UNIVERSITY OF CALIFORNIA, SAN DIEGO

Expanding the Utility of Proteins as Platforms for Coordination Chemistry

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Chemistry

by

Robert John Radford

Committee in charge:

Professor F. Akif Tezcan, Chair
Professor Clifford P. Kubiak
Professor Arnold L. Rheingold
Professor Douglas E. Smith
Professor Yitzhak Tor

2011

The Dissertation of Robert John Radford is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2011

Dedication

This thesis is dedicated to my amazing family and friends whose encouragement and support, for better or worse, have made me who I am today. EAERMP, IAY.

Table of Contents

List of Abbreviations

AUC: analytical ultracentrifugation

bis-His: bis-histidine clamp

bZIP: basic leucine zipper motif

CD: circular dichroism

CPI: crystal packing interactions

Cu_2:MBPC-1_2: Cu-mediated MBPC-1 dimer

Cyt b_{562}: cytochrome b_{562}

Cyt cb_{562}: cytochrome cb_{562}

DFT: density functional theory

DMF: dimethylformamide

DMSO: dimethyl sulfoxide

DTT: dithiothreitol

ΔG_f: equilibrium free energy of folding

EDTA: ethylene diamine tetraacetic acid

EGTA: ethylene glycol tetraacetic acid

FPLC: fast protein liquid chromatography

GuHCl: Guanidine hydrochloride

HCM: hybrid coordination motif

HPhen: histidine:1,-10-phenanthroline HCM

HQuin: histidine:8-hydroxyquinoline HCM

IA: iodoacetamide

ICP-OES: inductively coupled plasma optical emission spectroscopy

IMAC: immobilized metal affinity column

IR: Infrared

K_d: dissociation constant

MALDI: matrix-assisted laser-desorption ionization

MBP: metal binding protein

MBPC: metal binding protein complex

MDPSA: metal-dependent protein self-assembly

MOPS: 3-(N-morpholino)propanesulfonic acid

Ni_2:MBPC-1_3: Ni-mediated MBPC-1 trimer

PAGE: poly-acrylamide gel electrophoresis

Pd_2:MBPC-1_2: Pd-mediated MBPC-1 dimer

PEG: polyethylene glycol

Phen: 1,10-Phenanthroline

PPIs : protein-protein interactions

Quin: 8-Hydroxyquinoline

Tris: tris(hydroxymethyl)aminomethane

SASA: solvent accessible surface area

SE: sedimentation equilibrium

SV: sedimentation velocity

UV-vis: Ultraviolet-visible

List of Figures

List of Tables

Acknowledgements

It is a common misconception that scientists work in isolation. In point of fact, a successful scientific undertaking requires collaboration, support and communication between many parties with varying interests, expectations and motivations. From this perspective, I have been extremely fortunate to belong to a community here at UCSD, which typifies the type of collaborative, interdisciplinary and supportive environment in which you would want to conduct your Ph.D.

My advisor Akif Tezcan, has been exceptionally encouraging of me throughout my Ph.D. Akif was always supportive, allowing me to grow and explore, while still providing guidance and grounding. He taught me how to think, act, and even write like a scientist (though the latter is still a work in progress).

The Tezcan Lab members, both past and present, always made coming to work an entertaining experience. You find out a lot about a person when you spend 60 hrs/week together in a small windowless lab, and I have been fortunate to have some "forced socialization" with an outstanding group of people.

My projects literally would not have been possible without the help of many of the department's faculty, students and post-docs including: Arnold Rheingold, Curtis Moore, Cliff Kubiak, John Goeltz, Eric Benson, Jon Smieja, Elizabeth Komives, Mela Mulvihill, Nick, Treuheit, Brian Fuglestad, Seth Cohen, Jody Major Jourden, Matthieu Rouffet, Arpita Agrawal, Sergio Garibay, J. Andrew McCammon, Morgan Lawrenz, Joshua Figueroa, Treffly Ditri and Yongxuan Su.

Lastly, I need to thank two phenomenal undergraduates, Phuong Nguyen and Kang Du. These two made it fun to come to lab each day. They worked harder and smarter than I ever could have as an undergraduate and they will be exceptionally successful in whatever endeavor they decide to undertake.

Chapters 1 and 7 are reproduced, in part, with permission from: Radford, R. J., Brodin, J. D., Salgado, E. N., Tezcan, F. A. 2010. *Coord. Chem. Rev.* DOI:10.1016/j.ccr.2010.10.010. Copyright 2011, Elsevier B.V.

Chapter 2 is reproduced, in part, with permission from: R.J. Radford, F.A. Tezcan, *J. Am. Chem. Soc.*, 131 (2009) 9136-9137. Copyright 2009, The American Chemical Society.

Chapter 3 is reproduced, in part, with permission from: R.J. Radford, M. Lawrenz, P.C. Nguyen, J.A. McCammon, F.A. Tezcan, *Chem. Commun*., 47 (2011) 313-315. Copyright 2011, Royal Society of Chemistry

Chapter 4 is reproduced, in part, with permission from: R.J. Radford, P.C. Nguyen, T.B. Ditri, J.S. Figueroa, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 4362-4369. Copyright 2010 American Chemical Society

Chapter 5 is reproduced, in part, with permission from: R.J. Radford, P.C. Nguyen, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 7106-7115. Copyright 2010 American Chemical Society

The dissertation author is primary author on all reprinted materials. He would also like to acknowledge the NIH Heme and Blood training program for funding part of this work.

Vita

EDUCATION

2011 Doctor of Philosophy in Chemistry, University of California, San Diego

2009 Master of Science in Chemistry, University of California, San Diego

2006 Bachelor of Science in Biochemistry, University of California, Santa Barbara

AWARDS AND HONORS

2007-2010 NIH Heme and Blood Training Program

2008 Barbara and Paul Saltman award for excellence in teaching at the University of California, San Diego

2005 NRW International Graduate School Fellowship

2002-2006 Paul Orfalea Family Foundation Scholarship

PUBLICATIONS

1) Robert J. Radford, Jeffery D. Brodin, Eric N. Salgado and F. Akif Tezcan, "-Expanding the Utility of Proteins as Platforms for Coordination Chemistry," *Coordination Chemistry Reviews,* 2010, doi:10.1016/j.ccr.2010.10.010.

2) Robert J. Radford, Morgan Lawrenz, Phuong C. Nguyen, J. Andrew McCammon and F. Akif Tezcan, "Porous Protein Frameworks with Unsaturated Metal Centers in Sterically Encumbered Coordination Sites," *Chem. Commun.,* 2011, *47*, 313-315.

3) Robert J. Radford, Phuong C. Nguyen and F. Akif Tezcan, "Modular and Versatile Hybrid Coordination Motifs on α-Helical Protein Surfaces," *Inorg. Chem.,* 2010, *49*, 7106-7115

4) Robert J. Radford, Phuong C. Nguyen, Treffly B. Ditri, Joshua S. Figueroa and F. Akif Tezcan, "Controlled Protein Dimerization Through Hybrid Coordination Motifs," *Inorg. Chem.* 2010, *49*, 4362-4369

5) Eric N. Salgado, Robert J. Radford, and F. Akif Tezcan, "Protein Self-Assembly Directed by Coordination Chemistry," *Acc. Chem. Res.,* 2010, *43*, 661-672

6) Robert J. Radford, Mark D. Lim, Roberto Santana da Silva and Peter C. Ford, "Photochemical Cleavage of Nitrate Ion Coordinated to a Cr(III) Porphyrin," *J. Coord. Chem.,* 2010, *63*, 2743 — 2749.

7) Robert J, Radford and F. Akif Tezcan, "A Superprotein Triangle Driven by Nickel(II) Coordination: Exploiting Non-Natural Metal Ligands in Protein Self-Assembly," *J. Am. Chem. Soc*., 2009, *131*, 9136-9137

8) Sarah A. Cummings, Robert Radford, Gerhard Erker, Gerald Kehr, Roland Fröhlich "Formation of a Dynamic η^2-(O,N)-Hydroxylaminato Zirconocene Complex by Nitrosoarene Insertion into a Zr-C σ-Bond," *Organometallics*, 2006, *25*, 839-842

FIELDS OF STUDY

Major Field: Chemistry and Biochemistry

Sub-discipline Bioinorganic Chemistry, Inorganic Chemistry, Structural Biology and Biophysics.

ABSTRACT OF THE DISSERTATION

Expanding the Utility of Proteins as Platforms for Coordination Chemistry

by

Robert John Radford

Doctor of Philosophy in Chemistry

University of California, San Diego, 2011

Professor F. Akif Tezcan, Chair

Whether for constructing advanced materials and complex biological devices, or for building sophisticated coordination complexes with diverse metal-based functions, proteins are Nature's favorite building blocks. Yet, our ability to control the assembly of proteins or to use them as ligand platforms for inorganic chemistry has been somewhat limited. The goal of this thesis was to exploit the utility of a protein scaffold in both regards. First, by considering proteins as "simple" ligand platforms and controlling the metal coordination chemistry on their surfaces, we show how their self-assembly can be readily dictated by metal binding. Next, we show how having this measure of control can lead to the site-specifically localization of functional metal complexes on the surface of a proteins and peptides. While on one hand our studies have pointed out the challenges of using proteins as ligands, they have also revealed how extensive and chemically-rich protein surfaces can be.

Chapter 1: Metal Directed Protein Self-Assembly (MDPSA)

1.1 Introduction

Proteins are Nature's most versatile ligands. They feature a number of metal binding functionalities, and have the unique ability to form three-dimensional platforms that provide unparalleled control over both the coordination environment and reactivity of their associated metal center. It is therefore no coincidence that a large fraction (~35% and growing) of known proteins are metalloproteins [1]. Metalloproteins readily perform impressive, and highly coveted, chemical transformations such as N_2 fixation, H_2O oxidation, CO_2 reduction, and selective functionalization of organic substrates [2]. Moreover, metalloporteins execute these reactions under the ambient temperatures, pressures and neutral pHs typically associated with physiological conditions.

Understandably, there is a tremendous, and ongoing, effort dedicated to understanding the structure and function of metalloproteins [2]. Appropriately, much of this attention, thus far, has been directed toward the study of the reactive metal centers within the protein scaffold. Since metal-centered activities are generally established once the metal ions are firmly placed within the protein framework, a traditional avenue of study has been dedicated toward characterizing, mimicking and modifying metal coordination sites within protein interiors [3, 4]. Yet, recent advances in our understanding of cellular inorganic chemistry are indicating that the interactions of metals with protein surfaces are just as significant [5]. Metals, or their complexes (e.g., metallodrugs), under physiological conditions are in constant contact with protein

surfaces [6]. When properly controlled, metal protein-surface interactions mediate active transport of essential metal ions to their proper cellular destinations [7]. However, when dysregulated, the uncontrolled interaction of metal ions with protein surfaces can be exceptionally harmful, leading to protein aggregation [8], or the generation of reactive oxygen species [9]. Outside the cellular realm, metal-protein surface interactions have revolutionized protein biochemistry in the form the immobilized metal ion affinity chromatography (IMAC) [10], which allows for the facile purification of almost any protein of interest. In addition, metal complexes have served as invaluable spectroscopic/functional probes and catalytic sites [11, 12].

Clearly, there is a need for a deeper understanding of the inorganic chemistry of metal ions on protein surfaces. Herein, it will be shown that the surface of protein molecules can be appropriately engineered to achieve the same control and affinity as their small-molecule counterparts; allowing them to be literally used as metal ligands in a traditional synthetic inorganic chemical sense. This approach has not only enabled the metal-guided engineering of discrete protein architecture, but has also provided a path to construct novel metal coordination sites within protein interfaces. This work draws considerable inspiration from earlier works on fundamental inorganic coordination chemistry, supramolecular chemistry, and protein design/engineering.

1.2 Metal Directed Protein Self-Assembly (MDPSA)

Our primary motivation to study the inorganic chemistry of protein surfaces stems from our desire to use metal coordination to control protein-protein interactions (PPIs)

and protein self-assembly (PSA), which lie at the heart of nearly all cellular process and complex protein architectures. The design and engineering of PPIs are plagued by the fact that they constitute extensive molecular surfaces made up of numerous weak interactions [13, 14]. The additive nature of these weak bonds enable a given protein the ability form a tight and specific interaction with its appropriate partner within the context of the crowded cellular environment. However, these same extensive surfaces make the engineering of PPIs exceptionally difficult, often requiring a rigorous redesign of the protein interface through a combination of iterative computational and biological "evolution" (Figure 1.1a). While, these methods have had some resounding success [15, 16], their approaches are inherently labor-intensive and lack the ability to be readily generalizable. As an alternative, it was theorized that control over proteins supramolecular architecture could be achieved by the appropriate placement of a few select metal-binding residues on the proteins surface (Figure 1.1b), thereby obviating the need for extensive redesign of the proteins surface.

Coordination chemistry provides numerous advantages that make them ideally suited to control and direct PPIs/PSA. First, metal-ligand bonds are much stronger (10-50 kcal/mol) than weaker hydrogen, electrostatic, or van der Waal bonding interactions (2-10 kca/mol) that govern most natural PPIs. Yet, they simultaneously provide a reversibility, which allows for the formation of thermodynamic (rather than kinetic) products. Moreover, the inherent sterochemical preference of metal ions provides directionality to self-assembled systems, a requirement when constructing discrete architectures. Lastly, decades of research into coordination complexes have supplied a rich library of ligands specifically designed to tune and control the reactivity of metal

ions, as well as adding the potential for the introduction of non-natural reactivity. These advantages have been extensively utilizes in small-molecule based supramolecular systems leading to the ability to design molecules with: discrete geometric shapes [17], stimulus responsive assemblies [18], porous frameworks [19] that display heterogeneous chemical functionalities [20], as well a assemblies with chemical environments designed to promote reactivity [21].

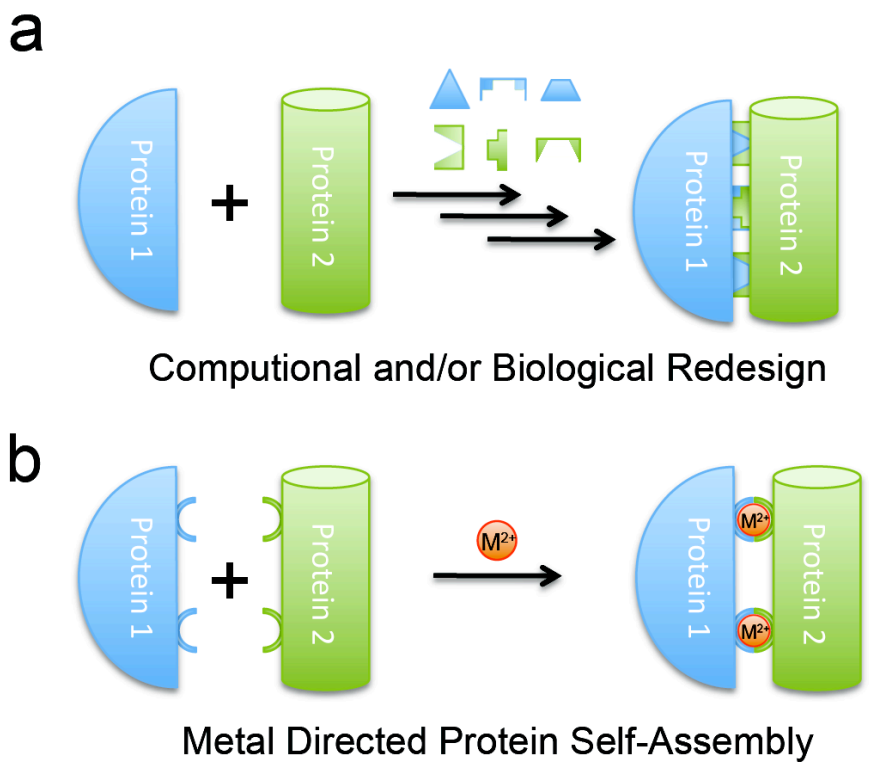


Figure 1.1. (a) Cartoon illustrating the intensive process of engineering novel PPIs. Complementarities are formed via an iterative design process traditionally involving either *in silico* interface redesign of a known protein complex and/or "biological" redesign via directed evolution. (b) A cartoon schematic illustrating the more facile metal directed protein self-assembly, whereby appropriately placed metal binding residues can drive protein self-assembly upon the addition of metal.

1.3 Choosing A Model System

In choosing a model system to demonstrate the feasibility of MDPSA, candidate proteins were subjected to several criteria: 1) the model protein had to be easily expressed in large quantities, 2) the protein needed to be thermodynamically stable and readily amenable to modification (i.e. site-directed mutagenesis and/or bioconjugation), 3) the protein needed to remain monomeric even at millimolar concentrations and 4) it should be small with a relatively featureless tertiary structure. For these reasons, the hemoprotein cytochrome cb_{562} (cyt cb_{562}) was chosen (Figure 1.2a). This small (106 residue) predominately α-helical protein that is engineered to have a "c-type" linkage of the b-type heme, which covalently attaches the heme group to the protein scaffold [22]. This modification results is an increase in the thermal and chemically stability of the protein over the native cyt b_{562}. Moreover, cyt cb_{562} can be readily expressed on the gram scale and is non self-associating up to millimolar concentrations. Cyt cb_{562} is also roughly cylindrical, readily crystallizes and has no free cysteine residues.

1.4 Metal Binding Protein Complex-1 (MBPC-1)

An inherent difficulty of MDPSA stems from the chemical heterogeneity of protein surfaces, which is replete with carboxylates, amines, imidazoles, and thiol groups (Figure 1.2a), making the site-selective localization of metals very challenging. Taking a cue from small-molecule metal ligands, bidentate motifs were engineered on the protein surface, thereby providing a metal-binding site that could readily outcompete the

monodentate functionalities listed above. The α-helix is a particularly well-suited architecture for the installation of metal-chelating motifs, owing to the regular spacing of its amino acid constituents and its prevalence in the protein kingdom. In fact, nature frequently uses α-helices as scaffolds for metal chelation: amino acid residues (His, Glu, Asp) placed in *i*/*i*+*3* and *i*/*i*+*4* patterns (thus pointing in the same direction) are quite regularly used to construct bidentate metal binding sites, such as those in Zn-finger domains, and di-iron and di-copper centers to name a few [23]. Inspired by these natural examples, chemists have employed *i*/*i*+*3* and *i*/*i*+*4* bidentate motifs consisting of natural or unnatural metal ligands, with iminodiacetic acid (IDA) exemplifying the latter, to stabilize α-helical protein folds [24, 25], build *de novo* metalloproteins with stable mono- and dinuclear metal centers [23, 26], build substrate-selective metallopeptide catalysts [27], and to facilitate purification by IMAC [28]. The imidazole sidechain of His is an ideal component for a surface chelating motif, as its "borderline soft" imide nitrogens possess a high affinity for most transition metals (relative to the "hard" carboxylates of the more common Asp and Glu residues); it also does not suffer from oxidation and covalent dimerization common with Cys residues. When installed on an α-helical platform in an *i*/*i*+4 pattern, the bis-His motif provides metal dissociation constants that are in the low μM regime for late first-row transition metals (Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}) [25, 29]. In initial studies by Salgado, Tezcan *et al.*, two such bis-His motifs (His59/His63 and His73/His77) were incorporated near the ends of Helix3 of cyt cb_{562} to make the construct MBPC-1 (Figure 1.2b), with the idea that metal coordination would lead to the oligomerization of this natively monomeric protein [30]. Indeed, the addition of equimolar Ni^{2+}, Cu^{2+} and Zn^{2+} to MBPC-1 results in the formation of discrete

superprotein architectures that are commensurate with the stereochemical preferences of these metal ions (Figures 1.3b and 1.3c): octahedral Ni^{2+} coordination promotes a C_3-symmetric trimer (Ni_2:MBPC-1_3) with the Ni's coordinated to three bis-His motifs donated by all three protomers; tetragonal Cu^{2+} produces a C_2-symmetric dimer (Cu_2:MBPC-1_2) with two bis-His motifs forming the equatorial coordination plane; and finally, tetrahedral Zn^{2+} coordination yields a D_2-symmetric tetramer (Zn_4:MBPC-1_4), where the Zn ligand set consists of a bis-His motif (H73/H77) from one protomer, a single His (H63) from a second, and an Asp (D74) from a third [30, 31]. Ni^{2+}, Cu^{2+} and Zn^{2+} are all exchange labile ions that allow the metal-directed protein self-assembly to proceed under thermodynamic control. In the presence of Pd^{2+}, multiple oligomeric products are formed, as would be expected from the relative substitution inertness of this ion. Nevertheless, the predominant product is a dimer (Pd_2:MBPC-1_2) with an identical structure to Cu_2:MBPC-1_2, driven by the square-planar coordination of Pd^{2+} (Figure 1.2c) [32].

These findings demonstrate that, in the absence of specific protein-protein interactions, the protein self-assembly is largely dictated by the stereochemical preference of the driving metal ion. Nevertheless, the coordination environment and the resulting supramolecular geometry of Zn_4:MBPC-1_4, for example, was in contrast to the expectations that a dimeric structure with Zn:His_4 coordination (one bis-His from each monomer) would be produced. Analytical ultracentrifugation studies revealed that a dimeric species indeed is formed at low Zn^{2+} and MBPC-1 concentrations, but it is subsequently replaced by the tetrameric architecture as the concentrations are increased [30].

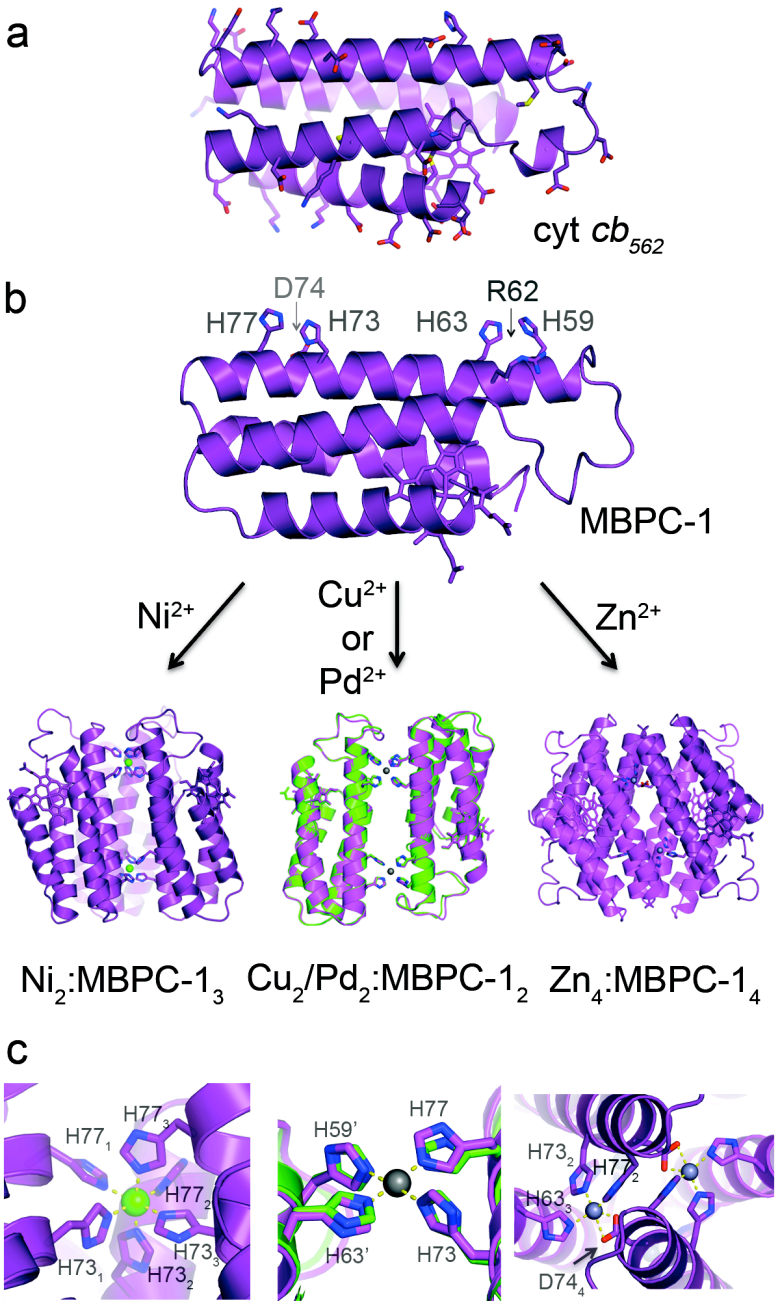


Figure 1.2. (a) Cartoon representation of the four-helix-bundle hemeprotein, cytochrome cb_{562} (cyt cb_{562}), with amino acid side chains capable of coordinating metal ions shown as sticks. (b) The cyt cb_{562} variant, MBPC-1, which can self-assemble into discrete structures in a metal-dependent fashion based on the stereochemical preference of the added metal ion (a C_3 trimer with Ni^{2+}, a C_2 dimer with Cu^{2+} or Pd^{2+}, a D_2 tetramer with Zn^{2+}). (c) Close-up view of the interfacial metal centers in each metal-mediated complex.

1.5 The Involvement of Secondary Interactions in Metal-Directed Protein Self-Assembly

The crisscrossed Zn_4:MBPC-1_4 architecture results in extensive surface contacts among the protomers, with a total buried surface area of nearly 5000 $Å^2$ in the complex. To probe if these contacts have a collective influence on the thermodynamics of self-assembly without introducing extensive surface mutations, Salgado *et al.* instead made a small perturbation within the metal coordination sphere. Each Zn in Zn_4:MBPC-1_4 is coordinated by an Asp74 located within the 73/77 bis-His clamp in the *i*+1 position. Ligation by Asp74's is central to the observed supramolecular architecture of Zn_4:MBPC-1_4, in that they crosslink the MBPC-1 monomers at the Helix3 C-termini to yield the V-shaped dimers. If non-covalent interactions between protein monomers had negligible effect and metal coordination was the sole determinant of protein self-assembly, then the whole oligomeric assembly could be inverted by simply moving the coordinating Asp residue from within the C-terminal 73/77 bis-His clamp into the N-terminal 59/63 bis-His clamp motif at the *i*+3 position. Towards this end, MBPC-2, the R62D/ D74A variant of MBPC-1 was engineered [33].

MBPC-2 forms a tetramer upon binding one molar equivalent of Zn according to analytical ultracentrifugation analyses. The crystal structure of Zn_4:MBPC-2_4 reveals a D_2-symmetrical architecture, which indeed is the "inverse" of Zn_4:MBPC-1_4 (Figure 1.3a) [33]. Whereas the V shapes are joined at the Helix3 C-termini in the latter, they are crosslinked at the N-termini in the former. Unexpectedly, the newly engineered Asp62 is not involved in Zn binding. Instead, each Zn ion in the assembly is ligated by the 73/77

bis-His motif from one protomer, His59 from a second, and His63 from a third, again yielding a tetrahedral Zn coordination geometry. In this arrangement, coordination by His59 and His63 from two protomers–instead of the planned Asp62 and His63 coordination–stabilize the V's by splaying apart to bind two Zn ions and, in turn, joining the Helix3 N-termini (Figure 1.3b).

Similarities between Zn_4:MBPC-1_4 and Zn_4:MBPC-2_4 structures suggest that tetrahedral Zn coordination again nucleates self-assembly and enforces D_2 supramolecular symmetry. Nevertheless, it is clear that metal coordination is not the only determinant, as MBPC-1 and MBPC-2 form distinctly different tetramers despite differing in only two residues situated far from the coordination sites. To probe the contribution of non-covalent interactions in metal-directed self-assembly, Salgado *et al.* took a closer look at the interprotomeric interfaces in Zn_4:MBPC-1_4 and Zn_4:MBPC-2_4. This inspection revealed a set of salt-bridging interactions that stood out in each complex: between Arg34 and Asp66 in Zn_4:MBPC-1_4, and between Arg34 and Asp62 in Zn_4:MBPC-2_4 (Figure 1.3c). Significantly, when these interactions are abolished through the Arg34Ala (MBPC-1) and Arg34Asp (MBPC-2) mutations, the tetrameric assemblies are replaced by heterogeneous ensembles that contain higher order aggregates [33]. These findings demonstrate that Arg34 interactions clearly are the guiding factor for the metal-induced oligomerization of MBPC-1 and 2.

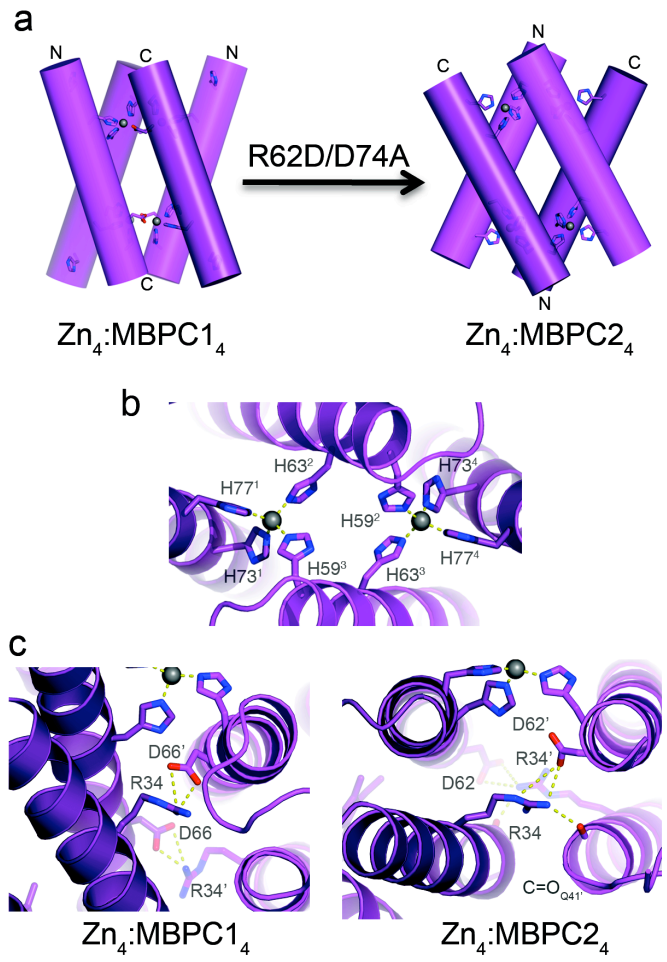


Figure 1.3. (a) The conversion between the two Zn^{2+} mediated tetramers Zn_4:MBPC-1_4 and Zn_4:MBPC-2_4, shown in cylindrical cartoon representation. (b) Close-up view of interfacial metal centers in Zn_4:MBPC-2. (c) Key secondary interactions in the interfaces of Zn_4:MBPC-1_4 and Zn_4:MBPC-2_4.

1.6 Expanding the Utility of Protein Surface Coordination Chemistry

The extent secondary interactions (i.e., H-bonds and salt bridges) contribute to the overall stability of proteins and PPIs can vary from system to system [34-37]. Yet, these geometry-dependent interactions are important in limiting the number of possible low

energy docking conformations, thus playing a major role in determining specificity [38, 39]. Our previous findings indicated that salt-bridging and H-bonding interactions can work in conjunction with metal-ligand interactions to dictate the geometric alignment of protein partners. This cooperativity leads to the population of discrete metal-directed supramolecular structures over other conformations of similar energy. Such sensitivity to secondary, non-covalent interactions makes our goal of predictably forming discrete superprotein architectures challenging.

In response to this apparent complication, I sought to expand the control of metal ions on the protein surface by introducing non-natural functionalities. More specifically, by capitalizing on the large "toolbox" of multidentate ligands, devised by synthetic chemists to control the geometry and reactivity of metal ions, an increase the ability to site-specifically localize a metal ion on the protein surface was sought. Similar work on peptide surfaces has been shown to allow for tuning of metal binding affinity over several orders on magnitude [40, 41], template the formation of tertiary structures [42, 43], and introduction non-natural enzymatic activity [44, 45]. It was hypothesized that by introducing multidentate binding motifs on the proteins surface through cysteine specific iodoacetamide functionalities, it would limit the possible numbers of metal-mediated oligomers, thereby gaining much needed control over our self-assembled systems. The results of these experiments will be discussed herein.

Chapters 1 is reproduced, in part, with permission from: Radford, R. J., Brodin, J. D., Salgado, E. N., Tezcan, F. A. 2010. *Coord. Chem. Rev.*

DOI:10.1016/j.ccr.2010.10.010. The dissertation author was the primary author of this paper.

1.7 References

[1] K.J. Waldron, N.J. Robinson, *Nature Rev. Microbiol.*, 7 (2009) 25-35.

[2] I. Bertini, H.B. Gray, E.I. Stiefel, J.S. Valentine, Biological Inorganic Chemistry, Structure & Reactivity, University Science Books, Sausalito, 2007.

[3] R.H. Holm, E.I. Solomon, *Chem. Rev.*, 104 (2004) 347-348.

[4] Y. Lu, S.M. Berry, T.D. Pfister, *Chem. Rev.*, 101 (2001) 3047-3080.

[5] A.E. Palmer, K.J. Franz, *Chem. Rev.*, 109 (2009) 4533-4535.

[6] A.V. Klein, T.W. Hambley, *Chem. Rev.*, 109 (2009) 4911-4920.

[7] A.K. Boal, A.C. Rosenzweig, *Chem. Rev.*, 109 (2009) 4760-4779.

[8] W. Maret, Y. Li, *Chem. Rev.*, 109 (2009) 4682-4707.

[9] D.G. Smith, R. Cappai, K.J. Barnham, *Biochim Biophys Acta*, 1768 (2007) 1976-1990.

[10] F.H. Arnold, B.L. Haymore, *Science*, 252 (1991) 1796-1797.

[11] J.R. Winkler, H.B. Gray, *Chem. Rev.*, 92 (1992) 369-379.

[12] S.R. Adams, R.E. Campbell, L.A. Gross, B.R. Martin, G.K. Walkup, Y. Yao, J. Llopis, R.Y. Tsien, *J. Am. Chem. Soc.*, 124 (2002) 6063-6076.

[13] B.A. Shoemaker, A.R. Panchenko, *PLos Comput. Biol.*, 3 (2007) 595-601.

[14] T. Kortemme, D. Baker, *Curr. Opin. Chem. Biol.*, 8 (2004) 91-97.

[15] I. Andre, P. Bradley, C. Wang, D. Baker, *Proc. Natl. Acad. Sci. USA*, 104 (2007) 17656-17661.

[16] J.W. Chin, A. Schepartz, *Angew Chem Int Ed*, 40 (2001) 3806-3809.

[17] S. Leininger, B. Olenyuk, P.J. Stang, *Chem. Rev.*, 100 (2000) 853-907.

[18] B.J. Holliday, C.A. Mirkin, *Angew. Chem., Int. Ed. Eng.*, 40 (2001) 2022-2043.

[19] N.W. Ockwig, O. Delgado-Friedrichs, M. O'Keeffe, O.M. Yaghi, *Acc. Chem. Res.*, 38 (2005) 176-182.

[20] K.K. Tanabe, S.M. Cohen, *Chem Soc Rev*, 40 (2011) 498-519.

[21] M.D. Pluth, R.G. Bergman, K.N. Raymond, *Science*, 316 (2007) 85-88.

[22] J. Faraone-Mennella, F.A. Tezcan, H.B. Gray, J.R. Winkler, *Biochemistry*, 45 (2006) 10504-10511.

[23] A. Lombardi, C.M. Summa, S. Geremia, L. Randaccio, V. Pavone, W.F. DeGrado, *Proc. Natl. Acad. Sci. USA*, 97 (2000) 6298-6305.

[24] F.Q. Ruan, Y.Q. Chen, P.B. Hopkins, *J. Am. Chem. Soc.*, 112 (1990) 9403-9404.

[25] M.R. Ghadiri, C. Choi, *J. Am. Chem. Soc.*, 112 (1990) 1630-1632.

[26] T.M. Handel, S.A. Williams, W.F. Degrado, *Science*, 261 (1993) 879-885.

[27] B.V. Popp, Z.T. Ball, *J. Am. Chem. Soc.*, 132 (2010) 6660-6662.

[28] R.J. Todd, R.D. Johnson, F.H. Arnold, *J. Chromatogr. A*, 662 (1994) 13-26.

[29] B.A. Krantz, T.R. Sosnick, *Nat. Struct. Biol.*, 8 (2001) 1042-1047.

[30] E.N. Salgado, J. Faraone-Mennella, F.A. Tezcan, *J. Am. Chem. Soc.*, 129 (2007) 13374-13375.

[31] E.N. Salgado, R.A. Lewis, S. Mossin, A.L. Rheingold, F.A. Tezcan, *Inorg. Chem.*, 48 (2009) 2726-2728.

[32] E. N. Salgado and F. A. Tezcan, unpublished results.

[33] E.N. Salgado, R.A. Lewis, J. Faraone-Mennella, F.A. Tezcan, *J. Am. Chem. Soc.*, 130 (2008) 6082-6084.

[34] N. Froloff, A. Windemuth, B. Honig, *Prot. Sci.*, 6 (1997) 1293-1301.

[35] S. Kumar, R. Nussinov, *Biophys. J.*, 83 (2002) 1595-1612.

[36] G.I. Makhatadze, V.V. Loladze, D.N. Ermolenko, X.F. Chen, S.T. Thomas, *J. Mol. Biol.*, 327 (2003) 1135-1148.

[37] R. Radhakrishnan, L.J. Walter, A. Hruza, P. Reichert, P.P. Trotta, T.L. Nagabhushan, M.R. Walter, *Structure*, 4 (1996) 1453-1463.

[38] Z.S. Hendsch, B. Tidor, *Prot. Sci.*, 3 (1994) 211-226.

[39] F.B. Sheinerman, R. Norel, B. Honig, *Curr. Opin. Struct. Biol.*, 10 (2000) 153-159.

[40] R.P. Cheng, S.L. Fisher, B. Imperiali, *J. Am. Chem. Soc.*, 118 (1996) 11349-11356.

[41] M.D. Shults, D.A. Pearce, B. Imperiali, *J. Am. Chem. Soc.*, 125 (2003) 10591-10597.

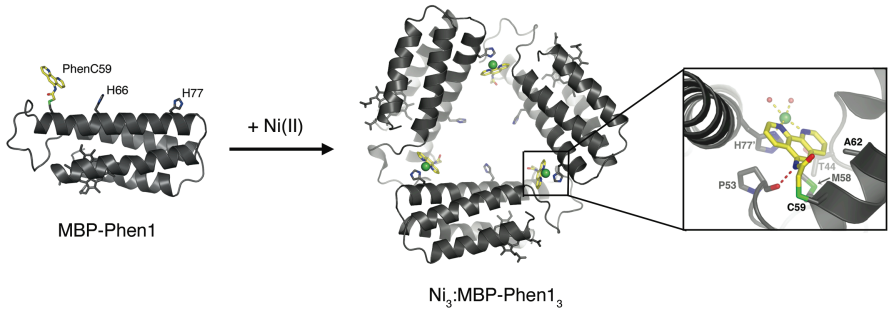
[42] M.R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.*, 114 (1992) 825-831.

[43] M.A. Case, G.L. McLendon, *Acc. Chem. Res.*, 37 (2004) 754-762.

[44] R.R. Davies, M.D. Distefano, *J Am Chem Soc*, 119 (1997) 11643-11652.

[45] C.H.B. Chen, L. Milne, R. Landgraf, D.M. Perrin, D.S. Sigman, *ChemBioChem*, 2 (2001) 735-740.

Chapter 2: A Superprotein Triangle Driven by Nickel(II) Coordination: Exploiting Non-Natural Metal Ligands in Protein Self-Assembly



2.1 Abstract

Recently, a strategy (Metal-Directed Protein Self-Assembly, MDPSA) was devised that utilizes the simultaneous stability, lability and directionality of metal-ligand bonds to drive protein-protein interactions. Here it is shown that both the structural and the functional scope of MDPSA can be broadened by non-natural metal chelating ligands incorporated onto protein surfaces. A cytochrome cb_{562} variant, MBPPhen1, which features a covalently attached phenanthroline (Phen) group on its surface, self-assembles into an unusual triangular architecture (Ni_3:MBPPhen1_3) upon binding Ni, owing to specific Phen-protein interactions. The crystal structure of Ni_3:MBPPhen1_3 reveals that the Phen group is buried in a small pocket on the protein surface, which result in an unsaturated Ni coordination environment.

2.2 Introduction

Nature utilizes proteins as building blocks to construct a wide variety of self-assembled nanoscale architectures, whose structures and dimensionalities are optimized for function. Some examples include: hollow 3-D cages for storage and compartmentalization, tubular structures for selective molecular transport, 2-D arrays for structural organization and templating, and 1-D filaments for mechanical motion and vectorial transport [3]. Despite advances in protein design and engineering [4-6], attaining the structural and functional sophistication of biological multi-protein architectures remains a distant goal. In order to circumvent the immense challenge of controlling the non-covalent interactions that hold these assemblies together, a strategy was devised ("Metal-Directed Protein Self-Assembly" (MDPSA)) that simultaneously utilizes the stability, lability and directionality of metal-ligand bonds in order to drive protein self-assembly [1, 7, 8]. The use of metal coordination to control protein self-assembly is attractive from both structural and functional perspectives: whereas the directionality and symmetry inherent in metal coordination can govern the overall supramolecular assembly, the resulting interfacial metal centers introduced the potential for new reactivities within biological scaffolds. With these advantages in mind, it was postulated that the structural and functional scope of MDPSA could be further augmented with non-natural metal ligands. In this chapter, the Ni-dependent self-assembly properties of a cyt cb_{562} variant, MBPPhen1, which features a covalently attached phenanthroline (Phen) group on its surface (Figure 2.1) is described. MBPPhen1 not only forms an unusual supramolecular architecture owing to specific Phen-protein interactions, but also

presents coordinatively unsaturated Ni centers, which are capable of binding anions within the assembly.

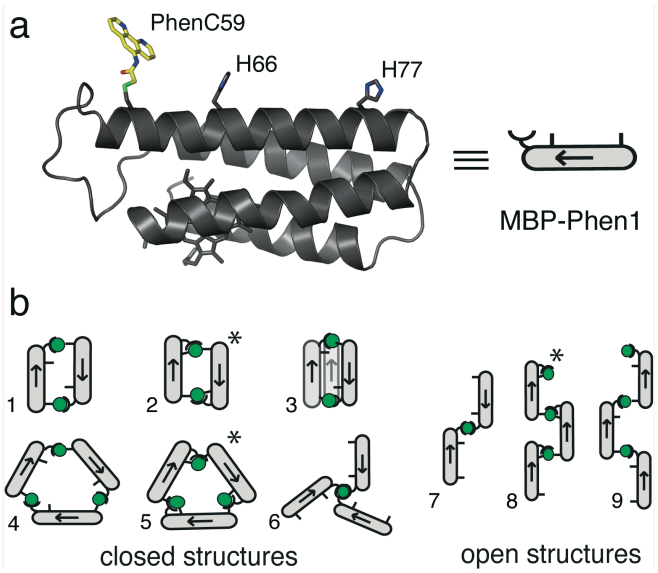


Figure 2.1. (a) MBPPhen1 structure with PhenC59 modeled in the extended form, and its cartoon representation. (b) Some possible Ni-induced oligomerization modes of MBPPhen1, where each Ni ion is coordinated by at least one Phen group. Structures that feature *i*/*i* +7 His66-PhenC59 coordination of Ni are labeled with an asterisk.

2.3 Results and Discussion

In order to site-selectively nucleate metal binding on protein surfaces, the Tezcan Group has adopted the strategy of employing multidentate motifs to outcompete the mostly monodentate sidechain functionalities for metal coordination. It was previously

shown by Salgado *et al.*, that a construct (MBPC-1), which contains two *i*/*i*+4 bis-His motifs near the termini of a single helix indeed self-assembles into different superstructures, using one or two of these bis-His motifs for Ni(II), Cu(II) and Zn(II) coordination [1, 7, 8]. Based on this work, it was imagined that non-natural multidentate ligands such as Phen with a single-point attachment would offer more structural flexibility than a bis-His motif while allowing non-proteinaceous functionalities to be incorporated into protein assemblies [10]. A cysteine-specific derivative of Phen (5-iodoacetamido-1,10-phenanthroline, IA-Phen) has previously been used to generate stable metal binding sites on proteins [11, 12]. Using IA-Phen, MBPPhen1, which contains a single Phen group covalently bound to Cys59 (PhenC59) near the N-terminus of Helix3 and His77 incorporated at the opposite end (Figure 2.1a), was constructed. Additionally, MBPPhen1 contains a His at position 66 (two helix turns away from Cys59) with the idea that this residue could form an *i*/*i* +7 tridentate acceptor motif together with PhenC59.

As an initial experiment, metal-binding titrations with Ni^{2+} were performed on MBPPhen1. It was hypothesized that the engineering of a tri-dentate motif, formed between the C59-Phen and the H66 residue, would create a high-affinity metal-binding site that would readily out-compete monodentate metal-binding functionalities replete on the protein's surface. Ni^{2+}, in particular, was chosen because of its ability (amongst late first-row transition metals) to form stable octahedral complexes [13]. Metal-binding titrations could be followed via UV-visible spectroscopy as a result of the non-natural Phen moiety, which in addition to providing a high-affinity binding site also provides a convenient spectroscopic handle for following metal binding interactions. Upon addition

of sequential aliquots of Ni^{2+}, a distinct 6 nm red shift in the Phen π-π^* absorbance band was observed ($\lambda_{max\text{-}free}$: 268, $\lambda_{max\text{-}bound}$: 274 nm) [14] with a clean isosbestic point at 272 nm, suggestive of a two-state process (Figure 2.2). The sharp transition observed in the binding isotherm, which occurs when the [Ni^{2+}] = [MBPPhen1], indicates that the metal-binding interaction is strong and can be reasonably assumed to be less than 1 µM [15].

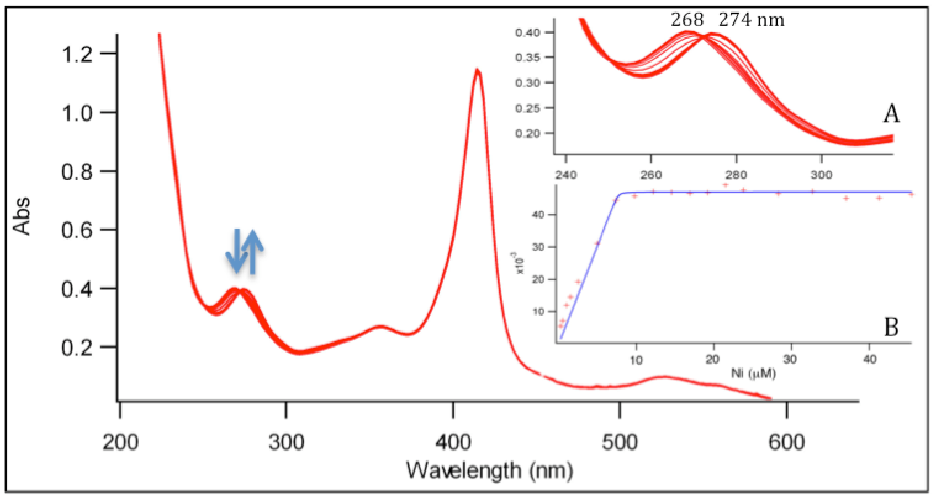


Figure 2.2. Titration of MBPPhen1 with Ni(II) as monitored by UV-vis spectroscopy. (Inset A) The redshift of the phenanthroline λ_{max} from 268 nm to 274 nm upon metal binding. (Inset B) The binding isotherm for MBPPhen1 to Ni(II), fit to a two-state model (blue line), where (+) represents $\Delta A_{280\ nm}$. The shape of the isotherm confirms 1:1 Ni:Phen binding and indicates a K_d < 1 µM.

While the metal-binding titrations definitively conclude the coordination of the Ni^{2+} to the Phen moiety, the experiment does not conclusively report on the coordination of the engineered H66 in the (*i*+7) position. In order to elucidate whether-or-not the H66 ligand was coordinating, the stability of the protein in the presence and absence of Ni^{2+} was determined. Previous work with engineered bis-His (*i*/*i*+4) motifs have shown that

the coordination of a metal ion by an intra-helix chelation motif can stabilize the protein fold [16], and it was hypothesized that determining the stability of MBPPhen1 in the presence and absence of Ni^{2+} would provide a convenient handle for determining coordination of the H66 residue to the Ni bound Phen moiety. As can be seen in Figure 2.3, a shift in the $GuHCl_{mid\text{-}point}$ (defined as the amount of GuHCl (M) at which [Unfolded Protein]=[Folded Protein]), by ~ 0.5 M in the Ni^{2+}:MBPPhen1 case, as compared to the Ni^{2+}:CM-MBP variant (were the Phen moiety is replace by a S-carboxymethyl group to neutralize cysteine reactivity), indicates the coordination of the H66 residue via the proposed *i*/*i*+7 tridentate motif. Interestingly, a shift of ~ 0.1 M GuHCl by the metal-free MBPPhen1 (as compared to CM-MBP) along with a concomitant increase in the slope of unfolding (Table 2.1), indicated the possibility of unforeseen Phen-protein interactions. Due to this unexpected complication, the assumption that the unfolding pathway follows a simple two-state model cannot be confirmed. Therefore no definitive thermodynamic assignments ascribing a $\Delta\Delta G_{metal}$ for the metal-mediated stabilization are made. However, a tentative assignment for the $\Delta G_{unfolding}$ is given in Table 2.1.

With the establishment of a multi-dentate metal binding site on the surface of MBPPhen1, the self-assembly properties of MBPPhen1 in the presence of Ni^{2+} was explored. The multitude of possible metal-mediated protein oligomerization

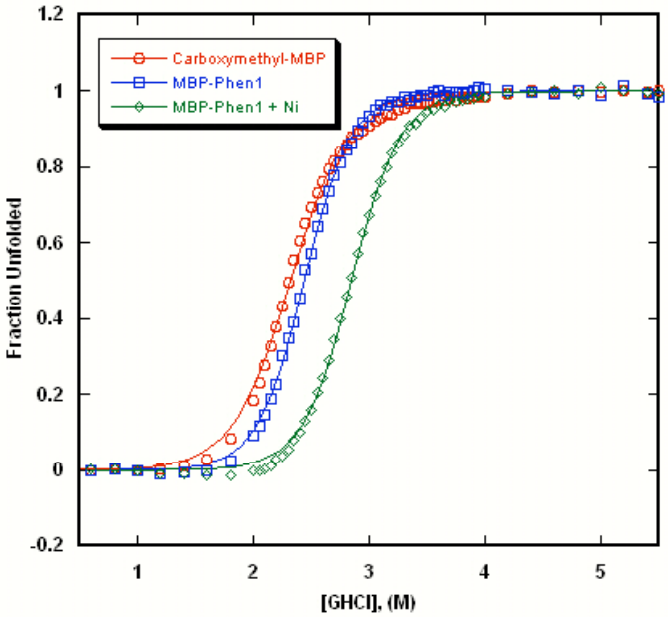


Figure 2.3. GuHCl unfolding curves, and fits, for CM-MBP (with EDTA), MBPPhen1 (with EDTA) and MBPPhen1 (in the presence of 1 mM Ni^{2+}). The parameters obtained from the fits are listed in Table 2.1.

Table 2.1. Parameters obtained from fits to the unfolding curves shown in Figure 2.3. The free energies for unfolding (ΔG_{unf}) obtained from these parameters are 5.4 kcal/mol (CM-MBP), 7.0 kcal/mol (MBPPhen1 in the absence of metal), and 7.9 kcal/mol (MBPPhen1 in the presence of Ni^{2+}).

A

Carboxymethyl-MBP	Value
Mid-point (M of GuHCl)	2.32 ±0.0052
Slope (kcal•mol^{-1}•M^{-1})	2.34 ±0.048
χ^2	0.018
R	0.998

B

MBP-Phen1 (Ni^{2+})	Value
Mid-point (M of GuHCl)	2.85 ± 0.0030
Slope (kcal•mol^{-1}•M^{-1})	2.76 ± 0.034
χ^2	0.00498
R	0.999

C

MBP-Phen1 (EDTA)	Value
Mid-point (M of GuHCl)	2.43 ± 0.0023
Slope (kcal•mol^{-1}•M^{-1})	2.875 ± 0.030
χ^2	0.00498
R	0.999

modes (Figure 2.1b) and the non-negligible interactions that could be formed between protein surfaces in these states make the *a priori* assignment of the thermodynamically preferred superprotein assembly challenging. Therefore, it was first attempted to determine the crystal structure of the Ni adduct of MBPPhen1. Crystals of MBPPhen1 were obtained in two space groups ($P2_1$ and $P6_322$) from two similar but different solution conditions containing equimolar protein and Ni, and their structures were determined at 2.4 and 3.0 Å resolution, respectively. Both structures (PDB IDs: 3FOO and 3FOP) reveal a unique triangular assembly, Ni_3:MBPPhen1_3, with perfect C_3 symmetry (Figure 2.4). Each vertex of this triangle is formed by a Ni ion coordinated to PhenC59 from one protein monomer and His77 (N_∂) from another (Figure 2.4c), whereby the three Ni's lie on the plane of the triangle, approximately 30 Å from one another. Ni-protein coordination appears to be the primary driving force for self-assembly, as the docking interactions between each protein monomer are minimal and non-specific. The flat shape and dimensions of Ni_3:MBPPhen1_3 are reminiscent of the 3-fold symmetrical components of biological cages such as the 432-octahedral ferritin shell [17] and the 532-icosahedral bacteriophage MS2 capsid [18].

The Phen group, instead of extending into the solvent, is positioned in a small hydrophobic crevice underneath the 50's loop, further stabilized by a H-bond between the PhenC59 amide nitrogen and the Pro53 backbone carbonyl (Figures 2.4c and d). These favorable interactions appear to explain the apparent increase in stability observed in the metal-free MBPPhen1 unfolding titration (Figure 2.3 and Table 2.1). The placement of Phen is the key to the open Ni_3:MBPPhen1_3 architecture: it protects the Ni ion from coordination by a second Phen group and allows only one other protein monomer to

coordinate through His77 in the *cis* position, which ultimately results in a severely unsaturated Ni coordination geometry. Apparently, the burial of Phen group is sufficiently favorable to overcome the tridentate Ni coordination by the *i*/*i*+7 His66-PhenC59 motif.

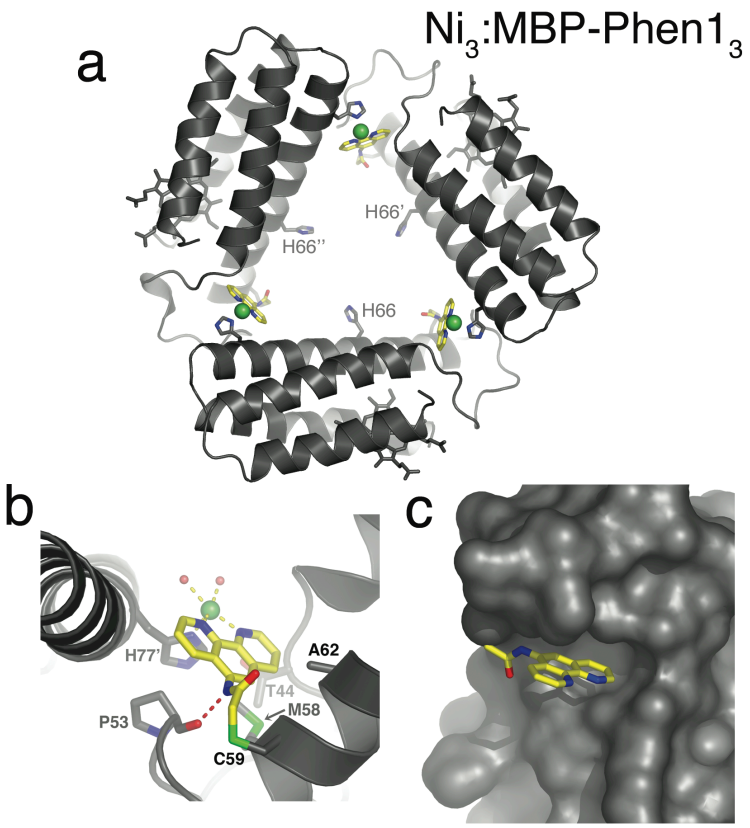


Figure 2.4. (a) Top-view of Ni_3:MBPPhen1_3. Ni ions are shown as green spheres, and PhenC59 is highlighted in yellow. (b) Coordination environment of Ni-PhenC59. The H-bond between the P53 carbonyl and the PhenC59 amide nitrogen is indicated with a red dashed line. Aquo/chloride ligands are shown as red spheres. (c) Burial of PhenC59 under the 50's loop.

While His77 and PhenC59 are clearly defined in the electron density maps, other coordination sites cannot be unambiguously assigned due to the resolution limits. Because the tunability of the beamline (SSRL BL9-1) is limited, the anomalous diffraction data for the Ni centers was not obtained. The fact that Ni is the only transition

metal in crystallization solutions, and the refined B-factors for the modeled Ni-ions (on average: 32 $Å^2$ in $P2_1$ crystals, 31 $Å^2$ in $P6_322$ crystals) match well with those of the coordinating Phen groups (on average: 29 $Å^2$ in $P2_1$ crystals, 33 $Å^2$ in $P6_322$ crystals) gives strong credence to their assignment. The assignment for the diffuse electron density near Ni trans to the Phen group is somewhat less certain (Figures 2.5 and 2.6). The broadness of the electron density is most consistent with two atoms; therefore, two water molecules were initially modeled into this density. The average B-factors for the coordinated waters (which showed a ± 0.3 Å variation on their distance from the Ni ions, Figure 2.5e) after refinement were 28 $Å^2$ for the $P2_1$ crystals (24 waters) and 4 $Å^2$ for the $P6_322$ crystals. Since the lower thermal factors for coordinated ligands compared to Ni is not reasonable, chloride ions were considered as an alternative possibility. Both crystal forms emerged from solutions containing 200 mM Cl^-; all other possible candidates (PEG fragments, TRIS) were ruled out due to their shapes and sizes that are incompatible with the electron density. The average refined B-factors for coordinated Cl^- ions were 58 $Å^2$ for the $P2_1$ crystals (24 ions) and 39 $Å^2$ for the $P6_322$ crystals. Based on these observations, the electron density is assigned to a mixture of water and chloride ligands that are in rapid exchange. Although the observation of bound chloride ligands in an aqueous environment is unexpected, it is not without precedence [19, 20]. For reporting purposes, the final submitted structure for the $P2_1$ crystals (3FOO) contains 24 water ligands, and that for $P6_322$ crystals (3FOP) contains 4 chloride ligands, for the sake of consistency with the observed B-factors.

The lattice packing interactions in both Ni_3:MBPPhen1_3 crystals are particularly noteworthy. In the $P2_1$ form, there are four crystallographically distinct, but identical

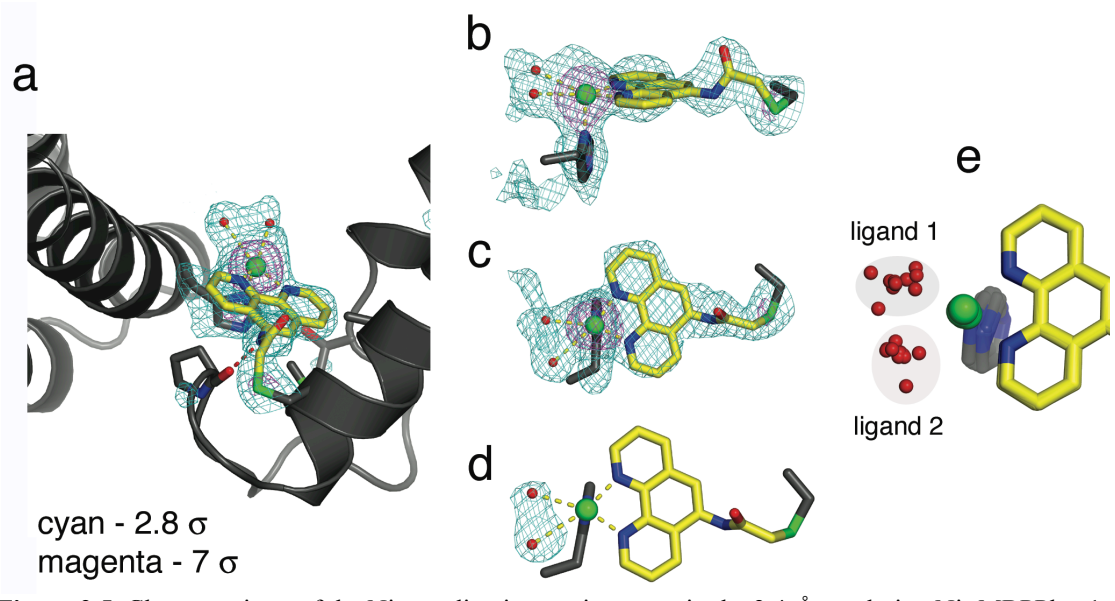


Figure 2.5. Close-up views of the Ni coordination environment in the 2.4-Å resolution Ni_3:MBPPhen1_3 structure ($P2_1$ form). (a) The view corresponding to Figure 4 showing the F_o-F_c omit difference density maps calculated using a model where the Ni-center and the coordinating ligands were omitted (cyan mesh – 2.8 σ, magenta mesh – 7 σ). (b), (c) Side and top-views of the Ni coordination environment, and the corresponding F_o-F_c omit difference density map. (d) F_o-F_c omit difference density maps calculated using a model where only the coordinating ligands trans to Phen were omitted (cyan mesh - 3 σ). (f) Superposition of the Ni coordination environments in the twelve crystallographically independent monomers in the asymmetric unit, showing the variation in the position of the two non-proteinaceous ligands modeled as water molecules.

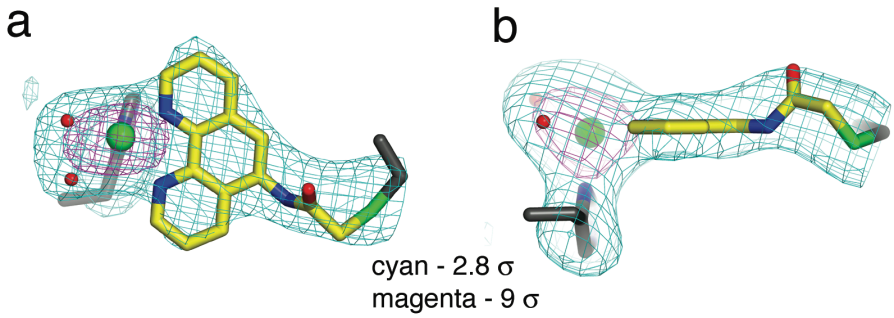


Figure 2.6. Ni coordination environment in the 3.15-Å resolution Ni_3:MBPPhen1_3 structure ($P6_322$ form) and the corresponding F_o-F_c omit difference density maps as viewed from the top (a) and the side (b).

(rmsd-C_α between all trimers = 0.3 Å) copies of the trimer in the asymmetric unit, which stack up along their threefold symmetry axes to form a tubular architecture (Figure 2.7a). Each of the four Ni_3:MBPPhen1_3 trimers adopts a different orientation around the long axis of the tube, giving rise to three distinct trimer-trimer interactions. In the lattice, the tubular units are further stacked end-on-end infinitely, which, due to the superposition of the four different trimer orientations, adopt an apparent hexagonal geometry. The resulting hexagonal tubes form a tightly packed 2-D array (50% solvent content) (Figure 2.7b); however, a possible six-fold symmetry is broken due to the staggered alignment of the trimers between some of the neighboring tubes. In the $P6_322$ crystals, the Ni_3:MBPPhen1_3 trimers are similarly arranged to form hexagonal tubular structures (Figure 2.7c). In contrast to the $P2_1$ lattice, the trimers of all adjacent tubes are now coplanar, which is required for generating the two-, three- and six-fold symmetry components of the $P6_322$ space group. Moreover, a central hexagonal tube is not accommodated in this lattice, leading to a large cavity of 6-nm diameter and an increased crystal solvent content of 64%.

Although these higher-order architectures are only formed in crystals, our observations suggest that open, symmetrical protein superstructures such as Ni_3:MBPPhen1_3 could be in principle be utilized as building blocks for porous protein frameworks. Analogous small-molecule based frameworks (commonly referred to as Metal-Organic Frameworks, or MOFs), have received great interest in recent years due to their unique chemical and structural properties which make then amenable to gas storage [21], gas separation [22] or heterogenous metal-based catalysis [23] (just to list a few). The porous Ni_3:MBPPhen1_3 crystals can be thought of as representing a new class of

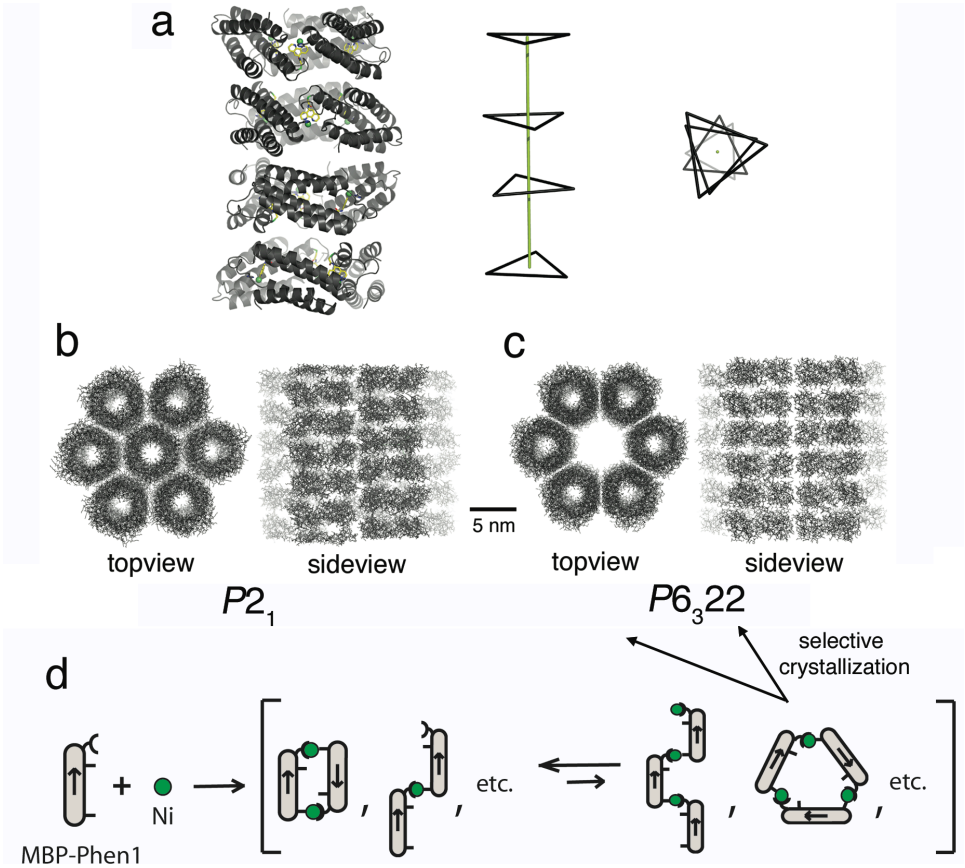


Figure 2.7. (a) Four Ni_3:MBPPhen1_3 trimers in the asymmetric unit of the $P2_1$ crystals, and their triangular representation (Ni's as vertices) viewed from the side and the top. (b,c) Lattice packing arrangement of Ni_3:MBPPhen1_3 in $P2_1$ and $P6_322$ crystal forms. (d) Suggested Ni-induced oligomerization behavior of MBPPhen1 in solution.

Bio-Inorganic Frameworks (or BIFs). A key advantage to BIFs would be the inherent functionality of the protein monomer, which could work separately, or in conjunction with the interfacial metal center to realize new metal-based functionality not readily seen in biological systems. The Ni_3:MBPPhen1_3 crystals, with their coordinatively unsaturated Ni centers, would seem to readily lend themselves to this type of reactivity. As an initial proof-of-concept to illustrate that the Ni center would be accessible to small-molecules within the protein framework, both Ni_3:MBPPhen1_3 crystals and similar metal-free

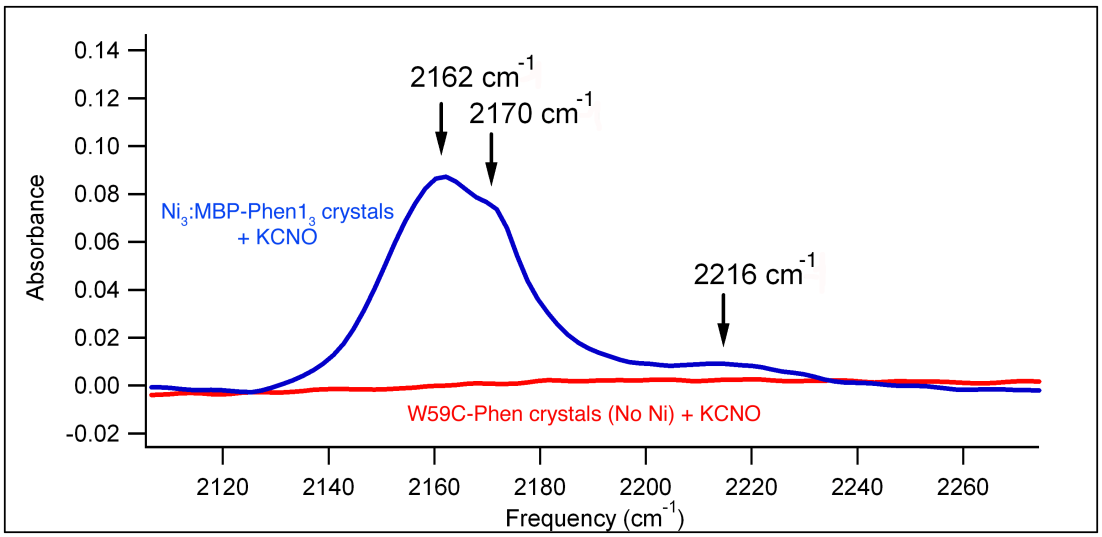


Figure 2.8. IR spectra of Ni_3:MBPPhen1_3 and W59C-Phen crystals soaked in a 20 mM KNCO solution. Only the protein-transparent region of the IR spectrum is shown for clarity. The fact that no NCO stretching frequencies are observed for W59C-Phen crystals indicates that NCO^- does not non-specially associate with protein crystals (*i.e.*, there is no free NCO^- within the crystals). As a reference, the similar Ni(terpy)$(NCO)_2(H_2O)$ complex exhibits two asymmetric C-N stretching bands (n_s) at 2170 and 2230 cm^{-1} attributed to two different modes of Ni-NCO interactions [2]. In comparison, the n_s stretching frequency in $Ni(NCO)_4^{2-}$ complex is observed at 2196 cm^{-1} [9].

crystals (an cyt cb_{562}-W59C-Phen variant crystallized in a metal free environment), were soaked in 20 mM KNCO for a period of 30 min. After extensive washing of the crystals, by exchanging into fresh mother liquor solutions, the resulting crystals were crushed in KBr and the IR spectrum of the crystals taken (Figure 2.8). The resulting IR spectra of the Ni_3:MBPPhen1_3 crystals showed 3 distinct bands at 2162, 2170 and 2216 cm^{-1}, respectively, indicative of the formation of a Ni(polypyridal)(NCO) complex [24]. Importantly, no such bands appeared in the metal free sample (Figure 2.8).

The fact that the identical Ni_3:MBPPhen1_3 structure is found in many different crystal packing environments provides strong evidence for its existence in solution. Yet, it has not been possible to observe the trimer at low protein concentrations (< 1 mM) by

hydrodynamic measurements (Figure 2.14), possibly because the lone linkage between each monomer in this structure is a single, and possibly labile, Ni-His bond. To capture Ni_3:MBPPhen1$_3$ in solution, chemical crosslinking and kinetic trapping experiments were attempted. A glance at the monomer-monomer interfaces in Ni_3:MBPPhen1$_3$ reveals that there are two pairs of Lys residues (K27/K42 and K51/K83) with an N_z-N_z distance of 9-11.5 Å (Figure 2.8) that could be linked across each interface using imidoester crosslinkers. Three such crosslinkers, DMS, DMP and DMA (Figure 2.9), whose spacer lengths are 11, 9.2 and 8.6 Å, respectively, were employed. While all three lead to the crosslinking of dimeric species, only the sufficiently long DMS and DMP yield appreciable quantities of trimer as expected from the crystal structure (Figure 2.10). Furthermore, SDS-PAGE results indicate that the presence of Ni^{2+} is required for the capture of both dimeric and trimeric species (Figure 2.10).

In an alternative strategy, it was envisioned that Ru(II) could serve as an exchange-inert surrogate for Ni(II) and permit the kinetic trapping of the crystallographically observed trimer. As predicted, incubation of MBPPhen1 with $Ru(DMSO)_4Cl_2$ yields various higher order species, including some trimeric forms (Figure 2.11). The trimeric species from Ru- as well as DMS-treated MBPPhen1 samples were isolated by size-exclusion and ion-exchange chromatography, and their metal contents determined, to further characterize their oligomeric composition. Each MBPPhen1 molecule contains a single Fe atom as a part of the covalently-attached heme group; therefore, the Ni:Fe and Ru:Fe ratios can be used to assign which superstructures shown in Figure 2.1b could be populated in solution. Inductively coupled plasma-optical emission spectroscopy

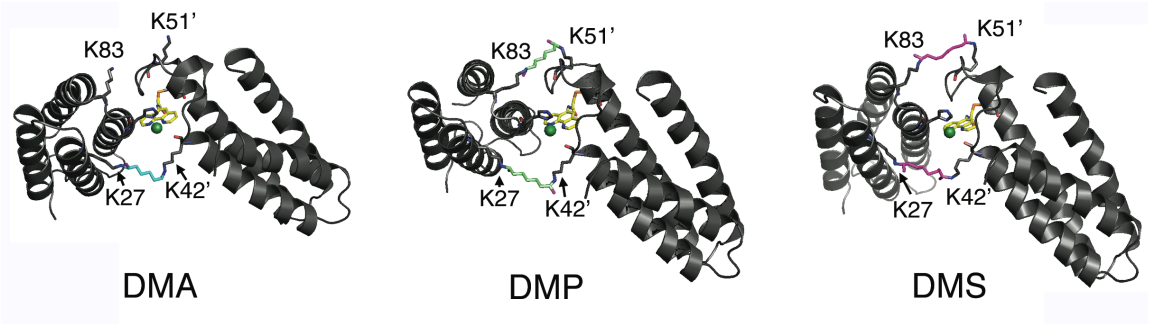


Figure 2.9. Models for DMA/DMP/DMS-mediated crosslinking of K83-K51 and K27-K42 across crystallographically-observed monomer-monomer interfaces in Ni_3:MBPPhen1$_3$ trimers. While the crosslinking of both Lys pairs appear to be readily accommodated (i.e., in low-energy, extended conformations of Lys, and no steric clashes with the intervening protein medium) with DMP and DMS, only the K27-K42 pair appears to be within reach for DMA.

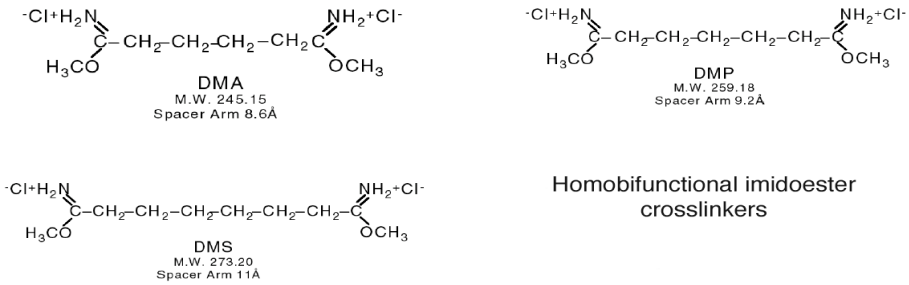


Figure 2.10. The amine-reactive, homobifunctional, imidoester crosslinkers used in this study (Figure adapted from Pierce).

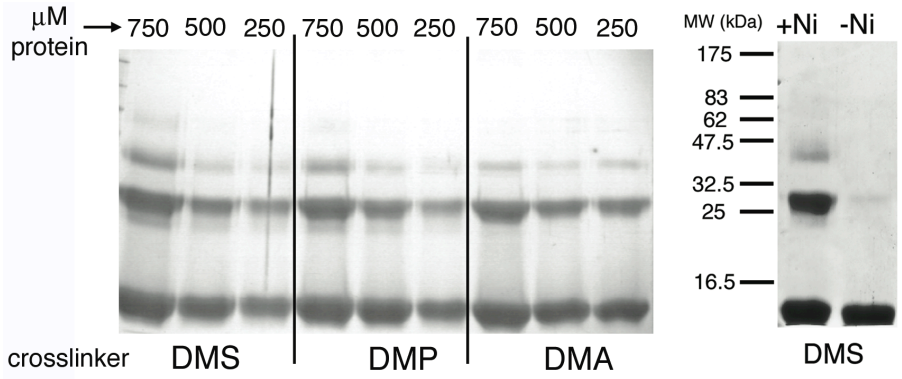


Figure 2.11. (left panel) Concentration-dependent crosslinking of MBPPhen1 by DMS, DMP and DMA in the presence of equimolar Ni as probed by SDS-PAGE electrophoresis. Crosslinking reactions were carried out using 250, 500 or 750 µM protein/$NiCl_2$ and 30-fold excess crosslinker at room temperature for 50 min. The monomeric mass is 12376 Da; the expected trimeric mass containing three DMS crosslinkers is 37737 Da. All crosslinkers yield dimeric, trimeric and some higher-order species in the presence of Ni; the yield of oligomeric species increases with crosslinker length. (right panel) The formation of oligomeric species is clearly metal dependent, as the yield for crosslinked dimeric and trimeric forms is markedly higher in the presence of Ni.

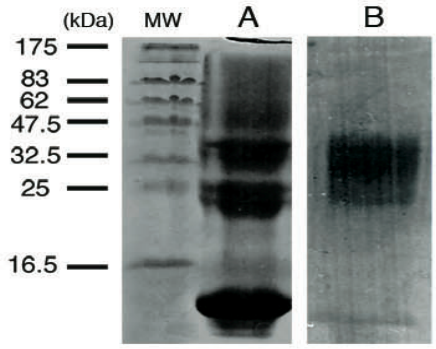


Figure 2.12. 12% Native-PAGE gel of Ru-crosslinked MBPPhen1 before (A) and after (B) purification. The broad band for the trimeric species indicates sample heterogeneity, which is also reflected by the somewhat broad sedimentation coefficient distribution (Figure 2.14).

(ICP-OES) measurements on thoroughly dialyzed samples yield a Ni:Fe ratio of 1.1 (± 0.1) and a Ru:Fe ratio of 1.4 (± 0.1) for the DMS- and Ru-crosslinked species, respectively (Table 2.2), consistent with the triangular species that contains one extrinsic metal per protein monomer. The higher-than-unity Ru:Fe ratio is ascribed to a fraction of trimers that contain Ru ions bound to other surface sites such as His66 not involved in protein crosslinking. Furthermore, a red shift in the Phen π-π* band from 268 to 274 nm in both Ni- and Ru-crosslinked species indicates that the Ni and Ru ions in solution are predominantly bound to PhenC59 (Figure 2.12).

Lastly, the un-crosslinked MBPPhen1 and the isolated trimeric species were subjected to sedimentation velocity measurements. As expected, in the absence of Ni (Figure 2.14, black line), MBPPhen1 yields a monomeric species (S_{max} = 1.8). A shift to a dimeric assembly (S_{max} = 2.3) is obtained at 500 μM protein and 500 μM Ni (Figure 2.14. blue line). This assembly is most consistent with an extended dimeric species (expected S_{max} = 2.4), such as structure 7 in Figure 2.1b. This result also indicates that significant

quantities of the trimeric species, in solution, require high protein concentrations (> 1mM).

Table 2.2. Metal analysis of the DMS- and Ru-crosslinked MBPPhen1 trimers, as determined by ICP-OES. All samples were run in triplicate, and signals were corrected by subtraction of buffer background. These analyses yield a Ni:Fe ratio of 1.1 (± 0.1) and a Ru:Fe ratio of 1.4 (± 0.1). A control sample of wild type cyt cb_{562} (not containing Phen) prepared in exactly the same fashion as the DMS-crosslinked species contains no Ni.

DMS-crosslinked trimer

Analyte (λ, nm)	Mean intensity	Conc. (mg/L)	Conc. µM
Ni (231.6)	3932.0	0.145 (±0.0084)	2.47
Fe (238.2)	7097.7	0.126 (±0.013)	2.25
Fe (239.56)	7513.1	0.129 (±0.014)	2.31

Ru-crosslinked trimer

Analyte (λ, nm)	Mean intensity	Conc. (mg/L)	Conc. µM
Ru (349.89)	23488.3	1.77 (±0.09)	15.7
Ru (240.27)	27807.1	1.74 (±0.094)	16.2
Fe (238.2)	42070.1	0.616 (±0.036)	11.0
Fe (239.56)	42430.4	0.596 (±0.032)	10.7

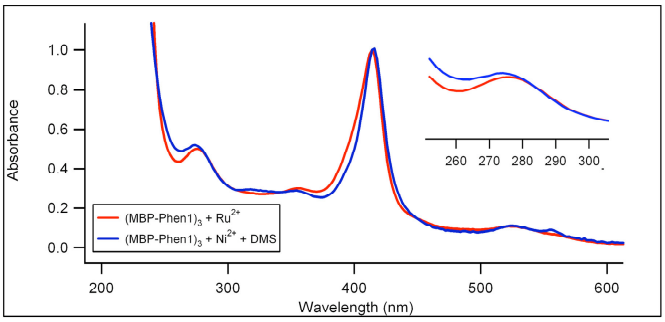


Figure 2.13. UV-visible spectra of the crosslinked trimeric species. Spectra of Ru(II)- (red) and DMS-crosslinked MBP-Phen1 in the presence of Ni(II) (blue) indicate that both species contain Phen groups that are fully metal-bound, with λ_{max} = 276 nm and λ_{max} = 274 nm for each species, respectively.

Such MBPPhen1 concentrations are not conducive to SV measurements due to the overwhelming intensity of heme absorption bands.

The sedimentation coefficient distributions for both the kinetically-trapped and chemically-crosslinked species are centered at 3.3 Svedbergs (Figure 2.14), in perfect agreement with calculations based on the Ni_3:$MBPPhen1_3$ crystal structure. The theoretical sedimentation coefficients for the extended and the closed trimers (structures

8 and 3 in Figure 2.1a), in contrast, are calculated to be 3.0 and 3.6 Svedbergs, respectively (Table 2.3). Crosslinking studies, taken together with the fact that the Ni_3:MBPPhen1$_3$ structure is retained in several lattice packing environments, lend strong credence to the formation of the crystallographically observed trimeric structure in solution, albeit at high (> 1mM) protein concentrations.

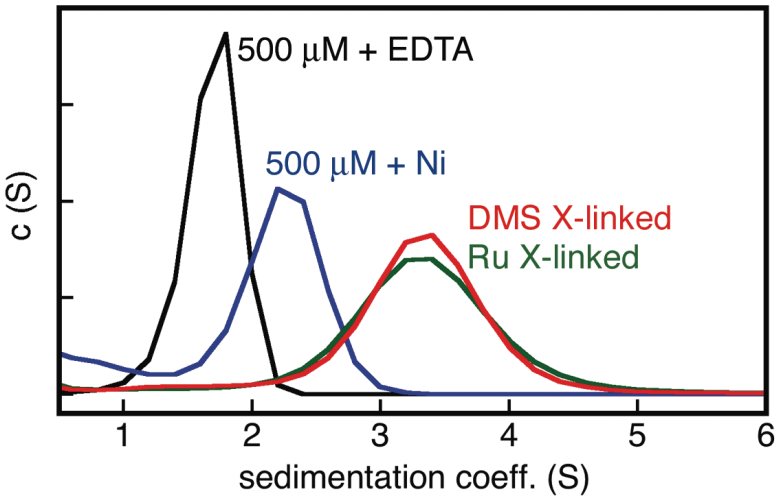


Figure 2.14. Normalized sedimentation coefficient distributions of MBPPhen1 under various conditions. In the absence of Ni (black line), only the monomeric form is observed (S_{max} = 1.8). The species with S_{max} = 2.3 (blue line) obtained at 500 µM protein and Ni is most consistent with an extended dimeric species (expected S_{max} = 2.4) such as structure 7 in Figure 2.1b. The population of the trimeric species in solution in significant quantities apparently requires high protein concentrations (> mM) such as those used for crystallization. Such MBPPhen1 concentrations are not conducive to SV measurements due to the overwhelming intensity of heme absorption bands. As detailed in Table 2.3, S_{max} = 3.3 observed for DMS- and Ru-crosslinked species is most consistent with a triangular species, whereas the theoretical sedimentation coefficients for the extended and the closed trimers (structures 8 and 3 in Figure 2.1b) are calculated to be 3.0 and 3.6 Svedbergs, respectively.

Table 2.3. Theoretical and experimentally determined sedimentation coefficients for various MBPPhen conformations (as illustrated in Figure 2.1b).[a]From Reference [1].

MBP-Phen1 Conformation	Mw (Da)	Calculated sedim. coeff. (S)	Experimental sedim. coeff. (S)
Monomer	12376	1.6	1.8
Closed dimer (Structure 1)	24752	2.67	2.6[a]
Extended Dimer (Structure 7)	24752	2.46	N/A
Closed Trimer (Structure 3)	37128	3.56	3.5[a]
Extended Trimer (Structure 9)	37128	3.0	N.A
DMS-crosslinked Ni_3:MBP-$Phen1_3$	37737	3.26	3.3
Ru_3:MBP-$Phen1_3$	37431	3.27	3.3

2.4 Conclusions

Synthetic metal coordinating functionalities have previously been employed for stabilizing coiled-coil assemblies [25], constructing reactive metal binding sites in protein interiors [11], and tuning the potentials of redox centers [26], among others [10]. It has been demonstrated here that such non-natural ligands incorporated onto protein surfaces can lead to novel biological architectures as well as potentially reactive metal coordination sites within these scaffolds. The wide array of functionalities available in the synthetic inorganic chemist's toolbox thus could provide a powerful means to generate structural and functional diversity in protein self-assembly.

2.5 Experimental Section

Site Directed Mutagenesis and Protein Expression/Purification/Characterization

Site directed mutagenesis was performed on the pETc-b562 plasmid (denoted as wild-type) [27], using a QuikChange kit (Stratagene) and employing primers obtained from Integrated DNA Technologies. The mutant plasmids were transformed into XL-1 blue *E. coli* cells and purified using the QIAprep Spin Miniprep kit (Qiagen). Four rounds of mutagenesis were executed to obtain the final mutant W59C R62A H63A D66H K77H (hereafter referred to as MBP). Sequencing of the mutant plasmid was performed at the UCSD Moore's Cancer Center.

The resulting mutant plasmid was transformed into protein expressing BL21(DE3) *E. coli* cells along with the *ccm* heme maturation gene cassette plasmid, pEC86 [28]. Cells were plated on LB agar containing the antibiotics ampicillin (100 μg/ml) and chloramphenicol (34 μg/ml), and grown overnight. A single colony was chosen and used to inoculate an LB solution containing the appropriate antibiotics. The resulting solution was incubated at 37^{o} C for 16 hours with rotary shaking at 250 rpm. No induction was necessary.

Mutant-expressing cells were harvested by centrifugation and frozen at -80 C for a period of 2 hours. Cells were subsequently thawed and lysed by sonication. The resulting lysate mixture was brought to pH 5 with the addition of 0.1M HCl and centrifuged at 16,000 g, 4^{o} C, for 1 hr. The soluble cell lystate fraction was then loaded onto a CM-Sepharose cation-exchange matrix (Amersham Biosciences), and purified using a NaCl gradient (0-0.5 M NaCl) in sodium acetate (pH 5) buffer. After exchange into 10 mM sodium phosphate (pH 8) buffer, the protein was further purified using an

Uno-Q (BioRad) anion exchange column on a DuoFlow chromatography workstation (BioRad) using a NaCl gradient (0-0.5 M NaCl in 10 mM sodium phosphate, pH 8). Protein purity was determined by SDS-PAGE gel electrophoresis. Verification of mutations was made through MALDI mass spectrometry (Found mass for MBP = 12137 amu; calc: 12136 amu)

Synthesis of Iodoacetamido-1,10-Phenanthroline (IA-Phen)

As a precursor, iodoacetic acid anhydride was freshly prepared by adding 2.64 g (12.8 mmol) of DCC to a stirred solution of 5.0 g (26.8 mmol) iodoacetic acid in 75 mL of ethyl acetate. Dicyclohexylurea precipitates immediately, but the mixture was allowed to stir for an additional 2 hrs, in the dark. The dicyclohexylurea was removed by filtration and the resulting solution was evaporated to near-dryness. 0.5 g (2.56 mmol) of 5-amino-1-10-phenanthroline (Polysciences) was dissolved in 90 mL of acetonitrile with slight heating. To this stirred solution, the iodoacetic acid anhydride, dissolved in 10 mL of acetonitrile, was added dropwise. The mixture was allowed to react, in the dark, overnight. The precipitated product was isolated by filtration and washed first with cold 5% sodium bicarbonate followed by water, and dried *in vacuo*. Both the ESI MS and NMR spectra correspond to previously reported literature values [29]. (Yield: 75%)

Functionalization of MBP with IA-Phen and iodoacetic acid

A solution of 0.3 mM MBP in degassed 0.1 M Tris buffer (pH 7.75) was treated with a 10-fold excess of dithiothreitol (DTT). The protein was allowed to reduce for a period of 1 hr. The protein was then dialyzed against 2 L of degassed 0.1 M Tris buffer

(pH 7.75), under an inert atmosphere, to remove the DTT. An approximate 10-fold excess of IA-Phen or iodoacetic acid was dissolved in 2 mL of degassed DMF, and added dropwise to the protein solution over the course of 1 min. The mixture was allowed to react in the dark at 25° C overnight, after which excess free label was removed using a 10-DG gel filtration column (Biorad). The labeled protein (MBPPhen1 or carboxymethyl-MBP (CM-MBP)) was dialyzed against 2 L of 10 mM sodium phosphate (pH 8) and subsequently purified on an Uno-Q anion-exchange column using an NaCl gradient (0-0.5 M NaCl in 10 mM sodium phosphate (pH 8) buffer). The purity of functionalized protein was determined by MALDI mass spectrometry and SDS-PAGE electrophoresis. Labeling yield: >90%. (Found mass for MBPPhen1 = 12376 amu; calc: 12372 amu).

Equilibrium Unfolding

An solution of unfolded protein (5 mL), containing 2 μM of protein (CM-MBP or MBPPhen1) and 1 mM of Ni^{2+} or EDTA was freshly prepared in a 20 mM Tris buffer (pH 7) with ~ 7.5 M guanidine HCl (GuHCl). In parallel, 3 mL of a folded protein solution containing 2 μM of protein (CM-MBP or MBPPhen1) in a 20 mM Tris (pH 7) buffer and 1 mM Ni^{2+} or EDTA was prepared. Over the course of the unfolding titration, the unfolded protein stock was titrated into the folded protein stock using an autotitrator (Microlab 500 Series), keeping the sample volume constant at 2 mL. The unfolding/folding reaction was monitored by CD spectroscopy, at 222 nm, on an Aviv 215 spectrometer. For every titration point, the solution was allowed to stir for 30 sec in order to reach equilibrium. This procedure was repeated for a minimum of 20 points

covering a GuHCl range of 0.1-6.5 M. Concentrations of the GuHCl stock solutions were calculated using the refractive indices of the folded and unfolded protein stock solutions [30]. Unfolding data were fit using Kaleidagraph (Synergy Software) assuming a two-state folding/unfolding equilibrium as described by Pace [31].

Ni^{2+} Binding Titrations

All absorption spectra were obtained on an HP 8452A spectrophotometer. MBPPhen1 concentration was determined based on heme absorption at λ = 415 nm (ε = 0.148 μM^{-1} cm^{-1}) [32]. A 1 mL solution of MBPPhen1 in 20 mM Tris buffer (pH 7) was freshly prepared from a concentrated stock using Hamilton syringes. To this solution, successive aliquots of a 500 μM $NiCl_2$ stock solution in 20 mM Tris (pH 7) were added. All data were baseline-corrected and adjusted for dilution. Absorbance at 280 nm, which displays the largest change upon Ni binding, was plotted as a function of Ni^{2+} concentration. The generated binding isotherm was fit to the following two-state equation [33] using IgorPro v. 4.0 (Wavemetrics, Inc.)

$$y = 0.5R\{A + B + x - \sqrt{(A + B + x)^2 - 4Bx}\}$$

where R is $\Delta\varepsilon$ (M^{-1} cm^{-1}), A is K_D (M), B is the concentration of MBPPhen1 (M), and x is concentration of Ni^{2+} (M).

Crosslinking of MBPPhen1 with lysine specific linkers dimethyl suberimidate (DMS), dimethyl pimelimidate (DMP), and dimethyl adipimidate (DMA) and ICP-OES/SV sample preparation

To a 0.5 mL solution containing 250, 500 or 750 μM MBPPhen1 and Ni^{2+} in 0.1 M HEPES buffer (pH 8), 10 μL of 0.3 M crosslinker (DMS, DMP or DMA, Pierce, Figure 2.4) in 0.1 M HEPES buffer (pH 8) was added. Aliquots of the reaction mixture were taken out every 10 min and quenched with 0.1 M Tris (pH 8.5). Progress of the reaction was monitored by SDS-PAGE electrophoresis. As a control, a parallel reaction was run under identical conditions, but containing 5 mM EDTA instead of Ni^{2+} (to ensure a metal free environment). After 50 min, any remaining free crosslinker was quenched with 0.1 M Tris (pH 8.5), and the reaction mixture was analyzed by SDS-PAGE gel electrophoresis (Figure 2.11).

The resulting protein mixtures were purified by size-exclusion chromatography on an ACA54 resin (Pall), with a running buffer of 20 mM Tris (pH 7) containing 100 μM $NiCl_2$ to prevent Ni^{2+} dissociation during purification. The trimer thus isolated was dialyzed against 2 × 2 L of metal-free Tris buffer to remove extraneous or weakly bound Ni^{2+}. The dialyzed products, and an aliquot of the corresponding dialysis buffer, were directly used for ICP-OES analysis (or for SV experiments, after further concentration). The UV-vis spectrum of the isolated (and dialyzed) trimeric species (Figure 2.13) shows a maximum at 274 nm for the Phen group, which suggests that the majority of Phen groups in solution are bound to Ni.

Crosslinking of MBPPhen1 with Ru^{2+} and ICP-OES/SV sample preparation

$Ru(DMSO)_4Cl_2$ was prepared as previously reported [34]. 2 mL of 1 mM MBPPhen1 in 0.1 M HEPES (pH 7) and 1 equivalent of $Ru(DMSO)_4Cl_2$ were allowed to react under an inert Ar atmosphere for 4 days at room temperature. The reaction was stopped by removal of excess $Ru(DMSO)_4Cl_2$ via a DG-10 (BioRad) gel filtration column. The trimeric species was isolated via size-exclusion chromatography on an ACA54 resin (Pall), using 20 mM Tris (pH 7) as the running buffer, and further purified on an Uno-Q anion-exchange column using an NaCl gradient. The isolated trimer (<10 mL total volume) was dialyzed against 2 × 2 L of metal-free Tris buffer to remove any unbound Ru^{2+}. The dialyzed product and an aliquot of the dialysis buffer were directly used for ICP-OES analysis (or for SV experiments, after further concentration). Identity of the trimer fractions was determined using native-PAGE gel electrophoresis (Figure 2.12). The UV-vis spectrum of the isolated trimeric species (Figure 2.13) shows a maximum at 276 nm for the Phen group, which suggests that all Phen groups in solution are bound to Ru.

Inductively coupled plasma-optical emission spectroscopy (ICP-OES) of Ru- and DMS-crosslinked MBPPhen1

Each MBPPhen1 molecule contains a single Fe atom as a part of the covalently-attached heme group; therefore, the Ni:Fe and Ru:Fe ratios can be used to assign the oligomeric composition of the isolated species.

All chemicals used for ICP-OES experiments were of analytical grade and water was deionized in a Milli-Q system (resistivity of 18.2 MΩ/cm). All glassware and

containers were rinsed 3× with 5% HNO_3. For standardization, a 30-mL solution containing 10 ppm of the metal analyte (Ni, Ru or Fe) was prepared by dilution from its certified 1000 ppm ICP stock solution (Ricca) with 5% HNO_3 in water. From this 10 ppm stock solution, 14 calibration solutions (0.0–5.0 ppm) were prepared by serial dilution. A series of 4 mL solutions, containing the crosslinked MBPPhen1 species, were prepared in a similar fashion. The Ni, Ru and Fe concentrations of the protein solutions were determined from their respective calibration curves. The ICP-OES results, as well as the specific wavelength used for the detection of each metal, are summarized in Table 2.

An inspection of both Ni_3:MBPPhen1$_3$ crystal structures indicates that there is a low affinity coordination site for Ni near the N-terminus of every monomer. The coordination sphere of this internally bound Ni consists of the backbone N and O of Ala1 and the carboxylate sidechain of Asp39. To ascertain that this binding site does not contribute to the "Ni-count" for the DMS-crosslinked sample, a control sample of wild-type cytochrome cb_{562} (which contains the low-affinity Ni site, but not the Phen groups for interprotein coordination) was prepared in exactly the same fashion as the DMS-crosslinked sample. ICP-OES analysis of this species does not show any trace of Ni.

Sedimentation Velocity Experiments

Sedimentation velocity (SV) experiments were performed in order to determine the solution-state oligomerization behavior of MBPPhen1 (Figures 2.14). All SV samples were prepared in 20 mM Tris buffer (pH 7). Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor

at 41,000 rpm for a total of 250 scans per sample. The following wavelengths were used for detection: 425 nm (10 μM protein) and 650 nm (500 μM protein).

All data were processed using SEDFIT [35]. Buffer viscosity, buffer density, and protein partial specific volume values were calculated at 25° C with SEDNTERP (http://www.jphilo.mailway.com). Partial specific volume (vbar) for MBPPhen1 mutant was calculated to be 0.7344 mg/ml, assuming a partial specific volume of heme of 0.82 mg/ml and 0.71 mg/ml for the phenanthroline [36]. All data were processed using fixed values for buffer density (r) (0.99764 g/ml) and buffer viscosity (0.0089485 poise).

Calculation of Theoretical Sedimentation Coefficients

Theoretical sedimentation coefficients (Table 2.3) were calculated using HYDROPRO v. 7.0 as previously reported [37]. Hypothetical oligomerization conformations, such as the extended dimer (structure 7 in Figure 2.1b) and the extended trimer (structure 9), were modeled using PYMOL [38]. The closed dimer (structure 1) and closed trimer (structure 3) conformations were approximated using Cu_2:MBPC-1_2 (PDB ID: 3DE8) and Ni_2:MBPC-1_3 (PDB ID: 3DE9) structures, respectively [39]. All values for vbar (0.7344 cm^3/g), buffer density (0.99764 g/cm^3) and viscosity ($8.99x10^{-3}$ P) were held constant.

Crystallography

All crystals were obtained by sitting drop vapor diffusion. MBPPhen1 was crystallized at 25° C with a precipitant solution consisting of 100 mM Tris (pH 8.5), 200 mM NaCl (for the $P2_1$ crystals) or 200 mM $MgCl_2$ (for the $P6_322$ crystals), 25% PEG

2000, and 3.3 mM $NiCl_2$. The drop consisted of 2 mL protein (3.3 mM in 20 mM Tris, pH 7) and 2 mL precipitation solution. Crystals appeared within one month, reaching a maximum size of ~ 100 μm × 100 μm × 50 μm. The crystals to be used for diffraction experiments were exchanged into a solution containing 20% glycerol, as a cryoprotectant, and frozen in liquid nitrogen or directly in the cryostream.

X-ray diffraction data for both crystal forms were collected at 100 K at Stanford Synchrotron Radiation Laboratory (BL 9-1) using 0.97 Å radiation. The data were processed using MOSFLM and SCALA [40]. The structure of Ni_3:MBPPhen1_3 was determined at 2.4 ($P2_1$ form) and 3.15 Å ($P6_322$ form) resolution, respectively, by molecular replacement with MOLREP, using the cyt cb_{562} structure (PDB ID 2BC5) [27] as the search model. The search model for each crystal form did not contain the heme or the Phen prosthetic groups; the observation of strong positive F_o-F_c density at expected positions for these groups confirmed the correct placement of protein monomers (12 in the $P2_1$ form and 2 in the $P6_322$ form) in the initial molecular replacement solutions. The topology and parameter files for the phenanthroline group were obtained using the Dundee ProDrg Server (http://davapc1.bioch.dundee.ac.uk/prodrg/index.html). Rigid-body, positional and thermal refinement with REFMAC [40, 41], along with manual rebuilding, and water placement with XFIT [42], produced the final models. For the $P6_322$ form, a two-fold non-crystallographic symmetry (NCS) restraint (tight main-chain and medium side-chain restraints) was applied throughout the positional/thermal refinement process. Similarly, for the $P2_1$ form, a twelve-fold NCS restraint (tight main-chain and loose side-chain restraints) was applied throughout the refinement. In addition, for the latter crystal form, TLS refinement using each twelve monomers as rigid bodies

was carried out during each refinement cycle. The Ramachandran plots were calculated with PROCHECK [43]. All figures were produced with PYMOL [38].

Table 2.4. X-ray data collection and refinement statistics for Ni:MBPPhen13

	Ni_3:MBP-Phen1$_3$ ($P2_1$)	**Ni_3:MBP-Phen1$_3$ ($P6_322$)**
Residues in complex	3 × (106 + 1 Heme + 1 Phen) + 6 Ni	3 × (106 + 1 Heme + 1 Phen) + 3 Ni
No. of complexes / asymmetric unit	4	2/3
Metal ions in asymmetric unit	24 Ni (12 intersubunit/12 intrasubunit)	4 Ni (2 intersubunit, 2 intrasubunit)
Waters in asymmetric unit	119	21
Unit cell dimensions (Å)	70.3 × 107.4 × 102.1	104.9 × 104.9 × 107.9
	$\alpha = \gamma = 90^o$, $\beta = 105.5^o$	$\alpha = \beta = 90^o$, $\gamma = 120^o$
Symmetry group	$P2_1$	$P6_322$
Resolution (Å)	98-2.4	90-3.0
X-ray wavelength (Å)	0.97	0.97
Number of Unique Reflections	55485	7479
Redundancy	2.7 (2.6)	13.1 (13.4)
Completeness (%)*	97.2 (97.0)	100.0 (100.0)
$\langle I / \sigma I \rangle$*	6.5 (1.3)	4.7 (1.2)
R_{sym}[‡] (%)*	9.6 (55.2)	15.1 (63.0)
R[§] (%)*	20.9 (27.1)	22.1 (26.9)
Free R[II] (%)*	26.8 (33.2)	27.7 (36.7)
Rms Bnd[¶] (Å)	0.009	0.013
Rms Ang[¶] (^{o})	1.25	1.44
Ramachandran plot (%)		
Residues in most favored regions	97.3	96.4
Residues in add.l allowed regions	2.7	3.6
Residues in generously allowed regions	0.0	0.0
Residues in disallowed regions	0.0	0.0

[‡] $R_{sym} = \Sigma\Sigma_j | I_j - \langle I \rangle | / \Sigma\Sigma_j | I_j |$.

[§] $R = \Sigma | | F_{obs} | - | F_{calc} | | / \Sigma | F_{obs} |$.

[II]Free R calculated against 7.2 and 7.0% of the reflections removed at random for $P2_1$ and $P6_322$ crystals, respectively.

[¶]Root mean square deviations from bond and angle restraints.

*Numbers in parentheses correspond to the highest resolution shell: 2.53 – 2.40 Å and 3.16 – 3.00 Å for data collection, and 2.46 – 2.40 Å and 3.08 – 3.00 Å for refinement.

As mentioned in the experimental section for the ICP-OES measurements, there is an internally coordinated Ni-ion for every MBPPhen1 monomer not involved in any interprotein contacts. This Ni ion is coordinated to the N-terminal amine and the carbonyl oxygen of Ala1, and Asp39. Its identity was established in an earlier study on a similar Ni-mediated superstructure (Ni_2:MBPC-1_3) using anomalous diffraction data collected at the Ni K-edge.[39] The average B-factors for these Ni ions (66 $Å^2$ in $P2_1$ crystals, 52 $Å^2$ in $P6_322$ crystals) are significantly higher than those of the interfacial Ni ions (32 $Å^2$ in $P2_1$ crystals, 31 $Å^2$ in $P6_322$ crystals) or the protein atoms (41 $Å^2$ in $P2_1$ crystals, 43 $Å^2$ in $P6_322$ crystals), indicating that they are weakly bound.

IR spectroscopy

Single crystals (~ 15) of Ni_3:MBPPhen1_3 grown from a 25% PEG 2000, 100 mM Tris (pH 8.5), 100 mM $MgCl_2$, 3.3 mM $NiCl_2$ solution were soaked in 100 µL of the the mother liquor plus 20 µL of a 100 mM potassium cyanate (KNCO) solution to give a final concentration of 20 mM KNCO. As a control single crystals (~5) of "metal free" W59C-Phen, a mutant that was crystallized in the presence of EDTA, were soaked in 100 µL of their respective precipitation solution (2 M NH_4SO_4, 100 mM Tris (pH 8.5) with 3 mM EDTA) along with 20 mM KNCO. After 30 minutes of soaking, the crystals were cleaned off of excess KCNO. The crystals were then ground with KBr and compressed into a pellet. Infrared spectra were collected on a Bruker Equinox 55 FTIR spectrometer.

Chapter 2 is reproduced, in part, with permission from: R.J. Radford, F.A. Tezcan, *J. Am. Chem. Soc.*, 131 (2009) 9136-9137.

2.6 References

[1] E.N. Salgado, R.A. Lewis, S. Mossin, A.L. Rheingold, F.A. Tezcan, *Inorg. Chem.*, 48 (2009) 2726-2728.

[2] R. Cortes, M.I. Arriortua, T. Rojo, J.L. Mesa, X. Solans, D. Beltran, *Acta Cryst. C*, 44 (1988) 986-990.

[3] J. Kyte, Structure in Protein Chemistry, 2nd ed., Garland Science, New York, 2007.

[4] G. Ghirlanda, J.D. Lear, N.L. Ogihara, D. Eisenberg, W.F. DeGrado, *J. Mol. Biol.*, 319 (2002) 243-253.

[5] I. Andre, P. Bradley, C. Wang, D. Baker, *Proc. Natl. Acad. Sci. USA*, 104 (2007) 17656-17661.

[6] D. Grueninger, N. Treiber, M.O.P. Ziegler, J.W.A. Koetter, M.S. Schulze, G.E. Schulz, *Science*, 319 (2008) 206-209.

[7] E.N. Salgado, J. Faraone-Mennella, F.A. Tezcan, *J. Am. Chem. Soc.*, 129 (2007) 13374-13375.

[8] E.N. Salgado, R.A. Lewis, J. Faraone-Mennella, F.A. Tezcan, *J. Am. Chem. Soc.*, 130 (2008) 6082-6084.

[9] A. Sabatini, I. Bertini, *Inorg. Chem.*, 4 (1965) 959-&.

[10] Y. Lu, *Curr. Opin. Chem. Biol.*, 9 (2005) 118-126.

[11] D.F. Qi, C.M. Tann, D. Haring, M.D. DiStefano, *Chem. Rev.*, 101 (2001) 3081-3111.

[12] C.H.B. Chen, L. Milne, R. Landgraf, D.M. Perrin, D.S. Sigman, *ChemBioChem*, 2 (2001) 735-740.

[13] H. Irving, R.J.P. Williams, *Nature*, 162 (1948) 746-747.

[14] R.P. Cheng, S.L. Fisher, B. Imperiali, *J. Am. Chem. Soc.*, 118 (1996) 11349-11356.

[15] K.A. Connors, Binding constants : the measurement of molecular complex stability, Wiley, New York, 1987.

[16] F.H. Arnold, J.-J. Zhang, *Trends Biotechnol.*, 12 (1994) 189-192.

[17] D.M. Lawson, P.J. Artymiuk, S.J. Yewdall, J.M.A. Smith, J.C. Livingstone, A. Treffry, A. Luzzago, S. Levi, P. Arosio, G. Cesareni, C.D. Thomas, W.V. Shaw, P.M. Harrison, *Nature*, 349 (1991) 541-544.

[18] R. Golmohammadi, K. Valegard, K. Fridborg, L. Liljas, *J. Mol. Biol.*, 234 (1993) 620-639.

[19] J.R. Su, D.J. Xu, *Acta Cryst. E*, 61 (2005) M1738-M1740.

[20] C.F. Ding, Y.F. Miao, B.Q. Tian, X.M. Li, S.S. Zhang, *Acta Cryst. E*, 62 (2006) M1062-M1063.

[21] M. Eddaoudi, J. Kim, N. Rosi, D. Vodak, J. Wachter, M. O'Keeffe, O.M. Yaghi, 295 (2002) 469-472.

[22] J.R. Li, R.J. Kuppler, H.C. Zhou, 38 (2009) 1477-1504.

[23] K.K. Tanabe, S.M. Cohen, *Inorg Chem*, 49 (2010) 6766-6774.

[24] The only other His residue (102) in this variant is coordinated to the heme iron in the protein core and thus unable to participate in PPI's.

[25] (a) Ghadiri, M. R.; Soares, C.; Choi, C., J. Am. Chem. Soc. 1992, 114, 4000-4002. (b) Lieberman, M.; Sasaki, T., J. Am. Chem. Soc. 1992, 113, 1470-1471. (c) Mutz, M. W.; McLendon, G. L.; Wishart, J. F.; Gaillard, E. R.; Corin, A. F., Proc. Natl. Acad. Sci. U. S. A. 1996, 93, 9521-9526.

[26] H.K. Privett, C.J. Reedy, M.L. Kennedy, B.R. Gibney, *J. Am. Chem. Soc.*, 124 (2002) 6828-6829.

[27] J.T. Faraone-Mennella, F. A.; Gray, H. B.; Winkler, J. R., *Biochemistry*, 45 (2006) 10504-10511.

[28] M.T.-M. Braun, L. , *Proc. Natl. Acad. Sci. USA*, 101 (2004) 12830-12835.

[29] F.N. Castellano, J.D. Dattelbaum, J.R. Lakowicz, *Anal. Biochem.*, 255 (1998) 165-170.

[30] Y. Nozaki, *Methods Enzymol.*, 26 (1972) 43-50.

[31] N.C. Pace, B.A. Shirley, J.A. Thomson, in: T.F. Creighton (Ed.) Protein Structure: A Practical Approach, IRL Press, Oxford, 1990, pp. 311-330.

[32] J. Faraone-Mennella, F.A. Tezcan, H.B. Gray, J.R. Winkler, *Biochemistry*, 45 (2006) 10504-10511.

[33] F.H. Stootman, D.M. Fisher, A. Rodger, J.R. Aldrich-Wright, *Analyst*, 131 (2006) 1145-1151.

[34] I.P. Evans, A. Spencer, G. Wilkinson, *J. Chem. Soc. Dalton Trans.*, (1973) 204-209.

[35] P. Schuck, *Biophys. J.* , 78 (2000) 1606-1619.

[36] Calculated using Advanced Chemistry Development (ACD/Labs) Software V8.14 for Solaris (© 1994-2009 ACD/Labs).

[37] J.G. de la Torre, M.L. Huertas, B. Carrasco, *Biophys. J.*, 78 (2000) 719-730.

[38] W.L. DeLano, The PYMOL Molecular Graphics System (http://www.pymol.org), 2003.

[39] E.N. Salgado, X.I. Ambroggio, J.D. Brodin, R.A. Lewis, B. Kuhlman, F.A. Tezcan, *Proc. Natl. Acad. Sci. USA*, 107 (2010) 1827-1832.

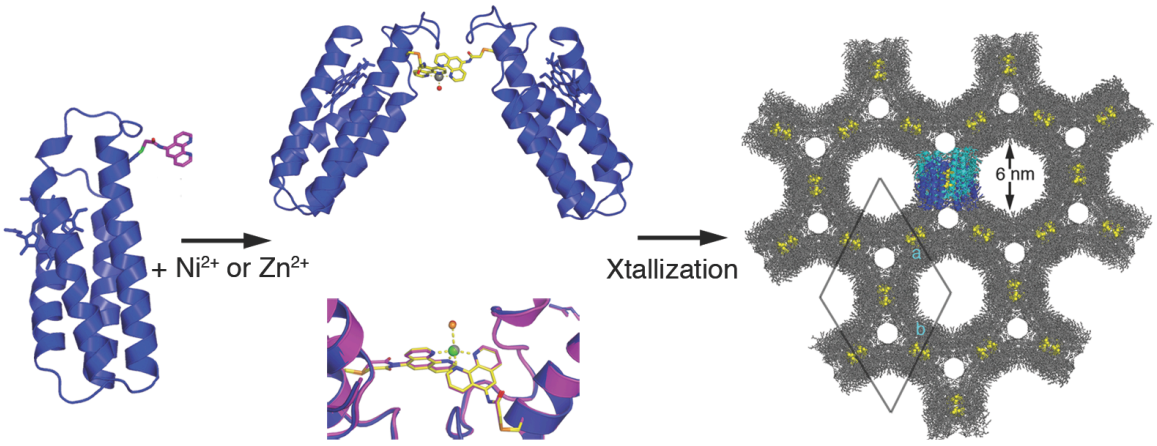
[40] CCP4, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 50 (1994).

[41] G. Murshudov, A. Vagin, E. Dodson, *Acta Cryst.*, D53 (1996) 240-255.

[42] D.E. McRee, *J. Mol. Graphics*, (1992) 44-46.

[43] R.A. Laskowski, Macarthur, M. W., Moss, D. S., and Thornton, J. M., *J. Appl. Crystallogr.*, 26 (1993) 283-291.

Chapter 3: Porous Protein Frameworks with Unsaturated Metal Centers in Sterically Encumbered Coordination Sites



3.1 Abstract

Described is an engineered metal-binding protein, MBPPhen2, which has the ability to form a porous crystalline framework featuring coordinatively unsaturated Zn- and Ni-centers

3.2 Introduction

To broaden the scope of Metal-Directed Protein Self-Assembly (MDPSA), in chapter 2, the incorporation of non-natural metal chelates onto the surface of a cyt cb_{562} variant, was reported. The resulting construct, MBPPhen1 (Figure 3.1), with a 1,10'-phenathroline (Phen) derivative attached to an engineered surface Cys (C59), was found to form Ni^{2+}-induced trimers, whereby the interfacial Ni centers were coordinated to a single Phen from one protomer and a histidine (H77) from a second, with two open coordination sites filled by solvent molecules (Figure 3.2). The open Ni coordination was a consequence of the Phen moiety being buried in a surface crevice under the overhang formed by the 50's loop, which prevented the formation of the saturated $Ni(MBPPhen1)_3^{2+}$ complex. Significantly, the open, trimeric Ni_3:$MBPPhen1_3$ architectures were found to stack up in the crystal lattice, giving rise to porous 3-D frameworks.

This work highlighted the potential advantages of using proteins as building blocks for functional materials. In fact, Nature predominantly uses amino acids to construct cellular machinery and advanced materials due to the structural and functional diversity these 20 building blocks provide, which has motivated chemists to follow suit. On the peptide level, there have been impressive engineering efforts over the last two decades. These efforts produced, among others, protein-like assemblies containing inorganic functional groups [1], materials with diverse applications in tissue engineering [2], nanoparticle assembly [3], and molecular electronics [4]. Yet, the mastery over the self-assembly of peptide-based structures has not yet reached a level where discrete,

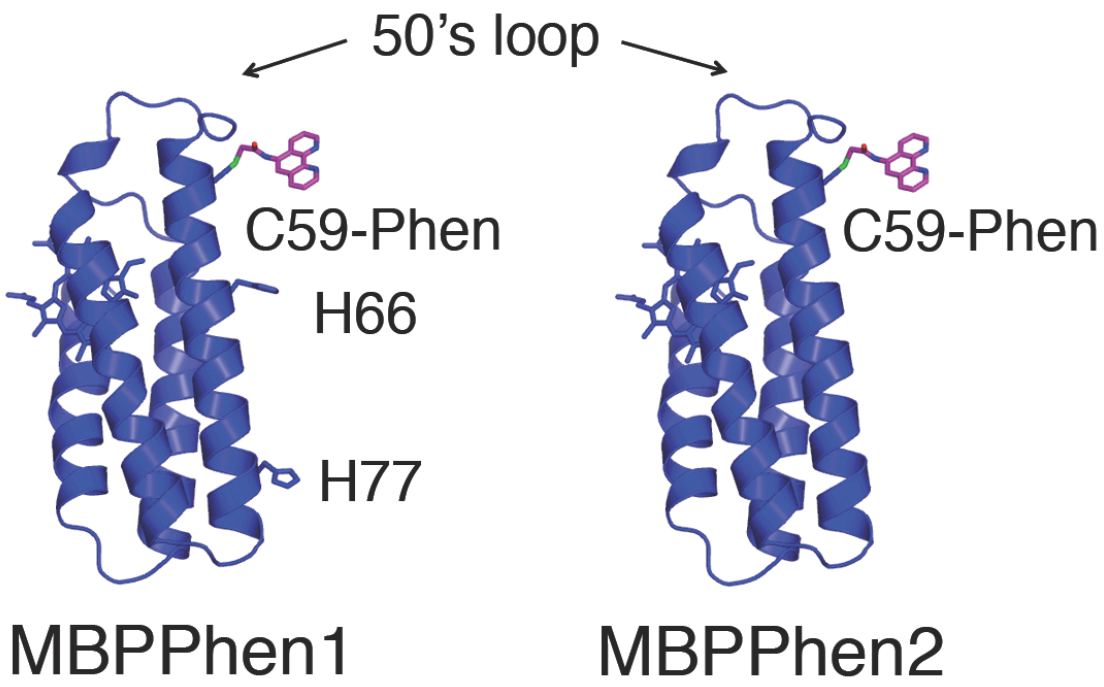


Figure 3.1. Cartoon representation of MBPPhen1 and MBPPhen2.

multi-dimensional architectures can be easily engineered beyond one-dimensional, fibrillar structures, owing to the fact that peptide building blocks are flexible and prone to heterogeneous aggregation [5]. However, these drawbacks could be circumvented if rigid, folded proteins were instead used as the primary building blocks. Indeed, the majority of biological machines and scaffolds are multi-protein architectures. The structural order, and the chemical functionalizability, of such multi-protein assemblies as virus capsids, protein cages and layers are now being widely exploited to generate hybrid materials towards light harvesting [6], catalysis [7], imaging [8], nanoparticle templating [9], and drug delivery [10].

In parallel, nano/meso-porous crystalline protein frameworks are showing promise towards applications such as molecular separation [11] and heterogeneous catalysis [12] in a similar fashion as zeolites and metal-organic frameworks (MOFs). Describe here is an in-depth characterization of the MBPPhen series of cyt cb_{562} variants, which form porous architectures and feature coordinatively unsaturated metal centers due to unforeseen interactions between the protein's surface and a non-natural metal chelate.

3.3 Results and Discussion

The goal of this study was to understand 1) if there is a thermodynamic bias towards the burial of Phen under the 50's loop in the MBPPhen1 construct (i.e., if this surface feature can be reproducibly used as steric bulk in analogy to smaller, sterically encumbered ligand frameworks), and

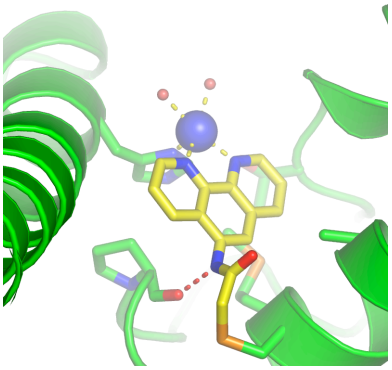


Figure 3.2. Inner-sphere coordination environment in the trimeric Ni:MBPPhen1$_3$ complex, highlighting the coordinatively unsaturated Ni^{2+} center.

2) if this burial can yield alternative protein oligomers/frameworks with unsaturated metal centers. To this end, a variant of MBPPhen1 (MBPPhen2, Figure 3.1) that lacks other likely surface ligands on Helix3 including H77 and H66, was made.

An important starting point was to determine whether the Phen burial would occur in the absence of metal coordination. MBP-Phen2 was thus crystallized in the presence of EDTA, and its structure determined at 2.1Å resolution (PDB ID: 3NMI). The asymmetric unit of the C_2 symmetric lattice contains six MBPPhen2 monomers arranged in three pairs (Figure 3.3). Two of the pairs are identical to each another and feature a head-to-head alignment of monomers, where the Phen moieties form π-stacking interactions with one another while still buried under the 50's loop (Figure 3.4a). Significantly, the conformation of these 4 Phen groups in the asymmetric unit is identical to that observed in the Ni_3:MBPPhen1$_3$ (Figure 3.4b): the aromatic moiety packs tightly into the cavity lined by the sidechains of T44, P53, M58, F61 and A62, and the amide nitrogen of the linker between C59 and Phen forms a H-bond with the backbone carbonyl of P53. The remaining pair of MBPPhen2 molecules also forms a head-to-head dimer, but in contrast,

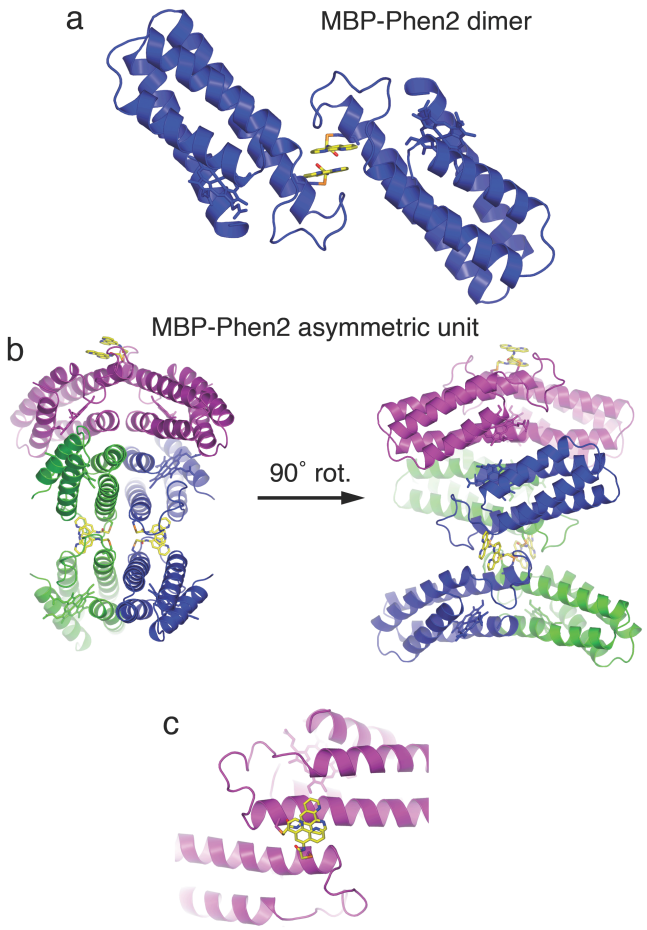


Figure 3.3. Various views of the metal free MBPPhen2 crystal structure. (a) The head-to-head aligned pair of MBPPhen molecules the feature π-stacked Phen groups, as shown in detail in Figure 2 in the Main Text. (b) Contents of the asymmetric unit in MBPPhen2 crystals. Each MBPPhen2 pair, whose Phen groups are found in association, are colored alike. (c) Close-up view of the two Phen groups found in the "out" conformation.

their Phen groups are now observed outside of the surface crevice. One Phen group is found pointing away from the 50's loop lying flat against the Helix3 surface, whereas the other one is significantly disordered as suggested by a weak corresponding electron density but still forming some π-stacking interactions with the former.

While the crystal structure of metal-free MBPPhen1 indicates a preference of the Phen moiety to be buried under the 50's loop, it also shows that it can explore other "out"

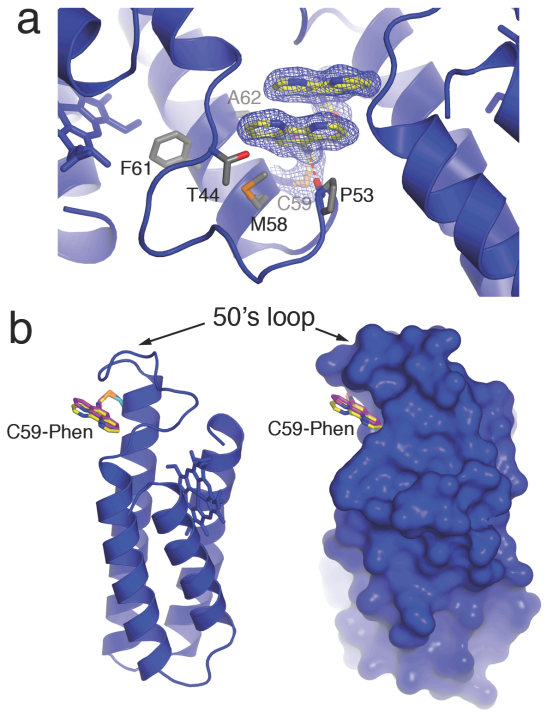


Figure 3.4. (a) Crystal structure of metal-free MBPPhen2, highlighting the Phen moiety buried under the 50's loop as well as its p-stacking interactions with a Phen group from another MBPPhen2 molecule in the asymmetric unit. The $2F_o–F_c$ map is contoured at 1.5σ. (b) Ribbon and surface representation of metal-free MBPPhen2, showing that its Phen conformation (yellow) is the same as that observed in the Ni_3:MBPPhen1$_3$ structure (magenta).

conformations. To gain further insight into the energetics of Phen-protein surface interactions, alchemical free energy calculations were carried out using the thermodynamic cycle in Figure 3.12. These calculations indicate that the buried Phen conformation (observed) is 4.2 ± 1.3 kcal/mol more favorable than an extended and fully solvent-exposed conformation (modeled).

We next examined how the steric encumbrance around the Phen group in MBPPhen2 affects its coordination behavior. Specifically, the solution oligomerization state of MBPPhen2 in the presence of Ni^{2+}, which could be expected to give a trimeric species for an exposed Phen group, was investigated. Sedimentation velocity (SV)

experiments with MBPPhen2 in the presence of a 1/3 equivalent of Ni^{2+} to drive the formation of the $Ni(MBPPhen)_3$ instead indicate that the predominant species in solution is a dimer (Figure 3.5), even at a protein/Ni concentration of 600/200 μM.

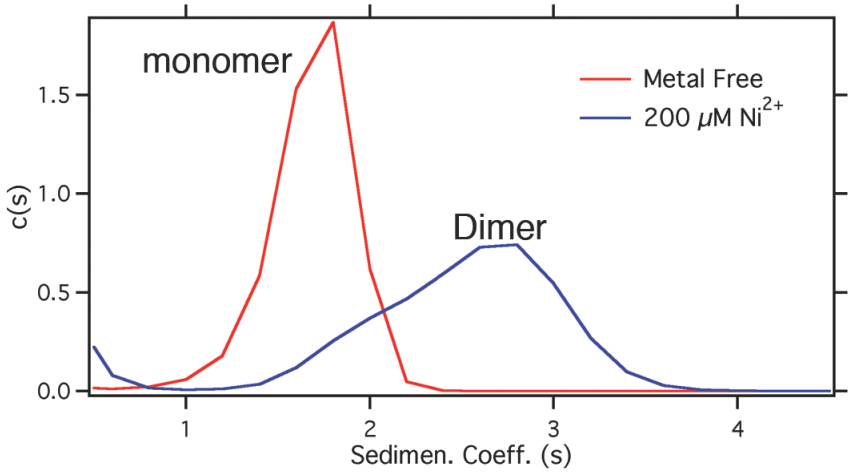


Figure 3.5. Sedimentation velocity distribution of MBPPhen2 in the absence of metal and presence of Ni^{2+} ([MBPPhen2] = 600 μM, [Ni^{2+}] = 200 μM).

The solution oligomerization behavior of MBPPhen2 is paralleled in the solid state. Single crystals of the Ni^{2+} and Zn^{2+} adducts of MBPPhen2 were obtained and determined their structures at 3.1 and 2.8-Å resolution, respectively (PDB IDs: 3NMJ and 3NMK). Remarkably, the hexagonal ($P6_5$ space group) crystals of the Ni^{2+} and Zn^{2+} complexes are entirely isostructural despite the distinct stereochemical preferences of the two ions. The Ni^{2+} and Zn^{2+}-induced MBPPhen2 dimers are shown in Figure 3.6a. These C_2-symmetric V-shaped dimers are mediated solely by Ni/Zn binding to Phen groups, in which one of the Phen's is buried, while the other in an extended conformation. Ni and Zn are found in identical distorted trigonal bipyramidal geometries (within the low resolution of the structures), which are completed by a water molecule (Figure 3.6b). In the asymmetric unit, pairs of Ni- and Zn-induced dimers are further interlaced, yielding

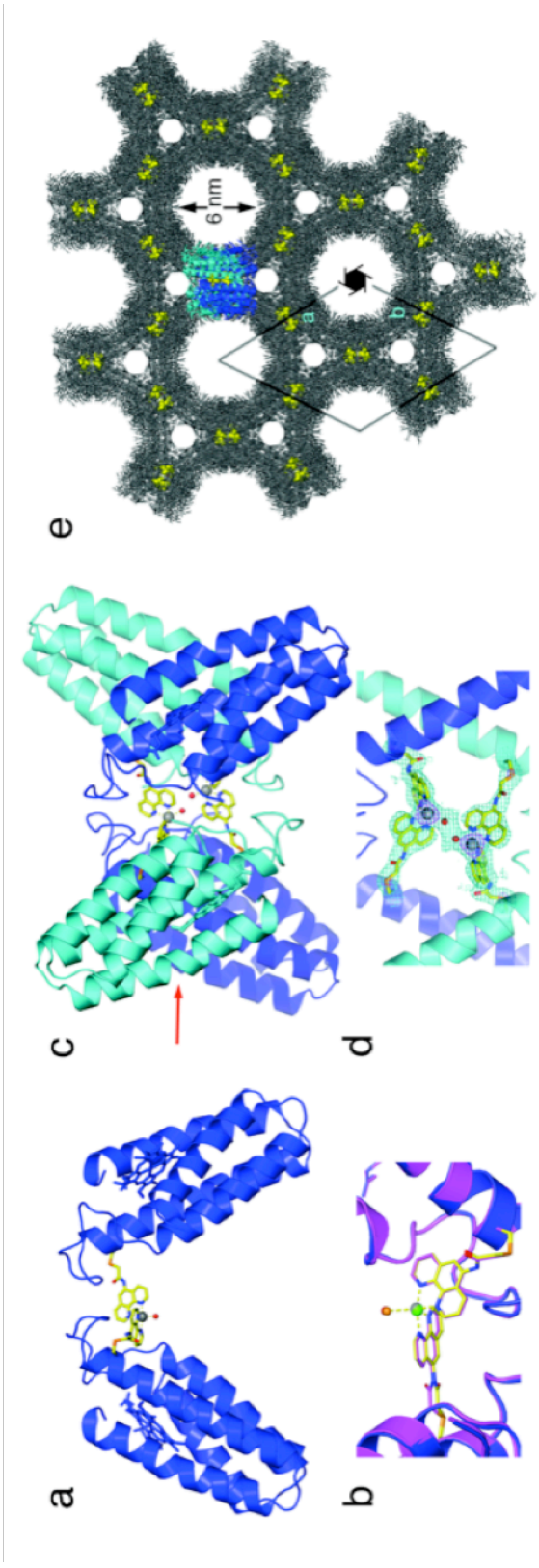


Figure 3.6. (a) Zn:MPBPhen$_2$ crystal structure. The analogous Ni complex is isostructural with the Zn complex (rms = 0.293 Å over all Cα's . (b) Close-up view of the metal coordination environment in Zn (blue) and Ni(magenta) complexes. (c) Dimer of Zn:MPBPhen$_2$ dimers in the asymmetric unit. (d) Close-up view of the metal coordination environments in the asymmetric unit of Zn:MPBPhen$_2$ crystals, showing the proximity of the neighboring metal centers. The $2F_o$–F_c map is contoured at 1.2 s. (e) Zn:MPBPhen$_2$ crystal lattice viewed down the *c*-axis. The highlighted dimer of dimers is viewed along the direction indicated with a red arrow in (c). Phen groups are shown as yellow spheres to highlight the relative positions of metal centers in the lattice

D_2-symmetric tetramers that hold the metal centers in close proximity (7 Å between metal centers, 3 Å between coordinated water molecules, Figure 3.6c, d). Clearly, the combination of the steric bulk around the Phen groups and lattice packing interactions are ultimately responsible for the formation of this particular supramolecular arrangement, and force Ni^{2+} and Zn^{2+} to adopt the same coordination geometry. The enforcement of identical coordination geometries on metal ions with distinct stereochemical preferences is typically reserved for rigid, bulky ligand platforms [13], which include highly evolved protein scaffolds with internal coordination sites. The fact that this can be achieved on the surface of a protein highlights the potential of crystalline protein frameworks being used as platforms for metal-based catalysis, as elegantly demonstrated in small-molecule crystal systems [14].

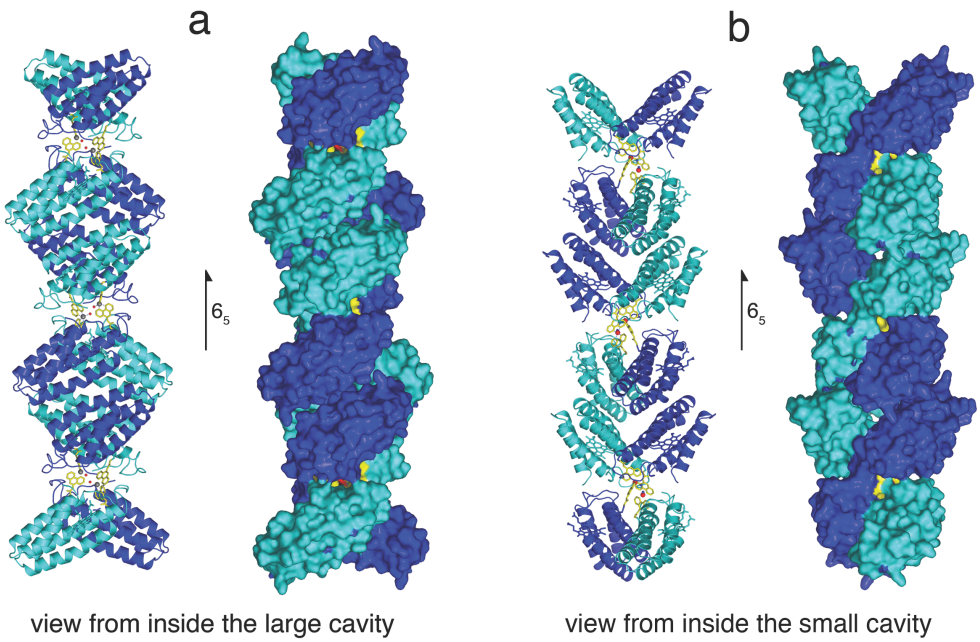


Figure 3.7. Columns of Zn:MBPPhen2$_2$ molecules that form the lining of the hexagonal pores in the crystal lattice, shown in ribbon and surface representations. The orientation of each column as seen from inside (a) the large pores (see Figure 6) and (b) the small pores.

Further in support for such potential applications, the crystals of Ni:MPBPhen2$_2$ and Zn:MPBPhen2$_2$ are highly porous (Figure 3.6e), with two sets of hexagonal channels 6 and 2 nm wide and an overall solvent content of 61%. A single asymmetric unit that contains a dimer of MBPPhen2 dimers, is highlighted within this lattice. These tetrameric units are arranged into columns using helix-helix packing interactions and form the lining of the two hexagonal channels (Figure 3.7). A surface representation of these MBPPhen2 columns indicates that the interfacial metal centers should be accessible from the solvent channels, yet protected enough to display selectivity. Investigation into whether the same isostructural units can be obtained in the presence of other metals, and testing the reactivity of Ni:MPBPhen2$_2$ and Zn:MPBPhen2$_2$ single crystals, is currently underway. As a prelude to these studies, these crystals were subjected to chemical crosslinking with 6% (v/v) glutaraldehyde for 30 min. Whereas unmodified crystals immediately dissolve upon transfer from the precipitation solution (30% PEG400, 0.2 M ammonium sulfate) into water, the crosslinked crystals are indefinitely stable in water, even after being kept at 98º C for 10 min, or in a 50% acetonitrile/water mixture (Figures 3.8-11).

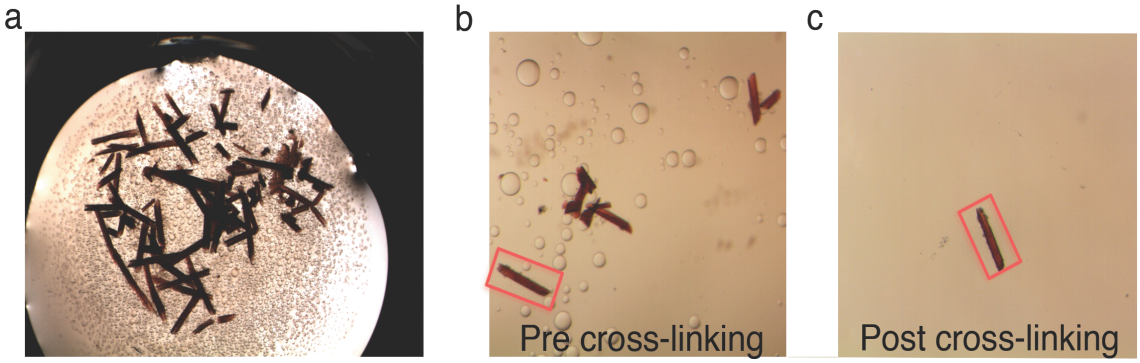


Figure 3.8. Crosslinking of Ni:MBPPhen2$_2$ crystals with glutaraldehyde. (a) Ni:MBPPhen2$_2$ crystals in the crystallization well. (b) Ni:MBPPhen2$_2$ crystals in precipitation solutions prior to crosslinking. (c) Sample Ni:MBPPhen2$_2$ crystal in water after cross-linking. The crosslinked crystal shown in (c) is highlighted with a red box in (b) for comparison.

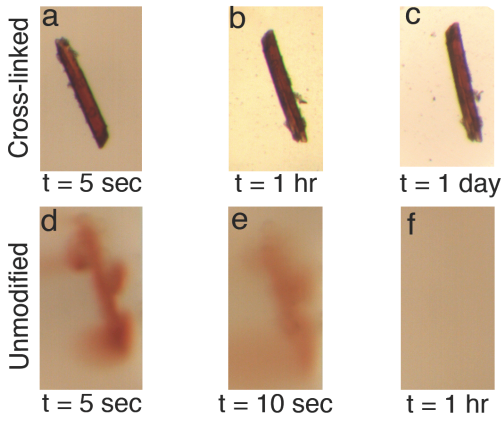


Figure 3.9. Comparing stability of crosslinked and unmodified Ni:MBPPhen2$_2$ crystals in H_2O (no precipitant). Crosslinked Ni:MBPPhen2$_2$ crystals reveal no noticeable deterioration after (a) 5 sec (b) 1 hr, or (c) 1 day. Conversely, unmodified Ni:MBPPhen2$_2$ crystals begin dissolving almost immediately under the same conditions ((d) 5 sec, (e) 10 sec) and are completely dissolved after a period of (f) 1 hr.

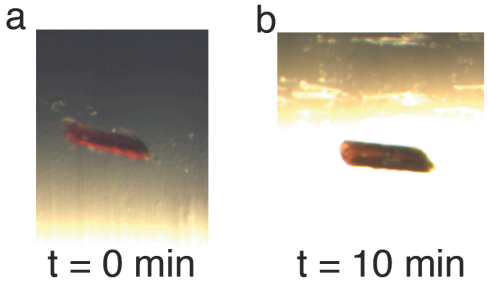


Figure 3.10. Monitoring the effect of high temperatures on crosslinked Ni:MBPPhen2$_2$ crystals. Ni:MBP-Phen2 crystals were placed in a thin walled Eppendorf tube and held at 98° C for 10 min. (a) prior to heating, (b) after heating.

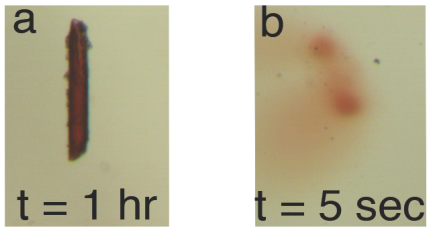


Figure 3.11. Effect of organic solvent on cross-linked and unmodified Ni:MBPPhen2$_2$ crystals. (a) No visible changes are seen after exposure of crosslinked Ni:MBPPhen2$_2$ crystals to 50% (v/v) H_2O/Acetonitrile. (b) In contrast, unmodified Ni:MBPPhen2$_2$ crystals immediately dissolve upon exposure to the same solvent.

3.4 Conclusions

Efforts to control protein self-assembly though metal coordination, have benefited from the deconstruction of proteins into simple, rigid ligand platforms for metal coordination. The study undertaken here has revealed strong parallels between the coordination complexes/frameworks of small organic molecules and those of proteins, suggesting that the approach of using proteins as metal ligands may have utilities beyond simply controlling their self-assembly.

3.5 Experimental Section

Site Directed Mutagenesis and Protein Expression/Purification/Characterization

Site directed mutagenesis, protein expression and purification were carried out as previously described in chapter 2. The sequence of the final plasmid, containing the mutations W59C R62A H63A D66H K77H (hereafter referred to as MBP2), was sequenced at the UCSD Moore's Cancer Center.

Synthesis of Iodoacetamido-1,10-Phenanthroline (IA-Phen)

IA-Phen was synthesized as described in chapter 2.

Functionalization of MBP2 with IA-Phen

The mutant MBP2 protein was labeled with IA-Phen as previously described in Chapter 2. The purity of functionalized protein was determined by MALDI mass

spectrometry and SDS-PAGE electrophoresis. Labeling yield: >90%. MW observed = 12300 amu, MW expected = 12298 amu).

Crystallography

General

All crystals were obtained by sitting drop vapor diffusion using a Hampton 24-well sitting drop plate. All reagents were of the highest possible purity commercially available. Solutions and reagents were filtered using a 0.2 micron cellulose membrane filter prior to use. Unless otherwise noted, the crystals to be used for diffraction experiments were exchanged into a solution containing 20% glycerol as a cryoprotectant. X-ray diffraction data for all crystals were collected at 100 K at Stanford Synchrotron Radiation Laboratory (BL 9-2 or BL 7-1) using 0.98 Å radiation. The data were processed using MOSFLM and SCALA.[15]

The initial molecular replacement solution for all crystals was determined by PHASER [16], using the cyt cb_{562} structure (PDB ID 2BC5) [17] as the search model. The search model for each crystal form did not contain the heme or the Phen prosthetic groups; the observation of strong positive F_o-F_c density at expected positions for these groups confirmed the correct placement of protein monomers. The topology and parameter files for the phenanthroline group were obtained using the Dundee ProDrg Server (http://davapc1.bioch.dundee.ac.uk/prodrg/index.html). Rigid-body, positional and thermal refinement with REFMAC [15, 18], along with manual rebuilding, and water placement with COOT [19], produced the final models. All figures were produced with PYMOL [20].

MBPPhen2

MBPPhen2 was crystallized at 25° C with a premade Hampton screen precipitation solution (HR2-144, #24) consisting of 2.8 M sodium acetate trihydrate (pH 7.0). The drop consisted of 2 μL of protein (2.6 mM in 20 mM Tris (pH 7) and 1 mM EDTA) and 1 μL of the precipitation solution. Crystals appeared within 2 months reaching a maximum size of ~ 300 μm × 300 μm × 500 μm. The crystals used in diffraction experiments were frozen directly in liquid N_2 without addition of a cryoprotectant. The structure was determined to 2.01-Å resolution (PDB: 3NMI) (Table 3.1).

Ni:MBP-Phen2

Ni:MBPPhen2$_2$ was crystallized at 25° C with a precipitation solution consisting of 0.1 M Tris (pH 8.5), 0.2 M ammonium sulfate, 30% PEG 400 and 6.2 mM nickel sulfate. The drop consisted of 2 μL of protein (3.1 mM in 20 mM Tris (pH 7)) and 1 μL of the precipitation solution. Crystals appeared within 1 months reaching a maximum size of ~ 100 μm × 200 μm × 200 μm. The crystals used in diffraction experiments were exchanged into cryoprotecting solution and frozen in liquid N_2. The structure was determined to 3.10-Å resolution (PDB: 3NMJ) (Table 3.1).

Zn:MBP-Phen2

Zn:MBPPhen2$_2$ was crystallized at 25° C with precipitation solution consisting of 0.1 M Tris (pH 8.5), 0.2 M ammonium sulfate and 30% PEG 400 and 12.4 mM zinc chloride. The drop consisted of 2 μL of protein (3.1 mM in 20 mM Tris (pH 7)) and 1 μL of the precipitation solution. Crystals appeared within 1 months reaching a maximum size

of ~ 100 μm × 200 μm × 200 μm. The crystals used in diffraction experiments were exchanged into cryoprotecting solution and frozen in liquid N_2. The structure was determined to 2.85-Å resolution (PDB: 3NMK) (Table 3.1).

Crosslinking Ni:MBPPhen2$_2$ Crystals with Glutaraldehyde

Ni:MBP-Phen2 crystals (~6) were added to a 12 μL of the precipitation solution (see above) and 4 μL of a 25% gultaraldehyde grade II solution (Sigma). The crystals were allowed to crosslink for 30 min, after which they were washed 3 times by placing them in successive 20 μL solutions of fresh precipitant solution. Although the unit cell was not determined, the integrity of the crystals was continuously checked by their ability to polarize light.

Table 3.1. X-ray data collection and refinement statistics for: MBPPhen2, Ni:MBPPhen2 and Zn:MBPPhen2.

X-ray data collection and refinement statistics			
	MBP-Phen2	**Ni:MBP-Phen2**	**Zn:MBP-Phen2**
Data collection location	SSRL BL 9-2	SSRL BL 7-1	SSRL BL 7-1
Unit cell dimensions (Å)	88.29 × 91.79 × 130.64 α = 90° ß= 108°, γ =90°	99.29 × 99.29 × 109.32 α =ß = 90°, γ = 120.0°	99.46 × 99.46 × 109.54 α =ß = 90°, γ = 120.0°
Symmetry group	*C*2	$P6_5$	$P6_5$
Resolution (Å)	2.01	3.10	2.80
X-ray wavelength (Å)	0.979	0.979	0.979
Number of unique reflections	65484	11200	15237
Redundancy	4.4	9.9	7.7
Completeness (%)*	99.7	100.0	99.9
$\langle I / \sigma I \rangle$*	1.7	1.4	1.9
R_{symm}[‡]* (%)	23.7	43.3	30.1
R_{work}[§]/R_{free}[II] (%)	19.8 / 23.5	19.3 / 23.8	18.8 / 23.3
Contents of asu			
Protein monomers	6	4	4
Ligands/ions	6 / 5	4 / 2	4 / 2
Water	285	4	32
R.m.s deviations[¶]			
Bond lengths (Å)	0.009	0.006	0.007
Bond angles (°)	0.985	0.742	0.814

*Numbers in parentheses correspond to the highest resolution shell: (2.120 - 2.010 Å) for EDTA:MBPPhen2, (3.270 – 3.100 Å) for Ni:MBPPhen2, and (2.950 – 2.800 Å) for Zn:MBPPhen2, respectively

[‡] $R_{sym} = \Sigma\Sigma_j | I_j - \langle I \rangle | / \Sigma\Sigma_j | I_j |$.

[§] $R = \Sigma | |F_{obs}| - |F_{calc}| | / \Sigma |F_{obs}|$

[II]Free R calculated against 7 of the reflections removed at random.

[¶]Root mean square deviations from bond and angle restraints.

Sedimentation Velocity Experiments

Sedimentation velocity (SV) experiments were performed in order to determine the solution-state oligomerization behavior of MBPPhen2. All SV samples were prepared in 20 mM Tris buffer (pH 7). Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor at 41,000 rpm for a total of 250 scans/sample at a wavelength of 664 nm (600 μM protein).

All sedimentation velocity data were processed using SEDFIT [21]. Buffer viscosity, buffer density, and protein partial specific volume values were calculated at 25° C with SEDNTERP (http://www.jphilo.mailway.com). Partial specific volume (vbar) for MBPPhen2 mutant was calculated to be 0.7360 mg/ml, assuming a partial specific volume of heme of 0.82 mg/ml and 0.71 mg/ml for the phenanthroline [22]. All data were processed using fixed values for buffer density (r) (0.99764 g/ml) and buffer viscosity (0.0089485 poise).

Computational Calculations and Simulations

Free Energy Calculation for Phen Burial

To determine the solvation free energy difference between the "in," stacked Phen conformation observed in MBPPhen2 complex and the more solvent exposed "out" conformation in the MBPPhen2 conformation, alchemical free energy calculations were performed using NAMD2.7β1. These calculations utilize unphysical intermediate states, in which the non-bonded potential energy terms for the Phen residue are linearly scaled to zero with the parameter λ, as represented in Figure 3.2. This decoupling [23] of the residue from its environment in both the "in" and "out" conformations allows

computation of a relative solvation free energy difference $\Delta\Delta G_{out \to in}$, as depicted in the thermodynamic cycle of Figure 3.2.

The decoupling steps for computation of ΔG_{in} and ΔG_{out} in Figure 3.2, were performed over 14 windows of λ (λ=0.0, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9,1.0), with electrostatic potential terms scaled to zero by window 7, and van der Waals potential terms annihilated by window 14. A softcore potential[24] was used for annihilation of the van der Waals potentials to improve convergence and accuracy of the calculation. For simulations of the complex at each window of λ, a 1 ns equilibration period precedes a 1 ns sampling period, during which values of $\left\langle \frac{dV}{d\lambda} \right\rangle_{\lambda}$ were computed every 5 timesteps for thermodynamic integration:[25]

$$\Delta G = \int_0^1 \left\langle \frac{dV}{d\lambda} \right\rangle_{\lambda} d\lambda$$

where V is the total potential energy of the system, and the average derivative is computed over the simulation time at each λ window. The integral is evaluated using cubic spline integration from λ=0, when the atoms of Phen are fully coupled, to λ=1, when the atoms do not interact with the rest of the system.

To restrict conformational sampling to the relevant conformational states of the Phen residue in the MBPPhen1 and MBPPhen2 complexes, a harmonic restraint potential $U_{COM} = k(r_{COM} - r_0)^2$ was applied to the Phen center of mass (COM), with a reference COM position r_0 and magnitude of the harmonic constant $k = \frac{3RT}{\left\langle \delta r_{COM}^2 \right\rangle}$ determined from

the average COM position and average fluctuation $\langle \delta r_{COM}^2 \rangle$ over 2 ns NPT (at T=300 K) simulations of both complexes. A relative free energy difference between these two conformations, $\Delta\Delta G_{out \to in}$ was found by subtracting ΔG_{in} from ΔG_{out}, as in Scheme SX. A bootstrapping re-sampling method was employed to obtain variances for the $\left\langle \frac{dV}{d\lambda} \right\rangle_{\lambda}$ values, which were then propagated in the integration to give the uncertainty reported for $\Delta\Delta G_{out \to in}$.

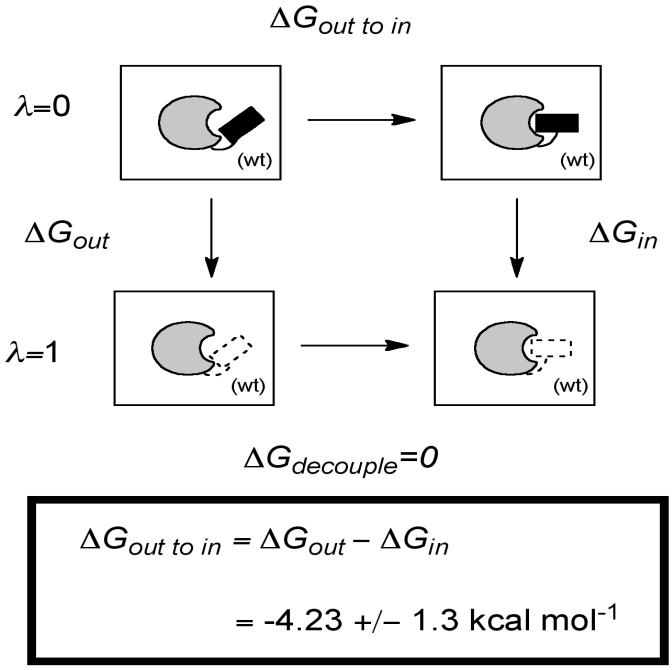


Figure 3.12. Cartoon representing the thermodynamic cycle used in computing the free energy difference between the two Phen residue conformations in MBPPhen2. At λ=1, the nonbonding potential terms of the Phen have been turned off, decoupling the residue from its environment.

Chapter 3 is reproduced, in part, with permission from: R.J. Radford, M. Lawrenz, P.C. Nguyen, J.A. McCammon, F.A. Tezcan, *Chem. Commun*., 47 (2011) 313-315.

3.6 References

[1] A. Lombardi, C.M. Summa, S. Geremia, L. Randaccio, V. Pavone, W.F. DeGrado, *Proc. Natl. Acad. Sci. USA*, 97 (2000) 6298-6305.

[2] T.C. Holmes, S. de Lacalle, X. Su, G.S. Liu, A. Rich, S.G. Zhang, *Proc. Natl. Acad. Sci. U. S. A.*, 97 (2000) 6728-6733.

[3] M.M. Stevens, N.T. Flynn, C. Wang, D.A. Tirrell, R. Langer, *Adv. Mater.*, 16 (2004) 915-918.

[4] M. Reches, E. Gazit, *Science*, 300 (2003) 625-627.

[5] E.H.C. Bromley, K. Channon, E. Moutevelis, D.N. Woolfson, *ACS Chem. Biol.*, 3 (2008) 38-50.

[6] R.A. Miller, A.D. Presley, M.B. Francis, *J. Am. Chem. Soc.*, 129 (2007) 3104-3109.

[7] S. Abe, K. Hirata, T. Ueno, K. Morino, N. Shimizu, M. Yamamoto, M. Takata, E. Yashima, Y. Watanabe, *J. Am. Chem. Soc.*, 131 (2009) 6958-6960.

[8] J.D. Lewis, G. Destito, A. Zijlstra, M.J. Gonzalez, J.P. Quigley, M. Manchester, H. Stuhlmann, *Nat. Med.*, 12 (2006) 354-360.

[9] J.C. Falkner, M.E. Turner, J.K. Bosworth, T.J. Trentler, J.E. Johnson, T.W. Lin, V.L. Colvin, *J. Am. Chem. Soc.*, 127 (2005) 5274-5275.

[10] M.L. Flenniken, L.O. Liepold, B.E. Crowley, D.A. Willits, M.J. Young, T. Douglas, *Chem. Commun.*, (2005) 447-449.

[11] L.Z. Vilenchik, J.P. Griffith, N. St Clair, M.A. Navia, A.L. Margolin, *J. Am. Chem. Soc.*, 120 (1998) 4290-4294.

[12] T. Koshiyama, N. Kawaba, T. Hikage, M. Shirai, Y. Miura, C.Y. Huang, K. Tanaka, Y. Watanabe, T. Ueno, *Bioconjugate Chem.*, 21 (2010) 264-269.

[13] E.E. Benson, A.L. Rheingold, C.P. Kubiak, *Inorg. Chem.*, 49 (2010) 1458-1464.

[14] Z. Huang, P.S. White, M. Brookhart, *Nature*, 465 (2010) 598-601.

[15] CCP4, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 50 (1994).

[16] A.J. McCoy, R.W. Grosse-Kunstleve, P.D. Adams, M.D. Winn, L.C. Storoni, R.J. Read, *J. Appl. Crystallogr.*, 40 (2007) 658-674.

[17] J.T. Faraone-Mennella, F. A.; Gray, H. B.; Winkler, J. R., *Biochemistry*, 45 (2006) 10504-10511.

[18] G. Murshudov, A. Vagin, E. Dodson, *Acta Cryst.*, D53 (1996) 240-255.

[19] P. Emsley, K. Cowtan, *Acta Cryst.*, D60 (2004) 2126-2132.

[20] W.L. DeLano, The PYMOL Molecular Graphics System (http://www.pymol.org), 2003.

[21] P. Schuck, *Biophys. Chem.*, 108 (2004) 187-200.

[22] Calculated using Advanced Chemistry Development (ACD/Labs) Software V8.14 for Solaris (© 1994-2009 ACD/Labs).

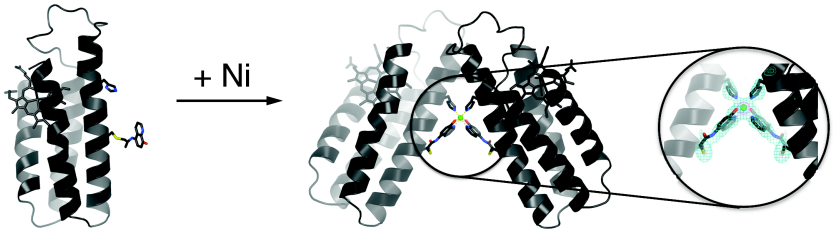
[23] M.K. Gilson, J.A. Given, B.L. Bush, J.A. McCammon, *Biophys. J.*, 72 (1997) 1047-1069.

[24] T.C. Beutler, A.E. Mark, R.C. van Schaik, P.R. Gerber, W.F. van Gunsteren, *Chem. Phys. Lett.*, 222 (1994) 529-539.

[25] J.G. Kirkwood, *J. Chem. Phys.*, 3 (1935) 300-313.

Chapter 4: Controlled Protein Dimerization through Hybrid Metal Coordination Motifs

4.1 Abstract



Protein homodimerization is the simplest form of oligomerization and is frequently utilized for the construction of functional biological assemblies and the regulation of cellular pathways. Despite its simplicity, dimerization still poses an enormous challenge for protein engineering and chemical manipulation, owing to the large molecular surfaces involved in this process. Reported here is the construction of a hybrid coordination motif – consisting of a natural (His) and a non-natural ligand (quinolate) – on the α-helical surface of cytochrome cb_{562}, which a) simultaneously binds divalent metals with high affinity, b) leads to a metal-induced increase in global protein stability, and importantly, c) enables the formation of a discrete protein dimer, whose shape is strictly dictated by the inner-sphere metal coordination geometry. The crystallographically-determined arrangement of metal-crosslinked α-helices closely approximate that of the DNA-binding domains of bZIP family transcription factors.

4.2 Introduction

Protein dimerization is an omnipresent process utilized for the construction of numerous functional biological assemblies and regulation of cellular pathways [3]. Given its broad biological significance, protein dimerization - and oligomerization in general - has been a subject of great fundamental interest [4], and represents a major target for protein engineering [5, 6] and chemical manipulation [7, 8]. Herein, it is shown that protein dimerization can be tightly controlled through a new class of high-affinity surface coordination motifs, which yield discrete, and biologically relevant, architectures dictated solely by metal binding, while simultaneously leading to the stabilization of the helical domains that they are installed on.

An important challenge in protein design is the engineering of protein interaction specificity, that is, the population of a single interaction geometry over other possible conformers, which often lie very close in energy. The previous two chapters (chapters 2 and 3), have detailed the expansion MDPSA to include non-natural ligands [9], which offer a far wider chemical spectrum than what is naturally available to control protein self-assembly through metal coordination [10-12]. Two cytochrome cb_{562} constructs (MBPPhen1 and MBPPhen2) both featuring a bidentate 1,10'-phenanthroline (Phen) functionality with a single-point surface attachment was observed to form porous frameworks upon the addition of metal ions. These particular supramolecular arrangements were ultimately enabled by the flexibility of the linker between Phen and the MBPPhen1/2 backbone, which allowed the Phen group to tuck into a hydrophobic pocket on the protein surface. While this example demonstrated that protein surface features can in principle be exploited - in analogy to synthetic ligand platforms - as steric

bulk [13], this strategy is not readily generalizable. It was thus envisioned that non-natural chelates may be combined with natural ligands incorporated elsewhere on the surface to form hybrid coordination motifs (HCMs), which would not only provide more rigid, high-denticity platforms for tighter metal binding and improved control of protein dimerization, but also be implemented more generally.

Because the α-helix is the most common secondary structure motif, it represents a particularly attractive platform for the incorporation of HCMs. Indeed, nature itself utilizes residues placed in *i*/*i*+3 and *i*/*i*+4 patterns quite regularly to construct stable metal binding sites [14]. In this study, an *i*/*i*+7 pattern - corresponding to a two-helix-turn separation – was employed to install a tridentate HCM on cyt cb_{562} with the idea that it would enforce protein dimerization upon binding a metal ion that prefers octahedral coordination (Figure 4.1.). A cyt cb_{562} variant (HQuin1) was thus constructed that features an HCM composed of His63 and an iodoacetamide-functionalized 5-amino-8-hydroxyquinoline (Quin) group that was covalently attached to Cys70. The Quin functionality was particularly chosen because it is (often) a monoanionic ligand with a high affinity for many metal ions [2], and importantly, its lack of internal symmetry (in contrast, for example, to Phen) can in principle be exploited to impose preferential dimerization geometries through metal coordination.

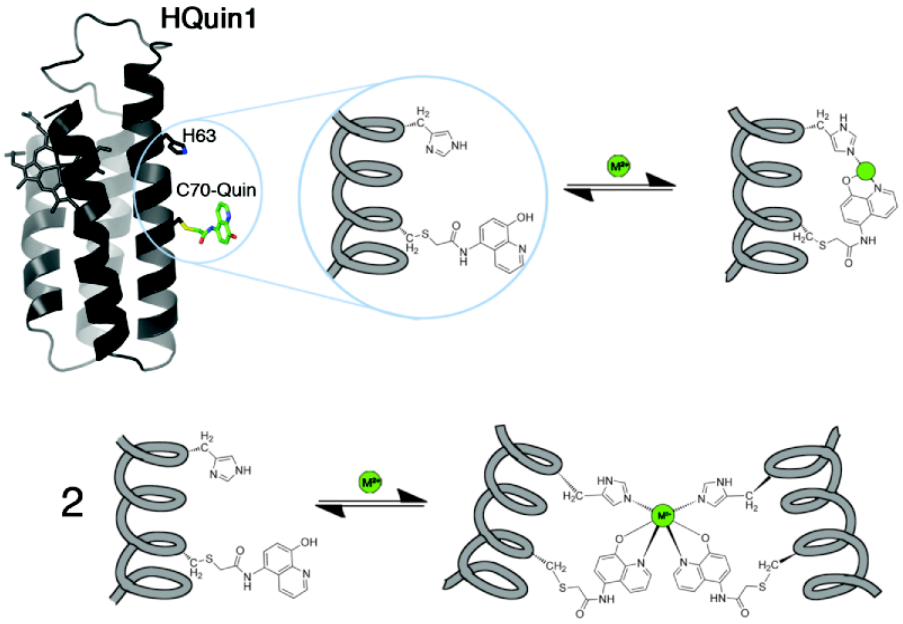


Figure 4.1. A cartoon schematic of the proposed mode of metal coordination, and dimerization, for the HQuin1 system.

4.3 Results and Discussion

Metal binding properties of HQuin1

The binding of HQuin1 to various divalent late-first-row transition metal ions (Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}), were investigated by taking advantage of the π-π* transition of Quin (λ_{max} = 244 nm) that undergoes a ~20-nm redshift upon metal coordination (Figure 4.1). Initial studies suggested that the affinity of HQuin1 for all of tested metal ions is too high (nanomolar or lower) to be reliably assessed by direct titrations; hence, EGTA was used as a competing ligand to determine dissociation constants (K_d) (Figure 4.2). As listed in Table 4.1, HQuin1 exhibits low K_d's that range from low nanomolar for Co^{2+} and Zn^{2+} to 85 fM for Cu^{2+} and follows the trend $K_{d\text{-}Co} > K_{d\text{-}Ni} << K_{d\text{-}Cu} >> K_{d\text{-}Zn}$,

roughly in line with the Irving-Williams series [15]. While the fact that the HQuin1 metal-binding affinities are several orders of magnitude higher than those for free Quin (Table 4.1) strongly suggests tridentate coordination, the H63A variant of HQuin1, (AQuin1), was prepared to ascertain the involvement of H63 in metal binding. The metal affinities for AQuin1 were found to be significantly lower than HQuin1 (Figure 4.3 and Table 4.1) and now in the range of those for free Quin, indicating that HQuin1 coordinates the tested metals in a tridentate fashion as planned. Evidently, the His-Quin HCM possesses sufficient internal flexibility to accommodate the various stereochemical preferences of Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+} coordination.

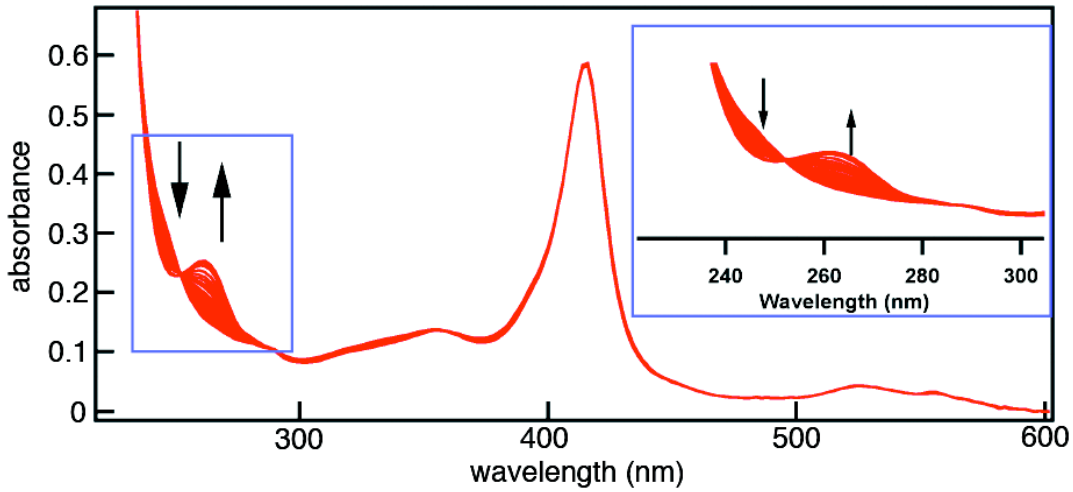


Figure 4.2. Representative changes in the absorption spectrum of HQuin1 upon M^{2+} binding. Shown here are a series of absorption spectra for a representative Zn^{2+} titration. (Inset) The observed changes in absorbance band of the Quin moiety upon metal binding.

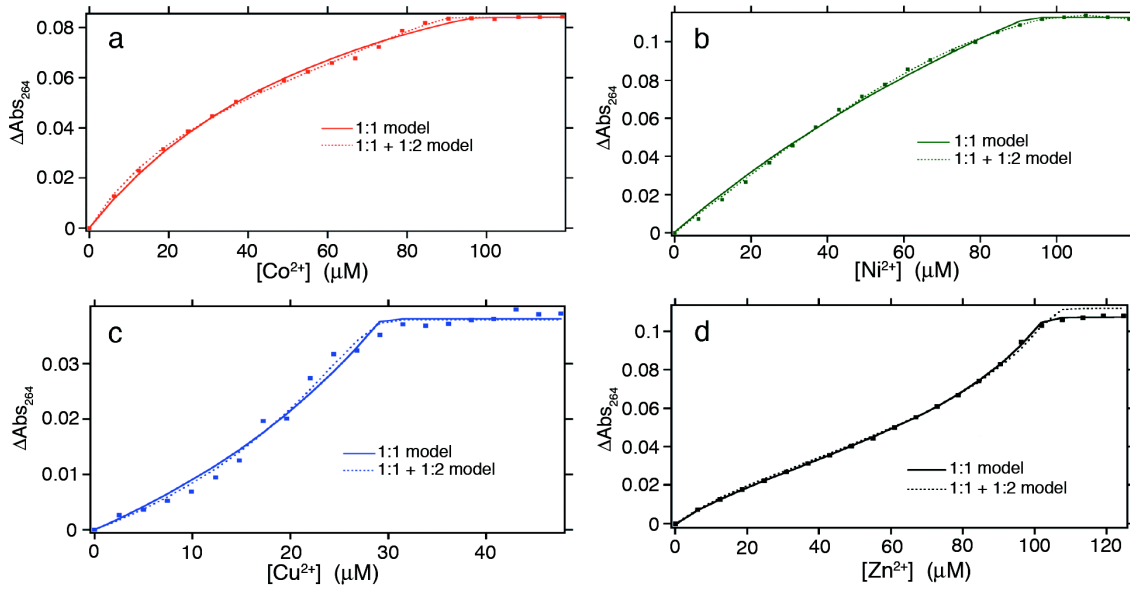


Figure 4.2. Metal binding titration data and fits for HQuin1 (3 – 5 μM) in the presence of EGTA (25-100 μM) as a competing ligand and various concentrations of (a) Co^{2+}, (b) $Ni^{2+,}$ (c) Cu^{2+}, (d) Zn^{2+}, as monitored by UV-vis spectroscopy. Regression analysis was performed using both a simple 1:1 (solid line) metal:protein binding model and a combined 1:1 and 1:2 (dotted line) model, the latter of which accounts for metal-induced protein dimerization. A minimum of two titrations, under identical conditions, was performed for each metal and their results averaged. Metal dissociation constants ($K_{d\text{-}metal}$) determined through these titrations are listed in Table 4.1, Additional figures and tables can be found in the appendix.

Table 4.1. Metal dissociation constants for HQuin1, AQuin1 and free 8-hydroxyquinoline (Quin).

Metal	HQuin1[a] (M)	AQuin1[a] (M)	Free Quin[b] (M)
Co^{2+}	$4.0\ (2) \times 10^{-9}$	$3.0\ (1) \times 10^{-7}$	6.48×10^{-7}
Ni^{2+}	$4.0\ (2) \times 10^{-10}$	$3\ (1) \times 10^{-8}$	1.55×10^{-7}
Cu^{2+}	$8.5\ (9) \times 10^{-14}$	$5.4\ (5) \times 10^{-9}$	2.29×10^{-10}
Zn^{2+}	$7.1\ (3) \times 10^{-9}$	$9\ (1) \times 10^{-7}$	8.74×10^{-7}

[a]See experimental section a detailed discussion on binding models. Additional tables can be found in the appendix section A4.2. [b]pH-adjusted dissociation constants reported in reference [2].

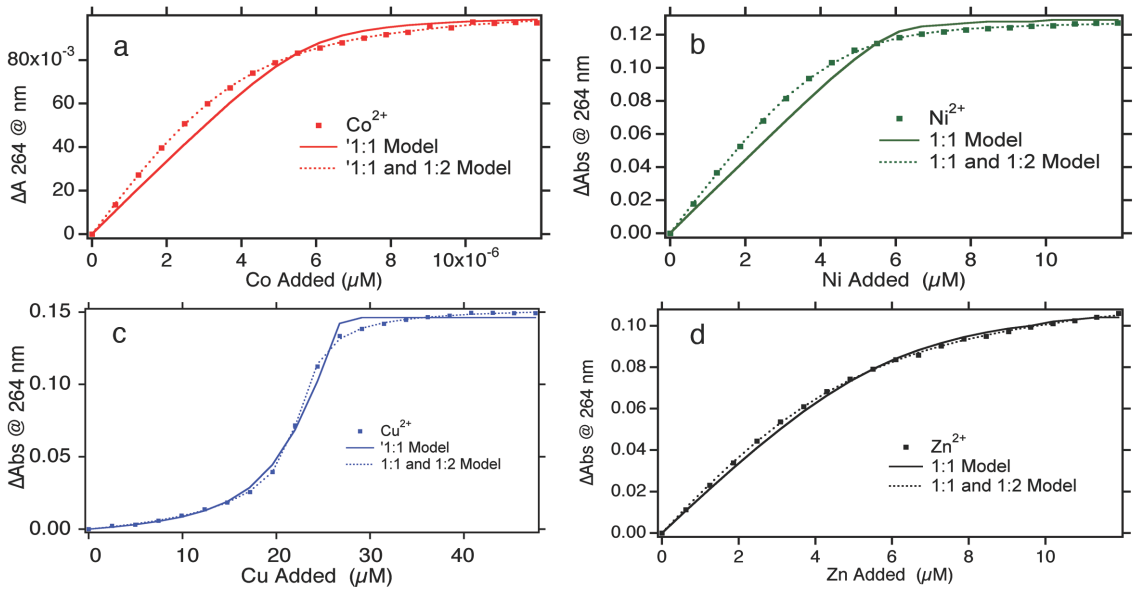


Figure 4.3. Metal binding titration data and fits for AQuin1. The lower metal affinity of AQuin1 (compared to HQuin1) allows for spectral changes due to M^{2+} ((a) Co^{2+} (b) Ni^{2+} (d) Zn^{2+}) binding to be directly monitored without the need for a competing ligand. Regression analysis for Co, Ni and Zn binding was preformed assuming both a simple 1:1 (solid) and 1:1/1:2 (dotted line) binding event. The notable exception is (c) Cu binding, for which the dissociation constant is still low enough to require both the inclusion of EGTA as a competing ligand AQuin1 concentrations ranged from 2 to 10 µM, and EGTA concentration was 25 µM for the Cu^{2+} titration.

Metal-Induced Stabilization

With the establishment of a high-affinity metal-binding site, it was then asked if the metal coordination by the *i*/*i*+7 His-Quin HCM would translate into increased protein stability. The stabilization of α-helical proteins/peptides through metallic or non-metallic crosslinking of *i*/*i*+4, *i*/*i*+7 and *i*/*i*+11 positions has been extensively documented [16]. Helical peptides stabilized through covalent hydrocarbon-stapling, in particular, have garnered recent attention due to their efficacy in inhibiting protein-protein interactions that are important pharmaceutical targets [17]. Guanidine hydrochloride (GuHCl) and

thermally induced unfolding experiments show that HQuin1 is indeed stabilized by metal binding, with Ni producing the largest and Cu the smallest effect (Ni > Zn ≈ Co > Cu) (Figures 4.3 and 4.4). In the case of Ni, the stabilization amounts to a ~1-M [GuHCl] or 12-K increase in the unfolding midpoint. It is suggested that the discrepancy between the trends in metal binding affinity (vide supra) and metal-induced stabilization is due to the differential interactions of Co, Ni, Cu and Zn with unfolded or partially folded conformations of HQuin1. Evidence for such metal cross-linked unfolded or partially folded states comes from the shallow unfolding transition of HQuin1 in the presence of metals (Figures 4.3 and A4.1.2), which can be attributed to deviations from two-state unfolding behavior [18].

In the absence of the Quin moiety (i.e., for the non-functionalized protein) or the coordinating H63 residue (i.e., for AQuin1) any enhancement in protein stability is eliminated (Figures 4.5a and 4.5b). Likewise, lowering the pH to 5.5, which is below the pKa of His (~6.5) but above that that of the Quin imine group, drastically diminishes the observed stabilization (Figure 4.5c). Taken together, the metal binding and protein unfolding titrations confirm that HQuin1 coordinates metal ions by engaging both H63 and Quin (C70).

Metal-Induced Self-Assembly of HQuin1

With the tridentate coordination mode established, the self-assembly properties of HQuin1 in response to metal coordination were examined using sedimentation velocity (SV) experiments. Under conditions where the metal:protein concentration ratio is greater than unity to ensure the full loading of the *i*/*i*+7 His-Quin HCM, the only HQuin1 species present in solution is monomeric (S_{max} = 1.8) (Figure 4.6). In contrast, at a metal:protein

ratio of 1:2, Co, Ni, Cu and Zn all induce the formation of dimeric species. Nevertheless, the dimers appear to have different shapes based on the varying sedimentation coefficients obtained with each metal (S_{max} = 2.4 for Cu, 2.5 for Zn, 2.6 for Co and Ni) (See Appendix A4.1.4).

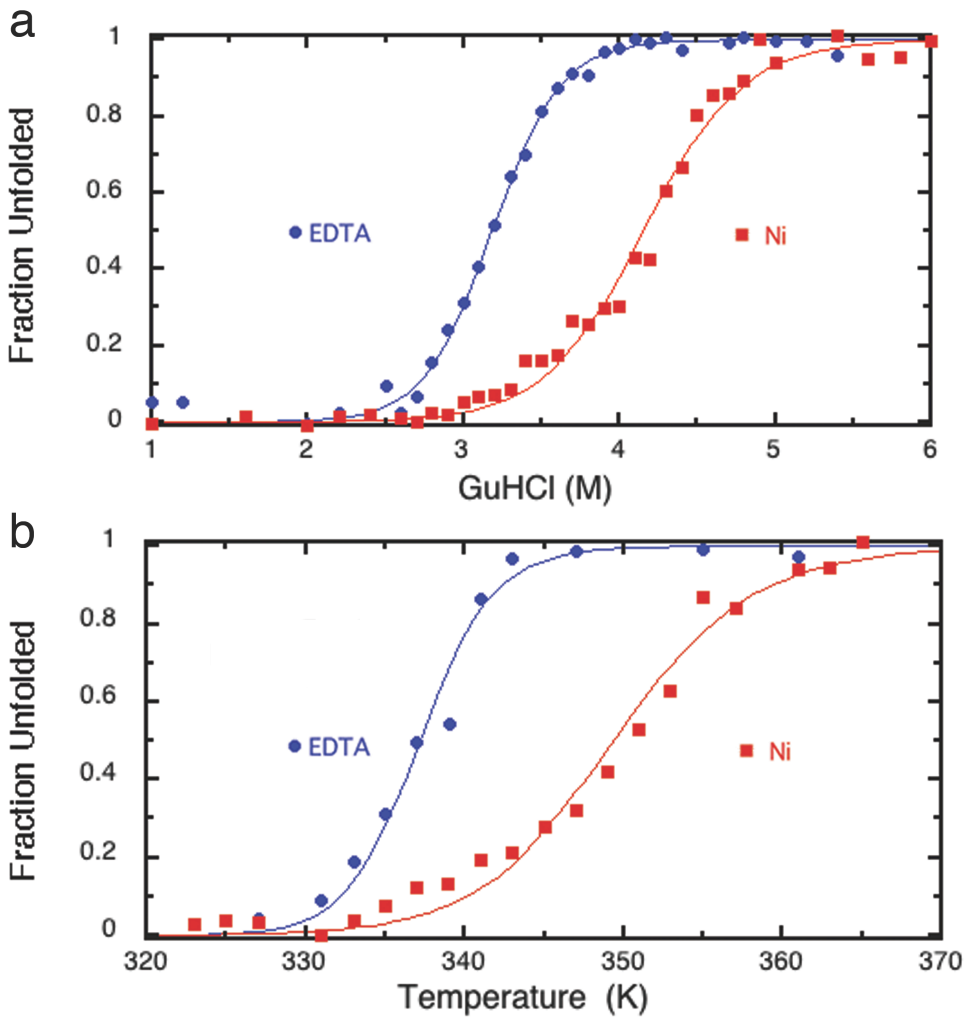


Figure 4.4. (a) GuHCl and (b) thermally induced unfolding curves for HQuin1 in the absence (blue dots and lines), and presence of Ni^{2+} (red dots and lines) (see Figure A4.1.2 for other metals). Although both sets of data were fit using a two-state unfolding model (see Experimental Section), the unfolding transitions in the presence of metals are shallower, suggesting metal cross-linked intermediate species.

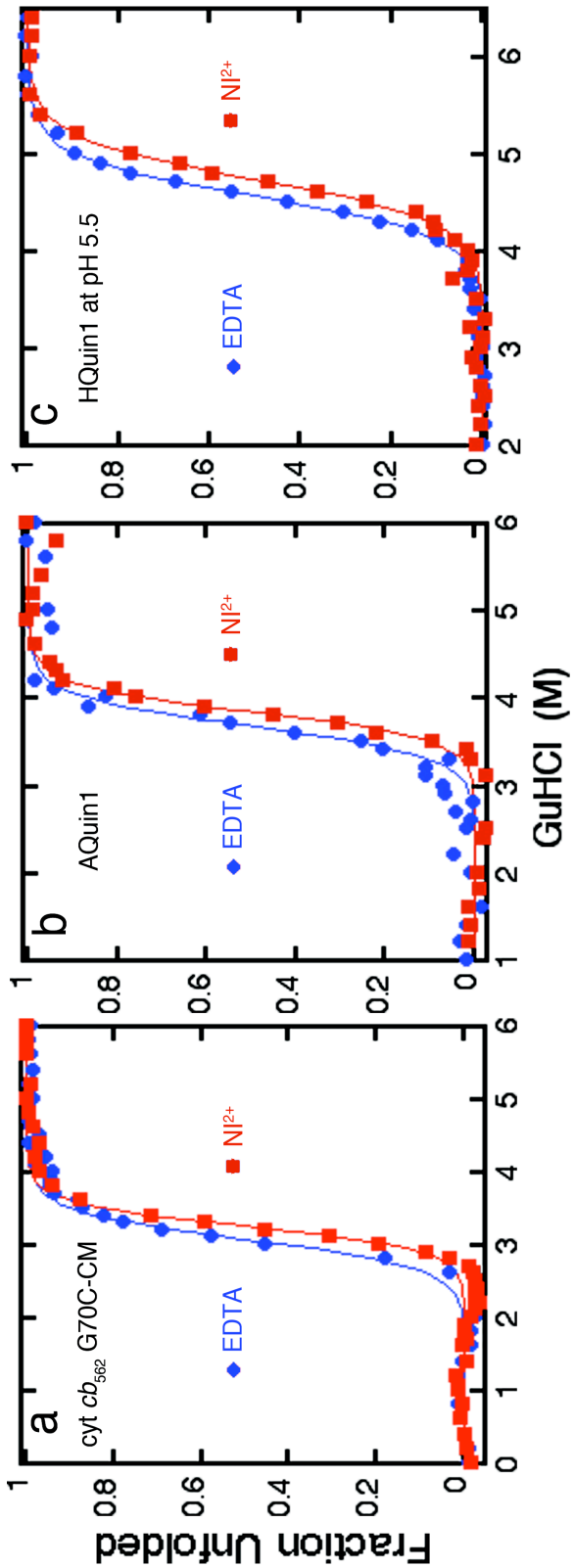


Figure 4.5. GuHCl induced unfolding titrations and fits for cyt cb_{562} variants in the absence (■) and presence) of 1 mM Ni^{2+} (•). (a) (CM)-G70C-cyt cb_{562} in 100 mM Tris buffer (pH 7.5) (b) AQuin1 in 100 mM Tris buffer (pH 7.5) (c) HQuin1 in 100 mM sodium acetate buffer (pH 5.5). Fitting parameters and statistics are listed in Table A4.3.4.

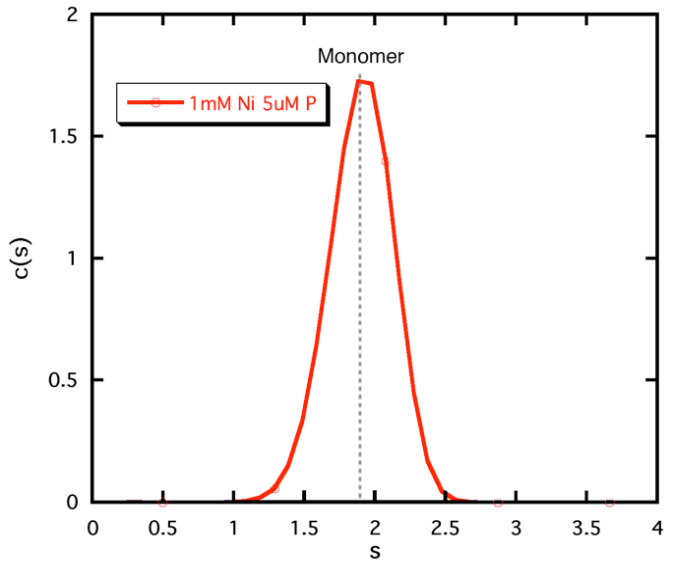


Figure 4.6. Sedimentation coefficient distribution for HQuin1 in the precence of excess metal. The sharp peak centered at a sedimentation coefficient of 1.9 reveals that HQuin1 remains monomeric in the presence of excess metal.

Given the octahedral coordination preference of Ni^{2+}, it was hypothesized that the metal ion would fully coordinate two tridentate His-Quin HCM's, thereby giving rise to a compact and rigid protein dimer (Figure 4.1.). To determine an apparent dissociation constant for Ni-induced protein dimerization ($K_{d\ Ni\text{-}dimer}$), a series of SV measurements were performed at varying HQuin1 concentrations while maintaining a 1:2 Ni:HQuin1 molar ratio. The SV distributions reveal a clear monomer-dimer transition as the HQuin1 concentration is increased from 5 μM to 25 μM (Figure 4.7), placing $K_{d\ Ni\text{-}dimer}$ roughly at 10 μM. This dissociation constant is in reasonable agreement with that obtained from sedimentation equilibrium (SE) measurements ($K_{d\ Ni\text{-}dimer}$ = 42 μM) run under similar conditions as the SV experiments (Figure A4.1.7).

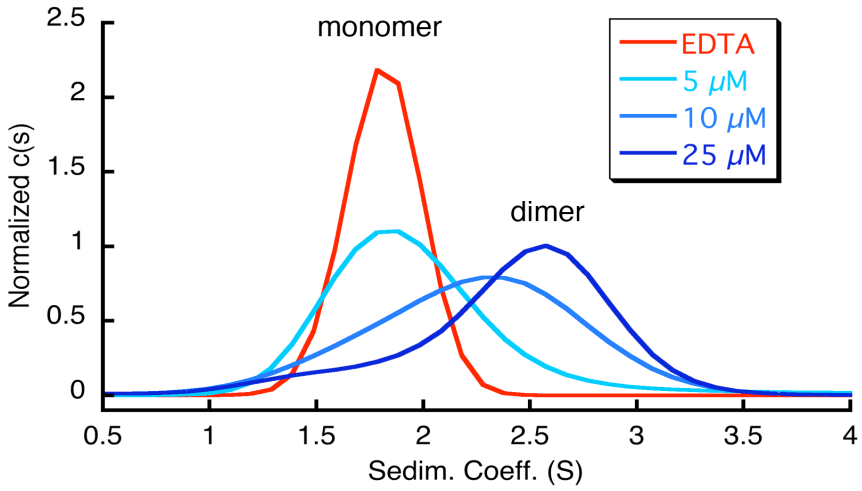


Figure 4.7. Sedimentation velocity experiments showing the change in sedimentation coefficient distributions for various concentrations of HQuin1 in the absence of metals (red trace) and in the presence of a half molar equivalent of Ni^{2+} (blue traces).

To elucidate the exact mode of Ni binding to His-Quin HCM and Ni-induced dimerization, the crystal structure of the Ni:HQuin1$_2$ complex was determined at 2.3 Å resolution (PDB ID: 3L1M). The asymmetric unit of the $P2_12_12$ crystals contains a single HQuin1 protomer coordinated to a half-occupied Ni^{2+} lying on a crystallographic two-fold symmetry axis. This two-fold symmetry produces a V-shaped dimer with a parallel arrangement of two HQuin1 protomers (Figure 4.8a). The acute angle (~50º) between the protomers results in minimal contact (~300 $Å^2$ buried surface) between their surfaces and is entirely enforced by Ni coordination to His-Quin HCM's in a distorted octahedral geometry. The Ni coordination sphere consists of a nearly ideal equatorial square plane formed by four nitrogens from two H63's and Quin's, and two axial phenolate oxygens that form a nonlinear O-Ni-O angle of 165º due in part to the small Quin bite angle of 80º (Figure 4.8b). The observed bond metrics closely approximate

those of Ni^{2+} complexes with free quinolate and amine-type ligands [19]. This suggests that Ni coordination in the Ni:HQuin1$_2$ complex is free from steric constraints that may be imposed by the covalent attachment of the His-Quin HCM to the protein surface. The conformational plasticity of the Cys70- Quin linker region – owing to four freely rotatable bonds – is evident in the electron density maps (Figure 4.8b): to accommodate octahedral Ni coordination, the linker adopts a somewhat strained conformation, whereby the S(Cys)-C1(linker) thioether and the C2(linker)-N(linker) amide bonds are found in a near-eclipsed configuration.

Stereochemical considerations suggest that there is one other possible isomer for His-Quin/Ni coordination that could lead to the formation of an alternative HQuin1 dimer in an antiparallel arrangement, again with little protein-protein contact (Figure 4.9). This alternative isomer would pose the Quin groups in a trans arrangement, whereby the equatorial plane would be formed by the Quin donor atoms and the axial positions would be occupied by the His ligands. . In order to determine if there is a thermodynamic basis for the exclusive population of the observed "*cis*-Quin" isomer, Density Functional Theory (DFT) calculations (ADF 2007.01, ZORA/TZ2P basis set) were carried out. Using BP86 and OLYP functionals in parallel, the optimized coordination geometries and corresponding energies for the *cis*-Quin and the *trans*-Quin isomers (Section A4.4) were calculated. These calculations bore out two important results: 1) The crystallographically-determined bond distances and angles closely approximate those of the calculated *cis*-Quin species (Tables A4.4.3 and A4.4.4), confirming the previous conclusion that Ni coordination is not strained by covalent linkages to the protein backbone. 2) The *cis*-Quin isomer is on average ~6 kcal/mol more stable than *trans*-Quin

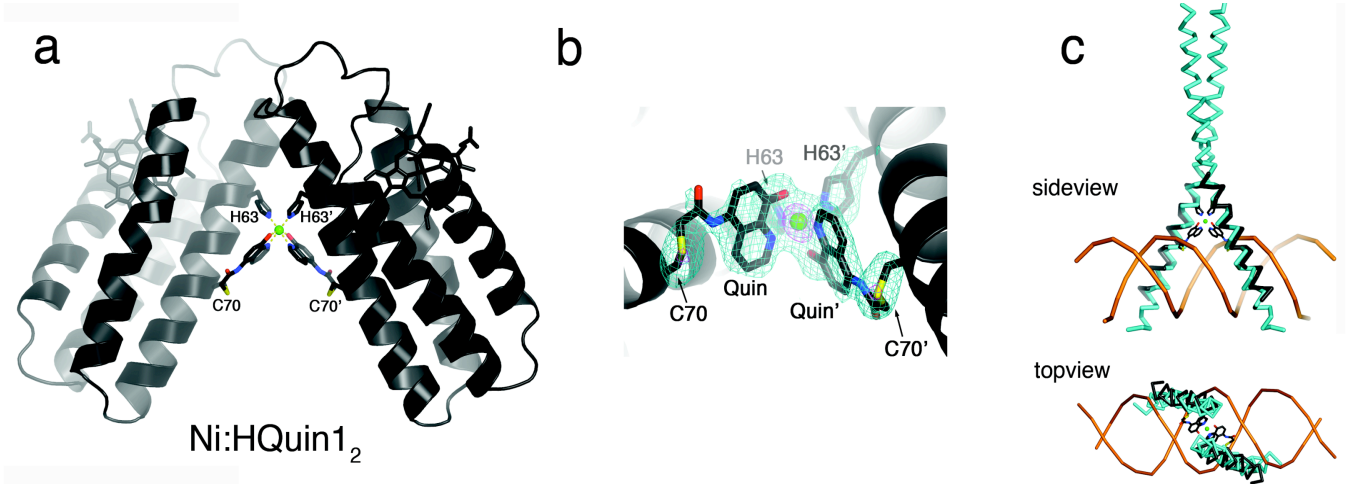


Figure 4.8. (a) Crystal structure of the Ni:HQuin1$_2$ dimer. (b) Close-up view of the Ni coordination environment and the corresponding Fo – Fc omit difference map (cyan – 3 s, purple – 9 s). (c) Backbone superposition of the Helix3 domains of Ni:HQuin1$_2$ (black) onto the basic domain of Jun bZip homodimer (cyan) complexed with cAMP responsive element (CRE) (brown) (PDB ID: 1JNM).

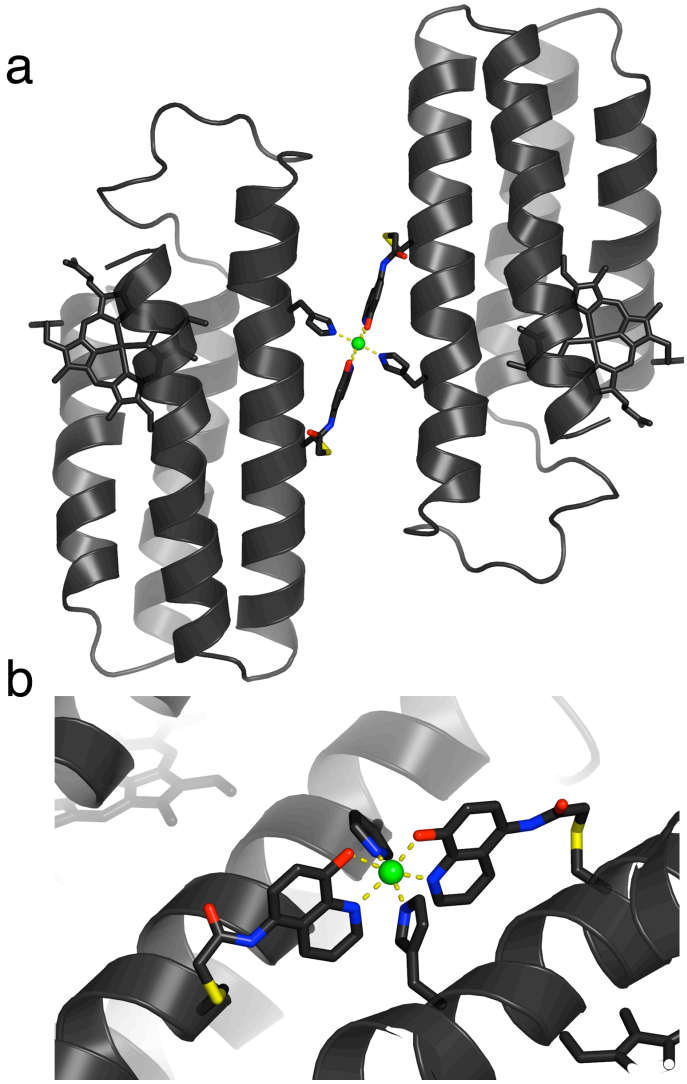


Figure 4.9. (a) Model for the alternative dimeric arrangement of Ni:HQuin1$_2$. (b) Close-up view of the corresponding octahedral Ni coordination environment. The modeling was done manually in XFIT [1], whereby a copy of the crystallographically determined HQuin1 protomer with the tridentate coordination motif was taken as a rigid unit and possible dimeric arrangements were explored where the octahedral coordination preference of Ni^{2+} would be fulfilled. The model illustrated here was found to be the only stereochemically allowed alternative arrangement.

(Table A4.2.8), due likely to the trans-directing effect of the imine ligands which would render a mutual trans orientation of the weaker-field phenolate ligands the least destabilized configuration. This significant energy difference should lead to the *cis*-Quin isomer being exclusively formed in solution, and ultimately to the formation of a single, discrete protein dimer. The engineering of a preferred dimeric protein conformation over others often requires implementation of negative design strategies (destabilization of undesired conformations that often lie close in energy) alongside the optimization of the desired target. While remarkable success has been achieved in some cases [6, 20-22], such design efforts are time- and labor-intensive and not readily generalizable. In the approach, dimerization specificity can be achieved through the consideration of only a few metal-ligand bonds, which eliminates the need to design large molecular surfaces and renders it readily applicable for other helical proteins.

A major objective of MDPSA is to access biologically functional geometries in a completely metal dependent fashion. This goal, in theory, would allow for the construction of biologically active structures with novel or expanded functionalities. Therefore, it was exciting to discover that structurally the V-shaped architecture of Ni:HQuin1$_2$ is closely reminiscent of the DNA-binding domains of the bZIP-family transcription factors. The bZIP proteins consist of a flexible, "proto-helical" basic domain that interacts with the DNA major groove and a helical leucine-zipper domain whose dimerization is necessary for the preorganization of the basic domain for stable DNA binding. Starting with the works of Kim [23] and Schepartz [24], it has been shown in many instances that DNA recognition by bZIP proteins is sensitively dependent on the dimeric orientation of the basic domains. A structural superposition of the Helix3 regions

(M58-N80) of the Ni:HQuin1_2 dimer with the DNA-bound basic domain of a representative bZIP protein reveals a very close match, with a root-mean-square-deviation of 1.6 Å over 46 Ca's (Figure 4.6c). This example demonstrates the potential utility of surface HCM's in directing the formation of rigid protein/peptide structures that are poised to recognize biological targets without the need for engineering extensive protein surfaces or peripheral oligomerization domains.

4.4 Conclusions

In summary, the proof-of-principle studies show that helical protein surfaces provide an ideal scaffold for the construction of multidentate coordination motifs that consist of natural and non-natural metal ligands. HCMs simultaneously combine the advantages of high metal binding affinity, helix stabilization with the ability to precisely control protein oligomerization, based solely on the inner-sphere coordination of the metal center. The functional versatility of HCMs, combined with their implicitly modular nature and ease of construction, should render them useful in many applications, ranging from selective protein labeling with metal reporters and protein purification to construction of superprotein architectures and probing protein-protein and protein-DNA interactions.

4.5 Experimental Section

Site Directed Mutagenesis and Protein Expression/Purification/Characterization

Site directed mutagenesis, protein expression and purification were performed as previously reported in chapter 2.

Point mutations were executed to obtain the following cyt cb_{562} variants: G70C-cyt cb562 and G70C/H63A-cyt cb_{562}. Sequencing of all mutant plasmids was carried out by Retrogen Inc. (San Diego, CA).

Protein purity was determined by SDS-PAGE gel electrophoresis. Verification of mutations was made through MALDI mass spectrometry (MW (G70C-cyt cb_{562}) = 12386 amu, MW (G70C/H63A-cyt cb_{562}) = 12320 amu).

Synthesis of Iodoacetamido-8-hydroxyquinoline (IA-Quin)

As a precursor, iodoacetic acid anhydride was freshly prepared by adding 660 mg (3.20 mmol) of dicyclohexylcarbodiimide (DCC) (Sigma) to a stirred solution of 1.19 g (6.43 mmol) iodoacetic acid (Sigma) in 25 mL of ethyl acetate. Dicyclohexylurea precipitates immediately, but the mixture was allowed to stir for 2 hrs in the dark. The dicyclohexylurea was removed by filtration and the resulting solution was evaporated to dryness. 500 mg (2.14 mmol) of 5-amino-8-hydroxyquinoline dihydrochloride (Sigma) was dissolved in 10 ml of acetonitrile by refluxing overnight with 975 μL (7 mmol) of triethylamine. The solution was filtered and the iodoacetic acid anhydride, dissolved in 5 ml of acetonitrile, was added. The mixture was allowed to react in the dark overnight. The product evaporated to dryness and washed extensively with cold 5% sodium

bicarbonate and water and dried in vacuo (Yield: 75%). Synthesis of IA-Quin was verified by mass spectrometry (ESI-MS, positive mode). Measured mass = 329.05 amu (expected mass = 328.9) (M + H+)

Functionalization of G70C-cyt cb562 and G70C/H63A-cyt cb562 with IA-Quin and iodoacetic acid to obtain HQuin1, AQuin1 and carboxymethylated (CM)-G70C-cyt cb562

A solution of 0.3 mM G70C-cyt cb_{562} or G70C/H63A-cyt cb_{562} in degassed 0.1 M Tris buffer (pH 7.75) was treated with a 10-fold excess of dithiothreitol (DTT) (Sigma). The protein was allowed to reduce for a period of 30 min. The protein was then dialyzed against 2 × 1 L of degassed 0.1 M Tris buffer (pH 7.75) under an inert atmosphere to remove the DTT. A 10-fold excess of IA-Quin or iodoacetic acid was dissolved in 2 mL of degassed dimethylformamide (DMF) and added dropwise to the protein solution over the course of 1 min. The mixture was allowed to react in the dark at 25° C overnight. The reaction mixture was then dialyzed again against 2 × 1 L of 10 mM sodium acetate (pH 5) and 1 mM ethylenediaminetetraacetic (EDTA). The crude labeled protein was subsequently purified on an Uno-S cation-exchange column (BioRad) using an NaCl gradient. The purity of the functionalized protein was determined by MALDI mass spectrometry and SDS-PAGE electrophoresis (Labeling yield: 60-95%). MW(HQuin1) = 12590 amu (exp. = 12589 amu); MW(AQuin1) = 12522 amu (exp. = 12523 amu); MW(CM-G70C-cyt cb_{562}) = 12447 amu (exp. = 12446 amu).

Chemical Unfolding Experiments

5 mL of an unfolded protein (CM-G70C-cyt cb_{562}, HQuin1 or AQuin1) solution containing 5 μM of protein and 1 mM of M^{2+} or EDTA was freshly prepared in ~ 8 M

guanidine HCl (GuHCl) in the appropriate buffer (either 0.1 M Tris buffer (pH 7.5) or 0.1 M sodium acetate (pH 5.5)). In parallel, 3 mL of a folded protein solution containing 5 μM protein in the appropriate buffer (either 0.1 M Tris buffer (pH 7.5) or 0.1 M sodium acetate (pH 5.5)) and 1 mM M^{2+} or EDTA was prepared. The unfolded protein stock was titrated into the folded protein stock at 25°C using an autotitrator (Microlab 500 Series), keeping the sample volume constant at 2 mL; protein unfolding/folding was monitored by CD spectroscopy (222 nm) on an Aviv 215 spectrometer. For every titration point, the solution was allowed to stir for 30 seconds in order to reach equilibrium. This procedure was repeated for a minimum of 20 points covering a GuHCl range of 0.1-6.5 M. GuHCl concentrations were calculated using the refractive indices of the folded and unfolded protein stock solutions.[25] Unfolding data were fit using Kaleidagraph (Synergy Software) with an expression that assumes a two-state folding/unfolding equilibrium as described by Pace (eq. 1):[26]

$$\text{Fraction Unfolded} = \frac{e^{\frac{(-m1 \times (m2 - [GuHCl])}{RT}}}{1 + e^{\frac{(-m1 \times (m2 - [GuHCl])}{RT}}} \qquad (1)$$

Where mL represents the slope of the unfolding transition and is defined as (∂ΔGH2O/∂[GuHCl]) and m2 represents the midpoint GuHCl concentration where 50% of the protein is unfolded. It is confirmed through sedimentation velocity experiments that all proteins, including HQuin1, remained in their monomeric form at under the conditions (5 μM of protein and 1 mM M^{2+}) used in chemical denaturation experiments (Figure A4.1.5). Fitting parameters for each titration is given in Table A4.2.5.

Thermal Unfolding

A 3 mL solution of HQuin1 containing 5 μM protein in 0.1 M Tris buffer (pH 7.5) and 1.5 M GuHCl was prepared. Addition of 1.5 M GuHCl was necessary to ensure that HQuin1 fully unfolds below 100°C. To the protein solutions either 1 mM Ni^{2+} or EDTA was added. The unfolding reaction was monitored by CD spectroscopy (222 nm). At each temperature, the solution was allowed to stir for 30 sec in order to reach equilibrium. This procedure was repeated for a minimum of 20 points covering a temperature range of 300-376 K. Although the thermal unfolding of HQuin1 is not completely reversible (i.e., the 222 nm CD signal does not return to pre-melting conditions), the curve was fit to a two-state model as described by John and Weeks (equation 2),[27] to obtain an apparent $\Delta T_{m\text{-metal}}$. Unfolding data were fit using Kaleidagraph (Synergy Software) to the following equation:

$$\text{Fraction Unfolded} = \frac{e^{\left(\left(\frac{\Delta H_{vH}}{R}\right)\times\left(\left(\frac{1}{T_m}\right)-\left(\frac{1}{T}\right)\right)\right)}}{1+e^{\left(\left(\frac{\Delta H_{vH}}{R}\right)\times\left(\left(\frac{1}{T_m}\right)-\left(\frac{1}{T}\right)\right)\right)}} \qquad (2)$$

where ΔHvH is the change in the van't Hoff transition enthalpy, Tm is melting point, T is the temperature in Kelvin and R is the universal gas constant. Table of thermal unfolding parameters is given in Table A4.2.6.

Metal Binding Titrations

General: Unless otherwise stated, all metal (M^{2+}) binding titrations were prepared by diluting a concentrated protein stock solution (HQuin1 or AQuin1) to a final volume of 2 mL with a final protein concentration ranging from 2 to 10 μM. All titrations were performed in 50 mM MOPS buffer (pH 7) previously treated with Chelex resin (Bio-Rad) to ensure a metal free environment. All pipette tips were rinsed 3x in an analytical grade 5% HNO3 (Fluka) solution before use. All further procedures followed to ensure a metal-free environment have been previously outline by Linse [28]. Titration data were fit using non-linear regression on Dynafit 3 (BioKin). All absorption spectra were obtained on an HP 8452A spectrophotometer. HQuin1 and AQuin1 concentrations were determined based on the Soret absorption maximum for cyt cb_{562} at 415 nm ($\varepsilon = 0.148\ \mu M^{-1}\ cm^{-1}$) [29]. All data were baseline-corrected and adjusted for dilution. Change in absorbance at 264 nm, (ΔA264 nm), which displays the largest increase upon M^{2+} binding, was plotted as a function of M^{2+} concentration. The change in extinction coefficient at 264 nm due to metal binding (Δε264 nm) was calculated based on titration data and held fixed during fitting. UV-vis absorbance profile for a representative titration (Zn binding) is shown in Figure 4.1. Due to the nature of the HQuin1 and AQuin1 system to self-associate, data were fit to a combined 1:1 and a 1:1/1:2 model. The latter model (1:2) was incorporated to account for both metal-induced protein dimerization. In all cases, the data were fit using both models. Titration curves for each metal for HQuin1 and AQuin1 can be found in Figures 4.2 and A4.2.1 respectively. The determined dissociation constants for HQuin1 and AQuin1 can be found in tables 4.1 and A4.2.1 – A4.2.4 respectively.

HQuin1 titrations: To a 2 ml solution of HQuin1, containing a known amount of EGTA (25-100 μM), successive aliquots of an appropriate M^{2+} stock (100 μM-2.5 mM) in 50 mM MOPS (pH 7) were added. The total amount of M^{2+} added never exceeded 100 μL (5% of the total volume). M^{2+} dissociation constants for EGTA were calculated using MaxChelator (http://maxchelator.stanford.edu) and fixed during data fitting. Despite the fact that the HQuin1 concentrations are sufficiently low to prevent a significant extent of dimer formation, the titration data were separately fit to two models, which do or do not take metal-induced protein dimerization into account (Tables A4.2.1 and A4.2.2). Changes in the extinction coefficient at 264 nm (Δε264) were calculated separately for each measurement and held fixed during final regression analysis. The Δε264 values used for different metals were 21560, 28920, 28970, 24090 M-1 cm-1 for Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+}, respectively.

AQuin1 Titrations: Due the lower affinity of AQuin1 (compared to HQuin1) for M^{2+}, inclusion of EGTA as a competing ligand was not necessary. The sole exception was Cu^{2+}, whose high affinity for AQuin1 required the addition of 25 μM EGTA. Titration data were handled in a similar manner to those of HQuin1. Once again, changes in extinction coefficient at 264 nm due to metal binding (Δε264 nm) were calculated for each titration and held fixed during the regression analysis. Titration data for AQuin1 were fit to both 1:1 and 1:1/1:2 models. Despite the fact that AQuin1 concentrations are sufficiently low to prevent a significant extent of dimer formation, in all cases data were notably better described by a model that included metal-induced dimerization (1:1/1:2). The heightened ability of AQuin1 to dimerize is attributed to the added flexibility the Quin moiety without the coordinating H63 residue. Regardless, dissociation constants

(KDmetal) for AQuin1 calculated unsing 1:1 and 1:1/1:2 models are listed in Tables A4.2.3 and A4.2.4.

Sedimentation Velocity Experiments

Sedimentation velocity (SV) experiments were performed in order to determine the solution-state oligomerization behavior of HQuin1. All SV samples were prepared in 20 mM Tris buffer (pH 7). Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor at 41,000 rpm for a total of 250 scans per sample. The following wavelengths were used for detection: 418 nm (5 mM protein), 425 nm (10 μM protein), 524 nm (25 mM protein) and 580 nm (200 μM protein).

All data were processed using SEDFIT [30]. Buffer viscosity, buffer density, and protein partial specific volume values were calculated at 25° C with SEDNTERP (http://www.jphilo.mailway.com). Partial specific volume (Vbar) for HQuin1 mutant was calculated to be 0.7347 mg/ml, assuming a partial specific volume of heme of 0.82 mg/ml and 0.75 mg/ml for the 8-hydroxyquinoline.[31] All data were processed using fixed values for buffer density (r) (0.99764 g/ml) and buffer viscosity (0.0089485 poise).

X-ray Crystallography

All crystals were obtained by sitting drop vapor diffusion. HQuin1 was crystallized at 25° C using a precipitant solution of 0.1 M Tris (pH 8.5), 25% PEG 1500, and 2.1 mM NiSO4. The drop consisted of 2 mL protein (2.1 mM in 20 mM Tris, pH 7) and 1 mL precipitation solution. Crystals appeared within one month, reaching a maximum size of ~ 500 mm × 500 mm × 700 mm. The crystals to be used for diffraction

experiments were exchanged into a solution containing 20% glycerol as a cryoprotectant and frozen in liquid nitrogen or directly in the cryostream.

X-ray diffraction data were collected at 100 K using an Bruker Apex II CCD detector and monochromatized Cu-Kα radiation (1.54 Å) produced by a Siemens sealed source. The data were processed using SAINT and Bruker SADABS. The structure of HQuin1:Ni was determined at 2.3 Å (P21212 spacegroup), by molecular replacement with PHASER,[32] using the monomeric cyt cb562 structure (PDB ID: 2BC5)[29] as the search model. The search model did not contain the heme or Quin prosthetic groups; the observation of strong positive Fo-Fc density at expected positions for these groups confirmed the correct placement of the protein monomer in the initial molecular replacement solutions. The topology and parameter files for the 5-acetamido-8-hydroxyquinoline group were obtained using the Dundee ProDrg Server (http://davapc1.bioch.dundee.ac.uk/prodrg/index.html). Rigid-body, positional and thermal refinement with CNS,[33] along with manual rebuilding and water placement with XFIT,[1] produced the final models. The Ramachandran plots were calculated with PROCHECK [34]. All figures were produced with PYMOL [35].

The final R and R_{free} values of 26.6 and 31.1% are somewhat higher than those for an average 2.3-Å resolution structure. An analysis of the diffraction data reveals no evidence for possible twinning, and molecular replacement and refinement of the data against other spacegroups either do not yield a solution or do not lead to better statistics. It was conclude that the high R-factors are likely due to the high solvent content (56%) and the lack of non-crystallographic symmetry within the asymmetric unit.

DFT Calculations

Density Functional Theory calculations were performed with the Amsterdam Density Functional (ADF) program suite [36, 37], version 2007.01 [38]. Crystallographic atomic coordinates were used as input where appropriate. Optimized geometries and molecular orbitals were visualized with the ADFView graphical routine of the ADF-GUI[39] and the Gaussview 3 program. For all atoms, the triple-ζ Slater-type orbital TZ2P ADF basis set was utilized without frozen cores. Relativistic effects were included by use of the zero-order regular approximation (ZORA) [40]. To ensure consistency over a range of exchange/correlation profiles, the molecular geometries and energies were evaluated with both the BP86 and OLYP functionals. For BP86, the local density approximation (LDA) of Vosko et al [41]. (VWN) was coupled with the generalized gradient approximation (GGA) corrections described by Becke [42] and Perdew [43, 44] for electron exchange and correlation, respectively. For OLYP, the parameterized (X = 0.67) exchange-only LDA was coupled with the GGA corrections described by Handy and Cohen (OTPX) [45] and Lee, Yang and Parr (LYP) [46] for electron exchange and correlation, respectively. All DFT calculations were performed on a home-built 72-CPU (1 × 8 master, 8 × 8 slave) Rocks 4.3 Linux cluster featuring Intel Xeon E5335 Quad-Core 2.00GHz processors. Job control was implemented with the Sun Grid Engine v. 5.3.

Chapter 4 is reproduced, in part, with permission from: R.J. Radford, P.C. Nguyen, T.B. Ditri, J.S. Figueroa, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 4362-4369.

4.6 References

[1] D.E. McRee, *J. Mol. Graphics*, 10 (1992) 44-46.

[2] A.E. Martell, R.M. Smith, Critical Stability Constants, Plenum Press, New York, 1974-1989.

[3] N.J. Marianayagam, M. Sunde, J.M. Matthews, *Trends Biochem. Sci.*, 29 (2004) 618-625.

[4] S. Jones, J.M. Thornton, *Prog. Biophys. Mol. Biol.*, 63 (1995) 31-59.

[5] B. Kuhlman, J.W. O'Neill, D.E. Kim, K.Y.J. Zhang, D. Baker, *Proc. Natl. Acad. Sci. USA*, 98 (2001) 10687-10691.

[6] D.N. Bolon, R.A. Grant, T.A. Baker, R.T. Sauer, *Proc. Natl. Acad. Sci. U.S.A.*, 102 (2005) 12724-12729.

[7] S. Gendreizig, M. Kindermann, K. Johnsson, *J. Am. Chem. Soc.*, 125 (2003) 14970-14971.

[8] T. Shahian, G.M. Lee, A. Lazic, L.A. Arnold, P. Velusamy, C.M. Roels, R.K. Guy, C.S. Craik, *Nat. Chem. Biol.*, 5 (2009) 640-646.

[9] R.J. Radford, F.A. Tezcan, *J. Am. Chem. Soc.*, 131 (2009) 9136-9137.

[10] M.R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.*, 114 (1992) 825-831.

[11] S. Burazerovic, J. Gradinaru, J. Pierron, T.R. Ward, *Angew. Chem. Int. Ed. Eng.*, 46 (2007) 5510-5514.

[12] D.E. Przybyla, J. Chmielewski, *J. Am. Chem. Soc.*, 130 (2008) 12610-12611.

[13] R.J. Radford, M. Lawrenz, P.C. Nguyen, J.A. McCammon, F.A. Tezcan, *Chem Commun*, 47 (2011) 313-315.

[14] A. Lombardi, C.M. Summa, S. Geremia, L. Randaccio, V. Pavone, W.F. DeGrado, *Proc. Natl. Acad. Sci. USA*, 97 (2000) 6298-6305.

[15] H. Irving, R.J.P. Williams, *Nature*, 162 (1948) 746-747.

[16] S. Fletcher, A.D. Hamilton, *J. R. Soc. Interface*, 3 (2006) 215-233.

[17] R.E. Moellering, M. Cornejo, T.N. Davis, C.D. Bianco, J.C. Aster, S.C. Blacklow, A.L. Kung, D.G. Gilliland, G.L. Verdine, J.E. Bradner, *Nature*, 462 (2009) 182-188.

[18] L. Regan, A. Rockwell, Z. Wasserman, W. DeGrado, *Prot. Sci.*, 3 (1994) 2419-2427.

[19] A. Crispini, D. Puccis, S. Sessa, A. Cataldi, A. Napoli, A. Valentini, M. Ghedini, *New J. Chem.*, 27 (2003) 1497-1503.

[20] J.J. Havranek, P.B. Harbury, *Nat. Struct. Biol.*, 10 (2003) 45-52.

[21] T. Kortemme, L.A. Joachimiak, A.N. Bullock, A.D. Schuler, B.L. Stoddard, D. Baker, *Nat. Str. Mol. Biol.*, 11 (2004) 371-379.

[22] G. Grigoryan, A.W. Reinke, A.E. Keating, *Nature*, 458 (2009) 859-864.

[23] R.V. Talanian, C.J. McKnight, P.S. Kim, *Science*, 249 (1990) 769-771.

[24] B. Cuenoud, A. Schepartz, *Science*, 259 (1993) 510-513.

[25] Y. Nozaki, *Methods Enzymol.*, 26 (1972) 43-50.

[26] N.C. Pace, B.A. Shirley, J.A. Thomson, in: T.F. Creighton (Ed.) Protein Structure: A Practical Approach, IRL Press, Oxford, 1990, pp. 311-330.

[27] D.M. John, K.M. Weeks, *Prot. Sci.*, 9 (2000) 1416-1419.

[28] Linse, S., Calcium-Binding Protein Protocols, Vol. 2: Methods and Techniques, ed. H.J. Vogel. Vol. 2. 2002, Totowa: Humana Press.

[29] J. Faraone-Mennella, F.A. Tezcan, H.B. Gray, J.R. Winkler, *Biochemistry*, 45 (2006) 10504-10511.

[30] P. Schuck, *Biophys. Chem.*, 108 (2004) 187-200.

[31] Calculated using Advanced Chemistry Development (ACD/Labs) Software V8.14 for Solaris (© 1994-2009 ACD/Labs).

[32] CCP4, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, 50 (1994).

[33] A.T. Brünger, P.D. Adams, G.M. Clore, W.L. DeLano, P. Gros, R.W. Grosse-Kunstleve, J.S. Jiang, J. Kuszewski, M. Nilges, N.S. Pannu, R.J. Read, L.M. Rice, T. Simonson, G.L. Warren, *Acta Crystallogr. D*, 54 (1998) 905-921.

[34] R.A. Laskowski, M.W. Macarthur, D.S. Moss, J.M. Thornton, *J. Appl. Crystallogr.*, 26 (1993) 283-291.

[35] W.L. DeLano, The PYMOL Molecular Graphics System (http://www.pymol.org), 2003.

[36] G. Te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler, *J. Comput. Chem.*, 22 (2001) 931-967.

[37] C.F. Guerra, J.G. Snijders, G. te Velde, E.J. Baerends, *J. Theor. Chem. Acc.*, 99 (1998) 391-403.

[38] ADF2007.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, www.scm.com.

[39] ADF-GUI 2007.01, SCM, Amsterdam, The Netherlands, www.scm.com Access date, February, 2008.

[40] E. van Lenthe, E.J. Baerends, J.G. Snijders, *J. Chem. Phys.*, 99 (1993) 4597-4610.

[41] S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.*, 58 (1980) 1200-1211.

[42] A.D. Becke, *Phys. Rev. A*, 38 (1988) 3098-3100.

[43] J.P. Perdew, W. Yue, *Phys. Rev. B*, 33 (1986) 8800-8802.

[44] J.P. Perdew, *Phys. Rev. B*, 34 (1986) 7406-7406.

[45] N.C. Handy, A.J. Cohen, *Mol. Phys.*, 99 (2001) 403-412.

[46] C.T. Lee, W.T. Yang, R.G. Parr, *Phys. Rev. B*, 37 (1988) 785-789.

Appendix to Chapter 4: Controlled Protein Dimerization through Hybrid Metal Coordination Motifs

Table of Contents

Section

Section A4.1. Experimental Figures

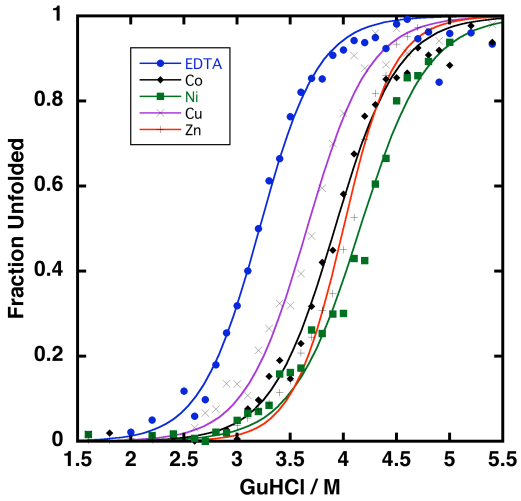


Figure A4.1.1. GuHCl induced unfolding data and fits for HQuin1 in the presence EDTA (blue), Co^{2+} (black), Ni^{2+} (green), Cu^{2+} (purple) and Zn (red). The parameters obtained from the fits are listed in Table S3.4.

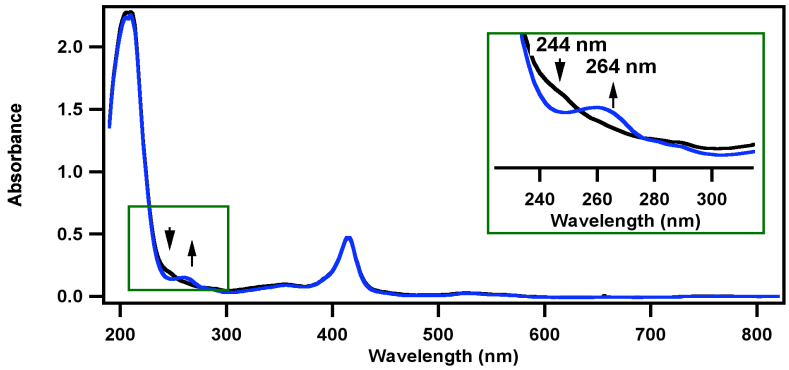


Figure A4.1.2. UV-visible spectra of HQuin1 in 100 mM sodium acetate (pH 5.5) with either 1 mM EDTA (black) or 1 mM Ni^{2+} (Blue). (Inset) The change in absorbance of the Quin group from the metal-free form (λ_{max}= 244 nm) to the metal bound form (λ_{max}= 264 nm) indicates the Quin group retains the its ability to bind metal even though no protein stabilization is observed at pH 5.5 (see Figure 4.5c).

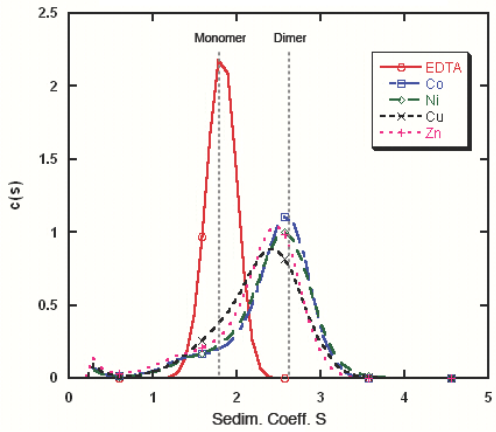


Figure A4.1.3. Sedimentation coefficient distributions of HQuin1 (25 µM) in the presence of 1 mM EDTA (red) or 12.5 µM M^{2+}: Co (blue), Ni (green), Cu (black), Zn (pink). Experiments were conducted at 25°C.

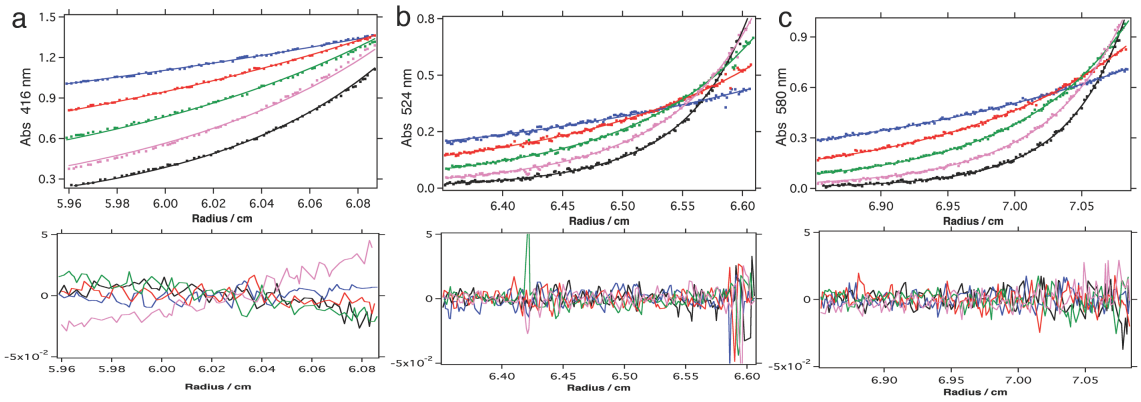


Figure A4.1.4. Sedimentation equilibrium profiles for HQuin1 in the presence of Ni^{2+}. (a) 5 µM protein at 15,000 (blue), 20,000 (red), 25,000 (green), 30,000 (pink) and 35, 000 (black) rpm. (b) 25 µM protein at 15,000 (blue), 20,000 (red), 25,000 (green), 30,000 (pink) and 35, 000 rpm (black). D) 200 µM protein at 15,000 (blue), 20,000 (red), 25,000 (green), 30,000 (pink) and 35, 000 rpm (black). All samples were in 20 mM TRIS buffer (pH 7) with a half equivalent of $NiSO_4$. Scans shown in a), b) and c) were globally fit to a monomer-dimer model yielding a minimized fit with a log $K_d^{dimer} = 4.38 \pm 0.023$ M^{-1} or K_d^{dimer} 41.6(2) µM. Experiments were conducted at 25°C.

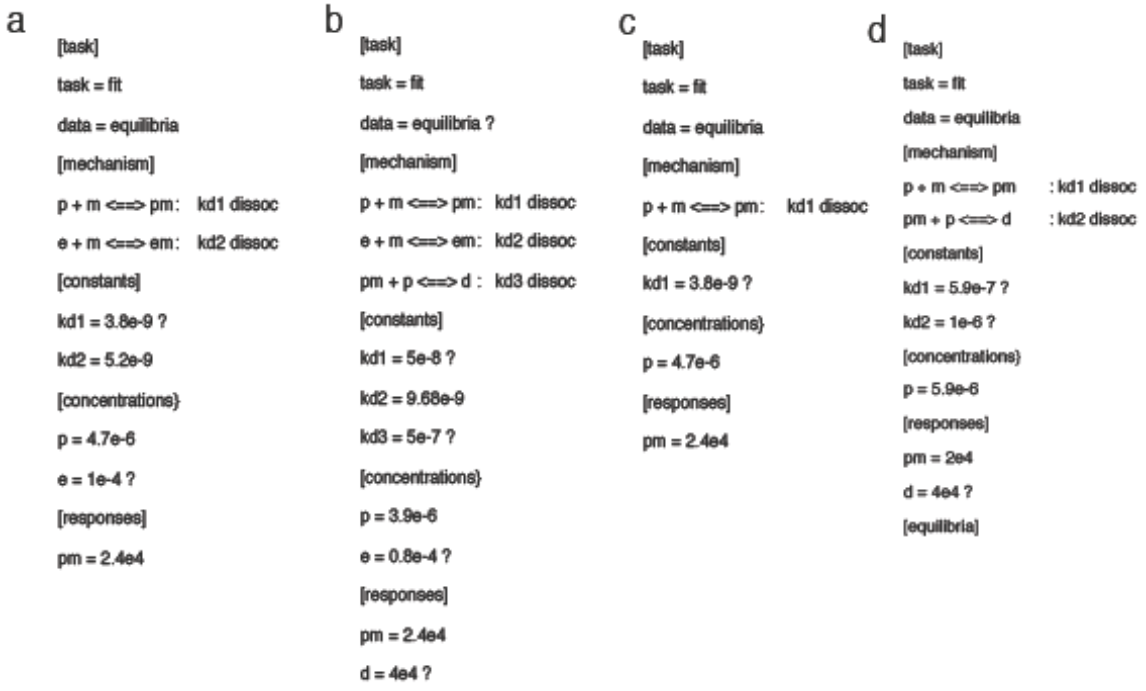


Figure A4.1.5. Example DynaFit scripts used for fitting HQuin1 and AQuin1 metal binding titrations. (a) 1:1 competitive binding model, (b) 1:1/1:2 competitive binding model, (c) 1:1 binding model and (d) 1:1/1:2 binding models. All models include the following variables: free protein (p), free metal (m), and metal-bound protein (pm). Variables for the concentration of free EGTA (e), EGTA-metal complex (em) and metal-induced metal dimer (d) were included where appropriate. The following dissociation constants were determined by non-linear regression analysis: kd1 (for initial protomer-metal binding), and kd3 (metal-induced dimerization). The metal dissociation constants for EGTA (kd2) were calculated using Maxchelator and held fixed. Values associated with "?" signify that the value was allowed to float during the fitting process.

Section A4.2. Experimental Tables

Table A4.2.1. Fitting parameters for metal binding titrations of HQuin1. Data fit using a 1:1 model. In all cases EGTA was used as competing ligand. EGTA:metal dissociation constants were determined using Maxchelator (http://maxchelator.stanford.edu) (25° C and a ionic strength of 0.05 M) and were held fixed during regression analysis.

Metal	K_D^{metal} (M)	K_D EGTA (M)
Co^{2+}	$4.0(2) \times 10^{-9}$	9.68×10^{-9}
Ni^{2+}	$4.0(2) \times 10^{-10}$	5.80×10^{-10}
Cu^{2+}	$8.5(9) \times 10^{-14}$	4.21×10^{-14}
Zn^{2+}	$7.1(3) \times 10^{-9}$	5.26×10^{-9}

Table A4.2.2. Fitting parameters for metal binding titrations of HQuin1 fit using a 1:1/1:2 (the latter representing protein dimerization) model.

Metal	K_D^{metal} (M)	K_D^{dimer} (M)	K_D EGTA (M)
Co^{2+}	$2\ (3) \times 10^{-8}$	$2\ (3) \times 10^{-6}$	9.68×10^{-9}
Ni^{2+}	$8.3(8) \times 10^{-10}$	$6\ (9) \times 10^{-4}$	5.80×10^{-10}
Cu^{2+}	$2\ (10) \times 10^{-12}$	$1(6) \times 10^{-6}$	4.21×10^{-14}
Zn^{2+}	$3(16) \times 10^{-9}$	$3\ (14) \times 10^{-6}$	5.26×10^{-9}

Table A4.2.3. Fitting parameters for metal binding titrations of AQuin1. Data were fit using a 1:1 model and determined via direct titration.

Metal	K_D^{metal} (M)
Co^{2+}	$2(1) \times 10^{-7}$
Ni^{2+}	$4(5) \times 10^{-8}$
Cu^{2+}	$6.0(1) \times 10^{-13}$
Zn^{2+}	$6(1) \times 10^{-7}$

Table A4.2.4. Fitting parameters for metal binding titrations of AQuin1. Data was fit using a 1:1/1:2 (the latter representing protein dimerization) model and determined via direct titration method. *Dissociation constants determined via competition assay using 25 μM EGTA as competing ligand.

Metal	K_D^{metal} (M)	K_D^{dimer} (M)
Co^{2+}	$3.0(1) \times 10^{-7}$	$2\ (1) \times 10^{-6}$
Ni^{2+}	$3(1) \times 10^{-8}$	$9(5) \times 10^{-7}$
Cu^{2+} *	$5.4(5) \times 10^{-9}$	$6.0(4) \times 10^{-9}$
Zn^{2+}	$9(1) \times 10^{-7}$	$6\ (3) \times 10^{-6}$

Table A4.2.5. Fitting parameters for chemical unfolding titrations of cyt cb_{562} variants: HQuin1, AQuin1, (CM)-G70C cyt cb_{562}.

cyt cb_{562} variant	Condition	Slope, m_1 (kcal/mol•M)	$[GuHCl]_m$, m_2 (M)	Figure
HQuin1	EDTA	2.1(1)	3.21(2)	1A, S2.4
HQuin1	Co	1.9(7)	3.91(1)	S2.4
HQuin1	Ni	1.8(8)	4.15(1)	1A, S2.4
HQuin1	Cu	1.9(1)	3.66(1)	S2.4
HQuin1	Zn	2.4(1)	3.99(1)	S2.4
(CM)-G70C	EDTA	2.9(1)	3.06(1)	S2.5a
(CM)-G70C	Ni	3.6(1)	3.24(1)	S2.5a
AQuin1	EDTA	3.4(2)	3.66(1)	S2.5b
AQuin1	Ni	3.8(2)	3.83(1)	S2.5b
HQuin1	EDTA (pH 5.5)	2.9(6)	4.55(1)	S2.5c
HQuin1	Ni (pH 5.5)	2.7 (1)	4.73(1)	S2.5c

Table A4.2.6. Fitting parameters for thermal unfolding titrations of HQuin1 (Figure 1B).

HQuin1	T_m (K)	slope (kcal/mol)	Figure
EDTA	337(1)	100(20)	1B
Ni	349(1)	55(4)	1B

Table A4.2.7. X-ray data collection and refinement statistics for Ni:HQuin1$_2$.

Contents of the asymmetric unit	106 residues + 1 Heme + 1 5-acetamide-8-hydroxyquinoline + 0.5 Ni + 46 waters
Unit cell dimensions	a = 88.44 Å , b = 34.51 Å, c = 46.90 Å $\alpha = \beta = \gamma = 90^o$
Symmetry group	$P2_12_12$
Resolution (Å)	50-2.30
X-ray wavelength (Å)	1.54
Number of Unique Reflections	6761
Redundancy	5.0 (3.2)
Completeness (%)*	99.0 (95.5)
$\langle I / \sigma I \rangle$*	14.3 (2.3)
R_{sym}[‡] (%)*	6.5 (40.9)
R[§] (%)*	26.6 (34.7)
Free R[II] (%)*	31.4 (39.4)
Rms Bnd[¶] (Å)	0.014
Rms Ang[¶] (^{o})	1.21
Ramachandran plot (%)	
Residues in most favored regions	90.8
Residues in add.l allowed regions	9.2
	0.0
Residues in generously allowed regions	0.0
Residues in disallowed regions	

[‡] $R_{sym} = \Sigma\Sigma_j | I_j - \langle I \rangle | / \Sigma\Sigma_j | I_j |$.
[§] $R = \Sigma | | F_{obs} | - | F_{calc} | | / \Sigma | F_{obs} |$.
[II]Free R calculated against 7.4 of the reflections removed at random.
[¶]Root mean square deviations from bond and angle restraints.
*Numbers in parentheses correspond to the highest resolution shell (2.39-2.30 Å)

Table A4.2.8. Total bonding energies for geometry optimized models. Lowest energy isomer and corresponding total bonding energy are in bold. $\Delta E = |$ *trans* isomer – *cis* isomer $|$.

Model	Total Bonding Energy (kcal mol^{-1})	ΔE (kcal mol^{-1})
***cis*-Quin BP86**	**-9438.29**	6.08
trans-Quin BP86	-9432.21	
***cis*-Quin OLYP**	**-9185.43**	5.21
trans-Quin OLYP	-9180.22	

Section A4.3. Input Files For Density Functional Calculations

A4.3.1. Input file for cis-Quin BP86

```
$ADFBIN/adf -n8 \
 <<< "
TITLE cisquinbp86 geo opt

MAXMEMORYUSAGE 11000

RELATIVISTIC ZORA
CHARGE  0 2
UNRESTRICTED

SCF
ITERATIONS 200
DIIS
END

XC
 LDA VWN
 GGA Becke Perdew
END

SYMMETRY NOSYM
ATOMS
 C                 2.27825000   -2.17357031   -2.44893750
 C                 1.40025000   -1.92357031   -1.38293750
 C                 2.84925000    2.64942969   -1.22293750
 C                 3.71425000    2.40542969   -2.28793750
 C                 3.08925000   -1.16257031   -2.96793750
 C                 2.04625000    1.60142969   -0.75393750
 C                 3.82425000    1.11442969   -2.85793750
 C                 3.01825000    0.08342969   -2.38193750
 C                 2.11925000    0.32242969   -1.32793750
 N                 1.33925000   -0.69257031   -0.84193750
```

```
 N                  4.64725000    0.91342969   -3.99193750
 O                  1.21325000    1.84042969    0.28106250
 C                  2.47925000   -2.90257031    4.89806250
 C                  2.07625000   -1.90757031    3.85106250
 C                  1.19725000   -1.99357031    2.82806250
 N                  2.57825000   -0.62757031    3.81106250
 C                  2.07125000    0.01942969    2.78206250
 N                  1.21325000   -0.78157031    2.17506250
 C                 -2.26775000    2.20342969   -2.44893750
 C                 -1.38975000    1.95342969   -1.38293750
 C                 -2.83875000   -2.61957031   -1.22293750
 C                 -3.70375000   -2.37557031   -2.28793750
 C                 -3.07875000    1.19242969   -2.96793750
 C                 -2.03575000   -1.57157031   -0.75393750
 C                 -3.81375000   -1.08457031   -2.85793750
 C                 -3.00775000   -0.05357031   -2.38193750
 C                 -2.10875000   -0.29257031   -1.32793750
 N                 -1.32875000    0.72242969   -0.84193750
 N                 -4.63675000   -0.88357031   -3.99193750
 O                 -1.20275000   -1.81057031    0.28106250
 Ni                -0.04275000   -0.12157031    0.66606250
 C                 -2.46875000    2.93242969    4.89806250
 C                 -2.06575000    1.93742969    3.85106250
 C                 -1.18675000    2.02342969    2.82806250
 N                 -2.56775000    0.65742969    3.81106250
 C                 -2.06075000    0.01042969    2.78206250
 N                 -1.20275000    0.81142969    2.17506250
 H                  2.32690405   -3.15559746   -2.87102217
 H                  0.78149673   -2.70882780   -1.00160422
 H                  2.79928463    3.61849928   -0.77204108
 H                  4.30458977    3.20661287   -2.68100807
 H                  3.74431586   -1.34858795   -3.79327753
 H                  5.04043829    1.78739467   -4.27757394
 H                  5.38283758    0.27648104   -3.76127036
 H                  3.21067741   -2.46393143    5.54421254
 H                  1.62078377   -3.18499581    5.47092512
 H                  2.89384888   -3.76876450    4.42613517
 H                  0.60053253   -2.84440442    2.57328891
 H                  3.23066890   -0.24538460    4.46549634
 H                  2.31219513    1.01978299    2.48856853
 H                 -2.31640405    3.18545683   -2.87102217
 H                 -0.77099673    2.73868717   -1.00160422
 H                 -2.78878463   -3.58863991   -0.77204108
 H                 -4.29408977   -3.17675349   -2.68100807
 H                 -3.73381586    1.37844732   -3.79327753
 H                 -5.02993829   -1.75753530   -4.27757394
 H                 -5.37233758   -0.24662167   -3.76127036
 H                 -3.20017741    2.49379080    5.54421254
 H                 -1.61028377    3.21485519    5.47092512
 H                 -2.88334888    3.79862387    4.42613517
 H                 -0.59003253    2.87426380    2.57328891
 H                 -3.22016890    0.27524397    4.46549634
 H                 -2.30169513   -0.98992362    2.48856853
END
```

```
GEOMETRY
 GO
 ITERATIONS 100
END

BASIS
 type TZ2P
 core none
END

END INPUT"
```

A4.3.2. Input file for *trans*-Quin BP86

```
$ADFBIN/adf -n8 \
 <<< "
TITLE transquinbp86 geo opt

MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE 0 2
UNRESTRICTED

SCF
 ITERATIONS 200
DIIS
END

XC
 LDA VWN
 GGA Becke Perdew
END

SYMMETRY NOSYM
ATOMS
 C                 2.53691018    3.25473656    1.32389056
 C                 1.59632173    2.58344734    0.52826116
 C                 2.18998762   -1.40165636    3.13598016
 C                 3.11703490   -0.73321961    3.93415911
 C                 3.16722171    2.61392680    2.39202222
 C                 1.57548671   -0.70313568    2.08852395
 C                 3.47508488    0.60937762    3.66671584
 C                 2.85433401    1.29102102    2.62287839
 C                 1.89585321    0.63866418    1.82910064
 N                 1.29721233    1.29672016    0.78805289
 N                 4.35522947    1.29071459    4.54243009
 O                 0.68202473   -1.35555377    1.31445786
 C                 4.27756800   -0.95916267   -3.76079374
 C                 3.22733586   -0.96584364   -2.69109994
 C                 2.36605195   -0.00965974   -2.27890050
 N                 2.95163532   -2.07696987   -1.92810866
```

```
 C                  2.00446978   -1.79726959   -1.05653870
 N                  1.61548027   -0.55029233   -1.25828081
 C                 -1.39989089    3.06610310   -2.79984282
 C                 -0.67933105    2.49894845   -1.73736029
 C                 -2.79537213   -1.73671627   -2.52415413
 C                 -3.50412440   -1.17768044   -3.58582034
 C                 -2.29510769    2.30381737   -3.55264449
 C                 -1.89757505   -0.92866963   -1.81415762
 C                 -3.36505804    0.19182972   -3.91564373
 C                 -2.46751495    0.98167578   -3.20168813
 C                 -1.72346979    0.42383092   -2.14743772
 N                 -0.85486935    1.20066202   -1.42848870
 N                 -4.03192544    0.71395961   -5.04981931
 O                 -1.21657554   -1.47253516   -0.78329150
 Ni                 0.13541406   -0.08537500   -0.01228125
 C                 -3.55155652    1.81850978    3.97566142
 C                 -2.79466238    1.09290088    2.90363250
 C                 -1.74218072    1.46785892    2.14320181
 N                 -3.08267017   -0.20146015    2.53800287
 C                 -2.28036426   -0.58000386    1.56460219
 N                 -1.44129154    0.40989863    1.31507457
 H                 -3.36549200    0.90118779   -5.77149243
 H                  5.32678139    1.36430741    4.31732794
 H                  4.00514105    1.69820870    5.38586749
 H                 -2.95874517    2.62946162    4.34421509
 H                 -4.46563646    2.20003950    3.57094802
 H                 -3.77036862    1.14378537    4.77676581
 H                 -3.79481470   -0.77006488    2.94975379
 H                 -2.30386062   -1.52485562    1.06300168
 H                 -1.23722323    2.41056410    2.17819989
 H                  1.95229620   -2.42859631    3.31982008
 H                  3.56462245   -1.24309972    4.76155725
 H                  3.86968030    3.13203658    3.01090361
 H                  2.77522390    4.27533838    1.10830181
 H                  1.11821456    3.09248173   -0.28241603
 H                  0.00718995    3.09654594   -1.17480704
 H                 -1.26079640    4.10019243   -3.03693344
 H                 -2.82978431    2.73157487   -4.37486370
 H                 -4.16386117   -1.79338935   -4.16075505
 H                 -2.93414546   -2.76332242   -2.25634490
 H                  5.12768120   -1.51812410   -3.42947870
 H                  3.88318285   -1.40365048   -4.65062009
 H                  4.57145927    0.04905521   -3.96579281
 H                  2.28269411    0.98206601   -2.67187825
 H                  1.61729250   -2.46365419   -0.31429527
 H                  3.40294458   -2.96445839   -2.02129711
 H                 -4.70131497    0.04789980   -5.37888159
END

GEOMETRY
 GO
 ITERATIONS 100
 FREQUENCIES
END
```

```
BASIS
 type TZ2P
 core none
END

END INPUT"
```

A4.3.3. Input file for *cis*-Quin OLYP

```
$ADFBIN/adf -n8 \
 <<< "
TITLE cisquinolyp geo opt

MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE  0 2
UNRESTRICTED

SCF
ITERATIONS 200
DIIS
END

XC
 GGA OLYP
END

SYMMETRY NOSYM
ATOMS
    1.C         2.243672   -2.200123   -2.447739
    2.C         1.360933   -1.893156   -1.394585
    3.C         2.799438    2.616302   -1.141645
    4.C         3.698600    2.399606   -2.207384
    5.C         3.067748   -1.211423   -2.947351
    6.C         2.005525    1.580376   -0.626074
    7.C         3.853038    1.168788   -2.829998
    8.C         3.043509    0.087600   -2.381129
    9.C         2.129828    0.305133   -1.300534
   10.N         1.305921   -0.681032   -0.855962
   11.N         4.730897    0.966737   -3.939152
   12.O         1.184926    1.701304    0.390447
   13.C         2.437509   -2.921593    4.885147
   14.C         2.064950   -1.925891    3.831134
   15.C         1.191833   -1.964906    2.765085
   16.N         2.627534   -0.660128    3.769787
   17.C         2.096791    0.004563    2.707069
   18.N         1.223016   -0.765627    2.083426
   19.C        -2.230928    2.201720   -2.454078
   20.C        -1.351280    1.888106   -1.400375
   21.C        -2.808029   -2.616503   -1.163603
   22.C        -3.702722   -2.393864   -2.231841
   23.C        -3.058864    1.218246   -2.957816
   24.C        -2.014977   -1.583542   -0.640540
```

```
   25.C        -3.852034   -1.160544   -2.850865
   26.C        -3.039884   -0.083284   -2.396887
   27.C        -2.130468   -0.307358   -1.314166
   28.N        -1.303268    0.674198   -0.864963
   29.N        -4.729238   -0.951703   -3.959357
   30.O        -1.201994   -1.707004    0.382077
   31.Ni       -0.013287   -0.005522    0.576671
   32.C        -2.427966    2.938778    4.892530
   33.C        -2.051687    1.931770    3.850598
   34.C        -1.180547    1.961229    2.782162
   35.N        -2.614942    0.665601    3.801692
   36.C        -2.087768   -0.007876    2.742510
   37.N        -1.214771    0.756352    2.110500
   38.H         2.246748   -3.204075   -2.869882
   39.H         0.667883   -2.626945   -0.986157
   40.H         2.726432    3.605812   -0.691999
   41.H         4.304043    3.239884   -2.555113
   42.H         3.722499   -1.414158   -3.794235
   43.H         5.203585    1.831942   -4.193106
   44.H         5.443249    0.261482   -3.746542
   45.H         3.191272   -2.508255    5.568697
   46.H         1.569987   -3.213524    5.492666
   47.H         2.859985   -3.835468    4.446808
   48.H         0.543447   -2.769587    2.440290
   49.H         3.318641   -0.290485    4.410766
   50.H         2.346571    1.016193    2.412637
   51.H        -2.229077    3.207099   -2.872524
   52.H        -0.653711    2.616459   -0.990083
   53.H        -2.740579   -3.607140   -0.715761
   54.H        -4.311256   -3.230525   -2.582830
   55.H        -3.712952    1.427587   -3.803558
   56.H        -5.210097   -1.812993   -4.211642
   57.H        -5.435512   -0.240471   -3.765605
   58.H        -3.184399    2.533601    5.578505
   59.H        -1.563133    3.239026    5.499372
   60.H        -2.849143    3.847595    4.441786
   61.H        -0.532818    2.762715    2.448374
   62.H        -3.307134    0.302855    4.445567
   63.H        -2.344242   -1.019720    2.454838
END

GEOMETRY
 GO
 ITERATIONS 100
END

BASIS
 type TZ2P
 core none
END

END INPUT"
```

A4.3.4. Input file for *trans*-Quin OLYP

```
$ADFBIN/adf -n8 \
 <<< "
TITLE transquinolyp geo opt

MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE 0 2
UNRESTRICTED

SCF
 ITERATIONS 200
DIIS
END

XC
 GGA OLYP
END

SYMMETRY NOSYM
ATOMS
    1.C         2.660094    3.296980    1.484037
    2.C         1.738059    2.579749    0.700783
    3.C         2.257117   -1.394138    3.241701
    4.C         3.190055   -0.752434    4.067420
    5.C         3.266868    2.666845    2.550695
    6.C         1.637437   -0.725907    2.166684
    7.C         3.566179    0.574659    3.901047
    8.C         2.969875    1.309970    2.838435
    9.C         2.010745    0.664650    1.979857
   10.N         1.409857    1.313495    0.941771
   11.N         4.549954    1.139830    4.797249
   12.O         0.789654   -1.284054    1.361309
   13.C         4.271415   -0.890660   -3.897429
   14.C         3.237762   -0.976303   -2.823462
   15.C         2.336222   -0.066632   -2.319924
   16.N         2.995539   -2.136439   -2.102367
   17.C         1.993737   -1.896444   -1.214652
   18.N         1.575301   -0.648081   -1.325296
   19.C        -1.514654    3.071739   -2.910007
   20.C        -0.789442    2.461934   -1.868925
   21.C        -2.807097   -1.762954   -2.574869
   22.C        -3.518142   -1.249418   -3.672321
   23.C        -2.374046    2.303150   -3.666783
   24.C        -1.881997   -0.980864   -1.857768
   25.C        -3.392434    0.065663   -4.104355
   26.C        -2.519290    0.923087   -3.383582
   27.C        -1.753039    0.394785   -2.285449
   28.N        -0.893921    1.172044   -1.563133
   29.N        -4.157770    0.600180   -5.188135
   30.O        -1.179946   -1.420262   -0.854001
   31.Ni        0.039236    0.014218   -0.039553
   32.C        -3.549225    1.892561    4.077658
```

```
   33.C        -2.831389    1.142506    3.004874
   34.C        -1.781773    1.471748    2.177268
   35.N        -3.173047   -0.148092    2.630069
   36.C        -2.347733   -0.544675    1.624896
   37.N        -1.495006    0.420915    1.328199
   38.H        -3.565574    0.928413   -5.952346
   39.H         5.331708    1.546791    4.280578
   40.H         4.151660    1.885909    5.369963
   41.H        -2.981103    2.787696    4.353740
   42.H        -4.550982    2.216638    3.755446
   43.H        -3.671025    1.284460    4.986371
   44.H        -3.906790   -0.712017    3.040877
   45.H        -2.378103   -1.508135    1.133792
   46.H        -1.213363    2.393423    2.153700
   47.H         1.997614   -2.437500    3.413017
   48.H         3.658199   -1.306046    4.881676
   49.H         3.981776    3.204932    3.173614
   50.H         2.881639    4.334600    1.237388
   51.H         1.260325    3.058790   -0.148918
   52.H        -0.119286    3.059263   -1.257207
   53.H        -1.400845    4.139580   -3.090999
   54.H        -2.971116    2.752330   -4.459387
   55.H        -4.199790   -1.913065   -4.210139
   56.H        -2.947632   -2.799467   -2.272773
   57.H         5.234607   -1.308101   -3.568100
   58.H         3.966793   -1.429609   -4.807801
   59.H         4.441946    0.156503   -4.171081
   60.H         2.188625    0.963858   -2.619489
   61.H         1.610311   -2.615600   -0.503834
   62.H         3.492001   -3.013268   -2.203625
   63.H        -4.790559   -0.102266   -5.565606
END

GEOMETRY
 GO
 ITERATIONS 100
 FREQUENCIES
END

BASIS
 type TZ2P
 core none
END

END INPUT
```

Section A4.4. Results From DFT Calculations

A4.4.1. Optimized cartesian coordinates and molecular structure for *cis*-Quin BP86

1.C	2.243672	-2.200123	-2.447739
2.C	1.360933	-1.893156	-1.394585
3.C	2.799438	2.616302	-1.141645
4.C	3.698600	2.399606	-2.207384
5.C	3.067748	-1.211423	-2.947351
6.C	2.005525	1.580376	-0.626074
7.C	3.853038	1.168788	-2.829998
8.C	3.043509	0.087600	-2.381129
9.C	2.129828	0.305133	-1.300534
10.N	1.305921	-0.681032	-0.855962
11.N	4.730897	0.966737	-3.939152
12.O	1.184926	1.701304	0.390447
13.C	2.437509	-2.921593	4.885147
14.C	2.064950	-1.925891	3.831134
15.C	1.191833	-1.964906	2.765085
16.N	2.627534	-0.660128	3.769787
17.C	2.096791	0.004563	2.707069
18.N	1.223016	-0.765627	2.083426
19.C	-2.230928	2.201720	-2.454078
20.C	-1.351280	1.888106	-1.400375
21.C	-2.808029	-2.616503	-1.163603
22.C	-3.702722	-2.393864	-2.231841
23.C	-3.058864	1.218246	-2.957816
24.C	-2.014977	-1.583542	-0.640540
25.C	-3.852034	-1.160544	-2.850865
26.C	-3.039884	-0.083284	-2.396887
27.C	-2.130468	-0.307358	-1.314166
28.N	-1.303268	0.674198	-0.864963
29.N	-4.729238	-0.951703	-3.959357
30.O	-1.201994	-1.707004	0.382077
31.Ni	-0.013287	-0.005522	0.576671
32.C	-2.427966	2.938778	4.892530
33.C	-2.051687	1.931770	3.850598
34.C	-1.180547	1.961229	2.782162
35.N	-2.614942	0.665601	3.801692
36.C	-2.087768	-0.007876	2.742510
37.N	-1.214771	0.756352	2.110500
38.H	2.246748	-3.204075	-2.869882
39.H	0.667883	-2.626945	-0.986157
40.H	2.726432	3.605812	-0.691999
41.H	4.304043	3.239884	-2.555113
42.H	3.722499	-1.414158	-3.794235
43.H	5.203585	1.831942	-4.193106
44.H	5.443249	0.261482	-3.746542
45.H	3.191272	-2.508255	5.568697
46.H	1.569987	-3.213524	5.492666
47.H	2.859985	-3.835468	4.446808
48.H	0.543447	-2.769587	2.440290
49.H	3.318641	-0.290485	4.410766
50.H	2.346571	1.016193	2.412637

51.H	-2.229077	3.207099	-2.872524
52.H	-0.653711	2.616459	-0.990083
53.H	-2.740579	-3.607140	-0.715761
54.H	-4.311256	-3.230525	-2.582830
55.H	-3.712952	1.427587	-3.803558
56.H	-5.210097	-1.812993	-4.211642
57.H	-5.435512	-0.240471	-3.765605
58.H	-3.184399	2.533601	5.578505
59.H	-1.563133	3.239026	5.499372
60.H	-2.849143	3.847595	4.441786
61.H	-0.532818	2.762715	2.448374
62.H	-3.307134	0.302855	4.445567
63.H	-2.344242	-1.019720	2.454838

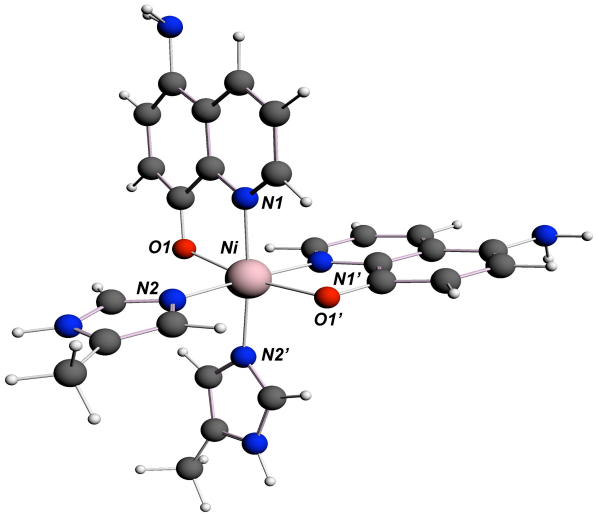


Figure A4.4.2. Optimized geometry of the model *cis*-Quin BP86

Table A4.4.3. Calculated bond distances for the crystallographically determined HQuin1 and the computationally minimized *cis*-Quin BP86 model.

Bond (Å)	HQuin1	Model	% Difference
Ni-O1	2.1(1)	2.08	2.4
Ni-N1	2.1(1)	2.05	0.9
Ni-N2	2.0(1)	2.09	5.4
Ni-O1’	2.1(1)	2.09	2.4
Ni-N1’	2.1(1)	2.06	1.4
Ni-N2’	2.0(1)	2.09	4.4

Table A4.4.4. Calculated bond angles for the crystallographically determined HQuin1 and the computationally minimized *cis*-Quin BP86 model.

Angle (°)	HQuin1	Model	% Difference
O1-Ni-N1	82(1)	81.13	1.1
O1-Ni-N2	95(1)	92.17	3.0
O1-Ni-O1’	165(1)	169.50	2.7
O1-Ni-N1’	87(1)	91.84	5.4
O1-Ni-N2’	94(1)	96.16	2.3
N1-Ni-N2	91(1)	91.93	1.0
N1-Ni-O1’	87(1)	91.56	5.1
N1-Ni-N1’	91(1)	91.27	0.3
N1-Ni-N2’	175.1)	176.96	1.1
N2-Ni-O1’	95(1)	95.61	0.6
N2-Ni-N1’	176(1)	175.20	0.5
N2-Ni-N2’	86(1)	86.78	0.9

A4.4.5. Optimized cartesian coordinates and molecular structure for *trans*-Quin BP86

1.C	2.660094	3.296980	1.484037
2.C	1.738059	2.579749	0.700783
3.C	2.257117	-1.394138	3.241701
4.C	3.190055	-0.752434	4.067420
5.C	3.266868	2.666845	2.550695
6.C	1.637437	-0.725907	2.166684
7.C	3.566179	0.574659	3.901047
8.C	2.969875	1.309970	2.838435
9.C	2.010745	0.664650	1.979857
10.N	1.409857	1.313495	0.941771
11.N	4.549954	1.139830	4.797249
12.O	0.789654	-1.284054	1.361309
13.C	4.271415	-0.890660	-3.897429
14.C	3.237762	-0.976303	-2.823462
15.C	2.336222	-0.066632	-2.319924
16.N	2.995539	-2.136439	-2.102367
17.C	1.993737	-1.896444	-1.214652
18.N	1.575301	-0.648081	-1.325296

19.C	-1.514654	3.071739	-2.910007
20.C	-0.789442	2.461934	-1.868925
21.C	-2.807097	-1.762954	-2.574869
22.C	-3.518142	-1.249418	-3.672321
23.C	-2.374046	2.303150	-3.666783
24.C	-1.881997	-0.980864	-1.857768
25.C	-3.392434	0.065663	-4.104355
26.C	-2.519290	0.923087	-3.383582
27.C	-1.753039	0.394785	-2.285449
28.N	-0.893921	1.172044	-1.563133
29.N	-4.157770	0.600180	-5.188135
30.O	-1.179946	-1.420262	-0.854001
31.Ni	0.039236	0.014218	-0.039553
32.C	-3.549225	1.892561	4.077658
33.C	-2.831389	1.142506	3.004874
34.C	-1.781773	1.471748	2.177268
35.N	-3.173047	-0.148092	2.630069
36.C	-2.347733	-0.544675	1.624896
37.N	-1.495006	0.420915	1.328199
38.H	-3.565574	0.928413	-5.952346
39.H	5.331708	1.546791	4.280578
40.H	4.151660	1.885909	5.369963
41.H	-2.981103	2.787696	4.353740
42.H	-4.550982	2.216638	3.755446
43.H	-3.671025	1.284460	4.986371
44.H	-3.906790	-0.712017	3.040877
45.H	-2.378103	-1.508135	1.133792
46.H	-1.213363	2.393423	2.153700
47.H	1.997614	-2.437500	3.413017
48.H	3.658199	-1.306046	4.881676
49.H	3.981776	3.204932	3.173614
50.H	2.881639	4.334600	1.237388
51.H	1.260325	3.058790	-0.148918
52.H	-0.119286	3.059263	-1.257207
53.H	-1.400845	4.139580	-3.090999
54.H	-2.971116	2.752330	-4.459387
55.H	-4.199790	-1.913065	-4.210139
56.H	-2.947632	-2.799467	-2.272773
57.H	5.234607	-1.308101	-3.568100
58.H	3.966793	-1.429609	-4.807801
59.H	4.441946	0.156503	-4.171081
60.H	2.188625	0.963858	-2.619489
61.H	1.610311	-2.615600	-0.503834
62.H	3.492001	-3.013268	-2.203625
63.H	-4.790559	-0.102266	-5.565606

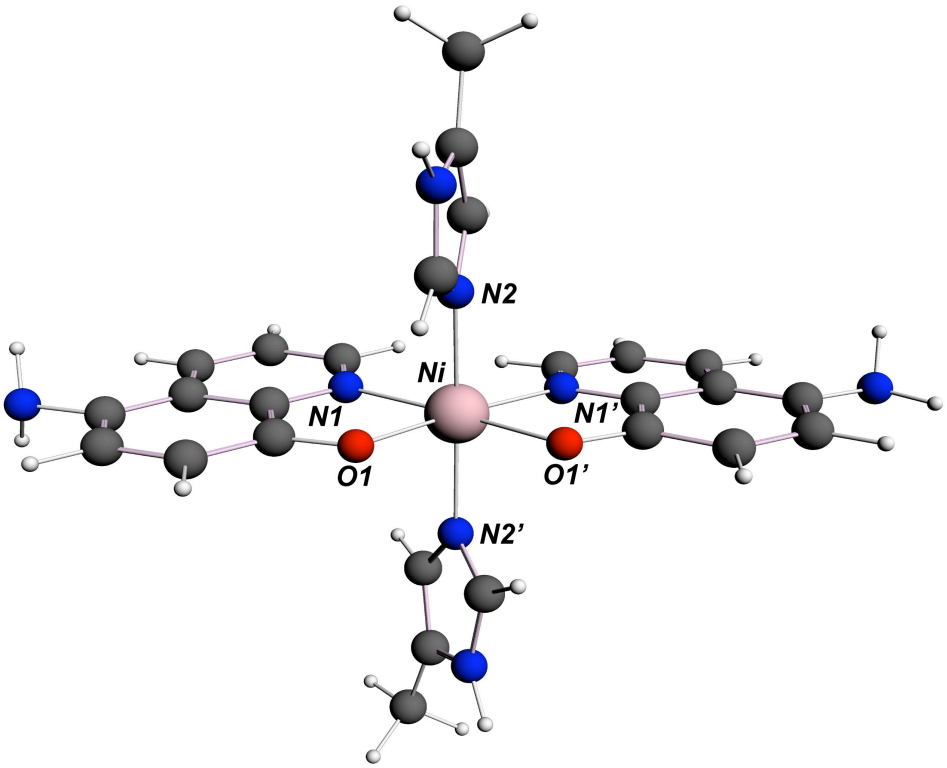


Figure A4.4.6. Optimized geometry of the model *trans*-Quin BP86.

Table A4.4.7. Calculated bond distances for the crystallographically determined HQuin1 and the computationally minimized *trans*-Quin BP86 model.

Bond (Å)	HQuin1	Model
Ni-O1	2.1(1)	2.05
Ni-N1	2.1(1)	2.12
Ni-N2	2.0(1)	2.11
Ni-O1’	2.1(1)	2.05
Ni-N1’	2.1(1)	2.13
Ni-N2’	2.0(1)	2.09

Table A4.4.8 Calculated bond angles for the crystallographically determined Ni:HQuin1$_2$ and the computationally minimized *trans*-Quin BP86 model.

Angle (°)	HQuin1	Model
O1-Ni-N1	82(1)	80.56
O1-Ni-N2	95(1)	87.23
O1-Ni-O1’	165(1)	92.23
O1-Ni-N1’	87(1)	173.16
O1-Ni-N2’	94(1)	86.84
N1-Ni-N2	91(1)	90.22
N1-Ni-O1’	87(1)	173.11
N1-Ni-N1’	91(1)	106.26
N1-Ni-N2’	175.1)	92.99
N2-Ni-O1’	95(1)	88.35
N2-Ni-N1’	176(1)	93.07
N2-Ni-N2’	86(1)	172.70

A4.4.9 Optimized Cartesian coordinates and molecular structure for *cis*-Quin OLYP.

1.C	2.324995	-2.177402	-2.478191
2.C	1.430639	-1.872074	-1.437539
3.C	2.867517	2.624964	-1.150409
4.C	3.791858	2.412759	-2.186910
5.C	3.155835	-1.191411	-2.965747
6.C	2.059987	1.586694	-0.668409
7.C	3.955621	1.189069	-2.817221
8.C	3.129550	0.107263	-2.401356
9.C	2.193538	0.318493	-1.339842
10.N	1.363031	-0.663660	-0.903398
11.N	4.872934	1.005640	-3.896680
12.O	1.223783	1.702959	0.319531
13.C	2.462496	-2.955878	4.946702
14.C	2.090652	-1.962924	3.890640
15.C	1.209821	-1.996593	2.833081
16.N	2.658789	-0.705850	3.817510
17.C	2.125453	-0.047926	2.757788
18.N	1.240713	-0.804782	2.142232
19.C	-2.321022	2.180219	-2.489718

20.C	-1.433889	1.875469	-1.442589
21.C	-2.874521	-2.620583	-1.161013
22.C	-3.794273	-2.408355	-2.201797
23.C	-3.149440	1.194365	-2.981272
24.C	-2.067926	-1.583016	-0.675747
25.C	-3.953595	-1.184690	-2.834169
26.C	-3.127310	-0.103913	-2.415523
27.C	-2.197214	-0.315534	-1.349243
28.N	-1.369674	0.668074	-0.906764
29.N	-4.866099	-1.000228	-3.917340
30.O	-1.236685	-1.698894	0.316490
31.Ni	-0.004617	-0.001512	0.524658
32.C	-2.445798	2.957243	4.956962
33.C	-2.079467	1.959405	3.903715
34.C	-1.201231	1.986651	2.843713
35.N	-2.652612	0.704539	3.834681
36.C	-2.124593	0.041883	2.775329
37.N	-1.238583	0.794535	2.155126
38.H	2.332305	-3.178639	-2.899605
39.H	0.741942	-2.610262	-1.040697
40.H	2.784443	3.605460	-0.689020
41.H	4.410482	3.251841	-2.503542
42.H	3.818956	-1.404636	-3.798553
43.H	5.372840	1.866692	-4.083130
44.H	5.566984	0.294987	-3.684009
45.H	3.226210	-2.552520	5.619777
46.H	1.601225	-3.233789	5.563869
47.H	2.868103	-3.875094	4.511245
48.H	0.556797	-2.800288	2.525495
49.H	3.360096	-0.339139	4.441713
50.H	2.393184	0.956409	2.469082
51.H	-2.324857	3.181201	-2.911998
52.H	-0.748996	2.614604	-1.040664
53.H	-2.795039	-3.600683	-0.698206
54.H	-4.412888	-3.246735	-2.520232
55.H	-3.807176	1.407234	-3.818459
56.H	-5.373118	-1.858109	-4.099187
57.H	-5.554419	-0.282322	-3.710728
58.H	-3.206499	2.557413	5.635667
59.H	-1.581418	3.237663	5.568570
60.H	-2.853113	3.874582	4.519000
61.H	-0.545810	2.787195	2.532916
62.H	-3.354637	0.340952	4.460554
63.H	-2.398686	-0.961382	2.489082

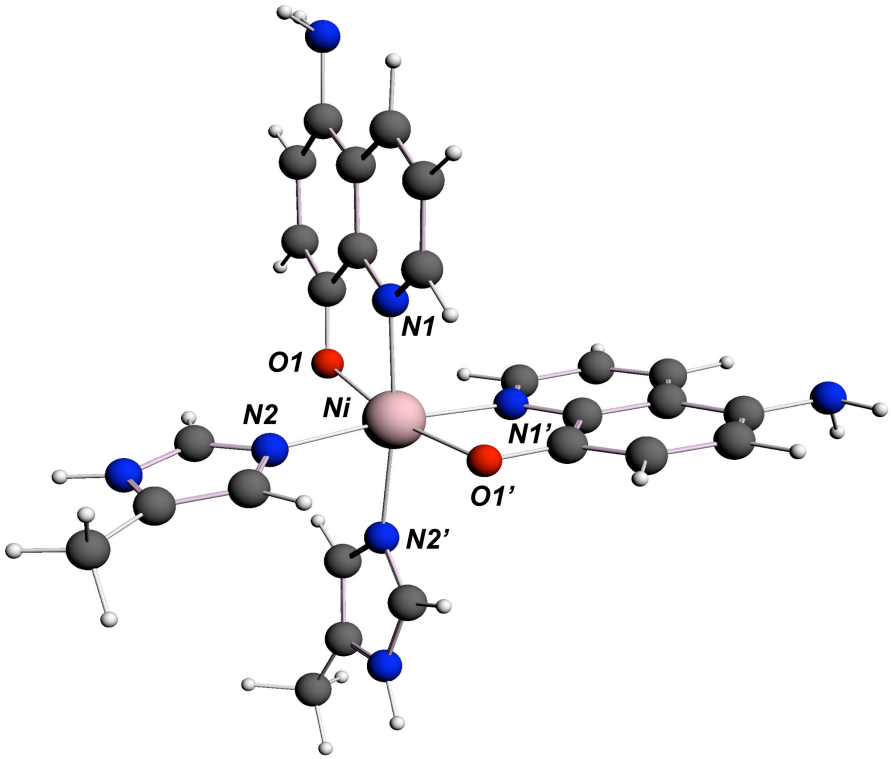


Figure A4.4.10. Optimized geometry of the model *cis*-Quin OLYP.

Table A4.4.11. Calculated bond distances for the crystallographically determined HQuin1 and the computationally minimized *cis*-Quin OLYP model.

Bond (Å)	HQuin1	Model	% Difference
Ni-O1	2.1(1)	2.11	0.5
Ni-N1	2.1(1)	2.09	0.5
Ni-N2	2.0(1)	2.19	9.1
Ni-O1’	2.1(1)	2.11	0.5
Ni-N1’	2.1(1)	2.09	0.5
Ni-N2’	2.0(1)	2.19	9.1

Table A4.4.12. Calculated bond angles for the crystallographically determined HQuin1 and the computationally minimized *cis*-Quin OLYP model.

Angle (°)	HQuin1	Model	% Difference
O1-Ni-N1	82(1)	78.94	3.8
O1-Ni-N2	95(1)	92.12	3.1
O1-Ni-O1’	165(1)	168.75	2.2
O1-Ni-N1’	87(1)	93.15	6.8
O1-Ni-N2’	94(1)	96.11	2.2
N1-Ni-N2	91(1)	90.94	0.1
N1-Ni-O1’	87(1)	93.44	7.1
N1-Ni-N1’	91(1)	93.5	2.7
N1-Ni-N2’	175.1)	173.17	1.1
N2-Ni-O1’	95(1)	96.3	1.4
N2-Ni-N1’	176(1)	173.68	1.3
N2-Ni-N2’	86(1)	84.5	1.8

A4.4.13 Optimized Cartesian coordinates and molecular structure for *trans*-Quin OLYP.

1.C	2.709784	3.303331	1.555393
2.C	1.790571	2.592947	0.767414
3.C	2.289799	-1.389086	3.281070
4.C	3.223172	-0.761191	4.110542
5.C	3.315579	2.668374	2.618011
6.C	1.678716	-0.704496	2.213458
7.C	3.608693	0.563278	3.958343
8.C	3.021263	1.312233	2.900064
9.C	2.062422	0.680296	2.033809
10.N	1.463331	1.328458	0.999943
11.N	4.587202	1.102247	4.871496
12.O	0.830521	-1.240239	1.410901
13.C	4.294914	-0.909563	-3.937319
14.C	3.263568	-0.999800	-2.864372
15.C	2.378459	-0.087877	-2.343424
16.N	3.004521	-2.162412	-2.163782
17.C	2.012228	-1.919896	-1.273139
18.N	1.611231	-0.667479	-1.358327
19.C	-1.570287	3.069186	-2.976251
20.C	-0.839189	2.467589	-1.937393
21.C	-2.818447	-1.768362	-2.624541

22.C	-3.536654	-1.269305	-3.718110
23.C	-2.427927	2.296483	-3.727214
24.C	-1.905327	-0.968248	-1.915853
25.C	-3.432751	0.043553	-4.157461
26.C	-2.568771	0.917531	-3.444059
27.C	-1.794019	0.404310	-2.346175
28.N	-0.940040	1.182888	-1.626097
29.N	-4.215082	0.547238	-5.244270
30.O	-1.200132	-1.389471	-0.920679
31.Ni	0.042965	0.025821	-0.044044
32.C	-3.577651	1.901728	4.131046
33.C	-2.870658	1.154878	3.051283
34.C	-1.834423	1.483623	2.211020
35.N	-3.203905	-0.135919	2.689531
36.C	-2.386046	-0.529861	1.684238
37.N	-1.542476	0.433380	1.369310
38.H	-3.630346	0.859583	-6.014688
39.H	5.383105	1.488017	4.370400
40.H	4.198257	1.862513	5.423793
41.H	-3.031844	2.815987	4.373913
42.H	-4.595760	2.188228	3.836772
43.H	-3.652415	1.309093	5.051278
44.H	-3.930100	-0.700429	3.101667
45.H	-2.427233	-1.499834	1.217143
46.H	-1.276069	2.406935	2.180955
47.H	2.022083	-2.429217	3.443630
48.H	3.682086	-1.324430	4.919682
49.H	4.024580	3.209322	3.238817
50.H	2.930447	4.340400	1.317614
51.H	1.312634	3.078691	-0.074194
52.H	-0.164718	3.067158	-1.338256
53.H	-1.459210	4.133912	-3.161808
54.H	-3.020935	2.747034	-4.517321
55.H	-4.209090	-1.945495	-4.246181
56.H	-2.944182	-2.803002	-2.318431
57.H	5.272032	-1.272730	-3.595223
58.H	4.016743	-1.490338	-4.826308
59.H	4.420960	0.130406	-4.247176
60.H	2.248942	0.945377	-2.629865
61.H	1.625514	-2.655558	-0.587134
62.H	3.482396	-3.042830	-2.276701
63.H	-4.818021	-0.181765	-5.607609

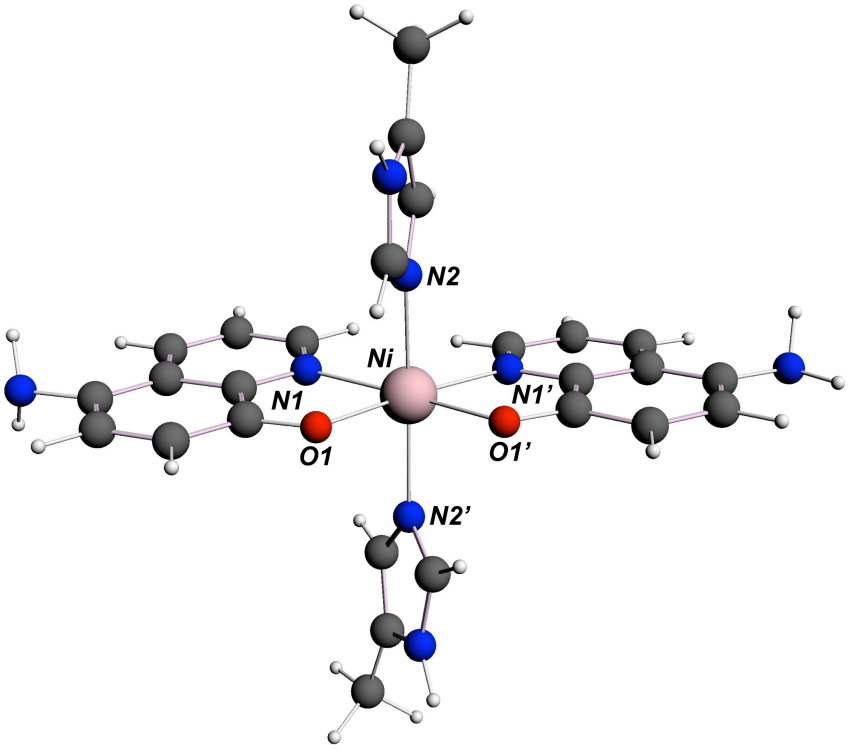


Figure A4.4.14. Optimized geometry of the model *trans*-Quin OLYP

Table A4.4.15. Calculated bond distances for the crystallographically determined HQuin1 and the computationally minimized *trans*-Quin OLYP model.

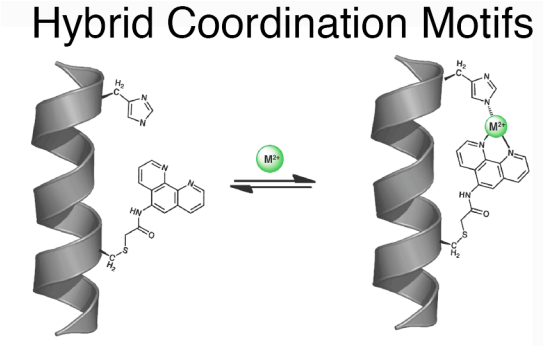
Bond (Å)	HQuin1	Model
Ni-O1	2.1(1)	2.07
Ni-N1	2.1(1)	2.19
Ni-N2	2.0(1)	2.16
Ni-O1’	2.1(1)	2.08
Ni-N1’	2.1(1)	2.19
Ni-N2’	2.0(1)	2.16

Table A4.4.16. Calculated bond angles for the crystallographically determined HQuin1 and the computationally minimized *trans*-Quin OLYP model.

Angle (°)	HQuin1	Model
O1-Ni-N1	82(1)	77.69
O1-Ni-N2	95(1)	87.65
O1-Ni-O1’	165(1)	96.14
O1-Ni-N1’	87(1)	173.46
O1-Ni-N2’	94(1)	88.02
N1-Ni-N2	91(1)	93.2
N1-Ni-O1’	87(1)	173.11
N1-Ni-N1’	91(1)	108.69
N1-Ni-N2’	175.1)	92.49
N2-Ni-O1’	95(1)	87.44
N2-Ni-N1’	176(1)	90.87
N2-Ni-N2’	86(1)	171.96

Chapter 5: Modular and Versatile Hybrid Coordination Motifs on a-Helical Protein Surfaces

5.1 Abstract



Reported here is the construction of phenanthroline (Phen) and terpyridine (Terpy) based hybrid coordination motifs (HCMs), which were installed on the surface of the four-helical bundle hemeprotein, cytochrome cb_{562}. The resulting constructs, termed HPhen1, HPhen2, HPhen3 and HTerpy1, feature HCMs that are composed of a histidine ligand and a Phen or Terpy functionality located two helix turns away, yielding stable tri- or tetradentate coordination platforms. Characterization of the tridentate HCMs indicates that they accommodate many divalent metal ions (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}) with nanomolar to femtomolar affinities, lead to significant stabilization of the α-helical protein scaffold through metal-mediated crosslinking, assert tight control over protein dimerization, and provide stable and high-affinity binding sites for substitution-inert metal probes. Analyses suggest that such tridentate HCMs may be used modularly on any α-helical protein surface in a sequence-independent fashion.

5.2 Introduction

The primary biological roles of metals including catalysis, electron transfer and structural stabilization are generally established once they are firmly placed within a protein scaffold [1]. Owing to the stability of the resulting complexes, the interactions between metals and the interiors of proteins are readily characterized and have justifiably formed the focus of Bioinorganic Chemistry. One could argue, on the other hand, that metals spend a good majority of their time interacting with protein surfaces, and that such transient, harder-to-characterize interactions carry *in vivo* and *in vitro* consequences that rival those of metal-protein interior interactions. The prevalence of metal-protein surface interactions become especially clear when picturing the behavior of metal ions and complexes within the crowded cellular environment, for example, as they are being passed on from one specific protein (e.g. a metallochaperone [2]) to another, or as they crosslink together multiple proteins whose aggregation may have dire consequences [3]. Similarly, outside the cellular realm, metal-protein surface interactions form the basis of immobilized metal ion affinity chromatography (IMAC [4]) as well as the functionalization of protein surfaces with metal complexes that have served as invaluable spectroscopic and functional probes [5, 6]. Given such broad importance and utility of metal-protein surface interactions, the need is clear for metal coordinating motifs that would enable a better control of inorganic chemistry on protein surfaces.

The original interest in metal-protein surface interactions stemed from a desire to use metal coordination chemistry to direct protein-protein interactions (PPIs) [7] and protein self-assembly more predictably and readily than computational design approaches. One caveat to utilizing metal coordination to control PPIs is the presence of

numerous metal binding sidechain functionalities on any given protein surface, which bring about the challenge of controlling metal localization. In chapters 2 and 3, attempts were made to circumvent this challenge by introducing a Cys-specific bidentate non-natural chelate (5-iodoacetamido-1,10 phenanthroline, IA-Phen) onto the surface of the four-helix-bundle protein, cytochrome cb_{562}. The resulting constructs, MBPPhen1/2, were found to form unqiue porous frameworks with coordinatively unsaturated metal centers as a result of thermodynamically favorably Phen/protein interactions [8, 9]. While the utility of these bioinorganic frameworks is just beginning to be explored, the approach is not yet generalizable. In an effort to exert more control over metal localization as well as metal-directed protein self-assembly, in chapter 4, another bidentate chelate (5-iodoacetamido-8-hydroxyquinoline, IA-Quin) was attached to a Cys (C70) in combination with a His (H63) located two helix turns away on the cyt cb_{562} surface, yielding the construct HQuin1 (Figure 5.1a) [10]. The resulting *i*/*i*+7 hybrid coordination motif (HCM) was shown to coordinate various divalent metal ions in a tridentate fashion, which led to: 1) high affinity divalent metal binding with dissociation constants (K_d's) ranging from nM to fM, 2) stabilization of the protein scaffold via metal-mediated crosslinking of a two-helix turn segment, and 3) tight control over protein dimerization via an octahedral metal coordination geometry. Several potential applications arise from these advantages including site-selective labeling of proteins with metal probes, improved protein separation with IMAC, stabilization of small helical peptides for pharmaceutical purposes, and manipulation of cellular pathways that depend on protein dimerization.

Given such possibilities and the ease of constructing an HCM via iodoacetamide (IA)-Cys coupling, the present study as sought to determine whether the advantageous properties of the HQuin1 HCM are generalizable, *i.e.*, whether the *i*/*i*+7 HCMs that consist of a His and a non-natural chelating ligand can be utilized in a modular fashion on any α-helical surface. Towards this end, a series of additional cyt cb_{562}-based constructs (Figure 5.1b-e) were created, which have been functionalized with various non-natural chelates (Figure 5.2): a) HPhen1, the phenanthroline (Phen)-derivatized counterpart of HQuin1, was constructed to probe the generality of the non-natural component; b) HPhen2, which features the opposite placement of His and the Cys-Phen group as in HPhen1 (Cys63 in the *i* and His70 in the *i*+7 position), was constructed to test the sensitivity of the *i*/*i*+7 HCM to the relative placement of the natural and non-natural ligands, c) HPhen3, which has the HCM motif located elsewhere on the cyt cb_{562} surface (His70, Cys77), was generated to test the sensitivity of the HCM to location, and, d) HTerpy1, the terpyridine-derivatized counterpart of HQuin1 and HPhen1, as a tetradentate HCM motif. Present here is the characterization of these constructs in terms of their metal binding affinities, metal-dependent stabilization and metal-dependent oligomerization properties (Figure 5.3). These results suggest that the *i*/*i*+7 HCMs may be modularly utilized on any α-helical protein surface towards a number of applications.

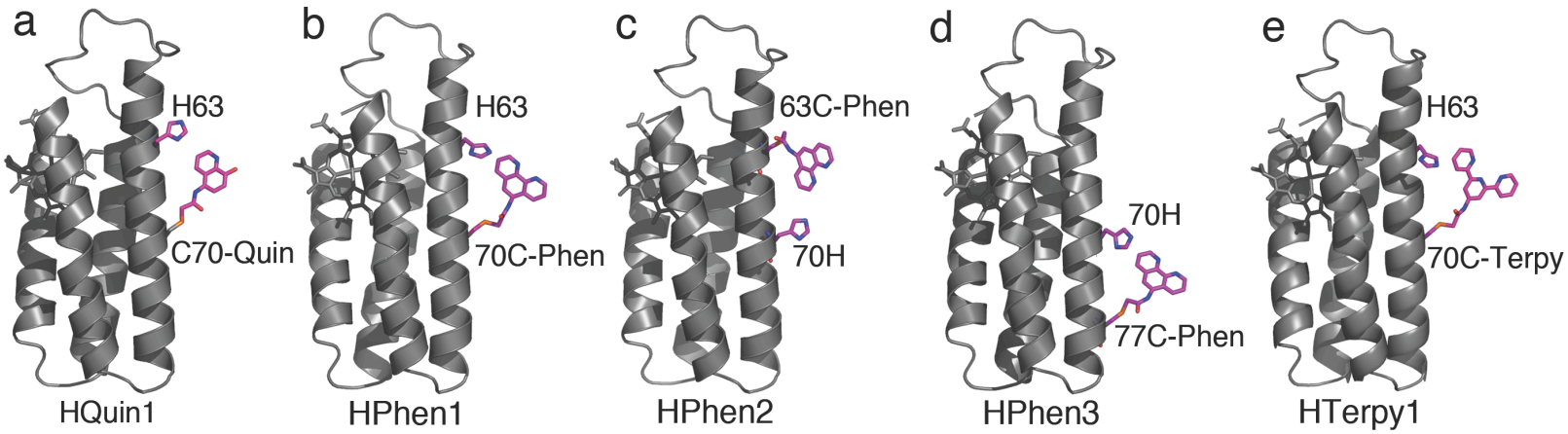


Figure 5.1. Cartoon representations for various HCM-bearing cyt cb_{562} variants. Functionalities that comprise the HCMs and the heme groups are shown as sticks and highlighted in magenta.

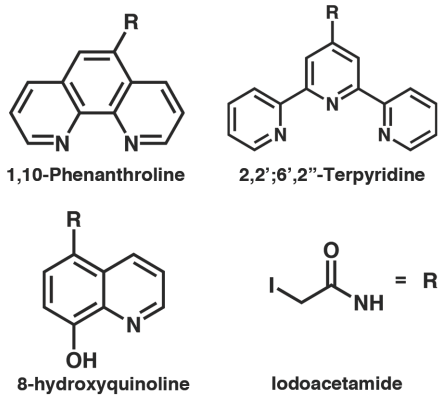


Figure 5.2. Line drawings of the Phenanthroline (Phen), terpyridine (Terpy) and hydroxyquinoline (Quin) derivatives used for the construction of HCMs. The iodoacetamide moiety is attached to the chelating functionalities through the amide nitrogen.

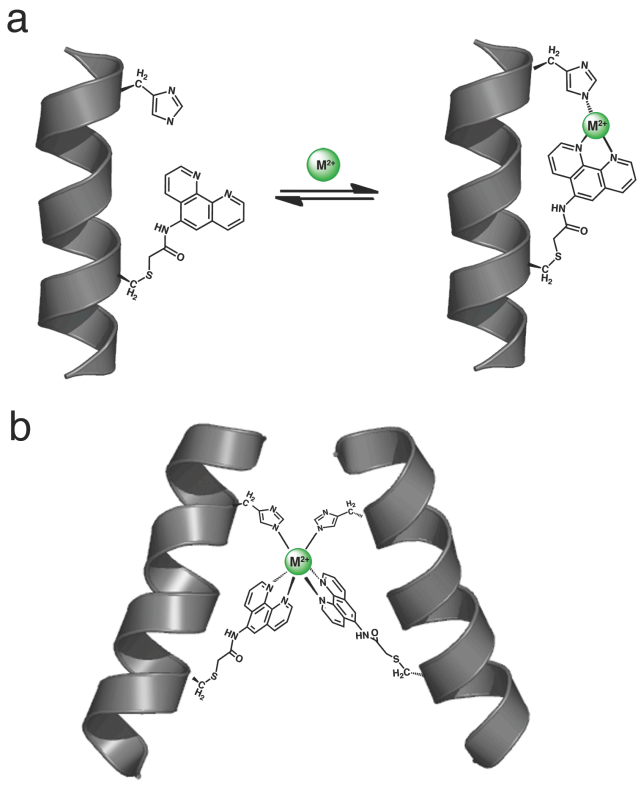


Figure 5.3. (a) Cartoon depiction for the proposed three coordinate *facial*-binding mode of a His-Phen HCM. (b) Proposed mode of metal-dependent dimerization mediated by a *i*/*i*+7 His-Phen HCM.

5.3 Results and Discussion

Construction of cyt cb_{562} variants with Quin, Phen and Terpy-bearing HCMs

One requisite for expanding the biological and chemical utility of HCMs is to demonstrate the modularity of the non-natural metal chelator within the HCM system and the ease of its incorporation. To this end, site-specifically labeled cyt cb_{562} variants were created bearing a single surface Cys residue with iodoacetamide (IA)-derivatized versions of the ubiquitous metal chelators Phen, Terpy and Quin to create the HCM variants shown in Figure 5.1. These non-natural ligands were chosen because they are commercially available as – or easily converted to – amino-functionalized precursors, their coordination chemistry has been extensively studied, and they represent a small but diverse set if ligands with variations in denticity and overall charge.

The amino-precursors of Phen, Quin and Terpy were converted in a one-pot reaction with iodoacetic anhydride or iodoacetyl chloride into IA-derivatives with 60-75% yield. Although IA-Phen, IA-Quin and IA-Terpy are sparingly soluble in water, they are easily introduced into cyt cb_{562} solutions after being solubilized in DMF or DMSO; no adverse effects have been found for these organic solvents on cyt cb_{562} up to a final volume fraction of 50% versus H_2O. Cys functionalization reactions proceed rapidly and specifically (provided that the solution pH is kept below 8 to prevent Lys labeling), with overall yields of modification ranging from 60% for IA-Terpy to 95% for IA-Phen after purification. In the case of cyt cb_{562}, the functionalized products are facilely separated from non-functionalized protein using anion-exchange chromatography (Figure A5.1.1).

Metal Binding Properties of Phen- and Terpy-based HCMs

The divalent metal-binding properties of HQuin1 were previously examined and confirmed that the *i*/*i*+7 His-Quin HCM was able to coordinate metals in a facial, tridentate geometry [10]. Here, similar metal-binding titrations for HPhen1, HPhen2 and HPhen3 using late first-row transition metals (Co^{2+}, Ni^{2+}, Cu^{2+}, Zn^{2+}) have been performed, to probe if the Phen functionality behaves similarly to Quin in the context of an HCM. It is important to note that the relative positions of the coordinating atoms to the point of protein attachment in the Phen derivative are equivalent to those in the Quin derivative (Figure 5.2). Metal binding by the His-Phen HCMs were monitored by the distinct 10-nm redshift in the π-π^* absorption band for Phen (metal-free λ_{max}= 272 nm; metal-bound λ_{max}= 282 nm) upon metal binding (Figure 5.4). It was confirmed through CD spectroscopy that the α-helical fold is not significantly perturbed by metal binding to the HCMs (Figure A5.1.19).

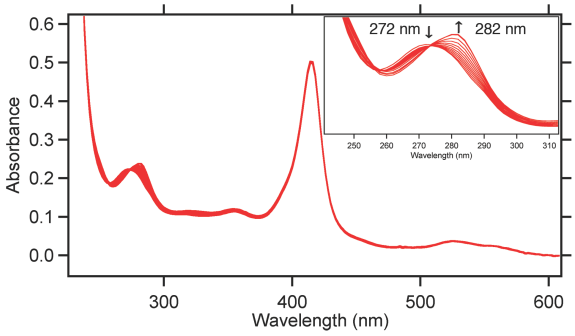


Figure 5.4. Spectral changes that accompany Zn^{2+} binding to HPhen1, as monitored by UV-vis spectroscopy. Spectra show a typical ferric heme spectrum with a Soret band at 415 nm (ε = 0.148 μM^{-1} cm^{-1}) along with a transition between metal-free (λ_{max}= 272 nm) and metal-bound (λ_{max}= 282 nm) Phen species. Inset: Close-up view of the UV region showing a clean isobestic point at 274 nm consistent with a 1:1 Zn^{2+}:Phen binding model.

As in the case of HQuin1, it was quickly established that Phen-based HCMs bind all tested divalent metals very tightly, which required all titrations to be performed in the presence of ethylene glycol tetraacetic acid (EGTA) as a competing ligand. Due to the inherent ability of Phen-based HCMs to undergo metal-mediated dimerization (Figure 5.3b), protein concentrations were kept sufficiently low (< 5 µM) to minimize dimer formation. In all cases (HPhen1-3 and all metals), the metal binding isotherms were satisfactorily described by a 1:1 binding model (Figures 5.5, A5.1.2 and A5.1.3).

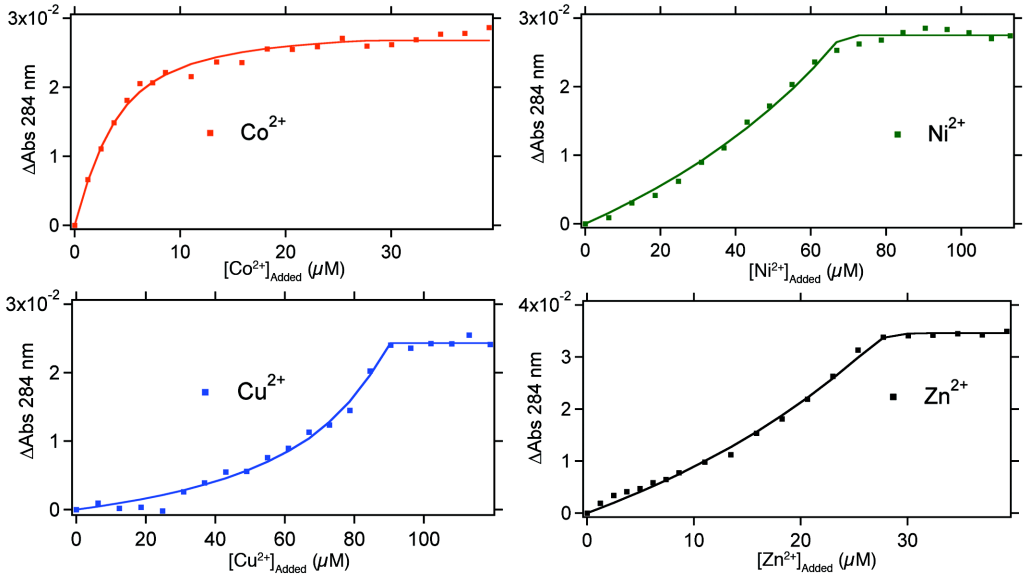


Figure 5.5. Metal-binding titration data and fits for HPhen1 as monitored by UV-visible spectroscopy. A typical titration sample contained 2-5 µM of HPhen1, 50 mM MOPS buffer (pH 7) and 20-100 µM of the competing ligand EGTA. A 1:1 binding model described all data satisfactorily. Binding isotherms for HPhen2, HPhen3 are shown in Figure A5.1.2 and A5.1.3. Dissociation constants (K_d) determined are listed in Table 5.1 and A5.1.

An analysis of the determined dissociation constants (Table 5.1) reveals that all Phen based HCMs display a significant increase in affinity over free Phen for the late first row transition metals, which strongly suggests the participation of the His component of the

HCMs in metal binding. Moreover, the affinities of HPhen1, 2 and 3 for divalent metals are similar and vary at most by six-fold, indicating that metal binding ability is not very sensitive to helix location or relative orientation of the HCM (see below for a discussion on the possible effects of intervening residues).

Table 5.1. Dissociation constants for HPhen-metal complexes compared to those for free 1,10-phenanthroline (Phen).

	Dissociation Constants (M)			
	HPhen1[a]	HPhen2[a]	HPhen3[a]	Free Phen[b]
Co^{2+}	$3.3\ (2) \times 10^{-10}$	$7.8\ (5) \times 10^{-10}$	$2.0\ (2) \times 10^{-9}$	8.0×10^{-8}
Ni^{2+}	$7.8\ (6) \times 10^{-10}$	$1.7\ (4) \times 10^{-10}$	$1.6\ (3) \times 10^{-10}$	3.9×10^{-8}
Cu^{2+}	$1.3\ (2) \times 10^{-13}$	$1.2\ (2) \times 10^{-13}$	$2.0\ (2) \times 10^{-13}$	2.5×10^{-9}
Zn^{2+}	$6\ (1) \times 10^{-9}$	$3.8\ (3) \times 10^{-9}$	$3.7\ (3) \times 10^{-8}$	3.9×10^{-7}

(a) Dissociation constants determined by competition with EGTA in 50 mM MOPS (pH 7). (*b*) pH-adjusted values based on reported K_d's [11].

In addition to the HPhen variants, a HTerpy1 construct was also investigated to determine whether (Figure 5.1e) can engage both Terpy and His in a tetradentate coordination motif. Metal binding titrations and sedimentation velocity (SV) experiments reveal that HTerpy1 almost exclusively forms a stable dimer with a saturation point reached upon addition of half an equivalent of M^{2+} (Figure A5.1.4), which has precluded the determination of the HTerpy1 metal binding affinities. While protein unfolding studies (see below) show evidence for metal coordination by both His and Terpy, the

unstrained, facial coordination geometry observed in HQuin1 and HPhen variants cannot be accommodated by the large Terpy group, leading to the formation of the thermodynamically and kinetically stable *bis*-Terpy adduct involving two proteins.

Metal-mediated protein stabilization through Phen- and Terpy-based HCMs

Next, the ability of the Phen- and Terpy-based HCMs to have any stabilizing effect on the protein scaffold was investigated. Since HCMs crosslink a ~7-Å long, two-helix-turn segment of cyt cb_{562} through metal coordination, an increase in the global stability of the protein should be expected. Metal crosslinking of both natural and non-natural residues at *i/i+4* positions has extensively been shown to induce α-helicity in peptides and significantly stabilize helical protein structures [12, 13]. Likewise, covalent cross-linking of sidechain functionalities in *i/i*+4, *i/i*+7 or *i/i*+11 positions can lock small peptides in α-helical conformations [14-16], which in turn have proven to be promising pharmaceutical agents that effectively disrupt protein-protein interactions and exhibit increased resistance to proteases in vivo [17].

In order to investigate the cross-linking ability of HCMs in the presence of metals, chemical and thermal unfolding studies were undertaken. In a typical chemical unfolding experiment, a solution of folded HPhen or HTerpy variant was titrated with increasing amounts of unfolded protein solution prepared in 8 M guanidine hydrochloride (GuHCl). The folding/unfolding transition was followed by CD spectroscopy, monitoring changes in ellipticity at 222 nm. Thermal unfolding measurements spanning 298 to 373 K were similarly monitored at 222 nm; because of the high stability of all variants, 1.5 M GuHCl was included in each sample to ensure that complete unfolding took place before 373 K.

In both chemical and thermal unfolding experiments, metals were present in large excess over protein to ensure full occupancy of HCMs, thereby preventing metal-induced dimerization.

The stability of all HPhen and HTerpy variants tested was found to increase in the presence of divalent metal ions. Figure 5.6 shows representative unfolding titrations of the variants, each of which display a particularly enhanced stability in the presence of Ni^{2+} (for other metals and thermal titrations, see Figures A5.1.5-9); a complete set of results is given in Table 5.2. At least in the case of the HPhen variants, the superior stabilizing effect of Ni^{2+} over other metals was attributed to the formation of an unstrained, facial coordination geometry by the His-Phen HCM, which was previously shown to be the case for the His-Quin HCM.

In order to establish that the observed protein stabilization is due to metal-mediated, intrahelical crosslinking, unfolding titrations of HPhen1 and HTerpy 1 were carried out at pH 5.5, where the His

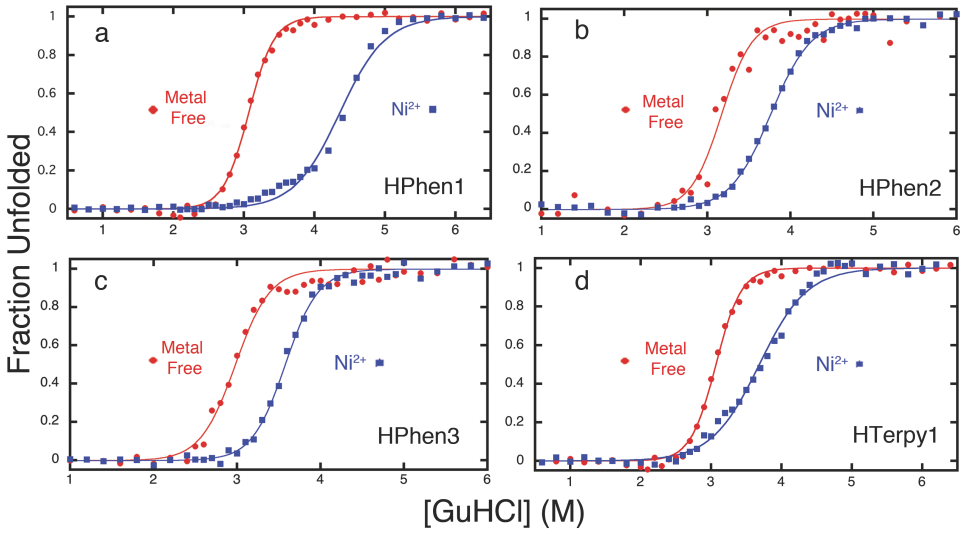


Figure 5.6. GuHCl induced unfolding titrations of (a) HPhen1 (b) HPhen2 (c) HPhen3 and (d) HTerpy1 in the presence, and absence, of Ni^{2+} as monitored by CD spectroscopy at 222 nm.

Table 5.2. Observed changes in the midpoint for the unfolding transition ($\Delta[GuHCl]_m$) for HPhen and HTerpy variants upon metal binding.

	$\Delta[GuHCl]_m$ (M)			
	Co^{2+}	Ni^{2+}	Cu^{2+}	Zn^{2+}
HPhen1	0.84	1.15	0.41	0.78
HPhen2	0.25	0.6	0.22	0.4
HPhen3	0.24	0.6	-0.02	0.29
HTerpy1	0.21	0.47	0.01	0.32
APhen1	0.02	0.12	-0.19	0.01

component of the HCM should be partially protonated and unable to fully coordinate metals (Figures 5.7a and 5.7b). Additional unfolding titrations were performed for variants of HPhen1 and HTerpy1, where either the Phen or the Terpy moiety was replaced by a carboxymethyl group (CM-G70C cyt cb_{562}, Figure A5.1.10) or the His63 residue is mutated to Ala (APhen1 and ATerpy1, Figures 5.7c, 7d and A5.1.11). The results

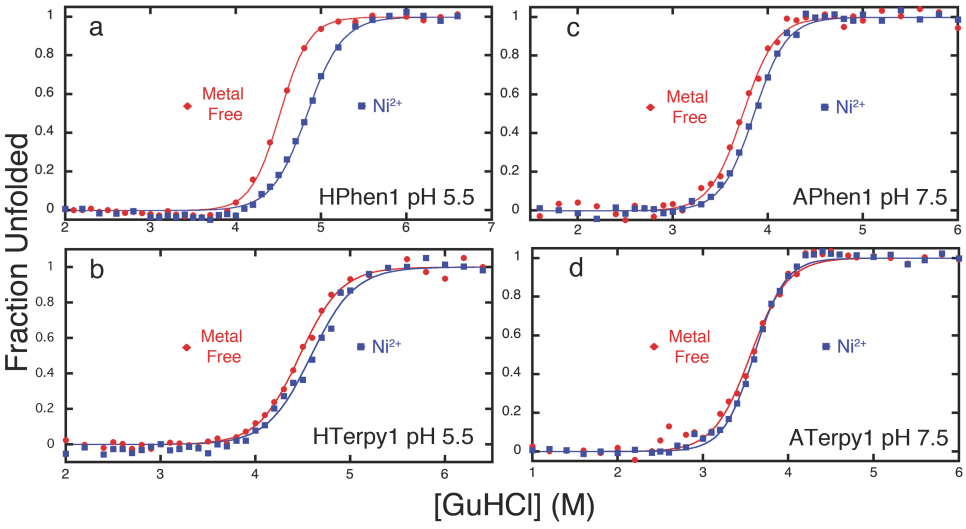


Figure 5.7. GuHCl induced unfolding titrations of (a) HPhen1 at pH 5.5, (b) HTerpy1 at pH 5.5, (c) APhen1 at pH 7.5, and (d) ATerpy1 at pH 7.5. The lack of significant protein stabilization in the presence of Ni^{2+} ions indicates that both His and Phen or Terpy moieties are involved in metal-mediated protein crosslinking.

indicate that, in all cases, Ni-induced stabilization is significantly diminished, confirming the involvement of both His and Phen (or Terpy) in metal coordination.

It is tempting to link the thermodynamics of metal binding by the HCMs (Table 5.1) to that of metal-induced protein stabilization (Table 5.2). Nevertheless, such a correlation is complicated by the fact that net protein stabilization is a function of metal binding not only to the folded but also to the unfolded state, which may display multiple modes of metal coordination (thus deviating from a two-state system). A presentation of free energies of unfolding – which assumes a two-state process – for the variants in the presence of metals, has been avoided. A good case in point is Cu^{2+}, which displays by far the highest affinity for any HCM, yet leads to the smallest extent of stabilization (Table 5.2). Regardless, the protein unfolding titrations clearly indicate that: 1) all Phen-based HCMs lead to a measurable metal-induced increase in protein stability, 2) this stabilization is not specific to a particular HCM location or orientation, and finally, 3) the Terpy-His HCM displays a diminished stabilizing effect due likely to an unfavorable metal coordination geometry.

Metal-Dependent Self-Assembly Properties of HPhen1

The ability to control protein self-assembly, both temporally and spatially, is an intensely pursued goal that is complicated by the necessity to design extensive molecular surfaces [18, 19]. Particularly challenging is to direct the self-assembly of proteins into discrete shapes that can recognize biological targets. One of the most exciting findings about HQuin1 was its ability to specifically dimerize upon Ni^{2+} binding into a rigid architecture that was shaped appropriately to bind major grooves of a double-helical

DNA [10]. Since HPhen1 is the closest in composition and chemical behavior to HQuin1 which has already been structurally characterized, its self-assembly properties were investigated as a representative of all Phen-bearing variants.

Sedimentation velocity (SV) experiments reveal that HPhen1 readily dimerizes in the presence of half an equivalent of Ni^{2+} with a sedimentation coefficient of 2.6 S similar to that of the Ni:HQuin1$_2$ complex (Figure 5.8a) [10]. The dissociation constant for the Ni:HPhen1$_2$ dimer ($K_{d\,(2mer\text{-}1mer)}$) was determined by sedimentation equilibrium (SE) experiments to be ~ 9 µM, which is lower than the $K_{d\,(2mer\text{-}1mer)}$ of 42 µM for Ni:HQuin1$_2$ (Figure 5.8b Figure A5.1.12) [10]. Significantly, the dimeric stability of Ni:Phen1$_2$ now closely approximates that of bZIP family transcription factors which use peripheral leucine zippers domains for dimerization, with K_d's in the low micromolar range [20].

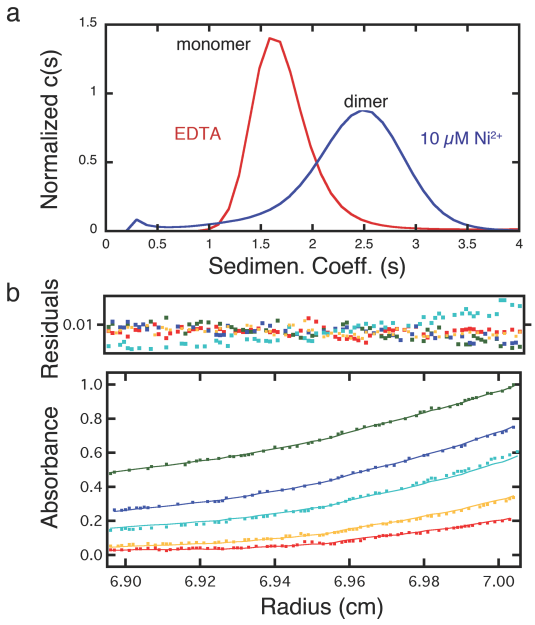


Figure 5.8. (a) Sedimentation velocity profiles of 20 µM HPhen1 in the absence and presence of Ni^{2+}. (b) Sedimentation equilibrium (SE) profiles for 20 µM HPhen1 in the presence of 10 µM Ni^{2+}, with equilibrium speeds of 20,000 (green), 25,000 (blue), 30,000 (cyan), 35,000 (yellow) and 41,000 rpm (red). Data were fit to a monomer-dimer self-association model with a log K = 5.05 (2) or 8.9 (1) µM.

Complete structural characterization of the Ni:HPhen1$_2$ dimer has remained elusive to this point. However, it is proposed – based on the similarities between HPhen1 and HQuin1 and the fact that dimerization in both cases is entirely dictated by metal coordination – that the structure of Ni:HPhen1$_2$ should closely resemble that of the Ni:HQuin1$_2$. In the case of Ni:HQuin1$_2$, it was determined through density functional theory (DFT) calculations that the most favored inner-sphere coordination arrangement would pose the Quin groups cis to one another in Λ configuration, whereby the two phenolate Quin oxygens would lie trans to each other, which is also the crystallographically observed configuration [10].

Similar DFT (BP86 and OLYP) calculations on Ni:HPhen1$_2$ were performed, to investigated the relative energies of two possible inner-sphere arrangements for the His-Phen HCM: one that presents Phen ligands *cis* to one another (*cis*-Phen) and one that presents Phen ligands in a *trans* configuration (*trans*-Phen) (Figure 5.9). These calculations suggest that the *cis*-Phen arrangement is ~5.2 kcal/mol more stable than the *trans*-Phen arrangement. In the case of Ni^{2+}:HQuin1$_2$, the higher stability of *cis*-Quin isomer was attributed to the trans-directing effect of the imine ligands which would render a mutual trans orientation of the weaker-field phenolate ligands the least destabilized configuration. With the *N, O* groups of Quin now replaced with the *N, N* groups in the Phen ligands, this argument cannot be made to explain the higher stability of the *cis*-Phen arrangement. Instead, a close inspection of the calculated structures reveals that in the *trans*-Phen arrangement, there would be considerable steric clashes

between the Phen hydrogens that lie on the Ni equatorial plane, which would be relieved in the *cis*-Phen arrangement.

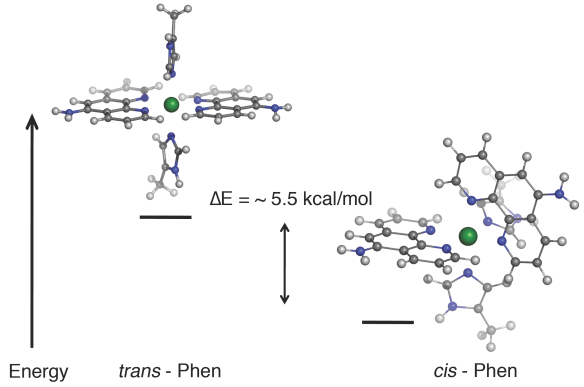


Figure 5.9. Energy minimized structures (BP86) for the proposed inner-sphere coordination geometries of the Ni:HPhen1_2 dimer.

Taking together the DFT results and solution studies, it is conclude that the *i*/*i*+7 His-Phen HCM would yield a Ni^{2+}-induced V-shaped dimer that is equivalent to the crystallographically characterized Ni:HQuin1_2 architecture (Figure 5.10) [10]. Although the reason for the specific formation of this V-shaped structure is different for His-Quin and His-Phen HCMs, both examples demonstrate that self-assembly of proteins can be programmed through a simple consideration of inner-sphere metal coordination, which is far more facile than designing extensive protein interfaces to the same end.

Using Phen-bearing HCMs for protein functionalization

Metal complexes site-specifically attached to protein surfaces have proven to be invaluable functional reporters. Among these, Ru-, Os- and Re-polypyridyl derivatives have been widely used due to their photophysical and photochemical properties [21, 22]. Similarly, bifunctional, As-based fluorescent reporters have been designed to specifically

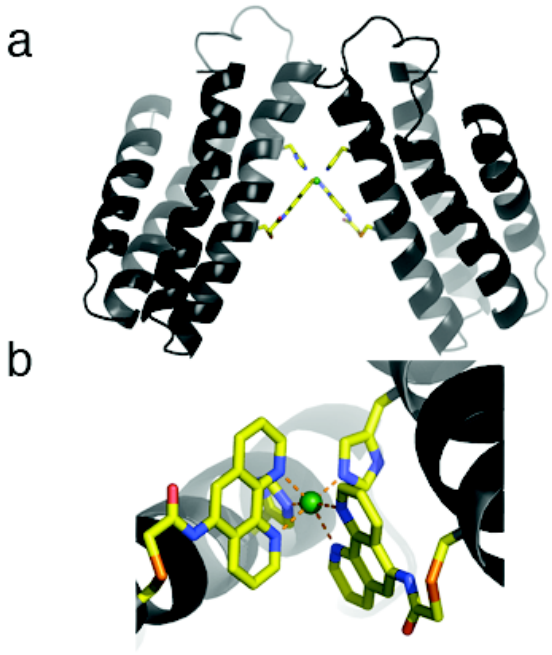


Figure 5.10. (a) The proposed Ni:HPhen1$_2$ architecture modeled after the crystallographically determined Ni:HQuin1$_2$ structure. (b) The corresponding inner coordination sphere.

bind bis-Cys patterns on proteins and are finding increasing use as target selective in vivo reporters [6]. Given the high affinity of Phen-based HCMs for divalent metals and their two-point attachment to the protein scaffold, it is envisioned that they could provide stable and specific target sites for functional metal-based probes on α-helical proteins. Moreover, it is surmised that if such probes are based on substitution-inert metals, they could result in the improved and irreversible stabilization of α-helical proteins/peptides and may be of value in terms of constructing helical peptide-based pharmaceutical agents. To investigate such possibilities, the interactions of HPhen1 with a *p*-cymene-capped Ru(II) compound, were explored. This particular piano-stool complex was chosen as a test case, because the Ru center is capped with an arene group (*p*-cymene), which

should prevent protein dimerization and accommodate facial binding by the His-Phen HCMs. Additionally, it is commercially available in a dimeric, chloro-substituted form (**I**), and weakly luminescent when bound to a polypyridines [23].

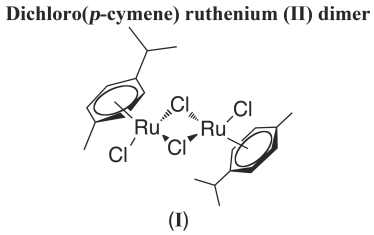


Figure 5.11. Line drawing of dichloro(*p*-cymene) ruthenium (II) dimer.

In a proof-of-principle study, a solution of HPhen1 was treated with 5-fold molar excess of compound **I** (*i.e.*, 10-fold excess Ru) dissolved in DMSO and stirred at room temperature for ~ 4 days. Reactions were quenched by removing unreacted **I** via gel filtration and subsequently purified by ion exchange chromatography. The FPLC chromatogram and corresponding mass spectra indicate that the only major product of the reaction is HPhen1 bound to a single Ru(*p*-cymene) adduct (Figure 5.11), with no discernible unlabeled or multiply labeled species. The absorbance spectra of the isolated Ru(*p*-cymene)(HPhen1) complex features the expected shift to ~286 nm in the Phen π-π^* band due to metal binding as well as a new band at 326 nm (Ru(II) $\rightarrow \pi^*$ arene MLCT) contributed by Ru-adduct (Figure A5.1.13 and A5.1.14). When excited at 326 nm, Ru(*p*-cymene)- HPhen1 displays a weak emission band centered at 442 nm. Both the absorbance (Figure A5.1.14) and the emission (Figure A5.1.15) features of Ru(*p*-cymene)-HPhen1 are similar to those of a highly analogous model complex, [(*p*-cymene)Ru(phen)(1-(4-cyanophenyl)imidazole)], in support of the intended mode of Ru coordination to the His-Phen HCM [23].

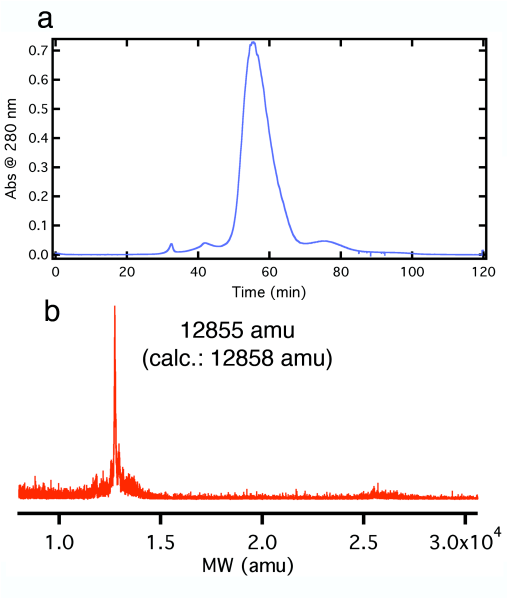


Figure 5.12. (a) Anion-exchange FPLC chromatogram of the crude Ru(*p*-cymene)-HPhen1 reaction mixture. Product eluted at 0.2-0.25 M NaCl using a linear NaCl gradient (0-0.5 M in 10 mM sodium phosphate, pH 8.0). (b) MALDI mass spectra of the major FPLC product identified as the Ru(*p*-cymene)-HPhen1 complex.

Next, the chemical unfolding behavior of the Ru(*p*-cymene)-HPhen1 complex was examined, to study the effects of HCM capping by a substitution-inert metal complex on protein stability. As shown in Figure 5.12, binding of Ru(*p*-cymene) to HPhen1 leads to a significantly higher stabilization compared to substitution-labile divalent metals (Figure 5.6 and Table 5.2), with a corresponding shift in the unfolding midpoint of ~1.5-M GuHCl. Under the reasonable assumption that Ru(*p*-cymene) is still bound to the His-Phen HCM upon denaturation (which is not necessarily the case for labile metals), the unfolding of Ru(*p*-cymene)-HPhen1 can now be treated as a two-state process, allowing the determination of the free energy of stabilization ($\Delta\Delta G_{folding}$) by Ru(*p*-cymene) binding to be 4.1 kcal/mol. The finding that global protein stability can be raised to such an extent by the metal-mediated crosslinking of a local fragment is particularly significant given

that the free energy of unfolding for natural proteins typically ranges from 5 to 15 kcal/mol [24].

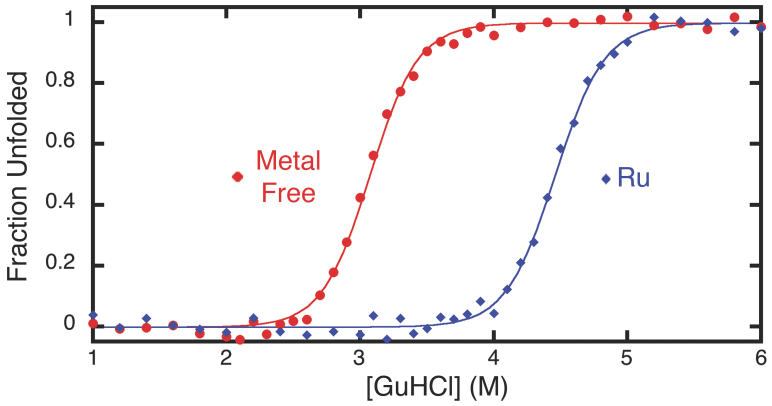


Figure 5.13. Chemical unfolding titrations (monitored by CD spectroscopy) showing the higher stability of the Ru(*p*-cymene)-HPhen1 complex (blue) with respect to HPhen1 in the absence of metals (red).

Effects of Intervening Residues in i/i+7 HCMs

To probe if tridentate *i*/*i*+7 HCMs may be used on any helical protein surface regardless of the amino acid content, a closer look was taken at the structural features of the variants with particular focus on the residues that lie between the coordinating His and the functionalized Cys. Figure 5.13a shows the Ni coordination mode of the His-Quin HCM in the previously determined Ni:HQuin1$_2$ structure, and Figure 5.13b shows the proposed conformation for HPhen1 modeled after the same structure. These structures clearly indicate that the only intervening residues of interest are at the *i*+3 and *i*+4 positions, regardless of the relative positions of His and Cys on the helix (*i.e.*, *i*/*i*+7 or *i*/*i*-7). Importantly, for both *i*+3 and *i*+4 positions (Asp66 and Ile67 for HPhen1 and

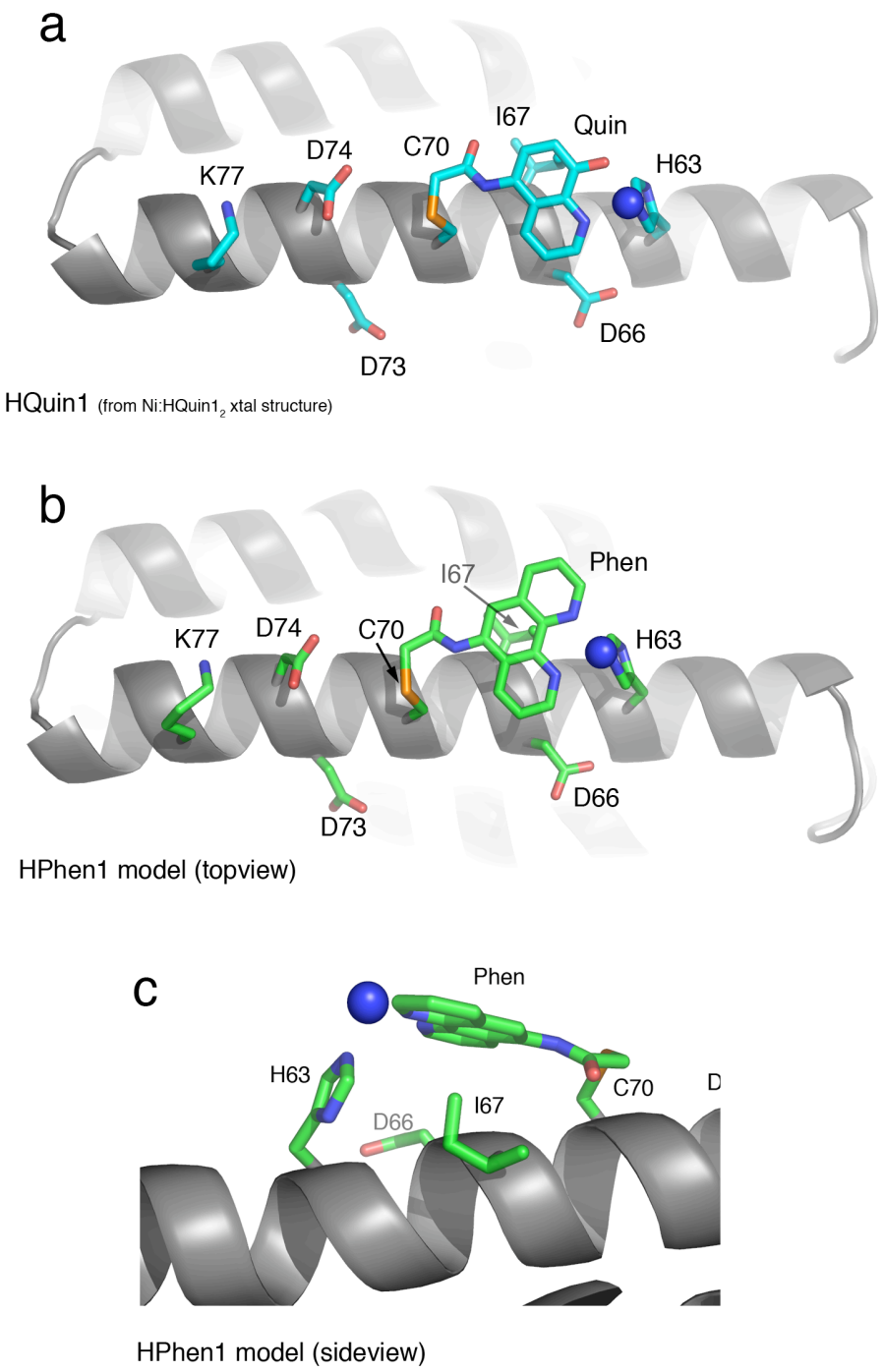


Figure 5.14. (a) Cartoon representation of the HQuin1 structure showing the Ni coordination mode by the *i*/*i*+7 His-Quin HCM. Other residues on Helix3 that are important either for the construction of HPhen3 HCM (positions 70 and 77) or those corresponding to the *i*+3 and *i*+4 positions for all variants are shown also as sticks. (b) The analogous representation of HPhen1 modeled after the HQuin1 structure. (c) Close-up view of the model for His-Phen HCM coordinated to a Ni^{2+}.

HQuin1), the $C_{backbone}$-C_a vectors that largely dictate the orientation of the sidechains are directed away from the coordinating groups. It could then be expected that the *i*/*i*+7 His-Phen or His-Quin HCMs may be universally installed on any regular a-helical surface to coordinate metals without significant interference by the intervening amino acids. This expectation is supported by the finding that HPhen1, HPhen2 and HPhen3 display more or less similar binding affinities for several divalent metal ions (Table 5.1) despite different sets of intervening residues: Asp66/Ile67 for HPhen1, Ile67/Asp66 for HPhen2 (inverse of HPhen1) and Asp73/Asp74 for HPhen3 (Figures 5.13a and b).

At the same time, a close inspection of the HQuin1 structure and the HPhen1 model (Figure 5.13c) shows that the side chain of Ile67 forms van der Waals contacts (d ~ 3.0 Å) with the Quin (or Phen) aromatic ring. These favorable interactions would be absent in the case of HPhen2 or HPhen3, which would respectively present Asp66 or Asp74 near the vicinity of Quin or Phen. Such differential interactions are likely culprits for the lack of any obvious trend in the metal binding affinities of HPhen1, HPhen2 and HPhen3 (Table 5.1). Nevertheless, it is envisioned that the *i*+3 and *i*+4 positions within *i*/*i*+7 HCMs may be exploited as an additional handles to fine-tune metal coordination by HCMs.

5.4 Conclusion

Cumulatively, the studies establish that *i*/*i*+7 HCMs that include a single His and a non-natural bidentate ligand like Phen and Quin can form tridentate chelating platforms on α-helices, extending the scope of coordination chemistry on protein surfaces. Such tridentate HCMs not only provide unprecedented metal binding affinities, but are also

able to stabilize α-helical structures, lead to the formation of discrete oligomers and provide high-affinity attachment sites for metal-based probes. These findings and analyses suggest that these HCMs may be utilized as modular units on any α-helical protein surface in a sequence independent fashion.

5.6 Experimental Section

Materials and Methods

Unless otherwise noted, all solvents, and buffers were purchased from Fisher Scientific or VWR and used without further purification. ACS reagent grade metal salts ($CoCl_2$, $NiSO_4$, $CuSO_4$ and $ZnCl_2$) were purchased from Sigma-Aldrich and used without further purification.

Mass Spectrometry

Protein mass spectrometry was carried out at Biomolecular/Proteomics Mass Spectrometry Facility at UCSD using a Voyager DE-STR MALDI-TOF mass spectrometer. Protein samples (100 μL) were first washed with 3× with 400 μL of nano-pure water (Millipore) using a centrifugal spin column (Millipore) equipped with a 10 KDa cutoff filter. In a typical experiment, 5 μL of a protein sample was mixed in a 1:1 ratio with sinapinic acid (Aligent) as a matrix. 1 μL of the resulting protein/matrix samples was plated on a standard 100 well plate and dried completely before use.

Mass spectrometry (MS) of small molecules was carried out at the Molecular Mass Spectrometry Facility at UCSD using either electrospray ionization (ESI) or an atmospheric pressure chemical ionization (APCI) source on a ThermoFinnigan

LCQDECA mass spectrometer equipped with a quadrupole ion trap mass analyzer and Xcalibur data system. The MS detector was operated under both positive and negative ion modes with a mass resolution range of 100 ppm.

Site Directed Mutagenesis

Site directed mutagenesis was performed has on the pETc-b562 plasmid (denoted as wild-type) [25] using the QuikChange kit (Stratagene) and employing primers obtained from Integrated DNA Technologies. The mutant plasmids were transformed into XL-1 Blue *E. coli* cells and purified using the QIAprep Spin Miniprep kit (Qiagen). Point mutations were introduced to obtain the following cyt cb_{562} variants: G70C-cyt cb_{562}, G70H/H63C-cyt cb_{562}, K77C/G70H/H63A/W59H-cyt cb_{562}, G70C/H63A-cyt cb_{562}. Sequencing of all mutant plasmids was carried out by Retrogen Inc. (San Diego, CA).

General Protein Expression and Purification Protocol

Protein expression and purification was carried out as reported in chapter 2.

Protein purity was determined by SDS-PAGE gel electrophoresis. Verification of mutations was made through MALDI mass spectrometry:

cyt cb_{562} variant	Calc. MW (amu)	Obs. MW (amu)
G70C	12386	12385
H63C/G70H	12386	12386
K77C/G70H/H63A/H59H	12280	12282
H63A/G70C	12320	12320

Synthesis of Iodoacetic anhydride

As a precursor, iodoacetic acid anhydride was freshly prepared by adding 2.64 g (12.8 mmol) of DCC to a stirred solution of 5.0 g (26.8 mmol) iodoacetic acid in 75 mL of ethyl acetate. Dicyclohexylurea precipitates immediately, but the mixture was allowed to stir for 2 h in the dark. The dicyclohexylurea was removed by filtration and the resulting solution was evaporated to dryness and used immediately.

Synthesis of 5-Iodoaceamido-1,10-phenanthroline (IA-Phen) and 5-Iodoacetamido-8-hydroxyquinoline (IA-Quin)

Please see chapters 2 and 4 for the synthetic details and procedures for IA-Phen and IA-Quin, respectively.

Synthesis of 4-Iodoacetamido-2,2':6',2"-terpyridine (IA-Terpy)

4-amino-2,2':6',2"-terpyridine (NH_2-Terpy) was prepared as previously described [26]. 0.1 g (0.4 mmol) of NH_2-Terpy was dissolved in 20 mL of dry dichloromethane. To this solution, 200 μL (1.4 mmol) of triethylamine was added and the reaction mixture was cooled to 0° C by stirring in an ice bath. Once the temperature equilibrated (approx. 20 min), 54 μL (0.6 mmol) of iodoacetyl chloride was added in a dropwise fashion. The mixture was left to react in the dark at 0° C for 30 min and then slowly brought up to room temperature. After a total period of 1 hour, the reaction volume was doubled with dichloromethane and washed extensively with cold 5% sodium bicarbonate followed by water. The organic fractions were evaporated dried in vacuo and used without further purification (Yield: ~60 %). In addition to IA-Terpy, small amounts of the amino

precursor and a chloroacetamide terpyridine adduct were present as impurities. However, since only IA-Terpy can efficiently modify the protein under relevant labeling conditions, the product was used without further purification. Product formation was verified by mass spectrometry. (ESI-MS, positive mode, Figure A5.1.18). Measured MW = 417.05 m/z (exp.: 417.22 m/z) (IA-Terpy + H^+); 325.29 m/z (exp.: 325.08 m/z) (ClA-Terpy + H^+); 249.44 m/z (exp.: 249.11 m/z) (NH_2-Terpy + H^+)

General protocol for functionalization of cyt cb_{562} variants with Phen and Terpy chelates

A solution of 0.3 mM of cyt cb_{562} protein solution in degassed 0.1 M Tris buffer (pH 7.75) was treated with a 10-fold excess of dithiothreitol (DTT) (Sigma). The protein was allowed to reduce for a period of 30 min. The protein was then dialyzed against 2 × 1 L of degassed 0.1 M Tris buffer (pH 7.75) under an inert atmosphere to remove DTT. A 10-fold excess of iodoacetamide label was dissolved in 2 mL of degassed DMF and added drop-wise to the protein solution over the course of 1 min. The mixture was allowed to react in the dark at 25° C overnight. The reaction mixture was then dialyzed again against 2 × 1 L of 10 mM sodium phosphate buffer (pH 8) and 1 mM EDTA. The crude labeled protein was subsequently purified on an Uno-Q anion-exchange column (BioRad) using a sodium chloride gradient. If further purification was necessary, the labeled fractions were combined and dialyzed against 2 × 1 L of 10 mM sodium acetate buffer (pH 5) and 1 mM EDTA. The protein mixture was then purified on an Uno-S (BioRad) cation-exchange column using a sodium chloride gradient. The final purity of

the functionalized protein was determined to be greater than 95% by MALDI mass spectrometry and SDS-PAGE electrophoresis. (Labeling yield: 60-95%).

Cyt cb_{562} variant	Calc. MW (amu)	Obs. MW (amu)
G70C-Phen (HPhen1)	12622	12622
G70H/H63C-Phen (HPhen2)	12622	12620
W59H/H63A/G70H/K77C-Phen (HPhen3)	12516	12513
G70C-Quin (HQuin1)	12589	12590
G70C-Terpy (HTerpy1)	12675	12680
H63A/G70C-Phen (APhen1)	12556	12552

Metal Binding Titrations

Unless otherwise stated, all metal (M^{2+}) binding titrations were performed as outlined in chapter 4.

Chemical and Thermal Unfolding Studies

Unless otherwise stated, all unfolding titrations were performed as outlined in chapter 4.

Sedimentation Velocity Experiments

SV experiments were performed in order to determine the solution-state oligomerization behavior of each HPhen/HTerpy variant. All SV samples were prepared in 20 mM Tris buffer (pH 7). Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor at 41,000 rpm for a total of 250 scans per sample. The following wavelengths were used for detection: 418 nm (5 mM protein), 420 nm (10 μM protein), 425 nm (20 mM protein), 540 nm (40 μM protein), 545 nm (60 μM protein) and 560 nm (100 μM protein).

All data were processed using SEDFIT.[27] Buffer viscosity, buffer density, and protein partial specific volume values were calculated at 25° C with SEDNTERP (http://www.jphilo.mailway.com). Partial specific volume (Vbar) for each variant was calculated assuming a partial specific volume of heme of 0.82 mg/mL and 0.71 mg/mL, 0.75 mg/mL, 0.87 mg/mL for Phen, Quin and Terpy respectively. All data were processed using fixed values for buffer density (r) (0.99764 g/mL) and buffer viscosity (0.0089485 poise).

DFT Calculations

DFT calculations were performed with Amsterdam Density Functional (ADF) program suite [28, 29], version 2007.01 [30], on a home-built 72-CPU (1 × 8 master, 8 × 8 slave) Rocks 4.3 Linux cluster featuring Intel Xeon E5335 Quad-Core 2.00GHz processors. Job control was implemented with the Sun Grid Engine v. 5.3. Crystallographic atomic coordinates were used as input where appropriate. Optimized geometries and molecular orbitals were visualized with the ADFView graphical routine of the ADF-GUI [31] and the Gaussview 3 program.

In ADF program suite calculations, the triple-ζ Slater-type orbital TZ2P ADF basis set was utilized without frozen cores for all atoms. Relativistic effects were included by use of the zero-order regular approximation (ZORA) [32]. To ensure consistency over a range of exchange/correlation profiles, the molecular geometries and energies were evaluated with both the BP86 and OLYP functionals.

In BP86 calculations, the local density approximation (LDA) of Vosko *et al* [33]. (VWN) was coupled with the generalized gradient approximation (GGA) corrections

described by Becke [34] and Perdew [35, 36] for electron exchange and correlation, respectively. In OLYP calculations, the parameterized (X = 0.67) exchange-only LDA was coupled with the GGA corrections described by Handy and Cohen (OTPX) [37] and Lee, Yang and Parr (LYP) [38] for electron exchange and correlation, respectively.

Chapter 5 is reproduced, in part, with permission from: R.J. Radford, P.C. Nguyen, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 7106-7115. Copyright 2010 American Chemical Society

5.7 References

[1] I. Bertini, H.B. Gray, E.I. Stiefel, J.S. Valentine, Biological Inorganic Chemistry, Structure & Reactivity, University Science Books, Sausalito, 2007.

[2] A.C. Rosenzweig, *Acc. Chem. Res.*, 34 (2001) 119-128.

[3] A.I. Bush, W.H. Pettingell, G. Multhaup, M.D. Paradis, J.P. Vonsattel, J.F. Gusella, K. Beyreuther, C.L. Masters, R.E. Tanzi, *Science*, 265 (1994) 1464-1467.

[4] F.H. Arnold, B.L. Haymore, *Science*, 252 (1991) 1796-1797.

[5] J.R. Winkler, H.B. Gray, *Chem. Rev.*, 92 (1992) 369-379.

[6] S.R. Adams, R.E. Campbell, L.A. Gross, B.R. Martin, G.K. Walkup, Y. Yao, J. Llopis, R.Y. Tsien, *J. Am. Chem. Soc.*, 124 (2002) 6063-6076.

[7] (a) Salgado, E. N.; Faraone-Mennella, J.; Tezcan, F. A., J. Am. Chem. Soc. 2007, 129, 13374-13375; (b) Salgado, E. N.; Lewis, R. A.; Faraone-Mennella, J.; Tezcan, F. A., J. Am. Chem. Soc. 2008, 130, 6082-6084; (c) Salgado, E. N.; Lewis, R. A.; Mossin, S.; Rheingold, A. L.; Tezcan, F. A., Inorg. Chem. 2009, 48, 2726-2728.

[8] R.J. Radford, F.A. Tezcan, *J. Am. Chem. Soc.*, 131 (2009) 9136-9137.

[9] R.J. Radford, M. Lawrenz, P.C. Nguyen, J.A. McCammon, F.A. Tezcan, *Chem Commun*, 47 (2011) 313-315.

[10] R.J. Radford, P.C. Nguyen, T.B. Ditri, J.S. Figueroa, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 4362-4369.

[11] A.E. Martell, R.M. Smith, Critical Stability Constants, Plenum Press, New York, 1974-1989.

[12] M.R. Ghadiri, C. Choi, *J. Am. Chem. Soc.*, 112 (1990) 1630-1632.

[13] F.Q. Ruan, Y.Q. Chen, P.B. Hopkins, *J. Am. Chem. Soc.*, 112 (1990) 9403-9404.

[14] H.E. Blackwell, R.H. Grubbs, *Angew. Chem. Int. Ed.*, 37 (1998) 3281-3284.

[15] C.E. Schafmeister, J. Po, G.L. Verdine, *J. Am. Chem. Soc.*, 122 (2000) 5891-5892.

[16] F.Z. Zhang, O. Sadovski, S.J. Xin, G.A. Woolley, *J. Am. Chem. Soc.*, 129 (2007) 14154-14155.

[17] G.L. Verdine, L.D. Walensky, *Clin. Cancer Res.*, 13 (2007) 7264-7270.

[18] T. Kortemme, D. Baker, *Curr. Opin. Chem. Biol.*, 8 (2004) 91-97.

[19] B.A. Shoemaker, A.R. Panchenko, *PLos Comput. Biol.*, 3 (2007) 595-601.

[20] J.J. Kohler, S.J. Metallo, T.L. Schneider, A. Schepartz, *Proc. Natl. Acad. Sci. USA*, 96 (1999) 11735-11739.

[21] B.R. Crane, A.J. Di Bilio, J.R. Winkler, H.B. Gray, *J. Am. Chem. Soc.*, 123 (2001) 11623-11631.

[22] F.N. Castellano, J.D. Dattelbaum, J.R. Lakowicz, *Anal. Biochem.*, 255 (1998) 165-170.

[23] S.K. Singh, M. Trivedi, M. Chandra, A.N. Sahay, D.S. Pandey, *Inorg. Chem.*, 43 (2004) 8600-8608.

[24] T. Alber, *Ann. Rev. Biochem.*, 58 (1989) 765-798.

[25] J. Faraone-Mennella, F.A. Tezcan, H.B. Gray, J.R. Winkler, *Biochemistry*, 45 (2006) 10504-10511.

[26] Z.G. Zhou, G.H. Sarova, S. Zhang, Z.P. Ou, F.T. Tat, K.M. Kadish, L. Echegoyen, D.M. Guldi, D.I. Schuster, S.R. Wilson, *Chem. Eur. J.*, 12 (2006) 4241-4248.

[27] P. Schuck, *Biophys. Chem.*, 108 (2004) 187-200.

[28] G. Te Velde, F.M. Bickelhaupt, E.J. Baerends, C. Fonseca Guerra, S.J.A. van Gisbergen, J.G. Snijders, T. Ziegler, *J. Comput. Chem.*, 22 (2001) 931-967.

[29] C.F. Guerra, J.G. Snijders, G. te Velde, E.J. Baerends, *J. Theor. Chem. Acc.*, 99 (1998) 391-403.

[30] ADF2007.01, SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands, www.scm.com.

[31] ADF-GUI 2007.01, SCM, Amsterdam, The Netherlands, www.scm.com Access date, February, 2008.

[32] E. van Lenthe, E.J. Baerends, J.G. Snijders, *J. Chem. Phys.*, 99 (1993) 4597-4610.

[33] S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.*, 58 (1980) 1200-1211.

[34] A.D. Becke, *Phys. Rev. A*, 38 (1988) 3098-3100.

[35] J.P. Perdew, W. Yue, *Phys. Rev. B*, 33 (1986) 8800-8802.

[36] J.P. Perdew, *Phys. Rev. B*, 34 (1986) 7406-7406.

[37] N.C. Handy, A.J. Cohen, *Mol. Phys.*, 99 (2001) 403-412.

[38] C.T. Lee, W.T. Yang, R.G. Parr, *Phys. Rev. B*, 37 (1988) 785-789.

Appendix to Chapter 5: Modular and Versatile Hybrid Coordination Motifs on α-Helical Protein Surfaces

Table of Contents

Section A5.1. Experimental Figures

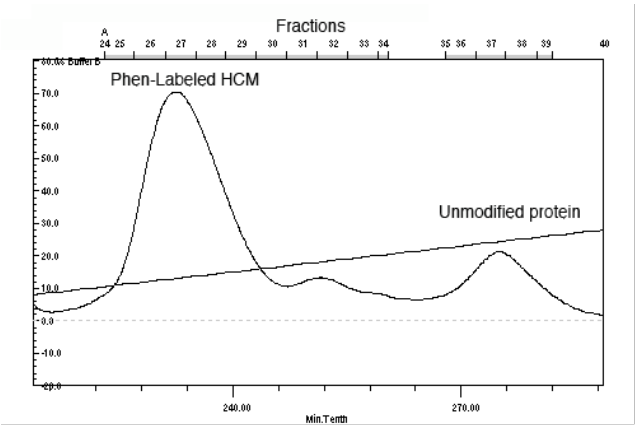


Figure A5.1.1 A representative FPLC trace (anion-exchange with a 5 mL Bio-Rad Uno-Q cartridge) for the purification of a HPhen2. Protein was prepared in 10 mM sodium phosphate buffer (pH 8) and eluted with a linear gradient of buffer B (0-50% 0.5 M NaCl in 10 mM sodium phosphate (pH 8)). Labeled protein generally elutes between 17.5 and 22% of buffer B. The trace shows three peaks with fractions 25-29 representing pure labeled protein, fractions 31-33 representing labeled protein as the major product along with a small unlabeled impurity, and fractions 36-39 corresponding to unlabeled protein.

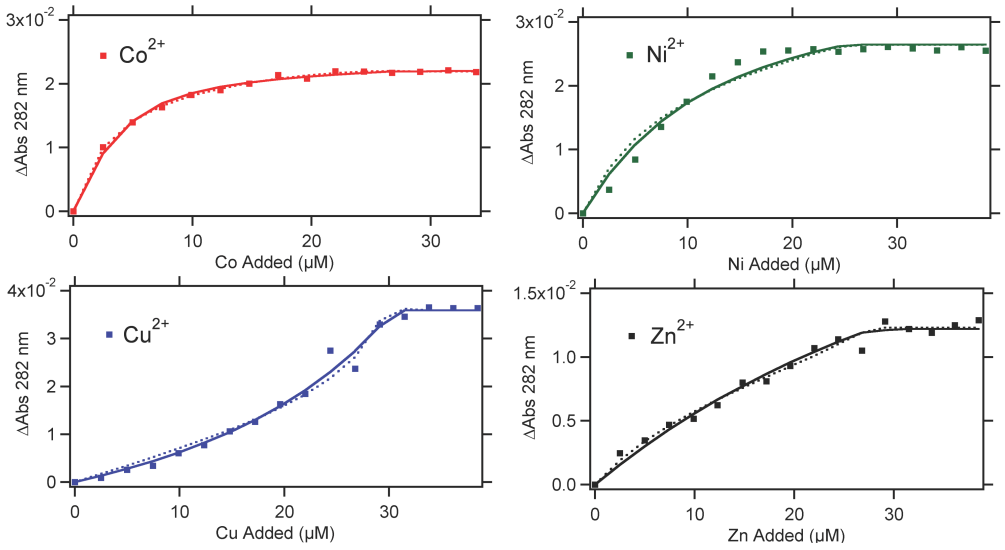


Figure A5.1.2 Titrations of HPhen2 with late first row transition metals as monitored by UV-vis spectroscopy. See Experimental Section for details. The data were fit to two different models. The first model assumes a simple 1:1 binding (solid line) and the second model uses a 1:1 and 1:2 mechanism (dotted line) the latter of which accounts for protein dimerization. In every case, a 1:1 model satisfactorily described the data. Corresponding K_d's can be found in Tables 5.1 and A5.2.1.

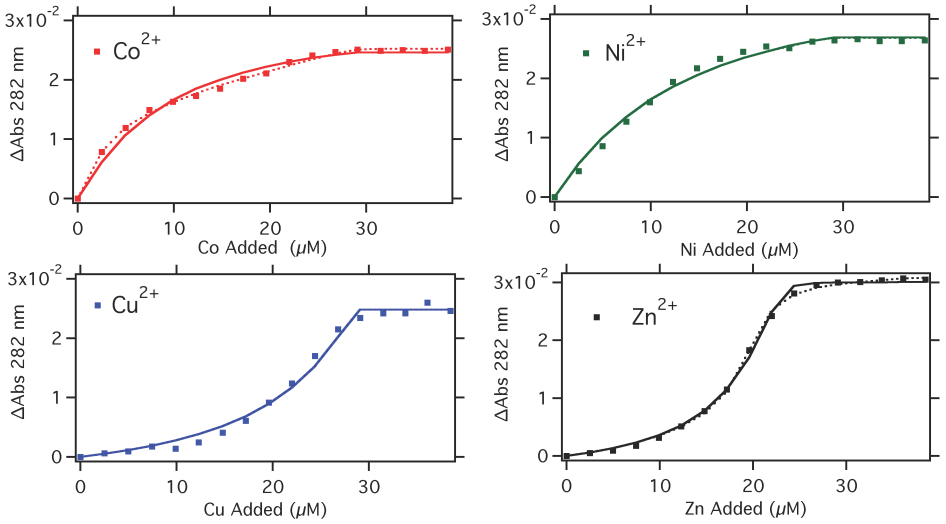


Figure A5.1.3. Titrations of HPhen3 with late first row transition metals as monitored by UV-vis spectroscopy. See Experimental Section for details. The data were fit to two different models. The first model assumes a simple 1:1 binding (solid line) and the second model uses a 1:1 and 1:2 mechanism (dotted line) the latter of which accounts for protein dimerization. In every case, a 1:1 model satisfactorily described the data. Corresponding K_d's can be found in Tables 5.1 and A5.2.1.

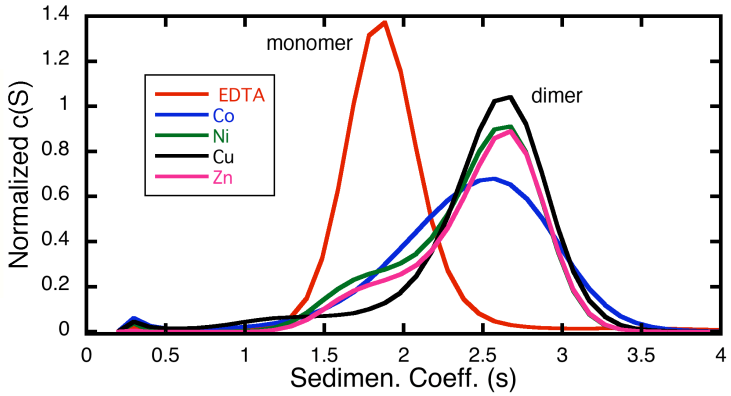


Figure A5.1.4. Sedimentation velocity profiles for HTerpy1 performed under similar conditions as metal binding titrations (5 μM HTerpy1 and 2.5 μM of M^{2+}). With every metal tested (Co^{2+}(blue), Ni^{2+} (green), Cu^{2+} (black), Zn^{2+}(pink) Metal free/EDTA(Red)), a dimeric species (2.6 S) is the major species present.

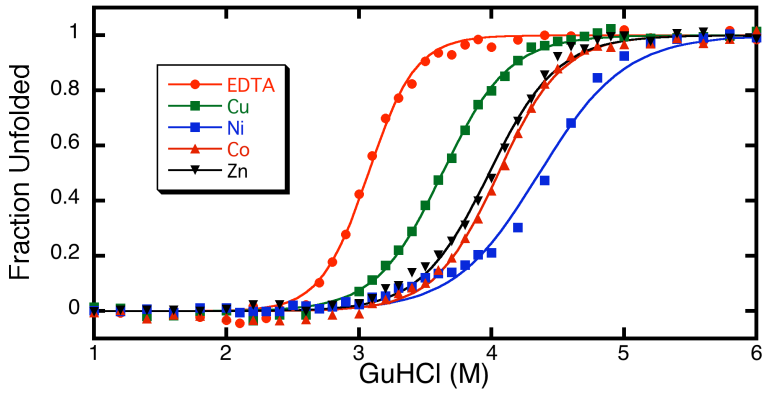


Figure A5.1.5. GuHCl induced unfolding titrations for HPhen1 in the presence and absence (EDTA) of late first-row transition metals as monitored by CD spectroscopy. Titration data were fit assuming a two-state unfolding model. Fitting parameters are given in Table A5.2.2.

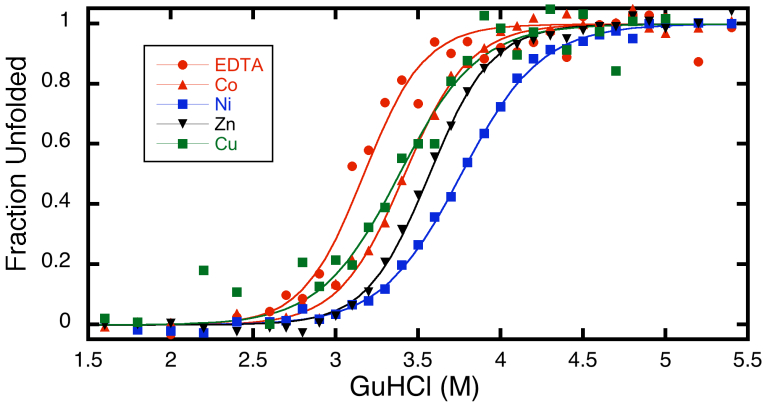


Figure A5.1.6. GuHCl induced unfolding titrations for HPhen2 in the presence and absence (EDTA) of late first-row transition metals as monitored by CD spectroscopy. Titration data were fit assuming a two-state unfolding model. Fitting parameters are given in Table A5.2.3.

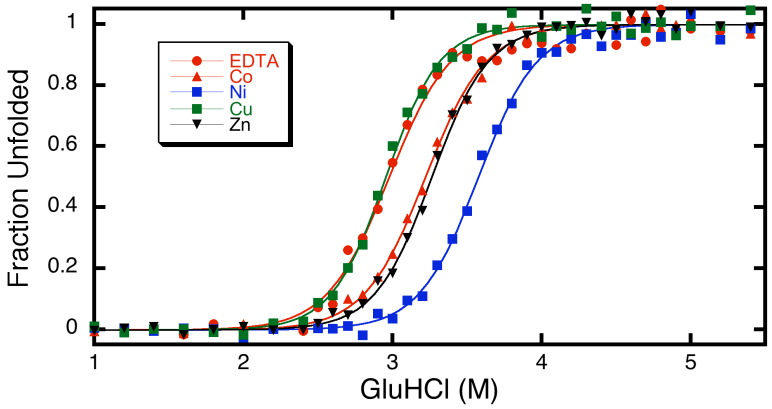


Figure A5.1.7. GuHCl induced unfolding titrations for HPhen3 in the presence and absence (EDTA) of late first-row transition metals as monitored by CD spectroscopy. Titration data were fit assuming a two-state unfolding model. Fitting parameters are given in Table A5.2.4.

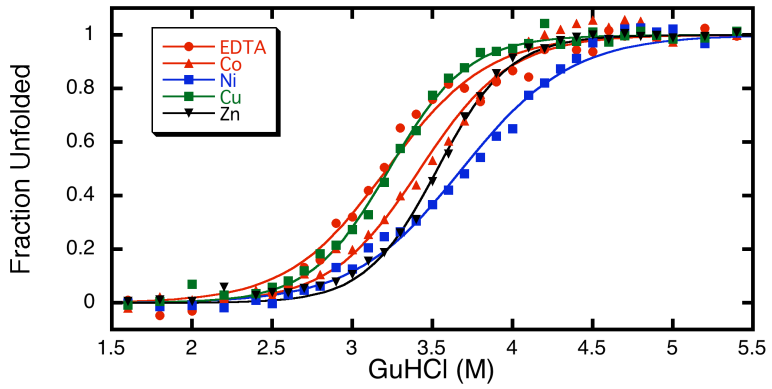


Figure A5.1.8. GuHCl induced unfolding titrations of HTerpy1 in the presence and absence (EDTA) of late first-row transition metals, as monitored by circular dichroism (CD) spectroscopy. Titration data were fit assuming a two-state unfolding model. Fitting parameters are given in Table A5.2.5.

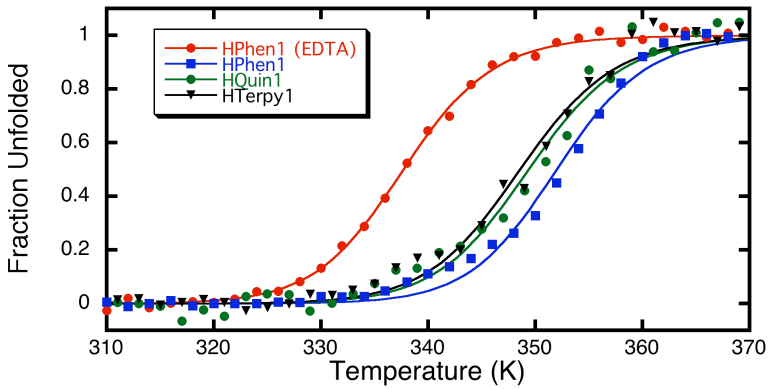


Figure A5.1.9. Thermally induced unfolding of 5 μM HPhen1, HQuin1 and HTerpy1 in the presence and absence (EDTA) of 1 mM Ni^{2+} as monitored by CD spectroscopy. Titrations were performed in 100 mM Tris buffer (pH 7.5) with 1.5 M GuHCl and fit to a two-state unfolding model. Fitting parameters are given in Table A5.2.6.

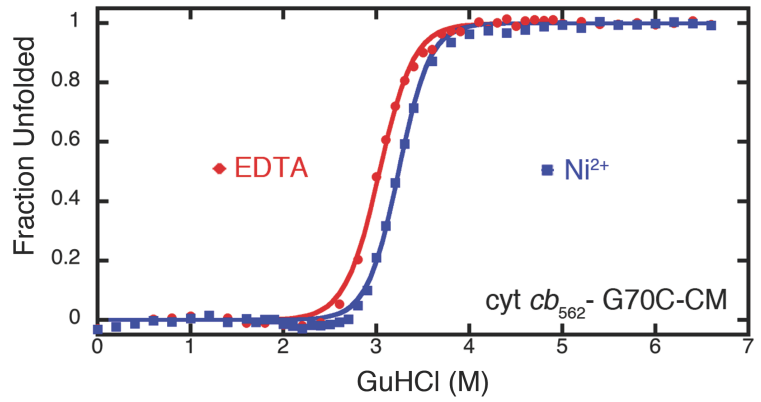


Figure A5.1.10. GuHCl induced unfolding titration of G70C-CM cyt cb_{562} in the presence and absence (EDTA) of Ni^{2+} as monitored by CD spectroscopy.

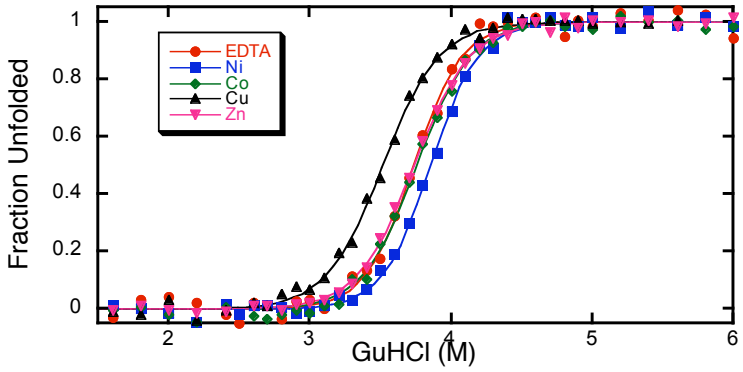


Figure A5.1.11. GuHCl induced unfolding titration of APhen1 in the presence and absence (EDTA) of M^{2+} as monitored by CD spectroscopy.

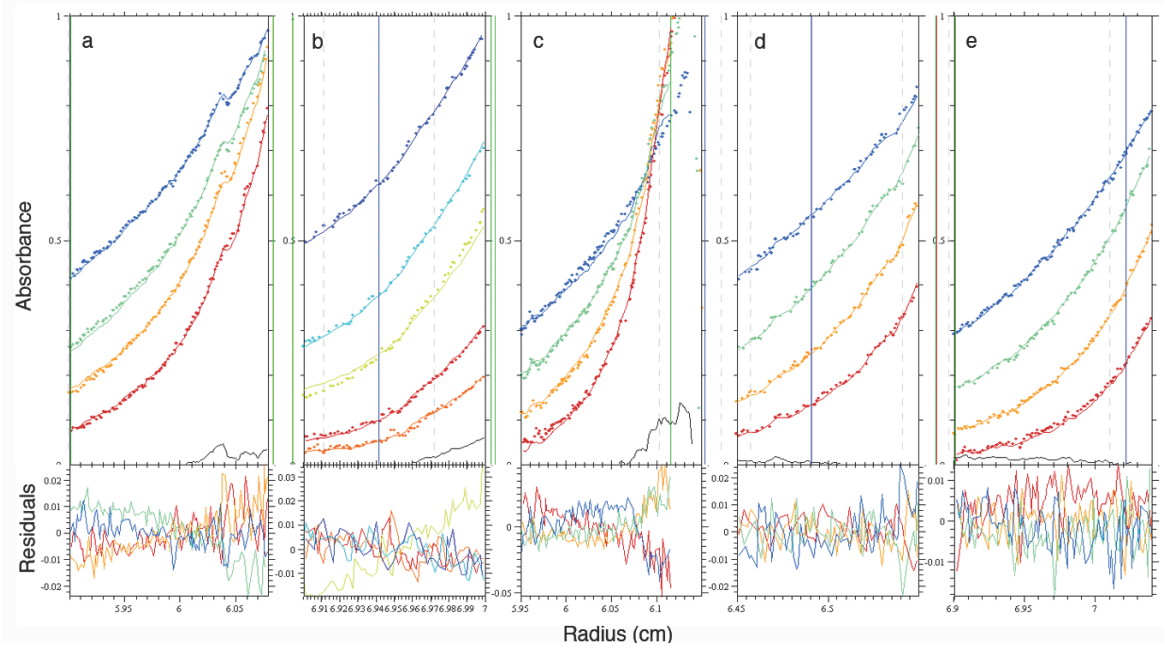


Figure A5.1.12. Sedimentation equilibrium profiles for HPhen1 in the presence of Ni^{2+}. (a) 10 µM (b) 20 µM (c) 40 µM (d) 60 µM and (e) 100 µM. Scans of each sample were taken after the samples were equilibrated at the following speeds for 14 hrs: 20,000 rpm (blue), 25,000 rpm (green), 30,000 rpm (yellow), 35,000 rpm (red) and 41,000 rpm (orange) rpm. All samples were in 20 mM TRIS buffer (pH 7) with a half equivalent of $NiSO_4$. Scans shown in a), b) and c) were globally fit to a monomer-dimer model yielding a minimized fit with a log $K_{d(2mer\text{-}1mer)} = 5.05 \pm 0.02$ M^{-1} or $K_{d(2mer\text{-}1mer)} = 8.9(1)$ µM. Experiments were conducted at 25°C.

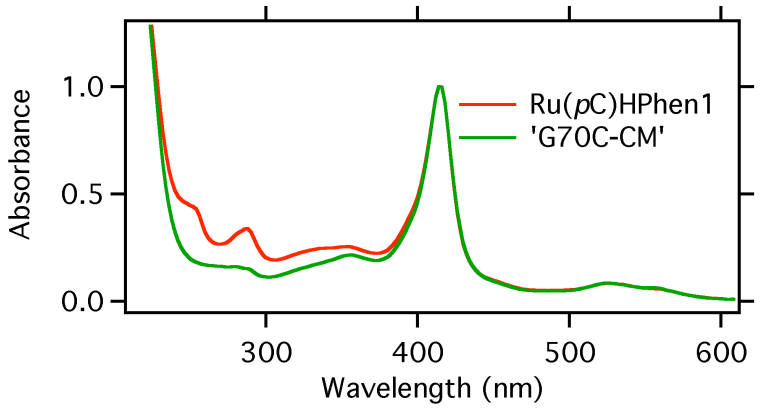


Figure A5.1.13. UV-visible spectra of Ru(*p*-cymene)-HPhen1 complex and G70C-CM cyt cb_{562}.

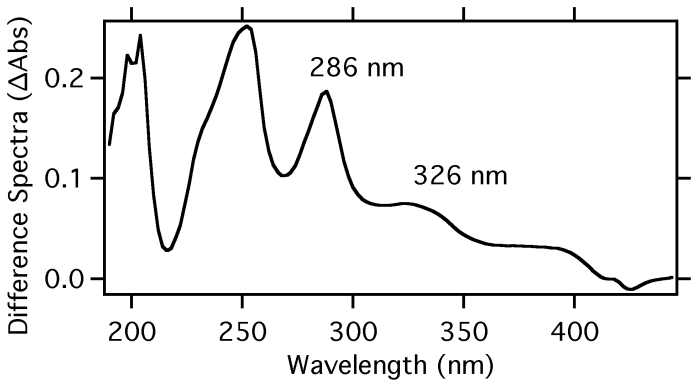


Figure A5.1.14. Difference spectrum (from Figure A5.1.13) highlighting the changes in absorbance upon the formation of the Ru(*p*-cymene)-HPhen1 complex.

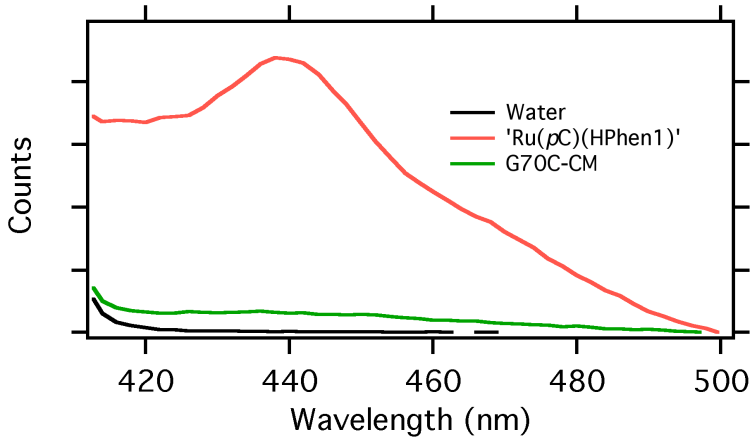


Figure A5.1.15. Fluorescence spectrum of 5 μM Ru(*p*-cymene)-HPhen1 as compared with that of 5 μM G70C-CM cyt cb_{562}. Spectra were taken in water and prepared anaerobically. (λ_{ex} = 326 nm; λ_{em} = 442 nm)

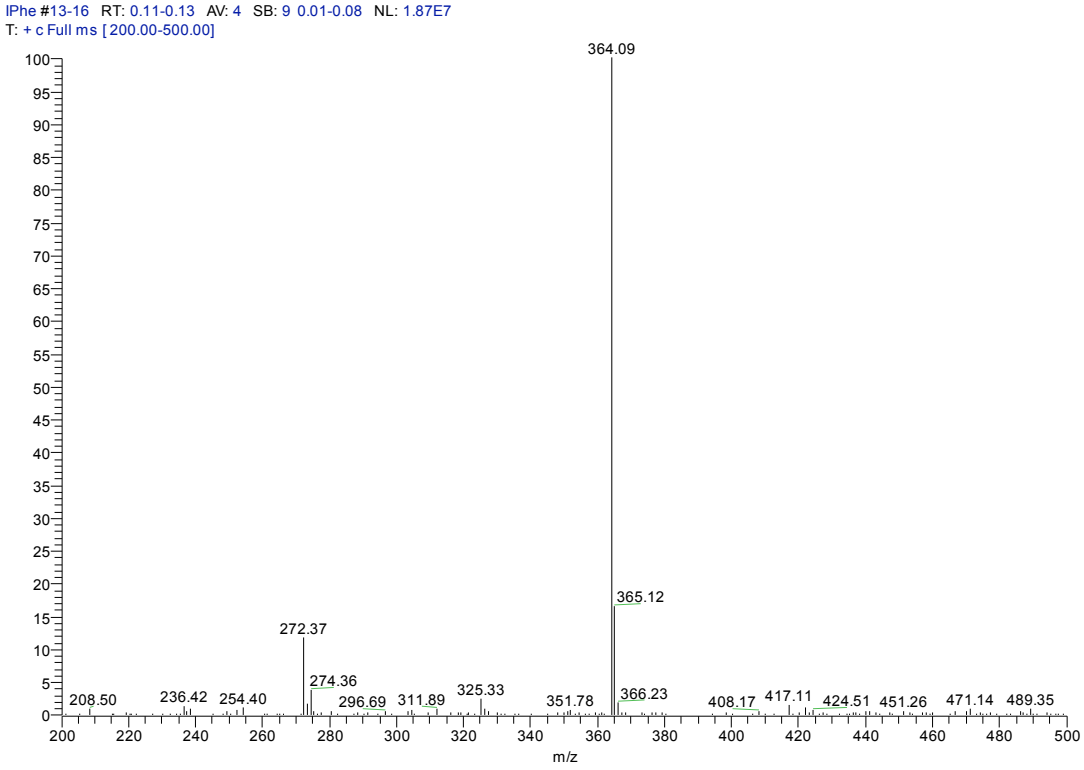


Figure A5.1.16. ESI-MS (positive mode) of 5-iodoacetamido-1,10-phenanthroline (IA-Phen). Measured MW = 364.09 m/z (exp.: 363.99) ($M + H^+$)

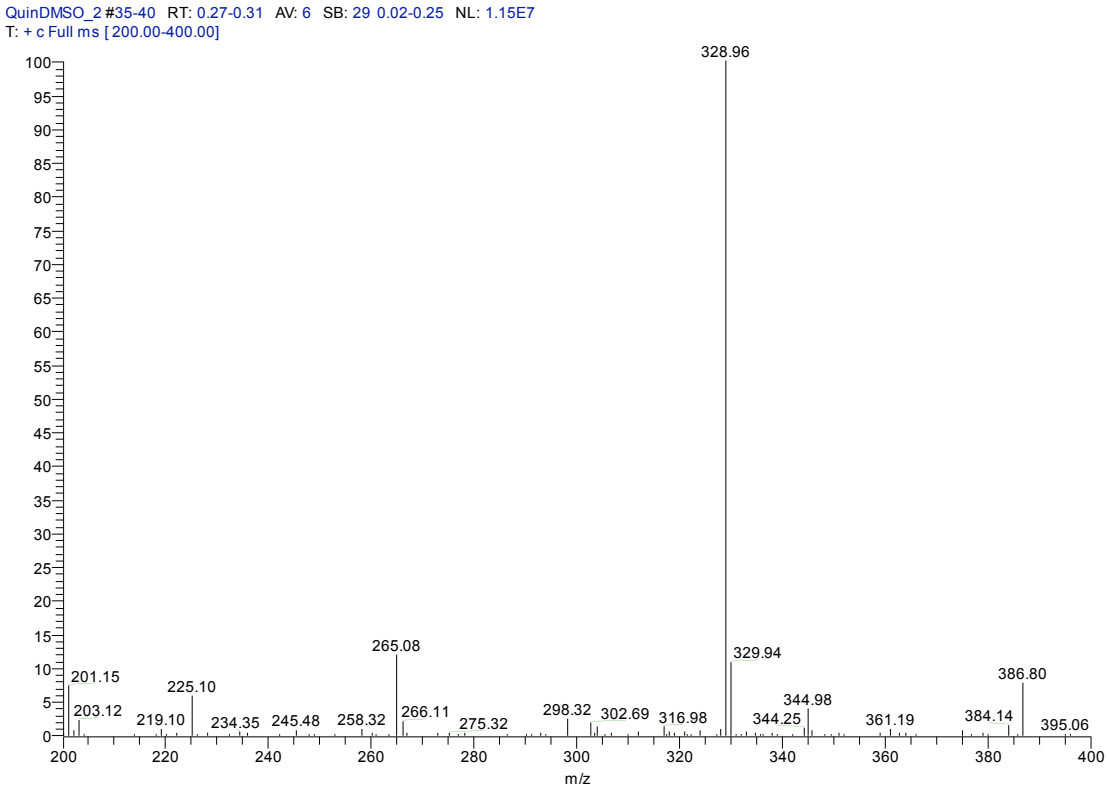


Figure A5.1.17. ESI-MS (positive mode) of 5-iodoacetamido-8-hydroxyquinoline (IA-Quin). Measured MW = 328.96 m/z (exp.: 363.99) ($M + H^+$)

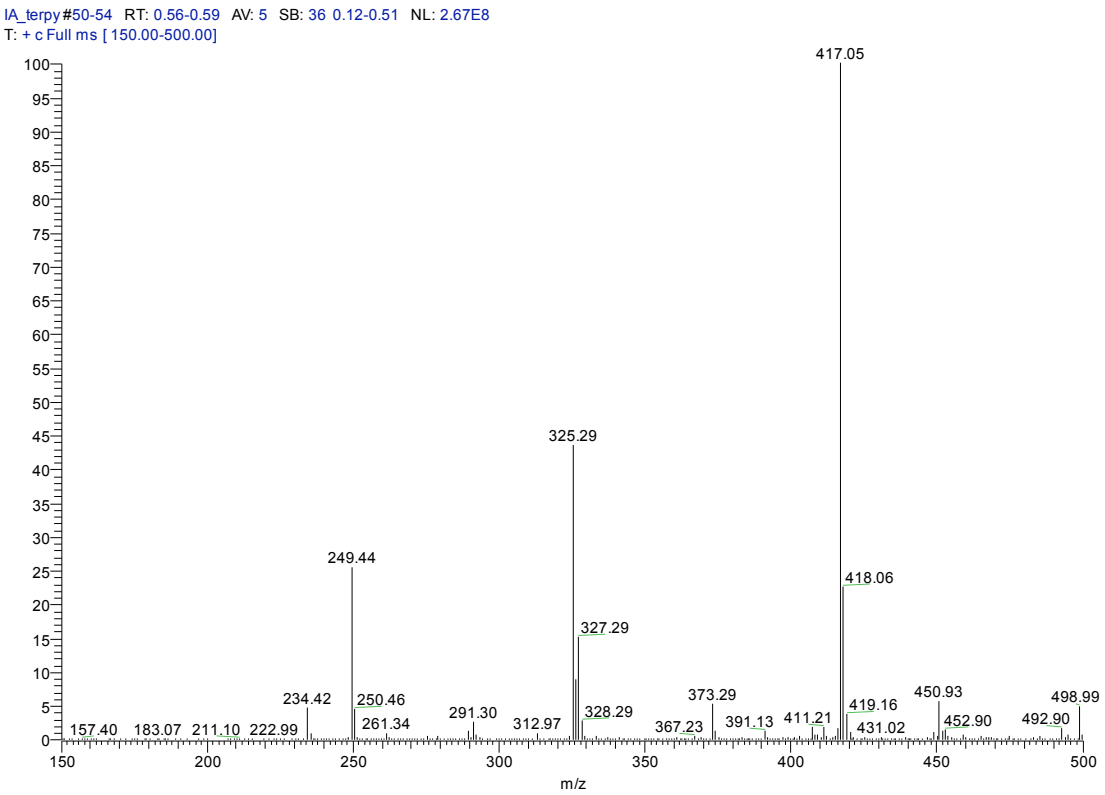


Figure A5.1.18. ESI-MS (positive mode) of 4-iodoacetamido-2,2':6',2"-terpyridine (ITerpy). Measured MW = 417.05 m/z (exp.: 417.22) ($M + H^+$). Observed peaks at 325.29 and 249.44 m/z correspond to 4-chloroacetamido-2,2':6',2"-terpyridine (ClA-Terpy) and 4-amino-2,2':6',2"-terpyridine (NH_2-Terpy) respectively.

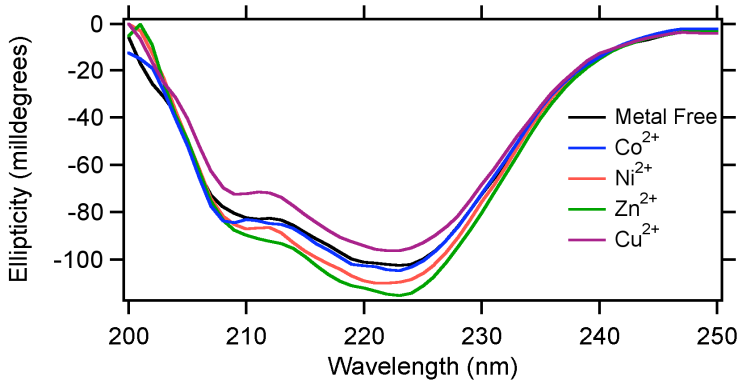


Figure A5.1.19. Wavelength scan of 6 µM HPhen1 in 20 mM sodium borate buffer (pH 7) at 25 °C and 200 µM of either EDTA (no metal) or M^{2+} as monitored by CD spectroscopy. Data shown are averages of 3 scans and was smoothed with a binomial function.

Section A5.2. Experimental Tables.

Table A5.2.1. Determined dissociation constants for metal-binding titrations for the HPhen series with M^{2+}. Data were fit using a 1:1/1:2 model, which accounts for both metal binding and protein dimerization. In all cases EGTA was used as competing ligand. EGTA:metal dissociation constants were determined using Maxchelator (http://maxchelator.stanford.edu) (25° C and a ionic strength of 0.05 M) and were held fixed during regression analysis.

	Dissociation Constants (K_d)		
	HPhen1 (M)	**HPhen2 (M)**	**HPhen3 (M)**
Co^{2+}	4 (4) × 10^{-10}	5 (2) × 10^{-9}	2 (6) × 10^{-10}
Ni^{2+}	2 (80) × 10^{-9}	6 (24) × 10^{-11}	2 (15) × 10^{-10}
Cu^{2+}	6 (4) × 10^{-13}	1 (15) × 10^{-11}	2 (11) × 10^{-11}
Zn^{2+}	3 (3) × 10^{-8}	5 (3) × 10^{-8}	2 (2) × 10^{-7}

Table A5.2.2. Fitting parameters for chemical unfolding titrations of cyt cb_{562} variants: HPhen1, APhen1, G70C-CM cyt cb_{562}.

Cyt cb_{562} variant	**Condition**	**Slope, m_1 (kcal/mol•M)**	**$[GuHCl]_m$, m_2 (M)**
HPhen1	EDTA	3.2(1)	3.07 (1)
HPhen1	Co	2.42 (5)	4.05 (1)
HPhen1	Ni	1.82 (7)	4.36 (1)
HPhen1	Cu	2.33 (7)	3.63 (1)
HPhen1	Zn	2.26 (6)	3.98 (1)
G70C-CM	EDTA	2.9(1)	3.06(1)
G70C-CM	Ni	3.6(1)	3.24(1)
APhen1	EDTA	3.4 (1)	3.73 (1)
APhen1	Co	3.1 (1)	3.75 (1)
APhen1	Ni	3.4 (1)	3.85 (1)
APhen1	Cu	2.9 (1)	3.51 (1)
APhen1	Zn	2.9 (1)	3.73 (1)
HPhen1	EDTA (pH 5.5)	3.4 (1)	4.51 (1)
HPhen1	Ni (pH 5.5)	2.8 (1)	4.83 (1)

Table A5.2.3. Fitting parameters for chemical unfolding titrations for HPhen2 with M^{2+}.

Cyt cb_{562} variant	Condition	Slope, m_1 (kcal/mol•M)	$[GuHCl]_m$, m_2 (M)
HPhen2	EDTA	3.1 (3)	3.16 (2)
HPhen2	Co	2.96 (1)	3.41 (1)
HPhen2	Ni	2.42 (1)	3.76 (1)
HPhen2	Cu	2.4 (2)	3.38 (2)
HPhen2	Zn	3.08 (1)	3.56 (1)

Table A5.2.4. Fitting parameters for chemical unfolding titrations for HPhen3 with M^{2+}.

Cyt cb_{562} variant	Condition	Slope, m_1 (kcal/mol•M)	$[GuHCl]_m$, m_2 (M)
HPhen3	EDTA	2.83 (2)	2.97 (1)
HPhen3	Co	2.86 (2)	3.21 (1)
HPhen3	Ni	2.96 (1)	3.57 (1)
HPhen3	Cu	3.22 (1)	2.95 (1)
HPhen3	Zn	3.07 (1)	3.26 (1)

Table A5.2.5. Fitting parameters for chemical unfolding titrations for HTerpy1 and ATerpy1.

Cyt cb_{562} variant	Condition	Slope, m_1 (kcal/mol•M)	$[GuHCl]_m$, m_2 (M)
HTerpy1	EDTA	1.89 (1)	3.21 (1)
HTerpy1	Co	2.41 (1)	3.42 (1)
HTerpy1	Ni	1.73 (1)	3.68 (1)
HTerpy1	Cu	2.40 (1)	3.22 (1)
HTerpy1	Zn	2.66 (1)	3.53 (1)
HTerpy1	EDTA (pH 5)	2.67 (1)	4.46 (1)
HTerpy1	Ni (pH 5)	2.54 (1)	4.58 (1)
ATerpy1	EDTA	2.72 (1)	3.56 (1)
ATerpy1	Ni	3.2 4(1)	3.59 (1)

Table A5.2.6. Fitting parameters for thermal unfolding titrations of HCM variants: HPhen1, HQuin1 and HTerpy1

HCM variant	Condition	T_m (K)	slope (kJ/mol)
HPhen1	EDTA	337 (1)	54 (1)
HPhen1	Ni	352 (1)	59 (2)
HQuin1	Ni	349 (1)	55 (4)
HTerpy1	Ni	348 (1)	54 (3)

Table A5.2.7. Total bonding energies for geometry optimized models. Lowest energy isomer and corresponding total bonding energy are in bold. $\Delta E = |$ *trans* isomer – *cis* isomer $|$.

Model	Total Bonding Energy (kcal mol^{-1})	ΔE (kcal mol^{-1})
***cis*-HPhen$_2$:Ni BP86**	**-10804.01**	5.35
trans-HPhen$_2$:Ni	-10798.66	
***cis*- Phen$_2$:Ni OLYP**	**-10517.02**	5.01
trans-HPhen$_2$:Ni OLYP	-10512.01	

Section A5.3. Input Files For Density Functional Calculations

A5.3.1. Input file for inner-coordination sphere of the *cis*-HPhen1$_2$:Ni^{2+} complex (BP86).

```
#$ -S /bin/bash
#$ -cwd

#$ -o output
#$ -e $JOB_ID.err
#$ -j y
#$ -M rradford@ucsd.edu
#$ -m beas
#$ -N phentransu
#$ -q adf.q
#$ -pe mpi 8

export cur_dir="`pwd`"
echo Running as user `whoami` on `hostname` at `date` in dir `pwd`
export temp_dir="/state/partition1/`whoami`.$JOB_ID"
mkdir $temp_dir
cd $temp_dir
echo With temp dir $temp_dir


# ADF enviroment variables, change as you need
```

```
export ADFHOME=/share/apps/adf2007.01/
export ADFBIN=/share/apps/adf2007.01/bin
export ADFRESOURCES=/share/apps/adf2007.01/atomicdata
export SCMLICENSE=/share/apps/adf2007.01/license
export SCM_TMPDIR=$temp_dir
export SCM_USETMPDIR=yes
export NSCM=8
export SCM_IOBUFFERSIZE=512


 #Put main code here ************************************

$ADFBIN/adf -n8 \
 <<< "
TITLE Nicis geo opt

MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE 2 3
UNRESTRICTED

SCF
 ITERATIONS 200
DIIS
END

XC
 LDA VWN
 GGA Becke Perdew
END

SYMMETRY NOSYM
ATOMS
 C                 0.00000000    0.00000000    0.00000000
 C                 1.49256122    0.00000000    0.00000000
 C                 2.42693295    1.01015667    0.00000000
 N                 2.24529015   -1.16582375    0.02218498
 C                 3.56299355   -0.83228914    0.02976203
 N                 3.70228499    0.48144472    0.01702533
 C                10.99154846    3.29300696   -0.91732029
 C                 9.76649252    2.57572479   -0.45595145
 C                 8.42465811    2.76330770   -0.70507607
 N                 9.80690694    1.49498046    0.41145927
 C                 8.53483249    1.07600525    0.64896160
 N                 7.67565567    1.83081541   -0.01410612
 Ni                5.65295941    1.28472803   -0.00859172
 H                10.74625400    3.95654088   -1.75289337
 H                11.43520195    3.90761729   -0.11919016
 H                11.76432279    2.59203105   -1.26701389
 H                10.64217050    1.07372238    0.79822072
 H                 8.25741807    0.24781706    1.28724959
 H                 7.96193740    3.50267040   -1.34772477
 H                -0.40717928   -0.69538441   -0.74906712
 H                -0.41369760   -0.28481885    0.98041905
```

```
 H                -0.37511359    1.00135511   -0.24127110
 H                 2.25463114    2.07933432   -0.00449593
 H                 4.38130839   -1.54009742    0.02399485
 H                 1.87848777   -2.10952611    0.01533945
 C                 4.69855903    3.62551319   -3.91832339
 C                 6.95525440   -2.59797330   -1.71575989
 C                 4.82493173    3.30322448   -2.55556408
 C                 6.59448059   -1.57173263   -0.83928322
 C                 5.03901876    2.65560725   -4.87776834
 C                 6.86023116   -2.36135354   -3.10540865
 C                 5.83549531    0.42080423   -5.40286890
 C                 6.29574577   -0.85678871   -5.02648906
 C                 5.49251641    1.39752677   -4.46428509
 C                 6.41742319   -1.13716794   -3.63288255
 C                 5.60236718    1.12009014   -3.07825958
 C                 6.06831524   -0.14251271   -2.65864971
 H                 6.37948554   -1.10621310   -7.06050337
 C                 4.55281559    5.02059536    2.00871462
 C                 6.83501076   -1.40804771    3.46938692
 C                 4.92317427    4.10133629    1.01120462
 C                 6.70708636   -0.93097127    2.16266644
 C                 4.64776446    4.63038406    3.35614212
 C                 6.48435915   -0.55161228    4.53602595
 C                 5.20505617    2.94175924    5.01201474
 C                 5.65999496    1.65871079    5.37622809
 C                 5.10422279    3.34686458    3.67888899
 C                 6.01551740    0.75629719    4.33136710
 C                 5.46545152    2.45734126    2.63553763
 C                 5.91620355    1.16093150    2.95677486
 H                 5.77972003    2.51833853    7.10855185
 H                 4.34972070    4.59109841   -4.21974834
 H                 7.29571029   -3.54148639   -1.34323765
 H                 4.57110116    4.02864600   -1.81109536
 H                 6.65641476   -1.72140320    0.21838538
 H                 4.95171578    2.87852403   -5.92064240
 H                 7.13584618   -3.14383389   -3.78117552
 H                 4.20540236    5.99829608    1.74736174
 H                 7.19049349   -2.39954196    3.65772485
 H                 4.85717991    4.37992148   -0.01978279
 H                 6.96446737   -1.55718134    1.33410385
 H                 4.37187037    5.31264037    4.13287167
 H                 6.57952556   -0.91303750    5.53863014
 N                 6.53384187   -1.54391308   -6.17473091
 N                 6.17197396   -0.40312850   -1.32792989
 N                 5.26394220    2.08581038   -2.18284070
 N                 5.36006422    2.87344891    1.34557680
 N                 6.26208552    0.31021958    1.95269374
 N                 5.63233111    1.60284057    6.73419909
 H                 6.35379620    0.99043632    7.05738502
 H                 6.86337833   -2.48716750   -6.13378188
 H                 5.74549280    0.65106919   -6.44391529
 H                 4.92724985    3.62946358    5.78323869
END

GEOMETRY
```

```
 GO
 ITERATIONS 100
 FREQUENCIES
END

BASIS
 type TZ2P
 core none
END

END INPUT
"
# end main code ******************************************

cp * $cur_dir/

# Optional, have to manually clean up otherwise
rm -rf $temp_dir
```

A5.3.2. Input file for inner-coordination sphere of the *trans*-HPhen1$_2$:Ni^{2+} complex (BP86).

```
#$ -S /bin/bash
#$ -cwd

#$ -o output
#$ -e $JOB_ID.err
#$ -j y
#$ -M rradford@ucsd.edu
#$ -m beas
#$ -N phentransu
#$ -q adf.q
#$ -pe mpi 8

export cur_dir="`pwd`"
echo Running as user `whoami` on `hostname` at `date` in dir `pwd`
export temp_dir="/state/partition1/`whoami`.$JOB_ID"
mkdir $temp_dir
cd $temp_dir
echo With temp dir $temp_dir


# ADF enviroment variables, change as you need

export ADFHOME=/share/apps/adf2007.01/
export ADFBIN=/share/apps/adf2007.01/bin
export ADFRESOURCES=/share/apps/adf2007.01/atomicdata
export SCMLICENSE=/share/apps/adf2007.01/license
export SCM_TMPDIR=$temp_dir
export SCM_USETMPDIR=yes
export NSCM=8
export SCM_IOBUFFERSIZE=512
```

```
 #Put main code here ***************************************

$ADFBIN/adf -n8 \
 <<< "
TITLE Nicis geo opt

MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE 2 2
UNRESTRICTED

SCF
 ITERATIONS 200
DIIS
END

XC
 LDA VWN
 GGA Becke Perdew
END

SYMMETRY NOSYM
ATOMS
 C                  0.00000000    0.00000000    0.00000000
 C                  1.49256122    0.00000000    0.00000000
 C                  2.42693295    1.01015667    0.00000000
 N                  2.24529015   -1.16582375    0.02218498
 C                  3.56299355   -0.83228914    0.02976203
 N                  3.70228499    0.48144472    0.01702533
 C                 10.99154846    3.29300696   -0.91732029
 C                  9.76649252    2.57572479   -0.45595145
 C                  8.42465811    2.76330770   -0.70507607
 N                  9.80690694    1.49498046    0.41145927
 C                  8.53483249    1.07600525    0.64896160
 N                  7.67565567    1.83081541   -0.01410612
 Ni                 5.65295941    1.28472803   -0.00859172
 H                 10.74625400    3.95654088   -1.75289337
 H                 11.43520195    3.90761729   -0.11919016
 H                 11.76432279    2.59203105   -1.26701389
 H                 10.64217050    1.07372238    0.79822072
 H                  8.25741807    0.24781706    1.28724959
 H                  7.96193740    3.50267040   -1.34772477
 H                 -0.40717928   -0.69538441   -0.74906712
 H                 -0.41369760   -0.28481885    0.98041905
 H                 -0.37511359    1.00135511   -0.24127110
 H                  2.25463114    2.07933432   -0.00449593
 H                  4.38130839   -1.54009742    0.02399485
 H                  1.87848777   -2.10952611    0.01533945
 C                  4.69855903    3.62551319   -3.91832339
 C                  6.95525440   -2.59797330   -1.71575989
 C                  4.82493173    3.30322448   -2.55556408
 C                  6.59448059   -1.57173263   -0.83928322
```

```
 C                  5.03901876    2.65560725   -4.87776834
 C                  6.86023116   -2.36135354   -3.10540865
 C                  5.83549531    0.42080423   -5.40286890
 C                  6.29574577   -0.85678871   -5.02648906
 C                  5.49251641    1.39752677   -4.46428509
 C                  6.41742319   -1.13716794   -3.63288255
 C                  5.60236718    1.12009014   -3.07825958
 C                  6.06831524   -0.14251271   -2.65864971
 H                  6.37948554   -1.10621310   -7.06050337
 C                  4.55281559    5.02059536    2.00871462
 C                  6.83501076   -1.40804771    3.46938692
 C                  4.92317427    4.10133629    1.01120462
 C                  6.70708636   -0.93097127    2.16266644
 C                  4.64776446    4.63038406    3.35614212
 C                  6.48435915   -0.55161228    4.53602595
 C                  5.20505617    2.94175924    5.01201474
 C                  5.65999496    1.65871079    5.37622809
 C                  5.10422279    3.34686458    3.67888899
 C                  6.01551740    0.75629719    4.33136710
 C                  5.46545152    2.45734126    2.63553763
 C                  5.91620355    1.16093150    2.95677486
 H                  5.77972003    2.51833853    7.10855185
 H                  4.34972070    4.59109841   -4.21974834
 H                  7.29571029   -3.54148639   -1.34323765
 H                  4.57110116    4.02864600   -1.81109536
 H                  6.65641476   -1.72140320    0.21838538
 H                  4.95171578    2.87852403   -5.92064240
 H                  7.13584618   -3.14383389   -3.78117552
 H                  4.20540236    5.99829608    1.74736174
 H                  7.19049349   -2.39954196    3.65772485
 H                  4.85717991    4.37992148   -0.01978279
 H                  6.96446737   -1.55718134    1.33410385
 H                  4.37187037    5.31264037    4.13287167
 H                  6.57952556   -0.91303750    5.53863014
 N                  6.53384187   -1.54391308   -6.17473091
 N                  6.17197396   -0.40312850   -1.32792989
 N                  5.26394220    2.08581038   -2.18284070
 N                  5.36006422    2.87344891    1.34557680
 N                  6.26208552    0.31021958    1.95269374
 N                  5.63233111    1.60284057    6.73419909
 H                  6.35379620    0.99043632    7.05738502
 H                  6.86337833   -2.48716750   -6.13378188
 H                  5.74549280    0.65106919   -6.44391529
 H                  4.92724985    3.62946358    5.78323869
END

GEOMETRY
 GO
 ITERATIONS 100
 FREQUENCIES
END

BASIS
 type TZ2P
 core none
```

```
END

END INPUT
"
# end main code ******************************************

cp * $cur_dir/

# Optional, have to manually clean up otherwise
rm -rf $temp_dir
```

A5.3.3. Input file for the inner-coordination sphere of the *cis*-HPhen1_2:Ni^{2+} complex (OLYP).

```
#$ -S /bin/bash
#$ -cwd

#$ -o output
#$ -e $JOB_ID.err
#$ -j y
#$ -M rradford@ucsd.edu
#$ -m beas
#$ -N phencisu
#$ -q adf.q
#$ -pe mpi 8

export cur_dir="`pwd`"
echo Running as user `whoami` on `hostname` at `date` in dir `pwd`
export temp_dir="/state/partition1/`whoami`.$JOB_ID"
mkdir $temp_dir
cd $temp_dir
echo With temp dir $temp_dir


# ADF enviroment variables, change as you need

export ADFHOME=/share/apps/adf2007.01/
export ADFBIN=/share/apps/adf2007.01/bin
export ADFRESOURCES=/share/apps/adf2007.01/atomicdata
export SCMLICENSE=/share/apps/adf2007.01/license
export SCM_TMPDIR=$temp_dir
export SCM_USETMPDIR=yes
export NSCM=8
export SCM_IOBUFFERSIZE=512


 #Put main code here ***************************************

$ADFBIN/adf -n8 \
 <<< "
TITLE phencisu geo opt
```

```
MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE 2 2
UNRESTRICTED

SCF
 ITERATIONS 200
DIIS
END

XC
 GGA OLYP
END

SYMMETRY NOSYM
ATOMS
   C         3.646596   -2.316668    4.277425
   C         2.833686   -1.391510    3.436131
   C         2.013071   -1.604722    2.352389
   N         2.746520   -0.034432    3.665771
   C         1.903356    0.513799    2.756207
   N         1.422419   -0.417429    1.942404
   C        -2.649755    2.580494    4.907523
   C        -2.308750    1.713671    3.742030
   C        -1.224684    1.686317    2.896374
   N        -3.137450    0.719682    3.258844
   C        -2.554965    0.144277    2.178770
   N        -1.382143    0.711045    1.920150
   Ni       -0.028201   -0.006653    0.424470
   H         4.670080   -1.941386    4.422577
   H         3.195299   -2.455701    5.271509
   H         3.714716   -3.299681    3.799774
   H         1.836712   -2.553224    1.866928
   H         3.247190    0.476222    4.383491
   H         1.694169    1.573972    2.712869
   H        -3.068163    1.996252    5.739281
   H        -1.752834    3.090709    5.275590
   H        -3.386792    3.350546    4.633735
   H        -0.353676    2.322632    2.952374
   H        -4.029160    0.445007    3.658921
   H        -3.008061   -0.669231    1.631719
   C         2.677497   -2.057177   -2.413996
   C        -1.592429   -4.106877    1.117160
   C         1.141254    4.244071   -0.034398
   C        -2.503340    1.345639   -3.033613
   C         1.654838   -1.806142   -1.479809
   C        -0.871760   -2.905274    1.240705
   C         0.489202    3.073904    0.374242
   C        -1.619907    1.326428   -1.945192
   C         3.464594   -1.015916   -2.859041
   C        -2.509793   -4.248525    0.097677
   C         2.221732    4.144150   -0.893671
   C        -3.245735    0.214422   -3.314965
   C         3.994734    1.422573   -2.768319
```

```
   C        -3.660573   -3.227535   -1.861728
   C        -3.868044   -2.157243   -2.715900
   C         3.735283    2.691459   -2.287016
   C        -2.723287   -3.173113   -0.801541
   C         3.233851    0.293062   -2.366919
   C        -3.109502   -0.933299   -2.507489
   C         2.635513    2.879400   -1.356579
   C        -1.953744   -1.994606   -0.588017
   C         2.186588    0.443244   -1.419586
   C        -2.170293   -0.865679   -1.441780
   C         1.895661    1.748291   -0.909867
   H         2.828423   -3.070699   -2.779410
   H        -1.413910   -4.910208    1.829379
   H         0.786443    5.210882    0.315460
   H        -2.591967    2.242985   -3.642007
   H         1.015973   -2.611876   -1.125248
   H        -0.142818   -2.777550    2.032455
   H        -0.370588    3.117726    1.037222
   H        -1.030623    2.207235   -1.699882
   H         4.256382   -1.187146   -3.587089
   H        -3.075666   -5.171376   -0.020914
   H         2.719557    5.051142   -1.230210
   H        -3.922647    0.228220   -4.167317
   N        -1.032116   -1.876233    0.406975
   N        -1.448888    0.256478   -1.172045
   N         0.857259    1.856834   -0.033405
   N         1.413365   -0.589434   -0.994559
   N         4.436660    3.791143   -2.719318
   N        -4.750490   -2.237688   -3.751645
   H        -5.038287   -1.424927   -4.275334
   H        -5.319812   -3.065768   -3.862516
   H         5.227603    3.628933   -3.330166
   H         4.555621    4.580847   -2.097873
   H         4.809494    1.272918   -3.476648
   H        -4.233844   -4.143327   -1.999235
END

GEOMETRY
 GO
 ITERATIONS 100
 FREQUENCIES
END

BASIS
 type TZ2P
 core none
END

END INPUT
"
# end main code ********************************************

cp * $cur_dir/

# Optional, have to manually clean up otherwise
```

```
rm -rf $temp_dir
```

A5.3.4. Input file for inner-coordination sphere of the *trans*-HPhen1_2:Ni^{2+} complex (OLYP).

```
#$ -S /bin/bash
#$ -cwd

#$ -o output
#$ -e $JOB_ID.err
#$ -j y
#$ -M rradford@ucsd.edu
#$ -m beas
#$ -N phentransu
#$ -q adf.q
#$ -pe mpi 8

export cur_dir="`pwd`"
echo Running as user `whoami` on `hostname` at `date` in dir `pwd`
export temp_dir="/state/partition1/`whoami`.$JOB_ID"
mkdir $temp_dir
cd $temp_dir
echo With temp dir $temp_dir


# ADF enviroment variables, change as you need

export ADFHOME=/share/apps/adf2007.01/
export ADFBIN=/share/apps/adf2007.01/bin
export ADFRESOURCES=/share/apps/adf2007.01/atomicdata
export SCMLICENSE=/share/apps/adf2007.01/license
export SCM_TMPDIR=$temp_dir
export SCM_USETMPDIR=yes
export NSCM=8
export SCM_IOBUFFERSIZE=512


 #Put main code here ************************************

$ADFBIN/adf -n8 \
 <<< "
TITLE phentransu geo opt

MAXMEMORYUSAGE 7000

RELATIVISTIC ZORA
CHARGE 2 2
UNRESTRICTED

SCF
 ITERATIONS 200
DIIS
END
```

```
XC
 GGA OLYP
END

SYMMETRY NOSYM
ATOMS
  C        -0.346320   -0.738165   -5.586200
   C        -0.225678   -0.176541   -4.209579
   C        -0.291713   -0.757351   -2.964659
   N         0.021772    1.161285   -3.956126
   C         0.096019    1.359410   -2.620295
   N        -0.085666    0.203620   -1.987482
   C        -0.459579   -1.582978    5.244033
   C        -0.185467   -0.866345    3.965354
   C        -0.326274   -1.221924    2.643056
   N         0.320230    0.420290    3.927617
   C         0.459001    0.810902    2.641828
   N         0.075946   -0.171756    1.831264
   Ni        0.005717   -0.037581   -0.084331
   H        -1.109517   -2.443260    5.068453
   H         0.465910   -1.945290    5.712392
   H        -0.962548   -0.924731    5.962669
   H         0.539070    0.993274    4.735904
   H         0.809418    1.788051    2.346498
   H        -0.679654   -2.158784    2.235337
   H        -1.120195   -0.218666   -6.168294
   H         0.601940   -0.656343   -6.136656
   H        -0.614800   -1.798301   -5.539212
   H        -0.457102   -1.795527   -2.718850
   H         0.254118    2.321523   -2.157662
   H         0.131708    1.885394   -4.658142
   C        -3.667215   -1.808998   -0.024867
   C         1.372982   -4.798541   -1.116702
   C        -1.470899    4.772514    0.879471
   C         3.716954    1.662150    0.146046
   C        -2.316492   -1.457624   -0.122741
   C         0.512624   -3.708703   -0.888785
   C        -0.627902    3.742759    0.440954
   C         2.360206    1.363205    0.043886
   C        -4.586869   -0.822324    0.248284
   C         2.731068   -4.593394   -1.027383
   C        -2.805933    4.476470    1.069416
   C         4.624213    0.626383    0.030811
   C        -5.083844    1.535971    0.743896
   C         4.617092   -3.040369   -0.656645
   C         5.112044   -1.779425   -0.406621
   C        -4.691558    2.838845    0.965628
   C         3.226241   -3.299476   -0.729788
   C        -4.158220    0.515396    0.423302
   C         4.178053   -0.685124   -0.220705
   C        -3.286309    3.173541    0.825214
   C         2.267267   -2.259641   -0.534326
   C        -2.766414    0.807059    0.274022
   C         2.766525   -0.923727   -0.290586
```

```
   C        -2.332139    2.179803    0.440134
   H        -3.963640   -2.847615   -0.148170
   H         0.961689   -5.776431   -1.359660
   H        -1.079214    5.771146    1.059921
   H         4.037321    2.686166    0.321973
   H        -1.540853   -2.198145   -0.291687
   H        -0.567106   -3.860646   -0.959516
   H         0.426374    3.959220    0.251335
   H         1.601208    2.136761    0.126080
   H        -5.645246   -1.058569    0.345307
   H         3.433727   -5.409357   -1.191583
   H        -3.475379    5.258726    1.421736
   H         5.684443    0.841191    0.144729
   N        -1.025593    2.490400    0.238543
   N        -1.882405   -0.199374    0.011860
   N         0.930944   -2.480418   -0.598479
   N         1.897898    0.121220   -0.139467
   N         6.461982   -1.546591   -0.283679
   N        -5.585278    3.798834    1.357908
   H        -5.389017    4.776922    1.198641
   H        -6.570461    3.562865    1.368737
   H         6.828704   -0.635127   -0.508371
   H         7.082638   -2.304932   -0.543188
   H        -6.136003    1.265597    0.828404
   H         5.307210   -3.867899   -0.819032
END

GEOMETRY
 GO
 ITERATIONS 100
 FREQUENCIES
END

BASIS
 type TZ2P
 core none
END

END INPUT
"
# end main code ******************************************

cp * $cur_dir/

# Optional, have to manually clean up otherwise
rm -rf $temp_dir
```

Section A5.4. Results From DFT Calculations

A5.4.1. Optimized cartesian coordinates and molecular structure for the inner-coordination sphere of the *cis*-HPhen1$_2$:Ni^{2+} complex (BP86).

<Feb24-2010> <20:18:19> Geometry Converged
Coordinates in Geometry Cycle 59

Atom	X	Y	Z (Angstrom)
1.C	3.646596	-2.316668	4.277425
2.C	2.833686	-1.391510	3.436131
3.C	2.013071	-1.604722	2.352389
4.N	2.746520	-0.034432	3.665771
5.C	1.903356	0.513799	2.756207
6.N	1.422419	-0.417429	1.942404
7.C	-2.649755	2.580494	4.907523
8.C	-2.308750	1.713671	3.742030
9.C	-1.224684	1.686317	2.896374
10.N	-3.137450	0.719682	3.258844
11.C	-2.554965	0.144277	2.178770
12.N	-1.382143	0.711045	1.920150
13.Ni	-0.028201	-0.006653	0.424470
14.H	4.670080	-1.941386	4.422577
15.H	3.195299	-2.455701	5.271509
16.H	3.714716	-3.299681	3.799774
17.H	1.836712	-2.553224	1.866928
18.H	3.247190	0.476222	4.383491
19.H	1.694169	1.573972	2.712869
20.H	-3.068163	1.996252	5.739281
21.H	-1.752834	3.090709	5.275590
22.H	-3.386792	3.350546	4.633735
23.H	-0.353676	2.322632	2.952374
24.H	-4.029160	0.445007	3.658921
25.H	-3.008061	-0.669231	1.631719
26.C	2.677497	-2.057177	-2.413996
27.C	-1.592429	-4.106877	1.117160
28.C	1.141254	4.244071	-0.034398
29.C	-2.503340	1.345639	-3.033613
30.C	1.654838	-1.806142	-1.479809
31.C	-0.871760	-2.905274	1.240705
32.C	0.489202	3.073904	0.374242
33.C	-1.619907	1.326428	-1.945192
34.C	3.464594	-1.015916	-2.859041
35.C	-2.509793	-4.248525	0.097677
36.C	2.221732	4.144150	-0.893671
37.C	-3.245735	0.214422	-3.314965
38.C	3.994734	1.422573	-2.768319
39.C	-3.660573	-3.227535	-1.861728
40.C	-3.868044	-2.157243	-2.715900
41.C	3.735283	2.691459	-2.287016
42.C	-2.723287	-3.173113	-0.801541
43.C	3.233851	0.293062	-2.366919
44.C	-3.109502	-0.933299	-2.507489
45.C	2.635513	2.879400	-1.356579
46.C	-1.953744	-1.994606	-0.588017

47.C	2.186588	0.443244	-1.419586
48.C	-2.170293	-0.865679	-1.441780
49.C	1.895661	1.748291	-0.909867
50.H	2.828423	-3.070699	-2.779410
51.H	-1.413910	-4.910208	1.829379
52.H	0.786443	5.210882	0.315460
53.H	-2.591967	2.242985	-3.642007
54.H	1.015973	-2.611876	-1.125248
55.H	-0.142818	-2.777550	2.032455
56.H	-0.370588	3.117726	1.037222
57.H	-1.030623	2.207235	-1.699882
58.H	4.256382	-1.187146	-3.587089
59.H	-3.075666	-5.171376	-0.020914
60.H	2.719557	5.051142	-1.230210
61.H	-3.922647	0.228220	-4.167317
62.N	-1.032116	-1.876233	0.406975
63.N	-1.448888	0.256478	-1.172045
64.N	0.857259	1.856834	-0.033405
65.N	1.413365	-0.589434	-0.994559
66.N	4.436660	3.791143	-2.719318
67.N	-4.750490	-2.237688	-3.751645
68.H	-5.038287	-1.424927	-4.275334
69.H	-5.319812	-3.065768	-3.862516
70.H	5.227603	3.628933	-3.330166
71.H	4.555621	4.580847	-2.097873
72.H	4.809494	1.272918	-3.476648
73.H	-4.233844	-4.143327	-1.999235

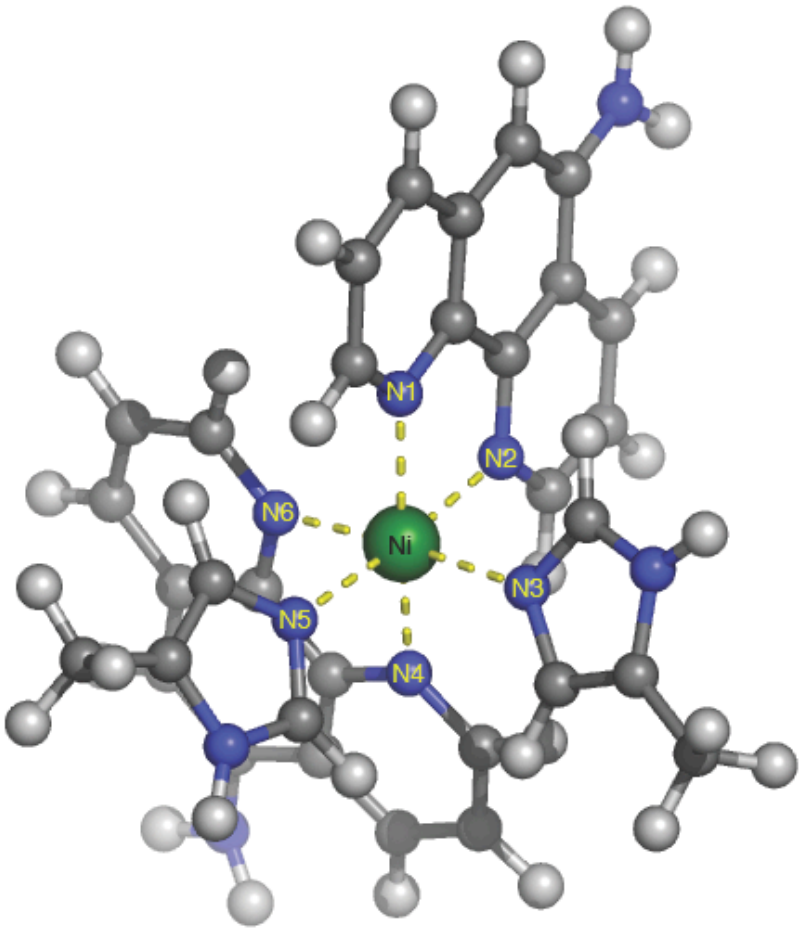


Figure A5.4.2. Computationally optimized geometry for the inner-sphere of the model *cis*-HPhen1$_2$:Ni^{2+} complex (BP86).

Table A5.4.3. Calculated bond distances for the computationally minimized inner-coordination sphere of *cis*-HPhen1$_2$:Ni^{2+} complex.

Bond (Å)	*Cis*-HPhen1$_2$:Ni^{2+} Model
Ni-N1	2.12
Ni-N2	2.15
Ni-N3	2.14
Ni-N4	2.11
Ni-N5	2.14
Ni-N6	2.10

Table A5.4.4. Calculated bond angles the inner-coordination sphere of the computationally minimized *cis*-HPhen1$_2$:Ni^{2+} complex (BP86).

Angle (°)	*Cis*-HPhen1$_2$:Ni^{2+} Model
N1-Ni-N2	77.8
N1-Ni-N3	90.1
N1-Ni-N4	166.7
N1-Ni-N5	99.0
N1-Ni-N6	94.3
N2-Ni-N3	93.4
N2-Ni-N4	90.4
N2-Ni-N5	175.5
N2-Ni-N6	89.2
N3-Ni-N4	96.8
N3-Ni-N5	89.8
N3-Ni-N6	175.3

A5.4.5. Optimized cartesian coordinates and molecular structure for the inner-coordination sphere of the *trans*-HPhen1$_2$:Ni^{2+} complex.

<Feb25-2010> <21:17:12> Geometry Converged
  Coordinates in Geometry Cycle 26

Atom	X	Y	Z (Angstrom)
1.C	0.065415	0.502308	0.026248
2.C	1.538608	0.274729	0.063979
3.C	2.619108	1.124516	-0.009279
4.N	2.100447	-0.983148	0.192705
5.C	3.449462	-0.867699	0.187827
6.N	3.801323	0.406944	0.067866
7.C	10.761681	3.644221	-1.145690
8.C	9.674188	2.765409	-0.627640
9.C	8.313574	2.733377	-0.837653
10.N	9.908357	1.732628	0.262025
11.C	8.732863	1.129878	0.560269
12.N	7.736940	1.714418	-0.095784
13.Ni	5.748174	1.111962	-0.029650
14.H	10.364085	4.344997	-1.886160
15.H	11.221799	4.231618	-0.338292
16.H	11.556113	3.057582	-1.627718
17.H	10.812639	1.471646	0.639743
18.H	8.650923	0.299739	1.246552
19.H	7.726311	3.377660	-1.478770
20.H	-0.410788	-0.088967	-0.768204
21.H	-0.410095	0.229276	0.978936
22.H	-0.149425	1.557599	-0.167191
23.H	2.608414	2.201708	-0.111290
24.H	4.116481	-1.714970	0.260784
25.H	1.586983	-1.854751	0.266162
26.C	4.602335	3.542898	-3.650128
27.C	7.086733	-2.711133	-2.020289
28.C	4.765395	3.097498	-2.328030
29.C	6.761954	-1.736063	-1.074192
30.C	4.990623	2.726897	-4.687452
31.C	6.945095	-2.405659	-3.358750
32.C	5.867591	0.530127	-5.415866
33.C	6.339681	-0.738507	-5.132009
34.C	5.515620	1.443521	-4.395251
35.C	6.491645	-1.128452	-3.741182
36.C	5.664111	1.095919	-3.021578
37.C	6.168549	-0.207783	-2.698417
38.H	6.544714	-1.348098	-7.094767
39.C	4.610770	4.984522	1.904002
40.C	6.579978	-1.412344	3.708425
41.C	4.990698	4.002458	0.972856
42.C	6.442552	-0.987706	2.381731
43.C	4.608350	4.667815	3.244171
44.C	6.309324	-0.510445	4.718710
45.C	5.146553	3.025875	5.015518
46.C	5.605309	1.789312	5.412660
47.C	5.019220	3.377609	3.646906
48.C	5.905216	0.796455	4.394951
49.C	5.350467	2.439305	2.625921

50.C	5.778903	1.131503	3.010477
51.H	5.616521	2.198872	7.418898
52.H	4.168265	4.523643	-3.833763
53.H	7.446194	-3.685401	-1.695523
54.H	4.416262	3.721672	-1.515641
55.H	6.889193	-1.958353	-0.022354
56.H	4.878881	3.041467	-5.723575
57.H	7.191974	-3.164027	-4.099185
58.H	4.335450	5.976367	1.554624
59.H	6.884643	-2.433542	3.923655
60.H	5.053669	4.267213	-0.075362
61.H	6.650957	-1.694879	1.586046
62.H	4.322470	5.403923	3.994217
63.H	6.389897	-0.832573	5.755382
64.N	6.647464	-1.616858	-6.126208
65.N	6.299557	-0.522743	-1.382094
66.N	5.315593	1.925927	-2.001498
67.N	5.309467	2.753708	1.298257
68.N	6.065443	0.236030	2.020178
69.N	5.731526	1.457417	6.738334
70.H	6.397917	0.749342	7.015132
71.H	7.014368	-2.537382	-5.948192
72.H	5.742397	0.838611	-6.452824
73.H	4.905059	3.779652	5.764288

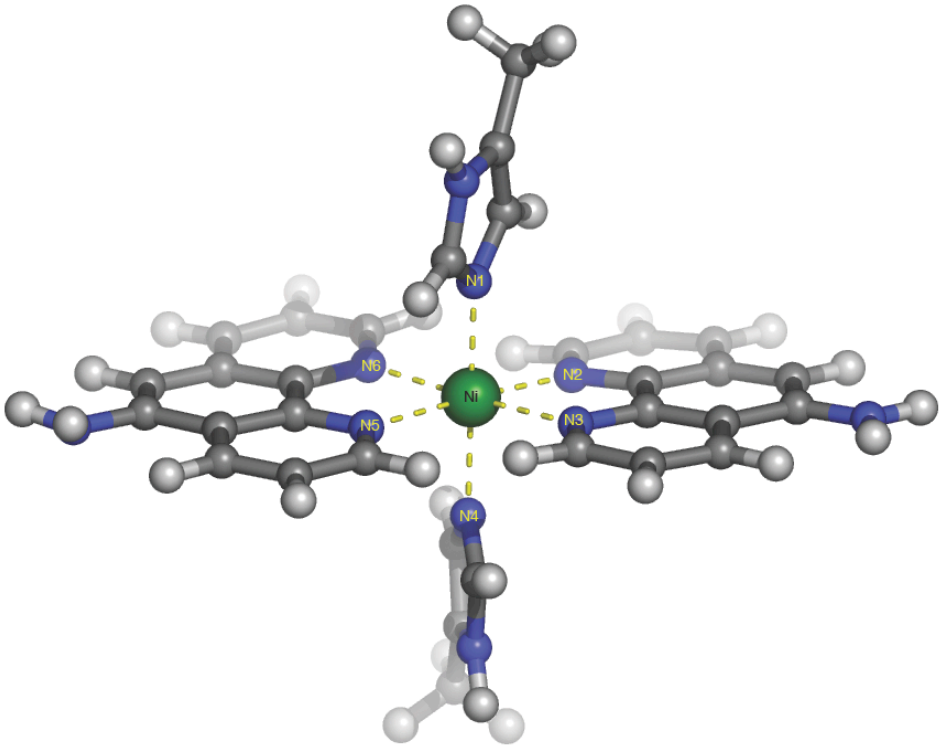


Figure A5.4.6. Optimized geometry of the inner-coordination sphere of the *trans*-HPhen1$_2$:Ni^{2+} complex (BP86).

Table A5.4.7. Calculated bond distances for the computational minimized inner-coordination sphere of the *trans*-HPhen1$_2$:Ni^{2+} complex (BP86).

Bond (Å)	***trans*-HPhen1$_2$:Ni^{2+} Model**
Ni-N1	2.08
Ni-N2	2.18
Ni-N3	2.19
Ni-N4	2.25
Ni-N5	2.16
Ni-N6	2.07

Table A5.4.8 Calculated bond angles for the computationally minimized inner-coordination sphere of the HPhen1$_2$:Ni^{2+} complex (BP86).

Angle (°)	**Model**
N1-Ni-N2	93.0
N1-Ni-N3	87.4
N1-Ni-N4	90.4
N1-Ni-N5	89.6
N1-Ni-N6	176.8
N2-Ni-N3	76.7
N2-Ni-N4	176.6
N2-Ni-N5	103.4
N2-Ni-N6	89.0
N3-Ni-N4	103.6
N3-Ni-N5	177.0
N3-Ni-N6	90.6

.

A5.4.9 Optimized cartesian coordinates for the inner-coordination sphere of the *cis*-HPhen1$_2$:Ni^{2+} complex (OLYP).

<Feb27-2010> <07:11:25> Geometry Converged
  Coordinates in Geometry Cycle 13

Atom	X	Y	Z (Angstrom)
1.C	3.713803	-2.392940	4.257187
2.C	2.896331	-1.451127	3.440179
3.C	2.081145	-1.619910	2.346650
4.N	2.803770	-0.105677	3.721713
5.C	1.967507	0.472253	2.833562
6.N	1.490425	-0.421308	1.981880
7.C	-2.716451	2.614526	4.972299
8.C	-2.371958	1.764641	3.796451
9.C	-1.262008	1.700191	2.990533
10.N	-3.220818	0.816594	3.269257
11.C	-2.627312	0.236173	2.203865
12.N	-1.422789	0.747912	1.995832
13.Ni	-0.031127	-0.003427	0.429434
14.H	4.751508	-2.054229	4.347301
15.H	3.307549	-2.493390	5.270728
16.H	3.723395	-3.384484	3.801801
17.H	1.915609	-2.532706	1.798518
18.H	3.287325	0.375940	4.466211
19.H	1.760717	1.527576	2.841783
20.H	-3.001167	2.008995	5.839607
21.H	-1.860154	3.227312	5.260596
22.H	-3.550549	3.289514	4.748782
23.H	-0.366970	2.289442	3.090818
24.H	-4.136423	0.579033	3.624756
25.H	-3.101715	-0.541999	1.630498
26.C	2.703242	-2.029957	-2.422951
27.C	-1.656310	-4.126880	1.122740
28.C	1.175843	4.276247	-0.068911
29.C	-2.555512	1.322035	-3.024046
30.C	1.688392	-1.776001	-1.485937
31.C	-0.957732	-2.917399	1.269678
32.C	0.523200	3.111178	0.343274
33.C	-1.673185	1.302255	-1.939374
34.C	3.480315	-0.988263	-2.876155
35.C	-2.526658	-4.281161	0.068030
36.C	2.255043	4.169924	-0.924041
37.C	-3.276674	0.184755	-3.319009
38.C	4.008859	1.442241	-2.793994
39.C	-3.647744	-3.269100	-1.906241
40.C	-3.865236	-2.199774	-2.752213
41.C	3.764517	2.711391	-2.315571
42.C	-2.728531	-3.208198	-0.832873
43.C	3.247893	0.318245	-2.386570
44.C	-3.125101	-0.969652	-2.526835
45.C	2.667328	2.904911	-1.384195
46.C	-1.976154	-2.025429	-0.600407
47.C	2.207108	0.470635	-1.434986

48.C	-2.191393	-0.896494	-1.455953
49.C	1.923177	1.779218	-0.930121
50.H	2.855910	-3.042910	-2.781057
51.H	-1.500637	-4.923780	1.842979
52.H	0.822342	5.243079	0.273968
53.H	-2.658160	2.223268	-3.618983
54.H	1.060908	-2.582621	-1.127127
55.H	-0.284864	-2.772206	2.103493
56.H	-0.337114	3.167551	0.997728
57.H	-1.099298	2.185053	-1.692237
58.H	4.266931	-1.158423	-3.605769
59.H	-3.072191	-5.210556	-0.070791
60.H	2.749618	5.074807	-1.259615
61.H	-3.947819	0.201225	-4.170843
62.N	-1.088137	-1.899402	0.421176
63.N	-1.486175	0.230433	-1.177310
64.N	0.888426	1.895507	-0.055899
65.N	1.442788	-0.562921	-1.003362
66.N	4.480483	3.798913	-2.762165
67.N	-4.728298	-2.296529	-3.807449
68.H	-5.101777	-1.473540	-4.246055
69.H	-5.322769	-3.107583	-3.869982
70.H	5.303879	3.599058	-3.310092
71.H	4.613284	4.577822	-2.137835
72.H	4.815076	1.285258	-3.504867
73.H	-4.204924	-4.189548	-2.053730

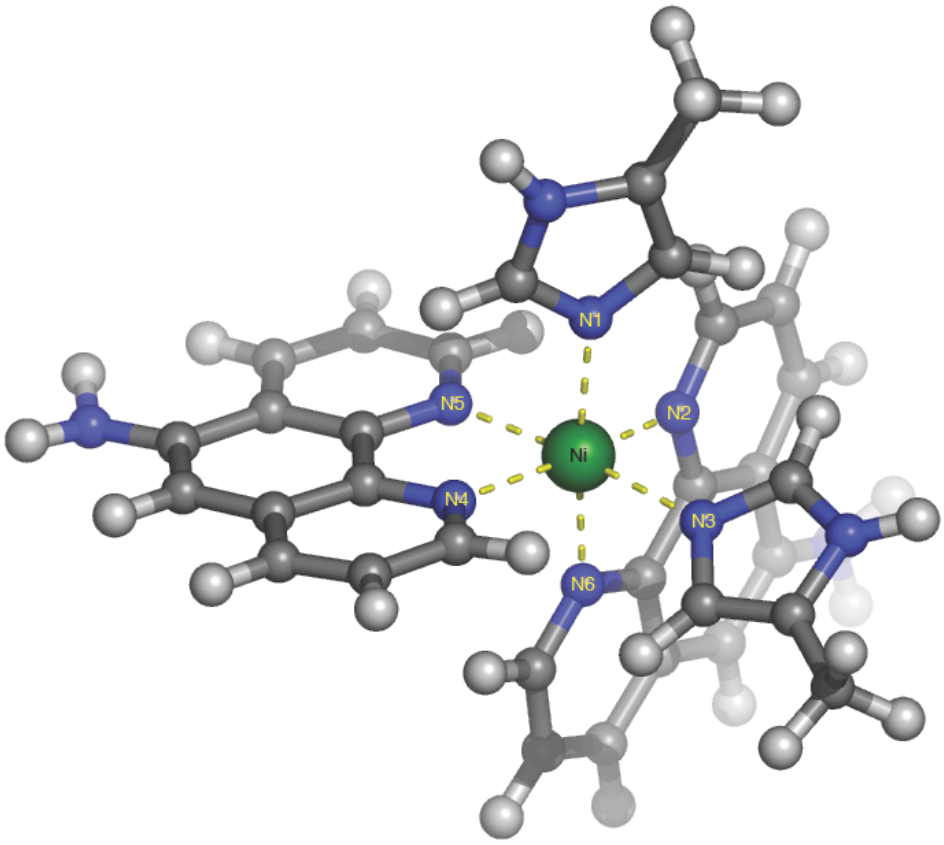


Figure A5.4.10. Optimized geometry of the inner-coordination sphere for the $HPhen1_2:Ni^{2+}$ complex (OLYP).

Table A5.4.11. Calculated bond distances for the computationally minimized inner-coordination sphere of the $HPhen1_2:Ni^{2+}$ complex (OLYP).

Bond (Å)	Model
Ni-N1	2.23
Ni-N2	2.16
Ni-N3	2.21
Ni-N4	2.17
Ni-N5	2.18
Ni-N6	2.13

Table A5.4.12. Calculated bond angles for the computationally minimized inner-coordination sphere for the $HPhen1_2:Ni^{2+}$ complex (OLYP).

Angle (°)	Model
N1-Ni-N2	97.3
N1-Ni-N3	89.9
N1-Ni-N4	89.6
N1-Ni-N5	93.7
N1-Ni-N6	173.9
N2-Ni-N3	91.7
N2-Ni-N4	166.4
N2-Ni-N5	91.4
N2-Ni-N6	77.6
N3-Ni-N4	99.9
N3-Ni-N5	174.77
N3-Ni-N6	86.9

A5.4.13 Optimized cartesian coordinates for inner-coordination sphere of the *trans*-$HPhen1_2:Ni^{2+}$ complex (OLYP).

```
<Feb27-2010> <10:39:01>  Geometry Converged
  Coordinates in Geometry Cycle 17
    Atom         X           Y           Z   (Angstrom)
    1.C       -0.378602   -0.741591   -5.837344
    2.C       -0.229027   -0.168909   -4.469181
    3.C       -0.307592   -0.721247   -3.214199
    4.N        0.054664    1.159724   -4.234325
    5.C        0.133734    1.365555   -2.901788
    6.N       -0.079810    0.236484   -2.243113
    7.C       -0.527704   -1.757903    5.379347
    8.C       -0.229236   -1.001967    4.129398
    9.C       -0.374093   -1.297311    2.795375
   10.N        0.319648    0.262845    4.138104
   11.C        0.481092    0.681448    2.865477
   12.N        0.067513   -0.245462    2.014534
   13.Ni      -0.006733   -0.044517   -0.122052
```

14.H	-1.085650	-2.667195	5.150417
15.H	0.390813	-2.048931	5.901143
16.H	-1.130335	-1.160807	6.072516
17.H	0.551011	0.799250	4.961693
18.H	0.889996	1.646975	2.620557
19.H	-0.765535	-2.204433	2.364435
20.H	-1.179313	-0.246482	-6.397915
21.H	0.545662	-0.641967	-6.417704
22.H	-0.620723	-1.804085	-5.778470
23.H	-0.512041	-1.749996	-2.968731
24.H	0.333620	2.333772	-2.474468
25.H	0.172949	1.871274	-4.941635
26.C	-4.050286	-1.839631	-0.039377
27.C	1.188421	-4.453450	-0.861949
28.C	-1.274739	4.462314	0.755860
29.C	4.052239	1.773969	0.097208
30.C	-2.685858	-1.576292	-0.207221
31.C	0.442512	-3.288670	-0.631410
32.C	-0.538406	3.334844	0.383518
33.C	2.676488	1.556365	0.028591
34.C	-4.880948	-0.810739	0.329920
35.C	2.559550	-4.376736	-0.861394
36.C	-2.626663	4.318086	0.981258
37.C	4.895555	0.691122	-0.029420
38.C	-5.162442	1.586295	0.835925
39.C	4.583750	-2.981532	-0.628216
40.C	5.196463	-1.763468	-0.435821
41.C	-4.659426	2.857988	0.983647
42.C	3.177400	-3.125882	-0.637850
43.C	-4.340007	0.481732	0.512627
44.C	4.362164	-0.591798	-0.244728
45.C	-3.232803	3.059629	0.815279
46.C	2.329500	-1.998286	-0.445537
47.C	-2.939435	0.647192	0.319831
48.C	2.940817	-0.714344	-0.268027
49.C	-2.383634	1.962403	0.469449
50.H	-4.425637	-2.844044	-0.204101
51.H	0.673787	-5.393868	-1.031793
52.H	-0.782833	5.423257	0.867869
53.H	4.436432	2.775893	0.253859
54.H	-2.031560	-2.377978	-0.503633
55.H	-0.633803	-3.367705	-0.617764
56.H	0.520231	3.452180	0.189053
57.H	2.002962	2.394558	0.117857
58.H	-5.945653	-0.974594	0.468231
59.H	3.169670	-5.259334	-1.030955
60.H	-3.203934	5.180911	1.294970
61.H	5.965214	0.849458	0.048589
62.N	-1.049595	2.114191	0.258564
63.N	-2.125157	-0.386292	-0.024207
64.N	0.974384	-2.085140	-0.434002
65.N	2.120385	0.360954	-0.128353
66.N	6.563553	-1.657262	-0.368438
67.N	-5.471103	3.909545	1.338297
68.H	-5.231645	4.838762	1.034076

69.H	-6.464936	3.737826	1.313441
70.H	7.006986	-0.795827	-0.636422
71.H	7.096068	-2.472285	-0.632700
72.H	-6.228364	1.412843	0.947778
73.H	5.190052	-3.868274	-0.787588

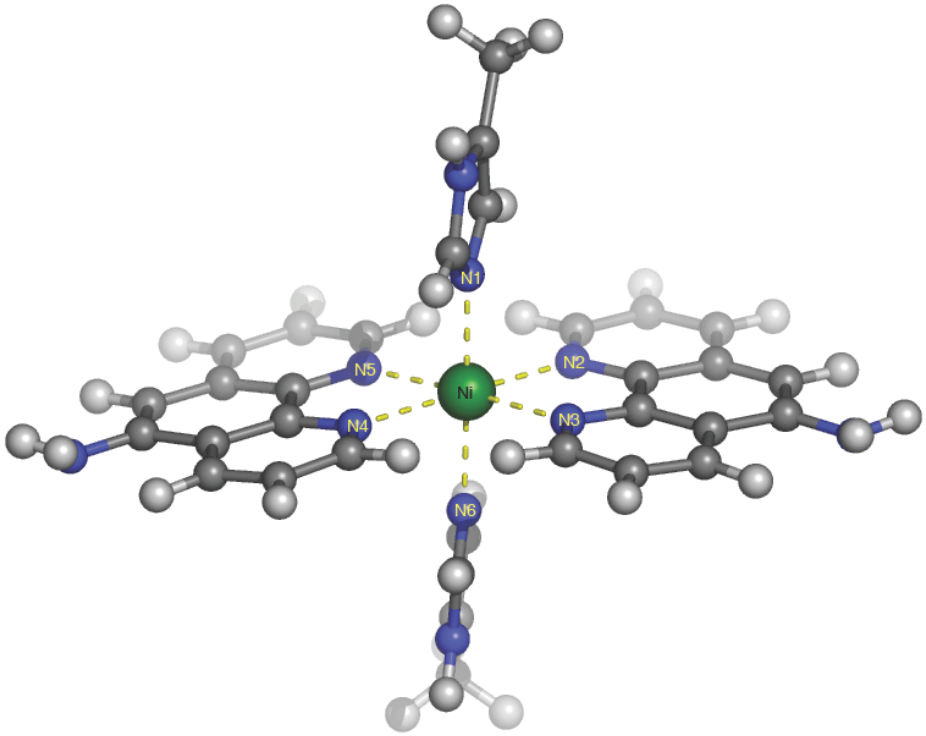


Figure A5.4.14. Optimized geometry for the computationally minimized inner-coordination sphere of the $HPhen1_2:Ni^{2+}$ complex (OLYP).

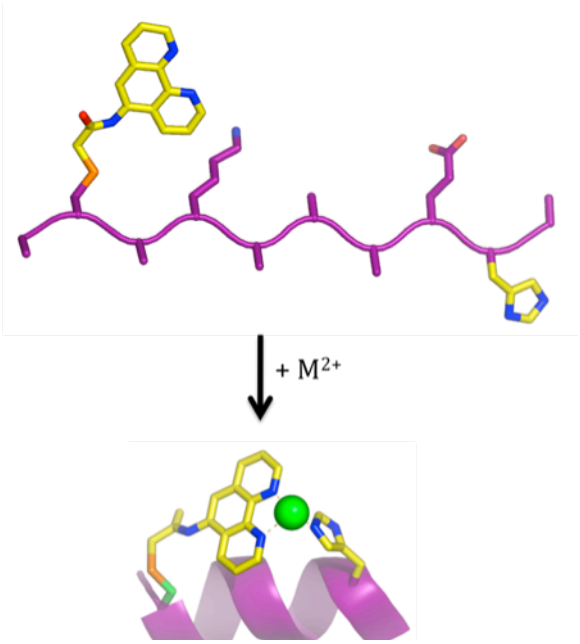
Table A5.4.15. Calculated bond distances for computationally minimized inner-coordination sphere of the *trans*-HPhen1_2:Ni^{2+} complex (OLYP).

Bond (Å)	Model
Ni-N1	2.14
Ni-N2	2.29
Ni-N3	2.17
Ni-N4	2.43
Ni-N5	2.15
Ni-N6	2.15

Table A5.4.16. Calculated bond angles for the computationally minimized inner-coordination sphere of the HPhen1_2:Ni^{2+} complex (OLYP).

Angle (°)	Model
N1-Ni-N2	89.8
N1-Ni-N3	90.3
N1-Ni-N4	91.3
N1-Ni-N5	91.8
N1-Ni-N6	177.8
N2-Ni-N3	75.2
N2-Ni-N4	178.8
N2-Ni-N5	106.7
N2-Ni-N6	92.1
N3-Ni-N4	104.8
N3-Ni-N5	177.0
N3-Ni-N6	89.2

Chapter 6: Site-Specific Localization of A Fluorescent Metal Center on a Small Metal-Stabilized Peptide



6.1 Abstract

Small peptides and proteins hold great potential for a new class of biomolecular pharmaceuticals. Their diverse chemical functionality, originating from the 20 natural (and numerous non-natural) amino acids, make them ideal platforms to target large macromolecular complexes. Moreover, recent advances in peptide stabilization chemistry, most notably hydrocarbon stapling, has lead to improved pharmacokinetic properties of these traditionally unstable molecules. In an attempt to further functionalize the peptide platform, reported here is the design of a small 10-mer peptide outfitted with a His/Phen hybrid coordination motif. The resulting peptide, 10aa-HPhen1, is found to be up to 45% helical (4 °C) in the presence of Ni^{2+}, and can also be used to site-specifically localize a fluorescent rhenium carbonyl complex on its surface. This

proof-of-concept study opens up the possibility for expanded incorporation of metal-based functionality into peptide systems.

6.2 Introduction

The use of peptide and proteins as bioactive compounds has long been heralded as being the foundation for the next wave of pharmaceutical agents [1, 2]. However, these compounds suffer from poor pharmacokinetic properties due to the instability of small peptides, which are traditionally unstructured and thus susceptible to proteolytic cleavage *in vivo*. Recently, advances in peptide and peptidomimetic chemistry have reinvigorated the field [3, 4]. The ability to stabilize the tertiary structure of a peptide, or peptidomimetic, through hydrocarbon stapling [5, 6], disulfide bond formation [7, 8], lactam bridges [9, 10], or hydrogen bond surrogates [11, 12], to name only a few, has demonstrated the potential of using peptide motifs to target protein-protein interactions that were once thought "undruggable" [13-15].

Concurrently, metal ions and metal complexes have also long been touted to be the foundation of an impending renaissance in pharmaceutical agents and diagnostic tools [18]. While metal complexes have an inherent structural and functional diversity that impart numerous advantages for medical and analytical purposes [19, 20], their intrinsic reactivity and lack of specificity renders most metal complexes too cytotoxic for general use and have, thus far, prevented metal-based therapeutics (other than cisplatin and its derivatives) and diagnostics from gaining wide acceptance in the medical and pharmaceutical communities. Therefore, it would be of great utility to design a platform

that combines the specificity and biological activity of a protein/peptide with the functional diversity of a metal complex. This new platform would create a new class of metal-based therapeutics or diagnostic agents that could specifically localize metal-based functionality (MBFs) in a cellular environment.

To this end, reported here the generation of a new class of metal-binding peptides, which are able to site-specifically localize metal ions and metal complexes on their surface as the result of an installed hybrid coordination motif (HCM) on its surface. As a direct consequence of the installed HCM, two small (10 and 20 amino acid) peptides were found to undergo structural changes consistent with the acquisition of an α-helical fold upon the addition of late first-row transition metals. In addition, the same motif can be utilized to localize a fluorescent tricarbonylrhenium(I) polypyridyl species on the peptides surface creating a coordinatively saturated metal center that simultaneously induces structure in the short 10-mer peptide, the results of which are discussed herein.

6.3 Results and Discussion

Beginning with the pioneering work of Ghadiri [21] and Hopkins [22], it has been repeatedly shown [23-27] that metal ions can serve as nucleation sites for the formation

20aa-HPhen1: AC-A-(EAAAK)$_2$-HAAAAAAC-A-CONH$_2$
10aa-HPhen1: AC-AHAKAAAEC-A-CONH$_2$

Figure 6.1. The amino acid sequence for the peptides 20aa-HPhen1 and 10aa-HPhen1. These alanine rich peptides have the requisite histidine and cysteine (red) for HCM formation, as well as a minimal number of salt bridges (blue) to aid in solubility and structure stabilization.

of helical peptides. However, these systems, which utilize both natural and unnatural amino acids, suffer from low metal-ion affinity and/or a lack or generalizability making their utility *in vivo* (where metal ion distribution and concentrations are tightly controlled) unlikely. Recently, our group reported the development of a protein-based metal-binding motif that utilizes a non-natural bidentate metal chelator in concert with a natural histidine residue to create a high-affinity metal-binding site on the surface of an α-helical protein [28, 29]. This binding motif, which is called a Hybrid Coordination Motif (HCM), was shown to: 1) have high-affinity (nM-fM) for late first-row transition metals, 2) stabilize the tertiary structure of the helical protein as a result of the *i*/*i*+7 positioning of the coordinating ligands, 3) direct the formation of discrete and biologically relevant protein dimers, and 4) to be modular with respect to both the non-natural bidentate chelator, as well as the general location and orientation of the HCM. To increase the biological utility of HCMs the feasibility of a peptide scaffold was explored. The ability to install a HCM on a peptide would allow for the construction of a modular and readily adaptable system for the localization MBFs in a cellular environment.

In order to test the ability of the His/Phen HCM to bind metal on a peptide framework, the motif was first install on a simple alanine-rich model peptide. Based on previous studies [21, 30], the 20-mer peptide (Figure 6.1) was expected to show partial helicity in solution (due to the presence of two Glu-Lys salt bridges), thereby allowing us to determine if metal binding would have a stabilizing or de-stabilizing effect. The peptide was produced using standard FMOC-based solid-state synthesis, on a MilliPore 9500 automated peptide synthesizer. The peptide sequence contains a single histidine and cysteine residue in the *i*/*i*+7 positions, with the histidine residing toward the N-terminus

of the peptide in a fashion analogous to the HPhen1 system [29]. The peptide thus obtained, which is called 20aa, was subsequently labeled with a cysteine-specific iodoacetamide-derivatized 1,10'-phenanthroline moiety and purified via high-pressure liquid chromatography (HPLC) to greater than 85% purity, as determined by analytical HPLC. The UV-visible absorption spectra (Figure 6.3a) of 20aa-HPhen1 exhibits the expected π-π* transition originating from the phenanthroline group, centered at λ_{max}= 268 nm (ε = 16 mM^{-1} cm^{-1}) in 20 mM sodium borate buffer (pH 7).

For our initial studies, the effect metal ions would have on the helicity of 20aa-HPhen1, was explored. Previously, it was shown that the two-point attachment of the *i*/*i*+7 HCM system served to stabilize the protein fold by cross-linking a 7 residue two-helical-turn segment of the proteins surface [28, 29]. In the same vein, the His/Phen HCM on 20aa-HPhen1 should be able to serve as a metal-mediated cross-link, leading to an increase in helicity in the presence of M^{2+} ions. As see in Figure 6.2, addition of late first-row transition metals (Co^{2+}, Ni^{2+} or Zn^{2+}) to 20aa-HPhen1 causes a decrease in the molar ellipticity at 222 and 208 nm, as monitored by circular dichroism (CD), indicating that the coordination of metal to the HCM is indeed serving to stabilize α-helix formation. Similar to the cyt cb_{562}-HPhen1, the maximal effect is observed with Ni^{2+}, which displays an increase in helicity from a mean residue ellipticity of -11858 deg•cm^2•$dmol^{-1}$ at 222 nm (~32 % helical) for the apo-peptide to -17537 deg•cm^2•$dmol^{-1}$ (~48% helical) in the presence of a 3 molar excess of Ni^{2+} (at 4°C in 20 mM sodium borate (pH 7)). Taken together, the shift in the phenanthroline absorbance from a metal-free λ_{max}=268 nm to a metal-bound λ_{max}= 274 nm along with an induction of helicity suggests that the 20aa-HPhen1 HCM is binding in an analogous fashion to the cyt cb_{562}-

HPhen1, which adopted a *facial* 3-coordinate octahedral binding geometry. This point is further strengthened by the observation that Cu^{2+}, which binds tightly to metal ligands but is less likely to bind in an octahedral geometry, has a minimal effect of the peptides structure, indicating the His/Phen HCM is still optimized to bind metals with a preference for octahedral binding geometries.

Next, the metal binding ability of 20aa-HPhen1 was investigated. The ability to create a high-affinity (nM to fM) metal-binding site on the surface of a protein was a main advantage of the protein-based HCM system, allowing for the site-specific localization of a Ru(*p*-cymene) complex on the surface of protein [29]. This feature, if imparted on the peptide-based system, would make the HCM system unique over other metal-mediated crosslinks that traditionally use natural metal-binding amino acids and display metal-binding dissociation constants ($K_{d\text{-}metal}$) in the mid-high μM range [21]. Thus a determined preliminary dissociate constant ($K_{d\text{-}Ni}$) for 20aa-HPhen1 with Ni^{2+}.

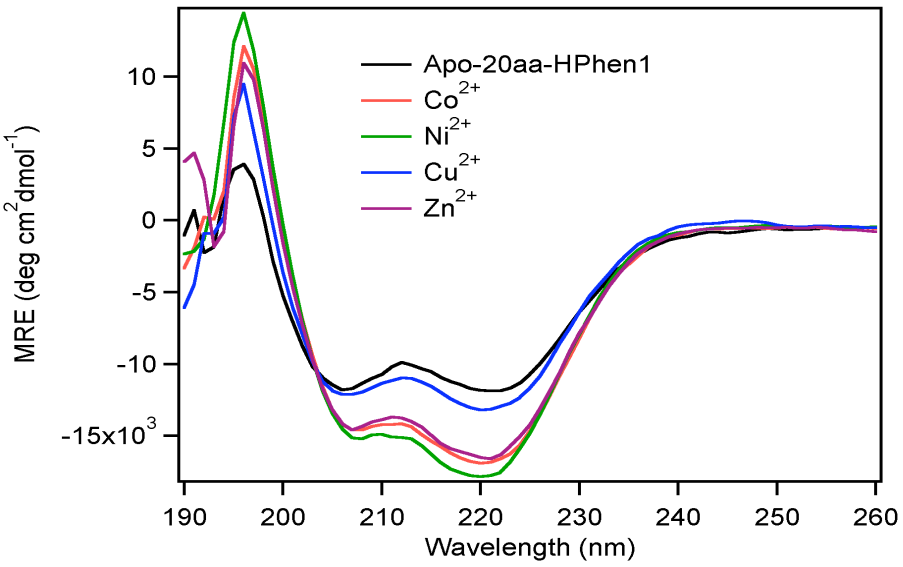


Figure 6.2. The circular dichroism (CD) spectra of 20aa-HPhen1 in the presence, and absence (black), of the late first-row transition metals (Co^{2+}(red), Ni^{2+}(green), Cu^{2+}(blue) and Zn^{2+} (purple)).

The coordination of metal to the Phen moiety was followed spectroscopically by monitoring a 6-nm redshift in the Phen UV absorption band from λ_{max}=268 nm for the metal-free species to λ_{max}= 274 nm for the metal-bound form. The changes in the Phen absorbance was paralleled by a decrease in molar ellipticity at 222 and 208 nm, respectively, as monitored by CD spectroscopy (Figure 6.3), indicating coordination of the histidine residue in the HCM motif. From these data, a preliminary dissociation constant (K_d) for metal binding can be obtained and was found to be ~ 1 nM for Ni^{2+}, in good agreement (within an order-of-magnitude) of the cyt *cb*$_{562}$-HPhen1 system.

With the ability to incorporated His/Phen HCMs onto a model 20aa peptide, we

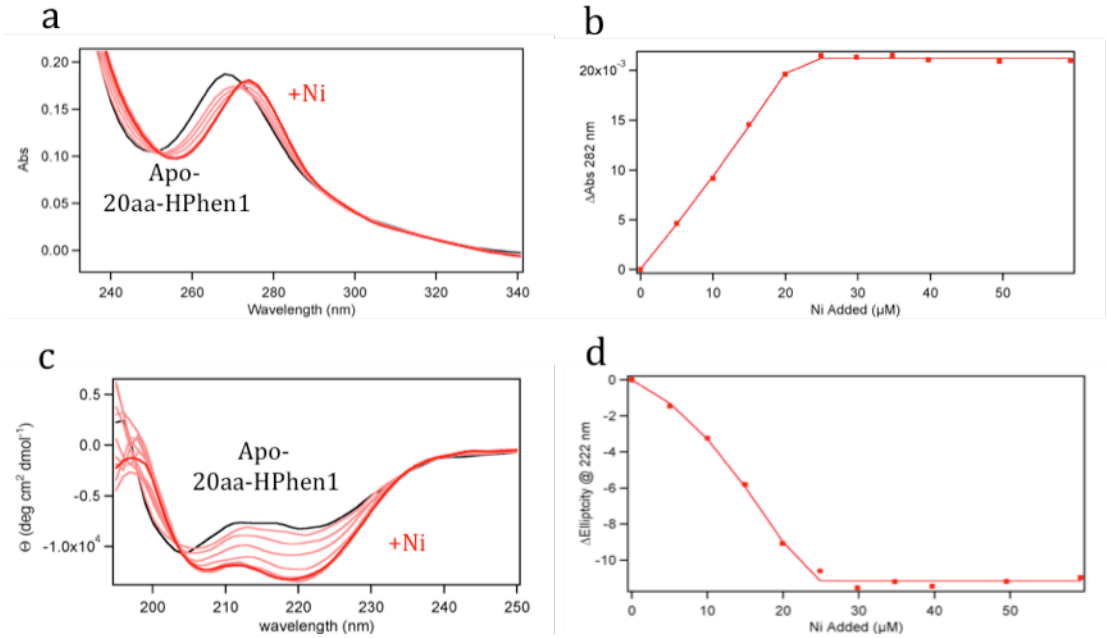


Figure 6.3. The observed spectroscopic changes (CD and UV-vis) in the 20aa-HPhen1 sprectra upon the sequential addition of Ni^{2+}. (a) A diagnostic 6 nm redshift in the phenanthroline π-π* absorption band (λ_{free}= 268 nm; λ_{bound}: 274 nm), indicates the Ni^{2+} is coordinating to the Phen group. Plotting the changes in absorbance (ΔA 282 nm) vs. Ni^{2+} added, yields the binding isotherm in (b). (c) Decreases in the elliptical signal at 222 and 208 nm, respectively, suggest the addition of Ni^{2+} is increasing the helicity of the peptide. Plotting the changes in ellipticity at 222 nm, as a function of Ni^{2+} added, yields the binding isotherm in (d). Both binding isotherms were fit assuming a simple 1:1 model and report binding constants of 8 (1) × 10^{-10} M and 2(1) × 10^{-9} M for the UV-vis and CD isotherm, respectively.

next attempted to make the system even smaller. As of late, there is considerable interest [11, 13, 15, 31, 32] in making small well-folded peptides that retain biological activity but have improved pharmacokinetics properties over longer peptides, which suffer from short circulation half-lives and enzymatic proteolysis due to their unstructured nature. To date, most strategies for cross-linking short peptides into helical conformations have relied on make stable, but non-functional, covalent bonds [3]. Therefore, it would be of great utility if the HCM system could stabilize a short peptide while simultaneously introducing MBF into the system. With this objective in mind a 10-amino-acid analog of 20aa-HPhen1, 10aa-HPhen1 (Figure 1), was synthesized. Similar to the 20 amino acid peptide, 10aa-HPhen1 is an alanine-rich peptide with a single histidine and cysteine residue and also retains a lone Glu-Lys salt bridge on the helical face opposite to the His\Cys residues. The peptide was subsequently labeled with the iodoacetamide functionalized phenanthroline derivative and purified (>85%) via HPLC chromatography.

Initial spectroscopic experiments suggested that 10aa-HPhen1 is primarily in random coil conformation with a strong negative signal at ~200 nm, as determined by CD spectroscopy (Figure 6.4). Upon the addition of excess Ni^{2+}, a large decrease in ellipticity at 205 and 218 nm, respectively, is observed, suggesting the formation of an α-helical structure in solution (Figure 6.4a). While, once again, the largest change is observed in the presence of Ni^{2+}, all other metals tested (Co^{2+}, Cu^{2+}, Zn^{2+}) also induce structural changes to various extents (Figure 6.4b), indicating the 10aa-HPhen1 systems is sufficiently flexible to accommodate the different stereochemical preference of the metal ions tested.

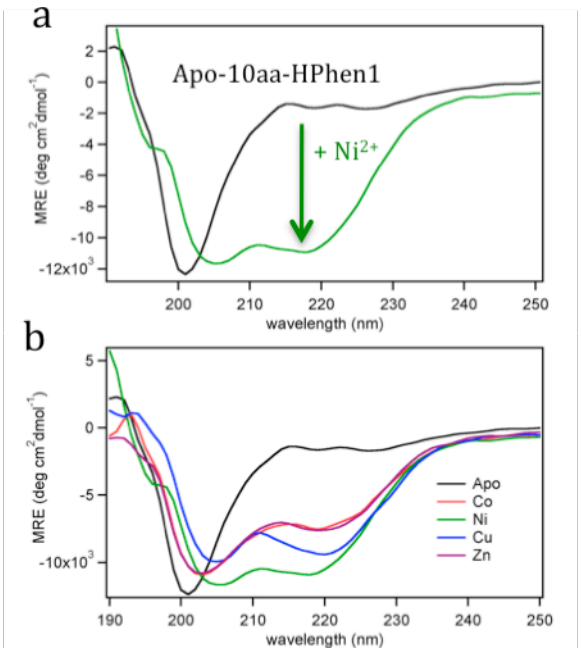


Figure 6.4. (a) The spectroscopic changes in the CD spectrum of 10aa-HPhen1 upon addition of Ni^{2+}. (b) Changes in the CD spectrum for 10aa-HPhen1 in the presence, and absence (black), of the late first-row transition metals Co^{2+}(red), Ni^{2+} (green), Cu^{2+}(blue), and Zn^{2+}(purple).

With the establishment of the His/Phen HCM on a short peptide, the ability to introduce MBF on the peptide surface, was explored. As a proof-of-concept, a $[Re(Phen)(CO)_3]^+$ species was chosen. The rhenium complex has well-established photophysical properties [33-35] and has been used in biological systems as both a fluorescent probe [35-37], and as a powerful oxidant to induce electron transfer in biological systems [16, 38]. Moreover, the thermodynamic preference of tricarbonyl groups to bind in a *facial* geometry in octahedral complexes makes the tricarbonyl rhenium species ideally suited for the tridentate HCM systems. As a precursor, a previously reported [35] cysteine-specific derivative of the rhenium compound, $Re(IPhen)(CO)_3Cl$, was synthesized. The iodoacetamide group allows for facile incorporation of the metal complex onto the peptide, obviating the need for the long

reaction times (>7 days) that plague the traditional routes of incorporation [39]. The rhenium analog of 10aa-HPhen1, 10aa-Re(HPhen1)$(CO)_3$, was prepared in high yields (>90%) by reacting Re(IPhen)$(CO)_3$Cl with the unlabeled peptide in a solution consisting of 50/50 (v/v) mixture of acetonitrile in 0.1 M HEPES buffer (pH 7.5). The mixture was allowed to react at R.T., in the dark, under anaerobic conditions for a period of 4 hours, at which point the peptide was exchanged into water via a G-10 (Bio-Rad) gel-filtration column and purified via HPLC.

After purification of 10aa-Re(HPhen1)$(CO)_3$, the mass of the complex revealed

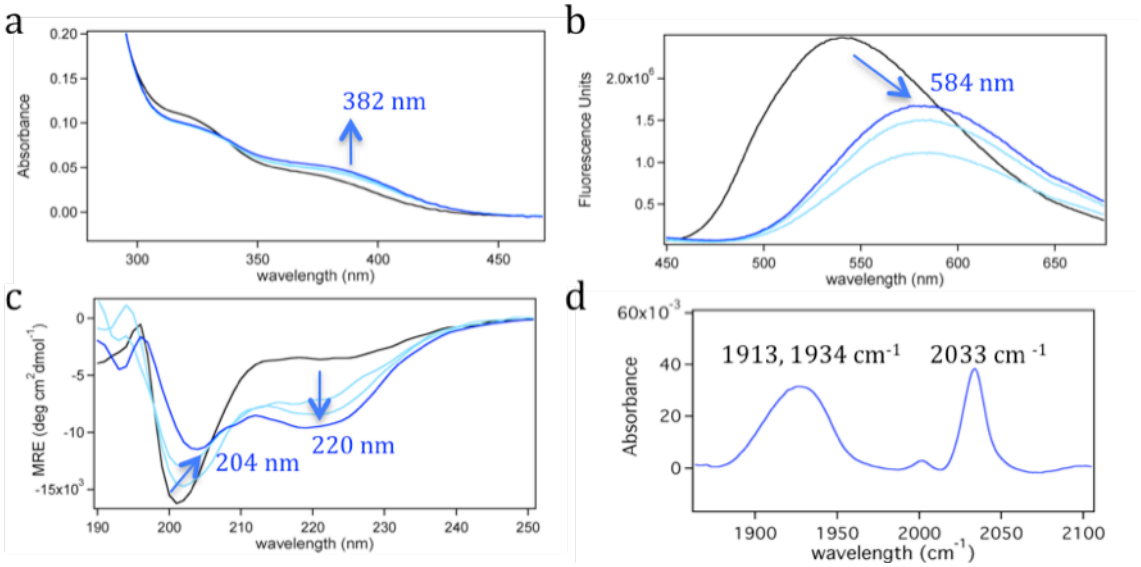


Figure 6.5. The spectroscopic changes (UV-vis, CD, fluorescence and IR) observed in 10aa-Re(HPhen1)$(CO)_3$ upon heating at 65°C for 6 hours. (a) The initial UV-visible absorption spectrum of 10aa-Re(HPhen1)$(CO)_3$ displays 3 main absorption bands, 282 nm (not shown), which corresponds to the π-π* transition from the Phen group, along with two broad MLCT bands at 325 and 380 nm, respectively. Upon heating, an increase in the MLCT at 380 nm is observed with a concomitant decrease in the MLCT band at 325 nm. A clean isosbestic point is observed at 334 nm, indicating a simple two-state process. (b) A change in fluorescence emission maximum, and intensity, is observed for 10aa-Re(HPhen1)$(CO)_3$ upon heating, moving from a λ_{em}: 534 nm to a λ_{em}: 584 nm; the latter of which is diagnostic of a $[Re(Phen)(His)(CO)_3]^+$ species [16]. (c) A decrease in ellipticity at 220 and 204 nm, respectively, along with an increase in the CD signal at 200 nm that is observed upon heating, suggest a more structured (helical) peptide. (d) The IR spectrum of the carbonyl region of the 10aa-Re(HPhen1)$(CO)_3$ after heating. The three observed bands at 1913, 1934 and 2033 cm^{-1}, respectively, are indicative of 3 CO groups bound in a *facial* manner on an octahedral rhenium complex [17].

that a single Re(Phen)$(CO)_3$ entity was successfully incorporated onto the peptide (determined MW = 1488 amu; expected MW = 1490 amu). The UV-visible absorption spectra of the peptide had three prominent bands centered at 282, 325 and 380 nm, respectively (Figure 6.5a). These absorption bands correspond to the π-π* transition from the Phen group, as well as a broad MLCT band [16], which, when excited (λ_{ex} = 330 nm), exhibit a strong fluorescent signal (λ_{em} = 532 nm) (Figure 6.5b). Finally, the CD spectrum of the peptide displays a strong minimum at 200 nm, indicating that the histidine did not coordinate during the labeling reaction and the peptide was still random coil (Figure 6.5c).

Upon heating (65°C) 10aa-Re(HPhen1)$(CO)_3$ for as little as 2 hours, the peptide undergoes dramatic spectroscopic changes. The MLCT band centered at 380 nm increases with a concomitant decrease in the absorption band at 325 nm. The intensity of the emission spectra is diminished and red-shifted to a λ_{em}= 584 nm, characteristic of a $[Re(Phen)(His)(CO)_3]^+$ species [16]. Furthermore, a large decrease in molar ellipticity is observed at 204 and 220 nm, respectively, with a simultaneous increase in the ellipticity at 200 nm. Heating was continued for a period of 6 hours, after which time no significant spectral changes were observed. After heating, the mean residue ellipticity at 220 nm was -9654 deg•cm^2•$dmol^{-1}$ (at 4°C in 20 mM sodium borate (pH 7)), which corresponds to a peptide that is 36% helical in solution. In order to ensure that the carbonyl groups remained in tact during the heating process, the peptide was lyophilized, mixed with dry KBr and the infrared (IR) spectrum taken. The IR spectrum shows three main absorption features in the carbonyl region at 1913, 1934 and 2033 cm^{-1}, respectively (Figure 6.5c). These bands coincide with the expected transitions for 3 carbonyl groups, coordinated in

a *facial* geometry on an octahedral rhenium center [17]. Taken together, these data strongly indicates the formation of a coordinatively saturated $[Re(Phen)(His)(CO)_3]^+$ on what is now a partially helical peptide (Figure 6.6).

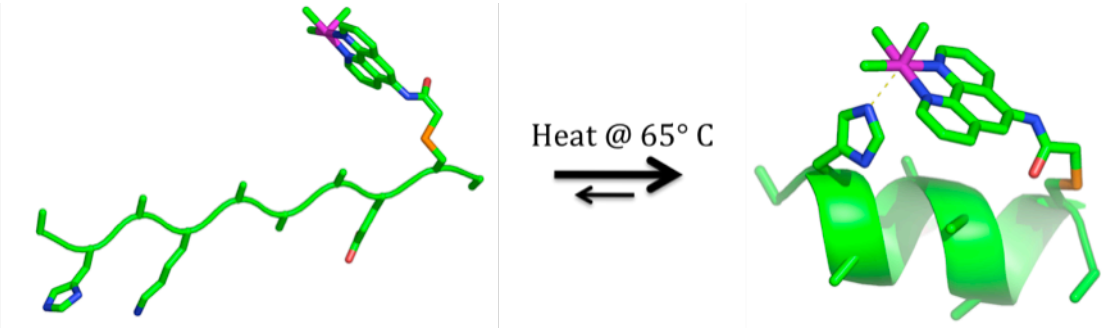


Figure 6.6. A cartoon schematic illustrating the proposed method of metal-mediated stabilized of the $Re(Phen)(CO)_3$ complex by the His\Phen HCM.

6.4 Conclusions

The ability to make small, functional, metal-peptide complexes will expand the utility peptide based therapeutics and diagnostics by allowing for the facile incorporation of metal-based functionality into bioactive compounds. The small stabilized peptides, 20aa-HPhen1 and 10aa-HPhen1, represent the first in a class of metal binding peptides, which, as a function of the incorporated HCM, can 1) tightly coordinate labile metal ions and/or 2) site-specifically localize metal complexes; all while stabilizing the helical fold of the peptide. This proof-of-concept study could lead to the possibility of novel metallopeptides that can synergistically utilize both a functional metal complex, as well

as the native bioactivity of a peptide framework to achieve a multivalent therapeutic, the consequences of which are currently being explored in our laboratory.

6.5 Experimental Section

Chemicals and Reagents

Unless otherwise noted, all chemicals were purchased from Fisher Scientific, VWR or Sigma Aldrich and were used without further purification. N^{α}-FMOC amino acids, with appropriate side-chain protecting groups, and PEG-PAL-PS resin were purchased from Applied Biosystems. Synthesis of 5-iodoacetamide-1,10-Phenanthroline [40] and Re(IPhen)$(CO)_3$Cl [35] was prepared as described previously.

Synthesis of 20aa and 10aa peptides

Peptides were synthesized with N-acylations and C-terminal amides on a MilliPore 9050 peptide synthesizer, using standard FMOC chemistry. Briefly, for each amino acid added, the FMOC protecting group on N-terminus of the growing peptide was removed with a solution of 20% (v/v) 4-methylpiperidine in dimethylformamide (DMF). After exchange into DMF, the resin was continually washed, for 30 min, with a solution containing a 4× molar equivalent of a free-acid amino acid and (2-(7-Aza-1H-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate) (HATU), dissolved in a minimal volume of 10% (v/v) 2,4,6-trimethylpyridine (collidine). After the 30 min reaction time, the resin is further washed with DMF, and the process repeats itself until all amino acids have been incorporated. Once all amino acids have been added, the N-

terminus of the peptide is acylated with a mixture containing: 0.5 M acetic acid anhydride, 0.5 M N-hydroxybenzaldehyde (HOBt), 10% (v/v) dichloromethane (DCM) in a DMF solution.

The peptide resin is subsequently exchanged into DCM, isolated and dried *in vacuo*. Once dry, the peptide is cleaved using "Reagent R", a common cleavage cocktail, containing: 5% (v/v) thioanisole, 3% (v/v) ethane dithiol and 0.2 % (v/v) anisole dissolve in trifluoroacetic acid (TFA). Reagent R is allowed to cleave the peptide from the resin for a period of 2 hours, at which point the mixture is filtered. The filtrate, which contains the cleaved peptide, is added to a solution of cold ether and allowed to precipitate overnight at -80C. The next day, the peptide is isolated via filtration, dissolved in a mixture of water and acetonitrile and purified via HPLC (Waters).

Peptide	Determined Mass	Expected Mass
10aa	983 amu	983 amu
20aa	1810 amu	1810 amu

Functionalization of 10aa or 20aa with IPhen

15 mg of either 10aa or 20aa was dissolved in 2 mL of degassed 0.1 M HEPES buffer (pH 7.5) under an inert atmosphere. Separately, a 3× molar excess of IPhen was dissolved in a minimal volume of DMF (typically 1-2 mLs), and added drop-wise to the dissolved peptide solution. The reaction mixture was allowed to stir at room temperature, in the dark, for a period of 4 hours. Upon completion of the reaction, the mixture was exchanged into water using a G-10 column (Bio-Rad) gel filtration column with a 700 MW cut-off. The isolated solution was loaded directly onto the HPLC (Waters C18 column (300 mm × 19 mm)) and purified using a gradient of 5-40% B over 30 min with a

10 ml/min flow rate. (Buffer A = H_2O with 0.1% (v/v) TFA; Buffer B = acetonitrile). The peptide identity and initial purity were determined by MALDI-TOF mass spectrometry. Fractions containing the pure labeled peptide were pooled and lyophilized. The purity of the peptide was verified using an analytical HPLC was determined to be greater than 85%.

Peptide	Determined Mass	Expected Mass
10aa-HPhen1	1219 amu	1219 amu
20aa-HPhen1	2049 amu	2049 amu

Functionalization of 10aa with Re(IPhen)(CO)$_3$Cl

The labeling and purification of 10aa with Re(IPhen)(CO)$_3$Cl occurred under identical conditions as the IPhen reaction (*vide supra*), with the exception that the Re(IPhen)(CO)$_3$Cl label was dissolved in degassed acetonitrile. MALDI mass of 10aa-Re(Phen)(CO)$_3$: 1488 amu (expected: 1490 amu).

CD Spectroscopy

General

Unless otherwise noted, all CD spectra were taken on an Aviv 215 spectrometer at 4°C. Every reported spectra is the average of 5 wavelength scans (190-260 nm) measured at 1 nm intervals with an averaging time of 5 seconds. After completion of the wavelength scans, the spectra were averaged, smoothed with a binomial function and corrected for any buffer signal. All CD data is reported in terms of mean residue

ellipticity. Fraction helicity was calculated based on the method reported by Fairlie [41], assuming a finite length correction faction (*k*) of 3.

Wavelength scans with late first-row transition metals (Co^{2+}, Ni^{2+}, Cu^{2+} and Zn^{2+})

For each metal:peptide sample, an aliquot of the appropriate peptide (from a concentrated peptide stock) was mixed with an aliquot of a 5 mM M^{2+} or EDTA (to ensure a metal-free environment) solution in 20 mM sodium borate buffer (pH 7). Each sample was further diluted with 20 mM sodium borate buffer (pH 7) to a final volume of 2.5 mL. The final concentration of all peptides ranged from 10-15 μM along with 50 μM of M^{2+} or EDTA. All concentrations were verified using the phenanthroline group as a spectroscopic handle. (20aa-HPhen1: ε_{268} = 16 mM^{-1} cm^{-1}; (Co:20aa-HPhen1): ε_{274} = 13.8 mM^{-1} cm^{-1}; (Ni:20aa-HPhen1): ε_{274} = 14.2 mM^{-1} cm^{-1}; (Cu:20aa-HPhen1): ε_{274} = 12.7 $mM^{-1}cm^{-1}$); (Zn:20aa-HPhen1): ε_{274} = 14.3 mM^{-1} cm^{-1}).

Formation of 6-coordinate 10aa-Re(HPhen)$(CO)_3$

A 2.5 mL solution containing 25 μM 10aa-Re(HPhen1)$(CO)_3$ was heated at 65°C in an Aviv 215 spectrometer for 1 hour intervals. After each hour, the temperature was dropped to 4°C and a wavelength scans were taken. The sample was allowed to warm to 25°C, at which point the UV-visible and emission spectrum were taken. Then sample was then placed back in the Aviv 215 and heated for an addition 1-hour interval. This process was repeated until no significant spectroscopic changes were observed (~ 6 hours).

Infrared spectrum of 10aa-Re(HPhen1)(CO)$_3$

The solution of 10aa-Re(HPhen1)(CO)$_3$ used for spectroscopic studies (*vide supra*) was evaporated to dryness using a speed-vac concentrator (Savant). The resulting solid was mixed with potassium bromide (KBr) and a pellet made by placing the mixture under pressure. The infrared spectra were collected on a Bruker Equinox 55 FTIR spectrometer.

6.6 References

[1] S.D. Nuttall, R.B. Walsh, *Curr Opin Pharmacol*, 8 (2008) 609-615.

[2] A.J. Dorner, S. Projan, *Curr Opin Pharmacol*, 8 (2008) 598-599.

[3] L.K. Henchey, A.L. Jochim, P.S. Arora, *Curr. Opin. Chem. Biol.*, 12 (2008) 692-697.

[4] M.J.I. Andrews, A.B. Tabor, *Tetrahedron*, 55 (1999) 11711-11743.

[5] C.E. Schafmeister, J. Po, G.L. Verdine, *J. Am. Chem. Soc.*, 122 (2000) 5891-5892.

[6] H.E. Blackwell, R.H. Grubbs, *Angew. Chem. Int. Ed.*, 37 (1998) 3281-3284.

[7] D.Y. Jackson, D.S. King, J. Chmielewski, S. Singh, P.G. Schultz, *J Am Chem Soc*, 113 (1991) 9391-9392.

[8] F.Z. Zhang, O. Sadovski, S.J. Xin, G.A. Woolley, *J. Am. Chem. Soc.*, 129 (2007) 14154-14155.

[9] C. Bracken, J. Gulyas, J.W. Taylor, J. Baum, *J Am Chem Soc*, 116 (1994) 6431-6432.

[10] J.C. Phelan, N.J. Skelton, A.C. Braisted, R.S. McDowell, *J Am Chem Soc*, 119 (1997) 455-460.

[11] L.K. Henchey, S. Kushal, R. Dubey, R.N. Chapman, B.Z. Olenyuk, P.S. Arora, *J Am Chem Soc,* 132 (2010) 941-943.

[12] J. Liu, D. Wang, Q. Zheng, M. Lu, P.S. Arora, *J. Am. Chem. Soc.*, 130 (2008) 4334-4337.

[13] R.E. Moellering, M. Cornejo, T.N. Davis, C.D. Bianco, J.C. Aster, S.C. Blacklow, A.L. Kung, D.G. Gilliland, G.L. Verdine, J.E. Bradner, *Nature*, 462 (2009) 182-188.

[14] G.L. Verdine, L.D. Walensky, *Clin. Cancer Res.*, 13 (2007) 7264-7270.

[15] L.D. Walensky, A.L. Kung, I. Escher, T.J. Malia, S. Barbuto, R.D. Wright, G. Wagner, G.L. Verdine, S.J. Korsmeyer, *Science*, 305 (2004) 1466-1470.

[16] W.B. Connick, A.J. DiBilio, M.G. Hill, J.R. Winkler, H.B. Gray, *Inorg Chim Acta*, 240 (1995) 169-173.

[17] K. Nakamoto, Infrared and Raman Spectra of Inorganic and Coordination Compounds - Part B: Applications in Coordination, Organometallic and Bioinorganic Chemistry, 5 ed., Wiley & Sons, Inc., New York, 1997.

[18] T.W. Hambley, *Science*, 318 (2007) 1392-1393.

[19] J.C. Dabrowiak, Metals in Medicine, John Wiley & Sons, Ltd., West Sussex, 2009.

[20] Z.J. Guo, P.J. Sadler, *Angew Chem Int Ed*, 38 (1999) 1513-1531.

[21] M.R. Ghadiri, C. Choi, *J. Am. Chem. Soc.*, 112 (1990) 1630-1632.

[22] F.Q. Ruan, Y.Q. Chen, P.B. Hopkins, *J. Am. Chem. Soc.*, 112 (1990) 9403-9404.

[23] R.L. Beyer, H.N. Hoang, T.G. Appleton, D.P. Fairlie, *J. Am. Chem. Soc.*, 126 (2004) 15096-15105.

[24] G. Goch, M. Maciejczyk, M. Oleszczuk, D. Stachowiak, J. Malicka, A. Bierzynski, *Biochemistry*, 42 (2003) 6840-6847.

[25] K.J. Kise, Jr., B.E. Bowler, *Biochemistry*, 41 (2002) 15826-15837.

[26] M.M. Lopez, D.H. Chin, R.L. Baldwin, G.I. Makhatadze, *Proc Natl Acad Sci U S A*, 99 (2002) 1298-1302.

[27] D.J. Cline, C. Thorpe, J.P. Schneider, *J Am Chem Soc*, 125 (2003) 2923-2929.

[28] R.J. Radford, P.C. Nguyen, T.B. Ditri, J.S. Figueroa, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 4362-4369.

[29] R.J. Radford, P.C. Nguyen, F.A. Tezcan, *Inorg. Chem.*, 49 (2010) 7106-7115.

[30] S. Marqusee, R.L. Baldwin, *Proc. Natl. Acad. Sci. USA*, 84 (1987) 8898-8902.

[31] R.S. Harrison, N.E. Shepherd, H.N. Hoang, G. Ruiz-Gomez, T.A. Hill, R.W. Driver, V.S. Desai, P.R. Young, G. Abbenante, D.P. Fairlie, *Proc Natl Acad Sci U S A*, 107 (2010) 11686-11691.

[32] J.W. Chin, A. Schepartz, *Angew Chem Int Ed*, 40 (2001) 3806-3809.

[33] R.J. Lin, K.S. Lin, I.J. Chang, *Inorg. Chim. Acta.*, 242 (1996) 179-183.

[34] M. Busby, A. Gabrielsson, P. Matousek, M. Towrie, A.J. Di Bilio, H.B. Gray, A. Vlcek, *Inorg. Chem.*, 43 (2004) 4994-5002.

[35] J.D. Dattelbaum, O.O. Abugo, J.R. Lakowicz, *Bioconjug Chem*, 11 (2000) 533-536.

[36] K.L. Haas, K.J. Franz, *Chem. Rev.*, 109 (2009) 4921-4960.

[37] K.K.W. Lo, *Top Organometal Chem*, 29 (2010) 115-158.

[38] C. Shih, A.K. Museth, M. Abrahamsson, A.M. Blanco-Rodriguez, A.J. Di Bilio, J. Sudhamsu, B.R. Crane, K.L. Ronayne, M. Towrie, A. Vlcek, Jr., J.H. Richards, J.R. Winkler, H.B. Gray, *Science*, 320 (2008) 1760-1762.

[39] A.M. Blanco-Rodriguez, A.J. Di Bilio, C. Shih, A.K. Museth, I.P. Clark, M. Towrie, A. Cannizzo, J. Sudhamsu, B.R. Crane, J. Sykora, J.R. Winkler, H.B. Gray, S. Zalis, A. Vlcek, Jr., *Chem. Eur. J.*, 17 (2011) 5350-5361.

[40] C.H.B. Chen, L. Milne, R. Landgraf, D.M. Perrin, D.S. Sigman, *ChemBioChem*, 2 (2001) 735-740.

[41] N.E. Shepherd, H.N. Hoang, G. Abbenante, D.P. Fairlie, *J Am Chem Soc*, 127 (2005) 2974-2983.

Chapter 7: Concluding Remarks

Throughout this thesis, there has been one unifying theme: inorganic chemists, can use proteins and peptides as "new" ligand platforms for coordination chemistry. Although this idea is not close to being new, (nature has been using proteins as metal platforms for billions of years, and chemists have exploited the interiors of stable proteins and peptide assemblies to the same end), proteins had not been considered as building blocks for inorganic chemistry in the traditional sense. It has been shown that proteins can indeed be treated as such: with the appropriate placement of metal-binding sidechain functionalities on the surface, they form complexes dictated by the stereochemical preferences of the metal ions. They can be exploited as sterically bulky ligands to create coordinatively unsaturated metal centers within interfaces. They can be arranged into ordered metal-driven frameworks just like their small organic counterparts. All this comes with the caveat that metal coordination on protein surfaces is not nearly as predictable as with synthetic ligands because of the chemical heterogeneity of the former. Nevertheless, this chemical heterogeneity also translates into powerful handles–*i.e.*, extensive non-covalent and covalent interactions–to control coordination chemistry, which makes the potential of proteins as novel ligand platforms, very exciting.

Chapters 7 is reproduced, in part, with permission from: Radford, R. J., Brodin, J. D., Salgado, E. N., Tezcan, F. A. 2010. *Coord. Chem. Rev.* DOI:10.1016/j.ccr.2010.10.010. Copyright 2011, Elsevier B.V.