STRUCTURAL, SPECTROSCOPIC, AND MECHANISTIC INSIGHTS INTO THE THREE PHASES OF NITRIFICATION METALLOBIOCHEMISTRY

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy | Chemistry and Chemical Biology

By

Meghan Anne Smith

August 2018

STRUCTURAL, SPECTROSCOPIC, AND MECHANISTIC INSIGHTS INTO THE THREE PHASES OF NITRIFICATION METALLOBIOCHEMISTRY

Meghan Anne Smith, Ph. D.

Cornell University 2018

Biological ammonia (NH_3) oxidation, referred to as nitrification, is a critical part of the biogeochemical nitrogen cycle. Nitrification is mediated by both bacteria and archaea to ultimately oxidize NH_3 to nitrite (NO_2^-), though there are also complete NH_3-oxidizing (comammox) bacteria capable of oxidizing NH_3 completely to nitrate (NO_3^-). In addition to these products, nitrification is also a major source of the by-products and environmental pollutants nitric oxide (NO), nitrous oxide (N_2O) and nitrogen dioxide (NO_2). Many steps of biological nitrification, including those leading to the production of these harmful products, are not currently clear; however, the work presented in this dissertation describes recent efforts and discoveries towards a complete understanding of the nitrification pathway. This process begins in both bacteria and archaea with the enzyme ammonia monooxygenase (AMO), which oxidizes NH_3 to hydroxylamine (NH_2OH). There exist two metal-binding sites in AMO of interest as these are highly conserved in AMOs and related enzymes. The true active site of this enzyme remains in debate, but here we show that both sites must remain intact for effective catalysis. In bacteria, the formed NH_2OH is further oxidized to NO by the enzyme NH_2OH oxidoreductase (HAO), though prior convention stated that HAO was able to oxidize NH_2OH fully to NO_2^-. There exists another enzyme in

NH_3–oxidizing bacteria (AOB) known as cytochrome (cyt) P460 that can oxidize NH_2OH to NO and N_2O. Here we present structural and mechanistic studies that describe how the unusual P460 cofactor and surrounding amino acids allow for this catalysis. The recent discovery that the true product of HAO is NO and not NO_2^- presents a challenge to find the enzyme in AOB which can complete this oxidation to the final product NO_2^-. Here we present a potential candidate, nitrosocyanin (NC), and describe preliminary experiments on its interaction with NO.

BIOGRAPHICAL SKETCH

Meghan A. Smith attended Creighton University in Omaha, NE and graduated *cum laude* with honors in 2013. She completed a B.S. in Chemistry on the Biochemistry track with a minor in Cognitive and Behavioral Neuroscience. During her time at Creighton, she received a fellowship from the Institutional Development Award Program (IDeA) Networks of Biomedical Research Excellence (INBRE) program in Nebraska. As part of this fellowship she completed a summer of research in the lab of Prof. James R. Alfano at the University of Nebraska, Lincoln and three years of research in the lab of Prof. Karin van Dijk at Creighton University. This work was focused on effector protein interactions of the type three secretion system (TTSS) that lead to virulence in the bacteria *Pseudomonas syringae* pv. *tomato*. After graduating from Creighton, Meghan pursued her PhD at Cornell University where she worked for Prof. Kyle M. Lancaster on the enzymology supporting ammonia (NH_3) oxidation in both bacteria and archaea. The following describes her work completed at Cornell University under the guidance of Prof. Lancaster.

ACKNOWLEDGMENTS

I gratefully acknowledge all current and former members of the Lancaster Research Group for their support, advice, and, importantly, friendship. Specifically, I would like to recognize Dr. Jonathan D. Caranto for his help in designing and performing experiments on projects concerning AMO, cyt P460, and nitrosocyanin. As well, I would like the thank him for providing general guidance and always making sure we were on track and had set goals in mind. I would also like to acknowledge Dr. Avery C. Vilbert for her support and help in setting up various cyt P460 experiments. I'd like to acknowledge Dr. Sudipta Chatterjee, Ida M. DiMucci, and Sean H. Majer for help in experimental setup and techniques. I'd also like to thank Rachael Coleman, Ben Looker, and Dr. Richard Walroth for their support. In addition, I wish luck to Sean, Rachael, and Ben for their continued work on these projects! I'd finally like to acknowledge Prof. Kyle M. Lancaster for his continued guidance, advice, support, and music tastes.

I gratefully acknowledge Dr. Michael Fenwick for assistance with X-ray data collection and Prof. Brian Crane for advice about structural refinement. Further, Boris Dzikhovsky for his help collecting EPR spectra. In addition, Dr. Nicholas Coleman for supplying the pMycoFos vector and Prof. Harry B. Gray for providing the pEC86 plasmid containing the cyt c maturation genes ccmABCDEFGH.

Finally, I'd like to acknowledge all my friends and family, without whom none of this would have been possible!

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SCHEMES

CHAPTER 1: INTRODUCTION

Reprinted (adapted) with permission from:

"Alternative Bioenergy: Updates to and Challenges in Nitrification Metalloenzymology"

Kyle M. Lancaster, Jonathan D. Caranto, Sean H. Majer, and Meghan A. Smith

Joule, **2018,** *2(3)*, 421-441.

A. NITRIFICATION

Nitrification—the overall biological oxidation of ammonia (NH_3) to nitrate (NO_3^-)—represents a crucial component of the biogeochemical nitrogen cycle. This process has conventionally been split into two distinct steps: the oxidation of NH_3 to nitrite (NO_2^-) carried out by NH_3-oxidizing bacteria (AOB)[1] and archaea (AOA)[2] and the oxidation of NO_2^- to NO_3^- completed by NO_2^--oxidizing bacteria (NOB).[3] However, organisms have been recently discovered which are capable of completing the full 8-electron oxidation of NH_3 to NO_3^-, and are thus called "complete NH_3-oxidizing" or comammox bacteria.[5, 6]

Oxidation of NH_3 by the nitrification pathway is a major source of economic, energetic, and chemical waste. NH_3 produced for agriculture via the Haber-Bosch process requires ca. 3% of the energy produced worldwide annually, but crops assimilate only 10% of this fixed nitrogen.[7] Further, the products NO_2^- and NO_3^- are aquatic pollutants which lead to eutrophication and dead zones. As the population demands more fertilizer use for increased food stock, production of NO_2^- and NO_3^- are also expected to increase dramatically. NH_3-oxidizing organisms have also been shown to be sources of potent greenhouse gases nitrous oxide (N_2O)[8] and nitric oxide (NO).[9] Indeed, the environmental protection agency (EPA) has cited current agricultural practices as the dominant source of global N_2O production.[10] Thus, an understanding of the nitrification process is critical to reducing the environmental damage inflicted by increased fertilization.

Organisms that rely on nitrification for energy generation and thus survival can be found in every ecosystem. Despite this prevalence and the impact on the

environment, relatively little is known about the chemical processes involved in nitrification. Much of this is due to the difficulty in obtaining biochemically useful quantities of relevant enzymes in the process. Nitrifying microorganisms generally have very long doubling times (ca. 26 hrs for the NH_3-oxidizing archaeon *Nitrosomopumilus maritimus*)[11] and grow to low cell densities. Complicating matters, many of the enzymes involved in the nitrification pathway remain unknown, while many known proteins have eluded heterologous expression. What is currently known can be summarized in **Figure 1.1**. Briefly, NH_3-oxidation begins with hydroxylation of NH_3 to form hydroxylamine (NH_2OH) via the enzyme NH_3 monooxygenase (AMO). From there, the pathways for AOB and AOA diverge.

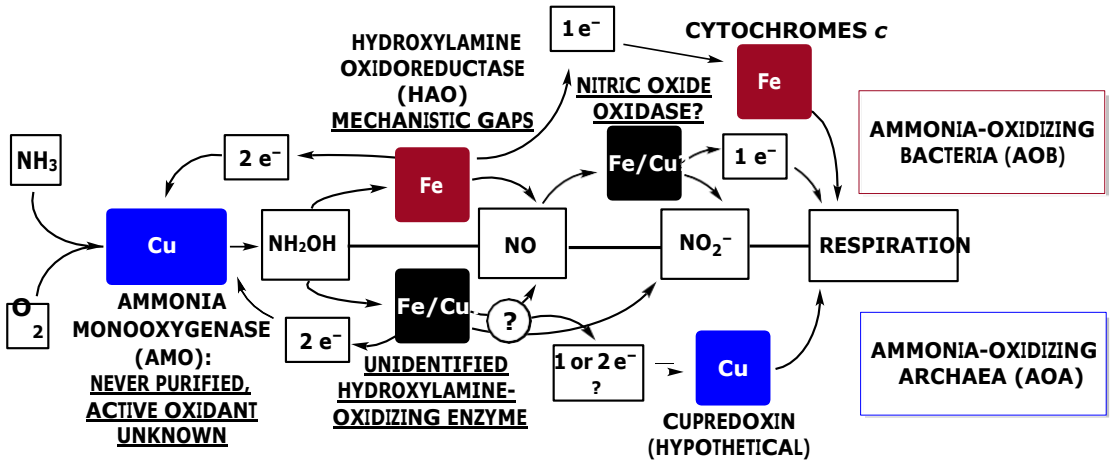
In AOB, the multiheme enzyme NH_2OH oxidoreductase (HAO) then oxidizes NH_2OH to NO, liberating 3 electrons in the process. As nitrifying organisms have been shown to produce stoichiometric amounts of NO_2^- from NH_3, there must exist a third, presently unknown enzyme in the pathway capable of oxidizing NO to NO_2^-. The enzymology of AOA is entirely unknown. In fact, AOA lack the cellular machinery required to form c type hemes present in HAO; thus, a completely novel enzyme must exist that is capable of transforming NH_2OH. The current state of knowledge of nitrification enzymology will be summarized below (**Figure 1.1**).

B. AMMONIA MONOOXYGENASE

AOB, AOA, and comammox bacteria initiate nitrification through the hydroxylation of NH_3 to hydroxylamine (NH_2OH) using the enzyme ammonia monooxygenase (AMO).[12-14] AMO is a member of the copper membrane monooxygenase (CuMMO) family,[15] which includes the methanotrophic enzyme

particulate methane monooxygenase (pMMO) family. The selective hydroxylation of methane to methanol effected by pMMO is essentially the same process carried out by AMO. Indeed, AMO has a broad substrate scope and can oxidize methane, alkenes,[16] methanol,[17] halogenated hydrocarbons,[18-19] aromatic compounds,[20] and sulfides.[21]

Figure 1.1: Summarized current understanding and critical knowledge gaps in the biochemistry of nitrification with distinctions made for ammonia oxidation by bacteria and archaea and nitrite oxidation by bacteria.



Detailed understanding of the mechanisms used by CuMMOs during the selective oxidations of challenging substrates would tremendously advance the development of controlled synthetic oxidation reactions. For AMO, however, little mechanistic insight into activity is available beyond substrate scope. This lack of information can be attributed to the difficulty in expressing and isolating the enzyme, as discussed earlier. AMO exists as an integral membrane protein and is therefore intrinsically difficult to purify in an active state.[22] A soluble form of AMO has also

been identified, based on NH_3 oxidation in soluble fractions of *Nitrosomonas europaea*.[23] This soluble AMO contains Cu similar to that observed in pMMO, as well as some Fe and possible Zn.[24] However, it should be noted that no activity was reported for the purified soluble AMO. Such loss of activity invariably impedes AMO biochemistry—membrane-bound AMO from AOB remains active in cell lysates only under strict conditions, including the presence of exogenous stabilizing agents and excess Cu.[25] Moreover, recombinant DNA technology has failed to advance AMO studies; AMO genes are toxic to *Escherichia coli*,[26] and no alternative recombinant expression systems have been reported. Thus, all available data concerning AMO activity are based on whole-cell or cell lysate experiments.

NH_2OH was first suggested as an obligate intermediate of NH_3 oxidation after the observation that *Nitrosomonas* cell suspensions can convert NH_2OH to NO_2^-.[27] This hypothesis was verified by the observation that *Nitrosomonas* cell suspensions containing hydrazine failed to produce NO_2^- and accumulated NH_2OH. Together, these results suggested that NH_2OH is an intermediate of AOB metabolism and that its conversion to NO_2^- is inhibited by hydrazine. NH_3 oxidation to NH_2OH can be selectively inhibited by acetylene or allylthiourea in both AOB and AOA, and both compounds are commonly used as nitrification inhibitors.[28-29, 30]

NH_3 oxidation was definitively pinpointed to AMO through the combined inactivation and labeling of the NH_3 oxidation inhibitor acetylene.[31] However, the existence of an enzyme at this point in the nitrification reaction was recognized before then and known to be found in the cell membrane and likely contain Cu.[32] AMO activity requires dioxygen (O_2), and isotopic labeling studies have shown that the O in

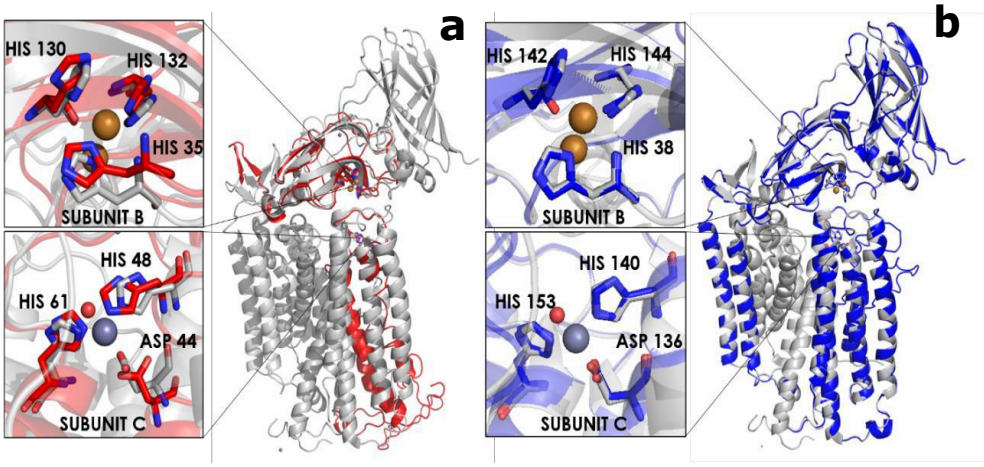
NH_2OH originates from O_2.[28, 33] Thus, AMO hydroxylation of NH_3 to NH_2OH requires O_2 activation, in which one O is incorporated into NH_2OH and the other is reduced to H_2O. Consistent with this requirement, the addition of NH_3 to *Nitrosomonas europaea* cell suspensions increases the rate of O_2 consumption.[34] This process is believed to be Cu-dependent, as O_2-activation activity can be restored to metal-depleted AOB membrane fractions via the addition of Cu ions.[25] The conversion of NH_3 to NH_2OH is a 2-electron oxidation, whereas O_2 is a 4-electron oxidant. Thus, O_2 activation likely requires the input of 2 electrons (**Eqn. 1.1**), and NH_2OH oxidation to NO_2^- acts as the electron source (*vide infra*).[21, 35]

$$NH_3 + O_2 + 2H^+ + 2e^- \rightarrow NH_2OH + H_2O \qquad \textbf{(1.1)}$$

Further mechanistic insights into the O_2 activation and NH_3 hydroxylation of AMO have thus far eluded biochemists. However, some understanding may be gleaned from studies of related enzymes. The reactivity, genetics, and structures of pMMOs have been characterized. These enzymes adopt an $\alpha_3\beta_3\gamma_3$ trimer quaternary structure, and analysis of the AMO gene cluster suggests that AMO assembles in a similar manner. AMO from AOA may also include a fourth subunit, amoX, which is postulated to assume the same yet unconfirmed role as the exogenous helix found in some pMMO structures.[36] Crystal structures of pMMOs reveal two metal-binding sites, the residues of which are highly conserved among all CuMMOs (**Figure 1.2**).[36, 37]

Although archaeal AMOs (aAMOs) and bacterial AMOs (bAMOs) likely have the same overall chemistry, they exhibit kingdom-specific NH_3 oxidation characteristics. Both AOA[29] and AOB[21] produce NH_2OH as an NH_3 oxidation

Figure 1.2: Structural homology in CuMMOs.



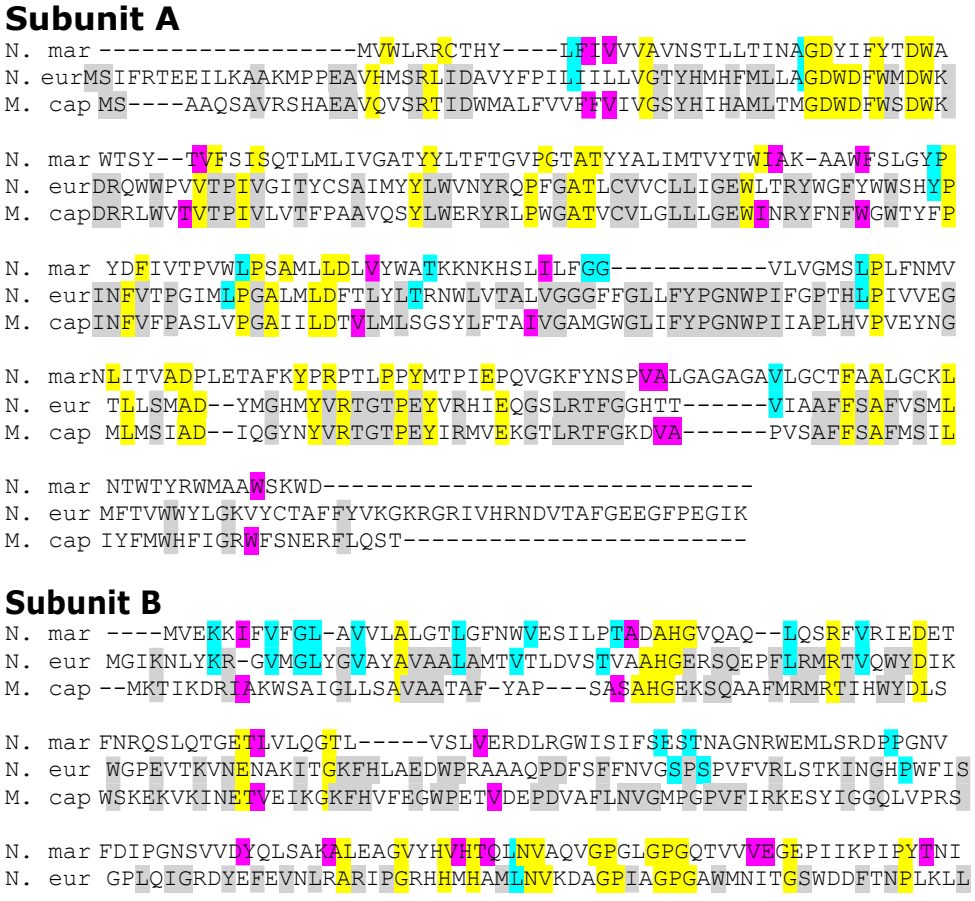
(a) Structural homology model of *N. maritimus* AMO subunits B and C threaded on to the structure of homologous subunits of *M. capsulatus* (Bath) pMMO (grey, PDB ID: 3RGB). Labels indicate the residue numbers of *N. maritimus* AMO. Orange spheres are copper ions, grey sphere is a zinc ion, red sphere is a water molecule, all from the 3RGB structure. (b) Structural homology model of *N. europaea* AMO subunits B and C threaded on to the structure of homologous subunits of *M. capsulatus* (Bath) pMMO (grey, PDB ID: 3RGB). Labels indicate the residue numbers of *N. europaea* AMO. Orange spheres are copper ions, grey sphere is a zinc ion, red sphere is a water molecule, all from the 3RGB structure. Models were generated using SWISS-MODEL Workspace.[4]

intermediate, which suggests that aAMOs and bAMOs have similar activities. Amino acid sequence homology and structural homology modeling of these enzymes also show that both types likely contain the same active site(s). However, although bAMOs share ~70% amino acid identity with pMMOs, only ~40% amino acid identity is found between bAMOs and aAMOs (**Figure 1.3**).[38] Therefore, bAMOs are more closely related to another bacterial enzyme with a different function than they are to their archaeal counterparts. Given that the substrates and competitive inhibitors for AMO are largely non-polar, the active site of AMO is likely hydrophobic. This arrangement

is consistent with NH_3 rather than ammonium (NH_4^+) as the substrate for bAMO.[28] However, some AOA (but not AOB) thrive at low pH (2.5), and thus, the substrate for these aAMOs could be NH_4^+, which has a pKa of 9.26.[38] In addition to their potential differences in substrate, aAMOs and bAMOs have different activity-based inhibition profiles, which has been suggested to indicate the possibility for different enzymatic intermediates.[39–30] A study comparing the inhibition of *Nitrososphaera viennensi*s (AOA) and *Nitrosospira multiformis* (AOB)[39] showed that the effective concentration 50 of allylthiourea was 1000 times higher for the archaeon. The linear terminal alkynes from C_1 to C_9 also inhibit or partially inhibit bAMOs, whereas aAMOs are relatively unaffected by C_6 to C_9 1-alkynes.[40] Indeed, 1-octyne can be used in culture to select for the growth of AOA over AOB. For example, *Nitrosopumilus maritimus* (AOA) is completely unaffected by ≤20 μM 1-octyne, whereas NH_3 oxidation is completely and irreversibly inhibited at 1 μM 1-octyne in the AOB *N. europaea* and *N. multiformis.* However, without additional data, it is impossible to confirm whether this outcome indicates differences in enzymatic intermediates, differences in substrate pockets, or Cu lability that ultimately affect activity. Additionally, a study comparing the temperature effects on nitrification activity in AOA and AOB found that maximum, minimum, and optimum temperatures for activity were different between the two kingdoms, suggesting the possibility for differences in thermodynamic properties of the respective AMOs.[41] These differences correspond to differences in heat capacity, which can be related to conformational states in enzymes. The authors suggest that this difference in heat capacity, and thus conformational states or "flexibility," accords with the differences observed in the

substrate range of aAMO and bAMO. Of course, more information is needed to determine how/if these observations relate specifically to AMO and no other factors, though as the authors report, this is an important reason why models predicting nitrification effects on the environment have not been very successful at accurately incorporating temperature.

Figure 1.3: Sequence homology between AOA *Nitrosopumilus maritimus* (N. mar) and AOB *Nitrosomonas europaea* (N. eur) AMO sequences, and CH_4-oxidizing bacterium *Methylococcus capsulatus* (M. cap) pMMO sequence. Identity between all three sequences is highlighted in yellow, identity between AOA and AOB AMO sequences is highlighted in cyan, identity between AOA AMO and pMMO is highlighted in purple, and identity between AOB AMO and pMMO is highlighted in grey. Sequence alignment generated using the National Center for Biotechnology Information (NCBI) Nucleotide BLAST.[42, 43]



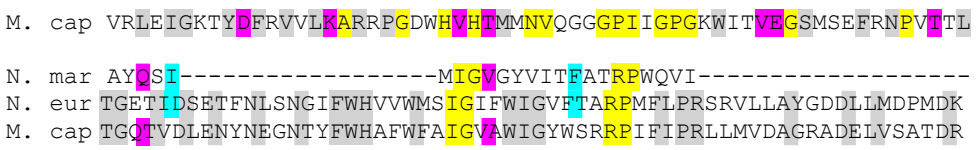
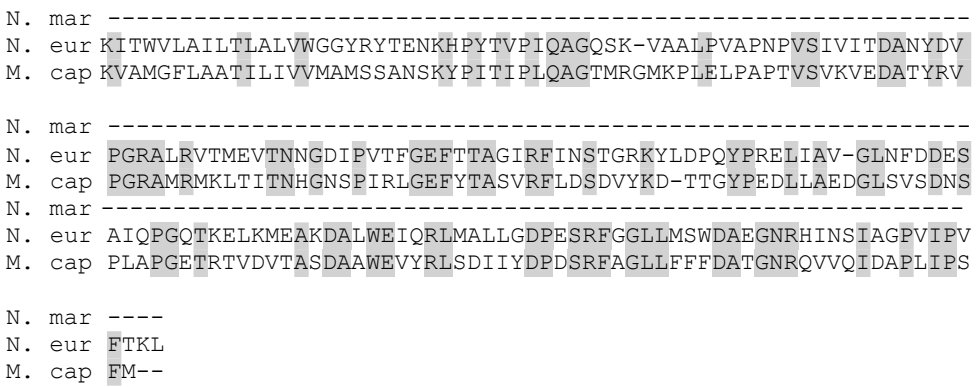
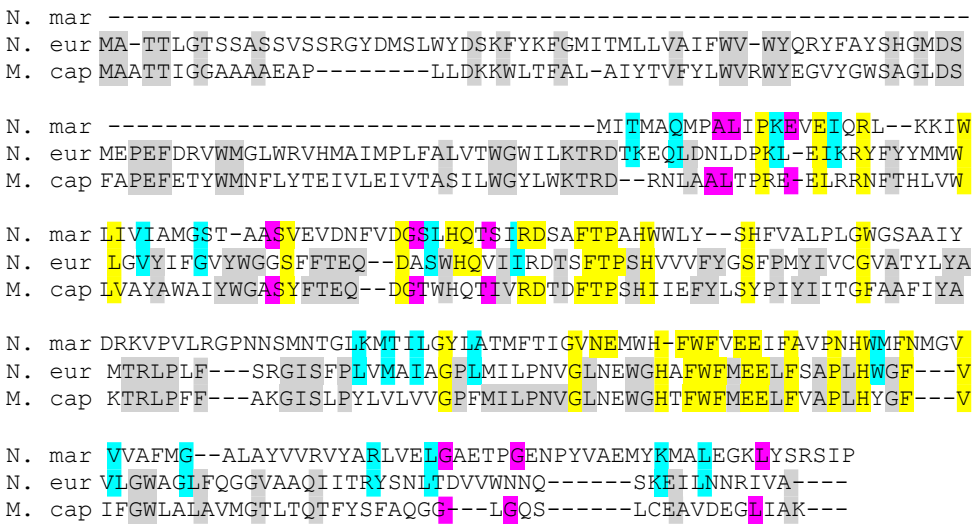


Figure 1.3 (Continued)



Subunit C



AOA have been increasingly discovered in new and diverse habitats, possibly (but not necessarily) due to advantages dictated by distinct AMOs. AOA can grow in various NH_3 concentrations, historically reported as much lower than those required for certain AOB growth. Compared with AOB *Nitrosomonas europaea* and *Nitrosococcus oceani*, AOA can have a specific affinity *that is 200 times higher*

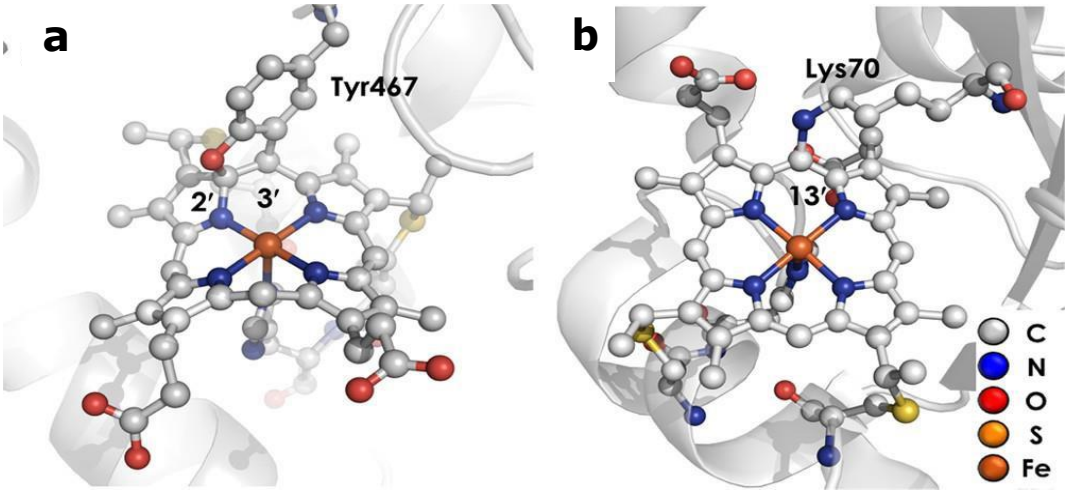
(where specific affinity is a^o_s, = V_{max}/K_m^{44}).[11] In a study comparing the O_2 uptake of N. maritimus (AOA), N. europaea, and N. oceani, the addition of 0.2 μM ammonium chloride was adequate to elicit 50% max O_2 uptake in N. maritimus, whereas no perturbed O_2 uptake was observed in the AOB.[11] The high affinity may explain the success of AOA in various environments. However, these studies were performed in whole cells; thus, the specific affinity for NH_3 may not reflect that of the aAMO. Indeed, one hypothesis is that AOA merely have a larger surface area to volume ratio, which would affect the apparent substrate affinity in experiments performed with whole cells. Moreover, the recent insights into comammox bacteria suggest that these NH_3-oxidizers may be even more environmentally competitive than AOA. Kits and co-workers[45] showed that the specific affinity for *Nitrospira inopinata* was higher than several AOA, including *Nitrososphaera gargensis*, *Nitrososphaera viennensis*, and 'Candidatus (Ca.) *Nitrosotenuis uzonensis*.' These authors further showed that AOB from *Nitrosomonas* gene cluster 6A also had specific affinities similar to non-marine AOA, disproving the generality that all AOA have higher affinities than AOB. Regardless of which kingdom is the more competitive NH_3-oxidizer, these observations emphasize differences between the nitrification pathways of archaea and bacteria, as further evidenced by the downstream enzymology leading to NO_2^- production. Without more information, it is presently unclear if these affinity differences are due to specific features of individual AMOs or other microbial factors. Either way, these differences highlight the importance of understanding not only AOB but also AOA and comammox biochemistry to fully appreciate the effects of nitrification on the environment.

C. HYDROXYLAMINE OXIDOREDUCTASE AND CYTOCHROME P460

HAO is a periplasmic, homotrimeric enzyme that contains 8 c-type hemes per monomer.[46] Seven of these c-types hemes are coordinatively-saturated and used for electron transfer. The eighth heme is a highly ruffled and unusual c-type heme where catalysis occurs. Due to the presence of a UV-vis feature at 460 nm in the ferrous state, this heme has been termed the "heme P460." In addition to the two thioether linkages common to all c-type hemes, the P460 heme cofactor contains additional covalent attachments to the protein backbone, via a tyrosine cross-linked to 2′ pyrrole α-C and 3′ *meso* C in HAO (**Figure 1.4**). This results in a highly distorted heme cofactor capable of performing the next step in nitrification, conversion of NH_2OH to NO.

There exists another enzyme with the heme P460 cofactor named cytochrome (cyt) P460. This enzyme forms a dimer with just 1 heme—the P460 heme—per monomer, making it much more amenable to spectroscopic analysis.[47] The heme P460 cofactor of cyt P460 also contains additional covalent attachments to the protein backbone, but unlike HAO contains only 1 additional cross-link via a lysine to the 13′ *meso* C (**Figure 1.4**). Until recently, both HAO and cyt P460 were reported to produce NO_2^- from NH_2OH.[47-52] However, HAO was never able to produce NO_2^- in stoichiometric amounts and required the presence of O_2. Anaerobic studies using cyt P460 showed that the true products of NH_2OH-oxidation for this enzyme are NO and N_2O and not NO_2^-.[53] N_2O formation is preceded by formation of $\{FeNO\}^7$ (following Enemark-Feltham notation,[54] where the number 7 indicates the number of Fe d-orbital electrons + the number of electrons in the NO π* orbital) and $\{FeNO\}^6$

Figure 1.4: Heme P460 Cofactors

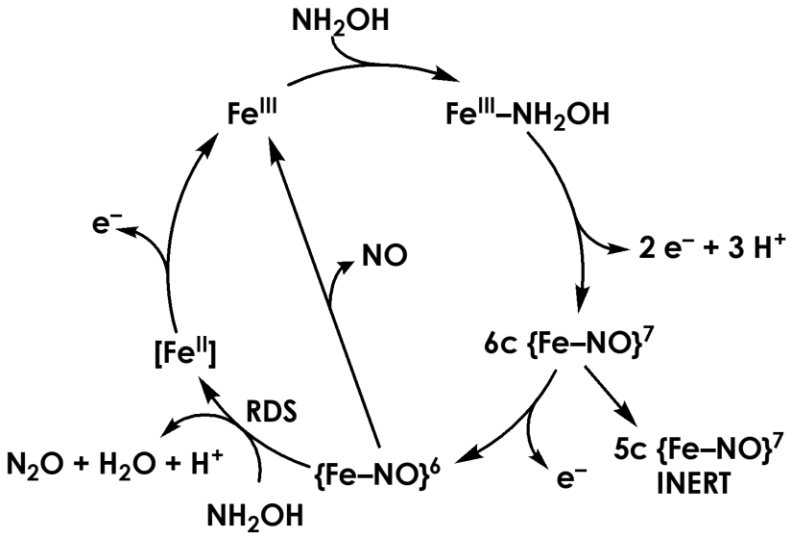


Heme P460 cofactors from *N. europaea* (a) hydroxylamine oxidoreductase (PDB ID: 4FAS) and (b) cytochrome P460 (PDB ID: 2JE3).

intermediates.[53, 55] The {FeNO}^{6} intermediate is then attacked by a second molecule of NH_2OH to form N_2O. This discovery in the true nature of cyt P460 catalysis prompted a revisit to HAO chemistry. Indeed, it was shown that anaerobically HAO produced no NO_2^-, and instead exclusively produced NO.[56] Isotope-labeling studies of AOB have shown that one O in the final product NO_2^- originates from O_2 and the other from H_2O.[57] It is known that the O in NH_2OH is from O_2,[58] thus the second O in NO_2^- must derive from H_2O, and would be inconsistent with HAO requiring O_2 for NO_2^- production. Presumably the NO_2^- that had been formed in the initial aerobic studies was the result of NO reacting with O_2 and would explain why stoichiometric NO_2^- production was never observed. This finding also agrees with studies on the HAO from anammox bacteria *Kuenenia stuttgartiensis*, shown to produce NO and not NO_2^-.[59]

Given the similarities in their cofactors, it is assumed that HAO and cyt P460 follow similar mechanisms but must differ in the relative stabilities of the $\{FeNO\}^6$ intermediates. That is, HAO appears biased towards loss of NO at this stage, while cyt P460 has the additional ability to produce N_2O given the availability of additional NH_2OH. In cyt P460, Fe^{III}–NH_2OH, 6-coordinate (6c)$\{FeNO\}^7$, an inactive 5-coordinate (5c), and $\{FeNO\}^6$ intermediates have been explicitly characterized, and a working mechanism has been proposed[53, 55] (**Figure 1.5**). As mentioned above, the presence of 7 non-catalytic hemes per monomer of HAO can confound spectroscopic analysis of intermediates formed during HAO catalysis. There is some evidence for an $\{FeNO\}^6$ species,[60] however, and it is expected that the HAO mechanism will be similar to that for cyt P460 (**Figure 1.5**).

Figure 1.5: Proposed Cyt P460 Catalytic Mechanism Based on Anaerobic Studies of Cyt P460 from *Nitrosomonas europaea*.[55]



Interestingly, cyt P460s are found in a variety of microorganisms other than AOB, including methanotrophs, proteobacteria, planctomycetes, acidobacteria, and bacteroidetes.[61, 62] Though the majority of work on cyt P460 has been on the variant from *N. europaea*, we also obtained cyt P460 from a related AOB *Nitrosomonas* sp. AL212. This new cyt P460 showed the same spectroscopic characteristics as that from *N. europaea*; however, it was discovered that the wild type variant from *Nitrosomonas* sp. AL212 was unable to oxidize hydroxylamine, despite being able to form the Fe^{III}–NH_2OH and $\{FeNO\}^6$ intermediates (*vide infra*). As will be discussed in Chapter 4, it was found that a residue in the second coordination sphere (a glutamate in *N. europaea* cyt P460 and an alanine in *Nitrosomonas* sp. AL212) played a crucial role in catalysis. In fact, mutating the alanine in the *Nitrosomonas* sp. AL212 cyt P460 to a glutamate was sufficient to imbue the previously inactive protein with the ability to oxidize NH_2OH. This finding opens the possibility for different classes of cyt P460 proteins. That is, it is possible that not all cyt P460s oxidize hydroxylamine, but that the second coordination spheres of these enzymes are tuned towards different reactivities. This may explain the widespread nature of cyt P460s in a variety of different types of microorganisms and will likely be an exciting future area of study.

D. NITRIC OXIDE OXIDATION

Revision of the nitrification pathway in AOB to reflect the true enzymatic product of HAO catalysis as NO and not NO_2^- has left a knowledge gap for the enzyme responsible for final conversion to NO_2^-. As stoichiometric conversion of NH_3 to NO_2^- is observed in AOB, it is unlikely that the NO_2^- obtained is the result of uncatalyzed reaction of NO with O_2. Including a third enzyme in the pathway would

also permit the full 4 electrons resulting from oxidation of NH_2OH to be captured by the AOB and used for cellular respiration.

Candidates for this nitric oxide oxidase (NOO) have been proposed. One possibility is the Cu-containing NO_2^- reductase, NirK. This enzyme is generally assumed to reduce NO_2^- to NO as it does in other organisms, but the NirK of *Alcaligenes faecalis* was shown to also be competent for the reverse reaction, oxidizing NO_2^- to NO.[63] Another possibility is the unusual red Cu protein nitrosocyanin.[64] Nitrosocyanin contains protein folds that resemble those found in typical type 1 or blue Cu proteins used for electron transfer. However, the canonical type 1 active site, such as those found in azurin and plastocyanin contain two histidines, one cysteine, and one axially coordinated methionine in a trigonally distorted tetrahedral geometry is replaced with by two histidines, a cysteine, a glutamate, and a solvent molecule that is lost upon reduction from Cu(II) to Cu(I).[64-66] These structural properties suggest nitrosocyanin is likely a poor electron transfer protein, and may instead have a catalytic function. Currently the role of nitrosocyanin is unknown; however, nitrosocyanin transcripts are present in high amounts, similar to those of other metabolic enzymes AMO and HAO. It has also been shown elsewhere that nitrosocyanin is upregulated during NH_3-oxidation or when cells are exposed to NO.[67-69] In the AOB *Nitrosomonas eutropha*, the nitrosocyanin gene is headed by a fumarate-nitrate reduction (FNR) protein binding region in the promotor sequence.[69, 70] FNR is a regulatory protein that has been implicated in response to low oxygen or the presence of NO.[71, 72] Nitrosocyanin is only found in AOB, and the available transcript data would suggest it has a role relevant directly to nitrification, particularly

in regards to NO. Notably, an AOB lacking the gene for nitrosocyanin, *Nitrosomonas* sp. Is79A3, produces large amounts of NO during NH_3-oxidation, potentially because nitrosocyanin acts as a NOO. As will be discussed in Chapter 5, nitrosocyanin serves as a likely candidate for a NOO, however more experiments which can explicitly measure this activity must be completed to be certain.

E. ADDRESSED IN THIS DISSERTATION

This introduction was intended to both describe the current understanding of nitrification enzymology and highlight the areas of uncertainty that remain. The work discussed in this dissertation addresses these areas of uncertainty at various levels of detail, including a targeted molecular level (i.e. identifying key amino acid residues) and a more global metabolic level (i.e. identifying unknown enzymes in the pathway). The second chapter will describe a recombinant expression platform for an aAMO and mutagenesis work designed to explore the nature of conserved metal–binding sites as they relate to AMO–dependent O_2–consumption. The third and fourth chapter describe the nature of inner and second sphere factors which affect the ability of cyt P460 to oxidize NH_2OH. The fifth and final chapter includes a recombinant expression system for nitrosocyanin, a candidate for the missing NOO, as well as preliminary experiments which probe the interaction of nitrosocyanin with NO.

F. REFERENCES

(1) Klotz, M. G., and Stein, L. Y. (2007) Nitrifier genomics and evolution of the nitrogen cycle, *FEMS Microbiol. Lett. 278*, 146-156.

(2) Walker, C. B., de la Torre, J. R., Klotz, M. G., Urakawa, H., Pinel, N., Arp, D. J., Brochier-Armanet, C., Chain, P. S., Chan, P. P., Gollabgir, A., Hemp, J., Hugler, M., Karr, E. A., Konneke, M., Shin, M., Lawton, T. J., Lowe, T., Martens-Habbena, W., Sayavedra-Soto, L. A., Lang, D., Sievert, S. M., Rosenzweig, A. C., Manning, G., and Stahl, D. A. (2010) Nitrosopumilus maritimus genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea, *Proc Natl Acad Sci USA 107*, 8818-8823.

(3) Lucker, S., Wagner, M., Maixner, F., Pelletier, E., Koch, H., Vacherie, B., Rattei, T., Damste, J. S., Spieck, E., Le Paslier, D., and Daims, H. (2010) A Nitrospira metagenome illuminates the physiology and evolution of globally important nitrite-oxidizing bacteria, *Proc Natl Acad Sci USA 107*, 13479-13484.

(4) Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics 22*, 195-201.

(5) van Kessel, M. A., Speth, D. R., Albertsen, M., Nielsen, P. H., Op den Camp, H. J., Kartal, B., Jetten, M. S., and Lucker, S. (2015) Complete nitrification by a single microorganism, *Nature 528*, 555-559.

(6) Daims, H., Lebedeva, E. V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., Jehmlich, N., Palatinszky, M., Vierheilig, J., Bulaev, A., Kirkegaard, R. H., von Bergen, M., Rattei, T., Bendinger, B., Nielsen, P. H., and Wagner, M. (2015) Complete nitrification by Nitrospira bacteria, *Nature 528*, 504-509.

(7) Erisman, J. W., Sutton, M. A., Galloway, J., Klimont, Z., and Winiwarter, W. (2008) How a century of ammonia synthesis changed the world, *Nat. Geosci. 1*, 636-639.

(8) Goreau, T. J., Kaplan, W. A., Wofsy, S. C., McElroy, M. B., Valois, F. W., and Watson, S. W. (1980) Production of NO2 – and N2O by nitrifying bacteria at reduced concentrations of oxygen, *Appl. Environ. Microbiol. 40*, 526-532.

(9) Kampschreur, M. J., Picioreanu, C., Tan, N., Kleerebezem, R., Jetten, M. S., and van Loosdrecht, M. C. (2007) Unraveling the source of nitric oxide emission during nitrification, *Water Environ Res. 79*, 2499-2509.

(10) (2018) Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2016, ((EPA), U. S. E. P. A., Ed.), Washington, D. C. .

(11) Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R., and Stahl, D. A. (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria, *Nature 461*, 976-979.

(12) Kits, K. D., Sedlacek, C. J., Lebedeva, E. V., Han, P., Bulaev, A., Pjevac, P., Daebeler, A., Romano, S., Albertsen, M., Stein, L. Y., Daims, H., and Wagner, M. (2017) Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle, *Nature 549*, 269-272.

(13) Vajrala, N., Martens-Habbena, W., Sayavedra-Soto, L. A., Schauer, A., Bottomley, P. J., Stahl, D. A., and Arp, D. J. (2013) Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea, *Proc Natl Acad Sci USA 110*, 1006-1011.

(14) Hooper, A. B., Arciero, D., Bergmann, D., and Hendrich, M. P. (2004) Chapter 6: The Oxidation of Ammonia as an Engergy Source in Bacteria. , In *Respiration in Archaea and Bacteria. Advances in Photosynthesis and Respiration* (D., Z., Ed.), Springer, Dordrecht.

(15) Tavormina, P. L., Orphan, V. J., Kalyuzhnaya, M. G., Jetten, M. S., and Klotz, M. G. (2011) A novel family of functional operons encoding methane/ammonia monooxygenase-related proteins in gammaproteobacterial methanotrophs, *Environ. Microbiol. Rep. 3*, 91-100.

(16) Hyman, M. R., Murton, I. B., and Arp, D. J. (1988) Interaction of Ammonia Monooxygenase from Nitrosomonas europaea with Alkanes, Alkenes, and Alkynes, *Appl. Environ. Microbio. 54*, 3187-3190.

(17) Voysey, P. A., and Wood, P. M. (1987) Methanol and Formaldehyde Oxidation by an Autotrophic Nitrifying Bacterium, *J. Gen. Microbiol. 33*, 283-290.

(18) Aciero, D., Vannelli, T., Logan, M., and Hooper, A. B. (1989) Degradation of trichloroethylene by the ammonia-oxidizing bacterium Nitrosomonas europaea, *Biochem. Biophys. Res. Commun. 159*, 640-643.

(19) Vannelli, T., Logan, M., Aciero, D., and Hooper, A. B. (1990) Degradation of halogenated aliphatic compounds by the ammonia-oxidizing bacterium Nitrosomonas europaea, *Appl. Environ. Microbiol. 56*, 1169-1171.

(20) Keener, W. K., and Arp, D. J. (1994) Transformations of aromatic compounds by Nitrosomonas europaea, *Appl. Environ. Microbiol. 60*, 1914-1920.

(21) Arp, D. J., Sayavedra-Soto, L. A., and Hommes, N. G. (2002) Molecular biology and biochemistry of ammonia oxidation by Nitrosomonas europaea, *Arch. Microbiol. 178*, 250-255.

(22) Ma, P., Varela, F., Magoch, M., Silva, A. R., Rosario, A. L., Brito, J., Oliveira, T. F., Nogly, P., Pessanha, M., Stelter, M., Kletzin, A., Henderson, P. J., and Archer, M. (2013) An efficient strategy for small-scale screening and production of archaeal membrane transport proteins in Escherichia coli, *PLoS One 8*, e76913.

(23) Gilch, S., Meyer, O., and Schmidt, I. (2009) A soluble form of ammonia monooxygenase in Nitrosomonas europaea, *Biol. Chem. 390*, 863-873.

(24) Gilch, S., Meyer, O., and Schmidt, I. (2010) Electron paramagnetic studies of the copper and iron containing soluble ammonia monooxygenase from Nitrosomonas europaea, *BioMetals 23*, 613-622.

(25) Ensign, S. A., Hyman, M. R., and Arp, D. J. (1993) In vitro activation of ammonia monooxygenase from Nitrosomonas europaea by copper, *J. Bacteriol. 175*, 1971-1980.

(26) Klotz, M. G., and Norton, J. M. (1998) Multiple copies of ammonia monooxygenase (amo) operons have evolved under biased AT/GC mutational pressure in ammonia-oxidizing autotrophic bacteria, *FEMS Microbiol. Lett. 168*, 303-311.

(27) Nicholas, D. J. D., and Jones, O. T. G. (1960) Oxidation of hydroxylamine in cell-free extracts of Nitrosomonas europaea, *Nature 185*, 512-514.

(28) Bock, E., and Wagner, M. (2006) Oxidation of Inorganic Nitrogen Compounds as an Energy Source, *Prokaryotes: A Handbook on the Biology of Bacteria, Vol 2, Third Edition*, 457-495.

(29) Vajrala, N., Martens-Habbena, W., Sayavedra-Soto, L. A., Schauer, A., Bottomley, P. J., Stahl, D. A., and Arp, D. J. (2013) Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea, *Proc Natl Acad Sci USA 110*, 1006-1011.

(30) Taylor, A. E., Taylor, K., Tennigkeit, B., Palatinszky, M., Stieglmeier, M., Myrold, D. D., Schleper, C., Wagner, M., and Bottomley, P. J. (2015) Inhibitory effects of C2 to C10 1-alkynes on ammonia oxidation in two Nitrososphaera species, *Appl. Environ. Microbiol. 81*, 1942-1948.

(31) Hyman, M. R., and Wood, P. (1985) Suicidal inactivation and labelling of ammonia mono-oxygenase by acetylene, *Biochem J. 227*, 719-725.

(32) Hyman, M. R., and Wood, P. (1983) Methane oxidation by *Nitrosomonas europaea*, *Biochem J. 212*, 31-37.

(33) Dua, R., Bhandari, B., and Nicholas, D. (1979) Stable isotope studies on the oxidation of ammonia to hydroxylamine by Nitrosomonas europaea, *FEBS Lett. 106*, 401-404.

(34) Meyerhof, O. (1917) Untersuchungen über den Atmungsvorgang nitriflzierender Bakterien, *Pflügers Archiv European J. Physiol. 166*, 240-280.

(35) Whittaker, M., Bergmann, D., Arciero, D., and Hooper, A. B. (2000) Electron transfer during the oxidation of ammonia by the chemolithotrophic bacterium Nitrosomonas europaea, *Biochim. Biophys. Acta. Bioenergetics 1459*, 346-355.

(36) Culpepper, M. A., and Rosenzweig, A. C. (2012) Architecture and active site of particulate methane monooxygenase, *Crit. Rev. Biochem. Mol. Biol. 47*, 483-492.

(37) Liew, E. F., Tong, D., Coleman, N. V., and Holmes, A. J. (2014) Mutagenesis of the hydrocarbon monooxygenase indicates a metal centre in subunit-C, and not subunit-B, is essential for copper-containing membrane monooxygenase activity, *Microbiology 160*, 1267-1277.

(38) Stahl, D. A., and de la Torre, J. R. (2012) Physiology and diversity of ammonia-oxidizing archaea, *Annu. Rev. Microbiol. 66*, 83-101.

(39) Shen, T., Stieglmeier, M., Dai, J., Urich, T., and Schleper, C. (2013) Responses of the terrestrial ammonia-oxidizing archaeon Ca. Nitrososphaera viennensis and the ammonia-oxidizing bacterium Nitrosospira multiformis to nitrification inhibitors, *FEMS Microbiol. Lett. 344*, 121-129.

(40) Taylor, A. E., Vajrala, N., Giguere, A. T., Gitelman, A. I., Arp, D. J., Myrold, D. D., Sayavedra-Soto, L., and Bottomley, P. J. (2013) Use of Aliphatic n-Alkynes To Discriminate Soil Nitrification Activities of Ammonia-Oxidizing Thaumarchaea and Bacteria, *Appl. Environ. Microbiol. 79*, 6544-6551.

(41) Taylor, A. E., Giguere, A. T., Zoebelein, C. M., Myrold, D. D., and Bottomley, P. J. (2017) Modeling of soil nitrification responses to temperature reveals thermodynamic differences between ammonia-oxidizing activity of archaea and bacteria, *Isme J 11*, 896-908.

(42) Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool, *J. Mol. Biol. 215*, 403-410.

(43) Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R., and Schäffer, A. A. (2008) Database indexing for production MegaBLAST searches, *Bioinformatics 24*, 1757-1764.

(44) Button, D. K. (1998) Nutrient uptake by microorganisms according to kinetic parameters from theory as related to cytoarchitecture, *Microbiol. Mol. Biol. Rev. 62*, 636-645.

(45) Kits, K. D., Sedlacek, C. J., Lebedeva, E. V., Han, P., Bulaev, A., Pjevac, P., Daebeler, A., Romano, S., Albertsen, M., and Stein, L. Y. (2017) Kinetic analysis of a complete nitrifier reveals an oligotrophic lifestyle, *Nature 549*, 269.

(46) Cedervall, P., Hooper, A. B., and Wilmot, C. M. (2013) Structural studies of hydroxylamine oxidoreductase reveal a unique heme cofactor and a previously unidentified interaction partner, *Biochemistry 52*, 6211-6218.

(47) Pearson, A. R., Elmore, B. O., Yang, C., Ferrara, J. D., Hooper, A. B., and Wilmot, C. M. (2007) The crystal structure of cytochrome P460 of Nitrosomonas europaea reveals a novel cytochrome fold and heme-protein cross-link, *Biochemistry 46*, 8340-8349.

(48) Hooper, A. B., and Nason, A. (1965) Characterization of hydroxylamine-cytochrome c reductase from the chemoautotrophs Nitrosomonas europaea and Nitrosocystis oceanus, *J. Biol. Chem. 240*, 4044-4057.

(49) Rees, M. K. (1968) Studies of the hydroxylamine metabolism of Nitrosomonas europaea. I. Purification of hydroxylamine oxidase, *Biochemistry 7*, 353-366.

(50) Hooper, A. B., and Terry, K. R. (1979) Hydroxylamine oxidoreductase of Nitrosomonas. Production of nitric oxide from hydroxylamine, *Biochim. Biophys. Acta. 571*, 12-20.

(51) Erickson, R. H., and Hooper, A. B. (1972) Preliminary characterization of a variant co-binding heme protein from Nitrosomonas, *Biochim. Biophys. Acta. 275*, 231-244.

(52) Zahn, J. A., Duncan, C., and DiSpirito, A. A. (1994) Oxidation of hydroxylamine by cytochrome P-460 of the obligate methylotroph Methylococcus capsulatus Bath, *J. Bacteriol. 176*, 5879-5887.

(53) Caranto, J. D., Vilbert, A. C., and Lancaster, K. M. (2016) Nitrosomonas europaea cytochrome P460 is a direct link between nitrification and nitrous oxide emission, *Proc Natl Acad Sci USA 113*, 14704-14709.

(54) Enemark, J. H., and Feltham, R. D. (1974) Principles of Structure, Bonding, and Reactivity for Metal Nitrosyl Complexes, *Coord. Chem. Rev. 13*, 339-406.

(55) Vilbert, A. C., Caranto, J. D., and Lancaster, K. M. (2018) Influences of the heme-lysine crosslink in cytochrome P460 over redox catalysis and nitric oxide sensitivity, *Chem. Sci. 9*, 368-379.

(56) Caranto, J. D., and Lancaster, K. M. (2017) Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc Natl Acad Sci USA 114*, 8217-8222.

(57) Andersson, K. K., and Hooper, A. B. (1983) O2 and H2O are each the source of one O in NO− 2 produced from NH3 by Nitrosomonas: 15N-NMR evidence, *FEBS Lett. 164*, 236-240.

(58) Hollocher, T., Tate, M., and Nicholas, D. (1981) Oxidation of ammonia by Nitrosomonas europaea. Definite 18O-tracer evidence that hydroxylamine formation involves a monooxygenase, *J. Biol. Chem. 256*, 10834-10836.

(59) Maalcke, W. J., Dietl, A., Marritt, S. J., Butt, J. N., Jetten, M. S. M., Keltjens, J. T., Barends, T. R. M., and Kartal, B. (2014) Structural Basis of Biological NO Generation by Octaheme Oxidoreductases, *J. Biol. Chem. 289*, 1228-1242.

(60) Hendrich, M. P., Upadhyay, A. K., Riga, J., Arciero, D. M., and Hooper, A. B. (2002) Spectroscopic characterization of the NO adduct of hydroxylamine oxidoreductase, *Biochemistry 41*, 4603-4611.

(61) Elmore, B. O., Bergmann, D. J., Klotz, M. G., and Hooper, A. B. (2007) Cytochromes P460 and c'-beta; a new family of high-spin cytochromes c, *FEBS Lett. 581*, 911-916.

(62) Bergmann, D. J., Zahn, J. A., Hooper, A. B., and DiSpirito, A. A. (1998) Cytochrome P460 genes from the methanotroph Methylococcus capsulatus bath, *J. Bacteriol*. *180*, 6440-6445.

(63) Wijma, H. J., Canters, G. W., de Vries, S., and Verbeet, M. P. (2004) Bidirectional catalysis by copper-containing nitrite reductase, *Biochemistry 43*, 10467-10474.

(64) Arciero, D. M., Pierce, B. S., Hendrich, M. P., and Hooper, A. B. (2002) Nitrosocyanin, a red cupredoxin-like protein from Nitrosomonas europaea, *Biochemistry 41*, 1703-1709.

(65) Lieberman, R. L., Arciero, D. M., Hooper, A. B., and Rosenzweig, A. C. (2001) Crystal structure of a novel red copper protein from Nitrosomonas europaea, *Biochemistry 40*, 5674-5681.

(66) Basumallick, L., Sarangi, R., DeBeer George, S., Elmore, B., Hooper, A. B., Hedman, B., Hodgson, K. O., and Solomon, E. I. (2005) Spectroscopic and density functional studies of the red copper site in nitrosocyanin: role of the protein in determining active site geometric and electronic structure, *J. Am. Chem. Soc. 127*, 3531-3544.

(67) Kartal, B., Wessels, H. J., van der Biezen, E., Francoijs, K. J., Jetten, M. S., Klotz, M. G., and Stein, L. Y. (2012) Effects of nitrogen dioxide and anoxia on global gene and protein expression in long-term continuous cultures of Nitrosomonas eutropha C91, *Appl. Environ. Microbiol. 78*, 4788-4794.

(68) Stein, L. Y., Campbell, M. A., and Klotz, M. G. (2013) Energy-mediated vs. ammonium-regulated gene expression in the obligate ammonia-oxidizing bacterium, Nitrosococcus oceani, *Front. Microbiol. 4*, 277.

(69) Schmidt, I., Steenbakkers, P. J., op den Camp, H. J., Schmidt, K., and Jetten, M. S. (2004) Physiologic and proteomic evidence for a role of nitric oxide in biofilm formation by Nitrosomonas europaea and other ammonia oxidizers, *J. Bacteriol. 186*, 2781-2788.

(70) Klotz, M. G., and Stein, L. Y. (2008) Nitrifier genomics and evolution of the nitrogen cycle, *FEMS Microbiol. Lett. 278*, 146-156.

(71) Vanspanning, R. J. M., Deboer, A. P. N., Reijnders, W. N. M., Spiro, S., Westerhoff, H. V., Stouthamer, A. H., and Vanderoost, J. (1995) Nitrite and

Nitric-Oxide Reduction in Paracoccus-Denitrificans Is under the Control of Nnr, a Regulatory Protein That Belongs to the Fnr Family of Transcriptional Activators, *FEBS Lett. 360*, 151-154.

(72) Cruz-Ramos, H., Crack, J., Wu, G. G., Hughes, M. N., Scott, C., Thomson, A. J., Green, J., and Poole, R. K. (2002) NO sensing by FNR: regulation of the Escherichia coli NO-detoxifying flavohaemoglobin, Hmp, *Embo J. 21*, 3235-3244.

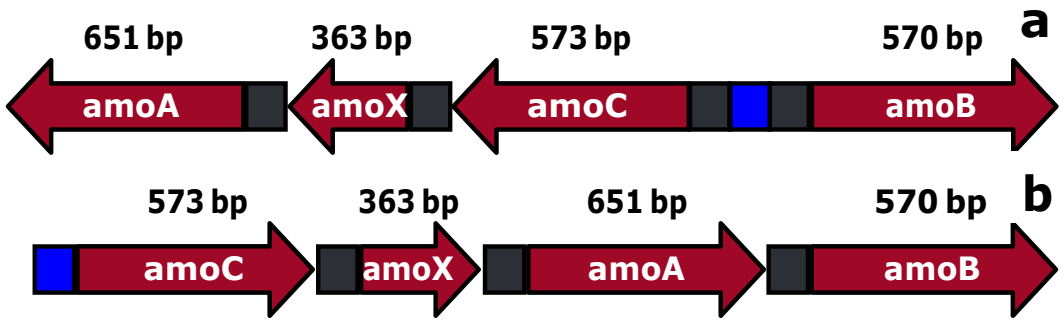
CHAPTER 2: RECOMBINANT EXPRESSION OF ACTIVE ARCHAEAL AMMONIA MONOOXYGENASE VARIANTS IN HETEROTROPHIC *MYCOBACTERIUM SMEGMATIS*

As discussed in Chapter 1: Introduction, AMO is a ubiquitous, powerful, and promiscuous enzyme. AMO effects the hydroxylation of NH_3 to NH_2OH using an active oxidant derived from activation of O_2.[1-3] This potent oxidant is necessarily capable of breaking the 107 kcal/mol N–H bond. Despite the prevalence of AMO in nature and broad interest in identifying the active intermediate, the fundamental biochemistry of AMO remains largely unknown.

Given the difficulty of cultivating ammonia oxidizing bacteria and archaea and, consequently, the challenge of isolating active AMO from these organisms, a recombinant expression system for AMO is desirable. We adapted the strategy of Coleman and co-workers,[4] allowing us to express an archaeal AMO (aAMO) gene cluster in rapidly-growing *Mycobacterium smegmatis* (doubling time 3 to 4 hrs[5] vs. 26 hrs for aAMO *Nitrosopumilus maritimus*[6]) through use of the pMycoFos vector. We rearranged the native *Nitrosopumilus maritimus* aAMO gene cluster (**Figure 2.1a**) such that the four putative subunits *amoA*, *amoB*, *amoC*, and *amoX* share a single orientation downstream of the pMycoFos promoter (**Figure 2.1b**).

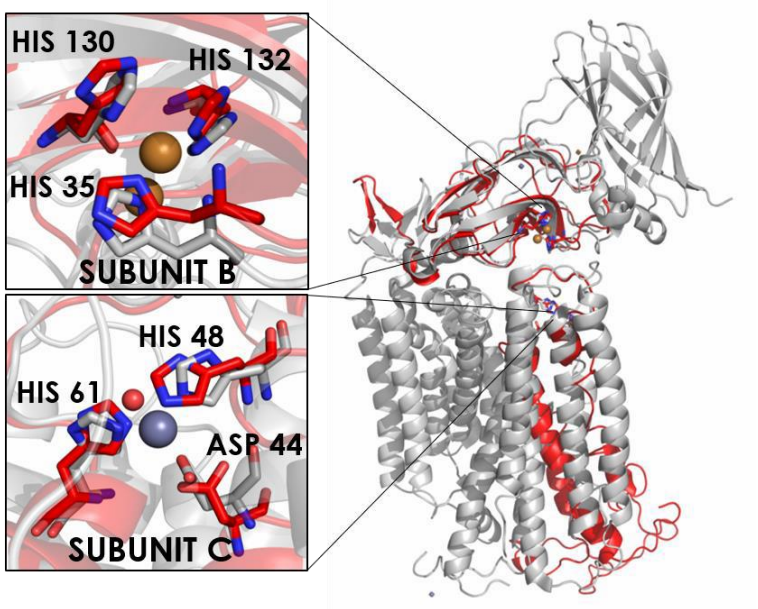
The assembly of a recombinant expression system also affords the unprecedented ability to carry out site-directed mutagenesis studies on aAMO. Taking inspiration from Lawton and Rosenzweig, we carried out structural homology modeling studies to identify putative active sites in *N. maritimus* aAMO.[7] The closest structural homologue we identified was the *Methylococcus capsulatus* (Bath) pMMO, with PDBID 3RGB. Our modeling efforts preserve two important metal binding sites in the amoB and amoC subunits (**Figure 2.2**). We systematically mutated likely metal-coordinating residues in both amoB and amoC subunits to explore the role of each

Figure 2.1: AMO Gene Cluster



(a) Native *N. maritimus* aAMO gene cluster. (b) *M. smegmatis* optimized *N. maritimus* aAMO gene cluster. Non-coding regions were not codon optimized, although Shine-Dalgarno sequences have been changed to those for *M. smegmatis*.

Figure 2.2: Structural Homology Model of aAMO



Structural homology model of *N. maritimus* amoB and amoC (red) aligned with *M. capsulatus* (Bath) pMMO (grey, PDBID: 3RGB). Insets show predicted metal binding sites in both subunits, with putative metal-coordinating residues labeled. Nitrogen atoms are dark blue, oxygen atoms are red, copper ions are represented as orange spheres, zinc ion is represented as a grey sphere, water molecule is represented as a red sphere.

conserved metal-binding site.

A. MATERIALS AND METHODS

General Considerations

Unless otherwise stated, "buffer" is defined as 50 mM phosphate, 0.05% TWEEN-80, pH 7.3. Protein concentration was determined using the detergent compatible Lowry assay[8] (Bio-Rad).

Plasmids and Bacterial Strains

The pMycoFos fosmid (Addgene plasmid # 84577)[4] was generously provided by Dr. Nicholas Coleman. The native *N. maritimus* aAMO gene cluster contains four putative genes with the open reading frames of the amoC, amoX, and amoA genes reading in one direction and that of the amoB gene adjacent to amoC but reading in the opposite direction (**Figure 2.1**). The cluster was codon optimized for protein expression in *M. smegmatis*, synthesized, and cloned into pMycoFos by GenScript. The resulting parent expression vector will thus be subsequently referred to as pMycoFos-CXAB. The gene cluster was synthesized so that the open reading frames were read in the same direction downstream from the acetamidase promoter of pMycoFos. Single amino acid substitutions were introduced by site-directed mutagenesis using the primers shown in Table 2.1. All genetic manipulations were performed on the aAMO gene cluster inserted into the cloning vector pUC57. The variant gene cluster was then amplified using the primers ClonEZ EcoRI CXAB and ClonEZ BamHI CXAB. PCR reagents were removed using a PCR cleanup kit (Qiagen) and the vector was inserted into pMycoFos double-digested with EcoRI and

BamHI using the ClonEZ PCR kit (Genscript). All mutations were verified by sequencing.

Chemically-competent *E. coli* EPI300 (Epicentre) cells were used as a cloning strain to construct pMycoFos-CXAB genetic variants. *E. coli* EPI300 cultures were grown in lysogeny broth (LB) medium with 50 µg /mL kanamycin at 37°C with shaking at 200 rpm. The pMycoFos fosmid was induced to high copy number in EPI300 according to Epicentre protocols. *Mycobacterium smegmatis* strain mc^{2}155 (ATCC) was rendered electrocompetent by washing with cold 10% glycerol using a protocol described in the literature.[9] Electrocompetent *M. smegmatis* were electroporated in a 1 mm gap width cuvette at a resistance of 1000

Table 2.1 – Site-directed mutagenesis and pMycoFos cloning primers
Red indicates mutagenesis sites; bolded text homologous to pMycoFos vector.

Primer Name	Sequence
B-H35A-F	CCCGACCGCCGACGCGGCTGGCGTCCAGGCCCA
B-H35A-R	CTGGGCCTGGACGCCAGCCGCGTCGGCGGTCGGG
B-H130A-F	GAGGCCGGCGTCTACGCTGTGCACACCCAGCTG
B-H130A-R	CAGCTGGGTGTGCACAGCGTAGACGCCGGCCTC
B-H132A-F	GGCGTCTACCACGTGGCTACCCAGCTGAACGTC
B-H132A-R	GACGTTCAGCTGGGTAGCCACGTGGTAGACGCC
B-H130A_H132A-F	GCCGGCGTCTACGCTGTGGCTACCCAGCTG
B-H130A_H132A-R	GGTAGCCACAGCGTAGACGCCGGCCTCGAG
C-D44A-F	GTGGACAACTTCGTCGCTGGCTCGCTGCACCAG
C-D44A-R	CTGGTGCAGCGAGCCAGCGACGAAGTTGTCCAC
C-H48A-F	GACGGCTCGCTGGCTCAGACCAGCATCCGC
C-H48A-R	GCGGATGCTGGTCTGAGCCAGCGAGCCGTC
C-H61A-F	CCTTCACGCCGGCGGCTTGGTGGCTGTACTC
C-H61A-R	GAGTACAGCCACCAAGCCGCCGGCGTGAAGG
EcoRI CXAB	**CATCTGTAAGAATTC**ATGATCACCATGGCCCAGATGCCC
BamHI CXAB	**TTGAGACACGGATCC**TCAGATGACCTGCCAGGGGCG

Ω, capacitance of 25 µF, and voltage at 1.25 kV using a Bio-Rad Gene Pulser II. Electroporated cells were immediately immersed in 1 mL of ice cold 7H9 Middlebrook/ADC medium (HiMedia Laboratories) and incubated on ice for 10 minutes. The 1 mL of recovered cells was then pipetted into another 1 mL of 7H9 Middlebrook/ADC medium and incubated for 3 hours at 37°C with shaking at 200 rpm. Cells were pelleted, resuspended in 100 µL of Middlebrook 7H9/ADC medium, and plated on Middlebrook 7H10/OADC agar plates (HiMedia Laboratories) containing 50 µg/mL kanamycin. Plates were incubated at 37°C for 5-7 days resulting in waxy, raised white colonies. Single colonies were picked and used to inoculate Middlebrook 7H9/ADC medium at 30°C with shaking at 125 rpm. Cells were grown for 3-5 days to an OD_{600} of 0.6-1.0. These cells were used to make 25% glycerol stocks and stored at -80°C. Induced transformant cultures were grown in 1600-mL Middlebrook 7H9/ADC cultures containing 50 µg/mL kanamycin and 0.2% acetamide in a 4L flask. These cultures were inoculated with a 1.6-mL glycerol stock and incubated at 30°C with shaking at 125 rpm for 4-5 days to an OD_{600} of 0.6-2.0. Cells were then incubated at 4°C with no shaking overnight and collected for membrane fractions and/or assayed for O_2 consumption activity as described below. Induced 50-mL *M. smegmatis* transformant cultures were grown under similar conditions except they were prepared in 125-mL Erlenmeyer flasks and inoculated with 500 µL glycerol stocks.

O_2-Consumption Assay

Cells were pelleted by centrifuging at 4000xg for 10 minutes. The cells were then washed by 2x resuspension/centrifugation cycles with 50 mM phosphate, 0.05%

tween-80, pH 7.3. The cells were resuspended in the same buffer and stored on ice until assayed. A 1-mL aliquot of the washed cells was pipetted into the water-jacketed assay vessel of the Oxygraph Plus (Hansatech) set at 30°C and containing 1 mL of the assay buffer and 18 µL of 20% glucose (final concentration of 10 mM glucose). The O_2 concentration was monitored for 5-10 minutes, at which point NH_4Cl was added to the cell sample to a final concentration of 10 mM. The O_2 concentration was monitored for an additional 5-10 minutes. The slope of O_2 concentration vs. time was determined before and after addition of NH_4Cl by linear least-squares regression. The percent change in O_2-consumption rate was calculated by **Eqn. 2.1**:

$$(rate_f - rate_i)/rate_i \times 100\% \qquad (\mathbf{2.1})$$

where $rate_i$ and $rate_f$ are the measured O_2 consumption rates before and after addition of NH_4Cl, respectively.

Preparation of Membrane Fractions for SDS-PAGE Analysis

Cells were cultured in 4L Erlenmeyer flasks as described above and harvested by centrifugation at 4000 × g in a Beckman-Coulter centrifuge with a JLA-16.25 rotor. Harvested cells were washed 2x with 250 mL of buffer to remove media components and finally resuspended in 20 mL of the same buffer. The resuspended cells were sonicated for 3 minutes at 50% power and a 50% pulse rate in a Branson Sonifier 250 (VWR). This was followed by sonicating the cells for 5 minutes at 75% power and a 50% pulse rate. The sonicated suspension was centrifuged at 4000 x g for 10 minutes at 4 °C and the supernatant collected. The membrane fractions were separated from the soluble fraction by ultracentrifugation at 120,000 x g and 4 °C for

1 hr. This resulted in a reddish, translucent pellet. The supernatant was discarded, and the pellet was resuspended in 50 mM MOPS, 250 mM NaCl, pH 7.3. The pellet was then centrifuged again. Two more wash cycles were performed, and the pellet was finally resuspended in 1 mL of 50 mM MOPS, 250 mM NaCl, pH 7.3 using a Dounce Homogenizer. The cell fractions were then aliquoted into 1.5-mL Eppendorf tubes and stored at -80°C. Protein was quantified using the detergent-compatible Lowry assay (Bio-Rad). 20 µg of protein from membrane fractions were mixed with 2x Laemmli buffer and run on SDS-PAGE gels. The gels were then stained with silver stain (Bio-Rad).

Allylthiourea Inhibition Assays

Cells were pelleted by centrifuging at 4000 x g for 10 minutes. The cells were then washed by 2x resuspension/centrifugation cycles with 50 mM phosphate with 0.05% TWEEN-80 (assay buffer), pH 7.3. For pre-incubation with allylthiourea (ATU), cells were resuspended in the same buffer with ATU to a final concentration of 2.5 mM and stored on ice for 30 minutes prior to assay. Cells were assayed as described above (O_2-consumption assay). To test the effect of ATU on cells after the addition of NH_4Cl, 2.5 mM ATU was added to the assay vessel approximately 45-60 seconds after NH_4Cl addition. The O_2 concentration was monitored for an additional 1-5 minutes, and the percent change in O_2-consumption rate was calculated as described above.

Colony PCR

Incorporation of pMycoFos-CXAB by transformed cells was verified using colony PCR to amplify aAMO genes.1 µL of cell culture was added to a standard

PCR mixture containing Taq DNA polymerase, dNTPS, $MgCl_2$, and reaction buffer (GoTaq Green, Promega). ClonEZ EcoRI CXAB and ClonEZ BamHI CXAB primers were used to amplify the "CXAB insert" containing the aAMO genes. PCR products were run on a 0.8% agarose gel containing ethidium bromide; representative images of each variant can be visualized below (**Figure 2.3**).

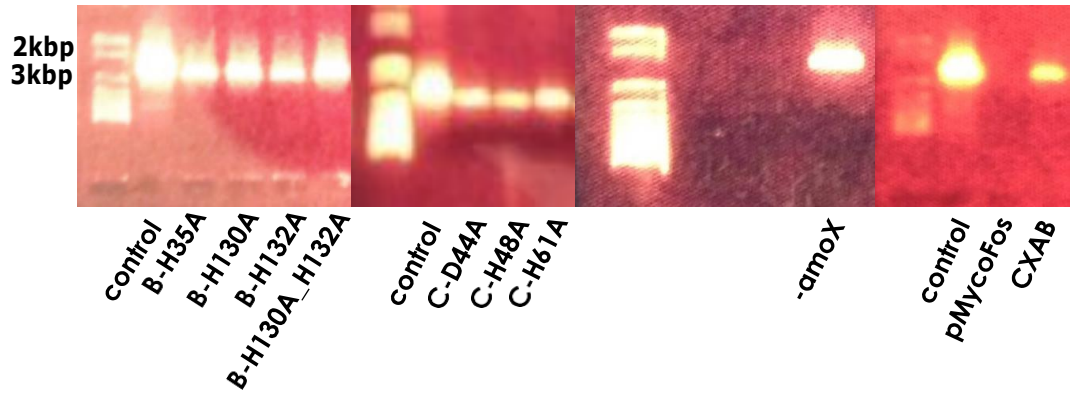
Reverse Transcription (RT)-PCR

Cultures of *M. smegmatis* transformed with pMycoFos-CXAB or empty pMycoFos vector were grown as described above in either the presence or absence of 0.2% acetamide. RNA was extracted from harvested cells using a RNeasy kit (Qiagen). Superscript II reverse transcriptase was used following the manufacturer's protocol (Life Technologies) to generate cDNA from the RNA samples. The cDNA collected was then used as template DNA for PCR following the same procedure as the colony PCR to amplify the CXAB insert. PCR products were run on a 0.8% agarose gel containing ethidium bromide.

B. RESULTS

Colony PCR was used to verify correct transformation of *M. smegmatis* cells. PCR products were visualized via agarose gel electrophoresis (**Figure 2.3**). Reverse-transcription PCR (RT-PCR) verified complete transcription of the synthetic aAMO gene cluster following acetamide induction of *M. smegmatis* transformed with pMycoFos-CXAB (**Figure 2.4a**). Initial SDS-PAGE analysis suggested the presence of new protein bands not present in cells transformed with just the pMycoFos vector (**Figure 2.4b**); however, these results could not reliably be repeated. SDS-PAGE gels of membrane fractions were run for each variant, and once again, consistent results

Figure 2.3: *M. smegmatis* Colony PCR



Agarose gel images depicting representative results of colony PCR on grown *M. smegmatis* cultures transformed with the indicated vector. PCR amplified the CXAB insert (2656 bp). Control was performed on pMycoFos-CXAB vector stock.

were elusive. This is likely due to improper solubilization of the membrane fractions. The use of various detergents was attempted to mitigate these issues, but unfortunately never resulted in reproducible results.

We used respirometry as a probe for aAMO activity. Rates of O_2 consumption by quantified *M. smegmatis* suspensions grown, pelleted, washed, and resuspended in 50mM sodium phosphate, 0.05% Tween-80, pH 7.3 to an OD_{600} ~1.0 were compared before and after addition of NH_4Cl to final concentrations of 10 mM. *M. smegmatis* transformed with "empty" pMycoFos exhibits no response to NH_4Cl. *M. smegmatis* transformed with pMycoFos-CXAB show a 44 ± 15% increase in the rate of O_2 consumption after 15 trials (including 5 separate *M. smegmatis* colonies, **Figure 2.5**). We found this response to be robust across multiple colonies, each representing independent transformations.

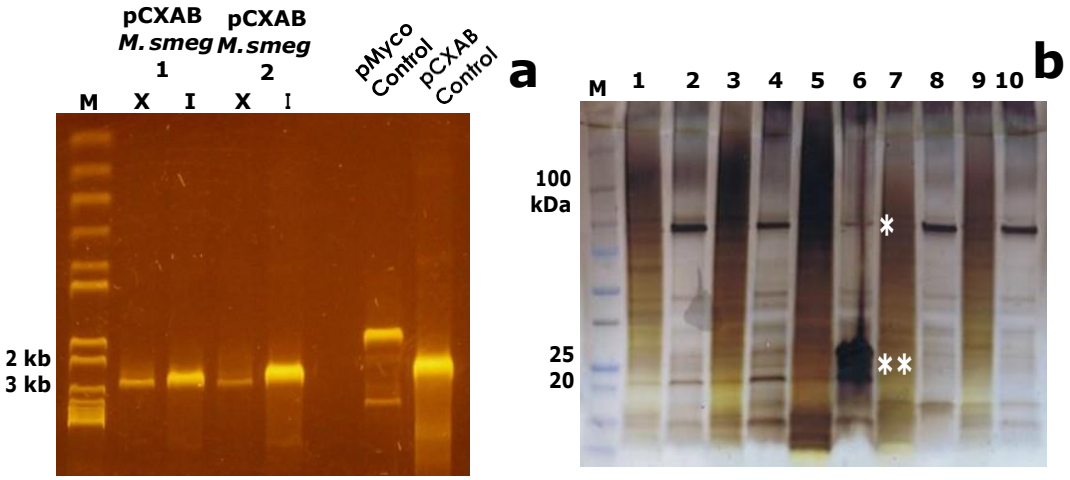
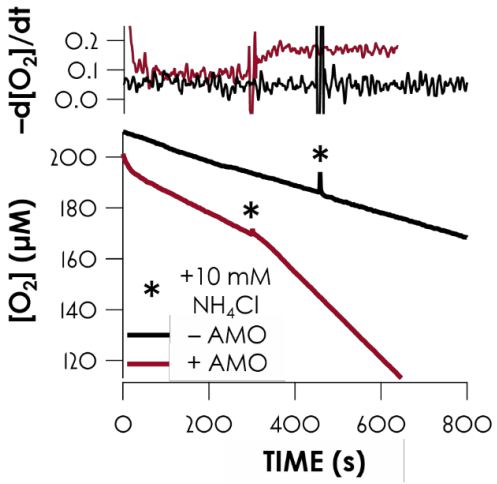


Figure 2.4: Detection of AMO in Cells

(a) RT-PCR results confirming expression of the aAMO gene cluster in acetamide induced (I; vs. uninduced, X) CXAB-transformed *M. smegmatis*. (b) SDS-PAGE gel showing pMycofos- (odd numbered lanes) or CXAB-transformed (even numbered lanes) *M. smegmatis* membrane fractions. Lanes 1/2 untreated; 3/4 treated with Brij-35; 5/6 treated with Tween-20; 7/8 treated with n-octyl-D-glucopyranoside; 9/10 treated with CHAPS. * Indicates a ~100kDa band present in CXAB-transformed membrane fractions not found in pMycoFos-transformed fractions. Treatment with Tween-20 resulted in the appearance of bands between 20-30kDa (**).

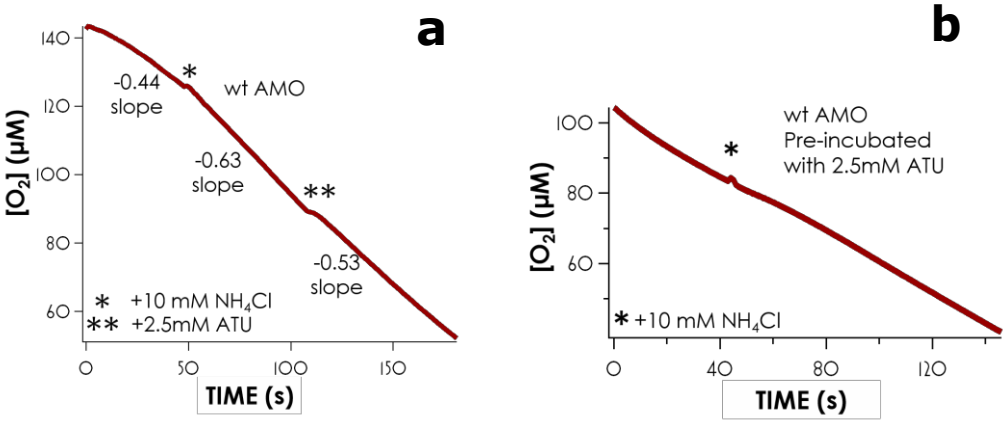
NH_4Cl-induced changes in O_2 consumption are totally abolished when pMycoFos-CXAB transformed *M. smegmatis* is incubated with the known AMO inhibitor allylthiourea (ATU)[2, 10, 11] prior to respirometry. The NH_4Cl induced increase in O_2 consumption of untreated *M. smegmatis* can also be partially arrested by addition of ATU (**Figure 2.6**). Specifically, following the addition of 10 mM NH_4Cl, which induces the ~44% rate increase in pMycoFos-CXAB *M. smegmatis* cells, addition of ATU immediately reduced the rate increase to ~20% (compared to the rate before NH_4Cl addition). ATU is likely an inhibitor of AMO because it can act as a Cu chelator. Thus, we suspect the reason cells with ATU added after NH_4Cl still show some

Figure 2.5: O_2 Consumption by *M. smegmatis*



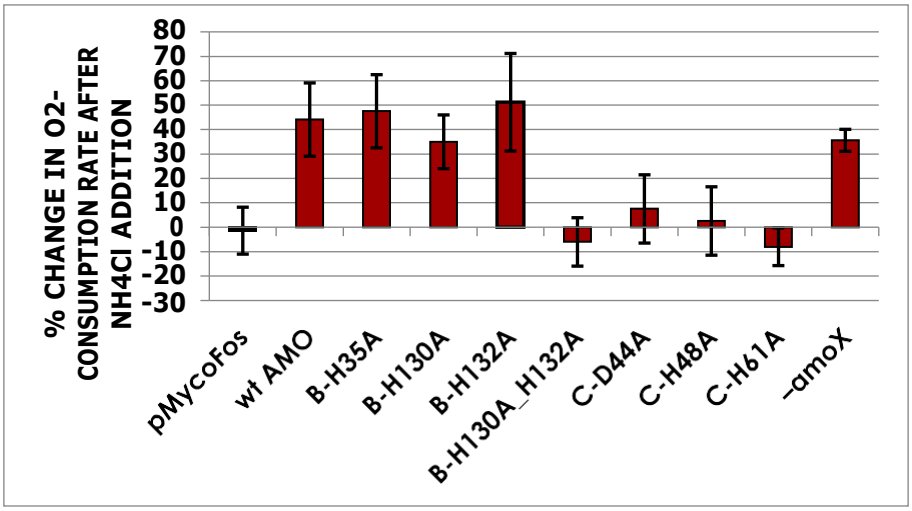
O_2 consumption by *M. smegmatis* expressing aAMO increases by ca. 40-50% after buffer supplementation with 10 mM NH_4Cl. *M. smegmatis* transformed with empty pMycoFos shows no change in O_2 consumption after the addition of NH_4Cl. (*) indicates the time of NH_4Cl addition.

Figure 2.6. Effect of Allylthiourea on O_2-Consumption



O_2 consumption by *M. smegmatis* expressing aAMO either treated with ATU after being treated with NH_4Cl (a), or pre-incubated with 2.5mM ATU before being supplied NH_4Cl (b).

Figure 2.7. NH_4Cl-Induced Changes in O_2 Consumption by *M. smegmatis*



Values include 3-5 trials per colony. Error bars indicate one standard deviation.

increased O_2 consumption is because the ATU is unable to sequester all Cu from the AMO metal-binding sites within the time of the assay. ATU has no effect on the O_2 consumption rate of *M. smegmatis* transformed with empty pMycoFos.

We systematically mutated putative metal-coordinating residues in both the amoB and amoC subunits (**Figure 2.2**) to alanine. Colony PCR for the CXAB insert was used to verify transformation of the AMO genes. Plasmids were also isolated from *M. smegmatis* cells and sequenced to confirm the presence of the correct mutation. These variants were then assayed for NH_4Cl-dependent changes in O_2 consumption, and the results of these studies are aggregated in **Figure 2.7**. Variants with single amoB mutations showed no changes in behavior relative to wild-type pMycoFos-CXAB.

However, amoC mutations D44A, H48A, and H61A and the amoB H130A_H132A double mutant abolish NH_4Cl-induced increases in O_2 consumption rate.

We also generated a variant construct that lacks the amoX hypothetical gene. A correctly transformed *M. smegmatis* culture (verified by colony PCR, **Figure 2.3**) exhibited an increase in O_2-consumption upon addition of NH_4Cl (~35%, **Figure 2.7**). This suggested that the hypothetical gene is not involved in AMO-mediated O_2 activation.

C. DISCUSSION

Nitrification is driven by both bacteria and archaea, though mounting evidence suggests archaea are the dominant nitrifiers,[12-15] making this an essential enzyme to understand. Though, even among more characterized CuMMOs the active site has remained intensely controversial both in terms of location (e.g. subunit B or C metal-binding site) and metal occupancy. Rosenzweig et al. have shown enzymatic activity to be dependent upon copper,[16] but in addition to their proposed dicopper model, both trinuclear[17] and mononuclear[18, 19] models have also been suggested. In a recent report using quantum refinement to improve crystallographic resolution of pMMO (PDB ID: 3RGB), the authors conclude that only a monocopper site is possible in the B subunit, however.[20] In fact, a recent structure of pMMO from *Methylomicrobium alcaliphilum* 20Z published in 2018 contains just one Cu at the pmoB site.[21] The C subunit metal occupancy still holds some ambiguity. Depending on crystallization conditions, the C subunit site has been shown to contain either Zn or Cu.[22, 23] It was determined that the presence of the Zn in this site was due to Zn from the crystallization buffer. Interestingly, it has also been shown that Zn can inhibit pMMO, possibly due to its

ability to replace Cu in the subunit C site.[23] Adding to the current ambiguity, the most recent pMMO structure from *Mm. alcaliphilum* 20Z contains a disordered pmoC site with no clear metal bound.[21]

Rosenzweig et al. have shown that a truncated, soluble pMMO subunit B is able to hydroxylate methane to form methanol,[24] suggesting the subunit B site is where substrate hydroxylation occurs. Likewise, various Cu-containing synthetic complexes similar to this site are able to oxidize methane.[25, 26] Coleman et al., on the other hand, showed through mutagenesis that residues in the C subunit are more important to the activity of related CuMMO hydrocarbon monooxygenase (HMO).[27] The ability of Zn to inhibit pMMO by binding to the pmoC site also shows this site plays some crucial role. Metal-coordinating residues at both sites are highly conserved among CuMMOs, and so both sites can be viewed as potential candidates for the active site.

Our preliminary mutagenesis work suggests that both AMO metal-binding sites are likely relevant to NH_4^+-dependent O_2-activation. At present it is unclear if O_2 activation occurs at both the amoB and amoC sites, or if these results are due to other factors. For instance, it is possible that single mutations to the amoB site are not enough to disrupt Cu binding at this site, hence the discrepancy between the amoB single and double mutants. Although, the observed effect of the double mutant could also be related to a greater degree of imposed structural changes that affect reactivity but are not related directly to the integrity of the amoB Cu-binding site. It would be beneficial to have more structural information for the aAMO and each of these mutants, but as there is still no ideal method to visualize or purify the enzyme, this will require much more future work.

In the same vein, though it appears that the amoC site may play a role in AMO activity, it is possible that mutations to the subunit C residues affect a structural aspect of the enzyme—such as an O_2 or substrate channel used to access the amoB site—that is not directly related to the mechanistic action of AMO. This would still result in no/reduced substrate-driven increase in O_2-consumption, even if the true active site is the amoB site. This could also help explain the location of a highly conserved pocket of residues at the interface of the three subunits near the subunit C site, as well as a specific residue (A56 in *N. maritimus* AMO) that seems to convey substrate preference.[27] This may also explain why the C subunit site appears to have many variations, and may in fact be more flexible.

Another possibility is that both sites serve a distinct purpose in overall activity and both contribute some role either in substrate binding, O_2-activation, or electron/proton delivery. Because both sites are so highly conserved among CuMMOs, it would follow that both sites serve some essential function. Given that a soluble pmoB subunit is able to hydroxylate methane (though to the greatly reduced degree), it seems most likely that this site is truly where substrate hydroxylation occurs, but that the C subunit site is essential for another component, either related to O_2-activation, electron/proton delivery, or in some structural manner. Of course, more information will need to be gathered in order to determine the exact role of each metal-binding site, as well as to confirm that the engineered mutations do not affect other factors, such as relative levels of protein expression or proper folding.

Importantly, it should be noted that we were unable to obtain consistent SDS-PAGE gel analysis of *M. smegmatis* membrane fractions. This result leaves the

possibility that instead of any changes pertinent to the AMO mechanism, any changes in the observed O_2-consumption may be due to differences in protein expression or improper folding of the protein in the cells due to the imposed mutation. Though it is unlikely that a single mutation could result in altered protein expression that is consistent across multiple mutations (e.g. for each individual amoC mutation), and from colony PCR it is clear that no changes in vector concentration are observed, without appropriate protein detection we cannot discount this possibility and make any assertions as to the nature of the active site. Regardless, we are still confident that with the observed new ability for *M. smegmatis* cells to increase O_2-consumption in response to NH_4Cl that we have expressed an active AMO.

Our results also show that the hypothetical amoX gene likely does not participate directly in O_2 activation. It is still possible that this gene contributes to the overall structure or stability of the enzyme *in vivo*, though it does not seem essential for activity in this recombinant system. Again, without proper visualization of the protein from cells, we cannot discount the possibility that the lower percent change in rate observed for these variants compared to the wildtype AMO is not due to a decrease in protein expression for this variant. Regardless, we do still observe some activity and thus can conclude that amoX is not essential for activity but may still have some function related to AMO activity.

D. CONCLUSION

We have recombinantly expressed active AMO in *M. smegmatis*. Incorporation of the aAMO gene cluster sensitizes transformed *M. smegmatis* to NH_4Cl, manifesting as an increase in the rate of O_2-consumption. Though further work is needed to confirm

the validity of our mutagenesis studies, our preliminary work suggests that two active sites need to be intact for proper function of AMO, and that the amoB site is robust against single amino acid residue mutations.

This work shows that recombinant expression of an active aAMO is possible; however, either better purification protocols for aAMO expressed in *M. smegmatis* or a more robust expression system with an optimum cell membrane structure for AMO expression are needed for future studies to explore the nature of the AMO metal-binding sites.

E. REFERENCES

(1) Arp, D. J., Sayavedra-Soto, L. A., and Hommes, N. G. (2002) Molecular biology and biochemistry of ammonia oxidation by Nitrosomonas europaea, *Arch. Microbiol. 178*, 250-255.

(2) Vajrala, N., Martens-Habbena, W., Sayavedra-Soto, L. A., Schauer, A., Bottomley, P. J., Stahl, D. A., and Arp, D. J. (2013) Hydroxylamine as an intermediate in ammonia oxidation by globally abundant marine archaea, *Proc Natl Acad Sci USA 110*, 1006-1011.

(3) Crossman, L. C., Moir, J. W. B., Enticknap, J. J., Richardson, D. J., and Spiro, S. (1997) Heterologous expression of heterotrophic nitrification genes, *Microbiology 143*, 3775-3783.

(4) Ly, M. A., Liew, E. F., Le, N. B., and Coleman, N. V. (2011) Construction and evaluation of pMycoFos, a fosmid shuttle vector for Mycobacterium spp. with inducible gene expression and copy number control, *J. Microbiol. Methods 86*, 320-326.

(5) Klann, A. G., Belanger, A. E., Abanes-De Mello, A., Lee, J. Y., and Hatfull, G. F. (1998) Characterization of the dnaG locus in Mycobacterium smegmatis reveals linkage of DNA replication and cell division, *J. Bacteriol. 180*, 65-72.

(6) Martens-Habbena, W., Berube, P. M., Urakawa, H., de la Torre, J. R., and Stahl, D. A. (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria, *Nature 461*, 976-979.

(7) Arnold, K., Bordoli, L., Kopp, J., and Schwede, T. (2006) The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling, *Bioinformatics 22*, 195-201.

(8) Lowry, O. H., Rosebrough, N. J., Farr, A. L., and Randall, R. J. (1951) Protein measurement with the Folin phenol reagent, *J. Biol. Chem. 193*, 265-275.

(9) Au - Goude, R., and Au - Parish, T. (2008) Electroporation of Mycobacteria, *JoVE*, e761.

(10) Juliette, L. Y., Hyman, M. R., and Arp, D. J. (1993) Inhibition of Ammonia Oxidation in Nitrosomonas europaea by Sulfur Compounds : Thioethers Are Oxidized to Sulfoxides by Ammonia Monooxygenase *Appl. Environ. Microbiol. 59*, 3718-3727.

(11) Hatzenpichler, R. (2012) Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea, *Appl. Environ. Microbiol. 78*, 7501-7510.

(12) Stahl, D. A., and de la Torre, J. R. (2012) Physiology and diversity of ammonia-oxidizing archaea, *Annu. Rev. Microbiol. 66*, 83-101.

(13) Leininger, S., Urich, T., Schloter, M., Schwark, L., Qi, J., Nicol, G. W., Prosser, J. I., Schuster, S. C., and Schleper, C. (2006) Archaea predominate among ammonia-oxidizing prokaryotes in soils, *Nature 442*, 806-809.

(14) Sims, A., Horton, J., Gajaraj, S., McIntosh, S., Miles, R. J., Mueller, R., Reed, R., and Hu, Z. (2012) Temporal and spatial distributions of ammonia-oxidizing archaea and bacteria and their ratio as an indicator of oligotrophic conditions in natural wetlands, *Water Res. 46*, 4121-4129.

(15) Wuchter, C., Abbas, B., Coolen, M. J., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G. J., Middelburg, J. J., Schouten, S., and Sinninghe Damste, J. S. (2006) Archaeal nitrification in the ocean, *Proc Natl Acad Sci USA 103*, 12317-12322.

(16) Balasubramanian, R., Smith, S. M., Rawat, S., Yatsunyk, L. A., Stemmler, T. L., and Rosenzweig, A. C. (2010) Oxidation of methane by a biological dicopper centre, *Nature 465*, 115-119.

(17) Chan, S. I., and Yu, S. S.-F. (2008) Controlled oxidation of hydrocarbons by the membrane-bound methane monooxygenase: The case for a tricopper cluster, *Acc. Chem. Res. 41*, 969-979.

(18) Lee, J. Y., and Karlin, K. D. (2015) Elaboration of copper-oxygen mediated C-H activation chemistry in consideration of future fuel and feedstock generation, *Curr. Opin. Chem. Biol. 25*, 184-193.

(19) Quinlan, R. J., Sweeney, M. D., Lo Leggio, L., Otten, H., Poulsen, J. C., Johansen, K. S., Krogh, K. B., Jorgensen, C. I., Tovborg, M., Anthonsen, A., Tryfona, T., Walter, C. P., Dupree, P., Xu, F., Davies, G. J., and Walton, P.

H. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components, *Proc Natl Acad Sci USA 108*, 15079-15084.

(20) Cao, L. L., Caldararu, O., Rosenzweig, A. C., and Ryde, U. (2018) Quantum Refinement Does Not Support Dinuclear Copper Sites in Crystal Structures of Particulate Methane Monooxygenase, *Angew. Chem. Int. Ed. Engl. 57*, 162-166.

(21) Ro, S. Y., Ross, M. O., Deng, Y. W., Batelu, S., Lawton, T. J., Hurley, J. D., Stemmler, T. L., Hoffman, B. M., and Rosenzweig, A. C. (2018) From micelles to bicelles: Effect of the membrane on particulate methane monooxygenase activity, *J. Biol. Chem. 293*, 10457-10465.

(22) Balasubramanian, R., and Rosenzweig, A. C. (2007) Structural and mechanistic insights into methane oxidation by particulate methane monooxygenase, *Acc. Chem. Res. 40*, 573-580.

(23) Sirajuddin, S., Barupala, D., Helling, S., Marcus, K., Stemmler, T. L., and Rosenzweig, A. C. (2014) Effects of zinc on particulate methane monooxygenase activity and structure, *J. Biol. Chem. 289*, 21782-21794.

(24) Amin, S. A., Moffett, J. W., Martens-Habbena, W., Jacquot, J. E., Han, Y., Devol, A., Ingalls, A. E., Stahl, D. A., and Armbrust, E. V. (2013) Copper requirements of the ammonia-oxidizing archaeon Nitrosopumilus maritimus SCM1 and implications for nitrification in the marine environment, *Limnol. Oceanogr. 58*, 2037-2045.

(25) Himes, R. A., and Karlin, K. D. (2009) Copper-dioxygen complex mediated C-H bond oxygenation: relevance for particulate methane monooxygenase (pMMO), *Curr. Opin. Chem. Biol. 13*, 119-131.

(26) Woertink, J. S., Smeets, P. J., Groothaert, M. H., Vance, M. A., Sels, B. F., Schoonheydt, R. A., and Solomon, E. I. (2009) A [Cu2O]2+ core in Cu-ZSM-5, the active site in the oxidation of methane to methanol, *Proc Natl Acad Sci USA 106*, 18908-18913.

(27) Liew, E. F., Tong, D., Coleman, N. V., and Holmes, A. J. (2014) Mutagenesis of the hydrocarbon monooxygenase indicates a metal centre in subunit-C, and not subunit-B, is essential for copper-containing membrane monooxygenase activity, *Microbiology 160*, 1267-1277.

CHAPTER 3: THE EPONYMOUS COFACTORS IN CYTOCHROME P460S FROM AMMONIA-OXIDIZING BACTERIA ARE IRON PORPHYRINOIDS WHOSE MACROCYCLES ARE DIBASIC

Reprinted (adapted) with permission from:

"The Eponymous Cofactors in Cytochrome P460s from Ammonia-Oxidizing Bacteria Are Iron Porphyrinoids Whose Macrocycles Are Dibasic"

Meghan A. Smith and Kyle M. Lancaster

Biochemistry, **2018**, *57*, 334-343

Cytochrome (cyt) P460s are small, ca. 17-23 kDa proteins found in diverse classes of organisms, including ammonia-oxidizing bacteria (AOB), methanotrophs, proteobacteria, planctomycetes, acidobacteria, and bacteroidetes.[1-2] Proposed to act in the detoxification of hydroxylamine (NH_2OH)[3] and/or nitric oxide (NO),[4] cyt P460 from *Nitrosomonas europaea* has been shown to be a biological source of the potent greenhouse gas nitrous oxide (N_2O).[3] Cyt P460 contains a unique porphyrinoid cofactor named "heme P460" due to the characteristic Soret maximum of the Fe^{II} form near 460 nm.

Heme P460 cofactors are unusual porphyrinoids found in both cyt P460 and hydroxylamine oxidoreductase (HAO), despite these two proteins bearing no sequence or structural homology.[5] In addition to two covalent Cys cross-links characteristic of *c*-type hemes, hemes P460 feature additional cross-links to amino acid side chains. The active site of HAO contains a Tyr phenolate side chain which forms two covalent attachments to the heme at the 2′ and 3′ carbons (**Figure 1.4a**. The *N. europaea* cyt P460 cofactor bears a cross-link between Lys Nε and the 13′ *meso* carbon (**Figure 1.4b**).[4, 6-9] The formation of this cross-link is expected to be autocatalytic, as recombinant expression of cyt P460 in both *E. coli* and *P. aeruginosa* yield the correct Lys-heme attachment.[6] Mutations to this Lys in *N. europaea* cyt P460 result in heme proteins with spectroscopic properties more typical of cyt *c* (i.e. a sharp Soret band with a maximum around 400 nm) and abolish the NH_2OH oxidase activity.[6,10] Thus, the cross-link defines the unusual spectroscopic properties and reactivity of hemes P460.

Prior studies have argued that the heme P460 cofactor in HAO should be

assigned as a tribasic phlorin (**Figure 3.1**).[8] However, the assignment of HAO's active site cofactor as comprising an electron-rich, tribasic macrocycle was thought to be at odds with the role of HAO in oxidative chemistry. It is interesting to note that though the crosslink has been shown to be essential for NH_2OH oxidation activity in *N. europaea* cyt P460, the identity and placement of the cross-link seem unimportant as these attributes are different in cyt P460 and HAO. Both Tyr and Lys have similarly alkaline side-chain pK_a values, but it is unlikely that this relates to the ability of both enzymes to oxidize NH_2OH. However, cyt P460 and HAO seem to differ in the binding of NO to the ferric heme. The cyt P460 mechanism has been shown to include an $\{FeNO\}^6$ (following Enemark-Feltham notation[11]) intermediate,[3] which is also proposed by computational models of HAO catalysis.[12] However, whereas the cyt P460 $\{FeNO\}^6$ persists indefinitely in the absence of O_2 or NH_2OH, an HAO $\{FeNO\}^6$ is believed to be short-lived.[13] As NO was recently shown to be the enzymatic product of HAO,[14] facile NO dissociation accords with the function of the enzyme. Since both HAO and cyt P460 from *N. europaea* are competent for NH_2OH dehydrogenation, only the presence but not the nature of the cross-link appears to be mandatory. However, the nature of the cross-link may influence Fe-NO binding affinities.

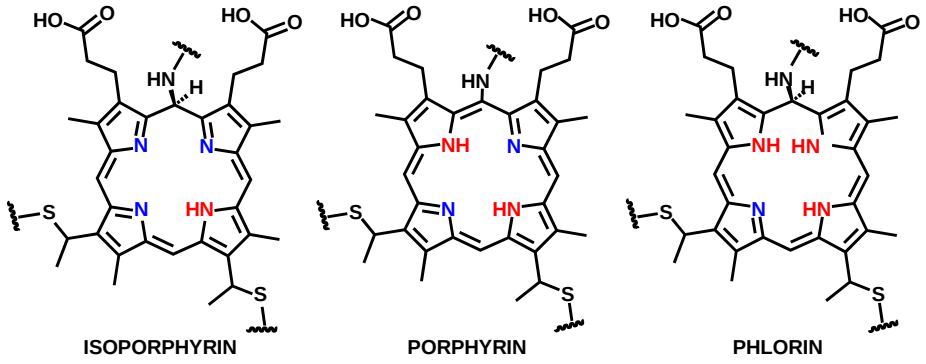
In a recent 2.1 Å crystal structure of HAO,[15] it was concluded that the heme carbons participating in the crosslink (C2´ and C3´) are in fact sp^3 hybridized, as evidenced by the elongation of the C—C bond distances of the heme. Although this structure remains consistent with a dibasic macrocycle, the aromaticity is disrupted. Consequently, the cofactor was described as a tetrapyrrole and not a true porphyrin. In

a 1.8 Å crystal structure of *N. europaea* cyt P460, the authors also assign the *meso* C participating in the crosslink as sp^3 hybridized by analogy.[4] If a single macrocycle C is truly sp^3 hybridized, this necessitates either a tribasic phlorin or monobasic isoporphyrin configuration (**Figure 3.1**) for the cyt P460 cofactor, and thus is distinct from the dibasic HAO heme P460 macrocycle. This is at odds with the similarity of spectroscopic properties between the heme P460 cofactors of these two proteins.

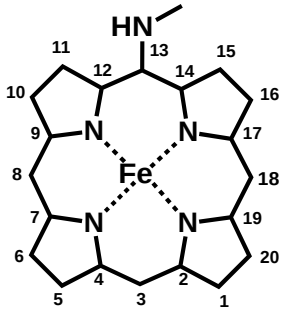
Herein we report a 1.45 Å X-ray crystal structure of cyt P460 from *Nitrosomonas* sp. AL212. This structure, with higher resolution and greater

---

Figure 3.1 A single sp^2 meso C in a porphyrin backbone can give rise to either isoporphyrin or phlorin electronic configurations.



Scheme 3.1. Atom numbering used for the *Nitrosomonas* sp. AL212 cyt P460 cofactor.



---

completeness, advances the structural understanding of cyt P460s by providing a model of the polypeptide loop enclosing the cofactor binding site. Our structural analysis, coupled to UV/vis absorption spectroscopy and time-dependent density functional theory (TDDFT) indicate that the *meso* C involved in the cross-link remains sp^2 hybridized. Distortions in heme planarity and decreased symmetry, enforced by the crosslink, appear sufficient to red-shift the Soret absorption maximum to values characteristic of P460 centers.

A. MATERIALS AND METHODS

General Considerations

Milli-Q water (18.2 MΩ; Millipore) was used in the preparation of all buffers and solutions. UV- visible (UV-vis) absorption spectra were obtained using a Cary 60 UV-vis spectrometer. Data were fit using Igor Pro version 6.37 (WaveMetrics). For the generation of the $\{FeNO\}^6$ species, the NO-donor disodium 1-(Hydroxyl-NNO-azoxy)-L-proline (PROLI-NONOate) was used, whereas for the generation of the $\{FeNO\}^7$ species, the HNO-donor disodium diazen-1-ium-1,2,2 triolate ($Na_2N_2O_3$, Angeli's salt) was used; both were purchased form Cayman Chemicals. All other chemicals were purchased from VWR International.

Protein Overexpression and Purification

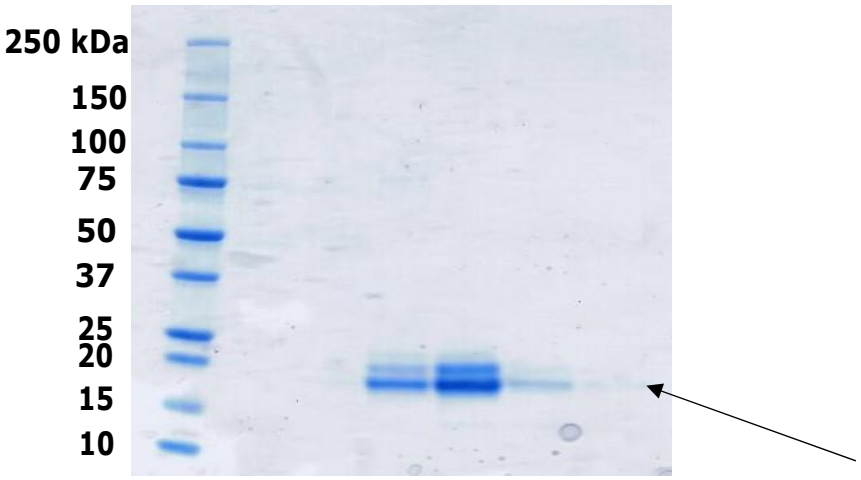
A codon-optimized gene for cyt P460 from *Nitrosomonas* sp. AL212 was synthesized and cloned into the *NcoI* and *XhoI* sites of pET-22b(+) vector by GenScript, Inc. The vector was designed to include a C-terminal His-tag found in the parent pET-22b(+) plasmid. This plasmid was co-transformed into *E. coli* BL21(DE3) cells along with a pEC86 plasmid containing the cyt *c* maturation genes

ccmABCDEFGH[16] (provided by H. B. Gray) and plated on LB agar supplemented with 100 μg mL^{-1} each ampicillin and 37 μg mL^{-1} chloramphenicol. Protein expression and purification followed the protocol for *N. europaea* cyt P460 expression and purification described previously.[3] Briefly, cells were grown on a 5 L scale in Terrific Broth containing 0.5% glycerol (0.1% inoculum), 100 μg mL^{-1} ampicillin, and 37 μg mL^{-1} chloramphenicol. Cultures were grown at 30 °C for 18 hrs before protein expression was induced by adding isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.4 mM. Cells were harvested 6 to 8 hours post-induction and lysed via sonication and centrifuged for 45 min at 35,000 × g relative centrifugal force. The supernatant was applied to a HisTrap HP Ni affinity column (GE Lifesciences), equilibrated with 20 mM Tris (pH 8.0) containing 150 mM NaCl. Proteins were eluted on a gradient over 10 column volumes to 20 mM Tris (pH 8.0) containing 150 mM NaCl and 330 mM imidazole. Green-colored fractions were combined, concentrated and further purified to homogeneity using a HiLoad Superdex 75 size-exclusion column equilibrated with a running buffer of 50 mM 3-(*N*-morpholino)propanesulfonic acid (pH 7.2) containing 150 mM NaCl.

Protein purity was assessed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE, **Figure 3.2**), as well as UV-visible absorption spectroscopy. A typical protein purification yielded Reinheitszahl values (Rz), known as the purity ratio of A440 nm/A280 nm, of 1.3. In all purifications, two bands were observed for cyt P460, as has been seen before in *N. europaea*.[1] One band matches the predicted protein mass (~21 kDa), whereas the other (~19 kDa) matches the expected mass of the protein minus the initial 27 residues. This region is possibly a cleavable

periplasmic tag sequence; thus, the two bands likely represent cleaved and non-cleaved versions of cyt P460.

Figure 3.2: SDS-PAGE Gel of Fractions Following Size-Exclusion Chromatography



EPR Spectroscopy

X-band (9.40-GHz) EPR spectra were collected on a sample containing 170 µM *Nitrosomonas* sp. AL212 cyt P460 in 200 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 8.0 with 25% (v/v) glycerol. The measurements were obtained using a Bruker Elexsys-II spectrometer equipped with a liquid He cryostat maintained at 10.0 K. EPR data were simulated using SpinCount.[17]

Crystallization and Data Collection

Purified Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 was screened for crystallization conditions using the PACT Suite (Qiagen). Crystals suitable for diffraction were obtained using the sitting drop method. 3 µL of 650 µM *Nitrosomonas* sp. AL212 cyt P460 in 50 mM 3-(*N*-morpholino)propanesulfonic acid

(MOPS), pH 8.0 were mixed with 1–4 µL of a crystallization buffer containing 0.2 M sodium chloride, 0.1 M sodium acetate pH 5.0, and 17–22% (w/v) PEG 6000 and allowed to equilibrate above 250 µL of this buffer. Crystals grew after 2-3 days at room temperature. To avoid oxidative damage, crystals used in diffraction experiments were grown, looped, and frozen in an anaerobic Coy chamber. Crystals were soaked in the aforementioned crystallization solution containing 30% ethylene glycol before being frozen in liquid N_2 for storage and transport.

X-ray diffraction experiments were conducted at beamline NE-CAT 24-ID-C of the Advanced Photon Source (APS). Crystals were irradiated at 100 K using X-rays with a wavelength (λ) of 0.979 Å. An initial model was generated in Phenix using the molecular replacement method and the cyt P460 structure from *N. europaea* (PDB ID 2JE2). The autosol function was then applied in Phenix[18] to produce an initial working *Nitrosomonas* sp. AL212 cyt P460 model. Refinements and building to completion were then conducted using Phenix and Coot,[19] respectively. PyMol[20] was used for figure-generation.

Calculations

Electronic structure calculations were performed using version 3.03 of the ORCA quantum chemical computation suite.[21] Calculations were carried out on cofactors including amino acid sidechains serving as inner-sphere ligands to Fe where applicable. Amino acid residues in the second- and outer-sphere of the cofactor binding pocket were excluded. Calculations employed crystallographic coordinates whose H-atom positions were optimized using the BP86 density functional[22, 23] and the scalar relativistically recontracted Ahlrich's def2-TZVP(-f)-ZORA functional on

all atoms.[24] These calculations included the zeroth-order regular approximation (ZORA)[25] for relativistic corrections, and solvation was modeled using COSMO[26] with a dielectric of 36.6 (acetonitrile). TDDFT calculations using the Tamm-Dancoff approximation[27] were initiated from single point calculations that employed the B3LYP hybrid density functional,[28, 29] the CP(PPP) basis set on Fe,[30] def2-TZVP(-f)-ZORA basis set on all other atoms, ZORA, and COSMO with a dielectric of 36.6 (acetonitrile). The RIJCOSX procedure was used to accelerate hybrid DFT calculations.[31, 32] For energy diagrams, quasi-restricted orbitals (QROs)[33] were generated from unrestricted Kohn-Sham spinorbitals. Orbital images were generated using Chimera.[34]

The heme out-of-plane distortions were analyzed using an online version of the normal-mode structural decomposition script developed following the normal-coordinate structural decomposition procedure originally outlined by Walter Jentzen and John Shelnutt.[35-36]

B. RESULTS

Spectroscopic Characteristics

Nitrosomonas sp. AL212 cyt P460 expressed in *E. coli* exhibits UV/vis absorption spectroscopic properties similar to *N. europaea*.[3] Resting, oxidized *Nitrosomonas* sp. AL212 cyt P460 features a broad Soret absorption band with λ_{max} = 440 nm (ε = 41 $mM^{-1}\,cm^{-1}$) and Q-bands at 576 nm and 628 nm (ε = 3.5 $mM^{-1}\,cm^{-1}$ and 3.9 $mM^{-1}\,cm^{-1}$, respectively) (**Figure 3.3**, **Figure 3.4a**). An $\{FeNO\}^6$ can be generated by treating 5 µM Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 with 50 µM NO-donor PROLI-NONOate (**Figure 3.4a**). Similarly, an $\{FeNO\}^7$ species can be formed

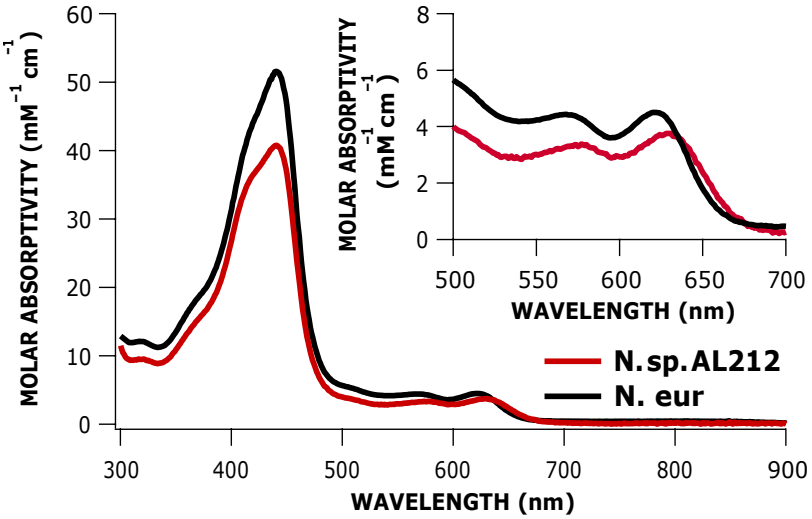
by treatment of 5 µM Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 with 50 µM HNO-donor $Na_2N_2O_3$ (Angeli's salt) (**Figure 3.4a**). The continuous wave X-band (9.40 GHz) EPR spectrum of resting oxidized *Nitrosomonas* sp. AL212 cyt P460 is characteristic of a slightly rhombic high spin ($S = {}^5/_2$) Fe^{III} system, with g_{eff}-values of 6.39, 5.13, 1.97, and an ${}^E/_D$ value of 0.03 (**Figure 3.4b**). These features closely match those of the *N. europaea* Fe^{III} cyt P460, whose g_{eff}-values are 6.57, 5.09, and 1.97 with an ${}^E/_D$ of 0.03.[3]

Crystal Structure

Diffraction quality, green, rod-like crystals of *Nitrosomonas* sp. AL212 Fe^{III} cyt P460 were grown via the sitting drop method in an anaerobic Coy chamber. These crystals tolerated soaking with cryoprotectant solution containing 30% ethylene

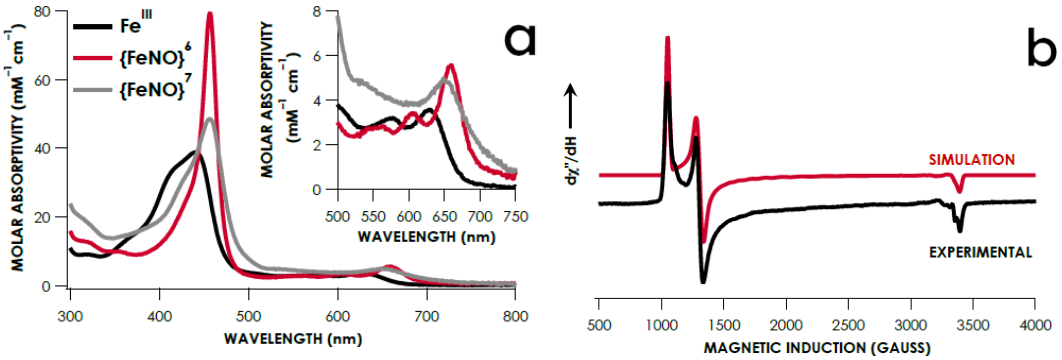
---

Figure 3.3 Comparison of UV-vis Absorption Profiles of *N. europaea* and *Nitrosomonas* sp. AL212 Fe^{III} cyt P460



UV-vis absorption spectra of resting Fe^{III} *N. europaea* (black) and *Nitrosomonas* sp. AL212 (red) cyt P460.

Figure 3.4 Spectroscopic characterization of various *Nitrosomonas* sp. AL212 cyt P460 species



(a) UV-vis absorption spectra of resting Fe^{III} (black), $\{FeNO\}^6$ (red), and $\{FeNO\}^7$ (gray) forms of *Nitrosomonas* sp. AL212 cyt P460. (b) 10 K X-band (9.40 GHz) EPR spectrum of Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 (black) recorded at 633 μW microwave power. The corresponding SpinCount simulation is shown in red.

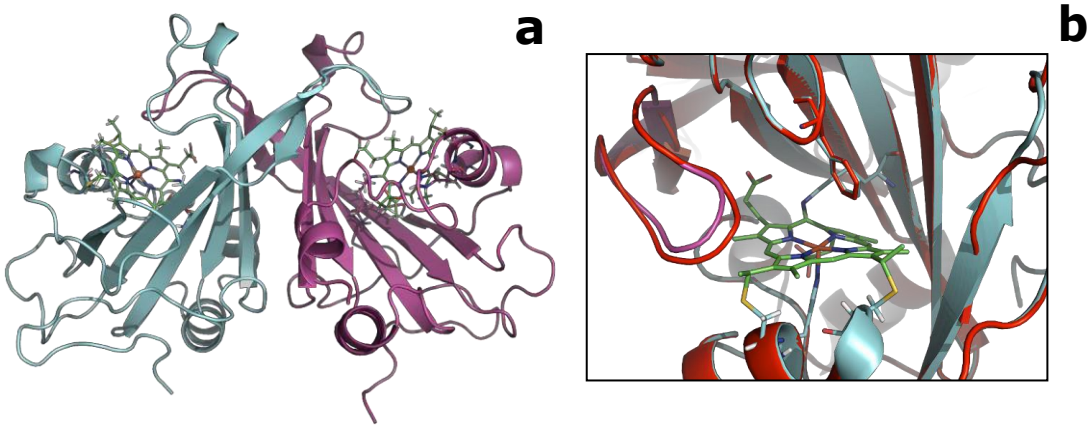
glycol. X-ray diffraction data were obtained to high, 1.45 Å resolution using highly attenuated 0.979 Å X-rays at beamline 24-IDC at APS.

The structure of the resting Fe^{III} form of *Nitrosomonas* sp. AL212 cyt P460 was solved via molecular replacement using the structure of *N. europaea* cyt P460 (PDB ID 2JE2) as a model, and the electron density was anisotropically refined[38] to a resolution of 1.45 Å. Our *Nitrosomonas* sp. AL212 cyt P460 structure includes residues 37–194 for chain A (80.6% complete) and 35–194 for chain B (81.6% completeness). As discussed, the absence of at least the first 27 residues may be attributable to cleavage of a periplasmic localization tag as suggested by SignalP.[39] Previous structures of cyt P460 from *N. europaea* include breaks in the electron density, resulting in missing loops around the heme P460 active site.[4] Our structure has no breaks in electron density, and thus affords a model of the full distal heme

pocket as well as a loop of residues 113–122 that wraps around from one monomer toward the heme site of another, partially blocking this site and exposing possible functional significance to the dimeric structure of the enzyme (**Figure 3.5, 3.6**).

The *N. europaea* structure features a phosphate ion coordinated to Fe *trans* to His140,[4] while HAO structures feature H_2O bound in this position.[15] By contrast, the *Nitrosomonas* sp. AL212 Fe is 5-coordinate. While we cannot directly rule out the possibility that photoreduction of cyt P460 occurred during data collection; however, no photoreduction was observed for the *N. europaea* cyt P460 crystal structure, which shows an almost identical heme structure.[4] Moreover, prolonged exposure to high-flux X-rays in Fe K-edge XAS experiments does not result in any spectral changes that would be consistent with Fe reduction.[10] We considered the possibility that Phe76 in the *Nitrosomonas* sp. AL212 structure disfavors axial ligand coordination (**Figure**

Figure 3.5. Crystal Structure of *Nitrosomonas* sp. AL212 Fe^{III} Cyt P460



(a) Full crystal structure of *Nitrosomonas* sp. AL212 Fe^{III} cyt P460 (PDB ID 6AMG). Subunit A shown in cyan, subunit B shown in purple. The heme P460 cofactors are shown in green. (b) Homology model of *N. europea* cyt P460 (red) with *Nitrosomonas* sp. AL212 cyt P460 (PDB ID: 6AMG; subunit A is cyan, subunit B is purple, P460 cofactor is green).

Table 3.1. X-ray Data Collection and Crystal Structure Refinement Statistics.

	Fe^{III} cyt P460
Wavelength (Å)	0.979
Temperature (K)	100
Space Group	P 2_1 2_1 2_1
a (Å)	47.3
b (Å)	80.1
c (Å)	109.0
α (deg)	90
β (deg)	90
γ (deg)	90
Reflections	475,694 (12,166)
Number of Reflections in R_{work} Set	66,399
Number of Reflections in R_{free} set	3,441
Resolution (Å)	64.62–1.45
R_{merge} (%)	4.1 (72.7)
$CC_{1/2}$	0.999 (0.748)
Completeness (%)	97.0 (90.4)
Redundancy	6.6 (4.6)
I/σ(I)	23.8 (1.4)
R_{work}	14.1 (25.0)
R_{free}	15.8 (23.4)
RMSD from Ideality	
Bonds (Å)	0.028
Angles (deg)	2.3
Average *B* Factors (Å^2)	29.2
Ramachandran Plot	
Allowed Regions (%)	100
Disallowed Regions (%)	0
PDBID	6AMG

3.5b). Phe76 is part of the loop comprising residues 65–80 that is missing from the *N. europaea* structure. Based on mass spectrometry evidence, the absence of this loop was attributed to oxidative damage to cyt P460 that occurred during crystallization.[4] Thus, it is possible that the *N. europaea* active site is similarly hindered from distal

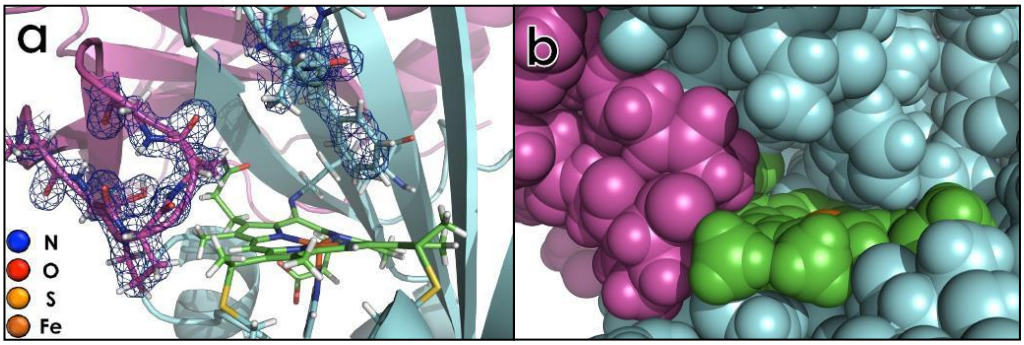
axial ligand coordination when the full loop is present. Structural homology modeling reveals that this loop is similar in both isozymes (**Figure 3.5b**), including in the placement of the "capping" Phe sidechain distal to the heme cofactor. The similarity of g-values in the $S = {}^5/_2$ EPR signals characteristic of resting Fe^{III} cyt P460 in both *N. europaea* and *Nitrosomonas* sp. AL212 and the similar UV-vis absorption profiles (**Figure 3.3**) of both proteins strongly suggests that both proteins are 5-coordinate in their resting Fe^{III} form.

Heme Geometry and Normal-Coordinate Structural Decomposition

Refinements of the Fe^{III} X-ray datasets were carried out with effectively no restraints imposed on cofactor structural parameters. Distances and angles about the Lys N–C13′ cross-link resulting from these refinements are shown for the A subunit in **Figure 3.6**, with complete heme metrical parameters for cofactors in both subunit reproduced in **Table 3.2**. Metrical parameters about the cross-linked 13′ *meso* C are inconsistent with sp^3 hybridization. To further probe this geometry, we imposed restraints on these distances to enforce sp^2–sp^3-like distances ca. 1.5 Å. This leads to the generation of negative difference density in the $F_{obs} - F_{calc}$ omit map clearly visible at 2.5 σ (**Figure 3.7**), further supporting a model with sp^2–sp^2 C–C distances (~1.4 Å).

The macrocycles in *Nitrosomonas* sp. AL212 and *N. europaea* cyt P460, as well as HAO are highly distorted away from planarity, as determined via normal-coordinate structural decomposition (NCSD).[35, 36] It has been shown that the type of heme distortion is generally conserved among proteins with the same functionality.[37,40] The cyt P460 Fe^{III} heme experiences almost equal distortion from saddling (B_{2u}) and ruffling (B_{1u}) (**Table 3.3**).

Figure 3.6. Cofactor Binding Pocket of *Nitrosomonas* sp. AL212 Cyt P460



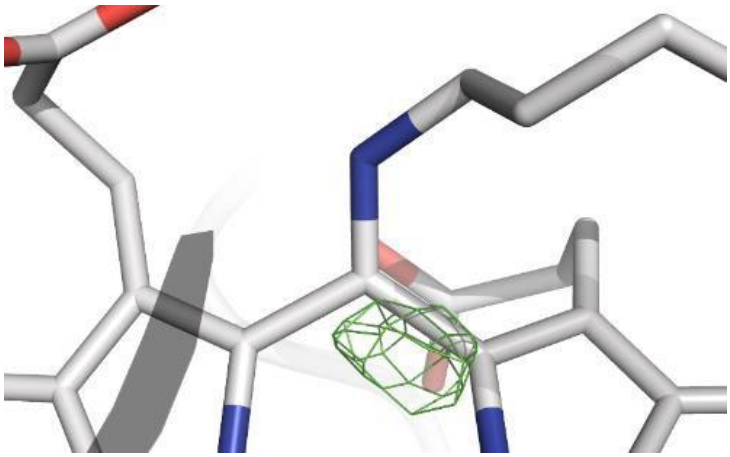
Binding pocket depicted using (a) sticks or (b) space-filling spheres. The $2F_o–F_c$ simulated annealing omit map at a level of 2.5 σ is displayed for the loops that cover the distal side of the heme P460 cofactor. Subunit A shown in cyan, subunit B shown in purple. The heme P460 cofactors are shown in green.

Table 3.2. Geometric parameters obtained from *Nitrosomonas* sp. AL212 cyt P460 crystal structure.[b]

C2–C3	1.34 Å (1.33 Å)
C3–C4	1.30 Å (1.32 Å)
C7–C8	1.35 Å (1.32 Å)
C8–C9	1.35 Å (1.31 Å)
C12–C13	1.39 Å (1.39 Å)
C13–C14	1.41 Å (1.40 Å)
C17–C18	1.39 Å (1.33 Å)
C18–C19	1.30 Å (1.35 Å)
C13–N_{Lys}	1.37 Å (1.37 Å)
∠C2–C3–C4	133.4° (132.8°)
∠C7–C8–C9	130.1° (135.5°)
∠C12–C13–C14	123.0° (121.8°)
∠C17–C18–C19	127.5° (123.6°)
∠C12–C13–N_{Lys}	116.0° (118.1°)
∠C14–C13–N_{Lys}	120.9° (118.9°)
Dihedral ∠$C_{13\ plane}$–N_{Lys}	175.6° (166.7°)

[b]Values outside of parentheses are given for the A subunit, values in parentheses correspond to the B subunit.

Figure 3.7 Calculated Difference Density During Heme Structure Refinement



Difference density (shown in green, $F_{obs} - F_{calc}$ omit map at 2.5 σ) as a result of refinement constrained to a structure more akin to phlorin or isoporphyrin configuration with C–C bonds elongated to 1.5 Å.

Calculated Spectroscopy

TDDFT calculations were used to predict UV-vis absorption spectra corresponding to the heme P460 derived from the Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 crystal structure. Variants with a protonated, "forced" sp^3 13′ *meso* C were generated by geometry optimization of H-atom positions using the BP86 functional and ZORA-def2-TZVP basis set on all atoms. The positions of non-H atoms were taken directly from the subunit A in the crystal structure, then H atoms were added as necessary to produce either a porphyrin, isoporphyrin, or phlorin configuration. Resultant structures were used as coordinates for the subsequent TDDFT calculations, which included sufficient roots to saturate the Soret region of the absorbance profile. The heme structure from horse heart metmyoglobin (PDBID: 1WLA[42]) was also subjected to this procedure to facilitate comparisons. A key feature of the spectra calculated for the three cyt P460 macrocycle variants are absorption profiles that are

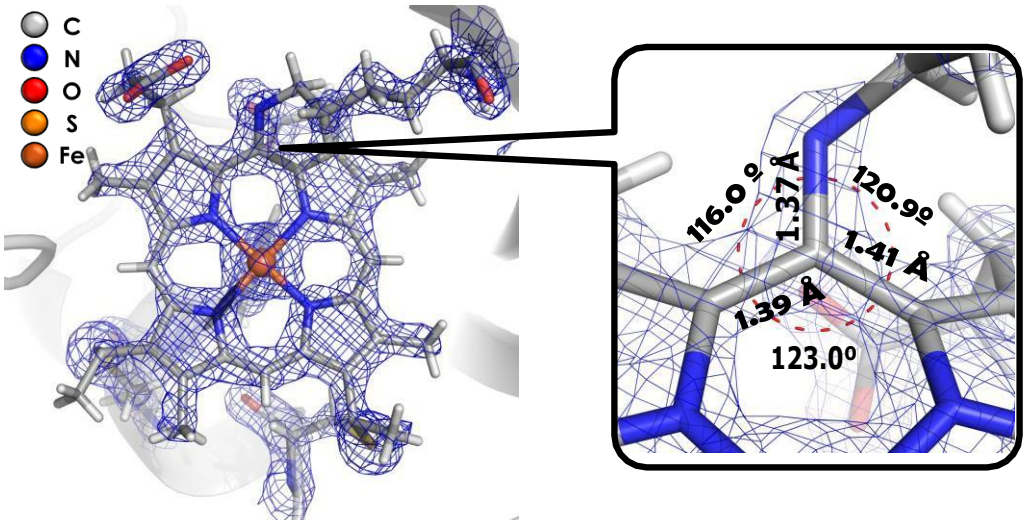
Table 3.3. Normal-Coordinate Structural Decomposition for Representative Heme Cofactors.

	Saddling (b_{2u})	Ruffling (b_{1u})	PDBID	Reference
***Nitrosomonas* sp. AL212 Fe^{III} cyt P460**	0.82	–0.89	6AMG	This work
***N. europea* cyt P460**	0.78	–0.76	2JE3	This work
***N. europaea* HAO**	0.14	2.36	4FAS	This work
Metmyoglobin	0.03	–0.20	1WLA	This work
Horse heart cyt C	–0.23	–1.01	1HRC	This work
Nitrophorin 4—H_2O	0.30	–0.55	1X8Q	Ref. 41
Nitrophorin 4—NO	0.37	–0.81	1X80	Ref. 41

highly red-shifted and diminished in intensity relative to that of metmyoglobin, in accord with the experimental spectra (**Figure 3.9**). These calculations predict an intense feature near 700 nm for the phlorin variant; while appreciable error is expected in the energy values calculated via TDDFT for these systems, the lack of any appreciable absorbance past the Q-band region in experimental Fe^{III} cyt P460 spectra (cf. **Figure 3.3, 3.4**) disfavors the phlorin assignment. We note that making a concrete assignment based on TDDFT is tenuous, nevertheless these data strongly suggest that sp^3 hybridization of the 13′ C is not necessary to give rise to an absorption profile with a red-shifted, weak Soret characteristic of Fe^{III} cyt P460.

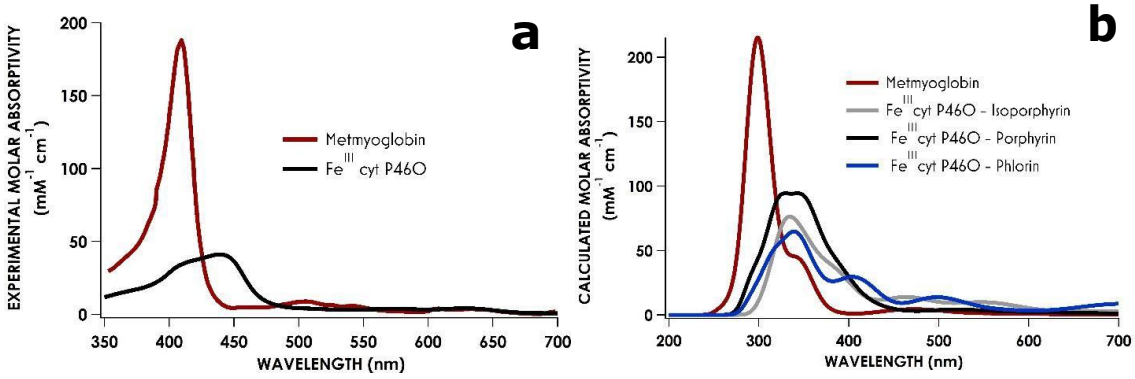
An additional factor that could perturb the heme P460 UV-vis absorption profile relative to canonical *c* type hemes is simply the presence of a *meso* substitution. In particular, π-donation from the Lys N lone pair to the porphyrin ring could profoundly influence the absorption spectrum. The influence of *meso* substitution was explored via calculation of spectra corresponding to two additional

Figure 3.8 $2F_o–F_c$ simulated annealing omit map of the heme P460 cofactor from *Nitrosomonas* sp. AL212 Fe^{III} cyt P460



$2F_o–F_c$ simulated annealing omit map shown at 2.5 σ and corresponding stick representation of the heme P460 cofactor of subunit A from *Nitrosomonas* sp. AL212 Fe^{III} cyt P460. The inset shows the 13′ *meso* C participating in the cross-link with the Lys106 sidechain and corresponding metrical parameters about this atom.

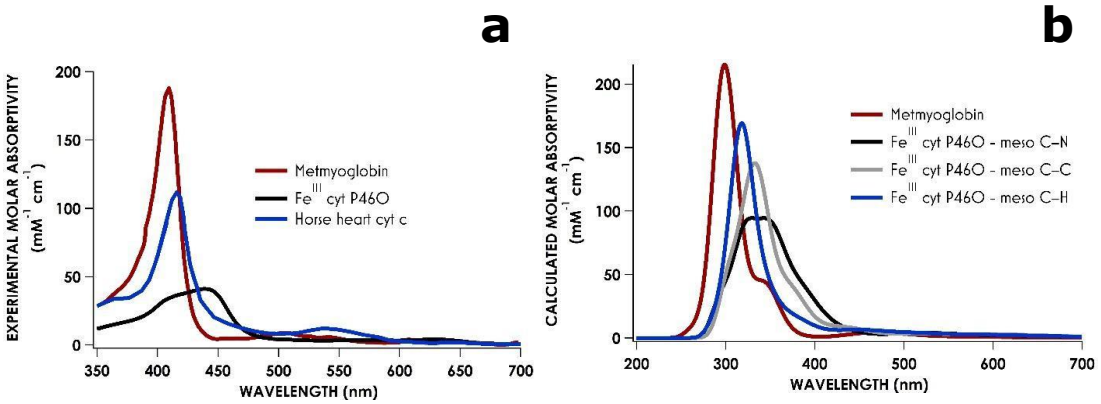
Figure 3.9 Comparing UV-vis spectra for different *Nitrosomonas* sp. AL212 cyt P460 basicity



(a) Experimental and (b) TDDFT-calculated (B3LYP/ZORA-def2-TZVP(-f) with CP(PPP) on Fe) UV-vis absorption spectra of horse heart metmyoglobin and Fe^{III} *Nitrosomonas* sp. AL212 cyt P460. In (b), spectra are calculated for isoporphyrin, porphyrin, and phlorin macrocycle configurations.

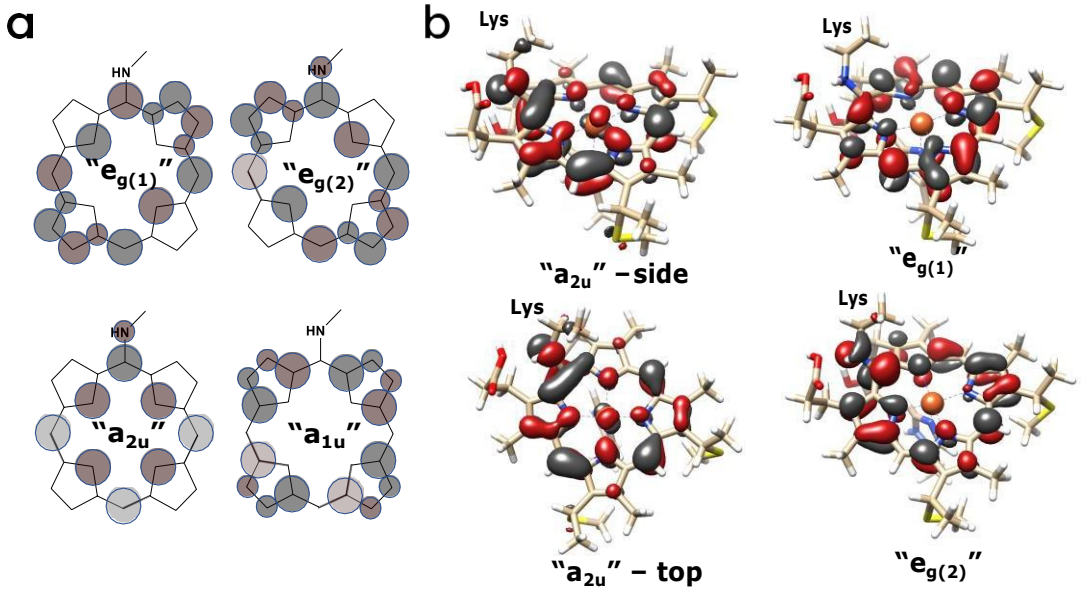
heme P460 variants: one whose cross-linking Lys N (of the 'meso C–N' variant) was replaced with a methylene (data for this structure hereby referred to as 'meso C–C') and a spectrum where the cross-link was removed entirely, leaving a standard *meso* unsubstituted porphyrin ('meso C–H'). These calculated spectra (**Figure 3.10**) show that the *meso* unsubstituted heme P460 variant Soret maximum is red-shifted by 20 nm relative to the calculated Soret maximum for metmyoglobin. Substitution by either a C- or N-based linkage leads to a 35 nm red shift, with the N-cross-linked variant showing the broadest profile and the least intense Soret absorption. Thus, *meso*-substitution—whether *meso* C–C or *meso* C–N bonds—compounds the red-shift imposed by heme distortion from planarity. These electronic structure calculations can be used further to explain the curiously broad Soret absorption profile of the Fe^{III}

Figure 3.10 Comparing UV-vis spectra for *Nitrosomonas* sp. AL212 cyt P460 for different Meso-X atoms



(a) Experimental UV-vis spectra for metmyoglobin (planar, no cross-link), horse heart cyt *c* (ruffled, no cross-link), and *Nitrosomonas* sp. AL212 (ruffled, has cross-link). (b) TDDFT-calculated (B3LYP/ZORA-def2-TZVP(-f) with CP(PPP) on Fe) UV-vis absorption spectra of horse heart metmyoglobin and Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 variants where the 13′ meso C is bound to H, C (from RCH_2), or N (from RNH).

Figure 3.11 Gouterman Four-Orbital Picture for *Nitrosomonas* sp. AL212

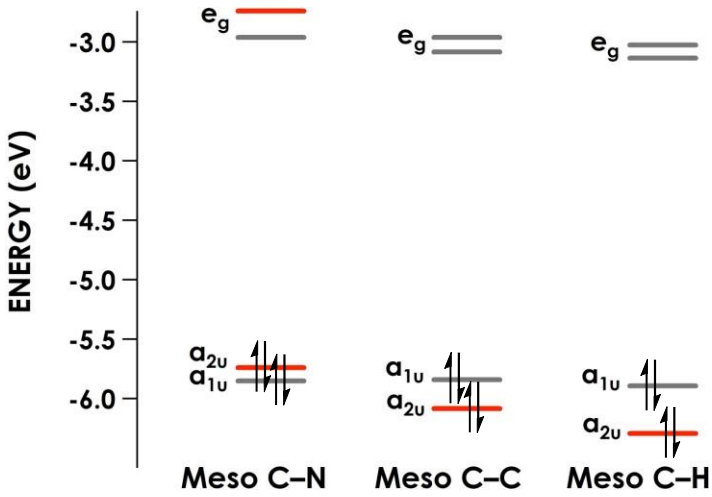


(a) Simplified, top down depiction of the frontier molecular orbitals of a porphyrin π-system modified to show the *meso*-C cross-link and labeled with corresponding D_{4h} irreducible representations. (b) Quasi-restricted orbital (QROs) plots depicting the porphyrin π-system molecular orbitals of the Fe^{III} cyt P460 cofactor with Lys N lone-pair admixture. Calculations were carried out using B3LYP/ZORA-def2-TZVP(-f) with CP(PPP) on Fe, and orbitals are plotted at an isovalue of 0.03 au.

heme P460 cofactor. Following the four-orbital model of porphyrin electronic structure defined by Gouterman and assuming—for convenience—rigorous D_{4h} symmetry (**Figure 3.11**),[43] Lys N lone pair mixing into the porphyrin ring π system is expected to most directly influence the composition and energetics of the a_{2u}-like orbital, causing it to increase in energy, which would result in a red-shifted Soret. This prediction is supported by the electronic structure calculations (**Figure 3.12**). Interestingly, the a_{2u}-like orbital's energy increase raises it to be nearly isoenergetic with the a_{1u}-like orbital; this accidental degeneracy should lead to a sharper Soret

instead of the broad, split Soret that is observed for cyt P460. However, the Lys N lone pair also contributes to one of the e_g orbitals, breaking the degeneracy of the e_g-like orbitals, thus providing an explanation for the observed broadening of the cyt P460 Soret band.

Figure 3.12 Energy Diagram for Meso-X Cross-linked Porphyrin



Simplified energy diagram depicting frontier porphyrin π-system molecular orbitals of the Fe^{III} cyt P460 cofactor as a dibasic porphyrin with Lys cross-link (meso C–N) and dibasic porphyrin variants where the cross-link is modified to a meso C–C linkage or removed entirely (meso C–H). The red lines indicate molecular orbitals whose energies are influenced by orbital admixture from the cross-link.

C. DISCUSSION

The structural parameters for the heme macrocycles in Fe^{III} *Nitrosomonas* sp. AL212 cyt P460 are most consistent with those of canonical porphyrin structures, where typical C–C distances and angles are 1.39 Å and 125°, respectively.[44] For isoporphyrin or phlorin structures, typical distances about the sp^3 hybridized C are

elongated to approximately 1.5 Å.[45–46] The geometry about the cross-link also agrees with a more planar *meso* C, consistent with an sp^2—and not sp^3—hybridized C.

The 13′ C geometry indicates that the cyt P460 heme macrocycle is dibasic. Moreover, the TDDFT-predicted spectra show that a hybridization change of the 13′ *meso* C is not necessary for the unusual spectroscopic characteristics of cyt P460. That is, an equally red-shifted Soret is predicted for all three macrocycle variants: porphyrin, isoporphyrin, and phlorin. The observed red-shift is thus best attributable to a combination of *meso* substitution and distortions of the macrocycle away from planarity. While the exact role of the cross-link remains unclear, a corrected electronic structural formulation now opens the possibility for detailed spectroscopic studies informed by high-level calculations.

We speculate that in addition to N lone pair participation in the porphyrin π system, another possible role for the cross-link is to enforce and/or promote the observed heme distortion. Distortions in the heme that lead to out-of-plane deformations are common in many heme-bearing proteins.[40] The most common types of deformation are ruffling (alternating rotations around the Fe–N bonds), saddling (movement of the pyrrole rings up and down), and doming (bending of all pyrrole rings in one direction).[36, 37, 40] Ruffling has been observed in the P460 cofactors of both HAO and cyt P460. Many properties of certain heme enzymes have been attributed to ruffling, including a lower reduction potential[37] and stronger bonding between the Fe and axial ligand.[49] A lower reduction potential for the cyt P460 and the HAO active sites may be why these centers avoid auto-reduction upon NO binding via reductive nitrosylation.[13] Moreover, maintaining coordination of axial histidine Fe

during enzymatic turnover has been shown to be important to the activity of cyt P460.[10] Thus, the common characteristics of ruffling seem to agree with the requirements of cyt P460 and HAO activity; however, there are problems with ascribing the properties of these enzymes entirely to ruffling, which is present in many enzymes besides cyt P460 and HAO (**Table 3.3**) that have not been shown to be competent for NH_2OH oxidation.

Ruffling has been offered as the reason for favoring $\{FeNO\}^6$ over $\{FeNO\}^7$ formation in the enzyme nitrophorin.[48] In certain blood-sucking insects, nitrophorin is used to release NO into the blood stream of the victim, leading to vasodilation and inhibition of platelet aggregation. This behavior relies on $\{FeNO\}^6$ formation, as an $\{FeNO\}^7$ species has a NO dissociation constant in the picomolar range (compared to micromolar for the $\{FeNO\}^6$).[49] The ruffling observed in nitrophorin is suggested to stabilize the protein against reductive nitrosylation, favoring the formation of the $\{FeNO\}^6$ necessary for release of NO in the victim, which is triggered by changes in pH or dilution of the concentration of NO around the protein.[48] The degrees of ruffling observed for cyt P460 and NO-bound nitrophorin are similar, so this agrees with the observation of an $\{FeNO\}^6$ species formed by cyt P460, which persists in anaerobic solution unless NH_2OH is introduced or if NO is removed from solution. The appreciable difference in the ruffling of HAO vs. cyt P460 and nitrophorin may also explain differential NO binding affinities between these two enzymes.[3] Thus, not only the presence of ruffling, but the degree to which the heme is distorted out of the plane seem important to reactivity, and may be heavily influenced by the cross-link. These structure-function relations will be explored in future work informed by enhanced

definition of the cofactor afforded by our aforementioned structure.

D. CONCLUSION

We have obtained the structure of *Nitrosomonas* sp. AL212 cyt P460, which includes a complete active site pocket absent in prior cyt P460 structures. The metrical parameters about the cross-link indicate that the cross-link does not impose any hybridization change on the 13′ *meso* C from sp^2 to sp^3. Our TDDFT-predicted spectra also show that such a hybridization change is not necessary to red shift the cyt P460 UV-vis spectra relative to a standard heme. The cyt P460 cofactor macrocycle is therefore best formulated as a dibasic porphyrin. Though the complete role of the cross-link remains undefined, we have shown that it subtly alters the heme electronic structure. This arises both by virtue of substituting the *meso* position C–H for a C–N linkage, and furthermore the cross-link contributes an anti-bonding interaction between the Lys N lone pair and the π system of the macrocycle. This interaction destabilizes key molecular orbitals comprising transitions that give rise to the Soret absorption feature. We hypothesize that the cross-link may promote and/or enforce heme distortions essential to function. These possibilities will be explored in future work employing our more complete structural analysis of the heme P460 cofactor that will seek to understand the exact effects of heme cross-links and distortions as they relate to Fe-nitrosyls and selective redox transformations of energetic substrates such as NH_2OH.

E. REFERENCES

(1) Elmore, B. O., Bergmann, D. J., Klotz, M. G., and Hooper, A. B. (2007) Cytochromes P460 and c'-beta; a new family of high-spin cytochromes c, *FEBS Lett. 581*, 911-916.

(2) Bergmann, D. J., Zahn, J. A., Hooper, A. B., and DiSpirito, A. A. (1998) Cytochrome P460 genes from the methanotroph Methylococcus capsulatus bath, *J. Bacteriol. 180*, 6440-6445.

(3) Caranto, J. D., Vilbert, A. C., and Lancaster, K. M. (2016) Nitrosomonas europaea cytochrome P460 is a direct link between nitrification and nitrous oxide emission, *Proc Natl Acad Sci USA 113*, 14704-14709.

(4) Pearson, A. R., Elmore, B. O., Yang, C., Ferrara, J. D., Hooper, A. B., and Wilmot, C. M. (2007) The crystal structure of cytochrome P460 of Nitrosomonas europaea reveals a novel cytochrome fold and heme-protein cross-link, *Biochemistry 46*, 8340-8349.

(5) Bergmann, D. J., and Hooper, A. B. (1994) The primary structure of cytochrome P460 of Nitrosomonas europaea: presence of a c-heme binding motif, *FEBS Lett. 352*, 324-326.

(6) Bergmann, D. J., and Hooper, A. B. (2003) Cytochrome P460 of Nitrosomonas europaea. Formation of the heme-lysine cross-link in a heterologous host and mutagenic conversion to a non-cross-linked cytochrome c', *Eur. J. Biochem. 270*, 1935-1941.

(7) Arciero, D. M., and Hooper, A. B. (1997) Evidence for a crosslink between c-heme and a lysine residue in cytochrome P460 of Nitrosomonas europaea, *FEBS Lett. 410*, 457-460.

(8) Arciero, D. M., and Hooper, A. B. (1998) Consideration of a phlorin structure for haem P-460 of hydroxylamine oxidoreductase and its implications regarding reaction mechanism, *Biochem. Soc. T. 26*, 385-389.

(9) Elmore, B. O., Pearson, A. R., Wilmot, C. M., and Hooper, A. B. (2006) Expression, purification, crystallization and preliminary X-ray diffraction of a novel Nitrosomonas europaea cytochrome, cytochrome P460, *Acta Crystallogr. F 62*, 395-398.

(10) Vilbert, A. C., Caranto, J. D., and Lancaster, K. M. (2018) Influences of the Heme-Lysine Crosslink in Cytochrome P460 Over Redox Catalysis and Nitric Oxide Sensitivity, *Chem. Sci. 9*, 368-379.

(11) Enemark, J., and Feltham, R. (1974) Principles of structure, bonding, and reactivity for metal nitrosyl complexes, *Coord. Chem. Rev. 13*, 339-406.

(12) Attia, A. A., and Silaghi-Dumitrescu, R. (2014) Computational investigation of the initial two-electron, two-proton steps in the reaction mechanism of hydroxylamine oxidoreductase, *J. Phys. Chem. B 118*, 12140-12145.

(13) Hendrich, M. P., Upadhyay, A. K., Riga, J., Arciero, D. M., and Hooper, A. B. (2002) Spectroscopic characterization of the NO adduct of hydroxylamine oxidoreductase, *Biochemistry 41*, 4603-4611.

(14) Caranto, J. D., and Lancaster, K. M. (2017) Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc Natl Acad Sci USA 114*, 8217-8222.

(15) Cedervall, P., Hooper, A. B., and Wilmot, C. M. (2013) Structural studies of hydroxylamine oxidoreductase reveal a unique heme cofactor and a previously unidentified interaction partner, *Biochemistry 52*, 6211-6218.

(16) Thöny-Meyer, L., Fischer, F., Künzler, P., Ritz, D., and Hennecke, H. (1995) Escherichia coli genes required for cytochrome c maturation, *J. Bacteriol. 177*, 4321-4326.

(17) Golombek, A. P., and Hendrich, M. P. (2003) Quantitative analysis of dinuclear manganese(II) EPR spectra, *J. Magn. Reson. 165*, 33-48.

(18) Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., and Grosse-Kunstleve, R. W. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D 66*, 213-221.

(19) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr. D 66*, 486-501.

(20) Schrodinger, LLC. (2015) The PyMOL Molecular Graphics System, Version 1.8.

(21) Neese, F. (2015) ORCA-An ab initio, Density Functional and Semiempirical Program Package (v. 3.0. 3) Max-Planck Institute for Bioinorganic Chemistry, *Mülheim an der Ruhr, Germany*.

(22) Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A 38*, 3098-3100.

(23) Perdew, J. P. (1986) Density-functional approximation for the correlation energy of the inhomogeneous electron gas, *Phys. Rev. B 33*, 8822-8824.

(24) Pantazis, D. A., Chen, X.-Y., Landis, C. R., and Neese, F. (2008) All-Electron Scalar Relativistic Basis Sets for Third-Row Transition Metal Atoms, *J. Chem. Theory Comput. 4*, 908-919.

(25) van Lenthe, E., van der Avoird, A., and Wormer, P. E. S. (1998) Density functional calculations of molecular hyperfine interactions in the zero order regular approximation for relativistic effects, *J. Chem. Phys. 108*, 4783-4796.

(26) Klamt, A., and Schüürmann, G. (1993) COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient, *J. Chem. Soc., Perk. T. 2*, 799-805.

(27) Hirata, S., and Head-Gordon, M. (1999) Time-dependent density functional theory within the Tamm–Dancoff approximation, *Chem. Phys. Lett. 314*, 291-299.

(28) Stephens, P., Devlin, F., Chabalowski, C., and Frisch, M. J. (1994) Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields, *J. Phys. Chem. 98*, 11623-11627.

(29) Becke, A. D. (1992) Density-functional thermochemistry. I. The effect of the exchange-only gradient correction, *J. Chem. Phys. 96*, 2155-2160.

(30) Neese, F. (2002) Prediction and interpretation of the 57 Fe isomer shift in Mössbauer spectra by density functional theory, *Inorg. Chim. Acta 337*, 181-192.

(31) Neese, F., Wennmohs, F., and Hansen, A. (2009) Efficient and accurate local approximations to coupled-electron pair approaches: An attempt to revive the pair natural orbital method, *J. Chem. Phys. 130*, 114108.

(32) Izsák, R., and Neese, F. (2011) An overlap fitted chain of spheres exchange method, *J. Chem. Phys. 135*, 144105.

(33) Neese, F. (2006) Importance of Direct Spin-Spin Coupling and Spin-Flip Excitations for the Zero-Field Splittings of Transition Metal Complexes: A Case Study, *J. Am. Chem. Soc. 128*, 10213-10222.

(34) Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. (2004) UCSF Chimera—A visualization system for exploratory research and analysis, *J. Comput. Chem. 25*, 1605-1612.

(35) Graves, A. B., Graves, M. T., and Liptak, M. D. (2016) Measurement of Heme Ruffling Changes in MhuD Using UV-vis Spectroscopy, *J. Phys. Chem. B 120*, 3844-3853.

(36) Jentzen, W., Song, X. Z., and Shelnutt, J. A. (1997) Structural characterization of synthetic and protein-bound porphyrins in terms of the lowest-frequency normal coordinates of the macrocycle, *J. Phys. Chem. B 101*, 1684-1699.

(37) Liptak, M. D., Wen, X., and Bren, K. L. (2010) NMR and DFT investigation of heme ruffling: functional implications for cytochrome c, *J. Am. Chem. Soc. 132*, 9753-9763.

(38) Inclusion of B-factor anisotropy on non-H atoms decreased Rfree from 18% to 15.4%.

(39) Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions, *Nat. Methods 8*, 785-786.

(40) Jentzen, W., Ma, J. G., and Shelnutt, J. A. (1998) Conservation of the conformation of the porphyrin macrocycle in hemoproteins, *Biophys. J. 74*, 753-763.

(41) Kondrashov, D. A., Roberts, S. A., Weichsel, A., and Montfort, W. R. (2004) Protein functional cycle viewed at atomic resolution: Conformational change and mobility in nitrophorin 4 as a function of pH and NO binding, *Biochemistry 43*, 13637-13647.

(42) Maurus, R., Overall, C. M., Bogumil, R., Luo, Y., Mauk, A. G., Smith, M., and Brayer, G. D. (1997) A myoglobin variant with a polar substitution in a conserved hydrophobic cluster in the heme binding pocket, *Biochim. Biophys. Acta 1341*, 1-13.

(43) Gouterman, M. (1961) Spectra of Porphyrins, *J. Mol. Spectrosc. 6*, 138-163.

(44) Fleischer, E. B. (1970) The Structure of Porphyrins and Metalloporphyrins, *Acc. Chem. Res. 3*, 105-112.

(45) Abhilash, G. J., Bhuyan, J., Singh, P., Maji, S., Pal, K., and Sarkar, S. (2009) NO_2-Mediated meso-Hydroxylation of Iron (III) Porphyrin, *Inorg. Chem. 48*, 1790-1792.

(46) Pistner, A. J., Yap, G. P., and Rosenthal, J. (2012) A Tetrapyrrole Macrocycle Displaying a Multielectron Redox Chemistry and Tunable Absorbance Profile, *J Phys. Chem. C 116*, 16918-168924.

(47) Kleingardner, J. G., Bowman, S. E. J., and Bren, K. L. (2013) The Influence of Heme Ruffling on Spin Densities in Ferricytochromes c Probed by Heme Core C-13 NMR, *Inorg. Chem. 52*, 12933-12946.

(48) Walker, F. A. (2005) Nitric oxide interaction with insect nitrophorins and thoughts on the electron configuration of the {FeNO}6 complex, *J. Inorg. Biochem. 99*, 216-236.

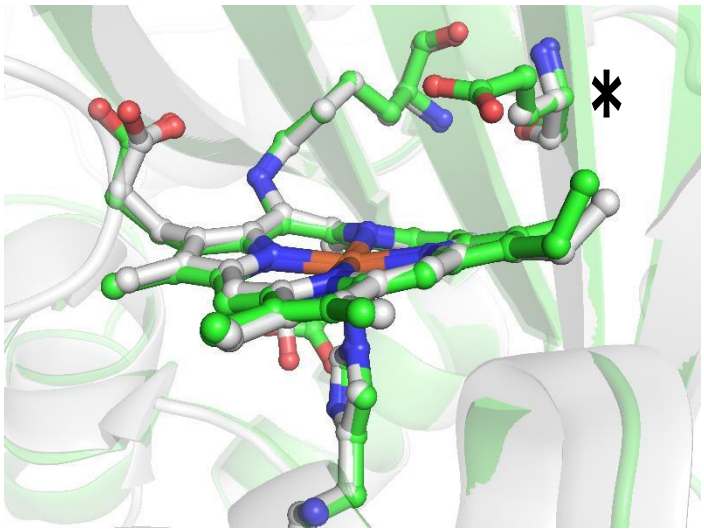
(49) Weichsel, A., Andersen, J. F., Roberts, S. A., and Montfort, W. R. (2000) Nitric oxide binding to nitrophorin 4 induces complete distal pocket burial, *Nat. Struct. Biol. 7*, 551-554.

CHAPTER 4: OUTER-SPHERE GATING OF SUBSTRATE OXIDATION AND N–N BOND FORMTAION IN CYTOCHROME P460

P460 cofactors are modified c-type hemes found in enzymes from aerobic ammonia oxidizing (AOB), anaerobic ammonia oxidizing (anammox), methanotrophic, and—possibly—denitrifying bacteria.[1, 2] P460 hemes are characterized by additional cofactor/amino acid cross-links. In octaheme enzymes such as HAO, 2 cross-links are contributed by a tyrosine, which binds through its hydroxy oxygen and C_3 to meso- and alpha-pyrrolic positions on the macrocycle (**Figure 1.4a**). In monoheme cytochrome P460 enzymes, a lysine sidechain forms a single N–C bond to a macrocyclic meso C (**Figure 1.4b**). The cofactor remains a dibasic macrocycle in either case,[3, 4] although the cross-link and ruffling result in a highly red-shifted UV-vis absorption profile with diminished absorptivity relative to standard c-hemes.[4, 5] To date, heme P460 cofactors have been shown to effect selective oxidation of NH_2OH to nitric oxide NO. In HAO, this NO rapidly dissociates and is ultimately oxidized to NO_2^-.[6] In monoheme cytochrome P460 enzymes, the NO remains bound to ferric heme P460 for a sufficient duration to allow nucleophilic attack by a second equivalent of NH_2OH to form N_2O.[7] The establishment of selective oxidation of NH_2OH to NO by hemes P460 overturns decades of convention, which incorrectly identified these cofactors as the source of NO_2^- produced by AOB as the stoichiometric product of NH_3 oxidation.

We recently reported the 1.45 Å crystal structure of cyt P460 from *Nitrosomonas* sp. AL212 (ref. 4, Chapter 3). We noted that, despite exhibiting identical inner-sphere structural features as well as spectroscopic features to the *N. europaea* variant, the AL212 cyt P460 was not competent for oxidation of NH_2OH. Expanding scope to include the outer coordination sphere surrounding AL212's P460 cofactor, we found that a key difference to the *N. europaea* variant is the substitution of an alanine residue

Figure 4.1 Comparison of cyt P460 active sites from *Nitrosomonas europaea* (green, PDB ID: 2JE3) and *Nitrosomonas* sp. AL212 (white, PDB ID: 6AMG).



Cyt P460 from *Nitrosomonas europaea* (green, PDB ID: 2JE3) and *Nitrosomonas* sp. AL212 (white, PDB ID: 6AMG). Residue 131 (97 in *N. europaea*) indicated by *.

for a glutamate at position 131 (Position 97 in *N. europaea*, **Figure 4.1**). We hypothesized that this glutamate operates as a proton relay during what should be a proton coupled oxidation of Fe^{III}–NH_2OH to $\{FeNO\}^7$. To explore this possibility, we generated several mutant variants of cyt P460 and analyzed their physical properties and ability to oxidize NH_2OH.

A. MATERIALS AND METHODS

General Considerations

Milli-Q water (18.2 MΩ; Millipore) was used in the preparation of all buffers and solutions. UV- visible (UV-vis) absorption spectra were obtained using a Cary 60 UV-vis spectrometer with temperature control set to 25°C. Data were fit using Igor Pro version 6.37 (WaveMetrics). For the generation of the $\{FeNO\}^6$ species, the NO-

donor disodium 1-(Hydroxyl-NNO-azoxy)-L-proline (PROLI-NONOate, Cayman Chemicals) was used. All other chemicals were purchased from VWR International.

Plasmids and Mutagenesis

Codon-optimized genes for cyt P460 from *Nitrosomonas* sp. AL212 or *Nitrosomonas europaea* were synthesized and cloned into the *NcoI* and *XhoI* sites of pET-22b(+) vector by GenScript, Inc. These vectors were designed to include a C-terminal His-tag found in the parent pET-22b(+) plasmid. Mutant variations of either cyt P460 gene were generated using site-directed mutagenesis (primers for each variant can be found in **Table 4.1**).

Table 4.1 – Site-directed mutagenesis primers; Red indicates mutagenesis sites
R = reverse primer, F= forward primer

Primer Name	Sequence (5´→ 3´)
Ala131X R	GAT GCC GTT AAA CTC GCC CGG GAA ATA GCC
Ala131Glu F	GAA GCG ATG GTG AAG GAT AGC AAA CGT TAC CCG
Ala131Gln F	CAG GCG ATG GTG AAG GAT AGC AAA CGT TAC CCG
Ala131Leu F	CTG GCG ATG GTG AAG GAT AGC AAA CGT TAC CCG
Ala131Asp F	GAT GCG ATG GTG AAG GAT AGC AAA CGT TAC CCG
Glu97Ala R	CAG ACC AAT GTA ATC GCC CAT AAA ATA ACC
Glu97Ala F	GCG GCG AGC GTG AAA GAC TCT CAG CGT

Protein Expression, Purification, and Crystallization

Protein expression and purification for each variant was the same as for wild type (WT) *N.* sp. AL212 as described in Chapter 3. The same crystallization conditions as those described in Chapter 3 were also used. For the soaking experiments, 0.5 M NH_2OH or 200 μM NO from the NO-donor PROLI-NONOate

were added to the cryoprotectant solution.

Crystallographic Data Collection

X-ray diffraction experiments were conducted at beamline F1 of the Cornell High Energy Synchrotron Source (CHESS) and beamline 24-ID-E of the Advanced Photon Source (APS) Northeastern Collaborative Access Team (NE-CAT). The Eiger 16M detector on 24-ID-E beam line is funded by a NIH-ORIP HEI grant (S10OD021527). Crystals were irradiated at 100 K using X-rays with a wavelength (λ) of 0.979 Å. X-ray diffraction data were indexed, integrated, scaled, and merged using the programs XDS[8] and CCP4.[9] An initial model was generated in Phenix[10] using the molecular replacement method and the cyt P460 structure from WT *Nitrosomonas* sp. AL212 [Protein Data Bank (PDB) entry 6AMG]. Refinements and building to completion were then conducted using Phenix and Coot[11] respectively. PyMol[12] was used to create figures.

Steady-state Activity Assays

All assays were performed in septum-sealed cuvettes flushed with N_2 gas. Anaerobic solutions of NH_2OH were prepared and assayed by the method of Frear and Burrell[13] for determination of the stock NH_2OH concentration. Final concentrations of 50 µM 2,6-dichlorophenolindophenol (DCPIP), 6 µM phenazine methosulfate (PMS), and 1 µM cyt P460 were added to 2 mL of deoxygenated 50 mM sodium phosphate (pH 8.0). The reaction was initiated by adding an appropriate volume of the NH_2OH stock solution to the reaction mixture through the septum with a Hamilton syringe. The reaction was monitored by following the absorption of DCPIP at 605 nm. The rate of the first 10% of the total oxidant consumption was determined through linear regression.

This rate was converted to the rate of oxidant consumed by using $\varepsilon_{605\ nm}$ = 20.6 mM^{-1} cm^{-1}.[14] At least three trials were performed for each concentration of NH_2OH.

Determination of NH_2OH and NO K_d

Cyt P460 variants were titrated with NH_2OH or NO added to septum-sealed cuvettes with a Hamilton syringe and monitored by UV-vis spectrometry. The K_d for NH_2OH was determined by following the disappearance of the shoulder feature at about 414 nm and using **Eqn. 4.1**:

$$A_{414\ nm} = \frac{\Delta A_{414\ nm}[NH_2OH]_0}{K_d + [NH_2OH]_0} \quad (\mathbf{4.1})$$

For the NO K_d, a similar process was repeated, but following the formation of the feature around 455 nm. Experiments were performed anaerobically in 50 mM sodium phosphate, pH 8.0 buffer, with temperature kept at 25 °C.

EPR Spectroscopy

X-band (9.40-GHz) EPR spectra were collected on a sample containing 170 μM cyt P460 in 50 mM sodium phosphate, pH 8.0 with 25% (v/v) glycerol. NH_2OH bound spectra were prepared with the addition of 200 mM NH_2OH. The measurements were obtained using a Bruker Elexsys-II spectrometer equipped with a liquid He cryostat maintained at 10.0 K. EPR data were simulated using SpinCount.[15]

Spectroelectrochemical Potentiometric Titrations

The spectroelectrochemical titrations were performed anaerobically using a 1 mm pathlength Basi cell with a mesh Pt working electrode, Pt counter and Ag/AgCl reference electrode. Bulk electrolysis experiments were conducted and monitored using a WaveNow potentiostat, Pine Research instrument, on solutions containing 200 μM cyt P460 variant and 20 μM methyl viologen (used as an electrochemical

mediator) in 100 mM phosphate, 100 mM NaCl pH 8.0. Potentials between −50 mV and −650 mV were applied in −50 mV increments from −300 mV to −500 mV, then −25 mV increments from −500 mV to −650 mV vs Ag/AgCl. Following each addition of a potential, the solution was allowed to equilibrate for approximately 10 mins before a UV/vis absorption spectrum was collected over the range of 200-800 nm. The resulting full wavelength spectra were analyzed and fit to the linearized Nernst equation:

$$E = E^{\circ} - \frac{RT}{nF} ln(Q) \quad \textbf{(4.2)}$$

Preliminary Kinetic Isotope Effect Determination

Preliminary data for a potential kinetic isotope effect (KIE) were determined by performing steady-state assays (as described above) in D_2O and 50 mM tribasic sodium phosphate buffer (pD 8.0), and with deuterated NH_2OH (ND_2OD) added. KIE was calculated by comparing the turnover frequency (TOF) in D_2O vs. H_2O as described by **Eqn. 4.3**:

$$\text{KIE} = \frac{\text{TOF}_{\text{H2O}}}{\text{TOF}_{\text{D2O}}} \quad \textbf{(4.3)}$$

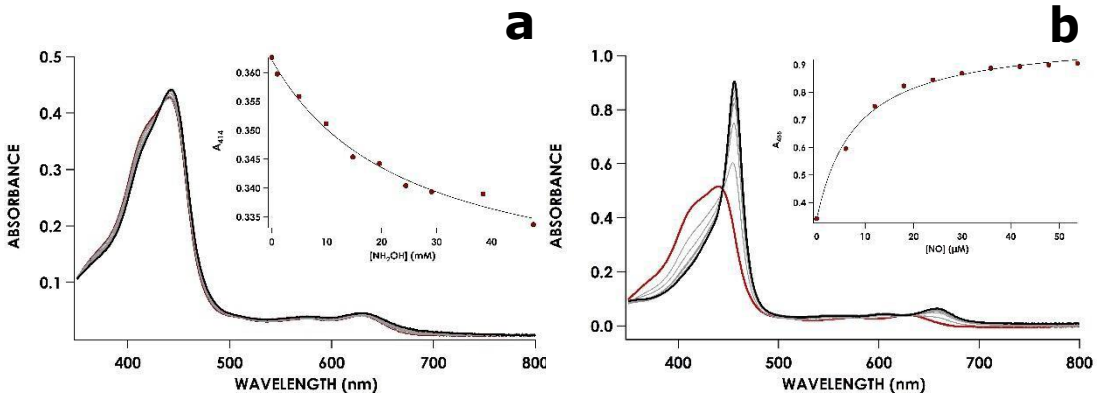
B. RESULTS

WT AL212 cyt P460, although incompetent for turnover, supports stable Fe^{III}–NH_2OH and $\{FeNO\}^6$ intermediates (when exogenous NH_2OH or the NO-donor PROLI-NONOate are added to solution, respectively) that were previously identified as pathway intermediates in the cyt P460 catalytic cycle for NH_2OH oxidation to NO/N_2O. We explored whether any significant differences in NH_2OH or NO binding affinities

distinguish catalytically inactive cyts P460**.** The 298 K K_d for NH_2OH binding to Fe^{III} cyt P460 at pH 8.0 is 9 ± 1 mM for the *N. europaea* variant. The WT AL212 variant exhibits a modest, 2-fold diminution in K_d to 18 ± 1 mM. Meanwhile, binding affinities for NO at pH 8.0 are essentially identical: for *N. europaea* this K_d is 10 μM ± 2 μM, while for AL212 it is 8 μM ± 1 μM. Representative examples of each titration can be found in **Figure 4.2**, and aggregate binding affinities are compiled in **Table 4.2**. The $Fe^{II/III}$ reduction potential for *N. europaea* cyt P460 is -400 ± 5 mV vs. NHE, while it is -424 ± 7 mV for AL212 (**Figure 4.3, Table 4.2**). Overall, these similarities are not surprising given that the overall cofactor structures in the two cyt P460 variants are effective superimposable (RMSD = 0.074 Å, **Figure 4.8**) and exhibit very similar UV-vis features (**Figure 4.4**). Despite these similarities, the AL212 Fe^{III}–NH_2OH adduct is redox inactive using either 2,6-dichlorophenolindophenol (DCPIP, +217 mV),

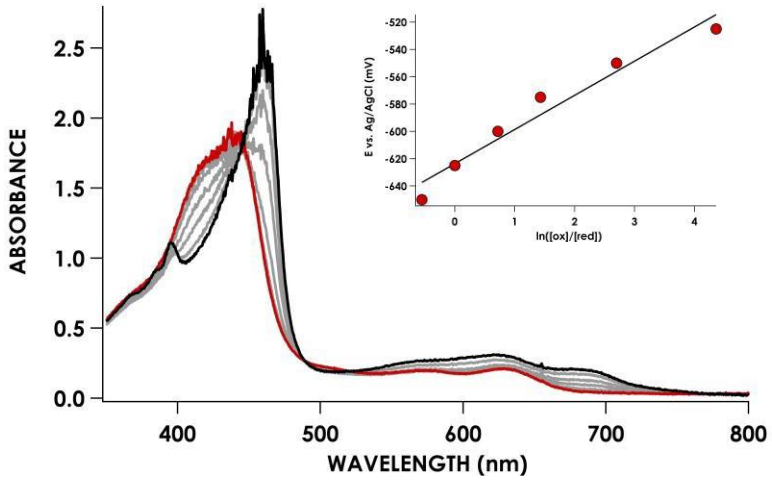
---

Figure 4.2 Determination of NH_2OH and NO K_d.



Representative example of WT *N.* sp. AL212 cyt P460 titration curve with NH_2OH (a) or NO (b). Insets show the plot of hyperbolic equations to determine K_d.

Figure 4.3 Spectroelectrochemical Potentiometric Titration of Cyt P460



UV/vis absortption spectra of WT AL212 cyt P460 at pH 8.0 as a function of applied potential. No potenial was applied in the red spectrum and the following spectrums in gray are in increments of −50 mV from −300 mV to −500 mV, then −25 mV increments from −500 mV to −650 mV vs Ag/AgCl. The inset is the linearized Nernst plot of the spectroelectrochemical data.

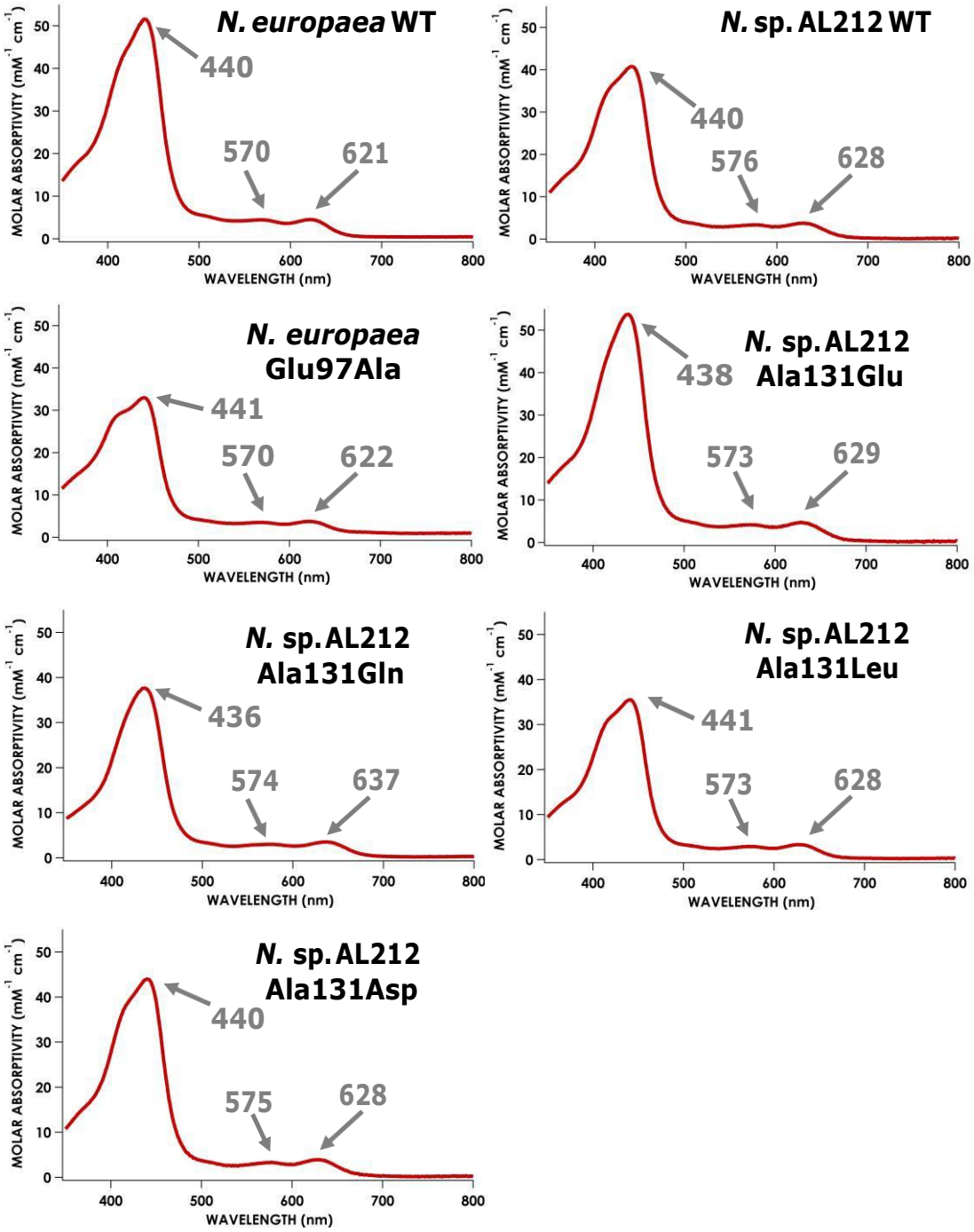
phenazine methosulfate (PMS, +80 mV), or hexaammonium ruthenium (III) chloride ($[Ru(NH_3)_6]Cl_3$, –8.3 mV) as oxidant, nor does its $\{FeNO\}^6$ undergo attack by NH_2OH to form the N–N bond of N_2O.

Introduction of a glutamate at position 131 in the AL212 cyt P460 restored catalytic competence for Fe^{III}–NH_2OH oxidation to NO and generation of N_2O by the AL212 protein. When 1 mM NH_2OH and 70 μM of the oxidant DCPIP were added to WT *N.* sp. AL212 cyt P460, no cyt P460 intermediates (*i.e.* $\{FeNO\}^6$, $\{FeNO\}^7$) were observed. Only basal, background consumption of DCPIP—monitored by following its absorbance at 605 nm—occurs, but this can also be observed in the absence of any cyt

Table 4.2: Properties of Variant Cyt P460s

				Fe^{III} component 1		Fe^{III} component 2		Fe–NH_2OH
Variant	**Reduction Potential (mV vs. NHE)**	**NH_2OH K_d**	**NO K_d**	**g_{eff} values**	**E/D**	**g_{eff} values**	**E/D**	**g_{eff} values**
WT *N. europaea*	-400 ± 5	9 ± 1 mM	10 ± 2 μM	6.57, 5.09, 1.97	0.03	N/A	N/A	2.75, 2.28, 1.54
WT *N.* sp. AL212	-424 ± 7	18 ± 1 mM	8 ± 1 μM	6.39, 5.13, 1.97	0.03	6.00, 5.52, 1.99	0.012	2.84, 2.25, 1.44
AL212 Ala131Glu	-428 ± 2	16 ± 5 mM	5 ± 1 μM	6.40, 5.14, 1.97	0.03	6.00, 5.51, 1.99	0.012	2.86, 2.27, 1.46
AL212 Ala131Gln	-406 ± 2	15 ± 3 mM	6 ± 1 μM	6.51, 5.12, 1.97	0.03	6.03, 5.53, 1.99	0.012	2.78, 2.28, 1.49
AL212 Ala131Leu	-381 ± 10	12 ± 3 mM	2 ± 0.5 μM	6.40, 5.11, 1.98	0.03	6.00, 5.48, 1.99	0.012	2.80, 2.27, 1.46
AL212 Ala131Asp	-388 ± 7	19 ± 7 mM	4 ± 1 μM	6.40, 5.12, 1.97	0.03	6.03, 5.50, 1.99	0.012	2.86, 2.25, 1.44

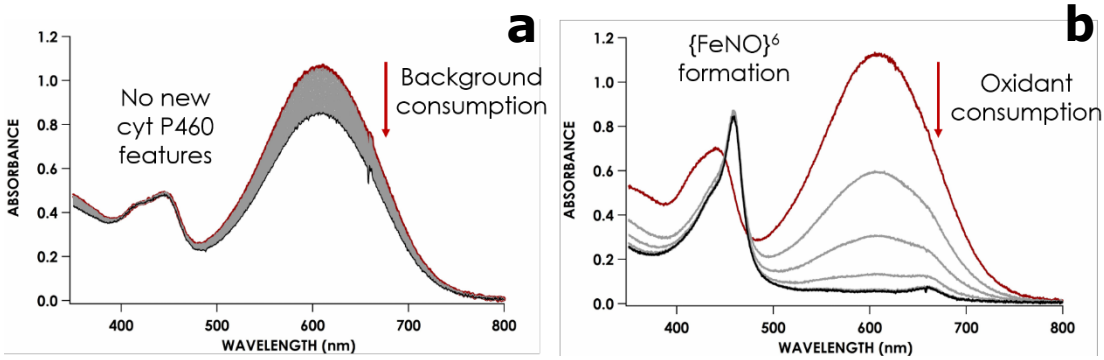
Figure 4.4 UV-vis Characteristics of Cyt P460 Variants



P460 (0.44 ± 0.19 μM DCPIP · mM^{-1} NH_2OH · min^{-1} background vs. 0.43 ± 0.02 μM DCPIP · μM^{-1} *N.* sp. AL212 WT cyt P460 · mM^{-1} NH_2OH · min^{-1}). Under the same conditions with the Ala131Glu variant, however, rapid formation of the cyt P460 $\{FeNO\}^6$ species concomitant with rapid oxidant consumption were observed (**Figure 4.5**). Using DCPIP as oxidant, Ala131Glu AL212 cyt P460 oxidizes NH_2OH with a turnover frequency of 2.1 ± 0.05 μM DCPIP · μM^{-1} cyt P460 · mM^{-1} NH_2OH · min^{-1}, ca. half of that of the wild-type *N. europaea* variant: 4.5 ± 0.06 μM DCPIP · μM^{-1} cyt P460 · mM^{-1} NH_2OH · min^{-1} (**Figure 4.6**).

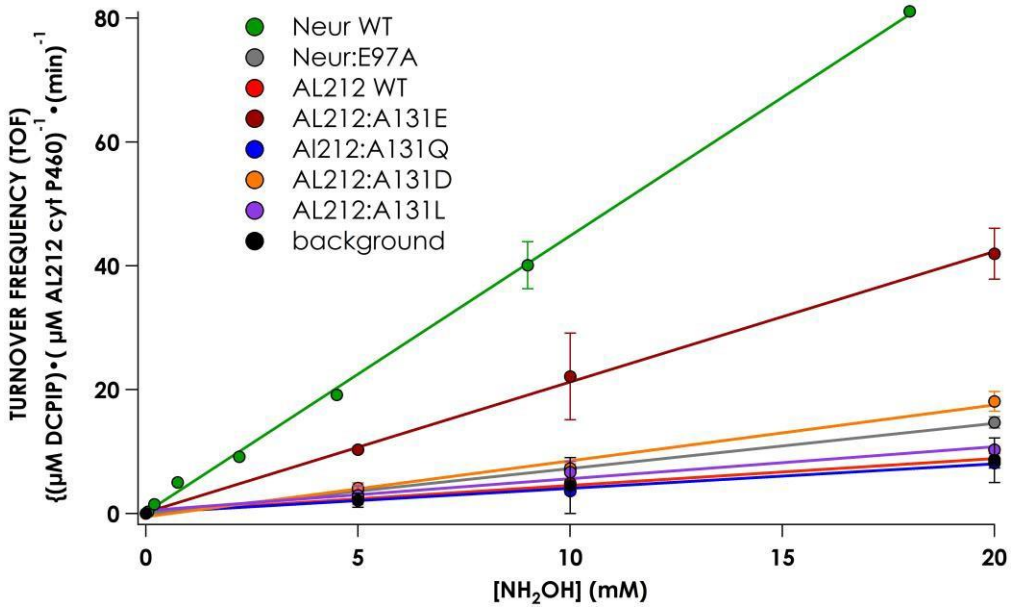
To explore the role of position 131 in NH_2OH oxidation catalysis, we generated several other Ala131X variants of AL212 cyt P460: Ala131Gln, Ala131Asp, and Ala131Leu. None of these variants exhibited significant NH_2OH oxidation activity above basal levels using DCPIP as oxidant, emphasizing the importance of a Glu residue in this position (**Figure 4.6**). Though interesting that the carboxylate-containing Ala131Asp variant was also unable to catalyze NH_2OH oxidation, we hypothesize that this is due to the inability of the relatively shorter side chain to effectively interact with bound NH_2OH (*vide infra*). In further accord with our hypothesis, Glu97Ala substitution obviates NH_2OH oxidation in the *N. europaea* protein (**Figure 4.6**). We note that these additional variants are all able to form Fe^{III}–NH_2OH and $\{FeNO\}^6$ adducts when treated with either NH_2OH or the NO-donor PROLI-NONOate, respectively. Again, these variants show similar NH_2OH/NO binding affinities to WT variants (**Table 4.2**). All variants also exhibit similar reduction potentials and spectral characteristics (**Table 4.2**).

Figure 4.5 Ala131Glu Mutation Allows for NH_2OH Oxidation with DCPIP as Oxidant



Reaction of 12 μM WT *N.* sp. AL212 cyt P460 (a) or AL212 Ala131Glu (b) with 1 mM NH_2OH and 70 μM DCPIP. Reaction was initiated with the addition of NH_2OH.

Figure 4.6 Steady-state NH_2OH Oxidase Activity Plot for Cyt P460 Variants

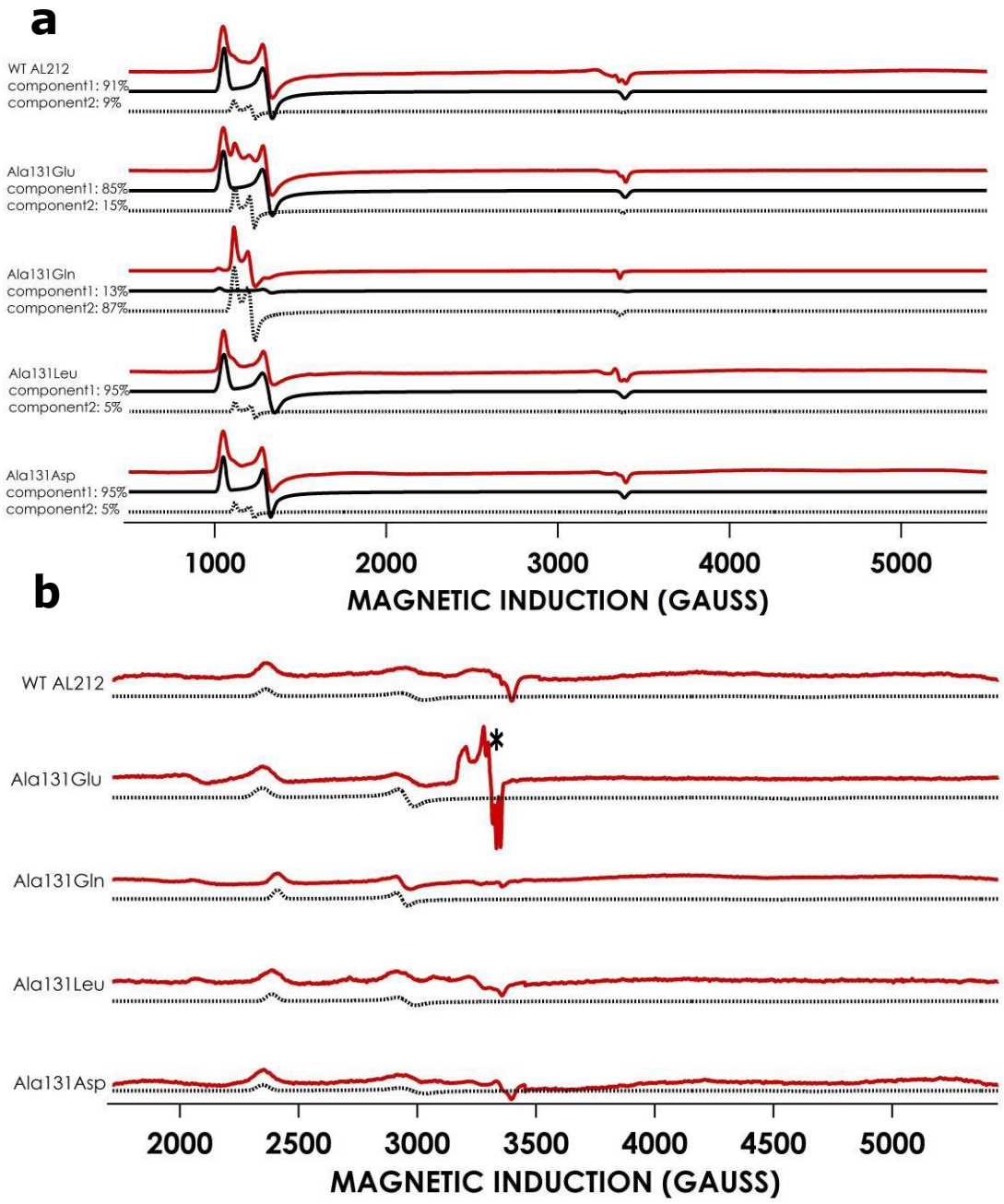


Assay conditions were 1 μM cyt P460, 6 μM phenazine methosulfate (PMS), and 70 μM DCPIP with various NH_2OH concentrations in anaerobic 50 mM sodium phosphate, pH 8.0, at 25 °C. Each data point is the average of at least three trials, with error bars representing one standard deviation. Background indicates the consumption of DCPIP under the same conditions with no cyt P460 present.

To further explore the nature of the AL212 cyt P460 variants, we also obtained EPR spectra for each in the resting Fe^{III} form (**Figure 4.7a**). Though all variants are high-spin (similar to WT AL212), it is clear that the AL212 cyt P460 variants contain two components. The components appear to be almost identical in each variant, where the only difference is the relative concentration of each component compared to the other (**Figure 4.7a**). We also obtained EPR spectra of the NH_2OH–bound forms of each AL212 variant, which exhibit single spin systems with spin Hamiltonian parameters similar to WT *N. europaea* cyt P460[7] (**Figure 4.7b**). Interestingly, the EPR of the Fe–NH_2OH species for the Ala131Glu variant also shows trace amounts of an $\{FeNO\}^7$ intermediate, which we observed in earlier characterization of the *N. europaea* cyt P460 Fe^{III}–NH_2OH and attributed to O_2 contamination during sample preparation. We do not observe this $\{FeNO\}^7$ intermediate in any of the other variants. Together, these results indicate that a basic residue in the distal pocket of a cyt P460 is essential for NH_2OH oxidation catalysis.

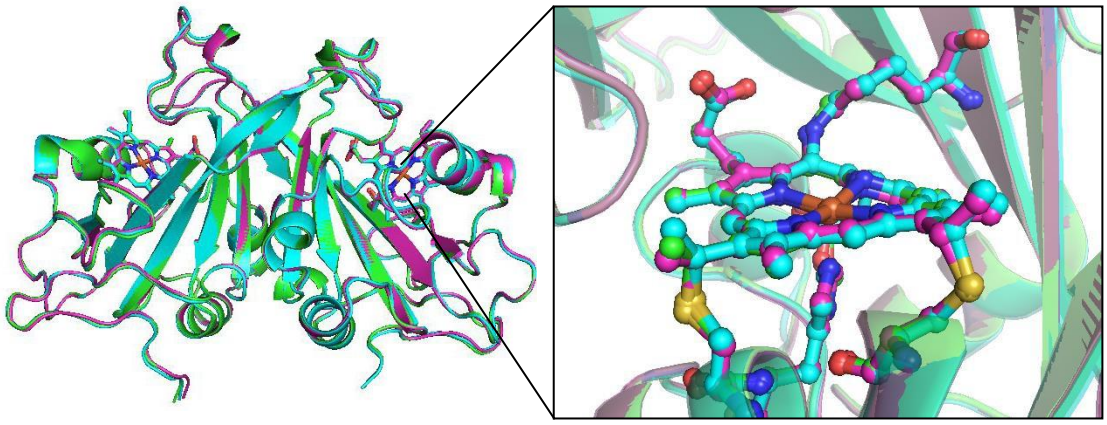
Crystal structures for Fe^{III} Ala131Glu and Ala131Gln were also obtained to further explore the effects of the engineered mutations. In addition, we were able to obtain structures of the AL212 cyt P460 Ala131Gln Fe^{III}–NH_2OH and the Ala131Glu $\{FeNO\}^6$ adducts. (**Figure 4.9**). Comparing the overall structure of Fe^{III} Ala131Glu, Ala131Gln, and WT AL212 shows that the mutation does not affect any overall structure that may affect reactivity (**Figure 4.8**). The global structures of all variants including the resting Fe^{III} WT AL212 cyt P460 are superimposable, with the largest pairwaise RMSD values being 0.294 between WT AL212 and Ala131Gln. As expected, the largest structural deviations are encountered in the distal pocket above the heme P460 cofactor.

Figure 4.7 EPR Spectra of Cyt P460 Variants



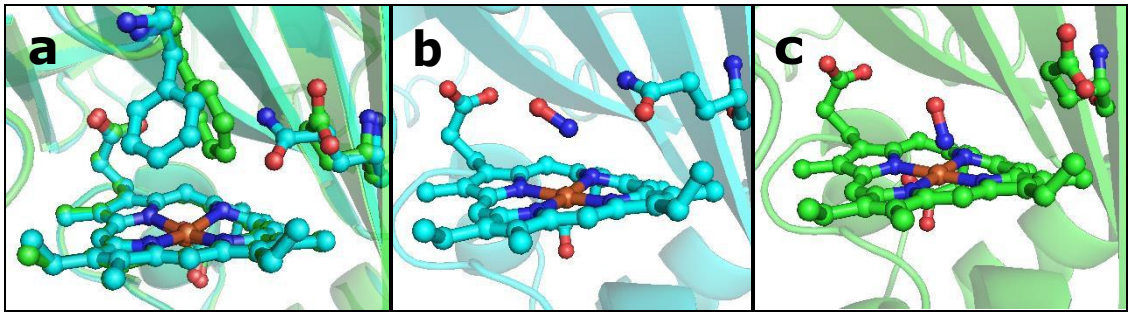
10 K X-band (9.40 GHz) EPR spectrum of Fe^{III} (a) and NH_2OH–bound (b) variants recorded at 633 μW microwave power. Experimental data is in red, simulations are in black solid and dashed lines. *Indicates the presence of the $\{FeNO\}^7$ intermediate formed as a result of O_2 contamination.

Figure 4.8 Overlay of WT Al212, Ala131Glu, and Ala131Gln Dimers



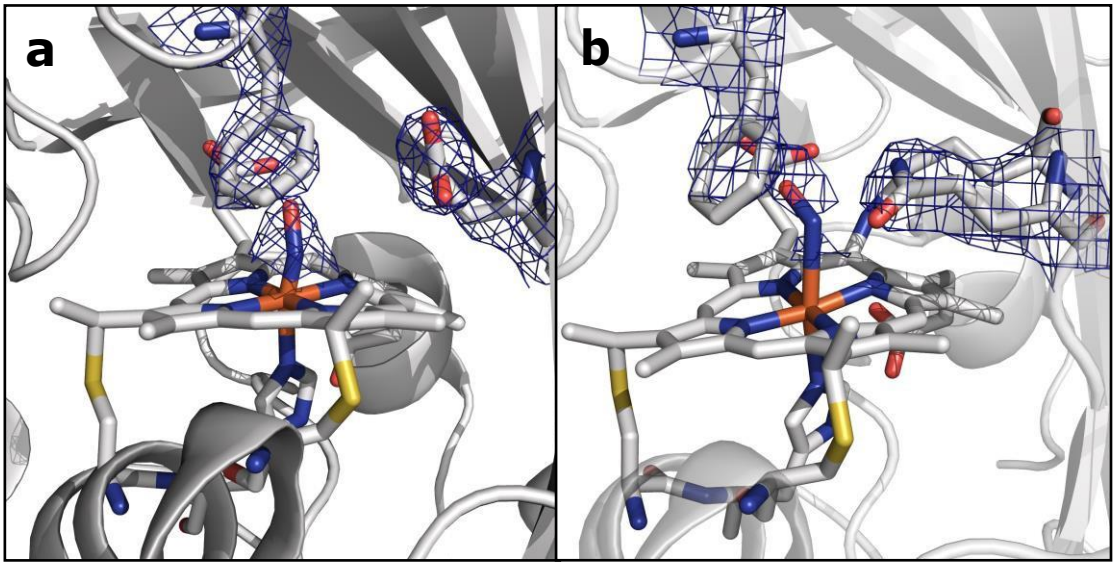
Structural homology between Fe^{III} WT AL212 (purple, PDB ID: 6AMG), Ala131Glu (green, PDB ID: 6EOX), and Ala131Gln (cyan, PDB ID: 6EOZ).

Figure 4.9 Active Site Configuration of Various Species



Active site view showing orientation of residue 131. (a) Comparison of Ala131Glu (green) and Ala131Gln (cyan) active sites with Phe76 and residue 131 highlighted in each. (b) Structure of Ala131Gln with NH_2OH bound (PDB ID: 6EOY). (c) Structure of Ala131Glu with NO bound (PDB ID 6E17).

Figure 4.10 Electron Density of Bound Substrates to Ala131Glu and Ala131Gln



$2F_o - F_c$ simulated annealing composite omit maps generated on final structures for Ala131Glu–NO (a) and Ala131Gln–NH_2OH (b). Structures represented as sticks shown in grey. $2F_o - F_c$ simulated annealing composite omit maps to σ level 1.0 represented as blue mesh.

Table 4.3 Bond Distances and Angles

	Ala131Glu–NO	Ala131Gln–NH_2OH
PDB ID	6E17	6EOY
Fe–ligand distance in Å (chain A, B, C, D)	1.50, 1.39, 1.57, 1.78	2.70, 2.75, 2.70, 2.70
Fe–ligand angle in deg. (chain A, B, C, D)	146.36, 152.19, 134.99, 108.40	121.22, 133.54, 141.47, 137.43

Preliminary kinetic isotope effect (KIE) data also suggest the glutamate residue may play a potential role in catalysis that relates to the initial proton coupled electron transfer (PCET) steps. When in deuterated buffer with deuterated NH_2OH (ND_2OD)

added, the observed KIE for Ala131Glu is approximately 2.5 (TOF = 0.84 ± 0.16 μM DCPIP · μM^{-1} cyt P460 · mM^{-1} NH_2OH · min^{-1}). This implies that there is at least one proton transfer affected by this mutation.

C. DISCUSSION

In order to dismiss the possibility that structural changes or other factors related to the mutation are in fact the cause of the changes in activity and not the presence of a carboxylate-containing residue, we also prepared Ala131Gln, Ala131Leu, and Ala131Asp variants in addition to the Ala131Glu. These variants showed spectral and chemical properties similar to WT AL212 and Ala131Glu (**Table 4.2**) and could also form Fe^{III}–NH_2OH intermediates (**Table 4.2, Figure 4.7**), however lacked the ability to oxidize NH_2OH, behaving as the WT AL212 cyt P460. Though interesting that the carboxylate-containing Ala131Asp variant was also unable to oxidize NH_2OH, this is most likely due to the inability of the relatively shorter side chain to effectively interact with bound NH_2OH. Likewise, the corresponding mutation to the *N. europaea* in which Glu97 was replaced with Ala lost its ability to oxidize NH_2OH.

Due to the chemical differences in the sidechains of alanine and glutamine, it is hypothesized that this residue may serve as proton relay during PCET events in the cyt P460 catalytic cycle (**Figure 1.5**). To test this theory, we performed preliminary steady-state experiments in deuterated buffer with deuterated NH_2OH (ND_2OD) added. The resulting KIE does indeed suggest that PCET events are an important part of the cyt P460 catalytic cycle and likely to be affected by the glutamate in the second coordination sphere. This would also explain why the corresponding Ala131Gln variant, which is structurally very similar to glutamate but unable to accept protons in the same

manner, is not able to oxidize NH_2OH, despite being able to make a Fe^{III}–NH_2OH intermediate (**Figure 4.5, Figure 4.9, Figure 4.10**). This shows the importance of this residue to NH_2OH oxidation and implies that likely this residue serves as a proton relay during catalysis. This hypothesis will be explored further in future experiments performed by the Lancaster research group.

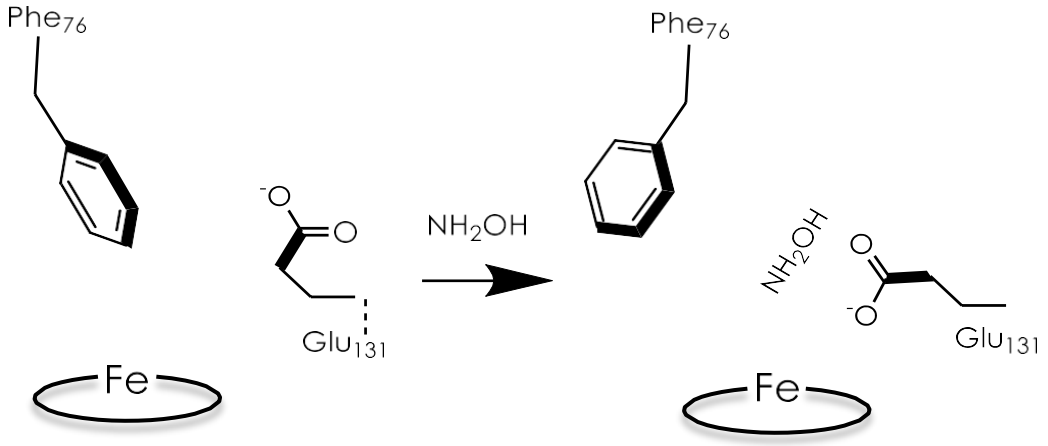
The necessity of the Glu residue in catalysis can be further supported by crystallographic evidence. We obtained crystal structures of Ala131Glu and Ala131Gln as isolated and Ala131Gln soaked with NH_2OH. The Fe–N bond distances for the Ala131Gln–NH_2OH are longer than anticipated (~2.5-2.7 Å instead of the anticipated distance closer to 2.0 Å), but we attribute this to the resolution of the structure, at which the e.s.d. in distances is estimated to be 0.4 Å. The composite omit map clearly suggests electron density reminiscent of a bound NH_2OH (**Figure 4.10**). Though we were unable to obtain quality data for a structure of Ala131Glu soaked with NH_2OH, we can still draw relevant insight from the comparison of the structures, as by EPR these intermediates are expected to be very similar (**Figure 4.7**). Comparing the active sites of the Ala131Glu/WT AL212 and Ala131Gln with NH_2OH bound, it is possible to see molecular motions which may help to accommodate substrate binding and enforce the involvement of residue 131. Specifically, with no substrate bound, Phe76 sits directly above the P460 cofactor. When NH_2OH binds, this Phe76 can reorient as is seen in the Ala131Gln structures. The residue in position 131 (Gln, in this case) can also then reposition to interact with the bound substrate (**Figure 4.9, Scheme 4.1**). This is reminiscent of a conserved, carboxylate-containing Asp residue in *N. europaea* HAO (NeHAO) as well as in the multi-heme, P460-containing anammox enzyme kustc1061

from *Kuennenia stuttgartiensis*.[16] In the NH_2OH-soaked structures of NeHAO and kustc1061, this Asp and a nearby His residue are shown to interact with the bound substrate. The authors suggested in fact that these residues likely participate in shuttling protons during catalysis. This further supports the notion that carboxylate-containing residues are important for heme P460 NH_2OH oxidation catalysis.

Interestingly, the resting Fe^{III} form of Ala131Gln also has the Phe76 residue moved to the side and Gln131 sidechain reoriented towards the heme. While it is unclear why the Gln and not the Glu variant prefers this orientation in the crystal structure, it could explain why a NH_2OH–bound Ala131Gln structure was more easily attainable, as these residues were already in an ideal conformation to accept substrate. From the EPR, it is unlikely that H_2O is bound in the Ala131Gln cyt P460 as it is also high spin, though the distribution of the two components in all AL212 variants seems to be related to the mutation, and thus may be related to the relative orientation preference of

---

Scheme 4.1 Proposed *N.* sp. AL212 cyt P460 molecular motions accompanying substrate binding



Phe76/residue131. That is, where WT AL212 is almost entirely component 1, which would correspond to Phe76 sitting directly over the Fe, Ala131Glu is more of a mixture of the two, and could suggest that the larger Glu residue can occasionally push the Phe76 out of the way and reorient towards the Fe center (component 2), and Ala131Gln is almost entirely component 2, as confirmed by the crystal structure. Confirmation of this will require more work, however, and likely involve mutation of the Phe76 residue.

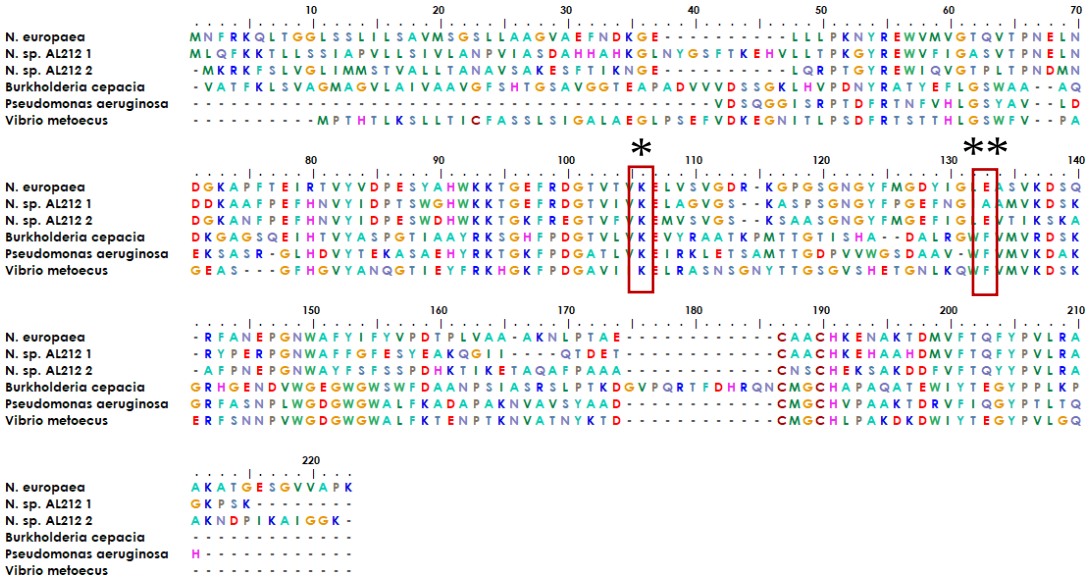
Despite our inability to obtain a structure of NH_2OH–bound Ala131Glu, we were able to get a structure with NO bound. Though crystals of Ala131Glu were soaked in an NO-donor (leading to an $\{FeNO\}^6$), the NO bond distances and angles (**Table 4.1**) in some chains agree more closely with an $\{FeNO\}^7$ configuration.[17-19] This suggests that perhaps the $\{FeNO\}^6$ unit was photoreduced upon data collection, as has been proposed for similar heme-nitrosyl crystal structures in the NO-sensing protein nitrophorin.[20, 21] Given that the bond lengths and angles are also not the same in all chains, it may suggest that more photoreduction occurred at one site than at the other, or that the resolution is not high enough to accurately describe the $\{FeNO\}^6$ intermediate. Additionally, the Fe–N distances are much shorter than anticipated. This is likely due to the resolution of the structure and not a reflection of the true bond distance. Still, this structure represents the first cyt P460-nitrosyl and can be used to further investigate intermediates in the cyt P460-catalyzed NH_2OH oxidation pathway. Of note, the Glu131 does not appear to be interacting with the bound NO, however this may not represent the conformation of the residues when the $\{FeNO\}^6$ is formed during NH_2OH oxidation instead of when exogenous NO is added.

The *Nitrosomonas* sp. AL212 genome contains two cyt P460 sequences (**Figure**

4.11), one that contains Ala131, which the present study is about, and one that is predicted to have a Glu in this position (though the latter also contains a CXXCH heme-binding motif different from the CAACH motif observed in the former AL212 cyt P460 and *N. europaea* cyt P460).[22] It is likely that this second cyt P460 behaves in a manner similar to *N. europaea* cyt P460 and is responsible for some of the N_2O produced by the enzyme, and likely serves to detoxify buildup of NH_2OH or NO. This other cyt P460 then must have some other function in *Nitrosomonas* sp. AL212. This may help explain the breadth of microorganisms other than AOB that have been shown to contain cyt

---

Figure 4.11 Sequence Homology Between Various Cyt P460 Genes



Sequence alignment generated using MEGA X software[23] showing homology between cyt P460 genes from *N. europaea*, the two cyt P460 sequences from *N.* sp. AL212, and genes predicted from mammalian pathogenic bacteria *Burkholderia cepacian*, *Pseudomonas aeruginosa*, and *Vibrio metoecus*. * indicates conserved Lys cross-link, **indicates the second-sphere Ala/Glu residue.

P460. We would benefit from understanding different regulatory patterns for each of the cyt P460s, as this may provide a clue as to when this other type of cyt P460 is operative, now that it is clear the two enzymes must serve different roles in the organism *Nitrosomonas* sp. AL212.

The observation of a cyt P460 that is unable to oxidize NH_2OH also allows for the possibility of different classes of cyt P460 that are tuned toward different substrate specificities. The hallmark of cyt P460 enzymes then may be the presence of the cross-link, whose role may partially be to alter the electronic structure of the active site such that the substrate is directly oxidized preferentially over the Fe. This would contrast with certain other Fe-heme containing enzymes that lack the cross-link and instead employ very reactive high valent Fe intermediates (e.g. compound I and compound II of P450[24]). Prior TD-DFT analysis indeed showed that π-donation from the Lys N raises the energy of orbitals associated with the P460 cofactor[4] (Chapter 3). Thus, it is possible that the higher energy heme P460 orbitals leads to increased NH_2OH character in the SOMO, which will interact with oxidant. In this manner, the characteristic heme P460 cross-link defines the oxidative capabilities of the cyt P460 enzyme, and the substrate preference/specificity is tuned by residues in second coordination sphere. This theory agrees quite well with the evolutionary trajectory of HAO into oxidative chemistry. HAO belongs to a larger family of multiheme cytochrome *c* containing enzymes, including hydrazine oxidoreductase (HZO), octaheme cytochrome *c* nitrite reductase (ONR), and pentaheme cytochrome *c* nitrite reductase (NrfA). It is believed that the addition of additional protein cross-links to the catalytic heme are what distinguished enzymes with oxidative chemistries from those with reductive chemistries.[25]

Comparative homology of currently annotated cyt P460 sequences would suggest distinct classes of cyt P460 which contain different residues (e.g. glutamate/aspartate, alanine, phenylalanine) in this crucial second sphere position. This may allude to distinct groups of cyt P460 which perform oxidative chemistry but are tuned for different substrates. It will be interesting in the future to consider the evolutionary origins of this reactivity, and may provide evidence for the function of cyt P460s which cannot naturally oxidize NH_2OH.

D. CONCLUSION

In addition to inner-sphere factors like the Lys cross-link,[5] cyt P460 NH_2OH activity is dependent on second-sphere factors including a nearby Glu residue. This Glu most likely functions as a proton-relay during catalysis, though the exact nature of this will need to be explored in future experiments involving kinetic isotope effects. This discovery also allows for the possibility of a broader range of cyt P460 catalysis other than NH_2OH oxidation and can help to explain the breadth of organisms other than AOB with genes for cyt P460s.

E. REFERENCES

(1) Elmore, B. O., Bergmann, D. J., Klotz, M. G., and Hooper, A. B. (2007) Cytochromes P460 and c'-beta; a new family of high-spin cytochromes c, *FEBS Lett. 581*, 911-916.

(2) Bergmann, D. J., Zahn, J. A., Hooper, A. B., and DiSpirito, A. A. (1998) Cytochrome P460 genes from the methanotroph Methylococcus capsulatus bath, *J. Bacteriol. 180*, 6440-6445.

(3) Cedervall, P., Hooper, A. B., and Wilmot, C. M. (2013) Structural studies of hydroxylamine oxidoreductase reveal a unique heme cofactor and a previously unidentified interaction partner, *Biochemistry 52*, 6211-6218.

(4) Smith, M. A., and Lancaster, K. M. (2018) The Eponymous Cofactors in Cytochrome P460s from Ammonia-Oxidizing Bacteria Are Iron Porphyrinoids Whose Macrocycles Are Dibasic, *Biochemistry 57*, 334-343.

(5) Vilbert, A. C., Caranto, J. D., and Lancaster, K. M. (2018) Influences of the heme-lysine crosslink in cytochrome P460 over redox catalysis and nitric oxide sensitivity, *Chem. Sci. 9*, 368-379.

(6) Caranto, J. D., and Lancaster, K. M. (2017) Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc Natl Acad Sci USA 114*, 8217-8222.

(7) Caranto, J. D., Vilbert, A. C., and Lancaster, K. M. (2016) Nitrosomonas europaea cytochrome P460 is a direct link between nitrification and nitrous oxide emission, *Proc Natl Acad Sci USA 113*, 14704-14709.

(8) Kabsch, W. (2010) Xds, *Acta Crystallogr. D Bio.l Crystallogr. 66*, 125-132.

(9) Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments, *Acta Crystallogr. D Biol. Crystallogr. 67*, 235-242.

(10) Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy,

A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D Biol. Crystallogr. 66*, 213-221.

(11) Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot, *Acta Crystallogr. D Biol. Crystallogr. 66*, 486-501.

(12) (2015) The PyMOL Molecular Graphics System, version 1.8, Schrodinger, LLC, Portland, OR.

(13) Frear, D. S., and Burrell, R. C. (1955) Spectrophotometric Method for Determining Hydroxylamine Reductase Activity in Higher Plants, *Anal. Chem. 27*, 1664-1665.

(14) Williamson, G., and Engel, P. C. (1984) Butyryl-CoA dehydrogenase from Megasphaera elsdenii. Specificity of the catalytic reaction, *Biochem. J. 218*, 521-529.

(15) Golombek, A. P., and Hendrich, M. P. (2003) Quantitative analysis of dinuclear manganese(II) EPR spectra, *J. Magn. Reson. 165*, 33-48.

(16) Maalcke, W. J., Dietl, A., Marritt, S. J., Butt, J. N., Jetten, M. S. M., Keltjens, J. T., Barends, T. R. M., and Kartal, B. (2014) Structural Basis of Biological NO Generation by Octaheme Oxidoreductases, *J. Biol. Chem. 289*, 1228-1242.

(17) Goodrich, L. E., Paulat, F., Praneeth, V. K., and Lehnert, N. (2010) Electronic structure of heme-nitrosyls and its significance for nitric oxide reactivity, sensing, transport, and toxicity in biological systems, *Inorg. Chem. 49*, 6293-6316.

(18) Lehnert, N., Scheidt, W. R., and Wolf, M. W. (2014) Structure and Bonding in Heme-Nitrosyl Complexes and Implications for Biology, *Struct. Bond. 154*, 155-223.

(19) Scheidt, W. R., and Ellison, M. K. (1999) The Synthetic and Structural Chemistry of Heme Derivatives with Nitric Oxide Ligands, *Acc. Chem. Res. 32*, 350-359.

(20) Weichsel, A., Andersen, J. F., Roberts, S. A., and Montfort, W. R. (2000) Nitric oxide binding to nitrophorin 4 induces complete distal pocket burial, *Nat. Struct. Biol. 7*, 551-554.

(21) Walker, F. A. (2005) Nitric oxide interaction with insect nitrophorins and thoughts on the electron configuration of the $\{FeNO\}^6$ complex, *J. Inorg. Biochem. 99*, 216-236.

(22) Suwa, Y., Norton, J. M., Bollmann, A., Klotz, M. G., Stein, L. Y., Laanbroek, H. J., Arp, D. J., Goodwin, L. A., Chertkov, O., Held, B., Bruce, D., Detter, J. C., Detter, J. C., Tapia, R., and Han, C. S. (2011) Genome sequence of Nitrosomonas sp. strain AL212, an ammonia-oxidizing bacterium sensitive to high levels of ammonia, *J. Bacteriol. 193*, 5047-5048.

(23) Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018) MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms, *Mol. Biol. Evol. 35*, 1547-1549.

(24) Sligar, S. G., Makris, T. M., and Denisov, I. G. (2005) Thirty years of microbial P450 monooxygenase research: peroxo-heme intermediates--the central bus station in heme oxygenase catalysis, *Biochem. Biophys. Res. Commun. 338*, 346-354.

(25) Klotz, M. G., Schmid, M. C., Strous, M., den Camp, H. J. M. O., Jetten, M. S. M., and Hooper, A. B. (2008) Evolution of an octahaem cytochrome c protein family that is key to aerobic and anaerobic ammonia oxidation by bacteria, *Environ. Microbiol. 10*, 3150-3163.

CHAPTER 5: INTERACTIONS OF THE RED COPPER PROTEIN NITROSOCYANIN WITH NITRIC OXIDE

Conventionally, bacterial nitrification was thought to be composed of two steps mediated by the enzymes AMO (**Eqn. 5.1**) and HAO (**Eqn. 5.2**):

$$NH_3 + O_2 + 2e^- \rightarrow NH_2OH \quad (\mathbf{5.1})$$

$$NH_2OH + H_2O \rightarrow NO_2^- + 5H^+ + 4e^- \quad (\mathbf{5.2})$$

It was recently shown, however, that in fact the true enzymatic product of HAO catalysis is NO, and not NO_2^-.[1] Thus, the correct second step of nitrification is as follows:

$$NH_2OH \rightarrow NO + 3H^+ + 3e^- \quad (\mathbf{5.3})$$

It is known that AOB produce a stoichiometric amount of NO_2^- from NH_3, and stable isotope studies have shown that one O in the NO_2^- produced comes from H_2O and not O_2.[2] Given that the other O is known to originate from O_2[3] (**Eqn. 1**), these observations suggest that the final conversion of NH_2OH to NO_2^- is performed by a currently unknown NO oxidase capable of catalyzing the final reaction in the nitrification pathway:

$$NO + H_2O \rightarrow NO_2^- + 2H^+ + 1e^- \quad (\mathbf{5.4})$$

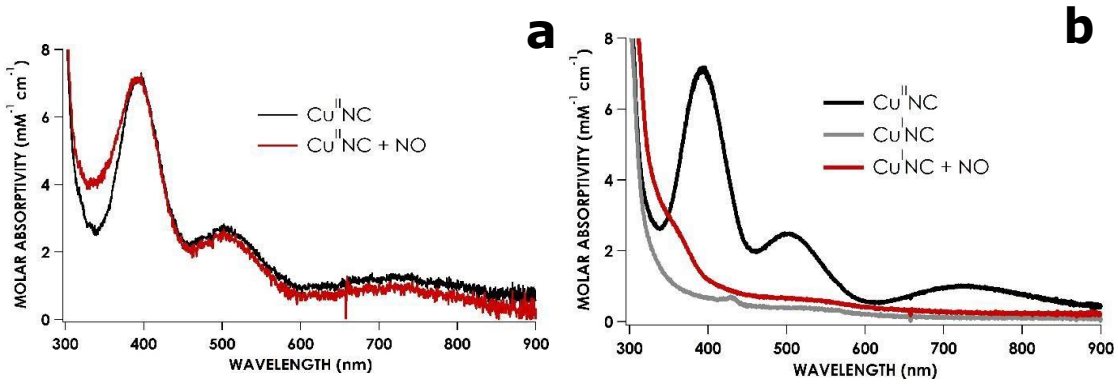
Mounting evidence has suggested a crucial role for the unusual protein nitrosocyanin (NC) in AOB.[4] NC exhibits primary and tertiary sequence homology to blue copper "cupredoxins," though the spectral properties and active site structure are distinct. [5-7] Canonical blue copper (or type 1) sites, such as those found in azurin and plastocyanin contain two histidines, one cysteine, and one axially coordinated methionine in a trigonally distorted tetrahedral geometry.[8, 9] These centers typically exhibit intense absorption features around 600 nm due to S(Cys) $\pi \rightarrow$ Cu $3d_{x2-y2}$ ligand-to-metal charge transfer (LMCT). Additionally, the electron paramagnetic

resonance (EPR) spectra for type 1 blue copper sites exhibit small $A_{\parallel}$ (< 70×10^{-4} cm^{-1}), differing from "normal" or type 2 copper sites and inorganic copper compounds where these values tend to be in excess of 150×10^{-4} cm^{-1}.[10] Blue copper proteins are involved in electron transfer processes, and have reduction potentials in the range of +184 to +680 mV.[11] Their efficacy in electron transfer is attributable to low inner-sphere reorganization energies imposed by the high degree of Cys–Cu covalency and second sphere constraints over the active site geometry.

In contrast to blue copper sites, the active site of NC is coordinated by two histidines, a cysteine, a glutamate, and a water molecule that is lost upon reduction from Cu^{II} to Cu^{I}.[5] This altered geometry gives rise to an absorption feature at 390 nm, as a result of S(Cys) σ to Cu $3d_{x2-y2}$ LMCT (**Figure 5.1**). Moreover, the EPR spectrum shows $A_{\parallel}$ ca. 140×10^{-4} cm^{-1} (**Figure 5.1**) and the reduction potential at pH 7.0 is +85 mV vs. SHE.[6] The presence of a solvent molecule that is bound in the oxidized form of the protein and lost upon reduction also indicates that NC would likely be a poor electron transfer protein due to increased inner-sphere reorganization energy. Taken together, these data suggest a role for NC other than electron transfer.

Despite its peculiarities, the role of NC remains unknown. Given the discrepancies between NC and blue copper proteins used for electron transfer, it is possible that NC has some catalytic role, possibly relevant to nitrification. Indeed, NC is only found in AOB; notably, AOB that lack the gene for NC, including *Nitrosomonas* sp. Is79A3, produce large amounts of NO during NH_3-oxidation. NC transcripts are present in high amounts,[4] similar to those of other metabolic enzymes AMO and HAO (**Table 5.1**). It has also been shown elsewhere that NC is upregulated

Figure 5.1 UV-vis Absorption Spectra of $Cu^{II}NC$ and $Cu^{I}NC$ before and after Addition of NO



UV-vis features of oxidized (a) and reduced (b) NC treated with NO-donor PROLI-NONOate. The features of the oxidized protein match those reported in literature,[6] with bands appearing at 390 nm, 500 nm, and 720 nm. When reduced, these features disappear. A new feature around 345 nm appears, but this is partially due to the presence of PROLI-NONOate which also absorbs around this region.

during NH_3-oxidation or when cells are exposed to NO.[12-14] In the AOB *Nitrosomonas eutropha*, the NC gene is headed by a fumarate-nitrate reduction (FNR) protein binding region in the promotor sequence.[14, 15] FNR is a regulatory protein that has been implicated in response to low oxygen or the presence of NO.[16, 17] Likewise, a similarly unusual purple copper protein from the NH_3-oxidizing archaea (AOA) *Nitrosopumilus maritimus* known as Nmar_1307, which also coordinates a solvent molecule in the active site has been shown to interact with NO and produce NO_2^-.[18] In addition, a mutated azurin variant designed to mimic the NC red copper site also reacted with NO to form an S-nitrosylated product.[19] Thus, it is clear that NC has a

Table 5.1 RNA Transcripts of AOB Taken from Environmental Samples and Cultured in Laboratory Fermenters

Read Rank	Name	RNA Reads
1	amoC	68104
2	amoB	22074
4	amoA	15721
9	cco protein	6306
21	cyt c551/c552	2464
22	HAO	2401
32	Nitrosocyanin	1734
56	amoC	940
67	cyt c554	811
87	amoC	630
88	cyt cm552	616
104	cyt c554	549
153	NorB	391
268	cyt P460	247
428	NorC	171
502	HaoB	146

*Data obtained by Prof. James P. Shapleigh (Cornell University), Wang Lan, and Liangwei Deng (Biogas Institute of Ministry of Agriculture, BIOMA)

crucial role in AOB that is likely related to NO. As a result, we investigated it as a potential candidate for the missing NO oxidase.

A. MATERIALS AND METHODS

General Considerations

Milli-Q water (18.2 MΩ; Millipore) was used in the preparation of all buffers and solutions. Disodium 1-(Hydroxyl-NNO-azoxy)-L-proline (PROLI-NONOate), used as a source of NO, was purchased from Cayman Chemicals. All other chemicals were purchased from VWR International and used as obtained. A stock solution of PROLI-

NONOate was prepared by dissolving 10 mg of PROLI-NONOate in 0.01 M NaOH. NO release from PROLI-NONOate was quantified following the method used by Hayashi et al.,[20] where $Na_2[Fe^{II}(EDTA)]$ was added to a solution containing diluted PROLI-NONOate stock to form an Fe(EDTA)–NO complex ($\varepsilon 440\ nm = 900\ M^{-1}\ cm^{-1}$; EDTA = ethylenediaminetetraacetic acid). Buffers were deoxygenated by bubbling with N2 for at least 30 min. All reactions were prepared in an anaerobic chamber (Coy) with deoxygenated buffers unless otherwise noted. UV– visible (UV–vis) absorption spectra measurements were performed with a Cary 60 UV-vis spectrometer. Data were fit using Igor Pro version 6.37 (WaveMetrics). Reduced (Cu^I) NC was prepared by titrating protein with sodium dithionite; excess dithionite was washed from solution using Amicon Ultra Centrifugal Filters (MIlliporeSigma). Protein activity was assessed by treating oxidized or reduced NC with NO from PROLI-NONOate and various oxidants including potassium ferricyanide ($K_3[Fe(CN)_6]$), 2,6-dichlorophenolindophenol (DCPIP), hexaammineruthenium(III) chloride ($Cl_3[Ru(NH_3)_6]$), azurin (grown in lab) and rusticyanin (provided by Prof. Jeffrey Warren). Concentration of NO_2^- in solution was determined using the Griess diazotization assay[21] (reagents purchased through Cayman Chemical).

Protein Expression

The NC coding sequence was codon optimized and cloned into the pET22b (+) vector. This plasmid was transformed into Lemo21(DE3) competent *E.coli* with selection on Luria-Bertani (LB) plates supplemented with 100 $\mu g\ ml^{-1}$ ampicillin and 34 $\mu g\ ml^{-1}$ chloroamphenicol. Single colonies were used to inoculate 3 mL LB starter cultures and were incubated with shaking for 6-8 h at 37 °C. Starter cultures were

pelleted and resuspended in 1 mL fresh LB media before being used to inoculate three 4 L culture flasks containing 2 L Terrific Broth media containing 0.5% glycerol (0.1% inoculum). The liter-volume cultures were grown at 30 °C overnight. Expression was induced by addition of 0.4 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and the temperature was brought down to 25 °C. 1 mM $CuSO_4$ was added at the time of induction, as well as 8–10 hrs post-induction. Cells were harvested by centrifugation 24 h post induction.

Osmotic Shock

Final cell pellets were thoroughly resuspended in a 20% sucrose solution buffered at pH 8.1 with 300 mM Tris and 1 mM EDTA. The volume was raised to ca. 400 mL divided evenly between two centrifuge bottes. Cells were permitted to osmotically equilibrate in this solution for 45 minutes at room temp, then the suspensions were spun for 20 minutes at 7500 × g. The supernatant was decanted and pellets were partially resuspended in the residual solution. Bottles were then transferred to an ice bucket. 50 mL of ice-cold MilliQ H_2O containing 500 μM $MgCl_2$ were then added quickly to each bottle and the pellets were thoroughly resuspended. The suspensions were transferred to one 250 mL beaker containing a magnetic stir bar, covered in parafilm, and allowed to stir gently at 4 °C for 14 minutes. The suspension was then spun at 20,000 × g in Oak Ridge tubes for 30 minutes. Supernatant was decanted and preserved.

Protein Purification

Periplasmic fraction obtained from osmotic shock was treated with 40% (m/v) of $(NH_4)_2SO_4$, causing formation of a precipitate. The precipitate was pelleted by

centrifugation at 30,000 × g in Oak Ridge tubes for 30 minutes. The supernatant was collected and the $(NH_4)_2SO_4$ percentage brought up to 50% (m/v) resulting in further precipitation. The precipitate was again pelleted by centrifugation at 30,000 × g in Oak Ridge tubes for 30 minutes, resulting in a pink precipitate and a clear supernatant. This second supernatant was discarded, and the pellet dissolved in 5-10 mL 20 mM Tris, pH 8.0 and concentrated in an Amicon centrifugal filter with molecular weight cutoff of 3 kDa (MilliporeSigma). This concentrated solution was diluted to 10X the starting volume with 25 mM Na_2HPO_4, pH 7.5 with 2 M $(NH_4)_2SO_4$ and concentrated again. This process was repeated a second time to ensure buffer exchange. This solution was loaded on to a 20-mL octyl sepharose column (GE Healthcare Life Sciences). A 20-column volume linear gradient of 25 mM Na_2HPO_4, pH 7.5 from 2 to 0 M $(NH_4)_2SO_4$ was applied to the column. Red fractions eluted around 0.1 to 0 M $(NH_4)_2SO_4$ and were pure by SDS-PAGE. These fractions were pooled and concentrated and buffer exchanged by 3X dilution/concentration cycles with 50 mM Na_2HPO_4, pH 8.0 in Amicon centrifugal filters. To remove adventitiously bound Cu, EDTA was added to concentrated solutions to a final concentration of 10 mM EDTA and incubated at 4 °C overnight. A PD10 column equilibrated with 50 mM Na_2HPO_4, pH 8.0 was used to remove excess Cu-EDTA. Concentrated protein was stored at 4°C aerobically or at room temperature anaerobically.

NO-selective Electrode

All reactions were prepared and sealed in 5 mL headspace gas chromatography (GC) vials (Wheaton). The concentration of formed NO was analyzed with a NO microsensor housed within a septum-piercing needle (Unisense).

Electron Paramagnetic Resonance (EPR)

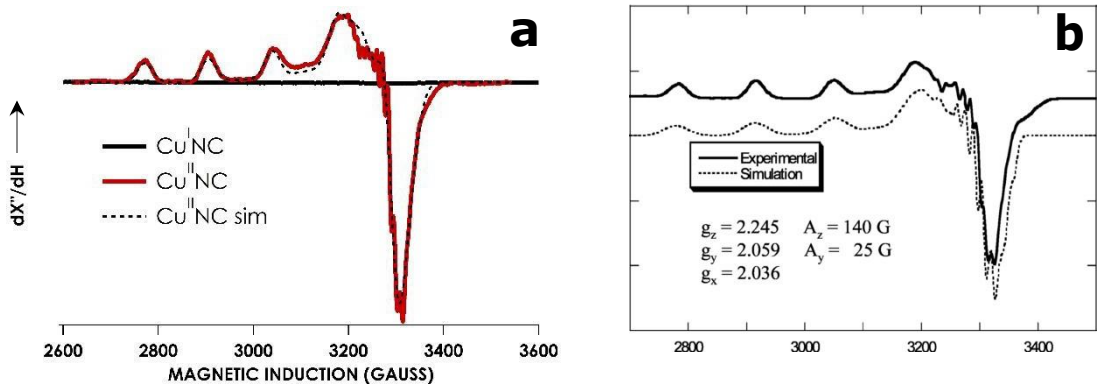
The protein concentration used was 175 μM in 50 mM sodium phosphate (pH 8.0) buffer with 25% (v/v) glycerol. The EPR spectra were recorded at X-band (9.40 GHz) on a Bruker Elexsys-II spectrometer equipped with a liquid He cryostat maintained at 10 or 77 K. EPR data were simulated, and spin concentrations were determined by using SpinCount.[22]

B. RESULTS

NC was expressed recombinantly in *E. coli* cells. Correct expression was confirmed by mass spectrometry (**Appendix A.I**). This recombinant protein exhibited UV-vis and EPR features (**Figure 5.1, Figure 5.2**) that match those reported for NC isolated from the native organism *Nitrosomonas europaea*.[6] NO was initially added to both the oxidized ($Cu^{II}NC$) and reduced ($Cu^{I}NC$) protein to see if any spectral changes occurred. No changes were observed for $Cu^{II}NC$, however, a shoulder feature at around 350 nm was observed when NO was added to $Cu^{I}NC$ (**Figure 5.1**), though this feature is in part due to the presence of PROLI-NONOate. No new EPR feature was observed for this species, and the spectrum obtained looked identical to that of $Cu^{I}NC$ (e.g. no signal). Additionally, when monitoring the NO in solution via a NO-sensitive electrode, it is clear that NO binds to the $Cu^{I}NC$, but $Cu^{II}NC$ has no effect on NO in solution (**Figure 5.3**).

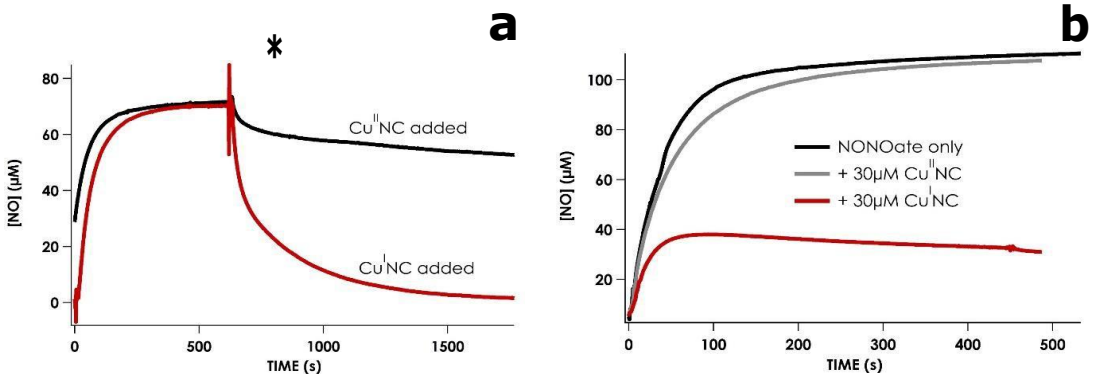
$Cu^{II}NC$ added to solution with a source of NO and the oxidant potassium ferricyanide ($K_3[Fe(CN)_6]$), showed a modest increase (approximately 6%, going from 0.249 ± 0.15 μM $K_3[Fe(CN)_6] \cdot min^{-1}$ to 1.67 ± 0.016 μM $K_3[Fe(CN)_6] \cdot min^{-1}$) in oxidant consumption upon addition of 10 μM $Cu^{II}NC$ (**Figure 5.4a**). However, as

Figure 5.2 EPR Spectra of Recombinantly Expressed and Native NC



(a) EPR spectra of recombinantly expressed oxidized and reduced NC. Simulated g_{eff} values for oxidized NC are 2.029, 2.062, and 2.245. (b) EPR spectrum of oxidized NC isolated from native organism. Reprinted with permission from Basumallick, L., et al. (2005). Spectroscopic and density functional studies of the red copper site in nitrosocyanin: Role of the protein in determining active site geometric and electronic structure. *Journal of the American Chemical Society*, *127*(10), 3531–3544. © 2005 American Chemical Society.

Figure 5.3 NO Concentration with either Cu^{II} or Cu^{I} NC Present



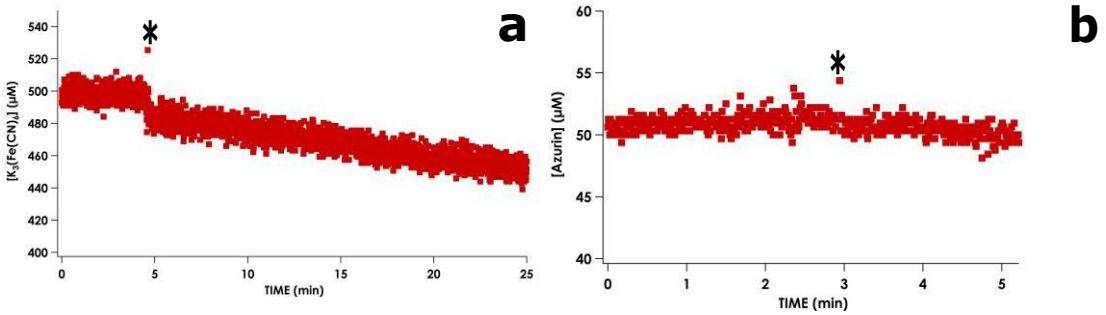
(a) Measured concentration of NO where oxidized or reduced NC was added to solutions of PROLI-NONOate which had been allowed to pre-decay (addition at *). (b) Measured concentration of NO released from PROLI-NONOate decay when added to solutions already containing either oxidized or reduced NC.

$K_3[Fe(CN)_6]$ will also react with NO to produce NO_2^-, it was not possible to distinguish NC-derived NO_2^- production from background reactivity. Increased oxidant consumption was not observed with Cu^{II}NC and azurin or rusticyanin used as the oxidant (**Figure 5.4b**).

Design of experiments in which Cu^{I}NC, oxidant, and NO are mixed under turnover conditions are challenging as there is a tendency for many oxidants to re-oxidize Cu^{I}NC; however, it appears that additional oxidation events are occurring (**Figure 5.5**). Following the UV-vis absorption features for a reaction of Cu^{I}NC, NO from PROLI-NONOate, and the $2e^-$ oxidant 2,6-dichlorophenolindophenol (DCPIP), one can see the return of Cu^{II}NC features rapidly reappear at 390 nm, concomitant with DCPIP consumption matching ½ the concentration of NC originally in solution (given that DCPIP is a $2e^-$ oxidant, this corresponds to the observed NC oxidation). DCPIP is rapidly consumed in this phase, with k_{obs} of 2.72 ± 0.14 μM DCPIP · μM^{-1} Cu^{I}NC · min^{-1}. Following this phase, however, DCPIP continues to be consumed at a much slower rate (initial rate of 0.125 ± 0.011 μM DCPIP · μM^{-1} Cu^{I}NC · min^{-1}), with no new NC features appearing. This does not occur in reactions where no NC is present.

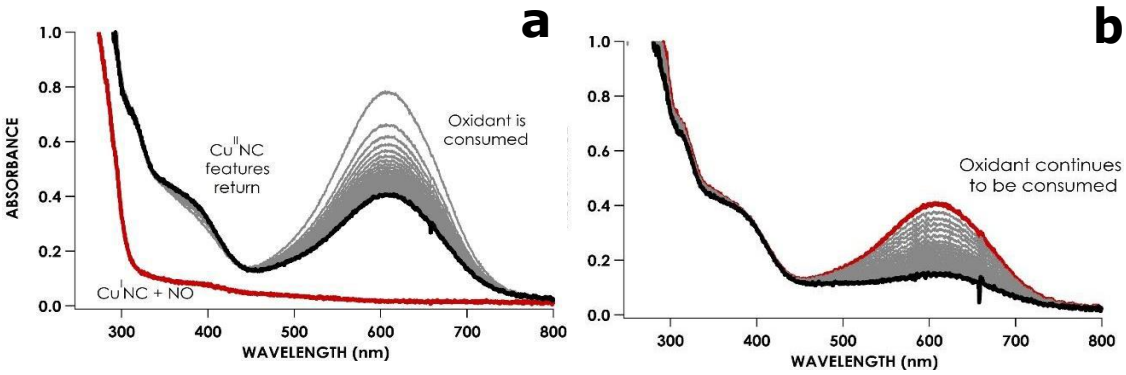
Detection of NO_2^- in this system has remained complicated. When determining the concentration of the NO_2^- in solution using the Griess diazotization assay, control samples containing no NC always have a higher concentration of NO_2^-. This is attributable to contamination of the PROLI-NONOate. This is specifically a problem with Cu^{I}NC. Addition of DCPIP to a solution of Cu^{I}NC and NO_2^- results in oxidation of Cu^{I}NC and release of bound NO_2^- as measured by the Griess assay (**Figure 5.6a**).

Figure 5.4 $Cu^{II}NC$ Interaction with NO and Oxidants Ferricyanide and Azurin



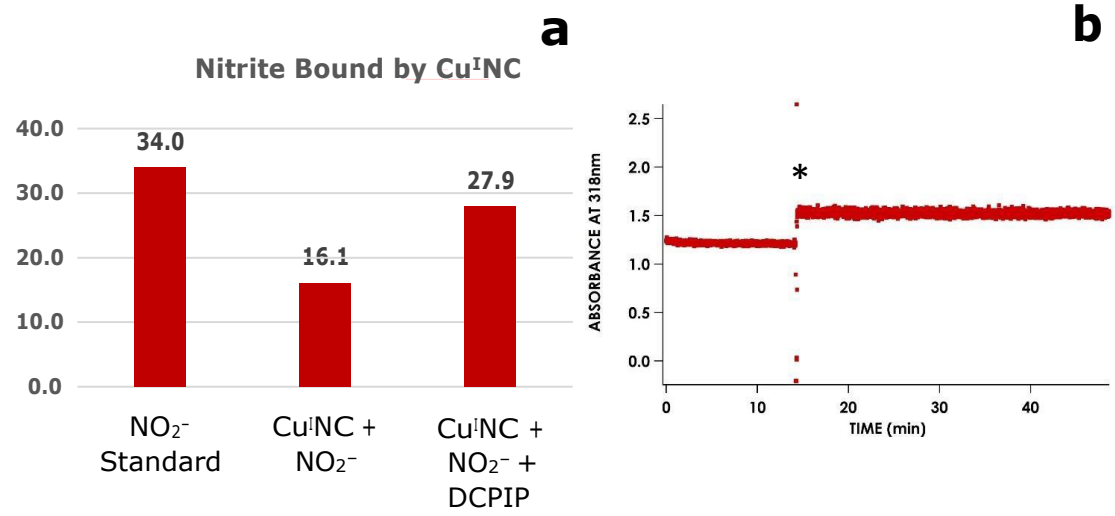
Oxidant consumption assays following addition of 10 μM $Cu^{II}NC$ to anaerobic solutions oxidant in buffer with ~ 1 atm NO gas applied to septum-sealed cuvette. Oxidant trials shown here are ferricyanide (a) and azurin (b). Ferricyanide consumption was monitored by UV-vis absorption at 424 nm, while azurin consumption was monitored at 625 nm. *Indicates addition of $Cu^{II}NC$.

Figure 5.5 Reaction of $Cu^{I}NC$, NO, and DCPIP.



The reaction was initiated with the addition of the DCPIP. (a) Initial phase of the reaction in which features consistent with $Cu^{II}NC$ rapidly return and necessary oxidant (DCPIP, the oxidized form of which has a UV-vis feature at 605 nm) is consumed. UV-vis scans were taken every 3 seconds. (b) Second phase of the reaction showing continued oxidant consumption with no observed changes in NC features. UV-vis scans were taken every 5 minutes. Griess assay results showed less NO_2^- in this sample than in the controls containing no NC.

Figure 5.6 $Cu^{I}NC$ Interaction with NO_2^-



(a) NO_2^- concentration determined by Griess diazotization assay for a sample containing only buffer and 35 µM NO_2^-, a sample containing 35 µM NO_2^- and 25 µM $Cu^{I}NC$, and a final sample containing 35 µM NO2–,25 µM $Cu^{I}NC$, and 25 µM of the oxidant 2,6-dichlorophenolindophenol (DCPIP). (b) Sodium dithionite consumption monitored by UV-vis of dithionite peak at 318 nm. Sample contained 150 µM sodium dithionite and 200 µM NO_2^- in 50 mM sodium phosphate buffer, pH 8.0. 20 µM $Cu^{I}NC$ was added where indicated (*). This was performed in an anaerobic environment.

After making this observation, we explored the possibility that NC may function bidirectionally and reduce NO_2^- as well as oxidize NO, as the Cu-containing nitrite reductase NirK has been shown to do.[23] However, no reductant (sodium dithionite, methyl viologen, or ascorbate) was consumed in the presence of $Cu^{I}NC$ and nitrite (**Figure 5.6b**).

C. DISCUSSION

Given the high expression of NC in AOB, both shown here (**Table 5.1**) and elsewhere,[4, 12-15] it follows that NC should have a crucial role in AOB. However, the

unusual nature of its active site provides few clues to the nature of its activity. We have been able to show here that Cu^{I}NC does indeed interact with NO (**Figure 5.3**) and appears to effect some kind of redox process (**Figure 5.5**) but have thus far been unable to detect a product from this reaction. It is possible that these difficulties are due to the apparent interaction of Cu^{I}NC with NO_2^-, though while the oxidant DCPIP was able to recover any bound NO_2^- applied exogenously (**Figure 5.6**), it does not appear to have this same effect during turnover conditions with Cu^{I}NC, NO, and excess DCPIP. In the former experiment, Cu^{I}NC can bind NO_2^- but once it is oxidized to Cu^{II}NC by DCPIP, releases it. Ideally in the case of the turnover conditions the DCPIP could both aid in the oxidation of NO to NO_2^- and the release of any bound NO_2^-. However, from the UV-vis of the turnover conditions (**Figure 5.5**), the Cu^{II}NC features appear immediately, so likely this scenario is not the case. Instead, it appears that the first step is the oxidation of the Cu in NC, to form some Cu^{II}NC–X species from Cu^{I}NC–NO. It is then unclear if this X species is released, leaving the Cu^{II}NC and an X species in solution that can react with DCPIP in the second stage, or if Cu^{II}NC–X reacts with DCPIP. In either case, no NO_2^- can be detected, so it is unclear what the products in either case would be. If the catalytic activity of NC truly is to oxidize NO to NO_2^-, then likely the problem is merely in the choice of using DCPIP as oxidant. As the Cu^{I}NC and not Cu^{II}NC is the form that clearly interacts with NO, true turnover conditions should ultimately return the enzyme to the resting Cu^{I}NC form. However, it also seems that the only way to release bound NO_2^- from Cu^{I}NC is to oxidize it. In fact, the same experiment was repeated with exogenous NO_2^- added to

Cu^{I}NC and hexaaminеruthenium(III) chloride ($Cl_3[Ru(NH_3)_6]$) as oxidant was unable to oxidize the NC and no bound NO_2^- was recovered.

There exist other discrepancies in the available data as well. For instance, if NO were bound to Cu^{I}, we would expect a $\{CuNO\}^{11}$ species that should have a distinctive EPR signal;[24, 25] however, assuming we are able to capture a NO–bound Cu^{I}NC, this species is EPR silent at temperatures from 77 to 10 K, 9.40 GHz, at a microwave power of 633 μW. Similarly, according to the concentration provided by the NO–selective electrode, it would appear as though two molecules of NO bind to every Cu center in NC. This would explain the absence of an EPR signal but would certainly be an unusual species in terms of Cu–based protein biochemistry. If we are in fact forming a Cu dinitrosyl, NC may instead serve as a kind of NO relay in order to keep NO build-up from becoming toxic, or to traffic NO to the actual NO oxidase. Of course, another possibility is that the apparent interaction with NO and NC is only coincidental, as there are other Cu–containing enzymes that also appear to bind NO.[26]

D. CONCLUSION

NC is likely an important enzyme unique to AOB biochemistry, yet its function remains elusive. Though possible that this protein serves a role in AOB N–metabolism, we are not yet able to definitively conclude that this protein functions as the missing NO oxidase. It is possible that if this is the true role of NC in AOB, that we were unable to reproduce the proper turnover conditions to effectively monitor NO_2^- production. However, it is also possible that NC does not serve as the NO oxidase, but instead as a NO–binding protein, or that it has a different role entirely. Future work on this protein should focus on exploring more oxidant choices, or if

possible, identifying any redox partners *in vivo*. Though the structural characteristics of NC allude to it being a poor electron transfer (ET) protein, because of it's homology with other ET proteins, it would also be informative to obtain and compare the electron self-exchange rate of NC to type 1 Cu proteins such as azurin or plastocyanin.

E. REFERENCES

(1) Caranto, J. D., and Lancaster, K. M. (2017) Nitric oxide is an obligate bacterial nitrification intermediate produced by hydroxylamine oxidoreductase, *Proc Natl Acad Sci USA 114*, 8217-8222.

(2) Andersson, K. K., and Hooper, A. B. (1983) O2and H2O are each the source of one O in NO−2produced from NH3byNitrosomonas:15N-NMR evidence, *FEBS Lett. 164*, 236-240.

(3) Hollocher, T. C., Tate, M. E., and Nicholas, D. J. D. (1981) Oxidation of ammonia by Nitrosomonas europaea. Definite 18O-tracer evidence that hydroxylamine formation involves a monooxygenase, *J. Biol. Chem. 256*, 10834-10836.

(4) Zorz, J. K., Kozlowski, J. A., Stein, L. Y., Strous, M., and Kleiner, M. (2018) Comparative Proteomics of Three Species of Ammonia-Oxidizing Bacteria, *Front. Microbiol. 9*, 938.

(5) Lieberman, R. L., Arciero, D. M., Hooper, A. B., and Rosenzweig, A. C. (2001) Crystal structure of a novel red copper protein from Nitrosomonas europaea, *Biochemistry 40*, 5674-5681.

(6) Arciero, D. M., Pierce, B. S., Hendrich, M. P., and Hooper, A. B. (2002) Nitrosocyanin, a red cupredoxin-like protein from Nitrosomonas europaea, *Biochemistry 41*, 1703-1709.

(7) Basumallick, L., Sarangi, R., DeBeer George, S., Elmore, B., Hooper, A. B., Hedman, B., Hodgson, K. O., and Solomon, E. I. (2005) Spectroscopic and density functional studies of the red copper site in nitrosocyanin: role of the protein in determining active site geometric and electronic structure, *J. Am. Chem. Soc. 127*, 3531-3544.

(8) Adman, E. T., and Jensen, L. H. (1981) Structural Features of Azurin at 2.7 Å Resolution, *Isr. J. Chem. 21*, 8-12.

(9) Colman, P. M., Freeman, H. C., Guss, J. M., Murata, M., Norris, V. A., Ramshaw, J. A. M., and Venkatappa, M. P. (1978) X-ray crystal structure analysis of plastocyanin at 2.7 Å resolution, *Nature 272*, 319-324.

(10) Solomon, E. I., Penfield, K. W., and Wilcox, D. E. (1983) Active sites in copper proteins. An electronic structure overview., *Struct. Bond. 53*, 1-57.

(11) Gray, H. B., Malmström, B. G., and Williams, R. J. P. (2000) Copper coordination in blue proteins, *JBIC 5*, 551-559.

(12) Kartal, B., Wessels, H. J., van der Biezen, E., Francoijs, K. J., Jetten, M. S., Klotz, M. G., and Stein, L. Y. (2012) Effects of nitrogen dioxide and anoxia on global gene and protein expression in long-term continuous cultures of Nitrosomonas eutropha C91, *Appl. Environ. Microbiol. 78*, 4788-4794.

(13) Stein, L. Y., Campbell, M. A., and Klotz, M. G. (2013) Energy-mediated vs. ammonium-regulated gene expression in the obligate ammonia-oxidizing bacterium, Nitrosococcus oceani, *Front. Microbiol. 4*, 277.

(14) Schmidt, I., Steenbakkers, P. J., op den Camp, H. J., Schmidt, K., and Jetten, M. S. (2004) Physiologic and proteomic evidence for a role of nitric oxide in biofilm formation by Nitrosomonas europaea and other ammonia oxidizers, *J. Bacteriol. 186*, 2781-2788.

(15) Klotz, M. G., and Stein, L. Y. (2008) Nitrifier genomics and evolution of the nitrogen cycle, *FEMS Microbiol. Lett. 278*, 146-156.

(16) Vanspanning, R. J. M., Deboer, A. P. N., Reijnders, W. N. M., Spiro, S., Westerhoff, H. V., Stouthamer, A. H., and Vanderoost, J. (1995) Nitrite and Nitric-Oxide Reduction in Paracoccus-Denitrificans Is under the Control of Nnr, a Regulatory Protein That Belongs to the Fnr Family of Transcriptional Activators, *FEBS Lett. 360*, 151-154.

(17) Cruz-Ramos, H., Crack, J., Wu, G. G., Hughes, M. N., Scott, C., Thomson, A. J., Green, J., and Poole, R. K. (2002) NO sensing by FNR: regulation of the Escherichia coli NO-detoxifying flavohaemoglobin, Hmp, *Embo. J. 21*, 3235-3244.

(18) Hosseinzadeh, P., Tian, S., Marshall, N. M., Hemp, J., Mullen, T., Nilges, M. J., Gao, Y. G., Robinson, H., Stahl, D. A., Gennis, R. B., and Lu, Y. (2016) A Purple Cupredoxin from Nitrosopumilus maritimus Containing a Mononuclear Type 1 Copper Center with an Open Binding Site, *J. Am. Chem. Soc. 138*, 6324-6327.

(19) Tian, S., Liu, J., Cowley, R. E., Hosseinzadeh, P., Marshall, N. M., Yu, Y., Robinson, H., Nilges, M. J., Blackburn, N. J., Solomon, E. I., and Lu, Y. (2016) Reversible S-nitrosylation in an engineered azurin, *Nat. Chem. 8*, 670-677.

(20) Hayashi, T., Caranto, J. D., Wampler, D. A., Kurtz, D. M., and Moenne-Loccoz, P. (2010) Insights into the Nitric Oxide Reductase Mechanism of Flavodiiron Proteins from a Flavin-Free Enzyme, *Biochemistry 49*, 7040-7049.

(21) Green, L. C., Wagner, D. A., Glogowski, J., Skipper, P. L., Wishnok, J. S., and Tannenbaum, S. R. (1982) Analysis of nitrate, nitrite, and [15N]nitrate in biological fluids, *Anal. Biochem. 126*, 131-138.

(22) Golombek, A. P., and Hendrich, M. P. (2003) Quantitative analysis of dinuclear manganese(II) EPR spectra, *J. Magn. Reson. 165*, 33-48.

(23) Wijma, H. J., Canters, G. W., de Vries, S., and Verbeet, M. P. (2004) Bidirectional catalysis by copper-containing nitrite reductase, *Biochemistry 43*, 10467-10474.

(24) Ruggiero, C. E., Carrier, S. M., Antholine, W. E., Whittaker, J. W., Cramer, C. J., and Tolman, W. B. (1993) Synthesis and Structural and Spectroscopic Characterization of Mononuclear Copper Nitrosyl Complexes - Models for Nitric-Oxide Adducts of Copper Proteins and Copper-Exchanged Zeolites, *J. Am. Chem. Soc. 115*, 11285-11298.

(25) Usov, O. M., Sun, Y., Grigoryants, V. M., Shapleigh, J. P., and Scholes, C. P. (2006) EPR-ENDOR of the Cu(I)NO complex of nitrite reductase, *J. Am. Chem. Soc. 128*, 13102-13111.

(26) Gorren, A. C., de Boer, E., and Wever, R. (1987) The reaction of nitric oxide with copper proteins and the photodissociation of copper-NO complexes, *Biochim. Biophys. Acta 916*, 38-47.

APPENDIX A: MOLECULAR BIOLOGY

A. pMycoFos Sequence

TCGATCTGATTCTGTGGATAACCGTATTACCGCCTTTGAGTGAGCTGATACCGCTCGCCG
CAGCCGAACGACCGAGCGCAACGCGTGAGCCCACCAGCTCCGTAAGTTCGGGTGCTGTGT
GGCTCGTACCCGCGCATTCAGGCGGCAGGGGGTCTAACGGGTCTAAGGCGGCGTGTACGG
CCGCCACAGCGGCTCTTAGCGGCCCGGAAACGTCCTCGAAACGACGCATGTGTTCCTCCT
GGTTGGTACAGGTGGTTGGGGGTGCTCGGCTGTCGCTGGTGTTTCATCATCAGGGCTCGA
CGGGAGAGCGGGGGAGTGTGCAGTTGTGGGGTGGCCCCTCAGCGAAATATCTGACTTGGA
GCTCGTGTCGGACCATACACCGGTGATTAATCGTGGTTTATTATCAAGCGTGAGCCACGT
CGCCGACGAATTTGAGCAGCTCTGGCTGCCGTACTGGTCCCTGGCAAGCGACGATCTGCT
CGAGGGGATCTACCGCCAAAGCCGCGCGTCGGCCCTAGGCCGCCGGTACATCGAGGCGAA
CCCAACAGCGCTGGCAAACCTGCTGGTCGTGGACGTAGACCATCCAGACGCAGCGCTCCG
AGCGCTCAGCGCCCGGGGGTCCCATCCGCTGCCCAACGCGATCGTGGGCAATCGCGCCAA
CGGCCACGCACACGCAGTGTGGGCACTCAACGCCCCTGTTCCACGCACCGAATACGCGCG
GCGTAAGCCGCTCGCATACATGGCGGCGTGCGCCGAAGGCCTTCGGCGCGCCGTCGATGG
CGACCGCAGTTACTCAGGCCTCATGACCAAAAACCCCGGCCACATCGCCTGGGAAACGGA
ATGGCTCCACTCAGATCTCTACACACTCAGCCACATCGAGGCCGAGCTCGGCGCGAACAT
GCCACCGCCGCGCTGGCGTCAGCAGACCACGTACAAAGCGGCTCCGACGCCGCTAGGGCG
GAATTGCGCACTGTTCGATTCCGTCAGGTTGTGGGCCTATCTTCCCGCCCTCATGCGGAT
CTACCTGCCGACCCGGAACGTGGACGGACTCGGCCGCGCGATCTATGCCGAGTGCCACGC
GCGAAACGCCGAATTTCCGTGCAACGACGTGTGTCCCGGACCGCTACCGGACAGCGAGGT
CCGCGCCATCGCCAACAGCATTTGGCGTTGGATCACAACCAAGTCGCGCATTTGGGCGGA
CGGGATCGTGGTCTACGAGGCCACACTCAGTGCGCGCCATGCGGCCATCTCGCGGAAGGG
CGCAGCAGCGCGCACGGCGGCGAGCACAGTTGCGCGGCGCGCAAAGTCCGCGTCAGCCAT
GGAGGCATTGCTATGAGCGACGGCTACAGCGACGGCTACAGCGACGGCTACAACTGGCAG
CCGACTGTCCGCAAAAAGCGGCGCGTGACCGCCGCCGAAGGCGCTCGAATCACCGGACTA
TCCGAACGCCACGTCGTCCGGCTCGTGGCGCAGGAACGCAGCGAGTGGTTCGCCGAGCAG
GCTGCACGCCGCGAACGCATCCGCGCCTATCACGACGACGAGGGCCACTCTTGGCCGCAA
ACGGCCAAACATTTCGGGCTGCATCTGGACACCGTTAAGCGACTCGGCTATCGGGCGAGG
AAAGAGCGTGCGGCAGAACAGGAAGCGGCTCAAAAGGCCCACAACGAAGCCGACAATCCA
CCGCTGTTCTAACGCAATTGGGGAGCGGGTGTCGCGGGGGTTCCGTGGGGGGTTCCGTTG
CAACGGGTCGGACAGGTAAAAGTCCTGGTAGACGCTAGTTTTCTGGTTTGGGCCATGCCT
GTCTCGTTGCGTGTTTCGTTGCGTCCGTTTTGAATACCAGCCAGACGAGACGGGGTTCTA
CGAATCTTGGTCGATACCAAGCCATTTCCGCTGAATATCGTGGAGCTCACCGCCAGAATC
GGTGGTTGTGGTGATGTACGTGGCGAACTCCGTTGTAGTGCTTGTGGTGGCATCCGTGGC
GCGGCCGCGGTACCAGATCTTTAAATCTAGATAAAGAAGTGACGCGGTCTCAAGCGTCGA
GCGTCGTCAGCGTGTCGAGGATGTCGAAGTCGTAGCCGTCGGCGCTGGCGATGTAGACCT
GCTGGTCGAATTGACTGTCGCGCATACACATCGGGCCCCGGGGCCCGTCGAACCCGACAT
CGTGCGCGGATGCCATCAGGTCCGGTATCTCGGGGGAGTGGGCCCGCTGGAAGATGGCCT
CGAGCGCAACGAGACCCTCGTAACAGGATTCGGCCATCGCGTTGAGCGGTGGCGCGTCGG
CGCCGTAGCGGGCGACGTAGCTGCCCATCAGGTCCATGGCACCCGCGGTGGCCAGTGAAC
TGAAGTACGCCGCGGCGACATAGAGGTTTTCGGTGGAGCCGGCGCCGCTGGCCAGCAGCA
TGTTCTCCTCCATCAGCGGGCTGAACCGCGCCATGCGGTCGTGCCCGCCGGCGCGCGCGA
ACTCGCGGTTGAACAACACGGCGTCCTGGCCGACGAGCAGCATCAACACGGCCTGCGCCC
CCGACGCGATGGCCTTGCGGACCGGTGCGCGGAAATCGTCGGTGCCGTACGGGACGTAGA
TCTCCCGTCTGAGCTCGAGGTCCAGATCTCGGCAGTACGCGCGGGCGGCCGCGGCGGAAC
GGCGCGGCCAGATGTAGTCATCGCCGACCAGGCACCAGGACCGGATGCCGAAGTGGTCGC
GCAGCCAGGCGAGCGCGGGCGCGATCTGGATCTGCGGTGTCTCGCCTGTGCAGAACACGC
CCGGTGTGCGTTCACCGCCCTCGTACAACGAGGTGTAGACGTACGGGATGCGGTCGCGGA
CCACCGGGGAGATGCGGTTGCGCACGGCCGAGATGTGCCAGCCGGTCACGGCGTCGAGAC

CGTGACCTCGCAACCGGTCGGCGACGGTCCGGGCGACGTCGTCGCCGGGCCGTCCGCCGT
CGAGCACCTCGATGGTGACCTTGCGGCCCTGCAGGCCGCCTCGGTCGTTGACCTCCTTGG
CCGCGAGCTCGGCCACGGCCTCGCACGAAGGCGCGAAGATTCCCGCTGGCCCTTGAAGCG
GAATCACCAGCCCGACGCGGAACTCAACCTCGCCGTCCTGCACTCCAGATCACCGTCGAT
CCCGTGTAGTCTGCGCTTCAAAGCTTTCTAGCAGAAATAATTCATTCTGAACAGACCCCG
CCGTCGACACGAGGAGACACCCACCATGGCCGCCGGACAGCAGCGCCGCCCCAACCTCCT
GCTGCCGTTGGTGCGTCTGACCCACCTCGCGGAGTCGGCGATCGAACGCGTGCTCGCGGA
CTCGTCGCTCAAGATCGAGGACTGGCGCGTGCTCGACGAGTTGGCCGGACGGCGCACCGT
GCCCATGAGCGATCTCGCGCAGGCCACGCTGATCACGGGTCCGACTCTCACCAGAACCGT
CGATCGCCTTGTGTCGCAAGGGATCATCTACCGGACTGCCGATCTGCATGACCGCCGGCG
GGTGCTCGTGGCGTTGACCCCGCGGGGGCGGACGCTGCGCAACCGCCTGGTGGACGCGGT
AGCCGAGGCCGAGTGTGCGGCTTTTGAATCGTGCGGGCTGGACGTCGACCAGTTGCGCGA
ACTCGTCGACACCACCTCGAATTTGACTTCGTAACCACCCGCGCCCGGCCGGCGTTCACC
CTTGACTTTTATTTTCATCTGGATATATTTCGGGTGAATGGAAAGGGGTGACCATGCCGA
CCTACACATTCCGTTGTTCCCACTGCGGTCCCTTCGATCTCACCTGCGCGATCTCCGAGC
GCGATGCGGCGGCGACCTGTCCGGAGTGCCGGACGCCGGCGCGCCGGGTCTTCGGTTCGG
TAGGGCTGACGACATTCACCGCGGGACATCACCGCGCATTCGACGCGGCGTCCGCGAGCG
CCGAAAGTCCCACGGTGGTGAAGTCGATTCCCGCAGGCGCGGACCGCCCGCGGGCCCCGC
GCCGCAATCCCGGTCTACCGAGTCTGCCGAGGTACTAGCGACATGGGTGGCGTCGGGCTC
TTCTACGTGGGTGCGGTGCTCATCATCGACGGGCTGATGCTGCTGGGCCGCATCAGCCCA
CGAGGCGCAACACCGCTGAACTTCTTCGTCGGCGGACTGCAGGTGGTGACGCCTACGGTG
CTGATCCTGCAGTCCGGCGGAGACGCGGCCGTGATCTTCGCGGCCTCCGGGCTCTACCTG
TTCGGCTTCACCTACCTGTGGGTGGCCATCAACAACGTGACCGACTGGGACGGAGAAGGT
CTCGGATGGTTCTCGCTGTTCGTCGCGATCGCCGCACTCGGCTACTCGTGGCACGCGTTC
ACCGCCGAGGCCGACCCGGCGTTCGGGGTGATCTGGCTGCTGTGGGCAGTGCTGTGGTTC
ATGCTGTTCCTGCTGCTCGGCCTGGGGCACGACGCACTGGGGCCCGCCGTCGGGTTCGTC
GCGGTGGCCGAAGGCGTGATCACCGCCGCCGTGCCGGCCTTCCTGATCGTGTCGGGCAAC
TGGGAAACCGGCCCGCTCCCCGCCGCGGTCATCGCCGTGATCGGTTTTGCCGCAGTTGTT
CTCGCATACCCCATCGGGCGCCGTCTCGCAGCGCCGTCAGTCACCAACCCTCCACCGGCC
GCGCTCGCGGCCACCACCCGATAAGAGAAAGGGAGTCCACATCTGTAAGAATTCTTAATT
AAGCTAGCATTTAAATGGATCCGTGTCTCAAAATCTCTGATGTTACATTGCACAAGATAA
AAATATATCATCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGT
TATGAGCCATATTCAACGGGAAACGTCTTGCTCGAGGCCGCGATTAAATTCCAACATGGA
TGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAAT
CTATCGCTTGTATGGGAAGCCCCATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAG
CGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCC
TCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGC
GATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATAT
TGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCC
TTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTT
GGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAA
AGAAATGCATAATCTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTC
ACTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGT
CGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTC
TCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAA
ATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAATCAGAATTGGTTAATTGGTTGTA
ACACTGGCAGAGCATTACGCTGACTTGACATCGATTCCCGGTATCAACAGGGACACCAGG
ATTTATTTATTCTGCGAAGTGATCTTCCGTCACAGGTATTTATTCGCGATAAGCTCATGG
AGCGGCGTAACCGTCGCACAGGAAGGACAGAGAAAGCGCGGATCTGGGAAGTGACGGACA
GAACGGTCAGGACCTGGATTGGGGAGGCGGTTGCCGCCGCTGCTGCTGACGGTGTGACGT

TCTCTGTTCCGGTCACACCACATACGTTCCGCCATTCCTATGCGATGCACATGCTGTATG
CCGGTATACCGCTGAAAGTTCTGCAAAGCCTGATGGGACATAAGTCCATCAGTTCAACGG
AAGTCTACACGAAGGTTTTTGCGCTGGATGTGGCTGCCCGGCACCGGGTGCAGTTTGCGA
TGCCGGAGTCTGATGCGGTTGCGATGCTGAAACAATTATCCTGAGAATAAATGCCTTGGC
CTTTATATGGAAATGTGGAACTGAGTGGATATGCTGTTTTTGTCTGTTAAACAGAGAAGC
TGGCTGTTATCCACTGAGAAGCGAACGAAACAGTCGGGAAAATCTCCCATTATCGTAGAG
ATCCGCATTATTAATCTCAGGAGCCTGTGTAGCGTTTATAGGAAGTAGTGTTCTGTCATG
ATGCCTGCAAGCGGTAACGAAAACGATTTGAATATGCCTTCAGGAACAATAGAAATCTTC
GTGCGGTGTTACGTTGAAGTGGAGCGGATTATGTCAGCAATGGACAGAACAACCTAATGA
ACACAGAACCATGATGTGGTCTGTCCTTTTACAGCCAGTAGTGCTCGCCGCAGTCGAGCG
ACAGGGCGAAGCCCTCGGCTGGTTGCCCTCGCCGCTGGGCTGGCGGCCGTCTATGGCCCT
GCAAACGCGCCAGAAACGCCGTCGAAGCCGTGTGCGAGACACCGCGGCCGGCCGCCGGCG
TTGTGGATACCTCGCGGAAAACTTGGCCCTCACTGACAGATGAGGGGCGGACGTTGACAC
TTGAGGGGCCGACTCACCCGGCGCGGCGTTGACAGATGAGGGGCAGGCTCGATTTCGGCC
GGCGACGTGGAGCTGGCCAGCCTCGCAAATCGGCGAAAACGCCTGATTTTACGCGAGTTT
CCCACAGATGATGTGGACAAGCCTGGGGATAAGTGCCCTGCGGTATTGACACTTGAGGGG
CGCGACTACTGACAGATGAGGGGCGCGATCCTTGACACTTGAGGGGCAGAGTGCTGACAG
ATGAGGGGCGCACCTATTGACATTTGAGGGGCTGTCCACAGGCAGAAAATCCAGCATTTG
CAAGGGTTTCCGCCCGTTTTTCGGCCACCGCTAACCTGTCTTTTAACCTGCTTTTAAACC
AATATTTATAAACCTTGTTTTTAACCAGGGCTGCGCCCTGTGCGCGTGACCGCGCACGCC
GAAGGGGGGTGCCCCCCCTTCTCGAACCCTCCCGGTCGAGTGAGCGAGGAAGCACCAGGG
AACAGCACTTATATATTCTGCTTACACACGATGCCTGAAAAAACTTCCCTTGGGGTTATC
CACTTATCCACGGGGATATTTTTATAATTATTTTTTTTATAGTTTTTAGATCTTCTTTTT
TAGAGCGCCTTGTAGGCCTTTATCCATGCTGGTTCTAGAGAAGGTGTTGTGACAAATTGC
CCTTTCAGTGTGACAAATCACCCTCAAATGACAGTCCTGTCTGTGACAAATTGCCCTTAA
CCCTGTGACAAATTGCCCTCAGAAGAAGCTGTTTTTTCACAAAGTTATCCCTGCTTATTG
ACTCTTTTTTATTTAGTGTGACAATCTAAAAACTTGTCACACTTCACATGGATCTGTCAT
GGCGGAAACAGCGGTTATCAATCACAAGAAACGTAAAAATAGCCCGCGAATCGTCCAGTC
AAACGACCTCACTGAGGCGGCATATAGTCTCTCCCGGGATCAAAAACGTATGCTGTATCT
GTTCGTTGACCAGATCAGAAAATCTGATGGCACCCTACAGGAACATGACGGTATCTGCGA
GATCCATGTTGCTAAATATGCTGAAATATTCGGATTGACCTCTGCGGAAGCCAGTAAGGA
TATACGGCAGGCATTGAAGAGTTTCGCGGGGAAGGAAGTGGTTTTTTATCGCCCTGAAGA
GGATGCCGGCGATGAAAAAGGCTATGAATCTTTTCCTTGGTTTATCAAACGTGCGCACAG
TCCATCCAGAGGGCTTTACAGTGTACATATCAACCCATATCTCATTCCCTTCTTTATCGG
GTTACAGAACCGGTTTACGCAGTTTCGGCTTAGTGAAACAAAAGAAATCACCAATCCGTA
TGCCATGCGTTTATACGAATCCCTGTGTCAGTATCGTAAGCCGGATGGCTCAGGCATCGT
CTCTCTGAAAATCGACTGGATCATAGAGCGTTACCAGCTGCCTCAAAGTTACCAGCGTAT
GCCTGACTTCCGCCGCCGCTTCCTGCAGGTCTGTGTTAATGAGATCAACAGCAGAACTCC
AATGCGCCTCTCATACATTGAGAAAAAGAAAGGCCGCCAGACGACTCATATCGTATTTTC
CTTCCGCGATATCACTTCCATGACGACAGGATAGTCTGAGGGTTATCTGTCACAGATTTG
AGGGTGGTTCGTCACATTTGTTCTGACCTACTGAGGGTAATTTGTCACAGTTTTGCTGTT
TCCTTCAGCCTGCATGGATTTTCTCATACTTTTTGAACTGTAATTTTTAAGGAAGCCAAA
TTTGAGGGCAGTTTGTCACAGTTGATTTCCTTCTCTTTCCCTTCGTCATGTGACCTGATA
TCGGGGGTTAGTTCGTCATCATTGATGAGGGTTGATTATCACAGTTTATTACTCTGAATT
GGCTATCCGCGTGTGTACCTCTACCTGGAGTTTTTCCCACGGTGGATATTTCTTCTTGCG
CTGAGCGTAAGAGCTATCTGACAGAACAGTTCTTCTTTGCTTCCTCGCCAGTTCGCTCGC
TATGCTCGGTTACACGGCTGCGGCGAGCGCTAGTGATAATAAGTGACTGAGGTATGTGCT
CTTCTTATCTCCTTTTGTAGTGTTGCTCTTATTTTAAACAACTTTGCGGTTTTTTGATGA
CTTTGCGATTTTGTTGTTGCTTTGCAGTAAATTGCAAGATTTAATAAAAAAACGCAAAGC
AATGATTAAAGGATGTTCAGAATGAAACTCATGGAAACACTTAACCAGTGCATAAACGCT

GGTCATGAAATGACGAAGGCTATCGCCATTGCACAGTTTAATGATGACAGCCCGGAAGCG
AGGAAAATAACCCGGCGCTGGAGAATAGGTGAAGCAGCGGATTTAGTTGGGGTTTCTTCT
CAGGCTATCAGAGATGCCGAGAAAGCAGGGCGACTACCGCACCCGGATATGGAAATTCGA
GGACGGGTTGAGCAACGTGTTGGTTATACAATTGAACAAATTAATCATATGCGTGATGTG
TTTGGTACGCGATTGCGACGTGCTGAAGACGTATTTCCACCGGTGATCGGGGTTGCTGCC
CATAAAGGTGGCGTTTACAAAACCTCAGTTTCTGTTCATCTTGCTCAGGATCTGGCTCTG
AAGGGGCTACGTGTTTTGCTCGTGGAAGGTAACGACCCCCAGGGAACAGCCTCAATGTAT
CACGGATGGGTACCAGATCTTCATATTCATGCAGAAGACACTCTCCTGCCTTTCTATCTT
GGGGAAAAGGACGATGTCACTTATGCAATAAAGCCCACTTGCTGGCCGGGGCTTGACATT
ATTCCTTCCTGTCTGGCTCTGCACCGTATTGAAACTGAGTTAATGGGCAAATTTGATGAA
GGTAAACTGCCCACCGATCCACACCTGATGCTCCGACTGGCCATTGAAACTGTTGCTCAT
GACTATGATGTCATAGTTATTGACAGCGCGCCTAACCTGGGTATCGGCACGATTAATGTC
GTATGTGCTGCTGATGTGCTGATTGTTCCCACGCCTGCTGAGTTGTTTGACTACACCTCC
GCACTGCAGTTTTTCGATATGCTTCGTGATCTGCTCAAGAACGTTGATCTTAAAGGGTTC
GAGCCTGATGTACGTATTTTGCTTACCAAATACAGCAATAGTAATGGCTCTCAGTCCCCG
TGGATGGAGGAGCAAATTCGGGATGCCTGGGGAAGCATGGTTCTAAAAAATGTTGTACGT
GAAACGGATGAAGTTGGTAAAGGTCAGATCCGGATGAGAACTGTTTTTGAACAGGCCATT
GATCAACGCTCTTCAACTGGTGCCTGGAGAAATGCTCTTTCTATTTGGGAACCTGTCTGC
AATGAAATTTTCGATCGTCTGATTAAACCACGCTGGGAGATTAGATAATGAAGCGTGCGC
CTGTTATTCCAAAACATACGCTCAATACTCAACCGGTTGAAGATACTTCGTTATCGACAC
CAGCTGCCCCGATGGTGGATTCGTTAATTGCGCGCGTAGGAGTAATGGCTCGCGGTAATG
CCATTACTTTGCCTGTATGTGGTCGGGATGTGAAGTTTACTCTTGAAGTGCTCCGGGGTG
ATAGTGTTGAGAAGACCTCTCGGGTATGGTCAGGTAATGAACGTGACCAGGAGCTGCTTA
CTGAGGACGCACTGGATGATCTCATCCCTTCTTTTCTACTGACTGGTCAACAGACACCGG
CGTTCGGTCGAAGAGTATCTGGTGTCATAGAAATTGCCGATGGGAGTCGCCGTCGTAAAG
CTGCTGCACTTACCGAAAGTGATTATCGTGTTCTGGTTGGCGAGCTGGATGATGAGCAGA
TGGCTGCATTATCCAGATTGGGTAACGATTATCGCCCAACAAGTGCTTATGAACGTGGTC
AGCGTTATGCAAGCCGATTGCAGAATGAATTTGCTGGAAATATTTCTGCGCTGGCTGATG
CGGAAAATATTTCACGTAAGATTATTACCCGCTGTATCAACACCGCCAAATTGCCTAAAT
CAGTTGTTGCTCTTTTTTCTCACCCCGGTGAACTATCTGCCCGGTCAGGTGATGCACTTC
AAAAAGCCTTTACAGATAAAGAGGAATTACTTAAGCAGCAGGCATCTAACCTTCATGAGC
AGAAAAAAGCTGGGGTGATATTTGAAGCTGAAGAAGTTATCACTCTTTTAACTTCTGTGC
TTAAAACGTCATCTGCATCAAGAACTAGTTTAAGCTCACGACATCAGTTTGCTCCTGGAG
CGACAGTATTGTATAAGGGCGATAAAATGGTGCTTAACCTGGACAGGTCTCGTGTTCCAA
CTGAGTGTATAGAGAAAATTGAGGCCATTCTTAAGGAACTTGAAAAGCCAGCACCCTGAT
GCGACCACGTTTTAGTCTACGTTTATCTGTCTTTACTTAATGTCCTTTGTTACAGGCCAG
AAAGCATAACTGGCCTGAATATTCTCTCTGGGCCCACTGTTCCACTTGTATCGTCGGTCT
GATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGAC
CACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCCCACTCGTAT
CGTCGGTCTGATAATCAGACTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTA
GTCTGGGACCATGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGACCACGGTCC
CACTCGTATCGTCGGTCTGATTATTAGTCTGGAACCACGGTCCCACTCGTATCGTCGGTC
TGATTATTAGTCTGGGACCACGGTCCCACTCGTATCGTCGGTCTGATTATTAGTCTGGGA
CCACGATCCCACTCGTGTTGTCGGTCTGATTATCGGTCTGGGACCACGGTCCCACTTGTA
TTGTCGATCAGACTATCAGCGTGAGACTACGATTCCATCAATGCCTGTCAAGGGCAAGTA
TTGACATGTCGTCGTAACCTGTAGAACGGAGTAACCTCGGTGTGCGGTTGTATGCCTGCT
GTGGATTGCTGCTGTGTCCTGCTTATCCACAACATTTTGCGCACGGTTATGTGGACAAAA
TACCTGGTTACCCAGGCCGTGCCGGCACGTTAACCGGGCTGCATCCGATGCAAGTGTGTC
GCTGTCGACGAGCTCGCGAGCTCGGACATGAGGTTGCCCCGTATTCAGTGTCGCTGATTT
GTATTGTCTGAAGTTGTTTTTACGTTAAGTTGATGCAGATCAATTAATACGATACCTGCG

TCATAATTGATTATTTGACGTGGTTTGATGGCCTCCACGCACGTTGTGATATGTAGATGA
TAATCATTATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACGGGGCGGCGACCT
CGCGGGTTTTCGCTATTTATGAAAATTTTCCGGTTTAAGGCGTTTCCGTTCTTCTTCGTC
ATAACTTAATGTTTTTATTTAAAATACCCTCTGAAAAGAAAGGAAACGACAGGTGCTGAA
AGCGAGCTTTTTGGCCTCTGTCGTTTCCTTTCTCTGTTTTTGTCCGTGGAATGAACAATG
GAAGTCCGAGCTCATCGCTAATAACTTCGTATAGCATACATTATACGAAGTTATATTCGA
TGCGGCCGCAAGGGGTTCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCTGGCTTAACTAT
GCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATGCGGTGTGAAATACCGCACAGA
TGCGTAAGGAGAAAATACCGCATCAGGCGCCATTCGCCATTCAGCTGCGCAACTGTTGGG
AAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTG
CAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTAT

B. pMycoFos–CXAB Insert Sequence

Putative Nitrosopumilus maritimus ammonia monooxygenase coding regions (Nmar_1500, amoA; Nmar_1501, amoX; Nmar_1502, amoC and _1503, amoB) inserted into the EcoRI and SwaI restriction sites of pMycoFos. AMO coding regions were codon-optimized for Mycobacterium smegmatis, but the intergenic regions were not codon optimized. Coding regions are highlighted in yellow.

ATGATCACCATGGCCCAGATGCCCGCCCTCATCCCGAAGGAGGTCGAGATCCAGCGCCTGAAG
AAGATCTGGCTCATCGTCATCGCCATGGGCTCCACCGCCGCGTCGGTCGAGGTGGACAACTTC
GTCGACGGCTCGCTGCACCAGACCAGCATCCGCGACTCCGCCTTCACGCCGGCGCACTGGTGG
CTGTACTCGCACTTCGTGGCCCTGCCCCTCGGCTGGGGCAGCGCCGCGATCTACGACCGCAAG
GTCCCGGTGCTGCGGGGCCCCAACAACTCGATGAACACCGGCCTCAAGATGACGATCCTGGGC
TACCTCGCGACCATGTTCACGATCGGCGTCAACGAGATGTGGCACTTCTGGTTCGTCGAGGAG
ATCTTCGCCGTGCCGAACCACTGGATGTTCAACATGGGCGTGGTCGTGGCGTTCATGGGCGCC
CTGGCGTACGTCGTGCGCGTCTACNCCCGGCTGGTGGAGCTCGGCGCGGAGACCCCGGGCGAG
AACCCCTACGTGGCCGAGATGTACAAGATGGCGCTGGAGGGCAAGCTCTACTCGCGGAGCATC
CCCTGAACAAAATGTGCAATCGCACAATTTTTTCTTATTTTATATTTTTAAAAAAGACGTAGA
TCGTACCCTCTTCAGCAGTTTAAGAAAGACTTAAAAGGATTCTGAATTAGAATAATTCCAGAA
TGACCCTCCCCAAGGGCTTCGGCAGCGGCGGCGGCGGCGGCAGCAGCAGCGCGGACGTGGAGC
GGATGATCGGCCGGCGGGTGGAGAACATGACCGGCATGATCACGGCCAGCTACTGGGCCGCGC
TGATCGCCACCTTCGCGGGCACGGCCGCGGGCTACTTCTACTACCCGTGGGCCTACCCCACCG
CGTCGGGCCACTTCGCGTTCATCGTCCTCGCCATCATCGAGGCGATCGGCTACATCTTCTGCG
TGAAGGTCATGGAGGAGGGCTCGAACAAGAACAGCAACGGCATCGTCGGCGCCTCCATCGCGG
GCACGGTGGCCTTCGTCCTGTTCGTGTCGCTCTTCGTGGGCTGGTGATCAGAAGGTAAATCCT
TCAGAACACTTCTTTTTCATTTTAATTTTTGACAAATACCTGTGTGTATTTGCCGAATAAAGC
GACGATCTCCATCACACTCTAAAATTTTATATACTGACCGTTTCTTCAACTTGATATGGTGTG
GCTCCGGCGCTGCACGCACTACCTCTTCATCGTGGTGGTGGCGGTGAACTCGACCCTGCTGAC
GATCAACGCGGGCGACTACATCTTCTACACCGACTGGGCGTGGACCTCGTACACGGTGTTCTC
GATCAGCCAGACCCTGATGCTCATCGTCGGCGCCACCTACTACCTGACCTTCACGGGCGTGCC
GGGCACCGCGACGTACTACGCCCTGATCATGACCGTCTACACGTGGATCGCGAAGGCCGCGTG
GTTCAGCCTGGGCTACCCCTACGACTTCATCGTGACCCCGGTCTGGCTGCCCTCCGCGATGCT
GCTCGACCTCGTGTACTGGGCCACCAAGAAGAACAAGCACTCCCTGATCCTCTTCGGCGGCGT
GCTGGTCGGCATGTCGCTGCCGCTCTTCAACATGGTGAACCTGATCACCGTCGCGGACCCCCT
CGAGACGGCCTTCAAGTACCCGCGCCCCACCCTGCCGCCCTACATGACGCCGATCGAGCCCCA
GGTCGGCAAGTTCTACAACTCGCCGGTCGCCCTGGGCGCCGGCGCGGGCGCCGTCCTCGGCTG
CACCTTCGCCGCGCTGGGCTGCAAGCTCAACACCTGGACGTACCGGTGGATGGCCGCGTGGAG
CAAGTGGGACTGAGGTGTATATTGAGAATATAAAGTAAGAGATTGAAATCAGGTTGTTTTTCT
TCCAATTTCTAACTATTTTTCAAATGATTTGTGAAATTATTTCATTTTCTATTTTTTTCATTG
TAAGAATAATAACTGTTTTTTGTTATTTTCAATGTTTAACATCGATGATTAACATCACTATTA
GTATCGATCCTCGCAATGAAGGTAAAACATATAACGACGTGTATTATCTTTCACAATATTAGG
GATAACTATGGTGGAGAAGAAGATCTTCGTGTTCGGCCTGGCGGTGGTCCTGGCGCTGGGCAC
GCTGGGCTTCAACTGGGTGGAGAGCATCCTCCCGACCGCCGACGCGCACGGCGTCCAGGCCCA
GCTGCAGTCGCGCTTCGTGCGGATCGAGGACGAGACCTTCAACCGCCAGAGCCTCCAGACCGG
CGAGACGCTGGTCCTCCAGGGCACGCTGGTCAGCCTCGTGGAGCGCGACCTGCGGGGCTGGAT
CTCCNTCTTCTCGGAGAGCACCAACGCGGGCAACCGCTGGGAGATGCTGTCGCGGGACCCGCC
CGGCAACGTGTTCGACATCCCGGGCAACTCCGTGGTCGACTACCAGCTGTCGGCCAAGGCGCT

CGAGGCCGGCGTCTACCACGTGCACACCCAGCTGAACGTCGCGCAGGTCGGCCCGGGCCTCGG
CCCCGGCCAGACCGTGGTCGTGGAGGGCGAGCCGATCATCAAGCCGATCCCCTACACGAACAT
CGCCTACCAGTCGATCATGATCGGCGTCGGCTACGTGATCACCTTCGCGACGCGCCCCTGGCA
GGTCATCTGA

C. CXAB amoB Mutants

ATGGTGGAGAAGAAGATCTTCGTGTTCGGCCTGGCGGTGGTCCTGGCGCTGGGCACGCTGGGC
TTCAACTGGGTGGAGAGCATCCTCCCGACCGCCGACGCGCACGGCGTCCAGGCCCAGCTGCAG
TCGCGCTTCGTGCGGATCGAGGACGAGACCTTCAACCGCCAGAGCCTCCAGACCGGCGAGACG
CTGGTCCTCCAGGGCACGCTGGTCAGCCTCGTGGAGCGCGACCTGCGGGGCTGGATCTCCATC
TTCTCGGAGAGCACCAACGCGGGCAACCGCTGGGAGATGCTGTCGCGGGACCCGCCCGGCAAC
GTGTTCGACATCCCGGGCAACTCCGTGGTCGACTACCAGCTGTCGGCCAAGGCGCTCGAGGCC
GGCGTCTACCACGTGCACACCCAGCTGAACGTCGCGCAGGTCGGCCCGGGCCTCGGCCCCGGC
CAGACCGTGGTCGTGGAGGGCGAGCCGATCATCAAGCCGATCCCCTACACGAACATCGCCTAC
CAGTCGATCATGATCGGCGTCGGCTACGTGATCACCTTCGCGACGCGCCCCTGGCAGGTCATC
TGA

Relevant histidine genes highlighted in red above in order: His35, His130, His132

D. CXAB amoC Mutants

ATGATCACCATGGCCCAGATGCCCGCCCTCATCCCGAAGGAGGTCGAGATCCAGCGCCTGAAG
AAGATCTGGCTCATCGTCATCGCCATGGGCTCCACCGCCGCGTCGGTCGAGGTGGACAACTTC
GTCGACGGCTCGCTGCACCAGACCAGCATCCGCGACTCCGCCTTCACGCCGGCGCACTGGTGG
CTGTACTCGCACTTCGTGGCCCTGCCCCTCGGCTGGGGCAGCGCCGCGATCTACGACCGCAAG
GTCCCGGTGCTGCGGGGCCCCAACAACTCGATGAACACCGGCCTCAAGATGACGATCCTGGGC
TACCTCGCGACCATGTTCACGATCGGCGTCAACGAGATGTGGCACTTCTGGTTCGTCGAGGAG
ATCTTCGCCGTGCCGAACCACTGGATGTTCAACATGGGCGTGGTCGTGGCGTTCATGGGCGCC
CTGGCGTACGTCGTGCGCGTCTACGCCCGGCTGGTGGAGCTCGGCGCGGAGACCCCGGGCGAG
AACCCCTACGTGGCCGAGATGTACAAGATGGCGCTGGAGGGCAAGCTCTACTCGCGGAGCATC
CCCTGA

Relevant histidine genes highlighted in red above in order: Asp44, His48, His61

E. *Nitrosomonas* sp. AL212 cyt P460 Sequence

ATGGGCCTGCAGTTCAAGAAAACCCTGCTGAGCAGCATTGCGCCGGTGCTGCTGAGCATTGTT
CTGGCGAACCCGGTGATTGCGAGCGATGCGCACCATGCGCACAAGGGTCTGAACTACGGCAGC
TTCACCAAGGAGCACGTTCTGCTGACCCCGAAAGGTTATCGTGAATGGGTTTTTATTGGCGCG
AGCGTGACCCCGAACGAGCTGAACGACGATAAAGCGGCGTTCCCGGAATTTCACAACGTGTAC
ATTGACCCGACCAGCTGGGGTCACTGGAAGAAAACCGGCGAGTTCCGTGATGGCACCGTGATC
GTTAAGGAACTGGCGGGTGTTGGCAGCAAAGCGAGCCCGAGCGGTAACGGCTATTTCCCGGGC
GAGTTTAACGGCATCGCGGCGATGGTGAAGGATAGCAAACGTTACCCGGAACGTCCGGGTAAC
TGGGCGTTCTTTGGCTTTGAGAGCTATGAAGCGAAGCAGGGTATCATTCAAACCGACGAGACC
TGCGCGGCGTGCCACAAAGAACATGCGGCGCACGATATGGTTTTCACCCAATTTTATCCGGTG
CTGCGTGCGGGCAAGCCGAGCAAACTCGAG

***Nitrosomonas* sp. AL212 cyt P460 Mutants**

Mutations to residue 131 replace the above alanine codon highlighted in red to one of the following variations:

Ala131Glu: GAA
Ala131Gln: CAG
Ala131Leu: CTG
Ala131Asp: GAT

G. *Nitrosomonas europaea* cyt P460 Sequence

ATGGCTGGCGTCGCGGAATTTAACGATAAAGGTGAACTGCTGCTGCCGAAAAATTATCGTGAA
TGGGTCATGGTGGGCACCCAGGTTACGCCGAACGAACTGAATGATGGTAAAGCTCCGTTTACC
GAAATTCGCACGGTTTATGTCGACCCGGAAAGCTACGCCCATTGGAAGAAAACCGGCGAATTC
CGTGATGGTACCGTGACGGTTAAAGAACTGGTCAGTGTGGGTGACCGTAAAGGTCCGGGTTCC
GGTAACGGTTATTTTATGGGCATTACATTGGTCTGGAAGCGAGCGTGAAAGACTCTCAGCGTT
TCGCCAACGAACCGGGTAATTGGGCATTTTATATCTTCTACGTTCCGGATACCCCGCTGGTCG
CGGCAGCAAAAAACCTGCCGACGGCCGAATGCGCTGCGTGTCACAAAGAAAATGCAAAAACCG
ACATGGTGTTTACGCAATTCTACCCGGTTCTGCGCGCCGCAAAGCTACCGGCGAAAGCGGTGT
GGTTGCGCCGAAACTCGAG

H. *Nitrosomonas europaea* cyt P460 Glu97Ala
Mutation is highlighted in red

ATGGCTGGCGTCGCGGAATTTAACGATAAAGGTGAACTGCTGCTGCCGAAAAATTATCGTGAA
TGGGTCATGGTGGGCACCCAGGTTACGCCGAACGAACTGAATGATGGTAAAGCTCCGTTTACC
GAAATTCGCACGGTTTATGTCGACCCGGAAAGCTACGCCCATTGGAAGAAAACCGGCGAATTC
CGTGATGGTACCGTGACGGTTAAAGAACTGGTCAGTGTGGGTGACCGTAAAGGTCCGGGTTCC
GGTAACGGTTATTTTATGGGCATTACATTGGTCTGGCGGCGAGCGTGAAAGACTCTCAGCGTT
TCGCCAACGAACCGGGTAATTGGGCATTTTATATCTTCTACGTTCCGGATACCCCGCTGGTCG
CGGCAGCAAAAAACCTGCCGACGGCCGAATGCGCTGCGTGTCACAAAGAAAATGCAAAAACCG
ACATGGTGTTTACGCAATTCTACCCGGTTCTGCGCGCCGCAAAGCTACCGGCGAAAGCGGTGT
GGTTGCGCCGAAACTCGAG

I. Nitrosocyanin Gene Sequence

ATGGAGCACAATTTTAATGTTGTTATCAATGCGTATGACACCACCATCCCGGAACTGAACGTG
GAGGGCGTGACCGTGAAGAATATCCGTGCGTTCAACGTTCTGAACGAGCCGGAAACCCTGGTG
GTTAAGAAAGGTGACGCGGTGAAGGTGGTTGTGGAGAACAAAAGCCCGATCAGCGAAGGTTTC
AGCATTGATGCGTTTGGCGTTCAGGAAGTGATCAAGGCGGGTGAAACCAAAACCATTAGCTTC
ACCGCGGACAAGGCGGGCGCGTTTACCATCTGGTGCCAACTGCATCCGAAGAATATCCATCTG
CCGGGCACCCTGAATGTGGTTGAATAA

J. Nitrosocyanin Mass Spectrometry Analysis

Accession	Description	Score	Coverage
P00761	Trypsin Pig - [TRYP_PIG]	517.99	36.77
P00001	Nitrosocyanin WP_011110790.1 [Nitrosomonas europaea]	508.02	61.03
15802236	cold shock protein [Escherichia coli O157:H7 EDL933]	48.67	78.26
90111711	ketoacid-binding protein [Escherichia coli K12]	48.55	51.56
24052838	PTS system protein HPr [Shigella flexneri 2a str. 301]	29.03	50.59
24053779	50S ribosomal subunit protein L6 [Shigella flexneri 2a str. 301]	27.62	45.76
24052039	murein lipoprotein [Shigella flexneri 2a str. 301]	23.47	33.33
56383612	50S ribosomal subunit protein L25 [Shigella flexneri 2a str. 301]	19.31	56.38
24053054	30S ribosomal subunit protein S16 [Shigella flexneri 2a str. 301]	19.20	46.34
24053787	30S ribosomal subunit protein S17 [Shigella flexneri 2a str. 301]	18.45	32.14

Top 10 results of mass spectrometry analysis of nitrosocyanin band cut from SDS-PAGE gel and digested with the protease trypsin. Analysis performed by Cornell University Institute of Biotechnology Proteomics Facility.

APPENDIX B: CRYSTALLOGRAPHIC DATA

	Fe^{III} cyt P460	Fe^{III} Ala131Glu	Ala131Glu –NO	Fe^{III} Ala131Gln	Ala131Gln–NH_2OH
Wavelength (Å)	0.979	0.979	0.979	0.979	0.979
Temperature (K)	100	100 K	100 K	100 K	100 K
Space Group	P 2_1 2_1 2_1	P1 2_1 1	P1 2_1 1	P1 2_1 1	P1 2_1 1
a (Å)	47.3	48.1	48.7	48.0	48.0
b (Å)	80.1	80.4	80.0	80.9	80.9
c (Å)	109.0	120.0	119.1	120.6	120.6
α (deg)	90.0	90.0	90.0	90.0	90.0
β (deg)	90.0	95.9	92.6	96.0	96.0
γ (deg)	90.0	90.0	90.0	90.0	90.0
Reflections	475,694 (12,166)	62608	63916	39644	41095
Number of Reflections in R_{work}	66,399	5549	2777	3858	4146
Number of Reflections in R_{free}	3,441	323	117	147	211
Resolution (Å)	64.6–1.45	119.6-1.97	48.6-1.97	67.1-2.30	29.4-2.25
R_{merge} (%)	4.1 (72.7)	5.3(111)	4.3 (120)	6.6 (98.8)	7.7 (165)
$CC_{1/2}$	0.99 (0.75)	0.99 (0.72)	0.93 (0.68)	0.87 (0.66)	0.99 (0.59)
Completeness (%)	97.0 (90.4)	96.1(84.9)	98.8 (96.6)	96.7 (98.3)	94.8 (57.8)
Redundancy	6.6 (4.6)	1.9 (1.9)	2.0 (1.9)	2.0 (2.0)	1.9 (1.8)
I/σ(I)	23.8 (1.4)	32.0 (1.1)	38.5 (0.9)	28.0 (2.1)	13.6 (0.5)
R_{work}	14.1 (25.0)	20.6 (43.4)	20.8 (41.6)	24.2 (33.2)	19.5 (34.9)
R_{free}	15.8 (23.4)	24.9 (48.4)	24.0 (45.0)	29.9 (41.5)	24.2 (40.0)
RMSD from Ideality					
Bonds (Å)	0.028	0.01	0.01	0.01	0.01
Angles (deg)	2.3	1.04	1.01	1.07	1.02
Average *B* Factors (Å^2)	29.2	53.0	57.1	50.4	52.4
Ramachandran Plot					
Allowed Regions (%)	100	100	100	100	100
Disallowed Regions (%)	0	0	0	0	0
PDBID	6AMG	6EOX	6E17	6EOZ	6EOY

APPENDIX C: ORCA INPUT FILES AND FINAL GEOMETRY OPTIMIZED STRUCTURE COORDINATES

A. Geometry Optimization ORCA Input File

```
!UKS BP86 ZORA-def2-TZVP(-f) def2/J PAL8 ZORA
!TightSCF SlowConv NormalPrint CPCM(Acetonitrile) Grid4
!OPT

%maxcore 4000

%geom optimizehydrogens true
        end

%SCF Directresetfreq 1
        DIIS MaxEq 15
        end
        Shift Shift 0.5
        Erroff 0.1
        end
        MaxIter 500
        end

#%method SpecialGridAtoms 26
#       SpecialGridIntAcc 14

* xyz CHARGE SPIN_MULTIPLICITY
COORDINATES
*
```

B. TDDFT ORCA Input File

```
!UKS B3LYP RIJCOSX ZORA-def2-TZVP(-f) def2/J PAL4 ZORA
!MORead noiter TightSCF SlowConv NormalPrint CPCM(Acetonitrile) Grid4

%maxcore 2


%basis newgto Fe "CP(PPP)" end
        end

%TDDFT NRoots 120
        MaxDim 800
        Triplets  false
        end

%method SpecialGridAtoms 26
        SpecialGridIntAcc 14
        end

* xyzfile CHARGE SPIN_MULTIPLICITY
COORDINATES

%eprnmr gtensor true
      nuclei = all Fe {aiso, adip, aorb, fgrad, rho}
      nuclei = all N {aiso, adip, fgrad}
      nuclei = all H {aiso, adip, fgrad}
      end
```

C. Cyt P460 Isoporphyrin Final Geometry Optimized Structure Coordinates

Charge: 2 Spin Multiplicity: 6

C	15.31845614467721	24.36973753631062	44.85235483475593
C	8.40130868311562	19.35816003826992	49.88036428168061
C	15.60901884329539	16.89568931007192	52.27571502430617
C	18.94586663343672	20.52640315786882	46.16886233385745
N	13.58515410314785	21.71631029927803	46.76573010894997
C	15.87945613745772	25.40392863792027	45.79905582839113
C	7.21951935966267	19.01994968217957	49.01349073000361
C	15.19846508089978	15.60810465118076	53.03250650889300
C	19.69340593673817	19.36789884116466	45.49040028997217
N	11.70163761908184	20.52302152370997	48.52259566031832
C	16.87737933252417	24.72906966975862	46.76748118017523
C	18.89347126900265	18.69270923348188	44.56347660645668
N	15.67146512730267	20.08919059518210	47.95716990083923
C	15.84114276146823	21.33093919149155	45.78782761729075
C	11.26377758731099	22.39224839159889	47.05099975177569
C	11.39587541577926	18.76570075986539	50.17679492385971
C	16.14023251754112	18.67261240709721	49.85907141292695
C	12.25946633922170	24.83456841832136	45.33595346451327
C	8.28516055116862	21.86489613740630	47.75263010771398
C	12.17723181345940	16.84685047650821	52.38565385207857
C	19.06109508904295	18.86326024992161	48.79828432937403
C	14.68159241675678	22.10986581331453	45.98281818182392
C	10.83827276432991	21.44065495333069	47.82009439878704
C	12.60851186110549	18.43734981509631	50.50006339322118
C	16.53106760423927	19.36246810405011	48.76983404247380
O	16.87884512980578	25.16409798001343	47.96483418568013
O	18.68650336980832	17.42667490139568	44.62386995131400
C	14.32453833273104	23.39934938920747	45.41958955552884
C	9.51011098487242	21.10240649154136	48.24552154729700
C	13.19645741921595	17.61150713713233	51.48519987806095
C	17.80484453093231	19.49686822215621	48.16235157643496
O	17.65058773917408	23.82330374551199	46.40917285471864
O	18.34656889778629	19.30352550548775	43.60708218826859
C	13.02215338188329	23.62620440391913	45.71542267949288
C	9.55582040650648	20.06269101484155	49.12258675223487
C	14.50239995759365	17.61343591396035	51.45554586576831
C	17.77748061100861	20.15083247847204	46.97693521093295
C	12.55684142493509	22.58553430291035	46.58252620242553
C	10.87919944916323	19.66543282760055	49.32157380778537
C	14.89618105052232	18.42731652367316	50.33540310195843
C	16.42914709593608	20.58607538418504	46.79902978005529
N	13.70424753986643	18.97679089961860	49.85827165420437
Fe	13.69841110642130	20.48496678349760	48.41050425765864
H	14.84334334770947	24.90025769853163	44.00917869676427

H	16.15234641403414	23.83882945586754	44.37059889036495
H	19.67712970172795	21.01620157961783	46.83127380509146
H	18.69805930985094	21.29463489209942	45.42584598331417
H	15.12368052503348	25.90600434126399	46.40961198971538
H	16.43409470300504	26.16454274721628	45.22762924180626
H	6.43358305344650	18.53317393451707	49.61000380247999
H	7.51699138342690	18.31410999344621	48.22308332433061
H	6.78316318939136	19.90346468648141	48.53718372740116
H	14.69855185540642	14.89615655962647	52.35795902536219
H	14.52736632495341	15.82277954931868	53.87102237219186
H	16.10158804922035	15.12992505911879	53.43286952472244
H	20.07718532785522	18.65996968685356	46.23026200210534
H	20.56742472276759	19.80774588993675	44.97735684630797
H	10.52374262676154	23.09652188841912	46.66211641498630
H	10.63138039336911	18.23487970599410	50.75601511798960
H	16.94875526514338	18.16388737946591	50.38709504762501
H	12.23603104921846	25.57559213315910	46.15153191083282
H	11.21260605392246	24.59318763413725	45.10081653945562
H	12.69427993354256	25.32573527359601	44.45611837736502
H	8.56313435620840	22.86773710132059	47.40530254784875
H	7.54339783457157	21.96587364733330	48.55446533304671
H	7.81383284984258	21.33813191818204	46.90901071856958
H	11.53362710441249	17.56431989056336	52.91369109400977
H	12.65822179454968	16.20664111377348	53.12788368116232
H	11.53539084438189	16.21913430897268	51.75137240304378
H	19.97748353387351	19.36695049151152	48.46754557149028
H	19.12883588922804	17.80078995637996	48.52281831517038
H	19.01634347976759	18.92894971484032	49.89264044273617
C	14.44140928876941	20.72131124845589	42.99117240525198
C	15.70604260962551	21.49098916597555	43.27422621925103
N	16.39961272313074	21.20190495086499	44.54758430689195
C	9.19460010142235	21.31444642226029	51.92134402821305
S	7.71497618346991	20.58718033068374	51.14976052630653
C	14.74639726813445	18.66005488160034	54.36790625165029
S	16.25469112737779	18.11661136308794	53.54639619416535
C	14.90296589870598	23.77316118383538	52.99273496987824
C	14.67410261420742	23.23530496828712	51.60376042303769
C	14.06611039306059	22.10471227119953	51.15607792596068
N	15.12986400089860	23.89218444679009	50.47936418201329
C	14.80961968735152	23.19508319497746	49.40044693693191
N	14.15631463923946	22.11082174531941	49.78089616893988
H	15.63165018852023	24.77759720953144	50.47697037151564
H	13.57448826489121	21.31291880631043	51.70672007962593
H	15.03118131915744	23.50351615968939	48.38803393263174
H	15.07703299184525	19.36062616701776	55.14582755501797
H	14.21544366200614	17.82966909908670	54.85055510470492

H	14.07580752698497	19.18766229698902	53.67895992915699
H	14.46783392315512	24.77667462755777	53.10458063874830
H	15.97751555621660	23.84137746958337	53.21600047207472
H	14.43578488652844	23.10807339044902	53.72792589474859
H	8.81050139107837	22.02119351274159	52.66868104528784
H	9.79722492905298	21.86627404417094	51.19037040035664
H	9.80074818238621	20.55741432446009	52.43333038456120
H	8.77073798742287	18.47909687927323	50.41616949375328
H	16.47066636982770	16.70022997435375	51.62898552636277
H	15.51712871300974	22.56427619108646	43.19040647176525
H	16.45549923536710	21.27243701555966	42.50104947227535
H	17.05756585906698	20.43009436379069	44.46470484245455
H	13.63980204078129	20.95065414315232	43.70562305899062
H	14.61386986600095	19.63607153903795	42.99653546863917
H	14.09270355621828	21.00811081493393	41.98965728040827
H	18.78140217350073	20.16199931457996	43.40068835223424
H	17.64490146695697	23.68100595998187	45.43548766542651
H	16.75988303808004	21.57624049227694	47.32676345922054

D. Cyt P460 Porphyrin, Meso C–N Final Geometry Optimized Structural Coordinates

Charge: 1 Spin Multiplicity: 6

C	15.33125154717211	24.36772013409857	44.86867487596499
C	8.40283230769547	19.35649816719800	49.88149626257398
C	15.60511408454263	16.88831073082587	52.28726869729194
C	18.95428660399856	20.52143198428420	46.18860003987297
N	13.59324132296888	21.71392469050244	46.77726262797926
C	15.89123234750261	25.40093309256898	45.81704583904207
C	7.22231774385078	19.01955763812234	49.01239416902708
C	15.19256453221343	15.60042640841191	53.04246326541687
C	19.70232111166660	19.36298457709881	45.51058677531364
N	11.70609742411934	20.52046096676692	48.53011209563702
C	16.88714326267612	24.72483658841487	46.78667873774464
C	18.90357342634417	18.68889301642226	44.58184208647491
N	15.67662967559827	20.08480162203285	47.97107668568154
C	15.85066019591348	21.32797462616508	45.80289535815466
C	11.27176535777398	22.39096357536824	47.05909181820096
C	11.39656304023895	18.76215321270021	50.18256006840429
C	16.14139645359318	18.66662868645195	49.87277155011847
C	12.27171090338045	24.83392630399493	45.34744086432879
C	8.29167476780473	21.86478919494875	47.75532856085572
C	12.17311974140226	16.84131812672061	52.39138651187682
C	19.06415011101258	18.85638280555155	48.81704505615633
C	14.69121842398486	22.10741414513499	45.99647691053502
C	10.84443191666977	21.43907009728954	47.82680005624795
C	12.60846873396940	18.43289630938479	50.50764232882728
C	16.53445377933388	19.35702827313364	48.78467902364116
O	16.88683293957898	25.15902494066965	47.98433744436434
O	18.69579597724544	17.42293286766568	44.64099807371956
C	14.33583595090712	23.39749194251765	45.43355189653363
C	9.51536556485521	21.10126801404146	48.24974967267842
C	13.19429003130177	17.60603420920641	51.49318911953761
C	17.80932841168284	19.49114013271292	48.17943966722141
O	17.66044822102313	23.81888877058817	46.42903925655902
O	18.35862653076463	19.30068566175214	43.62495565379404
C	13.03308098137686	23.62486946606824	45.72734765349871
C	9.55901424480120	20.06091280349787	49.12616100654136
C	14.50028159998521	17.60725188713015	51.46573854626894
C	17.78432953062292	20.14595004953813	46.99443797049592
C	12.56572535507378	22.58385299718059	46.59293498367633
C	10.88183343346392	19.66277370387165	49.32710044428349
C	14.89640641037396	18.42169634020955	50.34683266336787
C	16.43654159659457	20.58207306792864	46.81456469693858
N	13.70558677809832	18.97217279104293	49.86807773376005
Fe	13.70303552644075	20.48136575013733	48.42136107235865
H	14.85984805944462	24.90169124965136	44.02476363898583

H	16.16572276743591	23.83657996738533	44.38501677282287
H	19.68926489654313	21.01564728713140	46.84514278507289
H	18.69480325779341	21.28369094900433	45.44183504536207
H	15.13173280883242	25.89730684398534	46.42785128919964
H	16.44683466215525	26.16504465022832	45.25108017535639
H	6.43074186155875	18.53551422413728	49.60431761088633
H	7.52102663553902	18.31237993076971	48.22326879688648
H	6.79209156947896	19.90426557658092	48.53244443515438
H	14.69246194404693	14.88930211579872	52.36670411771462
H	14.51838377545021	15.81533624954033	53.87862013234780
H	16.09279962990267	15.11893442297058	53.44613560399008
H	20.08002682414580	18.65262894248963	46.25171291784307
H	20.58074173339454	19.79339860901914	44.99597221319947
H	10.53124525551180	23.09343562531327	46.66729103730945
H	10.62981063763712	18.23319265827068	50.76115169936451
H	16.94757923662145	18.15842683482591	50.40638514170311
H	12.27273447848516	25.59110844295543	46.14890660596908
H	11.21714101223207	24.60070910683450	45.13870343984649
H	12.69015128550559	25.30886019241234	44.44984716557828
H	8.57323498452886	22.86463306612865	47.40062178914876
H	7.55328269417666	21.97772419017983	48.55955380914939
H	7.80907920625571	21.33931400702874	46.91719064175053
H	11.50862301760423	17.55577539659373	52.89745150475533
H	12.65113853782047	16.22284621088768	53.15465219848771
H	11.54919937455575	16.18959320074857	51.76284908576496
H	19.97397671928750	19.40831962115284	48.54536147325074
H	19.18802338590715	17.81444265923878	48.48422145001170
H	18.98933530196950	18.85210389133581	49.91185856733068
C	14.45530516552351	20.72108999090912	43.00345666211696
C	15.71988904188514	21.48986126813105	43.28918264543370
N	16.41114886656807	21.19949644832615	44.56350526209550
C	9.19377426209999	21.31090912287490	51.92518279856532
S	7.71504728933243	20.58501315957836	51.15059542620025
C	14.73995197640820	18.65169307862464	54.37923982659129
S	16.24932506228387	18.10798010916775	53.55989307907141
C	14.90169906458436	23.76567313998050	53.00792227582927
C	14.67487798958786	23.22891859029871	51.61818677700225
C	14.06700948134742	22.09898077006698	51.16868660491898
N	15.13290234532604	23.88633023584807	50.49502177787439
C	14.81408835318705	23.19016470314914	49.41507724830802
N	14.15953632853470	22.10600324286542	49.79366352723581
H	15.63951585789271	24.76861778433484	50.49427707395099
H	13.57615426665894	21.30421379998455	51.71605287495182
H	15.05730317187727	23.48088500996394	48.40257322704412
H	15.06517982338439	19.37012615487460	55.14312792070102
H	14.21791082680259	17.82546357237391	54.87891547952015

H	14.06219959800328	19.15755678463678	53.68083096159232
H	14.46027937611891	24.76608566794100	53.12470706885959
H	15.97567366333916	23.84028893927660	53.23256752873154
H	14.43934681733501	23.09625347077678	53.74255688806203
H	8.81127821181246	22.03458181760766	52.65704593934154
H	9.80855882360406	21.84053501428739	51.18769800877842
H	9.78988556188891	20.55645618995088	52.45267269113356
H	8.77066052631649	18.47803039803196	50.41951073928043
H	16.46952075976441	16.69423937208215	51.64399130464172
H	15.52570834694771	22.56283895347873	43.20637962059522
H	16.46757873217417	21.27341808742760	42.51265275245503
H	17.00944369099459	20.38474345825530	44.48555830767250
H	13.65978859049151	20.94335380107652	43.72760750334833
H	14.62869770518522	19.63522099830780	43.00346677504422
H	14.09186688240739	21.00687981988870	42.00618666733717
H	18.75179714786337	20.18848156695133	43.46641882317250
H	17.58614293903090	23.60605898671964	45.47120896733447

E. Cyt P460 Phlorin Final Geometry Optimized Structure Coordinates

Charge: 0 Spin Multiplicity: 6

N	11.45814301769548	20.21883725003570	48.53107958446778
N	15.46289107918954	19.76012634119198	47.96168319010378
C	14.49986910889060	21.77780038051245	46.01893561332789
C	10.63255235674787	21.05769498744453	47.81661282226158
C	12.44417544525770	18.21508403852281	50.58483495323287
C	16.28756497180186	19.07316226251987	48.85224218843981
C	14.17050872281402	23.05833352405235	45.41972229829599
C	9.27439072019070	20.72116114624292	48.19744865949995
C	12.93077704751309	17.37617582075111	51.66931173252765
C	17.63756737563331	19.10879679658136	48.31222396061023
C	12.87265028813118	23.31681602367432	45.72291742505044
C	9.31993193538930	19.77357163733655	49.12987244711302
C	14.25771821610474	17.27059221552963	51.50820855904824
C	17.57494365173135	19.91459291160628	47.05371456999278
C	12.36510259746751	22.20560880771215	46.49325817314636
C	10.71276095038783	19.42671511841606	49.36262098688792
C	14.62715165052316	18.17705604246929	50.43174757039199
C	16.19013129208535	20.28838766398107	46.91598215745249
C	15.14697445472045	24.00626296251794	44.71026867055840
C	8.15660878754082	19.09868897198139	49.92049729321738
C	15.31035298359896	16.50953354284563	52.36171587836411
C	18.80311641386653	20.29947263571087	46.21291181544811
C	15.65152649673039	25.02036606886023	45.76186941276955
C	7.02135611383283	18.66083684930967	48.98926483606995
C	14.78475760295468	15.30230223674148	53.18699890312142
C	19.54146612544188	19.06341400763771	45.68413986675576
C	16.66325620408624	24.43624402983469	46.75448537306651
C	18.72706128977191	18.33546394458939	44.64272185543330
C	15.67250470759499	21.07325537622292	45.92447351194071
C	11.05302585515415	22.03202641174530	46.91519154806164
C	11.14032472529229	18.52191853541707	50.31305949878840
C	15.91624652919881	18.38665751666469	49.99783672678429
C	12.03926502263380	24.53748720126013	45.32268265786061
C	7.98674097393557	21.43515236525322	47.65391975953655
C	12.01837820093197	16.54608361498431	52.59283503011785
C	18.93027531330701	18.50274653834479	48.94881823025509
N	13.38563908753325	21.29197320843666	46.67667371944497
N	13.49942650430992	18.74497365266159	49.87412531393847
O	16.65452973319123	24.90154215391320	47.91876820041866
O	18.61042228024362	17.08344440050718	44.77296549696879
O	17.45677223823617	23.51405143302753	46.39990235560023
O	18.18037440704612	19.00032051619544	43.69414123802653
Fe	13.51714707254595	20.22678613112776	48.52102377927591
C	14.31106645786103	20.38568493284188	43.07309591968077

C	15.60045492965399	21.12687013801969	43.32450765609106
N	16.23133473577267	20.81573974648988	44.65330054755770
H	16.36261282383472	20.82856067118742	42.58980018351873
H	15.45125047993197	22.20130990754624	43.19822427685857
C	8.86266084333144	21.03211874502397	51.91165017125288
S	7.47123278505527	20.24633081942789	51.09121413553863
N	13.55819544629576	19.87509000978294	56.22481040393060
C	14.79795850180884	19.45084793388315	55.57451644560038
C	15.59155528472780	20.64275412884519	55.05821637560065
O	16.79661816761168	20.75098912505963	55.29924673908816
C	14.49342873140916	18.48469250683928	54.43993885757413
S	16.00013138167736	17.94938256712489	53.67556958211659
H	12.85007789026056	20.17343579444037	55.55055458704236
H	15.44183669630266	18.96367752892484	56.32028535646119
H	13.94802124722208	17.63089906684736	54.86935500500321
H	13.83657427653583	18.96927262379726	53.69872649828704
N	14.92501657476886	21.58050836047899	54.37980488087907
C	15.62761232630814	22.76844119567263	53.89650498192828
C	16.20906164875738	23.57543377672560	55.06497776934194
O	17.35629583905040	24.03488345315689	55.00122273649913
C	14.68202139756076	23.61545164062879	53.03204671147738
C	14.44668878837502	23.04740629439889	51.65770866909626
N	14.92496968739671	23.65984640910127	50.52806602963103
C	13.82009631558726	21.91798233680208	51.23804165627227
C	14.59556712859467	22.93766185324703	49.46263727914217
N	13.94502930997242	21.86674834059111	49.87000663309902
H	13.90875517731348	21.57830273068599	54.37399083430057
H	16.49284873194002	22.45222464386756	53.29628903672432
H	13.72490311141829	23.74780558087519	53.56016869146336
H	15.11392917215958	24.62087415812790	52.92543561413335
H	15.43112079726122	24.54183214748215	50.52278621482500
H	13.31264631508857	21.14011329306166	51.79439327943529
H	14.82692781195982	23.18479795358634	48.43508139774338
H	14.65593903451256	24.54355595749199	43.88794204943515
H	16.02181499979548	23.48279840376727	44.31646530182048
H	8.54412901291568	18.21541869079894	50.44309113559060
H	16.22428798986130	16.28643937521748	51.80574914681990
H	18.50007495019254	20.93978023927689	45.37774356863407
H	19.50396706871952	20.87320621548847	46.84104257936843
H	14.80868351830616	25.47340607097594	46.30547966682295
H	16.18133801925356	25.85431294951760	45.26369553040464
H	6.55177490496356	19.50925208322539	48.48261835010757
H	6.24097852188277	18.12828116606115	49.55485028953161
H	7.41476230640729	17.97129857777309	48.22905466550854
H	14.06567211319295	15.60345568800320	53.95317644828410
H	15.62946257508955	14.80135748018886	53.67762384677637

H	14.29702752130357	14.58004463508604	52.51489659925771
H	20.47699192055126	19.37711822479363	45.18969203596502
H	19.82879844889137	18.39014180848105	46.50333243742553
H	10.29723152882906	22.71589114438346	46.52726744450013
H	10.37456221903475	18.04008642522281	50.91646324502050
H	16.71977223433078	17.89246428941810	50.54208704455679
H	12.51759467167630	25.08941997317729	44.50264395286276
H	11.91209806513135	25.23402196661140	46.16608021779809
H	11.03442784206187	24.23861233450147	44.99118896570413
H	8.24340981997641	22.41515412235893	47.23016279993521
H	7.26713930574070	21.58346202562959	48.46838385438506
H	7.50973578728586	20.84116477195744	46.86163943012675
H	10.96744908508259	16.82974480926820	52.45253188090243
H	12.25251690853115	16.70083802419153	53.65312921699903
H	12.10538418457734	15.47334262442041	52.37705563384295
H	19.79780537316104	19.14325380857380	48.74007060395030
H	19.14154646994693	17.49679238481634	48.55419100479331
H	18.82352455690722	18.41730399771441	50.03860247392300
H	16.03328929454333	25.65495054696656	48.00804327459458
H	19.09309327221363	16.75184298072939	45.55757459058642
H	16.64348601292795	19.88669111788255	44.58851350575495
H	13.72698228570306	20.65432955460430	56.86373019607026
H	15.56025952802003	23.72712595501941	55.95661493113188
H	8.41619728667066	21.75289411966089	52.60938414376265
H	9.49703419024422	21.57153756581343	51.19685715854077
H	9.46845448517761	20.31699838814408	52.48319321187883
H	13.90927805798692	20.61827443355314	42.07199113387934
H	13.54004805366387	20.65516785952850	43.80858149949506
H	14.46401505000421	19.29663539684976	43.12510427823587
H	16.33020116067241	21.97872412391185	46.22083763414771

F. Metmyoglobin Final Geometry Optimized Structure Coordinates

Charge: 1 Spin Multiplicity: 6

C	9.28897850642263	30.57590764789195	2.62311241604898
C	10.51936351733631	30.54526975272064	3.50179839491705
N	10.72144925416969	31.42071779902308	4.58198591399937
C	11.62711383819520	29.75914493110037	3.45261517252573
C	11.91005658084354	31.16827709738319	5.11039892374998
N	12.47269942024251	30.16811359693119	4.48543251367781
Fe	14.39188639631337	29.39926883540069	4.95270446519012
C	15.58278602930120	32.61903348860463	5.17240119857480
C	13.79872972757637	29.38760898827010	8.32992821063653
C	13.27293461393890	26.25155947733351	4.73227147698716
C	15.30495214610685	29.31513629501687	1.63038062984452
N	14.67644146415583	30.71952056500265	6.42943388828881
C	15.05487233823906	32.04990480561606	6.28344470591801
C	14.92946984952410	32.69095894573973	7.58449420858340
C	14.42454275095924	31.81001050915982	8.44640109677433
C	14.28261414566128	30.54336034550289	7.74513963832979
C	14.01294239235211	32.03273482367042	9.91831584404887
C	15.26166657175453	34.19206812227735	7.80547968565653
C	14.02418157131935	35.10995651604109	8.06535999418672
C	13.12403431151161	35.08210470187704	6.79062739986240
O	13.70599175729734	35.44725022805958	5.71661283917659
O	11.91781802137736	34.69307623645371	6.95874378972980
N	13.72392913249072	28.07789970863645	6.26783401425741
C	13.45571286658767	28.25380874682848	7.62414889116555
C	12.79493064949954	27.11894856157678	8.16941875467571
C	12.65352431900392	26.23347536472393	7.19840769803142
C	13.22671078487817	26.80769858316744	5.99550330584499
C	12.40037703829624	26.99163805838407	9.71415546106086
C	12.00650066600409	24.84796627115000	7.22255202574809
C	12.61447024287459	23.82381782000901	8.15671364851108
N	14.31345177616352	28.04219389511418	3.44557088592861
C	13.82207531084685	26.74374088841589	3.56524107721013
C	14.06434417781641	26.01273167742262	2.33148448735906
C	14.58832751307504	26.90270831564470	1.44864075883879
C	14.77512794049174	28.17936640119515	2.19004397440241
C	13.62115281640558	24.56054737071062	2.09619709537035
C	15.06684689967650	26.69902043533161	-0.00106005264677
C	15.24535600539174	25.30747104675258	-0.48343833316283
N	15.21881067285534	30.70774952086711	3.65204480992129
C	15.56795568852896	30.46687577965933	2.33593098386173
C	16.22078437556912	31.62374366089444	1.76965153565940
C	16.28511110895698	32.53514524804313	2.76044778834319
C	15.65398636446574	31.98689568859490	3.93619477497854
C	16.67405905029879	31.78904449247693	0.30811682921713

C	16.99415030568379	33.91570404922253	2.73248679001661
C	18.46307991517213	33.64154294269228	3.15972851579993
C	19.46227002745465	34.74330891154256	3.15245737029303
O	19.83018000177278	35.23530200897504	2.06188181299821
O	19.93962816616817	35.12894462912189	4.26615180987577
O	16.32103569466152	28.42960751491791	5.15087335201796
H	11.88382705760795	28.95059039008287	2.78028754764988
H	12.31844179440620	31.69723659114519	5.96354935330651
H	15.91644617832901	33.65216064439703	5.22939051296779
H	13.55948060179067	29.42738527077327	9.39098985730180
H	12.88416837666127	25.23411242794914	4.66003538087161
H	15.61137721261977	29.28135941634012	0.58616733644839
H	14.45386325936009	32.95849081544868	10.30774933043650
H	14.34609110734598	31.19923808316198	10.55130905952873
H	12.91878734918454	32.10507272283027	10.00509432607181
H	15.80566190312255	34.58866561804646	6.94252159023024
H	15.93123513477273	34.27727853602243	8.67553006335697
H	13.47796321031476	34.79311471048142	8.96594569923857
H	14.36400814237156	36.14409878154260	8.21086184478007
H	11.74263758260553	34.50227960600714	7.90246622129965
H	11.62929040818398	26.22327545959732	9.83936918525887
H	12.01386048488754	27.94785020927773	10.08639283557726
H	13.28489069671681	26.71164965433330	10.30195142965093
H	11.99497881044129	24.44501903456412	6.20099147575691
H	10.94388997171817	24.98515803244296	7.49116918868649
H	13.66550276824553	23.61909727319431	7.90205979036789
H	12.06307363439078	22.87496983586421	8.08820361951132
H	12.58593402375062	24.14977698038471	9.20563909791876
H	12.88942136814952	24.50267735586976	1.27968888073879
H	13.15422643059217	24.14463782524646	2.99695461333980
H	14.47149329054059	23.91808366655093	1.83538097048153
H	14.36103538439577	27.23674745851878	-0.66087800838731
H	16.01722915932291	27.24769582135714	-0.10500653627181
H	14.30367516861445	24.74547034603069	-0.53261408112658
H	15.94320309471826	24.73547080574769	0.14608386020127
H	15.66481780012260	25.31619528477422	-1.50087395708766
H	15.83768247794809	31.59537066136070	-0.37843179385076
H	17.48145772936629	31.08530632226768	0.05798183085064
H	17.03833731573915	32.80784149291827	0.13229423206423
H	16.51052752530612	34.62403867288815	3.41706943288836
H	16.96255819064494	34.34454704460860	1.72241661336088
H	18.86466761488946	32.87965820681704	2.47422662780557
H	18.43493837690371	33.16304468132169	4.15285899259999
H	19.55248075216162	34.64245656463459	5.02197854654326
H	16.22991442660070	27.70359005564842	5.79941471175495
H	10.07613293492636	32.15050121111427	4.87209310970848

H	9.37475782649983	29.79732454839294	1.85630898310191
H	9.18280302894910	31.54959519363692	2.12308061845071
H	8.37536301863054	30.39279561368287	3.20715076700719
H	17.01833266046178	28.99939977157672	5.53295436732475

G. Cyt P460 Porphyrin, Meso C–C Final Geometry Optimized Structure Coordinates

Charge: 1 Spin Multiplicity: 6

C	15.32887429176159	24.36878195768493	44.86996742029274
C	8.40111020229028	19.35700154729516	49.88314365914517
C	15.60369851970375	16.88751937879049	52.28666352576826
C	18.95184109541857	20.52181502226930	46.18812436907476
N	13.59104146726113	21.71468114332524	46.77830044358976
C	15.88917100838439	25.40170403672806	45.81847398265713
C	7.22036929242960	19.02039024855013	49.01421998430732
C	15.19119548996333	15.59949362000331	53.04163055829727
C	19.69960381013298	19.36345927136492	45.50965865274809
N	11.70418515143125	20.52097311217057	48.53128693658001
C	16.88523630474625	24.72527857394513	46.78771041746833
C	18.90057650957576	18.68967560398705	44.58092489640479
N	15.67454518523783	20.08506482270807	47.97124426198857
C	15.84820168363462	21.32875159830266	45.80333069669280
C	11.26970164907343	22.39187737838645	47.06082648601786
C	11.39485095964069	18.76228648002991	50.18337463124729
C	16.13961088247332	18.66638156635784	49.87247908636448
C	12.26948889512649	24.83517145681760	45.34954987759319
C	8.28970932529708	21.86582565896261	47.75761653905506
C	12.17172779166734	16.84083732445490	52.39155460648967
C	19.06213851169962	18.85610736245007	48.81613932263406
C	14.68888166197306	22.10825661369121	45.99736694988822
C	10.84244414339495	21.43983996363212	47.82839959088097
C	12.60680302613152	18.43283867187381	50.50809719787429
C	16.53247990061724	19.35701119062739	48.78446945420870
O	16.88524161108248	25.15916378330042	47.98547728605048
O	18.69269685779417	17.42372128517811	44.63982062330695
C	14.33349877475117	23.39850968714042	45.43483557381844
C	9.51344777627254	21.10206503787577	48.25156950889497
C	13.19276731987524	17.60566915121123	51.49331001621589
C	17.80723489413961	19.49114626341091	48.17897176316161
O	17.65837661981082	23.81934146942767	46.42967111239541
O	18.35547787070160	19.30176414292220	43.62431981930165
C	13.03082792019315	23.62594397678152	45.72898862000508
C	9.55718643936525	20.06148508795354	49.12771342669371
C	14.49875114075444	17.60676897122528	51.46556249029170
C	17.78202836908896	20.14625037967946	46.99413870906763
C	12.56357407033713	22.58475635795719	46.59441890634108
C	10.88001349779367	19.66316706423325	49.32825374837586
C	14.89470519650520	18.42145588217583	50.34676268970894
C	16.43423980819842	20.58254564312900	46.81467343018014
N	13.70382960317051	18.97215945822719	49.86841977683461
Fe	13.70109648978666	20.48171578987426	48.42207192290008
H	14.85425464591870	24.90532923713250	44.02961009326737

H	16.15874558576050	23.83827654276403	44.37957203442895
H	19.68793235502677	21.00078945229206	46.85592925189997
H	18.71850858213030	21.29269215642253	45.45108345650058
H	15.13114828580625	25.90090487174378	46.42883146911075
H	16.44533194387600	26.16477757233434	45.25145166883159
H	6.42965032456999	18.53609154337524	49.60695361088428
H	7.51867511904043	18.31339981215660	48.22484262252744
H	6.78926093252514	19.90504733025087	48.53496205656009
H	14.69153100415728	14.88831505699036	52.36573326067910
H	14.51731311464526	15.81400620400778	53.87804605321471
H	16.09188904240223	15.11878394189732	53.44501025578339
H	20.07539371456851	18.65357297523360	46.25237454093305
H	20.57821981670081	19.79331217567843	44.99588384302914
H	10.53016166050142	23.09675077518349	46.67153937840816
H	10.62811120071847	18.23064611940508	50.75948330510375
H	16.94541948066370	18.15642262738481	50.40500270650362
H	12.25950566865336	25.58537746359060	46.15755850485657
H	11.21783763276605	24.59919278023662	45.12956306892472
H	12.69573002274658	25.31769661749813	44.46007334558870
H	8.57000957686794	22.86736722784584	47.40699124361361
H	7.54984892217934	21.97523957110721	48.56087328774649
H	7.80970761225719	21.34219649578078	46.91688064126770
H	11.50548912094856	17.55510711003424	52.89527364855851
H	12.65021372959796	16.22510610193253	53.15655943168432
H	11.55081937514502	16.18638420977848	51.76304201230254
H	19.97344341092270	19.39829339115946	48.53142414545253
H	19.17355450634208	17.81031488899024	48.49137042774164
H	18.99427893511080	18.86253617133814	49.91137096805785
C	14.45215452619872	20.72269606547762	43.00405957018899
C	15.71687422512928	21.49127304498307	43.28968050411191
C	16.40839655772897	21.20053222441061	44.56378132519932
C	9.19270625741137	21.31083150135957	51.92712424878300
S	7.71373908319531	20.58526593677428	51.15270282963670
C	14.73918705983090	18.65046844243984	54.37926475169994
S	16.24832730180648	18.10681344883940	53.55943675132647
C	14.90112170095259	23.76477010529494	53.00916673739386
C	14.67393261586109	23.22838427827649	51.61935642165874
C	14.06585024826101	22.09861405083264	51.16971693827950
N	15.13176042585482	23.88602622572978	50.49624385151414
C	14.81263647919930	23.19015308087043	49.41620525948512
N	14.15806738915410	22.10596354499859	49.79466764120973
H	15.63847481616552	24.76829275452866	50.49554669991754
H	13.57395695661251	21.30495151339713	51.71772118636920
H	15.05535238772420	23.48242378201758	48.40425235609168
H	15.06516272711493	19.36804899881168	55.14360053175184
H	14.21687553842321	17.82410129799609	54.87834321450668

H	14.06146075505774	19.15781296013432	53.68181004437938
H	14.46127099810067	24.76592809216826	53.12534233848421
H	15.97521587137594	23.83777240191684	53.23384980368701
H	14.43745504528023	23.09611294334537	53.74359157946974
H	8.80977893438799	22.02915760970168	52.66402521867053
H	9.80425263462127	21.84709237571429	51.19180815452094
H	9.79166092037519	20.55519570845827	52.44963900205210
H	8.76885424053524	18.47805131791534	50.42041263371582
H	16.46744257353472	16.69280512521047	51.64270650498268
H	15.50113599742495	22.56089402184300	43.16404396748909
H	16.44141824161640	21.27063271735907	42.48799144076000
H	13.64983112011319	20.95234512837377	43.72089384933129
H	14.62184219518284	19.63434795673389	43.04020791250357
H	14.07331411525973	20.96417366307454	42.00085480141833
H	18.65350484458390	20.23319577038048	43.54134507447156
H	17.61132267112342	23.64364082038905	45.46343547246180
H	17.36997744459254	21.75641661299833	44.47975430811407
H	16.72373875897321	20.15650101758740	44.47840377008917

H. Cyt P460 Porphyrin, Meso C–H

Charge: 1 Spin Multiplicity: 6

C	15.33169568854011	24.35931048226725	44.86311973853766
C	8.40237377116292	19.35593070599260	49.88253034220220
C	15.60419136998979	16.88681362810246	52.28873326980871
C	18.95337245872210	20.51265955690876	46.18573640635451
N	13.59301011379983	21.70809508666110	46.77468412539464
C	15.89238692850176	25.39316931883719	45.81037174848926
C	7.22148361142689	19.01869308838918	49.01405245804680
C	15.19126813406520	15.59983385864568	53.04525279982268
C	19.70070692348434	19.35323436713503	45.50862690376425
N	11.70580347118324	20.51713771307434	48.52916378973554
C	16.88825260033586	24.71755501465299	46.78037877090070
C	18.90141079144806	18.67862031831509	44.58072764949215
N	15.67599082826236	20.07917930830411	47.96949656082326
C	15.84999978035883	21.32021618720404	45.80009319000501
C	11.27191655048013	22.38643600338931	47.05648641090439
C	11.39591852047550	18.76053269643719	50.18336154050120
C	16.14063326326170	18.66260538234132	49.87240692709828
C	12.27249053889206	24.82733696648307	45.34225704957912
C	8.29176716830189	21.86225375648007	47.75401323807842
C	12.17220849385346	16.84145123954737	52.39380572841206
C	19.06318760104714	18.85005266403979	48.81573432955044
C	14.69095826584405	22.10035574499233	45.99324103706871
C	10.84435621971341	21.43546450477973	47.82521136721657
C	12.60776790331203	18.43105108251392	50.50843284757141
C	16.53370227310549	19.35179865663662	48.78355905559633
O	16.88845321833086	25.15287010871715	47.97762668801331
O	18.69309286482610	17.41280964080313	44.64114257916845
C	14.33600353288844	23.39005983148376	45.42918248845847
C	9.51525807321452	21.09865762338147	48.24883297317944
C	13.19348082534358	17.60485462338803	51.49461115335507
C	17.80848143269078	19.48476701885891	48.17785494219100
O	17.66106433038367	23.81092197534871	46.42339321091396
O	18.35648967396317	19.28975583943331	43.62341237765926
C	13.03342280041258	23.61829852355446	45.72311149120364
C	9.55866709494977	20.05910937710724	49.12621713529312
C	14.49946392102219	17.60546680640402	51.46681449660791
C	17.78345897632413	20.13846470824669	46.99224145324438
C	12.56583873142735	22.57830765102248	46.58980147039394
C	10.88136289674614	19.66057073829055	49.32718445785318
C	14.89565999601585	18.41867943498130	50.34702823946014
C	16.43581566866168	20.57501440481581	46.81230472399571
N	13.70495838593866	18.96922588808836	49.86807174003080
Fe	13.70269698101936	20.47705505664762	48.41992258094632
H	14.85288167704890	24.89611185829242	44.02560373984788

H	16.14800867469553	23.80967228853633	44.36087615739162
H	19.68127409772648	21.03619096305209	46.82581688188706
H	18.66876587106548	21.25444184156428	45.42338902125924
H	15.13056092410359	25.88905621981622	46.41849808625886
H	16.44731430827075	26.15703524805473	45.24359208970051
H	6.43125019519455	18.53488989462392	49.60774134976158
H	7.51950579459313	18.31107713819999	48.22518423328079
H	6.79003146947493	19.90296286733869	48.53441494007723
H	14.69116271439662	14.88806382131533	52.37035173438087
H	14.51792749138643	15.81556447168540	53.88177900132548
H	16.09206749735002	15.11945815616492	53.44868091393204
H	20.07836625682546	18.64437398894966	46.25164559076759
H	20.57832424223172	19.78320483410823	44.99357879708293
H	10.53382103383830	23.09408103801635	46.66977068949218
H	10.62903140316144	18.22993791628519	50.76022681866746
H	16.94695470394366	18.15664687436178	50.40779935149768
H	12.25657496469482	25.57385083692231	46.15351704158114
H	11.22257516890591	24.59127447099205	45.11425955087505
H	12.70434046064481	25.31346034644807	44.45765462364987
H	8.57229179730069	22.86330194327554	47.40232911686786
H	7.55195292992295	21.97245138906048	48.55712344943065
H	7.81225929858958	21.33741345140784	46.91380242246903
H	11.50653336331459	17.55671565720883	52.89672114905450
H	12.65073357058863	16.22651996738014	53.15933336972632
H	11.55133439904199	16.18654911453611	51.76583132021781
H	19.97293144621424	19.39693053283483	48.53511008951797
H	19.17867912548728	17.80613105841398	48.48640948364110
H	18.99213635250352	18.85117031118430	49.91074780143091
C	9.19472538271609	21.31192418263925	51.92414756824724
S	7.71547713131188	20.58594922470932	51.15064721496387
C	14.74037094349715	18.65256077110075	54.37925967572000
S	16.24929154486447	18.10740211117041	53.56002611532003
C	14.90403447069369	23.76516621607810	53.00305013993914
C	14.67660726503162	23.22720042907270	51.61388959210053
C	14.06811514542106	22.09710582793566	51.16562166946056
N	15.13462228326037	23.88334153326799	50.48997581679098
C	14.81521619907109	23.18629041622048	49.41078056988498
N	14.16028389862910	22.10278140729224	49.79056120015616
H	15.64178620201149	24.76537732886098	50.48801664464968
H	13.57582558082411	21.30462526949165	51.71494570378223
H	15.05909851456303	23.47770185406847	48.39898301398507
H	15.06703502758939	19.37005026621274	55.14337159925912
H	14.21746366127586	17.82676154362578	54.87860109538026
H	14.06291617033454	19.16060385680237	53.68202102208332
H	14.46489770461682	24.76679201492946	53.11768701684641
H	15.97820211430612	23.83768424396477	53.22747158414059

H	14.43987232047332	23.09779098456656	53.73829843858334
H	8.81171949725549	22.02917656292916	52.66205658350163
H	9.80512889130443	21.84991850358114	51.18918341511820
H	9.79450624065491	20.55641658378130	52.44587318818655
H	8.76982712824488	18.47738449321813	50.42063907801951
H	16.46765134252899	16.69078380282425	51.64477747731757
H	18.59916625585888	20.23932811203952	43.57453261699303
H	17.59764192695351	23.62076740849868	45.46147446616312
H	16.49227325046478	21.61133996838450	44.96587719141105

I. Heme Restraints Files

Aromatic

```
data_comp_list
loop_
_chem_comp.id
_chem_comp.three_letter_code
_chem_comp.name
_chem_comp.group
_chem_comp.number_atoms_all
_chem_comp.number_atoms_nh
_chem_comp.desc_level
HEC        HEC 'Unknown                ' ligand 73 43 .
#
data_comp_HEC
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.charge
_chem_comp_atom.partial_charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
HEC        CAA    C   CH2    0    .     26.7957   21.7332   23.4902
HEC        CAB    C   CSP    0    .     34.5133   21.1849   21.5785
HEC        CAC    C   C1     0    .     31.0303 14.1833 21.3988
HEC        CAD    C   CH2    0    .     26.4967 16.6802 23.7745
HEC        NA     N   N      0    .     28.3657 20.7245  20.5594
HEC        CBA    C   CH2    0    .     25.5733   22.0660   22.5895
HEC        CBB    C   CH3    0    .     35.2964   21.3388   20.2555
HEC        CBC    C   CH3    0    .     31.2287   12.6802   21.0785
HEC        CBD    C   CH2    0    .     26.6598   16.8673   25.3055
HEC        NB     N   N      0    .     31.6513 20.4597 19.7524
HEC        CGA    C   C      0    .     24.4243   22.6094   23.4624
HEC        CGD    C   C      0    .     25.4431   16.2587   26.0301
HEC        ND     N   N      0    .     28.4622 17.3212  21.1319
HEC        CHA    C   C      0    .     27.2842 19.2110 22.3734
HEC        CHB    C   C1     0    .     30.3615 22.1772 21.2828
HEC        CHC    C   CR16   0    .     33.3692 18.6270 20.8308
HEC        CHD    C   C1     0    .     30.0717 15.3693 20.6185
HEC        CMA    C   CH3    0    .     28.4030 23.2362 22.9518
HEC        CMB    C   CH3    0    .     33.0227 22.8689 21.8822
HEC        CMC    C   CH3    0    .     34.1513 15.2537 21.0630
```

```
HEC      CMD   C   CH3   0   .     28.0279 14.6214 23.2514
HEC      C1A   C   CR56  0   .     27.5128 20.3115     21.7956
HEC      C1B   C   CR5   0   .     31.3939 21.6793     20.7639
HEC      C1C   C   CR56  0   .     32.7615 17.2261     20.7714
HEC      C1D   C   CR5   0   .     29.1111 16.0373     21.4257
HEC      O1A   O   O     0   .     24.5548 23.7131  24.0617
HEC      O1D   O   O     0   .     24.2803 16.6777 25.7719
HEC      C2A   C   CR5   0   .     27.9355 21.1148     22.6157
HEC      C2B   C   CR5   0   .     32.3395 21.5275     21.5088
HEC      C2C   C   CR5   0   .     32.7406  15.8885    21.0832
HEC      C2D   C   CR5   0   .     28.5818 15.9894     22.7819
HEC      O2A   O   OC    0   .     23.3313 21.9805     23.5294
HEC      O2D   O   OC    0   .     25.6142 15.3824     26.9234
HEC      C3A   C   CR5   0   .     28.6916 21.8286     22.3627
HEC      C3B   C   CR5   0   .     33.0526 20.7252     21.3635
HEC      C3C   C   CR5   0   .     31.7991 15.1083     20.4339
HEC      C3D   C   CR5   0   .     27.8284 17.0583     23.0599
HEC      C4A   C   CR5   0   .     29.1707 21.8804     21.2453
HEC      C4B   C   CR56  0   .     32.9407 19.9113     20.4209
HEC      C4C   C   CR5   0   .     30.9414 15.7670     19.5084
HEC      C4D   C   CR56  0   .     27.6624 17.9630     22.1489
HEC      NC    N   N     0   .     31.5611   17.1264   19.7194
HEC      FE    FE  FE    0   .     29.9321   18.8680   19.9935
HEC      HMD3  H   HCH3  0   .     27.1509 14.3616     22.6587
HEC      HMD2  H   HCH3  0   .     27.7505 14.6835     24.3015
HEC      HMD1  H   HCH3  0   .     28.7918 13.8573     23.1228
HEC      HMC3  H   HCH3  0   .     34.7785 15.7389     21.8097
HEC      HMC2  H   HCH3  0   .     34.0745 14.1915     21.2885
HEC      HMC1  H   HCH3  0   .     34.5942 15.3843     20.0765
HEC      HMB3  H   HCH3  0   .     32.2972 23.5250     22.3598
HEC      HMB2  H   HCH3  0   .     33.4047 23.3457     20.9798
HEC      HMB1  H   HCH3  0   .     33.8469 22.6780     22.5689
HEC      HMA3  H   HCH3  0   .     28.4785 23.1977     24.0392
HEC      HMA2  H   HCH3  0   .     29.1293 23.9487     22.5625
HEC      HMA1  H   HCH3  0   .     27.4003 23.5493     22.6687
HEC      HBD2  H   HCH2  0   .     26.7257 17.9288     25.5363
HEC      HBD1  H   HCH2  0   .     27.5676 16.3668     25.6376
HEC      HBC3  H   HCH3  0   .     30.2867 12.1519     21.2180
HEC      HBC2  H   HCH3  0   .     31.9815 12.2628     21.7464
HEC      HBC1  H   HCH3  0   .     31.5578 12.5692     20.0459
HEC      HBB3  H   HCH3  0   .     35.9001 20.4488     20.0854
HEC      HBB2  H   HCH3  0   .     35.9459 22.2119     20.3172
HEC      HBB1  H   HCH3  0   .     34.5969 21.4649     19.4328
HEC      HBA2  H   HCH2  0   .     25.8576 22.8176     21.8548
HEC      HBA1  H   HCH2  0   .     25.2440 21.1641     22.0769
HEC      HAD2  H   HCH2  0   .     25.6954 17.3239     23.4135
```

```
HEC      HAD1   H   HCH2   0   .     26.2521   15.6421   23.5602
HEC      HAA2   H   HCH2   0   .     26.4979   21.0183   24.2570
HEC      HAA1   H   HCH2   0   .     27.1552   22.6429   23.9634
HEC      HHD   H   H     0   .     29.7796 14.3277 20.4948
HEC      HHC   H   HCR6   0   .     34.4510 18.5575 20.7343
HEC      HHB   H   H     0   .     30.4643 23.2614 21.3276
HEC      HAC   H   H     0   .     31.3476 14.3812 22.4216
#
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
_chem_comp_bond.value_dist_neutron
HEC   CAA    CBA  single       1.554 0.020    1.554
HEC   CAA    C2A  single       1.564 0.020    1.564
HEC   CAA    HAA2  single       0.939 0.020    1.096
HEC   CAA    HAA1  single       0.939 0.020    1.096
HEC   CAB    CBB  single       1.545 0.020    1.545
HEC   CAB    C3B  single       1.546 0.020    1.546
HEC   CAC    CBC  single       1.550 0.020    1.550
HEC   CAC    C3C  single       1.542 0.020    1.542
HEC   CAC    HAC  single       0.951 0.020    1.111
HEC   CAD    CBD  single       1.551 0.020    1.551
HEC   CAD    C3D  single       1.558 0.020    1.558
HEC   CAD    HAD2  single       0.939 0.020    1.096
HEC   CAD    HAD1  single       0.939 0.020    1.096
HEC   NA     C1A  aromatic     1.558 0.020    1.558
HEC   NA     C4A  aromatic     1.567 0.020    1.567
HEC   NA    FE    single       2.494 0.500    2.494
HEC   CBA    CGA  single       1.542 0.020    1.542
HEC   CBA    HBA2  single       0.939 0.020    1.096
HEC   CBA    HBA1  single       0.939 0.020    1.096
HEC   CBB    HBB3  single       0.939 0.020    1.096
HEC   CBB    HBB2  single       0.939 0.020    1.096
HEC   CBB    HBB1  single       0.939 0.020    1.096
HEC   CBC    HBC3  single       0.939 0.020    1.096
HEC   CBC    HBC2  single       0.939 0.020    1.096
HEC   CBC    HBC1  single       0.939 0.020    1.096
HEC   CBD    CGD  single       1.541 0.020    1.541
HEC   CBD    HBD2  single       0.939 0.020    1.096
HEC   CBD    HBD1  single       0.939 0.020    1.096
HEC   NB     C1B aromatic     1.605 0.020    1.605
HEC   NB     C4B aromatic     1.552 0.020    1.552
```

```
HEC   NB    FE    single      2.355 0.500    2.355
HEC   CGA   O1A   deloc       1.263 0.020    1.263
HEC   CGA   O2A   deloc       1.263 0.020    1.263
HEC   CGD   O1D   deloc       1.263 0.020    1.263
HEC   CGD   O2D   deloc       1.263 0.020    1.263
HEC   ND    C1D   aromatic    1.468 0.020    1.468
HEC   ND    C4D   aromatic    1.444 0.020    1.444
HEC   ND    FE    single      2.418 0.500    2.418
HEC   CHA   C1A   aromatic    1.264 0.300    1.264
HEC   CHA   C4D   aromatic    1.323 0.300    1.323
HEC   CHB   C1B   aromatic    1.258 0.300    1.258
HEC   CHB   C4A   aromatic    1.228 0.300    1.228
HEC   CHB   HHB   single      0.951 0.200    1.111
HEC   CHC   C1C   aromatic    1.528 0.300    1.528
HEC   CHC   C4B   aromatic    1.415 0.300    1.415
HEC   CHC   HHC   single      0.951 0.020    1.111
HEC   CHD   C1D   aromatic    1.422 0.300    1.422
HEC   CHD   C4C   aromatic    1.465 0.300    1.465
HEC   CHD   HHD   single      0.951 0.020    1.111
HEC   CMA   C3A   single      1.553 0.020    1.553
HEC   CMA   HMA3  single      0.939 0.020    1.096
HEC   CMA   HMA2  single      0.939 0.020    1.096
HEC   CMA   HMA1  single      0.939 0.020    1.096
HEC   CMB   C2B   single      1.551 0.020    1.551
HEC   CMB   HMB3  single      0.939 0.020    1.096
HEC   CMB   HMB2  single      0.939 0.020    1.096
HEC   CMB   HMB1  single      0.939 0.020    1.096
HEC   CMC   C2C   single      1.547 0.020    1.547
HEC   CMC   HMC3  single      0.939 0.020    1.096
HEC   CMC   HMC2  single      0.939 0.020    1.096
HEC   CMC   HMC1  single      0.939 0.020    1.096
HEC   CMD   C2D   single      1.549 0.020    1.549
HEC   CMD   HMD3  single      0.939 0.020    1.096
HEC   CMD   HMD2  single      0.939 0.020    1.096
HEC   CMD   HMD1  single      0.939 0.020    1.096
HEC   C1A   C2A   aromatic    1.223 0.020    1.223
HEC   C1B   C2B   aromatic    1.213 0.020    1.213
HEC   C1C   C2C   aromatic    1.374 0.020    1.374
HEC   C1C   NC    aromatic    1.599 0.020    1.599
HEC   C1D   C2D   aromatic    1.457 0.020    1.457
HEC   C2A   C3A   aromatic    1.070 0.020    1.070
HEC   C2B   C3B   aromatic    1.083 0.020    1.083
HEC   C2C   C3C   aromatic    1.384 0.020    1.384
HEC   C2D   C3D   aromatic    1.337 0.020    1.337
HEC   C3A   C4A   aromatic    1.217 0.020    1.217
HEC   C3B   C4B   aromatic    1.250 0.020    1.250
```

```
HEC  C3C      C4C  aromatic     1.423 0.020     1.423
HEC  C3D      C4D  aromatic     1.295 0.020     1.295
HEC  C4C      NC   aromatic     1.509 0.020     1.509
HEC  NC       FE   single       2.400 0.500     2.400
#
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_angle_esd
HEC  HAA1     CAA    HAA2        109.46 3.000
HEC  HAA1     CAA     C2A        109.48 3.000
HEC  HAA2     CAA     C2A        109.46 3.000
HEC  HAA1     CAA     CBA        109.47 3.000
HEC  HAA2     CAA     CBA        109.46 3.000
HEC   C2A     CAA     CBA        109.49 3.000
HEC   C3B     CAB     CBB        112.91 3.000
HEC   HAC     CAC     C3C        109.48 3.000
HEC   HAC     CAC     CBC        109.45 3.000
HEC   C3C     CAC     CBC        112.88 3.000
HEC  HAD1     CAD    HAD2        109.46 3.000
HEC  HAD1     CAD     C3D        109.47 3.000
HEC  HAD2     CAD     C3D        109.47 3.000
HEC  HAD1     CAD     CBD        109.47 3.000
HEC  HAD2     CAD     CBD        109.46 3.000
HEC   C3D     CAD     CBD        109.49 3.000
HEC  FE       NA      C4A        109.02 3.000
HEC  FE       NA      C1A        109.06 3.000
HEC   C4A     NA      C1A         97.44 3.000
HEC  HBA1     CBA    HBA2        109.47 3.000
HEC  HBA1     CBA     CGA        109.46 3.000
HEC  HBA2     CBA     CGA        109.47 3.000
HEC  HBA1     CBA     CAA        109.47 3.000
HEC  HBA2     CBA     CAA        109.48 3.000
HEC   CGA     CBA     CAA        109.48 3.000
HEC  HBB1     CBB    HBB2        109.47 3.000
HEC  HBB1     CBB    HBB3        109.47 3.000
HEC  HBB2     CBB    HBB3        109.47 3.000
HEC  HBB1     CBB     CAB        109.48 3.000
HEC  HBB2     CBB     CAB        109.47 3.000
HEC  HBB3     CBB     CAB        109.47 3.000
HEC  HBC1     CBC    HBC2        109.47 3.000
HEC  HBC1     CBC    HBC3        109.47 3.000
HEC  HBC2     CBC    HBC3        109.47 3.000
```

```
HEC  HBC1    CBC    CAC       109.48 3.000
HEC  HBC2    CBC    CAC       109.47 3.000
HEC  HBC3    CBC    CAC       109.47 3.000
HEC  HBD1    CBD   HBD2       109.47 3.000
HEC  HBD1    CBD    CGD       109.47 3.000
HEC  HBD2    CBD    CGD       109.47 3.000
HEC  HBD1    CBD    CAD       109.48 3.000
HEC  HBD2    CBD    CAD       109.47 3.000
HEC   CGD    CBD    CAD       109.48 3.000
HEC  FE      NB     C4B       108.87 3.000
HEC  FE      NB     C1B       109.38 3.000
HEC   C4B    NB     C1B        97.48 3.000
HEC   O2A    CGA    O1A       119.98 3.000
HEC   O2A    CGA    CBA       119.97 3.000
HEC   O1A    CGA    CBA       119.98 3.000
HEC   O2D    CGD    O1D       119.98 3.000
HEC   O2D    CGD    CBD       119.97 3.000
HEC   O1D    CGD    CBD       119.98 3.000
HEC  FE      ND     C4D       112.57 3.000
HEC  FE      ND     C1D       112.63 3.000
HEC   C4D    ND     C1D       119.50 3.000
HEC   C4D    CHA    C1A       133.79 8.000
HEC   HHB    CHB    C4A       109.46 3.000
HEC   HHB    CHB    C1B       109.46 3.000
HEC   C4A    CHB    C1B       133.44 8.000
HEC   HHC    CHC    C4B       109.47 3.000
HEC   HHC    CHC    C1C       109.47 3.000
HEC   C4B    CHC    C1C       134.47 8.000
HEC   HHD    CHD    C4C       109.44 3.000
HEC   HHD    CHD    C1D       109.44 3.000
HEC   C4C    CHD    C1D       134.74 8.000
HEC  HMA1    CMA   HMA2       109.46 3.000
HEC  HMA1    CMA   HMA3       109.47 3.000
HEC  HMA2    CMA   HMA3       109.46 3.000
HEC  HMA1    CMA    C3A       109.48 3.000
HEC  HMA2    CMA    C3A       109.47 3.000
HEC  HMA3    CMA    C3A       109.48 3.000
HEC  HMB1    CMB   HMB2       109.47 3.000
HEC  HMB1    CMB   HMB3       109.46 3.000
HEC  HMB2    CMB   HMB3       109.47 3.000
HEC  HMB1    CMB    C2B       109.48 3.000
HEC  HMB2    CMB    C2B       109.48 3.000
HEC  HMB3    CMB    C2B       109.47 3.000
HEC  HMC1    CMC   HMC2       109.47 3.000
HEC  HMC1    CMC   HMC3       109.47 3.000
HEC  HMC2    CMC   HMC3       109.47 3.000
```

HEC HMC1 CMC C2C 109.47 3.000
HEC HMC2 CMC C2C 109.48 3.000
HEC HMC3 CMC C2C 109.47 3.000
HEC HMD1 CMD HMD2 109.46 3.000
HEC HMD1 CMD HMD3 109.47 3.000
HEC HMD2 CMD HMD3 109.47 3.000
HEC HMD1 CMD C2D 109.47 3.000
HEC HMD2 CMD C2D 109.48 3.000
HEC HMD3 CMD C2D 109.48 3.000
HEC C2A C1A CHA 109.14 3.000
HEC C2A C1A NA 99.71 3.000
HEC CHA C1A NA 133.85 3.000
HEC C2B C1B CHB 115.84 3.000
HEC C2B C1B NB 99.60 3.000
HEC CHB C1B NB 133.80 3.000
HEC NC C1C C2C 94.43 3.000
HEC NC C1C CHC 112.41 3.000
HEC C2C C1C CHC 152.86 3.000
HEC C2D C1D CHD 139.36 3.000
HEC C2D C1D ND 93.13 3.000
HEC CHD C1D ND 126.58 3.000
HEC C3A C2A C1A 121.57 3.000
HEC C3A C2A CAA 112.54 3.000
HEC C1A C2A CAA 112.50 3.000
HEC C3B C2B C1B 121.56 3.000
HEC C3B C2B CMB 112.52 3.000
HEC C1B C2B CMB 112.51 3.000
HEC C3C C2C C1C 116.91 3.000
HEC C3C C2C CMC 112.51 3.000
HEC C1C C2C CMC 112.50 3.000
HEC C3D C2D C1D 111.84 3.000
HEC C3D C2D CMD 116.21 3.000
HEC C1D C2D CMD 116.19 3.000
HEC C4A C3A C2A 121.58 3.000
HEC C4A C3A CMA 112.51 3.000
HEC C2A C3A CMA 112.55 3.000
HEC C4B C3B C2B 121.62 3.000
HEC C4B C3B CAB 112.52 3.000
HEC C2B C3B CAB 112.52 3.000
HEC C4C C3C C2C 117.00 3.000
HEC C4C C3C CAC 112.58 3.000
HEC C2C C3C CAC 112.56 3.000
HEC C4D C3D C2D 118.99 3.000
HEC C4D C3D CAD 112.51 3.000
HEC C2D C3D CAD 112.52 3.000
HEC C3A C4A CHB 111.35 3.000

```
HEC   C3A     C4A     NA          99.69 3.000
HEC   CHB     C4A     NA         133.63 3.000
HEC   C3B     C4B     CHC        110.20 3.000
HEC   C3B     C4B     NB          99.73 3.000
HEC   CHC     C4B     NB         134.19 3.000
HEC   NC      C4C     C3C         94.50 3.000
HEC   NC      C4C     CHD        112.48 3.000
HEC   C3C     C4C     CHD         74.90 3.000
HEC   C3D     C4D     CHA        125.19 3.000
HEC   C3D     C4D     ND          96.54 3.000
HEC   CHA     C4D     ND         134.18 3.000
HEC  FE       NC      C4C        113.01 3.000
HEC  FE       NC      C1C        112.89 3.000
HEC   C4C     NC      C1C        117.16 3.000
HEC   NC      FE      ND          90.13 3.000
HEC   NC      FE      NB          89.04 3.000
HEC   ND      FE      NB         157.52 3.000
HEC   NC      FE      NA         172.83 3.000
HEC   ND      FE      NA          89.29 3.000
HEC   NB      FE      NA          88.77 3.000
#
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
HEC CONST_01      C4D     CHA     C1A     NA          -1.17 0.0 0
HEC CONST_02      C3A     C2A     C1A     NA          -0.19 0.0 0
HEC CONST_03      C1B     CHB     C4A     NA           1.50 0.0 0
HEC CONST_04      C2A     C3A     C4A     NA           0.09 0.0 0
HEC CONST_05      C1B     NB      FE      NA          31.73 0.0 0
HEC CONST_06      C4B     NB      FE      NA         137.17 0.0 0
HEC CONST_07      C1D     ND      FE      NA        -164.15 0.0 0
HEC CONST_08      C4D     ND      FE      NA         -25.47 0.0 0
HEC CONST_09      C1C     NC      FE      NA         -16.35 0.0 0
HEC CONST_10      C4C     NC      FE      NA         119.51 0.0 0
HEC CONST_11      C4A     CHB     C1B     NB           3.27 0.0 0
HEC CONST_12      C3B     C2B     C1B     NB          -0.07 0.0 0
HEC CONST_13      C1C     CHC     C4B     NB           1.28 0.0 0
HEC CONST_14      C2B     C3B     C4B     NB           0.09 0.0 0
HEC CONST_15      C1A     NA      FE      NB        -134.91 0.0 0
```

```
HEC CONST_16      C4A    NA    FE      NB        -29.60 0.0 0
HEC CONST_17      C1D    ND    FE      NB        -79.10 0.0 0
HEC CONST_18      C4D    ND    FE      NB         59.58 0.0 0
HEC CONST_19      C1C    NC    FE      NB         55.92 0.0 0
HEC CONST_20      C4C    NC    FE      NB       -168.22 0.0 0
HEC CONST_21      C4C    CHD    C1D    ND         16.81 0.0 0
HEC CONST_22      C3D    C2D    C1D    ND         -0.07 0.0 0
HEC CONST_23      C1A    CHA    C4D    ND         -2.12 0.0 0
HEC CONST_24      C2D    C3D    C4D    ND          0.08 0.0 0
HEC CONST_25      C1A    NA    FE      ND         22.70 0.0 0
HEC CONST_26      C4A    NA    FE      ND        128.01 0.0 0
HEC CONST_27      C1B    NB    FE      ND        -53.43 0.0 0
HEC CONST_28      C4B    NB    FE      ND         52.01 0.0 0
HEC CONST_29      C1C    NC    FE      ND       -101.62 0.0 0
HEC CONST_30      C4C    NC    FE      ND         34.24 0.0 0
HEC CONST_31      C3A    C4A    NA     C1A        -0.16 0.0 0
HEC CONST_32      NC    FE      NA     C1A        -62.63 0.0 0
HEC CONST_33      C3D    C4D    CHA    C1A         149.49 0.0 0
HEC CONST_34      C4A    C3A    C2A    C1A          0.07 0.0 0
HEC CONST_35      C3B    C4B    NB     C1B         -0.11 0.0 0
HEC CONST_36      NC    FE      NB     C1B        -141.44 0.0 0
HEC CONST_37      C3A    C4A    CHB    C1B         129.99 0.0 0
HEC CONST_38      C4B    C3B    C2B    C1B         -0.01 0.0 0
HEC CONST_39      C3B    C4B    CHC    C1C         128.80 0.0 0
HEC CONST_40      C4C    C3C    C2C    C1C         -0.14 0.0 0
HEC CONST_41      C3C    C4C    NC     C1C         -0.34 0.0 0
HEC CONST_42      C3D    C4D    ND     C1D         -0.14 0.0 0
HEC CONST_43      NC    FE      ND     C1D          8.71 0.0 0
HEC CONST_44      C3C    C4C    CHD    C1D         129.78 0.0 0
HEC CONST_45      NC    C4C    CHD    C1D          41.14 0.0 0
HEC CONST_46      C4D    C3D    C2D    C1D         -0.00 0.0 0
HEC CONST_47      C4A    NA    C1A    C2A          0.19 0.0 0
HEC CONST_48      C4D    CHA    C1A    C2A        -126.69 0.0 0
HEC CONST_49      C4B    NB    C1B    C2B          0.10 0.0 0
HEC CONST_50      C4A    CHB    C1B    C2B        -132.89 0.0 0
HEC CONST_51      C4B    CHC    C1C    C2C        -171.38 0.0 0
HEC CONST_52      C4C    NC    C1C    C2C          0.28 0.0 0
HEC CONST_53      NC    C4C    C3C    C2C          0.28 0.0 0
HEC CONST_54      C4D    ND    C1D    C2D          0.13 0.0 0
HEC CONST_55      C4C    CHD    C1D    C2D        -148.85 0.0 0
HEC CONST_56      NC    C1C    C2C    C3C         -0.08 0.0 0
HEC CONST_57      NC    FE      NA     C4A         42.69 0.0 0
HEC CONST_58      NC    FE      NB     C4B        -36.01 0.0 0
HEC CONST_59      NC    C1C    CHC    C4B         17.72 0.0 0
HEC CONST_60      NC    FE      ND     C4D        147.39 0.0 0
HEC CONST_61      NA    C1A    C2A    CAA         137.68 0.0 0
```

```
HEC CONST_62     CHA    C1A    C2A    CAA      -78.86   0.0 0
HEC CONST_63     C4A    C3A    C2A    CAA      -137.78   0.0 0
HEC CONST_64     C1B    C2B    C3B    CAB      137.89   0.0 0
HEC CONST_65     NB     C4B    C3B    CAB      -137.81   0.0 0
HEC CONST_66     CHC    C4B    C3B    CAB      77.43   0.0 0
HEC CONST_67     C1C    C2C    C3C    CAC      -132.86   0.0 0
HEC CONST_68     CHD    C4C    C3C    CAC      20.90   0.0 0
HEC CONST_69     NC     C4C    C3C    CAC      132.99 0.0 0
HEC CONST_70     C1D    C2D    C3D    CAD      134.71   0.0 0
HEC CONST_71     ND     C4D    C3D    CAD      -134.64   0.0 0
HEC CONST_72     CHA    C4D    C3D    CAD      65.43   0.0 0
HEC CONST_73     CMA    C3A    C4A    NA       -137.83   0.0 0
HEC CONST_74     CMB    C2B    C1B    NB       137.76 0.0 0
HEC CONST_75     CMD    C2D    C1D    ND       136.59   0.0 0
HEC CONST_76     CMB    C2B    C1B    CHB      -72.71   0.0 0
HEC CONST_77     CMA    C3A    C4A    CHB      77.25   0.0 0
HEC CONST_78     CMC    C2C    C1C    CHC      -39.20   0.0 0
HEC CONST_79     CMD    C2D    C1D    CHD      -54.90   0.0 0
HEC CONST_80     C1A    C2A    C3A    CMA      137.97   0.0 0
HEC CONST_81     C4B    C3B    C2B    CMB      -137.84   0.0 0
HEC CONST_82     NC     C1C    C2C    CMC      132.37 0.0 0
HEC CONST_83     C4C    C3C    C2C    CMC      -132.59 0.0 0
HEC CONST_84     C4D    C3D    C2D    CMD      -136.65 0.0 0
HEC CONST_85     HHB    CHB    C4A    NA       146.71   0.0 0
HEC CONST_86     HHC    CHC    C4B    NB       148.69   0.0 0
HEC CONST_87     HHD    CHD    C1D    ND       -131.01 0.0 0
HEC Var_01       C1B    CHB    C4A    NA         1.50 30.0 1
HEC Var_02       C1B    NB     FE     NA        31.73 30.0 1
HEC Var_03       C1D    ND     FE     NA      -164.15 30.0 1
HEC Var_04       C1C    NC     FE     NA       -16.35 30.0 1
HEC Var_05       C4A    CHB    C1B    NB         3.27  30.0 1
HEC Var_06       C1C    CHC    C4B    NB         1.28  30.0 1
HEC Var_07       C1A    NA     FE     NB      -134.91  30.0 1
HEC Var_08       C4C    CHD    C1D    ND        16.81 30.0 1
HEC Var_09       C3D    C4D    CHA    C1A      149.49  30.0 1
HEC Var_10       C3C    C4C    CHD    C1D      129.78  30.0 1
HEC Var_11       C4B    CHC    C1C    C2C     -171.38  30.0 1
HEC Var_12       C1A    C2A    CAA    CBA      -41.99  30.0 2
HEC Var_13       C2B    C3B    CAB    CBB     -105.38  30.0 2
HEC Var_14       C2C    C3C    CAC    CBC     -118.61  30.0 2
HEC Var_15       C2D    C3D    CAD    CBD       97.56  30.0 2
HEC Var_16       HMB3   CMB    C2B    C1B       58.64  30.0 2
HEC Var_17       HMC3   CMC    C2C    C1C       61.58  30.0 2
HEC Var_18       HMD3   CMD    C2D    C1D      -65.32  30.0 2
HEC Var_19       HMA3   CMA    C3A    C2A       67.14  30.0 2
HEC Var_20       C2A    CAA    CBA    CGA      177.17 30.0 2
```

```
HEC Var_21       C3D    CAD    CBD    CGD      -169.89  30.0 2
HEC Var_22      HBB3    CBB    CAB    C3B       -98.56  30.0 3
HEC Var_23      HBC3    CBC    CAC    C3C      -140.77  30.0 3
HEC Var_24       O1A    CGA    CBA    CAA        65.72  30.0 3
HEC Var_25       O1D    CGD    CBD    CAD       -58.61  30.0 3
#
loop_
_chem_comp_plane_atom.comp_id
_chem_comp_plane_atom.plane_id
_chem_comp_plane_atom.atom_id
_chem_comp_plane_atom.dist_esd
HEC plan-1   CBA 0.020
HEC plan-1   CGA 0.020
HEC plan-1   O1A 0.020
HEC plan-1   O2A 0.020
HEC plan-2   CAA 0.020
HEC plan-2    NA 0.020
HEC plan-2   CHA 0.020
HEC plan-2   CHB 0.020
HEC plan-2   CMA 0.020
HEC plan-2   C1A 0.020
HEC plan-2   C2A 0.020
HEC plan-2   C3A 0.020
HEC plan-2   C4A 0.020
HEC plan-2    FE 0.020
HEC plan-3   CAD 0.020
HEC plan-3    ND 0.020
HEC plan-3   CHA 0.020
HEC plan-3   CHD 0.020
HEC plan-3   CMD 0.020
HEC plan-3   C1D 0.020
HEC plan-3   C2D 0.020
HEC plan-3   C3D 0.020
HEC plan-3   C4D 0.020
HEC plan-3    FE 0.020
HEC plan-4   CAC 0.020
HEC plan-4   CHC 0.020
HEC plan-4   CHD 0.020
HEC plan-4   CMC 0.020
HEC plan-4   C1C 0.020
HEC plan-4   C2C 0.020
HEC plan-4   C3C 0.020
HEC plan-4   C4C 0.020
HEC plan-4    NC 0.020
HEC plan-4 FE 0.020
HEC plan-5   CAB 0.020
```

HEC plan-5 NB 0.020
HEC plan-5 CHB 0.020
HEC plan-5 CHC 0.020
HEC plan-5 CMB 0.020
HEC plan-5 C1B 0.020
HEC plan-5 C2B 0.020
HEC plan-5 C3B 0.020
HEC plan-5 C4B 0.020
HEC plan-5 FE 0.020
HEC plan-6 CBD 0.020
HEC plan-6 CGD 0.020
HEC plan-6 O1D 0.020
HEC plan-6 O2D 0.020

Forced Single Bonds

```
data_comp_list
loop_
_chem_comp.id
_chem_comp.three_letter_code
_chem_comp.name
_chem_comp.group
_chem_comp.number_atoms_all
_chem_comp.number_atoms_nh
_chem_comp.desc_level
HEC        HEC 'Unknown                ' ligand 73 43 .
#
data_comp_HEC
#
loop_
_chem_comp_atom.comp_id
_chem_comp_atom.atom_id
_chem_comp_atom.type_symbol
_chem_comp_atom.type_energy
_chem_comp_atom.charge
_chem_comp_atom.partial_charge
_chem_comp_atom.x
_chem_comp_atom.y
_chem_comp_atom.z
HEC        CAA    C   CH2    0    .      26.7957 21.7332 23.4902
HEC        CAB    C   CSP    0    .      34.5133 21.1849 21.5785
HEC        CAC    C   C1     0    .      31.0303 14.1833 21.3988
HEC        CAD    C   CH2    0    .      26.4967 16.6802 23.7745
HEC        NA     N   N      0    .      28.3657 20.7245 20.5594
HEC        CBA    C   CH2    0     .      25.5733 22.0660   22.5895
HEC        CBB    C   CH3    0     .      35.2964 21.3388   20.2555
HEC        CBC    C   CH3    0     .      31.2287 12.6802   21.0785
```

```
HEC      CBD   C   CH2   0   .     26.6598 16.8673 25.3055
HEC      NB    N   N     0   .     31.6513 20.4597 19.7524
HEC      CGA   C   C     0   .     24.4243 22.6094 23.4624
HEC      CGD   C   C     0   .     25.4431 16.2587 26.0301
HEC      ND    N   N     0   .     28.4622 17.3212 21.1319
HEC      CHA   C   C     0   .     27.2842 19.2110 22.3734
HEC      CHB   C   C1    0   .     30.3615 22.1772 21.2828
HEC      CHC   C   CR16  0   .     33.3692 18.6270 20.8308
HEC      CHD   C   C1    0   .     30.0717 15.3693 20.6185
HEC      CMA   C   CH3   0   .     28.4030 23.2362     22.9518
HEC      CMB   C   CH3   0   .     33.0227 22.8689     21.8822
HEC      CMC   C   CH3   0   .     34.1513 15.2537     21.0630
HEC      CMD   C   CH3   0   .     28.0279 14.6214     23.2514
HEC      C1A   C   CR56  0   .     27.5128 20.3115     21.7956
HEC      C1B   C   CR5   0   .     31.3939 21.6793     20.7639
HEC      C1C   C   CR56  0   .     32.7615 17.2261     20.7714
HEC      C1D   C   CR5   0   .     29.1111 16.0373     21.4257
HEC      O1A   O   O     0   .     24.5548 23.7131 24.0617
HEC      O1D   O   O     0   .     24.2803 16.6777 25.7719
HEC      C2A   C   CR5   0   .     27.9355 21.1148     22.6157
HEC      C2B   C   CR5   0   .     32.3395 21.5275     21.5088
HEC      C2C   C   CR5   0   .     32.7406 15.8885     21.0832
HEC      C2D   C   CR5   0   .     28.5818 15.9894     22.7819
HEC      O2A   O   OC    0   .     23.3313 21.9805     23.5294
HEC      O2D   O   OC    0   .     25.6142 15.3824     26.9234
HEC      C3A   C   CR5   0   .     28.6916 21.8286     22.3627
HEC      C3B   C   CR5   0   .     33.0526 20.7252     21.3635
HEC      C3C   C   CR5   0   .     31.7991 15.1083     20.4339
HEC      C3D   C   CR5   0   .     27.8284 17.0583     23.0599
HEC      C4A   C   CR5   0   .     29.1707 21.8804     21.2453
HEC      C4B   C   CR56  0   .     32.9407 19.9113     20.4209
HEC      C4C   C   CR5   0   .     30.9414 15.7670     19.5084
HEC      C4D   C   CR56  0   .     27.6624 17.9630     22.1489
HEC      NC    N   N     0   .     31.5611   17.1264   19.7194
HEC      FE    FE  FE    0   .     29.9321   18.8680   19.9935
HEC      HMD3  H   HCH3  0   .     27.1509 14.3616     22.6587
HEC      HMD2  H   HCH3  0   .     27.7505 14.6835     24.3015
HEC      HMD1  H   HCH3  0   .     28.7918 13.8573     23.1228
HEC      HMC3  H   HCH3  0   .     34.7785 15.7389     21.8097
HEC      HMC2  H   HCH3  0   .     34.0745 14.1915     21.2885
HEC      HMC1  H   HCH3  0   .     34.5942 15.3843     20.0765
HEC      HMB3  H   HCH3  0   .     32.2972 23.5250     22.3598
HEC      HMB2  H   HCH3  0   .     33.4047 23.3457     20.9798
HEC      HMB1  H   HCH3  0   .     33.8469 22.6780     22.5689
HEC      HMA3  H   HCH3  0   .     28.4785 23.1977     24.0392
HEC      HMA2  H   HCH3  0   .     29.1293 23.9487     22.5625
```

```
HEC       HMA1    H   HCH3   0    .     27.4003 23.5493     22.6687
HEC       HBD2    H   HCH2   0    .     26.7257 17.9288     25.5363
HEC       HBD1    H   HCH2   0    .     27.5676 16.3668     25.6376
HEC       HBC3    H   HCH3   0    .     30.2867 12.1519     21.2180
HEC       HBC2    H   HCH3   0    .     31.9815 12.2628     21.7464
HEC       HBC1    H   HCH3   0    .     31.5578 12.5692     20.0459
HEC       HBB3    H   HCH3   0    .     35.9001 20.4488     20.0854
HEC       HBB2    H   HCH3   0    .     35.9459 22.2119     20.3172
HEC       HBB1    H   HCH3   0    .     34.5969 21.4649     19.4328
HEC       HBA2    H   HCH2   0    .     25.8576 22.8176     21.8548
HEC       HBA1    H   HCH2   0    .     25.2440 21.1641     22.0769
HEC       HAD2    H   HCH2   0    .     25.6954 17.3239     23.4135
HEC       HAD1    H   HCH2   0    .     26.2521 15.6421     23.5602
HEC       HAA2    H   HCH2   0    .     26.4979 21.0183     24.2570
HEC       HAA1    H   HCH2   0    .     27.1552 22.6429     23.9634
HEC       HHD     H   H      0    .     29.7796 14.3277 20.4948
HEC       HHC     H   HCR6   0    .     34.4510 18.5575  20.7343
HEC       HHB     H   H      0    .     30.4643 23.2614 21.3276
HEC       HAC     H   H      0    .     31.3476 14.3812 22.4216
#
loop_
_chem_comp_bond.comp_id
_chem_comp_bond.atom_id_1
_chem_comp_bond.atom_id_2
_chem_comp_bond.type
_chem_comp_bond.value_dist
_chem_comp_bond.value_dist_esd
_chem_comp_bond.value_dist_neutron
HEC   CAA    CBA   single       1.554 0.020     1.554
HEC   CAA    C2A   single       1.564 0.020     1.564
HEC   CAA    HAA2  single       0.939 0.020     1.096
HEC   CAA    HAA1  single       0.939 0.020     1.096
HEC   CAB    CBB   single       1.545 0.020     1.545
HEC   CAB    C3B   single       1.546 0.020     1.546
HEC   CAC    CBC   single       1.550 0.020     1.550
HEC   CAC    C3C   single       1.542 0.020     1.542
HEC   CAC    HAC   single       0.951 0.020     1.111
HEC   CAD    CBD   single       1.551 0.020     1.551
HEC   CAD    C3D   single       1.558 0.020     1.558
HEC   CAD    HAD2  single       0.939 0.020     1.096
HEC   CAD    HAD1  single       0.939 0.020     1.096
HEC   NA     C1A  aromatic      1.558 0.020     1.558
HEC   NA     C4A  aromatic      1.567 0.020     1.567
HEC   NA     FE    single       2.494 0.500     2.494
HEC   CBA    CGA  single        1.542 0.020     1.542
HEC   CBA    HBA2  single       0.939 0.020     1.096
```

```
HEC  CBA   HBA1  single      0.939 0.020    1.096
HEC  CBB   HBB3  single      0.939 0.020    1.096
HEC  CBB   HBB2  single      0.939 0.020    1.096
HEC  CBB   HBB1  single      0.939 0.020    1.096
HEC  CBC   HBC3  single      0.939 0.020    1.096
HEC  CBC   HBC2  single      0.939 0.020    1.096
HEC  CBC   HBC1  single      0.939 0.020    1.096
HEC  CBD   CGD   single      1.541 0.020    1.541
HEC  CBD   HBD2  single      0.939 0.020    1.096
HEC  CBD   HBD1  single      0.939 0.020    1.096
HEC  NB    C1B   aromatic    1.605 0.020    1.605
HEC  NB    C4B   aromatic    1.552 0.020    1.552
HEC  NB    FE    single      2.355 0.500    2.355
HEC  CGA   O1A   deloc       1.263 0.020    1.263
HEC  CGA   O2A   deloc       1.263 0.020    1.263
HEC  CGD   O1D   deloc       1.263 0.020    1.263
HEC  CGD   O2D   deloc       1.263 0.020    1.263
HEC  ND    C1D   aromatic    1.468 0.020    1.468
HEC  ND    C4D   aromatic    1.444 0.020    1.444
HEC  ND    FE    single      2.418 0.500    2.418
HEC  CHA   C1A   single      1.550 0.300    1.264
HEC  CHA   C4D   single      1.550 0.300    1.323
HEC  CHB   C1B   aromatic    1.258 0.300    1.258
HEC  CHB   C4A   aromatic    1.228 0.300    1.228
HEC  CHB   HHB   single      0.951 0.200    1.111
HEC  CHC   C1C   aromatic    1.528 0.300    1.528
HEC  CHC   C4B   aromatic    1.415 0.300    1.415
HEC  CHC   HHC   single      0.951 0.020    1.111
HEC  CHD   C1D   aromatic    1.422 0.300    1.422
HEC  CHD   C4C   aromatic    1.465 0.300    1.465
HEC  CHD   HHD   single      0.951 0.020    1.111
HEC  CMA   C3A   single      1.553 0.020    1.553
HEC  CMA   HMA3  single      0.939 0.020    1.096
HEC  CMA   HMA2  single      0.939 0.020    1.096
HEC  CMA   HMA1  single      0.939 0.020    1.096
HEC  CMB   C2B   single      1.551 0.020    1.551
HEC  CMB   HMB3  single      0.939 0.020    1.096
HEC  CMB   HMB2  single      0.939 0.020    1.096
HEC  CMB   HMB1  single      0.939 0.020    1.096
HEC  CMC   C2C   single      1.547 0.020    1.547
HEC  CMC   HMC3  single      0.939 0.020    1.096
HEC  CMC   HMC2  single      0.939 0.020    1.096
HEC  CMC   HMC1  single      0.939 0.020    1.096
HEC  CMD   C2D   single      1.549 0.020    1.549
HEC  CMD   HMD3  single      0.939 0.020    1.096
HEC  CMD   HMD2  single      0.939 0.020    1.096
```

```
HEC   CMD    HMD1 single       0.939 0.020    1.096
HEC   C1A     C2A  aromatic     1.223 0.020    1.223
HEC   C1B     C2B aromatic     1.213 0.020    1.213
HEC   C1C     C2C  aromatic     1.374 0.020    1.374
HEC   C1C     NC   aromatic     1.599 0.020    1.599
HEC   C1D     C2D  aromatic     1.457 0.020    1.457
HEC   C2A     C3A  aromatic     1.070 0.020    1.070
HEC   C2B     C3B aromatic     1.083 0.020    1.083
HEC   C2C     C3C  aromatic     1.384 0.020    1.384
HEC   C2D     C3D  aromatic     1.337 0.020    1.337
HEC   C3A     C4A  aromatic     1.217 0.020    1.217
HEC   C3B     C4B aromatic     1.250 0.020    1.250
HEC   C3C     C4C  aromatic     1.423 0.020    1.423
HEC   C3D     C4D  aromatic     1.295 0.020    1.295
HEC   C4C     NC   aromatic     1.509 0.020    1.509
HEC   NC      FE    single      2.400 0.500    2.400
#
loop_
_chem_comp_angle.comp_id
_chem_comp_angle.atom_id_1
_chem_comp_angle.atom_id_2
_chem_comp_angle.atom_id_3
_chem_comp_angle.value_angle
_chem_comp_angle.value_angle_esd
HEC  HAA1    CAA   HAA2       109.46 3.000
HEC  HAA1    CAA    C2A       109.48 3.000
HEC  HAA2    CAA    C2A       109.46 3.000
HEC  HAA1    CAA    CBA       109.47 3.000
HEC  HAA2    CAA    CBA       109.46 3.000
HEC   C2A    CAA    CBA       109.49 3.000
HEC   C3B    CAB    CBB       112.91 3.000
HEC   HAC    CAC    C3C       109.48 3.000
HEC   HAC    CAC    CBC       109.45 3.000
HEC   C3C    CAC    CBC       112.88 3.000
HEC  HAD1    CAD   HAD2       109.46 3.000
HEC  HAD1    CAD    C3D       109.47 3.000
HEC  HAD2    CAD    C3D       109.47 3.000
HEC  HAD1    CAD    CBD       109.47 3.000
HEC  HAD2    CAD    CBD       109.46 3.000
HEC   C3D    CAD    CBD       109.49 3.000
HEC  FE      NA     C4A       109.02 3.000
HEC  FE      NA     C1A       109.06 3.000
HEC   C4A    NA     C1A        97.44 3.000
HEC  HBA1    CBA   HBA2       109.47 3.000
HEC  HBA1    CBA    CGA       109.46 3.000
HEC  HBA2    CBA    CGA       109.47 3.000
```

```
HEC  HBA1   CBA   CAA      109.47 3.000
HEC  HBA2   CBA   CAA      109.48 3.000
HEC   CGA   CBA   CAA      109.48 3.000
HEC  HBB1   CBB  HBB2      109.47 3.000
HEC  HBB1   CBB  HBB3      109.47 3.000
HEC  HBB2   CBB  HBB3      109.47 3.000
HEC  HBB1   CBB   CAB      109.48 3.000
HEC  HBB2   CBB   CAB      109.47 3.000
HEC  HBB3   CBB   CAB      109.47 3.000
HEC  HBC1   CBC  HBC2      109.47 3.000
HEC  HBC1   CBC  HBC3      109.47 3.000
HEC  HBC2   CBC  HBC3      109.47 3.000
HEC  HBC1   CBC   CAC      109.48 3.000
HEC  HBC2   CBC   CAC      109.47 3.000
HEC  HBC3   CBC   CAC      109.47 3.000
HEC  HBD1   CBD  HBD2      109.47 3.000
HEC  HBD1   CBD   CGD      109.47 3.000
HEC  HBD2   CBD   CGD      109.47 3.000
HEC  HBD1   CBD   CAD      109.48 3.000
HEC  HBD2   CBD   CAD      109.47 3.000
HEC   CGD   CBD   CAD      109.48 3.000
HEC  FE     NB    C4B      108.87 3.000
HEC  FE     NB    C1B      109.38 3.000
HEC   C4B   NB    C1B       97.48 3.000
HEC   O2A   CGA   O1A      119.98 3.000
HEC   O2A   CGA   CBA      119.97 3.000
HEC   O1A   CGA   CBA      119.98 3.000
HEC   O2D   CGD   O1D      119.98 3.000
HEC   O2D   CGD   CBD      119.97 3.000
HEC   O1D   CGD   CBD      119.98 3.000
HEC  FE     ND    C4D      112.57 3.000
HEC  FE     ND    C1D      112.63 3.000
HEC   C4D   ND    C1D      119.50 3.000
HEC   C4D   CHA   C1A      133.79 8.000
HEC   HHB   CHB   C4A      109.46 3.000
HEC   HHB   CHB   C1B      109.46 3.000
HEC   C4A   CHB   C1B      133.44 8.000
HEC   HHC   CHC   C4B      109.47 3.000
HEC   HHC   CHC   C1C      109.47 3.000
HEC   C4B   CHC   C1C      134.47 8.000
HEC   HHD   CHD   C4C      109.44 3.000
HEC   HHD   CHD   C1D      109.44 3.000
HEC   C4C   CHD   C1D      134.74 8.000
HEC  HMA1   CMA  HMA2      109.46 3.000
HEC  HMA1   CMA  HMA3      109.47 3.000
HEC  HMA2   CMA  HMA3      109.46 3.000
```

HEC HMA1 CMA C3A 109.48 3.000
HEC HMA2 CMA C3A 109.47 3.000
HEC HMA3 CMA C3A 109.48 3.000
HEC HMB1 CMB HMB2 109.47 3.000
HEC HMB1 CMB HMB3 109.46 3.000
HEC HMB2 CMB HMB3 109.47 3.000
HEC HMB1 CMB C2B 109.48 3.000
HEC HMB2 CMB C2B 109.48 3.000
HEC HMB3 CMB C2B 109.47 3.000
HEC HMC1 CMC HMC2 109.47 3.000
HEC HMC1 CMC HMC3 109.47 3.000
HEC HMC2 CMC HMC3 109.47 3.000
HEC HMC1 CMC C2C 109.47 3.000
HEC HMC2 CMC C2C 109.48 3.000
HEC HMC3 CMC C2C 109.47 3.000
HEC HMD1 CMD HMD2 109.46 3.000
HEC HMD1 CMD HMD3 109.47 3.000
HEC HMD2 CMD HMD3 109.47 3.000
HEC HMD1 CMD C2D 109.47 3.000
HEC HMD2 CMD C2D 109.48 3.000
HEC HMD3 CMD C2D 109.48 3.000
HEC C2A C1A CHA 109.14 3.000
HEC C2A C1A NA 99.71 3.000
HEC CHA C1A NA 133.85 3.000
HEC C2B C1B CHB 115.84 3.000
HEC C2B C1B NB 99.60 3.000
HEC CHB C1B NB 133.80 3.000
HEC NC C1C C2C 94.43 3.000
HEC NC C1C CHC 112.41 3.000
HEC C2C C1C CHC 152.86 3.000
HEC C2D C1D CHD 139.36 3.000
HEC C2D C1D ND 93.13 3.000
HEC CHD C1D ND 126.58 3.000
HEC C3A C2A C1A 121.57 3.000
HEC C3A C2A CAA 112.54 3.000
HEC C1A C2A CAA 112.50 3.000
HEC C3B C2B C1B 121.56 3.000
HEC C3B C2B CMB 112.52 3.000
HEC C1B C2B CMB 112.51 3.000
HEC C3C C2C C1C 116.91 3.000
HEC C3C C2C CMC 112.51 3.000
HEC C1C C2C CMC 112.50 3.000
HEC C3D C2D C1D 111.84 3.000
HEC C3D C2D CMD 116.21 3.000
HEC C1D C2D CMD 116.19 3.000
HEC C4A C3A C2A 121.58 3.000

```
HEC   C4A     C3A     CMA        112.51 3.000
HEC   C2A     C3A     CMA        112.55 3.000
HEC   C4B     C3B     C2B        121.62 3.000
HEC   C4B     C3B     CAB        112.52 3.000
HEC   C2B     C3B     CAB        112.52 3.000
HEC   C4C     C3C     C2C        117.00 3.000
HEC   C4C     C3C     CAC        112.58 3.000
HEC   C2C     C3C     CAC        112.56 3.000
HEC   C4D     C3D     C2D        118.99 3.000
HEC   C4D     C3D     CAD        112.51 3.000
HEC   C2D     C3D     CAD        112.52 3.000
HEC   C3A     C4A     CHB        111.35 3.000
HEC   C3A     C4A     NA         99.69 3.000
HEC   CHB     C4A     NA         133.63 3.000
HEC   C3B     C4B     CHC        110.20 3.000
HEC   C3B     C4B     NB         99.73 3.000
HEC   CHC     C4B     NB         134.19 3.000
HEC   NC      C4C     C3C        94.50 3.000
HEC   NC      C4C     CHD        112.48 3.000
HEC   C3C     C4C     CHD        74.90 3.000
HEC   C3D     C4D     CHA        125.19 3.000
HEC   C3D     C4D     ND         96.54 3.000
HEC   CHA     C4D     ND         134.18 3.000
HEC  FE       NC      C4C        113.01 3.000
HEC  FE       NC      C1C        112.89 3.000
HEC   C4C     NC      C1C        117.16 3.000
HEC   NC      FE      ND         90.13 3.000
HEC   NC      FE      NB         89.04 3.000
HEC   ND      FE      NB         157.52 3.000
HEC   NC      FE      NA         172.83 3.000
HEC   ND      FE      NA         89.29 3.000
HEC   NB      FE      NA         88.77 3.000
#
loop_
_chem_comp_tor.comp_id
_chem_comp_tor.id
_chem_comp_tor.atom_id_1
_chem_comp_tor.atom_id_2
_chem_comp_tor.atom_id_3
_chem_comp_tor.atom_id_4
_chem_comp_tor.value_angle
_chem_comp_tor.value_angle_esd
_chem_comp_tor.period
HEC CONST_01      C4D     CHA     C1A     NA          -1.17   0.0 0
HEC CONST_02      C3A     C2A     C1A     NA          -0.19   0.0 0
HEC CONST_03      C1B     CHB     C4A     NA           1.50   0.0 0
```

```
HEC CONST_04      C2A    C3A    C4A    NA          0.09 0.0 0
HEC CONST_05      C1B    NB     FE     NA         31.73 0.0 0
HEC CONST_06      C4B    NB     FE     NA        137.17 0.0 0
HEC CONST_07      C1D    ND     FE     NA       -164.15 0.0 0
HEC CONST_08      C4D    ND     FE     NA        -25.47 0.0 0
HEC CONST_09      C1C    NC     FE     NA        -16.35 0.0 0
HEC CONST_10      C4C    NC     FE     NA        119.51 0.0 0
HEC CONST_11      C4A    CHB    C1B    NB          3.27 0.0 0
HEC CONST_12      C3B    C2B    C1B    NB         -0.07 0.0 0
HEC CONST_13      C1C    CHC    C4B    NB          1.28 0.0 0
HEC CONST_14      C2B    C3B    C4B    NB          0.09 0.0 0
HEC CONST_15      C1A    NA     FE     NB       -134.91 0.0 0
HEC CONST_16      C4A    NA     FE     NB        -29.60 0.0 0
HEC CONST_17      C1D    ND     FE     NB        -79.10 0.0 0
HEC CONST_18      C4D    ND     FE     NB         59.58 0.0 0
HEC CONST_19      C1C    NC     FE     NB         55.92 0.0 0
HEC CONST_20      C4C    NC     FE     NB       -168.22 0.0 0
HEC CONST_21      C4C    CHD    C1D    ND         16.81 0.0 0
HEC CONST_22      C3D    C2D    C1D    ND         -0.07 0.0 0
HEC CONST_23      C1A    CHA    C4D    ND         -2.12 0.0 0
HEC CONST_24      C2D    C3D    C4D    ND          0.08 0.0 0
HEC CONST_25      C1A    NA     FE     ND         22.70 0.0 0
HEC CONST_26      C4A    NA     FE     ND        128.01 0.0 0
HEC CONST_27      C1B    NB     FE     ND        -53.43 0.0 0
HEC CONST_28      C4B    NB     FE     ND         52.01 0.0 0
HEC CONST_29      C1C    NC     FE     ND       -101.62 0.0 0
HEC CONST_30      C4C    NC     FE     ND         34.24 0.0 0
HEC CONST_31      C3A    C4A    NA     C1A        -0.16 0.0 0
HEC CONST_32      NC     FE     NA     C1A       -62.63 0.0 0
HEC CONST_33      C3D    C4D    CHA    C1A        149.49 0.0 0
HEC CONST_34      C4A    C3A    C2A    C1A         0.07 0.0 0
HEC CONST_35      C3B    C4B    NB     C1B        -0.11 0.0 0
HEC CONST_36      NC     FE     NB     C1B      -141.44 0.0 0
HEC CONST_37      C3A    C4A    CHB    C1B        129.99 0.0 0
HEC CONST_38      C4B    C3B    C2B    C1B        -0.01 0.0 0
HEC CONST_39      C3B    C4B    CHC    C1C        128.80 0.0 0
HEC CONST_40      C4C    C3C    C2C    C1C        -0.14 0.0 0
HEC CONST_41      C3C    C4C    NC     C1C        -0.34 0.0 0
HEC CONST_42      C3D    C4D    ND     C1D        -0.14 0.0 0
HEC CONST_43      NC     FE     ND     C1D         8.71 0.0 0
HEC CONST_44      C3C    C4C    CHD    C1D        129.78 0.0 0
HEC CONST_45      NC     C4C    CHD    C1D        41.14 0.0 0
HEC CONST_46      C4D    C3D    C2D    C1D        -0.00 0.0 0
HEC CONST_47      C4A    NA     C1A    C2A         0.19 0.0 0
HEC CONST_48      C4D    CHA    C1A    C2A        -126.69 0.0 0
HEC CONST_49      C4B    NB     C1B    C2B         0.10 0.0 0
```

```
HEC CONST_50      C4A    CHB    C1B    C2B        -132.89 0.0 0
HEC CONST_51      C4B    CHC    C1C    C2C        -171.38 0.0 0
HEC CONST_52      C4C    NC     C1C    C2C         0.28 0.0 0
HEC CONST_53      NC     C4C    C3C    C2C         0.28 0.0 0
HEC CONST_54      C4D    ND     C1D    C2D         0.13 0.0 0
HEC CONST_55      C4C    CHD    C1D    C2D        -148.85 0.0 0
HEC CONST_56      NC     C1C    C2C    C3C        -0.08 0.0 0
HEC CONST_57      NC     FE     NA     C4A         42.69 0.0 0
HEC CONST_58      NC     FE     NB     C4B        -36.01 0.0 0
HEC CONST_59      NC     C1C    CHC    C4B         17.72 0.0 0
HEC CONST_60      NC     FE     ND     C4D        147.39 0.0 0
HEC CONST_61      NA     C1A    C2A    CAA         137.68 0.0 0
HEC CONST_62      CHA    C1A    C2A    CAA        -78.86 0.0 0
HEC CONST_63      C4A    C3A    C2A    CAA        -137.78 0.0 0
HEC CONST_64      C1B    C2B    C3B    CAB         137.89 0.0 0
HEC CONST_65      NB     C4B    C3B    CAB        -137.81 0.0 0
HEC CONST_66      CHC    C4B    C3B    CAB         77.43 0.0 0
HEC CONST_67      C1C    C2C    C3C    CAC        -132.86 0.0 0
HEC CONST_68      CHD    C4C    C3C    CAC         20.90 0.0 0
HEC CONST_69      NC     C4C    C3C    CAC        132.99 0.0 0
HEC CONST_70      C1D    C2D    C3D    CAD         134.71 0.0 0
HEC CONST_71      ND     C4D    C3D    CAD        -134.64 0.0 0
HEC CONST_72      CHA    C4D    C3D    CAD         65.43 0.0 0
HEC CONST_73      CMA    C3A    C4A    NA         -137.83 0.0 0
HEC CONST_74      CMB    C2B    C1B    NB         137.76 0.0 0
HEC CONST_75      CMD    C2D    C1D    ND          136.59 0.0 0
HEC CONST_76      CMB    C2B    C1B    CHB        -72.71 0.0 0
HEC CONST_77      CMA    C3A    C4A    CHB         77.25 0.0 0
HEC CONST_78      CMC    C2C    C1C    CHC        -39.20 0.0 0
HEC CONST_79      CMD    C2D    C1D    CHD        -54.90 0.0 0
HEC CONST_80      C1A    C2A    C3A    CMA         137.97 0.0 0
HEC CONST_81      C4B    C3B    C2B    CMB        -137.84 0.0 0
HEC CONST_82      NC     C1C    C2C    CMC        132.37 0.0 0
HEC CONST_83      C4C    C3C    C2C    CMC        -132.59 0.0 0
HEC CONST_84      C4D    C3D    C2D    CMD        -136.65 0.0 0
HEC CONST_85      HHB    CHB    C4A    NA          146.71 0.0 0
HEC CONST_86      HHC    CHC    C4B    NB          148.69 0.0 0
HEC CONST_87      HHD    CHD    C1D    ND         -131.01 0.0 0
HEC Var_01        C1B    CHB    C4A    NA          1.50 30.0 1
HEC Var_02        C1B    NB     FE     NA          31.73 30.0 1
HEC Var_03        C1D    ND     FE     NA         -164.15 30.0 1
HEC Var_04        C1C    NC     FE     NA         -16.35 30.0 1
HEC Var_05        C4A    CHB    C1B    NB          3.27 30.0 1
HEC Var_06        C1C    CHC    C4B    NB          1.28 30.0 1
HEC Var_07        C1A    NA     FE     NB         -134.91 30.0 1
HEC Var_08        C4C    CHD    C1D    ND          16.81 30.0 1
```

```
HEC Var_09       C3D    C4D    CHA    C1A       149.49 30.0 1
HEC Var_10       C3C    C4C    CHD    C1D       129.78 30.0 1
HEC Var_11       C4B    CHC    C1C    C2C      -171.38 30.0 1
HEC Var_12       C1A    C2A    CAA    CBA       -41.99 30.0 2
HEC Var_13       C2B    C3B    CAB    CBB      -105.38 30.0 2
HEC Var_14       C2C    C3C    CAC    CBC      -118.61 30.0 2
HEC Var_15       C2D    C3D    CAD    CBD        97.56 30.0 2
HEC Var_16      HMB3    CMB    C2B    C1B        58.64 30.0 2
HEC Var_17      HMC3    CMC    C2C    C1C        61.58 30.0 2
HEC Var_18      HMD3    CMD    C2D    C1D       -65.32 30.0 2
HEC Var_19      HMA3    CMA    C3A    C2A        67.14 30.0 2
HEC Var_20       C2A    CAA    CBA    CGA       177.17 30.0 2
HEC Var_21       C3D    CAD    CBD    CGD      -169.89 30.0 2
HEC Var_22      HBB3    CBB    CAB    C3B       -98.56 30.0 3
HEC Var_23      HBC3    CBC    CAC    C3C      -140.77 30.0 3
HEC Var_24       O1A    CGA    CBA    CAA        65.72 30.0 3
HEC Var_25       O1D    CGD    CBD    CAD       -58.61 30.0 3
#
loop_
_chem_comp_plane_atom.comp_id
_chem_comp_plane_atom.plane_id
_chem_comp_plane_atom.atom_id
_chem_comp_plane_atom.dist_esd
HEC plan-1    CBA 0.020
HEC plan-1    CGA 0.020
HEC plan-1    O1A 0.020
HEC plan-1    O2A 0.020
HEC plan-2    CAA 0.020
HEC plan-2     NA 0.020
HEC plan-2    CHA 0.020
HEC plan-2    CHB 0.020
HEC plan-2    CMA 0.020
HEC plan-2    C1A 0.020
HEC plan-2    C2A 0.020
HEC plan-2    C3A 0.020
HEC plan-2    C4A 0.020
HEC plan-2     FE 0.020
HEC plan-3    CAD 0.020
HEC plan-3     ND 0.020
HEC plan-3    CHA 0.020
HEC plan-3    CHD 0.020
HEC plan-3    CMD 0.020
HEC plan-3    C1D 0.020
HEC plan-3    C2D 0.020
HEC plan-3    C3D 0.020
HEC plan-3    C4D 0.020
```

HEC plan-3 FE 0.020
HEC plan-4 CAC 0.020
HEC plan-4 CHC 0.020
HEC plan-4 CHD 0.020
HEC plan-4 CMC 0.020
HEC plan-4 C1C 0.020
HEC plan-4 C2C 0.020
HEC plan-4 C3C 0.020
HEC plan-4 C4C 0.020
HEC plan-4 NC 0.020
HEC plan-4 FE 0.020
HEC plan-5 CAB 0.020
HEC plan-5 NB 0.020
HEC plan-5 CHB 0.020
HEC plan-5 CHC 0.020
HEC plan-5 CMB 0.020
HEC plan-5 C1B 0.020
HEC plan-5 C2B 0.020
HEC plan-5 C3B 0.020
HEC plan-5 C4B 0.020
HEC plan-5 FE 0.020
HEC plan-6 CBD 0.020
HEC plan-6 CGD 0.020
HEC plan-6 O1D 0.020
HEC plan-6 O2D 0.020