Protein interface redesign facilitates conversion of zero-dimensional protein nanomaterials to their one- and two-dimensional analogues

Xiaorong Zhang  
China Agricultural University

Yu Liu  
China Agricultural University

Bowen Zheng  
China Agricultural University

Jiachen Zang  
China Agricultural University

Chenyuan Lv  
China Agricultural University

Tuo Zhang  
Michigan State University

Hongfei Wang  
RIKEN

Guanghua Zhao (✉️ gzhao@cau.edu.cn)  
China Agricultural University  https://orcid.org/0000-0001-8587-9680

Article

Keywords: protein nanomaterials, protein assemblies

DOI: https://doi.org/10.21203/rs.3.rs-170719/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Although various artificial protein nanoarchitectures have been constructed, controlling conversion between protein assemblies with different dimensions has largely been unexplored. Here, we describe a simple, effective approach to regulate conversion between 0D protein nanomaterials and their 1D or 2D analogues by adjusting the geometric arrangement of dimeric protein building blocks. *Thermotoga maritima* ferritin (TmFtn) naturally occurs as a dimeric protein, twelve of which interact with each other in a head-to-side manner to generate 0D 24-meric protein nanocage in the presence of Ca\(^{2+}\). By tuning two contiguous dimeric proteins to interact in a fully or partially side-by-side fashion through protein interface redesign, we can render the conversion of the inherent salt-mediated 0D protein nanocage into 1D or 2D nanomaterials in response to multiple external stimuli. Thus, one kind of dimeric protein building block can generate three protein materials with different dimensions in a manner that highly resembles natural pentamer building blocks from viral capsids that form different protein assemblies.

Introduction

Shape transformation phenomena are ubiquitous in nature. Many living organisms by shape transformation perform shape-to-function activities in response to the external environment\(^1\). For instance, *Amoeba proteus* undergoes multidirectional shape transformation to form pseudopod for navigation and rapid path alteration. These phenomena have triggered tremendous interest in mimicking the structure-property relationship of living systems. At a molecular level, shape-shifting related to DNA\(^2\), RNA\(^3\), peptides\(^4\) and small molecules\(^5\) has been reported in recent years owing to their relatively simple and controllable structure. In viral capsids, a single protein fold can be evolved to form multiple oligomeric states with different symmetries\(^6\), but to construct “smart” protein architectures artificially whose structure and shape transformation could be modulated by external stimuli remains challenging.

Proteins, as Nature’s most versatile building blocks, are largely responsible for the complexity of living organisms\(^7\). During evolution, proteins have acquired self-assembly properties to construct a variety of large, complex, and symmetric architectures such as 1D actin filaments\(^8\), 2D bacterial surface layers (S-layers)\(^9\), and 3D light-harvesting protein complexes of phycobilisomes\(^10\), thereby endowing their hosts with plenty of functions. It is well-known that protein-protein interactions (PPIs) at protein interfaces are the chief contributors to construct the diversified protein nanostructures\(^11\)-\(^13\). Following Nature’s inspiration to assemble protein building blocks into exquisite nanostructures, various self-assembly strategies, such as symmetry-directed design\(^14\)-\(^17\), metal-coordination\(^7\),\(^18\),\(^19\), host-guest interactions\(^20\),\(^21\) and the use of bifunctional ligands\(^22\),\(^23\) have been applied mainly to construct one-, two- and three-dimensional hierarchical protein nanostructures. In contrast to the above 1D, 2D and 3D protein architectures, Nature has also evolved a series of protein nanocages to fulfill a wide range of functions such as CO\(_2\) fixation by carboxysomes\(^24\), iron metabolism by ferritin\(^25\), DNA protection by Dps\(^26\), and nucleic acid storage and transport by viral capsids\(^27\). These protein nanocages represent a class of 0D nanomaterials that hold much promise for various applications. For example, by taking advantages of
their well-defined architectures, isolated interiors and high biocompatibility, scientists have subverted the above natural functions of the protein nanocages and explored them as nanocontainers for encapsulation and delivery of bioactive cargo molecules, as bio-templates for preparation of various nanomaterials, and as reaction centers for multienzyme catalysis. Recently, protein shell-like nanocages represent particularly interesting targets for designed assembly, and different strategies such as de novo design, fusion protein, directed evolution, key interface redesign, have been built to construct a variety of artificial protein nanocages that rival the size, property, and functionality of their natural analogues. Despite these advances, rendering conversion between protein nanoarchitectures with different dimensions, especially conversion between 0D protein materials to 1D or higher order analogues, controllable by design of PPIs in the laboratory has largely been inaccessible.

Herein, we introduce a straightforward protein interface redesign strategy that could be used to convert 0D protein nanomaterials into 1D and 2D analogues. We show here that a naturally occurring dimeric protein—Thermotoga maritima ferritin (TmFtn)—self-assembles into 0D 24-meric protein nanocage through interactions between two adjacent dimeric proteins in a head-to-side manner induced by calcium ions (Fig. 1a). In contrast, upon tuning these two contiguous dimeric proteins to interact with each other in a fully side-by-side manner by protein interface redesign, the above formed 0D protein nanocage converts into 1D nanostructure in the presence of calcium ions or PEG (Fig. 1b). Differently, when we adjusted the two adjacent dimeric proteins to interact in a partially side-by-side manner also by protein interface redesign, the assembly state of the dimeric protein molecules converts from the inherent protein nanocage into 2D nanoribbon in the presence of PEG (Fig. 1c). The dimensions of these 1D and 2D protein arrays collectively span nearly the entire nano- and micrometer scale (50 nm to 4.0 µm). In-depth characterization by X-ray crystallography, transmission electron microscopy (TEM), and atomic force microscope (AFM) revealed the structural basis of TmFtn self-assembly in atomic detail. This strategy opens up an avenue for dimensional conversion between 0D materials and 1D or 2D nanoarchitectures.

Results

Natural dimeric TmFtn assembly into 0D protein nanocage.

The four helix bundle structure is widely distributed in Nature, which has been utilized as building blocks to construct a number of proteins and enzymes to perform a considerably wide range of functions such as iron storage by ferritin, DNA protection by Dps protein, copper storage by Csp, electron transfer by cytochrome, ribonucleotide reduced into deoxynucleotide by R2 subunit of ribonucleotide reductase, methane oxidized into methanol by methane monoxygenase, and so on. Four helix bundles are also attractive for synthetic chemists because its interfaces are dominated by side chain and side chain interactions, which can be more tunable than β-strands. In this study, we focus on Thermotoga maritima ferritin (TmFtn), which is a naturally dimeric protein consisting of two antiparallel four helix bundles as shown in Supplementary Figure 1. In contrast, most of known ferritins are usually composed of 24 identical or similar subunits that assemble into a shell-like structure. Interestingly, the dimeric
TmFtn can convert into 24-meric protein nanocage in the presence of salts such as MgCl$_2$. Consistent with these recent findings, our results show that in solution, purified TmFtn molecules (Supplementary Fig. 2) indeed exist as dimers, but they can self-assemble into 24-meric protein nanocage in the presence of Ca$^{2+}$; addition of EDTA causes the formed protein nanocage disassembly into its dimeric form (Fig. 2a), indicative of a reversible process of protein assembly (Fig. 2b).

To gain insight into the role of calcium ions in the formation of the above 24-meric protein nanocage, we optimized protein crystals suitable for X-ray diffraction in the presence of calcium ions by screening a wide range of solution conditions, and subsequently we solved the crystal structure at resolution of 2.2 Å (Supplementary Tables 1 and 2). As expected, the presence of calcium ions facilitates the generation of the 24-meric shell-like protein. Careful analyses of the crystal structure revealed that nearby each C$_3$-C$_4$ interface, one calcium ion is coordinated with two acidic residues (Glu51 and Glu132) and three water molecules (Fig. 2c, d). Further structure analyses showed that a group of acidic amino acid residues are lined along the C$_3$-C$_4$ interface, thereby producing electrostatic repulsion along this interface in the absence of calcium ions. This might be an important reason why TmFtn naturally exists in the form of dimer in solution in the absence of salts. In contrast, the existence of calcium ions nearby the C$_3$-C$_4$ interface can largely eliminate such electrostatic repulsion through their binding with the acidic residues, facilitating the formation of the generation of 0D 24-meric protein nanocage. It is evident that in the crystal structure, each protein nanocage is composed of twelve dimers, and any of two adjacent dimers interact with each other in a head-to-side manner (Fig. 2b, d). It is reasonable to believe that calcium ions facilitate such interaction manner, thereby leading to the 0D shell-like assembly of dimeric TmFtn.

**Engineering dimeric TmFtn for assembly in a fully side-by-side manner into 1D nanomaterials.**

1D filaments. The above head-to-side interaction manner is ubiquitous in nearly all known ferritins from animal, plant, and bacteria. Based on this interesting phenomenon, we envisioned that if the head-to-side interaction manner between two adjacent dimeric TmFtn were adjusted to the fully side-by-side manner by protein interface redesign, the 0D 24-meric TmFtn nanocages would convert into 1D protein nanomaterials. Therefore, we describe a three-step computational method for realizing this idea: (1) Adjacent dimers were manually placed in a fully side-by-side manner as shown in Supplementary Figure 3a; (2) RosettaDock algorithm was carried out to optimize such fully side-by-side interactions into a more complementary protein-protein interface; (3) building new interactions at the new interface to drive self-assembly. After completing the first two steps, predicted surface residues include sites 114 and 147 and sites 114' and 147' that lie across the newly designed interface from each other as appropriate locations for installing new interactions (Supplementary Fig. 3b, c). We made a mutant named FLAL where Asn147 was replaced by aromatic phenylalanine (Phe) residue to create π-π interactions, while Glu112, Glu113 and Lys114 were replaced by three hydrophobic residues Leu, Ala and Leu to decrease electrostatic repulsion and increase the hydrophobic interactions between two adjacent dimers at the same time (Supplementary Fig. 4). Subsequently, this mutant was purified to homogeneity as characterized by SDS and native PAGE (Supplementary Fig. 5). If the above designed interactions worked, one would expect
that FLAL molecules assemble into 1D nanoarchitecture in solution. However, this is not the case. We found that similar to wild-type TmFtn, mutant FLAL also occurs as a dimer in solution under different experimental conditions (pH = 6.0-10.0; the concentration of NaCl = 0-500 mM) (Supplementary Fig. 6). Upon the elaborate inspection of the newly designed protein-protein interface (Supplementary Fig. 7), we found several acidic amino acid residues from D-helix of each subunit, which could impede adjacent FLAL molecules assembly into 1D nanostructure through electrostatic repulsion.

Inspired by the fact that Ca$^{2+}$ has the ability to induce natural TmFtn dimers assembly into 24-meric protein nanocage through coordination with acidic residues (such coordination not only eliminates electrostatic repulsion from acidic residues, but also produces covalent attraction) (Fig. 2c), we deemed that calcium ions might facilitate mutant FLAL to assemble into 1D nanomaterials through a similar mechanism. To confirm this idea, we examined FLAL self-assembly behavior in solution with respect to Ca$^{2+}$ and FLAL concentrations, these two parameters closely associated with Ca$^{2+}$-FLAL interactions. The microscopic characteristics of FLAL self-assembly as a function of the calcium concentration at [FLAL] = 48.0 mM at pH 8.0 were monitored by transmission electron microscopy (TEM). We found that at [Ca$^{2+}$] = 80.0 mM, FLAL molecules self-assemble into two kinds of protein arrays: cuboid-like and linear-shaped protein arrays as shown in Supplementary Figure 8a. Since FLAL concentration used here is significantly lower than that (at least 144.0 mM or 6.0 mg/mL) needed during conventional crystallization procedures, the formation of these two kinds of protein arrays correspond to a protein self-assembly process in solution. Enlargement of the small cuboid-like assembly revealed that 12 of dimeric FLAL molecules first assemble into 24-meric protein nanocages, which further assemble into well-organized 3D superlattices (Supplementary Fig. 8b, c) as we previously reported$^{16,17}$. This result is not surprising for two reasons: a) The above genetic modification for native dimeric TmFtn is only involved in amino acid residues nearby the $C_3$ and $C_4$ interfaces of the salt-mediated 24-meric protein nanocage rather than its $C_3$-$C_4$ interfaces, thus hardly affecting its inherent assembly property; b) substitution of Asn147 with Phe in native TmFtn can also have the possibility to make two adjacent protein nanocages join together along the 4-fold channels through π-π interactions as reported recently$^{16}$, resulting in the formation of such 3D protein superlattices. In contrast, enlargement of the linear-shaped species revealed the generation of 1D filaments by FLAL mutant molecules (Supplementary Fig. 8d). Differently, wild-type TmFtn dimers can only assemble into discrete protein nanocages under identical experimental conditions (Supplementary Fig. 9), these findings approving our design.

To visualize the morphology of the formed 1D filaments in detail, they were observed by TEM in different visual fields. On addition of excess Ca$^{2+}$, 1D filaments were formed rapidly within 30 mins (Fig. 3a). As time goes on, the 1D filaments become longer and more numerous (Figs. 3a-d), suggesting that the formation of the filament is thought to proceed through growth events. The length of these filaments grew up to tens of nanometers with the longest filament as 1.0 µm, and finally the formed 1D filaments were tangled together (Fig. 3d). In addition, the filaments have a tendency to stack together, and thus individual filament is hardly observed. Usually, two or four filaments stacked together, and these filament assemblies exhibit widths on the order of 7 nm or 17 nm, respectively, with the pitch of two filaments
approximately 3 nm (Fig. 3e, f). The width of the individual filament is about 2 nm, which is in good agreement with the thickness of FLAL dimer (Supplementary Fig. 1).

To obtain detailed structural information on the FLAL filament, we tried to crystallize this protein and eventually obtained qualified single crystals suitable for X-ray diffraction. The crystal structure was solved at resolution of 2.1 Å (Supplementary Tables 1 and 2). The packing pattern of FLAL molecules in the crystal is pronouncedly different from wild-type TmFtn. The side view of the crystal structure revealed that FLAL molecules arrange in a repeating, side-by-side fashion to form 1D filaments (Fig. 4a). In contrast, native TmFtn exhibits a different packing pattern in its crystal where twelve of TmFtn dimers assemble into a 24-mer protein cage (Fig. 2). Such difference in protein assembly between TmFtn and mutant FLAL confirms our design. The crystal structure shows that the width of the filament is about 5 nm from the side view and about 2 nm from the top view, these findings being in accordance with the above TEM observation showing that the width of 1D filaments is about 2 nm (Fig. 3).

Further crystal analyses reveal that two adjacent FLAL molecules have opposite orientations, forming a new interface along their D helices, which is exactly consistent with our design. In the structure with resolutions that permit detailed analysis of side-chain configurations, Phe147 and Leu114 side chains at the designed interface adopt suitable conformations to generate π-π stacking interactions and hydrophobic interactions, respectively (Fig. 4b, c), again approving our design at an atomic level. Except for the designed noncovalent interactions, Leu112 and Val119 also produce hydrophobic interactions (Fig. 4d), which is unexpected. Additionally, one calcium ion occurs nearby the newly designed protein interface, which is bound to three acidic residues (Glu108, Asn105 and Asp127) and one water molecule by coordination bonds (Fig. 4e). We believe that the cooperation of these noncovalent interactions and metal coordination along the inter-building-block interfaces promotes the formation of the 1D filaments. In the crystal structure, the formed 1D filaments further arrange in the vertical direction to create 2D protein assemblies (Supplementary Fig. 10a). Notably, these filaments are parallel-displaced from the side view (Supplementary Fig. 10b), resulting in two adjacent filaments being connected by electrostatic attraction. For example, Lys10 and Glu164 from two contiguous filaments are in close proximity, leading to the generation of electrostatic attraction (Supplementary Fig. 10c). This can explain structurally why several filaments stack together as observed by TEM (Fig. 3).

**1D nanorods.** One overarching goal of synthetic biomimetic nanomaterials is to construct hierarchical assemblies that are able to respond to various external stimuli so that the fabricated assemblies can be controlled. Therefore, besides calcium ions, we also investigated the effect of polyethylene glycol (PEG) on the assembly behavior of FLAL molecules because PEG is a hydrophilic nonionic polymer widely used in many biochemical and pharmaceutical applications due to its mild action on the biological activity of cell components. After trying different molecular weights and concentrations of PEG, we finally obtained 1D nanorods formed in the presence of 30% PEG with an average molecular weight of 1500 (PEG1500). As shown detailed in Figure 5a, 1D nanorods of several micrometers in length (up to 4 mm) can be visualized through TEM. Enlargement of the small 1D component showed that FLAL molecules are well-organized and arranged in a linear manner (Fig. 5b). On the basis of TEM characterization of
FLAL nanorods, the corresponding Fast Fourier Transform (FFT) image was recorded (Fig. 5c), and the real map from the inverted FFT in Figure 5d provided an excellent view of the structure of the 1D nanorod, revealing that FLAL molecules tend to assemble side-by-side with the designed driving force to form 1D nanorods.

To determine the structural basis of such 1D nanorods, we set out to grow single crystals large enough for X-ray diffraction analyses in the presence of PEG1500. We solved the crystal structure at 2.6 Å resolution (Supplementary Tables 1 and 2). As expected, the designed interface is responsible for driving self-assembly, but no electron densities of the PEG were observed (Supplementary Fig. 11) most likely due to the high flexibility of PEG and its weak ability to bind to protein. The crystal structure proved remarkably similar to the design model: protein molecules connect together through designed π-π stacking interactions and hydrophobic interactions in a repeating side-by-side pattern with adjacent FLAL molecules antiparallel with each other (Supplementary Fig. 11a, e). Thus, the crystal structure is in good agreement with the above results observed by TEM (Fig. 5), confirming our design in atomic detail.

Further analyses of the crystal structure showed that the assembly process of FLAL molecules can be summarized as follows: Firstly, FLAL molecules pack along the x-axis in such a way mentioned above to form 1D arrays (Supplementary Fig. 11a), such assembly pattern was also observed by TEM shown in Fig. 5b; secondly, just like the stacking of 1D filaments of FLAL molecules in the presence of Ca$^{2+}$, these 1D arrays stack along the y-axis through electrostatic interaction to form 2D arrays (Supplementary Fig. 11b), which is likewise visualized by TEM (Figs. 5b-d); finally, the formed 2D arrays further arrange in the vertical direction to create 3D protein assemblies (Supplementary Fig. 11c, d), and the weak interaction involved in this step mainly comes from their E helix (Supplementary Fig. 11f). Taken together, all these findings demonstrate that our protein interface redesign approach yields assembly of dimeric FLAL protein building blocks into the 1D nanomaterials induced by Ca$^{2+}$ or PEG rather than the inherent Ca$^{2+}$-mediated 24-meric protein shell-like assembly.

**Engineering dimeric TmFtn for assembly in a partially side-by-side manner into 2D nanoribbons.**

The above results demonstrated that tuning the interaction model of two adjacent protein dimers from the head-to-side manner to the fully side-by-side manner facilitates conversion of the inherent 0D protein nanocage into the 1D nanostructures. Bioinspired by the structure of α-keratin protofilaments that are formed from two staggered rows of head-to-tail coiled coils, we wonder what if we rearrange two adjacent protein dimers in a staggered pattern, namely a partially side-by-side pattern. To answer this question, we use the method mentioned above for protein docking as well, but do not need to preset the dimers to search for their possible arrangement. To this end, we simulated the orientations of two TmFtn dimers through global docking using RosettaDock. 1000 independent docking trajectories were carried out, and by combining energy simulation with visual inspection, a model with dimers in a partially side-by-side manner was selected to assist interface redesign (Supplementary Fig. 12a). We found that several hydrophobic residues including Leu107, Val119, Val126 and Val130 are distributed along the new interface (Supplementary Fig. 12b). To effectively utilize these hydrophobic residues, we plan to design...
hydrophobic interactions as the driving forces to construct such partially side-by-side protein assembly. Based on the fact that FLAL mutant contains three more hydrophobic amino acid residues located on its outer surface than wild-type TmFtn, therefore, to build stronger hydrophobic interactions at this newly designed interface (Supplementary Fig. 12c), we chose FLAL mutant instead of TmFtn as starting materials for further engineering.

To enable two dimers to interact with each other in a partially side-by-side fashion, we made a new mutant named FLAL-L where Arg137 in FLAL was replaced with Leu (Supplementary Fig. 4) to increase the area of the hydrophobic patch as shown in Supplementary Figure 12d. To gain insight into the assembly behavior of FLAL-L mutant, we purified it to homogeneity as suggested by SDS-PAGE and native-PAGE (Supplementary Fig. 13). We next investigated the assembly of FLAL-L molecules in the presence of PEG in solution. Upon screening the concentration and molecular weight of PEG, we found that FLAL-L molecules can likewise assemble into 1D filaments in the presence of 15% or 20% PEG1500, as shown in Supplementary Table 3. The 1D filaments with different sizes ranging from 100 nm to several micrometers can be visualized (Fig. 6a, b). Enlargement of the assembly revealed that the FLAL-L molecules are well-organized in a linear fashion (Fig. 6b). To obtain more insights into the structural information, the formed 1D filaments were further characterized by atomic force microscopy (AFM) with an intelligent mode. As shown in Figures 6c, d, the height of the array is 6 ± 0.5 nm, which is comparable with the height of FLAL-L molecules (∼5 nm) in the crystal (Supplementary Fig. 1). To gain deep insight into the above process, we investigated the formation of the 1D filaments as a function of time. Short 1D filaments were rapidly formed at 5 min. With an increase in time, the length and number of the formed 1D filaments increased gradually (Figs. 7a-e). After 15 days, except to the 1D filament, 2D ribbon-like nanomaterials appeared (Fig. 7f). A high-magnification TEM view (Fig. 7g) seized the periodic parallel arrangement of protein arrays, and the real map from invert FFT (Fig. 7h) gave an excellent view on the structure of the 2D nanoribbon, which reveals that two kinds of lines interlaced to form such structure (Fig. 7i). Further examination of the nanoribbons using AFM revealed that the height of the nanoribbon is nearly 10 nm (Supplementary Fig. 14).

To shed light on structural information on the above 1D and 2D nanomaterials, FLAL-L was also crystallized by vapor diffusion, and eventually qualified single crystals suitable for X-ray diffraction were obtained (Supplementary Fig. 15a). We solved the crystal structure at resolution of 2.3 Å (Supplementary Tables 1 and 2). The crystal structure analyses revealed that FLAL-L molecules are perfectly aligned with each other in a partially side-by-side manner to form 1D filaments (Fig. 8a), the top view of which is in good agreement with the 1D filament observed under TEM (Fig. 6a, b). As expected, intermolecular associations between two adjacent FLAL-L molecules are mediated entirely through hydrophobic interactions, as the contact region of adjacent FLAL-L molecules is rich in hydrophobic amino acid residues (Fig. 8b). The engineered Leu114 and Leu137 side chains cooperated with the original V119, L107, V126 and V130 side chains provide the bulk of the hydrophobic core, whereas the engineered Leu112, Ala113 and Phe147 are not involved in such interactions. The crystal structure provides unambiguous evidence to confirm our design in atomic detail. The above formed 1D filaments coalesce in parallel along the y axes through hydrogen bonds (Supplementary Fig. 15b, c), leading to the
generation of a 2D protein layer (Fig. 8c), corresponding to the 2D ribbons observed by TEM (Fig. 7). The arrangement of the 2D protein layer in the crystal perfectly matches the pattern seen in the 2D nanoribbons observed by TEM (Figs. 7f-i). Based on the observation that the thickness of each protein layer is about 2 nm as suggested by the crystal structure (Fig. 8c) and the height of the 2D nanoribbons detected by AFM (Supplementary Fig. 14) is around ~10 nm, it is reasonable to believe that the observed 2D nanoribbons are of five layers. In crystals, the formed 2D protein layers further arrange along the z axes to form 3D protein frameworks (Fig. 8d) through π-π interactions and electrostatic interactions (Supplementary Figs. 15d-f). All these results demonstrate that the hydrophobic interactions between two contiguous dimeric protein building blocks in a partially side-by-side fashion carried out by protein interface redesign facilitate conversion of inherent 0D 24-meric nanocages into 2D nanoribbons through 1D filaments.

Discussion

Particle polymorphism is a popular phenomenon in virus capsids, which possess inherent switches that allow the efficient disassembly and reassembly in response to solution conditions. For example, an icosahedral plant virus, cowpea chlorotic mottle virus (CCMV), can convert into a tubular nanostructure, they need dsDNA to act as a template and take advantage of the nonspecific binding of CP (capsid proteins) and DNA\textsuperscript{50,51}. Simian virus 40 (SV40) represents another interesting example of shape transformation between protein nanostructures with different dimensions. The SV40 capsid is mainly composed of 72 pentamers of VP1. Electron microscopy observations revealed that at pH 5.0, long and tubular structures are formed, whereas at high salt concentrations, small (20 nm) particles predominate\textsuperscript{52}. However, how to convert zero-dimensional protein cage into one- or two-dimensional architecture by a simple, effective method in the laboratory has yet to be explored. The challenge lies in the fact that the construction of 0D protein nanocage belongs to intramolecular assembly at the protein quaternary level, while the formation of the 1D or 2D nanomaterials by using the same building blocks are mainly associated with intermolecular assembly. Therefore, switching the intramolecular assembly of protein building blocks to its intermolecular analogue is the key to realize 0D®1D nanomaterials. We believe that spatially directed assembly of protein building blocks might be a solution to solve such problems. To confirm this idea, we have built here a protein interface redesign approach that is able to manipulate the directed assembly of protein building blocks, resulting in conversion of 0D protein nanocage into 1D or 2D protein arrays with nano and microscale long-range order. This approach enables shape transformation between protein nanoarchitectures with different dimensions to become possible in the laboratory by genetic modification. It has allowed us not only to construct the protein-protein interface, but also to adjust the geometric arrangement of protein building blocks that control protein assembly with different dimensions.

It is noteworthy that there is only one amino acid difference in amino acid sequence between FLAL and FLAL-L molecules, but their assembly behavior is markedly different from each other. FLAL molecules self-assemble into the 1D protein arrays, whereas FLAL-L has the ability to assemble into the 2D
nanoribbons. The large difference in self-assembly behavior between these two dimeric protein molecules reflects the importance of amino acid residue Arg137 in FLAL, suggesting that Arg137 could act as a switch to control conversion of 1D nanomaterials to 2D analogues. The occurrence of assemblies (1D and 2D protein arrays) with the dimeric protein as building blocks recalls among natural proteins the case of actin filaments and S-layers in terms of their dimensions and structural uniformity. Thus, our reported construction of 1D and 2D protein arrays from the dimeric protein as building blocks provides a model to study the assembly mechanism of natural protein architectures.

Controlling self-assembly is critical to the advancement of nanotechnology. Nowadays, scientists are able to accurately control over the protein self-assembly behaviors to construct multifarious supramolecular structures through rational design of PPIs. Though a variety of intricate protein nanostructures such as 0D polyhedral cages, 1D filaments/tubules, 2D nanosheets, and 3D crystalline frameworks have been created\textsuperscript{18-23, 35-40}, rendering directed assembly of protein building blocks into the custom-tailored nanoarchitectures remains challenging. Our reported protein engineering approach could tune the inherent head-to-side interaction manner of two adjacent dimeric protein building blocks to the fully or partially side-by-side manner by designing new protein interfaces, yielding directed assembly of the building blocks, thereby facilitating conversion of 0D protein nanocage into 1D or 2D nanomaterials. PPIs at the designed protein interfaces are mainly contributed from the cooperation of the hydrophobic interactions and aromatic π–π interactions, which effectively stabilize the newly designed protein interface. The above protein engineering approach that focuses on adjustment for the geometric arrangement of protein building blocks is conceptually and operationally simple. It is interesting that the head-to-side interaction manner between two contiguous dimeric protein building blocks occurs not only in all known ferritins, but also in other cage-like protein such as Dps, where six dimeric protein building blocks also assemble into a shell-like structure\textsuperscript{26}. Therefore, our engineering approach should in principle be applicable to some other protein architectures. This would produce a variety of protein nanomaterials with different geometries.

Methods

Protein preparation. The gene encoding \textit{Thermotoga maritima} ferritin (TmFtn) were synthesized by Synbio Technologies, which have been inserted into the plasmid pET-3a. Mutagenesis of the TmFtn cDNA was performed with the fast site-directed mutagenesis kit (TIANGEN Biotech Co., Ltd.). Polymerase chain reaction amplification was carried out using the pET-3a plasmid with the TmFtn gene as a template. Plasmids sequences were verified by DNA sequencing. And TmFtn as well as mutants were purified as previously described\textsuperscript{46}. Briefly, the plasmids corresponding to FLAL and FLAL-L were transformed into BL21 (DE3) \textit{E. coli} cells, respectively, and then cultured at 37 °C in 1-liter of LB media containing 100 µg/mL ampicillin. After the cell density reached an absorbance as 0.6 at 600 nm, protein expression was likewise induced with 200 mM IPTG for 10 h at 37 °C. Cells were harvested by centrifugation (10000 rpm) and the precipitate was re-suspended in 50 mM Tris–HCl (pH 8.0), followed by sonication. The supernatant was collected from lysed cell samples and subjected to heating at 90 °C for 10 min. Then
thermal-treated supernatant was collected after centrifugation at 10000 rpm for 30 min, and passed through a membrane filter. Finally, the protein solution was applied to an ion-exchange column (DEAE Sepharose Fast Flow, GE Healthcare), followed by gradient elution with 0-1.0 M NaCl. The purified protein was then dialyzed against 50 mM Tris-HCl (pH 8.0) at 4 °C to exclude NaCl from the solution, and protein concentrations were determined according to the Lowry method with bovine serum albumin (BSA) as standard.

Docking Simulations. 1000 independent docking trajectories were carried out using RosettaDock53. The two TmFtn or FLAL subunits that form the newly designed interface were used as the starting structure for the docking simulations. One of the monomers was randomly spun along the axis connecting the centers of mass of both partners and the same monomer was also allowed to search a space of up to 3 Å normal to that axis, 8 Å in the plane perpendicular to the axis, and with up to an 8° tilt from the axis and an 8° additional spin around the axis.

Polyacrylamide gel electrophoresis (PAGE). The purity and molecular weight of protein samples was estimated by polyacrylamide gel electrophoresis. Gel electrophoresis under denaturing conditions was carried out using a 15% polyacrylamide-SDS gel as reported by Laemmli54, and samples need to be heated in a water bath for 5 min. For native PAGE, a 4-20% polyacrylamide gradient gel was used and run at 5 mA for 10 h at 4 °C. Gels were stained with Coomassie brilliant blue R250.

High-Resolution Gel Filtration Chromatography Analyses. High-resolution gel filtration chromatography analyses were performed using an ÄKTA pure system coupled to a Superdex 200 Increase column (GE Healthcare) in buffer (50 mM Tris, 100 mM NaCl, pH = 8.0) with a flow rate of 0.5 mL/min.

Transmission electron microscopy (TEM) imaging. Protein samples (10 μL) were deposited on carbon-coated copper grids and excess solution was removed with filter paper after a 2-min incubation. Then protein samples were stained using 2% uranyl acetate for 5 min. Transmission electron micrographs were obtained at 80 kV through a Hitachi H-7650 transmission electron microscope.

Atomic force microscope (AFM) measurements. For AFM sample preparation, protein samples (10 μL) were pipetted on freshly cleaved mica (Beijing Zhongxingbairui Technology Co., Ltd.) and dried at room temperature. The AFM images were collected using a Nanoman VS (Bruker) with tapping mode at a resolution of 256 lines per image and a scan rate of 1 Hz. The AFM images were processed using NanoScope Analysis.

Crystallization, data collection and structure determination. Purified proteins were concentrated to 6 mg/mL in a buffer consisting of 20 mM Tris–HCl at pH 8.0, and crystals were obtained using the hanging drop vapor diffusion method under different conditions which were shown in Supplementary Tables 1. X-ray diffraction data were collected at Shanghai Synchrotron Radiation Facility (SSRF) (BL17U and BL19U) with merging and scaling by HKL-3000 software55. Data processing statistics are shown in Supplementary Tables 2. The structures were determined by molecular replacement using the Molrep program in CCP4 using the structure of *Thermotoga maritima* ferritin (TmFtn) (PDB code 1VLG) as a
search model. Structure refinement was conducted using the Refmac5 program and PHENIX software\textsuperscript{56}. The structure was rebuilt using COOT\textsuperscript{57}, which made the model manually adjusted. All Figures of the resulting structures were produced using the PyMOL\textsuperscript{58} program and UCSF\textsuperscript{59} Chimera package.

**Declarations**

**Data availability**

Coordinates and structure factors are deposited in the Protein Data Bank under the accession PDB IDs: 7DYA (24-meric TmFtn), 7DY8 (FLAL\textsubscript{r}), 7DY9 (FLAL\textsubscript{r}), and 7DYB (FLAL-L). Other data are available from the corresponding authors upon reasonable request. The source data underlying Fig. 2b and Supplementary Figs.1, 4, and 11 are provided as a Source Data file.

**Acknowledgements**

This work was supported by the National Natural Science Foundation of China (Nos. 31972018 and 31730069). The Shanghai Synchrotron Radiation Facility (SSRF) is especially acknowledged for beam time. We thank the staffs from BL17U1/BL18U1/19U1 beamline of the National Center for Protein Sciences Shanghai (NCPSS) at Shanghai Synchrotron Radiation Facility for assistance during data collection.

**Author contributions**

G.Z. and T.Z. conceived and directed the project and wrote the paper. X.Z. designed and performed experiments, analyzed data, and co-wrote the paper. B.Z. performed experiments. Y.L. performed the crystal data collection. J.Z., C.L. performed the experiments and co-wrote the paper.

**Competing interests**

The authors declare no competing interests.

**References**

1. Lay, C. L., Lee, M. R., Lee, H. K., Phang, I. Y. & Ling, X. Y. Transformative two-dimensional array configurations by geometrical shape-shifting protein microstructures. *ACS Nano* 10, 9708-9717 (2015).
2. Rothemund, Paul, W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* 440, 297-302 (2006).
3. Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* 333, 470-474 (2011).
4. Aggeli, A. et al. Responsive gels formed by the spontaneous self-assembly of peptides into polymeric \( \beta \)-sheet tapes. *Nature* 386, 259-262 (1997).
5. Waller, P. J. et al. Chemical conversion of linkages in covalent organic frameworks. *J. Am. Chem. Soc.* **48**, 15519-15522 (2016).

6. Zhao, X., Fox, J. M., Olson, N. H., Baker, T. S. & Young, M. J. In vitro assembly of cowpea chlorotic mottle virus from coat protein expressed in *Escherichia coli* and in vitro-transcribed viral cDNA. *Virology* **207**, 486-494 (1995).

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

8. Reisler, E. & Egelman, E. H. Actin Structure and Function: What We Still do not Understand. *J. Biol. Chem.* **282**, 36133-36137 (2007).

9. Chung, S., Shin, S. H., Bertozzi, C. R. & De Yoreo, J. J. Self-catalyzed growth of S layers via an amorphous-to-crystalline transition limited by folding kinetics. *Proc. Natl. Acad. Sci. USA* **107**, 16536-16541 (2010).

10. Ma, J. et al. Structural basis of energy transfer in *Porphyridium purpureum* phycobilisome. *Nature* **579**, 146-151 (2020).

11. Kortemme, T. & Baker, D. Computational design of protein-protein interactions. *Curr. Opin. Chem. Bio.* **8**, 91-97 (2004).

12. Schreiber, G. & Fleishman, S. J. Computational design of protein–protein interactions. *Curr. Opin. Chem. Bio.* **23**, 903-910 (2013).

13. Jones, S. & Thornton, J. M. Principles of protein-protein interactions. *Proc. Natl. Acad. Sci. USA* **93**, 13-20 (1996).

14. Padilla, J. E., Colovos, C. & Yeates, T. O. Nanohedra: Using Symmetry to Design Self Assembling Protein Cages, Layers, Crystals, and Filaments. *Proc. Natl. Acad. Sci. USA* **98**, 2217-2221 (2001).

15. Sinclair, J. C., Davies, K. M., Venien-Bryan, C. & Noble, M. E. Generation of Protein Lattices by Fusing Proteins with Matching Rotational Symmetry. *Nat. Nanotechnol.* **6**, 558-562 (2011).

16. Zhou, K. et al. On-Axis Alignment of Protein Nanocage Assemblies from 2D to 3D through the Aromatic Stacking Interactions of Amino Acid Residues. *ACS Nano* **12**, 11323-11332 (2018).

17. Zheng, B., Zhou, K., Zhang, T., Lv, C. & Zhao, G. Designed Two- and Three-Dimensional Protein Nanocage Networks Driven by Hydrophobic Interactions Contributed by Amyloidogenic Motifs. *Nano Lett.* **19**, 4023-4028 (2019).

18. Bailey, J. B., Zhang, L., Chiong, J. A., Ahn, S. & Tezcan, F. A. Synthetic modularity of protein-metal-organic frameworks. *J. Am. Chem. Soc.* **139**, 8160-8166 (2017).

19. Gu, C. et al. Structural insight into binary protein metal-organic frameworks with ferritin nanocages as linkers and nickel clusters as nodes. *Chem. Eur. J.* **26**, 3016-3021 (2020).

20. Kitagishi, H. et al. Supramolecular hemoprotein linear assembly by successive interprotein heme-heme pocket interactions. *J. Am. Chem. Soc.* **129**, 10326-10327 (2007).

21. Carlson, J. C. T. et al. Chemically controlled self-assembly of protein nanorings. *J. Am. Chem. Soc.* **128**, 7630-7638 (2006).
22. Yang, R. et al. 2D Square Arrays of Protein Nanocages through Channel-Directed Electrostatic Interactions with Poly(ω, L-Lysine). *Chem. Commun.* **50**, 2879-2882 (2014).

23. Yang, G. et al. Highly Ordered Self-Assembly of Native Proteins into 1D, 2D, and 3D Structures Modulated by the Tether Length of Assembly-Inducing Ligands. *Angew. Chem., Int. Ed.* **56**, 10691-10695 (2017).

24. Tanaka, S. et al. Atomic-Level Models of the Bacterial Carboxysome Shell. *Science* **319**, 1083-1086 (2008).

25. Arosio, P., Ingrassia, R. & Cavadini, P. Ferritins: a family of molecules for iron storage, antioxidation and more. *Biochim. Biophys. Acta, Gen. Subj.* **1790**, 589-599 (2009).

26. Grant, R., Filman, D., Finkel, S., Kolter, R. & Hogle, J. The Crystal Structure of Dps, a Ferritin Homolog that Binds and Protects DNA. *Nat. Struct. Biol.* **5**, 294-303 (1998).

27. Douglas, T. & Young, M. Viruses: Making Friends with Old Foes. *Science* **312**, 873-875 (2006).

28. Liang, M. et al. H-ferritin-nanocaged doxorubicin nanoparticles specifically target and kill tumors with a single-dose injection. *Proc. Natl. Acad. Sci. USA* **111**, 14900-14905 (2014).

29. Wen, A. M. & Steinmetz, N. F. Design of Virus-Based Nanomaterials for Medicine, Biotechnology, and Energy. *Chem. Soc. Rev.* **45**, 4074-4126 (2016).

30. Jordan, P. C. et al. Self-Assembling Biomolecular Catalysts for Hydrogen Production. *Nat. Chem.* **8**, 179-185 (2016).

31. King, N. P. et al. Computational Design of Self-Assembling Protein Nanomaterials with Atomic Level Accuracy. *Science* **336**, 1171-1174 (2012).

32. King, N. P. et al. Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **510**, 103-108 (2014).

33. Hsia, Y. et al. Design of a hyperstable 60-subunit protein icosahedron. *Nature* **535**, 136-139 (2016).

34. Lai, Y. T. et al. Structure of a designed protein nanocage that self-assembles into a highly porous cube. *Nat. Chem.* **6**, 1065-1071 (2014).

35. Cannon, K. A., Nguyen, V. N., Morgan, C. & Yeates, T. O. Design and Characterization of an Icosahedral Protein Nanocage Formed by a Double-Fusion Protein Containing Three Distinct Symmetry Elements. *ACS Synth. Biol.* **9**, 517-524 (2020).

36. Woersdorfer, B., Woycechowsky, K. J. & Hilvert, D. Directed evolution of a protein container. *Science* **331**, 589-592 (2011).

37. Worsdorfer, B., Pianowski, Z. & Hilvert, D. Efficient in vitro encapsulation of protein cargo by an engineered protein container. *J. Am. Chem. Soc.* **134**, 909-911 (2012).

38. Butterfield, G. L. et al. Evolution of a designed protein assembly encapsulating its own RNA genome. *Nature* **552**, 415-420 (2017).

39. Zhang, S. et al. "Silent" amino acid residues at key subunit interfaces regulate the geometry of protein nanocages. *ACS Nano* **10**, 10382-10388 (2016).
40. Zhang, S. et al. Conversion of the native 24-mer ferritin nanocage into its non-native 16-mer analogue by insertion of extra amino acid residues. *Angew. Chem. Int. Ed.* **55**, 16064-16070 (2016).

41. Vita, N. et al. A four-helix bundle stores copper for methane oxidation. *Nature* **525**, 140-143 (2015).

42. Wittung-Stafshede, P., Lee, J. C., Winkler, J. R. & Gray, H. B. Cytochrome $b_{562}$ folding triggered by electron transfer: approaching the speed limit for formation of a four-helix-bundle protein. *Proc. Natl. Acad. Sci. USA* **96**, 6587-6590 (1999).

43. Kolberg, M., Strand, K. R., Graff, P. & Andersson, K. K. Structure, function, and mechanism of ribonucleotide reductases. *Biochim. Biophys. Acta, Proteins Proteomics* **1699**, 1-34 (2004).

44. Lee, S. J., McCormick, M. S., Lippard, S. J. & Cho, U. Control of substrate access to the active site in methane monoxygenase. *Nature* **494**, 380-384 (2013).

45. Ceci, P., Forte, E., Di Cecca, G., Formara, M. & Chiancone, E. The Characterization of *Thermotoga Maritima* Ferritin Reveals an Unusual Subunit Dissociation Behavior and Efficient DNA Protection from Iron-Mediated Oxidative Stress. *Extremophiles* **15**, 431-439 (2011).

46. Chakraborti, S. et al. Three-Dimensional Protein Cage Array Capable of Active Enzyme Capture and Artificial Chaperone Activity. *Nano Lett.* **19**, 3918-3924 (2019).

47. Wang, Y. & Annunziata, O. Comparison between protein-polyethylene glycol (peg) interactions and the effect of peg on protein-protein interactions using the liquid-liquid phase transition. *J. Phys. Chem. B.* **111**, 1222-1230 (2007).

48. Cattani, G., Vogeley, L. & Crowley, P. B. Structure of a pegylated protein reveals a highly porous double-helical assembly. *Nat. Chem.* **7**, 823-828 (2015).

49. Mckittrick, J. et al. The structure, functions, and mechanical properties of keratin. *JOM* **64**, 449-468 (2012).

50. Mukherjee, S., Pfeifer, C. M., Johnson, J. M., Liu, J. & Zlotnick, A. Redirecting the coat protein of a spherical virus to assemble into tubular nanostructures. *J. Am. Chem. Soc.* **128**, 2538-2539 (2006).

51. Xu, Y. et al. DNA-templated CMV viral capsid proteins assembly into nanotubes. *Chem. Commun.* 49-51 (2008).

52. Kanesashi, S. et al. Simian virus 40 vp1 capsid protein forms polymorphic assemblies in vitro. *J. Gen. Virol.* **84**, 1899-1905 (2003).

53. Gray, J. J. et al. Protein-protein docking with simultaneous optimization of rigid-body displacement and side-chain conformations. *J. Mol. Biol.* **331**, 281-299 (2003).

54. Laemmli, U. K. Determination of protein molecular weight in polyacrylamide gels. *Nature* **227**, 680-685 (1970).

55. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307 (1997).

56. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr.* **66**, 213-221 (2010).
57. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr D Biol Crystallogr.* **66**, 486-501 (2010).

58. DeLano, W. L. The PyMOL Molecular Graphics System. *DeLano Scientific; San Carlos, CA, USA.* 2002.

59. Pettersen, E. F. et al. A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **25**, 1605-1612 (2004).