UC San Diego

UC San Diego Electronic Theses and Dissertations

Title

In Vitro and In Vivo Assembly of Supramolecular Metalloprotein Scaffolds via Templated Disulfide Bonds

Permalink

https://escholarship.org/uc/item/8bn9b853

Author

Medina, Annette Marie

Publication Date

2013

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

UNIVERSITY OF CALIFORNIA SAN DIEGO

In Vitro and In Vivo Assembly of Supramolecular Metalloprotein Scaffolds via Templated Disulfide Bonds

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

in

Chemistry

by

Annette Marie Medina

Committee in charge:

Professor Faik Akif Tezcan, Chair
Professor Seth Cohen
Professor Michael Gilson
Professor Elizabeth Komives
Professor William Trogler

2013

© Copyright

Annette Marie Medina, 2013

All rights reserved.

The Dissertation of Annette Marie Medina is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____

Chair

University of California, San Diego

2013

Dedication

I would like to dedicate this thesis to my family, especially to my mother Aida Morales. All of my successes are the fruit of her love, support, and encouragement.

Table of Contents

Signature Page iii

Dedication iv

Table of Contents v

List of Abbreviations vii

List of Figures x

List of Tables xiv

Acknowledgements xv

Vita xvii

Abstract of the Dissertation xviii

CHAPTER 1
- Introduction 1
- Goals and Experimental Approach 13
- References 15

CHAPTER 2
Zn(II)-Assisted Templating of Covalent Protein Macrocycles
- Introduction 22
- Results and Discussion 28
- Conclusions 46
- Materials and Methods 52
- References 58

CHAPTER 3
In vitro and cellular Self-Assembly of a Protein Cryptand via Templated Disulfide Bonds
- Introduction 64
- Results and Discussion 66
- Conclusions 95
- Materials and Methods 99
- References 106

CHAPTER 4
Structural and Spectroscopic Characterization of Engineered Protein Macrocycles Binding Copper
Introduction 115
Results and Discussion 122
Conclusions 141
Materials and Methods 148
References 151

CHAPTER 5
Conclusions 155
Future Directions 158
Materials and Methods 161
References 162

APPENDIX
1. Calculation of the periplasmic concentration of $^{C81/C96}RIDC1_4$ 164
2. Calculation of the redox potential of GSH-GSSG buffer solutions 165
3. Use of the Electron Paramagnetic Resonance Spectrometer 166
4. Matlab code for Easyspin simulations 171
References 172

List of Abbreviations

$^{A74/C38/C96}$MBPC1	Asp74Ala, Cys38, Cys96 variant of MBPC1
$^{A74/C38/C96}$RIDC1	Asp74Ala, Cys38, Cys96 variant of RIDC1
Asp	Aspartate
Az	Azurin
BCB	Blue Copper Binding
$^{C25/C96}$RIDC1	Cys25, Cys96 variant of RIDC1
$^{C38/C96}$MBPC1	Cys38, Cys96 variant of MBPC1
$^{C38/C96}$MBPC1$_4$	Cys38, Cys96 MBPC1 tetramer
$^{C38/C96}$RIDC1	Cys38, Cys96 variant of RIDC1
$^{C38/C96}$RIDC1$_4$	Cys38, Cys96 RIDC1 tetramer
$^{C81/C96}$RIDC1$_4$	Cys81, Cys96 RIDC1 tetramer
^{C96}RIDC1$_2$	Cys96 disulfide mediated RIDC1 dimer
^{C96}RIDC1$_4$	Cys96 disulfide and hydrophobic-interface mediated RIDC1 tetramer
C_α-C_α	alpha carbon- alpha carbon
Cu(I)$_6$:^{C96}RIDC1$_4$	^{C96}RIDC1$_4$ tetramer bound to 6 Cu(I)
Cu$_2$:^{C96}RIDC1$_4$	^{C96}RIDC1$_4$ tetramer bound to 2 Cu(II)
Cu$_2$:MBPC1$_2$	MBPC1 dimer bound to 2 Cu(II)
Cu$_2$:RIDC1$_2$	RIDC1 dimer bound to 2 Cu(II)
Cu$_4$:^{C96}RIDC1$_4$	^{C96}RIDC1$_4$ tetramer bound to 4 Cu(II)
Cys	Cysteine
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol

E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EPR	Electron Paramagnetic Resonance
Fo-Fc	Difference electron density map
GdnHCl	Guanidine Hydrochloride
Glu	Glutamate
GSH	Reduced glutathione
GSSG	Oxidized glutathione dimer
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
His	Histidine
i1	interface 1
i2	interface 2
i3	interface 3
ICP-OES	Inductively Coupled Plasma Optical Emission Spectroscopy
Kd	Dissociation constant
LB	Luria Bertani
MBPC1	Metal Binding Protomer Cytochrome -1
MCBP	Multi Copper Binding Proteins
MeTIR	Metal Templated Interface Redesign
MOPS	3-morpholinopropane-1-sulfonic acid
NCS	Non Crystallographic Symmetry
PDI	Protein Disulfide Isomerase
PPI	Protein Protein Interaction

RIDC1	Rosetta Interface Redesigned Cytochrome-1
r.m.s.d.	Root Mean Square Deviation
SDS	Sodium Dodecyl Sulfate
SDS-PAGE	SDS Polyacrylamide Gel Electrophoresis
Thr	Threonine
Trp	Tryptophan
TMBZ	3,3'-5,5'-tetramethylbenzidine
TRIS	2-Amino-2-Hydroxymethyl-propane-1,3-diol
UV-Vis	Ultraviolet Visible

List of Figures

Figure 1.1. Natural metalloenzymes and the chemical reactions which they catalyze 3

Figure 1.2. Different Zinc-binding proteins and their zinc coordination motif 5

Figure 1.3. The zinc-binding site of MMP9 6

Figure 1.4. MMP9's role in inflammation and the response to tissue injury 7

Figure 1.5. Examples of engineered metalloproteins using a top-down approach 9

Figure 1.6. Designed non-covalent and hydrophobic interactions 11

Figure 1.7. Summary of the MeTIR approach 12

Figure 2.1. Metal Templated Interface Redesign (MeTIR) of a stable tetrameric protein assembly that can accommodate Zn(II) 24

Figure 2.2. Overlay of tetramers of Zn_4:MBPC1$_4$, Zn_4:RIDC1$_4$, and Zn_4:^{C96}RIDC1$_4$ 26

Figure 2.3. Quantification of product formation after incubation of samples with or without Zn(II) after 16 hours 30

Figure 2.4. Cartoons of different oligomeric species that can be formed after adding EDTA or Zn(II) $^{C38/C96}$MBPC1 or $^{C38/C96}$RIDC1.... 32

Figure 2.5. Sedimentation coefficient distribution plots for the tetrameric samples of $^{C38/C96}$MBPC1$_4$ and $^{C38/C96}$RIDC1$_4$ 34

Figure 2.6. Crystal structures of Zn_4:$^{C38/C96}$MBPC1$_4$ and Zn_4:$^{C38/C96}$RIDC1$_4$ 36

Figure 2.7. Chemical denaturation titrations of tetramer samples 38

Figure 2.8. Competitive Zn binding by Fura-2 and $^{C38/C96}$RIDC1$_4$ monitored by changes in Fura-2 fluorescence excitation spectrum 41

Figure 2.9. Competitive Zn(II) binding of MagFura-2 and $^{C38/C96}MBPC1_4$ monitored by changes in the MagFura-2 fluorescence excitation spectrum 42

Figure 2.10. Competitive Cu(II) binding by Fura-2 and $^{C38/C96}RIDC1_4$ monitored by changes in Fura-2 fluorescence excitation spectrum 44

Figure 2.11. X-Band EPR spectrum of $^{C38/C96}RIDC1_4$ incubated with Cu(II) 46

Figure 3.1. Structural features of cyt cb_{562} central to the design of $^{C81/C96}$RIDC1 68

Figure 3.2. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 at room temperature 71

Figure 3.3. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 at 35 °C 72

Figure 3.4. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 in the presence of Zn(II) at room temperature 73

Figure 3.5. Proposed scheme for the disulfide-mediate self-assembly of $^{C81/C96}$RIDC1 74

Figure 3.6. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 in the presence of 1:19 GSH:GSSG 76

Figure 3.7. Bulk purification of $^{C81/C96}RIDC1_4$ 80

Figure 3.8. Sedimentation velocity profile of $^{C81/C96}RIDC1_4$ in the presence and absence of equimolar Zn(II) 81

Figure 3.9. Crystal structures of apo, and Zn-bound $^{C81/C96}RIDC1_4$ 82

Figure 3.10. Backbone superposition of the cryptand-like cores of apo- (cyan) and Zn-bound (dark gray) $^{C81/C96}RIDC1_4$ 83

Figure 3.11. The central cavity in Zn_4:$^{C81/C96}RIDC1_4$ 84

Figure 3.12. Competitive Zn binding by Fura-2 and $^{C81/C96}RIDC1_4$ monitored by changes in Fura-2 fluorescence excitation spectrum 85

Figure 3.13. Zn-binding isotherm of $^{C81/C96}RIDC1_4$ determined using Fura-2 as competing ligand 86

Figure 3.14. Structural features of Zn_4: $^{A74/C81/C96}RIDC1_4$ 87

Figure 3.15. Competitive Zn binding by Mag-Fura-2 and $^{A74/C81/C96}RIDC1_4$ monitored by changes in Mag-Fura-2 fluorescence excitation spectrum 88

Figure 3.16. Overlay of the cryptand-like cores of Zn_4:$^{A74/C81/C96}RIDC1_4$ and Zn_4:$^{C81/C96}RIDC1_4$ crystal structures 89

Figure 3.17. *In vivo* disulfide-mediated self-assembly of $^{C81/C96}RIDC1$ and its Zn binding properties 91

Figure 3.18. *In vivo* disulfide-mediated self-assembly of $^{C81/C96}RIDC1$ and various control variants characterized by SDS-PAGE 92

Figure 3.19. UV-vis spectrum of the periplasmic extract of $^{C81/C96}RIDC1$-overexpressing cells 93

Figure 3.20. MALDI-TOF spectrum of $^{C81/C96}RIDC1_4$ extracted from the periplasm 94

Figure 4.1. (a) Bar graph containing the number of structures of copper complexes (b) Classification of copper Types 1, 2, and 3 117

Figure 4.2. Characterization of Cu_2:$MBPC1_2$ 119

Figure 4.3. Characterization of Cu_2:$RIDC1_2$ 121

Figure 4.4. Crystal structures of $^{C96}RIDC1_4$ crystallized in the presence of Cu(II) or Cu(I) 124

Figure 4.5. Crystal structure of dimer of Cu_2:$MBPC1_2$ is shown in cyan overlayed on a half of Cu_4:$^{C96}RIDC1_4$ 126

Figure 4.6. Crystal structure of tetramers of Cu_2:$^{C96}RIDC1_4$ shown in overlayed with the $Cu(I)_6$:$^{C96}RIDC1_4$ 128

Figure 4.7. EPR characterization and simulations of $^{C96}RIDC1_4$ and $^{A81/C96}RIDC1_4$ in the presence of Cu(II) 130

Figure 4.8. Crystal structure of Cu_4:(^{C96}MBPC1_2)$_2$ and SV distribution of samples with and without Cu(II) 133

Figure 4.9. Competition assay between Fura-2 and $^{C81/C96}$RIDC$_4$ Cu(II) 135

Figure 4.10. Crystal structures and EPR spectrum of $^{C81/C96}$RIDC1_4 in the presence of Cu(II) 138

Figure 4.11. Crystal structures and EPR spectrum of $^{A59/C81/C96}$RIDC1_4 in the presence of Cu(II) 140

Figure 5.1. Schematic Presentation of domain organization of the Multi Copper Binding Proteins 157

Figure 5.2. Crystal structure of Zn_4:$^{A59/A77/C81/C96}$RIDC1_4 159

List of Tables

Table 2.1. Percentage of each oligomeric species that was quantified fromthe templating reactions of $^{C38/C96}$MBPC1 and $^{C38/C96}$RIDC1in the absence and presence of Zn(II) 48

Table 2.2. Percentage of each oligomeric species that was quantified from the templating reactions of $^{A74/C38/C96}$MBPC1 and $^{74/C38/C96}$RIDC1 in the absence and presence of Zn(II) 48

Table 2.3. Midpoints of chemical denaturation titrations of $^{C38/C96}$MBPC1$_4$, ^{C96}RIDC1$_4$ and $^{C38/C96}$RIDC1$_4$ tetramers in the absence and presence of Zn(II) 49

Table 2.4. Dissociation constants for $^{C38/C96}$RIDC1$_4$ binding to Zn(II) determined by a competition assay with Fura-2 and fit to a four-Zn equilibrium binding model 49

Table 2.5. Equivalents per tetramer of either Cu(II), Ni(II), and Co(II) that were quantified to bind to $^{C38/C96}$RIDC1$_4$ 49

Table 2.6. Dissociation constants for $^{C38/C96}$RIDC1$_4$ binding to Cu(II) determined by a competition assay with Fura-2 50

Table 2.7. Crystallization conditions 50

Table 2.8. Data collection and refinement statistics 51

Table 3.1. Primers for site-directed mutagenesis 98

Table 3.2. Crystallization conditions of mutant variants at 25 ˚C 98

Table 3.3. X-Ray data collection and refinement statistics 99

Table 4.1. Dissociation constants for $^{C81/C96}$RIDC1$_4$ and Cu(II) measured by competition assay with Fura-2 137

Table 4.2. X-Ray data collection and refinement statistics for ^{C96}RIDC1$_4$ with 6 Cu(I) 98 143

Table 4.3. X-Ray data collection and refinement statistics for ^{C96}MBPC1$_4$ with 4 Cu(II) 144

Table 4.4. X-Ray data collection and refinement statistics for $^{C96/C81}RIDC1_4$ with 2 Cu(II) 145

Table 4.5. X-Ray data collection and refinement statistics for $^{C96/C81}RIDC1_4$ with 4 Cu(II) 146

Table 4.6. X-Ray data collection and refinement statistics for $^{A59/C81/C96}RIDC1_4$ with 3 Cu(II) 147

Table 4.7. Crystallization conditions 148

Table 5.1. X-Ray data collection and refinement statistics for Zn_4:$^{A59/A74/C81/C96}RIDC1_4$ 160

Acknowledgements

First and foremost I want to thank my family. In particular, my mother Aida Morales who has always supported me and whose yearly visits to San Diego were filled the warmth of the island. My brother and sister, Luis and Rubí Medina, have also always given me their support and friendship. My father Angel Medina has always backed me up in my decision to pursue postgraduate studies and always challenges me intellectually. My boyfriend Octavio Mondragón has accompanied me during five of the six years of my graduate studies, and he is constantly encouraging me to be the best scientist I can be and to give my best work forward.

I would also like to thank my labmates who made working in the lab very much like being in a small family: including all the arguments, dysfunctions, and enjoyable moments. Particularly Dustin Huard has been my lab-partner-in-crime during many endeavors and kept me good company during failed experiments.

Finally I need to thank my advisor Akif Tezcan for all of his assistance, especially for his motivation and contagious passion for science. He has always been kind, helpful and understanding. I admire his discipline, respect for family, and work ethic. I aspire to become as excellent a writer as he is. I also would like to thank my committee members for their advice.

The dissertation was funded by both NSF-AGEP, and the Cellular and Molecular Training Grant (CMG-NIH).

Chapter 2 is currently being prepared for publication of the material Medina-Morales, A. M., Brodin, J. D., Tezcan, F. A. The dissertation author was the primary author was the primary investigator and author of this paper.

Chapter 3 was reproduced in part with permission from Medina-Morales, A. M., Perez, A., Brodin, J. D., Tezcan, F. A. **2013**. In Vitro and Cellular Self-Assembly of a Protein Cryptand via Templated Disulfide Bonds. *J. Am. Chem. Soc.* The dissertation author was the primary investigator and author of this paper.

Vita

Education

2007	B.S. in Chemistry, University of Puerto Rico, Rio Piedras
2009	M.S. in Chemistry, University of California, San Diego
2013	Ph.D. in Chemistry, University of California, San Diego

Awards and Honors

January 2012	Carl Storm Underrepresented Minority (CSURM) Fellowship
June 2009-June2011	Cellular and Molecular Genetics Training Grant (NIH)
September 2008-June 2009	San Diego Fellowship for Graduate Studies
Summer 2007	Alliance for Graduate Education and the Professoriate (AGEP-NSF) Fellowship

Publications

In vitro and in vivo self-assembly of a protein cryptand via templated disulfide bonds. A Medina-Morales, A Perez, JD Brodin, FA Tezcan. (2013) J. Am. Chem. Soc.

Metal-Templated Assembly of Covalently Linked Protein Macrocycles. A Medina-Morales, JD Brodin, FA Tezcan (in preparation).

Evolution of metal selectivity in templated protein interfaces. JD Brodin, A Medina-Morales, T Ni, EN Salgado, XI Ambroggio, FA Tezcan. J. Am. Chem. Soc., 2010, 132 (25), pp 8610-8617.

Covalent modification of a metal-organic framework with isocyanates: probing substrate scope and reactivity. E Dugan, Z Wang, M Okamura, A Medina-Morales, SM Cohen. Chem. Comm. 2008, 3366.

ABSTRACT OF THE DISSERTATION

In Vitro and In Vivo Assembly of Supramolecular Metalloprotein Scaffolds via Templated Disulfide Bonds

by

Annette Marie Medina

Doctor of Philosophy in Chemistry

University of California, San Diego, 2013

Professor Faik Akif Tezcan, Chair

Metalloproteins are some of the best catalysts in Nature. The synergistic combination of transition metals and the appropriate protein scaffold creates catalysts that can accomplish even the most difficult chemical transformations with unmatched efficiency. Several approaches have been applied to the *de novo* design of metalloenzymes. In this dissertation we present a strategy where we combine lessons from supramolecular chemistry, and protein design to engineer new macrocyclic metalloprotein complexes. The primary goal of this dissertation project was to devise a strategy for developing stable protein scaffolds that are pre-organized for binding different transition metals, a highly challenging goal in protein engineering. This goal was achieved by the refinement of a rational design technique that we have termed Metal Templated Interface Redesign (MeTIR). In a stepwise strategy, monomeric protein building block (cytochrome cb_{562}) was first engineered to present metal-

binding motifs on its surface, then modified appropriately with self-complementary hydrophobic residues, and finally, with multiple, strategically placed cysteines on its surface. It was found that through structural templating either by zinc(II) coordination or by self-complementary hydrophobic interactions, the surface cysteines could correctly and selectively form disulfide bonds to stabilize multiple interfaces during self-assembly, resulting in the formation of a new class of macrocyclic, supramolecular protein assemblies. These assemblies were found to stably scaffold various divalent transition metal ions in their protected interiors. In particular, ^{C81C96}RIDC1$_4$ was found to correctly self-assemble in vivo in the periplasm of bacterial cells and successfully compete for Zn(II) binding. By combining principles of inorganic and supramolecular chemistry with those of protein design, this research opens up new avenues for constructing novel supramolecular or bioinorganic scaffolds in living systems that can be subjected to selective pressure for the evolution of new and improved biological functions.

Chapter 1

Introduction

From the perspective of a Bioinorganic Chemist, the periodic table can be visualized as an intricate map that can guide us through the diversity of nature. Very fitting is the fact that transition metals are located at the center of the periodic table, since their roles are literally central to our studies. Within nature, metals are present in as many colors as functions.[1-3] Notably, the heme iron in hemoglobin is responsible for the deep red color of life-sustaining blood.[4] We know that metals are crucial for the execution of complex chemical reactions such as photosynthesis[5,6] and respiration,[7] and that biological systems rely on metals to perform a wide variety of tasks.[8,9] Recent bioinformatics studies have identified that as many as 50% of the proteins of some organisms are associated with metals.[10,11]

Metalloproteins are some of Nature's best catalysts and perform chemical reactions that chemists are actively attempting to replicate with high efficiency.[12] However, most of the reaction schemes developed by chemists require the use of organic solvents, anaerobic environments, and high amounts of pressure and temperature to produce relatively low yields of product (compared to the efficiency of the natural enzyme).[12] These chemical reactions are not trivial, and as shown in Figure 1.1, include the hydration of carbon dioxide, the production of methanol from methane gas, the functional and specific oxidation of substrates, the challenging conversion of nitrogen into ammonia, and the production of hydrogen gas.[13-18] Our current economy relies heavily on some of these processes, like the Haber-Bosch process for the production of ammonia, in order to sustain the worldwide agricultural system.[19,20] The use of hydrogen gas a clean fuel would also require that the

production of hydrogen is available by alternative synthetic approaches.[21] Consequently, it is imperative for chemists to create artificial catalysts that can perform these reactions in a clean, inexpensive and efficient manner.

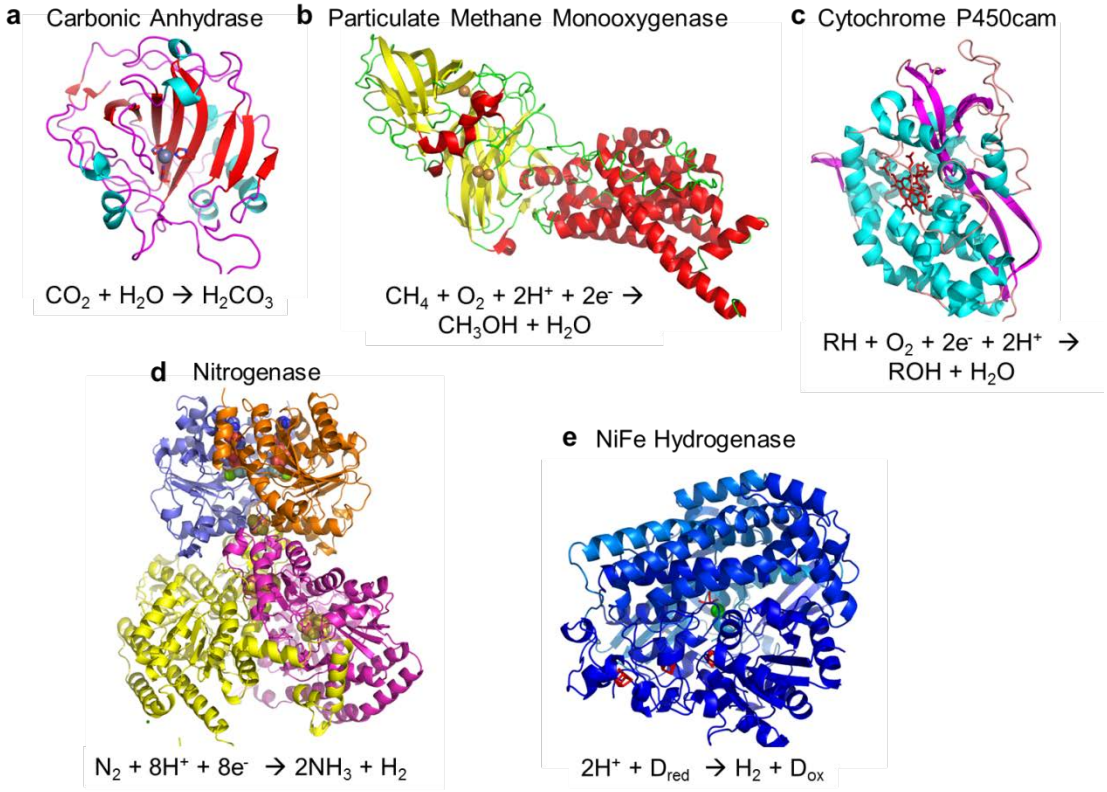

Figure 1.1. Natural metalloenzymes and the chemical reactions which they catalyze. The protein structures are shown as cartoons and the chemical reaction that is catalyzed is shown underneath. (**a**) The reduction of carbon dioxide into bicarbonate is catalyzed by a three-coordinate zinc center (gray sphere) located within carbonic anhydrase.[13] (**b**) Particulate Methane Monooxygenase (pMMO) catalyzes the oxidation of methane into methanol via an active site composed of a dinuclear copper center (orange spehres).[14] (**c**) The oxygenation of substrates by cytochrome P450 is mediated by a heme cofactor (red sticks).[15] (**d**) Nitrogenase catalyzes the reduction of nitrogen into ammonia by use of several metallic cofactors (shown as spheres).[22] (e) NiFe Hydrogenase mediates the reversible formation of hydrogen from protons by the use of a nickel cofactor (shown as green spheres).[23]

Our research aims to construct protein-based catalysts by designing protein scaffolds that can be functionalized with reactive metal centers to generate highly efficient enzymes with different chemical reactivities. Owing to the chemical diversity of amino acids, proteins are able to form hundreds of different structures. Nature has evolved metalloproteins because of the many advantages of the protein scaffold.[24-27] Mainly, since the environment around the reactive metal site can dictate the protein's reactivity, the metal's inner and outer sphere coordination can be fine-tuned by the careful selection of the coordinating amino acids with different chemical functionalities.[28,29,30-33] We can observe the variability of metalloprotein scaffolds and functionalities by describing the structures of the zinc proteins shown in Figure 1.2. These zinc proteins have evolved to possess very different structural conformations and functions, but all share the same tetrahedral zinc binding site. The amino acids that coordinate to the zinc site can be different in some of these proteins and this will influence the reactivity of the metal and chemical reaction performed. These proteins can have structural roles, like the zinc-finger protein (Figure 1.2a) which is responsible for DNA-binding,[34,35] or can perform chemical reactions such as lactamase hydrolysis by β-lactamases,[36] and the degradation of the extracellular matrix performed by matrix metalloproteinases.[37,38]

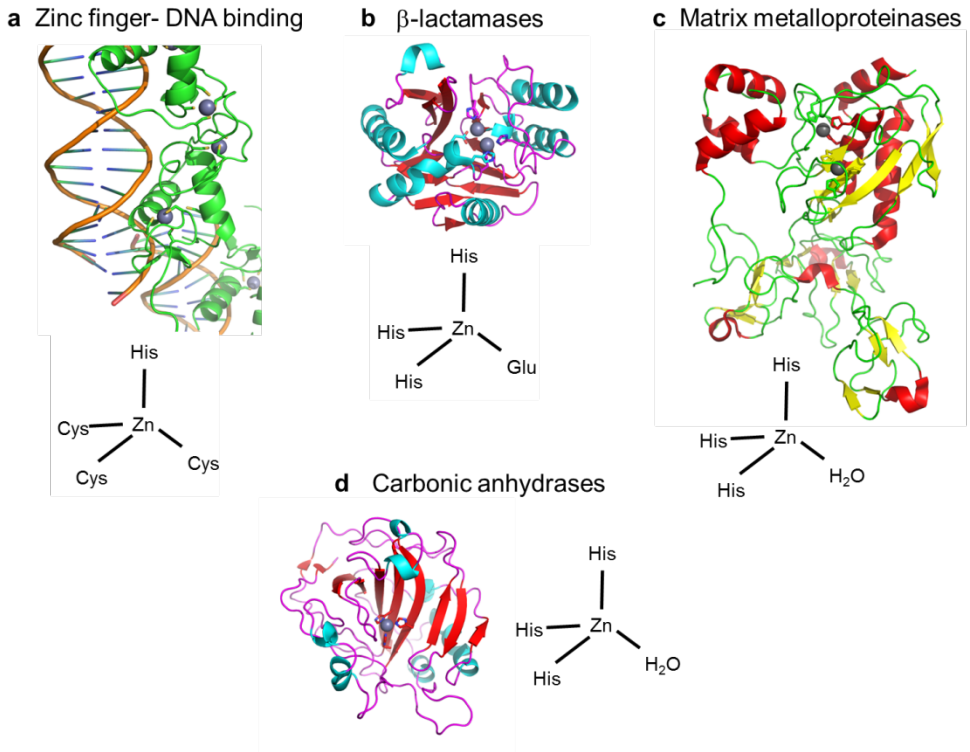

Figure 1.2. Different Zinc-binding proteins and their zinc coordination motifs. (**a**) The crystal structure of a Zinc finger is shown (green cartoon) bound to a DNA double helix. The zinc ions are coordinated via 3Cys and 1His residues.[35] (**b**) The crystal structure of a β-lactamase is shown as cartoon with the zinc active site highlighted as gray spheres coordinated by 3His and 1Glu.[36] The crystal structures of (**c**) Matrix metalloproteinase[38] and (**d**) carbonic anhydrase[39] share the same three-coordinate zinc binding motif. Their crystal structures are shown as cartoons and the zinc ions represented by gray spheres.

The influence of the protein scaffold on the reactivity of the metal active can be described by the example on Figure 1.3. Here is shown the zinc-binding site of the crystal structure of Matrix Metalloproteinase-9 (MMP9), which is a protein that is involved in the degradation of the extracellular matrix and is misregulated in some diseases.[37] This protein has evolved to possess a distinct fold composed of alpha helix (red), beta sheet structure (yellow) and random coil (green). If we explore the reactive site, we can observe that the zinc ion (gray sphere) is buried within the protein (Figure

1.3b). This shielding effect allows for the protection of the metal from the solvent and protects the cell from unwanted side reactions. In addition, the tuning of inner-sphere and outer-sphere interactions can be affected by altering the coordinating ligands by hydrophobic interactions, electrostatic interactions or hydrogen bonding effects. By further observation of the crystal structure, we can observe that it possesses substrate access channels built within the protein (Figure 1.3c). These likely evolved in order to enhance substrate specificity and selectivity. Finally, the protein has to be described within the context of the cell. Figure 1.4 shows MMP9 as part of a regulatory cascade that involves many other cellular components.[37] Since MMP9 can interact through protein-protein interactions it can transiently interact with cellular components and perform its native function.

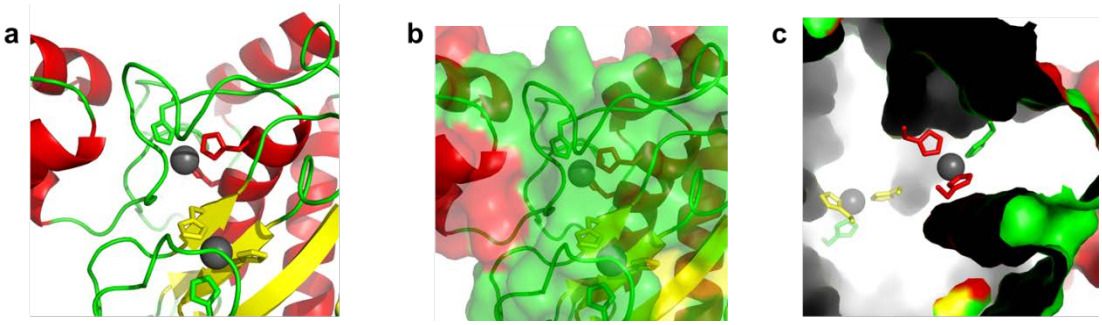

Figure 1.3. The zinc-binding site of MMP9. (**a**) The catalytic zinc ion is shown as a gray sphere bound to the corresponding 3His residues (shown as sticks). (**b**) Surface overlay of the crystal structure over the zinc site demonstrate that it is buried and protected from solvent. (**c**) The substrate access channels are shown near the catalytic site.

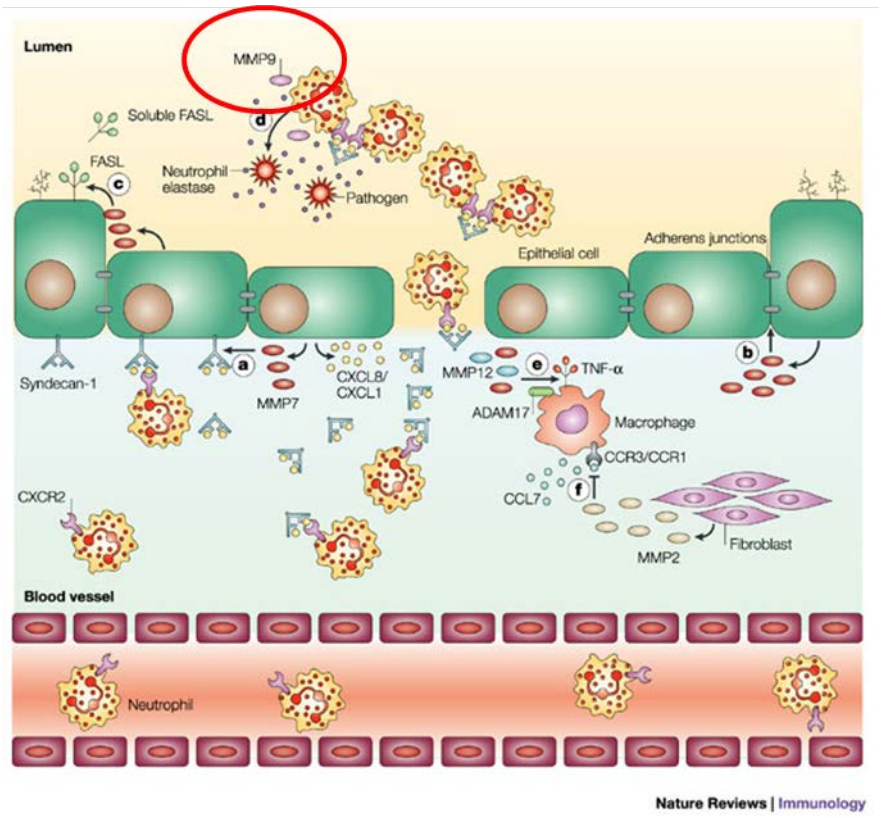

Figure 1.4. MMP9's role in inflammation and the response to tissue injury. MMP9 is shown as part of a programmed, coordinated response to repair damaged tissue and defend against infection (from reference 37).

A central goal of Bioinorganic Chemistry is to design proteins and artificially evolve them in order to learn about their functions, as well as to construct novel materials and technologies.[12,40-43] The field of protein design is greatly expanding due to the ability and ease of recombinant expression of protein mutants.[44,45] Also, the advent and improvement of computational methodologies makes protein engineering more widespread within the scientific community.[46,47] Within this field there are two primary methods: a top-down approach or a bottom-up approach. In Figure 1.5 are shown three examples of metalloprotein design through a top-down approach. In the

first example the authors took an existing metalloprotein and replaced the metal-binding cofactor (Figure 1.5a). Here the copper protein Azurin was redesigned by loop-directed mutagenesis and the metal coordination site is replaced with that of purple CuA.[48] Another approach was taken by the Dutton group where authors took an alpha-helical protein scaffold and inserted heme cofactors within it. The authors structurally characterized this heme maquette and engineered different axial ligands to the heme sites in order to influence its spectroscopic characteristics.[49] Finally, in Figure 1.5c is a similar approach taken by the Pecoraro group.[50] The authors constructed a three-coiled alpha-helical bundle that is held together stably by an internalized mercury ion that coordinates to engineered cysteine ligands. Within this structure, the authors engineered histidine amino acids that bind zinc. This three-coordinate zinc binding site was characterized as a hydrolytic enzyme.

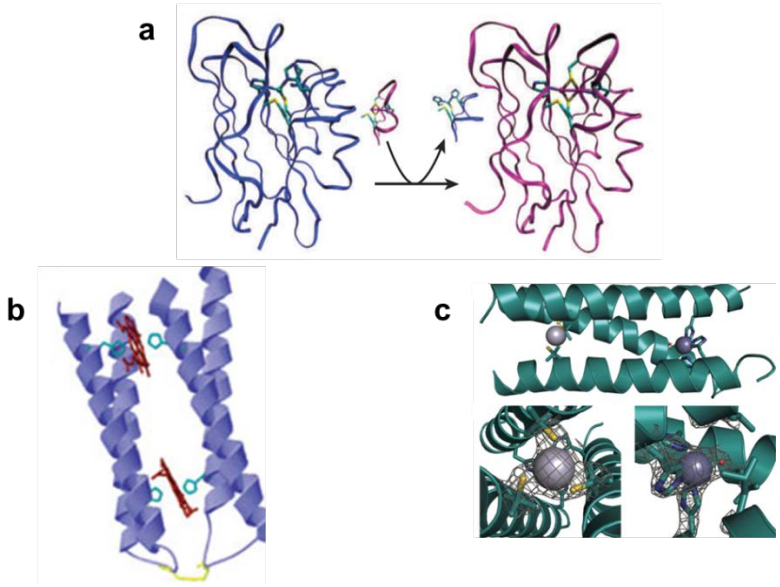

Figure 1.5. Examples of engineered metalloproteins using a top-down design approach. (**a**) The catalytic copper site of azurin was replaced by that of a purple CuA enzyme rendering it with the same spectroscopic characteristics of CuA.[48] (**b**) The cartoon of a designed of a heme maquette is shown in light blue with the heme cofactors shown as red sticks.[49] (**c**) A catalytic 3-coordinate zinc site was inserted within a mercury-mediated three-stranded coiled-coil peptide.[51]

One limitation of metalloprotein design is the difficulty in correctly modeling and predicting the coordination geometry of the metal site. This can be circumvented by evolving the protein around the metal site. In a bottom-up, or *de novo* approach, the protein scaffold is selected, and the metal binding functionality is introduced within it. To do this several computational protein design approaches can be utilized so that ideal protein-protein interactions are engineered. In addition, an increased versatility and variability is introduced by being able to select the starting protein building block. In our laboratory we take inspiration from Supramolecular Chemistry, particularly the field of metal-templated ligand design,[52-55] as well as from protein engineering, to design metalloproteins with metal sites at their interfaces.

The field of Supramolecular Chemistry has been inundated in the past three decades with research involving the design of small molecules that can self-assemble by employing non-covalent interactions and metal-assisted templating.[55-61] For example, metal-assisted self-assembly is a technique first introduced in the 1960's by Busch and co-workers (Figure 1.6a).[61] The formation of a tetradentate macrocycle was guided by the Ni(II) ion guiding the reactants to the correct positioning and coordination geometry to form the desired product (Figure 1.6b).[57,62] Another approach for the construction of self-assembled materials involves the use of hydrophobic interactions to drive the assembly of molecules. Protein chemists have applied these principles to engineer or stabilize proteins (Figure 1.6c).[26,63] On the left panel of Figure 1.6c is shown a model of an alpha-helical peptide that was designed to have hydrophobic interactions across one interface and upon aggregation it self-associates to form a four-helix bundle (right panel).[63]

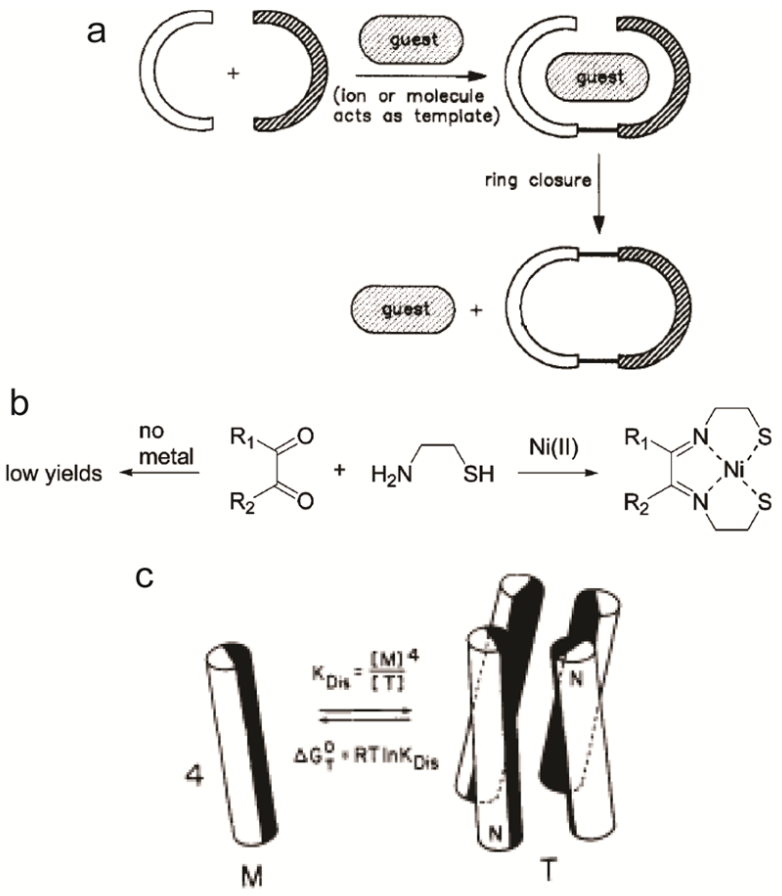

Figure 1.6. Designed non-covalent and hydrophobic interactions. (**a**) General approach for template synthesis where a guest molecule guides the assembly of ligands which crosslink to form a macrocycle whose structure is independent of the guest molecule binding (from reference 61). (**b**) Ni(II)-assisted templating of a macrocycle. In the absence of metal the diketone and sulfidoethanediol molecules do not react with one another but upon adding Ni(II) the molecules are guided to the appropriate coordination geometry and distances to react and form the tetradentate ligand (from reference 57). (**c**) Left: Monomeric peptide with a designed hydrophobic surface (shaded in black) which self-associates to form a four-helix bundle shown on the right (from reference 63).

By a combination of supramolecular design with protein engineering our laboratory has developed a technique that we termed Metal Templated Interface Redesign (MeTIR) (Figure 1.7).[64] MeTIR is a rational design strategy which starts with the engineering of metal-binding amino acids within a monomeric protein

(cytochrome cb_{562}).[65] Next, this protein will oligomerize after the addition of metal, and we take advantage of computational re-design methods to improve the protein-protein interactions across the oligomeric interfaces.[66] Finally, the introduction of covalent crosslinks across the protein interfaces allows us to stabilize the protein oligomer.[64] This final construct will be that of a stable protein oligomer with pre-formed metal-binding sites which we can be further engineered to bind different transition metals allowing us to go from an initial structural metal-binding site, to a catalytic metal-binding site.

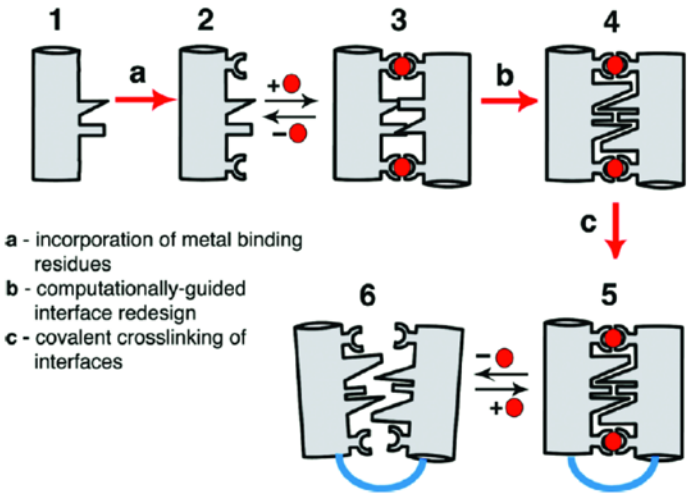

Figure 1.7. Summary of the MeTIR approach. A monomeric protein (1) is engineered to contain metal binding amino acids (2) and upon addition of metal (red spheres) it oligomerizes (3). After computationally-guided interface redesign (b) the protein can bind metal more selectively (4). Finally, we can covalently crosslink the interfaces of the protein by using disulfide bonds (5) and this protein oligomer stays intact independently of the presence of metal (6) (from reference 64).

Goals and Experimental Approach

The two main goals of my research dissertation are:

1) *To create a stable protein scaffold that is able to coordinate transition metals in a highly specific fashion and,*

2) *To enforce specific coordination geometries of metals within these new protein scaffolds.*

By utilizing the aforementioned MeTIR strategy, we combine the directionality and strength of metal coordination, with computational redesign of protein interfaces, to create stable multi-protein scaffolds. We demonstrated that we can engineer tetrameric protein constructs with hydrophobic interfaces, as well as with covalent linkages in the form of disulfide bonds.[64] Chapters 2 and 3 describe the design of additional disulfide bonds to the multi-protein assembly to create high affinity metal binding sites. In Chapter 2 we will demonstrate how to use a metal ion as a template to guide the correct formation of a protein tetramer which contains four interfacial disulfide bonds. The engineered $^{C38/C96}$MBPC1$_4$ and $^{C38/C96}$RIDC1$_4$ proteins form tetrameric macrocycles which, as shown through X-Ray crystallography, bind Zn(II) with the desired coordination geometry. We also establish that the $^{C38/C96}$RIDC1$_4$ construct bound to Zn(II) is much more stable than the previously designed species of ^{C96}RIDC1$_4$. Next, in Chapter 3, we show how to take advantage of designed hydrophobic interactions to guide an isostructural protein tetramer containing disulfide bonds at different interfaces. Particularly, $^{C81/C96}$RIDC1$_4$ was shown to self-assemble specifically and exclusively in the absence of metal and through engineered

hydrophobic interactions. We also demonstrate that we can selectively obtain the four-disulfide bonded tetrameric protein of $^{C81/C96}$RIDC1$_4$ in the periplasmic space of *E. coli*, a feat that has not been previously achieved by protein engineers. The construction of this specific disulfide-bonded tetramer in the periplasmic space of the cell opens the doors for performing cellular screening methods of a catalytic metalloenzyme *in vivo*.

Finally, we have focused on the design of redox active metal binding sites. We will show in Chapter 4 how we use a multi-protein assembly to create saturated and unsaturated Cu(II) or Cu(I) binding sites that have coordination geometries reminiscent of Type 1 and Type 2 copper enzymes. These copper binding sites have been characterized crystallographically to reveal the atomic coordinates of the copper-bound species, as well as spectroscopically, by measuring the Electron Paramagnetic Resonance (EPR) spectrum of Cu(II). After being able to use the pre-formed oligomeric scaffolds to bind copper, we aim to use this redox-capable protein construct to attain a catalytic enzyme.

References

1 Bertini, I., Gray, H. B., Stiefel, E. I. & Valentine, J. S. *Biological Inorganic Chemistry, Structure & Reactivity*. (University Science Books, 2007).

2 Holm, R. H., Kennepohl, P. & Solomon, E. I. Structural and functional aspects of metal sites in biology. *Chemical Reviews* **96**, 2239-2314 (1996).

3 Lu, Y., Berry, S. M. & Pfister, T. D. Engineering novel metalloproteins: Design of metal-binding sites into native protein scaffolds. *Chemical Reviews* **101**, 3047-3080 (2001).

4 Adair, G. The hemoglobin system VI. The oxygen dissociation curve of hemoglobin. *Journal of Biological Chemistry* **63**, 529-545 (1925).

5 Cammack, R., Rao, K. & Hall, D. Metalloproteins in the evolution of photosynthesis. *BioSystems* **14**, 57-80 (1981).

6 Wieghardt, K. The Active Sites in Manganese-Containing Metalloproteins and Inorganic Model Complexes. *Angewandte Chemie International Edition in English* **28**, 1153-1172 (1989).

7 Gray, H. B. & Malmstroem, B. G. Long-range electron transfer in multisite metalloproteins. *Biochemistry* **28**, 7499-7505 (1989).

8 Kennedy, M. L. & Gibney, B. R. Metalloprotein and redox protein design. *Current opinion in structural biology* **11**, 485-490 (2001).

9 Waldron, K. J. & Robinson, N. J. How do bacterial cells ensure that metalloproteins get the correct metal? *Nature Reviews Microbiology* **7**, 25-35 (2009).

10 Cvetkovic, A. *et al.* Microbial metalloproteomes are largely uncharacterized. *Nature* **466**, 779-782 (2010).

11 Waldron, K. J., Rutherford, J. C., Ford, D. & Robinson, N. J. Metalloproteins and metal sensing. *Nature* **460**, 823-830 (2009).

12 Que, L. & Tolman, W. B. Biologically inspired oxidation catalysis. *Nature* **455**, 333-340 (2008).

13 Khalifah, R. G. The carbon dioxide hydration activity of carbonic anhydrase I. Stop-flow kinetic studies on the native human isoenzymes B and C. *Journal of Biological Chemistry* **246**, 2561-2573 (1971).

14 Lieberman, R. L. & Rosenzweig, A. C. Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. *Nature* **434**, 177-182 (2005).

15 Poulos, T. L., Finzel, B. C. & Howard, A. J. High-resolution crystal structure of cytochrome P450cam. *Journal of molecular biology* **195**, 687-700 (1987).

16 Thorneley, R. N. F. & Lowe, D. J. in *Molybdenum Enzymes* (ed T. G. Spiro) 221-284 (John Wiley and Sons, 1985).

17 Armstrong, F. A. & Albracht, P. J. [NiFe]-hydrogenases: spectroscopic and electrochemical definition of reactions and intermediates. *Philosophical Transactions of the Royal Society a-Mathematical Physical and Engineering Sciences* **363**, 937-954 (2005).

18 Fontecilla-Camps, J. C., Volbeda, A., Cavazza, C. & Nicolet, Y. Structure/function relationships of NiFe - and FeFe -hydrogenases. *Chemical Reviews* **107**, 4273-4303, (2007).

19 Smil, V. Detonator of the population explosion. *Nature* **400**, 415-415 (1999).

20 Howard, J. B. & Rees, D. C. Structural basis of biological nitrogen fixation. *Chem. Rev.* **96**, 2965-2982 (1996).

21 Crabtree, G. W., Dresselhaus, M. S. & Buchanan, M. V. The hydrogen economy. *Physics Today* **57**, 39 (2004).

22 Tezcan, F. A. *et al.* Nitrogenase complexes: Multiple docking sites for a nucleotide switch protein. *Science* **309**, 1377-1380 (2005).

23 Volbeda, A. *et al.* Structure of the [NiFe] Hydrogenase Active Site: Evidence for Biologically Uncommon Fe Ligands⊥. *Journal of the American Chemical Society* **118**, 12989-12996 (1996).

24 Shoemaker, B. A. & Panchenko, A. R. Deciphering protein-protein interactions. Part II. Computational methods to predict protein and domain interaction partners. *Plos Computational Biology* **3**, 595-601 (2007).

25 Shoemaker, B. A. & Panchenko, A. R. Deciphering protein-protein interactions. Part I. Experimental techniques and databases. *Plos Computational Biology* **3**, 337-344 (2007).

26 Kortemme, T. *et al.* Computational redesign of protein-protein interaction specificity. *Nature Structural & Molecular Biology* **11**, 371-379 (2004).

27 Schulz, G. E. & Schirmer, R. H. *Principles of protein structure*. (Springer-Verlag KG., 1979).

28 Gray, H. B. & Winkler, J. R. Electron transfer in proteins. *Annual Review of Biochemistry* **65**, 537-561 (1996).

29 Vallee, B. L. & Williams, R. Metalloenzymes: the entatic nature of their active sites. *Proceedings of the National Academy of Sciences of the United States of America* **59**, 498 (1968).

30 Mauk, A. G. & Moore, G. R. Control of metalloprotein redox potentials: what does site-directed mutagenesis of hemoproteins tell us? *JBIC Journal of Biological Inorganic Chemistry* **2**, 119-125 (1997).

31 Shifman, J. M., Gibney, B. R., Sharp, R. E. & Dutton, P. L. Heme redox potential control in de novo designed four-α-helix bundle proteins. *Biochemistry* **39**, 14813-14821 (2000).

32 Tezcan, F. A., Winkler, J. R. & Gray, H. B. Effects of ligation and folding on reduction potentials of heme proteins. *Journal of the American Chemical Society* **120**, 13383-13388 (1998).

33 Marshall, N. M. *et al.* Rationally tuning the reduction potential of a single cupredoxin beyond the natural range. *Nature* **462**, 113-116 (2009).

34 Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A. & Wright, P. E. Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* **245**, 635-637 (1989).

35 Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252**, 809-817 (1991).

36 Ambler, R. The Structure of β-Lactamases. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences* **289**, 321-331 (1980).

37 Parks, W. C., Wilson, C. L. & López-Boado, Y. S. Matrix metalloproteinases as modulators of inflammation and innate immunity. *Nature Reviews Immunology* **4**, 617-629 (2004).

38 Rowsell, S. *et al.* Crystal structure of human MMP9 in complex with a reverse hydroxamate inhibitor. *Journal of molecular biology* **319**, 173-181 (2002).

39 Ippolito, J. A. *et al.* Structure of His94->Asp carbonic anhydrase II in a new crystalline form reveals a partially occupied zinc binding site. *Protein Engineering* **8**, 975-980 (1995).

40 Brodin, J. D. *et al.* Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **4**, 375-382, (2012).

41 Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855-862 (2009).

42 Matzapetakis, M. & Pecoraro, V. L. Site-selective metal binding by designed alpha-helical peptides. *Journal of the American Chemical Society* **127**, 18229-18233 (2005).

43 Lombardi, A. *et al.* Retrostructural analysis of metalloproteins: Application to the design of a minimal model for diiron proteins. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 6298-6305 (2000).

44 Estell, D. A., Graycar, T. P. & Wells, J. A. Engineering an enzyme by site-directed mutagenesis to be resistant to chemical oxidation. *Journal of Biological Chemistry* **260**, 6518-6521 (1985).

45 Muir, T. W., Sondhi, D. & Cole, P. A. Expressed protein ligation: a general method for protein engineering. *Proceedings of the National Academy of Sciences* **95**, 6705-6710 (1998).

46 Mandell, D. J. & Kortemme, T. Computer-aided design of functional protein interactions. *Nature Chemical Biology* **5**, 797-807 (2009).

47 Kitano, H. Computational systems biology. *Nature* **420**, 206-210 (2002).

48 Hay, M. T. *et al.* Spectroscopic characterization of an engineered purple CuA center in azurin. *Inorganic chemistry* **37**, 191-198 (1998).

49 Kalsbeck, W. A. *et al.* Structural and Electronic Properties of the Heme Cofactors in a Multi-Heme Synthetic Cytochrome. *Biochemistry* **35**, 3429-3438 (1996).

50 Zastrow, M. L., Peacock, F. A., Stuckey, J. A. & Pecoraro, V. L. Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nat Chem* **4**, 118-123, (2012).

52 Beves, J. E., Blight, B. A., Campbell, C. J., Leigh, D. A. & McBurney, R. T. Strategies and Tactics for the Metal-Directed Synthesis of Rotaxanes, Knots, Catenanes, and Higher Order Links. *Angewandte Chemie International Edition* **50**, 9260-9327, (2011).

53 Linton, B. & Hamilton, A. D. Formation of Artificial Receptors by Metal-Templated Self-Assembly. *Chemical Reviews* **97**, 1669-1680, (1997).

54 Sadownik, J. W. & Ulijn, R. V. Dynamic covalent chemistry in aid of peptide self-assembly. *Current Opinion in Biotechnology* **21**, 401-411, (2010).

55 McMurry, T. J., Raymond, K. N. & Smith, P. H. Molecular Recognition and Metal-Ion Template Synthesis. *Science* **244**, 938-943 (1989).

56 Anderson, S., Anderson, H. L. & Sanders, J. K. M. Expanding roles for templates in synthesis. *Accounts of Chemical Research* **26**, 469-475, (1993).

57 Thompson, M. C. & Busch, D. H. Reactions of Coördinated Ligands. II. Nickel (II) Complexes of Some Novel Tetradentate Ligands. *Journal of the American Chemical Society* **84**, 1762-1763 (1962).

58 Singh, R. & Whitesides, G. M. in *Sulphur-Containing Functional Groups* 633-658 (John Wiley & Sons, Inc., 1993).

59 Whitesides, G. M. & Grzybowski, B. Self-assembly at all scales. *Science* **295**, 2418-2421 (2002).

60 Fujita, M., Tominaga, M., Hori, A. & Therrien, B. Coordination assemblies from a Pd(II)-cornered square complex. *Accounts of Chemical Research* **38**, 369-378 (2005).

61 Hoss, R. & Vögtle, F. Template syntheses. *Angewandte Chemie International Edition in English* **33**, 375-384 (1994).

62 Thompson, M. C. & Busch, D. H. Reactions of Coordinated Ligands .9. Utilization of Template Hypothesis to Synthesize Macrocyclic Ligands in Situ. *Journal of the American Chemical Society* **86**, 3651-3656 (1964).

63 Eisenberg, D. *et al.* The design, synthesis, and crystallization of an alpha-helical peptide. *Proteins: Structure, Function, and Bioinformatics* **1**, 16-22 (1986).

64 Brodin, J. D. *et al.* Evolution of Metal Selectivity in Templated Protein Interfaces. *Journal of the American Chemical Society* **132**, 8610-8617, (2010).

65 Salgado, E. N., Faraone-Mennella, J. & Tezcan, F. A. Controlling Protein-Protein Interactions through Metal Coordination: Assembly of a 16-Helix Bundle Protein. *J. Am. Chem. Soc.* **129**, 13374-13375 (2007).

66 Salgado, E. N. *et al.* Metal-Templated Design of Protein Interfaces. *Proc. Natl. Acad. Sci. USA* **107**, 1827-1832 (2010).

Chapter 2

Zn(II)-Assisted Covalent Templating of Protein Macrocycles

Introduction

Nature confines catalytic metal active sites within protein scaffolds in order to enhance their reactivity. Metalloenzymes are able to fine-tune the catalytic efficiency of the reactive metal through features that have evolved over millions of years.[1,2] Some of these characteristics include shielding or de-shielding the metal from solvent,[3-5] rigid environments around the metal center,[6,7] and secondary sphere interactions that modulate the reduction potential of the metal active site.[8,9] Although much effort has been placed into the synthesis of biomimetic metal complexes that can be used as catalysts, it remains a main disadvantage that much of the chemistry is done with organic solvents, at very low temperatures, and with oxygen-free atmospheres.[10,11] This is in sharp contrast to metalloenzymes, which function in aqueous environments and under ambient conditions.

Our goal is to design rigid protein assemblies that contain internalized metal binding sites that can be utilized to obtain specific coordination geometries and reactive metal centers. The use of metals to direct the assembly of complex molecules has been well-exploited over the past five decades with the construction of a wide variety of two- and three-dimensional macrocyclic compounds.[12] By using transition metals to direct the specific arrangement of molecules, chemists can take advantage of the directionality and lability of metal-ligand bonds in order to obtain the most thermodynamically stable products.[12-16] Here, we establish the construction of macrocyclic metal-binding protein assemblies through the rational design of metal-

protein bonds, hydrophobic interactions, and disulfide crosslinks by applying the method of metal-templated ligand synthesis[15] to protein design.

We have previously described the construction of a metal binding protein termed Metal Binding Protein Cytochrome 1, or MBPC1, which upon binding Zn(II) assembles into a tetrameric oligomer that contains four tetrahedral Zn(II) binding sites (His63, Asp74, His73, and His77) (Figure 2.1a, 2.2b).[17] This tetrameric assembly can be described as possessing three interfaces (*i1*, *i2* and *i3*) which are highlighted in the tetramer overlay of Figure 2.2a. In order to further stabilize the tetrameric assembly we devised a strategy that we termed Metal Templated Interface Redesign (MeTIR) which incorporates the inclusion of hydrophobic and covalent interactions across the interface of MBPC1 in order to impart stability.[18] The resulting Rosetta Interface Designed Cytochrome, RIDC1, was designed utilizing the MeTIR strategy by the addition of mostly hydrophobic interactions across interface *i1* of the tetrameric oligomer (Figure 2.1b).[19] Upon addition of Zn(II) the Zn_4:$RIDC1_4$ tetramer was characterized crystallographically and in solution and was found to be isostructural to the parent Zn_4:$MBPC1_4$ assembly (see overlays in light gray and dark gray in Figures 2.2a, 2.2b).

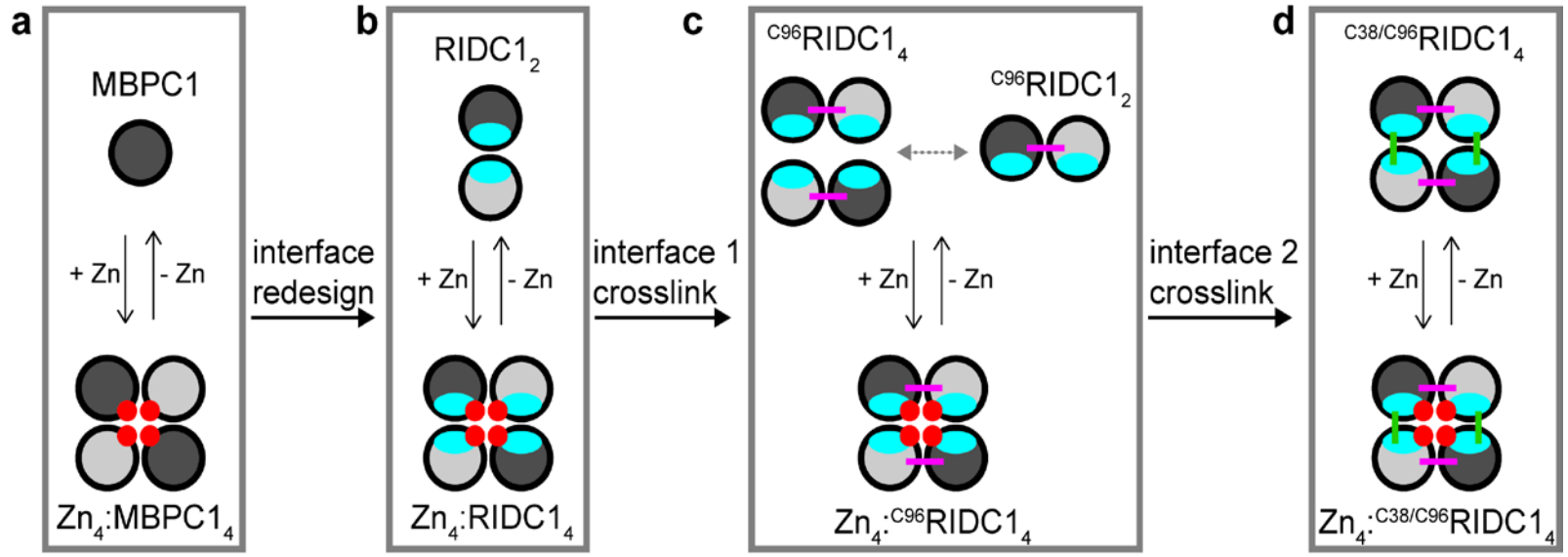

Figure 2.1. Metal Templated Interface Redesign (MeTIR) of a stable tetrameric protein assembly that can accommodate Zn(II). (**a**) Monomeric MBPC1 (gray circle, top) forms a tetramer (dark and gray circles, bottom) upon the addition of Zn(II) (red circles); (**b**) Hydrophobic interactions designed to stabilize interface *i1* are introduced and are highlighted in cyan; the metal-free $RIDC1_2$ (top) shown to form a tetramer upon the addition of Zn(II) (bottom); (**c**) Introduction of crosslink across interface *i2* in the form of a disulfide bond (magenta lines). The $^{C96}RIDC1_4$ metal-free tetramer is under equilibrium with its covalently bound $^{C96}RIDC1_2$ dimer (top); Upon Zn(II) binding a stable tetramer is also formed (bottom); (**d**) Interfacial crosslinks across interface *i1* are added and the four-disulfide bonded $^{C38/C96}RIDC1_4$ tetramers are shown in the absence and presence of Zn(II) (top and bottom respectively).

Since the hydrophobic mutations were designed across just one interface, RIDC1 has a high affinity for dimerization with itself across the interfacial mutations on *i1* (Figure 2.1b).[19] In order to stabilize the tetrameric species we proceeded to introduce a covalent linkage across interface *i2* in the form of a disulfide bond (Figure 2.1c). This was achieved after identifying a pair of symmetrically-related amino acids that were at optimal distances from each other for disulfide pairing. For this purpose was chosen residue 96.[18] The resulting ^{C96}RIDC1 forms a tetramer in the absence of metal and is able to accommodate Zn(II) with high affinity and specificity.[18] After characterizing ^{C96}RIDC1$_4$ in the metal-free and Zn(II)-bound states, it was identified that it can undergo a conformational rearrangement upon Zn(II) binding (Figure 2.1c). Not surprisingly, the metal-free form exists under equilibrium of a mixture of the ^{C96}RIDC1$_2$ dimer and a ^{C96}RIDC1$_4$ tetramer because of the fluidity of the designed hydrophobic interface *i1* (Figure 2.1c). The metal binding site is also rearranged upon Zn(II) binding.[18] Although this feature has the consequence of inducing flexibility on the tetrameric scaffold through the rearrangement of the hydrophobic interactions across interface *i1*, in order to accommodate a metal rigidly, it is undesirable for the metal binding amino acids to undergo such a large rearrangement (~16 Å).

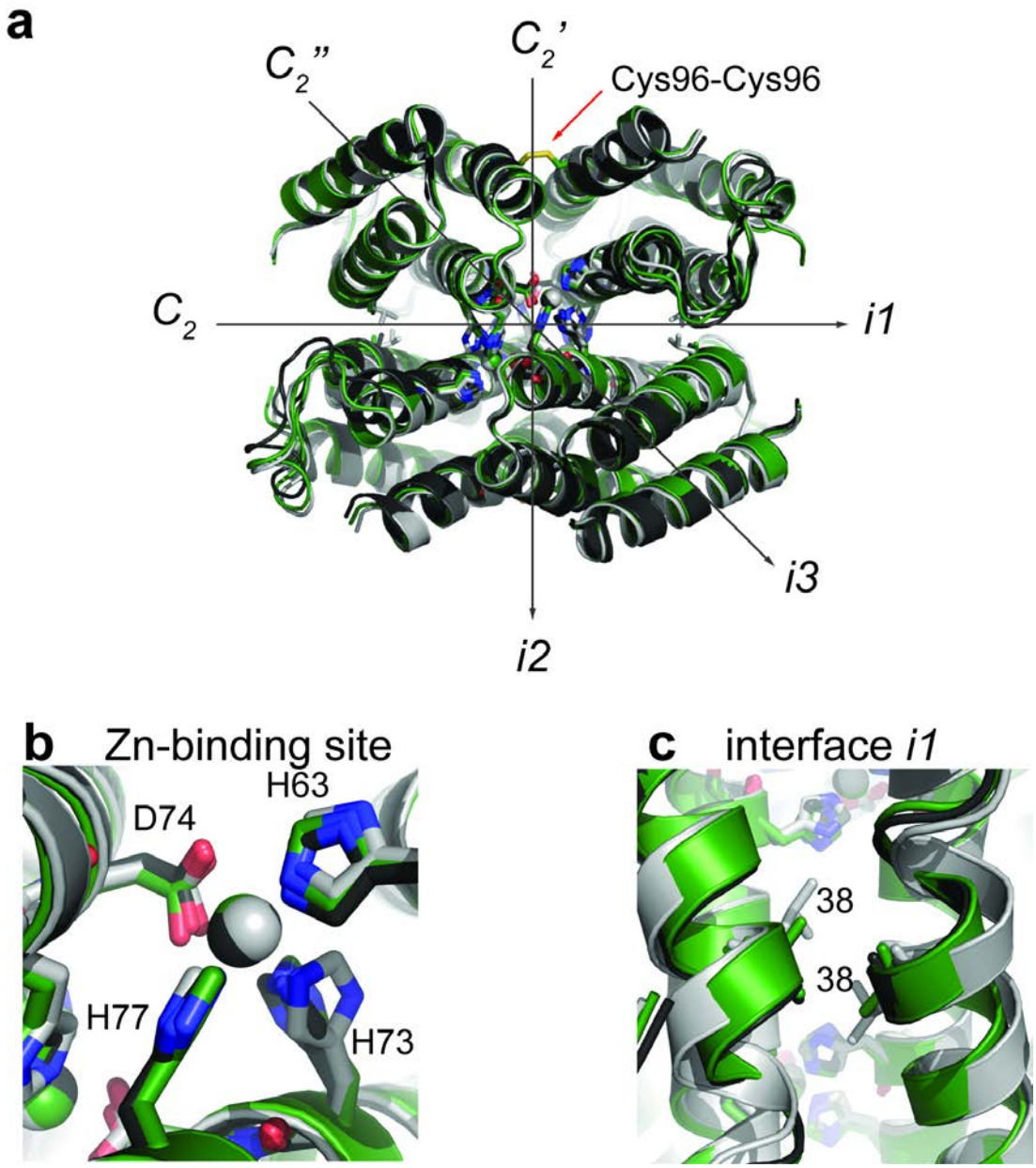

Figure 2.2. Overlay of (**a**) tetramers of Zn_4:MBPC1$_4$ (light gray), Zn_4:RIDC1$_4$ (dark gray), and Zn_4:^{C96}RIDC1$_4$ (green and Cys96-Cys96' disulfide bond shown as sticks and highlighted by red arrow); (**b**) Close-up view of the Zn(II)-binding sites (shown as sphere) and metal binding amino acids are shown as sticks (Asp74, His63, His73, His77); (**c**) View of the overlay of the three crystal structures' interface *i1*. Residue 38 is shown as sticks to highlight the proximity of the symmetric pair and that they are located at appropriate disulfide bonding distance.

To stabilize the protein tetramer further, here we demonstrate the introduction of additional covalent bonds across interface *i2* by introducing a cysteine at position 38 (Figure 2.2c). The final constructs, termed $^{C38/C96}$MBPC1$_4$ and $^{C38/C96}$RIDC1$_4$, are isostructural tetrameric protein assemblies (Figure 2.1d). Since we aim to use Zn(II) as the templating agent for the intended tetramer formation, we are also interested in studying the effect of the hydrophobic interactions across interface *i1*, which is why the "non-evolved" form $^{C38/C96}$MBPC1 was included in this study. Both new constructs contain four interfacial disulfide bonds, are able to accommodate Zn(II), and also possess high chemical stability ascribed to Zn(II) binding, covalent bonds and designed interfacial hydrophobic interactions. This is the first time the strategy of metal-templated macrocycle synthesis is applied to protein engineering.

Results and Discussion

Zn(II)-Assisted Self-Assembly.

Starting with the monomeric (disulfide reduced) forms of $^{C38/C96}$MBPC1 and $^{C38/C96}$RIDC1, we are interested in forming the desired tetramer with disulfide bonds across Cys96-Cys96' and Cys38-Cys38' after the addition of Zn(II). For this, we initially added excess DTT to the purified proteins in order to fully reduce the cysteines that might have become oxidized during the purification process. After removing the reducing agent by using a desalting column, we made 50 μM samples and added to these either one equivalent of Zn(II) or an excess of EDTA (1 mM) and incubated them at 35 ˚C for 16 hours. To determine the distribution of oligomeric species formed after incubation, aliquots of these samples were run on a non-reducing 12% denaturing SDS-PAGE gel and the amount of each oligomeric species was quantified by analysis of the density of the bands in each lane using ImageJ[20] (Figure 3.3a upper and lower panels, Table 2.1**)**. The bar graph on Figure 2.3a (lower panel) details the percentage of tetramer as a function of total product. There is an increase in the formation of tetramer upon addition of Zn(II). In the case of $^{C38/C96}$MBPC1, the amount of tetramer formed in the absence of Zn is only 10.4±1.3%; however, upon addition of one equivalent of metal the percentage of tetramer increases to 34.8±2.9%. Similarly, for $^{C38/C96}$RIDC1 in the absence of metal we measured a total of 20.7±0.8% of tetramer; however, for the samples containing one equivalent of Zn(II) this value doubles to 43.7±2.3%. From these data we show that the addition of Zn(II) to both $^{C38/C96}$MBPC1 and $^{C38/C96}$RIDC1, increases the formation of tetrameric product. Also,

we can take note that the formation of tetramer is higher for the "evolved" construct $^{C38/C96}$RIDC1 than for the parent $^{C38/C96}$MBPC1 both in the absence and presence of metal.

In order to demonstrate that Zn(II) binding drives the correct tetramer assembly through the aforementioned metal binding site (Figure 2.2b, Table 2.2**)**, we substituted the metal binding residue Asp74 for a non-chelating alanine in order to disrupt the metal binding event and consequently hinder the assembly of the tetramer. Indeed, upon addition of Zn(II) to $^{A74/C38/C96}$MBPC1 and $^{A74/C38/C96}$RIDC1 we observed a decrease in the formation of tetramer. Particularly, for $^{A74/C38/C96}$MBPC1 in the absence and presence of Zn(II) the percentages of tetrameric product are almost identical and are 11.4±0.7% and 14.6±0.3% respectively. In the "evolved" $^{A74/C38/C96}$RIDC1 there is also a decrease in the tetrameric product from the previously mentioned 43.7% in the presence of Zn(II) to only 12.0±1.2% in the absence of metal and of 22.6±1.7% in the presence of Zn(II). We can attribute the higher percentage of tetramer in the $^{A74/C38/C96}$RIDC1 construct as compared to the MBPC1 version to the hydrophobic interactions having an effect in the arrangement of the monomers to the appropriate conformation by allowing the formation of the dimer intermediate to form a tetramer.

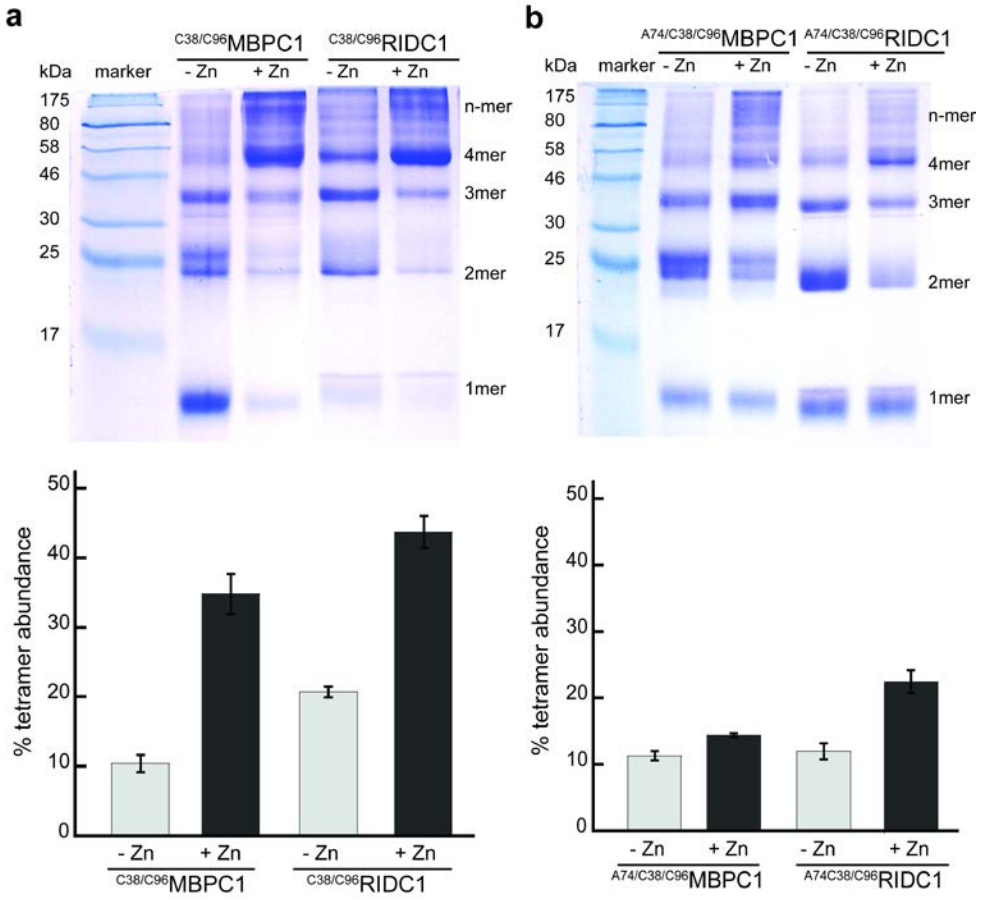

Figure 2.3. Quantification of product formation after incubation of samples with or without Zn(II) after 16 hours. (**a**) 12% non-reducing SDS-PAGE gels of the products of reactions of $^{C38/C96}$MBPC1 (lane 2 and 3) and $^{C38/C96}$RIDC1 (lane 4 and 5) in the presence and absence of Zn(II). Bottom panel is a bar graph showing the quantification of tetrameric product for each sample. (**b**) 12% non-reducing SDS-PAGE gels of reactions of $^{A74/C38/C96}$MBPC1 (lane 2 and 3) and $^{A74/C38/C96}$RIDC1 (lane 4 and 5) in the presence and absence of Zn(II). Bottom panel shows a bar graph with the tetrameric product for each sample. (See tables 2.1 and 2.2 for percentage values).

Given our results we propose the scheme shown in Figure 2.4 for the formation of the Zn(II)-driven tetramer. In the absence of templating metal for $^{C38/C96}$MBPC1 it is found that most of the sample forms monomers (33.5%) and dimers (26.9%). Consequently it is necessary for Zn(II) to be present to direct these lower molecular

weight species together to form the correct tetrameric species. Conversely, for metal-free $^{C38/C96}$RIDC1 the oligomeric species identified are of a small amount of monomer (only 7.2%), a large amount of dimer (27.1%) and a trimer species (21.7%). The dimeric species could be an intermediate that can form the correct tetramer, or it could also consist of the wrongly disulfide bonded dimer. Also, the trimeric species observed in the absence or presence of metal could consist of a kinetically trapped closed species, an incorrectly bonded open trimer or a correctly bonded trimeric intermediate to form the correct tetramer. Once metal is added to $^{C38/C96}$RIDC1 the amount of monomer and dimer decrease substantially (3.5 and 7.6% respectively), the trimer persists at a lower concentration (9.9%) and the tetramer is the major product (43.7%) followed by a large amount of high molecular weight species (35.2%). The large amount of high molecular weight species formed in the presence of Zn(II) for both $^{C38/C96}$MBPC1 and $^{C38/C96}$RIDC1 can be composed of large aggregates of proteins which can include correct Cys96-Cys96' and Cys38-Cys38' disulfide bonds and/or incorrect Cys38-Cys96 disulfide bonds which lead to large oligomers that are kinetically trapped since the covalent disulfide bond is not reversible under oxidizing conditions.

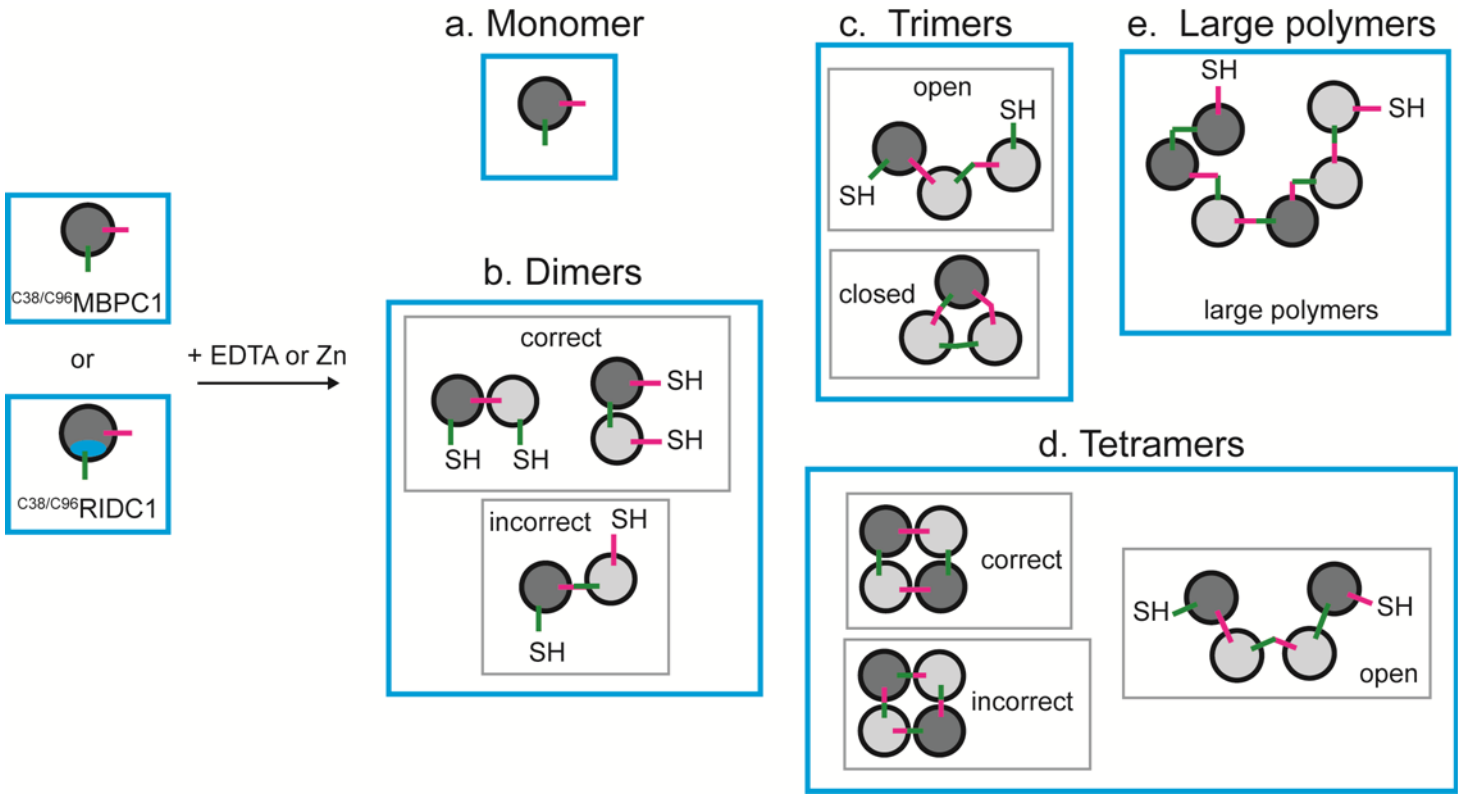

Figure 2.4. Cartoons of different oligomeric species that can be formed after adding EDTA or Zn(II) to $^{C38/C96}$MBPC1 or $^{C38/C96}$RIDC1. The red spheres indicating Zn and the cyan interface are omitted in the oligomeric scheme for clarity. Isolated into light blue boxes are the correct, incorrect, closed and open configurations that are possible to be formed for the different polymers. Starting with the monomeric $^{C38/C96}$MBPC1 or $^{C38/C96}$RIDC1 and upon the addition of Zn or a metal chelator (EDTA) different oligomeric species can form and these can include unreacted monomers (**a**) correct and incorrect dimers (**b**), open and closed trimers (**c**), correct or incorrect tetramers (**d**) and large polymers (**e**).

Structural and Biophysical Characterization of tetramer constructs.

In order to identify that the correct tetrameric products were in fact obtained, we sought to characterize the Zn(II)-templated tetrameric proteins with more detail. We proceeded to perform sedimentation velocity experiments on the Zn(II)-templated $^{C38/C96}$MBPC1$_4$ and $^{C38/C96}$RIDC1$_4$ tetrameric species in the presence of Zn(II) or absence of metal (after removing the templating zinc with EDTA). The sedimentation coefficient distributions for both $^{C38/C96}$MBPC1$_4$ and $^{C38/C96}$RIDC1$_4$ are shown in Figure 2.5a and b and their maxima occurs at 3.9 and 4.3 S in the absence and presence of Zn(II) respectively for both constructs. Since both of these tetramers are designed to be covalently linked through equivalent disulfide bonds, and since the difference in their sequence is only minimal (6 mutations), we expect them to be isostructural with each another and their sedimentation coefficients to be very similar to one another and this is what we can ascertain from their sedimentation coefficient values being identical in the absence and presence of metal. Importantly, the narrow distribution of tetrameric species for these in the absence and presence of Zn(II) leads us to assume that only one type of tetramer is isolated in solution. Upon addition of Zn(II), the sedimentation coefficients of each of the mutants increases by a small amount (0.3 S) which can be a consequence of an increase in the molecular weight of the tetramer because of Zn(II)-binding and/or because of a reconfiguration of the tetramer in order to accommodate metal with the preferred coordination geometry. This observation leads us to hypothesize that even though these tetramers are covalently crosslinked through the four interfaces, they are still slightly flexible and

that the disulfide bonds can allow small reconfigurations. These values are also similar to the previously studied ^{C96}RIDC1$_4$ tetramer which in the absence of metal has a sedimentation coefficient of 4.3 S and that increases to 4.5 S after the addition of Zn(II).[18]

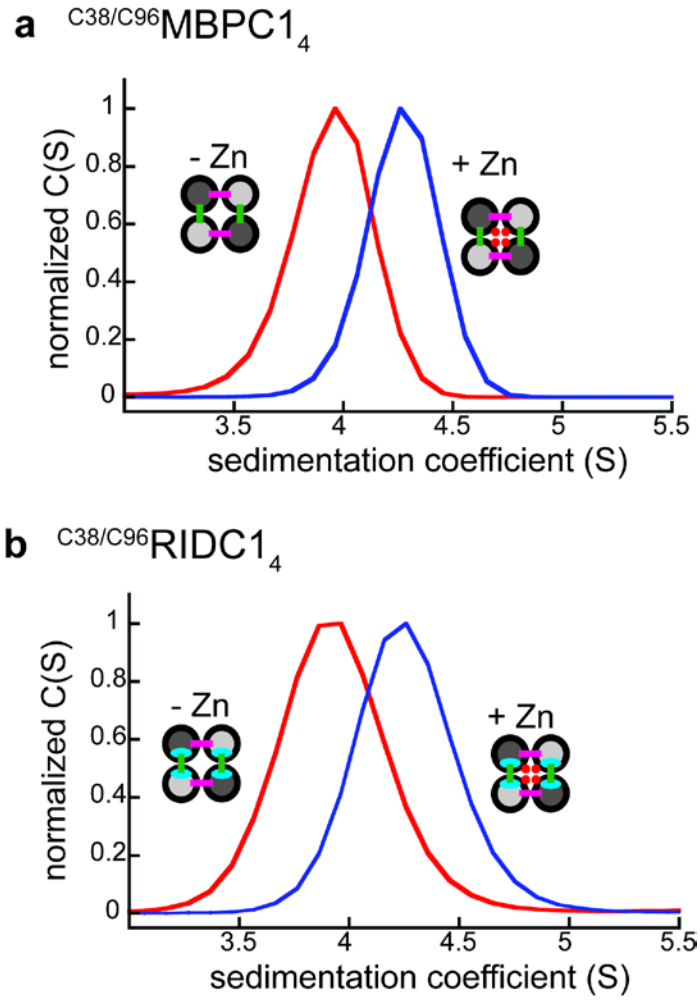

Figure 2.5. Sedimentation coefficient distribution plots for the tetrameric samples (1.25 µM) of (**a**) $^{C38/C96}$MBPC1$_4$ and (**b**) $^{C38/C96}$RIDC1$_4$ in the presence of excess of EDTA (red curves) and in the presence of 5 µM Zn(II) (blue curves) For both constructs the sedimentation coefficients increase from 3.9 S in the absence of metal to 4.3 S upon the addition of Zn(II).

Next, the crystal structures of the Zn(II)-bound tetramers of Zn_4:$^{C38/C96}$MBPC1$_4$ and Zn_4:$^{C38/C96}$RIDC1$_4$ were solved to 1.8 Å and 2.2 Å resolution respectively. The two structures are isostructural with one another with an r.m.s.d. of 0.481 Å over α-carbon. In addition, Zn_4:$^{C38/C96}$MBPC1$_4$ and Zn_4:$^{C38/C96}$RIDC1$_4$ are also isostructural to the parent Zn_4:^{C96}RIDC1$_4$ with r.m.s.d. values of 1.294 Å and 0.747 Å respectively . The four monomers are antiparallel to each other as expected, and the four Zn(II) binding sites all feature tetrahedral coordination geometries with the preserved ligand motif of His63, His77, His73, and Asp74. The crystal structures reveal that the disulfide bonds are formed as intended through the Cys38-Cys38' and Cys96-Cys96' crosslinks as shown by the omit F_o-F_c electron density maps (Figure 2.6 lower panels). We can infer from these crystal structures and from the narrow distribution of tetrameric species in the sedimentation coefficient distribution experiments, that the major tetrameric product obtained after the addition of Zn(II) corresponds to the designed, closed tetramer and not to an open tetramer.

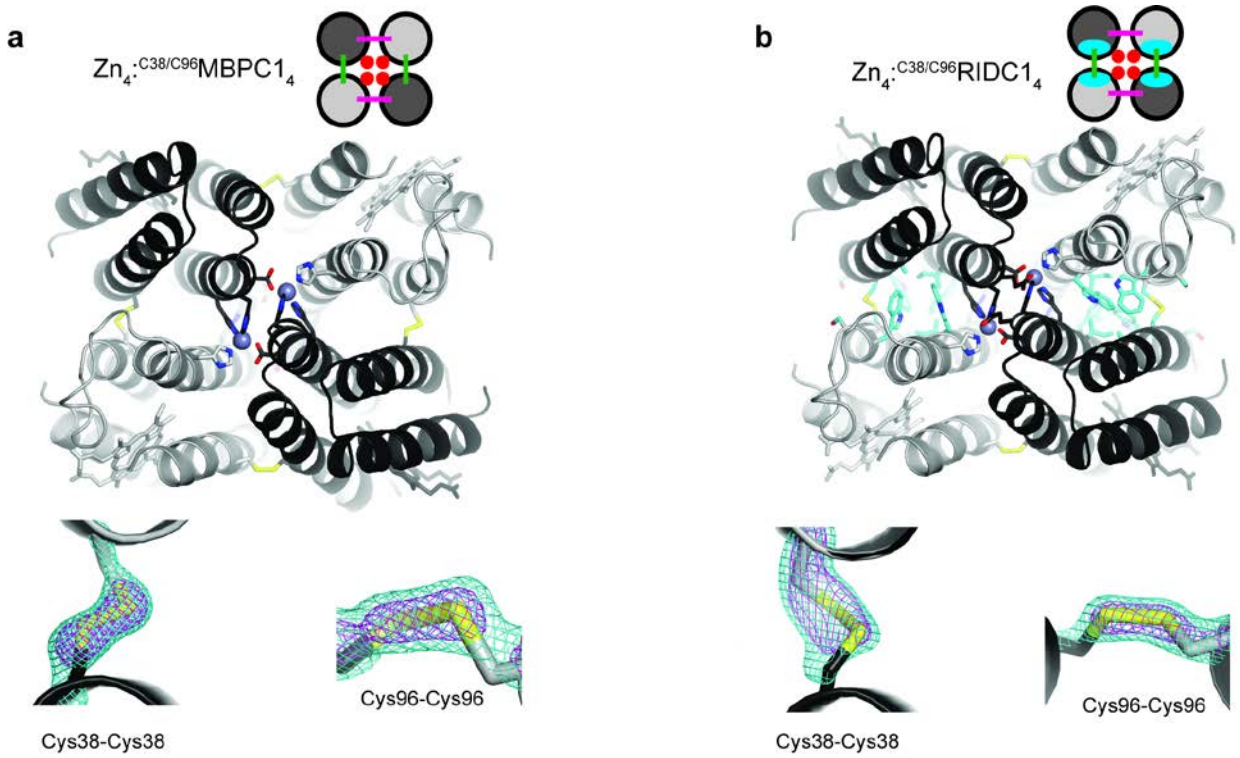

Figure 2.6. Crystal structures of (**a**) Zn_4:$^{C38/C96}$MBPC1_4 (solved at 1.8 Å) and (**b**) Zn_4:$^{C38/C96}$RIDC1_4 (solved at 2.2 Å). The structures are light gray and dark gray cartoons with the Zn(II) atoms as gray spheres and the metal-binding amino acids shown as sticks. The hydrophobic interactions of Zn_4:$^{C38/C96}$RIDC1_4 (**b**) are shown as cyan sticks. For (**a**) the lower panel highlights the disulfide crosslinks Cys38-Cys38 and Cys96-Cys96 with overlayed omit F_o-F_c electron density maps (C38-C38 are 5.5σ cyan and 7σ cyan; C96-C96 shown at 4σ cyan, 5.5 σ magenta). Disulfide crosslinks of (**b**) are also shown as sticks and are overlayed with omit electron density maps F_o-F_c (C38-C38 3σ cyan, 4σ magenta; C96-C96 5σ cyan, 6.5σ magenta). Both crystal structures have high homology to one another with r.m.s.d. of 0.481 Å over α-carbon.

Chemical stability and Zn(II) affinity of disulfide bonded tetramers.

To assess the effects that the additional disulfide crosslinks have on the stabilities of the pre-templated protein scaffolds, we proceeded to perform chemical denaturation titrations and measured the unfolding of the protein as a function of GdnHCl concentration (Figure 2.7, Table 3). These experiments were also performed on the two-disulfide crosslinked tetramer $^{C96}RIDC1_4$ in order to compare the effect of the additional disulfide bonds on the stabilities of the four-disulfide tetramers $^{C38/C96}MBPC1_4$ and $^{C38/C96}RIDC1_4$. It has been demonstrated that protein stability can be improved by the incorporation of disulfide bonds into a native scaffold.[21-25] In the absence of Zn(II) both $^{C38/C96}MBPC1_4$ and $^{C38/C96}RIDC1_4$ have $[GdnHCl]_{1/2}$ midpoints at 2.0 M, which contrasts sharply to the lower value of 1.0 M calculated for $^{C96}RIDC1_4$ in the absence of metal. Upon adding Zn(II) to the samples the $[GdnHCl]_{1/2}$ denaturation midpoints increased for all species. $^{C38/C96}MBPC1_4$ and $^{C38/C96}RIDC1_4$ had their midpoints augment to 2.1 and 3.1 M respectively, and the denaturation midpoint of $^{C96}RIDC1_4$ increased to 2.1 M after the addition of metal. The four-disulfide crosslinked tetramers are more stable in the absence of metal than the two-disulfide $^{C96}RIDC1_4$. This can be attributed to the fluidity of the hydrophobic interface across *i1* for $^{C96}RIDC1_4$ and the tetramer being able to dissociate into $^{C96}RIDC1_2$ dimers because of interruption of the hydrophobic interactions by the chemical denaturant.

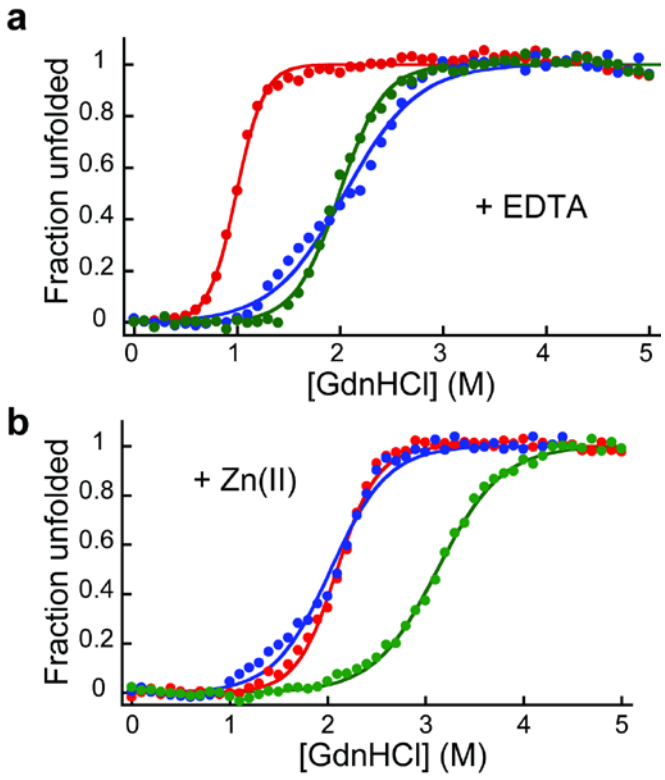

Figure 2.7. Chemical denaturation titrations of tetramer samples (2.5 μM) of ^{C96}RIDC1$_4$ (red), $^{C38/C96}$MBPC1$_4$ (blue), and $^{C38/C96}$RIDC1$_4$ (green) in the presence of (**a**) EDTA (1 mM) and (**b**) Zn(II) (5 μM). Data points of samples are shown in circles and fits for all samples are shown as line of the corresponding color. The denaturation midpoint values for each sample in the presence of EDTA and Zn(II) are shown in Table 2.3.

Upon adding Zn(II) to ^{C96}RIDC1$_4$ the tetramer is stabilized as indicated by the increase in denaturation midpoint. After comparing both of the four-disulfide-bonded tetramers in the absence of metal, the effect of the hydrophobic interface on $^{C38/C96}$RIDC1$_4$ versus $^{C38/C96}$MBPC1$_4$ appears to be negligible since the denaturation midpoint is the same for both species. Also, it is observed that for $^{C38/C96}$MBPC1$_4$ in the increase in chemical denaturation midpoint is almost negligible after the addition of Zn(II). This contrasts with the greater increase of the midpoint of $^{C38/C96}$RIDC1$_4$ in

the presence of metal. It is well documented that metal binding to an alpha helix in an *i*/*i+3* and *i*/*i+4* positions increases protein stability.[26-29] We can argue that the effect on the stability imparted by the addition of the hydrophobic interactions across interface *i1*, in addition to the Zn-binding event, add to a greater, cumulative stability for the $^{C38/C96}$RIDC1$_4$ scaffold.

This last observation is congruent with the different Zn(II) affinities measured for $^{C38/C96}$MBPC1$_4$ and $^{C38/C96}$RIDC1$_4$. In order to determine the dissociation constant of spectroscopically-silent Zn(II) for these tetramers, we performed competition assays between the tetramers and a chelating competitor which exhibits fluorescence upon binding Zn(II). Initially, we performed competition assays with Fura-2 and were able to obtain a binding curve for $^{C38/C96}$RIDC1$_4$ since it successfully competed in the same order of magnitude of the competitor ($K_{d,\ Zn\text{-}Fura}$ = 5.7 nM).[18] We determined that this construct binds with high-affinity to four Zn(II) sites which can be fit best to a four consecutive Zn-binding model (1 + 1 + 1 + 1 + 1) and have dissociation constants of low nM values for the first three binding events and of 10^{-8} M for the last binding event (Figure 2.8, Table 2.4). When comparing these values to those of the parent ^{C96}RIDC1$_4$ we find that the affinity for Zn(II) is very similar to the previously calculated values and that the additional two disulfide bonds do not increase the affinity for the metal for the protein. However, as discussed earlier, the disulfide bonds do add to the overall protein stability as determined with the chemical denaturation titrations. For $^{C38/C96}$MBPC1$_4$ it was determined that it could not compete with Fura-2 for Zn(II), so we proceeded to use the weaker Zn-chelator MagFura-2 for

the competition assay ($K_{d,\ Zn\text{-}MagFura2}$ = 47 nM).[30] After performing competition assay for $^{C38/C96}MBPC1_4$ with MagFura-2 for Zn(II) we were only able to observe one tight binding event which we could not fit successfully but can assume to be in the same order of magnitude as the competitor (Figure 2.9). However, we should note that on separate experiments we were able to identify four Zn(II) ions bound per tetramer after incubating $^{C38/C96}MBPC1_4$ (50 μM) with excess Zn(II). Even though the $^{C38/C96}MBPC1_4$ is pre-templated to bind Zn(II), the metal binding affinity is several orders of magnitude lower than the corresponding $^{C38/C96}RIDC1_4$. These results indicate very different dissociation constants between the tetramers and we can attribute these differences to the hydrophobic interactions across *i1* successfully stabilizing the Zn-bound form of the tetramer.

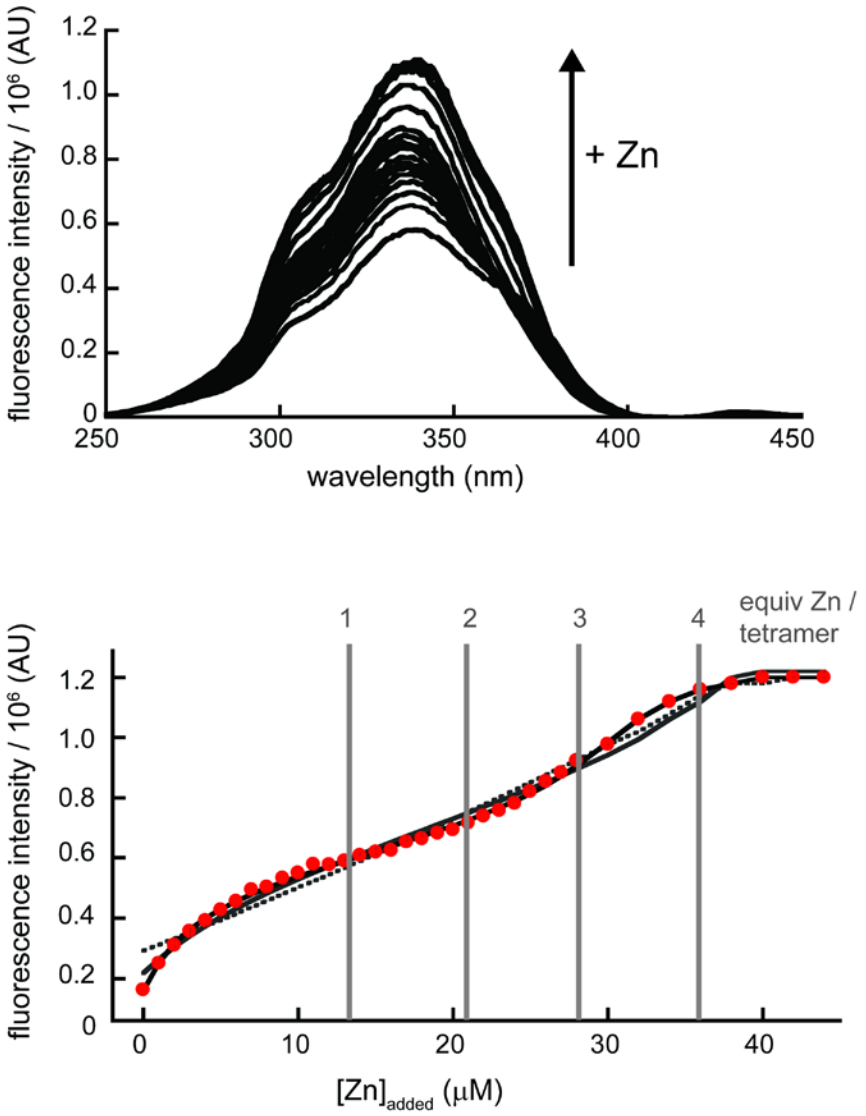

Figure 2.8. Competitive Zn binding by Fura-2 (5 μM) and $^{C38/C96}$RIDC1$_4$ (8.25 μM) monitored by changes in Fura-2 fluorescence excitation spectrum. **Top**: Fura-2 excitation spectra of different titration points showing the increase in emission at 505 nm. **Bottom**: Zn-binding isotherm plot of the emission of Fura-2 at 337 nm versus concentration of Zn(II) added. Gray vertical lines indicate the number of equivalents of Zn(II) binding per tetramer and demonstrate saturation after 4 equivalents of Zn(II) are added. The single equilibrium model (4 × 1, dotted line), or the two consecutive binding equilibrium model (2 + 2, solid gray line) do not appear to be adequate. The four consecutive Zn binding equilibrium model fits properly (black solid line).

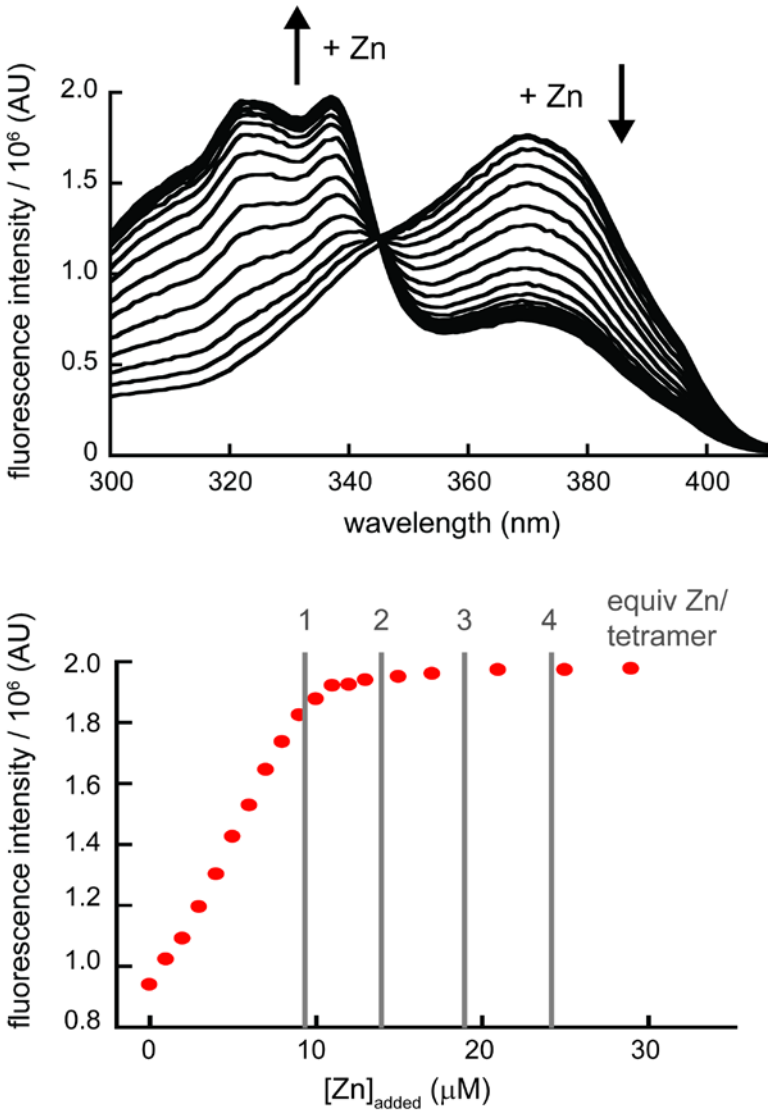

Figure 2.9. Competitive Zn(II) binding of MagFura-2 (5 μM) and $^{C38/C96}$MBPC1$_4$ (4.5 μM) monitored by changes in the MagFura-2 fluorescence excitation spectrum. **Top**: MagFura-2 excitation spectra of different titration points showing the increase in emission at 505 nm during the addition of Zn(II). **Bottom**: Zn-binding isotherm plot of the emission of Fura-2 at 337 nm versus concentration of Zn(II) added. Gray vertical lines indicate the concentration at which 1, 2, 3 or 4 equivalents of Zn(II) binding per tetramer are expected and since the endpoint occurs at 11 μM this corresponding to only ~1 competing Zn(II) binding site per tetramer.

$^{C38/C96}$RIDC1$_4$ tetramer can accommodate different transition metals.

In order to demonstrate that the $^{C38/C96}$RIDC1$_4$ metal-templated macrocycle can be used to bind different transition metals we incubated the purified tetramer (50 μM) in the presence of different equivalents of either Co(II), Ni(II), or Cu(II), and after using a desalting column to remove any unbound metal we proceeded to quantify metal content by ICP-OES analysis. The results shown in Table 2.5 demonstrate that the tetrameric scaffold is able to accommodate up to ~1.5 equivalents of Co(II), ~4 equivalents of Ni(II) and ~5 equivalents of Cu(II). Since we are able to remove the templating Zn(II) ion and substitute it for a different transition metal, we can validate our strategy as being similar to the design of small-molecule macrocycles which are able to accommodate different cations in order to obtain different properties and functionalities.

Encouraged by these results and to validate that our strategy allows for use of the pre-templated macrocycles to accommodate redox-active transition metals we proceeded to further characterize $^{C38/C96}$RIDC1$_4$ bound to Cu(II). First, we measured the dissociation constant of Cu(II) for the tetramer by performing Fura-2 competition assays with Cu(II) (Figure 2.10). We found that the protein is able to bind four Cu(II) ions with high affinity with dissociation constants measured to be ~10^{-13}-10^{-14} M (Figure 2.10, Table 2.6). These results contrast with the parent ^{C96}RIDC1$_4$ tetramer, which we determined to be able to bind only two Cu(II) sites with high affinity.[18] These differences can be due to the four disulfide bonded macrocyclic scaffold of $^{C38/C96}$RIDC1$_4$ being more rigid than the fluid ^{C96}RIDC1$_4$.

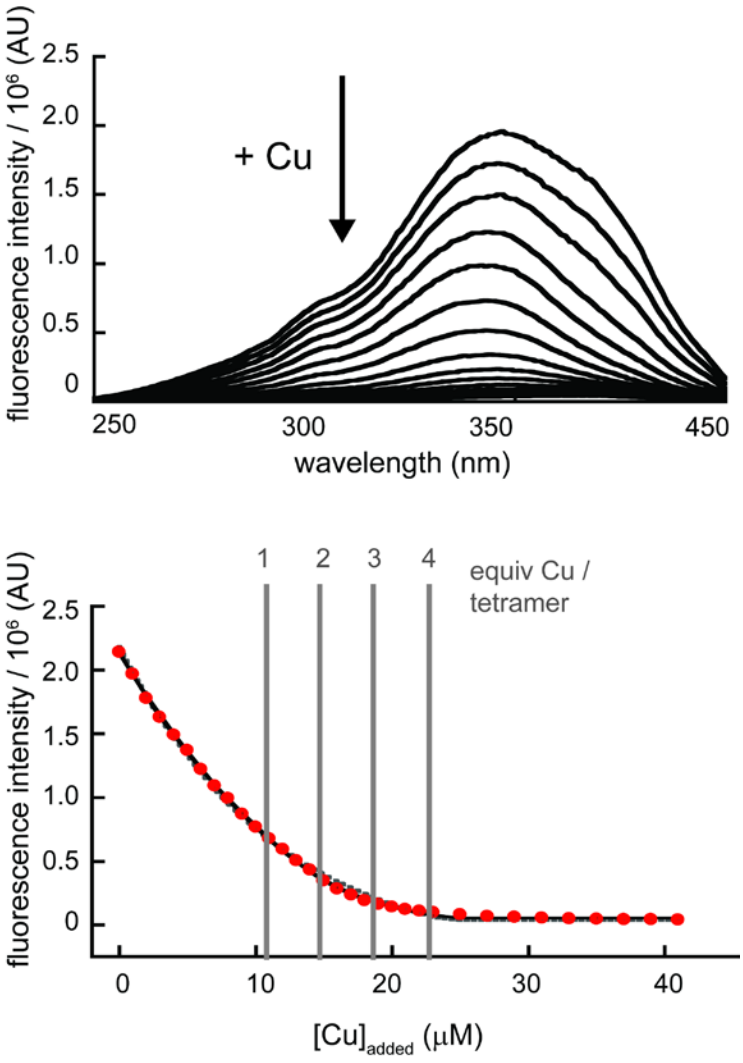

Figure 2.10. Competitive Cu(II) binding by Fura-2 (6.25 μM) and $^{C38/C96}$RIDC1$_4$ (4.25 μM) monitored by changes in Fura-2 fluorescence excitation spectrum. **Top**: Fura-2 excitation spectra of different titration points showing the increase in emission at 505 nm. **Bottom**: Cu(II)-binding isotherm plot of the emission of Fura-2 at 350 nm versus concentration of Cu(II) added. Gray vertical lines indicate the number of equivalents of Cu(II) binding per tetramer and demonstrate saturation after 4 equivalents of Cu(II) are added. The single equilibrium model (4 × 1, dotted line) does not appear to be adequate but the two consecutive binding equilibrium model (2 + 2, solid gray line) equilibrium model fits properly (black solid line) and so does the four consecutive binding equilibrium model (1 + 1 + 1 + 1) which is not shown since it is superimposable with the two consecutive binding model.

To characterize the Cu(II) site bound in solution we measured the Electron Paramagnetic Resonance spectrum of $^{C38/C96}RIDC1_4$ incubated with copper and identified it as having an axial spectrum with measured values of $g_{\|}$= 2.265, g_{perp} = 2.060 and $A_{Cu,\|}$ = 555 MHz, $A_{Cu,perp}$ = 45 MHz (Figure 2.11). These values correspond to a copper site being bound in a type-2 fashion with tetragonal coordination geometry. Superhyperfine splittings are identified which can correspond to the nitrogen atoms of histidines. The first derivative of the spectrum was taken and is shown as the inset in Figure 2.11 in order to highlight the superhyperfine splitting region. Although we cannot determine the exact coordination geometry of this construct we can model it successfully to a Cu(II) ion bound to three nitrogens (Figure 2.11 red simulation) which we can hypothesize can correspond to His59, His63, His73, and/or His77.

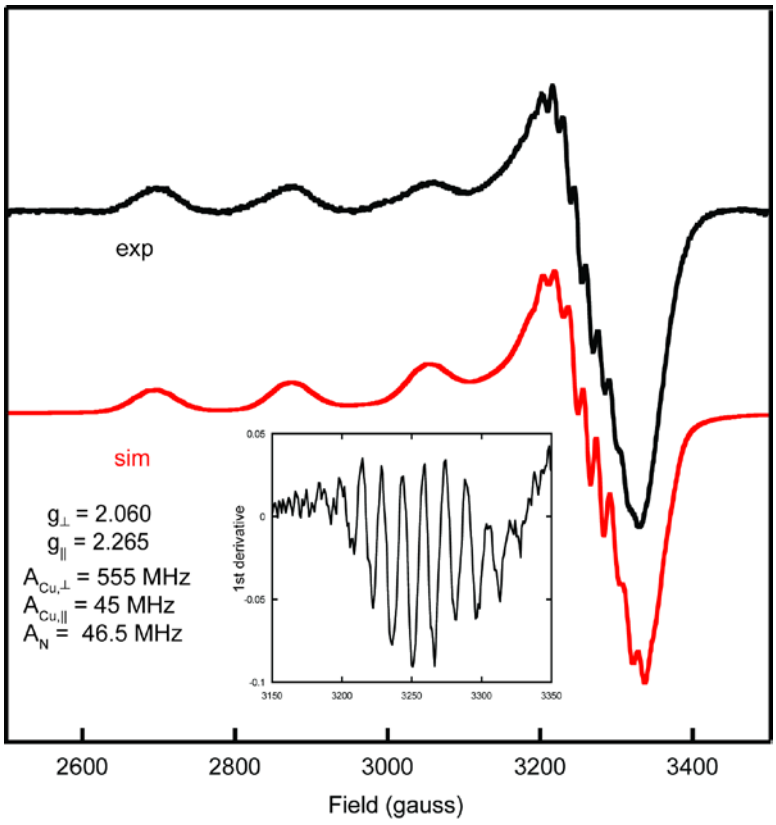

Figure 2.11. X-Band EPR spectrum of $^{C38/C96}$RIDC1$_4$ (300 µM) incubated with Cu(II). The experimental (black) spectrum is shown on top the simulation of an axial type of Cu(II) bound to 3 nitrogens is shown on the bottom (red) with simulated parameters for Cu(II) and nitrogen inside the spectrum. The inset is the first derivative of the intensity plotted against magnetic field in order to highlight the superhyperfine splittings.

Conclusions

The design and execution of multiple disulfide crosslinks on a protein multimer is very challenging since a protein with two or more thiols can form large undesirable polymeric assemblies when there is no direction to the crosslinking. Nature employs many strategies to ensure that disulfide bonds are formed correctly by using chaperone proteins like Protein Disulfide Isomerase (PDI).[31-35] It has been previously demonstrated that engineered disulfide bonds have to be carefully placed

because although they can stabilizing in some cases, they can also be deleterious to protein activity since they can place too much strain on the folded protein.[21,25,36,37] Also, the folding of recombinantly expressed proteins containing disulfide crosslinks requires the use of redox buffers or redox-active metals to catalyze disulfide bond formation and the correct folding of thermodynamically favored proteins.[38,39]

Using our metal templating strategy we were able to specifically guide the formation of the correct disulfide bonds on a macrocyclic protein assembly. The rigidity of the covalently crosslinked tetramers allowed us to place metals in the binding sites with minimal rearrangement of the metal-binding amino acids which we confirmed by obtaining the crystal structures of the Zn(II)-bound constructs. Also, we found that the hydrophobic interactions engineered into $^{C38/C96}$RIDC1$_4$ increase the affinity of the metal binding site for Zn(II) as opposed to lower affinity calculated for $^{C38/C96}$MBPC1$_4$. With these engineered scaffolds we are starting to explore how they can accommodate different transition metals at their core with the aim of getting reactive metal centers. We were able to accommodate different transition metals at the core of the macrocyclic protein assembly and particularly, we were able to obtain a high-affinity Cu(II)-bound tetramer. We propose that our protein design strategy will allow us to come up with new approaches into the design of stable metalloproteins which can be optimized to become catalytic assemblies. Our strategy should be generalizable to the application of the engineering of protein superstructures.

Table 2.1. Percentage of each oligomeric species (± s.d.) that was quantified from the templating reactions of $^{C38/C96}$MBPC1 and $^{C38/C96}$RIDC1 in the absence and presence of Zn(II) (as shown in Figure 2.3a).

	% monomer	% dimer	% trimer	% tetramer	% high MW species
$^{C38/C96}$MBPC1 + EDTA	33.5 ± 4.5	26.9 ± 1.1	17.7 ± 1.3	10.4 ± 1.3	11.5 ± 2.0
$^{C38/C96}$MBPC1 + Zn(II)	5.9 ± 0.9	9.7 ± 0.9	12.5 ± 0.4	34.7 ± 2.9	37.1 ± 3.7
$^{C38/C96}$RIDC1 + EDTA	7.2 ± 1.5	27.1 ± 3.2	21.7 ± 0.6	20.7 ± 0.8	23.2 ± 4.1
$^{C38/C96}$RIDC1 + Zn(II)	3.5 ± 0.6	7.6 ± 1.3	9.9 ± 0.6	43.7 ± 2.3	35.2 ± 4.1

Table 2.2. Percentage of each oligomeric species (± s.d.) that was quantified from templating reactions of $^{A74/C38/C96}$MBPC1 and $^{A74/C38/C96}$RIDC1 in the absence and presence of Zn(II) (as shown in Figure 2.3b).

	% monomer	% dimer	% trimer	% tetramer	% high MW species
$^{A74/C38/C96}$MBPC1 + EDTA	14.3 ± 1.5	40.6 ± 1.0	19.4 ± 0.8	11.4 ± 0.7	14.2 ± 1.8
$^{A74/C38/C96}$MBPC1 + Zn(II)	12.4 ± 0.7	22.4 ± 1.4	19.1 ± 0.8	14.6 ± 0.3	31.5 ± 1.4
$^{A74/C38/C96}$RIDC1 + EDTA	24.4 ± 2.3	38.7 ± 2.3	18.1 ± 0.8	12.0 ± 1.2	6.7 ± 3.4
$^{A74/C38/C96}$RIDC1 + Zn(II)	23.7 ± 1.5	17.7 ± 1.0	14.0 ± 1.3	22.6 ± 1.7	22.0 ± 3.9

Table 2.3. Midpoints of chemical denaturation titrations of $^{C38/C96}$MBPC1$_4$, ^{C96}RIDC1$_4$ and $^{C38/C96}$RIDC1$_4$ tetramers in the absence and presence of Zn (from figure 2.7).

	Midpoint in presence of EDTA (M)	Midpoint in presence of Zn (M)
$^{C38/C96}$MBPC1$_4$	2.0	2.1
^{C96}RIDC1$_4$	1.0	2.1
$^{C38/C96}$RIDC1$_4$	2.0	3.1

Table 2.4. Dissociation constants for $^{C38/C96}$RIDC1$_4$ binding to Zn(II) determined by a competition assay with Fura-2 and fit to a four-Zn equilibrium binding model (as shown in figure 2.8).

K_{d1}	$(3.8 \pm 1.6) \times 10^{-9}$
K_{d2}	$(2.9 \pm 1.2) \times 10^{-9}$
K_{d3}	$(2.3 \pm 0.9) \times 10^{-9}$
K_{d4}	$(3.5 \pm 1.0) \times 10^{-8}$

Table 2.5. Equivalents per tetramer of either Cu(II), Ni(II), and Co(II) that were quantified to bind to $^{C38/C96}$RIDC1$_4$ after incubation of samples in the presence of the metal and quantifying the remaining metal through ICP-OES.

Metal added	Metal equivalents / tetramer
Cu	4.9 ± 0.4
Ni	3.8 ± 0.4
Co	1.5 ± 0.3

Table 2.6. Dissociation constants for $^{C38/C96}$RIDC1$_4$ binding to Cu(II) determined by a competition assay with Fura-2 and fit to two consecutive (2 + 2) or four consecutive (1 + 1 + 1 + 1) equilibrium binding models (as shown in figure 2.10).

	Two-site model	Four-site model
K_{d1}	$(3.0 \pm 1.0)\text{x}10^{-13}$	$(2.2 \pm 0.6)\text{x}10^{-14}$
K_{d2}	$(4.4 \pm 0.2)\text{x}10^{-12}$	$(3.6 \pm 1.0)\text{x-}10^{-13}$
K_{d3}	- - -	$(5.5 \pm 1.8)\text{x}10^{-12}$
K_{d4}	- - -	$(4.8 \pm 1.6)\text{x}10^{-12}$

Table 2.7. Crystallization conditions.

	Protein	Precipitant
$^{C38/C96}$MBPC1$_4$	2 mM in 20 mM TRIS (pH 7) with 2 mM Zn(II)	PEG400, HEPES pH 7.5
$^{C38/C96}$RIDC1$_4$	2 mM in 20 mM TRIS (pH 7) with 2 mM Zn(II)	PEG1900 MME

Table 2.8. Data collection and refinement statistics.

	Zn_4:$^{C38/C96}$MBPC1_4	Zn_4:$^{C38/C96}$RIDC1_4
Data collection		
Space group	$P12_11$	$P6_122$
Cell dimensions		
a, *b*, *c* (Å)	57.338, 56.698, 126.377	52.487, 52.487, 255.666
α, β, γ (˚)	90.00, 101.73, 90.00	90.00, 90.00, 120.00
Resolution (Å)	1.80 (61.90)	2.22 (63.92)
R_{sym} or R_{merge} (%)*	5.7 (14.2)	5.4 (19.4)
$I/\sigma I$*	12.9 (5.4)	32.5 (12.6)
Completeness (%)*	98.6 (100.0)	99.8 (100.0)
Redundancy	3.7	15.2
Refinement		
Resolution (Å)	1.80	2.22
No. unique reflections	77,772	13,731
R_{work}/R_{free}	0.198/0.243	0.237/0.279
No. atoms		
Protein	6873	1724
Ligand/ion	8	7
Water	587	43
B-factors		
Protein	30.48	52.85
Ligand/ion	18.86	57.11
Water	33.2	45.06
R.m.s. deviations		
Bond lengths (Å)	0.0236	0.0226
Bond angles (˚)	3.9915	2.3271

Materials and Methods

Protein Mutagenesis, Purification, Tetramer Templating Reactions and Isolation:

The $^{C38/C96}$MBPC1, $^{C38/C96}$RIDC1, $^{A74/C38/C96}$MBPC1 and $^{A74/C38/C96}$RIDC1 cyt cb_{562} constructs were generated, expressed, and purified according to published procedures.[17,30,40] To the purified proteins was added excess DTT to reduce any disulfide bonds formed during the purification procedure. Removal of DTT and exchange into 50 mM TRIS (pH 7) and 250 mM NaCl, was done by using a 10DG desalting column (BioRad). Proteins were diluted to 100 µM concentration and incubated in the presence of one equivalent (100 µM) of $ZnCl_2$. The samples were incubated overnight in a rotary shaker at 35 ˚C and the crosslinked samples were later run on a size exclusion column (Superdex75) using an elution buffer of 20 mM TRIS (pH 7) and 150 mM NaCl in order to isolate the tetrameric fractions. Tetramer purity was determined by running the eluted fractions on non-reducing 12% SDS-PAGE gels. The tetramers were concentrated and excess EDTA was added to remove any metal ions bound during templating process and the samples were exchanged into Chelex-treated buffer of 20 mM MOPS (pH 7) and 150 mM NaCl, flash-frozen in 200 µL aliquots and stored at -80 ˚C until further use.

Zn or EDTA Templating for Quantification of Oligomeric Species:

The purified, reduced, monomeric proteins were exchanged into 50 mM TRIS (pH 7) 150 mM NaCl, diluted to 50 µM concentration, and 50 µL aliquots were incubated in the presence of either one equivalent of Zn(II) or excess (1 mM) of EDTA. After incubation, aliquots of the products were run on non-reducing 12%

SDS-PAGE gels for 40 minutes at 400 mV which were stained with Coomassie Brilliant Blue and destained with MeOH/HNO_3 solution. The gels were imaged using a scanner and the density of the bands was quantified and integrated using ImageJ.[20]

Analytical Ultracentrifugation:

Samples of either 1.25 μM $^{C38/C96}MBPC1_4$ or $^{C38/C96}RIDC1_4$ for sedimentation velocity experiments were made in 20 mM MOPS (pH 7), 150 mM NaCl and contained either 5 μM Zn(II) or 1 mM EDTA. Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor at 41,000 rpm for 300 scans per sample at 25 ˚C by detection at 415 nm. The data were processed using Sedfit[41]. Buffer viscosity (0.01002 poise), density (1.007 g/mL) and protein partial specific volume (0.7316 mL/g) were calculated at 25 ˚C with SEDNTERP (http://www.jphilo.mailway.com).

X-Ray Crystallography:

All crystals were obtained by sitting-drop vapor diffusion with growth conditions detailed in Table 2.7. The sitting drops each consisted of 2 μL protein solution and 1 μL precipitant solution. The wells contained 500 μL precipitant solution. Crystals grew at room temperature typically within a two-day period and matured over a few weeks. X-ray diffraction data were collected at 100 K at the Stanford Synchrotron Radiation Lightsource (SSRL). Diffraction data were processed using MOSFLM and SCALA on the CCP4 suite.[42] Structures were solved with molecular replacement by MOLREP[43] using the a single chain of the structure of the $^{C96}RIDC1$ (PDB ID:2IQ6) as a search model. Rigid-body, positional, thermal and TLS

refinement was carried out using REFMAC,[44] along with iterative manual model building and water/ligand/metal placement with COOT.[45] All structural figures were produced with PYMOL.[46] Ramachandran plots were determined using PROCHECK.[47]

Metal content analysis by ICP-OES:

Protein samples of 50 μM $^{C38/C96}$RIDC1$_4$ were incubated in 50 mM TRIS (pH 7), 150 mM NaCl with 1, 2, or 3 equivalents of either Zn(II), Cu(II), Co(II) and Ni(II) for 16 hours at room temperature on a shaker. The samples were run on desalting column (BioRad) to remove any unbound metal and the eluted protein was diluted to 2 mL by addition of Chelex-treated water and 69% HNO_3 to a final concentration of 3% HNO_3. ICP-OES standard metal solutions were prepared from 1,000-ppm certified metal stock solutions (Ricca) by serial dilutions, and a standard curve was prepared within a range of 0.1-10 ppm of metals in 3% nitric acid and Chelex-treated water. Data were collected on a Perkin-Elmer Optima 3000 DV ICP-OES spectrometer located at the Analytical Facility of the Scripps Institution of Oceanography. Because each protein monomer contains one Fe atom as part of the covalently linked heme molecule, the experimentally determined Zn:Fe ratios directly yielded the Zn:monomeric protein ratios. Wavelengths used for the detection of metal ions were Fe (234.349, 238.204, 239.562, 259.939, and 273.055 nm), Zn (202.548, 206.2, and 213.857 nm), Cu (327.393, 324.752, 224.7, 213.597, 222.778, and 221.459), Ni (231.604, 221.648, 341.476), and Co (238.892, 230.786, 236.38, and 231.16). All values reported for each metal are the averages of indicated wavelengths from 3 separate experiments.

Metal affinity competition assays:

$^{C38/C96}$MBPC1$_4$ with MagFura-2 and Zn(II): A 1-mL solution containing 5 μM MagFura-2 (Life Technologies) and 4.5 μM $^{C38/C96}$MBPC1$_4$ was titrated in a 1-cm cuvette with 1-2 μL aliquots of 1 mM $ZnCl_2$. After each addition, the sample was equilibrated for at least 5 minutes at room temperature, upon which the excitation scan was recorded on a Horiba Fluorolog 2 fluorimeter using an excitation range of 250-450 nm and detection at 510 nm. The change in excitation intensity at 335 nm was plotted versus Zn(II) concentration.

$^{C38/C96}$RIDC1$_4$ with Fura-2 and Zn(II): A similar protocol was used in the case of $^{C38/C96}$RIDC1$_4$ (8.25 μM) in competition titration experiments with Fura-2 (5 μM, Life Technologies). A 1-mL solution was titrated in a 1-cm cuvette with 1-2 μL aliquots of 1 mM $ZnCl_2$ and after each addition the sample was equilibrated for at least 5 minutes. The excitation scan was recorded on as described and the change in excitation intensity at 335 nm was plotted versus Zn(II) concentration. , and the resulting curve was fit alternatively to 1×4 Zn, 2×2 Zn or 4×1 Zn binding models using Dynafit[48] as described previously.[18]

$^{C38/C96}$RIDC1$_4$ with Fura-2 and Cu(II): Similarly, samples of $^{C38/C96}$RIDC1$_4$ (4.25 μM) and Fura-2 (6.25 μM, Life Technologies) were titrated in a 1-cm cuvette with 1-2 μL aliquots of 1 mM $CuCl_2$ and after each addition the sample was equilibrated for at least 5 minutes. The excitation scan was recorded on as described and the change in excitation intensity at 350 nm was plotted versus Zn(II)

concentration. , and the resulting curve was fit alternatively to 1×4 Cu(II), 2×2 Cu(II) or 4×1 Cu(II) binding models using Dynafit[48] as described previously.[18]

Denaturation curves:

The ^{C96}RIDC1$_4$, $^{C38/C96}$MBPC1$_4$, and $^{C38/C96}$RIDC1$_4$ tetrameric proteins were diluted to a concentration of 2.5 μM with 10 mM sodium borate (pH 7), 150 mM NaCl buffer. Samples contained either 1 mM EDTA or a two-fold excess of Zn in order to ensure that the protein was fully loaded. Denaturation curves were measured in a 1 cm pathlength quartz cuvette using an autotitrator (Hamilton) in an Aviv 215 circular dichroism spectrometer. The samples were equilibrated following 60 s of equilibration at 25 ˚C. The ellipticity was measured at 222 nm and plotted against concentration of denaturant (guanidine hydrochloride).

Electron Paramagnetic Resonance Measurements:

EPR samples had been previously incubated with 4 equivalents of Cu(II) per tetramer and the unbound metal was washed using a 10DG desalting column (BioRad). The samples consisted of 300 μM $^{C38/C96}$RIDC1$_4$ in 20 mM MOPS buffer (pH 7) with 150 mM NaCl. X-band spectra were obtained with a Bruker ELEXSYS E500 spectrometer equipped with a Bruker ER4131VT variable-temperature unit. Spectra were recorded at 125 K at a microwave frequency of 9873 MHz, 0.10 mT modulation amplitude and 0.63 mW power. The simulation parameters were obtained with EasySpin[49] software by using the program pepper which calculates continuous wave EPR spectra for single-crystal and powder samples. Broadening parameters and script is included in Appendix.

Chapter 2 is currently being prepared for publication of the material Medina-Morales, A. M., Brodin, J. D., Tezcan, F. A. The dissertation author was the primary author was the primary investigator and author of this paper.

References

1 Bergdoll, M., Eltis, L. D., Cameron, A. D., Dumas, P. & Bolin, J. T. All in the family: Structural and evolutionary relationships among three modular proteins with diverse functions and variable assembly. *Protein Science* **7**, 1661-1670 (1998).

2 Choi, M. & Davidson, V. L. Cupredoxins—A study of how proteins may evolve to use metals for bioenergetic processes. *Metallomics* **3**, 140-151 (2011).

3 Mauk, A. G. & Moore, G. R. Control of metalloprotein redox potentials: what does site-directed mutagenesis of hemoproteins tell us? *JBIC Journal of Biological Inorganic Chemistry* **2**, 119-125 (1997).

4 DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F. & Lombardi, A. De novo design and structural characterization of proteins and metalloproteins. *Annual review of biochemistry* **68**, 779-819 (1999).

5 Shifman, J. M., Gibney, B. R., Sharp, R. E. & Dutton, P. L. Heme Redox Potential Control in de Novo Designed Four-α-Helix Bundle Proteins†. *Biochemistry* **39**, 14813-14821 (2000).

6 Vallee, B. L. & Williams, R. Metalloenzymes: the entatic nature of their active sites. *Proceedings of the National Academy of Sciences of the United States of America* **59**, 498 (1968).

7 Marshall, N. M. *et al.* Rationally tuning the reduction potential of a single cupredoxin beyond the natural range. *Nature* **462**, 113-116 (2009).

8 Lancaster, K. M., Sproules, S., Palmer, J. H., Richards, J. H. & Gray, H. B. Outer-sphere effects on reduction potentials of copper sites in proteins: The curious case of high potential type 2 C112D/M121E Pseudomonas aeruginosa azurin. *Journal of the American Chemical Society* **132**, 14590-14595 (2010).

9 Harding, M. M. Geometry of metal-ligand interactions in proteins. *Acta Crystallographica Section D: Biological Crystallography* **57**, 401-411 (2001).

10 Que, L. & Tolman, W. B. Biologically inspired oxidation catalysis. *Nature* **455**, 333-340 (2008).

11 Tye, J. W., Darensbourg, M. Y. & Hall, M. B. De novo design of synthetic di-iron (I) complexes as structural models of the reduced form of iron-iron hydrogenase. *Inorganic chemistry* **45**, 1552-1559 (2006).

12 McMurry, T. J., Raymond, K. N. & Smith, P. H. Molecular Recognition and Metal-Ion Template Synthesis. *Science* **244**, 938-943 (1989).

13 Philp, D. & Stoddart, J. F. Self-assembly in natural and unnatural systems. *Angewandte Chemie International Edition in English* **35**, 1154-1196 (1996).

14 Linton, B. & Hamilton, A. D. Formation of Artificial Receptors by Metal-Templated Self-Assembly. *Chemical Reviews* **97**, 1669-1680 (1997).

15 Anderson, S., Anderson, H. L. & Sanders, J. K. M. Expanding roles for templates in synthesis. *Accounts of Chemical Research* **26**, 469-475 (1993).

16 Corbett, P. T. *et al.* Dynamic Combinatorial Chemistry. *Chemical Reviews* **106**, 3652-3711 (2006).

17 Salgado, E. N., Faraone-Mennella, J. & Tezcan, F. A. Controlling Protein-Protein Interactions through Metal Coordination: Assembly of a 16-Helix Bundle Protein. *J. Am. Chem. Soc.* **129**, 13374-13375 (2007).

18 Brodin, J. D. *et al.* Evolution of Metal Selectivity in Templated Protein Interfaces. *Journal of the American Chemical Society* **132**, 8610-8617 (2010).

19 Salgado, E. N. *et al.* Metal-Templated Design of Protein Interfaces. *Proeedings of the National Academy of Sciences USA* **107**, 1827-1832 (2010).

20 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675 (2012).

21 Matsumura, M. & Matthews, B. W. Control of Enzyme-Activity by an Engineered Disulfide Bond. *Science* **243**, 792-794 (1989).

22 Hodges, R., Zhou, N., Kay, C. & Semchuk, P. Synthetic model proteins: contribution of hydrophobic residues and disulfide bonds to protein stability. *Peptide research* **3**, 123 (1990).

23 Betz, S. F. Disulfide bonds and the stability of globular proteins. *Protein Science* **2**, 1551-1558 (1993).

24 Wells, J. A. & Powers, D. In vivo formation and stability of engineered disulfide bonds in subtilisin. *Journal of Biological Chemistry* **261**, 6564-6570 (1986).

25 Pace, C. N., Grimsley, G. R., Thomson, J. A. & Barnett, B. J. Conformational Stability and Activity of Ribonuclease-T1 with Zero, One, and 2 Intact Disulfide Bonds. *Journal of Biological Chemistry* **263**, 11820-11825 (1988).

26 Ghadiri, M. R. & Choi, C. Secondary Structure Nucleation in Peptides - Transition-Metal Ion Stabilized Alpha-Helices. *Journal of the American Chemical Society* **112**, 1630-1632 (1990).

27 Kellis, J. T., Todd, R. J. & Arnold, F. H. Protein stabilization by engineered metal chelation. *Nature Biotechnology* **9**, 994-995 (1991).

28 Radford, R. J., Nguyen, P. C., Ditri, T. B., Figueroa, J. S. & Tezcan, F. A. Controlled Protein Dimerization through Hybrid Coordination Motifs. *Inorganic Chemistry* **49**, 4362-4369 (2010).

29 Arnold, F. H. & Zhang, J.-J. Metal-mediated protein stabilization. *Trends Biotechnol.* **12**, 189-192 (1994).

30 Salgado, E. N., Brodin, J. D., To, M. M. & Tezcan, F. A. Templated Construction of a Zn-Selective Protein Dimerization Motif. *Inorganic Chemistry* **50**, 6323-6329 (2011).

31 Ferrari, D. M. & Soling, H. D. The protein disulphide-isomerase family: unravelling a string of folds. *Biochemical Journal* **339**, 1-10 (1999).

32 Gilbert, H. F. Thiol/Disulfide Exchange Equilibria and Disulfide Bond Stability. *Biothiols, Pt A* **251**, 8-28 (1995).

33 Riemer, J., Bulleid, N. & Herrmann, J. M. Disulfide formation in the ER and mitochondria: two solutions to a common process. *Science* **324**, 1284-1287 (2009).

34 Wilkinson, B. & Gilbert, H. F. Protein disulfide isomerase. *Bba-Proteins Proteom* **1699**, 35-44 (2004).

35 Xiao, R. Y. *et al.* The contributions of protein disulfide isomerase and its homologues to oxidative protein folding in the yeast endoplasmic reticulum. *J Biol Chem* **279** (2004).

36 Hodges, R. S. De novo design of alpha-helical proteins: Basic research to medical applications. *Biochemistry and Cell Biology* **74**, 133-154 (1996).

37 Zhou, N. E., Kay, C. M. & Hodges, R. S. Disulfide Bond Contribution to Protein Stability - Positional Effects of Substitution in the Hydrophobic Core

of the 2-Stranded Alpha-Helical Coiled-Coil. *Biochemistry-Us* **32**, 3178-3187 (1993).

38 Ahmed, A. K., Schaffer, S. W. & Wetlaufer, D. B. Nonenzymic Reactivation of Reduced Bovine Pancreatic Ribonuclease by Air Oxidation and by Glutathione Oxidoreduction Buffers. *Journal of Biological Chemistry* **250**, 8477-8482 (1975).

39 Lilie, H., Schwarz, E. & Rudolph, R. Advances in refolding of proteins produced in E-coli. *Current opinion in biotechnology* **9**, 497-501 (1998).

40 Salgado, E. N., Lewis, R. A., Faraone-Mennella, J. & Tezcan, F. A. Metal-mediated self-assembly of protein superstructures: Influence of secondary interactions on protein oligomerization and aggregation. *Journal of the American Chemical Society* **130**, 6082-6084 (2008).

41 Schuck, P. A model for sedimentation in inhomogeneous media. I. Dynamic density gradients from sedimenting co-solutes. *Biophysical Chemistry* **108**, 187-200 (2004).

42 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Biological Crystallography* **67**, 235-242 (2011).

43 Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallographica Section D-Biological Crystallography* **66**, 22-25 (2010).

44 Murshudov, G., Vagin, A. & Dodson, E. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica* **D53**, 240-255 (1996).

45 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **D60**, 2126-2132 (2004).

46 DeLano, W. L. *The PYMOL Molecular Graphics System (http://www.pymol.org).* (2003).

47 Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - A Program To Check The Stereochemical Quality Of Protein Structures. *Journal of Applied Crystallography.* **26**, 283-291 (1993).

48 Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase. *Analytical Biochemistry* **237**, 260-273 (1996).

49 Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of Magnetic Resonance* **178**, 42-55 (2006).

Chapter 3

In Vitro and Cellular Self-Assembly of a Protein Cryptand via Templated Disulfide Bonds

Introduction

Chemical templating, generally defined as the use of non-covalent interactions to spatially organize molecules to promote the selective formation of covalent bonds between them,[1,2] is pervasive in nature. Templating has also been a powerful strategy in the synthesis of complex, covalently linked architectures.[3,4] In coordination chemistry, metal ions have been exploited to direct the formation of macrocylic scaffolds,[5,6] ranging from phthalocyanines as initial examples[7] to the more complex polycyclic compounds that serve as selective hosts for metal ions.[8,9] Likewise, in supramolecular chemistry, templating has enabled the synthesis of a diverse array of challenging synthetic targets, including interlocked structures such as rotaxanes, catenanes and molecular knots.[10-13] The scope of such synthetic efforts in chemical templating has been expanded through the use of biological building blocks such as DNA,[14] the development of self-replicating peptides,[15-17] and broadly, the advent of dynamic combinatorial/covalent chemistry.[18-20] These advances in synthetic templating strategies prompted us to explore whether they can also be employed toward the construction of protein folds or assemblies, which execute most of the fundamental cellular tasks, and thus are primarily responsible for biological complexity. Specifically, we asked if we could template the selective formation of multiple disulfide bonds on a protein building block to control its assembly into discrete supramolecular architectures.

Tertiary and quaternary protein structures are largely shaped by noncovalent interactions, yet many require multiple disulfide bonds between cysteine residues for

stability and function.[21,22] Even though disulfide bonds are covalent and stable (60 kcal/mol) under oxidizing conditions, they are reversible through two-electron reduction and responsive to external stimuli such as pH and ionic strength.[23] This unique combination of stability, reversibility and stimuli-responsiveness not only accounts for the prevalence of disulfide bonds in biological systems, but also for their increasing use in synthetic supramolecular chemistry.[18-20] Despite these advantages, however, disulfide bonds are rarely exploited in *de novo* protein design. There are several reasons: first, this goal still requires the careful programming of non-covalent interactions for positioning (i.e., templating) several cysteines correctly, a grand challenge on its own.[24] Second, it requires the consideration of the multiple folding/assembly pathways involving alternative cysteine pairs.[21] Third, protein constructs with multiple thiol groups are difficult to produce and handle in experimental settings.[25] To our knowledge, there are no examples of designed proteins with multiple disulfide bonds.

Recently, we have applied design strategies of supramolecular coordination chemistry to control the assembly of monomeric proteins into closed or infinite protein superstructures.[26-28] Our approach is based on the premise that the strength and directionality of metal coordination interactions built onto protein surfaces can direct the assembly of symmetrical protein superstructures, whose interfaces can be subsequently stabilized through the introduction of favorable non-covalent interactions around the metal nuclei. Here, we show that such re-engineered non-covalent interactions can template the selective formation of multiple disulfide bonds that link

together a novel superprotein assembly. This assembly, $^{C81/C96}$RIDC1$_4$, features four interconnected disulfide bonds in a novel, cryptand-like topology. In analogy to natural metalloenzymes, $^{C81/C96}$RIDC1$_4$ is highly pre-organized for metal coordination, has a large internal cavity surrounding the metal sites, and is resistant to alterations in inner-sphere metal coordination, all of which constitute challenging targets for protein design. What is more, $^{C81/C96}$RIDC1$_4$ correctly self-assembles and incorporates Zn(II) ions in bacterial cells, thus forging a new link between synthetic supramolecular chemistry and synthetic biology.

Results and Discussion

Design Strategy for $^{C81/C96}$RIDC1.

The protein construct $^{C81/C96}$RIDC1 (the C81/C96 double mutant of Rosetta Interface Designed Cytochrome 1, Figure 3.1a), which is the monomeric building block for the $^{C81/C96}$RIDC1$_4$ tetramer, is a derivative of the four-helix-bundle hemeprotein, cytochrome *cb*$_{562}$.[29] We previously reported a first-generation cytochrome *cb*$_{562}$ mutant (MBPC1; Metal Binding Protomer of Cytochrome 1), which was engineered on its surface with metal coordinating residues (orange sticks in Figure 3.1a and 3.1b).[30] Upon binding various divalent transition metal ions, MBPC1 was observed to assemble into discrete assemblies whose compositions and supramolecular symmetries were dictated by the stereochemical preferences of the metal ions.[31] In particular, tetrahedral coordination of Zn(II) by residues His63, Asp74, His73, and His77 led to the formation of a tetrameric, D_2 symmetric assembly

(Figure 3.1b).[30] Owing to its dihedral symmetry, Zn_4:MBPC1_4 possesses an interlocked architecture with particularly extensive (~5000 $Å^2$ total), but non-favorable interfacial interactions, which are distributed over three pairs of orthogonal, C_2-symmetric interfaces (*i1*, *i2* and *i3*) (Figure 3.1b). These three interfaces are interconnected but non-overlapping. Consequently, the stabilization of any of these interfaces should also favor the formation of the other two interfaces (as well as the entire assembly), yet, from a protein engineering perspective, they should be independently addressable for redesign.

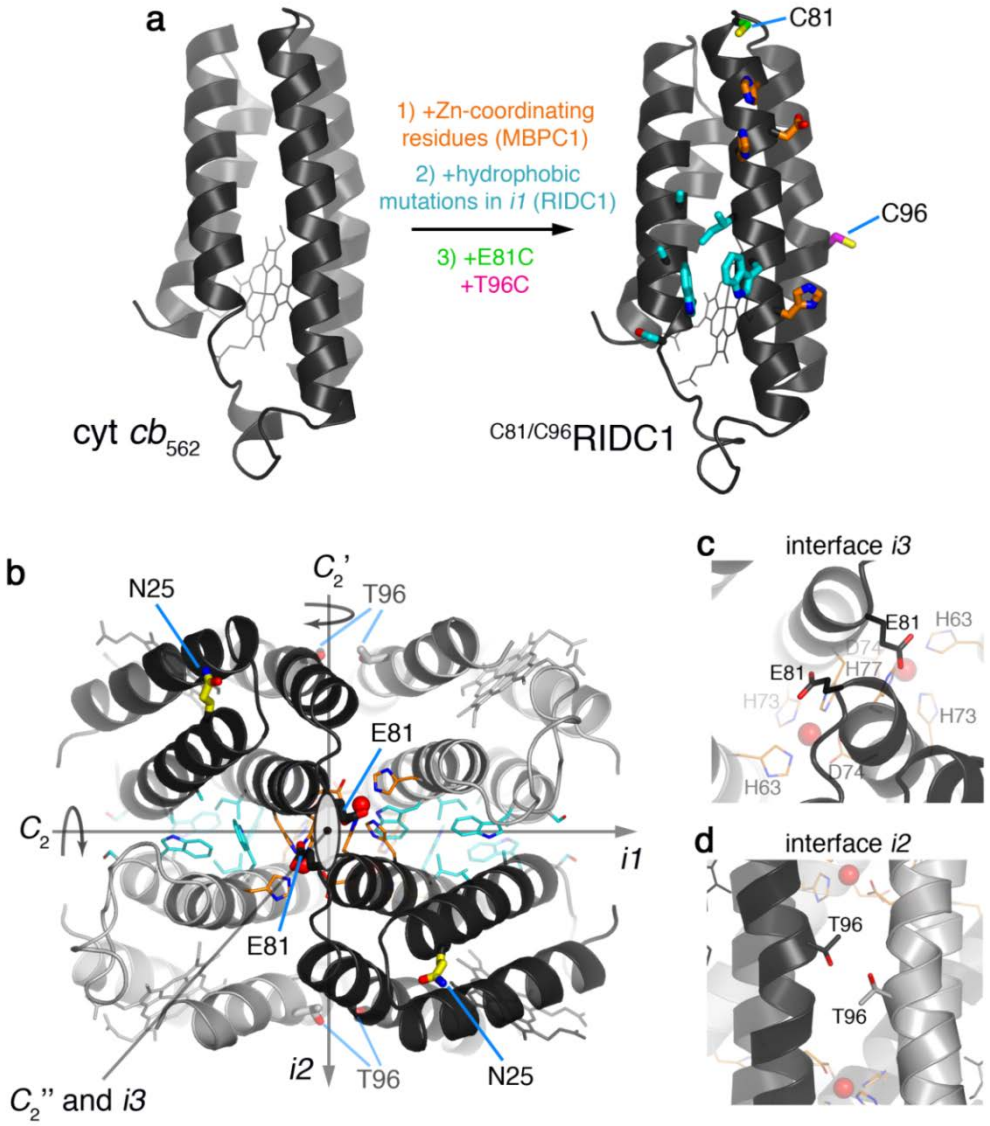

Figure 3.1. Structural features of cyt cb_{562} central to the design of $^{C81/C96}$RIDC1. (**a**) Three stages of mutations made on cyt cb_{562} to convert it to $^{C81/C96}$RIDC1. The mutated residues to construct MBPC1 and RIDC1 are shown as orange and cyan-colored sticks, respectively. Two engineered surface cysteines, C81 and C96, are highlighted in green and magenta, respectively. (**b**) *D2* symmetric structure of Zn_4-$RIDC1_4$ viewed down the C_2" axis, along which interface *i3* is formed. *i1* and *i2* are formed along the orthogonal C_2 and C_2' axes, which lie parallel to the plane of view. Hydrophobic residues inserted into *i1*, as well as the symmetrically related pairs of Glu81 and Thr96 residues in *i2* and *i3*, which are mutated to Cys, are highlighted. (**c**, **d**) Close-up views of Glu81 and Thr96.

We had shown that *i1*, the most extensive and close-packed of the three interfaces, can be stabilized through the incorporation of six mostly hydrophobic mutations (cyan sticks in Figure 3.1a and 3.1b) onto the MBPC1 monomer surface to generate RIDC1 and subsequently its Zn-mediated tetramer Zn_4:$RIDC1_4$, which is isostructural with Zn_4:$MBPC1_4$.[32] Indeed, in the absence of Zn(II) binding, RIDC1 was able to dimerize ($K_{d,\ dimer}$ = 26 μM) via the engineered hydrophobic interactions in an antiparallel fashion, as was observed in the Zn-driven tetrameric architecture.[32] *i3* is too small for stabilization through noncovalent interactions, and though extensive, *i2* features too wide a clearance between monomers for efficient hydrophobic packing. Nevertheless, both *i2* and *i3* present two pairs of symmetrically related residue positions (96-96' in *i2* and 81-81' in *i3*) across from one another in close proximity for potentially forming Cys-Cys disulfide bonds (C_α-C_α distances < 5.6 Å) (Figure 3.1c and 3.1d). Indeed, 96-96' disulfide bonds were shown to be readily accommodated in the Zn_4:$^{C96}RIDC1_4$ tetramer.[33] Since ^{C96}RIDC1 has only one surface Cys, selective disulfide formation between monomers is not an issue. Here, we set out to take advantage of the dihedral symmetry of the Zn-templated tetramer, and probe whether complementary, non-covalent interactions built into the *i1* interfaces can template the correct formation of four disulfide bonds (2 × Cys96-Cys96' and 2 × Cys81-Cys81') across the other two interfaces. We thus prepared the $^{C81/C96}$RIDC1 mutant through Glu81Cys and Thr96Cys mutations (green and magenta sticks in Figure 3.1a).

Oxidative Self-Assembly of $^{C81/C96}$RIDC1.

In the first set of experiments, solutions of isolated $^{C81/C96}$RIDC1 (typically 50 μM in concentration) were subjected to air oxidation at room temperature, and the time course for the formation of disulfide-containing products was determined by gel electrophoresis (SDS-PAGE) under denaturing and non-reducing conditions. The major products were a distinct tetrameric species (48±2% in abundance) and various higher order aggregates (40±3%) (Figure 3.2a, top panel). These products formed over the course of 24 h following a noticeable induction phase, during which a dimeric intermediate was transiently populated (Figure 3.2a, bottom panel). A small, but persistent population (5-8%) of trimeric species was also detected. At 35 °C, the rates of monomer consumption and tetramer/polymer production were considerably higher, and the induction period was largely eliminated, although the product distribution did not change significantly (Figure 3.3a and 3.3b). Inclusion of equimolar Zn(II) in the reaction mixture altered neither the rate of formation significantly nor the distribution of products (Figure 3.4a and 3.4b), suggesting that Zn(II) coordination does not contribute to the stability of the self-assembly intermediates.

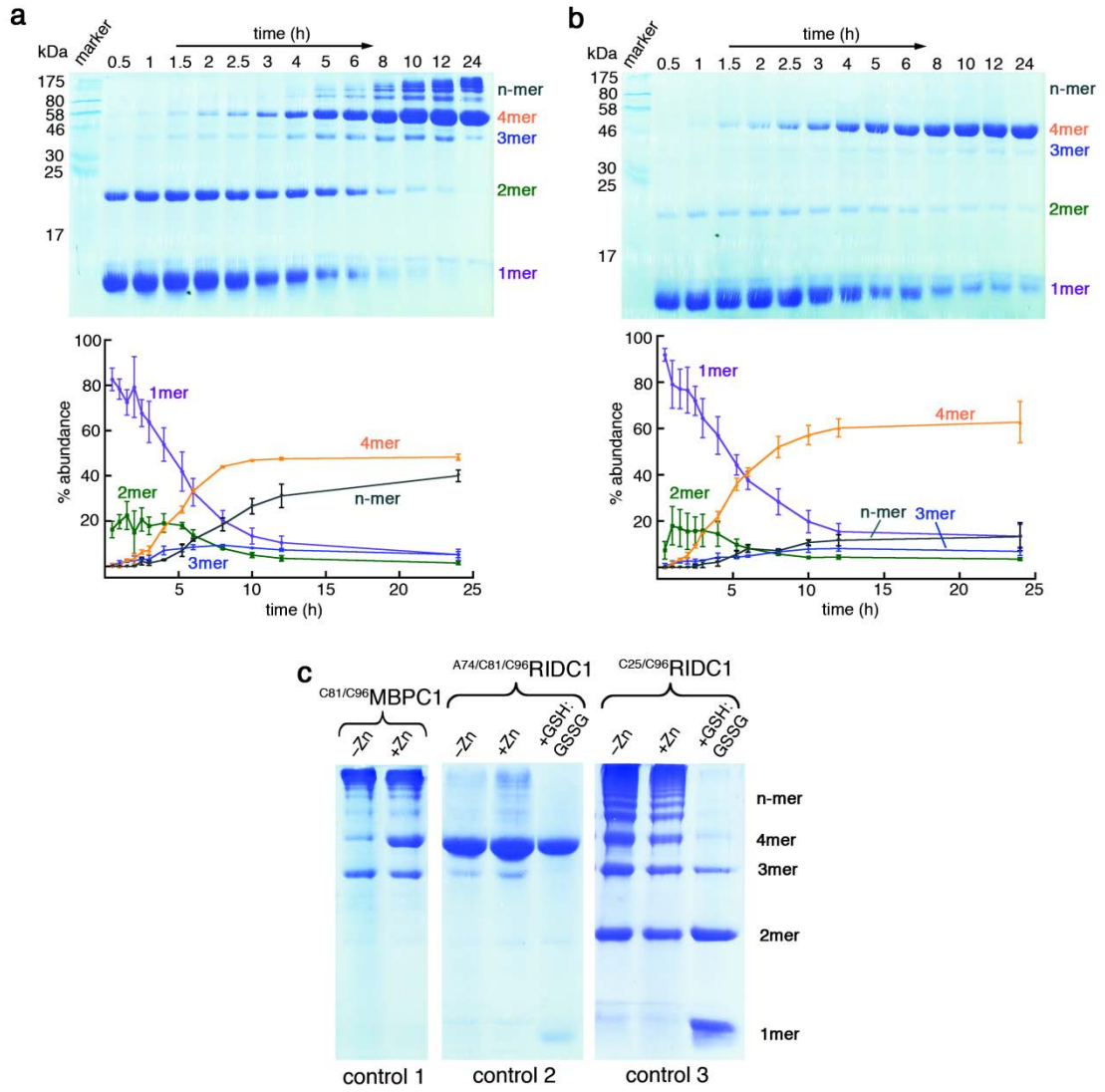

Figure 3.2. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 at room temperature. (**a**) Time-course of the oxidative self-assembly of $^{C81/C96}$RIDC1 (50 μM) at pH 7 as characterized by SDS-PAGE under non-reducing conditions. (**b**) Time-course of the disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 (50 μM) at pH 7 in the presence of 950 μM GSH and 50 μM GSSG. In (**a**) and (**b**), the upper panels represent the images of the actual gels, and the bottom panels show the relative integrated intensities of the observed bands in the gels. The data are the mean of three measurements ±1 s.d. (**c**) Disulfide-mediated self-assembly of the three control variants, $^{C81/C96}$MBPC1, $^{A74/C81/C96}$RIDC1 and $^{C25/C96}$RIDC1, in the presence or absence of Zn or GSH:GSSG redox buffer, as monitored by SDS-PAGE electrophoresis under non-reducing conditions.

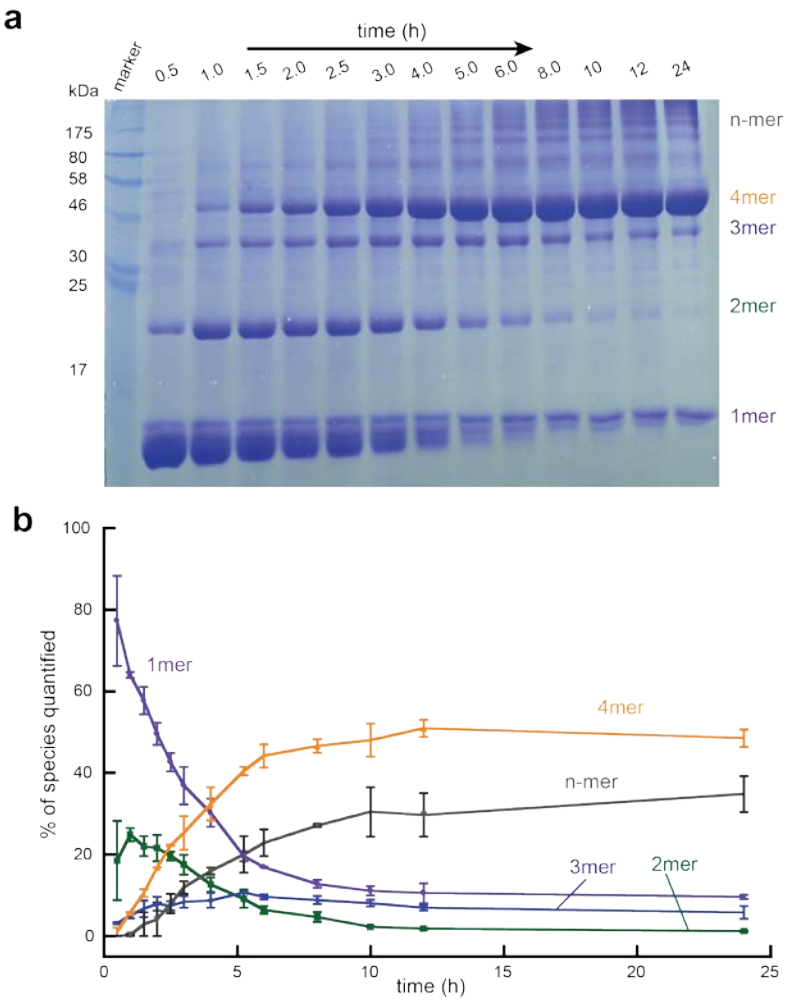

Figure 3.3. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 at 35 °C. (**a**) Time-course of the oxidative self-assembly of $^{C81/C96}$RIDC1 (50 μM) at pH 7 as characterized by SDS-PAGE electrophoresis under non-reducing conditions. (**b**) Relative integrated intensities of the observed bands in (**a**). The data are the mean of three measurements ±1 s.d.

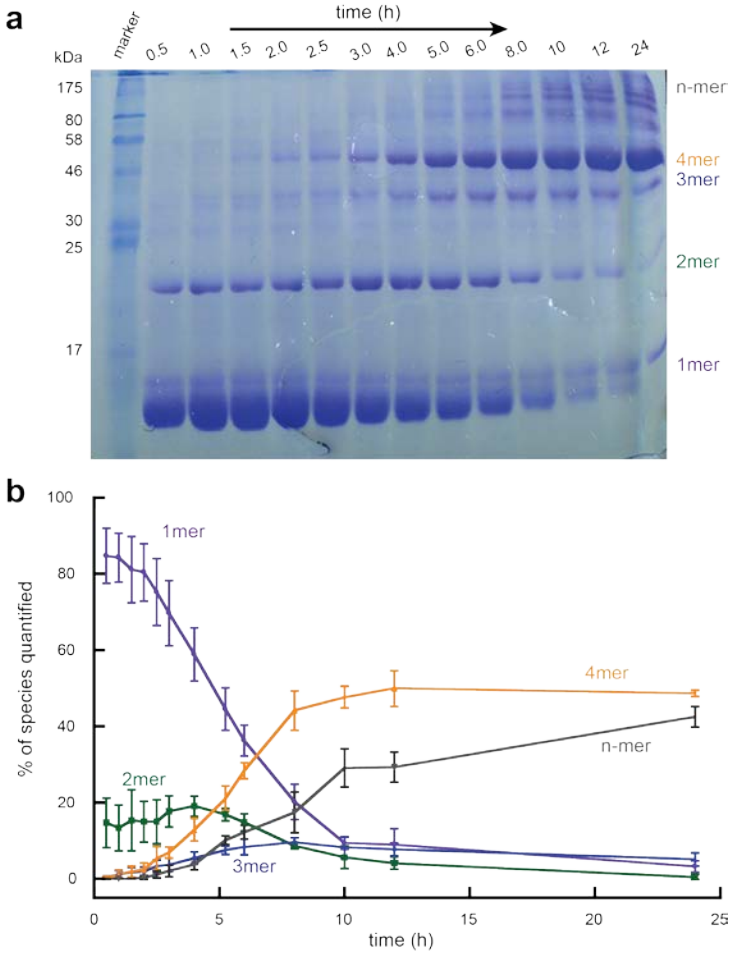

Figure 3.4. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 in the presence of Zn(II) at room temperature. (**a**) Time-course of the oxidative self-assembly of $^{C81/C96}$RIDC1 (50 μM) at pH 7 in the presence of equimolar Zn(II) as characterized by SDS-PAGE electrophoresis under non-reducing conditions. (**b**) Relative integrated intensities of the observed bands in (**a**). The data are the mean of three measurements ±1 s.d.

These initial results can be explained in terms of a proposed self-assembly model shown in Figure 3.5. According to this model, the predominant pathway proceeds through the reversible dimerization of $^{C81/C96}$RIDC1 (Species **1**) through the complementary interactions in *i1* that yield a dimer (**2**). This proposed dimer has all

four Cys residues correctly oriented for the formation of the closed tetramer (**5**), which is supported by the previously observed antiparallel arrangement of the metal-independent RIDC1$_2$ dimer.[32] The remaining fraction of monomers can form disulfide-linked dimers (detectable as such in SDS-PAGE) that are either correctly paired (**3**) to form the closed tetramer through complementary *i1* interactions or incorrectly paired (**4**), which would lead higher order aggregates (**6**) or a closed trimer (**7**).

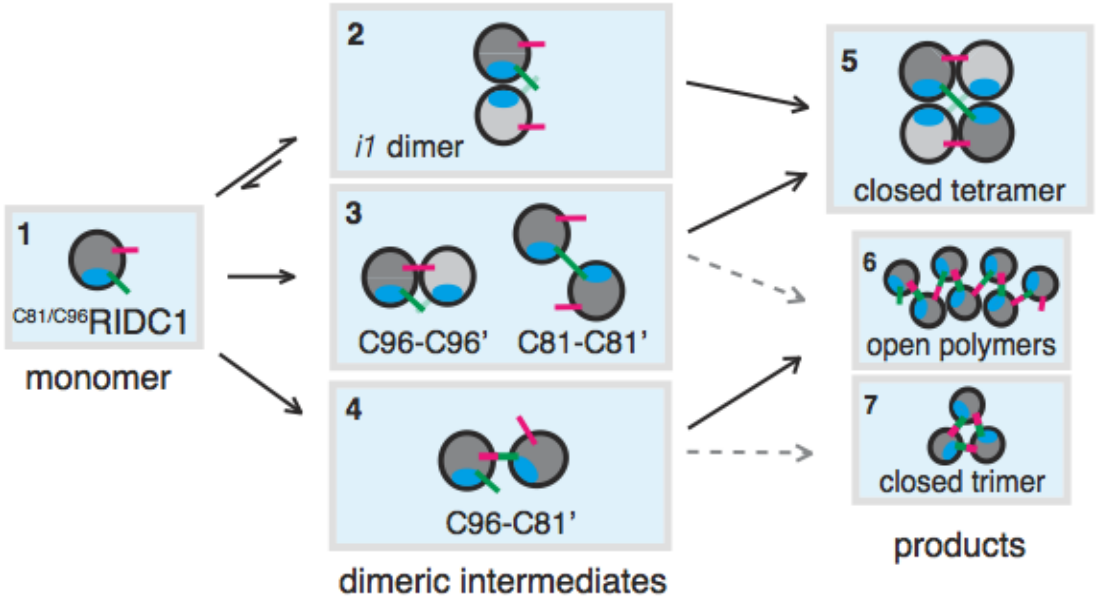

Figure 3.5. Proposed scheme for the disulfide-mediate self-assembly of $^{C81/C96}$RIDC1. The dark or light-grey-shaded circles correspond to the views of the four-helix protein building blocks from top or bottom as shown in Figure 1b. Residues C81 and C96 are illustrated as green and magenta sticks, and the hydrophobic residues in *i1* as cyan patches, matching the coloring scheme in Figure 1.

Given that disulfide bond formation is largely irreversible under these experimental conditions, the templated formation of the closed $^{C81/C96}RIDC1_4$ tetramer should be under kinetic control, whereby the non-covalent *i1* interactions geometrically favor the formation of the four, correct disulfide bonds. Under the assumption that the closed tetramer–which has all of its interfaces correctly formed–is also the thermodynamically most favored product among those shown in Figure 3.5, we surmised that its formation should be favored under conditions that promote reversible disulfide bond formation (*i.e.*, where all steps in Figure 3.5 operate under equilibrium). Toward this end, we utilized a redox buffer system containing reduced and oxidized forms of glutathione (GSH and GSSG, respectively). At a molar ratio of 19:1 GSH:GSSG (combined concentration of 1 mM, $E'_{pH7} = -187$ mV)[34], the yield for the tetramer after 1 d of oxidation indeed rose from 48% to 63±9 % (Figure 3.2b). As expected from a system under thermodynamic control, this increase in the yield compared to the non-redox-buffered system came primarily at the expense of the less stable, polymeric species, whose abundance decreased from 40% to 13±6%. Under more oxidizing conditions (1:19 GSH:GSSG at 1 mM total concentration, $E'_{pH7} = -77$ mV)[34], the product distribution did not change significantly, although the reaction proceeded more rapidly as expected (Figure 3.6a and 3.6b). Treatment with the strong reducing agent dithiothreitol (DTT, $E'_{pH7} = -330$ mV)[35] led to complete dissociation of all oligomeric products. Our results thus indicate that the disulfide bonds in $^{C81/C96}RIDC1_4$ have reduction potentials (–200 to –300 mV) typical of natural systems,

and that even small amounts of GSH are sufficient to catalyze reversible disulfide exchange and ensure thermodynamic control.

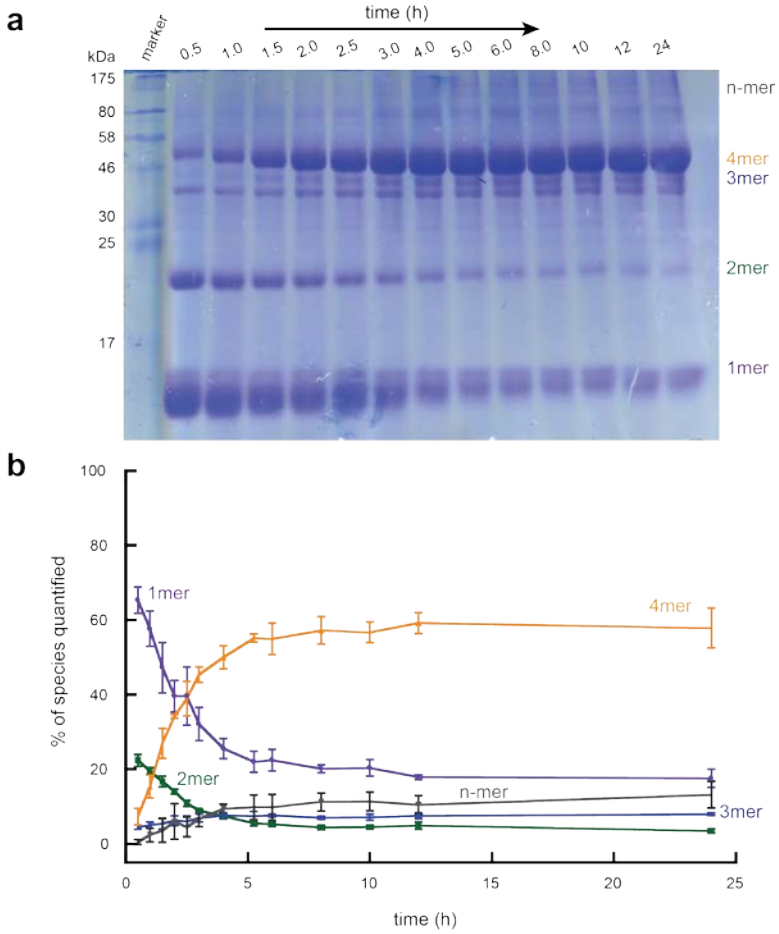

Figure 3.6. *In vitro* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 in the presence of 1:19 GSH:GSSG (1 mM total). (**a**) Time-course of the oxidative self-assembly of $^{C81/C96}$RIDC1 (50 μM) at pH 7 and room temperature as characterized by SDS-PAGE electrophoresis under non-reducing conditions. (**b**) Relative integrated intensities of the observed bands in (**a**). The data are the mean of three measurements ±1 s.d.

To further probe the validity of our model, we prepared three control variants and examined their self-assembly under oxidative conditions. The first variant, $^{C81/C96}$MBPC1, had the same two Cys residues but lacked the hydrophobic surface residues that stabilize and preorient the *i1* interface. In contrast to $^{C81/C96}$RIDC1, $^{C81/C96}$MBPC1 did not display any preference for tetramer formation: after ~16 hours of air oxidation, the main products were higher-order aggregates (~80% in abundance) with only ~7% in tetrameric form (Figure 3.2c, left panel). In this case, however, the addition of equimolar Zn(II) did yield a noticeable increase in the tetramer population (~20%), suggesting that Zn(II) coordination may help template the formation of the closed tetramer in the absence of the directing *i1* interactions.

To conclusively establish that the templated formation of the $^{C81/C96}$RIDC1$_4$ tetramer is independent of Zn(II) coordination, we prepared a second control variant, $^{A74/C81/C96}$RIDC1, which featured an alanine substituted for one of the Zn-coordinating residues, Asp74. The elimination of the Zn-Asp74 linkage was chosen because this bond is critical in the stabilization of the *i3* interface in the parent tetramer Zn$_4$:MBPC1$_4$. The yield of tetramer formation for $^{A74/C81/C96}$RIDC1 was not only unaffected in the presence or absence of Zn(II), but it was consistently higher than that of $^{C81/C96}$RIDC1 under all conditions (Figure 3.2c, middle panel). In the presence of 1 mM GSH:GSSG buffer, the tetramer abundance reached 77±3% after 16 h, with the remaining fraction mostly in the monomeric form. Clearly, the D74A mutation must either stabilize a productive self-assembly intermediate (species **2** or **3**) or destabilize

an unproductive one (species **4**), although the exact molecular details of this effect are not immediately apparent.

Our third control variant was an alternative double-Cys mutant $^{C25/C96}$RIDC1, which could form the complementary *i1* interactions, but had a Cys installed in position 25, which is not compatible for self-pairing in the closed tetramer (Figure 3.1b). In the absence of any structural bias, two Cys residues on a given protein monomer should give rise to multiple modes of disulfide-crosslinking and, in turn, the formation of open oligo- and polymers, with no particular enrichment of any intermediate n-mer. This was indeed our observation with $^{C25/C96}$RIDC1 under all conditions tested (± Zn(II), ± GSH/GGSG), which formed a "ladder" of species with no distinct population of an intermediate (Figure 3.2c, right panel). Taken together, our observations clearly show that the hydrophobic interactions engineered into *i1* are necessary and sufficient for directing the formation of the four desired disulfide bonds.

Structural and Chemical Characterization of $^{C81/C96}$RIDC1$_4$ and its variants.

The $^{C81/C96}$RIDC1$_4$ tetramer was purified in high yield by size-exclusion chromatography for further characterization (Figure 3.7a and 3.7b). Sedimentation velocity experiments with $^{C81/C96}$RIDC1$_4$ revealed a single, narrow peak with a maximum at 4.5 S, which slightly moved (to 4.6 S) upon addition of 4 equivalents of Zn^{II}, indicating the formation of a discrete, closed tetramer that does not undergo a large conformational change upon metal binding (Figure 3.8). This was confirmed through the crystal structures of apo- and Zn-bound $^{C81/C96}$RIDC1$_4$, which we determined at 2.1 and 1.2 Å resolution, respectively (Figure 3.9). An inspection of the

core architectures of both structures (Figure 3.10) reveals their striking topological resemblance to cryptands,[36] in particular spheroidal cryptands,[37] which are highly preorganized, covalently linked macropolycylic compounds designed to stably capture cations and anions in their cavities. The tetramers feature a macrocylic base structure that is formed by the linkage of Helices 4 from each monomer by the two pairs of C96-C96' and C81-C81' disulfide bonds. This covalently linked macrocycle is capped from the top and the bottom by Helices 3 from each monomer, which are linked through the hydrophobic *i1* interactions and help form a central cavity lined by the four Zn coordination sites (Figure 3.11). Like cryptands, this large cavity (740 $Å^3$ in apo- $^{C81/C96}RIDC1_4$, 1220 $Å^3$ in Zn_4:$^{C81/C96}RIDC1_4$) and the Zn coordination sites are relatively well isolated from the bulk owing to the packing of side chains in *i1* and *i2* interfaces, particularly in the case of the apo structure. Again, in analogy to cryptands, $^{C81/C96}RIDC1_4$ is highly pre-organized for Zn(II) coordination.

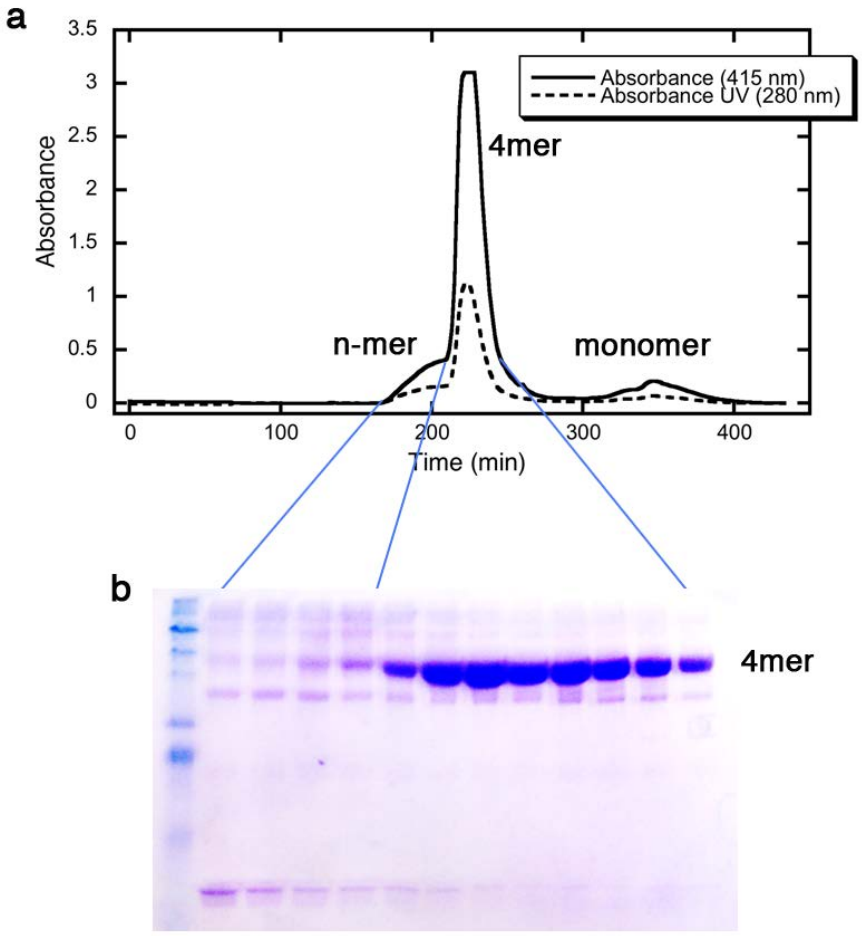

Figure 3.7. Bulk purification of $^{C81/C96}RIDC1_4$. (**a**) Size-exclusion chromatogram for the purification of $^{C81/C96}RIDC1_4$ monitored at 280 nm and 415 nm. (**b**) Non-reducing SDS-PAGE gel of the fractions corresponding to the higher-order oligomers and tetrameric species.

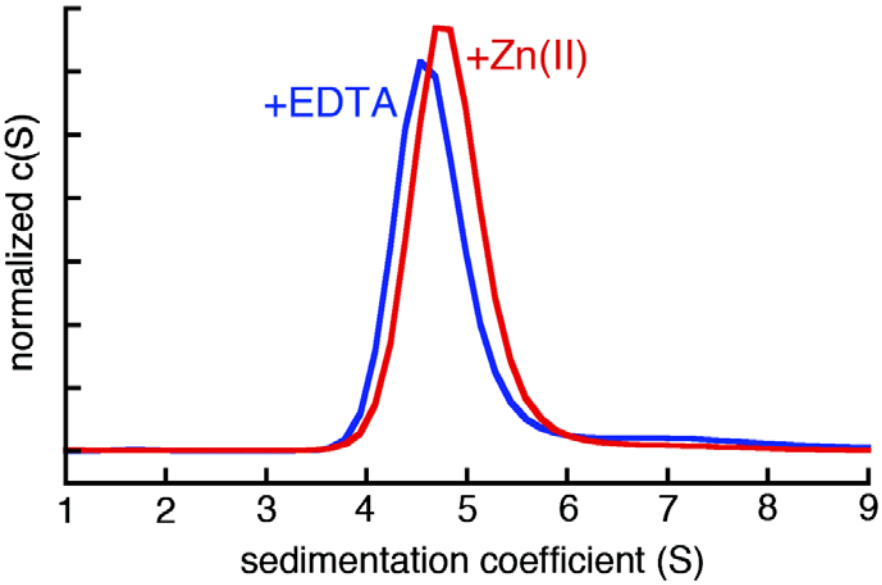

Figure 3.8. Sedimentation velocity profile of $^{C81/C96}$RIDC1$_4$ (5 µM) in the presence and absence of equimolar Zn(II). The sedimentation coefficient value increases from 4.5 S in the presence of EDTA to 4.6 S upon addition of Zn(II).

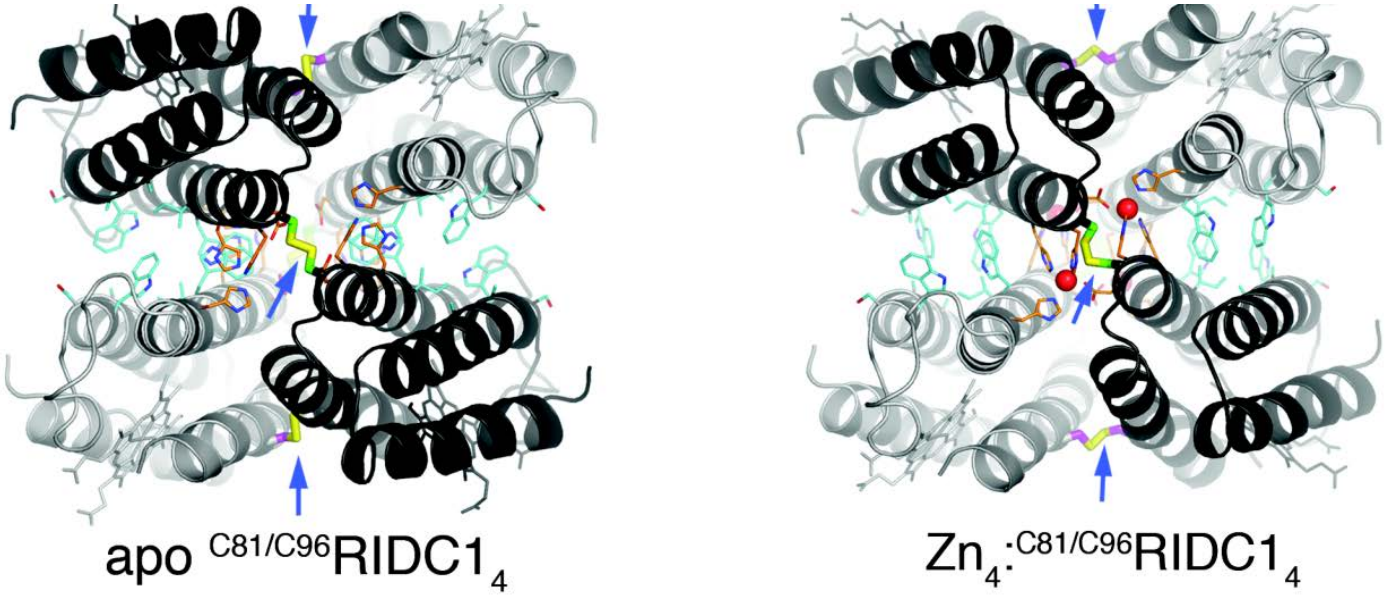

Figure 3.9. Crystal structures of apo, and Zn-bound $^{C81/C96}RIDC1_4$. The C81-81 and C96-C96 disulfide bonds are shown as green/yellow and magenta/yellow sticks, respectively. The hydrophobic residues in the *i1* interface are highlighted in cyan. Interfacial disulfide bonds, except one C81-C81 bond in the background, are indicated with blue arrows.

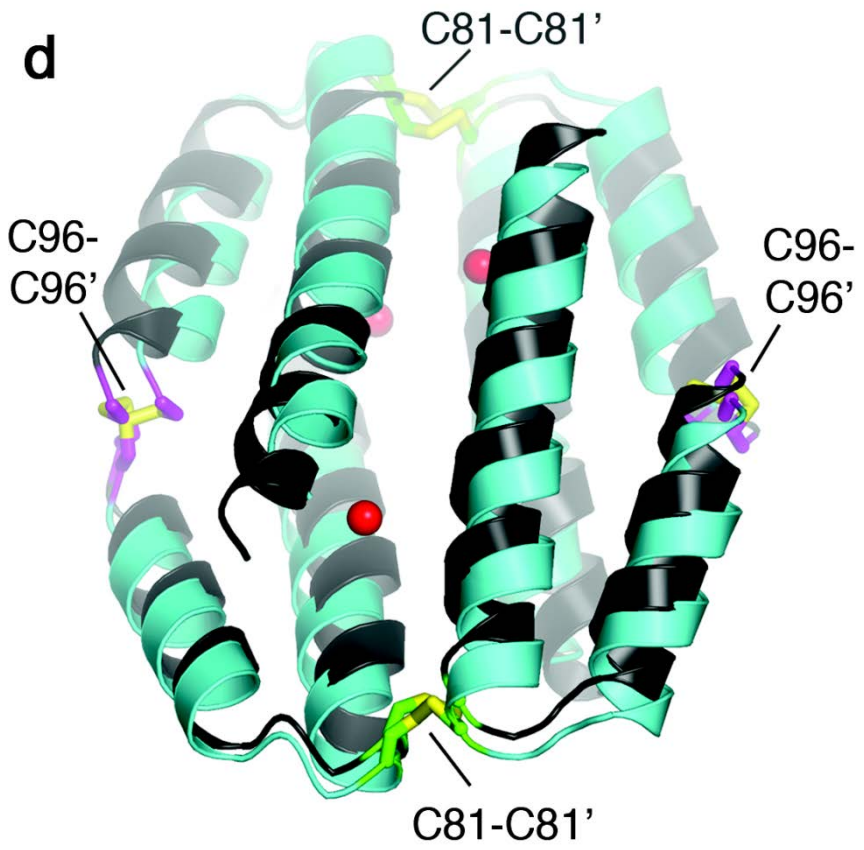

Figure 3.10. Backbone superposition of the cryptand-like cores of apo- (cyan) and Zn-bound (dark gray) $^{C81/C96}$RIDC1$_4$. The cryptand-like core is composed of residues 60-96 of each of the four protein monomers. The helical segments 81-96 from the outer, covalently linked periphery, whereas the helical segments 60-81 form part of the hydrophobic *i1* interface that cap the core cavity from top and bottom.

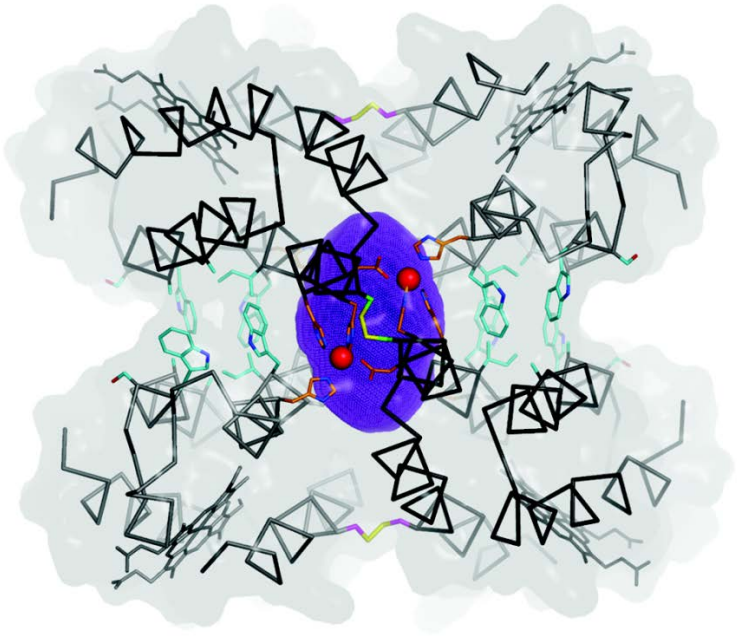

Figure 3.11. The central cavity in Zn_4:$^{C81/C96}$RIDC1$_4$ lined by the four Zn coordination sites is highlighted in magenta.

Structures of apo-$^{C81/C96}$RIDC1$_4$ and Zn_4:$^{C81/C96}$RIDC1$_4$ display a small deviation from each other, with a root-mean-square deviation (rmsd) of 2.2 Å measured over the backbone α-carbons of the cryptand-like core (Figure 3.10). The small rearrangement of the metal coordinating residues (His63/Asp74/His73/His77) upon Zn(II) binding is readily accommodated by a small twist of Helices 4 pivoted by the C81-C81' disulfide bonds. This rearrangement is accompanied by a widening of the *i1* interface in the Zn-bound structure, which is compensated by the rotation of the Trp41 and Trp66 sidechains to extend fully into the interface and preserve their

original hydrophobic contacts. The Zn(II) affinity of $^{C81/C96}$RIDC1$_4$ was determined by a competition titration using the fluorescent metal indicator Fura-2 (Figure 3.12), whose Zn(II)-complex has a dissociation constant ($K_{d,\ Zn\text{-}Fura}$) of 5.7 nM.[33] These titrations revealed that $^{C81/C96}$RIDC1$_4$ bound four Zn(II) ions as expected, and the isotherm was well fit by a 2 + 2 Zn equilibrium model (i.e., two pairs of independent sites), with $K_{d1,\ 2Zn}$= 2.6 ± 0.3 nM and $K_{d2,\ 2Zn}$= 25 ± 4 nM (Figure 3.13).

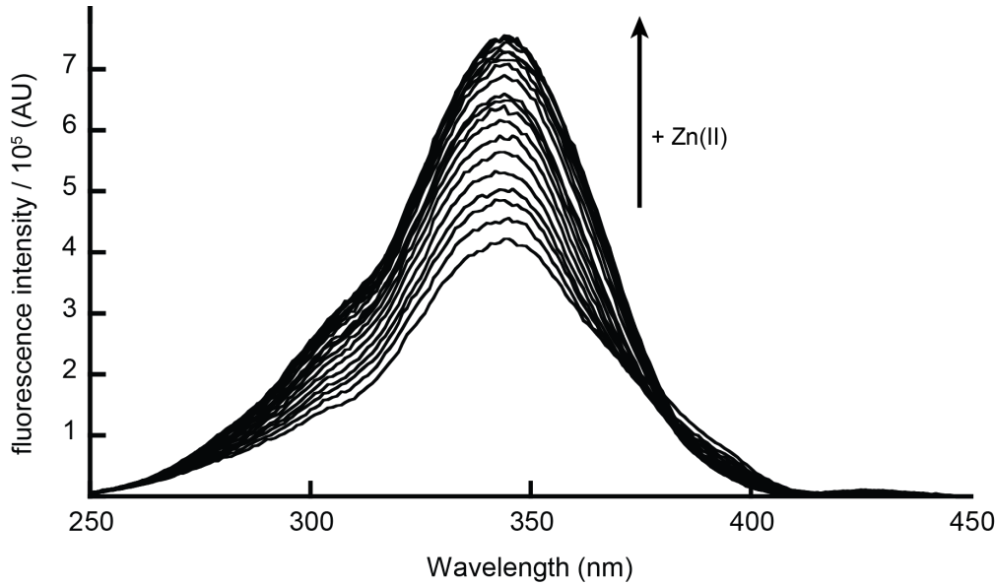

Figure 3.12. Competitive Zn binding by Fura-2 and $^{C81/C96}$RIDC1$_4$ monitored by changes in Fura-2 fluorescence excitation spectrum. The corresponding titration curve is shown in Fig. 3.13. The sample contained 4.2 μM Fura-2 and 2.75 μM $^{C81/C96}$RIDC1$_4$. The Fura-2 emission was detected at 510 nm.

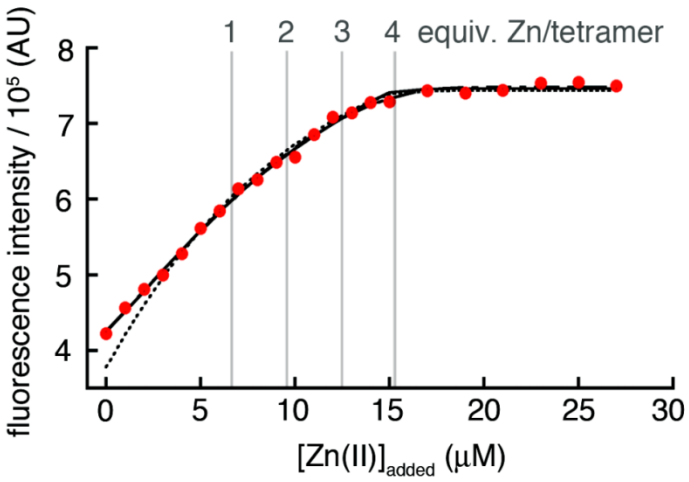

Figure 3.13. Zn-binding isotherm of $^{C81/C96}$RIDC1$_4$ (2.75 μM) determined using Fura-2 (4.2 μM) as a competing ligand. Nearly equally good fits are obtained using a four consecutive Zn binding equilibrium model (1 + 1 + 1 + 1, dashed line) or a two consecutive binding equilibrium model (2 + 2, solid line), whereas a single binding equilibrium model (4 × 1, dotted line) does not appear to be adequate.

The fact that the self-assembly of our cryptand-like scaffold was entirely uncoupled from Zn(II) coordination suggested that it could be utilized to construct unsaturated metal coordination sites. To examine this possibility, we crystallized the Zn-complex of the aforementioned $^{A74/C81/C96}$RIDC1 variant, which has the Zn-coordinating Asp74 carboxylate eliminated and can be isolated in its tetrameric form in high yield. The 2.3-Å structure of the tetramer shows that the four Zn ions are indeed coordinated only by the three original His side chains (His63, H73 and His77) and the fourth coordination site is occupied by a water molecule, closely mimicking the Zn coordination environment of carbonic anhydrases[38] and matrix metalloproteinases[39] (Figures 3.14a, 3.14b).

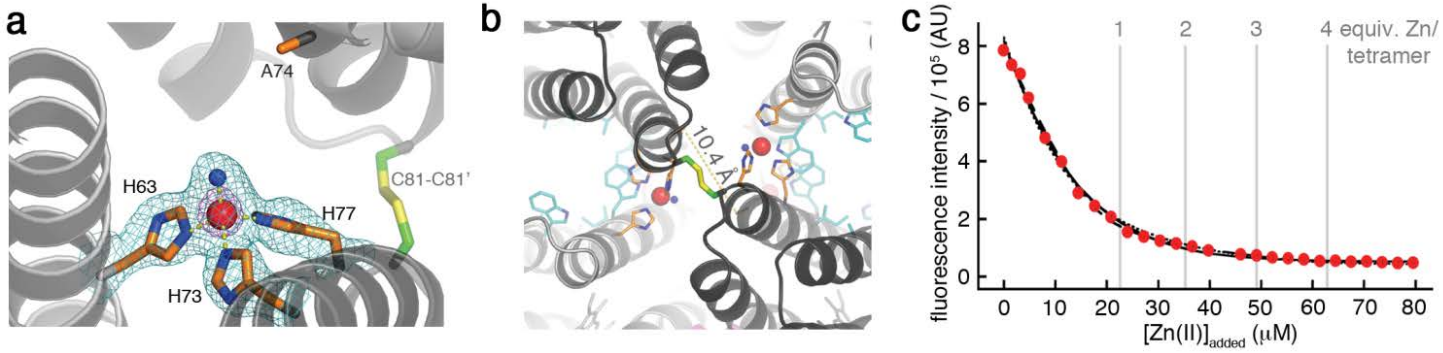

Figure 3.14. Structural features of Zn_4:$^{A74/C81/C96}$RIDC1$_4$. (**a**) Close-up view of one of the four identical 3His-Zn coordination sites, along with the corresponding $2F_o–F_c$ electron density map contoured at 1.4 σ (cyan) and 7 σ (magenta). The Zn-coordinated water molecule is shown as a blue sphere. (**b**) View of Zn_4:$^{A74/C81/C96}$RIDC1$_4$ down the C_2'' axis and the *i3* interface. The C_β-C_β distance (10.4 Å) between position 74 residues across the *i3* interface is highlighted with a dashed line. The corresponding distance in Zn_4:$^{C81/C96}$RIDC1$_4$ is 8.8 Å. (**c**) Zn-binding isotherm of $^{A74/C81/C96}$RIDC1$_4$ (13.5 μM) determined using Mag-Fura-2 (8.7 μM) as a competing ligand. Nearly equally good fits are obtained using a four consecutive Zn binding equilibrium model (1 + 1 + 1 + 1, dashed line), a two consecutive binding equilibrium model (2 + 2, solid line) model and a single binding equilibrium model (4 × 1, dotted line).

Competition titrations using Mag-Fura-2 ($K_{d,\ Zn\text{-}MagFura2}$ = 47 nM)[40] (Figures 3.14c, 3.15) could be adequately described by a 4 × 1 Zn equilibrium model (i.e., four independent, equivalent binding sites), yielding $K_{d,\ 4Zn}$ = 480 ± 35 nM. This dissociation constant is approximately one to two orders of magnitude higher than those determined for $^{C81/C96}$RIDC1$_4$, consistent with the loss of a coordinating ligand. The elimination of the D74-Zn coordination leads to a slight widening of the *i3* and subsequently the *i1* interfaces (Figure 3.14b**)**, all of which are reflected in an increased cavity volume of 2100 Å^3. Despite this potentially drastic perturbation to the inner Zn-coordination sphere, the crpytand-core of Zn$_4$:$^{A74/C81/C96}$RIDC1$_4$ is superposable onto that of Zn$_4$:$^{C81/C96}$RIDC1$_4$ with an rmsd of 1.6 Å (Figure 3.16), further highlighting the structural rigidity afforded by covalently interconnected topology.

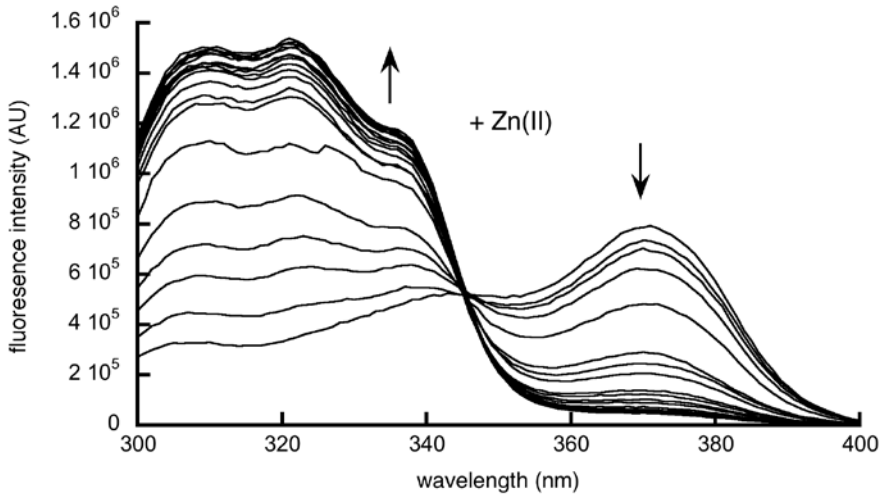

Figure 3.15. Competitive Zn binding by Mag-Fura-2 and $^{A74/C81/C96}$RIDC1$_4$ monitored by changes in Mag-Fura-2 fluorescence excitation spectrum. The corresponding titration curve is shown in Figure 3.15 The sample contained 8.7 μM Mag-Fura-2 and 13.5 μM $^{A74/C81/C96}$RIDC1$_4$. The Mag-Fura-2 emission was detected at 505 nm.

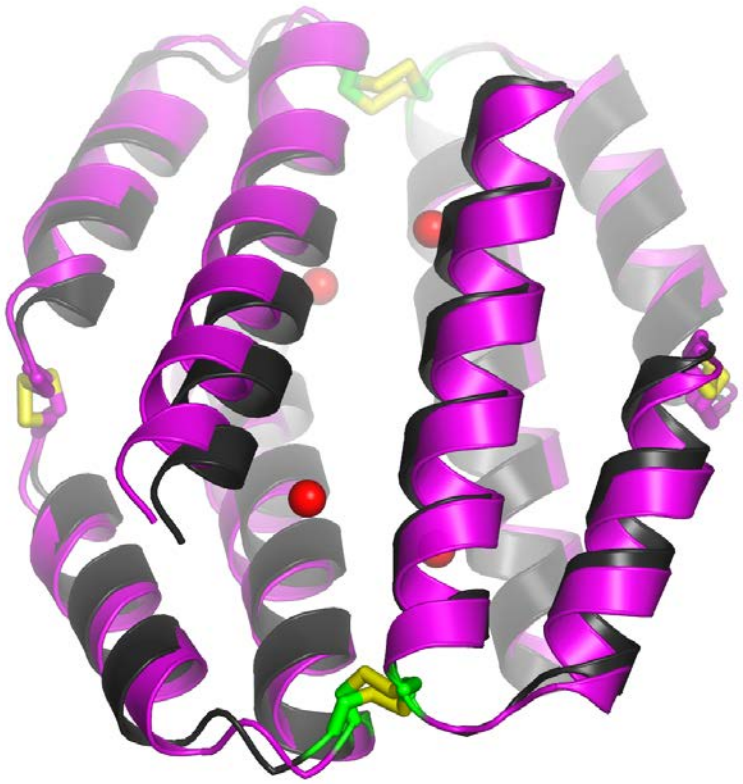

Figure 3.16. Overlay of the cryptand-like cores of Zn_4:$^{A74/C81/C96}$RIDC1$_4$ and Zn_4:$^{C81/C96}$RIDC1$_4$ crystal structures. Zn_4:$^{A74/C81/C96}$RIDC1$_4$ and Zn_4:$^{C81/C96}$RIDC1$_4$ are shown in magenta and dark grey, respectively.

In vivo self-assembly of $^{C81/C96}$RIDC1$_4$.

The gene encoding for $^{C81/C96}$RIDC1 is equipped with an engineered N-terminal signal sequence, which enables the translocation of the protein upon translation into the *E. coli* periplasm, where it is processed for the covalent (*c*-type) linking of the heme cofactor to the protein backbone.[29,41] The periplasmic space is an oxidizing environment and contains the biological machinery for disulfide bond formation and exchange.[21] This raised the possibility that $^{C81/C96}$RIDC1 might

correctly self-assemble into the cryptand-like tetramer *in vivo* and, given its relatively high Zn-affinity, compete for Zn ions in bacterial cells. To test this possibility, we performed a periplasmic extraction of the $^{C81/C96}$RIDC1-overexpressing *E. coli* cells through osmotic shock, which lyses the outer but not the inner membrane, such that any disulfide-linked products remain isolated from the cytosolic reductants. Solutions used for lysis or any other step thereafter also contained 50 mM iodoacetic acid for alkylating any free cysteines and disabling any post-lysis disulfide bond formation. Crude periplasmic lysates of $^{C81/C96}$RIDC1-overexpressing cells were immediately analyzed by non-reducing, SDS-PAGE, using 3,3',5,5'-tetramethylbenzidine (TMBZ) as the staining agent to detect heme-containing proteins[42] (Figure 3.17a) or Coomassie Blue to image the entire protein content (Figure 3.18). The TMBZ-stained gels indicated a prominent band (~63% abundance among all stained species, Figure 3.17a, left panel) near the 46-kDa molecular weight marker that would correspond to the $^{C81/C96}$RIDC1$_4$ tetramer (expected mass = 49,108 Da). When SDS-PAGE was carried out after the treatment of the extracts with DTT to reduce disulfide bonds, the main band moved to where a monomeric species would be expected (Figure 3.17a, right panel), establishing that $^{C81/C96}$RIDC1 primarily exists as a disulfide-linked tetramer in the periplasm. A similar SDS-PAGE profile was obtained in the case of the cells that overexpressed $^{A74/C81/C96}$RIDC1 (Figure 3.17a), with a tetramer abundance of ~60%. In contrast, periplasmic extracts of MBPC1 and RIDC1, which do not self-assemble into tetramers *in vitro*, consisted almost entirely of monomers. Extracts of $^{C81/C96}$MBPC1 contained a distribution of n-mers with no appreciable population of a tetrameric

species (Figure 3.17a), paralleling the findings of the self-assembly experiments carried out *in vitro* (Figure 3.1c) and reaffirming the necessity/sufficiency of hydrophobic *i1* interactions in templating the formation of $^{C81/C96}$RIDC1$_4$. It remains to be determined whether disulfide-exchange enzymes such as DsbA (E'≈ –100 mV)[43] are directly involved in the self-assembly of $^{C81/C96}$RIDC1$_4$ in the periplasm.

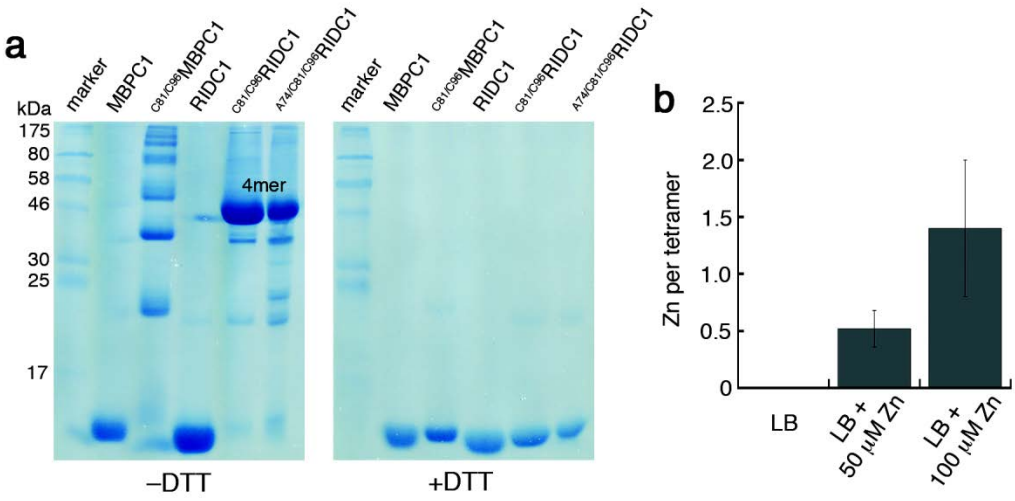

Figure 3.17. *In vivo* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 and its Zn binding properties. (**a**) SDS-PAGE analysis of the periplasmic extracts of MBPC1 (lane 2), $^{C81/C96}$MBPC1 (lane 3), RIDC1 (lane 4), $^{C81/C96}$RIDC1 (lane 5) and $^{A74/C81/C96}$RIDC1 (lane 6). Electrophoresis was carried out under non-reducing conditions using untreated extracts (left panel) or DTT-treated extracts (right panel). The gels were stained with TMBZ for heme detection. For the corresponding Coomassie-stained gels, see Supplementary Figure S8. (**b**) ICP-OES analysis of the Zn content of self-assembled $^{C81/C96}$RIDC1$_4$ tetramers isolated from the periplasmic extracts of E. coli cells grown in LB media under various conditions.

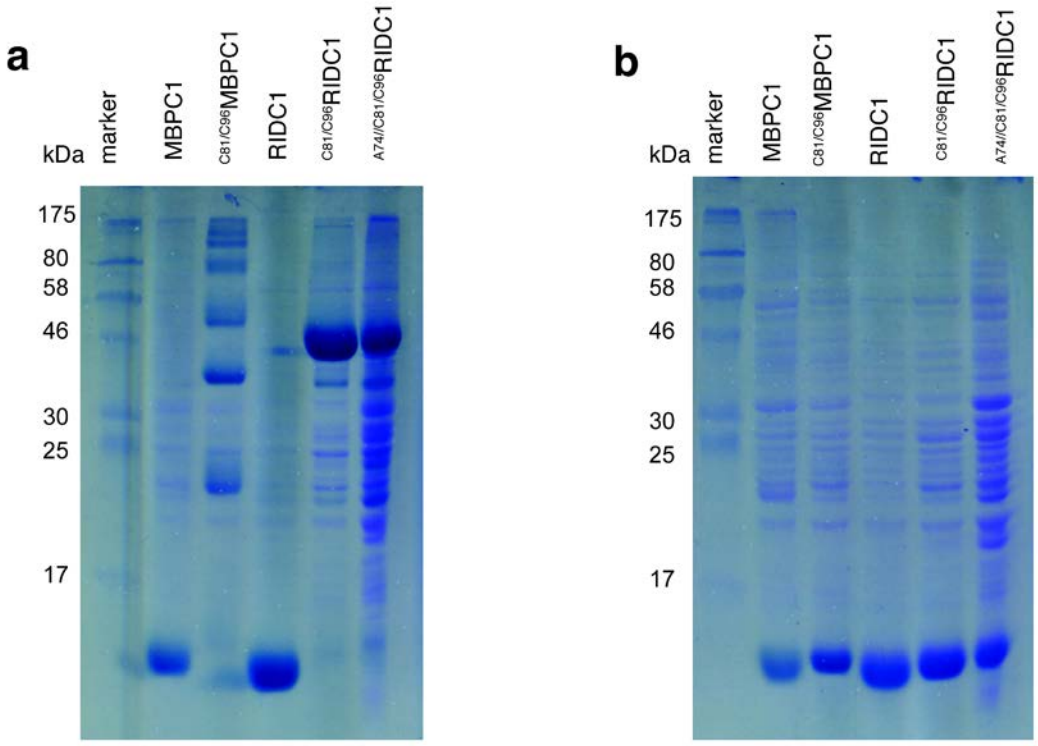

Figure 3.18. *In vivo* disulfide-mediated self-assembly of $^{C81/C96}$RIDC1 and various control variants characterized by SDS-PAGE. Electrophoresis was carried out under non-reducing conditions using untreated extracts (**a**) or DTT-treated extracts (**b**). These gels were stained with Coomassie Blue, and correspond to the gels shown in Figure 3.18, which were stained with TMBZ to only highlight heme-containing proteins.

The red-colored periplasmic extract displayed a UV-vis spectrum that was typical of reduced cyt cb_{562}, with *Q*-band maxima at 525 nm and 556 nm,[29] respectively (Figure 3.19). The species corresponding to $^{C81/C96}$RIDC1$_4$ was isolated by size exclusion chromatography. The MALDI mass spectrum displayed a single peak at 49,133 Da in the MALDI mass spectrum (Figure 3.20), slightly (25 amu) above the expected mass of the $^{C81/C96}$RIDC1$_4$ tetramer. These results establish that

$^{C81/C96}$RIDC1$_4$ correctly self-assembles in *E. coli* cells; we estimate the periplasmic concentration of $^{C81/C96}$RIDC1$_4$ to be 10-20 μM (See Appendix).

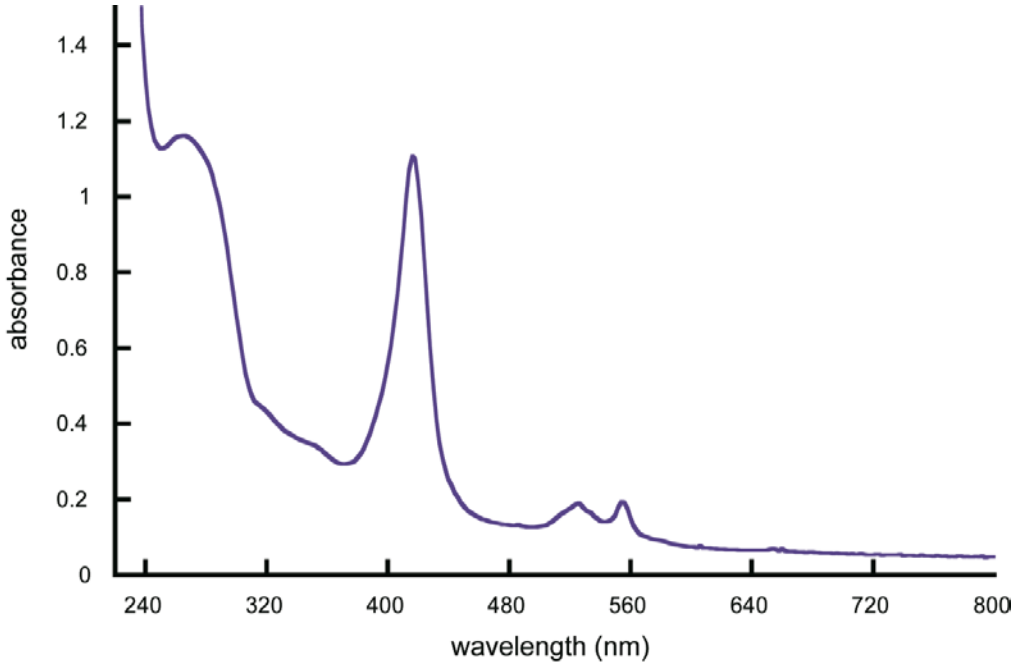

Figure 3.19. UV-vis spectrum of the periplasmic extract of $^{C81/C96}$RIDC1-overexpressing cells.

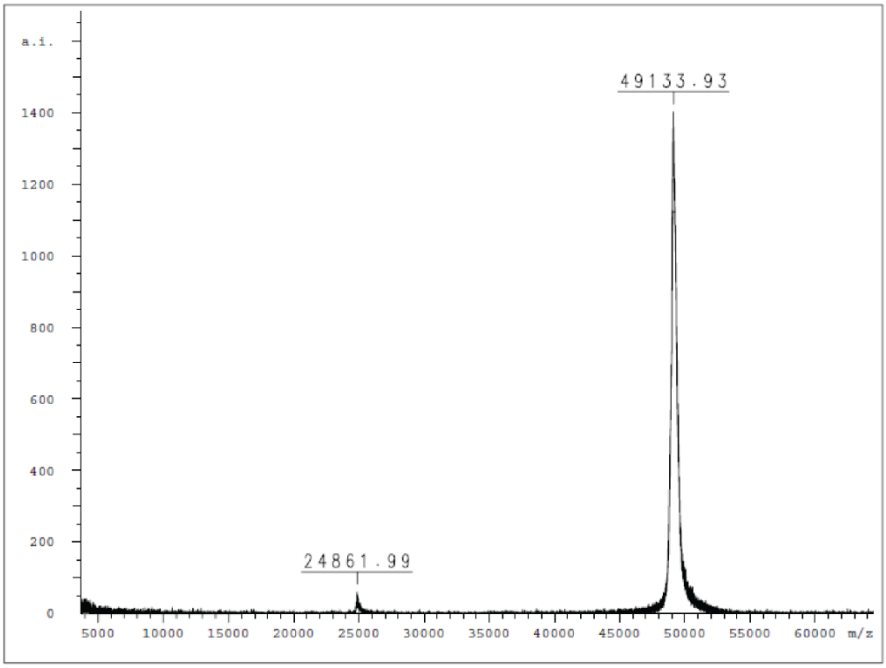

Figure 3.20. MALDI-TOF spectrum of $^{C81/C96}$RIDC1$_4$ extracted from the periplasm. The expected mass of $^{C81/C96}$RIDC1$_4$ is 49108 Da.

We next tested the ability of $^{C81/C96}$RIDC1$_4$ to compete for Zn(II) ions *in vivo*. In terms of metal homeostasis, the periplasmic space of Gram-negative bacteria serves as a buffer zone that screens the cytoplasm from large environmental fluctuations in metal ion concentrations as part of the so-called envelope stress response.[44] Though not to the extent seen in the cytoplasm, competition for Zn(II) in the periplasm is still keen. The periplasmic free Zn(II) concentration is tightly regulated by transport systems such ZnuABC ($K_{d,\ Zn\text{-}ZnuA}$ <20 nM),[45] and ZIP proteins,[46] and the periplasm contains many essential Zn-containing proteins (~20 known), including β-lactamase and superoxide dismutase.[47] In a typical experiment, $^{C81/C96}$RIDC1-expressing *E. coli* cells were grown in Luria-Bertani (LB) broth, the periplasmic $^{C81/C96}$RIDC1$_4$ tetramers

were isolated using strictly metal-free buffer solutions and glassware, and their zinc and iron contents were determined via ICP-OES. Here, the iron content served as a convenient indicator of protein concentration, as each $^{C81/C96}$RIDC1 monomer contains one iron atom in the covalently bound heme cofactor. When the cells were grown in unsupplemented LB medium ($[Zn]_{LB} \approx 10$ μM),[48,49] the isolated $^{C81/C96}$RIDC1$_4$ tetramers contained only 0.04 equivalents of zinc (Figure 3.17b). However, when the medium was supplemented with 50 or 100 μM $ZnCl_2$, the tetramers accumulated 0.52 ± 0.16 or 1.4 ± 0.6 Zn equivalents, respectively (Figure 3.17b). Under the same conditions, the non-self-assembling variants MBPC1 and RIDC1 did not contain any detectable Zn, indicating that Zn sequestration necessitates the formation of the closed tetramer. Competition by natural Zn-proteins and active regulation of the periplasmic Zn concentration likely prevent all $^{C81/C96}$RIDC1$_4$ tetramers from attaining stoichiometric Zn(II) binding. A more definitive picture would require knowledge of the concentrations and Zn-affinities of each Zn-binding protein in the periplasm alongside $^{C81/C96}$RIDC1$_4$. Nevertheless, our results indicate that $^{C81/C96}$RIDC1$_4$ is able to successfully vie for Zn(II) *in vivo* and sequester it from the environment.

Conclusions

Despite their ubiquity in natural protein assemblies and the structural and functional advantages they confer, disulfide bonds are rarely utilized in *de novo* protein design. Here, we have described a protein engineering strategy inspired by supramolecular chemistry, in which non-covalently interacting protein surfaces template the correct formation of multiple disulfide bonds to yield a novel, cryptand-like protein assembly. In designing $^{C81/C96}$RIDC1$_4$, we exploited the D_2 symmetry of the Zn-templated tetrameric parent architecture (Zn$_4$:MBPC1$_4$), which presented three sets of interconnected interfaces with C_2 symmetry that facilitated the pairing of multiple Cys residues. Indeed, D_2-symmetric tetramers are significantly overrepresented among all known tetrameric protein structures, in which the connectivity of the interfaces leads to more compact/stable structures, and allows facile incorporation of function or allostery.[50,51]

$^{C81/C96}$RIDC1$_4$ is noteworthy from several viewpoints. It represents the first multi-disulfide protein architecture engineered in a bottom-up fashion. Owing to its highly interconnected topology, $^{C81/C96}$RIDC1$_4$ is well pre-organized for metal binding in its interior, and possesses a large, protected cavity in its center. Both of these structural features are prominent characteristics of natural metalloenzymes, but difficult to achieve in *de novo* designed systems.[52] In fact, most successful metalloprotein/enzyme design approaches largely bypass the protein folding problem by utilizing either existing protein folds[53-55] or canonical folding motifs (e.g., α-helical coiled coils)[56,57] whose interiors are engineered/repurposed for constructing metal

binding sites. Although our building block (cyt cb_{562}) also is a folded protein, $^{C81/C96}$RIDC1$_4$ is distinct from previous designs in that the internal metal binding sites and the surrounding cavity are built entirely on the outside surface of this building block. Notably, cyt cb_{562} (a putative electron transfer protein) and $^{C81/C96}$RIDC1 (a protein that self-assembles into a metalloenzyme-like supramolecular architecture) are separated by only 11 surface point mutations; *i.e.*, they share ~90% sequence identity. This feature recapitulates, in a synthetic system, how nature can take advantage of a limited number of protein domains and folding motifs to create entirely new structures and functions.[58] Ultimately, $^{C81/C96}$RIDC1$_4$ is a synthetic (designed) supramolecular complex, yet it consists entirely natural components, and is able to correctly self-assemble *in vivo* and incorporate the desired metal ion in a highly competitive environment. This raises the tantalizing prospect of constructing novel metalloprotein scaffolds that can be subjected to selective pressure for the evolution of new biological functions.

Table 3.1. Primers for site-directed mutagenesis

Variant(s)	**Mutation**	**Primer Sequence (5'-3')**
$^{C81/C96}$RIDC1	E81C	For: CTGGCAAATTGCGGTAAAGTAAAAG
		Rev: CTTTTACTTTACCGCAATTTGCCAG
$^{A74/C81/C96}$RIDC1	D74A	For: CAGATTCACGCGGCGCTGCACCTG
		Rev: CAGGTGCAGCGCCGCGTGAATCTG
$^{C25/C96}$MBPC1 or $^{C25/C96}$RIDC1	Q25C	For: GGATAACGCGGCGTGCGTCAAAGAC Rev: CGTCTTTGACGCACGCCGCGTTATCC

Table 3.2. Crystallization conditions of mutant variants at 25 ˚C

Protein Crystal	**Protein**	**Precipitant**
apo $^{C81/C96}$RIDC1$_4$	2.1 mM in 20 mM MOPS (pH 7)	100 mM Tris pH 8.5, 30% methyl-pentanediol, 200 mM NaCl
Zn_4:$^{C81/C96}$RIDC1$_4$	2.1 mM in 20 mM MOPS (pH 7) with 5 mM Zn(II)	100 mM Tris pH 8.5, 35% methyl-pentanediol, 200 mM $(NH_4)_2SO_4$
Zn_4:$^{A74/C81/C96}$RIDC1$_4$	1.5 mM in 20 mM MOPS (pH 7) with 4 mM Zn(II)	100 mM HEPES pH 7.5, 35% methyl-pentanediol, 200 mM NaCl

Table 3.3. X-Ray data collection and refinement statistics. *Highest resolution shell is shown in parentheses.

	apo $^{C81/C96}$RIDC1$_4$	Zn$_4$:$^{C81/C96}$RIDC1$_4$	Zn$_4$:$^{A74/A59/C81/C96}$RIDC1$_4$
Data collection			
Space group	$C222_1$	$P12_11$	$P3_112$
Cell dimensions			
a, *b*, *c* (Å)	56.17, 87.79, 96.05	48.85, 64.79, 74.18	54.94, 54.94, 145.66
α, β, γ (˚)	90.00, 90.00, 90.00	90.00, 98.18, 90.00	90.00, 90.00, 120.00
Resolution (Å)	48.0-2.12	56.8-1.15	48.55-2.16
R_{sym} or R_{merge} (%)*	4.0 (49.9)	4.6 (56.8)	9.4 (47.1)
$I/\sigma I$*	17.6 (2.5)	10.9 (1.2)	6.5 (1.6)
Completeness (%)*	97.5 (84.4)	94.4 (83.0)	97.2 (99.0)
Redundancy	5.6 (4.8)	2.9 (2.6)	5.2 (5.7)
Refinement			
Resolution (Å)	48.02-2.12	31.94-1.22	48.55-2.30
No. unique reflections	13425	129615	13468
R_{work}/R_{free}	0.23/0.30	0.13/0.18	0.24/0.31
No. atoms			
Protein	1732	4094	1753
Ligand/ion	0	4 Zn	3 Zn
Water	54	582	53
B-factors			
Protein	59.9	54.125	62.5
Ligand/ion	---	11.465	51.4
Water	55.7	40.3	59.0
R.m.s. deviations			
Bond lengths (Å)	0.016	0.030	0.017
Bond angles (˚)	1.676	3.948	1.598

Materials and Methods

Site-Directed Mutagenesis.

Plasmid DNA's coding for the parent constructs ^{C96}RIDC1 and MBPC1 were available in our laboratory and have been previously reported.[30,33] The variants below were prepared employing Quickchange mutagenesis (Stratagene) with primers obtained from Integrated DNA Technologies (see table 3.1 for primers). The mutant plasmids were then transformed into XL-1 Blue *E. coli* cells and purified using the QIAprep Spin Miniprep kit (Qiagen). Sequencing of mutant plasmids was performed by Retrogen.

Expression and purification of mutants.

Variant DNA plasmids were transformed into BL21(DE3) *E. coli* cells containing the *ccm* heme maturation gene cassette plasmid (pEC86)[41] and plated onto Luria-Bertani (LB) agar plates containing antibiotics (34 μg/mL of chloramphenicol, 60 μg/mL of ampicillin) for 16 hours at 37 ˚C. Starter cultures were grown from individual colonies in 3 mL of LB broth and antibiotics in a rotary shaker at 250 rpm for 4-6 hours to an OD_{600} of 0.6-0.8. Later, 125 μL of starter culture was added for every liter of LB/antibiotic culture medium. Typically, 12-18 L of cell cultures were grown overnight for 16 hours. Cells were pelleted by centrifugation (4,000 rpm, 4 ˚C for 10 minutes) and resuspended in 100 mL of 50 mM sodium acetate (pH 5). Cells were sonicated for 12 minutes in an ice bath in the presence of lysozyme. The lysate was brought to pH 5 using acetic acid and centrifuged (16,000 rpm, 4 ˚C for 10 minutes); the red-colored supernatant contained the protein. The supernatant was

diluted to 2 L with 10 mM sodium acetate (pH 5) and was loaded onto a CM sepharose column (Sigma-Aldrich) using a running buffer solution of 10 mM sodium acetate (pH 5). The protein was eluted with a NaCl gradient (0-0.5 M) and concentrated using an Amicon stirred cell and a 10-kDa cutoff membrane. The protein solution was exchanged into a buffer solution of 10 mM sodium phosphate (pH 8), and further purification was performed by fast protein liquid chromatography (FPLC) on a DuoFlow workstation (Bio-Rad), using a Macroprep Q cartridge (BioRad). The protein was eluted using a NaCl gradient (0-0.5 M) in a buffer solution containing 10 mM sodium phosphate (pH 8) and 2 mM dithiothreitol (DTT).

Analytical ultracentrifugation.

Samples for sedimentation velocity experiments contained 1.25 μM $^{C81,C96}RIDC_4$, 20 mM MOPS (pH 7), 150 mM NaCl and either 5 μM Zn(II) or 1 mM EDTA. Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor at 41,000 rpm for 400 scans per sample at 25 ˚C by detection at 415 nm. The data were processed using Sedfit.[59] Buffer viscosity (0.01002 poise), density (1.007 g/mL) and protein partial specific volume (0.7316 mL/g) were calculated at 25 ˚C with SEDNTERP.

Metal content analysis by Inductively Coupled Plasma- Optical Emission Spectrometry (ICP-OES).

Concentrated protein solutions extracted from the periplasm (0.5 mL total volume) were run on an analytical size exclusion column (Econopack, BioRad, 25 × 1 cm) packed with Ultrogel Aca54 resin (Pall Corporation) using a running buffer

solution of 20 mM HEPES (pH 7). The eluted protein was diluted to 2 mL by addition of Chelex-treated water and 69% HNO_3 to a final concentration of 3% HNO_3. ICP-OES standard metal solutions were prepared from 1,000-ppm certified metal stock solutions (Ricca) by serial dilutions, and a standard curve was prepared within a range of 0.1-10 ppm of metals in 3% nitric acid and Chelex-treated water. Data were collected on a Perkin-Elmer Optima 3000 DV ICP-OES spectrometer located at the Analytical Facility of the Scripps Institution of Oceanography. Because each protein monomer contains one Fe atom as part of the covalently linked heme molecule, the experimentally determined Zn:Fe ratios directly yielded the Zn:monomeric protein ratios. Wavelengths used for the detection of metal ions were Fe (234.349, 238.204, 239.562, 259.939, and 273.055 nm), Zn (202.548, 206.2, and 213.857 nm). All values reported for each metal are the averages of indicated wavelengths from 3 separate experiments.

X-ray crystallography.

All crystals were obtained by sitting-drop vapor diffusion; detailed growth conditions were as follows in Table 3.2. The sitting drops each consisted of 2 μL protein solution and 1 μL precipitant solution. The wells contained 500 μL precipitant solution. Crystals grew at room temperature typically within a two-day period and matured over a few weeks. X-ray diffraction data were collected at 100 K at beamline 12-2 of the Stanford Synchrotron Radiation Lightsource (SSRL) using 0.98-Å or 1.00-Å radiation. Diffraction data were processed using MOSFLM and SCALA on the CCP4 suite.[60] Structures were solved with molecular replacement by MOLREP[61]

using the a single chain of the structure of the ^{C96}RIDC1 (PDB ID:2IQ6) as a search model. Rigid-body, positional, thermal and TLS refinement was carried out using REFMAC,[62] along with iterative manual model building and water/ligand/metal placement with COOT.[63] For the apo-$^{C81/C96}$RIDC1$_4$ and Zn_4:$^{A74/C81/C96}$RIDC1$_4$ structures, non-crystallographic symmetry (NCS) restraints (tight main-chain and medium side-chain restraints) were applied throughout the positional/thermal refinement process. For the Zn_4:$^{C81/C96}$RIDC1$_4$ structure, these restraints, as well as TLS refinement, were lifted in the final stages of refinement. For the same structure, anisotropic B-factor refinement was carried out. All structural figures and movies were produced with PYMOL.[64] Ramachandran plots were determined using PROCHECK.[65].

Disulfide-mediated $^{C81/C96}$RIDC1 self-assembly.

Upon purification, $^{C81/C96}$RIDC1 was concentrated to 2-4 mM in an Amicon stirred cell using a 10-kDa cutoff membrane. To the $^{C81/C96}$RIDC1 stock solution was added an excess (100 mM) of dithiothreitol (DTT) to dissociate any disulfide bonds. After a 30-min incubation, the protein solution was eluted through a 10DG desalting column (BioRad) to remove DTT, using a running buffer solution of 50 mM TRIS (pH 7) and 150 mM NaCl. The protein concentration was adjusted to 50 μM using the same buffer solution. Immediately, several 20-μL aliquots of the protein solution were transferred to 0.65-mL centrifuge tubes, and to these aliquots were added either 1 molar equivalent of $ZnCl_2$, 1 mM 1:19 GSH:GSSG, 1 mM 19:1 GSH:GSSG or nothing. Samples were incubated under air exposure, either at room temperature or in

an Eppendorf thermocycler at 35 ˚C. 1-μL aliquots from each sample were taken at desired time intervals over a period of 24 h; any remaining free cysteines in these solutions were passivated by addition of 50 mM iodoacetic acid and non-reducing SDS-PAGE loading buffer (bromophenol blue, 10% SDS, glycerol, 1.5 M Tris, pH 6.8). The aliquots were applied to 12% SDS-PAGE gels and run for 40 minutes at 400 V in a BioRad MiniProtean TetraCell box. Gels were stained with Coomassie Blue for 10 minutes and then destained for 1-2 hours in a 10% nitric acid/30% methanol solution. They were then imaged using a scanner, and protein band intensities were integrated using ImageJ.[66]

Preparation and isolation of tetrameric $^{C81/C96}$RIDC1$_4$ and $^{A74/C81/C96}$RIDC1$_4$.

As described above, purified and concentrated $^{C81/C96}$RIDC1 (2-4 mM) was first treated with DTT to reduce any preformed disulfide bonds. After removal of DTT, to the protein stock solution was added 1 mM of 1:19 GSH:GSSG and incubated at 35 ˚C for 16-20 hours to form the $^{C81/C96}$RIDC1$_4$ tetramer in high yield. The tetramer was isolated from other oligomeric species by fast protein liquid chromatography on a DuoFlow workstation (Bio-Rad), using a XK26/100 column (GE Healthcare) packed with Superdex 75 (GE Healthcare) and a running buffer solution of 20 mM TRIS (pH 7) and 150 mM NaCl. The tetramer fraction was identified by SDS-PAGE under non-reducing conditions. The tetramer was concentrated, and excess EDTA (>10 mM) was added to remove any metal ions bound to the protein during purification. The protein stock solution was exchanged into a Chelex-treated buffer solution of 20 mM

MOPS (pH 7) and 150 mM NaCl, flash-frozen in 200-μL aliquots and kept at -80 ˚C until further use.

Periplasmic protein extraction and heme staining.

A 1-L cell culture expressing the protein variant of interest was harvested after 16 hours of growth by centrifugation, and the cell pellet (~7 g) was washed three times with a resuspension buffer (10 mM TRIS, pH 7.4, 33 mM NaCl). Cold osmotic shock was performed by first suspending the cell pellet in 10 volumes of the resuspension buffer (70 mL), followed by the addition of 10 volumes of 40 % sucrose (70 mL) and excess iodoacetic acid (>50 mM). After incubation in an ice bath for 10 minutes, the cell suspension was centrifuged at 6,000 rpm, and the cell pellet was resuspended with 20 volumes (140 mL) of ice-cold solution of 10 mM TRIS (pH 7.4), 0.5 mM $MgCl_2$ and 50 mM iodoacetic acid. After another 10 min incubation, the cell suspension was once again centrifuged (20 minutes, 8,000 rpm), whereby the supernatant contained the red-colored periplasmic extract. The supernatant was concentrated using an Amicon stirred cell, and 15 μL of the concentrated sample was run on a non-reducing, 12% SDS-PAGE gel for 40 minutes at 400 mV. The gel was immersed in a solution of 30 mL of 6.3 mM tetramethylbenzidine (TMBZ) and 70 mL of 0.25 M sodium acetate (pH 5) in the dark for 1-2 h, after which 30 mM hydrogen peroxide was added to the solution. The gel was incubated for another 30 minutes to allow full development of heme-containing bands, which were imaged on a scanner and integrated using ImageJ.[66]

Competitive Zn(II) binding titrations.

A 1-mL solution containing 4 μM Fura-2 (Life Technologies) and 2-3 μM $^{C81/C96}RIDC_4$ was titrated in a 1-cm cuvette with 1-2 μL aliquots of 1 mM $ZnCl_2$. After each addition, the sample was equilibrated for at least 5 minutes at room temperature, upon which the excitation scan was recorded on a Horiba Fluorolog 2 fluorimeter using an excitation range of 250-450 nm and detection at 510 nm. The change in excitation intensity at 335 nm was plotted versus Zn(II) concentration, and the resulting curve was fit alternatively to 1×4 Zn, 2×2 Zn or 4×1 Zn binding models using Dynafit[67] as described previously.[33] A similar protocol was used in the case of $^{A74/C81/C96}RIDC1_4$ (13.5 μM) in competition titration experiments with Mag-Fura-2 (Life Technologies) (8.7 μM, excitation range of 300-400 nm, emission detection at 505 nm). The change in excitation intensity at 370 nm was plotted versus Zn(II) concentration, and the resulting curve was fit alternatively to 1×4 Zn, 2×2 Zn or 4×1 Zn binding models by Dynafit.

Measurement of the internal cavity within Zn_4:$^{C81/C96}RIDC1_4$.

The volumes and surface areas of the internal cavities of the apo $^{C81/C96}RIDC1_4$, Zn_4: $^{C81/C96}RIDC1_4$ and Zn_4:$^{A74/C81/C96}RIDC1_4$ structures were measured using the CASTp[68] web server with a probe radius of 2 Å for apo $^{C81/C96}RIDC1_4$ and Zn_4:$^{C81/C96}RIDC1_4$, and a probe radius of 3 Å for Zn_4:$^{A74/C81/C96}RIDC1_4$. For visualization of the cavities, the Voss Volume Voxelator (3V)[69] web server's Solvent Extractor Program was used with an inner probe radius of 3 Å and outer probe radius of 10 Å.

Chapter 3 was reproduced in part with permission from Medina-Morales, A. M., Perez, A., Brodin, J. D., Tezcan, F. A. **2013**. In Vitro and Cellular Self-Assembly of a Protein Cryptand via Templated Disulfide Bonds. *J. Am. Chem. Soc.* The dissertation author was the primary investigator and author of this paper.

References

1 Busch, D. H. Structural definition of chemical templates and the prediction of new and unusual materials. *Journal of Inclusion Phenomena and Molecular Recognition in Chemistry* **12**, 389-39 (1992).

2 Anderson, S., Anderson, H. L. & Sanders, J. K. M. Expanding roles for templates in synthesis. *Accounts of Chemical Research* **26**, 469-475 (1993).

3 McMurry, T. J., Raymond, K. N. & Smith, P. H. Molecular Recognition and Metal-Ion Template Synthesis. *Science* **244**, 938-943 (1989).

4 Furlan, R. L. E., Otto, S. & Sanders, J. K. M. Supramolecular templating in thermodynamically controlled synthesis. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 4801-4804 (2002).

5 Thompson, M. C. & Busch, D. H. Reactions of Coordinated Ligands .9. Utilization of Template Hypothesis to Synthesize Macrocyclic Ligands in Situ. *Journal of the American Chemical Society* **86**, 3651-3656 (1964).

6 Day, V. W., Marks, T. J. & Wachter, W. A. Large metal ion-centered template reactions. Uranyl complex of cyclopentakis(2-iminoisoindoline). *Journal of the American Chemical Society* **97**, 4519-4527 (1975).

7 Byrne, G. T., Linstead, R. P. & Lowe, A. R. Phthalocyanines part II The preparation of phthalocyanine and some metallic derivatives from o-cyanobenzamide and phthalimide. *Journal of the Chemical Society*, 1017-1022 (1934).

8 McMurry, T. J., Rodgers, S. J. & Raymond, K. N. Ferric ion sequestering agents. 16. Template and stepwise syntheses of a macrobicyclic catechoylamide ferric ion sequestering agent. *Journal of the American Chemical Society* **109**, 3451-3453 (1987).

9 Creaser, I. I. *et al.* Synthesis and Reactivity of Aza-Capped Encapsulated Co(III) Ions. *Journal of the American Chemical Society* **104**, 6016-6025 (1982).

10 Sauvage, J. P. Interlacing molecular threads on transition metals: catenands, catenates, and knots. *Accounts of Chemical Research* **23**, 319-327 (1990).

11 Cantrill, S. J., Chichak, K. S., Peters, A. J. & Stoddart, J. F. Nanoscale Borromean Rings. *Accounts of Chemical Research* **38**, 1-9 (2004).

12 Ponnuswamy, N., Cougnon, F. B. L., Clough, J. M., Pantos, G. D. & Sanders, J. K. M. Discovery of an Organic Trefoil Knot. *Science* **338**, 783-785 (2012).

13 Beves, J. E., Blight, B. A., Campbell, C. J., Leigh, D. A. & McBurney, R. T. Strategies and Tactics for the Metal-Directed Synthesis of Rotaxanes, Knots, Catenanes, and Higher Order Links. *Angewandte Chemie International Edition* **50**, 9260-9327 (2011).

14 Aldaye, F. A. & Sleiman, H. F. Guest-Mediated Access to a Single DNA Nanostructure from a Library of Multiple Assemblies. *Journal of the American Chemical Society* **129**, 10070-10071 (2007).

15 Lee, D. H., Granja, J. R., Martinez, J. A., Severin, K. & Ghadiri, M. R. A self-replicating peptide. *Nature* **382**, 525-528 (1996).

16 Yao, S., Ghosh, I., Zutshi, R. & Chmielewski, J. A pH-Modulated, Self-Replicating Peptide. *Journal of the American Chemical Society* **119**, 10559-10560 (1997).

17 Carnall, J. M. A. *et al.* Mechanosensitive Self-Replication Driven by Self-Organization. *Science* **327**, 1502-1506 (2010).

18 Cougnon, F. B. L. & Sanders, J. K. M. Evolution of Dynamic Combinatorial Chemistry. *Accounts of Chemical Research* **45**, 2211-2221 (2011).

19 Corbett, P. T. *et al.* Dynamic Combinatorial Chemistry. *Chemical Reviews* **106**, 3652-3711 (2006).

20 Sadownik, J. W. & Ulijn, R. V. Dynamic covalent chemistry in aid of peptide self-assembly. *Current Opinion in Biotechnology* **21**, 401-411 (2010).

21 Mamathambika, B. S. & Bardwell, J. C. Disulfide-Linked Protein Folding Pathways. *Annual Review of Cell and Developmental Biology* **24**, 211-235 (2008).

22 Fass, D. Disulfide Bonding in Protein Biophysics. *Annual Review of Biophysics* **41**, 63-79 (2012).

23 Singh, R. & Whitesides, G. M. in *Sulphur-Containing Functional Groups (1993)* 633-658.

24 Wu, C., Leroux, J.-C. & Gauthier, M. A. Twin disulfides for orthogonal disulfide pairing and the directed folding of multicyclic peptides. *Nature Chemistry* **4**, 1044-1049 (2012).

25 Craik, D. J. Protein folding: Turbo-charged crosslinking. *Nature Chemistry* **4**, 600-602 (2012).

26 Salgado, E. N., Radford, R. J. & Tezcan, F. A. Metal-Directed Protein Self-Assembly. *Accounts of Chemical Research* **43**, 661-672 (2010).

27 Radford, R. J., Brodin, J. D., Salgado, E. N. & Tezcan, F. A. Expanding the utility of proteins as platforms for coordination chemistry. *Coordination Chemistry Reviews* **255**, 790-803 (2011).

28 Brodin, J. D. *et al.* Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chemistry* **4**, 375-382, (2012).

29 Faraone-Mennella, J., Tezcan, F. A., Gray, H. B. & Winkler, J. R. Stability and folding kinetics of structurally characterized cytochrome cb_{562}. *Biochemistry* **45**, 10504-10511 (2006).

30 Salgado, E. N., Faraone-Mennella, J. & Tezcan, F. A. Controlling Protein-Protein Interactions through Metal Coordination: Assembly of a 16-Helix Bundle Protein. *Journal of the American Chemical Society* **129**, 13374-13375 (2007).

31 Salgado, E. N., Lewis, R. A., Mossin, S., Rheingold, A. L. & Tezcan, F. A. Control of Protein Oligomerization Symmetry by Metal Coordination: C_2 and C_3 Symmetrical Assemblies through Cu^{II} and Ni^{II} Coordination. *Inorganic Chemistry* **48**, 2726-2728 (2009).

32 Salgado, E. N. *et al.* Metal-Templated Design of Protein Interfaces. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1827-1832 (2010).

33 Brodin, J. D. *et al.* Evolution of Metal Selectivity in Templated Protein Interfaces. *Journal of the American Chemical Society* **132**, 8610-8617 (2010).

34 Jones, D. P. in *Methods in Enzymology* Vol. Volume 348 (eds Sies Helmut & Packer Lester) 93-112 (Academic Press, 2002).

35 Cleland, W. W. Dithiothreitol, a New Protective Reagent for SH Groups. *Biochemistry* **3**, 480-482 (1964).

36 Lehn, J. M. Cryptates: the chemistry of macropolycyclic inclusion complexes. *Accounts of Chemical Research* **11**, 49-57 (1978).

37 Graf, E. & Lehn, J. M. Cryptates. XVII. Synthesis and cryptate complexes of a spheroidal macrotricyclic ligand with octahedrotetrahedral coordination. *Journal of the American Chemical Society* **97**, 5022-5024 (1975).

38 Christianson, D. W. & Fierke, C. A. Carbonic Anhydrase: Evolution of the Zinc Binding Site by Nature and by Design. *Accounts of Chemical Research* **29**, 331-339 (1996).

39 Lovejoy, B. *et al.* Structure of the catalytic domain of fibroblast collagenase complexed with an inhibitor. *Science* **263**, 375-377 (1994).

40 Salgado, E. N., Brodin, J. D., To, M. M. & Tezcan, F. A. Templated Construction of a Zn-Selective Protein Dimerization Motif. *Inorganic Chemistry* **50**, 6323-6329 (2011).

41 Braun, M. & Thony-Meyer, L. Biosynthesis of artificial microperoxidases by exploiting the secretion and cytochrome c maturation apparatuses of Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 12830-12835 (2004).

42 Thomas, P. E., Ryan, D. & Levin, W. An improved staining procedure for the detection of the peroxidase activity of cytochrome P-450 on sodium dodecyl sulfate polyacrylamide gels. *Analytical Biochemistry* **75**, 168-176 (1976).

43 Wunderlich, M. & Glockshuber, R. Redox properties of protein disulfide isomerase (dsba) from escherichia coli. *Protein Science* **2**, 717-726 (1993).

44 Raivio, T. L. MicroReview: Envelope stress responses and Gram-negative bacterial pathogenesis. *Molecular Microbiology* **56**, 1119-1128 (2005).

45 Yatsunyk, L. A. *et al.* Structure and metal binding properties of ZnuA, a periplasmic zinc transporter from Escherichia coli. *Journal of Biological Inorganic Chemistry* **13**, 271-288 (2008).

46 Guerinot, M. L. The ZIP family of metal transporters. *Biochimica et Biophysica Acta (BBA) - Biomembranes* **1465**, 190-198 (2000).

47 Hantke, K. Bacterial zinc uptake and regulators. *Current Opinion in Microbiology* **8**, 196-202 (2005).

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

49 Taylor, P. K., Parks, B. Y. A., Kurtz, D. M. & Amster, I. J. Analysis of metal incorporation during overexpression of Clostridium pasteurianum rubredoxin by electrospray FTICR mass spectrometry. *Journal of Biological Inorganic Chemistry* **6**, 201-206 (2001).

50 Goodsell, D. S. & Olson, A. J. Structural symmetry and protein function. *Annual Review of Biophysics and Biomolecular Structure* **29**, 105-153 (2000).

51 Janin, J. l. Dicey Assemblies. *Science* **319**, 165-166 (2008).

52 Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855-862 (2009).

53 Park, H.-S. *et al.* Design and Evolution of New Catalytic Activity with an Existing Protein Scaffold. *Science* **311**, 535-538 (2006).

54 Khare, S. D. *et al.* Computational redesign of a mononuclear zinc metalloenzyme for organophosphate hydrolysis. *Nature Chemical Biology* **8**, 294-300, (2012).

55 Koehler, V. *et al.* Synthetic cascades are enabled by combining biocatalysts with artificial metalloenzymes. *Nature Chemistry* **5**, 93-99 (2013).

56 Reig, A. J. *et al.* Alteration of the oxygen-dependent reactivity of de novo Due Ferri proteins. *Nature Chemistry* **4**, 900-906 (2012).

57 Zastrow, M. L., Peacock, F. A., Stuckey, J. A. & Pecoraro, V. L. Hydrolytic catalysis and structural stabilization in a designed metalloprotein. *Nature Chem*istry **4**, 118-123 (2012).

58 Chothia, C., Gough, J., Vogel, C. & Teichmann, S. A. Evolution of the Protein Repertoire. *Science* **300**, 1701-1703 (2003).

59 Schuck, P. A model for sedimentation in inhomogeneous media. I. Dynamic density gradients from sedimenting co-solutes. *Biophysical Chemistry* **108**, 187-200 (2004).

60 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Biological Crystallography* **67**, 235-242 (2011).

61 Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallographica Section D-Biological Crystallography* **66**, 22-25 (2010).

62 Murshudov, G., Vagin, A. & Dodson, E. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Crystallographica* **D53**, 240-255 (1996).

63 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **D60**, 2126-2132 (2004).

64 DeLano, W. L. *The PYMOL Molecular Graphics System (http://www.pymol.org).* (2003).

65 Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - A Program To Check The Stereochemical Quality Of Protein Structures. *Journal of Applied Crystallography* **26**, 283-291 (1993).

66 Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nature Methods* **9**, 671-675 (2012).

67 Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase. *Analytical Biochemistry* **237**, 260-273 (1996).

68 Dundas, J. *et al.* CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Research* **34**, W116-W118 (2006).

69 Voss, N. R. & Gerstein, M. 3V: cavity, channel and cleft volume calculator and extractor. *Nucleic Acids Research* **38**, W555-W562 (2010).

Chapter 4

Structural and Spectroscopic Characterization of Engineered Protein Macrocycles Binding Copper

Introduction

One of the goals of Bioinorganic Chemistry is the design of metal-binding sites within engineered proteins to create functional and catalytic metalloenzymes.[1] By designing metalloproteins, scientists comprehend complex enzymatic reactions. This understanding influences the design of small molecule biomimetic complexes with catalytic functions which can serve as catalysts that access chemistry that are challenging to attain.[2-4] The research within our laboratory explores the design of metal-binding proteins with the goal of creating protein scaffolds with the ability to accommodate different transition metals. Chapters 2 and 3 cover the discussion of the technique of Metal Templated Interface Redesign (MeTIR) and the use of Zn(II) as a templating metal to construct protein macrocycles.[5] Here we prove that the pre-templated macrocyclic proteins accommodate the redox-active metal copper.

Copper is a redox active metal that accesses oxidation states of +2 and +1 within natural proteins. The functions of copper proteins include electron transfer, substrate oxidation or oxygenation, and/or substrate reduction.[6-8] The oxidation state of the metal influences the copper coordination geometry and the metal coordinating ligands. In the plot in Figure 4.1a is graph that classifies copper complexes by their coordination geometry and oxidation states.[9] As expected, the coordination geometry of copper in its oxidized (+2) state follows tetragonal coordination. Oxidized copper is usually square pyramidal, square bipyramidal or trigonal bipyramidal. Reduced copper (+1) tends to have tetrahedral, linear, bent or trigonal coordination geometries.[6,10,11] There are, however, some exceptions, for example some Cu(II)

proteins have tetrahedral coordination geometries, and these proteins usually belong to the electron transfer cupredoxin family.[12] The different types of copper proteins are classified as Type 1, Type 2 or Type 3 (Figure 4.1b).[10]

As expected, the coordination geometry and type of classification determines the protein's function. Most Type 1 tetrahedral copper proteins function as electron transfer proteins because of their ability to maintain the copper site in an entatic state, which allows it to reach +2/+1 oxidation states with minimal reorganization of its ligands and thus not affect the rate of electron transfer.[13,14] Type 2 proteins mediate oxidation processes or are structural sites, and Type 3 proteins, which contain dinuclear copper centers, are usually responsible for substrate oxidations and oxygenations where electron transfer requires mediation by multiple metal sites.[7,11] Some examples include laccases and multi-copper oxidases.

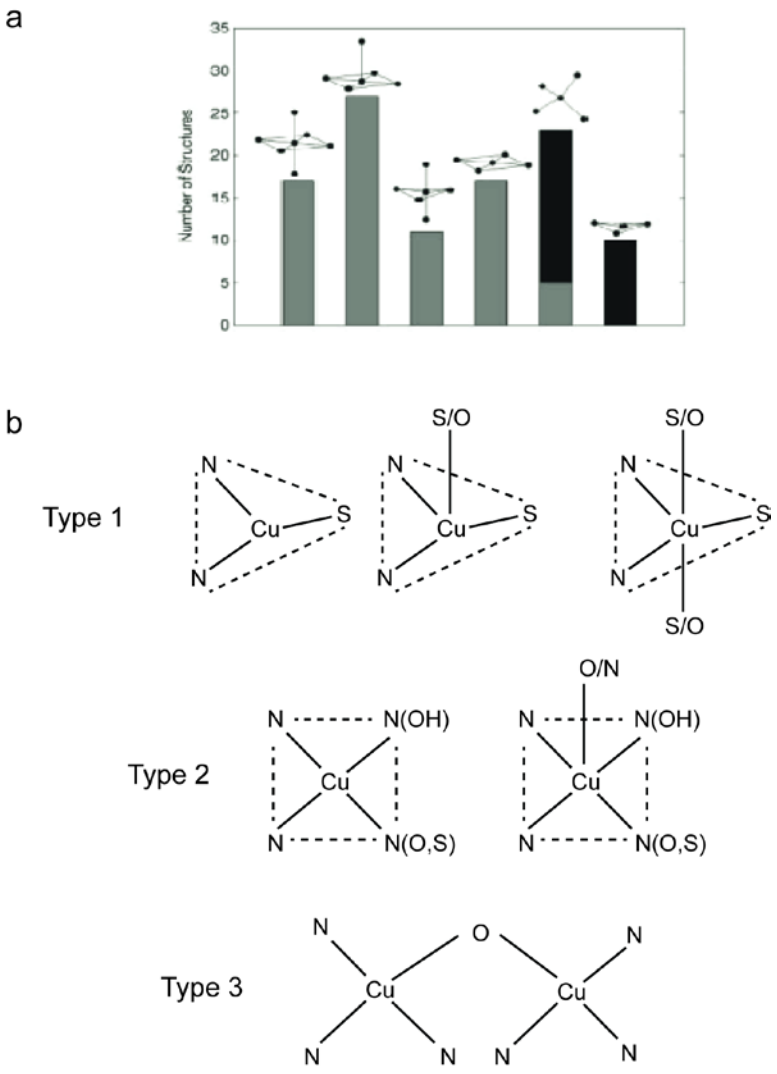

Figure 4.1. (**a**) Bar graph containing the number of structures of copper complexes classified by their coordination geometry and oxidation state (gray for +1, black for +2) (reference 9); (**b**) Classification of copper types 1, 2, and 3 by the active site's coordination geometry and ligand atoms.

Initial work: Cu(II) binds to MBPC1 and forms a supramolecular dimer.

The first effort in our laboratory to create an engineered protein was the construction of the Metal Binding Protomer Cytochrome-1 (MBPC1) which is a four-helix bundle cytochrome with engineered metal-binding histidine amino acids across its third helix.[15] The crystal structure of Cu_2:$MBPC1_2$, contains two square bipyramidal copper-binding sites that include the metal binding residues of His59, His63, His73, and His77, and two axial water ligands (Figures 4.2a, 4.2b, and 4.2c).[16] This protein forms a dimer in solution as well as crystallographically (Figure 4.2a, 4.2b). Electron Paramagnetic Resonace (EPR) studies of the protein in the presence of copper show an axial spectrum with parameters of $g_{\perp}$ = 2.06 and $g_{\|}$= 2.26, and hyperfine and superhyperfine coupling parameters of (Cu I=3/2) A_z=540 MHz; A_x=A_y=60 MHz and nitrogen (I=1) A_{iso} = 43 MHz; these parameters describe a square planar coordination geometry and type 2 copper binding (Figure 4.2d). The average Cu(II)-to-N_{His} distance of 2.08 Å compares favorably with Cu(II) tetrakis hexaimidazole complexes and the axial aquo ligand is subject to Jahn-Teller distortion because of the elongated Cu-H_2O distance of 2.55 Å.[17] The N_{His}-Cu-N_{His} angles of ~90˚ between the histidine residues at cis positions for copper indicate a near ideal square planar metal coordination environment. The metal coordination geometry in this assembly is unaltered by the supramolecular structure of the protein, which suggests that the *i* and *i*+4 His_2 metal coordination is the primary determinant of the dimeric supramolecular geometry.

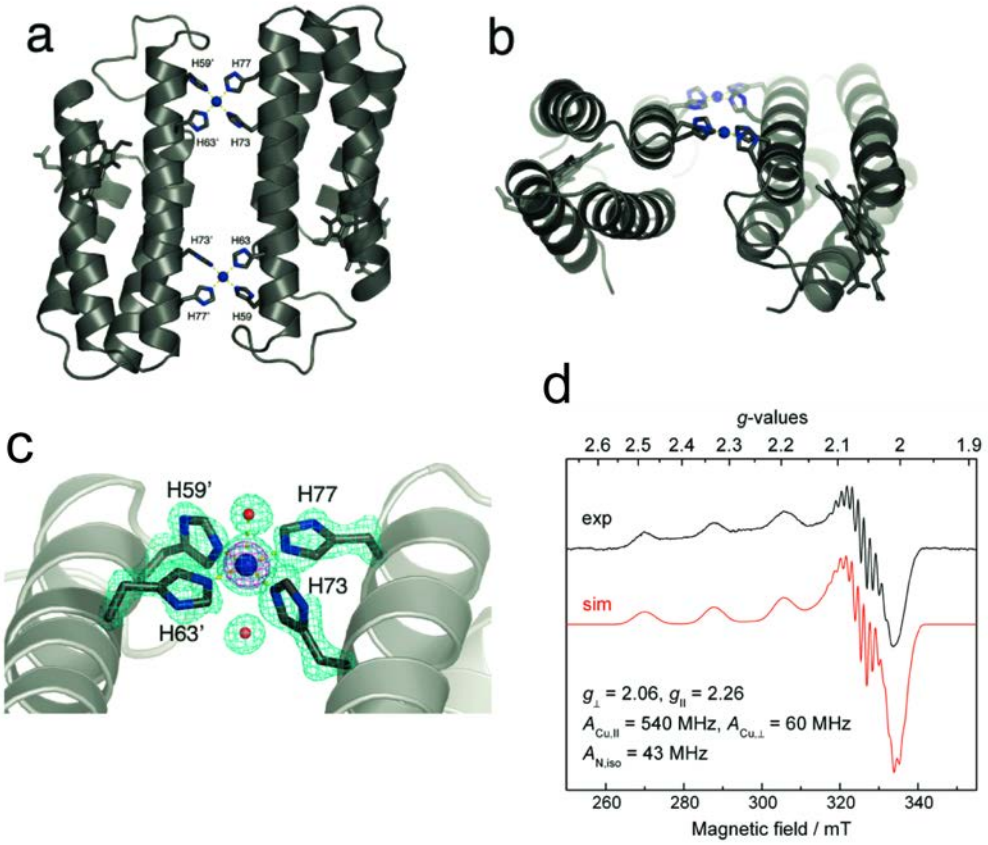

Figure 4.2. Characterization of Cu_2:MBPC1$_2$ (from reference 16). (**a**,**b**) Views of the Cu_2:MBPC1$_2$ crystal structure (PDB ID: 3DE8) showing the protein dimer as a cartoon in black, the metal binding histidines as sticks and the copper atoms are shown as blue spheres (**c**) Copper coordination environment showing His59, His63, His73, and His77 and water molecules shown as red spheres are positioned at 2.6 and 3.5 A respectively from copper (blue sphere) and Fo-Fc omit electron density maps (cyan at 3.2σ, magenta at 8σ) (**d**) Experimental (black) and simulated (red) X-band EPR spectra for Cu_2:MBP1$_2$ and the parameters used for the simulation.

Cu(II) can bind the RIDC1 protein dimer.

Furthermore, the effects on copper-binding of the engineered hydrophobic interface on the evolved Rosetta Interface Designed Complex- 1 (RIDC1) protein were investigated.[18] Although the goal of engineering of the RIDC1 protein was to stabilize the tetrameric zinc construct of Zn_4:RIDC1$_4$ (see Chapter 2), the hydrophobic

interactions across only one interface allows the RIDC1 protein to form a dimer in solution. Since the aforementioned Cu_2:$MBPC1_2$ is also a dimer, we hypothesized that the engineered hydrophobic interactions on RIDC1 will allow for more stable copper binding. Consequently, Salgado et al. characterized the dimeric construct of Cu_2:$RIDC1_2$.[18] After obtaining the crystal structure for this construct, the coordination geometry of the copper binding site was found to be largely influenced by the hydrophobic interface that was engineered onto the protein (Figure 4.3). The crystal structure of Cu_2:$RIDC1_2$ is of a dimer that possesses a crisscrossed monomer-to-monomer alignment which is significantly distorted in comparison to the antiparallel arrangement of the monomers seen in Cu_2:$MBPC1_2$ (Figures 4.3a, 4.3b). As a consequence of this arrangement of monomers, His59 is pushed out of the Cu(II) coordination sphere, thus leaving His63, His73, and His77 as the primary ligands (Figure 4.3c). The two Cu(II) coordination sites in the dimeric interface show a slightly distorted square-pyramidal geometry with one equatorial and one axial water molecule, which are both hydrogen-bonded to the Glu81 carboxylate from a symmetry-related dimer in the crystal lattice. The second copper site is in an approximately trigonal bipyramidal geometry whereby the crystallographically related Glu81 side chain is now directly coordinated. In addition, the spectroscopic signature of the EPR spectrum describes an axial type of copper containing nitrogen ligands because of the superhyperfine splittings (Figure 4.3d, shown by asterisks on inset). Contrasting the copper-bound dimers of MBPC1 and RIDC1, it is clear that the

hydrophobic interface influences a distortion of the dimer and thus alters the metal coordination geometry.

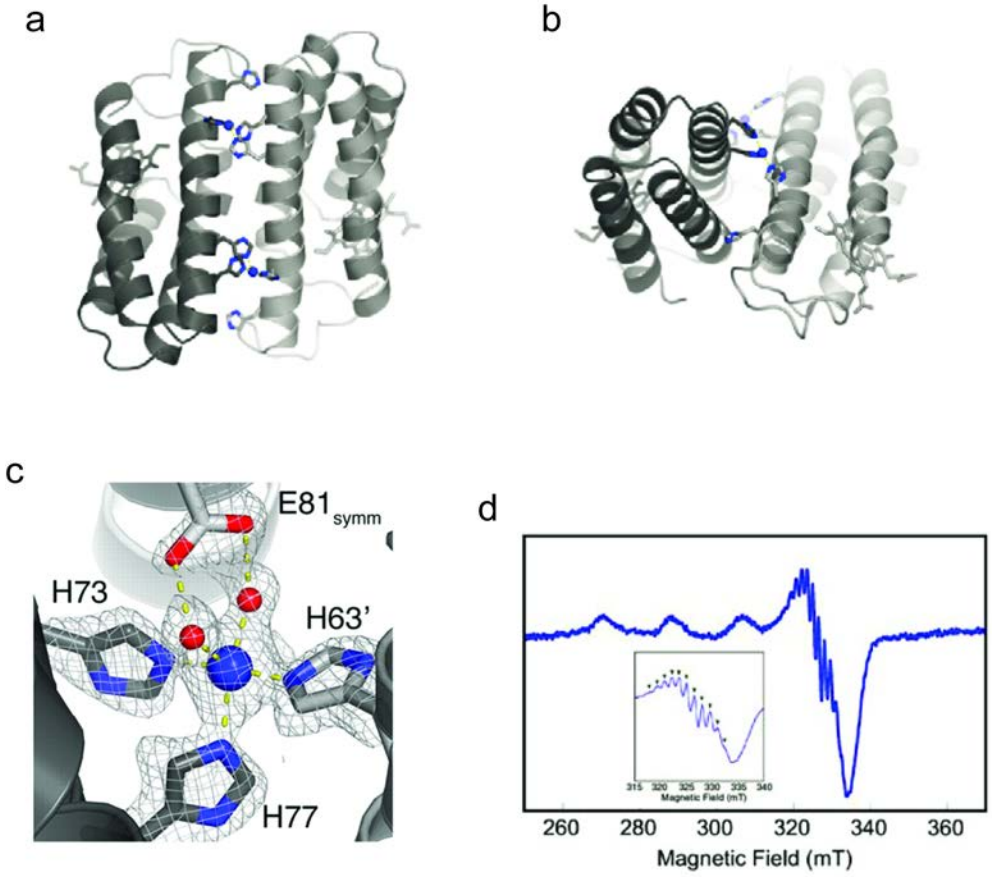

Figure 4.3. Characterization of Cu_2:$RIDC1_2$ (from reference 18). (**a**,**b**) Views of the Cu_2:$RIDC1_2$ crystal structure (PDB ID: 3HNL) showing the protein dimer as a cartoon in black and gray with the metal binding histidines as sticks and the copper atoms are shown as blue spheres (**c**) Copper coordination environment showing His63, His73, His77 and water molecules shown as red spheres that are hydrogen-bonding with Glu81 and the F_o-F_c omit electron density map is contoured at 3σ. (**d**) Experimental X-band EPR spectra for Cu_2:$RIDC1_2$ and the superhyperfine splitting pattern is highlighted by asterisks in the inset.

Results and Discussion

^{C96}RIDC1$_4$ binds Cu(II) in two forms.

After having explored the ability of the monomeric engineered proteins MBPC1 and RIDC1 to bind Cu(II) and consequently drive the formation of supramolecular dimers, we were interested in studying whether a pre-templated tetrameric scaffold can also accommodate copper, and whether the copper coordination geometry of the site could be predicted and stabilized. For this goal we used the Zn(II)-templated macrocyclic protein scaffolds to explore copper binding (See Chapters 2, and 3). We previously described the introduction of covalent crosslinks, in the form of disulfide bonds, across interface 2 of the Zn$_4$:RIDC1$_4$ tetramer (Figures 2.1 and 2.2).[5] The resulting ^{C96}RIDC1$_4$ tetramer forms in the absence of metal through the interactions of the hydrophobic surface across interface *i1*, and through the covalent disulfide bonds across interface *i2*.[5] Brodin et al. reported the ability of the ^{C96}RIDC1$_4$ tetramer to bind copper and identified through competition assays that it binds two Cu(II) ions in solution with very low dissociation constants in the range of 10^{-12}-10^{-13} M order of magnitude.[5] However, since the dimers of MBPC1 and RIDC1 bound to two coppers each, we expected the tetramer to bind to four Cu(II) ions and to involve the same metal binding amino acids as we had already identified (His59, His63, His73, His73 and/or Glu81).

To explore the coordination of copper to ^{C96}RIDC1$_4$ we obtained the crystal structures of the protein in the presence of Cu(II) in two different forms (Figure 4.4a, 4.4b, 4.4c, 4.4d). In both crystal structures the protein is crystallized as a tetramer.

However, the overall arrangement of the protomers within the ^{C96}RIDC1$_4$ scaffold is altered because of the fluidity of interface *i1* (Figure 4.4a, 4.4c). The first structure is the two-copper structure Cu$_2$:^{C96}RIDC1$_4$, which features two near-identical copper-binding sites with coordination to two nitrogens from two His77 ligands and to two oxygen atoms from two Glu81 carboxylates (Figure 4.4b). The coordination geometry of the binding site is distorted square pyramidal with the N_{His}-Cu-N_{His} bond angles averaging 82˚ and the O_{Glu}-Cu-O_{Glu} bond angles from the Glu81 carboxylate ligands averaging 83˚. The bond distances of the carboxylate oxygen from Glu81 to copper averages 2.0±0.1 Å and so do the N_{His} to Cu(II) bond distances. Also, if the Cu$_2$:^{C96}RIDC1$_4$ structure is superimposed with the Zn$_4$:^{C96}RIDC1$_4$ structure, there is a 2.9 Å root mean square deviation (r.m.s.d.) over C_α, which suggests that a preference for the hydrophobic interface to be in this particular arrangement.

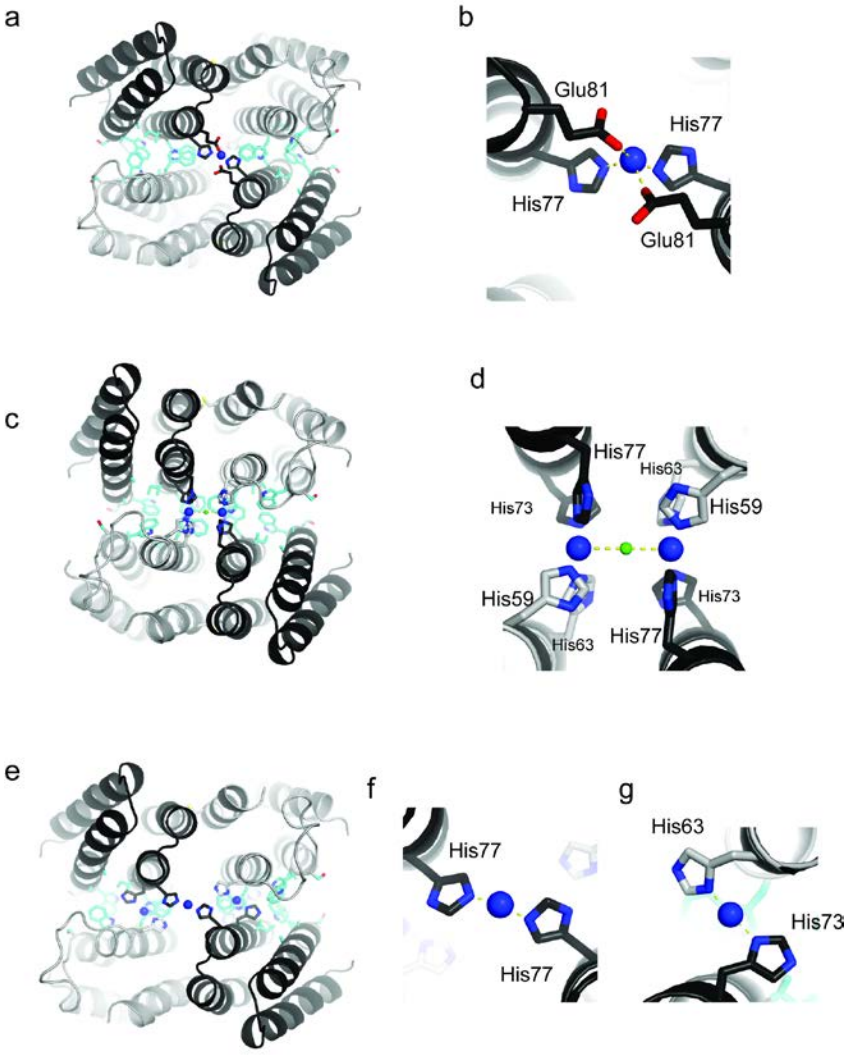

Figure 4.4. Crystal structures of ^{C96}RIDC1$_4$ crystallized in the presence of Cu(II) or Cu(I). Structures are shown as cartoons in light gray and dark gray with copper atoms shown as blue spheres and coordinating histidine amino acids shown as sticks and labeled accordingly. The hydrophobic mutations across interface *i1* are shown as cyan sticks. (**a**) Cu$_2$:^{C96}RIDC14 is shown and its copper site is shown in (**b**) which includes two His77 and two His81 amino acids as ligands. (**c**) Cu$_4$:^{C96}RIDC1$_4$ is shown with the copper sites shown in figure (**d**) and the dinuclear copper sites are shown in blue spheres with histidine ligands as sticks and bridging chloride ion is shown as a green sphere. (**e**) Cu(I)$_6$:^{C96}RIDC1$_4$ is shown and its two types of copper binding sites are shown in (**f**) and (**g**).

The second crystal structure was obtained with a different set of crystallization conditions and is of Cu_4:^{C96}RIDC1$_4$. Here we identify four Cu(II)-binding sites each bound to the previously described four histidine ligand motif of His59, His63, His73, and His77 (Figure 4.4c, 4.4d). Interestingly, to bind to all four histidines, as well as to maintain the engineered hydrophobic interactions, the copper binding sites have to come in close contact with one another and form two dinuclear copper sites which are bridged by a chloride ion (Figure 4.4d). As expected for a copper ion in its oxidized form, the coordination geometry identified crystallographically is tetragonal with all of the metal sites being square pyramidal and the four histidines at the plane with the bridging chloride as an axial ligand and an average bond distance for Cu-Cl-Cu of ~5 Å. The average N_{His}-Cu bond distance is 2.0±0.1 Å and the average bond angles between N_{His}-Cu-N_{His} of 89±3˚ agree with near-ideal square planar coordination geometry. Each half of the Cu_4:^{C96}RIDC1$_4$ tetramer can be superimposed with the dimer of Cu_2:MBPC1$_2$ with low a r.m.s.d. of 0.591 Å over C_α (Figure 4.5). This observation suggests a preferential coordination for the coppers with the histidine nitrogens as square planar ligands.

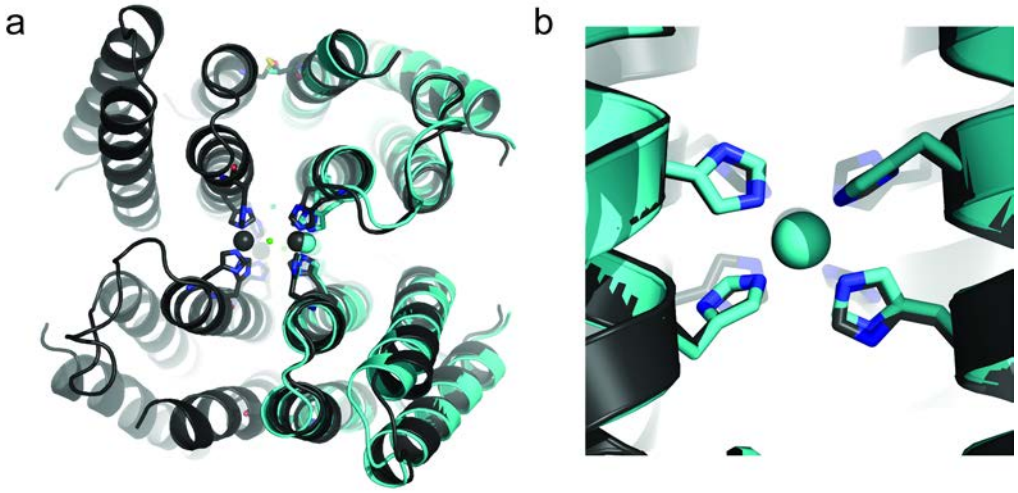

Figure 4.5. Crystal structure of dimer of Cu_2:MBP$C1_2$ is shown in cyan overlayed on the half of Cu_4:^{C96}RID$C1_4$ (shown in black). There is a high homology shown by the agreement of 0.591 Å of r.m.s.d. over backbone α-carbon.

Effects of oxidation state on metal binding: Cu(I) binding ^{C96}RID$C1_4$.

Furthermore, using the ^{C96}RID$C1_4$ construct, we studied the effect of the oxidation state of copper on the coordination geometry and ligand preferences. Since we identified that the Cu(II) ion can bind up to four histidine nitrogens as ligands, we wanted to learn if this tetragonal coordination geometry can be distorted to tetrahedral and maintain the four His ligands per copper site. Alternatively, the coordination geometry can also be altered to bind less than four histidine ligands in trigonal, linear, or bent coordination geometries. After obtaining the crystal structure of this $Cu(I)_6$:^{C96}RID$C1_4$ we found it to bind to 6 Cu(I) ions where each has bent coordination geometries (Figure 4.4e, 4.4f, 4.4g). There are two types of Cu(I) coordination sites present. Four of the Cu(I) are bound through the δ-nitrogens of His63 and His73 with

average N_{His}-Cu distances of 2.0±0.1 Å and bond angles of 167±10˚ (Figure 4.4g); whereas the remaining two Cu(I) ions bind two nitrogens from His77 with average bond distances of 2.0±0.1 Å and bent bond angles of 120±11˚ (Figure 4.4f). Although the coordination of up to four ligands to Cu(I) in a tetrahedral fashion is possible, it is not observed in the crystal structure. Also, if we superimpose this $Cu(I)_6$:^{C96}RIDC1$_4$ tetramer to the Cu_2:^{C96}RIDC1$_4$ tetramer (Figure 4.6) it is also superimposable with the low r.m.s.d. value of 0.579 Å over C_α. In the overlay in figure 4.6a, we observe that amino acids His63 and His73 are in the same configuration in both structures. Also, an examination at the hydrophobic interface across *i1*, reveals that the engineered hydrophobic amino acids are superimposable in both structures and that the tryptophans Trp66 and Trp41 of each tetramer are pi-stacking with their corresponding pairs in both crystal structures. This finding reveals a preference of the hydrophobic core to form Trp-Trp pi-stacking arrangements. After examination of the overlayed crystal structures of Cu_2:^{C96}RIDC1$_4$ and Cu_6:^{C96}RIDC1$_4$, we can hypothesize that the interfacial interactions across the engineered hydrophobic interface influence the arrangement of the tetramer and thus affect the coordination to copper in both oxidation states.

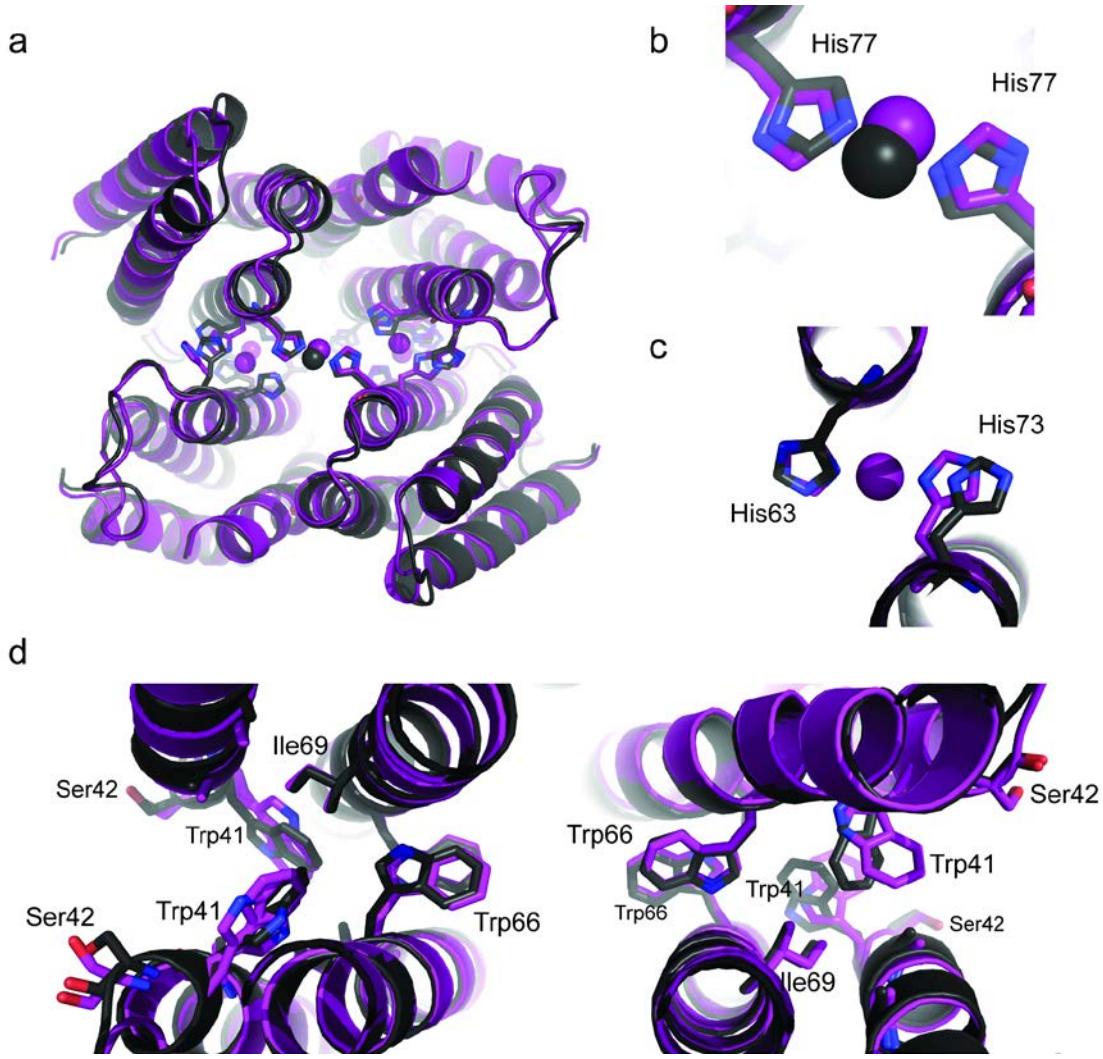

Figure 4.6. Crystal structure of tetramers of Cu_2:^{C96}RIDC1$_4$ shown in black overlayed with the Cu(I)$_6$:^{C96}RIDC1$_4$ (shown in magenta). (**a**) Tetramer configuration is shown with high degree of homology with 0.579 Å of r.m.s.d. over backbone α-carbon; (**b**) Cu(II) and Cu(I) binding to His77 are shown in black and magenta respectively for both crystal structures; (**c**) Cu(I) binding shown in magenta is overlayed with pre-oriented His63 and His73 ligands in the Cu_2:^{C96}RIDC1$_4$ structure (in black); (**d**) Interface *i2* of both crystal structures are overlayed to show the engineered hydrophobic amino acids Trp41, Ser42, Trp66, Ile69 overlaying in both crystal structures.

To characterize the Cu(II) binding in solution and to identify which coordination geometry the metal adopts in solution, we proceeded to perform EPR studies on the copper-incubated sample of ^{C96}RIDC1$_4$. After incubating the sample in the presence of copper, we performed elemental analysis using ICP-OES to ensure that there were four equivalents of Cu(II) bound per tetramer. An EPR spectrum of the Cu$_4$:^{C96}RIDC1$_4$ sample is shown in Figure 4.7a and displays very broad hyperfine splittings which suggests a mixture of species in solution. With this result we can hypothesize that the copper is bound with two different coordination geometries in solution since we were able to identify two different forms crystallographically. Although, the EPR spectrum cannot be easily simulated because of the multiple coordination geometries, we can identify superhyperfine splittings that can be correspond to the nitrogen atoms from histidine ligands binding to the coppers. The nitrogen ligands cannot be fully identified but can correspond to the aforementioned His59, His63, His73, and/or His77.

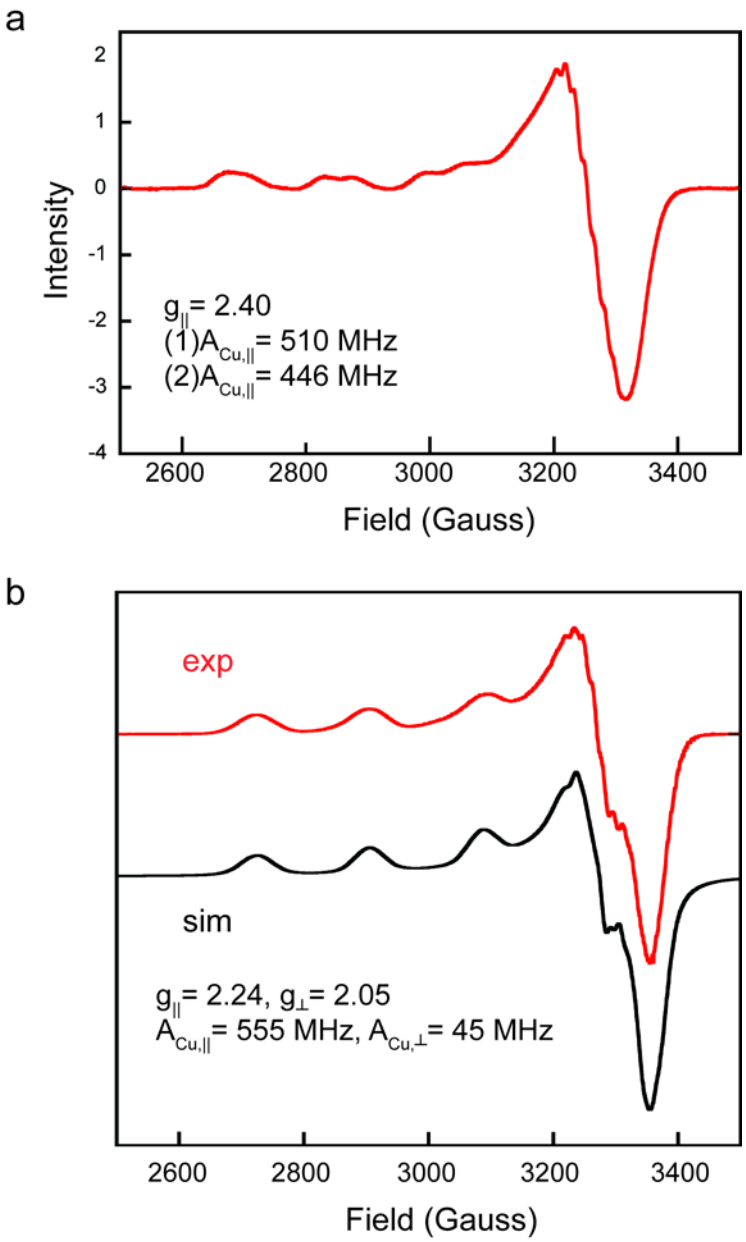

Figure 4.7. EPR characterization and simulations of ^{C96}RIDC1$_4$ and $^{A81/C96}$RIDC1$_4$ in the presence of Cu(II). (**a**) Experimental X-Band EPR spectrum of ^{C96}RIDC1$_4$ incubated with Cu(II). The hyperfine splitting pattern shows broadening which is indicative of multiple species present in solution. (**b**) Experimental (red) and simulated (black) spectra of $^{A81/C96}$RIDC1$_4$ incubated with Cu(II) and simulated parameters are shown in the inset.

Spectroscopic characterization of Cu(II) binding to $^{A81/C96}RIDC1_4$.

To isolate a single type of copper-binding form on $^{C96}RIDC1_4$, we substituted the metal-coordinating glutamate at position 81 for a non-chelating alanine with the hypothesis that this mutation would force the metal to only bind to the remaining histidine binding sites. We performed EPR studies of this construct in the presence of Cu(II) (Figure 4.7b). From this spectrum we can identify only one type of Cu(II) coordination as indicated by the single type of hyperfine splittings. We can simulate the spectrum to a Cu(II) bound to four nitrogen atoms with axial g-tensors of $g_\perp = 2.05$ and $g_\parallel = 2.24$ and spectroscopic parameters of $A_{Cu,\parallel} = 555$ MHz and $A_{Cu\perp} = 45$ MHz. After many trials and different crystallographic conditions, every effort to obtain the crystal structure of copper-loaded $^{A81/C96}RIDC1_4$, led to obtaining of metal-free tetramers that were isostructural to the reported apo $^{C96}RIDC1_4$.[5] Since we were not able to obtain its crystal structure we cannot fully identify the copper-coordinating amino acids but can hypothesize that they can include His59, His63, His73 and/or His77 since we can model the EPR spectrum to contain four nitrogen ligands.

Cu(II) binds to $^{C96}(MBPC1_2)_2$ and drives tetramer formation.

To better understand the effect of the hydrophobic interactions on the coordination geometry of copper to $^{C96}RIDC1_4$, we studied the same construct, without the hydrophobic interactions, in the presence of copper. Therefore we engineered in the Cys96-Cys96 disulfide bond onto the previously studied MBPC1 construct (Figure 4.8). The resulting $^{C96}MBPC1_2$ forms a dimer in solution because of its covalent disulfide crosslink. After adding Cu(II) to the sample the sedimentation coefficient

increases from 2.6 to 4.2 S which is indicative of dimerization of the covalently crosslinked dimer (Figure 4.8d). Furthermore, the crystal structure of Cu_4:^{C96}MBPC1$_4$ was obtained and solved to 2.40 Å resolution (Figure 4.8a). This construct forms a tetramer with two disulfide-bonded ^{C96}MBPC1$_2$ dimers that are interacting exclusively through the four Cu(II) binding sites. The Cu(II) atoms all bind with near-identical coordination geometries and contain the four histidine ligands of His59, His63, His73, His77, and some of the copper ions contain axial water ligands with average distances of Cu-H_2O of 2.0±0.2 Å (Figure 4.8b). The average Cu-N_{His} distance of 2.1±0.1 Å and the average bond angles between N_{His}-Cu-N_{His} of 90±6˚, indicate tetragonal coordination geometry. This is the same coordination motif observed for the parent Cu_2:MBPC1$_2$ and an overlay of the Cu_2:MBPC1$_2$ dimer with the Cu_4:^{C96}MBPC1$_4$ tetramer has a low r.m.s.d. of 2.25 Å over C_α (Figure 4.8c). In addition, we note that in this Cu_4:^{C96}MBPC1$_4$ crystal structure the copper ions are separated by 14.0±0.1 Å, with no other contacts between the disulfide-bonded dimers, which contrasts sharply to the shorter distance between the dinuclear copper centers observed for Cu_4:^{C96}RIDC1$_4$ (~5 Å), which is attributed to the engineered interactions forcing the hydrophobic interface into place.

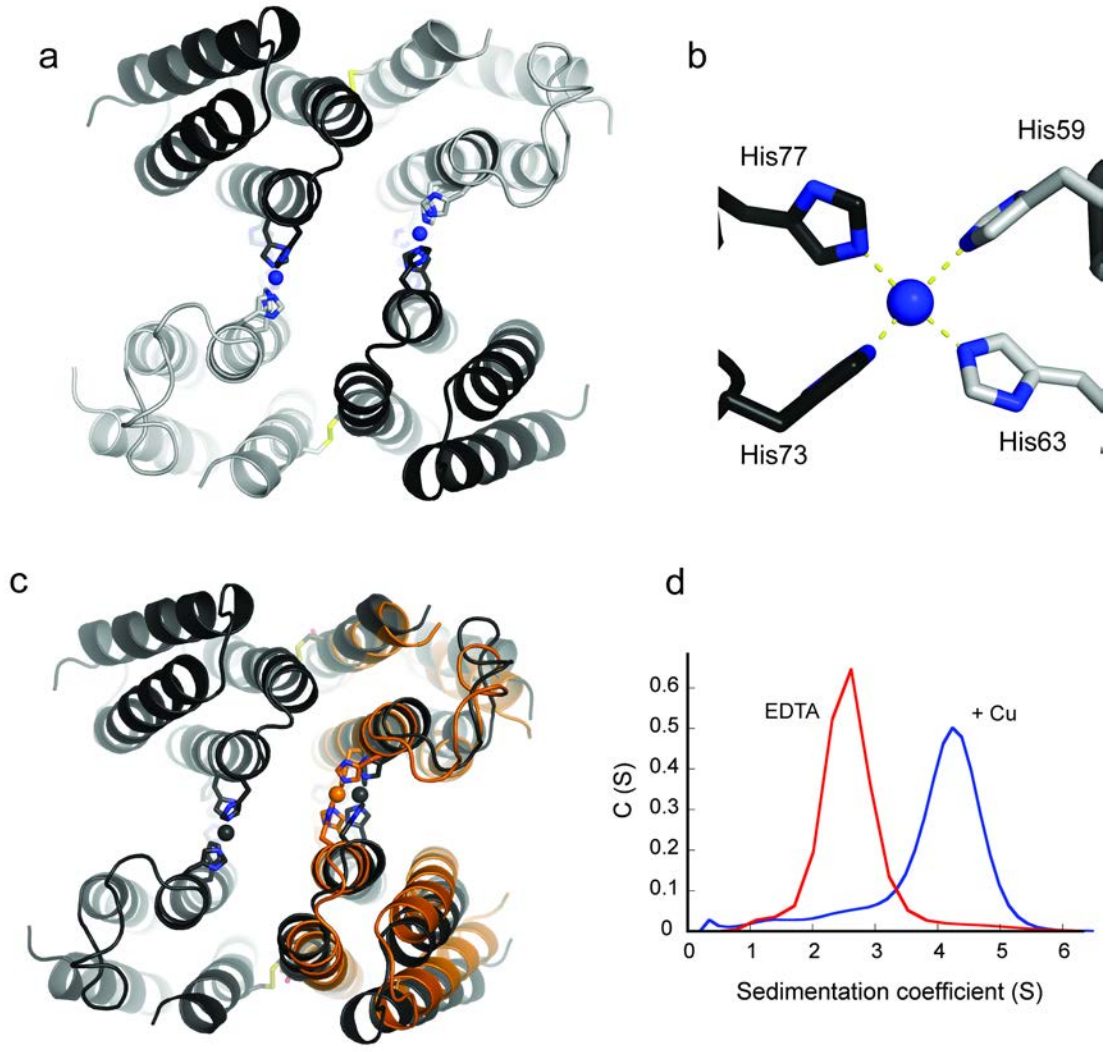

Figure 4.8. Crystal structure of Cu_4:(^{C96}MBPC1$_2$)$_2$ is shown (**a**) in cartoon in dark gray and light gray with disulfide bonds shown as yellow sticks and metal-binding histidines as sticks with coordinating coppers shown as blue spheres. (**b**) Copper coordination to His59, His63, His73 and His77. (**c**) Overlay of Cu2:MBPC12 (orange) with one half of Cu_4:(^{C96}MBPC1$_2$)$_2$ with r.m.s.d. of 2.25 Å over α-carbon. (**d**) Sedimentation velocity distribution of samples of ^{C96}MBCP1$_2$ in the presence of excess EDTA (1 mM) (shown as red) and the increase of sedimentation coefficient shown by the blue line.

Therefore, we determine that without the hydrophobic interactions present the coordination of the copper to the four histidine ligands in a tetragonal fashion is preferred. This preference for coordination through the *i*, *i+4* amino acids (His59, His63 and His73, His77) as ligands is what drives the dimerization of the protomers of MBPC1 and ^{C96}MBPC1$_2$. However, when the hydrophobic interface is introduced, the coordination of the metal site is altered to favor the most thermodynamically favorable arrangement of the hydrophobic interactions.

Use of pre-templated $^{C81/C96}$RIDC1$_4$ tetramer to bind Cu(II).

Since we found the ^{C96}RIDC1$_4$ scaffold to be too fluid to allow predictable and consistent copper coordination, we proceeded to use the engineered covalently crosslinked macrocyclic assembly of $^{C81/C96}$RIDC1$_4$ to accommodate copper. Differently from ^{C96}RIDC1$_4$, the engineered $^{C81/C96}$RIDC1$_4$ tetramer binds up to four Cu(II) in solution with high affinity as identified from its competition assay with Fura-2 ($K_{d,Cu}$ = 3×10^{-13}).[5] The measured dissociation constants are in the picomolar order of magnitude and the data fits best to a two-event binding model (2 + 2, dotted black line) (Figures 4.9a, 4.9b and Table 1). Additionally, we fit the data to an alternative one-site model (1 × 4, gray line) which was not appropriate, but the fit to a four-site model (1 + 1 + 1 + 1, red line) is as good as the two-site model.

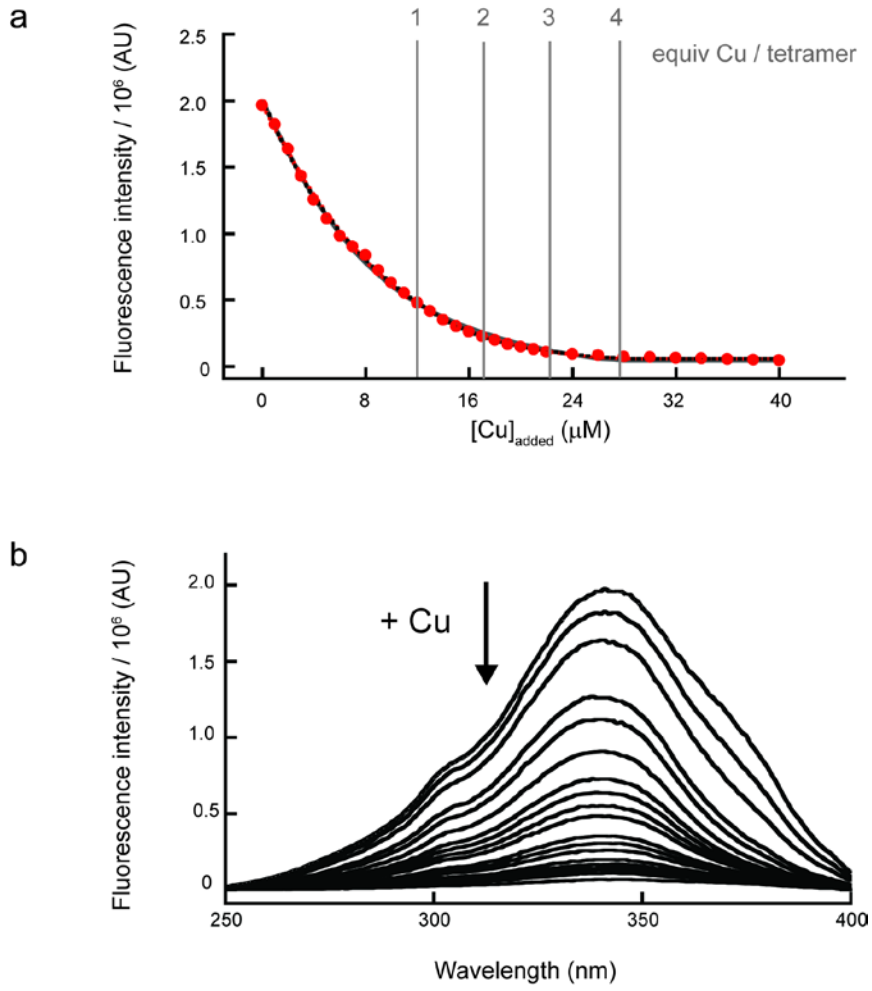

Figure 4.9. Competition assay between Fura-2 (6.25 μM) and $^{C81/C96}RIDC_4$ (5.25 μM) for Cu(II). (**a**) Excitation intensity at 343 nm plotted versus Cu(II) concentration and fits to 1 × 4 Cu (gray dotted line), 2 + 2 Cu (black line) and or 4 × 1 Cu (red line) binding models; (**b**) Excitation scans using excitation range of 250-400 nm and detection at 510 nm for individual titration plots are plotted to demonstrate the decrease in fluoresecence intensity after the addition of copper.

We obtained the crystal structure of $^{C81/C96}RIDC1_4$ in the presence of Cu(II) in two different forms (Figures 4.10a, and 4.10d). In the first structure, we identify two Cu(II) with tetragonal coordination geometries involving the His59, His63, His73, and His77 as ligands (Figure 4.10a, 4.10b, and 4.10c). These histidine sites all have

distorted tetragonal coordination with bond angles between the N_{His}-Cu-N_{His} that average 90±9˚ and Cu(II)-N_{His} bond distances that average 2.2 Å for His59, His73 and His77 and a longer bond distance of 2.5 Å for His63. Also, the half of the tetramer that is bound to copper can be superimposed with the Cu_2:MBPC1_2 dimer with an r.m.s.d. value of 1.580 Å over C_α. In the other half of the structure that is not bound to metal, the empty histidine sites are too far to allow for metal binding (Figure 4.10b) and this half can be superimposed to the Cu_2:MBPC1_2 dimer with the higher r.m.s.d. value of 3.036 Å over C_α.

In the second structure, we identified four Cu(II) ions bound to the $^{C81/C96}$RIDC1_4 tetramer (Figure 4.10d, 4.10e, 4.10f). There are two types of Cu(II) coordination sites observed within this structure. One set binds to the previously mentioned four histidines set with near-ideal tetragonal coordination geometry (Figure 4.10e). However, the second type of Cu(II) site contains three histidines at appropriate metal binding distances, but the fourth ligand, His63, is too distant to bind the metal (3.8 Å, Figure 4.10f). In this site there is a distortion from tetragonal coordination geometry as indicated by the bond angles N_{His}-Cu-N_{His} of 107˚, 101˚, and 150˚. The EPR spectrum of this construct is of a single type of copper site containing a Type 2 copper which we can model to contain 4 N_{His} as ligands. The simulated spectrum can be identified as axial with parameters of $g_\perp$ = 2.06 and $g_\parallel$ = 2.26 and hyperfine splitting parameters of $A_{Cu,\perp}$ = 555 MHz and $A_{Cu,\parallel}$ = 45 MHz. These findings lead us to believe that this scaffold is ideal to enforce tetrahedral coordination

geometry on the metal site since the copper coordination geometry is being influenced by the requirements of the supramolecular scaffold.

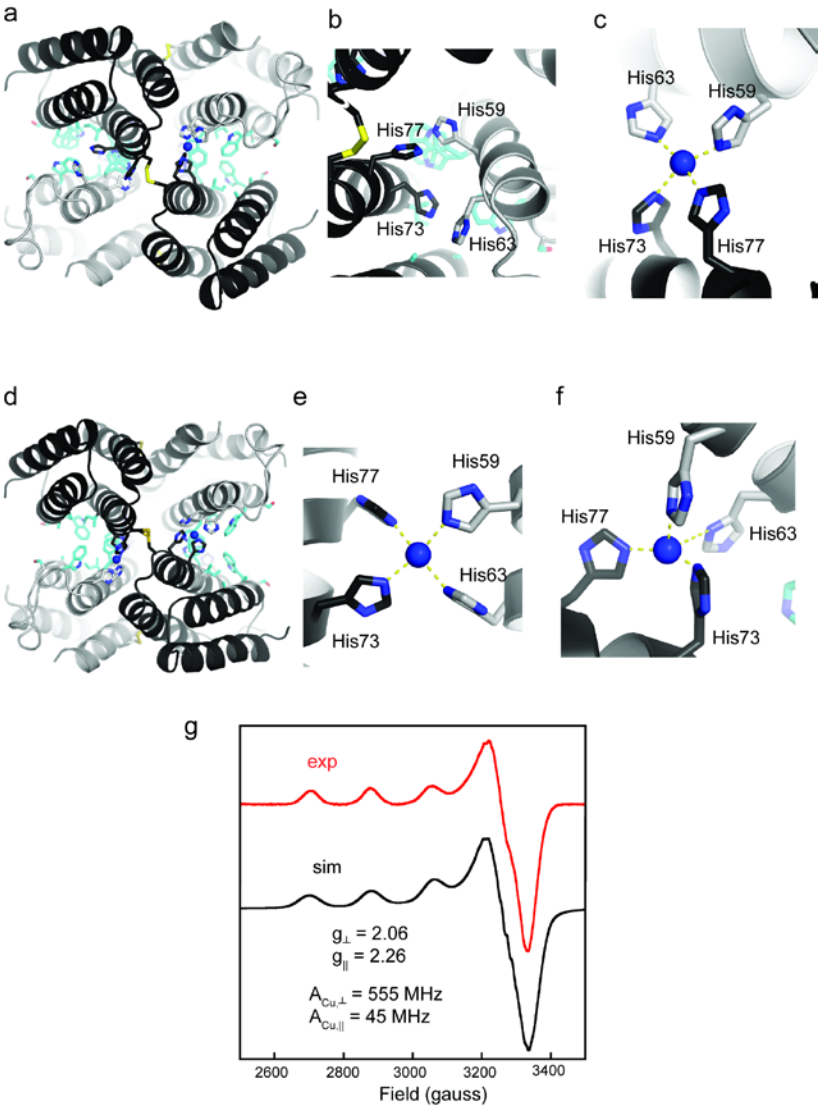

Figure 4.10. Crystal structures and EPR spectrum of $^{C81/C96}$RIDC1$_4$ in the presence of Cu(II). (**a**) Cu$_2$:$^{C81/C96}$RIDC1$_4$ crystal structure shown as dark gray and light gray cartoon with hydrophobic mutations shown as cyan sticks; (**b**) empty binding site of two-copper crystal structure; (**c**) Cu(II) binding site with copper shown as blue sphere and metal binding amino acids shown as sticks and labeled as His59, His63, His73 and His77; (**d**) Cu$_4$:$^{C81/C96}$RIDC1$_4$ crystal structure shown as dark gray and light gray cartoon with hydrophobic mutations shown as cyan sticks; (**e**) Cu(II) binding site with metal binding amino acids shown as sticks and labeled as His59, His63, His73 and His77; (**f**) Second type of Cu(II) binding site with binding sites of His59, His73 and His77other copper binding site; (**g**) Experimental (red) and simulated (black) X-Band EPR spectra of copper-loaded sample (300 µM) with four equivalents of Cu(II) bound per tetramer. Simulated parameters are shown in inset.

Catalytic copper proteins with a 3His coordination motif are present in nature. One example is nitrite reductase which is able to reduce NO^{2-} to NO gas. We wanted to investigate whether we can controllably enforce a 3His coordination motif on Cu(II) by using our macrocyclic protein scaffold. For this purpose we substituted His59 for a non-chelating alanine and studied the copper coordination on this structure. Upon crystallizing the $^{A59/C81/C96}RIDC1_4$ in the presence of copper we were pleased to find that it is able to bind to copper through the three remaining histidine nitrogens of His63, His73, and His77, and to have an open coordination site bound by waters (Figure 4.11a, 4.11b). We observe in the structure that there are two fully occupied copper sites per tetramer, a third partially occupied (occupancy ~ 0.5) copper site and an empty binding site. The Cu(II) coordination geometries of the fully occupied sites have average Cu-N_{His} distances of 2.1±0.1 Å and bond angles of N_{His}-Cu-N_{His} of 95±3˚, which suggests a distorted tetragonal site. Furthermore, we obtained the EPR spectrum of the copper-incubated sample of $^{A59/C81/C96}RIDC1_4$ as shown in Figure 4.11c. The axial spectrum can be simulated by parameters that are very similar to the parent scaffold of $g_{\perp} = 2.06$ and $g_{\parallel} = 2.26$ and hyperfine splitting parameters of $A_{Cu,\perp}$ = 550 MHz and $A_{Cu,\parallel}$ = 45 MHz. We are currently exploring whether this construct is able to bind to Cu(I) in the same coordination geometry in order to use it as an artificial nitrite reductase scaffold.

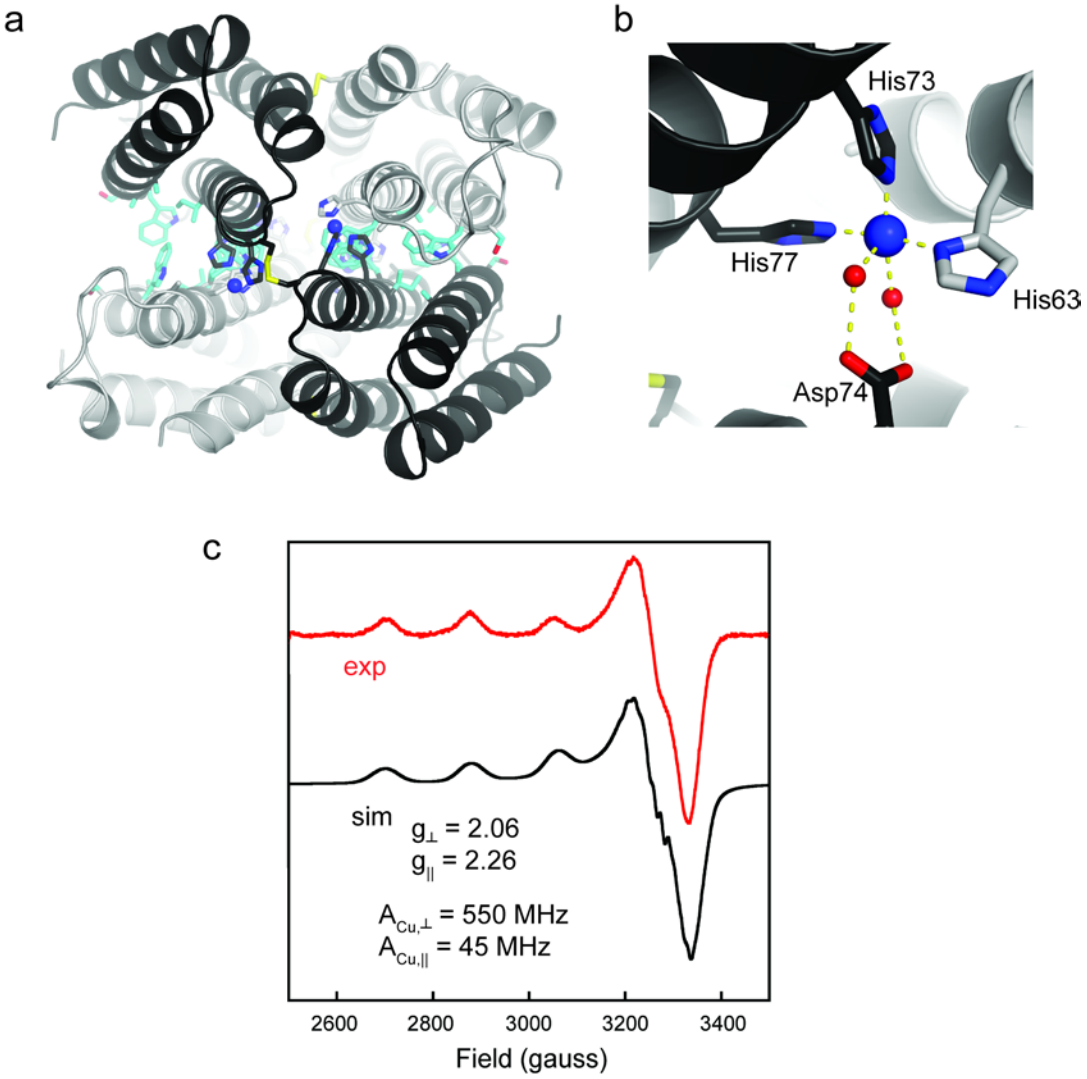

Figure 4.11. Crystal structures and EPR spectrum of $^{A59/C81/C96}$RIDC1$_4$ in the presence of Cu(II). (**a**) Cu$_3$:$^{A74/C81/C96}$RIDC1$_4$ crystal structure shown as dark gray and light gray cartoon with hydrophobic mutations shown as cyan sticks; (**b**) Cu(II) binding site with copper shown as blue sphere and metal binding amino acids shown as sticks and labeled as His63, His73 and His77 with two water molecules (red spheres) that are hydrogen bonding to Asp74; (**c**) Experimental (red) and simulated (black) X-Band EPR spectra of copper-loaded sample (300 µM) with four equivalents of Cu(II) bound per tetramer. Simulated parameters are shown in inset.

Conclusions

Our MeTIR approach has led us to construct several protein scaffolds with engineered binding sites. By using these scaffolds as macrocyclic ligands we studied Cu(II) binding and the influence of the hydrophobic interactions upon the metal site's coordination geometry. We hypothesize that Cu(II) prefers a four His coordination motif involving histidines His59, His63, His73, and His77 with an *i, i+4* coordination motif. Upon introduction of the hydrophobic interaction across *i1* we altered the coordination geometry of the metal and influenced a distortion on the tetragonal coordination. Furthermore, by using covalently crosslinked tetramers we are now able to control the coordination of the metal and we are currently studying a three coordinate copper site on $^{A59/C81/C96}RIDC1_4$ which has an open binding site that can be used as a catalytically active nitrite reductase mimic.

Table 4.1. Dissociation constants for $^{C81/C96}RIDC1_4$ and Cu(II) measured after competition assay with Fura-2 and fit to 1-site, 2-site and 4-site binding models.

	1-site	2-site	4-site
Kd_1	$(2.2 \pm 0.6) \times 10^{-13}$	$(1.0 \pm 0.2) \times 10^{-12}$	$(9.0 \pm 3.0) \times 10^{-13}$
Kd_2	---	$(5.1 \pm 1.0) \times 10^{-12}$	$(6.3 \pm 1.4) \times 10^{-13}$
Kd_3	---	---	$(1.0 \pm 0.4) \times 10^{-11}$
Kd_4	--	---	$(7.8 \pm 1.8) \times 10^{-12}$

Table 4.2. X-Ray data collection and refinement statistics for ^{C96}RIDC1$_4$ with 6 Cu(I). *Highest resolution shell is shown in parentheses.

Data collection	
Space group	$P12_11$
Cell dimensions	
a, *b*, *c* (Å)	48.05, 87.16, 51.75
α, β, γ (˚)	90.00, 95.80, 90.00
Resolution (Å)	51.48-1.80
R_{sym} or R_{merge} (%)	7 (10)
$I/\sigma I$*	16.2 (9.0)
Completeness (%)*	96.1 (100.0)
Redundancy*	5.6 (5.7)
Refinement	
Resolution (Å)	2.15
No. unique reflections	1505
R_{work}/R_{free}	0.189/0.276
No. atoms	
Protein	3487
Ligand/ion	6
Water	157
B-factors	
Protein	38.59
Ligand/ion	51.88
Water	39.11
R.m.s. deviations	
Bond lengths (Å)	0.015
Bond angles (˚)	3.362

Table 4.3. X-Ray data collection and refinement statistics statistics for ^{C96}MBPC1$_4$ with 4 Cu(II). *Highest resolution shell is shown in parentheses.

Data collection	
Space group	$P12_11$
Cell dimensions	
a, *b*, *c* (Å)	48.285, 96.029, 97.826
α, β, γ (˚)	90.00, 101.31, 90.00
Resolution (Å)	96.03-2.40
R_{sym} or R_{merge} (%)	7.2 (35.1)
$I/\sigma I$*	11.8 (3.5)
Completeness (%)*	99.3 (98.8)
Redundancy*	3.8 (3.8)
Refinement	
Resolution (Å)	2.40
No. unique reflections	2218
R_{work}/R_{free}	0.207/0.277
No. atoms	
Protein	6888
Ligand/ion	8
Water	76
B-factors	
Protein	48.39
Ligand/ion	45.05
Water	40.02
R.m.s. deviations	
Bond lengths (Å)	0.013
Bond angles (˚)	3.225

Table 4.4. X-Ray data collection and refinement statistics for $^{C96/C81}$RIDC1$_4$ with 2 Cu(II). *Highest resolution shell is shown in parentheses.

Data collection	
Space group	*P*32
Cell dimensions	
a, *b*, *c* (Å)	93.275, 93.275, 48.828
α, β, γ (˚)	90.00, 90.00, 120.00
Resolution (Å)	2.10 (80.95)
R_{sym} or R_{merge} (%)*	7.8 (19.8)
$I/\sigma I$*	12.0 (6.3)
Completeness (%)*	99.0 (100.0)
Redundancy	4.2 (4.2)
Refinement	
Resolution (Å)	2.10
No. unique reflections	22322
R_{work}/R_{free}	0.232/0.302
No. atoms	
Protein	3470
Ligand/ion	2
Water	51
B-factors	
Protein	24.29
Ligand/ion	26.87
Water	21.41
R.m.s. deviations	
Bond lengths (Å)	0.017
Bond angles (˚)	3.475

Table 4.5. X-Ray data collection and refinement statistics for $^{C96/C81}$RIDC1$_4$ with 4 Cu(II). *Highest resolution shell is shown in parentheses.

Data collection	
Space group	$P6_122$
Cell dimensions	
a, *b*, *c* (Å)	53.51, 53.51, 275.27
α, β, γ (˚)	90.0, 90.0, 120.0
Resolution (Å)	3.03 (39.32)
R_{sym} or R_{merge} (%)*	5.1 (13.0)
$I/\sigma I$*	37.2 (19.0)
Completeness (%)*	99.9 (100.0)
Redundancy	15.6 (16.5)
Refinement	
Resolution (Å)	3.0
No. unique reflections	9291
R_{work}/R_{free}	0.263/0.370
No. atoms	
Protein	1716
Ligand/ion	2
Water	14
B-factors	
Protein	86.52
Ligand/ion	47.73
Water	30.0
R.m.s. deviations	
Bond lengths (Å)	0.012
Bond angles (˚)	2.957

Table 4.6. X-Ray data collection and refinement statistics for $^{A59/C81/C96}RIDC1_4$ with 3 Cu(II). *Highest resolution shell is shown in parentheses.

Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c* (Å)	60.808, 92.545, 85.802
α, β, γ (˚)	90.00, 90.00, 90.00
Resolution (Å)	1.70 (62.92)
R_{sym} or R_{merge} (%)*	n/a
$I/\sigma I$*	n/a
Completeness (%)*	n/a
Redundancy	n/a
Refinement	
Resolution (Å)	2.00
No. unique reflections	31643
R_{work}/R_{free}	0.278/0.362
No. atoms	
Protein	3464
Ligand/ion	3
Water	164
B-factors	
Protein	34.55
Ligand/ion	48.50
Water	31.76
R.m.s. deviations	
Bond lengths (Å)	0.016
Bond angles (˚)	3.423

Materials and Methods:

X-Ray Crystallography: All crystals were obtained by sitting-drop vapor diffusion at room temperature with growth conditions as follows:

Table 4.7. Crystallization conditions.

Protein crystal	Protein	Precipitant
Cu_2:^{C96}RIDC1$_4$	3 mM protein, 3 mM CuSO4, in 20 mM TRIS (pH 7)	25% PEG400
Cu_4:^{C96}RIDC1$_4$	3 mM protein, 5 mM $CuSO_4$, in 20 mM TRIS (pH 7)	15% PEG1900 MME
$Cu(I)_6$:^{C96}RIDC1$_4$	2 mM protein in 20 mM TRIS; 2 mM Cu(I) with 5 mM sodium ascorbate inside an anaerobic chamber	25% PEG1900 MME
Cu_4:(^{C96}MBPC1$_2$)$_2$	2 mM protein in 20 mM TRIS (pH 7)	25% PEG400
Cu_2:$^{C81/C96}$RIDC1$_4$	3 mM protein in 20 mM MOPS, 150 mM NaCl; 3 mM $CuSO_4$	30% methylpentandiol
Cu_4:$^{C81/C96}$RIDC1$_4$	4 mM protein in 20 mM MOPS, 150 mM NaCl; 6 mM $CuSO_4$	35% methylpentandiol
Cu_3:$^{A59/C81/C96}$RIDC1$_4$	3 mM protein in 20 mM MOPS, 150 mM NaCl; 3 mM $CuSO_4$	35% methylpentandiol

The sitting drops each consisted of 2 μL protein solution and 1 μL precipitant solution. The wells contained 500 μL precipitant solution. Crystals grew at room temperature typically within a two-day period and matured over a few weeks. X-ray diffraction data were collected at 100 K at Stanford Synchrotron Radiation Lightsource (SSRL). Diffraction data were processed using MOSFLM and SCALA on the CCP4 suite.[19] Structures were solved with molecular replacement by MOLREP[20] using a single chain of the structure of the ^{C96}RIDC1 (PDB ID:2IQ6) or of MBPC1 (PDB ID:2QLA) as a search model. Rigid-body, positional, thermal and TLS refinement was carried out using REFMAC,[21] along with iterative manual model

building and water/ligand/metal placement with COOT.[22] All structural figures were produced with PYMOL.[23] Ramachandran plots were determined using PROCHECK.[24]

Metal content analysis by ICP-OES:

Protein samples were incubated in 50 mM TRIS (pH 7), 150 mM NaCl in the presence of one equivalent of Cu(II) for 16 hours at room temperature on a shaker. The samples were run on desalting column (BioRad) to remove any unbound metal and the eluted protein was diluted to 2 mL by addition of Chelex-treated water and 69% HNO_3 to a final concentration of 3% HNO_3. ICP-OES standard metal solutions were prepared from 1,000-ppm certified metal stock solutions (Ricca) by serial dilutions, and a standard curve was prepared within a range of 0.1-10 ppm of metals in 3% nitric acid and Chelex-treated water. Data were collected on a Perkin-Elmer Optima 3000 DV ICP-OES spectrometer located at the Analytical Facility of the Scripps Institution of Oceanography. Because each protein monomer contains one Fe atom as part of the covalently linked heme molecule, the experimentally determined Cu:Fe ratios directly yielded the Cu:monomeric protein ratios. Wavelengths used for the detection of metal ions were Fe (234.349, 238.204, 239.562, 259.939, and 273.055 nm), and Cu (327.393, 324.752, 224.7, 213.597, 222.778, and 221.459).

Metal affinity competition assays:

A 1-mL solution containing 6.25 μM Fura-2 (Life Technologies) and 5.25 μM $^{C81/C96}RIDC_4$ was titrated in a 1-cm cuvette with 1-2 μL aliquots of 1 mM $CuCl_2$. After each addition, the sample was equilibrated for at least 5 minutes at room temperature, upon which the excitation scan was recorded on a Horiba Fluorolog 2

fluorimeter using an excitation range of 250-400 nm and detection at 510 nm. The change in excitation intensity at 343 nm was plotted versus Cu(II) concentration, and the resulting curve was fit alternatively to 1 × 4 Cu, 2 × 2 Cu or 4 × 1 Cu binding models using Dynafit[25] as described previously.[5]

Electron Paramagnetic Resonance Measurements:

EPR samples had been previously incubated with 1-2 equivalent of Cu(II) and the unbound metal was washed using a 10DG desalting column. The samples consisted of 300-500 μM of ^{C96}RIDC1$_4$, $^{A81/C96}$RIDC1$_4$, $^{C81/C96}$RIDC1$_4$ or $^{A59/C81/C96}$RIDC1$_4$ in 20 mM MOPS buffer (pH 7) with 150 mM NaCl. X-band spectra were obtained with a Bruker ELEXSYS E500 spectrometer equipped with a Bruker ER4131VT variable-temperature unit. Spectra were recorded at 125 K at a microwave frequency of 9873 MHz, 0.10 mT modulation amplitude and 0.63 mW power.

Analytical ultracentrifugation:

Samples for sedimentation velocity experiments contained 5 μM (^{C96}MBPC1$_2$)$_2$ in 20 mM MOPS (pH 7), 150 mM NaCl and either 5 μM Cu(II) or 1 mM EDTA. Measurements were made on a Beckman XL-I Analytical Ultracentrifuge (Beckman-Coulter Instruments) using an An-60 Ti rotor at 41,000 rpm for 400 scans per sample at 25 ˚C by detection at 415 nm. The data were processed using Sedfit.[26] Buffer viscosity (0.01002 poise), density (1.007 g/mL) and protein partial specific volume (0.7316 mL/g) were calculated at 25 ˚C with SEDNTERP (http://www.jphilo.mailway.com).

References

1 Lu, Y., Yeung, N., Sieracki, N. & Marshall, N. M. Design of functional metalloproteins. *Nature* **460**, 855-862 (2009).

2 Que, L. & Tolman, W. B. Biologically inspired oxidation catalysis. *Nature* **455**, 333-340 (2008).

3 Gloaguen, F. & Rauchfuss, T. B. Small molecule mimics of hydrogenases: hydrides and redox. *Chemical Society Reviews* **38**, 100-108 (2009).

4 Halcrow, M. A. & Christou, G. Biomimetic chemistry of nickel. *Chemical Reviews* **94**, 2421-2481 (1994).

5 Brodin, J. D. *et al.* Evolution of Metal Selectivity in Templated Protein Interfaces. *Journal of the American Chemical Society* **132**, 8610-8617 (2010).

6 Solomon, E. I., Chen, P., Metz, M., Lee, S. K. & Palmer, A. E. Oxygen binding, activation, and reduction to water by copper proteins. *Angewandte Chemie-International Edition* **40**, 4570-4590 (2001).

7 Klinman, J. P. Mechanisms whereby mononuclear copper proteins functionalize organic substrates. *Chemical Reviews* **96** (1996).

8 Holm, R. H., Kennepohl, P. & Solomon, E. I. Structural and functional aspects of metal sites in biology. *Chemical Reviews* **96**, 2239-2314 (1996).

9 Gray, H. B., Malmström, B. G. & Williams, R. Copper coordination in blue proteins. *JBIC Journal of Biological Inorganic Chemistry* **5**, 551-559 (2000).

10 Solomon, E. I., Baldwin, M. J. & Lowery, M. D. Electronic structures of active sites in copper proteins: contributions to reactivity. *Chemical Reviews* **92**, 521-542 (1992).

11 Solomon, E. I., Sundaram, U. M. & Machonkin, T. E. Multicopper oxidases and oxygenases. *Chemical Reviews* **96** (1996).

12 Solomon, E. I., Hare, J. W. & Gray, H. B. Spectroscopic studies and a structural model for blue copper centers in proteins. *Proceedings of the National Academy of Sciences* **73**, 1389-1393 (1976).

13 Solomon, E. I., Szilagyi, R. K., George, S. D. & Basumallick, L. Electronic structures of metal sites in proteins and models: contributions to function in blue copper proteins. *Chemical Reviews-Columbus* **104**, 419-458 (2004).

14 Vallee, B. L. & Williams, R. Metalloenzymes: the entatic nature of their active sites. *Proceedings of the National Academy of Sciences of the United States of America* **59**, 498 (1968).

15 Salgado, E. N., Faraone-Mennella, J. & Tezcan, F. A. Controlling Protein-Protein Interactions through Metal Coordination: Assembly of a 16-Helix Bundle Protein. *Journal of the American Chemical Society* **129**, 13374-13375 (2007).

16 Salgado, E. N., Lewis, R. A., Mossin, S., Rheingold, A. L. & Tezcan, F. A. Control of Protein Oligomerization Symmetry by Metal Coordination: C_2 and C_3 Symmetrical Assemblies through Cu^{II} and Ni^{II} Coordination. *Inorganic Chemistry* **48**, 2726-2728 (2009).

17 Su, C. C. *et al.* Bonding Properties of Copper(Ii)-Imidazole Chromophores - Structures and Electronic-Properties of Tetrakis Imidazole Copper(Ii) Complexes - Molecular-Structures of Cu(N-Methylimidazole)(4)(Clo4)(2) and [Cu(N-Methylimidazole)(4)(H2o)(2)]Cl-2(H2o). *Polyhedron* **14**, 3011-3021 (1995).

18 Salgado, E. N. *et al.* Metal-Templated Design of Protein Interfaces. *Proceedings of the National Academy of Sciences of the United States of America* **107**, 1827-1832 (2010).

19 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Biological Crystallography* **67**, 235-242 (2011).

20 Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallographica Section D-Biological Crystallography* **66**, 22-25 (2010).

21 Murshudov, G., Vagin, A. & Dodson, E. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* **D53**, 240-255 (1996).

22 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **D60**, 2126-2132 (2004).

23 DeLano, W. L. *The PYMOL Molecular Graphics System (http://www.pymol.org).* (2003).

24 Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - A Program To Check The Stereochemical Quality Of Protein Structures. *Journal of Applied Crystallography* **26**, 283-291 (1993).

25 Kuzmic, P. Program DYNAFIT for the analysis of enzyme kinetic data: Application to HIV proteinase. *Analytical Biochemistry* **237**, 260-273 (1996).

26 Schuck, P. A model for sedimentation in inhomogeneous media. I. Dynamic density gradients from sedimenting co-solutes. *Biophysical Chemistry* **108**, 187-200 (2004).

Chapter 5

Conclusions and Future Directions

Conclusions

Designing a protein with two or more thiols in order to form specific disulfide crosslinks is a very challenging endeavor since large unstructured polymeric assemblies can form when there is no directionality upon crosslinking. Using our metal templating strategy we were able to specifically guide the formation of the correct disulfide bonds on a macrocyclic protein assembly. The rigidity of the covalently crosslinked tetramers allowed us to place metals in the binding sites with minimal rearrangement of the metal-binding amino acids which we confirmed by obtaining the crystal structures of the Zn(II)-bound constructs of $^{38/C96}$RIDC1$_4$ and $^{C38/C96}$MBPC1$_4$. With these engineered scaffolds we are starting to explore how they can accommodate different transition metals at their core with the aim of getting reactive metal centers. Similarly, by utilizing a parallel approach we designed the $^{C81/C96}$RIDC1$_4$ protein scaffold to form a novel cryptand-like macrocyclic assembly that binds Zn(II) and Cu(II) with high affinity. Additionally, this protein forms a tetrameric construct in the periplasmic envelope of *E. coli*. This finding will allow us to subject it to selective pressure for the evolution of new biological functions in the cell.

After using Cu(II) to bind the engineered macrocyclic proteins of ^{C96}RIDC1$_4$ and $^{C81/C96}$RIDC1$_4$, we were able to study the metal's preferred coordination geometry and the influence of the hydrophobic interactions upon its coordination geometry. We found that copper prefers a thermodynamically stable 4His coordination motif involving histidines His59, His63, His73, and His77 with an *i, i+4* coordination motif.

Upon introduction of the hydrophobic interaction across *i1* we altered the coordination geometry of the metal and influenced a distortion on the tetragonal coordination. Furthermore, by using covalently crosslinked tetramers we are now able to control the coordination of the metal and we are currently studying a three coordinate copper site on $^{A59/C81/C96}RIDC1_4$ which has an open binding site that can be used as a catalytically active nitrite reductase mimic.

It has been suggested that metals were responsible for the evolution of some multi-subunit protein complexes in nature.[1,2] One study refers to the Multi-Copper Oxidase (MCO) family of enzymes, which are responsible for reactions such as the reduction of nitrite into nitric oxide (by nitrite reductase) or phenol oxidations (by laccase) (see Figure 5.1).[1] This family of enzymes has been proposed to have evolved from a single domain Cupredoxin enzyme which has a Type I Cu(II) metal site. After a domain duplication event, and numerous rounds of evolution, multi-domain proteins evolved to contain several copper centers (Type I, Type II, or Type III), which are either structural or catalytic, within the protein's interface. Importantly, similarly to the MeTIR approach that we have described in this dissertation, it has been proposed that metal binding interactions were the main driving force to guide these protein assemblies, and this family of enzymes evolved through evolution and optimization of the interfacial interactions within the protein.

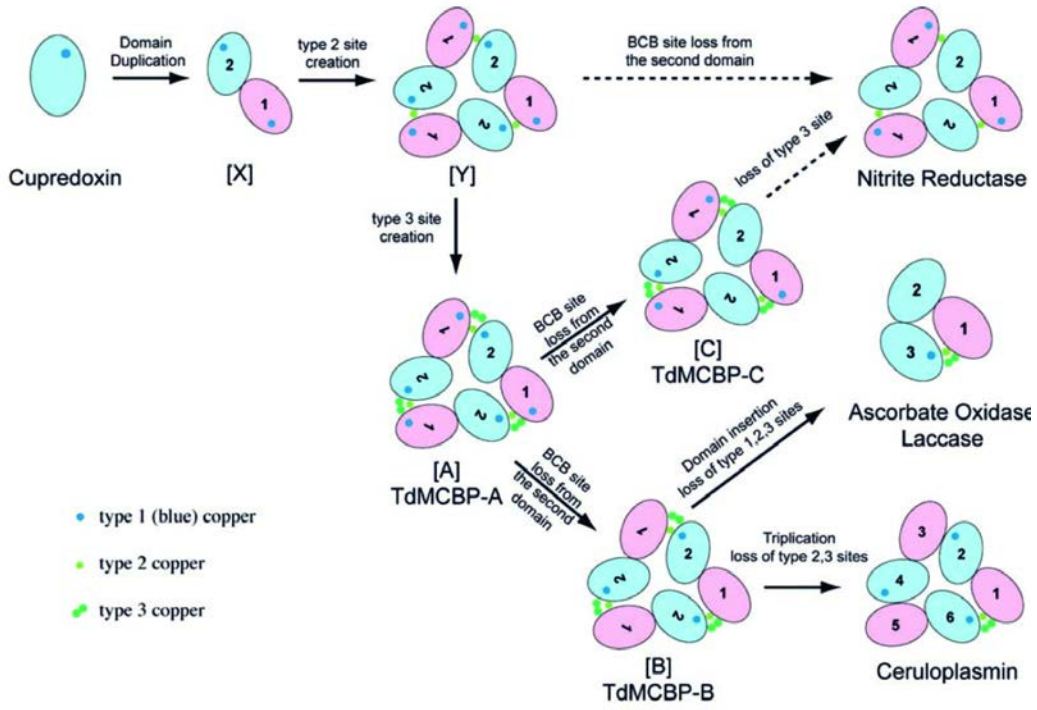

Figure 5.1. Schematic presentation of domain organization of the Multi Copper Binding Proteins (MCBPs). An oval shape indicates a Blue Copper Binding domain. Pink and blue colors show the domain classes IV and V. Blue and green dots indicate type 1 and 2/3 copper atoms, respectively. Solid arrows indicate the postulated evolutionary pathways. Conventional MCBPs are shown on the right. Five hypothetical proteins are shown with bracketed labels. Three of them ([X], [Y], and [A]) have BCB sites in both domains, whereas [B] and [C] have a BCB site only in the first and second domain, respectively. The hypothetical proteins [A], [B], and [C] have trinuclear interfacial copper binding sites (from reference 1).

Future Directions

We are currently exploiting the ability of the engineered $^{C81/C96}RIDC1_4$ self-assemble in the absence of metal to construct His to Ala mutants at positions 59, 63, 73, and 77 with the purpose of using these constructs to bind to Zn(II) and to Cu(II) with unsaturated coordination geometries. For example, we were able to obtain the crystal structure of the mutant protein of $^{A59/A77/C81/C96}RIDC1_4$ in the presence of Zn(II). This structure is a tetrameric protein linked through the engineered disulfide bonds of Cys96-Cys96 and Cys81-Cys81 that have been described in Chapter 3. However, the Zn(II) binding site is composed of two histidines at positions 63 and 73 bound to the metal and two water ligands which are hydrogen bonded through Asp74 (Figure 5.1b). The proposed coordination geometry was supposed to include Asp74 as a ligand but the position of the carboxylate is too distant to allow to metal-binding. This raises the prospect of engineering other amino acids at position 74 that will be able to bind to the Zn(II) ion. We are currently constructing other mutant versions that include His74 and Glu74 as ligands.

As part of this research we will also investigate whether the Zn(II) bound assemblies are able to be catalytic carbonic anhydrase mimics. To do this we will perform para-nitrophenol acetate hydrolysis assays with the Zn(II)-bound mutant constructs. After identifying the catalytically active ones we will also investigate if they are able to perform other type of hydrolysis reactions. Similarly, we will explore if these mutants bound to Cu(II) are also able to possess nitrite reductase activity.

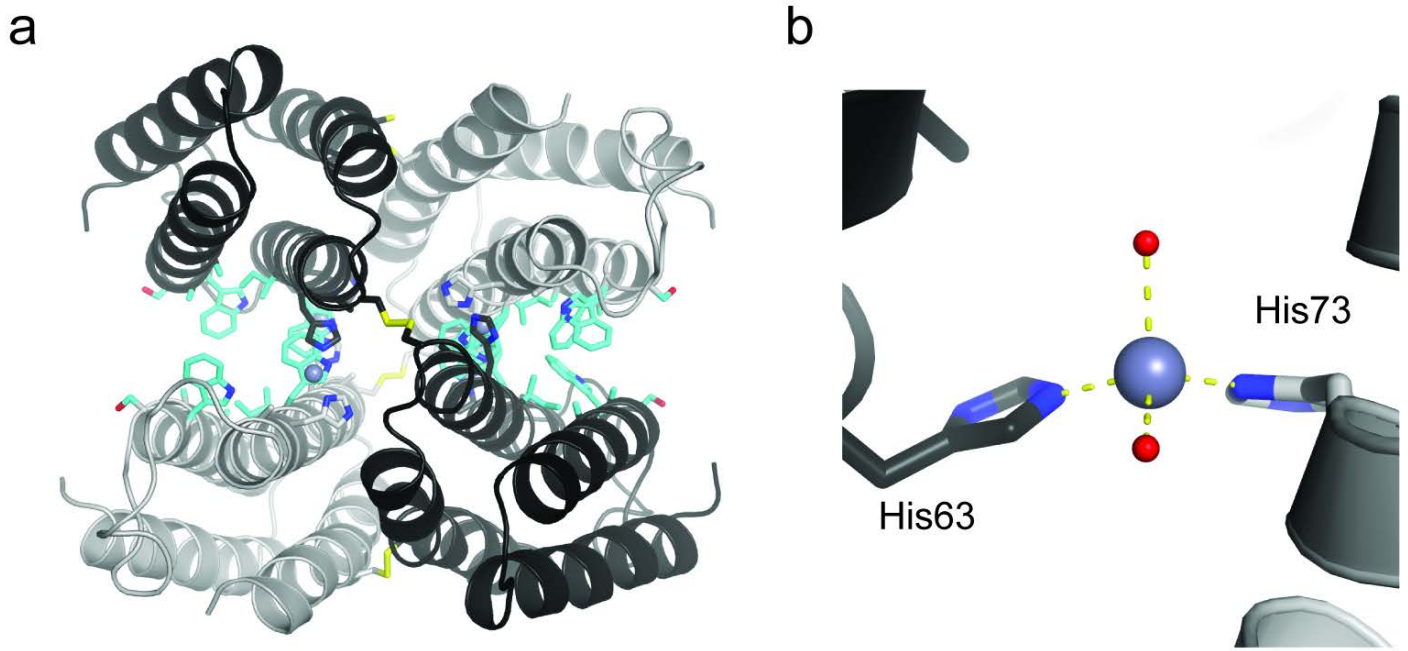

Figure 5.2. Crystal structure of Zn_4:$^{A59/A77/C81/C96}$RIDC1_4. (**a**) Topview of tetramer shown in dark and light gray cartoon. Hydrophobic amino acids are shown as cyan sticks. Zn(II) ions are shown as gray spheres coordinated by His63 and His73 (shown as sticks). (**b**) Close-up view of Zn(II) coordination to His73, His63, (shown as sticks) and two H_2O ligands (shown as red spheres).

Table 5.1. X-ray data collection and refinement statistics for Zn_4:$^{A59/A74/C81/C96}$RIDC1_4.

Data collection	
Space group	$P12_11$
Cell dimensions	
a, *b*, *c* (Å)	48.255, 90.774, 52.788
α, β, γ (˚)	90.00, 93.09, 90.00
Resolution (Å)	2.14
R_{sym} or R_{merge} (%)	3.8 (19.1)
$I/\sigma I$*	14.7 (4.0)
Completeness (%)*	88.6 (66.1)
Redundancy*	2.7 (2.6)
Refinement	
Resolution (Å)	2.30
No. unique reflections	17469
R_{work}/R_{free}	0.19667/0.27667
No. atoms	
Protein	3431
Ligand/ion	6
Water	66
B-factors	
Protein	47.2
Ligand/ion	54.4
Water	42.8
R.m.s. deviations	
Bond lengths (Å)	0.014
Bond angles (˚)	3.415

Materials and Methods

X-Ray Crystallography

All crystals were obtained by sitting-drop vapor diffusion. The sitting drops each consisted of 2 μL protein solution (2 mM) and 1 μL precipitant solution. The wells contained 500 μL precipitant solution which consisted of 35% methyl-pentanediol with 200 mM NaCl (pH 7.5). Crystals grew at room temperature typically within a two-day period and matured over a few weeks. X-ray diffraction data were collected at 100 K at beamline 12-2 of the Stanford Synchrotron Radiation Lightsource (SSRL) using 0.98-Å radiation. Diffraction data were processed using MOSFLM and SCALA on the CCP4 suite.[3] Structures were solved with molecular replacement by MOLREP[4] using the a single chain of the structure of the ^{C96}RIDC1 (PDB ID:2IQ6) as a search model. Rigid-body, positional, thermal and TLS refinement was carried out using REFMAC,[5] along with iterative manual model building and water/ligand/metal placement with COOT.[6] For the apo-$^{C81/C96}$RIDC1$_4$ and Zn_4:$^{A74/C81/C96}$RIDC1$_4$ structures, non-crystallographic symmetry (NCS) restraints (tight main-chain and medium side-chain restraints) were applied throughout the positional/thermal refinement process. For the Zn_4:$^{C81/C96}$RIDC1$_4$ structure, these restraints, as well as TLS refinement, were lifted in the final stages of refinement. For the same structure, anisotropic B-factor refinement was carried out. All structural figures and movies were produced with PYMOL.[7] Ramachandran plots were determined using PROCHECK.[8].

References

1 Nakamura, K., Kawabata, T., Yura, K. & Go, N. Novel types of two-domain multi-copper oxidases: possible missing links in the evolution. *FEBS Letters* **553**, 239-244 (2003).

2 Liu, C. & Xu, H. The metal site as a template for the metalloprotein structure formation. *Journal of Inorganic Biochemistry* **88**, 77-86 (2002).

3 Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallographica Section D-Biological Crystallography* **67**, 235-242 (2011).

4 Vagin, A. & Teplyakov, A. Molecular replacement with MOLREP. *Acta Crystallographica Section D-Biological Crystallography* **66**, 22-25 (2010).

5 Murshudov, G., Vagin, A. & Dodson, E. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst.* **D53**, 240-255 (1996).

6 Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallographica Section D-Biological Crystallography* **D60**, 2126-2132 (2004).

7 DeLano, W. L. *The PYMOL Molecular Graphics System (http://www.pymol.org).* (2003).

8 Laskowski, R. A., Macarthur, M. W., Moss, D. S. & Thornton, J. M. Procheck - A Program To Check The Stereochemical Quality Of Protein Structures. *Journal of Applied Crystallography* **26**, 283-291 (1993).

Appendix

Appendix 1

Calculation of the periplasmic concentration of $^{C81/C96}$RIDC1.

1) *Amount of protein expressed per liter of cell culture:*

The yield of $^{C81/C96}$RIDC1 expressed is ~ 1.5 mg/L culture and molecular weight of protein = 12,227 g/mol.

(0.0015 g/L culture) / (12227 g/mol) = 1.2×10^{-7} moles $^{C81/C96}$RIDC1 per L culture

2) *Number of E. coli. cells per L culture:*

The cell pellet wet weight ~ 7 g, and wet weight of one *E. coli.* cell = 9.5×10^{-13} g/cell.[1]

(7 g) / (9.5×10^{-13} g/ cell) = 7.4×10^{12} cells

3) *Volume of total periplasmic volume per L culture:*

The periplasmic volume of an *E. coli* cell ~ 0.2-0.4 fL.[2,3]

(0.2-0.4 $\times 10^{-15}$ L) $\times$ 7.4×10^{12} cells = 1.5-3 $\times 10^{-3}$ L of *E. coli* periplasm per L of culture

4) *Concentration of periplasmic $^{C81/C96}$RIDC1:*

(1.2×10^{-7} moles) / (1.5-3 $\times 10^{-3}$ L) = 40-80 $\times 10^{-6}$ M $^{C81/C96}$RIDC1

$\Rightarrow$ 10-20 $\times 10^{-6}$ M $^{C81/C96}$RIDC1$_4$

Appendix 2

Calculation of the redox potential of GSH-GSSG buffer solutions. The redox potential of the GSH-GSSG couple is dependent on the pH and concentration of the solution. The simplified Nernst equation for the GSH/GSSG couple is:[4]

$$E_h(in\ mV) = E_0 + 30 \times \log\left(\frac{[GSSG]}{[GSH]^2}\right)$$

where GSH and GSSG are molar concentrations and E_o is taken as -240 mV at pH 7.0.[5]

For a 19:1 molar ratio of GSH:GSSG at a total concentration of 1 mM:

$$E_h(in\ mV) = -240 + 30 \times \log\left(\frac{[50 \times 10^{-6}]}{[950 \times 10^{-6}]^2}\right) = -187mV$$

Similarly, for a 19:1 molar ratio of GSH:GSSG at a total concentration of 1 mM:

$$E_h(in\ mV) = -240 + 30 \times \log\left(\frac{[950 \times 10^{-6}]}{[50 \times 10^{-6}]^2}\right) = -77mV$$

Appendix 3

Use of the Electron Paramagnetic Resonance Spectrometer:

Setup:

First unplug the opening of the cavity and remove the cover; replace with appropriate tube holder.

Turn on 1-5:

1) Water (2 lines, open is parallel to the hose)

2) Turn on liquid nitrogen (open is parallel to the hose)

3) Turn on box #3 (chiller

4) Turn on box #4

5) Turn on #5 on power only and then the "on" button, (magnetic field)

Let the nitrogen gas flow for approximately 15 minutes to remove any water from the air inside the cavity.

On computer:

1) Click Acquisition, Connect to E500. Panel shows up.

2) Acquire new experiment, Exp. Name.

3) Check the box that says "Temperature unit"

4) Acquisition: Spectrometer configuration

Misc. Choose temp controller 4131VT

Proportionality band: 95%

Integration time: 60.00 s

Derivation time: 1.80 s

5) Click Apply, then close.

Attach the liquid nitrogen tank to the instrument

1) Fill dewar with liquid nitrogen

2) Insert metal heater rod into dewar slowly and seal with cork slowly (take glass cork out of venting outlet). Add any grease if necessary to make a proper seal. Clamp down cork using springs.

3) Attach transfer arm to connector at cavity (glass on glass, and then attach with black spring cap). Let rest on post, make sure it is well-greased.

4) Move dewar gently under transfer arm attachment, place down lightly on transfer arm and attach springs.

To connect to the VTU (Variable Temperature Unit) go back to computer:

1) Click on the thermometer symbol

2) Click on the temperature tab (it should be at 295 K)

3) Check "use VTU"

4) Set gas flow to 535 l/h

5) Turn on both heater at 30% and liquid nitrogen heater at 2% (make sure both stay on at the same time)

6) Once at 295 K change temp to 2%

7) Watch out for heater power not over 10, it should be around 1%

8) To bring down to temperature change temperature to 125 K and then change evaporator heater power to 23% and keep lowering stepwise.

9) Check for leaks, if there are bubbles, wait for them to freeze using a iceal seal around the cork

10) Wait for stabilization ~15 minutes

11) Error messages might pop up, but this is fine, make sure the temperature does not change too much; if it does there might be leaks.

12) Change the temperature and power until you reach the desired temperature, keep checking for leaks with water bottle

13) Wait until temperature is stable for ~15 minutes.

Sampling:

Controls:

- Modulation amplitude: 0.1
- Attenuation at 40
- Check g-factor value and set it at 2.0077000 for DPPH, collect for 100 gauss
- Control Si-P alloy, g-value set at 1.99891
- Temperature should be set to 125 K
- Click "Send parameters to experiment" button

For liquid protein samples:

- Freeze slowly: Take 1-2 minutes to freeze and clean the outside of the tube with a Kimwipe before inserting into cavity
- Insert into cavity and secure carefully (too tight may break the EPR tube)
- For copper, scan around 2500-3500, modulation amplitude of 3 is fine
- Attenuation: 20-30 dB
- Modulation frequency: 100.00 kHz
- Receiver gain: 60.0

Tuning:

1) Open tuning panel

2) Click **"Tune"**

3) Turn **OFF Reference Arm**

4) Adjust **Frequency** to get dip close

5) Click **"Operate**

6) Set Attenuation to **30 dB**

7) Adjust **Frequency** so **Diode current = 0** and **lock offset = 0**

8) Check down to **10 dB** using **Iris**

9) Turn **ON Reference Arm**

10) **Attenuation= 60 dB**

11) Adjust **BIAS** so **Diode Current= 200**

12) Change **Attenuation = 30dB**

13) Adjust **Signal Phase** so at **Maximum**

14) Use **Iris** so that the **Diode Current = 200**

15) Check down to **10 dB** using **Iris**

*****When taking samples in/out be sure to be in tune mode*****

Shut Down:

1) Take out sample and cavity holder and clean out ice around the cavity

2) Go to the thermometer symbol and set the temperature to 250 K, evaporator heater power to 10%

3) Temperature to 295 K, 2% evaporator heater power

4) Heat at 330 K for 15 minutes

5) Turn off heater power and evaporator heater by checking the Off checkmark

6) Use hair dryer to melt ice around connection between dewar and transfer arm carefully. Remove transfer arm from sample chamber attachment by disconnecting the black jacket, then gently wiggling.

7) Bring transfer arm to counter, let dewar heater sit beside container.

If restarting the instrument, refill the liquid nitrogen dewar and go through the procedure described before. Remember to bring the heating coil back to room temperature by using the blow dryer in the cool setting. Also be sure that there is no water condensed on any of the parts attached. If there is, bring everything to room temperature, dry, and then restart. Reapply grease if necessary

If shutting down:

8) Uncheck VTU box

9) Disconnect instrument from spectrometer

10) Turn off 5-1 (make sure water and gas lines are off)

Appendix 4

Matlab code for EasySpin[6] simulation for Cu(II) binding to 3 nitrogens

```
% Cu2+ with four imidazole nitrogens
%============================================================
% Simulation of the EPR spectrum of a S=1/2 Cu2+ coordinated by
% 3 imidazoles, so that the couplings to 3 nitrogens are
% resolved.

clear
Sys.Nucs = 'Cu,N,N,N';
Sys.g = [2.06 2.265];
A_Cu = [0.0015 0.0185]; % in cm^-1
A_N = [0.00155 0.00155]; % in cm^-1
Sys.A = [A_Cu;A_N;A_N;A_N]*30e3; % cm^-1 -> MHz
Sys.Apa = [0 0 0; 0 1 0; 0 1 1; 0 1 2]*pi/2;
Sys.lwpp = [0 1];

Exp.mwFreq = 9.3923;
Exp.Range = [250 350];
Exp.Temperature = 125; % temp in Kelvin
Opt.Method = 'perturb';
Opt.nKnots = [61 0];
Sys.HStrain = [25 25 25]; % [along x, along y, along z], in MHz
%Sys.gStrain = [0.005 0.01 0.005];
%Sys.AStrain = [10 10 30]; % in MHz
[Field,Spec] = pepper(Sys,Exp,Opt);


[B1,spec1,pars1] = eprload('R19638Cpolyfit3.DTA');
hold on;
plot(Field*10,Spec*300)
plot(B1,spec1,'g');
hold off;

```

References

1 Neidhardt, F. C. *et al. Escherichia coli and Salmonella typhimurium. Cellular and molecular biology. Volumes I and II.* (American Society for Microbiology, 1987).

2 Azam, T. A., Iwata, A., Nishimura, A., Ueda, S. & Ishihama, A. Growth phase-dependent variation in protein composition of the Escherichia coli nucleoid. *Journal of Bacteriology* **181**, 6361-6370 (1999).

3 Stock, J. B., Rauch, B. & Roseman, S. Periplasmic space in Salmonella typhimurium and Escherichia coli. *Journal of Biological Chemistry* **252**, 7850-7861 (1977).

4 Jones, D. P. in *Methods in Enzymology* Vol. Volume 348 (eds Sies Helmut & Packer Lester) 93-112 (Academic Press, 2002).

5 Rost, J. & Rapoport, S. Reduction potential of glutathione. *Nature* **201**, 185-186, doi:10.1038/201185a0 (1964).

6 Stoll, S. & Schweiger, A. EasySpin, a comprehensive software package for spectral simulation and analysis in EPR. *Journal of Magnetic Resonance* **178**, 42-55 (2006).