nature chemistry

ARTICLES

PUBLISHED ONLINE: 4 MARCH 2012 | DOI: 10.1038/NCHEM.1290

Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays

Jeffrey D. Brodin[1], X. I. Ambroggio[2], Chunyan Tang[1], Kristin N. Parent[1], Timothy S. Baker[1,3] and F. Akif Tezcan[1]*

Proteins represent the most sophisticated building blocks available to an organism and to the laboratory chemist. Yet, in contrast to nearly all other types of molecular building blocks, the designed self-assembly of proteins has largely been inaccessible because of the chemical and structural heterogeneity of protein surfaces. To circumvent the challenge of programming extensive non-covalent interactions to control protein self-assembly, we have previously exploited the directionality and strength of metal coordination interactions to guide the formation of closed, homoligomeric protein assemblies. Here, we extend this strategy to the generation of periodic protein arrays. We show that a monomeric protein with properly oriented coordination motifs on its surface can arrange, on metal binding, into one-dimensional nanotubes and two- or three-dimensional crystalline arrays with dimensions that collectively span nearly the entire nano- and micrometre scale. The assembly of these arrays is tuned predictably by external stimuli, such as metal concentration and pH.

Proteins are nature's most versatile building blocks, programmed at the genetic level to perform myriad functions and largely responsible for the complexity of an organism. The exquisite, error-free self-assembly of proteins into ordered yet dynamic nano- and microscale architectures, such as one-dimensional (1D) microtubules, 2D bacterial surface layers (S-layers) and 3D virus capsids is a cornerstone of life[1], and has served as a major inspiration for synthetic supramolecular chemistry[2] and nanoscience[3]. In fact, such self-assembled protein arrays themselves have been exploited widely as well-ordered and genetically or chemically modifiable structural templates in numerous nanotechnological applications[4–6]. Significantly, crystalline protein arrays form the basis of diffraction-based methods for structure determination, and their acquisition generally represents the rate-limiting step in this process. Thus, a bottom-up strategy for arranging proteins into periodic structures would have wide-ranging applications in crystallography and in nanotechnology, where it would allow access to novel biological polymers that incorporate the wide range of chemical functionalities afforded by the proteins.

In contrast to simpler building blocks, such as DNA[7,8], RNA[9,10] and peptides[11,12], the designed self-assembly of proteins has largely been inaccessible, plagued by the chemical heterogeneity and size of the molecular surfaces required for protein–protein interactions[13]. Previous efforts in designed protein self-assembly had limited success: generally, they required building blocks with high pre-existing symmetry, allowed little or no external control over self-assembly and yielded static protein architectures that did not display long-range order beyond a few hundred nanometres[14–18].

Inspired by efforts in supramolecular coordination chemistry[19–21], we adopted the strategy of using directional metal coordination bonds to control protein self-assembly, which led to the facile generation of discrete superprotein architectures[22,23]. In this study, we extend further into the protein realm the concepts of supramolecular coordination chemistry to create extended 1D, 2D and 3D molecular arrays. The premise of our approach is that metal-bonding interactions can capture all the salient features of protein–protein interactions (stability, specificity, directionality, symmetry and reversibility) on a much smaller surface than is needed by non-covalent interactions, and thus require 'little design'. The reversibility of metal coordination allows for the self-healing necessary to produce assemblies with long-range order. Importantly, chemical control and tunability are inherent in this approach: metal-directed protein self-assembly is, by default, dependent on external stimuli (that is, metal binding or any stimulus that affects metal coordination) and all the defining features of protein self-assembly listed above can be dictated by the choice of the metal ion and the coordination motif.

Using this approach, we show here that a monomeric protein—cytochrome (cyt) *cb* 562—with properly oriented metal coordinating motifs on its surface can self-assemble with Zn^{2+} into 1D helical nanotubes, and 2D or 3D crystalline arrays. The dimensions of these arrays collectively span nearly the entire nano- and micrometre scale (5 nm to 200 μm). In-depth characterization by X-ray crystallography, cryoelectron microscopy (cryo-EM) and image reconstruction revealed the structural basis of cyt *cb*-562 self-assembly in atomic detail, and guided the design of a variant whose 3D assembly was further controlled by surface-attached functional groups.

Results and discussion

Design of a self-assembling building block, Rosetta interface designed cytochrome 3 (RIDC3). Previously, we focused on the use of the four-helix-bundle haem protein cyt *cb*-562 as a building block, because it is stable, monomeric at millimolar concentrations and has a compact, cylindrical shape[22]. We showed that the cyt *cb*-562 variant MBPC1, with just two metal-chelating bis-histidine motifs (His59/His63 and His73/His77) on its surface, assembled into discrete oligomeric states with

[1]Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093, USA, [2]Rosetta Design Group LLC, Fairfax, Virginia 22030, USA, [3]Division of Biological Sciences, University of California, San Diego, La Jolla, California 92093, USA. *e-mail: tezcan@ucsd.edu

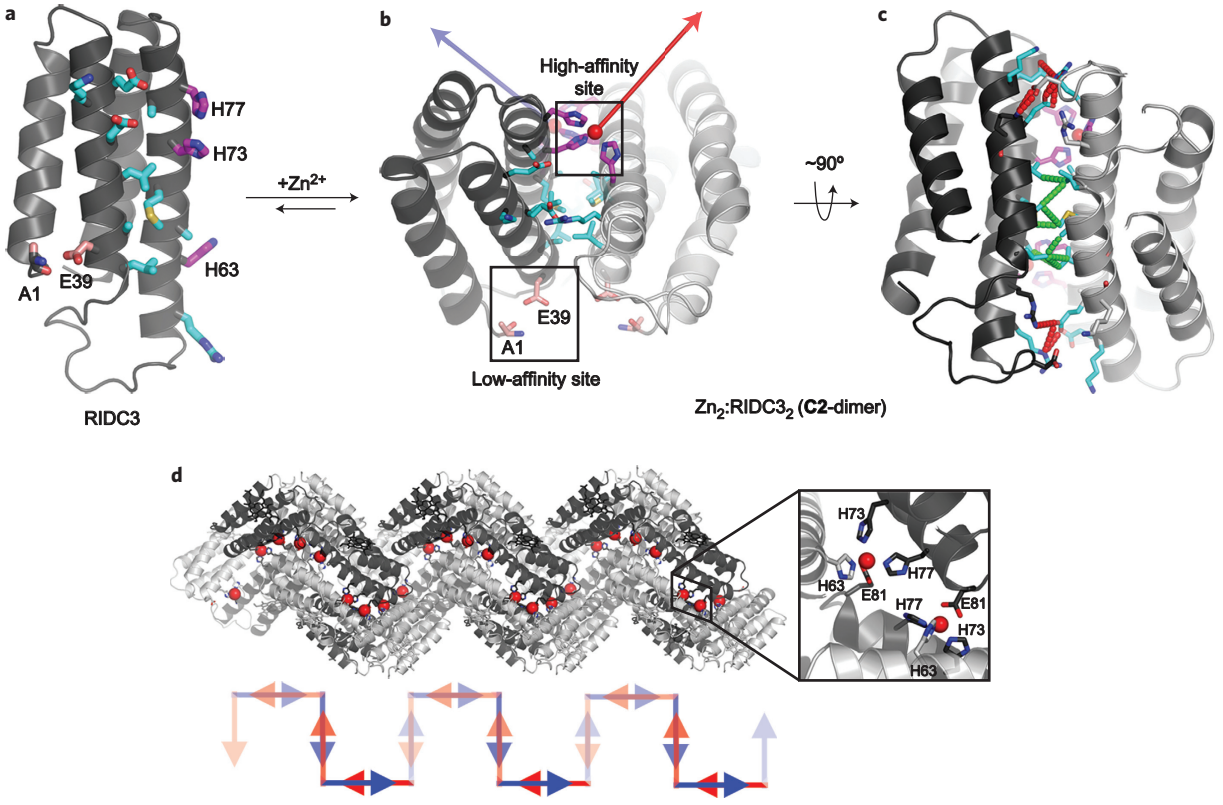

Figure 1 | Design and initial crystallographic characterization of RIDC3. a, Cartoon representation of the RIDC3 monomer. Surface residues predicted by RosettaDesign to stabilize the Zn-induced RIDC3 dimer, the high-affinity and the low-affinity Zn binding sites are highlighted as cyan, magenta and pink sticks, respectively. **b**, Zn (red spheres) binding by RIDC3 produces a C_2-symmetric dimer (Zn_2:$RIDC3_2$ or C2-dimer) with two orthogonal coordination vectors (red and blue arrows) that originate from the two high-affinity Zn sites. The C2-dimer geometry is derived from the structure of the Zn-mediated tetramer of MBPC1 (Zn_4:$MBPC1_4$), as illustrated in Supplementary Fig. S1. **c**, Rotated view of the C2-dimer showing the crystallographically observed hydrophobic (green segments) and polar (red segments) interactions in the dimeric interface. **d**, Pairwise Zn coordination interactions between neighbouring C2-dimers that involve the high-affinity sites and E81 (inset) in the lattice leads to a helical 1D chain. Shown below is a cartoon representation of the C2-dimer units in this chain, illustrated as pairs of blue and red arrows that approximately represent the direction of Zn coordination vectors.

stoichiometric Zn^{2+}, Cu^{2+} or Ni^{2+} binding[24,25]. When Zn^{2+} was added in a molar excess greater than 1.5-fold, MBPC1 formed large, heterogeneous aggregates, which we attributed to multiple and indiscriminate modes of Zn-mediated interprotein coordination. Combining the symmetry- and directional bonding-based self-assembly strategies of others[17,19,20], we posited that, if two or more selective Zn coordination sites on the MBPC1 surface could be stabilized in appropriate orientations relative to each other, it should be possible to dictate the self-assembly of the protein into regular, extended arrays through metal coordination.

Here we accomplished this using RIDC3 (Fig. 1a), which incorporates ten Rosetta-prescribed[26] surface mutations (Lys27Glu, Asp28Lys, Thr31Glu, Arg34Leu, Leu38Ala, Asn41Leu, His59Arg, Asp66Ala, Val69Met and Leu76Ala) into MBPC1 (see Supplementary Information for computational design). The mutations stabilize a C_2-symmetric dimer (Zn_2:$RIDC3_2$, the '**C2**-dimer') that forms on Zn^{2+} binding by His73 and His77 from one monomer and His63 from the other (the 'high-affinity site', Fig. 1b). This particular Zn coordination mode and the resulting **C2**-dimer geometry were derived from the structure of the Zn-directed tetramer of MBPC1 (ref. 24) (Supplementary Fig. S1). Each RIDC3 monomer also possesses a weaker metal binding site (the 'low-affinity site') formed by the *N*-terminal alanine and Asp39, which we have seen in various cyt *cb*-562 structures (Protein Data Bank identifiers: 3QVY, 3QVZ and 3DE9). **C2**-dimers contain three key design elements: (1) The tricoordinate high- and low-affinity sites provide stable and selective, yet labile, Zn^{2+} binding motifs that allow metal-directed self-assembly to proceed under thermodynamic control. (2) Each metal-binding motif leaves one Zn coordination site open for binding to another **C2**-dimer. Coordination vectors that originate from the high-affinity sites are nearly orthogonal to one another, which suggests that they could promote the self-assembly of **C2**-dimers in two perpendicular directions. (3) Small, hydrophobic residues are incorporated into the centre of the dimeric interface for enhanced stability, and polar residues are incorporated onto the periphery for geometric specificity (Fig. 1c), which generates a pattern commonly observed in transient, but specific protein–protein interactions[27].

Sedimentation velocity experiments showed that RIDC3 was monomeric, but that it converted fully into a dimeric form on binding equimolar Zn^{2+} at 5 μM concentration (Supplementary Fig. S2). The 2.3 Å crystal structure of the Zn-mediated dimer (Zn_2:$RIDC3_2$), obtained using polyethylene glycol (PEG)-precipitated crystals, confirmed that it possessed the desired **C2**-dimer geometry. The overall root mean squared deviation in the α-C positions between the crystallographically determined and Rosetta-predicted **C2**-dimer conformations is <0.6 Å (Supplementary Fig. S3). In the lattice, **C2**-dimers formed 1D coordination polymers in which each Zn bound to a high-affinity site is ligated to E81 from a neighbouring dimer (Fig. 1d). This arrangement provided the first evidence that **C2**-dimers might, indeed, self-assemble into extended arrays in solution, particularly in the presence of excess Zn^{2+}.

Zn-directed self-assembly of RIDC3 in solution. We examined RIDC3 self-assembly in solution with respect to Zn and RIDC3 concentrations and pH, the parameters expected to govern

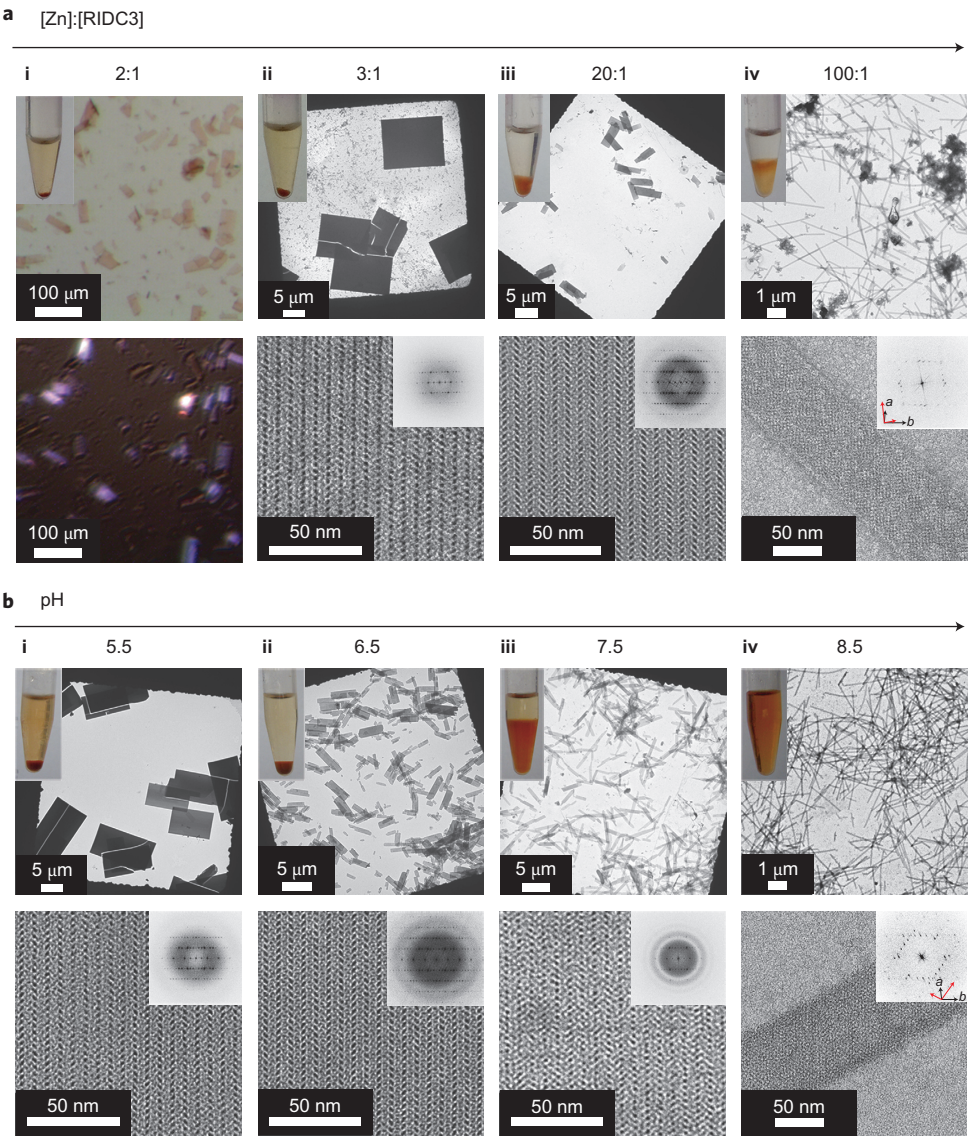

Figure 2 | Zn-induced RIDC3 self-assembly in solution characterized by light and electron microscopy. a, Dependence of the morphology of self-assembled RIDC3 arrays on the [Zn]:[RIDC3] ratio at [RIDC3] = 50 μM and pH = 5.5. **b**, Dependence of the morphology of self-assembled RIDC3 arrays on pH at [RIDC3] = 100 μM and [Zn] = 300 μM. These microscopy images illustrate a trend of decreasing array size with either an increase in the [Zn]:[RIDC3] ratio (**a**) or an increase in pH (**b**). For [Zn]:[RIDC3] = 2 at pH = 5.5 (column (**i**) in (**a**)), only macroscopic crystalline arrays were observed and were viewed by light microscopy with (top) and without (bottom) polarizers. The birefringence of the RIDC3 arrays is indicative of crystalline order. For all other experiments, the top and the bottom rows show low- and high-magnification TEM images, respectively, with the computed Fourier transforms of the high-magnification images shown as insets. Discrete reflections in the Fourier transforms with identical spacings under all conditions demonstrate that all the RIDC3 arrays are crystalline and contain the same underlying 2D lattice. The overlapping lattice axes in the tubular structures (column (**iv**)) are highlighted with black and red arrows in inset Fourier transform.

Zn–RIDC3 interactions. The macroscopic and microscopic characteristics of RIDC3 self-assembly as a function of the [Zn]:[RIDC3] ratio at [RIDC3] = 50 μM and pH = 5.5 were monitored by light and electron microscopy (EM) (Fig. 2a). Importantly, RIDC3 concentrations used here are significantly lower than those (>2 mM) needed during conventional cyt *cb*-562 crystallization procedures that require precipitants such as PEG. On Zn addition, RIDC3 formed precipitates rapidly, which over the course of several days evolved into planar, rectangular crystalline arrays (Supplementary Fig. S4). We refer to these as 'arrays' rather than 'crystals' to emphasize that they are obtained by bottom-up self-assembly. These 2D arrays decreased in size as the [Zn]:[RIDC3] ratio was increased (Fig. 2a). The lattice parameters for all 2D arrays were identical within experimental error ($a = 37 \pm 1$ Å, $b = 137 \pm 1$ Å, $\alpha = 90 \pm 1^{\circ}$), which indicates that they share the same underlying 2D arrangement of RIDC3 molecules. At 100-fold excess Zn, the arrays converted into uniformly thin, unbranched tubes up to 15 μm long with widths distributed around 80 nm (Fig. 2a and Supplementary Fig. S5). Fourier transforms computed from transmission electron microscopy (TEM) images of these tubes indicated the presence of two mirror-symmetric lattices (interlattice angle ~85°). Such overlapping lattices are typical of hollow, helical tubes flattened during specimen preparation for negative-stain TEM[28].

The same morphological trend was observed at a higher [RIDC3] (400 μM), except that the transitions occurred at lower [Zn]:[RIDC3] ratios (Supplementary Fig. S6). Similarly, when pH was raised from 5.5 to 8.5 at a fixed [Zn]:[RIDC3] ratio, the morphology of the assembled arrays gradually changed from 2D plates into helical nanotubes (Fig. 2b). Under these conditions,

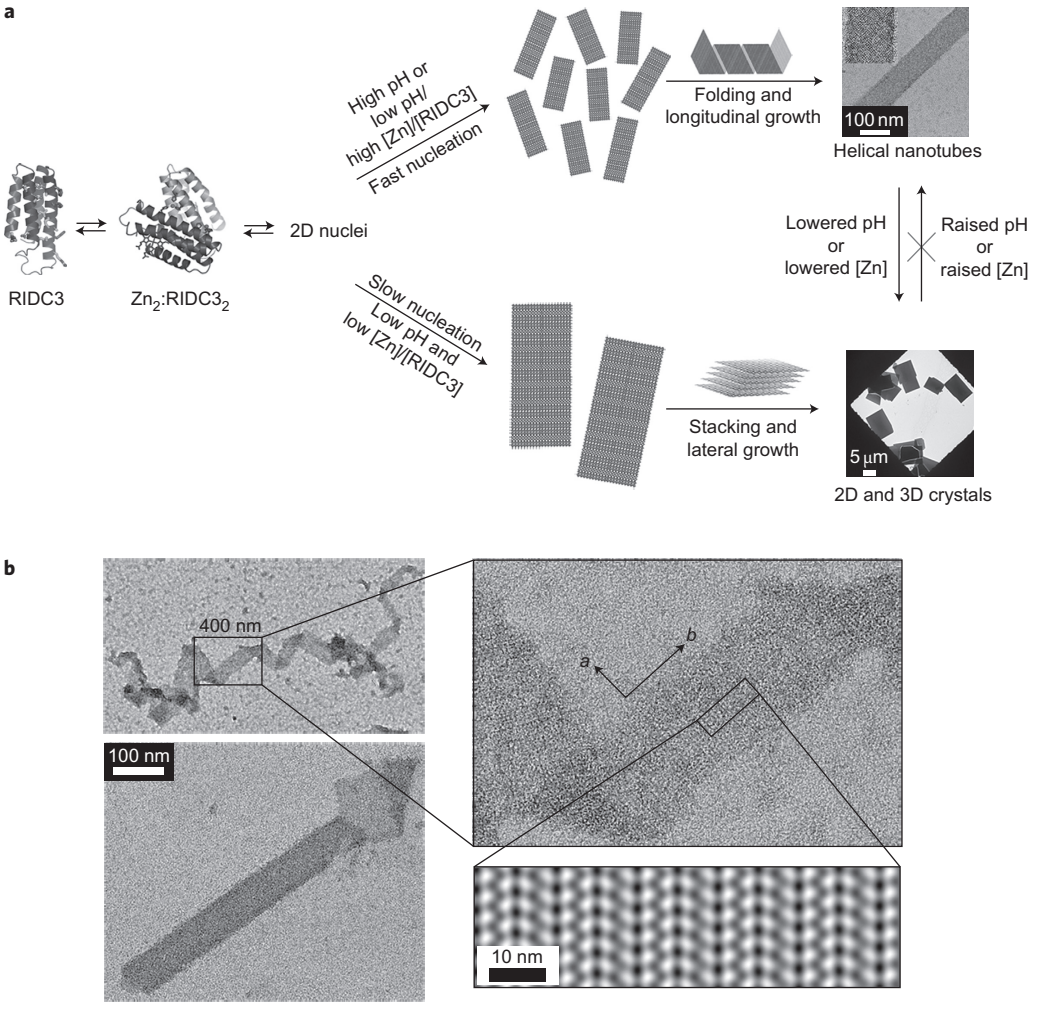

Figure 3 | Model for RIDC3 self-assembly. a, Hypothetical model for Zn-mediated RIDC3 self-assembly under fast and slow nucleation conditions. Fast nucleation/growth conditions (high pH or high [Zn]:[RIDC3] ratio) promote the formation of many 2D nuclei, which 'roll up' into helical nanotubes. Under slow nucleation conditions (low pH and low [Zn]:[RIDC3]), a smaller number of large 2D nuclei form, which can stack up in the third dimension because of the lack of repulsive interactions near the isoelectric point of RIDC3 (pI = 5.3). **b**, TEM images of a 2D RIDC3 ribbon and a tubular structure with frayed ends, which illustrate the interconversion between tubular and sheet-like morphologies. Shown in the bottom right corner is the 2D image reconstruction of a single-layered portion of the ribbon, which indicates the molecular arrangement of RIDC3 molecules.

the lattice constants remained the same as those at pH = 5.5. However, the nanotubes obtained at pH = 8.5 featured a different interlattice angle (~60°) and greater width (~90 nm, Supplementary Fig. S5), which indicates pH-dependent polymorphism. All RIDC3 assemblies were stable for at least six months at room temperature and remained intact on exchange into a solution with no Zn, but could be dissolved by adding ethylenediaminetetraacetic acid (EDTA). Nanotubes obtained at pH = 8.5 converted fully into 2D arrays on lowering the pH to 5.5, but this transition was not reversible.

Based on these results, we propose a mechanism for the Zn-mediated self-assembly of RIDC3 (Fig. 3a), in which all observed arrays originate from 2D nuclei that have the same arrangement of RIDC3 molecules. At high [Zn]:[RIDC3] ratios or at high pH, where the high-affinity sites are fully deprotonated, Zn–RIDC3 interactions are favoured and nucleation is rapid, which results in a large number of small nuclei that 'roll up' into helical nanotubes. This hypothesis is supported by the observation of monolayered 2D sheets at intermediate maturation times at pH = 8.5 and near the frayed ends of the nanotubes (Fig. 3b). At pH = 5.5, where the high-affinity sites are partially protonated, and at low [Zn]:[RIDC3] ratios, nucleation is slow and yields a smaller number of large 2D sheets. As RIDC3 (isoelectric point (pI) = 5.3) is essentially charge neutral at this pH, the repulsion between RIDC3 molecules is minimized, which encourages stacking of the 2D sheets and additional lateral growth. As demonstrated clearly by the structural studies described below, the crystalline RIDC3 arrays are products of Zn-directed self-assembly and do not result from chemically indiscriminate protein precipitation caused by changes in pH or [Zn].

Structural basis of RIDC3 self-assembly. To help determine the structural basis of Zn-mediated RIDC3 self-assembly, we set out to grow 3D arrays large enough for X-ray diffraction analysis. According to our model above, this could be achieved by further retarding the nucleation of 2D RIDC3 arrays. Large 3D RIDC3 crystalline arrays (≥200 μm × 20 μm × 5 μm) (Supplementary Fig. S7) were obtained in a pH = 6.0 solution that contained 2.5 mM RIDC3, 8 mM Zn^{2+} and 200 mM 2-(bis-(2-hydroxyethyl)-amino)-2-hydroxymethylpropane-1,3-diol (bis-TRIS) as a dual pH and Zn buffer. Given the pK_a (6.46) and Zn^{2+} association constant K_{Zn} (240 M^{-1}) of bis-Tris at 25 °C (ref. 29), we calculated $[Zn]_{free}$ to be <1 mM in the solution. Thus, bis-TRIS can act as an internal 'drip hose' for Zn delivery and promote growth of 3D arrays under

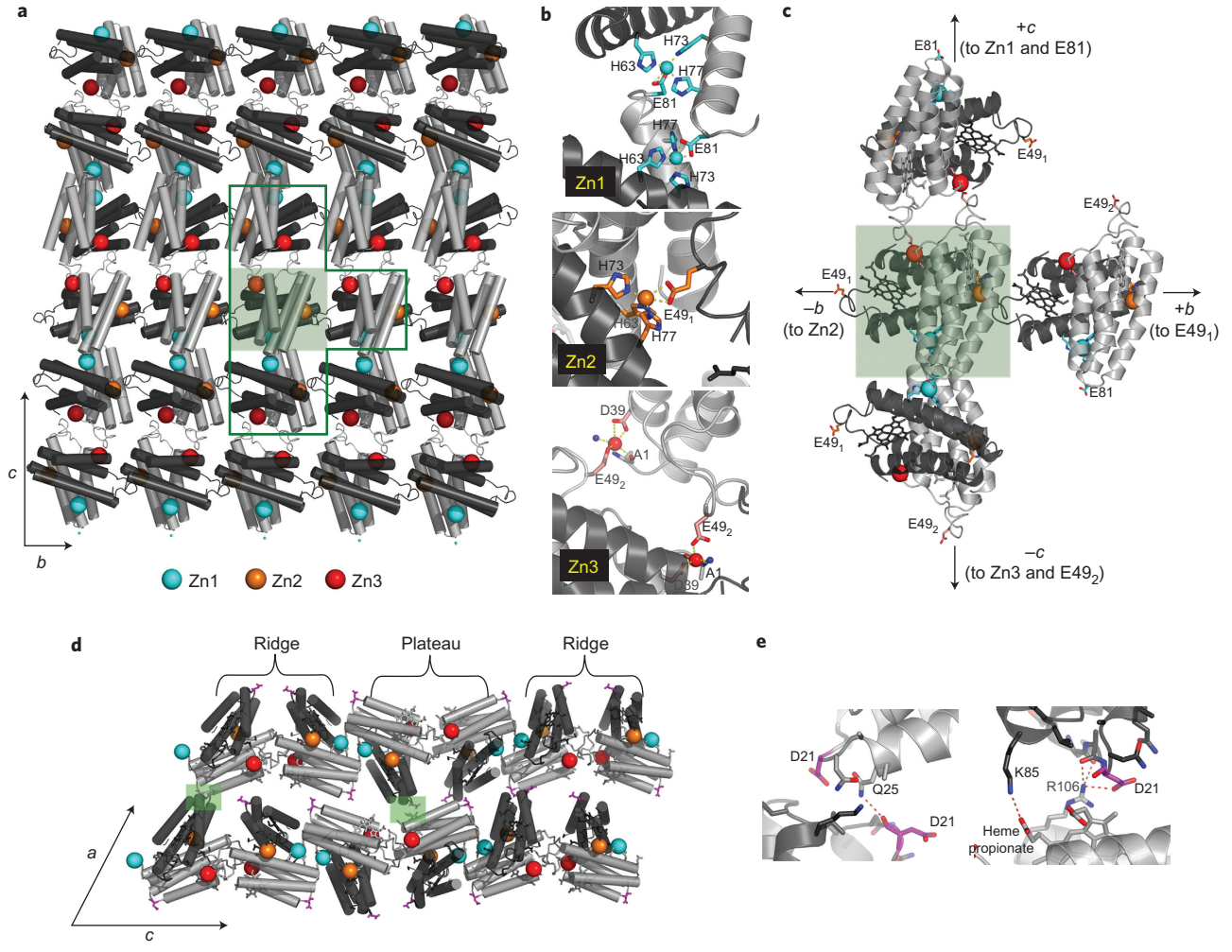

Figure 4 | Structural basis of Zn-mediated RIDC3 assembly. a, Crystallographically determined molecular arrangement in 2D Zn-RIDC3 sheets, viewed normal to the *b*–*c* plane of 3D crystalline arrays. A single **C2**-dimer is highlighted in the shaded green box. **b**, Close-up view of the three different Zn coordination environments that enable the self-assembly of RIDC3 self-assembly in two dimensions. Zn1 and Zn2 sites are formed by the high-affinity coordination motif described in Fig. 1, whereas Zn3 is formed by the low-affinity coordination motif. **c**, Close-up view of the T-shaped boxed area in (**a**). **d**,**e**, Contacts between each 2D RIDC3 layer, responsible for growth in the third dimension, are highlighted in shaded green boxes (**d**) and detailed in (**e**).

slow nucleation conditions. The X-ray crystal structure of 3D Zn–RIDC3 arrays, determined at 2.3 Å resolution (Supplementary Table S1), corroborated our self-assembly model and revealed the central role of Zn coordination. The *b* and *c* axes of these 3D arrays (C_2 space group, $a = 95.6$ Å, $b = 37.8$ Å, $c = 138.5$ Å, $\beta = 112.6^\circ$) match the lattice constants determined for the 2D arrays. We can conclude that, indeed, the 3D arrays form by the stacking of the 2D RIDC3 layers.

Intermolecular associations between RIDC3 molecules in the *b*–*c* plane are mediated entirely through three different types of Zn coordination at the previously described high- and low-affinity sites and involve no other surface contacts (Fig. 4a,b). The layers are built through the repetition of the stabilized **C2**-dimers (depicted as a pair of dark- and light-grey monomers). The observed arrangement of **C2**-dimers perfectly matches the pattern seen in monolayered 2D sheets (Fig. 3b, see Supplementary Fig. S8 for an overlay). Geometric relationships between the Zn-mediated protein interactions in the *b*–*c* layer are best understood by considering a single **C2**-dimer as the origin of nucleation (green shaded box in Fig. 4a,c). The orthogonal Zn1 and Zn2 coordination vectors that originate from the two high-affinity sites are the key determinants of crystal growth along the *c* and *b* cell axes, respectively, and yield the 90° angle between them. Along the *c* axis, two **C2**-dimers interconnect through pairwise Zn1–Glu81 interactions (same mode of interaction as in Fig. 1d). Along the *b* axis, neighbouring **C2**-dimers connect in a head-to-tail fashion through Zn2 and Glu49 ('E49$_1$'), which gives rise to an infinite chain. Finally, a second set of pairwise interactions mediated by Zn3 bound to a low-affinity site and a second Glu49 ('E49$_2$') introduces two-fold symmetry normal to the *b*–*c* plane, which allows propagation along the *c* axis and formation of the 2D array. As opposed to these strong Zn-mediated interactions, growth along the *a* axis involves two small (<250 Å^2) surface patches of polar interactions between the 2D layers (Fig. 4d,e). These only promote crystal growth along the *a* axis when the interlayer repulsions are minimized at pH = 5.5.

To complete our characterization of RIDC3 arrays, we determined the structure of non-flattened RIDC3 nanotubes using cryo-TEM. Images of the pH 5.5 nanotubes (Fig. 5a) that have C_9 helical symmetry were processed using helical, real-space reconstruction methods[30], which yielded a final map at near-nanometre resolution (Fig. 5b,c). Each subunit of the tube contains four discrete densities, each of which has a volume consistent with that of a **C2**-dimer (Supplementary Fig. S9). The outer and inner surfaces of the nanotubes display the same pattern of ridges and plateaus (Fig. 5b,c) seen in each 2D layer of the 3D RIDC3 arrays (Fig. 4). This finding

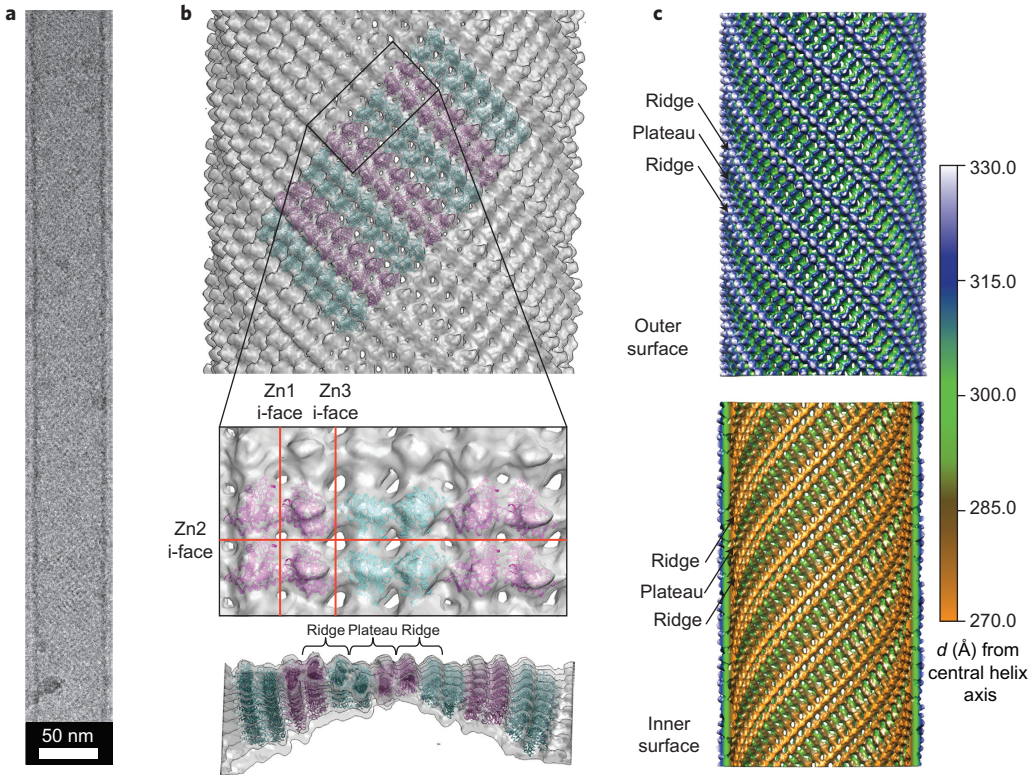

Figure 5 | RIDC3 nanotube structure and assembly. a, A typical vitrified, unstained nanotube used for helical cryo-EM reconstruction (see Supplementary Fig. S14 for collective views of RIDC3 nanotubes). **b**, Tetrameric RIDC3 units (Zn1-linked dimer of **C2**-dimers) modelled into the reconstructed tube-density map. Top, view of the outer surface showing the alternating orientations (cyan and magenta) of tetrameric units; middle, close-up view of the tube surface with modelled tetramers; bottom, lateral view of the tube, highlighting its curvature. Zn1, Zn2 and Zn3-mediated interfaces (i-faces) are indicated with red lines in the middle panel. **c**, Radially colour-coded representations of the outer (top) and inner (bottom) surfaces of the reconstructed nanotube reveal ridges and plateaus consistent with the X-ray crystal structure.

is again in accord with our model that both nanotubes and 3D arrays emerge from the same 2D arrangement of RIDC3 molecules.

Conversion of a 2D crystalline sheet into a helical tube arises when curvature is induced along both lattice vectors of the sheet. Using the Zn1-linked dimer of **C2**-dimers as a rigid unit (see Supplementary Information for details), we built a pseudo-atomic model of the tube to help identify the structural origins of this curvature (Fig. 5b). In this model, the crystallographically observed Zn2 coordination geometry appeared to be maintained (Supplementary Figs S10 and S11). In contrast, the Zn3 interface must simultaneously bend and twist to enable the transformation of the 2D sheet into a hollow tube (Supplementary Figs S10 and S12). Such a significant conformational change precludes the Zn3–$Glu49_2$ coordination observed in crystals. Instead, the EM model indicates that other nearby acidic residues (Asp2 and Asp50) may participate in an alternative mode of Zn coordination (Supplementary Fig. S12).

Directed stacking of RIDC3 sheets in the third dimension through chemical functionalization. As RIDC3 self-assembly is mediated by easily identifiable intermolecular interactions in all three dimensions (Fig. 4), we can correlate crystal morphology with the strength of these interactions. Strongest are the pairwise interactions between **C2**-dimers mediated by Zn1 and Zn3 along the *c* axis, followed by a singular Zn2 interaction between adjacent **C2**-dimers along the *b* axis. Consequently, the 2D sheets observed under fast nucleation conditions are rectangular (Fig. 3b), with their long edges parallel to *c* (*b* in the 2D designation) and short edges parallel to *b* (*a* in the 2D designation). 3D crystalline arrays are thinnest along the *a* axis, which is consistent with the weakest interactions occurring in this direction.

Given these considerations, we asked whether RIDC3 crystal morphology can be directed further through designed interactions. To test this possibility, we quantitatively labelled the D21C variant of RIDC3 with Rhodamine Red C2 maleimide at the surface position 21, which is involved in interlayer interactions in 3D arrays (Figs 4e and 6a). Rhodamine was chosen for functionalization because its derivatives are known to dimerize with sub-millimolar affinities[31], which should be strong enough to promote interlayer interactions and 3D growth, but weak enough to permit self-healing during self-assembly such that crystallinity is achieved. Rhodamine has the added benefit of exhibiting distinct spectral changes on dimerization[32], which can be used to monitor self-assembly.

As a critical test, we examined the Zn-mediated assembly of rhodamine–^{D21C}RIDC3 (R-^{C21}RIDC3) under fast nucleation conditions (pH = 8.5 and ten-fold excess [Zn]), which normally yield nanotubes with unmodified RIDC3. As desired, R-^{C21}RIDC3 instead assembled into multilayered crystalline arrays with the previously observed 2D lattice parameters ($a = 36 \pm 1$ Å, $b = 137 \pm 1$ Å, $\alpha = 90 \pm 1^{\circ}$) (Fig. 6b). At pH = 5.5, which is near the p*I* of RIDC3, the 3D growth of R-^{C21}RIDC3 was amplified further (Fig. 6c). The ultraviolet–visible spectrum of a suspension of Zn-mediated R-^{C21}RIDC3 crystals showed the rhodamine dimer band emerging at 534 nm (Fig. 6d), which is not pronounced when the protein is monomeric in solution. The circular dichroism (CD) spectrum of the same crystal suspension featured a distinct signature in the 520–630 nm region associated with the formation of a rigid environment around the rhodamine dimers (Fig. 6e), and this signature emerged only on full crystal maturation (Supplementary Fig. S13). These observations suggest that the R-^{C21}RIDC3 arrays also display order along the third dimension.

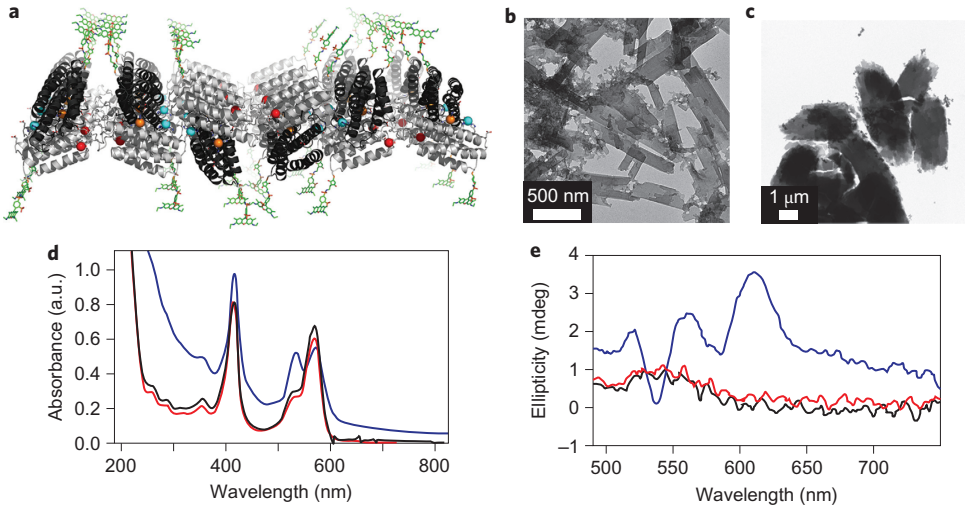

Figure 6 | Rhodamine-directed stacking of 2D RIDC3 arrays. a, Structural model of a 2D RIDC3 sheet uniformly labelled with Rhodamine Red C2 (green sticks) at Cys21. **b**,**c**, TEM images of fully matured, negatively stained R-^{C21}RIDC3 crystals obtained at pH 8.5 (**b**) and pH 5.5 (**c**). The formation of 3D structures at pH 8.5 rather than nanotubes evidences that rhodamine labels promote favourable interactions between 2D, Zn-mediated RIDC3 layers. **d**, Ultraviolet–visible absorption spectrum of a 50 μM R-^{C21}RIDC3 (1 mm path length) sample obtained prior to Zn addition (black), after crystal maturation (red) and on dissolution of crystals by addition of EDTA (blue). The intense band at 415 nm results from haem Soret absorption. **e**, CD spectra of the same R-^{C21}RIDC3 sample. The increase in absorption at 534 nm on Zn addition (**d**) indicates the formation of rhodamine dimers. The emergence of a CD signal in the 520–630 nm region (**e**) during the maturation of R-^{C21}RIDC3 crystals suggests that the rhodamine dimers are in discrete, rigid environments.

Conclusions

Our results demonstrate that the self-assembly of a protein can be directed in all three dimensions to yield periodic arrays with nano- and microscale long-range order. The 1D and 2D RIDC3 arrays described here resemble natural assemblies, such as microtubules[33], helical viruses[34] and S-layers[35], in terms of their dimensions and structural uniformity. More importantly, they recapitulate the ability of these macromolecules to assume distinct conformational states in response to environmental cues such as pH and metal concentrations[36]. A particular advantage of metal-directed protein self-assembly lies in its inherently chemically tunable and modular nature. Self-assembling proteins can be chosen based on their built-in functions (for example, electron transfer for cyt *cb*-562) and/or be augmented functionally through derivatization (for example, by rhodamine) or through the interfacial metals themselves. Together, these features create a versatile platform for engineering functional biomaterials and for studying the fundamentals of protein self-assembly and crystallization under highly controllable conditions.

The practical challenge in the general application of metal-directed protein self-assembly stems from the broad and relatively smooth energy landscape of metal–protein interactions, particularly when exchange-labile metal ions like Zn^{2+} are utilized. That is, owing to the abundance of coordinating residues on any protein's surface, potentially many metal-mediated aggregation states lie close in energy, separated by small barriers[22]. The breadth of the landscape for metal-directed protein self-assembly makes it difficult to design from scratch or re-engineer a protein that self-assembles into discrete supramolecular arrays rather than a heterogeneous mixture of metal cross-linked aggregates. Our findings suggest that the search space for finding discrete protein assemblies can be significantly reduced by (1) constructing symmetrical protein building blocks and (2) installing properly oriented metal coordination motifs on the protein surface that selectively, stably and reversibly bind metal ions. Once these design criteria are met, the smoothness of the landscape for metal-directed protein self-assembly can help ensure that protein self-assembly proceeds under thermodynamic control to yield the conformational states of lowest energy, which can interconvert in response to external stimuli. Indeed, such stimulus–responsive switching between nearly isoenergetic states is a hallmark of biological self-assembly.

Methods

Protein biochemistry. The generation of plasmids, expression of RIDC3 and ^{C21}RIDC3, and sedimentation velocity analysis were performed according to previously published procedures[24]. After purification, mutations were verified by matrix-assisted laser desorption/ionization (MALDI) mass spectrometry (RIDC3: molecular mass, calculated = 12,235 Da, measured = 12,241 Da; ^{C21}RIDC3: molecular mass, calculated =12,222 Da, measured = 12,210 Da).

Prior to labelling with Rhodamine Red C2 maleimide, ^{C21}RIDC3 was reduced with 10 mM dithiothreitol (DTT) and exchanged into a buffer that contained 20 mM HEPES (pH 7.5). A mixture that contained 95 μM ^{C21}RIDC3 and 330 μM Rhodamine Red C2 maleimide was then allowed to react in the dark for one hour at room temperature with gentle stirring. After quenching the reaction with 10 mM DTT, singly labelled protein was purified by anion-exchange chromatography, and quantitative labelling of the single free cysteine was verified by MALDI mass spectrometry (molecular mass, calculated =12,903 Da, measured = 12,889 Da).

Macromolecular crystallography. The PEG-precipitated RIDC3 crystals (Type 1, Fig. 1) and those formed through Zn-directed self-assembly (Type 2, Fig. 4) were obtained by sitting drop vapour diffusion at room temperature (see Supplementary Table S1 for solution composition). Crystals suitable for diffraction experiments were transferred to a solution of mother liquor containing 20% glycerol as the cryoprotectant and flash-frozen in liquid nitrogen. X-ray diffraction data were collected at 100 K and subsequently integrated using MOSFLM and scaled with SCALA[37]. Structures were then determined by molecular replacement using PHASER[38] and subjected to rigid-body, positional, thermal and TLS refinement in REFMAC[39], along with manual rebuilding in COOT[40]. All figures of the resulting structures were produced using PYMOL[41].

Preparation and imaging of RIDC3 nanostructures. 2D RIDC3 arrays or 1D nanotubes were produced at the indicated concentrations in 1 ml or 200 μl volume solutions buffered with either 20 mM MES (pH 5.5 and 6.5), MOPS (pH 7.5) or CHES (pH 8.5). After incubating at room temperature, 3 μl aliquots were deposited on home-made, carbon-coated Cu grids and stained with uranyl acetate. After being allowed to dry, grids were imaged on an FEI Sphera transmission electron microscope equipped with an LaB_6 electron gun operated at 200 keV. Images were recorded on a Gatan $2K^2$ CCD (charge coupled device) using objective-lens underfocus settings that ranged from 400 to 800 nm. Where indicated, 2D image reconstructions were generated using the 2dx software package[42].

For cryo-EM samples, solutions that contained 450 μM RIDC3 and 4.5 mM Zn were incubated at room temperature for one week in a 20 mM MES buffer (pH 5.5) to allow the formation of protein nanotubes. A 3 μl aliquot of this solution was then deposited onto a home-made, lacey carbon grid covered with a thin layer of continuous carbon, or a Quantifoil grid (see Supplementary Fig. S14 for views of

collections of nanotubes), plunged into liquid ethane after blotting and transferred into a precooled FEI Polara multispecimen holder. Images were then recorded on a Gatan 4K^2 CCD camera on an FEI Polara microscope using low-dose conditions ($\sim$12 e$^-$ Å^{-2}) at a nominal magnification of $\times$39,000 (calibrated pixel size = 2.883 Å), with objective-lens underfocus settings that ranged from 1.5 to 4 μm.

Image processing and 3D reconstruction. Image pre-processing was carried out essentially as previously described[30]. Briefly, helical particles were boxed using the helixboxer routine in EMAN[43] and padded to 8192^2 pixels, after which IHRSR++ was used for the following steps. A diffraction pattern was calculated for each tube using 'helixautomation', and tubes with the presence of specific layer lines (Supplementary Fig. S15) were included in the reconstruction. Each tube was then cut into segments using 'helixcutstk', diameters were computed for each segment using 'get_tube_diameters' and segments with consistent diameters were selected for further processing (diameter range 518–635 Å). Finally, the images were corrected for the contrast transfer function (estimated using RobEM (http://cryoEM.ucsd.edu/programs.shtm)) and merged into a single image stack for the final helical reconstruction.

The initial parameters as input for reconstruction using IHRSR++ were estimated from images of negatively stained tubes based on the interlattice angle and tube width. Tubes that yielded consistent reconstructions, as indicated by alternating dimers of **C2**-dimers in the 3D density map, convergence to the same refined helical parameters and a computed Fourier transform consistent with an averaged spectrum from the raw images, were chosen for inclusion in the final reconstruction. Ultimately, the starting values for the axial rise (ΔZ) and the rotation ($\Delta\phi$) per subunit were varied to ensure the global minimum was reached in the final reconstruction.

Received 15 November 2011; accepted 1 February 2012; published online 4 March 2012

References

1. Mann, S. Life as a nanoscale phenomenon. *Angew. Chem. Int. Ed.* **47,** 5306–5320 (2008).
2. Nguyen, S. T., Gin, D. L., Hupp, J. T. & Zhang, X. Supramolecular chemistry: functional structures on the mesoscale. *Proc. Natl Acad. Sci. USA* **98,** 11849–11850 (2001).
3. Seeman, N. C. & Belcher, A. M. Emulating biology: building nanostructures from the bottom up. *Proc. Natl Acad. Sci. USA* **99,** 6451–6455 (2002).
4. Shenton, W., Pum, D., Sleytr, U. B. & Mann, S. Synthesis of cadmium sulphide superlattices using self-assembled bacterial S-layers. *Nature* **389,** 585–587 (1997).
5. McMillan, R. A. *et al.* Ordered nanoparticle arrays formed on engineered chaperonin protein templates. *Nature Mater.* **1,** 247–252 (2002).
6. Lee, Y. J. *et al.* Fabricating genetically engineered high-power lithium-ion batteries using multiple virus genes. *Science* **324,** 1051–1055 (2009).
7. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440,** 297–302 (2006).
8. Zheng, J. P. *et al.* From molecular to macroscopic via the rational design of a self-assembled 3D DNA crystal. *Nature* **461,** 74–77 (2009).
9. 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).
10. Chworos, A. *et al.* Building programmable jigsaw puzzles with RNA. *Science* **306,** 2068–2072 (2004).
11. Aggeli, A. *et al.* Responsive gels formed by the spontaneous self-assembly of peptides into polymeric beta-sheet tapes. *Nature* **386,** 259–262 (1997).
12. Banwell, E. F. *et al.* Rational design and application of responsive [alpha]-helical peptide hydrogels. *Nature Mater.* **8,** 596–600 (2009).
13. Kortemme, T. & Baker, D. Computational design of protein–protein interactions. *Curr. Opin. Chem. Biol.* **8,** 91–97 (2004).
14. Dotan, N., Arad, D., Frolow, F. & Freeman, A. Self-assembly of a tetrahedral lectin into predesigned diamond-like protein crystals. *Angew. Chem. Int. Ed.* **38,** 2363–2366 (1999).
15. Ringler, P. & Schulz, G. E. Self-assembly of proteins into designed networks. *Science* **302,** 106–109 (2003).
16. Ballister, E. R., Lai, A. H., Zuckermann, R. N., Cheng, Y. & Mougous, J. D. *In vitro* self-assembly from a simple protein of tailorable nanotubes building block. *Proc. Natl Acad. Sci. USA* **105,** 3733–3738 (2008).
17. 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).
18. Sinclair, J. C., Davies, K. M., Venien-Bryan, C. & Noble, M. E. M. Generation of protein lattices by fusing proteins with matching rotational symmetry. *Nature Nanotech.* **6,** 558–562 (2011).
19. Caulder, D. L. & Raymond, K. N. Supermolecules by design. *Acc. Chem. Res.* **32,** 975–982 (1999).
20. Leininger, S., Olenyuk, B. & Stang, P. J. Self-assembly of discrete cyclic nanostructures mediated by transition metals. *Chem. Rev.* **100,** 853–907 (2000).
21. Holliday, B. J. & Mirkin, C. A. Strategies for the construction of supramolecular compounds through coordination chemistry. *Angew. Chem. Int. Ed.* **40,** 2022–2043 (2001).
22. Salgado, E. N., Radford, R. J. & Tezcan, F. A. Metal-directed protein self-assembly. *Acc. Chem. Res.* **43,** 661–672 (2010).
23. Radford, R. J., Brodin, J. D., Salgado, E. N. & Tezcan, F. A. Expanding the utility of proteins as platforms for coordination chemistry. *Coord. Chem. Rev.* **255,** 790–803 (2011).
24. Salgado, E. N., Faraone-Mennella, J. & Tezcan, F. A. Controlling protein–protein interactions through metal coordination: assembly of a 16-helix bundle protein. *J. Am. Chem. Soc.* **129,** 13374–13375 (2007).
25. Salgado, E. N., Lewis, R. A., Mossin, S., Rheingold, A. L. & Tezcan, F. A. Control of protein oligomerization symmetry by metal coordination: C_2 and C_3 symmetrical assemblies through Cu^{II} and Ni^{II} coordination. *Inorg. Chem.* **48,** 2726–2728 (2009).
26. Liu, Y. & Kuhlman, B. RosettaDesign server for protein design. *Nucl. Acids Res.* **34,** W235–238 (2006).
27. Jones, S. & Thornton, J. M. Principles of protein–protein interactions. *Proc. Natl Acad. Sci. USA* **93,** 13–20 (1996).
28. Polyakov, A., Severinova, E. & Darst, S. A. Three-dimensional structure of *E. coli* core RNA polymerase: promoter binding and elongation conformations of the enzyme. *Cell* **83,** 365–373 (1995).
29. Scheller, K. H. *et al.* Metal ion/buffer interactions. *Eur. J. Biochem.* **107,** 455–466 (1980).
30. Parent, K. N. *et al.* Cryo-reconstructions of P22 polyheads suggest that phage assembly is nucleated by trimeric interactions among coat proteins. *Phys. Biol.* **7,** 045004 (2010).
31. Hamman, B. D. *et al.* Tetramethylrhodamine dimer formation as a spectroscopic probe of the conformation of *Escherichia coli* ribosomal protein L7/L12 dimers. *J. Biol. Chem.* **271,** 7568–7573 (1996).
32. Chambers, R. W., Kajiwara, T. & Kearns, D. R. Effect of dimer formation of electronic absorption and emission-spectra of ionic dyes – rhodamines and other common dyes. *J. Phys. Chem.* **78,** 380–387 (1974).
33. Li, H. L., DeRosier, D. J., Nicholson, W. V., Nogales, E. & Downing, K. H. Microtubule structure at 8 Ångstrom resolution. *Structure* **10,** 1317–1328 (2002).
34. Glucksman, M. J., Bhattacharjee, S. & Makowski, L. Three-dimensional structure of a cloning vector: X-ray diffraction studies of filamentous bacteriophage M13 at 7 Å resolution. *J. Mol. Biol.* **226,** 455–470 (1992).
35. Sara, M. & Sleytr, U. B. S-layer proteins. *J. Bacteriol.* **182,** 859–868 (2000).
36. Ge, P. & Zhou, Z. H. Hydrogen-bonding networks and RNA bases revealed by cryo electron microscopy suggest a triggering mechanism for calcium switches. *Proc. Natl Acad. Sci. USA* **108,** 9637–9642 (2011).
37. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Cryst. D* **67,** 235–242 (2011).
38. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
39. Murshudov, G., Vagin, A. & Dodson, E. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Cryst. D* **53,** 240–255 (1996).
40. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Cryst. D* **60,** 2126–2132 (2004).
41. DeLano, W. L. The PYMOL molecular graphics system (DeLano Scientific, San Carlos, California (http://www.pymol.org), 2003).
42. Stahlberg, H., Gipson, B., Zeng, X. & Zhang, Z. Y. 2dx – User-friendly image processing for 2D crystals. *J. Struct. Biol.* **157,** 64–72 (2007).
43. Ludtke, S. J., Baldwin, P. R. & Chiu, W. EMAN: semiautomated software for high-resolution single-particle reconstructions. *J. Struct. Biol.* **128,** 82–97 (1999).

Acknowledgements

This work was primarily supported by the US Department of Energy (DOE) (Division of Materials Sciences, Office of Basic Energy Sciences, Award DE-FG02-10ER46677 to F.A.T.). Additional support was provided by the Agouron Foundation, Beckman Foundation, Sloan Foundation, National Science Foundation (CHE-0908115 to F.A.T., protein design), National Institutes of Health (EM characterization, R37GM-033050 and 1S10 RR.020016 to T.S.B. and F32 AI078624 to K.N.P.) and the University of California, San Diego. Portions of this research were carried out at Stanford Synchrotron Radiation Lightsource, operated by Stanford University on behalf of the DOE.

Author contributions

J.D.B. designed and performed most of the experiments and data analysis, and co-wrote the paper. X.I.A. performed computational interface design calculations. C.T. and K.N.P. provided guidance and assistance with EM data collection and analysis. T.S.B. guided EM data analysis and co-wrote the paper. F.A.T. initiated and directed the project, analysed data and co-wrote the paper.

Additional information

The authors declare no competing financial interests. Supplementary information accompanies this paper at www.nature.com/naturechemistry. Reprints and permission information is available online at http://www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.A.T.