

## REVIEW ARTICLE

# Synthetic biology principles for the design of protein with novel structures and functions

 Weijun Zhou, Tamara Šmidlehner and Roman Jerala 

Department of Synthetic Biology and Immunology, National Institute of Chemistry, Ljubljana, Slovenia

## Correspondence

R. Jerala, Department of Synthetic Biology and Immunology, National Institute of Chemistry, Ljubljana, Slovenia  
 Email: roman.jerala@ki.si

Weijun Zhou and Tamara Šmidlehner contributed equally

(Received 18 February 2020, revised 29 March 2020, accepted 3 April 2020, available online 9 May 2020)

doi:10.1002/1873-3468.13796

Edited by Alfonso Valencia

Nature provides a large number of functional proteins that evolved during billions of years of evolution. The diversity of natural proteins encompasses versatile functions and more than a thousand different folds, which, however, represents only a tiny fraction of all possible folds and polypeptide sequences. Recent advances in the rational design of proteins demonstrate that it is possible to design *de novo* protein folds unseen in nature. Novel protein topologies have been designed based on similar principles as natural proteins using advanced computational modelling or modular construction principles, such as oligomerization domains. Designed proteins exhibit several interesting features such as extreme stability, designability of 3D topologies and folding pathways. Moreover, designed protein assemblies can implement symmetry similar to the viral capsids, while, on the other hand, single-chain pseudosymmetric designs can address each position independently. Recently, the design is expanding towards the introduction of new functions into designed proteins, and we may soon be able to design molecular machines.

**Keywords:** coiled coils; *de novo* protein design; modular design; protein cages; synthetic biology;  $\beta$ -sheets

Long-lasting evolutionary processes often produce complex and efficient cellular processes, biostructures and their mutual interactions. Hence, nature has served scientists as a fertile ground for the development of novel technologies, methods and scientific research fields such as synthetic biology. Combining biological, chemical and engineering aspects, synthetic biology enables scientists to reach beyond naturally evolved processes and architectures [1,2]

Synthetic biology provides opportunities not only to redesign already existing biological systems but also allows engineering novel biological systems at the level of individual molecules, pathways, networks, whole-cell or multiple cell communication [2,3]. Many ideas that were proposed two decades ago are becoming reality nowadays, and recently, we have

been witnessing great advances in the field of synthetic biology such as microbial production of biofuels [4], pharmaceuticals [5–7] and materials [8,9], construction of artificial cells [10–13], or synthetic viruses [14,15], and artificial photosynthesis [16], just to name a few.

*De novo* protein design can generate protein nanostructures of different functions and architectures that were imagined in scientists' minds rather than being created by evolution. Already in 2005, Benner and Sismour [2] foresaw *de novo* protein design as a challenging field. Since then, our understanding, tools and techniques have increased considerably and computational techniques have become an essential part in *de novo* protein design. This review will focus on pointing out selected recently published achievements

## Abbreviations

CCs, coiled coils; CCPO, coiled-coil protein origami; LOCKR, latching orthogonal cage-key proteins; SwitCCh, a peptide-based conformational switch; SPOC, split-protease-cleavable orthogonal-CC-based; MnCO, manganese-carbonyl moieties; TMV, tobacco mosaic virus.

in novel modular protein folds and assemblies and their diverse applications.

## Building blocks in modular *de novo* protein design

*De novo* protein design involves design of amino acid sequence that self-assembles into a selected structure. Folding principles of natural proteins that govern efficient packing of the compact hydrophobic core and protein domain interaction interface have been used to construct diverse protein assemblies that often exhibit extremely high stability [17–20]. Modular design based on multiple fused interacting (di- and oligomerizing) domains can bypass some of the problems of the complex *de novo* protein design which as to combine multiple weak long-range cooperative interactions related to the packing of residues in the protein core and is applicable for the construction of new protein folds based on geometric and topological arrangement of interacting modules. Smallest modules typically comprise only individual secondary structure elements (helices or strand) but can also encompass larger folding units. Arrangement of modules prone to adopt either  $\alpha$ -helix or  $\beta$ -strand can yield novel protein folds. Such approach has been used to design bionanomaterials with different functions and properties [21] or proteins based on repeats [22]. The following sections will provide some examples of applications of secondary structure building blocks applicable for modular *de novo* protein design.

### Coiled coils

Coiled coils (CCs) are naturally occurring secondary structural motifs first identified by Francis Crick [23,24], found in leucine zipper oligomers [25] and have been later found in numerous other proteins where they have different functions, for example CCs serve as molecular spacers between distant protein domains transmitting conformational changes between them [26] or functioning as a molecular ruler defining the acyl chain length of bacterial lipopolysaccharides [27].

Coiled coils are composed of two or more  $\alpha$ -helices twisted around each other (in parallel or antiparallel orientation) forming a superhelix (Fig. 1A–D) [26,28].

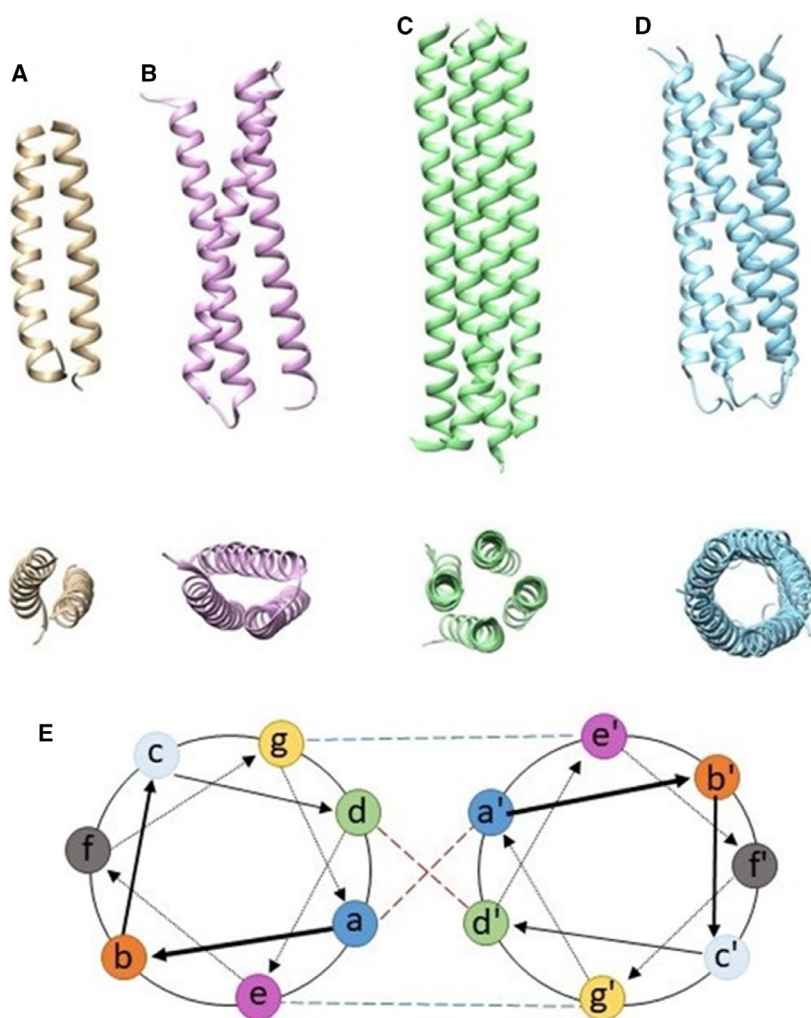
In the canonical dimeric coiled coils, the repetitive structure is clearly visible in the primary sequence, which is composed of heptad repeats (seven amino acid residues labelled abcdefg). Heptad typically comprise hydrophobic residues such as Leu or Ile at positions a and d, while residues that form interchain salt bridges are found at positions e and g (Fig. 1E) [28].

This simple pattern of hydrophobic and charged residues provides a variety of possibilities to design numerous CCs *de novo*. Sequences can be designed to define the mutual orientation of chains of CCs, either in the antiparallel [29] or parallel [30] orientation, to produce orthogonal homo- and heterometric CCs [30,31] of different oligomerization states [32–34]. Orthogonality means that a selected peptide binds only to its designated partner but no to others present in the same or other chains. The length of CCs also can be finely tuned by the number of heptad repeats or (de)stabilizing interactions [35,36]. Furthermore, changing solvent-exposed residues can yield in pH-sensitive oligomerization switches [37], while mutating inner residues can generate CCs with catalytic activity [38,39]. Another important property of CCs is their propensity to oligomerize into different oligomerization states; thus, CCs can be used to combine different protein domains of interest. For example, a multivalent display of agonist nanobodies that can trigger cancer cell death [40] was formed based on the oligomeric CC assembly. Additionally, CCs can serve as a delivery system as well, as the inner cavity of higher oligomer CCs (as shown in Fig. 1) can accommodate small molecules [41], cytotoxic drugs [42] or even encapsulate carbon nanotubes [43].

Detailed description of properties and design of CCs is provided in a review by Woolfson [44] and further comprehensive overview of a naturally occurring and *de novo* designed coiled-coil properties in review papers by Apostolovic *et al.* [45] and Wu and Collier [46], respectively. Classification of coiled coil with known tertiary structures is available in the periodic table webpage ([http://coiledcoils.chm.bris.ac.uk/ccplus/search/periodic\\_table](http://coiledcoils.chm.bris.ac.uk/ccplus/search/periodic_table)).

### $\beta$ -sheet-based proteins

$\beta$ -sheets, as another common secondary structural motif found in nature, seems to be less straightforward for the *de novo* design than CCs, both from the structural and geometric considerations and the rules governing their orthogonality.  $\beta$ -sheets are formed when several  $\beta$ -strands self-assemble into parallel, antiparallel or mixed arrangements.  $\beta$ -strands are typically shorter than CC building modules, comprising four to ten residues (in contrast to tens or hundreds of residues building CCs) and are stabilized by a network of hydrogen bonds and salt bridges between strands. The pattern of hydrogen bonds between the residues on two  $\beta$ -strands follows the pattern distinct from the one occurring in CCs. However,  $\beta$ -strand sequence also often contains a repetitive pattern with a periodicity of



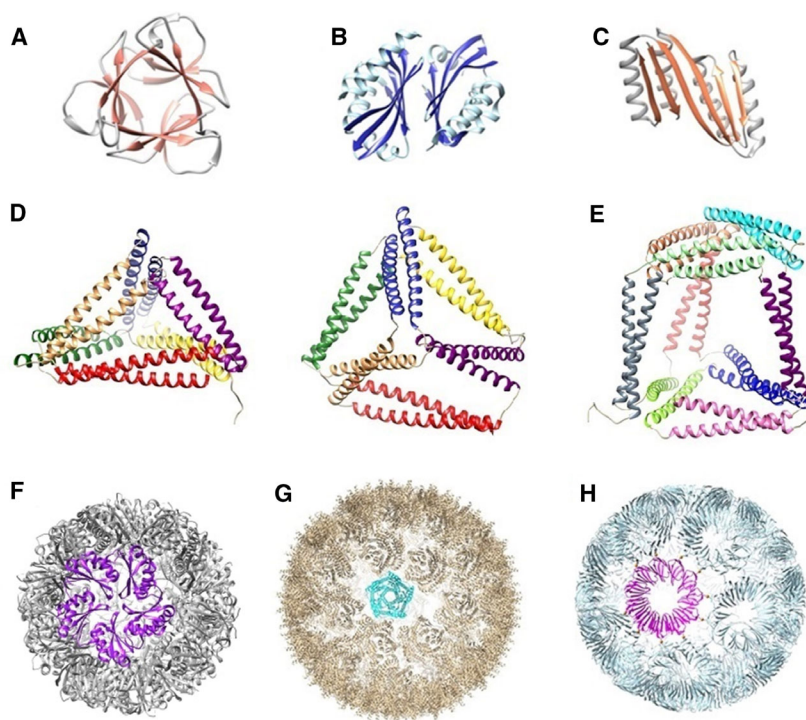
**Fig. 1.** Naturally occurring coiled coils. (A) Dimeric CC: GCN4 leucine zipper (PDB: 4DMD). (B) Trimeric CC: N-terminal domain of syntaxin-1A (PDB: 1EZ3). (C) Tetrameric CC from the surface layer protein tetrabrachion (PDB: 1FE6). (D) Pentameric CC from cartilage oligomeric matrix protein (PDB: 1MZ9). (E) Helical wheel diagram for parallel CC orientation showing interacting residues within heptad abcdefg between helices.

approx. 2, with a prominent role of hydrophobic residues such as Ile, Leu, Phe, Trp, Tyr, Thr and Val [28,47]. In addition to these observed patterns in natural proteins, it was shown that inserting charged residue into newly designed  $\beta$ -sheet proteins will hamper their aggregation [48].

Besides  $\beta$ -sheets,  $\beta$ -strands can assemble into other structures such as  $\beta$ -hairpins,  $\beta$ -turns and  $\beta$ -barrels [28] or be combined with  $\alpha$ -helices to form  $\alpha/\beta$  mixed structures [49]. Supported by computational methods [50], great progress has been made in developing novel protein folds based on  $\beta$ -strands. A unique ‘top-down symmetric deconstruction’ [51] strategy was applied to produce protein folding into the symmetric  $\beta$ -trefoil fold by gene duplication and fusion processes (Fig. 2A). The inspiration for an example of this design came from the deconstruction of human fibroblast growth factor-1, a 140 amino acid single-domain globular protein exhibiting the characteristic threefold

symmetry [52]. Similar work on  $\beta$ -trefoil was done by Meiering group whose design was inspired by the ricin protein. Using computational reconstruction and consensus sequence of  $\beta$ -trefoil subdomain module, highly thermally stable and soluble new threefold symmetric globular protein ThreeFoil was produced that has multivalent carbohydrate-binding properties [53].

Using Rosetta method, Baker group succeeded to produce proteins composed of  $\beta$ -sheets [54]. Afterwards, a *de novo* type of functional  $\beta$ -barrels was designed that is capable of selective binding of a small organic molecule confirmed by fluorescence [55]. The same group focused on designing novel  $\beta$ -sheet proteins based on their interactions with ligands. Inspired by cystatin, naturally occurring protein with four-stranded antiparallel  $\beta$ -sheets with  $\beta$ -bulges at the edges, novel protein folds were designed and more than thirty monomeric proteins were produced with melting temperature above 95 °C, twelve of which



**Fig. 2.** Structural diversity of designed proteins. (A–C) Structures of newly designed  $\beta$ -sheet-based protein folds: (A) threefold symmetric  $\beta$ -trefoil fold (PDB: 3Q7Y); (B) homodimer with curved beta-sheets (PDB: 5TPH); (C)  $\alpha/\beta$  mixed structures (PDB: 5CW9). (D,E) Single-chain CCPOs: (D) tetrahedron, side and top view; (E) trigonal prism (each CC dimer is presented in a different colour [61]). (F–H) Multiple chain protein cages: (F) dodecahedral cage (PDB: 5HPN), pentameric unit coloured purple, (G) icosahedral cage (PDB: 5MQ7), one building block coloured in cyan; (H), gold-coordinated protein cage (PDB: 6RVV), 11-mer ring coloured in magenta and Au coloured in orange.

folded into a proposed structure. The stability of novel proteins was additionally improved by disulfide bridge introduction and homodimerization. This approach presents the largest *de novo* curved  $\beta$ -sheet protein with a cavity (Fig. 2B) [56].

Further,  $\alpha/\beta$  mixed structures with six- and seven-stranded  $\beta$ -sheets were constructed by merging two  $\beta$ -sheets to form a single extended sheet, followed by amino acid sequence optimization at the newly formed strand–strand, strand–helix and helix–helix interfaces (Fig. 2C) [57].

### Assembling higher order protein nanostructures

Constructing protein architectures of a higher order by combining the presented secondary structure building blocks still requires computational modelling but also experimental validation and screening. Here are some of the latest accomplishments within this direction.

### Protein nanostructures composed of CCs

Due to their aforementioned easily controlled properties as well as their rigid and elongated shape, CCs were used as building blocks in many different 2D and 3D nanostructures.

Keating's group produced nanotriangles designed from the set of three mutually orthogonal dimeric CCs. Design includes generating three polypeptide chains that each consists of two fused nonassociating pair of CCs. Three fusion proteins assemble in a triangle shape in the two-step thermal annealing process [58]. In the design of another triangular and rectangular 2D nanostructures, ultrastable trimeric CCs were chosen as edges. Each edge was expressed separately and then ligated together into a single-chain cyclized structure through split intein chemistry. Those 2D nanostructures showed to be extremely thermostable and resistant in a harsh chemical environment [59].

When designing complex tertiary protein nanostructure built of single polypeptide chain with dimeric CCs as edges, Jerala's group designed the double Eulerian trail where CC pairs are arranged in a way that each edge is traversed by a single polypeptide chain twice, folding into a stable polyhedral structure. Between each CC-forming segment, a flexible linker of four to six amino acid residues was inserted. Such design successfully yielded novel coiled-coil protein origami (CCPO) cages that adopted tetrahedral cage-like nanostructure [60]. The second-generation CCPOs generated tetrahedra with improved folding properties and two new topologies: a four-sided pyramid and a triangular prism (Fig. 2D,E). All topologies successfully self-assemble *in vitro* and *in vivo* without triggering an immune response or any adverse effect in mice [61].



One of the advantages of a single-chain protein design over multichain assembly is the possibility to address individual desired position in order to introduce desired modifications or functionalities without symmetric constraints within the nanostructure.

Protein origamis, of tetrahedral [62] or octahedral shape, have also been produced by applying symmetry-directed oligomerizing module strategy. The octahedral protein cage can be designed by combining C3 and C4 symmetry elements. As a C3 symmetry element, trimeric esterase (TriEst) was chosen, which has been fused to a domain of CC-forming tetramer (C4 symmetry element). In this case, CCs were used as the oligomerization domains, rather than edges of octahedra, protruding out of each apex of the octahedra. The protein cages self-assembled in bacteria while retaining esterase activity [63]. As CCs in this octahedral cage protrude out into the solution, C terminus of the CC domains was fused to a maltose-binding protein that yielded enhanced production yields while retaining octahedral topology. This strategy shows high robustness and could be used in the polyvalent display of antigens for vaccine development or construction of artificial multienzyme complexes [64].

Another interesting design using CCs as building blocks was presented by Woolfson group where homotrimeric CCs were linked via disulfide bonds to a helix that forms a heterodimeric CC pair; thus, two complementary hubs were produced. When two hubs are mixed, heterodimeric CCs assembled forming a hexagonal network, which is flexible enough to eventually close up into 100 nm sphere [65]. This scaffold is particularly useful as a virus-like particle for vaccines.

## Large multichain protein cages

Multichain protein cages self-assemble in most cases as homo-oligomeric 3D structures such as virus capsids, ferritins, chaperonins and heat shock proteins. Due to their well-defined inner cavity, these supramolecular nanostructures are potentially applicable as the delivery systems. Surface of cages can be functionalized, not only to form interactions with other subunits but as well with other structures such as cell membrane; thus, the whole delivery complex is transferred to the desired location [66,67].

As those supramolecular architectures are self-assembled of multiple copies of protein monomers, selection of the monomer unit is the first step in *de novo* modular design of protein cages. The importance of symmetry in interacting monomers' conserved junctions should be emphasized since the junction needs to link the identical preceding unit and provide

the same interacting surface for the following unit [22,66,67]. There are several strategies to design those junctions, such as designing new protein–protein interfaces, genetically fusing two interacting domains, introducing bifunctional ligands or metal ions that can join interacting modular units together [67].

In an example, a circular permutation of a naturally found protein PduA (propanediol utilization), where its C terminus is fused to N terminus, changes protein assembly from 2D to a 3D structure. This protein self-assembles into symmetric hexamers, forming a 2-nm-thick molecular layer in the outer shell of the bacterial microcompartment. However, after a circular permutation, hexamer is no longer formed but rather an asymmetric pentamer that assembles into 3D dodecahedral cage with a diameter of 13 nm and 7-nm-wide inner cavity (Fig. 2F) [68]. While designing artificial protein cage, Sasaki *et al.* were inspired by the virus-like enzyme lumazine synthase from the hyperthermophilic bacterium *Aquifex aeolicus* that spontaneously assembles into icosahedral symmetric particles containing 60 identical lumazine synthase subunits (dodecahedron composed of 12 pentamers) [69]. To modify the selectivity of the capsid from its natural guest (riboflavin synthase), four residues that protrude into the capsid core were mutated to negatively charged glutamates accompanied by the introduction of additional seven negatively charged amino acids. These modifications created much larger sponge-like nanoparticles containing 180 and 360 subunits while retaining its high thermostability (Fig. 2G). The loading capacity of positively charged guest proteins has been further improved by directed evolution [70].

The symmetry-directed oligomerizing design presented in the previous section was also utilized to produce a large protein cage with Mr of approximately 2 MDa. TriEst was again used to introduce C3-symmetric building block, while the C5 symmetry component represented a *de novo* designed pentameric CCs. Two symmetry-introducing elements were genetically fused via flexible glycine-rich linkers of different lengths. This design yielded in icosahedral geometry cage composed out of 20 esterase trimers that retained its enzymatic activity, while the cage itself binds DNA and shows high thermal stability [71].

Another ultrastable and monodisperse protein cage of high molecular weight was presented by Malay *et al.* The assembly of this cage is controlled by Au(I) coordination through cysteine residues at protein–protein interfaces of the 11-mer protein rings (Fig. 2H). The disassembly of this protein cage is easily triggered by reducing agents [72]. Similarly, different ratios of zinc and iron ions that mediate monomer interactions

can generate dodecameric and hexameric cages [73]. Although multiple chain protein cage design presents challenges, recent advances in designing various *de novo* protein nanostructures provide a bright future in this field.

## Applications of *de novo* protein assemblies

*De novo* designed proteins with a defined topology structure and size can contribute to the understanding of native protein behaviour, folds and interactions [74–76]. They also show great prospects at numerous applications as functional devices for biology, biomedicine and synthetic circuits.

## Protein switch and *in vivo* designed logic circuits

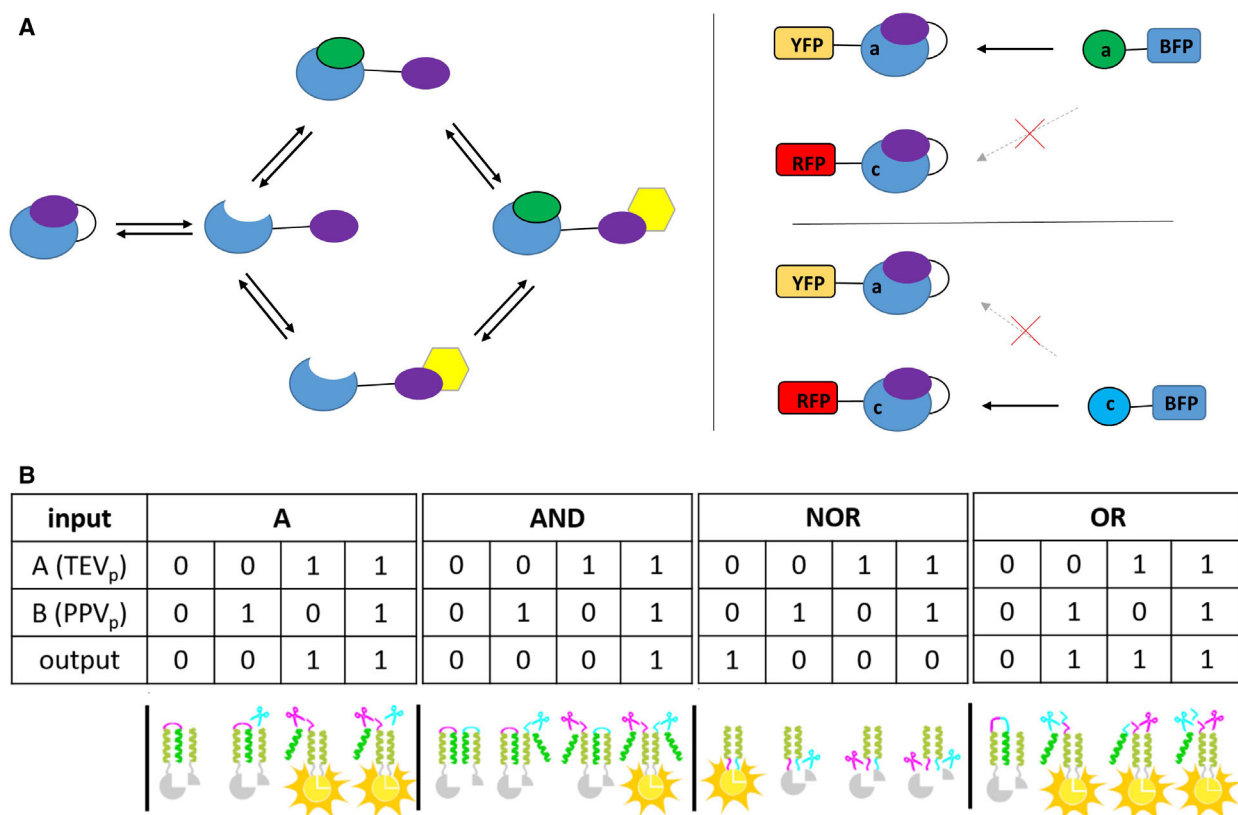
Allosteric regulation of proteins plays a key role in their biological function, while the mechanisms of conformational changes and protein interactions are still not demystified. *De novo* proteins open up new perspectives in protein structure and functional research. Recently, a bioactive *de novo* protein switch has been constructed utilizing ‘latching orthogonal cage-key proteins (LOCKR)’ strategy. The protein switch can interact either intramolecularly with a terminal ‘latch’ helix or intermolecularly with a peptide ‘key’. This *de novo* protein has two states with similar free energies, so the conformation switching can be triggered by an external input (Fig. 3A). This protein switch can be used for controlling gene expression in living cells [77]. By introducing metal-binding sites into CC peptides, a peptide-based conformational switch (SwitCCh) that can reversibly assemble into homodimer CC conformation in the presence of Zn(II) ions or at low pH has been designed [78]. The protein switch with accurately controllable topology and conformational variation shows large possibility to synthesize biomaterials with physiological activity; moreover, it also provides a promising way to perform complex functions and enable external control *in vivo*. Molecular logic devices attracted significant attention due to their application prospect in the areas of molecular computing and multivariate analysis. Although various molecular logic gates and logic circuits have been reported [79–81], fabricating molecular logic devices in living cells still face challenges. CC peptide and *de novo* designed proteins are excellently suited building blocks for *in vivo* genetically encoded logic circuits, due to their orthogonality, biocompatibility and production in bacteria and mammalian cells. Using CC peptides, a novel fast-responsive split-protease-

cleavable orthogonal-CC-based (SPOC) logic circuits that function in mammalian cells have been designed. In this design, proteolysis-responsive CCs fused with split proteins were used as information-processing modules in combination with orthogonal split proteases, which can be activated by small input molecules. The SPOC logic circuits achieved the full set of Boolean logic gates in mammalian cells and could respond to a signal within 15 min (Fig. 3B) [82]. Protein-level logic circuits were also designed by combination of protease cleavage sites with degrons. Engineered split viral proteases were used as mediators to separate the functional protein domain components from the degron, preventing the degradation of the fluorescent protein output. Such programmable protein circuits can implement various logical functions in mammalian cells [83]. Based on the ‘LOCKR’ technology mentioned above, a tunable protein degradation system was constructed to implement negative and positive feedback controls in mammalian cells [84]. All these logic platforms are driven by protein–protein interactions and can realize multiple functions and signal responses, which offer more options and large freedom in synthetic biology, diagnosis and therapy.

## Drug and small molecular delivery

Highly efficient *in vivo* molecular delivery has always been one of the important goals of biomaterial, biology and medical science [85–88]. Various *de novo* protein cages have been designed as containers for siRNA, small molecular and enzyme delivery, which showed great potential in disease diagnosis, drug delivery and cancer therapy [89–94]. Recently, a positively supercharged protein cage was created by introducing arginine residues into the luminal surface of a nonviral self-assembling protein cage (named O3-33). Thus, the protein cage can encapsulate negatively charged siRNA due to strong electrostatic affinity, therefore delivering and releasing the target gene for oligonucleotide therapeutics in mammalian cells. [95] Similarly, Azuma *et al.* presented an encapsulation strategy to efficiently package target enzymes into negatively charged protein cages through fusing target protein to a positively supercharged fluorescent protein, while the activity of the enzyme was not affected [96]. Morris *et al.* utilized self-assembling peptide cages as modular scaffolds to deliver subunit vaccines, successfully driving antigen-specific immune responses *in vivo* [97].

*De novo* protein cage can also assemble with inorganic or organic materials and molecules to construct complex nanodevice and achieve multiple functions. By either chemical conjugating or genetic insertion



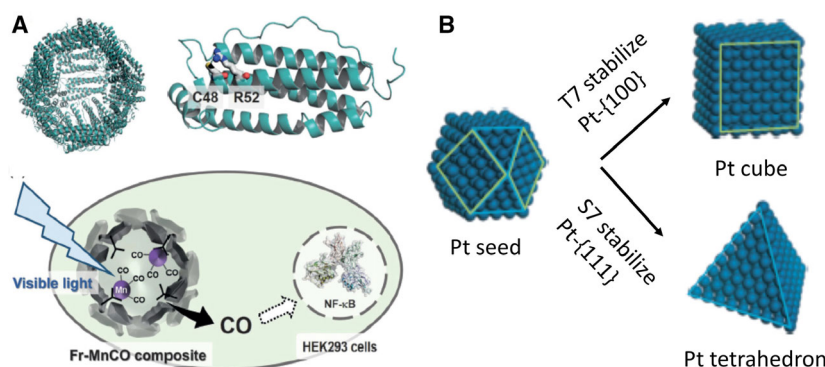
**Fig. 3.** Design of protein switches and logic circuits. (A) The design of dose-response protein switch based on decon LOCKR system. The switch consists of a cage (blue) and latch (purple). Several thermodynamic equilibria exist among the closed state and open state, which can bind to the key (green) and target (yellow). Two orthogonal switches can be expressed constitutively in the same cell and solely triggered by their own key [77]. (B) The logic circuits using CC-fused split proteases as signal transducers to achieve logic gates in mammalian cells [82].

hepatocellular carcinoma cell-binding peptide to the surface of encapsulin protein cage, Kang's group developed a multifunctional platform for drug delivery, cancer diagnostic or bioimaging [98].

Although carbon monoxide (CO) is toxic and insoluble, it is related to a series of defence mechanisms under the physiological and pathological conditions; therefore, *in vivo* delivery of CO is a challenging task [99,100]. The engineered ferritin (Fr) variant, whose Arg52 is replaced with a Cys, has a strong affinity towards the manganese-carbonyl moieties (MnCO), while the wild-type of Fr can barely retain Mn moieties. Triggered by light irradiation, MnCO can release CO molecule, and the release can be adjusted by the length of irradiation. Based on this principle, a Fr variant protein cage was fabricated for cargo and release of CO ligands *in vivo*, which demonstrated a promising approach for target loading and delivering, indicating the prospect of multifunctional protein assemblies and extended applications (Fig. 4A) [101].

## Biosensor and catalyst

Similar to DNA duplex, various controllable programmed CC-based biosensors have been constructed that relay on coiled-coil interactions and displacement for target detection [102–105] and enzyme activity regulation [106,107]. Besides coiled coils, *de novo* designed proteins also can serve as building blocks of biomimetic materials for *in vivo* biosensing, bioimaging and catalysis [108,109]. Apoferritin is a hollow spherical protein, which is stable under the alkaline conditions but can disassemble into 24 subunits under the acidic solution (pH < 2). Utilizing biotin functionalized apoferritin as a signal and amplification unit, a signal-on photoelectrochemical biosensor was constructed for microRNA detection [110]. A highly ordered virus-like protein cage with enzyme cargos has been fabricated by Douglas's group. Through hierarchically assembling cargo-scaffolding protein and coat protein, the virus-like protein cage was built by the bottom-up strategy.



**Fig. 4.** Examples of designed protein carriers and nanocrystals. (A) The crystal structure of ferritin (Fr, PDB: 1DAT) and its modified monomer peptide subunit, as well as engineered-Fr intracellular photoactive CO-releasing system. The release of CO from MnCO coordinated within the Fr cage is triggered by visible light irradiation (456 nm, 15 mWcm<sup>-2</sup>) [101]; (B) platinum nanocubes stabilized by T7 peptide (Thr-Leu-Thr-Thr-Leu-Thr-Asn) and nanotetrahedrons stabilized by S7 peptide (Ser-Ser-Phe-Pro-Gln-Pro-Asn) [118].

The exterior surface and interaction of protein monomers can be modified without varying its overall morphology. Moreover, this protein cage can be used as building blocks to form large size superlattices, due to the interaction between protein surfaces for the high-order self-assembly. During the assembly process, the protein cage remained high catalytic activity of the cargo enzyme in a multistep reaction of isobutyraldehyde conversion [111]. With similar hierarchical bottom-up assembly, a series of programmed protein nanoreactors were also developed for enzyme encapsulation. This kind of protein device can not only deliver and regulate the catalytic activity of the target enzyme *in vivo* but also efficiently maintain the activity of enzymes when they are embedded, immobilized or in harsh environments [112,113]. Taking advantage of *de novo* designed protein cage and sortase-based enzymatic conjugation approach, McConnell *et al.* developed a multienzyme platform to increase enzyme activity and high efficient concerted catalysis. The 24-subunit designed protein cage was used as a scaffold to immobilize cellulase enzymes in a highly ordered spatial arrangement [114]. In these studies, *de novo* designed proteins showed high potential as scaffolds to encapsulate, assemble and integrate small molecule and enzymes for target analysis or catalysis regulation.

## Nanomaterial synthesis

*De novo* designed proteins with different topology and abundant residues are also ideal building blocks to assemble long-range defined and highly ordered nanostructures or can be used as templates for nanomaterial synthesis, which indicates possibilities in bionic and biomaterial areas [115,116]. Recently, a

dimeric 4-helix *de novo* protein was designed as a building block to fabricate various supramolecular nanostructures such as nanopolyhedron and nanochain [117]. Huang's group used facet-selective binding peptides as regulating agents to synthesize the designed shape platinum nanocrystals with selectively exposed crystal surfaces and particular topology. It has been discovered that a short peptide T7 (Thr-Leu-Thr-Thr-Leu-Thr-Asn) could stabilize and selectively bind platinum nanocrystals with (100) crystal facet, while S7 (Ser-Ser-Phe-Pro-Gln-Pro-Asn) can only selectively combine platinum nanocrystals with (111) crystal facet. Therefore, T7 and S7 can particularly direct the growth of platinum nanocubes and nanotetrahedra, respectively (Fig. 4B) [118]. Moreover, the flexible T7 peptide can spontaneously transform into a large-area  $\beta$ -sheet structure at high concentration and thus lead to the self-assembly of cubic Pt nanocrystals into long-range one-dimensional or two-dimensional structure [119]. Tobacco mosaic virus (TMV) coat protein could assemble to a nanotube after its Thr103 was replaced by Cys, owing to disulfide bonds between the modified subunits [120]. TMV protein nanotube can strongly interact with metal ions like Au(III), due to the stable Au-S bond. Taking advantage of this principle, Wang's group synthesized a novel type of protein-templated gold nanoparticles. Depending on reaction conditions, the structure of nanoparticle could be shifted towards the formation of nanochains or nanorods [121]. These strategies indicated that *de novo* proteins could become a bridge between organic biomaterials and inorganic crystals. In this way, the advantages of both materials are combined together to achieve highly ordered structures and multiple physiology functions.



## Conclusions and perspectives

There seem to be less than 2000 natural protein folds, despite the fact that only a small fraction of all polypeptide sequences have their tertiary structure known. Recent advances in protein structure prediction with no related homology extend the number of protein folds and good molecular models with a topology [122]. Therefore, modular protein design provides a strategy that enables the production of *de novo* protein folds based on modules and interactions between those modules in a rational way.

Additionally, *de novo* protein design offers researchers new insights into protein folding and construction mechanisms, as well as the possibility to organize protein structures in a controllable manner and to create new types of biomimetic assemblies. Novel proteins not only possess the advantages of excellent orthogonality with natural cellular components and biocompatibility without interference in the normal cellular processes, but they can also be easily engineered and produced by bacteria or mammalian cells, which make them highly suitable for the construction of *in vivo* logic circuits and biosensors. Diverse multifunctional delivery and catalysis platforms can be fabricated using *de novo* proteins as vehicles and scaffolds. Moreover, *de novo* proteins with precise topology and active domains have been identified as highly useful building blocks for stabilizing biocompatible nanoparticles with various crystal conformation and high biocompatibility. Generally speaking, the efforts of *de novo* protein design have opened a new pathway to study protein structures and functions, and above all offered novel possibilities to construct artificial devices, biomaterials and biomimetics. Development of computational predictions and methods along with the techniques for determination of 3D protein structures, notably cryo-EM, have accelerated *de novo* protein design. *De novo* designed proteins already demonstrated many interesting properties such as extremely high stability, tunable size and shape of cavities. While there are many additional strategies for the construction of new folds, such as the formation of knotted structures, additional exciting challenges lie ahead of the introduction of new protein functions and the design of molecular machines.

## Acknowledgements

This research was supported by grants from the Slovenian Research Agency (P4-0176, J1-9173, J1-1711, N4-0080), ERC grant MaCChines to RJ, Horizon2020 CSA Bioroboost and ERANET project MediSurf. TŠ

is supported by the complementary MSCA scheme from the Ministry of Education, Science and Sports of Slovenia.

## References

- 1 Endy D (2005) Foundations for engineering biology. *Nature* **438**, 449–453.
- 2 Sismour AM and Benner SA (2005) Synthetic biology. *Nat Rev Genet* **6**, 533–543.
- 3 Drubin DA, Way JC and Silver PA (2007) Designing biological systems. *Genes Dev* **21**, 242–254.
- 4 Liao JC, Mi L, Pontrelli S and Luo S (2016) Fuelling the future: microbial engineering for the production of sustainable biofuels. *Nat Rev Microbiol* **14**, 288–304.
- 5 Chang MCY and Keasling JD (2006) Production of isoprenoid pharmaceuticals by engineered microbes. *Nat Chem Biol* **2**, 674–681.
- 6 Breitling R and Takano E (2015) Synthetic biology advances for pharmaceutical production. *Curr Opin Biotechnol* **35**, 46–51.
- 7 Sanchez-Garcia L, Martín L, Mangués R, Ferrer-Miralles N, Vázquez E and Villaverde A (2016) Recombinant pharmaceuticals from microbial cells: a 2015 update. *Microb Cell Fact* **15**, 33.
- 8 Fahnestock SR, Yao Z and Bedzyk LA (2000) Microbial production of spider silk proteins. *Rev Mol Biotechnol* **74**, 105–119.
- 9 Lee JW, Kim HU, Choi S, Yi J and Lee SY (2011) Microbial production of building block chemicals and polymers. *Curr Opin Biotechnol* **22**, 758–767.
- 10 Pohorille A and Deamer D (2002) Artificial cells: prospects for biotechnology. *Trends Biotechnol* **20**, 123–128.
- 11 Lentini R, Yeh Martín N and Mansy SS (2016) Communicating artificial cells. *Curr Opin Chem Biol* **34**, 53–61.
- 12 Xu C, Hu S and Chen X (2016) Artificial cells: from basic science to applications. *Mater Today* **19**, 516–532.
- 13 Buddingh' BC and van Hest JCM (2017) Artificial cells: synthetic compartments with life-like functionality and adaptivity. *Acc Chem Res* **50**, 769–777.
- 14 Zuber G, Dauty E, Nothisen M, Belguise P and Behr JP (2001) Towards synthetic viruses. *Adv Drug Deliv Rev* **52**, 245–253.
- 15 Wimmer E, Mueller S, Tumpey TM and Taubenberger JK (2009) Synthetic viruses: a new opportunity to understand and prevent viral disease. *Nat Biotechnol* **27**, 1163–1172.
- 16 Rosenbaum M, He Z and Angenent LT (2010) Light energy to bioelectricity: photosynthetic microbial fuel cells. *Curr Opin Biotechnol* **21**, 259–264.

- 17 Koga N, Tatsumi-Koga R, Liu G, Xiao R, Acton TB, Montelione GT and Baker D (2012) Principles for designing ideal protein structures. *Nature* **491**, 222–227.
- 18 Molski MA, Goodman JL, Chou F-C, Baker D, Das R and Schepartz A (2013) Remodeling a  $\beta$ -peptide bundle. *Chem Sci* **4**, 319–324.
- 19 Bhardwaj G, Mulligan VK, Bahl CD, Gilmore JM, Harvey PJ, Cheneval O, Buchko GW, Pulavarti SVSRK, Kaas Q, Eletsky A *et al.* (2016) Accurate de novo design of hyperstable constrained peptides. *Nature* **538**, 329–335.
- 20 Geiger-Schuller K, Sforza K, Yuhas M, Parmeggiani F, Baker D and Barrick D (2018) Extreme stability in de novo-designed repeat arrays is determined by unusually stable short-range interactions. *PNAS* **115**, 7539–7544.
- 21 Lin CY and Liu JC (2016) Modular protein domains: an engineering approach toward functional biomaterials. *Curr Opin Biotechnol* **40**, 56–63.
- 22 Parmeggiani F and Huang PS (2017) Designing repeat proteins: a modular approach to protein design. *Curr Opin Struct Biol* **45**, 116–123.
- 23 Crick FHC (1952) Is  $\alpha$ -keratin a coiled coil? *Nature* **170**, 882–883.
- 24 Crick FHC (1953) The Packing of  $\alpha$ -Helices: Simple Coiled-Coils. *Acta Cryst* **6**, 689–698.
- 25 Landschulz WH, Johnson PF and McKnight SL. (1988) The leucine zipper: A hypothetical structure common to a new class of DNA binding proteins. *Science* **240**, 1759–1764.
- 26 Truebestein L and Leonard TA (2016) Coiled-coils: the long and short of it. *BioEssays* **38**, 903–916.
- 27 Hagelueken G, Clarke BR, Huang H, Tuukkanen A, Danciu I, Svergun DI, Hussain R, Liu H, Whitfield C and Naismith JH (2015) A coiled-coil domain acts as a molecular ruler to regulate O-antigen chain length in lipopolysaccharide. *Nat Struct Mol Biol* **22**, 50–56.
- 28 Boyle AL (2018) Applications of de novo designed peptides. In *Peptide Applications in Biomedicine, Biotechnology and Bioengineering* (Koutsopoulos S, ed), pp. 51–86. Elsevier, Duxford.
- 29 Gurnon DG, Whitaker JA and Oakley MG (2003) Design and characterization of a homodimeric antiparallel coiled coil. *J Am Chem Soc* **125**, 7518–7519.
- 30 Gradišar H and Jerala R (2011) De novo design of orthogonal peptide pairs forming parallel coiled-coil heterodimers. *J Pept Sci* **17**, 100–106.
- 31 Chen Z, Boyken SE, Jia M, Busch F, Flores-Solis D, Bick MJ, Lu P, VanAernum ZL, Sahasrabudhe A, Langan RA *et al.* (2019) Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106–111.
- 32 Lombardi A, Bryson JW and DeGrado WF (1996) De novo design of heterotrimeric coiled coils. *Biopolymers* **40**, 495–504.
- 33 Ogihara NL, Weiss MS, Eisenberg D and DeGrado WF (2008) The crystal structure of the designed trimeric coiled coil coil-VaLd: implications for engineering crystals and supramolecular assemblies. *Protein Sci* **6**, 80–88.
- 34 Thomson AR, Wood CW, Burton AJ, Bartlett GJ, Sessions RB and Brady RL & Woolfson DN (2014) Computational design of water-soluble  $\alpha$ -helical barrels. *Science* **346**, 485–488.
- 35 De Crescenzo G, Litowski JR, Hodges RS and O'Connor-McCourt MD (2003) Real-time monitoring of the interactions of two-stranded de novo designed coiled-coils: effect of chain length on the kinetic and thermodynamic constants of binding. *Biochemistry* **42**, 1754–1763.
- 36 Thomas F, Boyle AL, Burton AJ and Woolfson DN (2013) A set of de novo designed parallel heterodimeric coiled coils with quantified dissociation constants in the micromolar to sub-nanomolar regime. *J Am Chem Soc* **135**, 5161–5166.
- 37 Lizatović R, Aurelius O, Stenström O, Drakenberg T, Akke M, Logan DT and André I (2016) A de novo designed coiled-coil peptide with a reversible pH-induced oligomerization switch. *Structure* **24**, 946–955.
- 38 Burton AJ, Thomas F, Agnew C, Hudson KL, Halford SE, Brady RL and Woolfson DN (2013) Accessibility, reactivity, and selectivity of side chains within a channel of de novo peptide assembly. *J Am Chem Soc* **135**, 12524–12527.
- 39 Tolbert AE, Ervin CS, Ruckthong L, Paul TJ, Jayasinghe-Arachchige VM, Neupane KP, Stuckey JA, Prabhakar R and Pecoraro VL (2020) Heteromeric three-stranded coiled coils designed using a Pb(ii)(Cys) 3 template mediated strategy. *Nat Chem* **12**, 405–411.
- 40 Anuar K, Khairil Anuar INA, Banerjee A, Keeble AH, Carella A, Nikov GI and Howarth M (2019) Spy&Go purification of SpyTag-proteins using pseudo-SpyCatcher to access an oligomerization toolbox. *Nat Commun* **10**, 1–13.
- 41 Thomas F, Dawson WM, Lang EJM, Burton AJ, Bartlett GJ, Rhys GG, Mulholland AJ and Woolfson DN (2018) De novo-designed  $\alpha$ -helical barrels as receptors for small molecules. *ASC Synth Biol* **7**, 1808–1816.
- 42 Eriksson M, Hassan S, Larsson R, Linder S, Ramqvist T, Lövborg H, Vikinge T, Figgemeier E, Müller J, Stetefeld J *et al.* (2009) Utilization of a right-handed coiled-coil protein from archaeobacterium *Staphylothermus marinus* as a carrier for cisplatin. *Anticancer Res* **29**, 11–18.
- 43 Mann FA, Horlebein J, Meyer NF, Meyer D, Thomas F and Kruss S (2018) Carbon nanotubes encapsulated in coiled-coil peptide barrels. *Chem – A Eur J* **24**, 12241–12245.

- 44 Woolfson DN (2017) Coiled-coil design: updated and upgraded. In *Fibrous Proteins: Structures and Mechanisms*. Subcellular Biochemistry (DP & Squire J, eds), pp. 35–61. Springer, Cham.
- 45 Apostolovic B, Danial M and Klok H-A (2010) Coiled coils: attractive protein folding motifs for the fabrication of self-assembled, responsive and bioactive materials. *Chem Soc Rev* **39**, 3541–3575.
- 46 Wu Y and Collier JH (2017)  $\alpha$ -Helical coiled-coil peptide materials for biomedical applications. *Wiley Interdiscip Rev Nanomed Nanobiotechnol* **9**, e1424.
- 47 Reeb J and Rost B (2019) Secondary Structure Prediction. In *Encyclopedia of Bioinformatics and Computational Biology* (Ranganathan S, Gribskov M, Nakai K, & Schönbach C, eds), pp. 488–496. Elsevier, Amsterdam.
- 48 Colley CS, Griffiths-Jones SR, George MW and Searle MS (2000) Do interstrand hydrogen bonds contribute to  $\beta$ -hairpin peptide stability in solution? IR analysis of peptide folding in water. *Chem Commun*, 593–594.
- 49 Yennamalli RM (2019) Protein design. In *Encyclopedia of Bioinformatics and Computational Biology* (Ranganathan S, Gribskov M, Nakai K, & Schönbach C, eds), pp. 644–651. Elsevier, Amsterdam.
- 50 Oi C, Treado JD, Levine ZA, Lim CS, Knecht KM, Xiong Y, O'Hern CS and Regan L (2018) A threonine zipper that mediates protein–protein interactions: Structure and prediction. *Protein Sci* **27**, 1969–1977.
- 51 Longo LM and Blaber M (2014) Symmetric protein architecture in protein design: Top-down symmetric deconstruction. In *Protein Design. Methods in Molecular Biology (Methods and Protocols)* (Köhle V, ed), pp. 161–182. Humana Press Inc, New York.
- 52 Lee J and Blaber M (2010) Experimental support for the evolution of symmetric protein architecture from a simple peptide motif. *PNAS* **108**, 126–130.
- 53 Broom A, Doxey AC, Lobsanov YD, Berthin LG, Rose DR, Howell PL, McConkey BJ and Meiering EM (2012) Modular evolution and the origins of symmetry: reconstruction of a three-fold symmetric globular protein. *Cell* **20**, 161–171.
- 54 Marcos E, Chidyausiku TM, McShan AC, Evangelidis T, Nerli S, Carter L, Nivón LG, Davis A, Oberdorfer G, Tripsianes K *et al.* (2018) De novo design of a non-local  $\beta$ -sheet protein with high stability and accuracy. *Nat Struct Mol Biol* **25**, 1028–1034.
- 55 Dou J, Vorobieva AA, Sheffler W, Doyle LA, Park H, Bick MJ, Mao B, Foight GW, Lee MY, Gagnon LA *et al.* (2018) De novo design of a fluorescence-activating  $\beta$ -barrel. *Nature* **561**, 485–491.
- 56 Marcos E, Basanta B, Chidyausiku TM, Tang Y, Oberdorfer G, Liu G, Swapna GVT, Guan R, Silva D-A, Dou J *et al.* (2017) Principles for designing proteins with cavities formed by curved  $\beta$  sheets. *Science* **355**, 201–206.
- 57 King IC, Gleixner J, Doyle L, Kuzin A, Hunt JF, Xiao R, Montelione GT, Stoddard BL, DiMaio F and Baker D (2015) Precise assembly of complex  $\beta$  sheet topologies from de novo designed building blocks. *Elife* **4**, e11012.
- 58 Park M, Bedewy M, Berggren KK and Keating AE (2017) Modular assembly of a protein nanotriangle using orthogonally interacting coiled coils. *Sci Rep* **7**, 10577.
- 59 Bai W, Sargent CJ, Choi J-M, Pappu RV and Zhang F (2019) Covalently-assembled single-chain protein nanostructures with ultra-high stability. *Nat Commun* **10**, 3317.
- 60 Gradišar H, Božič S, Doles T, Vengust D, Hafner-Bratkovič I, Mertelj A, Webb B, Šali A, Klavžar S and Jerala R (2013) Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat Chem Biol* **9**, 362–366.
- 61 Ljubetič A, Lapenta F, Gradišar H, Drobnak I, Aupič J, Strmšek Ž, Lainšček D, Hafner-Bratkovič I, Majerle A, Krivec N *et al.* (2017) Design of coiled-coil protein-origami cages that self-assemble in vitro and in vivo. *Nat Biotechnol* **35**, 1094–1101.
- 62 Badieyan S, Sciore A, Eschweiler JD, Koldewey P, Cristie-David AS, Ruotolo BT, Bardwell JCA, Su M and Marsh ENG (2017) Symmetry-directed self-assembly of a tetrahedral protein cage mediated by de novo-designed coiled coils. *ChemBioChem* **18**, 1888–1892.
- 63 Sciore A, Su M, Koldewey P, Eschweiler JD, Diffley KA, Linhares BM, Ruotolo BT, Bardwell JCA, Skinotis G, Neil E *et al.* (2016) Flexible, symmetry-directed approach to assembling protein cages. *Proc Natl Acad Sci* **113**, 8681–8686.
- 64 Cristie-David AS, Koldewey P, Meinen BA, Bardwell JCA and Marsh ENG (2018) Elaborating a coiled-coil-assembled octahedral protein cage with additional protein domains. *Protein Sci* **27**, 1893–1900.
- 65 Fletcher JM, Harniman RL, Barnes FRH, Boyle AL, Collins A, Mantell J, Sharp TH, Antognozzi M, Booth PJ, Linden N *et al.* (2013) Self-assembling cages from coiled-coil peptide modules. *Science* **340**, 595–599.
- 66 Heddle JG, Chakraborti S and Iwasaki K (2017) Natural and artificial protein cages: design, structure and therapeutic applications. *Curr Opin Struct Biol* **43**, 148–155.
- 67 Aumiller WM, Uchida M and Douglas T (2018) Protein cage assembly across multiple length scales. *Chem Soc Rev* **47**, 3433–3469.
- 68 Jorda J, Leibly DJ, Thompson MC and Yeates TO (2016) Structure of a novel 13 nm dodecahedral nanocage assembled from a redesigned bacterial microcompartment shell protein. *Chem Commun* **52**, 5041–5044.
- 69 Ladenstein R, Fischer M and Bacher A (2013) The lumazine synthase/riboflavin synthase complex: shapes

- and functions of a highly variable enzyme system. *FEBS J* **280**, 2537–2563.
- 70 Sasaki E, Böhringer D, Van De Waterbeemd M, Leibundgut M, Zschoche R, Heck AJR, Ban N and Hilvert D (2017) Structure and assembly of scalable porous protein cages. *Nat Commun* **8**, 1–10.
  - 71 Cristie-David AS, Chen J, Nowak DB and Bondy AL (2019) Coiled-coil-mediated assembly of an icosahedral protein cage with extremely high thermal and chemical stability. *J Am Chem Soc* **141**, 9207–9216.
  - 72 Malay AD, Miyazaki N, Biela A, Chakraborti S, Majsterkiewicz K, Stupka I, Kaplan CS, Kowalczyk A, Piette BMAG, Hochberg GKA *et al.* (2019) An ultra-stable gold-coordinated protein cage displaying reversible assembly. *Nature* **569**, 438–442.
  - 73 Golub E, Subramanian RH, Esselborn J, Alberstein RG, Bailey JB, Chiong JA, Yan X, Booth T, Baker TS and Tezcan FA (2020) Constructing protein polyhedra via orthogonal chemical interactions. *Nature* **578**, 172–176.
  - 74 Brunette TJ, Parmeggiani F, Huang PS, Bhabha G, Ekiert DC, Tsutakawa SE, Hura GL, Tainer JA and Baker D (2015) Exploring the repeat protein universe through computational protein design. *Nature* **528**, 580–584.
  - 75 Rocklin GJ, Chidyausiku TM, Goreshnik I, Ford A, Houliston S, Lemak A, Carter L, Ravichandran R, Mulligan VK, Chevalier A, *et al.* (2017) Global analysis of protein folding using massively parallel design, synthesis, and testing. *Science* **357**, 168–175.
  - 76 Doyle L, Hallinan J, Bolduc J, Parmeggiani F, Baker D, Stoddard BL and Bradley P (2015) Rational design of  $\alpha$ -helical tandem repeat proteins with closed architectures. *Nature* **528**, 585–588.
  - 77 Langan RA, Boyken SE, Ng AH, Samson JA, Dods G, Westbrook AM, Nguyen TH, Lajoie MJ, Chen Z, Berger S *et al.* (2019) De novo design of bioactive protein switches. *Nature* **572**, 205–210.
  - 78 Aupic J, Lapenta F and Jerala R (2018) SwitCCCh: Metal-site design for controlling the assembly of a coiled-coil homodimer. *ChemBioChem* **19**, 2453–2457.
  - 79 Ling J, Daly B, Silversen VAD and De Silva AP (2015) Taking baby steps in molecular logic-based computation. *Chem Commun* **51**, 8403–8409.
  - 80 Erbas-Cakmak S, Kolen S, Sedgwick AC, Gunnlaugsson T, James TD, Yoon J and Akkaya EU (2018) Molecular logic gates: The past, present and future. *Chem Soc Rev* **47**, 2228–2248.
  - 81 Yang J, Song Z, Liu S, Zhang Q and Zhang C (2016) Dynamically arranging gold nanoparticles on DNA origami for molecular logic gates. *ACS Appl Mater Interfaces* **8**, 32.
  - 82 Fink T, Lončarić J, Praznik A, Plaper T, Merljak E, Leben K, Jerala N, Lebar T, Strmšek Ž, Lapenta F *et al.* (2019) Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat Chem Biol* **15**, 115–122.
  - 83 Gao XJ, Chong LS, Kim MS and Elowitz MB (2018) Programmable protein circuits in living cells. *Science* **361**, 1252–1258.
  - 84 Ng AH, Nguyen TH, Gómez-Schiavon M, Dods G, Langan RA, Boyken SE, Samson JA, Waldburger LM, Dueber JE, Baker D *et al.* (2019) Modular and tunable biological feedback control using a de novo protein switch. *Nature* **572**, 265–269.
  - 85 Patel A, Patel M, Yang X and Mitra AK (2014) Recent advances in protein and peptide drug delivery: a special emphasis on polymeric nanoparticles. **19**, 1102–1120.
  - 86 Wiraja C, Zhu Y, Chin D, Lio S, Yeo DC, Xie M, Fang W, Li Q, Zheng M, Van SM *et al.* (2019) Framework nucleic acids as programmable carrier for transdermal drug delivery. *Nat Commun* **10**, 1147.
  - 87 Ouyang C, Zhang S, Xue C, Yu X, Xu H, Wang Z, Lu Y and Wu Z-S (2020) Precision-guided missile-like DNA nanostructure containing warhead and guidance control for aptamer-based targeted drug delivery into cancer cells in vitro and in vivo. *J Am Chem Soc* **142**, 1265–1277.
  - 88 Woolfson DN (2018) Modifying self-assembled peptide cages to control internalization into mammalian cells. *Nano Lett* **18**, 5933–5937.
  - 89 Murata M, Narahara S, Kawano T, Hamano N, Piao JS, Kang JH, Ohuchida K, Murakami T and Hashizume M (2015) Design and function of engineered protein nanocages as a drug delivery system for targeting pancreatic cancer cells via neuropilin-1. *Mol Pharm* **12**, 1422–1430.
  - 90 Lee EJ, Lee NK and Kim IS (2016) Bioengineered protein-based nanocage for drug delivery. *Adv Drug Deliv Rev* **106**, 157–171.
  - 91 Maccone A, Masciarelli S, Palombarini F, Quaglio D, Boffi A, Cardoso Trabuco M, Baiocco P, Fazi F and Bonamore A (2019) Ferritin nanovehicle for targeted delivery of cytochrome c to cancer cells. *Sci Rep* **9**, 11749.
  - 92 Giordano-Attianese G, Gainza P, Gray-Gaillard E, Cribioli E, Shui S, Kim S, Kwak M-J, Vollers S, Corria Osorio ADJ, Reichenbach P *et al.* (2020) A computationally designed chimeric antigen receptor provides a small-molecule safety switch for T-cell therapy. *Nature Biotechnol* **38**, 426–432.
  - 93 Fabian Sesterhenn A, Yang C, Bonet J, Cramer JT, Wen X, Wang Y, Chiang C-I, Abriata LA, Kucharska I, Castoro G *et al.* (2020) De novo protein design enables precise induction of functional antibodies in vivo. *bioRxiv*. <https://doi.org/10.1101/685867>
  - 94 Beesley JL, Baum HE, Hodgson LR, Verkade P, Banting GS and Woolfson DN (2018) Modifying self-assembled peptide cages to control internalization into mammalian cells. *Nano Lett* **18**, 5933–5937.



- 95 Edwardson TGW, Mori T and Hilvert D (2018) Rational engineering of a designed protein cage for siRNA delivery. *J Am Chem Soc* **140**, 10439–10442.
- 96 Azuma Y, Zschoche R, Tinzl M and Hilvert D (2016) Quantitative packaging of active enzymes into a protein CAGE. *Angew Chemie Int Ed* **55**, 1531–1534.
- 97 Morris C, Glennie SJ, Lam HS, Baum HE, Kandage D, Williams NA, Morgan DJ, Woolfson DN and Davidson AD (2019) A modular vaccine platform combining self-assembled peptide cages and immunogenic peptides. *Adv Funct Mat* **29**, 1–12.
- 98 Moon H, Lee J, Min J and Kang S (2014) Developing genetically engineered encapsulin protein cage nanoparticles as a targeted delivery nanoplatform. *Biomacromol* **15**, 3794–3801.
- 99 Johnson TR, Mann BE, Clark JE, Foresti R, Green CJ and Motterlini R (2003) Metal carbonyls: a new class of pharmaceuticals? *Angew Chemie – Int Ed* **42**, 3722–3729.
- 100 García-Gallego S and Bernardes GJL (2014) Carbon-monoxide-releasing molecules for the delivery of therapeutic co in vivo. *Angew Chemie – Int Ed* **53**, 9712–9721.
- 101 Fujita K, Tanaka Y, Abe S and Ueno T (2016) A photoactive carbon-monoxide-releasing protein cage for dose-regulated delivery in living cells. *Angew Chemie – Int Ed* **55**, 1056–1060.
- 102 Dai Y, Abbasi K, Bandyopadhyay S and Liu CC (2019) Dynamic control of peptide strand displacement reaction using functional biomolecular domain for biosensing. *ACS Sensors* **4**, 1980–1985.
- 103 Jeong W, Choi S-H, Lee H and Lim Y (2019) A fluorescent supramolecular biosensor for bacterial detection via binding-induced changes in coiled-coil molecular assembly. *Sensors Actuators B Chem* **290**, 93–99.
- 104 Cambay F, Henry O, Durocher Y and De Crescenzo G (2019) Impact of N-glycosylation on Fcγ receptor/IgG interactions: unravelling differences with an enhanced surface plasmon resonance biosensor assay based on coiled-coil interactions. *MAbs* **11**, 435–452.
- 105 De Franco S, Vandenamee J, Brans A, Verlaine O, Bendak K, Damblon C, Matagne A, Segal DJ, Galleni M, Mackay JP *et al.* (2019) Exploring the suitability of RanBP2-type Zinc Fingers for RNA-binding protein design. *Sci Rep* **9**, 1–13.
- 106 Inobe T and Genmei R (2015) Inhibition of the 26S proteasome by peptide mimics of the coiled-coil region of its ATPase subunits. *Biochem Biophys Res Commun* **468**, 143–150.
- 107 Lechner H, Ferruz N and Höcker B (2018) Strategies for designing non-natural enzymes and binders. *Curr Opin Chem Biol* **47**, 67–76.
- 108 Zhang Y, Ardejani MS and Orner BP (2016) Design and applications of protein-cage-based nanomaterials. *Chem – An Asian J* **11**, 2814–2828.
- 109 Woalder (2017) Protein nanoparticles as multifunctional biocatalysts and health assessment sensors. *Physiol Behav* **176**, 139–148.
- 110 Yin H, Wang M, Zhou Y, Zhang X, Sun B, Wang G and Ai S (2014) Photoelectrochemical biosensing platform for microRNA detection based on in situ producing electron donor from apoferritin-encapsulated ascorbic acid. *Biosens Bioelectron* **53**, 175–181.
- 111 Uchida M, McCoy K, Fukuto M, Yang L, Yoshimura H, Miettinen HM, LaFrance B, Patterson DP, Schwarz B, Karty JA *et al.* (2018) Modular self-assembly of protein cage lattices for multistep catalysis. *ACS Nano* **12**, 942–953.
- 112 Patterson DP, Prevelige PE and Douglas T (2012) Nanoreactors by programmed enzyme encapsulation inside the capsid of the bacteriophage P22. *ACS Nano* **6**, 5000–5009.
- 113 Patterson DP, Schwarz B, El-Boubbou K, Van Der Oost J, Prevelige PE and Douglas T (2012) Virus-like particle nanoreactors: Programmed encapsulation of the thermostable CelB glycosidase inside the P22 capsid. *Soft Matter* **8**, 10158–10166.
- 114 McConnell SA, Cannon KA, Morgan C, McAllister R, Amer BR, Clubb RT and Yeates TO (2020) Designed protein cages as scaffolds for building multienzyme materials. *Synth Biol* **9**, 381–391.
- 115 Chotera A, Sadihov H, Cohen-Luria R, Monnard PA and Ashkenasy G (2018) Functional assemblies emerging in complex mixtures of peptides and nucleic acid-peptide chimeras. *Chem – A Eur J* **24**, 10128–10135.
- 116 Sim S, Miyajima D, Niwa T, Taguchi H and Aida T (2015) Tailoring micrometer-long high-integrity 1D array of superparamagnetic nanoparticles in a nanotubular protein jacket and its lateral magnetic assembling behavior. *J Am Chem Soc* **137**, 4658–4661.
- 117 Kimura N, Mochizuki K, Umezawa K, Hecht MH, Arai R, Kimura N, Mochizuki K, Umezawa K, Hecht MH and Arai R (2020) Hyperstable de novo protein with a dimeric bisecting topology hyperstable de novo protein with a dimeric bisecting. *ACS Synth Biol* **9**, 254–259.
- 118 Chiu CY, Li Y, Ruan L, Ye X, Murray CB and Huang Y (2011) Platinum nanocrystals selectively shaped using facet-specific peptide sequences. *Nat Chem* **3**, 393–399.
- 119 Zhu E, Wang S, Yan X, Sobani M, Ruan L, Wang C, Liu Y, Duan X, Heinz H and Huang Y (2019) Long-range hierarchical nanocrystal assembly driven by molecular structural transformation. *J Am Chem Soc* **141**, 1498–1505.

- 120 Zhou K, Li F, Dai G, Meng C and Wang Q (2013) Disulfide bond: dramatically enhanced assembly capability and structural stability of tobacco mosaic virus nanorods. *Biomacromolecules* **14**, 2593–2600.
- 121 Zhou K, Zhang J and Wang Q (2015) Site-selective nucleation and controlled growth of gold nanostructures in tobacco mosaic virus nanotubulars. *Small* **11**, 2505–2509.
- 122 Marcos E and Silva DA (2018) Essentials of de novo protein design: Methods and applications. *Wiley Interdiscip Rev Comput Mol Sci* **8**, 1–19.