De Novo Designed Metallopeptides to Investigate Metal Ion Homeostasis, Electron Transfer, and Redox Catalysis

by

Alison G. Tebo

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemical Biology)
in the University of Michigan
2015

Doctoral Committee:

Professor Vincent L. Pecoraro, Co-Chair
Professor Ally Aukauloo, Co-Chair, Université Paris-Sud XI
Professor David P. Ballou
Professor Carol A. Fierke
Professor Nicolai Lehnert

To my parents

Acknowledgements

First and foremost, I would like to acknowledge my advisor, Vince. I have learned a lot from working with you over the past four and a half years. I've gotten to learn so many techniques and have benefitted a lot from the breadth of opportunity I've had in this lab. I have always really appreciated your insight and patience (especially when I fall off mountains). It's been a pleasure "rowing the boat in the same direction" with you. Secondly, I would like to thank my co-advisor, Ally. Your support and advocacy has allowed me to have many opportunities would I not have had otherwise. I would also like to thank my committee, Professor Carol Fierke, Professor Dave Ballou, and Professor Nicolai Lehnert. Discussions with you have always been extrememly helpful and have given me many ideas and strategies to address challenging topics.

I would like to thank all of my collaborators who have helped me gather, interpret and understand this data as well as train me. They are: Dr. Jean-Marc Latour, Dr. Ricardo García, Dr. Olivier Sénéque, Dr. Cédric Tard, Dr. Matteo Duca, Prof. James Penner-Hahn, Dr. Aniruddha Deb, Prof. Nicolai Lehnert, Amy Speelman, Ashley McQuarters, Dr. Lars Hemmingsen, and Prof. José J. G. Moura. Extra special thanks needs to be given to Dr. Annamaria Quaranta for help collecting and analyzing the photophysical data—Chapter 4 would not have happened without her help. I would also like to thank Dr. Winfreid Leibl, Dr. Hervé Bottin, and Dr. Christian Herrero for all of the support at CEA Saclay.

Finally, I would like to thank all of my labmates, former and current, in all of the labs that I have worked in over the past years. Thanks for the camaraderie and helpful discussions. You've been instrumental in my success. In particular the Pecoraro lab: Cathy, Fangting, Leela, Jeff, Ginny, David, Jake, Casey, Ted, Joe, Tu and Evan. To my cohort, PCB 2010: we are the best cohort. Period.

To my friends in Ann Arbor and around the world—you guys are awesome. The people who have really kept me grounded and sane over the years share in this success with me. In no particular order or level of completeness: Brooke, Cathy, JP, Jon, Nerissa, Gaëlle, Maddy,

Fangting, ChrisTom, Rachel, Alina, Erica, JennChen, Bernie, Becca, Mauri, Pallavi. Thank you to my partner, JB. You're awesome. I would not have made it through this without your support. I can't say enough good things about you.

And finally, thank you to my family who has been supportive and rigorous in their expectations. Thanks Martin, for taking the pressure off and distracting everyone with your theater stuff. To my parents, you're the best parents that anyone, especially a Ph.D. student could ask for. Thank you for setting the best of examples of what excellent people, parents, and scientists look like. Thank your for being an awesome example of scientists who do what they love and what they believe in, do it well, and with integrity. I could not have better role models.

Table of Contents

List of Tables

List of Figures

Abstract

Protein design is a powerful way to interrogate the basic requirements for function of metal sites by systematically incorporating elements important for function. Single-stranded three-helix bundles with either thiolate-rich sites for spectroscopic characterization and electron transfer, or histidine-rich sites for redox catalysis are described.

Using a previous design, two constructs were designed to incorporate a fourth cysteine residue to investigate thiolate-rich sites involved in metal ion homeostasis and electron transfer. Rational re-design replaced a putative coordinating histidine with a cysteine. A second construct embedded a CXXC binding motif into the helical scaffold. These two constructs show different UV-visisble, ^{113}Cd NMR, and ^{111m}Cd PAC, which indicate that they form different proportions of CdS_3O and CdS_4. The spectroscopy of these sites sheds light on how Cd(II) bindis to CadC and suggests a dynamic site in fast exchange with the solvent.

Previous attempts at the design of a rubredoxin site have focused on reproducing the peptide fold around or using flexible loop regions to define the site in addition to canonical CXXC motifs. However, the use of CXXC motifs embedded in an α-helical scaffold produces a rubredoxin site that reproduces the Mössbauer, MCD, and EPR of rubredoxin without the use of loop regions. This successful design is the largest deviation from consensus rubredoxin and zinc finger folds reported.

Electron transfer rates through a *de novo* designed scaffold were studied by the design and synthesis of a ruthenium trisbipyridine derivative appended to an exterior cysteine residues. A redox-active tyrosine in the 70th position is implicated as a relay amino acid from the iron center and absence of the tyrosine decreases the rate of electron transfer from the metal site. This is the first photo-generated tyrosine radical in a designed protein.

A construct, which was previously reported for CO_2 hydration, is substituted with copper and its spectroscopic and nitrite reductase activity are studied. This is the first demonstration of nitrite reductase activity in a single-stranded designed peptide.

This thesis provides insight into designed proteins and their applications and lays the groundwork for further studies to progress towards a unified multifunctional redox protein.

Chapter 1. Introduction

Overview

Evolutionary selective forces have succeeded, over an enormous timescale, in generating proteins with highly specialized structures and folds that feature highly efficient catalysis, specific interactions, and fine-tuned regulation. Protein design seeks to separate evolutionary artifacts from function to distill out larger questions of how the biochemical world around us functions. Rather than understand the specific interactions, regulation, and catalysis of a single system, protein design strives to understand the greater principles by which proteins function. This understanding can allow us to arrive at useful and robust systems on the human timescale, which can be applied towards effective modeling of native systems or the development of novel functions.

By isolating certain specific features in a suitably complex system, we can hope to understand the fundamentals of these (isolated) processes better and then gradually expand the complexity of our system to generate compact multifunctional systems with the desired functions. For example, our laboratory has spent many years studying the finer details of heavy metal complexation by thiolate-rich sites in proteins and has contributed to the development and interpretation of spectroscopic techniques to interrogate these processes. Thiol-coordinated metals play important roles in regulation and toxicity, as well as acting as structurally stabilizing sites in a protein fold. More recently, our laboratory's efforts have expanded to developing catalytic activity in our designed peptides, resulting in some of the most successful mimics for both hydrolytic and redox catalysis achieved in the protein design field. However, less effort has been devoted to developing and understanding electron transfer capability, which is important not only for multisite systems, but also to optimize redox catalysis. Thus, developing a series of tractable electron transfer centers that are compatible with other, successful catalytic systems is an attractive goal for protein designers.

The main goal of my thesis has been to explore the reactivity and coordination of thiolate-rich and histidine rich sites within a single chain three-helix bundle scaffold, α_3**D**. Initially, my

efforts were a natural expansion of our previous work in thiolate-coordination using the construct α_3**DIV**,[1] and allowed for an interpretation of how Cd(II) binds to metal sensing proteins, like CadC. Next, I incorporated iron into tetrathiolate sites in my design, characterized its spectroscopy, and its ability to cycle between two redox states. These spurred the expansion of these studies to understand the limits and features of electron transfer in our system more fully. Finally, I present preliminary work with a different peptide, α_3D**H$_3$**,[2] which when coordinated to copper produces a Type 2 copper center, a rather enigmatic site that can function as either an electron transfer site or as a redox catalytic site. In this first chapter, I will detail the relevant work in our laboratory and others' that has both inspired and informed our research. First, I will introduce protein design and some of its principles. Then, I will discuss metal homeostasis, native biological electron transfer systems, and electron transfer theory before moving on to previous attempts at protein design of electron transfer sites. Finally, I will place my work in the larger context of this laboratory's work and introduce the development of modified α_3**D** peptides for the purpose of metal coordination.

Protein Design

The interplay between primary, secondary, and tertiary structure has been studied by the likes of chemists such as Linus Pauling, who famously predicted with great accuracy the α-helix and β-sheet based only on idealized hydrogen bonding interactions. Understanding the relationship between the structure and function of proteins is not straightforward; what may be important in one instance, may not constitute a general rule that we can apply broadly. While the relationship between sequence and fold has been interrogated by numerous groups and with substantial progress, the more subtle questions of how a protein functions are difficult to target broadly. Features such as substrate specificity, protein-protein interactions, and allosteric regulation are certainly modulated by structures, but understanding how we might distill general understanding from studying what Nature came up with through trial and error is a formidable task. Essentially, there are two ways to approach the fundamental question of, "how do proteins work?" One approach is the top-down method of classic protein biochemistry—using mutations, inhibitors, and other tools to perturb or abrogate function to understand key structural features that are required for folding, a specific interaction, or catalysis. The second, bottom-up approach is protein design, where features are incorporated into a design with the goal of understanding to

what extent a certain structural element may contribute to a feature found in native proteins. Protein design also allows for the exploration of features beyond those found in nature i.e., features that are possible, yet unsampled.

Protein structure is defined and stabilized by several types of interactions, namely hydrophobic interactions, hydrogen bonding interactions, van der Waals interactions, disulfide bonds, covalent peptide bonds, and salt bridges. Protein folding is a balance between the favorable enthalpy of the formation of specific interactions as well as shielding hydrophobic residues from solvent and the unfavorable entropy of the folding process.[3] Setting aside those proteins that require assistance by chaperones, polypeptide chains can sample a variety of conformations, some of which may represent local minima on the protein folding landscape and act as intermediates in the folding process. In general, overall energy minimization favors the adoption of a unique, well-defined conformation, which has the correct oligomeric state, overall topology, correct spatial organization of secondary structure elements, and finally, a well-defined, hydrophobic interior in which single resides populate a single rotomeric state.[4]

An estimated one-half to one-third of native proteins require a metal ion for proper folding or catalytic function.[5] The wide variety of activity produced by a limited number of metals has been a focus of scientific research as we seek to understand how these proteins function and apply this to pressing questions in both the biomedical and the environmental realms. Consequently, the field of metalloprotein design has evolved alongside the greater protein design field. Protein design can be broadly classified into protein redesign and *de novo* protein design. Protein redesign uses native proteins as scaffolds and then builds new functions or binding sites into the pre-existing scaffold. The conversion of one (metal) site into another in the same scaffold is a prime way to examine different functionalities within the same framework and context, thus eliminating the variables created by comparing different sites within different proteins. Furthermore, by using a minimal number of changes to convert a known protein to a metalloprotein with a specific cofactor and activity, the basic requirements for activity can be examined in the absence of complications from the limits of stability and symmetry of the scaffold. Indeed, many native proteins are robust and modifiable, as well as easy to crystallize, and thus are excellent candidates for use as a scaffold for redesign. One can also seek to improve upon the native function of a metalloprotein by making systematic changes, favoring greater stability, altering inhibition profiles, or affecting reactivity as needed for a particular purpose.

Thus, protein redesign has seen great popularity in basic science research, as it has been an effective tool for understanding the underlying principles of biological activity.

De novo protein design seeks to generate a protein "from scratch" and features scaffolds whose primary sequence bears no relation to native proteins, thus identifying the basic, minimal features for function. It is essentially starting from a blank canvas and adding only the features necessary to achieve a particular function. *De novo* design has the advantage of being a truly robust test of the understanding of first principles and is not constrained to sequences related to known proteins. By using basic constructs that contain only the desired attributes, one is able to systematically study isolated metal-protein binding processes and spectroscopy. Similar to protein redesign, previously characterized *de novo* designed proteins can serve as scaffolds to build upon earlier designs or extended the complexity of a construct, which is especially useful for metalloprotein design. To date, most established *de novo* designed protein scaffolds consist of α-helical secondary structures, which either self-assemble to form coiled coils or fold as helix-loop-helix motifs into a helical bundle[6], although catalytic metalloenzymes have recently been reported using β-sheet constructs[7]. By isolating different types of sites and functions, we can gain a greater level of understanding of the fundamental processes associated with metal binding, catalysis, and electron transfer; as well as learn how to control these properties, in order to tune a construct's attributes for a given application. Knowledge gained through the design and characterization of these metal sites would then comprise a "tool box" of functions we can apply to the preparation of designer enzymes for a wide variety of applications. Our laboratory has long studied heavy metal complexation by thiolate-rich sites in proteins using de novo designed peptides and more recently has expanded this work to create a stabilizing structural site in a designed hydrolytic zinc enzyme, which is capable of carbonic anhydrase activity.[8] Recently, our laboratory has also reported redox catalytic activity in the form of a designed nitrite reductase.[9,10] Essential elements of this toolbox are electron transfer sites, which act to not only transfer electrons as part of a larger process such as respiration, but also to and from catalytic sites and substrates, in the context of a redox reaction.

Many of the functional de novo designed metalloproteins, including those form our lab, utilize an α-helical scaffold, so more attention will be given to these constructs except where examples of particular interest are worth noting.

Metal Homeostasis

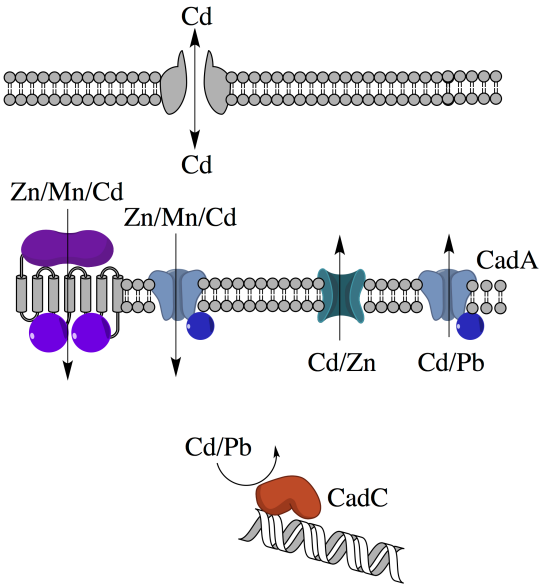


Figure 1.1 General schematic for some aspects of Cd(II) homeostasis in Gram-negative bacteria. It is thought that Cd(II) enters the cell by transport along with other divalent cations, where it can bind adventitiously to thiol-rich sites in essential proteins. Metalloregulatory proteins bind Cd(II) and regulate the expression of detoxification systems. Methods of efflux include but are not limited to cation diffusion facilitators (turquoise) and P-type ATPases (blue). The best characterized detoxification system for Cd(II) is the CadC metalloregulatory protein and the CadA ATPase (labeled). Adapted from ref. 2.

Metal homeostasis ensures that the right metal gets to the right protein, but also senses whether the levels of any metal are too high, as even the "right" metal can become toxic at high levels. Furthermore, some metals are toxic at nearly all concentrations, so cells and organisms need to be able to sense and remove these metals to prevent fatal toxicity. The formation of specific protein-metal interactions are what govern metal homeostasis by the specific uptake, efflux, storage, and trafficking of the metal of interest (Figure 1.1). The proteins involved play several different roles: metal transporters, which effect the directional transfer of a metal or metal-ligand complex across a barrier; metallochaperones, which traffic the metal intracompartmentally and effect the transfer of the metal to a specific protein; and metalloregulatory proteins, which sense the presence of the metal and interact with transcriptional elements on the DNA to either up- or down-regulate metal uptake and efflux. Naturally, for these systems to be effective, they must be selective for the metal ion of interest and have affinities that are appropriate to ensure the directional transfer of the metal ion towards its target.[11]

Metalloregulatory proteins can be separated into two broad classes based on function.[12] The first class comprises members that regulate genes required for metal efflux or sequestration to effectively lower the concentration of the metal ion of interest in the cytosol. The second class

does the opposite; they control gene expression for proteins involved in the uptake of metals. Proteins that control the gene expression for metal uptake bind that metal as a co-repressor such that high intracellular levels of the metal as sensed by the metal sensor indicate that uptake is no longer required and the system is down-regulated.[13] For most proteins that control metal sequestration or efflux to lower intracellular metal concentrations, the model is very similar to the lac operon—transcription is repressed until binding of the ligand (in this case the metal), causes dissociation of the repressor allowing for transcription of the operon. What is striking about metal-responsive transcription factors is their lack of distinctness—they are orthologs of other ligand-binding transcription factors, but they have evolved to have very specific interactions with certain metals.[12] The selectivity is unrelated to their global fold or mechanism, but rather selectivity for a particular metal must lie in the immediate proximity of the metal itself, a fact that is of particular interest to bioinorganic chemists.

This discussion on metal homeostasis will primarily focus on the sensing and regulation of heavy metals that have no known biological role and are toxic to cells and microorganisms. Detoxification of these metals can vary in the specifics, but in the end the toxic metal is exported out of the cell rather than being sequestered (as in the case of high iron levels). Efflux of toxic heavy metals is mostly carried out by cation diffusion facilitators (CDFs), P-type ATPases, and RNDs (resistance-nodulation-cell division).[12] These efflux proteins must also bind these metals and often exhibit some selectivity over which metals are exported, thus meriting a brief overview of their metal-binding characteristics. The best characterized CDF is a H+ antiporter termed YiiP that is mostly known for its role in Zn(II) export; however, it can also transport Cd(II) as the two metals are isovalent.[14] YiiP is a transmembrane homodimer with four metal binding sites per monomer, one of which is located in the lipid bilayer-spanning portion of the protein and is required for transport. Mutational analysis indicates that YiiP discriminates between Zn or Cd and other metals based on the tetrahedral geometry of the binding site and the enforcement of n=4 coordination number.

P-type ATPases involved in heavy metal transport share a mechanism with the other members of the family in which phosphorylation of an aspartate residue in the P domain, ATP binding in the N domain, and substrate binding at the transmembrane helices prime the protein for a conformational change that exposes the substrate (metal, in this case) to the extracytosolic space.[15] They are distinguished from other P-type counterparts by the metal-binding motifs in the

transmembrane helices, as well as by cytosolic metal-binding domains, which are unique to this class and appear to have a regulatory role.[12,15] Metal binding motifs at the transmembrane region are responsible for metal recognition; however, generally less is known about the coordination sphere of metals in proteins where binding and vectorial release at physiologically relevant rates are part of the protein's function. Subtypes tend to transport only a small subset of structurally related metals, e.g. Zn(II) is transported along with Cd(II) and Pb(II) and Cu(I) is transported along with Ag(I). The transmembrane metal binding site in Zn(II)/Cd(II)/Pb(II) transporters is characterized by Lys, Asp, and Gly residues as well as a conserved Cys-Pro-Cys sequence.[16] Overall, very little work has been done to understand this binding site. In Cu(I)/Ag(I) transporters the metal binding site is consists of Asn, Tyr, Met, Ser and a Cys-Pro-Xxx motif; the metal is coordinated in a trigonal planar arrangement with oxygen or nitrogen ligands in the coordination sphere as opposed to the typical thiol-saturated, digonal coordination in copper chaperones.[17] It is hypothesized that some transporters function by direct transfer from the relevant cytosolic chelator (such as glutathione) to the transmembrane metal binding site. In other systems, particularly Cu(I), where free Cu(I) levels are low and mostly bound to metallochaperones,[18] it appears that transfer is mediated by interaction between the chaperone and the transmembrane metal binding site.[17]

There are cytosolic metal binding domains that are located at either the N or C terminus of the P-type ATPase. The N-terminal domains are typically 60-70 residues and contain a highly conserved Cys-X-X-Cys metal binding sequence.[19] The cytosolic metal binding domains themselves are homologous to Cu(I) metallochaperones. These domains use the CXXC motif to present a digonal, trigonal planar, or equilibrium structure for Cu(I) binding, similar to Cu(I) coordination in metallochaperones.[20,21] Additionally, there is a lot of evidence that cytosolic metal binding domains can accept Cu(I) directly from chaperones;[22] however, subsequent transfer to the transmembrane metal binding site for transport has only been observed in specific cases,[23] and thus their role is hypothesized to be mostly regulatory. Interestingly enough, Zn(II) ATPases also have these cytosolic metal binding domains, despite the fact that there are no known metallochaperones for Zn(II). They share the properties of having cysteine-rich sequences with a Cys-X-X-Cys motif. The solution structure of the cytosolic metal binding domain of a particular member of this class that confers resistance to Zn(II), Cd(II), and Pb(II) was solved

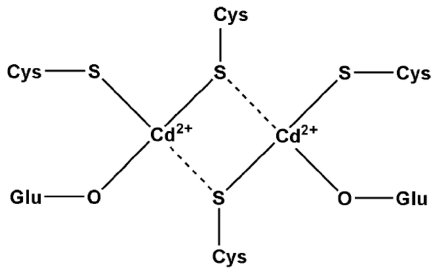


Figure 1.2 Coordination sphere of bimetallic Cd(II) center from CadA.

and showed that the metal binding motif, DCXXC was responsible for coordinating Zn(II) with two cysteines and an aspartate.[24]

CadA, another P-type ATPase responsible for Cd(II) and Zn(II) resistance also has this CXXC binding motif, however the protein forms a homodimer with a bimetallic center in which the coordination sphere of Cd(II) is completed by a Glu residue on an adjacent helix (Figure 1.2).[25] This complex was characterized by UV-visible spectroscopy, ^{113}Cd NMR, and Cd(II) XAS so the details of Cd(II) coordination are well-understood. While the presence of anionic residues here and in the transmembrane binding site may help with selectivity for divalent cations over monovalent cations, binding geometry seems to be important for functional selectivity, as Co(II), Cu(II), and Ni(II) all bind with similarly high affinity but show no functional activity.[26] Much still remains to be understood about how these exporters function and how selectivity is obtained, although it is clear that export of heavy metals is dominated by proteins rich in cysteine residues but with important distortions and asymmetry, perhaps to modulate affinities and kinetics of transport.

Heavy metals are typically sensed by members of one of two highly studied families: the ArsR/SmtB family and the MerR family, which both regulate the expression of genes involved in detoxification or storage. These regulators are sensitive to a small subset of metals, and while the coordination sphere of the bound metals is fairly well-defined, the details of proper function remain elusive due to uncharacterized allosteric states. Essentially, how the information of coordination of the "right" metal is transmitted to the overall quaternary conformation of the regulator is not well-understood.[12] The ArsR/SmtB family is the most extensively studied family.[13] It is named for the founding members As(III)- and Sb(III)-sensing ArsR and Zn(II)-sensing SmtB, and contains CadC, which will be discussed in detail in this section.

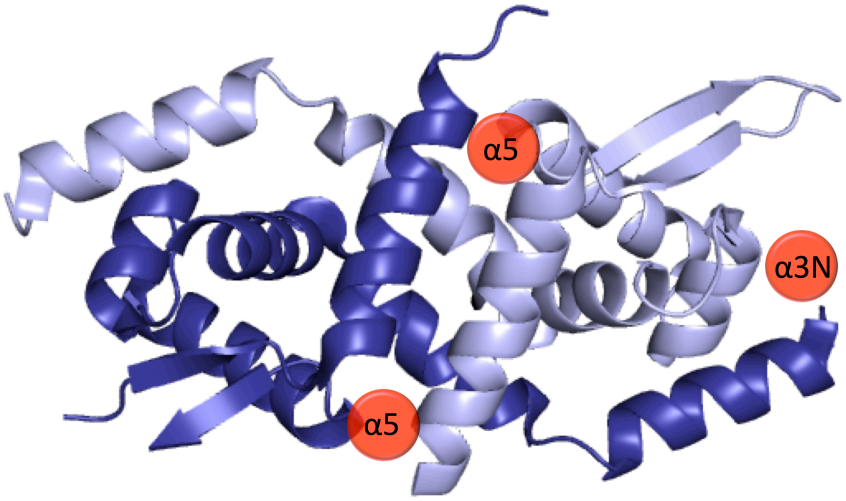


Figure 1.3 Ribbon representation of the two best-characterized binding sites (red) on ArsR/SmtB family metal sensors. The binding site (α4C) for Cd(II) on CmtR is not pictured here. This figure is based on the crystal structure of *S. aureus* pI258 CadC, which was not able to resolve the flexible N and C termini. Each monomer is colored differently. (PDB: 1U2W).

The ArsR/SmtB family of transcriptional repressors functions by de-repression of genes for detoxification upon metal binding. In low metal concentrations, the regulator sits on the promoter region of DNA preventing transcription of downstream genes. When metal concentrations increase, the repressor binds the metal, introducing a conformational change to a low DNA affinity conformation that dissociates from DNA, thus allowing transcription of the downstream detoxification genes. The ArsR/SmtB family of proteins has nine structurally distinct metal binding sites; certain sites tend to bind particular classes of metals, thus the relevant site for a particular repressor varies based on its target metal(s) (Figure 1.3).[12,27] Several crystal and NMR structures of apo- and metal-bound exemplars of this family have been solved and all exhibit a dimeric, winged helix-turn-helix topology used for DNA binding.[12,28,29] The metal binding sites are formed by ligands from both monomers to form symmetry related pairs of metal binding sites. These proteins use cysteine-rich binding sites to bind Cd(II), Pb(II), and As(III). Cd(II), Pb(II), and sometimes Zn(II) bind at the sites termed α3N (or type 1) that consist of cysteines donated from α-helices in the core of the protein and from the N terminal loop (Figure 1.3).[12,28] A second common binding site is α5, which is found in the center of the protein and uses N/O ligands to bind borderline metals such as Zn(II) and Co(II).[12,30] CadC from *S. aureus* pI258 confers resistance to Cd(II), Pb(II), Zn(II), and Bi(III) and regulates the expression of CadA. CadC binds Cd(II) tetrahedrally and Pb(II) trigonally at the α3N site.[28,30,31] Mutational analysis studies have revealed that the cysteines that coordinate Cd(II) and Pb(II) in CadC are nonequivalent—while removal of a single cysteine can still result in metal binding, the presence

of the bound metal is not "sensed" and the repressor does not dissociate from DNA.[30] While there is variation in coordination and metal specificity in proteins from different organisms, there are features common to specific homologous sites that inform which ligands will complex which metals (Table 1.1).

Table 1.1 Summary of ArsR/SmtB sensors and their selectivity and coordination[27]

Protein	Metal	Sensing Sites	Ligand Set	Coordination Environment
SmtB	Zn	α5	His_2, Asp, Glu	Tetrahedral
CzrA	Zn, Co	α5	His_3, Asp	Tetrahedral
ZiaR	Zn	α5 and α3N	α3N: Cys_3, His α5: His_2, Asp, Glu	Tetra-coordinate
ArsR	As, Sb	α3/α3N-2/α5-4[a]	Cys_3	Trigonal
CadC	Cd, Pb, Zn	α3N	Cys_4	Tetrahedral (Cd, Zn) or trigonal (Pb)
AztR	Cd, Pb, Zn	α3N	Cys_3, His	Tetrahedral (Cd, Zn) or trigonal (Pb)
BxmR	Cu, Ag, Cd, Zn	α3N (α5 for Zn)	Cys_4 (α5: His_2, Asp, Glu)	Tetrahedral (Cd, Zn) or trigonal (Cu)
CmtR	Cd, Pb	α4C	Cys_3	Trigonal
NmtR	Ni, Co	α5C	His_5, Asp	Octahedral
KmtR	Ni, Co	α5-3	His_4, Glu, Asp?	Penta- or octa-coordinate

[a]based on the organism: *E. coli, Corynebacterium glutamicum,* and *Acidithiobacillus ferrooxidans* respectively

Furthermore, while a protein will have multiple binding sites and even bind metals in multiple places, not all sites are metalloregulatory. For example, CadC from *S. aureus* can bind metals at both the α3N and α5 sites, but only α3N induces dissociation from the DNA.[30] Furthermore, in CadC the α3N is selective for Cd(II) over Zn(II), while the converse is true of the α5 site.[30,32] Crystallographic studies of CadC revealed that a single residue around the α5 binding site is likely responsible for whether metal binding can regulate dimerization or DNA binding.[32] This is the opposite case to other homologous proteins, such as SmtB, which bind Co(II) and Zn(II) at α5, and in this case the site α3N is not metalloregulatory. Metal selectivity between these two sites seems to be largely understood on the basis of hard/soft acid/base chemistry, although the coordination geometry can vary quite a bit between individual proteins and metals. This suggests that metal selectivity has co-evolved with alternate allosteric pathways to couple metal binding and DNA binding.[28,30] In this way, we may see different intermediates in the evolution of ArsR/SmtB metal sensors with some members only possessing a type 1 site (i.e. α3N) (such as ArsR), some that possess both, although only one is metalloregulatory (CadC), and some that possess only type 2 sites (i.e. α5) (SmtB).

It has also been proposed that kinetics may define the biological response to toxic metals, rather than affinity. Most metalloregulatory proteins bind metals at affinities in excess of 10^{12}

M^{-1}, which would imply that metal detoxification response would be triggered at less than one free metal ion per cell.[18] Thus, biologically relevant behavior of metalloregulation likely relies on the kinetics of metal binding and conformational change, as well as metal-ligand transient interactions rather than the absolute value of metal affinity.[18,31] Metalloregulatory proteins then would participate in detoxification through both de-repression of their operon and through sequestration in the cytosol. Pulsed alkylation mass spectrometry was used to probe the reactivity of the cysteine ligands in the presence of Pb(II), Cd(II), and Bi(III) and the results concurred with previous studies that suggested that the cysteine ligands are non-equivalent.[30,33] Specifically, Cys11, which does not coordinate Pb(II), was identified as a "weak point" in the chelation of all the metals tested and was suggested to be an entry site for metal-ligand complex exchange with CadC.[33] On the other hand, Cys7 was shown to be particularly non-reactive to this technique, forming strong coordination bonds with all tested metals, which is consistent with its previously suggested role as a critical ligand involved in regulating the allosteric response of the protein. These observations correspond well with the ligand environment of Cd(II) and Pb(II). Cd(II) has been shown to coordinate tetrahedrally with a CdS_4 primary coordination sphere based on optical spectroscopy and XAS, while Pb(II) is thought to coordinate to three cysteines trigonally. However, ^{113}Cd NMR revealed a chemical shift of 622 ppm, which is significantly lower than that of bonafide Cd-substituted S_4 proteins (zinc fingers and rubredoxins). Cd-substituted S_4 proteins typically have chemical shifts of 750-700 ppm (vs. 0.1 M $CdClO_4$).[30] In light of the functional asymmetry of the site, the current hypothesis is that Cd(II) assumes a distorted tetrahedral polyhedron based on available spectroscopy and pulsed alkylation mass spectrometry, although no high resolution structures of Cd(II)-bound CadC yet exist.

Another member of the ArsR/SmtB family, CmtR, binds Cd(II) at a different site that does not correspond either to the α3N site or the α5 site.[34] It also does not rely on four cysteines to coordinate Cd(II), but rather will bind Cd(II) and Pb(II) at the same site with only three cysteines.[34,35] The protein has a total of six cysteine residues; however, it was found through mutational analysis that only three (57, 61, and 102) were required for *in vivo* Cd(II)-based response forming a site termed α4C.[34] Also unusual for ArsR/SmtB metal sensors that often sense Cd(II) and Pb(II), CmtR has no biologically relevant response to Zn(II), although it is capable of tight in vitro binding.[34,35] The low molar extinction coefficient (14000-16000 $M^{-1}cm^{-1}$), relative to higher thiolate coordination, led to questions as to whether all three cysteine are

bound to Cd(II). Mutation of Cys102 led to a decrease in the molar extinction coefficient, implying that it is coordinated, although not at the magnitude that would be expected of a full-strength Cd(II)-thiolate bond. ^{113}Cd NMR showed a chemical shift of 480 ppm, which is upfield of proteins with trigonal planar or trigonal pyramidal coordination of Cd(II) by cysteine residues, suggesting that perhaps two cysteines coordinate more strongly than the third.[35] ^{111m}Cd PAC spectroscopy of the wildtype protein revealed a $CdS_3(O)$ coordination in a trigonal pyramidal or distorted tetrahedral environment with a higher frequency (and, therefore, greater distortion) than that found in CadC.[35] When C102S was measured, two signals were detected, both of which were similar to the frequency and coordination of wild type CmtR and it was suggested that the second signal corresponded to CdS_2O_2. Based on UV-vis spectroscopy complemented with ^{113}Cd NMR and ^{111m}Cd PAC, it is hypothesized that CmtR binds Cd(II) with two strong thiolate bonds, with a third coordinating more weakly. A more recent NMR structure of the metallated and apo-forms of CmtR confirmed that the chemical shifts of all three cysteines are perturbed upon metal binding, suggesting that Cys102 does indeed involve coordinating Cd(II).[36] It has been suggested that Cys102 plays an analogous role to that of Cys11 in CadC in weakly coordinating the bound Cd(II). However, Cys102 has also been shown to be the key allosteric ligand, much like Cys7 in CadC, suggesting a different mechanism of allostery between CmtR and CadC. Parallels can be drawn between the putative CdS_2—S coordination in CmtR and the putative CdS_3—S coordination in CadC; it is interesting to note the existence of distorted Cd(II) complexes with weakly coordinating ligands as being important in metal resistance and selectivity.

A discussion of metalloregulation would not be complete without mentioning MerR, a canonical transcriptional activator for Hg(II) resistance genes. This family of proteins also included members that sense metals other than Hg(II), such as Zn(II), Au(I), Cu(I), Cd(II), and Pb(II). In contrast with the ArsR/SmtB family, the MerR family all binds metals in the same region and metals are distinguished by the primary and secondary coordination spheres.[37,38] Two members of this class, CadR and PbrR confer selective resistance to Cd(II) and Pb(II), respectively. In the context of coordination chemistry of toxic metals, this class is interesting based on its lower coordination number as well as its lack of chelate motifs. CadR has cysteine residues located at positions 77, 112, and 119, which are thought to form the coordination environment, much as homologous ligands in MerR coordinate Hg(II) trigonally.[39-41] Unfortunately, very little spectroscopic characterization has been carried out on CadR (and

PbrR), so other than their existence we know very little about what the spectroscopy of such sites might look like.

Any extensive discussion on metal resistance and regulation reveals that subtle differences in coordination number, geometry, and ligand identity have consequences for the efficacy and directionality of metal transfer through cells. Understanding these processes requires detailed spectroscopic and structural characterization of members of each class of protein involved in a given process, in addition to functional *in vitro* and *in vivo* studies. In focusing primarily on Cd(II)-sensing and transporting proteins as a way to generally understand toxic heavy metals with significant soft acid character several themes begin to emerge. In general, these proteins use cysteine residues arranged in chelate motifs to strongly coordinate thiophilic metals with functional asymmetry, although we do not understand whether this is a feature or why. The number of cysteine residues may vary to aid in distinguishing between metals, as for example, Cu(I) is often coordinated linearly (or digonally) while Cd(II) is coordinated tetrahedrally. The identity of the ligands may also be varied to confer important functionality, such as altered binding affinity or kinetics. For example, heavy metal transporters have cysteine ligands combined with harder ligands, perhaps to facilitate directional release of the toxic metal. To put it more simply, if you were designing a resistance scheme you would want the tightest binding to occur at the level of sensing so that the small changes in intracellular concentrations are detected immediately. Then, chaperones should have an intermediate level of affinity—not enough to interfere with sensing, but enough to bind strongly when there are toxic levels present. Finally, transporters should bind well (although if chaperones are involved maybe it is less important) but not too tightly or inertly—you want the protein to turn over and actually transport the toxic metal out of the cell. As it is, heavy metal transporters have slower turnover than analogous transporters for other species, which is likely due to the difficulty in balancing binding and release thermodynamics and kinetics. This general scheme is essentially what you see happening—CadC and its ilk bind tightly (~10^{12} M^{-1}) and the corresponding transporters bind less tightly (~10^{8} M^{-1}).[27] So then, understanding these systems requires not only studies on the systems themselves, but also studies on analogous systems to help draw parallels and define the relevant spectroscopy. While small molecules can be a crystallizable way to help define these systems, proteins often behave much differently, so researchers have turned to using Cd(II)-

substituted proteins and protein design to generate relevant metal binding sites in protein environments.

Electron Transfer Sites in Native Proteins

Cysteine can form a part of the coordination sphere of any biologically relevant metal. It is present both as a part of larger clusters, such as in metallothioneins, FeMoCo (nitrogenase), Fe-S clusters, P clusters (nitrogenase), and H clusters (hydrogenases, all of them); and also as part of the coordination sphere of mononuclear sites such as zinc fingers, cupredoxins, and rubredoxin. Cysteine has a relatively high pK_a (8.5 in solution) and is a stronger σ and π donor in its anionic form, resulting in highly covalent complexes with characteristic, low-energy charge transfer

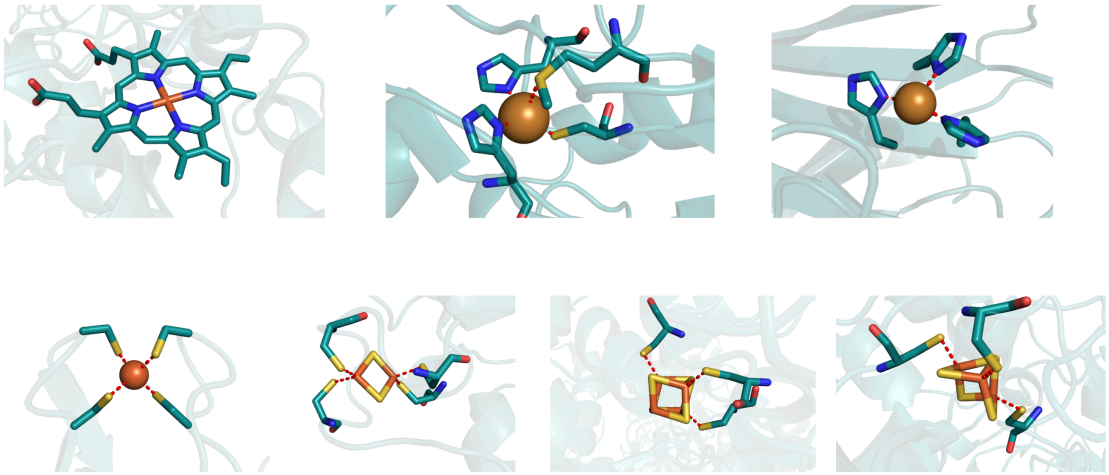


Figure 1.4 Representations of electron transfer centers found in native proteins. *Clockwise from top left:* heme , Type 1 Cu center, Type 2 Cu center, 3Fe-4S center, site-differentiated 4Fe-4S center, 2Fe-2S ferredoxin, and 1Fe rubredoxin. PDB codes: 3ZCF, 3USE, 1PHM, 2B3Y, 3AV8, 2DSX.

(CT) transitions. This behavior is responsible for the spectroscopic signatures of both cupredoxins[42] and iron-sulfur centers[43]. Similar to cysteine, inorganic sulfide also forms highly covalent complexes in biological contexts that display analogous low-energy CT transitions. The non-redox metals Zn(II), Cd(II), and Hg(II) bind to cysteine to form structurally stabilizing sites, interfere with cysteine reactivity, as well as providing sensing mechanisms of detoxification pathways for these metals. These sites are often difficult to characterize spectroscopically, and the mechanisms by which seemingly identical sites can distinguish between metals is still not well-understood. This topic will be covered in more depth in Chapter 2.

Electron transfer metalloproteins come in three major flavors (Figure 1.4): cytochromes, iron-sulfur centers, and mononuclear copper proteins.[5] Cytochromes are heme-containing proteins in which the heme iron is six coordinate and cycles between the Fe(II) and Fe(III) forms. Copper proteins are generally classified by their spectroscopic and functional characteristics. Types 1, 2, and 0 are all mononuclear sites that are distinguished primarily by their UV-visible and electron paramagnetic resonance (EPR) spectra. Binuclear copper centers such as Cu_A sites make up a fourth class of copper-based electron transfer site. Type 1 copper sites are characterized by a strong absorption band near 600 nm (~5000 cm^{-1}) which is attributed to a cysteinate-sulfur π to Cu(II) ligand-to-metal charge transfer (LMCT) transition. This highly covalent interaction, arising from the presence of cysteine in the primary coordination sphere, also perturbs the hyperfine splitting of the EPR signal leading to smaller $A_{||}$ hyperfine coupling constants.[44] These sites are classically understood to be electron transfer sites, with a relatively high reduction potential (+250 to + 350 mV vs. NHE). Type 2 copper sites, or "normal" copper sites, have lower reduction potentials, exhibit very weak d-d transitions, and have similar $A_{||}$ to aqueous copper complexes; these sites are typically coordinated by three histidines and have been implicated in both electron transfer and catalytic roles.[45] Type 0 copper centers are a recently discovered site formed by replacing the soft sulfur ligands with hard carboxylate ligands, resulting in a pseudo-tetrahedral coordination sphere.[46] These sites lack the strong absorption band at 600 nm, but maintain a high reduction potential and the compressed $A_{||}$ of typical Type 1 copper proteins. Cu_A sites, or purple copper sites, are a binuclear copper electron transfer site found in terminal oxidases with two cysteines that bridge the copper atoms forming a mixed valence site capable of one electron transfers.[5,47]

Iron-sulfur centers are ancient catalysts and ubiquitous across all branches of life, as they would have formed relatively easily in an anaerobic atmosphere and have been readily acquired by primitive enzymes; for these reasons they are ubiquitous across all kingdoms of life. They feature prominently in all metabolic pathways, especially respiration and photosynthesis, both of which feature multiple membrane-associated complexes with multiple Fe-S centers to facilitate electron transfer. The multiple oxidation states available to iron, combined with the chemical versatility of sulfur, make these centers ideal for electron transfer and storage. Iron-sulfur centers are also involved in certain catalytic roles, radical stabilization, and sensing.[48] The most common forms of iron-sulfur centers are characterized by their iron-sulfur cores: mononuclear $Fe(Cys)_4$,

binuclear 2Fe-2S , trinuclear 3Fe-4S, and tetranuclear 4Fe-4S, as well as higher order complex clusters (Figure 1.4). The $Fe(Cys)_4$ site found in the rubredoxin domain family and the small, soluble protein, desulforedoxin, is the simplest Fe-S center, containing no inorganic sulfide. There are two types of 2Fe-2S centers, ferredoxin and Rieske, which differ in their protein ligands and reduction potentials, but still function in electron transfer systems. Ferredoxin clusters are ligated to the protein by four cysteinate residues while Rieske centers are ligated by two cysteine residues and two histidine residues; the replacement of two anionic ligands for two neutral ligands is partially responsible for the +300 mV shift in reduction potential.[49] Trinuclear

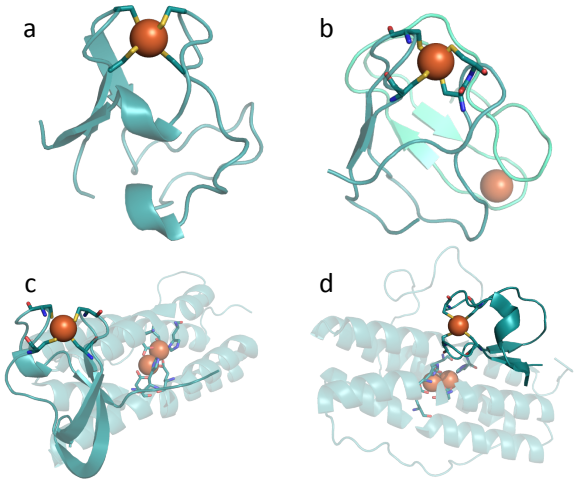


Figure 1.5 Crystal structures of important members of the rubredoxin family **a)** Rubredoxin (2DSX), **b)** Desulforedoxin (1DXG) **c)** Rubrerythrin (1LKO), and **d)** Nigrerythrin (1YV1)

clusters are known to perform electron transfer reactions, as well as undergo metal substitution reactions, which are part of the lifecycle of site-differentiated 4Fe-4S clusters.[50,51] Tetranuclear clusters are largely involved in electron transfer as either ferredoxin clusters or as the cofactor in High Potential Iron Proteins (HiPIPs), which differ in their active redox states and reduction potentials. While both types of cluster have the same resting state, $[Fe_4S_4]^{2+}$, the physiologically relevant reaction for HiPIPs is the oxidation at -150 mV to -350 mV to the state $[Fe_4S_4]^{3+}$, while in ferredoxins reduction at -250 mV to -500 mV results in $[Fe_4S_4]^{+}$.[48,52,53] While the sites are geometrically identical, HiPIP clusters are buried in the hydrophobic interior of the protein and this differing solvent access has been shown to affect the covalence of the cluster and consequently their relative reduction potentials and biologically relevant redox states.[53]

Rubredoxin was first described as a small, soluble protein by Lovenburg and Sobel in 1965 in extracts from a nitrogen-fixing anaerobe, *Clostridium pasteurianum* (Figure 1.5a).[54] Rubredoxins have been found in many anaerobic and micro-aerophilic bacteria, and have been extensively studied in sulfate-reducers in particular. Rubredoxins are small, redox-active electron transfer proteins (6-7 kDa) that coordinate iron pseudotetrahedrally with four thiolates from cysteine residues. A pair of canonical Cys-X-X-Cys motifs are responsible for forming the iron binding site. The overall fold of the protein is defined as alpha + beta, with 2-3 α-helices and 2-3 β-sheets. The high spin iron center exists in either the +2 or +3 oxidation state and performs one-electron transfer reactions, exchanging between the two available oxidation states. In fact, the soluble protein rubredoxin could be considered a member of a larger class of proteins containing rubredoxin-like domains. The domain retains the 3-stranded anti-parallel β-sheet and iron coordination and is found as a component of rubrerythrin (Figure 1.5c)[55-57], flavorubredoxin[58-60], and nigerythrin (Figure 1.2d)[61-63]. The reduction potentials of soluble rubredoxin vary between -90mV and 40mV.[54,64-66] However, the rubredoxin domain of rubrerythrin has a much higher reduction potential, +230 mV.[55] The idea that this might be due to distortions around the iron site

Table 1.2 Compared redox properties of Fe-S clusters

Cluster type	Cluster charge[a]	Species	Redox Potential (mV vs. NHE)
$Fe(SCys)_4$	-1/-2	Fe(III)/Fe(II)	-90 to 40
			(+230 for rubrerythrin)
2Fe-2S	-2/-3	Fe(III), Fe(II)/Fe_2(II)	-150 to -400
3Fe-4S	-2/-3	Fe_3(III)/Fe_2(III), Fe(II)	-70 to -400
4Fe-4S	-2/-3	Fe_2(III), Fe_2(II)/Fe(III), Fe_3(II)	-200 to -500
	-1/-2	Fe_3(III), Fe(II)/Fe_2(III), Fe(II)	+100 to +400

[a]based on coordination with thiolates [b]from refs [5,48,69] and references therein

was quickly discarded as Mössbauer and EPR data showed nearly identical electronic structures.[55] Indeed, the idea that distortion is not necessarily responsible for large shifts in the reduction potential is borne out by the biophysical characterization of desulforedoxin (Figure 1.5b), which also possesses a $Fe(Cys)_4$ site, but one that is significantly distorted relative to rubredoxin and has a reduction potential of -35 mV.[67,68] Recent computational studies on multiple high resolution X-Ray crystallographic structures of rubredoxin and rubrerythrin suggest that several amino acids serve as structural determinants of this perturbed reduction potential, and in particular, the positive shift of reduction potential observed in rubrerythrin is

due to polar residues near the iron site that are capable of hydrogen bonding to the site itself or to the backbone of the coordinating cysteines.[57]

Table 1.3 Comparison of spectroscopic properties of monomeric rubredoxin proteins

Protein	UV-Vis λ_{nm} ($\Delta\varepsilon$ $M^{-1}cm^{-1}$)	Redox Potential (vs. SHE)	Mössbauer: ox	Mössbauer: red
Rubredoxin[74,75]	750 (200-300)	-90 to 40 mV	δ=0.24	δ=0.70
	490 (8850)		ΔE_Q= 0.5	ΔE_Q= 3.25
	380 (10800)		D=+1.9	D= +7.6
	280 (21300)		E/D=0.23	E/D=0.28
			η=0.2	η=0.65
			$A_{x,y,z}$= -16, -15.9, -16.9	
Rubredoxin mutant: C42S	645 (810)	-273 mV	δ=0.26	Form B
FeS_3O[76]	440 (4060)		ΔE_Q= -0.7	δ=0.79
	331 (8910)		D=+1.22	ΔE_Q= -3.27
	280 (14600)		E/D=0.33	D=+6.2
			η=0.3	E/D=0.15
			$A_{x,y,z}$= -25, -22.1, -24.5	η=1.0
Desulforedoxin[68]	507 (7000)	-35 ± 15 mV	δ=0.25	δ=0.70
	370 (11900)	pH 7.6	ΔE_Q= -.75	ΔE_Q= 3.55
	310 (6500)		D= +2.2	D= -6
	285 (9400)		E/D=0.08	E/D=0.19
			η=0.6	η=0.35
			$A_{x,y,z}$= -15.4, -15.4, -15.4	

A lot of work has been done to understand the structural determinants of reduction potential in Fe-S proteins. In addition to taking advantage of the natural variation in reduction potential (i.e. HiPIP vs ferredoxins and rubrerythrins vs rubredoxins (Table 1.2), these proteins have also been subjected to mutational analysis and structural characterization, complemented with computational studies, to try to correlate structural features and perturbations in the reduction potential. Rubredoxin is one of the best-characterized proteins both structurally and spectroscopically, and has been studied by many techniques including UV-visible spectroscopy, Mössbauer spectroscopy, electron paramagnetic resonance (EPR) spectroscopy, X-Ray absorption spectroscopy (XAS), magnetic circular dichroism (MCD) spectroscopy and NMR

spectroscopy.[54,70-73] Additionally, desulforedoxin provides an electronically perturbed cluster to study, while mutation of the iron ligands has allowed the characterization of FeS_3O centers all within related folds (Table 1.2). This system is then primed to be studied in alternate folds, to assess the effect that the overall topology and local secondary structures have on the spectroscopy, reduction potential, and reactivity of a thiolate-coordinated iron.

A detailed series of studies on rubredoxin mutants were undertaken to understand the structural determinants of the reduction potential of rubredoxin. Among rubredoxins from different species, the reduction potential varies by ~ 100 mV. These perturbations in reduction potential can be reproduced by single mutations at particular positions. These residues comprise the structural determinants of the rubredoxin system, meaning the locations and residues that have the largest effect on (or "determine") the functionality and reactivity of the protein.[57,77,78] In rubredoxin, mutational analysis and sequence alignment have identified several residues that are particularly important.[57,79] Specifically, the identity of the residues near the S_γ of the iron-coordinating residues; varying the identity can change the redox potential by up to 126 mV.[80] Residues 8, 41, and 44 (numbered according to *C. pasteurniam* rubredoxin) are the residues that have been identified as being particularly important. However, the cause or causes for the perturbation are not immediately obvious. For example, the mutation of Val44 to Ala raises the reduction potential by 50 mV.

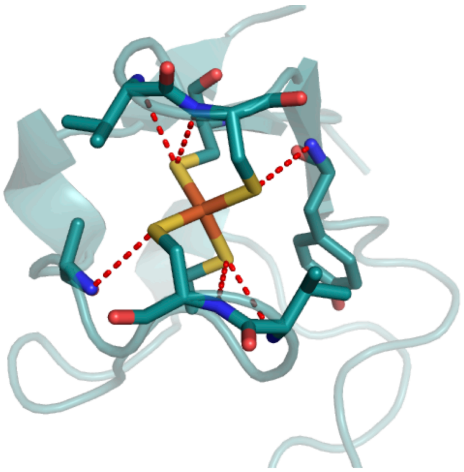


Figure 1.6 Representation of hydrogen bonding around the iron site in rubredoxins. The Sγ of the coordinating cysteine can hydrogen bond to backbone amides. PDB: 2DSX

The various factors that are likely, or have been hypothesized to modulate the redox potential are: the electrostatic environment imposed by the protein, solvent (water) accessibility, specific hydrogen bonds or dipoles to (or near) the site, and the covalency of the Fe-S bonds. S

K-edge EXAFS was used to investigate the variations in the covalency of the Fe-S bonds and it was discovered that there was no direct correlation between the covalency of the Fe-S bonds and the reduction potential.[81] Indeed, later studies suggested that bonding in the metal center remains relatively unperturbed amongst mutants and thus, the changes in redox potential are centered in movements of the protein backbone.[82,83] A series of combined crystallographic and computational studies investigated the electrostatics, hydrogen bonding, and solvent access to the site. Crystallographic evidence suggests that Leu41 can act as a "water gate," providing solvent access to S_γ, which helps to stabilize the higher charge (-2 vs. -1) on the reduced form.[84] Reducing the size of this residue by a Leu to Ala mutation further stabilizes the reduced form and causes a positive shift in the reduction potential. The effect of hydrogen bonds and specific dipole interactions has also been studied as there are six hydrogen bonding interactions in the immediate vicinity of the iron site (Figure 1.6).

Mutation of either of two glycine residues near the iron site to alanine results in a negative shift in the reduction potential, without changing the properties of the iron site itself. This shift was crystallographically determined to be due to strain around the site, and in extreme cases, inversion of nearby peptide bonds that alter the hydrogen bonding to the $Fe(Cys)_4$ site.[83] The change in reduction potential as a result of mutation of Val44 to Ala can be explained by shifts in the backbone and the formation of a new hydrogen bond to S_γ-42.[79,85] A detailed NMR study on a series of mutants at residue 8 and 44 used the amount of the electron spin delocalization to measure the strengths of nearby hydrogen bonds, and found that the average strengths of the hydrogen bonds around the iron site correlate linearly with the reduction potentials of the mutants.[80] Overall, the shorter the average hydrogen bonds around the site, the higher the reduction potential, as would be expected by more efficient electron delocalization. Furthermore, the electrostatics of the protein itself are important for its redox properties. In a computational study, calculation of the classical electrostatic potential from high resolution crystal structures combined with molecular dynamics simulations revealed that the peptide backbone is a major contributor to the reduction potential.[86] An additional study to investigate this phenomenon further found that the polarization response of the protein is best represented by a two component model that includes both a permanent component and a linear response component.[82] A comparison of simulations where the redox site is uncharged as well as the response of models which include only the backbone versus the backbone and polar residues

pinpoints the permanent potential component as being associated primarily with the peptide backbone, a result that is consistent with similar studies in [4Fe-4S] ferredoxins and HiPIPs.[77,78,87] The authors posit that the combination of a low dielectric constant inside the protein combined with the permanent potential at the redox site results in electron transfer reactions with relatively low activation energies and favorable driving forces.[82]

Rubrerythrin is an interesting case in that the rubredoxin domain has a similar fold to soluble rubredoxin, but it has a significantly elevated reduction potential relative to all isolated

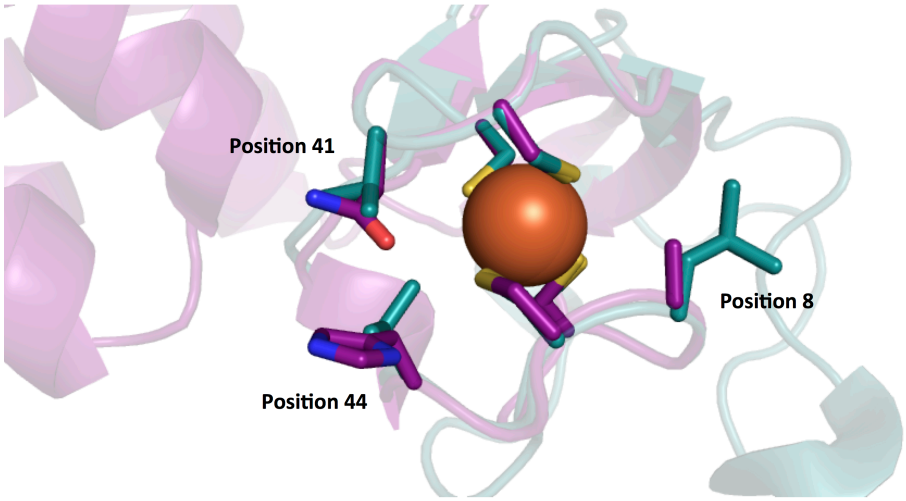


Figure 1.7 Overlay of the rubredoxin domains of rubrerythrin (purple) and rubredoxin (teal). Highlighted are positions 8, 41, and 44, which have been shown to be determinants of the reduction potential of the tetrathiolate site.

rubredoxins. Residues 8, 41, and 44 (using the numbering from *C. Pasteurianum*), which are important for the reduction potential of rubredoxin (*vide infra*), deviate from their conserved properties in rubrerythrin (Figure 1.7).[88] In rubrerythrin, polar groups (Asn and His) replace Val at positions 8 and 44, while Ala replaces Leu at position 41. A computational study examined the effects of these groups on the reduction potential and found that residues 8 and 44 are particularly important in modulating the reduction potential, while residue 41 is less important as its major role is water access in soluble rubredoxin and the rubredoxin domain sits at the interface of the rubrerythrin homodimer.[57] The polar residues in positions 8 and 44 allowed for multiple side chain conformations in both the reduced and oxidized states, which shift based on the state of the site and change the polarization response. Only changing two amino acids can have a striking effect on the reduction potential, due in part to the behavior of the specific side chains.

Previous Designs of Electron Transfer Proteins

Being able to design proteins that can bind multiple redox states of metals is important to understand both electron transfer as well as redox catalysis. Indeed, designed metal-binding proteins have been used to investigate long-range electron transfer processes, heme-binding proteins, and redox-active metalloproteins. For example, the Gray group has created a series of ruthenium-modified native heme proteins specifically for the study of intramolecular electron transfer rates.[89-91] Herein, I provide a concise review of protein design for electron transfer, followed by efforts to design rubredoxins, and finally a brief overview of the work that the Pecoraro laboratory has undertaken to place this work in context of the greater field of protein design. The metalloprotein design field has been exhaustively reviewed recently and I refer readers who wish any further detail to this recent review.[6]

The most progress in designing *de novo* proteins has come from working in α-helical scaffolds, which are known to form the secondary structures of numerous heme proteins in nature. Thus, heme proteins were some of the first targets for *de novo* design. This convenience, coupled with the wide array of accessible reactions catalyzed by heme has made them the subject of study for those particularly interested in electron transfer, oxygenases, and peroxidases. A great deal of progress has been made on the *de novo* design of heme proteins. As far back as 1989, amphipathic helices were used to construct a heme peptide capable of hydrolytic activity.[92] Efforts quickly began to focus on imparting native-like folding and native-like activity to designed proteins. The collaborative effort between Dutton and DeGrado is noteworthy in their success in developing heme-binding "maquettes" with cooperative binding, physiologically relevant redox potentials, and conformational specificity. These efforts were focused initially on hemes involved in electron transfer, which are six coordinate and have no open coordination sites, as *de novo* design worked to overcome the challenges posed by self-assembly and the rigorous symmetry this imposes. The progress in heme systems followed in lockstep with developments in *de novo* design as greater and greater control was achieved in these systems. The first redox-active *de novo* heme construct was published in 1994[93] and progressed quickly with structural characterization and a demonstration of proton-coupled electron transfer properties[94]. Investigations in this field mirrored those on native proteins as these constructs were examined on surfaces and electrodes, and in monolayers to understand potential applications.[95-

[101] These investigations were driven by a desire to understand both soluble electron transfer proteins as well as membrane-bound electron transfer centers. A model of cytochrome c oxidase, with both heme a and heme b cofactors was reported by Gibney et al. and is capable of enzymatic reduction by NAD(P)H and subsequent oxidation by O_2.[102] Cytochrome bc_1, which has an unusual perpendicular arrangement of the ligating histidines with respect to the hemes, was also successfully modeled, which has an unusual perpendicular arrangement of the ligating histidines with respect to the hemes. This water-soluble model reproduced the b subunit of cytochrome bc_1 by mimicking the residues around the heme-binding sites. Incorporation of a glycine residue allows close contact of the heme with the peptide backbone and addition of a threonine residue adds hydrogen-bonding to fix the position of the histidine.[103] More recently, these designs were combined to synthesize a transmembrane diporphyrin protein; a potentiometric titration showed two redox signals with apparent $E_{1/2}(Fe^{III}Fe^{III}/Fe^{II}Fe^{III})$ and $E_{1/2}(Fe^{III}Fe^{II}/Fe^{II}Fe^{II})$ of -97 (±3 mV) and -168 (±3 mV) vs NHE, respectively.[104] Furthermore, a c-type cytochrome maquette was designed that was capable of interfacing with the cellular heme machinery to incorporate covalently-attached heme c.[105] Titration of heme b into holopeptide adds a heme b at a second binding site. The hemes have reduction potentials of -188 mV and -250 mV for the heme c and b, respectively, meaning that electron transfer from b to c is thermodynamically favorable.

After success with the design and characterization of an O_2 transport protein[106], the limits of self-assembling constructs were explored by designing and characterizing a parallel series of single-stranded α-helical peptides for a diverse range of functions.[107] These constructs were capable of binding hemes, zinc porphyrins, zinc chlorins, and flavins, and were capable of functions such as electron transfer with native cytochrome c, photochemical charge separation, and light-activated electron transfer between different cofactors. This work demonstrated that the interhelical interfaces of pairs of helices in a four-stranded bundle were capable of functioning as domains and could be engineered separately and combined to yield a diverse range of functions.

Marcus Theory

Electron transfer is an important process that plays a key role in sustaining life, via photosynthesis and respiration, cycling nutrients and minerals, as well as playing a key role in the development of new technologies for green energy. Understanding its basic principles can

help us predict what types of processes are important in regulating the bioavailability of nutrients, what pathways or proteins are important within an organism, and how to design new technologies for the synthesis of fuels and industrially relevant compounds. Typically, bimolecular electron transfer can be thought of as occurring in the following steps:

1. Formation of a precursor complex,
2. Electron transfer, and
3. Dissociation of a successor complex. [108]

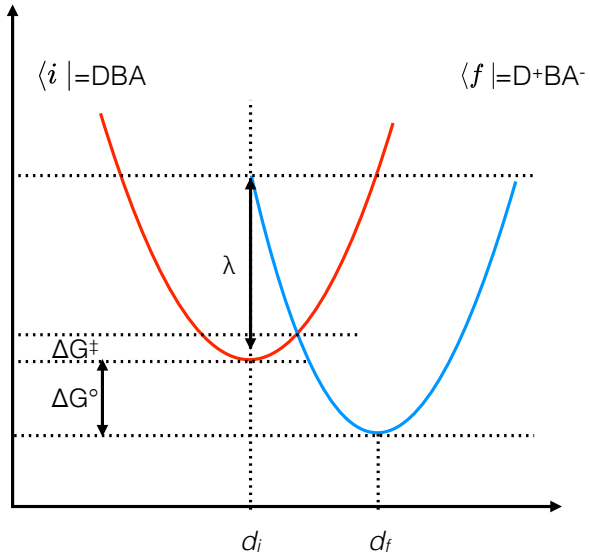


Figure 1.8 Reaction states can be approximated by parabolas for each the reactant (red) and product (blue) surface. The parabola represents both the nuclear coordinates as well as solvent molecules in each state. Reorganization energy, λ, can be considered as the energy required to force the reactants and the products to have the same nuclear coordinates. Modified from ref 111

The exact details of these steps may vary based on whether the electron transfer occurs at an electrode surface, as a bimolecular reaction, or as an intramolecular reaction, as in between two redox sites in a protein. The rate constant of an electron transfer reaction can be classically modeled by an equation that nominally takes the form of the empirical Arrhenius equation for chemical kinetics:

1. $$k_{ET} = \kappa A \sigma^2 e^{-\Delta G^*/RT}$$

Rudolph Marcus partitioned equation 1 by applying the Franck-Condon principle to electron transfer reactions such that the transition state (and activation energy) are a consequence of the inner and outer sphere effects of many factors, including vibrational states, relative positions of reactants, positions of solvent molecules, dipoles, and the dielectric constant of the environment.[109,110] Thus, the rate of reaction is dependent on the distance between centers, the nature of the centers, and the environment around and between the centers at the transition state. The semi-classical form of this equation then follows as:

2. $$k_{ET} = \frac{2\pi}{\hbar} H_{AB}^2 \frac{1}{\sqrt{4\pi\lambda RT}} e^{-(\Delta G^\circ + \lambda)^2 / 4\lambda RT}$$

This transition state occurs at the intersection of reactant and product free energy surfaces, which can be approximated by parabolas (Figure 1.8)[111]. Marcus related the free energy of activation, ΔG*, to the free energy of the reaction and a parameter termed reorganization energy (λ) (3).[109] He also introduced terms that take into account the distance dependence of the electron transfer rate.

3. $$\Delta G^* = \frac{(\Delta G + \lambda)^2}{4\lambda}$$

Reorganization energy is comprised of both an inner sphere and outer sphere term (4) and can be thought of as the energy required for the products to have the same nuclear coordinates as the reactants without electron transfer (Figure 1.8):

4. $$\lambda = \lambda_i + \lambda_o$$

The outer sphere term, using a two-sphere model of the reactants where *a* is the radius of each species and *Δe* is the charge transferred, (5) is dominated by the separation distance of the reactants (*R*) and the dielectric constant of the solvent (D_{op} and D_s). The inner sphere term arises from the changes in bond lengths in the different redox states (6).

5. $$\lambda_o = (\Delta e)^2 \left(\frac{1}{2a_1} + \frac{1}{2a_2} - \frac{1}{R}\right)\left(\frac{1}{D_{op}} - \frac{1}{D_s}\right)$$

6. $$\lambda_i = \frac{1}{2}\sum_j k_j (Q_j^r - Q_j^p)^2$$

An important consequence of Marcus theory is the relationship between ΔG*, ΔG°, and λ.[109,110] In the normal region, as the magnitude of ΔG° gets larger, the reaction gets faster up to the point where -ΔG°=λ and ΔG* is zero (Figure 1.9). This is the maximal rate of the reaction. If the magnitude of ΔG° continues to get larger, Marcus predicted an "inverted" region, where the rate of reaction will begin to slow again as the ΔG* begins to increase again.

So far we have not considered the transition state explicitly and the nature of the charge transfer step itself, which is contained in the κ term of eq. 1 and shown explicitly in eq. 2. For an

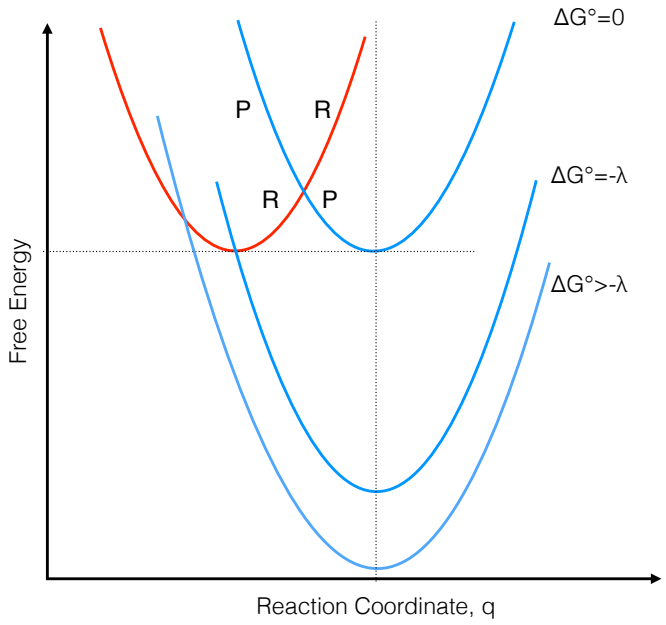


Figure 1.9 Overlay of reactant (R) and product (P) curves as the magnitude of ΔG° changes. As ΔG° becomes more negative (moving vertically along the dashed line) the ΔG* decreases, reaching a minimum when the magnitude of ΔG° equals λ . As ΔG° becomes more negative, the ΔG* begins to increase again in what is termed the "inverted" region. Figure modified from 110

electron transfer to occur there must be some degree of electronic coupling between the donor and acceptor, and this is represented by H_{AB} (or V_R, depending on the source). In a real system, the parabolas do not intersect as previously pictured, but rather are split according to the

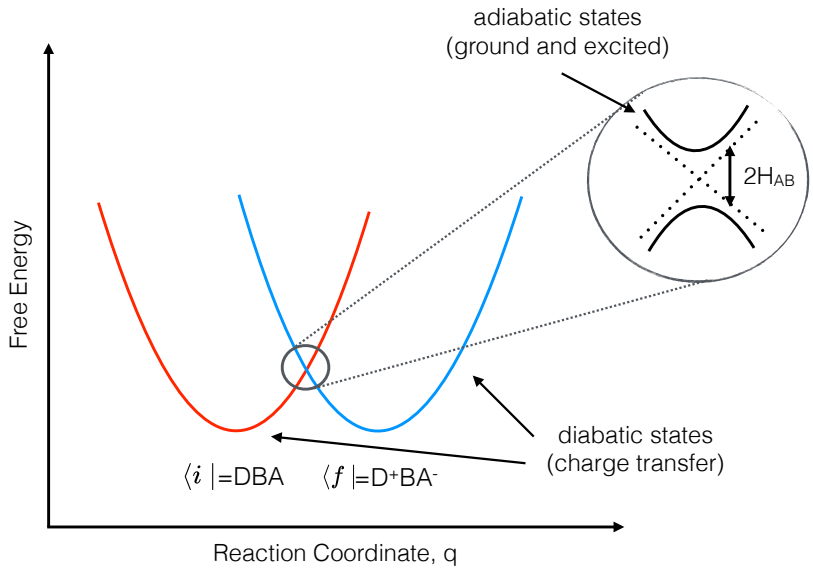


Figure 1.10 The general picture of two parabolas is reflective of a diabatic system where coupling is infinitesimally small. In a real system the parabolas do not actually cross—rather there is a splitting between the excited and ground states of the system. The magnitude of this splitting is $2H_{AB}$. Figure from 111

magnitude of the electronic coupling between the donor and acceptor (Figure 1.10).

If H_{AB} is small (non-adiabatic), once the transition state nuclear configuration is formed there is a low probability that the reaction will jump to the product potential energy surface. If H_{AB} is large (adiabatic) then the reactants will always cross to the product potential energy surface, which is most likely to happen when orbitals can overlap and the probability of successful electron transfer approaches 1.[108] However, in metalloproteins, the distances covered by the electron are large and the electronic coupling between the donor and acceptor is small. The donor/acceptor (D/A) interaction is weak and the transition state must be reached many times before an electron is transferred. Proteins also lower reorganization barriers by sequestering redox sites in hydrophobic environments so there is low energy required to repolarize the medium after electron transfer. A key factor in the rate of biological electron transfer is the degree to which the intervening protein medium can propagate the donor and acceptor electronic wave functions across the separation during electron tunneling. H_{AB} at a particular distance, H_{AB}^{R}, is typically expressed as,

7. $$H_{AB}^{R} = H_{AB}^{o} e^{-\frac{1}{2}\beta R}$$

where the maximum electronic coupling, H_{AB}^{o}, is modulated by the distance and a parameter, beta, which represents the effect of the intervening medium in facilitating electron transfer. The effect that the protein has on electron tunneling is then contained in the beta parameter. Systems in which the electron transfer proceeds by superexchange, that is, electron transfer processes that abide by a two-state model, show a strong distance dependence where the rate of electron transfer drops dramatically with distance.[90,109,112-115] There has been quite a bit of debate as to the protein's role in mediating this electronic coupling.[112-114,116,117] One possibility is that the intervening medium acts as a simple homogenous barrier to electron transfer and beta will have a constant value for all proteins.[112] Another possibility is that the structure of the intervening medium has an effect and this will modulate beta and thus the degree of coupling between two separated redox centers.[113] In this case, the electron transfer rate is pathway-dependent and Nature may choose certain secondary structures or amino acids to facilitate faster electron transfer. This is distinct from systems in which hopping is the primary mechanism and amino acids act as relays, as in ribonucleotide reductase.[90,118,119] These systems show a weaker distance dependence and faster electron transfer rates for electron transfer over 20 Å than a simple superexchange model.[114] The rates of these reactions are dependent on the number of hops,

which can each be evaluated as a single superexchange step following a single exponential kinetic expression.

Long-range electron transfer theory

One of the major ways that the effect of the protein medium has been studied is by time-resolved spectroscopy to isolate the kinetic parameters of different electron transfer steps. These reactions can be monitored in native systems like photosynthetic reaction centers or in semi-artificial systems like proteins modified with ruthenium chromophores. Comparative analysis and computational studies have yielded slightly different interpretations of the role of the protein medium and how the rate of an electron transfer reaction varies over distance. The rates measured in native systems are constrained by the parameters extant within the system; in other words, the driving force, distance, and orientation are hard to change to test various hypotheses. However, redox centers like quinones have been replaced with different types of non-native quinones to change driving force while leaving distance constant.[120,121] Other ways to assess the driving force and reorganization energy include varying temperature, pressure, isotopic composition, pH, ionic strength, magnetic and electric fields, and genetic manipulation.[122] In a separate series of experiments to study the effect of the protein medium on electron transfer rate, researchers appended ruthenium chromophores to the exterior of electron transfer proteins.[113] By this method, the rates of electron transfer have been measured in over thirty ruthenated metalloproteins.[116] Additionally, varying the location of where the ruthenium chromophore is placed allows for varying the distance of the electron transfer. This creates a system that can be laser-triggered with a reaction that has a high driving force and that can be studied over a range of distances. These electron transfers are studied at k_{max}, in the regime where $-\Delta G^{\circ}=\lambda$ and ΔG^{*} is zero.[110,112] Having a good estimation of the optimal rate for these reactions is important, since then the rates are largely dependent on the distance and electronic coupling (and by extension, β) (eq. 7).

Based on analysis of electron transfer rates in native systems as well as in semisynthetic and synthetic systems, Dutton et al. proposed a uniform barrier model to describe electron transfer via superexchange.[112] In this model, all protein media exert uniform effects on the electronic coupling, and a single β can be used for all intraprotein electron transfers. The argument then follows that distance and positioning are the forces by which nature engineers

proteins for electron transfer, rather than the selection of particular pathways or secondary structures. A plot of the logarithm of the k_{max} for electron transfer versus the distances between donor and acceptor in native and ruthenated proteins shows a linear trend to which both native and semisynthetic systems conform. The slope of this plot is β, which was found to be 1.4 $Å^{-1}$ and represents a 10-fold change in rate for every 1.7 Å change in distance.[112] The semi-classical equation (2) can be simplified to yield in the absence of any knowledge of the protein structure,

8. $$logk_{ET} = 15 - 0.6R - 3.1\frac{(\Delta G° + \lambda)^2}{\lambda}$$

When some details about the protein structure are known, the packing density (ρ) of the protein can be used to weight β:

9. $$logk_{ET} = 13 - (1.2 - 0.8\rho)(R - 3.6) - 3.1\frac{(\Delta G° + \lambda)^2}{\lambda}$$

An analysis of reactions within photosynthetic reactions centers shows that electron transfers within these proteins conform to this model.[117] Dutton suggested that given the inherent uncertainties in experimental techniques, this model adequately predicts electron transfer rates in proteins. A comparison of packing densities versus possible electron transfer pathways in structurally characterized oxidoreductases reveals that both the packing and pathways are essentially equivalent along biologically productive electron transfer routes and nonproductive electron transfer routes, suggesting that there is no correlation between path and function.[117] Furthermore, a survey of donor-acceptor distances in native proteins suggests a limit of ~ 14Å for single step electron transfer reactions. The authors argue that within this regime, there is no selective pressure for increased rates on the pathway itself; all the selective pressure for faster rates is on the distance between cofactors.

However, the success of some computational work, as well as a series of studies of ruthenated proteins, have led others to propose that the pathway through which an electron tunnels affects the degree to which distance modulates the electron transfer rate. The classic superexchange model, which described coupling between two distant redox sites through a simple bridge was modified for proteins to account for the rate scatter attributable to protein structural complexity.[123,124] This algorithm classifies three important linkages in proteins through which electrons can tunnel: covalent bonds, hydrogen bonds, and through-space jumps (eq.10).

10. $$k_{ET} \propto \left\{\prod_i \varepsilon_i^{bond} \prod_j \varepsilon_j^{bond}\, e^{-\beta^{space}(r_j - r_o)/2}\right\}^2$$

Based on this model, different secondary and tertiary structures would have subtly different coupling efficiencies.[124] For example, electron transfer through a β-strand, which is mediated by

the backbone covalent bonds in an extended arrangement is more efficient than coupling through an α-helix where the electron travels along the covalent backbone. The calculated β for a β-strand was 1.0 $Å^{-1}$,[124] whereas that of an α-helix was 1.97 $Å^{-1}$ [125]. However, if an electron travels through the hydrogen bonds that run parallel to the helical axis, the coupling becomes better and β equals 1.26. The strength of the hydrogen bonding and how close the coupling through the hydrogen bonds behaves as a covalent bond determines whether the coupling in α-helices is competitive with β-strands (β=0.9).[116,124] Extending this analysis to larger structural elements, electron transfer across helices or through β-sheets becomes highly dependent on the degree of orbital overlap in the interactions that stablize certain structural features.(Beratan; Gray 1999) Combined with this computational analysis, electron transfer rate measurements in ruthenated proteins were used to understand the coupling between redox centers in proteins. In particular, a series of six ruthenated azurins, where the ruthenium chromophore has been appended to histidine residues at different locations on the surface provided a detailed metric of electronic coupling along β-strands and through β-sheets.[125,126] The electronic coupling through a β-strand was found to be 1.1 $Å^{-1}$, which is very close to the predicted 1.0 $Å^{-1}$.[125] Comparison over a larger group of ruthenated proteins reveals that the distance dependent electron transfer kinetics for most fall within the bounds of β = 1.0-1.3 $Å^{-1}$ (β-sheet and α-helix, respectively)(Figure 1.11).[116]

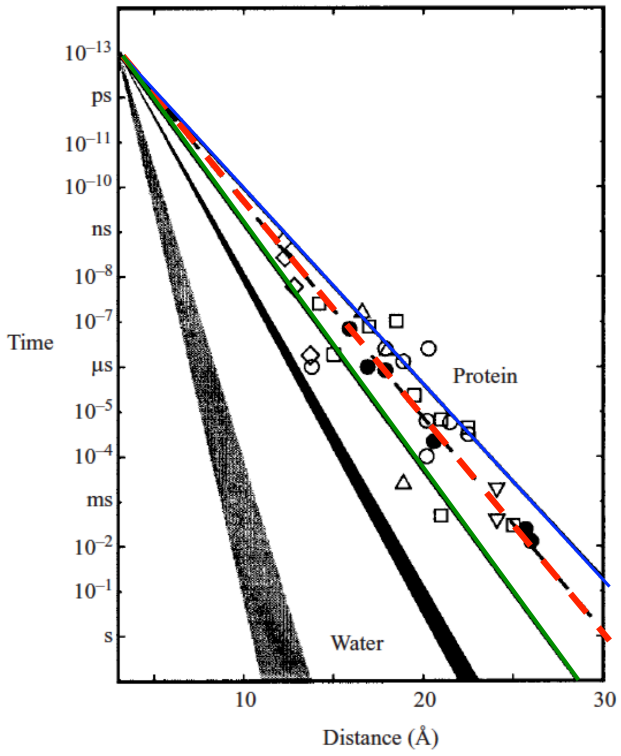


Figure 1.11 Tunneling timetable for electron transfer in ruthenium-modified proteins where azurin (•), cyt *c* (O), myoglobin (Δ), cyt b_{562} (□), HiPIP (open rhombus), and Fe:Zn-cyt *c* cystals (upside down triangle) are represented. The green line represents tunneling pathway prediction for an α-helix based with β= 1.3 $Å^{-1}$. The blue line is the predition for β-sheet (β = 1.0 $Å^{-1}$) and the red dashed line is average protein (β = 1.1 $Å^{-1}$). The grey wedge is the tunneling prediction through vacuum and the black wedge is the tunneling prediction through water. Modified from ref 114

While it would seem that then an average value could be applied to electron tunneling in proteins, a closer inspection reveals that at the same tunneling distance, there are differences of three orders of magnitude for the rate of electron transfer between donor and acceptor (Figure 1.11). This would indicate that there are some structural features that differ that are not accounted for by an uniform barrier model.[114,116] To address this issue and incorporate newer data on cyt b_{562}, Beratan and coworkers proposed that some protein structures dynamically average multiple coupling pathways and so appear to have a uniform barrier, while others favor a certain pathway that may have electronic coupling that is quite different from the "average" coupling through the protein medium.[127] Although each of the datasets on cytochrome b_{562}, cytochrome *c*, and myoglobin have one or two anomalously slow rates, nonetheless the authors showed that the rates all conform to a full quantum mechanical treatment that takes into account conformational averaging, solvation, and multiple pathways. Thus, current theories are capable of predicting and understanding this behavior. To understand more fully why these data points were slower than predicted by a simple distance dependence, a coherence parameter ($0<C<1$) was calculated to describe the coupling interactions between the donor and acceptor such that a low coherence parameter indicates that numerous pathways contribute to the overall coupling while a large coherence parameter suggests a single pathway dominates the coupling.[127] When compared to the measured rates, those rates that were anomalously slow had higher (~0.5) coherence parameters.[127] An examination of the pathways revealed that those that were dominated by a single pathway were coupled to the heme iron through the axial ligand, while systems where multiple pathways contributed were coupled along the heme edge. Thus, there exist two regimes: in the first, termed the small coherence/multiple pathway regime, there is a strong distance dependence observed for the rate constant of electron transfer and an average β can be applied to describe the rate. In this regime, the evolutionary linkage between fold and electron transfer is weak because the distance term dominates the coupling function. Second, is the high coherence/dominant pathway regime where either a strongly or weakly coupled pathway dominates the coupling and the structure can produce an effect on the electron transfer rate that varies several orders of magnitude.

While most of this theory has been directed at understanding processes in native proteins, the principles laid out in these works can be extended to guidelines and hypotheses for the design and engineering of novel proteins. These considerations become important as we consider how

best to design multi-site proteins that are capable of electron transfer. By and large, the two factors that need to be controlled in single electron transfer steps are the driving force of the reaction and the tunneling distance between redox cofactors. However, even unfavorable driving forces can be overcome in multi-step reactions when the distances between two redox cofactors are small enough and the driving force for subsequent transfer is favorable.[128,129] In general, productive electron transfers occur over distances between 4 and 14 Å, so two-site constructs where only a single step is intended should have the sites spaced accordingly. Additionally, it is likely that electron transfer reactions will not be subject to a strong fold and pathway dependence, so in principle, a fold only need stabilize both redox states of the associated cofactors for efficient electron transfer to occur.

Type 2 Copper Sites

Biogeochemical cycles regulate the bioavailability and paucity of nutrients over geographical distribution, phase distribution, and time. Nitrogen is both a major component of the atmosphere as well as present in all biomolecules. Fixing of nitrogen is the process by which non-bioavailable $N_2(g)$ is converted into NH_4^+, which is bioavailable. From here, NH_4^+ can be converted into biomolecules or undergo nitrification to produce NO_3^-. Denitrification is the process by which NO_3^- is converted back to N_2:[130]

$2NO_3^- \rightarrow 2NO_2^- \rightarrow 2NO \rightarrow N_2O \rightarrow N_2$

Nitrite reductases that catalyze the conversion of nitrite into nitric oxide are either copper proteins or cytochromes *cd*$_1$. Copper nitrite reductases have two copper sites, one of which is a type 1 copper that transfers an electron to the catalytic type 2 site.[131] In its resting state, type 2 copper is coordinated by three histidine residues and a water molecular in a distorted tetrahedral coordination.[132,133] A current accepted mechanism states that nitrite binding displaces this bound water, binding in a bidentate fashion with both oxygens coordinated to the copper.[132] Reduction of the copper and rearrangement of the nitrite cleaves one of the nitrogen-oxygen bonds, releasing water, leaving NO^+ bound to Cu(I) with nitrogen coordinated to the copper, in a nominally side-on mode. Reduction and water displacement of the NO returns the protein to the Cu(II) starting state.

Type 2 copper sites can also function as electron transfer sites, as in peptidylglycine monooxygenase (PHM). PHM has two type 2 copper sites, one with His_3 coordination, denoted

Cu_H, and other with His_2Met, known as Cu_M, and these play different roles in the protein.[134,135] In PHM, electron transfer occurs from the tris-histidine-ligated copper (Cu_H) to the catalytic site, Cu_M. The $Cu(His)_3$ site can also play catalytic roles in other proteins such as copper amine oxidase and quercertin 2,3 dioxygenase.[136]

Protein Design in the Pecoraro Lab

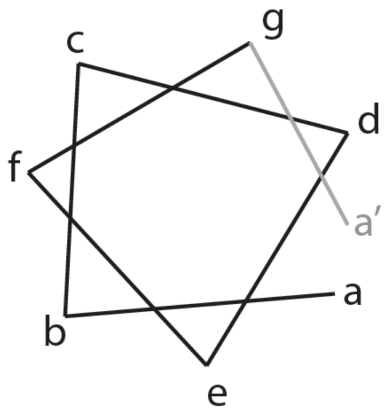


Figure 1.12 Schematic illustrating the spatial relationship of the heptad repeat within the helix wheel. Residues are labeled **a-g** where **a'** is the **a** residue of the following repeat.

The α-helical regions of the *de novo* designed α-helical proteins are based on the heptad repeat strategy. Thus, seven amino acids form repeats (with residues **a**-**g** see Figure 1.12[137]) in which the **a** and **d** positions are occupied by hydrophobic residues that face inwards in the coiled coil or helix bundle, driving the folding and association of the peptides. Substitution of a residue in the **a** or **d** position with a coordinating amino acid provides a metal-binding site. The **TRI** family of peptides consists of four heptad repeats with the sequence LKALEEK, which self-assemble into three-stranded coiled-coils above pH 5.5.[138,139] This can be expanded to five repeats to generate the **GRAND** series or contracted to three repeats to generate **BABY**.[140] The interior-facing leucine residues in **TRI** can be substituted for metal-chelating residues to generate metal binding sites. In particular, leucine to cysteine substitutions have been used to examine heavy metal binding to thiolate-rich sites. The kinetics and thermodynamics of metals such as Cd(II), Hg(II), Pb(II), and As(III) have been examined in detail. To review briefly, it was found that uncommon coordination geometries could be generated in these three-stranded coiled coils that mimic those types of distorted metal sites often found in metalloproteins. Due to the rotation of the peptide backbone, the **a** and **d** sites have slightly different coordination characteristics,

which orient the side chain either directly into the hydrophobic cavity of the peptide or angled toward the inter-helical interface. This results in a size difference between the two cavities for metal binding, which affects the binding affinities of metal ions to **a** and **d** sites. For example, Cd(II) binds to **TRI**L12C, a **d** site peptide, with a binding constant of 2.6 x 10^7 M^{-1} while it binds

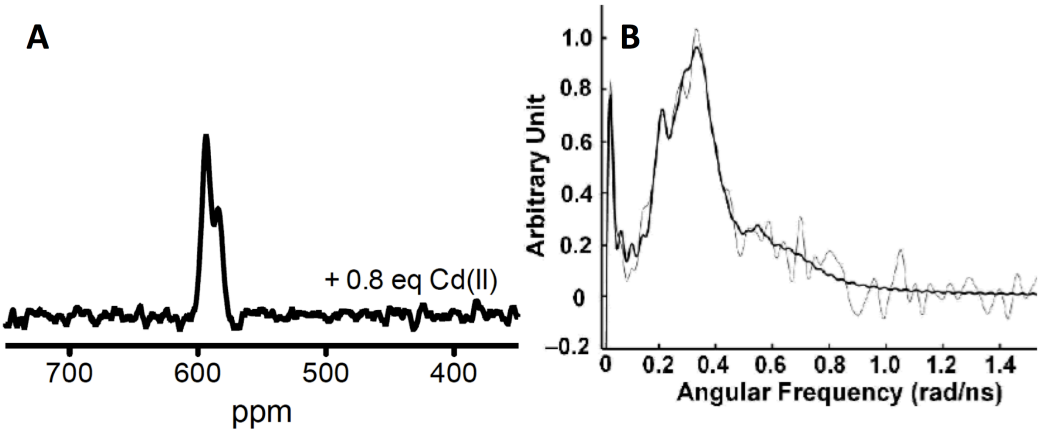


Figure 1.13 a) ^{113}Cd NMR data on α_3**DIV** shows two resonances at 595 and 583 ppm b) perturbed angular correlation spectroscopy (PAC) can be fit to three coordination spheres: CdS_3O (two conformations) and CdS_3N. From ref 149.

to **TRI**L16C with a binding constant of 1.6 x 10^8 M^{-1}.[141] Additionally, the protonation equilibrium associated with metal binding is different for the two metal binding sites; the pK_{a2} for Cd(II) binding is approximately one unit lower for **a** sites than **d** sites.[142,143] A solely three-coordinate, trigonal planar CdS_3 can be enforced by using penicillamine instead of cysteine to coordinate the Cd(II), thus rearranging the hydrophobic packing around the site.[144] Additionally, the packing around the site can be manipulated by engineering a "hole" above the site by substituting a leucine for an alanine, which enforces a tetrahedral Cd(II) coordination.[143] Furthermore, a similar effect can be achieved by using D-amino acids like D-leucine to enforce a trigonal planar CdS_3 site.[145] It is also possible to combine these two strategies to produce peptides that can selectively coordinate two Cd(II) in two different environments—three-coordinate and four-coordinate.[146] Much has been learned from these studies on heavy metal binding to three-stranded coiled coils. In an effort to extend our understanding of how heavy metal binding is affected by the dynamics of the protein, work has been initiated to generate single-stranded three-helix bundles for similar studies.

A three-helix bundle protein, designated α_3**D**, was designed by adding loop regions and varying the sequence of the heptads of CoilSer, the crystallographic analogue of **TRI**, to create a

protein that exhibits native-like folding and conformational specificity.[147,148] This designed protein exhibits native-like folding and stability, and its solution structure has been reported.[147] Our group then modified α_3**D** to incorporate three cysteine residues (α_3**DIV**) and was shown to bind mercury, cadmium and lead, analogously to S_3-type sites in the self-assembling coiled coil peptides.[1] In α_3**DIV**, cadmium binding was studied using UV-vis spectroscopy, ^{113}Cd NMR spectroscopy, and ^{111m}Cd perturbed angular correlation spectroscopy (PAC) (Figure 1.13).

Through UV-vis spectroscopy, it was found that Cd(II) binding to the three thiolates exhibits a pK_{a2} of 10.6 ± 0.1, which represents a metal-ligand deprotonation equilibrium with the simultaneous dissociation of 2 protons. ^{113}Cd NMR spectroscopy showed two resonances at 595 ppm and 583 ppm, similar to those observed in coiled-coil peptides with a CdS_3O coordination environment.[143] ^{111m}Cd PAC was performed to gain further information about the coordination environment on a faster timescale. Three species were observed and were assigned to CdS_3O in both the exo and endo conformers as well as CdS_3N with the nitrogen likely originating from a histidine residue at position 72.[149] With the successful generation of a tris-thiolate site in a three-helix bundle, this construct provides an ideal platform to deviate from the three-fold symmetric site and incorporate a tetrahedral thiolate site. Chapter 2 focuses on the design and characterization of tetrathiolate sites while Chapter 3 focuses on biophysical and electron transfer studies of iron in these de novo designed proteins. Chapter 4 details the catalytic conversion of nitrite to NO and water, which involves an expansion of the α_3D framework to include a Type 2 copper site, to enable of both electron transfer and redox catalysis.

References

(1) Chakraborty, S.; Kravitz, J. Y.; Thulstrup, P. W.; Hemmingsen, L.; DeGrado, W. F.; Pecoraro, V. L. Design of a Three-Helix Bundle Capable of Binding Heavy Metals in a Triscysteine Environment. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2049–2053.

(2) Cangelosi, V. M.; Deb, A.; Penner-Hahn, J. E.; Pecoraro, V. L. A De Novo Designed Metalloenzyme for the Hydration of CO2. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7900–7903.

(3) Robson, B.; Vaithilingam, A. Protein Folding Revisited. *Prog Mol Biol Transl Sci* **2008**, *84*, 161–202.

(4) Hill, R. B.; Raleigh, D. P.; Lombardi, A. De Novo Design of Helical Bundles as Models for Understanding Protein Folding and Function. *Acc. Chem. Res.* **2000**, 745–754.

(5) Holm, R. H.; Kennepohl, P.; Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **1996**, *96*, 2239–2314.

(6) Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayyum, H.; Pecoraro, V. L. Protein Design: Toward Functional Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3495–3578.

(7) Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stöhr, J.; Smith, T. A.; Hu, X.; DeGrado, W. F.; Korendovych, I. V. Short Peptides Self-Assemble to Produce Catalytic Amyloids. *Nat Chem* **2014**, *6*, 303–309.

(8) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. Hydrolytic Catalysis and Structural Stabilization in a Designed Metalloprotein. *Nat Chem* **2012**, *4*, 118–123.

(9) Tegoni, M.; Yu, F.; Bersellini, M.; Penner-Hahn, J. E.; Pecoraro, V. L. Designing a Functional Type 2 Copper Center That Has Nitrite Reductase Activity Within A-Helical Coiled Coils. *Proc. Natl. Acad. Sci. U. S. A* **2012**, *109*, 21234–21239.

(10) Yu, F.; Penner-Hahn, J. E.; Pecoraro, V. L. De Novo-Designed Metallopeptides with Type 2 Copper Centers: Modulation of Reduction Potentials and Nitrite Reductase Activities. *J. Am. Chem. Soc.* **2013**, *135*, 18096–18107.

(11) Pennella, M. A.; Giedroc, D. P. Structural Determinants of Metal Selectivity in Prokaryotic Metal-Responsive Transcriptional Regulators. *Biometals* **2005**, *18*, 413–428.

(12) Ma, Z.; Jacobsen, F. E.; Giedroc, D. P. Coordination Chemistry of Bacterial Metal Transport and Sensing. *Chem. Rev.* **2009**, *109*, 4644–4681.

(13) Giedroc, D. P.; Arunkumar, A. I. Metal Sensor Proteins : Nature's Metalloregulated Allosteric Switches. *Dalton Trans.* **2007**, *0*, 3107–3120.

(14) Lu, M.; Fu, D. Structure of the Zinc Transporter YiiP. *Science* **2007**, *317*, 1746–1748.

(15) Argüello, J. M.; Eren, E.; González-Guerrero, M. The Structure and Function of Heavy Metal Transport P1B-ATPases. *Biometals* **2007**, *20*, 233–248.

(16) Dutta, S. J.; Liu, J.; Hou, Z.; Mitra, B. Conserved Aspartic Acid 714 in Transmembrane Segment 8 of the ZntA Subgroup of P1B-Type ATPases Is a Metal-Binding Residue. *Biochemistry* **2006**, *45*, 5923–5931.

(17) González-Guerrero, M.; Argüello, J. M. Mechanism of Cu+-Transporting ATPases: Soluble Cu+ Chaperones Directly Transfer Cu+ to Transmembrane Transport Sites. *Proc. Natl. Acad. Sci. U. S. A* **2008**, *105*, 5992–5997.

(18) Outten, C. E.; O'Halloran, T. V. Femtomolar Sensitivity of Metalloregulatory Proteins Controlling Zinc Homeostasis. *Science* **2001**, *292*, 2488–2492.

(19) Arnesano, F.; Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Molteni, E.; Huffman, D. L.; O'Halloran, T. V. Metallochaperones and Metal-Transporting ATPases: a Comparative Analysis of Sequences and Structures. *Genome Res.* **2002**, *12*, 255–271.

(20) Arnesano, F.; Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S.; Huffman, D. L.; O'Halloran, T. V. Characterization of the Binding Interface Between the Copper Chaperone Atx1 and the First Cytosolic Domain of Ccc2 ATPase. *J. Biol. Chem.* **2001**, *276*, 41365–41376.

(21) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; D'Onofrio, M.; Gonnelli, L.; Marhuenda-Egea, F. C.; Ruiz-Dueñas, F. J. Solution Structure of the N-Terminal Domain of a Potential Copper-Translocating P-Type ATPase From Bacillus Subtilis in the Apo and Cu(I) Loaded States. *J. Mol. Biol.* **2002**, *317*, 415–429.

(22) Walker, J. M.; Huster, D.; Ralle, M.; Morgan, C. T.; Blackburn, N. J.; Lutsenko, S. The N-Terminal Metal-Binding Site 2 of the Wilson's Disease Protein Plays a Key Role in the Transfer of Copper From Atox1. *J. Biol. Chem.* **2004**, *279*, 15376–15384.

(23) Morin, I.; Gudin, S.; Mintz, E.; Cuillel, M. Dissecting the Role of the N-Terminal Metal-Binding Domains in Activating the Yeast Copper ATPase in Vivo. *FEBS J.* **2009**, *276*, 4483–4495.

(24) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Finney, L. A.; Outten, C. E.; O'Halloran, T. V. A New Zinc-Protein Coordination Site in Intracellular Metal Trafficking: Solution Structure of the Apo and Zn(II) Forms of ZntA(46-118). *J. Mol. Biol.* **2002**, *323*, 883–897.

(25) Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Su, X.-C.; Miras, R.; Bal, N.; Mintz, E.; Catty, P.; Shokes, J. E.; Scott, R. A. Structural Basis for Metal Binding Specificity: the N-Terminal Cadmium Binding Domain of the P1-Type ATPase CadA. *J. Mol. Biol.* **2006**, *356*, 638–650.

(26) Liu, J.; Dutta, S. J.; Stemmler, A. J.; Mitra, B. Metal-Binding Affinity of the Transmembrane Site in ZntA: Implications for Metal Selectivity. *Biochemistry* **2006**, *45*, 763–772.

(27) Osman, D.; Cavet, J. S. Bacterial Metal-Sensing Proteins Exemplified by ArsR–SmtB Family Repressors. *Nat. Prod. Rep.* **2010**, *27*, 668–680.

(28) Ye, J.; Kandegedara, A.; Martin, P.; Rosen, B. P. Crystal Structure of the Staphylococcus Aureus pI258 CadC Cd(II)/Pb(II)/Zn(II)-Responsive Repressor. *J. Bacteriol.* **2005**, *187*, 4214–4221.

(29) Arunkumar, A. I.; Campanello, G. C.; Giedroc, D. P. Solution Structure of a Paradigm ArsR Family Zinc Sensor in the DNA-Bound State. *Proc. Natl. Acad. Sci. U. S. A* **2009**, *106*, 18177–18182.

(30) Busenlehner, L. S.; Weng, T.-C.; Penner-Hahn, J. E.; Giedroc, D. P. Elucidation of Primary (Alpha(3)N) and Vestigial (Alpha(5)) Heavy Metal-Binding Sites in Staphylococcus Aureus pI258 CadC: Evolutionary Implications for Metal Ion Selectivity of ArsR/SmtB Metal Sensor Proteins. *J. Mol. Biol.* **2002**, *319*, 685–701.

(31) Busenlehner, L. S.; Giedroc, D. P. Kinetics of Metal Binding by the Toxic Metal-Sensing Transcriptional Repressor Staphylococcus Aureus pI258 CadC. *J. Inorg. Biochem.* **2006**, *100*, 1024–1034.

(32) Kandegedara, A.; Thiyagarajan, S.; Kondapalli, K. C.; Stemmler, T. L.; Rosen, B. P. Role of Bound Zn(II) in the CadC Cd(II)/Pb(II)/Zn(II)-Responsive Repressor. *J. Biol. Chem.* **2009**, *284*, 14958–14965.

(33) Apuy, J. L.; Busenlehner, L. S.; Russell, D. H.; Giedroc, D. P. Ratiometric Pulsed Alkylation Mass Spectrometry as a Probe of Thiolate Reactivity in Different Metalloderivatives of Staphylococcus Aureus pI258 CadC. *Biochemistry* **2004**, *43*, 3824–3834.
(34) Cavet, J. S.; Graham, A. I.; Meng, W.; Robinson, N. J. A Cadmium-Lead-Sensing ArsR-SmtB Repressor with Novel Sensory Sites. *J. Biol. Chem.* **2003**, *278*, 44560–44566.
(35) Wang, Y.; Hemmingsen, L.; Giedroc, D. P. Structural and Functional Characterization of Mycobacterium tuberculosisCmtR, a Pb II/Cd II-Sensing SmtB/ArsR Metalloregulatory Repressor †. *Biochemistry* **2005**, *44*, 8976–8988.
(36) Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S.; Cavet, J. S.; Dennison, C.; Graham, A. I.; Harvie, D. R.; Robinson, N. J. NMR Structural Analysis of Cadmium Sensing by Winged Helix Repressor CmtR. *J. Biol. Chem.* **2007**, *282*, 30181–30188.
(37) Changela, A.; Chen, K.; Xue, Y.; Holschen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragón, A. Molecular Basis of Metal-Ion Selectivity and Zeptomolar Sensitivity by CueR. *Science* **2003**, *301*, 1383–1387.
(38) Ibáñez, M. M.; Checa, S. K.; Soncini, F. C. A Single Serine Residue Determines Selectivity to Monovalent Metal Ions in Metalloregulators of the MerR Family. *J. Bacteriol.* **2015**, *In Press*.
(39) Lee, S. W.; Glickmann, E.; Cooksey, D. A. Chromosomal Locus for Cadmium Resistance in Pseudomonas Putida Consisting of a Cadmium-Transporting ATPase and a MerR Family Response Regulator. *App. Env. Microbiol.* **2001**, *67*, 1437–1444.
(40) Brocklehurst, K. R.; Megit, S. J.; Morby, A. P. Characterisation of CadR From Pseudomonas Aeruginosa: a Cd(II)-Responsive MerR Homologue. *Biochem. Biophys. Res. Comm.* **2003**, *308*, 234–239.
(41) Utschig, L. M.; Bryson, J. W.; O'Halloran, T. V. Mercury-199 NMR of the Metal Receptor Site in MerR and Its Protein-DNA Complex. *Science* **1995**, *268*, 380–385.
(42) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. Electronic Structures of Active Sites in Copper Proteins: Contributions to Reactivity. *Chem. Rev.* **1992**, 521–542.
(43) Orme-Johnson, W. H. Iron-Sulfur Proteins: Structure and Function. *Annu. Rev. Biochem.* **1973**, *42*, 159–204.
(44) Solomon, E. I. Spectroscopic Methods in Bioinorganic Chemistry: Blue to Green to Red Copper Sites. *Inorg. Chem.* **2006**, *45*, 8012–8025.
(45) Solomon, E. I.; Szilagyi, R. K.; DeBeer George, S.; Basumallick, L. Electronic Structures of Metal Sites in Proteins and Models: Contributions to Function in Blue Copper Proteins. *Chem. Rev.* **2004**, *104*, 419–458.
(46) Lancaster, K. M.; George, S. D.; Yokoyama, K.; Richards, J. H.; Gray, H. B. Type-Zero Copper Proteins. *Nat Chem* **2009**, *1*, 711–715.
(47) Hay, M.; Richards, J. H.; Lu, Y. Construction and Characterization of an Azurin Analog for the Purple Copper Site in Cytochrome C Oxidase. *Proc. Natl. Acad. Sci. U. S. A* **1996**, *93*, 461–464.
(48) Beinert, H.; Holm, R. H.; Münck, E. Iron-Sulfur Clusters: Nature's Modular, Multipurpose Structures. *Science* **1997**, *277*, 653–659.
(49) Link, T. A. The Structures of Rieske and Rieske-Type Proteins. *Adv. Inorg. Chem.* **1999**, 83-157.
(50) Venkateswara Rao, P.; Holm, R. H. Synthetic Analogues of the Active Sites of Iron-

Sulfur Proteins. *Chem. Rev.* **2004**, *104*, 527–559.
(51) Krebs, C.; Henshaw, T. F.; Cheek, J.; Huynh, B. H.; Broderick, J. B. Conversion of 3Fe-4S to 4Fe-4S Clusters in Native Pyruvate Formate-Lyase Activating Enzyme: Mössbauer Characterization and Implications for Mechanism. *J. Am. Chem. Soc.* **2000**, *122*, 12497–12506.
(52) C W Carter, J.; Kraut, J.; Freer, S. T.; Alden, R. A.; Sieker, L. C.; Adman, E.; Jensen, L. H. A Comparison of Fe4S4 Clusters in High-Potential Iron Protein and in Ferredoxin. *Proc. Natl. Acad. Sci. U. S. A* **1972**, *69*, 3526–3529.
(53) Dey, A.; Jenney, F. E.; Adams, M. W. W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP Versus Ferredoxin. *Science* **2007**, *318*, 1464–1468.
(54) Lovenberg, W.; Sobel, B. E. Rubredoxin: a New Electron Transfer Protein From Clostridium Pasteurianum. *Proc. Natl. Acad. Sci. U. S. A* **1965**, *54*, 193–199.
(55) LeGall, J.; Prickril, B. C.; Moura, I.; Xavier, A. V.; Moura, J. J.; Huynh, B. H. Isolation and Characterization of Rubrerythrin, a Non-Heme Iron Protein From Desulfovibrio Vulgaris That Contains Rubredoxin Centers and a Hemerythrin-Like Binuclear Iron Cluster. *Biochemistry* **1988**, *27*, 1636–1642.
(56) Jin, S.; Kurtz, D. M.; Liu, Z.-J.; Rose, J.; Wang, B.-C. X-Ray Crystal Structure of Desulfovibrio Vulgaris Rubrerythrin with Zinc Substituted Into the [Fe(SCys)4] Site and Alternative Diiron Site Structures. *Biochemistry* **2004**, *43*, 3204–3213.
(57) Luo, Y.; Ergenekan, C. E.; Fischer, J. T.; Tan, M.-L.; Ichiye, T. The Molecular Determinants of the Increased Reduction Potential of the Rubredoxin Domain of Rubrerythrin Relative to Rubredoxin. *Biophys. J.* **2010**, *98*, 560–568.
(58) Gomes, C. M.; Giuffrè, A.; Forte, E.; Vicente, J. B.; Saraiva, L. M.; Brunori, M.; Teixeira, M. A Novel Type of Nitric-Oxide Reductase. Escherichia Coli Flavorubredoxin. *J. Biol. Chem.* **2002**, *277*, 25273–25276.
(59) Gardner, A. M.; Helmick, R. A.; Gardner, P. R. Flavorubredoxin, an Inducible Catalyst for Nitric Oxide Reduction and Detoxification in Escherichia Coli. *J. Biol. Chem.* **2002**, *277*, 8172–8177.
(60) Gomes, C. M.; Vicente, J. B.; Wasserfallen, A.; Teixeira, M. Spectroscopic Studies and Characterization of a Novel Electron-Transfer Chain From Escherichia Coli Involving a Flavorubredoxin and Its Flavoprotein Reductase Partner. *Biochemistry* **2000**, *39*, 16230–16237.
(61) Pierik, A. J.; Wolbert, R. B.; Portier, G. L.; Verhagen, M. F.; Hagen, W. R. Nigerythrin and Rubrerythrin From Desulfovibrio Vulgaris Each Contain Two Mononuclear Iron Centers and Two Dinuclear Iron Clusters. *Eur. J. Biochem.* **1993**, *212*, 237–245.
(62) Lumppio, H. L.; Shenvi, N. V.; Garg, R. P.; Summers, A. O.; Kurtz, D. M. A Rubrerythrin Operon and Nigerythrin Gene in Desulfovibrio Vulgaris (Hildenborough). *J. Bacteriol.* **1997**, *179*, 4607–4615.
(63) Iyer, R. B.; Silaghi-Dumitrescu, R.; Kurtz, D. M.; Lanzilotta, W. N. High-Resolution Crystal Structures of Desulfovibrio Vulgaris (Hildenborough) Nigerythrin: Facile, Redox-Dependent Iron Movement, Domain Interface Variability, and Peroxidase Activity in the Rubrerythrins. *J. Biol. Inorg. Chem.* **2005**, *10*, 407–416.
(64) Moura, I.; Xavier, A. V.; Cammack, R.; Bruschi, M.; Le Gall, J. A Comparative Spectroscopic Study of Two Non-Haem Iron Proteins Lacking Labile Sulphide From

Desulphovibrio Gigas. *Biochim. Biophys. Acta* **1978**, *533*, 156–162.
(65) Jenney, F., Jr. Rubredoxin From Pyrococcus Furiosus. *Meth. Enzymol.* **2001**, 45-55.
(66) Bönisch, H.; Schmidt, C. L.; Bianco, P.; Ladenstein, R. Ultrahigh-Resolution Study on Pyrococcus Abyssi Rubredoxin: II. Introduction of an O–H···Sγ–Fe Hydrogen Bond Increased the Reduction Potential by 65 mV. *J. Biol. Inorg. Chem.* **2007**, *12*, 1163–1171.
(67) Moura, I.; Bruschi, M.; Le Gall, J.; Moura, J. J. G.; Xavier, A. V. Isolation and Characterization of Desulforedoxin, a New Type of Non-Heme Iron Protein From Desulfovibrio Gigas. *Biochem. Biophys. Res. Comm.* **1977**, *75*, 1037–1044.
(68) Moura, I.; Huynh, B. H.; Hausinger, R. P.; Le Gall, J.; Xavier, A. V.; Münck, E. Mössbauer and EPR Studies of Desulforedoxin From Desulfovibrio Gigas. *J. Biol. Chem.* **1980**, *255*, 2493–2498.
(69) Stephens, P. J.; Jollie, D. R.; Warshel, A. Protein Control of Redox Potentials of Iron−Sulfur Proteins. *Chem. Rev.* **1996**, *96*, 2491–2514.
(70) Atherton, N.; Garbett, K.; Gillard, R.; Mason, R. Spectroscopic Investigation of Rubredoxin and Ferredoxin. *Nature* **1966**, 590-593.
(71) Peisach, J.; Blumberg, W.; Lode, E. An Analysis of the Electron Paramagnetic Resonance Spectrum of Pseudomonas Oleovorans Rubredoxin. *J. Biol. Chem.* **1971**, 5877-5881.
(72) Eaton, W. A.; Lovenberg, W. Near-Infrared Circular Dichroism of an Iron-Sulfur Protein. D Leads to D Transitions in Rubredoxin. *J. Am. Chem. Soc.* **1970**, *92*, 7195–7198.
(73) Bunker, B.; Stern, E. A. The Iron-Sulfur Environment in Rubredoxin. *Biophys. J.* **1977**, *19*, 253–264.
(74) Wegner, P.; Bever, M.; Schünemann, V.; Trautwein, A. X.; Schmidt, C.; Bönisch, H.; Gnida, M.; Meyer-Klaucke, W. Iron–Sulfur Proteins Investigated by EPR-, Mössbauer-and EXAFS-Spectroscopy. *Hyperfine Interact.* **2004**, *156*, 293–298.
(75) Xiao, Z.; Lavery, M. J.; Ayhan, M.; Scrofani, S. D. B.; Wilce, M. C. J.; Guss, J. M.; Tregloan, P. A.; George, G. N.; Wedd, A. G. The Rubredoxin From Clostridium Pasteurianum: Mutation of the Iron Cysteinyl Ligands to Serine. Crystal and Molecular Structures of Oxidized and Dithionite-Treated Forms of the Cys42Ser Mutant. *J. Am. Chem. Soc.* **1998**, *120*, 4135–4150.
(76) Yoo, S. J.; Meyer, J.; Achim, C.; Peterson, J.; Hendrich, M. P.; Münck, E. Mössbauer, EPR, and MCD Studies of the C9S and C42S Variants of Clostridium Pasteurianum Rubredoxin and MDC Studies of the Wild-Type Protein. *J Biol Inorg Chem* **2000**, *5*, 475–487.
(77) Perrin, B. S., Jr.; Ichiye, T. Fold Versus Sequence Effects on the Driving Force for Protein-Mediated Electron Transfer. *Proteins* **2010**, *78*, 2798–2808.
(78) Perrin, B. S.; Ichiye, T. Characterizing the Effects of the Protein Environment on the Reduction Potentials of Metalloproteins. *J Biol Inorg Chem* **2013**, *18*, 103–110.
(79) Xiao, Z.; Maher, M. J.; Cross, M.; Bond, C. S.; Guss, J. M.; Wedd, A. G. Mutation of the Surface Valine Residues 8 and 44 in the Rubredoxin From Clostridium Pasteurianum: Solvent Access Versus Structural Changes as Determinants of Reversible Potential. *J. Biol. Inorg. Chem.* **2000**, *5*, 75–84.
(80) Lin, I.-J.; Gebel, E. B.; Machonkin, T. E.; Westler, W. M.; Markley, J. L. Changes in Hydrogen-Bond Strengths Explain Reduction Potentials in 10 Rubredoxin Variants.

Proc. Natl. Acad. Sci. U. S. A **2005**, *102*, 14581–14586.
(81) Kendra Rose; Susan E Shadle; Marly K Eidsness; Donald M Kurtz, Jr; Robert A Scott; Britt Hedman; Keith O Hodgson, A.; Edward I Solomon. Investigation of Iron−Sulfur Covalency in Rubredoxins and a Model System Using Sulfur K-Edge X-Ray Absorption Spectroscopy. *J. Am. Chem. Soc.* **1998**, 10743–10747.
(82) Dolan, E. A.; Yelle, R. B.; Beck, B. W.; Fischer, J. T.; Ichiye, T. Protein Control of Electron Transfer Rates via Polarization: Molecular Dynamics Studies of Rubredoxin. *Biophys. J.* **2004**, *86*, 2030–2036.
(83) Maher, M. J.; Xiao, Z.; Wilce, M. C.; Guss, J. M.; Wedd, A. G. Rubredoxin From Clostridium Pasteurianum. Structures of G10A, G43A and G10VG43A Mutant Proteins. Mutation of Conserved Glycine 10 to Valine Causes the 9-10 Peptide Link to Invert. *Acta Crystallogr. D Biol. Crystallogr.* **1999**, *55*, 962–968.
(84) Min, T.; Ergenekan, C. E.; Eidsness, M. K.; Ichiye, T.; Kang, C. Leucine 41 Is a Gate for Water Entry in the Reduction of Clostridium Pasteurianum Rubredoxin. *Protein Sci.* **2001**, *10*, 613–621.
(85) Eidsness, M. K.; Burden, A. E.; Richie, K. A.; Kurtz, D. M.; Scott, R. A.; Smith, E. T.; Ichiye, T.; Beard, B.; Min, T.; Kang, C. Modulation of the Redox Potential of the [Fe(SCys)(4)] Site in Rubredoxin by the Orientation of a Peptide Dipole. *Biochemistry* **1999**, *38*, 14803–14809.
(86) Swartz, P.; Beck, B. Structural Origins of Redox Potentials in Fe-S Proteins: Electrostatic Potentials of Crystal Structures. *Biophys. J.* **1996**, 2958–2969.
(87) Perrin, B. S.; Niu, S.; Ichiye, T. Calculating Standard Reduction Potentials of [4Fe-4S] Proteins. *J Comput Chem* **2013**, *34*, 576–582.
(88) deMaré, F.; Kurtz, D. M.; Nordlund, P. The Structure of Desulfovibrio Vulgaris Rubrerythrin Reveals a Unique Combination of Rubredoxin-Like FeS4 and Ferritin-Like Diiron Domains. *Nat. Struct. Mol. Biol.* **1996**, 539–546.
(89) Moser, C. C.; Chobot, S. E.; Page, C. C.; Dutton, P. L. Distance Metrics for Heme Protein Electron Tunneling. *Biochim. Biophys. Acta* **2008**, *1777*, 1032–1037.
(90) Giese, B.; Graber, M.; Cordes, M. Electron Transfer in Peptides and Proteins. *Curr. Op. Chem. Biol.* **2008**, *12*, 755–759.
(91) Ener, M. E.; Lee, Y.-T.; Winkler, J. R.; Gray, H. B.; Cheruzel, L. Photooxidation of Cytochrome P450-BM3. *Proc. Natl. Acad. Sci. U. S. A* **2010**, *107*, 18783–18786.
(92) Sasaki, T.; Kaiser, E. T. Helichrome: Synthesis and Enzymic Activity of a Designed Hemeprotein. *J. Am. Chem. Soc.* **1989**, *111*, 380–381.
(93) Choma, C. T.; Lear, J. D.; Nelson, M. J. Design of a Heme-Binding Four-Helix Bundle. *J. Am. Chem. Soc.* **1994**, 856–865.
(94) Shifman, J. M.; Moser, C. C.; Kalsbeck, W. A.; Bocian, D. F.; Dutton, P. L. Functionalized De NovoDesigned Proteins: Mechanism of Proton Coupling to Oxidation/Reduction in Heme Protein Maquettes †. *Biochemistry* **1998**, *37*, 16815–16827.
(95) Chen, X.; Moser, C. C.; Pilloud, D. L.; Dutton, P. L. Molecular Orientation of Langmuir−Blodgett Films of Designed Heme Protein and Lipoprotein Maquettes. *J Phys Chem B* **1998**, *102*, 6425–6432.
(96) Pilloud, D. L.; Rabanal, F.; Gibney, B. R. Self-Assembled Monolayers of Synthetic Hemoproteins on Silanized Quartz. *J. Am. Chem. Soc.* **1998**, 1926–1937.
(97) Chen, X.; Moser, C. C.; Pilloud, D. L. Engineering Oriented Heme Protein Maquette

Monolayers Through Surface Residue Charge Distribution Patterns. *J. Am. Chem. Soc.* **1999**, 9029–9037.

(98) Chen, X.; Discher, B. M.; Pilloud, D. L. De Novo Design of a Cytochrome B Maquette for Electron Transfer and Coupled Reactions on Electrodes. *J. Am. Chem. Soc.* **2002**, 617–624.

(99) Ye, S.; Strzalka, J. W.; Discher, B. M.; Noy, D.; Zheng, S.; Dutton, P. L.; Blasie, J. K. Amphiphilic 4-Helix Bundles Designed for Biomolecular Materials Applications. *Langmuir* **2004**, *20*, 5897–5904.

(100) Topoglidis, E.; Discher, B. M.; Moser, C. C.; Dutton, P. L.; Durrant, J. R. Functionalizing Nanocrystalline Metal Oxide Electrodes with Robust Synthetic Redox Proteins. *ChemBioChem* **2003**, *4*, 1332–1339.

(101) Ye, S.; Discher, B. M.; Strzalka, J.; Xu, T.; Wu, S. P.; Noy, D.; Kuzmenko, I.; Gog, T.; Therien, M. J.; Dutton, P. L.; et al. Amphiphilic Four-Helix Bundle Peptides Designed for Light-Induced Electron Transfer Across a Soft Interface. *Nano Lett.* **2005**, *5*, 1658–1667.

(102) Gibney, B. R.; Isogai, Y.; Rabanal, F.; Reddy, K. S. Self-Assembly of Heme a and Heme B in a Designed Four-Helix Bundle: Implications for a Cytochrome C Oxidase Maquette. *Biochemistry* **2000**, 11041–11049.

(103) Ghirlanda, G.; Osyczka, A.; Liu, W.; Antolovich, M.; Smith, K. M.; Dutton, P. L.; Wand, A. J.; DeGrado, W. F. De Novo Design of a D2-Symmetrical Protein That Reproduces the Diheme Four-Helix Bundle in Cytochrome Bc1. *J. Am. Chem. Soc.* **2004**, *126*, 8141–8147.

(104) Korendovych, I. V.; Senes, A.; Kim, Y. H.; Lear, J. D.; Fry, H. C.; Therien, M. J.; Blasie, J. K.; Walker, F. A.; DeGrado, W. F. De Novo Design and Molecular Assembly of a Transmembrane Diporphyrin-Binding Protein Complex. *J. Am. Chem. Soc.* **2010**, *132*, 15516–15518.

(105) Anderson, J. L. R.; Armstrong, C. T.; Kodali, G.; Lichtenstein, B. R.; Watkins, D. W.; Mancini, J. A.; Boyle, A. L.; Farid, T. A.; Crump, M. P.; Moser, C. C.; et al. Constructing a Man-Made C-Type Cytochrome Maquette in Vivo: Electron Transfer, Oxygen Transport and Conversion to a Photoactive Light Harvesting Maquette. *Chem. Sci.* **2014**, *5*, 507–514.

(106) Koder, R. L.; Anderson, J. L. R.; Solomon, L. A.; Reddy, K. S.; Moser, C. C.; Dutton, P. L. Design and Engineering of an O2 Transport Protein. *Nature* **2009**, *458*, 305–309.

(107) Farid, T. A.; Kodali, G.; Solomon, L. A.; Lichtenstein, B. R.; Sheehan, M. M.; Fry, B. A.; Bialas, C.; Ennist, N. M.; Siedlecki, J. A.; Zhao, Z.; et al. Elementary Tetrahelical Protein Design for Diverse Oxidoreductase Functions. *Nat. Chem. Biol.* **2013**, *9*, 826–833.

(108) Bertini, I.; Gray, H. B.; Stiefel, E. I.; Valentine, J. S. *Biological Inorganic Chemistry*; University Science Books: Sausalito, CA.

(109) Marcus, R. A.; Sutin, N. Electron Transfers in Chemistry and Biology. *Biochim. Biophys. Acta* **1985**, *811*, 265–322.

(110) Marcus, R. A. Electron Transfer Reactions in Chemistry: Theory and Experiment (Nobel Lecture). *Angew. Chem. Int. Ed. Engl.* **1993**, 69–92.

(111) Andrews, D. L. Biological Energy 5: Electron transfer theory http://biologicalphysics.iop.org/cws/article/lectures/53592 (accessed Apr 9, 2015).

(112) Moser, C. C.; Keske, J. M.; Warncke, K.; Farid, R. S.; Dutton, P. L. Nature of

Biological Electron Transfer. *Nature* **1992**, *355*, 796–802.
(113) Karpishin, T. B.; Grinstaff, M. W.; Komar-Panicucci, S.; McLendon, G.; Gray, H. B. Electron Transfer in Cytochrome C Depends Upon the Structure of the Intervening Medium. *Structure* **1994**, *2*, 415–422.
(114) Gray, H. B.; Winkler, J. R. Long-Range Electron Transfer. *Proc. Natl. Acad. Sci. U. S. A* **2005**, *102*, 3534–3539.
(115) Gray, H. B.; Winkler, J. R. Electron Flow Through Proteins. *Chem. Phys. Lett.* **2009**, *483*, 1–9.
(116) Gray, H. B.; Winkler, J. R. Electron Tunneling Through Proteins. *Quart. Rev. Biophys.* **1999**, *36*, 341–372.
(117) Page, C. C.; Moser, C. C.; Chen, X.; Dutton, P. L. Natural Engineering Principles of Electron Tunnelling in Biological Oxidation-Reduction. *Nature* **1999**, *402*, 47–52.
(118) Gray, H. B.; Malmström, B. G.; Williams, R. Copper Coordination in Blue Proteins. *J. Biol. Inorg. Chem.* **2000**.
(119) Migliore, A.; Polizzi, N. F.; Therien, M. J.; Beratan, D. N. *Chem. Rev.* **2014**, *114*, 3381.
(120) Gunner, M. R.; Dutton, P. L. Temperature and -dG Dependence of the Electron Transfer From BPh^{-} To QA in Reaction Center Protein From Rhodobacter Sphaeroides with Different Quinones as Q_A. *J. Am. Chem. Soc.* **1989**, 3400–3412.
(121) Giangiacomo, K. M.; Dutton, P. L. In Photosynthetic Reaction Centers, the Free Energy Difference for Electron Transfer Between Quinones Bound at the Primary and Secondary Quinone-Binding Sites Governs the Observed Secondary Site Specificity. *Proc. Natl. Acad. Sci. U. S. A* **1989**, *86*, 2658–2662.
(122) Boxer, S. G. Mechanisms of Long-Distance Electron Transfer in Proteins: Lessons From Photosynthetic Reaction Centers. *Annu Rev Biophys Biophys Chem* **1990**, *19*, 267–299.
(123) Beratan, D. N.; Onuchic, J. N.; Hopfield, J. J. Electron Tunneling Through Covalent and Noncovalent Pathways in Proteins. *J Chem Phys* **1987**, *86*, 4488–4498.
(124) Beratan, D. N.; Betts, J. N.; Onuchic, J. N. Protein Electron Transfer Rates Set by the Bridging Secondary and Tertiary Structure. *Science* **1991**, *252*, 1285–1288.
(125) Langen, R.; Chang, J.; Germanas, J. P.; Richards, J. H. Electron Tunneling in Proteins: Coupling Through a SS Strand. *Science* **1995**, 1733–1735.
(126) Regan, J. J.; Di Bilio, A. J.; Langen, R.; Skov, L. K.; Winkler, J. R.; Gray, H. B.; Onuchic, J. N. Electron Tunneling in Azurin: the Coupling Across a Beta-Sheet. *Chem. Biol.* **1995**, *2*, 489–496.
(127) Prytkova, T. R.; Kurnikov, I. V.; Beratan, D. N. Coupling Coherence Distinguishes Structure Sensitivity in Protein Electron Transfer. *Science* **2007**, *315*, 622–625.
(128) Jean Alric; Jérôme Lavergne; Fabrice Rappaport; André Verméglio; Katsumi Matsuura; Keizo Shimada, A.; Kenji V P Nagashima. Kinetic Performance and Energy Profile in a Roller Coaster Electron Transfer Chain: a Study of Modified Tetraheme-Reaction Center Constructs. *J. Am. Chem. Soc.* **2006**, 4136–4145.
(129) Moser, C. C.; Anderson, J. L. R.; Dutton, P. L. Guidelines for Tunneling in Enzymes. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **2010**, *1797*, 1573–1586.
(130) Moura, I.; Moura, J. J. Structural Aspects of Denitrifying Enzymes. *Curr. Op. Chem. Biol.* **2001**, *5*, 168–175.
(131) Libby, E.; Averill, B. A. Evidence That the Type 2 Copper Centers Are the Site of

Nitrite Reduction by Achromobacter Cycloclastes Nitrite Reductase. *Biochem. Biophys. Res. Comm.* **1992**, *187*, 1529–1535.
(132) Tocheva, E. I.; Rosell, F. I.; Mauk, A. G.; Murphy, M. E. P. Side-on Copper-Nitrosyl Coordination by Nitrite Reductase. *Science* **2004**, *304*, 867–870.
(133) Ferroni, F. M.; Guerrero, S. A.; Rizzi, A. C.; Brondino, C. D. Overexpression, Purification, and Biochemical and Spectroscopic Characterization of Copper-Containing Nitrite Reductase From Sinorhizobium Meliloti 2011. Study of the Interaction of the Catalytic Copper Center with Nitrite and NO. *J. Inorg. Biochem.* **2012**, *114*, 8–14.
(134) Kline, C. D.; Mayfield, M.; Blackburn, N. J. HHM Motif at the CuH-Site of Peptidylglycine Monooxygenase Is a pH-Dependent Conformational Switch. *Biochemistry* **2013**, *52*, 2586–2596.
(135) Chufán, E. E.; Prigge, S. T.; Siebert, X.; Eipper, B. A.; Mains, R. E.; Amzel, L. M. Differential Reactivity Between Two Copper Sites in Peptidylglycine A-Hydroxylating Monooxygenase. *J. Am. Chem. Soc.* **2010**, *132*, 15565–15572.
(136) Itoh, S. Mononuclear Copper Active-Oxygen Complexes. *Curr. Op. Chem. Biol.* **2006**, *10*, 115–122.
(137) Tebo, A. G.; Pecoraro, V. L. Artificial Metalloenzymes Derived From Three-Helix Bundles. *Curr. Op. Chem. Biol.* **2015**, *25C*, 65–70.
(138) Dieckmann, G. R.; McRorie, D. K.; Tierney, D. L.; Utschig, L. M.; Singer, C. P.; O'Halloran, T. V.; Penner-Hahn, J. E.; DeGrado, W. F.; Pecoraro, V. L. De Novo Design of Mercury-Binding Two-and Three-Helical Bundles. *J. Am. Chem. Soc.* **1997**.
(139) Dieckmann, G. R.; McRorie, D. K.; Lear, J. D.; Sharp, K. A.; DeGrado, W. F.; Pecoraro, V. L. The Role of Protonation and Metal Chelation Preferences in Defining the Properties of Mercury-Binding Coiled Coils. *J. Mol. Biol.* **1998**, *280*, 897–912.
(140) Ghosh, D.; Pecoraro, V. L. Understanding Metalloprotein Folding Using a De Novo Design Strategy. *Inorg. Chem.* **2004**, *43*, 7902–7915.
(141) Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. Peptidic Models for the Binding of Pb(II), Bi(III) and Cd(II) to Mononuclear Thiolate Binding Sites. *J. Biol. Inorg. Chem.* **2006**, *11*, 876–890.
(142) Peacock, A. F. A.; Iranzo, O.; Pecoraro, V. L. Harnessing Natures Ability to Control Metal Ion Coordination Geometry Using De Novo Designed Peptides. *Dalton Trans* **2009**, 2271–2280.
(143) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. The Correlation of 113Cd NMR and 111mCd PAC Spectroscopies Provides a Powerful Approach for the Characterization of the Structure of Cd II-Substituted Zn IIProteins. *Chem. Eur. J.* **2009**, *15*, 3761–3772.
(144) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. Switching the Chirality of the Metal Environment Alters the Coordination Mode in Designed Peptides. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 7371–7374.
(145) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. Using Diastereopeptides to Control Metal Ion Coordination in Proteins. *Proc. Natl. Acad. Sci. U. S. A* **2008**, *105*, 16566–16571.
(146) Iranzo, O.; Cabello, C.; Pecoraro, V. L. Heterochromia in Designed Metallopeptides: Geometry-Selective Binding of CdII in a De Novo Peptide. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6688–6691.

(147) Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. From Coiled Coils to Small Globular Proteins: Design of a Native-Like Three-Helix Bundle. *Protein Sci.* **1998**, *7*, 1404–1414.
(148) Walsh, S. T.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. Solution Structure and Dynamics of a De Novo Designed Three-Helix Bundle Protein. *Proc. Natl. Acad. Sci. U. S. A* **1999**, *96*, 5486–5491.
(149) Chakraborty, S. Designed Metalloproteins: From Structurally Characterized Scaffolds to Helical Bundles. Ph.D. Thesis. University of Michigan. Ann Arbor, MI. Dec. 2010.

Chapter 2. Modeling Cd(II) homeostasis in *de novo* designed proteins

Overview

There are many examples of metals sites with greater than three cysteines. Metals with cysteine in the primary coordination sphere figure prominently in redox reactions as well as in metal homeostasis. The first goal of my thesis is to expand our laboratory's previous work on tris-thiolate environments to sites with greater than three cysteines to help understand how heavy metal homeostasis functions, as well as to explore the principles behind the design of tetrathiolate sites in single-stranded constructs. Given the experience that our laboratory has in characterizing Cd(II) coordination in thiolate-rich sites, I characterized my designs using techniques specific for Cd(II), in particular, ^{111m}Cd perturbed angular correlation (PAC) spectroscopy and ^{113}Cd NMR. These systems where a fourth cysteine was incorporated provide models for prokaryotic heavy metal sensors that use metal binding to thiol-rich environments to regulate the expression of detoxification systems. In this introduction, I will provide some relevant background for these metal-sensing systems and the open questions that remain, particularly regarding Cd(II) coordination spheres, as well as some background on Cd(II) spectroscopy before reviewing some work that has previously been done on designed Cd(II)-binding proteins. This work is still unpublished. I would like to acknowledge my collaborator, Dr. Lars Hemingsen at the University of Copenhagen for the ^{111m}Cd PAC measurements and fitting.

Designed Proteins and Cd(II) Spectroscopy

Few members of the ArsR/SmtB and MerR families have been extensively characterized and lack structural studies to help resolve questions about unusual, intermediate, or distorted coordination spheres. Therefore, our group has turned to protein design methods and specialized spectroscopic techniques to help understand and define the relevant coordination chemistry for toxic heavy metals. Many of the suspected Cd(II)-binding sensors utilize 3-4 cysteines to

generate CdS_3, CdS_3O, or distorted CdS_4 sites. Also, because DNA-binding regulators often have α-helices that donate metal-binding ligands, designed peptides such as **TRI** and α_3D can be excellent first approximation models. Thus, the **TRI** scaffold and its inherent three-fold symmetry was used as a tool to help define the spectroscopy of Cd(II) coordination with three cysteines in the primary coordination sphere, which is the minimal cysteine coordination for CadC and related sensors. Furthermore, the heptad repeat structure used in designed α-helices places potential substitutions either three or four residues away from the next hydrophobic residue, creating the motifs, CXXL or CXXXL. Substitution in α_3D, which is inherently asymmetric, could then easily lead to a controlled number CXXC (or potentially CXXXC) motifs as are found in most native Cd(II)-binding proteins.

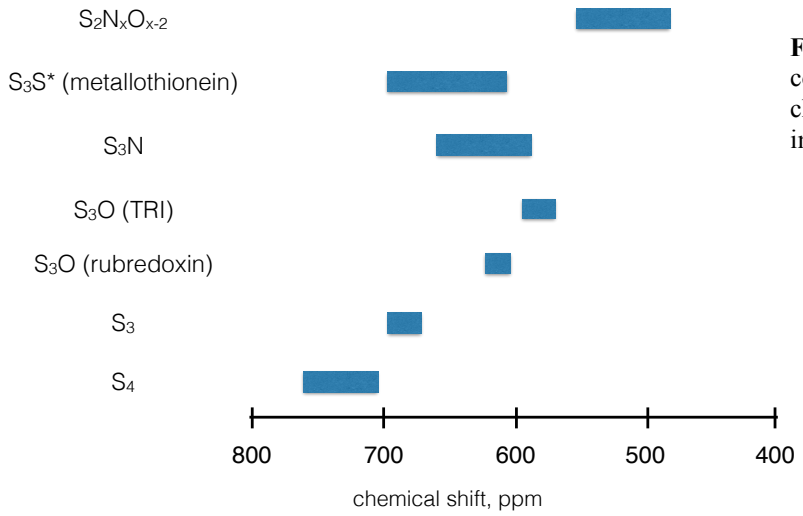


Figure 2.1 Correlation between coordination sphere and ^{113}Cd NMR chemical shift for cysteine-rich sites in proteins. Adapted from ref2 2.

Cadmium, lead, and mercury interactions with cysteine residues are characterized by high energy LMCT bands in the UV region. In particular, the number of thiolates coordinated to cysteine has been tied to a linear increase in the extinction coefficient, allowing for an assessment of cysteine binding stoichiometry from the UV spectrum alone.[1] These metals have other available spectroscopies to examine the primary coordination sphere that are unique on a metal-by-metal basis. In particular, Cd(II) binding can be characterized in some detail by using ^{113}Cd NMR and ^{111m}Cd perturbed angular correlation spectroscopy (PAC). Both of these methods give information as to the electrostatic environment around the bound metal. They have been used to as fingerprinting tools to help predict the likely coordination sphere of Cd(II) in both proteins and small molecules, and their use together is particularly powerful.

Both techniques belong to the same general class of hyperfine spectroscopic methods as EPR and Mössbauer spectroscopies. ^{113}Cd is extremely sensitive to the nature, number, and geometry of its ligands; at natural abundance, ^{113}Cd has a sensitivity about 8-fold that of naturally abundant ^{13}C.[2] The chemical shift for ^{113}Cd NMR spans more than 900 ppm, which results from the sensitivity of the nucleus to the deshielding effect of its ligands. The deshielding effect increases on the order of S>N>O, so proteins with greater numbers of cysteines coordinating the Cd(II) have larger chemical shifts.[2,3]Unfortunately, this high level of sensitivity can also lead to chemical exchange broadening, especially in the case of a metalloenzyme with an open coordination site where solvent or ligand may exchange, but can also include metal binding processes and protein conformational changes. Some care should be exercised to ensure that conditions are equivalent between those data being compared because common ions such as chloride can exert a deshielding effect on the nucleus and result in a downfield shift of the signal.[2] By extensive study of many proteins and systems several guidelines have been established with respect to Cd(II)-thiolate coordination: 1) there is a shift of ~200 ppm per thiol, 2) terminal cysteines are more extensively deshielded than bridging cysteines, and 3) cysteine is more deshielded than methionine by ~100 ppm. Variability in the signal is then introduced by mixed coordination spheres, geometry, coordination number, and the charge on ligands. Thus, sites with four coordinating cysteines have been found to have chemical shifts ranging from 600 ppm (including metallothioneins) to higher than 700 ppm with the latter being most common (Figure 2.1).

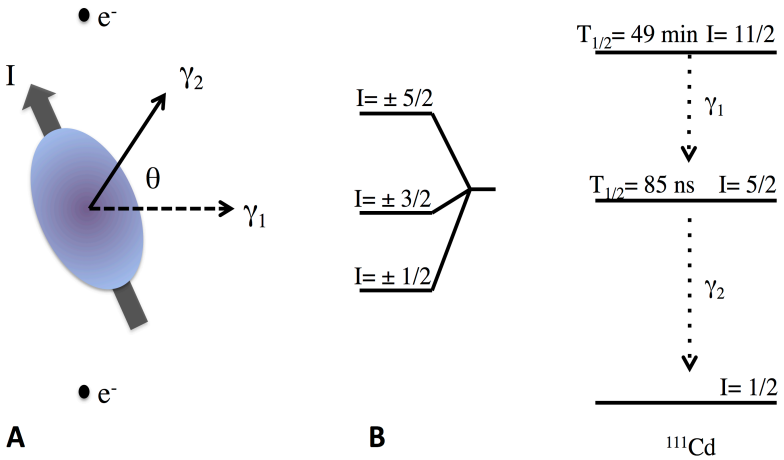


Figure 2.2 A) Graphical depiction of nucleus with spin **I** oriented with respect to two point charges, e^-. The origin of the nuclear quadrupole interaction (NQI) is the interaction of the positive nuclear charge distribution and the surrounding charge distribution (here, two point charges). Emission of the second gamma ray is related to the direction of emission of the first by angle θ. B) ^{111m}Cd decays by successive gamma ray emission and the hyperfine splitting of the intermediate energy level caused by the NQI is measured. Adapted from ref.4.

This fingerprint of the Cd(II) coordination can be combined with perturbed angular correlation (PAC) spectroscopy to develop a solid understanding of the site. PAC, as applied to bioinorganic chemistry, concerns itself with the interaction of the nuclear charge distribution and the surrounding charge distribution, called the NQI. It relies on the concept that particles that are emitted as a result of nuclear decay are not emitted randomly with respect to each other due to the conservation of angular momentum during the nuclear decay (Figure 2.2). In the case of ^{111m}Cd, the second gamma ray is usually emitted either in the same direction or exactly opposite to the first gamma ray, while emission at an angle of 90 degrees is least favored.[4] This would be entirely uninteresting to chemists if it weren't for the fact that the emission of this second gamma ray is perturbed by extranuclear fields. To measure this interaction, or NQI, the nucleus must be non-spherical, as measured by the nuclear electric quadrupole moment. Additionally, the electric field gradient (EFG) tensor at the position of the nucleus must also not be spherical, reflecting that the surrounding charge distribution is not spherical. The NQI leads to a splitting of the energy sub levels in nuclei with nuclear spin I >1; the interaction of the NQI and this I >1 state (Figure 2.2) results in a perturbation function that describes the periodic fluctuation over time of the probability of detecting the second gamma ray at a particular angle. PAC measures differences in this sub level splitting by fitting a Fourier transformed perturbation function to the

measured periodic fluctuation and uses it to fingerprint the local structure of the site. PAC is also capable of sensing differences in the NQI that happen on the same order of the lifetime of the intermediate state, which for ^{111m}Cd is 85 ns.[4]

The Pecoraro lab has studied Cd(II) coordination in *de novo* designed peptides using, in addition to other techniques, ^{111m}Cd PAC and ^{113}Cd NMR, and has generated a series of peptides that can act as 'standards' for various coordination geometries and stoichiometries. Our group has been able to control the coordination sphere—both geometry and coordination number—for Cd(II) and Hg(II) within the hydrophobic interior of peptides. Additionally, our group has demonstrated the ability to bind Cd(II) in a site-selective manner based on ion recognition.[5-7] However, most of these studies have been carried out in either the **TRI** scaffold or a longer (**GRAND**) or shorter (**BABY**) derivative. More recently, our group has sought to move this chemistry into a single-stranded scaffold, such as α_3D, which adopts more of a globular fold like that of a native protein. In part this has been to understand how the spectroscopy might change in a more native-like peptide, but also because a single-stranded construct is capable of enforcing asymmetry, which is essential to understanding how most native proteins function.

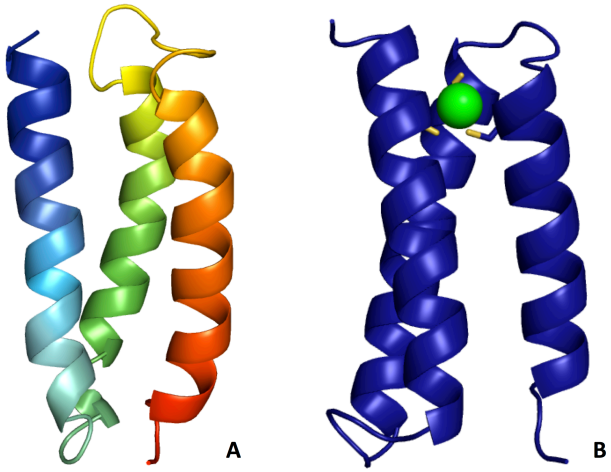


Figure 2.3 A) PyMol representation of α_3D (PDB: 2A3D) B) PyMol representation of α_3DIV based on apo-structure of α_3D**IV** (PDB: 2MTQ)

A former graduate student, Dr. Saumen Chakraborty, worked on adapting DeGrado's α_3D scaffold to bind metals by introducing cysteine residues, analogous to **TRI** (Figure 2.3). α_3D was chosen as a scaffold in part because its solution NMR structure had been solved and it had been proven to be highly stable and tolerant to mutations in the interior of the bundle.[8,9] In

collaboration with Dr. Joslyn Kravitz, four potential locations for a metal binding site were computationally modeled to assess the tolerance of the protein for introducing a metal site.[10,11] The result was α_3D**IV**, the first single-stranded antiparallel metal-binding designed peptide, which replaces three leucine residues in the interior of the three helix bundle with cysteine residues. The NMR structure of α_3D showed that the selected location had a well-ordered backbone and the cysteine binding site was largely sequestered from solvent. In fact, the metal binding site occupies a "box" in which the sides are formed by the helices themselves and the top and bottom defined by nearby hydrophobic residues: Phe31, Ile14, and Ile63 below the site and Tyr70 and Leu21 above the site. Despite the introduction of the metal binding site, α_3DIV was still stable and showed a ΔG of unfolding of 2.5 kcal/mol from denaturation studies with guanidine hydrochloride. This construct was characterized with respect to its heavy metal binding abilities and shown to bind Pb(II), Hg(II), and Cd(II) in a similar fashion to our previously published systems. Cd(II) was titrated directly into the protein, which yielded a λ_{max} of 232 nm and molar extinction coefficient of 18,200 $M^{-1}cm^{-1}$, in the presence of TCEP. These UV-vis characteristics are consistent with tris-thiolate coordination of Cd(II). Moreover, these titrations, which showed tight binding, were fitted to yield a K_a of $2.0x10^7$ M^{-1}. The pH

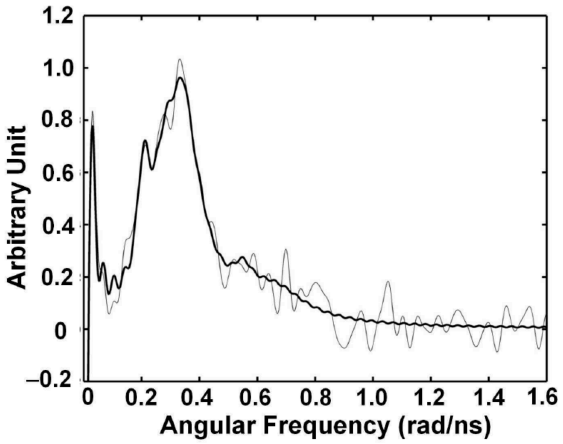


Figure 2.4 ^{111m}Cd PAC spectroscopy on α_3DIV with data (fine line) is overlaid with fit (thick line), which incorporates three NQIs with angular frequencies, ω_0 = 0.35 (η = 0), 0.268 (η = 0.18) and 0.17 (η = 0.5) rad/ns. The three frequencies were fit to CdS_3O in *exo* and *endo* conformations as well as a CdS_3N coordination mode. Data from ref. 10.

dependent behavior of the LMCT showed that the titration curve has a single transition that fits well to the simultaneous deprotonation of two cysteine thiols with a pK_{a2} of 10.6 ± 0.1. ^{113}Cd NMR was also performed and two signals were detected—one at 595 ppm and the other at 583

ppm. It was initially thought that these two signals corresponded to the *endo* and *exo* conformations of Cd(II) with respect to the plane of the coordinating cysteine.

To gain further insight on the coordination environment of α_3D**IV**, ^{111m}Cd PAC studies were performed. At pH 8.1, the PAC spectrum showed three NQIs with with angular frequencies ω_0 = 0.35 (η = 0), 0.268 (η = 0.18) and 0.17 (η = 0.5) rad/ns. The first two signals fit very well to pure, four-coordinate CdS_3O species in the *exo* and *endo* conformations, respectively. The third signal has a higher asymmetry parameter, which indicates that the ligands are likely different from the other two coordinations observed. This was best fit to a CdS_3N coordination, which invoked the coordination of a nearby histidine at position 72 to complete the coordination sphere. As a result of this analysis, the resonance at 595 ppm in the NMR was assigned to this coordination. More recently, our laboratory solved the NMR structure for apo-α_3D**IV**, which revealed that the introduction of cysteine residues in the interior of the bundle did not alter the overall structure by very much.[12]

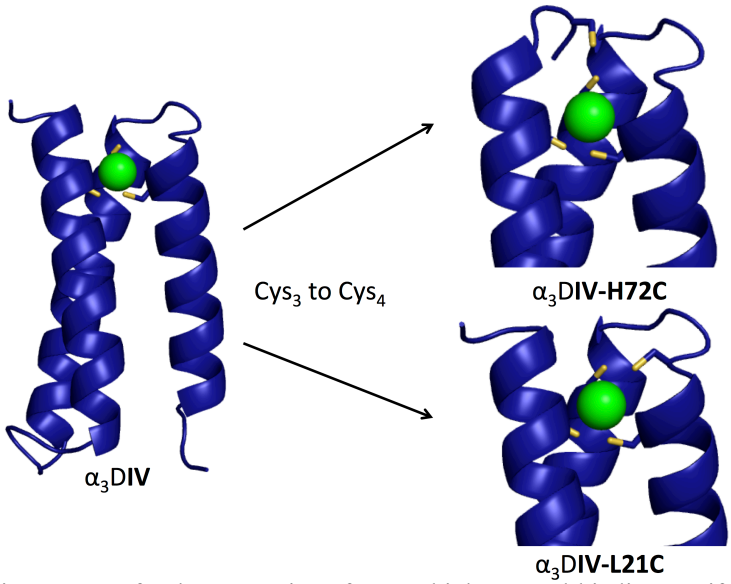


Figure 2.5 Design strategy for the generation of a tetrathiolate metal-binding motif in the α_3D**IV** scaffold. Here, the substitution of His72 (top right), which coordinates Cd in α_3D**IV** could allow for tetrathiolate coordination of Cd(II). An alternative approach substitutes Leu21 to create a CXXC binding motif. Models are based on PDB: 2MTQ

In looking to incorporate a four cysteine binding site in α_3D, we felt a logical first step would be to create His72Cys in α_3D**IV**, which ^{111m}Cd PAC and ^{113}Cd NMR data revealed can coordinate Cd(II), with the other three ligands provided by the cysteine residues already introduced into

α_3D**IV** (Figure 2.5). As an alternate approach, we also identified another residue, Leu21, which is located two residues away from Cys18, one of the cysteines incorporated into α_3D**IV**. This site would incorporate a CXXC motif, much as is found in zinc fingers, rubredoxin, and Cd(II)-sensing proteins. These two constructs represent two reasonable approaches to generating S_4 sites, and characterizing them could also provide some insights into the basic requirements for tetrahedral sites.

Materials and Methods

Protein Production and Purification

The gene for α_3D**IV**-H72C-GSGC was ordered from Celtek genes (Franklin, TN) subcloned into vector pET15b with ApR as a selective marker. The construct α_3D**IV**-L21C-GSGC was generated by sequenctial site directed mutagenesis of α_3D**IV**-H72C-GSGC by QuikChange kit (Stratagene) to replace the histidine in the 72nd position (C72H) and incorporating the L21C change. The plasmids were transformed into *E. coli* BL21 (DE3) and plated on LB Amp plates. A single colony was grown overnight in a starter colony and used to inoculate auto-induction media.[13] The cells were grown for 15-20 hours at 25 °C. The cells were resuspended in 1X PBS with lysozyme and 5 mM DTT then lysed by microfluidizer. After 30 min of heat denaturation at 55°C, the cell lysate was acidified and the soluble fraction collected after ultracentrifugation at 15K rpm at 4°C. The pure proteins were isolated by HPLC on a C18 reverse phase column and lyophilized to yield pure, white powder. Identity of the proteins was confirmed by ESI-MS.

UV-visible Spectroscopy

All experiments were carried out using a Cary 100 instrument. Purified, lyophilized proteins were resuspended and the concentration determined using the calculated molar extinction coefficient based on the aromatic residue content. Cadmium titrations were done in 50 mM CHES buffer at pH 8.6 at a peptide concentration of 20 μM. A stock solution of 0.0146 M $CdCl_2$, which was standardized by ICP, was titrated into the peptide solution anaerobically. Direct metal binding titrations were fitted to the following equation. The equation was derived beginning with Beer's Law and the relationship,

$$K_d = \frac{[M][pep]}{[pepM]}$$

which was evaluated with the quadratic equation yields the following binding isotherm:

$$Abs = \frac{\varepsilon \times \mathrm{b} \times (K_d + [M] + [pep]) - \sqrt{(K_d + [M] + [pep])^2 - 4 \times [\mathrm{M}] \times [pep]}}{2}$$

The pH dependence of the LMCT was assessed by binding Cd(II) anaerobically before titrating in small aliquots of concentrated acid or base in the presence of air. The pH was allowed to stabilize before spectra were measured. Data were fit to the following n proton equation as described previously[14]:

$$Y = \frac{\varepsilon_{dep} + \varepsilon_{prot} \times 10^{n(pK_a - pH)}}{1 + 10^{n(pK_a - pH)}}$$

^{113}Cd Nuclear Magnetic Resonance Spectroscopy

Samples were prepared anaerobically in 10% D_2O with 0.8 equivalents of enriched ^{113}Cd with peptide concentrations of 2-3 mM, and the pH was adjusted to 8.6. Samples were added to a Shigemi solvent-matched NMR tub and sealed with parafilm. The spectra were collected on a 500 MHz Varian spectrophotometer.

^{111m}Cd Perturbed Angular Correlation Spectroscopy

All perturbed angular correlation (PAC) experiments were performed using six detectors at a temperature of 2 °C, which was controlled by a Peltier element. The radioactive cadmium was produced on the day of the experiment at the University Hospital cyclotron in Copenhagen and extracted as described previously,[15] except that the HPLC separation of zinc and cadmium was omitted in order to avoid chloride contamination of the sample. This procedure may lead to zinc contamination of the sample, but the level of contamination (a few micromolar) should not interfere with the experiment. The ^{111m}Cd solution (10-40 μL) was mixed with nonradioactive cadmium acetate and TRIS buffer. The α_3D**IV**-H72C-GSGC or α_3D**IV**-L21C-GSGC peptide was then added (dissolved in ion-exchanged water), and the sample was left to equilibrate for 10 min to allow for metal binding. Finally, sucrose was added to produce a 55% w/w solution, and the pH of the solution was adjusted with H_2SO_4 or KOH. To measure the pH, a small volume of sample was removed from the solution to avoid chloride contamination of the sample. [Note: The pH was measured at room temperature the following day.] Because of the pH dependence on the temperature of TRIS solutions, the pH of the solution at 1 °C was calculated using pH(1 °C))=0.964[pH(25 °C)] + 0.86.] The samples were either used immediately after preparation or left on ice until the measurement was started. All buffers were purged with Ar and treated so as to lower metal contamination. Time resolution of the measurement was 0.860 ns, tpc 0.562 ns. All

fits were carried out with 300 data points, disregarding the 5 first points due to systematic errors in these. Two NQIs were included in the analysis. For the minor species the linewidth ($\Delta\omega_0/\omega_0$) and the asymmetry parameter (η) were fixed at the value obtained from the other spectrum (where it is the major species).

X-Ray Absorption Spectroscopy

Samples were dissolved in 25 mM CHES buffer in the absence of chloride with 0.66 eq of Cd(II) to ensure the absence of free cadmium in the sample. The samples were mixed with 30% glycerol as a glassing agent, loaded into a sample cell, and frozen in liquid nitrogen. Measurements were carried out at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 7-3 with a Si (220) double-crystal monochromator and a flat Rh-coated harmonic rejection mirror. Samples were maintained below 10 K with an Oxford Instruments liquid helium cryostat. Data were measured as fluorescence excitation spectra using a 30-element Ge array detector normalized to incident intensity measured with an N_2-filled ion chamber. Data were measured with steps of 0.25 eV in the XANES region (1 s integration time) and 0.05 $Å^{-1}$ in the EXAFS region to k = 12.5 $Å^{-1}$ (1–20 s integration, k2 weighted). Data reduction and fitting were performed in EXAFSPAK.

Results and Discussion

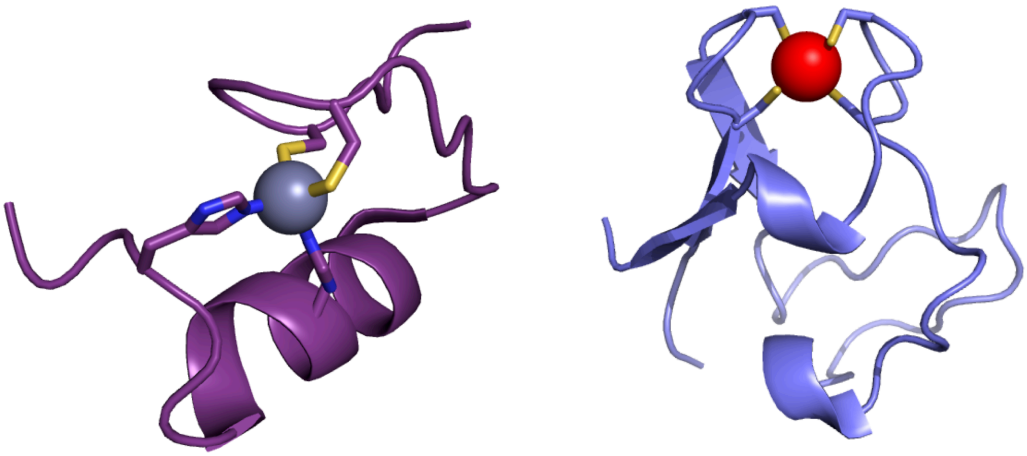


Figure 2.6 Zinc finger proteins (left) and rubredoxin (right) bind metals using flexible loop regions. The apo-CadC structure could not detect density for the flexible N-terminal region that provides two of the four coordinating cysteine and is not pictured here. (PDB: 3ZNF, 1DSX)

Our goal was to engineer a rubredoxin-type site into α_3D**IV**. Constructing such a site in α_3D**IV** presents a significant challenge because native sites that feature tetrathiolate coordination

typically utilize at least one loop or hairpin secondary structural element to coordinate the metal. Rubredoxin and zinc fingers use the canonical pair of Cys-X-X-Cys motifs found in the β-loop region of the protein to bind the metal (Figure 2.6). As described earlier, the ArsR/SmtB family are largely α-helical, and the coordination sphere for heavy metals consists of two cysteine residues from one helix and two from a flexible loop region.

Initially, I also pursued the creation of a tetrahedrally coordinated metal using an exogenous ligand and α_3D**IV**. There is some precedent for the use of exogenous ligands in completing a coordination sphere.[16] Thiophenol was chosen as a potential candidate for an exogenous ligand and was converted to sodium thiophenolate under air-free, basic conditions. Binding studies of thiophenolate to Cd(II)-bound α_3D**IV** monitored by UV-vis were inconclusive, as no obvious binding behavior was observed and absorbance from the phenyl ring

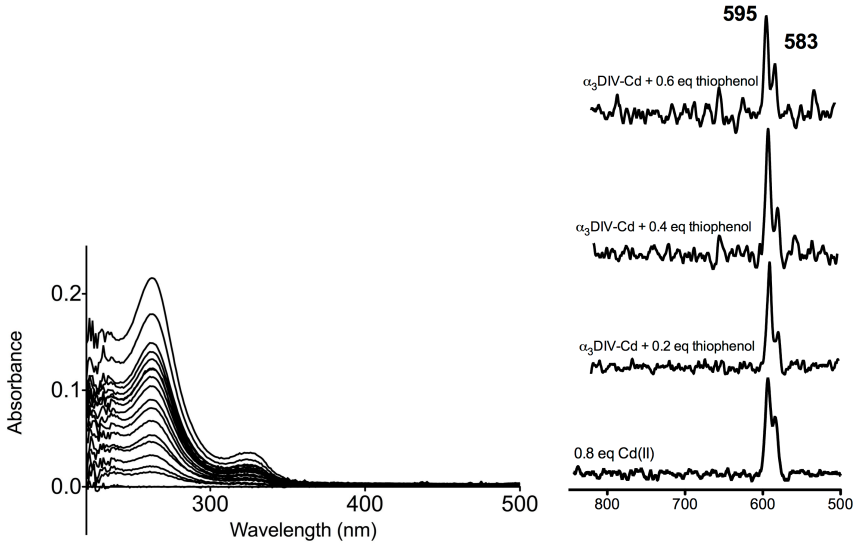


Figure 2.7 (*left*) UV-visible titration of thiophenolate into 2 μM Cd-α_3D**IV** shows only transitions that are associated with the ligand and no additional S—Cd LMCT. (*right)* Titration of thiophenolate into ^{113}Cd NMR sample of Cd-α_3D**IV** in which the same sample was used for all measurements, with small amounts of thiophenolate titrated in. No additional chemical shifts appear to indicate any coordination of Cd in a tetrathiolate environment.

obscured small perturbations (Figure 2.7). Thus, the sensitivity of ^{113}Cd NMR to the primary coordination sphere was exploited to probe the coordination geometry around Cd(II) in the presence of thiophenolate. Initial studies with thiophenol displayed a new feature in the NMR spectra at 602 ppm, but later repetitions revealed this to be an artifact. Based on NMR data, there appears to be no perturbation of the primary coordination sphere from the addition of thiophenolate; thus, thiophenolate is not a suitable exogenous ligand (Figure 2.7).

There were two likely options for designing a tetrathiolate iron site based on the three metal-binding cysteine residues from α_3D**IV**. One involved replacing a histidine in a loop region that previously had been implicated in binding Cd(II), and the other incorporated a cysteine two residues away from one of α_3D**IV** cysteine residues to generate a single CXXC binding motif (Table 2.1). These two constructs are named according to the mutations required to generate them, so replacement of His72 becomes α_3D**IV**-H72C and the chelate motif becomes α_3D**IV**-L21C. The gene for α_3D**IV**-H72C was synthesized by Celtek Genes and subcloned into vector pET15b for overexpression. Due to the inadvertent misplacement of the stop codon, four residues were added to the end of the protein, although this turned out to advantageous as this produced higher expression yields.[17]

The average yield for α_3D**IV**-H72C-GSGC is about 9 mg/L, and that of α_3D**IV**-L21C-GSGC is about 30 mg/L. Both exhibit the classic double well CD spectrum at 208 nm and 222 nm characteristic of a coiled coil or helix bundle protein.

Table 2.1 Sequences of α_3**D** derivatives

Peptide	Sequence
α_3**D**	MGS WAEFKQR LAAIKTR LQAL GG SEAELAAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE LQAY RLN
α_3**D**IV	MGS WAEFKQR LAAIKTR **C**QAL GG SEAE**C**AAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE **C**QAY RHN
α_3**D**IV-H72C	MGS WAEFKQR LAAIKTR LQA**C** GG SEAE**C**AAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE **C**QAY R**C**N GSG*X*
α_3**D**IV-L21C	MGS WAEFKQR LAAIKTR **C**QA**C** GG SEAE**C**AAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE **C**QAY RHN GSG*X*

X is either Ala or Cys depending on the intended application of the peptide

Cd(II) binds tightly to both α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC to form 1:1 complexes. The UV-visible spectra are characterized by LMCT in the UV region. The complex formed by Cd(II) binding to α_3D**IV**-H72C-GSGC is characterized by a single broad peak at 232 nm with a molar extinction coefficient of 25,000 $M^{-1}cm^{-1}$ (Figure 2.8). That of α_3D**IV**-L21C-GSGC shows a strong transition at 224 nm with a shoulder at 243 nm and molar extinction coefficients of 33,000 $M^{-1}cm^{-1}$ and 18,500 $M^{-1}cm^{-1}$, respectively (Figure 2.8, Table 2.1).

Both of these constructs have much higher molar extinction coefficients than previously designed Cd(II)-binding peptides from our labs, but are consistent with native Cd(II)-binding proteins.[18] Previous studies have elucidated that molar extinction coefficients increase linearly with respect to the increasing number of thiolates coordinated to Cd(II),[1] so the increase in molar extinction coefficient observed for these peptides compared to others in our lab is consistent with

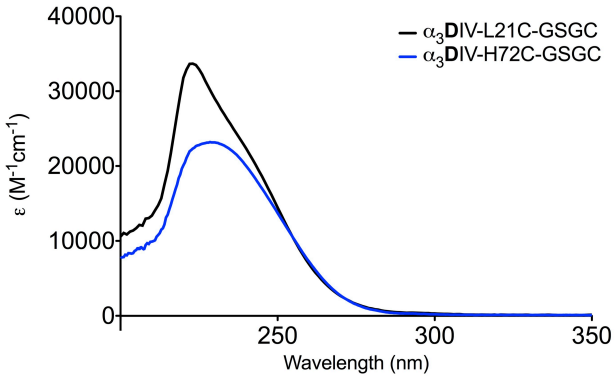


Figure 2.8 UV-visible data for α_3**D**IV-H72C-GSGC (blue) and α_3**D**IV-L21C-GSGC (black) shows different absorption features for the two constructs, which implies the two proteins feature different Cd(II) coordination environments.

increased cysteine coordination. CadC shows a single, broad peak at 238 nm with a molar extinction coefficient of 25,000 $M^{-1}cm^{-1}$. However, the absorption of Cd-substituted rubredoxin is centered at 232 nm with three resolvable transitions at 245 nm, 229 nm, and 213 nm. The molar extinction coefficient for the transition at 245 nm is 26,000 $M^{-1}cm^{-1}$, while that of the peak at 229 nm is estimated to be around 40000 $M^{-1}cm^{-1}$ (based on visual inspection of the figure).[1] This study concluded that the extinction coefficient of the lowest energy transition could be used to estimate the number of thiolates bonded to the Cd(II) based on ~5500 $M^{-1}cm^{-1}$per thiolate. Other studies have used this rule of thumb even in CadC, which only has a single broad peak and the lowest energy transition was not resolvable. [18] The extinction coefficient of the shoulder at 243 nm of α_3D**IV**-L21C-GSGC is lower than the 22000 $M^{-1}cm^{-1}$ benchmark, which could signal that less than 100% of the protein in the sample binds to Cd(II) with four thiolates. However, there appears to be no transition associated with the CdS_3O-type signal found in Cd-substituted TRI peptides and it is unlikely that in the absence of a fourth cysteine coordinating that the Cd(II) is not bound at all. Additionally, this rule of thumb was developed by looking at Cd-substituted rubredoxin, Cd_7-metallothionein, and Cd-substituted S-100b protein, each of which have 4, 20, and one cysteines, respectively.[1] A later paper concurred with this analysis on the basis of their study of a zinc finger protein substituted with Cd(II), which had a resolved low

energy transition at 248 nm with a molar extinction coefficient of 24000 $M^{-1}cm^{-1}$. It is unclear whether this comparison of four native proteins substituted with a non-native metal can be extended to designed peptides although some combination of lack of full coordination and some inherent differences between bonding of the studied native systems and this designed system is most likely. It's worth noting that I was unable to resolve a third, distinct transition at wavelengths shorter than 224 nm, although absorption does exist in this region. The lack of multiple resolvable transitions in CadC and α_3D**IV**-H72C-GSGC is interesting and suggests that perhaps α_3D**IV**-H72C-GSGC is in some ways a good spectroscopic model for CadC. The band structure found in rubredoxins and zinc fingers likely arises from the tetrahedral geometry of the site, and the lack of this structure in CadC is evidence that the Cd(II) coordination could be highly distorted with one very long Cd—S bond, or exchange between CdS_3O and CdS_4, creating the broader transition at an intermediate energy (238 nm vs. 229 nm and 245 nm as found in rubredoxin). This appears to also be the case in α_3D**IV**-H72C-GSGC, which features a similar broad band at an intermediate energy (232 nm vs. 224 nm and 243 nm as found in α_3D**IV**-L21C-GSGC). In comparison to α_3D**IV**, α_3D**IV**-H72C-GSGC has a higher molar extinction coefficient (18,200 vs. 25,000), which supports the concept that this may represent a mixture of two coordination environments and that some greater amount of thiolate coordination is present.

For comparison with α_3D**IV**, the binding constants were determined by fitting the direct titration with Cd(II). Both α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC showed statistically identical dissociation constants (Table 2.2). In fact, the dissociation constants are within an order of magnitude of that of α_3D**IV**. Not only are they within in an order of magnitude of α_3D**IV**, but they are within ~1 order of magnitude of every Cd(II) binding constant determined in this lab by direct metal binding titrations (Table 2.2). To test the results of the fitting I performed the titrations at 1/10th the concentration of a typical experiment and the K_b increased by 1 order of magnitude, indicating that the peptides bind too tightly for the dissociation constants to be accurately determined by direct metal binding titrations. The results of direct binding titrations therefore represent conservative lower limits for the affinity, and the K_b for all peptides in this lab are on the order of nanomolar or tighter. The binding constants for Cd(II) to CadC and other metal sensors have been determined by competition with EDTA, which binds Cd(II) tightly (log β_{11}=16.5). EDTA was titrated into a solution of Cd(II)-α_3D**IV**-H72C-GSGC and the absorbance at 232 nm monitored. Unfortunately, the resulting curve was sigmoidal (Figure 2.9), indicating

the presence of a ternary complex, and the affinity could not be determined satisfactorily even using modeling software such as Hyperquad. Nonetheless, these studies have shown that the affinities of these peptides are much higher than previously reported, and further investigation into accurate methods to determine the affinity are warranted. These other methods may comprise other chelators for direct competition or metal-buffered titrations, or dialysis of the peptide followed by ICP-MS.

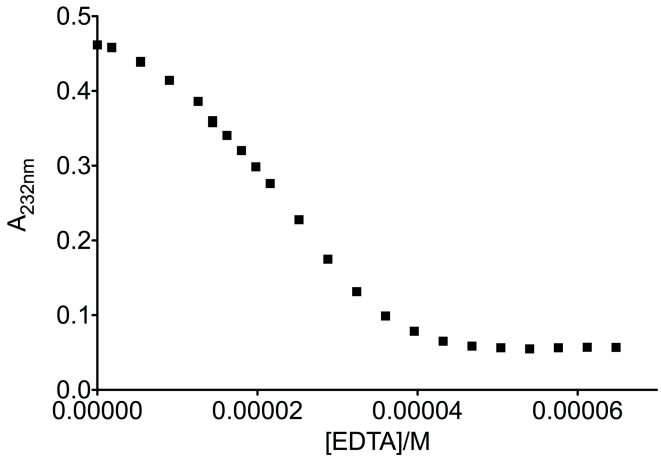


Figure 2.9 Example of titration of EDTA into 20 μM Cd(II)-α_3D**IV**-H72C-GSGC. Titration was monitored by the disappearance of the LMCT at 232 nm. Attempts to globally fit the data, taking into account the pK_as of EDTA, and with the inclusion of a ternary complex were unsuccessful.

Table 2.2 Summary of parameters for Cd(II)-bound *de novo* designed peptides

Peptide	UV-vis	pK_{a2}	^{113}Cd NMR	K_b	EXAFS Cd-S distance
TRIL9C[a]	232 (22,000)	13.4	615	1 x 10^8	2.49
TRIL12C[a]	231 (20,000)	14.6	619	2.6 x 10^7	2.49
TRIL16C[a]	232 (22,600)	13.4	625	1.6 x 10^8	
α_3**D**IV[b]	232 (18,200)	10.6	583, 595	2.0 x 10^7	
α_3**D**IV-H72C-GSGC	232 (25,000)	N/A	595	2.6 x 10^6	
α_3**D**IV-L21C-GSGC	224 (33,000)	N/A	685	2.7 x 10^6	
	243 (18,500)				

[a]from ref[19] and references therein [b]from ref[11]

Typically Cd(II)-bound to **TRI** and **CS** peptides in our lab have also been studied by examining the pH dependence of the LMCT. These experiments can give insight into the dynamics of the Cd(II) within these coiled coils. What has been found previously is that a single thiolate—Cd(II) bond forms at low pH with the other two cysteines being in the thiol form. Simultaneous deprotonation of these two thiols leads to the full LMCT transition. The pK_{a2} values, which are the constants that describe the stoichiometry of two protons for **TRI**L16C, **TRI**L12C, and **TRI**L9C are 13.4, 14.6, and 13.4, respectively (Table 2.2).[20] The pK_{a2} for α_3D**IV**

is several units lower, at 10.6.[11] It was suggested that the single-stranded nature of α_3D**IV** was possibly responsible for this behavior as the folding of the peptide may favor the formation of the deprotonated cysteine complex. Carrying out the same analysis on α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC has proved to be very difficult. Titrations under the same conditions have yielded spectra with shifting baselines, shifting λ_{max}, and molar extinctions that do not match with the UV-visible titrations under constant pH. I have tried beginning at high pH with pre-formed, deprotonated complex and lowering the pH in the presence of TCEP as a reductant, as well as carrying out the entire titration anaerobically, but still see issues with these inconsistencies. I also tried the experiment with α_3D**IV**-L21C-GSGA to see if the exterior cysteine in these constructs could be producing the problem and saw the same behavior from this peptide. From a single experiment on α_3D**IV**-H72C-GSGC that looked more consistent with the expected spectra and was reversible, fitting to an *n* proton equation yielded pK_a of 5.1 as a 1-proton process. A similar single experiment on α_3D**IV**-L21C-GSGC yielded a pK_a of 4.7 for 1.3 protons. It is probable that the models used in previous designs of Cd(II)-binding peptides are not appropriate for this peptide system since a two proton process does not fit well for these systems, especially given the difficulty of obtaining reproducible data during experiments. It is possible that metal binding introduces a global change in the fold, which may be partially responsible for the baseline shifts. Previous studies on zinc finger peptides with Zn(II) bound to four cysteines, showed a single proton process at a pK_a of 4.2, although global fitting was required because the spectra lacked a clean isosbestic point.[21] It seems likely that this type of process is also likely to be relevant in the case of α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC, as both seem to contain at least some proportion of tetrathiolate coordination.

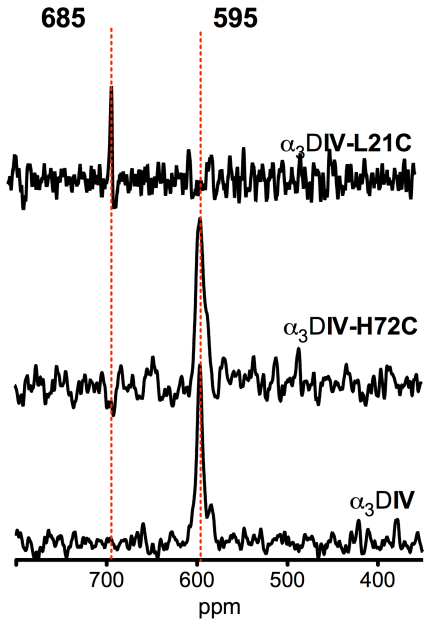


Figure 2.10 ^{113}Cd NMR reveals different coordination environments for α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC, while the coordination environment between α_3D**IV** and α_3D**IV**-H72C-GSGC is similar. All experiments were run with 2-3 mM peptide and 0.8 eq. of enriched ^{113}Cd in 10% D_2O.

The coordination of Cd(II) in α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC was assessed using ^{113}Cd NMR. Spectra from α_3D**IV**-H72C-GSGC with 0.8 eq ^{113}Cd(II) in 10% D_2O show a single peak at 595 ppm, which can also be seen in Cd(II)-bound α_3DIV (Figure 2.10). As a result of this data, the original assignments for α_3DIV ^{113}Cd NMR have been revisited. Initially it was thought that the signal at 595 ppm represented a CdS_3N species while the signal at 583 ppm, reflected a CdS_3O species. Yet comparison of ^{113}Cd NMR data from α_3D**IV**-H72C-GSGC and α_3D**IV** reveals that the signal at 595 ppm actually must represent the CdS_3O species. While very few native proteins exhibit chemical shifts in this region, studies in our lab have shown that this chemical shift range is often associated with CdS_3O. However, a series of studies on rubredoxin mutants with cysteine to serine mutations measured the chemical shifts of these extensively characterized S_3O sites to be between 605-645 ppm. I suggest that the species with chemical shifts <600 ppm in our designed peptides are similar or equivalent to those species that fall between 650-600 ppm in native proteins. By extension, α_3D**IV**-H72C-GSGC is a reasonable model for CadC as CadC has a chemical shift of 622 ppm, although its coordination sphere is still under dispute. On the other hand, the chemical shift of α_3D**IV**-L21C-GSGC is 685 ppm (Figure 2.10). In our lab, this chemical shift has been more typical of CdS_3 sites, although native proteins with pure S_3 coordination of Cd(II) are not known. The chemical shift of ^{113}Cd-substituted rubredoxin has been reported at 730 ppm.[22] However, there are Cd-substituted S_4 sites in native proteins that have shown chemical shifts lower than 700 ppm, although many of

these tend to be S_4 bimetallic bridged structures. However, the TLS protein has a chemical shift of 674 ppm and the single Cd(II) is coordinated to four cysteine residues.[23] The origin of the chemical shift of 685 ppm for α_3D**IV**-L21C-GSGC could be due to a distorted site or to an equilibrium between CdS_3O and CdS_4 forms that exchange faster than the NMR timescale, leading to coalescence of the signals from the two species to a single, intermediate one. It could also be that our designed proteins give ~ 30 ppm depressed chemical shifts, and that by taking this into consideration, this shift represents a designed protein with an S_4 site.

Since ^{111m}Cd PAC can give resolution of multiple species on a faster timescale, we turned to this technique to understand the Cd coordination in these two peptides. These experiments were performed in collaboration with Dr. Lars Hemmingsen and his lab, who measured and analyzed the data. The data sets for α_3D**IV**-H72C-GSGC and α_3D**IV**-L21C-GSGC were initially analyzed with one NQI, capturing the major species, but this did not give satisfactory fits, and thus two NQIs each were included (Figure 2.11, Table 2.4).

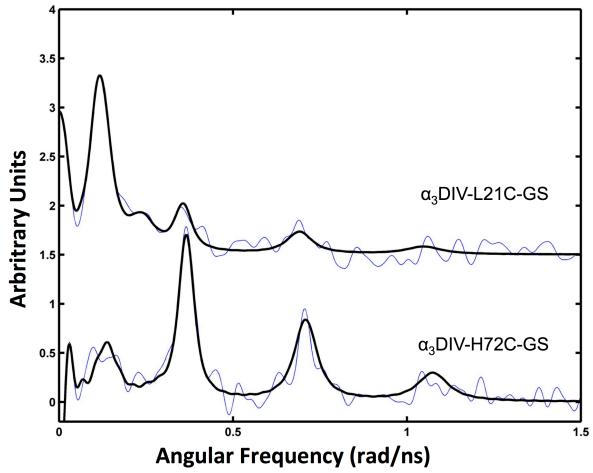


Figure 2.11 ^{111m}Cd PAC spectra for α_3D**IV**-L21C-GSGC (top) and α_3D**IV**-H72C-GSGC (bottom) with the data (fine line) and fit (thick line).

Table 2.3 Parameters fitted to PAC-data. The numbers in parenthesis are the standard deviations of the fitted parameters.

Peptide	pH	ω_0	η	$\Delta\omega_0/\omega_0$	$1/\tau_c$	A	χ_r^2
	(@ 1°C)	(rad/ns)		×100	us^{-1}	×100	
α_3D**IV**-H72C-GSGC	8.6	0.082(3) 0.356(1)	0.79f 0.16(1)	15f 4.2(5)	4.3(7) 4.3(7)	1.7(2) 6.3(4)	1.26
α_3D**IV**-L21C-GSGC	8.6	0.071(2) 0.347(3)	0.79(6) 0.16f	15 (3) 4f	5.2(8) 5.2(8)	5.0(2) 1.9(2)	1.28

f: fixed

One was the NQI recorded as the major species for α_3D**IV**-H72C-GSGC, and the other the major species recorded for α_3D**IV**-L21C-GSGC. This resulted in acceptable fits, and agrees with visual

inspection of the spectra (Figure 2.11). In other words, for α_3D**IV**-H72C-GSGC the major species (ω_0 = 0.356 rad/ns) is accompanied by a minor species (ω_0 = 0.082 rad/ns), which is highly similar to the major species recorded for α_3D**IV**-L21C-GSGC. Conversely, the minor species recorded for α_3D**IV**-L21C-GSGC resembles the major species recorded for α_3D**IV**-H72C-GSGC. The NQI with ω_0 around 0.350 rad/ns is comparable to that reported previously in the literature reflecting a CdS_3O coordination geometry, although with slightly higher frequency.[5,20] The NQI with ω_0 of about 0.080 rad/ns is highly similar to that reported in the literature for the structural (Cys_4) site of horse liver alcohol dehydrogenase[24] and for a peptide that reproduces this site[25]. Thus, the dominating species for α_3D**IV**-H72C-GSGC is most likely a CdS_3O coordination geometry, accompanied by a minor fraction, which is most likely a CdS_4 coordination geometry. Conversely, the dominating species for α_3D**IV**-L21C-GSGC is most likely a CdS_4 site accompanied by a minor fraction, which is a CdS_3O site. The relative populations for each site are given by the amplitudes (A) in Table 2.3.

Taken together, the NMR and PAC spectroscopic data suggest that α_3D**IV**-H72C-GSGC adopts a primarily CdS_3O coordination with a small proportion of CdS_4. The UV-visible spectroscopy supports this model as the spectrum is characterized by a single broad transition. The molar extinction coefficient is higher than that those previously reported for these sites by our group, but this could be due to the small proportion of CdS_4, which can contribute up to 6000 $M^{-1}cm^{-1}$ to the molar extinction coefficient. The ^{113}Cd NMR chemical shift suggests a CdS_3O coordination sphere and aligns well with one of the signals previously seen for α_3D**IV**. Based on this consistency in the chemical shift, it appears that the S_4 site is not detected in the ^{113}Cd NMR spectrum for α_3D**IV**-H72C-GSGC. Unlike ^{1}H NMR, ^{113}Cd NMR signals cannot be integrated, and the strength of the signal is not correlated with the proportion of the species observed so small a proportion of the S_4 may not be sufficient for detection under these conditions. The PAC and UV-visible spectroscopy indicate that the major species for α_3D**IV**-L21C-GSGC is CdS_4. The chemical shift for this peptide is smaller than that for a Cd-substituted classically tetrahedral protein, such as rubredoxin, and this smaller chemical shift could be due to coalescence of the signals for the major and minor species, or be indicative of a distortion in the site, or perhaps a quirk of these types of designed peptides. Given the two species in the PAC data and the good agreement in energy for the LMCT, it seems most likely that coalescence of the signals causes

the smaller chemical shift in the ^{113}Cd NMR or that the nature of the designed peptide shifts the chemical shift upfield slightly.

Taking into account the relative amplitudes of the PAC signals, one can quantitatively analyze the NMR spectroscopy to yield further insight into the observed behavior. By calculation of the relative percentages of the species from the PAC amplitudes, we find that α_3D**IV**-H72C-GSGC is 21.3% CdS_4 and 78.8% CdS_3O. Conversely, α_3D**IV**-L21C-GSGC is 72.5% CdS_4 and 27.5% CdS_3O. Thus we can write the following equations,

$$685 = 0.7246x + 0.2754y$$
$$595 = 0.2125x + 0.7875y$$

where x is the actual chemical shift of the CdS_4 species and y is the actual chemical shift of the CdS_3O species. Solving this system of equations yields a chemical shift of 734 ppm for CdS_4 and 557 ppm for CdS_3O, both of which are reasonable values that fall in line with what has previously been observed. The observed chemical shift for the CdS_3O species is slightly lower (~ 30 pp) than has been observed for both our previous designed peptides, as well as for native proteins with this center, but compared over the ~900 ppm range accessible by ^{113}Cd NMR, this is likely due to systematic error in the calculation of the PAC amplitudes. Indeed, the value of 557 ppm is within error for that previously determined in our lab based on a comparison of S_3 and S_3O sites.[26] These calculations suggest that the observed behavior is due to coalescence of the NMR signals caused by rapid exchange. Assuming the molar extinction coefficient for the resolved low energy transition in α_3D**IV**-L21C-GSGC is only due to CdS_4, a similar type of correction can be done to calculate the real molar extinction coefficient for the CdS_4 species. This calculation yields ε_{243nm} of 25,500 $M^{-1}cm^{-1}$, which reflects the expected molar extinction coefficient for a CdS_4 site.

We can then apply this rationale and calculated values for the pure species to CadC. The NMR signal for CadC, 622 ppm, is then made up of two signals at differing proportions. The equation for this system can be written as,

$$622 = 734x + 557(1 - x)$$

where x is the proportion of S_4 species. This analysis suggests that the NMR signal for CadC is produced by coalescence between two the two forms where 40% exists as S_4 and 60% exists as S_3O.

In this chapter I have detailed strategies that I have pursued to achieve a tetrathiolate site in our designed protein, α_3D**IV**, to help understand Cd(II) coordination in Cd(II)-sensing proteins

such as CadC. In addition to utilizing an exogenous thiolate ligand to bind Cd(II) in an S_4 environment, I pursued two protein designs. In one, replacement of a coordinating histidine in α_3D**IV**, yielded a protein that binds Cd(II) as primarily CdS_3O. The other utilized a chelate motif, CXXC, to create a binding site that binds primarily as CdS_4. In systems like CadC, it appears that the Cd(II) is bound as a distorted S_4 or a rapid exchange between S_4 and S_3O. While no PAC data has been published on CadC, the ^{113}Cd NMR shows a chemical shift of 622 ppm. This is greater than CdS_3O peptides that have been prepared in our lab, but about the same as cysteine to serine mutants of rubredoxin, which show chemical shifts between 605 and 645 ppm. However, molar extinction coefficients of the S—Cd(II) LMCT and reactivity studies of cysteine in the presence of bound Cd(II) suggest that four cysteines coordinate Cd(II). Thus it seems reasonable to invoke either a highly distorted center or a center that can exchange between CdS_3O and CdS_4. Based on the NMR chemical shifts determined for the pure S_4 and S_3O species determined as a result of these studies, CadC likely exists as 40% S_4 and 60% S_3O. In this sense, my designed proteins are good models for either end of that spectrum of CadC as α_3D**IV**-H72C-GSGC has primarily S_3O and α_3D**IV**-L21C-GSGC has primarily S_4. The chemical shift of α_3D**IV**-H72C-GSGC and the rest of its spectroscopy are excellent models for CadC and suggests that CadC is likely primarily a CdS_3O coordination with minor species that are S_4, which is borne out by a quantitative analysis of the spectroscopy. By comparison, I have demonstrated that a system like α_3D**IV**-L21C-GSGC is less consistent with CadC, which supports the conclusion that CadC exists as primarily a CdS_3O coordination sphere.

References:

(1) Henehan, C. J.; Pountney, D. L.; Vasák, M.; Zerbe, O. Identification of Cysteine Ligands in Metalloproteins Using Optical and NMR Spectroscopy: Cadmium-Substituted Rubredoxin as a Model [Cd(CysS)4]2- Center. *Protein Sci.* **1993**, *2*, 1756–1764.

(2) Armitage, I. M.; Drakenberg, T.; Reilly, B. Use of 113Cd NMR to Probe the Native Metal Binding Sites in Metalloproteins: an Overview. In *Cadmium: From Toxicity to Essentiality*; Sigel, A., Sigel, H., Sigel, R. K., Eds.; Metal Ions in Life Sciences; Springer Netherlands: Dordrecht, 2012; Vol. 11, pp 117–144.

(3) Vasák, M. Application of 113Cd NMR to Metallothioneins. *Biodegradation* **1998**, *9*, 501–512.

(4) Hemmingsen, L.; Butz, T. *Perturbed Angular Correlations of Γ-Rays (PAC) Spectroscopy*; John Wiley & Sons, Ltd: Chichester, UK, 2011; pp 1–17.

(5) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. Using Diastereopeptides to Control Metal Ion Coordination in Proteins. *Proc. Natl. Acad. Sci. U. S. A* **2008**, *105*, 16566–16571.

(6) Iranzo, O.; Cabello, C.; Pecoraro, V. L. Heterochromia in Designed Metallopeptides: Geometry-Selective Binding of CdII in a De Novo Peptide. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6688–6691.

(7) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro, V. L. Controlling and Fine Tuning the Physical Properties of Two Identical Metal Coordination Sites in De Novo Designed Three Stranded Coiled Coil Peptides. *J. Am. Chem. Soc.* **2011**, *133*, 239–251.

(8) Bryson, J. W.; Desjarlais, J. R.; Handel, T. M.; DeGrado, W. F. From Coiled Coils to Small Globular Proteins: Design of a Native-Like Three-Helix Bundle. *Protein Sci.* **1998**, *7*, 1404–1414.

(9) Walsh, S. T.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. Solution Structure and Dynamics of a De Novo Designed Three-Helix Bundle Protein. *Proc. Natl. Acad. Sci. U. S. A* **1999**, *96*, 5486–5491.

(10) Chakraborty, S. Designed Metalloproteins: From Structurally Characterized Scaffolds to Helical Bundles. Ph.D. Thesis. University of Michigan, Ann Arbor, MI. Dec. 2010.

(11) Chakraborty, S.; Kravitz, J. Y.; Thulstrup, P. W.; Hemmingsen, L.; DeGrado, W. F.; Pecoraro, V. L. Design of a Three-Helix Bundle Capable of Binding Heavy Metals in a Triscysteine Environment. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2049–2053.

(12) Plegaria, J. S.; Dzul, S.; Zuiderweg, E. R. P.; Stemmler, T. L.; Pecoraro, V. L. Apoprotein Structure and Metal Binding Characterization of a De Novo Designed Peptide, α3DIV, That Sequesters Toxic Heavy Metals. *Biochemistry* **2015**. *In Press.*

(13) Studier, F. W. Protein Production by Auto-Induction in High-Density Shaking Cultures. *Protein Express. Purif.* **2005**, *41*, 207–234.

(14) Yu, F.; Penner-Hahn, J. E.; Pecoraro, V. L. De Novo-Designed Metallopeptides with Type 2 Copper Centers: Modulation of Reduction Potentials and Nitrite Reductase Activities. *J. Am. Chem. Soc.* **2013**, *135*, 18096–18107.

(15) Hemmingsen, L.; Bauer, R.; Bjerrum, M.; Adolph, H.; Zeppezauer, M.; CedergrenZeppezauer, E. The Protein Conformation of Cd-Substituted Horse Liver Alcohol Dehydrogenase and Its Metal-Site Coordination Geometry in Binary and Ternary Inhibitor Complexes. *Eur. J. Biochem.* **1996**, *241*, 546–551.

(16) Frank, P.; Angove, H. C.; Burgess, B. K.; Hodgson, K. O. Determination of Ligand

Binding Constants for the Iron-Molybdenum Cofactor of Nitrogenase: Monomers, Multimers, and Cooperative Behavior. *J. Biol. Inorg. Chem.* **2001**, *6*, 683–697.

(17) Cangelosi, V. M.; Deb, A.; Penner-Hahn, J. E.; Pecoraro, V. L. A De Novo Designed Metalloenzyme for the Hydration of CO2. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7900–7903.

(18) Busenlehner, L. S.; Weng, T.-C.; Penner-Hahn, J. E.; Giedroc, D. P. Elucidation of Primary (Alpha(3)N) and Vestigial (Alpha(5)) Heavy Metal-Binding Sites in Staphylococcus Aureus pI258 CadC: Evolutionary Implications for Metal Ion Selectivity of ArsR/SmtB Metal Sensor Proteins. *J. Mol. Biol.* **2002**, *319*, 685–701.

(19) Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. Peptidic Models for the Binding of Pb(II), Bi(III) and Cd(II) to Mononuclear Thiolate Binding Sites. *J. Biol. Inorg. Chem.* **2006**, *11*, 876–890.

(20) Matzapetakis, M.; Farrer, B. T.; Weng, T.-C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. Comparison of the Binding of Cadmium(II), Mercury(II), and Arsenic(III) to the De Novo Designed Peptides TRI L12C and TRI L16C. *J. Am. Chem. Soc.* **2002**, *124*, 8042–8054.

(21) Sénèque, O.; Bonnet, E.; Joumas, F. L.; Latour, J.-M. Cooperative Metal Binding and Helical Folding in Model Peptides of Treble-Clef Zinc Fingers. *Chem. Eur. J.* **2009**, *15*, 4798–4810.

(22) Xiao, Z.; Lavery, M. J.; Ayhan, M.; Scrofani, S. D. B.; Wilce, M. C. J.; Guss, J. M.; Tregloan, P. A.; George, G. N.; Wedd, A. G. The Rubredoxin From Clostridium Pasteurianum: Mutation of the Iron Cysteinyl Ligands to Serine. Crystal and Molecular Structures of Oxidized and Dithionite-Treated Forms of the Cys42Ser Mutant. *J. Am. Chem. Soc.* **1998**, *120*, 4135–4150.

(23) Iko, Y.; Kodama, T. S.; Kasai, N.; Oyama, T.; Morita, E. H.; Muto, T.; Okumura, M.; Fujii, R.; Takumi, T.; Tate, S.-I.; et al. Domain Architectures and Characterization of an RNA-Binding Protein, TLS. *J. Biol. Chem.* **2004**, *279*, 44834–44840.

(24) Hemmingsen, L.; Bauer, R.; Bjerrum, M. J.; Zeppezauer, M.; Adolph, H. W.; Formicka, G.; Cedergren-Zeppezauer, E. Cd-Substituted Horse Liver Alcohol Dehydrogenase: Catalytic Site Metal Coordination Geometry and Protein Conformation. *Biochemistry* **1995**, *34*, 7145–7153.

(25) Heinz, U.; Hemmingsen, L.; Kiefer, M.; Adolph, H.-W. Structural Adaptability of Zinc Binding Sites: Different Structures in Partially, Fully, and Heavy-Metal Loaded States. *Chem. Eur. J.* **2009**, *15*, 7350–7358.

(26) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. The Correlation of 113Cd NMR and 111mCd PAC Spectroscopies Provides a Powerful Approach for the Characterization of the Structure of Cd II-Substituted Zn IIProteins. *Chem. Eur. J.* **2009**, *15*, 3761–3772.

Chapter 3. *De novo* design of electron transfer proteins: spectroscopic characterization of a designed rubredoxin

Overview

I have previously described my research into the introduction of tetrathiolate centers in α-helical designed proteins and its application to metal homeostasis. Our long-term goal as a lab and as a field is to be able to build proteins with multiple sites and functions. We have shown some ability in this respect with our carbonic anhydrase mimic, which utilizes a Hg(II) structural site and a Zn(II) catalytic site.[1] However, this system does not include any redox-active metals nor redox catalysis, which are subjects of great interest as many important biological processes involve redox reactions. An ideal system to develop would be an electron transfer site coupled to a redox catalytic site, as is seen in many native proteins. The generation of such a system also involves an understanding of how to gate electron transfer within a designed protein. In this chapter I will address some of these issues through the application of my tetrathiolate peptide, α_3D**IV**-L21C, to make a tetrathiolate-coordinated iron protein that reproduces the characteristics of rubredoxin. We are capable of designing and characterizing a rubredoxin site in a fold that is highly divergent from that of native rubredoxin, yet the spectroscopic characteristics of this peptide compare favorably both with other designed proteins as well as with rubredoxin. In this chapter I will detail the characterization of the iron site in my tetrathiolate peptide. This work is still unpublished.

Previous Designs of Iron-Sulfur Sites

The redesign of native proteins has been aided by the development of rational design algorithms. These automated algorithms search proteins of known structure for locations where side chain rotamers can be introduced to form a site of known coordination geometry.[2] This technique has been successful in the design of several metalloproteins. Immunoglobulin G binding domain B1 was designed to have a tetrahedrally-coordinated iron site, mimicking

Figure 3.1 Model of protein redesign effort to make a rubredoxin. The scaffold is an immunoglobulin G binding domain B1, which was computationally searched to find an appropriate site for the incorporation of four cysteine residues. PDB: 2GB1

rubredoxin, but was not robust and decomposed in air due to auto-oxidation reactions (Figure 4.1).[3]

A more successful mimic resulted from the conversion of a disulfide bond in thioredoxin to a rubredoxin site with spectroscopic properties that mimic rubredoxin and which is capable of several cycles of oxidation and reduction.[4] Previously, the same group had demonstrated the success of such an algorithm by designing a catalytically competent Fe superoxide dismutase from thioredoxin by incorporating an iron axially coordinated by histidine, equatorially by (N-His)$_2$(O-Asp) ligands and with open coordination sites for substrate binding.[5] Similarly, thioredoxin has been redesigned to incorporate a self-assembled Fe_4S_4 iron-sulfur cluster, resulting in high potential iron protein (HiPIP)-like spectral characteristics, including an EPR silent resting state.[6]

The ubiquity of iron-sulfur (Fe-S) clusters in proteins from all branches of life and their use in energetically interesting systems such at photosystem I and hydrogenases have promoted the study of their chemical and biochemical properties.[7] Biological centers consisting of iron and sulfur include clusters that are Fe(S-Cys)$_4$ (rubredoxin) sites, 2Fe-2S, Rieske 2Fe-2S, 3Fe-4S, and 4Fe-4S.[8] While rubredoxin sites represent structurally the simplest type of site, they are not necessarily the easiest to study. 4Fe-4S clusters can self-assemble in solution from inorganic sulfide and iron, and for this reason are thought to have comprised some of the earliest metal catalysts that were incorporated into primordial proteins in an anaerobic environment.[7] There are two categories of 4Fe-4S clusters: ferredoxin-type and high potential iron protein (HiPIP). Both are similar, but the protein fold and environment influence the redox potential drastically, with

ferredoxin centers accessing the -2/-1 couple at around -250mV to -400mV and the HiPIPs accessing the -3/-2 couple at +150mV to +350mV.[9] The first instance of a 4Fe-4S cluster incorporated into a *de novo* designed protein was reported by Gibney et al.[10] They synthesized both a protein with a ferredoxin site and one with a heme and a ferredoxin cluster. The reduction potentials of the 4Fe-4S clusters (-350mV) correlates with that of a typical ferredoxin cluster. A new ferredoxin maquette based on the consensus binding motif of *Clostridial* ferredoxins was designed that displays a pH-dependent equilibrium midpoint reduction potential.[11] The pH dependence of the reduction potential was demonstrated to be a one proton-one electron process where protonation may occur at the bridging sulfides. This study was also the first to incorporate 4Fe-4Se clusters into a designed protein.

The 4Fe-4S clusters in Photosystem I (PSI) are at even lower potentials, below -700mV. Recently, a small peptide model was designed that mimics two of the clusters by incorporating the cluster binding residues and nearby active site residues and was successful in reproducing the more negative reduction potentials found in PSI.[12] Further attempts have been made to incorporate 4Fe-4S clusters into helical bundles[13]—a non-natural fold for the site, as well as synthesize 3Fe-4S clusters in *de novo* designed proteins[14]. The 4Fe-4S clusters, when

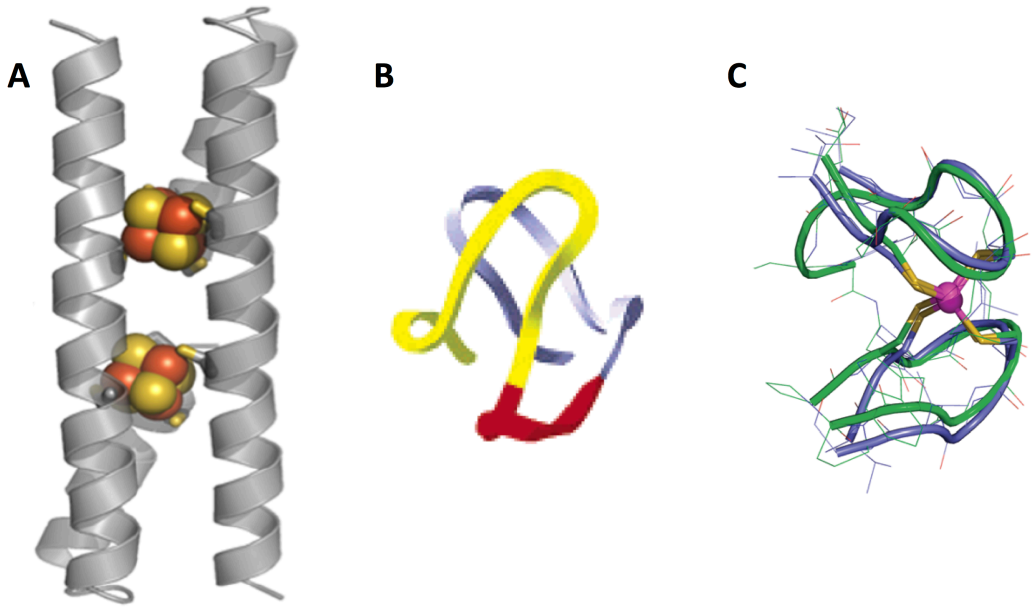


Figure 3.2 a) PyMol model of three-helix homodimer, DSD-Fdm, depicting the two bound 4Fe-4S clusters. From ref 9. **b)** Model of construct reported by Nanda *et al* of structure to mimic hairpin secondary structure. From ref 18. **c)** NMR solution structure (green) of L_{ZR} at pH 6.3 overlaid with the active site of rubredoxin. From ref. 19.

incorporated into the helical bundles exhibited rhombic EPR spectra that are typical of low-potential $[4Fe\text{-}4S]^{+}$ clusters. However, the reduction is irreversible and the cluster is not stable. Hoppe et al. formed 3Fe-4S clusters by mutation of cluster-binding residues, although the clusters were shown to be intermediates in the slow formation of the [4Fe-4S] clusters.

Recently, a three-helix homodimer that coordinates two iron-sulfur clusters in its hydrophobic cluster was reported (Figure 3.2a).[15] The incorporated metal clusters imparted increased thermal stability to the protein and were characterized by UV-visible spectroscopy, gel filtration, analytical ultracentrifugation, and EPR spectroscopy. Pulsed EPR revealed coupling between the two clusters consistent with a 29-34 Å separation between clusters. The second generation design has the clusters within 12 Å of each other—a more biologically relevant distance between clusters; however, they are not electronically coupled despite the reduced distance between the centers.[16] The reduction potential was measured with cyclic voltammetry and found to be -0.479V vs. NHE. The reduced peptide was incubated with native oxidized cytochrome c_{550} and the electron transfer from the 4Fe-4S cluster to the heme was monitored by UV-visible spectroscopy of the Soret bands of cytochrome c_{550}. Furthermore, the triplet state of zinc porphyrin was quenched by the holoprotein, although electron transfer could not be confirmed.[16]

An early attempt to *de novo* design a rubredoxin site was only marginally successful, as the Fe(III) form was not stable in aqueous solution and the UV-vis spectrum had to be measured in methanol instead (Table 3.1).[17] More recently, two designs that largely mimic the hairpin secondary structure around the native rubredoxin metal site have been reported. Nanda *et al* designed a C_2-symmetric metal site inspired by three strand sections of *P. furiosus* rubredoxin linked by a highly stable hairpin motif (Figure 3.2b).[18] This construct is capable of binding Zn(II), Co(II), and Fe(III); the reduction potential is 55 mV (vs. NHE) with iron bound and the center remains stable for sixteen oxidation and reduction cycles.

Jacques et al. recently reported a cyclic peptide model for rubredoxins (Figure 3.2c).[19] The fold and spectroscopic features of rubredoxin are almost perfectly reproduced but the reduction potential, is +140 mV (vs. NHE), rather than the -90 mV to +50 mV generally found in native rubredoxins. This is likely due to the small size of the model's and the greater relative solvent accessibility. This model was then compared with a series of zinc finger peptides with different topologies for their iron-binding capability.[20] The previously published zinc ribbon fold

was compared to a treble clef, a loosened zinc ribbon fold, and the linear peptide version of the zinc ribbon fold. All four peptide are able to form stable 1:1 Fe(II) complexes that can be air oxidized to Fe(III). The fold of the peptide had a large effect on the oxidation rate of the Fe(II) complexes and the stability of the Fe(III) complex. The zinc ribbon fold was found to provide a stronger affinity and better isolates the iron site from external redox agents.[20]

Table 3.1 Properties of designed rubredoxin sites

Construct	**UV-Vis λ_{nm} ($\Delta\varepsilon$ $M^{-1}cm^{-1}$)**	Reduction Potential vs. NHE
Farinas Fe(III) construct[3]	495 (~3000) 350 (~4500), 280 (~15500)	N/A
Nanda Fe(III) form[18]	750 (~500) 600 (~1500) 490 (~2900) 370 (~6800)	55 mV
Jacques Fe(III) form[19]	700 (220) 570 (2900) 491 (5250) 360 (7730)	144 mV

Materials and Methods

Protein Production and Purification

As described previously in Chapter 2.

Absorption Spectroscopy

UV-visible spectroscopy was carried out on a Cary 100 instrument over the range of 700 nm to 250 nm. An Fe(II) stock was prepared in a glovebox under N_2 atmosphere with ferrous ammonium sulfate and the concentration was determined by ICP. Peptides were degassed and loaded with iron under N_2 atmosphere in 20 mM Tris buffer at pH 8.5. Spectra of the oxidized form were generated by exposing the Fe(II) complex to air and scanning until the spectra stabilized. Concentration of the peptide was determined by the tryptophan and tyrosine content of the peptide.

X-Ray Absorption Spectroscopy

Samples were made in 25 mM Tris buffer in the absence of chloride with 0.66 eq of Fe(II) to ensure the absence of free iron in the sample. The samples were mixed with 30% glycerol as a glassing agent, loaded into a sample cell, and frozen in liquid nitrogen.

Measurements were carried out at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 7-3 with a Si (220) double-crystal mono-chromator and a flat Rh-coated harmonic rejection mirror. Samples were maintained below 10 K with an Oxford Instruments liquid helium cryostat. Data were measured as fluorescence excitation spectra using a 30-element Ge array detector normalized to incident intensity measured with an N_2-filled ion chamber. Data were measured with steps of 0.25 eV in the XANES region (1 s integration time) and 0.05 $Å^{-1}$ in the EXAFS region to k = 12.5 $Å^{-1}$ (1–20 s integration, k^2 weighted). Data reduction and fitting were performed in EXAFSPAK.

Mössbauer Spectroscopy

A 10 mM stock solution of ^{57}Fe ferrous ammonium sulfate in 10% sulfuric acid was prepared in an Ar atmosphere glove box. Peptide concentration was determined by DTNB assay and Trp absorption, where the average of the two methods was taken. Samples were made in either unbuffered solution, which was adjusted to pH 8.6 anaerobically, or in 50 mM Tris buffer at pH 8.5 and added to Mössbauer caps either inside the glovebox (for Fe(II)) or exposed to air (for Fe(III)). Fe(III) samples were monitored by UV-vis spectrophotometry before freezing in liquid N_2 to ensure full conversion to oxidized form.

Magnetic Circular Dichroism

Samples were prepared as 1.5 mM peptide with 1.0 mM ferrous ammonium sulfate in 50 mM Tris buffer at pH 8.5. The samples were allowed to air oxidize before adding to the sample holder in the presence of 50% ethylene glycol as a glassing agent. Spectra were collected from -1 to -7 T and from 2K to 50K.

Reduction Potential Determination

Solution electrochemistry was used to determine the reduction potential at pH 8.5. Iron-bound peptide was prepared in a glovebag (Sigma Aldrich) under Ar atmosphere and added to an Ar(g) degassed 5 mL electrochemical cell with 100 mM Tris buffer and 100mM Na_2SO_4 at pH 8.5. Temperature was maintained at 20 °C by using a circulating water temperature controller. Voltammograms were measured on an Autolab potentiostat using a highly ordered graphite

(HOPG) carbon electrode as the working electrode, standard calomel as the reference electrode, and platinum wire as a counter electrode. Measurements were taken at 20 mV/s.

Results and Discussion

Iron-sulfur proteins are among the earliest biologically generated metalloredox centers. The simplest, and one of the most carefully studied, of these sites is the single iron $Fe(Cys)_4$ center found in the rubredoxin family of electron transfer proteins. For this reason, rubredoxin and desulforedoxin have been attractive targets for clarifying structure-function relationships in metalloproteins using protein design.[21] The creation of an FeS_4 center in a designed protein would allow for the investigation of the general requirements for function, beginning with the primary coordination sphere. Deeper analysis would assess whether the classic rubredoxin fold is necessary to obtain the desired tetrahedral geometry or to generate reduction potentials in a physiologically relevant range. In the long-term, such a detailed understanding opens opportunities for the design and implementation of a complex, multi-site, multi-function protein capable of electron transfer and redox catalysis. While there are several notable designs that have been published,[4,18-20] the creation of such a site in the α_3D would represent the largest deviation from the consensus rubredoxin structure and sequence to date.

As I have previously detailed in Chapter 2, I designed tetrathiolate binding sites into two derivatives of α_3D**IV**—one that replaces a histidine in a loop region and the other that creates a single CXXC binding motif. Thus, two potential binding sites are created in a secondary structure that is far removed from the canonical pair of Cys-X-X-Cys motifs found in the β-loop

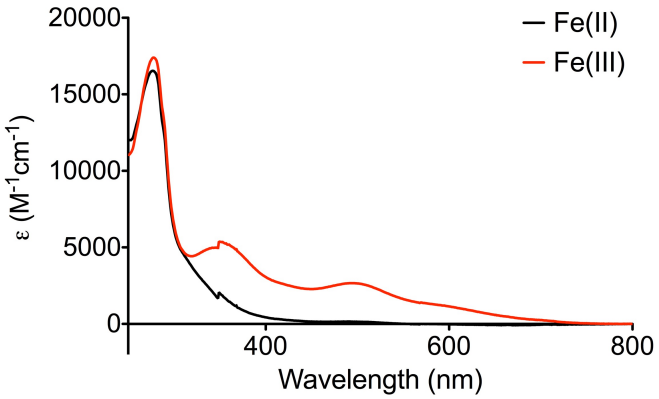


Figure 3.3 UV-visible spectra of α_3D**IV**-L21C with Fe(II) (black trace) and Fe(III) (red trace) incorporated. Measured in 50 mM Tris buffer in the presence of 50 μM TCEP, pH 8.5.

region of the native protein. These proteins were expressed in *E. coli* BL21 (DE3) and purified as previously described.[22] The cell lysate of the peptide containing the chelate motif, termed

α_3D**IV**-L21C, was a dark red color as is often observed during expression of iron-sulfur proteins, while that of the other, α_3D**IV**-H72C, showed no difference in color from typical cell lysate. Both variants of the terminal cysteine (GSGA and GSGC) of α_3D**IV**-L21C were used for these studies and found to have equivalent metal binding sites.

After final separation by HPLC, which removes all associated metal ions, the proteins were anaerobically reconstituted with ferrous ammonium sulfate. Iron binding was monitored by UV-visible spectroscopy and the Fe(III) state was formed by exposure to air. α_3D**IV**-H72C showed some absorbance that was indicative of poor or incomplete binding of iron while that of α_3D**IV**-L21C produced spectra that correspond well with native rubredoxin (Figure 3.2). Both of the spectra for the Fe(II) and Fe(III) forms of the metallated protein correspond well with rubredoxin in terms of the energy of the λ_{max} particularly in the low energy region. In the absence of reductant, the molar extinction coefficients are about three-fold larger in native rubredoxins (Table 3.2). While the rubredoxin values include that of the protein backbone, the contribution of the protein to characteristic absorptions at lower energy, particularly those around 490 nm, is very small. The difference in molar extinction coefficients that is observed likely arises from slight differences in the orbital overlap, or incomplete incorporation of iron into the peptide, such that not all the peptide is bound and contributes to the ferric signal. To test whether the molar extinction coefficient is lower because cysteines are being oxidized and the iron is dissociating from the protein, the reductant TCEP was added during the experiment. Small amounts of TCEP yield molar extinction coefficients that are ~1000 M^{-1} cm^{-1} higher at 490 nm (Table 3.2), while large amounts of TCEP interferes and produces a species with a different absorption profile. Thus, at least part of the reason for the lower molar extinction coefficient is from incomplete binding of the iron in the Fe(III) form. All high concentration spectroscopy that is sensitive to the iron site, i.e. Mössbauer, EPR, XAS, and MCD, were performed with excess peptide to ensure all iron is bound. The lack of any other type of absorbance profile suggests that dissociated iron is likely ferric (hydr)oxy clusters rather than complexed in a different manner.

Table 3.2 Spectroscopic parameters for rubredoxin and designed proteins

Protein	UV-Vis λ_{nm} ($\Delta\varepsilon$ $M^{-1}cm^{-1}$)	Redox Potential (vs NHE)	Mössbauer (ox)	Mössbauer (red)
Rubredoxin	750 (350)[23]	-100 to +50 mV[24]	δ=0.24 [25]	δ=0.70 [25]
	570 (3200)		ΔE_Q= -0.5	ΔE_Q= -3.25
	490 (6600)		D=+1.9	D=+7.4
	370 (7710)		E/D=0.23	E/D=0.28
	350 (7000, sh)		η=0.2	η=0.65
			$A_{x,y,z}$= -16, -15.9, -16.9	$A_{x,y,z}$= -20.1, -8.3, -30.1
RM1[a] [18]	750 (~500) [a]	+55 mV	N/A	N/A
	600 (~1500)			
	490 (~2900)			
	370 (~6800)			
L_{ZR}[19]	700 (220)	+144 mV	δ=0.24	δ=0.69
	570 (2900)		ΔE_Q= -0.5	ΔE_Q= -3.24
	491 (5250)		D=+1.9	D=+7.6
	360 (7730)		E/D=0.23	E/D=0.28
			η=0.13	η=0.84
			$A_{x,y,z}$= -15.8, -15.6, -19.9	$A_{x,y,z}$= -18.3, -8.3, 33
Trx[Rd] [4]	~500 (~4500) [a]			
	~360 (~7000)			
α_3**DIV-**L21C-GSGA	595 (1200) *(900)*[b]	-75 mV (pH 8.5)	δ=0.25(6)	
	491 (2700) *(1700)*[b]		ΔE_Q= -0.5(3)	
	345 (5000) *(3500)*[b]		D=+0.4(2)	
			E/D=0.14(3)	
			η=0.0	
			$A_{x,y,z}$= -16.6, -15.8, -17	
α_3**DIV-**L21C-GSGC	595 (900)[b]	-80 mV (pH 8.5)	δ=0.26(6)	δ=0.72
	491 (1300)		ΔE_Q= -0.5(3)	ΔE_Q= -3.27
	345 (2500)		D=+0.5(2)	
			E/D=0.15(3)	
			η=0.0	
			$A_{x,y,z}$= -15.9, -16, -17	
α_3**DIV-**H72C	~490 (~1400)			
	~350 (~3200)			

[a]From visual inspection of published spectra [b] molar extinction coefficients in absence of reductant

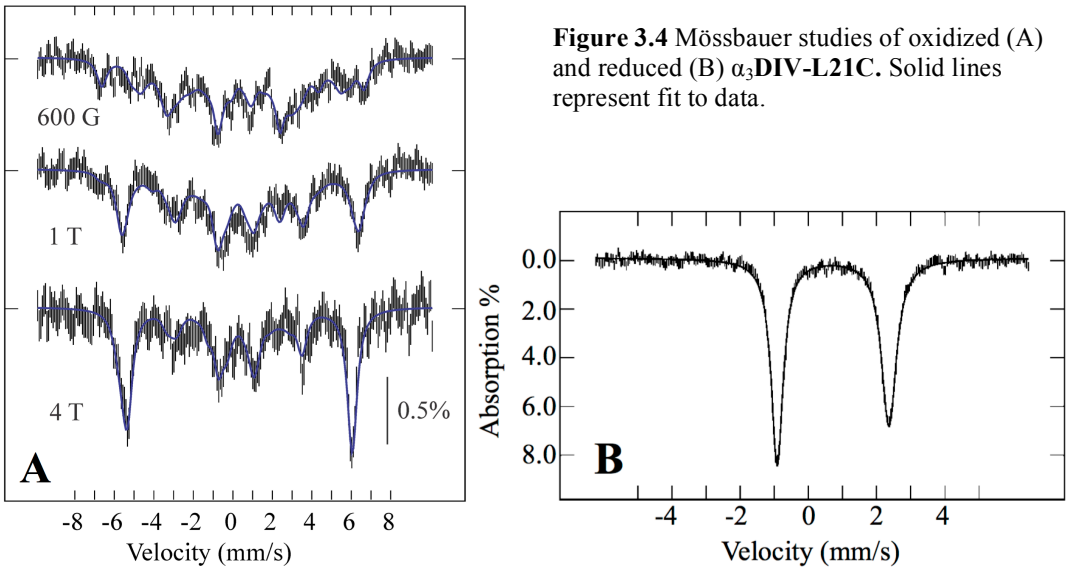


Figure 3.4 Mössbauer studies of oxidized (A) and reduced (B) α_3**DIV-L21C.** Solid lines represent fit to data.

The electronics of the iron site in α_3D**IV**-L21C were also investigated with Mössbauer spectroscopy (Figure 3.4). Mössbauer of the Fe(II) form revealed an isomer shift of 0.72 mm/s, which is typical of Fe(II) in tetrahedral sulfur coordination and fully consistent with native rubredoxin. The large line width and the line asymmetry observed at 4.2 K are a result of the hyperfine coupling between the nuclear spin and the electronic S=2 ground quintet. For the Fe(III) form, the isomer shift also clearly indicates a tetrahedral sulfur environment with a value of 0.25(6) mm/s . The hyperfine and zero field splitting (ZFS) parameters are consistent with a rubredoxin-type center. All three spectra show diffuse absorption under base line, which is not accounted for by the simulation. Nevertheless, these values for the FeS_4 agree well with both published values for rubredoxin as well as other rubredoxin mimics and indicate the successful production of a tetrahedrally-coordinated iron in both ferrous and ferric oxidation levels. Mössbauer was collected for Fe(II)-α_3D**IV**-L21C-GSGC, Fe(III)-α_3D**IV**-L21C-GSGC, and Fe(III)-α_3D**IV**-L21C-GSGA. There were several failed samples for each GSGC and GSGA peptide in each redox state, which was attributed to large pH fluctuations upon freezing and was remedied by increasing the buffer concentration in the samples. The Fe(III)-α_3D**IV**-L21C-GSGC and Fe(III)-α_3D**IV**-L21C-GSGA samples resulted in identical fits. Analysis is currently being undertaken by my collaborators on variable field samples of Fe(II)-α_3D**IV**-L21C-GSGC.

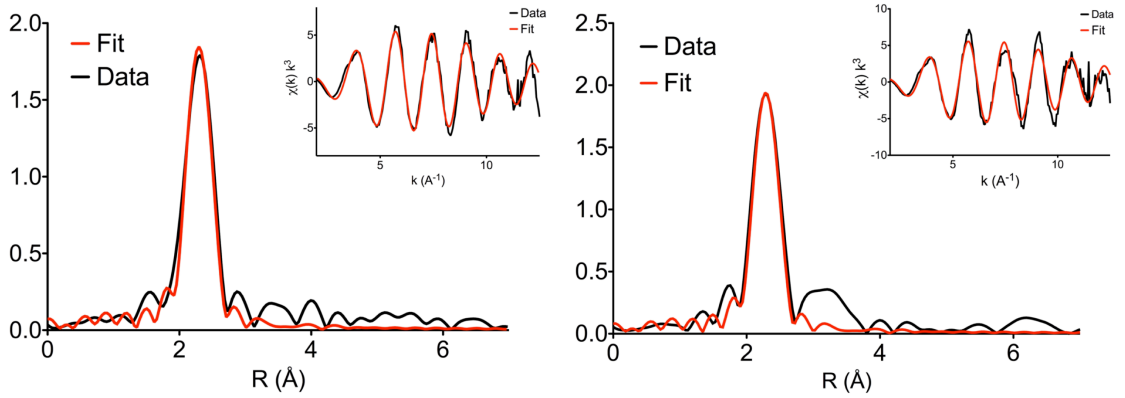


Figure 3.5 EXAFS fits show that α3D**IV**-L21C-GSGA (left) and α3D**IV**-L21C-GSGC (right) have equivalent first coordination spheres. Insets are χ(k) comparisons.

X-Ray absorption spectroscopy (XAS) was performed using the iron K-edge on α_3D**IV**-L21C to give a detailed understanding of the structural features of the iron site in the reduced form of the protein. Previous attempts to run XAS on the oxidized form of the protein resulted in immediate photoreduction of the site. The extended X-Ray fine structure (EXAFS) region was used to analyze the coordination of the reduced iron. The Fourier transform of both Fe(II)-α_3D**IV**-L21C-GSGC and Fe(II)-α_3D**IV**-L21C-GSGA are dominated by a single well-resolved peak at 2.3Å with no other significant backscatterers. For Fe(II)-α_3D**IV**-L21C-GSGA, the spectrum is well reproduced by a single shell of four sulfurs coordinated to the iron center at a distance of 2.32 Å with σ^2 of 0.00597 (Figure 3.5, Table 3.3). This corresponds well with previous EXAFS studies of native reduced rubredoxins, which have found distances of 2.32 Å.[25,26] The fit to the data of the Fe(II)-α_3D**IV**-L21C-GSGC sample produces essentially the same fit, with the exception of oscillations at ~3 Å. These could belong to the β methylene carbons, although there is no precedent for resolution of these scatterers in sulfur-ligated metals.

Table 3.3 Comparison of results of EXAFS fits for α3D**IV**-L21C

Shell	Fit result
GSGA Fe-S	R=2.32(5) σ^2=0.00597 ΔE_0=-14.866 F=130.3
GSGC Fe-S	R=2.31(9) σ^2=0.00555 ΔE_0=-15.278 F=249

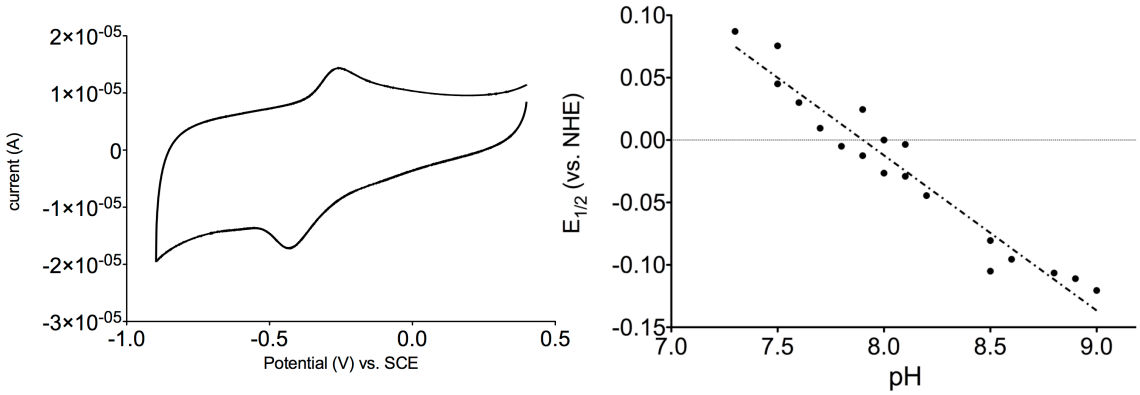


Figure 3.6 (left) Representative cyclic voltammagram of α_3**DIV-L21C** in solution with 100 mM Tris buffer, 100 mM Na_2SO_4, pH 8.5, scan rate 20 mV s^{-1}. (right) Dependence of apparent *E*° on pH. Dotted line is fit to Nernst equation (eq. 1).

The electron transfer properties of this construct were also investigated. Solution electrochemistry of α_3D**IV**-L21C-GSGA revealed a reversible one electron transfer with an apparent reduction potential, $E_{1/2}$, of -75 mV (vs. NHE) at pH 8.5 (Figure 3.6), the same pH at which the spectroscopic properties of this peptide were measured. This is very similar to the range of native rubredoxins, which have potentials of -90 to +50 mV vs NHE.[27-30] The pH dependence of the redox potential was investigated between pH 7.4 and 9.0 and fit to the Nernst equation (eq. 1).

$$E_{1/2} = E_o + \frac{0.05916}{n} logQ - \frac{0.05916}{n} pH \quad (1)$$

Fitting the pH dependence over this range yields a slope of 0.124 V, which indicates a 2-proton/1-electron process (Figure 3.6).[31] This is surprising, as native rubredoxin does not exhibit any pH dependence for its reduction potential. The thiolates in rubredoxin are hydrogen bonded to other residues and cannot be protonated. Rubredoxins that have been mutated to introduce an FeS_3O center do show pH dependence; however, the rest of the spectroscopy of α_3D**IV**-L21C precludes this as a reasonable explanation for this behavior in α_3D**IV**-L21C. It is more likely that the high number of charged residues around the site, which are derived from residues that make up the inter-helical interface in α_3D**IV** are responsible for the pH dependence of this site. In particular, the histidine that is known to coordinate in α_3D**IV** as well as Ser24, Arg71 and Ser75 (not pictured) could all contribute to this pH dependence (Figure 3.7). Extending this trend of pH dependence to pH 7.0 results in a reduction potential of +110 mV, which is about 50 mV higher than that of native rubredoxin. Based on the findings of comparative studies between

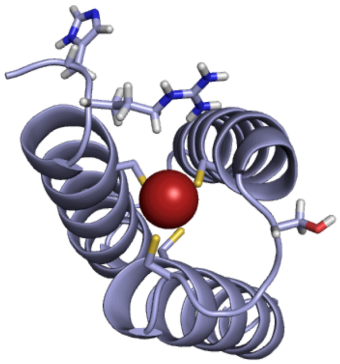


Figure 3.7 PyMol representation of some of the charged residues around the metal site in α_3D**IV**-L21C that may contribute to the pH dependence of the FeS_4 site (PDB: 2MTQ)

rubrerythrin and rubredoxin—which share the same fold and same site but have vastly different reduction potentials (+230 mV vs. ~0 mV)[27,32,33], I suggest that the pH dependence is due to the presence of these same polar groups surrounding the redox site (*vide infra*). In rubredoxin and rubrerythrin, the difference in these reduction potentials was due to the presence of two polar residues in rubrerythrin, Asn and His, which allowed for multiple side chain conformations in both the reduced and oxidized states and changed the polarization response of the environment of the redox site.[32] Given the high number of polar groups around one side of the redox site, a similar process could be occurring in this system α_3D**IV**-L21C. The acquisition of a metallated structure for α_3D**IV**-L21C could be used as the basis for similar computational analysis as in the comparison between rubrerythrin and rubredoxin to understand the pH dependence of this system. It should also be noted that I performed equivalent experiments on α_3D**IV**-L21C-GSGC and found the same pH dependence behavior and reduction potential.

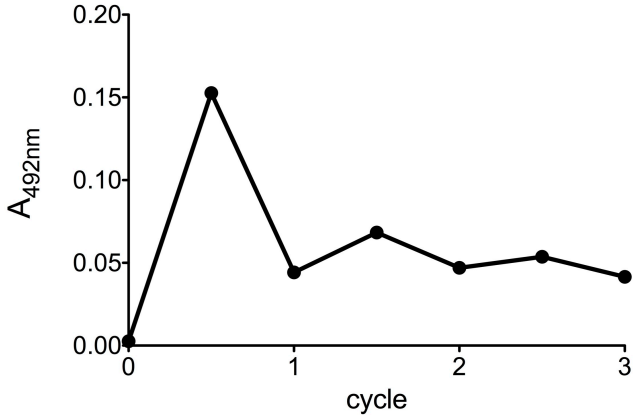


Figure 3.8 The redox reversibility of α_3D**IV**-L21C was tested by repeated exposure to air followed by reduction with sodium dithionite

An important benchmark for function is having two stable oxidation/ reduction states and the ability to cycle between them as native electron transfer proteins can. To test this, α_3D**IV**-L21C-Fe(II) was exposed to air and allowed to oxidize and then re-reduced with sodium

dithionite. The peptide was capable of sustaining three oxidation and reduction cycles, as measured by the absorbance at 492 nm, before gradually being irreversibly bleached (Figure 3.8). Since no precipitate was observed, the protein was likely not denatured to an unfolded state, but iron may have been displaced upon cysteine oxidation.

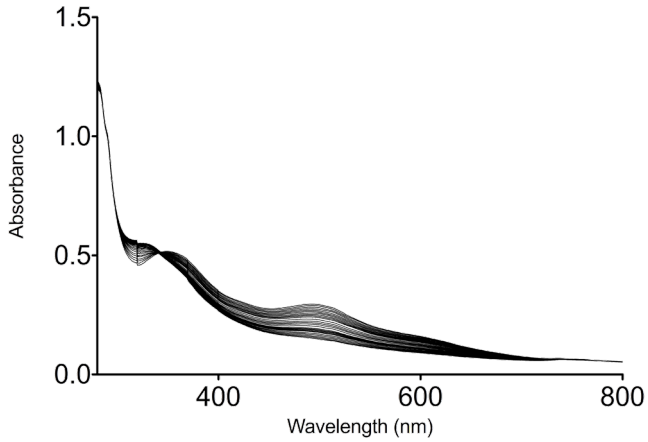


Figure 3.9 Reaction of α_3D**IV**-L21C with ascorbate monitored over the course of 2.5 hours at room temperature.

Reduction of Fe(III)-α_3D**IV**-L21C with other reductants have also been performed. DTT has largely been avoided, due to its ability to coordinate metal centers directly, but β-mercaptoethanol and ascorbate have both been used. Both result in the slow and (particularly in the case of ascorbate) incomplete reduction of the iron center, although the isosbestic point, particularly at early timepoints, indicates a clean conversion from one species to another and supports the bistability of this system (Figure 3.9).

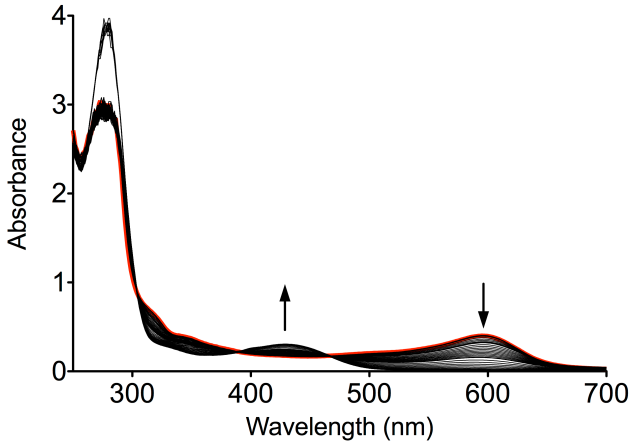


Figure 3.10 α_3D**IV**-L21C-GSGA can equilibrate with indigo trisulfonate in the presence of methyl viologen, xanthine, and xanthine oxidase.

This can also been seen in early experiments to clarify the reduction potential, where xanthine oxidase activity is coupled to a colorimetric redox indicator at low enough concentrations of xanthine oxidase so as to enforce equilibrium conditions at all times.[34] When the redox indicator is within 30 mV of the protein of interest so that the two can equilibrate and clean isosbestic points exist, the proportion of reduced and oxidized species can be calculated and their relative proportions related to the known potential of the dye. Flavin mononucleotide, methylene blue, and indigo trisulfonate were all tested at pH 8.5, and the highest potential dye, indigo trisulfonate ($E^{o'}$= -70 mV) was the only one which was capable of equilibrating with α_3D**IV**-L21C-GSGA (Figure 3.10). Calculation of the reduction potential by this method was complicated by the difficulty of isolating an appropriate isosbestic point for the protein and I eventually turned to cyclic voltammetry to measure the reduction potential directly. It is worth noting that the measured reduction potential by CV agrees well with the range of reduction potential given by the interaction with indigo trisulfonate.

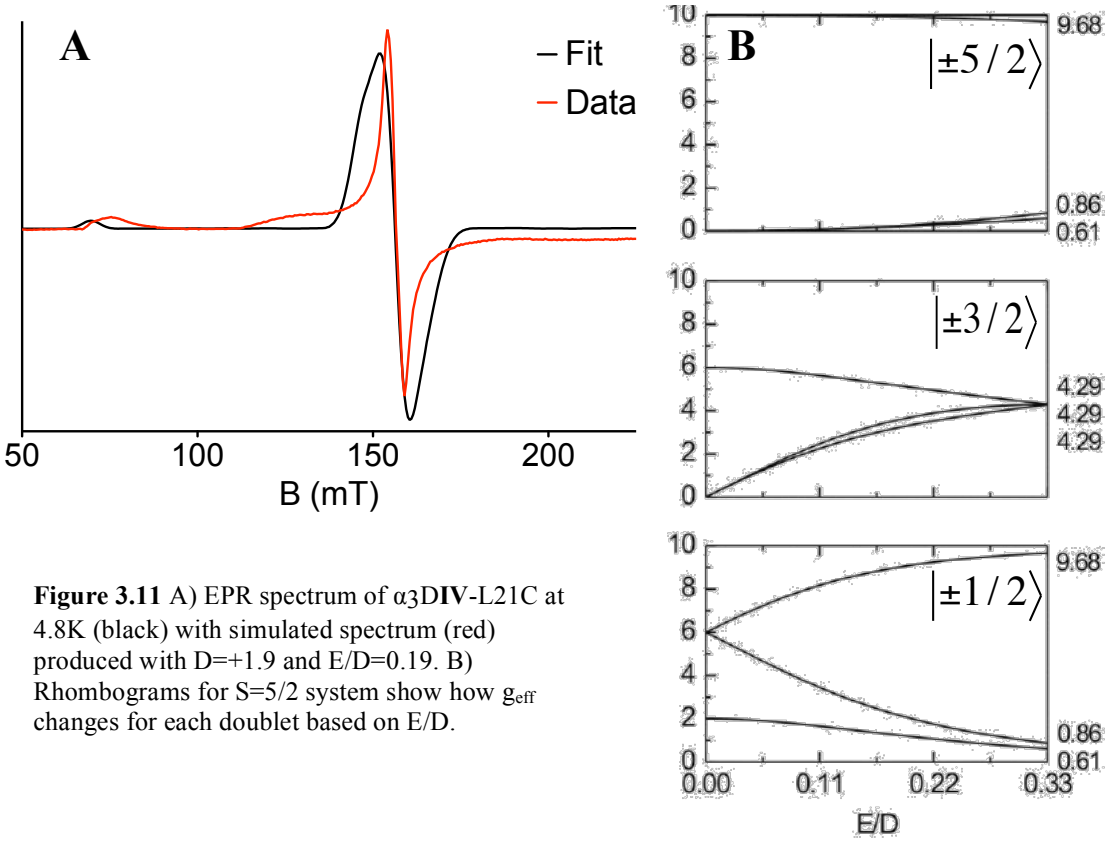


Figure 3.11 A) EPR spectrum of α_3D**IV**-L21C at 4.8K (black) with simulated spectrum (red) produced with D=+1.9 and E/D=0.19. B) Rhombograms for S=5/2 system show how g_{eff} changes for each doublet based on E/D.

Techniques such as magnetic circular dichroism (MCD) and EPR can help refine the electronic and magnetic parameters of iron sites. These rhombic, high spin iron systems can be measured by these techniques only at low temperatures. To clarify the spectroscopy of Fe-α_3D**IV**-L21C-GSGA, in particular the zero field parameters D and E/D, the EPR spectra were measured for the same sample at three different temperatures: 4.8K, 12.5K, and 17K (Figure

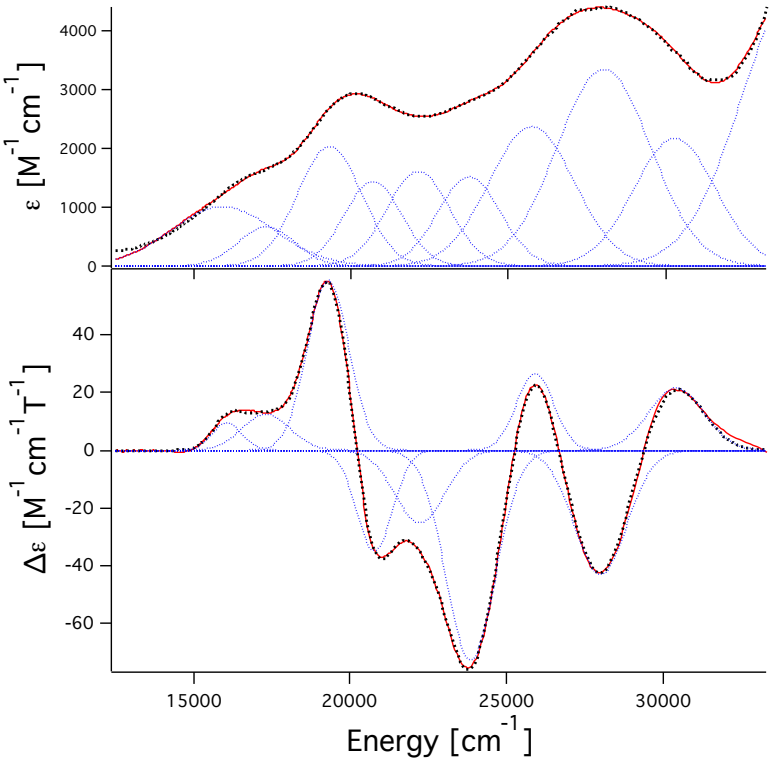


Figure 3.12 Deconvolution of absorption (top) and MCD (bottom) spectra for α_3D**IV**-L21C-Fe was performed in Peakfit. The MCD spectrum was measured at -1 T and 4 K.

3.11). Simple comparison with published spectra of rubredoxin and desulforedoxin shows that the spectra of α_3D**IV**-L21C are similar to those of rubredoxin. However, the low field resonance occurs at slightly lower g-value than that of rubredoxin (9.4 vs. ~8.9), which could be consistent with the lower E/D (0.25 in C. pasteurianum vs. 0.15 in α_3D**IV**-L21C) value calculated from the Mössbauer experiments (Figure 3.9). However, initial attempts to fit the EPR using the parameters from the results of the Mössbauer fits were not successful in simulating the spectra entirely. The best fit was obtained with D = +1.9 and E/D = 0.19 with a large E/D strain (0.1 cm^{-1}) (Figure 3.11). The intensity of the low field resonance was inversely proportional to temperature from 4.8 to 17.5 K, which indicates that it arises from transitions between a ground state Kramers doublet. While previous analysis of rubredoxins have yielded both negative and positive D values depending on the technique and fitting method used, most studies agree that D

is positive for this system.[35,36] The confusion arises partially from the dependence of the effective g-value on E/D. As can be seen in the rhombograms for S = 5/2 systems (Figure 3.11b), in highly rhombic systems the g-values for the $|\pm 5/2\rangle$ doublet and the $|\pm 1/2\rangle$ doublet are identical and thus both the position of the band and its temperature dependence at E/D above ~ 0.26 are insensitive to the sign of D. As the low-field g_{eff} is lower than that of a perfectly rhombic system, it is most likely that the low field resonance arises from the $|\pm 1/2\rangle$ doublet as it does in other rubredoxins, so we can conclude that D is positive, which agrees with the results of the Mössbauer fit. Nevertheless, some features are visible near the g_{eff} ~ 4.3 signal, suggesting partial resolution of the other directional tensors of the doublet. This, coupled with the lower g_{eff} at ~8.9 suggest that E/D is between the 0.15 and 0.29 calculated from other methods, as is likely closer to 0.2. More analysis is required to assess the fitting parameters for this system.

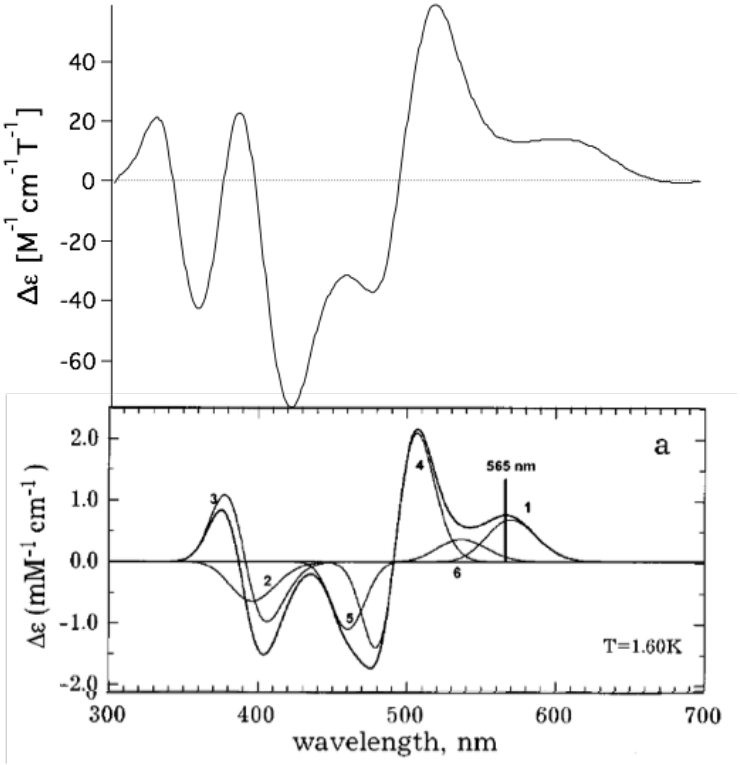


Figure 3.13 Comparison of MCD spectra from *D. gigas* rubredoxin and α_3D**IV**-L21C-Fe. From ref 35.

MCD data was collected and variable temperature-variable field studies were carried out over the temperature range 4K-50K and magnetic fields -1 to -7 T. The absorption and MCD spectra can be fit together to deconvolute the shape of the bands (Figure 3.12). With the exception of an extra MCD band at high energy that was not measured, all of the transitions from the absorption profile are reproduced in the MCD for a total of nine transitions. The saturation

behavior of the bands can be plotted as a function of magnetic field at different temperatures. The analysis of the saturation and the assignment of the bands is still under investigation with collaborator, Dr. Nicolai Lehnert, to understand how closely the electronic structure of α_3D**IV**-L21C-Fe corresponds with that of rubredoxin, although visually they match well (Figure 3.13).

The data reported here indicate that α_3D**IV**-L21C-Fe is a good mimic of a rubredoxin site in a de novo designed scaffold that is structurally far removed from a native rubredoxin fold. Our efforts show that simply placing four cysteines at reasonable distances and distribution around the metal is not capable of producing a rubredoxin site. In fact, all successful designed rubredoxin mimics with two stable oxidation states, including ours, have incorporated at least one CXXC motif. Thus, these motifs appear to form an important functional aspect of successful rubredoxin site design, independent of the secondary structure of the protein around the site. Previous attempts at designing tetrathiolate sites have been focused on two strategies: 1) reproducing the hairpin loops/zinc ribbon fold as closely as possible[18-20] and, 2) redesign of the loop regions of stable α/β proteins[3,4]. Our approach is a complete divergence from these strategies as our site is embedded within α-helices and does not rely on the inherent flexibility of loop regions. Despite this, Fe-α_3D**IV**-L21C does not show deviations from tetrahedral symmetry, as is found in desulforedoxin;[37] thus, not only is our site tetrathiolate-coordinated iron, but it also adopts the pseudotetrahedral symmetry found in native rubredoxins. Therefore, we have shown that tetrathiolate sites can be constructed within a relatively rigid framework by incorporating a CXXC motif.

As might be expected, the two most successful previous designs of a tetrathiolate-coordinated iron site relied on a secondary structure around the metal that exactly mimicked that of rubredoxin.[18-20] Nanda *et al.* designed two hairpin turns fused by a tryptophan zipper motif to generate a 40 residue construct with two β-turns, each with a CXXC motif, just as in rubredoxin. Jacques et al. investigated a series of zinc ribbon folds using cyclic peptide models to incorporate iron into tetrathiolate sites similarly formed in loop regions of the cyclic peptides with CXXC chelate motifs. These two models are the most successful in terms of replicating the spectroscopic features of rubredoxin as well as being able to cycle multiple times between the Fe(II) and Fe(III) states. The Nanda model, RM1, exhibited a UV-visible spectrum in which the energies of the charge transfer bands closely match that of rubredoxin, especially in the low energy region, although the molar extinction coefficients were three-fold smaller than that found

in rubredoxin (Table 2).[18] RM1 had a reduction potential of +55mV, which corresponds to that native rubredoxins (-100 mV to +50 mV) and cycled 16 times, which is the highest number of cycles reported for a designed protein. The L series of peptides, the most successful of which is a zinc ribbon fold, were designed to mimic zinc fingers.[19,20] The zinc ribbon fold peptide, L_{ZR}, impressively exhibited UV-vis data that are nearly identical to those of native rubredoxins in both energy and molar extinction coefficient (Table 2). Additionally, Mössbauer and EPR data were obtained, providing a detailed view of the electronic structure of the site, which nearly exactly matched that of native rubredoxin. The reduction potential of this site was measured at +144 mV vs. NHE and the construct was able to cycle between the Fe(II) and Fe(III) states seven times.

A better direct comparison for the success of our system is a thioredoxin redesign, which introduces a tetrathiolate site into a protein with an α/β fold consisting of a 4 stranded antiparallel β-sheet between three 3 α-helices.[4] A tetrathiolate iron site was computationally designed to take advantage of an extant CXXC motif that usually exists as a disulfide bond with two other cysteine residues incorporated in nearby loop regions. Reconstitution with Fe(II) followed by oxidation with air generated a UV-visible spectrum with characteristic transitions at 350 nm and 500 nm[4], although the relative extinction coefficients and band structure of shoulders at low energy is more reminiscent of spectra of desulforedoxin rather than that of rubredoxin. No detailed spectroscopy capable of distinguishing electronic symmetry structure was performed to examine the root of the electronic absorption spectrum further. While rubredoxin is highly rhombic, desulforedoxin is much less so and has an E/D= 0.08, which is evident in the lower g-values for the lowest Kramer's doublet and the higher g-value of the middle Kramer's double.[37] The UV-visible spectrum is characterized by a sharper charge transfer band at ~380 nm than rubredoxin and a lower energy charge transfer band at 507 nm (vs. 490 nm).[38,39] More characteristic is the relative intensities of the two absorbances—in rubredoxin the two have nearly the same molar extinction coefficient, while in desulforedoxin the band at 380 nm is 1.6X that of the signal at 507 nm. While the two proteins clearly have related iron sites, Mössbauer and EPR revealed the significant difference in symmetry.[36,37,40-42] This thioredoxin redesign construct was capable of cycling three times between the Fe(II) and Fe(III) when mediated by β-mercaptoethanol, but the reduction potential was not reported.

The spectroscopic parameters for α_3D**IV**-L21C-Fe correspond well with those of native rubredoxin and with other previous successful designs that were more reliant on the structural parameters exhibited by the native protein. In contrast, our design reproduces the spectroscopic characteristics of rubredoxin in the largest deviation from the consensus rubredoxin structure yet reported. While all other successful designs rely on loop regions to help define the metal-binding site, our construct uses only a CXXC motif embedded in an α-helical bundle. The α_3D**IV**-L21C-Fe exhibits UV-visible spectra that match in energy to that of rubredoxin and have molar extinction coefficients equivalent to that of the previously reported RM1. And in fact, this construct exhibits energies and relative molar extinction coefficients that better reproduce rubredoxin-type spectra than the other successful, bistable non-hairpin loop construct. Furthermore, Mössbauer, EPR, and MCD data of α_3D**IV**-L21C-Fe correspond well with those reported for both rubredoxin and L_{ZR}. While our construct can only sustain three cycles of oxidation and reduction, no other non-loop construct has been reported to cycle between the oxidized and reduced forms. It is important to note that among the folds most distant from native rubredoxin, the thioredoxin redesign accessed a desulforedoxin center whereas α_3D**IV**-L21C-Fe yields a rubredoxin type site.

One advantage of a *de novo* design strategy is that it allows investigation of well-studied metal centers unhindered by evolutionary constraints in order to assess what are essential factors for specific functions. Clearly, CXXC chelate motifs and the resulting stability and symmetry constraints they impose favor the formation of sites with two stable oxidation states. However, I have demonstrated that hairpin loops and β structure are not required to attain the relevant physical structure (geometry and bond distances), electronic structure, and reduction potential, although the native folds clearly confer some advantage for cycling compared with a non-optimized a-helical environment. This illustrates that *de novo* protein design is able to achieve native-like metalloredox centers even in highly distorted or non-natural folds and suggests that the overall fold is not essential for certain spectroscopic features. Rather, folds that can accommodate certain architectural features (i.e. chelate motifs) and that present ligands at appropriate spatial coordinates may produce similar characteristics, indicating that novel constructs combining metal sites with desired features (e.g., reduction potential or catalytic efficiency) originating from drastically different secondary environments may be combined into

a single, functional homogeneous protein motif that could be dissimilar from one or both of the initial targeted structures. Adapting metal sites from different folds to a generalized framework could allow us to tune these sites for alternative functions.

References

(1) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. Hydrolytic Catalysis and Structural Stabilization in a Designed Metalloprotein. *Nat Chem* **2012**, *4*, 118–123.

(2) Hellinga, H. W.; Richards, F. M. Construction of New Ligand Binding Sites in Proteins of Known Structure. I. Computer-Aided Modeling of Sites with Pre-Defined Geometry. *J. Mol. Biol.* **1991**, *222*, 763–785.

(3) Farinas, E.; Regan, L. The De Novo Design of a Rubredoxin-Like Fe Site. *Protein Sci.* **1998**, *7*, 1939–1946.

(4) Benson, D.; Wisz, M.; Liu, W. Construction of a Novel Redox Protein by Rational Design: Conversion of a Disulfide Bridge Into a Mononuclear Iron-Sulfur Center. *Biochemistry* **1998**, *37*, 7070-7076.

(5) Pinto, A. L.; Hellinga, H. W.; Caradonna, J. P. Construction of a Catalytically Active Iron Superoxide Dismutase by Rational Protein Design. *Proc. Natl. Acad. Sci. U. S. A* **1997**, *94*, 5562–5567.

(6) Coldren, C. D.; Hellinga, H. W.; Caradonna, J. P. The Rational Design and Construction of a Cuboidal Iron-Sulfur Protein. *Proc. Natl. Acad. Sci. U. S. A* **1997**, *94*, 6635–6640.

(7) Holm, R. H.; Kennepohl, P.; Solomon, E. I. Structural and Functional Aspects of Metal Sites in Biology. *Chem. Rev.* **1996**, *96*, 2239–2314.

(8) Venkateswara Rao, P.; Holm, R. H. Synthetic Analogues of the Active Sites of Iron-Sulfur Proteins. *Chem. Rev.* **2004**, *104*, 527–559.

(9) Dey, A.; Jenney, F. E.; Adams, M. W. W.; Babini, E.; Takahashi, Y.; Fukuyama, K.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. Solvent Tuning of Electrochemical Potentials in the Active Sites of HiPIP Versus Ferredoxin. *Science* **2007**, *318*, 1464–1468.

(10) Gibney, B. R.; Mulholland, S. E.; Rabanal, F.; Dutton, P. L. Ferredoxin and Ferredoxin-Heme Maquettes. *Proc. Natl. Acad. Sci. U. S. A* **1996**, *93*, 15041–15046.

(11) Kennedy, M. L.; Gibney, B. R. Proton Coupling to [4Fe-4S](2+/+) and [4Fe-4Se](2+/+) Oxidation and Reduction in a Designed Protein. *J. Am. Chem. Soc.* **2002**, *124*, 6826–6827.

(12) Antonkine, M. L.; Koay, M. S.; Epel, B.; Breitenstein, C.; Gopta, O.; Gärtner, W.; Bill, E.; Lubitz, W. Synthesis and Characterization of De Novo Designed Peptides Modelling the Binding Sites of [4Fe-4S] Clusters in Photosystem I. *Biochim. Biophys. Acta* **2009**, *1787*, 995–1008.

(13) Grzyb, J.; Xu, F.; Weiner, L.; Reijerse, E. De Novo Design of a Non-Natural Fold for an Iron-Sulfur Protein: Alpha-Helical Coiled-Coil with a Four-Iron Four-Sulfur Cluster Binding Site in Its Central Core. *Biochim. Biophys. Acta* **2010**, *1797*, 406-413.

(14) Hoppe, A.; Pandelia, M.-E.; Gärtner, W.; Lubitz, W. [Fe4S4]- and [Fe3S4]-Cluster Formation in Synthetic Peptides. *Biochim. Biophys. Acta* **2011**, *1807*, 1414–1422.

(15) Roy, A.; Sarrou, I.; Vaughn, M. D.; Astashkin, A. V.; Ghirlanda, G. De Novo Design of an Artificial Bis[4Fe-4S] Binding Protein. *Biochemistry* **2013**, *52*, 7586–7594.

(16) Roy, A.; Sommer, D. J.; Schmitz, R. A.; Brown, C. L.; Gust, D.; Astashkin, A.; Ghirlanda, G. A De Novo Designed 2[4Fe-4S] Ferredoxin Mimic Mediates Electron Transfer. *J. Am. Chem. Soc.* **2014**, *136*, 17343–17349.

(17) Lombardi, A.; Marasco, D.; Maglio, O.; Di Costanzo, L.; Nastri, F.; Pavone, V. Miniaturized Metalloproteins: Application to Iron-Sulfur Proteins. *Proc. Natl. Acad. Sci.*

U. S. A **2000**, *97*, 11922–11927.
(18) Nanda, V.; Rosenblatt, M. M.; Osyczka, A.; Kono, H.; Getahun, Z.; Dutton, P. L.; Saven, J. G.; DeGrado, W. F. De Novo Design of a Redox-Active Minimal Rubredoxin Mimic. *J. Am. Chem. Soc.* **2005**, *127*, 5804–5805.
(19) Jacques, A.; Clemancey, M.; Blondin, G.; Fourmond, V.; Latour, J.-M.; Sénèque, O. A Cyclic Peptide-Based Redox-Active Model of Rubredoxin. *Chem. Commun.* **2013**, *49*, 2915-2917.
(20) Jacques, A.; Latour, J.-M.; Sénèque, O. Peptide-Based FeS4 Complexes: the Zinc Ribbon Fold Is Unsurpassed to Stabilize Both the FeII and FeIII States. *Dalton Trans* **2014**, *43*, 3922–3930.
(21) Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayyum, H.; Pecoraro, V. L. Protein Design: Toward Functional Metalloenzymes. *Chem. Rev.* **2014**, *114*, 3495–3578.
(22) Chakraborty, S.; Kravitz, J. Y.; Thulstrup, P. W.; Hemmingsen, L.; DeGrado, W. F.; Pecoraro, V. L. Design of a Three-Helix Bundle Capable of Binding Heavy Metals in a Triscysteine Environment. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2049–2053.
(23) Xiao, Z.; Lavery, M. J.; Ayhan, M.; Scrofani, S. D. B.; Wilce, M. C. J.; Guss, J. M.; Tregloan, P. A.; George, G. N.; Wedd, A. G. The Rubredoxin From Clostridium Pasteurianum: Mutation of the Iron Cysteinyl Ligands to Serine. Crystal and Molecular Structures of Oxidized and Dithionite-Treated Forms of the Cys42Ser Mutant. *J. Am. Chem. Soc.* **1998**, *120*, 4135–4150.
(24) Meyer, J.; Moulis, J.-M. Rubredoxin. In *Handbook of Metalloproteins*; Messerschmidt, A., Huber, R., Eds.; Rubredoxin, vol, 2006; Vol. 1.
(25) Wegner, P.; Bever, M.; Schünemann, V.; Trautwein, A. X.; Schmidt, C.; Bönisch, H.; Gnida, M.; Meyer-Klaucke, W. Iron–Sulfur Proteins Investigated by EPR-, Mössbauer- and EXAFS-Spectroscopy. *Hyperfine Interact.* **2004**, *156*, 293–298.
(26) George, G. N.; Pickering, I. J.; Prince, R. C.; Zhou, Z. H.; Adams, M. W. X-Ray Absorption Spectroscopy of Pyrococcus Furiosus Rubredoxin. *J Biol Inorg Chem* **1996**, *1*, 226–230.
(27) Lovenberg, W.; Sobel, B. E. Rubredoxin: a New Electron Transfer Protein From Clostridium Pasteurianum. *Proc. Natl. Acad. Sci. U. S. A* **1965**, *54*, 193–199.
(28) Bönisch, H.; Schmidt, C. L.; Bianco, P.; Ladenstein, R. Ultrahigh-Resolution Study on Pyrococcus Abyssi Rubredoxin: II. Introduction of an O–H···Sγ–Fe Hydrogen Bond Increased the Reduction Potential by 65 mV. *J. Biol. Inorg. Chem.* **2007**, *12*, 1163–1171.
(29) Moura, I.; Xavier, A. V.; Cammack, R.; Bruschi, M.; Le Gall, J. A Comparative Spectroscopic Study of Two Non-Haem Iron Proteins Lacking Labile Sulphide From Desulphovibrio Gigas. *Biochim. Biophys. Acta* **1978**, *533*, 156–162.
(30) Jenney, F., Jr. [5] Rubredoxin From Pyrococcus Furiosus. *Meth. Enzymol.* **2001**, 45–55.
(31) Migliore, A.; Polizzi, N. F.; Therien, M. J.; Beratan, D. N. Biochemistry and Theory of Proton-Coupled Electron Transfer. *Chem. Rev.* **2014**, *114*, 3381–3465.
(32) Luo, Y.; Ergenekan, C. E.; Fischer, J. T.; Tan, M.-L.; Ichiye, T. The Molecular Determinants of the Increased Reduction Potential of the Rubredoxin Domain of Rubrerythrin Relative to Rubredoxin. *Biophys. J.* **2010**, *98*, 560–568.
(33) LeGall, J.; Prickril, B. C.; Moura, I.; Xavier, A. V.; Moura, J. J.; Huynh, B. H. Isolation and Characterization of Rubrerythrin, a Non-Heme Iron Protein From Desulfovibrio Vulgaris That Contains Rubredoxin Centers and a Hemerythrin-Like Binuclear Iron

Cluster. *Biochemistry* **1988**, *27*, 1636–1642.
(34) Massey, V. A Simple Method for Redox Potential Determination. In *Flavins and Flavoproteins*; Curti, B., Ronchi, S., Zanetti, G., Eds.; Berlin, 1990; pp 59–66.
(35) Oganesyan, V. S.; George, S. J.; Cheesman, M. R.; Thomson, A. J. A Novel, General Method of Analyzing Magnetic Circular Dichroism Spectra and Magnetization Curves of High-Spin Metal Ions: Application to the Protein Oxidized Rubredoxin, Desulfovibrio Gigas. *J Chem Phys* **1999**, *110*, 762–777.
(36) Börger, B.; Suter, D. Magnetic and Optical Anisotropy of Clostridium Pasteurianum Rubredoxin From Optically Detected Electron Paramagnetic Resonance. *J Chem Phys* **2001**, *115*, 9821–9827.
(37) Moura, I.; Huynh, B. H.; Hausinger, R. P.; Le Gall, J.; Xavier, A. V.; Münck, E. Mössbauer and EPR Studies of Desulforedoxin From Desulfovibrio Gigas. *J. Biol. Chem.* **1980**, *255*, 2493–2498.
(38) Moura, I.; Bruschi, M.; Le Gall, J.; Moura, J. J. G.; Xavier, A. V. Isolation and Characterization of Desulforedoxin, a New Type of Non-Heme Iron Protein From Desulfovibrio Gigas. *Biochem. Biophys. Res. Comm.* **1977**, *75*, 1037–1044.
(39) Yu, L.; Kennedy, M.; Czaja, C.; Tavares, P.; Moura, J. J.; Moura, I.; Rusnak, F. Conversion of Desulforedoxin Into a Rubredoxin Center. *Biochem. Biophys. Res. Comm.* **1997**, *231*, 679–682.
(40) Bertrand, P.; Gayda, J. P. A Ligand Field Analysis of the Spectroscopic Differences Between Rubredoxin and Desulforedoxin in the Reduced State. *Biochim. Biophys. Acta* **1988**, *954*, 347–350.
(41) Tavares, P.; Ravi, N.; Moura, J. J.; LeGall, J.; Huang, Y. H.; Crouse, B. R.; Johnson, M. K.; Huynh, B. H.; Moura, I. Spectroscopic Properties of Desulfoferrodoxin From Desulfovibrio Desulfuricans (ATCC 27774). *J. Biol. Chem.* **1994**, *269*, 10504–10510.
(42) Mathies, G.; Blok, H.; Disselhorst, J. A. J. M.; Gast, P.; van der Meer, H.; Miedema, D. M.; Almeida, R. M.; Moura, J. J. G.; Hagen, W. R.; Groenen, E. J. J. Continuous-Wave EPR at 275GHz: Application to High-Spin Fe3+ Systems. *J. Mag. Res.* **2011**, *210*, 126–132.

Chapter 4. Laser flash photolysis studies of *de novo* designed proteins: implications for design of long-rang of electron transfer

Overview

For important processes like photosynthesis and respiration, electron transfer in native systems must occur over large distances (~ 30 Å) with a sufficiently high rate generally considered to be in the millisecond to microsecond timescale, to sustain life. For example, the membrane-embedded proteins that play major roles in respiration and photosynthesis have to transfer electrons from one site of the membrane to the other, often using multiple redox sites to facilitate this process. In photosystems, the excitation of the chlorophyll must produce charge separation and electron transfer processes rapidly enough and with low enough back reactions for this to be a productive event. These proteins all form a matrix in which various redox sites are held at certain positions and orientations with respect to each other. How proteins facilitate electron transfer over long distances—the role of the protein medium—has been a subject of intense study over the last forty years and still there is much to be learned in how biological redox processes work.

While many studies on the factors affecting electron transfer rates have been performed on native proteins, considerably fewer have been performed in designed systems. Protein design methods allow us to define and create redox sites in nonnative folds, incorporate nonnative chromophores, as well as test hypotheses about the role of the protein medium by incorporating particular residues between chromophores. Furthermore, while designed proteins are intended to function like native proteins, these types of studies can test whether designed structures such as α-helices constructed by the heptad repeat method behave similarly to native proteins.

In this chapter of my thesis I will introduce some basic electron transfer theory and the notable previous studies on electron transfer rate in native and designed proteins. Then I will describe the design and synthesis of a ruthenium bipyridine chromophore that can be appended to exterior cysteine residues, as well as the characterization of electron transfer rates within α_3D

derivatives. I will demonstrate that the electron transfer rate to a bound metal is increased by the presence of intervening amino acids that can facilitate electron transfer and that we can characterize these transient electron relay amino acids. This work is still unpublished and was undertaken in collaboration with Dr. Ally Aukauloo and Dr. Annamaria Quaranta, who did the laser flash photolysis measurements and helped in fitting the transient spectra.

Photo-triggered electron transfer: designed proteins and $Ru(II)(bpy)_3^{2+}$

The vast majority of electron transfer work in designed proteins has focused on porphyrin cofactors and on the modification of peptides with chelating residues. Of particular interest are a short series of α-helical peptides that have been modified with non-natural amino acids to incorporate metals. Initial work used the complexation of ruthenium tris(bipyridine) and bipyridine derivatives to assemble three[1]- and four-helix bundles[2]. Cobalt tris-bipyridine complexes were used to assemble a three-helix bundle using Co(III) coordination.[3] Modification

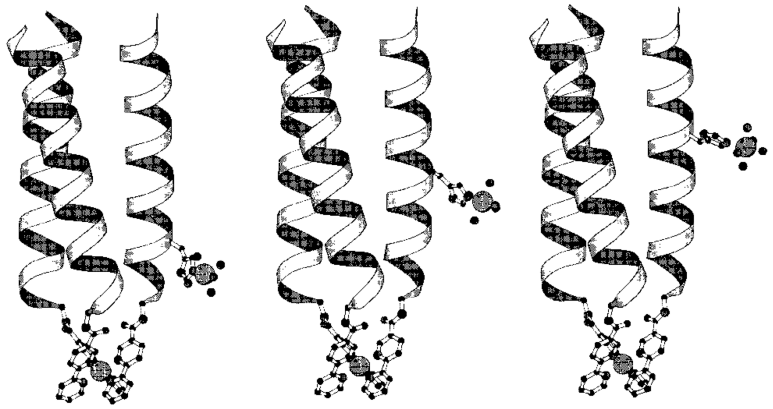


Figure 4.1 Cartoon depiction of construct used for studying electron transfer rate. The position of the $Ru(NH_3)_5$ can be varied by changing the position of the coordinating histidine residue. Electron transfer occurs between Ru(II) bipyridine and $Ru(NH_3)_5$. From ref. 3

of the peptides with viologen derivatives allowed for the measurement of electron transfer rates in different solvent conditions to understand how protein conformation affects electron transfer rates. By replacing the Co(III) coordination with Ru(II) and by the addition of histidine on the exterior of one of the helices, the authors were able to measure the electron transfer rate between the $Ru(II)(bpy)_3$ and the $Ru(NH_3)_5$ coordinated to the histidine.[3] By changing the displacement of the $Ru(NH_3)_5$ species (Figure 4.1)., the distance dependence of electron transfer could be investigated by pulse radiolysis and flash photolysis.[4]

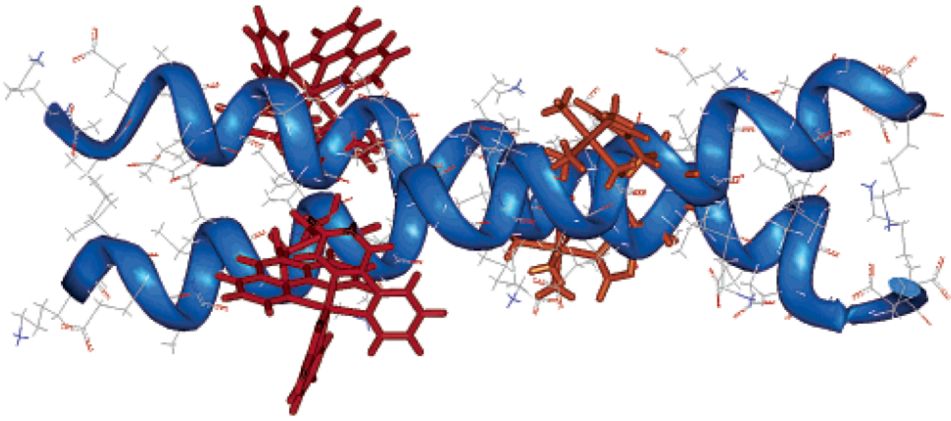


Figure 4.2 Computer rendering of intramolecular electron transfer construct comprised of a Ru(II) polypyridyl moiety conjugated to cysteine residues and a $Ru(NH_3)_5$ coordinated to histidines. From ref 8

A similar technique was used to investigate the electron transfer rate between $Ru(bpy)_3$ and $Ru(NH_3)_6$ across the helix-helix interface of a heterodimer. The observed k_{et} was found to be $380 \pm 80\ s^{-1}$ over 24 Å, which falls within the range reported for modified proteins over comparable distances.[5] Selective assembly of asymmetric four-helix bundles modified with ruthenium were created by chemoselective coupling to a cyclic peptide template.[6] Covalent attachment of Ru(II) bipyridine moieties at various distances to a heme coordinated to the interior of the bundle allowed measurement of through-helix electron transfer at distances of 27 and 23 Å. A technique was developed for the modification of exterior cysteine residues to attach polypyridine Ru(II) complexes.[7] When investigated for electron transfer, quenching of the fluorescence emission of the Ru(II) polypyridyl moiety was attributed to intramolecular electron transfer and the half-life τ, for the reaction was 72 ± 5 ns (Figure 4.2).[8] There was no difference between peptides with transposed positions of the two metal sites, which implies that there is no difference in the two directions for electron transfer. Overall, these studies showed that designed proteins are versatile scaffolds for the incorporation of redox-active complexes and that the scaffolds themselves display very similar electron transfer characteristics to native proteins.

These successes with relatively simple coordination complexes and cofactors encouraged researchers to design functional models for complicated natural processes like photosynthesis. By modularly combining an electron donor, a photosensitizer (porphyrin), an electron carrier, and a hydrogen evolution catalyst, photoinduced hydrogen evolution can be attained. A peptide dendrimer that binds multiple zinc mesoporphyrins (Zn MPs) was used as the photosensitizer.[9] The Zn MPs were coordinated axially by two histidine residues oriented in the interhelical space

between two adjacent helices. Triethanolamine acted as the electron donor and was photoelectrically oxidized by the Zn MP peptide dendrimers, which passed the electrons to an electron carrier, either methyl viologen or propylviologen sulfonate (PVS). Hydrogen evolution is accomplished by hydrogenase purified from *D. vulgaris*. Hydrogenase is purified in its oxidized, inactive form so it must first receive electrons from methyl viologen (or PVS) in order to catalyze proton reduction. Hydrogen production was on the order of 10^{-6} mol after two hours of irradiation as measured by gas chromatography.

There have been very few designed proteins subjected to the same rigor of study of their electron transfer rates as their native counterparts. Reaction centers involved in photosynthesis were some of the first proteins studied for which the rates of electron transfers were measured. The *de novo* designed proteins that mimic native proteins that have been studied for their electron transfer kinetics are much simpler models. In particular, of those that are more "native"-like, Dutton's group has designed a series of heme-binding proteins and have characterized light-activated electron transfer in this system.[10] The conversion of one of two heme binding sites to a site selective for Zn porphyrin or Zn chlorin allowed for the measurement of electron transfer rates within the protein. The presence of the heme increased the rate of decay decay of the Zn porphyrin excited state by ten-fold. Thus, it was found that the electron transfer from the Zn porphyrin to the heme over 19.3 Å had a rate constant of 1.1×10^3 s^{-1}, which is consistent with its driving force of 0.52 eV and reorganization energy of 0.9 eV. When a Zn chlorin is substituted for the Zn porphyrin, the rate increases by an order of magnitude, which the authors attribute to the greater size, and thus shorter distance between the two chromophores. It is worth noting that the transient kinetics for these processes were not fit directly; rather, the electron transfer rates were extracted from fits to the fluorescence quenching. In a different study from the Ghirlanda group, a *de novo* designed 4Fe-4S-binding protein was characterized with respect to its photo-driven intermolecular electron transfer kinetics.[11] When Zn porphyrin was added to the holo-peptide the triplet excited state was quenched by a factor of 14, but no signals for the oxidized Zn porphyrin were detected, so it is unclear whether electron transfer took place or quenching of the triplet state occurred by some other means.

Much of the work in this chapter has been undertaken under the additional supervision of my co-Chair, Prof. Ally Aukauloo and in collaboration with Dr. Winfried Leibl's group at the Commissariat à l'énergie atomique (CEA), Saclay. The work ongoing in their laboratories is

largely focused on charge separation and energy conversion using ruthenium chromophores. Of particular interest are the steps in PSII and methods towards functional analogues that might be used for energy conversion. Recently Dr. Aurélie Baron, a postdoc in Prof. Aukauloo's lab developed a series of modular "clickable" ruthenium bipyridine derivatives and demonstrated this as a convenient method to synthesize a series of ruthenium containing compounds.[12,13] Ruthenium trisbipyridine is a well-characterized and widely used photosensitizer with a long-lived excited state the capacity of being reduced or oxidized to generate either low or high potential intermediates. Ruthenium trisbipyridine complexes are characterized by strong MLCT transitions around 450 nm arising from t^2 to π transitions (Figure 4.3).[14]

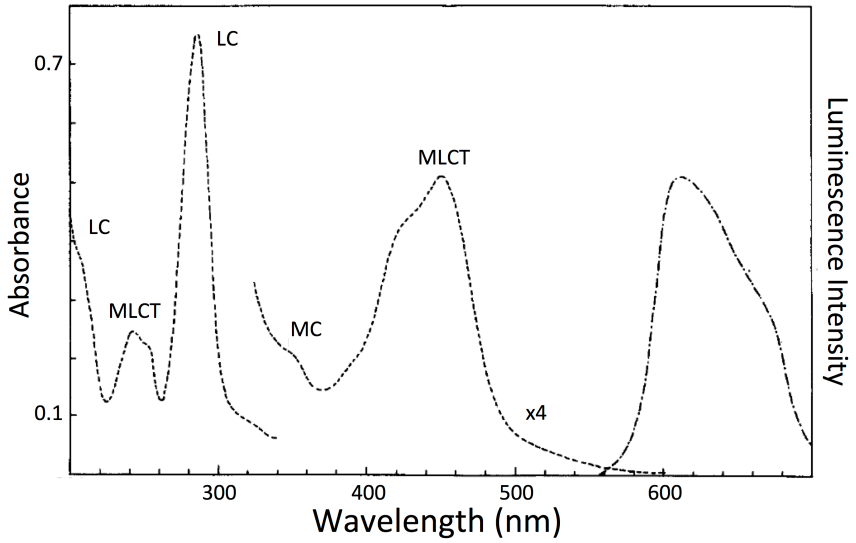


Figure 4.3 Absorbance (dotted) and luminescence (broken dashes) spectra for Ru(II)$(bpy)_3$ in aqueous solution at room temperature. Spectra from ref 14.

The luminescence emission spectrum of Ru(II)$(bpy)_3^{2+}$ is broad, and typically centered around 600 nm, depending on bipyridine substituents and solvent used. Ruthenium polypyridine complexes have accessible excited states that may be ligand-centered, MLCT, or metal-centered, and all decay to the lowest excited state (^{3}MLCT) with an efficiency of 1.[15] The excited state has oxidation and reduction potentials of -0.86 V and +0.84 (vs. SCE) and so can be said to be both a good oxidant and a good reductant. Thus, the excited state can decay by both reductive and oxidative quenching, in addition to energy transfer. The spectra of the excited and different redox states of $Ru(bpy)_3$ have been well-characterized as part of flash photolysis experiments.[14] The excited state of Ru(II)$(bpy)_3^{2+}$ has very little absorption at 450 nm, so this area of the spectrum bleaches from flash photolysis. Likewise, Ru(III)$(bpy)_3^{3+}$ has low absorbance at 450 nm. Flash

photolysis experiments are recorded as differences from the ground state spectrum, induced shining a laser into a particular band (for $Ru(II)(bpy)_3^{2+}$, 460 nm). Data may be collected either as a snapshot of the changes to a spectrum taken at a certain time after the laser pulse, or as the change in absorbance over time at a single wavelength. Consequently, examination of a kinetic trace measured at 450 nm may show phases that pertain to both the decay of the Ru* as well as the disappearance of Ru(III) as the Ru(III) recombines to regenerate the Ru(II) ground state. Alternatively, if Ru(I) is produced, a strong absorbance at 510 nm will evolve before the system recombines to yield the Ru(II) ground state. In this work, I have focused mostly on experimental procedures that use a reversible electron acceptor (e.g. methyl viologen) that oxidizes the Ru* to yield Ru(III), which can then react with the peptide species. Herein, I will detail my work on flash photolysis studies of inter- and intramolecular electron transfer in α_3D derivatives.

Materials and Methods

Synthesis:

All chemicals with the exception of propargyl maleimide (Kerafast) were acquired from Sigma-Aldrich or a subsidiary and used without prior purification, unless noted. Steps were based on published literature methods.[13] Please see Appendix I for NMR data.

2,2'-Bipyridine-N-oxide

To a solution of 2,2'-bipyridyl (3 g, 0.0188 mol, 1.0 eq.) in trifluoroacetic acid (15.0 mL, 21 g, 0.50 mol, 10 eq.) was added 30% hydrogen peroxide (3 mL, 0.95 g, 0.075 mol, 1.5 eq.). After stirring at room temperature for 4 h, the reaction mixture was neutralized by addition of aqueous 6N NaOH, then extracted with dichloromethane (4 x 50 mL). The combined organic layers were washed with aqueous saturated NaCl, dried over Na_2SO_4, filtered and concentrated to afford a colourless oil, which solidified under vacuum to a white solid (1.71g, 53%). ESI+ m/z=194.9 $[M+Na]^+$

4'-Nitro-2,2'-bipyridine-N'-oxide

2,2'-Bipyridine-N'-oxide (**1**) (3.0 g, 17.0 mmol, 1 eq.) was dissolved in concentrated sulphuric acid under stirring. A sodium nitrate (0.93 g, 18.7 mmol, 1.1 eq.) was added slowly and heated to 100°C overnight. Once cooled, the solution was poured into ice (150 g) and neutralized, in an

ice-bath, to pH 8 using 6N NaOH. The light yellow precipitate was filtered and washed with water. The solid was dissolved in methylene chloride, water was added and the mixture was extensively extracted with CH_2Cl_2. The combined organic layers were dried over sodium sulfate, filtered and concentrated to yield 0.8 g (7.0 mmol, 37 %) of 4'-nitro-2,2'-bipyridine-N'-oxide as a beige solid. *HRMS:* calculated 218.0560, found 218.0565 for $C_{10}H_8N_3O_3$. Calculated 240.0380, found 240.0381 for $C_{10}H_8N_3O_3Na$.

4'-Azido-2,2'-bipyridine-N-oxide

4'-Nitro-2,2'-bipyridine-N'-oxide (**2**) (0.8 g, 3.68 mmol, 1.0 eq.) and sodium azide (0.86 g, 13.25 mmol, 3.6 eq.) were suspended in anhydrous DMF (50 mL) and heated at 80°C for 48 hours under an argon atmosphere. After evaporation of the solvent, water was added and the mixture was extensively extracted with methylene chloride. The combined organic layers were dried over sodium sulfate, filtered and concentrated. This was then purified by flash chromatography on silica in methanol:dichloromethane (5:95). The product was isolated as the second spot, which was a cakey, yellow solid (0.5 g, 64%). *HRMS*: calculated 236.0543, found 236.0541 for $C_{10}H_7N_5ONa$.

4'-Azido-2,2'-bipyridine

4'-Azido-2,2'-bipyridine-N-oxide (**3)** (0.5 g, 2.36 mmol, 1.0 eq) was dissolved in dry dichloromethane (40 mL) and the solution was cooled to 0°C. Phosphorous tribromide (7.1 mL, 7.1 mmol, 3 eq) was added carefully. The reaction was allowed to stir under Ar(g) for 1 hour and then heated to reflux overnight. The solution was poured into ice and neutralized with 6N NaOH. The mixture was extensively extracted with CH_2Cl_2. The combined organic layers were dried over sodium sulfate, filtered and concentrated to yield 0.35 g (72.4 %). *ESI*+ m/z =220.1 $[M+Na]^+$

Bpy-mal: (1-((1-([2,2'-bipyridin]-4-yl)-1*H*-1,2,3-triazol-4-yl)methyl)-1*H*-pyrrole-2,5-dione)

4'-Azido-2,2'-bipyridine (**5**) (100 mg, 0.51 mmol, 1.0 eq.) was suspended in CH_2Cl_2) under an argon atmosphere. Propargyl maleimide (69 mg, 1.0 eq.) was added, followed by successive addition of water, sodium ascorbate (100 mg, 1 eq.) and copper sulfate pentahydrate (127 mg, 1 eq.). The reaction mixture was diluted with CH_2Cl_2/H_2O (1:1) and extensively extracted with

CH_2Cl_2. The organic layers were combined and washed with water, dried over sodium sulfate, filtered and concentrated to yield 90 mg of bpy-mal as a dark, orange oil (53%). *HRMS* calculated 333.1095 found 333.1098 for $C_{17}H_{13}N_6O_2$. Calculated 355.0914, found 355.0897 for $C_{17}H_{13}N_6O_2Na$.

Ru-bpymal

$[Ru(bpy)_2Cl_2]$ (120 mg, 0.25 mmol, 1.0 eq.) was reacted with silver nitrate (84 mg, 0.50 mmol, 2.0 eq.) in methanol (17.0 mL) for 3 hours at room temperature under an argon atmosphere. The suspension was filtered in order to remove the silver salt, and the filtrate was added to bpy-mal (90 mg, 0.28 mmol, 1.2 eq.). The solution was heated at reflux in the dark overnight under an argon atmosphere. The reaction mixture was allowed to go to room temperature and the solvent was evaporated. The remaining solid was re-dissolved in a minimum amount of methanol, and the desired compound was precipitated by drop wise addition of a saturated aqueous solution of ammonium hexafluorophosphate. The precipitate was filtered and dried under vacuum to yield 180 mg (0.17 mmol, 70 %) of the desired hexafluorophosphate ruthenium complex as a red solid. ESI+ m/z= 891.2 $[M+PF_6]^+$

Peptide Labeling:

L21C-based peptides were dissolved in 50 mM Tris and the concentration was taken using the A_{280}. 1.0 eq of Fe(II) stock (from Chapter 3) was bound to the peptide and the pH was adjusted to 8-8.5. The sample was then removed from the glovebox and allowed to air oxidize to confirm binding of the iron. Then, 1.1 eq of Ru-bpymal in DMSO were added and the solution was allowed to react for 30 min. The peptide was then purified via HPLC and the mass confirmed via ESI-MS.

α_3D**H$_3$** was prepared in phosphate buffer, pH 7.0. After the concentration was determined, 1.5 eq of Ru-bpymal in DMSO was added and allowed to react for 30 min. The peptide was purified via HPLC and the mass confirmed via ESI-MS.

UV-visible Spectroscopy:

Ground state absorption spectra were measured in a Specord spectrophotometer with 1 cm quartz cells.

Transient Absorption:

For transient absorption kinetics and spectral measurements in the time range 10 ns to 1 ms we used an Edinburgh Instruments LP920 Flash Photolysis Spectrometer system that incorporated a Continuum Surelite OPO for sample excitation (~7 ns pulse duration). The OPO was pumped by a Continuum Q-switched Nd:YAG laser operating at 355 nm. The LP920 system uses a 450 W Xenon arc lamp as the probe for the transient absorption kinetics measurements. Detection of the signal was performed either by a PMT or a water-cooled ICCD camera. The presented transient absorption spectra were typically the average of 20-50 measurements.

Results and Discussion

While there have been many studies of electron transfer and electron transfer rates in native proteins, relatively few have been done in designed proteins. Keeping in mind the debate over the role of the protein medium in modulating the electronic coupling between two redox sites, the extent to which designed proteins conform to the same behavior as native proteins is unknown. In a larger effort to design multi-site, multi-function proteins, understanding how to properly gate electron transfer between two redox sites is crucial. Our lab has undertaken a series of studies to characterize the electron transfer rate inter- and intramolecularly in the α_3D scaffold. My project has focused on such studies in Fe-α_3D**IV**-L21C.

Characterizing intramolecular processes in Fe-α_3D**IV**-L21C involves appending a photosensitizer to the protein, analogously to the studies that Gray's group has carried out on ruthenated proteins. We have focused on applying "click" chemistry techniques [12,13] to selectively attach a ruthenium chromophore to the peptide. Furthermore, we sought to develop a system that could be used with the GSGC series of peptides because there are several peptides in our lab with this terminal sequence that have been studied as a carbonic anhydrase model,[16] a blue copper model,[17] rubredoxin (Chapter 3), and nitrite reductase (Chapter 5). This led us to conceive of a ruthenium bipyridine chromophore modified with a maleimide moiety, which has been shown in many studies to react specifically with exterior cysteine residues.

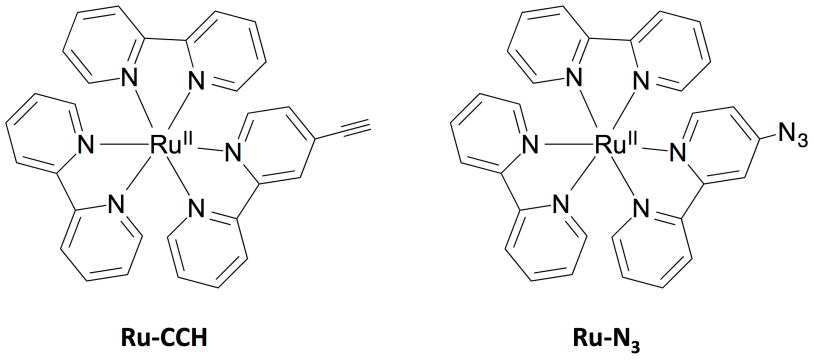


Figure 4.4 Clickable $Ru(bpy)_3$ complexes designed and synthesized by Baron *et al.*

"Click" chemistry is a term coined by Sharpless in 2001 to describe simple, stereospecific, high yielding reactions that can be used to combine different functional modules efficiently.[18] The classic reaction is the copper-catalyzed azide alkyne addition reaction, which uses ascorbate-generated Cu(I) to catalyze the cycloaddition of an azido group and an alkyne to

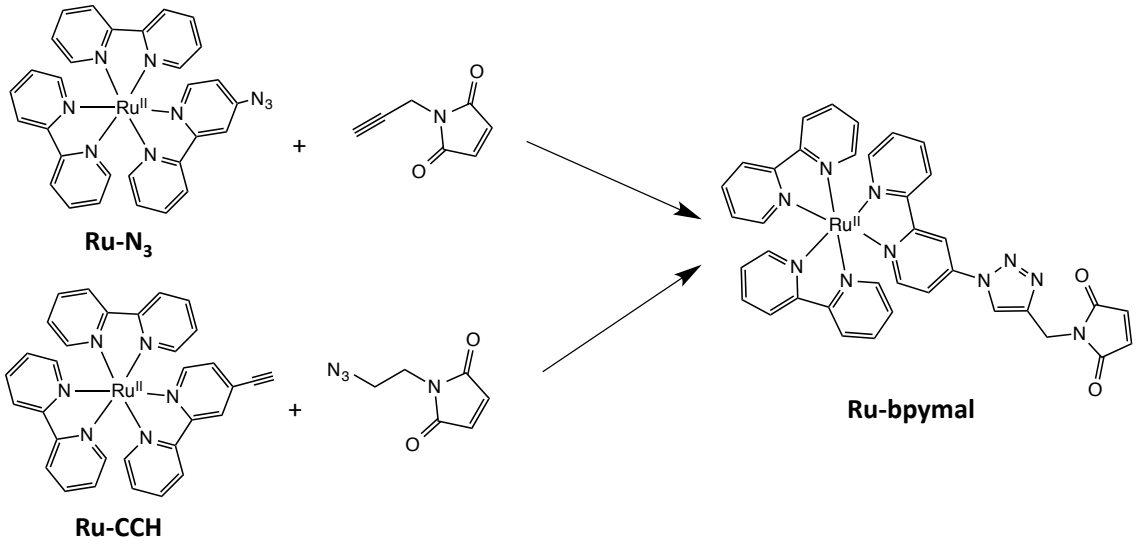


Figure 4.5 Original scheme to generate **Ru-bpymal** from small molecules at **Ru-CCH** and **Ru-N_3**

yield the 1,3-triazole product.[19,20] Based on the work by Dr. Aurélie Baron to generate Ru-CCH and Ru-N3 (Figure 4.4),[12,13]we sought to find a complementary maleimide link with the opposing azide or alkyne that upon reacting with an exterior cysteinyl residue would minimize the amount of distance placed between the chromophore and the peptide itself (and consequently the metal center). Thus a maleimide with a one or two carbon spacer seemed ideal, and my

original scheme involved synthesizing this small molecule, “clicking” it to the chromophore and then labeling the protein (Figure 4.5).

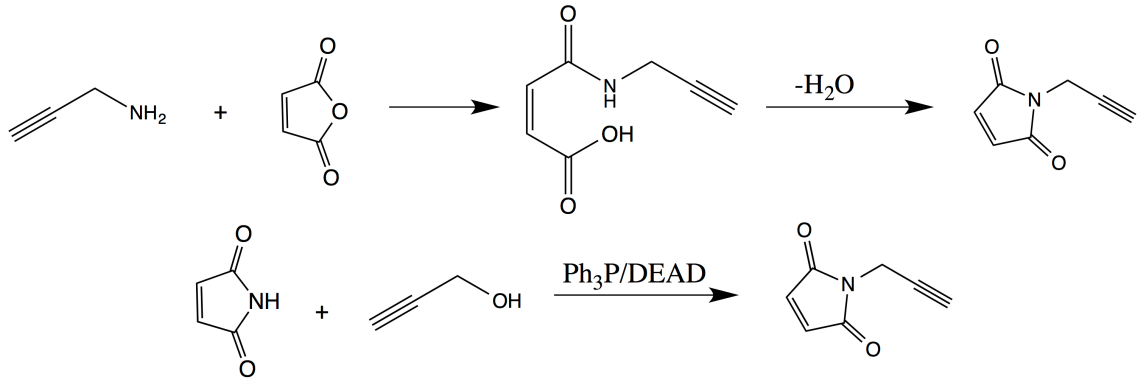


Figure 4.6 Two possible routes towards N-substituted maleimides, pictured here for synthesis of N-propargyl maleimide.

However, working with maleimide proved to be more difficult than predicted. Other groups have had success beginning with maleic anhydride, performing an amino displacement, and finally a high temperature step to close the ring (Figure 4.6).[21] Unfortunately, in my hands and in the hands of my collaborators, we were never able to close the ring with the small spacer groups we had appended. Furthermore, any purification of a maleimide group involving extensive flash chromatography always resulted in significant loss of product, as the maleimide

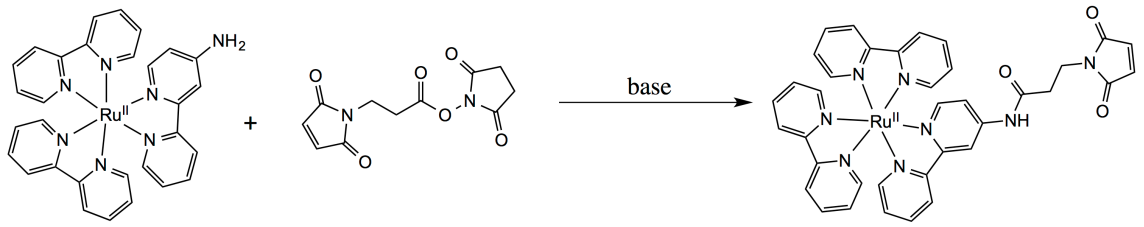


Figure 4.7 Scheme for synthesis with NHS-ester derivative of propargyl maleimide.

was hydrolyzed on the column. The highest yielding reaction to generate maleimide derivatives was a modified Mitsonobu reaction (Figure 4.6).[22] Unfortunately, the triphenylphosphine oxide produced as a by-product of this reaction reduces any azide groups that are present and is difficult to remove, creating problems in subsequent reactions where azide might be present, such as the click reaction itself. Propargyl maleimide is available commercially, but **Ru-N3** is difficult to make and so a maleimide with an azide is preferred for the coupling reaction. I also explored the idea of using N-hydroxysuccinamide esters of maleimide, which are highly

activated, good leaving groups, to react directly with a ruthenium derivative (Figure 4.7). Unfortunately, groups on the bipyridines, once they are coordinated to ruthenium, are highly deactivated so that performing this reaction, even in the presence of hydride, was unsuccessful.

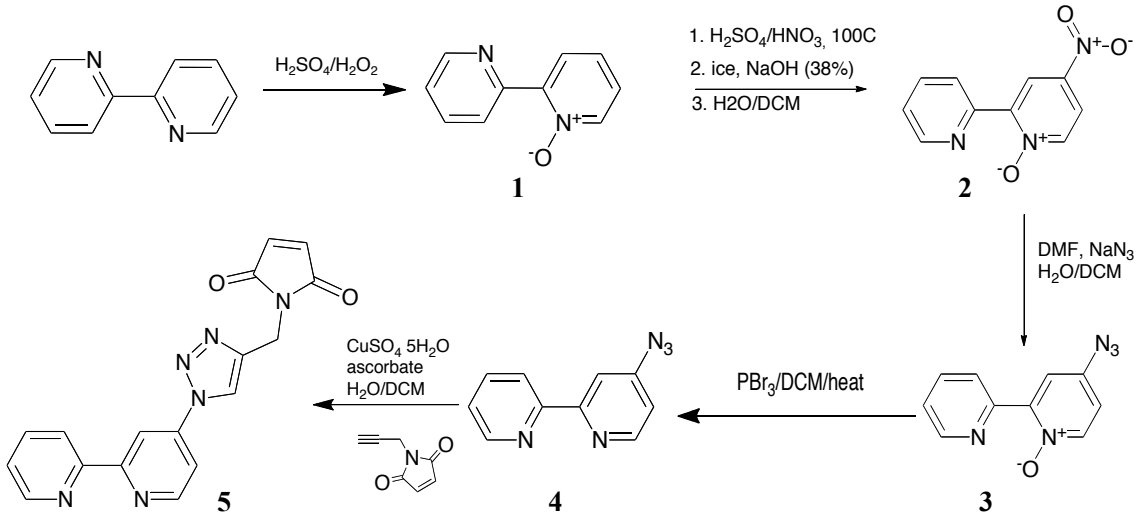


Figure 4.8 Scheme for synthesis of **bpy-mal** chromophore

Finally, after about a year and half of effort on the above methods, I chose to abandon the modular concept that originally inspired this synthesis and began to synthesize the bipyridine ligand from scratch (Figure 4.8). The most facile route proved to be synthesis of the 4-azido-2,2'-bipyridine (**4**) followed by a click reaction with commercial propargyl maleimide to yield **bpymal** (**5**). This intact chromophore was then used according to standard procedures to generate **Ru-bpymal**.[13] After the introduction of the maleimide group, I did not run columns so as not to lose any product. The peptides were labeled with the crude **Ru-bpymal** and purified via HPLC, which removes unsuccessful label and byproducts from the reactions.

Table 4.1 Sequences of peptides used in these studies

Peptide	Sequence
α_3D-GSGC	MGS WAEFKQR LAAIKTR LQAL GG SEAELAAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE LQAY RLN *GSGC*
α_3D**IV**-L21C-GSGA	MGS WAEFKQR LAAIKTR **C**QA**C** GG SEAE**C**AAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE **C**QAY RHN GSGA
α_3D**IV**-L21C-GSGC	MGS WAEFKQR LAAIKTR **C**QA**C** GG SEAE**C**AAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE **C**QAY RHN GSGC
α_3D**IV**-L21C-YF-GSGC	MGS WAEFKQR LAAIKTR **C**QA**C** GG SEAE**C**AAFEKE IAAFESE LQA***F*** KGKG NPE VEALRKE AAAIRDE **C**QA***F*** RHN GSGC
α_3D**H$_3$**-L21C-GSGC	MGS WAEFKQR LAAIKTR **H**QAL GG SEAE**H**AAFEKE IAAFESE LQAY KGKG NPE VEALRKE AAAIRDE **H**QAY RHN GSGC

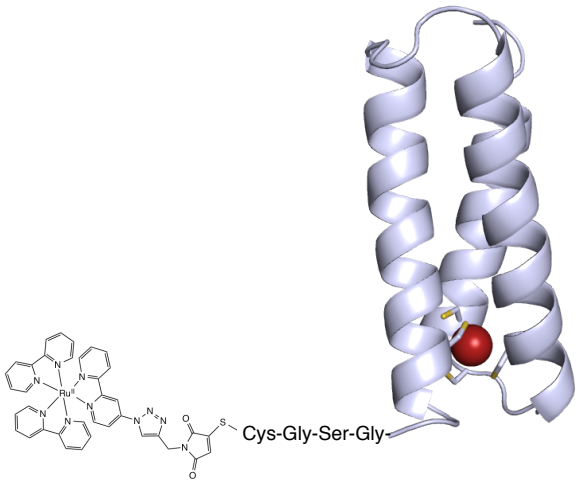


Figure 4.9 Model for **Rubpymal** conjugated to peptide. The maleimide moiety is conjugated to the free thiol of the cysteine in the GSGC tail.

For α_3D**IV**-L21C-GSGC and derivatives (Table 4.1), the preference of maleimide for solvent-exposed cysteines was exploited: iron was bound in the active site, which was confirmed by air oxidation to observe the colored form, followed by near-stoichiometric addition of **Ru-bpymal** (1.1 eq) and HPLC purification. For α_3D**H$_3$**-GSGC, **Ru-bpymal** was added super-stoichiometrically since there is only one cysteine in this construct (Table 4.1). Estimating the distance between the metal site and the ruthenium is difficult in the absence of crystal structures and the potential for flexibility in the intervening structure, but it is likely on the order of 15-20 Å, which is quite long for electron transfer that occurs by super exchange (Figure 4.9).

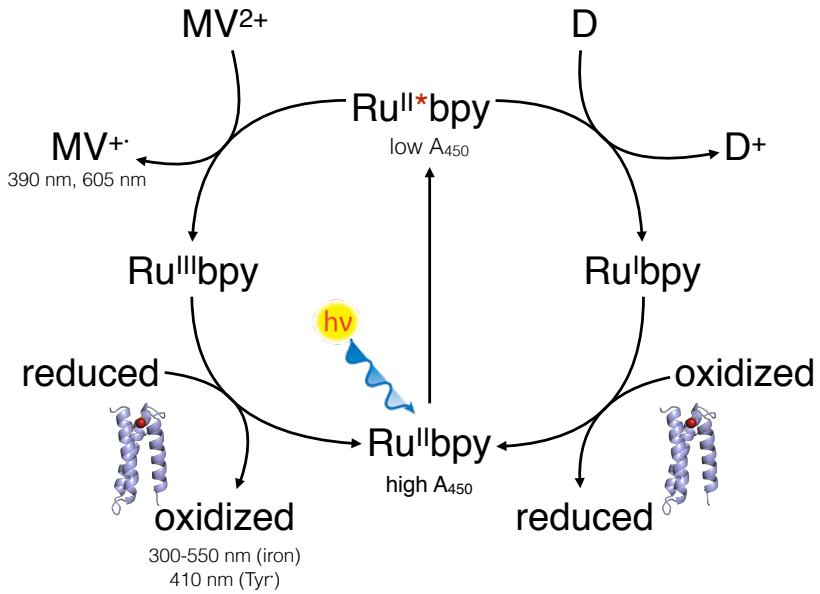


Figure 4.10 Possible routes for reaction with Ru(bpy)$_3$ and peptides. If the excited state does not react directly, an electron donor (D) can be used to generate Ru(I)(bpy)$_3$, which then reacts with oxidized peptide, to generate reduce peptide (*right)*. The reduced peptide can transfer an electron to D^+ to restore the initial state. Alternatively, an acceptor (here, methyl viologen) can be used to generate Ru(III)(bpy)$_3$, which then reacts with reduced peptide. Restoration of the initial state occurs when $MV^{+\bullet}$ reacts with oxidized peptide.

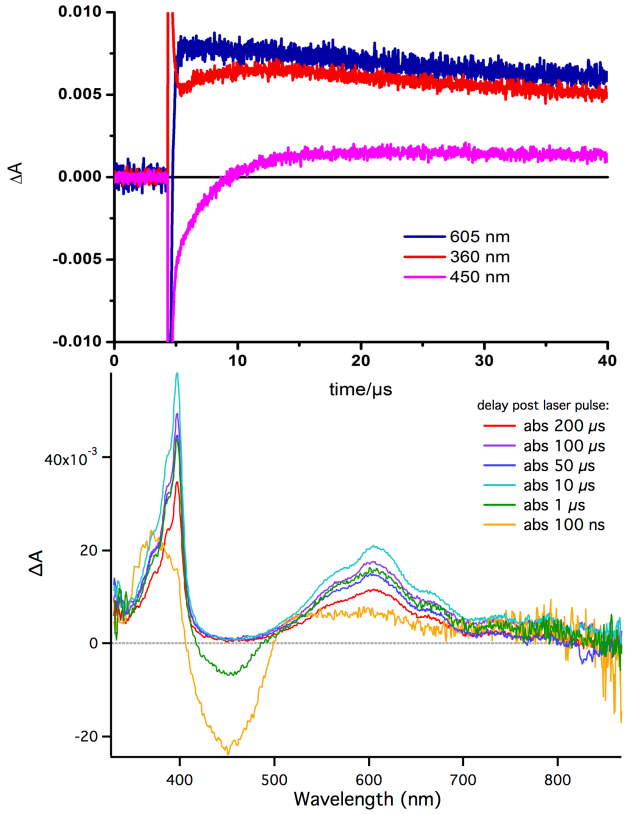


Figure 4.11 100 μM Fe-α_3D**IV**-L21C, 30 uM Ru(II)$(bpy)_3$, 20 mM MV, 100 mM Tris, pH 8.5, 460 nm excitation wavelength. *Top*: Plot of kinetic traces at 605 nm, 450 nm and 360 nm. A comparison of traces at 450 nm and 605 nm shows that the reduction of Ru(III), seen as decay of bleaching, occurs in a different kinetic regime from the decay of the signal at 650 nm. The traces at 360 nm and 450 nm share the same kinetics implying the same species is involved in both processes. *Bottom:* Camera spectra collected at various times post-laser pulse confirm that bleaching at 450 nm due to Ru(III) occurs on a different timescale from the decay of the reduced MV signal at 390 nm and 605 nm.

The intermolecular electron transfer kinetics of α_3D**IV**-L21C-GSGA with ruthenium trisbipyridine were studied by nanosecond laser flash photolysis. In reactions where the excited state of ruthenium trisbipyridine $(Ru^*(bpy)_3)^{2+}$ doesn't react directly with the metal center, a reversible donor or acceptor can be used to produce either $Ru(I)(bpy)_3^{+}$ or $Ru(III)(bpy)_3^{3+}$, which can then react with the metal center in the peptide (Figure 4.10, top). The most common reversible electron acceptors are methyl viologen and ruthenium hexaamine, while the most common reversible donor is ascorbate. Unfortunately, Fe-α_3D**IV**-L21C can react with ascorbate (Chapter 3) and ruthenium hexaamine directly, which precludes their use in these experiments. Thus, we proceeded to characterize the bimolecular electron transfer between α_3D**IV**-L21C-GSGA and $Ru(II)(bpy)_3^{2+}$ in the presence of methyl viologen (MV^{2+}). The reaction is characterized by biphasic kinetics and by tracking wavelengths that act as signatures for different species, we can extract information about the species involved in the reaction and their kinetics. Monitoring the absorbance at 450 nm gives information about the excited state, which appears as a bleaching effect. Methyl viologen has absorbance maxima at 390 nm and 605 nm in the

reduced state ($MV^{+\bullet}$), whereas the ferric form of the iron peptide absorbs strongly in the region of 300-550 nm (Chapter 3).

When the peptide and the Ru(II)$(bpy)_3$ are not linked, the Ru(III) is reduced back to Ru(II) before $MV^{+\bullet}$ re-oxidizes, as evidence by the recovery of absorbance at 450 nm as compared to the disappearance of absorbance at 605 nm (Figure 4.11, top). This implies that another species is reducing the Ru(III) before final charge recombination returns the system to

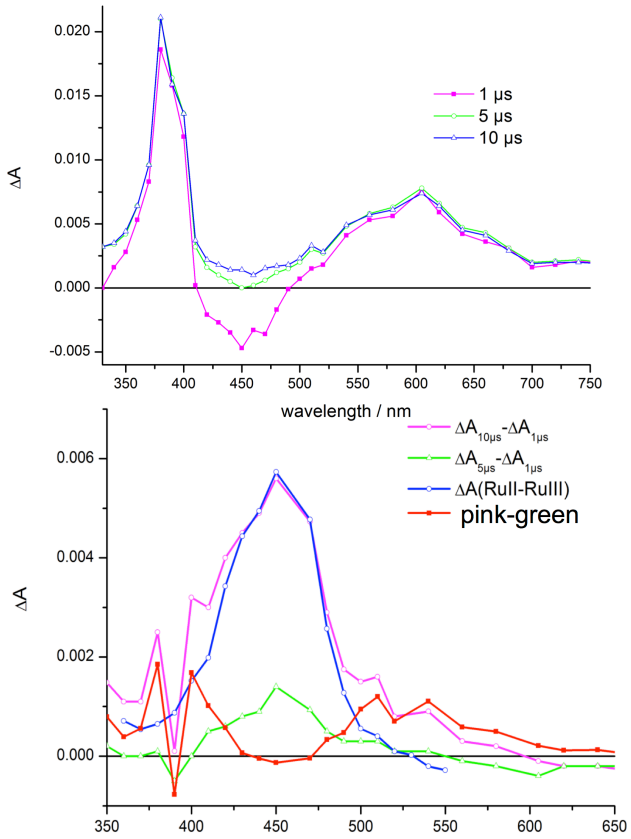


Figure 4.12 100 µM Fe-α_3D**IV**-L21C, 30 uM Ru(II)$(bpy)_3$, 20 mM MV, 100 mM Tris, pH 8.5, 460 nm excitation wavelength. *Top* Spectra can be calculated based on fits to kinetic traces at a series of wavelengths to yield the transient spectrum at a given time delay, in this case, 1 µs (pink) 5 µs (green), and 10 µs (blue). *Bottom:* Taking the difference of the transient spectra can be helpful in isolating which species are responsible for absorbances. The difference between 1 and 5 µs and between 5 and 10 µs are pictured in green and pink, respectively. The blue spectrum is the contribution from the difference between Ru(II) and Ru(III) only. The red trace is the difference between 5 µs and 10 µs corrected by the contribution from Ru(II)-Ru(III) to try to isolate the absorbing species, in this case, the iron. We take the point at 390 nm as being unreliable because it is a λ_{max} for $MV^{+\bullet}$. Taking into account uncertainties where other species absorb strongly, the red spectrum represents the oxidized iron species.

the ground state (Figure 4.10). This can be more easily observed in camera spectra, which take snapshots of the whole spectrum at specific delays post-laser flash. The bleaching at 450 nm occurs within 100 ns and disappears within 10 µs while the peaks at 390 nm and 605 nm persist for hundreds of microseconds (Figure 4.11, bottom). Furthermore, traces at 450 nm and at other wavelengths where iron absorbs show a growth of absorbance at 10 µs post-laser flash that matches the kinetics of the disappearance of Ru(III)(Figure 4.11, top). To determine the species involved in this electron transfer, one can work out a differential spectrum based on transient spectra at different times by the results of the exponential fit to the kinetic traces at a particular time. Unfortunately, in this case, this has proved to be very challenging, as the absorbance of the

Fe(III)-α_3D**IV**-L21C is in the same region as strong absorptions from both the ruthenium trisbipyridine and reduced methyl viologen. However, the spectra do not correspond to a simple recombination event between Ru(III)$(bpy)_3^{3+}$ and $MV^{+\bullet}$, which would show concurrent depletion of the Ru(III) signal at 450 nm and the $MV^{+\bullet}$ at 605 nm with no positive absorbance observed, as is seen the 10 μs trace (Figure 4.12, top, blue). Difference spectra of these transient kinetic traces followed by correction for the contribution of the difference between Ru(II) and Ru(III) (blue trace) yields a spectrum that has apparent λ_{max} at ~380 nm and ~500 nm (red trace) (Figure 4.12). It is important to note firstly that the point at 390 nm is anomalous due to this being the λ_{max} of $MV^{+\bullet}$ and thus small differences in absorbance are hard to measure; and secondly, the kinetic traces are taken every 10 nm so fine grain resolution of the spectra is impossible. Nevertheless, this spectrum is similar to that of Fe(III)-α_3D**IV**-L21C, especially taking into account the large nearby absorbances of $MV^{\bullet+}$ and Ru(II). It should be noted that the apparent λ_{max} from this calculated spectrum appear to be at 380 nm and 510 nm, versus 360 nm and 492 nm as measured directly on the peptide, which could be an artifact from the high molar extinction coefficients for absorbances of the ruthenium (ca. 14500 $M^{-1}cm^{-1}$) and methyl viologen (ca. 14000 $M^{-1}cm^{-1}$).[14,23] However, an alternative species that could be represented by this spectrum is oxidized tyrosine, which has a λ_{max} of 400 nm.[24] However, oxidized tyrosine has no significant absorption above 500 nm and while this construct has one Tyr near the iron site, there is no guarantee that a bimolecular reaction would proceed via this path. Thus, we conclude that these absorption characteristics are actually due to the Fe(III) peptide. Using the concentration of peptide in the experiment, the rate constant for this electron transfer is 4 x 10^9 $M^{-1}s^{-1}$.

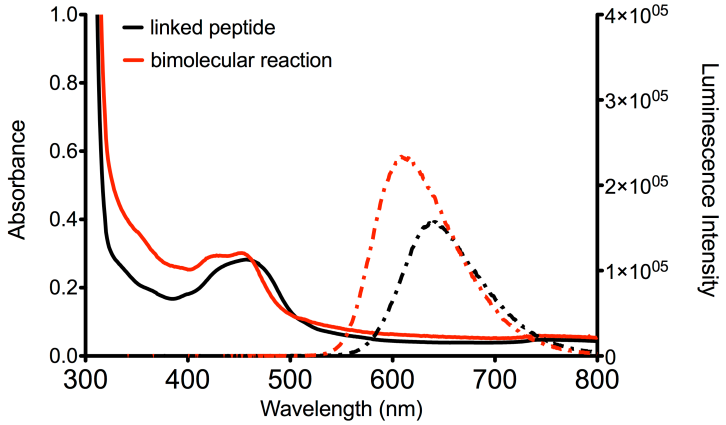


Figure 4.13 Comparison of **Ru-bpymal** appended to a peptide (black) and $Ru(bpy)_3$ with peptide in bimolecular reaction (red). The behavior is similar but the absorbance and luminescence are shifted to slightly lower energy. The emission intensities should not be directly compared as they were taken at different delays post-excitation.

To investigate intramolecular electron transfer, **Ru-bpymal**$^{2+}$ was appended to the terminal cysteine in Fe-α_3D**IV**-L21C-GSGC and the electron transfer rates between iron and ruthenium were measured. This adduct results in slight differences in the ground state absorption and the fluorescence emission spectra (Figure 4.13). In the absence of an external electron acceptor or donor there is slight quenching of the ruthenium excited state, which results in a shorter luminescence lifetime than when the peptide is not linked, as in the previous experiment (440 ns vs. 535 ns)(Figure 4.14). However, no particular absorption was observed as a result of

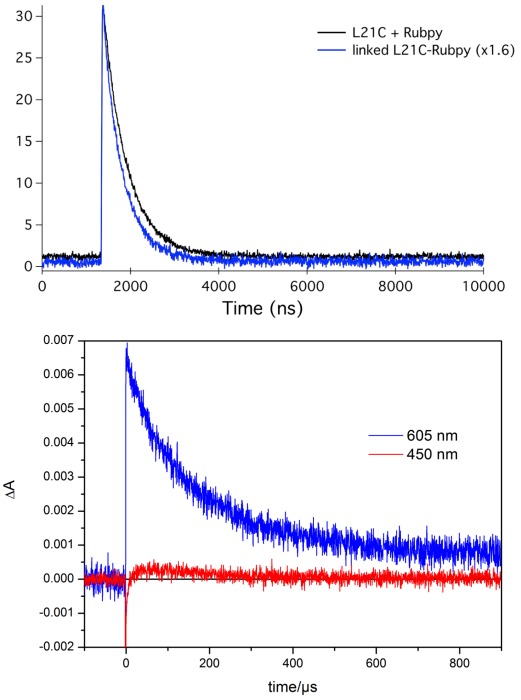


Figure 4.14 35 μM L21C-Rubpymal, 100 mM Tris, pH 8.5, 460 nm excitation wavelength *top:* The emission of **Ru-bpymal** is slightly quenched in the presence of the peptide and iron (435 ns vs. 535 ns). *Bottom:* methyl viologen is used to generate Ru(III). Comparison of the kinetics at 450 nm and 605 nm shows that negative signal at 450 nm disappears faster than the reduced methyl viologen at 605 nm. Furthermore, this process that reduces Ru(III) also produces a small positive absorbance, which is visible at ~25 μs.

this low level of quenching, so methyl viologen was again used to drive the oxidation of [Ru(II)*(bpy)2(bpy-mal)]$^{2+}$ to [Ru(III)(bpy)2(bpy-mal)]$^{3+}$. Once again, the Ru(III) is reduced faster than the methyl viologen is oxidized, suggesting an intramolecular electron transfer acts to reduce the Ru(III) (Figure 4.14). The apparent rate constant for this reaction was found to be 2 x 10^5 s^{-1}. There were similar difficulties in isolating the spectrum of the electron transfer species as in the intermolecular reaction, with the added challenge of not being able to compensate at all for the differences in molar extinction coefficients between the ruthenium(II) and iron(III) since the two are covalently linked (Figure 4.13). As ruthenium(II) has a molar extinction coefficient of 14,000 $M^{-1}s^{-1}$ and iron(III) has a molar extinction coefficient of 3000 $M^{-1}cm^{-1}$, signal can still be detected for the iron(III) in the bimolecular system by running the experiment with an excess of

iron-peptide. However, in the covalently linked system this is not the case as the two concentrations are directly correlated.

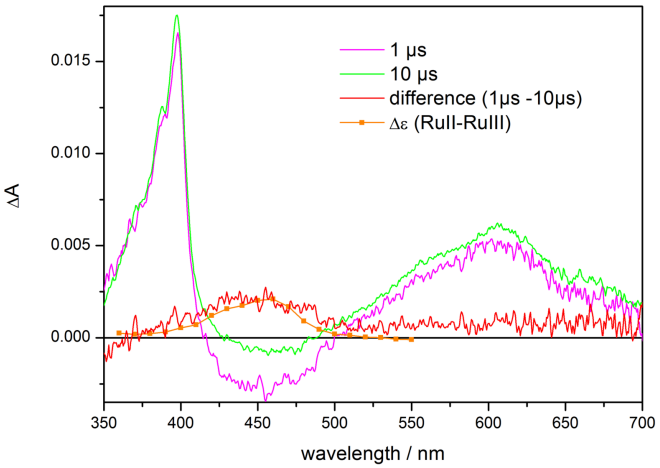


Figure 4.15 35 µM L21C-Rubpymal, 100 mM Tris, pH 8.5, 460 nm excitation wavelength Difference spectra of 1 and 10 µs traces (red) show some diffuse absorption that is mostly due to Ru(II)-Ru(III) (orange).

To understand what might be happening in this linked system better, we also examined the system in the absence of iron. Without iron coordinating the cysteines, methyl viologen forms a complex with the peptide, which we have observed in both the bimolecular and in this

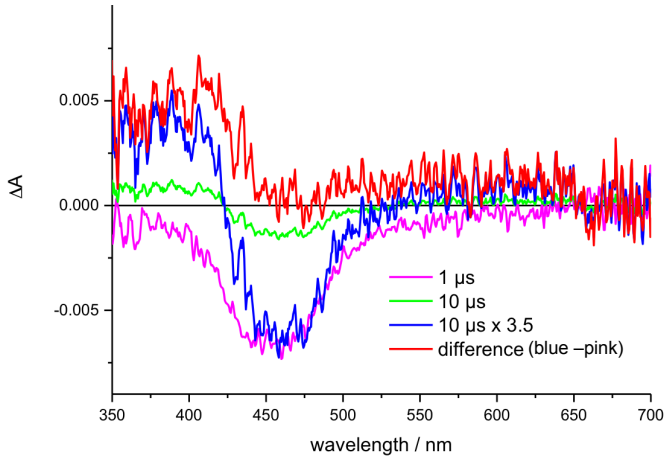


Figure 4.16 Simple examination of the camera spectra reveal an apparent absorbance at 400 nm (green), which is not obscured by the absorbance of an acceptor. Difference spectra (red) of the transient spectra at 1 µs (pink) and 10 µs (blue) normalized to the Ru(III) signal show a clear absorbance with an apparent λ_{max} of 410 nm.

linked system. Thus, we used ruthenium hexaamine as the electron acceptor for these measurements, which has the added advantage of being colorless, so the absorbance features of the peptide itself are less obscured. Examining camera spectra at different delay times shows that after 10 µs, a new feature can be observed between 350 nm and 400 nm (Figure 4.15).
By creating a difference of the transient spectra at 10 µs and 1 µs, we can better isolate the signal and see that the signal is indeed localized between 350 nm and 400 nm (Figure 4.16). The apparent rate constant for the formation of this signal is 4 x 10^5 s^{-1} (Table 4.2), while the decay of the system back to the ground state is on the order of 10^4 s^{-1}. The signal we observe is

reminiscent of spectra that have been observed for tyrosine radicals, which have a λ_{max} of 410 nm (Figure 4.17).[24,25]

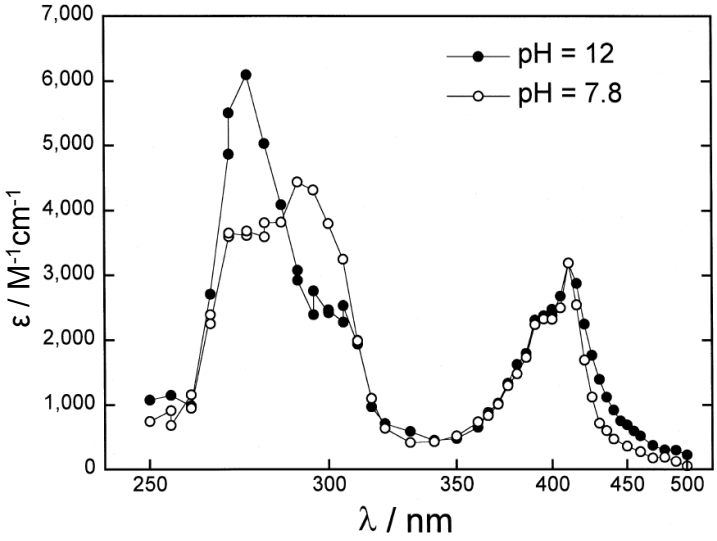


Figure 4.17 Transient absorption spectra of tyrosine radicals, $Y^{\bullet-}$, at pH 12 and 7.8 show a maximum in the visible region at 410 nm. From ref 24.

Given the position of Tyr70, we propose that Tyr70 serves as a relay for an electron to hop from the Fe(II) to the Ru(III) in this system, which also accounts for the similar rates observed in the apo- and iron peptide. The similarity in the rate constants implies that the rate-limiting step for this process is oxidation of the tyrosine. This model also agrees with theory, which states that most electron transfer that occur by super exchange (in a single step) occur at less than 14 Å, while the metal-to-metal distance in Fe-α_3D**IV**-L21C-**Rubpymal** is likely around 20 Å. Thus, electron transfer via a relay amino acid to shorten the distance is highly likely.

To test this hypothesis, I designed and expressed two more peptides, termed α_3D**IV**-L21C-YF-GSGC (hereon YF) and α_3D-GSGC (Table 4.1). These control for incorporating the metal binding site, but with all tyrosines in the peptide removed (YF), and not having the binding site at all but with tyrosine present (α_3D-GSGC). In the absence of the tyrosine, intramolecular electron transfer in YF-**Rubpyma**l was very difficult to analyze. At any wavelength, the kinetic traces do not follow simple exponential behavior. A global fit to the recorded wavelengths obtains a three-component exponential decay (Figure 4.18). The three components are characterized by half lives of (11 ± 1) μs, (141 ± 4) μs, and (810 ± 90) μs; the formation of the new absorbing species is associated with τ_2, 141 ± 4 μs. The fast component at 11 μs does not correspond to spectral changes as compared to the initial charge separation, so it is difficult to attribute to a specific reaction. The τ_2 component is assigned to the intramolecular electron

transfer, which is 100X slower than when the tyrosine is present. Finally, τ_3 corresponds to charge recombination with reduced methyl viologen to restore the system to the initial state.

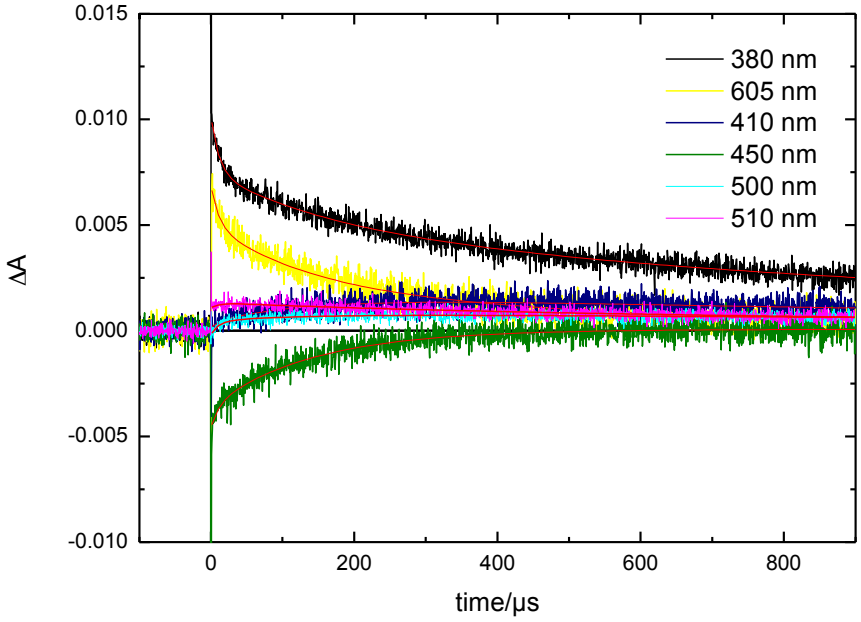


Figure 4.18 35 μM YF-Rubpymal, 20 mM MV, 50 mM Tris, pH 8.5, 460 nm excitation wavelength. Global fits over various wavelengths result in three-component exponential kinetics. At wavelengths where $MV^{+\bullet}$ and Ru(III) have small contributions (410, 500, and 510 nm) the formation of a new signal is evident.

The intramolecular electron transfer has a k_{app} of 7 x 10^3 s^{-1}, which is two orders of magnitude slower than when the tyrosine is there. Data collected on a shorter time scale, such as 40 μs instead of 1 ms, could help resolve the decay of the shorter half-life species at this wavelength. Clearly, the behavior of Fe-YF-**Rubpymal** is significantly different from that of Fe-α_3D**IV**-L21C-**Rubpymal** such that there is no reaction with the iron or it occurs with a low probability and thus competes with MV to reduce Ru(III). These findings imply that the presence of the tyrosine is vital for productive intramolecular electron transfer over the approximate 20 Å between the redox centers in this system.

Transient absorption measurements with α_3D-**Rubpymal**, which does not have a metal binding site, but does retain Tyr70 were measured to assess the role of the tyrosine. Contrary to Fe-YF-**Rubpymal**, Ru(III) is reduced extremely quickly in this system, on the order of 1 µs.

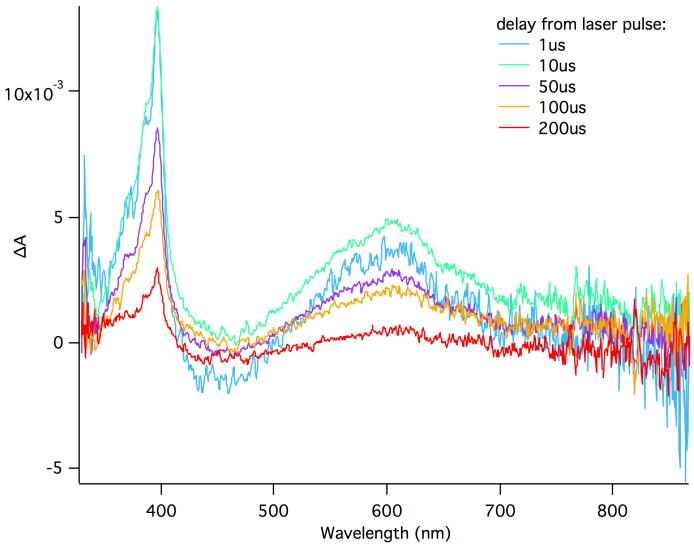


Figure 4.19 35 µM α_3D-Rubpymal, 20 mM MV, 50 mM Tris, pH 8.5, 460 nm excitation wavelength. Camera species of α_3D-**Rubpymal** over a range of time delay reveals that the Ru(III) reacts more quickly than the reduced methyl viologen

Camera spectra from this species show that the bleaching at 450 nm disappears very quickly, while absorbance from $MV^{+\bullet}$ remains at 605 and 390 nm (Figure 4.19, light blue trace). The kinetics at the relative λ_{max} of Ru(II)-Ru(III) (450 nm) and methyl viologen 605 nm can be used to calculate the relative amounts of each species formed in the reaction, knowing the molar extinction coefficients of the transient species.

A global fit analysis of the measured wavelengths at short timescales yields a biphasic growth with τ_1 = 80 ns and τ_2 = 750 ns (Figure 4.20, left). The first half-life is associated with the initial charge separation, while the second is associated with the intramolecular electron transfer and gives absorption at 405 nm and 410 nm, which would not have positive absorption if Ru(III) were the only species involved in this process. Absorption at these wavelengths is consistent with a phenoxyl radical (*vide infra*). The apparent rate corresponding to this intramolecular electron transfer is 1.3 x 10^6 s^{-1} (Table 4.2). A global fit analysis is not possible for longer timescales to measure the decay. Kinetics at 405 nm and 410 nm must be fit to a biexponential function in which τ_1 = 1500 ns and τ_2 = 306 µs, where the shorter half-life is consistent with the intramolecular electron transfer (Figure 4.20, right). The decay of the $MV^{+\bullet}$ signal at 605 nm occurs with a $\tau_{1/2}$ of 128 µs. This suggests that the tyrosine radical is decaying more slowly than the reduced MV, which could be due to some screening effect of the peptide or some stable

photo-product is accumulated in solution. Storing this sample overnight and remeasuring the following day results in a much longer-lived decay at 605 nm (~500 μs), which could be due to the extended contact that the methyl viologen has with the peptide such that the methyl viologen is also partially shielded from recombination or that the methyl viologen then interacts more closely with the tyrosine. However, the absorbance of the methyl viologen partially obscures the expected signal at 410 nm. To resolve some of the questions surrounding the recombination,

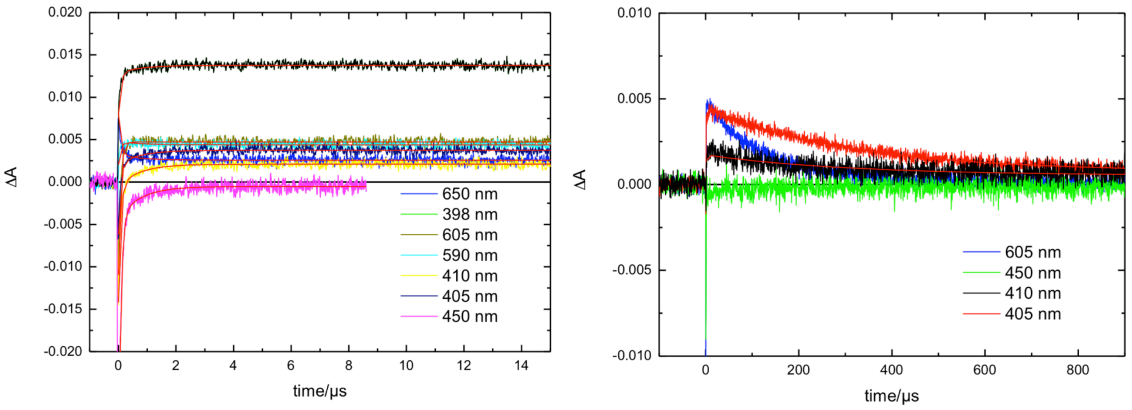


Figure 4.20 *left* A global fit analysis of the short timescales reveals a biphasic process with half-lives τ_1=80 ns and τ_2=750 ns. *Right*, at longer timescales the decay of the tyrosine radical and reduced methyl viologen can be observed, but cannot be fit globally. The methyl viologen (blue) decays more quickly than the radical (black and red).

these data should be re-collected with ruthenium hexaamine as the electron acceptor, as was done for apo-α_3D**IV**-L21C-**Rubpymal**. The combined fast phase observed in α3D-**Rubpymal** and very slow reaction in Fe-YF-**Rubpymal** suggests that tyrosine does play a role in mediating the electron transfer to the metal site.

To characterize intramolecular electron transfer in a peptide without a thiolate binding site, I appended **Ru-bpymal** to α_3DH_3, which contains histidines instead of cysteines at the same layer as α_3D**IV**. This peptide has previously been investigated with Zn(II) bound and shown to be an active carbonic anhydrase mimic.[16] In Chapter 5 of my thesis I will detail my work in characterizing its copper binding and nitrite reductase activity, but here I will detail the photophysical measurements performed on this peptide. A major problem in characterizing copper bound to this peptide is the extremely low molar extinction coefficient that copper has in type II sites. Cu(II) has d-d transitions that are on the order of 1-200 $M^{-1}cm^{-1}$, which is much too weak to be seen by this technique. However, this is advantageous in terms of investigating the

role of the tyrosine and whether it acts as an electron relay in this system as this peptide still possesses the tyrosine, but can be analyzed both in its apo- form or with copper bound. In the absence of an external electron acceptor, this linked peptide behaves very similarly to apo-α_3D**IV**-L21C-**Rubpymal**, with a luminescence emission wavelength of 640 nm and a half-life of 400 ns. The kinetics at 450 nm also decay at the same rate as the luminescence emission, suggesting that there is not efficient charge separation and the species responsible for the slight quenching of the excited state cannot be detected under these conditions. Another possibility is that energy transfer occurs to quench the excited state. To increase the efficiency of charge separation and creation of Ru(III), methyl viologen was added. A comparison of camera spectra show that all of the Ru(III) has decayed by 6 µs, while the reduced methyl viologen is still present (Figure 4.21, blue trace).

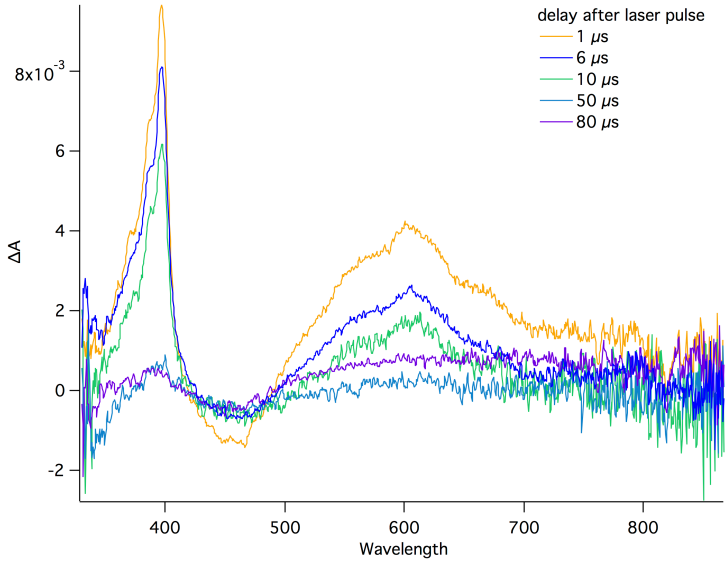


Figure 4.21 30 µM α_3DH_3-**Rubpymal**, 20 mM MV, 100 mM phosphate buffer, pH 6.0, 460 nm excitation wavelength. Camera spectra of α_3D**H_3-Rubpymal** show that all of the Ru(III) has decayed by 10 µs (green), while some reduced methyl viologen is still present.

Global fitting, as in the case of α_3D-**Rubpymal**, gives a fast and a slow phase in which the $\tau_{1/2}$ of the fast phase is 100 ns and 1800 ns for the slow phase (Figure 4.22). This is similar to what was observed in apo- α_3D**IV**-L21C-**Rubpymal** and in the longer wavelength spectra of α_3D-**Rubpymal**. The decay of the methyl viologen takes place on a much longer timescale, at about 700 µs. As was observed for α_3D-**Rubpymal**, the kinetics wavelengths around 410 nm fit to this faster process and result in positive absorbance, which is consistent with a phenoxyl radical. The apparent rate for this process is 5.6 x 10^5 s^{-1} (Table 4.2). When copper is present in solution, the fast, intramolecular electron transfer happens on the same timescale while the recombination with methyl viologen happens with a $\tau_{1/2}$ of 21 µs, rather than 700 µs. This can be seen in the

camera spectra, as the peaks at 390 nm and 605 nm noticeably decrease in the delay after the laser pulse (data not shown). The initial species produced has the same characteristics as was previously seen for the apo- and tyrosine-containing peptides, namely, there are clear absorption characteristics around 410 nm.

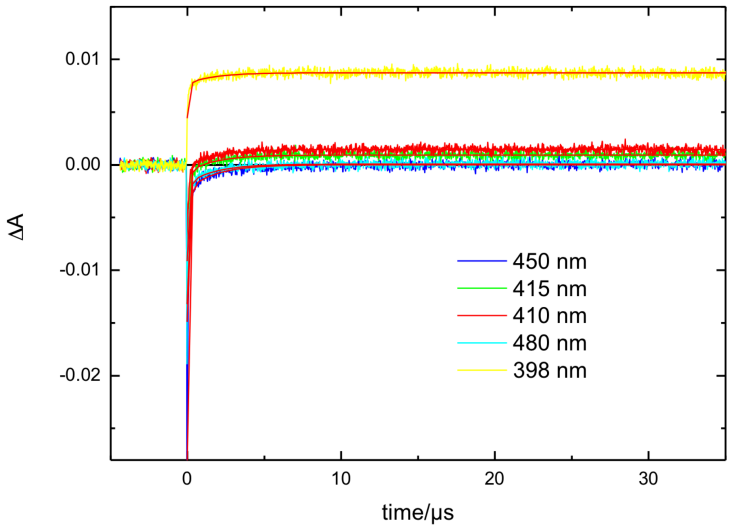


Figure 4.22 The result of the global fit at single wavelengths of α_3D**H$_3$-Rubpymal** in the presence of methyl viologen yields two phases τ_1=100 ns and τ_2= 1800 ns.

The kinetics at 410 nm and 415 nm can be fit to a biexponential function where the growth from the intramolecular electron transfer occurs at τ= 2.5 μs and the decay occurs at 14 μs (Figure 4.23). The significant positive absorption seen here cannot be attributed to either the MV or ruthenium species and is consistent with tyrosine radicals as seen in the literature and previously in other constructs. The shorter length of decay of methyl viologen when copper is present suggests that the metal site may oxidize the methyl viologen, increasing the rate of decay by over an order of magnitude slower.

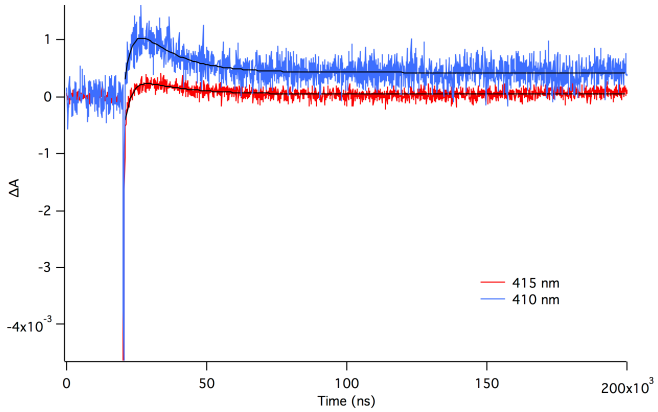


Figure 4.23 Single wavelength kinetics for Cu-α_3D**H$_3$-Rubpymal** show a strong absorbance followed by decay at wavelengths that are typical of tyrosine radicals.

Table 4.2 Apparent rate constant for intramolecular electron transfer

Peptide[a]	k_{app} (s^{-1})	Species Formed
α_3D	1.3 x 10^6	Tyrosine radical
Fe-α_3D**IV**-L21C	2.0 x 10^5	Tyrosine radical, Fe(III)
Fe-**YF**	7.0 x 10^3	Fe(III)
α_3D**IV**-L21C	4.0 x 10^5	Tyrosine radical
α_3D**H$_3$**	5.6 x 10^5	Tyrosine radical
Cu-α_3D**H$_3$**	5.0 x 10^5	Tyrosine radical

[a]**Ru-bpymal** omitted for clarity

Across the series of peptides measured for intramolecular electron transfer all have the same k_{app} for intramolecular electron transfer, with the exception of α_3D and YF. In the case of the YF peptide, the removal of the tyrosines greatly affected the electron transfer capabilities to the point where no electron transfer is observed on the same timescale as these other systems. This suggests that the tyrosine is required for electron transfer to occur. On the other hand, α_3D exhibits a slightly greater k_{app} for intramolecular electron transfer than the other peptides studied, despite not having a metal bound. All of the peptides studied, with the exception of α_3D, have mutations from leucine to metal-binding residues towards the C terminus. As detailed earlier in Chapters 1 and 2, these changes were made in an area of the protein that was determined to be amenable to mutation. However, these changes may have affected the stability of peptides with mutations in the same positions as α_3D**IV**, which can affect the dynamics of the protein. This difference in protein dynamics likely has an affect on the reorganization energy of the reaction. A smaller reorganization energy resulting from less dynamic movement through the protein could easily explain this difference of 2-3X k_{app} between the parent peptide, α_3D, and its derivatives.

Electron transfer from tyrosine to the metal center is likely very rapid, and separating the electron transfer from the tyrosine to the Ru(III) from the electron transfer from the reduced metal to the tyrosine may be impossible. I estimate the separation from the tyrosine to the metal site is on the order of 4-5Å. In the case of the iron system, overlapping absorbance between of the tyrosine (λ_{max} = 410 nm) and the iron (λ_{max}= 360 nm, 490 nm) makes distinguishing these two species difficult, especially when they both fall in the same time regime. In the copper system, very low molar extinction coefficients at characteristic λ_{max} make specific detection of the oxidized copper species impossible. However, insight into the series of electron transfers can still be gleaned from the information available from these studies. We propose that when metal is

bound, the rate-limiting electron transfer occurs from the tyrosine to the Ru(III), followed by rapid electron transfer from the metal center (Figure 4.23). At this point, reduced methyl viologen, reduces the metal bringing the system back to the ground state. In the case of the iron peptide, the calculated spectrum is quite a bit different than the spectrum of tyrosine alone and has absorbance at wavelengths greater that 500 nm, which strongly suggests that the Fe is involved. Furthermore, in this system, methyl viologen is reoxidized within 20 μs. When iron is not bound to the protein, we observe the spectrum of the tyrosine, but the τ associated with reoxidation of methyl viologen is 150 μs. The lengthening of the timescale for methyl viologen reoxidation in the absence of iron suggests that the system returns to the ground state via a different process. A parallel comparison can be made in the case of copper bound to α_3D**H$_3$** as when copper is bound, the reoxidation of methyl viologen also occurs on the timescale of ~ 20 μs, which lengthens to ~200 μs when copper is not present. Therefore, in peptides without a metal, tyrosine is the final location of the hole generated by reduction of Ru(III), while when a metal is present, a hopping mechanism takes place.

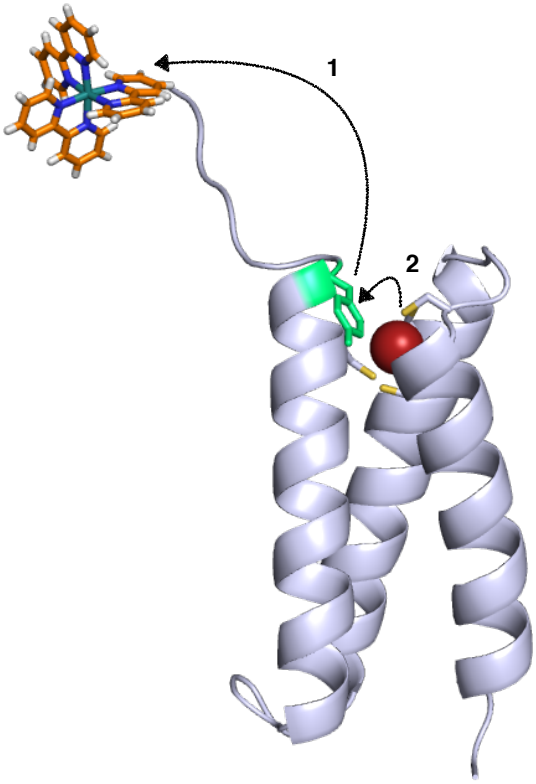


Figure 4.24 Model for electron transfer via hopping mechanism (pictured is iron-bound derivative). Ru(III) is reduced by tyrosine (green) with $k_{app} = 10^5\ s^{-1}$ (**1**). Generation of the oxidized tyrosine radical is quickly followed by electron transfer from the reduced metal site (dark red) to the oxidized tyrosine (**2**). The metal center is re-reduced by $MV^{+\bullet}$ to generate the ground state (not pictured).

As described in Chapter 1, a distance decay factor β describes the distance dependence of electron transfer rate by modulating the electronic coupling between the two sites in a superexchange pathway. While it seems that in most cases, a β of 1.1 $Å^{-1}$ is sufficient to describe

most proteins as averaging of multiple electron transfer pathways occurs, causing the protein to appear as if it represents a uniform barrier to electron tunneling. These studies do not allow us to interpret β for the α_3D scaffold, as multiple labeling sites and distance dependence would need to be measured. In order to determine the reorganization energy, the rate dependence on driving force would also need to be measured. By taking into account the size of the chromophore and the intervening residues, a reasonable estimate for the distance between the chromophore and the metal site is ~20 Å. The distance between the ruthenium and the tyrosine can be estimated at ~16 Å. The reorganization energy of iron sulfur clusters has been measured and calculated to be relatively low, on the order of 0.2 to 0.5 eV.[26,27] Using Dutton's simplified equation (11) and using the potential of the iron site (Chapter 3) with a reorganization energy of 0.4, the electron transfer between the iron site (where distance is 20 Å) the rate is predicted to be much less than 1 s^{-1}. However, the electron transfer to a tyrosine at 16 Å should occur at a rate of 1.2 x 10^5 s^{-1}, which is very similar to the observed k_{app} in this study. This designed protein conforms to current electron transfer theory and these calculations support the model in which tyrosine is required for electron transfer to the nearby metal center as the absence of tyrosine results in a decrease in rate of over two orders of magnitude.

Tyrosine radicals are vital to long distance electron transfer in many important biological systems such as Photosystem II (PSII), ribonucleotide reductase (RNR), and cytochrome c oxidase.[28] These radicals can play both oxidative and reductive roles as their reduction potential is highly sensitive to their protonation state. While the importance of tyrosine radicals in natural processes is well-known, the details of their behavior can be hard to study because of the complexity of the systems of which they are a part. In RNR and PSII tyrosines act as relays, whereas when engineered in azurin to be between a rhenium compound and the copper active site it was completely inactive even though the same system with a tryptophan replacing the tyrosine showed tryptophan acting as a relay.[29,30] The Tommos group has recently reported a three-helix bundle, α_3Y that has been used specifically to study tyrosine in peptides, but all the studies have been electrochemical, but were unable to spectroscopically characterize the tyrosine radical.[31,32] Our system rather adventitiously is capable of providing spectroscopic characterization of a tyrosine radical in the absence of other complex processes that exist in systems such as PSII or RNR.

Through these studies we have observed that a designed tetrathiolate site with iron coordinated can interact in intermolecular electron transfer with $Ru(III)(bpy)_3$. Furthermore, by attaching a ruthenium bipyridine derivative, we can observe intramolecular electron transfer via tyrosine as a relay. This is the first time that such detailed transient absorption studies on a *de novo* designed protein have been undertaken. This also represents the first kinetic study of non-heme intramolecular electron transfer in a designed protein as well as the first time a relay amino acid has been characterized in a designed protein. The kinetic parameters extracted from these studies are consistent the theoretical rate for electron transfer over this distance between ruthenium and tyrosine. This series of studies shows that designed proteins can be engineered to produce long distance electron transfer that mimics native processes and suggests strategies of how this scaffold is capable of being engineered to create a multi-site multi-function redox protein.

References:

(1) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 825–831.
(2) Ghadiri, M. R.; Soares, C.; Choi, C. *J. Am. Chem. Soc.* **1992**, *114*, 4000–4002.
(3) Mutz, M. W.; McLendon, G. L.; Wishart, J. F.; Gaillard, E. R.; Corin, A. F. *Proc. Natl. Acad. Sci. U. S. A* **1996**, *93*, 9521–9526.
(4) Mutz, M. W.; Case, M. A.; Wishart, J. F.; Ghadiri, M. R.; McLendon, G. L. *J. Am. Chem. Soc.* **1999**, *121*, 858–859.
(5) Fedorova, A.; Shin, Y.; Ogawa, M. *J. Am. Chem. Soc.* **2000**, *122*, 7999–8006.
(6) Rau, H. K.; DeJonge, N.; Haehnel, W. *Proc. Natl. Acad. Sci. U. S. A* **1998**, *95*, 11526–11531.
(7) Fedorova, A.; Ogawa, M. Y. *Bioconjugate Chem.* **2002**, *13*, 150–154.
(8) Fedorova, A.; Chaudhari, A.; Ogawa, M. Y. *J. Am. Chem. Soc.* **2003**, *125*, 357–362.
(9) Sakamoto, M.; Kamachi, T.; Okura, I.; Ueno, A.; Mihara, H. *Biopolymers* **2001**, *59*, 103–109.
(10) Farid, T. A.; Kodali, G.; Solomon, L. A.; Lichtenstein, B. R.; Sheehan, M. M.; Fry, B. A.; Bialas, C.; Ennist, N. M.; Siedlecki, J. A.; Zhao, Z.; Stetz, M. A.; Valentine, K. G.; Anderson, J. L. R.; Wand, A. J.; Discher, B. M.; Moser, C. C.; Dutton, P. L. *Nat. Chem. Biol.* **2013**, *9*, 826–833.
(11) Roy, A.; Sommer, D. J.; Schmitz, R. A.; Brown, C. L.; Gust, D.; Astashkin, A.; Ghirlanda, G. *J. Am. Chem. Soc.* **2014**, *136*, 17343–17349.
(12) Baron, A.; Herrero, C.; Quaranta, A.; Charlot, M.-F.; Leibl, W.; Vauzeilles, B.; Aukauloo, A. *Chem. Commun.* **2011**, *47*, 11011–11013.
(13) Baron, A.; Herrero, C.; Quaranta, A.; Charlot, M.-F.; Leibl, W.; Vauzeilles, B.; Aukauloo, A. *Inorg. Chem.* **2012**, *51*, 5985–5987.
(14) Kalyanasundaram, K. *Coord. Chem. Rev.* **1982**, *46*, 159–244.
(15) Campagna, S.; Puntoriero, F.; Nastasi, F.; Bergamini, G.; Balzani, V. Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 2007; Vol. 280, pp 117–214.
(16) Cangelosi, V. M.; Deb, A.; Penner-Hahn, J. E.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7900–7903.
(17) Plegaria, J. S. Ph. D. Thesis. University of Michigan. Ann Arbor, MI. April 2015.
(18) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed.* **2001**, *40*, 2004–2021.
(19) Wu, P.; Feldman, A. K.; Nugent, A. K.; Hawker, C. J.; Scheel, A.; Voit, B.; Pyun, J.; Fréchet, J. M. J.; Sharpless, K. B.; Fokin, V. V. *Angew. Chem. Int. Ed.* **2004**, *43*, 3928–3932.
(20) Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. *J. Am. Chem. Soc.* **2004**, *127*, 210–216.
(21) Cava, M. P.; Deana, A. A.; Muth, K.; Mitchell, M. J. *Org. Synth.* **1961**, *5*, 93–94.
(22) Walker, M. A. *J. Org. Chem.* **1995**, *60*, 5352–5355.
(23) Darwent, J. R.; Kalyanasundaram, K. *J. Chem. Soc., Faraday Trans. 2* **1981**, *77*, 373.
(24) Candeias, L. P.; Turconi, S.; Nugent, J. *Biochim. Biophys. Acta* **1998**, *1363*, 1–5.
(25) Cordes, M.; Köttgen, A.; Jasper, C.; Jacques, O.; Boudebous, H.; Giese, B. *Angew. Chem. Int. Ed.* **2008**, *47*, 3461–3463.
(26) Sigfridsson, E.; Olsson, M.; Ryde, U. *J. Phys. Chem. B* **2001**.
(27) Kümmerle, R.; Zhuang-Jackson, H.; Gaillard, J. *Biochemistry* **1997**.
(28) Migliore, A.; Polizzi, N. F.; Therien, M. J.; Beratan, D. N. *Chem. Rev.* **2014**, *114*, 3381–3465.

(29) Crutchley, R. J.; Ellis, W. R., Jr; Gray, H. B. *J. Am. Chem. Soc.* **1985**, *107*, 5002–5004.
(30) Gray, H. B.; Winkler, J. R. *Quart. Rev. Biophys.* **1999**, *36*, 341–372.
(31) Valentine, K. G.; Hay, S.; Tommos, C. *J. Am. Chem. Soc.* **2011**, *133*, 17786–17795.
(32) Berry, B. W.; Martínez-Rivera, M. C.; Tommos, C. *Proc. Natl. Acad. Sci. U. S. A* **2012**, *109*, 9739–9743.

Chapter 5. Nitrite reductase activity in the α_3D scaffold

Overview

Having described my research using *de novo* protein design to better understand homeostasis and electron transfer, a key function remaining is to be able to build a multi-site, multi-function protein redox catalysis. The development of such capability in the same scaffold containing a developed electron transfer site would then allow for future studies to concentrate on combining the two into a unified construct. However, α_3D derivatives have not yet been used for redox catalysis. In this chapter I will describe my work building on the work of a former graduate student, Dr. Fangting Yu, who investigated nitrite reductase activity within the TRI scaffold.[1] Much as she took a tris-histidine site that had been studied for carbonic anhydrase activity and substituted in a copper, I have replaced the Zn(II) in the previously reported α_3D**H$_3$** (Figure 5.1) with a copper and have begun investigation into its redox catalytic properties. The folding and electrostatics of α_3D**H$_3$** are known to be different from histidine-substituted TRI derivatives. The protein scaffold has been shown to affect the carbonic anhydrase activity of Zn(II)-α_3D**H$_3$** relative to TRI. It seems likely that copper substitution in this peptide to give Cu(I/II)-α_3D**H$_3$** would result in different properties than those observed in TRI derivatives given the difference in protein topology and local environment of the metal site in the protein. Herein, I present preliminary work on the nitrite reductase activity of α_3DH$_3$.

Introduction

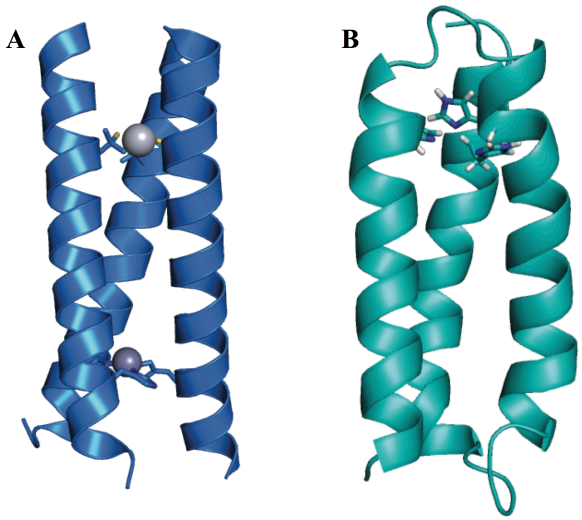


Figure 5.1 PyMol representations of TRI and α_3D peptides with trishistidine sites. A) crystal structure of Hg_SZn_HTRIL9CL23H, which has been shown to be an effective carbonic anhydrase mimic. B) model of α_3D**H$_3$** based on NMR structure of α_3D**IV** (2MTQ). This peptide has also been shown to have carbonic anhydrase activity when Zn is bound. As substitution of TRIL23H with copper resulted in a series of effective nitrite reductase mimics, substitution of α_3D**H$_3$** with copper could also be capable of redox catalysis.

Redox reactions are an important class of reactions in both biology and industry. From a design perspective, redox-active metalloproteins have an extra layer of complexity in that the designed construct must bind all redox states of the target metal. In some cases, this means only that multiple bond lengths must be accommodated, but in others, as with copper proteins, different binding geometries must be accessed with sufficient affinity to keep the metal bound in both oxidation states during turnover. Even when high stability for both oxidation levels is achieved, to obtain the desired reduction potential, the site must have the appropriate relative stabilities for these oxidation states. Copper centers play many roles in known proteins, participating in electron transfer and oxygen transport as well as catalyzing oxygen activation and metabolizing small molecule conversions.[2] Copper proteins are generally classified by their spectroscopic and functional characteristics. As covered in Chapter 1, the mononuclear copper centers termed types 0, 1, and 2 are largely distinguished by their spectroscopic characteristics. Type 2 sites are typically coordinated by histidines and display spectroscopy (particularly A_{II}) consistent with aqueous copper coordination. Type 2 copper sites are a versatile class of mononuclear copper that can function as both electron transfer centers and as redox catalysts. The type 2 copper centers in peptidylglycine -hydroxylating monooxygenase (PHM) and copper nitrite reductase (NiR) both exhibit $Cu(His)_3$ ligation, but perform drastically different roles: a type 2 site in PHM transfers an electron to a type 1 copper site,[3] while in NiR it accepts an electron from a type 1 copper site and catalyzes the proton-dependent reduction of nitrite to NO and water[4,5]. The studies outlined in Chapter 4 have begun to address this His_3 site as an electron transfer site in α_3DH_3, while this chapter will focus on its redox catalytic activity. Histidine-

ligated coppers are also present in type 3 copper sites, which is a binuclear site in which the two coppers are antiferromagnetically coupled when bridged.[6] Clearly, histidine-ligated copper complexes play diverse roles in nature without obvious selection for one role over another.

Protein design is an excellent way to examine the functional switch of these structurally related sites and how the protein environment may mediate this change, particularly since the $Cu(His)_3$ site may be prepared in the absence of the second copper site. The TRI scaffold was utilized to design a functional model for the catalytic site in NiR and was shown to catalyze the reduction of nitrite with ascorbate as a sacrificial electron donor with a pseudo-first order rate constant of 5.2×10^{-4} s^{-1} at pH 5.8.[7] The copper, when in its reduced state, is bound in a trigonal planar arrangement by the three histidine residues. However, the redox potential of this construct was 402 mV vs. NHE at pH 5.8, which is much higher than that of the type 2 copper center in NiR (218 mV at pH 6.0 in *R. sphaeroides*[8]).

Table 5.1: Spectroscopic and kinetic values for *de novo* designed NiR mimics based on the **TRI** scaffold. All values are reported for pH 5.8[9]

peptide	Δ_{charge}	d-d λ_{max} ($\varepsilon/M^{-1}cm^{-1}$)	reduction potential[a]	Cu^{I} K_d, $[M^{-1}]$	Cu^{II} K_d, $[M^{-1}]$	rate $[M^{-1}min^{-1}]$	pK_a^W
TRI-H	0	644 (143)	402±8	$(3.1\pm0.7) \times 10^{-12}$	$(4.0\pm0.8) \times 10^{-8}$	$(2.2\pm0.2) \times 10^{-6}$	8.53±0.02
TRI-EHE27K	0	671 (80)	600±10	$(1.4\pm0.6) \times 10^{-14}$	$(3\pm1) \times 10^{-7}$	$(1.1\pm0.7) \times 10^{-6}$	9.6±0.2
TRI-EHE27Q	-3		560±10	$(2\pm1) \times 10^{-14}$	$(1.3\pm0.3) \times 10^{-7}$	$(1.3\pm0.9) \times 10^{-6}$	
TRI-EH	-6	659 (110)	590±10	$(4.2\pm0.2) \times 10^{-14}$	$(7\pm3) \times 10^{-7}$	$(2.0\pm0.3) \times 10^{-6}$	9.86±0.05
TRI-EHK24Q	-9		490±20	$(5\pm3) \times 10^{-13}$	$(2\pm1) \times 10^{-7}$	$(3.8\pm0.8) \times 10^{-6}$	
TRI-EHK24E	-12	665 (101)	520±12	$(3.52\pm0.09) \times 10^{-14}$	$(5\pm2) \times 10^{-7}$	$(3.6\pm0.6) \times 10^{-6}$	9.8±0.2

[a]reduction potentials reported are calculated from the relative affinities of Cu(I) and Cu(II). Please see ref 1.

Given that the surrounding environment affects the redox potential of a site, the electrostatics of the site were altered to optimize the redox potential for increased catalytic rates.

The previously reported parent peptide, TRI-H (corresponding to TRIL23H) was altered by inverting the position of a lysine above the active site and a glutamate below the active site to give a construct named TRI-EHE27K (corresponding to K22E and E27K), which does not change the overall charge, only its distribution around the active site (Figure 5.2).[9] This inversion had the greatest effect on the Cu(I) binding affinity, resulting in 100-fold tighter binding regardless of pH (Table 5.1). To examine how the charge around the site could be used to alter the Cu(II) stability constant and consequently, the redox potential and activity; the change at K22E (TRI-EH) was kept constant and the residues below the active site (at positions 24 and 27) were varied to produce peptides with overall Δcharge of -3, -6, -9, and -12 (relative to the parent peptide, TRI-H).[9] This series of peptides was studied to examine the pH dependence of Cu(II) binding, affinity for Cu(I) and Cu(II), Cu(II) EPR spectra, and NiR activities produced by varying the secondary coordination sphere (Table 5.1).

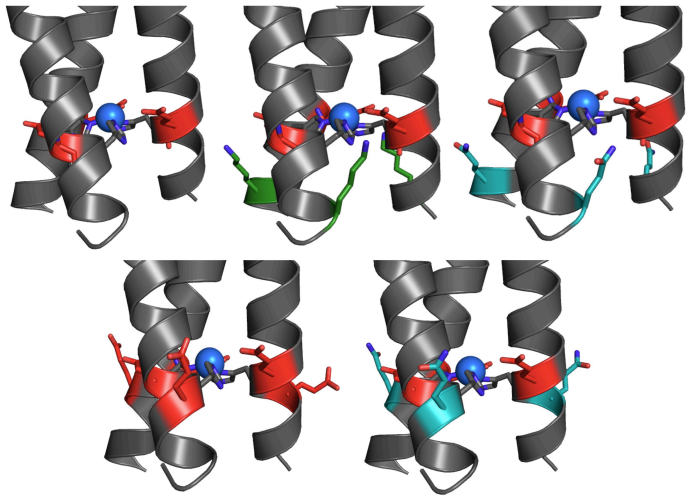


Figure 5.2 Representation of charge series around the copper site in **TRI-EH**. From top left, clockwise, **TRI-EH** (corresponds to **TRI-H**K22E, Δ_{charge} -6), **TRI-EH**E27K (Δ_{charge} 0)**, TRI-EH**K27Q (Δ_{charge} -3)**, TRI-EH**K24Q (Δ_{charge} -9)**, TRI-EH**K24E (Δ_{charge} -12). Red represents changes to negatively charged residues, green to positively charged residues, and teal to neutrally charged residues. Models are based on the crystal structure of $Hg^{II}_{S}Zn^{II}_{N}$(**CS**L9CL23H)$_3$ (PDB: 3PBJ).

Careful examination of the pH dependence of the d-d visible band and EPR signals revealed that the presence of K22E introduces a hydrogen bond to the Cu-coordinating imidazole and raises the pK_a of an exogenous water bound to the copper center (denoted as pK_a^{W}). The increase in local negative charge across the series resulted in less positive reduction potentials, as expected; however, this increase was due to a destabilization of Cu(I) binding, rather than a higher stability constant for Cu^{II}.[9] This study showed that, by careful choice of the pH conditions and the peptide sequence, one could vary the reduction potential of this site by nearly 200 mV.

Further studies on this system have included altering the packing above the layer and changing the electronics of the coordinating histidines. The sterics of the hydrophobic layer above the copper at the 19th position were modified by mutating the leucine to either an alanine, which may confer more space above the copper, or a bulkier isoleucine, that reduces the space above the copper. This series of peptides is: TRIW-HL19I, TRIW-H, and TRIW-HL19A. Subsequent studies on this series of peptides showed that changes at the 19th position lead to altered Cu(I) coordination number and minor changes in the Cu(II) coordination environment. Cu(II)/(I)(TRIW-HL19A)$_3^{2+/+}$ exhibits ~75-fold increase in rate of nitrite reduction as compared to Cu(II)/(I)(TRIW-H)$_3^{2+/+}$, which is likely due to the different coordination environment observed in this peptide, which may include better solvent/substrate access. Methylating the delta nitrogen of the imidazole ring altered the electronics of the coordinating histidine residue. The series of peptides for this study was TRIW-$_M$H, TRIW-$_M$HL19A, and TRIW-$_M$HK24E, all of which have a common feature of $_M$H. Cu(II)/(I)(TRIW-$_M$H)$_3^{2+/+}$ exhibits NiR rates increased by ~450-fold compared to Cu(II)/(I)(TRIW-H)$_3^{2+/+}$. Since the L19A and K24E mutations were shown to increase the NiR rates, therefore they were incorporated together to achieve a better enzyme. The combined mutation peptide, Cu(II)/(I)(TRIW-MHL19A)$_3^{2+/+}$ has an NiR rate that is ~700-fold higher than that of Cu(II)/(I)(TRIW-H)$_3^{2+/+}$.

These studies provide a well-characterized system with which to compare Cu-α_3DH$_3$, particularly with respect to its copper coordination spheres and nitrite reductase activity. In considering how copper introduced into this peptide may change its activity or coordination, it is important to note the difference between the two scaffolds. While both TRI and α_3D, have three helices, their overall topology is drastically different, which could affect both the copper binding properties and the catalytic activity. TRI is a three-fold symmetric system in which three helices self- associate to form a parallel three-stranded coiled coil. On the other hand, the α_3D system is a single-stranded construct, which folds into an asymmetric three helix bundle with short loop regions to generate an antiparallel construct. Therefore, in the TRI series the histidines are highly symmetric and the side chains will be oriented similarly with respect to the helical axis, while in α_3D, the side chains will be oriented opposite to each other on two α-helices that are antiparallel to each other. The comparison of how hydrolytic activity is affected by this topological difference was performed by comparison of carbonic anhydrase activity.[10] However, substitution

of copper into α_3DH_3 requires that the scaffold stabilize more than one oxidation state, and thus coordinating copper to a site in an antiparallel bundle may have implications into its activity.

Materials and Methods

Protein Production and Purification

Peptides are generally expressed and purified as described previously in Chapter 2, with the difference that lysis buffer is 1X PBS and lysozyme only (no DTT).

Absorption Spectroscopy

UV-visible spectroscopy was carried out on a Cary 100 instrument. $CuCl_2$ and $Cu(ACN)BF_4$ stocks were prepared in a glovebox under N_2 atmosphere and the concentration was determined by ICP in the case of Cu(II) and for Cu(I) by titration with a selective Cu(I) chelator as previously described.[9] In general, samples were prepared by dissolution of apo-peptide in water or buffer and the concentration was measured using a previously determined molar extinction coefficient.[10]

pH titration:

Copper was bound to the peptide and diluted to 400 μM of Cu- $\alpha_3D\mathbf{H_3}$ in water. The pH was titrated by adding small aliquots of sulfuric acid to the solution. The pH was monitored and allowed to stabilize before a measurement was taken. This was performed simultaneous on three cuvettes containing the same solution. The data were converted to molar extinction coefficients and each transition over the course of the pH experiment was fit independently to the following equation:

$$Y = \frac{\varepsilon_{dep} + \varepsilon_{prot} \times 10^{n(pK_a - pH)}}{1 + 10^{n(pK_a - pH)}}$$

where Y is the observed molar extinction coefficient, ε_{dep} is the molar extinction coefficient of the deprotonated form, ε_{prot} is the molar extinction coefficient of the protonated form, and n is the number of protons.

Cu(I) binding:

The affinity of the $(His)_3$ site in the peptide for Cu(I) was determined using a competitive chelation assay as has previously been described. The well-defined Cu(I) chelator bathocuproine disulfonate (BCS^{2-}) was used, which binds Cu(I) into a colored species $[Cu(BCS)_2]^{3-}$ (λ_{max} = 483 nm; ε_{483} = 13300 $M^{-1}cm^{-1}$; log β_2 = 19.8).[43] The concentration of peptide stock solution was determined from Trp absorbance at 280 nm. A solution of 40 μM Cu(I) and 80 μM $(TRIW\text{-}H)_3$ was prepared in an inert atmosphere box in an aqueous 50 mM buffer solution (MES pH 5.9) and titrated with a 5.00 mM bathocuproine disulfonate solution. The formation of $[Cu(BCS)_2]^{3-}$ was monitored by the increase of absorbance at 483 nm with UV-visible spectroscopy. All titrant additions were carried out in an inert atmosphere box. Hyperquad was used to fit the spectrophotometric data taking into account the protonation equilibrium of BCS^{2-} at pH 5.9.

Fluorescence Spectroscopy

Cu(II) affinity for the peptide was performed according to previously reported procedures. Briefly, Cu(II) quenching of tryptophan fluorescence was used. A 500 nM solution of peptide at pH 5.8 was titrated with small aliquots of $CuCl_2$ and the fluorescence intensity quenching was monitored. The curve was fit to the following equation to yield the K_d:[1]

$$Y = I_o - \left[\frac{(I_o - I_f)}{2[P]}\left((K_d + [Cu(II)] + [P]) - \sqrt{(K_d + [Cu(II)] + [P])^2 - (4\times[Cu(II)]\times[P])}\right)\right]$$

Cyclic Voltammetry

Solution electrochemistry was used to determine the reduction potential at pH 8.5. Copper-bound peptide was prepared in a glovebag (Sigma Aldrich) under Ar atmosphere and added to an Ar(g) degassed 5 mL electrochemical cell with 100 mM HEPES buffer and 100 mM Na_2SO_4 at pH 7.5 and 8.0. Temperature was maintained at 20 °C by using a circulating water temperature controller. Voltammograms were measured on an Autolab potentiostat using a highly ordered graphite (HOPG) carbon electrode as the working electrode, standard calomel as the reference electrode, and platinum wire as a counter electrode. Measurements were taken at 20 mV/s.

X-Ray Absorption Spectroscopy

Samples were made in 25 mM Tris buffer in the absence of chloride with 0.66 eq of Fe(II) to ensure the absence of free iron in the sample. The samples were mixed with 30%

glycerol as a glassing agent, loaded into a sample cell, and frozen in liquid nitrogen. Measurements were carried out at the Stanford Synchrotron Radiation Lightsource (SSRL) beamline 7-3 with a Si (220) double-crystal mono-chromator and a flat Rh-coated harmonic rejection mirror. Samples were maintained below 10 K with an Oxford Instruments liquid helium cryostat. Data were measured as fluorescence excitation spectra using a 30-element Ge array detector normalized to incident intensity measured with an N_2-filled ion chamber. Data were measured with steps of 0.25 eV in the XANES region (1 s integration time) and 0.05 $Å^{-1}$ in the EXAFS region to k = 12.5 $Å^{-1}$ (1–20 s integration, k^2 weighted). Data reduction and fitting were performed in EXAFSPAK. The data were calibrated to a copper foil at 8980 eV. The data were converted to k space using 9000 eV.

Michaelis-Menten Kinetics: Nitrite Reductase activity

All kinetic reactions were prepared anaerobically in total volumes of 400 µl and measured in 0.1 cm path length cuvettes. Different concentrations of $NaNO_2$ were incubated with the peptides. Samples contained 100 µM Cu(II)(3SCC), 50 µM apo-peptide in 50 mM potassium phosphate buffer. Controls contained 50 µM apo-peptide. The pH was adjusted to 5.8. The reactions were initiated by injecting NaAsc to the reaction mixture. The final NaAsc concentration was 1.5 mM. The rates of the reaction were calculated by two times the difference between rates of NaAsc consumption (decrease of absorbance at 251 nm) between the sample and the control. The rates were correlated to the substrate concentrations through fits to the Michaelis-Menten equation in Prism 5 (GraphPad software).

Results and Discussion

While many classes of enzymes use zinc as part of the active site or for structural stabilization, perhaps the most well known role for Zn^{II} is in the active site of carbonic anhydrase (CA). Carbonic anhydrase is a hydrolase enzyme that catalyzes the reversible hydration of carbon dioxide. In α-CAs, the most extensively-studied family of CA, Zn^{II} is tetrahedrally coordinated to three histidine residues and an exogenous water ligand, which is deprotonated to form the catalytically competent species.[11] A former graduate student, Dr. Melissa Zastrow, examined zinc hydrolytic catalysis in the TRI system. This bifunctional construct, termed $Hg(II)_SZn(II)_N$(**TRI**L9CL23H)$_3$ was characterized for its hydrolytic activity towards the non-

native substrate p-nitrophenylacetate as well as CO_2.[12] X-Ray crystallography showed that this is an excellent structural model for the primary coordination sphere of Zn(II) in carbonic anhydrase. This construct exhibited a kinetic pK_a of 8.8 (corresponding to the Zn-bound water) and a catalytic efficiency towards CO_2 hydration of 1.8×10^5 $M^{-1}s^{-1}$, which is the fastest rate for hydration of CO_2 reported for a carbonic anhydrase mimic. This system can be compared to Zn(II)- α_3DH_3, in which leucine residues were replaced with histidines (L18H, L28H, L67H) to create a tris-histidine site and a nearby histidine was changed to a valine (H72V) to prevent competition for Zn^{II} binding.[10] This construct, $\alpha_3D\mathbf{H_3}$, is the first reported *de novo* designed, single-stranded three-helix bundle that functions as a CA mimic. This protein catalyses CO_2 hydration with a kinetic pK_a of 9.4 and a catalytic efficiency of 3.2×10^4 (M^{-1} s^{-1}), which is 5.6-fold slower than $Hg^{II}{}_S Zn^{II}{}_N(\mathbf{TRI}L9CL23H)_3$. A comparison of the other kinetic parameters reveals that TRIL9CL23H has a k_{cat} that is an order of magnitude higher and a K_m that is about 2.5-fold higher than, $\alpha_3D\mathbf{H_3}$. Additionally, Zn(II) was found to bind more tightly to $\alpha_3D\mathbf{H_3}$ by two orders of magnitude. These differences in catalytic activity and binding affinity were attributed to the differences in the folds of the two peptides. An antiparallel bundle is expected to have a weaker dipole than the coiled-coil because adjacent helices pointing opposite directions will largely cancel out each other's dipole. Also, in an antiparallel construct, the imidazole rings will be in less-symmetric orientations.

In considering how substituting copper for zince into $\alpha_3D\mathbf{H_3}$, one must take into account the major differences between the chemistries of copper and zinc. Zn(II) is not redox-active and is promiscuous in its ligand preference, and is found bound to both histidines and cysteines in native systems.[2] Zn(II) generally prefers tetrahedral geometry and the coordination number often changes during its catalytic cycle as substrates bind. On the other hand, Cu(II) is d^9 with a strong Jahn-Teller distortion and thus prefers axially distorted square planar coordination.[2] Adding an electron to the system produces d^{10} Cu(I), which is isoelectronic with Zn(II) and can adopt a linear, trigonal, or tetrahedral coordination environments. This strong preference for a different coordination geometry in two different redox states creates a high inner sphere reorganization energy during oxidation and reduction events as both bond lengths and geometry will change to accommodate the different preferences of Cu(II) and Cu(I). A copper redox protein must accommodate both oxidation states of copper with minimal perturbations so as to lower this reorganization geometry. The "entatic" state model is typically used to describe copper bound to

proteins (particularly type 1 sites), in that the protein enforces a high energy geometry to lower the reorganization energy during redox processes.[2,6] A copper enzyme then must balance the stability of both oxidation states as well as bind substrate.

Dr. Fangting Yu studied type 2 copper sites in the TRI scaffold and their ability to catalyze the reduction of nitrite to nitric oxide.[1] These studies included spectroscopic and kinetic characterization of the Cu(I) and Cu(II) states, as well as the kinetic profiles of several series of mutants. While the base peptide was catalytic, modifications to the charge environment, steric environment, and ligands have improved the catalytic parameters and provided a wealth of data on how small changes in the surrounding environment can perturb both the function and spectroscopy of the type 2 copper sites. Thus I can compare α_3D**H$_3$** directly to these previous studies and use them to benchmark the properties of α_3D**H$_3$**. A designed protein with a coupled electron transfer and redox catalyst system needs a catalytic site, and by utilizing protocols in use in our lab I can and assess its potential for a scaffold for redox catalysis.

To examine the ability of α_3D**H$_3$** to bind and utilize copper, I began by examining the UV-visible spectroscopy of copper bound to the peptide. Type 2 copper sites show low molar extinction coefficients between 600 and 700 nm. At pH 5.8, the copper shows a broad absorbance centered at 670 nm with molar extinction coefficient of 160 $M^{-1}cm^{-1}$. The energy of the d-d band is lower than those of both TRIW-H and TRIW-EH, which could indicate some difference in the environment around the metal site or in the coordination of the Cu(II) The peptide TRIW-H has a λ_{max} at 643 nm with a molar extinction coefficient of 135 $M^{-1}cm^{-1}$, while TRIW-EH has a λ_{max} of 659 nm and a molar extinction coefficient of 110 $M^{-1}cm^{-1}$.[7,9] Furthermore, the peptides with $_M$H and L19A modifications show similarly low energy d-d bands, although the molar extinction coefficients are also much lower. The λ_{max} suggests that α_3D**H$_3$** has coordination numbers similar to these peptides, although the molar extinction coefficient suggests that the geometry may be slightly different, so as to slightly increase the transition.

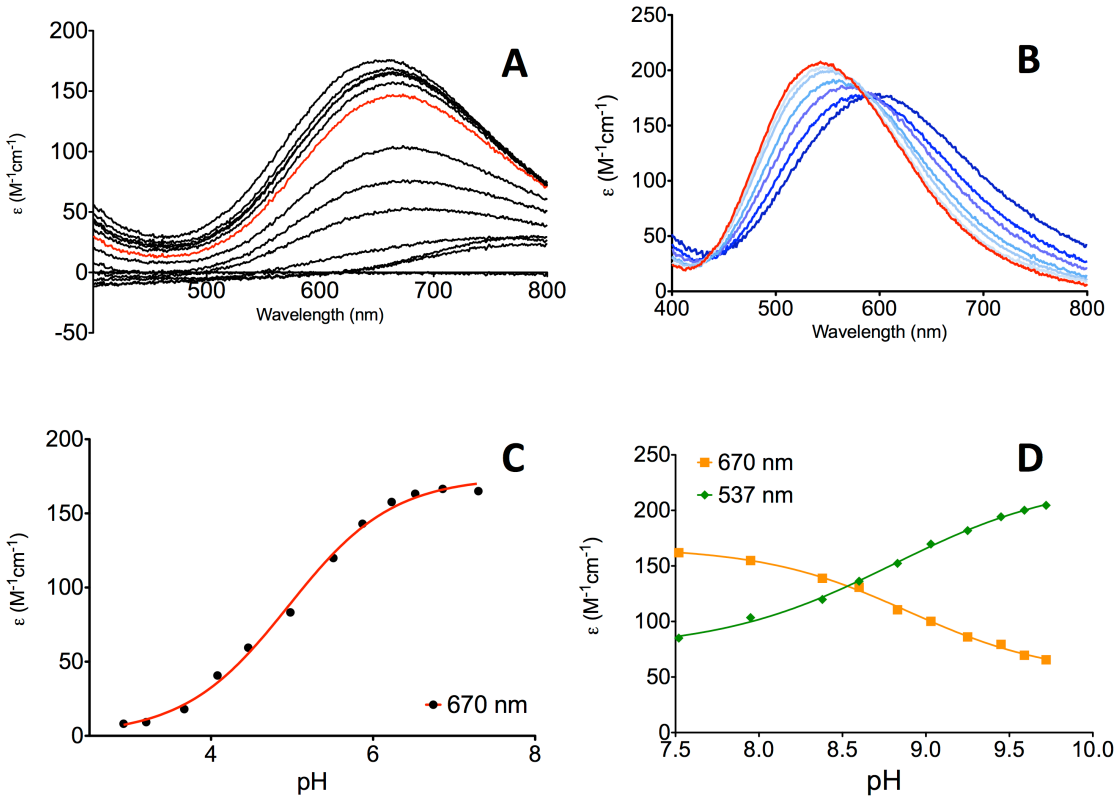


Figure 5.3 The d-d bands of Cu-α_3D**H$_3$** are pH dependent. At low pH from pH 3.5 to 7.5 (A) the intensity of the transition increases and is associated with copper binding. At pH 5.8 (red line) the absorbance is centered at 670 nm. As the pH is raised further from 7.5 to 9.8 (B), the absorbance shifts to 630 nm and then transitions to 537 nm at pH 9.8 (red line) with isosbestic points at 484 nm and 588 nm. The pH dependence associated with Cu(II) binding (C) can be fit to pK_a of 4.8. The deprotonation at higher pH (D) can be fitted at multiple wavelengths to a pK_a of 8.8

As the pH changes from low to high, the spectrum shifts where at low pHs the λ_{max} is 670 nm, which then shifts to higher energy ending at 538 nm at pH 9.7 with a molar extinction coefficient of 200 $M^{-1}cm^{-1}$(Figure 5.3, Table 5.2). The pH dependence of the copper binding process results in a pK_a of 4.8 ± 0.1. As the pH is raised past 8, a process that is likely associated with water bound to the copper center takes place and the spectrum converts to that of the high pH form with isosbestic points at 588 nm and 434 nm. The pK_a of the deprotonation is 8.83 ± 0.05 and corresponds to a 1-proton process. The pK_a of TRIW-H was 8.5, but the pK_as of the TRIW-EH series are all a pH unit higher—between 9.6 and 9.9. This would suggest that the lower energy d-d bands in α_3D**H$_3$** are not directly correlated with the pK_a, as the TRIW-EH series of peptides all had lower energy d-d bands, but much higher pK_as than α_3D**H$_3$**.

Table 5.2 Compared UV-vis and affinity parameters at pH 5.8

peptide	d-d λ_{max} ($\varepsilon/M^{-1}cm^{-1}$)	reduction potential	Cu^{I} K_d, [M^{-1}]	Cu^{II} K_d, [M^{-1}]	pK_a
TRI-H[7]	644 (143)	402±8	(3.1±0.7) x 10^{-12}	(4.0±0.8) x 10^{-8}	8.53±0.02
TRIW-HL19A[13]	652 (94)	600±10	(1.4±0.6) x 10^{-14}	(1.3±0.2) x 10^{-5}	9.08±0.03
α_3D**H$_3$**	670 (160)	200±30	2 x 10^{-9} [a]	(2±1) x 10^{-7}	8.83±0.05

[a]calculated based on reduction potential at pH 7.5 and estimated affinity of Cu(II) at pH 7.5 based on trends in TRI

To determine the Cu(II) binding affinity, I used the quenching of tryptophan fluorescence as has previously been reported.[7,9] At pH 5.8, the K_d of α_3D**H$_3$** for Cu(II) is 2±1 x 10^{-7} M^{-1}, which is equivalent to that for the TRIW-EH series, and an order of magnitude higher than $_M$H and L19A, and an order of magnitude smaller than TRIW-H (Figure 5.4, Table 5.2). Overall, the Cu(II) binding affinity for this peptide is unremarkable in the context of peptides studied in our lab. To determine the Cu(I) binding affinity I used bathocuproinedisulfonate (BCS^{2-}) as a competitive chelator as has been previously been reported.[1,7,9] The spectrum of $Cu(BCS)_2^{2-}$ is characterized by a λ_{max} of 483 nm with a molar extinction coefficient of 13300 $M^{-1}cm^{-1}$ and a logβ_2 of 19.8. The peptide was bound with 1 eq. Cu(I) and 5 mM BCS was titrated into the solution, causing an increase in the absorbance at 483 nm (Figure 5.4).

The spectra were fit with global fitting program, Hyperquad, taking into account the pK_a of BCS. Unfortunately, the results of the titration could not be accurately fitted unless a peptide-BCS-Cu(I) ternary complex was included in the model (Figure 5.4). The β for this species was fitted to 19.1 (± 0.2). This high affinity ternary complex precludes the ability to fit the affinity of the peptide for Cu(I). Another chelator used in our lab is bicinchoninic acid (BCA), which has lower affinity than BCS. The log β_2 for BCA has been reported at 11.4, 14.7, and 17.2, depending on the method used to determine the affinity.[14-16] Nonetheless, this chelator has been previously used in our lab to determine the Cu(I) affinity to peptides. Another possibility for a chelator with lower affinity than either BCA or BCS is (5,6-diphenyl-3-(2-pyridyl)- 1,2,4-triazine-4,4''-disulfonic acid) (Fz), which has a log β_2 of 11.6 and has been used to determine the affinity of Aβ peptides.[17] The combined Cu(I) and Cu(II) affinities have previously been used to

calculate the reduction potential of the site based on the relative affinities and the known reduction potential of free copper. In this case, the ternary complex precludes an accurate determination of the Cu(I) binding constant so the reduction potential cannot be measured in this way.

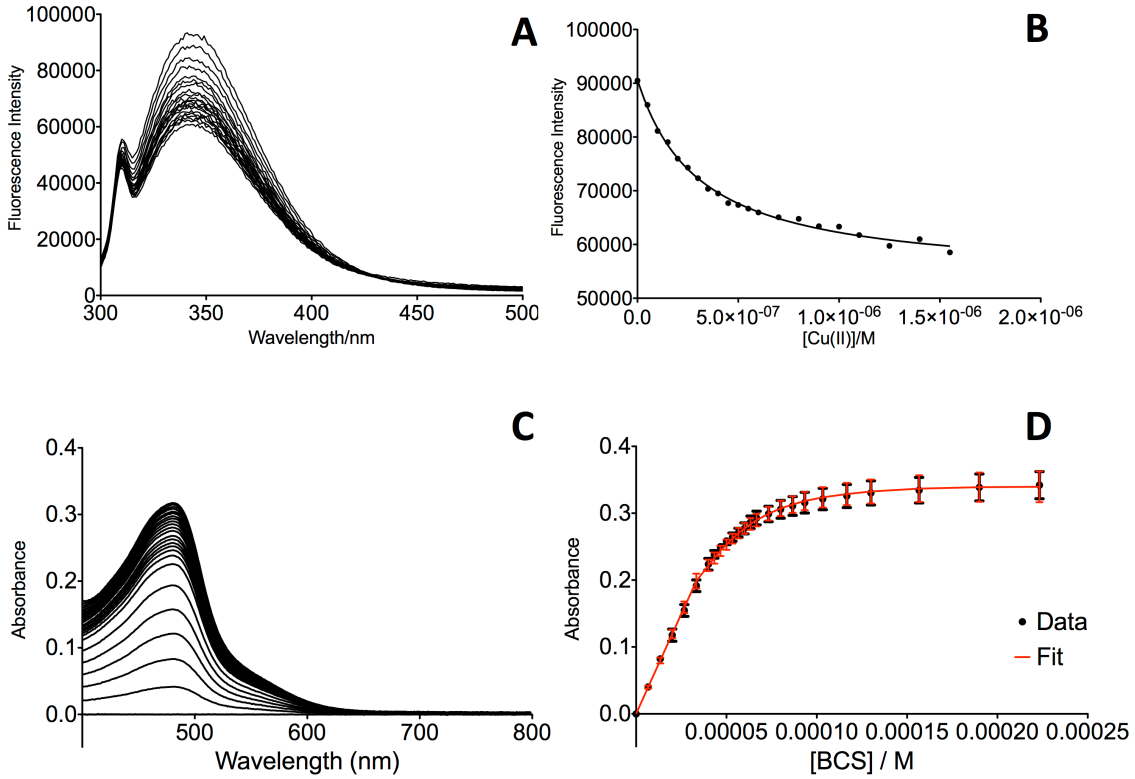


Figure 5.4 Cu(II) affinity was measured by the quenching of tryptophan fluorescence (A) and fit to yield a K_d (B, representative). Cu(I) chelation by BCS^{2-} gives a strong absorption at 483 nm (C) unfortunately global fitting analysis (D, in quadruplicate) revealed a high affinity ternary complex, which precludes accurate K_d determination.

Some initial studies had been done with $\alpha_3D\mathbf{H_3}$-GSGC in which I conjugated the final cysteine residue with a maleimide (Table). This construct, termed $\alpha_3D\mathbf{H_3}$-mal, should present the same coordination environment as the peptide used for most of these studies. For comparison, the λ_{max} of $\alpha_3D\mathbf{H_3}$-mal at pH 8.0 is 620 nm, while that of $\alpha_3D\mathbf{H_3}$ is 617 nm. With this peptide I performed some solution electrochemistry experiments. These experiments were carried out in 100 mM HEPES with 100 mM Na_2SO_4 as supporting electrolyte. At pH 7.5 with a scan rate of 20 mV/s, the peptide showed semi-reversible oxidation and reduction processes where the midpoint potential was 195 mV vs. NHE. At pH 8.0 and 20 mV/s the midpoint potential was 200 mV vs. NHE (Figure 5.5). However, the reduction and oxidation peaks are separated by nearly 100 mV, suggesting very slow kinetics or another process occurring at the same time, such as a

change in coordination. Given the difference in preferred coordination between Cu(I) and Cu(II) and how the TRI peptides change coordination number from 4 or 5 to 3 upon reduction, it is likely that this process occurs in this system too. The peak separation increases with higher scan

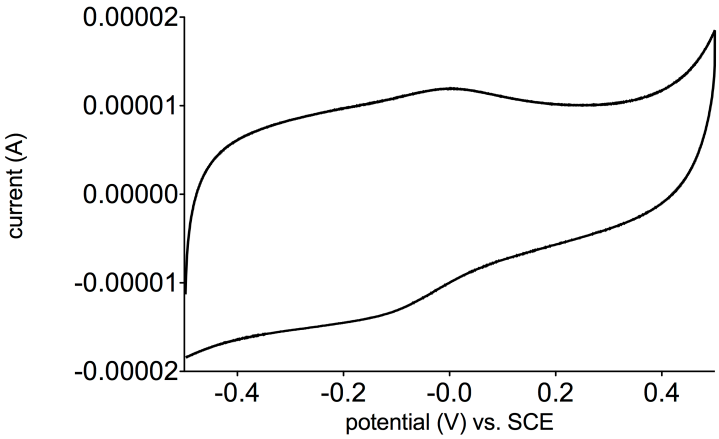


Figure 5.5 Solution cyclic voltammetry on α_3D**H$_3$**-mal at pH 7.5 in 100 mM HEPES and 100 mM Na_2SO_4 with HOPG working electrode, standard calomel reference, platinum wire counter electrode. Peaks were determined by fitting in Autolab software.

rates, suggesting that this electrochemical behavior has a kinetic component. Thus, I've used the slower scan rates to determine a more accurate reduction potential. Assuming that the Cu(II) affinities follow the same trend as those in TRI ,it should bind with a K_d of 10^{-9} at pH 7.4. Based on the relationship between reduction potential and relative affinities of each redox state, then the Cu(I) binding affinity at pH 7.5 is ~2 x 10^{-9} M, which is two orders of magnitude lower than that seen for most of the TRI series.[1] Also striking is the difference in reduction potential between the Cu-TRIW-H series and Cu-α_3D**H$_3$**-mal. The reduction potentials for the TRI series are much higher—all have reduction potentials between +400 and +600 mV vs. NHE. This dramatic difference of 200 mV in the reduction potential could have major implications in the kinetic behavior of this peptide.

Table 5.3 Compared EPR parameters

Protein	k_{app} (s^{-1})	G	A_{II} (mT)	Reference
Nitrite Reductase	2.05	2.3	14	18
Peptidylglycine monooxygenase	2.05	2.29	16.81	19
TRIW-H	2.05	2.27	18.58	7
α_3D**H$_3$**	2.045	2.27	18.40	This work

EPR taken at X-band on the α_3D**H$_3$** peptide at pH 5.8 showed a type 2 axially distorted EPR signal (Figure 5.6). The superfhyperfine splitting from the imidazole nitrogens is resolved, and supports the conclusion that this site is a type 2 copper site. Also, the g values and A_{II} are consistent with those observed in type 2 copper sites in native proteins and with those observed

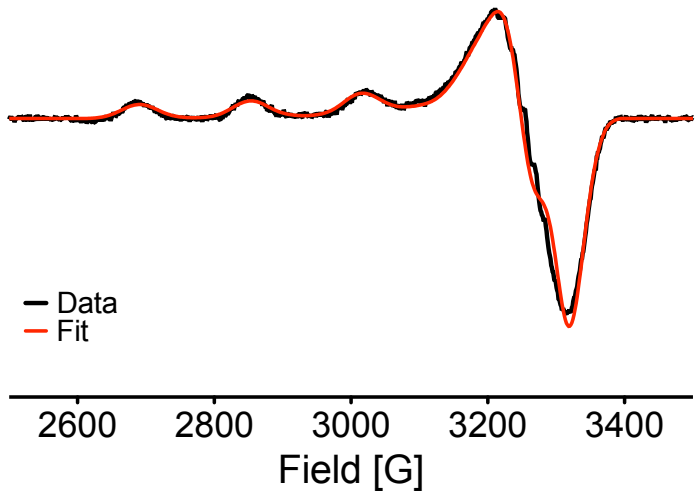
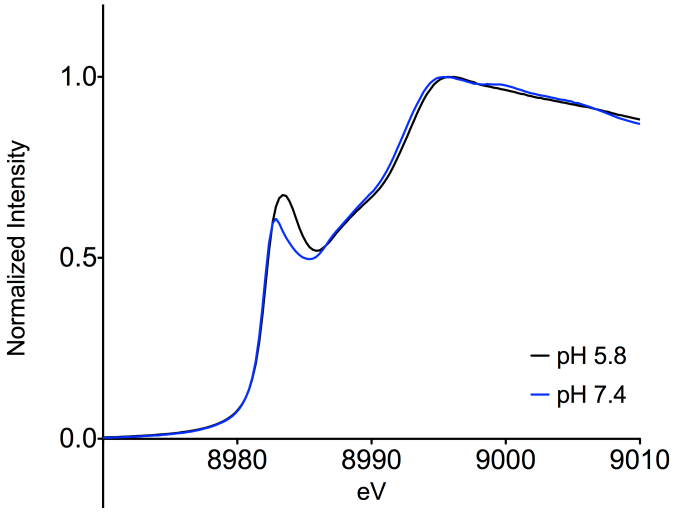


Figure 5.6 EPR Cu(II)-α_3D**H$_3$** at pH 5.8 (black). Simulated spectrum (red) corresponds to $g_\perp = 2.045$, $g_{II} = 2.27$, and $A_{II} = 18.40$ mT. Recorded at 120K, 9.316 GHz, 5G modulation amplitude, 20.51 mW microwave power.

for model Cu-peptides in our lab. The $g_\perp$ was found to 2.055 and the g_{II} was 2.265, which are slightly lower than those observed in nitrite reductase[8] although comparable with those observed in the TRI series[1,9]. The A_{II} coupling constant is 16.15 mT, which is slightly lower than that of the TRI series peptides, but equal to that found in nitrite reductase. While these slight differences may indicate small structural or geometric changes, the EPR confirms that α_3D**H$_3$** binds copper in a type 2 site.

Figure 5.7 XANES spectra of Cu(I)-α_3D**H$_3$** at pH 5.8 (black) and pH 7.4 (blue) shows highly similar environments between the two pHs.



X-Ray absorption spectroscopy was performed on Cu(I)-α_3D**H$_3$** at pH 5.8 and pH 7.4 at the copper K edge. An examination of the X-Ray absorption near-edge structure (XANES) can give information about the coordination number and the oxidation state of the metal. In particular, the intensity of the pre-edge peak, which has been attributed to a 1s→4p transition, is indicative of coordination number. A very intense feature such that the ratio of the edge to pre-

edge is above 0.9 reflects a coordination number of two, while a moderately intense peak at ~0.5 indicates a coordination number of three, and four coordinate complexes have very low resolution of this feature.[20] In the case of both the pH 5.8 and pH 7.4 complexes, the intensity of this transition at ~8993 eV is consistent with a three coordinate site (0.69 and 0.59, respectively) (Figure 5.7). Extended X-Ray absorption fine structure (EXAFS) can be used to fit bond distances. EXAFS data was only collected on the pH 5.8 sample, however given the overall similarity of the XANES for both pH 5.8 and 7.4, the pH 5.8 sample is probably representative of the coordination environment over this pH range. Given the information about the coordination number and the design of the protein, it is reasonable to fit the EXAFS data beginning with three histidines. The first shell scattering is well reproduced by three nitrogen atoms at 1.95 Å with σ^2 of 0.007 and ΔE_o of -10 (Figure 5.8). This is consistent, although on the longer side, of what has previously been seen for copper bound to tris-histidine sites at pH 5.8.[1,7,9] The oscillations at R ~2-4 Å are consistent with backscattering from the second sphere nitrogen of the imidazole ring. Multi-scattering fitting that takes these effects into account is still underway, and will allow me to refine all of the fit parameters.

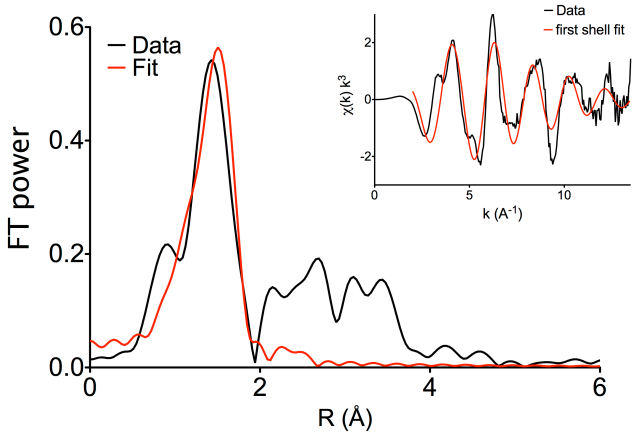


Figure 5.8 Preliminary first shell fits to EXAFS of Cu(I)-α_3D**H$_3$** at pH 5.8 yields R= 1.95 Å, with σ^2 of 0.007, ΔE_o of -10 and F= 90.9.

The ability of α_3D**H$_3$** to act as a catalyst for nitrite reductase activity was also examined. The assay relies on ascorbate as a source of electrons for the reaction and the disappearance of the ascorbate is monitored at 251 nm. The sample reaction has peptide-copper complex plus additional apo-peptide to ensure all the copper is bound. The control reaction has only apo-peptide. In initial trials I had issues distinguishing the rate between the control reaction so I modified the reaction conditions to have higher copper-peptide concentrations while maintaining low free copper levels (below 0.5%) based on the Cu(II) affinity. At pH 5.8, α_3D**H$_3$** catalyzes

nitrite reduction with a k_{cat} of 1400 min^{-1}, k_{cat}/K_m of 2600 $min^{-1}M^{-1}$, and K_m of 0.5 M (Figure 5.9).

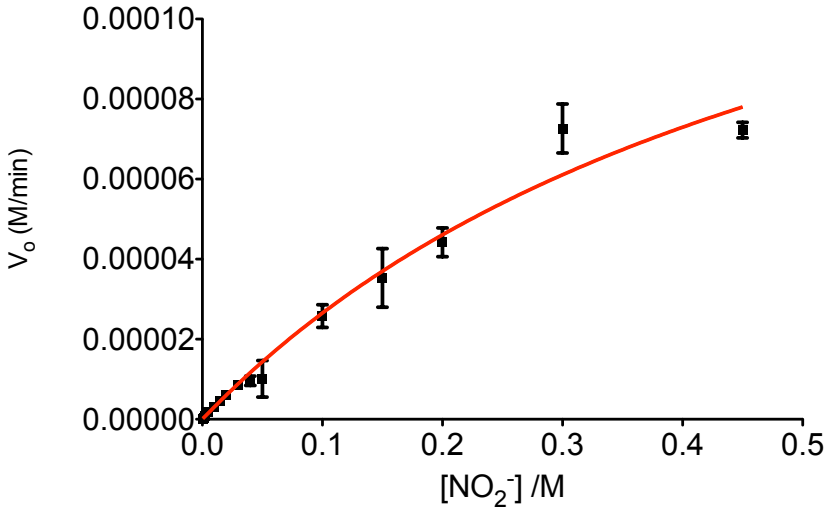


Figure 5.9 Michaelis Menten kinetics of nitrite reductase activity at pH 5.8.

Table 5.4 Comparison of Michaelis-Menten kinetics

Peptide	k_{cat} (min^{-1})	K_M (M)	k_{cat}/K_M ($min^{-1}M^{-1}$)
Cu(TRIW-HL19A)$_3^{n+}$	14 ± 2	0.24 ± 0.05	60 ± 15
Cu(TRIW-$_M$H)$_3^{n+}$	89 ± 7	0.18 ± 0.02	(4.9 ± 0.7) × 10^2
Cu(TRIW-$_M$HL19A)$_3^{n+}$	88 ± 3	0.13 ± 0.01	(6.8 ± 0.6) × 10^2
α_3D**H$_3$**	1.8 ± 0.3	0.5 ± 0.1	3.1 ± 0.3

TRI data from [13]

Compared to the TRI system, α_3D**H$_3$** has both a lower k_{cat} and k_{cat}/K_m value and a much higher K_m (Table 5.3). However, the parameters for k_{cat} and K_m are ill defined under these conditions because substrate levels did not reach or exceed K_m. The apparently higher K_m in α_3D**H$_3$** suggests that the affinity for the substrate is much lower in this system than in TRI. This is opposite to the trend observed in the carbonic anhydrase mimics, Zn_NHg_S-TRIL9CL23H and Zn-α_3D**H$_3$** (Figure 5.1). In α_3D**H$_3$**, the K_m was 3.5 ± 0.6 mM at the fastest pH measured, compared to 10 ± 2 mM in TRIL9CL23H, a decrease of 2.5-fold.[10] This suggests that space is not the only factor controlling the K_m in these systems, as one might expect that access of both bicarbonate and nitrite into the bundle would be similar in Zn(II)-α_3D**H$_3$** and Cu(II)-α_3D**H$_3$** based

on size. However, at this pH nitrite will be negatively charged, which could affect its entry into the site relative to carbon dioxide.

In TRIW-H and related peptides, it is hypothesized that the Cu(II) state has 3-5 ligands, with one or two waters coordinated. In particular, TRIW-H based on its relatively higher molar extinction coeffiecient (135 $M^{-1}cm^{-1}$ vs. 94 $M^{-1}cm^{-1}$ for L19A), was suggested to be 5-coordinate,

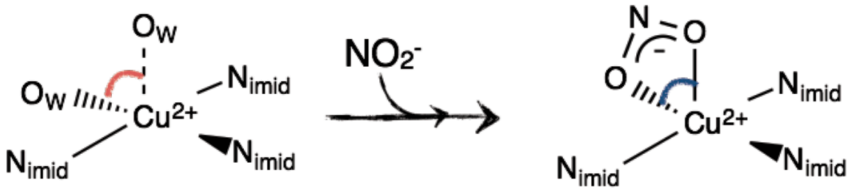


Figure 5.10 Proposed binding mode for nitrite upon displacement of water (O_W) in TRIW-H.

with three histidines bound and two waters. However, only one pK_a was detected that corresponds to a water, so either only one water is capable of deprotonating, which is typical of metal sites with multiple waters bound, or only one water is bound. This is the same situation in $\alpha_3D\mathbf{H_3}$, which also only shows a one 1-proton process in the pH titration d-d transitions. The explanation for this was that one of the waters was deprotonated upon binding, which means this process is included in the low pH, Cu(II) binding dependence pK_a (4.8 in this study, 4.5 for TRIW-HL19A).[1] Due to the rapid reduction of Cu(II) by ascorbate, Cu(I) is the resting state during turnover conditions. The cuprous ion is a three-coordinate species in the absence of substrate. If access to the copper site is perturbed in the case of nitrite, it seems reasonable that access to the copper site for ascorbate could be affected as well, which is larger and also negatively charged. This could affect the kinetics of both the initial binding of nitrite as well as the re-reduction of the Cu(II) peptide after one round of catalysis. The lower reduction is better matched with the potential of nitrite, which has a standard potential E′° +0.37 V, which is lower than the potentials of the TRI series.[21,22] This does not increase the efficiency of catalysis, which could be due to low substrate affinity.

The lowered reduction potential in this peptide as compared to the TRI series is likely a consequence of the antiparallel topology of α_3D. Since TRI is a parallel construct, the side chains all point towards the C terminus. However, in α_3D, two of the side chains will point in the same direction, while the other will point in the opposite direction as a consequence of the antiparallel nature of the bundle. The different orientations of the side chains in an antiparallel construct may

enforce something closer to an entatic state than in a three-fold symmetric construct like TRI, thus affecting the reduction potential. This distortion arising from the direction of the side chains in the site could explain why the catalytic rates in α_3D**H$_3$** are less than for TRI. These distortions could affect binding of copper and zinc as well as the binding of the substrate. It is also possible that the overall antiparallel topology and helical dipole orientation could have some effect on the metal site and its anion-binding capabilities. Further characterization of Cu-α_3D**H$_3$** can help understand how distortions in the site may affect how copper binds and indicate how its reactivity is altered with respect to hydrolytic catalysis and the TRI scaffold.

In conclusion, this chapter present preliminary work on the nitrite reductase activity of α_3D**H$_3$** and has demonstrated that this is the best nitrite reductase model ever reported. While further characterization is warranted, these initial investigations give some insight into designing peptides capable of nitrite reductase activity. The histidine ligands present in α_3D**H$_3$** are introduced at the same layer as the cysteine ligands in α_3D**IV**, and between these two constructs Cd(II), Hg(II), Pb(II), Zn(II), Cu(I/II), and Fe(II/III) can all be bound at this site.[10,23] In particular, the size of the heavy metals Cd, Hg, and Pb suggest that this site has some level of flexibility and space, such that when a harder metal, like Cu(II) is bound, a substrate may bind for a catalytic reaction. The initial design for α_3D**IV** indicated that this side of the protein is more flexible,[23] and thus more amenable to substitutions. However, moving this site to the other end of the protein or further towards the center of the bundle should provide less space for water at the type 2 copper site and leave the region open for an electron transfer site, such as that described in Chapter 3. This is one potential route to a multi-site, multi-functional protein. Overall, these studies confirm that the α_3D scaffold is capable of accommodating redox catalysis and that this system could be a departure for further development into a multifunctional protein.

References:

(1) Yu, F. De novo designed metallopeptides with a type 2 copper center: a structural and functional model for nitrite reductase. Ph.D. Thesis. University of Michigan. Ann Arbor, MI, 2014.

(2) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chem. Rev.* **1996**, *96*, 2239–2314.

(3) Chufán, E. E.; Prigge, S. T.; Siebert, X.; Eipper, B. A.; Mains, R. E.; Amzel, L. M. *J. Am. Chem. Soc.* **2010**, *132*, 15565–15572.

(4) Libby, E.; Averill, B. A. *Biochem. Biophys. Res. Comm.* **1992**, *187*, 1529–1535.

(5) Tocheva, E. I.; Rosell, F. I.; Mauk, A. G.; Murphy, M. E. P. *Science* **2004**, *304*, 867–870.

(6) Solomon, E. I.; Baldwin, M. J.; Lowery, M. D. *Chem. Rev.* **1992**, 521–542.

(7) Tegoni, M.; Yu, F.; Bersellini, M.; Penner-Hahn, J. E.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A* **2012**, *109*, 21234–21239.

(8) Jacobson, F.; Pistorius, A.; Farkas, D.; De Grip, W.; Hansson, O.; Sjölin, L.; Neutze, R. *J. Biol. Chem.* **2007**, *282*, 6347–6355.

(9) Yu, F.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2013**, *135*, 18096–18107.

(10) Cangelosi, V. M.; Deb, A.; Penner-Hahn, J. E.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7900–7903.

(11) Silverman, D. N.; Lindskog, S. *Acc. Chem. Res.* **1988**, *21*, 30–36.

(12) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Nat Chem* **2012**, *4*, 118–123.

(13) Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayyum, H.; Pecoraro, V. L. *Chem. Rev.* **2014**, *114*, 3495–3578.

(14) Yatsunyk, L. A.; Rosenzweig, A. C. *J. Biol. Chem.* **2007**, *282*, 8622–8631.

(15) Xiao, Z.; Donnelly, P. S.; Zimmermann, M.; Wedd, A. G. *Inorg. Chem.* **2008**, *47*, 4338–4347.

(16) Miras, R.; Morin, I.; Jacquin, O.; Cuillel, M.; Guillain, F.; Mintz, E. *J. Biol. Inorg. Chem.* **2008**, *13*, 195–205.

(17) Alies, B.; Badei, B.; Faller, P.; Hureau, C. *Chem. Eur. J.* **2011**, *18*, 1161–1167.

(18) Ferroni, F. M.; Guerrero, S. A.; Rizzi, A. C.; Brondino, C. D. *J. Inorg. Biochem.* **2012**, *114*, 8–14.

(19) Chen, P.; Bell, J.; Eipper, B. A.; Solomon, E. I. *Biochemistry* **2004**, *43*, 5735–5747.

(20) Kau, L. S.; Spira-Solomon, D. J.; Penner-Hahn, J. E.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1987**, *109*, 6433-6442.

(21) Averill, B. A. *Chem. Rev.* **1996**, *96*, 2951–2964.

(22) Berks, B. C.; Ferguson, S. J.; Moir, J.; Richardson, D. J. *Biochim. Biophys. Acta* **1995**, *1232*, 97–173.

(23) Chakraborty, S.; Kravitz, J. Y.; Thulstrup, P. W.; Hemmingsen, L.; DeGrado, W. F.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2011**, *50*, 2049–2053.

Chapter 6. Conclusions

Protein design allows us to test the realms of possibility for particular structures and functions. Initially conceived as a method to understand structure-function relationships, the field has grown beyond just the creation of an α-helix or a β-sheet to encompass functional catalytic mimics,[1-4] proteins that can interface with cellular machinery,[5] and engineered quaternary structures to create structural or catalytic arrays of proteins[6,7]. Since metals bound to proteins often have known geometry and ligands, they are attractive targets for protein design because they provide a fixed point of reference for required structural features and accessible characterization methods. The field has moved from simple incorporation of native-like metal sites into both *de novo* and native scaffolds to functional constructs that allow us to test basic hypotheses about what elements are required for a particular function, effectively decoupling the evolutionary history of a site or protein from its function.

The Pecoraro group has a long history of investigating thiolate sites in the TRI family.[8-11] This modular scaffold is made up of a heptad repeat to create a highly stable α-helical motif that self-associates above pH 5.5 to form a three-stranded coiled coil. This work has allowed for a better interpretation of heavy metal ion homeostasis by providing tractable models for spectroscopic features and kinetics.[12-15] More recently this work has broadened to encompass hydrolytic[1,16,17] and redox catalysts[18,19]. This could be seen as building up a molecular toolbox with which one can construct a multi-site, multi-function protein from well-understood units. However, the TRI scaffold has several limitations, the most important of which is that it is three-fold symmetric. While methods are being developed to address this issue, the adoption of single-stranded scaffolds like α_3D presents another attractive strategy towards more complex designed proteins. Three-helix bundles like α_3D are formed by a single strand folding into an antiparallel α-helical bundle. Expanding the previous work into α_3D is not only a test for how broadly applicable some of the spectroscopic models based on TRI are, but can also give some insight into whether the antiparallel or single-stranded nature of α_3D fundamentally changes how metal

sites can be constructed in such a scaffold. Moving forward, understanding how metal sites can be built into a scaffold like α_3D lays the groundwork to explore asymmetric sites, sites with higher protein ligand coordination, and multifunctional constructs made up of one or more of these elements.

In this thesis, I have described single-stranded three helix bundles with either thiolate-rich sites for spectroscopic characterization and electron transfer, or histidine-rich sites for redox catalysis. In Chapter 2, I characterized these constructs with respect to cadmium coordination and applied them to understand ongoing debates about how heavy metals are coordinated in metal sensing proteins. In the absence of a structure of CadC bound to Cd(II), these peptides provide insight into how subtle changes in coordination can produce different spectral features and allow us to draw conclusions as to how Cd(II) is coordinated to this metalloregulatory protein. In Chapter 3, I described experiments showing that when one of these constructs is substituted with iron, it is an excellent spectroscopic model for rubredoxin. This work describes the successful incorporation of a rubredoxin site in the largest deviation from rubredoxin consensus structure ever reported. This is also the most extensive characterization of designed rubredoxin sites undertaken. In Chapter 4, I extended this work to study electron transfer reactivity in designed proteins with the synthesis of a ruthenium trisbipyridine chromophore that can be conjugated to designed peptides using exterior cysteine residues. The kinetics of electron transfer from a ruthenium attached to a terminal cysteine to the metal center were examined and it was found that given the distance of this transfer, a tyrosine in the intervening space is essential for fast electron transfer. Several α_3D derivatives were examined and the tyrosine in the 70^{th} position was found to be redox active in all derivatives where it was present. These findings reveal that these designed proteins conform to similar behavior as seen in native proteins and suggest that when engineering electron transfer sites that are farther than ~ 14 Å, the incorporation of redox active residues to act as relays can sustain much higher rates of electron transfer. Finally, in Chapter 5, I described studies in which a peptide with carbonic anhydrase activity was substituted with Cu(II) to carry out redox catalysis of nitrite reduction. This work confirms that redox catalytic sites can be incorporated into the α_3D scaffold and is a proof of concept for redox catalytic incorporation. Furthermore, the data provide some insight into how the antiparallel nature of α_3D may affect the reduction potential, the binding of Cu(I), and the catalytic efficiency of nitrite reduction.

These studies can form the basis for incorporating more than one redox active site in a single scaffold. The sites can be studied separately and later incorporated together in a unified construct. From work on designing tetrathiolate sites described in Chapter 3, I have shown that elements like chelate motifs strongly help to stabilize the site and increase binding efficiency in both redox states. A major improvement of the site would be to increase the stability to carry out a greater number of reduction-oxidation cycles. The irreversible oxidation process that occurs, which limits the number of cycles, is likely tied to solvent access. Thus, during reduction/oxidation cycles the oxidized iron could interact with the cysteine thiolates to generate a disulfide. While this process could be reversed, if water is nearby it could then coordinate the iron and in the presence of air, generate insoluble hydroxides, resulting in a slow but irreversible oxidation process. Incorporating bulkier, more non polar residues could reduce the solvent access, a situation that should be measurable by the observation of lower proportions of CdS_3O in the ^{111m}Cd PAC and by being able to cycle longer between the Fe(II) and Fe(III). The bulkier side chains should also have the effect of modulating the reduction potential, particularly if polar residues near the iron are substituted for bulky, nonpolar side chains. The incorporation of these residues would also more closely mimic the environment in rubredoxin itself, as Val and Leu residues near the iron active site have been shown to be important in protecting the site from water. Another method of increasing the hydrophobicity of the site and limiting solvent access is to create α_3D**IV**-F31C. This peptide introduces a chelate motif where the second chelating residue is further towards the center of the bundle rather than towards the edge of the helix, while the layer of coordinating cysteines from α_3D**IV** is retained (Figure 6.1). The result is that the site is moved further towards the center of the peptide with Leu21 in place, providing more shielding from the solvent.

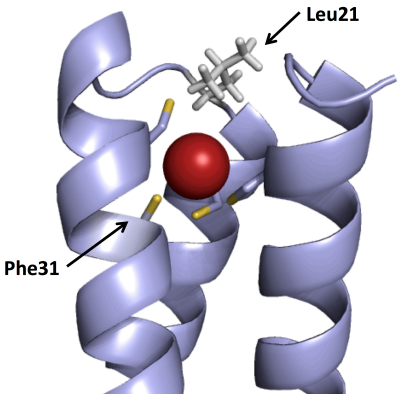


Figure 6.1 PyMol representation of α_3D**IV**-F31C, which moves the iron site towards the middle of the helices and retains Leu21 for shielding from solvent. Based on PDB: 2MTQ

Furthermore, the tolerance of this site to localization within the three-helix bundle should be explored. So far, all constructs within α_3D have been generated at the same layer of amino acids. To incorporate more than one site, other areas of the protein need to be explored for their ability to accommodate a metal binding site as well as the metal site's tolerance to incorporation in other areas of the protein. When α_3D**IV** was first designed, four layers of hydrophobic residues were identified as possible locations for incorporation of a tris-thiolate site. These sites can be converted into candidate tetrathiolate sites with chelate motifs for iron binding and electron transfer (Figure 6.2). Moving the site towards the center of the protein (site **II** or **III**) is another way to further shield the site from water. In removing stabilizing nonpolar interactions from the center of the bundle, compensating mutations may need to be made to accommodate the creation of the site and to promote proper folding. In L21C (**IV**), the site exists at the C terminal end of the protein, and moving the metal towards the center of the protein will likely decrease the flexibility of the site. Based on analysis of the structure of α_3D, a gradient of dynamic behavior

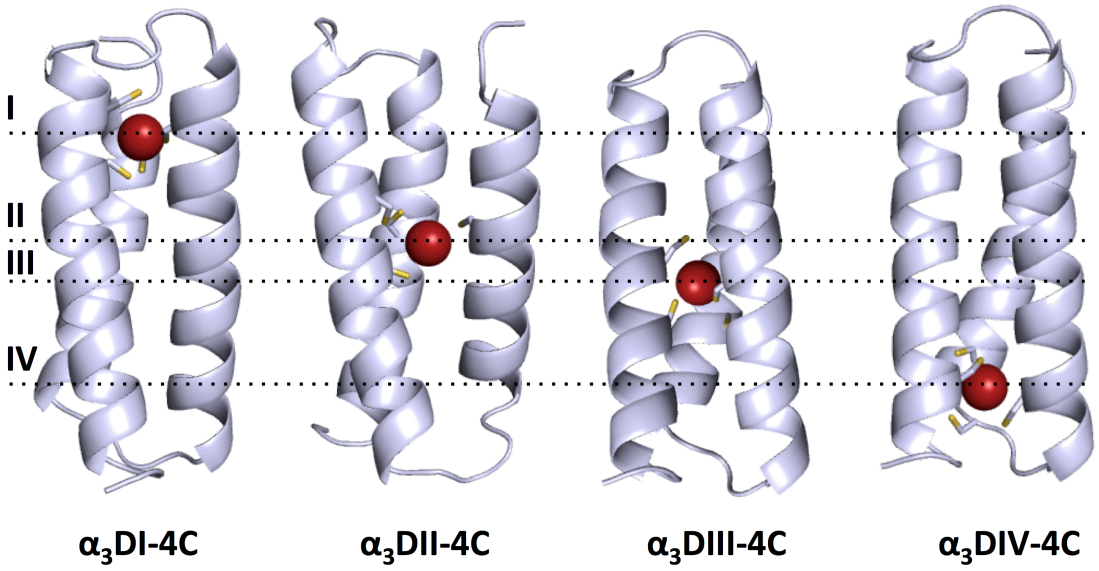


Figure 6.2 Representation of four layers where tetrathiolate sites incorporating a chelate motif can be placed along α_3D, as originally designed for α_3DIV. α_3DIV-4C is α_3DIV-L21C. Based on PDB: 2MTQ

and malleability is predicted moving from the N terminus to the C terminus. This suggests that the sites at the C terminus are more tolerant to mutation. Thus, site **III** would be a good candidate for incorporating a second site. However, this could place the metals, at sites **III** and **IV**, as close as 5 Å apart, which could result in coupling the two redox sites or complicating the incorporation of two metals. Thus, sites **II** and **I** should also be explored for their ability to

incorporate a metal-binding site. The proteins for tetrathiolate coordination at these sites have already been designed and the plasmids are in the lab.

In the case of incorporating a new site at position **I** or **II** in an eventual two-site construct, the sites will be separated by up to 20 Å as measured between sites **I** and **IV** (Figure 6.3). In this case, the lessons from studying the kinetics of electron transfer in α_3D derivatives can be applied to ensure efficient electron transfer between the two sites. This work was a proof of principle for

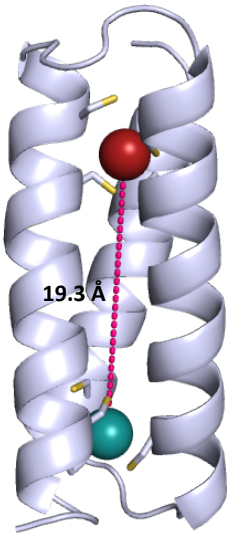


Figure 6.3 Estimate of distance between site **I** and site **IV** is ~19 Å. PDB: 2MTQ

measuring electron transfer rates between a ruthenium bipyridine chromophore and a bound metal. This work should be taken further and the electron transfer rates for different locations and distances in α_3D should be measured. Plasmids with cysteine residues to append the ruthenium chromophore at positions 2, 55, and 61 have already been designed and are in the lab. However, based on this work it is likely that the incorporation of a redox-active amino acid residue can facilitate electron transfer over longer distances, much as in seen in native proteins. However, a particular challenge in this is the driving force of the reaction. Redox-active residues such as tryptophan and tyrosine have high reduction potentials and kinetics that are complicated by the coupling of proton transfer with electron transfer. Between two metal sites at typical potentials found in proteins, the driving force of the single step to the tyrosine (or tryptophan) is uphill. The potential of tyrosine at pH 7 is 0.93 V vs. NHE, which drops to 0.7 V vs. NHE if deprotonated.[20,21] In native proteins, single uphill steps are often observed as part of a multi-step process in which the overall ΔG is negative. In these cases the kinetics of the electron transfer overcome the negative driving force through proximity between electron transfer sites and the overall ΔG of the process, although the electron transfer is most efficient if the potential of the

relay is within 200 mV of the acceptor.[22] The systems studied in Chapter 4 vary slightly from this regime in that the entire reaction has such a high driving force (Ru(III) is ~ 1.5V vs. NHE) that careful matching of reduction potentials is not necessary for rapid kinetics. However, in a two site construct with lower driving force, hydrogen-bonding residues can be incorporated to address the problem of protonation state, as nearby histidines are thought to participate in a "rocking" motion at tyrosines Z and D in PSII.[23] This rocking motion effectively deprotonates the tyrosine radical during electron transfer an re-protonates it after reduction of the tyrosine. Thus, placing hydrogen bonding residues near a tyrosine intended as a relay is crucial to controlling the potential of the tyrosine so as to facilitate fast electron transfer.

I have detailed ways to test the site dependence of the rubredoxin site, but the same principle of layers for substitution can be applied to α_3D**H$_3$**. The histidine residues in this construct were designed at the same layer as in α_3D**IV** and thus the layers at positions **I**, **II**, and **III** are also possible candidates for alternate locations or a tris-histidine site. An important consideration in this case is that histidine is much larger than cysteine so the tolerance of the protein towards mutation becomes crucial. In this sense, retaining the tris-histidine site at layer **IV** and engineering the residues around the site to improve the catalytic activity makes the most sense as this end of the protein is more tolerant to mutation. Also worth considering is that some significant reengineering may need to be done to stabilize the scaffold towards these mutations. Further consideration should be also given to the composition of these other positions. As α_3D was originally derived from CoilSer, which also follows the heptad repeat strategy, the concept of **a** and **d** sites (Figure 1.12) can also be applied. The **IV** position is an analogous substitution to previously reported TRI substitutions—Leu to X substitutions at an **a** site. However, due to the antiparallel nature of α_3D, substitutions at **I**, **II**, and **III**, all involve at least one **d** site residue and substitutions of residues such as isoleucine. Each turn in an α-helix comprise of 3.6 residues, meaning that residues at positions *i* and *i+3* are not directly on top of each other. The **a** vs. **d** site distinction is important in that it affects the direction that the side chain points. This has been shown previously in the TRI series to affect the coordination and the spectroscopy of heavy metal ions.[12,24,25] For distorted sites, where over the course of a catalytic reaction different conformations are sampled, this may be advantageous. As of now, the affect of **a** vs. **d** substitution in antiparallel bundles is unknown.

As a thought experiment, I explored a possible method to engineer some of these considerations α_3D, where the two metal binding sites are located at positions **IV** and **I** (Figure 6.4). In this model the rubredoxin site is located at **I** and the nitrite reductase site is located as in α_3DH_3 where Ile14 and Phe31 form the hydrogen bonded pair. An alternative would be to incorporate a hydrogen bonded tyrosine and histidine at Leu11 and Phe38, which would move the relay towards the rubredoxin site at **I**. The potentials of the two sites (-75 mV and 200 mV) are well-matched for electron transfer. However, this does not take into account how to control which metal binds to which site. Such a mixed metal system is many steps of design beyond where our abilities are now. Copper(I) has extraordinarily high affinity for cysteine residues and so would displace the iron bound to a rubredoxin site. A more tenable solution is to pair a copper

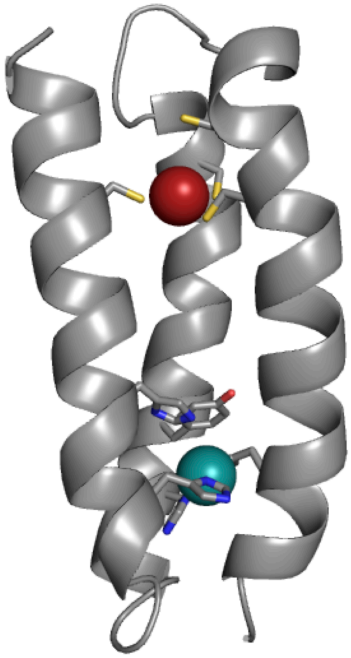


Figure 6.4 Model for two site construct with rubredoxin site at **I** and nitrite reductase at **IV** with I14H and F31Y to form hydrogen-bonded tyrosine. (PDB: 2A3D)

electron transfer site with a copper catalytic site and vice versa for iron for a site like that in cysteine dioxygenase, in which iron is coordinated by three histidines, as this strategy removes the problem of metal selectivity. Within the context of what has been done in the lab, pairing two copper sites is closest to realization between the work detailed here in Chapter 5 and the work of a fellow graduate student, Jefferson Plegaria, on copper-based electron transfer proteins. Focusing on a copper-based system has the added advantage of examining the fascinating functional switch that occurs in type 2 sites as some type 2 copper sites are catalytic, while others function as electron transfer sites. The question of whether a certain function is selected for by

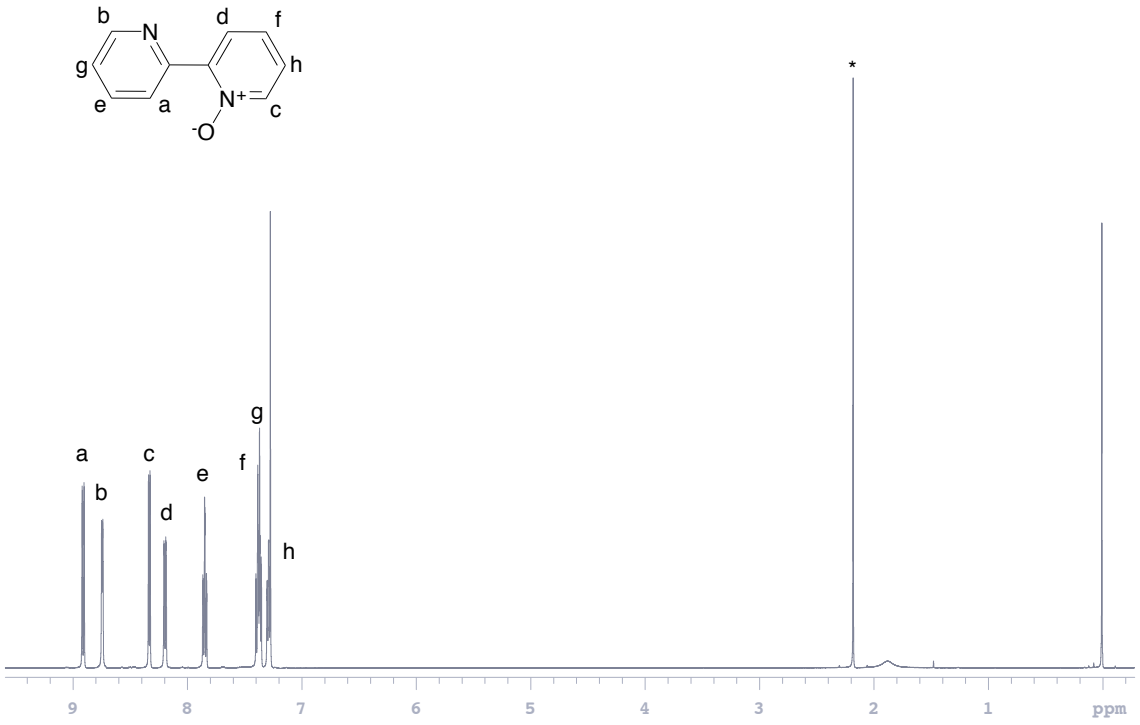
the reduction potential of the site or its solvent (and substrate) accessibility could be addressed directly in such a multi-site construct.

Overall, this thesis describes significant advances towards features for a designed multifunctional redox protein as well as addressing current questions in the fields of metal homeostasis, protein design, and electron transfer. The design and characterization of tetrathiolate sites can be applied to a diverse variety of fields, and in this thesis I have addressed their applicability to metal ion homeostasis and electron transfer. I have described an excellent model for CadC coordination to Cd(II) that suggests a dynamic, non-homogenous coordination sphere. This construct can be extended to other metals that bind at the cysteine site in ArsR/SmtB metal sensors. I have also applied these tetrathiolate designs to the electron transfer protein rubredoxin, which has resulted in a successful rubredoxin mimic that has the largest deviation from the consensus metal binding motif in rubredoxins and zinc fingers. This demonstrates that binding at β-turns is not required to achieve the same electronic environment for the iron and suggests that chelate motifs and the microenvironment around the site are important for stability. The work describing electron transfer kinetics represents the first detailed study of electron transfer kinetics through a *de novo* designed peptide scaffold. Finally, the characterization of copper nitrite reductase activity suggests that the topology of a protein fold should be matched to the requirements of the metal when designing catalytic sites. This model also has the greatest catalytic efficiency for a designed nitrite reductase reported to date. My doctoral work provides insight into designed proteins and their applications and serves as the basis for progression towards a unified multifunctional redox protein.

References:

(1) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Nat Chem* **2012**, *4*, 118–123.
(2) Cangelosi, V. M.; Deb, A.; Penner-Hahn, J. E.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 7900–7903.
(3) Rufo, C. M.; Moroz, Y. S.; Moroz, O. V.; Stöhr, J.; Smith, T. A.; Hu, X.; DeGrado, W. F.; Korendovych, I. V. *Nat Chem* **2014**, *6*, 303–309.
(4) Yu, F. Ph.D. Thesis. University of Michigan, Ann Arbor, MI. 2014.
(5) Anderson, J. L. R.; Armstrong, C. T.; Kodali, G.; Lichtenstein, B. R.; Watkins, D. W.; Mancini, J. A.; Boyle, A. L.; Farid, T. A.; Crump, M. P.; Moser, C. C.; Dutton, P. L. *Chem. Sci.* **2014**, *5*, 507–514.
(6) Oohora, K.; Onoda, A.; Kitagishi, H.; Yamaguchi, H.; Harada, A.; Hayashi, T. *Chem. Sci.* **2011**, *2*, 1033.
(7) Radford, R. J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2009**, *131*, 9136–9137.
(8) Tebo, A. G.; Pecoraro, V. L. *Curr. Op. Chem. Biol.* **2015**, *25C*, 65–70.
(9) Plegaria, J. S.; Dzul, S.; Zuiderweg, E. R. P.; Stemmler, T. L.; Pecoraro, V. L. *Biochemistry* **2015**.
(10) Plegaria, J. S.; Pecoraro, V. L. *Is. J. Chem.* **2015**.
(11) Yu, F.; Cangelosi, V. M.; Zastrow, M. L.; Tegoni, M.; Plegaria, J. S.; Tebo, A. G.; Mocny, C. S.; Ruckthong, L.; Qayyum, H.; Pecoraro, V. L. *Chem. Rev.* **2014**, *114*, 3495–3578.
(12) Matzapetakis, M.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2005**, *127*, 18229–18233.
(13) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2011**, *133*, 239–251.
(14) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. *Chem. Eur. J.* **2009**, *15*, 3761–3772.
(15) Chakraborty, S.; Iranzo, O.; Zuiderweg, E. R. P.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2012**, *134*, 6191–6203.
(16) Zastrow, M. L.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2013**, *135*, 5895–5903.
(17) Zastrow, M. L.; Pecoraro, V. L. *Biochemistry* **2014**, *53*, 957–978.
(18) Tegoni, M.; Yu, F.; Bersellini, M.; Penner-Hahn, J. E.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A* **2012**, *109*, 21234–21239.
(19) Yu, F.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2013**, *135*, 18096–18107.
(20) Harriman, A. *J. Phys. Chem.* **1987**, *91*, 6102–6104.
(21) Warren, J. J.; Winkler, J. R.; Gray, H. B. *FEBS Lett.* **2012**, *586*, 596–602.
(22) Gray, H. B.; Winkler, J. R. *Chem. Phys. Lett.* **2009**, *483*, 1–9.
(23) Migliore, A.; Polizzi, N. F.; Therien, M. J.; Beratan, D. N. *Chem. Rev.* **2014**, *114*, 3381–3465.
(24) Lee, K.-H.; Matzapetakis, M.; Mitra, S.; Marsh, E. N. G.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2004**, *126*, 9178–9179.
(25) Pecoraro, V. L.; Peacock, A.; Iranzo, O. Long, E., Ed.; ACS Symposium Series: *Bioinorganic Chemistry.* Washington, DC, 2009. pp. 183-197

Appendix I



All assignments based on literature values: see Baron, A *et al. Inorg. Chem.* **2012**
Spectra taken in $CDCl_3$ = 7.26 ppm, unless noted
* = solvent

Figure A.1 2,2'-bipyridyl-N-oxide

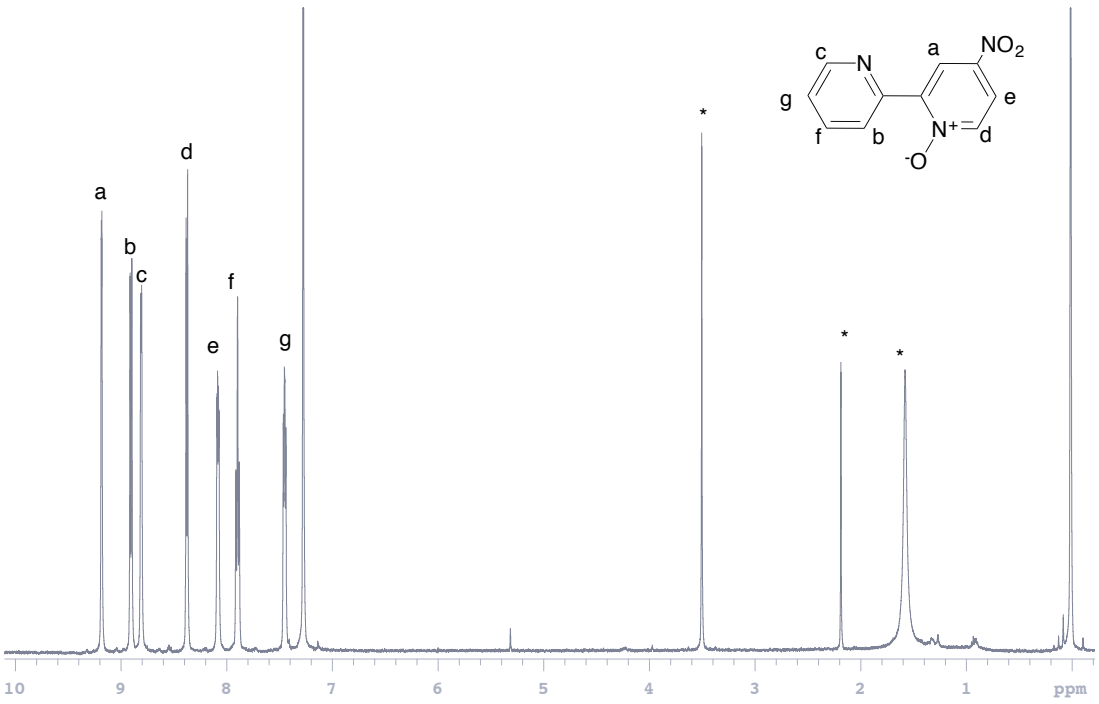


Figure A.2 4'-Nitro-2,2'-bipyridyl-N-oxide

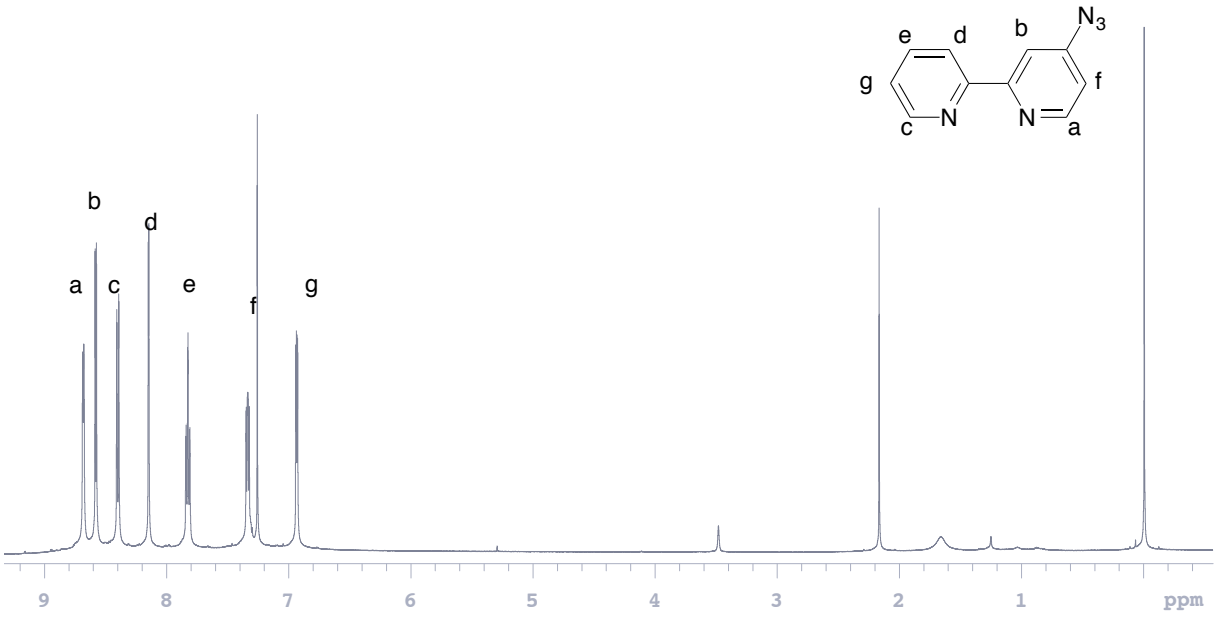


Figure A.3 4'-azido-2,2'-bipyridyl-N-oxide

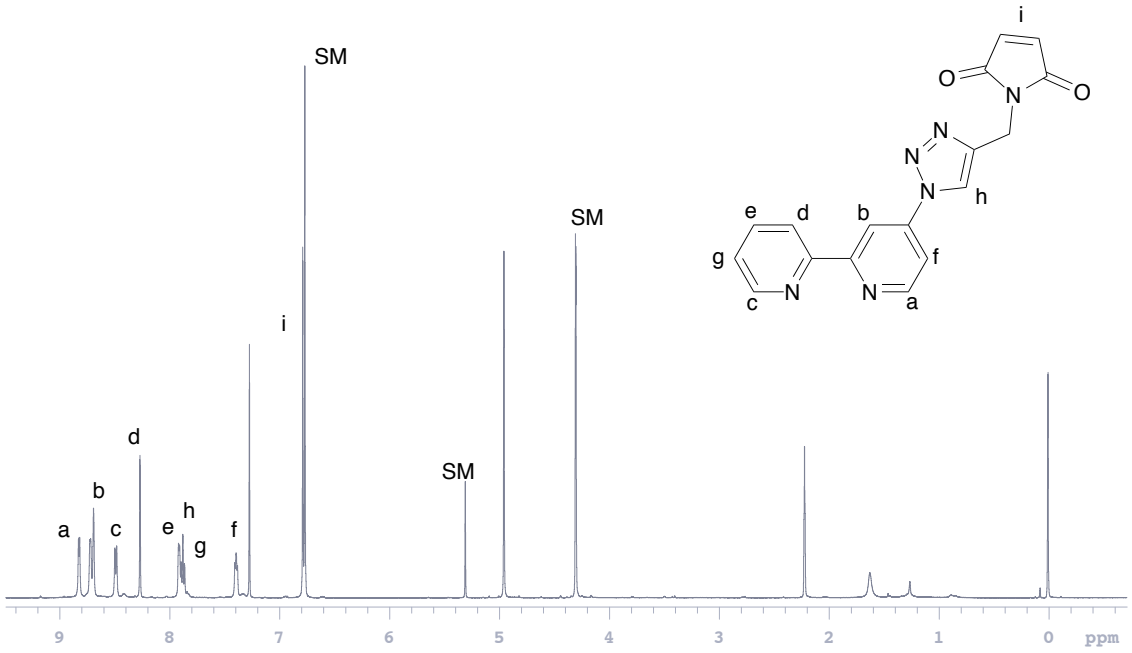
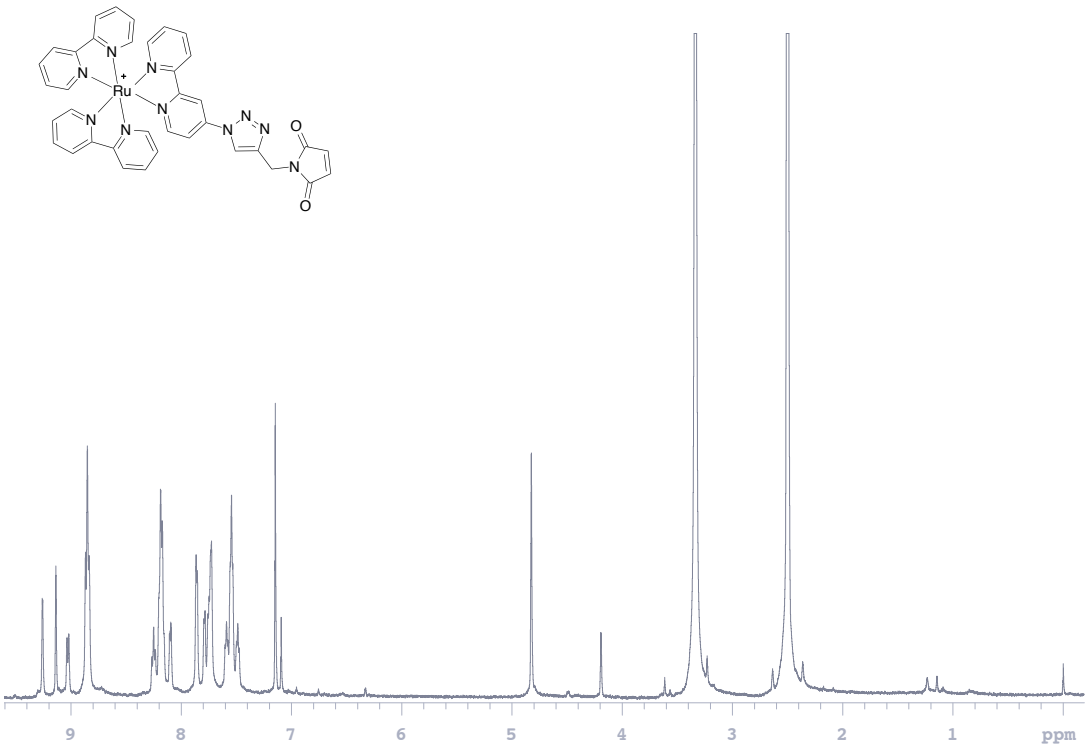


Figure A.4 Bpy-mal: (1-((1-([2,2'-bipyridin]-4-yl)-1*H*-1,2,3-triazol-4-yl)methyl)-1*H*-pyrrole-2,5-dione)



Spectra taken in DMSO-d_6

Figure A.5 Ru-bpymal