

Protein assemblies: nature-inspired and designed nanostructures

Article

Accepted Version

Hamley, I. W. (2019) Protein assemblies: nature-inspired and designed nanostructures. *Biomacromolecules*, 20 (5). pp. 1829-1848. ISSN 1525-7797 doi:
<https://doi.org/10.1021/acs.biomac.9b00228> Available at
<http://centaur.reading.ac.uk/83140/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1021/acs.biomac.9b00228>

Publisher: ACS Publications

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online



Protein Assemblies: Nature-Inspired and Designed Nanostructures

Ian W.Hamley

Dept of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, UK

Abstract

Ordered protein assemblies are attracting interest as next generation biomaterials with a remarkable range of structural and functional properties, leading to potential applications in biocatalysis, materials templating, drug delivery and vaccine development. This Review covers ordered protein assemblies including protein nanowires/nanofibrils, nanorings, nanotubes, designed 2D and 3D ordered protein lattices and protein-like cages including polyhedral virus-like cage structures. The main focus is on designed ordered protein assemblies, in which the spatial organization of the proteins is controlled by tailored non-covalent interactions (including metal-ion binding interactions, electrostatic interactions and ligand-receptor interactions among others) or by careful design of modified (mutant) proteins or *de novo* constructs. The modification of natural protein assemblies including bacterial S-layers and cage-like and rod-like viruses to impart novel function, e.g. enzymatic activity, is also considered. A diversity of structures have been created using distinct approaches, and this Review provides a summary of the state-of-the-art in the development of these systems, which have exceptional potential as advanced bionanomaterials for a diversity of applications.

Keywords: Protein Assemblies, Protein lattices, Protein Cages, Nanowires, Nanotubes

1. Introduction

Nature exploits protein assemblies of different types, ranging from viruses to microtubules and bacterial pili to large protein assemblies and bacterial S-layers. Extended protein assemblies are structural components of the extracellular matrix and biofilms for example, as well as cell motility structures. Many viruses and bacterial microcompartments (BMCs) comprise ordered protein assemblies forming cages (discussed further in Section 5 below) around nucleic acids in the case of viruses or proteins in the case of BMCs, although others classes of viruses have extended assemblies (several are mentioned in Sections 2 and 3 below). Proteins can also assemble around metal centres, for example in heme proteins or form large multi-component assemblies such as ribosomes. Advances in the understanding of computational protein design and genetic engineering methods have recently enabled the rational design of protein subunit structures to form regular 1D-, 2D- and 3D- superstructures including nanowires, nanotubes, 2D and 3D lattices and cage structures. This Review describes research in the field of engineered protein nanostructures, and summarizes the various methods that have been employed to fabricate such ordered protein assemblies. We do not consider natural protein assemblies, although modified variants are discussed. Also not included is a discussion of natural protein crystal structures, obviously a huge separate subject in its own right.

Engineered protein assemblies are attractive for future applications due to their potential for large scale biosynthesis and their biofunctionality and biocompatibility. New emergent properties are expected from the nanostructured materials. Nanowire and nanotube structures may be created that have potentially valuable structural or (opto) electronic properties.

Assemblies based on enzymes may have enhanced catalytic behaviour. Synthetic vaccines can be designed inspired by, or based on, existing virus structures. Other applications for protein cages may exploit their potential to encapsulate cargo.

In the most typical approach, fusion protein assemblies are created with pre-defined symmetry elements in order to define the directionality of the superstructure.¹⁻¹² Less commonly, proteins or peptides may be covalently linked (via flexible peptide or other spacers) in order to pre-impose directionality of assembly.¹³⁻¹⁴

In a different approach, non-covalent interactions can be employed by modification of specific residues at the protein surface to enable π - π stacking¹⁵⁻¹⁸ or metal ion coordination,¹⁹⁻²² for example. Again, the position and relative orientations of modification sites have to be chosen to design protein assemblies with specific symmetries which lead to defined superstructures. Alternatively, host-guest ligand-receptor interactions may be employed.^{15-17, 23} In the simplest case, binary structures can be formed through electrostatic interactions.²⁴⁻³⁴

The topic of ordered protein assemblies has been the focus of a number of previous reviews,³⁵⁻⁴¹ and has been touched upon in broader reviews of protein-based materials.⁴² Here, the state-of-the-art in this emerging field is summarized and exemplified by selected works which elegantly highlight the precision control of nanoscale protein assemblies that can be achieved using advanced design and synthesis methods.

2. Nanofibril and Nanoring Structures

Linear arrays of proteins, i.e. nanofibrils or nanowires result from one-dimensional assembly through suitable binary head-tail interactions such as lock-and-key or ligand-receptor interactions or complementary binding interactions. As an example of assembly guided by ligand-receptor/lock-and-key interactions, linear structures have been constructed through cucurbit[8]uril, CB[8], host-guest interactions.²³ Dimers of glutathione S-transferase (GST) were modified at the two N-termini with the tripeptide FGG which forms a host-guest complex with CB[8] with high binding constant. The resulting nanowires were characterized by diameters of 5 nm and were tens of nm in length.

Using complementary binding peptide interactions, Usui *et al.* created so-called “nanolego” building blocks based on pairing of homotetrameric proteins modified at each of the four sub-unit C-termini by a peptide (Fig.1a).¹ Two protein fusion building blocks each containing four corner peptides were produced, the two types having complementary binding peptide units (Nanolego A and B, Fig.1a). Self-assembly of a mixture of the two leads to linear aggregation and nanofibril formation (Fig.1b). The tetrameric protein scaffold was a superoxide reductase (SOR) and the peptide units were either a mouse PDZ (signalling protein) domain or a PDZ-binding peptide (these two reversibly associate with high binding affinity). In a further development, the PDZ peptides were modified with cysteine mutations to enable step-wise extension of the aggregates based on disulfide bond formation. Finally, the fabrication of finite length aggregates (3-mers) was studied by introducing capping units with only two of the four binding units.¹

Protein nanorings have been created using chemical inducers of protein dimerization, for example using a dimeric methotrexate $\text{MTX}_2\text{-C}_9$ which has high binding affinity for dihydrofolate reductase, DHFR (which was modified with extended linkers between the two subunits).⁴³ Methotrexate is a therapeutic molecule used in chemotherapy and the treatment of autoimmune diseases including arthritis, which acts by inhibiting DHFR. Fig.2a,b show the proposed assembly mechanism along with the DHFR_2 variants prepared, while Fig.2c shows a representative TEM image of nanorings (20-28 nm in outer diameter). The closure into ring structures with defined sizes (i.e. number of subunits) is proposed to result from the balance between conformational flexibility and the entropy of oligomerisation.⁴³ The same group has demonstrated ring structures with 8-30 nm diameter using fusion proteins DHFR-Hint1 (dihydrofolate reductase-histidine triad nucleotide binding) with a variable length peptide spacer between the Hint1 unit and the DHFR protein.² These fusion proteins were polymerized with a dimeric enzyme inhibitor molecule. The ratio of intra- to inter-molecular polymerization could be controlled via adjustment of fusion protein concentration, leading to oligomers containing 2-12 monomers. Inter-molecular cyclization was also favoured by reduction of the length of the linking peptide.²

Small tetramer and trimer oligomers ('nanorings') can be made from complementary coiled-coil forming peptides joined by a disordered flexible $(\text{GN})_x$ peptide linker (Fig.3).¹⁴ Parallel dimeric coiled-coil formation favors the formation of the trimeric and tetrameric structures, which were detected by analytical ultracentrifugation. The peptides with the shortest linker GN_1 formed fibril structures (Fig.3), as imaged by TEM.¹⁴ Peptide fibril structures are not considered further in this section, as this subject has been extensively reviewed elsewhere.⁴⁴

The rod-shaped M13 bacteriophage has been used as a nanowire template scaffold for a variety of applications. In one example, photocatalytic structures based on chemical grafting of photosensitizer and catalyst molecules to the M13 major coat protein p8 was reported.⁴⁸ The M13 bacteriophage comprises approximately 2700 copies of α -helical coat proteins arranged around the viral DNA. Proteins with either chemically linked zinc porphyrin photosensitizer or an iridium oxide catalyst (attached non-covalently using an IrO₂-binding peptide) were co-assembled, producing fibre (nanowire) structures. Light-driven water splitting was observed to be catalysed by the assemblies.⁴⁸ The same strategy was used to produce cobalt oxide nanowires using p8 coat proteins modified with metal ion-binding tetra-glutamate sequences.⁴⁹ Additional incorporation of gold-binding peptide produced hybrid gold-cobalt oxide wires as electrodes improved the charge storage capacity of model lithium ion batteries.⁴⁹ In a further development, the M13 bacteriophage has been genetically engineered to incorporate a terminal peptide that binds single wall carbon nanotubes (SWCNTs) and to incorporate peptides within the major coat protein that bind amorphous iron phosphate. These iron phosphate-based nanowire materials demonstrated excellent performance as cathodes for lithium ion batteries.⁵⁰ An engineered M13 bacteriophage was also developed to produce a Co-Pt hybrid material with superparamagnetic properties.⁵¹ Another recent example involves the use of M13 bacteriophages as scaffolds for metal deposition onto nanofoam meshes prepared by glutaraldehyde cross-linking of M13 modified with a glutamic-acid rich peptide at the N terminus.⁵² The free-standing metal nanofoams prepared may have use in the development of novel electrodes.

Nanorings can be created by co-assembly of modified M13 bacteriophages and linker molecules.⁵³ The M13 bacteriophage was genetically engineered to incorporate anti-

streptavidin and hexahistidine peptides at opposite ends. Stoichiometric addition of streptavidin-NiNTA (Ni(II)-nitriloacetic acid hexahistidine-binding motifs) linkers led to the reversible formation of nanorings.⁵³ Work on M13 bacteriophage and other viruses as scaffolds for nanomaterials development has been reviewed.⁵⁴

Electrostatic interactions between oppositely charged proteins can drive assembly into nanoring structures. In a recent example, supercharged Cerulean and GFP (Green Fluorescent Protein) variants were mixed to form toroid (nanoring) structures comprising two stacked octameric rings, as revealed by high resolution cryo-EM.³⁰

3. Nanotube Structures

In this section, protein nanotubes created by design or modification of natural nanotube structures (rod-like viruses in particular) are considered. Peptide nanotube structures are not discussed in this section, this topic having been the subject of several recent reviews.⁵⁵⁻⁶⁰

A subunit of cytochrome *c* that comprises a four-helix-bundle haem protein has been used as a stable building block to produce nanotube and two- or three-dimensional crystalline structures via metal-coordination interactions.¹⁹ A variant of the protein modified to present two metal-binding bis-histidine motifs on its surface forms a C_2 -symmetric dimer structure stabilized by Zn^{2+} binding. Furthermore, one Zn^{2+} binding site is left open to binding by another dimer, the binding sites being positioned to favour orthogonal assembly, leading to helical chains. Appropriate solution conditions (high pH or low pH and high relative Zn^{2+} ion concentration) leads to fast formation of nanotubes, whereas the opposite conditions lead to slow nucleation into 2D and 3D crystals. Fig.4 shows a cryo-EM image with a model for the helical arrangement of the tetrameric (Zn^{2+} -linked pair of dimers) building unit.¹⁹ The same

cytochrome *c* variant was later coupled via cysteine cross-linking to produce a dimer which was used to construct a tetrameric aggregating unit via Zn^{2+} coordination.¹⁹ The tetrameric dimer wraps helically to form the walls of nanotubes, of two types, which were observed according to the solution conditions (pH, buffer and Zn^{2+} excess). Slow nucleation conditions lead to the formation of 2D- and 3D- crystals in this system, as discussed in Section 4.

The formation of nanotubes was observed by inducing the association of the homotetrameric protein soybean agglutinin (SBA) using a ligand containing both a galactose-based sugar unit to bind the protein and an aromatic motif (Rhodamine B, RhB) to drive π - π stacking interactions (Fig.5).¹⁵ A model for the helical wrapping of the proteins to form the nanotube wall could be obtained, based on electron tomographic imaging. The nanotube growth kinetics can be changed by temperature adjustment and the nanotubes could be dissociated by adding β -cyclodextrin which binds the RhB.¹⁵ These structures resemble protein microtubule structures which are formed from dimers of the protein tubulin and have an outer diameter 24 nm and lengths up to tens of micrometers. Microtubules are involved in mitosis and this, for example, is the basis of the activity of the anti-cancer drug taxol which hinders microtubule depolymerisation, promoting the arrest of mitosis and death of cancer cells.⁶¹⁻⁶²

Tobacco mosaic virus (TMV) has a rod-like structure, consisting of an array of coat proteins wrapped around an RNA core, leading to a nanotube capsid structure. This has been exploited to produce protein nanotubes by modifying the coat proteins, which has benefits because it is suggested that the display of antigens in a regular array leads to an enhanced immunogenicity compared to that induced by free proteins.⁶³ Palmer and co-workers modified TMV coat proteins to enable biotinylation which then allows binding of streptavidin-tagged proteins, exemplified with GFP and an N-terminal fragment of the canine oral papillomavirus L2

protein. In both cases, nanotube structures which elicited higher immunogenic response than unconjugated (and un-assembled) protein were observed.⁶³ In a similar fashion, an S123C mutant TMV coat protein has been used as a handle to attach electron donor and acceptor fluorescence chromophores, creating a scaffold for light harvesting.⁶⁴⁻⁶⁵ Unmodified TMV has a negatively charged surface and this has been used to template the deposition of cationically-modified gold nanoparticles via electrostatic interactions, leading to the formation of twisted fibre bundle structures.³⁴ Magnetic alignment of these structures yielded a plasmonic polarizer.

It is possible to design peptides that adsorb to particular species or surfaces. Using this inherent design flexibility, DeGrado's group have developed antiparallel homo-hexameric coiled coil peptides that assemble into nanotubes around SWCNTs.⁶⁶ The peptides incorporate hydrophobic units, for example the C_β methyl of Ala, to facilitate binding to the hexagonal array of carbon atoms at the nanotube surface.

4. 2D and 3D Crystal Structures

Planar assembled protein structures exist in nature, for example S-layers are monolayers of (glyco)protein structure in the membrane of archaea and certain bacteria.⁶⁷ S-layer protein structures have been reconstructed using tetrameric fusion proteins comprising one of a number of S-protein fragments and three streptavidin units.³ It was shown that these engineered constructs can form planar (oblique) lattice structures on flat surfaces or on cell wall polymer-containing cell wall fragments. The regular display of the streptavidin units enables binding of biotin and biotinylated protein such as ferritin in a regular pattern.³ It has been suggested that the periodic structure of S-layers is ideal for the development of affinity

matrices used in DNA, protein or antibody detection chips.⁶⁸ S-layer proteins have also been considered as potential systems in vaccine development, due to the ability to present antigens at surfaces, along with adjuvant properties.⁶⁸ S-layer structures additionally have potential in the development of immobilized biocatalysts, this having been demonstrated with fusion constructs incorporating extremophile enzymes.⁶⁹⁻⁷⁰ The fluorescent protein GFP has been incorporated into S-layer fusion proteins, enabling the creation of fluorescent biomarkers, pH indicators and the fluorescence imaging of the uptake of S-layered liposomes into cells.⁶⁸ Nanoparticle arrays can also be templated using the periodic structure of S layers, modified to display gold-binding cysteine residues,⁷¹ or utilising S-layer pores to grow cadmium sulfide quantum dots.⁷² Further details on these applications can be found in a review on S-layer structures.⁶⁸

Small cross-shaped aggregates (Fig.6) have been created using the C_4 -symmetric tetrameric catalytic protein RhuA, L-rhamnose-phosphate aldolase.⁷³ Each aldolase subunit was modified with a His₆ tag for oriented binding to a planar surface as well as two tethered biotin uses to bind streptavidin with defined orientation. TEM revealed the presence of the expected cross-shaped network aggregates (Fig.6) on lipid monolayers when mixing biotinylated RhuA (^bR) and ^bR modified with four streptavidins (^bR.S₄). The size of the small aggregates could be expanded by adding spacers of bis-biotinylated streptavidin (^{bb}S, Fig.6). Rod spacers created by mixing ^{bb}S and S led to extended string-like structures.⁷³ In a further development, the authors incorporated a Ca²⁺-binding β -helix fragment of the enzyme serralyisin between two PGAL (6-phospho- β -galactoside) proteins in a PGAL- β -PGAL construct and showed Ca²⁺-dependent switching, with a change in the separation of the two domains in the dumbbell-shaped fusion construct.⁷³

The importance of designing the correct interface between proteins in multi-protein assemblies has been emphasized by several groups.^{24, 74-76} Grueninger *et al.* emphasized the importance of rigid side chain contacts and they designed mutants of proteins with such enhanced contacts.⁷⁴ In particular, they modified monomeric PGAL to favour dimer formation by enriching contacts across local twofold axes and also produced tetramers from dimeric O-acetylserine sulfhydrylase (Oas) and urocanase (Uro) and modified tetrameric RhuA to favour pairwise association at the C_4 axis surfaces. They also adapted the mycobacterial porin MypA to give a D_8 -symmetric unit forming tail-tail dimers. Fig.7 shows the modified protein structures along with the symmetry axes (and Fig.7d shows a TEM image showing linear association of the modified RhuA tetramer shown in Fig.7b).⁷⁴ RhuA was used as a building block for 2-dimensional lattice assembly in a study where aggregation was controlled via several types of interaction via selective protein mutations.²⁰ Specifically, single-disulfide, double-disulfide or double-histidine (metal coordinating) mutants were prepared. The self-assembly process is reversible via oxidation/reduction (of disulphide interactions between cysteines) or using EDTA, a zinc ion chelator in the case of the double-histidine variant. Fig.8 shows the 2D lattices resulting from the assembly process which have different symmetries and a high degree of regularity. The C^{89} RhuA variant was observed to form a number of defect-free 2D lattice polymorphs as a result of the dynamic single disulfide bond flexibility. This material shows ideal auxetic behaviour, undergoing longitudinal expansion upon transverse stretching.²⁰ Recent all-atom molecular dynamics simulations suggest that the free-energy landscape of these lattices is governed by solvent reorganization entropy.⁷⁷

Two-dimensional crystals as well as nanoribbon and nanowire structures were observed using the homotetrameric protein LecA from *Pseudomonas aeruginosa*.¹⁶ LecA is galactose-

specific (which influences its infectivity) and mixing LecA assembly-inducing ligands containing galactopyranoside derivatives with pendant rhodamine B (RhB) units induces association of the proteins due to π - π stacking of the RhB units (cf. Fig.5 and associated discussion in the preceding section). Several stacking modes are possible depending on the ligand spacer length, which influences the geometry of π - π stacking interactions (Fig.9). This leads to assembly into the observed ribbon, 2D crystal or nanowire structures.¹⁶ Three-dimensional crystal structures can be formed by using the sugar-lectin binding and π - π stacking interactions, using concanavalin A (con A), a homotetrameric protein with D_2 symmetry and mannose or lactose-based ligands incorporating aromatic Rhodamine B units to drive dimerization via π - π stacking (cf. for example Fig.5 and Fig.9).¹⁷ Platelet-shaped crystals were noted, and single crystal x-ray diffraction enabled determination of the distinct structures of the crystals formed with different ligand linkers.¹⁷

The methods discussed so far rely on protein site modification, chemical coupling, or production of fusion proteins with defined geometries to drive self-assembly. In contrast, Sinclair *et al.* developed a class of fusion protein comprising units taken from protein assemblies with different rotational symmetries, linked at their termini along one symmetry axis.⁴ The fusion constructs are suitable for high-level, soluble expression in *E. coli*. These can aggregate into 1D or 2D structures termed crysalins. The components may be homologous (comprising only one type of subunit) or heterologous (combined two types of subunit). Using streptavidin/Streptag I as heterologous D_2 assemblies and DsRed as a homologous D_2 assembly, linear assemblies were observed (Fig.10a) whereas combination of *E.coli* ALAD (ALAD: aminolevulinic acid dehydrogenase) as D_4 homologous assembly and streptavidin/Streptag1 as D_2 heterologous unit led to 2D lattices (Fig.10b). The same D_4 homologous unit with Lac21E/Lac21K (heterotetrameric coiled coil peptides based on a Lac

repressor protein sequence, stabilized by Glu/Lys interactions⁷⁸) as a C_2 building block led to a different 2D lattice (Fig.10c).⁴

An alternative approach is *de novo* design of proteins to create 2D lattices. Gonen *et al.* used the Rosetta protein modelling software to design proteins to form 2D lattices with defined symmetries.⁷⁵ Specifically, from among the 17 distinct 2D lattice structures that can be formed from 3D objects, they selected a subset with two unique interfaces and building blocks with internal point symmetry. The designed proteins were then expressed in genetically engineered *E. coli* and the structures assembled in solution were observed by cryo-TEM.⁷⁵ Fig. 11 shows the targeted 2D lattices along with representations of the protein packings and the designed interface structures along with TEM images and projection maps with overlaid design models. The construct for the $P321$ lattice is a trimer of β -helices, that for the $P42_12$ lattice is based on tetrameric α -helices and that $P6$ is based on α -helical hexamers.⁷⁵ The same concept was used to design protein cage structures, as discussed further in Section 5.

A novel route to porous 2D crystal structures was developed based on screening the protein data bank for small oligomeric proteins with defined rotational symmetry, with a central pore smaller than 5 nm, interfaces engineered to avoid steric clashes, flexible loops and termini oriented such that the C-terminus in one oligomer subunit can be linked via a short spacer to the N-terminus of a subunit in different oligomer (favouring inter-oligomer association).⁷⁶ This concept was exploited with a dimeric protein from *S. typhimurium*, STM4215 which forms a hexameric aggregate with a ~3 nm pore (Fig.12a). The protein was engineered such that the subunits were linked with a six-residue linker (Fig.12b), leading to the formation of the 2D honeycomb lattice shown in Fig.12c. Self-assembly was induced by addition of

calcium ions, since each subunit coordinates one Ca^{2+} ion. TEM observations confirmed the formation of the expected 2D honeycomb lattice (Fig.12d). Lanci *et al.* computationally designed a three-helix coiled coil peptide to form honeycomb (p6 symmetry) lattices, by designing a charged outer interface in a homo-trimeric coiled coil (Fig.13a) to favour pairwise complementary electrostatic interactions between helices in neighbouring peptides.²⁴ A single crystal structure for the designed protein confirmed the intended designed structure (Fig.13b).²⁴ Honeycomb lattices are discussed further in Section 5, since such structures have been observed (in the case of building blocks with flexible linkers) to curve into cage structures.⁷⁶

In a pioneering paper, Dotan *et al.* showed that cubic structures based on diamond lattices can be prepared by creating dimers of the lectin concanavalin A, which has a tetrameric structure.⁷⁹ The proteins were dimerized using bis-mannopyranoside, leading to a dumbbell shaped dimer which packs into a diamond lattice due to the imposed configuration of the protein subunits. As mentioned in Section 3, a modified cytochrome protein comprising dimers of C_2 -symmetric dimers via histidine-mediated Zn^{2+} -coordination that forms nanotubes can also assemble into 2D and 3D crystal structures under slow nucleation conditions.¹⁹

The protein ferritin is interesting for protein nanomaterials design due to its highly symmetric cage-like structure. Ferritins comprise 24 subunits which assemble into a pseudo-spherical shell with octahedral symmetry. Each subunit consists of a four α -helix bundle and a fifth short E-helix.²⁵ The E-helices form the C_4 -symmetric channels (Fig.14), of which there are six in a ferritin shell (along with eight C_3 axes). The channels in ferritin have sizes between 0.3 nm and 0.4 nm and allow the transport of small ions and molecules. Yang *et al.* expanded

the C_4 -symmetric pore size (Fig.14) in mature soybean seed ferritin (mSSF) by E helix deletion from the H-1 half of the subunits (Fig.14).²⁵ The expanded pore was able to accommodate poly(L-lysine) (degree of polymerization = 15), leading to tethering of the ferritin cages into a square array via electrostatic interactions (ferritin is rich in acidic residues).²⁵ Zhou *et al.* have exploited the symmetry within the subunits of a ferritin protein in order to substitute aromatic residues located near the C_4 symmetry axes (Fig.15a,b).¹⁸ The substitution of phenylalanine (F) or tyrosine (Y) residues at a single site within each of the 24 protein subunits was designed to induce directional aromatic stacking π - π interactions. These were shown to lead to the formation of planar 2D lattices of ferritin molecules in the case of F-substituted proteins (Fig.15c) or 3D cubic lattices in the case of Y-substituted proteins (Fig.15d) as shown schematically in Fig.15a.¹⁸

Three-dimensional protein crystals can be developed as the basis of a new type of metal-organic framework (MOF).²¹⁻²² The Tezcan group has developed MOFs based on ferritin arrays, initially substituting a Zn^{2+} -binding histidine residue at residue 122 near to the C_3 axis.²¹ The ferritin proteins were then non-covalently linked using benzene-1,4-dicarboxylic acid, leading to bcc crystal structures.²¹ This work was extended to other divalent metal ions and alternative dihydroxamate linkers, leading to a range of MOFs with body-centred cubic or tetragonal lattices.²²

Native ferritin (human heavy chain) forms a fcc lattice (Fig.16a-c) mediated by Ca^{2+} K86Q interactions (Fig.16d). Tezcan and coworkers have shown that this lattice can be used as a template for polymerization of a polymer hydrogel network and that by appropriate choice of a responsive polymer, polymer gels containing embedded ferritin lattices that can be reversibly swollen by change of ionic strength or pH (Fig.16e).⁸⁰ The polymer chosen was

poly(acrylate-*co*-acrylamide) which was prepared by free radical polymerization in the presence of APS (ammonium persulfate) and TEMED (tetramethylenediamine) as initiators and with *N,N'*-methylenebis(acrylamide) as cross-linker, and NaCl was included to limit swelling during polymerization. Post-polymerization swelling was initiated by placement in deionized water (Fig. 16e), leading to an increase in lattice parameter from $a = 19$ nm to $a = 23$ nm, determined by SAXS. The expansion was isotropic as confirmed by isotropic expansion of the faceted polyhedral gel crystals observed by optical microscopy. The gels exhibited self-healing behaviour, for example cracks in the polyhedral crystals induced by ion-induced contraction were observed to spontaneously seal due to the dynamic nature of the bonds between polymer chains and ferritin proteins.⁸⁰

Complexes of a zinc phthalocyanine with eight cationic groups with tetra-anionic pyrene derivative can bind to the anionic surface patches on ferritin (apoferritin), inducing crystallization into fcc packed co-crystals (Fig. 17), with a lattice spacing $a = 20$ nm.⁸¹ The crystals retain the photoactivity of the phthalocyanine dye molecules including fluorescence and light-induced singlet oxygen production.

Simply using electrostatic interactions between oppositely charged proteins, it is possible to produce co-crystals of avidin (which has net negative charge) and cowpea chlorotic mottle virus (CCMV) with a net positive charge, by mixing in aqueous solution.³² The crystals had a bcc structure with lattice spacing $a = 35$ nm. The use of avidin further enabled the pre- or post- assembly functionalization of the crystals with biotinylated molecules such as fluorescent dyes, enzymes or gold nanoparticles.³² In another example, binary crystals have been produced by co-crystallization of oppositely charged ferritin, surface modified with either basic (arginine or lysine) or acidic (glutamic acid or aspartic acid) residues.²⁷ The binary crystal had tetragonal symmetry. Metal oxide nanoparticles could be sequestered

within either or both protein cages. This concept was developed to show that the crystallization could be modulated by metal ion (Mg^{2+}) concentration, the binary tetragonal lattice at low Mg^{2+} concentration being replaced with a unitary cubic lattice at high $[Mg^{2+}]$.⁸² In a related study, binary crystal structures were fabricated using mixtures of anionic proteins and cation-coated gold nanoparticles.³¹ The anionic proteins were cage-like ferritin (apoferritin or magnetoferritin) or CCMV, and these can encapsulate RNA or superparamagnetic iron oxide particles.

Electrostatic interactions lead to the formation of a fcc lattice, as exemplified in mixtures of a P22 bacteriophage coat protein (virus-like particle, VLP) and a G6 (sixth generation) PAMAM dendrimer.^{28, 83} PAMAM dendrimers are cationic poly(amido amine) particles. The P22 bacteriophage coat protein (CP) was modified with a short anionic peptide (VAALEKE)₂ at the C terminus, producing an anionic particle. Mixing in appropriate proportions in suitable ionic strength conditions leads to the formation of a cubic lattice. Alternatively, amorphous aggregates could be prepared using a ditopic protein linker that binds the CP at multiple symmetry-specific sites. This linker can also be used to “cement” the ordered cubic structures formed in mixtures with PAMAM dendrimers, stabilizing the assembly against increase in ionic strength.²⁸ Fusion proteins of the P22 CP with the enzymes ketoisovalerate decarboxylase (KivD) or alcohol dehydrogenase A (AdhA) formed capsid structures similar to those of the unmodified CP.⁸³ The enzymatic activity was found to be retained in the G6 dendrimer-modified CP assemblies, enzymes being confined within the VLPs. In a similar fashion, PAMAM dendrimers can be used to produce binary crystal structures (with hcp or fcc structure) with ferritin.³³ The lattice constant is controlled by the size of the dendrimer (i.e. the generation number).

5. Cage Structures and Polyhedral Nanoparticles

Many viruses and also some proteins such as ferritins⁸⁴ or carboxysomes⁸⁵ (involved in carbon fixation by bacteria) naturally form pseudo-spherical polyhedral cage structures. Clathrin-coated vesicles also have a cage structure, built from triskelion (three-arm) subunits of the Clathrin heavy chain (with bound light chains).⁸⁶ Clathrin can form tetrahedral mini-coat, hexagonal barrel or soccer ball structures *in vitro*.⁸⁶ A discussion of these structures, and those of viruses, is outside the scope of the present review, although examples of virus-like protein nanoparticle assemblies and of virus-derived assemblies are considered.

Controlling the association of coiled coil peptides by design has enabled the assembly of cage structures. Woolfson and coworkers designed a two-component system comprising a homotrimeric coiled coil linked to one of two heterodimeric coiled coils (containing complementary charged residues) through an external disulfide bond between cysteine residues (Fig.18a).⁸⁷ The building blocks are expected to form a honeycomb lattice, however due to the inherent conformational flexibility, closed shell structures termed SAGES, self-assembled cage-like particles, were observed with a diameter of approximately 100 nm.⁸⁷ Later, Ryadnov's group developed cysteine-linked homodimeric coiled-coils with three different faces such that complementary electrostatic interactions between neighbouring dimers would favour formation of a honeycomb lattice (Fig.18b) or so-called tecto-dendrimer unit.²⁶ Again, curling up into virus-like cages was observed in practice, with a diameter of approximately 12- 18 nm. The cage-like particles were able to transfect RNA and DNA. In related work, Castelletto *et al.* have prepared covalently-linked "triskelion" three-arm peptides containing the self-complementary β -sheet sequence RRWTWE, based on a sequence from lactoferrin.¹³ These associate to give honeycomb lattices which curve into cage structures or capsules, able to encapsulate and deliver siRNA, and with additional

antimicrobial activity. This was ascribed to membrane pore formation, as imaged by AFM using model supported lipid bilayers.¹³

Attaching coiled-coil peptides to the free C-terminus of a trimeric aldolase protein (KDPG aldolase from *T. maritima*) enables the design of cage-like assemblies by mixing homologues with complementary heterodimer-forming coiled coils.⁸⁸ After expression of the C-terminal extended aldolase in *E. coli*, TEM and AUC (analytical ultracentrifugation) confirmed the presence of small assemblies in solution, with typical diameters 10-20 nm.⁸⁸⁻⁸⁹ A dimer was reported to be the most common assembled structure, although some tetrahedral and octahedral cages were detected.⁸⁹ This work was extended by using an esterase C₃-symmetric trimer linked via flexible spacers to C-terminally attached helical peptides, designed to form tetrameric coiled coils.⁵ The fusion protein was expected to form octahedral cage structures. The experiments confirmed the formation of such structures, provided the length of the spacer was sufficient, via mass spectrometry, AUC and TEM imaging.

Yeates and coworkers have produced nanocage structures from fusion proteins, using the concept shown in Fig.19. The fusion protein comprised trimeric bromoperoxidase and the dimeric M1 matrix protein of influenza virus, connected by a nine-residue helical linker. The fusion protein was expected to have a tetrahedral shape, favouring the formation of dodecameric cage structures, which indeed were observed by TEM, after recombinant expression in *E. coli* and preparation of aqueous solutions.⁶ A crystal structure for the dodecameric cage structure was later obtained.⁹⁰ The authors also reported a fusion protein that forms helical filaments based on the M1 protein fused to carboxylesterase linked by a 5-residue α -helical linker.⁶ In a similar fashion, fusion proteins designed to encode the information necessary to direct assembly have been used to produce 24-subunit cage

structures, based on positioning of trimeric building blocks along each of the three-fold symmetry axes of a tetrahedron.⁷ The protein structure and interaction modelling software Rosetta⁹¹ was then used to design the sequences at the interfaces of the building blocks, in order to enhance the stability of the interface through packing of suitable hydrophobic residues. In addition, structures were assembled from four trimeric and six dimeric building blocks aligned along the respective tetrahedral symmetry axes. After screening for solubility and compatibility with self-assembly, constructs were selected for experimental study. TEM images showed that the fusion proteins expressed in *E. coli* self-assembled into the designed structures in solution, and crystal structures were obtained for some of the assemblies.⁷

Developing this concept, icosahedral protein cages have been created by design of 60-subunit fusion proteins using trimeric protein scaffolds arranged with icosahedral symmetry (i.e. arranging the trimer three-fold symmetry axis to be coincident with the 3-fold axes of the icosahedron).⁸ The distance from the icosahedron centre and the rotation angle of each trimer about its axis were then optimized for close packing, minimising steric clashes. The hydrophobic interfaces between the trimer building blocks were then filled by computer-assisted design of amino acid sequences. Fig.20 shows cryo-TEM images along with reconstructions from the model design, confirming the icosahedral cage structure.⁸ In an extension of this work, this group also presented 120-subunit icosahedral protein cages with sizes 24-40 nm in diameter based again on designed fusion proteins, but using heteromeric components.⁹ Combinations of distinct building blocks among dimers, trimers and pentamers (according to the icosahedral symmetry elements) were used, for example 12 pentameric and 20 trimeric building blocks aligned along the five-fold and three-fold icosahedral symmetry axes can produce an icosahedral protein cage, which can also be constructed from combinations of pentamers and dimers or trimers and dimers.⁹ In a parallel development of

the helical oligomer fusion strategy, a cubic cage- forming structure was designed and expressed in *E. coli*.¹⁰ The intention was to create a porous material, resembling a MOF, although long-range cubic ordering was not observed. The fusion protein comprises trimeric KDPG aldolase (the same used by Patterson *et al.*⁸⁸) linked via a four-residue helical linker to the dimeric domain of protein FkpA (Fig.21a). The designed 24-subunit cage structure with octahedral symmetry is shown in Fig.21b. Single crystal x-ray diffraction and TEM confirmed the presence of the cubic cage structures after incubation of solutions of the fusion protein, although 12-mers, 18-mers and 24-mers were also detected by mass spectrometry analysis, TEM and SAXS.¹⁰

Ferritin, which is widely used to prepare protein lattices as discussed in the preceding section, is also a cage-like protein.⁹² Mutants have been engineered with Cys residues in metal-binding domains in order to sequester gold formed by reduction from Au³⁺ ions. A crystal structure of the cage with bound gold was obtained.⁹² Modification of ferritin nanocages by attachment of PEG facilitates penetration of the nanoparticles into tumor tissue and airway mucus.⁹³⁻⁹⁴ The PEG surface coating density was optimized by mixing highly PEGylated ferritin (attached via surface amines) with the native ferritin by disassembling the proteins and then reassembling by pH control. The anticancer drug doxorubicin was conjugated to PEGylated ferritin via an acid-labile linker as a therapeutic delivery vehicle.⁹³

The size of protein cages can be tuned by modification of the surface charge, as exemplified by recent work on the capsid-forming enzyme AaLS which in its native form adopts an icosahedral shape (60 sub units).⁹⁵⁻⁹⁶ Directed evolution led to a supercharged luminal capsid surface, able to better encapsulate oppositely (positively charged) cargo, in particular HIV protease, with an expansion in cage size corresponding to 180 or 240 sub units.⁹⁵ The structure of the expanded supercharged cages was investigated in detail using cryo-TEM and

was found to comprise tetrahedrally- and icosahedrally- arranged pentameric units.⁹⁶ By mixing negatively supercharged AaLS with cationically-supercharged ferritin, nested cage structures are obtained.²⁹ This is good example in which tuning of electrostatic interactions on protein surfaces can be used to create new assemblies, in this so-called Matryoshka-type structures.

Self-assembling peptide nanoparticles (SAPNs) have been designed based on peptides that contain two α -helical domains linked by a two-glycine residue spacer, one of the oligomerization domains comprises a coiled coil that forms pentamers, while the other is from a trimeric coiled-coil domain (Fig.22a).⁹⁷⁻⁹⁸ The peptides are positioned to lie on the C_5 and C_3 symmetry axes respectively of an icosahedron or dodecahedron. The nanoparticles contain 60 or 180 peptides were modelled based on an icosahedral structure (Fig.22b).⁹⁸ The former nanoparticle structure is favoured for a *de novo* designed sequence containing cysteine residues (for which there is the potential for disulphide cross-linking)⁹⁷ whereas the latter results from a modified construct with alanines replacing the cysteines and with extended terminal domains.⁹⁸ The systems form roughly spherical nanoparticles with a diameter of 16 nm (for the 60 subunit protein)⁹⁷ or 27 nm (for the 180 subunit protein).⁹⁸ In an extension of this research, variant SAPNs were prepared and characterized by SANS, STEM (which enables molar mass estimation) and DLS.⁹⁹ Based on the determined particle size (the core radius from SANS was 35- 37 nm) and molar mass, it was proposed that these larger nanoparticles contain 240, 300 or 360 peptides, which were modelled as virus-like polyhedra.⁹⁹

Fusion of a *de novo* designed protein that forms a dimeric folded four-helix bundle with a trimeric domain from T4 bacteriophage fibritin leads to oligomers comprising multiples of 6-

mers, as shown in Fig.23.¹¹ Fitting of SAXS data enabled the envelope shape of the aggregates in solution to be obtained, which indicated the presence of tetrahedral and barrel-shaped assemblies.¹¹

Virus-like particles (VLPs) have been modified to create nanoreactors, based on enzymes incorporated as fusion proteins with the scaffold proteins (SPs) which form the inner shell of viruses, which are surrounded with coat proteins (CPs). This is exemplified by the N-terminal conjugation of alcohol dehydrogenase (AdhD) to the SP of bacteriophage P22 (Fig.24).¹² A P22 VLP is composed of approximately 420 copies of a 46.6 kDa coat protein (CP) that assembles into an icosahedral capsid with the aid of approximately 100–330 copies of a 33.6 kDa scaffolding protein (SP). In the AdhD-SP conjugate, the C-terminal α -helical scaffold protein facilitates co-assembly with the P22 CP, leading to particles indistinguishable from those of native P22. The AdhD gene is inserted into the pET11 expression vector (Fig.24). The catalytic activity was maintained, furthermore since P22 undergoes structural transitions on heating which lead to expansion or pore formation, the accessibility of the tethered enzymes can be adjusted thermally.¹² In a parallel study, encapsulation of thermostable CelB glycosidase inside the P22 capsid was demonstrated using the same concept, again with no loss of enzyme activity and without impairing the ability of the P22 to undergo thermally-induced morphology changes.¹⁰⁰ The packaging of fluorescent proteins at the interior surface of P22 VLPs was demonstrated in a similar fashion.¹⁰¹ The concept was later extended to incorporate multiple (2 or 3) fused enzymes, including CelB and dimeric ADP-dependent glucokinase and also monomeric AT-dependent galactokinase in the 3-enzyme construct.¹⁰² These enzymes can catalyse a cascade of coupled reactions, demonstrated with lactose as substrate. The activity of all encapsulated enzymes was confirmed, and the kinetic parameters were measured.

6. Conclusions and Future Perspectives

In summary, a variety of approaches have been successfully demonstrated to assemble proteins into defined aggregates including cages and 1D, 2D or 3D structures. Protein assembly can be induced by non-covalent interactions such as metal-ion mediated pairing or hydrophobic side-chain interface engineering or electrostatic interactions using modified proteins or by *de novo* design of proteins. Protein mutants can be created exploiting C- or N-terminal modifications or site-selective modifications, utilising suitable residues such as cysteines located suitably with respect to protein sub-unit symmetry axes. A range of natural multi-subunit proteins can be used for this purpose, there being a range of proteins with suitable C_2 , C_3 , C_6 , D_2 and D_4 subunit symmetries, among others, which can be used to produce 2D and 3D lattices, whilst an essential element of large cage structures is the inclusion of pentameric proteins in the design. It is possible to produce cage and 2D and 3D lattices with a remarkable degree of precision in the ordering using protein assemblies, provided appropriate design rules are followed.

It will be interesting to follow further research developments that lead to the design and creation of novel lattice structures (and possibly aperiodic quasicrystals). Perhaps inspiration can be taken from the field of DNA origami, utilising stronger covalent interactions (such as multiple hydrogen bonds between nucleic acids) than have been exploited thus far. Other superstructures such as multi-ring (and interlinked) assemblies can be envisaged in analogy with the field of rotaxanes, with the related challenge to construct novel protein motors, inspired or distinct from natural ones.

Coiled coil proteins/peptides are an attractive design unit for simple *de novo* designed assemblies including polyhedral particles, ring structures, planar lattices and linear assemblies although coiled coils are combined with other elements to create cage and 3D lattice structures. On the other hand, assemblies based on natural proteins such as enzymes enables the potential exploitation of the native function, for example biocatalysis. Native and mutant proteins can be produced recombinantly (commonly using *E. coli* expression vectors), leading to the potential to scale up the synthesis.

An alternative method to produce functional protein-based materials is to use protein assemblies as templates or scaffolds, as exemplified by the modification of polyhedral or rod-like virus capsids with desired function by engineering of the coat or scaffold proteins. Another example is the use of bacterial S-layers to produce two-dimensional protein arrays, modified to enable metal templating or to create planar catalysts by positioning enzymes. Materials with remarkable catalytic and optoelectronic properties have been engineered in this way. These may have a role in addressing important challenges, for example in photocatalytic water oxidation or in CO₂ fixation, as discussed above, and related applications in clean energy generation can be envisaged, by choice of appropriate enzymes. Since enzyme cascades have important roles *in vivo*, their engineering using protein assemblies is also an exciting avenue for future developments.

As well as applications in biocatalysis, protein assemblies have potential in the creation of novel porous materials for separation and cage-like structures can be used to encapsulate and deliver cargo such as drug molecules, in a targeted manor (exploiting or modifying the protein coating to target particular cell functionalities). Alternatively, the intrinsic properties of such particles could be used to induce immunogenicity, with the potential additional

benefits arising from self-adjuvant properties. Another class of therapeutic approaches may involve the modification of the assembly pathway of protein superstructures such as microtubules, which is the basis of taxol's anticancer activity. There are many related examples of protein structures (e.g. extracellular protein assemblies, ion pumps etc) involved in disease progression which have not yet been targeted.

As yet, there are few examples of dynamic engineered protein assemblies, although in one recent example it has been demonstrated that a transition in 2D lattice structure of RhuA variant crystals (discussed in detail in Section 4) can be achieved by vigorous mixing and sedimentation (or by reversible Ca^{2+} - induced switching).⁷⁷ There is considerable scope to produce new responsive materials by incorporating biological motor protein elements (myosins, dyneins, ATPase etc). This is an area with great potential to produce innovative active biomaterials.

Considering the impressive examples outlined in this Review, it should be clear that protein materials are very promising components of next generation structural and functional biomaterials based on the unprecedented diversity of structures and properties that have evolved in natural proteins or can be designed into *de novo* constructs.

References

- (1) Usui, K.; Maki, T.; Ito, F.; Suenaga, A.; Kidoaki, S.; Itoh, M.; Taiji, M.; Matsuda, T.; Hayashizaki, Y.; Suzuki, H., Nanoscale elongating control of the self-assembled protein filament with the cysteine-introduced building blocks. *Protein Science* **2009**, *18* (5), 960-969.

- (2) Chou, T.-F.; So, C.; White, B. R.; Carlson, J. C. T.; Sarikaya, M.; Wagner, C. R., Enzyme nanorings. *ACS Nano* **2008**, *12*, 2519-2525.
- (3) Moll, D.; Huber, C.; Schlegel, B.; Pum, D.; Sleytr, U. B.; Sara, M., S-layer-streptavidin fusion proteins as template for nanopatterned molecular arrays. *Proc. Nat. Acad. Sci USA* **2002**, *99* (23), 14646-14651.
- (4) 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.* **2011**, *6* (9), 558-562.
- (5) Sciore, A.; Su, M.; Koldewey, P.; Eschweiler, J. D.; Diffley, K. A.; Linhares, B. M.; Ruotolo, B. T.; Bardwell, J. C. A.; Skiniotis, G.; Marsh, E. N. G., Flexible, symmetry-directed approach to assembling protein cages. *Proc. Nat. Acad. Sci USA* **2016**, *113* (31), 8681-8686.
- (6) Padilla, J. E.; Colovos, C.; Yeates, T. O., Nanohedra: Using symmetry to design self assembling protein cages, layers, crystals, and filaments. *Proc. Nat. Acad. Sci USA* **2001**, *98* (5), 2217-2221.
- (7) King, N. P.; Bale, J. B.; Sheffler, W.; McNamara, D. E.; Gonen, S.; Gonen, T.; Yeates, T. O.; Baker, D., Accurate design of co-assembling multi-component protein nanomaterials. *Nature* **2014**, *510* (7503), 103-108.
- (8) Hsia, Y.; Bale, J. B.; Gonen, S.; Shi, D.; Sheffler, W.; Fong, K. K.; Nattermann, U.; Xu, C. F.; Huang, P. S.; Ravichandran, R.; Yi, S.; Davis, T. N.; Gonen, T.; King, N. P.; Baker, D., Design of a hyperstable 60-subunit protein icosahedron. *Nature* **2016**, *535* (7610), 136-139.
- (9) Bale, J. B.; Gonen, S.; Liu, Y. X.; Sheffler, W.; Ellis, D.; Thomas, C.; Cascio, D.; Yeates, T. O.; Gonen, T.; King, N. P.; Baker, D., Accurate design of megadalton-scale two-component icosahedral protein complexes. *Science* **2016**, *353* (6297), 389-394.
- (10) Lai, Y. T.; Reading, E.; Hura, G. L.; Tsai, K. L.; Laganowsky, A.; Asturias, F. J.; Tainer, J. A.; Robinson, C. V.; Yeates, T. O., Structure of a designed protein cage that self-assembles into a highly porous cube. *Nature Chem.* **2014**, *6* (12), 1065-1071.

- (11) Kobayashi, N.; Yanase, K.; Sato, T.; Unzai, S.; Hecht, M. H.; Arai, R., Self-Assembling Nano-Architectures Created from a Protein Nano-Building Block Using an Intermolecularly Folded Dimeric de Novo Protein. *J. Am. Chem. Soc.* **2015**, *137* (35), 11285-11293.
- (12) Patterson, D. P.; Prevelige, P. E.; Douglas, T., Nanoreactors by Programmed Enzyme Encapsulation Inside the Capsid of the Bacteriophage P22. *ACS Nano* **2012**, *6* (6), 5000-5009.
- (13) Castelletto, V.; de Santis, E.; Alkassem, H.; Lamarre, B.; Noble, J. E.; Ray, S.; Bella, A.; Burns, J. R.; Hoogenboom, B. W.; Ryadnov, M. G., Structurally plastic peptide capsules for synthetic antimicrobial viruses. *Chem. Sci.* **2016**, *7* (3), 1707-1711.
- (14) Boyle, A. L.; Bromley, E. H. C.; Bartlett, G. J.; Sessions, R. B.; Sharp, T. H.; Williams, C. L.; Curmi, P. M. G.; Forde, N. R.; Linke, H.; Woolfson, D. N., Squaring the Circle in Peptide Assembly: From Fibers to Discrete Nanostructures by de Novo Design. *J. Am. Chem. Soc.* **2012**, *134* (37), 15457-15467.
- (15) Yang, G.; Zhang, X.; Kochovski, Z.; Zhang, Y. F.; Dai, B.; Sakai, F. J.; Jiang, L.; Lu, Y.; Ballauff, M.; Li, X. M.; Liu, C.; Chen, G. S.; Jiang, M., Precise and Reversible Protein-Microtubule-Like Structure with Helicity Driven by Dual Supramolecular Interactions. *J. Am. Chem. Soc.* **2016**, *138* (6), 1932-1937.
- (16) Yang, G.; Ding, H. M.; Kochovski, Z.; Hu, R. T.; Lu, Y.; Ma, Y. Q.; Chen, G. S.; Jiang, M., 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. Engl.* **2017**, *56* (36), 10691-10695.
- (17) Sakai, F.; Yang, G.; Weiss, M. S.; Liu, Y. J.; Chen, G. S.; Jiang, M., Protein crystalline frameworks with controllable interpenetration directed by dual supramolecular interactions. *Nat. Commun.* **2014**, *5*, 8.
- (18) Zhou, K.; Zang, J.; Chen, H.; Wang, W.; Wang, H.; Zhao, G., On-axis alignment of protein nanocage assemblies from 2D to 3D through the aromatic stacking interactions of amino acid residues. *ACS Nano* **2018**, *12* (11), 11323-11332.

- (19) Brodin, J. D.; Ambroggio, X. I.; Tang, C. Y.; Parent, K. N.; Baker, T. S.; Tezcan, F. A., Metal-directed, chemically tunable assembly of one-, two- and three-dimensional crystalline protein arrays. *Nature Chem.* **2012**, *4* (5), 375-382.
- (20) Suzuki, Y.; Cardone, G.; Restrepo, D.; Zavattieri, P. D.; Baker, T. S.; Tezcan, F. A., Self-assembly of coherently dynamic, auxetic, two-dimensional protein crystals. *Nature* **2016**, *533* (7603), 369-+.
- (21) Sontz, P. A.; Bailey, J. B.; Aln, S.; Tezcan, F. A., A Metal Organic Framework with Spherical Protein Nodes: Rational Chemical Design of 3D Protein Crystals. *J. Am. Chem. Soc.* **2015**, *137* (36), 11598-11601.
- (22) Bailey, J. B.; Zhang, L.; Chiong, J. A.; Ahn, S.; Tezcan, F. A., Synthetic Modularity of Protein-Metal-Organic Frameworks. *J. Am. Chem. Soc.* **2017**, *139* (24), 8160-8166.
- (23) Hou, C.; Li, J.; Zhao, L.; Zhang, W.; Luo, Q.; Dong, Z.; Xu, J.; Liu, j., Construction of protein nanowires through cucurbit[8]uril-based highly specific host-guest interactions: An approach to the assembly of functional proteins. *Angew. Chem., Int. Ed. Engl.* **2013**, *52*, 5590-5593.
- (24) Lanci, C. J.; MacDermaid, C. M.; Kang, S. G.; Acharya, R.; North, B.; Yang, X.; Qiu, X. J.; DeGrado, W. F.; Saven, J. G., Computational design of a protein crystal. *Proc. Nat. Acad. Sci USA* **2012**, *109* (19), 7304-7309.
- (25) Yang, R.; Chen, L.; Yang, S.; Lv, C.; Leng, X.; Zhao, G., 2D square arrays of protein nanocages through channel-directed electrostatic interactions with poly(alpha, L-lysine). *Chem. Comm.* **2014**, *50* (22), 2879-2882.
- (26) Noble, J. E.; De Santis, E.; Ravi, J.; Lamarre, B.; Castelletto, V.; Mantell, J.; Ray, S.; Ryadnov, M. G., A De Novo Virus-Like Topology for Synthetic Virions. *J. Am. Chem. Soc.* **2016**, *138* (37), 12202-12210.
- (27) Künzle, M.; Eckert, T.; Beck, T., Binary Protein Crystals for the Assembly of Inorganic Nanoparticle Superlattices. *J. Am. Chem. Soc.* **2016**, *138* (39), 12731-12734.

- (28) McCoy, K.; Uchida, M.; Lee, B.; Douglas, T., Templated Assembly of a Functional Ordered Protein Macromolecular Framework from P22 Virus-like Particles. *Acs Nano* **2018**, *12* (4), 3541-3550.
- (29) Beck, T.; Tetter, S.; Künzle, M.; Hilvert, D., Construction of Matryoshka-Type Structures from Supercharged Protein Nanocages. *Angew. Chem., Int. Ed. Engl.* **2015**, *54* (3), 937-940.
- (30) Simon, A. J.; Zhou, Y.; Ramasubramani, V.; Glaser, J.; Pothukuchy, A.; Gollihar, J.; Gerberich, J. C.; Leggere, J. C.; Morrow, B. R.; Jung, C.; Glotzer, S. C.; Taylor, D. W.; Ellington, A. D., Supercharging enables organized assembly of synthetic biomolecules. *Nature Chem.* **2019**, *11* (3), 204-212.
- (31) Kostiainen, M. A.; Hiekkataipale, P.; Laiho, A.; Lemieux, V.; Seitsonen, J.; Ruokolainen, J.; Ceci, P., Electrostatic assembly of binary nanoparticle superlattices using protein cages. *Nature Nanotechnology* **2013**, *8* (1), 52-56.
- (32) Liljestrom, V.; Mikkila, J.; Kostiainen, M. A., Self-assembly and modular functionalization of three-dimensional crystals from oppositely charged proteins. *Nat. Commun.* **2014**, *5*, 9.
- (33) Liljestrom, V.; Seitsonen, J.; Kostiainen, M. A., Electrostatic Self-Assembly of Soft Matter Nanoparticle Cocrystals with Tunable Lattice Parameters. *ACS Nano* **2015**, *9* (11), 11278-11285.
- (34) Liljestrom, V.; Ora, A.; Hassinen, J.; Rekola, H. T.; Nonappa; Heilala, M.; Hynninen, V.; Joensuu, J. J.; Ras, R. H. A.; Torma, P.; Ikkala, O.; Kostiainen, M. A., Cooperative colloidal self-assembly of metal-protein superlattice wires. *Nat. Commun.* **2017**, *8*, 10.
- (35) Papapostolou, D.; Howorka, S., Engineering and exploiting protein assemblies in synthetic biology. *Mol. Biosyst.* **2009**, *5* (7), 723-732.
- (36) Howorka, S., Rationally engineering natural protein assemblies in nanobiotechnology. *Curr. Opin. Biotechnol.* **2011**, *22* (4), 485-491.
- (37) Lai, Y. T.; King, N. P.; Yeates, T. O., Principles for designing ordered protein assemblies. *Trends Cell Biol.* **2012**, *22* (12), 653-661.

- (38) Sciore, A.; Marsh, E. N. G., Symmetry-directed design of protein cages and protein lattices and their applications. In *Macromolecular Protein Complexes*, Harris, J. R.; Marles-Wright, J., Eds. Springer: Cham, Switzerland, 2017; pp 195-224.
- (39) Luo, Q.; Hou, C. X.; Bai, Y. S.; Wang, R. B.; Liu, J. Q., Protein Assembly: Versatile Approaches to Construct Highly Ordered Nanostructures. *Chem. Rev.* **2016**, *116* (22), 13571-13632.
- (40) Kuan, S. L.; Bergamini, F. R. G.; Weil, T., Functional protein nanostructures: a chemical toolbox. *Chem. Soc. Rev.* **2018**, *47* (24), 9069-9105.
- (41) Aumiller, W. M.; Uchida, M.; Douglas, T., Protein cage assembly across multiple length scales. *Chem. Soc. Rev.* **2018**, *47* (10), 3433-3469.
- (42) de la Rica, R.; Matsui, H., Applications of peptide and protein-based materials in bionanotechnology. *Chem. Soc. Rev.* **2010**, *39* (9), 3499-3509.
- (43) Carlson, J. C. T.; Jena, S. S.; Flenniken, M.; Chou, T. F.; Siegel, R. A.; Wagner, C. R., Chemically controlled self-assembly of protein nanorings. *J. Am. Chem. Soc.* **2006**, *128* (23), 7630-7638.
- (44) Hamley, I. W., Peptide fibrillisation. *Angewa. Chem., Int. Ed. Engl.* **2007**, *46*, 8128-8147.
- (45) Palmer, A. M., Neuroprotective therapeutics for Alzheimer's disease: progress and prospects. *Trends Pharmacol. Sci.* **2011**, *32* (3), 141-147.
- (46) Aida, T.; Meijer, E. W.; Stupp, S. I., Functional Supramolecular Polymers. *Science* **2012**, *335* (6070), 813-817.
- (47) Wei, G.; Su, Z. Q.; Reynolds, N. P.; Arosio, P.; Hamley, I. W.; Gazit, E.; Mezzenga, R., Self-assembling peptide and protein amyloids: from structure to tailored function in nanotechnology. *Chem. Soc. Rev.* **2017**, *46* (15), 4661-4708.
- (48) Nam, Y. S.; Magyar, A. P.; Lee, D.; Kim, J. W.; Yun, D. S.; Park, H.; Pollom, T. S.; Weitz, D. A.; Belcher, A. M., Biologically templated photocatalytic nanostructures for sustained light-driven water oxidation. *Nature Nanotech.* **2010**, *5* (5), 340-344.

- (49) Nam, K. T.; Kim, D. W.; Yoo, P. J.; Chiang, C. Y.; Meethong, N.; Hammond, P. T.; Chiang, Y. M.; Belcher, A. M., Virus-enabled synthesis and assembly of nanowires for lithium ion battery electrodes. *Science* **2006**, *312* (5775), 885-888.
- (50) Lee, Y. J.; Yi, H.; Kim, W. J.; Kang, K.; Yun, D. S.; Strano, M. S.; Ceder, G.; Belcher, A. M., Fabricating Genetically Engineered High-Power Lithium-Ion Batteries Using Multiple Virus Genes. *Science* **2009**, *324* (5930), 1051-1055.
- (51) Lee, S. K.; Yun, D. S.; Belcher, A. M., Cobalt ion mediated self-assembly of genetically engineered bacteriophage for biomimetic Co-Pt hybrid material. *Biomacromolecules* **2006**, *7* (1), 14-17.
- (52) Ohmura, J. F.; Burpo, F. J.; Lescott, C. J.; Ransil, A.; Yoon, Y.; Records, W. C.; Belcher, A. M., Highly adjustable 3D nano- architectures and chemistries via assembled 1D biological templates. *Nanoscale* **2019**, *11* (3), 1091-1102.
- (53) Nam, K. T.; Peelle, B. R.; Lee, S. W.; Belcher, A. M., Genetically driven assembly of nanorings based on the M13 virus. *Nano Lett.* **2004**, *4* (1), 23-27.
- (54) Flynn, C. E.; Lee, S. W.; Peelle, B. R.; Belcher, A. M., Viruses as vehicles for growth, organization and assembly of materials. *Acta Mater.* **2003**, *51* (19), 5867-5880.
- (55) Gao, X.; Matsui, H., Peptide-based nanotubes and their applications in bionanotechnology. *Adv. Mater.* **2005**, *17*, 2037-2050.
- (56) Tsai, C. J.; Zheng, J.; Zanuy, D.; Haspel, N.; Wolfson, H.; Aleman, C.; Nussinov, R., Principles of nanostructure design with protein building blocks. *Proteins* **2007**, *68* (1), 1-12.
- (57) Scanlon, S.; Aggeli, A., Self-assembling peptide nanotubes. *Nano Today* **2008**, *3* (3-4), 22-30.
- (58) Valéry, C.; Artzner, F.; Paternostre, M., Peptide nanotubes: molecular organisations, self-assembly mechanisms and applications. *Soft Matter* **2011**, *7*, 9583-9594.
- (59) Hamley, I. W., Peptide Nanotubes. *Angew. Chem., Int. Ed. Engl.* **2014**, *53* (27), 6866-6881.
- (60) Adler-Abramovich, L.; Gazit, E., The physical properties of supramolecular peptide assemblies: from building block association to technological applications. *Chem. Soc. Rev.* **2014**, *43* (20), 6881-6893.

- (61) Schiff, P. B.; Fant, J.; Horwitz, S. B., Promotion of Microtubule Assembly In vitro by Taxol. *Nature* **1979**, 277 (5698), 665-667.
- (62) Jordan, M. A.; Wilson, L., Microtubules as a target for anticancer drugs. *Nat. Rev. Cancer* **2004**, 4 (4), 253-265.
- (63) Smith, M. L.; Lindbo, J. A.; Dillard-Telm, S.; Brosio, P. M.; Lasnik, A. B.; McCormick, A. A.; Nguyen, L. V.; Palmer, K. E., Modified Tobacco mosaic virus particles as scaffolds for display of protein antigens for vaccine applications. *Virology* **2006**, 348 (2), 475-488.
- (64) Miller, R. A.; Presley, A. D.; Francis, M. B., Self-assembling light-harvesting systems from synthetically modified tobacco mosaic virus coat proteins. *J. Am. Chem. Soc.* **2007**, 129 (11), 3104-3109.
- (65) Ma, Y. Z.; Miller, R. A.; Fleming, G. R.; Francis, M. B., Energy transfer dynamics in light-harvesting assemblies templated by the tobacco mosaic virus coat protein. *J. Phys. Chem. B* **2008**, 112 (22), 6887-6892.
- (66) Grigoryan, G.; Kim, Y. H.; Acharya, R.; Axelrod, K.; Jain, R. M.; Willis, L.; Drndic, M.; Kikkawa, J. M.; DeGrado, W. F., Computational Design of Virus-Like Protein Assemblies on Carbon Nanotube Surfaces. *Science* **2011**, 332 (6033), 1071-1076.
- (67) Sleytr, U. B.; Schuster, B.; Egelseer, E. M.; Pum, D., S-layers: principles and applications. *Fems Microbiology Reviews* **2014**, 38 (5), 823-864.
- (68) Ilk, N.; Egelseer, E. M.; Sleytr, U. B., S-layer fusion proteins - construction principles and applications. *Curr. Opin. Biotechnol.* **2011**, 22 (6), 824-831.
- (69) Tschiggerl, H.; Breitwieser, A.; de Roo, G.; Verwoerd, T.; Schaffer, C.; Sleytr, U. B., Exploitation of the S-layer self-assembly system for site directed immobilization of enzymes demonstrated for an extremophilic laminarinase from *Pyrococcus furiosus*. *J. Biotechnol.* **2008**, 133 (3), 403-411.
- (70) Schaffer, C.; Novotny, R.; Kupcu, S.; Zayni, S.; Scheberl, A.; Friedmann, J.; Sleytr, U. B.; Messner, P., Novel biocatalysts based on S-Layer self-assembly of *Geobacillus Stearothermophilus* NRS 2004/3a: A nano biotechnological approach. *Small* **2007**, 3 (9), 1549-1559.

- (71) Badelt-Lichtblau, H.; Kainz, B.; Vollenkle, C.; Egelseer, E. M.; Sleytr, U. B.; Pum, D.; Ilk, N., Genetic Engineering of the S-Layer Protein SbpA of *Lysinibacillus sphaericus* CCM 2177 for the Generation of Functionalized Nanoarrays. *Bioconjugate Chem.* **2009**, *20* (5), 895-903.
- (72) Shenton, W.; Pum, D.; Sleytr, U. B.; Mann, S., Synthesis of cadmium sulphide superlattices using self-assembled bacterial S-layers. *Nature* **1997**, *389* (6651), 585-587.
- (73) Ringler, P.; Schulz, G. E., Self-assembly of proteins into designed networks. *Science* **2003**, *302* (5642), 106-109.
- (74) Grueninger, D.; Treiber, N.; Ziegler, M. O. P.; Koetter, J. W. A.; Schulze, M. S.; Schulz, G. E., Designed protein-protein association. *Science* **2008**, *319* (5860), 206-209.
- (75) Gonen, S.; DiMaio, F.; Gonen, T.; Baker, D., Design of ordered two-dimensional arrays mediated by noncovalent protein-protein interfaces. *Science* **2015**, *348* (6241), 1365-1368.
- (76) Matthaei, J. F.; DiMaio, F.; Richards, J. J.; Pozzo, L. D.; Baker, D.; Baneyx, F., Designing Two-Dimensional Protein Arrays through Fusion of Multimers and Interface Mutations. *Nano Lett.* **2015**, *15* (8), 5235-5239.
- (77) Alberstein, R.; Suzuki, Y.; Paesani, F.; Tezcan, F. A., Engineering the entropy-driven free-energy landscape of a dynamic nanoporous protein assembly. *Nature Chem.* **2018**, *10* (7), 732-739.
- (78) Fairman, R.; Chao, H. G.; Lavoie, T. B.; Villafranca, J. J.; Matsueda, G. R.; Novotny, J., Design of heterotetrameric coiled coils: Evidence for increased stabilization by Glu(-)-Lys(+) ion pair interactions. *Biochemistry* **1996**, *35* (9), 2824-2829.
- (79) Dotan, N.; Arad, D.; Frolow, F.; Freeman, A., Self-assembly of a tetrahedral lectin into predesigned diamondlike protein crystals. *Angew. Chem., Int. Ed. Engl.* **1999**, *38* (16), 2363-2366.
- (80) Zhang, L.; Bailey, J. B.; Subramanian, R. H.; Tezcan, F. A., Hyperexpandable, self-healing macromolecular crystals with integrated polymer networks. *Nature* **2018**, *557* (7703), 86-91.

- (81) Mikkila, J.; Anaya-Plaza, E.; Liljestrom, V.; Caston, J. R.; Torres, T.; de la Escosura, A.; Kostianen, M. A., Hierarchical Organization of Organic Dyes and Protein Cages into Photoactive Crystals. *ACS Nano* **2016**, *10* (1), 1565-1571.
- (82) Künzle, M.; Eckert, T.; Beck, T., Metal-Assisted Assembly of Protein Containers Loaded with Inorganic Nanoparticles. *Inorg. Chem.* **2018**, *57* (21), 13431-13436.
- (83) Uchida, M.; McCoy, K.; Fukuto, M.; Yang, L.; Yoshimura, H.; Miettinen, H. M.; LaFrance, B.; Patterson, D. P.; Schwarz, B.; Karty, J. A.; Prevelige, P. E.; Lee, B.; Douglas, T., Modular Self-Assembly of Protein Cage Lattices for Multistep Catalysis. *ACS Nano* **2018**, *12* (2), 942-953.
- (84) Liu, X. F.; Theil, E. C., Ferritins: Dynamic management of biological iron and oxygen chemistry. *Acc. Chem. Res.* **2005**, *38* (3), 167-175.
- (85) Tanaka, S.; Sawaya, M. R.; Phillips, M.; Yeates, T. O., Insights from multiple structures of the shell proteins from the beta-carboxysome. *Protein Science* **2009**, *18* (1), 108-120.
- (86) Fotin, A.; Cheng, Y. F.; Sliz, P.; Grigorieff, N.; Harrison, S. C.; Kirchhausen, T.; Walz, T., Molecular model for a complete clathrin lattice from electron cryomicroscopy. *Nature* **2004**, *432* (7017), 573-579.
- (87) Fletcher, J. M.; Harniman, R. L.; Barnes, F. R. H.; Boyle, A. L.; Collins, A.; Mantell, J.; Sharp, T. H.; Antognozzi, M.; Booth, P. J.; Linden, N.; Miles, M. J.; Sessions, R. B.; Verkade, P.; Woolfson, D. N., Self-Assembling Cages from Coiled-Coil Peptide Modules. *Science* **2013**, *340* (6132), 595-599.
- (88) Patterson, D. P.; Desai, A. M.; Holl, M. M. B.; Marsh, E. N. G., Evaluation of a symmetry-based strategy for assembling protein complexes. *RSC Adv.* **2011**, *1* (6), 1004-1012.
- (89) Patterson, D. P.; Su, M.; Franzmann, T. M.; Sciore, A.; Skiniotis, G.; Marsh, E. N. G., Characterization of a highly flexible self-assembling protein system designed to form nanocages. *Protein Sci.* **2014**, *23* (2), 190-199.
- (90) Lai, Y. T.; Cascio, D.; Yeates, T. O., Structure of a 16-nm Cage Designed by Using Protein Oligomers. *Science* **2012**, *336* (6085), 1129-1129.

- (91) Leaver-Fay, A.; Tyka, M.; Lewis, S. M.; Lange, O. F.; Thompson, J.; Jacak, R.; Kaufman, K.; Renfrew, P. D.; Smith, C. A.; Sheffler, W.; Davis, I. W.; Cooper, S.; Treuille, A.; Mandell, D. J.; Richter, F.; Ban, Y. E. A.; Fleishman, S. J.; Corn, J. E.; Kim, D. E.; Lyskov, S.; Berrondo, M.; Mentzer, S.; Popovic, Z.; Havranek, J. J.; Karanicolas, J.; Das, R.; Meiler, J.; Kortemme, T.; Gray, J. J.; Kuhlman, B.; Baker, D.; Bradley, P., Rosetta3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. In *Methods in Enzymology, Vol 487: Computer Methods, Pt C*, Johnson, M. L.; Brand, L., Eds. 2011; pp 545-574.
- (92) Maity, B.; Abe, S.; Ueno, T., Observation of gold sub-nanocluster nucleation within a crystalline protein cage. *Nat. Commun.* **2017**, *8*, 9.
- (93) Huang, X. L.; Chisholm, J.; Zhuang, J.; Xiao, Y. Y.; Duncan, G.; Chen, X. Y.; Suk, J. S.; Hanes, J., Protein nanocages that penetrate airway mucus and tumor tissue. *Proc. Nat. Acad. Sci USA* **2017**, *114* (32), E6595-E6602.
- (94) Huang, X. L.; Zhuang, J.; Chung, S. W.; Huang, B. W.; Halpert, G.; Negron, K.; Sun, X. R.; Yang, J.; Oh, Y.; Hwang, P. M.; Hanes, J.; Suk, J. S., Hypoxia-tropic Protein Nanocages for Modulation of Tumor- and Chemotherapy-Associated Hypoxia. *ACS Nano* **2019**, *13* (1), 236-247.
- (95) Wörsdörfer, B.; Woycechowsky, K. J.; Hilvert, D., Directed Evolution of a Protein Container. *Science* **2011**, *331* (6017), 589-592.
- (96) Sasaki, E.; Bohringer, D.; van de Waterbeemd, M.; Leibundgut, M.; Zschoche, R.; Heck, A. J. R.; Ban, N.; Hilvert, D., Structure and assembly of scalable porous protein cages. *Nat. Commun.* **2017**, *8*, 10.
- (97) Raman, S.; Machaidze, G.; Lustig, A.; Aebi, U.; Burkhard, P., Structure-based design of peptides that self-assemble into regular polyhedral nanoparticles. *Nanomed.-Nanotechnol. Biol. Med.* **2006**, *2* (2), 95-102.
- (98) Yang, Y. K.; Ringler, P.; Muller, S. A.; Burkhard, P., Optimizing the refolding conditions of self-assembling polypeptide nanoparticles that serve as repetitive antigen display systems. *J. Struct. Biol.* **2012**, *177* (1), 168-176.

- (99) Indelicato, G.; Wahome, N.; Ringler, P.; Muller, S. A.; Nieh, M. P.; Burkhard, P.; Twarock, R., Principles Governing the Self-Assembly of Coiled-Coil Protein Nanoparticles. *Biophys. J.* **2016**, *110* (3), 646-660.
- (100) Patterson, D. P.; Schwarz, B.; El-Boubbou, K.; van der Oost, J.; Prevelige, P. E.; Douglas, T., Virus-like particle nanoreactors: programmed encapsulation of the thermostable CelB glycosidase inside the P22 capsid. *Soft Matter* **2012**, *8* (39), 10158-10166.
- (101) O'Neil, A.; Reichhardt, C.; Johnson, B.; Prevelige, P. E.; Douglas, T., Genetically Programmed In Vivo Packaging of Protein Cargo and Its Controlled Release from Bacteriophage P22. *Angew. Chem., Int. Ed. Engl.* **2011**, *50* (32), 7425-7428.
- (102) Patterson, D. P.; Schwarz, B.; Waters, R. S.; Gedeon, T.; Douglas, T., Encapsulation of an Enzyme Cascade within the Bacteriophage P22 Virus-Like Particle. *ACS Chem. Biol.* **2014**, *9* (2), 359-365.
- (103) Hempstead, P. D.; Yewdall, S. J.; Fernie, A. R.; Lawson, D. M.; Artymiuk, P. J.; Rice, D. W.; Ford, G. C.; Harrison, P. M., Comparison of the three-dimensional structures of recombinant human H and horse L ferritins at high resolution. *J. Molec. Biol.* **1997**, *268* (2), 424-448.

Figures

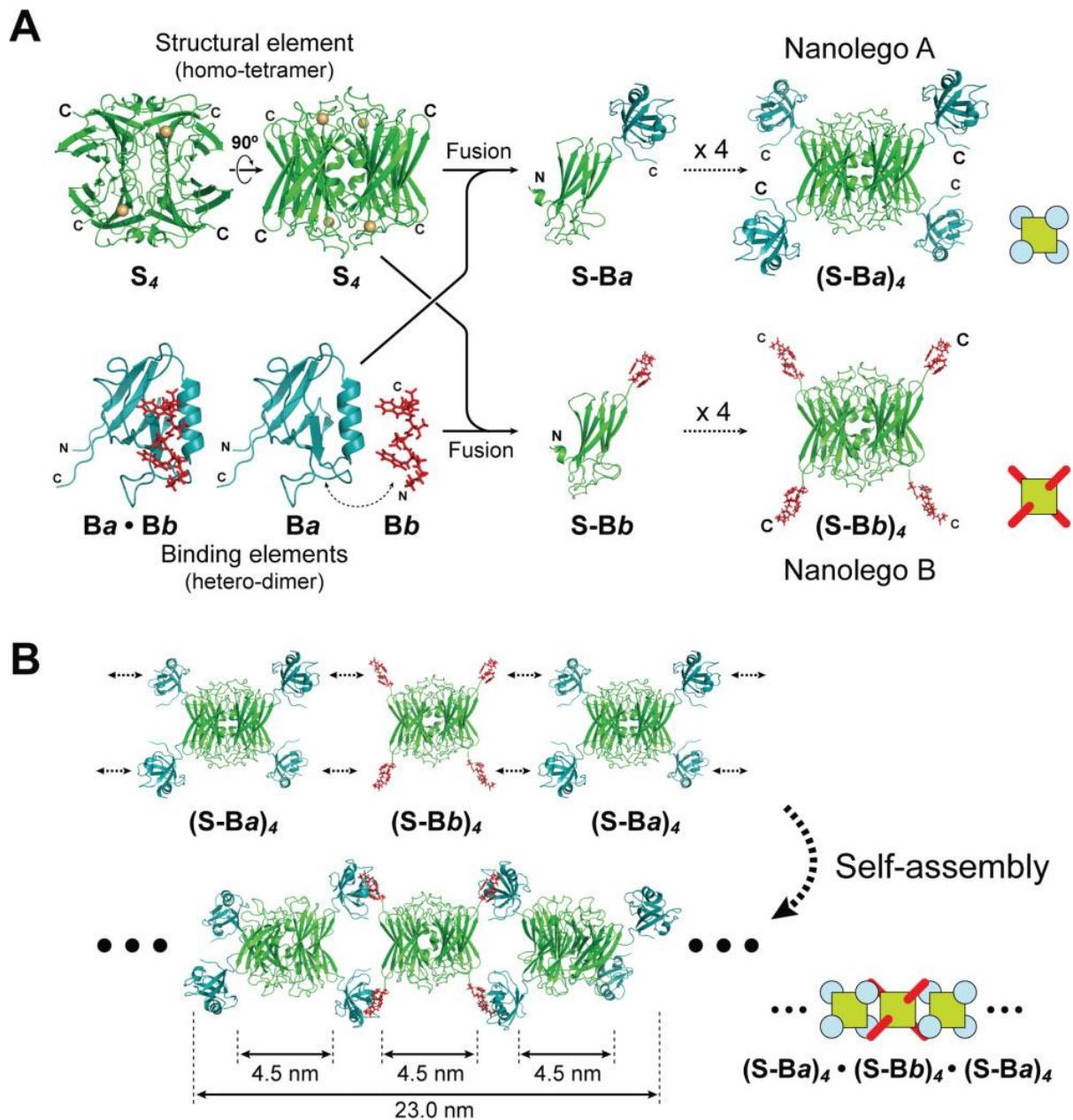


Fig.1. (a) Concept to construct Nanolego A and Nanolego B building blocks from the homotetrameric S_4 protein (superoxide reductase) modified with peptides Ba and Bb (PDZ domain peptide and PDZ-binding peptide) by creating fusion proteins using the subunit C-termini (shown). (b) Pairwise linear self-assembly leads to nanofibril formation. Reproduced from ref.¹ with permission of John Wiley and Sons.

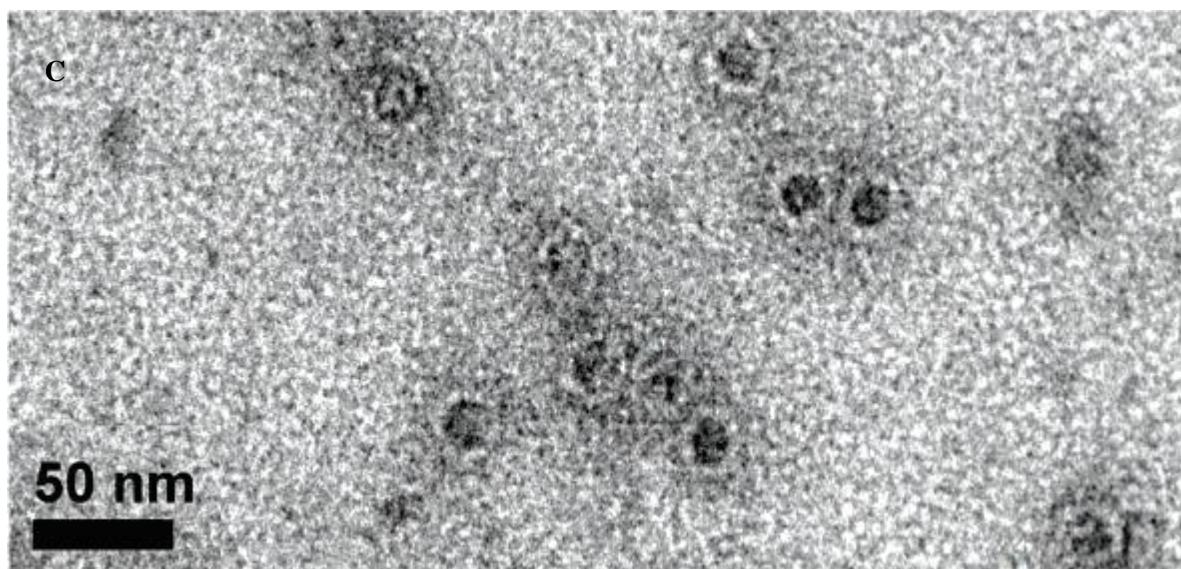
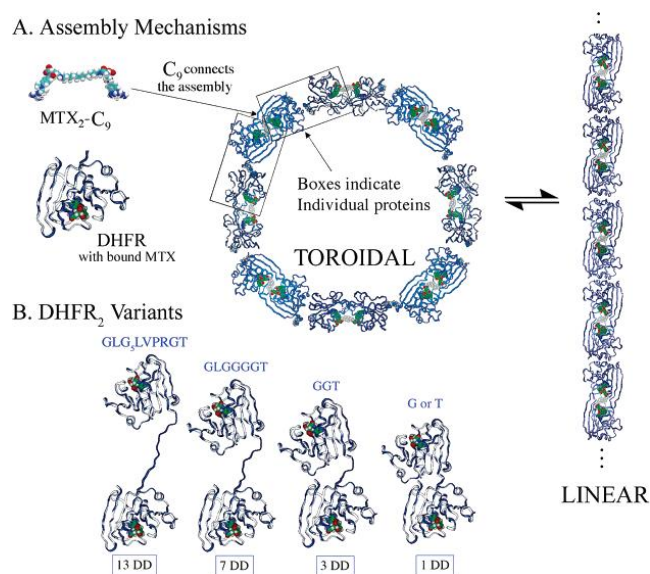


Fig.2. (a) Nanoring (toroid) formation using (bis-methotrexate MTX₂-C₉ binding to *E. coli* dihydrofolate reductase variants with extended spacers between the two subunits (b), (c) TEM image showing nanorings assembled in a solution of 1DD-G with MTX₂-C₉. Reproduced with permission from ref. ⁴³ Copyright 2006 American Chemical Society.

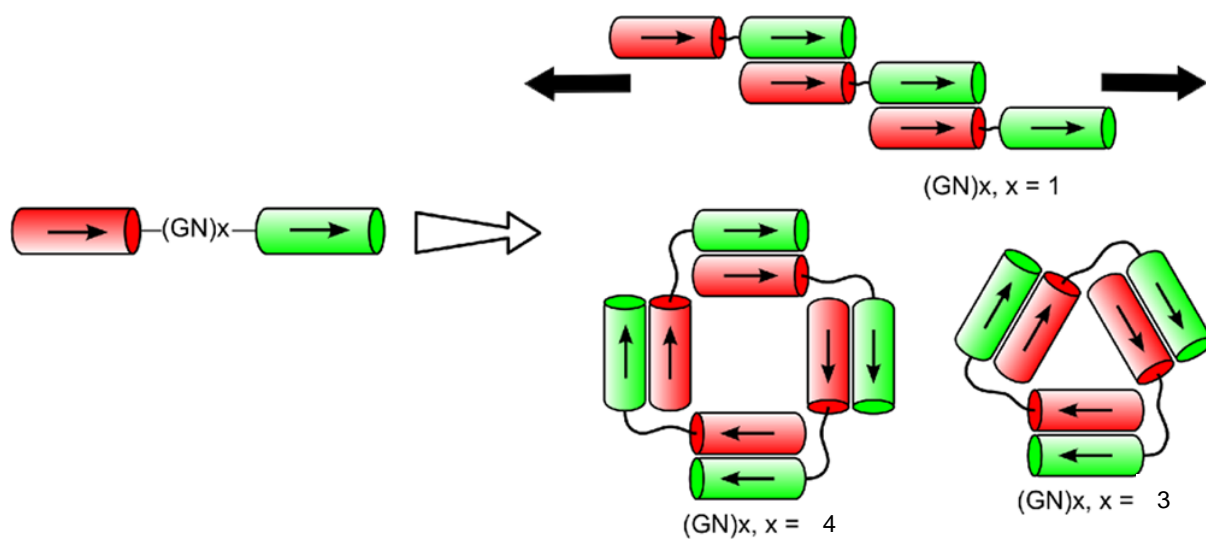


Fig.3. Dimeric coiled coil peptides with a flexible $(GN)_x$ linker between the two complementary helical peptides can form extended fibrils ($x=1$), triangular trimers ($x=3$) or square tetramers ($x=4$). Reproduced with permission from ref. ¹⁴ Copyright 2012 American Chemical Society.

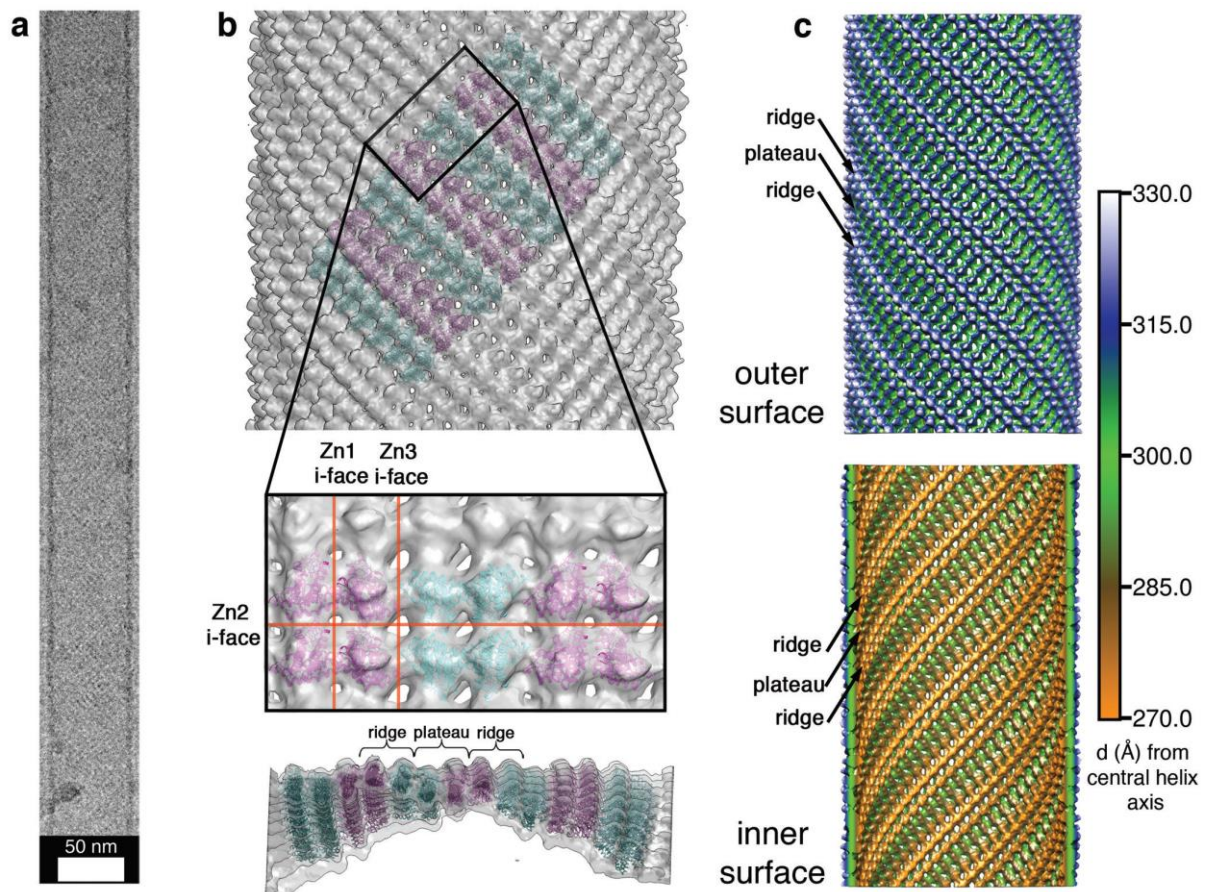


Fig.4. Nanotube structure from assembly of a designed cytochrome-based protein subunit.¹⁹

(a) Cryo-TEM image of a representative nanotube, (b) Helical arrangement of tetrameric proteins into a nanotube structure stabilized by zinc ion coordination at the interfaces (i-faces) shown, Zn1 for example denotes the dimer of C_2 -symmetric dimers forming the tetrameric building block. (c) Models for the outer (top) and inner (bottom) nanotube surfaces showing ridges and plateaus as shown side-on in (b). Reprinted by permission of Springer Nature from ref. ¹⁹ Copyright 2012.

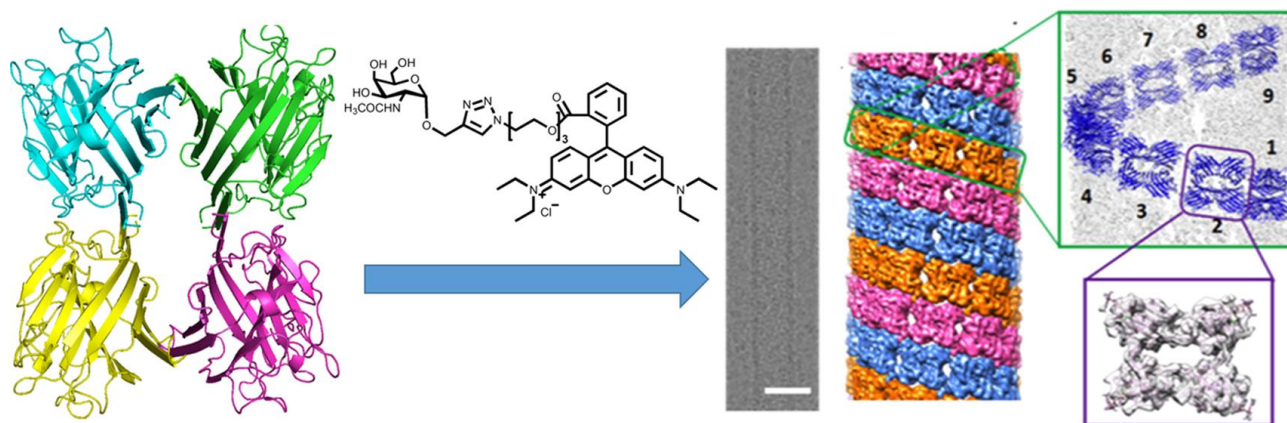


Fig.5. Soybean agglutinin (SBA) is a homotetrameric protein with D_2 symmetry. Addition of designed ligands incorporating a galactosamine (shown) or galactopyranoside unit able to bind SBA linked to a Rhodamine B unit able to undergo π - π stacking interactions which drives protein association, leading to helical wrapping into nanotubes (scale bar= 25 nm). Reproduced with permission from ref.¹⁵ Copyright 2016 American Chemical Society.

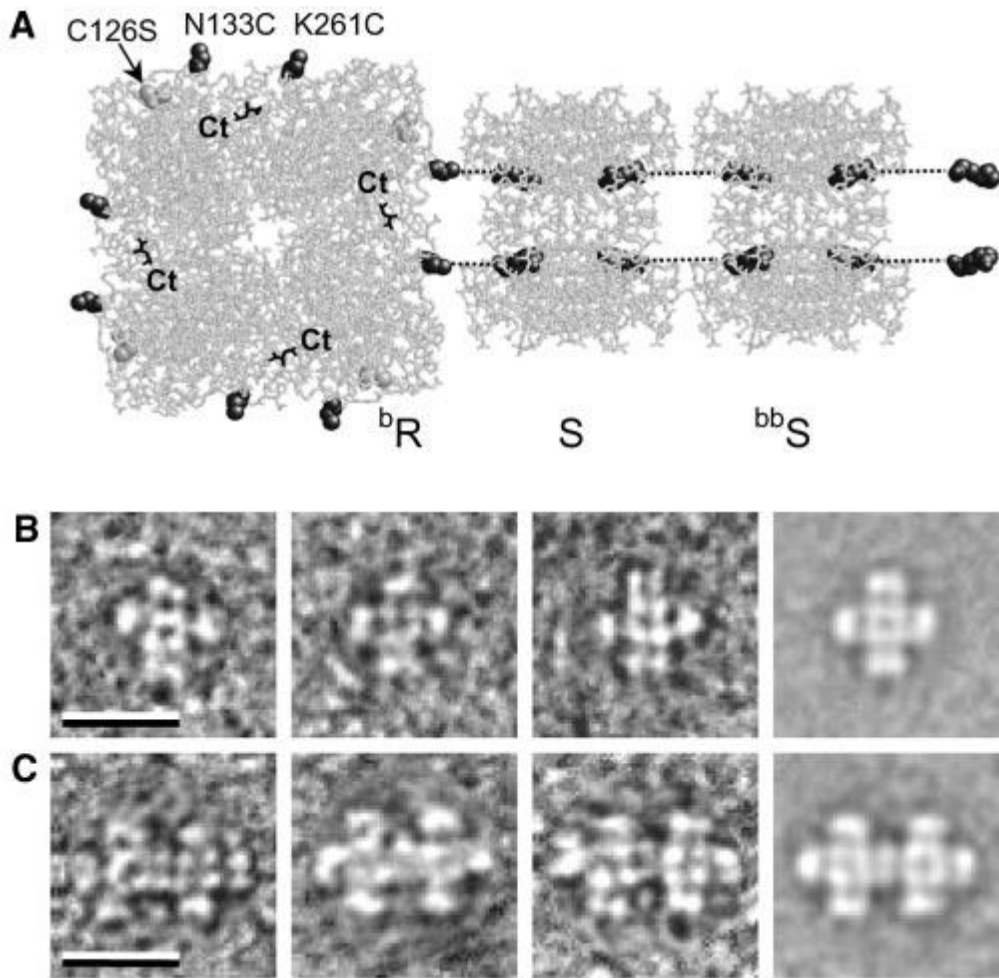


Fig.6. (a) Schematic to show modification of protein RhuA with eight biotins (bR), showing two of them binding to streptavidin (S), this in turn binding to biotinylated streptavidin linkers(^{bb}S). The C terminal (Ct) units are highlighted, these are sites for hexahistidine tagging. (b,c) Representative TEM images of aggregates of bR and $^bR.S_4$ on lipid monolayers. The scale bars indicate 200 nm. From ref.⁷³ Reprinted with permission from AAAS.

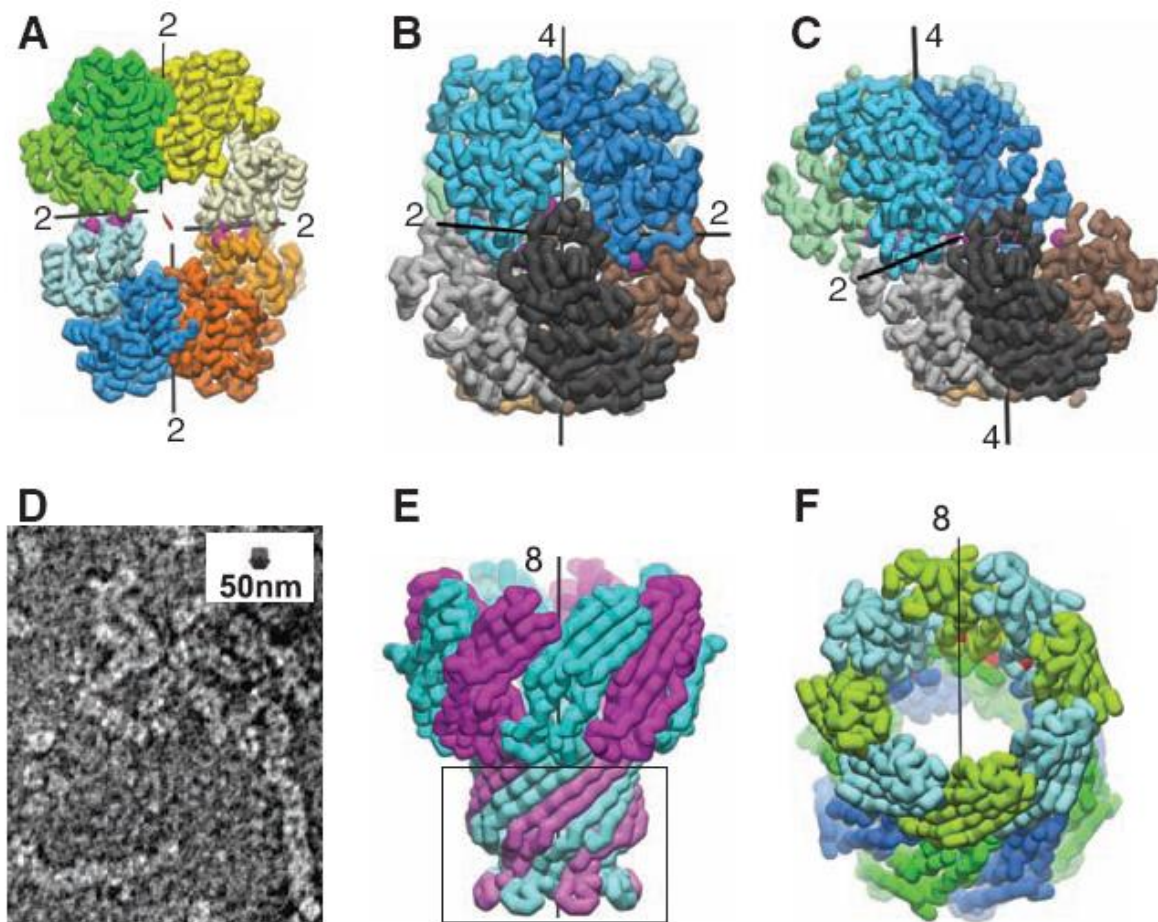


Fig.7. Modification of protein interfaces (mutations marked by purple spheres) to favour small oligomers.⁷⁴ (a) Crystal structure of C_2 -symmetric UroA tetramer showing the two-fold molecular symmetry axis (red) and four local twofold axes relating the cores (black lines), (b) Octamer formed by dimerization of RhuA dimer with D_4 symmetry, (c) RhuB octamer with C_2 symmetry, (d) Negative stain TEM image of RhuE showing assembly of fibers, the inset scale bar shows a RhuA octamer to scale, (e) Native mycobacterial porin, with the deleted membrane-immersion part indicated by the box, giving MypA, (f) D_8 -symmetric assembly of two MypA molecules (top and bottom rings). The positions of the 52-residue deletions are marked by red spheres. From ref. ⁷⁴ Reprinted with permission from AAAS.

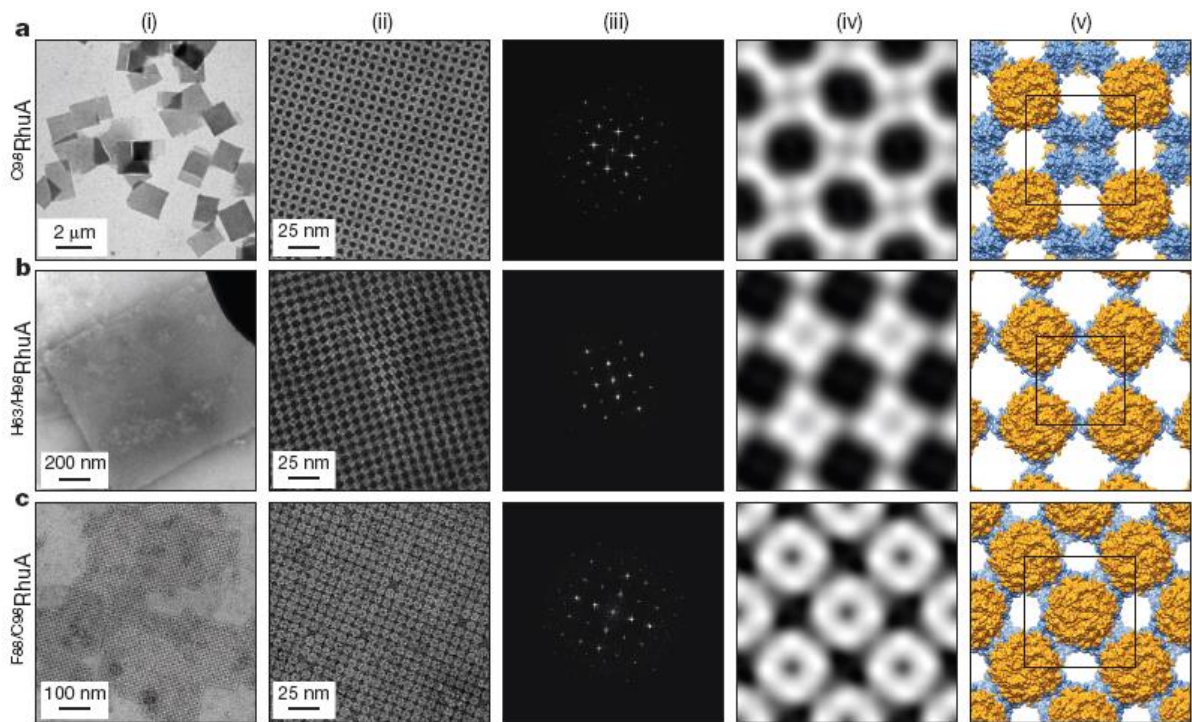


Fig.8. Distinct 2D lattices formed by the indicated RhuA mutants. (i) Low magnification TEM images, (ii) High magnification TEM images, (iii) Fourier transforms of images in column (ii), (iv) Reconstructed 2D images from the Fourier transforms, (v) Structural models based on (iv). Reprinted by permission of Springer Nature from ref.²⁰ Copyright 2016.

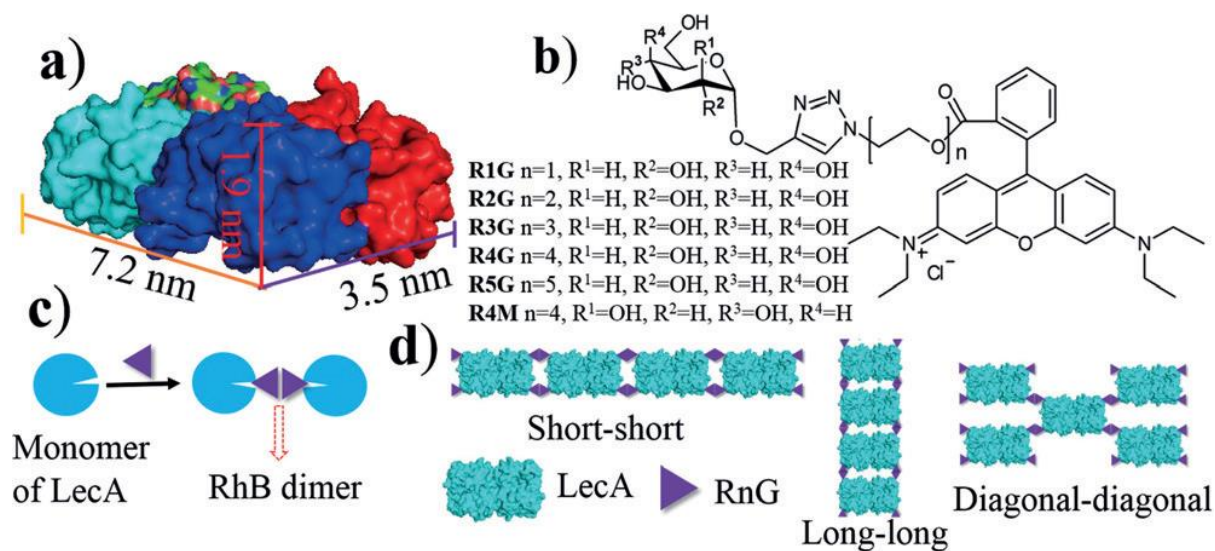


Fig.9. Galactose-based ligand-driven assembly of LecA homotetrameric proteins. (a) Structure of the LecA tetramer (protein data bank pdb id 4LKD), (b) Chemical structures of ligands RnG (n = 1 to 5) and R4M, (c) Illustration of dimerization, (d) Possible arrangements of LecA/RnG giving rise to different 1D and 2D nanostructures. Reproduced from ref. ¹⁶ with permission of John Wiley and Sons.

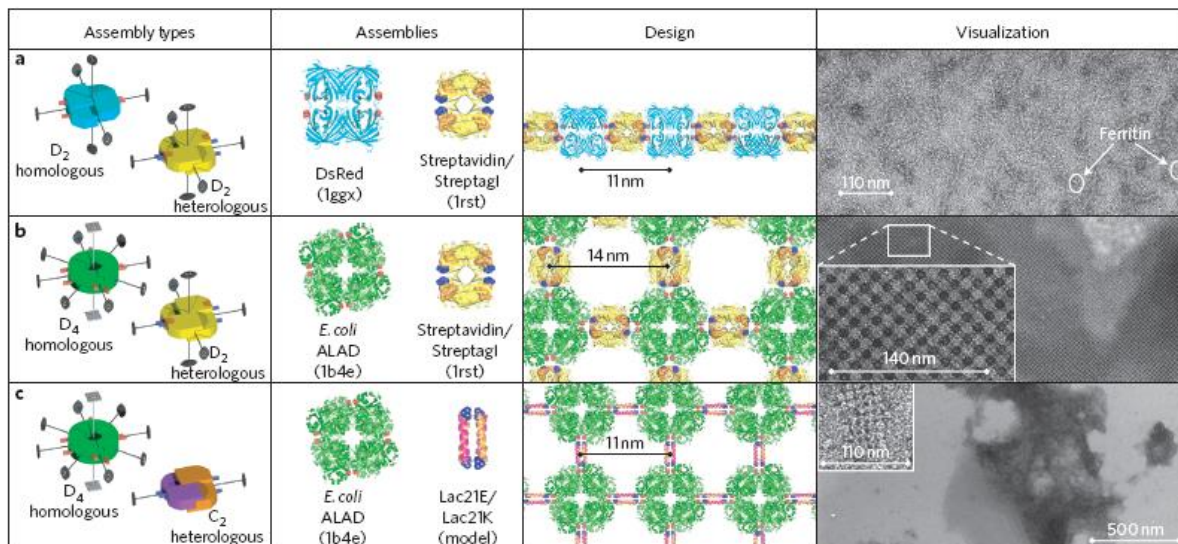


Fig.10. Fusion of protein assembly elements leads to 1D and 2D superstructures (termed crystalins). Left column: Schematic of assembly units showing symmetry elements, Second column: Protein/peptide components incorporated in fusion proteins, Third column: Designed structures, Right column: TEM images of observed structures. Reprinted by permission of Springer Nature from ref. ⁴ Copyright 2011.

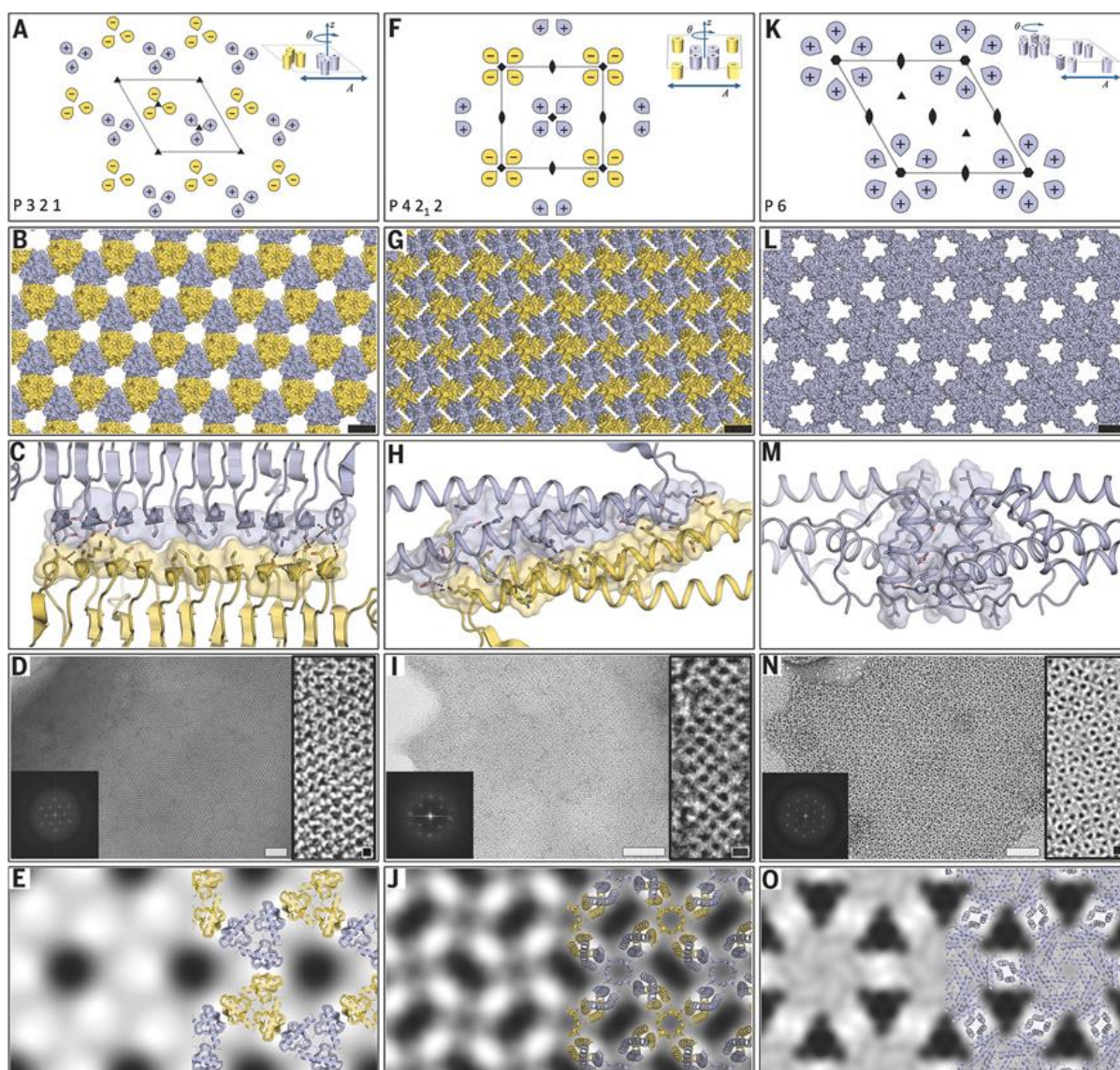


Fig.11. Creation of designed 2D lattices by protein design. (A, F, K) targeted lattices with (inset) protein subunit arrangements, (B, G, L) models for designed proteins packed into the lattices shown above, (C, H, M) designed interface structure for the corresponding lattice structure in the same column, (D, I, N) cryo-TEM images of expressed protein lattices (white scale bars = 50 nm, black scale bars = 5 nm), with inset Fourier transforms (E, J, O).

Calculated projection maps (14 Å or 15 Å resolution) with overlaid protein designs shown on the right of each image. From ref. ⁷⁵ Reprinted with permission from AAAS.

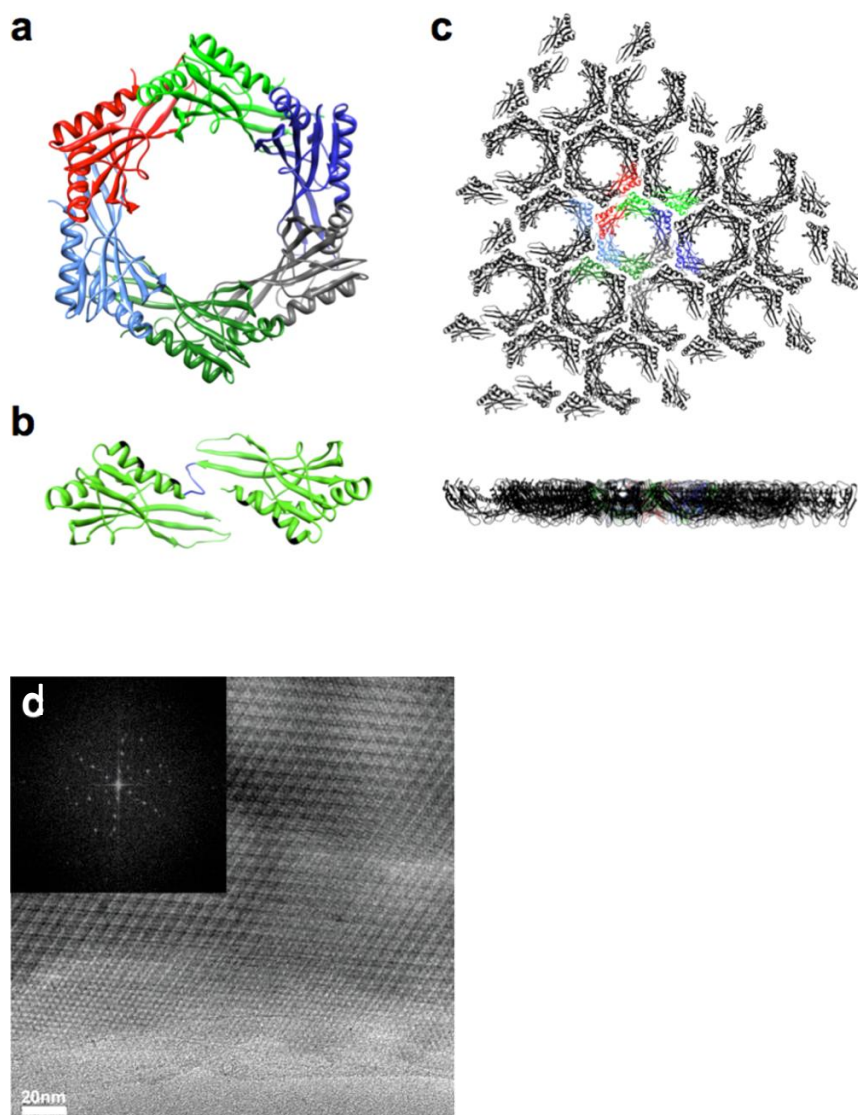


Fig.12. 2D lattices with $p6$ symmetry from a modified hexameric *S. typhimurium* STM4215 protein TTM. (a) Ribbon structure showing native hexameric structure, (b) A dimer linked by an introduced short six-residue sequence (blue) with Rosetta designed modified interfaces (shown in black), (c) Top and side views of the expected 2D lattice with individual TTM dimers shown in different colours, (d) TEM image after incubating a protein sample with CaCl_2 (inset: Fourier transform image showing hexagonal symmetry). Reproduced with permission from ref. ⁷⁶ Copyright 2015 American Chemical Society.

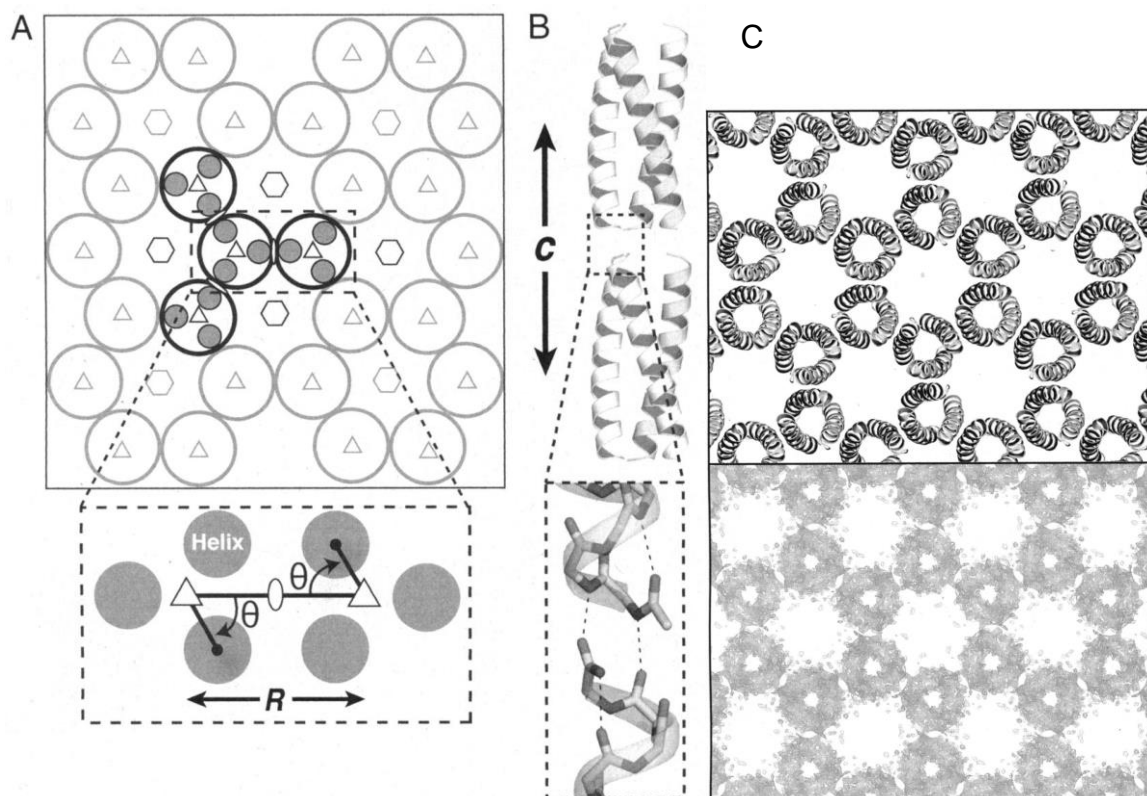


Fig.13. Design of a honeycomb lattice from homotrimeric coiled coils. (a) Schematic of the p6 structure along with lattice symmetry elements, the parameters θ and R adjusted in the peptide design are shown in the bottom scheme, (b) Two layers of the peptides showing H-bonding at the interlayer interface facilitated by fixing the unit cell length c , (c) Electron density map from a single crystal structure (bottom) compared to a model structure (top). Reproduced from ref. ²⁴ with permission of the authors.

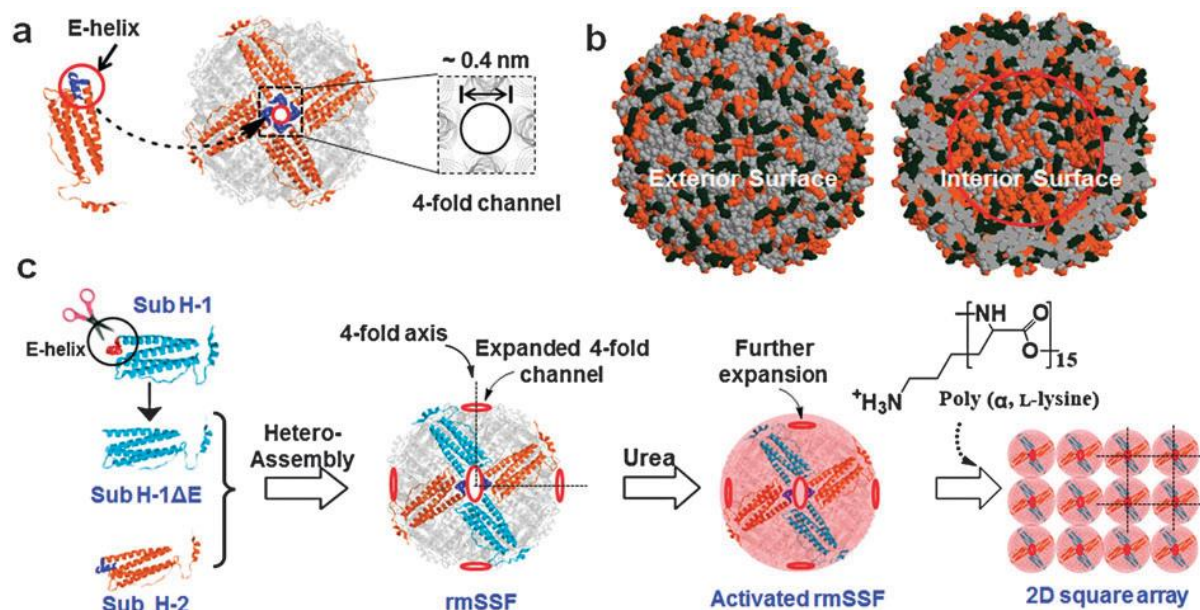


Fig.14. (a) Schematic showing a ferritin C₄-symmetry axis with pore, based on assembly of subunits shown on left with E-helix highlighted, (b) Surface charges on ferritin – orange spheres indicates negative charges and black spheres are positive charges, (c) Showing strategy to expand pore size by E-helix deletion from H-1 subunits in reconstructed mature soybean seed ferritin (rmSSF). The expanded pore size enables ingress of poly(L-lysine) which links proteins into a square array via electrostatic interactions. From Chemical communications by Royal Society of Chemistry (Great Britain). Reproduced with permission of Royal Society of Chemistry from ref. ²⁵

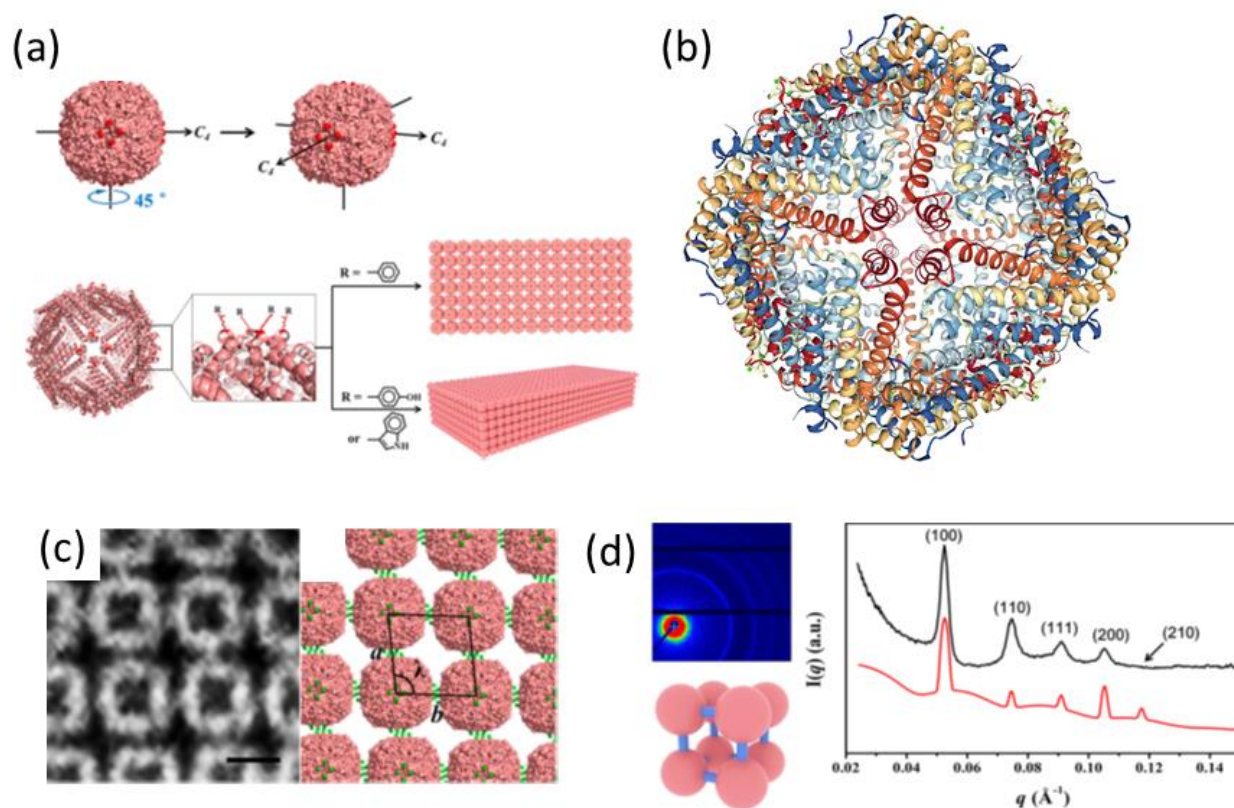


Fig.15. 2D and 3D assemblies by aromatic substitution within the 24-subunit ferritin protein cage structure. (a) Schematic of ferritin structure showing C_4 rotation axes and substitution sites near these axes, leading to 2D and 3D lattices depending on the aromatic residue substituted at Glu162. (b) Showing one of the 4-fold symmetry axes of human H-chain ferritin (pdb file 2FHA),¹⁰³ (c) reconstructed image from FFT of a TEM image of a 2D oblique (rhombic) lattice for the phenylalanine substituted protein assembly, (d) SAXS pattern and one-dimensional intensity profile with indexed reflections corresponding to a simple cubic packed 3D structure (shown) for the tyrosine-substituted protein assembly. Parts (a), (c), (d) reproduced with permission from ref. ¹⁸ Copyright 2018 American Chemical Society.

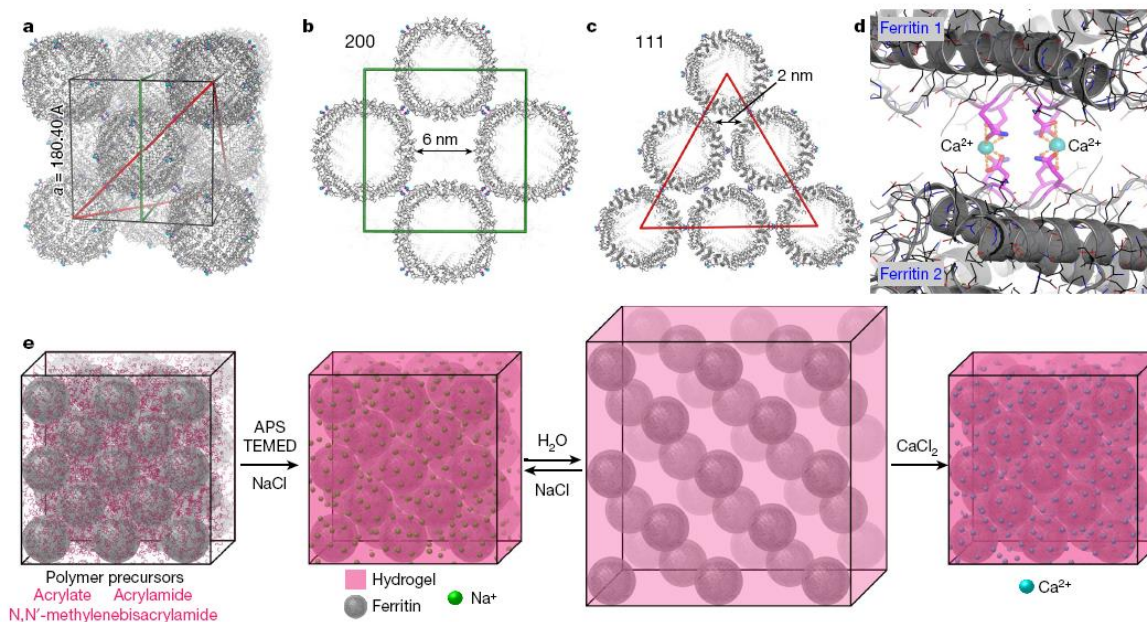


Fig.16. Concept of swellable protein-embedded polymer hydrogel crystals. (a)-(c) Showing fcc packing in ferritin crystals (Protein Data bank Identifier, pdb id 6B8F), (d) Ca^{2+} - mediated interactions leading to the packing of ferritin proteins in the crystal lattice, (e) Schematic of polymerization around the ferritin lattice scaffold to produce a reversibly swellable hybrid polymer-protein crystal hydrogel structure. Reprinted by permission of Springer Nature from ref. ⁸⁰ Copyright 2018.

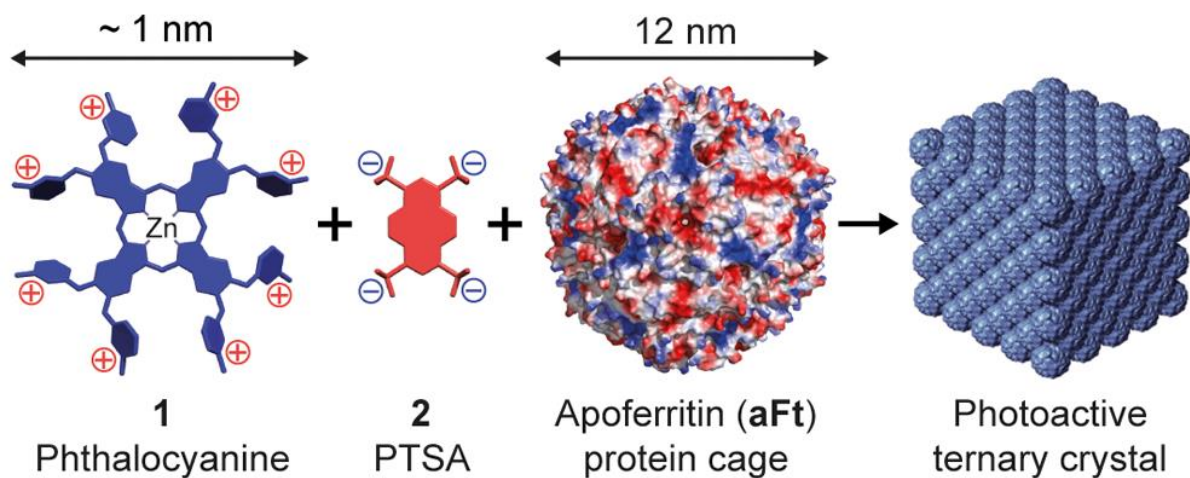
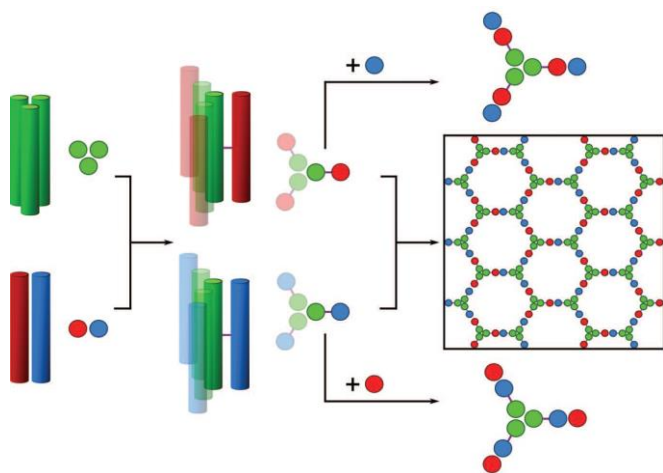


Fig.17. Complexes form between a zinc phthalocyanine derivative **1** and the tetra-anionic pyrene derivative PTSA (1,3,6-8-pyrene tetra-sulfonic acid) **2**. These complexes bind to anionic patches on the apoferritin protein surface, leading to the formation of cubic crystals which retain the photo-activity of the phthalocyanine dye. Reproduced with permission from ref. ⁸¹ Copyright 2016 American Chemical Society.

(a)



(b)

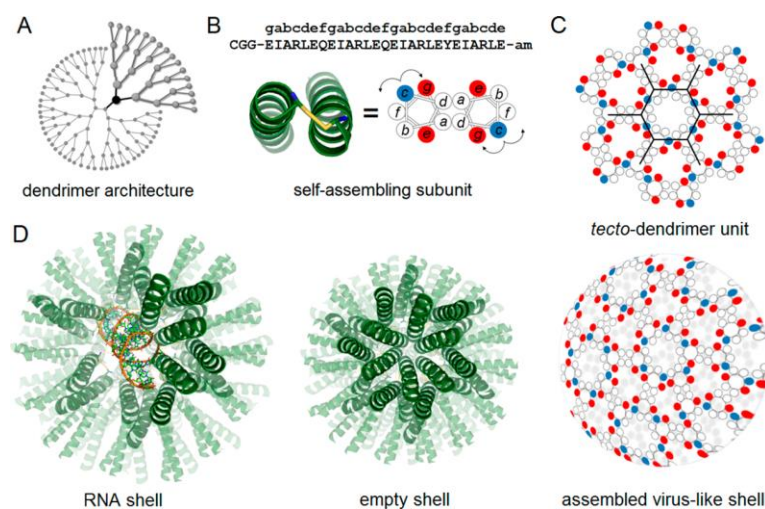


Fig.18. (a) Schematic for coiled coil peptide assembly designed to self-assemble into a honeycomb lattice (which is observed to curve into a cage structure).⁸⁷ Left: a homotrimeric coiled coil is linked via cysteine disulphide cross-linking to a homodimeric coiled coil. Mixing of either the top building block (centre, green and red) termed Hub A with coiled coil module B (basic coil peptide, blue) or Hub B (centre bottom 3-arm structure, green and blue) and module A (acidic coil peptide, red) leads to the formation of a honeycomb lattice (right). (b) Design of a dendrimer-like coiled coil peptide which forms a cage structure.²⁶ A) Dendrimer architecture, B) cysteine-linked (yellow connector) coiled coil dimer, red and blue circles indicate glutamate and arginine residues respectively (c) Expected honeycomb lattice, (D) Model for RNA-filled capsule, empty shell and observed virus-like cage structure. Part (a) from ref. ⁸⁷ Reprinted with permission from AAAS. Part (b) Reproduced with permission from ref. ²⁶ Copyright 2016 American Chemical Society.

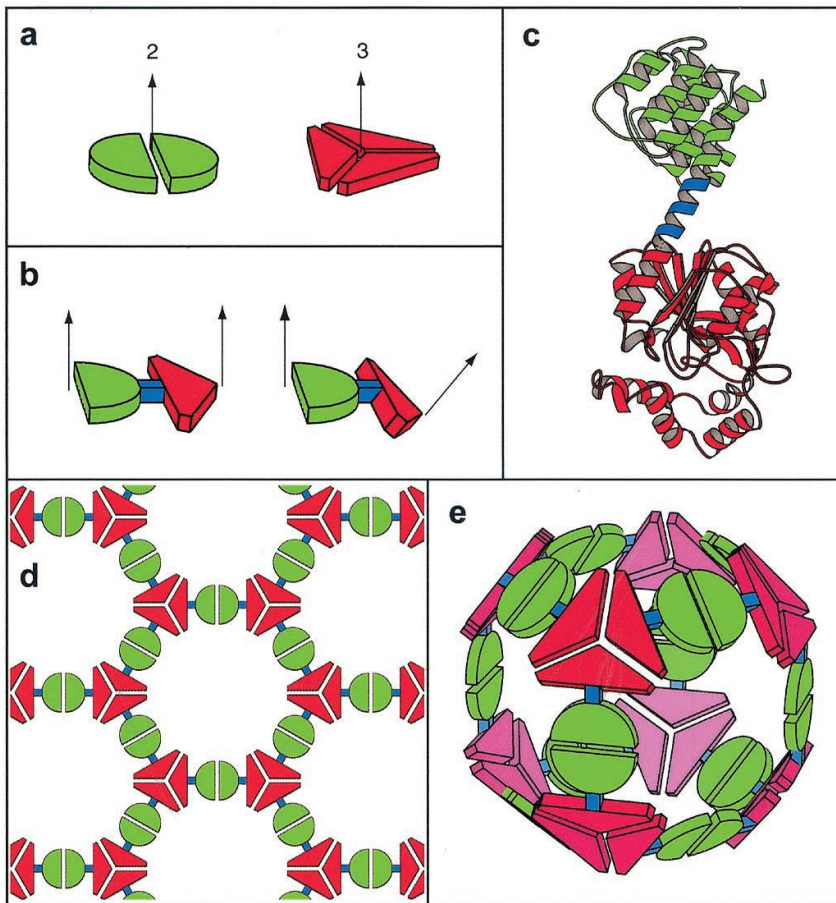


Fig.19. Fusion protein design. (a) Proteins with different sub-unit symmetries (here 2-fold and 3-fold rotation symmetry).⁶ (b) Fusion of two proteins (showing two possible geometries), (c) A ribbon diagram showing an example of a fusion construct where red and green proteins are linked by a short α -helix (blue). The fusion requires one protein to have an initial α -helix domain, the other protein must have a terminal α -helix, (d) Schematic of a 2D honeycomb lattice that assembles from flat fusion dimers, (e) Schematic of a cage structure formed when the two proteins are twisted, as shown in part (b), right. Copyright 2001 National Academy of Sciences.

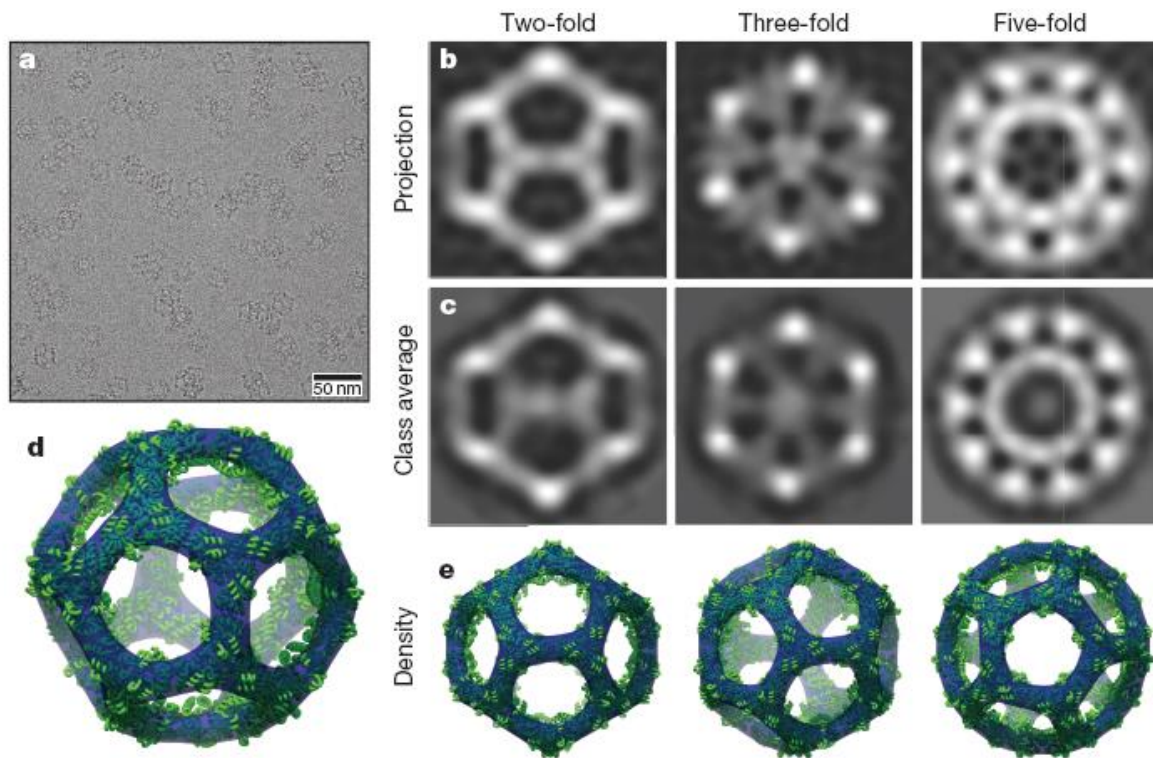


Fig.20. Icosahedral protein cages. (a) Low-magnification cryo-TEM image showing cages in different projections, (b) Back-projections of structure along different symmetry axes based on the model, (c) class averages from cryo-TEM images (bottom), (d) Three-dimensional model of the icosahedral structure, and (e) projections corresponding to images in (b,c).

Reprinted by permission of Springer Nature from ref. ⁸ Copyright 2016.

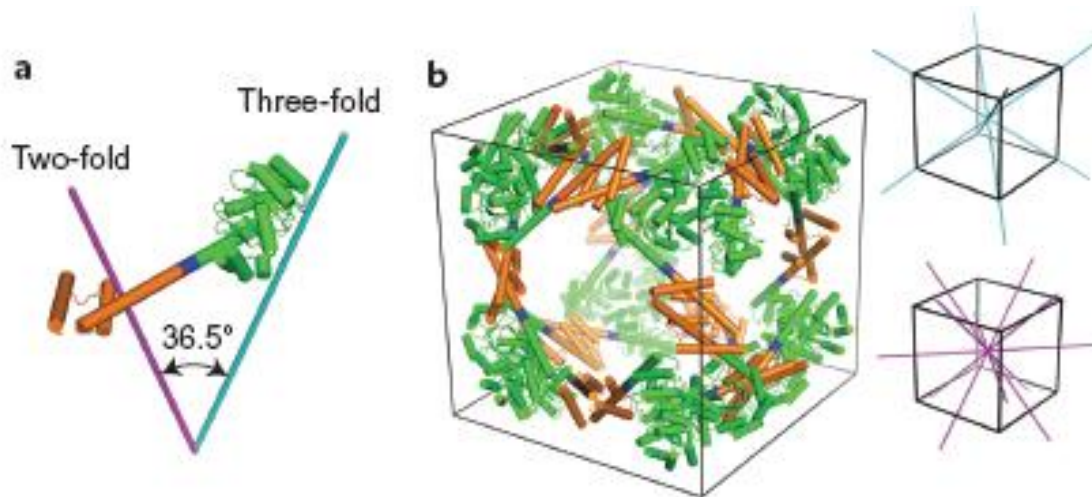


Fig.21. (a) Design of a fusion protein with appropriate oriented symmetry axes based on a trimeric protein (green) linked to a dimeric domain (orange) via a four-residue helical linker (blue), (b) Intended 24-subunit cubic cage structure with octahedral symmetry, the three-fold symmetry axes (cyan) and two-fold symmetry axes (magenta) of a cube being shown on the right. Reprinted by permission of Springer Nature from ref. ¹⁰ Copyright 2014.

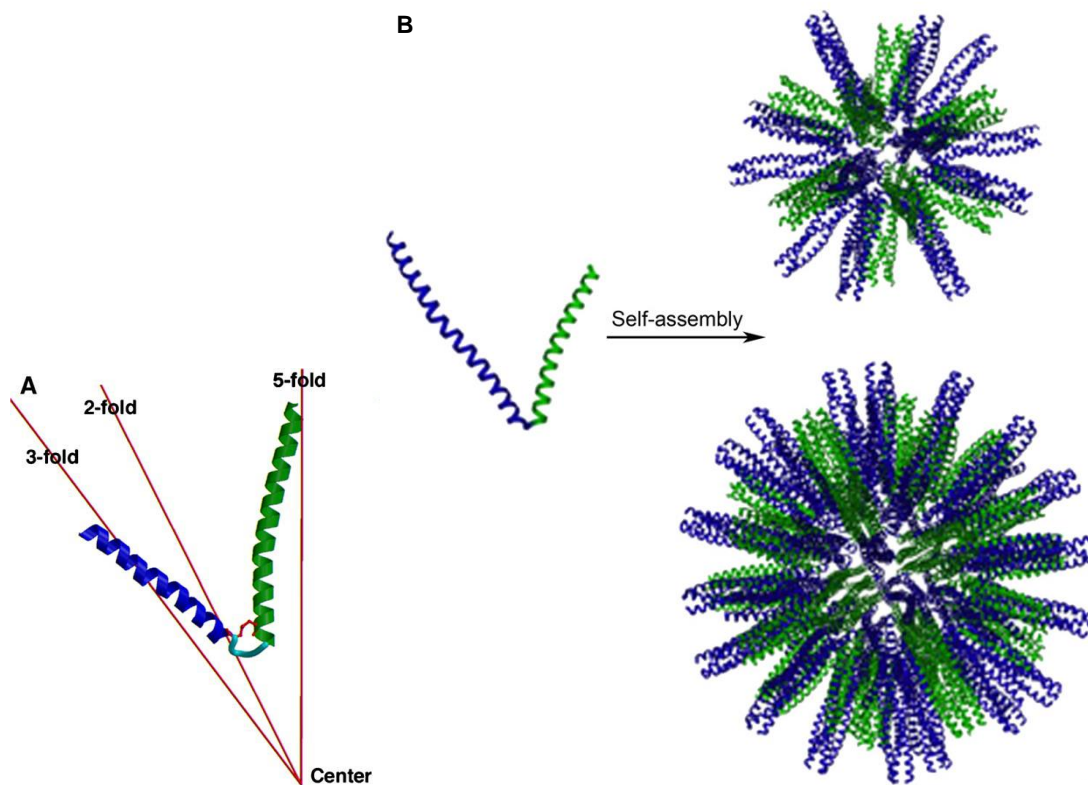


Fig.22. Polyhedral peptide nanoparticles based on (a) building block comprising two linked coiled-coil peptides designed to form pentamers (green) or trimers (blue),⁹⁷ with (b) models for their assembly into icosahedral particles, Top: Nanoparticle containing 60 peptides, Bottom: nanoparticle containing 180 peptides.⁹⁸ Reprinted from refs. ⁹⁷⁻⁹⁸ with permission from Elsevier.

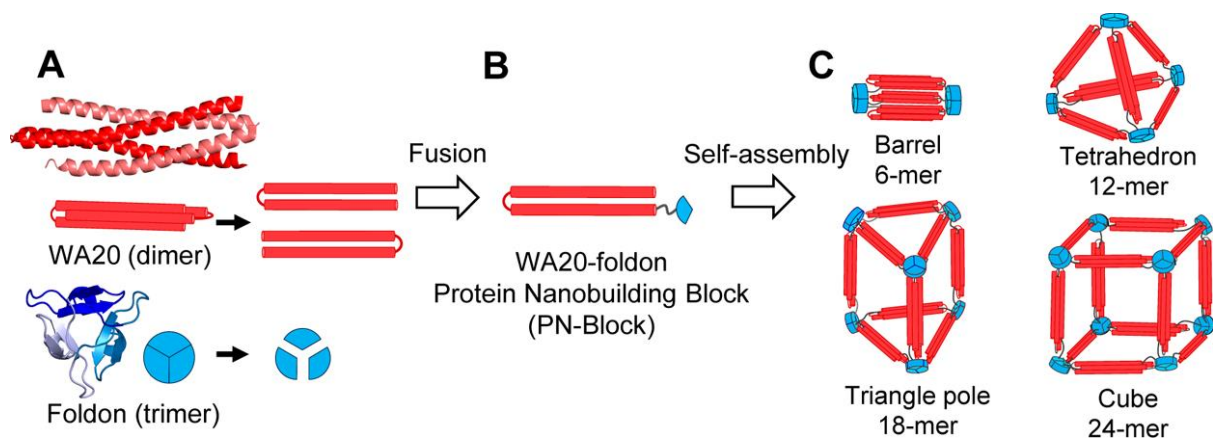


Fig.23. (a, b) Fusion protein from a designed four-helix dimer and a trimer from T4 phage fibrin, (c) Possible assemblies expected for the fusion protein, which are based on multiples of 6-mers. Reproduced with permission from ref. ¹¹ Copyright 2015 American Chemical Society.

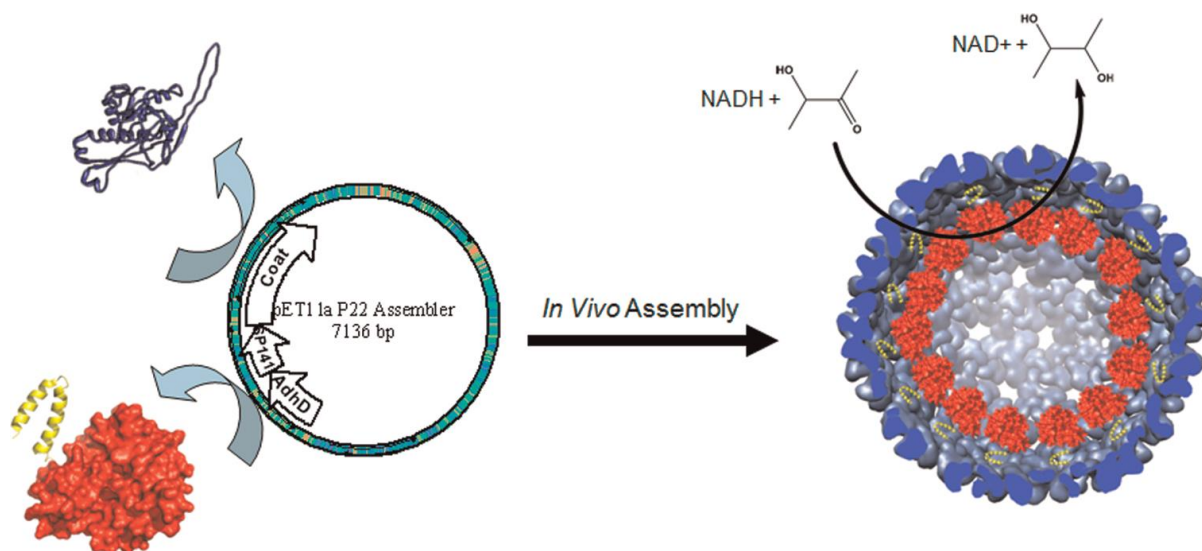


Fig.24. Incorporation of alcohol dehydrogenase into the pET expression vector for bacteriophage P22 produces the AlhD-SP conjugate (red, with C-terminal truncated scaffold protein shown in yellow) and the coat protein shown in blue leads to assembly of virus-like particles shown on the right, decorated with enzymes on the interior with model enzymatic activity shown (NAD: Nicotinamide adenine dinucleotide). Reproduced with permission from ref. ¹² Copyright 2012 American Chemical Society.

TOC ENTRY

