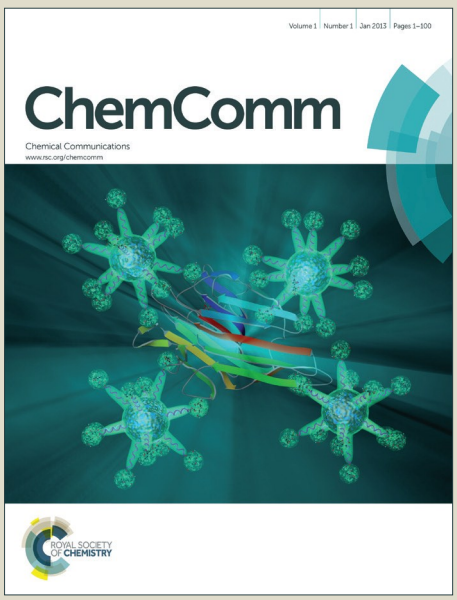
ChemComm

Accepted Manuscript



ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

COMMUNICATION

Comparison of Design Strategies for α-Helix Backbone Modification in a Protein Tertiary Fold†

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

www.rsc.org/

Nathan A. Tavenor, Zachary E. Reinert, George A. Lengyel‡, Brian D. Griffith and W. Seth Horne*

We report here the comparison of five classes of unnatural amino acid building blocks for their ability to be accommodated into an α-helix in a protein tertiary fold context. High-resolution structural characterization and analysis of folding thermodynamics yield new insights into the relationship between backbone composition and folding energetics in α-helix mimetics and suggest refined design rules for engineering the backbones of natural sequences.

Foldamers,[1] unnatural oligomers capable of adopting discrete folded structures reminiscent of those seen in nature, have found utility in a variety of applications.[2] In a body of research spanning more than 20 years, a wealth of secondary structures, including helices, turns, and sheets, have been shown to be accessible by backbones of diverse chemical compositions. A frontier challenge in the field is determining how to combine these secondary structures into more complex tertiary and quaternary folding topologies, either biologically inspired[3] or abiotic in design.[4] The significance of this as an objective stems from the prospect of expanding the range of functions accessible when diverse natural folding patterns are achievable by such agents.

While many foldamer structures have been developed through *de novo* design, backbone engineering of biological sequences has been shown as a viable strategy for recreating a variety of natural folds and functions.[5] In this approach, a portion of the α-amino acid residues in a designated mimetic target are replaced by analogues with an altered backbone, and modifications are made in a way that retains as many of the original side-chain functional groups as possible. The result is a "heterogeneous-backbone" oligomer consisting of a mixture of α-residues and unnatural counterparts that collectively display a native-like sequence of side chains (e.g., an α/β-peptide that blends α- and β-amino acid residues[6]). If modifications are made carefully, such analogues can show similar folding behaviour and biological function as the prototype natural sequence but improved stability to enzymatic degradation *in vitro*[7] and *in vivo*.[8]

Sequence-guided backbone engineering has found wide use in mimicry of helix[6a,8] and sheet[3a,9,10] secondary structures. We recently showed that this modification strategy is also capable of generating more complex tertiary folding patterns.[11] This goal was realized through the simultaneous modification of helix, loop, sheet, and turn secondary structures in a small bacterial protein with several unnatural residue classes. Examination of the effect of individual modifications revealed that changes to the α-helix were the most detrimental to tertiary fold stability. The prevalence of α-helices in proteins makes this an important limitation and reversing the destabilization resulting from helix backbone modification an important goal if backbone engineering is to prove a general method for developing foldamer analogues of a wider array of target tertiary folds.

Here, we compare five classes of unnatural-backbone units (Fig. 1A) for the ability to be accommodated into α-helical secondary structure in a protein context: β^3-residues, β^2-residues, the β^{cyc}-residue ACPC, the achiral C_α-Me-α-residue Aib, and chiral C_α-Me-α-residues (Fig. 1A). While each of these classes is known, no prior effort has compared them side-by-side for the ability to stabilize helical folds in a heterogeneous backbone. High-resolution structural characterization and biophysical analysis of folding thermodynamics in a series of variants of a common protein scaffold provide a robust picture of the relationship between backbone composition and folding propensity of α-helix mimetic foldamers.

The host sequence in the present work is protein GB1 (**1**, Fig. 1B,C), and our first-generation design for helix modification, previously reported,[11] makes use of β^3-residues in an $\alpha \rightarrow \beta^3$ substitution scheme that conserves the parent side chain at each site (**2**, Fig. 1D). Making one substitution in each turn of the GB1 helix resulted in an identical folded structure but destabilized the folded state considerably.[12] Protein **2** will serve as the benchmark for comparison of strategies for helix backbone modification examined herein.

Department of Chemistry, University of Pittsburgh, Pittsburgh, PA 15260, USA.
‡ Present address: Department of Chemistry, Slippery Rock University, Slippery Rock, PA 16057, USA.
† Electronic Supplementary Information (ESI) available: Figures S1-S7, Tables S1-S3, Scheme S1, and methods. See DOI: 10.1039/x0xx00000x

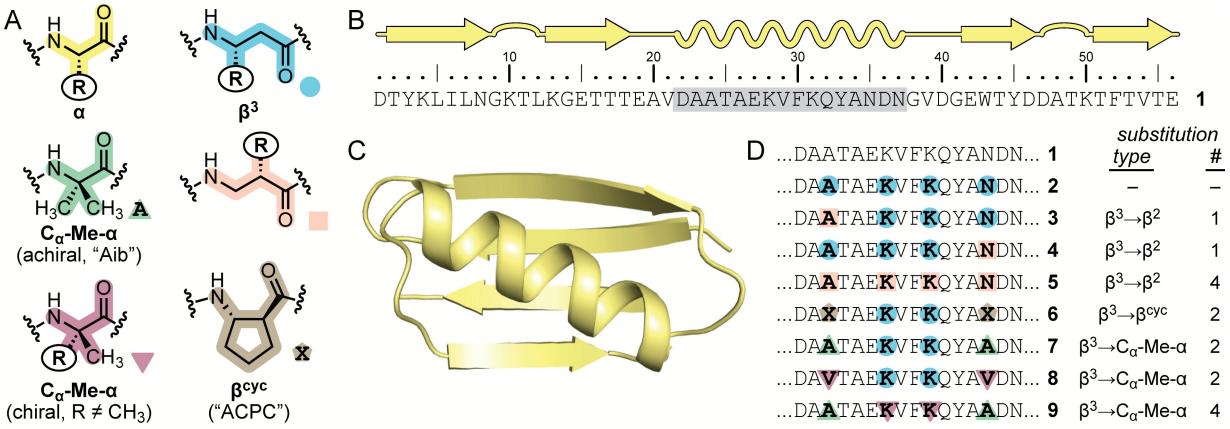


Fig. 1 (A) Structures of a natural α-residue and five classes of unnatural replacements compared herein. (B, C) Sequence and crystal structure (PDB 2QMT) of *Streptococcal* protein GB1 (**1**), the host sequence for helix modification; the crystal structure differs from the wild-type sequence at the N-terminus (MQ in crystal structure vs. DT in **1**). (D) Sequences for variants of protein **1** bearing heterogeneous-backbone helices; note, **1**-**9** are all 56-residue oligomers, but only the helical segment (gray shading in A) is shown. For β^3, β^2, and chiral C_α-Me-α-residues, the R group in the building block is that of the corresponding natural α-amino acid denoted by the single letter code in the sequence.

The first variable we investigated in the relationship between backbone composition and helix stability was the placement of β-residue side chains. β^3-residues, in which the side chain is attached adjacent to the amide nitrogen, are common foldamer building blocks due to their commercial availability. The regioisomeric β^2-residues, where the side chain is adjacent to the carbonyl, are less utilized; however, they have been studied in contexts including pure β-peptide helices[13] and mixed-backbone α/β-peptide sheets.[10a,10b] We were motivated to examine $\beta^3 \rightarrow \beta^2$ substitution in the helix of GB1 by two factors. First, results from a prior computational study suggest β^2-residues may be more predisposed than β^3 counterparts to support the backbone conformation adopted in heterogeneous-backbone α/β-peptide helices.[14] Second, the side chain movement from $\beta^3 \rightarrow \beta^2$ substitution restores a local orbital interaction involving an Asn side chain that may be important to folded stability (*vide infra*).

We designed proteins **3**-**5** (Fig. 1D) to ascertain the effect of β-residue side-chain placement on folded stability of helix-modified GB1 variants. Derivatives of β^2-Ala and β^2-Lys suitable for use in solid-phase synthesis were prepared by reported routes,[15] and a new protected form of β^2-Asn was prepared by adaptation of known methods (Scheme S1).[16] We synthesized proteins **3**-**5** by standard Fmoc solid phase methods and purified each by HPLC and ion exchange chromatography (Fig. S1, Table S1). Proteins were assayed by coupled thermal and chemical denaturation monitored by circular dichroism (CD) spectroscopy to probe folding thermodynamics (Fig. S4).[12,17] Each protein was also subjected to crystallization trials by hanging drop vapour diffusion, leading to single crystals of **3** and **4** that were analysed by X-ray diffraction and solved to 1.95 Å and 1.80 Å, respectively (Table S2).

Aside from the expected side-chain displacements, analogues **3** and **4** exhibit essentially identical tertiary folds as both natural backbone **1** and analogue **2** bearing an α/β^3 helix (Fig. 2). Analysis of folding thermodynamics revealed individual $\beta^3 \rightarrow \beta^2$ replacement was neutral to slightly destabilizing (Fig. 3), although magnitudes of the differences were small and similar within experimental uncertainty. The unfavourable effect on folding free energy is made up of an enthalpic stabilization that is more than overcome by an entropic penalty (Table S3). Though the exact origin of the entropy enthalpy compensation is not clear, changes in the sensitivity of the folded state to chemical denaturant and the heat capacity difference between the folded and unfolded states (Table S3) suggest a more compact denatured ensemble in β^2-residue containing variants vs. β^3 counterparts.[12]

Recent published findings highlight the importance of local orbital interactions in protein folding energetics.[18] One example is an intramolecular $n \rightarrow \pi^*$ overlap in Asn involving partial donation of a carboxamide lone pair into an antibonding orbital from the backbone carbonyl.[19] Based on the observed distance from side-chain C=**O** to backbone **C**=O,[19] evidence for this interaction is seen at α-Asn_{35} in wild-type **1** (PDB 2QMT[20]) and β^2-Asn_{35} in **4** (two of four chains in the asymmetric unit); however, the β^3-residue connectivity does

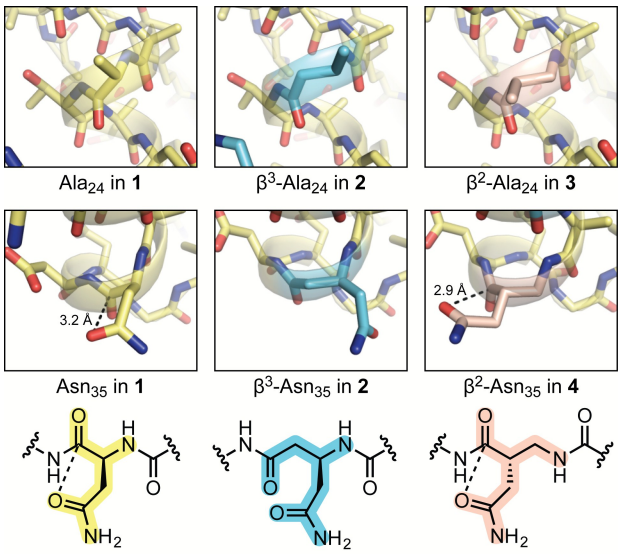


Fig. 2 Comparison of the corresponding α-, β^3-, and β^2-residues at two sites in the crystal structures of **1**-**4** (PDB 2QMT, 4KGR, 5HFY, and 5HG2). For position Asn_{35}, chemical structures and distances for putative side-chain to backbone $n \rightarrow \pi^*$ interactions are shown. In the structure for **1**, the side chain carboxamide of Asn_{35} is flipped relative to the reported structure (PDB 2QMT).

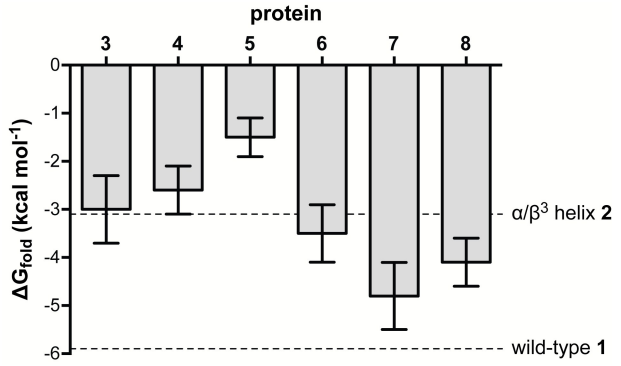


Fig. 3 Folding free energy at 298 K for proteins **1**-**8**. Error bars show parameter uncertainty from the fits (Fig S4, Table S3).

not allow the necessary side-chain orientation at β^3-Asn_{35} in **2** (Fig. 2). The $n\rightarrow\pi^*$ interaction has been suggested to be worth up to 1.2 kcal mol^{-1} in folding enthalpy.[19] An enthalpic stabilization of 0.4 kcal mol^{-1} was observed upon $\beta^3\rightarrow\beta^2$-Asn substitution in **4** vs. **2**; however, the change was comparable to that from $\beta^3\rightarrow\beta^2$-Ala replacement in **3** vs. **2**, where no functional side chain was involved. Collectively, the above data support the hypothesis that β^2 residues are superior to β^3 analogues in their ability to recapitulate side-chain to backbone orbital interactions; however, this difference does not appear to contribute significantly to folding energetics in this specific system.

Global $\beta^3\rightarrow\beta^2$ residue replacement in **2** to produce analogue **5** was significantly destabilizing, and the energetic penalty was almost entirely enthalpic in origin. This result was somewhat surprising given the enthalpic stabilization observed in variants **3** and **4**. Based on the data obtained for **2**-**4**, part of the ~1.6 kcal mol^{-1} difference in folding free energy between **2** and **5** can be attributed to the substitution of $\beta^3\rightarrow\beta^2$-Ala (0.1 kcal mol^{-1}) and $\beta^3\rightarrow\beta^2$-Asn (0.5 kcal mol^{-1}). The remaining 1.0 kcal mol^{-1} penalty comes from the $\beta^3\rightarrow\beta^2$-Lys replacements, and we suspect the dominant contribution arises from altered interactions involving Lys_{31}, as detailed below.

In wild-type protein **1**, Lys_{31} forms van der Waals contacts with Trp_{43} in the hydrophobic core and a salt bridge with Glu_{27}. In variants **2**-**4**, the same tertiary contacts are observed (Fig. S5). Crystallization attempts with **5** were not fruitful; however, movement of the side chain at Lys_{31} after $\beta^3\rightarrow\beta^2$ substitution should abolish these tertiary contacts (Fig. S5) and destabilize the fold. Supporting this hypothesis, removal of the side chain in question by $Lys_{31}\rightarrow$Ala mutation on the natural backbone results in a comparable degree of destabilization of 0.9 kcal mol^{-1} (Fig. S6, Table S3) as calculated for the remaining difference for **5** vs. **2** (*vide supra*).

From a design standpoint, the above results suggest that β^3- and β^2-residues are comparable in terms of fundamental folding propensity as components of heterogeneous-backbone α/β-peptide helices. Selection of the optimal regioisomer is context dependent and must take into account side-chain contacts important to folding and/or function. While the above examples show how $\beta^3\rightarrow\beta^2$ substitution can be detrimental, it stands to reason that an identical adjustment of side-chain placement could be beneficial in other systems. Thus, while the commercial availability of protected β^3 building blocks make them a good choice for backbone modification, the more synthetically challenging β^2 analogues are likely to be valuable in certain cases.

In situations where side chains make no important contacts related to folding or function, incorporation of β^{cyc}-residues like ACPC in place of acylic β^3-residues can improve helix folded stability.[21] Protein **6** is a previously reported variant of **2** in which rigidified β^{cyc} residues replace two β^3-residues.[12] In the GB1 tertiary fold, $\beta^3\rightarrow\beta^{cyc}$ substitution was structurally well accommodated but led to only a modest increase in folded stability. β^{cyc}-Residues limit energetically accessible backbone conformational space by incorporating an otherwise freely rotatable bond into a ring. Thus, cyclization of β-residues can be thought of in similar terms as another known strategy for helix stabilization: methylation of C_α in α-residues. C_α-Me-α-residues are strong helix promoters,[22] and the achiral variant Aib has previously been examined in protein contexts.[23] The achiral nature of Aib leads to no inherent preference for a left- vs. right-handed helix; however, the biological arrangement can be favoured with chiral C_α-Me-α-residues.[24]

The above precedents led us to investigate which is a superior means of stabilizing a helical fold through backbone rigidification: chiral C_α-Me-α-, achiral C_α-Me-α-, or β^{cyc}-residues. In order to probe this question experimentally, we prepared and characterized proteins **7**-**8**, following the methods described above (Fig. S2, Table S1). Protein **7** is a variant of **6** in which the two outer β^3-residues in the helix are replaced by Aib. In protein **8**, these two sites incorporate the chiral C_α-methylated analogue of Val. Collectively, the series **6**-**8** enable the quantification of the relative benefit to folded stability from helix rigidification by the β^{cyc}-residue ACPC, the achiral C_α-Me-α-residue Aib, and the chiral C_α-Me-α analogue of Val.

A crystal structure of **7** solved to 2.15 Å resolution showed that, as expected, C_α-Me-α-residues are well accommodated in a helix that also contains β^3-residues (Fig. 4). $\beta^3\rightarrow$Aib replacement stabilized the folded state considerably (Fig. 3), enhancing folding free energy by 1.7 kcal mol^{-1} in **7** vs. **2**. The stabilization was entirely enthalpic and partially offset by a surprisingly large unfavourable impact on folding entropy. Replacement of the Aib residues in **7** with the chiral C_α-Me-Val reduced the entropic penalty, suggesting it may be tied to the accessibility of left-handed helical dihedrals in the former leading to a greater number of conformational states in the unfolded ensemble. Unexpectedly, Aib→C_α-Me-Val replacement was also accompanied by an unfavourable impact on folding enthalpy, the origin of which is not clear. The net

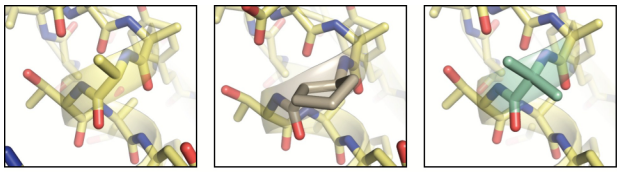


Fig. 4 Comparison of the corresponding α-, β^{cyc}-, and C_α-Me-α-residues at position 24 in the crystal structures of **1**, **6**, and **7** (PDB 2QMT, 4OZB, and 5HI1).

effect was a diminished overall folding free energy from $\beta^3 \rightarrow C_\alpha$-Me-Val vs. $\beta^3 \rightarrow$Aib substitution. The improved enthalphic stability of both C_α-Me-α variants **7-8** over all β-residue variants **2-6** suggests a more native-like helix geometry as an explanation for the improved folding stability. Comparing the folding free energy among **6**-**8** suggests that the relative ability to stabilize a helical fold in a protein context follows the trend Aib > C_α-Me-Val > ACPC.

Encouraged by the results for proteins **7** and **8**, we next replaced all four sites in the helix with C_α-Me-α-residues. Ala_{24} and Asn_{35} were replaced with Aib, while Lys_{28} and Lys_{31} were replaced with C_α-Me-Lys to produce GB1 analogue **9** (Fig. 1D). The synthesis and purification of **9** proved challenging, and the yield of purified material insufficient for full analysis of folding thermodynamics. A simple CD melt, however, showed **9** to be the most thermally stable among the mutants examined here (Fig. S7); its melting temperature (T_m) was only 2.5 °C lower than all-natural backbone **1**. The slightly lower T_m of **9** relative to **1** may result from a steric clash of the Lys_{31} C_α-Me group with Trp_{43}. Overall, these results further confirm chiral C_α-Me-α-residues are superior to β-residues (β^3, β^2, or β^{cyc}) as components of mixed-backbone helices; however, the synthetic challenge in their preparation and incorporation by solid-phase synthesis mitigates this utility somewhat.

In summary, we have reported here the comparative effect on folded structure and thermodynamics of five residue types as components of heterogeneous-backbone helix mimetics in a tertiary fold context. No significant difference was seen in helix folding propensity between β^3 vs. β^2 residues; however, the choice between them is important when a particular side chain is involved in tertiary contacts. In cases where side-chain functional groups are not critical, the C_α-Me-α-residue Aib was superior to C_α-Me-α-Val and the β^{cyc}-residue ACPC as a helix rigidifier. In all cases, C_α-Me-α-residues (chiral or achiral) proved better than β-residues in helix stabilization. The refined design rules reported here for α-helix modification, taken with recent efforts involving β-sheets, represent a significant advance toward the application of backbone engineering to the design of protein mimetics with both native-like tertiary folded structure and thermodynamic stability. The results are also interesting to consider in relation to the diversity of non-canonical amino acids isolated from meteorites and experiments that mimic prebiotic environments.[25]

Funding for this work was provided by the National Institutes of Health (GM107161 to W.S.H.). We thank Timothy Cunningham and Sunil Saxena for plasmids used in the expression of wild-type GB1 and the K31A mutant.

Notes and references

1 S. H. Gellman, *Acc. Chem. Res.*, 1998, **31**, 173-180.

2 (a) A. D. Bautista, C. J. Craig, E. A. Harker and A. Schepartz, *Curr. Opin. Chem. Biol.*, 2007, **11**, 685-692; (b) C. M. Goodman, S. Choi, S. Shandler and W. F. DeGrado, *Nat. Chem. Biol.*, 2007, **3**, 252-262; (c) G. Guichard and I. Huc, *Chem. Commun.*, 2011, **47**, 5933-5941; (d) W. S. Horne, *Expert Opin. Drug Discov.*, 2011, **6**, 1247-1262; (e) T. A. Martinek and F. Fulop, *Chem. Soc. Rev.*, 2012, **41**, 687-702; (f) J. Sun and R. N. Zuckermann, *ACS Nano*, 2013, **7**, 4715-4732.

3 (a) Z. Hegedüs, E. Wéber, É. Kriston-Pál, I. Makra, Á. Czibula, É. Monostori and T. A. Martinek, *J. Am. Chem. Soc.*, 2013, **135**, 16578-16584; (b) P. S. P. Wang, J. B. Nguyen and A. Schepartz, *J. Am. Chem. Soc.*, 2014, **136**, 6810-6813; (c) J. W. Checco, D. F. Kreitler, N. C. Thomas, D. G. Belair, N. J. Rettko, W. L. Murphy, K. T. Forest and S. H. Gellman, *Proc. Natl. Acad. Sci. USA*, 2015, **112**, 4552-4557.

4 (a) R. V. Mannige, T. K. Haxton, C. Proulx, E. J. Robertson, A. Battigelli, G. L. Butterfoss, R. N. Zuckermann and S. Whitelam, *Nature*, 2015, **526**, 415-420; (b) N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil and I. Huc, *Nature Chem.*, 2015, **7**, 334-341; (c) G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth and G. Guichard, *Nature Chem.*, 2015, **7**, 871-878.

5 Z. E. Reinert and W. S. Horne, *Org. Biomol. Chem.*, 2014, **12**, 8796-8802.

6 (a) L. M. Johnson and S. H. Gellman, *Methods Enzymol.*, 2013, **523**, 407-429; (b) H. M. Werner and W. S. Horne, *Curr. Opin. Chem. Biol.*, 2015, **28**, 75-82.

7 H. M. Werner, C. C. Cabalteja and W. S. Horne, *ChemBioChem*, in press (DOI: 10.1002/cbic.201500312).

8 R. W. Cheloha, A. Maeda, T. Dean, T. J. Gardella and S. H. Gellman, *Nat. Biotechnol.*, 2014, **32**, 653-655.

9 J. D. Pham, R. K. Spencer, K. H. Chen and J. S. Nowick, *J. Am. Chem. Soc.*, 2014, **136**, 12682-12690.

10 (a) G. A. Lengyel, R. C. Frank and W. S. Horne, *J. Am. Chem. Soc.*, 2011, **133**, 4246-4249; (b) G. A. Lengyel and W. S. Horne, *J. Am. Chem. Soc.*, 2012, **134**, 15906-15913; (c) G. A. Lengyel, G. A. Eddinger and W. S. Horne, *Org. Lett.*, 2013, **15**, 944-947; (d) G. A. Lengyel, Z. E. Reinert, B. D. Griffith and W. S. Horne, *Org. Biomol. Chem.*, 2014, **12**, 5375-5381.

11 Z. E. Reinert, G. A. Lengyel and W. S. Horne, *J. Am. Chem. Soc.*, 2013, **135**, 12528-12531.

12 Z. E. Reinert and W. S. Horne, *Chem. Sci.*, 2014, **5**, 3325-3330.

13 D. Seebach, S. Abele, K. Gademann, G. Guichard, T. Hintermann, B. Jaun, J. L. Matthews, J. V. Schreiber, L. Oberer, U. Hommel and H. Widmer, *Helv. Chim. Acta*, 1998, **81**, 932-982.

14 K. Möhle, R. Günther, M. Thormann, N. Sewald and H.-J. Hofmann, *Biopolymers*, 1999, **50**, 167-184.

15 Y. Chi, E. P. English, W. C. Pomerantz, W. S. Horne, L. A. Joyce, L. R. Alexander, W. S. Fleming, E. A. Hopkins and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 6050-6055.

16 (a) L. A. Carpino, H. G. Chao, S. Ghassemi, E. M. E. Mansour, C. Riemer, R. Warrass, D. Sadat-Aalaee, G. A. Truran, H. Imazumi, H. Wenschuh, M. Beyermann, M. Bienert, H. Shroff, F. Albericio, S. A. Triolo, N. A. Sole and S. A. Kates, *J. Org. Chem.*, 1995, **60**, 7718-7719; (b) C. Mazzini, J. Lebreton, V. Alphand and R. Furstoss, *J. Org. Chem.*, 1997, **62**, 5215-5218; (c) T. Ok, A. Jeon, J. Lee, J. H. Lim, C. S. Hong and H.-S. Lee, *J. Org. Chem.*, 2007, **72**, 7390-7393.

17 B. Kuhlman and D. P. Raleigh, *Protein Sci.*, 1998, **7**, 2405-2412.

18 G. J. Bartlett, A. Choudhary, R. T. Raines and D. N. Woolfson, *Nat. Chem. Biol.*, 2010, **6**, 615-620.

19 G. J. Bartlett, R. W. Newberry, B. VanVeller, R. T. Raines and D. N. Woolfson, *J. Am. Chem. Soc.*, 2013, **135**, 18682-18688.

20 H. L. Frericks Schmidt, L. J. Sperling, Y. G. Gao, B. J. Wylie, J. M. Boettcher, S. R. Wilson and C. M. Rienstra, *J. Phys. Chem. B*, 2007, **111**, 14362-14369.

21 W. S. Horne, J. L. Price and S. H. Gellman, *Proc. Natl. Acad. Sci. USA*, 2008, **105**, 9151-9156.

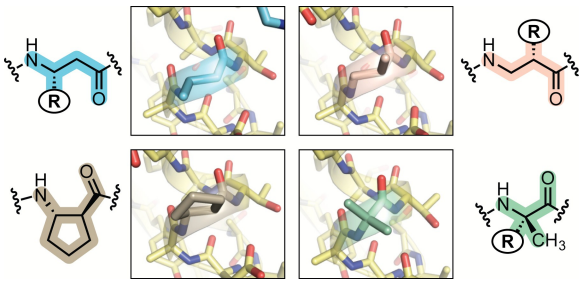
22 (a) I. L. Karle and P. Balaram, *Biochemistry*, 1990, **29**, 6747-6756; (b) G. R. Marshall, E. E. Hodgkin, D. A. Langs, G. D. Smith, J. Zabrocki and M. T. Leplawy, *Proc. Natl. Acad. Sci. USA*, 1990, **87**, 487-491; (c) K. T. O'Neil and W. F. DeGrado, *Science*, 1990, **250**, 646-651.

23 (a) V. De Filippis, F. De Antoni, M. Frigo, P. Polverino de Laureto and A. Fontana, *Biochemistry*, 1998, **37**, 1686-1696; (b) V. Y. Torbeev, H. Raghuraman, D. Hamelberg, M. Tonelli, W. M. Westler, E. Perozo and S. B. H. Kent, *Proc. Natl. Acad. Sci. USA*, 2011, **108**, 20982-20987.

24 M. Crisma and C. Toniolo, *Pept. Sci.*, 2015, **104**, 46-64.

25 (a) A. S. Burton, J. C. Stern, J. E. Elsila, D. P. Glavin and J. P. Dworkin, *Chem. Soc. Rev.*, 2012, **41**, 5459-5472; (b) J. L. Bada, *Chem. Soc. Rev.*, 2013, **42**, 2186-2196.

TOC Figure and Text



Structural and thermodynamic analysis of a family of synthetic proteins with heterogeneous backbones yields new insights into the ability of unnatural amino acids to be accommodated into α-helices.