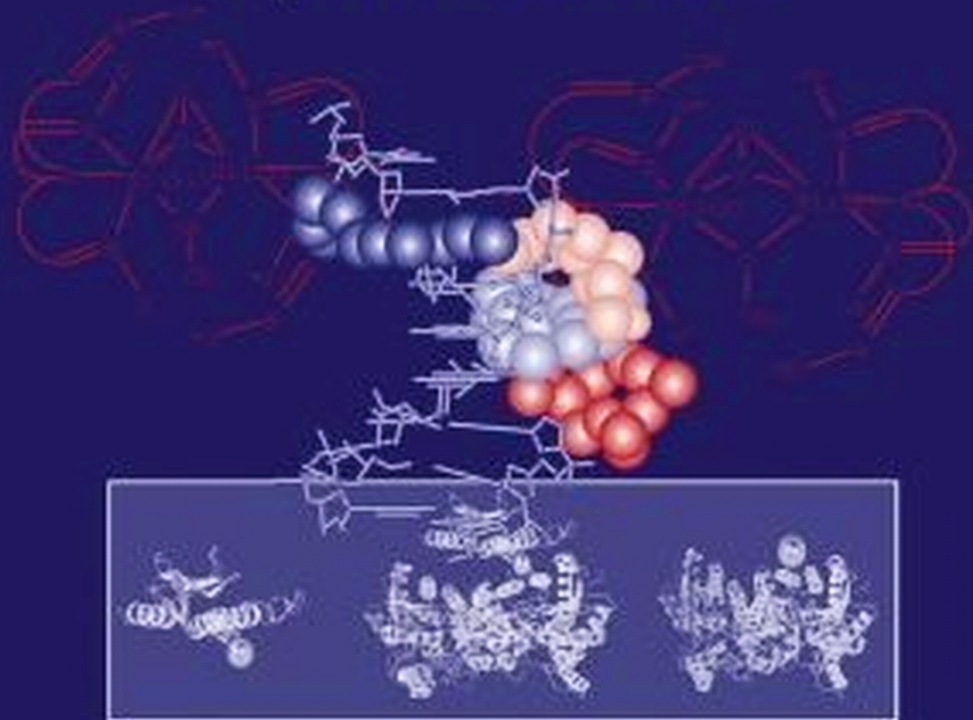


ACS SYMPOSIUM SERIES 1012

# Bioinorganic Chemistry

## Cellular Systems and Synthetic Models



EDITED BY

Eric C. Long and Michael J. Baldwin

# **Bioinorganic Chemistry**



ACS SYMPOSIUM SERIES **1012**

**Bioinorganic Chemistry**  
**Cellular Systems and Synthetic  
Models**

**Eric C. Long, Editor**

*Indiana University–Purdue University Indianapolis  
(IUPUI)*

**Michael J. Baldwin, Editor**

*University of Cincinnati*

**Developed from a symposium sponsored by the  
Central Region of the American Chemical Society at the 39<sup>th</sup>  
Annual Meeting of the ACS Central Region (CERMACS 2007),  
Covington, Kentucky, USA,  
May 20-23, 2007**



American Chemical Society, Washington, DC



ISBN: 978-0-8412-6975-0

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Copyright © 2009 American Chemical Society

Distributed by Oxford University Press

All Rights Reserved. Reprographic copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Act is allowed for internal use only, provided that a per-chapter fee of \$40.25 plus \$0.75 per page is paid to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. Republication or reproduction for sale of pages in this book is permitted only under license from ACS. Direct these and other permission requests to ACS Copyright Office, Publications Division, 1155 16th Street, N.W., Washington, DC 20036.

The citation of trade names and/or names of manufacturers in this publication is not to be construed as an endorsement or as approval by ACS of the commercial products or services referenced herein; nor should the mere reference herein to any drawing, specification, chemical process, or other data be regarded as a license or as a conveyance of any right or permission to the holder, reader, or any other person or corporation, to manufacture, reproduce, use, or sell any patented invention or copyrighted work that may in any way be related thereto. Registered names, trademarks, etc., used in this publication, even without specific indication thereof, are not to be considered unprotected by law.

PRINTED IN THE UNITED STATES OF AMERICA

# Foreword

The ACS Symposium Series was introduced in 1974 to provide a mechanism for publishing symposia quickly in book form. The purpose of the series is to publish timely, comprehensive books developed from ACS sponsored symposia based on current scientific research. Occasionally, books are developed from symposia sponsored by other organizations when the topic is of keen interest to the chemistry audience.

Before agreeing to publish a book, the proposed table of contents is reviewed for appropriate and comprehensive coverage and for interest to the audience. Some papers may be excluded to better focus the book; others may be added to provide comprehensiveness. When appropriate, overview or introductory chapters are added. Drafts of chapters are peer-reviewed prior to final acceptance or rejection, and manuscripts are prepared in camera-ready format.

As a rule, only original research papers and original review papers are included in the volumes. Verbatim reproductions of previously published papers are not accepted.

**ACS Books Department**

## Table of Contents

### Preface

*Eric C. Long and Michael J. Baldwin* ix-x

### Cellular Systems

#### **1 Iron Sulfur Cluster Biosynthesis: Scaffold and Donor Proteins, and Mechanistic Insights**

*J. A. Cowan* 3-16

#### **2 Molecular Interaction between Frataxin and Ferrochelatase During Heme Assembly: Frataxin's Role as a Potential Iron Chaperone During Heme Biosynthesis**

*Kalyan C. Kondapalli, Andrew Dancis, and Timothy L. Stemmler* 17-30

#### **3 Biosynthesis and Regulation of the Heme: A Biosynthetic Pathway**

*Eric L. Hegg* 31-46

#### **4 Assembly of Cytochrome c Oxidase: Synthesis and Insertion of the Metal Cofactors**

*Zhihong Wang and Eric L. Hegg* 47-61

#### **5 New Approaches to Analyzing the Site Selectivities and Crystal Structures of DNA Targeted Metal Complexes: Analysis of the Antitumor Agent Bleomycin**

*Eric C. Long, Millie M. Georgiadis, Kristie D. Goodwin, and Mark A. Lewis* 63-80

#### **6 Zn(II) Homeostasis in *E. coli***

*Thusitha Gunasekera, J. Allen Easton, Stacy A. Sugerbaker, Lindsey Klingbeil, and Michael W. Crowder* 81-95

### Synthetic Models

#### **7 N<sub>2</sub>S<sub>3</sub>X-Fe Models of Nitrile Hydratase**

*Martin G. O'Toole and Craig A. Grapperhaus* 99-113

#### **8 Studies into the Metal Chemistry of the Carbaporphyrinoids: Insights into the Biological Choice of Porphyrin**

*Christopher J. Ziegler* 115-131

#### **9 Bioinspired Aerobic Substrate Oxidation: A Ni(II)-Oximate Catalyst that Parallels Biological Alcohol and Amine Oxidation Chemistry**

*Michael J. Baldwin, Jeanette A. Krause, Michael J. Goldcamp, Micheal Haven, Sara E. Edison, and Leah N. Squires* 133-150

#### **10 Inorganic Models for Two-Electron Redox Chemistry in Biological Systems: Ligand-Bridged Molybdenum and Tungsten Dimers**

*Franklin A. Schultz, Richard L. Lord, Xiaofan Yang, and Mu-Hyun Baik* 151-166

**11 Metal-Mediated Peptide Assembly: From Discrete Molecular Species to Large-Scale Morphologies**

*Jing Hong, Olesya A. Kharenko, Mikhail V. Tsurkan, and Michael Y. Ogawa* 167-182

**12 Understanding the Biological Chemistry of Mercury Using a de novo Protein Design Strategy**

*Vincent L. Pecoraro, Anna F. A. Peacock, Olga Iranzo, and Marek Luczkowski* 183-197

**13 Understanding Oxotransferase Reactivity in a Model System Using Singular Value Decomposition Analysis**

*Brian W. Kail, Charles G. Young, Mitchell E. Johnson, and Partha Basu* 199-217

**14 DNA Minor Groove Recognition by Ni(II)· and Cu(II)·Gly-Gly-His Derived Metallopeptides: Models of Protein and Natural Product DNA Recognition**

*Eric C. Long, Ya-Yin Fang, and Mark A. Lewis* 219-241

**Indexes**

**Author Index** 245

**Subject Index** 247-253



# Preface

Bioinorganic chemistry has diversified and grown considerably as an interdisciplinary chemical science during the past several decades, paralleling an ever-increasing appreciation of the impact of inorganic elements on the fields of biology, medicine, biochemistry, and environmental science. Indeed, while many new bioinorganic research avenues have opened to investigation, other more “traditional” lines of research have shifted focus or broadened their original scope in the pursuit of deeper levels of understanding. For example, while studies of isolated metalloenzymes remain a mainstay of the field, many studies now seek also to determine the impact of a metalloenzyme, metal ion, or carrier thereof within a larger biological context involving multiple biomacromolecular “players”. Included here are studies of cellular systems such as metal ion influenced biosynthetic pathways, trafficking, exogenous drug activation, and regulatory phenomena. Similarly, in comparison to their predecessors, contemporary synthetic model systems that mimic metal ions in a biological environment have become increasingly sophisticated. Here, for example, our ability to synthesize ligands that are more protein- or peptide-like has increased dramatically, complementing the use of traditional low molecular weight organic ligands. In light of recent efforts and ongoing interest devoted to the bioinorganic chemistry of the two key areas highlighted above, *Bioinorganic Chemistry: Cellular Systems & Synthetic Models* has sought to provide an illustrative cross section of cutting edge research in these important pursuits. This book is intended not only to describe the individual efforts highlighted in each chapter, but also to showcase the evolution of two bioinorganic avenues of research from their traditional roots.

*Bioinorganic Chemistry: Cellular Systems & Synthetic Models* derives from a symposium entitled *Bioinorganic Chemistry* organized by one of us for the 39<sup>th</sup> Meeting of the Central Region of the American Chemical Society (CERMACS 2007) held in May 2007. The twelve participants invited to speak at this symposium were drawn from states within or adjacent to the area included in the ACS Central Region. The

excellent slate of talks presented during the symposium underscores the strength of the bioinorganic chemistry endeavor in the central part of the United States. Upon further perusal, this slate of talks was revealed to be a representative cross section of current work in the area of cellular systems and synthetic models. It included work ranging from metal ion homeostasis, the biosynthesis of metal-containing cellular components and the role of metal ions in drug structure, activation, and DNA targeting to synthetic models that impact our knowledge of catalysis, electron transfer, bio-related nanoassemblies and metal ion toxicity. We felt that the outstanding list of participants and the fascinating range of topics covered would be of interest to a much larger audience of chemists and biochemists than were able to attend the symposium, so this book project was born. While a few of the participants were unable to contribute chapters to this book, the majority of those participating in the symposium are represented in this volume. Additionally, several bioinorganic chemists who did not attend the meeting, but were from the same geographical area as the symposium participants, have been invited to contribute to this book, making it an even stronger representation of bioinorganic chemistry in the Central Region.

Overall, this volume is intended to represent the current state of bioinorganic chemistry research with regard to two areas of study: work devoted increasingly to the understanding of inorganic activities at the cellular level and also the development of ever more sophisticated model systems. These are two areas of traditional, yet continuously evolving interest to bioinorganic chemists from beginning graduate students to experienced and established researchers in the field.

**Eric C. Long**

Department of Chemistry & Chemical Biology  
Indiana University-Purdue University Indianapolis (IUPUI)  
Indianapolis, IN 46202  
eclong@iupui.edu

**Michael J. Baldwin**

Department of Chemistry  
University of Cincinnati  
Cincinnati, OH 45221  
baldwimj@email.uc.edu

## **Cellular Systems**



## Chapter 1

# Iron Sulfur Cluster Biosynthesis

## Scaffold and Donor Proteins, and Mechanistic Insights

J. A. Cowan

Evans Laboratory of Chemistry,  
The Ohio State University, 100 West 18<sup>th</sup> Avenue, Columbus, OH 43210

The structural and functional properties of the iron and sulfide donor proteins (frataxin and IscS/NifS, respectively) that mediate [2Fe-2S] cluster assembly on the scaffold protein ISU/IscU are reviewed in light of the developing mechanistic understanding of cluster biosynthesis. The structural information now available for each class of protein provides a context for understanding recent measurements of the kinetics of cluster assembly and the thermodynamics of iron binding and scaffold-partner interactions. The protein NFU is discussed as a mediator of persulfide bond cleavage to yield inorganic sulfide for cluster assembly.

Iron-sulfur proteins are widely distributed in nature and can be found in anaerobic, aerobic and photosynthetic bacteria, fungi, plants, and mammals. They serve a variety of biological roles, including electron transport, catalysis, structural, and sensory. Interest in the biological chemistry of Fe-S clusters has evolved from biological redox chemistry, to enzymology (aconitase and mutY) and gene regulatory function (IRP and soxR). Understanding the biosynthetic pathways for these complex cofactors forms part of a broader interest in the

biosynthesis of metal cofactors and the intracellular trafficking of metal ions and cofactors. Proteins that deliver discrete metal ions to target proteins, including copper, nickel, and manganese constitute a growing family of metal transporters (1). However, many proteins and enzymes contain complex metal cofactors that must themselves be the product of a biosynthetic assembly apparatus. Recent advances in this area include the characterization of operons that define gene products responsible for the biosynthesis of nitrogenase cofactors (2), heme (3), and iron-sulfur clusters (4).

While most soluble cellular iron is associated with an ill-defined "labile" iron pool, much of the iron that is used by the cell is found in the form of specific iron cofactors; especially as iron-sulfur clusters and hemes. The biosynthesis of iron-sulfur centers has only come of age within the past decade following the characterization of operons encoding cluster assembly genes by Dean and coworkers (4). In eukaryotes, homologous proteins are nuclear transcribed. While predominantly localized within the mitochondrial matrix, one group has reported cytosolic ISU (cluster assembly protein) and NFS (S-donor protein) in human cell lines (5). Alternative cytoplasmic cluster maturation proteins have also been proposed (6). Whether the cytosolic components represent an independent assembly apparatus is unclear, nevertheless, both models support a role for ISU-mediated assembly of [2Fe-2S] building blocks prior to delivery to a target protein (7-9).

Recent work has laid the foundation for a mechanistic understanding of iron sulfur cluster assembly that includes frataxin-mediated iron delivery to ISU/IscU (10) and subsequent delivery of sulfide via the cysteinyl persulfide bond of IscS/NifS (11, 12). The resulting cluster is then transferred to other cluster-dependent proteins. An alternative pathway (13), involving initial transfer of sulfur with subsequent uptake of iron, is precluded by the weak binding of iron to the persulfide labeled IscU/ISU (14). While iron-bound ISU has been converted to cluster-bound protein following delivery of sulfide (chemically or enzymatically) (14), reconstitution of sulfur-labeled ISU following addition of iron has not been established (13). Moreover, NFU-type proteins have been demonstrated to catalyze the reductive cleavage of the cysteinyl persulfide bonds of IscS-type proteins, yielding inorganic sulfide (11).

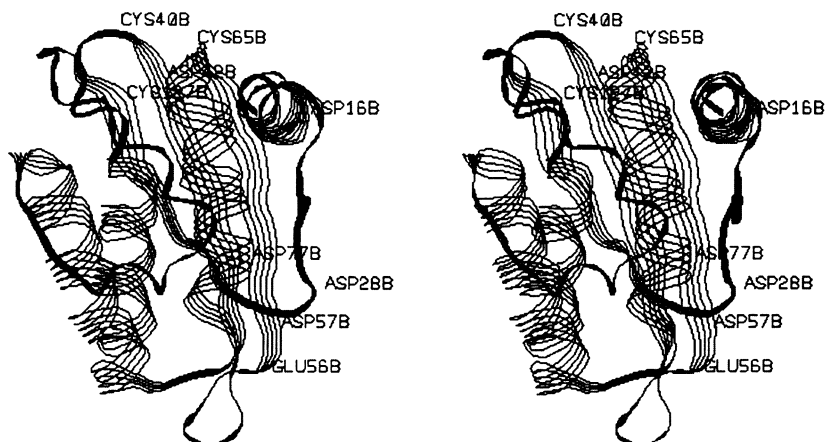
Throughout this work the term ISU will be used to denote the eukaryotic protein, or to refer to the general family of such proteins, while the term IscU will refer to the bacterial homologues. Also, the term IscS will generally denote the sulfur donor, since bacterial IscS and NifS proteins, and eukaryotic NFS, are very similar in sequence and function.

## Iron-Sulfur Cluster Assembly Protein ISU

ISU domains are among the most highly conserved in nature, spanning all three kingdoms of life. They appear to provide a platform for assembly of  $[2\text{Fe-2S}]^{2+}$  clusters prior to transfer to target apo proteins. Whether the cytosolic components represent an independent assembly apparatus is unclear. Nevertheless, both intra- and extra-mitochondrial Fe-S proteins appear to depend on mitochondrial components for the biosynthesis of their metallocofactors, with ISU (eukaryotes) or IscU (prokaryotes) serving as the principal scaffolding protein on which  $[2\text{Fe-2S}]$  building blocks are assembled prior to delivery to a target protein (8, 15). Inactivation of individual gene products typically leads to a build-up of cellular (or mitochondrial) iron and loss of enzymatic activity in certain cluster-dependent proteins. However, the importance of the ISU/IscU cluster scaffold protein is underscored by the retention of this gene in all organisms (16), and by the observation that double deletion of IscU homologs in *Saccharomyces cerevisiae* ( $\Delta\text{isu1}\Delta\text{isu2}$ ) is lethal (17). In bacteria the Isc operon is the principal Fe-S cluster assembly vehicle in the cell. In *E. coli* from which the entire *isc* operon was deleted, the activity of Fe-S proteins was found to be only 2–10% that in wild-type cells, with the residual activity ascribed to a contribution from the *SUF* system (an alternative operon that is utilized under conditions of oxidative stress to facilitate production of proteins that combat that stress) (18).

ISU-type proteins mediate cluster transfer to apo protein targets. Rate constants have been determined for cluster transfer from ISU to apo ferredoxin (Fd) for both *Homo sapiens* and *Schizosaccharomyces pombe* proteins, and cross reactions have also been examined (8, 15). Substitution of a key aspartate residue of ISU (D37) (Figure 1) is found to decrease the rate of cluster transfer by at least an order of magnitude (for wild-type *Hs* ISU cluster transfer to *Hs* apo Fd,  $k_2 \sim 540 \text{ M}^{-1}\text{min}^{-1}$ , relative to  $\sim 56 \text{ M}^{-1}\text{min}^{-1}$  for D37A ISU). This change in rate constant does not reflect any change in binding affinity of the ISU and Fd proteins. The pH dependencies of cluster transfer rates are similar for WT and D37A ISU, arguing against a role for Asp37 as a catalytic base, although evidence for general base catalysis mediating deprotonation of Cys from the apo target is supported by an observed  $\text{p}K_a$  of 6.9 determined from the pH profiles for both WT and D37A ISU. Such a  $\text{p}K_a$  value is at the lower limit for Cys and is common for solvent-accessible Cys thiols. The temperature dependence of the rate constant defining the cluster transfer reaction for wild type versus the aspartate derivative is distinct. Thermal activation parameters ( $\Delta H^*$  and  $\Delta S^*$ ) are

consistent with a solvent-accessible ISU-bound cluster, with desolvation as a principle barrier to cluster transfer. Experiments to determine the dependence of reaction rate constants on viscosity indicate cluster transfer to be rate-limiting. Fully oxidized cluster appears to be the natural state for transfer to target



*Figure 1. The crystal structure (chain B) of the ISU-type protein from Streptococcus pyogenes (Sp) (PDB ID 1SU0) shown in stereoview. Highlighted residues include Cys required for Fe-S cluster formation (C40, C65, C127) and highly conserved acidic residues (D16, D42, D77, E56, D57, and D28). Residue D42 is the homologous site for D37 found in human ISU.*

proteins. Reduced Fd does not readily reduce ISU-bound  $[2\text{Fe-2S}]^{2+}$  and does not promote cluster transfer to an apo Fd target.

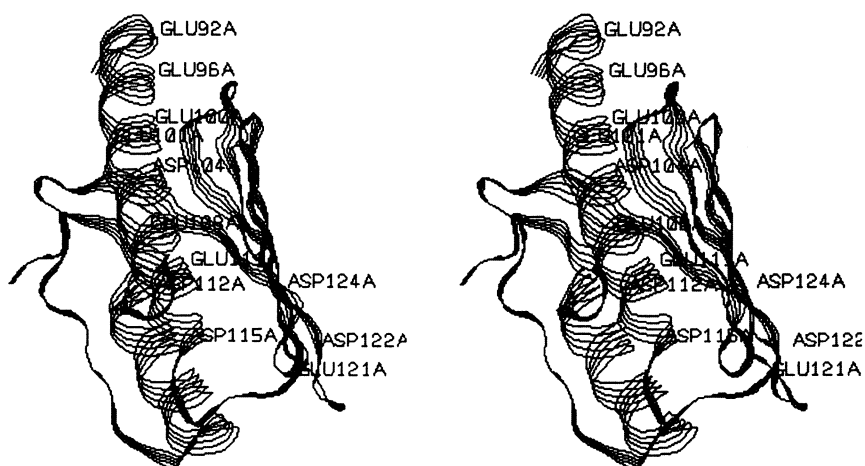
Important for the understanding of the functional properties of the iron-sulfur scaffold IscU is knowledge of the structure and dynamics of this protein class. Structural characterization of *Thermotoga maritima* IscU by CD (19) and high resolution NMR (20) yielded data indicating a high degree of secondary structure. However, the latter also revealed IscU to exist in a dynamic equilibrium between two or more distinct conformations, possibly existing in a molten globule state. A variety of experimental methods (19), including near-ultraviolet circular dichroism, 1-anilino-8-naphthalenesulfonic acid binding, free energy of unfolding, hydrodynamic radius measurements, and limited tryptic digestion suggested unusual dynamic behavior that is not fully consistent with typical protein states such as fully folded, fully unfolded, or molten globule. Interestingly, in those cases where an IscU structure has been determined crystallographically or by NMR, an adventitious bound metal appears to stabilize



a particular conformer. All of the available structural data are consistent with a structural model of a protein sampling multiple discrete tertiary conformations in which these structural transitions occur on a “slow” time scale, described by the term multiple discrete conformers. There is undoubtedly a wide spectrum of conformational lability across the family of IscU type proteins, depending on the activation barriers to conformational transition.

## Frataxin as an Iron Chaperone or Delivery Protein

Frataxin (Figure 2) has recently been identified as a likely physiological iron donor to ISU during assembly of the [2Fe-2S] cluster core (10, 21-23), and to ferrochelatase for the final step of heme biosynthesis (24, 25). Dysfunction of frataxin results in mitochondrial iron accumulation and oxidative damage to mitochondrial DNA (26, 27), as well as the potential to disrupt other cellular processes dependent on such iron cofactors. This is the first characterized example of an iron-delivery protein and it remains to be seen how relevant its functional chemistry will be, relative to cell metabolism across a spectrum of organisms. As noted later, the absence of sequence homologs in certain other organisms does not preclude the presence of structural and functional



*Figure 2. Crystal structure of human frataxin (PDB ID 1EKG) shown in stereoview. The substituted residues are highlighted with stick structures and were modified in the preparation of derivative A (E100A/E101A/D104A/E108A/E111A/D112A); derivative B (D115A/E121A/D122A/D124A); and derivative C (E92A/E96A/E100A/E101A).*

equivalents of frataxin. The bacterial protein IscA has also been suggested as an iron donor for cluster assembly on IscU (28), but the absence of critical control experiments make this difficult to assess. IscA has been better characterized as an alternative scaffolding protein that can assemble a [2Fe-2S] cluster and subsequently transfer it to a target protein such as ferredoxin (29, 30).

### Mapping Out a Role for Carboxylates in the Anionic Patch of Frataxin

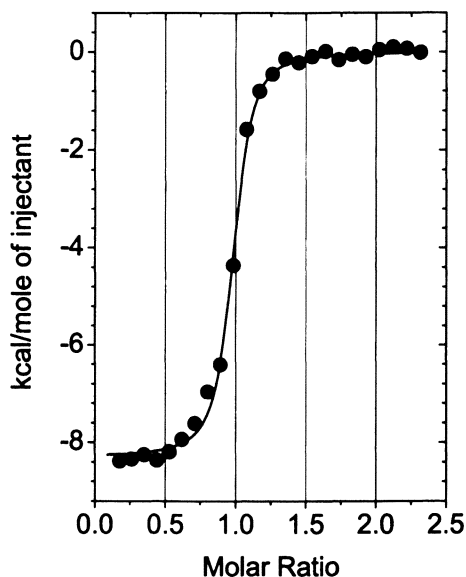
While frataxin seems likely to serve diverse cellular roles, many of these functions appear to involve delivery of iron from a negatively-charged patch on the protein surface (Figure 2), while other roles are promoted through a hydrophobic domain (31, 32). The anionic surface of human frataxin is defined by the twelve acidic residues located in the  $\alpha 1$  helix and  $\beta 1$  sheet (32), with direct binding of partner proteins to the same surface (33). Five acidic residues (E100, E108, E111, D112, D124) are completely conserved in other animals and in plants, yeast, and eubacteria. All frataxin proteins show very similar structures, with the anionic patch located at the boundary of the same  $\alpha 1$  and  $\beta 1$  secondary structural domains, suggestive of an essential role in frataxin function (34).

The most thoroughly characterized form of human frataxin is truncated (residues 78 – 210) beyond the point defined by mitochondrial protease sites, and is found to bind up to 7 iron ions (10) with a  $K_D$  of  $\sim 55 \mu\text{M}$  (for non-His-tagged frataxin). While relatively weak by physiological standards, the affinity for iron increases significantly when frataxin forms complexes with natural partner proteins, and  $K_D$ 's on the order of nM affinity are found (24, 25). Accordingly, the iron binding sites revealed for frataxin alone may best be viewed as a pool of sites that are available for iron binding in complexes with partner proteins.

Frataxin can bind to the cellular ISU target protein in the presence of iron and supports Fe-S cluster reconstitution; however, frataxin does not show significant binding to ISU in the absence of iron (10). This has been interpreted as reflecting a need for iron to bridge the two proteins (10), and is a viewpoint that is further supported by cellular studies on the yeast homologue (22).

Consistent with this model, both X-ray and NMR solution structures of ISU-type proteins from several organisms show five conserved acidic residues near the Fe-S cluster assembly site (Figure 1) (35) that may be responsible for binding to a partner iron delivery protein, with iron serving as a bridge to promote binding between frataxin and ISU through these two anionic surfaces. NMR studies of iron binding to bacterial, yeast, and truncated human frataxins, show that those residues demonstrating the most significant resonance shifts lie within the anionic patch (33, 36).

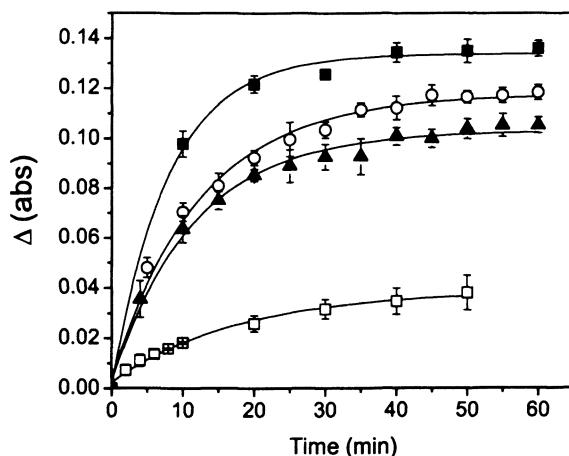
While prior work from this laboratory (10), and others (33, 36-38), has suggested a role for iron in promoting binding of the partner protein ISU to frataxin, most likely through iron-bridged carboxylates, the minimum Fe:frataxin stoichiometry had not been evaluated. To determine the number of iron ions required to promote binding between frataxin and human ISU, and to define their locations on the surface of frataxin, ITC binding experiments were performed with varying equivalents of iron in the presence of ISU and native or derivative frataxin. Figure 2 summarizes the derivatives of frataxin that were prepared, which consisted of multiple substitutions of surface carboxylates. Both native protein and derivatives were found to require only one iron center to promote binding to ISU with nM affinity, and the iron-dependent binding affinity of each frataxin derivative is relatively similar to that of native frataxin (39), as defined by isothermal titration calorimetry experiments (Figure 3).



*Figure 3. ITC measurements of holo frataxin binding to D37A human ISU were carried out at 27 °C on a MicroCal ultrasensitive titration calorimeter. All samples were in degassed 50 mM HEPES (pH 7.5) buffer. A 7  $\mu$ L aliquot of a 0.5 mM solution of titrant was delivered over 10 s to 0.03 mM holo frataxin with an adequate interval (10 min) between injections to allow complete equilibration.*

### Frataxin-Promoted Delivery of Iron to ISU

Prior reports have demonstrated the transfer of iron from holo frataxin to nucleation sites on ISU as a prerequisite step for [2Fe-2S] cluster formation on ISU (10). The time course of the cluster assembly reaction is conveniently monitored from the 456 nm absorbance of holo ISU formed during the [2Fe-2S] cluster assembly reaction (Figure 4). A kinetic rate constant  $k_{\text{obs}} \sim 0.126 \text{ min}^{-1}$  was determined with 100  $\mu\text{M}$  ISU, 2.4 mM  $\text{Na}_2\text{S}$ , and 40  $\mu\text{M}$  holo frataxin in 50 mM Hepes buffer (pH 7.5) with 5 mM DTT. Similar rates were obtained for IscS/Cys-mediated sulfur delivery, consistent with iron release from frataxin as a rate-limiting step in the cluster assembly reaction.



*Figure 4. Iron sulfur reconstitution assay with native truncated frataxin (filled square) and the three derivatives (A, filled triangle; B, open circle; C, open square). Reaction solutions contained 40  $\mu\text{M}$  frataxin, or its derivatives, 280  $\mu\text{M}$  ferrous ion, 5 mM DTT, 2.4 mM  $\text{Na}_2\text{S}$  and 100  $\mu\text{M}$  D37A ISU in 50 mM Hepes buffer (pH 7.5). Each solution component required for the reconstitution assay was prepared under anaerobic conditions and cluster formation was monitored by UV-vis absorption spectroscopy. Experiments were performed at least in triplicate and data points reflect average values with corresponding errors.*

To better define the location and roles of potential iron binding sites, three multiply-substituted frataxin derivatives (frataxin A, frataxin B, and frataxin C;

Figure 2) have been investigated to explore the functional roles of these acidic residues, including their role in binding iron ions and partner proteins, and/or iron delivery. Cluster reconstitution reactions mediated by these frataxin derivatives show a modest decrease for derivative C only (Table I) (39).

**Table I. [2Fe-2S] Reconstitution Rates for Native and Derivative Frataxins**

<i>frataxin</i>	$k_{obs}$ ( $min^{-1}$ )
Native	$0.12 \pm 0.01$
Derivative A	$0.10 \pm 0.01$
Derivative B	$0.10 \pm 0.01$
Derivative C	$0.034 \pm 0.004$

Relative to the position of acidic patches, the overall structural features of both frataxin and ISU appear to serve a minor role in dictating the relative binding affinity and positioning of each protein. Consequently, the reaction sequence is likely to involve initial binding of iron to a cluster of acidic residues that drives frataxin binding to ISU, with subsequent transfer to the conserved cysteine residues in the iron-sulfur cluster binding pocket of ISU. The latter transfer step may be facilitated by additional carboxylate residues. If the acidic residues are not positioned to facilitate iron binding and positioning in the neighborhood of the ISU cysteines, then the reconstitution rate will be compromised.

While there is a reasonable consensus that frataxin does play a direct role in the assembly of cellular Fe-S centers, this idea has been questioned on the basis that cellular deletion of the frataxin gene is non-lethal, although the resulting cells are very sickly and Fe-S cluster formation is significantly impaired (40, 41). While the universality of frataxin mediated delivery is uncertain, it is certain that there are frataxin-like structural homologues that lack sequence homology that would not readily be identified in typical sequence analyses (42). Such proteins appear to function as frataxin-like proteins.

### Sulfide-Donor

Members of the NifS/IscS family are pyridoxyl-5'-phosphate dependent proteins that deliver inorganic sulfide during assembly of the [2Fe-2S] cluster on the IscU scaffold protein. While [2Fe-2S] cluster reconstitution of apo Fe-S proteins can often be achieved *in vitro* by addition of excess iron and sulfide in

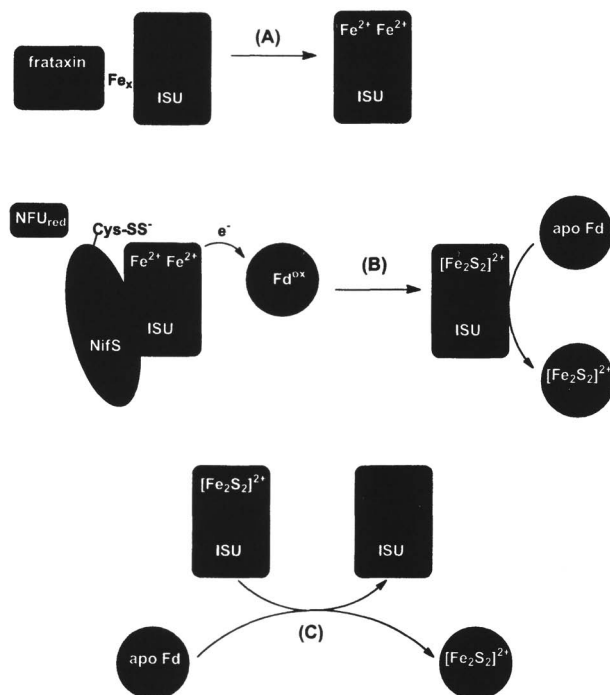
the presence of a reductant, the toxicity of free iron and sulfide requires specific carrier proteins for cellular assembly. Frataxin has been identified as a viable, though perhaps non-exclusive iron donor protein, and inorganic sulfide appears to be provided by IscS. IscS (or NifS-like proteins) belong to the large and diverse  $\alpha$  family of vitamin B<sub>6</sub> enzymes and are evolutionarily conserved (43). They contain the pyridoxyl-5'-phosphate (PLP) cofactor covalently linked through an imine bond to the  $\epsilon$ -amino group of the critical lysine residue within the active site to form the internal aldimine. Sulfur equivalents are provided to IscU/Isu via a persulfide intermediate through the decomposition of L-cysteine (44, 45). IscS and its homologs thereof have been shown to exhibit wide substrate specificity, as evident by the ability to reconstitute apo Fe-S proteins such as ferredoxin and Rieske protein of *Synechocystis* PCC 680 and reconstitution of nitrogenase of *Azotobacter vinelandii* (44, 46-48). IscS proteins display remarkable structural similarity. The X-ray crystal structure of the *Thermotoga maritima* NifS-like homolog (*Tm* NifS) reveals a two-domain structure (small and large) consisting of the putative active cysteine (C325) at the surface near the interface of the IscS dimer and the PLP cofactor buried within the large domain near the active pocket (12). The X-ray crystal structure of *E. coli* IscS reported to 2.1 Å is nearly identical to *Tm* NifS with an ordered loop containing the active surface cysteine (C328) (49), which was not resolved for *Tm* NifS.

The chemistry underlying reductive cleavage of the cysteinyl persulfide bonds of NifS-like proteins is poorly understood. The C-terminal domain of *Av* NifU was earlier described as thioredoxin-like, and a gene encoding the C-terminal homolog has been located in human chromosome 22q (and named NFU) (17). Human NFU forms a complex with NifS-like proteins and is a functionally competent reducing agent for cysteinyl persulfide bond cleavage, releasing inorganic sulfide for incorporation into the ISU-bound [2Fe-2S] cluster (11).

This domain has highly conserved homologs in both prokaryotes and eukaryotes (50, 51). Recent reports have suggested NFU to serve as an alternative scaffold protein on the Fe-S cluster assembly pathway (52); however, the evidence in support of such a role is relatively weak. The presence of a highly conserved thioredoxin-like Cys-X-X-Cys motif (53, 54), the absence of any other viable candidate for cysteinyl persulfide reduction, and the prevalence of a homologous, though functionally ill-defined motif at the C-terminus of NifU, supports the hypothesis of a role for NFU as a reductase that mediates cysteinyl persulfide bond cleavage (55). A similar role for a CXXC motif has previously been elucidated in many other protein disulfide isomerases (PDI) (56).

## Working Model for [2Fe-2S] Cluster Assembly and Transfer to Target Proteins

Figure 5 illustrates a working model for mitochondrial Fe-S cluster assembly that emerges from the previously summarized results from a number of laboratories working on eukaryotes and related pathways in prokaryotes (10, 57-59). While the ability of one bacterial IscU protein to support a [4Fe-4S] center has been demonstrated (58), the generality of this observation, and the



**Figure 5.** A working model for [2Fe-2S] formation on the ISU scaffold that is consistent with published data. (A) Frataxin mediates initial delivery of iron to apo ISU through an iron-bridged complex, establishing a nucleation site for [2Fe-2S] cluster formation. (B) Subsequent sulfide delivery is mediated through the persulfide form of NifS/IscS that forms a complex with ISU. Reductive cleavage of NifS/IscS persulfide bonds by reduced NFU yields the two equivalents of inorganic  $\text{S}^{2-}$  for cluster formation. The reduced [2Fe-2S] cluster initially formed on IscU may then be oxidized by  $\text{Fd}^{\text{ox}}$  to the more stable  $[\text{2Fe-2S}]^{2+}$  form (18), with subsequent transfer (C) of the cluster to an apo protein.

physiological relevance of such a center, remain to be established. A role in mitochondrial aconitase reconstitution has been suggested (60).

## References

1. Soriano, A.; Colpas, G. J.; Hausinger, R. P. *Biochemistry* **2000**, *39*, 12435-12440.
2. Rangaraj, P.; Shah, V. K.; Ludden, P. W. *Proc. Natl Acad. Sci. USA* **1997**, *94*, 11250-11255.
3. Dailey, H. A.; Dailey, T. A.; Wu, C. K.; Medlock, A. E.; Wang, K. F.; Rose, J. P.; Wang, B. C. *Cell. Mol. Life Sci.* **2000**, *57*, 1909-1926.
4. Zheng, L.; Cash, V. L.; Flint, D. H.; Dean, D. R. *J. Biol. Chem.* **1998**, *273*, 13264-13272.
5. Tong, W. H.; Rouault, T. *EMBO J.* **2000**, *19*, 5692-5700.
6. Roy, A.; Solodovnikova, N.; Nicholson, T.; Antholine, W.; Walden, W. E. *EMBO J.* **2003**, *22*, 4826-4835.
7. Agar, J. N.; Zheng, L.; Cash, V. L.; Dean, D. R.; Johnson, M. K. *J. Am. Chem. Soc.* **2000**, *122*, 2136-2137.
8. Mansy, S. S.; Wu, G.; Surerus, K. K.; Cowan, J. A. *J. Biol. Chem.* **2002**, *277*, 21397-21404.
9. Wu, G.; Mansy, S. S.; Wu, S.-P.; Surerus, K. K.; Foster, M. W.; Cowan, J. A. *Biochemistry* **2002**, *41*, 5024-5032.
10. Yoon, T.; Cowan, J. A. *J. Am. Chem. Soc.* **2003**, *125*, 6078-6084.
11. Liu, Y.; Cowan, J. A. *Chem. Commun.* **2007**, 3192-3194.
12. Kaiser, J. T.; Clausen, T.; Bourenkow, G. P.; Bartunik, H. D.; Steinbacher, S.; Huber, R. *J. Mol. Biol.* **2000**, *297*, 451-464.
13. Smith, A. D.; Agar, J. N.; Johnson, K. A.; Frazzon, J.; Amster, I. J.; Dean, D. R.; Johnson, M. K. *J. Am. Chem. Soc.* **2001**, *123*, 11103-11104.
14. Nuth, M.; Yoon, T.; Cowan, J. A. *J. Am. Chem. Soc.* **2002**, *124*, 8774-8775.
15. Wu, S.; Wu, G.; Surerus, K. K.; Cowan, J. A. *Biochemistry* **2002**, *41*, 8876-8885.
16. Hwang, D. M.; Dempsey, A.; Tan, K. T.; Liew, C. C. *J. Mol. Evol.* **1996**, *43*, 536-540.
17. Schilke, B.; Voisine, C.; Beinert, H.; Craig, E. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 10206-10211.
18. Tokumoto, U.; Kitamura, S.; Fukuyama, K.; Takahashi, Y. *J. Biochem.* **2004**, *136*, 199-209.
19. Mansy, S. S.; Wu, G.; Surerus, K. K.; Cowan, J. A. *J. Biol. Chem.* **2002**, *277*, 21397-21404.
20. Bertini, I.; Cowan, J. A.; Del Bianco, C.; Luchinat, C.; Mansy, S. S. *J. Mol. Bio.* **2003**, *331*, 907-924.



21. Chen, O. S.; Hemenway, S.; Kaplan, J. *Proc Natl Acad Sci USA* **2002**, *99*, 12321-12326.
22. Gerber, J.; Muehlenhoff, U.; Lill, R. *EMBO Rep.* **2003**, *4*, 906-911.
23. Puccio, H.; Koenig, M. *Curr. Opin. Genet. Dev.* **2002**, *12*, 272-277.
24. Lesuisse, E.; Santos, R.; Matzanke, B. F.; Knight, S. A.; Camadro, J. M.; Dancis, A. *Hum. Mol. Genet.* **2003**, *12*, 879-889.
25. Yoon, T.; Cowan, J. A. *J. Biol. Chem.* **2004**, *279*, 25943-25946.
26. Kispal, G.; Csere, P.; Prohl, C.; Lill, R. *EMBO J.* **1999**, *18*, 3981-3989.
27. Puccio, H.; Koenig, M. *Hum. Mol. Genet.* **2000**, *9*, 887-892.
28. Ding, H.; Clark, R. J.; Ding, B. *J. Biol. Chem.* **2004**, *279*, 37499 - 37504.
29. Ollagnier-De-Choudens, S.; Mattioli, T.; Takahashi, Y.; Fontecave, M. *J. Biol. Chem.* **2001**, *276*, 22604-22607.
30. Wu, S.; Cowan, J. A. *Biochemistry* **2003**, *42*, 5784-5791.
31. Musco, G.; Stier, G.; Kolmerer, B.; Adinolfi, S.; Martin, S.; Frenkiel, T.; Gibson, T.; Pastore, A. *Structure Fold. Des.* **2000**, *8*, 695-707.
32. Dhe-Paganon, S.; Shigeta, R.; Chi, Y. I.; Ristow, M.; Shoelson, S. E. *J. Biol. Chem.* **2000**, *275*, 30753-30756.
33. He, Y.; Alam, S. L.; Proteasa, S. V.; Zhang, Y.; Lesuisse, E.; Dancis, A.; Stemmler, T. L. *Biochemistry* **2004**, *43*, 16254-16262.
34. Bencze, K. Z.; Kondapalli, K. C.; Cook, J. D.; McMahon, S.; Millan-Pacheco, C.; Pastor, N.; Stemmler, T. L. *Crit. Rev. Biochem. Mol. Biol.* **2006**, *41*, 269-291.
35. Liu, J.; Oganessian, N.; Shin, D. H.; Jancarik, J.; Yokota, H.; Kim, R.; Kim, S. H. *Proteins* **2005**, *59*, 875-881.
36. Cook, J. D.; Bencze, K. Z.; Jankovic, A. D.; Crater, A. K.; Busch, C. N.; Bradley, P. B.; Stemmler, A. J.; Spaller, M. R.; Stemmler, T. L. *Biochemistry* **2006**, *45*, 7767-7777.
37. Acquaviva, F.; De Biase, I.; Nezi, L.; Ruggiero, G.; Tatangelo, F.; Pisano, C.; Monticelli, A.; Garbi, C.; Acquaviva, A. M.; Cocozza, S. *J. Cell. Sci.* **2005**, *118*, 3917-3924.
38. Foury, F.; Pastore, A.; Trincal, M. *EMBO Rep.* **2007**, *8*, 194-199.
39. Huang, J.; Dizin, E.; Cowan, J. A. *J. Biol. Inorg. Chem.* **2008**, *13*, 825-836.
40. Duby, G.; Foury, F.; Ramazzotti, A.; Herrmann, J.; Lutz, T. *Hum. Mol. Genet.* **2002**, *11*, 2635-2643.
41. Li, D. S.; Ohshima, K.; Jiralerspong, S.; Bojanowski, M. W.; Pandolfo, M. *FEBS Lett.* **1999**, *456*, 13-16.
42. Sazanov, L. A.; Hinchcliffe, P. *Science* **2006**, *311*, 1430-1436.
43. Christen, P.; Mehta, P. K. *The Chemical Record* **2001**, *1*, 436-447.
44. Zheng, L.; White, R. H.; Cash, V. L.; Jack, R. F.; Dean, D. R. *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 2754-2758.
45. Zheng, L.; White, R. H.; Cash, V. L.; Dean, D. R. *Biochemistry* **1994**, *33*, 4714-4720.

46. Jaschkowitz, K.; Seidler, A. *Biochemistry* **2000**, *39*, 3416-3423.
47. Schneider, D.; Jaschkowitz, K.; Seidler, A.; Rogner, M. *Ind. J. Biochem. Biophys.* **2000**, *37*, 441-446.
48. Gubernator, B.; Seidler, A.; Rogner, M.; Szczepaniak, A. *Prot. Exp. Purif.* **2003**, *29*, 8-14.
49. Cupp-Vickery, J. R.; Urbina, H.; Vickery, L. E. *J. Mol. Biol.* **2003**, *330*, 1049-1059.
50. Lorain, S.; Lecluse, Y.; Scamps, C.; Mattei, M.-G.; Lipinski, M. *Biochim.. Biophys. Acta* **2001**, *1517*, 376-383.
51. Magnaghi, P.; Roberts, C.; Lorain, S.; Lipinski, M.; Scambler, P. J. *Nat. Genet.* **1998**, *20*, 74-77.
52. Tong, W.; Jameson, G. N. L.; Huynh, B.; Rouault, T. A. *Proc. Natl. Acad. Sci. USA.* **2003**, *100*, 9762-9767.
53. Holmgren, A. *J. Biol. Chem.* **1989**, *264*, 13963-13966.
54. Yano, H.; Kuroda, S.; Buchanan, R. B. *Proteomics* **2002**, *2*, 1090-1096.
55. Frazzon, J.; Dean, D. R. *Curr. Opin. Chem. Biol.* **2003**, *7*, 166-173.
56. Ellgaard, L.; Ruddock, L. W. *EMBO J.* **2005**, *6*, 28-32.
57. Mansy, S. S.; Cowan, J. A. *Acc. Chem. Res.* **2004**, *37*, 719-725.
58. Agar, J. N.; Krebs, C.; Frazzon, J.; Huynh, B. H.; Dean, D. R.; Johnson, M. K. *Biochemistry* **2000**, *39*, 7856-7862.
59. Hoff, K. G.; Ta, D. T.; Tapley, T. L.; Silberg, J. J.; Vickery, L. E. *J. Biol. Chem.* **2002**, *277*, 27353-27359.
60. Chandramouli, K.; Unciuleac, M.-C.; Naik, S.; Dean, D. R.; Huynh, B. H.; Johnson, M. K. *Biochemistry* **2007**, *46*, 6804 -6811.

## Chapter 2

# Molecular Interaction between Frataxin and Ferrochelatase During Heme Assembly

### Frataxin's Role as a Potential Iron Chaperone During Heme Biosynthesis

Kalyan C. Kondapalli<sup>1</sup>, Andrew Dancis<sup>2</sup>,  
and Timothy L. Stemmler<sup>1,\*</sup>

<sup>1</sup>Department of Biochemistry and Molecular Biology,  
Wayne State University School of Medicine, Detroit, MI 48201

<sup>2</sup>Department of Medicine, Division of Hematology–Oncology,  
University of Pennsylvania, Philadelphia, PA 19104

Frataxin, a mitochondrial protein known to be important in cellular iron homeostasis, has been suggested to serve as an iron chaperone during heme synthesis. Here we report an extensive biophysical characterization of the intermolecular interaction between yeast frataxin (Yfh1) and yeast ferrochelatase (Hem15). Yfh1 iron-binding residues on the Hem15 binding surface were mutated and characterized regarding complex formation. Our results suggest redundancy in Yfh1 iron-binding sites for *in vivo* iron delivery during heme biosynthesis.

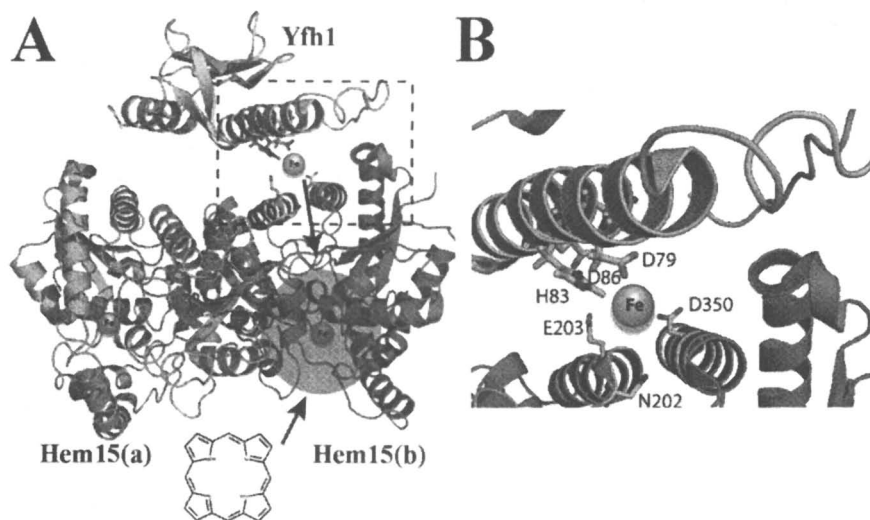
Iron is essential for most life forms and often utilized as a cofactor in critical enzymes involving oxidation and reduction. The reactivity of iron towards oxidation chemistry, if left unregulated, can be damaging to proteins, lipids and DNA in cells. One mechanism organisms utilize to regulate iron reactivity is by constraining metal at the center of heme cofactors. Heme containing proteins are

ubiquitous throughout the body and perform a variety of biological functions including oxygen transport (hemoglobins), catalysis (peroxidases) and electron transport (cytochromes) (1). In humans, about 70% of the total iron in the body is found in hemoglobin, so it is no surprise that numerous disorders are linked to an inability to produce adequate amounts of heme (*i.e.*, the anemias or porphyrias) (2). In eukaryotes, heme biosynthesis begins within the mitochondria, starting with the condensation of glycine and succinyl CoA by the enzyme  $\delta$ -aminolevulinate synthase. Heme synthesis also ends within the mitochondria with the insertion of Fe(II) into the porphyrin ring, a reaction catalyzed by the enzyme ferrochelatase. While the identity and role of proteins within the porphyrin biosynthetic pathway are well established, the identity of the protein that delivers Fe(II) to ferrochelatase has been elusive until recently.

Frataxin, a nuclear encoded mitochondrially targeted protein essential for cellular iron homeostasis (3), has been suggested to serve as the mitochondrial iron chaperone that delivers Fe(II) to ferrochelatase (4). Frataxin is part of a growing number of iron chaperone proteins responsible for protein assisted metal delivery (5). Humans with frataxin deficiency have the neurodegenerative disorder Friedreich's ataxia (FRDA), which affects 1 in 50,000 (6). Cellular phenotypes of affected tissues in FRDA include mitochondrial iron hyperaccumulation, elevated oxidative stress and deficiency in both heme (7) and iron-sulfur clusters (8, 9). *In vivo* studies in yeast have shown frataxin is required for correct metallation of porphyrin; in the absence of iron, zinc protoporphyrin is formed instead of heme (4). *In vitro* studies show frataxin binds to ferrochelatase with a nanomolar affinity (4) and that holo frataxin stimulates ferrochelatase catalyzed heme production, presumably by delivering the Fe(II) required to make heme (4, 10, 11). Given the substantial attraction between frataxin and ferrochelatase, the fact that frataxin stimulates Fe-cofactor production during *in vitro* activity assays, and that frataxin is required for heme production *in vivo*, a role for frataxin as an iron chaperone during heme biosynthesis seems highly likely.

Some insight into the molecular interaction between frataxin and ferrochelatase has come from recent structural studies (12). The conserved structure of the frataxin orthologs indicates they are members of the  $\alpha$ - $\beta$  sandwich motif family. All frataxin orthologs are highly acidic (pI's range between 4 and 6) and all have highly conserved acidic residues located on their surface-exposed helical plane. These residues have been shown to be important in iron-binding. NMR intermolecular surface mapping studies indicate ferrochelatase binds to the helical surface of frataxin, utilizing in part the iron-binding residues on frataxin's N-terminal  $\alpha$ -helix (12, 13). In the case of ferrochelatase, biochemical studies indicate that during heme production surface-

exposed helical polar residues initially accept the Fe(II) being delivered (14). Iron is then translocated down the ferrochelatase substrate channel to the enzyme's active site. Based on published data identifying the ferrochelatase initial Fe-binding site, frataxin's ferrochelatase binding surface identified by NMR, the frataxin to ferrochelatase protein/protein binding affinity and stoichiometry, and the structures of the frataxin and ferrochelatase orthologs, we generated a model for frataxin binding and iron delivery to ferrochelatase in yeast (Figure 1) (12). This model is also applicable to human proteins, although the human ferrochelatase, unlike the yeast protein, possesses a structurally important [2Fe-2S] cluster (15, 16).



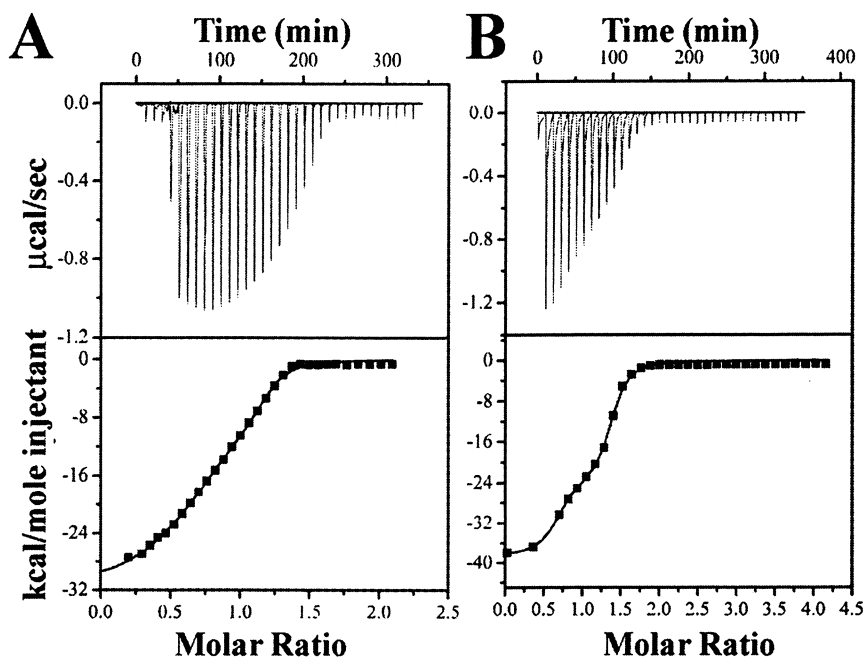
**Figure 1. Model for Fe(II) delivery during yeast heme biosynthesis (12).** (A) Monomeric Yfh1 (top) binds to the matrix exposed region of Hem15 dimer (monomer a and b, bottom) utilizing illustrated Yfh1 iron-binding (D79, D86) and additional helical plane residues (D82, E93, H95,). Iron transfer is accommodated using Hem15 residues (N202, E203, D350) at the ferrochelatase initial iron acceptance site. Following iron transfer, metal is free to translocate down the Hem 15 substrate channel (top arrow) to the enzyme's active site (shaded area). Protoporphyrin IX is delivered through the membrane (bottom arrow). (B) Expansion of iron transfer region between Yfh1 and Hem15 (boxed region in Figure 1A).

While our model is consistent with published data regarding how frataxin and ferrochelatase interact, additional studies are required to test our model and provide precise atomic details of the inter-protein interaction. The goal of this report is therefore to expand the general understanding of how frataxin and ferrochelatase interact, focusing on proteins only in the yeast system. Binding studies were utilized to address the role of iron in helping form the Yfh1/Hem15 intermolecular complex. Structural studies were used to provide the metrical details for the metal being transferred between Yfh1 and Hem15. Mutagenesis studies were utilized to help address the specific Yfh1 iron-binding site during transfer and delivery. When combined, these studies provided an expanded understanding of the iron delivery mechanism utilized by frataxin during heme biosynthesis.

## **Binding Affinity of Yeast Frataxin and Ferrochelatase**

Iron-binding and selective metal delivery to ferrochelatase are anticipated functions for frataxin in our heme biosynthesis model. While the iron-binding ability of the frataxin orthologs has been documented (10, 17-19), it remains unclear if binding to ferrochelatase is metal dependent. Human frataxin binds to ferrochelatase with a binding affinity of 17 nM and a monomeric binding stoichiometry of 1:2 (*i.e.*, one frataxin monomer to one ferrochelatase dimer) (10). Binding between the human proteins is completely dependent on the presence of iron. In contrast, binding between the yeast partners occurs in the absence of metal (4). Apo Yfh1 binds to Hem15 with a binding affinity of 40 nM, again at a 1:2 respective monomeric protein ratio, as measured by surface plasmon resonance spectroscopy. The interaction between yeast proteins has also been identified by NMR to occur in the absence of metal (13). In this report, we characterized the binding affinity of both apo and holo Yfh1 to apo Hem15 using isothermal titration calorimetry (ITC). Previous ITC studies showed monomeric apo Yfh1 binds two Fe(II) atoms with an average  $K_d$  of 2.5  $\mu$ M (18). In our current complex ITC experiment, anaerobic solutions of apo Hem15 were injected into either apo or holo Yfh1. Experiments were performed at 30 °C on a VP-ITC titration microcalorimeter (Micro Cal Inc.). Protein solution conditions utilized during the study were the following: 20 mM HEPES (pH 7.4), 150 mM NaCl and 0.05% deoxycholate. Experiments were repeated in triplicate on independent samples to ensure data reproducibility. The Origin 5.0 software package (Micro Cal Inc.) was used for data analysis.

Yfh1 binds to Hem15 in a manner that is modulated by, but not dependent on, the presence of iron. Binding between proteins is observed using both apo and holo Yfh1 (Figure 2). In both cases, Yfh1 binds to Hem15 with an approximate 1:2 respective monomeric protein ratio. Binding of apo Yfh1 to



*Figure 2. Raw isothermal titration calorimetry (top) and binding isotherm data (bottom) for the binding of Hem15 to apo (A) and holo (B) Yfh1. Data were collected by injecting small aliquots of a 550  $\mu$ M Hem15 into a 50  $\mu$ M Yfh1 solution. Spacing between injections was 10 at a 500 rpm stirring speed.*

Hem15 gave an average  $K_d$  from a two binding-site simulation model of  $480 \pm 62$  nM; thermodynamic parameters for the fit gave an overall  $\Delta G$  of  $-4.7 \pm 1.2$  kcal/mol, a  $\Delta H$  of  $-18.5$  kcal/mol and a  $T\Delta S$  of  $-13.8$  kcal/mol, indicating binding could be driven predominately by enthalpic contributions. Factors like water release due to solvation of metal ion or protonation state of amino acid side chains that mediate complex formation could contribute to the change in enthalpy, so further experiments using multiple buffers with distinct ionization enthalpies need to be done to address this issue. Binding of holo Yfh1 to Hem15 gave an average  $K_d$ , using again a two-site binding simulation model, of  $163 \pm 26$  nM. Though the data could be fit with a sequential binding model, a two-site simulation provided the best fit and the two  $K_d$  values were significantly distinct with a higher affinity for the first binding. Thermodynamic parameters gave an overall  $\Delta G$  of  $-9.9 \pm 1.2$  kcal/mol, a  $\Delta H$  of  $-30.9$  kcal/mol and a  $T\Delta S$  of  $-21.2$  kcal/mol, indicating again that binding could be driven by enthalpic contributions. Although apo Yfh1 binds to Hem15, in contrast to the situation

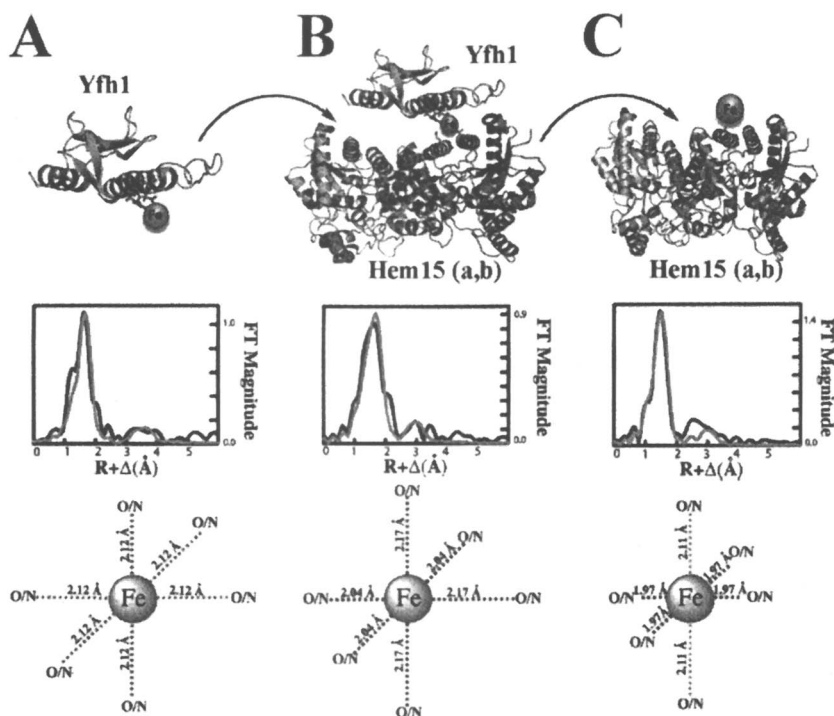
observed with the human orthologs, the presence of iron does improve the binding affinity between the two partners to a  $K_d$  value that is 4 times stronger than is observed in the absence of metal.

## Fe-Coordination Structure During Metal Transfer

Although structural details regarding the frataxin and ferrochelatase complex are lacking, accurate details regarding the coordination environments of the iron being delivered have already been provided by X-ray absorption spectroscopic (XAS) studies. XAS studies have been used to characterize the structure of iron bound to both individual proteins, providing metrical details of the metal at the starting and the end point during metal delivery (18, 20). X-ray absorption near edge spectral (XANES) analysis of iron bound to Yfh1 show bound Fe(II) is high spin (18). Simulations to the extended X-ray absorption fine structure (EXAFS) spectrum for holo Yfh1 indicate metal is coordinated in a moderately symmetric Fe- (O/N)<sub>6</sub> ligand environment (Figure 3 and Table I). The average nearest neighbor bond length for the oxygen and/or nitrogen ligands for holo Yfh1 is 2.12 Å, a value most consistent with six coordinate ferrous iron-oxygen crystallographically characterized small molecules (21). Long-range iron-carbon scattering is also observed. Fe(II) bound to Hem15 is also high spin and six coordinate, again utilizing only O and N based ligands in the metal's nearest neighbor coordination sphere. Previously published fits to the holo Hem15 iron EXAFS indicated two unique Fe-nearest neighbor environments constructed by 4 O/N ligands at an average bond length of 2.00 Å and 2 Fe-N<sub>imidazole</sub> interaction at 2.12 Å (20). In this report, Fe-EXAFS data for holo Hem15 were collected over a higher energy range (and hence higher resolution) than the published fits. Although we were unable to justifiably fit the Hem15 iron EXAFS using a N<sub>imidazole</sub> model in the manner previously published, we were able to again resolve two independent Fe-nearest neighbor ligand environments constructed by 4 O/N ligands at 1.97 Å and 2 Fe-O/N bonds at 2.11 Å (Figure 3 and Table I). The nearest neighbor bond lengths are consistent with the 0.13 Å resolution of our data. Again, long range carbon scattering is observed with this sample. Long range scattering observed in our Fe-Hem15 sample could also not be fit with a multiple scattering imidazole model.

Structural characterization of Fe bound in the Yfh1/Hem15 complex gave an average metal coordination geometry unique from either of the starting or end points outlined above. XANES analyses confirm iron is high spin and stable as Fe(II) and the metal ligand coordination geometry, evaluated based on the 1s→3d electronic absorption energies and areas, is again consistent with a moderately symmetric six coordinate Fe(II) complex (XANES data not shown). EXAFS fitting analysis for iron in the complex is best simulated with two unique Fe-O/N nearest neighbor ligand environments (Figure 3 and Table I). However,





*Figure 3. Summary of Fe(II) EXAFS characterization for Yfh1 (A), the Yfh1/Hem15 complex (B) and for Hem15 (C). Top: Structural model of proteins with iron bound; Middle: Fourier transforms (black) and simulations (grey) of Fe EXAFS; Bottom: graphical representation of approximate iron coordination geometry.*

in this case, the closer nearest neighbor environment is centered at an average iron-ligand bond length of 2.04 Å. This ligand interaction is fairly disordered, based on the low coordination number and high Debye-Waller factor from this best fit. A well resolved long Fe-O/N bond interaction, centered at the average bond length of 2.17 Å, is the structural basis for the second Fe-nearest neighbor ligand environment in our simulation. This long Fe(II)-O/N bond distance of 2.17 Å is most consistent with the longer Fe-N bond seen in small iron molecules (21). These data therefore suggest iron in the complex has a higher number of nitrogen ligands than the starting and end points. Long range carbon scattering is observed in the complex Fe EXAFS, however this could not be fit with an imidazole model.

Although the Fe EXAFS fitting results for the Yfh1/Hem15 complex were unique from either the starting or end points, it is impossible to determine which

**Table I. Summary of Fe EXAFS Fitting Analysis for Yfh1, the Yfh1/Hem15 Complex and Hem15: Best-Fit Parameters**

Sample	Nearest Neighbor Ligand Environment <sup>a</sup>				Long-Range Ligand Environment <sup>a</sup>				F <sup>f</sup>
	Atom <sup>b</sup>	R(Å) <sup>c</sup>	C.N. <sup>d</sup>	s <sup>2</sup> <sup>e</sup>	Atom <sup>b</sup>	R(Å) <sup>c</sup>	C.N. <sup>d</sup>	s <sup>2</sup> <sup>e</sup>	
Yfh1	O/N	2.12	5.0	5.78	C	3.07	2.0	4.30	0.92
					C	3.30	2.0	4.47	
Complex	O/N	2.04	2.0	3.69	C	3.13	1.0	5.05	0.64
	O/N	2.17	3.0	2.02	C	3.50	2.0	1.96	
Hem15p	O/N	1.97	4.0	3.39	C	3.54	2.0	1.97	0.61
	O/N	2.11	2.0	2.37					

<sup>a</sup> Independent metal-ligand scattering environment.

<sup>b</sup> Scattering atoms: O (oxygen), N (nitrogen), C (carbon).

<sup>c</sup> Average metal-ligand bond length for 2 independent samples.

<sup>d</sup> Average metal-ligand coordination number for 2 independent samples.

<sup>e</sup> Average Debye-Waller factor in Å<sup>2</sup> x 10<sup>3</sup> for 2 independent samples.

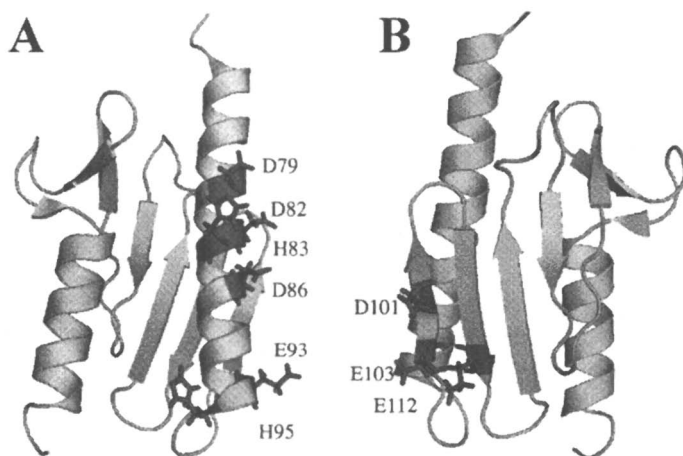
<sup>f</sup> Number of degrees of freedom weighted mean square deviation between data and fit.

step during delivery that we are trapping iron in the complex. Given the longer (2.17 Å) Fe-O/N bond length for iron in the complex sample only, and the correlation of this longer distance to small molecule nitrogen scattering, it is tempting to suggest we may have trapped the metal in a transfer transition state. In this state, nitrogen side chain residues, possibly from Yfh1 (H83) and Hem15 (N202), could be contributing the nitrogen ligands to iron. However, these data could also easily be attributed to having trapped iron during translocation down the Hem15 substrate channel or as being a mixture of many possible transition states. Mutagenesis studies are therefore required to confirm which species we are observing in our complex XAS studies. A method to achieve this would be to block iron translocation down the Hem15 substrate. To accomplish this, we have mutated Hem15 residues Y163 and W199, amino acids located directly in the Fe-substrate transport channel that are essential for Fe transport down the ferrochelatase substrate channel in the human ortholog (14, 22). Once completed, we will characterize the complex Fe-EXAFS with these mutants, and measure the binding affinity between Yfh1 and mutant Hem15, to provide a better understanding of iron delivery between the proteins.

## Frataxin Helix-1 Metal Binding Site

Our model for heme biosynthesis predicts Yfh1 helix-1 residues are utilized during iron delivery to Hem15. While only a small number of conserved acidic residues on frataxin's  $\beta$ -sheet surface have been implicated in iron-binding, there are multiple residues lining the surface exposed region of the protein's  $\alpha$ -helix-1 (Figure 4) that have been shown by NMR to interact with metal. Evidence for this interaction came from two independent NMR titrations of Fe(II) into  $^{15}\text{N}$  labeled Yfh1. This first characterization, performed in phosphate buffer, suggested residues dispersed along the middle and C-terminal sections of Yfh1's helix-1 were affected by the presence of iron (including residues H83, D86, E93 and H95) (13). The interaction between iron and specific residues was shown by chemical shift perturbations, and direct evidence for metal binding was not observed, *i.e.*, there was no spectral line broadening due to binding of the paramagnetic Fe(II). In our second titration, residues on helix-1 were line broadened beyond recognition indicating a direct effect coupled with the presence of the paramagnetic iron (18).

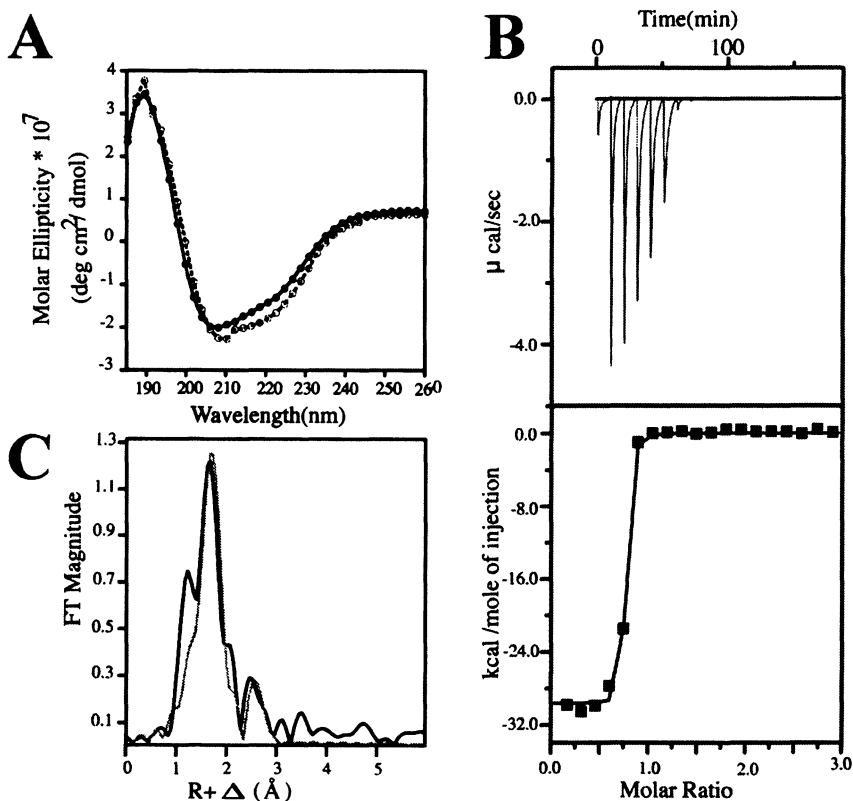
We prepared a Yfh1 mutant covering a subset of the helix-1 residues that were perturbed by the presence of Fe(II) in our initial titration results. Alanine



**Figure 4. Identification of Yfh1 residues affected by the presence of Fe(II) during both NMR titrations of anaerobic iron into  $^{15}\text{N}$  labeled protein at 600 MHz. (A) Helical plane residues H83, D86, E93 and H95, identified in the phosphate buffer titration with additional residues D79 and D82 identified in the subsequent HEPES buffer titration. (B) Summary of  $\beta$ -sheet plane residues also implicated in Fe-binding during the two separate titrations.**

was substituted at Yfh1 positions D86, E93 and H95 in a plasmid for bacterial expression of the mature processed Yfh1 (amino acids 52-174) using the QuickChange II mutant strategy (Stratagene). Verification of the mutations was performed at the nucleotide (DNA sequencing) and protein (mass spectroscopy) levels. The fold of the D86E93H95A Yfh1 mutant was identical to wild type protein, as measured by circular dichroism spectroscopy (Figure 5). Iron-binding was characterized by ITC (Figure 5) in both phosphate and HEPES buffer. Mutant protein bound only a single iron atom with a  $K_d = 35.46$  nM. Binding was again influenced towards the enthalpic energy contribution ( $\Delta G_{\text{binding}} = -10.27$  kcal/mol,  $\Delta H = -29.6$  kcal/mol,  $T\Delta S = -19.3$  kcal/mol). Characterization of the iron structure by XAS confirmed bound Fe(II) is high spin and held in a Fe-ligand coordination dominated by *ca.* six O/N ligands at an average bond length of 2.14 Å (Figure 5). Identical mutations were introduced into a plasmid designed for yeast expression, and the triple mutant form of Yfh1 was expressed from the native promoter on the chromosome of a  $\Delta yfh1$  strain, following replacement of the endogenous gene by plasmid shuffling. Surprisingly, this triple mutant strain showed no discernable heme or even general iron regulatory phenotypes, as judged by cell growth, cytochrome *c* levels, protoporphyrin levels, regulated high affinity cellular Fe uptake, ferric reductase or mitochondrial iron levels (data not shown). Therefore, while it appears that we can successfully block iron from interacting with the C-terminal helix-1 residues on Yfh1 *in vitro* (at positions D86A, E93A and H95A), the role of this abolished interaction *in vivo* had no effect. A possible explanation relates to the presence of additional iron-binding sites, perhaps mediated by other acidic residues, in the helix-1 region. Also, recent data indicates that there is large superabundance of Yfh1 in normal yeast cells, such that depletion of 80% or more of the mitochondrial Yfh1 protein can be achieved without causing much of a phenotype (23, 24). Establishment of Yfh1 structure-function correlations *in vivo* may require further evaluation of iron-binding site mutants at lower expression levels.

Our initial NMR titration confirms that multiple acidic residues on Yfh1's helix-1 are affected by the presence of iron. Experiments are in progress to further characterize the role of Yfh1 helix-1 acidic residues towards Fe-binding. Interestingly, when we repeated our Fe(II) to  $^{15}\text{N}$  Yfh1 NMR titration in HEPES buffer with Mg(II) salt in our second study (18), we again saw helix-1 residues (including E93 and H95) that underwent amide chemical shift perturbations coupled with an interaction with iron. However, we also saw a subset of helix-1 residues (D79, D82, H83 and D86) that were completely line broadened in the titration, presumably because they directly interact with the paramagnetic metal from the titrate. The identification of these additional helix-1 residues related to Yfh1 iron-binding may indicate that multiple binding sites for divalent metals are present on the protein, or that frataxin is able to accommodate the binding of a single metal utilizing a variety of residues in this structural region. Experiments



**Figure 5. Biophysical characterization of apo and holo D86E93H95A Yfh1.**  
 (A) Comparison of wt-(dotted line) and mutant (dashed line) apo Yfh1 CD spectra. (B) Raw (top) and binding isotherm (bottom) ITC data for anaerobic iron-binding to mutant Yfh1 (dotted) with simulations (solid line). (C) Fourier transform (black) and simulation (grey) of mutant Fe EXAFS.

are currently underway to test the role of these additional helix-1 residues towards iron-binding and delivery during Yfh1/Hem15 binding.

## Conclusions

Frataxin involvement in heme synthesis was suggested due to the striking phenotype of frataxin deficient yeast mutants almost completely lacking heme biosynthetic activity while retaining zinc chelatase activity (4). A direct interaction between frataxin and ferrochelatase was originally shown with the

yeast orthologs to be independent of the presence of iron. However, binding between the human orthologs is completely dependent on the presence of iron (10). Differences may reflect the unique properties of the mature yeast and human frataxin orthologs. Human frataxin has an additional 17 amino acid C-terminal tail that is essential for protein stability (25) and which structurally folds against frataxin's helical ferrochelatase binding plane (26). In addition, human frataxin/ferrochelatase studies were performed on an N-terminally truncated form of frataxin that may modulate binding between the protein partners (10). Finally, human ferrochelatase has an Fe-S cluster that is not present in the yeast ortholog, and this divergence might account for altered protein interactions (27). The fact that apo Yfh1 binds to Hem15 may suggest a transient interaction between protein partners that is highly supported when metal is present or that iron delivery is not the only function of frataxin during heme biosynthesis. A possible additional role of the apo Yfh1/Hem15 interaction could be to simply inhibit Zn loading of porphyrin by ferrochelatase.

Metal delivery is still a probable functional role of frataxin during heme biosynthesis. The enhanced Yfh1 to Hem15 binding affinity we see in the presence of metal indicates iron directly helps mediate binding between these protein partners. Given the weak ( $\mu\text{M}$ ) binding affinity of iron to frataxin, compared to the nM value obtained between holo Yfh1 and Hem15, and the fact that holo Yfh1 to Hem15 binding is enthalpically driven, may indicate that the selective transfer of iron to ferrochelatase is accomplished through numerous stabilizing energetic factors (H-bonding, Van der Waals interactions and ionic bonds) formed between the two proteins when iron is present. Protein complex Fe-EXAFS data shows metal coordination is distinct from the starting and end points in delivery, suggesting we may be monitoring the Fe structure during or immediately following metal transfer (*i.e.*, by trapping metal at the Hem15 initial metal acceptance site or during translocation, respectively). Hem15 mutants are currently being prepared to block metal translocation down the ferrochelatase substrate channel to address the identity of the metal structure we have observed in our complex EXAFS data.

Our initial model for Fe delivery by Yfh1 suggested the acidic residue D86 participates in Fe transfer to Hem15 during heme biosynthesis. Although our D86E93H95 mutant only binds a single Fe atom, the fact that this mutant is still active for *in vivo* heme biosynthesis could mean that there is a fair amount of accommodability in helix-1 Fe-binding site on Yfh1 or that the mutant still can bind multiple iron atoms, only at an affinity that exceeds the detectable limit of our technique (ITC). Our initial solution conditions may have obscured additional helix-1 Fe-binding sites. Alternatively, the second Fe-binding site we detected by NMR, located on the Yfh1  $\beta$ -sheet surface, may be utilized during heme delivery. Yfh1's  $\beta$ -sheet surface is crucial for only Fe-S bioassembly (23), and mutation of Yfh1  $\beta$ -sheet Fe-binding residues shows clear Fe-S cluster

assembly phenotypes (28). Heme synthesis in our triple mutant has not been directly addressed, especially under conditions of low protein expression designed to elicit relevant mutant phenotypes. Clearly a structural snapshot of each molecule in the holo Yfh1/Hem15 complex would be beneficial when trying to decipher the specific Yfh1 residues involved in iron delivery during heme biosynthesis. To this end, we have recently cocrystallized the holo Yfh1 monomer with the Hem15 dimer. While only in its early stages, the structures are not yet at a point where the exact molecular orientation and structural details of Yfh1 in the complex have been determined, however it is exciting to speculate that the molecular details of the frataxin/ferrochelatase is forthcoming.

### Acknowledgements

This work was supported by the American Heart Association for K.C.K. (0610139Z) and by the National Institutes of Health for T.L.S. (DK068139) and A.D. (DK53953). A portion of the research was carried out at the Stanford Synchrotron Radiation Laboratory (SSRL). SSRL is a national user facility operated by Stanford University on behalf of the U.S. Department of Energy, Office of Basic Energy Sciences. The SSRL Structural Molecular Biology Program is supported by the Department of Energy, Office of Biological and Environmental Research, and by the NIH, National Center for Research Resources, Biomedical Technology Program.

### References

1. Crichton, R. *Inorganic Biochemistry of Iron Metabolism: From Molecular Mechanisms to Clinical Consequences-2<sup>nd</sup> Ed.*; John Whiler & Sons, Ltd.: West Sussex, England, 2001.
2. Miret, S.; Simpson, R. J.; McKie, A. T. *Annu. Rev. Nutr.* **2003**, *23*, 283-301.
3. Babcock, M.; de Silva, D.; Oaks, R.; Davis-Kaplan, S.; Jiralerspong, S.; Montermini, L.; Pandolfo, M.; Kaplan, J. *Science* **1997**, *276*, 1709-1712.
4. Lesuisse, E.; Santos, R.; Matzanke, B. F.; Knight, S. A.; Camadro, J. M.; Dancis, A. *Hum. Mol. Genet.* **2003**, *12*, 879-889.
5. Shi, H.; Bencze, K. Z.; Stemmler, T. L.; Philpott, C. C. *Science* **2008**, *320*, 1207-1210.
6. Delatycki, M. B.; Williamson, R.; Forrest, S. M. *J. Med. Genet.* **2000**, *37*, 1-8.
7. Schoenfeld, R. A.; Napoli, E.; Wong, A.; Zhan, S.; Reutenauer, L.; Morin, D.; Buckpitt, A. R.; Taroni, F.; Lonnerdal, B.; Ristow, M.; Puccio, H.; Cortopassi, G. A. *Hum. Mol. Genet.* **2005**, *14*, 3787-3799.

8. Rotig, A.; de Lonlay, P.; Chretien, D.; Foury, F.; Koenig, M.; Sidi, D.; Munnich, A.; Rustin, P. *Nat. Genet.* **1997**, *17*, 215-217.
9. Stehling, O.; Elsasser, H. P.; Bruckel, B.; Muhlenhoff, U.; Lill, R. *Hum. Mol. Genet.* **2004**, *13*, 3007-3015.
10. Yoon, T.; Cowan, J. A. *J. Biol. Chem.* **2004**, *279*, 25943-25946.
11. Park, S.; Gakh, O.; O'Neill, H. A.; Mangravita, A.; Nichol, H.; Ferreira, G. C.; Isaya, G. *J. Biol. Chem.* **2003**, *278*, 31340-31351.
12. Bencze, K. Z.; Kondapalli, K. C.; Cook, J. D.; McMahon, S.; Millan-Pacheco, C.; Pastor, N.; Stemmler, T. L. *Crit. Rev. Biochem. Mol. Biol.* **2006**, *41*, 269-291.
13. He, Y.; Alam, S. L.; Proteasa, S. V.; Zhang, Y.; Lesuisse, E.; Dancis, A.; Stemmler, T. L. *Biochemistry* **2004**, *43*, 16254-16262.
14. Wu, C. K.; Dailey, H. A.; Rose, J. P.; Burden, A.; Sellers, V. M.; Wang, B. C. *Nat. Struct. Biol.* **2001**, *8*, 156-160.
15. Bencze, K. Z.; Yoon, T.; Millan-Pacheco, C.; Bradley, P. B.; Pastor, N.; Cowan, J. A.; Stemmler, T. L. *Chem. Commun.* **2007**, 1798-1800.
16. Dailey, H. A.; Dailey, T. A.; Wu, C. K.; Medlock, A. E.; Wang, K. F.; Rose, J. P.; Wang, B. C. *Cell. Mol. Life Sci.* **2000**, *57*, 1909-1926.
17. Bou-Abdallah, F.; Adinolfi, S.; Pastore, A.; Laue, T. M.; Chasteen, N. D. *J. Mol. Biol.* **2004**, *341*, 605-615.
18. Cook, J. D.; Bencze, K. Z.; Jankovic, A. D.; Crater, A. K.; Busch, C. N.; Bradley, P. B.; Stemmler, A. J.; Spaller, M. R.; Stemmler, T. L. *Biochemistry* **2006**, *45*, 7767-7777.
19. Yoon, T.; Cowan, J. A. *J. Am. Chem. Soc.* **2003**, *125*, 6078-6084.
20. Ferreira, G. C.; Franco, R.; Mangravita, A.; George, G. N. *Biochemistry* **2002**, *41*, 4809-4818.
21. Centre, C. C. D. *ConQuest 1.5*, Cambridge, 2002.
22. Karlberg, T.; Lecerof, D.; Gora, M.; Silvegren, G.; Labbe-Bois, R.; Hansson, M.; Al-Karadaghi, S. *Biochemistry* **2002**, *41*, 13499-13506.
23. Wang, T.; Craig, E. A. *J. Biol. Chem.* **2008**, *283*, 12674-12679.
24. Karthikeyan, G.; Santos, J. H.; Graziewicz, M. A.; Copeland, W. C.; Isaya, G.; Van Houten, B.; Resnick, M. A. *Hum. Mol. Genet.* **2003**, *12*, 3331-3342.
25. Adinolfi, S.; Nair, M.; Politou, A.; Bayer, E.; Martin, S.; Temussi, P.; Pastore, A. *Biochemistry* **2004**, *43*, 6511-6518.
26. Musco, G.; Stier, G.; Kolmerer, B.; Adinolfi, S.; Martin, S.; Frenkiel, T.; Gibson, T.; Pastore, A. *Structure* **2000**, *8*, 695-707.
27. Dailey, H. A.; Finnegan, M. G.; Johnson, M. K. *Biochemistry* **1994**, *33*, 403-407.
28. Kondapalli, K. C.; Cook, J. D.; Dancis, A.; Stemmler, T.L.; *manuscript in preparation.*



## Chapter 3

# Biosynthesis and Regulation of the Heme A Biosynthetic Pathway

Eric L. Hegg

Department of Biochemistry and Molecular Biology,  
Michigan State University, East Lansing, MI 48824

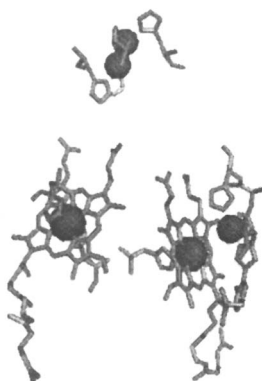
Heme A is an obligatory cofactor in all eukaryotic and most bacterial cytochrome *c* oxidase enzymes (CcO). Because of its importance to CcO and aerobic metabolism, considerable effort has recently been invested in understanding the mechanism and regulation of heme A biosynthesis. The activity of heme A synthase is strictly dependent on O<sub>2</sub>, and yet there is no incorporation of O<sub>2</sub> into the products. Heme A synthase is now known to utilize a unique electron-transfer mechanism when oxidizing heme O to heme A. Interestingly, the heme A biosynthetic pathway is regulated at least partly via a heme-dependent process in which heme A synthase is positively regulated by intracellular heme levels via Hap1.

Cytochrome *c* oxidase (CcO), an integral membrane, multi-component complex, is the last enzyme in the energy-transducing, electron-transfer chain in all plants, animals, aerobic yeasts, and some bacteria (1-3). Arguably one of the most important enzymes in aerobic metabolism, cytochrome *c* oxidase catalyzes the transfer of electrons from cytochrome *c* to molecular oxygen, thus reducing O<sub>2</sub> to water. Utilizing the energy released from this exergonic reaction, cytochrome *c* oxidase pumps up to four protons (eight charge equivalents)

across the inner membrane of either the mitochondria or bacteria, generating a considerable membrane potential and an appreciable chemical gradient in the process (4-7). The combination of these two gradients produces a proton-motive force of over 200 mV that is ultimately utilized to synthesize ATP. In fact, the proton gradient generated during this process is responsible for approximately 50% of the ATP formed during eukaryotic aerobic metabolism. It is therefore not surprising that a deficiency in cytochrome *c* oxidase activity is the leading cause of respiratory chain defects in humans. Mutations in either CcO or any one of the numerous accessory proteins required for its assembly lead to a variety of multisystemic disorders, including hepatic failure, severe infantile myopathy, Leigh syndrome, and cardiomyopathy (8-10).

One of the fascinating aspects of CcO involves the synthesis and insertion of the heme cofactors. Cytochrome *c* oxidase contains two heme molecules. The first heme is involved in electron transfer, accepting electrons from the dinuclear Cu<sub>A</sub> site and then shuttling the electrons to the second heme (1, 11-14). The second heme molecule interacts with a mononuclear copper ion, and together this Cu-heme heterobimetallic center constitutes the site of O<sub>2</sub> reduction (Figure 1). Interestingly, the hemes contained within CcO are not the more common B-type hemes, but rather they are modified hemes known as heme A (15, 16). Heme A differs from heme B in that one of the vinyl groups is transformed into a hydroxyethylfarnesyl moiety in a reaction catalyzed by heme O synthase (HOS), and a methyl substituent is oxidized to a formyl group by heme A synthase (HAS) (Figure 2). Significantly, both of these modifications are required to obtain stable and functional eukaryotic CcO.

What is the significance of these transformations? Why does Nature synthesize heme A—a heme whose only known function is to be a cofactor in the terminal heme-copper oxidases? Not surprisingly, the two modifications have drastically different effects on the properties of the heme as determined from thermodynamic measurements obtained using a designed “protein maquette” composed of a four- $\alpha$ -helix bundle and containing two heme binding sites (17). The effect of oxidizing the methyl substituent on pyrrole ring D to an electron-withdrawing aldehyde raises the midpoint potential of the maquette-bound heme by approximately +180 mV, presumably making the heme in CcO more easily reduced by the relatively weak reductant cytochrome *c*. The result of this change in redox potential, however, is to destabilize the binding of the ferric heme to the maquette by about 4.1 kcal/mol (17). Thus, the aldehyde affects both the redox potential and heme binding. Conversely, the addition of the farnesyl group to the vinyl substituent on pyrrole ring A stabilizes heme binding to the maquette by at least 40,000-fold, or 6.3 kcal/mol, without significantly altering the redox potential. These results suggest that the aforementioned modifications of the heme B periphery are related and complementary (17). We propose that the oxidation of the methyl substituent to an aldehyde makes heme A a better



*Figure 1. Structure of redox active metal cofactors in CcO (figure prepared in PyMOL). Electrons are delivered from cytochrome c to the dinuclear  $\text{Cu}_A$  site (upper) before being transferred to heme A (lower left). The electrons are then shuttled to the heterobimetallic heme A- $\text{Cu}_B$  site (lower right) where  $\text{O}_2$  is reduced to water. The original coordinates are available in reference 14.*

electron acceptor and therefore aids in CcO activity, while the farnesyl moiety serves to anchor heme A in the enzyme and prevent the heme from dissociating when oxidized to the ferric state.

Despite the critical importance of heme A to aerobic metabolism, many gaps remain in our understanding of heme A biosynthesis, regulation, and transport. This chapter summarizes recent progress in the field and our current understanding of the heme A biosynthetic pathway.

## **Oxidation of Heme O to Heme A**

The oxidation of heme O to heme A is an intriguing reaction for a couple of reasons. For starters, HAS is itself a heme-containing enzyme, utilizing a heme B cofactor (18). Thus, in HAS we observe the unusual situation of a heme cofactor acting on a heme substrate to generate a new heme product. In addition, the conversion of a methyl substituent to an aldehyde is an unusual transformation as a biosynthetic end-product, presumably because of the reactive nature of aldehydes. Not only can aldehydes react with free amines, but the C-H bond of an aldehyde is easier to oxidize than the C-H bond of a methyl group. Thus, if there is enough thermodynamic driving force to oxidize the methyl group, why is the aldehyde product not oxidized to the carboxylate?

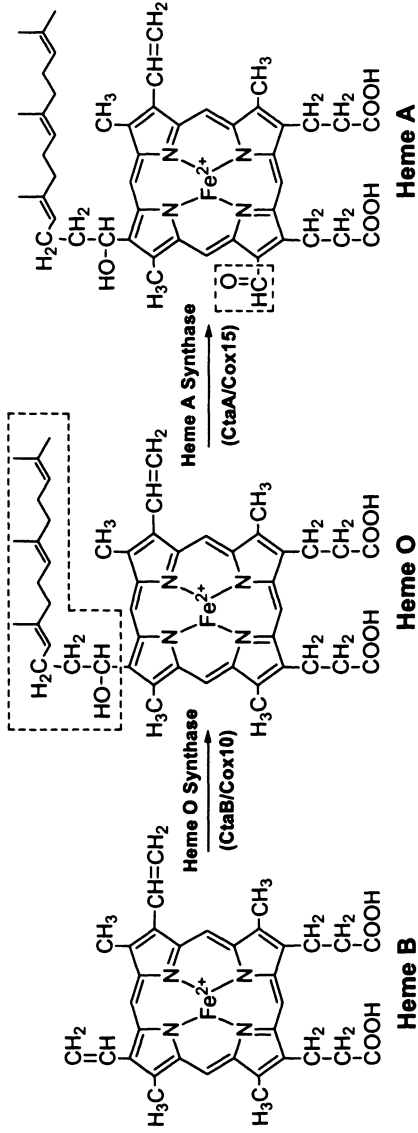


Figure 2. Conversion of heme B into heme A. The CtaB/CtaA nomenclature is used in most bacteria while the Cox10/Cox15 nomenclature is used in eukaryotes and R. sphaeroides. (Reproduced from reference 19. Copyright 2002 American Chemical Society.)

To ascertain the mechanism of heme A biosynthesis, we cloned HOS and HAS from *Bacillus subtilis* and heterologously expressed them in *Escherichia coli*. In addition to observing the production of both heme O and heme A, we also observed two additional previously unidentified heme products. Utilizing a number of techniques including optical spectroscopy, NMR spectroscopy, tandem mass spectroscopy, and chemical synthesis, we identified these two hemes as an alcohol intermediate and an overoxidized carboxylate product (19, 20). The carboxylate derivative, however, has only been identified when HAS is heterologously expressed in *E. coli*; no carboxylate derivative has been observed when HAS is expressed under native conditions.

To determine if HAS is utilizing successive monooxygenase reactions to convert heme O to heme A, we first tested whether the activity of HAS was dependent on O<sub>2</sub> by monitoring HOS and HAS activity under various growth conditions (20). As shown in Figure 3, the amount of heme A isolated from aerobically grown *E. coli* steadily increased for at least two hours after the induction of HOS and HAS. When *E. coli* was grown under nearly anaerobic conditions, however, there was essentially no heme A produced despite the fact that there was a considerable amount of heme O formed. In addition, analysis of protein levels revealed that HAS formation and stability is not altered by the presence or absence of O<sub>2</sub>. Together, these results strongly suggest that the activity of HAS is strictly dependent on O<sub>2</sub>.

If O<sub>2</sub> is required for HAS activity, it would be reasonable to hypothesize that this heme-containing enzyme oxidizes heme O to heme A via successive P450-like monooxygenase reactions. If this is the case, then O<sub>2</sub> should be incorporated into the oxidized products. When *in vivo* assays were performed using 100% <sup>18</sup>O<sub>2</sub>, however, high-resolution mass spectrometry revealed that no <sup>18</sup>O was incorporated into any of the products (Figure 4), including the alcohol intermediate (20). This surprising result indicates that HAS is almost certainly *not* utilizing successive monooxygenase reactions, and thus another mechanism appears to be occurring.

If O<sub>2</sub> is not incorporated into heme A by HAS, what is the source of the oxygen atom? The next most obvious source of the new oxygen atom in heme A is water, and indeed, subsequent experiments in which the *E. coli* was incubated with 95% H<sub>2</sub><sup>18</sup>O (20) indicated that water was incorporated into heme A (numerous controls were performed to correct for the fact that aldehydes can exchange with water). Thus, HAS presents a conundrum: The activity of HAS is strictly dependent on O<sub>2</sub>, and yet O<sub>2</sub> is not incorporated into any of the products. Instead, H<sub>2</sub>O is the source of the oxygen atom in the products.

So what is the mechanism of HAS? The simplest mechanism that is consistent with the available data is that HAS oxidizes heme O to heme A via an electron transfer mechanism. In this scenario, HAS would use O<sub>2</sub> to generate a high valent iron-oxo complex (compound I), but then oxidize heme O via

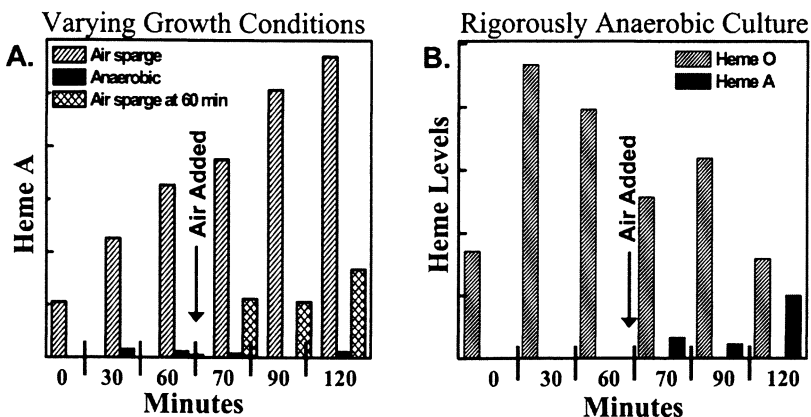


Figure 3. Heme biosynthesis under varying growth conditions.

electron transfer (Figure 5). The resulting carbocation would then be trapped by water, generating the oxidized product with incorporation of the O atom from  $H_2O$ . The fact that both hemes in HAS are known to be coordinated to His residues (like peroxidases) (18, 21) as opposed to Cys residues (like P450s) is consistent with a mechanism that utilizes electron transfer instead of hydrogen atom abstraction.

There are two possible routes by which electron transfer could result in the oxidation of the methyl substituent on heme O. The first is via outer-sphere electron transfer, as depicted in Figure 5. In this mechanism, the cofactor heme B binds and activates  $O_2$  to form compound I, and then heme O is oxidized via a peroxidase-type mechanism. In the second, related mechanism, HAS oxidizes heme O via autoxidation. In this case, heme O binds and activates  $O_2$  to form compound I, while heme B is presumably involved in shuttling electrons from a putative ferredoxin to the active site. Heme O would then be oxidized by internal electron transfer, similar to the mechanism of heme cross-linking elucidated by Ortiz de Montellano and coworkers (22). While the labeling experiments of HAS strongly suggest that heme O is oxidized via electron transfer, they do not allow us to distinguish between these two possible scenarios, and additional experiments are required.

### Transport of Heme O to HAS

So how might heme be transported through the heme A biosynthetic pathway? There are three limiting possibilities. The first possibility is via simple diffusion. Given the toxic nature of free hemes (23, 24), however, this scenario

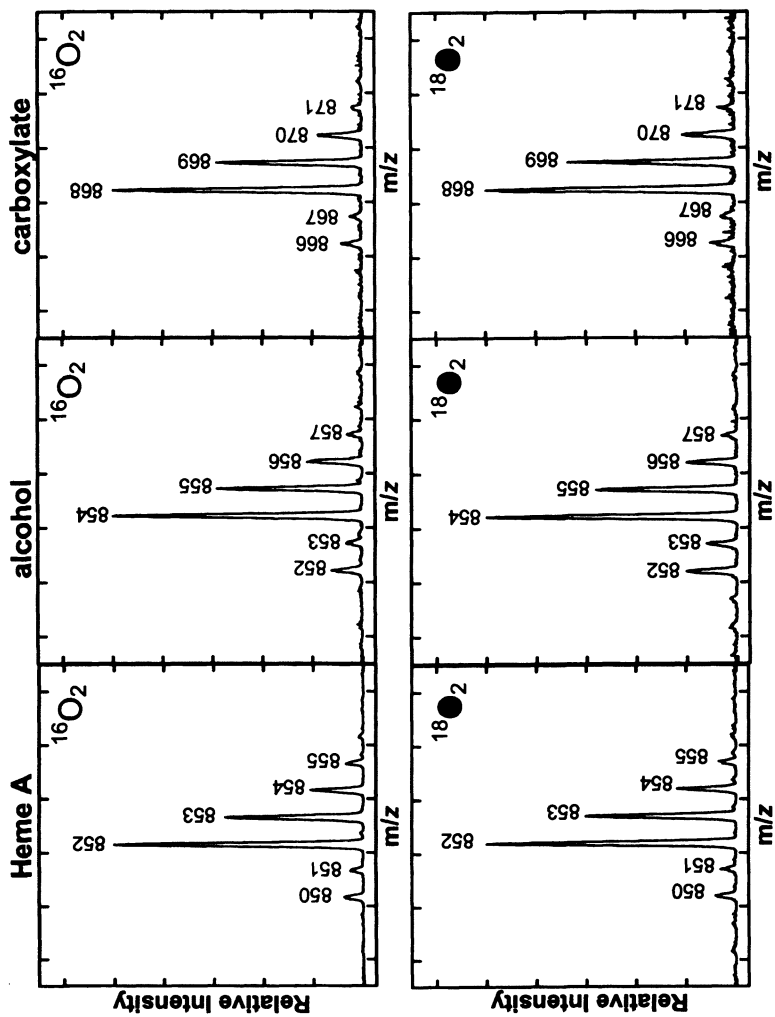


Figure 4. Mass spectral analysis of hemes isolated from cells grown under  $^{16}\text{O}_2$  (top) and cells grown under  $^{18}\text{O}_2$  (bottom). (Reproduced from reference 20. Copyright 2004 American Chemical Society.)

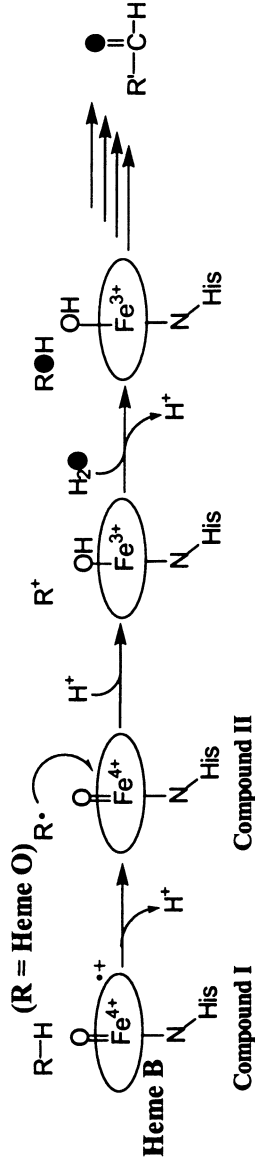


Figure 5. Potential mechanism for heme O oxidation via outer-sphere electron transfer. (Adapted from reference 20.)



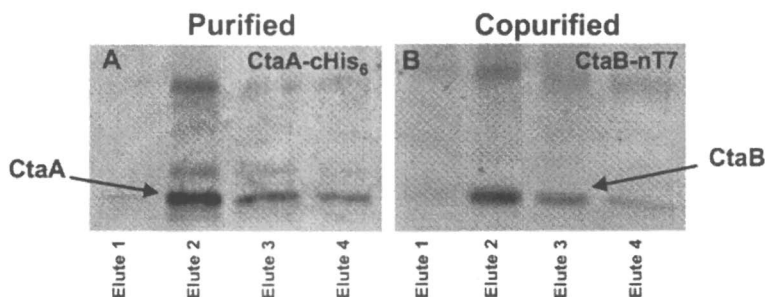
does not seem plausible, and it seems more likely that cells carefully control the transport of this redox-active cofactor. The second possibility is that a “heme chaperone” exists. This would be consistent with the chaperones identified in pathogenic bacteria (25-27) and in *E. coli* cytochrome *c* biogenesis (28, 29). However, to date no chaperone has been identified for the heme A biosynthetic pathway in any organism. The third limiting possibility is that HOS and HAS form a physiologically relevant complex, and that heme O is transferred directly from HOS to HAS without being freely released.

Early preliminary data seemed to support the notion that HOS and HAS interact in bacteria and that heme O is passed directly between them. When HOS was heterologously expressed in *E. coli* alone, approximately 70% of the total extractable heme in *E. coli* was converted into heme O. In the presence of both HOS and HAS, however, the total amount of farnesylated heme produced in *E. coli* decreased. In fact, the presence of HAS appeared to decrease the activity of HOS by approximately 30% (30). Our working hypothesis is that in the absence of HAS, heme O is released by HOS into the membrane. In the presence of its physiological partner, however, HOS must release heme O directly to HAS. Furthermore, either the conversion of heme O to heme A or the release of heme A is the rate-determining step, thus effectively inhibiting the activity of HOS in this “assembly-line” process.

If HOS and HAS form a stable, physiologically relevant complex, then HOS and HAS might be expected to copurify. To test this hypothesis, HOS and HAS from *B. subtilis* were heterologously expressed in *E. coli*. Significantly, when HAS containing a cHis<sub>6</sub> epitope tag was purified on a Ni-NTA column, HOS eluted in the same fractions, indicating that HAS and HOS copurified (Figure 6) (30). Likewise, when HOS was purified on an affinity column, HAS was found to copurify with HOS. Importantly, neither protein bound to the “opposite” affinity column in the absence of its physiological partner.

A number of control experiments were performed to establish that the copurification observed between HOS and HAS was *not* the result of nonspecific interactions (30). In the first experiment, HOS and HAS were expressed in different cells, and the detergent-solubilized cell lysates were then combined and incubated together prior to protein purification. Again, HOS and HAS were found to copurify. In a second experiment, *n*-dodecyl  $\beta$ -D-maltoside was used to solubilize HOS and HAS instead of the detergent Triton X-100, and this change was found to have no effect on the interaction between HOS and HAS. Because *n*-dodecyl  $\beta$ -D-maltoside has a lower micelle number as well as a lower critical micelle concentration (cmc), it seems unlikely that the copurification observed is an artifact of the two enzymes being trapped in the same micelle.

In an additional set of experiments, it was shown that similar results were obtained when HOS and HAS from *Rhodobacter sphaeroides* were heterologously expressed in *E. coli* (30). The presence of HAS seems to decrease the activity of HOS by approximately 30%, and HAS copurifies with



**Figure 6.** Western blot analysis of proteins purified from *E. coli* coexpressing *CtaA* (HAS) and *CtaB* (HOS). (A) *CtaA*-cHis<sub>6</sub> purified on a Ni-NTA column; (B) Same fractions as in gel A probed with an antibody recognizing the nT7 epitope tag on *CtaB*. These results demonstrate that HOS and HAS from *B. subtilis* copurify when expressed in *E. coli*. (Adapted from reference 30. Copyright 2004 American Chemical Society.)

HOS. Because *R. sphaeroides* is a member of the class of proteobacteria from which the mitochondrion is thought to have developed, both HOS and HAS (as well as the catalytic subunits of CcO) share a high degree of sequence homology with their eukaryotic counterparts. Thus, what we observe in *R. sphaeroides* may very well be more generally applicable to include eukaryotic organisms. However, no direct interaction has yet been observed in yeast between HOS and HAS despite numerous attempts, and any extrapolation between the work reported here and potential interactions between eukaryotic HOS and HAS must therefore be treated with caution.

Nevertheless, the copurification of HOS and HAS observed in bacterial systems support the notion that HOS and HAS form a physiologically relevant complex. Furthermore, the fact that the presence of HAS appears to decrease the activity of HOS is consistent with the hypothesis that heme O is transferred directly from HOS to HAS, and that either the conversion of heme O into heme A or the release of the heme A product is the rate-determining step in this heterologous expression system. Thus, the heme A biosynthetic pathway may be at least partially regulated by product release.

## Regulation of the Heme A Biosynthetic Pathway

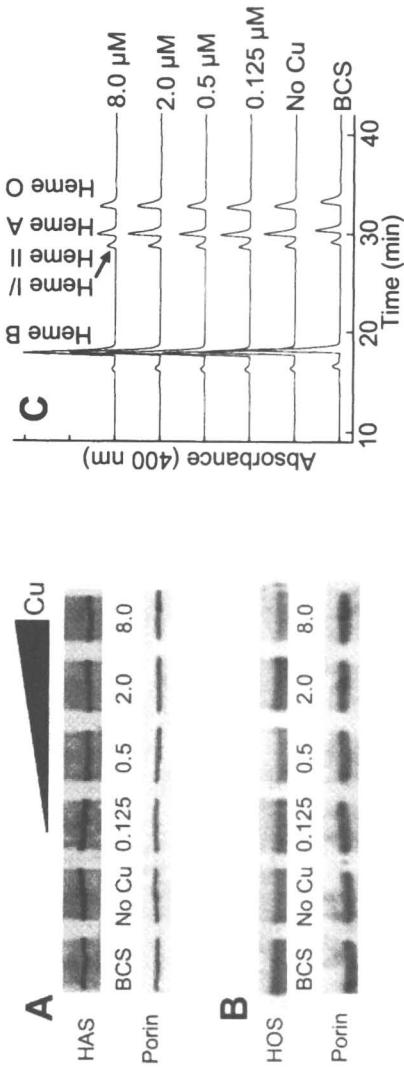
Another critical aspect of heme A biosynthesis involves the regulation of the heme A biosynthetic pathway. Because free hemes are toxic to cells (23, 24), there is tremendous evolutionary pressure to maintain tight control over the heme A biosynthetic pathway to ensure that sufficient, but not excessive, quantities of heme A are available. How is the flux of heme through the

HOS/HAS pathway regulated? In *B. subtilis*, the transcription of the operons containing HOS and HAS is controlled in part by the *resD* and *strC* gene products (31, 32). Currently it is not clear how HOS and HAS are regulated once they are transcribed, although as stated above, product release may play a partial role. In eukaryotes, however, very little is known about the regulation of the heme A biosynthetic pathway (33). Obvious factors that one might predict would affect the levels of proteins involved in CcO biosynthesis, such as O<sub>2</sub> concentrations (34), respiratory competency of the cell (35), fermentable versus nonfermentable carbon sources (36), and rich versus minimal media (37), were reported not to affect the transcription of either HOS or HAS in *S. cerevisiae*.

It was previously suggested that copper might play a role in the regulation of the heme A biosynthetic pathway. A regulatory role for copper makes intuitive sense because both copper and heme A are obligatory cofactors in CcO, and CcO is the only known physiological target of heme A. Why synthesize heme A if there is not a sufficient level of copper available? In addition, there are numerous examples where copper and iron/heme homeostasis are intertwined (38-40), as well as ample circumstantial evidence that copper might directly affect heme A homeostasis (41-43).

We sought to test the hypothesis that copper was involved in the regulation of the heme A biosynthetic pathway. *S. cerevisiae* was grown under various copper conditions, and the HOS and HAS protein levels were ascertained by Western blot analysis (44). The steady state levels of HOS and HAS were not altered as the intracellular copper levels changed, even when copper was essentially absent (Figure 7), indicating that copper does not regulate the transcription, translation, or stability of either HOS or HAS in *S. cerevisiae*. To ascertain if copper regulates the activity of either HOS or HAS, however, it was necessary to switch to a different organism because both CcO and its heme cofactors are degraded in *S. cerevisiae* in the absence of copper. HOS and HAS from *B. subtilis* were heterologously expressed in *E. coli*, and the activity of each enzyme was analyzed by monitoring intracellular heme levels (44). Intracellular copper concentrations had no effect on the production of either heme O, heme A, or the alcohol/carboxylate derivatives (Figure 7), indicating that copper does not directly regulate either HOS or HAS. One possibility, however, is that copper regulates HOS or HAS activity indirectly through another protein. However, because HOS and HAS were heterologously expressed, the possibility of indirect regulation could not be addressed in this experiment.

To determine if copper affects the activity of HOS or HAS either directly or indirectly, we performed analogous experiments in *R. sphaeroides* in which both enzymes were under native control. *R. sphaeroides* CcO is stable in the absence of both heme A and copper, and therefore any change in the amount of heme A observed with varying intracellular copper levels cannot be attributable to the absence of CcO, and must be the result of either direct or indirect regulation of heme A levels via copper. As in the case of *E. coli*, no variation in heme A levels in *R. sphaeroides* was detected as a function of copper concentration (44).



**Figure 7.** Effect of copper concentrations (expressed in  $\mu\text{M}$ ) on HOS and HAS. "No Cu" medium was Chelex-treated to remove divalent metal ions, while the "BCS" medium was Chelex-treated and had the copper-specific chelator bathocuproinedisulfonic acid ( $\pm \text{CuSO}_4$ ) added to the growth medium to a final concentration of  $100 \mu\text{M}$ . In all cases, high-purity metals ( $\pm \text{CuSO}_4$ ) were returned to the growth medium (see Ref. 43); panels A and B: Western blot analyses of HAS (A) and HOS (B) from cells grown in various medium [Cu] with 2% glucose as a fermentable carbon source (YPD); panel C: HPLC analyses of hemes produced by *B. subtilis* HOS and HAS when expressed in *E. coli* in the presence of varying [Cu] in the medium. Heme I and heme II refer to the alcohol intermediate and the carboxylate derivative of heme A, respectively. (Reproduced from reference 44. Copyright 2005 American Chemical Society.)

Together, these data indicate that copper is not involved in regulating heme A biosynthesis at any point along the pathway.

### Regulation of Heme A Biosynthesis via Intracellular Heme Levels

If copper does not influence the heme A biosynthetic pathway, then how is this critical pathway controlled? To address this question, we developed a new reporter plasmid termed pCOX15/lacZ (COX15 is the gene that codes for HAS in eukaryotes) in which the promoter region of the HAS gene from *S. cerevisiae* was fused to lacZ (Figure 8). The lacZ gene encodes  $\beta$ -galactosidase, whose activity can be easily and quantitatively established via a standard assay (45). Thus, the amount of  $\beta$ -galactosidase activity will be a measure of the level of lacZ transcription, which in turn will be a measure of how strongly the promoter is turned on. By transforming this new plasmid into *S. cerevisiae*, we were able to probe the effect that a variety of different growth conditions had on the HAS promoter.

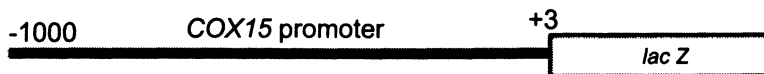


Figure 8. Reporter plasmid used to screen conditions that alter the regulation of COX15 (HAS).

Initial results indicated that intracellular heme levels might regulate the heme A biosynthetic pathway. To test this hypothesis in more detail, pCOX15/lacZ was transformed into an ALA ( $\delta$ -aminolevulinic acid) synthase deletion mutant (*hem1 $\Delta$* ) (46). ALA synthase is the first step in heme biosynthesis (47), and this deletion mutant can be overcome by the addition of either ALA or heme to the growth medium, allowing us to manipulate intracellular heme levels. As expected,  $\beta$ -galactosidase activity increased with increasing intracellular heme levels (46). Measuring the transcript level of genomic HAS via reverse transcriptase real-time PCR as well as measuring the corresponding protein levels yielded results that were consistent with those obtained using pCOX15/lacZ, *i.e.*, HAS is positively regulated by intracellular heme levels.

Hap1 (heme activator protein 1) is a heme-regulated transcription factor found in *S. cerevisiae* (48), and we hypothesized that this protein might play a role in the regulation of HAS transcription. Consistent with this hypothesis, the

$\beta$ -galactosidase activity of *S. cerevisiae hap1* $\Delta$  cells transformed with pCOX15/lacZ was independent of intracellular heme levels (46). Significantly, a putative Hap1 binding site was located in the promoter region of HAS, and mutation of this binding site completely obliterated regulation of HAS via heme. Together, these results indicate that the transcription of HAS is positively regulated by intracellular heme levels via Hap1.

Given the fact that HOS and HAS are part of the same pathway, and the fact that neither heme O nor heme A has another known function except as a precursor and cofactor for CcO, respectively, one might expect HOS and HAS to be regulated in a similar manner. However, when analogous experiments were performed using the promoter of HOS, no change in the transcript level or protein level of heme O could be observed (46). Furthermore, the levels of HOS were not influenced by the presence/absence of HAS. Thus, it appears as if HOS and HAS are differentially regulated, and the physiological significance of this finding remains to be elucidated.

## Conclusions

The regulation, transport, and biosynthesis of hemes along the heme A biosynthetic pathway is a complex process that is only beginning to be elucidated. Our data reveal that HAS utilizes a unique, O<sub>2</sub>-dependent electron transfer mechanism to oxidize the methyl substituent of heme O during the biosynthesis of heme A. Furthermore, we have established that HOS and HAS can copurify, and we hypothesize that HOS and HAS form a physiologically relevant complex *in vivo* (at least in bacteria), and that heme O is transferred directly from HOS to HAS without being freely released. Finally, we have established that in yeast, HAS is positively regulated by intracellular heme levels via the transcription factor Hap1.

## Acknowledgments

This work was supported by the National Institutes of Health (GM 66236 to E.L.H.).

## References

1. Ferguson-Miller, S.; Babcock, G. T. *Chem. Rev.* **1996**, *96*, 2889-2907.
2. Garcia-Horsman, J. A.; Barquera, B.; Rumbley, J.; Ma, J.; Gennis, R. B. *J. Bacteriol.* **1994**, *176*, 5587-5600.
3. Michel, H.; Behr, J.; Harrenga, A.; Kannt, A. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 329-356.

4. Malmström, B. G. In *Biological Systems and Artificial Supramolecular Systems*; Balzani, V., Ed.; Wiley-VCH: New York, NY, 2001; Vol. 3, p 39-55.
5. Namslauer, A.; Brzezinski, P. *FEBS Lett.* **2004**, *567*, 103-110.
6. Trumpower, B. L.; Gennis, R. B. *Annu. Rev. Biochem.* **1994**, *63*, 675-716.
7. Hosler, J. P.; Ferguson-Miller, S.; Mills, D. A. *Annu. Rev. Biochem.* **2006**, *75*, 165-187.
8. Antonicka, H.; Leary, S. C.; Guercin, G.-H.; Agar, J. N.; Horvath, R.; Kennaway, N. G.; Harding, C. O.; Jaksch, M.; Shoubridge, E. A. *Hum. Mol. Genet.* **2003**, *12*, 2693-2702.
9. Antonicka, H.; Mattman, A.; Carlson, C.; Glerum, D. M.; Hoffbuhr, K. C.; Leary, S. C.; Kennaway, N. G.; Shoubridge, E. A. *Am. J. Hum. Genet.* **2003**, *72*, 101-114.
10. Shoubridge, E. A. *Am. J. Med. Genet.* **2001**, *106*, 46-52.
11. Abramson, J.; Shimamura, T.; Iwata, S. In *Respiration in Archaea and Bacteria: Diversity of Prokaryotic Electron Transport Carriers*; Zannoni, D., Ed.; Kluwer Academic Publishers: Dordrecht, 2004; Vol. 15, p 115-128.
12. Abramson, J.; Svensson-Ek, M.; Byrne, B.; Iwata, S. *Biochim. Biophys. Acta* **2001**, *1544*, 1-9.
13. Brzezinski, P.; Larsson, G.; Ädelroth, P. In *Respiration in Archaea and Bacteria: Diversity of Prokaryotic Electron Transport Carriers*; Zannoni, D., Ed.; Kluwer Academic Publishers: Dordrecht, 2004; Vol. 15, p 129-153.
14. Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. *Science* **1996**, *272*, 1136-1144.
15. Barros, M. H.; Carlson, C. G.; Glerum, D. M.; Tzagoloff, A. *FEBS Lett.* **2001**, *492*, 133-138.
16. Mogi, T. In *The Iron and Cobalt Pigments: Biosynthesis, Structure, and Degradation*; Kadish, K. M.; Smith, K. M.; Guillard, R., Eds.; Academic Press: Amsterdam, 2003; Vol. 12, p 157-181.
17. Zhuang, J.; Reddi, A. R.; Wang, Z.; Khodaverdian, B.; Hegg, E. L.; Gibney, B. R. *Biochemistry* **2006**, *45*, 12530-12538.
18. Svensson, B.; Anderson, K. K.; Hederstedt, L. *Eur. J. Biochem.* **1996**, *238*, 287-295.
19. Brown, K. R.; Allan, B. M.; Do, P.; Hegg, E. L. *Biochemistry* **2002**, *41*, 10906-10913.
20. Brown, K. R.; Brown, B. M.; Hoagland, E.; Manye, C. L.; Hegg, E. L. *Biochemistry* **2004**, *43*, 8616-8624.
21. Hederstedt, L.; Lewin, A.; Throne-Holst, M. *J. Bacteriol.* **2005**, *187*, 8361-8369.
22. Colas, C.; Ortiz de Montellano, P. R. *Chem. Rev.* **2003**, *103*, 2305-2332.
23. Carlsen, C. U.; Møller, J. K. S.; Skibsted, L. H. *Coord. Chem. Rev.* **2005**, *249*, 485-498.
24. Rytter, S. W.; Tyrrell, R. M. *Free Radical Biol. Med.* **2000**, *28*, 289-309.

25. Eakanunkul, S.; Lukat-Rodgers, G. S.; Sumithran, S.; Ghosh, A.; Rodgers, K. R.; Dawson, J. H.; Wilks, A. *Biochemistry* **2005**, *44*, 13179-13191.
26. Genco, C. A.; Dixon, D. W. *Mol. Microbiol.* **2001**, *39*, 1-11.
27. Lansky, I. B.; Lukat-Rodgers, G. S.; Block, D.; Rodgers, K. R.; Ratliff, M.; Wilks, A. *J. Biol. Chem.* **2006**, *281*, 13652-13662.
28. Stevens, J. M.; Daltrop, O.; Allen, J. W. A.; Ferguson, S. J. *Acc. Chem. Res.* **2004**, *37*, 999-1007.
29. Thöny-Meyer, L. *Biochemistry* **2003**, *42*, 13099-13105.
30. Brown, B. M.; Wang, Z.; Brown, K. R.; Cricco, J. A.; Hegg, E. L. *Biochemistry* **2004**, *43*, 13541-13548.
31. Liu, X.; Taber, H. W. *J. Bacteriol.* **1998**, *180*, 6154-6163.
32. Paul, S.; Zhang, X.; Hulett, F. M. *J. Bacteriol.* **2001**, *183*, 3237-3246.
33. Barros, M. H.; Tzagoloff, A. *FEBS Lett.* **2002**, *516*, 119-123.
34. Kwast, K. E.; Lai, L.-C.; Menda, N.; James, D. T., III; Aref, S.; Burke, P. V. *J. Bacteriol.* **2002**, *184*, 250-265.
35. Epstein, C. B.; Waddle, J. A.; Hale, W. I. V.; Dave, V.; Thornton, J.; Macatee, T. L.; Garner, H. R.; Butow, R. A. *Mol. Biol. Cell* **2001**, *12*, 297-308.
36. Gasch, A. P.; Spellman, P. T.; Kao, C. M.; Carmel-Harel, O.; Eisen, M. B.; Storz, G.; Botstein, D.; Brown, P. O. *Mol. Biol. Cell* **2000**, *11*, 4241-4257.
37. Washburn, M. P.; Koller, A.; Oshiro, G.; Ulaszek, R. R.; Plouffe, D.; Deciu, C.; Winzeler, E.; Yates, J. R., III *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3107-3112.
38. Crisp, R. J.; Pollington, A.; Galea, C.; Jaron, S.; Yamaguchi-Iwai, Y.; Kaplan, J. *J. Biol. Chem.* **2003**, *278*, 45499-45506.
39. Lin, S.-J.; Pufahl, R. A.; Dancis, A.; O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **1997**, *272*, 9215-9220.
40. Nakayama, K.; Takasawa, A.; Terai, I.; Okui, T.; Ohyama, T.; Tamura, M. *Arch. Biochem. Biophys.* **2000**, *375*, 240-250.
41. Glerum, D. M.; Shtanko, A.; Tzagoloff, A. *J. Biol. Chem.* **1996**, *271*, 14504-14509.
42. Keyhani, E.; Keyhani, J. *Biochim. Biophys. Acta* **1980**, *633*, 211-227.
43. Williams, D. M.; Loukopoulos, D.; Lee, G. R.; Cartwright, G. E. *Blood* **1976**, *48*, 77-85.
44. Morrison, M. S.; Cricco, J. A.; Hegg, E. L. *Biochemistry* **2005**, *44*, 12554-12563.
45. Silverman, S. J.; Rose, M.; Botstein, D.; Fink, G. R. *Mol. Cell. Biol.* **1982**, *2*, 1212-9.
46. Wang, Z.; Hegg, E. L. *J. Biol. Chem.* in press.
47. Urban-Grimal, D.; Volland, C.; Garnier, T.; Dehoux, P.; Labbe-Bois, R. *Eur. J. Biochem.* **1986**, *156*, 511-519.
48. Zhang, L.; Hach, A. *Cell. Mol. Life Sci.* **1999**, *56*, 415-426.



## Chapter 4

# Assembly of Cytochrome *c* Oxidase

## Synthesis and Insertion of the Metal Cofactors

Zhihong Wang and Eric L. Hegg\*

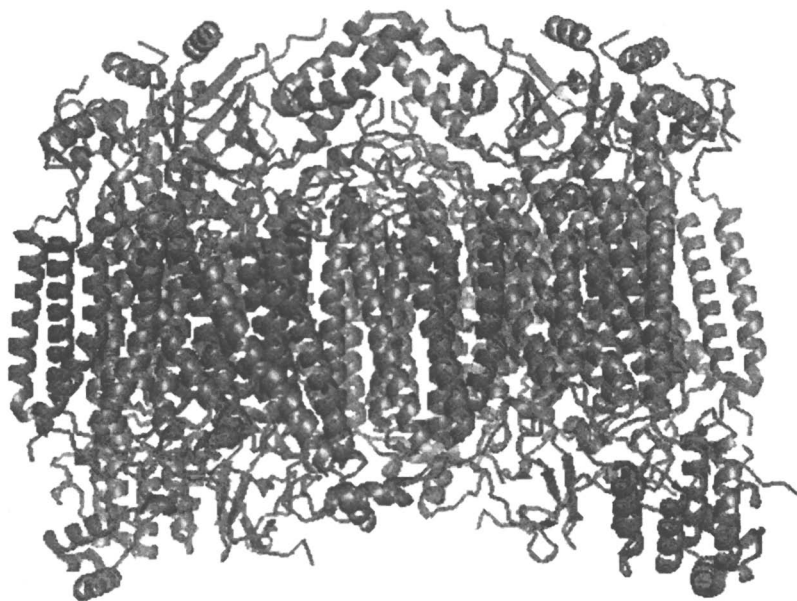
Department of Biochemistry and Molecular Biology,  
Michigan State University, East Lansing, MI 48824

Cytochrome *c* oxidase requires numerous metal cofactors for its activity. This chapter reviews the assembly of cytochrome *c* oxidase, with special attention to the synthesis and insertion of the obligatory heme A and copper metal cofactors as well as their functions in the assembly of the entire cytochrome *c* oxidase complex.

The respiratory chain is composed of four multiple-subunit complexes, NADH dehydrogenase (complex I), succinate dehydrogenase (complex II), cytochrome *c* reductase (complex III) and cytochrome *c* oxidase (complex IV, CcO) (1). The four complexes, located in the inner mitochondrial membrane of eukaryotes and the inner cytoplasmic membrane of prokaryotes, are electronically connected by ubiquinone and cytochrome *c*, which transfer electrons through complex I or complex II to complex III, and finally to complex IV, where molecular oxygen is reduced to water. Concurrently, protons are pumped across the inner mitochondrial membrane of eukaryotes or the cytoplasmic membrane of prokaryotes. The proton gradient is utilized by ATP synthase (complex V) to synthesize ATP. In many organisms, the respiratory complexes and complex V are assembled into supercomplexes which have been

identified and characterized by blue native polyacrylamide gel electrophoresis (PAGE) and electron microscopy (2, 3).

Complex IV, commonly called cytochrome *c* oxidase (abbreviated: CcO, COX, or complex IV), is the terminal oxidase of the respiratory chain (4). CcO accepts electrons from reduced cytochrome *c* and then transfers these electrons to molecular oxygen, reducing O<sub>2</sub> to water. This process is coupled to pumping protons across the membrane to generate a proton gradient which is ultimately utilized by ATP synthase (5). Eukaryotic CcO is composed of up to 13 subunits, three of which are encoded by mitochondrial DNA (mtDNA) while the remaining ten subunits are encoded by the nuclear genome (4) (Figure 1). In addition to these 13 subunits, eukaryotic CcO also contains multiple metal cofactors required for its assembly and function, including two heme A molecules (heme *a* and heme *a*<sub>3</sub>), two copper centers (Cu<sub>A</sub> and Cu<sub>B</sub>), magnesium, sodium, and zinc (6). The heme *a* molecule accepts electrons from the Cu<sub>A</sub> site and transfers electrons to the bimetallic heme *a*<sub>3</sub>-Cu<sub>B</sub> site, the catalytic site of CcO where molecular oxygen is reduced to water (Figure 2). The exact functions of the magnesium, sodium, and zinc ions remain unknown. This chapter reviews the current knowledge covering the assembly of CcO, especially the maturation and insertion of the heme A and copper cofactors.



*Figure 1. Crystal structure of cytochrome c oxidase from bovine heart, a homodimer as rendered in PyMOL (original coordinates published in Ref. 4).*

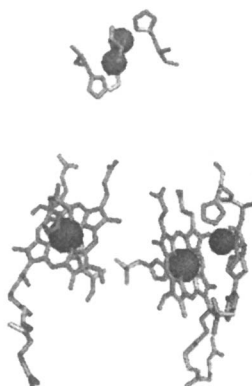


Figure 2. Structure of metal centers and ligands in bovine heart cytochrome *c* oxidase as rendered by PyMOL (original coordinates were published in Ref. 4).

## Assembly of Cytochrome *c* Oxidase

The yeast *Saccharomyces cerevisiae*, *Rhodobacter sphaeroides*, and human cell lines have been mainly utilized to elucidate the mechanism of CcO assembly and the individual roles of the accessory proteins. Although it has been hypothesized for more than 20 years that the assembly of CcO is a sequential process (7), the details concerning this process remain poorly understood. There are three major obstacles to understanding CcO assembly. First, in addition to the structural subunits, yeast CcO requires more than two dozen nuclear-encoded accessory proteins for its proper assembly and function (many of which have a human homolog), highlighting the complexity of CcO assembly (8); these accessory proteins are involved in the transcription, translation, and insertion of mtDNA-encoded subunits, in the import and processing of nuclear-encoded subunits, in the biosynthesis and insertion of the heme A cofactors, and in copper homeostasis. Second, the absence of any subunit or cofactor often causes the other subunits to be degraded, making it difficult to distinguish between primary and secondary phenotypes. In addition, unassembled mitochondrial proteins from yeast CcO are extremely prone to proteolysis. Human mitochondria, however, have a slower turnover rate of unassembled protein, and some accumulated assembly intermediates can therefore be detected with blue native PAGE. Third, defects in other respiratory complexes, such as complex III, often cause similar phenotypes as mutants with defects in CcO subunits or accessory proteins (e.g., decreased steady-state concentrations of structural subunits and cofactors), resulting in multiple pleiotropic effects.

High-resolution crystal structures of CcO from bovine heart as well as those of the *aa*<sub>3</sub>-type CcO from the protobacteria *Paracoccus denitrificans* and

*Rhodobacter sphaeroides* were recently solved (4, 9, 10). In addition, molecular genetics and biochemical studies have provided a wealth of information regarding the assembly of CcO. Many of these studies have been performed on yeast by analyzing the respiratory-deficient *Saccharomyces cerevisiae* mutants (commonly called *pet* mutants) which are conveniently selected for the ability to grow on glucose but not on non-fermentable carbon sources (8). By comparing the cytochrome spectra and the respiratory complex activity, the CcO-deficient strains from the *pet* mutant collection can be easily identified. *Rhodobacter sphaeroides* has one  $aa_3$ -type CcO which is highly homologous to the catalytic core of mammalian CcO. Given the fact that the  $aa_3$ -type CcO from protobacteria is simpler and more stable, *Rhodobacter sphaeroides* is often utilized to study eukaryotic CcO. Human cell lines from patients have also been extensively studied to elucidate the assembly of CcO and to identify the assembly intermediates.

It has been observed that most human diseases related to CcO deficiency are caused by mutations in one of the genes encoding CcO accessory proteins, including Cox10, Cox15, Surf1/Shyl, Sco1, and Sco2. Mutations in the mtDNA-encoded CcO genes are very rare, and no diseases associated with a mutation in a nuclear-encoded CcO structural subunit have yet been identified (11). Therefore, in this section additional details on the functions of the CcO accessory proteins Cox10, Cox15, Surf1/Shyl, Sco1, and Sco2 will be provided. The CcO accessory proteins discussed in this chapter are listed in Table I.

### Subunits of CcO

Mammalian CcO contains 13 subunits while the yeast *Saccharomyces cerevisiae* CcO has 11 subunits. In both organisms, CcO has three highly conserved subunits which are encoded by the mitochondrial DNA (mtDNA), named Cox1, Cox2, and Cox3. Cox1, the largest and most hydrophobic subunit, is composed of 12 transmembrane helices and contains two redox active metal centers, heme *a* and a heme  $a_3$ -Cu<sub>B</sub> binuclear metal center. Cox1 also contains one sodium ion, although the functional significance of this site is not currently known. Cox2, the smallest and the least hydrophobic core subunit with only two transmembrane helices, contains the binuclear mixed-valent Cu<sub>A</sub> site. The function of the magnesium ion found at the interface of Cox1 and Cox2 remains unknown, although it may help to prevent the back-leak of protons through the water-exit channel (12). Cox3 is composed of seven transmembrane helices and does not contain any metal cofactors. Unlike the mtDNA-encoded subunits, the functions of the nuclear-encoded subunits of CcO, which exist mainly in eukaryotes, remain largely unknown. The crystal structure from bovine-heart CcO reveals that subunit Vb (homologous to Cox4 of yeast CcO) contains one zinc ion, but no other metal ions have been identified in any of the nuclear-encoded subunits. Site-directed mutagenesis demonstrated that Zn(II) coordi-

**Table I. Proteins Involved in Yeast CcO Assembly**

<i>Protein</i>	<i>Description</i>
Cox10	Heme O synthase, converts heme B to heme O
Cox15	Heme A synthase, converts heme O to heme A
Shy1	May facilitate the maturation of heme $\alpha_3$ site
Mss51	Involved in Cox1 translation, regulation, and assembly
Cox14	Involved in Cox1 translation, regulation, and assembly
Coal	CcO accessory protein, may be involved in Cox1 synthesis
Cox17	Copper chaperone, involved in CcO assembly
Cox19	Copper chaperone, involved in CcO assembly
Cox23	Copper chaperone, required for CcO assembly
Sco1	Copper chaperone, delivers copper to Cu <sub>A</sub> site in Cox2
Sco2	Exact function unknown, may have a redundant function
Cox11	Copper chaperone, delivers copper to Cu <sub>B</sub> site in Cox1

nation to Cox4 is essential to the assembly and stabilization of CcO in yeast (13).

### **Cox10/Cox15 (Heme O Synthase/Heme A Synthase) and Related Heme A Insertion**

Heme A is an essential cofactor in cytochrome *c* oxidase. Deficient heme A biosynthesis has been linked to ageing and age-related diseases, such as Alzheimer's disease (AD) (14). When compared to healthy controls, heme A in Alzheimer's patients is decreased by 22%, while heme B and heme C levels remain largely unchanged, leading to the hypothesis that the decreased CcO activity in Alzheimer's patients and elderly patients might be due to the decreased heme A levels (15). In heme-deficient conditions, CcO activity is selectively decreased while other respiratory complexes remain intact, presumably because heme A biosynthesis is more susceptible to heme deficiency than other pathways (16, 17).

Heme A is derived from the more generally utilized heme B. Heme O synthase (Cox10) catalyzes the first reaction by the addition of the farnesyl moiety to heme B to generate heme O, while heme A synthase (Cox15) converts heme O to heme A by oxidizing the C-8 methyl substituent to an aldehyde (18-21) (Figure 3).

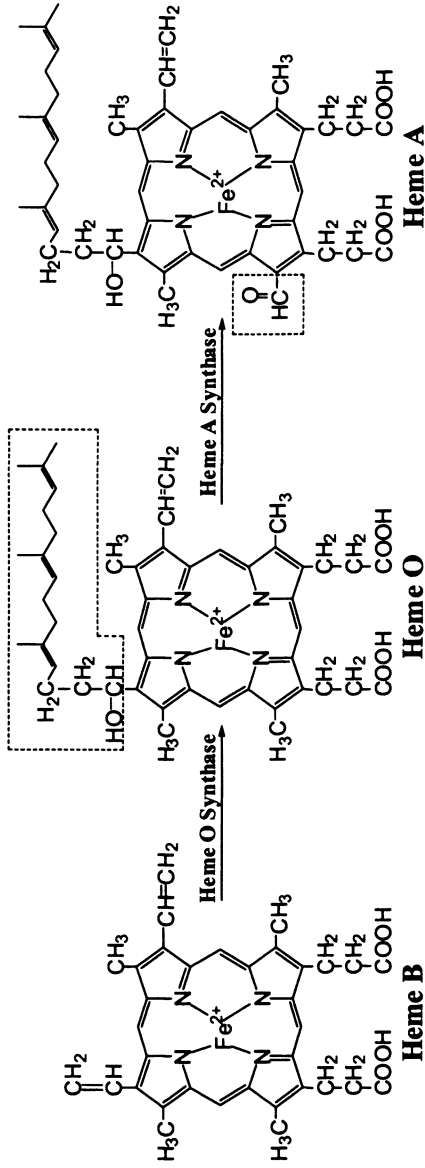


Figure 3. Transformation of heme B to heme A catalyzed by the enzymes heme O synthase (Cox10) and heme A synthase (Cox15).

One puzzling question is how heme A is incorporated into CcO, the final and only known target of heme A. It was previously demonstrated that the assembly of Cox1 in the mitochondrion is a heme A-dependent process (22). Mutations on *COX10* or *COX15* in humans cause CcO deficiency, and no subassembly complexes accumulate, indicating that heme A is required not only for functional CcO, but also for the stable assembly of the CcO complex (23). The identification of a heme A-Cox1 assembly intermediate and the lack of a Cox1-Cox4-Cox5 subassembly complex in *COX10*-deficient human cells support the notion that heme A insertion is the early step of CcO assembly (22-26). The crystal structure of CcO reveals that heme A is deeply buried inside the inner mitochondrial membrane. It was subsequently hypothesized that heme A may enter the hydrophobic pocket of Cox1 from the IMS (intermembrane space) side of the IM (inner membrane) (27, 28).

Interestingly, the deletion of the Surf1 homolog in *R. sphaeroides* led to a 50% loss of heme  $a_3$ , indicating that Surf1 facilitates the insertion of heme  $a_3$  either directly or indirectly in this organism (29). This was the first demonstration that Surf1 might be involved in metal cofactor maturation. Studies in a *shy1* (yeast homolog of Surf1) null strain demonstrate that overexpression of Cox15 (heme A synthase) did not increase the heme  $a_3$ -Cox1 assembly intermediate as other null strains did, which is consistent with the role of Shy1 in heme  $a_3$  maturation (24). The deletion of Surf1/Shy1 does not abolish the CcO activity completely, however, and it is therefore possible that there are some unidentified proteins that are also involved in heme A insertion. Another hypothesis is that HAS interacts with Cox1 transiently and directly transfers heme A to Cox1, a process which could be aided by accessory proteins such as Shy1/Surf1.

### **Sco1/Sco2 and Related Copper Homeostasis in the Mitochondrion**

Because of the toxicity of free copper ions, they are transferred from one protein to another via direct protein-protein interaction. The two copper sites, Cu<sub>A</sub> and Cu<sub>B</sub>, are essential to the function and assembly of CcO. In the absence of either cofactor, CcO is not functional and gets degraded in eukaryotes. The delivery and insertion of copper into CcO has been well studied, especially in the yeast *S. cerevisiae*, and a number of small and cysteine-rich proteins engaged in mitochondrial copper homeostasis have been identified including Cox17, Cox19, Cox23, Cox11, and the Sco family of proteins. The copper transport pathway in eukaryotic CcO involves Cox17 delivering copper ions to Cox11 and Sco1. Cox11 facilitates copper insertion into the Cu<sub>B</sub> site in subunit I of CcO, while Sco1 mediates maturation of the Cu<sub>A</sub> site in subunit II.

Cox17, a 8-kDa hydrophilic protein, is highly conserved in eukaryotic organisms. The *cox17* null mutant is CcO deficient but can be partially rescued by the addition of copper to growth medium (30). Cox17 has six conserved

cysteine residues involved in redox reactions and copper binding, and Cox17 can bind up to four copper ions per monomer when fully reduced (31). Mutational analysis of the conserved cysteine residues demonstrated that the CCXC motif is essential for Cox17 function (32). Because *cox17* and *sco1* mutants both cause the complete loss of Cox2 (the location of the Cu<sub>A</sub> site), and because *cox17* mutants can be suppressed by *Sco1* overexpression, it was initially hypothesized that Cox17 exclusively delivers copper to *Sco1* (33). Subsequent *in vitro* studies, however, revealed that Cox17 transfers Cu(I) to both Cox11 and *Sco1*, two proteins responsible for Cu<sub>B</sub> insertion and Cu<sub>A</sub> insertion, respectively (34). The docking interfaces between Cox17 and both Cox11 and *Sco1* have been proposed, although no stable interaction has been detected, presumably due to the transient nature of the interaction between these proteins.

Interestingly, Cox17 is localized both in the cytoplasm and in the intermembrane space (IMS), perhaps suggesting a secondary function of delivering cytoplasmic copper to the mitochondrion (35). However, Cox17 is still functional when tethered to the inner mitochondrial membrane by a heterologous transmembrane sequence, and the mitochondrial copper concentration is not decreased in the *cox17* null mutants (36). One possible explanation for these results is that cytosolic Cox17 localization simply results from incomplete mitochondrial uptake. A second possibility is that the primary function of Cox17 is shuttling copper ions to *Sco1* and Cox11 inside the mitochondrion, and that transporting copper ions from the cytosol to the mitochondrion is a secondary function in which Cox17 plays a redundant role.

Recently, two other proteins, Cox19 and Cox23, were identified as new copper chaperones. The two proteins are structurally homologous to Cox17 and indeed have a number of similarities to Cox17 including solubility, dual location to both the cytoplasm and the IMS, conserved cysteine residues, and CcO deficiency in the null mutants. Furthermore, Cox19 is also functional when it is tethered to the IM. Nevertheless, Cox19 and Cox17 play different roles because *COX17* is not a high-copy suppressor of *cox19* null mutants, nor can *COX19* rescue *cox17* null mutants (37). Interestingly, *cox19* null mutants cannot be rescued by adding exogenous copper to the medium, while the *cox23* null mutants can only be rescued by exogenous copper when Cox17 is overexpressed. Overexpression of Cox23 has no effect on *cox17* mutants (38). To date, no physical interaction between Cox17 and Cox19 or Cox23 has been detected, and the specific functions of both Cox23 and Cox19 remain unknown.

*Sco1* and *Sco2* are highly similar to each other (59% similarity between yeast *Sco* proteins and 39% between human *Sco* proteins), and each has the CXXXC copper-binding motif. The two *Sco* proteins are believed to deliver copper ions to the Cu<sub>A</sub> site, which is exposed in the IMS. In humans, both *SCO1* and *SCO2* are essential genes, and mutations of either of them are associated with CcO deficiency, although *SCO1* patients have a clinical presentation which is different from that of *SCO2* patients. In yeast, however, *sco1* mutants are



respiratory incompetent while *sco2* mutants are not (39). Sequence alignments reveal that the human *SCO* genes are not orthologous to the two yeast *SCO* genes (40), which presumably explains why mutations in the two organisms produce different phenotypes. In the following two paragraphs, the functions of human Sco and yeast Sco will therefore be discussed separately.

Studies on human Sco1 and Sco2 established that both can stably bind one Cu(I) or Cu(II) ion, and that the copper-binding ability is essential to their function *in vivo* (41). Overexpression of Sco1 in the *SCO2* patient cells and overexpression of Sco2 in the *SCO1* patient cells further decreased the CcO activity, indicating that human Sco1 and Sco2 have non-overlapping but cooperative functions in copper insertion into CcO (42). One model for Cu<sub>A</sub> maturation in humans is that Cox17 delivers copper to Sco2, which then shuttles it to Cox2 with the help of Sco1. The crystal structure of human Sco1 is similar to redox-active proteins such as thioredoxins and peroxiredoxin, suggesting a possible role in mitochondrial signaling (43). Recent studies demonstrated that Sco1 and Sco2 also act as the molecular signal to regulate the cellular copper homeostasis, and that Sco2 has a prominent role in the signaling pathway (44).

In yeast, Sco1 and Sco2 are anchored to the inner mitochondrial membrane and protrude into the IMS. The function of Sco2 in yeast remains unresolved, while the role of the Sco1 protein in the maturation of Cu<sub>A</sub> site has been well established (39, 45-49). The crystal structure of Sco1 from yeast show that yeast Sco1 also contains a CXXXC motif and a thioredoxin-like fold, similar to human Sco1 and *B. subtilis* Sco1 (45, 46, 50). The identified physical interaction between Sco1 and Cox2 clearly demonstrates that Sco1 transfers copper ions to the Cu<sub>A</sub> site of Cox2, and the interaction between them likely occurs via electrostatic forces (46, 49). The Cu<sub>A</sub> site contains two copper ions, yet Sco1 only shuttles one Cu(I) at a time. Thus, it is possible that the Cu<sub>A</sub> site requires an additional copper chaperone.

Cox11 is highly conserved in some bacteria and most eukaryotes. The aa<sub>3</sub>-type CcO purified from *R. sphaeroides* cells lacking Cox11 contains heme *a*, heme *a*<sub>3</sub>, and Cu<sub>A</sub>, but not Cu<sub>B</sub>, indicating that Cox11 specifically delivers copper to the Cu<sub>B</sub> site. Cox11 is an integral inner-mitochondrial membrane protein, with an “N-in, C-out” topology. The soluble C-terminus is located in the IMS and binds one Cu(I) per monomer via cysteine residues (51). Interestingly, Cox11 associates with the mitochondrial translation machinery, suggesting that the formation of the Cu<sub>B</sub> site might occur concurrently with the translation and insertion of Cox1 into the membrane. This scenario would explain how Cox11 delivers copper to the Cu<sub>B</sub> site, which is deeply buried in the hydrophobic pocket of Cox1 (51, 52).

## Surf1/Shy1

The mammalian Surf1 protein and its yeast homolog, Shy1, are the most extensively studied accessory proteins because most patients afflicted with

Leigh syndrome symptoms have mutations in the *SURF1* gene. As mentioned previously, Surf1/Shyl was proposed to facilitate the maturation of the heme  $a_3$  site. However, the exact function of Surf1/Shyl remains unknown.

An assembly intermediate containing Cox1, Cox4 (corresponding to Cox5a in yeast), and Cox5 (corresponding to Cox6 in yeast) was detected in those patients. In patients and yeast cells, a mutation in Surf1/Shyl reduces the total amount of CcO dramatically, although residual amounts of fully assembled complex are still detectable, suggesting that Shyl/Surf1 either has a redundant function or that it increases the efficiency of other accessory proteins (53, 54). An interaction between Surf1 and Cox2 in human fibroblasts was also suggested (54).

Recently, *MSS51* was identified as a suppressor of *shyl* null mutants. Mss51 is required for Cox1 translation and elongation and is part of a transient Cox14-Cox1-Mss51 complex which down-regulates Cox1 synthesis (24, 55, 56). It is proposed that Shyl promotes the full expression of Cox1 by catalyzing the release of free Mss51 from the Cox14-Cox1-Mss51 complex (56). Interactions between Shyl and Mss51, Cox14, Coa1, and subcomplexes of CcO were detected recently (57, 58), which is consistent with previous data in both yeast and humans showing that Shyl/Surf1 is part of a high molecular weight complex (54). Interestingly, Shyl and Cox14 were also found to associate with the CcO-cytochrome  $bc_1$  supercomplex in *S. cerevisiae* (57). Whether this part of the assembly process applies to other organisms is unknown, because Mss51, Coa1 and Cox14 are only weakly conserved in fungi.

*HAP4* is another multicopy suppressor of *shyl* null mutants and is believed to function by causing the overexpression of subunits Cox5a and Cox6 (59). Interestingly, the suppression of the phenotype observed in *shyl* null cells by the overexpression of Hap4 and Mss51 are additive. Copper supplementation to the media also partially restores the respiratory defect of the yeast *shyl* mutant. It is possible that copper supplementation promotes the formation of the  $Cu_B$  site, thereby increasing the stability of Cox1-Cox5a-Cox6 complex. In conclusion, Shyl/Surf1 is involved in the early steps of CcO assembly, and it may be involved in the insertion of heme  $a_3$ , full expression of Cox1, and/or the incorporation of Cox2 into the Cox1-Cox5a-Cox6 subassembly complex.

### CcO Assembly Line

The assembly of the multiple-subunit CcO complex is a sequential process (7), and more than two dozen accessory proteins are required to coordinate this process effectively (8). Metabolic labeling experiments on cultured human cells identified two CcO assembly intermediates, subcomplexes S2 and S3 (26). S2 is composed of Cox1 and Cox4, which is consistent with the fact that Cox1 and Cox4 interact in the crystal structure (4). S3 is composed of all the subunits except Cox6a and Cox7b. The homolog of Cox6a in yeast, encoded by *COX13*,

is not essential for CcO assembly (60), suggesting that Cox6a is incorporated into CcO late in the assembly process.

Investigations on the CcO subassembly complexes in fibroblast cultures from patients carrying *COX10*, *SURF1*, or *SCO1* mutations provide additional important information regarding CcO maturation. *COX10*-deficient fibroblasts lack the subassembly complex Cox1-Cox4-Cox5a, which was identified both in *SURF1*-deficient cells and *SCO1*-deficient cells, suggesting that the incorporation of heme A into Cox1 occurs earlier than the incorporation of Cox4 and Cox5a, consistent with the observation of a Cox1-heme A intermediate. In addition, *SCO1*-deficient cells do not accumulate the subassembly complex containing Cox1 and Cox2, suggesting that Cox2 associates with Cox1-Cox4-Cox5a after Cu<sub>A</sub> insertion. *SURF1*-deficient cells accumulate similar subassembly complexes as *SCO1*-deficient cells, indicating that SURF1 may also facilitate the association of Cox2 with Cox1-Cox4-Cox5a.

In the yeast *S. cerevisiae*, a subassembly complex containing Cox1, Cox3, Cox5, and Cox6 (the corresponding nomenclature of CcO subunits in mammals are Cox1, Cox3, Cox4, and Cox5a) was identified in *COX2* mutants, demonstrating that the assembly order is conserved between yeast and humans (61). In addition, the presence of an early, pro-oxidant heme A-Cox1 intermediate (as observed in humans) was supported by the fact that cells lacking Cox1 or heme A are peroxide-resistant while other mutants lacking Cox2, Cox11 or Sco1 are peroxide-sensitive (24). The proposed model for the assembly order of human CcO is shown in Figure 4 (62).

## Conclusions

In conclusion, the assembly of CcO is a sophisticated process which requires many proteins for transcription, translation, cofactor biosynthesis and transport, assembly, and regulation. The biosynthesis and/or insertion of the metal cofactors, including heme A and copper, is especially intriguing and complex. These events require multiple proteins acting in a highly orchestrated process to ensure proper CcO assembly. In the past three decades, much progress has been made towards understanding the genotype-phenotype relationship, especially in yeast. Continued studies are required to explore the remaining puzzles concerning the CcO accessory proteins and their related human homologs.

## Acknowledgments

This work was supported by the National Institutes of Health (GM 66236 to E.L.H.).

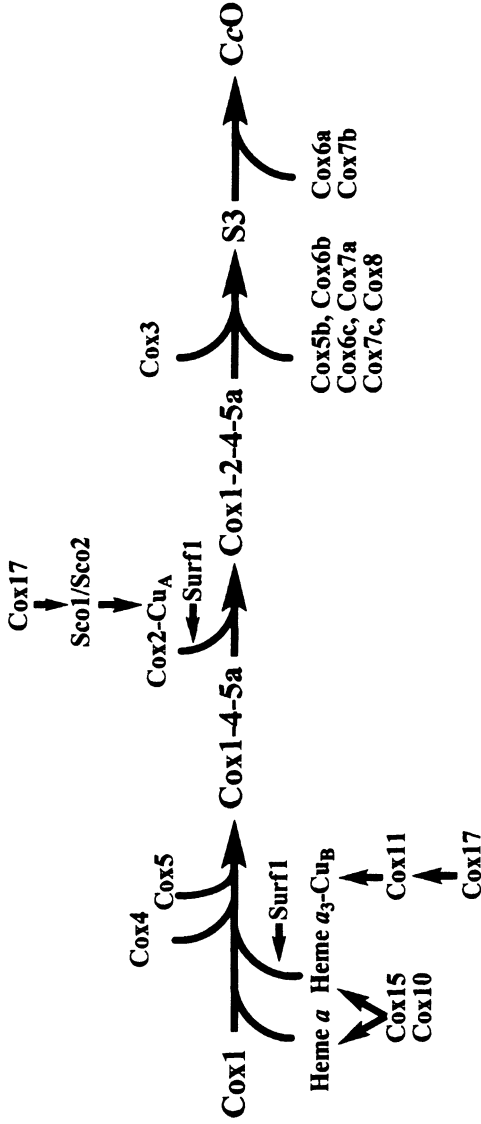


Figure 4. Proposed model of the assembly order of human cytochrome c oxidase (adapted from Ref. 62).

## References

1. Schagger, H. *Biochim. Biophys. Acta* **2002**, *1555*, 154-159.
2. Schafer, E.; Dencher, N. A.; Vonck, J.; Parcej, D. N. *Biochemistry* **2007**, *46*, 12579-12585.
3. Stroh, A.; Anderka, O.; Pfeiffer, K.; Yagi, T.; Finel, M.; Ludwig, B.; Schagger, H. *J. Biol. Chem.* **2004**, *279*, 5000-5007.
4. Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawaitoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. *Science* **1996**, *272*, 1136-1144.
5. Brown, B. M.; Wang, Z. H.; Brown, K. R.; Cricco, J. A.; Hegg, E. L. *Biochemistry* **2004**, *43*, 13541-13548.
6. Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawaitoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. *Science* **1995**, *269*, 1069-1074.
7. Wielburski, A.; Nelson, B. D. *Biochem. J.* **1983**, *212*, 829-834.
8. Tzagoloff, A.; Dieckmann, C. L. *Microbiol. Rev.* **1990**, *54*, 211-225.
9. Qin, L.; Hiser, C.; Mulichak, A.; Garavito, R. M.; Ferguson-Miller, S. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 16117-16122.
10. Iwata, S.; Ostermeier, C.; Ludwig, B.; Michel, H. *Nature* **1995**, *376*, 660-669.
11. Pecina, P.; Houstkova, H.; Hansikova, H.; Zeman, J.; Houstek, J. *Physiol. Res.* **2004**, *53 Suppl 1*, S213-223.
12. Schmidt, B.; McCracken, J.; Ferguson-Miller, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15539-15542.
13. Coyne, H. J., III; Ciofi-Baffoni, S.; Banci, L.; Bertini, I.; Zhang, L.; George, G. N.; Winge, D. R. *J. Biol. Chem.* **2007**, *282*, 8926-89234.
14. Atamna, H. *Ageing Res. Rev.* **2004**, *3*, 303-318.
15. Parker, W. D., Jr.; Parks, J.; Filley, C. M.; Kleinschmidt-DeMasters, B. K. *Neurology* **1994**, *44*, 1090-1096.
16. Ojaimi, J.; Masters, C. L.; Opeskin, K.; McKelvie, P.; Byrne, E. *Mech. Ageing Dev.* **1999**, *111*, 39-47.
17. Atamna, H.; Liu, J.; Ames, B. N. *J. Biol. Chem.* **2001**, *276*, 48410-48416.
18. Puustinen, A.; Wikström, M. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 6122-6126.
19. Brown, K. R.; Allan, B. A.; Do, P.; Hegg, E. L. *Biochemistry* **2002**, *41*, 10906-10913.
20. Barros, M. H.; Carlson, C. G.; Glerum, D. M.; Tzagoloff, A. *FEBS Lett.* **2001**, *492*, 133-138.
21. Mogi, T. In *The Iron and Cobalt Pigments: Biosynthesis, Structure, and Degradation*; Kadish, K. M.; Smith, K. M.; Guillard, R., Eds.; Academic Press: Amsterdam, 2003; Vol. 12, pp 157-181.
22. Wielburski, A.; Nelson, B. D. *FEBS Lett.* **1984**, *177*, 291-294.

23. Williams, S. L.; Valnot, I.; Rustin, P.; Taanman, J. W. *J. Biol. Chem.* **2004**, *279*, 7462-7469.
24. Khalimonchuk, O.; Bird, A.; Winge, D. R. *J. Biol. Chem.* **2007**, *282*, 17442-17449.
25. Herrmann, J. M.; Funes, S. *Gene* **2005**, *354*, 43-52.
26. Nijtmans, L. G. J.; Taanman, J. W.; Muijsers, A. O.; Speijer, D.; Van den Bogert, C. *Eur. J. Biochem.* **1998**, *254*, 389-394.
27. Cobine, P. A.; Pierrel, F.; Winge, D. R. *Biochim. Biophys. Acta* **2006**, *1763*, 759-772.
28. Carr, H. S.; Winge, D. R. *Acc. Chem. Res.* **2003**, *36*, 309-316.
29. Smith, D.; Gray, J.; Mitchell, L.; Antholine, W. E.; Hosler, J. P. *J. Biol. Chem.* **2005**, *280*, 17652-17656.
30. Glerum, D. M.; Shtanko, A.; Tzagoloff, A. *J. Biol. Chem.* **1996**, *271*, 14504-14509.
31. Palumaa, P.; Kangur, L.; Voronova, A.; Sillard, R. *Biochem. J.* **2004**, *382*, 307-314.
32. Heaton, D.; Nittis, T.; Srinivasan, C.; Winge, D. R. *J. Biol. Chem.* **2000**, *275*, 37582-37587.
33. Punter, F. A.; Glerum, D. M. *J. Biol. Chem.* **2003**, *278*, 30875-30880.
34. Horng, Y. C.; Cobine, P. A.; Maxfield, A. B.; Carr, H. S.; Winge, D. R. *J. Biol. Chem.* **2004**, *279*, 35334-35340.
35. Beers, J.; Glerum, D. M.; Tzagoloff, A. *J. Biol. Chem.* **1997**, *272*, 33191-33196.
36. Maxfield, A. B.; Heaton, D. N.; Winge, D. R. *J. Biol. Chem.* **2004**, *279*, 5072-5080.
37. Rigby, K.; Zhang, L.; Cobine, P. A.; George, G. N.; Winge, D. R. *J. Biol. Chem.* **2007**, *282*, 10233-10242.
38. Barros, M. H.; Johnson, A.; Tzagoloff, A. *J. Biol. Chem.* **2004**, *279*, 31943-31947.
39. Winge, D. R. *Structure* **2003**, *11*, 1313-1314.
40. Papadopoulou, L. C.; Sue, C. M.; Davidson, M. M.; Tanji, K.; Nishino, I.; Sadlock, J. E.; Krishna, S.; Walker, W.; Selby, J.; Glerum, D. M.; Coster, R. V.; Lyon, G.; Scalais, E.; Lebel, R.; Kaplan, P.; Shanske, S.; De Vivo, D. C.; Bonilla, E.; Hirano, M.; DiMauro, S.; Schon, E. A. *Nat. Genet.* **1999**, *23*, 333-337.
41. Horng, Y. C.; Leary, S. C.; Cobine, P. A.; Young, F. B.; George, G. N.; Shoubridge, E. A.; Winge, D. R. *J. Biol. Chem.* **2005**, *280*, 34113-34122.
42. Leary, S. C.; Kaufman, B. A.; Pellicchia, G.; Guercin, G. H.; Mattman, A.; Jaksch, M.; Shoubridge, E. A. *Hum. Mol. Genet.* **2004**, *13*, 1839-1848.
43. Williams, J. C.; Sue, C.; Banting, G. S.; Yang, H.; Glerum, D. M.; Hendrickson, W. A.; Schon, E. A. *J. Biol. Chem.* **2005**, *280*, 15202-15211.
44. Leary, S. C.; Cobine, P. A.; Kaufman, B. A.; Guercin, G. H.; Mattman, A.; Palaty, J.; Lockitch, G.; Winge, D. R.; Rustin, P.; Horvath, R.; Shoubridge, E. A. *Cell Metab.* **2007**, *5*, 9-20.

45. Imriskova-Sosova, I.; Ye, Q.; Hill, B. C.; Jia, Z. *Acta Crystallogr. D. Biol. Crystallogr.* **2003**, *59*, 1299-1301.
46. Balatri, E.; Banci, L.; Bertini, I.; Cantini, F.; Cioffi-Baffoni, S. *Structure* **2003**, *11*, 1431-4143.
47. Beers, J.; Glerum, D. M.; Tzagoloff, A. *J. Biol. Chem.* **2002**, *277*, 22185-22190.
48. Nittis, T.; George, G. N.; Winge, D. R. *J. Biol. Chem.* **2001**, *276*, 42520-42526.
49. Lode, A.; Kuschel, M.; Paret, C.; Rodel, G. *FEBS Lett.* **2000**, *485*, 19-24.
50. Abajian, C.; Rosenzweig, A. C. *J. Biol. Inorg. Chem.* **2006**, *11*, 459-466.
51. Banting, G. S.; Glerum, D. M. *Eukaryot. Cell* **2006**, *5*, 568-578.
52. Khalimonchuk, O.; Ostermann, K.; Rodel, G. *Curr. Genet.* **2005**, *47*, 223-233.
53. Tiranti, V.; Galimberti, C.; Nijtmans, L.; Bovolenta, S.; Perini, M. P.; Zeviani, M. *Hum. Mol. Genet.* **1999**, *8*, 2533-2540.
54. Nijtmans, L. G.; Artal Sanz, M.; Bucko, M.; Farhoud, M. H.; Feenstra, M.; Hakkaart, G. A.; Zeviani, M.; Grivell, L. A. *FEBS Lett.* **2001**, *498*, 46-51.
55. Barrientos, A.; Korr, D.; Tzagoloff, A. *EMBO J.* **2002**, *21*, 43-52.
56. Barrientos, A.; Zambrano, A.; Tzagoloff, A. *EMBO J.* **2004**, *23*, 3472-3482.
57. Mick, D. U.; Wagner, K.; van der Laan, M.; Frazier, A. E.; Perschil, I.; Pawlas, M.; Meyer, H. E.; Warscheid, B.; Rehling, P. *EMBO J.* **2007**, *26*, 4347-4358.
58. Pierrel, F.; Bestwick, M. L.; Cobine, P. A.; Khalimonchuk, O.; Cricco, J. A.; Winge, D. R. *EMBO J.* **2007**, *26*, 4335-4346.
59. Fontanesi, F.; Jin, C.; Tzagoloff, A.; Barrientos, A. *Hum. Mol. Genet.* **2008**, *17*, 775-788.
60. Taanman, J. W.; Capaldi, R. A. *J. Biol. Chem.* **1993**, *268*, 18754-18761.
61. Horan, S.; Bourges, I.; Taanman, J. W.; Meunier, B. *Biochem. J.* **2005**, *390*, 703-708.
62. Stiburek, L.; Hansikova, H.; Tesarova, M.; Cerna, L.; Zeman, J. *Physiol. Res.* **2006**, *55*, S27-S41.





## Chapter 5

# New Approaches to Analyzing the Site Selectivities and Crystal Structures of DNA Targeted Metal Complexes

### Analysis of the Antitumor Agent Bleomycin

Eric C. Long<sup>1,\*</sup>, Millie M. Georgiadis<sup>1,2,\*</sup>, Kristie D. Goodwin<sup>2</sup>,  
and Mark A. Lewis<sup>1</sup>

<sup>1</sup>Department of Chemistry and Chemical Biology,  
Purdue School of Science, and <sup>2</sup>Department of Biochemistry and Molecular  
Biology, Indiana University School of Medicine,  
Indiana University–Purdue University Indianapolis (IUPUI),  
Indianapolis, IN 46202

New approaches to determining the site-selectivities of DNA targeted metal complexes and their DNA-bound crystal structures are described as applied to bleomycin, a metal binding antitumor agent. These approaches include a fluorescent intercalator displacement assay and a host-guest DNA-ligand co-crystallization strategy. The application of these approaches, respectively, confirmed that the binding site-selectivities of Co(III)• and Fe(III)•bleomycin parallel the established cleavage selectivity of this drug and led to the first crystal structures of duplex DNA-bound Co(III)•bleomycin.

Cellular DNA continues to be an important target for compounds and metal complexes of biomedical significance (1, 2). Indeed, given the rapid accumulation of information concerning gene sequences, protein binding sites,

and other DNA-based targets being revealed through genome sequencing efforts, it is important to be able to develop and investigate DNA-targeted agents at a similar rate (3, 4). Towards this end, the use of combinatorial strategies to facilitate the discovery of nucleic acid-targeted metal complexes with select properties is occurring and likely to increase (5, 6). Unfortunately, these efforts make it apparent that our ability to generate DNA-targeted metal complexes has far out-paced the throughput capabilities of the traditional methodologies used to characterize DNA-metal complex binding events—our ability to determine the DNA site-selectivities of such systems and their DNA-bound structures continues to rely on methods such as footprinting (7), the inherent DNA cleavage properties of a metal complex to mark a location of DNA interaction (8), and established oligonucleotide co-crystallization or NMR protocols (9, 10); these strategies evolved at a time when the characterization of a single DNA binding agent, *e.g.*, bleomycin (Figure 1), required years of effort by multiple research groups (11-17) to elucidate a binding mechanism, site selectivity profile, or DNA-bound structure. Given this situation, new approaches are needed to expedite the analysis of DNA-targeted metal complexes.

In an effort aimed at developing a rapid method to determine the DNA site-selectivities of low molecular weight agents, a fluorescent intercalator displacement (FID) assay was developed recently by Boger (18, 19). This new approach complements existing, technically more demanding strategies (20) to determine the DNA binding site-selectivities of individual agents. The FID assay is also amenable to the screening of compound libraries to identify components with the ability to bind to a DNA sequence of interest. For the above, rank orders of DNA binding site preferences *or* compounds, respectively, can be generated in a short timeframe, often less than a day.

In addition to knowledge of the site-selectivity of an agent, structural information is often necessary to fully evaluate a DNA binding ligand. In this area, our laboratories have developed a host-guest (H-G) DNA crystallization strategy for the analysis of site-selective DNA binding ligands (21). The H-G methodology was developed originally in the Georgiadis laboratory to crystallize and analyze nucleic acid sequences of biological significance (22-25) and was subsequently extended by Georgiadis and Long to include DNA-ligand interactions (21, 26). As detailed later, the H-G crystallization methodology can facilitate the structural analysis of a DNA-ligand interaction by providing a DNA-protein crystal lattice “framework” that can accommodate a variety of DNA binding ligands. Further, H-G crystallizations can be micro-seeded leading to the rapid generation of diffraction-quality crystals. Indeed, we have applied the H-G crystallization strategy in tandem with the FID strategy to rapidly characterize a novel DNA minor groove binding organic cation (26).

Herein we describe the application of the strategies highlighted above to the study of DNA binding metal complexes. As an example, we describe the FID analysis and H-G crystallization of metalbleomycin (Figure 1), a clinically employed antitumor agent that has served as a paradigm for the development and

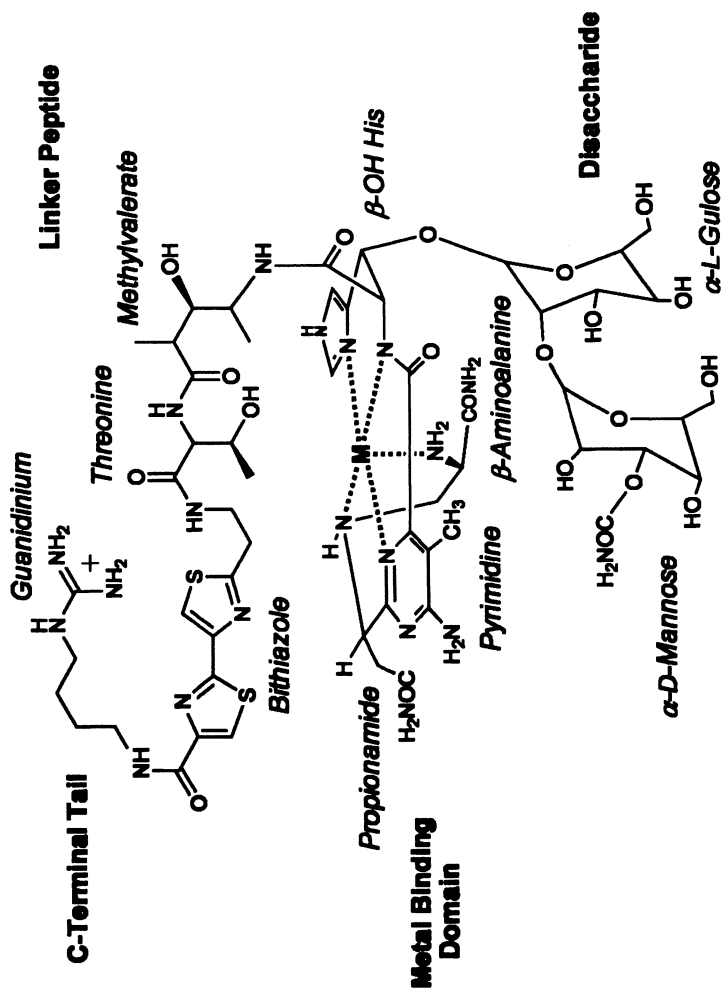


Figure 1. Structure of metallobleomycins.

discovery of DNA-targeted metal complexes for more than 40 years (2, 11-17). Bleomycins constitute a family of glycopeptide-derived natural products isolated from *Streptomyces verticillus* (27). The clinical efficacy of the bleomycins (28, 29) is thought to derive from their ability to mediate single- and double-strand DNA duplex cleavage (30, 31); however RNA may constitute another important cellular target (32, 33). The DNA cleavage activity of bleomycin is  $\text{Fe}^{2+}$  and  $\text{O}_2$  dependent (34) and results from selective C4'-H abstraction (35) from pyrimidine nucleotides contained within 5'-GT(C) sites (36, 37). Bleomycin also binds Cu *in vivo* (38), a species competent to induce DNA strand scission (39), and other metal ions *in vitro* including Co and Zn (11, 14). Bleomycin is modular in its overall structure (Figure 1): a metal binding domain is connected via a linker peptide to a positively-charged bithiazole/C-terminal "tail" that increases the overall drug-DNA binding affinity; the metal binding domain is also modified with a disaccharide moiety that contributes to the efficiency of DNA cleavage (13, 16).

Despite extensive study, a thorough, comparative examination of the DNA binding site-selectivity of metalbleomycin apart from its cleavage site-selectivity has not been carried out nor has a crystallographic analysis of its DNA-bound structure been achieved. Investigations towards both these ends are reviewed herein. Overall, metalbleomycin serves as an excellent example through which to showcase the FID assay and the H-G crystallization methodology to illustrate how these approaches can be applied to DNA-targeted metal complexes.

## Fluorescence Intercalator Displacement

The FID assay can be used to determine the site-selectivity of a DNA binding agent comprehensively through the use of libraries of synthetic hairpin oligonucleotide binding targets containing a variable sequence cassette (18, 19). Collectively, these cassettes can be designed to represent all possible 4- or 5-base pair (or longer) sequences nested in a common hairpin sequence; there are 136 and 512 unique hairpins for the 4- and 5-base pair libraries, respectively. Microplate display of the individual members of an oligonucleotide library in the presence of a DNA intercalating fluorophore (ethidium bromide or thiazole orange) results in nearly uniform fluorescence emission from each sample-well due to the essentially non-specific DNA intercalation of these reporter molecules. Upon addition of a DNA binding ligand to each microplate well, sequences that contain a preferred binding site for the ligand exhibit decreased fluorescence due to competitive displacement of the reporting fluorophore. While titrations are necessary to determine accurate affinity constants, the extent of decreased fluorescence ( $\Delta F$ ) correlates to the DNA affinity of the analyte. A typical sample well contains 1.5  $\mu\text{M}$  DNA (hairpin), 4.5  $\mu\text{M}$  ethidium bromide, buffer/0.1M NaCl, and a concentration of analyte that can be sub-stoichiometric or exceed the hairpin concentration depending upon its DNA affinity.

Results obtained in the above manner can be rank-ordered as a function of the % *F* decreases of all hairpins and cast in the form of a histogram. Such an analysis quickly reveals the DNA binding site preferences of an agent of interest—triplicate analyses can be performed in a few hours leading to data that cannot be obtained typically without extensive DNA footprinting or cleavage experiments using multiple DNA fragments to test all base pair sequences.

### Previous Applications

The FID assay has been used to assess the DNA binding site-selectivities of a variety of organic minor groove binding agents including netropsin, distamycin, berenil, Hoechst 33258, DAPI, hairpin polyamides, actinomycin, and a phenyl-benzimidazole diamidine (18, 19, 26, 40, 41). As applied to metal complexes, an FID-derived strategy was used recently to examine enantiomers of dinuclear polypyridyl Ru(II) complexes that target the minor groove non-intercalatively (42). In this case it was reported that intercalating fluorophores yielded results that were inconsistent with the findings of other techniques (NMR, DNA affinity chromatography, and equilibrium dialysis) during the analysis of the binding of these complexes to an oligonucleotide bearing a single-base bulge. To resolve these inconsistencies, a non-intercalating, groove binding fluorophore, DAPI, was used in lieu of an intercalator; this change in reporting fluorophore yielded FID results that now reflected the findings of the other techniques (42) most likely due to the similar DNA binding modes used by both DAPI and the analyte. These results emphasize that the DNA binding mode of an analyte needs to be considered, if known, when optimizing the FID assay for a new application.

### Metal Complexes of Bleomycin

Most of our understanding of the DNA site-selectivity of bleomycin is inferred through knowledge of the locations of drug-induced DNA cleavage. For example, DNA cleavage by Fe(II)•bleomycin + O<sub>2</sub> is predominantly focused to 5'-GC and 5'-GT sites, with lesser extents occurring at 5'-GA and 5'-GG sites; this suggests that drug binding occurs to similar relative extents at these locations. While it is likely that this assumption reflects the behavior of some metalbleomycins, it may not be applicable to all, nor DNA cleaving agents in general. Indeed, even in the comparison of two metalbleomycins (Fe and Co) thought to be structurally analogous (43-46), to rely upon sites of DNA strand scission alone as a means to report binding selectivity was found to be misleading due to inherent differences in their mechanisms of DNA cleavage (43): this study concluded that while Fe• and Co•bleomycin occupy similar DNA

binding sites, they produce different cleavage site-selectivities because irradiation of Co(III)•bleomycin likely leads to a diffusible oxidant along with an oxidant directed to the site of actual drug-DNA binding. Thus, the FID assay provides an opportunity to examine the site-selectivities of metallobleomycins and other agents in a fashion that reports *only* on binding activity.

### *Co(III)•Bleomycin “Green”*

Co(III)•bleomycin B<sub>2</sub> “green” contains a stable, metal-bound hydroperoxide and is thought to closely model the DNA recognition and structure of Fe(II)•bleomycin + O<sub>2</sub> (43-46); this metallobleomycin does not cleave DNA in the absence of irradiation. Co(III)•bleomycin “green” was examined using the FID assay (40) resulting in the merged-bar histogram shown in Figure 2. FID plot curvature was observed indicating that the metaldrug bound selectively to the 136 hairpin targets provided through a 4-base pair cassette library. Selectivity also was evident by the clear discrimination that occurred against the A/T-only hairpin cassettes (all appearing at the lowest end of the rank order) vs. the other cassettes. In comparison, apo-bleomycin B<sub>2</sub> and Co<sup>2+</sup> salts alone failed to decrease any sample fluorescence. Analysis of the FID data indicated that across the range of drug concentrations studied, among the top ranked 15 DNA hairpins, 14 contained at least two 5'-GT or 5'-GC sites (when the four base pair cassette and the flanking hairpin sequences were considered). Overall, the top 50% of the ranked 136 hairpins *all* contained at least one 5'-GT or 5'-GC site while only 16 of the lowest-ranked 50% contained a single 5'-GT or 5'-GC site. Therefore, overall drug binding site rank-order clearly depended upon the presence of 5'-GC and 5'-GT sites.

A further perusal of the top five ranked sequences (Figure 2) indicated that they were the same at all drug concentrations examined and that 5'-ACGT and 5'-ACGC were observed consistently within the top three. Notably, the 5'-ACGT site was ranked first at all drug concentrations examined, including drug concentrations that were less than one equivalent relative to hairpin (40); 5'-ACGC was ranked second at most concentrations. Considering their sequences, these cassettes contain two inverted 5'-GT(C) bleomycin target sites suggesting a drug binding preference for sites that may enable the association of two metallobleomycin equivalents. While current structural models (16, 47) indicate that these sites are too close to accommodate the full, simultaneous association of two metal binding domains, concurrent bithiazole intercalation by two drug equivalents could however occur in such a way as to allow minor groove association of one or the other drug metal binding domain, perhaps in rapid exchange. In support of this view, quantitative titrations indicated that the 5'-ACGT hairpin was able to bind two equivalents of Co(III)•bleomycin (40) and crystallographic analyses of DNA-bound Co(III)•bleomycin (47) indicated that bithiazole intercalation at some sites can be uncoupled from stable metal binding domain-minor groove association. For comparison to the above, 5'-GTAC, a site

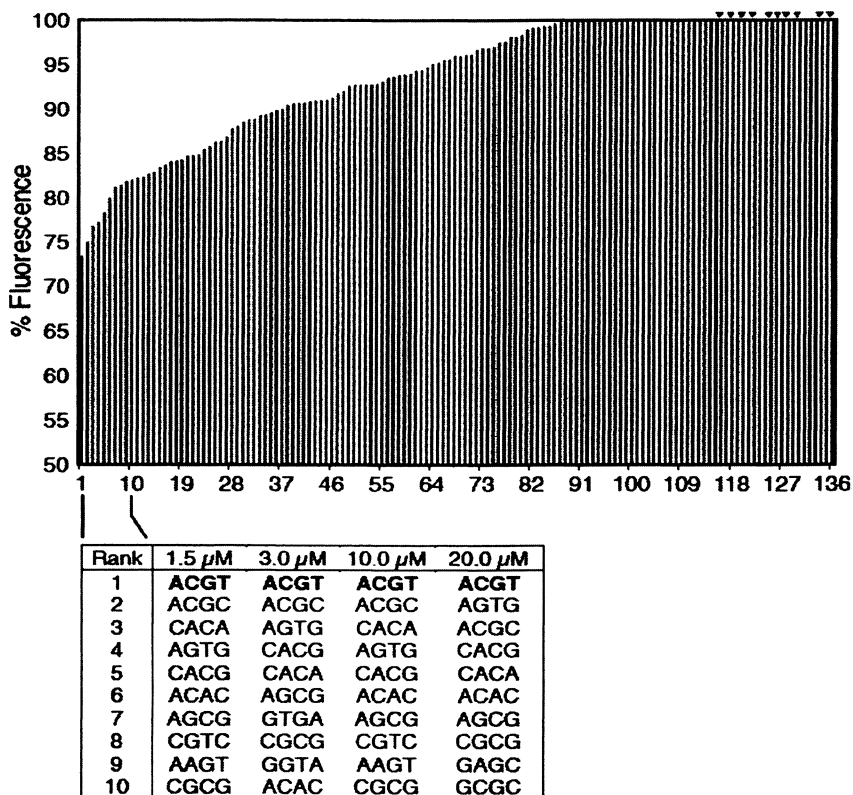


Figure 2. FID analysis of Co(III)•bleomycin "green". Results plotted in the histogram are from the analysis of 10  $\mu\text{M}$  drug; small triangles above the histogram indicate the position of A/T-only sequences in the rank-order.

with two oppositely-oriented 5'-GT sites, failed to yield a well-defined titration end-point (40) and simple tandemly-arranged sites such as 5'-GTGT did not confer an advantage to bleomycin binding beyond that of an isolated 5'-GT site.

The FID analysis described above succeeded in correlating the sites of preferred DNA binding by Co(III)•bleomycin "green" to the established DNA cleavage selectivity of Fe(II)•bleomycin + O<sub>2</sub>. These results support the notion that Co(III)•bleomycin can act as a structural model of cleavage-active Fe•bleomycin and that sites of drug-directed DNA cleavage reflect sites of DNA-drug recognition and binding. Interestingly, this study also revealed an interesting metallobleomycin preference for tetranucleotide sites with the potential to bind two drug equivalents simultaneously, albeit through an as yet uncharacterized means.

*Fe(III)•Bleomycin*

Employing the conditions used in the analysis of Co(III)•bleomycin (40), two sources of Fe(III)•bleomycin were analyzed by FID. The first source was generated through the direct admixture of equimolar amounts of Fe<sup>3+</sup> with bleomycin, while the other was generated through aerobic admixture of equimolar amounts of bleomycin and Fe<sup>2+</sup> (in the absence of DNA) followed by incubation for a time period sufficient to allow complete drug deactivation.

FID analyses of directly-formed Fe(III)•bleomycin (Figure 3) and Fe(III)•bleomycin formed via Fe(II)•bleomycin + O<sub>2</sub> (not shown) yielded results that were essentially identical to one another and very similar to those obtained for Co(III)•bleomycin (Figure 2); the FID plot curvatures were all quite similar as well as the final positioning of the A/T-only cassettes at the lowest end of each rank order. An examination of the rank-ordered hairpin cassettes indicated that across the Fe(III)•bleomycin concentrations studied (results in parentheses are for the product of Fe(II)•bleomycin + O<sub>2</sub>), among the top ranked 15 DNA hairpins, 13 (11) contained at least two 5'-GT or 5'-GC sites (when cassette and flanking sequences were considered). Overall, the top 50% of the ranked 136 hairpins *all* contained at least one 5'-GT or 5'-GC site while, in contrast, only 16 (19) of the lowest-ranked 50% contained a single 5'-GT or 5'-GC site. These results compared favorably to those obtained with Co(III)•bleomycin. Perhaps the strongest indicator of the similarities observed between these three metallobleomycin sources was the consistent presence of the top-ranked 5'-ACGT site. Similarly, 5'-ACGC was rank-ordered second or third at most concentrations. Thus, these three metallobleomycins consistently selected three hairpin sequences in common out of their top five ranked sequences. Underscoring this degree of similarity, minor shuffling of rank order is normally expected in the FID assay due to the often small differences between the absolute amounts of %*F* change that can, at times, be within the range of the experimental error of the assay (18, 19).

Overall, FID rank-order of the *binding* selectivity of Fe(III)•bleomycin correlated to known drug cleavage sites. These findings clearly support the notion that sites of drug induced cleavage reflect sites of specific [vs. non-specific (48)] DNA-drug binding.

*Cu(II)•Bleomycin*

Cu(II)•bleomycin was also analyzed using the FID assay (Figure 4). In stark contrast to the others, this metallobleomycin failed to produce an appreciable decrease in overall *F* and did not exhibit FID plot curvature. Notably also, Cu(II)•bleomycin did not discriminate against binding to the A/T-only cassettes; the A/T-only cassettes were distributed throughout the oligonucleotide rank order in contrast to the other metallobleomycins studied. These results indicated that Cu(II)•bleomycin did not bind to these hairpin substrates appreciably under



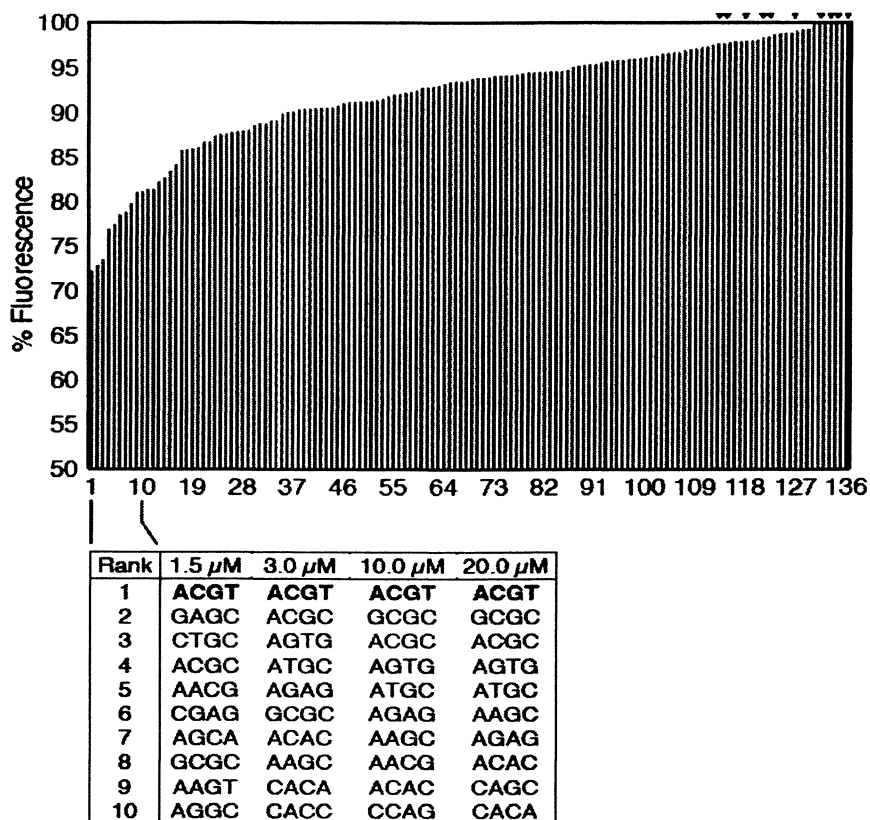


Figure 3. FID analysis of  $\text{Fe(III)}\cdot\text{bleomycin}$ . Results plotted in the histogram are from the analysis of  $10\ \mu\text{M}$  drug; small triangles above the histogram indicate the position of A/T-only sequences in the rank-order.

the FID assay conditions used for the other metallobleomycins; the lack of FID response observed is in accord with the relative DNA binding affinities of  $\text{Cu(II)}\cdot$  vs.  $\text{Co(III)}\cdot$  and  $\text{Fe(III)}\cdot\text{bleomycin}$ : a direct comparison (48) of oligonucleotide binding indicated that  $\text{Cu(II)}\cdot\text{bleomycin}$  bound 50- and 4-fold less well to a preferred 5'-GC site in comparison to  $\text{Co(III)}\cdot$  and  $\text{Fe(III)}\cdot\text{bleomycin}$ , respectively;  $\text{Cu(II)}\cdot\text{bleomycin}$  was also found to these oligonucleotides in fast exchange on the NMR time scale while the other metallobleomycins bound in slow exchange. Thus, in the presence of a competitive binding agent such as ethidium bromide it is not surprising that  $\text{Cu(II)}\cdot\text{bleomycin}$  binding cannot be detected in the FID assay. This analysis thus serves to illustrate the FID assay response to a weaker binding agent and also

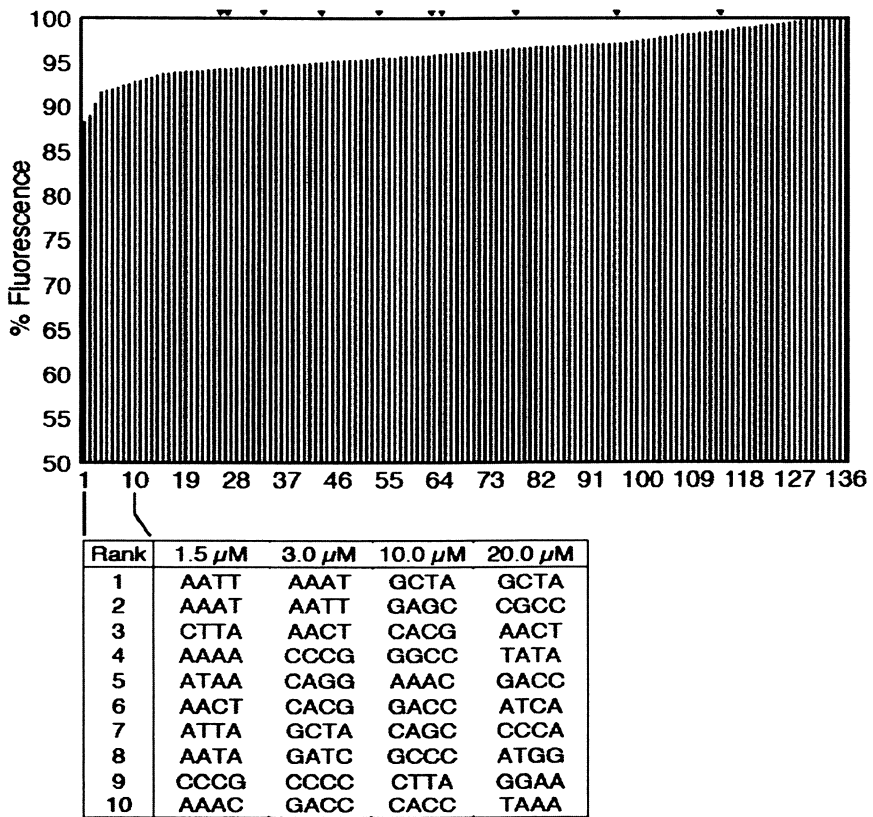


Figure 4. FID analysis of  $\text{Cu(II)}\cdot\text{bleomycin}$ . Results plotted in the histogram are from the analysis of  $10\ \mu\text{M}$  drug concentrations; small triangles above the histogram indicate the position of A/T-only sequences in the rank-order.

underscores the inherent differences in the DNA binding of various metallobleomycins in spite of their ability to induce selective DNA cleavage (39).

## Summary

The preceding FID analysis of several metallobleomycins verified that the DNA binding site-selectivities of Co(III)• and Fe(III)•bleomycin parallel each other quite closely and also correlate to the established cleavage site-selectivity of Fe(II)•bleomycin + O<sub>2</sub>. Notably different, Cu(II)•bleomycin did not exhibit an FID assay response under conditions used for the other metallobleomycins. This outcome was likely due to the weaker DNA binding and fast exchange of Cu(II)•bleomycin relative to other metallobleomycins (48) and was no doubt exacerbated by the presence of a competitive binding fluorophore. In spite of this limitation, the FID assay is likely to contribute to the study of additional metal complex-DNA interactions in the future.

## Host-Guest DNA Crystallization

Host-Guest DNA crystallization differs from other methods of DNA crystallization in that a DNA oligonucleotide *guest* is co-crystallized with a protein *host*, the N-terminal fragment of the Moloney murine leukemia virus reverse transcriptase (MMLV RT) (21-26, 47). In RT fragment-DNA host-guest structures, a conserved site in the RT protein binds in a non-sequence specific manner to the three terminal base pairs of each end of a DNA duplex oligonucleotide (Figure 5). The resulting structures of DNA have only minor disruptions in regions of contact with the protein and, importantly, the DNA is free of backbone-groove contacts commonly observed in oligonucleotide-only crystal lattices; backbone-groove contacts are often problematic and can block binding by low molecular weight agents when attempting to co-crystallize them with DNA through conventional means. Indeed, the Host-Guest system provides an approximately 10 Å thick solvent shell surrounding the surface of the central oligonucleotide allowing adequate room to accommodate a wide range of bulky DNA binding ligands in a native-like, aqueous environment.

For the DNA-ligand interactions studied so far (21, 26, 47), H-G crystallizations were pursued using 16-base pair DNA duplexes in which one protein molecule and one half of the 16-base pair duplex make up the unique repeating unit of the crystallographic lattice (Figure 5). Therefore, the 16-bp oligonucleotides used are self-complementary and contain two identical analyte binding sites. Ligands can be introduced by pre-complexation with the DNA prior to crystallization or by soaking into preformed RT-DNA crystals.

Further advantages of the host-guest method include the ability to: (1) obtain diffraction quality crystals of the desired complex overnight by employing microseeding techniques; (2) measure and process high resolution diffraction

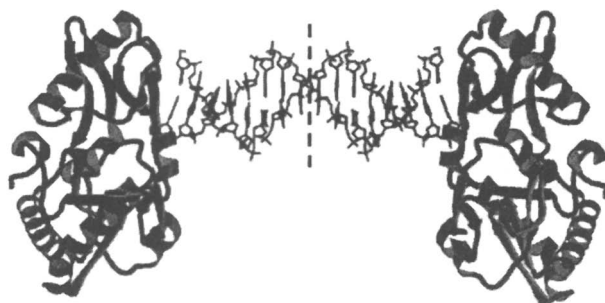


Figure 5. Unique repeating units of the H-G crystal lattice.

data (1.8-2.0 Å) using a home X-ray source and *R*-axis *IV* ++ image plate detector in one day; and (3) solve the phase problem by molecular replacement using the N-terminal fragment as the search model to obtain an unbiased electron density map for the DNA with bound ligand within an hour of completing the data collection. Thus, it is possible to ascertain whether a desired compound is bound in approximately two days. Traditional methods of obtaining crystals of DNA-ligand complexes routinely require weeks to months to obtain diffraction quality crystals. In addition to the above, preparation of large quantities of highly purified N-terminal RT fragment is straightforward as is the purification of the oligonucleotides—upon reagent preparation, the method is quite rapid. To date, the H-G crystallization strategy has been used to examine netropsin (21) and a cationic minor groove binding benzimidazole-diamidine (26, 41). In the latter case, we established that information obtained through FID analysis can be used to guide the design of oligonucleotides for H-G crystallization. As described next, the range of H-G crystallization applications now includes DNA-targeted metal complexes (47).

### Co(III)•Bleomycin B<sub>2</sub> “Green”

Our strategy for the H-G crystallization of Co(III)•bleomycin involved the design of 16-base pair oligonucleotides that placed isolated 5'-GT sites at various positions relative to the termini of the oligonucleotide (47). Despite being a top-ranked binding site, the 5'-ACGT site revealed by FID was not deemed compatible with crystallography due to the likely disordering of at least one of the two bound equivalents of drug. Thus, using the self-complementary deoxyoligonucleotide shown below containing two symmetry-related 5'-GT Co(III)•bleomycin binding sites:

ATTAGTTATAACTAAT  
TAATCAATATTGATTA

the desired drug-DNA complexes were obtained through the soaking of Co(III)•bleomycin into preformed RT-1 H-G crystals. Noteworthy, drug binding was observed to correlate with a change in the unit cell dimensions of the RT-1 crystals. What follows is an overview of the structure recently published and described in detail elsewhere (47).

In the structure obtained using RT-1 H-G crystals, Co(III)•bleomycin was found to occupy a 5-bp site involving residues G<sub>5</sub> through T<sub>7</sub> on one strand and A<sub>10</sub> through A<sub>14</sub> on the complementary strand (Figure 6). Ordered electron density was observed for the entire drug molecule (except for the propionamide and the peroxide ligand) with the drug bithiazole moiety clearly intercalated between residues A<sub>10</sub> and A<sub>11</sub>. To accommodate drug intercalation into the helical stack, both A<sub>8</sub> residues flanking the DNA dyad extrude. Consequently, a T<sub>9</sub>-T<sub>9</sub>' base pair forms that does not impact drug-DNA binding; indeed, the drug-bound and unbound DNA structures were reported to be quite similar.

Bound to **1**, the intercalated bithiazole is positioned with its C-terminal thiazole ring sandwiched between A<sub>10</sub> and A<sub>11</sub> while the N-terminal thiazole occupies a more central position between the base pairs involving these same residues; the helix is unwound by 8.9°. In this position, the guanidinium tail is threaded through to the major groove and forms a water-mediated H-bond to the N7 of A<sub>10</sub>. Connected to the bithiazole, the Thr-methylvalerate linker is well ordered and makes two specific H-bonds to the sugar-phosphate backbone; these interactions include the methyl valerate hydroxyl group to the O3' of T<sub>6</sub> and the bithiazole NH (amide link to Thr) to the O4' of T<sub>7</sub>.

With the C-terminal domains of bleomycin positioned as described above, the N-terminal metal binding domain resides in the minor groove and binds Co(III) in a distorted square-planar geometry via the ligands and screw sense depicted in Figure 1. Further, the face of the metal binding domain occupied by the β-aminoalanine axial ligand was observed to partially stack against the disaccharide moiety in a side-by-side arrangement in the minor groove; this interaction likely serves to stabilize the metal binding domain-minor groove interaction (49). The interface between these two domains is supported by three intramolecular H-bonds: the mannose C4-OH to the pyrimidine carbonyl and the β-aminoalanine NH<sub>2</sub> to both the β-O of His and to the mannose carbamoyl carbonyl. Noteworthy, this intramolecular structural arrangement was also observed in the crystal structure of Cu(II)•bleomycin bound to a bleomycin resistance protein (50). In this compact, side-by-side arrangement, the metal binding-disaccharide domain forms specific intermolecular H-bonds to the floor and walls of the minor groove. These include H-bonds from the drug pyrimidine moiety endocyclic N and C4-NH<sub>2</sub> to the N2 and N3 of G<sub>5</sub>, respectively; the pyrimidine C4-NH<sub>2</sub> to the O4' of G<sub>5</sub>; and the mannose C2-OH and carbamoyl amine to the O3' of A<sub>14</sub> and the O2 of T<sub>13</sub>, respectively. No specific interactions between the drug and the T nucleobase of the 5'-GT target site were observed: this suggests that the structure and flexibility of a 5'-G-Py step (in contrast to a 5'-G-Pu step) likely facilitates bleomycin-minor groove intercalation leading to overall 5'-GT(C) targeting. Noteworthy, bound as described above, modeling of

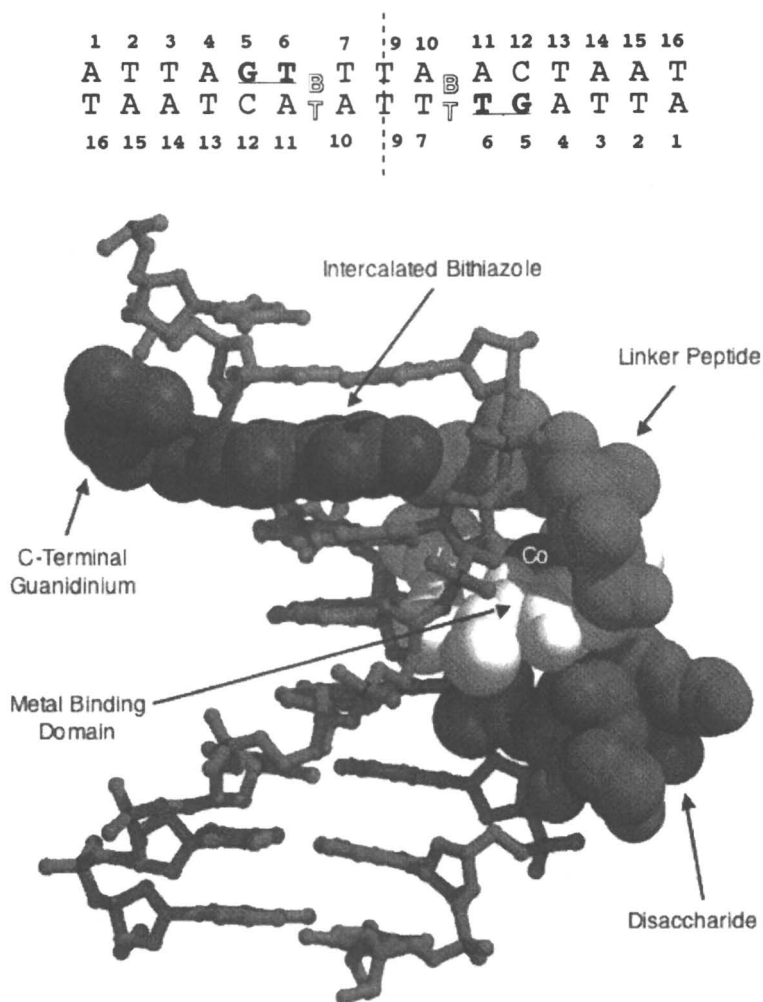
a metal-bound hydroperoxide places the distal O 2.5 Å from the target C4'-H of T<sub>6</sub>; this likely parallels the interaction of HOO-Fe(III)•bleomycin at a target site prior to C4'-H abstraction and concerted O-O bond cleavage (51).

## Summary

Initial structural information on the interactions of Co(III)•bleomycin with DNA was derived from 2D NMR data (45, 52-54). The recent X-ray crystallographic study (47) reviewed here provides the first 3-D structures with new insights on the detailed interactions of Co(III)•bleomycin B<sub>2</sub> with DNA, including the mode of bithiazole intercalation and interactions of the metal binding domain in the minor groove, as well as structural changes that occur within the DNA itself upon drug binding. H-G crystallographic studies in progress comparing the structures of DNA-bound Co(III)•bleomycin A<sub>2</sub> with those of B<sub>2</sub> are expected to provide further insights into the intricate nature of this DNA-ligand interaction. Overall, the example provided by Co(III)•bleomycin serves to illustrate the applicability of H-G crystallization to the investigation of DNA-targeted metal complexes.

## Conclusions

Appreciation of cellular DNA as a target for low molecular weight metal complexes and compounds continues to evolve along with our ability to generate and discover such agents through expedited means including combinatorial chemistry and high-throughput screening. Thus, given the wherewithal to increase the pace of DNA binding metal complex discovery and synthesis, it is important to be able to investigate the binding activities and structures of DNA-targeted complexes at a similar rate. Through the use of several metallobleomycins, two new approaches that facilitate the study of DNA binding metal complexes have been described: the FID assay and H-G crystallization. With the FID assay, a comprehensive evaluation of the DNA binding site-selectivity of a metal complex can be achieved in a very short period of time. Importantly, this technique, unlike indirect strategies that, for example, rely on acts of complex-induced biopolymer strand scission, reports only on the act of metal complex-DNA binding. For Co(III)• and Fe(III)•bleomycin, the FID assay verified the binding site preferences of these metallobleomycins for 5'-GT(C) sites in comparison to one another and also to the known cleavage site-selectivity of Fe(II)•bleomycin + O<sub>2</sub>; results obtained in this assay also revealed a drug preference for tetranucleotide cassettes that contain two inverted 5'-GT(C) dinucleotide sites, *e.g.*, 5'-ACGT and 5'-ACGC. In addition, as demonstrated using Cu(II)•bleomycin, FID assay results are obtained only when the



*Figure 6. Schematic of 1 in the presence Co(III)•bleomycin B<sub>2</sub> “green” and the nucleotide numbering referred to in the text; BT, intercalated bithiazole; dashed line, oligonucleotide dyad axis (top). H-G crystal structure of DNA-bound Co(III)•bleomycin B<sub>2</sub> “green” (bottom).*

complex/analyte binds with a reasonable affinity to a target nucleic acid and competes favorably with the reporting fluorophore used in the assay.

Along with knowledge of the DNA binding site-selectivity of a metal complex, information pertaining to its DNA-bound three dimensional structure is often essential for further progress or the synthesis of next generation complexes. Towards this end, H-G crystallography can expedite the crystallization and analysis of a site-selectively bound metal complex by providing a protein-oligonucleotide “framework” within a crystal lattice that eliminates problems associated with backbone-groove contacts that often makes the oligonucleotide co-crystallization of bulky, complex ligands problematic; as discussed, the H-G crystallization system can be microseeded and also expedites the analysis of ligand-bound crystals once they are obtained. Using Co(III)•bleomycin “green” we demonstrate that H-G crystallization is compatible with the study of metal complexes leading to the generation and analysis of the first duplex DNA-bound crystal structures of a metallobleomycin. These crystallizations revealed details concerning the structure of the DNA-bound drug and, equally as important, the impact of drug binding on the DNA substrate structure.

Overall, the FID assay and the H-G crystallization strategy, either singly or in combination (26), can assist in circumventing the usual “slow” steps associated with the analysis of DNA-ligand binding. Elimination of these historically slow steps would no doubt accelerate the discovery process and perhaps lead to the increased development and application of metal complexes towards biomedically-significant problems.

## Acknowledgements

This work was supported by the National Institutes of Health (E.C.L. and M.M.G.) and a fellowship from the American Cancer Society (K.D.G.). We also wish to thank Marianne Cuff, Steve Ginell, and Andrzej Joachimiak from the Structural Biology Center Collaborative Access Team at the Advanced Photon Source. Results shown in this report are derived from work performed at Argonne National Laboratory, Structural Biology Center at the Advanced Photon Source. Argonne is operated by UChicago Argonne, LLC, for the U.S. Department of Energy, Office of Biological and Environmental Research under contract DE-AC02-06CH11357.

## References

1. *DNA and RNA Binders: From Small Molecules to Drugs*; Demeunynck, M.; Bailly, C.; Wilson, W. D., Eds.; Wiley-VCH: Weinheim, 2003.
2. Zeglis, B. M.; Pierre, V. C.; Barton, J. K. *Chem. Comm.* **2007**, *44*, 4565-4579.
3. Thurston, D. E. *Br. J. Cancer* **1999**, *80* (Suppl. 1), 65-85.



4. Browne, M. J.; Thurlby, P. L. *Genomes, Molecular Biology and Drug Discovery*; Academic: London, 1996.
5. Huang, X.; Pieczko, M. E.; Long, E. C. *Biochemistry* **1999**, *38*, 2160-2166.
6. Leitheiser, C. J.; Smith, K. L.; Rishel, M. J.; Hashimoto, S.; Konishi, K.; Thomas, C. J.; Li, C.; McCormick, M. M.; Hecht, S. M. *J. Am. Chem. Soc.* **2003**, *125*, 8218-8227.
7. Hampshire, A. J.; Rusling, D. A.; Broughton-Head, V. J.; Fox, K. R. *Methods* **2007**, *42*, 128-140.
8. Pyle, A. M.; Barton, J. K. *Prog. Inorg. Chem.* **1990**, *38*, 414-475.
9. Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291-309.
10. Wang, A. H.-J.; Robinson, H. In *Nucleic Acid Targeted Drug Design*; Probst, C. J.; Perun, T. J., Eds.; Dekker: New York, NY, 1992; pp 17-64.
11. Stubbe, J.; Kozarich, J. W. *Chem. Rev.* **1987**, *87*, 1107-1136.
12. Burger, R. M. *Chem. Rev.* **1998**, *98*, 1153-1169.
13. Boger, D. L.; Cai, H. *Angew. Chem. Int. Ed.* **1999**, *38*, 448-476.
14. Claussen, C. A.; Long, E. C. *Chem. Rev.* **1999**, *99*, 2797-2816.
15. Hecht, S. M. *J. Nat. Prod.* **2000**, *63*, 158-168.
16. Chen, J.; Stubbe, J. *Nat. Rev. Cancer* **2005**, *5*, 102-112.
17. Hecht, S. M. In *Anticancer Agents from Natural Products*; Cragg, G. M.; Kingston, D. G. I.; Newman, D. J., Eds.; CRC: Boca Raton, FL, 2005; pp 357-381.
18. Boger, D. L.; Fink, B. E.; Brunette, S. R.; Tse, W. C.; Hedrick, M. P. *J. Am. Chem. Soc.* **2001**, *123*, 5878-5891.
19. Tse, W. C.; Boger, D. L. *Acc. Chem. Res.* **2004**, *37*, 61-69.
20. *Drug-DNA Interaction Protocols*; Fox, K. R., Ed.; Methods in Molecular Biology; Humana: Totowa, NJ, 1997; Vol. 90.
21. Goodwin, K. D.; Long, E. C.; Georgiadis, M. M. *Nucleic Acids. Res.* **2005**, *33*, 4106-4116.
22. Cote, M. L.; Yohannan, S. J.; Georgiadis, M. M. *Acta Cryst. D Biol. Cryst.* **2000**, *56*, 1120-1131,
23. Najmudin, S.; Cote, M. L.; Sun, D.; Yohannan, S.; Montano, S. P.; Gu, J.; Georgiadis, M. M. *J. Mol. Biol.* **2000**, *296*, 613-632,
24. Cote, M. L.; Georgiadis, M. M. *Acta Cryst. D Biol. Cryst.* **2001**, *57*, 1238-1250.
25. Cote, M. L.; Pflomm, M.; Georgiadis, M. M. *J. Mol. Biol.* **2003**, *330*, 57-74.
26. Goodwin, K. D.; Lewis, M. A.; Tanious, F. A.; Tidwell, R. R.; Wilson, W. D.; Georgiadis, M. M.; Long, E. C. *J. Am. Chem. Soc.* **2006**, *128*, 7846-7854.
27. Umezawa, H.; Maeda, K.; Takeuchi, T.; Okami, Y. *J. Antibiot. (Tokyo)* **1966**, *19*, 200-209.
28. Connors, J. M. *J. Clin. Oncol.* **2005**, *23*, 6400-6408.
29. Einhorn, L. H. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4592-4595.
30. Nagai, K.; Yamaki, H.; Suzuki, H.; Tanaka, N.; Umezawa, H. *Biochim. Biophys. Acta* **1969**, *179*, 165-171.
31. Povirk, L. F.; Wübker, W.; Köhnlein, W.; Hutchinson, F. *Nucleic Acids Res.* **1977**, *4*, 3573-3579.

32. Carter, B. J.; de Vroom, E.; Long, E. C.; van der Marel, G. A.; van Boom, J. H.; Hecht, S. M. *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 9373-9377.
33. Abraham, A. T.; Lin, J. J.; Newton, D. L.; Rybak, S.; Hecht, S. M. *Chem. Biol.* **2003**, *10*, 45-52.
34. Sausville, E. A.; Peisach, J.; Horwitz, S. B. *Biochem. Biophys. Res. Commun.* **1976**, *73*, 814-822.
35. Wu, J. C.; Kozarich, J. W.; Stubbe, J. *Biochemistry* **1985**, *24*, 7562-7568.
36. D'Andrea, A. D.; Haseltine, W. A. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 3608-3612.
37. Takeshita, M.; Grollman, A. P.; Ohtsubo, E.; Ohtsubo, H. *Proc. Natl. Acad. Sci. USA* **1978**, *75*, 5983-5987.
38. Kano, M.; Tomita, S.; Ishida, S.; Murakami, A.; Okada, H. *Chemotherapy (Tokyo)* **1973**, *21*, 1305-1310.
39. Ehrenfeld, G. M.; Shipley, J. B.; Heimbrook, D. C.; Sugiyama, H.; Long, E. C.; van Boom, J. H.; van der Marel, G. A.; Oppenheimer, N. J.; Hecht, S. M. *Biochemistry* **1987**, *26*, 931-942.
40. Lewis, M. A.; Long, E. C. *Bioorg. Med. Chem.* **2006**, *14*, 3481-3490.
41. Tanious, F. A.; Laine, W.; Peixoto, P.; Bailly, C.; Goodwin, K. D.; Lewis, M. A.; Long, E. C.; Georgiadis, M. M.; Tidwell, R. R.; Wilson, W. D. *Biochemistry* **2007**, *46*, 6944-6956.
42. Spillane, C. B.; Smith, J. A.; Morgan, J. L.; Keene, F. R. *J. Biol. Inorg. Chem.* **2007**, *12*, 819-824.
43. McLean, M. J.; Dar, A.; Waring, M. J. *J. Mol. Rec.* **1989**, *1*, 184-192.
44. Chang, C. H.; Dallas, J. L.; Meares, C. F. *Biochem. Biophys. Res. Commun.* **1983**, *110*, 959-966.
45. Vanderwall, D. E.; Lui, S. M.; Wu, W.; Turner, C. J.; Kozarich, J. W.; Stubbe, J. *Chem. Biol.* **1997**, *4*, 373-387.
46. Xia, C.; Forsterling, F. H.; Petering, D. H. *Biochemistry* **2003**, *42*, 6559-6564.
47. Goodwin, K. D.; Lewis, M. A.; Long, E. C.; Georgiadis, M. M. *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 5052-5056.
48. Li, W.; Zhao, C.; Xia, C.; Antholine, W. E.; Petering, D. H. *Biochemistry* **2001**, *40*, 7559-7568.
49. Iiyama, T.; Chikira, M.; Oyoshi, T.; Sugiyama, H. *J. Biol. Inorg. Chem.* **2003**, *8*, 135-140.
50. Sugiyama, M.; Kumagai, T.; Hayashida, M.; Maruyama, M.; Matoba, Y. *J. Biol. Chem.* **2002**, *277*, 2311-2320.
51. Decker, A.; Chow, M. S.; Kemsley, J. N.; Lehnert, N.; Solomon, E. I. *J. Am. Chem. Soc.* **2006**, *128*, 4719-4733.
52. Wu, W.; Vanderwall, D. E.; Turner, C. J.; Kozarich, J. W.; Stubbe, J. *J. Am. Chem. Soc.* **1996**, *118*, 1281-1294.
53. Zhao, C.; Xia, M.; Q.; Forsterling, H.; DeRose, E.; Antholine, W. E.; Subczynski, W. K.; Petering, D. H. *J. Inorg. Biochem.* **2002**, *91*, 259-268.
54. Hoehn, S. T.; Junker, H. D. Bunt, R. C.; Turner, C. J. Stubbe, J. *Biochemistry* **2001**, *40*, 5894-5905.

## Chapter 6

# Zn(II) Homeostasis in *E. coli*

**Thusitha Gunasekera, J. Allen Easton, Stacy A. Sugerbaker,  
Lindsey Klingbeil, and Michael W. Crowder\***

**Department of Chemistry and Biochemistry,  
Miami University, Oxford, OH 45056**

The maintenance of optimal zinc levels in bacteria is required for cell function and survival. In an effort to probe the cellular response to Zn(II) deficient and excess stresses and to identify any potential cytoplasmic Zn(II)-metallochaperones, DNA microarrays were used. Although many Zn(II)-responsive genes were identified, no plausible candidates for cytoplasmic Zn(II)-metallochaperones were identified. This result warrants a new model to explain intracellular Zn(II) transport in bacteria, and a novel hypothesis that involves a role of the ribosomes in Zn(II) storage and transport is offered.

Zn(II) is an essential trace element of many proteins and is required for life in all organisms (1). In fact, Zn(II) serves as a catalytic cofactor in members of all six major functional classes of enzymes and a structural cofactor for many other proteins, including zinc finger-containing proteins and transcription factors (2). Nonetheless, high levels of Zn(II) are toxic to cells, presumably due to binding to adventitious metal binding sites and to sites that would normally be occupied by other metal ions. Therefore, it is essential that all cells maintain homeostatic mechanisms that regulate Zn(II) absorption, intracellular uptake, distribution, and excretion (3-5). There are several excellent reviews on Zn(II)

homeostasis in eukaryotic cells (6-8). However, relatively little is known about Zn(II) utilization in bacterial cells.

O'Halloran has reported that *E. coli* cells have a quota of *ca.*  $2 \times 10^5$  zinc ions per cell (millimolar concentrations) (9). However, essentially all of this Zn(II) is bound by 100-200 Zn(II) metalloproteins (10), and there is not a persistent pool of "free" Zn(II) inside an *E. coli* cell at equilibrium (9). By using a cell assay and Zn(II) transcription regulators Zur and ZntR, O'Halloran and coworkers estimated that there is much less than one free Zn(II) ion (free or aquated Zn(II) complexes) in an *E. coli* cell (9).

### **How is Zn(II) Transported Into and Out of the *E. coli* Cytoplasm?**

Since *E. coli* is Gram-negative and has two cell membranes, there must be at least two steps in importing Zn(II) into the cytoplasm: (1) transport across the outer membrane into the periplasm and (2) transport across the inner membrane into the cytoplasm (Figure 1). It is not known how Zn(II) enters the periplasm or the chemical form of the Zn(II) (aquated, peptide-bound, siderophore-bound, *etc.*) that is transported across the outer membrane. It is possible that Zn(II) enters the periplasm via porins (11), which can transport low molecular weight species; however, this possibility has not been demonstrated. The transport of Zn(II) across the inner membrane is better understood. In low Zn(II) conditions, Zn(II) is bound by ZnuA, a periplasmic protein, that delivers the Zn(II) to ZnuB, which uses ATP-hydrolysis by ZnuC to transport Zn(II) across the inner membrane (12). Expression of ZnuABC is regulated by Zur, which is a Fur-like transcription regulator that represses transcription of ZnuABC in the presence of Zn(II) (13). ZupT is a Zn(II) influx pump that is a member of the ZIP family of proteins; it has recently been shown to be a nonspecific metal transporter (14, 15). There are also several other potential broad-spectrum metal ion importers, such as the Pit inorganic phosphate transport system, that may transport Zn(II) into *E. coli* cells (16). In addition, it has been proposed that the Zur-regulated, periplasmic protein YodA may have a BC partner (like ZnuABC described above) and can be involved with Zn(II) import (17, 18); however, the identity of the BC partners is not known.

For Zn(II) export, there are at least 3 membrane transporters that have been identified in *E. coli*. ZntA, the high specificity Zn(II) efflux pump, is transcriptionally-regulated by ZntR, a MerR homolog that binds Zn(II) (19-21). ZitB is a member of the cation diffusion facilitator (CDF) family that is specifically induced by Zn(II) (22); however, it is not clear if this transporter is specific only for Zn(II). A second CDF exporter, YiiP, is also present in *E. coli*, and a recent crystal structure on this exporter revealed details about how Zn(II) binds (23). Although work still remains to be performed to fully understand how

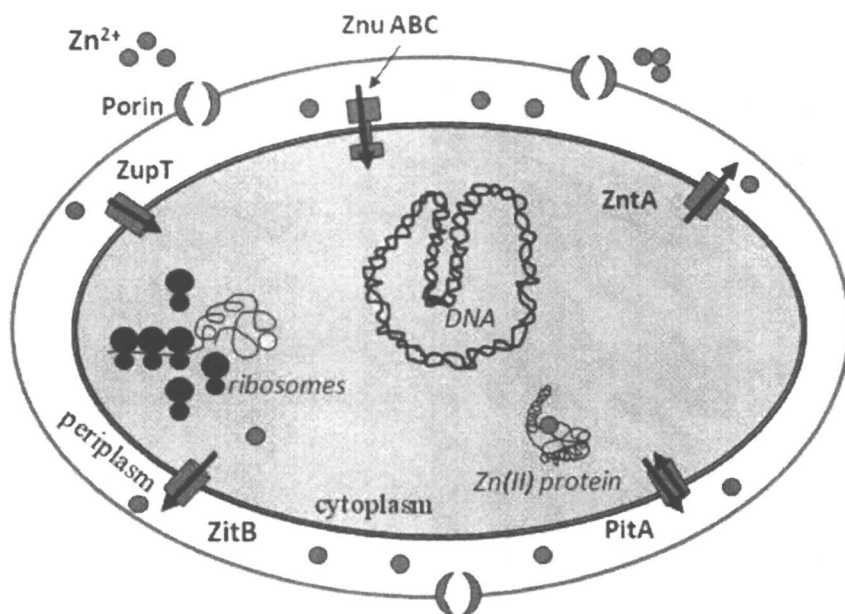


Figure 1. Zn(II) homeostasis in *E. coli*.

Zn(II) is transported across the cell membranes, there is a great deal known about this part of Zn(II) homeostasis in *E. coli*. However, it is not known how Zn(II) is transported from the intracellular side of the importers to cytoplasmic Zn(II) metalloproteins or how intracellular Zn(II) is delivered to the exporters.

### How is Intracellular Zn(II) Transported/Mobilized in Bacteria?

Intracellular transport of metal ions has been a very well studied research area in the last few years. The most often-cited case of this transport involves copper transport in yeast by the copper metallochaperones (24, 25). Cu(I) enters the cytoplasm of yeast via copper transport receptors, and Cu(I) is bound by one of three transport proteins: Lys7, Atx1, or Cox17. Lys7 delivers copper to CuZn superoxide dismutase, Atx1 delivers copper to ccc2 that activates an Fe(II) uptake system, and Cox17 delivers copper to the mitochondria for the ultimate uptake into cytochrome c oxidase. A similar copper transport system has been reported in humans (26), and there may be a system in bacteria as well (27). Metal ion transport systems are known for iron, nickel, and manganese (24, 25, 28). However, no cytoplasmic Zn(II) transporters have been identified in

bacteria. There is some speculation that metallothionein may play a role in Zn(II) transport in eukaryotes (29-31); however, there are no metallothioneins yet known in most bacteria (32). To explain Zn(II) distribution in bacteria, three hypotheses have been offered: (1) C<sup>+</sup>/C<sup>-</sup> system, (2) burst model, and (3) Zn(II) metallochaperones.

In *B. subtilis*, there are at least two paralogous proteins (called C<sup>+</sup> and C<sup>-</sup>) that appear to be involved in Zn(II) homeostasis (33-35). Similar paralogs are found in several different bacterial strains (36). Interestingly, these proteins are found on the surface of the ribosome in some cases and in the interior of the ribosome in others. The C<sup>+</sup> paralogs contain a Cys motif that is thought to bind Zn(II), and in time of sufficient Zn(II) levels in the cell, the C<sup>+</sup> motif-containing proteins bind to the ribosome. The C<sup>-</sup> paralogs lack the putative Zn(II) binding Cys motif, and the C<sup>-</sup> paralogs are thought to replace the C<sup>+</sup> paralogs during time of Zn(II) deficiency in *B. subtilis*, thereby releasing up to two equivalents of Zn(II) per ribosome. Unlike *B. subtilis* which contains two C<sup>+</sup>/C<sup>-</sup> paralogs, *E. coli* only contains one C<sup>+</sup>/C<sup>-</sup> paralog (L31/YkgM) (36). YkgM is transcriptionally-regulated by Zur (35). It is not clear how the unused Zn(II) from the C<sup>+</sup> paralog is distributed to the correct Zn(II)-metalloproteins during Zn(II) deficient conditions.

According to the "burst" model (37), the cell presumably moves away from equilibrium, and then bursts or pulses of "free" Zn(II) (or as complexes with water or small compounds like glutathione as ligands) enter the cytoplasm from the extracellular medium, and from the periplasm (prokaryotes) or subcellular compartments (eukaryotes) via the transporters mentioned above. This scenario is reminiscent of Zn(II) release into the synapse from nerve cells (38). The Zn(II) ions would then distribute into the cytoplasm and bind to newly-made, apo-Zn(II)-metalloproteins, to the ribosomes, *etc.*, as well as to non-specific sites on proteins, amino acids, or nucleic acids. The cell apparently returns to equilibrium by the activities of the membrane-spanning Zn(II) transporters listed above.

The "pulse" mechanism for intracellular Zn(II) mobilization in eukaryotes presents some significant questions. For example, how does the released Zn(II) from the vacuoles find its way to the newly-made Zn(II)-metalloproteins and how does the unused and non-specifically bound Zn(II) get returned to the transporters in the vacuolar membrane? In prokaryotes, the lack of intracellular compartmentalization presents the additional problem of many more non-specific Zn(II)-binding sites in the cytoplasm of these cells. It is possible that when the cell needs Zn(II), the non-specifically bound Zn(II) could be mobilized (by shifts in pH, for example) to provide the cell with the needed Zn(II). This latter scenario is not likely since the pH required to release Zn(II) from nucleic acids, for example, would be so low that the cell would not survive. Moreover, previous studies have shown that there are no significant concentrations of "free" or aquated transition metal ions present in the cytoplasm of a cell at equilibrium.

Also, the “burst” mechanism for transition metal ion transport/distribution is unprecedented in the literature.

The third hypothesis for Zn(II) transport/homeostasis involves the presence of soluble, cytoplasmic Zn(II)-metallochaperones. In this model, Zn(II) is brought into the cytoplasm by Zn(II) transporters (ZnuABC, ZupT, PitA), binds to chaperone proteins/peptides, and thus distributes to the places in the cell where Zn(II) is needed. The total cytoplasmic concentration of “free” Zn(II) would remain at equilibrium values. This mechanism requires specific, intracellular metal ion transporters, such as the transporters (metallochaperones) that are known for Cu, Mn, and Fe delivery (24, 25, 28). While, no such cytoplasmic Zn(II)-metallochaperones have been identified in any organism, several groups have hypothesized their existence in eukaryotes and prokaryotes (3, 4, 9, 39). In this work, we conducted experiments in an effort to identify Zn(II)-metallochaperones in *E. coli*.

### Search for Zn(II)-Metallochaperones

The discoveries of Ni-, Fe-, and Cu-metallochaperones have resulted from genetic studies. In prokaryotes, Fe- and Ni-metallochaperones were encoded on the same operon as the protein that is metallated. For example, the gene for UreE is on the same operon as urease (40). In eukaryotes, the discovery of Cu-metallochaperones was serendipitous as yeast cells were being screened for phenotypes not directly associated with Cu transport. The first soluble Cu-metallochaperone reported was Atx1, which was discovered by Culotta and coworkers who were screening yeast cells for suppressors of superoxide dismutase (SOD) deficiency (41). Tzagoloff and coworkers, who were screening respiratory-defective mutants of yeast, identified a copper delivery pathway involving Cox17 that ultimately loads cytochrome c oxidase with copper (42). Culotta subsequently reported that a Cu transport pathway is responsible for loading superoxide dismutase with copper (43). The proteins identified with these genetic screens were used to conduct sequence homology searches, and homologs were found in many species (25, 26, 44-46).

We believe that using the genetic approaches that were used to identify other metallochaperones to discover any Zn(II) metallochaperones in *E. coli* will not be successful. In yeast/*E. coli* there are less than ten copper-containing proteins (47), and most of these proteins are involved with oxygen/reactive oxygen species chemistry. In contrast, there could be almost 200 Zn(II)-metalloproteins that have numerous catalytic and structural roles (48). The choice of a deficient phenotype that could be linked directly to Zn(II) would be very challenging. The choice of superoxide dismutase deficiency (CuZn enzyme) would probably not be successful in *E. coli* since there are two other

versions of SOD that do not require Zn(II) (49). We spent considerable effort examining the genes located near known Zn(II)-metalloproteins in an effort to identify some potential Zn(II)-metallochaperone candidates. However, no potential candidate was discovered with these searches. Therefore, we proposed to identify Zn(II)-responsive proteins/genes in *E. coli* by using DNA microarrays.

### Zn(II)-Responsive Genes in *E. coli* Cells

In order to identify Zn(II)-responsive genes, *E. coli* BW21135 cells were stressed with Zn(II) deficiency by using 10  $\mu$ M TPEN and with Zn(II) excess by using 5 mM ZnSO<sub>4</sub>. We attempted to stress the cells with Zn(II) deficiency without the use of TPEN; however, we found that the cells were not stressed when grown in minimal medium not containing the chelator. TPEN is often referred to as a “Zn(II)-specific chelator”; however, the literature shows that TPEN can coordinate Cu(II) and Fe(II) (50). We reasoned that any genes that are involved with Zn(II) import would be up-regulated in Zn(II) deficient conditions and down-regulated in Zn(II) excess conditions. Conversely, genes involved with Zn(II) export would be up-regulated in Zn(II) excess conditions and down-regulated in Zn(II) deficient conditions. The response of two Zn(II) transcription regulators, Zur and ZntR, would allow for us to test whether we were able to realize Zn(II) deficient and excess conditions.

We chose DNA microarrays, instead of proteomic studies, as the method to conduct global analyses of the stressed *E. coli* cells because this genomic technique allows for screening of all *ca.* 4,200 genes in *E. coli* simultaneously and could possibly identify peptides, which are typically too small to identify in most proteomic techniques. In addition, several other groups (51-54) had conducted DNA arrays at different conditions examining the effect of Zn(II) excess on the gene profile of *E. coli*, and the results from these other studies could serve as comparisons for our data.

The first step in conducting global analyses is determining how the cells will be stressed and how long the stress should be imposed. The latter part of this step is rarely done in DNA microarray studies, and this omission can often lead to missed genes that are regulated. In an effort to determine the optimum stress time for the cells, we used real-time (quantitative) PCR (RT-PCR) studies and evaluated the time-dependent transcription of all genes (*znuA*, *zntA*, *zitB*, and *zupT*) (12-15) previously implicated in Zn(II) trafficking in *E. coli* or regulated by Zur (*ykgM* and *yodA*). The response of zur-regulated genes (*znuA*, *ykgM*, *yodA*) to TPEN was large, showing a > 1,000-fold increase at 30 minutes of stress. Previous cDNA array studies were conducted at 5 minutes and 5 hours.



The transcription levels for the Zn(II)-related genes at these times are approximately equal but are still over 100-fold less than at 30 minutes. In contrast, the transcription levels of *zupT*, *zntA*, *zitB*, and control 16S RNA were not greatly affected by the presence of TPEN (<4-fold), suggesting that low levels of intracellular Zn(II) do not greatly affect these genes. The absence of a strong response of *zntA* to TPEN is consistent with previous studies on ZntR, which activates transcription of *zntA* when bound to Zn(II). This result is also consistent with the recent report of ZntR's resistance to proteolysis when bound to DNA and Zn(II); in low Zn(II) conditions, ZntR does not bind to its promoter site or to Zn(II) and is lysed (55).

A concentration of 5 mM zinc sulfate was used to stress *E. coli* cells with Zn(II) excess (Crowder, unpublished). No significant (if compared to the TPEN results) effect was observed on any of the genes tested. The results on the Zur-regulated transcripts (*znuA*, *ykgM*, and *yodA*) are not unexpected since Zn(II)-bound Zur represses transcription of these genes (56). However, the fold changes of *zntA* are surprisingly low, as compared to the changes observed in the Zur-regulated genes. Zn(II) binding by ZntR was expected to result in a significant increase in expression of *zntA* (57); however, our data and previous microarray studies demonstrate only small changes in the transcription of *zntA* under high Zn(II) conditions (54). To determine the optimal time to harvest *E. coli* cells stressed with Zn(II) excess for cDNA arrays, Yamamoto *et al.* used an S1 nuclease assay and found that the *zntA* gene was most highly transcribed after 5 minutes of stress (53). In our study, approximately 5-fold changes were seen in *zntA* after 5 minutes of stress but quickly diminished to near baseline levels after 30-60 minutes.

In an effort to ensure that the observed effect of TPEN on expression was due to Zn(II) and not iron, iron/zinc-rescue experiments were performed. Cultures of control cells (BW25113) and  $\Delta znuA$  knock-out mutant cells (obtained from National BioResource Project, National Institute of Genomics, Japan) (58) were grown and starved for Zn(II) using 10  $\mu$ M TPEN. After 30 hours of growth, there is little or no growth in the  $\Delta znuA$  mutant cells in the presence of TPEN, while control cells without TPEN as well as control cells with TPEN show growth of at least an OD<sub>600</sub> of 1.5. After 30 hours of growth, FeSO<sub>4</sub> was added to the TPEN-treated cells. If the growth response had been due to iron deficiency, the effect should have been reversed by the addition of FeSO<sub>4</sub>. This was not the case, but growth was rescued upon the later addition of ZnSO<sub>4</sub> at 36 hours. These data suggest it was Zn(II), not Fe(II), causing the decreased growth in the  $\Delta znuA$  mutant cell line in the presence of TPEN.

The purpose for the aforementioned real-time PCR experiments was to determine the best time to harvest cells for cDNA microarray analysis of the zinc-starved condition. We prepared total RNA samples from *E. coli* that were

stressed for 30 minutes with TPEN and completed the subsequent microarray studies. After filtering criteria were implemented (genes that were not present in both conditions, genes that were always expressed below 10% of the average array signal (150 for the chips with target intensity of 1500), and genes with  $p$  values  $> 0.05$  were excluded from further analysis), 114 genes were found to be up-regulated, and 146 genes were down-regulated. The top 5 up- and down-regulated genes are shown in Table I. In our previous microarray studies (54), we identified 55 up-regulated transcripts, and 56% of these genes were also up-regulated in this study reported here. Most of the genes found in both studies are Zur- or Fur-regulated. All five Zur-regulated genes *yodA*, *ykgM*, *znuA*, *znuB*, and *znuC* were significantly up-regulated in the cells stressed with TPEN for 30 minutes. In this work, we report two additional genes that are up-regulated by Zn(II)-starvation that are in the same gene clusters as *yodA* and *znuA* and could possibly be regulated by Zur. The first of these is *yodB*, an integral membrane protein that is a predicted cytochrome, is located directly downstream of *yodA* in the *E. coli* chromosome (59). The other Zur-regulated candidate, *yebA*, is located directly downstream of *znuA*, a known Zur-regulated protein (59). Neither of these genes have been studied in enough detail to know if the proteins that they encode could possibly be involved in Zn(II) homeostasis and thus will be a target of future research. Of the remaining up-regulated genes, 22% of them belong in the inorganic ion transport and metabolism COG (cluster of orthologous group) (60), 9% have unknown function, 17% do not belong to an identified COG, and the remaining genes are spread among the other COG's.

**Table I. Genes Up- and Down-Regulated by TPEN**

<i>Gene</i>	<i>Description</i>	<i>P value</i> <sup>a</sup>	<i>Fold change</i>
<i>ykgM</i>	Ribosome paralog	2.0	+205
<i>yodA</i>	Periplasmic, lipocalin-like protein	2.0	+168
<i>znuA</i>	Periplasmic Zn(II) transporter	2.0	+42
<i>mntH</i>	Putative Mn(II) transporter	2.0	+21
<i>Fes</i>	Enterochelin esterase	2.0	+13
<i>sodB</i>	Iron superoxide dismutase	2.0	-15
<i>yohM</i>	Hypothetical protein	2.1	-9.9
<i>sdhD</i>	Succinate dehydrogenase subunit	2.0	-8.6
<i>acnB</i>	Aconitate hydratase B	2.0	-7.2
<i>sdhA</i>	Succinate dehydrogenase subunit	2.0	-7.1

<sup>a</sup> all values are  $\times 10^{-5}$

Our previous microarray study (54) only identified 46 down-regulated genes, while the current study revealed a 3-fold increase in the number of down-regulated genes. Only 32% of the genes in the former study were identified in the current data, and most of the genes found in both studies are involved in flagellar biosynthesis. *ZntA*, the high specificity exporter of Zn(II) (21), was down-regulated (2.5 fold), and this value is similar to that found in the qPCR studies. Of the remaining genes that were down-regulated in *E. coli* cells stressed for 30 minutes with TPEN, 33% are members of the energy production and conversion COG, 16% belong to the cell motility COG (flagellar biosynthesis), 11% belong to the amino acid transport and metabolism COG, 5% is not a member of any identified COG, and the remaining genes are spread among the other COG's (60).

### **There are No Zn(II)-Responsive, Cytoplasmic Zn(II)-Metallochaperones in *E. coli***

The studies above identified a number of Zn(II)-responsive genes in *E. coli*. Surprisingly, the Zur-regulated transcripts (*ykgM*, *znuABC*, *yodA*) were up-regulated much higher than the *ZntR*-regulated transcript (*zntA*). In addition, the expression levels of other proposed Zn(II) transporters such as *zitB*, *zupT*, *yjiP*, and *zraP* were not affected by Zn(II) levels. Almost all of the Zn(II)-responsive genes that were identified in this study have known biological roles. While it is possible that some of these proteins could have dual roles (known role plus a Zn(II) transport role), these data suggest that there are not cytoplasmic Zn(II)-metallochaperones in *E. coli*. Interestingly, very few of the genes shown to exhibit differential expression in our study were identified in previous DNA microarray studies (51-54), presumably due to different times and methods to stress the cells, different microarray slides, different filtering criteria, and different *E. coli* cell lines used. This result supports our conclusion that there are no Zn(II)-responsive, cytoplasmic Zn(II)-metallochaperones in *E. coli*, but this result also points out a significant limitation of DNA microarrays and a caution about the use of this technology to probe global responses to external stresses.

### **A New Model for Explaining Intracellular Zn(II) Transport in Bacteria**

There are at least 100 Zn(II)-metalloproteins in *E. coli* (10), and if intracellular Zn(II) transport is similar to metal transport for other metal ions (24), then there must be at least 100 distinct Zn(II)-metallochaperones. The inability to identify even one Zn(II)-metallochaperone candidate argues against

the use of such proteins in Zn(II) homeostasis in *E. coli*. Therefore, other mechanisms must be in place to transport Zn(II) in this organism and presumably other bacteria. Any new models must account for all of the previously published data on Zn(II) homeostasis. For example: (1) there is 0.1-0.3 mM Zn(II) in *E. coli* (9); (2) all of the Zn(II) is bound (9); (3) Zn(II) must be delivered to the correct protein; (4) there are known membrane-spanning transporters that move Zn(II) in and out of the cytoplasm (4); and (5) there are no known storage molecules (metallothioneins) or storage “organelles” in *E. coli* (32). With these conditions in mind, we offer a novel model to explain Zn(II) homeostasis in *E. coli* that involves the ribosomes (Figure 2).

In this model, Zn(II) is presumably brought into the periplasm via porins. In low Zn(II) conditions, Zn(II) is bound by ZnuA, delivered to ZnuB, and brought into the cytoplasm (12). During conditions with normal Zn(II) levels, Zn(II) can also enter the cytoplasm by other transporters such as ZupT (14, 15). The Zn(II) is delivered to the ribosome in an unknown method; however, we speculate that the L31/YkgM paralog system (36) may lead to the ribosomes being directed to the export proteins, like ZntA or ZitB or the import proteins, like ZnuB or ZupT. In low Zn(II) conditions, YkgM is expressed at high levels, and it replaces L31 on the ribosome (36). YkgM-bound ribosomes are directed toward the Zn(II) import proteins. Once bound to the ribosome, Zn(II) can be co-translationally inserted into nascent polypeptides that have metal binding amino acids oriented towards the metal ion. For proteins that are not Zn(II)-metalloproteins, the Zn(II) diffuses back onto the ribosome; therefore, this model suggests that Zn(II) may be involved in translation of all proteins and could serve as a nucleation site for *in vivo* folding. When intracellular Zn(II) levels are too high, YkgM is not up-regulated so L31 binds in its place, and the ribosome is directed to the Zn(II) exporters ZntA and ZitB. Previous studies have suggested that *E. coli* ribosomes can bind Zn(II) (61); however, the crystal structures of the ribosomes (62), which were purified using buffers containing high concentrations of EDTA, did not apparently contain bound Zn(II). Experiments are underway to test the different parts of this speculative hypothesis.

## Conclusions

Our understanding of Zn(II) homeostasis in *E. coli* is far from complete. It is clear that intracellular Zn(II) transport in bacteria is different than the transport of other metal ions, since none of the known Zn(II) proteins have the genes for transport proteins on the same operon/gene cluster. In addition, DNA microarrays have failed to identify a single Zn(II)-metallochaperone candidate. It is possible that Zn(II)-metallochaperones are not Zn(II)-responsive, and therefore would not be up- or down-regulated by the presence/absence of Zn(II).

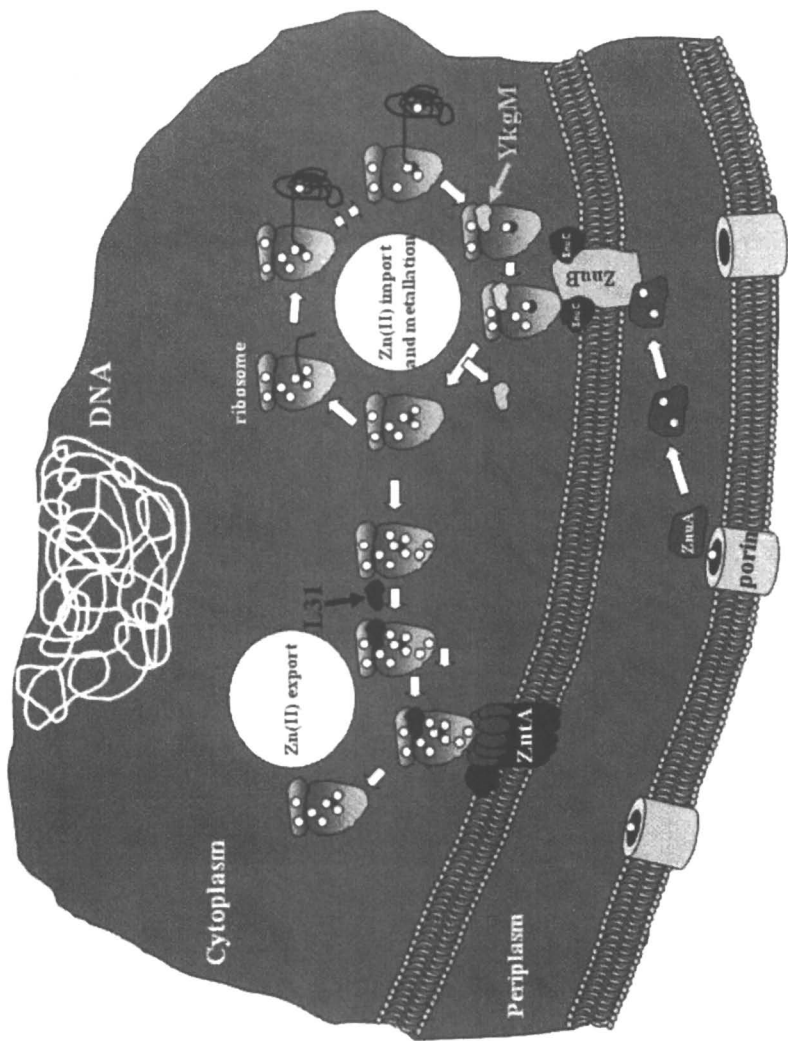


Figure 2. Model for Zn(II) homeostasis in *E. coli*.

However, the use of a metallochaperone to deliver Zn(II) to specific proteins would require at least 100 Zn(II)-metallochaperones in *E. coli*, if the situation for Cu-metallochaperones (63) is evident in this organism. A novel, unprecedented model for Zn(II) homeostasis and transport is offered here; however, there is much work needed to test the different parts of this hypothesis. The understanding of Zn(II) homeostasis in bacteria is highly-significant, as Zn(II) is an essential metal ion for survival and there appear to be different mechanisms for Zn(II) homeostasis in eukaryotes as compared to prokaryotes (8). These characteristics suggest that Zn(II) homeostatic pathways in bacteria may be good targets for inhibitor design efforts, and these inhibitors have the potential to be novel antibiotics. In addition, Zn(II) homeostasis appears to be involved in virulence of several bacterial strains. For example, Yang *et al.* demonstrated that a Zur mutant of *X. oryzae* (cause of bacterial blight) showed a reduction in virulence on rice (64). A ZnuB mutant of *Brucella abortus* has been tested as a live vaccine candidate in wild life and livestock (65). Lastly, neutrophil-derived calprotectin inhibits *S. aureus* infections by chelation of Zn(II) and Mn(II) (66). It is clear that our understanding of Zn(II) homeostatic pathways has significant biomedical implications.

## References

1. Vallee, B. L.; Auld, D. S. In *Interface Between Chemistry and Biochemistry*; Joiles, P.; Jornvall, H., Eds.; Verlag: Basel, Switzerland, 1995; pp. 259-277.
2. Vallee, B. L.; Falchuk, K. H. *Physiol. Rev.* **1993**, *73*, 79-118.
3. Blencowe, D. K.; Morby, A. P. *FEMS Microbiol. Lett.* **2003**, *27*, 291-311.
4. Chimienti, F.; Aouffen, M.; Favier, A.; Seve, M. *Curr. Drug Targets* **2003**, *4*, 323-338.
5. MacDiarmid, C. W.; Gaither, L. A.; Eide, D. *EMBO J.* **2000**, *19*, 2845-2855.
6. Eide, D. J. *Biochim. Biophys. Acta* **2006**, *1763*, 711-722.
7. Gaither, L. A.; Eide, D. J. *Biometals* **2001**, *14*, 251-270.
8. Sekler, I.; Sensi, S. L.; Hershinkel, M.; Silverman, W. F. *Mol. Med.* **2007**, *13*, 337-343.
9. Outten, C. E.; O'Halloran, T. V. *Science* **2001**, *292*, 2488-2492.
10. Andreini, C.; Banci, L.; Bertini, I.; Rosato, A. *J. Proteome Res.* **2006**, *5*, 3173-3178.
11. Nikaido, H. *Mol. Microbiol.* **1992**, *6*, 435-442.
12. Patzer, S. I.; Hantke, K. *Mol. Microbiol.* **1998**, *28*, 1199-1210.
13. Garrido, M. E.; Bosch, M.; Medina, R.; Llagostera, M.; Perez de Rozas, A. M.; Badiola, I.; Barbe, J. *FEMS Microbiol. Lett.* **2003**, *221*, 31-37.

14. Grass, G.; Wong, M. D.; Rosen, B. P.; Smith, R. L.; Rensing, C. *J. Bacteriol.* **2002**, *184*, 864-866.
15. Grass, G.; Franke, S.; Taudte, N.; Nies, D. H.; Kucharski, L. M.; Maguire, M. E.; Rensing, C.; Wong, M. D.; Rosen, B. P.; Smith, R. L. *J. Bacteriol.* **2005**, *187*, 1604-1611.
16. Harris, R. M.; Webb, D. C.; Howitt, S. M.; Cox, G. B. *J. Bacteriol.* **2001**, *183*, 5008-5014.
17. David, G.; Blondeau, K.; Schiltz, M.; Penel, S.; Lewit-Bentley, A. *J. Biol. Chem.* **2003**, *278*, 43728-43735.
18. Puskarova, A.; Ferienc, P.; Kormanec, J.; Homerova, D.; Farewell, A.; Nystrom, T. *Microbiology* **2002**, *148*, 3801-3811.
19. Rensing, C.; Mitra, B.; Rosen, B. P. *Proc. Natl. Acad. Sci. U S A* **1997**, *94*, 14326-14331.
20. Sharma, R.; Rensing, C.; Rosen, B. P.; Mitra, B. *J. Biol. Chem.* **2000**, *275*, 3873-3878.
21. Brocklehurst, K. R.; Hobman, J. L.; Lawley, B.; Blank, L.; Marshall, S. J.; Brown, N. L.; Morby, A. P. *Mol. Microbiol.* **1999**, *31*, 893-902.
22. Grass, G.; Fan, B.; Rosen, B. P.; Franke, S.; Nies, D. H.; Rensing, C. *J. Bacteriol.* **2001**, *183*, 4664-4667.
23. Lu, M.; Fu, D. *Science* **2007**, *317*, 1746-1748.
24. Rosenzweig, A. C. *Chem. Biol.* **2002**, *9*, 673-677.
25. O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **2000**, *275*, 25057-25060.
26. Arnesano, F.; Banci, L.; Bertini, I.; Ciofi-Baffoni, S.; Molteni, E.; Huffman, D. L.; O'Halloran, T. V. *Genome Res.* **2002**, *12*, 255-271.
27. Solioz, M.; Stoyanov, J. V. *FEMS Microbiol. Lett.* **2003**, *27*, 183-195.
28. Luk, E.; Jensen, L. T.; Cullota, V. C. *J. Biol. Inorg. Chem.* **2003**, *8*, 803-809.
29. Beyersmann, D.; Schmidt, C. In *Metals and Genetics*; Sarkar, B., Ed.; Kluwer Academic/Plenum: New York 1999; pp 145-158.
30. Beyersmann, D.; Haase, H. *BioMetals* **2001**, *14*, 331-341.
31. Maret, W. *J. Nutrition* **2003**, *133*, 1460S-1462S.
32. Robinson, N. J.; Whitehall, S. K.; Cavet, J. S. *Adv. Microb. Physiol.* **2001**, *44*, 183-213.
33. Natori, Y.; Nanamiya, H.; Akanuma, G.; Kosono, S.; Kudo, T.; Ochi, K.; Kawamura, F. *Mol. Microbiol.* **2007**, *63*, 294-307.
34. Akanuma, G.; Nanamiya, H.; Natori, Y.; Nomura, N.; Kawamura, F. *J. Bacteriol.* **2006**, *188*, 2715-2720.
35. Nanamiya, H.; Akanuma, G.; Natori, Y.; Murayama, R.; Kosono, S.; Kudo, T.; Kobayashi, K.; Ogawawara, N.; Park, S.-M.; Ochi, K.; Kawamura, F. *Mol. Microbiol.* **2004**, *52*, 273-283.
36. Makarova, K. S.; Ponomarev, V. A.; Koonin, E. V. *Genome Biol.* **2001**, *2*, 1-14.

37. Ramsay, L. M.; Gadd, G. M. *FEMS Microbiol. Lett.* **1997**, *152*, 293-298.
38. Lopez-Garcia, C.; Molowny, A.; Ponsoda, X.; Nacher, J.; Sancho-Bielsa, F. *Rev. Neurol.* **2001**, *33*, 341-347.
39. Suhy, D. A.; Simon, K. D.; Linzer, D. I. H.; O'Halloran, T. V. *J. Biol. Chem.* **1999**, *274*, 9183-9192.
40. Mulrooney, S. B.; Hausinger, R. P. *FEMS Microbiol. Rev.* **2003**, *27*, 239-261.
41. Lin, S. J.; Culotta, V. C. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 3784-3788.
42. Glerum, D. M.; Shtanko, A.; Tzagoloff, A. *J. Biol. Chem.* **1996**, *271*, 14504-14509.
43. Culotta, V. C.; Klomp, L. W. J.; Strain, J.; Casareno, R. L. B.; Krems, B.; Gitlin, J. D. *J. Biol. Chem.* **1997**, *272*, 23469-23472.
44. Puig, S.; Thiele, D. *Curr. Opin. Chem. Biol.* **2002**, *6*, 171-180.
45. Harrison, M. D.; Jones, C. E.; Solioz, M.; Dameron, C. T. *Trends Biochem. Sci.* **2000**, *25*, 29-32.
46. Huffman, D. L.; O'Halloran, T. V. *Annu. Rev. Biochem.* **2001**, *70*, 677-148.
47. Rensing, C.; Grass, G. *FEMS Microbiol. Rev.* **2003**, *27*, 197-213.
48. Katayama, A.; Tsujii, A.; Wada, A.; Nishino, T.; Ishihama, A. *Eur. J. Biochem.* **2002**, *269*, 2403-2413.
49. Hopkin, K.; Papazian, M.; Steinman, H. *J. Biol. Chem.* **1992**, *267*, 24253-24258.
50. Anderegg, v. G.; Hubmann, E.; Podder, N. G.; Wenk, F. *Helv. Chem. Acta* **1977**, *60*, 123-140.
51. Brocklehurst, K. R.; Morby, A. P. *Microbiology* **2000**, *146*, 2277-2282.
52. Lee, L. J.; Barrett, J. A.; Poole, R. K. *J. Bacteriol.* **2005**, *187*, 1124-1134.
53. Yamamoto, K.; Ishihama, A. *J. Bacteriol.* **2005**, *187*, 6333-6340.
54. Sigdel, T. K.; Easton, J. A.; Crowder, M. W. *J. Bacteriol.* **2006**, *188*, 6709-6713.
55. Pruteanu, M.; Neher, S. B.; Baker, T. A. *J. Bacteriol.* **2007**, *189*, 3017-3025.
56. Gaballa, A.; Helmann, J. D. *J. Bacteriol.* **1998**, *180*, 5815-5821.
57. Outten, C. E.; Outten, F. W.; O'Halloran, T. V. *J. Biol. Chem.* **1999**, *274*, 37517-37524.
58. Baba, T.; Ara, T.; Hasegawa, M.; Takai, Y.; Okumura, Y.; Baba, M.; Datsenko, K. A.; Tomita, M.; Wanner, B. L.; Mori, H. *Mol. Systems Biol.* **2006**, *2* (Open Access Article Number: 2006.0008).
59. Riley, M.; Abe, T.; Arnaud, M. B.; Berlyn, K. B.; Blattner, F. R.; Chaudhuri, R. R.; Glasner, J. D.; Horiuchi, T.; Keseler, I. M.; Kosuge, T.; Mori, H.; Perna, N. T.; Plunkett, G.; Rudd, K. E.; Serres, M. H.; Thomas, G. H.; Thomson, N. R.; Wishart, D.; Wanner, B. L. *Nucl. Acids Res.* **2006**, *34*, 1-9.
60. Tatusov, R. L.; Koonin, E. V.; Lipman, D. J. *Science* **1997**, *278*, 631-637.
61. Nomura, M. *Bacteriol. Rev.* **1970**, *34*, 228-277.



62. Schuwirth, B. S.; Borovinskaya, M. A.; Hau, C. W.; Zhang, W.; Vila-Sanjurjo, A.; Holton, J. M.; Cate, J. H. D. *Science* **2005**, *310*, 827-834.
63. Rosenzweig, A. C. *Acc. Chem. Res.* **2001**, *34*, 119-128.
64. Yang, W.; Liu, Y.; Chen, L.; Gao, T.; Hu, B.; Zhang, D.; Liu, F. *Curr. Microbiol.* **2007**, *54*, 307-314.
65. Yang, X.; Becker, T.; Walters, N.; Pascual, D. W. *Infect. Immun.* **2006**, *74*, 3874-3879.
66. Corbin, B. D.; Seeley, E. H.; Raab, A.; Feldmann, J.; Miller, M. R.; Torres, V. J.; Anderson, K. L.; Dattilo, B. M.; Dunman, P. M.; Gerads, R.; Caprioli, R. M.; Nacken, W.; Chazin, W. J.; Skaar, E. P. *Science* **2008**, *319*, 962-965.



## **Synthetic Models**



## Chapter 7

# **N<sub>2</sub>S<sub>3</sub>X-Fe Models of Nitrile Hydratase**

**Martin G. O'Toole and Craig A. Grapperhaus\***

**Department of Chemistry,  
University of Louisville, Louisville, KY 40292**

A series of N<sub>2</sub>S<sub>3</sub>X-Fe iron complexes employing the pentadentate ligand 4,7-bis-(2'-methyl-2'-mercaptopropyl)-1-thia-4,7-diazacyclononane, (bmmp-TASN)<sup>2-</sup>, has been developed as synthetic analogues of the active site of the enzyme nitrile hydratase (NHase). The spin state of the complex is controlled by the identity of the variable ligand (X = Cl, CN, NO, OFe(bmmp-TASN)). The low-spin complexes serve as effective structural and electronic mimics of NHase due to similarities in donor ability of thioethers and sulfur-oxygenates. The importance of the spin state on the oxygen sensitivity of iron-thiolates and the possible role of sulfur oxygenation in the native enzyme are examined.

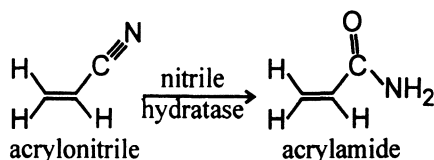
As highlighted in previous reviews of nitrile hydratase, nitriles from both natural and industrial sources are a prevalent chemical functionality in our environment (1-3). In industry, nitriles find use in the production of fine chemicals, as pesticides, and as feedstock for the production of polyacrylamide and other polymers. Historically, waste products from these processes were dumped into open waterways or pumped into deep pressure wells, leading to deleterious environmental effects. Naturally occurring nitriles include

indoleacetonitriles, which serve as growth hormones in plants, and cyanoglycosides, which are used as chemical defense agents in plants and bacteria. The discovery of the ability of some organisms to assimilate nitriles eventually led to the elucidation of two pathways of nitrile metabolism. One involves the enzyme nitrilase, which hydrolyzes the nitrile to produce the corresponding carboxylic acid and ammonia. The second pathway utilizes the hydrolytic enzyme nitrile hydratase (NHase).

As shown in Scheme 1, NHase catalyzes the conversion of nitriles to amides (4). The active site contains either a non-corrin cobalt(III) or non-heme iron(III). Amino acid sequence comparisons have shown that the primary coordination sphere is conserved regardless of the identity of the metal center and consists of a -C-S-L-C-S-C- motif (5). EPR studies on Fe-NHase revealed the iron center maintains a low-spin Fe(III) state throughout the catalytic cycle, and that the iron center has a variable coordination site for substrate interaction (6). These findings are consistent with the hypothesis that the enzyme functions solely as a hydrolytic (*i.e.*, redox-inactive) catalyst. Incubation of the enzyme with nitric oxide in the dark inactivates the enzyme. Exposure to light was found to reinstate activity with concomitant loss of NO, thus revealing a novel photo-regulatory mechanism (7-10).

X-ray diffraction studies have revealed an active site consisting of a single metal center coordinated by two deprotonated amides derived from the peptide backbone, three cysteine sulfurs, and a sixth coordination site occupied by a variable ligand, Figure 1 (11, 12). In the resting form of the enzyme, the sixth position has been determined to be occupied by water/hydroxide derived from the solvent. Structural studies on NHase<sub>dark</sub>, the NO-inactivated form of the enzyme, (at 1.7 Å resolution) further revealed that two of the cysteine derived sulfurs are post-translationally modified in an asymmetric fashion (12). Cys-112 was found to be modified to a sulfinato (RSO<sub>2</sub><sup>-</sup>) moiety, while Cys-114 had undergone single oxygen atom addition to form a sulfenato (RS(O)<sup>-</sup>) ligand. The third thiolate in the active site, Cys-109, remains unmodified and occupies a position trans to the variable site. Culturing of the enzyme under anaerobic conditions resulted in no sulfur-oxygenation and inactive enzyme (13). Exposure to aerobic conditions restored enzymatic activity, suggesting that sulfur-oxygenation is essential for catalytic turnover. While the exact function of this post-translational modification remains unclear, several proposed roles for the sulfur oxygenation (*e.g.*, increasing the Lewis acidity of the iron center (14), facilitating the reversible binding of NO for photo-regulation (15), serving as a proximal base for catalytic turnover (16), or improving the ligand binding affinity for iron) have received some experimental verification.

The mechanism of NHase may proceed through interaction of either the nitrile or water substrate at the metal center, Figure 2. Recent computational



Scheme 1.

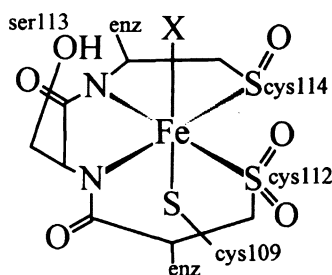


Figure 1. Active site of nitrile hydratase (NHase).

studies suggest an active role for sulfur-oxygenates in facilitating catalysis, either by providing hydrogen bond stabilization to incoming substrates or by acting as a proximal acid/base for the hydrolysis reaction (16-18). In light of these findings, efforts to synthesize Fe sulfur-oxygenate mimics, while maintaining a low-spin Fe(III) state have been undertaken (1-3, 19). A major hurdle in their preparation is the controlled oxygen sensitivity towards sulfur-oxygenation as opposed to disulfide formation.

In this chapter, we will describe our efforts to model the structural, electronic, and functional aspects of NHase using a series of model complexes employing the  $\text{N}_2\text{S}_3$ -pentadentate donor 4,7-bis-(2'-methyl-2'-mercaptopropyl)-1-thia-4,7-diazacyclononane, (bmmp-TASN)<sup>2-</sup>, Scheme 2. This ligand incorporates two tertiary amine nitrogens, two *cis*-thiolate sulfurs, and a thioether sulfur. The parent iron(III) complex [(bmmp-TASN)FeCl] (**1**) is prepared in a one pot synthesis employing  $\text{H}_2\text{TASN}$  (20), isobutylene sulfide (21), and  $[\text{Bu}_4\text{N}][\text{FeCl}_4]$  (22, 23). Additional iron derivatives are prepared by exchanging the chloride for other monodentate donors (23, 24). The choice of monodentate ligand influences the spin-state of the iron. While complex **1** contains high-spin iron(III), exchange of chloride for cyanide yields the low-spin complex [(bmmp-TASN)FeCN] (**2**). The  $\{\text{Fe-NO}\}^6$  nitrosyl complex

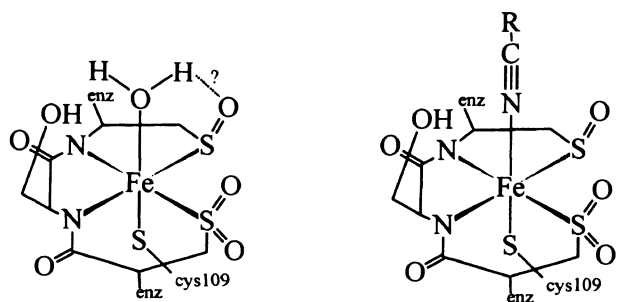
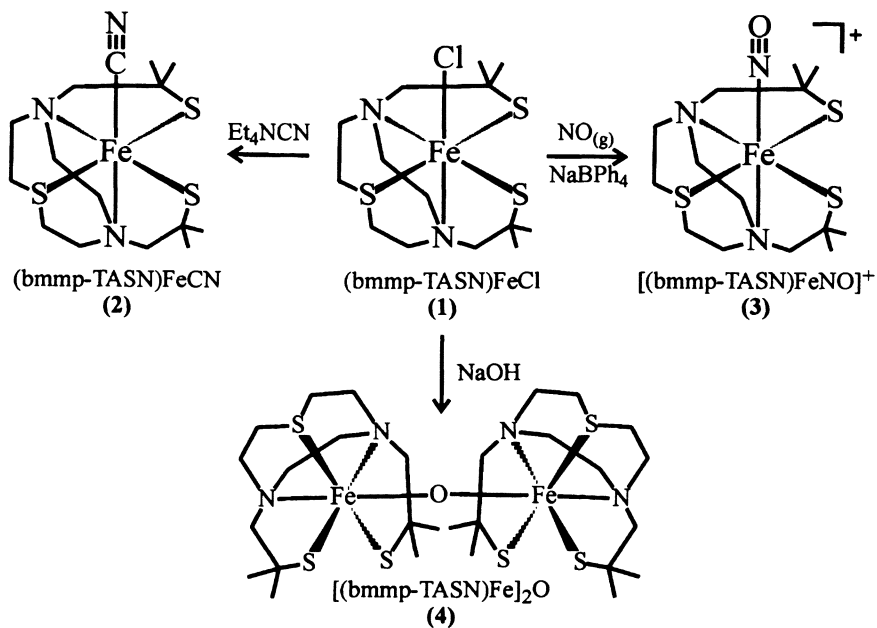


Figure 2. Proposed substrate binding motifs of NHase.



Scheme 2.



$[(\text{bmmp-TASN})\text{FeNO}]^+$  (**3**), which is obtained when nitric oxide is purged through an acetonitrile solution of **1**, also displays a low-spin state. In the presence of strong base, **1** yields the  $\mu$ -oxo dinuclear complex  $[(\text{bmmp-TASN})\text{Fe}]_2\text{O}$  (**4**), which contains two anti-ferromagnetically coupled high-spin iron(III) centers.

## Structural Modeling

We have previously reported the X-ray crystal structures of **2**, **3**, and **4** (23-25). In each case, a pseudo-octahedral  $\text{N}_2\text{S}_3\text{X-Fe}$  complex is observed with a similar arrangement of the  $(\text{bmmp-TASN})^{2-}$  ligand. As highlighted in Figure 3, the structures of our model complexes reproduce key geometric features of the NHase active site. In NHase, the two amido nitrogen donors,  $\text{N}_1$  and  $\text{N}_2$ , and the lone thiolate sulfur,  $\text{S1}$ , occupy one face of the octahedron, while the two post-translationally modified sulfur-oxygenates,  $\text{S2}$  and  $\text{S3}$ , reside on the opposing face that contains the variable ligand,  $\text{X}$ . Since this orientation facilitates possible second-coordination sphere interactions between the oxygenated cysteines and the substrate and/or intermediates, it is a key feature of models attempting to address the effects of S-oxygenation. In the series of  $(\text{bmmp-TASN})\text{FeX}$  model complexes, the facially coordinated TASN ring fulfills the role of the invariant  $\text{N}_2\text{S}$  face. The rigid facial coordination of the TASN ring forces the pendant thiolate donors and the variable ligand onto the opposing face of the octahedron, thus mimicking the arrangement of the oxygenated thiolates and the substrate binding site on the  $\text{S}_2\text{X}$  face of the complex. Unlike NHase, which orientates the three sulfur donors in a facial arrangement with the thiolate trans to the variable site, our model complexes invariably coordinate the three sulfur donors in a meridional arrangement. This forces the amine donor  $\text{N}_1$  trans to the  $\text{S}_3$  thiolate and the amine donor  $\text{N}_2$  trans to the variable site.

Selected bond distances and angles for  $\text{NHase}_{\text{dark}}$  and complexes **2-4** are provided in Table I (12, 23-25). Atoms are labeled as shown in Figure 3. The low-spin nitrosyl complex **3** accurately reproduces the metal-ligand bond distances of  $\text{NHase}_{\text{dark}}$ . The average  $\text{Fe-N}_{\text{amine}}$  distances of 2.044(5) Å in **3** are slightly shorter than the  $\text{Fe-N}_{\text{amido}}$  distances in  $\text{NHase}_{\text{dark}}$  of 2.07 Å. The  $\text{Fe-NO}$  bond in **3** of 1.609(5) Å is 0.04 Å shorter than in the enzyme. Most interestingly, the average  $\text{Fe-S}$  bond distances in **3** and NHase are statistically indistinguishable with values of 2.287(3) and 2.30 Å, respectively. Thus, the  $\text{Fe-S}$  bond distance alone provides no insight as to the identity of the sulfur donor as a thiolate, sulfur-oxygenate, or thioether.

A closer inspection of the  $\text{Fe-S}$  bond distances in **3** reveals the  $\text{Fe-S}$  bond distances to the neutral thioether, 2.285(1) Å, and anionic thiolates, 2.284(1) and 2.291(1) Å, are equivalent within experimental error. As shown in Table I, the  $\text{Fe-S}_{\text{thioether}}$  and  $\text{Fe-S}_{\text{thiolate}}$  bond distances in the low-spin complex **2** are also similar. This trend, however, does not apply to the high-spin complex **4**. In **4**,

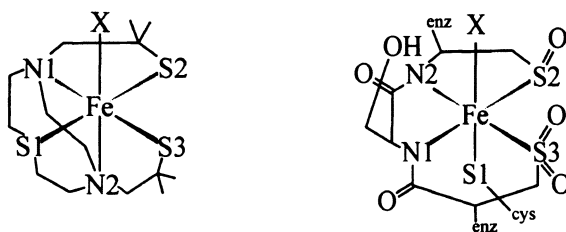


Figure 3. Atom labeling schemes for 1-4 (left) and NHase (right).

Table I. Selected Bond Distances (Å) for 2-4 and NHase<sub>dark</sub>

Bond <sup>f</sup>	2 <sup>b</sup>	3 <sup>c</sup>	4 <sup>d</sup>	NHase <sub>dark</sub> <sup>e</sup>
Fe-S1	2.286(4)	2.285(2)	2.6221(6)	2.30
Fe-S2	2.311(4)	2.284(2)	2.3970(5)	2.33
Fe-S3	2.179(4)	2.291(2)	2.3374(5)	2.26
Fe-N1	2.069(11)	2.024(5)	2.297(1)	2.09
Fe-N2	2.078(12)	2.063(5)	2.377(1)	2.05
Fe-X <sup>f</sup>	1.922(15)	1.609(6)	1.7954(3)	1.65

<sup>a</sup>Atoms labeled as shown in Figure 3. <sup>b</sup>Ref. 25. <sup>c</sup>Ref. 24. <sup>d</sup>Ref. 25. <sup>e</sup>Ref 12. <sup>f</sup>X = NO for NHase<sub>dark</sub> and 3; X = CN for 2; X = OFe(bmmp-TASN) for 4.

the Fe-S<sub>thiolate</sub> bond distances, 2.3374(5) and 2.3970(5) Å, are significantly shorter than the Fe-S<sub>thioether</sub> bond distance, 2.6221(6) Å, consistent with expectations based on relative  $\sigma$ -donor ability of thiolates and thioethers. In the low-spin complexes 2 and 3, additional electronic effects must be present to counter the expected  $\sigma$ -bonding effect.

## Electronic Modeling

Figure 4 compares the relative  $\sigma$ - and  $\pi$ -donor ability of three distinct sulfur donors: thiolate, sulfinate, and thioether (17, 23, 24, 26-28). Based on simple charge considerations, both the thiolate and sulfinate donors are anionic, while the thioether is neutral. In terms of  $\sigma$ -donor ability, the thiolato sulfur should be stronger than the sulfinate since the oxygen atoms of the latter withdraw significant electron density away from sulfur. The neutral thioether is the poorest  $\sigma$ -donor. In terms of  $\pi$ -donor ability, only the thiolate sulfur possess the

requisite p-type lone pair to interact with the  $t_{2g}$  orbitals on iron. Both the sulfinate and thioether lack a lone pair of electrons in an orbital of proper symmetry for  $\pi$ -donation, although they may be  $\pi$ -acceptors (29, 30).

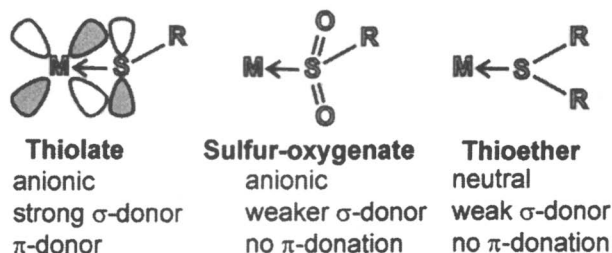


Figure 4. Relative  $\sigma$ - and  $\pi$ -donor abilities of various sulfur ligands.

In low-spin iron(III) complexes, both the  $\sigma$ - and  $\pi$ -donor ability of the sulfur ligand must be considered. When coordinated to the  $\pi$ -donating thiolate, the “ $t_{2g}^5$ ” configuration of low-spin iron(III) results in occupation of both the  $\pi$  and  $\pi^*$  molecular orbitals. A “ $\pi$ -repulsion” occurs since the energy minimum of the  $\pi^*$ -interaction occurs at a longer Fe-S distance than the energy minimum of the  $\pi$ -interaction resulting in a lengthening of the Fe-S bond. This results in a longer Fe-S<sub>thiolate</sub> bond than predicted based only on  $\sigma$ -donor ability. In the case of thioether and sulfinate donors, the ligands are not  $\pi$ -donors and the Fe-S distances are more or less a function of  $\sigma$ -donor ability. This results in nearly equal Fe-S<sub>thiolate</sub> and Fe-S<sub>thioether</sub> bond distances for low-spin Fe(III). The Fe-S<sub>sulfinate</sub> bond distances are slightly shorter.

In high-spin iron(III) complexes, the “ $t_{2g}^3 e_g^{*2}$ ” electron configuration results in a larger ionic radius with respect to low-spin iron(III). Additionally, the “ $t_{2g}$ ” orbitals are only half-occupied. These two effects combine to reduce the significance of  $\pi$ -interactions resulting in Fe-S bond distances that reflect the  $\sigma$ -donor ability of the sulfur donor. As a result, Fe-S<sub>thiolate</sub> bonds are significantly shorter than Fe-S<sub>thioether</sub> bond distances. To our knowledge, no high-spin iron sulfinate complexes have been reported.

As noted above, for low-spin iron(III) complexes the Fe-S bond distances for thiolates, sulfur-oxygenates, and thioethers are similar. However, bond distance is a function of the sum of all interactions and the similar bond distances do not imply these bonds are electronically equivalent. Simple charge arguments would suggest thiolates are better electronic mimics of sulfur-oxygenates than thioethers. While this rudimentary analysis may hold for high-spin iron(III) complexes, for low-spin iron(III) the combined  $\sigma$  and  $\pi$ -interactions described above clearly show that thioethers and sulfur-oxygenates

are more electronically similar than thiolates and sulfur-oxygenates. As such, thioethers can reasonably be employed to mimic the electronic effect of sulfur-oxygenation in low-spin iron complexes. The thioether does not provide the same second coordination sphere interactions as sulfur-oxygenates allowing the electronic effect to be studied independently of other interactions.

The nitrosyl model complex **3** is structurally and electronically similar to the inactive form of the enzyme, NHase<sub>dark</sub>. As described above, the sulfur-oxygenate donors of the enzyme can be modeled electronically by the thioether sulfur incorporated into the TASN ring. As shown in Table II, key spectroscopic features of NHase<sub>dark</sub> are reproduced by **3** despite differences in sulfur donor types in the first coordination sphere.

The electron density at iron can be probed indirectly by IR spectroscopy. The  $\nu_{\text{NO}}$  stretching frequencies of **3** and NHase<sub>dark</sub> differ by  $4\text{ cm}^{-1}$  with values of  $1856$  and  $1852\text{ cm}^{-1}$ , respectively (8, 24). A more direct probe of electron density at iron can be obtained from Mössbauer spectroscopy (24, 31). The isomer shift, which measures the electron density at iron, and the quadrupole splitting, which is a function of anisotropy of the electron density, are strikingly similar between **3** and NHase<sub>dark</sub> consistent with similar electronic environments.

For hydrolytic enzymes, such as NHase, the reduction potential, as an indicator of Lewis acidity, is also of interest. Table III summarizes measured potentials for complexes **1-4** (23, 24). Complexes **1**, **2** and **4** each display a reversible reduction at negative potentials of  $-0.66$ ,  $-0.88$ , and  $-0.67\text{ V}$  (vs. Ag/AgCl), respectively, assigned to a  $\text{Fe}^{\text{III/II}}$  couple. Likewise, an irreversible sulfur-centered oxidation was observed at positive potentials of  $+0.58$ ,  $+0.29$ , and  $+0.66\text{ V}$ . Interestingly, the high-spin complexes **1** and **4** most accurately mimic the  $\text{Fe}^{\text{III/II}}$  reduction potential of NHase ( $-0.48\text{ V}$  vs. NHE;  $-0.68\text{ V}$  vs. Ag/AgCl) (19). Not surprisingly, the cyano complex **2** is more difficult to reduce. Complex **3** displays two reversible events at  $-0.234$  and  $-1.25\text{ V}$ , which are assigned as  $\{\text{FeNO}\}^{6/7}$  and  $\{\text{FeNO}\}^{7/8}$  redox couples, and an irreversible thiolate oxidation ( $+0.96\text{V}$ ).

The redox potentials provide interesting insight into an additional spectroscopic parameter: the charge transfer band displayed by nitrile hydratase at  $710\text{ nm}$  ( $14,800\text{ cm}^{-1}$ ) (19). While the ability of a model complex to mimic this parameter is desirable, it can also be misleading. In a limiting sense, the charge transfer band can be regarded as a coupled metal-centered reduction and ligand-centered oxidation. As such, the energy of the band is related to the difference in the redox potentials between thiolate oxidation and metal reduction, not their absolute values.

As shown in Table III, the high-spin complex **1** and the low-spin nitrosyl complex **3** display charge transfer bands at  $631\text{ nm}$  ( $15,800\text{ cm}^{-1}$ ) and  $650\text{ nm}$  ( $15,400\text{ cm}^{-1}$ ), respectively. This is a difference of only  $400\text{ cm}^{-1}$  even though the spin-states are different and the reduction potentials differ by  $\sim 300\text{ mV}$ . The similarities result from nearly identical potential differences between thiolate oxidation and metal reduction,  $1.24$  and  $1.19\text{ V}$ , respectively. Conversely, the

**Table II. Spectroscopic Comparison of NHase<sub>dark</sub> and 3**

<i>Parameter</i>	<i>NHase<sub>dark</sub><sup>a,b</sup></i>	<i>3<sup>c</sup></i>
IR: $\nu_{\text{NO}}$	1852 $\text{cm}^{-1}$	1856 $\text{cm}^{-1}$
UV-vis: $\lambda_{\text{max}}$	710 nm	650 nm
Mössbauer	0.03 mm/s	0.06 mm/s
$\delta$ ( $\Delta E_Q$ )	(1.47 mm/s)	(1.75 mm/s)

<sup>a</sup>Ref. 8. <sup>b</sup>Ref. 31. <sup>c</sup>Ref. 24.

**Table III. Redox Potentials<sup>a</sup> and Charge Transfer Band Energy**

	<i>1<sup>b</sup></i>	<i>2<sup>b</sup></i>	<i>3<sup>c</sup></i>	<i>4<sup>b</sup></i>	<i>NHase<sub>dark</sub><sup>d</sup></i>
$E_{\text{red}}$ (V)	-0.66 <sup>b</sup>	-0.88	-0.23 -1.25	-0.67 <sup>b</sup>	-0.68
$E_{\text{ox}}$ (V)	+0.58 <sup>e</sup>	+0.29 <sup>e</sup>	+0.96 <sup>e</sup>	+0.66 <sup>e</sup>	not reported
$\Delta E^f$ (V)	1.24	0.96	1.19	1.33	N/A
$\lambda_{\text{max}}$ ( $\text{cm}^{-1}$ )	15,800	15,200	15,400	18,900	14,800

<sup>a</sup>Potentials for 1-4 recorded in  $\text{CH}_2\text{Cl}_2$  vs.  $\text{Ag}/\text{Ag}^+$  ( $\text{Ag}/\text{Ag}^+ = 0.197$  V vs. NHE).

<sup>b</sup>Ref. 24. <sup>c</sup>Ref. 23. <sup>d</sup>Ref. 19. <sup>e</sup>Irreversible event measured at a scan rate of 200 mV/s. <sup>f</sup>Calculated as the absolute value of  $E_{\text{red}} - E_{\text{ox}}$ .

two high-spin complexes **1** and **4** display nearly identical reduction potentials, but vastly different charge transfer band energies of 631 nm ( $15,800 \text{ cm}^{-1}$ ) and 529 nm ( $18,900 \text{ cm}^{-1}$ ). This shift is attributable to a less accessible sulfur-oxidation in **4** as compared to **1**. Thus the charge-transfer band is not a reliable indicator of how well a model reproduces the absolute electronic environment of the enzyme, but rather a measure of how well it separates the oxidized and reduced forms of the complex.

## Functional Modeling

Functional models of nitrile hydratase should reproduce the key reactions occurring at the active site. In addition to the catalytic hydrolysis of nitriles to amides, these also include photo-induced cleavage of coordinated nitric oxide and thiolate oxygenation under aerobic conditions. The nitrosyl complex **3** displays significant structural and electronic features in common with NHase<sub>dark</sub>. These spectroscopic similarities, however, do not translate to a functional model of photo-regulated NO dissociation.

Photolysis of **3**, with a W bulb or Hg vapor lamp source, in various solvents results in no detectable spectroscopic change (23). This observation underscores the labilizing effect of the trans thiolate ligand, which is present in NHase<sub>dark</sub> but not in **3**, for promoting rapid ligand exchange at the variable site (3). As such, the Fe-N<sub>nitrosyl</sub> bond distance in NHase<sub>dark</sub>, 1.65 Å, is slightly longer than in **3**, 1.609(5) Å. Synthetic models that display photo-dissociation possess a trans thiolate and a longer Fe-N<sub>nitrosyl</sub> bond distance (32). While the absence of a trans-thiolate has implications on ligand exchange rates, it does not preclude our models from promoting hydrolysis or sulfur-oxygenation.

The photoreactivity of nitrile hydratase models was addressed through a computational analysis of **3** and other models by Richards, *et al.* (32). Since photo-induced cleavage requires excitation of an electron into an Fe-NO antibonding orbital, the electronic structure of several complexes were investigated. In that study, it was concluded that the low energy band of **3** in the visible region is primarily a d-d transition. Our previously reported DFT calculations reveal significant sulfur character in the HOMO to HOMO-4 and Fe-NO antibonding character in the LUMO and LUMO+1 orbitals, Figure 5, similar to the orbitals described in photo-reactive complexes. The cause for the discrepancy in these two sets of calculations is unclear.

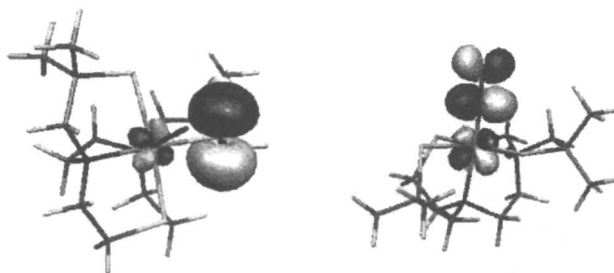


Figure 5. Representation of the HOMO (left) and LUMO (right) of **3**.

Prior to the discovery of oxygenated cysteine residues at the active site of NHase, the oxygen sensitivity of iron-thiolates was well known to yield disulfide and iron-oxo decomposition products (1). Since 1998, a number of discrete iron sulfur-oxygenates have been obtained through direct aerobic oxidation (15, 33-39). Analysis of this literature reveals that in all cases of discrete S-oxygenation, the iron is either in a confirmed low-spin state or a low spin-state can be inferred. Since most iron-thiolate complexes prior to 1998 were high-spin, it is reasonable to suggest that the spin-state of iron dictates the oxygen sensitivity of the thiolate complex. Our series of complexes is well

suites to test this hypothesis since the spin-state can be regulated by changes in the variable ligand X.

Oxygen exposure of the high-spin complex **1** rapidly results in disulfide formation and decomplexation of iron (25). The ultimate iron containing products are insoluble orange iron-oxo clusters. The X-ray structure of one such cluster isolated from a derivative of **1** identifies the iron cluster as  $[\text{Fe}_{14}\text{O}_8(\text{OCH}_2\text{CH}_3)_{20}\text{Cl}_8]^{2-}$  (40). To our knowledge, this is the largest discrete iron-oxo cluster structurally characterized.

Exposure of the low spin complex **2** to aerobic conditions also leads to the formation of insoluble products, however the solid precipitate is green in color (25). The IR spectrum of the iron-containing product reveals strong absorption bands in the region between  $1000\text{-}1200\text{ cm}^{-1}$  consistent with the presence of sulfur-oxygenates. Confirmation of the IR assignment is gained through comparison of the spectrum with that of the structurally characterized nickel disulfonate, (bmmp-O<sub>6</sub>-TASN)Ni. Mass spectral analysis confirms the presence of iron bound sulfur oxygenates.

To account for the spin-state dependent reactivity, theoretical investigations of **1** and **2** have been conducted (25). A covalency analysis of the metal ligand bonds, based on the NBO/NLMO method described by Richards, reveals significant differences between **1** and **2**, Table IV (32, 41). Using a relative covalency scale of 0 to 1 in which 0 represents a strictly ionic bond and 1 a fully covalent bond, the high spin complex **1** displays iron-thiolate bonds with covalency values of 0.11 and 0.40, whereas the low-spin complex **2** displays values of 0.75 and 0.86.

Based on the covalency analysis, **1** can be regarded as having more ionic iron-thiolate bonds, while **2** has more covalent iron-thiolate bonds. From experimental results, **1** reacts with oxygen to yield disulfides and iron-oxo

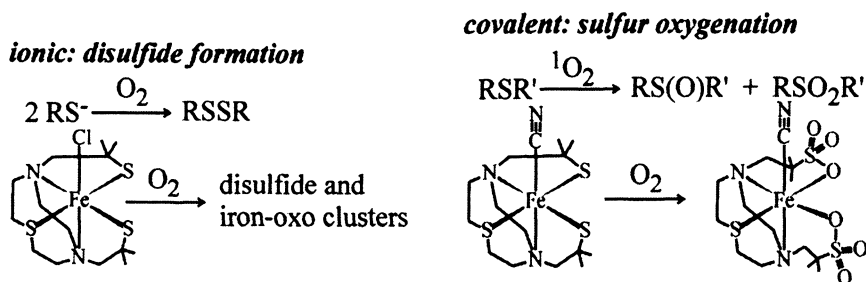
**Table IV. Calculated Bond Covalencies<sup>a</sup>**

<i>Bond<sup>b</sup></i>	<i>1</i>	<i>2</i>
Fe-S1	0.32	0.84
Fe-S2	0.11	0.75
Fe-S3	0.40	0.86
Fe-N1	0.08	0.13
Fe-N2	0.10	0.58
Fe-X	0.29	0.56
Total Covalency	1.30	3.73

<sup>a</sup>Calculated as described in refs. 25, 32, and 41.

<sup>b</sup>Atoms labeled as shown in Figure 3.

clusters, whereas oxygen exposure of **2** results in sulfur-oxygenation. This reactivity mirrors that observed with organic thiolates and sulfides, Scheme 3 (42). While thiols are difficult to oxidize, deprotonation to the ionic thiolate typically results in rapid oxidation to disulfide. With thioethers, excitation of oxygen to the singlet excited state, to overcome a spin-barrier, yields sulfoxides and sulfones. Since complex **2** is paramagnetic, there is no spin barrier with ground state oxygen and the reaction proceeds without excitation of oxygen.



Scheme 3.

While some functional hydrolytic models of NHase have been obtained with cobalt mimics of NHase, to our knowledge catalytic nitrile hydrolysis has not been observed with iron-based mimics (14). In order to achieve hydrolysis of nitriles, the iron center must be able to function as a Lewis acid to coordinate the nitrile and polarize the carbon nitrogen triple bond or bind water and lower its  $\text{pK}_a$ . In either event, an accessible coordination site on iron is required.

Since the spin-state of the (bmmp-TASN)FeX series of complexes is dependent on X, any of the complexes 1-4 could serve as a suitable precursor for substrate binding studies. However, the complexes 1-4 do not readily dissociate their variable ligand in solution. As noted above, this can be attributed to the presence of an amine, and not a thiolate, trans to the variable ligand.

Recent advances towards a functional mimic have focused on metathesis of the coordinated chloride for weakly coordinating anions, such as triflate. Alternately, the triflate complex can also be obtained from **4** by addition of triflic acid. Preliminary results with (bmmp-TASN)FeOTf reveal a weakly bound triflate anion that can be displaced in solution by ligands including acetonitrile and water. If the primary role of S-oxygenation is the tuning of the iron acidity, these complexes are likely to show hydratase activity. However, if S-oxygenation also serves to promote substrate reactivity via a second-coordination sphere interactions, further refinements will be necessary.



## Future Directions

The knowledge gained from experimental work and computational studies on the [(btmp-TASN)FeX] series of complexes has made it possible to understand the electronic factors that will lead to a successful mimic of NHase. Strong donor ligands are required in order to maintain iron in a low spin state in both six and five coordinate environments of the type seen during catalytic turnover. The low spin environment is in turn necessary for the production of sulfur oxygenates which can fine tune the Lewis acidity of the iron and potentially participate in the acid/base chemistry of the hydrolysis cycle. Alternately, thioethers can be employed as electronic mimics of sulfur-oxygenates to prepare functional models. Such an approach would avoid the complications of controlled sulfur-oxygenation. These lessons can now be applied to future production of catalytic mimics of the nitrile hydratase enzyme.

## Acknowledgements

This research was supported in part by the National Science Foundation (CHE-0238137 and CHE- 0749965).

## References

1. Mascharak, P. K. *Coord. Chem. Rev.* **2002**, *225*, 201-214.
2. Harrop, T. C.; Mascharak, P. K. *Acc. Chem. Res.* **2004**, *37*, 253-260.
3. Kovacs, J. A. *Chem. Rev.* **2004**, *104*, 825-848.
4. Sugiura, Y.; Kuwahara, J.; Nagasawa, T.; Yamada, H. *J. Am. Chem. Soc.* **1987**, *109*, 5848-5850.
5. Kobayashi, M.; Shimizu, S. *Nat. Biotechnol.* **1998**, *16*, 733-736.
6. Nelson, M. J.; Jin, H.; Turner, I. M., Jr.; Grove, G.; Scarrow, R. C.; Brennan, B. A.; Que, L., Jr. *J. Am. Chem. Soc.* **1991**, *113*, 7072-7073.
7. Endo, I.; Odaka, M.; Yohda, M. *Trends Biotechnol.* **1999**, *17*, 244-249.
8. Noguchi, T.; Honda, J.; Nagamune, T.; Sasabe, H.; Inoue, Y.; Endo, I. *FEBS Lett.* **1995**, *358*, 9-12.
9. Tsujimura, M.; Odaka, M.; Nagashima, S.; Yohda, M.; Endo, I. *J. Biochem. (Tokyo)* **1996**, *119*, 407-413.
10. Noguchi, T.; Hoshino, M.; Tsujimura, M.; Odaka, M.; Inoue, Y.; Endo, I. *Biochemistry* **1996**, *35*, 16777-16781.
11. Huang, W. J.; Jia, J.; Cummings, J.; Nelson, M.; Schneider, G.; Lindqvist, Y. *Structure* **1997**, *5*, 691-699.
12. Shigehiro, S.; Nakasako, M.; Dohmae, N.; Tsujimura, M.; Tokoi, K.; Odaka, M.; Yohda, M.; Kamiya, N.; Endo, I. *Nat. Struct. Biol.* **1998**, *5*, 347-351.

13. Murakami, T.; Nojiri, M.; Nakayama, H.; Odaka, M.; Yohda, M.; Dohmae, N.; Takio, K.; Nagamune, T.; Endo, I. *Protein Sci.* **2000**, *9*, 1024-1030.
14. Tyler, L. A.; Noveron, J. C.; Olmstead, M. M.; Mascharak, P. K. *Inorg. Chem.* **2003**, *42*, 5751-5761.
15. Lee, C. M.; Chen, C. H.; Chen, H. W.; Hsu, J. L.; Lee, G. H.; Liaw, W. F. *Inorg. Chem.* **2005**, *44*, 6670-6679.
16. Hopmann, K. H.; Guo, J. D.; Himoto, F. *Inorg. Chem.* **2007**, *46*, 4850-4856.
17. Dey, A.; Jeffrey, S. P.; Darensbourg, M. Y.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *Inorg. Chem.* **2007**, *46*, 4989-4996.
18. Yano, T.; Wasada-Tsutsui, Y.; Arii, H.; Yamaguchi, S.; Funahashi, Y.; Ozawa, T.; Masuda, H. *Inorg. Chem.* **2007**, *46*, 10345-10353.
19. Artaud, I.; Chatel, S.; Chauvin, A. S.; Bonnet, D.; Kopf, M. A.; Leduc, P. *Coord. Chem. Rev.* **1999**, *192*, 577-586.
20. Hoffman, P.; Steinhoff, A.; Mattes, R. *Z. Naturforsch., B: Chem. Sci.* **1987**, *42b*, 867-873.
21. Snyder, H. R.; Stewart, J. M.; Ziegler, J. B. *J. Am. Chem. Soc.* **1947**, *69*, 2672-2674.
22. Clausen, C. A.; Good, M. L. *Moess. Effect Meth.* **1968**, *4*, 187-200.
23. Grapperhaus, C. A.; Li, M.; Patra, A. K.; Poturovic, S.; Kozlowski, P. M.; Zgierski, M. Z.; Mashuta, M. S. *Inorg. Chem.* **2003**, *42*, 4382-4388.
24. Grapperhaus, C. A.; Patra, A. K.; Mashuta, M. S. *Inorg. Chem.* **2002**, *41*, 1039-1041.
25. O'Toole, M. G.; Kreso, M.; Kozlowski, P. M.; Mashuta, M. S.; Grapperhaus, C. A. *J. Biol. Inorg. Chem.* **2008**, *in press*.
26. Ashby, M. T.; Enemark, J. H.; Lichtenberger, D. L. *Inorg. Chem.* **1988**, *27*, 191-197.
27. Grapperhaus, C. A.; Mullins, C. S.; Kozlowski, P. M.; Mashuta, M. S. *Inorg. Chem.* **2004**, *43*, 2859-2866.
28. Farmer, P. J.; Reibenspies, J. H.; Lindahl, P. A.; Darensbourg, M. Y. *J. Am. Chem. Soc.* **1993**, *115*, 4665-4674.
29. Blake, A. J.; Holder, A. J.; Hyde, T. I.; Schröder, M. *J. Chem. Soc., Chem. Commun.* **1989**, 1433-1434.
30. Mullen, G. E. D.; Fässler, T. F.; Went, M. J.; Howland, K.; Stein, B.; Blower, P. J. *J. Chem. Soc., Dalton Trans.* **1999**, 3759-3766.
31. Popescu, V. C.; Münck, E.; Fox, B. G.; Sanakis, Y.; Cummings, J. G.; Turner, I. M., Jr.; Nelson, M. J. *Biochemistry* **2001**, *40*, 7984-7991.
32. Greene, S. N.; Richards, N. G. J. *Inorg. Chem.* **2004**, *43*, 7030-7041.
33. Chen, H. W.; Lin, C. W.; Chen, C. C.; Yang, L. B.; Chiang, M. H.; Liaw, W. F. *Inorg. Chem.* **2005**, *44*, 3226-3232.
34. Duff, S. E.; Hitchcock, P. B.; Davies, S. C.; Barclay, J. E.; Evans, D. J. *Acta Crystallogr. E* **2005**, *61*, M1313-M1315.
35. Galardon, E.; Giorgi, M.; Artaud, I. *Chem. Commun.* **2004**, 286-287.
36. Heinrich, L.; Li, Y.; Vaissermann, J.; Chottard, G.; Chottard, J. C. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 3526-3528.

37. Lee, C. M.; Hsieh, C. H.; Dutta, A.; Lee, G. H.; Liaw, W. F. *J. Am. Chem. Soc.* **2003**, *125*, 11492-11493.
38. Noveron, J. C.; Olmstead, M. M.; Mascharak, P. K. *J. Am. Chem. Soc.* **2001**, *123*, 3247-3259.
39. Sellmann, D.; Hein, K.; Heinemann, F. W. *Inorg. Chim. Acta* **2004**, *357*, 3739-3745.
40. Grapperhaus, C. A.; O'Toole, M. G.; Mashuta, M. S. *Inorg. Chem. Comm.* **2006**, *9*, 1204-1206.
41. Chang, C. H.; Boone, A. J.; Bartlett, R. J.; Richards, N. G. J. *Inorg. Chem.* **2004**, *43*, 458-472.
42. Oae, S. *The Organic Chemistry of Sulfur*; Plenum: New York, NY, 1977.



## Chapter 8

# Studies into the Metal Chemistry of the Carbaporphyrinoids

## Insights into the Biological Choice of Porphyrin

Christopher J. Ziegler

Department of Chemistry,  
University of Akron, Akron, OH 44325-3501

Porphyrin macrocycles are ubiquitous in biology; however the chemical reasons behind the biological choice of porphyrin remain unclear. Investigations into the analogous rings with skeletal modifications can provide insight into this question. In this chapter, the metal binding chemistry of the carbaporphyrins, porphyrin analogs with one or more carbon atoms at core positions, is discussed. This replacement significantly modifies the metal binding properties of the ring, as well as the electronic structure and reactivity of the resultant metal adducts

The importance of the porphyrin macrocycle to biology can not be overstated (1-3). Heme-based enzymes and proteins, which contain iron porphyrins in their active sites, are ubiquitous in biological systems. In mammals, key metabolic and catabolic pathways are mediated by these metallomacrocycles, including oxygen transport, storage and activation, catalysis, membrane transport, electron transfer, and substrate sensing (4). In

addition to higher organisms, heme enzymes are observed across the phyla, ranging from bacteria to green plants (5). The prevalence of porphyrinic macrocycles in biological systems raises the question of the reasons behind this apparent choice in cofactor. Biological systems persistently employ porphyrins for one of two possible reasons. The first interpretation is that porphyrin has thermodynamic and kinetic attributes that are optimal for its various applications in biology. Porphyrin is remarkably stable, and readily coordinates to the first row transition metal ions that are prevalent in biology (6). The second explanation is that organisms use porphyrin because it is available, and that the malleability of this macrocycle lends itself to a variety of applications. These two possible explanations raise fundamental questions about the ubiquity of porphyrin. Is there something about the structure of porphyrin that is optimal for the many roles it plays in biology, or do organisms use porphyrins in various proteins because it is available?

The porphyrin macrocycle is composed of four pyrrolic subunits linked together by single  $sp^2$  hybridized carbon linkers (Figure 1). The nitrogen atoms face inward, resulting in a square of coordinating heteroatoms directed toward the core of the ring, ideal for binding intermediate-sized transition metal ions. The aromatic pyrrole rings are conjugated to each other through the bridging meso carbons, resulting in a highly stable and aromatic 18  $\pi$  electron annulene ring. The ring current passes through the internal nitrogens, resulting in electronic interactions between the  $\pi$  system of the ring and bound metal ions; this results in readily observable changes to the absorption spectrum of porphyrins depending on the identity of the metal ion, the oxidation state, and the presence of axial ligands on the metal. Finally, porphyrin typically is a dianionic ligand for most metals, and generally stabilizes divalent oxidation states in the absence of additional anions.

One method for addressing why porphyrin is the macrocycle of choice for biological systems is to vary the structure of the ring and observe changes to its coordination chemistry. There have been many studies into the effects of peripheral substitution on metalloporphyrin properties and reactivity (7-9), where the electron donating or withdrawing attributes of the substituent can perturb the  $\pi$  system and subsequently the chemistry of the metal site. While this work has provided much insight into the relationship between the porphyrin  $\pi$  system and the chemistry of bound metals, peripheral modification does not inform on the reasons behind the selection of the specific twenty carbon-four nitrogen structure of porphyrin.

The specific attributes and characteristics of the porphyrin macrocycle are best determined by actually modifying the porphyrin skeleton itself, and observing the subsequent changes to the properties of metal adducts. There are several ways one can change the twenty carbon-four nitrogen skeleton of

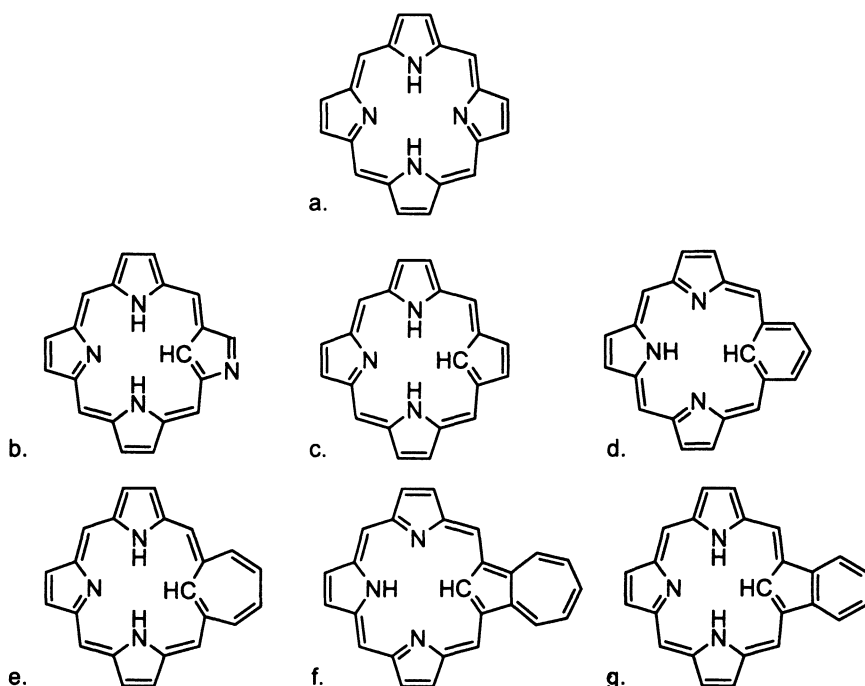


Figure 1. Porphyrin (a) and several carbaporphyrins: N-confused porphyrin (b); carbaporphyrin (c); *m*-benziporphyrin (d); tropiporphyrin (e); azuliporphyrin (f); benzocarbaporphyrin (g).

porphyrin to generate alternate structures (10). The different types of porphyrin variation fall into four categories: the isomers of porphyrin, contracted rings, expanded rings, and skeletal modified rings. One type of skeletal modification is via substitution of one or more of the nitrogen atoms at the core with other elements. This has been carried out with phosphorous (11), oxygen, sulfur, selenium (12-14), and carbon (15). These last types of carbon-modified porphyrin analogs and isomers have been dubbed the carbaporphyrins.

There are several well known examples where “modified” porphyrins are used in biology, and in these cases the changes to the macrocycle structure do result in improved properties. The chlorophylls and chlorins are both porphyrinoids that are reduced at  $\beta$  pyrrolic positions (16). The effect of this modification is an increase in excited state lifetimes and quantum yields of

fluorescence. The cobalamin macrocycle, found in the B-12 cofactor, is varied in two ways: the ring is contracted and significantly reduced versus normal porphyrin (17). This results in a ring with a single proton at the core, which helps stabilize the lower valent Co(I)/Co(II) redox couple. Although one can only speculate how additional alternative porphyrinic macrocycles might play roles in biology, recent fundamental advances in porphyrin synthetic chemistry can be used to examine the metal binding properties and resultant stability and reactivity of their metal adducts.

In this chapter, we shall consider only one type of change to the porphyrin macrocycle: the replacement of one or more interior nitrogen positions with carbon atoms. Figure 1 also shows several carbaporphyrinoids that are currently under investigation in several research groups. Carbaporphyrins can be either porphyrin isomers, as seen in the *N*-confused porphyrins (NCPs) (18, 19), or porphyrin analogs. The analogs include the benziporphyrins (20), expanded porphyrinoids where six member benzene rings have replaced a pyrrole ring, and azuliporphyrin (21), a "true carbaporphyrin" where the pyrrole has been substituted with an azulene ring. In addition to normal porphyrin derivatives, there are also carbon substituted phthalocyanine analogs known as the hemiporphyrazines, as shown in Figure 2. These rings have been known for more than five decades, but only recently has their metal binding chemistry been examined. In all of these macrocycles, the internal carbon position is an  $sp^2$  hybridized atom with a C-H bond that extends into the core of the ring.

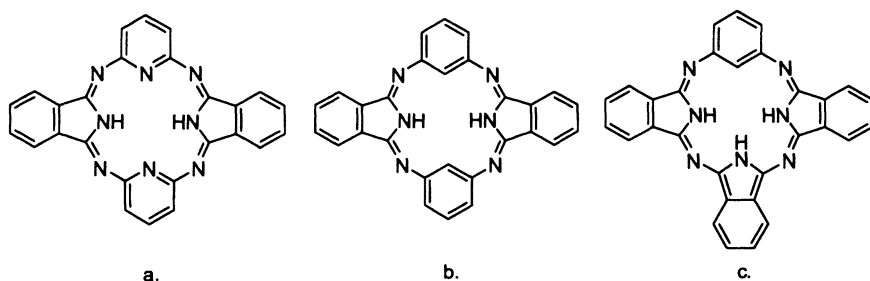


Figure 2. The hemiporphyrazines: bis-pyridyl hemiporphyrazine (a); dicarbahemiporphyrazine (b); and benziphthalocyanine (c).

The presence of a carbon in the core of the ring indicates that these macrocycles will form organometallic interactions upon metal binding. In biological systems, naturally occurring organometallic complexes are rare. The cobalamin cofactor is one of the more important examples, which possesses an axial metal-carbon bond in a cobalt porphyrinoid macrocycle (17). Another common example is seen in carbon-monoxide deactivated ferrous heme (22);



the oxygen storage and transport proteins hemoglobin and myoglobin are structured to minimize the binding of this toxic species at axial site. More recently, organometallic bonding has been observed several non-heme active site, such as the iron based hydrogenases (23).

Considering the above, why not a carbaporphyrin cofactor? As shall be shown below, the simple replacement of a carbon for a nitrogen greatly affects the properties of the macrocycle. For example, these changes in metal binding properties are readily observed in the isomer *N*-confused porphyrin, in which the modification to the ring structure can be thought of as a minimal change: simple swapping of whenever a carbon is placed at a core position, regardless of the nature of the ring structure.

### The Acidity of the Internal Carbon

In normal porphyrins, the macrocycle is readily deprotonated, affording a dianionic macrocycle with a pore size that can readily accommodate many transition metal ions. In contrast, the internal C-H bonds of the carbaporphyrins are not acidic, and are not removed even by relatively strong bases such as alkyl lithium reagents. The lack of acidity in the carbaporphyrins was first observed by both Furuta and Latos-Grazynski in *N*-confused porphyrin (18, 19). Although the external nitrogen site in this macrocycle can be readily deprotonated or protonated, the core C-H bond remains inert to bases. Later, similar observations were made in *m*-benzporphyrin and in azuliporphyrin (20, 21). More recently, in our group, we have isolated the lithium adduct of benziphthalocyanine (24). This compound shows deprotonation of the core of the macrocycle with the exception of the internal C-H bond, even though the complex is generated via reaction with the highly basic reagent lithium bis(trimethylsilyl)amide.

Due to this lack of acidity, the internal carbon often does not activate upon metalation with many transition metals. This results in a novel form of porphyrin metal binding where the internal C-H bond(s) remain intact but lie close to the bound metal. Several metals exhibit this type of bonding, including manganese (25-27), rhodium (28), iron (29), cobalt (30), zinc (31), palladium (32), silver (33), cadmium, and mercury (34). The C-H bond can reside very close to the metal ion, often within 2.5 Å. As a result of this proximity, this interaction has been referred to as "agostic," however, the key measurement that would define the agostic interaction, the  $^{13}\text{C}$ - $^1\text{H}$  coupling constant, has yet to be measured.

In *N*-confused porphyrins, this "agostic" coordination mode has been observed frequently. To date, monomeric complexes of this macrocycle with divalent manganese (35, 26), divalent iron (36, 29), lanthanides (37), rhodium carbonyl (38), and group 12 elements (34), have been structurally characterized that exhibit this peculiar type of metal binding. A generic structure for several of

these compounds and the monomeric  $\text{Mn}^{\text{II}}(\text{NCP})\text{py}$  complex are shown in Figure 3. In manganese, metalation with both manganese carbonyls and manganese bromide results in Mn(II) complexes that retain the internal C-H bond. An Fe(II) complex can be generated by metalation with iron bromide, and the structure is directly analogous to the Mn(II) bromide species, including a metal that resides above the plane of the macrocycle and a tilted pyrrole. In addition to the axial bromide, a Fe(II) thiolate species with an intact internal C-H bond can be generated by metathesis of the bromide. A monomeric zinc complex can be prepared by recrystallization of zinc *N*-confused porphyrin metalation products from pyridine. In all of these monomeric complexes, the axial ligand is tilted toward the inverted pyrrole, and the metal resides significantly above the plane of the three pyrrolic nitrogens.

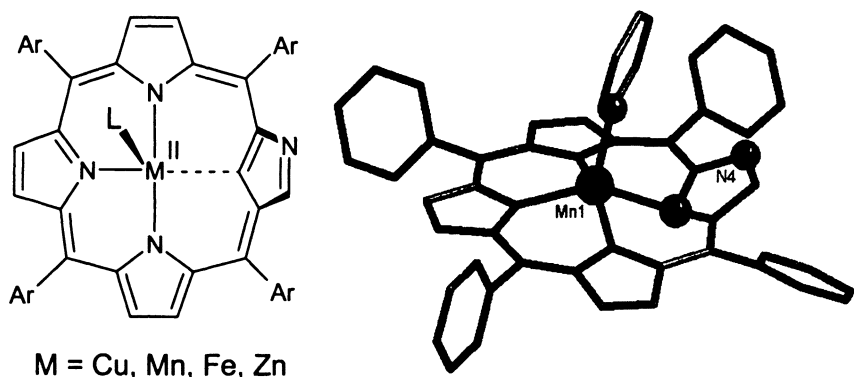


Figure 3. Monomeric *N*-confused porphyrin first row metal complexes that exhibit an agostic-type interaction (left) and structure of  $\text{Mn}^{\text{II}}(\text{NCP})\text{py}$  (right).

In the absence of coordinating species, *N*-confused porphyrins with intact C-H bonds in the core form dimers via self-coordination through the external nitrogen position. Dimeric compounds have been characterized by X-ray methods for Mn(II) (26), Fe(II) (35), Zn(II) (34, 39), Pd(II) (32), and Cd(II) complexes (34). These particular binding mode can form due to the tilt in the pyrrole which allows the external nitrogen to coordinate to the axial position of the neighboring macrocycle. The coordination of this macrocycle at this position results in a slip stacked dimeric structure. These complexes are typically quite sterically congested; in the Mn(II) and Fe(II) dimers, four of the meso phenyl rings are nearly co-planar with the *N*-confused porphyrin ring. In the cases of the Zn(II) and Cd(II) complexes, the dimers were isolated only upon use of the diphenyl *N*-confused porphyrin macrocycles, which have significantly reduced steric bulk.

The retention of the core hydrogens on  $sp^2$  hybridized carbons has been observed in other carbaporphyrins. In the *m*-benzporphyrin macrocycle, adducts with Fe(III) (40), Ni(II), Zn(II), Cd(II) and Hg(II) (41) results in complexes that exhibit the same type of agostic interaction with the internal C-H bond of the benzene ring. The Fe(III), Ni(II) and Cd(II) complexes have been structurally characterized; the tilts of the benzene ring are larger than that seen in the analogous *N*-confused porphyrin complexes, and the M-C distances are longer as well ( $>2.5$  Å). Similar chemistry is observed in the *p*-benzporphyrin macrocycle, and the Ni(II) and Cd(II) adducts of this ring have been elucidated by X-ray crystallography (42).

Recently we have seen identical chemistry in the carbahemiporphyrazines. As can be seen in Figure 4, we have observed this chemistry with the metals manganese, iron, cobalt, copper and silver with the dicarbahemiporphyrazine macrocycle (30, 33). With the exception of copper, which will be described below, both internal C-H bonds are retained upon metalation. The lack of C-H activation in these complexes distorts the macrocycle in these compounds from planarity. Since there are two intact C-H bonds at the core, the planar distortion is much greater than that seen in *N*-confused porphyrin. The C-M distances are longer than those seen in *N*-confused porphyrins ( $> 2.4$  Å), and, based on our  $^1\text{H}$  NMR measurement of the Ag(I) complex, the electronic structure of the C-H bond is not greatly perturbed when compared to the free base macrocycle (6.77 ppm in the silver adduct versus 6.91 ppm).

In the above complexes, the internal C-H bond can frequently be activated oxidatively to form the fully deprotonated core. In most cases, there has to be a redox active metal ion in the core of the ring or there has to be an external oxidant available such as atmospheric dioxygen. In the case of Mn(II) *N*-confused porphyrins, exposure to atmospheric dioxygen results in the formation of the higher valent Mn(III) macrocycles with a direct Mn-C bond (26). The  $\text{Mn}^{\text{III}}(\text{NCP})\text{Br}$  is shown in Figure 5. In a similar fashion, the  $\text{Fe}^{\text{II}}(\text{NCTPP})\text{Br}$  complex can react with dioxygen to afford the analogous Fe(III) complex where the internal carbon binds directly to the metal (43). Alternatively, the use of electron withdrawing groups at the periphery of *N*-confused porphyrin, such as perfluorophenyl rings, can help activate the internal C-H bond. This can be seen with copper binding; the complex has a direct Cu-C bond upon metal insertion (44). However, it is unclear whether the activation results from the increased acidity of the internal carbon, or from the change in redox potential of the ring.

With other metals, such as nickel (19), palladium (32), platinum (45), antimony (46), and occasionally silver (47) and cobalt (48), the internal C-H bond is broken upon initial metalation. In the case of cobalt and silver *N*-confused porphyrins, the fully deprotonated adducts are always observed upon metalation, which are shown in Figure 5. In the case of these two metals, the macrocycle takes on different charges: trianionic in the case of Ag(III) and dianionic in the case of the Co(III)  $\text{PPh}_3$  adduct. *N*-confused porphyrin can adopt multiple charges, as will be described below. In most of these complexes

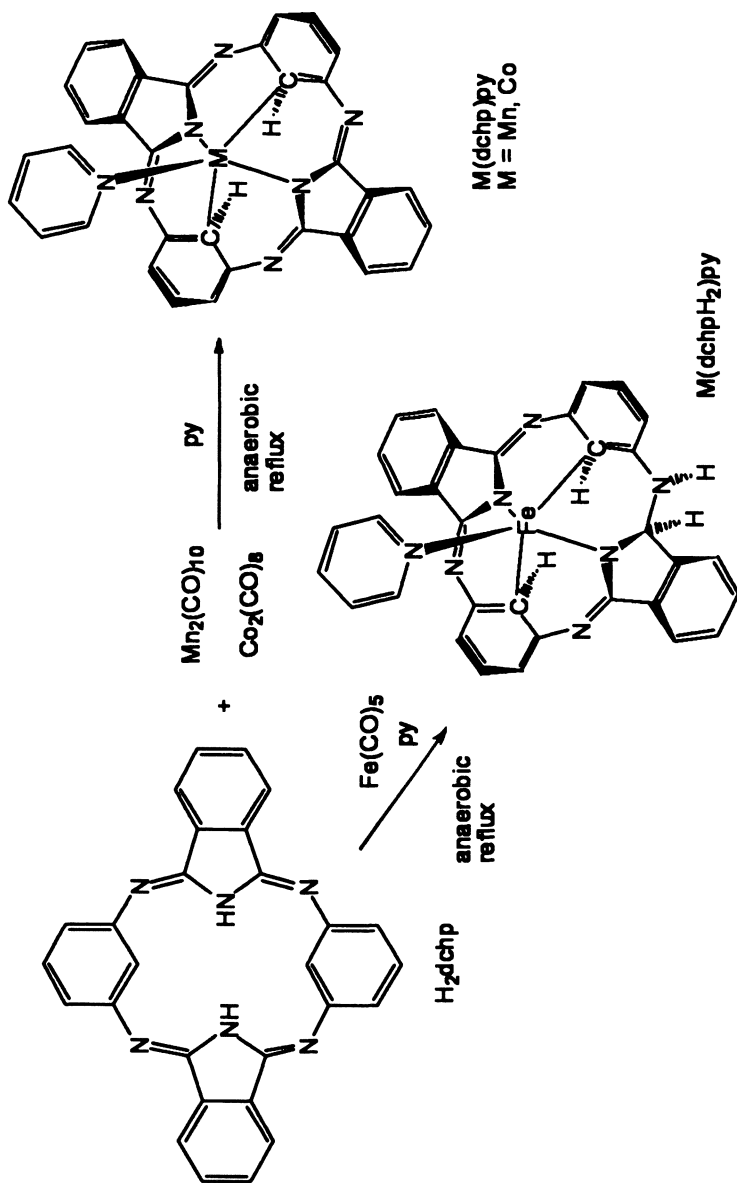


Figure 4. Examples of the metalation chemistry of dicarbahemiporphyrazine. These metal complexes retain the internal C-H bonds.

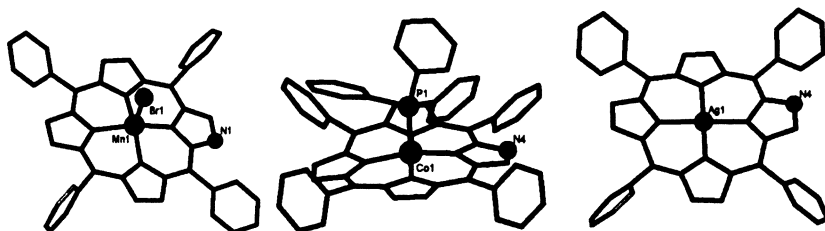


Figure 5. *N*-confused porphyrin metal complexes with direct M-C bonds.

in *N*-confused porphyrin, the C-M bond is indistinguishable from the C-N bond; this is readily observed in the disorder present in the structures of many of these complexes. In these macrocycles, these lengths range from 1.9 to 2.1 Å, which is in the same range as the M-N bonds.

There are also several examples of other carbaporphyrins that form direct C-M bonds upon metalation. The *m*-benzporphyrin macrocycle has been shown to form such bonds with Pd(II), Pt(II) and also Ni(II) (49), which can have the internal C-H bond activated upon metalation using a different methodology from that described above using NiCl<sub>2</sub>. Internal C-M bond formation has been observed with insertions various metal ions into a number of carbaporphyrinoids. Examples include the adducts of Ni(II) into dimethoxybenzporphyrin (50), Pd(II) into oxybenzporphyrin (51), Ni(II), Pd(II) and Pt(II) into azuliporphyrin (52, 53), Pd(II) into oxybenzocarbaporphyrin (54), and Ag(III) into benzocarbaporphyrin (55), oxybenzporphyrin (56), tropiporphyrin (57), *O*-confused porphyrin (58), and carbaporpholactone (59). In our recent work we have observed the C-H bond activation with the insertion of both Co(II) and Co(III) into benziphthalocyanine (60).

This discrepancy in metal binding structures is in agreement with the lack of acidity. Instead of deprotonation being the mechanism by which a direct metal carbon bond is formed, it is clear that oxidative addition is the preferred route. With regard to the use of carbaporphyrins by biological systems, the lack of C-H bond activation upon initial metalation with iron would certainly modify the biosynthesis of carbaporphyrin analogs of hemes. However, the lack of acidity of the internal carbon position is not the only complication associated with carbaporphyrins; as will be described in the next section, the reactivity of the internal carbon is significantly different from that of normal tetrapyrroles.

## Reactivity of the Internal Carbon

Although in many cases the internal carbon-hydrogen bond remains intact upon metalation, the close proximity of this group to the metal ion does sensitize this position for reaction. In addition, even when a direct M-C bond is formed, the carbon is reactive. In *N*-confused porphyrins, benziporphyrins and

carbahemiporphyrazines, these macrocycles show reactivity at the internal carbon position upon metalation (61). Specifically, this site becomes sensitive to both electrophilic and nucleophilic attack. The reactivity to electrophiles is due to the carbanionic nature of the core carbon, while nucleophiles can attack the internal position often through the reductive elimination from the metal.

Even in the free base, the internal C-H group of the carbaporphyrin *N*-confused porphyrin is highly reactive. Exposure of the free base to halogenating reagents such as elemental bromine or *N*-bromosuccinamide results in the formation of a C-X bond in the core of the macrocycle (62). This reaction results in inversion of the pyrrole ring, which leads to formation of *N*-fused porphyrin. Susceptibility to electrophilic attack in a metalated complex was first observed with the Ni(II) complex of *N*-confused porphyrin. This was noted in the reaction of simple alkyl halides, such as CH<sub>3</sub>I, which alkylates at the 21-C position as well as the expected external nitrogen (63). Dolphin and coworkers also noted that this position can also be substituted with a cyano group derived from an equivalent of dichlorodicyanoquinone (DDQ) (64).

The Dolphin group also observed that in Ni<sup>II</sup>NCP the 21-C could be oxidized with OsO<sub>4</sub> to produce an alkoxide at the internal carbon position (65). Structural analysis of the complex, which has a Ni(III) ion at the core, shows that the internal carbon has some sp<sup>3</sup> character. Such oxidations can be extremely common for *N*-confused porphyrins. In the Fe<sup>III</sup>(NCP)Br complex, the Fe-C bond can be attacked by molecular dioxygen, resulting in a ketone-type carbon-oxygen bond at the core of the ring (40). Such oxidation reactions can lead to the bleaching of *N*-confused porphyrin during aerobic metalation reactions. This has been noted by Furuta and coworkers, who observed the oxidation of *N*-confused porphyrin to a tripyrrone during the aerobic metalation of the freebase macrocycle with Cu(II) (66).

Other carbaporphyrins also exhibit reactivity at the internal benzene carbon position. The attempted metalation of *m*-benziporphyrin with silver acetate results in acetoxylation of the internal carbon atom but not the insertion of the metal ion (67). The reaction of tetraphenyl-*m*-benziporphyrin with CuCl<sub>2</sub> in THF at room temperature resulted in a tetranuclear compound (40). In this complex, where a Cu<sub>2</sub>Cl<sub>2</sub> core bridges two metalated *N*-confused porphyrins, the core carbon is substituted with a chloride, replacing the hydrogen atom. The Pd(II) complex of oxybenziporphyrin undergoes alkylation at the internal hydrogen position upon reaction with alkyl halides (51). When azuliporphyrin is metalated with copper acetate, the internal oxygen site is oxidated to form the acetoxy species in a reaction directly analogous to that seen in nickel and iron *N*-confused porphyrin (68). The reaction of benzocarbaporphyrin with FeCl<sub>3</sub> in CHCl<sub>3</sub>-MeOH solvent system result in a regioselective oxidation reaction at the inner carbon, forming a dimethyl ketal derivative (69) which can be extended to other alcohols such as isopropyl alcohol and ethylene glycol. Elongated reaction

times result in the halogenation of the internal carbon atom and the formation of a nonaromatic dione species.

Nucleophiles can also substitute at the internal carbon position. Reaction of *m*-benzporphyrin with silver tetrafluoroborate in pyridine under reflux conditions yielded a pyridine-appended macrocycle, with a C-N bond at the core of the ring. The pyridine then forms a second bond at the ortho position with an adjacent meso carbon (70). Recently, we noticed a directly analogous reaction upon the metalation of dicarbahemiporphyrazine with Cu(II) in pyridine (33). In this macrocycle, one of the two core carbon positions is substituted with a pyridine. In both cases, the formation of the C-N bond is believed to result from reductive elimination at the metal. Initial metalation results in the formation of a M-C bond, and the metal is subsequently coordinated by a solvent pyridine. The activation of the C-H bond at the core results in a high valent metal center (Ag(III) and Cu(III) respectively) which can then reductively eliminate to form a monovalent metal center and a C-N bond.

The sensitivity of the internal carbon to reaction would be a significant complicating factor for carbaporphyrins in protein active sites. Dioxygen binding and activation, for example, would readily result in oxidation at the core carbon position. In addition, the susceptibility of the internal carbon to nucleophiles could affect the coordination chemistry of the side chains of cysteine and histidine, frequently bound at the axial positions of heme sites. While novel inorganic structures do result from adducts formed upon substitution at the internal carbon, directly analogous chemistry to that seen in normal porphyrins is much less likely due to its susceptibility.

## The Electronic Effects of Organometallic Bonding

Another aspect of the organometallic bonding observed in the metal adducts of the carbaporphyrins is the increase in metal-ligand covalency upon the introduction of carbons at the core. The decreased electronegativity of carbon versus nitrogen results in a less polar bond; this covalency can stabilize different electronic states than seen in the more ionic bonding found in normal porphyrins. In addition, the incorporation of one or more covalent bonds increases the anisotropy at the metal center, through both the reduction in symmetry as well as through the overall change in binding pocket polarity.

Such effects can be readily observed in the Mn(III) adduct of *N*-confused porphyrin described above (26). The coordination geometry of *N*-confused porphyrin in this complex is identical to that seen in normal porphyrin Mn(III) complexes, with the only significant change being the inverted pyrrole ring. Several years ago, we carried out a high frequency and field EPR measurement on Mn<sup>III</sup>(NCP)py<sub>2</sub> in both polycrystalline and frozen solution, as shown in Figure 6 (71, 72). By fitting the resonance frequency to the quantum energy, we

were able to obtain a simultaneous fit to all experimental data and derive the spin Hamiltonian parameters for the complex. In regular porphyrins, Mn(III) complexes are typically axial, and their spin Hamiltonian parameters reflect the nearly zero zfs tensor, with an E value close to zero. In the *N*-confused porphyrin adduct, the metal has a high spin ground state and a highly rhombic zfs tensor, as reflected by the large ratio of  $E/D = 0.20$ . This number results not from a small E value ( $E = -0.608(3) \text{ cm}^{-1}$ ), but instead from a large negative D parameter ( $D = -3.084(3) \text{ cm}^{-1}$ ). This D parameter is a reflection of the large effect of the carbon on the zfs, most likely along the N-Mn-C vector. Thus, the inversion of a single pyrrole ring has a very significant and measureable change on the electronic structure of the metal ion in *N*-confused porphyrin.

Another example of the change in electronic structure induced by a simple inversion of a single pyrrole can be seen by comparing the Mo(II) complexes of *N*-confused porphyrin and normal porphyrin (73). Molybdenum can be inserted into *N*-confused porphyrin via reaction of the freebase with  $\text{Mo}^{\text{II}}(\text{CO})_4(\text{pip})_2$ , which results in a six coordinate Mo(II) complex with two axial piperidines (pip). The normal porphyrin analog can be generated in an identical manner, and the two complexes are nearly isostructural (with the exception of the two swapped atom positions), isoelectronic, and have the metals in the same oxidation state. This allows one to observe directly the electronic effects of the pyrrole inversion and the formation of the metal-carbon bond. Their UV-visible spectra, for example, exhibit marked differences. The spectrum of the TPP complex shows features typical for a d-hyper metalloporphyrin spectrum, with a red shifted intense Soret band and much weaker Q bands. The *N*-confused porphyrin complex, however, has a much less intense Soret band and more intense Q bands, and all of the bands are red shifted relative to those of normal metalloporphyrins. In particular, there is an intense low energy absorption at 686 nm. In addition, the d orbital configurations of the molybdenums are different in the two complexes. The  $\text{Mo}^{\text{II}}(\text{TPP})(\text{pip})_2$  has a  $\mu_{\text{eff}} = 2.79$ , which corresponds to a  $S = 1$  spin state. A similar susceptibility is seen in  $\text{Mo}^{\text{II}}(\text{TPP})\text{py}_2$  ( $\mu_{\text{eff}} = 2.83$ ). In contrast, the  $\text{Mo}^{\text{II}}(\text{NCP})(\text{pip})_2$  complex has lower susceptibility ( $\mu_{\text{eff}} = 1.68$ ), most likely indicating a thermally populated  $S = 1$  spin state mixed with a diamagnetic ground state.

## Conclusions

The metal chemistry of the carbaporphyrins has been shown to be potentially as rich as that of normal porphyrins; however, the chemistry is distinctly different. Although the replacement of a nitrogen with its neighbor on the periodic table may seem to be small modification, the effect of this change is readily observable on metal binding, reactivity, and electronic structure. Notably, there are striking similarities across the carbaporphyrins, despite their varied structures.



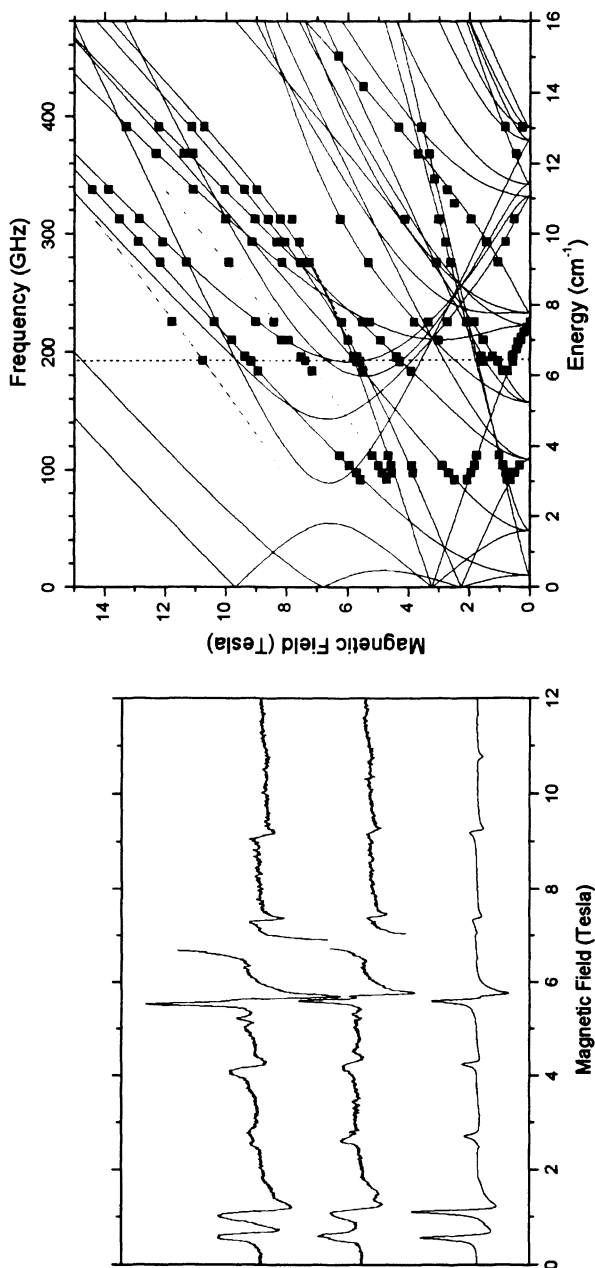


Figure 6. Experimental HFEPR spectra of polycrystalline  $\text{Mn}(\text{NCTPP})(\text{py})_2$  (Left, top trace) at 192 GHz and 10 K, and in pyridine solution in the same conditions (middle trace). The bottom trace represents a powder-pattern simulation using the following set of spin Hamiltonian parameters:  $S = 2$ ;  $D = -3.07 \text{ cm}^{-1}$ ;  $E = -0.61 \text{ cm}^{-1}$ ;  $g_{\text{iso}} = 2.00$ . Resonance field vs. quantum energy dependence of HFEPR signals in polycrystalline  $\text{Mn}(\text{NCTPP})(\text{py})_2$  at 10 K (Right) (71, 72). (Copyrights 2006 and 2007 American Chemical Society.)

With regard to the question of the biological choice of nitrogen over carbon in porphyrin, it is clear that the carbaporphyrins might not be useful cofactors for metal binding. The observed lack of activation of the internal C-H bond results in less stable metal adducts; this is clearly observed in the biologically relevant metals iron and manganese. In addition, when the internal C-H bond is activated to make a direct M-C bond, the resultant complexes exhibit very different properties versus their normal porphyrin analogs. Lastly, the internal carbon is reactive to both nucleophilic and oxophilic attack; in particular this has been observed with molecular dioxygen and the iron adducts of the carbaporphyrins. Hemes play key roles in oxygen binding and activation in biological systems, whereas iron carbaporphyrins would be reactive at the internal carbon position under similar conditions. Thus, the replacement of a metalloporphyrin for a heme would not result in a similarly functioned protein. However, this experiment has yet to be conducted; there is a good body of work on the replacement hemes with iron porphycenes in proteins, and these porphyrin isomers do affect the properties of the proteins (74-76). Future investigations into the chemistry of the carbaporphyrins will continue to highlight these differences, and hopefully uncover more novel structures and reactivity.

## References

1. *The Porphyrin Handbook*; Kadish, K. M.; Smith, K. M.; Guillard, R., Eds.; Academic Press: New York, NY, 2000; Vols. 1-20.
2. *Porphyrins and Metalloporphyrins*; Smith, K. M., Ed.; Elsevier: Amsterdam 1975.
3. Dolphin, D. *The Porphyrins*; Academic Press: New York, NY, 1979.
4. *Iron Porphyrins*; Lever, A. B. P.; Gray, H. B., Eds.; Physical Bioinorganic Chemistry; Wiley-VCH: New York, NY; 1989; Vol. 4.
5. Milgrom, L. R. *The Colours of Life: An Introduction to the Chemistry of Porphyrins and Related Compounds*; Academic Press: New York, NY, 1997.
6. Buchler, J. W. In *The Porphyrins*; Dolphin, D., Ed.; Academic Press: New York, NY, 1979; Vol. 1, pp 157-187.
7. Ogoshi, H.; Mizutani, T. *Acc. Chem. Res.* **1998**, *31*, 81-89.
8. Balaban, T. S. *Acc. Chem. Res.* **2005**, *38*, 612-623.
9. Medforth, C. J.; Senge, M. O.; Smith, K. M.; Sparks, L. D.; Shelnut, J. A. *J. Am. Chem. Soc.* **1992**, *114*, 9859-9869.
10. Sessler, J. L.; Weghorn, S. J. *Expanded, Contracted and Isomeric Porphyrins*; Pergamon: New York, NY, 1997.
11. Matano, Y.; Nakashima, M.; Nakabuchi, T.; Imahori, H.; Fujishige, S.; Nakano, H. *Org. Lett.* **2008**, *10*, 553-556.

12. Latos-Grażyński, L. In *The Porphyrin Handbook*; Kadish, K. M.; Smith, K. M.; Guillard, R., Eds.; Academic Press: New York, NY, 2000; Vol. 2., pp 361-416.
13. Broadhurst, M. J; Grigg, R. Johnson, A. W. *J. Chem. Soc. (C)* **1971**, 3681-3690.
14. Ulman, A.; Manassen, J. *J. Am. Chem. Soc.* **1975**, *97*, 6540-6544.
15. Lash, T. D. *Eur. J. Inorg. Chem.* **2007**, *46*, 5461-5481.
16. Galezowski, M.; Gryko, D. T. *Curr. Org. Chem.* **2007**, *11*, 1310-1338.
17. Brown, K. L. *Chem. Rev.* **2005**, *105*, 2075-2149.
18. Furuta, H.; Asano, T.; Ogawa, T.; *J. Am. Chem. Soc.* **1994**, *116*, 767-768.
19. Chmielewski, P. J.; Latos-Grażyński, L.; Rachlewicz, K.; Glowiak, T. *Angew. Chem. Int. Ed.* **1994**, *33*, 779-781.
20. Stępień, M.; Latos-Grażyński, L. *Acc. Chem. Res.* **2005**, *38*, 88-98.
21. Lash, T. D.; Chaney, S. T. *Angew. Chem. Int. Ed.* **1997**, *36*, 839-840.
22. Peng, S.-M.; Ibers, J. A. *J. Am. Chem. Soc.* **1976**, *98*, 8032-8036.
23. Schmidt, M.; Contakes, S. M.; Rauchfuss, T. B. *J. Am. Chem. Soc.* **1999**, *121*, 9736-9737.
24. Sripothongnak S.; Ziegler, C. J. *Inorg. Chem.* submitted.
25. Bohle, D. S.; Chen, W.-C.; Hung, C.-H. *Inorg. Chem.* **2002**, *41*, 3334-3336.
26. Harvey, J. D.; Ziegler, C. J. *Chem. Commun.* **2002**, 1942-1943.
27. Harvey, J. D.; Ziegler, C. J. *Chem. Commun.* **2003**, 2890-2981.
28. Toganoh, M.; Ishizuka, T.; Furuta, H. *Chem. Commun.* **2004**, 2464-2465.
29. Rachlewicz, K.; Wang, S.-L.; Peng, C.-H.; Hung, C.-H.; Latos-Grażyński, L. *Inorg. Chem.* **2003**, *42*, 7348-7350.
30. Cetin, A.; Durfee, W. S.; Ziegler, C. J. *Inorg. Chem.* **2007**, *46*, 6239-6241.
31. Furuta, H.; Ishizuka, T.; Osuka, A. *J. Am. Chem. Soc.* **2002**, *124*, 5622-5623.
32. Furuta, H.; Kibo, N.; Maeda, H.; Ishizuka, T.; Osuka, A.; Nanami, H.; Ogawa, T. *Inorg. Chem.* **2000**, *39*, 5424-5425.
33. Wu, R.; Çetin, A.; Durfee, W. S.; Ziegler, C. J. *Ang. Chem. Int. Ed.* **2006**, *45*, 5670-5673.
34. Furuta, H.; Morimoto, T.; Osuka, A. *Inorg. Chem.* **2004**, *43*, 1618-1624.
35. Hung, C.-H.; Chen, W.-C.; Lee, G.-H.; Peng, S.-M. *Chem. Commun.* **2002**, 1516-1517.
36. Chen, W.-C.; Hung, C.-H. *Inorg. Chem.* **2001**, *40*, 5070-5071.
37. Zhu, X.; Wong, W.-K.; Lo, W.-K.; Wong, W.-Y. *Chem. Commun.* **2005**, 1022.
38. Srinivasan, A.; Furuta, H.; Osuka, A. *Chem. Commun.* **2001**, 1666-1667.
39. Furuta, H.; Ishizuka, T.; Osuka, A. *J. Am. Chem. Soc.* **2002**, *124*, 5622-5623.
40. Hung, C.-H.; Chang, F.-C.; Lin, C.-Y.; Rachlewicz, K.; Stępień, M.; Latos-Grażyński, L.; Lee, G.-H.; Peng, S.-M. *Inorg. Chem.* **2004**, *43*, 4118.

41. Stepień, M.; Latos-Grażyński, L.; Szterenber L.; Panek, J.; Latajka Z. *J. Am. Chem. Soc.* **2004**, *126*, 4566-4580.
42. Stepień, M.; Latos-Grażyński, L.; Szterenber L. *Inorg. Chem.* **2004**, *43*, 6654-6662.
43. Rachlewicz, K.; Wang, S.-L.; Ko, J.-L.; Hung, C.-H.; Latos-Grażyński, L. *J. Am. Chem. Soc.* **2004**, *126*, 4420.
44. Maeda, H.; Ishikawa, Y.; Matsuda, T.; Osuka, A.; Furuta, H. *J. Am. Chem. Soc.* **2003**, *125*, 11822-11823.
45. Furuta H.; Youfu, K.; Maeda H.; Ishizuka T.; Osuka A. *Angew. Chem. Int. Ed.* **2003**, *42*, 2186-2188.
46. Ogawa, T.; Furuta, H.; Takahashi, M.; Morino, A.; Uno, H. *J. Organomet. Chem.* **2000**, *611*, 551-557.
47. Furuta, H.; Ogawa, T.; Uwatoko, Y.; Araki, K. *Inorg. Chem.* **1999**, *38*, 2676-2682.
48. Harvey, J. D.; Ziegler, C. J. *Chem. Commun.* **2004**, 1666-1667.
49. Stepień, M.; Latos-Grażyński, L. *Chem. Eur. J.* **2001**, *7*, 5113-5117.
50. Richter, D. T.; Lash T. D. *Tetrahedron*, **2001**, *57*, 3657-3671.
51. Stepień, M.; Latos-Grażyński, L.; Lash, T. D.; Szterenber, L. *Inorg. Chem.* **2001**, *40*, 6892-6900.
52. Graham, S. R.; Ferrence, G. M.; Lash, T. D. *Chem. Commun.* **2002**, 894-895.
53. Lash, T. D.; Colby, D. A.; Graham, S. R.; Ferrence, G. M.; Szczepura, L. F. *Inorg. Chem.* **2003**, *42*, 7326-7338.
54. Venkatraman, S.; Anannd, V. G.; Pushpan, S. K.; Sankar, J.; Chandrashekar, T. K. *Chem. Commun.* **2002**, 462-463.
55. Muckey, M. A.; Szczepura, L. F.; Ferrence, G. M.; Lash, T. D. *Inorg. Chem.* **2002**, *41*, 4840-4842.
56. Lash, T. D.; Rasmussen, J. M.; Bergman, K. M.; Colby, D. A. *Org. Lett.* **2004**, *6*, 559-552.
57. Bergman, K. M.; Ferrence G. M.; Lash, T. D. *J. Org. Chem.* **2004**, *69*, 7888-7897.
58. Pawlicki, M.; Latos-Grażyński, L. *Chem.-Eur. J.* **2003**, *9*, 4650-4660.
59. Pawlicki, M.; Latos-Grażyński, L. *J. Org. Chem.* **2005**, *70*, 9123-9130.
60. Cetin, A.; Sriphothongnak, S.; Wu, R. Durfee, W. S.; Ziegler, C. J. *Chem. Commun.* **2007**, 4289-4290.
61. Chmielewski, P. J.; Latos-Grażyński, L. *Coord. Chem. Rev.* **2005**, *249* 2510-2533.
62. Furuta, H.; Ishizuka, T.; Osuka, A.; Ogawa, T. *J. Am. Chem. Soc.* **2000**, *122*, 5748-5757.
63. Chmielewski, P. J.; Latos-Grażyński, L.; Głowiak, T. *J. Am. Chem. Soc.* **1996**, *118*, 5690-5701.
64. Xiao, Z.; Patrick, B. O.; Dolphin, D. *Chem. Commun.* **2003**, 1062-1063.
65. Xiao, Z.; Patrick, B. O.; Dolphin, D. *Inorg. Chem.* **2003**, *42*, 8125-8127.

66. Furuta, H.; Maeda, H.; Osuka, A. *Org. Lett.* **2002**, *4*, 181-184.
67. Stępień, M.; Latos-Grażyński, L. *Inorg. Chem.* **2003**, *42*, 6183-6193.
68. Colby, D. A.; Ferrence G. M.; Lash, T. D. *Angew. Chem. Int. Ed.* **2004**, *43*, 1346-1349.
69. Hayes, M. J.; Spence, J. D.; Lash, T. D. *Chem. Commun.* **1998**, 2409-2410.
70. Stępień, M.; Latos-Grażyński, L. *Org. Lett.* **2003**, *5*, 3379-3381.
71. Harvey, J. D.; Ziegler, C. J.; Telser, J.; Ozarowski, A.; Krzystek J. *Inorg. Chem.* **2005**, *44*, 4451-4453.
72. Harvey, J. D.; Ziegler, C. J.; Telser, J.; Ozarowski, A.; Krzystek, J. *Inorg. Chem.* **2006**, *45*, 8459.
73. Harvey, J. D.; Shaw, J. L.; Herrick, R. S.; Ziegler, C. J. *Chem. Commun.* **2005**, 4663-4665.
74. Hayashi, T.; Dejima, H.; Matsuo, T.; Sato, H.; Murata, D.; Hisaeda, Y. *J. Am. Chem. Soc.* **2002**, *124*, 11226-1127.
75. Matsuo, T.; Tsuruta, T.; Maehara, K.; Sato, H.; Hisaeda, Y.; Hayashi, T. *Inorg. Chem.* **2005**, *44*, 9391-9396.
76. Hayashi, T.; Murata, D.; Makino, M.; Sugimoto, H.; Matsuo, T.; Sato, H.; Shiro, Y.; Hisaeda, Y. *Inorg. Chem.* **2006**, *45*, 10530-10536.



## Chapter 9

# Bioinspired Aerobic Substrate Oxidation

## A Ni(II)-Oximate Catalyst that Parallels Biological Alcohol and Amine Oxidation Chemistry

Michael J. Baldwin<sup>\*</sup>, Jeanette A. Krause, Michael J. Goldcamp<sup>1</sup>,  
Micheal Haven<sup>2</sup>, Sara E. Edison<sup>3</sup>, and Leah N. Squires<sup>4</sup>

Department of Chemistry,  
University of Cincinnati, Cincinnati, OH 45221  
Current addresses: <sup>1</sup>Department of Chemistry, Wilmington College,  
Wilmington, OH 45177; <sup>2</sup>Intertek Caleb Brett, Detroit, MI 48209;  
<sup>3</sup>Hexion Specialty Chemicals, Cincinnati, OH 45245;  
<sup>4</sup>Idaho National Laboratory, Idaho Falls, ID 83401

The Ni(II) complex of tris(1-propan-2-onyl oximate)amine (TRISOX) catalyzes the aerobic oxidation of primary alcohols and amines. Parallels to several metalloenzymes guide studies of Ni(TRISOX) chemistry. Ni(TRISOX) catalyzes the same net reactions as galactose oxidase, Cu-dependent amine oxidase, and catalase. Spectroscopy of an observable catalytic intermediate shows reversible oxidation of ligand oximates to iminoxyl radicals. Galactose oxidase and amine oxidase also require both a redox-active organic component and a transition metal. Crystal structures of H<sub>n</sub>TRISOX ligands with Fe, Co, Cu and Zn extend this chemistry to other metals.

In the *biomimetic chemistry* approach to bioinorganic modeling chemistry, well-defined, “small molecule”, metal complexes that reproduce aspects of the structure and/or function of a metalloprotein are synthesized. These biomimetic complexes generally include the same metal in a similar coordination

environment to that of the protein active site. The goal is either to provide insight into how the biological system works through structural, spectroscopic or reactivity studies, or to produce synthetic metal complexes designed to accomplish useful chemistry. The *bioinspired chemistry* approach uses biological systems as conceptual models for the design and mechanistic understanding of a useful synthetic complex, without trying to produce a close structural mimic of the active site. This approach attempts to produce chemistry that parallels that of the biological system, without being constrained to the same metal and coordinating functional groups that are biologically available. Examples of bioinspired chemistry are found in fields as diverse as materials (1-3), organometallic chemistry (4) and molecular recognition (5, 6). As an example of applying this bioinspired chemistry approach to catalysis, we describe a complex that promotes the same chemistry as several enzymes that catalyze the oxidation of alcohol and amine substrates by the net transfer of two hydrogen atoms to O<sub>2</sub>. Like the enzymes, the complex includes a *metal-organic redox hybrid*, that is, the mechanistically important involvement of both a redox active metal and a redox-active organic component. Neither the metal (nickel) nor the coordinating functional groups (oximates) in the synthetic complex are among those found in the related enzymes. Yet, similarities between the chemistry of the synthetic and biological molecules guide our understanding of how the small molecule catalyst works.

### Choice of Synthetic System: Ni(II)(TRISOX)

Aerobic oxidation reactions are a good choice for applying bioinspired chemistry. Oxygen is inexpensive, readily available, and produces no environmentally damaging waste products. A wide variety of aerobic substrate oxidations are efficiently catalyzed by metalloenzymes. However, the biomimetic approach has produced few complexes that catalyze aerobic oxidation reactions in a commercially useful manner (7). To complement the numerous biomimetic complexes synthesized with copper or iron (the most common metals in biological oxygen chemistry), the system chosen for our work uses a metal in an oxidation state that does not typically react with oxygen. The rationale for this approach is that if a generally "oxygen-inactive" metal can be made to react with O<sub>2</sub>, then this unusual reaction will promote unusual substrate oxidations and provide new directions for catalyst development.



## Choice of Metal

Reactions of Ni(II), the metal chosen for this work, with O<sub>2</sub> are uncommon but not unprecedented. Ni(II) complexes with amidates and thiolates, both biologically available ligands, have been reported to react with O<sub>2</sub>. In thiolate complexes, oxygen preferentially attacks the ligand sulfur rather than the metal, resulting in irreversible oxygenation of the thiolate and an oxygen-inactive product (8, 9). Margerum showed that Ni(II)-peptide complexes react with O<sub>2</sub>, resulting in oxidative ligand degradation (10, 11). Later, Kimura developed Ni(II)-amidate complexes that react with O<sub>2</sub> (12). The main outcome of these reactions was shown to be ligand oxidation (13) driven by H-atom abstraction from the ligand (14). As with the thiolate complexes, this leaves a complex that does not react with oxygen.

Complexes of nickel in its Ni(I) oxidation state have been demonstrated by Riordan and coworkers to have a rich chemistry with O<sub>2</sub> (15, 16). Also, some Ni(II) complexes have been shown to react with H<sub>2</sub>O<sub>2</sub> (17, 18). Both of these types of reactions involve either the nickel or oxygen component in a reduced oxidation state relative to Ni(II)-catalyzed substrate oxidation by O<sub>2</sub>.

There are Ni-dependent enzymes involved in oxygen-related chemistry. In the Ni-dependent acireductone dioxygenase (ARD), which catalyzes the oxygenation of acireductone in a shunt out of the methionine salvage pathway (19, 20), the Ni(II) likely acts as a Lewis acid that activates the substrate rather than directly interacting with oxygen (21). Ni-dependent superoxide dismutases (NiSODs) catalyze the disproportionation of superoxide to hydrogen peroxide and O<sub>2</sub> (22, 23), but are not involved in organic substrate oxidation chemistry.

## Choice of Ligand

The ligand design was guided by the following criteria:

- It incorporates functional groups that provide access to high oxidation states of nickel.
- The ligand architecture does not include hydrogen atoms in positions that are prone to oxidative abstraction (24).
- The ligand chelates the metal in a manner that leaves two exchangeable coordination positions that are *cis* to each other.
- Substitution of the coordinating functional groups is synthetically facile so that the reactivity of the complex can be optimized.

Holm and coworkers found that thiolates, amidates and oximates were most effective at making higher nickel oxidation states accessible (25). However, the studies noted in the previous section showed that in O<sub>2</sub> reactions of their Ni(II) complexes, amidates and thiolates were themselves readily oxidized in a manner that precludes catalysis. Oximates have been shown to support stable nickel complexes in both the Ni(III) and Ni(IV) oxidation states (26-28). They also lack the  $\alpha$ - and  $\beta$ -hydrogen atoms that are most prone to oxidative abstraction by high valent metals, thereby additionally satisfying the second design criterion.

In order to provide *cis* exchangeable coordination positions, the common "tripodal amine" ligand motif was used. The *cis* relationship of the available coordination positions is potentially important in oxygen activation chemistry, allowing adjacent coordination of oxygen and substrate, or side-on ( $\eta^2$ ) coordination of O<sub>2</sub>. This ligand arrangement is especially important for Ni(II), as the alternative *trans* arrangement of the open coordination positions would result in a square planar geometry that is resistant to rapid ligand exchange. The tripodal amine offers facile substitution of one or more oximate groups by replacement of one or more of the branches off of the central tertiary amine. A potential drawback of this motif is the presence of hydrogens that are  $\beta$  to the metal-amine bond [*vide supra*], but the arrangement of the chelate orients these hydrogen atoms away from the metal, protecting them from abstraction.

The synthesis of the tris(oxime) tripodal amine compound, tris(1-propan-2-onyl oxime)amine, or H<sub>3</sub>TRISOX, was reported over a century ago by Matthaiopoulos (29), and again several decades ago (30). However, prior to our work, there had been no reported x-ray crystal structure of either the compound itself (31) or any metal complex with it. The neutral ligand and its arrangement upon chelating Ni(II) are shown in Figure 1. It is upon deprotonation of the oxime groups to form the tris(oximate) ligand, TRISOX, that this ligand is expected to provide access to higher nickel oxidation states and O<sub>2</sub> chemistry.

## Reactivity and the "Redox Hybrid" Biological Parallel

### Oxygen Reactivity and Substrate Oxidations

Nickel(II) complexes with the neutral tris(oxime) ligand, H<sub>3</sub>TRISOX, do not react with oxygen. As expected, their only reactivity is ligand exchange at the two available *cis* coordination positions. We have reported Ni(II)(H<sub>3</sub>TRISOX) complexes with a variety of ligands in these two *cis* positions (32-34). Addition of one equivalent of base per Ni(H<sub>3</sub>TRISOX) results in deprotonation of one oxime on the ligand, producing bis(oximate)-bridged dimers like the one in Figure 2 (32). These are also not oxygen-active. However, with deprotonation of the remaining oxime groups the Ni(II)(TRISOX) complex rapidly changes color

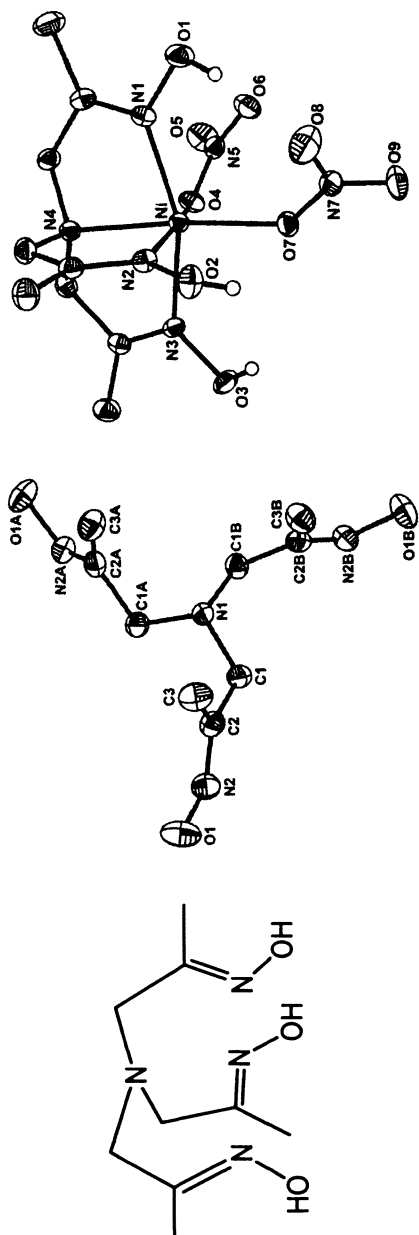


Figure 1. Line drawing of H<sub>3</sub>TRISOX (left); crystal structure of H<sub>3</sub>TRISOX (Data are from reference 31) (center); and crystal structure of Ni(H<sub>3</sub>TRISOX)(NO<sub>3</sub>)<sub>2</sub> (right).

from purple to brown upon exposure to air. While this “fully deprotonated” complex has not been structurally characterized, it is likely a  $\text{Ni}_2(\text{TRISOX})_2$  dimer as well (35).

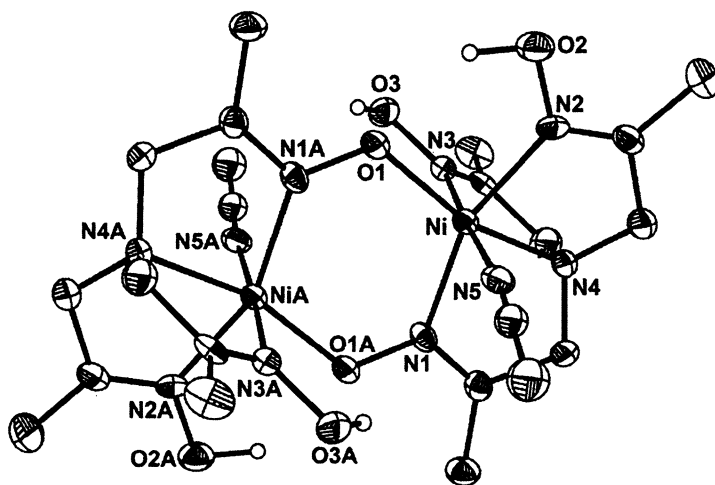


Figure 2. Crystal structure of  $[\text{Ni}(\text{H}_2\text{TRISOX})(\text{CH}_3\text{CN})]_2(\text{ClO}_4)_2$ . (Reproduced with permission from reference 32. Copyright 2002 American Chemical Society.)

The oxygen-promoted color change only occurs in the presence of appropriate aerobic oxidation substrates (36). The first substrate found for this reaction was methanol. Few examples of homogeneous catalysts for aerobic oxidation of methanol are reported in the literature (37, 38). Other primary alcohols, including benzyl alcohol and ethanol, are also substrates for the  $\text{Ni}(\text{TRISOX})$  catalyst, while secondary alcohols are not affected (36). The product in each case is the corresponding aldehyde (*i.e.*, formaldehyde from methanol) with no carboxylic acid detected. Some amines are also substrates for this reaction; benzylamine is oxidized to the imine, resulting in the Schiff base-like product benzylidenebenzylamine in the presence of excess benzylamine (36). Thus, the reaction involves the formal 2-electron oxidation of the alcohol or amine substrate by loss of two hydrogen atoms.

After several turnovers, the substrate oxidation slows substantially, with no more than a dozen turnovers observed over 24 hours (36). Thus, some percentage of the catalyst undergoes a side reaction during each catalytic cycle leading to an unreactive complex. Work is currently underway to identify this side reaction and circumvent it. However, it has been found that as the hydrogen atom abstraction oxidation chemistry slows, oxygen atom transfer to  $\text{PPh}_3$ , which

is not observed during the first hour of oxygen reactivity, begins to take place (35). Approximately two equivalents of the phosphine oxide are produced per nickel, with isotopic labeling studies showing that the inserted oxygen is derived from O<sub>2</sub> (32).

### Parallel Reactions in Metalloenzymes

The selective aerobic oxidation of primary alcohols to aldehydes, but not secondary alcohols to ketones, is reminiscent of the chemistry catalyzed by the Cu-dependent enzyme, galactose oxidase (39). Similarly, the Cu-binding  $\beta$ -amyloid protein relevant to Alzheimer's disease promotes aerobic oxidation of cholesterol, a primary alcohol (cholesterol oxidase activity) (40). The Cu-dependent amine oxidases catalyze the aerobic oxidation of amines to aldehydes (41), the hydration products of imines. Each of these enzymes that promotes aerobic oxidation of primary alcohols and amines to the same products as Ni(TRISOX) catalyze the net reaction in Equation 1. If the net reactions



promoted by the Ni(TRISOX) system are analogous to these biological examples, then hydrogen peroxide is expected to be produced as a coproduct in the nickel-catalyzed reactions as well. Colorimetric tests for H<sub>2</sub>O<sub>2</sub> using a titanium sulfate reagent gave negative results. However, careful determination of the amount of formaldehyde produced and net O<sub>2</sub> consumed gave a 2:1 ratio (36). This is consistent with the reaction in Eq. 1 *and* a negative peroxide test *if* the hydrogen peroxide is rapidly disproportionated.

Guided by the hypothesis that the Ni(TRISOX) reactions are analogous to the metalloenzyme reactions despite no H<sub>2</sub>O<sub>2</sub> accumulation, the nickel complex was tested for catalase-like activity (H<sub>2</sub>O<sub>2</sub> disproportionation). The Ni(TRISOX) complex was indeed able to disproportionate 400 equivalents of H<sub>2</sub>O<sub>2</sub> in fewer than 15 minutes, far faster than it would be produced in the alcohol or amine oxidation reactions (36). The substrate requirements can further be refined based on the net reaction for substrate oxidation in Eq. 1. The thermodynamic requirement is that the sum of the bond dissociation energies for the two hydrogens abstracted from the substrate must be less than the sum of the two O-H bond dissociation energies for H<sub>2</sub>O<sub>2</sub>, or 134.5 kcal/mol. Furthermore, lack of reactivity of some substrates that satisfy this requirement, for example secondary alcohols and certain amines, shows that branching at an  $\alpha$ -carbon prevents the catalytic reaction despite favorable thermodynamics (36).

Another common feature of the metalloenzymes that promote the same reactions as the Ni(TRISOX) complex is their catalytic use of a redox-active organic component in addition to the redox active copper that they require. In

galactose oxidase, it is a modified tyrosine that is cross-linked to a cysteine sulfur (39); in the cholesterol oxidase activity of the  $\beta$ -amyloid protein, it appears to be a tyrosine as well (42). In the Cu-dependent amine oxidases, the redox-active organic component is topaquinone (43). It is intriguing to consider, then, whether the reactivity analogy between the synthetic Ni(TRISOX) catalyst and the metalloenzymes extends beyond the selection of substrates and the products to a mechanism that includes a metal-organic redox hybrid. In fact, oximates can be oxidized to iminoxyl radicals capable of abstracting hydrogen atoms from various substrates (44-47).

### Spectroscopic Studies of an Observable Intermediate

When the reaction of the Ni(TRISOX) complex with oxygen in methanol is monitored by UV/visible absorption spectroscopy (Figure 3A), intense absorption bands grow in at around 350 nm, 420 nm, and 535 nm (35). The absorption maximizes after about 1 minute at room temperature with 1 mM Ni. The 350 and 420 nm features disappear over the course of 30-60 minutes, while significant intensity in the 535 nm region remains for a longer time. Under anaerobic conditions, spectroelectrochemistry experiments show strikingly similar spectra upon electrochemical oxidation of the Ni(II)(TRISOX) complex. In acetonitrile, which is not an oxidation substrate, features are seen at 360 and 500 nm, while in methanol the spectrum is nearly identical to that observed during the oxygen reaction. The reported spectrum of an oximate-derived iminoxyl radical has strong features at 350 nm and 400 nm, with a weaker absorption at 620 nm (46). This similarity suggests that the spectrum observed during the oxygen reaction is due to oxidation of oximate(s).

Electron paramagnetic resonance (EPR) spectroscopy confirms the formation of an organic radical. A strong feature near  $g = 2.0$  grows in and then disappears on a similar time scale to the intense features in the UV/visible spectrum (35). A similar EPR spectrum is observed upon electrochemical oxidation of an acetonitrile solution of Ni(TRISOX) by bulk electrolysis. In methanol, no EPR features are observed after bulk electrolysis due to reaction of the oxidized species with methanol. The location of the unpaired electron density responsible for these EPR features is confirmed by spectra in which either the oximate nitrogens or the amine nitrogen is substituted with  $^{15}\text{N}$  in air-exposed methanol (35). While the  $^{15}\text{N}$ -amine spectrum looks nearly identical to the natural abundance spectrum, the  $^{15}\text{N}$ -oximate spectrum has a very different shape, as shown in Figure 3B. As the  $g = 2$  EPR signal disappears, a classic Ni(III) signal grows in. This occurs on the same time scale as the conversion in reactivity from H-atom transfer chemistry with alcohol and amine substrates to O-atom transfer to a phosphine substrate. Later, as the oxygen chemistry ceases altogether, the Ni(III) EPR signal disappears.

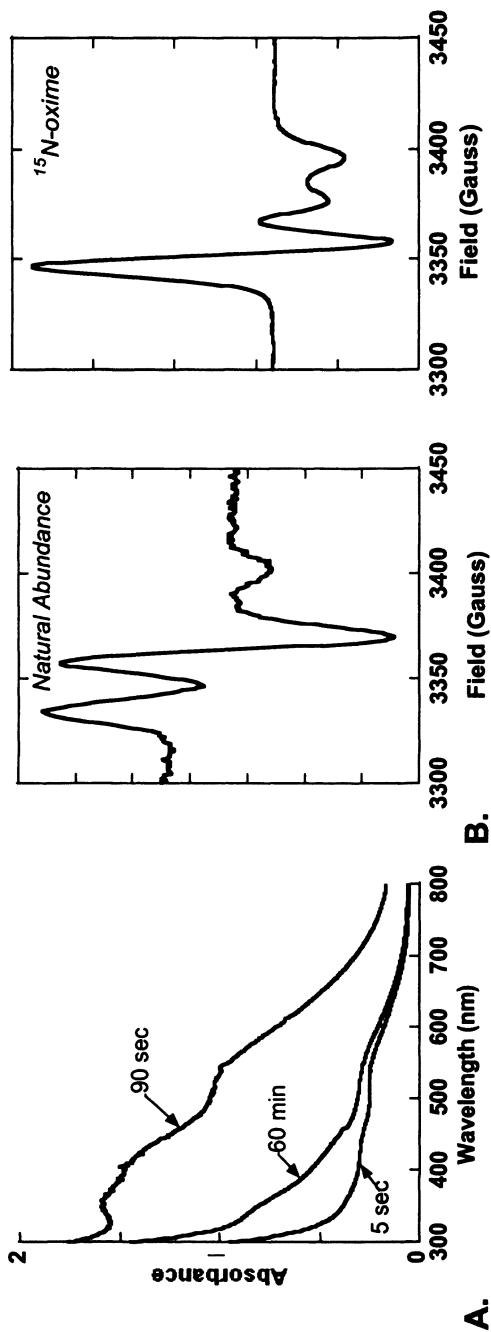


Figure 3. (A) UV/visible absorption spectra of the Ni(TRISOX)/methanol air oxidation system at various times after initiation of the reaction and (B) EPR spectra of the Ni(TRISOX)/methanol air oxidation system frozen at about 60 s after initiation, generated using natural abundance TRISOX (left) and generated using TRISOX enriched with  $^{15}\text{N}$  at the oximate nitrogens (right).

The UV/vis and EPR data demonstrate that an iminoxyl radical is formed during the oxygen reactivity of Ni(TRISOX) with methanol as substrate. Thus, the role of the ligand is not simply to support unusually high oxidation states of nickel, but it is intimately involved in the redox chemistry. This parallels the requirement for a redox-active organic cofactor in the enzymes described above. Additionally, the electrochemical generation of spectra that are so similar to those observed for the intermediate in catalytic turnover with O<sub>2</sub> suggests that an electrochemically generated model for the intermediate may be used to gain further insight into the relationship between the organic radical and the metal.

### TRISOX Ligation to Other Metals

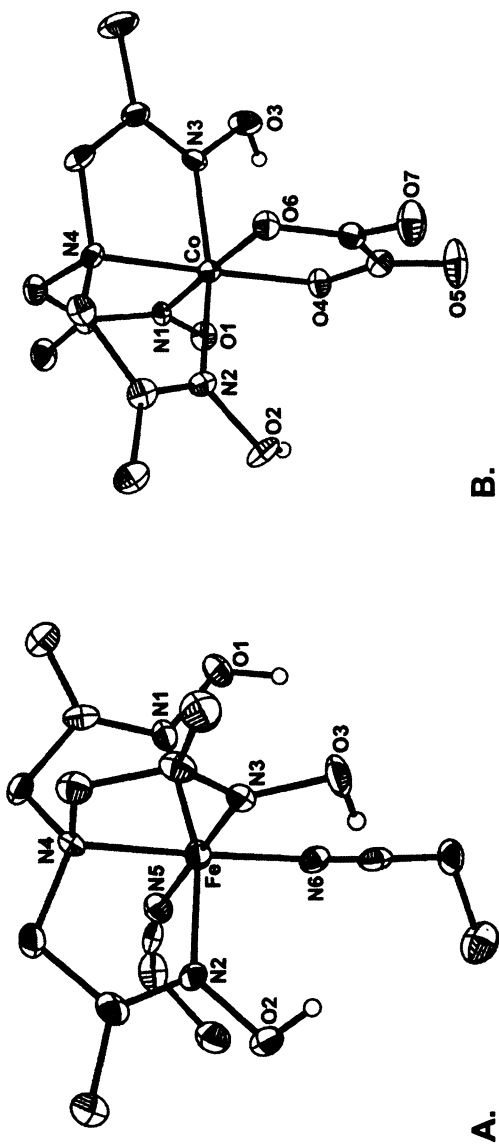
Although the H<sub>3</sub>TRISOX compound has been in the literature for a long time (29, 30) there have been no previous reports of structurally characterized metal complexes of H<sub>n</sub>TRISOX other than the Ni(II) complexes and a Zn(II) complex (48) associated with this work. In order to extend the TRISOX chemistry to other metals, we have determined structures of several new complexes of this ligand.

In [Fe(II)(H<sub>3</sub>TRISOX)(CH<sub>3</sub>CH<sub>2</sub>CN)<sub>2</sub>](ClO<sub>4</sub>)<sub>2</sub>, Figure 4A, the TRISOX ligand is coordinated in a similar manner to the Ni(II) complexes of H<sub>3</sub>TRISOX. That is, all three oximes are in their neutral, protonated state; solvent-derived ligands occupy the open *cis* coordination positions. In Co(III)(H<sub>2</sub>TRISOX)(ox), Figure 4B, the higher oxidation state of the metal results in deprotonation of one of the oximates in order to give a neutral coordination complex. The -2 charge on the oxalate ligand and the negative charge on the oximate arm of H<sub>2</sub>TRISOX<sup>-</sup> balance the +3 charge of Co(III). This is an example of a structure with a (deprotonated) oximate arm that does not form a bridge to a second metal.

[Zn<sub>2</sub>(H<sub>2</sub>TRISOX)<sub>2</sub>](ClO<sub>4</sub>)<sub>2</sub> in Figure 5 also has one deprotonated arm in the oximate form, but, like the nickel complex in Figure 2, the oximate bridges to a second metal. Unlike the nickel complex, the oximate bridge is formed in the absence of added base. This may be a consequence of the greater Lewis acidity of Zn(II) relative to Ni(II) and Fe(II) in the same coordination environment. Alternatively, the crystallized species may be the more insoluble component of a protonation equilibrium. When an electron rich ligand like acetate is coordinated as in Zn(H<sub>3</sub>TRISOX)(OAc)<sub>2</sub> (48) all of the oximes in the crystallized species maintain their protons in the absence of added base, possibly due to the decrease in Lewis acidity of the Zn.

Copper(II) has a Lewis acidity intermediate between those of Ni(II) and Zn(II). When Cu(II) was added to H<sub>3</sub>TRISOX in the absence of added base, a mixture of two complexes formed crystals. One, [Cu(H<sub>3</sub>TRISOX)(NO<sub>3</sub>)](NO<sub>3</sub>), is a blue monomer with the neutral H<sub>3</sub>TRISOX ligand, shown in Figure 6A. While the crystallographic data for the other, green crystals did not yield a





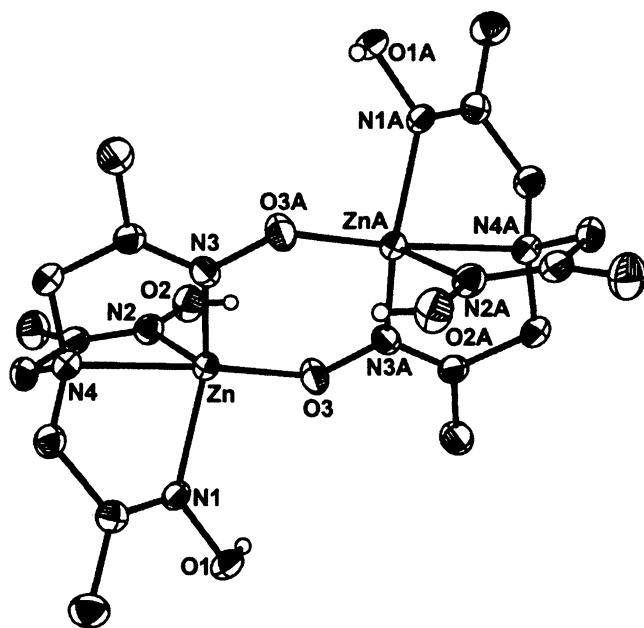
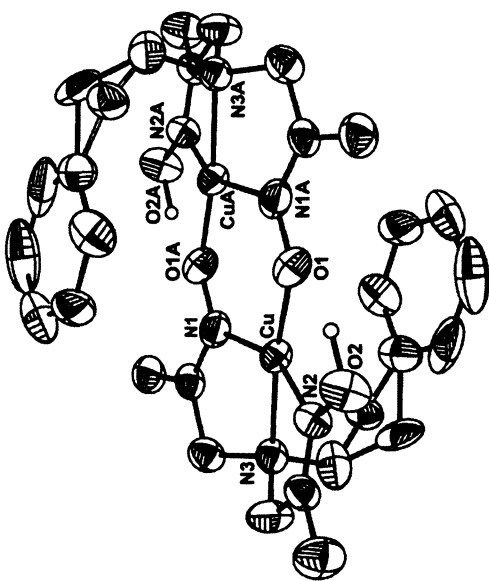


Figure 5. Crystal structure of  $[Zn_2(H_2TRISOX)_2](ClO_4)_2$ .

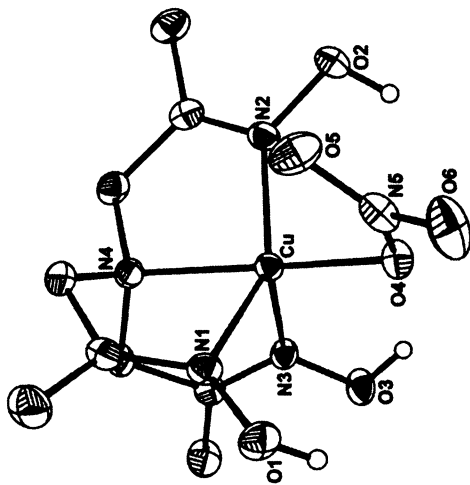
structure of sufficiently high resolution for publication, it is clear from the preliminary solution that it is a bis(oximate-bridged) dimer with the  $H_2TRISOX^-$  ligand like the complexes in Figures 2 and 5A. It is also similar to the structure of  $[Cu_2(HPEABO)_2](ClO_4)_2$  in Figure 6B. The  $H_2PEABO$  ligand is a derivative of  $H_3TRISOX$  in which one of the oxime arms is replaced by a non-coordinating phenylethyl arm (33).

## Conclusions

Explorations of the aerobic oxidation chemistry promoted by the Ni(II)-oximate redox hybrid described herein exemplifies the approach of bioinspired chemistry. While neither the metal nor the coordinating functional groups chosen mimic those in the enzymes that catalyze similar chemistry, studies by others on these enzymes has guided our understanding of the synthetic nickel catalyst. For example, the production of hydrogen peroxide as a co-product in the enzymes led us to determine the full net reaction catalyzed by Ni(TRISOX), and the thermodynamic parameters for substrate oxidation by this system. In examining this reaction, we found that the Ni(TRISOX) complex demonstrates “catalase-



B.



A.

Figure 6. (A) Crystal structure of  $[\text{Cu}(\text{H}_3\text{TRISOX})(\text{NO}_3)](\text{NO}_3)$ ; (B) Crystal structure of  $[\text{Cu}_2(\text{HPEABO})_2(\text{ClO}_4)_2]$ .

like” activity as well. Parallels to the redox-active organic component required in the enzyme systems aided in interpreting the spectroscopy of the observed intermediate in the oxygen chemistry of the nickel system. Finally, biomimetic Cu-O<sub>2</sub> chemistry that has been reported (49, 50) as part of model studies to understand various copper-dependent metalloenzymes has contributed to a proposed mechanism for methanol oxidation by the Ni(TRISOX) system (35). After O<sub>2</sub> and substrate binding equilibria to form the hypothetical complex on the left in Figure 7, two hydrogen atom transfers from substrate to catalyst produce the complex in the center that is in equilibrium with its tautomer on the right. This structure is consistent with the spectroscopic data obtained so far.

The unexpected chemistry that has been shown for the Ni(TRISOX) system suggests that exploring the chemistry of other metals with TRISOX will also reveal interesting reactivity. The structures reported here provide an initial basis for those studies, as they include complexes of both iron and copper, two metals that are frequently used in metalloenzymes for oxygen activation, as well as zinc which provides a useful non-redox active metal control for comparison. We have synthesized a number of derivatives of TRISOX in which one of the oxime arms is replaced by a different functional group (33). We are continuing to look at replacement of two oxime arms in order to determine the minimum ligand requirements for oxygen activity of Ni(II) with these ligands. The role of steric bulk is also being explored as a means of circumventing the conversion of the catalyst to an unreactive species.

The synthetic Ni(TRISOX) catalyst, multiple examples of enzymes that catalyze the same H-atom transfer chemistry as Ni(TRISOX), and the only other homogeneous metal catalysts reported to promote the multiple turnover aerobic oxidation of methanol (37) all incorporate the combination of a transition metal and a redox-active organic component. Having used information from the enzymes to elucidate the chemistry of Ni(TRISOX) thus far, it is intriguing to anticipate that understanding more details of the role of the ligand radical in the synthetic system will lead to new insight into the energetic reasons that the enzymes have evolved to use organic radicals in concert with transition metals to accomplish the two-H-atom transfer chemistry.

## Acknowledgments

NSF-MRI funding (CHE-0215950) for the SMART6000 diffractometer is gratefully acknowledged. The Ohio Crystallography Consortium was generously funded by the Ohio Board of Regents.

MJB dedicates this chapter to Dr. Gerard J. Colpas (1960-2008), bioinorganic chemist and great friend.

X-ray structural data are available in CIF format (CCDC 684747 - 684752) (free of charge) from [www.ccdc.cam.ac.uk/conts/retrieving.html](http://www.ccdc.cam.ac.uk/conts/retrieving.html) or from the

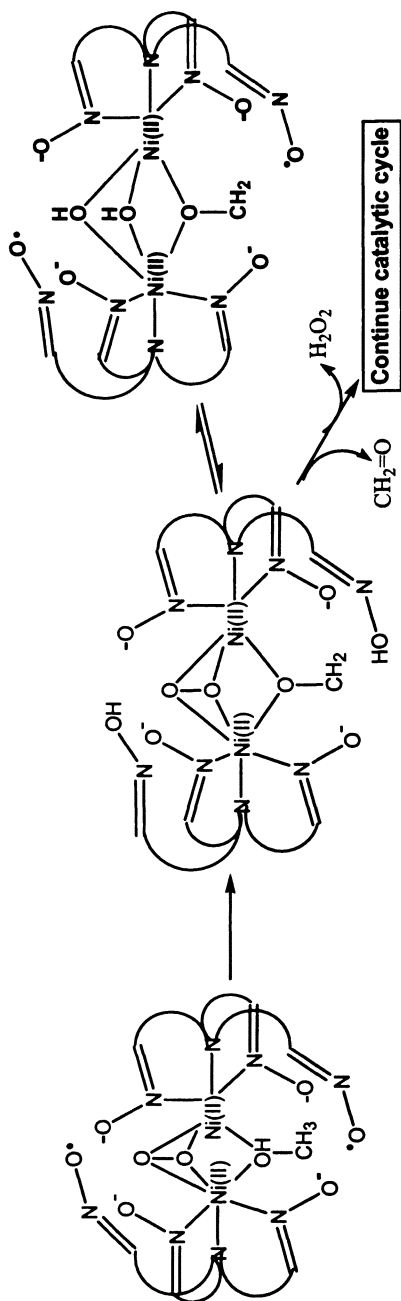


Figure 7. Key steps and structures in a proposed mechanism for aerobic methanol oxidation by the Ni(TRISOX) catalyst.

Cambridge Crystallographic Data Centre (CCDC), 12 Union Road, Cambridge CB2 1EZ, United Kingdom; e-mail: [deposit@ccdc.cam.ac.uk](mailto:deposit@ccdc.cam.ac.uk); Fax: +44(0)1223-336033.

## References

1. Barron, A. E.; Zuckerman, R. N. *Curr. Opin. Chem. Biol.* **1999**, *3*, 681-687.
2. Fricke, M.; Schadler, V. *Macromol. Biosci.* **2007**, *7*, 103-104.
3. Sanchez, C.; Arribart, H.; Giraud-Guille, M. M. *Nature Mater.* **2005**, *4*, 277-288.
4. Liebeskind, L. S.; Srogl, J.; Savarin, C.; Polanco, C. *Pure Appl. Chem.* **2002**, *74*, 115-122.
5. Hammes, B. S.; Luo, X.; Carrano, M. W.; Carrano, C. J. *Angew. Chem. Int. Ed.* **2002**, *41*, 3259-3261.
6. Park, H.; Kim, K. M.; Lee, A.; Ham, S.; Nam, W.; Chin, J. *J. Am. Chem. Soc.* **2007**, *129*, 1518-1519.
7. Sheldon, R. A. In *Biomimetic Oxidations Catalyzed by Transition Metal Complexes*; Meunier, B., Ed.; Imperial College Press: London, 2000; pp 613-662.
8. Kumar, M.; Colpas, G. J.; Day, R. O.; Maroney, M. J. *J. Am. Chem. Soc.* **1989**, *111*, 8323-8325.
9. Grapperhaus, C. A.; Darensbourg, M. Y. *Acc. Chem. Res.* **1998**, *31*, 451-459.
10. Paniago, E. B.; Weatherburn, D. C.; Margerum, D. W. *Chem. Commun.* **1971**, 1427-1428.
11. Bossu, F. P.; Paniago, E. B.; Margerum, D. W.; Kirskey, S. T. J.; Kurtz, J. L. *Inorg. Chem.* **1978**, *17*, 1034-1042.
12. Kimura, E.; Sakonaka, A.; Machida, R. *J. Am. Chem. Soc.* **1982**, *104*, 4255-4257.
13. Chen, D.; Martell, A. E. *J. Am. Chem. Soc.* **1990**, *112*, 9411-9412.
14. Cheng, C.-C.; Gulia, J.; Rokita, S. E.; Burrows, C. J. *J. Mol. Catal. A* **1996**, *113*, 379-391.
15. Fujita, K.; Schenker, R.; Gu, W.; Brunold, T. C.; Cramer, S. P.; Riordan, C. G. *Inorg. Chem.* **2004**, *43*, 3324-3326.
16. Schenker, R.; Kieber-Emmons, M. T.; Riordan, C. G.; Brunold, T. C. *Inorg. Chem.* **2005**, *44*, 1752-1762.
17. Shiren, K.; Ogo, S.; Fujinami, S.; Hayashi, H.; Suzuki, M.; Uehara, A.; Watanabe, Y.; Moro-oka, Y. *J. Am. Chem. Soc.* **2000**, *122*, 254-262.
18. Itoh, S.; Bandoh, H.; Nakagawa, M.; Nagatomo, S.; Kitagawa, T.; Karlin, K. D.; Fukuzumi, S. *J. Am. Chem. Soc.* **2001**, *123*, 11168-11178.
19. Dai, Y.; Pochapsky, T. C.; Abeles, R. H. *Biochemistry* **2001**, *40*, 6379-6387.

20. Chai, S. C.; Ju, T.; Dang, M.; Goldsmith, R. B.; Maroney, M. J.; Pochapsky, T. C. *Biochemistry* **2008**, *47*, 2428-2438.
21. Szajna-Fuller, E.; Rudzka, K.; Arif, A. M.; Berreau, L. M. *Inorg. Chem.* **2007**, *46*, 5499-5507.
22. Youn, H.-D.; Kim, E.-J.; Roe, J.-H.; Hah, Y. C.; Kang, S.-O. *Biochem. J.* **1996**, *318*, 889-896.
23. Fiedler, A. T.; Bryngelson, P. A.; Maroney, M. J.; Brunold, T. C. *J. Am. Chem. Soc.* **2005**, *127*, 5449-5462.
24. Collins, T. J. *Acc. Chem. Res.* **1994**, *27*, 279-285.
25. Kruger, H.-J.; Peng, G.; Holm, R. H. *Inorg. Chem.* **1991**, *30*, 734-742.
26. Baucom, E. I.; Drago, R. S. *J. Am. Chem. Soc.* **1971**, *93*, 6469-6475.
27. Singh, A. N.; Singh, R. P.; Mohanty, J. G.; Chakravorty, A. *Inorg. Chem.* **1977**, *16*, 2597-2601.
28. Singh, A. N.; Chakravorty, A. *Inorg. Chem.* **1980**, *19*, 969-971.
29. Matthaiopoulos, G. *Chem. Ber.* **1898**, *31*, 2396-2399.
30. Ogloblin, K. A.; Potekhin, A. A. *J. Org. Chem. USSR (Engl. Trans.)* **1965**, *1*, 399-405.
31. Goldcamp, M. J.; Krause Bauer, J. A.; Baldwin, M. J. *Acta Cryst.* **2002**, *E58*, o1354-o1355.
32. Goldcamp, M. J.; Robison, S. E.; Krause Bauer, J. A.; Baldwin, M. J. *Inorg. Chem.* **2002**, *41*, 2307-2309.
33. Goldcamp, M. J.; Edison, S. E.; Squires, L. N.; Rosa, D. T.; Vowels, N. K.; Coker, N. L.; Krause Bauer, J. A.; Baldwin, M. J. *Inorg. Chem.* **2003**, *42*, 717-728.
34. Jones, R. M.; Goldcamp, M. J.; Krause, J. A.; Baldwin, M. J. *Polyhedron* **2006**, *25*, 3145-3158.
35. Edison, S. E.; Conklin, S. D.; Kaval, N.; Cheruzel, L. E.; Krause, J. A.; Seliskar, C. J.; Heineman, W. R.; Buchanan, R. M.; Baldwin, M. J. *Inorg. Chim. Acta* **2008**, *361*, 947-955.
36. Edison, S. E.; Hotz, R. P.; Baldwin, M. J. *Chem. Commun.* **2004**, 1212-1213.
37. Chaudhuri, P.; Hess, M.; Muller, J.; Hildenbrand, K.; Bill, E.; Weyhermuller, T.; Wieghardt, K. *J. Am. Chem. Soc.* **1999**, *121*, 9599-9610.
38. Zhang, C. X.; Liang, H.-C.; Kim, E.-I.; Shearer, J.; Helton, M. E.; Kim, E.; Kaderli, S.; Incarvito, C. D.; Zuberbuhler, A. D.; Rheingold, A. L.; Karlin, K. D. *J. Am. Chem. Soc.* **2003**, *125*, 634-635.
39. Clark, K.; Penner-Hahn, J. E.; Whittaker, M.; Whittaker, J. W. *Biochemistry* **1994**, *33*, 12553-12557.
40. Puglielli, L.; Friedlich, A. L.; Setchell, K. D. R.; Nagano, S.; Opazo, C.; Cherney, R. A.; Barnham, K. J.; Wade, J. D.; Melov, S.; Kovacs, D. M.; Bush, A. I. *J. Clin. Invest.* **2005**, *115*, 2556-2563.
41. McGuirl, M. A.; Dooley, D. M. *Curr. Opin. Chem. Biol.* **1999**, *3*, 138-144.
42. Barnham, K. J.; Haeffner, F.; Ciccotosto, G. D.; Curtain, C. C.; Tew, D.; Mavros, C.; Beyreuther, K.; Carrington, D.; Masters, C. L.; Cherny, R. A.; Cappai, R.; Bush, A. I. *FASEB J.* **2004**, *18*, 1427-1429.

43. Janes, S. M.; Palcic, M. M.; Scaman, C. H.; Smith, A. J.; Brown, D. E.; Dooley, D. M.; Mure, M.; Klinman, J. P. *Biochemistry* **1992**, *31*, 12147-12154.
44. Thomas, J. R. *J. Am. Chem. Soc.* **1964**, *86*, 1446-1447.
45. Brokenshire, J. L.; Roberts, J. R.; Ingold, K. U. *J. Am. Chem. Soc.* **1972**, *94*, 7040-7049.
46. Everett, S. A.; Naylor, M. A.; Stratford, M. R. L.; Patel, K. B.; Ford, E.; Mortenson, A.; Furguson, A. C.; Vojnovic, B.; Wardman, P. *J. Chem. Soc., Perkin Trans. 2* **2001**, 1989-1997.
47. Pratt, D. A.; Blake, J. A.; Mulder, P.; Walton, J. C.; Korth, H.-G.; Ingold, K. I. *J. Am. Chem. Soc.* **2004**, *126*, 10667-10675.
48. Goldcamp, M. J.; Krause Bauer, J. A.; Baldwin, M. J. *J. Chem. Crystallog.* **2005**, *35*, 77-83.
49. Mahadevan, V.; Henson, M. J.; Solomon, E. I.; Stack, T. D. P. *J. Am. Chem. Soc.* **2000**, *122*, 10249-10250.
50. Lam, B. M. T.; Halfen, J. A.; Young Jr, V. G.; Hagadorn, J. R.; Holland, P. L.; Lledos, A.; Cucurull-Sanchez, L.; Novoa, J. J.; Alvarez, S.; Tolman, W. B. *Inorg. Chem.* **2000**, *39*, 4059-4072.



## Chapter 10

# Inorganic Models for Two-Electron Redox Chemistry in Biological Systems

## Ligand-Bridged Molybdenum and Tungsten Dimers

Franklin A. Schultz<sup>1,\*</sup>, Richard L. Lord<sup>2</sup>, Xiaofan Yang<sup>2</sup>,  
and Mu-Hyun Baik<sup>2,3,\*</sup>

<sup>1</sup>Department of Chemistry and Chemical Biology,  
Indiana University–Purdue University Indianapolis (IUPUI),  
Indianapolis, IN 46202

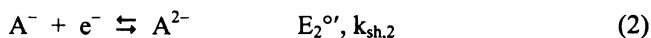
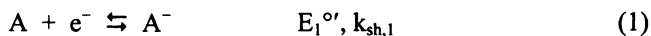
<sup>2</sup>Department of Chemistry and School of Informatics,  
Indiana University–Bloomington (IUB), Bloomington, IN 47405

<sup>3</sup>Institut für Chemie, Technische Universität Berlin, Berlin, Germany

We describe the energetics of concerted two-electron transfer and metal–metal bond cleavage in the sulfido- and phosphido-bridged dimers  $[M_2(\mu-ER_n)_2(CO)_8]^{0/2-}$  ( $M = Mo$  or  $W$ ;  $R_nE^- = RS^-$  or  $R_2P^-$ ). Reduction of the neutral dimers to their corresponding dianions occurs in a single two-electron step accompanied by cleavage of a metal–metal single bond and large changes in the M–M distance and M–E–M and E–M–E bridge angles. These structural reorganizations provide the driving force for concerted two-electron transfer via inversion of the 0/– and –/2– redox potentials.

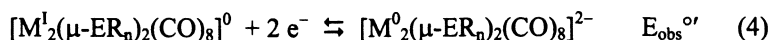
Concerted multielectron transfer is crucial to many chemical and biological processes including nitrogen fixation (1, 2), hydrogen production (3, 4), and oxygen evolution (5, 6). The ability to exchange two redox equivalents in a

single step correlates well the requirements of chemical bond formation and bypasses potentially harmful radical intermediates. Although essential, multielectron transfer is inherently paradoxical, because electrostatic considerations suggest that it should be more difficult to transfer a second unit of charge to a reactant than the first ( $E_1^{o'} > E_2^{o'}$ , equations 1 and 2) and that the one-electron intermediate,  $A^-$ , should be stable with respect to disproportionation ( $\Delta G_{\text{disp}} > 0$ , equation 3).



This normal sequence of events accords with the “convexity principle” (7), which states that the energy of stabilization upon addition of the first electron to an atomic or molecular system ( $-\Delta G_1$ ) is greater than the stabilization energy attending addition of the second ( $-\Delta G_2$ ). The converse set of conditions ( $-\Delta G_2 > -\Delta G_1$ ,  $E_2^{o'} > E_1^{o'}$ ,  $\Delta G_{\text{disp}} < 0$ ) may arise, however, when a change in structure or composition, e.g., ligation, protonation, ion-pair formation, accompanies electron transfer. This circumstance results in a multielectron event via an “inversion” of potentials and the destabilization of  $A^-$ . Numerous examples of such behavior have been identified and discussed (8-13).

In recent years, we have encountered and undertaken experimental studies (14-18) on two families of ligand-bridged binuclear complexes, which undergo the one-step, two-electron transfer shown in equation 4.



This reaction is noteworthy in that multielectron transfer occurs without a change in stoichiometry of the redox active unit. Moreover, a metal-metal single bond is cleaved and large structural deformations — expansion and contraction of the M-E-M and E-M-E angles by ca. 30° and an increase of >1.0 Å in the M-M distance — occur within the intact  $M_2E_2$  core as the complexes pass from the  $M(I)_2$  state to the  $M(0)_2$  state. These structural changes, which are illustrated in Figure 1, are presumed to provide the driving force for coincident two-electron transfer.

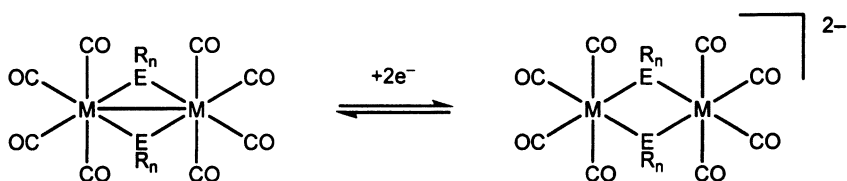


Figure 1. Structural changes accompanying the two-electron reduction of  $[M_2(\mu-ER_n)_2(CO)_8]^0$  ( $M = Mo, W$ ;  $R_nE^- = RS^-, R_2P^-$ ) complexes.

Concerted multielectron transfer accompanied by metal–metal bond cleavage within compositionally invariant redox centers is relevant to the biological processes of nitrogen fixation (1, 2) and hydrogen evolution (3, 4). The active site of  $N_2$  fixation is thought to be the iron–molybdenum cofactor (FeMoco), which is a polynuclear  $MoFe_7S_9X$  core within the MoFe protein of nitrogenase. The intermetal distances in the semi-reduced state of FeMoco are consistent with metal–metal bonding, but are too short to permit substrate binding. It has been suggested (19) that electron transfer accompanied by metal–metal bond cleavage could loosen the Mo–Fe–S framework of FeMoco and enable substrate binding as the enzyme is reduced to its catalytically competent level. Several pieces of evidence support this hypothesis. (1) A polynuclear Mo–Fe–S cluster has been prepared wherein an increase by two in the number of valence electrons lengthens an Fe–Fe distance by  $\sim 1.0 \text{ \AA}$  (20). (2) Electrochemical studies show that FeMoco (extracted from the MoFe protein of nitrogenase) undergoes two-electron reduction at a negative potential under an atmosphere of CO (21). (3) The  $Fe_8S_7$  P-cluster of nitrogenase, which supplies electrons to FeMoco, also exhibits two-electron redox behavior (22).

The  $[M_2(\mu-ER_n)_2(CO)_8]^{0/2-}$  complexes in Figure 1 also are similar to the sulfido-bridged di-iron carbonyls,  $Fe_2(\mu-SR)_2(CO)_6$ , which model the active site of the Fe-only hydrogenases (23, 24). Species such as  $Fe_2(\mu-SR)_2(CO)_6$  (25) and their  $Fe_2(\mu-PR_2)_2(CO)_6$  analogs (26, 27) undergo two-electron reduction accompanied by metal–metal bond cleavage via reactions that are comparable to equation 4. For these compounds changes in bridging and/or terminal ligand structure can sensitively adjust the balance between normal and inverted potential behavior (28–30). Moreover, many  $Fe_2(\mu-SR)_2$  model compounds catalyze the electrochemical reduction of  $H^+$  to  $H_2$  (31–33). The extent to which the one-electron intermediate is stabilized is an important factor in understanding the mechanism of such reactions.

Cognizant of the above parallels we have conducted detailed electrochemical studies of the kinetics and thermodynamics of the two-electron transfer reactions of  $[M_2(\mu-ER_n)_2(CO)_8]^{0/2-}$  complexes (14–18). We have sought

to understand the extent to which the nature of the metal and the bridging ligand influence the energetics of the overall process and to identify the molecular and structural features which enable potential inversion. The experimental findings are compared with the results of computational studies employing extended Hückel molecular orbital (EHMO) and density functional (DFT) theories. We have shown previously that DFT is capable of calculating redox potentials to within approximately 150 mV (34), providing a powerful basis for a deep conceptual understanding of multielectron redox behavior (35-37) through calculation of the disproportionation free energy in solution (equation 3). Moreover, specific molecular contributions to the energetics of concerted two-electron transfer can be identified by construction of a theoretical square scheme wherein energy differences are partitioned between charge addition and structural change for each redox step (36). Such detailed examination of experimental and computational results can provide insight to molecular features that control multielectron behavior in biological systems.

## Electrochemical Studies

Electrochemical experiments were conducted by use of equipment and procedures described in references 15-18. In general, electrochemical data were obtained by oxidizing  $[\text{Et}_4\text{N}]_2[\text{M}_2(\mu\text{-SR})_2(\text{CO})_8]$  and reducing  $[\text{M}_2(\mu\text{-PR}_2)_2(\text{CO})_8]$  forms at a glassy carbon or mercury working electrode in purified solvent plus supporting electrolyte. Syntheses of these compounds are described in references 14, 15, and 18. The procedure for extracting electrochemical parameters for a two-electron electrode reaction from variable sweep-rate cyclic voltammetry data is based on the analysis of Ryan (38) and is described in detail in reference 18. In all cases, parameter values were checked by comparison of experimental traces with those obtained by digital simulation (39). The uncertainty of measured potentials is 10–20 mV; that of measured rate constants is 10–20% (18). Potentials measured using different reference half-cells were corrected to values versus the normal hydrogen electrode (NHE) by use of the following offsets:  $E^{\circ}$  (ferrocene/ferricenium) = +0.64 V vs. NHE (40),  $E^{\circ}$  (SCE) = +0.241 V vs. NHE (41), and  $E^{\circ}$  (Ag/AgCl) = +0.232 V vs. NHE (42).

Figure 2 shows cyclic voltammograms and observed redox potentials  $[E_{\text{obs}}^{\circ}] = (E_1^{\circ} + E_2^{\circ})/2$  for  $[\text{W}_2(\mu\text{-SPh})_2(\text{CO})_8]^{0/2-}$  (15) and  $[\text{W}_2(\mu\text{-PPh}_2)_2(\text{CO})_8]^{0/2-}$  (18) at a sweep rate of 0.1 V s<sup>-1</sup>. The overall two-electron character of the reactions has been confirmed by controlled potential coulometry and comparative voltammetric peak current measurements. Only single reduction and oxidation peaks are observed at sweep rates ranging from 0.005 to 1000 V s<sup>-1</sup>; thus, the one-electron intermediate in reactions 1 and 2 is not detectable by simple electrochemical means.

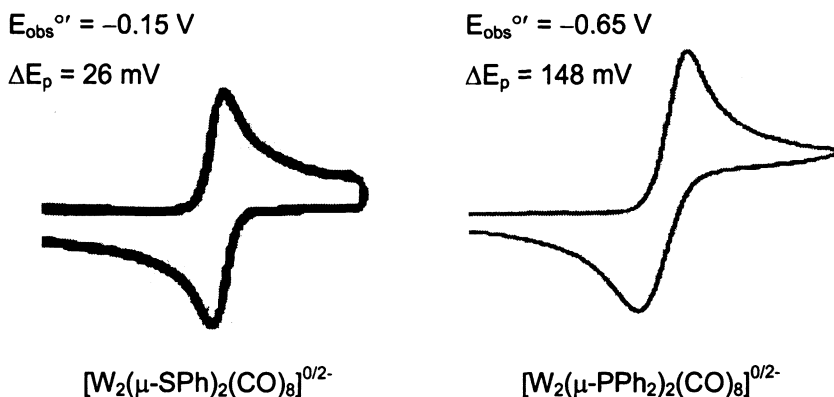


Figure 2. Cyclic voltammetric responses of 1.8 mM  $[\text{W}_2(\mu\text{-SPh})_2(\text{CO})_8]^{0/2-}$  (glassy carbon electrode, DMF/0.1 M  $\text{Bu}_4\text{NBF}_4$ ) and 0.3 mM  $[\text{W}_2(\mu\text{-PPh}_2)_2(\text{CO})_8]^{0/2-}$  (Hg electrode, acetone/0.3 M  $\text{Bu}_4\text{NPF}_6$ ) at 0.1  $\text{V s}^{-1}$ .

There are several notable differences and similarities in electrochemical parameters among the twelve  $[\text{M}_2(\mu\text{-ER}_n)_2(\text{CO})_8]^{0/2-}$  redox systems that have been examined (14-18). Most importantly the thermodynamics and kinetics of reaction 4 are very dependent on the nature of the bridging ligand donor atom (E), but are influenced only marginally by the bridging ligand substituent (R) and the identity of the metal (M). As shown in Figure 2 the phosphido-bridged complex is more difficult to reduce than the sulfido-bridged species by some 500 mV, and the apparent electron-transfer kinetics are more sluggish for the  $\text{Ph}_2\text{P}^-$  analog. The latter fact is apparent from the relative peak potential separations in Figure 2.  $\Delta E_{\text{p}}$  approximately equals the 29 mV value anticipated for a Nernstian two-electron transfer for  $[\text{W}_2(\mu\text{-SPh})_2(\text{CO})_8]^{0/2-}$  (41), whereas  $\Delta E_{\text{p}}$  is 148 mV for  $[\text{W}_2(\mu\text{-PPh}_2)_2(\text{CO})_8]^{0/2-}$ . In contrast,  $E_{\text{obs}}^{\text{o}'}$  values for homologous complexes with M = Mo or W differ by no more than 40 mV, and those with R = Ph, *t*-Bu, Bz exhibit a range of less than 200 mV. Electron-transfer kinetics likewise are influenced only slightly by such changes in composition.

We have used several experimental methodologies to ascertain the extent of potential inversion and the kinetic and thermodynamic parameters of reactions 1 and 2. A mechanistic study employing variable sweep rate cyclic voltammetry ( $\nu = 0.005$  to 40  $\text{V s}^{-1}$ ) and digital simulation (39) found  $\Delta E^{\text{o}'} = E_2^{\text{o}'} - E_1^{\text{o}'} = +20$  mV and heterogeneous electron-transfer rate constants of  $k_{\text{sh},1} = k_{\text{sh},2} \cong 0.1 \text{ cm s}^{-1}$  for  $[\text{W}_2(\mu\text{-SBz})_2(\text{CO})_8]^{0/2-}$  in DMF at a Hg electrode (16). Although no evidence of the electrode reaction intermediate was found in this electrochemical experiment, an infrared spectroelectrochemical investigation detected the one-electron reduced dimer via changes in the carbonyl absorptions at 1700–2100  $\text{cm}^{-1}$  and reported  $\Delta E^{\text{o}'} = +230$  mV in  $\text{CH}_2\text{Cl}_2/0.1 \text{ M Bu}_4\text{NPF}_6$  (17).

Accurate determination of electrochemical parameters is difficult for the sulfido-bridged complexes, because of their relatively large electron-transfer rate constants and the tendency of oxidized forms to undergo solvolysis reactions (15). Consequently, we selected the  $[M_2(\mu\text{-PPh}_2)_2(\text{CO})_8]^{0/2-}$  couples for detailed mechanistic study (18) employing variable sweep rate cyclic voltammetry and the guidelines of Ryan (38) for diagnosing two-electron electrode reactions. An important first goal is identification of the rate-limiting step of the overall reaction. Examination of the voltammogram in Figure 2 as well as those recorded at faster sweep rates (18) shows that the anodic peak is smaller and broader than the cathodic one during the reduction and reoxidation of  $[\text{Mo}_2(\mu\text{-PPh}_2)_2(\text{CO})_8]^0$ . This observation identifies the second electron transfer (equation 2) as rate-limiting and allows for extracting the relevant electrochemical parameters  $E_1^{o'}$ ,  $E_2^{o'}$ ,  $k_{sh,1}$ , and  $k_{sh,2}$ . These results are presented in Table I together with estimated values for the  $[M_2(\mu\text{-SR})_2(\text{CO})_8]^{0/2-}$  systems.

**Table I. Electrochemical Data for  $[M_2(\mu\text{-ER}_n)_2(\text{CO})_8]^{0/2-}$  Complexes<sup>a</sup>**

$M\text{-}E$ ( $R_n$ )	Solvent/Supporting Electrolyte	$E_{obs}^{o'}$ (V)	$\Delta E^{o'}$ (mV)	$k_{sh,1}$ ( $\text{cm s}^{-1}$ )	$k_{sh,2}$ ( $\text{cm s}^{-1}$ )	Ref.
Mo-S (Ph)	CH <sub>3</sub> CN/ 0.1 M Bu <sub>4</sub> NBF <sub>4</sub>	-0.13	N.A.	N.A.	N.A.	15
W-S (Ph)	DMF/ 0.1 M Bu <sub>4</sub> NBF <sub>4</sub>	-0.15, -0.34 <sup>b,c</sup>	+20 <sup>b</sup> +230 <sup>b,c</sup>	~ 0.1 <sup>b</sup>	~ 0.1 <sup>b</sup>	15- 17
Mo-P (Ph <sub>2</sub> )	Acetone/ 0.3 M Bu <sub>4</sub> NPF <sub>6</sub>	-0.69	+170	0.4	0.038	18
W-P (Ph <sub>2</sub> )	Acetone/ 0.3 M Bu <sub>4</sub> NPF <sub>6</sub>	-0.65	+180	0.8	0.014	18

<sup>a</sup> $E_{obs}^{o'} = (E_1^{o'} + E_2^{o'})/2$ ;  $\Delta E^{o'} = E_2^{o'} - E_1^{o'}$ ; potentials in V vs. NHE.

<sup>b</sup>Values reported are for the BzS<sup>-</sup> derivative (16, 17).

<sup>c</sup>From infrared spectroelectrochemistry in CH<sub>2</sub>Cl<sub>2</sub>/0.1 M Bu<sub>4</sub>NPF<sub>6</sub> (17).

Several important points are revealed by the data in Table I. (1) For the compounds investigated thus far, inverted potential behavior is observed with  $\Delta E^{o'} = E_2^{o'} - E_1^{o'}$  ranging from +20 to +230 mV. Because a normal ordering of potentials for molecules the size of  $[M_2(\mu\text{-ER}_n)_2(\text{CO})_8]^{0/2-}$  is anticipated to yield  $\Delta E^{o'} = -0.5$  to  $-0.7$  V (8), the structural reorganizations that accompany reaction

4 must provide a significant energy offset to overcome conventional electrostatics. (2) Rate constants for the second electron transfer,  $k_{sh,2}$ , are more than an order of magnitude smaller than  $k_{sh,1}$  for  $[M_2(\mu-PPh_2)_2(CO)_8]^{0/2-}$  couples. This fact is consistent with identification of equation 2 as the rate-limiting electrode reaction and suggests that greater inner-shell reorganization accompanies the second electron transfer, as solvent reorganization is expected to be approximately the same in all cases. (3) Electrode potentials are more negative and electron-transfer rate constants are smaller for  $R_2P^-$  compared to  $RS^-$  bridged species, whereas neither potentials nor rate constants exhibit a significant metal dependence. This finding contrasts with the pattern of Mo- and W-based potentials in higher oxidation state complexes where  $E_{Mo}^{o'} - E_W^{o'}$  is typically ca. +0.6 V (43). However, in polynuclear M-Fe-S cluster compounds, where the heterometal is in a lower oxidation state, this difference is on the order of 0.0–0.2 V (44).

Observation of greater ligand- rather than metal-dependent potentials and rate constants merits further comment and suggests that the two-electron reactions involve significant bridging ligand participation. Indeed, more negative potentials for reduction of the sulfido-bridged complexes are consistent with both the greater electronegativity of S versus P and reported values of –2.3 (45) and –2.5 (46) V vs. NHE for reduction of  $Ph_2S$  and  $Ph_3P$ , respectively. The role of ligand-centered electron transfer in biological processes is an important consideration. Bridging ligand participation in multimetal electron transfer reactions recently has been probed through studies of binuclear  $Cu_2N_2$ ,  $Cu_2P_2$  and  $Cu_2S_2$  complexes, which model the  $Cu_2(\mu-SR)_2$  active site of cytochrome *c* oxidase (47–49). Metal-ligand bond covalency is an important determinant of the redox properties of these systems. Hence, the bridging ligand character of the redox active molecular orbital in  $[M_2(\mu-ER_n)_2(CO)_8]^{0/2-}$  complexes is among the properties examined in the following computational studies.

## Computational Studies

Extended Hückel molecular orbital calculations were used to obtain an initial, qualitative picture of the electronic structure changes involved in two-electron transfer. EHMO calculations were carried out in  $D_{2h}$  symmetry on  $H_2P^-$  bridged analogs of the  $[M_2(\mu-PR_2)_2(CO)_8]^{0/2-}$  complexes using the CACAO program (50). Structural information and computational details are provided in reference 18.

Figure 3 contains drawings of the redox active molecular orbital in the oxidized and reduced states of the W complex and a Walsh diagram of the orbital energy changes attending the reaction. The redox active orbital is the  $b_{3u}$   $\sigma^*$  orbital, which arises from an antibonding interaction between the  $d_{x^2-y^2}$

orbitals of the two metals. In the  $M(I)_2$  state this orbital is the vacant LUMO and has significant metal (22%) plus phosphorus (12%) character. In the  $M(0)_2$  state it is the filled HOMO and exhibits increased metal (36%) and diminished phosphorus (8%) content.

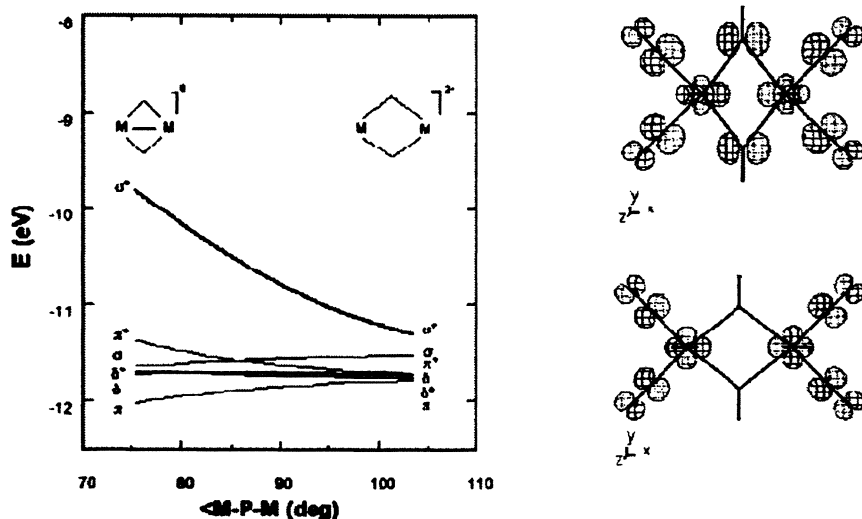


Figure 3. Walsh diagram (left) and drawings of the redox active  $\sigma^*$  orbital (right) in the oxidized and reduced forms of  $[W_2(\mu-PH_2)_2(CO)_8]^{0/2-}$ . Reproduced with permission from reference 18. Copyright 2002 American Chemical Society.

The Walsh diagram shows the remarkable dependence of the  $\sigma^*$  orbital energy on structure. This energy decreases dramatically as the  $M-P-M$  bridge angle increases from  $75^\circ$  to  $104^\circ$ , whereas the energies of the other  $\sigma$ ,  $\pi$  and  $\delta$  orbitals in the metal-metal bonding manifold remain relatively unchanged. Thus, electron transfer into the metal-plus-bridging ligand antibonding orbital cleaves the  $M-M$  single bond and initiates structural changes that lead to an inversion of potentials and an overall two-electron transfer.

Density functional theory provides a more quantitative means of calculating the energy changes accompanying a change in structure or composition of a molecular system. Hence, DFT calculations were performed using the Jaguar 5.5 suite of quantum chemistry programs (51). Geometry optimizations were performed at the B3LYP/6-31G\*\* level of theory (52, 53) with Mo and W represented using the Los Alamos LACVP basis (54-56). While this model chemistry generates reasonable structures, the energies so calculated are not



reliable for redox phenomena, as previously reported (34). Subsequent single-point energies were calculated with Dunning's correlation consistent triple- $\zeta$  basis set, cc-pVTZ(-f) (57), with transition metals represented using a decontracted version of LACVP to match the effective core potential with a triple- $\zeta$  quality basis.

Solvation energies were computed at the double- $\zeta$  level using a self-consistent reaction field approach based on numerical solutions of the Poisson-Boltzmann equation (58-60). These were computed at the optimized gas-phase geometry utilizing an appropriate dielectric constant for comparison to the experimental conditions ( $\epsilon = 37.5$  for acetonitrile;  $\epsilon = 20.7$  for acetone). The standard set of optimized radii in Jaguar were employed: Mo (1.526 Å), W (1.534 Å), H (1.150 Å), C (1.900 Å), O (1.600 Å). Vibrational analyses using analytical frequencies were computed at the double- $\zeta$  level, ensuring all stationary points to be minima.

Thermodynamic properties were obtained as summarized in equations 5-8, with standard approximations assumed for entropy corrections in the gas phase. Solution phase free energies were obtained by adding the free energy of solvation to the gas phase free energy. Zero-point energies and entropy corrections were derived using unscaled frequencies.

$$\Delta H(\text{gas}) = \Delta E(\text{SCF}) + \Delta ZPE \quad (5)$$

$$\Delta G(\text{gas}) = \Delta H(\text{gas}) - 298.15 \cdot \Delta S(\text{gas}) \quad (6)$$

$$\Delta G(\text{sol}) = \Delta G(\text{gas}) + \Delta G(\text{solv}) \quad (7)$$

$$E^{\circ} = -\Delta G^{\text{ET}}(\text{sol})/F \quad (8)$$

The terms in equations 5-8 are defined as  $\Delta H(\text{gas})$  = gas phase enthalpy,  $\Delta E(\text{SCF})$  = electronic energy,  $\Delta ZPE$  = zero-point energy correction,  $\Delta G(\text{gas})$  = gas phase Gibbs free energy,  $\Delta S(\text{gas})$  = gas phase entropy,  $\Delta G(\text{sol})$  = solution phase free energy,  $\Delta G(\text{solv})$  = free energy of solvation,  $\Delta G^{\text{ET}}(\text{sol})$  = solvated free energy of reduction, and  $E^{\circ}$  = standard reduction potential.

For two-electron redox chemistry the quantity of interest is the free energy change for the disproportionation of the one-electron intermediate ( $\Delta G_{\text{disp}}$ , eq 3). This quantity equals the energy difference between the dianion and monoanion (ET2) minus the energy difference between the monoanion and neutral (ET1);  $\Delta G_{\text{disp}} = \Delta \text{ET} = \text{ET2} - \text{ET1}$ .  $\Delta G_{\text{disp}} < 0$  indicates that the disproportionation of  $A^-$  is spontaneous and that two-electron transfer will ensue via inversion of potentials  $E_1^{\circ}$  and  $E_2^{\circ}$ .

We may examine multielectron transfer reactions in greater depth by partitioning the energy changes for each step into those resulting from electron

addition and those resulting from structural relaxation (36). Electron addition (E1, E2) is computed by taking the difference between the energies of the reduced and oxidized forms at the geometry of the oxidized form. The structural relaxation component (S1, S2) is captured by allowing structural relaxation and computing that energy gain. The overall energy changes for each electron transfer reaction are thus  $ET1 = E1 + S1$  and  $ET2 = E2 + S2$ . Free energy differences for isolated molecules in the gas phase,  $\Delta G(\text{gas})$ , include contributions from zero-point energy ( $\Delta ZPE$ ) and entropy ( $-T\Delta S$ ) differences. However, because these terms are no larger than 0.1–0.2 eV,  $\Delta G(\text{gas})$  consists almost entirely of the difference in electronic energies [ $\Delta E(\text{SCF})$ ]. Solvation energy differences,  $\Delta G(\text{solv})$ , which are calculated using a continuum solvation model (61, 62), are added to the gas-phase energy yielding  $\Delta G(\text{sol}) = \Delta G(\text{gas}) + \Delta G(\text{solv})$  as the free energy change attending each component of the electron transfer reaction in solution.

Results of the DFT calculations are summarized in Table II. The energy partitioning between electron addition and structural change is depicted in Figure 4 for  $[\text{Mo}_2(\mu\text{-SPh})_2(\text{CO})_8]^{0/-2-}$ . Here, horizontal arrows represent the energy change attending electron addition, vertical arrows represent the energy change attending structural relaxation, and diagonal arrows represent the overall free energy differences for each electron transfer reaction. Free energy differences are converted to electrode potentials by use of the relation  $E^{or} = -\Delta G(\text{sol})/F$  and addition of 4.43 V (63) to convert to the NHE scale.

The following salient points are derived from the data in Table II. Gas-phase addition of the first electron is favorable in all cases [ $E1(\text{gas}) < 0$ ], but addition of the second electron is not [ $E2(\text{gas}) > 0$ ]. Solvation favors both electron transfers, but stabilizes the second much more than the first. Values of  $\Delta G(\text{solv})$  are consistent with the Born model, which predicts this quantity to vary as the square of charge; hence,  $E2(\text{solv}) \approx 3 \cdot E1(\text{solv})$ . However, Table II shows that, if disproportionation free energies are calculated on the basis of electron addition and solvation alone, potential inversion would not occur in three of four cases [i.e.,  $E2(\text{sol}) - E1(\text{sol}) > 0$ ]. In other words, the convexity principle (7) would be obeyed without the structural reorganization. It is the disproportionate partition of the overall structural relaxation energy that ultimately affords the two-electron redox behavior. For example, in the case of  $[\text{Mo}_2(\mu\text{-SPh})_2(\text{CO})_8]^{0/2-}$  we found that the overall structural relaxation energy that accompanies the two-electron reduction process amounts to 1.325 eV, i.e., 30.55 kcal mol<sup>-1</sup>. Our energy partitioning protocol assigns 0.500 eV of that energy to be released in the first step [S1(solv) in Table II], whereas 0.825 eV is released following the second reduction. Thus, structural relaxation contributes 0.325 eV towards potential inversion. Similar trends can be seen for all  $[\text{M}_2(\mu\text{-ER}_n)_2(\text{CO})_8]^{0/2-}$  complexes enumerated in Table II.

**Table II. Calculated Energies (eV) Accompanying Electron Transfer and Structural Change in  $[M_2(\mu\text{-EPh}_n)_2(\text{CO})_8]^{0/1-2-}$  Complexes<sup>a</sup>**

<i>Complex</i>		<i>E1</i>	<i>S1</i>	<i>ET1</i>	<i>E2</i>	<i>S2</i>	<i>ET2</i>	$\Delta ET$
Mo <sub>2</sub> S <sub>2</sub>	$\Delta G_{\text{gas}}$	-2.440	-0.529	-2.969	0.674	-0.950	-0.276	2.693
	$\Delta G_{\text{solv}}$	-1.379	0.029	-1.350	-4.208	0.125	-4.083	-2.733
	$\Delta G_{\text{sol}}$	-3.819	-0.500	<b>-4.319</b>	-3.534	-0.825	<b>-4.359</b>	<b>-0.040</b>
W <sub>2</sub> S <sub>2</sub>	$\Delta G_{\text{gas}}$	-2.282	-0.561	-2.843	0.672	-0.925	-0.253	2.590
	$\Delta G_{\text{solv}}$	-1.390	0.029	-1.361	-4.217	0.149	-4.068	-2.707
	$\Delta G_{\text{sol}}$	-3.672	-0.532	<b>-4.204</b>	-3.545	-0.776	<b>-4.321</b>	<b>-0.117</b>
Mo <sub>2</sub> P <sub>2</sub>	$\Delta G_{\text{gas}}$	-1.628	-0.591	-2.219	0.946	-0.833	0.113	2.332
	$\Delta G_{\text{solv}}$	-1.277	0.000	-1.277	-3.742	0.039	-3.703	-2.426
	$\Delta G_{\text{sol}}$	-2.905	-0.591	<b>-3.496</b>	-2.796	-0.794	<b>-3.590</b>	<b>-0.096</b>
W <sub>2</sub> P <sub>2</sub>	$\Delta G_{\text{gas}}$	-1.490	-0.718	-2.208	0.862	-0.871	-0.009	2.199
	$\Delta G_{\text{solv}}$	-1.303	0.011	-1.292	-3.783	0.010	-3.773	-2.481
	$\Delta G_{\text{sol}}$	-2.793	-0.707	<b>-3.500</b>	-2.921	-0.861	<b>-3.782</b>	<b>-0.282</b>

<sup>a</sup> Calculations performed with Jaguar 5.5; B3LYP/LACVP\*\*//B3LYP/cc-pVTZ(-f)

Table III summarizes the results of the DFT calculations and compares them with experimental observations. The calculated two-electron redox potentials  $E_{\text{calc}}^{0'}$  differ from experimental values  $E_{\text{obs}}^{0'}$  by no more than 200 mV, which is within the approximate 150 mV uncertainty of the computational method that we observed in our previous study (34). More important than the absolute value of the redox potential, however, is the magnitude of potential inversion, quantified as  $\Delta E_{\text{calc}}^{0'}$  in Table III. In all cases our calculations successfully capture the potential inverted scenarios, with  $\Delta E_{\text{calc}}^{0'}$  in agreement with experimental values within ~100 mV, which is remarkable considering that this “error” corresponds to little more than a 2 kcal mol<sup>-1</sup> discrepancy between theory and experiment for the free energy. Some systematic error cancellations can be expected in these calculations. Moreover, our model calculations correctly predict sulfido-bridged complexes to be more easily reduced than phosphido-bridged species by  $\geq 0.5$  V and confirm the trend of little metal dependence among redox potentials mentioned above. The differences in  $E_{\text{calc}}^{0'}$  arise largely from the electron affinities of the compounds [E1(gas), E2(gas)] and are in accord with the greater electronegativity of S versus P and the realization that bonding within the M<sub>2</sub>E<sub>2</sub> core is highly covalent. Indeed, our DFT calculations show the redox active  $\sigma^*$  MO to be 19-40% metal and 21-36% bridging ligand in character in the neutral

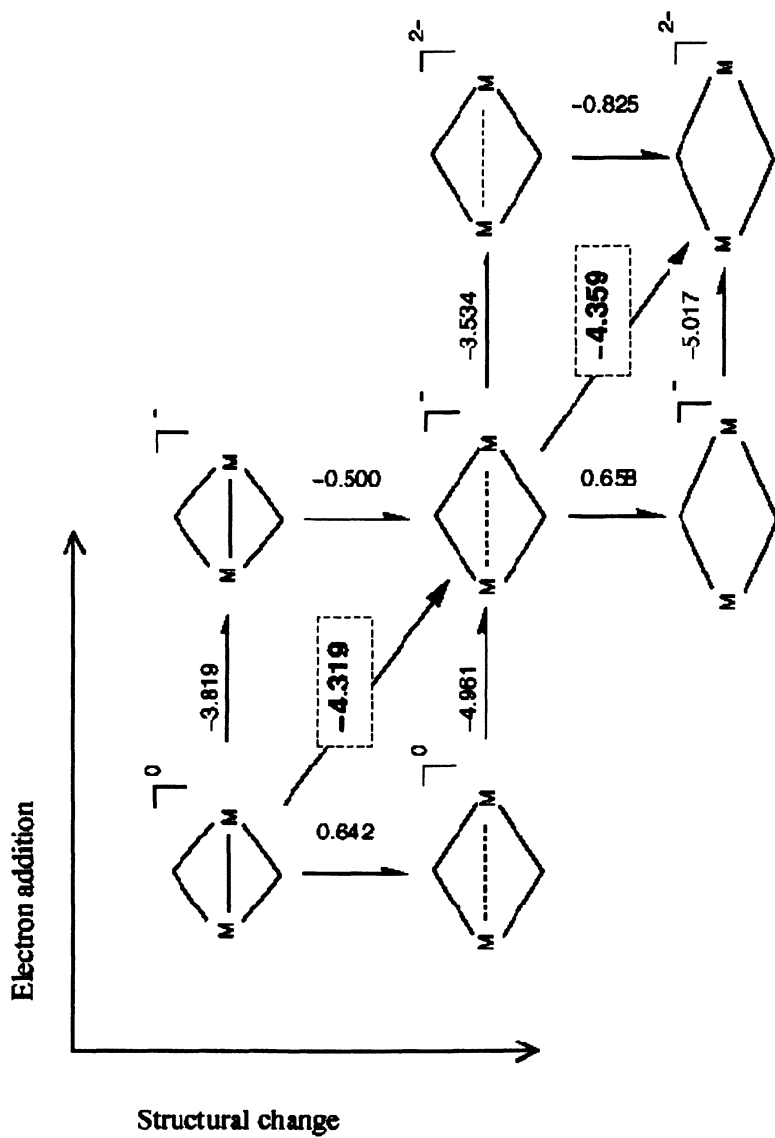


Figure 4. Free energy partitioning between electron addition and structural change for  $[\text{Mo}_2(\mu\text{-SPh})_2(\text{CO})_8]^{0/-2-}$  in acetonitrile. Energies are given in eV.

and monoanion forms versus 38-55% metal and 5-20% bridging ligand in the dianions.

**Table III. Comparison of Calculated and Experimental Potentials for  $[M_2(\mu\text{-EPh}_n)_2(\text{CO})_8]^{0/2-}$  Complexes**

<i>Complex</i>	$E_{obs}^{o/a}$ (V)	$E_{calc}^{o'}$ (V)	$\Delta E_{obs}^{o/a}$ (mV)	$\Delta E_{calc}^{o'}$ (mV)
Mo <sub>2</sub> S <sub>2</sub>	-0.13	-0.09	-	+40
W <sub>2</sub> S <sub>2</sub>	-0.15	-0.17	+20-230	+117
Mo <sub>2</sub> P <sub>2</sub>	-0.69	-0.89	+170	+96
W <sub>2</sub> P <sub>2</sub>	-0.65	-0.79	+180	+282

<sup>a</sup> From Table I.

In summary, our calculations demonstrate that potential inversion arises from the more favorable structural reorganization energy attending the second electron transfer. The driving force for potential inversion is diminished slightly by the solvation component attending structural change, because  $\Delta G(\text{solv})$  is less favorable for the slightly larger structurally relaxed forms. However, the net result is favorable, and the range of calculated  $\Delta E^{o'}$  values is in good agreement with experimental observations. Larger structural reorganization attending the second electron transfer also is consistent with the electrode kinetic measurements, which show  $k_{sh,2} < k_{sh,1}$ . Although difficult to quantify experimentally, greater potential inversion is predicted for Ph<sub>2</sub>P<sup>-</sup> versus PhS<sup>-</sup> bridged species (Table III). However, this thermodynamic offset produces a burden on apparent electron-transfer kinetics (9), which may contribute to the observation of smaller  $k_{sh,2}$  values for the phosphido-bridged complexes.

## Conclusions

Concerted two-electron transfer in ligand-bridged  $[M_2(\mu\text{-ER}_n)_2(\text{CO})_8]^{0/2-}$  [M = Mo, W; R<sub>n</sub>E<sup>-</sup> = RS<sup>-</sup>, R<sub>2</sub>P<sup>-</sup>] complexes has been probed by experimental and computational means. Reduction of the neutral compounds proceeds by electron transfer into a covalent metal-bridging ligand  $\sigma^*$  orbital that is accompanied by metal-metal bond cleavage and extensive structural reorganization within the M<sub>2</sub>E<sub>2</sub> core, which results in an inversion of the normal sequence of one-electron redox potentials. Occurrence of closely spaced or potential-inverted multielectron events within compositionally invariant, polynuclear metal-ligand centers is potentially characteristic of important bioinorganic redox processes

such as nitrogen fixation, hydrogen production, and oxygen evolution. For the model systems considered here, our calculations of the two-electron redox potentials are within 0.2 V of experimental values and accurately predict the following features: (1) a high degree of metal-ligand covalency in the redox active molecular orbital, (2) more negative potentials for phosphido- versus sulfido-bridged complexes, (3) an inversion of potentials for consecutive one-electron transfer reactions resulting in observation of a single two-electron event, and (4) confirmation that greater structural reorganization attends the second electron-transfer reaction.

## References

1. Howard, J. B.; Rees, D. C. *Chem. Rev.* **1996**, *96*, 2965-2982.
2. Burgess, B. K.; Lowe, D. J. *Chem. Rev.* **1996**, *96*, 2983-3012.
3. Adams, M. W. W.; Stiefel, E. I. *Science* **1998**, *282*, 1842-1843.
4. Evans, D. J.; Pickett, C. J. *Chem. Soc. Rev.* **2003**, *32*, 268-275.
5. Yachandra, V. K.; Sauer, K.; Klein, M. P. *Chem. Rev.* **1996**, *96*, 2927-2950.
6. Eisenberg, R.; Gray, H. B. *Inorg. Chem.* **2008**, *47*, 1697-1699.
7. Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
8. Evans, D. H.; Hu, K. *J. Chem. Soc., Faraday Trans.* **1996**, *92*, 3983-3990.
9. Evans, D. H. *Acta Chem. Scand.* **1998**, *52*, 194-197.
10. Kraiya, C.; Evans, D. H. *J. Electroanal. Chem.* **2004**, *565*, 29-35.
11. Pierce, D. T.; Geiger, W. E. *J. Am. Chem. Soc.* **1992**, *114*, 6063-6073.
12. DiMaio, A.-J.; Rheingold, A. L.; Chin, T. T.; Pierce, D. T.; Geiger, W. E. *Organometallics* **1998**, *17*, 1169-1176.
13. Jude, H.; Bauer, J. A. K.; Connick, W. B. *J. Am. Chem. Soc.* **2003**, *125*, 3446-3447.
14. Zhuang, B.; McDonald, J. W.; Schultz, F. A.; Newton, W. E. *Organometallics* **1984**, *3*, 943-945.
15. Smith, D. A.; Zhuang, B.; Newton, W. E.; McDonald, J. W.; Schultz, F. A. *Inorg. Chem.* **1987**, *26*, 2524-2431.
16. Fernandes, J. B.; Zhang, L. Q.; Schultz, F. A. *J. Electroanal. Chem.* **1991**, *297*, 145-161.
17. Hill, M. G.; Rosenhein, L. D.; Mann, K. R.; Mu, X. H.; Schultz, F. A. *Inorg. Chem.* **1992**, *31*, 4108-4111.
18. Uhrhammer, D.; Schultz, F. A. *J. Phys. Chem. A* **2002**, *106*, 11630-11636.
19. Rees, D. C.; Howard, J. B. *Curr. Opin. Chem. Biol.* **2000**, *4*, 559-566.
20. Coucouvanis, D.; Han, J.; Moon, N. *J. Am. Chem. Soc.* **2002**, *124*, 216-224.
21. Ibrahim, S. K.; Vincent, K.; Gormal, C. A.; Smith, B. E.; Best, S. P.; Pickett, S. J. *Chem. Commun.* **1999**, 1019-1020.

22. Musgrave, K. B.; Liu, H. I.; Ma, L.; Burgess, B. K.; Watt, G.; Hedman, B.; Hodgson, K. O. *J. Biol. Inorg. Chem.* **1998**, *3*, 344-352.
23. Capon, J.-F.; Gloaguen, F.; Schollhammer, P.; Talarmin, J. *Coord. Chem. Rev.* **2005**, *249*, 1664-1676.
24. Gloaguen, F.; Lawrence, J. D.; Rauchfuss, T. B.; Bernard, M.; Rohmer, M.-M. *Inorg. Chem.* **2002**, *41*, 6573-6582.
25. Darchen, A.; Mousser, H.; Patin, H. *J. Chem. Soc., Chem. Commun.* **1988**, 968-970.
26. Collman, J. P.; Rothrock, R. K.; Finke, R. G.; Moore, E. J.; Rose-Munch, F. *Inorg. Chem.* **1982**, *21*, 146-156.
27. Cheah, M. H.; Borg, S. J.; Bondin, M. I.; Best, S. P. *Inorg. Chem.* **2004**, *43*, 5635-5644.
28. Windhager, J.; Rudolph, M.; Bräutigam, S.; Görls, H.; Weigand, W. *Eur. J. Inorg. Chem.* **2007**, 2748-2760.
29. Capon, J.-F.; Ezzaher, S.; Gloaguen, F.; Pétilion, F. Y.; Schollhammer, P.; Talarmin, J.; Davin, T. J.; McGrady, J. E.; Muir, K. W. *New J. Chem.* **2007**, *31*, 2052-2064.
30. Ezzaher, S.; Capon, J.-F.; Gloaguen, F.; Pétilion, F. Y.; Schollhammer, P.; Talarmin, J. *Inorg. Chem.* **2007**, *46*, 9863-9872.
31. Chong, D.; Georgakaki, I. P.; Mejia-Rodriguez, R.; Sanabria-Chincilla, J.; Soriaga, M. P.; Darenbourg, M. Y. *Dalton Trans.* **2003**, 4158-4163.
32. Felton, G. A. N.; Glass, R. S.; Lichtenberger, D. L.; Evans, D. H. *Inorg. Chem.* **2006**, *45*, 9181-9184.
33. Cheah, M. H.; Tard, C.; Borg, S. J.; Liu, X.; Ibrahim, S. K.; Pickett, C. J.; Best, S. P. *J. Am. Chem. Soc.*, **2007**, *129*, 11085-11092.
34. Baik, M.-H.; Freisner, R. A. *J. Phys. Chem. A* **2002**, *106*, 7407-7412.
35. Baik, M.-H.; Ziegler, T.; Schauer, C. K. *J. Am. Chem. Soc.* **2000**, *122*, 9143-9154.
36. Baik, M.-H.; Schauer, C. K.; Ziegler, T. *J. Am. Chem. Soc.* **2002**, *124*, 11167-11181.
37. Yang, X.; Baik, M.-H. *J. Am. Chem. Soc.* **2006**, *128*, 10833-10839.
38. Ryan, M. D. *J. Electrochem. Soc.* **1978**, *125*, 547-555.
39. Rudolph, M.; Reddy, D. R.; Feldberg, S. *Anal. Chem.* **1994**, *66*, 589A-600A.
40. Connelly, N. G.; Geiger, W. E. *Chem. Rev.* **1996**, *96*, 877-910.
41. Bard, A. J.; Faulkner, L. R. *Electrochemical Methods. Fundamentals and Applications*; Second edition; John Wiley & Sons: New York, 2001.
42. Yu, Q. Y.; Salhi, C. A.; Ambundo, E. A.; Heeg, M. J.; Ochrymowycz, L. A.; Rorabacher, D. B. *J. Am. Chem. Soc.* **2001**, *123*, 5720-5729.
43. Bradbury, J. R.; Schultz, F. A. *Inorg. Chem.* **1986**, *25*, 4408-4416.
44. Zanello, P. *Coord. Chem. Rev.* **1988**, *87*, 1-54.
45. Gerdil, R. *J. Chem. Soc. B* **1966**, 1071-1075.
46. Santhanam, K. S. V.; Bard, A. J. *J. Am. Chem. Soc.* **1968**, *90*, 1118-1122.

47. Harkins, S. B.; Mankad, N. P.; Miller, A. J. M.; Szilagyi, R. K.; Peters, J. C. *J. Am. Chem. Soc.* **2008**, *130*, 3478-3485.
48. Rhee, Y. M.; Head-Gordon, M. *J. Am. Chem. Soc.* **2008**, *130*, 3878-3887.
49. Gamelin, D. R.; Randall, D. W.; Hay, M. T.; Houser, R. P.; Mulder, T. C.; Canters, G. W.; de Vries, S.; Tolman, W. B.; Hedman, B.; Hodgson, K. O.; Solomon, E. I. *J. Am. Chem. Soc.* **1998**, *120*, 5246-5263.
50. Mealli, C.; Proserpio, D. M. *J. Chem. Educ.* **1990**, *67*, 399-402.
51. Jaguar 5.5, Schrödinger, Inc., Portland, OR, 2003.
52. Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648-5652.
53. Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098-3100.
54. Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 270-283.
55. Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, *82*, 284-298.
56. Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299-310.
57. Dunning, T. H., Jr. *J. Chem. Phys.* **1989**, *90*, 1007-1023.
58. Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, *100*, 11775-11788.
59. Edinger, S. R.; Cortis, C.; Shenkin, P. S.; Friesner, R. A. *J. Phys. Chem. B* **1997**, *101*, 1190-1197.
60. Friesner, R. A.; Murphy, R. B.; Beachy, M. D.; Ringnalda, M. N.; Pollard, W. T.; Dunitz, B. D.; Cao, Y. X. *J. Phys. Chem. A* **1999**, *103*, 1913-1928.
61. Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, *99*, 2161-2200.
62. Tomasi, J.; Mannucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999-3094.
63. Reiss, H.; Heller, A. *J. Phys. Chem.* **1985**, *89*, 4207-4213.



## Chapter 11

# Metal-Mediated Peptide Assembly

### From Discrete Molecular Species to Large-Scale Morphologies

**Jing Hong, Olesya A. Kharenko, Mikhail V. Tsurkan,  
and Michael Y. Ogawa\***

**Department of Chemistry and Center for Photochemical Sciences,  
Bowling Green State University, Bowling Green, OH 43403**

Two classes of peptide-based nanoassemblies have been prepared by using metal coordination to direct the spatial orientation of self-assembling peptide ligands. Small-scale nanoassemblies have been prepared by binding metal ions to the hydrophobic faces of amphipathic polypeptides. This promotes formation of peptide assemblies having discrete molecular structures with dimensions on the order of several nanometers. In addition, large-scale nanoassemblies have been prepared by binding metal complexes to the solvent-exposed surfaces of the self-assembling peptide structures. This binding provides a type of non-covalent cross-link to these systems which enables the formation of globular and/or fibrillar structures having dimensions on the orders of tens to hundreds of nanometers.

Much inspiration for the design of biological nanoassemblies draws from the folding properties of proteins and nucleic acids in which combinations of hydrogen-bond, electrostatic, and hydrophobic interactions are used to drive the formation of discrete molecular architectures. In proteins, this can be demonstrated by the way different polypeptide sequences fold into the common

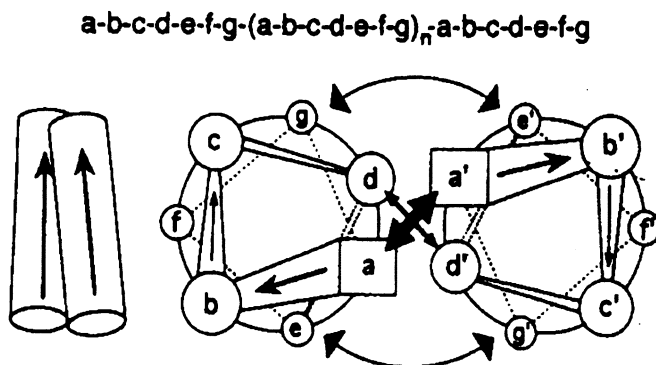
secondary structure motifs of  $\alpha$ -helices,  $\beta$ -sheets, and  $\beta$ -turns, and how these elements may in turn self-oligomerize into higher order structures such as  $\alpha$ -helical coiled-coils and  $\beta$ -sheet fibrils. Similarly, the base-pairing interactions of nucleic acids can result in the formation of a diverse array of structures including double-stranded helices, loops and junctions. Based on this knowledge, growing interest has focused on the fabrication of new types of bio-inspired materials using peptides, amino acids, and/or nucleic acids as self-assembling building blocks for supramolecular structures. Indeed, the ease of synthesis, relative stability, and functional variability of these systems make them particularly attractive for such use and an increasingly diverse array of self-assembling biological molecules has been studied for future application in such fields as microelectronics, tissue engineering, and medicine (1-4).

Our group has been investigating the design of biological nanostructures built from inorganic coordination compounds and self-assembling peptide structures (5-12). In particular, this work is attempting to exploit the directional bonding properties of inorganic transition metal compounds in ways that can direct the spatial orientation of self-assembling peptide ligands to create a new class of nanometer scale metal-peptide assemblies. It is anticipated that combining these different approaches towards molecular self-assembly will increase the degree of structural diversity currently available to other types of bio-inspired nanoassemblies, and may afford greater opportunity to control their three-dimensional structures. Indeed, as will be shown below, recent results from our group demonstrate the successful design of both small and large-scale metal-peptide nanostructures whose morphologies range from discrete molecular assemblies (*i.e.*, multihelix bundles) to large-scale peptide fibrils and globules.

### The $\alpha$ -Helical Coiled-Coil

The metal-peptide nanoassemblies studied in our laboratory are all based on peptide sequences that have been designed to self-assemble into two-stranded  $\alpha$ -helical coiled-coils. This important protein structural motif exists as a noncovalent supercoiling of two separate  $\alpha$ -helices that is stabilized by a specific packing of the hydrophobic side-chains of the component peptide strands (13). An examination of native coiled-coils shows that such interactions can arise from amino acid sequences that are based on a seven residue heptad repeat, (abcdefg)<sub>n</sub>, in which hydrophobic amino acids typically occupy positions "a" and "d" of the heptad, hydrophilic residues fill positions "b", "c", and "f", and oppositely-charged residues may occupy positions "e" and "g" in order to form stabilizing inter-chain salt bridges (Figure 1). Thus, the structural motif of  $\alpha$ -

helical coiled-coils has been shown to provide an attractive and versatile route towards the design of self-assembling peptide structures.



*Figure 1. Amino acid sequence consisting of multiple heptad repeats (top); schematic cartoon of a parallel coiled-coil (left); and a helical wheel diagram showing the view down helical axes of the coiled-coil (right).*

As will be discussed below, our group has worked to augment the inherent self-assembly properties of  $\alpha$ -helical coiled-coils by adding metal-binding sites to: 1) their hydrophobic cores in order to construct “small-scale” metal-peptide nanoassemblies having dimensions on the order of several nanometers (5, 8-11), and 2) their solvent-exposed surfaces for their construction of “large-scale” assemblies (6, 7, 12).

## Small-Scale Metal-Peptide Assemblies

### **Cd(II) Binding in the Hydrophobic Core Produces a 2-Stranded Coiled-Coil**

This section describes some work done by our group to incorporate metal-binding sites into the hydrophobic core of two-stranded  $\alpha$ -helical coiled-coils. The peptide sequence employed in this study was based on the (IEALEGK) heptad repeat which has been extensively used by our group in mechanistic electron-transfer studies (14-16). However, the sequence was modified to contain the Cys-X-X-Cys metal binding domain in a way that might create a thiolate-rich metal-binding domain within the hydrophobic core of the coiled-

coil structure. This new peptide, called C16C19-GGY, thus has the following sequence: Ac-K(IEALEGK)<sub>2</sub>(CEACEGK)-(IEALEGK)-GGY-amide which places cysteine residues at both the “a” and “d” positions of the third heptad repeat. The C16C19-GGY peptide was purified by reverse-phase HPLC and characterized by MALDI mass-spectroscopy. Figure 2 shows that the circular dichroism (CD) spectrum of the C16C19-GGY apopeptide consists of a large, negative maximum centered at 205 nm to indicate that it exists as a disordered random coil in aqueous solution. This was an unexpected result based on our previous experience with similar peptides and indicates that a two-stranded coiled-coil is unable to accommodate the presence of four cysteinyl side-chains within its hydrophobic core. However, the results presented in Figure 2 also show that the CD spectrum of C16C19-GGY changes substantially upon the addition of CdCl<sub>2</sub> (11). Negative maxima are now observed at 208 and 222 nm to indicate the presence of  $\alpha$ -helices. Further, the intensity of these signals increase with increasing amounts of Cd(II) added to solution, reaching a limiting value of  $\theta_{222} = -17,500 \text{ deg cm}^2 \text{ dmol}^{-1}$  which corresponds to a helical content of *ca.* 50%. Under these conditions, an ellipticity ratio of  $[\theta_{222}] / [\theta_{208}] = 0.97$  was observed which falls within the range generally regarded to indicate the presence of a coiled-coil structure (13). C16C19-GGY thus undergoes a significant random-coil to coiled-coil metal-induced folding event.

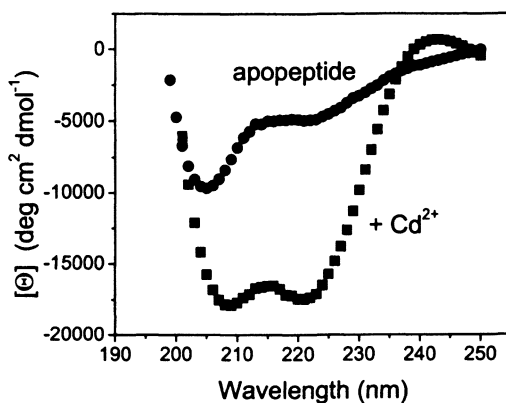
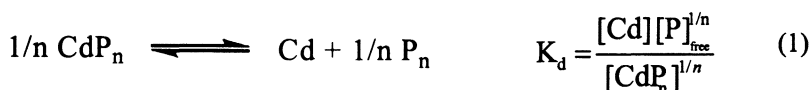


Figure 2. Circular dichroism spectra of C16C19-GGY before and after the addition of 1 equivalent of CdCl<sub>2</sub>.

The oligomerization state of this new metal-peptide assembly was studied by high performance size exclusion chromatography (11) which showed that the C16C19-GGY apopeptide and its Cd(II) adduct have apparent molecular masses of 3.3 and 5.3 kDa, respectively. These results indicate that the Cd(II) adduct

exists as a peptide dimer, and are consistent with our previously reported SDS-PAGE experiments (11). Recent dynamic light scattering measurements performed in our laboratory further confirm this observation.

The metal-peptide stoichiometry of the dimeric Cd peptide was studied by UV-Vis spectroscopy (11) as an absorption band at 238 nm is observed upon addition of Cd(II) to the peptide which is assigned to the ligand-to-metal charge-transfer (LMCT) transition of the newly formed Cd-S bonds. A Job plot demonstrated that the complex consists of 2 peptides and 1 metal ion. These results were supported by spectrophotometric titrations analyzed according to the following equilibrium (1) to yield  $n = 2$  and  $K_d = 0.65 \pm 0.08 \mu\text{M}$ .



In summary, the results show that the C16C19-GGY peptide undergoes a dramatic metal-induced folding process from a monomeric random coil to the organized structure of a two-stranded coiled-coil upon the binding of Cd(II).

### **A Luminescent 4-Stranded Coiled-coil is Produced Upon Cu(I) Binding in the Hydrophobic Core**

To further investigate the nature of the metal-induced protein folding properties of C16C19-GGY, we sought to determine if similar effects could be observed upon the binding of other soft metal ions to this peptide. As will be described below, a particularly interesting situation was observed in the case of Cu(I) addition (9, 10).

The addition of Cu(I) to a solution of C16C19-GGY produces an intense ( $\phi = 0.053$ ) ambient temperature luminescence centered at 600 nm and having an excitation maximum at 330 nm. This luminescence can be quenched by the addition of either ferricyanide, oxygen, or urea to indicate, respectively, that the emitting species is associated with the reduced Cu(I) state, has significant triplet character, and is quenched upon exposure to bulk solvent. It is noted that similar photoluminescent properties have been reported for Cu(I) derivatives of the cysteine rich metal-binding protein metallothionein (17, 18) as well as for the Cox17 copper chaperone (19), and the copper responsive transcription factors ACE1 (20). However, a common feature of this group of luminescent Cu(I) proteins is that they all contain a polynuclear copper(I) cluster buried within their hydrophobic core (21). As such, this observation is consistent with earlier work by Ford and co-workers who studied the photochemistry of Cu(I) compounds and found that photoluminescence occurs in polynuclear clusters for which metal-metal bonding interactions can stabilize the cluster-centered

photoexcited-state (22). A common feature of these systems is an extremely large (*ca.* > 200 nm) Stokes shifts in their emission spectra which is strikingly similar to that observed for Cu(I)/C16C19-GGY.

Thus, the observation of a 600 nm luminescence from Cu(I)/C16C19-GGY suggests that this metal-peptide assembly must also contain a polynuclear Cu(I) cluster. Circular dichroism studies showed that the addition of Cu(I)/C16C19-GGY results in a pronounced random-coil to coiled-coil conformational change of the peptide similar to that observed for Cd(II) addition. However, in contrast to the Cd-peptide assembly analytical ultracentrifugation experiments showed that the Cu(I)/C16C19-GGY adduct exists as a peptide tetramer. The metal binding stoichiometry of the Cu(I)/C16C19-GGY adduct was determined by monitoring the various spectral changes that can be observed upon the addition of Cu(I) to the peptide. Interestingly, the results indicate that four equivalents of Cu(I) are present in the peptide tetramer which is in contrast to the behavior previously observed for the Cd(II) protein.

The above observations of the metal-dependent assembly properties of C16C19-GGY were not anticipated from either the sequence of the apo-peptide, which was expected to form two-stranded coiled-coils, nor from the previously observed behavior of the Cd(II) assembly. These results, while not yet understood, therefore provide an important demonstration of how the conformational properties of this family of small-scale metal-peptide nanoassemblies are governed by the specific nature of their metal ion components. It is of further interest to note that the Cu(I)/C16C19-GGY system was shown to possess very unusual photophysical properties including bimolecular photoinduced electron-transfer reactions occurring in the inverted Marcus regime as discussed below (8).

### **Observation of Inverted Marcus Behavior for Collisional Electron-Transfer Reactions**

The strong room temperature luminescence of Cu(I)/C16C19-GGY suggests that this metal-peptide assembly might function as a photoinduced electron-transfer agent. The emission lifetime was thus measured and found to follow double exponential decay kinetics in which  $\tau_S = 1.1 \mu\text{s}$  and  $\tau_L = 7.7 \mu\text{s}$ . The source of these two emission components is presently unknown, but it is noted that the two components are of equal intensities. It is speculated that the samples studied might contain two separate types of metal-peptide assemblies, each having slightly different conformations and emission lifetimes.

Importantly, both lifetime components of Cu(I)/C16C19-GGY were found to be shortened in the presence of the electron-acceptor  $[\text{Ru}(\text{NH}_3)_6]^{3+}$  (8). It was further noted that the relative amplitudes of the fast and slow emission decay

components of this assembly remain approximately equal at all quencher concentrations studied. Thus, the two lifetimes components were separately analyzed according to the Stern-Volmer equation in which  $\tau^0$  and  $\tau$  are

$$\tau^0 / \tau = 1 + k_q \tau^0 [Q] \quad (2)$$

the emission lifetimes of either the short or long emission component measured in the absence ( $\tau^0$ ) or presence ( $\tau$ ) of the quencher, and  $k_q$  is the rate constant for the bimolecular quenching reaction (eq 2). Linear Stern-Volmer behavior was observed to demonstrate that emission quenching arises from a purely collisional mechanism. No evidence is seen for a ground-state protein-quencher complex. Linear Stern-Volmer behavior is also observed when the different members of a related series of ruthenium(III) ammine complexes are used to quench the Cu(I) protein emission. Significantly, in the cases where L = pyridine or 3,5-lutidine the emission quenching is accompanied by an increased transient absorption at 400 nm corresponding to the MLCT bands of the reduced  $[\text{Ru}(\text{NH}_3)_5\text{L}]^{2+}$  species to show that the quenching process occurs by an electron-transfer mechanism.

The thermodynamic driving force for these photoinduced electron-transfer reactions were determined by eq 3 in which  $E_{1/2}(\text{Cu(II)/Cu(I)}) = 0.343 \text{ V vs. NHE}$  is the reversible ground-state reduction potential of the copper site as:

$$\Delta G_{et}/nF = [E_{1/2}(\text{Cu(II)/Cu(I)}) - E^{0-0}(*\text{Cu(I)})] - E_{1/2}(\text{A/A}^-) \quad (3)$$

determined by redox potentiometry,  $E^{0-0}(*\text{Cu(I)}) = 2.04 \text{ eV}$  is the 0-0 excitation energy of the electron-donor as measured by 77K emission experiments, and  $E_{1/2}(\text{A/A}^-)$  is the ground state reduction potential of the various ruthenium ammine electron-acceptors. Significantly, Figure 3 shows that the values of

$$k_{et} = \sqrt{\frac{4\pi^3}{h^2 \lambda k_B T}} H_{DA}^2 \exp\left(\frac{-(\Delta G^0 + \lambda)^2}{4\lambda k_B T}\right) \quad (4)$$

both  $k_S$ , the quenching constant for the short lifetime component, and  $k_L$ , is the quenching constant for the long component, show a modest, albeit distinct decrease in magnitude with increasing driving force indicating the occurrence of inverted Marcus behavior for these reactions. The source of the scatter in the data is not presently known but might be related to different binding behavior of

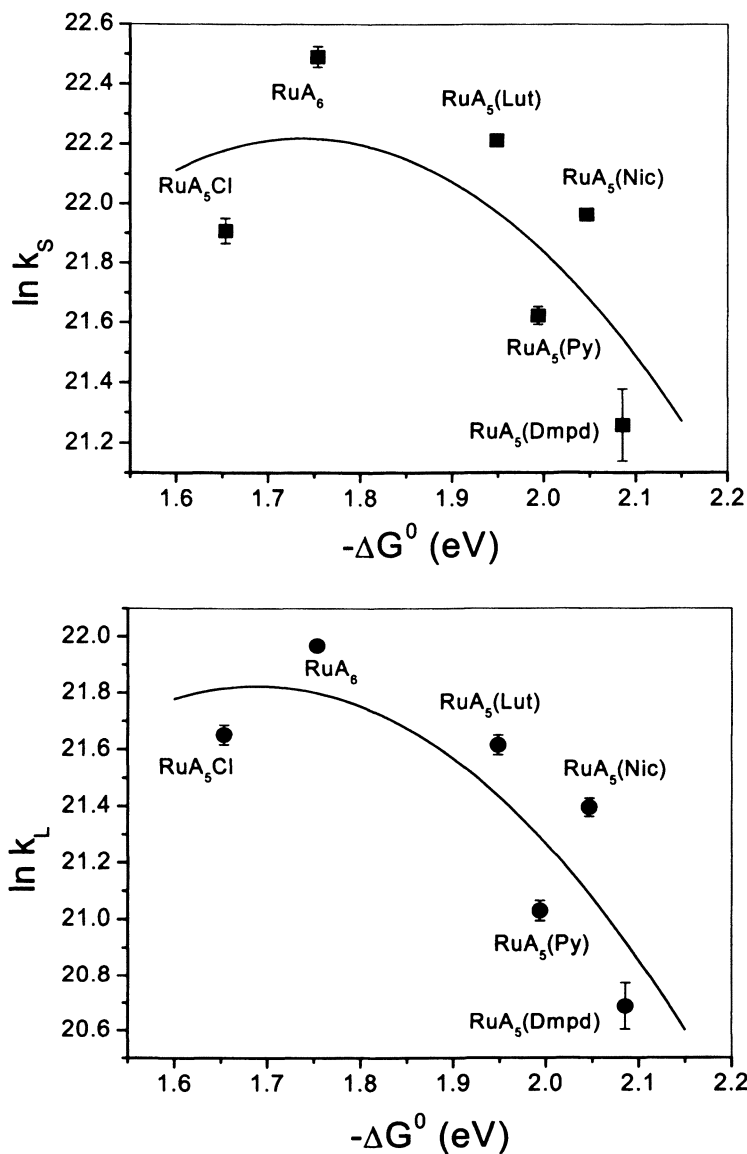


Figure 3. Free energy plot of  $k_{obs}$  values for the short (top) and long (bottom) components to the emission lifetime. Error bars represent the deviation from the mean of triplicate experiments. The solid lines are fits of the data to eq 4.



the various quenchers with Cu(I)/C16C19-GGY. Nonetheless, Figure 3 shows that the free energy dependence of both  $k_S$  and  $k_L$  can be fit to the Marcus equation (eq 4) to yield values of  $\lambda = 1.7 \pm 0.1$  eV and  $H_{DA} = 4.8 \pm 0.4$  cm<sup>-1</sup> for the short component, and  $\lambda = 1.7 \pm 0.1$  eV and  $H_{DA} = 3.9 \pm 0.4$  cm<sup>-1</sup> for the long component. A cross relation analysis of the  $\lambda$ -values yields a value of  $\lambda_{22} = 2.2$  eV for Cu(I)/C16C19-GGY which is very similar to that reported for [Cu(phen)<sub>2</sub>]<sup>2+/+</sup> ( $\lambda = 2.4$  eV) (23), assuming  $\lambda_{11} = 1.2$  eV for the ruthenium pentammine acceptors (24). The large reorganization energy suggests that the one-electron oxidation of Cu(I)/C16C19-GGY produces a large conformational change of the multinuclear Cu(I) cluster in the 4-helix assembly.

The observation of inverted Marcus behavior in this system is unexpected because bimolecular rates of high driving-force ET reactions typically saturate at the diffusion limit according to the Rehm-Weller equation (25). Apparently, the reactions involving the Cu(I)/C16C19-GGY occur at rates that are slightly less than diffusion. Further insight into this behavior can be obtained by examining the results reported by Ford and co-workers (26, 27) for the related small molecule Cu(I) cluster, [Cu<sub>4</sub>I<sub>4</sub>(pyridine)<sub>4</sub>]. The one-electron excited-state reduction potential of this cluster was calculated to be  $E_{1/2}(\text{Cu(II)/Cu(I)*}) = -1.65$  V vs. NHE which is very similar to that of Cu(I)/C16C19-GGY, for which  $E_{1/2}(\text{Cu(II)/Cu(I)*}) = -1.76$  V vs. NHE. However, no evidence for inverted Marcus behavior was observed for the [Cu<sub>4</sub>I<sub>4</sub>(pyridine)<sub>4</sub>] system as its high driving force quenching constants did indeed saturate at  $k_q = 10^{10}$  M<sup>-1</sup>s<sup>-1</sup>. Since the quenching reactions involving both the Cu-metalloprotein and isolated Cu-cluster occur with comparable driving forces, the results suggest that the slower quenching rates of Cu(I)/C16C19-GGY are due to electronic effects.

It is thus speculated that the encapsulation of the emissive Cu(I)<sub>4</sub>S<sub>4</sub>(N/O)<sub>4</sub> cluster within the hydrophobic core of protein serves to prohibit close approach between the donor and acceptor sites which reduces their degree of electronic coupling and slows their rates of electron-transfer. In order to test this hypothesis, new Cu(I) metalloproteins have been recently designed in which the Cu(I) binding sites have been moved closer to the solvent-exposed N-terminus of the protein. It is speculated that such changes to the protein structure will result in faster photoinduced electron-transfer rates and yield Rehm-Weller behavior at high driving forces.

## Large-Scale Metal-Peptide Assemblies

The above results show how combining the principles of supramolecular coordination chemistry with those of *de novo* protein design can produce a new family of metal-peptide nanoassemblies in which metal coordination is used to promote the formation of  $\alpha$ -helical coiled-coils having discrete molecular

architectures. The following sections will describe how simple modification of this approach leads to the production of larger-scale assemblies having spatial dimensions on the order of tens to hundreds of nanometers.

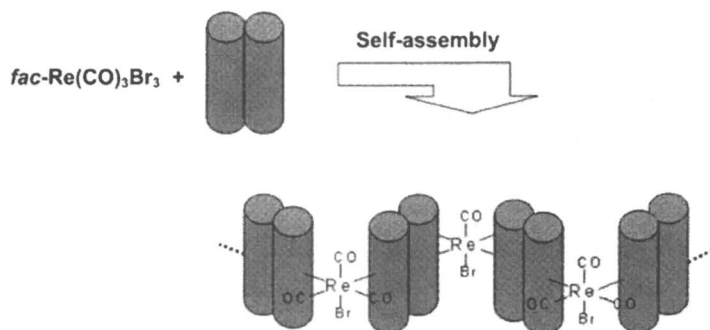
### Formation of Linear Metal-Peptide Chains

Encouraged by the above work, our group has been investigating a different route by which metal coordination can be used to create extended three-dimensional arrays of self-assembled peptides. Rather than designing metal-binding domains within the hydrophobic interior of a coiled-coil motif, this new approach places metal-binding 4-pyridyl alanine (Pal) residues at the solvent-exposed surfaces of these structures and subsequent metal-binding is used to crosslink these structures into large-scale metal-peptide nanoassemblies (Figure 4a).

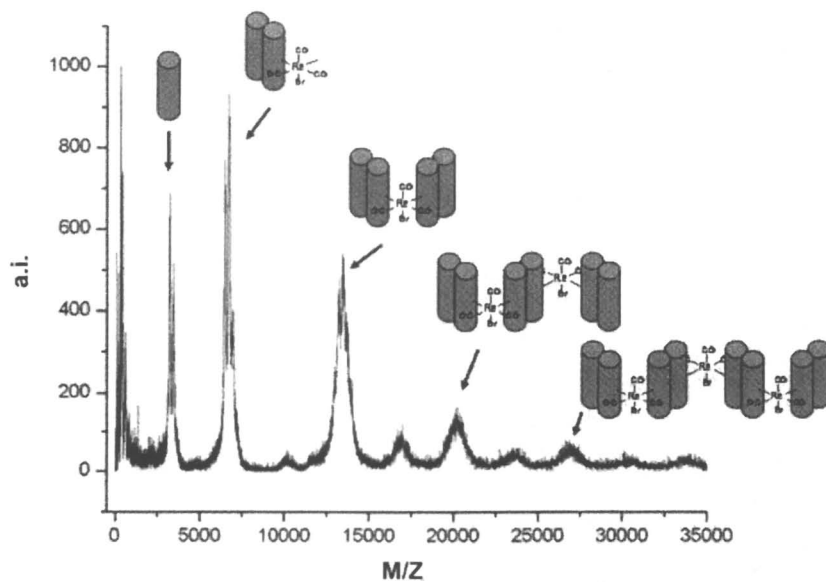
This work began with our preparation of a 30-residue polypeptide having the sequence: K(IEALEGK)(IEALEPaK)(IEACEGK)(IEALEGK)G which places the non-natural amino acid 4-pyridyl alanine (Pal) at position 14 of the sequence; this is the most solvent-exposed position of the second heptad repeat (12). An additional feature of this sequence is that it incorporates a cysteine residue at position 19 (a hydrophobic position) in order to allow for the introduction of an inter-chain disulfide bond crosslink to ensure that this peptide forms a stable two-stranded coiled-coil.

To create the desired metal-peptide assemblies, samples of this crosslinked peptide were incubated in methanol at 37°C in the presence of a limiting amount of  $[\text{NEt}_4][\text{ReBr}_3(\text{CO})_3]$  (12). After reacting overnight, a gel-like substance appeared at the bottom of the reaction vessel which could be resolubilized upon shaking. The UV-Vis spectrum of the resolubilized reaction mixture was used to confirm the formation of the metal-peptide conjugate and a combination of SDS-PAGE and MALDI-MS experiments were performed to help characterize the ability of the metal-peptides to form supramolecular nanostructures. The metal-peptide reaction mixture was found to contain several different higher molecular-weight species, showing a progression of higher molecular weight metal-peptide assemblies containing from 1 to at least 4 coiled-coil subunits (Figure 4b). However, it is noted that the major species observed in both experiments corresponds to a Re center coordinated to only one coiled-coil dimer. This observation highlights the difficulty of coordinating two coiled-coil peptides onto a single metal center and is perhaps due to the steric congestion involved in this process. The additional peaks seen in the spectrum in Figure 4b are attributed to the coordination of non-crosslinked Pal14C19 monomers to some of the Re centers. It is thought that these incompletely formed coiled-coils can serve as chain termination sites during oligomer formation to further limit the ability of these systems to form larger assemblies.

(a)



(b)



**Figure 4.** Nanoassembly formation by metal coordination to the hydrophilic exterior of coiled-coil peptides (a) and MALDI-MS of the metal-peptide reaction mixture (b).

## Formation of Large-Scale Metal-Peptide Fibrils and Spheres

The results above gave encouraging evidence that inorganic coordination chemistry could indeed be used to create a new class of metal-peptide nanoassemblies. However, they also highlighted the importance of minimizing steric interactions when forming these structures. An alternate assembly strategy was therefore attempted in which Pal-containing peptide monomers (called AQ-Pal14) were first coordinated to the *cis* positions of a single Pt(en) center (en = ethylene diamine) in order to create a metal-peptide synthon unit, and the resulting synthon units were then allowed to self-assemble through the formation of non-covalent coiled-coils as shown in Figure 5. Formation of the metal-peptide synthons was monitored by SDS-PAGE and 2D  $^1\text{H}$  NMR, both of which showed no evidence of unreacted starting materials after a 7-day incubation (6). Interestingly, analysis of the reaction mixture under non-denaturing conditions showed evidence for the formation of extremely large metal-peptide nanostructures. HPSEC was then used to monitor the time-dependent evolution of these species starting from the presence of peptide monomers at the start of the reaction, to a mixture of structures having a broad distribution of molecular weights after 3 days, to the presence of metal-peptide assemblies with  $\text{MW} > 600$  kDa traveling at the column void volume. These results were confirmed by static light scattering (SLS) experiments which showed the final reaction product to have a molecular mass of  $\text{MW} = (7 \pm 4) \times 10^6$  Da and a radius of  $R_z = 18 \pm 4$  nm. The morphology of the Pt-peptide assemblies was characterized by both transmission electron microscopy (TEM) and atomic force microscopy (AFM). Both techniques showed the presence of two distinct classes of large-scale assemblies (Figure 6). The dominant population consists of globular structures having radii in the range of 15 – 25 nm and heights in the range of 5 – 10 nm (6).

These structures have the appearance of collapsed spheres and it is noted that their diameters are consistent with the hydrodynamic radii determined by the static light scattering experiments. In addition to these spherical structures, Figure 6 shows that the assembly of the Pt-peptide building block units also forms a smaller population of one-dimensional fibrils which have lengths of the order  $10^2$  nm and thicknesses ranging from 4 – 10 nm. These dimensions are consistent with a model in which the  $\alpha$ -helical coiled-coils are lying flat on the solid substrate and chain propagation occurs through metal complexation.

At present, it is not understood why two distinct populations of large-scale assemblies were produced by the reaction of Pt(en) with AQ-Pal14. However, it is noted that under carefully controlled growth conditions, Moore *et al.* (28) found that large (micron-sized) peptide globules were the primary kinetic intermediates in the process of amyloid fiber formation. These workers thus proposed an *in vitro* growth model for amyloid production which involves the preliminary formation of large peptide globules which then partition into either amorphous precipitates, inert materials, and/or aggregation-competent species

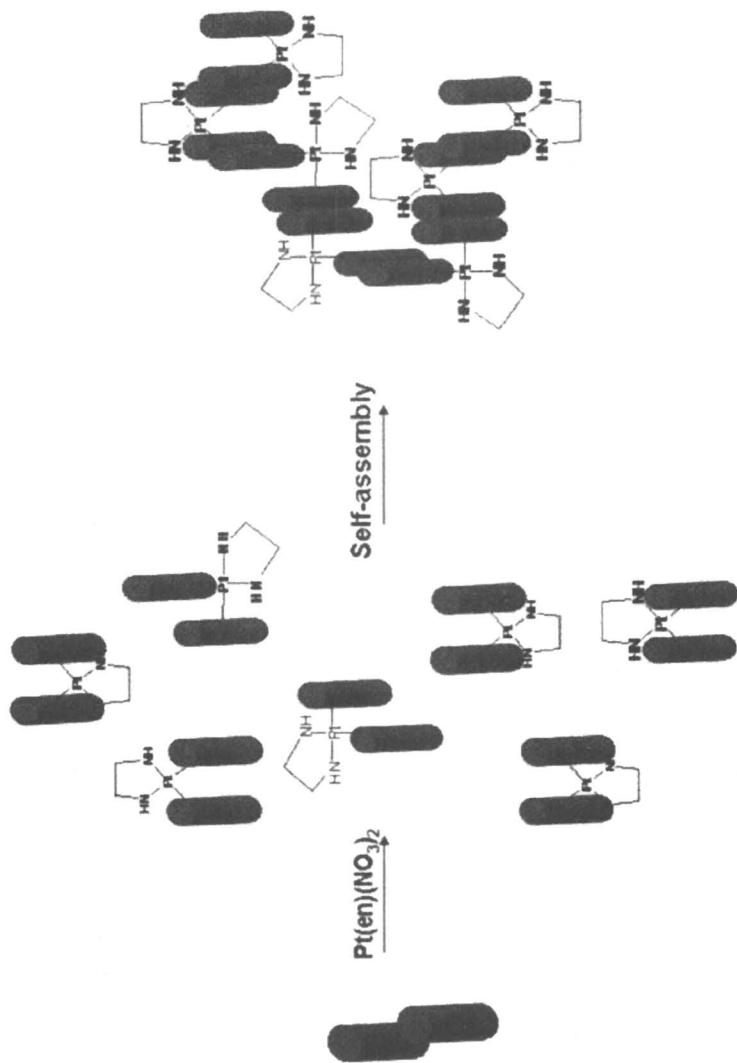
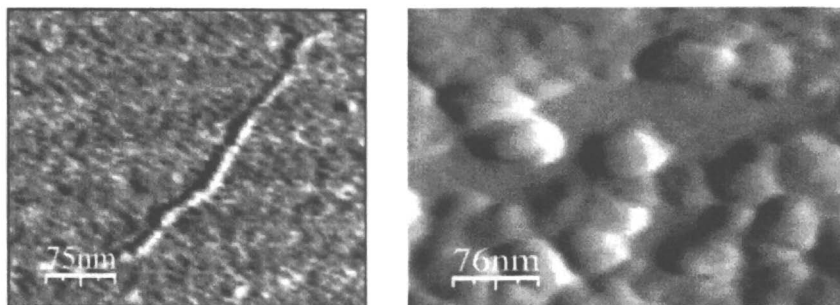


Figure 5. Scheme showing the formation of the metal-peptide building block units via complexation of two  $\alpha$ -helical P<sub>114</sub> peptides (cylinders) to a Pt(en) center and their subsequent self-assembly through the formation of  $\alpha$ -helical coiled-coils. (Reprinted with permission from reference 6. Copyright 2007 American Chemical Society.)



*Figure 6. Large-scale nanoassemblies formed by the reaction of the AQ-Pal14 peptide with  $Pt(en)^{2+}$ . (Reprinted with permission from reference 6. Copyright 2007 American Chemical Society.)*

that in turn can oligomerize to form fibrillar networks. Importantly, this growth model explicitly accounts for the presence of multiple morphological species in peptide samples involved in amyloid fiber formation, and is further consistent with the observed presence of micron-sized globules of A $\beta$  peptide in the brains of Alzheimer patients (29). As speculated in reference 6, a similar mechanism might be responsible for the simultaneous presence of the two different types of assemblies shown in Figure 6. Ongoing work in our group is thus geared towards developing methods of separating these different large-scale assemblies, further characterizing their conformational properties, and determining their metal content.

## Conclusions

Our group has found that two different classes of metal-peptide nanoassemblies can be prepared by combining the principles of protein folding with those of supramolecular coordination chemistry. Small-scale nanoassemblies can be prepared by binding metal complexes to the hydrophobic faces of amphipathic polypeptides. This promotes formation of peptide assemblies having discrete molecular structures (*i.e.*, metal-bridged coiled-coils) with spatial dimensions on the order of several nanometers. In addition, large-scale nanoassemblies can be prepared by binding metal ions to the solvent-exposed surfaces of previously-assembled peptide structures. This provides non-covalent cross-links to these systems which enable them to form globular and/or fibrillar structures having dimensions on the orders of tens to hundreds of nanometers.

## Acknowledgements

This work was supported by the National Science Foundation grant CHE-0455441, Petroleum Research Fund grant 40790-AC, and the National Institutes of Health grant GM61171.

## References

1. Sotiropoulou, S.; Sierra-Sastre, Y.; Mark, S. S.; Batt, C. A. *Chem. Mat.* **2008**, *20*, 821-834.
2. Gazit, E. *Chem. Soc. Rev.* **2007**, *36*, 1263-1269.
3. Woolfson, D. N.; Ryadnov, M. G. *Curr. Opin. Chem. Biol.* **2006**, *10*, 559-567.
4. Rajagopal, K.; Schneider, J. P. *Curr. Opin. Struct. Biol.* **2004**, *14*, 480-486.
5. Mukherjee, M.; Zhu, X.; Ogawa, M. Y. *Inorg. Chem.* **2008**, *47*, 4430-4432.
6. Tsurkan, M. V.; Ogawa, M. Y. *Biomacromol.* **2007**, *8*, 3908-3913.
7. Tsurkan, M. V.; Ogawa, M. Y. *Inorg. Chem.* **2007**, *46*, 6849-6851.
8. Hong, J.; Kharenko, O. A.; Fan, J. F.; Xie, F.; Petros, A. K.; Gibney, B. R.; Ogawa, M. Y. *Angew. Chem., Int. Ed.* **2006**, *45*, 6137-6140.
9. Hong, J.; Kharenko, O. A.; Ogawa, M. Y. *Inorg. Chem.* **2006**, *45*, 9974-9984.
10. Kharenko, O. A.; Kennedy, D. C.; Demeler, B.; Maroney, M. J.; Ogawa, M. Y. *J. Am. Chem. Soc.* **2005**, *127*, 7678-7679.
11. Kharenko, O. A.; Ogawa, M. Y. *J. Inorg. Biochem.* **2004**, *98*, 1971-1974.
12. Tsurkan, M. V.; Ogawa, M. Y. *Chem. Commun.* **2004**, 2092-2093.
13. Hodges, R. S. *Biochem. Cell Biol.* **1996**, *74*, 133-154.
14. Fedorova, A.; Chaudhari, A.; Ogawa, M. Y. *J. Am. Chem. Soc.* **2003**, *125*, 357-362.
15. Kornilova, A. Y.; Wishart, J. F.; Ogawa, M. Y. *Biochemistry* **2001**, *40*, 12186-12192.
16. Kornilova, A. Y.; Wishart, J. F.; Xiao, W. Z.; Lasey, R. C.; Fedorova, A.; Shin, Y. K.; Ogawa, M. Y. *J. Am. Chem. Soc.* **2000**, *122*, 7999-8006.
17. Salgado, M. T.; Bacher, K. L.; Stillman, M. J. *J. Inorg. Biochem.* **2007**, *12*, 294-312.
18. Roschitzki, B.; Vasak, M. *J. Inorg. Biochem.* **2002**, *7*, 611-616.
19. Heaton, D. N.; George, G. N.; Garrison, G.; Winge, D. R. *Biochemistry* **2001**, *40*, 743-751.
20. Casas-Finet, J. R.; Hu, S.; Hamer, D.; Karpel, R. L. *Biochemistry* **1992**, *31*, 6617-6626.
21. Stillman, M. J. *Coord. Chem. Rev.* **1995**, *144*, 461-511.
22. Ford, P. C.; Cariati, E.; Bourassa, J. *Chem. Rev.* **1999**, *99*, 3625-3647.

23. Gray, H. B.; Malmstrom, B. G.; Williams, R. J. P. *J. Biol. Inorg. Chem.* **2000**, *5*, 551-559.
24. Brown, G. M.; Sutin, N. *J. Am. Chem. Soc.* **1979**, *101*, 883-892.
25. Rehm, D.; Weller, A. *Isr. J. Chem.* **1970**, *8*, 259-271.
26. Tran, D.; Ryu, C. K.; Ford, P. C. *Inorg. Chem.* **1994**, *33*, 5957-5959.
27. Dossing, A.; Ryu, C. K.; Kudo, S.; Ford, P. C. *J. Am. Chem. Soc.* **1993**, *115*, 5132-5137.
28. Moore, R. A.; Hayes, S. F.; Fischer, E. R.; Priola, S. A. *Biochemistry* **2007**, *46*, 7079-7087.
29. Barghorn, S.; Nimmrich, V.; Striebinger, A.; Krantz, C.; Keller, P.; Janson, B.; Bahr, M.; Schmidt, M.; Bitner, R. S.; Harlan, J.; Barlow, E.; Ebert, U.; Hillen, H. *J. Neurochem.* **2005**, *95*, 834-847.



## Chapter 12

# Understanding the Biological Chemistry of Mercury Using a *de novo* Protein Design Strategy

Vincent L. Pecoraro<sup>\*</sup>, Anna F. A. Peacock, Olga Iranzo<sup>1</sup>,  
and Marek Luczkowski<sup>2</sup>

Department of Chemistry,  
University of Michigan, Ann Arbor, MI 48109–1055  
Current addresses: <sup>1</sup>Instituto de Tecnologia Química e Biológica,  
Avda. Da República (EAN), 2781–901, Oeiras, Portugal; and  
<sup>2</sup>University of Wrocław, 14 Joliot-Curie, 50–383 Wrocław, Poland

Hg(II) is a well known toxin that has a high affinity for protein thiolate functional groups. While an area of significance, a dearth of literature exists on the chemistry of Hg(II) with thiolate containing proteins. In this chapter we demonstrate the design of proteins that complex Hg(II) in linear, trigonal planar, and tetrahedral environments. Physical techniques such as <sup>199</sup>Hg NMR, <sup>199m</sup>Hg PAC and UV-vis spectroscopy to characterize Hg(II) sites in proteins are also described along with the application of our understanding of Hg(II) interactions with designed proteins to address the binding of Hg(II) in protein sites such as MerA (2-coordinate), MerR (3-coordinate) and Hg substituted rubredoxin (4-coordinate). Finally, knowledge from this system is used to predict the chemistry of Hg(II) bound forms of Hah1 at high pH.

Although Lewis Carroll's famed hatter is the most notorious literary character thought inflicted by mercury poisoning, the phrase "mad as a hatter" is known to predate *Alice in Wonderland* by over fifty years. Despite the fact that the toxic effects of mercurials have been known for at least two hundred years, this element still remains a concern in modern society. Chief among the culprits

for dangerous mercuric complexes are alkyl mercurials such as dimethyl mercury (1). Of particular interest has been the release of mercury through industrial stacks, the burning of fossil fuels, the presence of mercury in vaccines and silver amalgam fillings and the accumulation of mercury in top of the food chain fish such as tuna (2).

Given the broad deleterious health effects of this element, it is somewhat remarkable that so little is truly known about the specific biochemical sites of mercury intervention. Clearly Hg(II) has a high affinity for sulfhydryl groups of proteins; however, there are numerous targets that have been inferred, but not necessarily proven as the origin of the toxic effects (3, 4). Certainly, methyl mercury compounds, being more lipid soluble, easily cross the blood brain barrier, but once across, it is unclear which specific biochemical targets are impacted (*e.g.*, tubulin, acetylcholinesterase, *etc.*). The same can be said for immunologic suppression by mercury, the effects are well established but the molecular targets remain elusive. Detoxification in humans is associated with metallothioneins which sequester the metal (5), whereas bacteria reduce Hg(II) to the less toxic Hg(0) using a reductase (6).

In order to understand the biochemistry of mercury more completely, it is important to describe the chemistry of this element with sulfhydryl donors in the most common coordination environments. Ideally, one would examine mercuric complexes within a construct that was basically invariant, but which allowed for preparation of mercury compounds in the most common structures. The desired coordination modes include linear (or slightly bent) for 2-coordinate complexes, trigonal planar (or slightly T-shaped) for 3-coordinate compounds and tetrahedral for 4-coordinate complexes, Figure 1. Unfortunately, there are no known native systems that allow for this diversity of metal coordination geometry. However, given advances in peptide synthesis and the prediction of

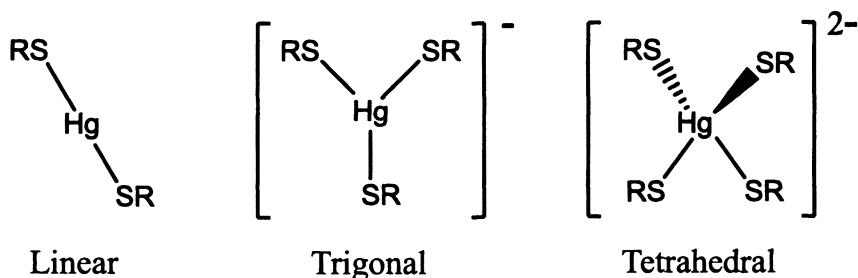


Figure 1. Common coordination modes for Hg thiolate complexes.

protein structure, one should be able to exploit the emerging field of *de novo* protein design in order to prepare well defined scaffolds that should sequester mercury into each of these desired structures. This article describes how this objective may be met.

## Preparation of 2-, 3- and 4-Coordinate Hg(II) Thiolate Complexes

We have shown that the *de novo* designed TRI peptide family associates into specific  $\alpha$ -helical aggregates depending on the pH, sequence composition and peptide length. The parent peptide TRI, with the linear sequence Ac-G(LKALEEK)<sub>4</sub>G-NH<sub>2</sub>, exploits the concept of a *heptad repeat* to place hydrophobic residues (in this case, leucine) in the 1<sup>st</sup> and 4<sup>th</sup> positions of a seven amino acid repeating sequence. The consequence of this sequence is that hydrophobes will occupy one face of an  $\alpha$ -helix while the remaining positions can contain residues that both solubilize the peptide and stabilize a specific aggregation state through specific salt bridges. In the case of TRI, one forms at low pH values (< 5.5) predominantly two-stranded coiled coils and at pH values > 5.5 parallel, three-stranded coiled coils (7, 8). Similar pH dependent aggregation state behavior is observed for the related peptides, BABY and GRAND, which have the same repeated heptad sequence but which are either shorter (three heptads) or longer (five heptads), respectively, than TRI. The stability of the aggregate is enhanced by lengthening the peptide (9). Thus, at the same concentrations, BABY peptides may be unassociated and unfolded, TRI peptides may be partially associated and folded and GRAND peptides fully associated and folded. In order to introduce metal binding sites into these peptides, one or more leucine residues are replaced by cysteine residues making, for example, TRIL16C (see Table I for this and related sequences). When associated as a two-stranded coiled coil, one can prepare a scaffold presenting two cysteines to a metal, whereas the three-stranded coiled coil architecture provides a trigonal plane of three cysteinyl sulfur atoms. Alternative constructs include di-substituted peptides in which adjacent layers of leucines are substituted by cysteine residues yielding peptides that provide four or six sulfur donor atoms depending on whether they form two- or three-stranded coiled coils.

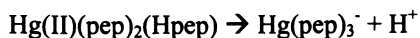
Our initial foray into mercury binding with these peptides utilized TRIL16C and TRIL12C. Based on sequence, these peptides differ only by the site of cysteine substitution with TRIL16C having a cysteine in the 1<sup>st</sup> or a position and TRIL12C placing cysteine in the 4<sup>th</sup> or d position of a heptad. As we will see in a

Table I. Derivatives of TRI Peptides

<i>Peptide</i>	<i>Sequence</i>
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH <sub>2</sub>
TRIL9C	Ac-G LKALEEK CKALEEK LKALEEK LKALEEK G-NH <sub>2</sub>
TRIL12C	Ac-G LKALEEK LKACEEK LKALEEK LKALEEK G-NH <sub>2</sub>
TRIL16C	Ac-G LKALEEK LKALEEK CKALEEK LKALEEK G-NH <sub>2</sub>
TRIL19C	Ac-G LKALEEK LKALEEK LKACEEK LKALEEK G-NH <sub>2</sub>
TRIL9CL12C	Ac-G LKALEEK CKACEEK LKALEEK LKALEEK G-NH <sub>2</sub>
TRIL12CL16C	Ac-G LKALEEK LKACEEK CKALEEK LKALEEK G-NH <sub>2</sub>

moment, this slight shift will have a profound impact on trigonal metal coordination; however, we will first explore the 2-coordinate species formed by both systems (7, 8). At pH values below 5.5, both peptides prefer to aggregate as two-stranded coiled coils. Thus, it was expected, and subsequently confirmed through spectroscopic studies, that 2-coordinate bis thiolato Hg(II) compounds would exist regardless of the peptide:metal ratio. More interesting was the behavior at higher pH values where the peptide has a three-stranded coiled coil aggregation state preference. Three distinct behaviors were observed. If the stoichiometry of peptide to mercury was 2:1, only a two-stranded coiled coil with 2-coordinate Hg(II) was observed for either peptide. This observation demonstrated that under these conditions, the stability of Hg(II) in a 2-coordinate structure exceeded that of the bundle to retain its preferred three-stranded coiled coil aggregation state. If the ratio of peptide to mercury was increased to 3:1 and the pH was maintained between 5.5 and 7, the predominant species for both peptides was a three-stranded coiled coil that contained a two coordinate Hg(II) species (10). This suggested that the third cysteine of the aggregate remained protonated and uncoordinated to the Hg(II). If the pH of these solutions was now raised (to 8.6 for TRIL16C and 9.5 for TRIL12C) one obtained three-stranded aggregates that contained fully 3-coordinate, trigonal planar Hg(II) (9). We were delighted with these observations as these mercurated peptides provided the first peptidic system to bind Hg(II) as a trigonal thiolato complex in aqueous solution.

Further analysis demonstrated that an equilibrium existed between 2- and 3-coordinate Hg(II) within the three-stranded coiled coil according to the equation:



Quantitative assessment of these equilibria for TRIL16C and TRIL12C yielded  $pK_a$ 's of 7.8 and 8.5, respectively (11). These data provided the first inkling that metal binding to cysteine residues was dependent on the **a** vs. **d** substitution pattern of these peptides. Subsequent studies with Cd(II) and Pb(II) derivatives have further demonstrated this principle (12, 13). It should be noted that the  $pK_a$  for the 2- to 3-coordinate conversion is not dependent on the stability of the peptide aggregate (e.g., BABYL9C, TRIL9C and GRANDL9C have the same  $pK_a$  values of  $7.6 \pm 0.2$ ) (9, 14), but instead tracks with the **a** or **d** site substitution pattern. A summary of these various equilibria is given in Figure 2.

In a subsequent study, we first demonstrated that the simple addition of small thiolates such as  $\beta$ -mercaptoethanol to existing  $Hg(pep)_2$  did not lead to trigonal thiolato mercury complexes (15). Thus, the capability of forming these desired structures is a direct consequence of peptide recognition. We next addressed the relationship between the peptide self association affinity and the ability to form trigonal Hg(II) species. Using eleven different peptides, we found a linear free energy correlation between the self association affinity to form a three-stranded coiled coil and the energy of the formation of  $Hg(pep)_3^-$  (15). These studies conclusively demonstrated that our ability to complex Hg(II) as a trigonal complex was a direct consequence of the designed peptide recognition and that we could titrate metal peptide affinities by controlling the self association affinities of the apo-peptides. We next challenged our design strategy by testing Hg(II) complexation in solutions containing mixtures of TRIL2WL9C and TRIL2WL23C. In theory, mixtures of these peptides could form anti-parallel three-stranded coiled coils that yielded trigonal Hg(II) (e.g.,  $Hg(TRIL2WL9C)_2(TRIL2WL23C)$ ). In fact, no such heterotrimeric complexes were observed by circular dichroism or  $^{199}Hg$  NMR spectroscopies (16). These studies demonstrated conclusively that we could define the coordination environment of the mercury ion while retaining exquisite control of protein aggregation state and orientation.

With 2- and 3-coordinate Hg(II) complexes in hand, our next objective was to prepare  $Hg(SR)_4^{2-}$  complexes. Because our protein design does not allow for the formation of four-stranded aggregates, we needed to shift our strategy to peptides containing dual cysteine substitution. The two peptides that we chose to examine were TRIL12CL16C and TRIL9CL12C. These constructs allowed us to compare a  $-Cys_a-X-X-X-Cys_a-$  binding motif found in TRIL12CL16C to the more common  $-Cys_a-X-X-Cys_d-$  sequence of native proteins found in TRIL9CL12C. Studies of these dual substituted peptide systems are complicated by perturbing two adjacent leucine layers which stabilize the peptide aggregates. Thus, rather than achieving pure two-stranded or three-stranded coiled coils at  $pH = 8.5$ , we observe a mixture of species (17). Nonetheless, addition of Hg(II) leads to well defined structures. In the case of TRIL12CL16C one obtains two-stranded coiled coils that yield spectroscopic parameters that are the hallmark of

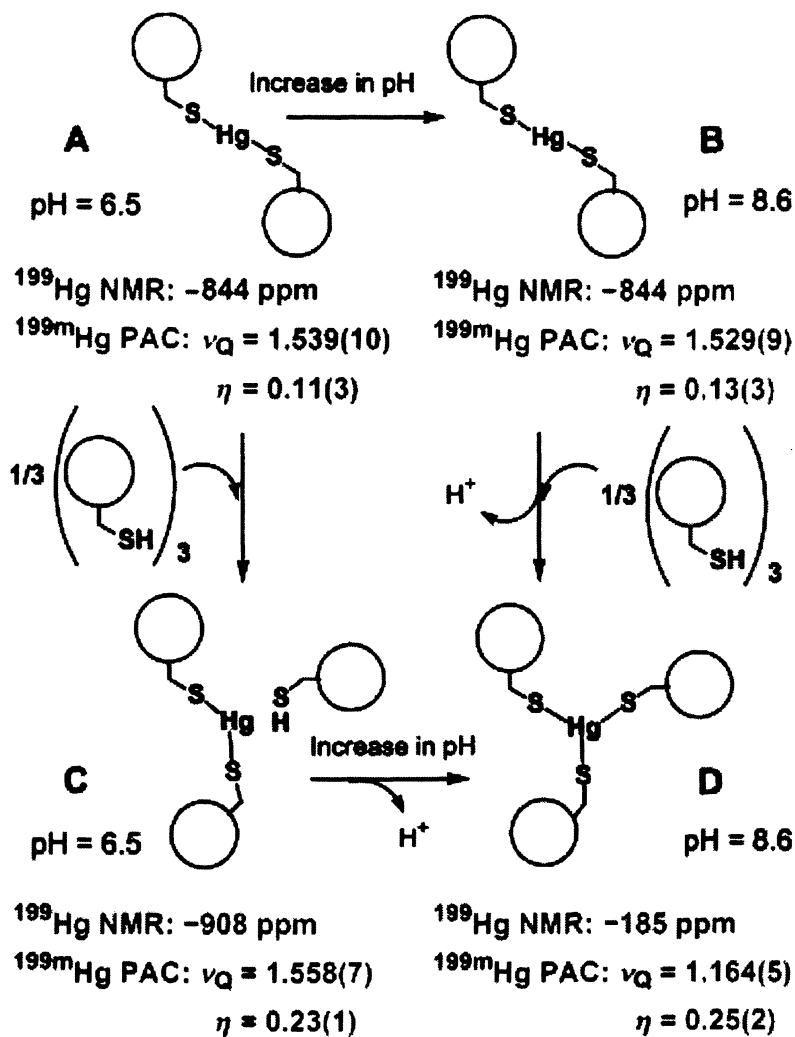


Figure 2. Species present at different TRIL9C/Hg<sup>II</sup> ratios and pH values (reproduced with permission from reference 10. Copyright Wiley-VCH Verlag GmbH & Co. KGaA).

distorted tetrahedral Hg(II) (*vide infra*). While solutions of TRIL9CL12C also yield tetrahedral Hg(II) complexes, mixtures of two- and three-stranded coiled coils exist simultaneously. These observations once again underscore the observation that cysteine substitution into **a** and **d** sites is important, in this case because of aggregate stability as well as topological differences that arise from having different numbers of amino acids in the intervening region between cysteines.

## Spectroscopic Characterization of Hg(II) Thiolate Complexes

A significant advantage for working with Hg(II) is that this ion can be probed by numerous spectroscopic techniques, each of which providing complementary information about structure over a broad range of timescales. In this section we will describe the spectroscopic signatures for Hg(II) as 2-, 3- or 4-coordinate species. The data provided in Tables II and III provide a useful compilation of parameters that should allow future workers to assign Hg(II) coordination geometries in biomolecules containing a homoleptic thiolate binding motif.

**Table II. Spectroscopic Values for Hg/TRILXC Complexes**

<i>Hg Coordination Mode</i>	$\lambda$ ( $\Delta\epsilon$ ) <sup>a</sup>	$pK_a$	$R_{HgS}$ ( $\text{\AA}$ ) <sup>b</sup>
Linear 2-coordinate	240 (2700) <sup>c</sup>	-	2.32 <sup>c</sup>
Trigonal 3-coordinate <b>a</b> site	247 (19200) <sup>d</sup> 265 (11900) <sup>d</sup> 295 (5800) <sup>d</sup>	7.6±0.2 <sup>e</sup>	2.44 <sup>d</sup>
Trigonal 3-coordinate <b>d</b> site	230 (21300) <sup>f</sup> 247 (15000) <sup>f</sup> 297 (5500) <sup>f</sup>	8.5±0.2 <sup>f</sup>	2.44 <sup>f</sup>
Linear 2-coordinate within a 3-stranded coiled coil	247 (2000) <sup>g</sup>	-	-
Tetrahedral 4-coordinate	230 (8100) <sup>h</sup> 289 (7100) <sup>h</sup>	-	-

<sup>a</sup>  $\lambda$  given in nm and  $\Delta\epsilon$  given as  $M^{-1} \text{ cm}^{-1}$ ; <sup>b</sup> Average Hg-S EXAFS bond lengths;

<sup>c</sup> Data for TRIL16C from Ref. 8; <sup>d</sup> Data for TRIL16C from Ref. 11;

<sup>e</sup> Data for TRIL9C from Ref. 14; <sup>f</sup> Data for TRIL12C from Ref. 11;

<sup>g</sup> Data for TRIL12C from Ref. 18; <sup>h</sup> Data for TRIL12CL16C from Ref. 17.

**Table III.  $^{199}\text{Hg}$  NMR and  $^{199\text{m}}\text{Hg}$  PAC Spectroscopic Values for Hg/TRILXC Complexes**

<i>Hg Coordination Mode</i>	$\delta$ ( $^{199}\text{Hg}$ ppm)	PAC ( $\nu_Q$ , $\eta$ ) <sup>a</sup>
Linear 2-coordinate	-844 <sup>b</sup>	1.53, 0.13 <sup>b</sup>
Trigonal 3-coordinate a site	-185 <sup>b</sup>	1.16, 0.25 <sup>b</sup>
Trigonal 3-coordinate d site	-316 <sup>c</sup>	-
Linear 2-coordinate within a 3-stranded coiled coil	-908 <sup>b</sup>	1.56, 0.23 <sup>b</sup>
Tetrahedral 4-coordinate	-500 <sup>d</sup>	-

<sup>a</sup>  $\nu_Q$  given in GHz and  $\eta$  is a unitless quantity; <sup>b</sup> Data for TRIL9C from Ref. 10;

<sup>c</sup> Data for TRIL19C from Ref. 10; and <sup>d</sup> Data for TRIL12CL16C from Ref. 17.

Despite having a closed shell  $d^{10}$  configuration, Hg(II) can have a rich ultraviolet absorption spectrum due to ligand to metal charge transfer (LMCT) excitations that occur at high energy. These LMCT transitions occur at successively lower energies as the Hg(II) coordination sphere accumulates additional thiolate ligands. Figure 3 compares these transitions for Hg(TRIL9C)<sub>2</sub>, Hg(TRIL9C)<sub>3</sub><sup>-</sup> and Hg(TRIL12CL16C)<sub>2</sub><sup>2-</sup> which represent Hg(SR)<sub>2</sub>, Hg(SR)<sub>3</sub><sup>-</sup> and Hg(SR)<sub>4</sub><sup>2-</sup> chromophores, respectively. The spectra shown correspond to both a site peptides; while, the UV spectra are slightly different when d site peptides are used, these perturbations are small compared to those observed for the different coordination numbers. The absorption maximum for Hg(SR)<sub>2</sub> species is always at shorter wavelength than 220 nm. In contrast, strong absorptions are observed for Hg(SR)<sub>3</sub><sup>-</sup> and Hg(SR)<sub>4</sub><sup>2-</sup> in the region between 225 and 340 nm. For Hg(SR)<sub>3</sub><sup>-</sup>, a  $\lambda_{\text{max}}$  is observed at 247 nm ( $\epsilon = 19,200 \text{ M}^{-1}\text{cm}^{-1}$ ) with shoulders at 265 nm ( $\epsilon = 11,900 \text{ M}^{-1}\text{cm}^{-1}$ ) and 295 nm ( $\epsilon = 5,800 \text{ M}^{-1}\text{cm}^{-1}$ ) (11). These absorption spectra are red shifted with a Hg(SR)<sub>4</sub><sup>2-</sup> chromophore ( $\lambda_{\text{max}} = 289 \text{ nm}$ ;  $\epsilon = 7,100 \text{ M}^{-1}\text{cm}^{-1}$ ) (17). Given the significant differences in these spectral signatures, the UV spectrum affords a good initial assessment of the Hg first coordination sphere in homoleptic thiolate complexes.

X-ray absorption spectroscopy is a powerful method to assess the structure of mercury thiolate complexes. X-ray Absorption Near Edge Structure (XANES) spectroscopy can cleanly differentiate Hg(0), Hg(I) and Hg(II) by the energy of the emission edge. Extended X-ray Absorption Fine Structure (EXAFS) spectroscopy provides information regarding the coordination number and Hg-S bond distances. Typically, the EXAFS spectrum can establish the coordination number (*i.e.*, number of sulfur atoms) to within 20%. More important, the bond length precision is 0.01 Å. Given that each increase of one



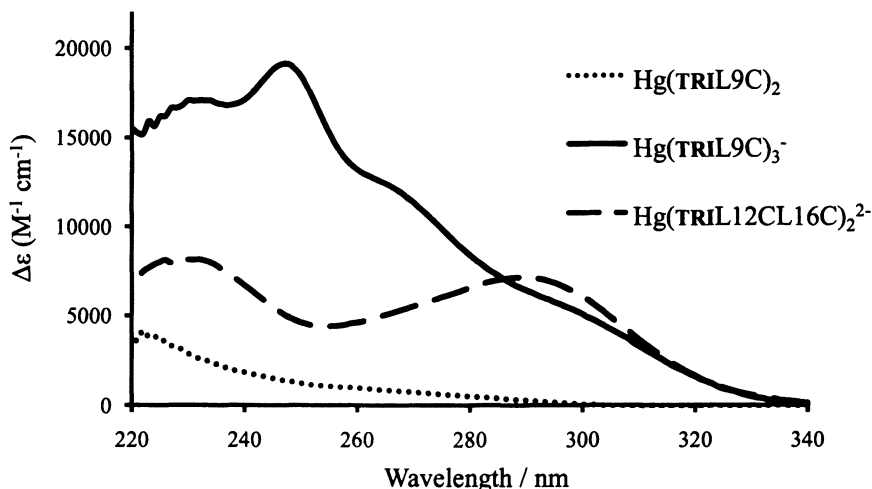


Figure 3. UV-visible spectra for  $\text{Hg}(\text{TRIL9C})_2$ ,  $\text{Hg}(\text{TRIL9C})_3^-$  and  $\text{Hg}(\text{TRIL12CL16C})_2^{2-}$  at pH 9.6 (References 7 and 17).

sulfur ligand to the coordination number causes an approximately 0.1 Å average increase in Hg-S distance. Thus, by precisely measuring the Hg-S distance in a biosample, one obtains an excellent estimate for the coordination number of the metal. As shown in Table II, the typical bond length for  $\text{Hg}(\text{SR})_2$  is 2.32 Å (8). As predicted, this value increases to 2.44 Å for  $\text{Hg}(\text{SR})_3^-$  (11). While we have not yet measured the  $\text{Hg}(\text{SR})_4^{2-}$  EXAFS spectrum, we estimate that the average Hg-S bond length for the tetrahedral complex should be ~2.50 Å. Unfortunately, EXAFS does not provide orientation of the sulfur ligands. Therefore, in the absence of an X-ray structure, other spectroscopic techniques are required to define further the coordination sphere of the Hg(II).

One of the most powerful tools in the modern chemical analysis arsenal is NMR spectroscopy. In addition to being diamagnetic, the  $^{199}\text{Hg}(\text{II})$  isotope benefits from having an  $I = \frac{1}{2}$  nuclear spin. With high sensitivity, a reasonable  $R_f$  (resonance frequency), a wide chemical shift range and relatively inexpensive access to the isotope,  $^{199}\text{Hg}$  NMR provides an attractive approach for clarifying the coordination environment and structure of Hg(II) complexes. The 1000 ppm range of chemical shifts allows a sensitivity that is unavailable with either UV or X-ray absorption spectroscopies to discriminate between different coordination modes and structures. In general, the furthest upfield shifts < -800 ppm correspond to 2-coordinate complexes whereas the most downfield shifts > -350 ppm are observed for 3-coordinate Hg(II) thiolate species. Tetrahedral Hg(II) compounds such as  $\text{Hg}(\text{TRIL12CL16C})_2^{2-}$  (-500 ppm) are observed between these two extremes (17). While these broad limits provide useful trends, one can

extract even more information by a detailed examination of related peptide systems. As an example,  $\text{Hg}(\text{TRIL9C})_2$  exhibits a single resonance at -844 ppm while  $\text{Hg}(\text{TRIL9C})_2(\text{HTRIL9C})$  has a -908 ppm signal (10). In both cases, UV and EXAFS spectroscopy predict 2-coordinate species, but certainly a greater than 60 ppm chemical shift difference suggests some perturbation in structure. We will explain the nature of these structural modifications below; however, it is useful to consider another example of the sensitivity of  $^{199}\text{Hg}$  NMR. If the pH of the solution of  $\text{Hg}(\text{TRIL9C})_2(\text{HTRIL9C})$  is raised from 6.5 to 8.6 there is a marked change in the UV signature, a corresponding increase in the Hg-S bond distance and the appearance of a new signal at -185 ppm indicating that a 3-coordinate  $\text{Hg}(\text{TRIL9C})_3^-$  was formed (14, 10). Comparison of this complex, formed from an a site peptide, to the Hg(II) complex of the corresponding d site peptide  $\text{Hg}(\text{TRIL12C})_3^-$  yields similar UV spectra but now the  $^{199}\text{Hg}$  NMR resonance occurs at -316 ppm, a 130 ppm upfield shift compared to  $\text{Hg}(\text{TRIL9C})_3^-$  (10). Examples of these  $^{199}\text{Hg}$  NMR spectra are shown in Figure 4. Clearly, NMR spectroscopy is reporting on structural differences that are less visible to the other spectroscopic techniques; however, the interpretation of these data requires the use of a more exotic spectroscopic technique.

Perturbed Angular Correlation (PAC) Spectroscopy measures the nuclear quadrupole interaction (NQI) of an appropriate metal nucleus, in this case  $^{199\text{m}}\text{Hg}$ , providing detailed information on the electronic and molecular structure of a metal site (19). Used in conjunction with  $^{199}\text{Hg}$  NMR spectroscopy, one can extract detailed structural and dynamic information about the desired chromophore (PAC typically works on the nanosecond timescale whereas NMR spectroscopy probes the millisecond time regime). Five NQI parameters are typically extracted from the fits to the PAC data, with two parameters  $\nu_Q$  and  $\eta$  being the most important for this discussion. The  $\nu_Q$  is used to evaluate the first coordination sphere ligands, while  $\eta$  provides information regarding the symmetry of the Hg environment. Given this introduction, we can revisit the two systems described by the  $^{199}\text{Hg}$  NMR experiments. The  $\text{Hg}(\text{TRIL9C})_2$  (-844 ppm) and  $\text{Hg}(\text{TRIL9C})_2(\text{HTRIL9C})$  (-908 ppm) both correspond to 2-coordinate "linear" species. The respective  $\nu_Q$  and  $\eta$  values are 1.539 GHz and 0.11 ( $\text{Hg}(\text{TRIL9C})_2$ ) and 1.558 GHz and 0.23 ( $\text{Hg}(\text{TRIL9C})_2(\text{HTRIL9C})$ ) (10). The closely similar  $\nu_Q$  values confirm a two coordinate structure; however, the variation in the asymmetry parameter  $\eta$  demonstrates that there is a significant change perturbation to the "linear" geometry depending on whether Hg(II) is in a two-stranded or three-stranded coiled coil. Since  $\eta$  has a range of values between 0 and 1, where 0 is pure axial symmetry, we conclude that the Hg environment in  $\text{Hg}(\text{TRIL9C})_2(\text{HTRIL9C})$  deviates more significantly from linearity. It appears that the two-stranded coiled coil allows for a highly symmetric linear Hg structure. In contrast, placing a linear Hg(II) within the three-stranded bundle causes a bending of the S-Hg-S angle. The origin of this effect is unknown but could be due to: A) the weak interaction of the third

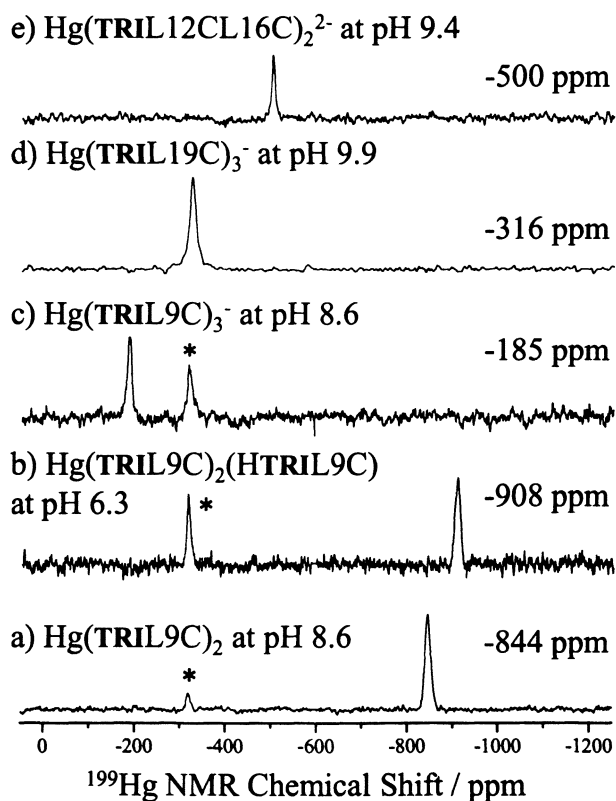


Figure 4.  $^{199}\text{Hg}$  NMR spectra of solutions containing the species  $\text{Hg}(\text{TRIL9C})_2$  (a);  $\text{Hg}(\text{TRIL9C})_2(\text{HTRIL9C})$  (b);  $\text{Hg}(\text{TRIL9C})_3^-$  (c);  $\text{Hg}(\text{TRIL19C})_3^-$  (d); and  $\text{Hg}(\text{TRIL12CL16C})_2^{2-}$  (e). Data from references 10 and 17. \* Corresponds to the external standard ( $\text{Hg}(\text{TRIL19C})_3^-$  at pH 9.6).

sulfhydryl atom with the  $\text{Hg}(\text{II})$ ; B) the averaging of the  $\text{Hg}(\text{II})$  environment by rapid transfer of the one remaining proton between the three sulfur atoms within the coiled coil; or C) structural deformation imposed by the twisting of the three-stranded coiled coil without the necessity of the third sulfur entering the metal coordination sphere. The second example from the  $^{199}\text{Hg}$  NMR section, in which we compare trigonal structures formed with **a** vs. **d** peptides, is presently under investigation. The  $\nu_Q$  and  $\eta$  values 1.164 GHz and 0.25 for  $\text{Hg}(\text{TRIL9C})_3^-$  are markedly different from those of the 2-coordinate species. We soon hope to collect data for  $\text{Hg}(\text{TRIL19C})_3^-$  to assess whether the  $\text{Hg}(\text{II})$  in a **d** site is more highly distorted than in an **a** site. We expect that this may be the case based on the models shown in Figure 5.

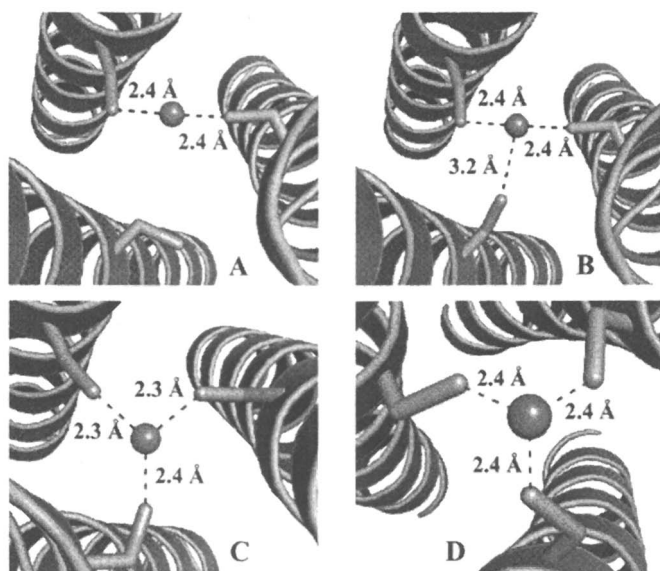


Figure 5. Models based on crystal structures<sup>20,21</sup> of Hg(II) binding within a coiled coil interior. A) Linear  $\text{HgS}_2$ , B) T-shaped  $\text{HgS}_3^-$  with reorientation of one Cys ligand and C) trigonal  $\text{HgS}_3^-$  with significant reorientation of Cys ligands, all in a d site, and D) trigonal  $\text{HgS}_3^-$  in an a site with no reorientation of Cys.

These structures have been derived based on the X-ray structures of the analogous Coil Ser (CS) peptides for  $\text{As}(\text{CSL9C})_3$  and  $\text{apo}(\text{CSL19C})_3$  (20, 21). It appears from these data that the a site provides a much more compact environment in which to encapsulate the Hg(II). We are hopeful that future crystallographic investigations with mercurated peptide will be able to assess this hypothesis. We also hope to collect data on our tetrahedral constructs, but based on previous literature values we anticipate that both the  $\nu_Q$  and  $\eta$  values for this geometry will be very small.

Considering all of these spectroscopic techniques together, one may summarize the properties that are expected for homoleptic mercury(II) thiolates in biological systems. Two coordinate complexes are expected to have featureless UV spectra above 220 nm,  $^{199}\text{Hg}$  NMR chemical shifts  $< -800$  ppm and  $\nu_Q$  and  $\eta$  values in the range 1.5-1.6 GHz and 0.11-0.25. Three coordinate species have rich UV signatures typically around 247 nm range with strong  $\epsilon$  coefficients. The  $^{199}\text{Hg}$  NMR chemical shifts  $> -350$  ppm and  $\nu_Q$  and  $\eta$  values in the range 1.1-1.2 GHz and  $\sim 0.25$ , respectively. Tetrahedral centers have red shifted UV spectra as compared to the 3-coordinate structures, have intermediate  $^{199}\text{Hg}$  NMR chemical shifts and are predicted to have low values for  $\nu_Q$  and  $\eta$ .

## Applications to Biological Systems

Given that we now have a considerable wealth of knowledge amassed for the spectroscopic properties of Hg(II), we are in a strong position to evaluate the binding of Hg(II) to a wide range of sulfur containing biomolecules. The first system we will consider is MerA, the mercuric ion reductase responsible for reduction of Hg(II) to Hg(0) using NADPH as a reductant. In 1991, Schiering *et al.* published the structure of this enzyme using Cd(II) as an analogue for Hg(II) (22). The authors identified four potential ligands to the Cd(II) (two tyrosines at  $\sim 3.1$  Å and two cysteines at 2.2 Å) and suggested that these tyrosines and possibly the intersubunit cysteines were bound to Hg(II). There are a few problems with this analysis. First, the ionic radius of oxygen is smaller than that of sulfur, Hg(II) is more thiophilic than Cd(II), and Hg(II) and Cd(II) have significantly different coordination number preferences. Thus, one might have predicted that thiolate coordination would be much more important for Hg(II) than Cd(II) and that the tyrosines, at such a long distance, would have little impact to the Hg(II). In 1999, Tröger examined MerA using  $^{199}\text{mHg}$  PAC and obtained  $\nu_Q$  and  $\eta$  values in the range 1.42 GHz and 0.15 (23). To our knowledge, there is no report of the  $^{199}\text{Hg}$  NMR for this protein; however, the NQI parameters alone are very revealing. These parameters are in reasonable agreement with a nearly linear, 2-coordinate bis(cysteinylyl) ligation of Hg(II) by MerA. The reduced  $\nu_Q$  probably reflects a slightly longer Hg(II)-SR distance in the protein.

Our second example is MerR, the protein that regulates the expression of the *mer* operon. Both O'Halloran and Walsh predicted that MerR contained the rare  $\text{Hg}(\text{SR})_3^-$  binding site (24, 25). Comparison of the spectroscopic features of MerR to those of our designed peptides is informative: MerR has a Hg-S distance of 2.43 Å (24); UV maxima at 240 nm ( $\epsilon = 16,620$ ), 260 nm ( $\epsilon = 11,150$ ) and 290 nm ( $\epsilon = 4,120$ ) (26);  $\nu_Q$  and  $\eta$  values 1.18 GHz and 0.25 (23); and a  $^{199}\text{Hg}$  NMR chemical shift of -106 ppm (27). Comparison to the corresponding values in Tables II and III for our peptides leaves little doubt that Hg(II) is bound in a slightly distorted trigonal planar environment by three cysteinylyl side chains.

Our last example will be that of 4-coordinate complexes typified by Hg(II) substituted rubredoxin (Rd). All the spectroscopic data taken together suggests that the mercurated protein (Hg-Rd) compares well to our model peptides, supporting the assignment of a  $\text{Hg}(\text{SR})_4^{2-}$  center. The Hg-Rd has a Hg-S distance of 2.534 Å (28);  $\lambda_{\text{max}} = 284$  nm ( $\epsilon = 20,000$ ), 257 nm ( $\epsilon = 12,000$ ) and 230 nm ( $\epsilon = 22,000$ ) (29);  $\nu_Q$  and  $\eta$  values 0.09 GHz and 0 (29); and a  $^{199}\text{Hg}$  NMR chemical shift of -241 ppm (30). What is most interesting about this comparison is that Hg-Rd incorporates the metal primarily within a  $\beta$ -sheet fold with the metal binding site formed within a peptidic loop whereas our peptides are  $\alpha$ -helical. This demonstrates that the parameters presented here are basically

invariant with respect to the secondary protein structure. Thus, we should be able to use a combination of these spectroscopies to unravel Hg(II) sites in unknown protein structure types. One system where the structure is well defined is the human copper metallochaperone Hah1. The X-ray structure of this mercurated protein at pH 6.5 is a homodimer which can be interpreted as encapsulating a Hg(II) ion in a trigonal environment with a fourth thiol closely associated, but unbound (31). An interesting question is what would be the structure of this system if the pH were raised to 8.5 or 9? We have shown that in designed peptides the last sulfur often requires a high pH to bind the metal. Our studies suggest that a combination of  $^{199}\text{Hg}$  NMR,  $^{199\text{m}}\text{Hg}$  PAC and UV spectroscopies may reveal whether Hg(Hah1)<sub>2</sub> actually forms a tetrahedral structure at higher pH.

### Acknowledgements

V.L.P. thanks the NIH (R01 ES0 12236) for support of this work and O.I. thanks the Margaret and Herman Sokol Foundation for a Postdoctoral Award.

### References

1. Davis, L. E.; Kornfield, M.; Mooney, H. S.; Fiedler, K. J.; Haaland, K. Y.; Orrison, W. W.; Cernichiari, E.; Clarkson, T. W. *Annal. Neurol.* **1994**, *35*, 680-688.
2. Vallee, B. L.; Ulmer, D. D. *Annu. Rev. Biochem.* **1972**, *41*, 91-128.
3. Rooney, J. P. K. *Toxicology* **2007**, *234*, 145-156.
4. Onyido, I.; Norris, A. R.; Buncel, E. *Chem. Rev.* **2004**, *104*, 5911-5929.
5. Satoh, M.; Nishimura, N.; Kanayama, Y.; Naganuma, A.; Suzuki, T.; Tohyama, C. *J. Pharmacol. Exp. Ther.* **1997**, *283*, 1529-1533.
6. Moore, M. J.; Distefano, M. D.; Zydowsky, L. D.; Cummings, R. T.; Walsh, C. T. *Acc. Chem. Res.* **1990**, *23*, 301-308.
7. Dieckmann, G. R.; McRorie, D. K.; Lear, J. D.; Sharp, K. A.; DeGrado, W. F.; Pecoraro, V. L. *J. Mol. Biol.* **1998**, *280*, 897-912.
8. 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. *J. Am. Chem. Soc.* **1997**, *119*, 6195-6196.
9. Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2004**, *43*, 7902-7915.
10. Iranzo, O.; Thulstrup, P. W.; Ryu, S.; Hemmingsen, L.; Pecoraro, V. L. *Chem. Eur. J.* **2007**, *13*, 9178-9190.
11. Matzapetakis, M.; Farrer, B. T.; Weng, T.-C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2002**, *124*, 8042-8054.

12. Matzapetakis, M.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2005**, *127*, 18229-18233.
13. Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Biol. Inorg. Chem.* **2006**, *11*, 876-890.
14. Farrer, B. T.; Harris, N. P.; Balchus, K. E.; Pecoraro, V. L. *Biochemistry* **2001**, *40*, 14696-14705.
15. Ghosh, D.; Lee, K.-H.; Demeler, B.; Pecoraro, V. L. *Biochemistry* **2005**, *44*, 10732-10740.
16. Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2005**, *45*, 9959-9973.
17. Łuczowski, M.; Stachura, M.; Schirf, V.; Demeler, B.; Hemmingsen, L.; Pecoraro, V. L. *submitted*.
18. Dieckmann, G. R. Ph.D. thesis, University of Michigan, Ann Arbor, MI, 1995.
19. Hemmingsen, L.; Narcisz, K.; Danielsen, E. *Chem. Rev.* **2004**, *104*, 4027-4061.
20. Touw, D. S.; Nordman, C. E.; Stuckey, J. E.; Pecoraro, V. L. *Proc. Nat. Acad. Sci.* **2007**, *104*, 11969-11974.
21. Touw, D. S.; Stuckey, J. E.; Pecoraro, V. L. *unpublished results*.
22. Schiering, N.; Kabsch, W.; Moore, M. J.; Distefano, M. D.; Walsh, C. T.; Pai, E. F. *Nature* **1991**, *352*, 168-172.
23. Tröger, W. *Hyp. Interact.* **1999**, *120/121*, 117-128.
24. Wright, J. G.; Tsang, H. T.; Penner-Hahn, J. E.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2434-2435.
25. Helmann, J. D.; Ballard, B. T.; Walsh, C. T. *Science* **1990**, *247*, 946-948.
26. Watton, S. P.; Wright, J. G.; MacDonnell, F. M.; Bryson, J. W.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2824-2826.
27. Utschig, L. M.; Bryson, J. W.; O'Halloran, T. V. *Science* **1995**, *268*, 380-385.
28. George, G. N.; Pickering, I. J.; Prince, R. C.; Zhou, Z. H.; Adams, M. W. W. *J. Biol. Inorg. Chem.* **2000**, *1*, 226-230.
29. Faller, P.; Ctortocka, B.; Tröger, W.; Butz, T.; Vašák, M. *J. Biol. Inorg. Chem.* **2000**, *5*, 393-401.
30. Blake, P. R.; Lee, B.; Summers, M. F. *New J. Chem.* **1994**, *18*, 387-395.
31. Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; O'Halloran, T. V.; Rosenzweig, A. C. *Nat. Struct. Biol.* **2000**, *7*, 766-771.





## Chapter 13

# Understanding Oxotransferase Reactivity in a Model System Using Singular Value Decomposition Analysis

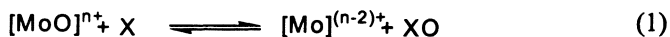
Brian W. Kail<sup>1</sup>, Charles G. Young<sup>2</sup>, Mitchell E. Johnson<sup>1</sup>,  
and Partha Basu<sup>1,\*</sup>

<sup>1</sup>Department of Chemistry and Biochemistry,  
Duquesne University, Pittsburgh, PA 15282

<sup>2</sup>School of Chemistry, University of Melbourne, Victoria 3010, Australia

Singular value decomposition (SVD) analysis on spectrophotometric data obtained from an oxygen atom transfer (OAT) reaction involving a molybdoenzyme model system is reported. Specifically, the rate of solvolysis reaction of a phosphoryl intermediate complex has been compared with independent measurements. The SVD derived reaction rates are consistent with other measurements. This generalized approach is applicable in examining other bioinorganic reactions, and data processing.

For several years, we have been investigating the molecular events of oxygen atom transfer (OAT) reactions involving oxo-molybdenum complexes as models for molybdenum oxotransferases (1-7). The overall reaction is shown in eq 1.



OAT reactions are fundamental reactions in chemistry and biology and constitute a major part of the broader class of atom transfer reactions (8-11). Such

reactions are thought to be involved in the mechanisms of several metalloenzymes such as cytochrome P<sub>450</sub>, sulfite oxidase, nitrate reductase, and dimethyl sulfoxide (DMSO) reductase (12).

A net chemical reaction, such as the OAT reaction, can be broken down into a set of elementary steps in which bonds are broken and made. We have been interested in defining how these individual steps play a role in OAT reactivity and in developing methods to evaluate the individual contributions of these steps. This is different from defining the overall reaction in a single step model where the two reactants pass through a single transition state. Such reactions have been extensively reviewed (13-16). Details of such reactions are garnered through kinetics studies using techniques such as NMR spectroscopy and electronic (UV-visible) spectroscopy. Electronic spectra of reaction mixtures as a function of time often exhibit 'isosbestic' points indicating the number of chemically relevant species. Because of the well-known linear relationship between spectral response (absorption) and concentration, the temporally resolved spectra can be used for quantitation. Herein, we show that, in the absence of other corroborating evidence, this approach may be misinterpreted especially where repeat scan measurements leading to isosbestic points are conducted at a single temperature. More than one chemical transformation operating at the same rate at a particular temperature can also give an isosbestic point; therefore, multiple species can appear as a singly transforming species. It is possible that a mechanistically important, albeit small, component can be missed in such a system. These issues can be addressed using suitable mathematical approaches.

A particularly convenient mathematical approach that has been applied to many biological reactions is singular value decomposition (SVD) (17-23). In biochemical systems, SVD analysis has been successfully applied in studying nanosecond infra-red spectroscopy (24), the photocycles of bacteriorhodopsin (25) and halorhodopsin (26), the reaction of myoglobin with hydrogen peroxide (27, 28), the reaction of cytochrome c oxidase with molecular oxygen (29), and the deconvolution of multiple components in complex NMR spectra (30). The usefulness of the SVD algorithm in providing information about the species and events in a well-defined OAT reaction that models sulfite oxidase reactivity using tertiary phosphine as a substrate is described here.

## SVD Analysis of a Molybdoenzyme Model System

The system we chose to explore is a well characterized inorganic model complex,  $\text{Tp}^{\text{iPr}}\text{Mo}^{\text{VI}}\text{O}_2\text{X}$  (X = SR, OR;  $\text{Tp}^{\text{iPr}}$  = hydrotris(3-isopropylpyrazol-1-

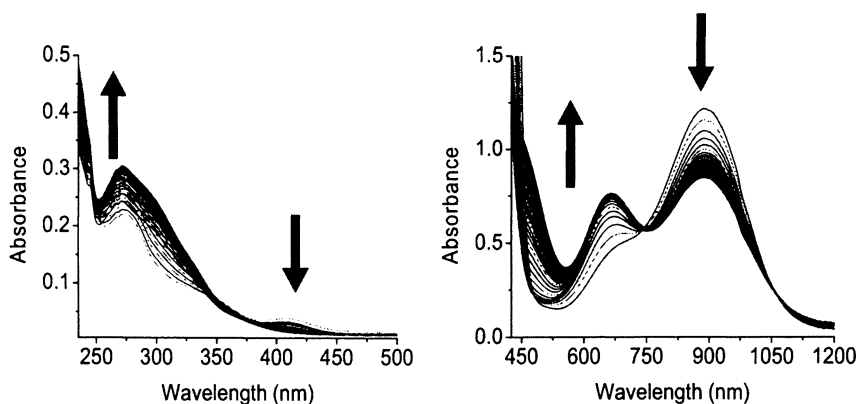
yl)borate). The OAT reactivity of this system has been developed to model sulfite oxidase, a prototype of oxotransferase where sulfite is transformed to sulfate at the dioxo-Mo(VI) center. Phosphines are commonly used as model substrates because of their versatility, which can aid structure-function correlation. We have demonstrated that the OAT reaction from  $\text{Tp}^{\text{iPr}}\text{MoO}_2\text{X}$  to tertiary phosphines (*e.g.*,  $\text{PEt}_3$ ) proceeds through at least two steps. In the first step, a phosphoryl intermediate complex,  $\text{Tp}^{\text{iPr}}\text{Mo}^{\text{IV}}\text{O}(\text{OPh})(\text{OPEt}_3)$ , **1**, is formed with concomitant two-electron reduction of the metal center, and in the second step the phosphine oxide is displaced by a solvent molecule. The solvolysis of **1** is the subject of this report. We have investigated this reaction in acetonitrile, using electronic and NMR spectroscopy, and applied SVD methodology to the electronic spectral data. Compound **1** was synthesized as previously reported (6), and the kinetic parameters from traditional single wavelength measurements for the solvolysis reaction have been reported (7).

## SVD Methodology

Time-dependent spectra of the solvolysis of **1** were collected in two regions: 1200-500 nm, denoted the low energy region, and 600-230 nm, denoted the high-energy region. For a typical low energy experiment, 3 - 5 mg ( $4.8 - 8.0 \times 10^{-6}$  mol) of **1** was dissolved in 750  $\mu\text{L}$  - 1 mL of dry-degassed solvent at 11°C. The final concentration ranged from 4.8 - 10.6 mM in four separate experiments with a typical dead time of ~ 1.5 min. Data were recorded with a resolution of 1-2 nm/data point and at a rate of 2.5 - 5 min per scan, over a 200 - 500 min reaction time. Experiments in the higher energy region were conducted similarly, with a final concentration range of 200 - 375  $\mu\text{M}$ . These data (collected in ASCII format) were arranged in a matrix format with variables such as wavelength, time interval and amplitude, and subjected SVD analysis using a suitable program.

The appropriate SVD-derived spectral and temporal eigenvectors were selected and the temporal vectors were modeled. Ideally, the temporal vectors are the kinetic traces of individual components, each one being associated with a spectrum of a pure component (*i.e.*, the spectral vector). Once the temporal vectors had been modeled, the pure component spectra were reconstructed as a function of the pre-exponential multiplier obtained from the analysis, SVD determined spectral eigenvectors, and the corresponding eigenvalues. After the spectra of the component species were determined, the extinction profile was calculated and used along with the calculated decay times to construct a linear combination of the pure component species' contributions to the observed

spectrum. This linear combination was then used to provide a simulation and a global fit to the observed, time-dependent absorption spectra (Figure 1). The rate of the reaction was also evaluated using  $^{31}\text{P}$  NMR spectroscopy (7).



*Figure 1. Time-dependent optical spectra of 1 in acetonitrile taken at 11 °C. High energy region (left); low energy region (right). Each spectrum was scanned for 2.5 - 5 min, and the spectral data were collected for 200 - 500 min of reaction time. The arrows indicate the direction of change.*

### Mathematical Description of Singular Value Decomposition

The details of the SVD methodology can be found elsewhere (31, 32). The observed absorption can be defined as a linear function of the absorption of the individual components (eq 2). Here  $A_{\text{obs}}$  represents the observed absorption at a

$$A_{\text{obs}} = x_1 * c_1 + x_2 * c_2 + x_3 * c_3 + \dots x_n * c_n \quad (2)$$

particular wavelength,  $c_1 \dots c_n$  represent the concentration of components, and  $x_1 \dots x_n$  represent the molar absorptivity of the corresponding components times the path length (1 cm). The spectral data matrix,  $A_{\text{obs}} = [A]$ , can be represented as a combination of two orthogonal eigenvectors termed the left and right singular vectors (U and V, respectively) such that eq 3 is maintained. Here U is an  $m \times n$  matrix, S is an  $n \times n$  diagonal matrix, and  $V^T$  is an  $n \times n$  matrix.

$$A = U * S * V^T \quad (3)$$

The eigenvalues obtained from the SVD analysis of [A] (singular values of A) indicate the relative spectral importance of its associated columns of U and V. The left singular vectors (columns of U) represent the basis spectral vectors associated with the matrix [A] such that when multiplied by the corresponding singular value, a mathematical representation of the pure component species is attained. The corresponding right singular vectors (columns of V) are representative of the change in the spectral amplitude of a given basis spectrum. Thus, when the product  $U_i * S_i$  is considered with respect to its column of  $V_i$ , the shape of the spectrum of the pure species can be predicted. In addition, the relative contribution of that spectrum ( $S_i$ ) to the overall spectrum and the extent of change of the spectrum ( $V_i$ ) of a particular basis species can also be calculated. Because the matrix [A] can be divided into the three separate matrices, as defined in eq 3, one can reconstruct the data matrix as in eq 4 such that  $|A - A_i|^2$  is minimized (eq 5) (33) (where  $m$  is a measure of wavelengths and

$$A_i = U_i * S_i * V_i^T \quad (4)$$

$$|A - A_i|^2 = m * n * \sigma^2 < \sum_{i=r}^n (S_{i,j})^2 \quad (5)$$

$n$  is the temporal portion of the data). The magnitude of  $\sigma^2$  is the difference between [A] and  $[A_i]$ . It provides a measure of how much of the original data contains useful information and permits the vectors predominantly representing noise to be discarded. Thus, SVD generates a 'new' data set with reduced noise but which contains all of the important information present in the original matrix. In addition, this new data matrix is compact in size which helps in data analysis. Autocorrelations of U and V (eq 6, where  $u_{ij}$  represents cross correlation

$$C_{U_j} = \sum_{i=1}^{m-1} u_{ij} * u_{i+1,j} \quad (6)$$

between the column elements of U) are often used to understand the significance of the results (*i.e.*, the extent of contribution of a species to the overall spectrum). In addition, numerical differences (eq 5) have been used as screening methods, as well as the magnitude of the singular values corresponding to columns of U and V. Shrager and Hendler (33) have suggested that accepted values of  $C_{U_j}$  should be higher than +0.5 and have upper and lower limits of 1 and -1, respectively. Also, any numerical distance (eq 5) greater than the sum of the singular values squared should be discarded. Henry and Hofrichter (20) have suggested a cross-correlation method that may be applied to extract minor

components (primarily noise), such that the autocorrelations in eq 6 are maximized (eq 7). When the autocorrelation order is not maintained, the

$$C_{U_i,k} = \sum_{i=1}^{m-1} u_{ij} * u_{i+1,k} \quad (7)$$

independent rotation of U and V is performed. In this case, matrix  $[X_{ij}]$  with eigenvectors comprised of the vectors of a rotation matrix  $[R_{ij}]$  is created. This  $[R_{ij}]$  matrix, when multiplied by the original source vector, produces new vectors  $[U_i^R]$  from  $U_i$  or  $[V_i^R]$  from  $V_i$  and filters the noise to the remaining vectors  $[U_{i+1}^R]$  and  $[V_{i+1}^R]$  as defined in eqs 8 and 9. In these new vectors signal is optimized with respect to noise. The rotation procedure defined in eq 8 results in a loss of orthogonality in the columns of V and U that can be remedied through the implementation of eq 9, followed by a second processing by SVD.

$$U_{kl} * R_{kl} = U_{kl}^R \quad (8)$$

$$A_{ij} = (U_{ij} * S_{ij} * U_k^R * S_k * V_{ij}^T * V_k^{RT}) \quad (9)$$

Once the columns of U, S and V have been calculated, the columns of V are simulated for the fewest numbers of kinetic processes (eq 10). Here,  $F_{ij}$  represents the change in amplitude of the  $i^{\text{th}}$  basis spectrum accompanying the  $j^{\text{th}}$  relaxation, for which maximum amplitude is observed in  $V_i$ . In this case, the data

$$V_{i(t)} = \sum_{j=1}^k F_{ij} * e^{-\lambda_j t} \quad (10)$$

set can further be reduced to a set of (e) exponential decays, (f) values of  $F_{ij}$  and (b) basis spectra, as provided in eq 10. A reconstructed, noise-free spectrum is calculated by eq 11, where  $U_i(\lambda)$  represents the basis spectra of a component and  $S_i$  represents its singular value.

$$D_k(\lambda) = \sum_{j=k} \sum_{i=1} U_i(\lambda) * S_i * F_{ij} \quad (11)$$

## Results From the SVD Analysis of Spectral Data

The spectral eigenvectors (U) (eq 3) provide a mathematical solution to the extent of each species' contribution to the observed spectra. These eigenvectors provide details of the absorption spectrum that each species contributes to the

overall spectrum at any given time. The first five vectors of  $U$ , obtained from the high energy region, and the autocorrelation values for the first five  $U$  and  $V$  vectors are tabulated in Table I.

**Table I. Representative SVD Results From High Energy Experiments<sup>a</sup>**

<i>Species</i>	<i>Autocorrelation</i> $U^b$	<i>Autocorrelation</i> $V^b$	<i>S value</i>	<i>%</i> <i>Contribution</i>
1	0.998	0.982	4.01	94.0
2	0.995	0.852	0.241	5.7
3	0.936	0.801	0.021	0.3
4	0.756	0.158	0.008	$2.0 \times 10^{-2}$
5	0.222	-0.148	0.005	$1.2 \times 10^{-3}$

<sup>a</sup>600-235 nm; 375 × 60 data matrix.

<sup>b</sup>Autocorrelation values depend upon the resolution in the data set. For example, converting this 375 × 60 matrix into a 375 × 30 matrix shows a significant drop in the magnitude of the autocorrelation values in spectral ( $U_1$ ,  $U_2$  and  $U_3$ ), as well as the temporal ( $V_1$ ,  $V_2$ , and  $V_3$ ) eigenvectors for the respective component.

When the spectral eigenvectors are plotted (Figure 2), vector A exhibits both positive and negative relative absorbance. The negative value (~ 300 nm) suggests that this species is likely the cause of a decrease in absorbance. This change can be independently observed in a single wavelength measurement, provided that its overall contribution is significant. In contrast, eigenvector B exhibits positive intensity. The positive intensity suggests that this species can contribute to an overall increase in absorption from ~ 300 - 600 nm, provided that it also has a measurable contribution to the total spectral matrix over time. Eigenvector C also exhibits positive intensity and behaves similarly to eigenvector B. All other eigenvectors exhibit very noisy spectral features.

Four of the first five spectral eigenvectors have auto-correlation values that are above 0.5 and were therefore considered. The significance of a given species to the overall observed spectrum was established by the eigenvalue ( $S$ ) that corresponds to the  $U$  and  $V$  vectors. Figure 3 shows the exponential decay in  $S$  values with increasing eigenvector number, as well as a corresponding decay in the autocorrelation values of the spectral and temporal eigenvectors. This rapid decay shows that at some value  $N$ , both the  $U$  and  $V$  eigenvectors will contribute a nearly constant amount, suggesting that these vectors represent spectral noise. In this case, when the  $S$  eigenvalues are considered, the first species contributes to ~ 94% of the total spectral variation, while the second and third contribute 5.7% and 0.3%, respectively (Table I). By the tenth set of eigenvectors, the contribution is  $9.8 \times 10^{-10}$  % and the shape of  $U$  and  $V$  (Figure 4) appear as noise. Because vectors for species 5 through 10 account for ~ 0.02% of the

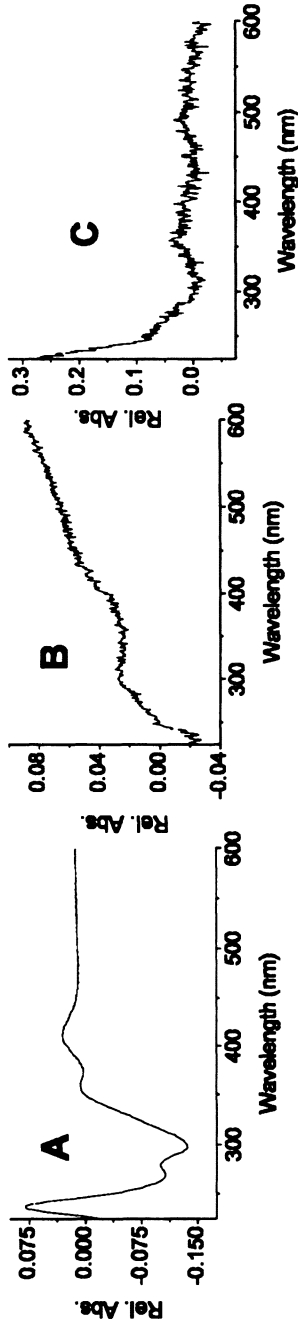


Figure 2. The first three spectral eigenvectors (denoted as A, B and C) for the solvolysis of I (from the high energy region data). These eigenvectors form the basis of the observed data [A].



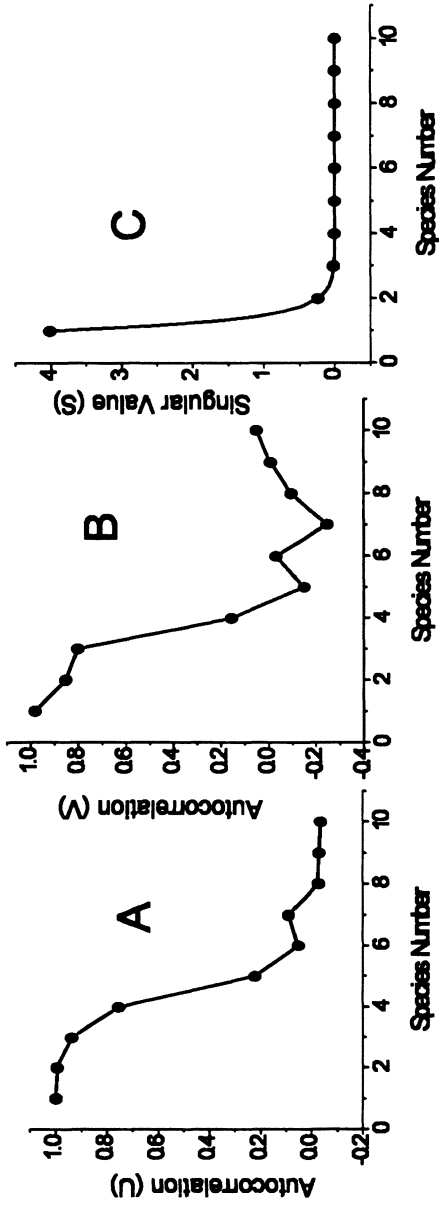


Figure 3. Autocorrelation plots for A) the first ten spectral eigenvectors; B) the first ten temporal eigenvectors; and C) singular value plot, in the high energy region of the spectra. Note that the first three eigenvectors make reasonable contribution.

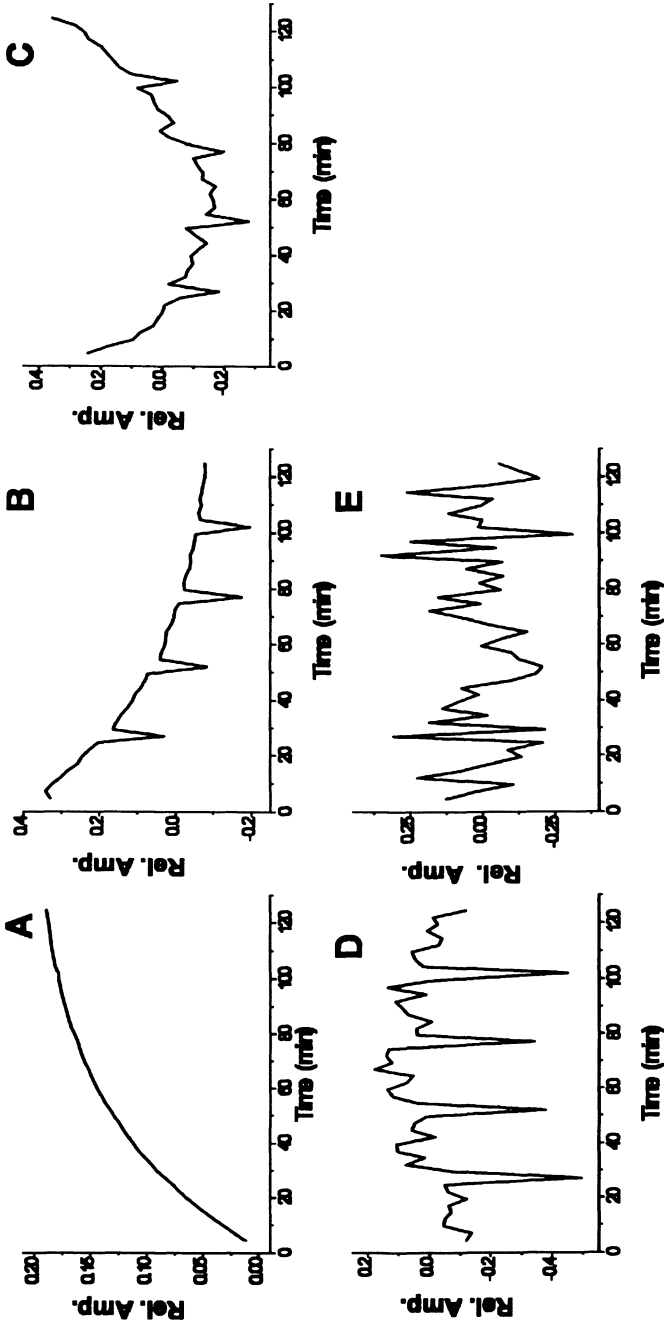


Figure 4. The first five temporal eigenvectors for the decay of I in the high energy region. The repeated spike in the spectra is considered noise from spectrophotometer. After the first three, a positive contribution from temporal eigenvectors could not be observed, thus they are considered noise.

spectral variation, they do not represent a significant contribution to [A], and are not considered further. These two criteria limited the plausible vectors to A, B, C, and possibly D.

Next, the results from the temporal eigenvectors (V), which detail how the  $i^{\text{th}}$  U vector changes with time, were evaluated. Figure 4 shows the first five V vectors, where the first four vectors (A, B, C, D) were considered to have proper shape. From these, the fourth V vector was discarded because of its low autocorrelation (0.158). Here panel A (Figure 4) shows a clean single exponential increase in amplitude while panel B reflects single exponential decay. However, panel C shows a biphasic behavior consistent with the presence of at least two detectable species. Thus, the species in panel A increases relative to the starting complex, while the second species, panel B, decreases. The third species, in panel C, appears to follow a more complex temporal pattern, which may be an artifact of its low contribution to the observed spectra ( $S_3 = 0.02$  0.34%) as well as its considerably lower autocorrelation values in its spectral eigenvector (0.936) and temporal eigenvector (0.801). Both the low spectral contribution and poor autocorrelations allow for this third species to be strongly influenced by the preceding, more significant, species. Thus, in the process, three eigenvectors (A, B, and C) are found to contribute to the overall spectral features. From NMR spectra, these three eigenvectors represent  $\text{Tp}^{\text{iPr}}\text{MoO}(\text{OPh})(\text{OPeEt}_3)$  (eigenvector A, species 1),  $\text{Tp}^{\text{iPr}}\text{MoO}(\text{OPh})(\text{MeCN})$  (eigenvector B, species 2), and a dinuclear species (eigenvector C, species 3) (7). The predicted spectra of these components are shown in Figure 5.

In order to obtain estimates of the rates of transformation of the components, the columns of the temporal eigenvectors were fit with a sum of exponentials equal to the number of components found in the SVD data reduction scheme. Thus, in all SVD based analyses, a fit based on a sum of three exponents was used (eq 12). The time constants for the three processes are:

$$y = y_0 + \left[ f_1 * \exp\left(-\tau_1/t\right) + f_2 * \exp\left(-\tau_2/t\right) + f_3 * \exp\left(-\tau_3/t\right) \right] \quad (12)$$

$\tau_1 \sim 39.9$  min,  $\tau_2 \sim 44.1$  min, and  $\tau_3 \sim 210.9$  min. The data indicate that  $\tau_1 \sim \tau_2 < \tau_3$ , suggesting formation of species 2 may originate from the decay of species 1. In addition, the rate constants obtained from the SVD analysis were employed to establish a concentration profile for the reaction using four possible mechanisms: 1)  $A \rightarrow B \rightarrow C$ ; 2)  $A \rightleftharpoons B \rightarrow C$ ; 3)  $A \rightarrow B \rightleftharpoons C$ ; and 4)  $A \rightleftharpoons B \rightleftharpoons C$ . For  $A \rightarrow B$  conversion, (*i.e.*, species 1 to species 2 conversion),  $\tau_{\text{avg}} = 42$  min leads to  $k_f = 0.023 \text{ min}^{-1}$  as the rate constant. For  $A \rightleftharpoons B$  conversions, with  $\tau_1 = 40.0$  min, leads to  $k_f = 0.025 \text{ min}^{-1}$  and  $\tau_2$  was taken to be  $k_r = 0.022 \text{ min}^{-1}$ , such that  $K_{\text{eq}} = (k_f/k_r) = 1.10$ . Lastly, for  $B \rightarrow C$  conversion,  $k_f$  was taken to be  $0.0048 \text{ min}^{-1}$ , while for the  $B \rightleftharpoons C$  conversions  $K_{\text{eq}}$  was taken to be 0.0048.

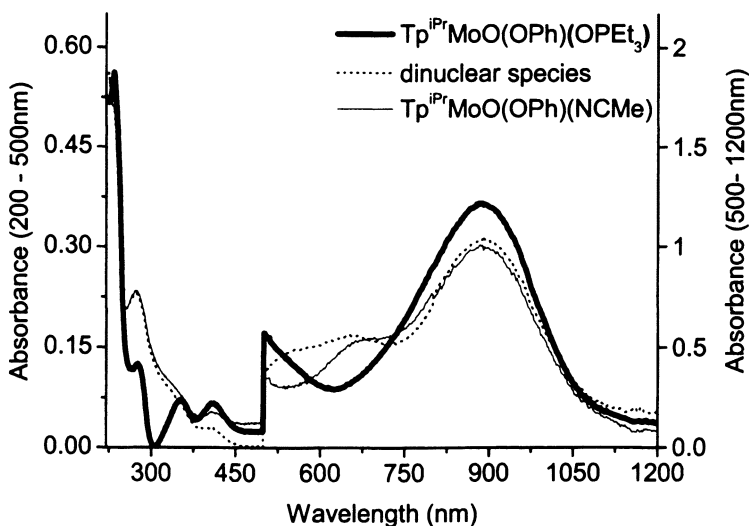
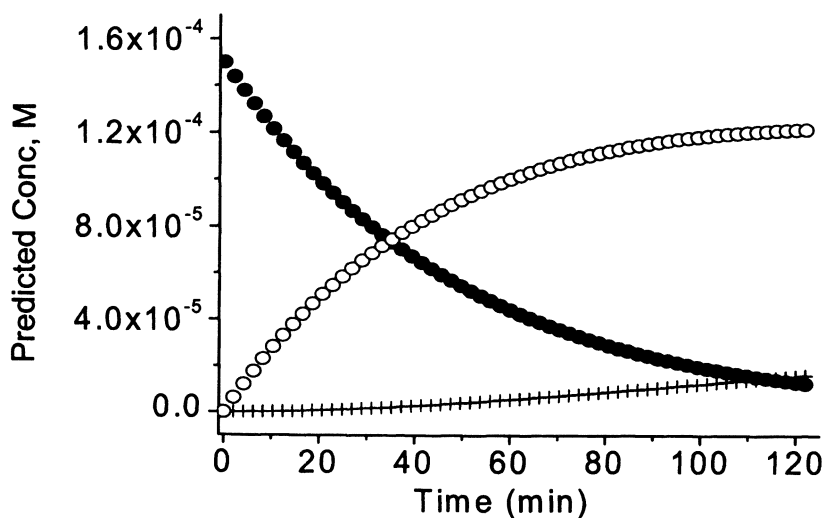


Figure 5. Calculated spectra of the pure components, as determined from SVD analysis. The low-energy spectra (500-1200 nm) were recorded at ten times the concentration of 1, relative to the concentration for the high-energy spectra accounting for the discontinuity at 500 nm. (Reproduced with permission from PhD dissertation, Duquesne University.)

After creating the concentration profiles (Figure 6), an absorption profile was generated from the pure component spectra using eq 9, and the spectral data matrix was reconstructed as a linear combination of the component spectra. The best fit was obtained using the simplest mechanism,  $A \rightarrow B \rightarrow C$ , and was taken as the plausible mechanism. Using this mechanism and the extinction profiles, absorption spectra at different time points were calculated (Figure 7). The SVD derived kinetic parameters were compared with those obtained (7) from the single wavelength measurement optical measurements and  $^{31}\text{P}$  NMR spectroscopy (Table II).

### Evaluation of SVD Derived Results

SVD analysis of the difference spectra (*i.e.*, absorption at time  $t$  minus the absorption at time zero,  $A_t - A_0$ ) was found to produce a better quality deconvolution than non-difference spectral data. This was especially true in regions of the spectrum where changes in low intensity transitions (*e.g.*, d-d transitions) do not lead to appreciable changes in the visible spectrum. However,



*Figure 6. Calculated concentration profile for the decay using the  $A \rightarrow B \rightarrow C$  mechanism. [A] is shown in solid circles, [B] is shown in open circles and [C] is shown in plus sign. Note that the decrease in concentration of species A, while concentrations of species B and C increase, albeit differently. In this experiment the reaction was followed for 120 min.*

**Table II. Comparison of  $\tau_{obs}$  Data by Complementary Methods**

<i>Method</i>	$\tau_{obs}$ (min) for the decay of species <i>1</i>	$\tau_{obs}$ (min) for the formation of species <i>2</i>	$\tau_{obs}$ (min) for the decay of species <i>2</i>
UV-Vis 301 nm	47.5 ± 0.2		
UV-Vis 910 nm	50.3 ± 0.3		
UV-Vis 680 nm		55.6 ± 2.3	431 ± 6.4
<sup>31</sup> P NMR [Mo- OPR <sub>3</sub> ]	50.0 ± 2.9 <sup>a</sup>	47.2 ± 4.1 <sup>b</sup>	
Average SVD V <sub>1</sub>	44.0 ± 4.0 <sup>c</sup>	55.0 ± 8.8 <sup>d</sup>	227.5 ± 20 <sup>e</sup>
Average UV-Vis		51.1 ± 4.1	
Average <sup>31</sup> P		49.5 ± 2.0	
Average SVD		49.8 ± 8.2	

<sup>a</sup>decay of Mo-bound phosphine oxide.

<sup>b</sup>Growth of free phosphine oxide, which is assumed to the formation of species **2**.

<sup>c</sup>Eigenvector, V1. <sup>d</sup>Eigenvector, V2. <sup>e</sup>Eigenvector, V3.

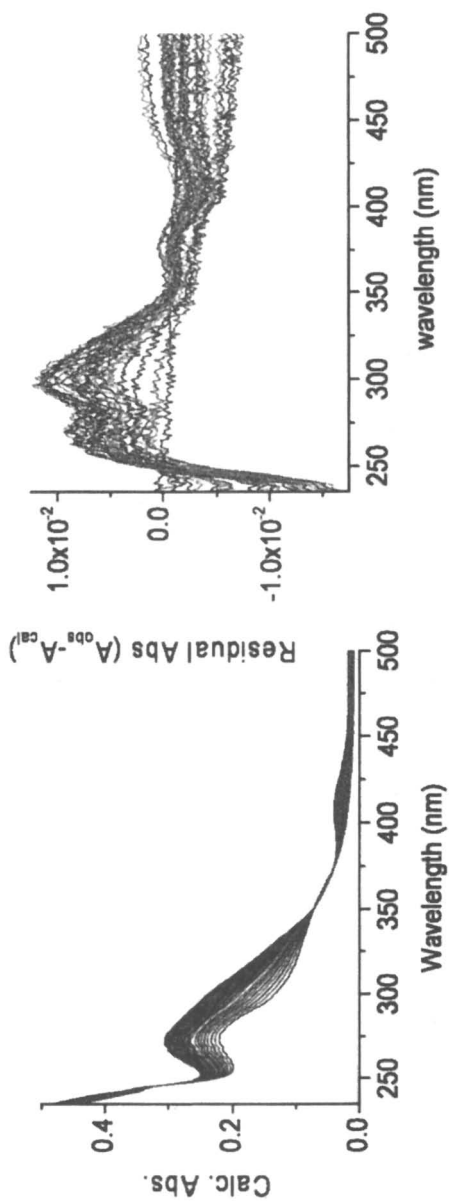


Figure 7. Calculated absorption spectra in the high energy region with mechanism 1 ( $A \rightarrow B \rightarrow C$ ).  
Right, the residual of calculated and the observed data difference.

the form of the data set is significant when higher intensity transitions are involved. It is important to take all data with systematic intervals such that the time domain may be assumed as a constant independent variable with only absorption changes at a given wavelength as the dependent variable. A data set that does not have this information is unlikely to produce a reliable data reduction in the temporal eigenvectors and may therefore lose any relevance to kinetic applications. Care should be taken in deriving the kinetic parameters using SVD, particularly for those species that are very minor contributors to the overall spectra. While the associated errors may be different, the individual rate constants agree well with the data obtained from other methods, for the first two species in our model system (Table II). Thus, the four trial average SVD derived rate constants provide a reasonable depiction of the true rate of the process, while a given single SVD measurement provides a less ideal representation.

The SVD derived data for the second process, *i.e.*, the conversion of species 2 to species 3, could only be compared with electronic spectral measurements that indicate that the SVD overestimated the rate by a factor of  $\sim 2$ . This process also has the largest associated error. The first process ( $A \rightarrow B$ ) is approximately ten times faster than the second process ( $B \rightarrow C$ ), as determined by single wavelength studies, and it is not optimal to collect data for a period long enough to observed significant production of species 3, while maintaining sufficient observation of the conversion of species 1 into species 2. A reasonable period of observation for the second process would require an approximate 33 h time frame and the size of the data set required to simultaneously account for the more rapid conversion of species 1 to species 2 and the species 2 to species 3 conversions would approach 1,000 spectra.

The SVD analysis of this model reaction revealed several important aspects. First, SVD analysis clearly revealed the presence of a second reaction (*i.e.*,  $B \rightarrow C$ ) which was not obvious from the repeat scan analysis where reasonable isosbestic points (Figure 1) were observed. Thus, kinetic data collected at wavelengths where the spectra of the three species are not distinctly different (*e.g.*, 300 nm, two of three species exhibit similar spectral features) can be taken as only for a pure reaction (*i.e.*,  $A \rightarrow B$ ). Second, SVD analysis revealed the spectral features of the pure component species (Figure 5) which permitted selection of proper wavelengths for single wavelength kinetic investigation (7). Third, SVD also revealed the rate of change in concentration of individual species that agree well with those obtained from other independent measurements (Table II). It is well-known that the formation of  $\mu$ -oxo-dinuclear compounds is pervasive among OAT model compounds. The SVD analysis reveals that under the experimental conditions used here such a reaction proceeds at a slower time scale. Such a dinucleation reaction, however, is biologically unimportant.



## Conclusions

Here we described the application of SVD toward understanding the OAT reactivity of a well-defined molybdoenzyme model system. The SVD algorithm produces rate constants that are comparable with the single wavelength measurement or NMR methods, provided that each step in the reaction occurs on a similar time scale. The algorithm does exceptionally well at the prediction of the pure component spectra and detection of key absorption bands.

In this chapter we have outlined the procedure of SVD analysis as it applies to a model oxotransferase. The implementation of SVD in a kinetic study is useful in predicting the spectra of unknown components, which can lead to greater mechanistic insights into bioinorganic systems, *e.g.*, identification of protonation state, reactive intermediates, and structural flexibility. Such approaches have been applied to investigating the reaction of myoglobin with hydrogen peroxide (27) and pH dependence of the reactivity of cytochrome c oxidase (22) or its binding to carbon monoxide (34). This approach has yet to be applied in understanding the reactivity of any oxotransferase, where the exact nature of the molecular events may vary, but the general principle of data analysis should still be applicable. Another potential area where bioinorganic chemistry can benefit from using SVD approach is the analysis of data containing regulation of gene or protein. With the advent of microarray and proteomic techniques, a large amount of data can be generated in a short period of time. Analysis of such data could be a daunting task and SVD or principal component analysis can provide a wealth of information.

## Acknowledgements

We thank Profs. Brett Larget and Andrew Pacheco for stimulating discussions and the National Institutes of Health (GM 651555) and Australian Research Council for support of this research.

## References

1. Nemykin, V. N.; Basu, P. *Inorg. Chem.* **2005**, *44*, 7494-7502.
2. Smith, P. D.; Millar, A. J.; Young, C. G.; Ghosh, A.; Basu, P. *J. Am. Chem. Soc.* **2000**, *122*, 9298-9299.
3. Nemykin, V. N.; Davie, S. R.; Mondal, S.; Rubie, N.; Kirk, M. L.; Somogyi, A.; Basu, P. *J. Am. Chem. Soc.* **2002**, *124*, 756-757.
4. Nemykin, V. N.; Laskin, J.; Basu, P. *J. Am. Chem. Soc.* **2004**, *126*, 8604-8605.

5. Nemykin, V. N.; Basu, P. *Dalton Trans.* **2004**, 1928-1933.
6. Millar, A. J.; Doonan, C. J.; Smith, P. D.; Nemykin, V. N.; Basu, P.; Young, C. G. *Chem. Eur. J.* **2005**, *11*, 3255-3267.
7. Kail, B. W.; Pérez, L. M.; Zarić, S. D.; Millar, A.; Young, C. G.; Hall, M. B.; Basu, P. *Chem. Eur. J.* **2006**, *12*, 7501-7509.
8. Taube, H. In *Mechanistic Aspects of Inorganic Reactions*; Rorabacher, D. B., Endicott, J. F., Eds.; ACS Symposium Series No. 198; ACS: Washington, D.C., 1982; pp. 151-179.
9. Sharpless, K. B. *Tetrahedron* **1994**, *50*, 4235-4258.
10. Holm, R. H. *Chem. Rev.* **1987**, *87*, 1401-1449.
11. McMaster, J.; Tunney, J. M.; Garner, C. D. *Prog. Inorg. Chem.* **2003**, *52*, 539-583.
12. Enemark, J. H.; Cooney, J. J. A.; Wang, J.-J.; Holm, R. H. *Chem. Rev.* **2004**, *104*, 1175-1200.
13. Woo, L. K. *Chem. Rev.* **1993**, *93*, 1125-1136.
14. Espenson, J. H. *Coord. Chem. Rev.* **2005**, *249*, 329-341.
15. Bakac, A. *Coord. Chem. Rev.* **2006**, *250*, 2046-2058.
16. Abu-Omar, M. M. *Chem. Commun.* **2003**, 2102-2111.
17. Hendler, R. W.; Bose, S. K.; Shrager, R. I. *Biophys. J.* **1993**, *65*, 1307-1317.
18. Bracher, A.; Schramek, S.; Bacher, A. *Biochemistry* **2001**, *40*, 7896-7902.
19. Vandegriff, K. D.; Shrager, R. I. *Methods Enzymol.* **1994**, *232*, 460-485.
20. Henry E. R.; Hofrichter, J. *Methods Enzymol.* **1992**, *210*, 129-192.
21. Szundi, I.; Cappuccio, J. A.; Borovok, N.; Kotlyar, A. B.; Einarsdóttir, O. *Biochemistry* **2001**, *40*, 2186-2193.
22. Petraccone, L.; Pagano, B.; Esposito, V.; Randazzo, A.; Piccialli, G.; Barone, G.; Mattia, C. A.; Giancola, C. *J. Am. Chem. Soc.* **2005**, *127*, 16215-16223.
23. Crampin, E. J.; Schnell, S.; McSharry, P. E. *Prog. Biophys. Mol. Biol.* **2004**, *86*, 77.
24. Sasaki, J.; Yuzawa, T.; Kandori, H.; Maeda, A.; Hamaguchi, H.-O. *Biophys. J.* **1995**, *68*, 2073-2080.
25. Kulcsár, A.; Saltiel, J.; Zimányi, L. *J. Am. Chem. Soc.* **2001**, *123*, 3332-3340.
26. Dioumaev, A. K. *Biophys. Chem.* **1997**, *67*, 1-26.
27. Egawa, T.; Yoshioka, S.; Takahashi, S.; Hori, H.; Nagano, S.; Shimada, H.; Ishimori, K.; Morishima, I.; Suematsu, M.; Ishimura, Y. *J. Biol. Chem.* **2003**, *278*, 41597-41606.
28. Cooper, C. E.; Jurd, M.; Nicholls, P.; Wankasi, M. M.; Svistunenko, D. A.; Reeder, B. J.; Wilson, M. T. *Dalton Trans.* **2005**, 3483-3488.
29. Szundi, I.; Liao, G.-L.; Einarsdóttir, O. *Biochemistry* **2001**, *40*, 2332-2339.
30. Xu, Q.; Sachs, J. R.; Wang, T.-C.; Schaefer, W. H. *Anal. Chem.* **2006**, *78*, 7175-7185.

31. Kail, B. W. Ph.D. Dissertation, Duquesne University, Pittsburgh, PA 2006.
32. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes: The Art of Scientific Computing*, 3<sup>rd</sup> Ed.; Cambridge University Press: Cambridge, U.K., 2007.
33. Shrager, R. I.; Hendler, R. W. *Anal. Chem.* **1982**, *54*, 1147-1152.
34. Szundi, I.; Ray, J.; Pawate, A.; Gennis, R. B.; Einarsdóttir, O. *Biochemistry* **2007**, *46*, 12568-12578.



## Chapter 14

# DNA Minor Groove Recognition by Ni(II)• and Cu(II)•Gly-Gly-His Derived Metallopeptides

## Models of Protein and Natural Product DNA Recognition

Eric C. Long<sup>1,\*</sup>, Ya-Yin Fang<sup>2</sup>, and Mark A. Lewis<sup>1</sup>

<sup>1</sup>Department of Chemistry and Chemical Biology,  
Purdue School of Science,  
Indiana University-Purdue University Indianapolis (IUPUI),  
Indianapolis, IN 46202

<sup>2</sup>Department of Biochemistry and Molecular Biology,  
Howard University College of Medicine, Washington, D.C. 20059

Metallopeptides derived from either Ni(II)• or Cu(II)•Gly-Gly-His serve as models to increase our understanding of protein- and natural product-DNA minor groove recognition events. Reviewed herein are studies aimed at understanding detailed aspects of M(II)•Gly-Gly-His derived metallopeptide-DNA minor groove recognition through non-DNA cleavage based techniques. These studies include the results of recent NMR, DNA fiber EPR, and molecular dynamics simulations.

Metallopeptides derived from M(II)•Gly-Gly-His model the amino-terminal Cu<sup>2+</sup> or Ni<sup>2+</sup> bound Xaa-Xaa-His tripeptide motifs found in the serum albumins (1, 2), histatins (3), neuromedins (1), and the human DNA binding and condensing protein protamine P2 (4). In addition, M(II)•Gly-Gly-His derived metallotripeptides, either alone or upon conjugation to other entities, have found diverse applications in the field of nucleic acid recognition and targeted cleavage agent design due to their unique structures and metal-supported reactivities (5). Within proteins or tripeptides, Xaa-Xaa-His motifs bind Cu<sup>2+</sup> or Ni<sup>2+</sup> with high affinities at physiological pH through coordination of the peptide terminal

amine, two intervening deprotonated peptide amide bonds, and the His imidazole side chain to yield approximately square-planar structures (Figure 1) (1, 6). The metal binding and resulting redox properties of these motifs have prompted their use in studies of both macromolecule- and low molecular weight agent-nucleic acid recognition and reactivity events (5). Indeed, early-on  $\text{Cu}^{2+}$ -bound Gly-Gly-His,  $\text{Cu(II)}\cdot\text{Gly-Gly-His}$ , was shown to induce DNA strand scission in the presence of ascorbate (7). This metal-based oxidative reactivity led to the synthetic (8-10) and biosynthetic (11, 12) incorporation of this tripeptide motif at the amino-termini of DNA binding proteins as an affinity cleavage appendage; further studies with these same systems revealed that the oxidative species generated in the presence of  $\text{Cu}^{2+}$  is non-diffusible and that analogous  $\text{Ni}^{2+}$  complexes, when appropriately activated, can also lead to a non-diffusible oxidant capable of deoxyribose-based DNA strand scission (8). Along with the modification of DNA binding proteins,  $\text{Cu(II)}\cdot$  and  $\text{Ni(II)}\cdot\text{Gly-Gly-His}$  derived metallotripeptides have been conjugated to a variety of other nucleic acid-targeted systems including peptides (13), drugs (14-16), oligonucleotides (17, 18), and peptide nucleic acids (PNAs) (19). These reactive motifs have also been exploited in the development of protein scission and cross-linking agents (20), and have been used as models of Ni-induced DNA damage and toxicity (4, 21-23).

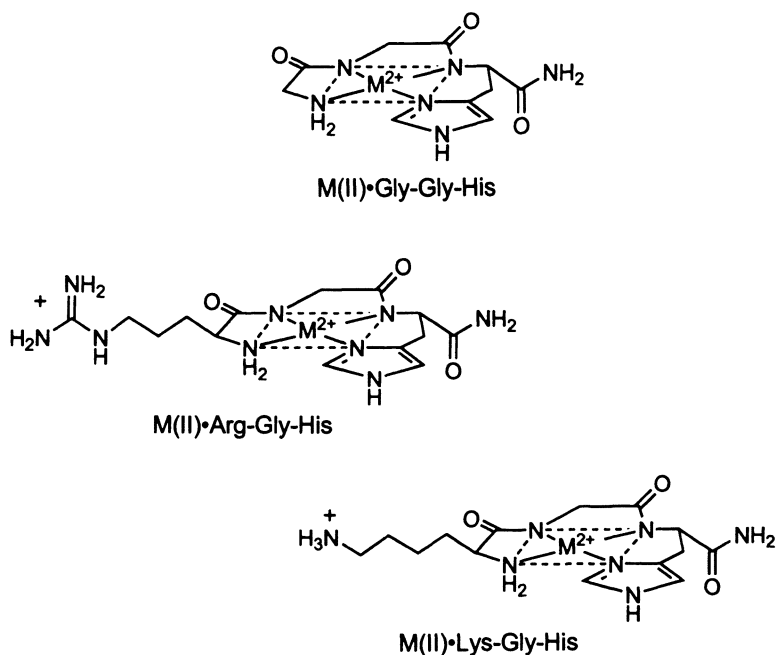


Figure 1. Metal-ligand coordination in  $\text{M(II)}\cdot\text{Gly-Gly-His}$ -derived metallopeptides.

Along with their activity in conjugates, as stand-alone metallotripeptides M(II)•Xaa-Xaa-His systems have afforded an opportunity to examine fundamental peptide and amino acid molecular recognition events with DNA and RNA (5, 24). M(II)•Xaa-Xaa-His metallopeptides are “drug-like” in their overall size, offer the advantage of being stable, structurally well-defined metal-ligand complexes at physiological pH, and allow the first two amino acids to be substituted for any  $\alpha$ -amino acid; in some instances, metal-ligand affinities can be further enhanced by particular substitutions (6). Importantly, this final attribute permits various amino acids to be incorporated into the tripeptide allowing their accompanying side chains to project from the periphery of the metal-peptide framework to potentially influence the molecular recognition of macromolecules such as DNA and RNA. Furthermore, as synthetic systems, the *orientation* of these same side chain functional groups relative to the main equatorial plane of the metallopeptide can be controlled through the inclusion of L- vs. D-amino acid residues during their solid-phase synthesis. Indeed, our laboratory has studied metallopeptides derived from either individually- or combinatorially-synthesized (25) peptides containing chemical functionalities and amino acids commonly employed by proteins and peptide-based natural product antitumor agents for DNA and RNA molecular recognition (*e.g.*, amides, amines, and guanidinium functional groups through Asn, Lys, and Arg residues, respectively).

Our previous work with M(II)•Xaa-Xaa-His systems has demonstrated that Ni(II)•Gly-Gly-His derived metallopeptides containing carboxy-terminal amides recognize and induce site-selective DNA (26) and RNA (27) cleavage as a function of their amino acid compositions, stereochemistries, and overall shape when activated with peroxide-based agents such as H<sub>2</sub>O<sub>2</sub>, organic peracids, or KHSO<sub>5</sub> (28). We determined that the inclusion of positively-charged amino acids such as Arg and Lys focus the cleavage activity of these metallopeptides to A/T-rich DNA regions and that the inclusion of D-amino acids can lead to alternative cleavage site-selectivities and efficiencies (26). These investigations indicated that metallotripeptides derived from M(II)•Gly-Gly-His interact with DNA via the minor groove; evidence in support of minor groove recognition, as discussed in detail in previous reviews (5, 24) includes: a 3'-asymmetric DNA cleavage pattern when DNA restriction fragments are employed as substrates; avoidance of homopolymeric A<sub>n</sub>•T<sub>n</sub> sites containing narrowed minor grooves; cleavage inhibition in the presence of the minor grooved targeted agent distamycin; a 40° orientation of the mean equatorial plane of Cu(II)•Xaa-Gly-His relative to a substrate DNA fiber axis (29); and direct DNA strand scission by Ni(II)•Xaa-Gly-His that involves abstraction of the minor groove accessible C4'-H (28).

These early studies suggested that Ni(II)•Gly-Gly-His derived metallopeptides likely recognize DNA through insertion of the amino-terminal nitrogen and His imidazole “edges” of their approximately square-planar equatorial

planes into the DNA minor groove (Figure 2). Such an interaction would permit the formation of intermolecular hydrogen bonds between the pyrrole N-H of the His imidazole and the amino-terminal N-H protons of the main equatorial plane of the metallopeptide to prominent hydrogen bond acceptors found on the floor of the minor groove of A/T-rich regions (*i.e.*, the O2 of T residues and the N3 of A residues). In the presence of Arg or Lys within the amino-terminal peptide position, additional interactions with the minor groove are possible.

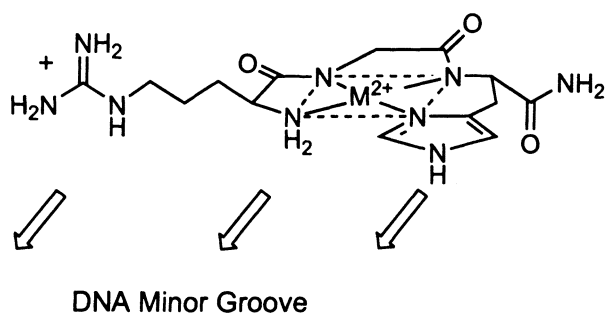


Figure 2. Potential DNA minor groove recognition elements present on the amino-terminal/His imidazole "edge" of an  $M(II)\cdot\text{Arg-Gly-His}$  metallopeptide.

Given their amino acid composition,  $M(II)\cdot\text{Xaa-Xaa-His}$  metallopeptides serve as synthetically accessible models of both protein- and drug-nucleic acid interactions. Perhaps of direct relevance to the former, these systems may mimic the amino-terminal  $\text{Cu}^{2+}$  or  $\text{Ni}^{2+}$  binding Arg-Thr-His motif found in the human DNA binding protein protamine P2 (4). Interestingly, metallopeptide sequences selected from a library based on their DNA cleavage propensities revealed  $\text{Ni(II)}\cdot\text{Arg-Thr-His}$  to be a preferred sequence; in general, the amino acid residues selected from this same library were found to be those that occur frequently in known DNA minor groove targeted protein motifs such as A/T hooks and SPKK motifs (25). Further, as likely minor groove recognizing agents, these systems also model features of DNA recognition used by peptide-based natural products such as netropsin and distamycin (30) and, given the presence of a metal center and ability to mediate  $\text{C4}'\text{-H}$  bond abstraction, the bleomycin class of minor groove binding antitumor "metallopeptides" (31-37). Indeed, similar to metal-bound bleomycin,  $M(II)\cdot\text{Xaa-Xaa-His}$  metallopeptides have been noted to mediate very efficient single- and double-strand DNA cleavage (28, 38, 39).



The above systems thus shed light on fundamental nucleic acid molecular recognition and reactivity events and may point the way toward the development of peptidomimetics with drug-like properties. This is an especially interesting prospect given their recently reported ability to recruit metal ions *in vivo* resulting in activity against HIV Rev response element RNA in mammalian cells (40). Perhaps related, early-on Cu(II)•Gly-Gly-His in the presence of ascorbate was noted as having activity against Ehrlich ascites tumor cells (41). Thus, further studies of M(II)•Xaa-Xaa-His metallopeptides are likely to impact the design of agents that function under physiological conditions.

To date, investigations of M(II)•Xaa-Xaa-His metallopeptide-DNA recognition *in vitro* have relied almost exclusively upon metallopeptide-induced DNA cleavage and the use of high-resolution DNA restriction fragment analyses to report on their site-selective DNA interactions (5, 24); however, to rely upon DNA scission chemistry alone as a reporter of site-selectivity may be misleading given that recognition and cleavage events may not always be tightly coupled (42). Accordingly, our recent studies have sought to assess the DNA recognition mode used by Ni(II)•Gly-Gly-His derived metallopeptides through a combination of experimental and computational approaches including the use of DNA-fiber EPR (43), 1D and 2D NMR (44), MD simulations (44-46), and comparative investigations using diastereomeric metallopeptides containing Arg and Lys residues with both L and D  $\alpha$ -carbon stereocenters (46). This chapter will therefore focus on the outcome of investigations directed towards understanding the DNA recognition of these metallopeptide systems *in the absence of cleavage reactivity*. In doing so the ability of these metallopeptides to act as models for both protein- and natural product-DNA interactions will be underscored.

## Evaluation of DNA Minor Groove Recognition

A combined experimental-computational approach using NMR spectroscopy and parallel, but independent, molecular dynamics simulations was pursued to define the DNA recognition mode used by Ni(II)•Xaa-Gly-His metallopeptides (44). As stated, all available evidence indicates a metallopeptide-minor groove interaction that occurs at A/T-rich DNA regions. A/T-rich regions of DNA are generally characterized as having narrow minor groove widths of reasonable depth that promote the binding of complementary-shaped, low molecular weight ligands such as the natural products netropsin and distamycin (30) or agents such as Hoechst 33285 (47). Such agents typically form hydrogen bonds with prominent acceptors found at the floor of this groove: the N3 of A residues and the O2 of T residues (47).

## NMR Analyses

Using the well-studied oligonucleotide: 5'-d(CGCGAATTCGCG)<sub>2</sub> (47), qualitative 500 MHz NMR analyses were carried out to define the location and orientation of the metallopeptide-oligonucleotide interaction. Initially, 1D NMR titrations, involving incremental increases in the metallopeptide-to-DNA ratio, indicated that each metallopeptide [Ni(II)•L-Arg-Gly-His and Ni(II)•D-Arg-Gly-His] produced observable changes in its respective 1D spectra: while the imino resonances of the G•C base pairs flanking the A/T-core were affected differently, those furthest away from the core being affected less than those closer to the core, the imino resonances of the A•T base pairs were broadened equally. Similarly, among the non-exchangeable nucleobase resonances, the C8-H proton resonances of the A residues were broadened with increasing metallopeptide concentration while the analogous C8-H guanine resonances were much less affected. In addition, the non-exchangeable C5-methyl resonances of the T residues were broadened upon increases in the metallopeptide-to-DNA ratio with the T<sub>7</sub> residue resonances being more perturbed than those of the T<sub>8</sub> methyls (residues numbered sequentially from the oligonucleotide 5'-terminus); these results indicated that the T<sub>7</sub> methyls and adenine C8-Hs are in closer proximity to the Ni center than the T<sub>8</sub> residues and the guanine C8-Hs. Taken together, these results suggest that the central Ni of the metallopeptide is located near the dyad axis of the oligonucleotide substrate.

Complementing these 1D experiments, 2D NMR (NOESY) analyses were also conducted. In these analyses a titration also was carried out to permit the straightforward differentiation of intramolecular DNA cross-peaks from those that occur due to either DNA-associated metallopeptide intramolecular or from DNA-metallopeptide intermolecular interactions. Of primary significance to this study, as a function of increasing metallopeptide concentration a moderate intensity cross-peak that connected 7.64 and 7.01 ppm – the positions of the adenylyl H-2 of A<sub>6</sub> and His imidazole ring C4 proton of Ni(II)•L-Arg-Gly-His [and Ni(II)•D-Arg-Gly-His] – appeared and increased in intensity (Figure 3). This observation indicated that the His imidazole ring of the metallopeptide was approximately 5-6 Å from the A<sub>6</sub> adenylyl C2H verifying the insertion of this metallopeptide moiety into the minor groove of the AATT core in close proximity to the minor groove floor.

## Computational Analyses/Molecular Dynamics Simulations

In an effort that was conducted in parallel but initially blind to the outcome of NMR investigations, computational examinations of the above metallo-peptides and DNA were carried out (44-46) using protocols similar to those used

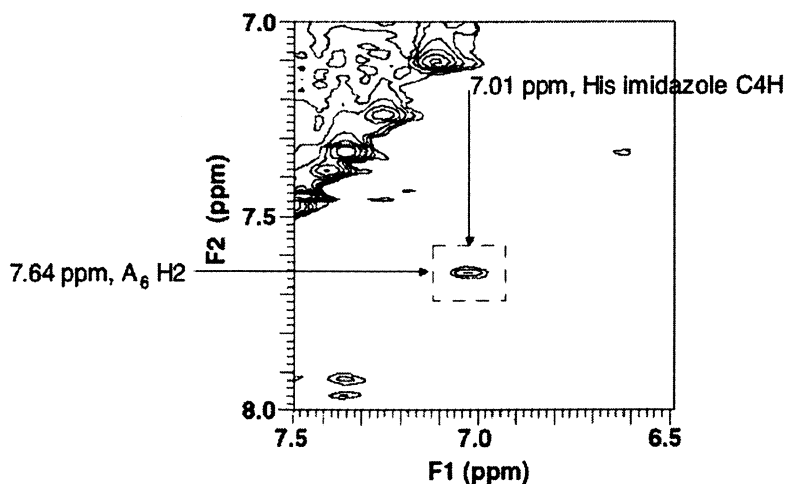


Figure 3. Expansion of the region around the DNA-metallopeptide intermolecular cross-peak observed in the 1:1 DNA: Ni(II)•L-Arg-Gly-His complex NOESY spectrum.

by others (48-50) in the analysis of minor groove binding ligands. In each of these investigations an available crystal structure (51) of netropsin-bound CGCGAATTCGCG (PDB reference code 1D86) was used to create an A/T-rich minor groove binding site by replacing netropsin with each metallopeptide intended for study. In each case the metallopeptide structures were assembled using an available crystal structure (52) to which appropriate side chain functionalities were appended; as described fully elsewhere (44), these structures were used also to develop appropriate force field parameters.

Given that each of the four “edges” of these approximately square planar metallopeptides present several combinations of potential hydrogen bond donors and acceptors, we sought to compare the relative stabilities of each of these four possible minor groove binding modes directly through unrestrained MD simulations (45). Ni(II)•Gly-Gly-His and Ni(II)•Arg-Gly-His were docked manually into the minor groove at the dyad of the DNA substrate described above using these four possible orientations (Figure 4). Subsequently, each assembled metallopeptide-DNA complex was imported into AMBER, counterions were added to achieve electroneutrality, and each complex was solvated explicitly using TIP3P water boxes to complete each system (45). After energy minimization, equilibration, and a gradual loosening of restraints resulting ultimately in a restraint-free system, production runs in excess of 1400 ps were carried out.

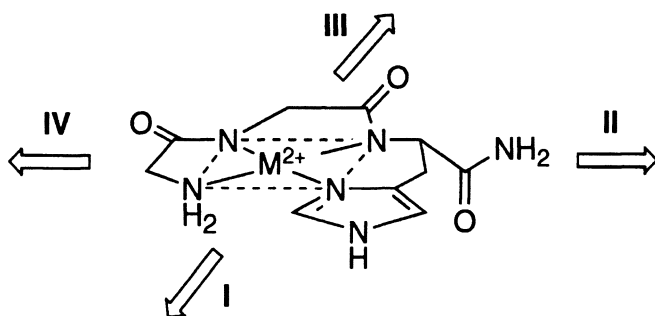


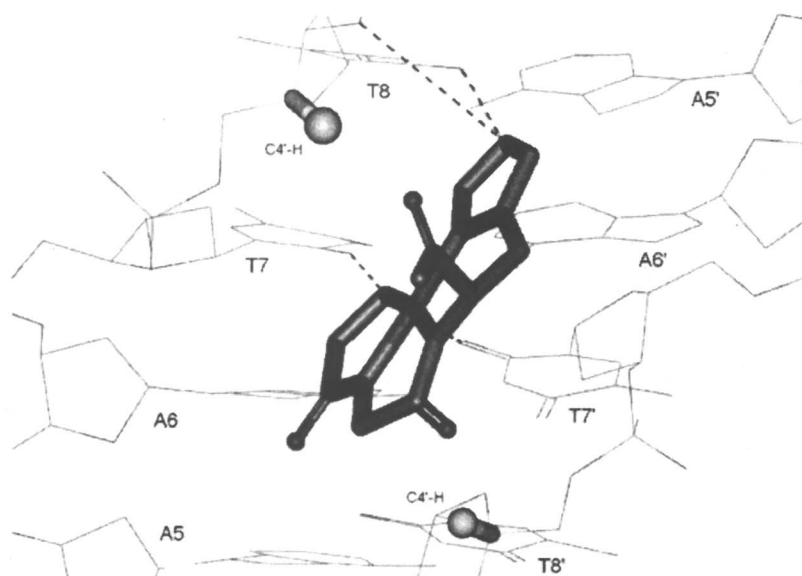
Figure 4. Binding orientations used during metallopeptide-minor groove docking and starting structure generation.

### *Ni(II)•Gly-Gly-His*

Analysis of the MD simulations resulting from the four docked orientations of Ni(II)•Gly-Gly-His indicated that orientation I maintained the key interactions found in its starting structure. In this orientation the metallopeptide remained located at the oligonucleotide dyad with its equatorial plane inserted into the minor groove and parallel to the minor groove walls. During the course of the simulation, the metallopeptide displayed dynamic motions relative to the DNA that involved a “rocking” of the metallopeptide in the minor groove, *i.e.*, clockwise and counterclockwise rotations about the C<sub>n</sub> axis of the metal center. Interestingly, the extremes of this motion resembled the starting structures found in orientations II and IV. In fact, it was observed that the initial docking orientations II and IV both immediately turned, respectively, to an orientation that resembled orientation I and, upon equilibration, commenced motions as described for this starting orientation. Distinctly different from orientations I, II, and IV, metallopeptides docked initially in orientation III were observed to move away from this position on the DNA indicating a very weak attraction or even a repulsion between the host and guest.

The above simulations indicate that the most favored orientation for the interaction between Ni(II)•Gly-Gly-His and the AATT core was through the metallopeptide amino terminus and His imidazole edge. In orientation I, the metallopeptide is able to form hydrogen bonds with the floor of the minor groove. These interactions included hydrogen bonds from the metallopeptide terminal amine N-H to the O2 of T7, O2 of T7', and O4' and O5' of T8; from the pyrrole N-H of the His side chain to O2 and O2P of T8; and from the metallopeptide carboxy-terminal amide to various hydrogen bond acceptors (Figure 5). While relatively stable in this orientation compared to the others, in the absence of other stabilizing interactions (such as a positively-charged side chain), this particular metallopeptide remained somewhat independent of the

DNA and is not held firmly in the groove resulting in the rocking motions described. These observations are consistent with the relatively random and inefficient DNA cleavage produced by this metallopeptide. In addition to the visual analysis of the MD trajectories, analysis of the relative binding energies of each of the four orientations (I – IV) agree with these qualitative observations: orientation I results in the lowest relative binding energy while orientation III leads to the highest relative binding energy ( $\Delta\Delta E$  for orientations II, III, and IV were 6.2, 90.1, and 0.2 kcal/mol, respectively;  $\Delta\Delta E = \Delta E_{\text{complex}} - \Delta E_{\text{DNA}} - \Delta E_{\text{ligand}}$ ).



*Figure 5. Illustration of the hydrogen bonding network observed in the average structure derived from simulations of Ni(II)•Gly-Gly-His bound to 5'-d(CGCGAATTCGCG)<sub>2</sub> via orientation I.*

#### *Ni(II)•L-Arg-Gly-His and Ni(II)•L-Lys-Gly-His*

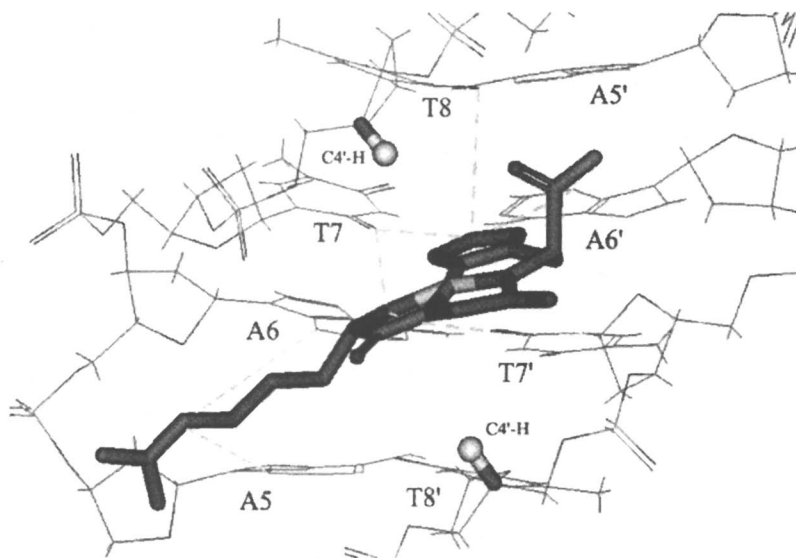
Simulations of Ni(II)•L-Arg-Gly-His were carried out using the same four docking orientations described for Ni(II)•Gly-Gly-His; the main difference in this case being the presence of the Arg side chain “extension” of the main metallopeptide equatorial plane – this “extension” could either complement or complicate metallopeptide-minor groove recognition. Visual analysis of the trajectories resulting from these simulations indicated that the system initiated from orientation I was stable and remained in its starting orientation throughout

the trajectory, *i.e.*, the metallopeptide remained associated with the minor groove and its motions paralleled and complemented those of the DNA substrate indicating favorable interactions. This particular orientation, unlike the others described below, also permitted a deep insertion of the metallopeptide into the groove and maintained a distance (3-5 Å) between the NOE partners noted earlier that was entirely compatible with the observation of an NOE by NMR. Similarly, the distances maintained between the Ni center and proximal C4'-H targets for strand scission were most compatible in orientation I. Noteworthy, the Arg side chain, while remaining flexible, was found to remain localized in the minor groove and interacted transiently with hydrogen bond acceptors located there; these observations were, in fact, quite reminiscent of results obtained from similar studies of Hoechst 33285 and netropsin (48-50).

In contrast to orientation I, the remaining orientations (II – IV) led to comparatively low levels of intermolecular interaction between the DNA groove and the metallopeptide; the interactions observed can be accounted for simply due to the presence of the positively-charged Arg residue which, in contrast to the charge neutral Gly-Gly-His system, assisted in maintaining this particular metallopeptide in proximity to the DNA. These qualitative observations were also supported by the relative binding energies calculated for the systems: orientation I is the most stable while all other orientations are significantly less stable ( $\Delta\Delta E$  for orientations II, III, and IV were 53.8, 129.6, and 100.9 kcal/mol, respectively;  $\Delta\Delta E = \Delta E_{\text{complex}} - \Delta E_{\text{DNA}} - \Delta E_{\text{ligand}}$ ).

Overall, these results indicate that the most favored orientation for Ni(II)•L-Arg-Gly-His minor groove association is I. In this orientation, the metallopeptide fits snugly in the minor groove with little independent motion indicative of a relatively strong intermolecular interaction. This association mode permits the metallopeptide equatorial plane to be inserted into the minor groove equidistant from, and parallel to, the groove walls and simultaneously associated with the floor of the minor groove via the amino-terminal N-H, the pyrrole N-H of the His imidazole to the O2 of T<sub>7</sub> and T<sub>8</sub> and N3 of A<sub>6</sub> residues (Figure 6); these interactions were found to be stable and long-lived (up to 80% of the simulation time) and maintain a His C4-H to A<sub>6</sub> C2-H distance of 3-5 Å. In addition, the side chain of Arg is extended in the minor groove like an “outrigger” and makes transient contacts within the minor groove via its guanidinium functionality (Figure 6). Due to the L-stereochemistry of the Arg residue, the overall curvature of this particular metallopeptide diastereoisomer appears to complement well the curvature of the minor groove producing an advantageous “isohelical” fit that allows it to maintain interactions and complement the motions of the DNA substrate.

Oriented in the fashion described above, and with a transient third point of interaction mediated by the Arg side chain, Ni(II)•L-Arg-Gly-His interacts strongly with the DNA minor groove in comparison to the Gly-only metallo-



*Figure 6. Illustration of the hydrogen bonding network observed in the average structure determined from simulation of Ni(II)•L-Arg-Gly-His bound to 5'-d(CGCGAATTCGCG)<sub>2</sub> via orientation I.*

peptide; however despite their differences, both Ni(II)•L-Arg-Gly-His and Ni(II)•Gly-Gly-His appear to insert their equatorial planes to a comparable depth into the minor groove leading to a similar positioning of their respective Ni centers relative to the C4'-H targets of their known chemical reactivity: the Ni centers of these systems were approximately 4 Å from the C4' protons of the T<sub>7</sub> and T<sub>8</sub> residues thus explaining their common pathways to DNA cleavage. Interestingly, abstraction of these DNA protons would lead to the 3'-asymmetric cleavage patterns observed during restriction fragment cleavage analyses (26). Despite their similarities, these metallopeptides cleave DNA with differing efficiencies: Ni(II)•Gly-Gly-His mediates the relatively non-selective, inefficient abstraction of C4'-H while Ni(II)•L-Arg-Gly-His cleaves only a subset of available nucleotides at a given A/T-rich site supporting the notion of a more selective and locally stabilized interaction within the minor groove.

Effectively identical to the behavior of Ni(II)•L-Arg-Gly-His, Ni(II)•L-Lys-Gly-His also targets A/T-rich regions of DNA through C4'-H abstraction chemistry but with slightly less efficiency in comparison to its Arg-containing analogue (26). Accordingly, simulations of this system were also carried out using a starting structure in orientation I. Analysis of the simulations resulting from this system indicated that Ni(II)•L-Lys-Gly-His moves and flexes with the

A/T core of the DNA substrate in a fashion that is virtually identical to the behavior of Ni(II)•L-Arg-Gly-His. Indeed a detailed analysis of the MD simulation of this system indicated a clear correspondence to Ni(II)•L-Arg-Gly-His except for the behavior of the Lys side chain: the Lys side chain was observed to alternate between two conformations consequently interacting less strongly with the DNA than the Arg side chain. Given that the main DNA recognition moieties of these metallopeptides are identical (terminal amine and His imidazole), these observations suggest a plausible rationale for why Ni(II)•L-Lys-Gly-His and Ni(II)•L-Arg-Gly-His display identical cleavage site-selectivities but that the former consistently displays weaker DNA cleavage.

## Summary

NMR studies verify the A/T-rich minor groove targeting and orientation displayed by M(II)•Xaa-Xaa-His metallopeptides during their recognition of DNA. When coupled to independent MD simulations, these studies support the evidence for minor groove recognition presented earlier and provide further insight into details of the interactions that occur: these metallopeptides insert their amino terminal N-H and His imidazole edges into the minor groove and interact via hydrogen bonding with the floor of the minor groove of A/T-rich regions; metallopeptides that possess a positively-charged Arg or Lys side chain in their first peptide position are able to make additional favorable groove interactions. These studies indicate the origins of the A/T-rich DNA selectivity displayed by these systems by demonstrating the hydrogen bonding potential of these regions of DNA and their close overall complementarity to the size and shape of these metallopeptides. In fact, the structure of AATT-bound Ni(II)•L-Arg-Gly-His averaged from MD simulations indicated a surprising correspondence to the crystal structures of netropsin and Hoechst 33285 bound to this same DNA site: the act of metal binding and minor groove association by Arg-Gly-His leads to an alignment of peptide N-H hydrogen bond donors that are very similar to those of netropsin and Hoechst dye. These three systems align (Figure 7) with an RMS deviation of 0.52 Å to yield structures that are very similar with regards to their curvature, thickness, and N-H donor positions.

In light of the similarity of M(II)•Arg-Gly-His metallopeptides to agents that target the DNA minor groove, knowledge of the presence of an amino-terminal Arg-Thr-His Cu<sup>2+</sup> binding motif in the human DNA binding and condensing protein protamine P2 (4), as noted earlier, lends itself naturally to speculation as to whether or not this protein motif is also a native minor groove binding structure. Indeed, this motif is known to influence the DNA binding of protamine upon metal binding (4); however whether the role of this motif is in DNA binding, metal sequestration, or some other yet-to-be-defined role has not been determined.



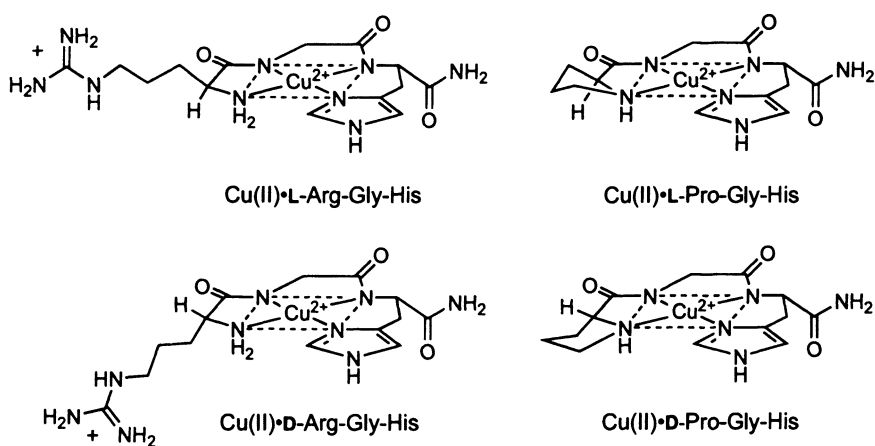


*Figure 7. Comparison of the average structure of Ni(II)•L-Arg-Gly-His (black structure) to those of netropsin and Hoechst 33285 emphasizing the close correspondence of their nitrogen atoms (dark spheres) involved in DNA minor groove recognition.*

### **Analysis of Metallopeptide Diastereoisomers**

The general recognition mode described above leads to a minor groove associated metallopeptide that directs an amino-terminal side chain into the minor groove while amino acid side chains present at the second Xaa position would be exposed at the surface of the minor groove and likely directed away from the minor groove floor. Consequently, the amino acid present in the first peptide position is poised to influence minor groove recognition and “fit” while the amino acid present in the second peptide position is likely to influence groove recognition to a far lesser extent. This overall view parallels the DNA cleavage behavior of these systems: L-Arg/Lys residues present in the amino-terminal position of these metallopeptides lead to increased site-selective DNA cleavage, *i.e.*, only a subset of nucleotides within a given A/T-rich recognition site are cleaved, whereas D-Arg/Lys residues in this same peptide position decrease site-selectivity, often manifested as uniform cleavage of all nucleotides within a given A/T-rich target site (46). This is likely due to a decreased structural complementarity between the D-Arg isomer and the minor groove and an interaction that does not allow the metallopeptide to discriminate easily between available cleavage sites. Indeed, the angle of side chain projection from the main equatorial plane of the L-Arg containing metallopeptide isomer is wholly different from that of the D-Arg substituted system (Figure 8). In further

support of the above view, substitution of L- and D-Pro residues in the amino-terminal peptide position, a compact, non-extended side chain substitution that minimizes the steric differences between these two stereochemistries (Figure 8), leads to virtually identical DNA cleavage patterns. In comparison to amino-terminal substituted systems, DNA cleavage patterns generated by systems substituted in the second peptide position [Ni(II)•Gly-L/D-Arg-His and Ni(II)•Gly-L/D-Lys-His] exhibited nearly identical DNA cleavage patterns. In fact, these patterns were most consistent with the cleavage pattern induced by Ni(II)•Gly-Gly-His, albeit with increased intensity due to the positive charge now present.



*Figure 8. Schematic structures of metallopeptide diastereoisomeric pairs emphasizing the differences in the projection of their side chain functional groups from the main equatorial metallopeptide plane.*

To examine the basis for the cleavage behavior noted above and to explore details of these stereoisomer-driven DNA recognition differences, we carried out DNA-fiber EPR (43) and additional MD simulations (46). These analyses permitted the comparison of diastereomeric metallopeptide pairs and an examination of the influence of their side chain substitutions on minor groove recognition.

### DNA Fiber EPR Analyses

DNA fiber EPR analyses of Cu(II)•Xaa-Xaa-His metallopeptides were conducted in collaboration with M. Chikira and coworkers (Chuo University,

Tokyo, Japan) (29, 43). DNA fiber EPR can provide unique information on the DNA binding stereospecificities and dynamics of paramagnetic complexes (53): if a paramagnetic complex binds stereospecifically to DNA fabricated into highly orientated fibers, the EPR spectra will change as a function of the angle between the DNA fiber axis and the static magnetic field. Simulations of the resulting spectra reveal information pertaining to the orientation and randomness of the DNA bound complexes. During the course of earlier DNA fiber EPR studies of Cu(II)•Gly-Gly-His, Cu(II)•L-Arg-Gly-His, and Cu(II)•L-Lys-Gly-His (29), we determined that the  $g_{\parallel}$  axes and mean equatorial planes of these systems were tilted approximately 50° and 40°, respectively, relative to the DNA fiber axis in a fashion consistent with the stereospecific orientation of these metallopeptides in the minor groove.

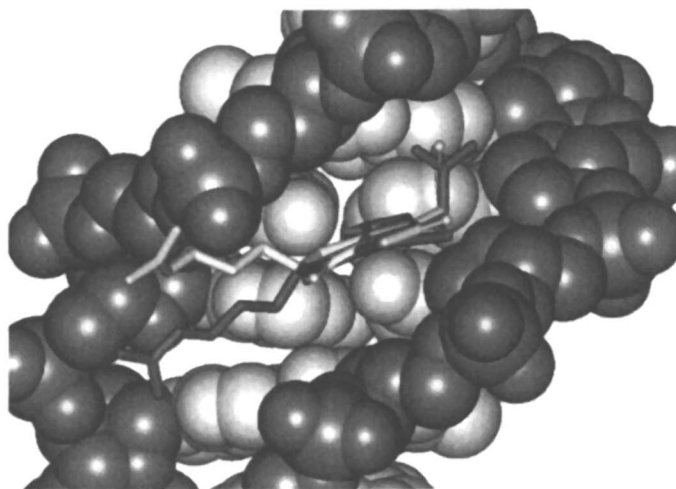
More recently, in light of the DNA cleavage behavior of diastereomeric metallopeptides, Cu(II)•D-Arg-Gly-His and Cu(II)•D-Lys-Gly-His were examined by DNA fiber EPR for comparison to Cu(II)•L-Arg-Gly-His and Cu(II)•L-Lys-Gly-His (43). The DNA fiber EPR analyses of Cu(II)•D-Arg-Gly-His and Cu(II)•D-Lys-Gly-His produced spectra that exhibited a somewhat less conspicuous angular dependence in comparison to the L-isomers, indicating that these systems were oriented more randomly on the DNA fibers. In comparison to Cu(II)•L-Arg-Gly-His and Cu(II)•L-Lys-Gly-His, analyses of Cu(II)•D-Arg-Gly-His and Cu(II)•D-Lys-Gly-His produced an estimation of 45° for the orientation of their mean equatorial planes relative to the DNA fiber axis; this is in relatively good agreement with the same angle measured for the L-isomers: 40°. In stark contrast to the L-isomers, however, analyses of these D-isomers indicated  $\Delta\theta$  values (a measure of the degree of metal complex orientational randomness on the DNA) that were *twice* that of the L-isomers ( $\Delta\theta = 40^\circ$  vs. 20°). These results indicate that the stereochemistry of the amino terminal metallopeptide residue is critical to the attainment of a stable, stereospecific interaction with the DNA minor groove and support the notion that the D-isomeric systems participate in an overall DNA orientation that while similar to their L-substituted counterparts, is more dynamic. These results clearly parallel the results of DNA cleavage analyses using these same peptide sequences.

## Computational Analyses/Molecular Dynamics Simulations

### *Ni(II)•D-Arg-Gly-His*

In comparison to the MD analysis of Ni(II)•L-Arg-Gly-His already described, simulations of Ni(II)•D-Arg-Gly-His initiated from a starting structure corresponding to orientation I indicated that this metallopeptide diastereoisomer retained the use of the main hydrogen bond donors already described: Ni(II)•D-Arg-Gly-His used its amino-terminal N-H, His imidazole N-H, and side chain

guanidinium functionality to interact with features of the A/T-rich minor groove. However, this particular system employed fundamentally different strategies for their use. Unlike the L-Arg isomer, Ni(II)•D-Arg-Gly-His was observed to be much less stable in its groove association and acted independent of the DNA. This metallopeptide shifted positions within the A/T-rich core and, quite noticeably, the D-Arg residue stereochemistry and resulting angle of Arg side chain projection from the main equatorial metallopeptide plane created a steric block causing the metallopeptide to form only partial contacts with the floor of the minor groove, in particular the critical contacts that form with the amino-terminal N-H protons. Lacking the complementary isohelical “fit” observed for the L-Arg isomer (Figure 9), Ni(II)•D-Arg-Gly-His maintained structural independence and alternated between two sets of contact points on opposite sides of the minor groove. Given this reduced “fit”, Ni(II)•D-Arg-Gly-His was also observed to slide along the minor groove floor placing it in proximity to the C4'-H positions of multiple target nucleotides. While visibly different in comparison to the L-Arg isomer, Ni(II)•D-Arg-Gly-His did not simply interact with DNA like Ni(II)•Gly-Gly-His because it still maintained three points of contact with the minor groove floor. The above observations are consistent both with the outcome of DNA fiber EPR studies that indicated dynamic behavior and also the DNA cleavage patterns described earlier.



*Figure 9. Overlay of space-filling models emphasizing the differences in the locations of the side chains of Ni(II)•L-Arg-Gly-His (dark metallopeptide) vs. Ni(II)•D-Arg-Gly-His (light metallopeptide) while bound in the minor groove of 5'-d(CGCGAATTCGCG).*

*Ni(II)•L-Pro-Gly-His and Ni(II)•D-Pro-Gly-His*

Metallopeptides containing amino-terminal Pro residues do not exhibit the distinct diastereoisomeric differences displayed by the Arg substituted systems: DNA cleavage patterns induced by the L-Pro and D-Pro substituted systems were found to be virtually identical (46). This observation is supported by MD simulations; in simulations of both Pro-substituted isomers it was found that the pyrrole N-H of the His imidazole was the only primary means of contact with the minor groove floor. Unlike other metallopeptides, both Pro metallopeptides are blocked from forming contacts with the minor groove floor via their terminal imine nitrogens due to the presence of the five-membered Pro ring. Consequently, both metallopeptide diastereoisomers use their His pyrrole N-H to mediate a single main point of contact with the DNA and in doing so maintained an overall similar minor groove interaction. This is reflected in the nearly identical DNA cleavage patterns observed. These findings support the notion that an extended side chain appendage to the main equatorial plane at the amino-terminal position can influence recognition to a much greater extent than a sterically compact side chain.

*Ni(II)•Gly-L-Arg-His and Ni(II)•Gly-D-Arg-His*

Having examined the influence of substitutions in the amino-terminal metallopeptide position, substitutions were also examined within the middle peptide position. In orientation I, DNA docked metallopeptides of this type permit the Arg side chain of either stereochemistry to project initially away from the surface of the DNA. Upon simulation, these systems were immediately noted to behave similar to one-another and to remain aligned parallel to the walls of the minor groove at the center of the A/T-rich region; the simulations indicate that these metallopeptides and the hydrogen bond contacts they form with the minor groove are similar to one another and very much like those observed for Ni(II)•Gly-Gly-His – the interactions with DNA that occurred were mainly through the amino-terminal N-H protons to the O2 of T<sub>7</sub> on two strands. In comparison to the behavior of the main equatorial plane, the guanidinium “tails” of both metallopeptides, now present in the middle peptide positions, bend towards the minor groove and interact with acceptors along its edge. In this position relative to the DNA groove, the stereochemistry of the Arg substitution does not appear to markedly influence its point of interaction with the DNA allowing the main equatorial planes of both diastereoisomers to interact with the minor groove of the DNA substrate similarly. In fact, this positioning of the positively-charged Arg side chain and the contacts it now forms outside of the minor groove serves to disrupt the His imidazole-minor groove interactions that would likely otherwise occur making the overall interactions that occur in the

minor groove very much like those of Ni(II)•Gly-Gly-His. This aspect provides a clear rationale for why the DNA cleavage patterns of these two diastereoisomers resemble one another and the pattern produced by Ni(II)•Gly-Gly-His (46).

### Summary

The additional studies described above further support the recognition model developed earlier and rationalize the diastereoisomeric differences noted during DNA cleavage analyses. DNA fiber EPR studies indicate that while the predominant orientation of these metallopeptides continues to occur via a stereospecific alignment in the minor groove, metallopeptides containing D-Arg or D-Lys residues in their terminal amino acid positions are much more dynamic in their DNA binding behavior. In the amino-terminal metallopeptide position, amino acid substitutions such as D-Arg or D-Lys with side chains that can extend into the minor groove produce metallopeptide structures that no longer complement the curvature of the minor groove as seen with L-Arg or L-Lys substitutions. These substitutions lead to minor groove-metallopeptide destabilization resulting in metallopeptides that can slide along the minor groove and interact with multiple sites. In contrast, amino acid substitutions such as Pro have less impact due to their constrained, compact nature leading to diminished steric differences between their respective diastereoisomeric metallopeptides. Similarly, in the middle position the presence of D-Arg or D-Lys vs. L-Arg or L-Lys substitutions have little impact beyond increasing electrostatic attraction, leading to interactions that are similar to those described for Ni(II)•Gly-Gly-His.

In addition to explaining aspects of M(II)•Xaa-Xaa-His metallopeptide-minor groove recognition, the above findings serve also to illustrate fundamental aspects of drug or small molecule-minor groove recognition. First, the particular location of a positively-charged functional group within a DNA-targeted molecule relative to its bound DNA substrate can either promote or complicate an overall DNA interaction. Second, subtle steric features of a minor groove targeted species can profoundly impact recognition activity. Third, the inclusion of a positively-charged functional group such as a guanidinium or a simple protonated amine into a DNA-targeted structure, while promoting similar overall interactions with the minor groove, appears to be consistently enhanced in the case of a guanidinium; this observation suggests one rationale for the predominance of guanidinium-containing functionalities in DNA-targeted natural products (30, 47).

### Conclusions

The studies described in this chapter support the model of metallopeptide-minor groove recognition proposed in light of DNA cleavage analyses (5, 24)

and provide further insight into this recognition mechanism. In general, metallopeptides recognize the DNA minor groove of A/T-rich regions through their ability to adopt an overall shape that is complementary to the minor groove width, depth, and curvature of A/T-rich regions and also through their ability to present hydrogen bond donor functionalities along one edge of their structures to hydrogen bond acceptors found at the floor of this DNA groove. As expected, the efficiency of this interaction is enhanced in the presence of positively-charged residues such as Arg or Lys. In these aspects of their DNA recognition activity, metallopeptides are reminiscent of natural product or synthetic minor groove binders such as netropsin, distamycin and Hoechst 33285 (but, as noted below, with lower affinities and increased exchange rates). In addition, in light of the metal-bound Arg-Thr-His motif present in human protamine P2, and the selection of this same sequence from a library of metallotriptides, it is possible that these metallopeptides directly model the activity of a native protein motif. While naturally-occurring L-amino acids, as would be found in native proteins, lead to metallopeptides such as M(II)•L-Arg-Xaa-His that complement well the curvature of the DNA minor groove and promote a relatively stable groove interaction, alternative structures that contain D-amino acids in the first, minor groove-directed peptide position either decrease this complementarity, if they contain extended side chains such as D-Arg, or have little influence if they are non-extended, compact side chains. Alternatively, amino acid substitutions in the middle peptide position lead to less influence on overall minor groove recognition, as expected based on this amino acid position relative to the DNA groove.

While these metallopeptide systems have been noted to recognize A/T-rich DNA and to promote very efficient single- and double-strand DNA cleavage (28, 38, 39), they likely do so during a relatively short-lived minor groove residence time; this notion is supported by observations indicating a lack of binding saturation reminiscent of enzyme-like behavior during kinetic analyses of DNA cleavage (38, 39). Similar support is derived from attempts to examine the site-selectivity of these systems via fluorescent intercalator displacement (FID) analyses (54-56) (as also discussed in Chapter 5). During the course of FID analyses it was determined that M(II)•Xaa-Xaa-His metallopeptides, including those with Arg or Lys substitutions, do not compete favorably with the strong DNA binding affinity of the reporting fluorescent intercalators ethidium bromide and thiazole orange; the metallopeptides failed to produce detectable decreases in sample fluorescence akin to a strong equilibrium driven DNA binder at reasonable levels of added metallopeptide. The results of these FID assays, using both Ni(II)• and Cu(II)•Xaa-Xaa-His metallopeptides, were, in fact, reminiscent of FID analyses of Cu(II)•bleomycin (see Chapter 5). Cu(II)•bleomycin, while competent to bind, unwind, and cleave DNA (57), was nonetheless unable to displace DNA hairpin-bound ethidium bromide. This observation is in accord with the finding that the interaction of Cu(II)•bleomycin with DNA is in fast

exchange relative to other metallobleomycins such as Co(III)• and Fe(III)•bleomycin (58). These results suggest that M(II)•Xaa-Xaa-His metallopeptides interact similarly with DNA.

Overall then, it is likely that the unique structural features of A/T-rich DNA regions provide an environment conducive to metallopeptide recognition and increased residence time to an extent commensurate with an act of C4'-H abstraction; the dynamic nature of this interaction and the extent of the residence time are also fine-tuned by structural aspects of the metallopeptide. While metallopeptide interactions unaccompanied by cleavage reactivity with other DNA sites cannot be ruled out, it appears clear from all experimental studies carried out so far that particular groove widths and hydrogen bonding features, and not simply groove accessibility or nucleotide composition, are major factors leading to metallopeptide-DNA recognition and strand scission. For example, along with the non-homopolymeric A/T-rich sites targeted most frequently by these systems, narrowed G/C sites such as CCT (59) promote and stabilize the interaction of some metallopeptides to an extent sufficient to allow C4'-H abstraction while homopolymeric polyA•polyT sites, narrowed to an extent greater than mixed A/T-rich regions (60), appear likely to block frequent metallopeptide association due to the less accessible nature of their minor grooves. In addition, like other minor groove targeted agents (30, 47), metallopeptides avoid canonical G/C-rich sites with relatively wide minor grooves and the presence of the exocyclic amine of G.

Thus, it is likely that the minor groove width and hydrogen bonding potential of mixed A/T-rich regions, and other DNA sites that resemble these regions, are neither too wide nor too narrow to accommodate a targeting metallopeptide. Curiously, the DNA cleavage patterns generated by these metallopeptides complement those produced by hydroxyl radicals (59) further supporting the notion that these metallopeptides do not simply target easily accessed groove sites as do hydroxyl radicals, but particular groove widths and features. Indeed the studies described herein likely describe the behavior of metallopeptides at such sites. Ultimately, such regions of DNA increase the residence time of the bound metallopeptide and also contribute to directing a metallopeptide catalyst towards a C4'-H target leading to increased reactivity as also influenced and fine-tuned by the overall shapes and chemical features (39) of these metallopeptide "catalysts".

## Acknowledgements

This work was supported by the National Institutes of Health (GM 62831 to E.C.L) and additional National Institutes of Health funding through the RCMI program, Division of Research Infrastructure, National Center for Research Resources (G12 RR003048 to Howard University).



## References

1. Harford, K; Sarkar, B. *Acc. Chem. Res.* **1997**, *30*, 123-130.
2. Camerman, N.; Camerman, A.; Sarkar, B. *Can. J. Chem.* **1976**, *54*, 1309-1316.
3. Melino, S.; Gallo, M.; Trotta, E.; Mondello, F.; Paci, M.; Petruzzelli, R. *Biochemistry* **2006**, *45*, 15373-15383.
4. Bal, W.; Wojcik, J.; Maciejczyk, M.; Grochowski, P.; Kasprzak, K. S. *Chem. Res. Toxicol.* **2000**, *13*, 823-830.
5. Long, E. C.; Claussen, C. A. In *DNA and RNA Binders: From Small Molecules to Drugs*; Demeunynk, M.; Bailly, C.; Wilson, W. D., Eds.; Wiley-VCH: New York, 2003; pp 88-125.
6. Kozlowski, H.; Bal, W.; Dyba, M.; Kowalik-Jankowska, T. *Coord. Chem. Rev.* **1999**, *184*, 319-346.
7. Chiou, S.-H. *J. Biochem.* **1983**, *94*, 1259-1267.
8. Mack, D. P.; Dervan, P. B. *Biochemistry* **1992**, *31*, 9399-9405.
9. Mack, D. P.; Dervan, P. B. *J. Am. Chem. Soc.* **1990**, *112*, 4604-4606.
10. Mack, D. P.; Iverson, B. L.; Dervan, P. B. *J. Am. Chem. Soc.* **1988**, *110*, 7572-7574.
11. Nagaoka, M.; Hagihara, M.; Kuwahara, J.; Sugiura, Y. *J. Am. Chem. Soc.* **1994**, *116*, 4085-4086.
12. Harford, C.; Narindrasorasak, S.; Sarkar, B. *Biochemistry* **1996**, *35*, 4271-4278.
13. Shullenberger, D. F.; Long, E. C. *Bioorg. Med. Chem. Lett.* **1993**, *3*, 333-336.
14. Grokhovsky, S. L.; Nikolaev, V. A.; Zubarev, V. E.; Surovaya, A. N.; Zhuze, A. L.; Chernov, B. K.; Sidorova, N.; Zasedatelev, A. S. *Mol. Biol. (Moscow)* **1992**, *26*, 1274-1297.
15. Morier-Teissier, E.; Boitte, N.; Helbecque, N.; Bernier, J. L.; Pommery, N.; Duvalet, J. L.; Fournier, C.; Hecquet, B.; Catteau, J. P.; Henichart, J. P. *J. Med. Chem.* **1993**, *36*, 2084-2090.
16. Steullet, V.; Dixon, D. W. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2935-2940.
17. De Napoli, L.; Messere, A.; Montesarchio, D.; Piccialli, G.; Benedetti, E.; Bucci, E.; Rossi, F. *Bioorg. Med. Chem. Lett.* **1999**, *7*, 395-400.
18. Truffert, J.-C.; Asseline, U.; Brack, A.; Thuong, N. T. *Tetrahedron* **1996**, *52*, 3005-3016.
19. Footer, M.; Egholm, M.; Kron, S.; Coull, J. M.; Matsudaira, P. *Biochemistry* **1996**, *35*, 10673-10679.
20. Van Dijk, J.; Lafont, C.; Knetsch, M. L. W.; Derancourt, J.; Manstein, D. J.; Long, E. C.; Chaussepied, P. *J. Muscle Res. Cell Motil.* **2004**, *25*, 527-537.
21. Bal, W.; Jezowska-Bojczuk, M.; Kasprzak, K. S. *Chem. Res. Toxicol.* **1997**, *10*, 906-914.

22. Bal, W.; Lukszo, J.; Kasprzak, K. S. *Chem. Res. Toxicol.* **1997**, *10*, 915-921.
23. Muller, J. G.; Hickerson, R. P.; Perez, R. J.; Burrows, C. J. *J. Am. Chem. Soc.* **1997**, *119*, 1501-1506.
24. Long, E. C. *Acc. Chem. Res.* **1999**, *99*, 827-836.
25. Huang, X.; Pieczko, M. E.; Long, E. C. *Biochemistry* **1999**, *38*, 2160-2166.
26. Liang, Q.; Eason, P. D.; Long, E. C. *J. Am. Chem. Soc.* **1995**, *117*, 9625-9631.
27. Brittain, I. J.; Huang, X.; Long, E. C. *Biochemistry* **1998**, *37*, 12113-12120.
28. Liang, Q.; Ananias, D. C.; Long, E. C. *J. Am. Chem. Soc.* **1998**, *120*, 248-257.
29. Nagane, R.; Koshigoe, T.; Chikira, M.; Long, E. C. *J. Inorg. Biochem.* **2001**, *83*, 17-23.
30. Zimmer, C.; Wahnert, U. *Prog. Biophys. Mol. Biol.* **1986**, *47*, 31-112.
31. Stubbe, J.; Kozarich, J. W. *Chem. Rev.* **1987**, *87*, 1107-1136.
32. Burger, R. M. *Chem. Rev.* **1998**, *98*, 1153-1169.
33. Boger, D. L.; Cai, H. *Angew. Chem. Int. Ed.* **1999**, *38*, 448-476.
34. Claussen, C. A.; Long, E. C. *Chem. Rev.* **1999**, *99*, 2797-2816.
35. Hecht, S. M. *J. Nat. Prod.* **2000**, *63*, 158-168.
36. Chen, J.; Stubbe, J. *Nat. Rev. Cancer* **2005**, *5*, 102-112.
37. Hecht, S. M. In *Anticancer Agents from Natural Products*; Cragg, G. M.; Kingston, D. G. I.; Newman, D. J., Eds.; CRC: Boca Raton, FL, 2005; pp 357-381.
38. Jin, Y.; Cowan, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 8408-8415.
39. Jin, Y.; Lewis, M. A.; Gokhale, N. H.; Long, E. C.; Cowan, J. A. *J. Am. Chem. Soc.* **2007**, *129*, 8353-8361.
40. Jin, Y.; Cowan, J. A. *J. Biol. Inorg. Chem.* **2007**, *12*, 637-644.
41. Kimoto, E.; Tanaka, H.; Gyotoku, J.; Morishige, F.; Pauling, L. *Cancer Res.* **1983**, *43*, 824-828.
42. McLean, M. J.; Dar, A.; Waring, M. J. *J. Mol. Rec.* **1989**, *1*, 184-192.
43. Hamada, H.; Abe, Y.; Nagane, R.; Fang, Y.-Y.; Lewis, M. A.; Long, E. C.; Chikira, M. *J. Inorg. Biochem.* **2007**, *101*, 1529-1536.
44. Fang, Y.-Y.; Ray, B. D.; Claussen, C. A.; Lipkowitz, K. B.; Long, E. C. *J. Am. Chem. Soc.* **2004**, *126*, 5403-5412.
45. Fang, Y.-Y.; Lipkowitz, K. B.; Long, E. C. *J. Chem. Theory Comput.* **2006**, *2*, 1453-1463.
46. Fang, Y.-Y.; Claussen, C. A.; Lipkowitz, K. B.; Long, E. C. *J. Am. Chem. Soc.* **2006**, *128*, 3198-3207.
47. Neidle, S. *Nat. Prod. Rep.* **2001**, *18*, 291-309.
48. Wellenzohn, B.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *J. Am. Chem. Soc.* **2000**, *122*, 3927-3931.
49. Wellenzohn, B.; Flader, W.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *Biophys. J.* **2001**, *81*, 1588-1599.

50. Flader, W.; Wellenzohn, B.; Winger, R. H.; Hallbrucker, A.; Mayer, E.; Liedl, K. R. *J. Phys. Chem. B* **2001**, *105*, 10379-10387.
51. Sriram, M.; van der Marel, G. A.; Roelen, H. L.; van Boom, J. H.; Wang, A. H. J. *Biochemistry* **1992**, *31*, 11832-11834.
52. Bal, W.; Djuran, M. I.; Margerum, D. W.; Gray, E. T.; Mazid, M. A.; Tom, R. T.; Nieboer, E.; Sadler, P. J. *J. Chem. Soc., Chem. Commun.* **1994**, 1889-1890.
53. Chikira, M. *J. Inorg. Biochem.* **2008**, *102*, 1016-1024.
54. Tse, W. C.; Boger, D. L. *Acc. Chem. Res.* **2004**, *37*, 61-69.
55. *Drug-DNA Interaction Protocols*; Fox, K. R., Ed.; Methods in Molecular Biology; Humana: Totowa, NJ, 1997; Vol. 90.
56. Lewis, M. A.; Long, E. C. *Bioorg. Med. Chem.* **2006**, *14*, 3481-3490.
57. Ehrenfeld, G. M.; Shipley, J. B.; Heimbrook, D. C.; Sugiyama, H.; Long, E. C.; van Boom, J. H.; van der Marel, G. A.; Oppenheimer, N. J.; Hecht, S. M. *Biochemistry* **1987**, *26*, 931-942.
58. Li, W.; Zhao, C.; Xia, C.; Antholine, W. E.; Petering, D. H. *Biochemistry* **2001**, *40*, 7559-7568.
59. Sy, D.; Savoye, C.; Begusova, M.; Michalik, V.; Charlier, M.; Spothheim-Maurizot, M. *Int. J. Radiat. Biol.* **1997**, *72*, 147-155.
60. Burkoff, A. M.; Tullius, T. D. *Cell* **1987**, *48*, 935-943.



## **Indexes**



# Author Index

- Baik, Mu-Hyun, 151  
Baldwin, Michael J., 133  
Basu, Partha, 199  
Cowan, J. A., 3  
Crowder, Michael W., 81  
Dancis, Andrew, 17  
Easton, J. Allen, 81  
Edison, Sara E., 133  
Fang, Ya-Yin, 219  
Georgiadis, Millie M., 63  
Goldcamp, Michael J., 133  
Goodwin, Kristie D., 63  
Grapperhaus, Craig A., 99  
Gunasekera, Thusitha, 81  
Haven, Micheal, 133  
Hegg, Eric L., 31, 47  
Hong, Jing, 167  
Iranzo, Olga, 183  
Johnson, Mitchell E., 199  
Kail, Brian W., 199  
Kharenko, Olesya A., 167  
Klingbeil, Lindsey, 81  
Kondapalli, Kalyan C., 17  
Krause, Jeanette A., 133  
Lewis, Mark A., 63, 219  
Long, Eric C., 63, 219  
Lord, Richard L., 151  
Łuczkowski, Marek, 183  
O'Toole, Martin G., 99  
Ogawa, Michael Y., 167  
Peacock, Anna F. A., 183  
Pecoraro, Vincent L., 183  
Schultz, Franklin A., 151  
Squires, Leah N., 133  
Stemmler, Timothy L., 17  
Sugerbaker, Stacy A., 81  
Tsurkan, Mikhail V., 167  
Wang, Zhihong, 47  
Yang, Xiaofan, 151  
Young, Charles G., 199  
Ziegler, Christopher J., 115





# Subject Index

- “Agostic coordination,” in porphyrins, 119-120
- Alcohol and amine oxidation, with nickel oximate catalyst, 134-145
- Alcohols, catalytic oxidation substrate, 137-139
- Alzheimer’s disease  
amyloid fibers, 180  
heme A biosynthesis, 51
- Amines, as oxidation substrate, 137
- Assembly protein ISU, rate constants, 5
- B**
- B-12 cofactor, 118
- Bacillus subtilis, zinc transport in, 84
- Benziporphyrins  
metal complexes, 121  
reactivity of internal carbon, 124
- Bioinspired chemistry and aerobic oxidation, 134-136  
*See also* Nickel oximate catalysts
- Bioinspired nanoassemblies, 167-181
- Bleomycin  
action in DNA, 66  
metal complexes, 67-73  
structure, 65*f*
- C**
- C–H bond, role in porphyrin–metal bonds, 119-120
- Cadmium binding in metal–peptide assemblies, 169-171
- Carbahemiporphyrazines, 121-122
- Carbaporphyrinoids, metal chemistry, 115-128
- Carbaporphyrins  
acidity of internal carbon, 119-123  
electronic effects of organometallic binding, 125-127  
reactivity of internal carbon, 123-125  
types, 117*f*, 118
- Carboxylates in anionic patch of frataxin, 8-9
- Chlorins, 118
- Chlorophylls, 118
- Co(III) bleomycin “green”  
DNA binding sites, 74-76  
fluorescence intercalator  
displacement analysis, 68-69  
nucleotide sequences cleaved by, 68  
structure, 75-76
- Cobalamin cofactor, 118
- Cobalt in active site of nitrile hydratase, 100
- “Convexity principle,” 152, 160
- Copper  
binding, in copper–metal complex, 171-172  
chaperones Cox19, Cox23, 54  
homeostasis and HAP4, 56  
homeostasis and Sco1/Sco2, 53-55  
insertion into cytochrome *c* oxidase, 53-54  
metallochaperones, mercury binding site, 196  
role in heme A biosynthesis, 41-44  
transport in yeast, 83
- Copper- and nickel-derived metalloproteins in DNA, recognition, 219-233
- Copper–peptide complexes  
inverted Marcus behavior, 172-175  
photoinduced electron transfer, 173

- Cox subunits of cytochrome *c* oxidase, 50-51, 57
- Cox1 translation and MSS51, 56
- Cox10/Cox15 and heme insertion, 51-53
- Cox11, in copper homeostasis, 55
- Cox17, in copper homeostasis, 54
- Cox19, Cox23 as copper chaperones, 54
- Cu(II) bleomycin, DNA binding sites, 70-73
- Cytochrome *c* oxidase
  - accessory proteins, 51*f*
  - assembly, 49, 56-58
  - Cox subunits, 57
  - crystal structure, 48*f*, 49-50
  - heme cofactors, 32-34, 53
  - insertion of copper, 53
  - role in cells, 32-33, 48
  - Surf1/Shy1, 56-57
  - zinc ion, 50

## D

- Density functional theory
  - calculations
    - for molybdenum and tungsten complexes, 160
    - redox potential, 154
    - vs. experimental procedures, 161
- DNA
  - cleavage by metallopeptides, 221-223
  - crystallization strategy for
    - bleomycin analysis, 64, 66
  - fiber, EPR analyses of copper metallopeptides, 232-233
  - recognition models, 219-233
  - site selectivities, fluorescent intercalator displacement assay, 64
  - targeted metal complexes, new analytic approaches, 63-78

- DNA binding ligands
  - and analysis of stressed *E. coli*, 86-88
  - and zinc-responsive proteins and genes, 86-89
  - fluorescent intercalator
    - displacement assay, 66-73
  - host-guest crystallization analysis, 64

## E

- E. coli*, zinc homeostasis, 81-92
- Electron paramagnetic resonance
  - spectroscopy of nickel complexes, 139-141
- Electron transfer in biological systems, 151-164
- Emission lifetime of copper-metal complexes, 172-173
- Extended Hückel molecular orbital calculations, 157-163

## F

- Fe(III) bleomycin, nucleotide sequences cleaved by, 70, 71*f*
- Fluorescence intercalator
  - displacement assay, 66-73
- Frataxin
  - and IscS/NifS donor proteins, 2
  - and yeast ferrochelatase, binding affinity, 20-22
  - as iron chaperone or delivery protein, 7-11
  - crystal structure, 7*f*
  - delivery of iron to ISU, 10-11
  - effects of deficiency, 18
  - Helix-1 metal binding site, 25-26
  - model of iron delivery, 18-20
  - role for carboxylates in the anionic patch, 8-9

role of iron in binding to ISU, 9  
structure, 18-19

*See also* Heme

Free energy of metal-peptide  
complexes, 174*f*, 175

Friedrich's ataxia, 18

## H

H<sub>3</sub> tris(1-propan-2-onyl  
oximate)amine, 136-137

HAP4, role in copper homeostasis, 56

$\alpha$ -Helical coiled coil, 168-169

Hem15

binding affinity to Yfh1, 20-22

Heme

activator protein 1, role in heme A  
biosynthesis, 43-44

and cytochrome *c* oxidase, 32-34

assembly, interaction between  
frataxin and ferrochelatase, 17-29

biosynthesis in yeast, model of iron  
delivery, 19*f*

biosynthesis, under varying growth  
conditions, 35-36

role of intracellular levels in heme  
A biosynthesis, 43-44

synthesis, coordination structure  
during metal transfer, 21-24

Yfh1 iron-binding sites, 19-29

Heme A

biosynthesis and regulation, 31-44

incorporation into cytochrome *c*  
oxidase, 53

regulation of biosynthetic pathway,  
40-44

role of copper in biosynthesis, 41-44

synthesis from heme O, 33-38

Heme A synthase, 33-38

Heme B, 32, 34*f*

Heme O

oxidation to heme A, 33-38

transport to heme A synthase, 36-40

Heme O synthase, 33-38

Heme O synthase/heme A synthase,  
and heme A insertion, 51-53

Hemiporphyrazines, 118

Host-guest DNA crystallization

crystal lattice, 74*f*

general description and advantages,  
73-74

Hückel molecular orbital theory,

calculation of redox potential, 154

Hydrogen peroxide, from nickel-

catalyzed reactions, 138

## I

Inverted Marcus behavior in copper-  
peptide complexes, 172-175

Iron

complexes in nitrile hydratase, 100,  
105-106

complexes with porphyrin, 120

coordination structure during  
transfer, 21-24

delivery to ferrochelatase, 18-20

delivery to ISU, 10-11

extended x-ray absorption fine  
structure analysis, 24-25

in formation of Yfh1/Hem15  
complex, 20-22

frataxin as iron chaperone, 7-11

length of bond to ligand, 23-24

physiologic roles in humans, 18

*See also* Heme

Iron molybdenum cofactor, 152,  
153

Iron sulfur clusters

assembly model, 13-14

assembly protein ISU, 5-7

biosynthesis, 3-14

oxygenates, 108-109, 110

reconstitution, 10-11

ISU assembly protein

rate constants, 5

structural characterization, 6-7

sulfide donors to, 11-12

## L

- Leigh syndrome and Surf1 protein, 55-56
- Ligand-bridged binuclear complexes, 152
- Ligand design for nickel oximate catalysts, 135-136
- Luminescent coiled coil in copper-peptide complex, 171-172

## M

- Manganese
  - complexes with porphyrin, 119-121
  - electronic effects of organometallic binding, 126-127
- Mercuric ion reductase, 195
- Mercury
  - and *de novo* protein design strategy, 183-196
  - binding in biological systems, 195-196
  - toxicity, 184
- Mercury-peptide complexes
  - perturbed angular correlation spectroscopy, 192-193
  - ultraviolet absorption spectrum, 190
- Mercury-substituted rubredoxin, 195
- Mercury-sulfhydryl complexes, 184-189
- Mercury-sulfur bond length, 190-191, 195
- Mercury thiolate complexes
  - nuclear magnetic resonance spectroscopy, 191-193
  - preparation, 185-189
  - spectroscopic characterization, 189-194
  - X-ray absorption spectroscopy, 190-191
- Metal complexes, DNA-targeted, 63-78

## Metalloenzymes

- biological parallel, 136-137
- and nickel-tris(1-propan-2-onyl oximate)amine reactions, 138-139

## Metallopeptides

- diastereoisomers, 231-233
- free energy, 174*f*, 175
- interaction with oligonucleotides, 224
- in molecular recognition, 219-223

## Metal-mediated peptide assembly, 167-181

- Metal-metal bond cleavage, 151-164
- Metal-peptide assemblies, large-scale
  - fibril and sphere formation, 178-180
  - linear chain formation, 176-177
- Metal-peptide assemblies, small-scale
  - cadmium binding, 169-171
  - inverted Marcus behavior, 172-175
  - luminescence produced with copper binding, 171-172

## Methanol

- as catalytic oxidation substrate, 137-138, 141
- oxidation by tris(1-propan-2-onyl oximate)amine, 146, 147*f*

## Molecular dynamics simulations, of DNA-metallopeptide interactions, 224-230

- Molybdenum and tungsten complexes
  - dimers, 151-164
  - optimization, 158-159
  - thermodynamic properties, 159

Molybdenum in *N*-confused porphyrins, 126

## Molybdoenzyme model system, 199-215

## MSS51 and Cox1 translation, 45

## N

- $N_2S_3X$ -Fe models of nitrile hydratase, 99-110

- Nanoassemblies, peptide-based, 167-181
- N*-confused porphyrins  
 complexes with direct metal–carbon bonds, 121, 123*f*  
 metal binding, 119-120  
 reactivity of internal carbon, 124
- Nickel  
 complexes, oxygen reactivity, 135-137  
 metallopeptide–DNA interactions, molecular dynamics simulations, 226-230  
 metalloptides, in DNA recognition, 219-233  
 reactions catalyzed by, 138
- Nickel oximate catalysts  
 ligand design criteria, 135-136  
 parallel to biological alcohol/amine oxidation, 134-145
- Nickel–tris(1-propan-2-onyl oximate)amine  
 air oxidation, 140-141  
 oxygen reaction in methanol, 139
- Nitrile hydratase  
 active site, 100, 101*f*  
 bond covalencies, 109  
 charge transfer band, 106-107  
 conversion of nitriles to amides, 100, 101*f*  
 electronic modeling, 104-107  
 functional modeling, 107-111  
 mimics, 110-111  
 reduction potential, 106-107  
 structural modeling, 103-104  
 sulfur ligand models, 104-106  
 thiolate complexes, oxygen sensitivity, 108-109
- Nitriles, environmental sources, 100-101
- Nitrogen fixation, 152, 153
- Nuclear magnetic resonance  
 of mercury thiolate complexes, 191-193  
 of metallopeptide–oligonucleotide interactions, 224
- Nucleophiles, in porphyrins, 125
- O**
- Organometallic bonding in carbaporphyrins, 125-127
- Organometallic interactions of porphyrins, 118-119
- Oxo-molybdenum complexes, and oxygen atom transfer, 200-215
- Oxotransferase reactivity, and singular value decomposition analysis, 199-215
- Oxygen  
 atom transfer reactions, in oxomolybdenum complexes, 200-215  
 chemistry, in nickel-dependent enzymes, 135  
 in heme synthase activity, 35-37  
 reactivity of nickel complexes, 136-137  
 sensitivity, of nitrile hydratase thiolate complexes, 108-109  
 source in heme A, 35
- P**
- Peptides  
 and nucleic acids, self-assembly, 168  
 $\alpha$ -helical coiled-coil, 168-169  
 metal mediated, 167-181  
 self-association with mercury, 187
- Perturbed angular correlation spectroscopy, of mercury–peptide complexes, 192-193
- pH-dependent aggregation of mercury with peptides, 186-188
- Phosphido- and sulfido-bridged dimers, 151-164

Phosphido-bridged complexes,  
electrochemical parameters, 156-  
157

Phosphines as model substrates,  
200

Photoinduced electron transfer, 173

Photoregulation of nitrile hydratase,  
100, 108

Porphyrins

complexes with manganese, 119-  
121

in biological systems, 115-118,  
128

nucleophiles at internal carbon  
position, 125

role of C-H bond in metal bonding,  
119-120

structure, 116-117

*See also N-confused porphyrins*

Proteins

cytochrome *c* oxidase accessory  
proteins, 51*t*

*de novo* protein design, and  
mercury, 183-196

design, and mercury complexes,  
187

heme activator protein 1, 43-44

IscS/NifS donor proteins, 2

ISU assembly protein, 5-7

Surf1, 55-56

thiolate-containing proteins, and  
mercury, 183-196

zinc-responsive proteins, 86-89

*See also* Frataxin, Sco proteins,

## R

Redox active molecular orbitals, in  
tungsten complexes, 157-158

Redox chemistry in biological systems

computational studies, 157-163

electrochemical studies, 154-157

“redox hybrid,” parallel to  
metalloenzymes, 136-137

Redox potential

of sulfido vs. phosphido-bridged  
complexes, 161, 163*t*

in two-electron systems, 154-155

RNA cleavage by metallopeptides,  
221-223

Ruthenium, free energy in peptide  
complexes, 174*f*, 175

## S

Sco proteins

differences between human and  
yeast, 55

Sco1/Sco2 and copper homeostasis,  
53-55

Self-assembly

of peptides and nucleic acids, 168

of peptides, three-dimensional  
arrays, 176-179

Shy1

homolog, and insertion of heme  $a_3$ ,  
53

yeast homolog of Surf1, 56-57

Singular value decomposition analysis

mathematical description, 202-204

oxotransferase activity, 199-215

results of spectral data analysis,  
204-213

use in biochemical systems, 200

Sulfide donors to ISU protein, 11-12

Sulfido- and phosphido-bridged  
dimers, 151-164

Sulfido-bridged complexes, electro-  
chemical parameters, 156-157

Sulfite oxidase as model enzyme, 201

Sulfur oxygenation, roles in nitrile  
hydratase, 100-101

Surf1 protein and Leigh syndrome,  
55-56

Surf1/Shy1 in cytochrome *c* oxidase  
assembly, 56-57

Surf homolog, and insertion of heme  
 $a_3$ , 53

## T

## TPEN

- a “zinc-responsive chelator,” 86
- regulation of genes by, 86-89

## TRI peptide family, 185-189

Tris(1-propan-2-onyl oximate)amine  
and metals other than nickel and  
zinc, 141-144

- choice of synthetic system for nickel  
complex, 134-136

methanol oxidation by, 146, 147*f*

nickel complex, 134-145

reactions parallel to metalloenzyme  
reactions, 138-139

## Tungsten and molybdenum

complexes, optimization, 158-159

dimers, 151-164

redox active molecular orbitals,  
157-158

## U

Ultraviolet absorption spectrum, of  
mercury-peptide complexes, 190

## V

Voltammograms in two-electron  
systems, 154-155

## X

## X-ray absorption

- near-edge spectral analysis, in  
heme assembly studies, 21

spectroscopy, of mercury thiolate  
complexes, 190-191

X-ray crystal structure of nitrile  
hydratase, 103

## Y

Yeast and discovery of  
metallochaperones, 85

## Yfh1

binding affinity to Hem15, 20-  
22

helix-1 residues, 25-26

iron binding sites on heme, 19-  
29

## Yfh1/Hem15 complex

extended X-ray absorption fine  
structure analysis, 24-25

role of iron, 20-22

## Z

## Zinc

homeostasis, as potential drug  
target, 72

homeostasis in *E. coli*, 81-92

intracellular mobilization, “pulse”  
mechanism, 84-85

ion, in cytochrome *c* oxidase,  
50

metallochaperones in yeast, 85-  
86

trafficking genes, 86-87

transport/mobilization in bacteria,  
82-85, 89-91

ZntR-regulated genes, 86-87, 89

Zur-regulated genes, 86-89