UNIVERSITY OF CALIFORNIA SAN DIEGO

SAN DIEGO STATE UNIVERSITY

Design of Metal-Controlled Protein-Protein Interactions

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

Doctor in Philosophy

in

Chemistry

by

Brian Maniaci

Committee in Charge:

University of California San Diego

Professor Katja Lindenberg
Professor Navtej Toor

San Diego State University

Professor John J. Love, Chair
Professor Christal Sohl
Professor Robert Zeller

2019

This dissertation of Brian Maniaci is approved, and is acceptable in quality and form for publication on microfilm and electronically:

_____

_____

_____

_____

_____
Chair

University of California San Diego

San Diego State University

2019

DEDICATION

This dissertation is dedicated to my grandparents, parents, and sister and brother for their incredible support, continuous encouragement, and endless love.

TABLE OF CONTENTS

LIST OF ABREVIATIONS

^{o}C	degree Celsius
Å	Angstrom
AUC	Analytical Ultracentrifugation
CCM	CheckMyMetal (Online Server)
CD	Circular Dichroism
FPLC	Fast Protein Liquid Chromatography
HPLC	High Pressure Liquid Chromatography
K_d	Equilibrium Dissociation Constant
kDa	kilo-Dalton
NMR	Nuclear Magnetic Resonance
MALS	Multi-Angle Light Scattering
MCD	Metal-Controlled Dimer
MeTIR	Metal-Templated Interface Redesign
MID1	Metal Interface Design 1
mM	millimolar
MW	Molecular Weight
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PISA	Proteins, Interfaces, Structures and Assemblies (Online Server)
rMeTIR	Reverse Metal-Templated Interface Redesign
RMSD	Root-mean-square-deviation
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

SE	Sedimentation Equilibrium
SV	Sedimentation Velocity
SEC	Size Exclusion Chromatography
SEC-MALS	Size Exclusion Chromatography with Multi-Angle Light Scattering
µL	microliter
Small molecules:	
EDTA	ethylenediaminetetraacetic acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
IPTG	Isopropyl β-D-1-thiogalactopyranoside
MOPS	3-morpholinopropane-1-sulfonic acid
TRIS	tris(hydroxymethyl)aminomethane
Proteins:	
Binding Partner A	Mutant of Gβ1 (L12H, T16L, V29E, Y33E, N37L)
Binding Partner B	Mutant of Gβ1 (L12R, T16L, V29H, Y33H, N37L)
Binding Partner C	Mutant of Gβ1 (L12H, T16V, V29E, Y33E, N37V)
cyt cb_{562}	Cytochrome cb_{562}
Ferritin	human H-maxiferritin
Gβ1	β1 domain of Protein G
Gβ1-3H	Mutant of Gβ1 (L12H, V29H, Y33H)
Gβ1-M2	Mutant of Gβ1 (E15V, T16L, T18I, N37L)
M2_11H	Mutant of Gβ1 (T11H, E15V, T16L, T18I, V29H, Y33H, N37L)
MCD_C1	Mutant of Gβ1 (L12H, T16L, V29H, Y33H, N37L)
MCD_C2	Mutant of Gβ1 (L12H, E15V, T16A, T18I, V29H, Y33H, N37A)

MCD_M2	Mutant of Gβ1 (L12H, E15V, T16L, T18I, V29H, Y33H, N37L)
MonA	Mutant of Gβ1 (T16F, T18A, V21E, T25L, K28Y, V29I, K31R, Q32A, Y33L, N35K, D36A, N37Q)
MonB	Mutant of Gβ1 (A23I, E27A, K28D, K31A, N35A, D40K, Y45A, D47E)

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

Chapter 2 is reproduced in full with permission from Maniaci, Brian; Lipper Colin H.; Anipindi, Deepthi L.; Erlandsen, Heidi; Cole, James L.; Stec, Boguslaw; Huxford, Tom; Love, JJ. Design of High-Affinity Metal-Controlled Protein Dimers. *Biochemistry*. **2019** 58 (17), 2199-2207. Copyright 2019 American Chemical Society. The dissertation author was the primary investigator and author of this material.

Chapter 3, in part is currently being prepared for submission for publication of the material. Maniaci, Brian; Mealka, Matt; Onofrei, David; Stengel, Dillan; Huxford, Tom; Love, JJ. Design of Metal-Controlled Protein Heterodimers (*in preparation*). The dissertation author was the primary investigator and author of this material.

I would like to thank Dr. Love for his support, scientific input, mentoring, and also for giving me the freedom to pursue my own scientific ideas. I would also like to that everyone in the Love group for their help and advice: Markus, Benjamin, Melissa, Peter, Ariana, Eric and Johnathan. I would also like to thank Dr. Sohl, Dr. Huxford, and Dr. Stumph for their scientific and professional advice and sharing their laboratory equipment. I would also like to thank Dr. Lindenberg, and Dr. Toor, and Dr. Zeller for being on my graduate committee and asking insightful questions. I would also like to thank Dr. Stec for determining the structures of the metal-controlled dimers. I would also like to thank Matt Mealka, David Onofrei, and Dillan Stengel for determining the structures of the metal-controlled heterodimers. To Josh Swider, Braden Silva, and Sean Maddox; thank you for being great friends. Finally, I would like to thank my family and friends for their support over my graduate studies. My work was supported by the California Metabolic Research Foundation and Dr. Love's research grants. I am grateful to be a recipient of an Arne N.

Wick Predoctoral Research Fellowship from the California Metabolic Research Foundation and an ARCS scholar.

VITA

Education

2009 B.S. Biochemistry
California Polytechnic State University, San Luis Obispo

2016 M.A Chemistry
San Diego State University

2019 Ph.D. Chemistry
University of California San Diego and San Diego State University

Employment History

2010-2010 Laboratory Technician
TechNeal, Pomona, CA

2010-2013 Research Associate
VivoScript, Costa Mesa, CA

2013-2019 Teaching Assistant, San Diego State University, San Diego, CA

Publications

Maniaci, B; Mealka, Matt; Onofrei; David; Stengel, Dillan; Huxford, Tom; Love, JJ., Design of Metal-Controlled Protein Heterodimers. (*in preparation*)

Maniaci, B., Lipper, C. H., Anipindi, D. L., Erlandsen, H., Cole, J. L., Stec, B., Huxford, T., Love, J. J., Design of High-Affinity Metal-Controlled Protein Dimers. *Biochemistry*, **2019** 58 (17), 2199-2207.

Yomogida, K.; Wu, S.; Baravati, B.; Avendano, C.; Caldwell, T.; **Maniaci, B**.; Zhu, Y.; Chu, C. Q. Cell penetrating recombinant Foxp3 protein enhances Treg function and ameliorates arthritis. *Biochem. Biophys. Res. Commun.* **2013**, *434* (2), 263–267.

Yomogida, K.; Chou, Y.; Pang, J.; Baravati, B.; **Maniaci, B.**; Wu, S.; Zhu, Y.; Chu, C. Q. Streptavidin suppresses T cell activation and inhibits IL-2 production and CD25 expression. *Cytokine* **2012**, *58* (3), 431–436.

Awards and Fellowships

2015-2019 Achievement Rewards for College Scientists (ARCS) Foundation Award

2017-2018 Laurel Amtowter Scholarship

2017-2018 Harry E. Hamber Memorial Scholarship

2017-2018 Arne N. Wick Predoctoral Research Fellowship

ABSTRACT OF THE DISSERTATION

Design of Metal-Controlled Protein-Protein Interactions

by

Brian Maniaci

Doctor of Philosophy in Chemistry

University of California San Diego, 2019

San Diego State University, 2019

Professor John J. Love, Chair

The field of protein design strives to engineer new molecules that interact in a specific, controlled manner to form novel functional complexes. Engineered proteins that generate specific complexes upon the addition of an exogenous agent, such as metal ions, will likely be integral elements of these efforts. Molecular control over protein assembly and disassembly is possible through the introduction of novel metal-binding sites at precise locations in protein complexes. These methods have primarily generated metal-mediated and metal-controlled

homodimers. The overall goals of my dissertation projects were to increase our understanding of metal-mediated associations and improve the utility of metal-controlled dimerization systems.

As a model system, we used the β1 domain of Protein G (Gβ1) as a scaffold to build novel dimer complexes. A published report from a different research group described the generation of a Gβ1 variant that forms a constitutive, symmetric homodimer. This small, well-characterized dimer was an excellent starting model for the design of metal-controlled dimers as considerable biophysical analysis was performed on this variant, and the high resolution three-dimensional structure was solved with X-ray crystallography. We used structure-based rational design to engineer histidine residues at the dimeric interface, which ultimately resulted in high-affinity, metal-controlled protein-protein interactions. Almost all of our design attempts, which contained various interfacial modifications, were shown to form metal-controlled homodimers that bind with moderate to extremely high affinity.

In addition to the successful design of high-affinity symmetric homodimers, we were also interested in generating novel heterodimers that also bind with relative high-affinity. To achieve this we reengineered one of the metal binding sites of a metal-controlled homodimer with oppositely charged side-chains to generate intermolecular protein salt-bridges between an arginine residue on one monomer and two glutamic acid residues on the other. The crystal structures revealed both the designed salt bridges and tetrahedral zinc coordination site. We also demonstrated that binding of the heterodimer complex, which contains the designed salt bridges, was disrupted upon addition of increasing concentrations of sodium chloride. The novel, high-affinity metal-controlled homo- and hetero-dimer proteins could potentially be used as highly effective building blocks for novel biomaterials.

Chapter 1: Introduction

1.1 Protein-Protein Interface Design

Interactions between proteins are essential for life. Some examples of protein-protein interactions includes assembly of macromolecular structures, transfer of genetic information, activation and inhibition of function, and molecular recognition. Since in any given cell there are thousands of different proteins, it is highly important that protein-protein interactions maintain and promote cellular functions. Improving our understanding of what drives and governs protein-protein interactions are important components in advancing basic and applied research. This sought-after goal could lead to the creation of new biomaterials and biotechnological tools. In this work, structure-guided design was used to generate high-affinity metal-controlled dimers.

Our understanding of what drives and governs protein-protein interactions is aided through the structural analysis and classification of naturally occurring protein-protein interactions. Atomic-level structural models are essential for understanding the structure-function relationship of proteins. The Protein Data Bank (PDB) is a public archive of macromolecular structures determined using NMR, X-Ray Crystallography, theoretical modeling, and Cryo-Electron Microscopy[1]. Structural models from the PDB may provide insights into the diversity of proteins, protein-protein interfaces, and assemblies[2]. Protein structures in the PDB might highlight the overall three-dimensional shape, different conformational states, metal ions and/or cofactors, the binding orientation and intermolecular contacts between subunits and other proteins.

The dynamic association and dissociation of protein complexes plays an important role in many biological systems[3]. A large fraction of proteins exist in multimeric forms, but some exist independently of other proteins. Monomers are characterized by their non-self-interacting surfaces. Proteins can interact and assemble into multimeric assemblies by a process called protein oligomerization. Evolutionary and functional advantages for protein oligomerization might include reduced genome size, the generation of new protein surfaces, and improved stability[4]. Discrete structures of oligomeric proteins have a defined composition and subunit stoichiometry. Oligomeric proteins are composed of identical (homo-oligomeric) or non-identical (hetero-oligomeric) subunits. The association between protein subunits can vary in duration and strength.

Protein complexes are held together by the sum of many non-covalent interactions spread over an extensive binding surface[5]. Non-covalent interactions commonly found in protein-protein interactions might include van der Waals, hydrophobic and electrostatic interactions, and hydrogen bonding. The process of protein assembly requires sufficient enthalpic and entropic contributions, which must overcome the loss of entropy associated with complex formation[6]. Specific protein-protein contacts may influence the affinity and specificity of protein-protein interactions.

During complex formation, two or more protein molecules interact and generate a new protein-protein interface. This newly generated interface is characterized by interacting and nearby residues, which are involved in stabilizing the protein-protein interaction. Most oligomeric protein complexes in the PDB have a contact area in the 500 – 1500 $Å^2$ range[3]. The protein-protein interface of complexes typically have a high degree of physical and chemical complementarity[7]. The analysis of numerous protein-protein interfaces have not led to

conclusive chemical and physical properties common in all transient and permanent protein complexes[8,9].

Oligomeric proteins are common in nature, comprising approximately 1/3 of all proteins[10]. The most prevalent complexes in the PDB are homodimers and heterodimers[5]. Homodimers are abundant in nature, often have large non-polar protein-protein interfaces, and tend to form permanent complexes[8]. Many homodimers form symmetric structures which might confer structural or functional advantages[11]. Approximately 75 % of all oligomers found in nature are characterized as homo-oligomers[10]. Heterodimers were found to have fewer hydrophobic groups and more polar groups at the protein-protein interface when compared with homodimers[9].

The association of homodimers and heterodimers use certain recognizable binding motifs at the protein-protein interface. The binding motifs of homodimers and heterodimers were classified using regular and non-regular secondary elements[12]. The four classes of secondary elements found at protein-protein interfaces included α, β, mixed αβ, and non-regular regions (loops, turns, and other non-rotameric confirmations). Analysis of the secondary structures elements in many protein-protein interfaces may provide additional insights for rationalizing and designing protein binding motifs.

Regular secondary structure segments in homodimers contribute almost half of the protein-protein interface, while in heterodimers they contribute about one-third. Secondary structure pairing at the interface shows that only α-helical interactions are twice as abundant for homodimer interfaces (22.4 %) compared with heterodimer interfaces (10.9 %)[12]. In comparison, only β-strand interactions are equally favored between homodimers (8.8 %) and heterodimers (8.4 %)[12]. The most common secondary structural segment at the interface of

homodimer (40.3 %) and heterodimers (46.3 %) was the non-regular region interacting with short segments of helices/strands[12].

1.2 Metal Coordination in Proteins

Metal ions are associated with an estimated 30 - 40% of all proteins[13,14]. The presence of a metal in a protein structure or protein-protein interface might imply important structural and/or functional roles. Metal ions may promote catalytic activities, conformational changes, activation and inhibition of function, and formation of transient protein-protein interactions[15]. Metal ions found at the protein-protein interface might form and stabilize natural and engineered protein-protein interactions[16,17]. In addition, metal coordination sites are responsive to external stimuli since they may be broken or formed through changes in pH or metal chelating agents.

Metal binding sites in proteins utilize the chemical functionalities of certain amino acids. Late-first row transition metals may form metal-ligand bonds with the side chains of histidine, cysteine, glutamate, and/or aspartate[18]. Metal coordination may occur by the nitrogen atom in histidine and the sulfur group in cysteine or by carboxylate ions in aspartate and glutamate. In the Lewis acid-base theory, transition metal ions are Lewis acids and nitrogen, sulfur, and oxygen atoms are Lewis bases. The Lewis acids and Lewis bases are an electron pair acceptor and donor, respectively. A metal coordination complex is a product of the Lewis acid-base reaction, where the central metal ion is bound by ligands from the macromolecule or environment.

Transition metal binding generally occurs via a few amino acid sequence motifs involving histidine and cysteine residues[19]. A bidentate ligand arrangement occurs when two metal binding amino acids are in close spatial proximity and donate two pairs of electrons to the Lewis acid. The “chelate effect” describes the close proximity of two metal-binding ligands,

which have enhanced affinity for metal ions. Bidentate metal chelation sites in proteins are engineered by specific amino acid sequence motifs: HX_3H (H is histidine and X is any amino acid) in an α-helix, HXH in a β-strand, and HX_2H in a reverse β-turn[19]. Novel metal binding sites have been engineered to generate proteins that self-assemble and form oligomeric structures[19,17,20,21]. Unfortunately, there is no canonical amino acid sequence or surface topology that will universally lead to a predictable oligomerization state[22].

Metal-ions and cofactors found at the protein-protein interface may mediate the assembly of protein oligomers and macromolecular assemblies. Within the PDB, analysis of metal-mediated oligomeric proteins may provide valuable insights into the diversity and positioning of metal ions and cofactors at the protein-protein interface[15]. This survey found approximately 22 % of oligomeric proteins are associated with a transition metal or cofactor[15]. Of these oligomeric protein, around 8 % contain a late-first-row transition metal (Mn, Fe, Co, Ni, Cu, Zn) or cofactor (iron-sulfur clusters and heme coordination complexes) at the protein-protein interface[15]. Metal ions and cofactors were characterized as interfacial if they were directly coordinated to residues of different molecules and share a buried surface. In the subset of oligomeric proteins, zinc and heme were found in approximately 40 % and 15 - 20 % of all cases, respectively[15]. The most predominant late-first-row transition metal or cofactor in the protein-protein interface is zinc, followed by heme.

It was found that only a small subset of oligomeric proteins contain a late-first-row transition metal or cofactor at the protein-protein interface[15]. Of this subset, a majority of oligomeric proteins are likely to contain binding sites for zinc or heme[15]. The prevalence of zinc in oligomeric proteins might be due to its high bioavailability and functional properties. Zinc is the second most abundant trace metal found in eukaryotic organisms[23]. The bioavailability of

zinc in *E. coli* is approximately 0.1 mM[24]. Important functional properties of zinc (II) include high binding stability, Lewis acid, flexible coordination geometry, and redox inactivity[18]. Zinc (II) binding sites often display distorted tetrahedral coordination geometry[15,25].

Zinc coordination sites often plays important structural and/or functional roles in proteins[26]. Structural roles of zinc vary depending upon whether it stabilizes tertiary or quaternary structures. Zinc plays an important functional role in the catalytic mechanism of zinc-dependent enzymes. The tetrahedral coordination geometry of the enzyme active site is characterized by an open coordination sphere exposed to an aqueous environment[25]. The fourth coordinating ligand is usually occupied by a water molecule or hydroxyl ion. The specificity of the biochemical reaction depends upon the arrangement of amino acids in the second and third coordination spheres of the active site.

1.3 Protein-Protein Interactions Driven By Metal Coordination

Selective incorporation of metal binding sites onto monomeric scaffolds have generated both metal-mediated and metal-controlled interactions[20,17,21]. Metal-mediated interactions are loosely defined as protein-protein interactions occurring in the presence of metal ions. In some metal-mediated interactions the protein-protein interaction occurs in the absence and presence of metal ions[20]. In comparison, the interface of metal-directed protein assemblies are devoid of favorable side chain interactions that might promote protein oligomerization in the absence of metal ions[17]. For a metal-directed interaction, the metal ion is responsible for bringing together non-interacting partners. Protein oligomerization of metal-directed interactions are driven by metal coordination and stabilized by favorable interfacial contacts[17].

The Tezcan group at U.C.S.D. demonstrated the practicality of metal-directed protein assembly of cytochrome B $_{562}$ (*cb*$_{562}$), a monomeric four-helix bundle[17]. The crystal lattice of

wild-type *cb*$_{562}$, demonstrates the α-helical positions of different molecules being in close spatial proximity, despite being monomeric in solution up to millimolar concentrations[17]. A variant called His^4-*cb*$_{562}$ was generated by incorporating a di-histidine motif at *i* and *i* + 4 positions on the α-helix at the crystal contact points of the individual *cb*$_{562}$ molecules. A metal chelating sites generated by the di-histidine motif enhances affinity for metal ions[17]. The incorporation of metal binding sites at the crystal contacts of *cb*$_{562}$ provides a stimuli responsive element to modulate and control protein-protein interactions. Metal-directed protein self-assembly of His^4-*cb*$_{562}$ was confirmed by adding EDTA or lowering the pH below 6, which disrupts the metal-controlled complex.

The supramolecular arrangement of His^4-*cb*$_{562}$ was controlled by the metal coordination geometry of zinc (II)[22]. At high protein concentrations, the dominant quaternary structure of His^4-*cb*$_{562}$ was a tetramer. The crystal structure of the 4 Zn: 4 His^4-*cb*$_{562}$ assembly shows a distorted tetrahedral coordination environment with metal binding sites from three His^4-*cb*$_{562}$ molecules[22]. The designed di-histidine sites of His^4-*cb*$_{562}$ were found to participate in intermolecular zinc (II) metal coordination and protein oligomerization. The metal-directed assembly of His^4-*cb*$_{562}$ was due to the strength and directionality of metal-ligand coordination bonds[17].

Oligomeric assemblies of His^4-*cb*$_{562}$ are influenced by the coordination geometry of different late-first row transition metals at the protein-protein interface[22,21]. Different oligomeric assemblies of His^4-*cb*$_{562}$ were controlled by the non-tetrahedral coordination preferences of nickel (II) and copper (II). Zinc (II) coordination yields a symmetric tetramer, while the square planar preferences of copper (II) and nickel (II) leads to an antiparallel dimer and a parallel

trimer, respectively. Different supramolecular assemblies of His^4-*cb*$_{562}$ protein can be generated through different late-first row transition metals.

A collaborative research project between Tezcan and the Khulman group at U.N.C. Chapel Hill introduced a new metal-directed design strategy. This strategy, called Metal Templated Interface Redesign (MeTIR), provides a route for engineering *de novo* protein interfaces and metal-independent protein-protein interactions. This design strategy involved an initial metal templating step followed by computational redesign of the protein-protein interface. This research project used the previously generated His^4-*cb*$_{562}$ variant, which formed a metal-directed protein assembly[17]. The computational redesign of the 4 Zn: 4 His^4-*cb*$_{562}$ complex focused on redesigning the protein interface by incorporating hydrophobic residues at selected positions while maintaining the metal binding sites.

Two redesigned variants of His^4-*cb*$_{562}$ (RIDC-1 and RIDC-2) were redesigned with an increase in surface hydrophobicity to form a well-packed core. Design strategies between RIDC-1 and RIDC-2 involve incorporation of hydrophobic bulk in the designed interface, with RIDC-2 having a smaller and less-than-optimally packed hydrophobic core[21]. A significant increase in surface hydrophobicity might improve the stability and binding affinities of *de novo* protein-protein interactions[21]. The monomer-dimer dissociation constants of RIDC-1 and RIDC-2 were 25 and 55 μM, respectively[21]. Two redesigned variants of His^4-*cb*$_{562}$ (RIDC-1 and RIDC-2) self-assemble in both the presence and absence of metal ions. A MeTIR design strategy converted a non-self-associating protein into a metal-independent protein-protein interaction with a binding affinity in the low micromolar range.

The MeTIR collaborative research project further explored metal-mediated oligomerization of the redesigned protein interfaces[20]. Metal-mediated oligomerization of

RIDC-1 differs due to the preferred coordination geometry of zinc (II) and copper (II) ions. Upon the addition of zinc (II), the His^4-cb_{562} and RIDC-1 variants both form tetrameric structures. Structural analysis of the metal-templated His^4-cb_{562} and RIDC-1 interfaces suggest the computationally redesigned interface has improved stability in terms of metal binding and oligomerization[21]. In comparison, a precise dimer of RIDC-1 was generated due to the square planar coordination of copper (II). The distortion of the copper (II) coordination environment suggests the RIDC-1 has a preference for zinc (II) coordination.

The Tezcan group also developed a reverse Metal-Templated Interface Redesign (rMeTIR) strategy by engineering a self-assembling protein into a metal-directed protein assembly[27]. The self-interacting surfaces of Human H-maxiferritin (Ferritin) assemble into a symmetric, cooperative 24-meric protein cage[28]. This metal-templated interface redesign of the Ferritin involved three distinct steps including the installation of metal coordination motifs, crystal structure determination, followed by the elimination of key interfacial interactions in the Ferritin scaffold[27]. The rMeTIR strategy generated a monomeric Ferritin molecule with metal coordination motifs. Assembly of the 24-meric protein cage was dependent on copper (II) coordination[27]. The chemically inducible nature of the redesigned protein might enable passive encapsulation of guest molecules in the Ferritin protein cage.

Using a computational design approach the Khulman group expanded the utility of metal-templated design strategies[20]. A Metal Interface Design (MID1) strategy was used to redesign a monomeric protein for metal-mediated self-assembly. Both MeTIR and MID1 design strategies generated a *de novo* protein-protein interaction. The MeTIR and rMeTIR design strategies involved three separate steps; an initial metal templating event, crystal structure determination,

and rationally redesign of the protein-protein interface[21,27]. The novelty of the MID1 approach is the simultaneous design of metal-binding sites and protein-protein contacts.

A monomeric protein scaffold was designed to promote a symmetric metal-mediated interaction with novel metal-binding sites and favorable intermolecular contacts[20]. This design strategy generated a zinc (II)-mediated symmetric homodimer with high binding affinity and orientation specificity[20]. This *de novo* protein-protein interaction had a binding affinity lower than 30 nM[20]. The computationally designed protein formed a symmetric homodimer in the presence and absence of zinc (II). A metal interface design strategy might be one methodology to increase affinity and selectivity of protein binding[20]. Another advantage of the MID1 design strategy is the ability to rationally control the orientation specificity of a design protein-protein interaction.

Metal-directed assembly strategies had limited success in generating metal-mediated interactions involving different binding partners[29]. Wild-type Ubiquitin has an open coordination sphere at histidine 68, which might be used for intermolecular zinc (II) coordination[29]. Using computational protein design, a monomeric scaffold was re-engineered with three histidine residues and favorable contact points at the protein-protein interface. The metal-mediated interaction involves three engineered histidine residues from the re-engineered scaffold and histidine 68 from Ubiquitin, representing a 3:1 zinc coordination arrangement model. The metal-mediated interaction between Ubiquitin and the designed protein contained a single metal coordination site, while other metal-directed strategies have multiple zinc coordination sites[17,20,21]. One metal coordination site provides fewer energetically favorable interactions between the different proteins.

In the design of a metal-mediated heterodimer, a single metal coordination site was designed between Ubiquitin and the re-engineered scaffold. The designing protein scaffold, called Speltzer, binds Ubiquitin with a Kd of 20 μM and 68 μM in the presence and absence of zinc, respectively[29]. NMR and mutagenesis suggests the binding arrangement between Spelter and Ubiquitin does not occur via the predicted 3:1 coordination model. The researchers suggested a 2:2 coordination arrangement (i.e., each protein contributes two ligands for metal coordination) might generate a more energetically favorable metal-mediated interaction[29].

The metal-directed assembly of monomeric proteins provides a method for generating novel protein-protein interactions through metal coordination. This metal-directed assembly has several advantages over other methodologies to generate novel protein-protein interactions. The strength of metal-ligand bonds are stronger than non-covalent bonds found at the protein-protein interface[17]. The preferred coordination geometry of different metal ions may change the arrangement and oligomeric states of engineered proteins[21]. Metal-ligand bonds can be broken or formed or broken through external ligands or changes in pH. In addition, interfacial metal coordination sites might provide intrinsic reactivity and have implications for rational enzyme design[30]. Collectively, these advantages make metal-directed assembly a viable methodology to generate high affinity protein-protein interactions. We plan on applying these attributes of metal-directed assembly to improve the binding affinity of a previously generated *de novo* protein-protein interaction[31].

1.4 Streptococcal Protein G

The β1 domain of Protein G (Gβ1) is the protein used in my structure-guided design project. Protein G is displayed on the cell surface on group G *Streptococcus*. The protein binds to the Fc region of immunoglobulin G (IgG), which might help the organism evade host defense

systems[32]. Gβ1 consists of 56 amino acids which folds into a four-stranded antiparallel β-sheet with an overlapping α-helix. Gβ1 has high thermal stability with a melting temperature of 87 ^{o}C[32]. It is important to highlight that wild-type Gβ1 contains no histidine or cysteine residues, since they may act as metal ligands.

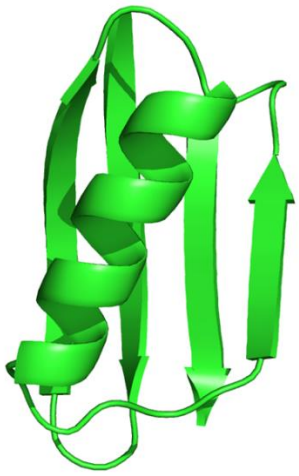

Figure 1.1: Cartoon representation of the β1 domain of Streptococcal protein G (Gβ1). PDB entry 1PGA[34].

In dilute systems, wild-type Gβ1 remains monomeric since intermolecular hydrogen bonds do not unable to overcome hydrogen bonding between the molecule and solvent[33]. Protein oligomerization may be favored at higher concentrations common in NMR and X-Ray crystallography. Favorable intermolecular associations of Gβ1 generally involve the outer β-strands rather than the α-helix. Gβ1 molecules have an intrinsic propensity to interact via unsatisfied hydrogen bond donors and acceptors on the outer β-strands (β-2 strand or β-3 strand). The intermolecular β-strand associations of Gβ1 are defined as head-to-tail (β-2 strand and β-3' strand) or head-to-head orientation (β-2 strand and β- 2' strand). The crystal structures of wild-type Gβ1 (PDB entry 1PGA) showed intermolecular β-strand interactions are stabilized by four to five hydrogen bonds in the head-to-tail association[34,35]. In comparison, a head-to-head orientation promoted inverted protein symmetry and six hydrogen bonds stabilizing the

intermolecular interaction[33]. In self-assembling systems, intermolecular interactions of Gβ1 may be stabilized by non-polar protein-protein interfaces[33,36,37].

Gβ1 is a model protein for studying the relationships between sequence, stability, and complex formation[33,37,36,31]. Modifications to Gβ1 have resulted in formation structures of side-by-side dimers, domain-swapped dimer, and an intertwined tetramer[33,36,37,38]. In the side-by-side dimers, hydrophobic mutations promote dimer formation through rearrangement of tyrosine at position 33 in a head-to-head association[33,37]. The dissociation constant of the side-by-side dimers were in the low micromolar range. Destabilizing mutations near the protein core of Gβ1 resulted in an intertwined tetramer[36]. Analysis of the intertwined tetramer showed how a single point mutation promoted further structural rearrangement into a domain-swapped dimer[38]. Core and surface residue mutations can change the overall topology and oligomerization state of Gβ1 variants[33,36,37,38].

1.5 Previous Studies in the Love Group

A *de novo* protein-protein interface strategy was explored using Gβ1 as a protein design scaffold. In a previous study, computational docking and amino acid sequence design were used to generate a heterodimeric complex[31]. Using a predefined protein-protein binding mode, *de novo* interactions of Gβ1 were restricted to a helix-to-helix arrangement. Backbone docking was followed by the use of the ORBIT suite of design algorithms to determine optimal amino acid substitutions at the newly created interface[39]. In addition to the hydrophobic interface, cross-dimer polar interactions were generated in the structural model between Monomer A and Monomer B. These mutations resulted in the creation of a Gβ1 heterodimer referred to as MonA/MonB. Heteronuclear Nuclear Magnetic Resonance (NMR) techniques were used to

verify that the two monomers did form a complex although analytical ultracentrifugation revealed the binding constant to be quite weak ~300 μM[31].

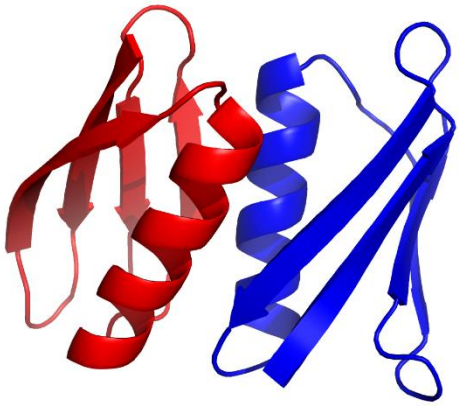

Figure 1.2: Computational Model of MonA/MonB Complex. The MonA/MonB complex are docked in a helix-to-helix arrangement[31].

In an attempt to increase the binding affinity of the designed MonA/MonB heterodimer, rational design was used to 'bootstrap' stronger intermolecular interactions between the monomers. In one example, pairs of metal-coordinating histidine residues were introduced at select interfacial positions using the computationally derived helix-face-to-helix-face model. This design did not give rise to increased heterodimer affinity but did result in the formation of MonA homo-complexes. The mutations that gave rise to the MonA homo-complexes are the following: MonA(A32H, A36H), MonA(F16H, A18H), and MonA(E21H, L25H). In three cases the mutations resulted in MonA variants that are monomeric in the absence of metal, yet form mixtures of higher-order complexes upon the addition of zinc.

The main structural feature of MonA homo-complexes MonA(A32H, A36H) and MonA(E21H, L25H) is that dimer and complex formation occurs via extension of the four-stranded β-sheet. The crystal structures of the MonA homo-complexes also reveal that zinc ions are, for the most part, bound by the engineered histidine pairs, but not at the experimentally

determined interface. Thus we believe that metal binding to histidine residues functions to induce intermolecular interactions that significantly promote crystallization. The experimentally determined binding arrangement of MonA/MonA variants are similar to other dimeric forms of Gβ1.

1.6 Design of Metal-Controlled Protein Dimers

The metal-binding variants of MonA and another research group's symmetric homodimer[33] called Gβ1-M2 (PDB entry 3FIL) both form head-to-head assemblies. Gβ1-M2 is a constitutive dimer generated by directed evolution from wild-type Gβ1[33]. The head-to-head association promotes a symmetric homodimer with favorable cross-strand and helical interactions[33]. Four thermally stabilizing mutations E15V, T16L, T18I, and N37L promote dimer formation by enlarging the hydrophobic core of the protein-protein interface. The Gβ1-M2 molecules show inverted symmetry and are self-contained allowing for formation of discrete oligomers. Gβ1-M2 provides an excellent starting model for the design of metal-controlled dimers. In addition to providing a protein design scaffold, GB1-M2 also forms a high-affinity symmetric homodimer with a binding dissociation constant in the range of 1 μM[33]. Structure-guided design was used to incorporate metal-binding sites along the Gβ1-M2 dimer interface.

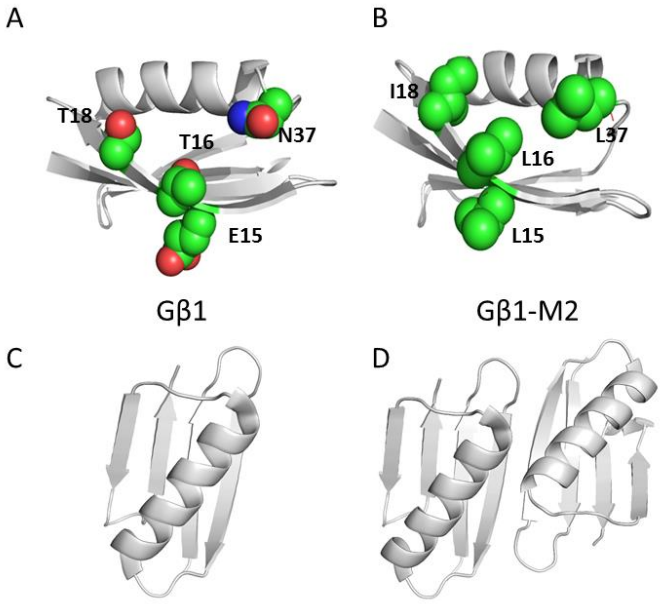

Figure 1.3: Comparison of Gβ1 and Gβ1-M2. A. Residues found in wild-type Gβ1 (PDB entry 1pga)[34]. B. Four thermally stabilizing mutations in Gβ1-M2 (PDB entry 3FIL)[33]. C. The biological assembly of wild-type Gβ1 is a monomer. D. The biological assembly of Gβ1-M2 is a symmetric dimer.

Using a structure-guided approach, novel metal binding sites were engineered on Gβ1 promoting formation of metal-controlled homodimers. In addition, we used hydrophobic mutations found in Gβ1-M2, a dimeric scaffold of Gβ1, to help promote a symmetric homodimer[33]. We incorporated three histidine residues at the protein-protein interface of Gβ1-M2, which allowed for the generation of two metal binding sites. The symmetric nature and small interface of Gβ1-M2 limited the sites for engineering novel metal binding sites. Histidine residues on the alpha helix and loop generated an intermolecular metal coordination site. The designed metal binding sites binding zinc and function to drive dimer formation.

Another goal of the project was to study the effects of interfacial modifications on binding affinity. After the initial generation of a metal-controlled homodimer, we further modified the interface to generate additional metal-controlled homodimers with varying binding affinity. Modifications to the metal-controlled homodimer included reversion of hydrophobic

interfacial residues back to wild-type Gβ1 and/or conservative replacements with a smaller hydrophobic residue. The metal-controlled dimers are generated by zinc (II) coordination with hydrophobic residues contributing to dimer stability and affinity. Modifications to the interface of Gβ1 led to the generation of three high-affinity metal-controlled homodimers. These variants were further modified to contain a GC linker at the C-terminus enabling fluorescent labeling of the metal-controlled dimers. The binding affinity of the fluorescently labeled metal-controlled dimers were assessed using analytical ultracentrifugation with fluorescence detection.

In a new research aim, our goal was to re-engineer a metal-controlled dimer into a metal-controlled heterodimer. This was accomplished by replacing one of the metal binding centers with an intermolecular salt bridge that functions to contribute favorable binding energy. The metal controlled association between two different proteins involved an intermolecular metal coordination site and salt bridges, where both sites contribute favorable binding energy and promote the formation of a heterodimer. The newly designed binding partners contains a portion of the metal coordination site, requiring both partners for the generation of an intermolecular metal coordination site. The designed binding partners replaced the other metal binding site with oppositely charged residues. The generation of "networked" salt bridges may provide heterodimer specificity and promote favorable interactions at the designed protein-protein interface. One binding partner contains the positively charged arginine residue, while the other binding partner has negatively charged glutamic acid residues. The binding affinity of the metal-controlled dimers was assessed using analytical ultracentrifugation.

The work discussed here addresses two fundamental applications of protein-protein interface design: incorporation of novel functionality using structure-guided design and re-engineering of protein-protein interactions. Using structure-guided design we generated a metal-

controlled homodimer. We then modified the interface to generate two additional metal-controlled homodimers with variable binding affinity. After structural validation of the metal-controlled homodimers, we redesigned the interface to generate two protein pairs. The metal interface of the protein pairs requires metal coordination to drive the designed protein-protein interaction. The metal-controlled heterodimer has two sites for interactions: an intermolecular metal coordination site and designed salt bridges.

1.7 References

(1) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Research*. **2000**, *28* (1), 235–242.

(2) Krissinel, E.; Henrick, K. *J. Mol. Biol.* **2007**, *372* (3), 774–797.

(3) Nooren, I. M. A.; Thornton, J. M. *J. Mol. Biol.* **2003**, *325* (5), 991–1018.

(4) Marianayagam, N. J.; Sunde, M.; Matthews, J. M. *Trends Biochem Science*. **2004**, 29 (11) 618–625.

(5) Jones, S.; Thornton, J. M. *Prog. Biophys. Mol. Biol.* **1995**, *63* (1), 31–65.

(6) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204* (1), 155–164.

(7) Chakrabarti, P.; Janin, J. *Proteins Struct. Funct. Genet.* **2002**, *47* (3), 334–343.

(8) Jones, S.; Thornton, J. M. *Proc. Natl. Acad. Sci.* **1996**, *93* (1), 13–20.

(9) Conte, L. Lo; Chothia, C.; Janin, È. *J. Mol. Biol*. **1999**. 285 (5), 2177-2198

(10) Goodsell, D. S.; Olson, A. J. *Annu Rev Biomol Struct*. **2000**. 29 (1) 105-153

(11) Swapna, L. S.; Srikeerthana, K.; Srinivasan, N. *PLoS One* **2012**, *7* (5).

(12) Guharoy, M.; Chakrabarti, P. *Bioinformatics* **2007**, *23* (15), 1909–1918.

(13) Andreini, C.; Bertini, I.; Cavallaro, G.; Holliday, G. L.; Thornton, J. M. *J. Biol. Inorg. Chem.* **2008**, *13*, 1205–1218.

(14) Andreini, C.; Bertini, I.; Rosato, A. *Acc. Chem. Res.* **2009**, *42* (10), 1471–1479.

(15) Song, W. J.; Sontz, P. A.; Ambroggio, X. I.; Tezcan, F. A. *Annu. Rev. Biophys.* **2014**, *43*,

409–431.

(16) Salgado, E. N.; Radford, R. J.; Tezcan, F. A. *Acc. Chem. Res.* **2010**, *43* (5), 661–672.

(17) Salgado, E. N.; Faraone-mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2007**, *129*, 13374–13375.

(18) Laitaoja, M.; Valjakka, J.; Jänis, J. *Inorg. Chem.* **2013**, *52* (19), 10983–10991.

(19) Haymore, F. H. A. and Zhang, Ji-Hu. I. *Trends in Biotechnology* **1991**, 12 (52) 1796–1797.

(20) Der, B. S.; MacHius, M.; Miley, M. J.; Mills, J. L.; Szyperski, T.; Kuhlman, B. *J. Am. Chem. Soc.* **2012**, *134* (1), 375–385.

(21) Salgado, E. N.; Ambroggio, X. I.; Brodin, J. D.; Lewis, R. A.; Kuhlman, B.; Tezcan, F. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (5), 1827–1832.

(22) Salgado, E. N.; Lewis, R. A.; Mossin, S.; Rheingold, A. L.; Tezcan, F. A. *Inorg. Chem.* **2009**, *48* (7), 2726–2728.

(23) Bertini, I.; Decaria, L.; Rosato, A. *J. Biol. Inorg. Chem.* **2010**, *15* (7), 1071–1078.

(24) Finney, L. A.; Halloran, T. V. O. *Science.* **2011**, *931* (2003), 931–937.

(25) McCall, K. A.; Huang, C.; Fierke, C. A. *J. Nutr.* **2000**, *130* (5), 1437S–1446S.

(26) Auld, D. S. *BioMetals* **2001**, *14* (3–4), 271–313.

(27) Huard, D. J. E.; Kane, K. M.; Akif Tezcan, F. *Nat. Chem. Biol.* **2013**, *9* (3), 169–176.

(28) Liu, X.; Theil, E. C. *Acc. Chem. Res.* **2005**, *38* (3), 167–175.

(29) Der, Bryan S., Jha, Ramesh K., Lewis, Steven M., Thompson,, Peter M., Guntas, Gurkan., Khulman, B. *Proteins* **2013**, *81* (7), 1245–1255.

(30) Der, B. S.; Edwards, D. R.; Kuhlman, B. *Biochemistry* **2012**, *51* (18), 3933–3940.

(31) Huang, P.-S.; Love, J. J.; Mayo, S. L. *Protein Sci.* **2007**, *16* (12), 2770–2774.

(32) Gronenborn, Angela M., Filpula, David R. Essig, Nina Z., Achari, Aniruddha., Whitlow, Mark., Wingfield, Paul T., Clore, G. M. *Science*. **1991**, 253 (5020), 657–661.

(33) Thoms, S.; Max, K. E. A.; Wunderlich, M.; Jacso, T.; Lilie, H.; Reif, B.; Heinemann, U.; Schmid, F. X. *J. Mol. Biol.* **2009**, *391* (5), 918–932.

(34) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721–4729.

(35) Barchi, J. J.; Grasberger, B.; Gronenborn, a M.; Clore, G. M. *Protein Sci.* **1994**, *3* (1), 15–21.

(36) Frank, K.; Dyda, F.; Dobrodumov, A.; Gronenborn, A. M. *Nat. Struct. Biol.* **2002**, *9* (11), 877–885.

(37) Jee, J.; Byeon, I. J. L.; Louis, J. M.; Gronenborn, A. M. *Proteins Struct. Funct. Genet.* **2008**, *71* (3), 1420–1431.

(38) Byeon, I. J. L.; Louis, J. M.; Gronenborn, A. M. *J. Mol. Biol.* **2003**, *333* (1), 141–152.

(39) Dahiyat, B. I.; Mayo, S. L. *Science.* **1997**, *278* (5335), 82–87.

Chapter 2: Design of High-Affinity Metal-Controlled Protein Dimers

2.1 Abstract

The ability to precisely control protein complex formation has high utility in the expanding field of biomaterials. Driving protein−protein binding through metal−ligand bridging interactions is a promising method of achieving this goal. Furthermore, the capacity to precisely regulate both complex formation and dissociation enables additional control not available with constitutive protein complexes. Here we describe the design of three metal-controlled protein dimers that are completely monomeric in the absence of metal yet form high-affinity symmetric homodimers in the presence of zinc sulfate. The scaffold used for the designed dimers is the β1 domain of streptococcal protein G. In addition to forming high affinity dimers in the presence of metal, the complexes also dissociate upon addition of EDTA. Biophysical characterization revealed that the proteins maintain relatively high thermal stability, bind with high affinity, and are completely monodisperse in the monomeric and dimeric states. High-resolution crystal structures revealed that the dimers adopt the target structure and that the designed metal-binding histidine residues successfully bind zinc and function to drive dimer formation.

2.2 Introduction

The ongoing production of novel biomaterials will continue to entail the design of synthetic proteins that self-assemble into complexes with a specific structure. Significant progress in this area has come from robust computational methods[1−5] and powerful directed evolution screens[6−8], both inevitably enhanced through rational design. The ability to chemically control protein assembly and disassembly is highly useful in this expanding field. This level of molecular control can be achieved through the substitution of metal-binding residues at select interfacial amino acid positions. This method was originally utilized to increase protein stability[9]

and has since been used to drive specific protein complex formation. Considerable research in this area has been advanced by the Tezcan group, which initially exploited crystal packing contacts observed for the protein cyt cb562 to introduce metal-binding residues between closely packed helices in the crystal lattice, which resulted in the formation of a tetramer (16-helix bundle protein) upon addition of Zn(II)[10]. Additional mutagenesis studies of that variant provided further insights into the particular forces that give rise to supramolecular specificity for this system as well as natural protein complexes[11]. In subsequent metal-mediated designs, hybrid coordination motifs were engineered between introduced histidine residues and residues conjugated with non-natural synthetic metal ligands[12−14]. The resulting complexes adopt discrete structures upon addition of metal and demonstrate that novel structures can be generated without the need for extensive reengineering of interfacial protein surfaces.

In another seminal project, Der et al. used Rosetta-Match to introduce two-residue zinc-binding sites on the surface of known monomeric protein scaffolds followed by exhaustive computational assessment of symmetric interface starting structures and rigorous filtering based on interfacial physicochemical parameters[15]. Biophysical characterization of variants of the best design, termed MID1, provided proof that metal-mediated high affinity protein−protein interactions were successfully generated using this primarily computational approach. In a joint project, members of the Kuhlman and Tezcan groups used a metal-templated interface redesign (MeTIR) approach to reengineer the noncomplementary surfaces of the cyt cb562 variant to form discrete complexes in the presence of various metals[16]. These efforts gave rise to a number of protein complexes that form in the presence and absence of metal ligands and provide unique perspectives and interesting conjecture with regard to how natural protein complexes may have originally evolved. A reverse approach to metal-driven complex formation was also reported for

the human H-maxiferritin protein, which naturally forms a constitutive 24-subunit cagelike complex that functions to encapsulate Fe(II) and O_2[17]. The reverse method, termed reverse metal-templated interface redesign (rMeTIR), was used to mutate interfacial residues to first abolish complex formation in the absence of metal and, more importantly, to then drive complex formation of cagelike complexes in the presence of divalent copper. The resulting engineered variants were also shown to effectively encapsulate various exogenous compounds under physiological conditions, which is not possible with wild-type H-maxiferritin.

Here we describe three engineered metal-controlled dimers (MCDs) that are completely monomeric in the absence of metal yet form monodisperse, high-affinity dimers in the presence of metal (zinc). The dimers that form upon addition of metal also dissociate in the presence of the metal-chelating compound EDTA. The overall design of the MCDs was inspired by the combination of three distinct protein design approaches: (1) an earlier purely computational *de novo* dimer design, (2) subsequent incorporation of key elements from a constitutive dimer generated by a different research group via directed evolution[18,19], and (3) structure-based rational design.

2.3 Materials and Experimental Details

2.3.1 Gene Synthesis, Protein Expression, and Purification.

The scaffold used for the MCD variants is the β1 domain of streptococcal protein G (Gβ1). The β1 domain [Protein Data Bank (PDB) entry 1PGA] consists of amino acids 228−282 of immunoglobulin protein G. The genes for all Gβ1 mutant variants and Gβ1-M2 (PDB entry 3FIL) were synthesized using standard polymerase chain reaction (PCR) methods and cloned into the NdeI (5′-end) and EcoRI (3′-end) sites of the pet21a vector (Novagen, San Diego, CA). Additional variants of Gβ1 and Gβ1-M2 were constructed using PCR-based methods and cloned

into the NdeI (5′-end) and EcoRI (3′-end) sites of pet21a. Correct gene sequences were verified using standard DNA sequencing. *Escherichia coli* strain BL21(DE3) was chemically transformed with the corresponding plasmids and grown to an OD600 of ∼0.8. Protein expression was induced for ∼3 h upon addition of isopropyl β-D-1-thiogalactopyranoside at a final concentration of 1 mM. Cells were pelleted by centrifugation at 6000 rpm for 10 min and stored at −80 °C. A freeze−thaw protein extraction cycle was utilized and included incubation of the cell pellet on ice for 30 min and room temperature for 15 min, followed by a dry ice/ethanol bath for 10 min, and repeated three times. The cell pellet was gently resuspended in ∼20 mL of phosphate-buffered saline (pH 6.8) for 60 min. The sample was centrifuged at 10000 rpm for 30 min, and acetonitrile was added to the resulting supernatant [to 30% (v/v)] to induce precipitation of impurities. The resulting supernatant was lyophilized to remove acetonitrile and water. The lyophilized sample was resuspended in water and purified using a Varian 10 μm C8 high-performance liquid chromatography (HPLC) preparative reverse-phase column with a linear 1%/min acetonitrile/water gradient containing 0.1% trifluoroacetic acid. The purified proteins were lyophilized, and the resulting dry protein was suspended in water. The pH was adjusted to ∼7 and buffered with 20 mM Tris, 100 mM sodium chloride, and 200 ppm sodium azide (pH 7.0). Protein purity was assessed using sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE), and protein concentration determined by absorbance at 280 nm using calculated extinction coefficients. The proteins were concentrated using centrifugal concentration to approximately 2 mM (∼12 mg/mL) and stored at −20 °C.

2.3.2 Size Exclusion Chromatography.

Stock concentrations of the three MonA homodimer variants were diluted to 500 μM and analyzed using size exclusion chromatography (SEC). SEC was performed using a FPLC

instrument (GE Healthcare) with ultraviolet (UV) detection at 280 nm; 100 μL of each sample was injected onto a 25 mL Superdex75 10/300 GL size exclusion column (GE Healthcare), and the elution volumes were compared with that of an AB-fusion tandem protein. The AB fusion functions as a standard for dimer formation, which corresponds to the expression of MonA fused to MonB at the level of the gene. The two genes are expressed in tandem with a $(Gly4Ser)_2$ intervening linker. The three MonA variants were injected individually in (1) 20 mM Tris and 100 mM sodium chloride (pH 7.0) and (2) 20 mM Tris, 100 mM sodium chloride, and 1 mM zinc sulfate (pH 7.0).

2.3.3 SEC−MALS Characterization.

Stock proteins of Gβ1, MCD variants, and controls were diluted to 1 mM in buffer containing the following four sets of solution components: (1) 20 mM Tris and 100 mM sodium chloride (pH 7.0), (2) 20 mM Tris, 100 mM sodium chloride, and 5 mM EDTA (pH 7.0), (3) 20 mM Tris, 100 mM sodium chloride, and 500 μM zinc sulfate (pH 7.0), and (4) 20 mM Tris, 100 mM sodium chloride, and 1 mM zinc sulfate (pH 7.0). Size exclusion chromatography was performed using a FPLC instrument (GE Healthcare) with UV detection at 280 nm. The FPLC instrument was connected inline with multiangle light scattering (Wyatt miniDAWN TREOS detector system). ASTRA software (Wyatt Technologies) was used to analyze elution peak(s) and determine the molecular weight calculated from the light scattering data.

2.3.4 Circular Dichroism (CD).

CD data were collected on an Aviv 420 instrument using a 2 mm path-length cuvette. Protein concentrations were ∼25 μM with buffer consisting of 20 mM MOPS, 10 mM sodium chloride, and 1.5 M guanidinium hydrochloride (pH 7.0). Metal-induced dimerization was measured with 20 mM MOPS, 10 mM sodium chloride, 1.5 mM guanidinium hydrochloride, and

25 μM zinc sulfate (pH 7.0). Far-UV scans confirmed that the Gβ1 and MCD variants have a mixed α-helical/β-sheet secondary structure with minima at 208 and 222 nm. The CD signal at 222 nm was monitored as a function of temperature and fit to a two-state unfolding model. Folded protein and unfolded states were measured between 25 and 95 °C.

2.3.5 X-ray Crystallography.

For crystallization purposes, the MonA homodimer variants and the MCD variants were expressed and purified as described above and concentrated to a minimum 2 mM. MCD variants were screened in the absence and presence of zinc sulfate. To study the effects of metal-controlled association, excess zinc sulfate was added to the MonA and MCD stock solutions before crystallization. A Mosquito LPC nanoliter crystallization robot was used for high-throughput screening of crystal formation conditions. Crystals were obtained by vapor diffusion at room temperature from hanging drops. The optimized crystallography conditions for the variants are listed in Table 2.8. Optimized crystallization buffer with 50% glycerol was used for storage and transport in liquid nitrogen. Data were collected from two sources: San Diego State University home source (Rigaku RU-H39) with a rotating anode generator and Berkley Lab Advanced Light Source (ALS-5.0.1). Data were processed using the program HKL2000[20]. Structures were determined by molecular replacement with the program PHASER[21] in the CCP4 suite[22]. The protein Gβ1 domain (PDB entry 1PGA) was used as a molecular replacement model. The refinement was carried out with REFMAC[23] using several cycles of the least-squares refinement fit interspaced with manual editing of the structure in COOT[24]. Several of the structures required inclusion of disordered residues and careful examination of the metal ion-binding sites as they required inclusion of counterions such as chlorine. The refinement procedure usually converged with fewer than 10 macrocycles of the refinement and manual

inspection. The structures were analyzed using PISA[25], SFCHECK[26], and PROCHECK[27]. The final validated structures (by internal PDB audit) were deposited as PDB entries 6NL6, 6NL7, 6NL8, 6NL9, 6NLA, and 6NLB. Some additional analysis of the metal ion-binding sites was performed using the CheckMyMetal[28] (CMM) Web site. Descriptive software tools such as Proteins, Interfaces, Structures, Assemblies (PISA) and CCM provide an estimate for binding energies and a secondary validation tool for organization of the lattice and the binding sites. Data collection and refinement statistics are listed in Table 2.9.

2.3.6 Fluorescent Labeling of MCD Variants.

MCD variants were fluorescently labeled to increase their sensitivity for analysis using an AUC instrument equipped with fluorescence detection. To generate fluorescently labeled MCD variants, the C-terminus of each variant was extended with a -Gly-Cys extension for directed fluorescent conjugation to the cysteine side chain with fluorescein 5-maleimide (Thermo Scientific). The MCD variants with the -Gly-Cys extensions were expressed and purified using the same methods described above. Following initial purification, the lyophilized protein was buffered with 20 mM sodium phosphate, 100 mM sodium chloride, and 10 mM EDTA (pH 7.3). A 10 mM fluorescein 5-maleimide stock solution was prepared using dimethylformamide (DMF) as the solvent. Fluorescent labeling entailed using a 20-fold molar excess of fluorescein 5-maleimide dye over the concentration of each MCD protein. The optimal labeling efficiency was found when the final volume of DMF was Optimal labeling efficiency occurred when the final volume of DMF was < 10% (v/v). The labeling reaction mixture was incubated at room temperature for 1 h, and followed by overnight incubation at 4 °C. The labeled proteins were purified using reverse-phased HPLC to separate free dye and unlabeled proteins. Following purification, the lyophilized protein was buffered in 20 mM Tris, 100 mM sodium chloride, at

pH 7.0. Labeling efficiencies for the MCD-Gly-Cys variants were measured using UV-absorbance and a molar extinction coefficient of 68,000 M^{-1} cm^{-1} at 495 nm for fluorescein-5-maleimide.

2.3.7 Sedimentation Velocity Measured Using Standard Analytical Ultracentrifugation.

MCD_C1 was initially subjected to sedimentation velocity analytical ultracentrifugation assessment at the Sanford Burnham Prebys Medical Discovery Institute using a Beckman Coulter XL1 analytical ultracentrifuge equipped with a UV absorption detector. MCD_C1 was analyzed in the presence and absence of an equimolar zinc concentration to determine the effects of metal-mediated dimerization. A 35 μM MCD_C1 sample was prepared in (1) 20 mM Tris and 100 mM sodium chloride (pH 7.0) and (2) 20 mM Tris, 100 mM sodium chloride, and 35 μM zinc sulfate (pH 7.0). Samples were loaded into both sectors of a double-sector cell equipped with sapphire windows. Data were collected at 21 °C for 24 h at a rotor speed of 42000 rpm with a 50 Ti rotor. Data were analyzed in Sedfit[29] using the c(s) model. The molecular masses, partial specific volumes, solvent density, and viscosity used in the data analysis were calculated in Sednterp[30].

2.3.8 Sedimentation Velocity with Fluorescence Detection.

Additional sedimentation velocity analytical ultracentrifugation assessment was performed using a Beckman Coulter XL-I analytical ultracentrifuge equipped with an AU-FDS fluorescence detector (AVIV Biomedical, Inc., Lakewood, NJ) at the Center for Open Research Resources and Equipment, University of Connecticut. Samples were prepared in 20 mM Tris, 100 mM sodium chloride, and 1 mM zinc sulfate (pH 7.0). Due to slow equilibration, MCD_C2 samples were diluted to their final concentrations and equilibrated at 4 °C for 24 h. Samples were loaded into both sectors of double-sector cells equipped with quartz windows. Data were

collected at 20 °C for 16 h at a rotor speed of 50000 rpm with an An50 Ti rotor. Data were analyzed in Sedfit[29] using the c(s) model. The molecular masses, partial specific volumes, solvent density, and viscosity used in the data analysis were calculated in Sednterp[30].

2.4 Results and Discussion

In a previous protein design study, a geometric recognition algorithm was modified to dock the backbone coordinates of the β1 domain of streptococcal protein G (Gβ1) to itself using surface complementarity as a measure of intermolecular fitness[31,32]. This resulted in an orientation in which the helices of the two monomers are positioned in the proximity of and relatively parallel to one another. This model is termed the helix face-to-helix-face orientation (Figure 2.1).

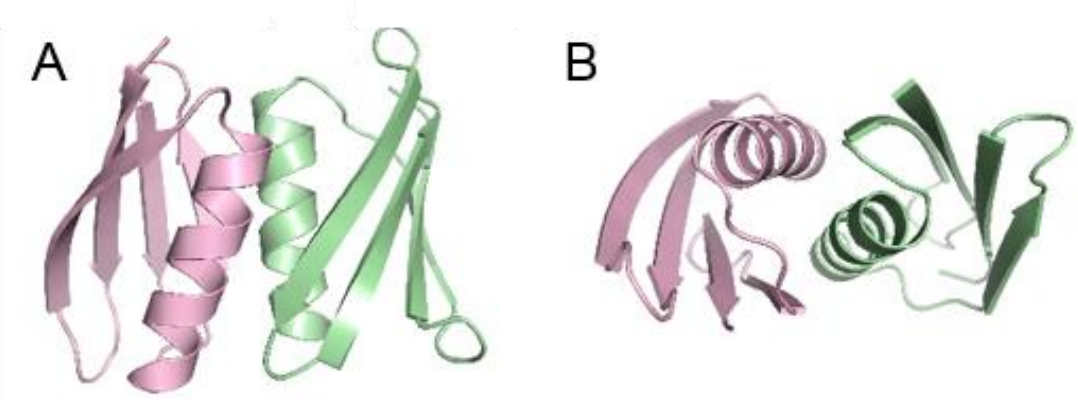

Figure 2.1: Intermolecular Orientation of the Helix-Face-to-Helix Face Computational Design Model. A) side view, B) bottom-up view

Backbone docking was followed by the use of the ORBIT suite of design algorithms[33] to determine optimal amino acid substitutions at the newly created interface. The mutations resulted in the creation of a Gβ1 heterodimer termed MonA/ MonB. Nuclear magnetic resonance (NMR) was used to verify that the two monomers did form a complex, although analytical ultracentrifugation revealed the binding constant to be quite weak (~300 μM)[31,32,34].

In an attempt to increase the binding affinity of the engineered MonA/MonB heterodimer, rational design was used to “bootstrap” stronger intermolecular interactions between the monomers. In one example, pairs of metal-coordinating histidine residues were introduced at select interfacial positions using the computationally derived helix-face-to-helix-face model. This design did not give rise to increased heterodimer affinity but did result in the formation of MonA homo-complexes. In three cases, the mutations resulted in MonA variants that are fully monomeric in the absence of metal yet form mixtures of higher-order complexes upon addition of zinc (SEC results for these variants are shown in Figure 2.7 A−C).

Structural characterization of two of these redesigns [i.e., MonA(A32H, A36H) and MonA(E21H, L25H)] provided important insights into metal-mediated self-association. Crystals were grown for these variants in the presence of 20 mM zinc sulfate, and the resulting structures revealed that the variants do not adopt the helix-face-to-helix-face model orientation but instead form complexes primarily through extension of the β-sheet (Figure 2.2). Because this structure is quite different from helix-face-to-helix-face model orientation, the designed metal-binding residues are generally not located at the interface. We believe that metal−histidine interactions drive dimer and higher-order associations as for example that depicted in Figure 2.2 D.

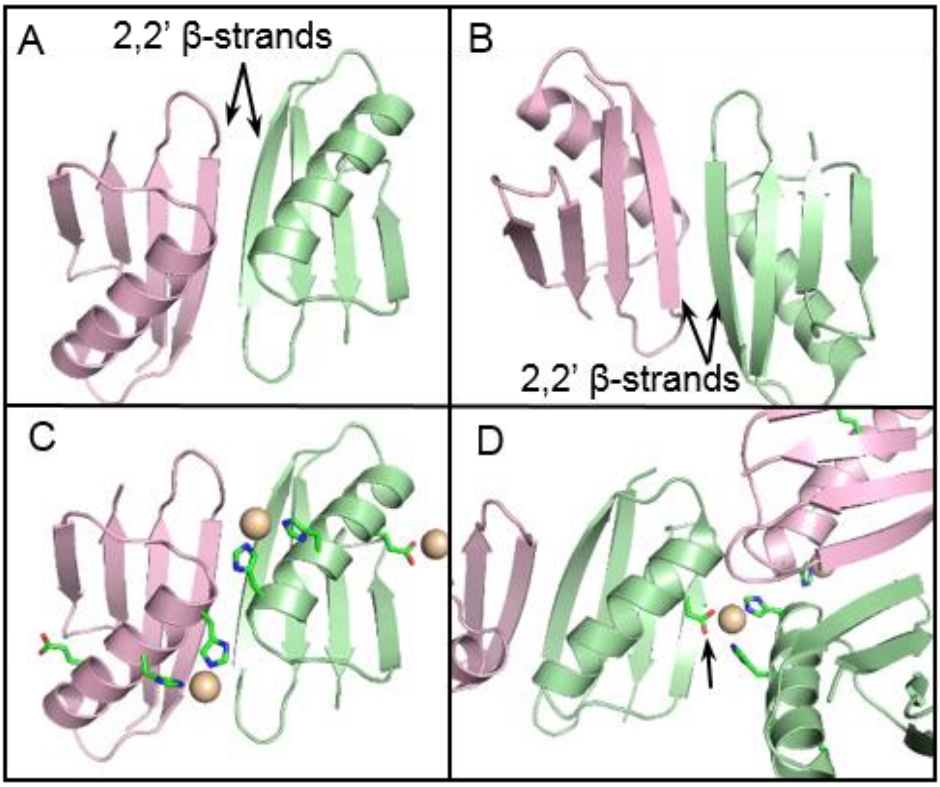

Figure 2.2: Structure of the MonA Homo-Complex A32H A36H Variant. A) and B) illustrate that complex formation occurs primarily through extension of the β-sheet. C) Designed metal-ligand histidine residues (32H, 36H) do bind zinc but not across the interface. D) Cross-dimer interaction between wild-type Glu27 (black arrow) and the introduced metal-ligand histidine residues 32H and 36H.

On the basis of these results, we pursued metal-controlled dimerization using the experimentally derived complex structures formed through β-sheet extension. The rationale for this is based on three factors: (1) favorable enthalpic interactions derived from cross-interface β-sheet hydrogen bonding, (2) the dimer orientation in the crystal structures of the MonA homo-complexes, and (3) an identical orientation of a crystal structure from an unrelated project in which a highly stable constitutive Gβ1 dimer was generated using directed evolution[18]. Wunderlich et al. used a directed evolution method termed "Proside" to select for stabilizing amino acids at four Gβ1 "boundary" positions[18]. Codon randomization of four boundary positions resulted in a library displayed on phage, incubated with immobilized antibodies, and

challenged with the protease chymotrypsin. They determined the crystal structure of a strongly stabilized variant, Gβ1-M2 (Figure 2.3A, C), which revealed that the four mutations induce dimer formation through the highly similar β-sheet extension observed for the crystal structures of the MonA metal-induced homo-complexes (Figure 2.2A, B).

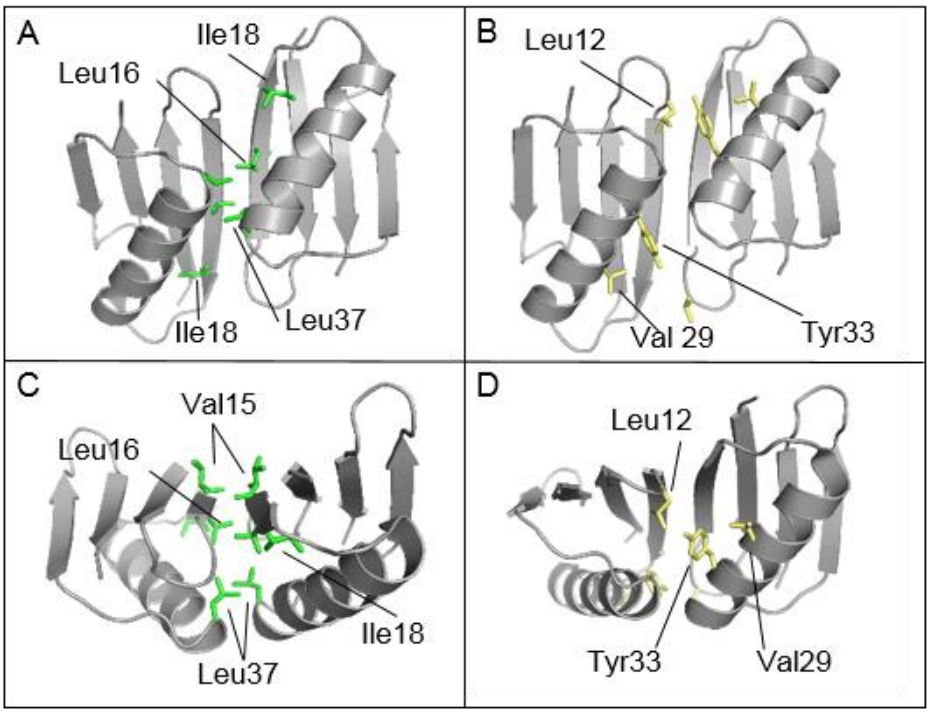

Figure 2.3: Structure of the Gβ1-M2 constitutive dimer generated with the Proside directed evolution method (PDB entry 3FIL)[18,19]. The amino acid positions of the four residues that give rise to Gβ1-M2 (E15V, T16L, T18I, N37L) are depicted in green in panels A & C. The three residues that, when mutated to histidine (described in this work), provide cross-interface metal ligands (L12H, V29H, Y33H) are depicted in yellow in panels B & D. Panel C & D are ~90° rotations of panels A & B respectively.

2.4.1 Design and SEC−MALS Characterization of Metal-Controlled Dimers.

The dimer structure derived from the Proside-generated Gβ1-M2 variant (Figure 2.3A, C) was an excellent starting model for the design of metal-controlled dimers because its X-ray crystal structure was determined to high resolution and it forms a high-affinity symmetric homodimer with a binding dissociation constant in the range of 1 μM[19]. It differs from wild-type Gβ1 at four positions: E15V, T16L, T18I, and N37L. To engineer metal-controlled dimers,

metal−ligand histidine residues were combined with variations of the four Gβ1-M2 mutations. For the work reported herein, structure-based rational design was used to identify two sets of three amino acid positions that, upon mutation to histidine, would potentially form cross-monomer interactions driven by metal binding. The two sets consist of positions 11, 29, and 33 and positions 12, 29, and 33. Positions 11 and 12 were chosen on the basis of the cross-interface proximity to positions 29 and 33. Positions 29 and 33 are located on the α-helix and are thus ideal as metal ligands as they are in the proximity of one another in standard i to i + 4 helical positions.

The first set of three mutations was introduced in the context of the four Gβ1-M2 mutations and consists of T11H, V29H, and Y33H. This variant (termed M2_11H) was subjected to SEC−MALS analysis, which revealed that it does not form a metal-controlled dimer but instead forms a constitutive dimer in the presence and absence of metal (Table 2.1).

Table 2.1: Biophysical Parameters of Metal-Controlled Dimers and Controls

variant	description	native state	state upon addition of metal[a]	T_m (°C)	T_m (°C) with metal
Gβ1	wild-type	monomer	monomer	49.0	49.3
Gβ1-M2[b]	M2 phage-derived dimer[22,23]	dimer	dimer	87.0	87.1
Gβ1_3H[c]	no M2 mutations,only 3H[c]	monomer	monomer	35.9	37.5
M2_11H	all four M2 mutations plus T11H, V29H, and Y33H	dimer	dimer	73.8	78.7
MCD_M2	all four M2 mutations plus 3H	monomer	dimer	63.6	73.9
MCD_C1	15 and 18 WT16 and 37 M2 plus 3H	monomer	dimer	52.2	64.9
MCD_C2	15 and 18 M216 and 37 Ala plus 3H	monomer	dimer	48.1	49.1

[a]Determined by SEC−MALS in 1 mM zinc sulfate. [b]M2: E15V, T16L, T18I, and N37L. [c]H: L12H, V29H, and Y33H.

The SEC−MALS-determined molecular weights of this variant are 12.2 kDa without metal and 12.3 kDa with zinc sulfate. The second variant, termed MCD_M2, consists of the four Gβ1-M2 mutations and the metal−ligand mutations L12H, V29H, and Y33H. SEC−MALS analysis revealed that this variant is monomeric in the absence of metal (6.7 kDa) yet forms a monodisperse dimer (12.2 kDa) in the presence of zinc (Figure 2.4). Initially, there was minor evidence of dimer formation in the absence of metal, but after the addition of EDTA to the MCD_M2 solution, the minor dimer peak was no longer observed (Figure 2.4 A). We believe

that residual metal carried over from bacterial expression may be the source of the observed minor dimerization. It is likely that MCD_M2 successfully functions as a metal-controlled dimer, as opposed to the constitutive dimer, M2_11H, because intermolecular hydrophobic contacts were reduced upon mutation of the wildtype leucine residue at position 12 to a metal-binding histidine residue (i.e., the L12H mutation in MCD_M2).

For the study reported herein, we were interested in designing metal-controlled dimers with variable affinities. To achieve this, we relied on the biophysical analysis originally performed on Gβ1-M2[19]. Thoms et al. used both X-ray crystallography and NMR spectroscopy to fully characterize the high-resolution structure of the Gβ1-M2 variant (PDB entry 3FIL). They also performed an extensive series of biophysical experiments on permutations of the four Gβ1-M2 mutations[19]. In short, they found that mutations to larger hydrophobic residues at positions 16 and 37 (T16L and N37L) provide the greatest contribution to thermal stabilization of the symmetric Gβ1-M2 dimer. To generate MCD variants with altered affinities, we designed two additional variants using different combinations of the four Gβ1- M2 mutations. These modifications essentially altered the degree of hydrophobic contact at the MCD interface. This approach, albeit at a more complex level, was used to convert the H-maxiferritin protein from one that spontaneously self-assembles into one that assembles only in the presence of divalent copper[17]. This was achieved by reducing intermolecular contacts at one of the complex interfaces combined with the introduction of metal-binding residues that drive complex formation. For one of the variants we examined, hydrophobic contacts were reduced by reversion back to wild-type Gβ1 residues at positions 15 and 18 (V15E and I18T, respectively). This variant, MCD_C1, is also a monodisperse monomer in the absence of metal (6.4 kDa) and forms a clean dimer (12.5 kDa) in the presence of zinc (Figure 2.4 B). To further study the effects of

interfacial mutations on binding affinity, a third MCD variant was designed in which positions 16 and 37 were mutated to alanine, whereas positions 15 and 18 consisted of Gβ1-M2 mutations. This construct, termed MCD_C2, contains the three metal−ligand mutations and E15V, T16A, T18I, and N37A at the dimer interface. SEC−MALS analysis in which the protein concentration was 1 mM (Figure 2.4 C) revealed that MCD_C2 also functions as a metal-controlled dimer. Mutation to residues with a smaller average surface area functioned to reduce thermal stability and binding affinity relative to those of MCD_M2 and MCD_C1.

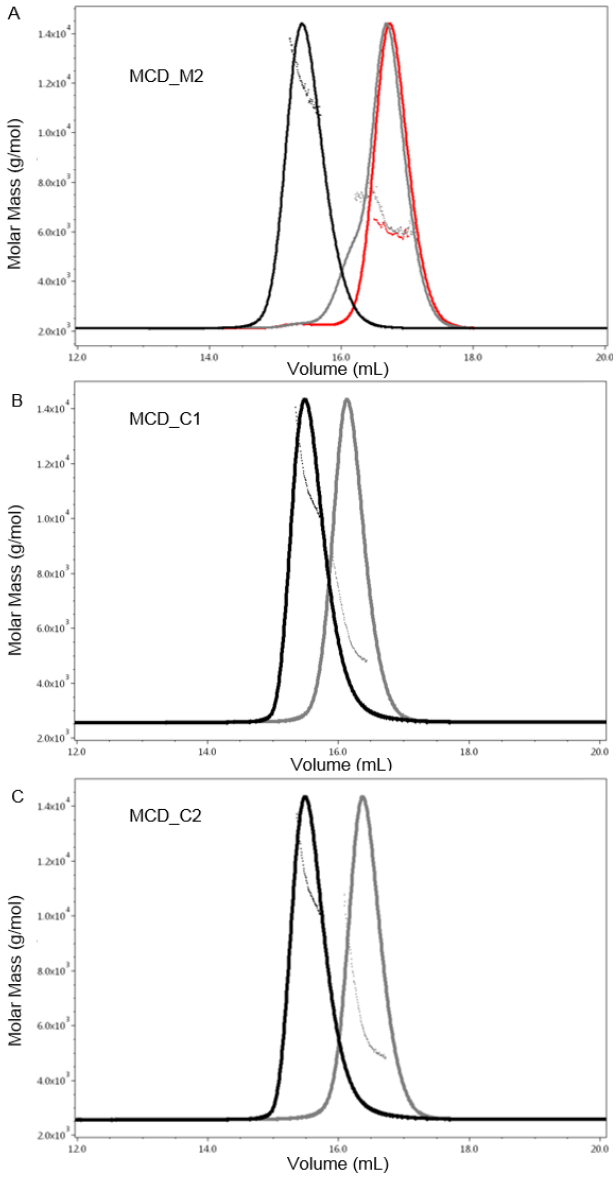

Figure 2.4: SEC-MALS Chromatographs. Overlaid SEC-MALS traces for A) MCD_M2, B) MCD_C1, and MCD_C2. Each sample was run in the presence of zinc (darker trace on left) and the absence of metal (lighter trace on right). In A) the broader (gray) trace on the right corresponds to the monomer with no added EDTA while the more symmetric trace (red) corresponds to the same variant after EDTA was added to the solution.

2.4.2 Thermal Characterization of the MCDs.

Circular dichroism was used to measure the thermal stabilities of the three designed MCD variants, in addition to associated controls (Table 2.1). The parent Gβ1-M2 variant is a hyperthermophile and thus does not denature below 100 °C[19]. Therefore, it was necessary to

record CD spectra in the same manner in which Gβ1-M2 was initially studied, i.e., in 1.5 M guanidinium hydrochloride[19]. Overall increases in the measured melting temperatures (Tm) reflect both stabilization via dimer formation and the degree of hydrophobic contact at the interfaces. In general, incorporating hydrophobic residues with smaller accessible surface areas at interfacial positions resulted in variants with lower melting temperatures.

The parent variant, Gβ1-M2, has the highest recorded Tm of 87 °C, which is higher than that of wild-type Gβ1 by ∼38 °C. The Tm for MCD_M2 (in the absence of metal) is 63.6 °C, which is reduced relative to that of Gβ1-M2 by 23.4 °C (Table 2.1). This is likely due to the loss of hydrophobic packing contributed by valine 29 and tyrosine 33, but especially leucine 12 as it is relatively close to the cross-monomer positions of isoleucine 18 and leucine 37 (Figure 2.3). The Tm for the constitutive dimer M2_11H is 73.8 °C, which is 10.2 °C higher than that of the metal-controlled dimer, MCD_M2. These variants differ only by substitution of a histidine for leucine at position 12. M2_11H does exhibit an increase in its Tm of ∼5 °C in the presence of metal, which indicates that stabilization by metal−ligand bridging likely occurs to some degree at position 11 even for this constitutive dimer.

To test whether the metal−ligand mutations alone could drive dimer formation, a Gβ1 control variant was generated that contained only the three metal−ligand mutations (i.e., L12H, V29H, and Y33H). This variant, termed Gβ1_3H, was subjected to CD-based thermal analysis in 1.5 M guanidine hydrochloride, which indicated that its Tm (35.9 °C) was reduced relative to that of wild-type Gβ1 by ∼13 °C. This demonstrates that wild-type residues L12, V29, and Y33 contribute to some extent to the thermal stability of Gβ1. SEC−MALS analysis performed on the Gβ1_3H control indicates that it does not form a dimer (or higher-order complexes) in the presence or absence of metal.

Of the three successful MCD variants, MCD_M2 exhibits the highest Tm of 63.6 °C, which increases by 10.3 °C upon addition of metal. The Tm for the MCD_C1 variant, which contains only the two dominant Gβ1-M2 mutations (T16L and N37L),[19] is reduced to 52.2 °C in the absence of metal yet increases 12.7 °C to a value of 64.9 °C in its presence. MCD_C2, which has alanine substitutions at the two critical positions 16 and 37, does form a metal-controlled dimer at 1 mM even though its Tm is reduced to 48.1 °C with no appreciable increase upon addition of metal (Table 2.1). The increases in Tm for MCD_M2 and MCD_C1, in the presence of metal, further demonstrate that metal-driven dimer formation is robust as these complexes form even in the presence of 1.5 M guanidinium hydrochloride.

2.4.3 Sedimentation Velocity Analytical Ultracentrifugation.

The successful MCD variants were subjected to analytical ultracentrifugation (AUC) analysis in which differences in sedimentation velocity were used in an attempt to measure the metal-controlled binding affinities. Initially, MCD_C1 was analyzed using sedimentation velocity with standard A280 absorbance detection. The protein concentration for this analysis was ~35 μM, and the results indicated that the protein is monodisperse in the presence and absence of zinc. In the absence of zinc, only a single monomer peak is observed (1.041 S, apparent MW of 6.7 kDa), and in the presence of an equimolar concentration of zinc (35 μM), only a single dimer peak is observed (1.514 S, apparent MW of 11.8 kDa). Because there is no evidence of dimer dissociation in the presence of zinc, we initially estimated the Kd to be < 200 nM. SEC−MALS analysis run at 1 mM for this sample indicated that, even at this relatively high concentration, there is no evidence of dimer formation or any higher-order aggregation in the absence of metal.

For enhanced detection sensitivity, all three MCDs were fluorescently labeled with fluorescein and analyzed using sedimentation velocity with fluorescence detection. For MCD_C1, the c(s) continuous sedimentation coefficient distributions show an ~1.8 S dimer with no evidence of dissociation down to the lowest concentration of 15 nM, indicating that the Kd < 75 pM. For MCD_C2 (the variant designed with the lowest degree of interfacial hydrophobicity), the distributions reveal a monomer peak at ~1.2 S whose relative amplitude increases as the concentration decreases. This behavior is characteristic of a reversible monomer−dimer equilibrium that is slow on the time scale of sedimentation. In an attempt to determine the dissociation constant, the signal average sedimentation coefficients were calculated and plotted as a function of the total protein concentration to produce a binding isotherm. The data do not fit well to a monomer−dimer equilibrium model, suggesting that equilibrium has not been achieved, and it is not possible to obtain a precise dissociation constant; thus, we estimate the Kd to be in the range of ~400 nM. Finally, MCD_M2 (the variant with all four Gβ1-M2 mutations) exhibits evidence of dissociation at concentrations of < 50 nM, and therefore the Kd appears to lie in the low nanomolar range and thus exhibits a binding affinity that is intermediate between those of MCD_C1 and MCD_C2. This is not unexpected as the hydrogen bonding register for MCD_M2 (described below) is different compared to those of the other variants and may reflect non-optimal packing of residues at the dimer interface.

2.4.4 X-ray Crystal Structures of Designed Dimers.

Crystals were grown for the two higher-affinity metal-controlled dimers, MCD_M2 and MCD_C1, in both the presence and absence of zinc (crystallographic data collection and refinement statistics are listed in Tables 2.8 and 2.9). The resulting structures very closely match the target model structure, Gβ1-M2, as well as the MonA homo-complex structures (Figure 2.5).

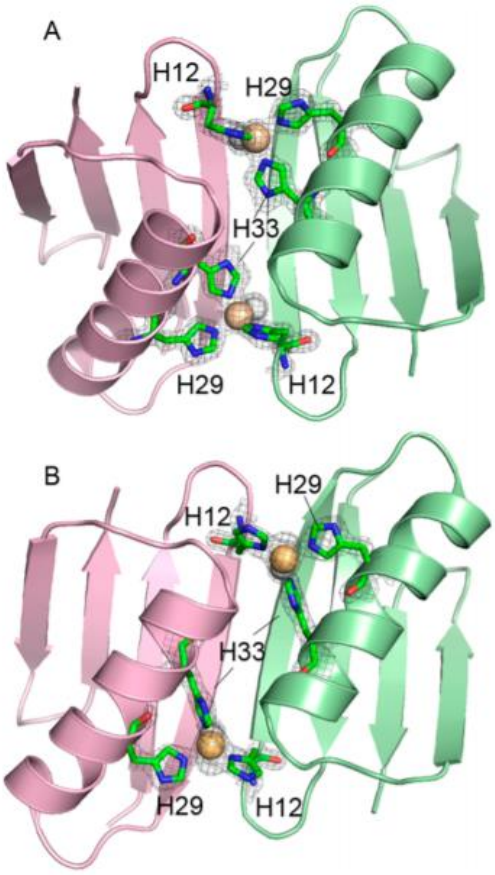

Figure 2.5: Crystal Structures of MCD_M2 (A) and MCD_C1 (B). The designed metal-ligand histidine residues are illustrated as stick-bonds. Zinc atoms are illustrated as amber spheres.

Root-mean-square deviation (RMSD) values for comparisons of all structures are relatively low (Table 2.4 and Figure 2.9) with an average value of 0.50 Å for Cα atom alignments and 0.56 Å for all atom alignments. As with the structure of the parent variant, Gβ1-M2, all structures adopt a head-to-head intermolecular orientation in which a dyad-symmetric dimer is formed through extension of the β-sheets. The structures revealed that the β-sheets form intermolecular hydrogen bonds with the symmetrically related counterparts resulting in a continuous extended β-sheet. The structures and cross-interface hydrogen bond register for Gβ1-M2, the two MonA homo-complexes, and MCD_C1 are highly similar, with the one exception being MCD_M2. For the MonA complexes, MCD_C1, and Gβ1-M2, there are six interfacial hydrogen bonds: two between positions 17 and 13′, two between positions 15 and 15′, and two

between positions 13 and 17′ (the numbering for cross-interface positions is denoted with a prime). For MCD_M2, there are only four interfacial hydrogen bonds: two between positions 17 and 15′ and two between the symmetrically related positions 15 and 17′. The alternative intermolecular hydrogen bond register for MCD_M2 implies that the interface is malleable, which is likely due to the predominantly hydrophobic nature of interfacial residues that, in combination with the metal−ligand histidine residues, drive dimer formation.

The designed zinc coordination sites are formed by two histidine residues located in i to i + 4 positions (29 and 33) on the Gβ1 α-helix and an additional cross-interface interaction involving a histidine at position 12, which is located on the turn between the first and second β-strands. Tetrahedral zinc ion coordination involves three histidine ligands and a chloride ion. Analysis of the crystal structure of the two MCD complexes, in the presence of zinc sulfate, demonstrates that the distances and bond angles between the zinc atom and the three histidine ligands are within range of standard tetrahedral values (Tables 2.6 and 2.7). In the crystal structure of MCD_C1, there is an additional ordered chloride ion in the proximity of the tetrahedral coordination site. It appears to be aligned with the histidine-coordinated zinc atom and the tetrahedral chloride ion and likely functions as a counterion.

2.5 Conclusion

The challenges associated with novel protein interface design are on par with standard protein design but with the additional criterion of incorporating protein−protein bridging interactions that do not cause nonspecific aggregation. Metal−ligand interactions, designed to bridge interfaces, can provide significant intermolecular binding energy and reduce the need for the precisely balanced size of hydrophobic patches on the surface of designed binding partners. In addition, various approaches have utilized structural motifs (e.g., β-strands) for anchoring and

enhancing designed protein−protein interactions[15,35,3]. Results from our previous de novo interface design, as well as the constitutive Gβ1-M2 dimer (generated in a different laboratory), indicate that the preferred intermolecular interaction for Gβ1 is also through β-strand pairing. By combining cross-monomer metal−ligand interactions with β-strand pairing, we aimed to generate high binding affinity and also provide insight into the energetics of molecular self-assembly. The metal-controlled variants we engineered are completely monomeric in the absence of metal yet form relatively high-affinity exceptionally clean dimers in the presence of zinc. The designed proteins are monodisperse in the presence and absence of metal and exhibit no evidence of multistate complexes or aggregation in either state. We plan to use tandemly expressed variations of the designed MCDs as building blocks of biomaterial that can be customized to specifically assemble and disassemble upon addition of metal and chelating agents, respectively.

2.6 Funding and Acknowledgments

Funding This work was supported by the California Metabolic Research Foundation and in part by the National Science Foundation under Career Grant 0448670. In addition, this material is based upon work supported by the U.S. Army Research Laboratory and the U.S. Army Research Office under Grant W911NF-13-1- 0155. B.M. was a recipient of an Arne N. Wick Predoctoral Research Fellowship from the California Metabolic Research Foundation.

Acknowledgements The authors thank J. Caldwell, A. Farah, and E. Farokhi for support with protein crystallization. The authors thank A. Bobkov for support with analytical ultracentrifugation. This research used resources of the Advanced Light Source, which is a U.S. Department of Energy (DOE) Office of Science User Facility operated by Lawrence Berkeley National Laboratory under Contract DE-AC02-05CH11231, and the Advanced Photon Source, a

DOE Office of Science User Facility operated by Argonne National Laboratory under Contract DE-AC02- 06CH11357.

2.7 Supporting Information

2.7.1 Complete Amino Acid Sequences for all Variants

Table 2.2: Complete Amino Acid Sequences for MCD Variants and Controls

Variant Name	Mutations	Sequence
Gβ1	Wild-type	MTYKLILNGKTLKGETTTEAVDAATAEKVFKQY ANDNGVDGEWTYDDATKTFTVTE
Gβ1-M2	E15V, T16L, T18I, N37L	MTYKLILNGKTLKG**VL**T**I**EAVDAATAEKVFKQY AND**L**GVDGEWTYDDATKTFTVTE
M2_11H	T11H, E15V, T16L, T18I, V29H, Y33H, N37L	MTYKLILNGK**H**LKGVLTIEAVDAATAEK**H**FKQ**H** ANDLGVDGEWTYDDATKTFTVTE
MCD_M2	L12H, E15V, T16L, T18I, V29H, Y33H, N37L	MTYKLILNGKT**H**KG**VL**T**I**EAVDAATAEK**H**FKQ**H** AND**L**GVDGEWTYDDATKTFTVTE
MCD_C1	L12H, T16L, V29H, Y33H, N37L	MTYKLILNGKT**H**KGE**L**TTEAVDAATAEK**H**FKQ**H** AND**L**GVDGEWTYDDATKTFTVTE
Gβ1_3H	L12H, V29H, Y33H	MTYKLILNGKT**H**KGETTTEAVDAATAEK**H**FKQ**H** ANDNGVDGEWTYDDATKTFTVTE
MCD_C2	L12H, E15V, T16A, T18I, V29H, Y33H, N37A	MTYKLILNGKT**H**KG**VA**T**I**EAVDAATAEK**H**FKQ**H** AND**A**GVDGEWTYDDATKTFTVTE

Table 2: All interfacial mutations from wild-type Gβ1 are bolded.

Table 2.3: Complete Amino Acid Sequences for MonA, MonB, MonA Homo-Complexes Variants and the Fusion Dimer Control

Variant Name	Protein Sequence
MonA	MTYKLILNGKTLKGE**F**T**A**EA**E**DAA**L**AE**YI**F**RAL**A**KAQ**GVDGEWTYDD ATKTFTVTE
MonB	MTYKLILNGKTLKGETTTEAVD**I**ATA**AD**VF**A**QYA**A**DNGV**K**GEWT**A**D**E** ATKTFTVTE
AB-fusion (dimer control)	MTYKLILNGKTLKGE**F**T**A**EA**E**DAA**L**AE**YI**F**RAL**A**KAQ**GVDGEWTYDD ATKTFTVTE**GGGGSGGGGSH-** MTYKLILNGKTLKGETTTEAVD**I**ATA**AD**VF**A**QYA**A**DNGV**K**GEWT**A**D**E** ATKTFTVTE
MonA (A32H, A36H)	MTYKLILNGKTLKGE**F**T**A**EA**E**DAA**L**AE**YI**F**RH**L**A**KHQ**GVDGEWTYDD ATKTFTVTE
MonA (F16H, A18H)	MTYKLILNGKTLKGE**H**T**H**EA**E**DAA**L**AE**YI**F**RAL**A**KAQ**GVDGEWTYDD ATKTFTVTE
MonA (E21H, L25H)	MTYKLILNGKTLKGE**F**T**A**EA**H**DAA**H**AE**YI**F**RAL**A**KAQ**GVDGEWTYDD ATKTFTVTE

Table 3: All interfacial mutations from wild-type Gβ1 are bolded

2.7.2 X-Ray Crystal Structures and Associated Omit Maps

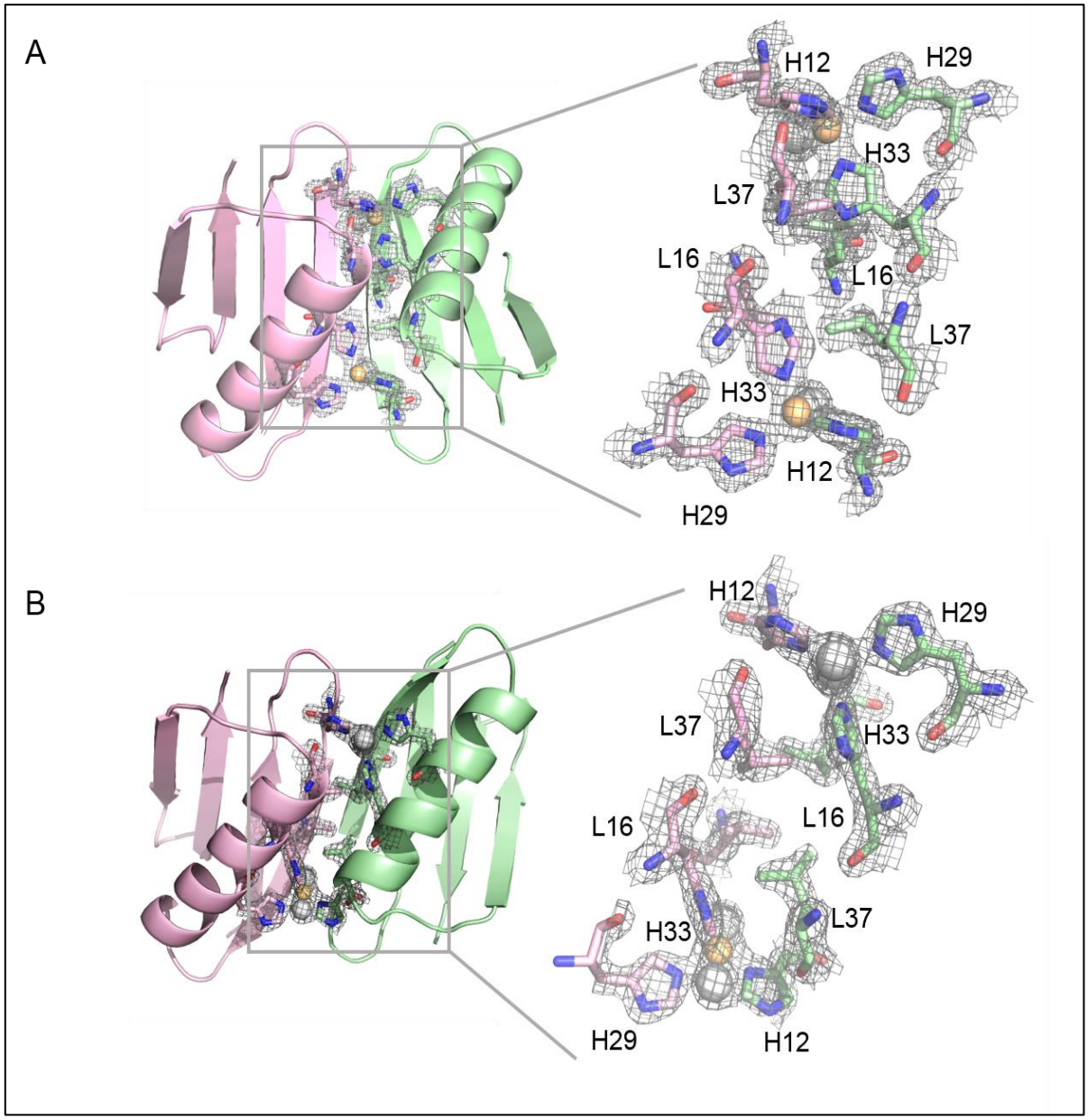

Figure 2.6: X-Ray Crystal Structures and Omit Maps. High-resolution crystal structures (left) for MCD_M2 (A) and MCD_C1 (B). The $2F_o$-F_c electron density map omit maps (contouring level=1.5σ) are shown on the right for the metal-ligand histidine residues (His 12, 29, 33) and also the hydrophobic residues, Leu 16 and Leu 37, that contribute to complex formation. Side-chains are illustrated as stick-bonds, zinc atoms as amber spheres, and chlorine atoms as light gray spheres.

2.7.3 Size Exclusion Chromatography Performed on MonA Homo-Complexes

In a project that followed the initial computational de novo protein-protein design project[32], the helix-face-to-helix-face model was used, in combination with rational design, to re-engineer the MonA/MonB complex with the goal of increasing the binding affinity. This did not result in an increase to the binding affinity for the MonA/MonB complex but did result in three MonA mutants that formed homo-complexes in the presence of the metal yet remained completely monomeric when no metal was present. The mutations that gave rise to these variants are the following: MonA(A32H, A36H), MonA(F16H, A18H), and MonA(E21H, L25H). The degree of complex formation, for these three MonA homo-complexes was assessed using size-exclusion chromatography (Figure 2.7, next page). In an attempt to ascertain the molecular weights of the MonA homo-complexes, and thus the degree of multimerization, the different MonA samples were subjected to SEC combined with multiangle-light scattering (SEC-MALS) analyses. The results of this analysis revealed that it is not be possible to determine the molecular weights of the MonA homo-complexes as the SEC-MALS measured values were extremely variable and did not correlate with SEC retention times. We believe this is likely due to that fact that accurate MW values can only be measured by MALS on samples that are exceptionally mono-dispersive. If there is even exceedingly minor bleed over between SEC peaks the poly-dispersive nature of the resulting solution precludes the ability to accurately determine MW values. This was the case for the MonA homo-complexes but definitely not for the Gβ1-M2 (the constitutive homo-dimer) as well as the three successful MCD designs as these are all highly mono-dispersive and generate very consistent SEC-MALS derived MW values. In Figure 2.7 the AB-fusion (dimer control) corresponds to the expression of MonA fused to MonB at the level of the gene. The two genes are expressed in tandem with a $(Gly_4Ser)_2$ intervening linker.

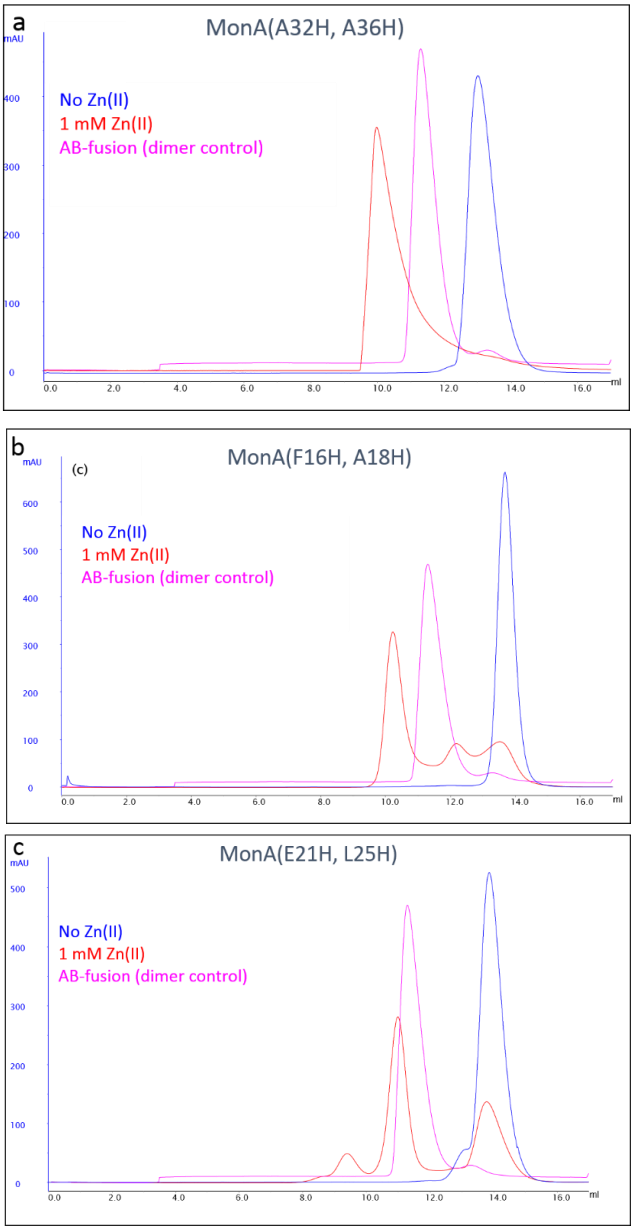

Figure 2.7: Size-Exclusion Chromatography of MonA Homo-Complexes. A) MonA(A32H, A36H, B) MonA(F16H, A18H), C) MonA(E21H, L25H). Blue trace - sample with no added zinc sulfate, pink trace - MonA/MonB fusion control, red trace - sample run with 1 mM zinc sulfate. The y-axis corresponds to absorbance at 280 nm and the x-axis corresponds to retention volume in mL

2.7.4 X-ray Crystal Structures for the MonA Homo-Complexes

The X-ray crystal structure of the MonA homo-complex variant A32H, A36H (in the presence of zinc) was solved by single wavelength anomalous dispersion (SAD) phasing using the absorption edge of zinc and refined to 1.49 Å resolution. This allowed for the direct calculation of zinc positions within the protein unit cell and calculation of experimental electron density maps from these initial zinc positions. The main structural feature of this MonA homo-complex is that dimer and complex formation occurs via extension of the four-stranded β-sheet (Figure 2.2 D in manuscript). This mode of dimerization is analogous to the constitutive homo-dimer Gβ1-M2 and the successful MCD complexes. The crystal structures of the MonA homo-complexes also reveal that zinc ions are, for the most part, bound by the engineered histidine pairs, but not at the experimentally determined interface. Both MonA variants crystallize much more readily in the presence of zinc although the structures revealed that metal-ligand histidine residues do not participate in cross-dimer interactions. Therefore we believe that metal binding to histidine residues (as well as to wild-type residues, e.g., Glu27, main manuscript, Figure 2.2 D) functions to induce intermolecular interactions which significantly promote crystallization.

The crystal structure of wild-type Gβ1 (PDB: 1PGA)[36] was used with molecular replacement to solve the structure of MonA homo-complex E21H L25H (Figure 2.8). The structure was refined to 1.48 Å resolution. The complex forms a tetramer within the asymmetric unit and it appears that the zinc atoms function as catalysts for crystallization in addition to driving self-assembly. For this structure, in addition to extension of the β-sheet between the 2nd β-strands from each monomer, as observed for all structures, there is also additional intermolecular interactions between 3rd β-strands in the structure MonA E21H L25H variant (Figure 2.8).

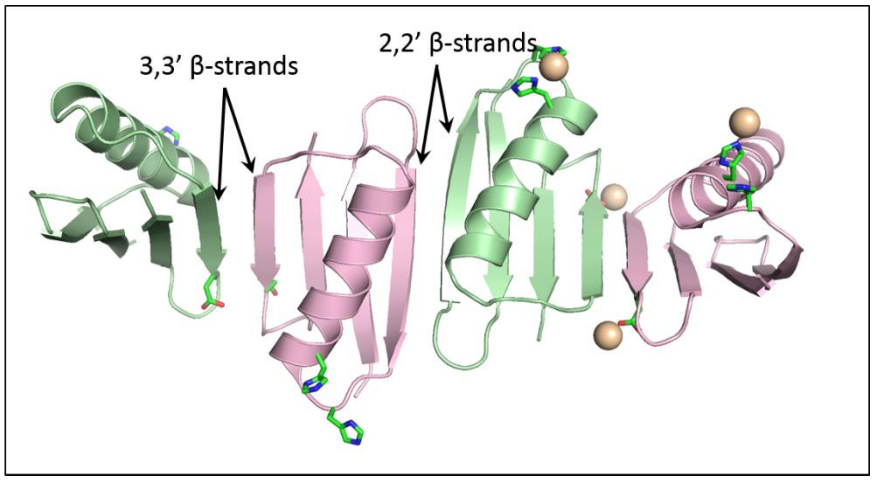

Figure 2.8: Structure of the MonA homo-complex E21H L25H variant. Metal-ligand histidine residues are shown as rods and zinc atoms as amber sphere. Black arrows illustrate that complex formation occurs through β-strand pairings.

2.7.5 RMSD Values for Structural Alignment

Table 2.4: RMSD Values for Structural Alignment (Pymol) in Angstroms

	MonA A32H, A36H	MonA E21H, L25H	Gβ1-M2	MCD_C1	MCD_M2
MonA A32H, A36H		0.72 (0.74)	0.45 (0.50)	0.52 (0.56)	0.40 (0.61)
MonA E21H, L25H	0.72 (0.74)		0.67 (0.60)	0.54 (0.53)	0.36 (0.38)
Gβ1-M2	0.45 (0.50)	0.67 (0.60)		0.60 (0.63)	0.39 (0.54)
MCD_C1	0.52 (0.56)	0.54 (0.53)	0.60 (0.63)		0.38 (0.46)
MCD_M2	0.40 (0.61)	0.36 (0.38)	0.39 (0.54)	0.38 (0.46)	

The first value in each box is the RMSD values for the alignment based on the Cα carbon atoms while the value in the parenthesis is the RMSD value for the alignment of all atoms.

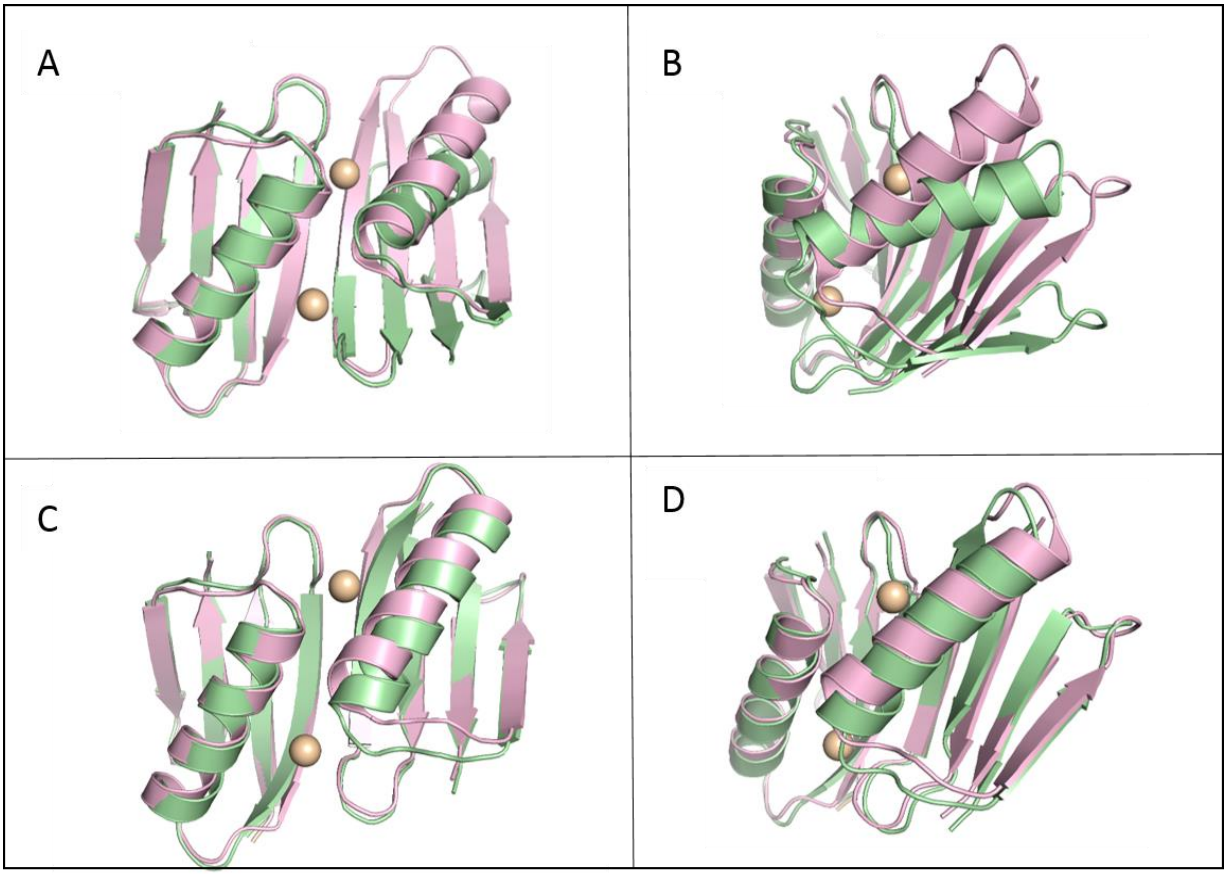

Figure 2.9: Structure Overlays of the MCDs relative to Gβ1-M2. In all four panels the structure of Gβ1-M2 is colored pink and the MCD variant is colored green. (A & B) MCD_M2, (C & D) MCD_C1. Panels C & D are approximately 90° rotations (along the Y axis) of the structures in panels A & B. Zinc atoms are represented as amber spheres.

2.7.6 SEC-MALS Analysis of the Successful MCD Variants

The three successful MCD variants, the constitutive homo-dimer Gβ1-M2, plus various Gβ1 mutant controls, were subjected to SEC-MALS analysis. It is believed that accurate SEC-MALS determined molecular weights were measured for these variants because they are all highly mono-dispersive in both the presence and absence of added metal (zinc). This fact is in agreement with the exceptionally clean data generated by both standard AUC as well as fluorescence-based AUC. In both cases the MCD variants ran as a single monomer peak in the

absence of metal and a single dimer peak when metal is present. Below are the SEC-MALS determined molecular weights for all variants.

Table 2.5: SEC-MALS Measured Molecular Weights (kDa)

Variant Name	EDTA	No EDTA Additive	500 µM Zn	1 mM Zn	2 mM Zn
Gβ1	7.3	6.8	6.8	6.9	6.6
Gβ1-M2	12.0	12.0	12.2	12.2	12.2
M2_11H	12.2	12.3	12.3	12.2	12.4
MCD_M2	6.2	6.7	12.1	12.2	12.3
Gβ1_3H	6.0	6.1	7.7	7.6	7.3
MCD_C1	6.4	6.5	12.5	12.5	12.7
MCD_C2	6.4	6.3	11.6	11.7	11.9

2.7.7 Physical Characteristics of the Zinc Tetrahedral Coordination Sites

Zinc binding sites found in protein structures show that standard zinc nitrogen (histidine) distances in zinc metalloproteins is 2.09 ± 0.14 Å[37]. The crystal structures for MCD_M2 and MCD_C1, in the presence of zinc, each contain one molecule in the asymmetric unit. The biological assembly of both the MCD_M2 and MCD_C1 is a dimer in the presence of zinc ions. Molecular symmetry, based on space groups and a single molecule in the asymmetric unit, was used to generate the structure of each metal-controlled dimer. Non-ideal differences in bond angles may arise from minor errors in unit cells, which may give rise to differences from ideal zinc coordination for the bond distances and angles.

Table 2.6: Distances between each Zinc Atom and the Side-Chain Nitrogen Atom

Residue	MCD_M2 (Å)	MCD_C1 (Å)
H12	2.05	1.72
H29	2.04	2.33
H33	1.78	2.17

Table 2.7: Angles at Zinc Atoms between pairs of Metal-Ligand Histidine Residues

Residue	MCD_M2 (°)	MCD_C1 (°)
H12 and H29	107.13	111.07
H12 and H33	99.60	109.21
H29 and H33	100.67	82.49

2.7.8 Thermal Denaturation of all Variants Monitored with Circular Dichroism

The thermal denaturation melting temperatures for wild-type Gβ1, the constitutive homo-dimer Gβ1-M2, various Gβ1 mutant controls, and the successful MCD variants were measured using circular dichroism. The increase in the CD signal at 222 nm was monitored as a function of increasing temperature under different experimental conditions. Normalized curves with both the folded and unfolded states for the Gβ1 and MCD variants are shown below.

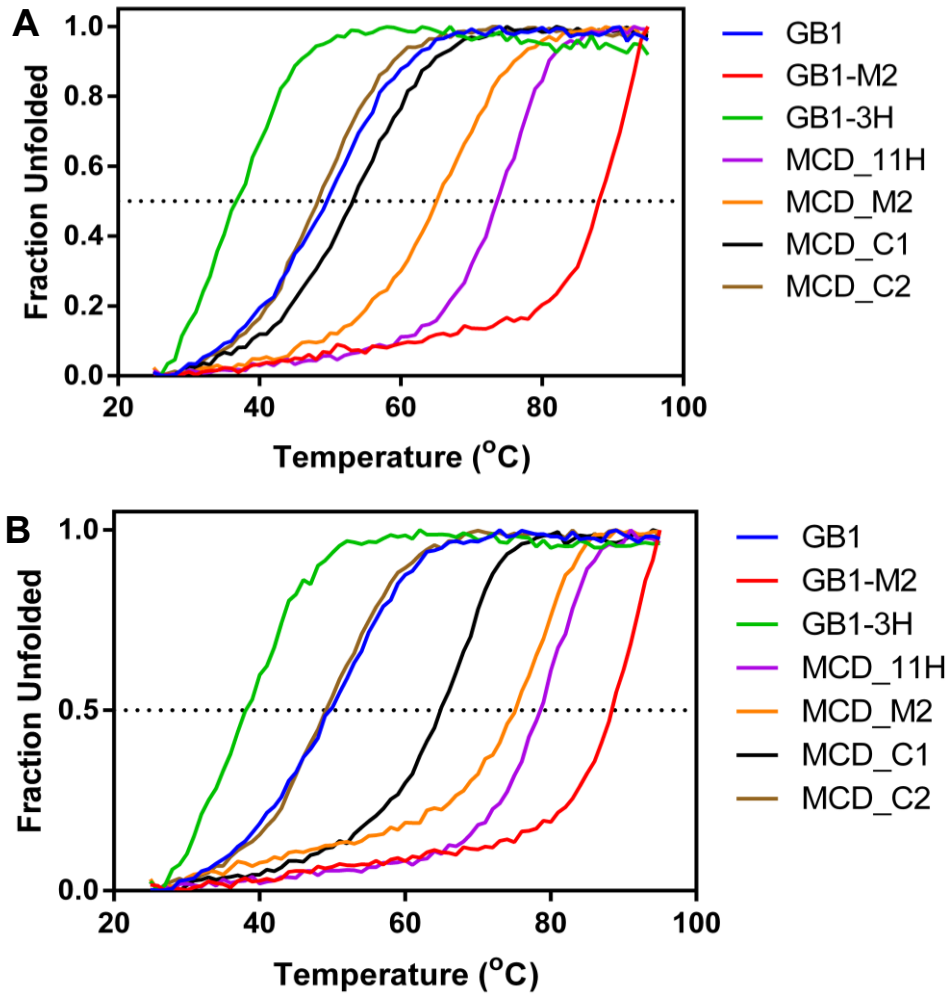

Figure 2.10: CD Thermal Melts for MCD Variants plus Controls. A) Solution containing no added metal, and B) CD collected in the presence of metal. Normalized curves are illustrated for the various proteins listed to the right. The buffer consisted of 20 mM MOPS, 10 mM sodium chloride, 1.5 M guanidinium hydrochloride, and pH 7.0. For B) the solutions with metal consisted of 25 µM zinc sulfate.

2.7.9 Summary of Crystal Formation

Table 2.8: Crystallization buffer, space group, cell parameters, and resolution

Protein	Crystallization buffer	Space group	Molecules in asymmetric unit	Resolution (Å)
MonA A32H A36H	70 mM acetic Acid pH 3.6, 30 mM acetic acid pH 5.8, 30% 2,4-methylpentanediol, 100 mM NaCl, and 20 mM zinc sulfate	$P3_2$	4	1.4
MonA E21H L25H	80 mM acetic acid pH 3.6, 20 mM acetic acid pH 5.8, 30% 2,4-methylpentanediol, 200 mM NaCl and 20 mM zinc sulfate	$P2_1$	4	1.4
MCD_M2, apo	31% PEG 4,000 0.1M Tris pH 8.5 200mM $MgCl_2$	$P3_1$	4	2.3
MCD_M2, zinc	4M NaCl 0.1M HEPES pH 7.5 50mM $MgCl_2$, 5mM zinc sulfate	$I4_1$	1	1.3
MCD_C1, apo	30% PEG 4000 0.1M HEPES pH 7.5 200 mM $MgCl_2$	$P2_1$	4	1.7
MCD_C1, zinc	18.5% PEG 400 0.1M HEPES pH 7.5 50 mM $MgCl_2$, 5mM zinc sulfate	$I4_1$	1	1.5

2.7.10 SDS-PAGE Run for all Reported Variants

Figure 2.11: SDS PAGE Run for All Reported Variants. All purified samples were subjected to SDS PAGE analysis in an 18% acrylamide gel. Lanes 2 through 6 correspond to various MonA variants. Lane 7 corresponds to the forced dimer in which the gene for MonB was cloned behind the gene for MonA (plus a 15 amino acid linker).

2.7.11 X-Ray data collection parameters and structure refinement statistics

Table 2.9: X-ray data collection parameters and structure refinement statistics.

Data collection	**MonA (A32H, A36H) twinned: 0.51-0.49 PDB: 6NL6**	**MonA (E21H, L25H) PDB: 6NL7**	**MCD_C1, zn PDB: 6NL8**	**MCD_C1, apo PDB: 6NL9**	**MCD-M2, zn PDB: 6NLA**	**MCD-M2, apo twinned: 0.52-0.48 PDB: 6NLB**
Space group	$P3_2$	$P2_1$	$I4_1$	$P2_1$	$I4_1$	$P3_1$
Unit cell (Å)	48.29, 48.29, 83.34	40.93, 63.06, 53.23	69.83, 69.83, 24.16	45.81, 52.03, 50.15	63.16, 63.16, 39.61	42.89, 42.89, 102.17
Wavelength (Å)	1.00002	1.098	1.0083	1.0083	1.0083	1.54
Resolution (Å)	1.4	1.4	1.5	1.7	1.3	2.3
Unique reflections	38,948	44,547	8,881	21,921	16,453	7,581
Completeness (%)	95.7 (97.4)	98.73 (80.1)	97.38 (81.27)	96.70 (76.28)	97.95 (83.64))	85.96 (33.65)
Redundancy (%)	9.5 (6.5)	9.2 (7.3)	6.9 (5.0)	3.5 (3.2)	7.0 (5.8)	9.9 (6.5)
Rmerge (%)[2]	0.115 (0.72)	0.13 (0.53)	0.054 (0.42)	0.098 (0.49)	0.057 (0.41)	0.10 (0.29)
$\langle I/\sigma(I)\rangle$	19.5 (2.0)	14.2 (3.1)	26.8 (3.0)	22.4 (2.6)	24.6 (3.1)	7.8 (2.1)
Refinement						
Resolution range (Å)	37.38 – 1.40	40.06 – 1.40	49.38 – 1.50	45.64 – 1.70	31.58 – 1.34	37.15 – 2.30
No of subunits No of atoms	4	4	1	4	1	1
total	2092	932	523	1277	1038	126
water/metal ion	264 / 6	309 / 4	57 / 2	145 / 1	107 / 2	88 / 1
R-cryst[3] / R-free[4]	0.149 / 0.214	0.113 / 0.158	0.165 / 0.235	0.185 / 0.223	0.110 / 0.146	0.192 / 0.258
Rmsd bond lengths (Å)	0.037	0.016	0.031	0.009	0.033	0.020
Rmsd bond angles (Å)	3.18	1.85	2.48	1.12	2.89	2.25
Average B-factor						
Protein	16.5	19.2	32.2	28.9	23.7	42.6
Metals	19.56	14.3	29.2	18.1	17.9	40.9
Ramachandran Plot (%)						
favored	98.1	98.1	98.1	97.6	97.9	98.1
allowed	1.9	1.9	1.9	2.4	2.1	1.9

[1] Highest-resolution shell (2.502 - 2.567 Å) information in parentheses.

[2] $R_{merge} = (\sum_h \sum_i | \langle I(h)\rangle - I(h)_i |) / \sum_h \sum_i I(h)_i$, where $I(h)_i$ is the i^{th} observation of reflection *h* and $\langle I(h)\rangle$ is the mean intensity of all observations of reflection *h*.

[3] Crystallographic R-factor = $(\sum_h || F_{obs}(h) | - | F_{calc}(h) ||) / \sum_h | F_{obs}(h) |$, where $F_{obs}(h)$ and $F_{calc}(h)$ are the observed structure factor amplitude and the structure factor amplitude calculated from the model, respectively.

[4] The free R-factor was monitored with 5% of the data excluded from refinement.

2.7.12 Analytical Ultracentrifugation Data for MCD Variants

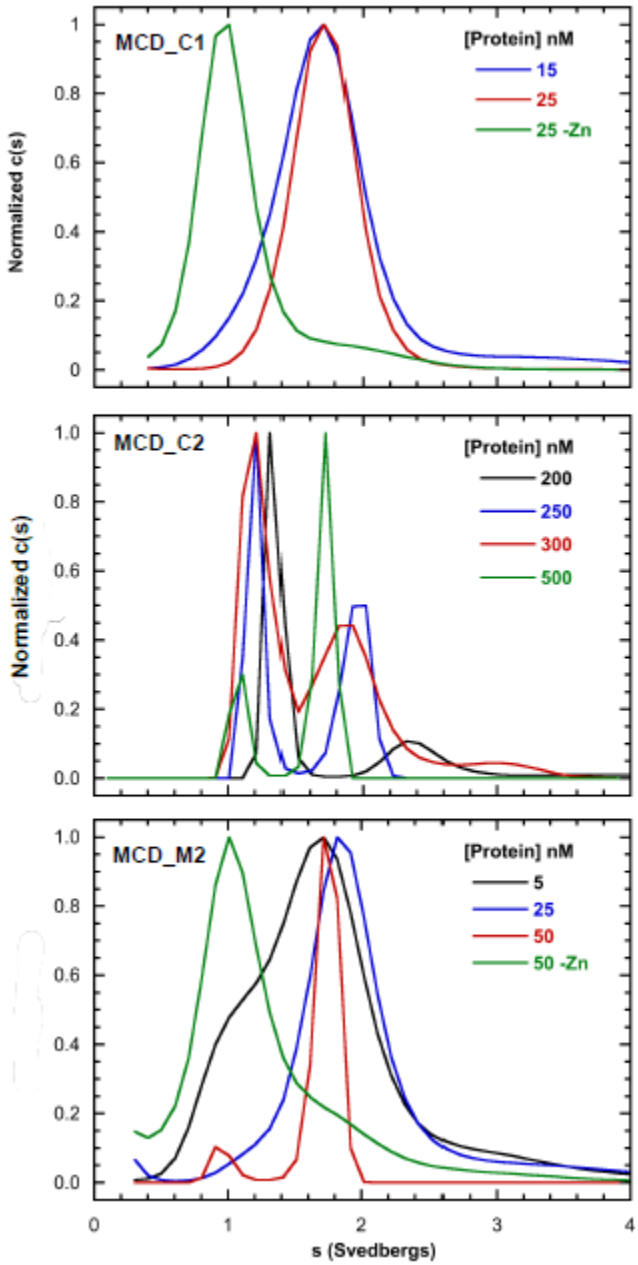

Figure 2.12: Sedimentation velocity analysis of the three MCD complexes monitored using fluorescence detection. Top, MCD_C1; middle, MCD_C2; bottom, MCD_M2. The c(s) sedimentation coefficient distributions were obtained at several concentrations for each construct and were normalized by maximum peak amplitude.

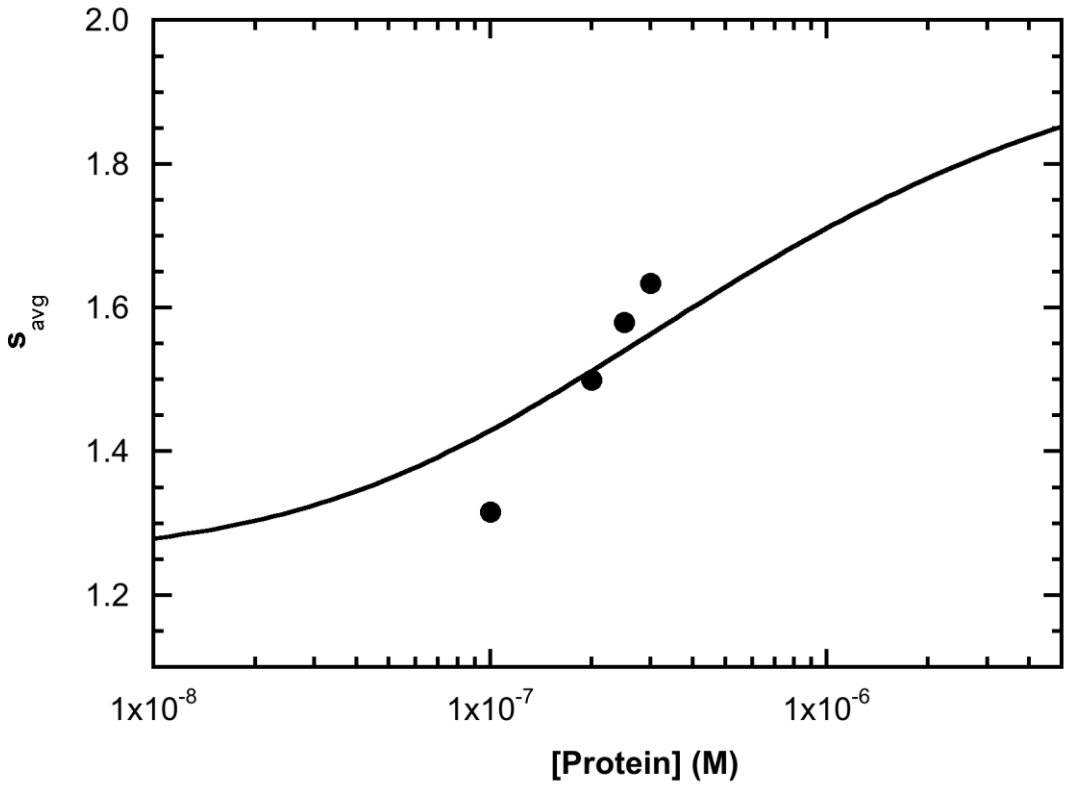

Figure 2.13: Analysis of MCD_C2 monomer-dimer equilibrium. MCD_C2 is the variant with the lowest degree of hydrophobicity at the dimer interface. The signal-average sedimentation coefficients are plotted as a function of protein concentration. The data were fit to monomer-dimer equilibrium model with the sedimentation coefficients of the monomer and dimer fixed at 1.25 S and 2.0 S, respectively. The data do not fit well to this model, suggesting that the system is not at equilibrium.

2.8 References

(1) Boyken, S. E.; Chen, Z.; Groves, B.; Langan, R. A.; Oberdorfer, G.; Ford, A.; Gilmore, J. M.; Xu, C.; Dimaio, F.; Pereira, J. H.; Sankaran, B.; Seelig, G.; Zwart, P. H.; Baker, D. *Science.* **2016**, *352* (6286), 680–687.

(2) Mills, J. H.; Sheffler, W.; Ener, M. E.; Almhjell, P. J.; Oberdorfer, G.; Pereira, J. H.; Parmeggiani, F.; Sankaran, B.; Zwart, P. H.; Baker, D. *Proc. Natl. Acad. Sci.* **2016**, *113* (52), 15012–15017.

(3) Mou, Y.; Huang, P.-S.; Hsu, F.-C.; Huang, S.-J.; Mayo, S. L. *Proc. Natl. Acad. Sci. U. S. A.* **2015**, *112* (34), 10714–10719.

(4) Fleishman, S. J.; Whitehead, T. A.; Ekiert, D. C.; Dreyfus, C.; Jacob, E. *Science.* **2011**, *332* (6031), 816–821.

(5) Fallas, J. A.; Ueda, G.; Sheffler, W.; Nguyen, V.; McNamara, D. E.; Sankaran, B.; Pereira, J. H.; Parmeggiani, F.; Brunette, T. J.; Cascio, D.; Yeates, T. R.; Zwart, P.; Baker, D. *Nat. Chem.* **2016**, *9*, 353–360.

(6) Barakat, N. H.; Barakat, N. H.; Carmody, L. J.; Love, J. J. *J. Mol. Biol.* **2007**, *366* (1), 103–116.

(7) Packer, M. S.; Liu, D. R. *Nat. Rev. Genet.* **2015**, *16* (7), 379–394.

(8) Sha, F.; Salzman, G.; Gupta, A.; Koide, S. *Protein Sci.* **2017**, *26* (5), 910–924.

(9) Arnold, F. H.; Zhang, J. H. *Trends Biotechnol.* **1994**, *12* (5), 189–192.

(10) Salgado, E. N.; Faraone-mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2007**, *129*, 13374–13375.

(11) Eric N. Salgado, R. A. L. J. F.-M.; Tezcan, F. A.; Salgado, E. N.; Lewis, R. A.; Faraone-Mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2008**, *130*, 6082–6084.

(12) Salgado, E. N.; Radford, R. J.; Tezcan, F. A. *Acc. Chem. Res.* **2010**, *43* (5), 661–672.

(13) Radford, R. J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2009**, *131* (26), 9136–9137.

(14) Radford, R. J.; Nguyen, P. C.; Ditri, T. B.; Figueroa, J. S.; Tezcan, F. A. *Inorg. Chem.* **2010**, *49* (9), 4362–4369.

(15) Der, B. S.; MacHius, M.; Miley, M. J.; Mills, J. L.; Szyperski, T.; Kuhlman, B. *J. Am. Chem. Soc.* **2012**, *134* (1), 375–385.

(16) Salgado, E. N.; Ambroggio, X. I.; Brodin, J. D.; Lewis, R. A.; Kuhlman, B.; Tezcan, F. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (5), 1827–1832.

(17) Huard, D. J. E.; Kane, K. M.; Akif Tezcan, F. *Nat. Chem. Biol.* **2013**, *9* (3), 169–176.

(18) Wunderlich, M.; Martin, A.; Staab, C. A.; Schmid, F. X. *J. Mol. Biol.* **2005**, *351* (5), 1160–1168.

(19) Thoms, S.; Max, K. E. A.; Wunderlich, M.; Jacso, T.; Lilie, H.; Reif, B.; Heinemann, U.; Schmid, F. X. *J. Mol. Biol.* **2009**, *391* (5), 918–932.

(20) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276* (January 1993), 307–326.

(21) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40* (4), 658–674.

(22) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67* (4), 235–242.

(23) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1997**, *53* (3), 240–255.

(24) Emsley, P.; Cowtan, K. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12 I), 2126–2132.

(25) Krissinel, E.; Henrick, K. *J. Mol. Biol.* **2007**, *372* (3), 774–797.

(26) Vaguine, A. A.; Richelle, J.; Wodak, S. J. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **1999**, *55* (1), 191–205.

(27) Laskowski, R. A.; MacArthur, M. W.; Moss, D. S.; Thornton, J. M. *J. Appl. Crystallogr.* **1993**, *26* (2), 283–291.

(28) Zheng, H.; Cooper, D. R.; Porebski, P. J.; Shabalin, I. G.; Handing, K. B.; Minor, W. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 223–233.

(29) Schuck, P. *Biophys. J.* **2000**, *78* (3), 1606–1619.

(30) Laue, T.; Shah, B.; Ridgeway, T.; Pelletier, S. *R. Soc. Chem. Ed Harding S, Rowe A Hort. J.* **1992**.

(31) Huang, P. S.; Love, J. J.; Mayo, S. L. *J. Comput. Chem.* **2005**, *26* (12), 1222–1232.

(32) Huang, P.-S.; Love, J. J.; Mayo, S. L. *Protein Sci.* **2007**, *16* (12), 2770–2774.

(33) Dahiyat, B. I.; Mayo, S. L. *Science.* **1997**, *278* (5335), 82–87.

(34) Shukla, U. J.; Marino, H.; Huang, P. S.; Mayo, S. L.; Love, J. J. *J. Am. Chem. Soc.* **2004**, *126* (43), 13914–13915.

(35) Stranges, P. B.; Machius, M.; Miley, M. J.; Tripathy, A.; Kuhlman, B. *Proc. Natl. Acad. Sci.* **2011**, *108* (51), 20562–20567.

(36) Gallagher, T.; Alexander, P.; Bryan, P.; Gilliland, G. L. *Biochemistry* **1994**, *33*, 4721–4729.

(37) Laitaoja, M.; Valjakka, J.; Jänis, J. *Inorg. Chem.* **2013**, *52* (19), 10983–10991.

Chapter 3: Design of a Metal-Controlled Heterodimer

3.1 Abstract:

The field of protein design strives to engineer new molecules that interact in a specific, controlled manner to form novel functional complexes[1,2,3]. Engineered proteins that generate specific complexes upon the addition of an exogenous agent, such as metal ions, will likely be integral elements of these efforts. Molecular control over protein assembly and disassembly is possible through the introduction of novel metal-binding sites at precise locations in monomeric proteins. These methods have primarily generated metal-mediated and metal-controlled homodimers[4,5,6,7,8]. To broaden the repertoire of biomaterial-based tools, we undertook the re-engineering of a metal-controlled homodimer into a metal-controlled heterodimer. The generation of a metal-controlled heterodimer would increase our understanding of metal-mediated associations and improve the utility of metal-induced protein dimerization systems.

3.2 Introduction:

Modifications of the β1 domain of streptococcal protein G (Gβ1) has led to the generation of a highly stabilized constitutive dimer[9], a *de novo* heterodimer[10], and metal-controlled homodimers[11]. In the previous chapter we described the successful design and biophysical characterization of three metal-controlled homodimer proteins[11]. This work was based on a constitutive homodimer of Gβ1 that was generated via directed evolution by another research group[9]. The thermally stabilized constitutive homodimer, termed Gβ1-M2 (PDB entry 3FIL), was utilized by our research group as a design scaffold for the generation of metazal-controlled homodimers[9]. Gβ1-M2 is a four-fold mutant of wild-type Gβ1 that forms a symmetric homodimer through extension of the β-sheet. The four mutations that cause Gβ1 to dimerize are

located at the dimer interface and consisted of the following mutations: E15V, T16L, T18I, and N37L[9].

We generated metal-controlled homodimers by incorporating hydrophobic mutations found in Gβ1-M2 (Figure 3.1 A) and metal binding residues L12H, V29H, and Y33H. Our method generated two intermolecular metal coordination sites at the symmetric protein-protein interface. These mutant variants of Gβ1 are completely monodisperse in the monomeric state, yet form clean, monodisperse dimers upon the addition of zinc (II). The metal-controlled dimer with the strongest binding affinity, MCD_C1, contains the followin0g mutations relative to wild-type Gβ1: L12H, T16L, V29H, Y33H, and N37L (PDB entry 6NL8, Figure 3.1 B)[11]. The metal-controlled homodimer interface of MCD_C1 contains hydrophobic interfacial residues and two tetrahedral zinc (II) coordination sites. The symmetric interface of MCD_C1 contains leucine residues at interfacial positions 16 and 37 and interface bridging histidine residues at positons 12, 29, and 33. These five mutations promoted a metal-controlled dimer variant with very high binding affinity (*i.e.*, < 75 pM)[11].

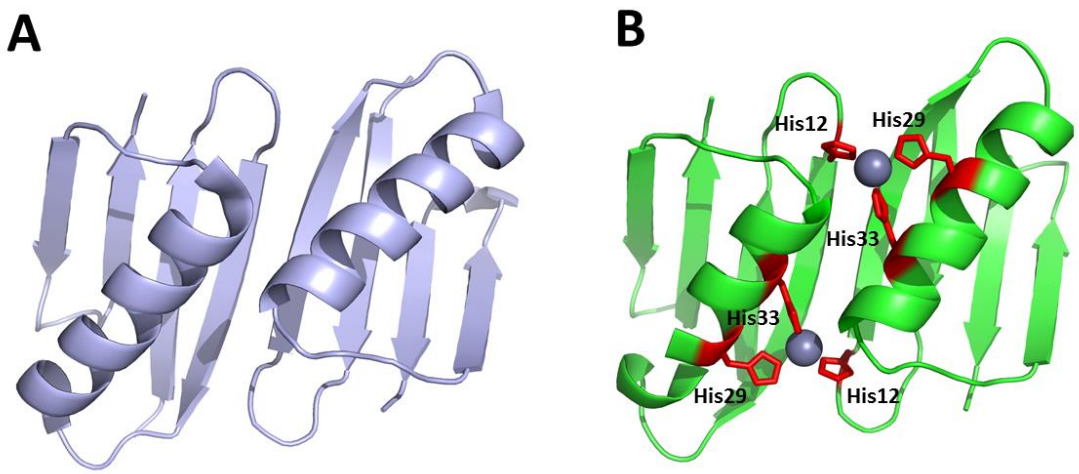

Figure 3.1: Comparison of Gβ1-M2 and MCD_C1. A) Gβ1-M2 is a constitutive dimer of Gβ1 (PDB entry 3FIL)[9]. B) MCD_C1 (PDB entry 6NL8)[11] contains two metal binding sites using histidine residues at positions 12, 29, and 33. The zinc (II) metal is colored grey.

A number of natural multimeric complexes[12] in addition to our metal-controlled homodimers form symmetric structures. Advantages associated with protein complexes that bind symmetrically might include enhanced binding enthalpy, greater thermal stability, enhanced protein folding, reduced aggregation, and the possibility for allosteric regulation[13]. The interface of a symmetric protein complex contains sets of pair-wise interactions, which double energetically favorable interactions. Strong binding for both homodimers and heterodimers may occur when hydrophobic residues are matched on interacting surfaces[14]. Relative to heterodimeric structures, homodimers can be more energetically favorable as symmetry elements can function to enhance self-association and interaction propensity[14].

The conversion of a homodimer protein to a heterodimer is a potentially challenging process due to the loss of favorable energetic contributions associated with symmetrically related elements. Therefore one can potentially expect concomitant loss of binding affinity upon conversion of favorable symmetry elements at the interface to asymmetric elements. In the

evolution of heterodimers from homodimers, nature invariably used both negative and positive selection for the generation of stable heterodimeric interactions. An example of negative selection, or design, which functions to prevent the formation of self-associated structures, is the positioning of charged residues across the interface on proteins that are meant to interact with one another. The opposite charges at the heterodimeric interface provides favorable binding energy (i.e., positive design), while also providing negative design elements that prevent homodimer formation through like charge repulsive forces. Therefore the incorporation of charged residues at the interacting surface of homodimers represents a positive and negative design strategy for the generation of heterodimers.

Successful design strategies for engineering a homodimer into a heterodimer have involved remodeling a symmetric complex with complementary steric and/or favorable electrostatic interactions[15,2,16]. In the design of complementary steric interactions, the surface of a symmetric protein may be modified to generate complementary knobs and holes. Knobs are created by replacing a small amino acid with a larger one, while holes are created by replacing a large amino acid with a smaller one. In the design of favorable electrostatic interactions, complementary charged residue pairs at the interface are involved in promoting and stabilizing heterodimer assembly. The engineering of charged residues at the interface of a symmetric assembly are thought to prevent self-association. Both design strategies required the alteration of hydrophobic residues at the heterodimeric protein-protein interface to prevent unwanted homodimer formation[16,15]. Another research group used computational protein design to redesign a homodimer into a heterodimer by incorporating both steric and electrostatic interactions at the protein-protein interface. This design strategy generated a new protein pairs

that approach 80 – 90% pure heterodimers, suggesting some amount of homodimerization remains[2].

To engineer a metal-controlled heterodimer, we used a metal-controlled homodimer, MCD_C1, as the starting design scaffold. The engineering of a metal-controlled heterodimer entailed the design of two distinct protein binding partners, instead of only one for the previous metal-controlled homodimer project. Difficulty associated with engineering a metal-controlled heterodimer is that each binding partners might exhibit a higher propensity to bind to itself, as opposed to selectively binding to the designed binding partner. In the conversion of a symmetric metal-controlled homodimer into a heterodimer it was necessary to not only redesign the metal-associated aspects of a homodimer, but also the metal independent hydrophobic interactions known to exist at the dimer interface. This was necessary to avoid unwanted interactions that can lead to metal-controlled homodimers, metal-independent heterodimers, metal-independent homodimer, as well as modifications that would abrogate dimer formation.

To generate a metal-controlled interaction between two different proteins, we used rational design to modify the recognition specificity of MCD_C1. Previous strategies to modify the recognition specificity of a protein-protein interface has led to the generation of protein pairs with altered interacting specificity[3]. For example, the recognition specificity of a DNase-inhibitor protein pair is a function of complementary charges and charged interactions at the protein-protein interface[17]. The re-engineering of the DNase-inhibitor protein pair involved altering charged interactions and sequence changes in both binding partners to generate a new protein pair[3]. In our project, we envision one of the metal-binding sites would be removed to change the recognition specificity and binding mode of MCD_C1. The site of metal-binding

interaction specificity would be replaced with favorable interactions that destabilize the original metal-controlled homodimer and stabilize the designed metal-controlled heterodimer.

Proteins, protein-protein interfaces, and macromolecular assemblies frequently contain salt bridges. These geometrically specific interactions may contribute to conformational stability, molecular recognition, and catalysis[18]. Salt bridges may contain basic side chain residues (e.g., histidine, arginine, and lysine) and acidic side chain residues (e.g., aspartic acid, glutamic acid). Oppositely charged residues were defined as interacting if the Nitrogen and Oxygen atoms of the side chains are within 4 Å[18]. A simple salt bridge is defined as an interaction between a single pair of oppositely charged side chain, while a complex salt bridge involves interactions among two or more charged side chains[19]. Complex salt bridges, sometimes referred to as "networked" salt bridges, often involve interactions with multiple acidic side chains and one basic side-chain in single or adjacent chains[18]. Complex salt bridges may play important roles in driving complex formation of protein subunits or joining proteins in protein assemblies[19].

We re-engineered the specificity of MCD_C1, a metal-controlled dimer, to generate a new metal-controlled heterodimer. Our design strategy required mutations that promote and stabilize the metal-controlled heterodimer (*i.e.* 'positive state' design) and simultaneously destabilize metal-controlled homodimeric states and metal-independent homodimeric states (*i.e.* 'negative state' design). Modification of the MCD_C1 interface involved replacing one the zinc tetrahedral coordination sites with complementary charged residues at the design protein-protein interface. Optimizing the specificity and chemical control of oligomeric association required elements of both 'positive and negative state' design.

3.3 Experimental Methods:

3.3.1 Gene Synthesis, Protein Expression, and Purification

The scaffold uses for the metal-controlled heterodimer is the β1 domain of streptococcal protein G (Gβ1). The β1 domain consists of amino acids 228-282 of immunoglobulin protein G (PDB entry 1PGA). The genes for variants of Gβ1 were synthesized using standard polymerase chain reaction (PCR) methods and cloned using restriction sites Nde1 (5'-end) and EcoR1 (3'-end) in the pet21a (Novagen, San Diego, CA). *E. coli* strain BL21(DE3) was chemically transformed with the corresponding plasmids and grown to an OD_{600} ~0.8. Protein expression was induced for ~3 hours upon addition of IPTG at a final concentration of 1 mM. Cells were pelleted by centrifugation at 6,000 rpm for 10 minutes and stored in a -80 °C freezer. A freeze-thaw protein extraction cycle was utilized and included incubation on ice for 30 minutes, room temperature for 15 minutes, followed with a dry-ice ethanol bath for 10 minutes, and repeated three times. The cell pellet was gently resuspended in PBS, pH 6.8 for 60 minutes. The sample was centrifuged at 10,000 rpm for 30 minutes and acetonitrile was added to the resulting supernatant (to 30% v/v) to induce precipitation of impurities. The resulting supernatant was lyophilized to remove the acetonitrile-water. The lyophilized sample was resuspended in water and purified using a Varian 10 micron C8 HPLC preparative reverse-phased column with a linear 1%/min acetonitrile/water gradient containing 0.1% trifluoroacetic acid. The purified protein was lyophilized and the resulting dry protein suspended in water. The pH was adjusted to ~7 and buffered with 20 mM Tris, 100 mM sodium chloride, 200 ppm sodium azide, pH 7.0. Protein purify was assessed using SDS-PAGE and protein concentration determined by absorbance at 280 nm using theoretical extinction coefficients. The proteins were concentrated using centrifugal filtration to approximately 2 mM (~12 mg/mL) and stored at -20 ^{o}C.

3.3.2 SEC-MALS Characterization

Stock proteins of the three binding partners were diluted to 1 mM in buffer containing the either 1) 20 mM Tris, 100 mM sodium chloride, pH 7.0 or 2) 20 mM Tris, 100 mM sodium chloride, 1 mM zinc sulfate, pH 7.0. Interactions between the different binding partners was performed by diluting both proteins to 0.5 mM, for an overall protein concentration of 1 mM. Interactions between the two binding partners, when measured in the presence of increasing concentrations of sodium chloride, were performed at a protein concentration of 100 µM, with both binding partners being 50 µM. Designed proteins were diluted in buffers containing 20 mM Tris, 1 mM zinc sulfate and varying concentrations of sodium chloride: 100, 300, 500, 1000, and 2000 mM. Size exclusion chromatography was performed using a FPLC (GE Healthcare) with UV detection at 280 nm. 100 µL of each sample was injected onto a 25 mL Superdex75 10/300 GL size exclusion column (GE Healthcare). The FPLC was connected in-line with multiangle light scattering (Wyatt miniDAWN TREOS detector system). ASTRA software (Wyatt Technologies) was used to analyze elution peaks and determine the molecular weight calculated by light scattering.

3.3.3 Circular Dichroism

CD data was collected on an Aviv 420 instrument using a 2 mm path-length cuvette. Protein concentrations were approximately 25 µM with buffer consisting of 20 mM MOPS, 10 mM sodium chloride, 25 µM zinc sulfate, 1.5 M guanidinium hydrochloride, pH 7.0. Far-UV scans confirmed that the binding partners have a well-defined alpha helical secondary structure element due to minima at 208 nm and 222 nm. The CD signal at 222 nm was monitored as a function of temperature and fit to a two-state unfolding model. Folded protein and unfolded states were measured between 25 and 95 °C.

3.3.4 X-ray Crystallography

For crystallization purposes, binding Partners A, B, and C, were expressed and purified as described above and concentrated to a minimum 2 mM. To study the effects of metal-controlled association, excess zinc sulfate was added to an equimolar solution containing 1 mM of binding Partner A and B or binding Partner B and C before crystallization. Crystals were obtained by vapor diffusion at room temperature from hanging drops. The 3 µL drop consisted of 2 mM Protein (1 mM Partner B and 1 mM Partner A or C supplanted with 4 mM zinc sulfate) was mixed with 3 µL of the crystallization solution. The optimized crystallography conditions for binding Partner A + B and B + C were 0.1 M HEPES pH 7.5 200 mM $MgCl_2$ 24 % PEG 3000 and 0.1 M HEPES pH 7.5 200 mM $MgCl_2$ 26 % PEG 10000, respectively. Optimized crystallization buffer with 5% glycerol was used for storage and transport in liquid nitrogen. Data were collected at the Berkley Lab Advanced Light Source (ALS-5.0.1). Data were processed using the programs XDS[20] and iMOSFLM[21]. Structures were determined by molecular replacement with the program PHENIX[22]. The protein Gβ1 domain (PDB entry 1PGA) was used as a molecular replacement model. The refinement was carried out with PHENIX refine using several cycles of the least-squares refinement fit interspaced with manual editing of the structure in COOT[23]. The refinement procedure usually converged with fewer than 10 macrocycles of the refinement and manual inspection. The structures were analyzed using PISA[24] and MolProbity[25]. The structures are currently being prepared for submission to the Protein Data Bank (PDB). Some additional analysis of the metal ion-binding sites was performed using the CheckMyMetal (CMM)[26] online server. Descriptive software tools such as PISA and CCM provide an estimate for binding energies and a secondary validation tool for organization of the lattice and the binding sites. Data collection and refinement statistics are listed in Table 3.9.

3.3.5 Analytical Ultracentrifugation: Sedimentation Velocity

Binding Partners B and C were subjected to sedimentation velocity analytical ultracentrifugation assessment at the *Sanford Burnham* Prebys Medical Discovery Institute. Binding Partners B and C were individually prepared in 20 mM Tris, 100 mM sodium chloride, pH 7.0 and 20 mM Tris, 100 mM sodium chloride, 25 µM zinc sulfate, pH 7.0. Binding Partners B and C were incubated together and analyzed using the previously mentioned conditions. Samples were analyzed using a Beckman Coulter ProteomeLab XL-1 analytical ultracentrifuge equipped with UV-absorption detector. Samples were loaded into both sectors of double sector cell equipped with sapphire windows. Data were collected at 21 ºC for 24 hours at a rotor speed of 42,000 rpm with a 50 Ti rotor. Data were analyzed in Sedfit[27] using the c(s) model. The molecular masses, partial specific volumes, solvent density and viscosity used in the data analysis were calculated in Sednterp[28].

3.5.6. Analytical Ultracentrifugation: Sedimentation Equilibrium

Sedimentation equilibrium experiments were performed using a Beckman Coulter ProteomeLab XL-I analytical ultracentrifuge at the *Sanford Burnham* Prebys Medical Discovery Institute. Samples were prepared in 20 mM Tris, 100 mM sodium chloride, 60 µM zinc sulfate, pH 7.0 buffer. Protein samples at concentrations 40 µM, 20 µM, and 10 µM were loaded in a 6-channel equilibrium cells and spun in an An-50 Ti Analytical 8-place titanium rotor at 21 ºC until equilibrium was reached. Data sets for Binding Partner B, Binding Partner C, and Binding Partners B + C were collected at 30,000 rpm and 35,000 rpm. HeteroAnalysis AUC software was used for the data analysis of the binding Partners. The analytical ultracentrifugation data sets determining apparent molecular weight were fitted to the single ideal species model. The data sets determining the association constant were fitted to a monomer-dimer model.

3.4 Results

3.4.1 Rational Design of Metal-Controlled Heterodimer

For the work reported herein a rational design approach was used to re-engineer a metal-controlled homodimer into a metal-controlled heterodimer. To engineer a metal-controlled heterodimer, we used a metal-controlled homodimer, MCD_C1, as the starting design scaffold[11]. MCD_C1 contains the following mutations relative to wild-type Gβ1: L12H, T16L, V29H, Y33H, and N37L (PDB entry 6NL8)[11]. The metal-controlled homodimer interface of MCD_C1 is formed from hydrophobic interfacial residues and two tetrahedral zinc (II) coordination sites. The engineering of a metal-controlled heterodimer entailed the design of two distinct protein binding partners, instead of only just one for the previous metal-controlled homodimer project.

The intermolecular binding interactions we previously designed to drive successful metal-controlled dimer formation was made up of histidine residues from one monomer (located in standard *i* to *i* + 4 α-helical positions) across from a histidine residue located on the other binding partner[11]. These residues generated two tetrahedral zinc coordination sites, where three ligands are histidine and the fourth ligand consisted of a chloride ion[11]. These symmetric metal-controlled homodimers have two metal binding sites per dimer. The monomer that provides two histidine residues at one side provides only one histidine at the other site, and vice versa (Figure 3.1 B).

To generate two distinct binding partners, structure-based design was utilized to eliminate one of the metal binding interaction and replace it with "networked" salt bridges. The goal was to create a metal-controlled heterodimer by providing two distinct sites of intermolecular interactions that are meant to bridge the dimer interface and drive complex formation: 1) a metal

binding site consisting of three histidine residues and 2) a site made up of residues that form intermolecular salt-bridges driven by favorable electrostatic interactions.

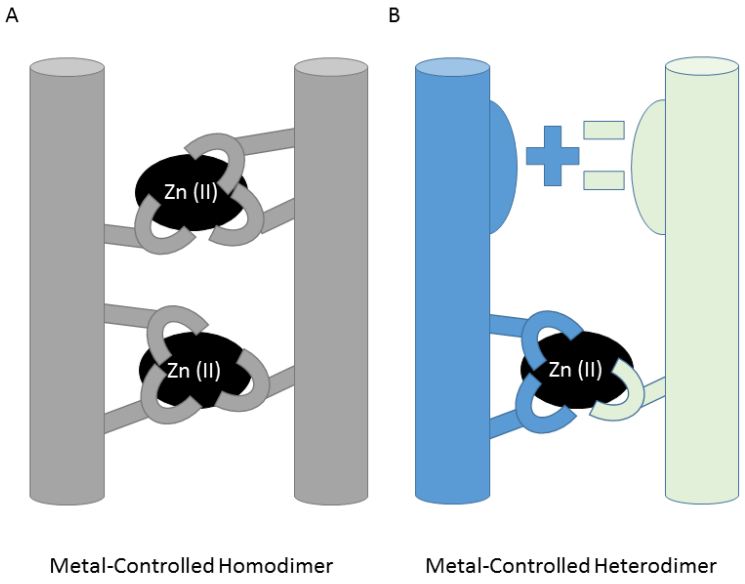

Figure 3.2: Cartoon Diagram of the Metal-Controlled Homodimer and Metal-Controlled Heterodimer. A. The metal-controlled dimer has two metal-coordination sites. B. A metal coordination site was replaced with charged residues to generate "networked" salt bridges.

In an attempt to enhance heterodimer formation and binding affinity, and inspired by naturally occurring "networked" salt bridges (i.e., 2:1 ratio of acidic to basic side-chains), we converted one of the metal binding sites to "networked" salt bridges consisting of an arginine residue on one monomer across from two glutamic acid residues on the other. To achieve this we substituted the single histidine from the metal binding center with a basic residue (arginine). Arginine is planar and has five possible sites for hydrogen bonding, while Lysine is nonplanar and has three possible sites. The arginine side chain has the ability to form two hydrogen bonds on one side of the side-chain and one hydrogen bond on the other. Arginine was found to be the most abundant charged residue at the interface of protein-protein interactions and complex salt bridges[19,29]. We chose to use an arginine residue, as opposed to histidine and lysine, due to its

prevalence in naturally occurring salt bridges, the number of sites possible for hydrogen bonds, as well as charge density.

To enhance the likelihood that the designed monomers of the heterodimeric pair form favorable hydrogen bonding interactions, the arginine residue was positioned across from two amino acids which have acidic side chains. This was done so that the engineered arginine residue could potentially form favorable hydrogen bonding interactions with the carboxyl functional groups of cross-monomer aspartic or glutamic acid residues. In natural protein complexes, the prevalence of glutamic and aspartic acid in salt bridges are approximately equal, with a slight preference for glutamic acid[19]. The longer chain length of glutamic acid might promote more conformational freedom, while also providing less steric hindrance upon interacting with arginine[19]. We ultimately choose to incorporate two glutamic acid residues in an attempt to generate intermolecular salt bridges. One might imagine many different possible complementary charged interactions may be designed in the engineering of intermolecular salt bridges.

We replaced a metal-binding site in the metal controlled dimer with charged residues to engineer an intermolecular salt bridge between an arginine residue on one binding partner and two glutamic acid residues on the other heterodimeric binding partner. The two distinct binding partners are referred to herein as binding Partner A and B. Binding Partner A consists of a histidine residue at position 12, which is positioned across the metal-binding histidine residues at positions 29 and 33 on binding Partner B. To generate "networked" salt bridges, positions 29 and 33 on binding Partner A were engineered to contain glutamic acid residues. Therefore, for binding Partner B the conserved metal-binding site consists of histidine residues at positions 29 and 33, and a positively charged arginine residue at position 12. These mutations ultimately

resulted in the conservation of one metal binding site, and the conversion of the other to novel intermolecular salt bridges made up of an arginine residue from binding Partner B, and two glutamic acid residues from binding Partner A.

The juxtaposition of the arginine and glutamic acid residues, across the dimer interface, were meant to generate favorable intermolecular salt bridges (Figure 3.3 A). Using a central arginine residue, the proposed orientation of the putative "networked" salt bridges contain two nitrogen groups contributed by arginine 12 (binding Partner B) that form cross-interface electrostatic interactions with the carboxyl group of glutamate 29 (binding Partner A). In addition, a third nitrogen from Arginine 12 would interact favorably with the carboxyl group from glutamate 33 (binding Partner A).

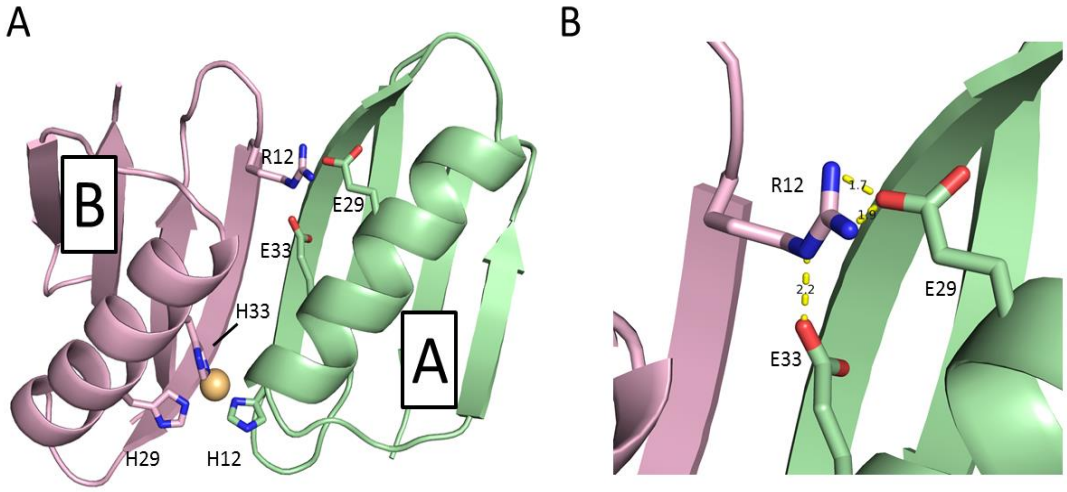

Figure 3.3: Designed interactions between binding partner A and B. A) Two distinct interactions between A and B are 1) a metal binding site consisting of three histidine residues and 2) a site consisting of residues that form intermolecular salt-bridges. B) Distances between Nitrogen and Oxygen atoms in binding Partners A and B. Zinc is represented by an amber sphere.

3.4.2 Analysis of Heterodimeric Binding Partners

Size exclusion chromatography (SEC) undertaken in the presence and absence of metal ions, was used to investigate metal-dependent and metal-independent interactions, respectively.

Molecular weights for the protein variants were determined from size-exclusion chromatography coupled with multi-angle light scattering (SEC-MALS). Analysis of binding Partner A by itself revealed that it is a monomer in the absence of metal, yet forms a metal-controlled homodimer upon the addition of 1 mM zinc. A bidentate ligand arrangement was formed through glutamic acid residues that are in close proximity such as the designed amino acids at positions 29 and 33. This arrangement may contribute to homodimer formation for binding partner A through the chelation of metal ions. The metal-controlled interface of binding Partner A contains interfacial metal binding sites, designed amino acids at positions 29 and 33, and leucine residues at positions 16 and 37 (located at the hydrophobic dimer interface). The combination of these interfacial residues appears to promote an unwanted metal-controlled homodimer. Upon subjecting binding Partner B to the same analysis, the results revealed that Binding Partner B remains monomeric in both the presence and absence of zinc (II) ions. This is likely due to the fact that arginine residues do not contain a metal-coordinating side chain.

To determine if these two newly designed variants would preferentially form a heterodimer, binding Partners A and B were mixed together in a one-to-one molar ratio in both the presence and absence of zinc (II) ions. The theoretical size of a dimeric complex between binding Partners A + B complex is 12.4 kDa. In the absence of metal, associations between binding Partners A + B resulted in an average molecular weight of 11.1 kDa. In comparison, for the sample in which metal was added, binding Partner A and B exhibited an average molecular weight of 13.4 kDa. Thus, metal-mediated and metal-independent interactions may occur between binding Partners A + B (Figure 3.4).

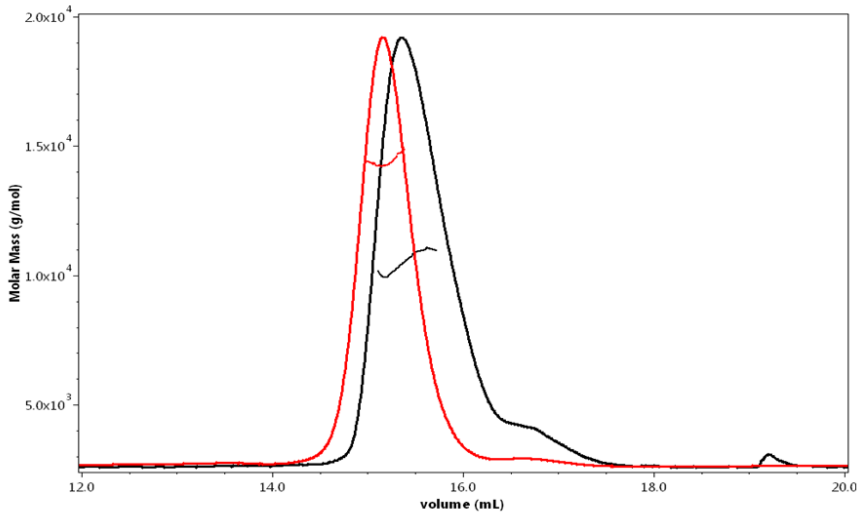

Figure 3.4: SEC-MALS Characterization of Binding Partners A and B. Interaction of Binding Partners A and B, 0.5 mM of each protein was loaded onto a Superdex 75 column. Multi-angle light scattering was collected in the absence of metal (black) or presence of 1 mM Zinc sulfate (red). The molecular weight from the interactions of binding Partner A and B were 11.1 and 13.4 kDa in the absence (black) and presence of zinc (red), respectively.

Based on these results we believe a metal-mediated heterodimer interaction occurs between binding Partner A and B for the following reasons. First, SEC-MALS results indicated an interaction occurs between binding Partners A + B in the absence of metal. Second, if metal-mediated interactions between binding Partners A + B did not occur, then we should observe equivalent monomeric and dimeric peaks in SEC-MALS, which would be associated with binding Partner B and binding Partner A, respectively. Our goal of generating a metal-controlled heterodimer is hindered by metal-independent and metal-dependent interactions between binding Partners A + B. To design a metal-controlled heterodimer, it is critical that each individual binding partner remains monomeric in the absence and presence of zinc.

3.4.3 Redesign of Binding Partner A

SEC-MALS results indicate that binding Partner A remains monomeric in the absence of metal, yet forms a metal-controlled homodimer in the presence of zinc. In an attempt to reduce the propensity of binding Partner A to form a homodimer in the presence of zinc, the designed

glutamic acids residues at positions 29 and 33 were left intact, and the hydrophobic nature of the dimer interface was modified to reduce its propensity to self-associate in the presence of zinc. Modification of the dimer interface might also reduce favorable interactions with binding Partner B, with interactions between binding Partners A and B occurring in the absence of zinc. In an attempt to reduce these unwanted interactions the amino acids at positions 16 and 37 were mutated to valine residues. This resulted in the design and production of the third variant, binding Partner C, which contains valine residues at positions 16 and 37, while retaining histidine at position 12 and the designed glutamic acid residues at positions 29 and 33.

Table 3.1: Amino Acid Mutations at the Heterodimeric Metal-Controlled Protein Interfaces

	Metal-Controlled Dimer	Binding Partner A	Binding Partner B	Binding Partner C
Hydrophobic Interface	T16L, N37L	T16L, N37L	T16L, N37L	T16V, N37V
Metal Coordination	L12H, V29H, Y33H	L12H	V29H, Y33H	L12H
Electrostatic Interactions		V29E, Y33E	L12R	V29E, Y33E

3.4.4 Biophysical Characterization of Binding Partners A, B, and C

A project goal was to generate a heterodimer in which each binding partner remains monomeric in the absence and presence of zinc, when incubated by themselves, yet form a heterodimer only with one another in the presence of zinc. Previous analysis indicated binding Partner B is monomeric in the absence and presence of zinc. SEC-MALS analysis performed on binding Partner C revealed this variant remains monomeric in the presence and absence of zinc (Figure 3.11). As opposed to binding Partner A, binding Partner C remains monomeric in the presence and absence of zinc and therefore no further modifications were needed for this variant.

In the absence of metal, the SEC-MALS retention time suggests binding Partner B + C do not interact, as the measured molecular weight, 7.0 kDa, reveals both these variants remain monomeric. Upon the addition of zinc (II), binding Partners B + C formed a metal-controlled heterodimer with a calculated molecular weight of 12.0 kDa. Binding Partners B and C do not form metal-mediated homodimers or exhibit evidence of interaction in the absence of zinc.

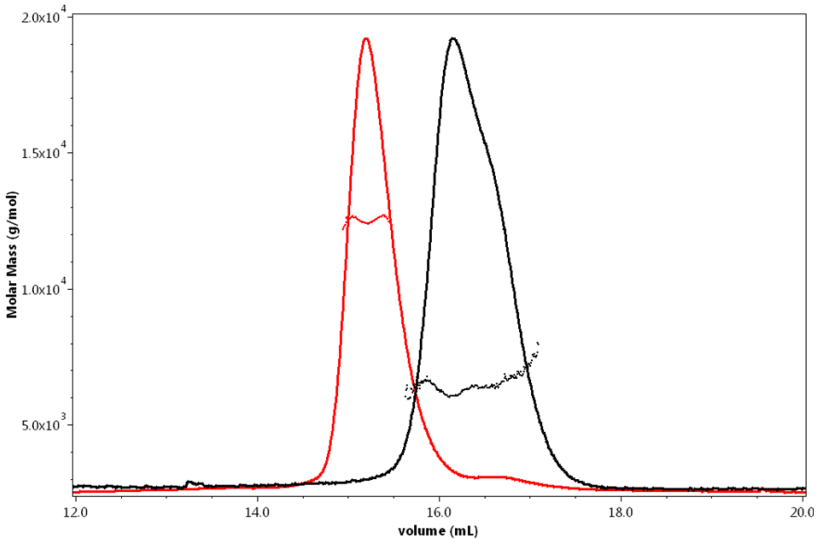

Figure 3.5: SEC-MALS Characterization of Binding Partner B and C. B. When incubated together binding Partners B and C do not exhibit evidence of interaction and have a calculated molecular weight of 7.0 kDa. In the presence of zinc, Binding Partners B + C form a metal-controlled heterodimer with a calculated molecular weight of 12.0 kDa.

Thermally-induced unfolding of the binding Partners was monitored by circular dichroism. The high thermal stability exhibited by the parent constative dimer, Gβ1-M2, as well as our designed metal-controlled dimers, require that the thermal unfolding experiments be performed in the presence of 1.5 M Guanidine-HCl[9,11]. In the presence of equimolar zinc (II), the thermal unfolding temperatures (Tm) were measured for the different binding partners. The measured melting temperatures for binding Partners A, B, and C were 73.5 ^{o}C, 78.0 ^{o}C, and 56.0 ^{o}C, respectively. The melting temperatures measured for metal-controlled interactions between binding Partners A + B and B + C were 76.9 ^{o}C and 68.8 ^{o}C, respectively. It was determined

that metal-controlled interactions between binding Partner A + B and B + C increased the thermal stability by ~1^{o}C.

Table 3.2: Biophysical Characterization of Binding Partners A, B, and C

	Binding Partner A	Binding Partner B	Binding Partner C	Binding Partner A + B	Binding Partner B + C
Molecular Weight of Native State (kDa)	6.8	7.1	6.9	11.1	7.0
Native State	monomer	monomer	monomer	dimer	monomer
Molecular Weight upon addition of metal (kDa)[a]	12.2	7.5	8.3	13.4	12.0
State upon addition of metal[a]	dimer	monomer	monomer	dimer	dimer
Molecular Weight upon addition of metal (kDa)[a]	12.2	7.5	8.3	13.4	12.0
Tm (°C) with zinc[b]	73.5	78.0	56.0	76.9	68.8

a[determined by SEC-MALS in 1 mM zinc sulfate]b[determined by CD in buffer containing 25 µM Zinc]

3.4.5 Analytical Ultracentrifugation Experiments

In addition to subjecting the metal-controlled complexes to SEC-MALS analysis, the molecular weights were also measured using two different analytical ultracentrifugation methods (i.e., sedimentation velocity and equilibrium experiments). For the velocity experiments, molecules migrate as a function of size, shape, and interaction with other macromolecules[30]. For the equilibrium experiments, molecules migrate and separate until an equilibrium state is reached, which provides a distribution of associated and non-associated states. When performing equilibrium experiments under different protein concentrations, one can accurately determine the molecular weights and binding affinities for self- and hetero-associations[30].

Binding Partners B and C were subjected to sedimentation velocity analysis to measure the molecular weight of the individual components, in addition to potential interaction between

Partners B + C. Binding Partner B and binding Partner C were evaluated individually in the absence and presence of zinc. The measured molecular weights for binding Partner B and C indicated the proteins remain monomeric under both experimental conditions. Interactions between binding Partners B + C were not detected in the absence of zinc (Figure 3.20). Binding Partners B + C form a metal-controlled heterodimer with an apparent molecular weight of 11.7 kDa (Figure 3.22).

Table 3.3: Apparent Molecular Weight and Binding Affinity Determined by Analytical Ultracentrifugation

Method	Parameter	Binding Partner B, zinc	Binding Partner C, zinc	Binding Partner B + C, zinc
AUC-SV	Apparent Molecular Weight (kDa)	7.1	6.5	11.7
AUC-SE	Binding Affinity at 30,000 rpm (μM)	2,800	571	18.3
	Binding Affinity at 35,000 rpm (μM)	1,500	1,456	19.4

For the sedimentation equilibrium experiments, binding Partners B, Partner C, and Partners B + C were evaluated over a broad concentration range in the presence of zinc (II). Binding Partner B remains monomeric and has an apparent molecular weight of 6.7 kDa. Binding Partner C has an apparent molecular weight of 7.8 kDa, which is ~ 25% higher than the monomer molecular weight. Fitting binding Partner C to a monomer-dimer equilibrium model yields a dissociation constant of 571 μM and 1,456 μM at 30,000 and 35,000 rpm, respectively. Binding Partner B + C have an apparent molecular weight of 11.7 kDa, and a binding affinity of ~18 and ~19 μM at 30,000 and 35,000 rpm, respectively. The specific nature and strength of the metal-controlled binding Partner B + C interaction was confirmed by similar dissociation constants at two different rotor speeds.

3.4.6 Salt Dependence on Dimer Formation

The designed salt bridges consisting of an arginine residue on binding Partner B and two glutamic acid residues on binding Partner C. To further confirm that the salt bridges contributes favorable enthalpy for complex formation we disrupted dimer formation under increasing concentrations of sodium chloride. All samples were measured using SEC-MALS in a buffer that consists of 20 mM Tris, 1 mM Zinc sulfate, pH 7.0 with increasing concentrations of sodium chloride. The AUC determined dissociation constant for the metal-controlled heterodimer (binding Partner B and C) is approximately 19 µM. Preliminary SEC-MALS analysis was performed with the protein concentration at 1 mM, which is approximately 50 times the dissociation constant. To ensure that the complexes were closer to an ideal state of dynamic equilibrium the concentrations for both binding partners were reduced to 100 µM, which is approximately 5 times the AUC determined dissociation constant.

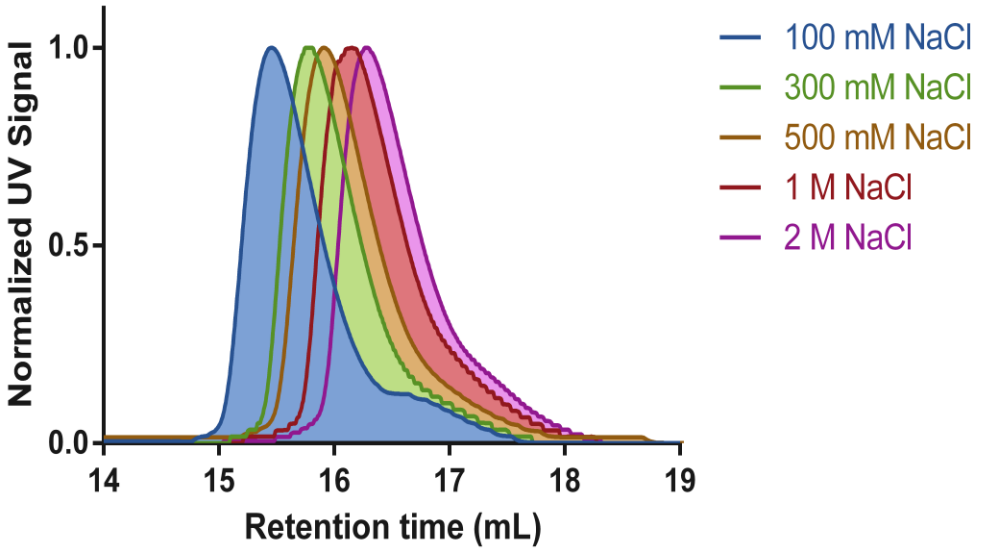

Figure 3.6: SEC-MALS measured Salt-Dependence on Dimer Formation. The intermolecular salt bridge of the metal-controlled heterodimer is disrupted upon increasing concentrations of sodium chloride.

In comparison to the preliminary SEC-MALS analysis, the protein concentration was reduced to 100 µM of Binding Partner B + C with increasing concentrations of sodium chloride. At 100 mM sodium chloride, the metal-controlled heterodimer has a molecular weight of 10.7 kDa, which is comparable with the molecular weight of 12.0 kDa determined at a higher protein concentration (1 mM). As the salt concentration was increased, the measured molecular weight is reduced from a dimer to a value observed for an individual binding partner. Binding Partner B + C are effectively monomeric when incubated in the presence of 1 M and 2 M sodium chloride. The results of this analysis indicates complex formation is dependence on salt concentrations. This provides additional evidence that the designed salt-bridges provides favorable binding energy and contributes to dimer formation.

Table 3.4: Salt-Titration Molecular Weights determined by SEC-MALS

Binding Partner B + C	Molecular Weight (kDa)
100 mM Sodium Chloride	10.7
300 mM Sodium Chloride	9.9
500 mM Sodium Chloride	8.4
1 M Sodium Chloride	7.5
2 M Sodium Chloride	6.9

Determined by SEC-MALS in [20 mM Tris, 1 mM zinc sulfate, additive, pH 7.0]

3.4.7 X-Ray Crystal Structures of Binding Partners A + B and B + C

Binding Partners A, B, and C were generated from the MCD_C1 protein design scaffold. As with the structure of the parent variant, MCD_C1, the structures of binding Partner A + B and B + C adopt a dimeric head-to-head assembly in which a heterodimer is formed through extension of the β-sheets. The structures and cross-interface hydrogen bond register for Gβ1-M2,

MCD_C1, and binding Partners A + B, and binding Partners B + C are very similar. For all the structures, there are six interfacial hydrogen bonds: two between positions 17 and 13′, two between positions 15 and 15′, and two between positions 13 and 17′ (the numbering for cross-interface positions is denoted with a prime).

The designed zinc coordination site of binding Partners A + B and B + C are formed by two histidine residues located in *i* to *i* + 4 positions (binding Partner B) on the Gβ1 α-helix and an additional cross-interface interaction involving a histidine at position 12 (binding Partner A and C), which is located on the turn between the first and second β-strands. Tetrahedral zinc coordination involves three histidine ligands and a chloride ion. Analysis of the crystal structure of the binding Partners A + B and B + C complexes, demonstrates that the distances and bond angles between the zinc atom and the three histidine ligands are within range of standard tetrahedral coordination geometry (Tables 3.7 and 3.8).

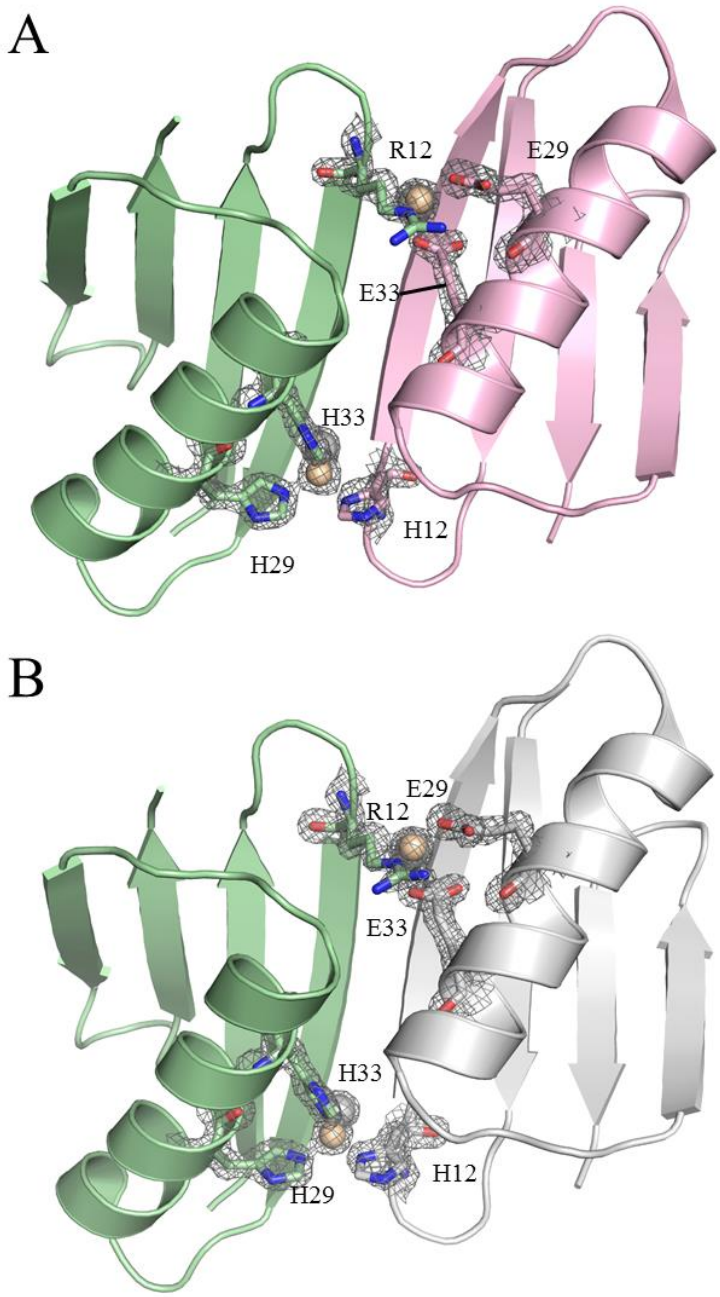

Figure 3.7: Crystal structures of Binding Partners A + B and B + C (A) Binding Partners A + B and (B) Binding Partners B + C. The designed metal–ligand histidine residues are illustrated as stick bonds, with the corresponding 2Fo – Fc omit electron density map (1.5σ). Zinc atoms are illustrated as amber spheres, and chlorine atoms as light grey spheres

In addition to the intermolecular zinc coordination site, the interface of binding Partners A + B contained "networked" salt bridges. These "networked" salt bridges contains a central arginine residue being surrounded by two glutamic acid residues. The side chain of arginine

contains a guanidinium group, which provides five possible sites for electrostatic interactions. Analysis of the crystal structure of binding Partners A + B demonstrates that the glutamic acid at position 29 forms electrostatic interactions with Nη and Nε sites of the arginine side chain (Figure 3.8 B). In addition, the arginine side chain is stabilized through additional electrostatic interactions with glutamic acid at positon 33. The designed salt bridges between binding Partners A + B was confirmed by the arrangement and distances between residues. Near the site of the "networked" salt bridges, there was an additional zinc (II) and chloride ion embedded at the protein-protein interface.

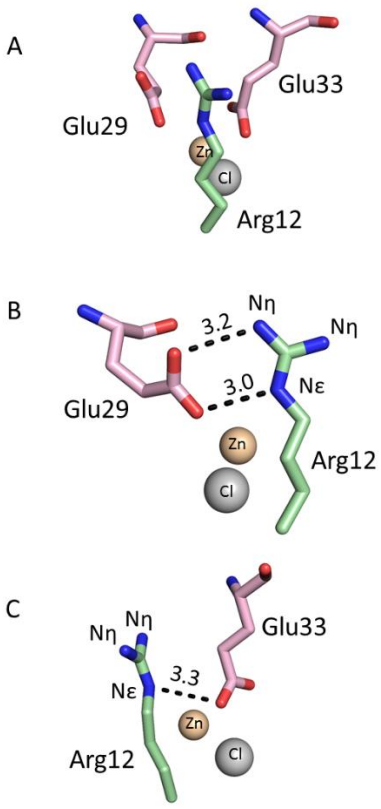

Figure 3.8: "Networked" Salt Bridges between binding Partners A + B. A. "Networked" salt bridges between binding Partner A (glutamic acid at positons 29 and 33) and B (arginine at position 12). B. Electrostatic interactions between glutamic acid (position 29) and arginine. C. An electrostatic interaction between glutamic acid (position 33) and arginine. Selected residues of binding Partners A and B are colored light pink and green, respectively.

Binding Partner A was redesigned to generate Binding Partner C, where leucine residues at positions 16 and 37 were modified to valine. Binding Partner C remains monomeric in the presence and absence of metal ions. The protein-protein interface of binding Partners B + C contains intermolecular salt bridges and a zinc coordination site. Analysis of the crystal structures of binding Partners B + C demonstrates the formation of "networked" salt bridges with a central arginine atom surrounded by two glutamic acid residues. Three hydrogen bonding interactions occur between glutamic acids at positions 29 and 33 and arginine at position 12. While a similar electrostatic interactions were observed for both structures, the distances between Nitrogen and Oxygen are slightly longer for binding Partners B + C. Variations in the electrostatic distances might be due to the crystallization conditions. Both binding Partner A + B and Partner B + C contain an additional zinc (II) and chloride ion near the "networked" salt bridges.

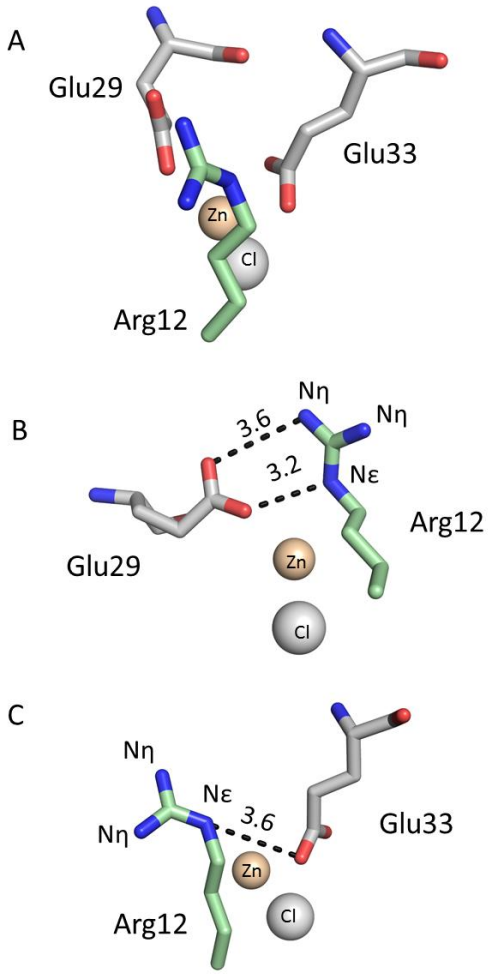

Figure 3.9: "Networked" Salt Bridges between binding Partners B + C. A. "Networked" salt bridges between binding Partner C (glutamic acid at positons 29 and 33) and B (arginine at position 12). B. Electrostatic interactions between glutamic acid (position 29) and arginine. C. An electrostatic interaction between glutamic acid (position 33) and arginine. Selected residues of binding Partners B and C are colored green and grey, respectively.

3.5 Discussion:

3.5.1 Metal-mediated interactions between two different proteins

There are many naturally occurring metal-mediated heterodimer structures in the Protein Data Bank (PDB)[31,32]. In one example, a metal-mediated heterodimer interaction occurs between human growth hormone and prolactin receptor complex (PDB entry 1BP3)[31]. The coordination bonds of the metal-mediated heterodimer are spread over two different subunits, where each subunit contributes two ligands for metal coordination. Upon addition of zinc (II), binding

between the hormone and receptor is enhanced by a factor of 10,000[31]. In the hormone-receptor complex one zinc (II) ion mediates protein assembly, suggesting a single metal coordination site is a viable strategy for engineering metal-mediated interactions between two different proteins.

Another zinc-mediated heterodimer interaction occurs between the N-terminal fragment of non-receptor protein tyrosine kinase (Lck) and the cytosolic C-terminal fragment of the CD4 receptor (PDB entry 1Q68)[32]. The structure of the Lck-CD4-Zn^{2+} complex demonstrated that Lck and CD4 both contribute two ligands for zinc (II) coordination. In addition, biophysical analysis suggests Lck and CD4 form both metal-mediated homo- and hetero-dimeric species[33]. In late 2018, another researcher group reported modifications to the CD4 cytoplasmic tail which enhanced metal-mediated heterodimer specificity[34]. The selective metal-mediated heterodimer might be used for protein heterodimerization assembly and molecular biology routes[34].

There has been limited success in engineering new metal-mediated interactions between two different monomeric protein scaffolds[35]. Wild-type Ubiquitin has an open coordination sphere at histidine 68, which might be used for intermolecular zinc (II) coordination[35]. Using computational protein design, a monomeric protein scaffold was re-engineered with three histidine residues and favorable contact points at the designed protein-protein interface. The metal-mediated interaction involves three engineered histidine residues from the monomeric protein scaffold and histidine 68 from Ubiquitin, representing a 3:1 zinc coordination arrangement model. The metal-mediated interaction between Speltzer, the designed protein scaffold, and Ubiquitin contained a single metal coordination site, which provides fewer energetically favorable interactions between the different proteins. Previously engineered metal-mediated and metal-controlled homodimers contained two metal coordination sites[4,5,6,7,8,11].

In the design of a metal-mediated heterodimer, a single metal coordination and favorable contact points were designed between Ubiquitin and Speltzer. Speltzer binds Ubiquitin with a binding dissociation constant of 20 μM and 68 μM in the presence and absence of zinc, respectively[35]. NMR and mutagenesis suggests the binding arrangement between Spelter and Ubiquitin does not occur via the predicted 3:1 coordination model. The researchers suggested a 2:2 coordination arrangement (i.e., each protein contributes two ligands for metal coordination) might generate a more energetically favorable metal-mediated interaction[33]. A single zinc coordination site, containing of a 2:2 coordination arrangement, was common in naturally occurring metal-mediated heterodimers[31,32].

3.5.2 Our metal-controlled dimers and its variants

In this work, we described the design of a metal-controlled heterodimer using a metal-controlled homodimer, MCD_C1 as the protein design scaffold. Our research goal was to re-engineer a high affinity metal-controlled dimer into a non-promiscuous protein pair. Ultimately, we designed a protein pair that had two unique sites for protein-protein interactions. Binding Partner C contains negatively charged glutamic acid residues at positions 29 and 33, while binding Partner B contains a positively charged arginine residue at position 12. The metal coordination site between binding Partners B and C are formed by the juxtaposition of histidine residues at positions 12, 29 and 33. Heterodimer recognition specificity was generated by having each binding partner contain a portion of the metal binding site and oppositely charged residues. Binding Partner B and C do not interact in the absence of metal ions or form metal-controlled homodimers.

In our metal-controlled heterodimer system, zinc (II) coordination controls the association of binding Partner B and C. The heterodimer structures of binding Partners B + C

and Partners A + B adopt a head-to-head intermolecular orientation similar to MCD_C1[11] (metal-controlled homodimer) and Gβ1-M2[9] (constitutive dimer of Gβ1). Both heterodimer structures revealed metal coordination bonds and "networked" salt bridges at the designed interface. Interactions between binding Partners B + C formed a metal-controlled heterodimer with a binding dissociation constant of ~ 19 μM. Binding Partners B and C represent an orthogonal protein pair since the engineered protein-protein interaction occurs only upon addition of zinc (II).

Table 3.5: Designed Oligomerization states of Gβ1

Terminology	Oligomerization State	Binding Affinity
Gβ1-wild type	monomer	N.A.
Gβ1-M2[a]	dimer	~ 1 μM range
MCD_C1[b]	metal-controlled dimer	< 75 pM
Binding Partners B and C	metal-controlled heterodimer	~ 19 μM

a: Dimer Formation of a Stabilized Gβ1 Variant: A Structural and Energetic Analysis
b: Design of High-Affinity Metal-Controlled Protein Dimers

3.5.3 Potential applications of the metal-controlled dimers and its variants

These designed binding Partners B and C are completely monodisperse in the monomeric state, yet form clean, monodisperse heterodimers upon the addition of zinc (II). The metal-controlled heterodimers are readily converted back to monomers upon the addition of metal-chelating agents such as EDTA. The ability of these proteins to switch between high affinity heterodimers and monodisperse monomers enables them to function as valuable biotechnological tools. For example, one can envision using these proteins to bring attached macromolecules

(such as proteins) together, and also induce them to separate under predefined experimental conditions. We envision using the our generated metal-controlled homodimers[11] and metal-controlled heterodimers as a functional tag for metal-directed protein assembly.

Attachment to the metal-controlled homodimers and heterodimers would likely occur at the level of the gene, where the gene for an attached passenger protein could be expressed in tandem with that of the metal-controlled dimer. Designing fusion proteins containing multiple protein domains remains a successful approach for engineering new protein functions[36]. This strategy involves the attachment of two different gene sequences to generate a fusion gene. Expression of the fusion gene would generate a fusion protein containing the engineered protein with metal binding sites, followed by a protein of interest, with an intervening structural linker.

A fusion protein with a metal-controlled dimer functional tag provides chemical control over the interaction between linked proteins. Depending on the external environment, these metal-controlled dimers could switch between high affinity dimers and monodisperse monomers. The separation distance of the functional proteins can be influenced by the length and flexibility of the structural linker. Metal coordination changes the orientation and distance of the fusion protein, which further enables the close spatial proximity of the proteins containing multiple domains.

The metal-directed assembly of identical fusion proteins remains relatively straightforward for metal-controlled homo- and hetero-dimers. In comparison, the directed assembly of non-identical fusion proteins using the metal-controlled homodimer scaffold would involve intrinsic complications. If two non-identical proteins (X and Y) were fused to the metal-controlled homodimer (MCD), the resulting assemblies would be 25 % MCD-X/MCD-X, 25 % MCD-X/MCD-Y, 25 % MCD-Y/MCD-X, and 25 % MCD-Y/MCD-Y. Operating under the

assumptions of no biases in protein binding and MCD-X/MCD-Y AND MCD-Y/MCD-X are functionally identical, a heterologous MCD-X/MCD-Y assembly would have a maximum yield of 50 %. A more efficient approach would be utilizing the metal-controlled interaction between binding Partners B and C, which may generate a metal-directed assembly of non-identical proteins with a yield closer to 100 %.

Metal-mediated and metal-controlled heterodimers would significantly improve the utility of metal-directed protein assembly. Interactions between engineered heterodimers have been utilized in applications extending to affinity purification[37], directed assembly[38], and biomaterials[39]. Metal-controlled protein dimerization systems provides a mechanism to regulate the physical distance between proteins. In addition to the previously mentioned applications, a metal-controlled interaction between two different proteins enables chemical control over assembly and disassembly[8] and may contain properties of metalloenzymes[40]. A selective metal-controlled heterodimer might be used for protein heterodimerization assembly and manipulation of cellular processes.

Metal-controlled protein assembly has been the subject of ongoing research to expand the utility of chemically-induced dimerization systems. These methods have generated high affinity and orientation specific metal-mediated and metal-controlled homodimers[7,11]. In our work, we generated a metal-controlled heterodimer from a metal-controlled homodimer, where heterodimerization assembly is controlled by zinc (II) coordination. A metal-controlled heterodimer assembly provides an engineering scaffold for directed and controllable assembly of different proteins. The interplay between metals and protein scaffolds might allow for new connections in biotechnology, synthetic biology, biomaterials, and supramolecular chemistry.

3.6 Supporting Information

3.6.1 Amino Acid Sequences of Gβ1 and its variants

Table 3.6: Amino Acid Sequences for Metal-Controlled Dimer and Binding Partners

	Mutations	Sequence
Gβ1	wild-type	MTYKLILNGKTLKGETTTEAVDAATAE KVFKQYANDNGVDGEWTYDDATKTFT VTE
Metal-Controlled Dimer	L12H, T16L, V29H, Y33H, N37L	MTYKLILNGKT**H**KGE**L**TTEAVDAATAE K**H**FKQ**H**AND**L**GVDGEWTYDDATKTFT VTE
Binding Partner A	L12H, T16L, V29E, Y33E, N37V	MTYKLILNGKT**H**KGE**L**TTEAVDAATAE K**E**FKQ**E**AND**L**GVDGEWTYDDATKTFT VTE
Binding Partner B	L12R, T16L, V29H, Y33H, N37L	MTYKLILNGKT**R**KGE**L**TTEAVDAATAE K**H**FKQ**H**AND**L**GVDGEWTYDDATKTFT VTE
Binding Partner C	L12H, T16V, V29E, Y33E, N37V	MTYKLILNGKT**H**KGE**V**TTEAVDAATAE K**E**FKQ**E**AND**V**GVDGEWTYDDATKTFT VTE

The mutations, relative to wild-type Gβ1, are bolded.

3.6.2 SEC-MALS Analysis of Binding Partners A, B, and C

Size exclusion chromatography with multi-angle light scattering (SEC-MALS) used to analyze Binding Partners A, B, and C. The individual samples were analyzed in the presence or absence of 1 mM zinc buffer.

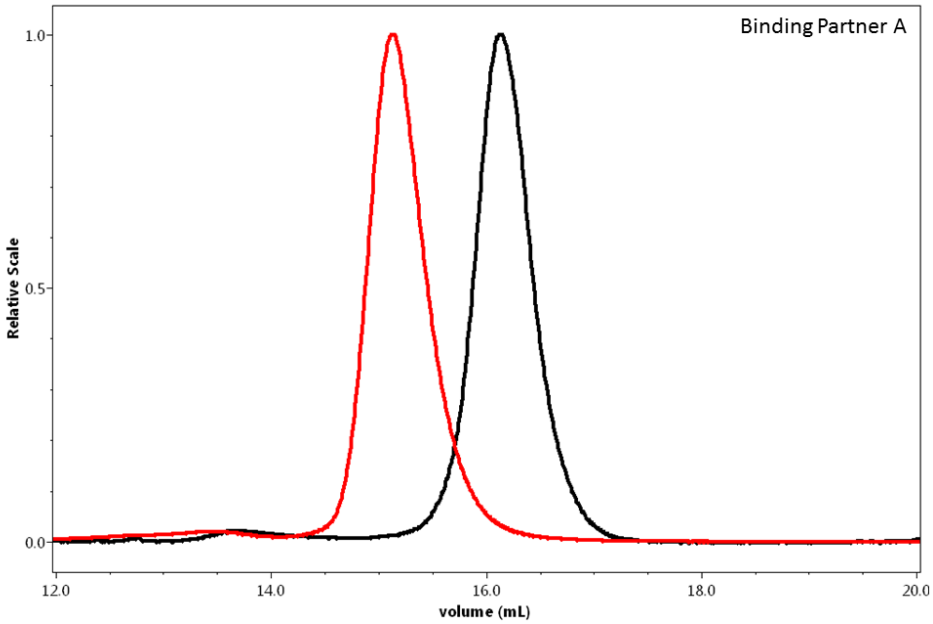

Figure 3.10: SEC-MALS Characterization of binding Partner A. Multi-angle light scattering was collected in the absence of metal (black) or presence of 1 mM Zinc sulfate (red). The molecular weight for binding Partner A was 6.8 and 12.2 kDa, respectively.

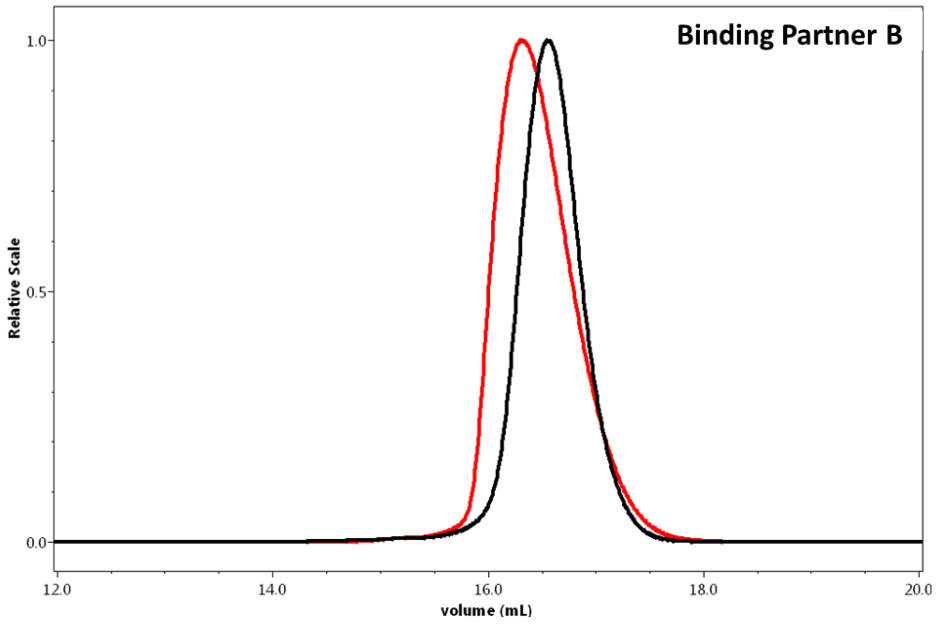

Figure 3.11: SEC-MALS Characterization of binding Partner B. Multi-angle light scattering was collected in the absence of metal (black) or presence of 1 mM Zinc sulfate (red). The molecular weight for binding Partner B was 7.1 and 7.5 kDa, respectively.

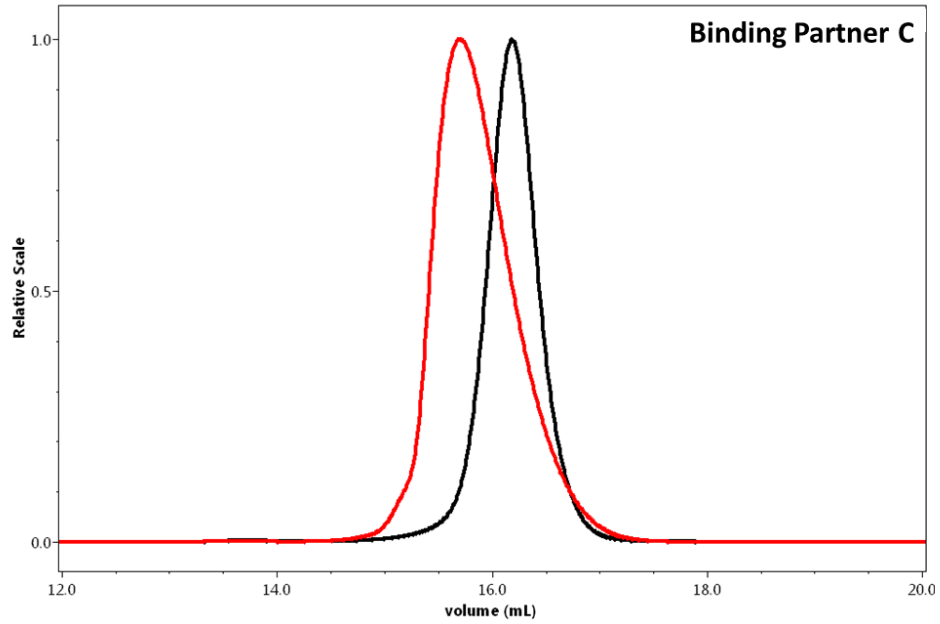

Figure 3.12: SEC-MALS Characterization of binding Partner C. Multi-angle light scattering was collected in the absence of metal (black) or presence of 1 mM Zinc sulfate (red). The molecular weight for binding Partner C was 6.9 and 8.3 kDa, respectively.

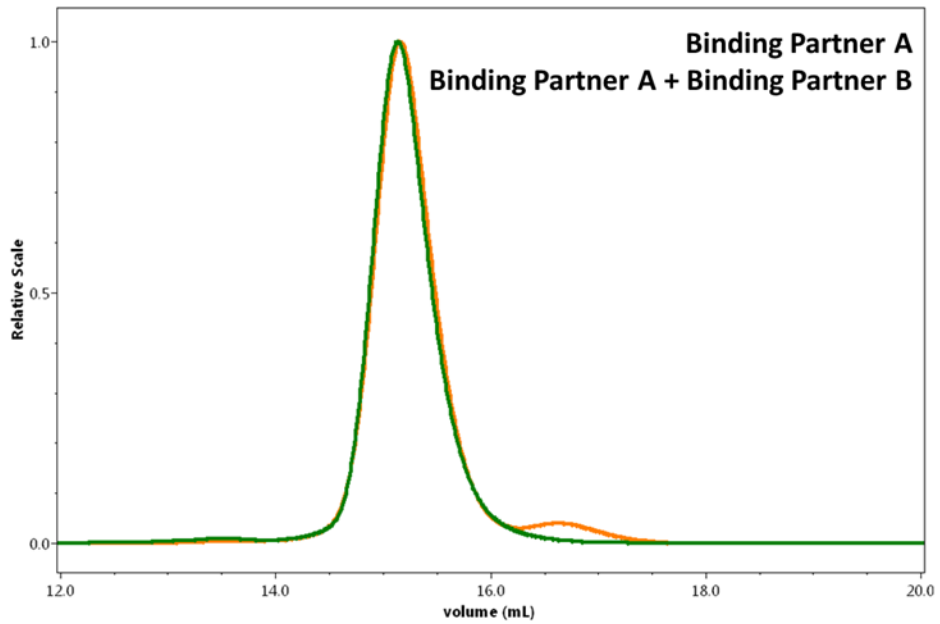

Figure 3.13: Metal-mediated interactions between binding Partner A and Partners A + B. Binding Partner A (green) and binding Partner A/binding Partner B (orange) in the presence of 1 mM zinc buffer. The calculated molecular weight of binding Partner A in 1 mM zinc buffer was 12.2 kDa. The calculated molecular weight of binding Partner A and B was 13.4 kDa.

3.6.3 SDS-PAGE analysis of Binding Partners

The purity of the Binding Partners was assessed using SDS-PAGE. An 18% SDS_PAGE gel was used to analyze the individual components. The migration of the individual proteins do not correspond to the theoretical molecular weight of around 6.5 kDa.

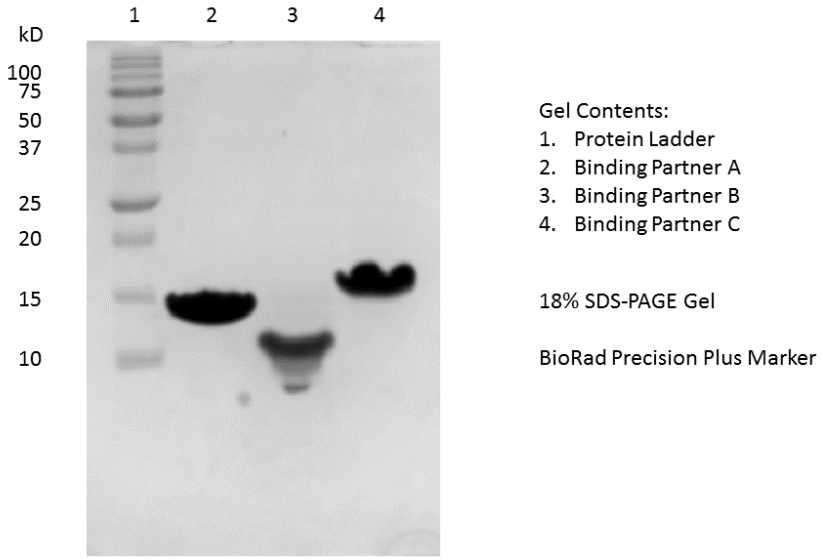

Figure 3.14: SDS-PAGE Gel of Binding Partners A, B, and C

3.6.4 Thermal Denaturation of all Variants Monitored with Circular Dichroism

The thermal denaturation melting temperatures for the Binding Partners were measured using circular dichroism. The increase in the CD signal at 222 nm was monitored as a function of increasing temperature under different experimental conditions. Normalized curves with both the folded and unfolded states for the Binding Partners are shown below.

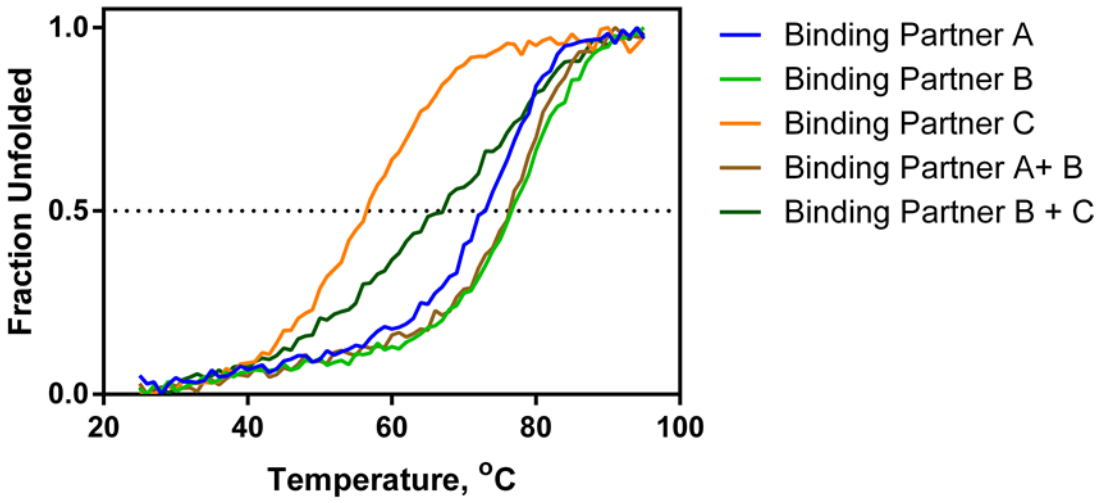

Figure 3.15: Thermal Denaturation of Binding Partner A, B, and C

3.6.5 Analytical Ultracentrifugation: Sedimentation Velocity of Binding Partners B and C

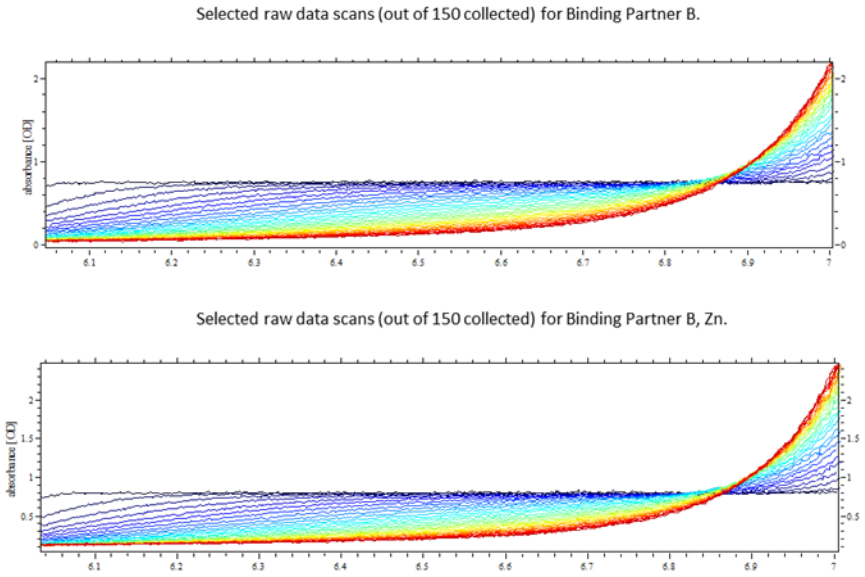

Figure 3.16: Selected Raw Sedimentation Velocity Data for Binding Partner B

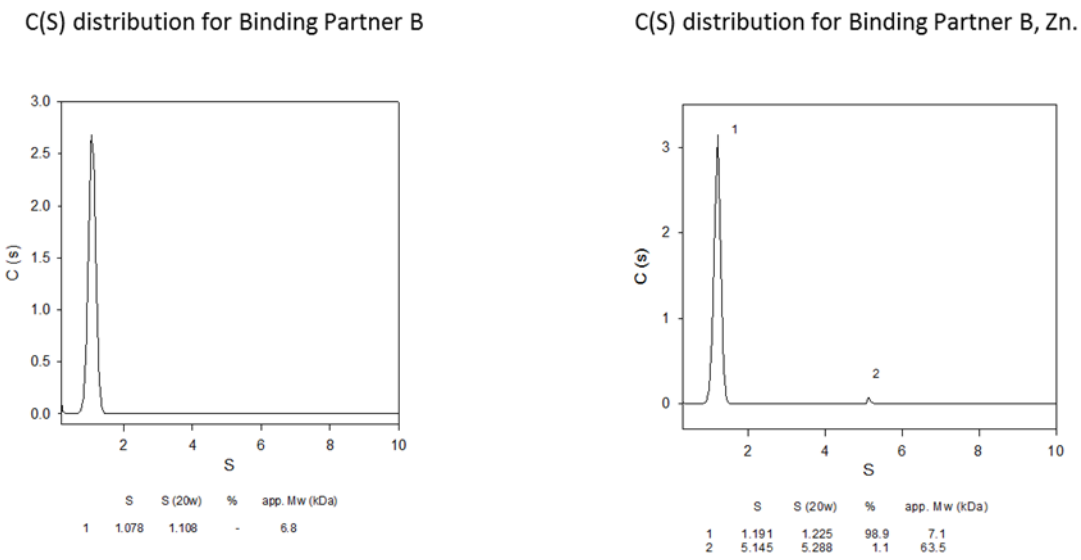

Figure 3.17: C(S) Distribution of Binding Partner B

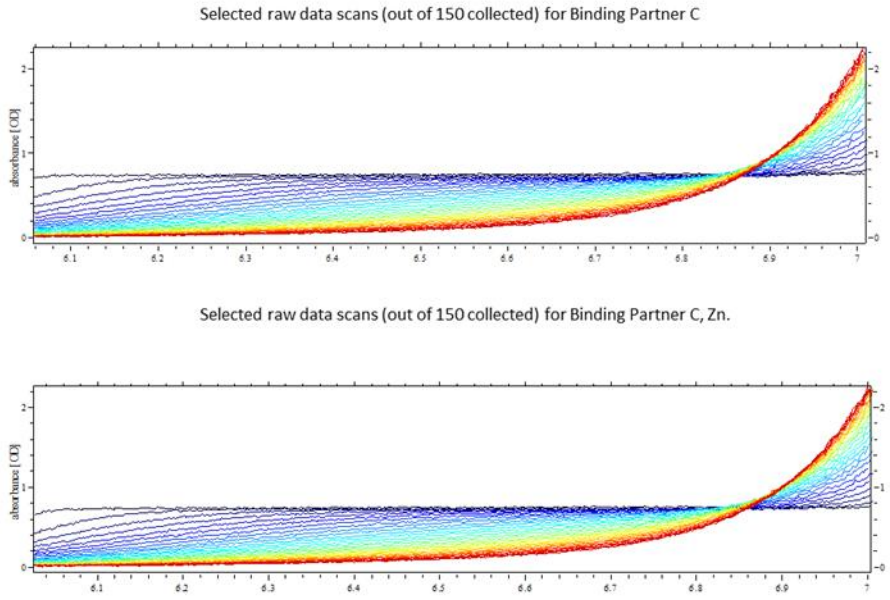

Figure 3.18: Selected Raw Sedimentation Velocity Data for Binding Partner C

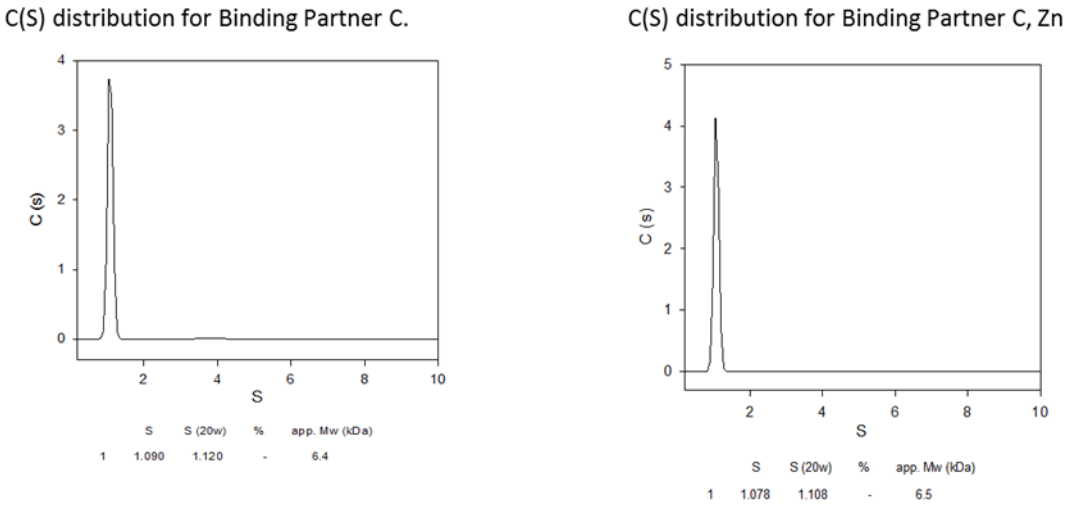

Figure 3.19: C(S) Distribution of Binding Partner C

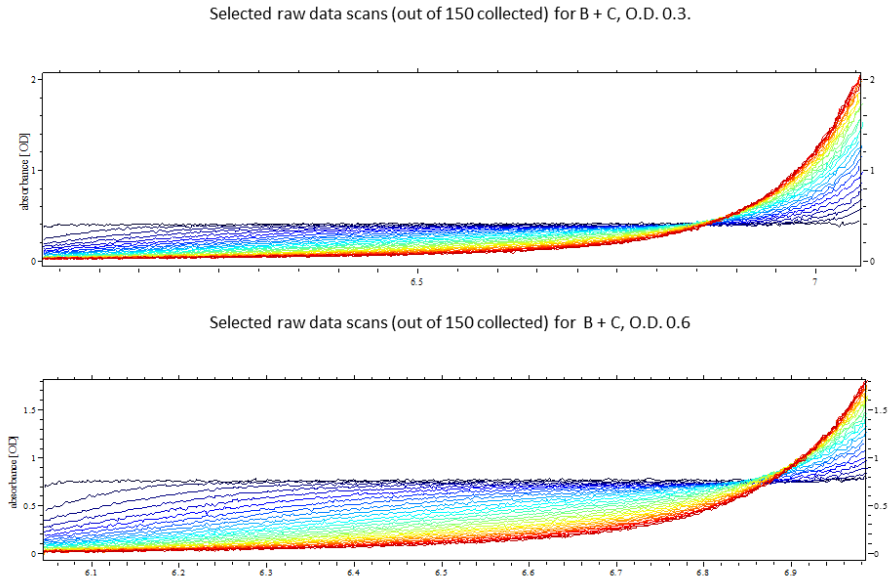

Figure 3.20: Selected Raw Sedimentation Velocity Data for Binding Partner B + C

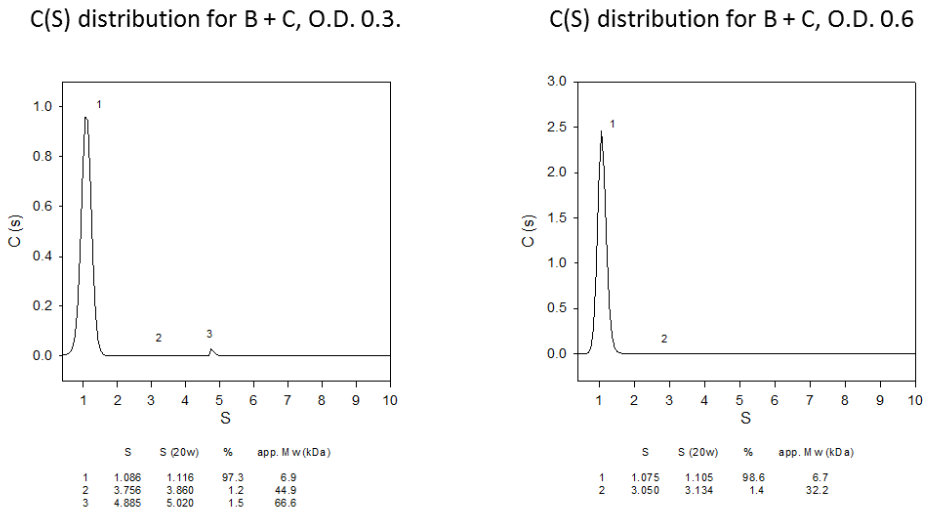

Figure 3.21: C(S) Distribution of Binding Partner B + C

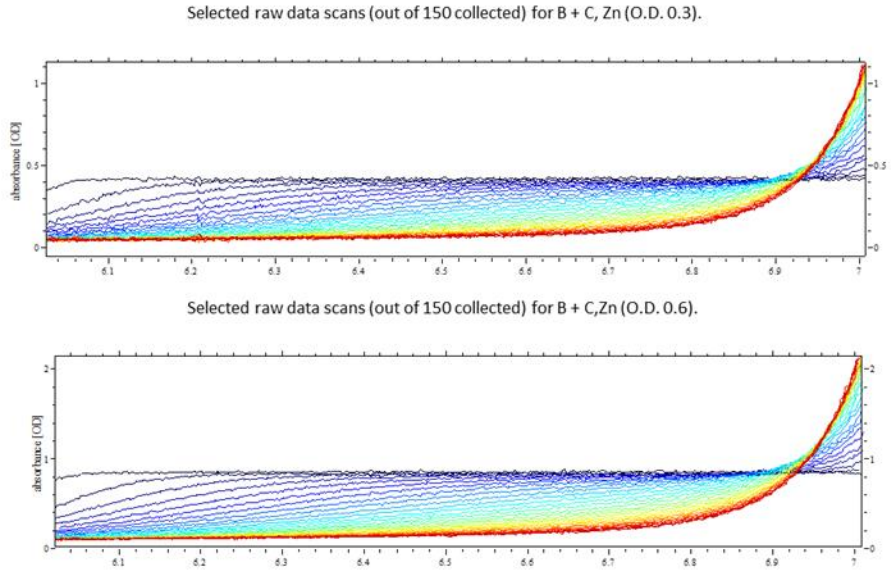

Figure 3.22: Selected Raw Sedimentation Velocity Data for Binding Partner B + C, zinc

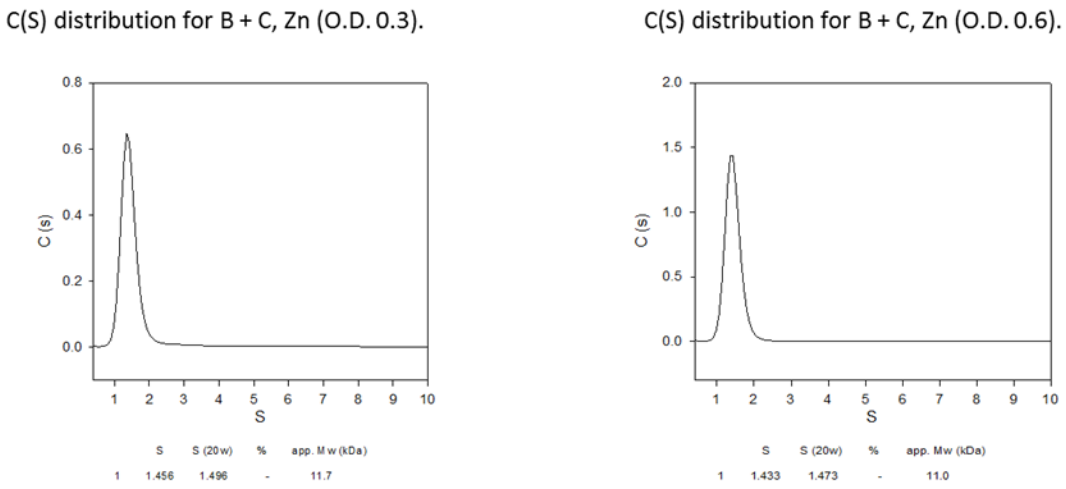

Figure 3.23: C(S) Distribution of Binding Partner B + C, zinc

3.6.6 Analytical Ultracentrifugation: Sedimentation Equilibrium of Binding Partner B and C

```
HeteroAnalysis - Version 1.1.58
12/3/2018, 3:07:59 PM

The current directory is C:\My_files\Analysis\xlidata\112818\145457

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA4
Dataset #1:: 6_4A30278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA4
Dataset #2:: 6_4B30278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA4
Dataset #3:: 6_4C30278.dat


=======================================
Model :: Ideal
12/3/2018, 3:09:55 PM

      File              Baseline            Ref.Conc.          Ref.Dist.  Ext.Coeff.
 1: 6_4A30278.dat      -3.39E-02±1.24E-02   7.01E-05±1.50E-06   5.9427    8.480E+03
 2: 6_4B30278.dat      -6.87E-03±5.99E-03   3.33E-05±7.39E-07   6.4470    8.480E+03
 3: 6_4C30278.dat      -2.80E-03±3.14E-03   1.55E-05±3.88E-07   6.9317    8.480E+03

Buoyant MW = 1806 ± 42 Da
MW = 6718 using the current values of vbar = 0.7290 and density = 1.00290
Sigma at Max. Speed of 30000 RPM = 0.732 cm-2

Number of Iterations: 5
Number of Points: 302
Degrees of Freedom: 295
RMS Deviation: 0.00538
Number of Parameters Held Constant: 0

Variance Ratio at the 95.46% confidence level:
F(NDF,NDF,1-P) = 1.2181
```

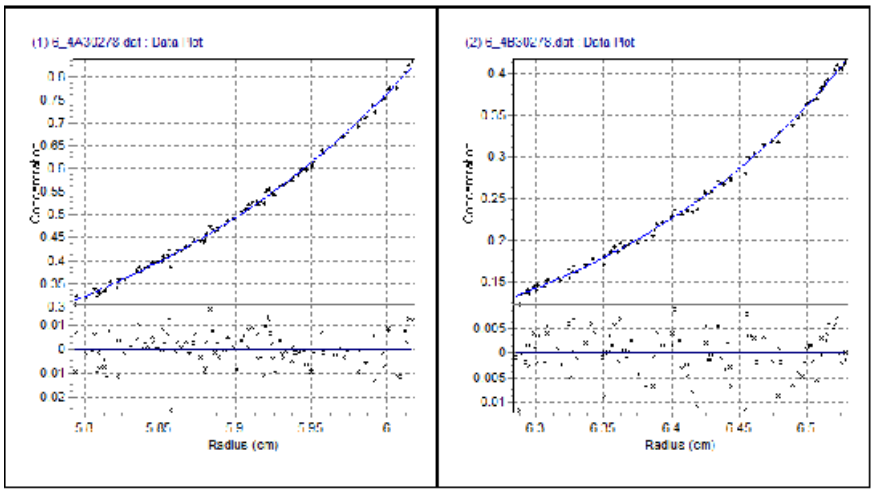

Figure 3.24: Sedimentation Equilibrium Data for Binding Partner B, zinc at 30,000 rpm

```
HeteroAnalysis - Version 1.1.58
12/3/2018, 3:36:39 PM

The current directory is C:\My_files\Analysis\xlidata\112818\145457

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA4
Dataset #1:: 12_4A35278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA4
Dataset #2:: 12_4B35278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA4
Dataset #3:: 12_4C35278.dat


========================================
Model :: Ideal
12/3/2018, 3:39:06 PM

       File              Baseline            Ref.Conc.          Ref.Dist.  Ext.Coeff.
 1: 12_4A35278.dat      -1.07E-02±6.26E-03   6.74E-05±7.91E-07   5.9503    8.480E+03
 2: 12_4B35278.dat       4.21E-03±3.07E-03   3.14E-05±4.00E-07   6.4523    8.480E+03
 3: 12_4C35278.dat       1.72E-04±1.86E-03   1.37E-05±2.35E-07   6.9280    8.480E+03

Buoyant MW = 1881 ± 24 Da
MW = 6997 using the current values of vbar = 0.7290 and density = 1.00290
Sigma at Max. Speed of 35000 RPM = 1.038 cm-2

Number of Iterations: 14
Number of Points: 315
Degrees of Freedom: 308
RMS Deviation: 0.00558
Number of Parameters Held Constant: 0

Variance Ratio at the 95.46% confidence level:
F(NDF,NDF,1-P) = 1.2130
```

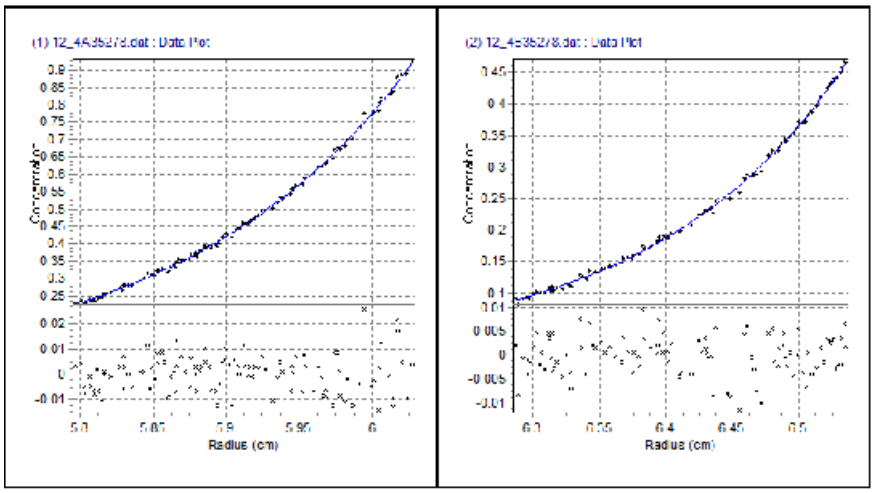

Figure 3.25: Sedimentation Equilibrium Data for Binding Partner B, zinc at 35,000 rpm

```
HeteroAnalysis - Version 1.1.58
12/3/2018, 2:55:41 PM

The current directory is C:\My_files\Analysis\xlidata\112818\145457

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA2
Dataset #1:: 6_2A30278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA2
Dataset #2:: 6_2B30278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA2
Dataset #3:: 6_2C30278.dat


========================================
Model :: Ideal
12/3/2018, 2:57:45 PM

      File              Baseline            Ref.Conc.           Ref.Dist.  Ext.Coeff.
 1: 6_2A30278.dat       2.94E-02±8.18E-03   7.27E-05±1.02E-06    5.9610    8.480E+03
 2: 6_2B30278.dat       3.39E-02±3.64E-03   3.07E-05±4.63E-07    6.4500    8.480E+03
 3: 6_2C30278.dat       1.30E-02±2.23E-03   1.36E-05±2.77E-07    6.9253    8.480E+03

Buoyant MW = 2113 ± 33 Da
MW = 7772 using the current values of vbar = 0.7260 and density = 1.00290
Sigma at Max. Speed of 30000 RPM = 0.856 cm-2

Number of Iterations: 16
Number of Points: 349
Degrees of Freedom: 342
RMS Deviation: 0.00603
Number of Parameters Held Constant: 0

Variance Ratio at the 95.46% confidence level:
F(NDF,NDF,1-P) = 1.2011
```

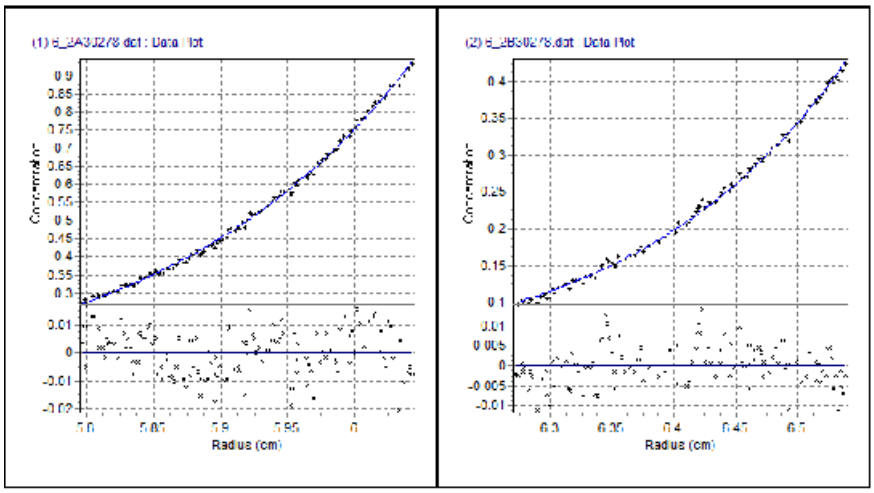

Figure 3.26: Sedimentation Equilibrium Data for Binding Partner C, zinc at 30,000 rpm

```
HeteroAnalysis - Version 1.1.58
12/3/2018, 3:02:47 PM

The current directory is C:\My_files\Analysis\xlidata\112818\145457

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA2
Dataset #1:: 12_2A35278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA2
Dataset #2:: 12_2B35278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA2
Dataset #3:: 12_2C35278.dat


=======================================
Model :: Ideal
12/3/2018, 3:04:03 PM

       File                Baseline             Ref.Conc.           Ref.Dist.  Ext.Coeff.
 1: 12_2A35278.dat       -7.21E-03±4.93E-03    7.19E-05±6.38E-07   5.9450     8.480E+03
 2: 12_2B35278.dat        1.14E-02±2.71E-03    3.56E-05±3.48E-07   6.4603     8.480E+03
 3: 12_2C35278.dat        3.33E-03±1.67E-03    1.33E-05±2.07E-07   6.9193     8.480E+03

Buoyant MW = 1904 ± 18 Da
MW = 7003 using the current values of vbar = 0.7260 and density = 1.00290
Sigma at Max. Speed of 35000 RPM = 1.049 cm-2

Number of Iterations: 16
Number of Points: 343
Degrees of Freedom: 336
RMS Deviation: 0.00564
Number of Parameters Held Constant: 0

Variance Ratio at the 95.46% confidence level:
F(NDF,NDF,1-P) = 1.2030
```

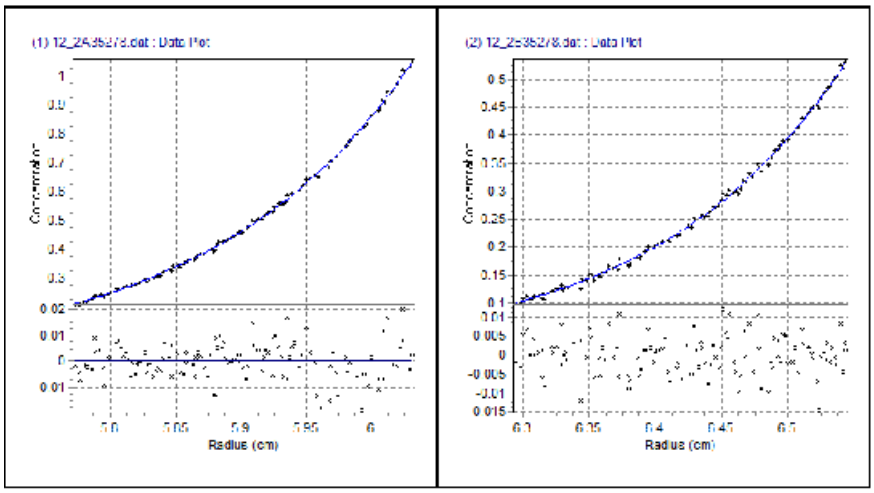

Figure 3.27: Sedimentation Equilibrium Data for Binding Partner C, zinc at 35,000 rpm

```
HeteroAnalysis - Version 1.1.58
12/3/2018, 2:37:13 PM

The current directory is C:\My_files\Analysis\xlidata\110218


The current directory is C:\My_files\Analysis\xlidata\112818\145457
Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA6
Dataset #1:: 6_6A30278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA6
Dataset #2:: 6_6B30278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00006.RA6
Dataset #3:: 6_6C30278.dat


=======================================
Model :: Ideal
12/3/2018, 2:41:17 PM

      File                Baseline            Ref.Conc.           Ref.Dist.  Ext.Coeff.
 1: 6_6A30278.dat         7.92E-03±4.36E-03   6.49E-05±5.85E-07   5.9437     8.480E+03
 2: 6_6B30278.dat         2.36E-02±2.17E-03   2.83E-05±2.87E-07   6.4437     8.480E+03
 3: 6_6C30278.dat         1.57E-02±1.45E-03   1.49E-05±1.86E-07   6.9530     8.480E+03

Buoyant MW = 2973 ± 28 Da
MW = 10994 using the current values of vbar = 0.7275 and density = 1.00290
Sigma at Max. Speed of 30000 RPM = 1.203 cm-2

Number of Iterations: 16
Number of Points: 350
Degrees of Freedom: 343
RMS Deviation: 0.00634
Number of Parameters Held Constant: 0

Variance Ratio at the 95.46% confidence level:
F(NDF,NDF,1-P) = 1.2007
```

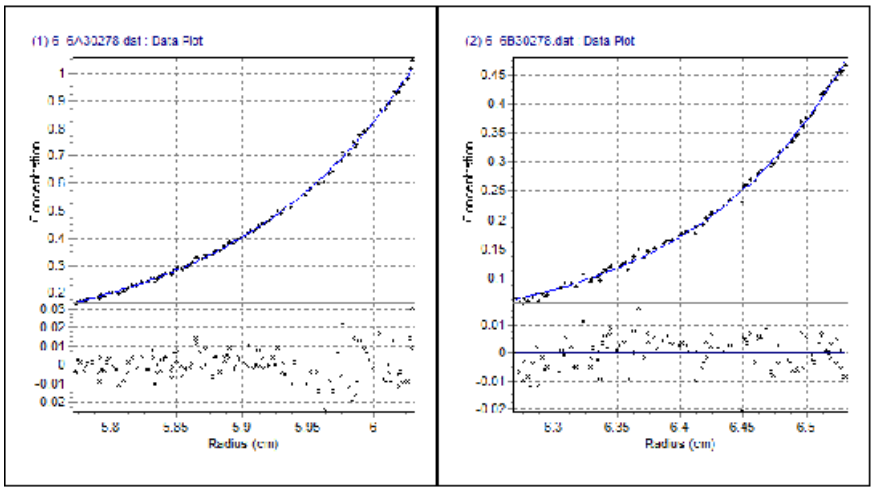

Figure 3.28: Sedimentation Equilibrium Data for Binding Partner B + C, zinc at 30,000 rpm

```
HeteroAnalysis - Version 1.1.58
12/3/2018, 2:48:25 PM

The current directory is C:\My_files\Analysis\xlidata\112818\145457

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA6
Dataset #1:: 12_6A35278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA6
Dataset #2:: 12_6B35278.dat

Source:: C:\My_files\Analysis\xlidata\112818\145457\00012.RA6
Dataset #3:: 12_6C35278.dat


========================================
Model :: Ideal
12/3/2018, 2:49:49 PM

       File              Baseline            Ref.Conc.           Ref.Dist.  Ext.Coeff.
 1: 12_6A35278.dat       1.42E-02±2.73E-03   6.43E-05±3.96E-07    5.9523    8.480E+03
 2: 12_6B35278.dat       2.60E-02±1.62E-03   3.10E-05±2.18E-07    6.4643    8.480E+03
 3: 12_6C35278.dat       1.75E-02±1.26E-03   1.14E-05±1.50E-07    6.9403    8.480E+03

Buoyant MW = 3000 ± 17 Da
MW = 11094 using the current values of vbar = 0.7275 and density = 1.00290
Sigma at Max. Speed of 35000 RPM = 1.654 cm-2

Number of Iterations: 16
Number of Points: 367
Degrees of Freedom: 360
RMS Deviation: 0.00748
Number of Parameters Held Constant: 0

Variance Ratio at the 95.46% confidence level:
F(NDF,NDF,1-P) = 1.1955
```

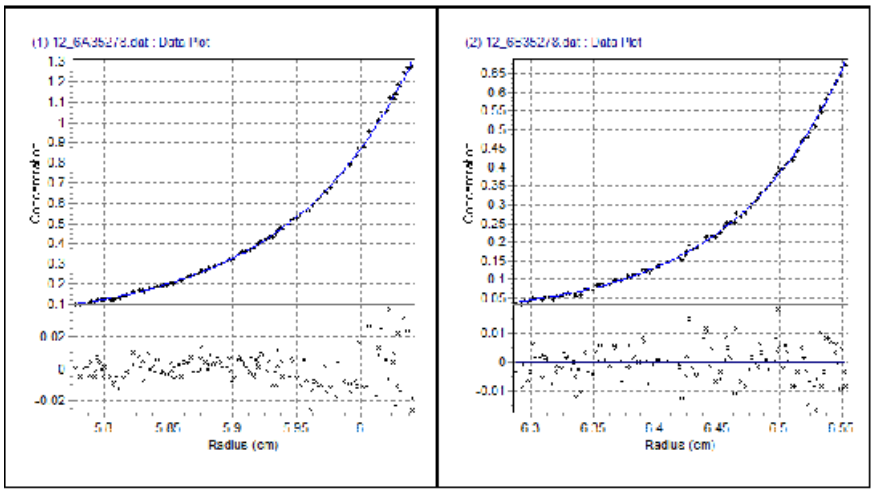

Figure 3.29: Sedimentation Equilibrium Data for Binding Partner B + C, zinc at 35,000 rpm

3.6.7 X-Ray Crystal Structures and Associated Omit Maps of Binding Partners A + B and B + C

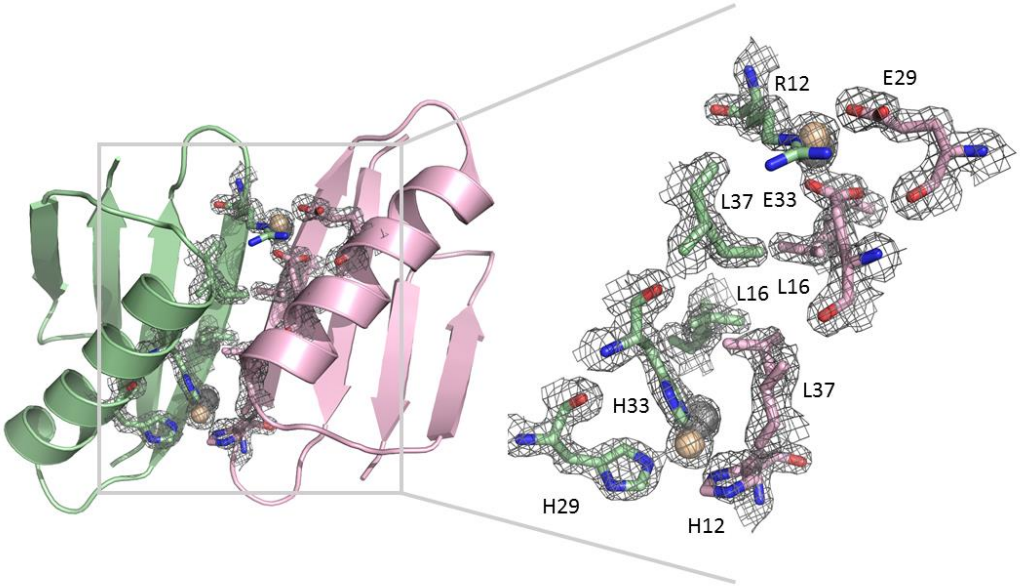

Figure 3.30: X-Ray Crystal Structure and Omit Maps of Binding Partners A + B. The 2Fo – Fc electron density maps (contour level – 1.5 σ) of binding Partner B (light green color) and binding Partner A (light pink color) are shown on the right for the metal-histidine residues (His 12, 29, 33) and also the hydrophobic residues, Leu 16 and Leu 37, that contribute to complex formation. Side-chains are illustrated as sticks, zinc atoms as amber spheres, and chloride ions as light grey spheres.

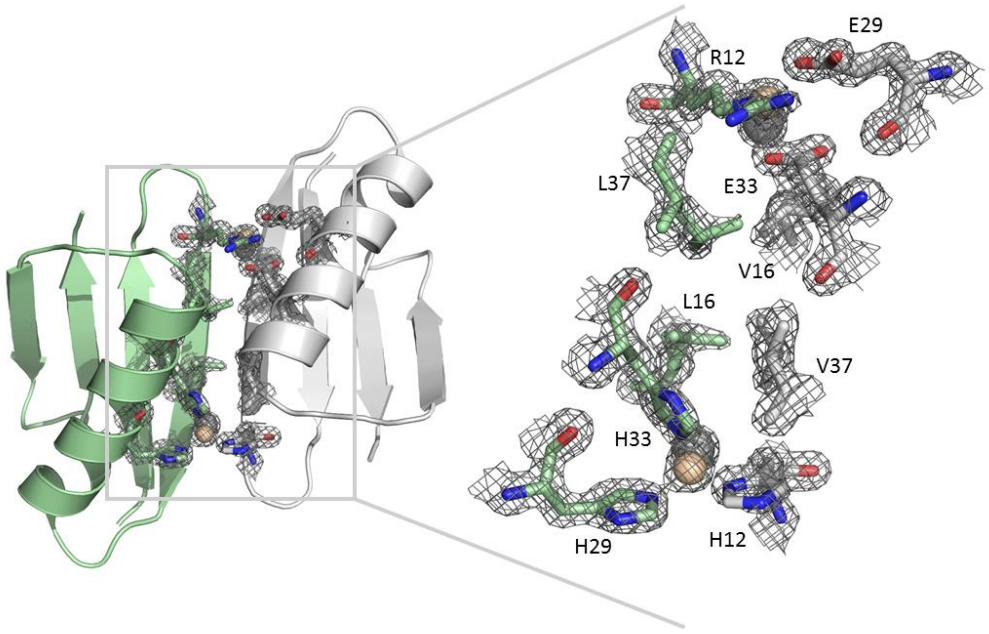

Figure 3.31: X-Ray Crystal Structure and Omit Maps of Binding Partners B + C. The 2Fo – Fc electron density maps (contour level – 1.5 σ) of binding Partner B (light green color) and binding Partner C (light grey color) are shown on the right for the metal-histidine residues (His 12, 29, 33) and also the hydrophobic residues, Leu 16 and Leu 37 (binding Partner B) and Val 16 and Val 37 (binding Partner C) that contribute to complex formation. Side-chains are illustrated as sticks, zinc atoms as amber spheres, and chloride ions as light grey spheres.

3.6.8 Physical Characteristics of the Zinc Tetrahedral Coordination Site

Zinc binding sites found in protein structures show that standard zinc nitrogen (histidine) distances in zinc metalloproteins is 2.09 ± 0.14 Å[41]. The crystal structures for binding Partners A + B and binding Partners B + C each contains four molecules in the asymmetric unit. The biological assembly of binding Partners A + B and B + C is a dimer, in the presence of zinc (II).

Table 3.7: Distances between each Zinc Atom and the Side-Chain Nitrogen Atom

	Binding Partners A + B		**Binding Partners B + C**	
Residue	Chains A and D	Chains B and C	Chains A and B	Chains B and C
H12	2.17	2.15	2.08	2.19
H29	1.99	2.03	2.05	1.98
H33	1.82	1.83	1.86	1.85

Table 3.8: Angles at Zinc Atoms between pairs of Metal-Ligand Histidine Residues

	Binding Partners A + B		**Binding Partners B + C**	
Residue	Chains A and D	Chains B and C	Chains A and B	Chains B and C
H12 and H29	104.5°	100.1°	97.3°	102.4°
H12 and H33	97.4°	96.9°	90.9°	88.8°
H29 and H33	103.9°	99.6°	98.3°	102.0°

3.6.9 X-ray data collection parameters and structure refinement statistics

Table 3.9: X-ray data collection parameters and structure refinement statistics

Data collection	Binding Partner A + B	Binding Partner B + C
Space Group	P1	P1
Unit Cell (Å)	25.68 86.33 50.82 75.36 50.82 75.36	25.73 86.35 51.04 75.40 51.04 75.40
Resolution (Å)	1.5	1.4
Observations	123665	132390
Source	ALS 5.0.1	ALS 5.0.1
Wavelength (Å)	0.977410	0.977410
<I/σ(I)>	23.2(8.8)	31.8(11.3)
Completeness	95.11 (93.58)	93.46% (88.69%)
Structure solution by MR		
Molecular replacement model	1pga	1pga
Z-score	22.3	78.7
LLG score	842.241	8093.472
Refinement		
Resolution range (Å)	34.77 - 1.5 (1.554 - 1.5)	34.930 - 1.400 (1.450 - 1.4)
Unique reflections	36666 (3612)	44838 (4288)
Protein residues	224	224
Ligand atoms	16	16
Water Atoms	370	356
R-work	0.1690 (0.2353)	0.1799 (0.2583)
R-free	0.1895 (0.2332)	0.2078 (0.2716)
Geometry		
RMS (bonds)	0.003	0.005
RMS (angles)	0.58	0.71
Clash Score	1.14	2.01
Ramachandran plot		
Favored (%)	98.15	98.15
Allowed (%)	1.85	1.85
MolProbity score	0.91	0.97

3.7 References

(1) Chen, Z.; Boyken, S. E.; Jia, M.; Busch, F.; Flores-solis, D.; Bick, M. J.; Lu, P.; Vanaernum, Z. L.; Sahasrabuddhe, A.; Langan, R. A.; Bermeo, S.; Brunette, T. J.; Dimaio, F.; Sgourakis, N. G.; Wysocki, V. H.; Baker, D.; Mulligan, V. K.; Lauren, P. *Nature* **2019**. 565 (7737) 106-111

(2) Leaver-Fay, A.; Froning, K. J.; Atwell, S.; Aldaz, H.; Pustilnik, A.; Lu, F.; Huang, F.; Yuan, R.; Hassanali, S.; Chamberlain, A. K.; Fitchett, J. R.; Demarest, S. J.; Kuhlman, B. *Structure* **2016**, *24* (4), 641–651.

(3) Kortemme, T.; Joachimiak, L. A.; Bullock, A. N.; Schuler, A. D.; Stoddard, B. L.; Baker, D. *Nat. Struct. Mol. Biol.* **2004**, *11* (4), 371–379.

(4) Salgado, E. N.; Ambroggio, X. I.; Brodin, J. D.; Lewis, R. A.; Kuhlman, B.; Tezcan, F. A. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107* (5), 1827–1832.

(5) Radford, R. J.; Nguyen, P. C.; Ditri, T. B.; Figueroa, J. S.; Tezcan, F. A. *Inorg. Chem.* **2010**, *49* (9), 4362–4369.

(6) Salgado, E. N.; Faraone-mennella, J.; Tezcan, F. A. *J. Am. Chem. Soc.* **2007**, *129*, 13374–13375.

(7) Der, B. S.; MacHius, M.; Miley, M. J.; Mills, J. L.; Szyperski, T.; Kuhlman, B. *J. Am. Chem. Soc.* **2012**, *134* (1), 375–385.

(8) Huard, D. J. E.; Kane, K. M.; Akif Tezcan, F. *Nat. Chem. Biol.* **2013**, *9* (3), 169–176.

(9) Thoms, S.; Max, K. E. A.; Wunderlich, M.; Jacso, T.; Lilie, H.; Reif, B.; Heinemann, U.; Schmid, F. X. *J. Mol. Biol.* **2009**, *391* (5), 918–932.

(10) Huang, P.-S.; Love, J. J.; Mayo, S. L. *Protein Sci.* **2007**, *16* (12), 2770–2774.

(11) Maniaci, B.; Lipper, C. H.; Anipindi, L. D.; Erlandsen, H.; Cole, J. L.; Stec, B.; Huxford, T.; Love, J. J. *Biochemistry* **2019**, *58* (17), 2199–2207.

(12) Goodsell, D. S.; Olson, A. J. *Annu Rev Biophys Biomol Struct*. **2000**. 29 (1) 105-153

(13) André, I.; Strauss, C. E. M.; Kaplan, D. B.; Bradley, P.; Baker, D. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105* (42), 16148–16152.

(14) Lukatsky, D. B.; Shakhnovich, B. E.; Mintseris, J.; Shakhnovich, E. I. *J. Mol. Biol.* **2007**, *365* (5), 1596–1606.

(15) Atwell, S.; Ridgeway, J. B. B.; Wells, J. A.; Carter, P. *J. Mol. Biol.* **1997**, *270*, 26–35.

(16) Gunasekaran, K.; Pentony, M.; Shen, M.; Garrett, L.; Forte, C.; Woodward, A.; Ng, S. Bin; Born, T.; Retter, M.; Manchulenko, K.; Sweet, H.; Foltz, I. N.; Wittekind, M.; Yan, W. *J. Biol. Chem.* **2010**, *285* (25), 19637–19646.

(17) Ko, T.; Liao, C.; Ku, W.; Chak, K.; Yuan, H. S. *Structure* **1999**, No. 7, 91–102.

(18) Donald, J. E.; Kulp, D. W.; Degrado, W. F. *Proteins* **2012**, *79* (3), 898–915.

(19) Musafia, B.; Buchner, V.; Arad, D. *J. Mol. Biol.* **1995**, *254* (4), 761–770.

(20) Kabsch, W. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 133–144.

(21) Musafia, B.; Buchner, V.; Arad, D. *J. Mol. Biol.* **1995**, *254* (4), 761–770.

(22) Adams, P. D.; Afonine, P. V.; Bunkóczi, G.; Chen, V. B.; Davis, I. W.; Echols, N.; Headd, J. J.; Hung, L. W.; Kapral, G. J.; Grosse-Kunstleve, R. W.; McCoy, A. J.; Moriarty, N. W.; Oeffner, R.; Read, R. J.; Richardson, D. C.; Richardson, J. S.; Terwilliger, T. C.; Zwart, P. H. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (2), 213–221.

(23) Emsley, P.; Cowtan, K. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2004**, *60* (12 I), 2126–2132.

(24) Krissinel, E.; Henrick, K. *J. Mol. Biol.* **2007**, *372* (3), 774–797.

(25) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2010**, *66* (1), 12–21.

(26) Zheng, H.; Cooper, D. R.; Porebski, P. J.; Shabalin, I. G.; Handing, K. B.; Minor, W. *Acta Crystallogr. Sect. D Struct. Biol.* **2017**, *73*, 223–233.

(27) Schuck, P. *Biophys. J.* **2000**, *78* (3), 1606–1619.

(28) Laue, T.; Shah, B.; Ridgeway, T.; Pelletier, S. *R. Soc. Chem.* S. Harding, A. Rowe, and J Horton, eds (Cambridge, UK: The Royal Society of Chemistry). **1992**.

(29) Janin, J.; Miller, S.; Chothia, C. *J. Mol. Biol.* **1988**, *204* (1), 155–164.

(30) Cole, J. L.; Lary, J. W.; Moody, T.; Laue, T. M. *Methods Cell Biol.* **2008**, *84* (07), 143–179.

(31) Somers, W.; Ultsch, M.; De Vos, A. M.; Kossiakoff, A. A. *Nature*. 1994, pp 478–481.

(32) Kim, P. W.; J. Sun, Z-Y; Blacklow, S. C.; Wagner, G.; Eck M.J. *Science.* **2003**, *301* (5640), 1725–1728.

(33) Davis, A. M.; Berg, J. M. *J. Am. Chem. Soc.* **2009**, *131* (32), 11492–11497.

(34) Kocyła, A.; Krężel, A. *Chem. Commun.* **2018**, 13539–13542.

(35) Der, Bryan S., Jha, Ramesh K., Lewis, Steven M., Thompson,, Peter M., Guntas, Gurkan., Khulman, B. *Proteins* **2013**, *81* (7), 1245–1255.

(36) Van Rosmalen, M.; Krom, M.; Merkx, M. *Biochemistry* **2017**, *56* (50), 6565–6574.

(37) Fernandez-Rodriguez, J.; Marlovits, T. C. *Protein Sci.* **2012**, *21* (4), 511–519.

(38) Wu, Z.; Goldstein, B.; Laue, T. M.; Liparoto, S. F.; Nemeth, M. J.; Ciardelli, T. L. *Protein Sci.* **1999**, 8 (3) 482–489.

(39) Woolfson, D. N. *Biopolymers* **2009**, *94* (1), 118–127.

(40) Der, B. S.; Edwards, D. R.; Kuhlman, B. *Biochemistry* **2012**, *51* (18), 3933–3940.

(41) Laitaoja, M.; Valjakka, J.; Jänis, J. *Inorg. Chem.* **2013**, *52* (19), 10983–10991.

Chapter 3, in part is currently being prepared for submission for publication of the material. Maniaci, Brian; Mealka, Matt; Onofrei, David; Stengel, Dillan; Huxford, Tom; Love, JJ. Design of Metal-Controlled Protein Heterodimers (*in preparation*). The dissertation author was the primary investigator and author of this material.