Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Structural Biology

Computational design of novel protein–protein interactions – An overview on methodological approaches and applications

Anthony Marchand[a,b], Alexandra K. Van Hall-Beauvais[a,b] and Bruno E. Correia[a,b]

Abstract

Protein–protein interactions (PPIs) govern numerous cellular functions in terms of signaling, transport, defense and many others. Designing novel PPIs poses a fundamental challenge to our understanding of molecular interactions. The capability to robustly engineer PPIs has immense potential for the development of novel synthetic biology tools and protein-based therapeutics. Over the last decades, many efforts in this area have relied purely on experimental approaches, but more recently, computational protein design has made important contributions. Template-based approaches utilize known PPIs and transplant the critical residues onto heterologous scaffolds. *De novo* design instead uses computational methods to generate novel binding motifs, allowing for a broader scope of the sites engaged in protein targets. Here, we review successful design cases, giving an overview of the methodological approaches used for templated and *de novo* PPI design.

Addresses

[a] Institute of Bioengineering, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[b] Swiss Institute of Bioinformatics (SIB), Lausanne, Switzerland

Corresponding author: Correia, Bruno E. (bruno.correia@epfl.ch)

(Van Hall-Beauvais A.K.)

(Correia B.E.)

Current Opinion in Structural Biology 2022, **74**:102370

This review comes from a themed issue on **Engineering and Design**

Edited by **Sarel J. Fleishman** and **Roy A. Mariuzza**

For a complete overview see the Issue and the Editorial

Available online xxx

https://doi.org/10.1016/j.sbi.2022.102370

Keywords

Protein–protein interactions, Computational design, *De novo* binders, Protein engineering.

Introduction

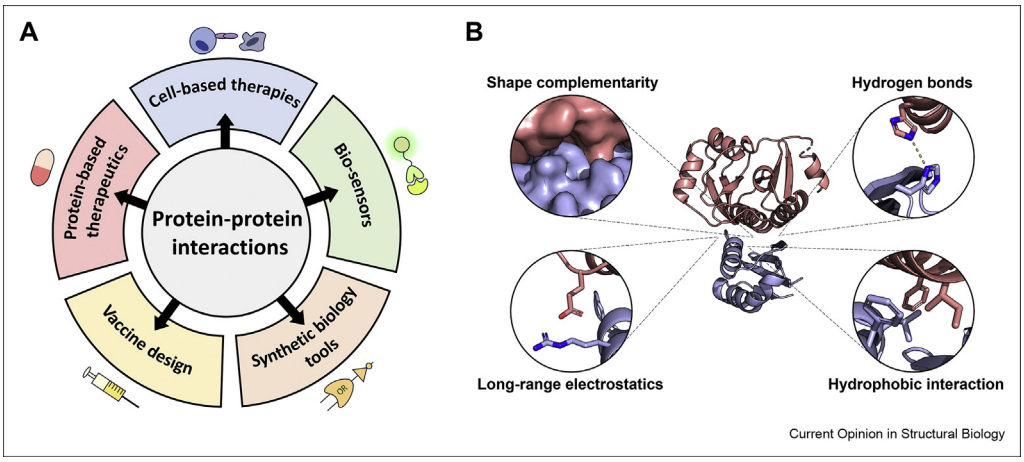
Proteins are among the most ubiquitous molecules of life and are likely the most versatile in terms of function, biophysical properties, and diversity. They perform primordial functions for cell signaling, structure, transport, catalysis, regulation, and defense, among others. Many fundamental protein functions involve association with other proteins, referred to as protein–protein interactions (PPIs) [1]. Native PPIs are involved in most cellular functions and their binding affinities span several orders of magnitude, with dissociation constants commonly ranging from picomolar to micromolar [2].

PPIs are involved in cell homeostasis processes that, if disrupted, can lead to numerous disease progressions, either pathogenic, degenerative or cancer-related [3]. Of the more than 645,000 disease-relevant PPIs, few have been successfully targeted by drugs [4]. A wide majority remain "undruggable" mainly due to featureless interfaces that lack defined binding pockets for small molecules [4]. In addition to studying PPIs as a source of potential druggable targets, PPIs are at the core of novel biotechnology tools such as protein-based therapeutics [5,6], cell therapies [7–9], bio-sensors [10–12], vaccine candidates [13–15] and other synthetic biology applications [16–19] (Figure 1A).

Similar to protein folding processes, protein association is driven by energy minimization. This process has several driving forces, including Van der Waals interactions, hydrophobicity, and electrostatic steering (also called long-range electrostatics) [20]. Hydrogen bonds and salt bridges stabilize the interaction and improve specificity [21–23]. The geometry of the molecular surface [24], with both shape and chemical complementarity of the interacting partners, plays a critical role for protein association [25] (Figure 1B).

In order to engineer novel PPIs, approaches such as *in vitro* evolution have been extensively used in the past decades [26–28]. However, one of the most important limitations of *in vitro* evolution is that it is "site agnostic," meaning that it is impossible to predict with certainty where the evolved binder will target the protein of interest. For the biological function of the binder, this is an important challenge that computational approaches attempt to solve. With the rise of

Figure 1



Overview of potential applications for novel PPIs and molecular features that drive protein association. A) Protein–protein interactions (PPI) have numerous applications for vaccine design, protein-based therapeutics (e.g. antibodies, inhibitors, etc.), cell-based therapies (e.g. CAR-T), bio-sensors (e.g. diagnostics), or as synthetic biology tools (e.g. ON/OFF-switch) B) Different structural features that can be designed by computational methods are necessary to engineer a strong PPI. These include good shape complementarity, hydrophobic patches, hydrogen bonds, and long-range electrostatic interactions (electrostatic steering) that stabilize the interaction and improve specificity.

computational methodologies numerous bioinformatics tools to predict, design, and engineer protein structures have been developed to address the limitations of the *in vitro* maturation techniques [29–31].

In this review, we will highlight successful design cases and discuss challenges in the computational design of PPIs. We group computational PPI design strategies in two categories: I) template-based design and II) *de novo* design. The first approach consists of transplanting a motif that mediates an existing PPI interface onto a new protein scaffold [32]. Despite its robustness and relatively high success rate, this strategy constrains PPI design to existing interfaces and precludes the possibility of targeting new sites. To explore a broader landscape of solutions, *de novo* design strategies aim to create completely new interactions starting from only the structure of the target protein [32]. However, engineering PPIs from scratch remains a non-trivial task requiring a detailed understanding of biomolecular interactions and stands as a stringent test of our understanding of the driving forces of PPIs.

Box 1. Key terms in the field of *de novo* PPI design.

Term	Definition
Binding motif	Continuous or discontinuous structural segments of amino acids that encompass the interface in a protein–protein interaction.
Hotspot	Key residues that have a large energetic contribution for the affinity of the protein–protein interaction.
One-sided design	Approach where the binder is designed and the target remains constant.
Two-sided design	Approach where both interfaces involved in the protein complex are designed.
Scaffold protein	Heterologous protein used as a recipient for the grafting of hotspot residues and/or binding motif(s).
De novo scaffold protein	Protein scaffolds that have been designed using computational approaches that model protein backbones and find the best sequences to stabilize the fold.
De novo PPI design	Design of novel protein–protein interactions without using explicit information of binding motifs used in native protein complexes.

Template-based design of protein–protein interactions

The template-based approach consists of transplanting the binding motif of an existing PPI into a new structural context (Figure 2). The motif is grafted onto a protein scaffold by sidechain grafting (i.e., backbone mimicry and then sidechain replacement) or backbone grafting (i.e., full motif transplantation including

sidechains and backbone). Alternatively, a *de novo* protein scaffold can be built around the binding motif.

One of the first cases of successful computational sidechain grafting design dates from the early 2000s by Liu and colleagues [33]. The Protein Data Bank (PDB) [34] was searched for scaffolds that contained three residues satisfying the geometric relationships of the C_α–C_β vectors of the three key residues of EPO required for binding to the receptor EPOR. Grafting only these three residues onto an appropriate scaffold resulted in a binder with 24 nM affinity to EPOR, highlighting the crucial contribution of hotspot residues in PPIs [35]. Several years later, a similar strategy [15,36] used backbone similarity searches to find host protein scaffolds onto which continuous viral epitopes were transplanted. To address higher structural epitopes, this approach was extended to transplant discontinuous backbone segments of a viral epitope [37]. In both cases, the epitope transplantation gave binding affinities to the antibody in the nanomolar range and high structural agreement to the original epitope.

Sidechain grafting has been successfully used to transplant helical motifs onto *de novo* designed scaffolds. Successful examples include the design of candidate protein-based inhibitors against influenza hemagglutinin (HA) and botulinum neurotoxin B (BoNT/B), using known HA binders or natural BoNT/B target respectively as a helical template motif for subsequent grafting on *de novo* miniprotein scaffolds [38]. Future research efforts in protein-based therapeutics will benefit from the generation of highly stable *de novo* scaffolds presenting functional motifs.

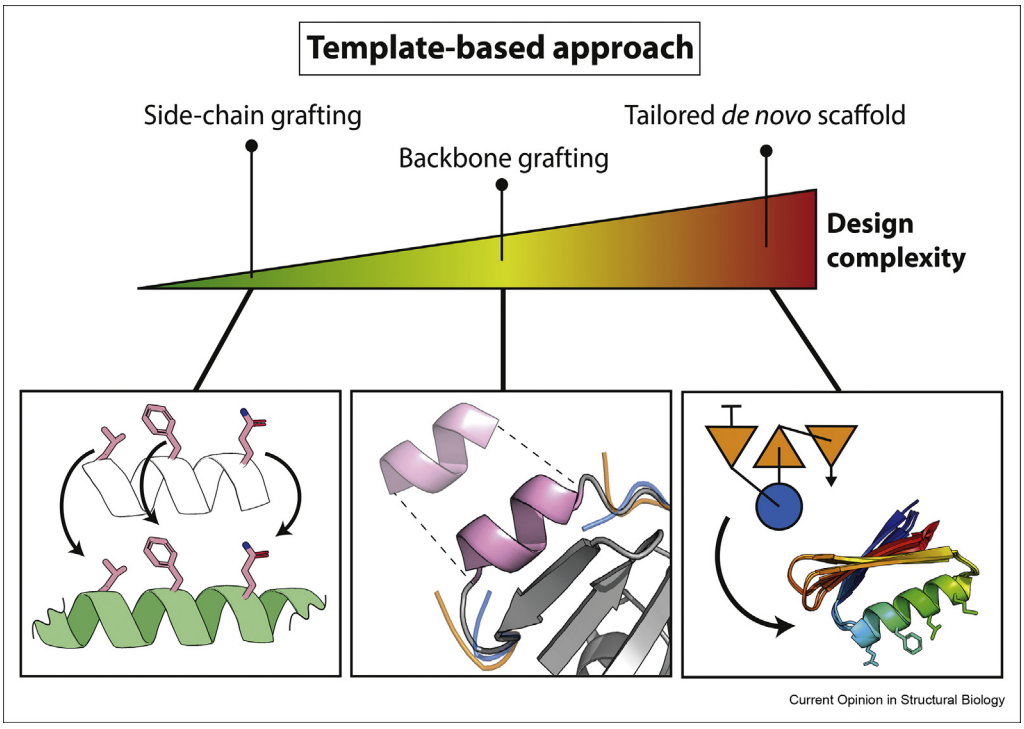
Recently, the sidechain grafting approach for PPI engineering demonstrated useful applications for synthetic biology and the design of small molecule-controlled switches. The underlying principle consists of repurposing an existing PPI that can be targeted by a known small molecule to control its dissociation. Giordano-Attianese and colleagues [7] repurposed the binding of BH3-motif to Bcl-XL by grafting the sidechains onto a globular scaffold protein. This led to a 3.9 pM affinity for Bcl-XL and created a protein switch controlled by a Bcl-XL inhibitor. The novel switch was incorporated into the chimeric antigen receptor (CAR) of T cells and was shown to turn off killing activity upon the addition of the small molecule. Work by Shui and colleagues has extended the logical behavior of this system creating a multidomain architecture that, upon the addition of a small molecule, triggers the association of the two protein subunits [16]. These applications demonstrate promising applications for translational research in the domain of cell engineering.

Motif grafting by sidechain replacement faces limitations when the motif is too complex to find a structurally compatible protein scaffold. Grafting approaches have been described where both sidechains and backbone are grafted onto protein scaffolds. Azoitei and colleagues designed epitope-scaffolds by selecting scaffolds based on N- and C- termini alignments to identify sites in proteins where the motif was grafted and the connection regions were further refined and designed [39]. Such strategy was also successfully utilized to transplant a complex binding site from an HIV epitope, composed of two discontinuous segments that were required to present a precise three-dimensional structure to mediate productive binding to the antibody B12 [40]. The two segments of the epitope were grafted in a stepwise fashion and multiple rounds of *in vitro* evolution were performed to optimize the binding affinity of the designed scaffold, highlighting the difficulty of grafting complex sites onto protein scaffolds.

To address more complex epitopes, the Fold From Loops (FFL) protocol was proposed as an alternative by folding *de novo* scaffold proteins to stabilize the binding motif of interest [13]. The FFL approach was first used to embed a viral epitope from RSV onto a *de novo* folded and designed three-helix bundle protein. Several of the designs bound with picomolar affinities to a site-specific monoclonal antibody and the designs showed, for the first time, the ability to elicit neutralizing antibodies in non-human primates. Further, the FFL protocol was utilized by Procko et al. to design a protein inhibitor against an Epstein Barr-Viral (EBV) Bcl2-homolog called BHRF1 [41]. Extensive *in vitro* maturation was necessary to stabilize and improve the affinity to BHRF1 and the success rate of functional designs was rather low. The FFL methodology was also used by Bryan et al. [42] to develop small, ultra-stable mini-protein scaffolds designed around a five amino acid stretch of PDL-2, one of the native binding partners of PD-1, resulting in a 100 nM binder for PD-1.

Two main shortcomings of the FFL protocol came to light: I) the lack of compatibility for multiple discontinuous motifs; II) the incorporation of the binding partner during the folding-design simulations for the optimization of additional contacts and as a constraint for the sampling the conformational/sequence space. Bonet et al. improved FFL, by developing a Rosetta framework called FunFolDes, which addressed these drawbacks [43]. This novel approach successfully functionalized "functionless" folds by incorporating the Respiratory Syncytial Virus protein F (RSVF) site IV on a *de novo* protein called TOP7. Another intrinsic limitation of the FFL approach was its reliance on existing structures, either native or *de novo* designed. To circumvent this drawback, Sesterhenn and colleagues proposed the TopoBuilder, a protocol for the assembly of *de novo* topologies conditioned to the structure of the motif of interest [14]. Upon the assembly of the topologies with the embedded functional/binding motif, the FunFolDes

Figure 2



PPI design methods using the template-based approach. The template-based approach can be subdivided (from lowest to highest complexity) in I) side-chain grafting, II) backbone grafting, or III) use of a tailored *de novo* scaffold. Side-chain grafting transplants binding motifs from an existing PPI onto a heterologous scaffold that stabilizes the interaction between these side-chains and the binding target. In backbone grafting the transplantation involves the full backbone and side chains of the binding motif involved in a PPI onto a heterologous scaffold. Backbone grafting often poses the difficulty of modeling realistic backbones and finding suitable stabilizing sequences in the connecting segments between the grafted motif and the scaffold. Finally, in a more tailored approach, a *de novo* scaffold could be built around the motif of interest by specifying the arrangement of secondary structure elements to generate a three-dimensional topology.

folding and design protocol is used for sequence generation. This work contributed to the development of different candidate vaccine immunogens that elicited neutralizing antibodies against specific viral epitopes and created a series of functional molecules that were used for different synthetic biology applications [14,44]. Other methods of grafting hotspots to *de novo* scaffolds led to rapid design of a nanomolar SARS-CoV-2 binder that neutralized SARS-CoV-2 [45] and the use of *de novo* peptides as a scaffold for PPI disruptors [46]. These methods highlight the potential uses for *de novo* scaffolds, albeit dependent on known interactions.

Overall, these methods allow for PPI design with various levels of complexity, however, they are limited to known binding interactions. To broaden the landscape of targetable protein interfaces, *de novo* approaches to generate motifs that can mediate novel PPIs is needed.

De novo design of protein–protein interactions

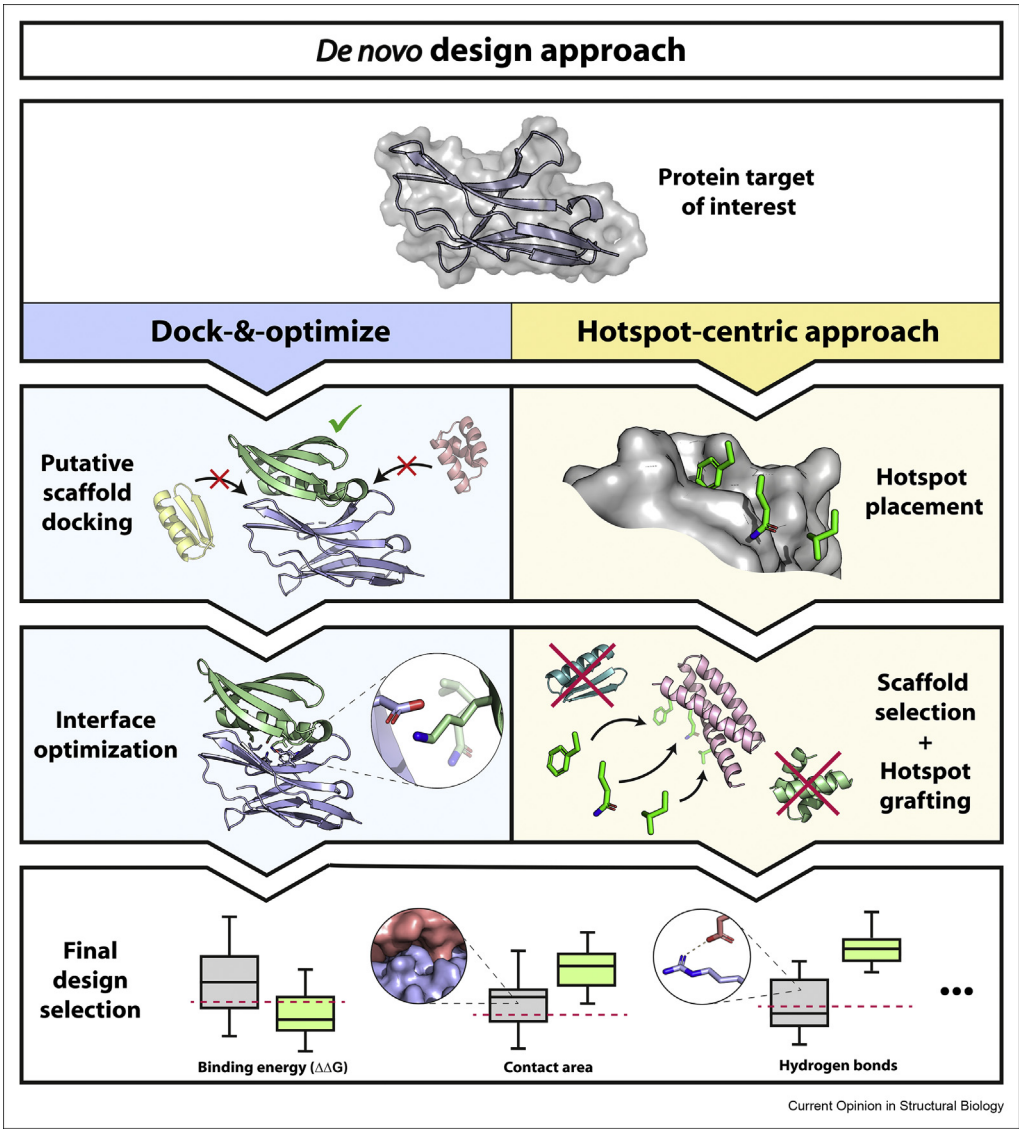
In the context of this review, *de novo* strategies for the design of protein interactions rely only on the structural information of the target, which we generally refer to as one-sided design. *De novo* design strategies are subdivided in: I) dock-&-optimize; II) hotspot-centric approaches [32] (Figure 3). The dock-&-optimize approach consists of two stages. First, hundreds of protein scaffolds are computationally docked on the target protein to find configurations with favorable shape complementarities. Second, interface residues of the best candidates are computationally designed to improve the binding propensity. Alternatively, the hotspot-centric approach first requires the placement of a few clustered hotspot residues before grafting onto a suitable scaffold protein that will be further refined [47].

One of the early publications [48] in this field introduced a Rosetta-based protocol, called DDMI (Docking, Design, Minimization and Interface), following the dock-&-optimize approach. The DDMI protocol is a two-step approach which uses rigid-body docking to find a suitable orientation for the partner scaffold and then iterates between sequence design and energy minimization to settle the interface to the lowest

energy state. Their best candidate, named "Spider Roll," used a single pre-selected scaffold to target the kinase domain of p21-activated kinase 1 (PAK1), and showed only weak affinity ($K_D \approx 100$ μM). This study and others using a dock-&-optimize approach [49,50] were strong demonstrations that more accurate energy force fields are needed, as well as larger pools of scaffold candidates and, due to all these limitations, *in vitro* evolution techniques may be required to further optimize the putative binders.

In an alternative route, hotspot-centric approaches were proposed. Fleishmann and colleagues were the first to implement a hotspot-centric method to target a conserved surface site on the stem of the influenza hemagglutinin (HA) [51]. The design approach consisted of docking disembodied residues, selecting suitable scaffolds, and refining the interface with RosettaDesign [52]. Out of 73 designs screened by yeast display, 2 showed binding to HA including one with an apparent affinity of 200 nM. Two rounds of

Figure 3



PPI design methods using the *de novo* approach. The *de novo* design approach consists of two alternative strategies: I) Dock-&-optimize or II) Hotspot-centric approach. The first is a two-step method that combines the docking of putative scaffolds and then an interface optimization aiming to minimize the binding energy between the target and the most appropriate scaffold. In the second method, hotspot residues are searched for, placed on the interface of interest, and then grafted on a scaffold which is suitable for both side-chain orientations and target interface. In both methodologies, a final selection based on different metrics (binding energy, contact area, hydrogen bonds, etc.) is needed to reduce the pool of designs to be tested.

affinity maturation were performed, providing insight into the sub-optimal features of the designed protein: I) void volumes at the interface should be minimized and backbone minimization can facilitate the choice of suitable residues; II) complementary electrostatic charges which remain outside of the hydrogen-bond range (∼3 Å) should not be underestimated; III) the energetic cost for the desolvation of charged residues in close contact with non-polar amino acids should not be neglected. In conclusion, the hotspot-centric strategy yielded a higher affinity binder than the dock-&-optimize approach, noting the fact that these were optimized by *in vitro* evolution.

Later on, Procko et al. [49] targeted the hen egg lysozyme (HEL) using the same approach with two polar hotspot residues. Scaffold candidates were docked, refined, and selected to satisfy both the disembodied hotspot residues and the complementarity for the target. Out of 21 designs, one showed a modest affinity of 7 μM and required two rounds of directed evolution and four mutations to obtain a final affinity of 8 nM. This experiment, as the previous one, had to rely partially on known interacting residues, as well as *in vitro* maturation techniques to improve binders to an acceptable affinity, although requiring only a few mutations. Despite these promising results, both studies showed that hotspot residue placement was a promising approach, however the need for *in vitro* optimization and the low success rates support that improved energy functions and methods are still necessary.

Recently, computational tools such as the rotamer interaction field (RIF) docking have been proposed to search for *de novo* hotspots for PPI and protein-ligand design without prior knowledge. Briefly, billions of disembodied residue conformations are docked on the target interface with the aim of introducing hydrogen bonds and hydrophobic packing interaction to create an energetically favorable interface. All RIF rotamers are stored and can be rapidly sampled for scaffold matching using a docking grid-based search algorithm [53]. RIF docking and a miniprotein scaffold library were used for the rapid generation of protein-based therapeutics against SARS-CoV-2 spike protein, with *de novo* designs having affinity lower than 1 nM after *in vitro* evolution optimization [54]. Ultimately, with the same strategy, the same group was also able to generalize the hotspot-centric approach proposed by Fleishman and colleagues [51] by generating at least one binder for 12 different target proteins [55]. These publications were among the first to demonstrate complete *de novo* hotspot generation for PPI engineering. Intriguingly, most binders designed so far rely on helical structures, limiting the landscape of binding motifs available for PPI designs, especially when working with disembodied residues. This approach still seems dependent on a large library screening (15′000−100′000 designs per target), although it undoubtedly pushed the frontier in *de novo* PPI design.

Challenges and perspectives

The methodology for designing novel PPIs has evolved rapidly in the past years. Templated design approaches take advantage of known binding partners and transplant either the sidechains of hotspot residues or the backbone and sidechains of the region of interest. Although this method is reliable and has led to the design of many successful binders, it is limited by relying on known binding partners. *De novo* design does not rely on such interactions and is a more difficult problem that poses a rigorous test to our understanding of the principles that drive protein−protein interactions. Recently, RIFDock has allowed for *de novo* hotspots to be predicted without prior knowledge of binding partners and from these hotspots, high affinity binders have been developed.

Despite these successes, there are still challenges that need to be addressed. There is a low success rate for *de novo* designs, and the designs that are successful often need rounds of *in vitro* evolution to improve the affinity. One plausible explanation could be the lack of proper energy functions to capture long-range interactions and the effect of water molecules. A study also found that poorly designed buried hydrogen bonds account for most of the failure in *de novo* PPI attempts [56]. Computational tools aiming to design broad hydrogen bond networks, such as HBnet [57], or the introduction of score penalty for buried unsatisfied polar atoms [58] may help future *de novo* design pipelines to tackle challenging polar interfaces. Additionally, more work must be done to extend these methods beyond helical motifs. Although helical binders can be successful, opening this strategy to more than one secondary structure would further increase the breadth of structural space that could be covered. Finally, it seems an emergent theme that most of the *de novo* PPI designs target known PPI interfaces, leaving unsolved challenges in targeting arbitrary target sites that may have low interface forming propensity. Despite these challenges, new tools for protein engineers are being developed that can address these difficulties. Newly introduced machine learning based software such as MaSIF [59] allows protein engineers to predict novel binding sites and possible binding partners. The introduction of Alphafold [30] and RoseTTAFold [60] allows for the prediction of three-dimensional protein structures with just the amino acid sequence. These tools and others will assist protein engineers in further studies.

Despite the difficulty of understanding and accurately designing novel PPIs, the number of computational methods available is expanding steadily and will undoubtedly lead to a higher success rates and benefit to translational research with biomedical applications.

Funding

We thank the support of the European Research Council (starting grant no. 716058), the Swiss National Science Foundation (grant no. 310030_188744), the NCCR Molecular Systems Engineering (www.nccr-mse.ch) and the NCCR Chemical Biology (www.nccr-chembio.ch).

Conflict of interest statement

Nothing declared.

References

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

- • of special interest
- •• of outstanding interest

1. Janin J, Bahadur RP, Chakrabarti P: **Protein–protein interaction and quaternary structure**. *Q Rev Biophys* 2008, **41**: 133–180.

2. Smith MC, Gestwicki JE: **Features of protein-protein interactions that translate into potent inhibitors: topology, surface area and affinity**. *Expet Rev Mol Med* 2012, **14**. e16–e16.

3. Gonzalez MW, Kann MG: **Chapter 4: protein interactions and disease**. *PLoS Comput Biol* 2012, **8**:e1002819.

4. Mabonga L, Kappo AP: **Protein-protein interaction modulators: advances, successes and remaining challenges**. *Biophys Rev* 2019, **11**:559–581.

5. Robertson NS, Spring DR: **Using peptidomimetics and constrained peptides as valuable tools for inhibiting protein–protein interactions**. *Molecules* 2018, **23**.

6. Zhang G, Andersen J, Gerona-Navarro G: **Peptidomimetics targeting protein-protein interactions for therapeutic development**. *Protein Pept Lett* 2018, **25**:1076–1089.

7. • 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.*: **A computationally designed chimeric antigen receptor provides a small-molecule safety switch for T-cell therapy**. *Nat Biotechnol* 2020, **38**:426–432.

Using sidechain grafting of an existing PPI, the authors were able to design a drug-controlled PPI that can lead to an OFF-switch system in T cell. This work highlighted the strength of PPI engineering for the development of synthetic cell-based therapies and other synthetic biology applications.

8. Wu C-Y, Roybal KT, Puchner EM, Onuffer J, Lim WA: **Remote control of therapeutic T cells through a small molecule-gated chimeric receptor**. *Science* 2015, **350**. aab4077–aab4077.

9. Zajc CU, Dobersberger M, Schaffner I, Mlynek G, Pühringer D, Salzer B, Djinović-Carugo K, Steinberger P, De Sousa Linhares A, Yang NJ, *et al.*: **A conformation-specific ON-switch for controlling CAR T cells with an orally available drug**. *Proc Natl Acad Sci USA* 2020, **117**:14926.

10. Scheller L, Strittmatter T, Fuchs D, Bojar D, Fussenegger M: **Generalized extracellular molecule sensor platform for programming cellular behavior**. *Nat Chem Biol* 2018, **14**:723–729.

11. Quijano-Rubio A, Yeh H-W, Park J, Lee H, Langan RA, Boyken SE, Lajoie MJ, Cao L, Chow CM, Miranda MC, *et al.*: **De novo design of modular and tunable protein biosensors**. *Nature* 2021, **591**:482–487.

12. Langan RA, Boyken SE, Ng AH, Samson JA, Dods G, Westbrook AM, Nguyen TH, Lajoie MJ, Chen Z, Berger S, *et al.*: **De novo design of bioactive protein switches**. *Nature* 2019, **572**:205–210.

13. Correia BE, Bates JT, Loomis RJ, Baneyx G, Carrico C, Jardine JG, Rupert P, Correnti C, Kalyuzhniy O, Vittal V, *et al.*: **Proof of principle for epitope-focused vaccine design**. *Nature* 2014, **507**:201–206.

14. •• Sesterhenn F, Yang C, Bonet J, Cramer JT, Wen X, Wang Y, Chiang C-I, Abriata LA, Kucharska I, Castoro G, *et al.*: **De novo protein design enables the precise induction of RSV-neutralizing antibodies**. *Science* 2020, **368**:eaay5051.

This work introduces a novel computational tool, the TopoBuilder, which aims to assemble tailored *de novo* topologies to stabilize complex motifs of interest. TopoBuilder proved itself to be a powerful tool for vaccines designs, which can be adapted to *de novo* PPI designs.

15. Correia BE, Ban Y-EA, Holmes MA, Xu H, Ellingson K, Kraft Z, Carrico C, Boni E, Sather DN, Zenobia C, *et al.*: **Computational design of epitope-scaffolds allows induction of antibodies specific for a poorly immunogenic HIV vaccine epitope**. *Structure* 2010, **18**:1116–1126.

16. Shui S, Gainza P, Scheller L, Yang C, Kurumida Y, Rosset S, Georgeon S, Di Roberto RB, Castellanos-Rueda R, Reddy ST, *et al.*: **A rational blueprint for the design of chemically-controlled protein switches**. *Nat Commun* 2021, **12**:5754.

17. Chen Z, Kibler RD, Hunt A, Busch F, Pearl J, Jia M, VanAernum ZL, Wicky BIM, Dods G, Liao H, *et al.*: **De novo design of protein logic gates**. *Science* 2020, **368**:78–84.

18. Edgell CL, Smith AJ, Beesley JL, Savery NJ, Woolfson DN: **De novo designed protein-interaction modules for in-cell applications**. *ACS Synth Biol* 2020, **9**:427–436.

19. Lajoie MJ, Boyken SE, Salter AI, Bruffey J, Rajan A, Langan RA, Olshefsky A, Muhunthan V, Bick MJ, Gewe M, *et al.*: **Designed protein logic to target cells with precise combinations of surface antigens**. *Science* 2020, **369**:1637–1643.

20. Shaul Y, Schreiber G: **Exploring the charge space of protein–protein association: a proteomic study**. *Protein Struct Funct Bioinform* 2005, **60**:341–352.

21. Schreiber G: **CHAPTER 1 protein–protein interaction interfaces and their functional implications**. In *Protein–protein interaction regulators*. The Royal Society of Chemistry; 2021: 1–24.

22. Nilofer C, Sukhwal A, Mohanapriya A, Kangueane P: **Protein-protein interfaces are vdW dominant with selective H-bonds and (or) electrostatics towards broad functional specificity**. *Bioinformation* 2017, **13**:164–173.

23. Xu D, Tsai CJ, Nussinov R: **Hydrogen bonds and salt bridges across protein-protein interfaces**. *Protein Eng Des Sel* 1997, **10**:999–1012.

24. Richards FM: **AREAS, volumes, packing, and protein structure**. *Annu Rev Biophys Bioeng* 1977, **6**:151–176.

25. Lawrence MC, Colman PM: **Shape complementarity at protein/protein interfaces**. *J Mol Biol* 1993, **234**:946–950.

26. Chao G, Lau WL, Hackel BJ, Sazinsky SL, Lippow SM, Wittrup KD: **Isolating and engineering human antibodies using yeast surface display**. *Nat Protoc* 2006, **1**:755–768.

27. Packer MS, Liu DR: **Methods for the directed evolution of proteins**. *Nat Rev Genet* 2015, **16**:379–394.

28. Kieke MC, Cho BK, Boder ET, Kranz DM, Wittrup KD: **Isolation of anti-T cell receptor scFv mutants by yeast surface display**. *Protein Eng Des Sel* 1997, **10**:1303–1310.

29. Leman JK, Weitzner BD, Lewis SM, Adolf-Bryfogle J, Alam N, Alford RF, Aprahamian M, Baker D, Barlow KA, Barth P, *et al.*: **Macromolecular modeling and design in Rosetta: recent methods and frameworks**. *Nat Methods* 2020, **17**:665–680.

30. •• 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.

By using multi-sequence alignments and a machine learning algorithm, the authors proposed Alphafold to predict protein structures. Alphafold outperformed current structure predictors and is rapidly expanding towards multimer prediction and protein design, both of interest for *de novo* PPI engineering.

31. Yang J, Yan R, Roy A, Xu D, Poisson J, Zhang Y: **The I-TASSER Suite: protein structure and function prediction**. *Nat Methods* 2015, **12**:7–8.

32. Schreiber G, Fleishman SJ: **Computational design of protein–protein interactions**. *Catal Regul Protein-Protein Interact* 2013, **23**:903–910.

33. Liu S, Liu S, Zhu X, Liang H, Cao A, Chang Z, Lai L: **Nonnatural protein–protein interaction-pair design by key residues grafting**. *Proc Natl Acad Sci USA* 2007, **104**:5330.

34. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE: **The protein Data Bank**. *Nucleic Acids Res* 2000, **28**:235–242.

35. Kortemme T, Baker D: **A simple physical model for binding energy hot spots in protein-protein complexes**. *Proc Natl Acad Sci USA* 2002, **99**:14116–14121.

36. Ofek G, Guenaga FJ, Schief WR, Skinner J, Baker D, Wyatt R, Kwong PD: **Elicitation of structure-specific antibodies by epitope scaffolds**. *Proc Natl Acad Sci USA* 2010, **107**: 17880–17887.

37. McLellan JS, Correia BE, Chen M, Yang Y, Graham BS, Schief WR, Kwong PD: **Design and characterization of epitope-scaffold immunogens that present the motavizumab epitope from respiratory syncytial Virus**. *J Mol Biol* 2011, **409**:853–866.

38. Chevalier A, Silva D-A, Rocklin GJ, Hicks DR, Vergara R, Murapa P, Bernard SM, Zhang L, Lam K-H, Yao G, *et al.*: **Massively parallel de novo protein design for targeted therapeutics**. *Nature* 2017, **550**:74–79.

39. Azoitei ML, Ban Y-EA, Julien J-P, Bryson S, Schroeter A, Kalyuzhniy O, Porter JR, Adachi Y, Baker D, Pai EF, *et al.*: **Computational design of high-affinity epitope scaffolds by backbone grafting of a linear epitope**. *J Mol Biol* 2012, **415**: 175–192.

40. Azoitei ML, Correia BE, Ban Yih-En A, Carrico C, Kalyuzhniy O, Chen L, Schroeter A, Huang P, McLellan JS, Kwong PD, *et al.*: **Computation-guided backbone grafting of a discontinuous motif onto a protein scaffold**. *Science* 2011, **334**:373–376.

41. Procko E, Berguig GY, Shen BW, Song Y, Frayo S, Convertine AJ, Margineantu D, Booth G, Correia BE, Cheng Y, *et al.*: **A computationally designed inhibitor of an Epstein-Barr viral Bcl-2 protein induces apoptosis in infected cells**. *Cell* 2014, **157**:1644–1656.

42. • Bryan CM, Rocklin GJ, Bick MJ, Ford A, Majri-Morrison S, Kroll AV, Miller CJ, Carter L, Goreshnik I, Kang A, *et al.*: **Computational design of a synthetic PD-1 agonist**. *Proc Natl Acad Sci USA* 2021, **118**. e2102164118.

The authors used the Fold From Loop (FFL) methodology to develop small, ultra-stable mini-protein scaffolds designed around a binding motif against the immune checkpoint PD-1. This work represents a strong proof-of-concept for the development of protein-based therapeutics and new tools for the generation of *de novo* scaffold libraries for PPI design.

43. Bonet J, Wehrle S, Schriever K, Yang C, Billet A, Sesterhenn F, Scheck A, Sverrisson F, Veselkova B, Vollers S, *et al.*: **Rosetta FunFolDes - a general framework for the computational design of functional proteins**. *PLoS Comput Biol* 2018, **14**. e1006623–e1006623.

44. Yang C, Sesterhenn F, Bonet J, van Aalen EA, Scheller L, Abriata LA, Cramer JT, Wen X, Rosset S, Georgeon S, *et al.*: **Bottom-up de novo design of functional proteins with complex structural features**. *Nat Chem Biol* 2021, **17**:492–500.

45. Linsky TW, Vergara R, Codina N, Nelson JW, Walker MJ, Su W, Barnes CO, Hsiang T-Y, Esser-Nobis K, Yu K, *et al.*: **De novo design of potent and resilient hACE2 decoys to neutralize SARS-CoV-2**. *Science* 2020, **370**:1208.

46. Fletcher JM, Horner KA, Bartlett GJ, Rhys GG, Wilson AJ, Woolfson DN: **De novo coiled-coil peptides as scaffolds for disrupting protein–protein interactions**. *Chem Sci* 2018, **9**: 7656–7665.

47. Fleishman SJ, Corn JE, Strauch E-M, Whitehead TA, Karanicolas J, Baker D: **Hotspot-centric de novo design of protein binders**. *J Mol Biol* 2011, **413**:1047–1062.

48. Jha RK, Leaver-Fay A, Yin S, Wu Y, Butterfoss GL, Szyperski T, Dokholyan NV, Kuhlman B: **Computational design of a PAK1 binding protein**. *J Mol Biol* 2010, **400**:257–270.

49. Procko E, Hedman R, Hamilton K, Seetharaman J, Fleishman SJ, Su M, Aramini J, Kornhaber G, Hunt JF, Tong L, *et al.*: **Computational design of a protein-based enzyme inhibitor**. *J Mol Biol* 2013, **425**:3563–3575.

50. Dang LT, Miao Y, Ha A, Yuki K, Park K, Janda CY, Jude KM, Mohan K, Ha N, Vallon M, *et al.*: **Receptor subtype discrimination using extensive shape complementary designed interfaces**. *Nat Struct Mol Biol* 2019, **26**:407–414.

51. Fleishman SJ, Whitehead TA, Ekiert DC, Dreyfus C, Corn JE, Strauch E-M, Wilson IA, Baker D: **Computational design of proteins targeting the conserved stem region of influenza hemagglutinin**. *Science* 2011, **332**:816.

52. Kuhlman B, Dantas G, Ireton GC, Varani G, Stoddard BL, Baker D: **Design of a novel globular protein fold with atomic-level accuracy**. *Science* 2003, **302**:1364–1368.

53. Dou J, Vorobieva AA, Sheffler W, Doyle LA, Park H, Bick MJ, Mao B, Foight GW, Lee MY, Gagnon LA, *et al.*: **De novo design of a fluorescence-activating β-barrel**. *Nature* 2018, **561**: 485–491.

54. •• Cao L, Goreshnik I, Coventry B, Case JB, Miller L, Kozodoy L, Chen RE, Carter L, Walls AC, Park Y-J, *et al.*: **De novo design of picomolar SARS-CoV-2 miniprotein inhibitors**. *Science* 2020, **370**:426.

The authors used a rotamer interaction field (RIF) docking pipeline to search for *de novo* hotspots for therapeutic PPI designs against SARS-CoV-2 spike protein. This represents the first generalization of the hotspot-centric approach without prior hotspot knowledge.

55. • Cao L, Coventry B, Goreshnik I, Huang B, Park JS, Jude KM, Marković I, Kadam RU, Verschueren KHG, Verstraete K, *et al.*: **Robust de novo design of protein binding proteins from target structural information alone**. *Nature* 2022, https://doi.org/10.1038/s41586-022-04654-9.

This paper extends the application of the rotamer interaction field (RIF) and is the most recent version of de novo PPI designs directed against 12 different targets.

56. Stranges PB, Kuhlman B: **A comparison of successful and failed protein interface designs highlights the challenges of designing buried hydrogen bonds**. *Protein Sci Publ Protein Soc* 2013, **22**:74–82.

57. Boyken SE, Chen Z, Groves B, Langan RA, Oberdorfer G, Ford A, Gilmore JM, Xu C, DiMaio F, Pereira JH, *et al.*: **De novo design of protein homo-oligomers with modular hydrogen-bond network–mediated specificity**. *Science* 2016, **352**: 680–687.

58. • Coventry B, Baker D: **Protein sequence optimization with a pairwise decomposable penalty for buried unsatisfied hydrogen bonds**. *PLoS Comput Biol* 2021, **17**:e1008061.

This publication proposes an algorithm that introduces a penalty for buried unsatisfied polar atoms and can be implemented in modeling suite such as Rosetta. It offers an interesting solution to deal with buried unsatisfied polar atoms, which remain a global challenge when designing novel PPIs.

59. • Gainza P, Sverrisson F, Monti F, Rodolà E, Boscaini D, Bronstein MM, Correia BE: **Deciphering interaction fingerprints from protein molecular surfaces using geometric deep learning**. *Nat Methods* 2020, **17**:184–192.

MaSIF is a geometric deep learning framework for rapid PPI interface and partner prediction. The algorithm demonstrated accurate predictions three to four orders of magnitude faster than other state-of-the-art algorithm, as well as promising potentials for an adaptation for novel PPI engineering.

60. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Schaeffer RD, *et al.*: **Accurate prediction of protein structures and interactions using a three-track neural network**. *Science* 2021, **373**: 871–876.