Protein design

https://doi.org/10.1038/s41594-025-01560-2

Flexing protein assemblies

Yang Li & Yang Zhang

Structural flexibility is often considered a challenge in protein design, but it brings opportunities for creating protein assemblies that have several defined structures. A study now suggests that the controllable design of oligomorphic protein assemblies could be achieved by modulating structural flexibility.

Protein assemblies are complex macromolecular structures that are formed from multiple protein monomers (subunits) through certain interactions (interfaces). Of these, polyhedral assemblies represent a specific type that can self-assemble into regular geometric configurations with symmetrical properties. Natural polyhedral protein assemblies have a variety of biomedical applications, such as targeted delivery, enzyme encapsulation and vaccine design[1]. Inspired by the rich functions of these naturally evolved polyhedral protein assemblies, great efforts have been made to design new assemblies. However, existing computational approaches focus primarily on designing rigid and precise building blocks, whereas natural assemblies can adopt different architectures through structural flexibility, with certain regions within proteins exhibiting numerous different conformations. Now, writing in *Nature Structural and Molecular Biology*, Khmelinskaia et al.[2] reveal how local structural flexibility within protein building blocks can lead to oligomorphism – the ability of designed proteins to form a small, defined set of distinct architectures rather than a single intended structure.

For more than a decade, computational protein designers have focused on creating protein assemblies with strict symmetry and rigidity[3–6]. These designed proteins typically form a single, predetermined structure. However, those approaches contrast with naturally occurring protein assemblies, such as viral capsids and clathrin, which exhibit structural flexibility to perform specialized functions, including adapting to cargoes of various sizes[7].

Now, Khmelinskaia et al.[2] have characterized three computationally designed proteins – KWOCA 18, KWOCA 70, and I32-10 – that adopted unexpected structures which deviated from their intended architectures. The former two proteins are from previous designs created using the Degreaser protocol, which was implemented to eliminate cryptic transmembrane domains and improve protein secretion[8]. I32-10 was designed to form an icosahedral nanoparticle by the RosettaScripts framework[9]. Through comprehensive experimental analyses, including native mass spectrometry (nMS) and cryo-electron microscopy (cryo-EM), Khmelinskaia et al.[2] found that these designed subunits assemble into a limited number of well-formed structures rather than

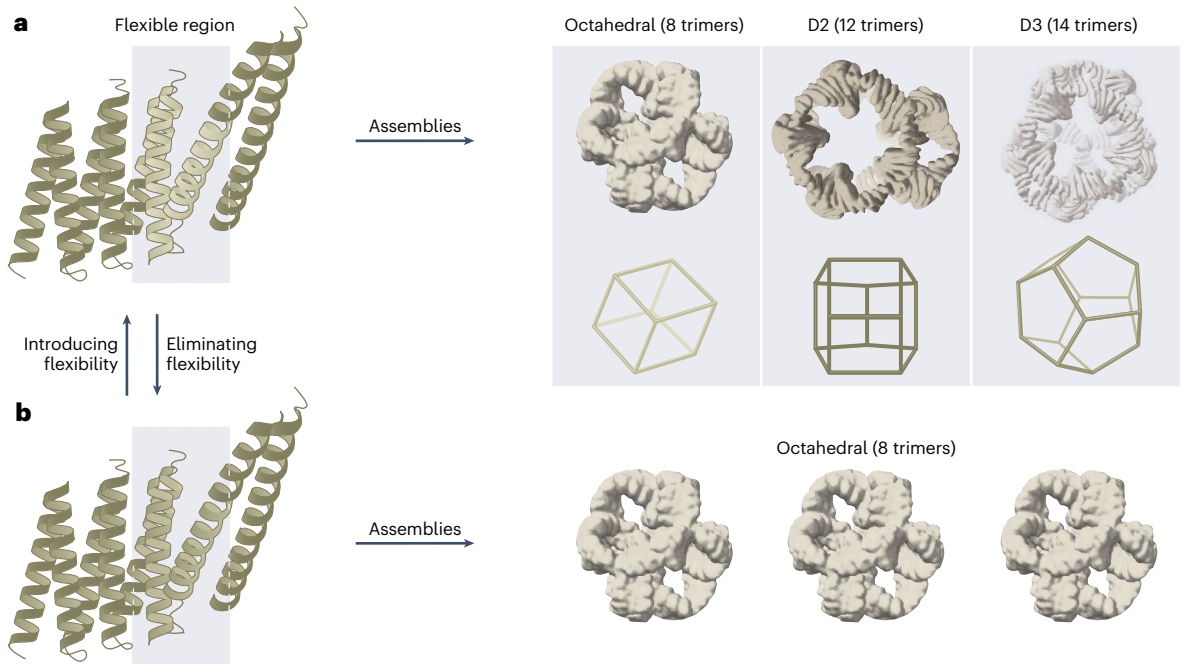

Fig. 1 | Controlling flexibility in protein assembly design. a, Khmelinskaia et al.[2] have found that the flexible junction region (boxed) in the trimeric building block of KWOCA 70 enables oligomorphic assembly, resulting in several distinct architectures with *D*2 and *D*3 symmetry, composed of 12 and 14 trimers, respectively. **b**, Redesigning the junction region to eliminate flexibility (KWOCA 70 D7) restricts the assembly to the originally intended monomorphic octahedral structure. Figure adapted from ref. 2, Springer Nature.

random aggregates. For example, KWOCA 70 – originally designed to form an octahedral assembly – produced two distinct architectures with D2 and D3 symmetry, comprising 12 and 14 trimeric building blocks, respectively (Fig. 1a). Similarly, KWOCA 18 formed two different structures, while I32-10 displayed even higher structural diversity with six different structures.

Khmelinskaia et al.[2] proposed that local flexibility might drive the formation of these alternative and heterogeneous conformations. To examine this hypothesis, they applied AlphaFold2[10], one of the most successful deep-learning-based models for predicting protein structure, and Rosetta software to identify potential flexible regions. AlphaFold 2 can estimate the quality of its structure predictions through predicted local distance difference test (pLDDT) scores. Lower pLDDT scores often indicate regions where a protein might be naturally flexible or disordered[11]. Their analysis revealed lower pLDDT scores and higher solvent accessibility in junction or hinge regions compared with in core domains (Fig. 1a), suggesting generally higher flexibility in the junction regions. Molecular dynamics simulations and cryo-EM density-fitting experiments further confirmed a substantial flexibility in these regions.

Next, the researchers wanted to further check whether eliminating flexibility in the junction region of the trimeric building blocks could recover the intended monomorphic assembly. To mitigate structural flexibility, Khmelinskaia et al.[2] used two cutting-edge artificial intelligence (AI)-based protein-design models: RFdiffusion[6], which can generate new protein backbone conformations using diffusion models, and ProteinMPNN[4], which optimizes the amino-acid sequence to stabilize these conformations. By redesigning the junction regions, the authors successfully stabilized the structure, leading to the KWOCA 70 D7 variant, which formed the intended octahedral assembly (Fig. 1b).

This research shows that protein flexibility can arise from specific structural features in junction regions and can be controlled through targeted redesign using advanced deep-learning models. Although these findings are based on specific cases, they open avenues for developing general models to regulate flexibility.

More importantly, the concept of flexibility control brings opportunities that existing rigid-design approaches cannot achieve. For example, precisely tuning structural flexibility could enable the development of adaptive cargo carriers capable of adjusting their size to encapsulate different molecules. This could lead to medical delivery systems that adjust their size to carry different drugs or target different cells, much like how natural transport systems in our cells work.

Finally, Khmelinskaia et al.[2] highlight the value of investigating and reporting unexpected 'failure' results in protein design. Rather than dismissing proteins that did not form their intended structures, these researchers investigated why and proposed a distinct design principle.

In summary, this study emphasizes the importance of controlled flexibility in computational protein design, especially as the field increasingly leverages deep-learning approaches to tackle complex tasks. While existing protein-design methods primarily ask, 'Will it fold?', this research proposes a new question: 'How will it flex?'.

Yang Li[1] & Yang Zhang[1,2,3] ✉

[1]Cancer Science Institute of Singapore, National University of Singapore, Singapore, Singapore. [2]Department of Computer Science, School of Computing, National University of Singapore, Singapore, Singapore. [3]Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Singapore.
✉e-mail: zhang@zhanggroup.org

Published online: 12 May 2025

References

1. Fries, C. N. et al. *Nat. Nanotechnol.* **16**, 1–14 (2021).
2. Khmelinskaia, A. et al. *Nat. Struct. Mol. Biol.* https://doi.org/10.1038/s41594-025-01490-z (2025).
3. Huang, P. S., Boyken, S. E. & Baker, D. *Nature* **537**, 320–327 (2016).
4. Dauparas, J. et al. *Science* **378**, 49–56 (2022).
5. Lutz, I. D. et al. *Science* **380**, 266–273 (2023).
6. Watson, J. L. et al. *Nature* **620**, 1089–1100 (2023).
7. Morris, K. L. et al. *Nat. Struct. Mol. Biol.* **26**, 890–898 (2019).
8. Wang, J. Y. et al. *Proc. Natl Acad. Sci. USA* **120**, e2214556120 (2023).
9. Bale, J. B. et al. *Science* **353**, 389–394 (2016).
10. Jumper, J. et al. *Nature* **596**, 583–589 (2021).
11. Ruff, K. M. & Pappu, R. V. *J. Mol. Biol.* **433**, 167208 (2021).

Competing interests

The authors declare no competing interests.