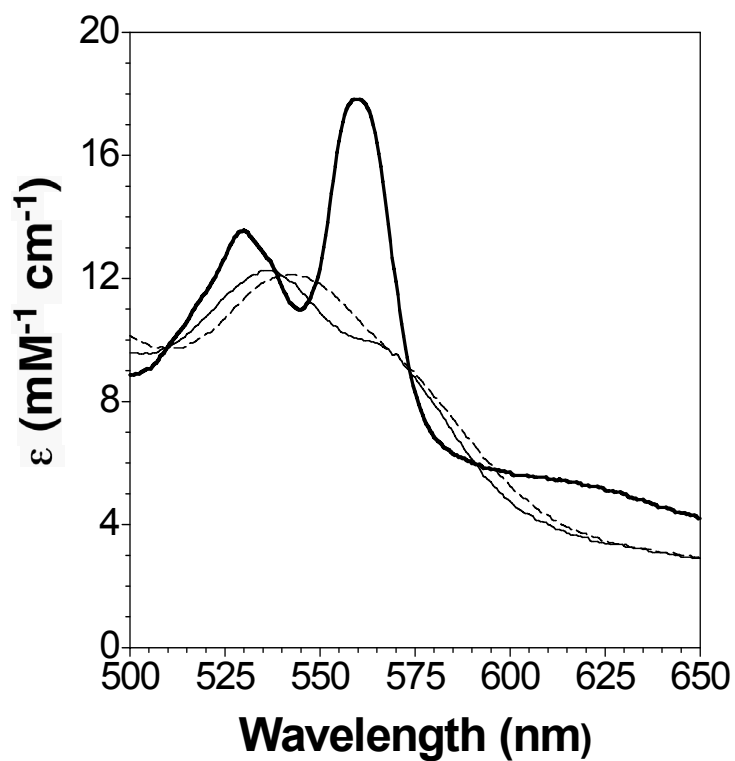


Figure 41. UV-Visible spectra of $\text{KatG}^{\text{Nterm}}$ charge transfer region. Absorption spectra for the α and β transitions (500 – 700 nm). Spectra were recorded for ferric (solid line), ferri-cyano (dashed line), and ferrous (bold line) $\text{KatG}^{\text{Nterm}}$. Ferrous $\text{KatG}^{\text{Nterm}}$ was obtained by addition of dithionite to ferric $\text{KatG}^{\text{Nterm}}$.

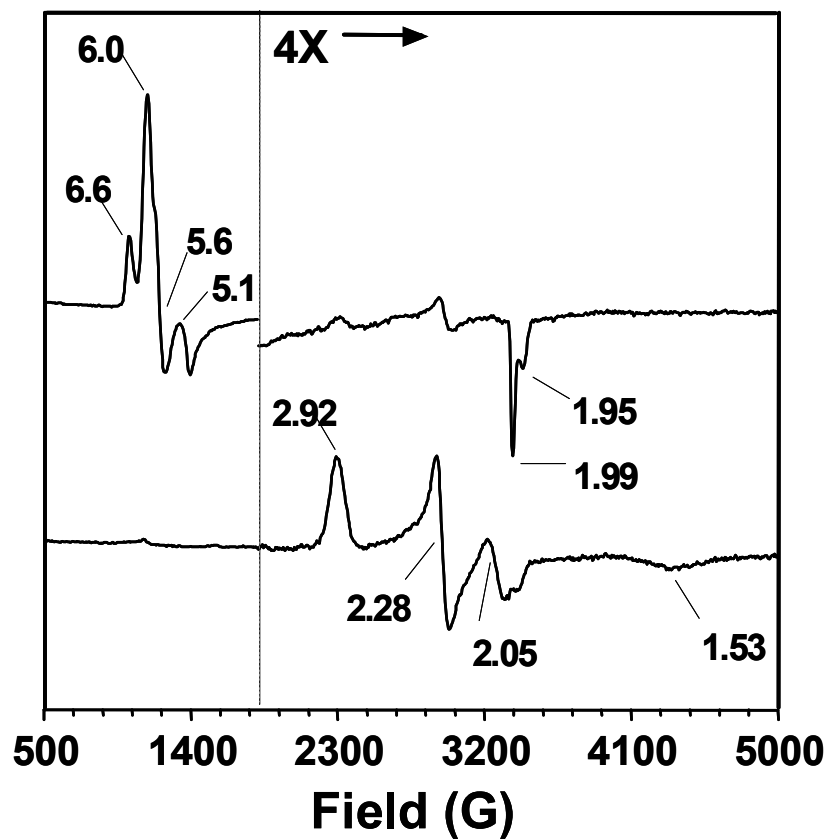


Figure 42. EPR of wtKatG and $\text{KatG}^{\text{Nterm}}$. EPR Spectra of wtKatG (upper spectrum) and $\text{KatG}^{\text{Nterm}}$ (lower spectrum).

($S = \frac{1}{2}$) heme environment, strongly contrasting that of wtKatG. Additionally, these signals strongly suggests histidine as the sixth ligand, with g values around 2.9, 2.26 and 1.6, rather than hydroxide which has characteristic g values around 2.7, 2.2, and 1.86 (82). Interestingly, KatG^{Nterm}, and Ca²⁺-deficient lignin peroxidase (LiP) and MnP (especially the D47A MnP mutant) all share nearly identical EPR spectra. The primary exception is that D47A MnP has a considerably greater population of high-spin species still present (83-85).

Active Site Histidine Substitution.

Because our spectroscopic studies showed the presence of a strong-field sixth ligand (likely histidine) to the heme iron in KatG^{Nterm}, we used site-directed mutagenesis to evaluate the possibility that the distal histidine was this sixth ligand [Figure 43]. Substitution of either His 106 (the distal histidine) or His 267 (the proximal histidine) with alanine led to a shift in the absorption maximum of the Soret band from 416 to 411 nm [Figure 44]. Significant broadening of the Soret band was also detected along with a sharp decrease in molar absorptivity. Additional spectral shifts were detected when both histidines (106 and 267) were substituted with alanine; the λ_{\max} shifted to 387 nm and there was a further decrease in molar absorptivity at 416 nm. In each case, the spectral shifts were consistent with a decrease in specifically bound heme and an increase of free hemin in solution. The effect was most dramatic when both active site histidines were

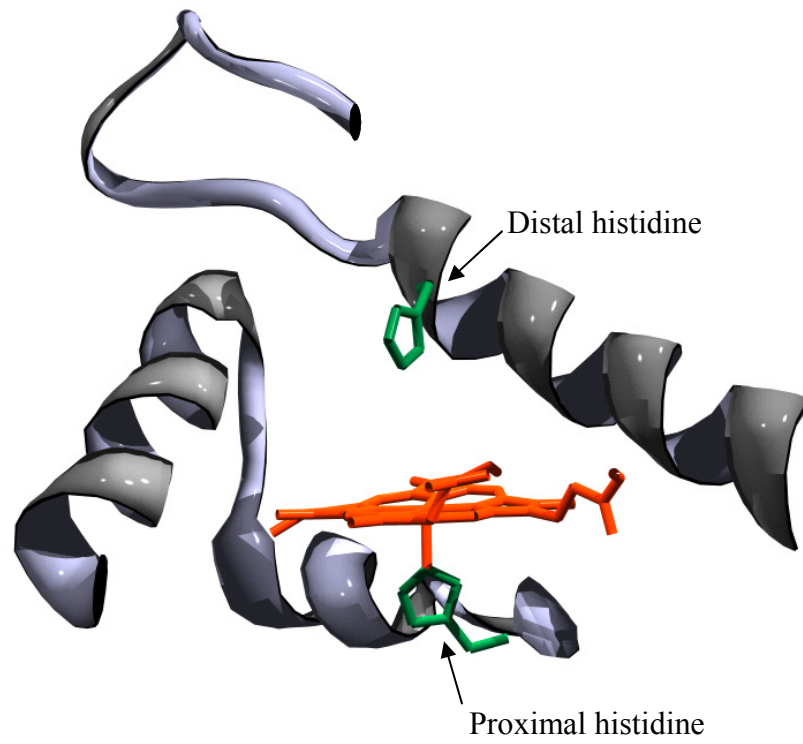


Figure 43. Active site histidines of catalase-peroxidase in MnP.

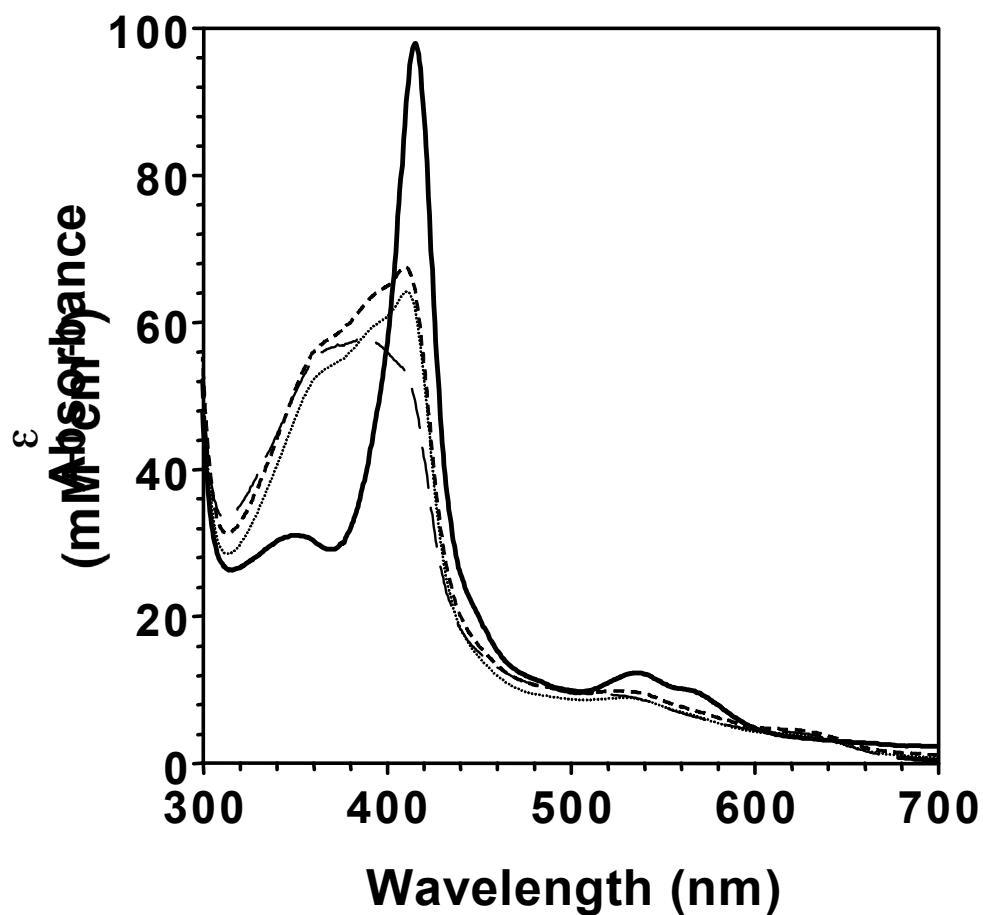


Figure 44. UV-Visible absorption spectra for KatG^{Nterm} and histidine variants.

Absorption spectra of KatG^{Nterm} (solid line), H106A KatG^{Nterm} (dashed line),

H267A KatG^{Nterm} (dotted line), and H106A/H267A KatG^{Nterm} (long dash line).

replaced by alanine. These spectra indicate that the sixth ligand to the heme iron is provided by a histidine in the distal cavity of the active site.

Cyanide Binding.

In order to probe the lack of KatG^{Nterm} activity, substrates were provided to the enzyme in order to gain insight to their binding kinetics. Initially, peroxyacetic acid (PAA) was introduced as a peroxide substrate. Stopped-flow studies showed that PAA resulted in bleaching of the heme. This bleaching occurred to an even greater extent when another peroxy- substrate, m-chloroperoxybenzoic acid, was used [Figure 45]. In view of these results, cyanide binding was used to mimic major steps involved in compound I formation. Consistent with the lack of activity, CN⁻ binding was virtually undetectable for KatG^{Nterm} on the same time scale as wtKatG. At higher [CN⁻] and over longer time periods, there was a slow shift in spectrum from a maximum at 416 nm to 423 nm, consistent with conversion to a Fe^{III}-CN complex [Figure 46]. The conversion followed single exponential kinetics [Figure 47], and the dependence of k_{obs} on CN⁻ concentration was a linear one [Figure 48]. Linear regression analysis of the data produced a rate constant (from the slope) for binding (k_{on}) of 120 M⁻¹ s⁻¹ and a rate constant for dissociation (k_{off}) of 0.01 s⁻¹. Together, these results indicate a K_D for the Fe^{III}-CN complex of KatG^{Nterm} of 83 μM. This was corroborated by monitoring the amplitude of each exponential trace as a function of CN⁻ concentration [Figure 49]. The amplitudes followed a hyperbolic dependence on CN⁻, and from a fit of the data to

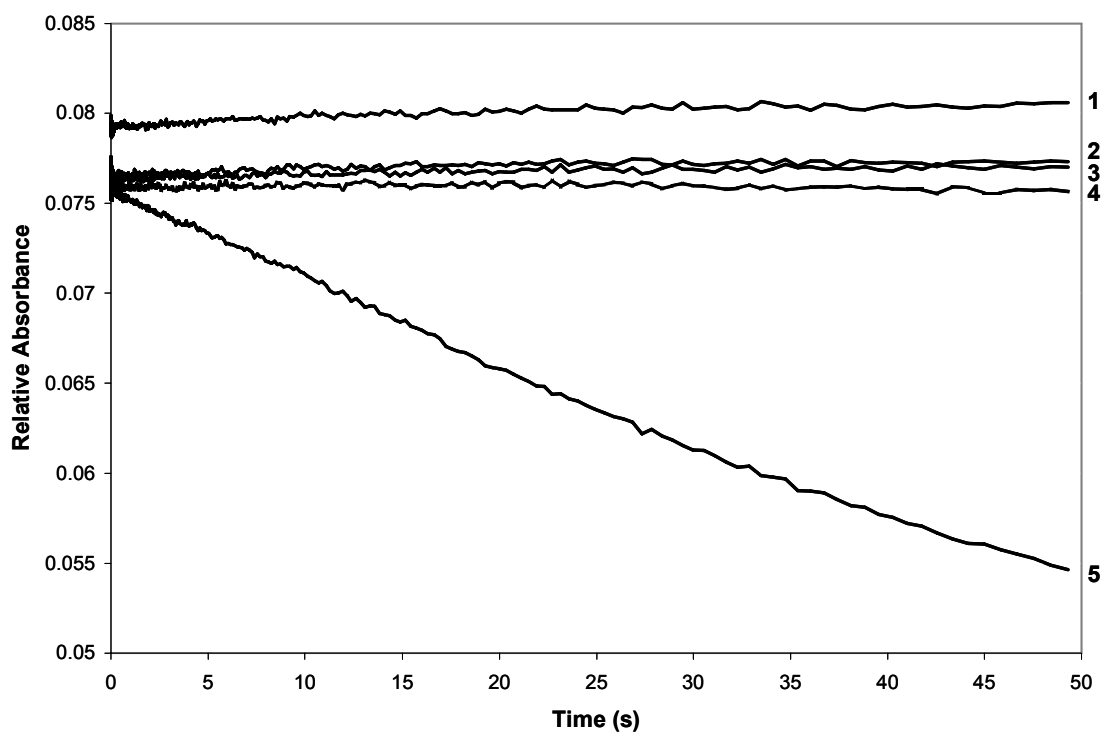


Figure 45. Stopped-flow reaction of KatG^{Nterm} with peroxyacetic acid (PAA). Relative absorbance over time of KatG^{Nterm} (trace 1), and reactions of KatG^{Nterm} and 1 μM, 2 μM, 5 μM, and 100 μM concentrations of PAA (traces 2-5, respectively).

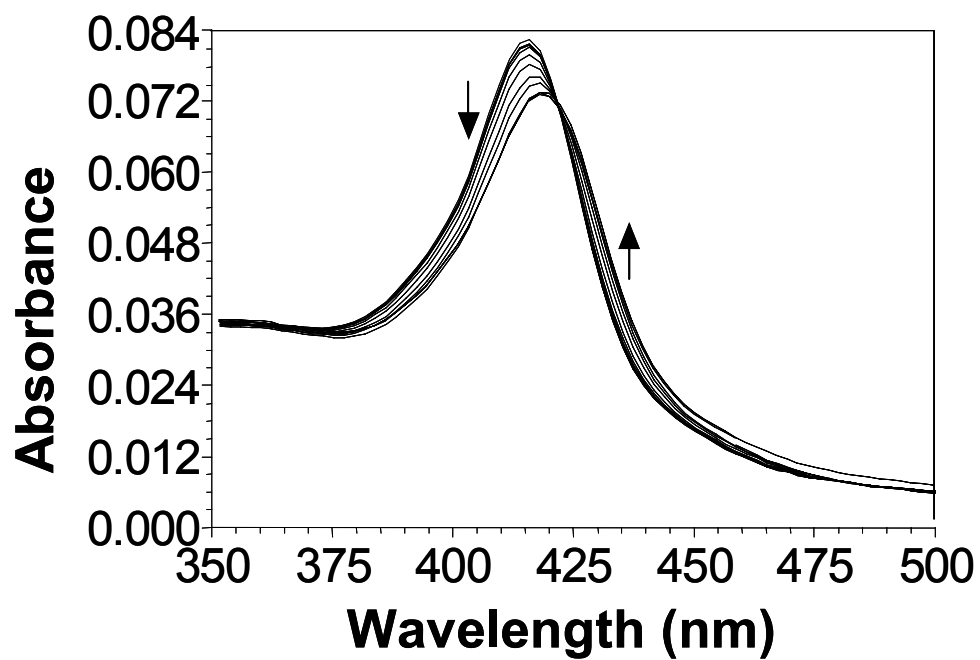


Figure 46. Cyanide binding spectra for KatG^{Nterm}. Spectra corresponding to formation of the ferri-CN complex of KatG^{Nterm} in the presence of 0.5 mM cyanide. The directions of absorbance changes with time are indicated by the arrows.

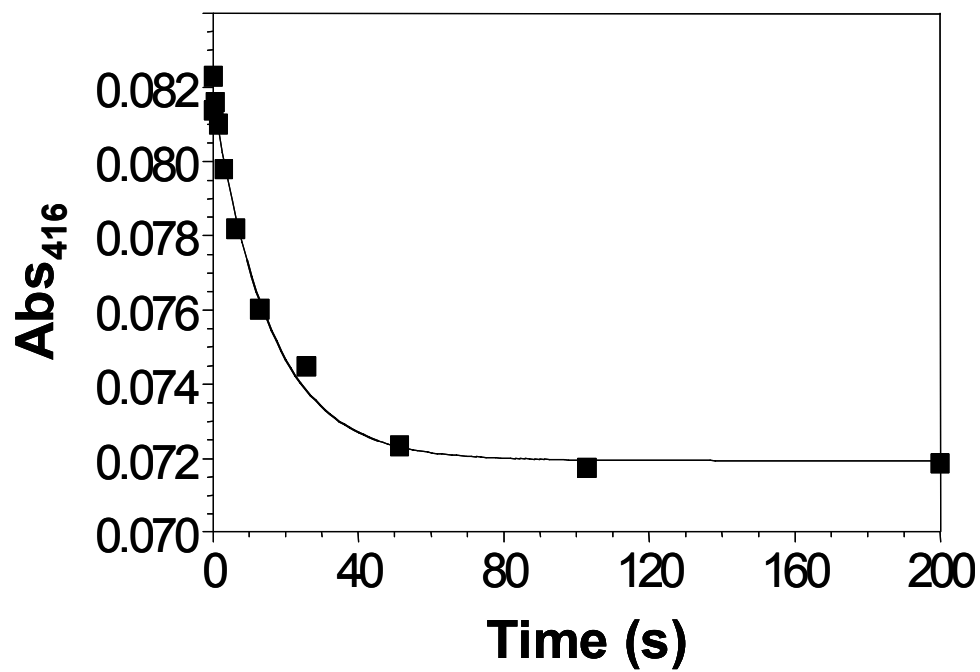


Figure 47. Time course for cyanide binding to KatG^{Nterm}.

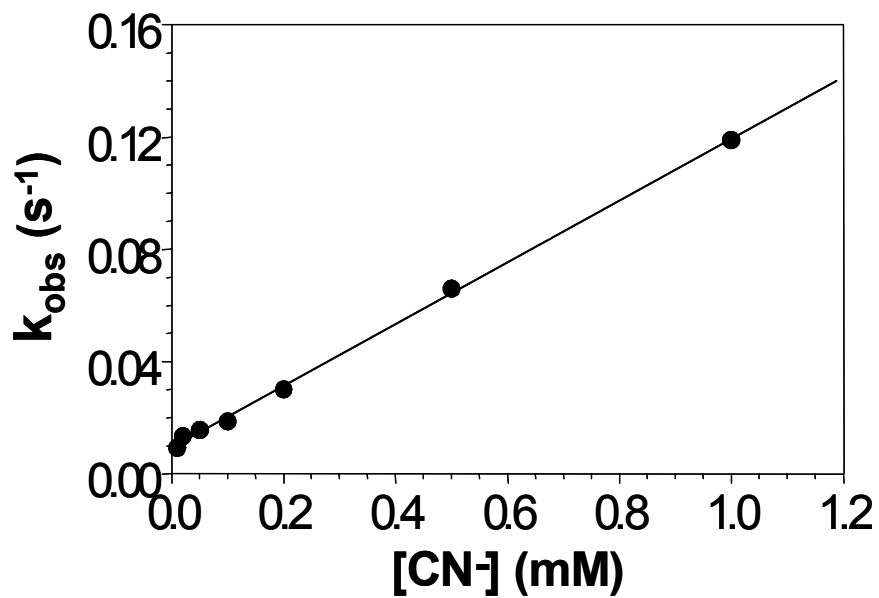


Figure 48. Effect of cyanide concentration on k_{obs} for cyanide binding to KatG^{Nterm}.

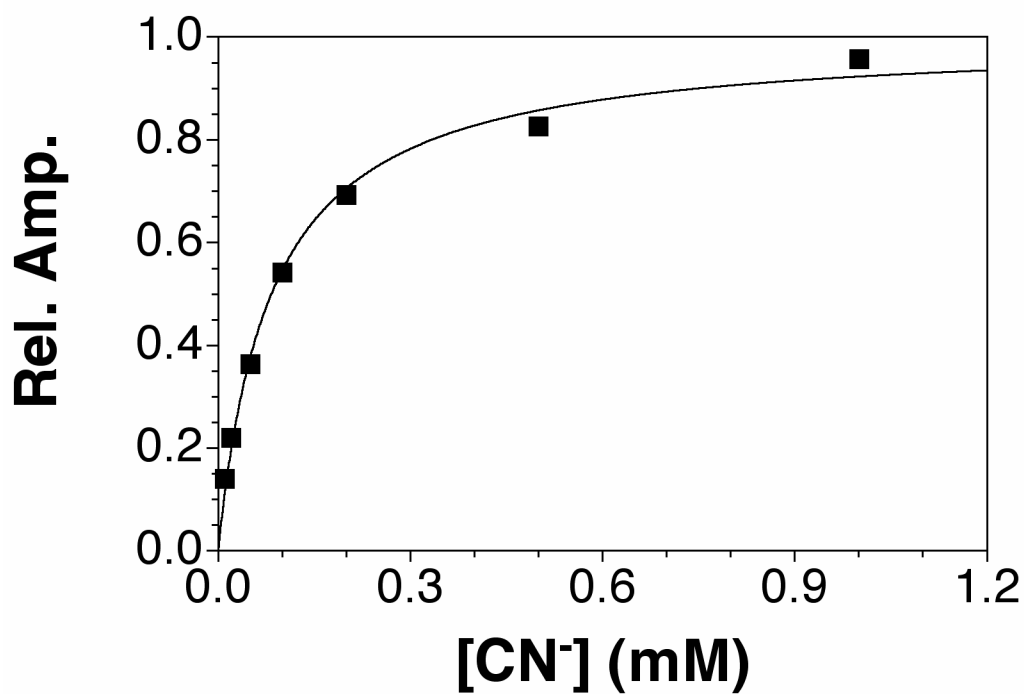


Figure 49. Effect of cyanide concentration on the amplitude of binding traces for $\text{KatG}^{\text{Nterm}}$

equation 2, a K_D value of $82 \mu\text{M}$ was obtained. Conversely, wtKatG binds CN^- according to a k_{on} of $5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$ and a K_D of about $1 \mu\text{M}$.

Comparison to MnP and LiP.

Very similar changes in active site structure and function are observed in certain peroxidases upon loss of Ca^{2+} . In the structures of MnP and LiP, Ca^{2+} is found on both the distal and proximal sides of the active site heme and the distal can be displaced leading to introduction of the distal His to the coordination sphere of the heme iron (81, 83-88) [Figure 50]. This rearrangement is highlighted by a hexacoordinate, low-spin heme complex and the loss of activity in both cases. Inspection of the MnP structure reveals that the distal calcium atom is coordinated to residues from helix B and loop BC. Importantly, helix B bears His 46 (the distal histidine) (87). This residue is a general base required for the catalytic function of peroxidases and catalase-peroxidases (84, 85). While no calcium has been observed in the currently available structures of catalase-peroxidases, the highly similar changes in heme environment and activity observed upon removal of the C-terminal domain suggests that one of its functions may be to support the architecture of the active site, preventing coordination of the heme iron by the distal histidine.

We measured the response of $\text{KatG}^{\text{Nterm}}$ to separately expressed and isolated C-terminal domain, $\text{KatG}^{\text{Cterm}}$. Far-UV CD spectra for $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ indicate

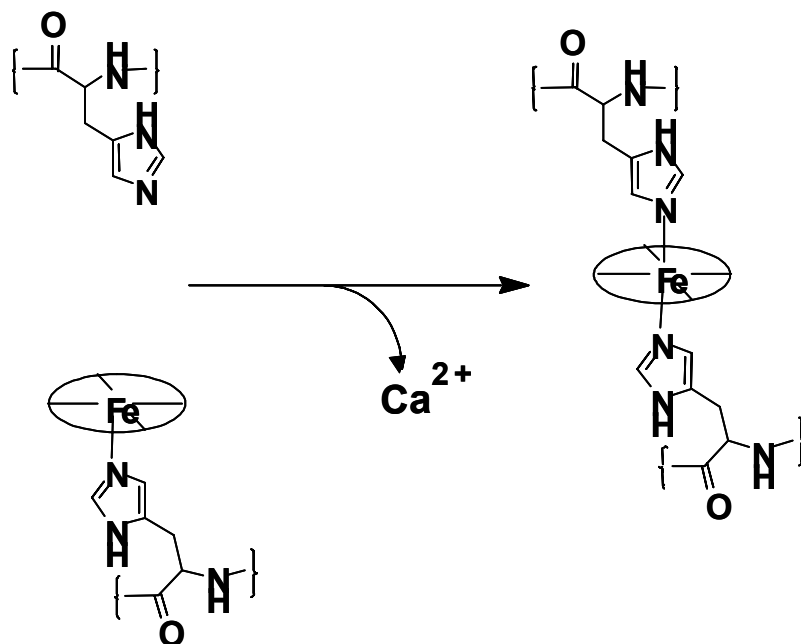


Figure 50. Changes in active site coordination in MnP and LiP upon loss of Ca^{2+} (84, 88).

secondary structural content substantially different from each other and wild-type KatG [Figure 51]. However, a CD spectrum for the physical mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ (1:1 ratio) was highly similar to that of wtKatG. These data suggest that: 1)

preparation of the separated domains, $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$, does not result in major disruption of secondary structural content from that which is expected for each domain from the wild-type enzyme, and 2) the physical mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ did not appreciably change the secondary structural content of either protein. The composite spectrum derived arithmetically from the separate spectra for $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ was superimposable on the spectrum obtained from the physical mixture of the domains [Figure 52]. An equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ had appreciable catalase and peroxidase activities [Figure 53 and 54]. Neither $\text{KatG}^{\text{Nterm}}$ nor $\text{KatG}^{\text{Cterm}}$ had either catalase or peroxidase activities on their own. Furthermore, incubation of $\text{KatG}^{\text{Nterm}}$ with a generic protein such as bovine serum albumin (BSA) did not restore either catalase or peroxidase activity, indicating that specific interactions between the domains were essential for restoration of the catalase-peroxidase active site in $\text{KatG}^{\text{Nterm}}$. Further examination revealed that catalase and peroxidase activities increased with time of coincubation of the two domains [Figure 55]. At 4°C, increases in peroxidase and catalase activities were observed over the first twenty-hours. Interestingly, both activities were observed to return at similar rates. Fits of the observed data to a single exponential equation produced rate constants for reactivation (k_{react}) of $5.6 \times 10^{-5} \text{ s}^{-1}$ and $4.2 \times 10^{-5} \text{ s}^{-1}$ for catalase and peroxidase activities, respectively.

We measured steady-state kinetic parameters for the catalase activity of the equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ following a 48-hour incubation at 4°C. By

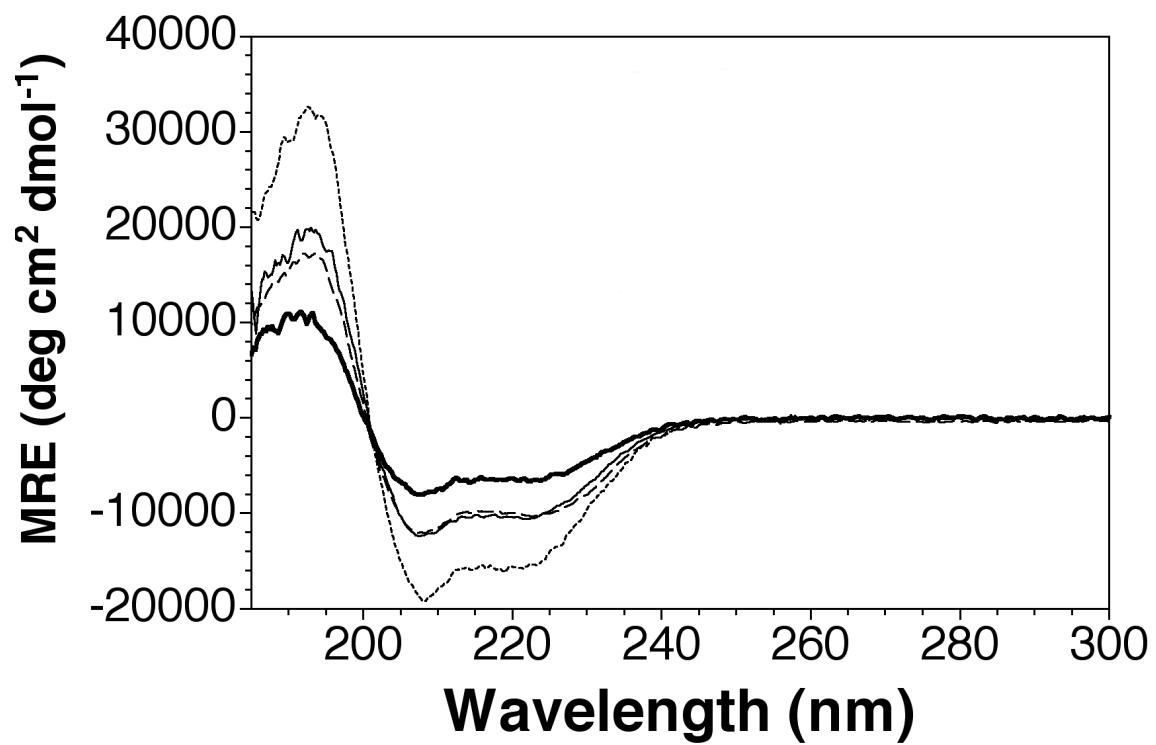


Figure 51. Far UV circular dichroism spectra for wtKatG and each of its separately expressed and isolated domains. Spectra were recorded for KatGNterm (bold line), KatGCterm (dotted line), wtKatG (dashed line), and an equimolar mixture of KatGNterm and KatGCterm (solid line).

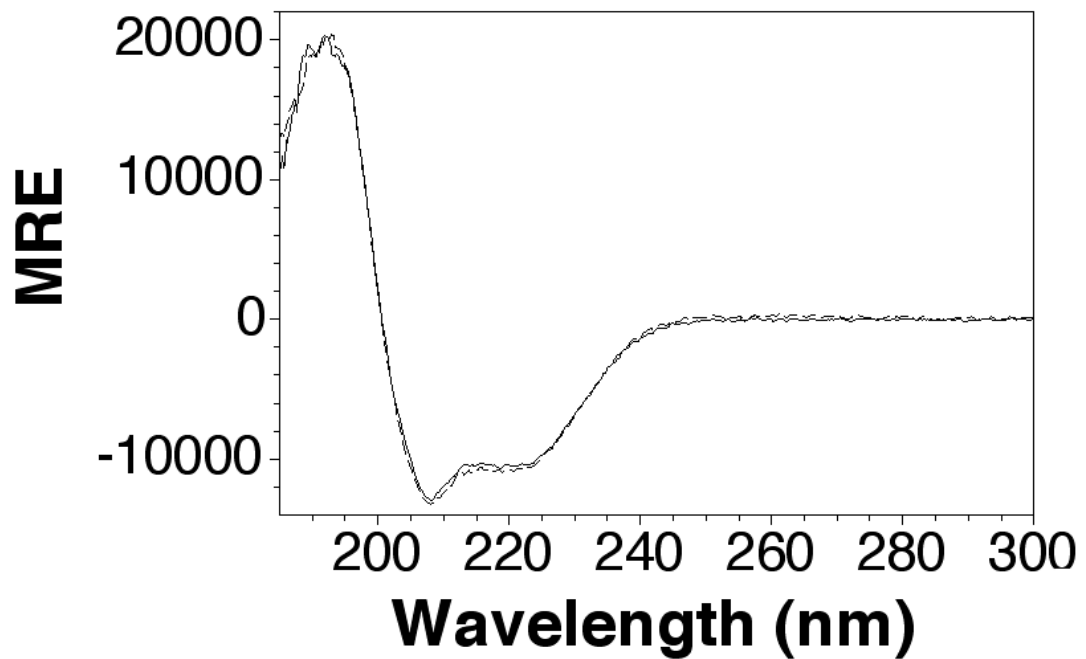


Figure 52. Physical and arithmetic mixtures of KatG^{Nterm} and KatG^{Cterm}. The arithmetic mixture is indicated by the dashed line and the physical mixture (both 1:1) is shown by the solid line.

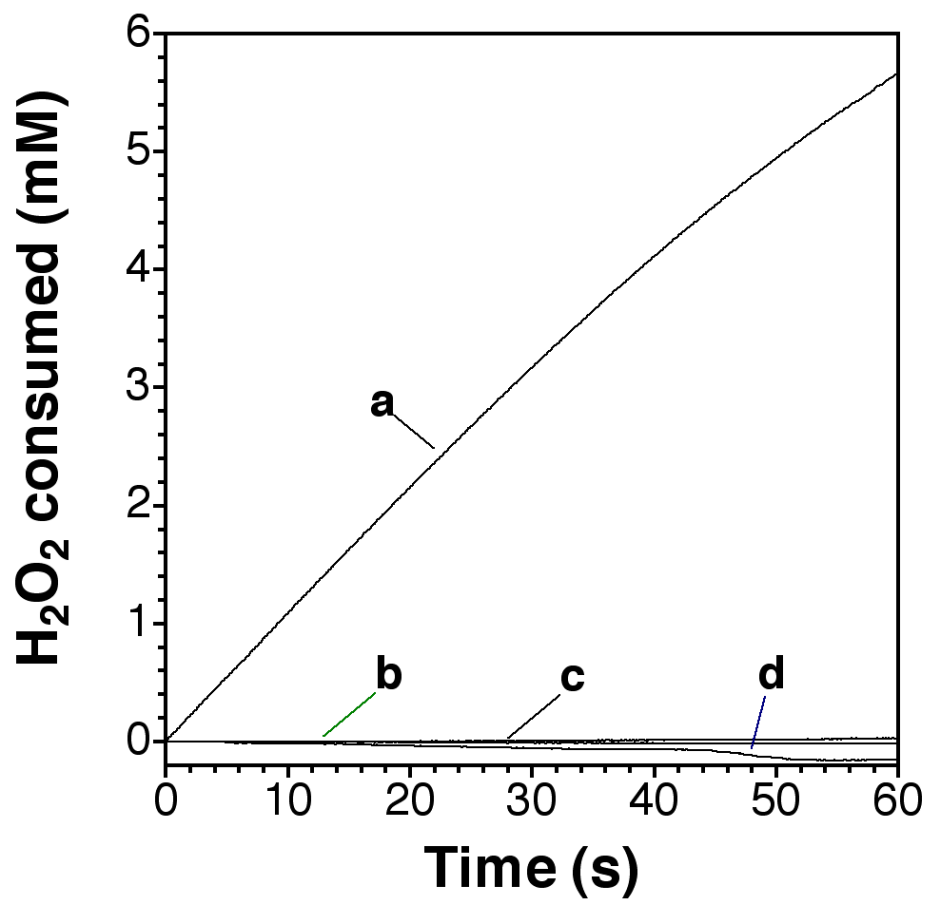


Figure 53. Reactivation of catalase activity for $\text{KatG}^{\text{Nterm}}$ by addition of $\text{KatG}^{\text{Cterm}}$. Catalase activity measured for an equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ (a), $\text{KatG}^{\text{Nterm}}$ alone (b), $\text{KatG}^{\text{Cterm}}$ alone (c), and an equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and bovine serum albumin (BSA).

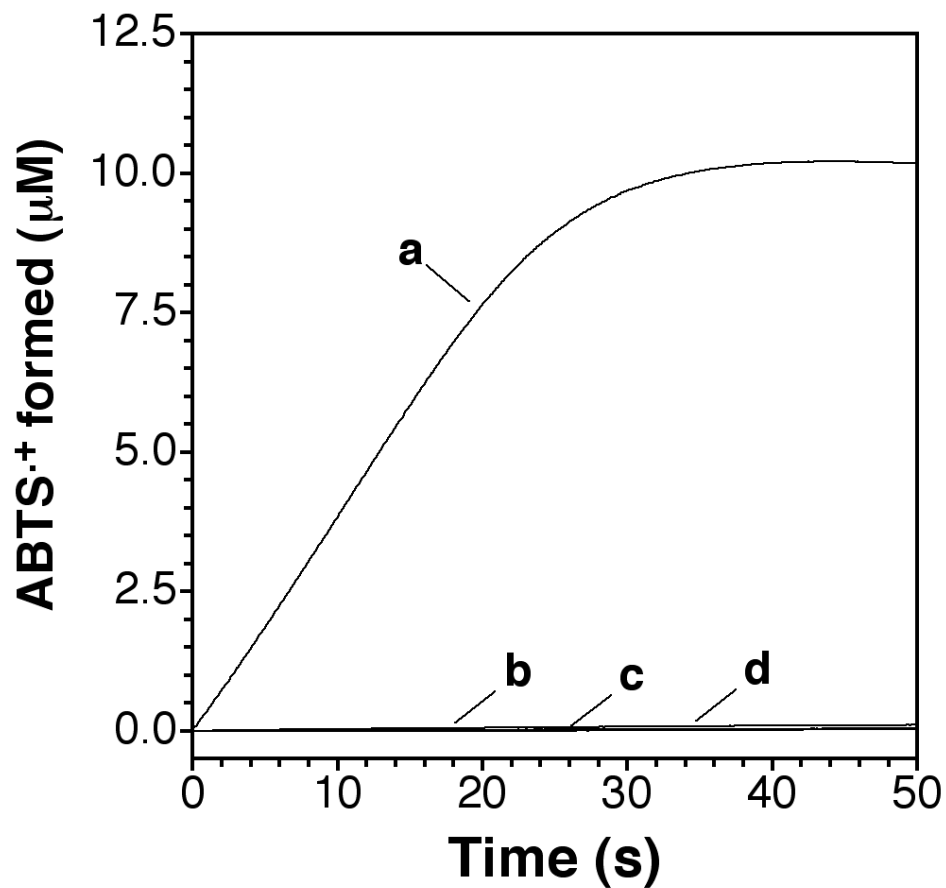


Figure 54. Reactivation of peroxidase activity for $KatG^{Nterm}$ by addition of $KatG^{Cterm}$. Peroxidase activity was measured for an equimolar mixture of $KatG^{Nterm}$ and $KatG^{Cterm}$ (a), $KatG^{Nterm}$ alone (b), $KatG^{Cterm}$ alone (c), and an equimolar mixture of $KatG^{Nterm}$ and bovine serum albumin.

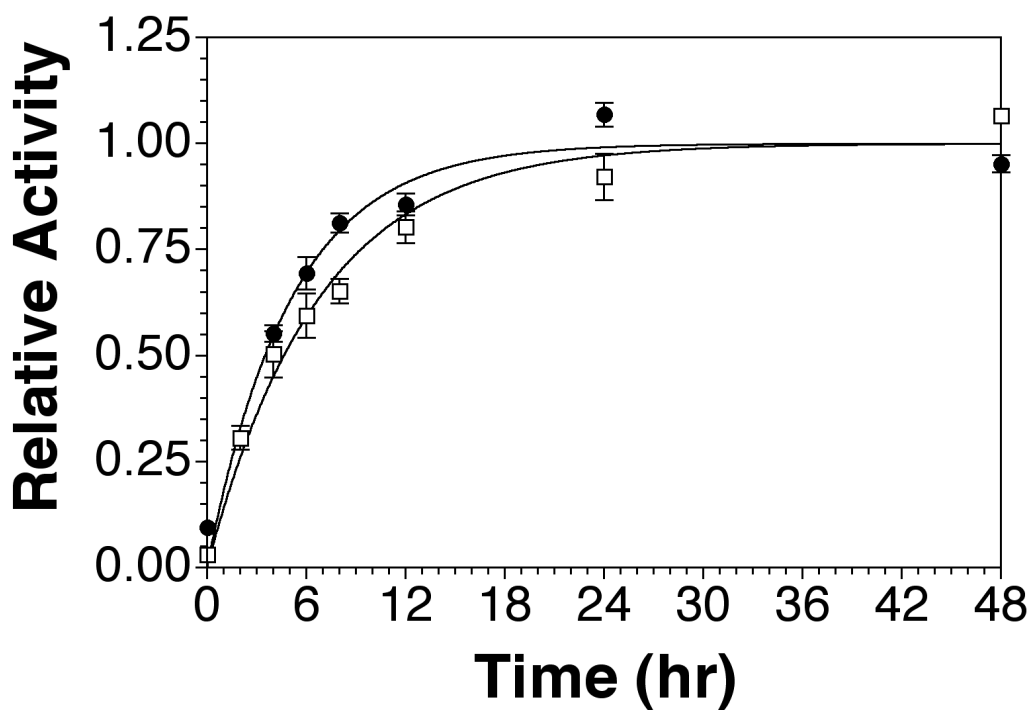


Figure 55. Catalase and peroxidase activities over time for an equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$. Catalase activity (closed circles) and peroxidase activity (open squares) was measured following incubation at the indicated time.

evaluating the effect of H_2O_2 concentration on catalase activity [Figure 56], we obtained a k_{cat} of 1740 s^{-1} . This represents 15% of the value obtained by our laboratory for wild-type KatG (56), and falls between 10% and 50% of the values obtained for catalase-

peroxidases from other organisms. An apparent K_M for H_2O_2 of 3.3 mM was determined, a value nearly identical to that obtained for wtKatG (56) and well within the range for that observed for other catalase-peroxidases. The apparent second-order rate constant obtained for the recovered catalase activity was $5.3 \times 10^5 M^{-1} s^{-1}$. The effect of H_2O_2 concentration on the peroxidase activity of the equimolar mixture of the separately expressed and isolated KatG domains was also evaluated [Figure 57]. At low concentrations of H_2O_2 , increases in peroxidase activity were observed, but a progressive decrease in peroxidase activity was observed at H_2O_2 concentrations exceeding 1 mM. An apparent second-order rate constant for peroxidase activity of $6 \times 10^3 M^{-1} s^{-1}$ was determined and represents 8.5% of the value that we have recorded for wild-type KatG (56).

Visible absorption spectra corresponding to the α , β , and charge transfer transitions for KatG^{Nterm} changed little immediately following the addition of an equimolar concentration of the C-terminal domain [Figure 58]. However, following a 48-hour incubation at 4°C, increases in absorption corresponding to the short- and long-wavelength charge transfer transitions (near 500 nm and 645 nm, respectively) were observed. Decreases in the intensities of the α and β transitions (536 and 568 nm, respectively) were also observed, and a 4 nm blue shift in the Soret (γ) band from 416 nm to 412 nm was also detected (data not shown). These spectral shifts are all consistent with

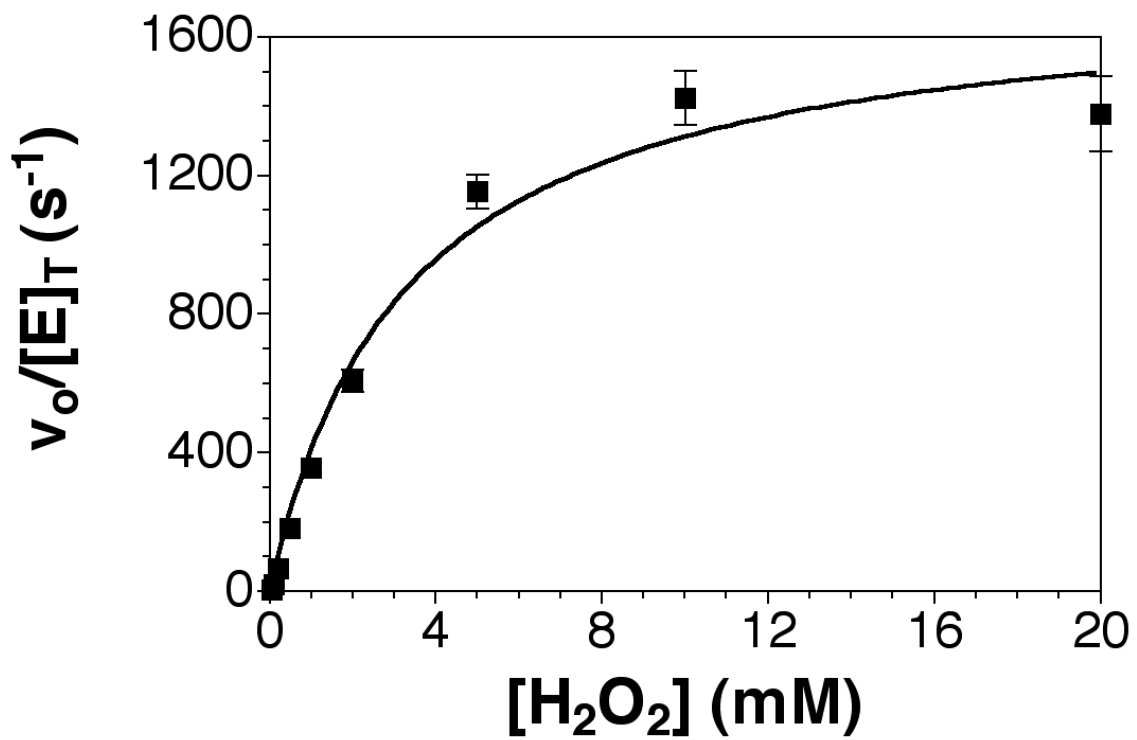


Figure 56. Effect of H_2O_2 on the catalase activity of an equimolar mixture of $KatG^{Nterm}$ and $KatG^{Cterm}$.

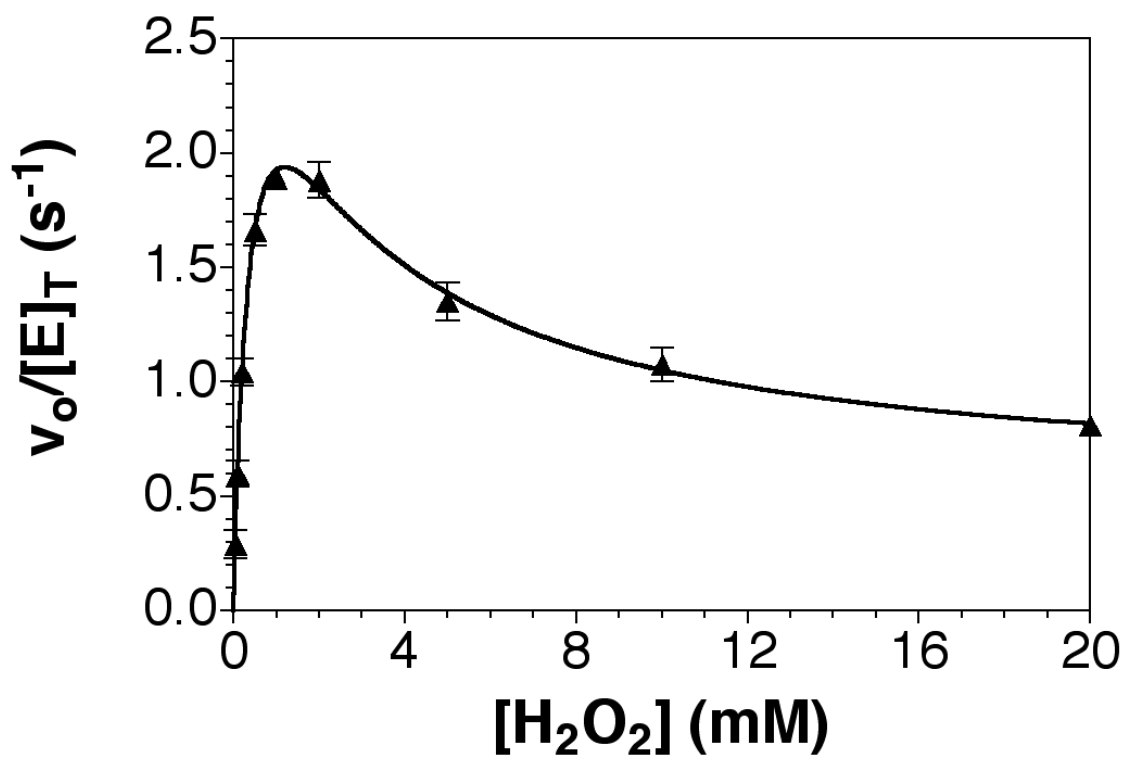


Figure 57. Effect of H_2O_2 on the peroxidase activity of an equimolar mixture of $KatG^{Nterm}$ and $KatG^{Cterm}$.

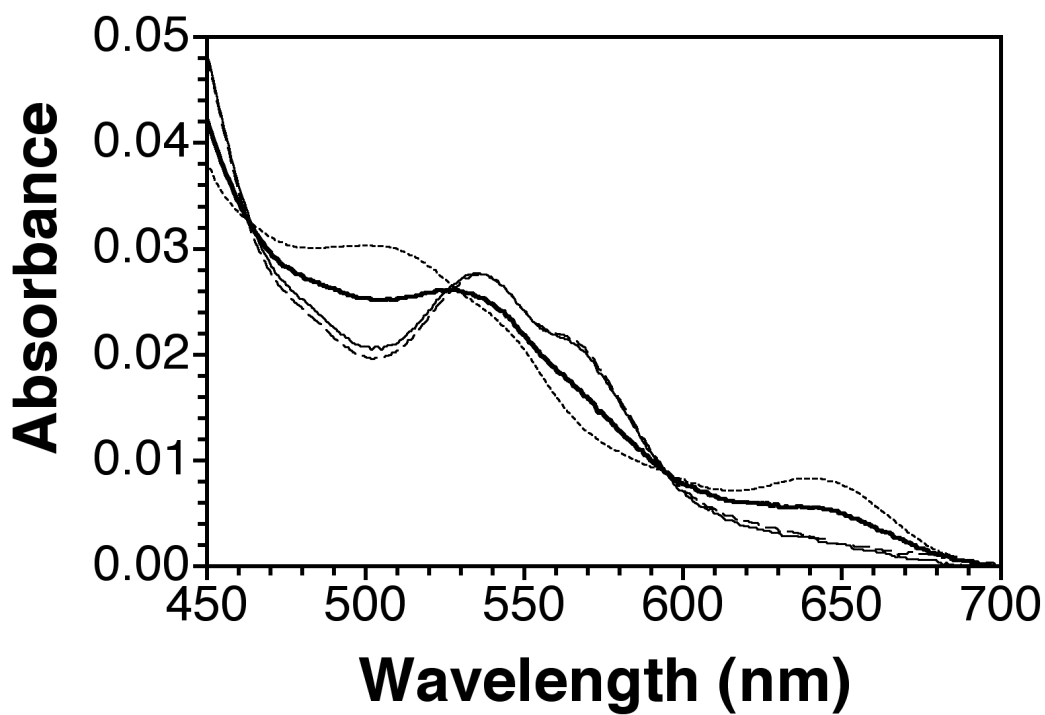


Figure 58. Effect of separately expressed and isolated $\text{KatG}^{\text{Cterm}}$ on the visible absorption spectrum of $\text{KatG}^{\text{Nterm}}$. A spectrum corresponding to the charge transfer and α and β transitions of the heme prosthetic group were recorded for $\text{KatG}^{\text{Nterm}}$ alone (solid line), an equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ immediately after mixing (dashed line), and an equimolar mixture of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ following a 48-hour incubation at 4°C (bold line). A typical spectrum for wtKatG is also shown (dotted line).

a loss of the hexacoordinate low-spin heme iron typical of $\text{KatG}^{\text{Nterm}}$ along with increases in high-spin species (hexacoordinate and/or pentacoordinate) that are typically observed for wild-type KatG. Addition of 100 equivalents of H_2O_2 to recombined samples

showed no significant changes in the spectra for the substrate deficient samples [Figure 59].

In the absence of the KatG^{Cterm}, the EPR spectrum of KatG^{Nterm} is dominated by signals consistent with a hexacoordinate low-spin heme complex [Figure 61-spectrum a]. Indeed, the g_o (2.92), g_{ee} (2.28), and g_o (1.53) values obtained from the spectrum suggest that the strong-field sixth ligand in the coordination sphere is a histidine residue. Data reported earlier in this dissertation [Figure 44] suggests that this histidine is the so-called distal histidine within the active site of catalase-peroxidases. Addition of an equimolar concentration of KatG^{Cterm} to KatG^{Nterm} resulted in the immediate appearance of EPR signals corresponding to high-spin pentacoordinate and hexacoordinate heme iron, as evident from a group of signals with g -values around 6 and 2 [spectrum b]; however, the hexacoordinate low-spin complex was still the dominant species present. Following a 48-hour incubation (4°C) of the two domains together, a dramatic increase in the high-spin species was observed along with the concomitant decrease in the hexacoordinate low-spin complex [spectrum c]. Though some hexacoordinate low-spin complex was still observable after the 48-hour incubation, the mixture of the separately expressed and isolated domains bore a strong resemblance to wild-type KatG (spectrum d). Additionally, recombined and incubated samples were exposed to peroxide substrate (100 eel.) resulting in a diminished EPR signal intensity (~30%). No new or obvious species were observed with the exception of a weak free radical signal at $g \sim 2$. [Figure 60].

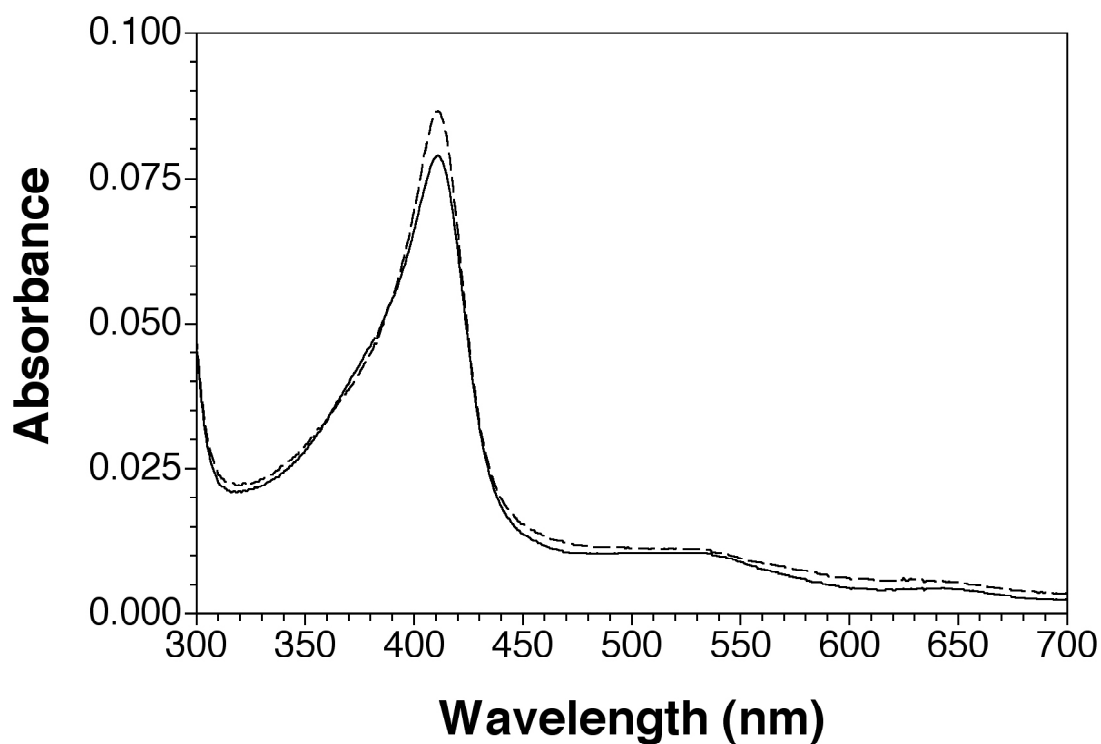


Figure 59. Effect of H₂O₂ on an equimolar mixture of KatG^{Nterm} and KatG^{Cterm}. A spectrum corresponding to the charge transfer and α and β transitions of the heme prosthetic group were recorded for an equimolar mixture of KatG^{Nterm} and KatG^{Cterm} with (solid line) and without (dashed line) H₂O₂.

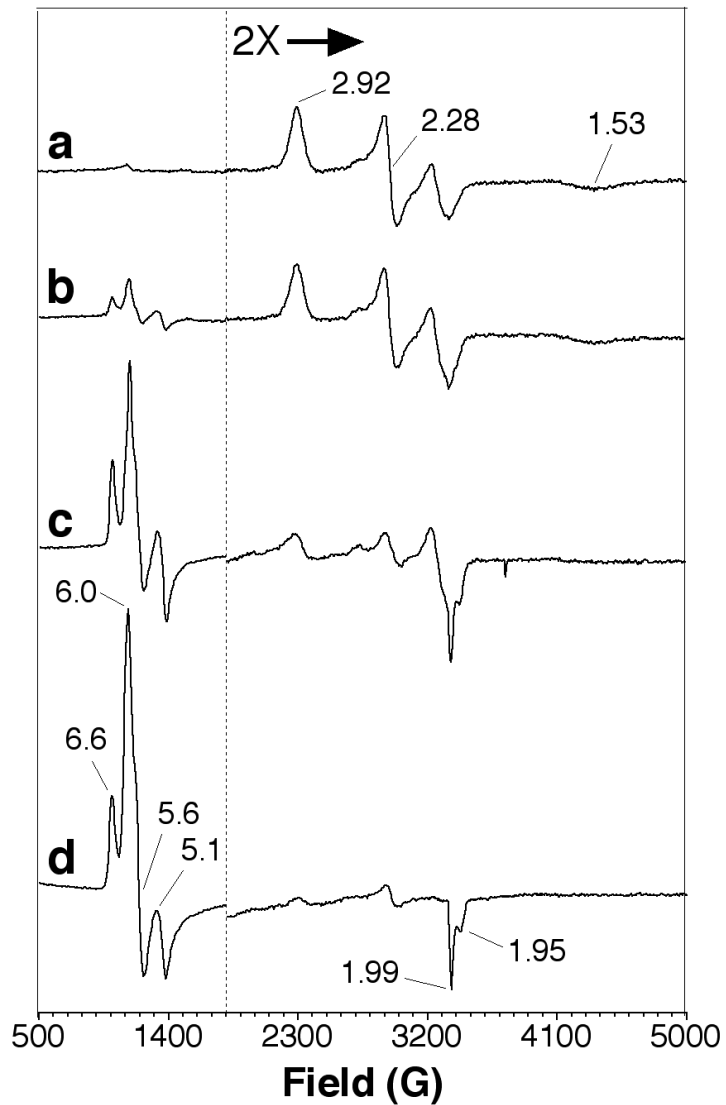


Figure 60. Effect of separately expressed and isolated KatG^{Cterm} on the EPR spectrum for KatG^{Nterm}. An EPR spectrum was recorded for KatG^{Nterm} alone (a), KatG^{Nterm} immediately following mixture with equimolar KatG^{Cterm} (b), KatG^{Nterm} following a 48-hour incubation (4°C) with equimolar KatG^{Cterm} (c), and wtKatG (d).

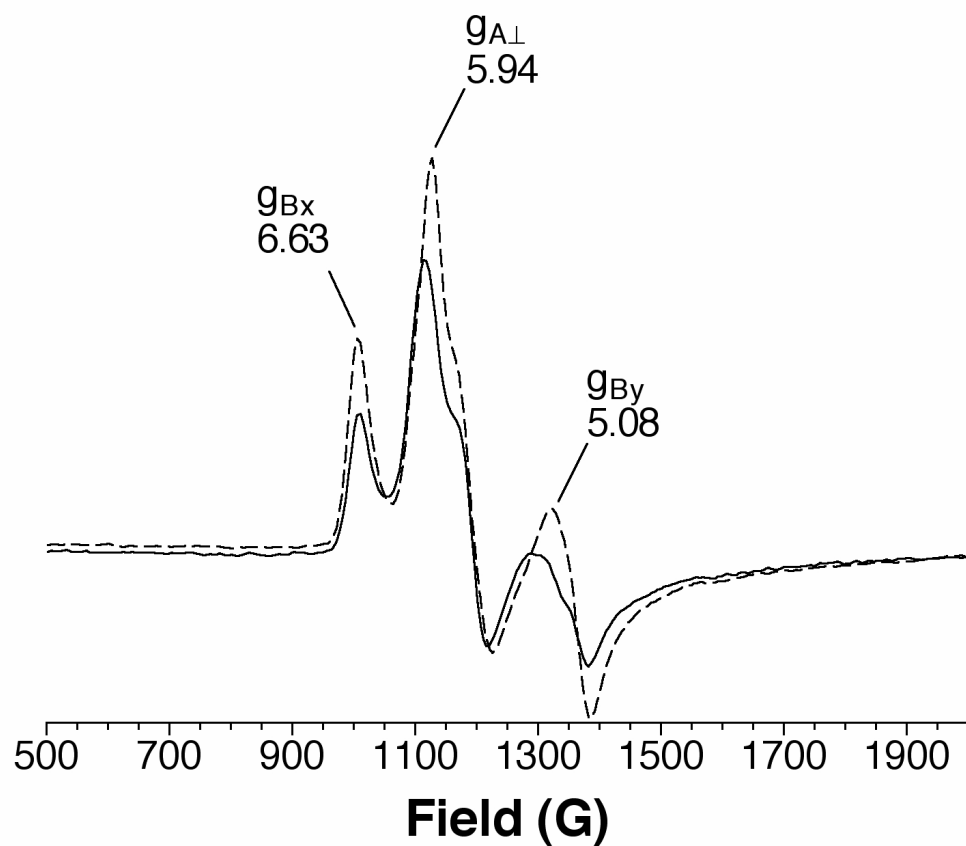


Figure 61. Effect of H_2O_2 on the EPR spectrum of the equimolar combination of $KatG^{Nterm}$ and $KatG^{Cterm}$. EPR spectrum was collected for an equimolar combination of $KatG^{Nterm}$ and $KatG^{Cterm}$ with (solid line) and without (dashed line) H_2O_2 .

IV. DISCUSSION

The broad requirement for oxygen in living systems requires the versatility of transition metals to activate oxygen for reactivity while at the same time strictly controlling the activated oxygen to prevent oxidative damage. One such transition metal, iron is found in proteins performing a variety of functions including, oxygen transport by hemoglobin and myoglobin, activation of less reactive organic compounds by cytochrome P450 and lipoxygenase removal of peroxides and radical oxygen species by proteins such as iron-superoxide dismutase, and catalase-peroxidases

The reactivity of these proteins is controlled structurally at four different levels; the active site iron cofactor level, the immediate active site level, a level peripheral to the active site and finally at a level far distant from the active site. Truly, there is a great body of knowledge concerning the first three levels of control, several examples of which are presented in this dissertation for different cofactor systems; however, there is very little knowledge about the last level due to the lack of a suitable model. With an appropriate model then a better understanding of how these distant structural features impact function becomes possible. Perfect examples of such protein systems are found among some hemoproteins. For instance considering monofunctional peroxidases and catalase-peroxidases they share a high degree of sequence homology and structural similarity within the first three levels of control but the ability of the catalase-peroxidases

to catalyze both catalase and peroxidase reaction using a single active site, can not be adequately be explained by comparing only the first three levels of structure between these two enzymes groups. Therefore, it is obvious that we must look further out to the distant interactions in order to fully understand the role of distant structural features in function. This dissertation addresses the functional roles of distant interactions, including novel purification methods and diagnostic methods for this purpose. In addition to providing new insight into this poorly understood aspect of enzyme structure/function relationships region this dissertation presents catalase-peroxidase proteins as model enzymes for the understanding of structural control on all four levels as well as being proteins uniquely suited for structural manipulation to tune function, particularly toward enzyme engineering.

A comparison of monofunctional peroxidases and catalase-peroxidases reveals three distinct areas of structural difference, found in the catalase-peroxidases that include two interhelical insertions between the D and E helices and the F and G helices respectively and the large C-terminal domain. Much insight into the roles of the two interhelical insertions was provided by the work of Yongjiang Li, a former member of the Goodwin laboratory (32). The focus of the work described in this dissertation is on the structural roles of the C-terminal domain. Despite its substantial distance from the active site (≥ 30 Å) and lack of any identifiable catalytic activity of its own, the C-terminal domain is clearly an integral component of catalase-peroxidase structure and catalysis.

Identifying the role of the C-terminal domain

Expression of the *E. Coli* catalase-peroxidase (KatG) N-terminal domain (KatG^{Nterm}) in the absence of the C-terminal domain produced insoluble protein. The method developed for purification of the heme receptor protein, also discussed in this dissertation, proved to be advantageous for purification of insoluble proteins and thus was ideal for purification of the insoluble KatG^{Nterm}. KatG^{Nterm}, had no observable activity, catalase or peroxidase. Inspection by far UV CD showed that the secondary structural content of the protein was as it was expected based on crystal structures of intact catalase-peroxidases (33) having an expected 8-10% lower α -helical content than that of the wild type. UV-visible absorption spectra indicated that the active site heme was specifically ligated in the active site and not adventitiously bound.

Consistent with the loss of activity CN⁻ binding by KatG^{Nterm} was dramatically reduced compared to wild-type KatG. Indeed, the rate for binding, k_{on} , was decreased by three orders of magnitude for KatG^{Nterm} ($120 \text{ M}^{-1} \text{ s}^{-1}$) as compared to wild-type KatG ($5 \times 10^5 \text{ M}^{-1} \text{ s}^{-1}$). Spectral investigation revealed that the active site heme was converted to an exclusively hexacoordinate low-spin state rather than the mixture of hexacoordinate and pentacoordinate high-spin heme environments typical for native catalase-peroxidases. The low spin state of KatG^{Nterm} indicates that the sixth ligand to the heme is a strong field ligand. According to EPR this ligand is likely supplied by a histidine residue.

Considering the histidines within the heme coordination sphere the most likely candidate for a sixth or distal ligand would be the histidine which normally functions as a

general base in catalysis, His 106 (*E. coli* KatG numbering). Because heterolytic cleavage of hydroperoxide substrates requires their direct access to the heme iron, and because the distal histidine is required as a general base for the reaction, the loss of catalase and peroxidase activity is a reasonable expectation for active site rearrangement of this type. Site directed mutagenesis of the His 106 indicates that it, in fact, does ligate the heme in the active site of KatG^{Nterm}. Therefore, removal of the C-terminal domain results in a complete loss of catalase and peroxidase activity and a conversion of the active site to a hexacoordinate low-spin heme environment.

Restructuring of the Active Site

These effects, resulting from removal of the C-terminal domain, are at least partially reversed by the reintroduction of the C-terminal domain as a separately expressed and purified protein. Evidence of this restructuring can clearly be seen in the visible spectra which reveal an increase in high-spin species, and a subsequent reduction of hexacoordinate low-spin species. Confirmations of these changes are seen in the EPR spectra as well. In fact, EPR reveals that the distribution of pentacoordinate and hexacoordinate high-spin species nearly identical to that of the wild-type enzyme. The mixed population of coordination states for ferric catalase-peroxidases is a peculiarity of these enzymes; one that is highly sensitive to subtle changes in purification, aging, buffer conditions, and so on (80, 91). It is striking that this distribution is so precisely reproduced by domain recombination.

The more exacting measure of active site reconstruction, however, is the presence of not only peroxidase activity, but catalase activity as well. Recent studies by our laboratory and others indicate that the catalase activity of the catalase-peroxidases depends upon highly ordered hydrogen bonding networks are required for catalase activity which are either dependent upon (at least in part) or act together with a novel covalent crosslink between active site Trp 105, Tyr 226, and Met 252 (*E. coli* KatG numbering). Many catalase-peroxidase variants created by site-directed mutagenesis and deletion mutagenesis exhibit little or no catalase activity, though modifications to the heme environment are subtle in response to these modifications (32, 45, 92-96). In fact, the vast majority of these variants retain peroxidase activity equal to or better than their wild-type counterparts (32, 45, 92-96). Therefore, it is evident that the substantial recovery of catalase activity observed from KatG^{Nterm} incubated with KatG^{Cterm} indicates that the C-terminal domain supports a very precise restructuring of the active site environment allowing for reestablishment of to hydrogen bonding networks, in a significant proportion of the KatGNterm population.

A comparison of the reported structures for catalase-peroxidases to the known behavior of similar monofunctional peroxidases illuminates the connection between the presence or absence of the C-terminal domain and the coordination state of the heme iron, particularly as it involves the distal histidine (His 106) (33). There are prominent *intrasubunit* interactions observed between the B'C' loop and E' helix of the C-terminal domain and the BC loop of the N-terminal domain [Figure 62]. The most prominent of these interactions between the C-terminal and N-terminal domains include an Arg117/Asp597 ion pair, contact between the Arg 484 guanidinium group and two

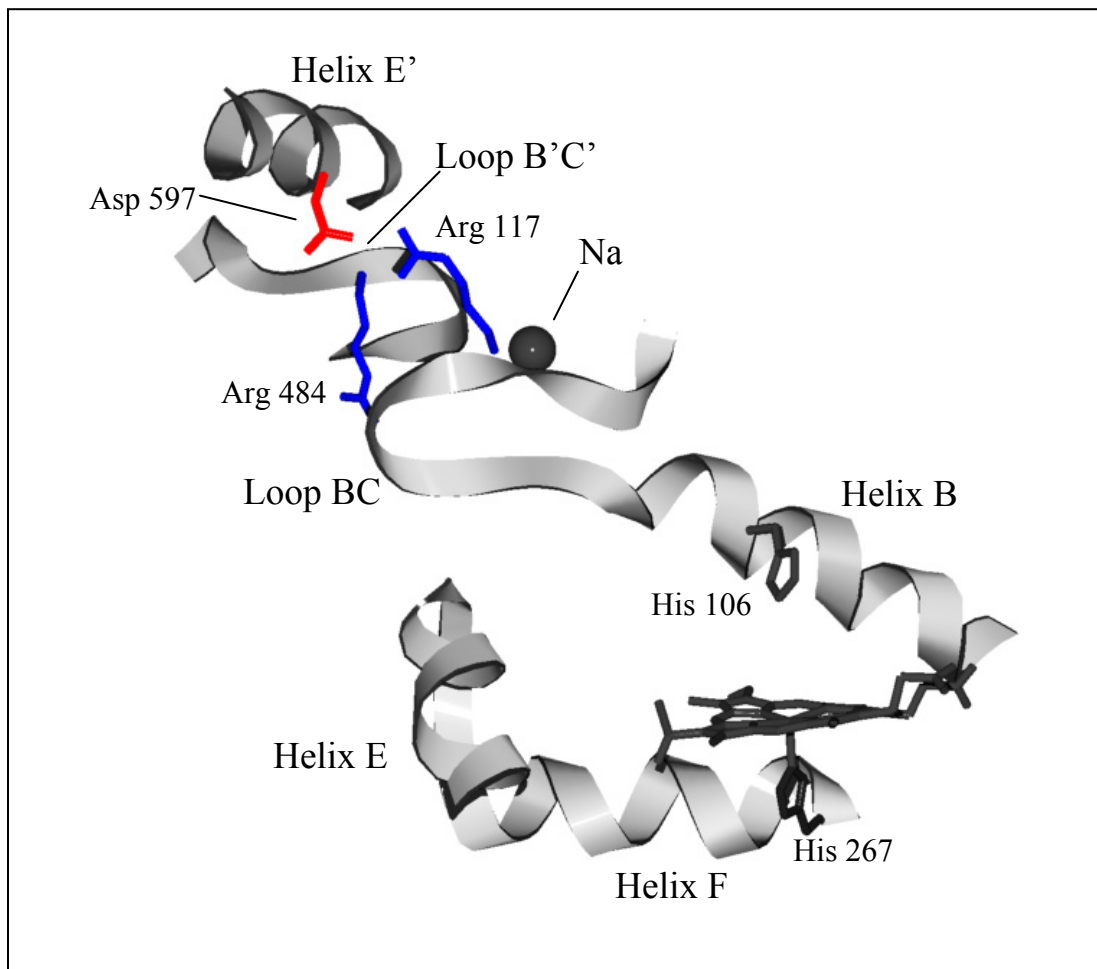


Figure 62. Structural representation of the major helices surrounding the active sites of catalase-peroxidase for *Burkholderia pseudomallei* (PDB accession No.1MWV) . Helices E, F and B, from the N-terminal domain are shown and Helix E' and Loop B'C', both from the C-terminal domain are shown. Additionally, the distal and proximal histidines are labeled His10s and His 267, respectively. The position of the sodium ion is shown by the grey sphere, the Arg 117 and 484 side chains are in blue, and the Asp 597 side chain is red. Amino acid numbering is in accordance to their appearance in the *E. Coli* KatG sequence.

backbone carbonyls and Ile114, a phenol oxygen of Tyr 111 in hydrogen bonding distance from Asp 482 (not shown) as well as legends to a bound sodium ion contributed by both domains. Substantial *intersubunit* interactions are also observed between the AB loop and C helix, of the N-terminal domain, and a structure encompassing helices G' through I' of the C-terminal domain. On both sides of these interdomain interfaces, many of the individual residues involved in the interactions are strictly conserved across the catalase-peroxidases.

Comparison to Monofunctional Peroxidases

Monofunctional peroxidases support the BC loop and B helix of the N-terminal domain in a different way. For instance, in class II peroxidases, the fungal secretory peroxidases, a calcium ion is coordinated by residues from the BC loop and B-helix (86, 87). The loss of this calcium ion results in loss of coordination by the BC loop thus allowing for the coordination of the distal histidine to the heme iron, yielding a hexacoordinate low-spin heme in the active site and the concomitant loss of peroxidase activity (84, 85, 88). Interestingly, this bishistidine coordinated heme of the Ca²⁺-deficient manganese peroxidase (a class II peroxidase) loses a histidine ligand with decreasing pH according to a pK_a of 5.7 (84). In our system k_{react} increases as pH decreases with an apparent pK_a of 6.2 which would be consistent with the protonation of a histidine residue to disfavor its coordination to the heme iron and thus achieve activation of KatG^{Nterm}.

In the class III plant secretory peroxidases, such as horseradish peroxidase and peanut peroxidase, the BC loop and B helix are provided additional stability via a disulfide bond that straddles the calcium coordination site (100, 101). Thus, less activity is lost, and a more subtle alteration of the heme environment occurs upon loss of Ca^{2+} from these enzymes (102). In fact, greater active site stability has been imparted to the class II manganese peroxidase by using site-directed mutagenesis to introduce an analogous disulfide bond to mimic those found in class III (103).

In Class II and Class III peroxidases, restoration of the active site and peroxidase activity occurs rapidly upon reconstitution with calcium (81, 104). In the case of Ca^{2+} -depleted MnP, for instance, activity is restored within minutes of addition of calcium ion (81). On the contrary, spectral shifts and activity assays for $\text{KatG}^{\text{Nterm}}$ indicate a 24 hour time frame for restructuring of the active site and recovery of activity. One reasoning for the distinct difference in the reactivation time is that reactivation of $\text{KatG}^{\text{Nterm}}$ was accomplished using only one equivalent of $\text{KatG}^{\text{Cterm}}$ while Ca^{2+} -depleted MnP was exposed to over 30 equivalents of Ca^{2+} (81). The slow recovery of activity and spectral shifts indicate slow restructuring of the active site in $\text{KatG}^{\text{Nterm}}$. Nevertheless, to first approximation, $\text{KatG}^{\text{Nterm}}$ reactivation occurs at about one-tenth the rate of Ca^{2+} -dependent MnP reactivation.

One possibility which would be consistent with a slow rearrangement would be the formation of the Trp-Tyr-Met covalent adduct unique to catalase-peroxidases (33). However, considering that the formation of this adduct is peroxide dependant and proceeds through formation of compound I it seems highly unlikely since in the case of $\text{KatG}^{\text{Nterm}}$ active site restructuring is accomplished by the addition of $\text{KatG}^{\text{Cterm}}$ alone

with no addition of peroxide (93, 97). Also, the hexacoordinated heme environment is not conducive to compound I formation. Thus, formation of the adduct could not precede the structural rearrangement necessary for reactivation.

A second possible explanation could be that initial structural perturbations of $\text{KatG}^{\text{Nterm}}$ and/or $\text{KatG}^{\text{Cterm}}$ are necessary prior to establishing the positive association leading to restructuring of the active site. This is consistent with the theory proposed by Carpena et. al. which suggests that the C-terminal domain may act as a platform on which the N-terminal domain folds into its proper conformation (106). This too seems unlikely since there are not large-scale changes in the CD spectra. The CD spectra of $\text{KatG}^{\text{Nterm}}$ and $\text{KatG}^{\text{Cterm}}$ are consistent with what is expected from the wild-type enzyme and no conformational changes are observed in the CD spectra upon incubation of the two domains together. In any case, it is clear that $\text{KatG}^{\text{Cterm}}$ is able to induce whatever structural changes in the N-terminal domain are necessary to restore its active conformation.

Interestingly, a rearrangement that would involve shifts in preexistent structural elements could account for the observed data. A complex connection between the first 30 amino acids of both N-terminal domains is seen in all known catalase-peroxidase structures. Associative contributions to the dimer interface also come from each C-terminal domain (33). Therefore, it seems only logical that with the removal of the C-terminal domain and subsequent structural changes in the N-terminal domain that the first amino acids of the N-terminal domain could adopt an alternative conformation that does not allow for the correct interaction across the dimer interface. If so then this alternative conformation could present a significant kinetic barrier to active site restructuring.

Slow conformational changes of this type have been observed in the substrate-induced interconversion of quaternary isoforms of human porphobilinogen synthase. Rate constants for this interconversion as measured by substrate-induced reactivation of the enzyme were on the order of $1.7 \times 10^{-2} \text{ s}^{-1}$ (reported as 1.05 h^{-1}), which is very similar to what we have observed for $\text{KatG}^{\text{Nterm}}$ reactivation (107). Whether measured spectroscopically (~50 %) or by catalytic activity (~20 %), active site restructuring by $\text{KatG}^{\text{Nterm}}$ is not complete. It is possible that a fraction of the N-terminal domain is locked in a conformational state that is noncompliant to the actions of $\text{KatG}^{\text{Cterm}}$. It seems likely that this fraction is indicated by that portion which remains in its hexacoordinate low-spin state (~50 %). At the present time the conditions which address this population are not known.

Investigation of Incomplete Reactivation

Whether measured spectroscopically (~50%) or by catalytic activity (~20%) active site restructuring by $\text{KatG}^{\text{Nterm}}$ is not complete. It is possible that a fraction of the N-terminal domain is locked in a conformational state that is noncompliant to the actions of $\text{KatG}^{\text{Cterm}}$. It seems likely that this fraction is indicated by that portion which remains in its hexacoordinate low-spin state (~50%). At the present time the conditions which address this population are not known. It appears that the proportion of $\text{KatG}^{\text{Nterm}}$ is spectroscopically responsive to $\text{KatG}^{\text{Cterm}}$ -dependent restructuring, either in less active conformation or produces a mixed population of active and inactive states. Two possible

explanations are currently under investigation in the Goodwin laboratory. One is that the restructuring of KatG^{Nterm} by KatG^{Cterm} produces a proportion of active sites, which though similar to that of wild-type KatG, do not react correctly with H₂O₂. Indeed, the fact that reaction with a low concentration of H₂O₂ produces an appreciable change in EPR spectrum evident a full hour after addition suggests that at least some proportion of restructured KatG^{Nterm} is compromised with respect to catalytic consumption of substrate.

A second possibility is that there may be some structural elements in KatG^{Nterm} outside of the active site which are important for catalysis but are not properly reoriented with addition of KatG^{Cterm}. Such effects are seen in a KatG^{Nterm} variant where the FG insertion has been removed. Addition of equimolar amounts of KatG^{Cterm} to this variant result in spectral shifts indicative of approximately the same active site restructuring as seen in the unmodified KatG^{Nterm} by KatG^{Cterm}. However, unlike KatG^{Nterm} the recovered peroxidase activity is comparable ($\geq 50\%$) to that of intact KatG lacking the same FG insertion (Li and Goodwin, unpublished observations). Interestingly, there is an *intersubunit* point of contact between the FG insertion, in the N-terminal domain, of one subunit to the C-terminal domain of the second subunit. Therefore it is possible that KatG^{Cterm} does not fully or correctly establish this contact when introduced to KatG^{Nterm}.

Future directions.

The essential nature of the C-terminal domain to catalase-peroxidase structure and catalysis is demonstrated by its ability to restore the N-terminal domain to its bifunctionally active conformation. Reactivation of the enzyme in this manner provides a unique opportunity to probe the mechanisms by which the C-terminal domain directs the folding and maintenance of this uniquely bifunctional active site. Understanding of these mechanisms could be applied to other hemoproteins (e. g. fungal peroxidases) to make them more amenable to expression and manipulation in bacterial systems. This would remove a common obstacle to the rapid progression of research with such enzymes.

This knowledge and future information from catalase-peroxidase structural studies could be used to engineer the catalase-peroxidase structure in order to manipulate its function. The C-terminal domain came about from a gene duplication and fusion event of a monofunction peroxidase and comparison to monofunctional peroxidases reveals that relatively few modifications have occurred (33, 35). Considering this it seems feasible that by reversing these modifications then one could convert the C-terminal domain into an active domain, like the N-terminal domain, thereby creating a dual-bifunctional enzyme. Also, by manipulating different modification one could modulate or tune the enzyme toward one function or the other, catalase or peroxidase, resulting in a variety of diverse proteins.

In order to accomplish this the active site must first be modified in order to facilitate heme binding. To do so the residue in place of the proximal histidine, Arg617, must be replaced so that heme can bind. Adjacent to this a methionine residue, Met616,

is oriented so that its large aliphatic side chain protrudes into the heme binding cavity. This position, in peroxidases, is normally occupied by Gly, Ser or Ala residue so in this case Met616 will be replaced with Ala. Finally, side chains of two residues, Phe579 and Arg580, block heme binding in the "active site" and will be replaced to allow for heme accommodation. Once heme is in place then modifications can be made to recover activity. Two residues, critical to peroxidase activity, must be substituted by site directed mutagenesis. The first of these is Ala473 which is in the place of the distal histidine, which plays the critical role as a general base in catalysis (45). The second residue is Ser469 which replaces an essential Arg residue (45).

In order to regain catalase activity the Trp-Tyr-Met crosslink must be reestablished. The Trp residue is already in place for the crosslink so it will only be necessary to replace the existing Leu602 with the required Met and then introduce the Tyr residue in proximity to the tryptophan residue. Procedures are already in place to accomplish these proposed mutagenesis as well as purification methods and techniques to characterize the novel enzymes produced in this study. In addition to engineering these new enzymes this project will most likely address questions regarding the Trp-Tyr-Met crosslink required for catalase activity. For instance, what structural attributes are actually necessary to ensure formation of the crosslink?

The C-terminal domain could also function as a starting point for engineering other novel hemoproteins. Catalase-peroxidases have been observed to catalyze reactions outside of catalase and peroxidase including dioxygenase and peroxynitritase activities (30, 31, 79) and catalase-peroxidases are found in a wide range of eubacteria and archaea making them ideal for a broad spectrum of possibilities for enzyme engineering.

Therefore, this future work should contribute a great deal of understanding about the structure and function relationship in catalase-peroxidases and other hemoproteins as well as establishing a foundation for new directions in enzyme engineering.

Summary

Catalase-peroxidases are a unique group of hemoproteins which have the ability to catalyze both catalase *and* peroxidase reactions within a single active site. Within metalloproteins, including hemoproteins, there are four main levels of structural control to tune the active site and direct catalysis: 1) the ligands to the metal, 2) immediate interactions with residues in the active site, 3) interactions with residues peripheral to the active site and 4) interactions from distant structural features. While catalase-peroxidases share a high degree of sequence similarity with monofunctional peroxidases, they do not share such similarities with monofunctional catalases. Thus, their ability to perform catalase activity is something of a mystery. A comparison of the active sites of catalase-peroxidase and monofunctional peroxidases reveals that they are practically superimposable. Therefore, within the first two levels of structural control reasons for bifunctionality are not clear. Looking to the third level, we see a similar conservation of interactions between the two protein types. It is clear from these similarities that in order to better understand the structural control and the essence of the bifunctional nature of catalase-peroxidases that one must look outward to the final level of control, the distant protein structures.

The comparison of catalase-peroxidases to monofunctional peroxidases reveals that there are three regions distant from the active site which are unique to catalase-peroxidases and not found in the monofunctional enzymes. These regions include two interhelical insertions, the DE and FG insertions. These are found in loops between the D and E helices and the F and G helices, respectively, and a large (300 amino acid) C-terminal domain. Due to the fact that these regions are not found in the functionally limited monofunctional peroxidases, it seems likely that they could play roles in the bifunctional nature of catalase-peroxidases. Much insight into of the roles of the DE and FG insertions has already been provided by a fellow group member Yongjiang Li (110). Conversely, the work of this dissertation has explored the role of the C-terminal domain in catalase-peroxidase structure and catalysis.

Removal of the C-terminal domain resulted in a protein composed only the N-terminal domain, KatG^{Nterm}. The size and secondary structure of his KatG^{Nterm} protein was as expected, however, it had neither catalase not peroxidase activity. Cyanide binding studies revealed a decrease in CN⁻ binding for KatG^{Nterm} by three orders of magnitude, as compared to the wild type, wtKatG, indicating that substrate access was greatly reduced and that the KatG^{Nterm} active site was occluded in some way. Spectroscopic evaluation of the protein revealed that the heme coordination environment of KatG^{Nterm} was exclusively hexacoordinate low-spin. This contrasts the predominately high-spin mixed pentacoordinate and hexacoordinate heme species typically seen in catalase-peroxidases. Furthermore, the EPR spectrum for KatG^{Nterm} indicated a histidine residue as the sixth ligand to the heme iron. Since histidine is a strong field ligand this would induce a low-spin state on the heme. Considering the available residues for

ligation within close proximity to the heme the distal histidine, which acts as a general base in catalysis, is the most likely candidate as the sixth ligand. Indeed, site directed mutagenesis of this distal histidine reveals that it in fact does ligate the heme in the active site. Therefore, removal of the C-terminal domain results in a shift in the coordination environment to high-spin hexacoordinate and results in complete loss in catalase and peroxidase function.

The work described in this dissertation demonstrated that reintroduction of a separately expressed and isolated C-terminal protein, KatG^{Cterm} restructured of the active site indicated by a decrease in the hexacoordinate low-spin spectra and a concomitant increase in the high-spin pentacoordinate. These shifts reaching an optimal recovery of ~50% of wild type character after 48 hr of incubating KatG^{Nterm} with KatG^{Cterm}. Far-UV CD spectrum of the incubated domains differed from spectra of the individual domains as well as that of wtKatG but was superimposable with the arithmetic combination of the individual KatG^{Nterm} and KatG^{Cterm} spectrum. This indicates that the recombination does not result in major disruptions of the secondary structural content from what is expected for the two domains together. With the restructuring of the active site appreciable catalase and peroxidase activity (~20%) was also recovered culminating at 24 hr.

Similar bis-histidine coordination and subsequent loss of activity occurs in some monofunctional peroxidases upon loss of Ca²⁺. In these enzymes a Ca²⁺, located on the distal side of the heme interacts with residues from the BC loop on the B helix where the distal histidine is found. Loss of these interactions, due to the removal of calcium ion, result in coordination to the heme iron by the distal histidine and loss of activity. In catalase-peroxidases, thus far, no calcium ion has been identified to stabilize the active

site structure; however, there are several specific and highly conserved *intrasubunit* interactions between the C-terminal domain and the BC loop of the N-terminal domain. There are additional strictly conserved *intersubunit* interactions between the N-terminal and C-terminal domain which may provide stability to the active site. Thus, the multiple interactions between the C-terminal and N-terminal domains suggest that at least one of the roles of the C-terminal domain is to precisely control the architecture of the active site for catalysis. Investigation of these interactions is expected to provide greater detail into how the C-terminal domain impacts the bifunctionality of catalase-peroxidases. These advances into the understanding of structure/function roles of regions distant to the active site, like the C-terminal domain, provide good models for probing distant interactions, not only in catalase-peroxidases, but in other proteins as well.

REFERENCES

1. Baker, R. D. Cook, C. O., Goodwin, D. C. (2005) Catalase-Peroxidase Active Site Restructuring by a Distant and "Inactive" Domain. *Biochemistry*
2. Scarpa, M. S., R.; Viglino, P. and Rigo, A. (1983) Superoxide ion as active intermediate in the autoxidation of ascorbate by molecular oxygen. *J. Biol. Chem.* 258, 6695-6697
3. LeMay, J. H., Eugene, H., Burdge, J. R. & Brown, T. L. (2002) *Chemistry : The Central Science*, Prentice Hall, Inc, NJ
4. Koppenol, W. H. a. L., J. F. (1984) The oxidizing nature of the hydroxyl radical. A comparison with the ferryl ion (FeO^{2+}). *J. Phys. Chem.* 88, 99-101
5. Letoffe, S., Redeker, V., and Wandersman, C. (1998) Isolation and characterization of an extracellular haem-binding protein for *Pseudomonas aeruginosa* that shares function and sequence similarities with the *Serratia marcescens* HasA haemophore. *Mol. Micro.* 28, 1223-1234
6. Walczyk, T. v. B., F. (2005) Deciphering the iron isotope message of the human body. *International Journal of Mass Spec.* 242, 117-134

7. Cook, J. D. (1990) Adaptation in iron metabolism. *Am. J. Clin. Nutrition* 51, 301-308
8. Gardner, H. W. (1989) Soybean lipoxygenase-1 enzymatically forms both (9S)- and (13S)-hydroperoxides from linoleic acid by a pH-dependent mechanism. *Biochim Biophys Acta* 1001, 274-281
9. Que, J. L. H., R. Y. N. (1996) Dioxygen activation by enzymes with mononuclear non-heme iron active sites. *Chem. Rev.* 96, 2607-2624
10. Sloane, D. L. L., R.; Craik, C. S. and Sigal, E. (1991) A primary determinant for lipoxygenase positional specificity. *Nature* 354, 149-152
11. Solomon, E. I. B., T. C.; Davis, M. I.; Kemsley, J. n.; Lee, S-K.; Lehnert, N.; Neese, F.; Skulan, A. J.; Yang, Y-S. and Zhou, J. (2000) Geometric and electronic structure/function correlations in non-heme iron enzymes. *Chem. Rev.* 100, 235-349
12. Sloane, D. L. (1996) In Lipoxygenase: Lipoxygenase Pathway Enzymes. In American Oil Chemists' Society (Piazza, G., ed) pp. 57-79, IL
13. Sigal, E., Craik, C. S., Highland, E.; Grunberger, D.; Costello, L. L.; Dixon, R. A. F. and Nadel, J. A. (1988) Molecular cloning and primary structure of human 15-lipoxygenase. *Biochem. Biophys. Res. Commun.* 157, 457

14. Solomon, E. I. Z., J.; Neese, F. and Pavel, E. G. (1997) New insights from spectroscopy into the structure/function relationships of lipoxygenases. *Chem. Biol.* 4, 795-808
15. Minor, W., Steczko, J., Stec, B., Otwinowski, Z., Bolin, J. T., Walter, R. & Axelrod, B. (1996) Crystal structure of soybean lipoxygenase L-1 at 1.4 Å Resolution. *Biochemistry.* 35, 10687-10701
16. Wallar, B. J. L., J. D. (1996) Dioxygen activation by enzymes containing binuclear non-heme iron clusters. *Chem. Rev.* 96, 2625-2657
17. Rosenzweig, A. C., Nordlund, P., Takahara, P. M., Fredrick, C. A. & Lippard, S. J. (1995) Geometry of the soluble methane monooxygenase catalytic diiron center in two oxidation states. *Chemistry & Biology* 2, 409-418
18. Siegbahn, P. E. M. C., R. H. (1997) Mechanism of C-H activation by diiron methane monooxygenase: quantum chemical studies. *J. Am. Chem. Soc.* 119, 3103-3113
19. Brandstetter, H., Whittington, D. A., Lippard, S. J. & Frederick, C. A. (1999) Mutational and structural analyses of the regulatory protein B of soluble methane monooxygenase from *Methylococcus capsulatus*. *Chem. Biol.* 7, 441-449
20. Brazeau, B. J., Austin, R. N., Tarr, C., Groves, J. T., and Lipscomb, J. D. (2001) Intermediate Q from soluble methane monooxygenase hydroxylates the

- mechanistic substrate probe norcaradiene: Evidence for a stepwise reaction. *J. Am. Chem. Soc.* 123, 11831-11837
21. Liu, Y. N., J. C.; Lee S.-K. and Lipscomb, J.D. (1995) Gating Effects of Component B on Oxygen Activation by the Methane Monooxygenase Hydroxylase Component. *J Biol Chem* 270, 24662-24665
 22. Kopp, D. A., and Lippard, S. J. (2002) Soluble methane monooxygenase: activation of dioxygen and methane. *Curr. Opin. Struct. Biol.* 6, 568-576
 23. Siegbahn, P. E. M. C., R. H.; Nordlund, P. J. Biol. (1998) Mechanism of methane monooxygenase - a structural and quantum chemical perspective. *J. Biol. Inorg. Chem.* 3, 314-317
 24. Day, M. W., Hsu, B. T., Joshua, L., Park, J. B., Zhou, Z. H., Adams, M. W. W. & Rees, D. C. (1992) *Protein Science* 1, 1494 -1507
 25. Beinert, H. K., M. C. (1989) Engineering of protein bound iron-sulfur clusters. A tool for the study of protein and cluster chemistry and mechanism of iron-sulfur enzymes. *Eur. J. Biochem.* 186, 5-15
 26. Lauble, H., Kennedy, M. C., Beinert, H., Stout, C. D. (1994) Crystal structure of aconitase with trans-aconitase and nitrocitrate bound. *J. Mol. Biol* 237, 437-451
 27. Robbins, A. H., Stout, C. D. (1989) Structure of activated aconitase: Formation of the [4Fe-4S] cluster in the crystal. *Proc. Natl. Acad. Sci.* 86, 3639-3643

28. Welinder, K. G. (1992) Superfamily of plant, fungal and bacterial peroxidases. *Curr Opin Struct Biol.* 2
29. Adachi, S., Negano, S., Ishimori, K., Wananabe, Y., Morishima, I., Egawa, T., Kitagawa, T., and Makino, R. (1991) Roles of proximal ligand in heme-proteins - replacement of proximal histidine of human myoglobin with cystine and tyrosine by site directed mutagenesis as models for P450, chloroperoxidase, and catalase. *Biochemistry* 32, 241-252
30. Wengenack, N. L. J., M. P.; Rusnak, F. Nad Stern, M. K. (1999) *Mycobacterium tuberculosis* KatG is a peroxynitritase. *Biochem. Biophys. Res. Commun.* 256, 485-487
31. Sagawa, T. Y., T.; Tsugawa, S.; Takeda, Y.; Ihara, H.; Mitamura, E.K.; Urabe, L. and Ohkubo, K. (1999) A novel versatility of catalase I as a dioxygenase for indole-ring-opening dioxygenation. *Chem. Lett.* 1999, 399-340
32. Li, Y. Goodwin, Douglas C. (2004) Vital roles of an interhelical insertion in catalase-peroxidase bifunctionality. *Biochem. Biophys. Res. Commun.* 318, 970-976
33. Yamada, Y., Fujiwara, T., Sato, T., Igarashi, N., and Tanaka, N. (2002) The 2.0 Å crystal structure of catalase-peroxidase from *Haloarcula marismortui*. *Nat Struct Biol* 9, 691-695

34. Wang, J., Mauro, M., Edwards, S. L., Oatley, S. J., Fishel, L. A., Ashford, V. A., Yuong, N., and Kraut, J. (1990) X-ray Structures of recombinant yeast cytochrome c peroxidase and three heme-cleft mutants prepared by site-directed mutagenesis. *Biochemistry* 29, 7160-7173
35. Welinder, K. G. (1991) Bacterial catalase-peroxidases are gene duplicated members of the plant peroxidase superfamily. *Biochim Biophys Acta* 1080, 215-220
36. Zamocky, M., Regelsberger, G., Jakopitsch, C., and Obinger, C. (2001) The molecular peculiarities of catalase-peroxidases. *FEBS Lett* 492, 177-182
37. Gouterman, M. (1978) In *The Porphyrins* (Dolphin, D., ed) Vol. 3, Academic, New York
38. Dunford, H. B. (1999) *Heme Peroxidases*, John Wiley & Sons, New York
39. Renganathan, V. a. G., M. H. (1986) *Biochemistry* 25, 1626-1631
40. Eaton, W. A., and Hochstrasser, R. M. (1967) *J. Chem. Phys.* 46, 2533
41. Cheesman, M. R., Greenwood, C. and Thompson, A.J. (1991) Magnetic circular dichroism of hemoproteins. *Adv. Inorg. Chem.* 36, 205-255
42. Braterman, P. S., Davies, R. C. and Williams, R. J. P. (1964) *Adv. Chem. Phys.*, 359

43. Tamura, M. a. H., H. (1972) Optical and magnetic measurements of horseradish peroxidase. 3. Electron paramagnetic resonance studies at liquid-hydrogen and -helium temperatures. *Biochim Biophys Acta* 284, 20-29
44. Maltempo (1972) Electron paramagnetic resonance analysis of horseradish peroxidase in situ and after purification. *Biochemistry* 18
45. Hillar, A., Peters, B., Pauls, R., Loboda, A., Zhang, H., Mauk, A. G., and Loewen, P. C. (2000) Modulation of the activities of catalase-peroxidase HPI of *Escherichia coli* by site-directed mutagenesis. *Biochemistry* 39, 5868-5875
46. Dunford, H. B., and Stillman, J. S. (1976) On the Function and Mechanism of Action of Peroxidases. *Coord Chem Rev*, 187-251
47. WHO (2005) Global tuberculosis control - surveillance, planning, financing. In, World Health Organization, London/Geneva
48. WHO (2005) Drug- and multidrug-resistant tuberculosis (MDR-TB). In, World Health Organization
49. Johnsson, D. S., King, and Schultz, P. G. (1995) The studies on the mechanism of action of isoniazid and ethionamide in the chemotherapy of tuberculosis. *J. Am. Chem. Soc.* 117, 5009-5010
50. Rattan, A., Kalia, A., and Ahmad, N. (1998) Multidrug-resistant *Mycobacterium tuberculosis*: molecular perspectives. *Emerg Infect Dis* 4, 195-209.

51. Martinez, M. D. A. M. B. (1999) Uptake of iron from different compounds by enteroinvasive *Escherichia coli*. *Rev. Microbiol.* 30, 149-152
52. Garcia, E., Nedialkov, Y. A., Elliott, J., Motin, V. L., and Brubaker, R. R. (1999) Molecular characterization of KatY (antigen 5), a thermoregulated chromosomally encoded catalase-peroxidase of *Yersinia pestis*. *J Bacteriol* 181, 3114-3122.
53. Brunder, W., Schmidt, H., and Karch, H. (1996) KatP, a novel catalase-peroxidase encoded by the large plasmid of enterohaemorrhagic *Escherichia coli* O157:H7. *Microbiology* 142, 3305-3315
54. Mehigh, R. J., and Brubaker, R. R. (1993) Major stable peptides of *Yersinia pestis* synthesized during the low- calcium response. *Infect Immun.* 61, 13-22
55. Bandyopadhyay, P., and Steinman, H. M. (2000) Catalase-Peroxidases of *Legionella pneumophila*: Cloning of the katA Gene and Studies of KatA Function. *J Bacteriol* 182, 6679-6686
56. Varnado, C. L., Hertwig, K. M., Thomas, R., Roberts, J. K., and Goodwin, D. C. (2004) Properties of a novel periplasmic catalase-peroxidase from *Escherichia coli* O157:H7. *Arch Biochem Biophys.* 421, 166-174
57. Bracken, C. S., Baer, M. T., Abdur-Rashid, A., Helms, W. and Stojiljkovic, I. (1999) Use of heme-protein complexes by the *Yersinia enterocolitica* HemR

- receptor: histidine residues are essential for receptor function. *J Bacteriol.* 181, 6063-6072
58. Beale, S. I., Neidhardt, F. C., Curtiss III, R., Ingraham, J. L., Lin, E. C. C., Low, K. B., Magasanik, B., Reznikoff, W. S., Riley, M., and Schaechter, M. (1996) Biosynthesis of hemes. In *Echerichia coli and Salmonella: cellular and molecular biology. 2nd edition.* (Umbarger, H. E., ed) pp. p731-748, ASM Press, Washington, D. C.
59. Crosa, J. H. (1984) The relationship of plasmid-mediated iron transport and bacterial virulence. *Ann. Rev. Microbiol.* 38
60. Litwin, M. C., and Calderwood, B. S. (1993) Role of iron in regulation of virulence genes. *Clin. Microbiol.* 6
61. Aisen, P., Leibman, A. & Zweier, J. (1978) Stoichiometric and site characteristics of the binding of iron in human transferrin. *J Biol Chem.* . 253, 1930–1937
62. Payne, S. M., and Torres, A. G. (1997) Haem iron transport system in enterohaemorrhagic Escherichia Coli O157:H7. *Molecular Microbiology* 23, 825-833
63. Wandersman, C., Stojiljkovic, I. (2000) Bacterial heme sources: the role of heme, hemoprotein receptors and hemophores. *Current Opinion in Microbiology* 3, 215-220
64. Braun, V. (2003) Iron uptake by Echerichia coli. *Front. Biosci.* 8, 1409-1421

65. Torres, A. G., Redford, P., Welch, R. A., and Payne, S. M. (2001) Ton-B dependent systems of uropathogenic *Escherichia coli*: Aerobactin and heme Transport and TonB are required for virulence in the mouse. *Infect Immun* . 69, 6179-6185
66. Law, D. (2000) Virulence factors of *Escherichia coli* 0157 and other Shiga toxin-producing *E. coli*. *Journal of Applied Microbiology* 88, 729-745
67. Bagg, A., Neilands, J. B. (1987) Molecular mechanism of regulation of siderophore-mediated iron assimilation. *Microbiological rev.* 51, 509-518
68. Ferguson, A. D., Hofmann, E., Coulton, J.W., Diedrichs, K, Wolfram, W. (1998) Siderophore-Mediated Iron Transport: Crystal Structure of FhuA with Bound Liposaccharide. *Science* 282, 2215-2220
69. Stojiljkovic (1999) Use of Heme-protein complexes by the *Yersinia enterocolitica* HemR receptor: Histidine residues are essential for receptor function. *J Bacteriol* 181, 6063-6072
70. Braun, M., Killmann, H., Braun, V. (1999) The β -barrel domain of FhuA Δ 5-160 is sufficient for TonB-dependent FhuA activities of *Escherichia coli*. *Molecular Microbiology* 33, 1037-1049
71. Banci, L., Bartalesi, I., Ciofi-Baffoni, S., and Tien, M. Unfolding and pH studies on manganese peroxidase: Role of heme and calcium on secondary structure stability. *Biopolymers* 72, 38-47

72. Law, D. (2000) Virulence factors of *Escherichia coli* 0157 and other Shiga toxin-producing *E. coli*. *Journal of Applied Microbiology*, 729-745
73. Pawelek, P. D., Croteau, N., Ng-Thow-Hing, C., Khursigara, C. M., Moiseeva, N., Allaire, M. & Coulton, J. W. (2006) Structure of TonB in complex with FhuA, *E. coli* outer membrane receptor. *Science* 312, 1399-
74. Shultis, D. D., Purdy, M. D., Banchs, C. N. & Wiener, M. C. (2006) Outer membrane active transport: Structure of the BtuB: TonB Complex. *Science* 312, 1396-1399
75. Postle, K., Letlain, Tracy E. (1997) TonB protein appears to transduce energy by shuttling between the cytoplasmic membrane and the outer membrane in *Escherichia coli*. *Molecular Microbiology* 24, 271-283.
76. Ghosh, A. (1996) *Chemical Biology* 3, 1011
77. Sambrook, J., Maniatis, T., and Fritsch, E. F. (1989) *Molecular cloning : a laboratory manual, 2nd ed.*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y
78. Claiborne, A., and Fridovich, I. (1979) 254 Purification of the o-dianisidine peroxidase from *Escherichia coli* B. Physicochemical characterization and analysis of its dual catalytic and peroxidatic activities. *J Biol Chem* 254, 4245-4252

79. Loewen, P. C. (1997) Bacterial catalases. In *Oxidative stress and the molecular biology of antioxidant defenses* pp. 273-308, Cold Spring Harbor Laboratory Press, Plainview, N.Y.
80. Chouchane, S., Giroto, S., Kapetanaki, S., Schelvis, J. P., Yu, S., and Magliozzo, R. S. (2003) Analysis of heme structural heterogeneity in Mycobacterium tuberculosis catalase-peroxidase (KatG). *J Biol Chem* 278, 8154-8162
81. Sutherland, G. R., and Aust, S. D. (1997) Thermodynamics of binding of the distal calcium to manganese peroxidase. *Biochemistry* 36, 8567-8573
82. Gadsby, P. M. A. a. T. A. J. (1990) Assignment of the axial ligands of ferric ion in low spin hemoproteins by near-infrared magnetic circular dichroism and electron paramagnetic resonance spectroscopy. *J. Am. Chem. Soc.* 112, 5003-5011
83. Banci, L., Bartalesi, I., Ciofi-Baffoni, S., and Tien, M. (2003) Unfolding and pH studies on manganese peroxidase: Role of heme and calcium on secondary structure stability. *Biopolymers* 72, 38-47
84. Sutherland, G. R., Zapanta, L. S., Tien, M., and Aust, S. D. (1997) Role of calcium in maintaining the heme environment of manganese peroxidase. *Biochemistry* 36, 3654-3662
85. Nie, G., and Aust, S. D. (1997) Spectral changes of lignin peroxidase during reversible inactivation. *Biochemistry* 36, 5113-5119

86. Poulos, T. L., Edwards, S. L., Wariishi, H., and Gold, M. H.. (1993) Crystallographic refinement of lignin peroxidase at 2 Å. *J Biol Chem.* . 268, 4429-4440
87. Sundaramoorthy, M., Kishi, K., Gold, M. H., and Poulos, T. L. (1994) The crystal structure of manganese peroxidase from *Phanerochaete chrysosporium* at 2.06-Å resolution. *J Biol Chem* 269, 32759-32767
88. Nie, G., and Aust, S. D. (1997) Effect of calcium on the reversible thermal inactivation of lignin peroxidase. *Arch Biochem Biophys.* 337, 225-231
89. Wada, K., Tada, T. Nakamura, Y., Kinoshita, T., Tamoi, M., Shigeoka, S, and Nishimura, K. (2002) Crystallization and preliminary X-ray diffraction studies of catalase-peroxidase from *Synechococcus* Pcc 7942. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 58, 157-159
90. Bertrand, T. E., N. A.; Jones, J. N.; Jesmin, N. A; Nagy, J. M.; Jamart-Gregoire, B.; Raven, E. L. and Brown K. A. (2004) Crystal structure of *Mycobacterium tuberculosis* catalase-peroxidase. *J. Biol. Chem.* 279, 38991-38999
91. Jakopitsch, C. I., A.; Schmuckenschlager, F.; Wanasinghe, A.; Poltl, G.; Furtmuller, P. G.; Ruker, F.; and Obinger, C. (2004) Influence of the unusual covalent adduct on the kinetics and formation of radical intermediates in *synechocystis* catalase-peroxidase: a stopped flow and EPR characterization of the MET275, TYR249, and ARG439 variants. *J. Biol. Chem.* 279

92. Jakopitsch, C., Auer, M., Regelsberger, G., Jantschko, W., Furtmuller, P. G., Ruker, F., and Obinger, C. (2003) Distal site aspartate is essential in the catalase activity of catalase-peroxidases. *Biochemistry* 42, 5292-5300
93. Ghiladi, R. A. M., K. F.; and Ortiz de Montellano, P. R. (2005) Role of the Met-Tyr-Trp cross-link in *Mycobacterium tuberculosis* catalase-peroxidase (KatG) as revealed by KatG-(M255I). *Biochemistry* 44
94. Jakopitsch, C., Auer, M., Regelsberger, G.; Jantschko, W.; Furtmuller, P. G.; Ruker, F.; and Obinger, C. (2003) The catalytic role of the distal site asparagine-histidine couple in catalase-peroxidases. *Eur. J. Biochem.* 270, 1006-1013
95. Jakopitsch, C., Auer, M., Ivancich, A., Ruker, F., Furtmuller, P. G., and Obinger, C. (2003) Total conversion of bifunctional catalase-peroxidase (KatG) to monofunctional peroxidase by exchange of a conserved distal side tyrosine. *J Biol Chem* 278, 20185-20191
96. Yu, S. G., S.; Zhao, X. and Magliozzo, R. S. (2003) Rapid formation of compound II and tyrosyl radical in the Y229F mutant of *Mycobacterium tuberculosis* catalase-peroxidase disrupts catalase but not peroxidase function. *J. Biol. Chem.* 278, 44121-44127
97. Ghiladi, R. A. K., G. M.; Medzizihradzky, K. F.; and Ortiz de Montellano, P. R. (2005) The Met-Tyr-Trp crosslink in *Mycobacterium tuberculosis* catalase-peroxidase (KatG): auto-catalytic formation and effect on enzyme catalysis and spectroscopic properties. *J. Biol. Chem.* 280, 22651-22663

98. Donald, L. J., Krokhin, O. V., Duckworth, H. W., Wiseman, B., Deemagarn, T., Singh, R., Switala, J., Carpena, X., Fita, I., and Loewen, P. C. (2003) Characterization of the catalase-peroxidase KatG from *Burkholderia pseudomallei* by mass spectrometry. *J Biol Chem* 278, 35687-35692
99. Eady, N. A. J. J., N. A.; Servos, S.; Cass, A. E. G.; Nagy, J. M. and Brown, K. A. (2005) Probing the function of *Mycobacterium tuberculosis* catalase-peroxidase by site-directed mutagenesis. *Dalton Trans.*, 3495-3500
100. Schuller, D. J., Ban, N., Huystee, R. B., McPherson, A., and Poulos, T. L. (1996) The crystal structure of peanut peroxidase. *Structure* 4, 311-321
101. Gajhede, M., Schuller, D. J., Henriksen, A., Smith, A. T., and Poulos, T. L. (1997) Crystal structure of horseradish peroxidase C at 2.15 Å resolution. *Nat Struct Biol* 4, 1032-1038
102. Shiro, Y., Kurono, M., and Morishima, I. (1986) Presence of endogenous calcium ion and its functional and structural regulation in horseradish peroxidase. *J Biol Chem* 261, 9382-9390
103. Reading, N. S., and Aust, S. D. (2000) Engineering a disulfide bond in recombinant manganese peroxidase results in increased thermostability. *Biotechnol Prog* 16, 326-333

104. Howes, B. D. F., A.; Raimondi, L.; Indiani, C.; Smulevich, G. (2001) The Critical Role of the Proximal Calcium Ion in the Structural Properties of Horseradish Peroxidase. *Journal of Biological Chemistry* 276, 40704-40711
105. Jakopitsch, C., Kolarich, D., Petutschnig, G., Furtmuller, P. G., and Obinger, C. (2003) Distal side tryptophan, tyrosine and methionine in catalase-peroxidases are covalently linked in solution. *FEBS Lett*, 135-140
106. Carpena, X. M.-A., W.; Loewen, P.C. (2004) Structure of the C-terminal domain of the catalase-peroxidase KatG from *Escherichia coli*. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* 60, 1824-1832
107. Tang, L. S., L. and Jaffe, E. K. (2005) Substrate-induced interconversion of protein quarternary structure isoforms. *J. Biol. Chem.* 280, 15786-15793
108. Erman, J. E., Vitello, L. B., Miller, M. A., Shaw, A., Brown, K. A., and Kraut, J. (1993) Histidine 52 is a critical residue for rapid formation of cytochrome c peroxidase compound I. *Biochemistry* 32, 9798-9806
109. Vitello, L. B. E., J. E., Miller, M. A., Wang, J. and Kraut, J. (1993) Effect of arginine-48 replacement on the reaction between cytochrome c peroxidase and hydrogen peroxide. *Biochemistry* 32, 9807-9818
110. Li, Y. a. G., D. C. (2004) Vital roles of an interhelical insertion in catalase-peroxidase bifunctionality. *Biochem. Biophys. Res. Commun.* 318, 970-976

