

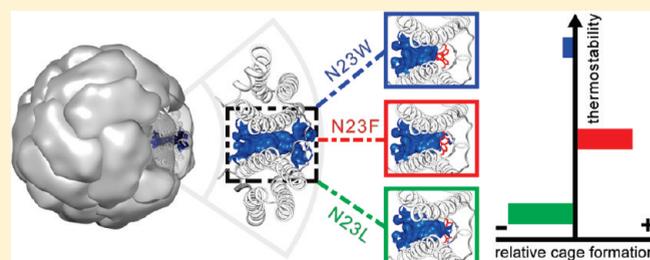
Stabilization of a Protein Nanocage through the Plugging of a Protein–Protein Interfacial Water Pocket

Maziar S. Ardejani, Noel X. Li, and Brendan P. Orner*

Division of Chemistry and Biological Chemistry, Nanyang Technological University, 21 Nanyang Link, Singapore 637371

S Supporting Information

ABSTRACT: The unique structural properties of the ferritin protein cages have provided impetus to focus on the methodical study of these self-assembling nanosystems. Among these proteins, *Escherichia coli* bacterioferritin (*EcBfr*), although architecturally very similar to other members of the family, shows structural instability and an incomplete self-assembly behavior by populating two oligomerization states. Through computational analysis and comparison to its homologues, we have found that this protein has a smaller than average dimeric interface on its 2-fold symmetry axis mainly because of the existence of an interfacial water pocket centered around two water-bridged asparagine residues. To investigate the possibility of engineering *EcBfr* for modified structural stability, we have used a semiempirical computational method to virtually explore the energy differences of the 480 possible mutants at the dimeric interface relative to that of wild-type *EcBfr*. This computational study also converged on the water-bridged asparagines. Replacing these two asparagines with hydrophobic amino acids resulted in proteins that folded into α -helical monomers and assembled into cages as evidenced by circular dichroism and transmission electron microscopy. Both thermal and chemical denaturation confirmed that, in all cases, these proteins, in agreement with the calculations, possessed increased stability. One of the three mutations shifts the population in favor of the higher-order oligomerization state in solution as evidenced by both size exclusion chromatography and native gel electrophoresis. These results taken together suggest that our low-level design was successful and that it may be possible to apply the strategy of targeting water pockets at protein–protein interfaces to other protein cage and self-assembling systems. More generally, this study further demonstrates the power of jointly employing *in silico* and *in vitro* techniques to understand and enhance biostructural energetics.



Protein–protein interactions are fundamental to instituting the specificity and integrity of the complex networks of macromolecules that function cooperatively to maintain and regulate integral cellular signaling and structure.^{1,2} Therefore, establishing a detailed and general understanding of protein–protein interactions will lead to an enhanced appreciation of greater biological processes.³ In addition, the rational design of various systems, whether they exist in the realm of structural biology^{4,5} or signal transduction,^{6,7} can emphatically demonstrate a top-to-bottom understanding of these systems or expose painfully a lack thereof.

Investigations involving the rational design of protein–protein interactions are rare, although the design of tertiary structure and small-molecule binding sites in proteins and structured peptides has become common.^{7–10} Much of the work in the design of protein–protein structure has focused on the synthetic–biological engineering of proteins to form novel nanoscale constructions or molecular devices through self-assembly.¹¹ A major thrust in this direction has been in the development of coiled-coil-based, α -helical polypeptides that can specifically assemble into nanotubes and nanofibers.^{12–14} Nanoscale cage structures, however, have been only rarely engineered, and few of these studies focus on the protein–protein interactions that control their assembly. Moreover,

many of these structures are multidomain virus capsid proteins that are structurally too complex to be applied as models for rational design.^{15,16}

Nanocage structures are formed by the assembly of the α -helical ferritin protein family. Although the structural biology of these proteins has been well-explored,¹⁷ there is still much not understood about their self-assembly and structural energetics.^{5,18} Moreover, although members of the ferritin family have been employed extensively in the formation and stabilization of inorganic nanomaterials,¹⁹ these applications are limited by the proteins' often nonideal stability and the lack of understanding of the general rules governing their assembly, an understanding of which could provide for the eventual rational engineering of proteins with novel physical properties and structures tailored for specific applications.

Because of the depth of the work on the ferritin family and their relatively simple structures made up of four-helix bundle monomers that assemble into discrete octahedral, 24-mer, homo-oligomers with C₄, C₃, and C₂ axes of symmetry, these proteins make up an ideal model system for the exploration of the

Received: February 10, 2011

Revised: April 13, 2011

Published: April 13, 2011

rational design of protein–protein interactions that shape self-assembly on a nanoscale. Such understanding could provide a step toward the de novo design of large protein suprastructures.

We⁵ and others²⁰ have previously demonstrated that *Escherichia coli* bacterioferritin (*EcBfr*) exists as an in vitro mixture of self-assembled dimers and 24-mers, and we have discovered a minidomain that, when removed, causes the protein to discretely populate the dimeric state. In addition, we have shown that it is possible to similarly shift the population to a discrete dimer and the overall stability through “alanine shaving” of single residues.¹⁸ These results clearly demonstrate that the self-assembly of this protein can be easily manipulated through subtle modification. In addition, other mutational studies of bacterioferritins have led to the production of variants with either crippled ability to assemble into cages or decreased stability compared to that of the wild-type parent.^{21,22} The only design to stabilize a bacterioferritin cage structure extended the C-terminus of *EcBfr*, resulting in a protein with a central cavity restricted by 60% and an inability to bind heme.²⁰

In this work, we first quantitatively compare the architectures of the bacterioferritin family to discover structural similarities among the bacterioferritins from various organisms and to identify the divergent architectural properties of *EcBfr* that might cause structural instability of this protein. Then we describe a minimal computational approach to engineering *EcBfr* for structural stability based on single-point mutations at protein–protein interfaces to plug water pockets, as opposed to the optimization of the monomer hydrophobic core, the engineering of metal ion or disulfide linkages, for example. Finally, we generate three of the computationally screened proteins and present characterization of their stability and self-assembly.

MATERIALS AND METHODS

The Supporting Information (Figure S1) contains the protein sequence of *EcBfr* and also highlights the positions of the individual mutations. Oligonucleotides were synthesized using solid-phase phosphoramidite technology by 1st Base Pte. Ltd. (Singapore). All concentrations of reagents in molecular biological reactions are initial concentrations before dilution unless otherwise stated.

Analysis of Bacterioferritin Family Protein–Protein Interfaces and Quaternary Structure. Comparative analysis of the protein–protein interfaces of the bacterioferritin family was performed using the web-based software “Protein interfaces, surfaces and assemblies service” (PISA, European Bioinformatics Institute).²³ The Protein Data Bank (PDB) structures used as input files are listed in Table S1 of the Supporting Information. Ligands and ions present in the crystal structures were omitted from calculations, and crystal structures were processed in “Auto” mode. The symmetry-related interfaces are defined in Figure S2 of the Supporting Information.

Free Energy Simulations of Virtual Mutants at the C2 Interface of *EcBfr*. All calculations for virtual mutation experiments with *EcBfr* were performed using FoldX 3.0 Beta3 (2009).²⁴ PDB entry 2vxi was chosen for computational mutagenesis experiments because of the several crystal structures available for *EcBfr*, this has the highest available resolution (1.9 Å). An *EcBfr* dimer was created by deleting all the subunits except two, chains C and D, which were centered at a C2 symmetry axis. The RepairPDB protocol was used to minimize the energy of the structure. The side chains were rotated while the backbones were

fixed, and the buffer sulfate and heme ligands were removed. This repaired PDB file was then used with the BuildModel protocol to explore the changes in unfolding free energies (relative to that of the wild type) of the 480 (24 interfacial residues \times 20 amino acids) possible point mutants (at 298 K, pH 7, and a 50 mM ionic strength). The calculation for each mutant was repeated at least five times, and energy minimization for each replicate was initiated from a different rotamer set. The $\Delta\Delta G$ values were averaged, and the standard deviation was calculated. The results of the calculation for some of the most stable mutants are compiled in Table S3 of the Supporting Information.

Analysis of Virtual Mutants at the C2 Interface of *EcBfr*. The wild-type and mutant C2 interfaces of *EcBfr* were further analyzed using the protein–protein interface comparison server, ProtorP, to extract detailed structural information.²⁵ The volume of “water accessible empty space” between the interfacial atoms, i.e., the pocket size, was measured using Computed Atlas of Surface Topography of proteins (CASTp)²⁶ with a probe radius of 1.4 Å. The pocket volume index is defined as the ratio of the pocket volume in cubic angstroms to the interface accessible surface area in square angstroms, which quantifies the relative importance of pockets in different protein interfaces. These results are compiled in Table S4 of the Supporting Information.

Preparation of Expression Constructs. Site-directed mutagenesis of wild-type *EcBfr* was performed, yielding the mutant constructs (QuickChange, Novagen). The sequences of the primers are presented in Figures S3–S7 of the Supporting Information. For each mutant, the mixing of reaction buffer (2 μ L of a 10 \times solution), the dsDNA template of wild-type *EcBfr*²⁷ (0.5 μ L of a 66.22 ng/ μ L solution), forward oligonucleotide primer (1 μ L of a 10 μ M solution), reverse oligonucleotide primer (1 μ L of a 10 μ M solution), dNTP mix (Fermentas, 2 μ L of a solution containing dATP, dTTP, dGTP, and dCTP at 2 mM each), and ddH₂O (13.5 μ L) was performed to bring the total volume to 20 μ L. *PfuTurbo* DNA polymerase (0.75 μ L of a 2.5 units/ μ L solution) was then added. The polymerase chain reaction (PCR) involved an initial step at 95 °C for 30 s followed by 18 cycles of 95 °C for 30 s, 55 °C for 1 min, and 68 °C for 6 min, which were followed by a subsequent extension step at 68 °C for 10 min. The *DpnI* restriction enzyme (0.5 μ L of a 20 μ g/ μ L solution) was added to 5 μ L of the PCR product together with NEBuffer 4 (New England Biolabs, 1 μ L) and ddH₂O (3.5 μ L) to bring the final volume to 10 μ L. The solution was incubated (37 °C, 1 h), and the mutated plasmids were transformed into electrocompetent NovaBlue *E. coli* and isolated by Miniprep (Qiagen). The success of the mutagenesis was determined through sequencing (Figures S3–S7 of the Supporting Information).

Gene Expression and Protein Purification. Plasmids harboring the desired genes were transformed into *E. coli* BL21-(DE3) (Novagen) cells by electroporation. Gene expression in the subsequent culture was induced (IPTG, Fermentas, 2 mL of a 0.1 M solution), and the culture was incubated (3 h, 30 °C). The cells were harvested by centrifugation and lysed by sonication. The soluble protein was applied to Ni-NTA resin (QIAGEN) and eluted either by affinity tag cleavage with the addition of enterokinase (New England Biolabs, 8 μ L of a 2 μ g/mL solution, 36 h, 4 °C) or buffers containing imidazole (250 mM) to elute the protein affinity fusions that were subsequently digested in solution and purified with Ni-NTA resin. The protein was concentrated via ultrafiltration (Sartorius Vivaspin 6) and further purified by size exclusion chromatography (Superdex 200

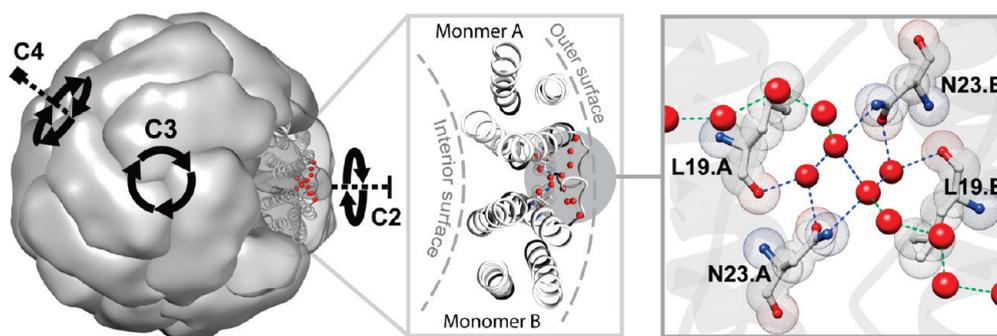


Figure 1. Structural hydration of the dimer interface in the *EcBfr* cage. Four water molecules buried at the interface make hydrogen bonds with the side chain of N23 and the backbone of L19 (shown by blue dashed lines) and interact with noninterfacial waters (shown by green dashed lines). The figure was made using UCSF Chimera and PDB entry 2vxi.

10/300 GL). The purity of the proteins was determined by SDS–PAGE (Figure S8 of the Supporting Information). The identity of the protein monomers was confirmed by MALDI TOF mass spectroscopy (Figure S9 and Table S5 of the Supporting Information).

Size Exclusion Chromatography (SEC). Purified protein solutions were equilibrated via gel filtration chromatography (GFC) running buffer [50 mM NaH_2PO_4 and 150 mM NaCl (pH 7.0)] through repeated ultrafiltration (Vivaspin, Sartorius) at 4 °C. The samples were subjected to size exclusion chromatography (AktaFPLC, Superdex 200 10/300 GL, flow rate of 0.5 mL/min, 4 °C). The column was calibrated using six proteins as standards (GE Biosystems Calibration Kit).⁵ Analytical SEC was performed at least five times for each protein using protein samples (10 μL , 300 $\mu\text{g}/\text{mL}$) equilibrated in GFC buffer (3 days, 4 °C).

Transmission Electron Microscopy (TEM). A solution of purified protein (10 μL , 100 $\mu\text{g}/\text{mL}$) was prepared in phosphate buffer [15.7 mM NaH_2PO_4 and 34.3 mM K_2HPO_4 (pH 7.2)] and applied as a drop to a piece of parafilm. A circular copper grid (300 mesh, Formvar/carbon-coated) was gently placed onto the protein drop and incubated (room temperature, 1 min). The excess solution was then removed from the grid by wicking with a piece of filter paper. Once the surface of the grid was dry, it was placed onto a drop of a uranyl acetate solution (10 μL , 1%, w/v) for 1 min. The excess stain was then removed by wicking, and the grid was left to dry. TEM data were obtained using a Jeol JEM-1400 transmission electron microscope operating at 100 keV. TEM micrographs were analyzed using ImageJ.²⁸

Dynamic Light Scattering (DLS). A solution of the purified protein (500 μL , 300 $\mu\text{g}/\text{mL}$) was prepared in phosphate buffer [50 mM NaH_2PO_4 and 150 mM NaCl (pH 7.0)]. A Brookhaven 90Plus Particle Size Analyzer was used to perform the DLS analysis using a 1 cm quartz cell. Each data set was collected over 3 min, and five data sets were collected for each protein. The data were used to calculate the average hydrodynamic diameter of the protein particles.

Native Polyacrylamide Gel Electrophoresis (native PAGE). The protein samples (10 μL , 300 $\mu\text{g}/\text{mL}$) were equilibrated in GFC buffer (3 days, 4 °C) and mixed with native sample buffer (10 μL , Bio-Rad) prior to analysis by native PAGE (4% stacking gel and 7% resolving gel). The gel was stained with Coomassie Brilliant Blue.

Circular Dichroism Spectroscopy and Thermal Denaturation. The protein solutions were equilibrated in phosphate buffer [15.67 mM NaH_2PO_4 and 34.33 mM K_2HPO_4 (pH 7.2)] through repeated ultrafiltration (Vivaspin, Sartorius, 0.2 μm). Protein concentrations were determined by BCA (Novagen),

and the solutions were diluted (300 $\mu\text{g}/\text{mL}$). CD spectra were recorded on a JASCO J-810 spectropolarimeter from 200 to 250 nm at temperatures ranging from 20 to 110 °C at 5 °C intervals (1 mm path length cuvette, 1 nm bandwidth, three accumulations, and 470 s equilibration time at each temperature point) (Figure S10 of the Supporting Information). Once the solutions reached 110 °C, they were cooled to 20 °C over 30 min and spectra were recorded from 200 to 250 nm for comparison with the spectra obtained at 20 °C before thermal unfolding.

Chemical Denaturation with Guanidine Hydrochloride. Guanidine stock solution (8 M) was made with the phosphate buffer used for CD (above). The concentration of the solution was determined by refractometry after it was filtered.²⁹ Chemical denaturation was initiated by diluting a concentrated protein stock into phosphate buffers containing varying concentrations of GuHCl (0–7.5 M). All samples contained 100 $\mu\text{g}/\text{mL}$ protein and were incubated at room temperature for 40 h before being measured. CD signals for each individual sample were measured 10 times at 222 nm and averaged as described above.

RESULTS

Comparative Analysis of Protein–Protein Interfaces and Quaternary Structure in the Bacterioferritin Family. Ferritins and bacterioferritins, the heme-containing ferritin from bacteria, essentially have the same quaternary and tertiary architectures. The structures of bacterioferritins from various organisms have been extensively studied.^{30–40} Bacterioferritins, because of their ability to host porphyrins and to perform bioinorganic catalysis, are candidate scaffolds for the design of synthetic biomolecular systems for a number of applications.⁴¹ Structural alignment (Table S2 of the Supporting Information) of *EcBfr* with the available crystal structures of the bacterioferritin family reveals that their backbones are highly superimposable (average root-mean-square deviation of 0.8 Å, z score of >20) and have similar tertiary structure even for proteins with a low level of sequence identity, suggesting that monomer structural analysis may not provide much direction in enhancing the stability of these proteins. Therefore, the higher-level structure was investigated.

Higher-order quaternary structure and the protein–protein interfaces of the protein cages of the bacterioferritin family were analyzed using PISA (Protein Interfaces, Surfaces and Assemblies). These calculations (Table S1 of the Supporting Information) show that the area buried upon formation of the cage for *EcBfr* is slightly smaller than the average. As the monomeric structure and the size of the 24-mer cage of the

bacterioferritins are very similar (Table S2 of the Supporting Information), these differences could be a result of variation of protein–protein interface size.

As the bacterioferritins assemble into 24-mer cages with octahedral symmetry, these data can be further parsed by comparing the buried surface area of the protein–protein interactions about the C2, C3, and C4 rotational axes (Figure S2A of the Supporting Information). This analysis reveals (Table S1 of the Supporting Information) that the total buried surface area of *EcBfr* is slightly below that of the average bacterioferritin because its C2 and C4 interfaces are smaller than the average even though the C3 interface is larger than the average (Figure S2B of the Supporting Information).

A deeper analysis reveals that a number of structural waters are bound at the *EcBfr* C2 dimeric interface (Figure 1), and the existence of this water pocket could explain why the C2 interface of *EcBfr* is smaller than average. These water molecules are coordinated to the side chains of N23 and the backbone of L19 from both chains of the dimer.³⁰ In the most recent crystal structure of *EcBfr* (PDB entry 2vxi), the water molecules in this pocket appear to be involved in interactions with other water molecules proximate to the interface and thus may be connected to the bulk solvent outside of the cage.³⁸ Exchange of these waters with the bulk has not yet been assessed. The presence of this water pocket suggests that a strategy for stabilizing *EcBfr* might be to plug the water pocket by optimizing the side chain packing at this region of the dimeric interface.⁴²

Computational Stabilization of *EcBfr*'s Dimeric Interface.

To redesign the dimeric interface using a minimal amount of computational power, we used a semiempirical force field, FoldX,⁴³ to explore the free energy differences of all possible point mutations of the C2, dimeric interface relative to the wild type. Each of the 24 residues at this interface was virtually mutated to each of the other 19 amino acids (and itself as a control), resulting in a total of 480 mutations in each of the two monomeric chains of the dimer. The free energy of unfolding (ΔG) was calculated for both the wild type and mutant for five replicates, and comparison of these energies ($\Delta\Delta G$) is an indication of the stabilizing or destabilizing effect of the mutation. In the free energy calculations, the effect of interactions with water molecules was investigated using the continuum solvation model implied in FoldX. This implicit solvation was included in all free energy simulations for the wild type and mutants. Basically, the extra stabilizing effect provided by a water molecule making more than one hydrogen bond with the protein that cannot be taken into account with nonexplicit solvent approximations is separately included in the FoldX force field.⁴³

Of the 23 mutants that exhibited the strongest stabilizing effect ($\Delta\Delta G < -1$ kcal/mol), most were positioned near the water pockets and many involved mutations from hydrophilic to hydrophobic residues. Mutations to methionine (6 of 23) were removed from the set as methionine, because of its conformational flexibility, tends to impart protein–protein interfaces with nonspecific plasticity^{44,45} and, because of its susceptibility to oxidation, can cause structural instability and aggregation.⁴⁶ Many of the predicted stabilizing mutations involved N23 (Table S3 of the Supporting Information), and it could be envisioned that the interface could be stabilized by using hydrophobic residues to bridge the water pocket. On the basis of this rationale, the three top predicted mutations, N23F, N23L, and N23W, were selected for further characterization in silico and in vitro.

Effects of Mutation on the Predicted Physical and Structural Properties of the Dimeric Interface. We examined the predicted structures of N23F, N23L, and N23W to determine how these

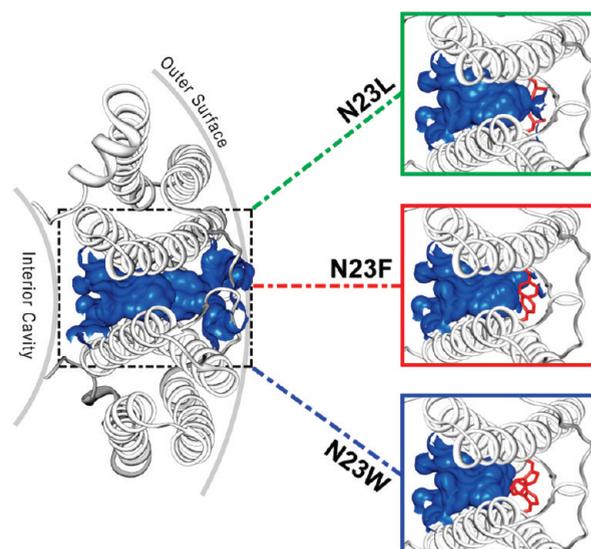


Figure 2. Plugging of the water pocket at the dimeric interface upon replacement of N23 with hydrophobic amino acids.

Table 1. Thermodynamic and Structural Analysis for the Wild Type and Redesigned Protein Variants

	water pocket volume (\AA^3)	dimeric interface area (\AA^2)	T_m ($^\circ\text{C}$)	ΔT_m ($^\circ\text{C}$)
wild type	214.0	913.2	58.9 ± 0.8	–
N23L	204.0	1036.7	60.1 ± 0.8	1.2
N23F	165.5	1069.4	69.5 ± 0.5	10.6
N23W	129.7	1137.8	80.1 ± 1.0	21.2

mutations affected the size of the water pocket and other physical properties of the C2 interface (Table S4 of the Supporting Information). In all cases, mutation caused the pocket to contract, with N23W having the strongest effect (Figure 2 and Table 1). Moreover, introduction of the mutations increased the amount of buried surface area, in general, and the amount of nonpolar surface area, in particular, at the interface. In all cases, the water was displaced, and because of new side chain interactions, the water-mediated hydrogen bonding interaction network around the wild-type Asn23 on *EcBfr* is replaced by contacts between the hydrophobic residues (Figure 3).^{47,48}

Cloning, Expression, and Purification of the *EcBfr* N23X Mutants. The N23F, N23W, and N23L mutants of *EcBfr* were cloned (Figures S3–S7 of the Supporting Information), expressed in *E. coli*, purified (Figure S8 of the Supporting Information), and confirmed by MALDI-TOF (Figure S9 and Table S5 of the Supporting Information).

Case Assembly of the *EcBfr* N23X Mutants (TEM). To determine whether mutation deleteriously affects the ability of the mutants to assemble into cage structures, we subjected the proteins to transmission electron microscopy (TEM) (Figure 4). TEM analysis of the *EcBfr* variants demonstrated detectable protein cages for all the proteins, indicating that the predicted stabilizing mutations do not cause significant change in the structure of the protein cage.

Solution Assembly of the *EcBfr* N23X Mutants (DLS). While TEM can show the inherent ability of proteins to form cages, the

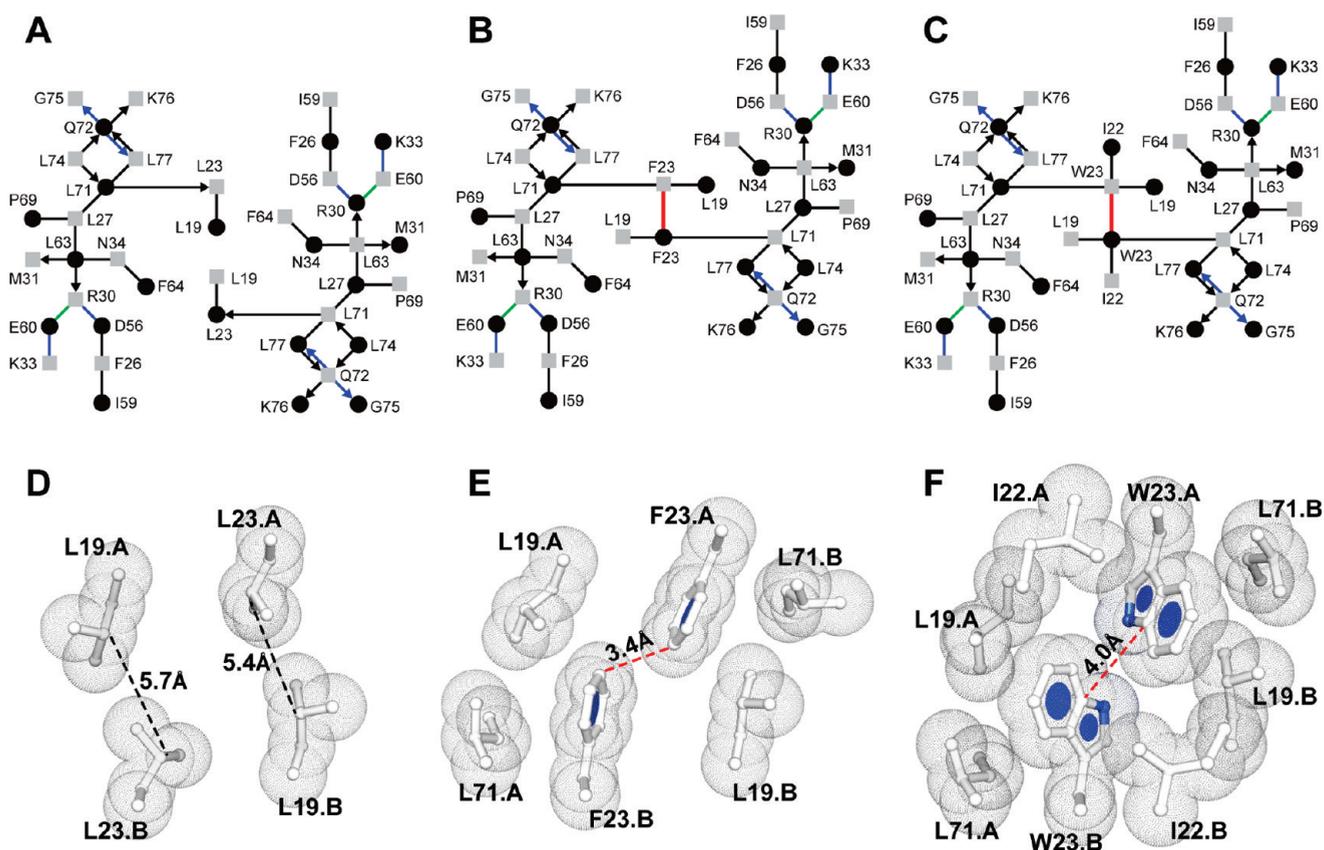


Figure 3. Computational prediction that mutation of N23 to hydrophobic residues repacks the dimeric interface. The interaction networks at the dimeric interface of mutant proteins, N23L (A), N23F (B), and N23W (C) are shown. Side chain–side chain interactions appear as lines, and backbone–side chain interactions are depicted as arrows pointing from the backbone to the side chain. Blue, green, and black lines or arrows represent hydrogen bonding, salt bridges, and van der Waals interactions, respectively. Gray squares and black circles depict the residues from chains A and B, respectively. The contact map was generated using Aquaprot and a symmetry assumption.⁴⁸ (D) The water-mediated hydrogen bonding interaction network around the wild-type Asn23 on *EcBfr* is replaced by the isosteric hydrophobic N23L mutation. (E) In the N23F mutant (E), residue L71 packs against the “offset stacked” F23 side chains to further stabilize the Phe–Phe aromatic interaction (plane–plane angle of 28.4° and centroid–centroid distance of 6.7 Å). (F) This network of hydrophobic residues expands in the N23W mutant by burial of I22 in the vicinity of the Trp–Trp aromatic interaction (plane–plane angle of 74.8°; $R_{Scen} = 7.7$ Å; $R_{6cen} = 6.3$ Å).

samples are in a dried, adsorbed state. Because these can be considered “forcing conditions”, TEM results sometimes are not indicative of what goes on in solution.¹⁸ To determine whether mutation affects the ability of the mutants to assemble in solution, dynamic light scattering was performed. All three mutants form particles that are similar in size to that of wild-type (WT) *EcBfr* (Figure 4F), suggesting that they do assemble in solution.

Determination of the Stoichiometry of Solution Assembly of the *EcBfr* N23X Mutants via Size Exclusion Chromatography (SEC) and Native PAGE. *EcBfr* exists as an ~1:1 mixture of 24-mer and dimer.^{5,18} To determine if the predicted stabilizing mutations can affect the ratio between these two populations in solution, we analyzed the proteins by size exclusion chromatography (SEC). Mutating the asparagine to phenylalanine (N23F), which was predicted to have the largest stabilizing role (Table S3 of the Supporting Information), shifts this ratio in favor of the 24-mer (Figure 5A,B), suggesting that our strategy was successful in discovering a protein with enhanced assembly properties. The other two mutants shift the ratio in favor of the dimer, suggesting that while this design was successful, forming hydrophobic interactions across the water pocket will not always enhance the assembly of the protein, implying that something more complicated may be occurring (see below). These trends were confirmed by native gel

electrophoresis (Figure 5C), although there is a minor band visible in the native gel that has not been characterized.

Thermal Stability of the *EcBfr* N23X Mutants. To determine if the predicted stabilizing mutations can influence the thermal stability of *EcBfr*, we subjected the proteins to thermal denaturation monitored by circular dichroism (CD). Although we would like to be able to measure directly the energetics of the protein–protein interactions, thermal denaturation measures the relative energy between the folded, oligomerized and unfolded, unoligomerized states that can be used as a metric, albeit an indirect one that could be complicated by additional oligomerization intermediates (see below). All the mutants fold into a helical secondary structure as evidenced by the characteristic CD minima at 208 and 222 nm (Figure 6A). To investigate their thermal stability, we monitored the 222 nm peak with respect to temperature (Figure 6B and Figure S10 of the Supporting Information), and the thermal transition temperatures (T_m) were determined. All of the mutants that were predicted to be more stable than the WT protein were indeed more stable, and the most stable mutant, N23W, was ~20 °C more stable than the WT protein.

Chemical Stability of the *EcBfr* N23X Mutants. To further characterize the relative stabilities of the proteins, we employed chemical denaturation with guanidinium hydrochloride (GuHCl).

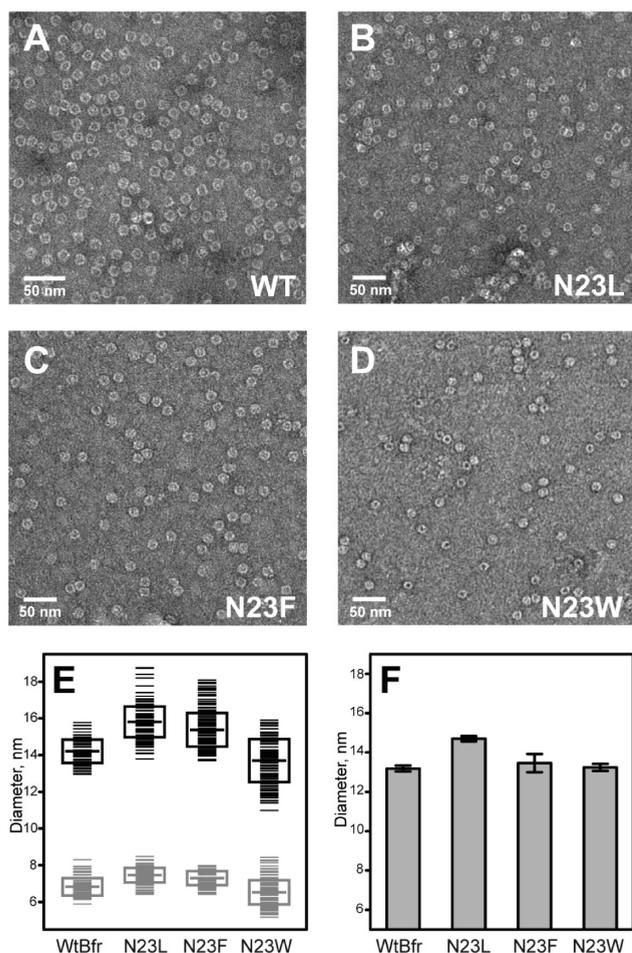


Figure 4. Redesigned proteins can assemble into nanostructures. TEM micrographs for wild-type *EcBfr* (A) and the N23L (B), N23F (C), and N23W (D) mutants. (E) Quantification of cage outer diameters (black) and cavity inner diameters (gray) from TEM micrographs. Boxes show the standard deviation, and the horizontal dash represents the mean. (F) Hydrodynamic diameters for protein cages as measured by DLS. Error bars represent the standard deviation.

The three mutants uphold the trend observed in the thermal denaturation, further suggesting that this simple virtual stabilization strategy can successfully predict mutations that result in enhanced stability (Figure 6C).

DISCUSSION

Rational designs for complex, folded proteins have had limited success, and the tailoring of protein quaternary structure to produce nanoarchitectures with new properties has been rare.¹⁶ As protein nanocages have been of interest to the materials community because of their nanoscale, aqueous solubility, and biocompatibility,⁴⁹ we reasoned that establishing and employing design strategies to these structures is timely. We investigated whether it would be possible to stabilize *E. coli* bacterioferritin (*EcBfr*) by focusing on protein–protein interactions that govern its self-assembly into an octahedral 24-mer. Through comparison with the other members of this structural family, we found that *EcBfr* has a smaller than average dimerization interface at its 2-fold rotational axis of symmetry. Further investigation revealed a pocket of bound waters at this interface centered around N23 and L19 (Figure 1). Concurrently, a virtual mutagenesis strategy was employed that computationally introduced every single-point mutation for all of the residues at this interface. Many of the stabilizing mutations predicted by this analysis were related to the water pocket in general and N23 specifically. Of these predicted stabilizing mutations, N23L, N23F, and N23W were selected for *in vitro* study (Figures 2 and 3). In all three cases, we found that they assemble into cages (Figure 4) and increase the thermal stability of the protein (Figure 6), and one of these pushed the oligomerization state toward the 24-mer at the expense of the dimer (Figure 5), suggesting that our relatively low-level computational strategy was successful and that the strategy of targeting water pockets at protein–protein interfaces may be able to be applied to other protein cage and self-assembling systems. More generally, this study further demonstrates the power of using *in silico* and *in vitro* studies in concert.⁵⁰

Along with the success of our relatively low-level computational approach, this research may provide additional insight. The lack of correlation between the overall stabilization provided by mutation and the influence mutation has on the oligomerization state is somewhat surprising, although we had previously found an alanine

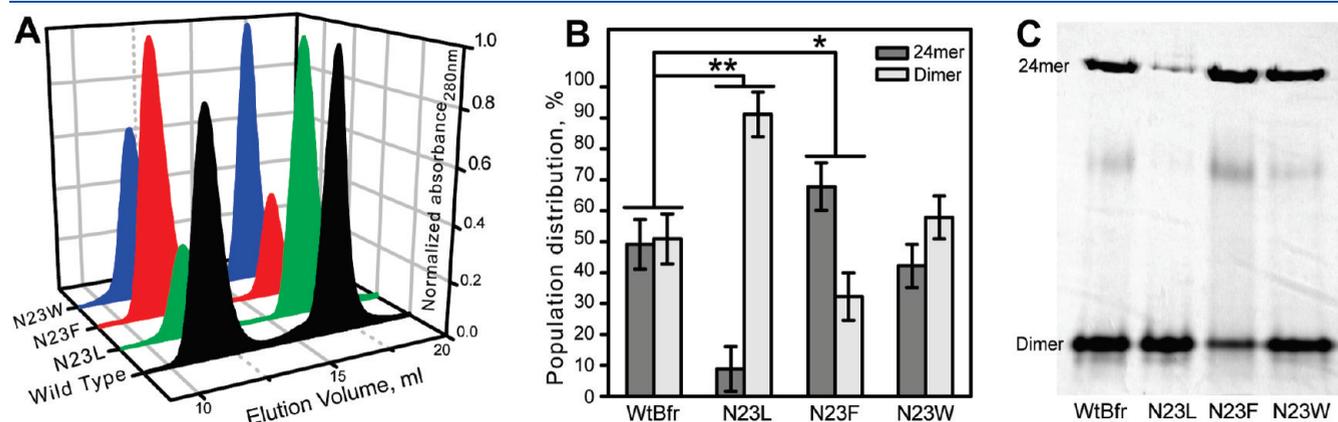


Figure 5. Redesigned proteins exhibit altered oligomerization propensities in solution. (A) Distribution of dimer vs 24-mer (left) determined by size exclusion chromatography (SEC); curves were normalized so that the height of the highest peak on all curves is the same. (B) Quantification of the SEC data determined by integrating the area under the SEC peaks for the mutant proteins (see Table S6 of the Supporting Information for values); one asterisk denotes that the two-tailed *P* value equals 0.0120, and two asterisks denote that the two-tailed *P* value equals 0.0005 according to the paired Student’s *t* test ($n = 6$). (C) Native PAGE shows the distribution of higher- and lower-order oligomerization states for each of the mutant proteins.

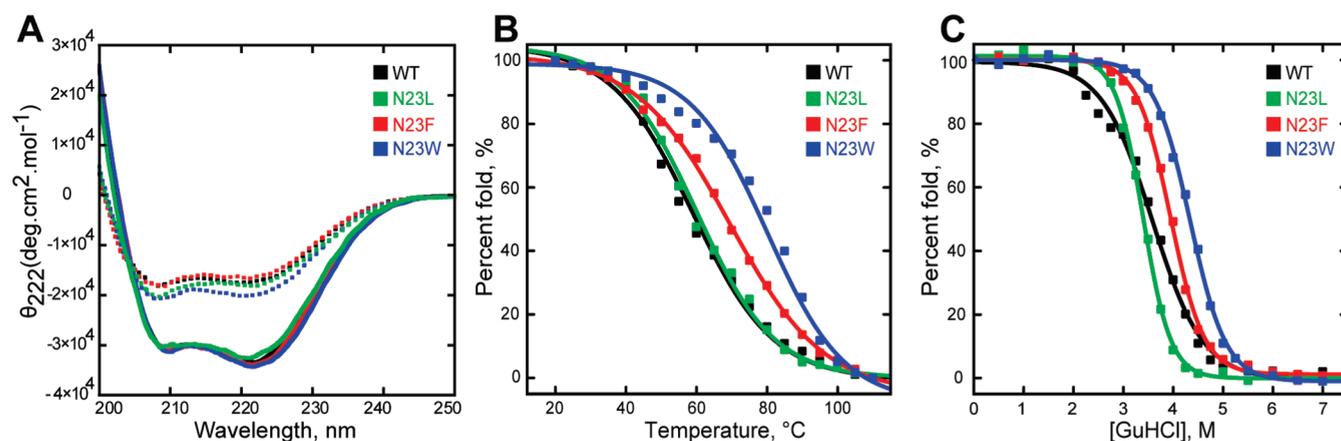


Figure 6. Redesigned proteins show partially reversible α -helical folds and enhanced thermal and chemical stability. (A) Circular Dichroism (CD) spectra of wild-type *EcBfr* and thermally more stable variants at 20 °C before thermal denaturation (—) and at 20 °C after thermal unfolding (···) as an indication of the native α -helical fold and the unfolding reversibility. (B and C) Percentage fold retained by the proteins as monitored by the CD signal at 222 nm during thermal (B) and GuHCl denaturation (C).

mutation in the dimeric interface of *EcBfr* that thermally stabilizes the protein but weakens its ability to assemble.¹⁸ However, it is possible to envision a system in which the size of the residue that can fit into a water pocket without changing the overall geometry of the dimer is limited by the size of the pocket itself. Increasing steric bulk over this limit may result in stabilization of the dimeric interface but may disrupt greater oligomerization. Because *EcBfr* naturally exists as a mixture of these two oligomerization states, it is especially sensitive to this effect and therefore is a useful tool for exploring it further. One could imagine that even subtle conformational changes caused by mutation at the dimeric interface may force the monomers to reposition slightly away from their native arrangement and consequently cause the dimer to become geometrically incompatible with the closed cage. The concept of a thermally stabilized but unproductive dimeric building block may have analogy with a masonry voussoir (the wedge-shaped stones used in constructing an arch); if its shape is altered to subtend too small or too large an angle, the assembled structure would not be able to span the native distance.⁵¹ In other words, geometrical changes at the dimer interface could result in the alteration of the projection of the other interfaces so that the subunits cannot make an enclosed cage. Along with the appropriate chemical complementarity, the partially assembled structure must provide an appropriate geometric environment for the docking of additional subunits with respect to the overall nanostructure. Thus, the elongation step of the self-assembly process becomes slower than the nucleation step, eventually resulting in kinetic trapping.⁵² In a sense, Fisher's "lock and key" for molecular recognition extends into "arch and keystone" for self-assembly.^{53–56} Intriguingly, expansion of this concept might result in the purposeful alteration of the geometry of the building blocks so that they interact in new ways, resulting in assemblies with novel oligomerization states or nanostructures.

■ ASSOCIATED CONTENT

S Supporting Information. Protein sequence alignment of the bacterioferritin protein family, a graph showing three symmetry-related interfaces in a typical maxiferritin, a table summarizing the results of structural analysis of bacterioferritin family assemblies, a table outlining the results of the structural alignment of the *EcBfr*

(PDB entry 2vxi) monomer with the other bacterioferritin monomers with available crystal structures, results of virtual mutation experiments, a table outlining the calculated properties of the C2, dimeric interfaces in wild-type *EcBfr* and designed variants, the sequencing results for all gene constructs, protein gels showing each step of the purification for all proteins, the results of MALDI-TOF mass spectrometry characterization of the purified proteins, and full CD spectra for the thermal denaturation experiments for all proteins. This material is available free of charge via the Internet at <http://pubs.acs.org>.

■ AUTHOR INFORMATION

Corresponding Author

*Telephone: +65 6316 8757. Fax: +65 6791 1961. E-mail: orner@ntu.edu.sg.

Funding Sources

M.S.A. is supported by SPMS graduate scholarships. N.X.L. thanks the CBC undergraduate summer research fund. This paper fulfills SPMS's "18-month rule" for M.S.A. This research was supported by an SPMS start-up grant, a Singapore Ministry of Education Academic Research Fund Tier 1 Grant (RG 53/06), and B.P.O.'s personal salary.

■ ACKNOWLEDGMENT

We thank Rongli Fan and Yu Zhang for their invaluable help and collaboration and Ai Hua Seow for use of the CBC teaching lab instruments. We also thank Indraneel Ghosh (University of Arizona, Tucson, AZ) and Zhang Dawei (Nanyang Technological University) for helpful discussions.

■ ABBREVIATIONS

Bfr, bacterioferritin; *EcBfr*, bacterioferritin from *E. coli*; PCR, polymerase chain reaction; Ek, enterokinase; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; EDTA, ethylenediaminetetraacetic acid; BCA, bicinchoninic acid; DTT, dithiothreitol; IPTG, isopropyl β -D-thiogalactopyranoside; SEC, size exclusion chromatography; CD, circular dichroism; TEM, transmission electron microscopy; DLS, dynamic light scattering; FPLC, fast protein liquid chromatography.

REFERENCES

- (1) Yamada, T., and Bork, P. (2009) Evolution of biomolecular networks: Lessons from metabolic and protein interactions. *Nat. Rev. Mol. Cell Biol.* 10, 791–803.
- (2) Beltrao, P., Kiel, C., and Serrano, L. (2007) Structures in systems biology. *Curr. Opin. Struct. Biol.* 17, 378–384.
- (3) Aloy, P., Böttcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.-C., Bork, P., Superti-Furga, G., Serrano, L., and Russell, R. B. (2004) Structure-based assembly of protein complexes in yeast. *Science* 303, 2026–2029.
- (4) Grueninger, D., Treiber, N., Ziegler, M. O. P., Koetter, J. W. A., Schulze, M.-S., and Schulz, G. E. (2008) Designed protein-protein association. *Science* 319, 206–209.
- (5) Fan, R., Boyle, A. L., Cheong, V. V., Ng, S. L., and Orner, B. P. (2009) A helix swapping study of two protein cages. *Biochemistry* 48, 5623–5630.
- (6) Yadav, S. S., Yeh, B. J., Craddock, B. P., Lim, W. A., and Miller, W. T. (2009) Reengineering the signaling properties of a Src family kinase. *Biochemistry* 48, 10956–10962.
- (7) Ulrich, S. M., Kenski, D. M., and Shokat, K. M. (2003) Engineering a “methionine clamp” into Src family kinases enhances specificity toward unnatural ATP analogues. *Biochemistry* 42, 7915–7921.
- (8) Orner, B. P., Salvatella, X., Sánchez Quesada, J., de Mendoza, J., Giralt, E., and Hamilton, A. D. (2002) De novo protein surface design: Use of cation- π interactions to enhance binding between an α -helical peptide and a cationic molecule in 50% aqueous solution. *Angew. Chem.* 114, 125–127.
- (9) Simon, M. D., Feldman, M. E., Rauh, D., Maris, A. E., Wemmer, D. E., and Shokat, K. M. (2006) Structure and properties of a re-engineered homeodomain protein–DNA interface. *ACS Chem. Biol.* 1, 755–760.
- (10) Gribbon, C., Channon, K. J., Zhang, W., Banwell, E. F., Bromley, E. H. C., Chaudhuri, J. B., Oreffo, R. O. C., and Woolfson, D. N. (2008) MagicWand: A single, designed peptide that assembles to stable, ordered α -helical fibers. *Biochemistry* 47, 10365–10371.
- (11) Bromley, E. H. C., Channon, K., Moutevelis, E., and Woolfson, D. N. (2008) Peptide and protein building blocks for synthetic biology: From programming biomolecules to self-organized biomolecular systems. *ACS Chem. Biol.* 3, 38–50.
- (12) Dong, H., Paramonov, S. E., and Hartgerink, J. D. (2008) Self-assembly of α -helical coiled coil nanofibers. *J. Am. Chem. Soc.* 130, 13691–13695.
- (13) Lazar, K. L., Miller-Auer, H., Getz, G. S., Orgel, J. P. R. O., and Meredith, S. C. (2005) Helix-turn-helix peptides that form α -helical fibrils: Turn sequences drive fibril structure. *Biochemistry* 44, 12681–12689.
- (14) Zhou, M., Bentley, D., and Ghosh, I. (2004) Helical supramolecules and fibers utilizing leucine zipper-displaying dendrimers. *J. Am. Chem. Soc.* 126, 734–735.
- (15) Mateo, R., Luna, E., Rincon, V., and Mateu, M. G. (2008) Engineering viable foot-and-mouth disease viruses with increased thermostability as a step in the development of improved vaccines. *J. Virol.* 82, 12232–12240.
- (16) Dalmau, M., Lim, S., and Wang, S.-W. (2009) Design of a pH-dependent molecular switch in a caged protein platform. *Nano Lett.* 9, 160–166.
- (17) Lee, S. S. C., and Richter, G. W. (1976) The monomers and oligomers of ferritin and apoferritin: Association and dissociation. *Biochemistry* 15, 65–70.
- (18) Zhang, Y., Raudah, S., Teo, H., Teo, G. W. S., Fan, R., Sun, X., and Orner, B. P. (2010) Alanine-shaving mutagenesis to determine key interfacial residues governing the assembly of a nano-cage maxi-ferritin. *J. Biol. Chem.* 285, 12078–12086.
- (19) Fan, R., Chew, S. W., Cheong, V. V., and Orner, B. P. (2010) Fabrication of gold nanoparticles inside unmodified horse spleen apoferritin. *Small* 6, 1483–1487.
- (20) Andrews, S. C., Smith, J. M. A., Hawkins, C., Williams, J. M., Harrison, P. M., and Guest, J. R. (1993) Overproduction, purification and characterization of the bacterioferritin of *Escherichia coli* and a C-terminally extended variant. *Eur. J. Biochem.* 213, 329–338.
- (21) Wong, S. G., Tom-Yew, S. A. L., Lewin, A., Le Brun, N. E., Moore, G. R., Murphy, M. E. P., and Mauk, A. G. (2009) Structural and mechanistic studies of a stabilized subunit dimer variant of *Escherichia coli* bacterioferritin identify residues required for core formation. *J. Biol. Chem.* 284, 18873–18881.
- (22) Kilić, M. A., Spiro, S., and Moore, G. R. (2003) Stability of a 24-meric homopolymer: Comparative studies of assembly-defective mutants of *Rhodobacter capsulatus* bacterioferritin and the native protein. *Protein Sci.* 12, 1663–1674.
- (23) Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. *J. Mol. Biol.* 372, 774–797.
- (24) Schymkowitz, J., Borg, J., Stricher, F., Nys, R., Rousseau, F., and Serrano, L. (2005) The FoldX web server: An online force field. *Nucleic Acids Res.* 33, W382–W388.
- (25) Reynolds, C., Damerell, D., and Jones, S. (2009) ProtorP: A protein-protein interaction analysis server. *Bioinformatics* 25, 413–414.
- (26) Binkowski, T. A., Naghibzadeh, S., and Liang, J. (2003) CASTp: Computed atlas of surface topography of proteins. *Nucleic Acids Res.* 31, 3352–3355.
- (27) Zhang, Y., Raudah, S., Teo, H., Teo, G. W. S., Fan, R., Sun, X., and Orner, B. P. (2010) Alanine-shaving mutagenesis to determine key interfacial residues governing the assembly of a nano-cage maxi-ferritin. *J. Biol. Chem.* 285, 12078–12086.
- (28) Gallagher, S. R. (2010) Digital Image Processing and Analysis with ImageJ, in *Current Protocols Essential Laboratory Techniques*, pp 3: A.3C.1–A.3C.24, John Wiley & Sons, Inc.
- (29) Nozaki, Y. (1972) The preparation of guanidine hydrochloride. In *Methods in Enzymology* (Hirs, C., Ed.) pp 43–50, Academic Press, San Diego.
- (30) Frolov, F., Kalb, A. J., and Yariv, J. (1994) Structure of a unique twofold symmetric haem-binding site. *Nat. Struct. Mol. Biol.* 1, 453–460.
- (31) Cobessi, D., Huang, L.-S., Ban, M., Pon, N. G., Daldal, F., and Berry, E. A. (2002) The 2.6 Å resolution structure of *Rhodobacter capsulatus* bacterioferritin with metal-free dinuclear site and heme iron in a crystallographic 'special position'. *Acta Crystallogr. D* 58, 29–38.
- (32) Macedo, S., Romao, C. V., Mitchell, E., Matias, P. M., Liu, M. Y., Xavier, A. V., LeGall, J., Teixeira, M., Lindley, P., and Carrondo, M. A. (2003) The nature of the di-iron site in the bacterioferritin from *Desulfovibrio desulfuricans*. *Nat. Struct. Mol. Biol.* 10, 285–290.
- (33) Swartz, L., Kuchinskas, M., Li, H., Poulos, T. L., and Lanzilotta, W. N. (2006) Redox-dependent structural changes in the *Azotobacter vinelandii* bacterioferritin: New insights into the ferroxidase and iron transport mechanism. *Biochemistry* 45, 4421–4428.
- (34) van Eerde, A., Wolterink-van Loo, S., van der Oost, J., and Dijkstra, B. W. (2006) Fortuitous structure determination of 'as-isolated' *Escherichia coli* bacterioferritin in a novel crystal form. *Acta Crystallogr. F* 62, 1061–1066.
- (35) Robert, J., Tamar, A.-N., and Manfred, S. W. (2008) Bacterioferritin from *Mycobacterium smegmatis* contains zinc in its di-nuclear site. *Protein Sci.* 17, 1138–1150.
- (36) Crow, A., Lawson, T. L., Lewin, A., Moore, G. R., and Brun, N. E. L. (2009) Structural basis for iron mineralization by bacterioferritin. *J. Am. Chem. Soc.* 131, 6808–6813.
- (37) Gupta, V., Gupta, R. K., Khare, G., Salunke, D. M., and Tyagi, A. K. (2009) Crystal structure of BfrA from *Mycobacterium tuberculosis*: Incorporation of selenomethionine results in cleavage and demetallation of haem. *PLoS One* 4, e8028.
- (38) Willies, S., Isupov, M., Garman, E., and Littlechild, J. (2009) The binding of haem and zinc in the 1.9 Å X-ray structure of *Escherichia coli* bacterioferritin. *J. Biol. Inorg. Chem.* 14, 201–207.
- (39) Nam, K. H., Xu, Y., Piao, S., Priyadarshi, A., Lee, E. H., Kim, H.-Y., Jeon, Y. H., Ha, N.-C., and Hwang, K. Y. (2010) Crystal structure of bacterioferritin from *Rhodobacter sphaeroides*. *Biochem. Biophys. Res. Commun.* 391, 990–994.
- (40) Weeratunga, S. K., Lovell, S., Yao, H., Battaile, K. P., Fischer, C. J., Gee, C. E., and Rivera, M. (2010) Structural studies of bacterioferritin B

from *Pseudomonas aeruginosa* suggest a gating mechanism for iron uptake via the ferroxidase center. *Biochemistry* 49, 1160–1175.

(41) Conlan, B., Cox, N., Su, J.-H., Hillier, W., Messinger, J., Lubitz, W., Dutton, P. L., and Wydrzynski, T. (2009) Photo-catalytic oxidation of a di-nuclear manganese centre in an engineered bacterioferritin 'reaction centre'. *Biochim. Biophys. Acta* 1787, 1112–1121.

(42) Dodge, M. M. (1896) *Hans Brinker, or, The silver skates, a story of life in Holland*, Scribner, New York.

(43) Guerois, R., Nielsen, J. E., and Serrano, L. (2002) Predicting changes in the stability of proteins and protein complexes: A study of more than 1000 mutations. *J. Mol. Biol.* 320, 369–387.

(44) Gellman, S. H. (1991) On the role of methionine residues in the sequence-independent recognition of nonpolar protein surfaces. *Biochemistry* 30, 6633–6636.

(45) O'Neil, K. T., and DeGrado, W. F. (1990) How calmodulin binds its targets: Sequence independent recognition of amphiphilic α -helices. *Trends Biochem. Sci.* 15, 59–64.

(46) Hu, D., Qin, Z., Xue, B., Fink, A. L., and Uversky, V. N. (2008) Effect of methionine oxidation on the structural properties, conformational stability, and aggregation of immunoglobulin light chain LEN. *Biochemistry* 47, 8665–8677.

(47) Swint-Kruse, L. (2004) Using networks to identify fine structural differences between functionally distinct protein states. *Biochemistry* 43, 10886–10895.

(48) Reichmann, D., Phillip, Y., Carmi, A., and Schreiber, G. (2008) On the contribution of water-mediated interactions to protein-complex stability. *Biochemistry* 47, 1051–1060.

(49) Kramer, R. M., Li, C., Carter, D. C., Stone, M. O., and Naik, R. R. (2004) Engineered protein cages for nanomaterial synthesis. *J. Am. Chem. Soc.* 126, 13282–13286.

(50) Yu, L., Yu, P. S., Yee Yen Mui, E., McKelvie, J. C., Pham, T. P. T., Yap, Y. W., Wong, W. Q., Wu, J., Deng, W., and Orner, B. P. (2009) Phage display screening against a set of targets to establish peptide-based sugar mimetics and molecular docking to predict binding site. *Bioorg. Med. Chem.* 17, 4825–4832.

(51) Macnab, R. M., and Aizawa, S. I. (1984) Bacterial motility and the bacterial flagellar motor. *Annu. Rev. Biophys. Bioeng.* 13, 51–83.

(52) Hagan, M. F., and Elrad, O. M. (2010) Understanding the concentration dependence of viral capsid assembly kinetics the origin of the lag time and identifying the critical nucleus size. *Biophys. J.* 98, 1065–1074.

(53) Fischer, E. (1894) Einfluss der configuration auf die wirkung der enzyme. *Ber. Dtsch. Chem. Ges.* 27, 2985–2993.

(54) Lichtenthaler, F. W. (1995) 100 Years "Schlüssel-Schloss-Prinzip": What made Emil Fischer use this analogy? *Angew. Chem., Int. Ed.* 33, 2364–2374.

(55) Friedrich, C. (2007) Emil Fischer's lock-and-key hypothesis after 100 years: Towards a supracellular chemistry. In *Perspectives in Supramolecular Chemistry* (Behr, J.-P., Ed.), pp 1–23, John Wiley & Sons, Inc., New York.

(56) The Constitution of the Commonwealth of Pennsylvania (1970) Philadelphia.