Published in final edited form as:

Org Biomol Chem. 2014 November 28; 12(44): 8796–8802. doi:10.1039/c4ob01769b.

Protein Backbone Engineering as a Strategy to Advance Foldamers toward the Frontier of Protein-Like Tertiary Structure

Zachary E. Reinert and W. Seth Horne

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA

W. Seth Horne: horne@pitt.edu

Abstract

A variety of non-biological structural motifs have been incorporated into the backbone of natural protein sequences. In parallel work, diverse unnatural oligomers of *de novo* design (termed “foldamers”) have been developed that fold in defined ways. In this Perspective article, we survey foundational studies on protein backbone engineering, with a focus on alterations made in the context of complex tertiary folds. We go on to summarize recent work illustrating the potential promise of these methods to provide a general framework for the construction of foldamer mimics of protein tertiary structures.

Introduction

Proteins are central to life, yet their covalent connectivity is simple – a sequence-specific oligomer of α -amino acid residues each bearing one of twenty common side chains. The diverse functions proteins carry out are made possible by their ability to adopt well-defined folded conformations encoded by different side-chain sequences. Proteins may seem privileged scaffolds in the myriad folds accessible from a common chemical backbone, but they are not unique in this capacity.¹ Gellman coined the term “foldamer” to refer to any “polymer with a strong tendency to adopt a specific compact conformation.”^{1a} Proteins (and other so-called “biofoldamers”) clearly fall within this rubric; however, the main focus in the field of foldamer design has been on establishing the full scope of non-biological backbones that can fold in defined ways.

Much early work centered on recreating peptide secondary structures, where tremendous success has been achieved.¹ In recent years, researchers have begun to embrace the challenge of pushing foldamers beyond secondary structure to more complex tertiary and quaternary folding patterns reminiscent of those found in proteins. Early advances toward this goal included reports of bundled quaternary assemblies of foldameric helices² as well as related helix-turn-helix motifs.³ While these examples represent important milestones, the fraction of known natural protein folding motifs successfully recreated in non-biological backbones remains very small. A significant limiting factor that has held back the field is the challenge of design. Great strides have been made in elucidating the relationship between primary sequence and folding in proteins.⁴ When the fundamental connectivity of the

Correspondence to: W. Seth Horne, horne@pitt.edu.

oligomer backbone changes, however, a chemist must return to the drawing board to conceive side-chain sequences that will specify a desired fold.

Natural proteins are surprisingly tolerant to chemical changes to the α -peptide backbone, and diverse non-biological structural motifs have been successfully incorporated in the context of folded tertiary structures. Such modifications are typically made in isolation; however, using them in concert offers a strategy to convert natural protein sequences into oligomers with substantial unnatural backbone content and sophisticated folded structures. In the present Perspective article, we review foundational work in protein backbone engineering, with a focus on chemical alterations made in complex tertiary folds. We then summarize recent work from our lab illustrating the potential promise of these methods to provide a general framework for the design of heterogeneous-backbone foldamer mimics of natural protein tertiary structures.

Protein Backbone Engineering

In the early 1990's, two key technologies emerged that expanded the scope of protein engineering beyond side-chain substitution to encompass modifications to the polypeptide backbone itself.⁵ The first was the invention of techniques for heterologous expression of proteins containing building blocks beyond the 20 canonical amino acids.^{5a} The second was the improvement in efficiency of protein chemical synthesis resulting from the development of native chemical ligation and semisynthesis.^{5b,5c} Enabled by these advances, chemists have explored folding and function in a variety of proteins with unnatural backbone elements incorporated at one or more points in a natural sequence (Fig. 1).

One of the first questions protein backbone alteration was applied to address was the role of hydrogen bonding in folding. While the importance of main-chain hydrogen bonds to protein structure has long been appreciated, their precise energetic role was and remains a source of debate. Efforts to address hypotheses related to backbone hydrogen bonding directly were hampered by the need for an experimental system that enabled the isolated perturbation of hydrogen bonds in a folded protein with minimal other structural changes. A number of reports have shown that backbone amide $\rightarrow$ ester substitution meets this need.⁶ The replacement of an α -amino acid residue in a protein by an α -hydroxy acid with the same side chain is isosteric but substitutes a hydrogen-bond donor ($-\text{NH}-$) for an acceptor ($\text{O}-$) and weakens the neighboring carbonyl oxygen's capacity to act as an acceptor.

In one of the earliest examples of backbone ester incorporation in a protein, total chemical synthesis was used to replace a single amide in the turkey ovomucoid third domain with an ester.^{6a} Pioneering studies in this and other host sequences set the stage for detailed studies on energetic consequences of amide $\rightarrow$ ester substitution in diverse protein model systems.⁶ While more than a decade of work has yet to yield a definitive answer as to the exact energetic contribution of backbone hydrogen-bonds to folding, a rich data set on structure, thermodynamics, and kinetics in depsipeptide analogues of natural proteins has emerged from these studies.

An alternate isosteric amide replacement employed to probe hydrogen-bond strength in proteins is the substitution of the carbonyl oxygen with sulfur. The resulting thioamides are

weaker hydrogen bonding acceptors than amides, have an altered capacity for neighboring-residue orbital interactions, and can be used as minimalist chromophores for studying protein dynamics by spectroscopic methods.⁷

Whereas the work described above largely sought to use protein backbone modification to *understand* nature, other researchers have used such alterations to *improve* on nature (e.g., enhance the thermodynamic stability of a folded protein). One focus in this area has been on β -turn secondary structure. Building on a body of research showing diverse moieties that can promote tight turns in short peptides, chemists began to apply such modifications to larger folds.⁸ Early examples included the application of a bicyclic turn mimetic or a dibenzofuran-based scaffold to replace β -turns in several small proteins.^{8a-f} More recently, a number of unnatural-backbone turn replacements were incorporated at a fixed position in a single protein, the WW domain of Pin1.^{8g,8h} Studies on folding kinetics and thermodynamics in the resulting family of variants yielded insights into the folding mechanism and how it can be controlled by rational backbone modification.

Another fundamental aspect of protein structure subject to alteration by chemical synthesis is stereochemistry. In the early 1990's, several labs reported the total chemical synthesis of enantiomers of natural proteins.⁹ These studies showed that D-proteins adopt mirror-image folds compared to their natural L-counterparts (Fig. 2) and, in the case of enzymes, mirror-image substrate specificity. These efforts provided incisive experimental support for fundamental paradigms of molecular chirality. One intriguing application of D-proteins is their use in combination with identical or closely related L-protein counterparts to produce racemic or quasi-racemic protein crystals.¹⁰ This technology has allowed for the structures of several previously elusive proteins to be solved.

Total synthesis and semisynthesis have also been used to generate diastereomers of natural proteins containing a single D-amino acid.¹¹ In one example, a glycine residue from a potassium ion channel found in a left-handed helical conformation was replaced with a D-residue predisposed to this backbone arrangement.^{11a} In another study, a glycine in the capping region of an α -helix in ubiquitin was replaced with a D-residue, leading to variants with similar folded structures and stabilities to the natural protein.^{11b,11c}

Aminoisobutyric acid (Aib), the C_{α} -methylated variant of alanine, was one of the first modified backbones shown able to be incorporated into proteins by heterologous expression^{5a} and has also been introduced through chemical synthesis.^{11d,12} Aib has long been known to be a strong helix stabilizer in short peptides, and this ability was shown to hold up in the context of a protein tertiary structure.¹² Semisynthesis of a mutant of the C-terminal domain of thermolysin with one or two Aib residues in an α -helix segment led to a measurable stabilization of the tertiary fold.¹² In another report, Aib was incorporated into a loop of the protein HIV protease by total chemical synthesis in order to probe structure and dynamics in a flexible "flap" region of the enzyme crucial to substrate recognition (Fig. 3).^{11d}

A related form of protein backbone modification to C_{α} -methylation is the alkylation of amide nitrogens. In one early example, *N*-methylation was used to block dimerization of

interleukin 8.¹³ Self-association of the wild type is mediated by formation of a β -sheet with two strands from each monomer; site-specific *N*-methylation at this interface disrupts interstrand hydrogen bonding and led to a well-folded but monomeric interleukin 8 analogue. Backbone nitrogen atoms can also serve as points of attachment for larger side chains. *N*-Alkyl glycine or “peptoid” residues have found widespread use in foldamer design,^{1a–e} and they have begun to be examined in heterogeneous backbones based on natural proteins.¹⁴ One report carried out “translocation mutagenesis” in ribonuclease, in which several side chains in the native protein were relocated from C $_{\alpha}$ to the adjacent nitrogen through semisynthesis.^{14a} A variant with three peptoid residues showed enzymatic efficiency within 10-fold of wild-type. Another recent paper has shown chemical ligation at serine can provide synthetic access to chimeric peptoid/protein oligomers.^{14b}

Like peptoid building blocks, β -amino acids have also found dual use in *de novo* foldamer design¹ and as tools to engineer the backbone of natural sequences. Examples of β -residue incorporation in folded proteins include the replacement of an entire α -peptide secondary structure segment by a β -peptide counterpart. Such “prosthesis” strategies have been applied to replace a β -turn in ribonuclease and a helix from interleukin-8.¹⁵ A complementary method for using β -residues to engineer natural peptides and proteins is the replacement of α -residues at regular intervals in biologically derived sequences with β^3 analogues bearing the same side chain. Application of such sequence-guided $\alpha \rightarrow \beta$ residue substitution has produced mixed α/β -peptides that fold to form helix-bundle quaternary structures, act as potent mimics of protein–protein binding interactions, and replace an α -helix in the enzyme chorismate mutase.^{2d,16}

Though not isosteric to an amide, 1,2,3-triazoles have similar physical properties and have found utility as amide replacements in proteins.¹⁷ In one study, four amides in the core of a tetrameric α -helix bundle were replaced with 1,4-disubstituted triazoles, demonstrating the capacity of the heterocycle to participate in backbone hydrogen bonding in a folded protein context.^{17a} Triazoles have also been used to probe the role of amide *cis/trans* isomerization in folding.^{17b} A tertiary amide in ribonuclease was replaced with a fixed analogue based on either a 1,4- or 1,5-disubstituted triazole; the latter proved a more effective substitution for the *cis*-amide found in the native. Triazoles have also been used to facilitate protein synthesis, taking advantage of the copper(I)-catalyzed reaction between azides and alkynes for non-native ligation of suitably functionalized peptide fragments.^{17c}

Control over backbone amide isomerization has also been exerted by replacement of amides with substituted alkenes. Two reports used this strategy to probe the proline-rich collagen triple helix.¹⁸ Unlike the case of ribonuclease described above, *cis*-amide linkages in collagen are unproductive with respect to folding. One group locked an isomerizable tertiary amide in collagen by replacement with an *E*-alkene,^{18b} while another performed a similar replacement in an already *trans*-favoring secondary amide.^{18a} In both cases, amide $\rightarrow$ alkene substitution was found to destabilize the fold considerably, perhaps due to the non-isosteric nature of the replacements as well as the large difference in polarity between an amide and an olefin.

Complementary to the use of protein backbone alteration to rigidify structure, chemical changes can also be applied to enhance flexibility.¹⁹ Amide→thioether modification in a surface-exposed loop of the Pin1 WW domain was used to probe the role of enhanced flexibility by replacing conformationally constrained amide pseudo-double bond with a freely rotatable single bond.^{19a} We have shown that three-residue loops in the B1 domain of *Streptococcal* protein G can be replaced by equal-backbone length oligomers of polyethylene glycol without substantially disrupting the protein's fold or function.^{19b}

Protein Backbone Engineering as a Strategy to Generate Foldamer Tertiary Structure

In considering the design challenge of recreating complex tertiary structures on foldameric backbones, as outlined in the Introduction, we saw a possible connection to work on protein backbone engineering described in the previous section. Increasing the fraction of such modifications to a critical threshold creates heterogeneous oligomeric backbones that begin to look increasingly like typical "foldamers" conceived of by *de novo* design. To look at it another way, a protein can be thought of an oligomer with two orthogonal sequences: a sequence of monomer side chains and a sequence of monomer backbone structures. In the last few years, we have been exploring this idea experimentally. The central hypothesis guiding the work is that any natural protein can serve as the starting point for the design of its own unnatural mimic, given systematic rules for backbone modification.

In surveying precedent on protein backbone engineering, examples of β -sheet mimicry were underrepresented compared to turns, helices, and loops.²⁰ As a result, we focused some early efforts on better understanding how to effectively reproduce sheet folds on unnatural backbones. Our approach was to apply diverse backbone modifications to hairpin-forming α -peptide sequences and compare the folded structure and thermodynamic stability of the resulting oligomers by multidimensional NMR analysis in aqueous buffer.²¹ Alterations evaluated included α -residue replacement by *N*-Me- α -residues as well as β -residues and γ -residues of systematically differing backbone connectivity (side-chain placement, stereochemistry, ring constraint).

Comparison of the folding behavior of the above building blocks in hairpin host sequences led to the identification of several strategies for mimicry of sheet folds by heterogeneous backbones (Fig. 4). Replacement of α -residues with *N*-Me- α analogues is minimally perturbative to sheet formation, although such modification is limited to terminal strands due to the loss of a hydrogen-bond donor.^{21d} Hydrogen bonding capacity can be maintained by replacing α -residues in each strand with β -residues; however, both the exact structure of the β -residue employed as well as the pattern of α -residue substitution are important. 1:1 $\alpha \rightarrow \beta$ replacement with β^2 , β^3 , or $\beta^{2,3}$ residues of appropriate stereochemistry is well tolerated but the accompanying backbone elongation alters how side chains are displayed in the folded state.^{21a} In contrast, either 2:1 $\alpha\alpha \rightarrow \beta$ or 2:2 $\alpha\alpha \rightarrow \beta\beta$ substitution in each strand leads to a native-like fold.^{21b} The cyclically constrained γ^{cyc} -residue Acc proved to have an especially high sheet folding propensity. Its incorporation into a hairpin-encoding sequence by 1:1 $\alpha \rightarrow \gamma$ substitution in each strand led to a heterogeneous oligomer with a *more* stable fold than the natural backbone on which it was based.^{21c}

In parallel to the above work on hairpin mimicry, we have explored wholesale tertiary structure modification. We chose the B1 domain of protein G from *Streptococcus* bacteria (GB1) as a host sequence for these studies. GB1 has been a workhorse in biophysics, structural biology, and computational work going back more than two decades since its structure was first reported.²² It has also served as a subject of protein backbone engineering in a prior study where one of its turns was replaced by a dibenzofuran-based mimic.^{8d} From the standpoint of peptide and protein mimicry by foldamers, we saw the tertiary structure of GB1 (all the common elements of secondary structure in a single chain lacking disulfides) as a significant vertical step in complexity. Moreover, we were motivated by the hypothesis that design principles developed for backbone modification of GB1 would prove applicable to other target tertiary folds.

We used GB1 to test the applicability of design principles for sheet mimicry developed in the simple hairpin model sequences described above to a more complex protein tertiary structure context.^{21d} The best method for sheet modification in the hairpin system (1:1 $\alpha \rightarrow \gamma^{Cyc}$ residue substitution with Acc in each strand) was compatible with the tertiary fold of GB1 but somewhat destabilizing. Optimization of the placement of the unnatural residues to minimize the impact of backbone elongation on packing of hydrophobic core residues resulted in a variant with all four strands of the sheet modified and a folding free energy within ~ 1 kcal mol⁻¹ of wild-type.

We next sought to establish whether diverse strategies for backbone alteration could be combined in a single host sequence (Fig. 5).²³ We used four classes of unnatural building blocks in concert, each employed individually in prior work on protein backbone engineering, to modify GB1: *N*-Me- α -residues, D-residues, β^3 -residues, and the C $_{\alpha}$ -Me- α -residue Aib. As detailed above, the most important consequence of backbone N-methylation is the loss of a hydrogen-bond donor. In comparison to a typical L- α -residue, Aib occupies a restricted conformational space and is known to promote both helix and turn secondary structures.^{11d,12} D- α -residues show mirror image backbone torsional preferences to their natural counterparts, and they are especially adept at promoting tight turns in heterochiral backbones.⁹⁻¹⁰ The enhanced conformational flexibility of β^3 -residues resulting from three freely rotatable backbone bonds makes them compatible with a variety of secondary structures, including α -helix, β -turn, and β -sheet.^{2d,15-16,21a,21b} Although both β^3 -residues and Aib are well accommodated in helices, only the former maintain native side-chain functionality.

Informed by the above considerations, we individually replaced each secondary structural element in GB1 with a modified backbone based on the natural sequence. In the helix and loops, we replaced a subset of α -residues with β^3 -counterparts bearing the same side chain. In the sheet, we modified the two capping strands by $\alpha \rightarrow N$ -Me- α replacement. In the turns, we used two different unnatural residues known to promote tight backbone reversal: Aib and a D- α -residue.

We first evaluated the individual impact of each backbone change on folding through a combination of X-ray crystallography and circular dichroism thermal denaturation studies. These experiments showed that each modification was structurally well tolerated but

reduced the thermodynamic stability of the fold. We next combined all four modified secondary structures into a single sequence, leading to a variant of GB1 with ~20% unnatural backbone content. The thermodynamic folded stability of the resulting oligomer, though low, was near that predicted by simple additivity of the observed changes to folding free energy for modification of the individual secondary structures. This is an important finding because it suggests continued optimization of strategies for backbone alteration of isolated secondary structures should also improve folded stability in highly modified and complex folds.

In an effort to improve the stability of the fold and further explore the dual-sequence concept (backbone and side chain), we made side-chain alterations in the heterogeneous-backbone GB1 mimic. When the ~20% modified backbone was combined with a known computationally optimized side-chain sequence, the resulting heterogeneous-backbone oligomer showed a significant improvement in thermodynamic folded stability. The finding that side-chain alterations have the same stabilizing effect when made on a natural vs. heterogeneous backbone provides strong evidence for the notion that backbone and side-chain sequences can be thought of as orthogonal from a design perspective.

In addition to our overarching goal to develop new design principles for foldamer mimicry of proteins, we have also worked to better understand fundamental relationships between backbone connectivity and folding thermodynamics (Fig. 6).²⁴ In our modification of GB1 secondary structures described above, we found $\alpha \rightarrow \beta^3$ substitution in the helix and loops was tolerated but somewhat destabilizing to the folded state. From a simple structural comparison, it would be reasonable to surmise an entropic penalty towards folding from $\alpha \rightarrow \beta^3$ residue replacement due to the additional backbone conformational freedom in the latter. This paradigm motivated pioneering work showing β -peptide folding behavior can be controlled through backbone cyclization, which restricts the conformational space available to β -residues. As an example, a *trans*-cyclopentane-based β^{cyc} -residue promotes a particular helical conformation in β -peptides²⁵ and also supports α -helix-like folds in heterogeneous α/β -peptides based on natural sequences.^{16b,16c}

We incorporated β^3 -residues and β^{cyc} -residues throughout GB1 in an effort to evaluate the thermodynamic impact of $\alpha \rightarrow \beta^3$ residue replacement on folding enthalpy and entropy, the ability of $\beta^3 \rightarrow \beta^{\text{cyc}}$ replacement to stabilize the fold, and the energetic consequences of backbone rigidification.²⁴ We characterized folding in the wild-type protein and modified variants by X-ray crystallography and biophysical analysis.

Crystal structures of several mutants confirmed they adopt identical folds to the natural protein (Fig. 6A). Both $\alpha \rightarrow \beta^3$ and $\beta^3 \rightarrow \beta^{\text{cyc}}$ alterations were found to be enthalpically unfavorable towards folding (Fig. 6B). Backbone rigidification by $\beta^3 \rightarrow \beta^{\text{cyc}}$ substitution was entropically favorable, consistent with theory; however, the enthalpic penalty outweighed any entropic stabilization in most cases. Surprising was the finding that $\alpha \rightarrow \beta^3$ replacement was entropically *favorable* to folding. At first glance, this result would seem to run counter to assumptions about different backbone flexibilities in α - vs. β^3 -peptides (i.e., that folding a more flexible backbone containing β^3 residues should be entropically more costly).

Close analysis of the thermodynamic data, however, revealed significantly altered heat capacities and sensitivity to chemical denaturant in the modified backbones. We interpreted these data to indicate the surprising favorable change in folding entropy accompanying $\alpha \rightarrow \beta^3$ residue substitution resulted from a change in solvation of the denatured state. Specifically, enhanced backbone torsional freedom leads to a denatured ensemble with less residual folded structure and greater solvent exposed surface area. Transition to the native state releases much of the ordered water solvating the protein to disordered bulk in an entropically favorable contribution to folding. Other things being equal, more flexible backbones will contribute a greater degree of this desolvation on a per-residue basis. In general terms, the above findings demonstrate that protein backbone alteration can have complex and unanticipated effects on folding thermodynamics.

Conclusions

In summary, significant progress has been made in protein secondary and quaternary structure mimicry by unnatural foldamer backbones; however, recreating the full array of tertiary structures found in nature remains a significant design challenge. A large and growing body of work has shown that, despite the intricate interplay of intermolecular forces that specify folding behavior in proteins, their backbones can be surprisingly tolerant to chemical modification. Although such backbone engineering has typically been performed at isolated sites, systematically combining modifications provides a potentially general framework for the design of heterogeneous-backbone foldamers with complex tertiary structures based on natural protein sequences.

Advances in total chemical synthesis, semisynthesis, and biological expression of proteins with modified backbones notwithstanding, production of chains that are long and highly modified remains a challenging endeavor. New technologies that address this need would both expand the scope of protein targets amenable to backbone engineering and push the boundaries of functions accessible to foldamers.

A number of unanswered questions remain for the fields of protein engineering and foldamer design. Among active avenues in our ongoing research program in these areas, we are establishing the scope and limitations of design principles for tertiary structure mimicry articulated in GB1 by applying them to new prototype sequences. We are also working to devise and test improved methods for backbone engineering that stabilize native folded states. Finally, we are examining how significant backbone modification influences the biological properties of proteins (e.g., function, biostability). We anticipate ongoing work in these areas will both enhance fundamental understanding of structure and function in natural proteins and inform the design of unnatural entities that mimic those structures and functions.

Acknowledgments

Funding supporting this work was provided by the University of Pittsburgh and the National Institutes of Health (R01GM107161).

Notes and References

1. (a) Gellman SH. *Acc Chem Res.* 1998; 31:173–180.(b) Hill DJ, Mio MJ, Prince RB, Hughes TS, Moore JS. *Chem Rev.* 2001; 101:3893–4011. [PubMed: 11740924] (c) Bautista AD, Craig CJ, Harker EA, Schepartz A. *Curr Opin Chem Biol.* 2007; 11:685–692. [PubMed: 17988934] (d) Goodman CM, Choi S, Shandler S, DeGrado WF. *Nat Chem Biol.* 2007; 3:252–262. [PubMed: 17438550] (e) Horne WS. *Expert Opin Drug Discov.* 2011; 6:1247–1262. [PubMed: 22647064] (f) Guichard G, Huc I. *Chem Commun.* 2011; 47:5933–5941.(g) Pilsel LA, Reiser O. *Amino Acids.* 2011; 41:709–718. [PubMed: 21462000] (h) Martinek TA, Fulop F. *Chem Soc Rev.* 2012; 41:687–702. [PubMed: 21769415]
2. (a) Raguse TL, Lai JR, LePlae PR, Gellman SH. *Org Lett.* 2001; 3:3963–3966. [PubMed: 11720580] (b) Cheng RP, DeGrado WF. *J Am Chem Soc.* 2002; 124:11564–11565. [PubMed: 12296699] (c) Qiu JX, Petersson EJ, Matthews EE, Schepartz A. *J Am Chem Soc.* 2006; 128:11338–11339. [PubMed: 16939241] (d) Horne WS, Price JL, Keck JL, Gellman SH. *J Am Chem Soc.* 2007; 129:4178–4180. [PubMed: 17362016]
3. (a) Lee BC, Zuckermann RN, Dill KA. *J Am Chem Soc.* 2005; 127:10999–11009. [PubMed: 16076207] (b) Delsuc N, Léger JM, Massip S, Huc I. *Angew Chem Int Ed.* 2007; 46:214–217.(c) Petersson EJ, Schepartz A. *J Am Chem Soc.* 2008; 130:821–823. [PubMed: 18166055] (d) Price JL, Hadley EB, Steinkruger JD, Gellman SH. *Angew Chem Int Ed.* 2010; 49:368–371.
4. Dill KA, MacCallum JL. *Science.* 2012; 338:1042–1046. [PubMed: 23180855]
5. (a) Ellman JA, Mendel D, Schultz PG. *Science.* 1992; 255:197–200. [PubMed: 1553546] (b) Dawson P, Muir T, Clark-Lewis I, Kent S. *Science.* 1994; 266:776–779. [PubMed: 7973629] (c) Muir TW, Sondhi D, Cole PA. *Proc Natl Acad Sci USA.* 1998; 95:6705–6710. [PubMed: 9618476]
6. (a) Lu W, Qasim MA, Laskowski M, Kent SBH. *Biochemistry.* 1997; 36:673–679. [PubMed: 9020764] (b) Chapman E, Thorson JS, Schultz PG. *J Am Chem Soc.* 1997; 119:7151–7152.(c) Koh JT, Cornish VW, Schultz PG. *Biochemistry.* 1997; 36:11314–11322. [PubMed: 9298950] (d) Shin I, Ting AY, Schultz PG. *J Am Chem Soc.* 1997; 119:12667–12668.(e) Belligere GS, Dawson PE. *J Am Chem Soc.* 2000; 122:12079–12082.(f) Wales TE, Fitzgerald MC. *J Am Chem Soc.* 2001; 123:7709–7710. [PubMed: 11480998] (g) Blankenship JW, Balambika R, Dawson PE. *Biochemistry.* 2002; 41:15676–15684. [PubMed: 12501196] (h) Silinski P, Fitzgerald MC. *Biochemistry.* 2003; 42:6620–6630. [PubMed: 12767246] (i) Deechongkit S, Nguyen H, Powers ET, Dawson PE, Gruebele M, Kelly JW. *Nature.* 2004; 430:101–105. [PubMed: 15229605] (j) Gao J, Bosco DA, Powers ET, Kelly JW. *Nat Struct Mol Biol.* 2009; 16:684–690. [PubMed: 19525973]
7. (a) Miwa JH, Pallivathural L, Gowda S, Lee KE. *Org Lett.* 2002; 4:4655–4657. [PubMed: 12489953] (b) Goldberg JM, Batjargal S, Petersson EJ. *J Am Chem Soc.* 2010; 132:14718–14720. [PubMed: 20886849] (c) Newberry RW, VanVeller B, Guzei IA, Raines RT. *J Am Chem Soc.* 2013; 135:7843–7846. [PubMed: 23663100]
8. (a) Baca M, Kent SBH, Alewood PF. *Protein Sci.* 1993; 2:1085–1091. [PubMed: 8358291] (b) Jean F, Buisine E, Melnyk O, Drobecq H, Odaert B, Hugues M, Lippens G, Tartar A. *J Am Chem Soc.* 1998; 120:6076–6083.(c) Viles JH, Patel SU, Mitchell JBO, Moody CM, Justice DE, Uppenbrink J, Doyle PM, Harris CJ, Sadler PJ, Thornton JM. *J Mol Biol.* 1998; 279:973–986. [PubMed: 9642075] (d) Odaert B, Jean F, Melnyk O, Tartar A, Lippens G, Boutillon C, Buisine E. *Protein Sci.* 1999; 8:2773–2783. [PubMed: 10631995] (e) Kaul R, Angeles AR, Jäger M, Powers ET, Kelly JW. *J Am Chem Soc.* 2001; 123:5206–5212. [PubMed: 11457382] (f) Kaul R, Deechongkit S, Kelly JW. *J Am Chem Soc.* 2002; 124:11900–11907. [PubMed: 12358534] (g) Liu F, Du D, Fuller AA, Davoren JE, Wipf P, Kelly JW, Gruebele M. *Proc Natl Acad Sci USA.* 2008; 105:2369–2374. [PubMed: 18268349] (h) Fuller AA, Du D, Liu F, Davoren JE, Bhabha G, Kroon G, Case DA, Dyson HJ, Powers ET, Wipf P, Gruebele M, Kelly JW. *Proc Natl Acad Sci USA.* 2009; 106:11067–11072. [PubMed: 19541614]
9. (a) Milton, RCd; Milton, SCF.; Kent, SBH. *Science.* 1992; 256:1445–1448. [PubMed: 1604320] (b) Zawadzke LE, Berg JM. *J Am Chem Soc.* 1992; 114:4002–4003.(c) Fitzgerald MC, Chernushevich I, Standing KG, Kent SBH, Whitman CP. *J Am Chem Soc.* 1995; 117:11075–11080.
10. (a) Zawadzke LE, Berg JM. *Proteins: Struct, Funct, Bioinf.* 1993; 16:301–305.(b) Hung LW, Kohmura M, Ariyoshi Y, Kim SH. *J Mol Biol.* 1999; 285:311–321. [PubMed: 9878408] (c)

- Pentelute BL, Gates ZP, Tereshko V, Dashnau JL, Vanderkooi JM, Kossiakoff AA, Kent SBH. *J Am Chem Soc.* 2008; 130:9695–9701. [PubMed: 18598029]
11. (a) Valiyaveetil FI, Sekedat M, MacKinnon R, Muir TW. *Proc Natl Acad Sci USA.* 2004; 101:17045–17049. [PubMed: 15563591] (b) Bang D, Makhataдзе GI, Tereshko V, Kossiakoff AA, Kent SB. *Angew Chem Int Ed.* 2005; 44:3852–3856. (c) Bang D, Gribenko AV, Tereshko V, Kossiakoff AA, Kent SB, Makhataдзе GI. *Nat Chem Biol.* 2006; 2:139–143. [PubMed: 16446709] (d) Torbeev VY, Raghuraman H, Hamelberg D, Tonelli M, Westler WM, Perozo E, Kent SBH. *Proc Natl Acad Sci USA.* 2011; 108:20982–20987. [PubMed: 22158985]
12. De Filippis V, De Antoni F, Frigo M, Polverino de Lauro P, Fontana A. *Biochemistry.* 1998; 37:1686–1696.
13. Rajarathnam K, Sykes BD, Kay CM, Dewald B, Geiser T, Baggiolini M, Clark-Lewis I. *Science.* 1994; 264:90–92. [PubMed: 8140420]
14. (a) Lee BC, Zuckermann RN. *ACS Chem Biol.* 2011; 6:1367–1374. [PubMed: 21958072] (b) Levine PM, Craven TW, Bonneau R, Kirshenbaum K. *Org Lett.* 2014; 16:512–515. [PubMed: 24393000]
15. (a) Arnold U, Hinderaker MP, Nilsson BL, Huck BR, Gellman SH, Raines RT. *J Am Chem Soc.* 2002; 124:8522–8523. [PubMed: 12121081] (b) David R, Gunther R, Baumann L, Luhmann T, Seebach D, Hofmann HJ, Beck-Sickinger AG. *J Am Chem Soc.* 2008; 130:15311–15317. [PubMed: 18942784]
16. (a) Horne WS, Boersma MD, Windsor MA, Gellman SH. *Angew Chem Int Ed.* 2008; 47:2853–2856. (b) Horne WS, Price JL, Gellman SH. *Proc Natl Acad Sci USA.* 2008; 105:9151–9156. [PubMed: 18587049] (c) Horne WS, Johnson LM, Ketas TJ, Klasse PJ, Lu M, Moore JP, Gellman SH. *Proc Natl Acad Sci USA.* 2009; 106:14751–14756. [PubMed: 19706443] (d) Haase HS, Peterson-Kaufman KJ, Lan Levengood SK, Checco JW, Murphy WL, Gellman SH. *J Am Chem Soc.* 2012; 134:7652–7655. [PubMed: 22548447] (e) Cheloha RW, Maeda A, Dean T, Gardella TJ, Gellman SH. *Nat Biotechnol.* 2014; 32:653–655. [PubMed: 24929976] (f) Mayer C, Müller MM, Gellman SH, Hilvert D. *Angew Chem Int Ed.* 2014; 53:6978–6981.
17. (a) Horne WS, Yadav MK, Stout CD, Ghadiri MR. *J Am Chem Soc.* 2004; 126:15366–15367. [PubMed: 15563148] (b) Tam A, Arnold U, Soellner MB, Raines RT. *J Am Chem Soc.* 2007; 129:12670–12671. [PubMed: 17914828] (c) Valverde IE, Lecaille F, Lalmanach G, Aucagne V, Delmas AF. *Angew Chem Int Ed.* 2012; 51:718–722.
18. (a) Jenkins CL, Vasbinder MM, Miller SJ, Raines RT. *Org Lett.* 2005; 7:2619–2622. [PubMed: 15957905] (b) Dai N, Wang XJ, Etzkorn FA. *J Am Chem Soc.* 2008; 130:5396–5397. [PubMed: 18366169]
19. (a) Ferguson N, Pires JR, Toepert F, Johnson CM, Pan YP, Volkmer-Engert R, Schneider-Mergener J, Daggett V, Oschkinat H, Fersht A. *Proc Natl Acad Sci USA.* 2001; 98:13008–13013. [PubMed: 11687614] (b) Reinert ZE, Musselman ED, Elcock AH, Horne WS. *ChemBioChem.* 2012; 13:1107–1111. [PubMed: 22539233]
20. Nowick JS. *Acc Chem Res.* 2008; 41:1319–1330. [PubMed: 18798654]
21. (a) Lengyel GA, Frank RC, Horne WS. *J Am Chem Soc.* 2011; 133:4246–4249. [PubMed: 21370877] (b) Lengyel GA, Horne WS. *J Am Chem Soc.* 2012; 134:15906–15913. [PubMed: 22946450] (c) Lengyel GA, Eddinger GA, Horne WS. *Org Lett.* 2013; 15:944–947. [PubMed: 23390979] (d) Lengyel GA, Reinert ZE, Griffith BD, Horne WS. *Org Biomol Chem.* 2014; 12:5375–5381. [PubMed: 24909436]
22. Gronenborn A, Filpula D, Essig N, Achari A, Whitlow M, Wingfield P, Clore G. *Science.* 1991; 253:657–661. [PubMed: 1871600]
23. Reinert ZE, Lengyel GA, Horne WS. *J Am Chem Soc.* 2013; 135:12528–12531. [PubMed: 23937097]
24. Reinert ZE, Horne WS. *Chem Sci.* 2014; 5:3325–3330. [PubMed: 25071931]
25. Appella DH, Christianson LA, Klein DA, Powell DR, Huang X, Barchi JJ, Gellman SH. *Nature.* 1997; 387:381–384. [PubMed: 9163422]

Fig. 1. Examples of unnatural backbone substitutions made in place of amides (A) and α -amino acid residues (B) in the context of folded protein tertiary structures.

Fig. 2.
(A) Universal L- α $\rightarrow$ D- α -residue replacement (A) generates D-proteins that adopt mirror image folds to their natural counterparts, as seen in the crystal structure of racemic monellin (B, PDB 1KRL).

Fig. 3. Crystal structures of backbone-modified variants of HIV protease where a Gly residue in a flap flanking the substrate binding pocket is replaced with either Aib or D-Ala (PDB 3HAU, 3IAW, 3HLO).

Fig. 4. (A) Backbone alterations compatible with folding in heterogeneous-backbone hairpin sequences. (B) Example NMR structures of hairpin folds containing the indicated unnatural building block(s). Gray spheres indicate a cluster of hydrophobic side chains important for folding.

Fig. 5.
Examples of backbone alteration carried out on different secondary structural elements in GB1 (PDB 4KGR, 4KGS, 4KGT).

Fig. 6.
(A) Incorporation of β^3 and β^{cyc} residues in helix (top) and loop (bottom) contexts of protein GB1 (PDB 2QMT, 4KGR, 4KGS, 4OZB, 4OZC). (B) Effects of these substitutions on folding thermodynamics.