Published in final edited form as:

J Am Chem Soc. 2018 August 08; 140(31): 10043–10053. doi:10.1021/jacs.8b05812.

Determining the structural and energetic basis of allostery in a *de novo* designed metalloprotein assembly

Lewis A. Churchfield, Robert G. Alberstein, Laura M. Williamson, and F. Akif Tezcan*

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0356 United States

Abstract

Despite significant progress in protein design, the construction of protein assemblies that display complex functions (*e.g.*, catalysis or allostery) remains a significant challenge. We recently reported the *de novo* construction of an allosteric supramolecular protein assembly (Zn-C38/C81/C96R1₄) in which the dissociation and binding of Zn^{II} ions were coupled over a distance of 15 Å to the selective hydrolytic breakage and formation of a single disulfide bond. Zn₄:C38/C81/C96R1₄ was constructed by Zn^{II}-templated assembly of a monomeric protein (R1, a derivative of cytochrome *cb*₅₆₂) into a tetramer, followed by progressive incorporation of non-covalent and disulfide bonding interactions into the protein-protein interfaces to create a strained quaternary architecture. The interfacial strain thus built allowed mechanical coupling between the binding/dissociation of Zn^{II} and formation/hydrolysis of a single disulfide bond (C38-C38) out of a possible six. While the earlier study provided structural evidence for the two end-states of allosteric coupling, the energetic basis for allosteric coupling and the minimal structural requirements for building this allosteric system were not understood. Toward this end, we have characterized the structures and Zn-binding properties of two related protein constructs (C38/C96R1 and C38R1) which also possess C38-C38 disulfide bonds. In addition, we have carried out extensive molecular dynamics (MD) simulations of C38/C81/C96R1₄ to understand the energetic basis for the selective cleavage of the C38-C38 disulfide bond upon Zn^{II} dissociation. Our analyses reveal that the local interfacial environment around the C38-C38 bond is key to its selective cleavage, but this cleavage is only possible within the context of a stable quaternary architecture which enables structural coupling between Zn^{II} coordination and the protein-protein interfaces.

Graphical Abstract

Corresponding Author: tezcan@ucsd.edu.

Notes

The authors declare no competing financial interest.

The Supporting Information is available free of charge via the Internet at <http://pubs.acs.org>.

Additional experimental data

Coordinate file for apo-C38/C96R1₈

Structure factor file for apo-C38/C96R1₈

Coordinate file for Zn-C38/C96R1₄

Structure factor file for Zn-C38/C96R1₄

Topology file for Cso and Csa residues

Parameter file for the Cso and Csa residues

INTRODUCTION

Nature uses a limited set of tertiary folds as building blocks to create a vast array of protein assemblies and complexes that drive the chemical diversity of biological systems.^{1–4} This supramolecular mode of construction enables modular combination of different functionalities inherent in the protein building blocks, while also creating new molecular interfaces between them.^{1, 2} These interfaces themselves can house functional sites or serve as active conduits that are central to the structural dynamics of protein assemblies, underlying key phenomena such as catalysis, signal transduction and allostery.^{5–7} The bottom-up design of synthetic protein assemblies to carry out such sophisticated functions would provide an opportunity to expand upon the current scope of natural evolution. Additionally, it would enable the fabrication of systems that are free of evolutionary constraints and thus more amenable to yielding insights about their molecular mechanisms of action. Accordingly, there has been growing interest and significant progress in the rational design of supramolecular protein architectures.^{8–17} Yet, the number of artificial protein assemblies that possess emergent functional properties such as catalysis or allostery remains limited to a few instances.^{18–20}

Given the active and dynamic roles of interfaces in natural protein assemblies,²¹ we have sought not only to tailor protein-protein interactions to obtain desired supramolecular structures, but also to design the interfaces in such a way that would allow chemical (re)activities to be incorporated into them and for their dynamics to be modulated by external stimuli.^{20, 22–26} One approach, termed Metal-Templated Interface Redesign (MeTIR),^{22, 26} has exploited the strength, reversibility and inherent tunability of metal-coordination interactions to direct the self-assembly of proteins into well-defined and chemically responsive architectures. Based on the structures of these architectures, the surfaces of the protein building blocks can subsequently be tailored to incorporate additional non-covalent,²⁶ covalent^{20, 23, 25} or metal-coordination¹⁹ interactions within the nascent interfaces (Figure 1).

Using MeTIR, we recently reported the first designed allosteric supramolecular protein assembly ($\text{Zn}^{\text{C38/C81/C96}}\text{R1}_4$) in which the dissociation and binding of Zn^{II} ions was coupled over a distance of 15 Å to the hydrolytic breakage and formation of a disulfide bond.²⁰ This was achieved by building strain at the quaternary structure level, which led to a semi-rigid/flexible protein scaffold that allowed mechanical coupling between the two events (Figure 2).²³ The process of constructing $\text{Zn}^{\text{C38/C81/C96}}\text{R1}_4$ was as follows: a monomeric protein, cytochrome *cb*₅₆₂,²⁷ was first tailored on its surface with bis-His metal-chelating motifs to yield a D_2 symmetric tetramer that dissociates in the absence of Zn^{II} .²⁴

Using the crystal structure of the Zn-dependent tetramer as a guide, the engineered variant was further modified on its surface with computationally designed hydrophobic residues (to yield R1)²⁶ and then with a single Cys residue (to yield C⁹⁶R1)²³, such that the variant self-assembled into a tetramer, Zn-C⁹⁶R1₄ (Figure 1). As intended by the templating strategy, Zn-C⁹⁶R1₄ was stabilized by the hydrophobic interactions in the designed protein-protein interfaces (highlighted in cyan in Figure 1) and the two pairs of C96-C96 disulfide bonds in a pair of orthogonal interfaces (highlighted in green in Figure 1). Owing to the concurrent stabilization of two sets of interfaces, the tetrameric state was maintained in the absence of Zn^{II}. Yet, this architecture underwent a large Zn-coupled conformational change, enabled by the flexibility of the designed hydrophobic interactions and the C96-C96 crosslinks (Figure 1).²³ We next incorporated two additional pairs of disulfide linkages to fully crosslink each protomer to the others: C81-C81 linkages in the remaining non-designed interface (yielding C81/C⁹⁶R1₄)²⁵ and then C38-C38 cross-links within the designed interface (yielding C38/C81/⁹⁶R1₄).²⁰ This crosslinking was expected to counteract the inherent flexibility of the scaffold, while giving rise to a strained quaternary architecture (Figure 2). Indeed, the resulting hexuply disulfide-bonded tetrameric assembly, C38/C81/C⁹⁶R1₄, bound Zn^{II} ions with decreased affinity by 5 kcal/mol (~20 kJ/mol) compared to the less extensively disulfide-bonded C⁹⁶R1₄ and C81/C⁹⁶R1₄ assemblies.²⁰ Importantly, the dissociation of Zn^{II} ions was accompanied by the remote hydrolysis of a single C38-C38 bond, relieving structural strain, thus providing an allosteric system (Figure 2).²⁰

Allostery in natural proteins manifests in diverse and often subtle ways, extending well beyond the classic example of hemoglobin^{28, 29} to proteins involved in the innate immune system,^{30, 31} transcriptional regulation^{32, 33} and intracellular transport.³⁴ A quantitative and atomic-level description of allostery in highly evolved systems remains challenging, due in part to the complex functional connectivity of networks of amino acids that can span an entire protein.^{35, 36} This argues for the utility of model systems, where structural/functional elements of interest can be designed from scratch and manipulated to interrogate allosteric coupling. In this regard, the C38/C81/⁹⁶R1₄ assembly, which is similar to the classical allosteric protein hemoglobin in terms of its *D*₂ symmetric architecture²⁸ and adaptive interfaces,²⁹ provides an advantageous test case. Its interfaces, which contain both of the allosterically coupled elements (*i.e.*, Zn^{II} binding sites and disulfide linkages), are built from scratch through a minimal set of interactions and mutations in these interfaces, interactions that have readily identifiable structural/chemical consequences. With these features in hand, we set out to examine in greater detail the basis of allostery between Zn^{II} release and disulfide bond hydrolysis in C38/C81/⁹⁶R1₄: What is the minimal extent of interfacial connectivity/quaternary strain that could yield an allosteric system? Why is a single C38-C38 bond the weakest link among all other disulfide bonds? What provides the driving force for C38-C38 bond hydrolysis and how is the allosteric coupling energy stored/dissipated in the protein structure? To address these questions, we have carried out extensive experimental analyses and molecular dynamics simulations. Our results highlight the importance of structural dynamics at both local (*i.e.*, a single interface containing C38-C38) and global (*i.e.*, the entire quaternary assembly) levels in establishing this allosteric function.

RESULTS AND DISCUSSION

Preparation of alternative C38-bearing R1 variants

Our earlier study established the presence of allostery through the characterization of two structural end states ($^{C38/C81/96}R1_4$ and $Zn-^{C38/C81/96}R1_4$) (Figure 2). However, this analysis did not explain why a single C38-C38 bond is selectively broken/formed rather than the other disulfide bonds (a second C38-C38 bond and pairs of C81-C81 and C96-C96 bonds) and whether this breakage is a consequence of the full covalent connectivity of all interfaces. We therefore prepared two additional variants: $^{C38}R1$, which would only contain C38-C38 linkages upon self-assembly, and $^{C38/C96}R1$, which would contain a pair of C96-C96 linkages in addition to the C38-C38 disulfide bonds.

$^{C38}R1$ only has a single surface-exposed Cys and was primarily obtained as a disulfide-linked dimer, $^{C38}R1_2$, when purifying the protein in the absence of reductant (Figure 3a). The sedimentation velocity (SV) profile of the isolated $^{C38}R1_2$ species showed a single symmetric peak at $S_{max} = 2.7$ S, consistent with a dimeric state (Figure 3a). Interestingly, $^{C38}R1_2$ showed little tendency to form a Zn-mediated tetramer, even at relatively high protein and Zn^{II} concentrations (200 μ M) (Figure 3a) that were previously found to induce the tetramerization of the parent R1 variant lacking any disulfide crosslinks.²⁶ This observation suggested that the C38-C38 disulfide bonds disfavored the formation of the tetrameric state, or conversely, that the C38-C38 bonds would be destabilized within the context of a covalently enforced tetramer.

Since $^{C38/C96}R1$ bears more than one cysteine, its self-assembly can involve the formation of aggregates with undesired disulfide bonds. Therefore, the $^{C38/C96}R1$ protein was first isolated as a monomer under reducing conditions, and its self-assembly was carried out under mildly oxidizing conditions in the presence of 1.5 equivalents of Zn^{II} to template the formation of the desired tetramer with two pairs of C38-C38 and C96-C96 linkages. Under these conditions, the tetrameric species, $^{C38/C96}R1_4$, was obtained in $34 \pm 2\%$ yield (Figure S1 and Table S1) and could be isolated to $>85\%$ purity via size-exclusion chromatography (SEC) (Figure 3B). The exclusion of Zn^{II} from the self-assembly solution or the use of a variant lacking the hydrophobic residues that stabilize the designed interfaces decreased yields to $27 \pm 2\%$ and $16 \pm 1\%$ (Figure S1 and Table S1), respectively. This finding highlights the role of the designed hydrophobic interface in templating the desired tetrameric species. The SV profile of the isolated $^{C38/C96}R1$ species (5 μ M) in the presence of EDTA confirmed the formation of a compact tetramer that persists in the absence of bound metal ($S_{max} = 4.0$ S) (Figure 3B). The addition of equimolar Zn^{II} caused the SV peak to shift slightly ($S_{max} = 4.3$ S), indicating that there was a small metal-induced conformational change.

Thermodynamics of Zn^{II} binding by $^{C38/C96}R1_4$

We next determined the metal binding properties of $^{C38/C96}R1_4$. To this end, we carried out competitive binding titrations of Zn^{II} to $^{C38/C96}R1_4$ in the presence of the fluorescent metal indicator Fura-2 ($K_{d,Zn} = 5.7$ nM)²³ as a competing ligand. These titrations showed the expected stoichiometry of four Zn^{II} ions per tetramer (Figure 4), as has been previously

observed for other disulfide-linked R1 tetramers.^{20, 23, 25} The data could be well described by a thermodynamic model with four different dissociation constants ($K_d = 4 \pm 2$ nM; 3 ± 1 nM; 2 ± 1 nM; and 40 ± 10 nM), yielding a total free energy for Zn^{II} coordination of $G_{4\text{Zn}} = -44.0$ kcal/mol (Table 1). This value is very similar to those determined for $\text{C}^{96}\text{R}_{14}$ (-44.5 kcal/mol) and $\text{C}^{81/\text{C}^{96}}\text{R}_{14}$ (-43.7 kcal/mol), but *ca.* 5 kcal/mol more favorable than that for the fully covalently crosslinked $\text{C}^{38/\text{C}^{81}/\text{C}^{96}}\text{R}_{14}$ (-39.2 kcal/mol). This finding implies that the reduced Zn^{II} binding affinity of $\text{C}^{38/\text{C}^{81}/\text{C}^{96}}\text{R}_{14}$ must result from the combination of crosslinking across all interfaces to create a quaternary-level strain, and cannot be attributed solely to the presence of the C38-C38 linkages.

In our earlier examination of $\text{C}^{38/\text{C}^{81}/\text{C}^{96}}\text{R}_{14}$ by isothermal titration calorimetry (ITC), we had determined that Zn^{II} binding by that variant ($H_{\text{ITC}} = 2.4$ kcal/mol) was significantly less endothermic than $\text{C}^{96}\text{R}_{14}$ ($H_{\text{ITC}} = 15$ kcal/mol) and $\text{C}^{81/\text{C}^{96}}\text{R}_{14}$ ($H_{\text{ITC}} = 14$ kcal/mol), both of which lacked C38-C38 linkages. Since all of these variants bound Zn^{II} ions with identical metal coordination geometries²⁰, this ~ 12 kcal/mol difference was attributed to the formation of a C38-C38 bond upon Zn^{II} coordination by $\text{C}^{38/\text{C}^{81}/\text{C}^{96}}\text{R}_{14}$.²⁰ We found this to be the case for $\text{Zn}-\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ as well (see below). We carried out analogous ITC measurements for Zn^{II} binding by $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ and fit the data using the dissociation constants determined by fluorescence competition titrations to obtain a H_{ITC} of 4 ± 2 kcal/mol (Figure S2 and Table 1). Based on the similarity of this value to that for $\text{C}^{38/\text{C}^{81}/\text{C}^{96}}\text{R}_{14}$ (also containing C38-C38 bonds) and its dissimilarity to the enthalpies of Zn^{II} binding to $\text{C}^{96}\text{R}_{14}$ and $\text{C}^{81/\text{C}^{96}}\text{R}_{14}$ (lacking C38-C38 bonds) (Table 1), we conclude that C38-C38 bond formation/hydrolysis is also allosterically coupled to Zn^{II} coordination/dissociation in the $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ complex.

Structural characterization of the $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ complex

To further investigate allosteric coupling between Zn^{II} coordination and C38-C38 bond formation in $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$, we determined the crystal structures of Zn-bound and apo- $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ at 2.2 and 2.6 Å resolution, respectively (Figure 5). The structure of the Zn-bound $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ complex confirmed the expected stoichiometry of four Zn^{II} ions per tetramer, as well as the presence of two pairs of intact C38-C38 and C96-C96 disulfide linkages (Figure 5a). The $\text{Zn}-\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ complex is nearly identical to all other Cys-crosslinked R1 complexes, $\text{Zn}-\text{C}^{96}\text{R}_{14}$, $\text{Zn}-\text{C}^{81/\text{C}^{96}}\text{R}_{14}$ and $\text{Zn}-\text{C}^{38/\text{C}^{81}/\text{C}^{96}}\text{R}_{14}$, in terms of overall topology (C α RMSD = 1.18 Å, 1.31 Å and 0.97 Å, respectively), Zn^{II} coordination geometry and the alignment of the engineered hydrophobic interactions (Figure S3). In contrast, the structure of apo- $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ revealed an unexpected topology, in which eight protein units form a self-contained (*i.e.*, closed) architecture wherein all C38-C38 and C96-C96 disulfide bonds and the hydrophobic protein-protein interactions are satisfied. The arrangement and connectivity of the $\text{C}^{38/\text{C}^{96}}\text{R}_1$ units in the crystal lattice is illustrated in Figure 5b. As indicated by SV experiments and non-reducing SDS-PAGE (Figure 3b), apo- $\text{C}^{38/\text{C}^{96}}\text{R}_{14}$ exists predominantly as a tetramer in solution, with only a small fraction of higher order species, likely impurities that have not been eliminated by SEC.

Therefore, the crystallographically observed octameric arrangement of $\text{C}^{38/\text{C}^{96}}\text{R}_1$ can only be explained by the presence of an unstable or dissociated C38-C38 bond in the apo-

C38/C96R1₄ tetramer, which would allow it to ‘butterfly’ open and dimerize via the coupling of the free C38 residues *in crystallo* (Figure 5c). This lattice arrangement would not be possible—and indeed was not observed²⁰—with apo-C38/C81/C96R1₄, which possesses the additional C81-C81 crosslinks that prevent the tetramer from assuming an open conformation. Thus, our crystallographic findings, combined with the thermodynamic measurements, strongly indicate that allosteric coupling is also present in the C38/C96R1₄ complex despite the fact that it lacks a fully disulfide crosslinked structure like C38/C81/C96R1₄. Furthermore, our findings also suggest that the placement of the C38-C38 linkages in the designed, hydrophobic interfaces and confinement in a covalently linked tetramer must underlie their susceptibility to selective cleavage.

Computational analysis of allostery in C38/C81/C96R1₄

The available crystal structures of C38/C81/C96R1₄ in the Zn-bound and Zn-free forms represent the two end states of allosteric coupling between Zn^{II} removal and C38-C38 bond breakage. However, these structures provide little direct information about the energetic basis of this coupling, nor do they explain why a single C38-C38 is selectively hydrolyzed/formed among several others. This embodies a major obstacle to the study of any allosteric coupling process, as the means by which information is communicated structurally typically involves complex (and transient) sequences of localized conformational changes, which can be difficult to infer purely from well-defined end states.³⁵ Zn-C38/C81/C96R1₄ presents the added challenge of decoupling structural changes induced by Zn^{II} removal from those caused by disulfide bond breakage. Therefore, we set out to examine the C38/C81/C96R1₄ complex through molecular dynamics (MD) simulations, which allowed us to evaluate thermodynamic parameters for ensembles of structural states that would not be captured by experiments.

We first aimed to understand the energetics of C38/C81/C96R1₄ allostery by examining the elementary steps that must occur between the Zn-bound and Zn-free states (Figure 6a and 7a), focusing first on the disulfide bond breakage event. To estimate the energetic cost of a disulfide bond hydrolysis, we carried out density functional theory (DFT) calculations; the free energy of dimethyl disulfide hydrolysis (G_{break}) to yield a thiol and a sulfenate was found to be thermodynamically uphill by 17 kcal/mol. Next, we reasoned that the allosteric transformation of C38/C81/C96R1₄ following Zn^{II} removal could be decomposed into a bond hydrolysis and a structural rearrangement step (Figure 6a). We postulated that a) disulfide bond breakage events were most likely preceded by the formation of an energized, metal-free (but fully disulfide-bonded) C38/C81/C96R1₄ intermediate (termed **X**, Figure 6a), and b) structural rearrangements of this unstable intermediate was the primary energetic driver of disulfide hydrolysis (*i.e.*, $G_{\text{break}} + G_{\text{rearr}} < 0$). For the purposes of MD simulations, we prepared **X** by using the Zn-C38/C81/C96R1₄ crystal structure as a starting point, removing the metal ions, and solvating the system in an $85 \times 85 \times 85 \text{ \AA}^3$ box of explicit water with 150 mM NaCl. The resulting system of 56,796 atoms was equilibrated for at least 5 ns. Because the structure of **X** could be maintained indefinitely *in silico*, it permitted the evaluation of downstream reaction pathways that would not be accessible by experimental methods (Figure 6a).

We first investigated the uniqueness of the C38-C38 disulfide bond hydrolysis by simulating the hypothetical structural ensembles that would arise from cleavage of a single C81-C81 or C96-C96 bond in intermediate **X** (Figure 6a). Comparing the crystal structures of Zn-bound and apo-^{C38/C81/C96}R1₄, we previously observed that the Ca-Ca distance between the C38 residues (d_{38-38}) increased from 5.5 Å to 10.6 Å following tetramer opening accompanying Zn^{II} removal and C38-C38 bond hydrolysis. Likewise, breaking a C81-C81 bond could permit rearrangement to an open state (C81-C81_{open}), and a single cleaved C96-C96 bond could give rise to a distinct open state (C96-C96_{open}) wherein d_{81-81} or d_{96-96} would increase (Figure 6a). We simulated bond breakage by substituting an intact disulfide in **X** with a thiol-sulfenate pair. We monitored the Ca-Ca distances (d_{81-81} , d_{96-96}) for the two structural ensembles during the course of 100 ns simulations (Figure 6b-c). The simulated C81-C81 ensemble exhibited rapid and reversible transitions between an open ($d_{81-81} = 9$ Å) and closed state ($d_{81-81} = 6$ Å). Similarly, C96-C96 showed a preference to adopt a more open conformation ($d_{96-96} = 11.5$ Å), but still made occasional excursions to a closed-like state ($d_{96-96} = 7.5$ Å). As these trajectories appeared well-equilibrated, we could estimate the potential of mean force (PMF) for both constructs (Figure S4). The PMF for C81-C81 breakage is a double-well potential, possessing minima at 6 Å and 9 Å with a slight energetic preference for opening ($G_{\text{rearr.}} = -0.7$ kcal/mol). These minima are separated by a small barrier of ~1 kcal/mol, consistent with the rapid reversibility of the transition observed in the trajectory. In contrast, the PMF of the C96-C96 ensemble possesses a single minimum centered at 11.5 Å, modestly favoring tetramer opening ($G_{\text{rearr.}} = -3.5$ kcal/mol) with respect to d_{96-96} . The preference for both bond breakage events to permit tetramer opening is consistent with the view of ^{C38/C81/C96}R1₄ as a globally strained architecture. However, in both cases the cost of disulfide bond hydrolysis (17 kcal/mol) would significantly outweigh the driving force provided by structural rearrangement of the protein assembly ($G_{\text{break.}} + G_{\text{rearr.}} > 0$), explaining why the cleavage of C81-C81 and C96-C96 bonds is not observed.

We next undertook MD simulations of the C38-C38 ensemble, corresponding to the experimentally observed bond hydrolysis event (Figure 7a). The structure underwent an initial relaxation followed by a steady increase in d_{38-38} over the course of 200 ns, reaching a final value of ~10 Å (Figure 7b). Importantly, we did not observe any propensity of C38-C38 to adopt a closed-like state, suggesting that the structural rearrangement of the protein assembly is quasi-irreversible ($G_{\text{rearr.}} \ll 0$) and strongly favors an open conformation (and thus offsets the penalty of C38-C38 bond cleavage). When we separately simulated C38-C38_{open} from the metal-free crystal structure (see below), we noted that the equilibrated open state ($d_{38-38} = 11.5$ Å) was not fully reached following simulated bond breakage, likely due to a kinetic barrier imposed by global structural rearrangements.

The close agreement between the MD simulations and experimental observations prompted us to further examine the thermodynamic relationship between the C38-C38 bond hydrolysis reaction and the structural rearrangement of the ^{C38/C81/C96}R1₄ assembly. We evaluated the thermodynamics of the two end states of the C38-C38 bond hydrolysis reaction: 1) the intermediate **X** bearing all six disulfide bonds, and 2) C38-C38_{open} construct in the crystallographically observed open state (C38-C38_{open}) (Figure 7c). Surprisingly,

comparison of the configurational entropies of the two states revealed the C38-C38_{open} state to be entropically disfavored ($T \Delta S_{\text{config.}} = -67$ kcal/mol) compared to **X** (Figure 7c).

This entropic penalty must be viewed in the context of the overall structural rearrangement and could arise from a restriction of interfacial residues upon repacking. We determined the root-mean squared fluctuation of the residues comprising the engineered interface ($\text{RMSF}_{\text{interface}}$) in the two states, and found a restriction in the flexibility of the interface containing the broken C38-C38 disulfide ($\text{RMSF}_{\text{interface}} = -3.2$ Å) that outweighed changes in the opposing interface bearing an intact C38-C38 disulfide ($\text{RMSF}_{\text{interface}} = 1.3$ Å) (Table S2). This entropic penalty implied the existence of a strong enthalpic driving force ($H_{\text{rearr.}} \ll 0$) to render the structural rearrangement to be favorable.

We examined this enthalpic driving force by evaluating the potential energy differences experienced by each set of engineered residues between the two end states of C38-C38 bond breakage. This change in structural enthalpy ($H_{\text{struct.}}$) was fully converged after 300 ns (Figure S5) and provided a qualitative assessment of residues that participated favorably in the overall enthalpy of rearrangement (Table 2). Looking first at the residues associated directly with allosteric coupling, we found that the disulfide bonds exhibited a modest enthalpic gain ($H_{\text{struct.}} = -2$ kcal/mol) while the metal-binding residues incurred a significant enthalpic penalty ($H_{\text{struct.}} = +10$ kcal/mol), ruling these out as the primary drivers of C38-C38 bond hydrolysis. In contrast, the engineered hydrophobic interfaces (which contain the C38-C38 bonds) exhibited a large enthalpic gain ($H_{\text{struct.}} = -39$ kcal/mol), implicating their rearrangement as crucial to the adoption of the open conformation ultimately observed in the absence of Zn^{II} ions.

Taken together, these results strongly indicate that the allosterically relevant conformational changes of $\text{C}^{38}/\text{C}^{81}/\text{C}^{96}\text{R}_{14}$ are governed by an entropy/enthalpy tradeoff. This tradeoff primarily involves the designed, hydrophobic interfaces of **R1**, which rearrange upon Zn^{II} dissociation to form favorable packing interactions that overcome the energetic cost incurred by reduced configurational freedom of the assembly. We conclude then that the direct embedding of the C38-C38 bonds in these hydrophobic interfaces must account for their selective cleavage over the C96-C96 or C81-C81 linkages, which are located in the non-designed interfaces that are free of structural constraints. We further conclude that the driving force provided by the structural rearrangement of the hydrophobic interfaces is only sufficient for the hydrolysis of a single C38-C38 bond, and their coupled rearrangement and relaxation precludes breaking both C38-C38 bonds. This energetic limitation ultimately leads to the emergence of directionality (or asymmetry) in a symmetric system.

Conclusions

An important principle underpinning all natural protein assemblies is that the whole is greater than the sum of its parts. The synthetic construct $\text{Zn}-\text{C}^{38}/\text{C}^{81}/\text{C}^{96}\text{R}_{14}$ contains multiple engineered protomers that act cooperatively to satisfy this design imperative. Fundamentally, the structural basis of this behavior arises from a strained covalently linked quaternary architecture, evidenced by the diminished propensity for $\text{C}^{38}\text{R}_{12}$ to oligomerize. The intermediate level of crosslinking in $\text{C}^{38}/\text{C}^{96}\text{R}_1$ permits allosteric coupling due to the

structural strain imposed on the C38-C38 disulfide bonds, while Zn^{II} binding remained unaffected. However, the cooperative action of individual components to perform a synergistic function or exhibit an emergent property often hinges upon the coordinated interconversion between multiple conformational or energetic states.^{21, 37, 38}

Here, we have delineated how all of the designed self-assembly motifs in C³⁸/C⁸¹/C⁹⁶R₁₄ participate synergistically to create an allosteric protein architecture. Clearly, the coupling between the Zn^{II} binding sites and disulfide crosslinks must involve both of these engineered features. However, our MD simulations revealed that the repacking of the engineered hydrophobic contacts at the tetramer core must be critical as well; these fluid interfaces provide the requisite driving force that renders C38-C38 bond breakage uniquely favorable relative to C81-C81 and C96-C96 breakage events. This process is defined by a delicate thermodynamic balance of counteracting entropic and enthalpic factors, and asymmetric variation in how the interfacial residues are altered by the structural rearrangement. Allosteric transitions in natural proteins are energetically complex processes and can be modeled as minimizations of competing energetic forces in response to perturbations such as ligand binding.³⁹ In an analysis of several globular proteins, it was found that allosteric motions largely arise from sterically constrained competition among different, sequence-specific modes of hydrophobic burial,⁴⁰ in analogy to what we have observed with C³⁸/C⁸¹/C⁹⁶R₁₄. Thus, our study shows that a *de novo* designed—and arguably rudimentary—model system such as C³⁸/C⁸¹/C⁹⁶R₁₄ can recapitulate the complex interplay of energetic factors and structural features that underlie the coordinated functions of highly evolved proteins.

A detailed conceptual description of allosteric phenomena in natural protein assemblies began with the Koshland-Nemethy-Filmer⁴¹ and Monod-Wyman-Changeux⁴² models, which view allosteric transitions as a toggling between two well-defined structural states. A more recent view, termed the ensemble model,³⁵ considers the statistical distributions of an ensemble of states, and is inclusive of proteins in which allostery occurs without a significant structural change⁴³ and disorder (rather than robust structure) can efficiently mediate remote coupling.⁴⁴ We find that the structural toggling exhibited by Zn-C³⁸/C⁸¹/C⁹⁶R₁₄ can still be adequately described by the two-state allosteric model, due to the fact that it involves the formation/breakage of relatively strong covalent and coordination bonds which involve large changes in free energy. With further advances in our ability to design dynamic protein assemblies, it should be possible to create allosteric systems that operate via conformational states that lie closer in energy³⁷ and can be gated by events—such as ligand binding—that involve the formation/breakage of non-covalent interactions.

Materials and Methods

Site directed mutagenesis and protein purification

Protein samples were expressed and purified as previously described.^{20, 22} Briefly, the C38 mutation was introduced into plasmids encoding cyt *cb*₅₆₂ variants (pET20b-[cybC])²⁷ via site directed mutagenesis. Proteins were expressed in BL21(DE3) *E. coli* cells harboring the *ccm* heme maturation cassette plasmid, pEC86.⁴⁵ Cells were cultured in lysogeny broth (LB) supplemented with ampicillin (60 µg/mL) and chloramphenicol (34 µg/mL), with

protein expression occurring by auto-induction at 37 °C. Typically, 20 L of culture was harvested after a 16 h growth.

Pelleted cells were resuspended in a solution of 5 mM sodium acetate (pH 5), frozen, thawed, and sonicated for 12 min in pulses of 30 sec on and 60 sec off in the presence of lysozyme (~100 mg). In the case of $C^{38}/C^{96}R1$, excess DTT (~300 mg) was also added prior to sonicating. For purification in the reduced state, all buffers were supplemented with 2 mM DTT prior to templating. The lysate was treated with sodium hydroxide to a pH of 10, acetic acid to a final pH of 5, and cleared by centrifugation at 4 °C. The cleared lysate was applied to a CM Sepharose column (GE Healthcare) equilibrated in 5 mM sodium acetate (pH 5) and eluted in a NaCl gradient. Visibly red eluate was concentrated and dialyzed into a 10 mM sodium phosphate (pH 8) solution (supplemented with 2 mM DTT for reduced-state purification of $C^{38}/C^{96}R1$). Samples were applied to a High Q-cartridge column (BioRad) and eluted with a NaCl gradient. Fractions with Reinheitszahl ratios (A_{421}/A_{280} reduced; A_{415}/A_{280} oxidized) above 3 were retained and concentrated.

Samples of $C^{38}R1$ and templated $C^{38}/C^{96}R1$ (see below) were purified by size-exclusion chromatography (Superdex 75 resin from GE Healthcare). Sample purity was assessed spectrophotometrically ($A_{415}/A_{280} > 4$) and by non-reducing SDS-PAGE. Reduction of $C^{38}/C^{81}R1$ was readily accomplished with the addition of DTT, while $C^{38}R1$ reduction was carried out by addition of tris(2-carboxyethyl)phosphine hydrochloride (TCEP) and extensive boiling (15 min). The purest fractions were combined, concentrated, treated with excess (>10 equivalents) of EDTA, and exchanged into a solution of 20 mM MOPS (pH 7) and 150 mM NaCl. Stock solutions of disulfide bonded apo-protein were flash frozen and stored at -80 °C.

Zn-templated assembly of $C^{38}/C^{81}R1_4$

Prior to the templated self-assembly reaction, concentrated solutions of reduced $C^{38}/C^{81}R1$ were treated with excess (>10 equivalents) of DTT and transferred into a glove tent under an anaerobic (<10 ppm O_2) atmosphere of argon with 10% hydrogen, and exchanged into a solution of 50 mM Tris (pH 7) and 150 mM NaCl. Protein concentration was determined spectrophotometrically ($\epsilon_{415,ox.} = 0.148 \mu M^{-1} cm^{-1}$; $\epsilon_{421,red.} = 0.162 \mu M^{-1} cm^{-1}$)²⁷ and diluted to a final concentration of 50 μM . Zn-templated assembly was carried out in the presence of 75 μM $ZnCl_2$ (1.5 equivalents). Templating reactions were shaken overnight at 37 °C under ambient atmosphere, and templating efficiency was assessed by non-reducing SDS-PAGE.

Analytical ultracentrifugation

Solutions of oxidized apo-protein were treated with either 1 equivalent of $ZnCl_2$ or 1 mM EDTA. Sedimentation velocity measurements were made in a solution of 20 mM MOPS (pH 7) and 150 mM NaCl at 25 °C on a Beckman XL-1 Analytical Ultracentrifuge equipped with an An-60 Ti rotor and at 41,000 rpm. Samples were monitored at 415 nm for 5 μM protein samples, 450 nm for 50 μM samples, and 590 nm for 200 μM protein samples for at least 12 h. Scans were processed in Sedfit⁴⁶ using buffer viscosity (0.01002 poise), density (1.007 g/mL) and partial specific volume (0.7316 mL/g) parameters calculated by SEDNTERP

(<http://www.jphilo.mailway.com/download.htm>). The SV profiles in Figure 2 are reported at a confidence level of 0.95 and were plotted in Igor (WaveMetrics).

Competitive Zn^{II} binding titrations

All solutions were prepared using 20 mM MOPS (pH 7) and 150 mM NaCl pre-treated with Chelex 100 resin (BioRad). The concentration of a light-protected stock of Fura-2 (Invitrogen) was determined spectrophotometrically ($\epsilon_{362} = 27,000 \text{ M}^{-1} \text{ cm}^{-1}$).⁴⁷ A sample of C³⁸/C⁹⁶R1₄ (8.25 μM tetramer; 33 μM monomer) and Fura-2 (5 μM) was titrated with ZnCl₂ while thermostatted at 22 °C with 5 min incubations. Fluorescence emission at 510 nm was monitored over a 250–450 nm excitation scan (Figure S6). Fluorescence intensity for the 335 nm excitation as a function of [Zn^{II}] was modeled using Dynafit⁴⁸ as previously reported.^{20, 23, 25} The data could be described by a sequential binding model with four dissociation constants (Table S3), and was plotted using Igor (WaveMetrics).

Isothermal titration calorimetry

All solutions were prepared using 20 mM MOPS (pH 7) and 150 mM NaCl pre-treated with Chelex 100 resin (BioRad). C³⁸/C⁹⁶R1₄ was dialyzed twice at 4 °C for >8 h ($V_{\text{sample}}/V_{\text{reservoir}} > 200$), and the second dialysis buffer was retained for preparing titrant and titrand solutions. A VP-ITC calorimeter (Microcal) containing 25 μM C³⁸/C⁹⁶R1₄ titrand (1.4125 mL cell volume) and water reference was thermostatted at 22 °C for 180 s. The sample was titrated with 1 μL of 1.4 mM ZnCl₂, followed by a schedule of 30 5.5 μL injections delivered at 5 min intervals. The initial injection was discarded, and integrated heats of binding were background subtracted using the final five injection points. The data were fit to a four-site sequential binding model in ORIGIN (OriginLab) using the dissociation constants determined by competitive titration to determine the apparent enthalpies of Zn^{II} binding (H_{ITC}) (Table S3 and Figure S2).

X-ray crystallography

Sitting-drop vapor diffusion screens of apo-C³⁸/C⁹⁶R1₄ and Zn-C³⁸/C⁹⁶R1₄ (pre-treated with 1.5 equivalents of ZnCl₂) were carried out at room temperature with drops of 2 μL of protein and 1 μL of precipitant solution against 500 μL reservoirs of precipitant solution. Crystals of Zn-C³⁸/C⁹⁶R1₄ were obtained with a stock of 375 μM tetramer and a precipitant solution of 0.12 M CaCl₂, 0.2 M Tris (pH 7.5) and 24% PEG 1900 MME (w/v). Crystals of apo-C³⁸/C⁹⁶R1 were obtained using a protein stock of 675 μM tetramer mixed and a precipitant solution of 45% 2-methyl-2,4-pentane diol (v/v), 0.2 M Bis-Tris (pH 6.5), and 0.2 M CaCl₂. Crystals of Zn-C³⁸/C⁹⁶R1₄ were cryoprotected in perfluoropolyether (~1 min soak). Harvested crystals were frozen and stored in liquid nitrogen. Diffraction data were collected at 100 K with 0.98 Å radiation, on beamline BL9-2 at the Stanford Synchrotron Radiation Lightsource (SSRL) for apo-C³⁸/C⁹⁶R1 and beamline BL7-1 for Zn-bound C³⁸/C⁹⁶R1. The data were integrated using iMosflm,⁴⁹ scaled with SCALA,⁵⁰ and modeled by molecular replacement using C⁹⁶R1₁ (PDB: 3IQ6) as the search model in PHASER.⁵¹ Automated rigid-body and positional refinements were carried out using REFMAC⁵² with non-crystallographic symmetry (NCS) restraints that were gradually relaxed, alongside manual model rebuilding (loop deletion, disulfide bond addition, and placement of water and other heteroatoms) carried out in COOT.⁵³ The final treatment of the Zn-C³⁸/C⁹⁶R1₄ dataset

included an intensity-based twin refinement procedure (twin fraction = 49.5 %) in REFMAC, which has been previously used for a cyt *cb*₅₆₂ crystal with similar lattice parameters.⁵⁴ Ramachandran plots were calculated using PROCHECK.⁵⁵ Structural alignments and structure figures were generated using PyMOL.⁵⁶ Data collection and refinement parameters are reported in Table S4.

DFT Calculations

Geometry optimization, single point energy calculations, and frequency calculations were performed using Gaussian09 (Rev D.01).⁵⁷ Calculations were performed with the B3LYP functional and the 6-311+g(d,p) basis set, and using an implicit continuum of water and the Solvation Model based on Density (SMD).⁵⁸ The model reactants for disulfide breakage consisted of dimethyl disulfide and a hydroxide ion, affording methan-ethiol and methylsulfanolate products. G_{break} was estimated from the difference in free energies of the products and reactants.

Parameterization of sulfenic acid and sulfenate

We generated CHARMM-compatible forcefield parameters for the sidechains of cysteine sulfenic acid (Cso) and cysteine sulfenate (Csa) using protonated and deprotonated forms of ethane-SO-thioperoxol (Figure S7). Atom types were assigned by analogy to cysteine (for C and S) and serine (for O and H) in the CHARMM27 forcefield.^{59, 60} Partial charge, bond, angle, and dihedral parameters for these atoms were generated using the Forcefield Tool Kit (ffTK)⁶¹ in VMD.⁶² Geometry optimization was carried out in Gaussian⁵⁷ at the MP2/6-31G* level of theory, and CHARMM-type partial charges for the β -carbon, sulfur, oxygen, and the terminal hydrogen (for Cso) (Table S5) were obtained via 2D optimization of water hydrogen-bonding interactions at the HF/6-31G* level of theory followed by molecular mechanics fitting in ffTK.⁶¹ Bond (Table S6) and angle (Table S7) parameters were obtained by determining the Hessian matrix of the molecule's potential energy surface at the MP2/6-31G* level of theory and performing molecular mechanics fitting in ffTK.⁶¹ Dihedral parameters (Table S8) were obtained by torsion angle scans at the MP2/6-31G* level of theory in Gaussian followed by molecular mechanics fitting using ffTK.⁶¹ Artefactual parameters incorporating the terminal methyl group were excluded. Topology and parameter files for Cso and Csa are available in the online supplementary materials.

Model building and molecular dynamics simulations

The crystal structure of Zn-C₃₈/C₈₁/C₉₆R₁₄ (PDB: 5L32) with the Zn^{II} ions removed was used as the starting model for metal-free C₃₈/C₈₁/C₉₆R₁₄ with six disulfide bonds (intermediate **X**). The crystal structure of metal-free C₃₈/C₈₁/C₉₆R₁₄ (PDB: 5L31) with an unmodeled loop (residues 46–50, repaired using PyMOL⁵⁶) was used to model C₃₈/C₈₁/C₉₆R₁₄ in the open state (C₃₈-C₃₈_{open}). Heteroatoms, including waters, were removed, the tetramer interiors were measured using a 1.4 Å rolling probe in VOIDOO, and the internal cavities were solvated using FLOOD.⁶³ The PSFGEN plugin of VMD⁶² was used to add hydrogen atoms, assign atom types, join disulfide-bonded cysteine residues, c-type heme linkages, and axial heme ligands. CHARMM27 parameters,^{59, 60} custom parameters for c-type heme (HEMO),⁶⁴ and custom parameters for Csa (see above) were

used, consistent with previously reported pK_a 's of sulfenic acid.^{65, 66} Structures were solvated in an $85 \times 85 \times 85 \text{ \AA}^3$ cube of TIP3P water using SOLVATE⁶¹, leaving at least $10 \text{ \AA}$ of space on all sides in both states. Ions were placed using the Autoionize plugin⁶² to neutralize the system and add 150 mM NaCl . For the six-disulfide $C38/C81/C96R14$ system, 71 sodium ions and 47 chloride ions were added, and 72 sodium ions and 47 chloride ions were added to the five-disulfide-bearing tetramer. Final builds of the **X** and $C38\text{-}C38_{\text{open}}$ systems contained 56,796 atoms and 56,811 atoms, respectively.

All molecular dynamics simulations were carried out on a home-built workstation using NAMD2.10 with multi-core and CUDA support.⁶⁷ Simulations were carried out in the isobaric-isothermal (NPT) ensemble, with constant pressure (1.01325 bar) maintained by the Nosé-Hoover method (piston period = 200 fs ; piston decay = 50 fs). Temperature was maintained at 300 K by Langevin dynamics with a damping coefficient of 1 ps^{-1} . Non-bonded interactions were evaluated out to a $12.0 \text{ \AA}$ cutoff, with a switching function applied for distances exceeding $10.0 \text{ \AA}$, and with long-range electrostatic interactions calculated using the Particle mesh Ewald method. Hydrogen bond lengths were held fixed using the SHAKE algorithm.⁶⁸ Periodic boundary conditions were used for all simulations, and the dynamics trajectory was propagated using the velocity Verlet algorithm with a time step of 2 fs . For initial minimizations, all protein and heme atoms were held fixed, and the systems were subjected to 1,000 steps of steepest-descent energy minimization, followed by a further 5,000-step minimization with no fixed atoms. The systems were equilibrated for at least 5 ns (2.5×10^6 steps), at which point the structures appeared to have relaxed. Simulations were carried out at a rate of about 20 ns simulated time per day, with atom coordinates saved every 500 steps, or at 1 ps intervals.

Simulated disulfide bond breakage and free energy estimation

Frames of the simulation of **X** ($C38/C81/C96R14$ bearing six disulfide bonds) were used as the starting point for simulated bond breakage. Using PSFGEN in VMD,⁶² one pair of disulfide-bonded Cys were mutated to a Cys thiol and a Cys, with the missing terminal H and O atoms added, respectively. Initial guesses for terminal sidechain dihedrals were manually adjusted to avoid clashes where needed. A single water molecule was replaced with a sodium ion to neutralize the system using the Autoionize plugin of VMD.⁶² The systems were subjected to 5,000 step minimizations with the sidechain atoms of the newly broken disulfide bond allowed to move. After a 5 ns equilibration, the protein coordinates from 100 ns ($100,000$ frames) of the $C81\text{-}C81$ and $C96\text{-}C96$ simulations, and 200 ns ($200,000$ frames) of $C38\text{-}C38$, were extracted to measure d_{81-81} , d_{96-96} and d_{38-38} , respectively. The d_{81-81} and d_{96-96} values were binned in $0.5 \text{ \AA}$ increments (Figure S4) to obtain the corresponding probability distributions, $p(d_{81-81})$ and $p(d_{96-96})$, from which we estimated the relative free energies of each state via:

$$\Delta G_d = -k_b T \ln p(d)$$

Using the above relationship, we determined the potential mean force (PMF) of structural opening with respect to d_{81-81} and d_{96-96} (Figure S4) to estimate G_{rearr} for both hypothetical bond breaking events.

Thermodynamic analysis of simulated C38-C38 breakage

The configurational entropy ($S_{\text{config.}}$) of $\text{C}^{38}/\text{C}^{81}/\text{C}^{96}\text{R}_{14}$ bearing six disulfide bonds (**X**) and the $\text{C}^{38}\text{-C}^{38}_{\text{open}}$ state was estimated using the quasiharmonic (QH) approximation⁶⁹ implemented within the cpptraj program of AmberTools 16.⁷⁰ Protein coordinates from three 50 ns segments (150,000 frames total) of a 300 ns simulation were extracted for analysis. All heavy atom coordinates for each protein were aligned to the first frame of their respective trajectory, from which an average structure was calculated. The coordinates for each frame were realigned to this average structure, and the mass-weighted covariance matrix was calculated and diagonalized to obtain the QH mode frequencies. The entropy of the first 9,000 modes was evaluated using the analytical expression for the entropy of a quantum harmonic oscillator and summed to yield final estimates of $S_{\text{config.}}$ (Figure S8). To calculate $\text{RMSF}_{\text{interface}}$, coordinates of A34, W41, S42, W66 and I69 in individual chains, or individual dimers (Figure S9), were aligned to the first frame of their respective trajectories to obtain the $\text{RMSF}_{\text{interface}}$ for each set of residues (Table S2). The structural enthalpy ($H_{\text{struct.}}$) was computed for engineered interfacial residues (A34, W41, S42, W66 and I69), the metal binding ligands (H63, H73, D74, and H77), and the disulfide bonds (C38, C81, and C96) using the NAMDEnergy plugin in VMD. The Cys38/Csa38 pair in $\text{C}^{38}\text{-C}^{38}_{\text{open}}$, its intact C38-C38 counterpart in **X**, and all solvent atoms in both ensembles were excluded from H_{struct} calculations. H_{struct} for the groups of engineered residues was well converged after 300 ns (300,000 frames) (Figure S5).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Jonathan Rittle for helpful discussions, and for his assistance in carrying out the DFT calculations. This work was funded by the NSF (CHE1306656 and CHE1607145 to F.A.T.), as well as by the NIH (T32GM008326) through traineeships awarded to L.A.C and R.G.A. Crystal structures reported in this work (PDB IDs: 6DHY and 6DHZ) were collected at SSRL, which is supported by the DOE, Office of Science, Office of Basic Energy Sciences and Biological and Environmental Research, as well as by the NIH.

References

1. Ali MH, Imperiali B. *Bioorg Med Chem.* 2005; 13:5013. [PubMed: 15993087]
2. Goodsell DS, Olson AJ. *Annu Rev Biophys Struct.* 2000; 29:105–153.
3. At the time of writing, the CATH database (<http://www.cathdb.info/>) contained over 53 million entries classified into 1,391 topologies.
4. Ahnert SE, Marsh JA, Hernández H, Robinson CV, Teichmann SA. *Science.* 2015;350. [PubMed: 26472912]
5. Bertini I, Gray HB, Stiefel EI, Valentine JS. *Biological Inorganic Chemistry, Structure & Reactivity.* University Science Books; Sausalito: 2007.
6. Lippard S, Berg J. *Principles of Bioinorganic Chemistry.* University Science Books; Mill Valley: 1994.
7. Song WJ, Sontz PA, Ambroggio XI, Tezcan FA. *Ann Rev Biophys.* 2014; 43:409. [PubMed: 24773016]
8. Butterfield GL, Lajoie MJ, Gustafson HH, Sellers DL, Nattermann U, Ellis D, Bale JB, Ke S, Lenz GH, Yehdego A, Ravichandran R, Pun SH, King NP, Baker D. *Nature.* 2017; 552:415. [PubMed: 29236688]

9. Der BS, Machius M, Miley MJ, Mills JL, Szyperski T, Kuhlman B. *J Am Chem Soc.* 2012; 134:375. [PubMed: 22092237]
10. Lai YT, Reading E, Hura GL, Tsai KL, Laganowsky A, Asturias FJ, Tainer JA, Robinson CV, Yeates TO. *Nat Chem.* 2014; 6:1065. [PubMed: 25411884]
11. Gonen S, DiMaio F, Gonen T, Baker D. *Science.* 2015; 348:1365. [PubMed: 26089516]
12. Suzuki Y, Cardone G, Restrepo D, Zavattieri PD, Baker TS, Tezcan FA. *Nature.* 2016; 533:369–373. [PubMed: 27135928]
13. Alberstein R, Suzuki Y, Paesani F, Tezcan FA. *Nat Chem.* 2018; 10:732–739. [PubMed: 29713036]
14. Sontz PA, Bailey JB, Ahn S, Tezcan FA. *J Am Chem Soc.* 2015; 137:11598–11601. [PubMed: 26305584]
15. Brodin JD, Auyeung E, Mirkin CA. *Proc Natl Acad Sci USA.* 2015; 112:4564. [PubMed: 25831510]
16. Oohora K, Burazerovic S, Onoda A, Wilson YM, Ward TR, Hayashi T. *Angew Chem Int End Engl.* 2012; 124:3884.
17. Brodin JD, Ambroggio XI, Tang CY, Parent KN, Baker TS, Tezcan FA. *Nat Chem.* 2012; 4:375–382. [PubMed: 22522257]
18. Song WJ, Yu J, Tezcan FA. *J Am Chem Soc.* 2017; 139:16772–16779. [PubMed: 28992705]
19. Song WJ, Tezcan FA. *Science.* 2014; 346:1525–1528. [PubMed: 25525249]
20. Churchfield LA, Medina-Morales A, Brodin JD, Perez A, Tezcan FA. *J Am Chem Soc.* 2016; 138:13163–13166. [PubMed: 27649076]
21. Marsh JA, Teichmann SA. *Ann Rev Biochem.* 2015; 84:551. [PubMed: 25494300]
22. Bailey JB, Subramanian RH, Churchfield LA, Tezcan FA. *Methods Enzymol.* 2016; 580:223. [PubMed: 27586336]
23. Brodin JD, Medina-Morales A, Ni T, Salgado EN, Ambroggio XI, Tezcan FA. *J Am Chem Soc.* 2010; 132:8610–8617. [PubMed: 20515031]
24. Salgado EN, Faraone-Mennella J, Tezcan FA. *J Am Chem Soc.* 2007; 129:13374–13375. [PubMed: 17929927]
25. Medina-Morales A, Perez A, Brodin JD, Tezcan FA. *J Am Chem Soc.* 2013; 135:12013. [PubMed: 23905754]
26. Salgado EN, Ambroggio XI, Brodin JD, Lewis RA, Kuhlman B, Tezcan FA. *Proc Natl Acad Sci USA.* 2010; 107:1827–1832. [PubMed: 20080561]
27. Faraone-Mennella J, Tezcan FA, Gray HB, Winkler JR. *Biochemistry.* 2006; 45:10504. [PubMed: 16939202]
28. Perutz MF, Rossmann MG, Cullis AF, Muirhead H, Will G, North AC. *Nature.* 1960; 185:416–422. [PubMed: 18990801]
29. Yuan Y, Tam MF, Simplaceanu V, Ho C. *Chem Rev.* 2015; 115:1702. [PubMed: 25607981]
30. Hayden JA, Brophy MB, Cunden LS, Nolan EM. *J Am Chem Soc.* 2013; 135:775. [PubMed: 23276281]
31. Nakashige TG, Zhang B, Krebs C, Nolan EM. *Nat Chem Biol.* 2015; 11:765. [PubMed: 26302479]
32. Campanello GC, Ma Z, Grosseohme NE, Guerra AJ, Ward BP, Dimarchi RD, Ye Y, Dann CE, Giedroc DP. *J Mol Biol.* 2013; 425:1143. [PubMed: 23353829]
33. Capdevila DA, Braymer JJ, Edmonds KA, Wu H, Giedroc DP. *Proc Nat Acad Sci USA.* 2017; 114:4424–4429. [PubMed: 28348247]
34. Rubino JT, Franz KJ. *J Inorg Biochem.* 2012; 107:129. [PubMed: 22204943]
35. Motlagh HN, Wrabl JO, Li J, Hilser VJ. *Nature.* 2014; 508:331. [PubMed: 24740064]
36. Suel GM, Lockless SW, Wall MA, Ranganathan R. *Nat Struct Biol.* 2003; 10:59. [PubMed: 12483203]
37. Levy Y. *Biochemistry.* 2017; 56:5040. [PubMed: 28809494]
38. Frauenfelder H, Sligar SG, Wolynes PG. *Science.* 1991; 254:1598. [PubMed: 1749933]
39. Ferreira DU, Hegler JA, Komives EA, Wolynes PG. *Proc Natl Acad Sci USA.* 2011; 108:3499. [PubMed: 21273505]
40. England JL. *Structure.* 2011; 19:967. [PubMed: 21742263]

41. Koshland DE, Némethy G, Filmer D. *Biochemistry*. 1966; 5:365. [PubMed: 5938952]
42. Monod J, Wyman J, Changeux JP. *J Mol Biol*. 1965; 12:88. [PubMed: 14343300]
43. Popovych N, Sun S, Ebright RH, Kalodimos CG. *Nat Struct Mol Biol*. 2006; 13:831. [PubMed: 16906160]
44. Hilser VJ, Thompson EB. 2007; 104:8311.
45. Arslan E, Schulz H, Zufferey R, Kunzler P, Thony-Meyer L. *Biochem Biophys Res Commun*. 1998; 251:744. [PubMed: 9790980]
46. Schuck P. *Biophys J*. 2000; 78:1606. [PubMed: 10692345]
47. Gryniewicz G, Poenie M, Tsien RY. *J Biol Chem*. 1985; 260:3440. [PubMed: 3838314]
48. Kuzmic P. *Anal Biochem*. 1996; 237:260. [PubMed: 8660575]
49. Battye TG, Kontogiannis L, Johnson O, Powell HR, Leslie AG. *Acta Crystallogr, Sect D: Biol Crystallogr*. 2011; 67:271. [PubMed: 21460445]
50. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. *Acta Crystallogr D*. 2011; 67:235. [PubMed: 21460441]
51. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. *J Appl Crystallogr*. 2007; 40:658. [PubMed: 19461840]
52. Murshudov GN, Vagin AA, Dodson EJ. *Acta Crystallogr D Biol Crystallogr*. 1997; 53:240. [PubMed: 15299926]
53. Emsley P, Cowtan K. *Acta Crystallogr D Biol Crystallogr*. 2004; 60:2126. [PubMed: 15572765]
54. Salgado EN, Brodin JD, To MM, Tezcan FA. *Inorg Chem*. 2011; 50:6323. [PubMed: 21648390]
55. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. *J Appl Crystallogr*. 1993; 26:283.
56. Schrodinger, LLC. The PyMOL Molecular Graphics System, Version 1.3. 2010.
57. Frisch MJ, GWTSchlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Scalmani G, Barone V, Mennucci B, Petersson GA, Nakatsuji H, Caricato M, Li X, Hratchian HP, Izmaylov AF, Bloino J, Zheng G, Sonnenberg JL, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Montgomery JA, JrPeralta JE, Ogliaro F, Bearpark M, Heyd JJ, Brothers E, Kudin KN, Staroverov VN, Kobayashi R, Normand J, Raghavachari K, Rendell A, Burant JC, Iyengar SS, Tomasi J, Cossi M, Rega N, Millam JM, Klene M, Knox JE, Cross JB, Bakken V, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Martin RL, Morokuma K, Zakrzewski VG, Voth GA, Salvador P, Dannenberg JJ, Dapprich S, Daniels AD, Farkas Ö, Foresman JB, Ortiz JV, Cioslowski J, Fox DJ. *Gaussian 09, Revision D.01*. Gaussian, Inc; Wallingford CT: 2016.
58. Marenich AV, Cramer CJ, Truhlar DG. *J Phys Chem B*. 2009; 113:6378. [PubMed: 19366259]
59. Mackerell AD, Feig M, Brooks CL. *J Comput Chem*. 2004; 25:1400. [PubMed: 15185334]
60. MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FT, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiorkiewicz-Kuczera J, Yin D, Karplus M. *J Phys Chem B*. 1998; 102:3586. [PubMed: 24889800]
61. Mayne CG, Saam J, Schulten K, Tajkhorshid E, Gumbart JC. *J Comp Chem*. 2013; 34:2757. [PubMed: 24000174]
62. Humphrey W, Dalke A, Schulten K. *J Mol Graphics*. 1996; 14:33.
63. Kleywegt GJ, Jones TA. *Acta Crystallogr, Sect D: Struct Biol*. 1994; 50:178.
64. Autenrieth F, Tajkhorshid E, Baudry J, Luthey-Schulten Z. *J Comp Chem*. 2004; 25:1613. [PubMed: 15264255]
65. Nelson KJ, Parsonage D, Hall A, Karplus PA, Poole LB. *Biochemistry*. 2008; 47:12860. [PubMed: 18986167]
66. Hugo M, Turell L, Manta B, Botti H, Monteiro G, Netto LE, Alvarez B, Radi R, Trujillo M. *Biochemistry*. 2009; 48:9416. [PubMed: 19737009]
67. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, Chipot C, Skeel RD, Kale L, Schulten K. *J Comput Chem*. 2005; 26:1781. [PubMed: 16222654]

68. Ryckaert JP, Ciccotti G, Berendsen HJC. *J Comput Phys.* 1977; 23:327.
69. Andricioaei I, Karplus M. *J Chem Phys.* 2001; 115:6289.
70. Case DA, IYB-SBrozell SR, Cerutti DS, Cheatham TE, , IICruzeiro VWD, Darden TA, Duke RE, Ghoreishi D, Gilson MK, Gohlke H, Goetz AW, Greene D, Harris R, Homeyer N, Izadi S, Kovalenko A, Kurtzman T, Lee TS, LeGrand S, Li P, Lin C, Liu J, Luchko T, Luo R, Mermelstein DJ, Merz KM, Miao Y, Monard G, Nguyen C, Nguyen H, Omelyan I, Onufriev A, Pan F, Qi R, Roe DR, Roitberg A, Sagui C, Schott-Verdugo S, Shen J, Simmerling CL, Smith J, Salomon-Ferrer R, Swails J, Walker RC, Wang J, Wei H, Wolf RM, Wu X, Xiao L, York DM, Kollman PA. AMBER 2016. University of California; San Francisco: 2016.

Figure 1.

Iterative engineering of monomeric cytochrome cb_{562} . The addition of bis-His metal chelating motifs (red inset), computationally prescribed hydrophobic interfacial contacts (blue inset) and disulfide bond crosslinks (green inset) affords the homotetrameric protein, $Zn\text{-}^{C96}R1_4$, which undergoes a significant rearrangement about its engineered interfaces upon removal of Zn^{II} . PDB IDs: $cyt\ cb_{562}$ (2BC5), $Zn\text{-}^{C96}R1_4$ (3IQ5) and $^{C96}R1_4$ (3IQ6).

Figure 2. C38/C81/C96R1_4 allosterically couples Zn-removal from its core binding sites to breakage of a peripheral C38-C38 disulfide bond. This thermodynamically unfavorable bond breaking event is accompanied by significant structural rearrangement. PDB IDs: $\text{Zn-C38/C81/C96R1}_4$ (5L32) and C38/C81/C96R1_4 (5L31). G_{break} was estimated from DFT calculations of a model reaction.

Figure 3.

SDS-PAGE in the presence and absence of reductant, and sedimentation velocity profiles in the presence and absence of Zn^{II} , for purified (a) C38R1 and (b) C38/C96R1.

Figure 4. Zn-binding isotherm of C³⁸/C⁹⁶R1 determined using Fura-2 fluorescence, and corrected for dilution (red circles). Data were fit to a model with four consecutive binding equilibria (solid line). Vertical lines indicate equivalents of Zn^{II} (with respect to tetramer) added in excess of Fura-2.

Figure 5.

X-ray crystal structures of (a) Zn-bound (PDB: 6DHY) and (b) metal-free C38/C81/C96R1 (PDB: 6DHZ). Insets show the interfacial disulfide bonds: C38-C38 (pink) and C96-C96 (green). $2F_o - F_c$ electron density maps (black mesh) are contoured at 1σ . (c) Proposed scheme for the formation of the crystallographic octamer C38/C96R1 from a tetrameric species that bears a single broken disulfide bond. Due to the low extent of cross-linking, an intermediate butterfly tetramer could form by breaking and reforming the designed interfacial contacts (cyan) during crystallization.

Figure 6.

(a) Schematic representation of the hypothetical C81-C81 and C96-C96 bond breakage events. Removal of Zn^{II} by EDTA presumably furnishes an intermediate state (X) of C38/C81/C96R1₄ bearing six disulfide bonds. Breaking a single disulfide bond (purple circles) permits adoption of an open state. (b–c) Plots of C α separation (d_{81-81} , d_{96-96}) from MD simulations of the C81-C81 and C96-C96 ensembles. One broken disulfide bond (blue traces and insets) is modeled as a thiol/sulfenate pair, while its counterpart (black traces and insets) remains intact. G_{rearr} was estimated from the PMFs of tetramer opening with respect to d_{81-81} and d_{96-96} for the C81-C81 and C96-C96 ensembles.

Figure 7.

(a) Schematic representation of C38-C38 bond breakage (purple circles) from a six-disulfide metal-free intermediate, **X**. (b) Plot of Ca separation (d_{38-38}) of the simulated C38-C38 ensemble for the broken (blue) and intact (black) C38-C38 disulfides. (c) Thermodynamic analyses of the end states for C38-C38 bond breakage. $T\Delta S_{\text{config}}$ and $\text{RMSF}_{\text{interface}}$ were determined from MD simulations.

Table 1Thermodynamics of Zn^{II} binding to disulfide-linked R1 variants.

Variant	G_{4Zn} (kcal/mol)	H_{ITC} (kcal/mol)
C ⁹⁶ R1 ₄	-44.5 ± 0.5^a	$+15 \pm 3^c$
C38/C ⁹⁶ R1 ₄	-44.0 ± 0.2	$+4 \pm 3$
C81/C ⁹⁶ R1 ₄	-43.7 ± 0.2^b	$+14 \pm 1^c$
C38/C81/C ⁹⁶ R1 ₄	-39.2 ± 0.2^c	$+2.4 \pm 0.3^c$

[†]Free energy of Zn^{II} binding (G_{4Zn}) was determined by competitive titration, and uncertainties are fitting errors of a single measurement.

[‡]Apparent Zn-binding enthalpies (H_{ITC}) were determined by ITC, and uncertainties are the standard deviations of at least three replicate measurements.

^aFrom ref. 23.

^bFrom ref. 25.

^cFrom ref. 20.

Table 2

Changes in structural enthalpy ($H_{\text{struct.}}$) of key designed residues in C38/C81/C96R14 following C38-C38 breakage.

Protein feature	$H_{\text{struct.}}$ (kcal/mol)
Disulfide bonds	-2
Zn ^{II} binding sites	10
Hydrophobic interfaces	-39