Proteomimetics as protein-inspired scaffolds with defined tertiary folding patterns

Horne, W. Seth; Grossmann, Tom N.

published in
Nature Chemistry
2020

DOI (link to publisher)
10.1038/s41557-020-0420-9

document version
Publisher's PDF, also known as Version of record

nature chemistry

PERSPECTIVE

https://doi.org/10.1038/s41557-020-0420-9

Check for updates

Proteomimetics as protein-inspired scaffolds with defined tertiary folding patterns

W. Seth Horne[1]✉ and Tom N. Grossmann[2]✉

Proteins have evolved as a variable platform that provides access to molecules with diverse shapes, sizes and functions. These features have inspired chemists for decades to seek artificial mimetics of proteins with improved or novel properties. Such work has focused primarily on small protein fragments, often isolated secondary structures; however, there has lately been a growing interest in the design of artificial molecules that mimic larger, more complex tertiary folds. In this Perspective, we define these agents as 'proteomimetics' and discuss the recent advances in the field. Proteomimetics can be divided into three categories: protein domains with side-chain functionality that alters the native linear-chain topology; protein domains in which the chemical composition of the polypeptide backbone has been partially altered; and protein-like folded architectures that are composed entirely of non-natural monomer units. We give an overview of these proteomimetic approaches and outline remaining challenges facing the field.

Proteins are defined by a hierarchy of structural complexity that is founded on their primary amino acid sequence. Local and long-range intramolecular interactions among backbone and side-chain atoms determine higher-order folding behaviour by facilitating the formation of secondary structures and their arrangement into tertiary structure (Fig. 1a)[1]. Protein folds have evolved to manifest varied and unique functional characteristics that are central to biological systems. The mimicry of protein structure and function represents a long-standing challenge to chemists and has been approached with artificial agents of varying size and complexity (Fig. 1b). At one end of this hierarchy are molecules intended to recreate features of a few amino acids or isolated secondary structure — generally referred to as peptidomimetics[2,3]. Early examples of peptidomimetics include mimics of primary structures found in small peptide hormones and peptide substrates of protease enzymes[2]. The latter were developed as inhibitors of protease function, and clinical use of these compounds revolutionized the treatment of diseases such as HIV and hepatitis C[4]. Later, the stabilization or recreation of secondary structure in the absence of a protein context became a central focus of endeavours towards peptidomimetics. The challenges inherent to the mimicry of these more complex folded structures led to the invention of a number of distinct approaches, including peptides containing artificial building blocks, macrocyclic agents and entirely artificial backbones with defined folding propensities (Fig. 1b)[3,5,6]. Over the past decade, peptidomimetics have become valuable bioactive agents and drug candidates of particular utility in targeting the extended and flat biomacromolecular surfaces involved in protein–protein interactions[7].

Building on the above precedent on peptidomimetics, there has been a recent growing interest in the design of artificial agents that mimic larger and more complex tertiary folds. While a daunting challenge, efforts along these lines have been enabled by advances in methodology for protein chemical synthesis[8] and bioconjugation[9] as well as the emergence of powerful protein-engineering approaches, including directed evolution[10], computational design[11] and non-canonical amino acid mutagenesis[12]. A critical mass of results now suggests a burgeoning field of proteomimetics that are conceptually linked to, yet distinct from, their smaller peptidomimetic counterparts. The term proteomimetic has found limited previous use in the literature, primarily to describe molecules that recreate isolated segments of protein secondary structure[5,13,14]. Herein, we propose a definition of proteomimetics as molecules that have a defined tertiary fold and recreate certain features of a protein, such as shape, recognition properties or enzymatic activity. In this Perspective, we highlight recent advances towards proteomimetic structures based on a classification of these agents into one of three categories (Fig. 1b). Most examples of proteomimetics reported to date were obtained by chemical protein engineering, an approach that can provide protein-like molecules with altered chain topology and/or partially artificial backbone composition introduced in a chain composed mainly of natural amino acids. Other work has addressed the challenge of proteomimetic design by applying non-natural building blocks to construct completely artificial backbones able to adopt large, complex folds.

Altered chain topology

In natural proteins, the linear topology of the main chain is often altered by disulfide bridges formed between two cysteines spatially aligned in the folded state but distant in primary sequence. Such intramolecular cross-links are frequently located between distinct secondary structures, thereby stabilizing the domain tertiary fold. While the installation of non-native disulfide bridges has been widely applied for the stabilization of tertiary structures, the short distance of the resulting cross-link (Cβ–Cβ distance $\leq$5.5 Å)[15] considerably restricts the number of possible modification sites. Seeking to overcome this limitation, work dating back to the 1960s has shown the use of biselectrophilic reagents to cross-link the amines of two lysine side chains, giving access to structurally diverse bridges[16–18]. This approach was used to probe the spatial proximity of lysine residues[16] and to investigate the effect of intramolecular cross-links on protein folding and tertiary structure stability[19]. Given the high abundance of surface lysine residues in proteins and the resulting selectivity issues, these macrocyclization reactions were limited to a few protein model systems such as ribonuclease A[16,17,19].

[1]Department of Chemistry, University of Pittsburgh, Pittsburgh, PA, USA. [2]Department of Chemistry and Pharmaceutical Sciences, VU University Amsterdam, Amsterdam, the Netherlands. ✉e-mail: horne@pitt.edu; t.n.grossmann@vu.nl

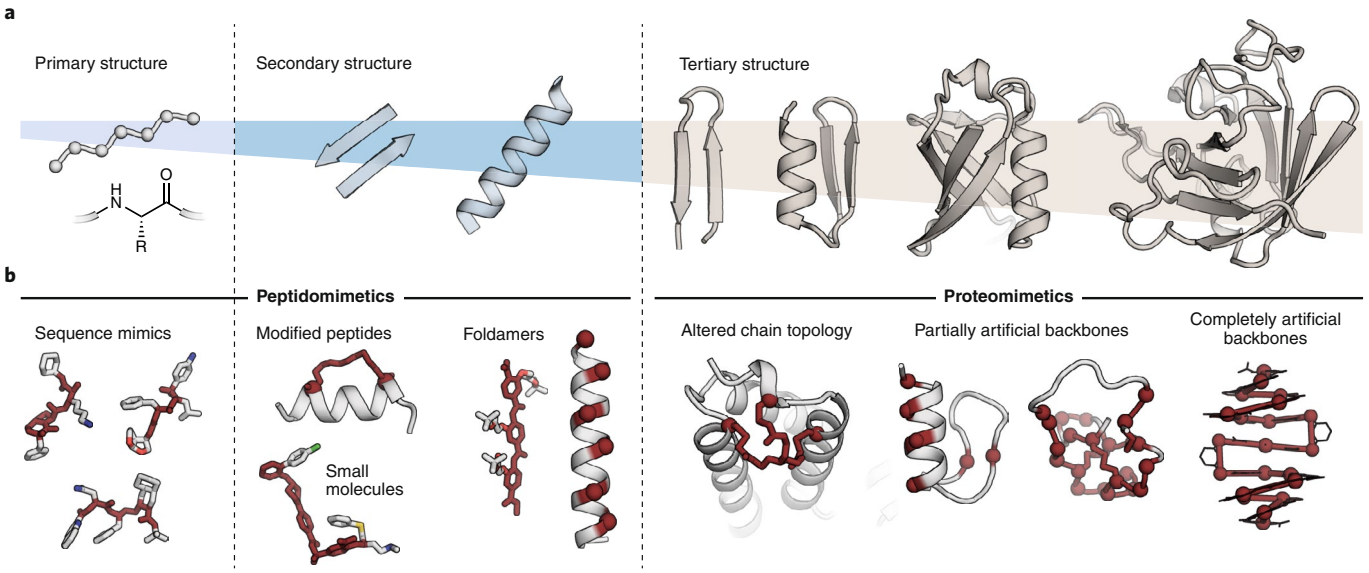

Fig. 1 | Overview of protein structures and their mimetics. a, Hierarchy of peptide and protein structure, which spans in complexity from primary sequence (light blue) to folded secondary structure (blue) and tertiary structure (tan). **b**, Corresponding hierarchy of peptide and protein mimetics. Peptidomimetics encompass agents that recreate primary sequence or isolated secondary structure and can be broadly classified into small molecules, peptides containing chemical modification (for example, macrocyclization shown) or artificial backbones with defined folds ('foldamers'). Proteomimetics, the focus of this Perspective, comprise agents that recreate protein tertiary structure and can be classified into molecules with altered chain topology, partially artificial backbone compositions or completely artificial backbone compositions. Example peptidomimetic and proteomimetic structures are shown with non-natural parts highlighted in red. (Coordinates for structures from: Protein Data Bank (PDB) entries 1A46, 1IJA, 1T3R, 1UBQ, 2AGH, 2QMT, 2YJ1, 2YXJ, 3OXC, 4N5T, 5KVN; Cambridge Structural Database (CSD) entry 697088; and refs. [67,85]).

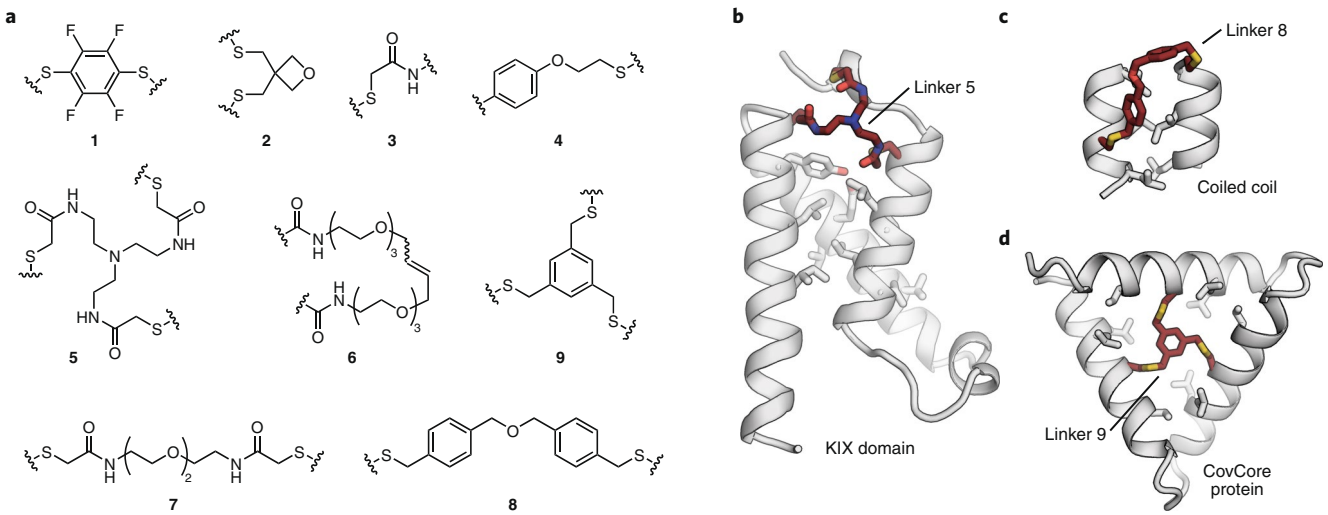

Fig. 2 | Examples of proteomimetics based on altered chain topology. a, Overview of bridging moieties including linker structures resulting from reactions of biselectrophilic cross-linkers or non-natural amino acids with cysteine (**1**–**4**, **7**, **8**), symmetric trivalent bridges used to cross-link three cysteine residues (**5**, **9**) and a linker resulting from a ring-closing metathesis between two modified asparagine residues (**6**). **b**, Model of the INCYPRO-stabilized KIX domain based on an NMR structure of the unmodified protein (PDB 2AGH)[25]. **c**, NMR structure of a chemically dimerized parallel coiled coil.[32] **d**, NMR structure (PDB 5V2G) of the de novo-designed tricyclic three-helix CovCore structure[37]. Proteins are depicted in cartoon representation with hydrophobic core residues shown as sticks (grey). Cross-links are highlighted (red) and shown as sticks.

Aiming for more general strategies to modify protein chain topology, disulfide mimetics have been obtained by cross-linking two reduced cysteine side chains with biselectrophilic agents. Examples include the installation of tetrafluorobenzene (**1**, Fig. 2a) for the cross-linking of H2 relaxin[20] and the application of oxetane-based scaffold **2** for the modification of antibody fragments and a diphtheria toxin-derived carrier protein[21]. These biselectrophilic agents connect the two cysteine sulfur atoms via a three- or four-carbon bridge. Depending of the location of the original disulfide, the introduction of the organic linker can result in altered protein function[20]. Cysteine-based protein cross-linking approaches have also been developed to introduce novel intradomain cross-links. For example, a chloroacetamide-modified lysine has been incorporated into chemically synthesized affibodies and albumin-binding

domains to capture an appropriately aligned cysteine side chain providing thioether linkage **3**[22,23]. Broadening the application of this strategy to larger proteins, non-canonical mutagenesis was used to introduce a tyrosine-derived electrophilic amino acid into a myoglobin-based cyclopropanation biocatalyst[24]. The resulting thioether cross-link **4** stabilized the enzyme towards thermal stress and high concentrations of organic cosolvent, enabling the cyclopropanation of water-insoluble substrates[24]. To avoid complications in heterologous expression of proteins that result from the use of non-natural amino acids, two or three solvent-exposed cysteines have been introduced and reacted with bis- or triselectrophilic agents resulting in the in situ cyclization of proteins (INCYPRO). The most pronounced stabilization effects from the INCYPRO strategy were obtained with a triselectrophilic cross-link (**5**) that resulted in bicyclic versions of the transcriptional coactivator domain KIX (Fig. 2b) and of the enzyme sortase A[25]. The stabilized sortase A variant was able to catalyse reactions under denaturing conditions. Notably, although INCYPRO requires the absence of additional solvent-accessible cysteines, the chemistry tolerates the presence of an active-site cysteine in sortase A[25].

The selective chemical modification of entire protein domains as described above is limited by the inherent selectivity issues when addressing natural amino acids. This renders the total chemical synthesis an appealing alternative strategy, in particular when aiming for smaller peptide-derived proteomimetics. Chemical synthesis allows the introduction of a broad range of non-natural building blocks and topologies, and the application of methodologies originally developed in the context of classic peptidomimetics. In the construction of such synthetic proteomimetics, secondary structures such as β-hairpins and α-helices have been utilized as assembly units to obtain higher-order folding patterns. A β-hairpin comprises two antiparallel β-strands connected by a short loop structure. Because hairpins typically exhibit a low folding propensity in isolation, modifications are required to stabilize the fold. Most approaches to this end involve macrocyclization either via side-chain-to-side-chain bridges or by head-to-tail connections using turn-inducing loop structures[3,26]. Recently, a β-sheet motif within a WW domain was stabilized using long polyethyleneglycol-based cross-linkers (**6**)[27]. A noteworthy aspect of that work was the finding that relatively long and flexible cross-link architectures can efficiently stabilize a complex tertiary structure.

In addition to stabilizing natural tertiary structures, chemical cross-links have been applied to link isolated secondary structures in nonlinear topologies and facilitate their artificial spatial arrangement into complex protein-like architectures. Early examples involve the use of cyclic β-hairpins[28] or porphyrins[29] as scaffolds for the covalent attachment of four helical peptides resulting in folded higher-order structures. In later work, natural and de novo-designed coiled coils[30,31] served as starting point for the design of cross-linked helix bundles[32–34]. Dimeric coiled coils (Fig. 2c) have been stabilized using various linkages such as ethylene glycol (**7**) and dibenzyl ether-based structures (**8**) installed via biselectrophilic cross-linkers and targeting two cysteine side chains[32,33]. Analogously, cross-linked trimeric helix bundles were created, revealing that robust assembly relies on carefully designed cross-linkers and supporting interhelical hydrophobic contacts[32,34]. Hydrophobic bis- and triselectrophiles, previously used for the generation of mono- and bicyclic peptides[35,36], have also been utilized to create covalent linkages in the hydrophobic core of designed artificial tertiary folds (termed the CovCore approach)[37]. These proteomimetics were de novo designed using computational approaches and resulted in structures of high thermal stability, as observed for a triangular, head-to-tail cyclized three-helix domain arranged around a central 1,3,5-trimethyl benzene linker (**9**) (Fig. 2d)[37].

The cross-linking approaches described above implement changes in protein topology via the covalent modification of amino acid side chains, resulting in permanent structural constraints. Alternatively, secondary structure elements, in particular α-helices, have been assembled using controllable moieties to devise dynamic proteomimetics. A switchable scaffold has been constructed via head-to-head ligation of a parallel coiled coil connected by a peptide loop containing a binding epitope[38]. Receptor binding to this epitope leads to opening of the coiled-coil structure, which can be detected in real time employing an appropriately aligned fluorophore–quencher pair. In addition, three-helix assemblies have been templated using nucleic acid triplex formation facilitating the dynamic self-assembly of protein-like structures[39] and highlighting the potential of orthogonal self-assembly systems in the design of proteomimetics.

Partially artificial backbones

The canonical L-α-peptide backbone connectivity in proteins is not unique in conveying the ability to adopt a protein-like folding pattern[40–42]. Alteration of backbone composition was initially conceived in the context of peptidomimetics, where diverse artificial backbones — termed 'foldamers'[43] — have been used to mimic helices and other isolated segments of protein secondary structure[3,5]. Though primarily applied to generate peptidomimetics, alteration of backbone chemical composition also has utility in recreating larger, more complex folds and thus proteomimetics. Changing the backbone of a protein can have dramatic effects, such as altered folding behaviour, molecular recognition, stability to degradation and dynamics. Building blocks that have been applied to protein backbone modification are diverse in structure (Fig. 3a). Some are close structural analogues of native L-α-residues, while others bear little resemblance to a natural protein backbone. In addition to structural diversity, there is considerable variation in the number and density of backbone modifications that have been applied in a given chain. Substitution can involve a single amino acid, a contiguous segment of the backbone or the replacement of amino acids interspersed throughout a domain.

Pioneering efforts at protein backbone alteration in tertiary folds involved single-atom substitution to modulate the forces that drive folding and assembly (for example, replacement of a natural α-residue (**10**) with an N-Me-α analogue (**11**))[44–49]. Later work broadened this approach — often referred to as 'protein prosthesis' — to encompass more extensive changes to backbone chemical composition at a single site, with a particular eye to controlling conformational preferences at β-turns[50–54]. The concept of isolated prosthetic modification has been leveraged more recently as a strategy to control dynamics in the study of an intrinsically disordered sequence through the replacement of one or two α-residues with rigidified Cα-methyl analogues (**12**) in the activation domain from the p160 transcriptional co-activator for thyroid hormone and retinoid receptors[55]. Collectively, the above precedents establish that backbone connectivity of a natural sequence can be changed without abolishing the ability to adopt a complex tertiary fold. An important characteristic shared by these examples is that the altered backbone connectivity was isolated to one or two sites. Although such limited modification can have profound effects, chemists have been motivated to produce complex tertiary folds from backbones that are more artificial in composition. To this end, a number of reports have demonstrated mimicry of prototype proteins by variants in which an α-helical segment is replaced by a helix mimic. In some cases, the prosthetic helix replacement was entirely artificial in chemical composition. For example, substitution of a helix in the chemokine interleukin-8 with a helix composed of β^3-residues (**13**)[56] and substitution of a helix in the transcription factor hypoxia inducible factor-1α with an aromatic oligoamide helix mimetic[57] both led to functional variants of the prototype protein. Related, the α-helix in a zinc finger domain was recently replaced by a helical oligomer of urea-based residues (**14**)[58]. Beyond adopting a native-like fold and metal-binding site (Fig. 3b), the chimeric zinc

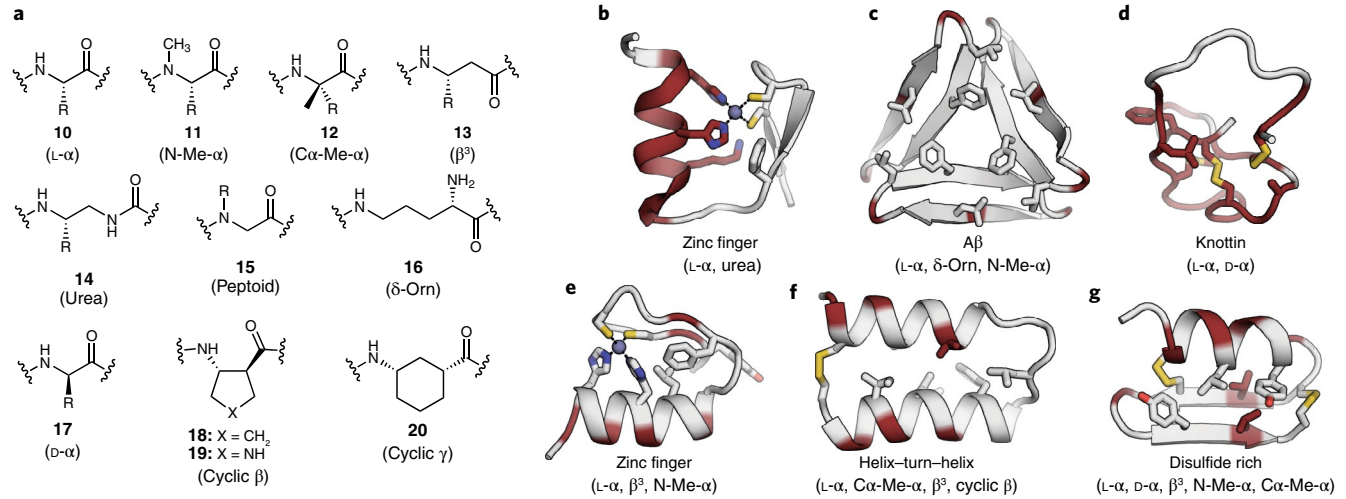

Fig. 3 | Examples of proteomimetics based on partially artificial backbone compositions. a, Chemical structures of selected residue types used to construct modified protein backbones. **b**–**g**, Experimentally determined folded structures of modified-backbone proteomimetics. Native L-α-residues are shown in grey and non-natural monomers in red. Each panel is labelled with the domain mimicked and a list of residue types in the chain. **b**, NMR structure of a zinc finger domain mimic with an artificial helix (PDB 5N14)[58]. **c**, Crystal structure of a trimeric assembly formed by a sequence derived from amyloid β (Aβ) with turn-inducing and sheet-disrupting modifications (PDB 4NTR)[94]. **d**, NMR structure of a knottin mimic with a heterochiral backbone comprising a D-α-peptide core and native loop[67]. **e**, NMR structure of a zinc finger domain mimic with backbone modifications in the helix and hairpin regions (PDB 5US3)[71]. **f**, Crystal structure of a helix–turn–helix scaffold modified in the helices (PDB 4WPB)[72] . **g**, NMR structure of a de novo disulfide-rich scaffold harbouring modifications throughout the helix, loop, turn and sheet (PDB 6E5J)[74].

finger domain was able to bind double-stranded DNA. This interaction is driven largely by contacts with the α-helix replaced in the prototype, underscoring the effective structural mimicry of this portion of the protein by the prosthetic segment.

An alternative approach to prosthetic protein backbone modification is to intersperse artificial units alongside natural α-residues to replace a local segment of secondary structure in a folded protein with a backbone that displays the natural side-chain sequence on a partially artificial main chain. An attractive characteristic of this approach is that it does not require a priori knowledge of a secondary structure mimetic compatible with a given prototype. Instead, careful modification enables the biological side-chain sequence to recapitulate key non-covalent contacts that drive folding and function on the newly artificial backbone. Early efforts towards the incorporation of locally heterogeneous backbones in protein tertiary folds involved replacement of a single secondary structure element. Examples of α-helix modifications include transposition of side chains from Cα to N in the enzyme ribonuclease A through incorporation of peptoid residues (**15**)[59] and the introduction of extra backbone methylene units in an engineered chorismite mutase through α- to β^3-residue substitution[60]. Localized backbone alterations in β-sheet-rich folds including β^3-residues, N-Me-α-residues and ornithine connected via its δ-amine (δ-Orn, **16**) have been applied to generate bioactive analogues of an antiangiogenic protein fragment[61] and to control the assembly behaviour of amyloidogenic sequences (Fig. 3c)[62–65]. The latter efforts were motivated by a desire to obtain fundamental insights into as well as potential inhibitors of amyloid formation related to diseases such as Alzheimer's and Huntington's. Protein domains in which the functionally relevant segment is a surface-exposed loop can also be subject to prosthetic modification, as illustrated by the introduction of a flexible spacer in an ion-channel-inhibitory conotoxin[66] and the grafting of an engineered receptor-binding peptide loop onto a disulfide-rich D-α-residue (**17**)-based knottin scaffold (Fig. 3d)[67]. In both cases, the modification improved stability to degradation by proteolytic enzymes.

The above examples illustrate that backbone alterations can be accommodated in any of the common secondary structures, but can such changes be applied globally? Answering this question requires consideration of two key challenges. The first relates to monomer selection. Because different backbones are predisposed to stabilize different local folds, it is difficult to create a proteomimetic consisting of an array of different secondary structures based on just a single α-residue replacement type. The second challenge relates to the plasticity of natural side-chain sequences with respect to functioning on backbones of altered composition. In many of the examples cited above, backbone modifications made to an isolated part of a protein had a small unfavourable effect on fold stability and/or function. This is not surprising, as the sequences being mimicked did not evolve to function on the types of employed backbones. Despite these challenges, substantial progress has been made in generating proteomimetics from more densely modified backbones.

One of the first examples to apply systematic backbone modification throughout a tertiary structure entailed the mimicry of a small, phage-derived peptide inhibitor of vascular endothelial growth factor signalling with an ordered but irregular folding pattern[68]. Substitution of selected α-residues in the native sequence with a combination of β^3-residues and cyclic β-residue **18** led to an analogue that showed an improvement in stability to degradation by protease enzymes, albeit at the expense of a modest loss in receptor affinity. Moving towards mimicry of more complex folds required the application of a wider assortment of α-residue replacements. An early illustration of this is seen in the development of heterogeneous-backbone variants of the B1 domain of streptococcal protein G (GB1), which contains helix, sheet, turn and loop secondary structures and is stabilized by a well-packed hydrophobic core[69]. The key to successful recreation of the GB1 tertiary fold was to employ a variety of different backbone modification types (for example, D-α-, N-Me-α-, Cα-Me-α- and β^3-residues, as well as cyclic β-residue **19** and cyclic γ-residue **20**). Subsequent work on the GB1 system yielded a set of general design principles for the application of diverse building block types in the construction of heterogeneous-backbone proteomimetics[70]. These design principles have since been shown to be applicable to a variety of other proteins.

Aiming for the reproduction of a zinc finger, it was found that some heterogeneous-backbone variants of the prototype containing a combination of β³- and N-Me-α-residues exhibited higher thermodynamic folded stability than the natural backbone (Fig. 3e)[71]. This was a noteworthy result in illustrating that a biologically derived side-chain sequence can fold more favourably on a heterogeneous backbone. This study also highlights challenges in developing proteomimetics based on heterogeneous backbones, as a metal-binding β-turn in the prototype proved completely intolerant to a number of backbone alterations known to effectively stabilize turns in other contexts. The ability to mimic isolated helical secondary structures with heterogeneous backbones has been leveraged to produce mimics of a bioactive helix–turn–helix tertiary scaffold in which the helices incorporate a combination of Cα-Me-α-, β³- and cyclic β-residues (Fig. 3f)[72]. Noteworthy here was the observation that receptor affinities of the variants were in some cases indistinguishable from the natural-backbone prototype. The functional potential of heterogeneous-backbone proteomimetics was further highlighted by a recent report of artificial variants of the protein ubiquitin that are able to engage in native multi-protein signalling pathways[73]. De novo-designed tertiary structures have also served as inspiration for the design of proteomimetics, as shown by the backbone modification of a computationally designed disulfide-rich scaffold to yield mimetics with identical folds and increased proteolytic stability (Fig. 3g)[74].

Completely artificial backbones

The design of proteomimetics based on completely artificial backbones is a particularly challenging endeavour as it requires the de novo design of a side-chain sequence able to specify a tertiary fold, and at the same time a backbone capable of manifesting the sequence-encoded fold. Notably, the resulting architectures, unhindered by natural constraints of protein-like folding patterns, offer the prospect of functions beyond those known from natural systems. Pioneering efforts that paved the way towards completely artificial proteomimetics involved a range of oligamide-based foldamers capable of adopting robust secondary structures[40–42,75]. While these molecules were inspired to some degree by biomacromolecules, many of the resulting folds depart considerably from nature, complicating their use for the creation of tertiary folding patterns. The difficulties associated with the design of completely artificial proteomimetics explain the smaller number of mimetics compared with the two categories discussed above. Already existing examples, however, highlight the potential of complex architectures that lack analogous prototypes in nature.

Early efforts towards complex folded structures composed entirely of artificial monomer units entailed the design of amphiphilic helices that assemble into defined helix bundles. Examples with folded structures characterized at high resolution include an oligomer of β³-residues (**13**) that forms an octameric bundle (Fig. 4a)[76] and an oligomer of urea-based monomers (**14**) that forms a hexameric assembly (Fig. 4b)[77]. Both these systems are formally an example of quaternary structure as each consists of an assembly of isolated secondary structure elements. Nevertheless, these prototypes were later advanced to create analogous helix-bundle assemblies with complex functions, including small-molecule recognition[78], catalysis[79,80] and metal binding[81]. The first examples towards unimolecular tertiary structure from completely artificial backbones involved the creation of helix–turn–helix motifs by a design analogous to that of abovementioned helix bundles[82]. Results here were encouraging, although the proposed tertiary structures were not confirmed by X-ray crystallography or NMR.

Oligoamides based on rigid aromatic building blocks (for example, **21**–**29**, Fig. 4c) have proven to be a particularly robust platform for creating complex folding patterns. A sophisticated example of a corresponding unimolecular folding motif is the helical molecular capsule that lacks an analogous architecture in natural proteins. Patterning aromatic amide building blocks of differing length, geometry and composition along a helical chain yielded a defined internal pocket that could selectively bind a monosaccharide guest (Fig. 4d)[83]. Notably, this architecture is formally at the level of secondary structure; however, it is able to accomplish a function (molecular encapsulation) that in a protein would require a higher-order tertiary fold. Design principles towards aromatic oligoamides have been advanced more recently to construct true tertiary folds comprising multiple secondary structures arrayed in a single entirely artificial chain. Corresponding helix–turn–helix (Fig. 4e)[84] and helix–sheet–helix motifs[85], as well as multi-stranded sheets[86] have been designed and characterized at atomic resolution. Most noteworthy, the helix–turn–helix system displays cooperativity in its folding behaviour[87] — a hallmark feature of natural proteins.

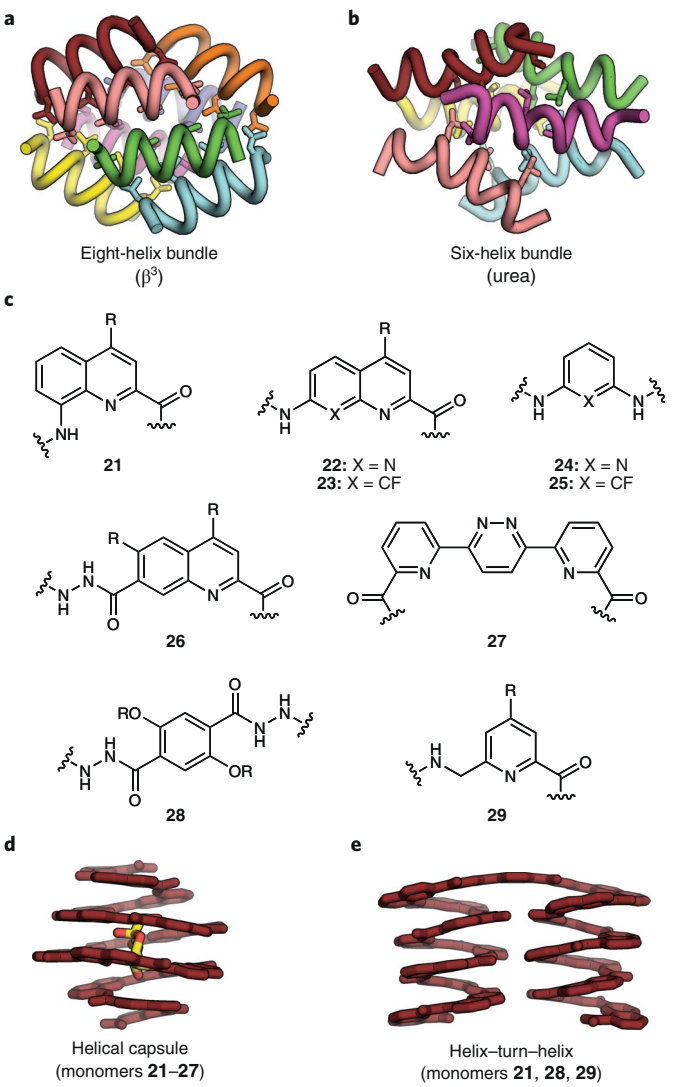

Fig. 4 | Examples of proteomimetics based on entirely artificial backbone compositions. a,**b**, Crystal structures of an octameric helix bundle assembly formed by a β³-peptide (**a**)[76] and a hexameric helix bundle assembly formed by an urea oligomer (**b**)[77]; structures are coloured by chain. **c**, Chemical structures of monomers used in the construction of aromatic oligoamide foldamers. **d**, Crystal structure of fructose (yellow) in complex with an aromatic oligoamide helical capsule[83]. **e**, Crystal structure of an aromatic oligoamide helix-turn-helix motif[84].

Conclusions and future directions

The unique folding properties and functions of proteins have stimulated diverse efforts to understand these phenomena at the molecular level, and to mimic or improve certain protein features. Initially, small protein fragments were the focus of such work, resulting in a multitude of approaches towards peptidomimetic agents. Recently, there has been an increasing interest in the mimicry of larger more complex folds resulting in tertiary structure mimetics, which we term proteomimetics herein. These efforts have been mainly motivated by basic questions regarding protein-folding and -recognition processes as well as the design of high-affinity ligands or artificial agents with enzymatic activity. The past has seen the emergence of novel mimicry approaches that have often been applied individually, but more recent results show a trend towards the use of different approaches in combination. Notably, while proteomimetic approaches in the past have predominantly focused on the characterization of folding properties of the resulting molecules, current efforts appear to focus more on their use as biomimetic agents, highly selective binders or catalysts. In addition, recent advances in allied fields will enable the more efficient development and construction of proteomimetics. For example, one can expect that the number of biocompatible reactions used for the construction of proteomimetics will continuously increase, and that computational de novo protein design will give access to entirely novel proteomimetic scaffolds. Regarding efforts towards development of proteomimetics as bio-inspired catalysts, the rapidly evolving field of artificial metalloenzymes[88,89] will provide a rich source of inspiration. Notably, covalently modified proteins have already been evolved by using metal-coordinating non-natural amino acids[90–92] or artificial post-translational modification[93]. Proteomimetic approaches may be diverse, but scientists working in this area are united by a shared curiosity about protein function and the search for artificial substitutes. The emerging field of proteomimetics research promises to benefit from this diversity and lead to important advances that address central questions associated with human health and a more sustainable society.

Received: 10 September 2019; Accepted: 9 January 2020;
Published online: 6 February 2020

References

1. Sali, A., Shakhnovich, E. & Karplus, M. How does a protein fold? *Nature* **369**, 248–251 (1994).
2. Gante, J. Peptidomimetics — tailored enzyme inhibitors. *Angew. Chem. Int. Ed.* **33**, 1699–1720 (1994).
3. Pelay-Gimeno, M., Glas, A., Koch, O. & Grossmann, T. N. Structure-based design of inhibitors of protein–protein interactions: mimicking peptide binding epitopes. *Angew. Chem. Int. Ed.* **54**, 8896–8927 (2015).
4. Clercq, E. D. The design of drugs for HIV and HCV. *Nat. Rev. Drug. Discov.* **6**, 1001–1018 (2007).
5. Azzarito, V., Long, K., Murphy, N. S. & Wilson, A. J. Inhibition of α-helix-mediated protein–protein interactions using designed molecules. *Nat. Chem.* **5**, 161–173 (2013).
6. Robertson, N. S. & Spring, D. R. Using peptidomimetics and constrained peptides as valuable tools for inhibiting protein–protein interactions. *Molecules* **23**, 959 (2018).
7. Scott, D. E., Bayly, A. R., Abell, C. & Skidmore, J. Small molecules, big targets: drug discovery faces the protein–protein interaction challenge. *Nat. Rev. Drug. Discov.* **15**, 533–550 (2016).
8. Kent, S. B. H. Novel protein science enabled by total chemical synthesis. *Protein Sci.* **28**, 313–328 (2019).
9. Krall, N., da Cruz, F. P., Boutureira, O. & Bernardes, G. J. L. Site-selective protein-modification chemistry for basic biology and drug development. *Nat. Chem.* **8**, 102–112 (2016).
10. Renata, H., Wang, Z. J. & Arnold, F. H. Expanding the enzyme universe: accessing non-natural reactions by mechanism-guided directed evolution. *Angew. Chem. Int. Ed.* **54**, 3351–3367 (2015).
11. Huang, P. S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
12. Chin, J. W. Expanding and reprogramming the genetic code. *Nature* **550**, 53–60 (2017).
13. Orner, B. P., Ernst, J. T. & Hamilton, A. D. Toward proteomimetics: terphenyl derivatives as structural and functional mimics of extended regions of an α-helix. *J. Am. Chem. Soc.* **123**, 5382–5383 (2001).
14. Rennie, Y. K., McIntyre, P. J., Akindele, T., Bayliss, R. & Jamieson, A. G. A TPX2 proteomimetic has enhanced affinity for Aurora-A due to hydrocarbon stapling of a helix. *ACS Chem. Biol.* **11**, 3383–3390 (2016).
15. Dombkowski, A. A., Sultana, K. Z. & Craig, D. B. Protein disulfide engineering. *FEBS Lett.* **588**, 206–212 (2014).
16. Marfey, P. S., Nowak, H., Uziel, M. & Yphantis, D. A. Reaction of bovine pancreatic ribonuclease A with 1,5-difluoro-2,4-dinitrobenzene. *J. Biol. Chem.* **240**, 3264–3269 (1965).
17. Hartman, F. C. & Wold, F. Bifunctional reagents. Cross-linking of pancreatic ribonuclease with a diimido ester. *J. Am. Chem. Soc.* **88**, 3890–3891 (1966).
18. Uy, R. & Wold, F. in *Protein Crosslinking: Biochemical and Molecular Aspects* (ed. Friedman, M.) 169–186 (Springer, 1977).
19. Lin, S. H., Konishi, Y., Denton, M. E. & Scheraga, H. A. Influence of an extrinsic cross-link on the folding pathway of ribonuclease A. Conformational and thermodynamic analysis of cross-linked (lysine^{7}-lysine^{41})-ribonuclease A. *Biochemistry* **23**, 5504–5512 (1984).
20. Luhmann, T., Mong, S. K., Simon, M. D., Meinel, L. & Pentelute, B. L. A perfluoroaromatic abiotic analog of H2 relaxin enabled by rapid flow-based peptide synthesis. *Org. Biomol. Chem.* **14**, 3345–3349 (2016).
21. Martinez-Saez, N. et al. Oxetane grafts installed site-selectively on native disulfides to enhance protein stability and activity in vivo. *Angew. Chem. Int. Ed.* **56**, 14963–14967 (2017).
22. Ekblad, T. et al. Synthesis and chemoselective intramolecular crosslinking of a HER2-binding affibody. *Biopolymers* **92**, 116–123 (2009).
23. Lindgren, J. & Karlstrom, A. E. Intramolecular thioether crosslinking of therapeutic proteins to increase proteolytic stability. *ChemBioChem* **15**, 2132–2138 (2014).
24. Moore, E. J., Zorine, D., Hansen, W. A., Khare, S. D. & Fasan, R. Enzyme stabilization via computationally guided protein stapling. *Proc. Natl Acad. Sci. USA* **114**, 12472–12477 (2017).
25. Pelay-Gimeno, M., Bange, T., Hennig, S. & Grossmann, T. N. In situ cyclization of native proteins: structure-based design of a bicyclic enzyme. *Angew. Chem. Int. Ed.* **57**, 11164–11170 (2018).
26. Robinson, J. A. β-hairpin peptidomimetics: design, structures and biological activities. *Acc. Chem. Res.* **41**, 1278–1288 (2008).
27. Xiao, Q. et al. Stapling of two PEGylated side chains increases the conformational stability of the WW domain via an entropic effect. *Org. Biomol. Chem.* **16**, 8933–8939 (2018).
28. Mutter, M. et al. Template-assembled synthetic proteins with 4-helix-bundle topology. Total chemical synthesis and conformational studies. *J. Am. Chem. Soc.* **114**, 1463–1470 (1992).
29. Sasaki, T. & Kaiser, E. T. Synthesis and structural stability of helichrome as an artificial hemeproteins. *Biopolymers* **29**, 79–88 (1990).
30. Lupas, A. N. & Bassler, J. Coiled coils — a model system for the 21st century. *Trends Biochem. Sci.* **42**, 130–140 (2017).
31. Beesley, J. L. & Woolfson, D. N. The de novo design of α-helical peptides for supramolecular self-assembly. *Curr. Opin. Biotechnol.* **58**, 175–182 (2019).
32. Wuo, M. G., Hong, S. H., Singh, A. & Arora, P. S. Synthetic control of tertiary helical structures in short peptides. *J. Am. Chem. Soc.* **140**, 16284–16290 (2018).
33. Wuo, M. G., Mahon, A. B. & Arora, P. S. An effective strategy for stabilizing minimal coiled coil mimetics. *J. Am. Chem. Soc.* **137**, 11618–11621 (2015).
34. Wang, C. et al. Site-specific isopeptide bridge tethering of chimeric gp41 N-terminal heptad repeat helical trimers for the treatment of HIV-1 infection. *Sci. Rep.* **6**, 32161 (2016).
35. Timmerman, P., Beld, J., Puijk, W. C. & Meloen, R. H. Rapid and quantitative cyclization of multiple peptide loops onto synthetic scaffolds for structural mimicry of protein surfaces. *ChemBioChem* **6**, 821–824 (2005).
36. Heinis, C. & Winter, G. Encoded libraries of chemically modified peptides. *Curr. Opin. Chem. Biol.* **26**, 89–98 (2015).
37. Dang, B. B. et al. De novo design of covalently constrained mesosize protein scaffolds with unique tertiary structures. *Proc. Natl Acad. Sci. USA* **114**, 10852–10857 (2017).
38. Mueller, C. & Grossmann, T. N. Coiled-coil peptide beacon: a tunable conformational switch for protein detection. *Angew. Chem. Int. Ed.* **57**, 17079–17083 (2018).
39. Lou, C. G. et al. Peptide-oligonucleotide conjugates as nanoscale building blocks for assembly of an artificial three-helix protein mimic. *Nat. Commun.* **7**, 12294 (2016).
40. Hill, D. J., Mio, M. J., Prince, R. B., Hughes, T. S. & Moore, J. S. A field guide to foldamers. *Chem. Rev.* **101**, 3893–4012 (2001).
41. Bautista, A. D., Craig, C. J., Harker, E. A. & Schepartz, A. Sophistication of foldamer form and function in vitro and in vivo. *Curr. Opin. Chem. Biol.* **11**, 685–692 (2007).
42. Goodman, C. M., Choi, S., Shandler, S. & DeGrado, W. F. Foldamers as versatile frameworks for the design and evolution of function. *Nat. Chem. Biol.* **3**, 252–262 (2007).
43. Gellman, S. H. Foldamers: a manifesto. *Acc. Chem. Res.* **31**, 173–180 (1998).

44. Rajarathnam, K. et al. Neutrophil activation by monomeric interleukin-8. *Science* **264**, 90–92 (1994).
45. Chapman, E., Thorson, J. S. & Schultz, P. G. Mutational analysis of backbone hydrogen bonds in staphylococcal nuclease. *J. Am. Chem. Soc.* **119**, 7151–7152 (1997).
46. Lu, W., Qasim, M. A., Laskowski, M. & Kent, S. B. H. Probing intermolecular main chain hydrogen bonding in serine proteinase–protein inhibitor complexes: chemical synthesis of backbone-engineered turkey ovomucoid third domain. *Biochemistry* **36**, 673–679 (1997).
47. Beligere, G. S. & Dawson, P. E. Design, synthesis, and characterization of 4-ester CI2, a model for backbone hydrogen bonding in protein α-helices. *J. Am. Chem. Soc.* **122**, 12079–12082 (2000).
48. Wales, T. E. & Fitzgerald, M. C. The energetic contribution of backbone–backbone hydrogen bonds to the thermodynamic stability of a hyperstable P22 Arc repressor mutant. *J. Am. Chem. Soc.* **123**, 7709–7710 (2001).
49. Deechongkit, S. et al. Context-dependent contributions of backbone hydrogen bonding to β-sheet folding energetics. *Nature* **430**, 101–105 (2004).
50. Baca, M., Kent, S. B. H. & Alewood, P. F. Structural engineering of the HIV-1 protease molecule with a β-turn mimic of fixed geometry. *Protein Sci.* **2**, 1085–1091 (1993).
51. Jean, F. et al. Synthesis and structural and functional evaluation of a protein modified with a β-turn mimic. *J. Am. Chem. Soc.* **120**, 6076–6083 (1998).
52. Viles, J. H. et al. Design, synthesis and structure of a zinc finger with an artificial β-turn. *J. Mol. Biol.* **279**, 973–986 (1998).
53. Kaul, R., Angeles, A. R., Jäger, M., Powers, E. T. & Kelly, J. W. Incorporating β-turns and a turn mimetic out of context in loop 1 of the WW domain affords cooperatively folded β-sheets. *J. Am. Chem. Soc.* **123**, 5206–5212 (2001).
54. Arnold, U. et al. Protein prosthesis: a semisynthetic enzyme with a β-peptide reverse turn. *J. Am. Chem. Soc.* **124**, 8522–8523 (2002).
55. Schmidtgall, B. et al. Dissecting mechanism of coupled folding and binding of an intrinsically disordered protein by chemical synthesis of conformationally constrained analogues. *Chem. Commun.* **53**, 7369–7372 (2017).
56. David, R. et al. Artificial chemokines: combining chemistry and molecular biology for the elucidation of interleukin-8 functionality. *J. Am. Chem. Soc.* **130**, 15311–15317 (2008).
57. Burslem, G. M. et al. Towards "bionic" proteins: replacement of continuous sequences from HIF-1α with proteomimetics to create functional p300 binding HIF-1α mimics. *Chem. Commun.* **52**, 5421–5424 (2016).
58. Lombardo, C. M. et al. Design and structure determination of a composite zinc finger containing a nonpeptide foldamer helical domain. *J. Am. Chem. Soc.* **141**, 2516–2525 (2019).
59. Lee, B.-C. & Zuckermann, R. N. Protein side-chain translocation mutagenesis via incorporation of peptoid residues. *ACS Chem. Biol.* **6**, 1367–1374 (2011).
60. Mayer, C., Müller, M. M., Gellman, S. H. & Hilvert, D. Building proficient enzymes with foldamer prostheses. *Angew. Chem. Int. Ed.* **53**, 6978–6981 (2014).
61. Hegedüs, Z. et al. Foldameric α/β-peptide analogs of the β-sheet-forming antiangiogenic anginex: structure and bioactivity. *J. Am. Chem. Soc.* **135**, 16578–16584 (2013).
62. Cheng, P.-N., Liu, C., Zhao, M., Eisenberg, D. & Nowick, J. S. Amyloid β-sheet mimics that antagonize protein aggregation and reduce amyloid toxicity. *Nat. Chem.* **4**, 927–933 (2012).
63. Kar, K. et al. β-Hairpin-mediated nucleation of polyglutamine amyloid formation. *J. Mol. Biol.* **425**, 1183–1197 (2013).
64. Kar, K. et al. Backbone engineering within a latent β-hairpin structure to design inhibitors of polyglutamine amyloid formation. *J. Mol. Biol.* **429**, 308–323 (2017).
65. Kreutzer, A. G. & Nowick, J. S. Elucidating the structures of amyloid oligomers with macrocyclic β-hairpin peptides: insights into Alzheimer's disease and other amyloid diseases. *Acc. Chem. Res.* **51**, 706–718 (2018).
66. Green, B. R. et al. Conotoxins containing nonnatural backbone spacers: cladistic-based design, chemical synthesis, and improved analgesic activity. *Chem. Biol.* **14**, 399–407 (2007).
67. Mong, S. K. et al. Heterochiral knottin protein: folding and solution structure. *Biochemistry* **56**, 5720–5725 (2017).
68. Haase, H. S. et al. Extending foldamer design beyond α-helix mimicry: α/β-peptide inhibitors of vascular endothelial growth factor signaling. *J. Am. Chem. Soc.* **134**, 7652–7655 (2012).
69. Reinert, Z. E., Lengyel, G. A. & Horne, W. S. Protein-like tertiary folding behavior from heterogeneous backbones. *J. Am. Chem. Soc.* **135**, 12528–12531 (2013).
70. George, K. L. & Horne, W. S. Foldamer tertiary structure through sequence-guided protein backbone alteration. *Acc. Chem. Res.* **51**, 1220–1228 (2018).
71. George, K. L. & Horne, W. S. Heterogeneous-backbone foldamer mimics of zinc finger tertiary structure. *J. Am. Chem. Soc.* **139**, 7931–7938 (2017).
72. Checco, J. W. et al. Targeting diverse protein–protein interaction interfaces with α/β-peptides derived from the Z-domain scaffold. *Proc. Natl Acad. Sci. USA* **112**, 4552–4557 (2015).
73. Werner, H. M., Estabrooks, S. K., Preston, G. M., Brodsky, J. L. & Horne, W. S. Exploring the functional consequences of protein backbone alteration in ubiquitin through native chemical ligation. *ChemBioChem* **20**, 7752–7755 (2019).
74. Cabalteja, C. C., Mihalko, D. S. & Horne, W. S. Heterogeneous-backbone foldamer mimics of a computationally designed, disulfide-rich miniprotein. *ChemBioChem* **20**, 103–110 (2019).
75. Guichard, G. & Huc, I. Synthetic foldamers. *Chem. Commun.* **47**, 5933–5941 (2011).
76. Daniels, D. S., Petersson, E. J., Qiu, J. X. & Schepartz, A. High-resolution structure of a β-peptide bundle. *J. Am. Chem. Soc.* **129**, 1532–1533 (2007).
77. Collie, G. W. et al. Shaping quaternary assemblies of water-soluble non-peptide helical foldamers by sequence manipulation. *Nat. Chem.* **7**, 871–878 (2015).
78. Miller, J. P., Melicher, M. S. & Schepartz, A. Positive allostery in metal ion binding by a cooperatively folded β-peptide bundle. *J. Am. Chem. Soc.* **136**, 14726–14729 (2014).
79. Wang, P. S. P., Nguyen, J. B. & Schepartz, A. Design and high-resolution structure of a β^3-peptide bundle catalyst. *J. Am. Chem. Soc.* **136**, 6810–6813 (2014).
80. Bécart, D. et al. Helical oligourea foldamers as powerful hydrogen bonding catalysts for enantioselective C–C bond-forming reactions. *J. Am. Chem. Soc.* **139**, 12524–12532 (2017).
81. Collie, G. W. et al. Molecular recognition within the cavity of a foldamer helix bundle: encapsulation of primary alcohols in aqueous conditions. *J. Am. Chem. Soc.* **139**, 6128–6137 (2017).
82. Lee, B.-C., Zuckermann, R. N. & Dill, K. A. Folding a nonbiological polymer into a compact multihelical structure. *J. Am. Chem. Soc.* **127**, 10999–11009 (2005).
83. Chandramouli, N. et al. Iterative design of a helically folded aromatic oligoamide sequence for the selective encapsulation of fructose. *Nat. Chem.* **7**, 334–341 (2015).
84. De, S. et al. Designing cooperatively folded abiotic uni- and multimolecular helix bundles. *Nat. Chem.* **10**, 51–57 (2018).
85. Lamouroux, A. et al. Controlling dipole orientation through curvature: aromatic foldamer bent β-sheets and helix–sheet–helix architectures. *J. Am. Chem. Soc.* **139**, 14668–14675 (2017).
86. Atcher, J. et al. Aromatic β-sheet foldamers based on tertiary squaramides. *Chem. Commun.* **55**, 10392–10395 (2019).
87. Mazzier, D., De, S., Wicher, B., Maurizot, V. & Huc, I. Interplay of secondary and tertiary folding in abiotic foldamers. *Chem. Sci.* **10**, 6984–6991 (2019).
88. Hayashi, T., Hilvert, D. & Green, A. P. Engineered metalloenzymes with non-canonical coordination environments. *Chem. Eur. J.* **24**, 11821–11830 (2018).
89. Schwizer, F. et al. Artificial metalloenzymes: reaction scope and optimization strategies. *Chem. Rev.* **118**, 142–231 (2018).
90. Lee, H. S. & Schultz, P. G. Biosynthesis of a site-specific DNA cleaving protein. *J. Am. Chem. Soc.* **130**, 13194–13195 (2008).
91. Drienovská, I., Rioz-Martínez, A., Draksharapu, A. & Roelfes, G. Novel artificial metalloenzymes by in vivo incorporation of metal-binding unnatural amino acids. *Chem. Sci.* **6**, 770–776 (2015).
92. Drienovská, I. et al. Design of an enantioselective artificial metallo-hydratase enzyme containing an unnatural metal-binding amino acid. *Chem. Sci.* **8**, 7228–7235 (2017).
93. Obexer, R. et al. Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. *Nat. Chem.* **9**, 50–56 (2016).
94. Spencer, R. K., Li, H. & Nowick, J. S. X-ray crystallographic structures of trimers and higher-order oligomeric assemblies of a peptide derived from Aβ17–36. *J. Am. Chem. Soc.* **136**, 5595–5598 (2014).

Acknowledgements

We thank P. Arora and B. Pentelute for providing coordinates of proteomimetic structures from their published work. T.N.G. is grateful for support by the European Research Council (ERC starting grant number 678623). W.S.H. thanks the National Institutes of Health (GM107161) for financial support.

Author contributions

W.S.H. and T.N.G developed the concept, researched and wrote the manuscript.

Competing interests

T.N.G. is listed as an inventor on a patent application related to the INCYPRO stabilization approach.

Additional information

Correspondence should be addressed to W.S.H. or T.N.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.