Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Biotechnology

Review

Advances in designed bionanomolecular assemblies for biotechnological and biomedical applications

Jaka Snoj[1,*], Weijun Zhou[1,*], Ajasja Ljubetič[1,2] and Roman Jerala[1,2]

Recent advances in protein engineering have revolutionized the design of bionanomolecular assemblies for functional therapeutic and biotechnological applications. This review highlights the progress in creating complex protein architectures, encompassing both finite and extended assemblies. AI tools, including AlphaFold, RFDiffusion, and ProteinMPNN, have significantly enhanced the scalability and success of *de novo* designs. Finite assemblies, like nanocages and coiled-coil-based structures, enable precise molecular encapsulation or functional protein domain presentation. Extended assemblies, including filaments and 2D/3D lattices, offer unparalleled structural versatility for applications such as vaccine development, responsive biomaterials, and engineered cellular scaffolds. The convergence of artificial intelligence–driven design and experimental validation promises strong acceleration of the development of tailored protein assemblies, offering new opportunities in synthetic biology, materials science, biotechnology, and biomedicine.

Addresses
[1] Department of Synthetic Biology and Immunology, National Institute of Chemistry, Ljubljana, Slovenia
[2] EN-FIST Centre of Excellence, Ljubljana, Slovenia

Corresponding authors: Ljubetič, Ajasja (ajasja.ljubetic@ki.si), Jerala, Roman (roman.jerala@ki.si)
[*] These authors contributed equally to this work.

Current Opinion in Biotechnology 2025, **92**:103256

This review comes from a themed issue on **Tissue, Cell and Pathway Engineering**

Edited by **Wilfried Weber** and **Fussenegger Martin**

For complete overview of the section, please refer to the article collection, "Tissue, Cell and Pathway Engineering (2025)"

Available online 18 January 2025

https://doi.org/10.1016/j.copbio.2024.103256

Introduction

Protein engineering has been established as a crucial field in biotechnology and synthetic biology, aiming to create novel proteins with desired structures and functions. Protein design and assembly into complex structures have been inspired by the intricate molecular machines found in nature, which perform a wide range of biological functions defined at the nanoscale. Most protein-based molecular machines are assembled from several components, which facilitates their versatility through combinations and evolution of modules. Through recent advances in computational design, experimental techniques, and a deeper understanding of protein interactions, scientists can tailor proteins for specific tasks, such as drug development, materials science, and biomedical and bioengineering applications [1,2].

Designed protein assembly into functional molecular scaffolds and machines can be achieved through various strategies, including computational modeling of new interfaces and fusion of existing oligomers [3]. Recent advances in protein design have focused on creating both finite and extended assemblies, such as protein cages or two-dimensional arrays that have applications in drug delivery and as structural scaffolds in synthetic biology [4]. Computational tools have been integral in designing protein interfaces and predicting how proteins can self-assemble. These tools have been particularly effective in designing protein complexes with precise geometries and binding properties using principles derived from the structural data of naturally occurring proteins. The combination of computational approaches and experimental validation is valuable for creating new materials with tailored properties.

Scope of the review

In this review, we focus on designed protein assemblies that have the potential for biomedical applications. We will discuss the role of computational design, modular protein design, and the integration of these techniques to create robust protein assemblies with potential for biotechnological innovation. For a general introduction to protein design, see Refs. [1,5,6], and for recent opinions on the benefits of computational protein design, see Woolfson et al. [7].

Recent advancement in single-chain design

Since 2021, deep neural network–based computational methodologies, such as AlphaFold2 [8,9], AlphaFold 3 [10], RosettaFold2 [11], and large language models like ESM2 [12], significantly advanced protein design by

enabling accurate protein structure prediction from amino acid sequences. The first artificial intelligence (AI) *de novo* single-chain proteins were less than 150 amino acid residues [13], and the size limit was increased to 300 residues by using RFDiffusion [14]. Recently, the limit of single-chain design was further raised by backpropagation through AF2 with relaxed sequence optimization enabling to reach impressive 1000 residues of length that have been also experimentally validated [15].

Protein assemblies

Creating larger structures typically involves assembling multiple components while ensuring specific interactions often combined with symmetry to align the building blocks as designed. The connections can be made through rigid or flexible genetic fusion, novel protein–protein interactions (PPIs), or metal/small molecule–mediated interfaces [16,17].

Protein assemblies are generally classified as finite or extended [2]. Finite assemblies are oligomeric structures of a defined stoichiometry and size that perform diverse biochemical functions, such as scaffolding, encapsulation, signaling, and catalysis, often relying on their assembly for stability and activity. In contrast, extended assemblies are open, polymeric structures with less defined oligomerization state, often with crystalline-like order, serving as cellular scaffolding, transport, and are characterized by their dimensionality (1D, 2D, or 3D).

Finite assemblies

The design of finite protein assemblies enables the creation of structures with applications, such as molecular encapsulation or spatially controlled antigen presentation. The design of such structures is most efficiently done computationally using strategies such as genetic fusion, interface design, or modular (coiled-coil) assembly and using more recent *de novo* design methods based on AI.

An approach of computational design of genetic fusion of self-oligomerizing domains into higher order assemblies uses a rapid and systematic check if overlaps of building blocks have low root mean square deviation (RMSD) and if they fulfill the required geometric criteria. The method (named WORMS) avoids the design of new PPI and enabled the development of novel architectures with helical bundle-repeat protein fusions [18]. Rigid fusion to create bivalent connectors was explored also to design complex asymmetric assemblies [19].

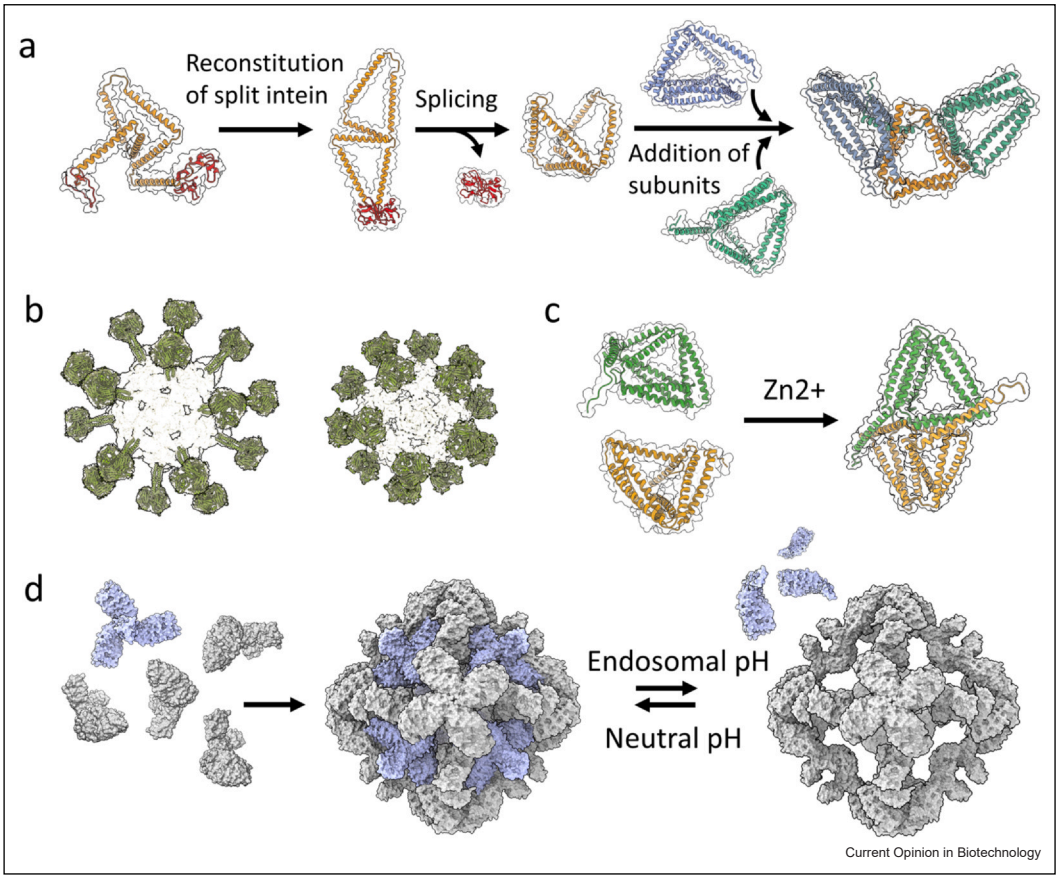
De novo–designed assemblies offer tunability of several parameters, such as stability, symmetry, and responsiveness, enabling the construction of diverse geometries and scales. Huddy et al. introduced a toolkit of twistless helix repeat (THR) blocks that facilitates modular protein assemblies with linear, curved, or angled configurations [20]. The assembly size is controlled by adjusting the number of THR repeats without altering interblock interfaces. This approach allows for predictable, scalable designs spanning finite nanocages to 1D, 2D, and 3D expandable nanostructures.

An alternative, a modular approach that uses orthogonal CC pairs to assemble chains forming coiled-coil protein origami (CCPO) structures has advanced since the first self-assembling tetrahedral protein cage [21]. Subsequent studies validated CCPO structures using SAXS [22–24] and high-resolution X-ray diffraction [25] and explored designed folding pathways [26], enabling the design of more complex multichain assemblies. Advancing modular designs from single-chain to larger CCPO assemblies required refining design principles, particularly topology and terminal CC segment positioning, to prevent fraying by less stable peptide pairs [24]. The precise building block arrangements are crucial for oligomeric CC-based nanostructures to achieve desired conformations, such as, for example, bipyramids formed by a pair of pseudo-symmetric chains [23]. Efforts to extend beyond two-chain assemblies to three-chain structures have highlighted the critical importance of structural constraints for regulating linker length and flexibility. Without these constraints, heterogeneous tertiary or quaternary structures may arise [27]. The modular designed three-chain CC-based assembly, an irregular concave octahedron, utilized a split intein for chain cyclization via spontaneous self-splicing, providing preorganization of building blocks as structural constraints to facilitate the correct assembly (Figure 1a) [28]. This strategy enables the construction of well-defined nanostructures with individually addressable vertices or edges, facilitating applications such as antigen presentation, receptor activation, catalytic center positioning, or binding site arrangement.

Precise control over geometry of *de novo* nanocages can be used to prepare structure-based vaccines, where the periodic presentation of antigens increases the immune response. Protein cages that present antigen Receptor-Binding Domain (RBD) in controllable configurations demonstrated that appropriate spacing of antigens improves the immune response and neutralization titers (Figure 1b) [29]. Additionally, designed cages can also be functionalized to house chlorophyll dimers in a precise geometry that enable reproduction of reaction centers of photoactive organisms [30]. Such assembled nanocages represent the first steps toward *de novo* design of photosynthetic compartments analogous to chromatophores.

Dawson et al. have shown how a library of designed self-assembling α-helical barrels (αHBs) can be used for sensing small molecules. Different members of αHBs libraries have lumens with diverse properties. When

Figure 1



Functional finite assemblies that are responsive to pH or Zn(II), can have tunable spacing, or can be assembled from many preorganized parts. **(a)** Heterotrimeric octahedral structure where the necessary structural constraints for the desired geometry were achieved by cyclization of the connecting (orange) subunit via spontaneous intein-mediated self-splicing. A rough structure was determined by SAXS (SASBDB ID: SASDUS2) [28]. **(b)** *De novo* cages for presenting antigen RBD at a tunable distance. Precise control over the spacing of the antigens was achieved by varying the number of heptades of the tethering helical bundles to maximize the immunity against influenza: 6 heptades (left, PDB ID: 8UR7) or 1 heptade (right, PDB ID: 8UR5) [29]. **(c)** Designed heterodimeric bipyramidal cage with zinc-dependent self-assembly. The rough structure of the dimeric assembly was determined by SAXS [22]. **(d)** Trimeric plugs (blue) designed to initiate cargo release from an octahedral cage (gray) in environments with low pH. A schematic representation is based on a cryo-EM density map (EMD-29602) and refined models [37].

loaded with a fluorescent hydrophobic dye, the analytes caused differential fluorophore displacement. The observed patterns of the response to different oligomeric peptide barrels can be connected to specific small molecules using machine learning, enabling the analysis of complex mixtures [31].

The interface design can be accomplished by modeling of elements crucial for the assembly, which can implement the desired geometry. The first step involves protein–protein docking using computational tools, such as nanohedra, which integrates symmetry-based geometric constraints with a fragment-based approach that leverages empirical patterns from known protein–protein interfaces to generate native-like docking poses [32]. The latest physics-based method for docking is RPXDock [33], which uses a hierarchical search algorithm and a sequence-independent scoring function to efficiently search through large conformational spaces. In the second interface design stage, amino acid sequences are engineered to stabilize the specified protein–protein interface. The deep learning–based tool ProteinMPNN [34] is now integral to the final stages of many protein assembly design workflows. Its utility has been demonstrated in recent studies, including the successful assembly of two-component cages in 13 of 27 tested configurations [35] and the design of new tetrahedral cages using a fragment-based approach combined with ProteinMPNN to enhance the experimental success [36].

Besides being able to accurately design the geometry of finite assemblies, to control the assembly with different chemical stimuli is an important step toward

functionalization of such structures. pH-dependent CC pairs allowed a controlled assembly of pseudo-symmetric chains into a bipyramid (Figure 1c) [22]. Reliable, reversible control of CC-based cage structures has potential for applications, such as cargo delivery and release in the cellular environments with low Zn(II) concentrations or low pH, for example, in the nucleus or endosomes. A recent example of a controllable disassembly are trimeric plugs that allow opening or closing of porous protein cages in different environments (Figure 1d). The pH of disassembly can be tuned from 5.1 to 6.7, making the cages useful for the endosomal delivery [37].

AI methods have been used to design higher order assemblies. Watson et al. demonstrated that cyclic oligomers and icosahedral cages can be designed using RFDiffusion [14]. Wicky et al. have shown that AF2 hallucination can be used to make large cyclic homo-oligomers [38]. The methods are currently limited by the large graphics processing unit (GPU) memory needed for the design since all units are explicitly modeled.

Larger assemblies can be created using pseudosymmetry (same backbone different sequence). An approach to protein assembly utilizing hetero-oligomers [39,40] has enabled the construction of T=4 nanocages with four components and dimensions up to 75 nm. Even larger cages can be created by combining pseudo- and quasi-symmetry (same sequence in different symmetry roles), where cages with sizes up to 96 nm have been designed [41]. A downside of this approach is that a distribution of cage sizes is obtained. The distribution can be partially controlled by the stoichiometric ratio of monomers.

With higher protein design success rates, oligomeric assemblies can be used as the components of complex protein machines. For example, oligomers have been designed to function as axles and rotors [42] and cryo-EM structures of various rotation states have been confirmed.

Extended assemblies

While extended assemblies enable larger and more versatile architectures with significant functional potential, they also pose challenges in stability, design complexity, and computational demands. The algorithms must handle increased degrees of freedom while maintaining structural fidelity across extensive assemblies. One solution is RPXDock, a sequence-independent docking method that aligns protein subunits into symmetric architectures using a residue-pair transform scoring system. This system streamlines rigid-body docking of cyclic oligomers, enabling the efficient exploration of configurations without prior sequence design [33].

One-dimensional assemblies

In 1D structures such as filaments or nanorods, controlling subunit flexibility is crucial for maintaining integrity. Different CC-mediated fibrils have been designed [43,44]. Mezgec et al. utilized engineered natural spectrin repeats fused to CC dimers to create rigid nanorods extending to several micrometers long (Figure 2a). These structures retained α-helix continuity and were validated using atomic force microscopy and electron microscopy [45]. Decorated with functional domains, the rods demonstrated precise and tunable protein spatial presentation along filament chains, enabling nanometer-scale spacing ideal for applications such as B-cell receptor stimulation in vaccine design.

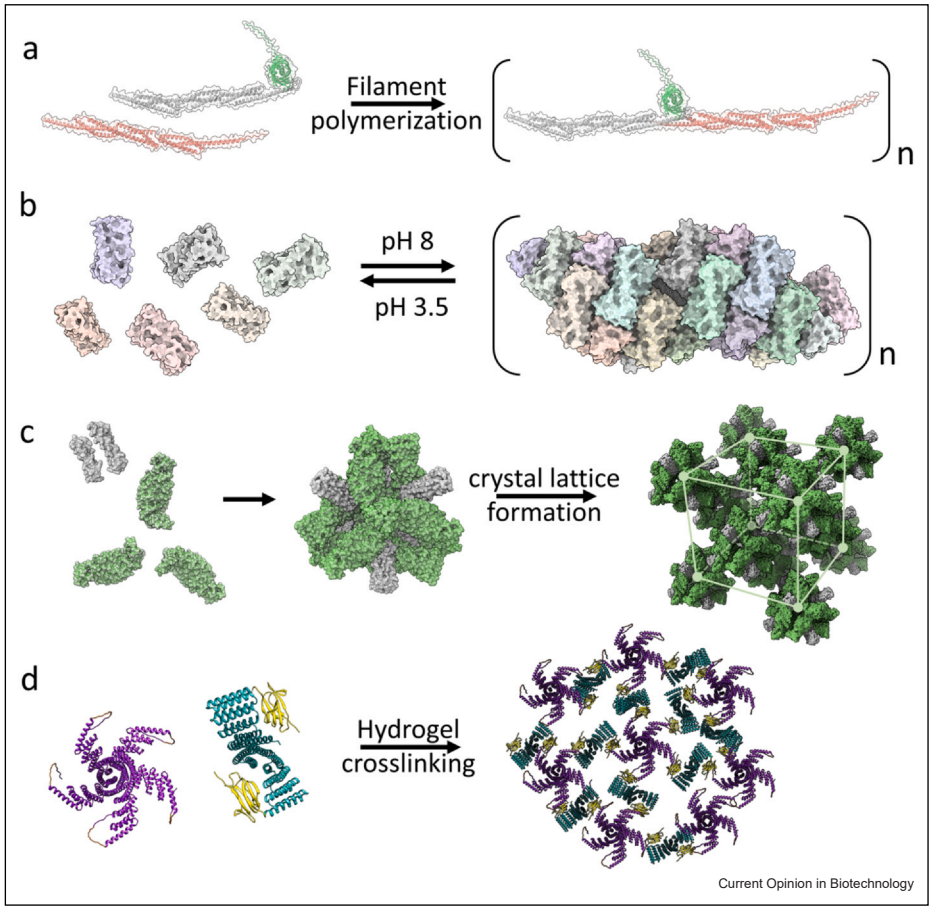
Biological systems often leverage environmental cues for the dynamic assembly, inspiring biomaterial engineering innovations. For example, Hao et al. designed helical protein filaments that respond to pH changes by embedding histidine residues into the subunit interface (Figure 2b). These filaments exhibit reversible assembly and disassembly within narrow pH ranges [46]. Similarly, Neville et al. used computational modeling and machine learning to create patterned nanofibers from designed repeat proteins. Their approach enables the formation of tunable, extendable fibers with controlled mechanical properties [47].

Two-dimensional assemblies

Designing 2D protein assemblies is challenging, as it requires the precise interface design to introduce symmetry to achieve ordered and functional assemblies. Zhang et al. made a significant breakthrough in controlling protein self-assembly by engineering patchy proteins capable of forming diverse, precisely patterned 2D crystalline-like structures [48]. Similarly, Ben-Sasson et al. developed a computational approach to design two-component arrays by creating rigid, noncovalent interfaces between dihedral protein-building blocks. This method enables the construction of arrays with a p6m symmetry, assembling into nearly crystalline micrometer-scale lattices both *in vitro* and within living cells [49]. The introduced dihedral symmetry was essential, as it could compensate the errors in the design of trimers.

The innovative use of 2D protein arrays marks a significant advancement in the study of cellular polarity and cytoskeletal asymmetry. Cell polarity is integral to processes like cell division, differentiation, and tissue organization. Watson et al. explored how engineered Par protein complexes mimic and manipulate polarity, uncovering mechanisms by which Par complexes induce cytoskeletal asymmetry. This work revealed critical insights into the spindle orientation and symmetry breaking during mitosis, offering a high-resolution

Figure 2



Examples of extended protein assemblies. **(a)** Extended head-to-tail polymerization into single-chain filaments facilitated by orthogonal CC-linking motifs [45]. **(b)** Reversible assembly and disassembly of pH-dependent protein filaments (PDB ID: 8UAO) [46]. **(c)** Cyclic protein-building blocks docked into a two-component cage, which is then arrayed in a 3D lattice (PDB ID: 8CWY) [51]. **(d)** Protein fusions consisting of self-assembling oligomeric cores fused to crosslinking units (Spycatcher and Spytag, respectively) and form a tunable hydrogel network [52].

framework for studying polarity and its effects on asymmetric cell division in mammals [50].

Three-dimensional assemblies

A significant milestone in *de novo* protein design has been achieved by the creation of synthetic crystals [51]. Designing crystals has been a longstanding challenge due to the risk of off-target states during the assembly. Li et al. addressed this by adopting a hierarchical strategy: first, they designed protein cages using RPXDock and subsequently docked these cages into crystal lattices (Figure 2c) [51]. This approach effectively reduced the degrees of freedom, streamlining the design process. The resulting crystal structures were experimentally validated, and their lattices were shown to be tunable. Remarkably, two-component crystals were found in cell lysates, highlighting the robustness of the method.

In addition to ordered assemblies, less precisely defined 3D arrangements such as protein hydrogels offer exciting potentials. For instance, Mout et al. engineered protein hydrogels with programmable elastic properties (Figure 2d) [52]. They combined *de novo* protein dimers and oligomers linked by either covalent or noncovalent interactions. By systematically varying the linker lengths, they achieved a precise control over the elastic properties of the protein networks, showcasing a powerful platform for designing functional biomaterials. Disordered assemblies are also engineered liquid protein condensates, inspired by liquid–liquid phase separation [53], which play roles in various natural processes and have been constructed *de novo* from designed oligomerizing building blocks. These building blocks feature defined stoichiometry and tunable affinity, enabling the formation of protein compartments within bacterial or eukaryotic cells. Such compartments can be used to

sequester specific components and regulate their assembly or disassembly in response to chemical or biological signals [54].

Challenges

Despite the remarkable success of deep neural network–based methods in protein structure prediction and design, several challenges and opportunities remain in the design of assemblies. One problem is that AF 2/3 cannot predict entire assemblies due to size limitations. There have been many attempts at improving the algorithms [55–57], but all these methods work by first predicting the structure of dimers and trimers and then trying to assemble the full assembly from the components. This cannot consider the cooperative effects and is not predictive of *de novo*–designed fibers and cages. AF 2/3 has problems predicting even some single-chain *de novo* designs such as CCPO that have complex folding topology and are not homologous to any natural assemblies. The likely reason for this limitation is the absence of CCPO structures in the training data sets used by these algorithms as well as difficulties in capturing folds that differ from natural proteins. Those limitations may be overcome by the expansion of the experimentally determined data set.

De novo–designed assemblies can supersede many properties of natural assemblies and may be able to function as orthogonal to some natural systems, such as cytoskeleton or cellular compartments or even replace some of the components. It is however challenging to incorporate weak numerous interactions that may play a role in biological systems. Despite some limitations, the tunability and modularity of designed assemblies offer adaptability to align with emerging biological insights. One of the most important challenges for biomedical applications of synthetic systems is the risk of immunogenicity both via antibodies as well as T cell–mediated immunity. For monomeric therapeutics like minibinders, this can be mitigated through computationally optimized sequences that enhance fold stability and solubility, reducing major histocompatibility complex (MHC) class II presentation [58]. While so far immune response to minibinders has been reported as rather weak, as the solvent exposed is highly charged, this remains to be investigated in more detail and for long-term *in vivo* exposure.

Outlook

The spectacular advances in AI-based methods have significantly enhanced the success rates of experimental single-chain protein designs and pushed the boundary of reliably predicted monomers and assemblies toward larger sizes. Techniques such as AlphaFold2 [9], ProteinMPNN [34], and RPXDock [33] have made it easier to design *de novo* protein assemblies. The development of AI systems with an inherent understanding of symmetry — potentially reducing memory demands — is likely to be solved in the short- to mid-term. These innovations are poised to drive higher success rates, enabling the design of assemblies tailored to precise specifications without the need for extensive experimental optimization and accelerating the creation of functional protein architectures in the years to come. The design of functional protein assemblies is likely to transform biotechnology and biomedicine, with the *de novo* design of new molecular machines. While the optimization of catalytic efficiency still requires large-scale screening and directed evolution due to the insensitivity to point mutations, large language models can be used to significantly accelerate the iterative cycle of functional optimization [59].

The design of tunable, modular, and controllable assemblies enables not only the creation of versatile geometries but, more importantly, the development of programmable functions [60] with significant potential for biomedical and biotechnological applications. Designed nanocages enhance immune responses in vaccines through precise antigen presentation [29] and mimic photosynthetic systems by housing chlorophyll dimers [30]. pH-dependent CC assemblies offer possibilities for targeted cargo delivery [22], while pH-controlled disassembly of trimeric plugs to enable endosomal release [37]. Tunable nanorods could support protein presentation with tunable spacing, which may be beneficial for vaccine development [45]. Programmable protein hydrogels hold promise for designing extracellular matrices and synthetic organelles [52] while some *de novo* helical proteins can controllably assemble in cells forming compartments that improve enzymatic activity [61]. The most exciting applications are however likely to arise from the open access and user-friendly modeling platforms, such as Colab, AlphaFold3 and others, that will democratize protein design and make its use accessible to specialists in diverse areas of biotechnology.

CRediT authorship contribution statement

Jaka Snoj: Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Weijun Zhou**: Conceptualization, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Ajasja Ljubetič**: Conceptualization, Investigation, Supervision, Funding acquisition, Writing – original draft, Writing – review & editing. **Roman Jerala**: Conceptualization, Funding acquisition, Supervision, Project administration, Writing – review & editing.

Data Availability

No data were used for the research described in the article.

Declaration of Competing Interest

The authors declare no conflict of interest.

Acknowledgements

We would like to thank Dr. Rubul Mout for the 3D models of the hydrogel.

This work was funded by the Slovenian Research and Innovation Agency project N1-0323, N1-0377, J7-4640 and program P4-0176, ERC AdG project 101141584 PROFI and European Innovation Council project LoopOfFun, Project 101070817.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Pan X, Kortemme T: **Recent advances in de novo protein design: principles, methods, and applications**. *J Biol Chem* 2021, **296**:100558.

2. Zhu J, Avakyan N, Kakkis A, Hoffnagle AM, Han K, Li Y, Zhang Z, Choi TS, Na Y, Yu C-J, *et al*.: **Protein assembly by design**. *Chem Rev* 2021, **121**:13701-13796.

3. Woolfson DN: **A brief history of de novo protein design: minimal, rational, and computational**. *J Mol Biol* 2021, **433**:167160.

4. Li C, Zhang R, Wang J, Wilson LM, Yan Y: **Protein engineering for improving and diversifying natural product biosynthesis**. *Trends Biotechnol* 2020, **38**:729-744.

5. Huang PS, Boyken SE, Baker D: **The coming of age of de novo protein design**. *Nature* 2016, **537**:320-327.

6. Chu AE, Lu T, Huang P-S: **Sparks of function by de novo protein design**. *Nat Biotechnol* 2024, **42**:203-215.

7. Woolfson DN, Colwell LJ, Chen Z, Vorobieva AA, Polizzi NF, Stein A, Liu H, Parmeggiani F, Peacock A, Singh R, *et al*.: **How do you anticipate computational protein design will change biotechnology and therapeutic development?** *Cell Syst* 2024, **15**:994-999.

8. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, *et al*.: **Highly accurate protein structure prediction with AlphaFold**. *Nature* 2021, **596**:583-589.

9. Yang Z, Zeng X, Zhao Y, Chen R: **AlphaFold2 and its applications in the fields of biology and medicine**. *Signal Transduct Target Ther* 2023, **8**:115.

10. Abramson J, Adler J, Dunger J, Evans R, Green T, Pritzel A, Ronneberger O, Willmore L, Ballard AJ, Bambrick J, *et al*.: **Accurate structure prediction of biomolecular interactions with AlphaFold 3**. *Nature* 2024, **630**:493-500.

11. Baek M: **Efficient and accurate prediction of protein structures and interactions using RoseTTAFold**. *Acta Crystallogr Sect A Found Adv* 2022, **78**:a235-a235.

12. Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y, *et al*.: **Evolutionary-scale prediction of atomic-level protein structure with a language model**. *Science* 2023, **379**:1123-1130.

13. Anishchenko I, Pellock SJ, Chidyausiku TM, Ramelot TA, Ovchinnikov S, Hao J, Bafna K, Norn C, Kang A, Bera AK, *et al*.: **De novo protein design by deep network hallucination**. *Nature* 2021, **600**:547-552.

14. •• Watson JL, Juergens D, Bennett NR, Trippe BL, Yim J, Eisenach HE, Ahern W, Borst AJ, Ragotte RJ, Milles LF, *et al*.: **De novo design of protein structure and function with RFdiffusion**. *Nature* 2023, **620**:1089-1100.
A generative deep learning protein design method with high experimental success rates that can be used to design monomers, oligomers, or heterodimeric protein dimers.

15. •• Frank C, Khoshouei A, Fuβ L, Schiwietz D, Putz D, Weber L, Zhao Z, Hattori M, Feng S, de Stigter Y, *et al*.: **Scalable protein design using optimization in a relaxed sequence space**. *Science* 2024, **386**:439-445.
Design of largest single chain proteins to date, reaching 1000 amino acids in size with confirmed cryo-EM structure. AlphaFold 2 back-propagation coupled with ProteinMPNN.

16. Mallik BB, Stanislaw J, Alawathurage TM, Khmelinskaia A: **De novo design of polyhedral protein assemblies: before and after the AI revolution**. *Chembiochem* 2023, **24**:e202300117.

17. Li Y, Tian R, Shi H, Xu J, Wang T, Liu J: **Protein assembly: controllable design strategies and applications in biology**. *Aggregate* 2023, **4**:e317.

18. Hsia Y, Mout R, Sheffler W, Edman NI, Vulovic I, Park Y-J, Redler RL, Bick MJ, Bera AK, Courbet A, *et al*.: **Design of multi-scale protein complexes by hierarchical building block fusion**. *Nat Commun* 2021, **12**:2294.

19. Sahtoe DD, Praetorius F, Courbet A, Hsia Y, Wicky BIM, Edman NI, Miller LM, Timmermans BJR, Decarreau J, Morris HM, *et al*.: **Reconfigurable asymmetric protein assemblies through implicit negative design**. *Science* 2022, **375**:eabj7662.

20. • Huddy TF, Hsia Y, Kibler RD, Xu J, Bethel N, Nagarajan D, Redler R, Leung PJY, Weidle C, Courbet A, *et al*.: **Blueprinting extendable nanomaterials with standardized protein blocks**. *Nature* 2024, **627**:898-904.
An innovative approach that uses straight repeat proteins to easily tune the dimension of assemblies. Presents examples of circular, polyhedral, and 1D assemblies whose dimensions can be easily adjusted.

21. Gradišar H, Božič S, Doles T, Vengust D, Hafner-Bratkovič I, Mertelj A, Webb B, Šali A, Klavžar S, Jerala R: **Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments**. *Nat Chem Biol* 2013, **9**:362-366.

22. Aupič J, Lapenta F, Strmšek Ž, Merljak E, Plaper T, Jerala R: **Metal ion-regulated assembly of designed modular protein cages**. *Sci Adv* 2022, **8**:eabm8243.

23. Lapenta F, Aupič J, Vezzoli M, Strmšek Ž, Da Vela S, Svergun DI, Carazo JM, Melero R, Jerala R: **Self-assembly and regulation of protein cages from pre-organised coiled-coil modules**. *Nat Commun* 2021, **12**:939.

24. Ljubetič A, Lapenta F, Gradišar H, Drobnak I, Aupič J, Strmšek Ž, Lainšček D, Hafner-Bratkovič I, Majerle A, Krivec N, *et al*.: **Design of coiled-coil protein-origami cages that self-assemble in vitro and in vivo**. *Nat Biotechnol* 2017, **35**:1094-1101.

25. • Satler T, Hadži S, Jerala R: **Crystal structure of de novo designed coiled-coil protein origami triangle**. *J Am Chem Soc* 2023, **145**:16995-17000.
The first ever high-resolution structure of a CC protein origami structure. Interestingly, analysis of a crystal structure of this triangular CCPO showed the polypeptide chain to fold into a trefoil-type protein knot topology, and AlphaFold2 fails to predict the correct fold.

26. Aupič J, Strmšek Ž, Lapenta F, Ljubetič A, Jerala R, Pahovnik D, Pisanski T, Drobnak I: **Designed folding pathway of modular coiled-coil- based proteins**. *Nat Commun* 2021, **12**:940.

27. Cristie-David AS, Marsh ENG: **Metal-dependent assembly of a protein nano-cage**. *Protein Sci* 2019, **28**:1620-1629.

28. Snoj J, Lapenta F, Jerala R: **Preorganized cyclic modules facilitate the self-assembly of protein nanostructures**. *Chem Sci* 2024, **15**:3673-3686.

29. Ellis D, Dosey A, Boyoglu-Barnum S, Park Y-J, Gillespie R, Syeda H, Hutchinson GB, Tsybovsky Y, Murphy M, Pettie D, *et al*.: **Antigen spacing on protein nanoparticles influences antibody responses to vaccination**. *Cell Rep* 2023, **42**:113552.

30. Ennist NM, Wang S, Kennedy MA, Curti M, Sutherland GA, Vasilev C, Redler RL, Maffeis V, Shareef S, Sica AV, *et al.*: **De novo design of proteins housing excitonically coupled chlorophyll special pairs**. *Nat Chem Biol* 2024, **20**:906-915.

31. Dawson WM, Shelley KL, Fletcher JM, Scott DA, Lombardi L, Rhys GG, LaGambina TJ, Obst U, Burton AJ, Cross JA, *et al.*: **Differential sensing with arrays of de novo designed peptide assemblies**. *Nat Commun* 2023, **14**:383.

32. Laniado J, Meador K, Yeates TO: **A fragment-based protein interface design algorithm for symmetric assemblies**. *Protein Eng Des Sel* 2021, **34**:gzab008.

33. Sheffler W, Yang EC, Dowling Q, Hsia Y, Fries CN, Stanislaw J, Langowski MD, Brandys M, Li Z, Skotheim R, *et al.*: **Fast and versatile sequence-independent protein docking for nanomaterials design using RPXDock**. *PLoS Comput Biol* 2023, **19**:e1010680.

34. •• Dauparas J, Anishchenko I, Bennett N, Bai H, Ragotte RJ, Milles LF, Wicky BIM, Courbet A, de Haas RJ, Bethel N, *et al.*: **Robust deep learning-based protein sequence design using ProteinMPNN**. *Science* 2022, **378**:49-56.
A protein message parsing neural network that is used to design the side chains, given an existing backbone structures. Makes soluble and stable proteins.

35. de Haas RJ, Brunette N, Goodson A, Dauparas J, Yi SY, Yang EC, Dowling Q, Nguyen H, Kang A, Bera AK, *et al.*: **Rapid and automated design of two-component protein nanomaterials using ProteinMPNN**. *Proc Natl Acad Sci USA* 2024, **121**:e2314646121.

36. Meador K, Castells-Graells R, Aguirre R, Sawaya MR, Arbing MA, Sherman T, Senarathne C, Yeates TO: **A suite of designed protein cages using machine learning and protein fragment-based protocols**. *Structure* 2024, **32**:751-765 e11..

37. Yang EC, Divine R, Miranda MC, Borst AJ, Sheffler W, Zhang JZ, Decarreau J, Saragovi A, Abedi M, Goldbach N, *et al.*: **Computational design of non-porous pH-responsive antibody nanoparticles**. *Nat Struct Mol Biol* 2024, **31**:1404-1412.

38. Wicky BIM, Milles LF, Courbet A, Ragotte RJ, Dauparas J, Kinfu E, Tipps S, Kibler RD, Baek M, DiMaio F, *et al.*: **Hallucinating symmetric protein assemblies**. *Science* 2022, **378**:56-61.

39. Kibler RD, Lee S, Kennedy MA, Wicky BIM, Lai SM, Kostelic MM, Carr A, Li X, Chow CM, Nguyen TK, *et al.*: **Design of pseudosymmetric protein hetero-oligomers**. *Nat Commun* 2024, **15**:10684.

40. Lee S, Kibler RD, Ahn G, Hsia Y, Borst AJ, Philomin A, Kennedy MA, Huang B, Stoddard B, Baker D: **Four-component protein nanocages designed by programmed symmetry breaking**. *Nature* 2024, **636**:1-7.

41. Dowling QM, Park Y-J, Fries CN, Gerstenmaier NC, Ols S, Yang EC, Wargacki AJ, Dosey A, Hsia Y, Ravichandran R, *et al.*: **Hierarchical design of pseudosymmetric protein nanocages**. *Nature* 2024, https://doi.org/10.1038/s41586-024-08360-6

42. Courbet A, Hansen J, Hsia Y, Bethel N, Park Y-J, Xu C, Moyer A, Boyken SE, Ueda G, Nattermann U, *et al.*: **Computational design of mechanically coupled axle-rotor protein assemblies**. *Science* 2022, **376**:383-390.

43. Bromley EHC, Channon KJ: **Alpha-helical peptide assemblies giving new function to designed structures**. *Prog Mol Biol Transl Sci* 2011, **103**:231-275.

44. Jorgensen MD, Chmielewski J: **Recent advances in coiled-coil peptide materials and their biomedical applications**. *Chem Commun* 2022, **58**:11625-11636.

45. Mezgec K, Snoj J, Ulčakar L, Ljubetič A, Tušek Žnidarič M, Škarabot M, Jerala R: **Coupling of Spectrin repeat modules for the assembly of nanorods and presentation of protein domains**. *ACS Nano* 2024, **18**:28748-28763.

46. Shen H, Lynch EM, Akkineni S, Watson JL, Decarreau J, Bethel NP, Benna I, Sheffler W, Farrell D, DiMaio F, *et al.*: **De novo design of pH-responsive self-assembling helical protein filaments**. *Nat Nanotechnol* 2024, **19**:1016-1021.

47. Bethel NP, Borst AJ, Parmeggiani F, Bick MJ, Brunette TJ, Nguyen H, Kang A, Bera AK, Carter L, Miranda MC, *et al.*: **Precisely patterned nanofibres made from extendable protein multiplexes**. *Nat Chem* 2023, **15**:1664-1671.

48. Zhang S, Alberstein RG, De Yoreo JJ, Tezcan FA: **Assembly of a patchy protein into variable 2D lattices via tunable multiscale interactions**. *Nat Commun* 2020, **11**:3770.

49. Ben-Sasson AJ, Watson JL, Sheffler W, Johnson MC, Bittleston A, Somasundaram L, Decarreau J, Jiao F, Chen J, Mela I, *et al.*: **Design of biologically active binary protein 2D materials**. *Nature* 2021, **589**:468-473.

50. Watson JL, Krüger LK, Ben-Sasson AJ, Bittleston A, Shahbazi MN, Planelles-Herrero VJ, Chambers JE, Manton JD, Baker D, Derivery E: **Synthetic Par polarity induces cytoskeleton asymmetry in unpolarized mammalian cells**. *Cell* 2023, **186**:4710-4727 e35..

51. Li Z, Wang S, Nattermann U, Bera AK, Borst AJ, Yaman MY, Bick MJ, Yang EC, Sheffler W, Lee B, *et al.*: **Accurate computational design of three-dimensional protein crystals**. *Nat Mater* 2023, **22**:1556-1563.

52. • Mout R, Bretherton RC, Decarreau J, Lee S, Gregorio N, Edman NI, Ahlrichs M, Hsia Y, Sahtoe DD, Ueda G, *et al.*: **De novo design of modular protein hydrogels with programmable intra- and extracellular viscoelasticity**. *Proc Natl Acad Sci USA* 2024, **121**:e2309457121.
The first example of designed 3D protein crystals that form in the designed crystallographic group. The authors have developed a hierarchical approach, where they first design a polyhedral assembly that is then docked in the desired crystal symmetry.

53. Yuan C, Li Q, Xing R, Li J, Yan X: **Peptide self-assembly through liquid-liquid phase separation**. *Chem* 2023, **9**:2425-2445.

54. Ramšak M, Ramirez M, Hough LE, Shirts MR, Vidmar S, Eleršič Filipič K, Anderluh G, Jerala R: **Programmable de novo designed coiled coil-mediated phase separation in mammalian cells**. *Nat Commun* 2023, **14**:7973.

55. Shor B, Schneidman-Duhovny D: **CombFold: predicting structures of large protein assemblies using a combinatorial assembly algorithm and AlphaFold2**. *Nat Methods* 2024, **21**:477-487.

56. Bryant P, Pozzati G, Zhu W, Shenoy A, Kundrotas P, Elofsson A: **Predicting the structure of large protein complexes using AlphaFold and Monte Carlo tree search**. *Nat Commun* 2022, **13**:6028.

57. Mirabello C, Wallner B, Nystedt B, Azinas S, Carroni M: **Unmasking AlphaFold to integrate experiments and predictions in multimeric complexes**. *Nat Commun* 2024, **15**:8724.

58. Berger S, Seeger F, Yu T-Y, Aydin M, Yang H, Rosenblum D, Guenin-Macé L, Glassman C, Arguinchona L, Sniezek C, *et al.*: **Preclinical proof of principle for orally delivered Th17 antagonist miniproteins**. *Cell* 2024, **187**:4305-4317 e18..

59. Jiang K, Yan Z, Di Bernardo M, Sgrizzi SR, Villiger L, Kayabolen A, Kim BJ, Carscadden JK, Hiraizumi M, Nishimasu H, *et al.*: **Rapid in silico directed evolution by a protein language model with EVOLVEpro**. *Science* 2024, **386**, https://doi.org/10.1126/science.adr6006 eadr6006.

60. Kortemme T: **De novo protein design-From new structures to programmable functions**. *Cell* 2024, **187**:526-544.

61. Hilditch AT, Romanyuk A, Cross SJ, Obexer R, McManus JJ, Woolfson DN: **Assembling membraneless organelles from de novo designed proteins**. *Nat Chem* 2024, **16**:89-97.