Porter Goff, K., Nicol, D., Williams, C., Crump, M., Zieleniewski, F., Samphire, J., Baker, E., & Woolfson, D. (2019). Stabilizing and Understanding a Miniprotein by Rational Redesign. *Biochemistry*, *58*(28), 3060-3064. https://doi.org/10.1021/acs.biochem.9b00067

Stabilizing and Understanding a Miniprotein by Rational Redesign

Kathryn L. Porter Goff,[†] Debbie Nicol,[†,] Christopher Williams,[†,‡] Matthew P. Crump,[†,‡] Francis Zieleniewski,[†] Jennifer L. Samphire,[†] Emily G. Baker,[†] Derek N. Woolfson*[†,‡,§]

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

[‡] BrisSynBio, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol, BS8 1TQ, UK

[§] School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, UK

Supporting Information Placeholder

ABSTRACT: Miniproteins reduce the complexity of the protein-folding problem allowing systematic studies of contributions to protein folding and stabilization. Here, we describe the rational redesign of a miniprotein, PPα, comprising a polyproline-II helix, loop and α helix. The redesign provides a *de novo* framework to interrogate non-covalent interactions. Optimized PPα has significantly improved thermal stability with a midpoint unfolding temperature (T_M) of 51°C. Its NMR structure indicates a higher density of stabilizing non-covalent interactions than the parent peptide, specifically increased CH–π interactions. In part, we attribute this to improved long-range electrostatic interactions between the two helical elements. We probe further sequence-to-stability relationships in the miniprotein through a series of rational mutations.

The primary sequence of a protein determines its three-dimensional structure and function. Understanding sequence-to-structure relationships is known as the protein-folding problem.[1] Studying miniproteins reduces the complexity of the problem, allowing contributions to protein folding and stability to be dissected more easily.[2, 3] Prominent examples of miniproteins include: the Trp-cage,[4] βαβ motifs,[5] villin headpiece,[6] pancreatic peptides[7-10] and the Trp-Plexus.[11] The formation of many weak and cooperative non-covalent interactions is central to protein stability, allowing the loss of conformational entropy upon folding to be overcome. This is particularly important for miniproteins with fewer interactions possible.[12]

α Helices are common building blocks in proteins. They are often stabilized by tertiary and quaternary interactions, as exemplified in α-helical coiled coils, a ubiquitous folding motif comprising two or more α helices supercoiled around each other.[13] Most coiled-coil sequences have a seven-residue heptad repeat, ***abcdefg***, with hydrophobic residues at ***a*** and ***d***, and polar residues at the remaining sites. This pattern leads to the association of helices *via* "knobs-into-holes" packing interactions, where side chains of one helix dock into diamond-shaped holes of another.[14, 15] α Helices can also be stabilized by the interdigitation of proline (Pro, P) from polyproline-II helices between aromatic residues on α helices, which are analogous to knobs-into-holes packing.[2, 7, 16]

Elsewhere, we describe the fragment-based design of a 34-residue miniprotein, PPα.[17] This comprises a polyproline-II helix, loop and α helix adapted from natural proteins (Figure 1A,B).[7, 16] PPα is monomeric in aqueous solution and unfolds reversibly with a T_M of 39 ˚C. A solution NMR structure shows the stabilizing effect of Pro and tyrosine (Tyr, Y) interdigitation and associated CH–π interactions. In the latter, protons of C–H bonds interact with aromatic rings.[17-20] Here, we explore the optimization of PPα by rational redesign, revealing sequence-to-stability relationships for the miniprotein fold and delivering a *de novo* framework of significantly enhanced thermal stability.

In the design of α-helical coiled coils, charged residues often flank the hydrophobic core and form inter-helical Coulombic interactions.[2] We reasoned that PPα might be stabilized by introducing similar interactions between the ***2*** and ***3*** positions of the polyproline-II helix and the ***e*** and ***g*** positions of the α helix (Figure 1A,C).

There are two extreme possibilities for pairing charged residues in PPα: a lysine (Lys, K)-based polyproline-II helix and a glutamic acid (Glu, E)-based α helix or *vice versa* (PPα-KE and PPα-EK, respectively, Table 1). We optimized helix propensity for the solvent-exposed face of the α helix by placing alanine (Ala, A) at ***b*** and ***c*** and Lys at ***f*** in both designs.

Otherwise, the helix-helix interface was maintained, with Pro at ***1***, leucine (Leu, L) at ***a*** and Tyr at ***d***., along with the PPα loop. We adopted the coiled-coil heptad nomenclature for the α-helical sequence as it has a 7-residue repeat.

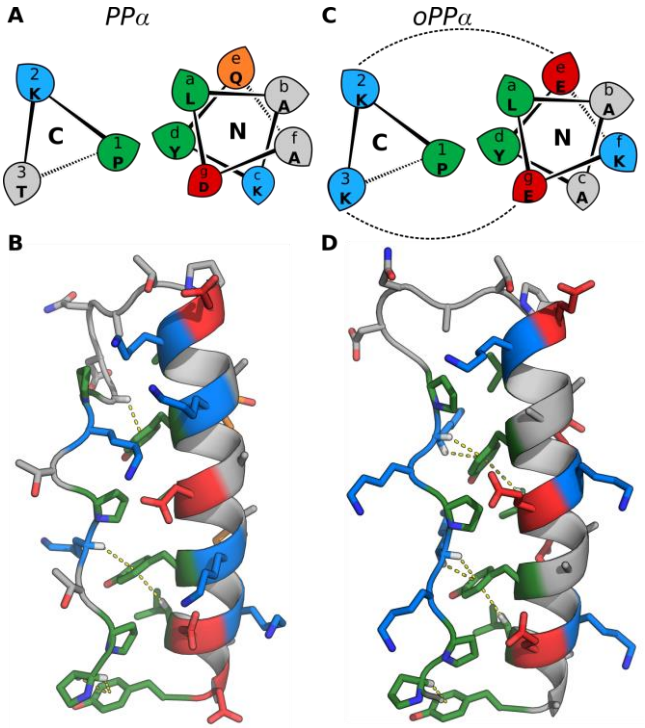

Figure 1. Helical wheel representations for the two helices in PPα (A) and PPα-KE (oPPα) (C). Leaf tips indicate the direction of the Cα-Cβ bond vectors. Lowercase letters on the α helix and numbers on the polyproline-II helix denote the helical register. C and N refer to the helix termini nearest the viewer. Dashed black lines: long-range interhelical electrostatic interactions. Representative models from the NMR ensembles of PPα (B, PDB entry 5LO2, model 14) and PPα-KE (D, PDB entry 6GWX, model 8). Dashed yellow lines: CH–π interactions. Key: Lys (blue), Glu/Asp (red), and Tyr/Pro/Leu (green), Gln (orange).

PPα-KE and PPα-EK were synthesized by solid-phase peptide synthesis, purified by reverse-phase HPLC, and confirmed by MALDI-TOF mass spectrometry (Figure S1A,B). Circular dichroism (CD) spectroscopy revealed both peptides were 50% α helical at 5 °C like PPα (Figure 2A). Sedimentation equilibrium analytical ultracentrifugation (AUC) showed that both peptides were monomeric (Figure S3A,B). PPα-KE had cooperative and reversible thermal unfolding transitions when followed by far- and near-UV CD (Figure S2U,V), and the T_M (51 °C) was substantially increased over PPα (39 °C). However, PPα-EK was less stable than PPα-KE by 10 °C (Figure 2B). We posit that the reduced stability of PPα-EK stems from the large charge on the α helix—+6 compared to -2 for PPα-KE—and different propensities of Glu and Lys for α and polyproline-II helices (Table S1,2).[21,22]

High-resolution NMR spectroscopy was used to determine the solution-phase structure of PPα-KE (Figure 1D), with 95.8% of the ^{1}H NMR resonances assigned. The core of the structure was well defined, while several of the solvent-exposed residues were more dynamic. The root-mean-square deviations (RMSDs) across the 20 structures of the ensemble were 0.52 ± 0.13 Å (backbone) and 1.05 ± 0.15 Å (all atom), comparable to those for the structure of PPα (Table S3).[17] A representative structure matched PPα with backbone and all-atom RMSDs of 0.5 Å and 0.9 Å, respectively.

The average distances between N_ζ of Lys and C_δ of Glu introduced in PPα-KE were 8.7 ± 1.8 Å and 11.5 ± 2.4 Å for ***2:e*** and ***3:g***, respectively. These fall outside of any accepted definition of salt-bridge interactions,[23] suggesting longer-range electrostatic interactions may stabilize the folded state of PPα-KE. To test this, the thermal stability of PPα-KE was measured at different salt concentrations in phosphate buffer (Figure S2C,D). With increased concentrations of NaCl from 0 to 700 mM the T_M fell from 61 to 57 °C (Figure S2W), indicative of an electrostatic component to thermostability.

As PPα is stabilized by CH–π interactions,[17] we searched for these in PPα-KE. This revealed 87 CH–π interactions across the ensemble (4.35 per structure), an increase over PPα (68 CH–π interactions; 3.4 per structure) (Figure 1B,D, Table S5 and Figure S8). A shift in CH donors between the two miniproteins was also observed. In PPα-KE fewer CH donors emanated from Pro residues (PPα-KE, 5; PPα, 22), but more arose from Leu (PPα-KE, 25; PPα, 4) and Lys residues (PPα-KE, 57; PPα, 15). On closer inspection, the side chains of Lys4 and Lys7 of PPα-KE lie across the faces of Tyr27 and Tyr20, respectively, leading to CH–π interactions between the C_α and C_β protons of the former and the aromatic rings of the latter. To test for CH–π interactions experimentally, we replaced all of the electron-rich Tyr residues in PPα-KE by more-electron-poor phenylalanine (Phe, F) giving PPα-KE-Phe (Table 1). As with PPα,[17] this reduced the T_M by approximately 20 °C (Figure S2E,F). This indicates that CH–π interactions are at play in PPα-KE in addition to van der Waals' forces.

Overall, we posit that the improved stability of PPα-KE arises from increased long-range electrostatic interactions between the polyproline-II and α helix,[24-27] which lead to an increased density of shorter-range CH–π interactions.

Tyr residues at the ***d*** sites of the 7-residue repeats of PPα and related natural folds are critical for folding and stability.[7, 16, 17] These residues form two vertices of 4-residue 'holes' defined by side chains at successive ***d***, ***a***, ***g***, and ***d*** sites, which accommodate Pro 'knobs' from the polyproline-II helix. In PPα and PPα-KE, the ***g*** site is fixed as aspartic acid (Asp, D) or Glu, respectively. The ***a*** sites are all Leu, which is the preferred residue in natural sequences.[7, 28] To explore possible substitutes for Leu at ***a***, we made mutants with all three ***a*** sites replaced by: β-branched, isoleucine (Ile, I) or valine (Val, V); charged,

Table 1. Designed peptides and summary of biophysical data.

Peptide	Sequence and helical register *321321321321321 efgabcd efgabcd efgabcd efgabcd efgabcd*	TM / °C	AUC *
Parent PPα	PPTKPTKP GDNAT PEKLAKY QADLAKY QKDLADY	39	0.9
PPα-KE/oPPα	PPKKPKKP GDNAT PEKLAAY EKELAAY EKELAAY	51	1.0
PPα-EK	PP**EE**P**EE**P GDNAT PEKLAAY **KK**K**K**LAAY **KK**K**K**LAAY	40	0.9
PPα-KE-Phe	PPKKPKKP GDNAT PEKLAA**F** EKELAA**F** EKELAA**F**	33	1.0
PPα-KE-E@a	PPKKPKKP GDNAT PEK**E**AAY EKE**E**AAY EKE**E**AAY	-	-
PPα-KE-A@a	PPKKPKKP GDNAT PEK**A**AAY EKE**A**AAY EKE**A**AAY	-	1.0
PPα-KE-K@a	PPKKPKKP GDNAT PEK**K**AAY EKE**K**AAY EKE**K**AAY	19	1.0
PPα-KE-I@a	PPKKPKKP GDNAT PEK**I**AAY EKE**I**AAY EKE**I**AAY	40	1.0
PPα-KE-V@a	PPKKPKKP GDNAT PEK**V**AAY EKE**V**AAY EKE**V**AAY	34	1.0
PPα-KE-A@g	PPKKPKKP GDNAT PEKLA**E**Y EK**A**LA**E**Y EK**A**LA**E**Y	21	1.0
PPα-KE-L@g	PPKKPKKP GDNAT PEKLA**E**Y EK**L**LA**E**Y EK**L**LA**E**Y	76	1.2
PPα-KE-a↔g	PPKKPKKP GDNAT PEKLAAY **AALEKE**Y **AALEKE**Y	19	1.1
oPPα-2	PPKKP GDNAT PEKLAAY EKELAAY	19	0.9
oPPα-4	PPKKPKKPKKP GDNAT PEKLAAY EKELAAY EKELAAY EKELAAY	66	1.0
oPPα-5	PPKKPKKPKKPKKP GDNAT PEKLAAY EKELAAY EKELAAY EKELAAY EKELAAY	72	1.0
oPPα-5-skip	PPKKPKKPPKKPKKP GDNAT PEKLAAY EKELAAY EKELAAY EKELAAY EKELAAY	68	1.0

Peptides were capped with *N*-terminal acetyl and *C*-terminal amide groups. *Oligomeric state (× monomer mass) as determined by AUC. Mutations to the PPα-KE/oPPα interface are shown in bold.

Glu or Lys; or Ala (Table 1).

AUC confirmed that these PPα-KE-X@***a*** mutants were monomeric in solution (Figure S3). However, CD spectroscopy revealed a range of stabilities. PPα-KE-E@***a*** was not folded, possibly due to the proximal glutamates at ***e*** if the α helix were folded. The stabilities of PPα-KE-A@***a*** and PPα-KE-K@***a*** were compromised compared to PPα-KE (Figure 2C,D). By contrast, both PPα-KE-I@***a*** and PPα-KE-V@***a*** were folded, albeit with reduced thermal stabilities compared with PPα-KE. This suggests that the extra steric bulk of the β carbon may hinder docking of Pro into the hole, although we note that α-helical propensities of both Ile and Val are also lower than that of Leu.[21]

We also probed the ***g*** position of the hole, initially *via* an Ala mutation to determine the contribution of this residue to stability. To maintain the introduced electrostatic interactions between the helices, Glu was shifted to the ***c*** position, although this may introduce favorable $K_i \rightarrow E_{i+4}$ interactions. The resulting PPα-KE-A@g was monomeric, but significantly less stable than PPα-KE (Figure S2I,J). Next, we mutated the ***g*** position to Leu, giving a fully hydrophobic hole. The thermal unfolding transition of PPα-KE-L@***g*** was broad, with a T_M of 76 °C (Figure S2K,L). However, the CD signal was concentration dependent, indicating aggregation (Figure S2M,N). AUC experiments returned a mass of 1.2 × monomer mass, suggesting PPα-KE-L@***g*** forms higher-order species, which could involve the wider hydrophobic interface. Our interests lie in fully monomeric miniproteins, so the ***g*** positions were not probed further.

As a final exploration of hole-residue preferences, we swapped the ***a*** and ***g*** residues (PPα-KE-***a*↔*g***). This tests an alternate helical register, LXXXYXX, as opposed to LXXYXXX in PPα-KE. Residues at ***e*** and ***c***, and ***b*** and ***f*** were swapped to avoid introducing charged interactions in the α helix of PPα-KE-***a*↔*g*** not present

in PPα-KE. PPα-KE-***a*↔*g*** was less folded and less thermally stable (T_M 19 °C) than PPα-KE. Thus, PPα-KE possesses the optimal arrangement of residues in the diamond-shaped hole.

The results of this mutagenesis study reveal that Leu at ***a***, Tyr at ***d*** and Glu at ***g*** provide the optimal 'hole' to accommodate Pro in PPα-like miniproteins. On this basis, we renamed PPα-KE as optimised-PPα, oPPα.

Using oPPα, we investigated the effect of helix length on stability. Variants were synthesised with 1-, 3- and 4-unit polyproline-II-helical repeats and corresponding 2-, 4- and 5- heptad α-helical repeats (Table 1, Figure S15). Like oPPα, which completes this series with two polyproline-II and three

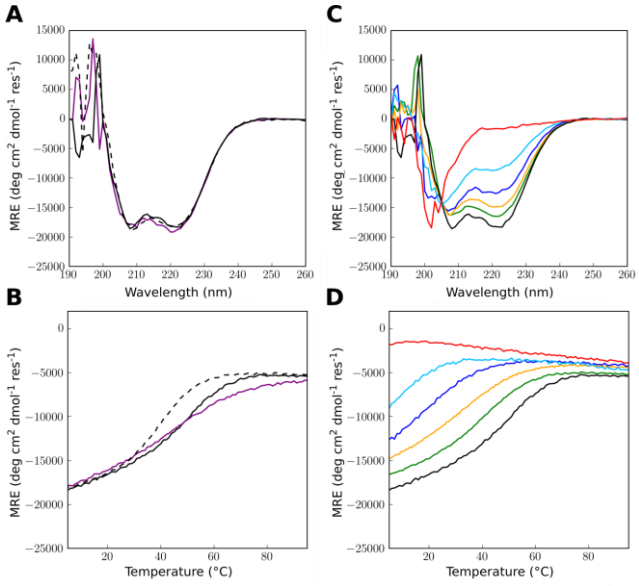

Figure 2. Folding and stability of PPα-KE and mutants. Far-UV CD spectra at 5 °C for: (A) parent PPα (black, dash), PPα-KE (black), and PPα-EK (purple); and (C) PPα-KE-E@a (red), PPα-KE-A@a (sky blue), PPα-KE-K@a (blue), PPα-KE-I@a (green), PPα-KE-V@a (orange), and PPα-KE (black). (B&D) Temperature dependence of the mean residue ellipticity (MRE) at 222 nm, with the key as for panels A&C. Conditions: 100 μM peptide, PBS, pH 7.4.

α-helical repeats, these peptides were monomers (Figure S3M,N,O). However, only oPPα, oPPα-4 and oPPα-5 were appreciably and stably folded (Figure 3), with the T_M increasing from 51 to 66 to 72 °C, respectively. As observed in α-helical coiled coils, folding and stability increases in a non-linear, cooperative manner.[29, 30]

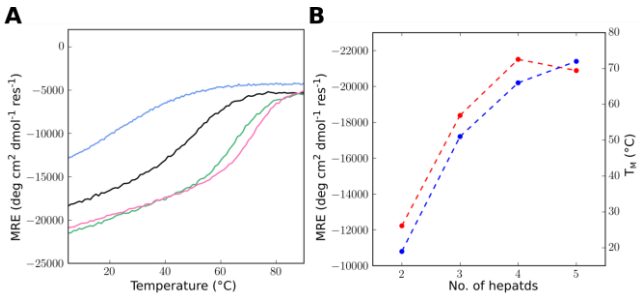

Figure 3. Folding and stability of oPPα variants of different lengths. (A) Thermal unfolding curves for oPPα-2 (blue), oPPα (black), oPPα-4 (green), and oPPα-5 (pink). (B) Plot of heptad length vs T_M (blue) and MRE_{222} at 5 °C (red). Conditions: 100 μM peptide, PBS, pH 7.4.

There is a periodicity mismatch between the 3- and 7- residue sequence repeats of polyproline-II and α helices: the former span ≈ 9.3 Å and the latter ≈ 10.5 Å. For oPPα-5 this could result in a ≈6 Å discrepancy over the lengths of the helices in the interface. The bacterial adhesin AgI/II from *S. mutans* overcomes this with Pro insertions that break the PXX repeat.[16] Therefore, we introduced an extra Pro to the middle of the polyproline-II repeat of oPPα-5. This peptide, oPPα-5-skip, was monomeric, folded and stable (Figure S2S,T) but had a slightly lower T_M than the parent. This suggests that the helix-helix interface of the oPPα fold is plastic and accommodates minor mismatches in periodicities, perhaps not possible for longer, fibrous assemblies.

In summary, a miniprotein fold has been optimized by rational protein redesign to give oPPα, a completely *de novo* framework with significantly enhanced thermal stability relative to the parent design. We propose that complementary long-range electrostatic interactions and increased shorter-range noncovalent interactions are responsible for the enhanced stability. The specific sequence-to-structure/stability relationships that we have explored should help guide further designs of PPα-like miniproteins. oPPα should be a useful addition to the toolkit of synthetic peptide building blocks for constructing more-complex systems through modular assembly.[31, 32]

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website.

Experimental details and characterization data.

AUTHOR INFORMATION

Corresponding Author

*d.n.woolfson@bristol.ac.uk

ORCID

Kathryn L. Porter Goff: 0000-0002-0533-7791
Debbie Nicol: 0000-0001-6160-659X
Christopher Williams: 0000-0001-5806-9842
Matthew P. Crump: 0000-0002-7868-5818
Francis Zieleniewski: 0000-0003-1529-8594
Jennifer L. Samphire: 0000-0002-5945-415X
Emily G. Baker: 0000-0002-1593-5525
Derek N. Woolfson: 0000-0002-0394-3202

Funding

KLPG, DN, JLS and FZ are funded by the EPSRC-funded Bristol Chemical Synthesis Centre for Doctoral Training (EP/G036764/1). EGB and DNW are supported by a European Research Council Advanced Grant (340764). CW and the 700 MHz NMR spectrometer are supported by BrisSynBio, a BBSRC/EPSRC-funded Synthetic Biology Research Centre (BB/L01386X1). DNW holds a Royal Society Wolfson Research Merit Award (WM140008). We thank the reviewers of the original manuscript for insightful comments.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Dill, K. A., Ozkan, S. B., Shell, M. S., and Weikl, T. R. (2008) The protein folding problem., *Annu. Rev. Biophys. 37*, 289-316.

(2) Baker, E. G., Bartlett, G. J., Porter Goff, K. L., and Woolfson, D. N. (2017) Miniprotein Design: Past, Present, and Prospects., *Acc. Chem. Res. 50*, 2085-2092.

(3) Gelman, H., and Gruebele, M. (2014) Fast protein folding kinetics., *Q. Rev. Biophys. 47*, 95-142.

(4) Neidigh, J. W., Fesinmeyer, R. M., and Andersen, N. H. (2002) Designing a 20-residue protein., *Nat. Struct. Biol. 9*, 425-430.

(5) Liang, H. H., Chen, H., Fan, K. Q., Wei, P., Guo, X. R., Jin, C. W., Zeng, C., Tang, C., and Lai, L. H. (2009) De novo design of a βαβ motif., *Angew. Chem. Int. Ed. 48*, 3301-3303.

(6) McKnight, C. J., Matsudaira, P. T., and Kim, P. S. (1997) NMR structure of the 35-residue villin headpiece subdomain., *Nat. Struct. Biol. 4*, 180-184.

(7) Blundell, T. L., Pitts, J. E., Tickle, I. J., Wood, S. P., and Wu, C. W. (1981) X-ray analysis (1. 4-Å resolution) of avian pancreatic polypeptide: Small globular protein hormone., *Proc. Natl. Acad. Sci. USA 78*, 4175-4179.

(8) Chin, J. W., and Schepartz, A. (2001) Concerted evolution of structure and function in a miniature protein., *J. Am. Chem. Soc. 123*, 2929-2930.

(9) Hodges, A. M., and Schepartz, A. (2007) Engineering a monomeric miniature protein., *J. Am. Chem. Soc. 129*, 11024-11025.

(10) Zondlo, N. J., and Schepartz, A. (1999) Highly specific DNA recognition by a designed miniature protein., *J. Am. Chem. Soc. 121*, 6938-6939.

(11) Craven, T. W., Cho, M. K., Traaseth, N. J., Bonneau, R., and Kirshenbaum, K. (2016) A miniature protein stabilized by a cation-π interaction network., *J. Am. Chem. Soc. 138*, 1543-1550.

(12) Pace, C. N., Scholtz, J. M., and Grimsley, G. R. (2014) Forces stabilizing proteins., *FEBS Lett. 588*, 2177-2184.

(13) Woolfson, D. N. (2017) Coiled-coil design: updated and upgraded., *Subcell. Biochem. 82*, 35-61.

(14) Crick, F. H. (1953) The packing of alpha-helices: simple coiled coils., *Acta Crystallogr. 6*, 689-697.

(15) Walshaw, J., and Woolfson, D. N. (2001) SOCKET: a program for identifying and analysing coiled-coil motifs within protein structures., *J. Mol. Biol. 307*, 1427-1450.

(16) Larson, M. R., Rajashankar, K. R., Patel, M. H., Robinette, R. A., Crowley, P. J., Michalek, S., Brady, L. J., and Deivanayagam, C. (2010) Elongated fibrillar structure of a streptococcal adhesin assembled by the high-affinity association of α- and PPII-helices, *Proc. Natl. Acad. Sci. USA 107*, 5983-5988.

(17) Baker, E. G., Williams, C., Hudson, K. L., Bartlett, G. J., Heal, J. W., Porter Goff, K. L., Sessions, R. B., Crump, M. P., and Woolfson, D. N. (2017) Engineering protein stability with atomic precision in a monomeric miniprotein., *Nat. Chem. Biol. 13*, 764-770.

(18) Kang, Y. K., and Byun, B. J. (2012) Strength of CH···π interactions in the C-terminal subdomain of villin headpiece., *Biopolymers 97*, 778-788.

(19) Nishio, M., Umezawa, Y., Fantini, J., Weiss, M. S., and Chakrabarti, P. (2014) CH-π hydrogen bonds in biological macromolecules., *Phys. Chem. Chem. Phys. 16*, 12648-12683.

(20) Zondlo, N. J. (2013) Aromatic-proline interactions: electronically tunable CH/π interactions, *Acc. Chem. Res. 46*, 1039-1049.

(21) Pace, C. N., and Scholtz, J. M. (1998) A helix propensity scale based on experimental studies of peptides and proteins., *Biophys. J. 75*, 422-427.

(22) Brown, A. M., and Zondlo, N. J. (2012) A Propensity Scale for Type II Polyproline Helices (PPII): Aromatic Amino Acids in Proline-Rich Sequences Strongly Disfavor PPII Due to Proline-Aromatic Interactions., *Biochemistry 51*, 5041-5051.

(23) Barlow, D. J., and Thornton, J. M. (1983) Ion-Pairs in Proteins., *J. Mol. Biol. 168*, 867-885.

(24) Schreiber, G., and Fersht, A. R. (1996) Rapid, electrostatically assisted association of proteins., *Nat. Struct. Biol. 3*, 427-431.

(25) Schreiber, G., Haran, G., and Zhou, H. X. (2009) Fundamental aspects of protein-protein association kinetics., *Chem. Rev. 109*, 839-860.

(26) Hemsath, L., Dvorsky, R., Fiegen, D., Carlier, M. F., and Ahmadian, M. R. (2005) An electrostatic steering mechanism of Cdc42 recognition by Wiskott-Aldrich syndrome proteins., *Mol. Cell 20*, 313-324.

(27) Blochliger, N., Xu, M., and Caflisch, A. (2015) Peptide binding to a PDZ domain by electrostatic steering via nonnative salt bridges., *Biophys. J. 108*, 2362-2370.

(28) Li, X., Sutcliffe, M. J., Schwartz, T. W., and Dobson, C. M. (1992) Sequence-specific ^{1}H NMR assignments and solution structure of bovine pancreatic polypeptide., *Biochemistry 31*, 1245-1253.

(29) Su, J. Y., Hodges, R. S., and Kay, C. M. (1994) Effect of chain-length on the formation and stability of synthetic α-helical coiled coils., *Biochemistry 33*, 15501-15510.

(30) Litowski, J. R., and Hodges, R. S. (2001) Designing heterodimeric two-stranded α-helical coiled-coils: the effect of chain length on protein folding, stability and specificity., *J. Pept. Res. 58*, 477-492.

(31) Fletcher, J. M., Boyle, A. L., Bruning, M., Bartlett, G. J., Vincent, T. L., Zaccai, N. R., Armstrong, C. T., Bromley, E. H. C., Booth, P. J., Brady, R. L., Thomson, A. R., and Woolfson, D. N. (2012) A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology., *ACS Synth. Biol. 1*, 240-250.

(32) Fletcher, J. M., Harniman, R. L., Barnes, F. R. H., Boyle, A. L., Collins, A., Mantell, J., Sharp, T. H., Antognozzi, M., Booth, P. J., Linden, N., Miles, M. J., Sessions, R. B., Verkade, P., and Woolfson, D. N. (2013) Self-assembling cages from coiled-coil peptide modules., *Science 340*, 595-599.