



Towards functional *de novo* designed proteins

William M Dawson¹, Guto G Rhys¹ and Derek N Woolfson^{1,2,3}

Our ability to design completely *de novo* proteins is improving rapidly. This is true of all three main approaches to *de novo* protein design, which we define as: minimal, rational and computational design. Together, these have delivered a variety of protein scaffolds characterised to high resolution. This is truly impressive and a major advance from where the field was a decade or so ago. That all said, significant challenges in the field remain. Chief amongst these is the need to deliver functional *de novo* proteins. Such designs might include selective and/or tight binding of specified small molecules, or the catalysis of entirely new chemical transformations. We argue that, whilst progress is being made, solving such problems will require more than simply adding functional side chains to extant *de novo* structures. New approaches will be needed to target and build structure, stability and function simultaneously. Moreover, if we are to match the exquisite control and subtlety of natural proteins, design methods will have to incorporate multi-state modelling and dynamics. This will require more than black-box methodology, specifically increased understanding of protein conformational changes and dynamics will be needed.

Addresses

¹ School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

² School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK

³ BrisSynBio, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK

Corresponding author: Woolfson, Derek N (D.N.Woolfson@bristol.ac.uk)

Current Opinion in Chemical Biology 2019, 52:102–111

This review comes from a themed issue on **Synthetic biomolecules**

Edited by **Johan Winne** and **Annemieke Madder**

For a complete overview see the [Issue](#) and the [Editorial](#)

Available online 20th July 2019

<https://doi.org/10.1016/j.cbpa.2019.06.011>

1367-5931/© 2019 Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Introduction

De novo protein design is said to have come of age [1]. From the early *de novo* proteins confirmed by high-resolution structures [2–4], the field has advanced rapidly with new scaffolds covering all- α [5,6*,7], all- β [8], and mixed- α/β and $\alpha + \beta$ structural space [9,10*,11]. In addition, side-chain constellations can be controlled exquisitely to introduce networks of hydrogen bonds throughout target

structures [12], which, in turn, can improve the design and characterisation of *de novo* membrane proteins [13].

However, the ability to design functional *de novo* proteins from scratch, or to embellish existing *de novo* scaffolds with new functions, is still in its infancy. Herein, we use terms like ‘functional protein design’ for any stably folded *de novo* protein frameworks that incorporate interactions with small or large molecules, catalytic activity and so on. With notable exceptions—for example, reports of a functional ion transporter [14], a *de novo* designed catalytic triad [15**], and a highly efficient *de novo* enzyme [16**]—general design principles for functional protein design are sparse. Indeed, it may be that overoptimised *de novo* proteins, which are often hyperthermally stable, may not make good platforms for functional design, as it is known that dynamics play essential roles in ligand binding and catalysis [17–19].

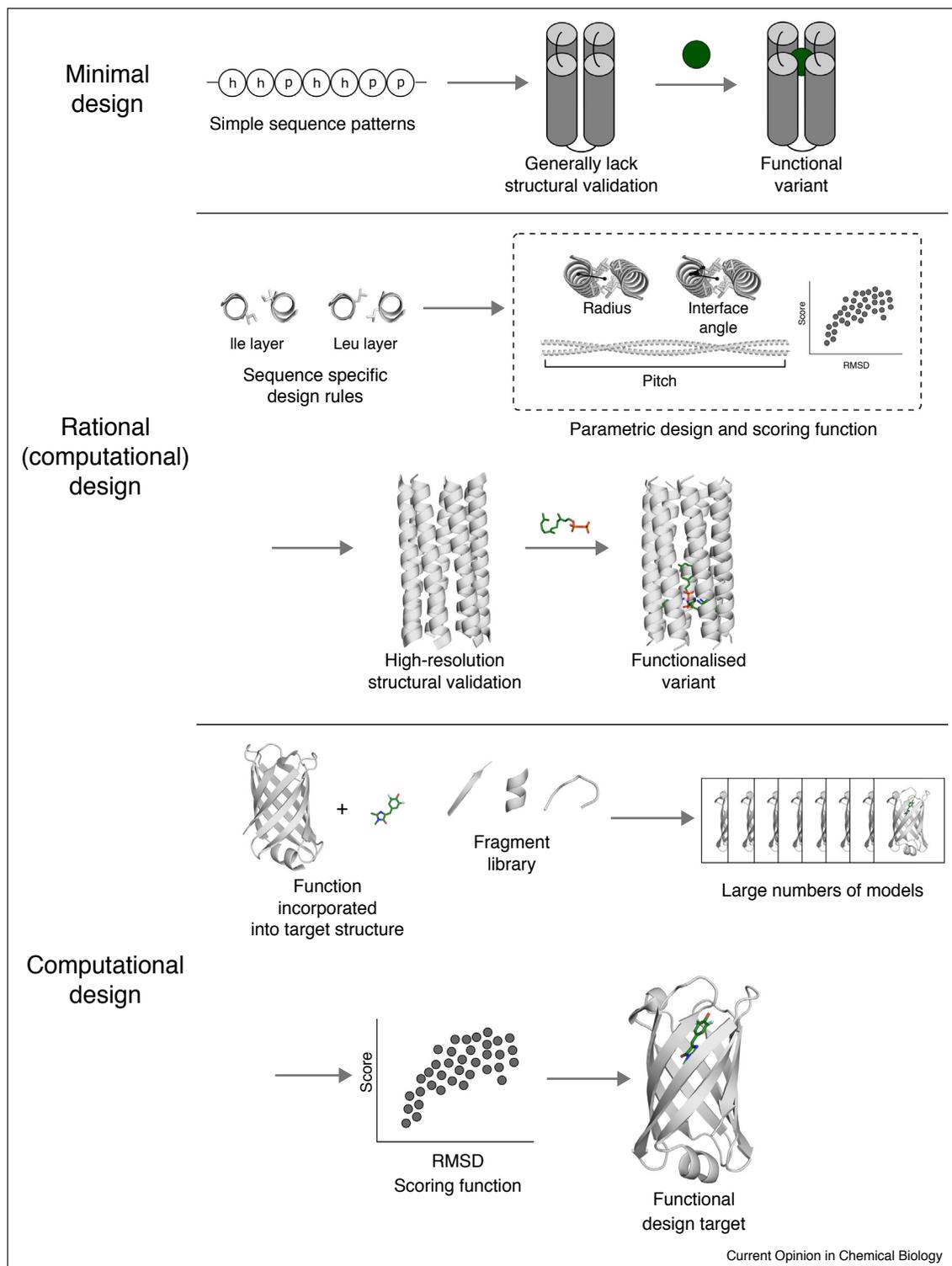
Herein, we focus on truly *de novo* proteins rather than those achieved through protein engineering or redesign—that is, where functions are improved in or introduced to natural proteins. Of course, the latter have led to novel enzymes and ligand-binding proteins [20–22]. Whilst impressive, protein engineering relies on the inherent stability of natural scaffolds and their tolerance to mutation, and often uses the randomness of directed evolution to access the targeted function [23]. By contrast, *de novo* protein design removes the dependence on naturally evolved scaffolds, and has the potential for a deeper understanding of the contribution that every side chain makes towards the structure, stability and function of *de novo* proteins. Of course, this is an extremely challenging approach and its goals are ambitious.

Notable advances have also been made in introducing metal-binding and protein–protein interactions into *de novo* proteins (see recent reviews Refs. [24–26]). However, these pose different challenges to those laid out herein, and are only mentioned in passing in this review.

From minimal, through rational, to computational design

There is no single approach to protein design. However, the field can be split broadly into three different approaches (Figure 1). In minimal design binary patterns of polar (p) and hydrophobic (h) residues are used to define a target structure [27,28]. α -Helices lend themselves to this as they can be directed to fold and assemble with sequence patterns of the type hpphppp. As a result, the vast majority of work in this area has targeted four-helix bundles. Rational design goes a step further by incorporating more-specific sequence-to-structure

Figure 1



Overview of minimal, rational and computational design approaches. **Minimal design** relies on binary patterning of hydrophobic (h) and polar (p) to define a target structure. Despite considerable effort, few of these have been validated through to high-resolution structures. Nonetheless, such minimal scaffolds have been modified to introduce ligand binding and catalysis. The vast majority of minimal protein design has focused on four-helix bundle proteins. In **rational design**, which can incorporate computational methods, binary patterns are supplemented by specific sequence-to-structure relationships for the target; for example, subtly different combinations of Ile and Leu side chains in coiled-coil interfaces can direct alternate oligomer states. Such rules can be very powerful when coupled with parametric design to build, score and rank multiple models for a target. This approach has now led to many high-resolution structures including for structures not known or rare in biology; for example, a

relationships, or design rules, often garnered from inspection of the sequences and structures of natural proteins [29]. In both minimal and rational designs, extant stable scaffolds are then modified to produce functional variants. Computational design generally uses databases of structural motifs, for example, short peptide fragments, to construct the target scaffolds and to fit many primary sequences onto these [1]. In this way, large numbers of models are built and scored with an energy function. This allows variants to be ranked ahead of experimental studies. This approach also facilitates the introduction of functionality early in the design process; that is, stable proteins can be built around a target function [30**]. Another advantage of computational design over the minimal and rational approaches is that it allows access to more-complex structures [8]. That said, the combination of rational and computational approaches, particularly using parametric design to generate the backbone scaffolds, is proving powerful in delivering a variety of *de novo* proteins that both mimic natural protein structures and expand upon them.

The sections below build on these ideas emphasising functional designs that have been achieved thus far within each approach.

Minimal design of functional four-helix bundles

DeGrado, Hecht and Dutton have pioneered the concepts of minimal and rational *de novo* protein design (reviewed extensively in Refs. [27,33,34]). In short, these combine chemical intuition about protein structure and basic sequence-to-structure relationships to deliver straightforward designed protein scaffolds (Figure 1). Key targets in these endeavours have been four-helix bundles, which involve the coalescence of amphipathic helices encoded by self-associating peptides or within single polypeptide chains. For some time, these have been adapted to deliver functional designs.

In the ‘maquette’ approach [35], Dutton and coworkers [36,37] iteratively develop a minimal four-helix scaffold that is characterised at each step (Figure 2). Sheehan *et al.* use this to design a biliverdin-binding protein [37]: starting from a molten-globule state with promiscuous binding [38], potential binding sites are probed experimentally through cysteine-ligation scanning, and the resulting binding site is stabilised further by rational design.

Recently, Watkins *et al.* demonstrate how powerful minimalistic design can be in functional *de novo* design.

The authors reposition heme C binding sites within a foregoing four-helix maquette (Figure 2) [39]. The resulting construct shows activity for oxidation and oxidative dehalogenation [16**]. Impressively, the kinetic analysis reveals that this *de novo* catalyst is as proficient as natural oxidoreductase enzymes, but with enhanced chemical and thermal stability.

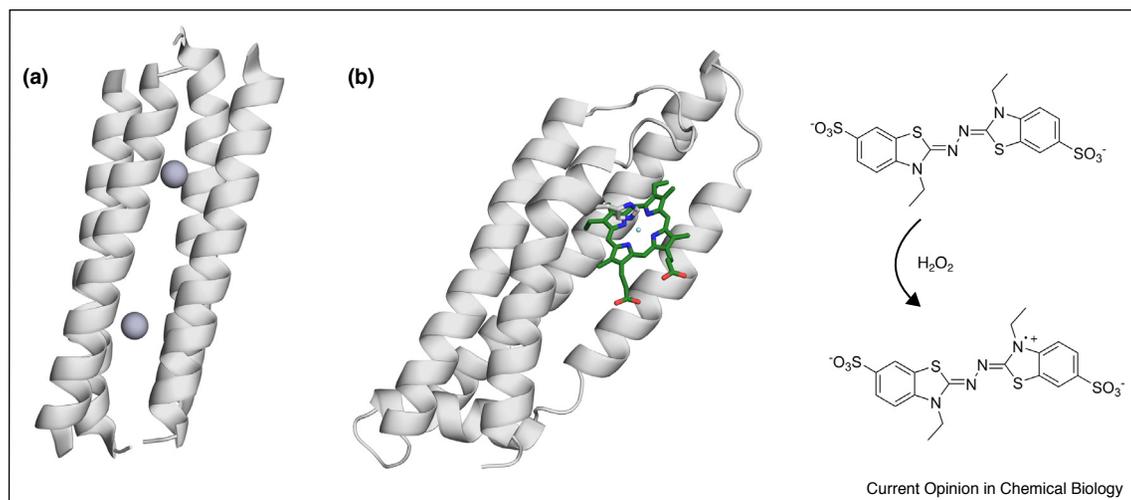
Similarly, Donnelly *et al.* apply binary patterns of polar (*p*) and hydrophobic (*h*) residues—for example, *phpphhpphpphhp* sequences—to produce a catalytic four-helix bundle from two helix-loop-helix monomers [40*]. Building on previous work to select enzyme-like functions from libraries of *de novo* sequences [41], the authors find one construct that hydrolyses ferric enterobactin with enantiomeric selectivity. Further investigations show that five polar/charged amino acids in the core are key to activity. This is the first example of a *de novo* protein that is essential for maintaining living cells, though it was achieved through selection rather than rational design.

Following a tradition established by Lear *et al.* [42], Lalaurie *et al.* employ minimal design to deliver a *de novo* membrane protein [43]. By analysing a small subset of natural membrane proteins, the authors develop a low-complexity leucine-rich sequence. This embeds in membranes and binds heme, although attempts at using this in catalysis appear to result in degradation of the cofactor.

Overall, the minimalistic approach to design has been successful for four-helix bundles. However, the lack of high-resolution structures for many of these designs emphasises the need to consider the stereochemical arrangement of the residues, that is, side-chain packing, to achieve well-ordered protein cores and, with these, better-defined 3D structures. For those cases where structural data have been obtained it has been for apoproteins, that is, the protein scaffold without ligand or catalytic residues/prosthetic groups present, rather than functional *de novo* four-helix bundles [44–47]. Arai *et al.* present the structure of a minimally designed four-helix bundle with primitive esterase and lipase activity [48]. However, this protein is shown to form a domain-swapped dimeric species, rather than the expected monomeric species. Further highlighting the limitations of minimal design, computational approaches have led to high-resolution structural data of both inert and functional four-helix bundles [30**,49].

(Figure 1 Legend Continued) heptameric coiled coil, CC-Hept (PDB: 4PNA). Functions have then been incorporated into these scaffolds; for example, a model of farnesyl diphosphate (green) bound in a heptameric coiled coil [31*]. **Computational design** often uses databases of protein fragments to assign thousands of potential amino-acid sequences to the design target. Energy functions are used to rank the designs with the most favourable being taken forward for experimental validation. Increasingly, functionality is being incorporated in the initial design stage rather than being appended to a stable scaffold; for example, the design of a fluorescence-activating β -barrel (PDB: 6CZ1) [32**].

Figure 2



Minimalistic protein design. **(a)** A non-functional 'maquette' with mercury ions (grey spheres) bound in the core (PDB: 1M3W) [45]. This simple protein fold is used as an example due to the lack of structural data of functional variants. **(b)** A computational model of the *de novo* C45 protein and the reaction catalysed by this peroxidase [16**]. The ligand, heme C ($C_{34}H_{34}N_4O_4Fe$, green), is activated upon addition of hydrogen peroxide, enabling oxidation of ABTS as shown.

Rational parametric design of functional assemblies

Rules-based or rational protein design and computational design do not have to be mutually exclusive (Figure 1). By incorporating design rules into computational design algorithms, the number of models that need to be built and scored can be reduced dramatically. Parametric design lends itself to this. Here, target protein folds are described mathematically with a minimal number of parameters. Not surprisingly given their simplicity and potential regularity, *de novo* four-helix bundles have been designed parametrically [50]. However, except for a single example [30**], high-resolution structural data validating the models remain elusive.

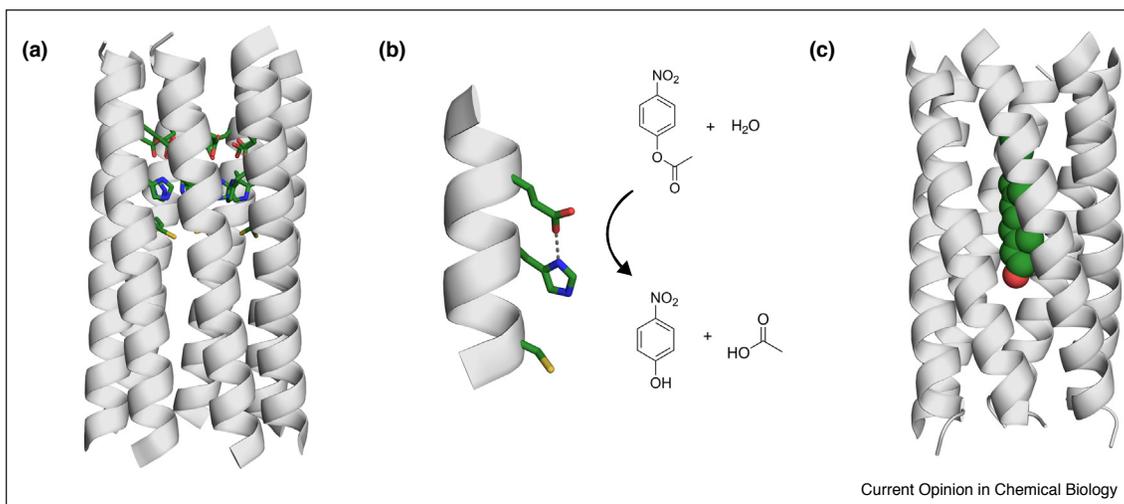
Coiled-coil proteins also lend themselves to parameterisation. Before moving onto computational coiled-coil design, it is worth highlighting the designability of these structures because of their relatively straightforward sequences and structures. For example, Harbury *et al.* describe variants of the GCN4 leucine zipper with combinations of Ile and Leu residues in the core to produce parallel dimeric, trimeric and tetrameric structures, and to deliver rules for oligomer-state selection [51]; *n.b.*, wild-type GCN4 leucine zipper is a parallel homodimer. Fletcher *et al.* use these rules to design fully *de novo* homomeric dimers to tetramers, and Thomas *et al.* supplement the rules to deliver heterodimers with a range of dissociation constants [52,53]. These designs have proven useful as highly stable and robust building blocks for supramolecular assembly in materials science and

synthetic biology [54–57]. However, they have no inherent function.

Crick was the first to describe coiled-coil structures parametrically [58]. Starting with the tight geometry of the α helix, he reasoned that coiled-coil structures could be defined by the radius and pitch of a superhelical assembly with two or more α helices plus a parameter (the interface angle) for the relative twist between helices (Figure 1). Numerous implementations of Crick's equations are now available to generate coiled-coil scaffolds computationally and to build *de novo* sequences into these [59–65]. Thomson *et al.* adapt coiled-coil design principles and rules and combine them with parametric computational design to target larger discrete coiled-coil assemblies [5]. By increasing the size of the hydrophobic interface presented by the component α helices, *de novo* pentamers, hexamers and heptamers are achieved. These are termed α -helical barrels as they possess a central channel. Whilst these structures are not functional themselves, the fully accessible channels are prime targets for functionalisation [66]. Similarly, Huang *et al.* use parametric design within Rosetta to create (hyperstable) trimeric, tetrameric and pentameric coiled coils [67]. Similarly, these *de novo* assemblies are not functional themselves.

Burton *et al.* use rational design to introduce hydrolase activity into the heptameric coiled-coil scaffold, CC-Hept [15**]. In this design, each helix contributes a Cys-His-Glu catalytic triad to the lumen of the barrel (Figure 3). Kinetic analysis shows CC-Hept-CHE to be on par with other *de novo* and engineered catalysts, although these are all

Figure 3



Rationally designed functional proteins. **(a)** The first example of a catalytically active triad in a fully *de novo* protein (PDB: 5E2C) [15[•]]. **(b)** A single helix showing the Cys-His-Glu triad used to catalyze the hydrolysis of 4-nitrophenol acetate. **(c)** Farnesol (green, C₁₅H₂₆O) bound in the lumen of a hexameric coiled-coil assembly (PDB: 6EIZ) [31[•]].

poor compared with natural esterases and design or engineered systems that incorporate metals [68]. This heptameric hydrolase is the first example of a functional catalytic triad incorporated into a completely *de novo* designed scaffold.

In addition to hydrolysing substrates, the α -helical barrels can bind other small molecules. Thomas *et al.* perform a systematic study to probe the size and shape of molecules that can be sequestered within the hydrophobic channels [31[•]]. Without modification, the pentamer, hexamer and heptamer all bind small, hydrophobic molecules with low μ M affinities. Specificity for negatively or positively charged molecules has been added through the rational placement of ionisable side chains in the lumen (Figure 3).

Fragment-based computational design beyond protein engineering

As protein structures increase in complexity, more sophisticated approaches are needed to access more-elaborate architectures. By harnessing the power of computers, thousands of designs can be generated and analysed *in silico* at scales beyond minimal and rational design. The most widespread approach is fragment-based design, which has three aspects: libraries of fragments or motifs are taken from structural databases, algorithms are developed to combine these to assemble target structures, and scoring functions are used to assess both the assembled structures and sequences that best fit onto them (Figure 1) [69–72]. This is epitomised by the Rosetta suite for computational protein design developed by the Baker group [73].

There are numerous examples of new functions being engineered into natural proteins using these methods [74–76], including opioid binders [77], an amino-acid binder [78] and Schiff-base-forming enzyme [79]. A related approach mimics nature by combining larger protein fragments [80] and has proven successful for generating non-functional *de novo* proteins [81,82]. Whilst relying heavily on the evolutionary traits of the parent enzymes, these chimeric proteins have activities that match their natural counterparts. Lapidoth *et al.* adapt this approach in an automated fashion to create TIM barrels, a ubiquitous fold consisting of eight α -helices and eight β -strands arranged in tandem, that is, $(\beta\alpha)_8$, with hydrolase and lactonase activity [83[•]].

Huang *et al.* and Marcos *et al.* have designed *de novo* proteins incorporating cavities with potential for catalysis or small-molecule binding [10[•],11]. In the first study, a *de novo* four-fold symmetric $(\beta\alpha)_8$ -barrel is designed using RosettaRemodel developed for repeat proteins [10[•]]. This is of interest as TIM barrels are the most common enzyme topology found in nature. The second study develops design principles for curved β sheets [11]. Applying analyses of bulges and register shifts in naturally curved β sheets, the authors use RosettaDesign to obtain nine *de novo* scaffolds with pockets that could be modified for ligand binding. Serendipitously, the crystal structure of one scaffold has a ligand bound in the cavity, highlighting the potential for functionalisation.

Despite these successes, fragment-based design might be considered a ‘black-box approach’ with few design rules or general principles being gleaned. For example,

impressively, Rocklin *et al.* apply a massive-scale approach to protein design, coupling stability against protease degradation with yeast display to deliver a large number of stable, *de novo* mini-proteins [84]. That said, a design rule to emerge from this study is that certain charged side chains near the termini of helices stabilise the constructs, which is in agreement with conclusions drawn from a previous study combining bioinformatics and rational design of single α helices [85].

Designing in function from the beginning

Most of the studies described above focus on the design of the *de novo* scaffold before functionalisation or improving an already established functional *de novo* protein. Though there are clear examples to the contrary [15^{••},31[•],39], extant *de novo* designed scaffolds may not have suitable cavities or sites for every targeted function, or the post-incorporation of the necessary functional residues may prove problematic. Indeed, for natural proteins it is well documented that small changes, even distal to the binding site/active site, can have large effects on the function [86]. Therefore, design strategies that incorporate, or at least consider, the functional aspect at an early stage could ultimately lead to more successful outcomes.

Polizzi *et al.* describe such an approach to design a porphyrin-binding four-helix bundle. Of course, four-helix bundles that bind porphyrins have been designed previously. In fact, tight binding to a porphyrin cofactor in *de novo* four-helix bundles is common due to the hydrophobicity of the ligand and the strengths of side chain-metal interactions [87,88]. However, the lack of structural data from these studies has precluded validation of these designs. With this in mind, Polizzi *et al.* simultaneously design a well-folded hydrophobic core and a ligand-binding site into a four-helix bundle (Figure 4) [30^{••}]. By factoring the long-range influence of residues distal to the ligand binding site, the

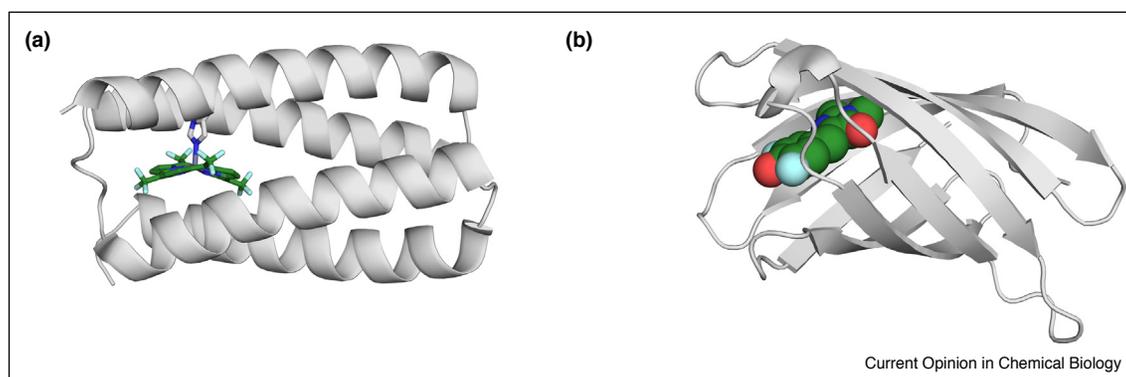
authors improve on earlier designs [89] and obtain a high-resolution structure.

Dou *et al.* take a similar approach to design ligand-binding β -barrel proteins [32^{••}]. Recognising irregularities in sheets, the authors use a 2D map of side-chain interactions and ‘kinks’ in the structure caused by glycine residues to direct 3D model building. This results in the successful design, characterisation and crystallisation of the first *de novo* water-soluble β -barrel protein. Furthermore, β barrels that bind small molecules are targeted to incorporate an environment-sensitive fluorophore that only fluoresces when held in a specific conformation by the *de novo* scaffold (Figure 4). The fluorophore is introduced early in the design strategy, rather than by embellishing a non-functional variant. This study is impressive for two reasons: firstly, accessing soluble β -rich proteins has proven challenging in protein design; secondly, the *de novo* proteins activate fluorescence of the small molecules *in vivo*. That said, before library screening is used to improve the designs, the low μ M affinity of the small molecule is similar to previously reported binding constants to α -helical barrels [31[•]].

Challenges ahead

The robust and routine design of functional *de novo* proteins remains an unsolved problem. For instance, to our knowledge, there are no examples of tight binding of small, polar molecules by *de novo* proteins. The change in approach in the last few years to incorporate the functional aspect of the design at an early stage shows clear potential, which we envisage will become more evident as design algorithms improve. However, as stated above, accessing functional *de novo* proteins that work on a par with natural proteins will likely require the incorporation of conformational changes and dynamics into the design process [90]. Such design targets will need improved abilities to build and

Figure 4



Computationally designed functional proteins. **(a)** The first high-resolution structure of a porphyrin-binding *de novo* four-helix bundle [30^{••}]. The interaction between the zinc atom in the unnatural porphyrin ring ($C_{24}H_8F_{12}N_4Zn$) and the histidine side chain is shown (PDB: 5TGY). **(b)** An example of a fully *de novo* water soluble β -barrel (PDB: 6CZH) [32^{••}]. The environmental sensitive ligand, DFHBI ($C_{12}H_{10}F_2N_2O_2$) bound to the cavity (green spheres).

score *in silico* models that access multiple states. Thankfully, methods for multistate design are being developed [91–94]. For instance, Grigoryan *et al.* use such an approach to design leucine zippers that selectively bind a single partner from 20 members of the bZIP family by modelling potential off-target interactions as part of the design process [95]; Löffler *et al.* engineer a $(\beta\alpha)_8$ -barrel into a retro-aldolase with measurable, albeit low, catalytic efficiency [94]; and Feng *et al.* use conformation ensembles to engineer ligand-binding G-protein-coupled receptors [96*].

Natural allosteric proteins can be engineered to bind different small molecules [76,97]. Similarly, existing allosteric systems can be used to control new functions [98,99]. Switchable *de novo* coiled-coil systems, both reversible and irreversible, can be controlled through temperature [100], pH [101,102] and metal binding [103–105]. However, fully *de novo* allosteric proteins that respond to small-molecule inducers have yet to be reported.

Arguably more progress is being made utilising dynamic multistate *de novo* design. Davey *et al.* recently give an example of an, albeit engineered, dynamic protein that accesses two conformations that exchange on a millisecond timescale [106*]. Focusing on *de novo* proteins, Rhys *et al.* present a *de novo* α -helical barrel that is hexameric in solution but crystallises as an octameric assembly [6*]. Joh *et al.* present a *de novo* zinc-ion transporter by designing a membrane-spanning four-helix bundle with two distinct coordination sites each of which destabilises the other upon metal binding [14]. However, the challenge of incorporating dynamics to improve catalysis or small-molecule binding has yet to be met.

Overall, despite considerable and encouraging advances in *de novo* protein design there are many challenges ahead for the *de novo* design of functional proteins. These are being actively targeted by the field as a whole. If advances continue at the current rate of delivery of *de novo* protein scaffolds, then protein design will indeed have come of age.

Conflict of interest statement

Nothing declared.

Acknowledgements

WMD, GGR and DNW were supported by a European Research Council Advanced Grant (340764). DNW is supported by the EPSRC and BBSRC through the BrisSynBio Synthetic Biology Research Centre (BB/L01386X1). DNW holds a Royal Society Wolfson Research Merit Award (WM140008). We thank Ross Anderson for providing a computational model of the C45 protein.

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. Huang P-S, Boyken SE, Baker D: **The coming of age of *de novo* protein design.** *Nature* 2016, **537**:320-327.
 2. Hill CP, Anderson DH, Wesson L, Degrado WF, Eisenberg D: **Crystal-structure of α_1 – implications for protein design.** *Science* 1990, **249**:543-546.
 3. Nautiyal S, Alber T: **Crystal structure of a designed, thermostable, heterotrimeric coiled coil.** *Protein Sci* 1999, **8**:84-90.
 4. 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.
 5. Thomson AR, Wood CW, Burton AJ, Bartlett GJ, Sessions RB, Brady RL, Woolfson DN: **Computational design of water-soluble α -helical barrels.** *Science* 2014, **346**:485-488.
 6. Rhys GG, Wood CW, Lang EJM, Mulholland AJ, Brady RL, Thomson AR, Woolfson DN: **Maintaining and breaking symmetry in homomeric coiled-coil assemblies.** *Nat Commun* 2018, **9**:4132.
- This paper describes an empirical investigation of coiled coils with greater than five helices. A stable, multi-state assembly is presented, and also the first structure of an octameric coiled coil.
7. Zhang S-Q, Huang H, Yang J, Kratochvil HT, Lolicato M, Liu Y, Shu X, Liu L, DeGrado WF: **Designed peptides that assemble into cross- α amyloid-like structures.** *Nat Chem Biol* 2018, **14**:870-875.
 8. Marcos E, Chidyausiku TM, McShan AC, Evangelidis T, Nerli S, Carter L, Nivón LG, Davis A, Oberdorfer G, Tripsianes K *et al.*: **De novo design of a non-local β -sheet protein with high stability and accuracy.** *Nat Struct Mol Biol* 2018, **25**:1028-1034.
 9. Liang H, Chen H, Fan K, Wei P, Guo X, Jin C, Zeng C, Tang C, Lai L: **De novo design of a $\beta\alpha\beta$ motif.** *Angew Chem Int Ed* 2009, **48**:3301-3303.
 10. Huang PS, Feldmeier K, Parmeggiani F, Velasco DAF, Hocker B, Baker D: **De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy.** *Nat Chem Biol* 2016, **12**:29-34.
- The first example of a structurally validated computationally designed TIM barrel. Although not functional, this fold is the most common topology for natural enzymes and shows the potential for the scaffold to be functionally embellished in the future.
11. Marcos E, Basanta B, Chidyausiku TM, Tang Y, Oberdorfer G, Liu G, Swapna GVT, Guan R, Silva D-A, Dou J *et al.*: **Principles for designing proteins with curved β sheets.** *Science* 2017, **355**:201-206.
 12. 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.
 13. Lu P, Min D, DiMaio F, Wei KY, Vahey MD, Boyken SE, Chen Z, Fallas JA, Ueda G, Sheffler W *et al.*: **Accurate computational design of multipass transmembrane proteins.** *Science* 2018, **359**:1042-1046.
 14. Joh NH, Wang T, Bhate MP, Acharya R, Wu Y, Grabe M, Hong M, Grigoryan G, DeGrado WF: **De novo design of a transmembrane Zn^{2+} -transporting four-helix bundle.** *Science* 2014, **346**:1520-1524.
 15. Burton AJ, Thomson AR, Dawson WM, Brady RL, Woolfson DN: **Installing hydrolytic activity into a completely *de novo* protein framework.** *Nat Chem* 2016, **8**:837-844.
- The first example of an active catalytic triad being introduced into a fully *de novo* framework/scaffold. A Cys-His-Glu triad is incorporated into a heptameric coiled coil establishing hydrolase activity in this framework.
16. Watkins DW, Jenkins JMX, Grayson KJ, Wood N, Steventon JW, Le Vay KK, Goodwin MI, Mullen AS, Bailey HJ, Crump MP *et al.*: **Construction and in vivo assembly of a catalytically proficient and hyperthermostable *de novo* enzyme.** *Nat Commun* 2017, **8**:358.
- An artificial oxidoreductase obtained through minimal design. This catalytic maquette shows rates of catalysis on a par with natural enzymes.
17. Stank A, Kokh DB, Fuller JC, Wade RC: **Protein binding pocket dynamics.** *Acc Chem Res* 2016, **49**:809-815.

18. Kay LE: **New views of functionally dynamic proteins by solution NMR spectroscopy.** *J Mol Biol* 2016, **428**:323-331.
19. Palmer AG III: **Enzyme dynamics from NMR spectroscopy.** *Acc Chem Res* 2015, **48**:457-465.
20. Jiang L, Althoff EA, Clemente FR, Doyle L, Röthlisberger D, Zanghellini A, Gallaher JL, Betker JL, Tanaka F, Barbas CF *et al.*: **De novo computational design of retro-aldol enzymes.** *Science* 2008, **319**:1387-1391.
21. Siegel JB, Zanghellini A, Lovick HM, Kiss G, Lambert AR, St. Clair JL, Gallaher JL, Hilvert D, Gelb MH, Stoddard BL *et al.*: **Computational design of an enzyme catalyst for a stereoselective bimolecular diels-alder reaction.** *Science* 2010, **329**:309-313.
22. Tinberg CE, Khare SD, Dou J, Doyle L, Nelson JW, Schena A, Jankowski W, Kalodimos CG, Johnsson K, Stoddard BL *et al.*: **Computational design of ligand-binding proteins with high affinity and selectivity.** *Nature* 2013, **501**:212-216.
23. Giger L, Caner S, Obexer R, Kast P, Baker D, Ban N, Hilvert D: **Evolution of a designed retro-aldolase leads to complete active site remodeling.** *Nat Chem Biol* 2013, **9**:494-498.
24. Peacock AFA: **Incorporating metals into de novo proteins.** *Curr Opin Chem Biol* 2013, **17**:934-939.
25. Tebo AG, Pecoraro VL: **Artificial metalloenzymes derived from three-helix bundles.** *Curr Opin Chem Biol* 2015, **25**:65-70.
26. Schreiber G, Fleishman SJ: **Computational design of protein-protein interactions.** *Curr Opin Struct Biol* 2013, **23**:903-910.
27. DeGrado WF, Summa CM, Pavone V, Nastro F, Lombardi A: **De novo design and structural characterization of proteins and metalloproteins.** *Annu Rev Biochem* 1999, **68**:779-819.
28. Hecht MH, Das A, Go A, Bradley LH, Wei Y: **De novo proteins from designed combinatorial libraries.** *Protein Sci* 2004, **13**:1711-1723.
29. Woolfson DN, Bartlett GJ, Bruning M, Thomson AR: **New currency for old rope: From coiled-coil assemblies to α -helical barrels.** *Curr Opin Struct Biol* 2012, **22**:432-441.
30. Polizzi NF, Wu Y, Lemmin T, Maxwell AM, Zhang SQ, Rawson J, Beratan DN, Therien MJ, DeGrado WF: **De novo design of a hyperstable non-natural protein-ligand complex with sub-Å accuracy.** *Nat Chem* 2017, **9**:1157-1164.
- Computational design is used to design stability and binding for an unnatural porphyrin into a four-helix bundle simultaneously. This is the first example of a high-resolution structure for this class of *de novo* protein with its ligand bound.
31. Thomas F, Dawson WM, Lang EJM, Burton AJ, Bartlett GJ, Rhys GG, Mulholland AJ, Woolfson DN: **De novo-designed α -helical barrels as receptors for small molecules.** *ACS Synth Biol* 2018, **7**:1808-1816.
- A systematic study of ligand binding to the channels of *de novo* pentameric, hexameric and heptameric α -helical barrels. Ionisable amino acids are introduced to the lumens of the barrels to enable binding of positively and negatively charged molecules.
32. 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 beta-barrel.** *Nature* 2018, **561**:485-491.
- De novo* designed water-soluble β barrels are described for the first time. Variants that bind an environmentally sensitive fluorophore are characterised both *in vitro* and *in vivo*.
33. Moffet DA, Hecht MH: **De novo proteins from combinatorial libraries.** *Chem Rev* 2001, **101**:3191-3203.
34. Discher BM, Koder RL, Moser CC, Dutton PL: **Hydrophilic to amphiphilic design in redox protein maquettes.** *Curr Opin Chem Biol* 2003, **7**:741-748.
35. Grayson KJ, Anderson JLR: **The ascent of man(made oxidoreductases).** *Curr Opin Struct Biol* 2018, **51**:149-155.
36. Kodali G, Mancini JA, Solomon LA, Episova TV, Roach N, Hobbs CJ, Wagner P, Mass OA, Aravindu K, Barnsley JE *et al.*: **Design and engineering of water-soluble light-harvesting protein maquettes.** *Chem Sci* 2017, **8**:316-324.
37. Sheehan MM, Magaraci MS, Kuznetsov IA, Mancini JA, Kodali G, Moser CC, Dutton PL, Chow BY: **Rational construction of compact de novo-designed biliverdin-binding proteins.** *Biochemistry* 2018, **57**:6752-6756.
38. Farid TA, Kodali G, Solomon LA, Lichtenstein BR, Sheehan MM, Fry BA, Bialas C, Ennist NM, Siedlecki JA, Zhao Z *et al.*: **Elementary tetrahelical protein design for diverse oxidoreductase functions.** *Nat Chem Biol* 2013, **9**:826-833.
39. Watkins DW, Armstrong CT, Beesley JL, Marsh JE, Jenkins JMX, Sessions RB, Mann S, Ross Anderson JL: **A suite of de novo c-type cytochromes for functional oxidoreductase engineering.** *Biochim Biophys Acta* 2016, **1857**:493-502.
40. Donnelly AE, Murphy GS, Diganantonio KM, Hecht MH: **A de novo enzyme catalyzes a life-sustaining reaction in Escherichia coli.** *Nat Chem Biol* 2018, **14**:253-255.
- A minimal protein design is embellished using combinatorial libraries to produce a four-helix bundle capable of mimicking the function of an essential enzyme.
41. Fisher MA, McKinley KL, Bradley LH, Viola SR, Hecht MH: **De novo designed proteins from a library of artificial sequences function in Escherichia coli and enable cell growth.** *PLoS One* 2011, **6**:e15364.
42. Lear JD, Wasserman ZR, DeGrado WF: **Synthetic amphiphilic peptide models for protein ion channels.** *Science* 1988, **240**:1177-1181.
43. Lalaurie CJ, Dufour V, Meletiou A, Ratcliffe S, Harland A, Wilson O, Vamasiri C, Shoemark DK, Williams C, Arthur CJ *et al.*: **The de novo design of a biocompatible and functional integral membrane protein using minimal sequence complexity.** *Sci Rep* 2018, **8**:14564.
44. Skalicky JJ, Gibney BR, Rabanal F, Bieber Urbauer RJ, Dutton PL, Wand AJ: **Solution structure of a designed four- α -helix bundle maquette scaffold.** *J Am Chem Soc* 1999, **121**:4941-4951.
45. Huang SS, Gibney BR, Stayrook SE, Leslie Dutton P, Lewis M: **X-ray structure of a maquette scaffold.** *J Mol Biol* 2003, **326**:1219-1225.
46. Wei Y, Kim S, Fela D, Baum J, Hecht MH: **Solution structure of a de novo protein from a designed combinatorial library.** *Proc Natl Acad Sci U S A* 2003, **100**:13270.
47. Go A, Kim S, Baum J, Hecht MH: **Structure and dynamics of de novo proteins from a designed superfamily of 4-helix bundles.** *Protein Sci* 2008, **17**:821-832.
48. Arai R, Kobayashi N, Kimura A, Sato T, Matsuo K, Wang AF, Platt JM, Bradley LH, Hecht MH: **Domain-swapped dimeric structure of a stable and functional de novo four-helix bundle protein, WA20.** *J Phys Chem B* 2012, **116**:6789-6797.
49. Murphy GS, Sathyamoorthy B, Der BS, Machius MC, Pulavarti SV, Szyperski T, Kuhlman B: **Computational de novo design of a four-helix bundle protein—DND_4HB.** *Protein Sci* 2015, **24**:434-445.
50. Grigoryan G, DeGrado WF: **Probing designability via a generalized model of helical bundle geometry.** *J Mol Biol* 2011, **405**:1079-1100.
51. Harbury P, Zhang T, Kim P, Alber T: **A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants.** *Science* 1993, **262**:1401-1407.
52. Fletcher JM, Boyle AL, Bruning M, Bartlett GJ, Vincent TL, Zaccai NR, Armstrong CT, Bromley EHC, Booth PJ, Brady RL *et al.*: **A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology.** *ACS Synth Biol* 2012, **1**:240-250.
53. Thomas F, Boyle AL, Burton AJ, Woolfson DN: **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* 2013, **135**:5161-5166.
54. Burgess NC, Sharp TH, Thomas F, Wood CW, Thomson AR, Zaccai NR, Brady RL, Serpell LC, Woolfson DN: **Modular design of self-assembling peptide-based nanotubes.** *J Am Chem Soc* 2015, **137**:10554-10562.

55. 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.
56. Hussey BJ, McMillen DR: **Programmable T7-based synthetic transcription factors.** *Nucleic Acids Res* 2018, **46**:9842-9854.
57. Lee MJ, Mantell J, Hodgson L, Alibhai D, Fletcher JM, Brown IR, Frank S, Xue WF, Verkade P, Woolfson DN *et al.*: **Engineered synthetic scaffolds for organizing proteins within the bacterial cytoplasm.** *Nat Chem Biol* 2018, **14**:142-147.
58. Crick FHC: **The Fourier transform of a coiled-coil.** *Acta Crystallogr* 1953, **6**:685-689.
59. Harbury PB, Plecs JJ, Tidor B, Alber T, Kim PS: **High-resolution protein design with backbone freedom.** *Science* 1998, **282**:1462-1467.
60. Offer G, Sessions R: **Computer modeling of the α -helical coiled-coil – packing of side-chains in the inner-core.** *J Mol Biol* 1995, **249**:967-987.
61. Wood CW, Bruning M, Ibarra AA, Bartlett GJ, Thomson AR, Sessions RB, Brady RL, Woolfson DN: **CCBuilder: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies.** *Bioinformatics* 2014, **30**:3029-3035.
62. Wood CW, Woolfson DN: **CCBuilder 2.0: powerful and accessible coiled-coil modeling.** *Protein Sci* 2018, **27**:103-111.
63. Wood CW, Heal JW, Thomson AR, Bartlett GJ, Ibarra AA, Brady RL, Sessions RB, Woolfson DN: **ISAMBARD: an open-source computational environment for biomolecular analysis, modelling and design.** *Bioinformatics* 2017, **33**:3043-3050.
64. Szczepaniak K, Ludwiczak J, Winski A, Dunin-Horkawicz S: **Variability of the core geometry in parallel coiled-coil bundles.** *J Struct Biol* 2018, **204**:117-124.
65. Guzenko D, Strelkov SV: **Optimal data-driven parameterization of coiled coils.** *J Struct Biol* 2018, **204**:125-129.
66. Burton AJ, Thomas F, Agnew C, Hudson KL, Halford SE, Brady RL, Woolfson DN: **Accessibility, reactivity, and selectivity of side chains within a channel of de novo peptide assembly.** *J Am Chem Soc* 2013, **135**:12524-12527.
67. Huang P-S, Oberdorfer G, Xu C, Pei XY, Nannenga BL, Rogers JM, DiMaio F, Gonen T, Luisi B, Baker D: **High thermodynamic stability of parametrically designed helical bundles.** *Science* 2014, **346**:481-485.
68. Studer S, Hansen DA, Pianowski ZL, Mittl PRE, Debon A, Guffy SL, Der BS, Kuhlman B, Hilvert D: **Evolution of a highly active and enantiospecific metalloenzyme from short peptides.** *Science* 2018, **362**:1285-1288.
69. Boas FE, Harbury PB: **Potential energy functions for protein design.** *Curr Opin Struct Biol* 2007, **17**:199-204.
70. Gainza-Ciraoui P, Correia BE: **Computational protein design—the next generation tool to expand synthetic biology applications.** *Curr Opin Biotechnol* 2018, **52**:145-152.
71. Mackenzie CO, Grigoryan G: **Protein structural motifs in prediction and design.** *Curr Opin Struct Biol* 2017, **44**:161-167.
72. Li Z, Yang Y, Zhan J, Dai L, Zhou Y: **Energy functions in de novo protein design: current challenges and future prospects.** *Annu Rev Biophys* 2013, **42**:315-335.
73. Leaver-Fay A, Tyka M, Lewis SM, Lange OF, Thompson J, Jacak R, Kaufman KW, Renfren PD, Smith CA, Sheffler W *et al.*: **Rosetta3: an object-oriented software suite for the simulation and design of macromolecules.** *Methods Enzymol* 2011, **487**:545-574.
74. Dou J, Doyle L, Greisen P Jr, Schena A, Park H, Johnsson K, Stoddard BL, Baker D: **Sampling and energy evaluation challenges in ligand binding protein design.** *Protein Sci* 2017, **26**:2426-2437.
75. Jester BW, Tinberg CE, Rich MS, Baker D, Fields S: **Engineered biosensors from dimeric ligand-binding domains.** *ACS Synth Biol* 2018, **7**:2457-2467.
76. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing MA, Cascio D, Rogers JK, Isaacs FJ, Kosuri S, Baker D *et al.*: **Engineering an allosteric transcription factor to respond to new ligands.** *Nat Methods* 2016, **13**:177-183.
77. Bick MJ, Greisen PJ, Morey KJ, Antunes MS, La D, Sankaran B, Reymond L, Johnsson K, Medford JI, Baker D: **Computational design of environmental sensors for the potent opioid fentanyl.** *eLife* 2017, **6**:e28909.
78. Banda-Vázquez J, Shanmugaratnam S, Rodríguez-Sotres R, Torres-Larios A, Höcker B, Sosa-Peinado A: **Redesign of LAOBP to bind novel L-amino acid ligands.** *Protein Sci* 2018, **27**:957-968.
79. Garrabou X, Wicky BIM, Hilvert D: **Fast Knoevenagel condensations catalyzed by an artificial Schiff-base-forming enzyme.** *J Am Chem Soc* 2016, **138**:6972-6974.
80. Höcker B: **Design of proteins from smaller fragments—learning from evolution.** *Curr Opin Struct Biol* 2014, **27**:56-62.
81. Eisenbeis S, Proffitt W, Coles M, Truffault V, Shanmugaratnam S, Meiler J, Höcker B: **Potential of fragment recombination for rational design of proteins.** *J Am Chem Soc* 2012, **134**:4019-4022.
82. Jacobs TM, Williams B, Williams T, Xu X, Eletsky A, Federizon JF, Szyperski T, Kuhlman B: **Design of structurally distinct proteins using strategies inspired by evolution.** *Science* 2016, **352**:687-690.
83. Lapidoth G, Khersonsky O, Lipsh R, Dym O, Albeck S, Rogotner S, Fleishman SJ: **Highly active enzymes by automated combinatorial backbone assembly and sequence design.** *Nat Commun* 2018, **9**:2780.
- Computational design is used to combine fragments from enzyme families and install xylanase and lactonase activity. Activities matching the natural proteins are obtained without the need to evolve the sequences further.
84. Rocklin GJ, Chidyausiku TM, Goresnik I, Ford A, Houlston S, Lemak A, Carter L, Ravichandran R, Mulligan VK, Chevalier A *et al.*: **Global analysis of protein folding using massively parallel design, synthesis, and testing.** *Science* 2017, **357**:168-175.
85. Baker EG, Bartlett GJ, Crump MP, Sessions RB, Linden N, Faul CFJ, Woolfson DN: **Local and macroscopic electrostatic interactions in single α -helices.** *Nat Chem Biol* 2015, **11**:221-228.
86. Tyukhtenko S, Rajarshi G, Karageorgos I, Zvonok N, Gallagher ES, Huang H, Vemuri K, Hudgens JW, Ma X, Nasr ML *et al.*: **Effects of distal mutations on the structure, dynamics and catalysis of human monoacylglycerol lipase.** *Sci Rep* 2018, **8**:1719.
87. Negron C, Fufezan C, Koder RL: **Geometric constraints for porphyrin binding in helical protein binding sites.** *Proteins Struct Funct Bioinform* 2009, **74**:400-416.
88. Reedy CJ, Gibney BR: **Heme protein assemblies.** *Chem Rev* 2004, **104**:617-650.
89. Fry HC, Lehmann A, Sinks LE, Asselberghs I, Tronin A, Krishnan V, Blasie JK, Clays K, DeGrado WF, Saven JG *et al.*: **Computational de novo design and characterization of a protein that selectively binds a highly hyperpolarizable abiological chromophore.** *J Am Chem Soc* 2013, **135**:13914-13926.
90. Osuna S, Jiménez-Osés G, Noey EL, Houk KN: **Molecular dynamics explorations of active site structure in designed and evolved enzymes.** *Acc Chem Res* 2015, **48**:1080-1089.
91. Leaver-Fay A, Jacak R, Stranges PB, Kuhlman B: **A generic program for multistate protein design.** *PLoS One* 2011, **6**:e20937.
92. Negron C, Keating AE: **Multistate protein design using CLEVER and CLASSY.** *Methods Enzymol* 2013, **523**:171-190.
93. Davey JA, Chica RA: **Improving the accuracy of protein stability predictions with multistate design using a variety of backbone ensembles.** *Proteins Struct Funct Bioinform* 2014, **82**:771-784.
94. Löffler P, Schmitz S, Hupfeld E, Sterner R, Merkl R: **Rosetta:MSF: a modular framework for multi-state computational protein design.** *PLoS Comput Biol* 2017, **13**:e1005600.
95. Grigoryan G, Reinke AW, Keating AE: **Design of protein-interaction specificity gives selective bZIP-binding peptides.** *Nature* 2009, **458**:859-864.

96. Feng X, Ambia J, Chen K-YM, Young M, Barth P: **Computational design of ligand-binding membrane receptors with high selectivity.** *Nat Chem Biol* 2017, **13**:715-723.

A combination of protein design, homology modelling and ligand docking is used to predict GPCR-ligand conformations and to engineer new ligand-binding specificity into a receptor of unknown structure.

97. Schmidt K, Gardill BR, Kern A, Kirchweger P, Borsch M, Muller YA: **Design of an allosterically modulated doxycycline and doxorubicin drug-binding protein.** *Proc Natl Acad Sci U S A* 2018, **115**:5744-5749.
98. Korendovych IV, Kulp DW, Wu Y, Cheng H, Roder H, DeGrado WF: **Design of a switchable eliminase.** *Proc Natl Acad Sci U S A* 2011, **108**:6823-6827.
99. Raymond EA, Mack KL, Yoon JH, Moroz OV, Moroz YS, Korendovych IV: **Design of an allosterically regulated retroaldolase.** *Protein Sci* 2015, **24**:561-570.
100. Ciani B, Hutchinson EG, Sessions RB, Woolfson DN: **A designed system for assessing how sequence affects α to β conformational transitions in proteins.** *J Biol Chem* 2002, **277**:10150-10155.
101. Lizatović R, Aurelius O, Stenström O, Drakenberg T, Akke M, Logan Derek T, André I: **A *de novo* designed coiled-coil peptide with a reversible pH-induced oligomerization switch.** *Structure* 2016, **24**:946-955.
102. Zhang Y, Bartz R, Grigoryan G, Bryant M, Aaronson J, Beck S, Innocent N, Klein L, Procopio W, Tucker T *et al.*: **Computational design and experimental characterization of peptides intended for pH-dependent membrane insertion and pore formation.** *ACS Chem Biol* 2015, **10**:1082-1093.
103. Cerasoli E, Sharpe BK, Woolfson DN: **ZiCo: a peptide designed to switch folded state upon binding zinc.** *J Am Chem Soc* 2005, **127**:15008-15009.
104. Ambroggio XI, Kuhlman B: **Computational design of a single amino acid sequence that can switch between two distinct protein folds.** *J Am Chem Soc* 2006, **128**:1154-1161.
105. Aupič J, Lapenta F, Jerala R: **SwitCCh: metal-site design for controlling the assembly of a coiled-coil homodimer.** *ChemBioChem* 2018, **19**:2453-2457.
106. Davey JA, Damry AM, Goto NK, Chica RA: **Rational design of proteins that exchange on functional timescales.** *Nat Chem Biol* 2017, **13**:1280-1285.

A natural protein is engineered to switch between two distinct conformations on a millisecond timescale. This is the first example of dynamics being incorporated by rational protein redesign.