EDGE ARTICLE

Coiled coils 9-to-5: rational *de novo* design of α-helical barrels with tunable oligomeric states†

William M. Dawson, [ID] *[a] Freddie J. O. Martin,[a] Guto G. Rhys [ID] [ab]
Kathryn L. Shelley, [ID] [ac] R. Leo Brady [ID] [c] and Derek N. Woolfson [ID] *[acd]

The rational design of linear peptides that assemble controllably and predictably in water is challenging. Short sequences must encode unique target structures and avoid alternative states. However, the non-covalent forces that stabilize and discriminate between states are weak. Nonetheless, for α-helical coiled-coil assemblies considerable progress has been made in rational *de novo* design. In these, sequence repeats of nominally hydrophobic (*h*) and polar (*p*) residues, *hpphppp*, direct the assembly of amphipathic helices into dimeric to tetrameric bundles. Expanding this pattern to *hpphhph* can produce larger α-helical barrels. Here, we show that pentameric to nonameric barrels are accessed by varying the residue at one of the *h* sites. In peptides with four L/I–K–E–I–A–x–Z repeats, decreasing the size of Z from threonine to serine to alanine to glycine gives progressively larger oligomers. X-ray crystal structures of the resulting α-helical barrels rationalize this: side chains at Z point directly into the helical interfaces, and smaller residues allow closer helix contacts and larger assemblies.

Introduction

Most commonly, natural coiled-coil (CC) peptides form dimers, trimers and tetramers with consolidated hydrophobic cores.[1,2] Control over oligomeric state is achieved by different combinations of mainly isoleucine (Ile, I) and leucine (Leu, L) residues in the core.[3,4] Larger oligomers are rare in nature.[5,6] Interestingly, some of these larger structures are α-helical barrels (αHBs) with accessible central channels making them appealing scaffolds for functional design, *e.g.* binding, catalysis, delivery, and transport.[7–13] Variants of a natural dimer and *de novo* tetramer serendipitously form heptameric and hexameric αHBs, respectively.[14,15] To automate the design of αHBs, we have developed computational-design tools to deliver 5-, 6- or 7-helix αHBs.[16] These oligomers can be rationalized retrospectively to advance further sequence-to-structure relationships for CC design.

Most αHBs are Type-2 CCs based on *hpphhph* sequence repeats, labelled *abcdefg* (Fig. 1).[17] Typically, αHBs have L/IxxIAxA repeats; *i.e.*, *a* = Leu or Ile and *d* = Ile. β-Branched

residues at *d* are particularly important for maintaining open αHBs.[18] The hexameric and heptameric αHBs (CC-Hex2 and CC-Hept, systematically named CC-Type2-$(S_gL_aI_d)_4$ and CC-Type2-$(A_gL_aI_d)_4$) have *a* = Leu, *d* = Ile and *e* = alanine (Ala, A), but differ at *g*, which is Ala in the heptamer and the slightly larger serine (Ser, S) in the hexamer (Table 1). Another variant, CC-Pent (CC-Type2-$(I_gL_aI_dE_e)_4$) has Ile at *g*, although it differs from the other examples having *e* = glutamic acid (Glu, E).[16] A second series with all-Ile cores (*a* = *d* = Ile) has been characterized.[16,18] In these, another hexamer, CC-Hex3 (CC-Type2-$(S_gI_aI_d)_4$), follows the design rules above (Table 1);[16] and a peptide with Ala at *g* (CC-Type2-$(A_gI_aI_d)_4$) forms an octamer when crystallized in the presence of isopropanol.[18]

These previously described αHBs show a trend: increasing the size of side chains at *g* decreases the oligomer state formed. Therefore, we reasoned that a series of minimal changes, solely

[a]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK. E-mail: w.dawson@bristol.ac.uk; d.n.woolfson@bristol.ac.uk

[b]Department of Chemistry, University of Bayreuth, Universitätsstraße 30, 95447 Bayreuth, Germany

[c]School of Biochemistry, University of Bristol, Biomedical Sciences Building, University Walk, Bristol BS8 1TD, UK

[d]Bristol BioDesign Institute, University of Bristol, Life Sciences Building, Tyndall Avenue, Bristol BS8 1TQ, UK

† Electronic supplementary information (ESI) available: Methods and ESI data. See DOI: 10.1039/d1sc00460c
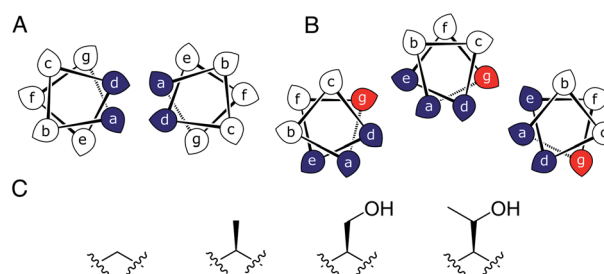
Fig. 1 (A and B) Type-N (A) and Type-2 (B) CC interfaces. The *g* position is highlighted in the Type-2 interface. (C) Side-chain structures of glycine, alanine, serine and threonine.

Table 1 Sequences of *de novo* αHBs and summary of biophysical characterization[a]

Heptad repeat (***abcdefg***)	Systematic name	BUDE oligomer score	DPH K_D (μM)	XRD αHB oligomeric state(s)	AUC SV	AUC SE
LKEIAxT	CC-Type2-$(T_gL_aI_d)_4$	5	6.8 ± 1.3	5	4.8	4.8
IKEIAxT	CC-Type2-$(T_gI_aI_d)_4$	5	0.84 ± 0.19	5	5.4	5.2
LKEIAxS[16]	CC-Type2-$(S_gL_aI_d)_4$	6	1.6 ± 0.2	6	5.7	6.5
IKEIAxS[16]	CC-Type2-$(S_gI_aI_d)_4$	6	3.8 ± 0.8	6	6.2	6.1
LKEIAxA[16]	CC-Type2-$(A_gL_aI_d)_4$	7	1.3 ± 0.3	7	6.9	7.0
IKEIAxA[16,18]	CC-Type2-$(A_gI_aI_d)_4$	6	2.2 ± 0.3	8	6.1	5.7
LKEIAxG	CC-Type2-$(G_gL_aI_d)_4$	9	0.076 ± 0.0065	9 (& collapsed 6)	6.5	6.3
IKEIAxG	CC-Type2-$(G_gI_aI_d)_4$	8	0.66 ± 0.12	6 & 7	5.0	5.1

[a] BUDE: Bristol University Docking Engine. DPH: diphenylhexatriene. XRD: X-ray diffraction. AUC: analytical ultracentrifugation. SV: sedimentation velocity. SE: sedimentation equilibrium.

at the ***g*** position—*i.e.*, at Z in L/IxxIAxZ repeats—might direct oligomeric state systematically and reliably.

Specifically, we considered the addition of a single heavy atom (C or O) to the side chain through the series glycine (Gly, G), Ala, Ser and threonine (Thr, T) (Fig. 1C). Our aim was to make a series of peptides with minimal changes to give a robust family of αHBs with tuneable oligomeric states for applications in synthetic biology and protein design.[13] N.B. To maximise further uses of the peptide assemblies, we avoided larger hydrophobic residues at ***g***, *e.g.* valine, as our experience is that these can lead to difficulties in purification. We also avoided cysteine because of potential issues with redox chemistry and to reserve it for introducing function in follow-on studies.[11,19,20]

Results and discussion

To supplement foregoing designs with Ser and Ala at ***g***, we designed four peptides with Gly or Thr at this position; *i.e.*, CC-Type2-$(G_gL_aI_d)_4$, CC-Type2-$(G_gI_aI_d)_4$, CC-Type2-$(T_gL_aI_d)_4$ and CC-Type2-$(T_gI_aI_d)_4$, Tables 1 and S2.† For simplicity, we refer to these as Gly@***g*** and Thr@***g*** peptides. Our hypothesis was that these should direct larger and smaller oligomers, respectively.

We built and optimized parametric models for Gly@***g*** and Thr@***g*** in ISAMBARD.[21] The sequences were modelled as parallel αHBs of oligomer state 5 to 10, and scored using BUDE[22,23] (Table S3†). For the historical designs, the most-favoured states were indeed those observed experimentally,[16] the exception being CC-Type2-$(A_gI_aI_d)_4$, which predicted as a hexamer as observed in solution, but crystallized as an octamer.[18] Encouragingly, the new Thr@***g*** sequences consistently scored best as pentamers: for the ***a*** = ***d*** = Ile variant the pentameric assembly was favoured outright; while the ***a*** = Leu, ***d*** = Ile variant scored equally well as pentamer or hexamer. Conversely, both Gly variants scored more favourably as larger oligomeric states: the ***a*** = ***d*** = Ile variant as an octamer; and ***a*** = Leu, ***d*** = Ile as a nonamer. Although, there was less discrimination between models for the Gly@***g*** sequences than for Thr@***g***, Fig. S1.† Thus, modelling supports the hypothesis that oligomeric states of Type-2 CCs can be tuned by side chains at ***g***.

The Gly@***g*** and Thr@***g*** peptides were synthesized, purified by HPLC, and confirmed by mass spectrometry (Fig. S2–S8†).

Circular dichroism (CD) spectroscopy indicated that all four peptides were α helical at low μM concentrations, Fig. 2A. CD spectra recorded at increasing temperatures showed that both Thr@***g*** variants were hyperthermostable, Fig. 2B. Whereas, the Gly@***g*** variant with ***a*** = Leu, ***d*** = Ile had the beginnings of a thermal unfolding curve consistent with the anticipated destabilizing effect of Gly on α-helical structures.[24–26] The Gly@***g*** variant with ***a*** = ***d*** = Ile unfolded at 38 °C and did not fully refold on cooling. Indeed, peptide sequences with ***a*** = ***d*** = Ile are more unstable than their ***a*** = Leu, ***d*** = Ile counterparts. Whilst β-branched residues like Ile clearly impart oligomer-
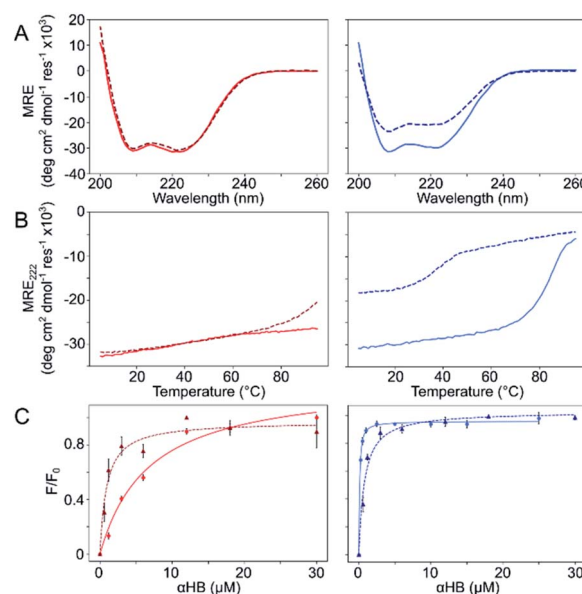

Fig. 2 Solution-phase biophysical characterization of Gly@***g*** (right) and Thr@***g*** (left) peptides. (A) CD spectra at 20 °C. (B) Thermal denaturation following the CD signal at 222 nm (MRE_{222}). (C) Saturation binding curves with DPH. Key: CC-Type2-$(T_gL_aI_d)_4$ (red), CC-Type2-$(T_gI_aI_d)_4$ (dark red, dashed), CC-Type2-$(G_gL_aI_d)_4$ (blue) and CC-Type2-$(G_gI_aI_d)_4$ (navy, dashed). Conditions: (A and B) 10 μM peptide. (C) 0–300 μM peptide, 1 μM DPH, 5% v/v DMSO. All experiments were performed in phosphate buffered saline at pH 7.4 (PBS; 8.2 mM Na_2HPO_4, 1.8 mM KH_2PO_4, 137 mM NaCl, 2.4 mM KCl).
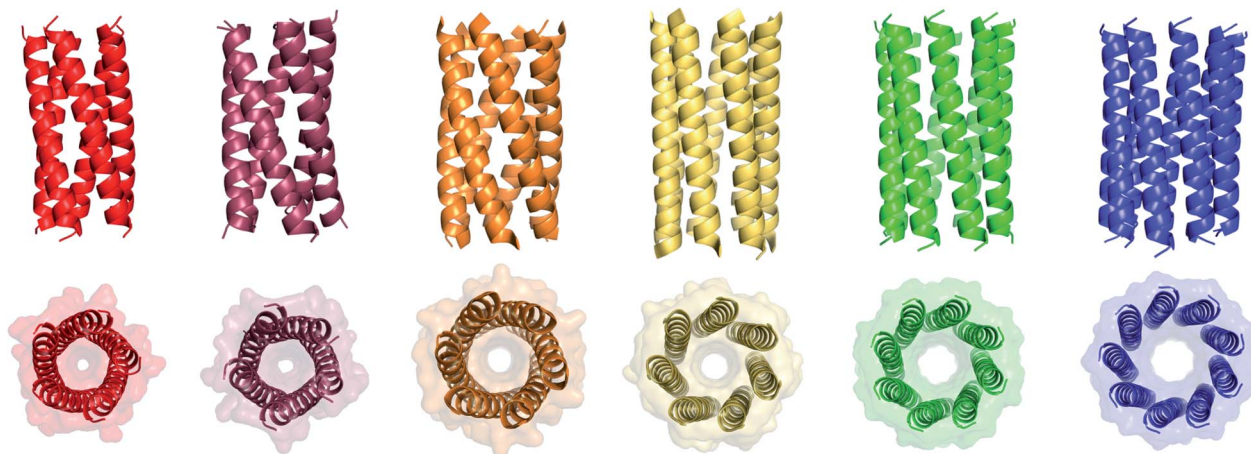
Fig. 3 X-ray crystal structures for pentameric through to nonameric αHBs. (Left to right) CC-Type2-$(T_gL_aI_d)_4$ (red, PDB 7BAS), CC-Type2-$(T_{gl}a_Id)_4$ (dark red, PDB 7BAU), CC-Type2-$(S_gL_aI_d)_4$ (orange, PDB 4PN9), CC-Type2-$(A_gL_aI_d)_4$ (yellow, PDB 4PNA), CC-Type2-$(A_{gl}a_Id)_4$ (green, PDB 6G67) and CC-Type2-$(G_gL_aI_d)_4$ (blue, PDB 7BIM).‡[1]

state specificity in CC systems,[3,4,18] we posit that this may be at the expense of some loss in overall thermal stability.

Next, dye-binding assays were used to assess the presence of accessible channels, Fig. 2C and S9† and Table 1. Binding of diphenylhexatriene (DPH) is a reliable indicator that CC peptides form αHBs in solution with a 1 : 1 correlation with X-ray crystal structures.[16,18] All four new peptides bound DPH consistent with the formation of αHBs.

We determined X-ray protein crystal structures for all Gly@g and Thr@g variants. As predicted, the Thr@g peptides formed parallel pentamers (Fig. 3 and Tables S4 and S5†). Comparison with the foregoing computationally designed pentamer, CC-Type2-$(I_gL_aI_dE_e)_4$,[16] revealed similar CC parameters for the structures. Significantly, Thr@g with $a = d = $ Ile has a wider channel of ≈ 9 Å than previous designs (≈ 7 Å), which increases the scope to functionalize this variant.

The Gly@g peptide with $a = $ Leu, $d = $ Ile crystallized in two forms. Gratifyingly, one was an all-parallel nonamer, which is a new αHB with an exceptionally large channel of diameter $\approx 9.5–11.5$ Å, Fig. 3 (Table S4†). Attempts to model solvent into density observed in the channel were inconsistent. Therefore, representative solvent molecules were included where they matched the density; though these did not make any stabilizing contacts with protein. Although there are natural nonameric protein assemblies,[27,28] this is the first stand-alone, water-soluble α-helical CC of this size. The second crystal form revealed a collapsed C2-symmetric 6-helix bundle (Fig. 4A and B, Table S4†). This was surprising, as the β-branched Ile residues at d would be expected to prevent collapse.[18] We posit that the small size of Gly relaxes this design rule allowing access to other nearby regions of the CC free-energy landscape.[29]

Gly@g with $a = d = $ Ile also crystallized in two forms (Fig. 4C and D, Table S5†), However, both solved as αHBs; namely, a hexamer and a heptamer. Thus, β-branched Ile at both a and d maintains the open assembly[18] even with Gly residues at g. Although a larger oligomer was predicted *in silico*, the

calculated internal energies for Gly@g were similar for the different oligomers (Fig. S1†).

Because of the apparent structural duality with Gly@g, we examined the oligomeric states of all peptides in solution by analytical ultracentrifugation (AUC; Table 1, Fig. S10–S13†). Both sedimentation velocity (SV) and sedimentation equilibrium (SE) measurements for the Thr@g variants returned pentameric molecular weights consistent with the X-ray crystal structures. Indeed, for the pentamers through heptamers for these and previous designs, the correlation between the solution-phase and crystal-state oligomers was good, Table 1.

However, where larger oligomers (octamer[18] and nonamer) were observed in crystals, smaller oligomers were consistently observed in solution, Table 1. This suggests that the solution
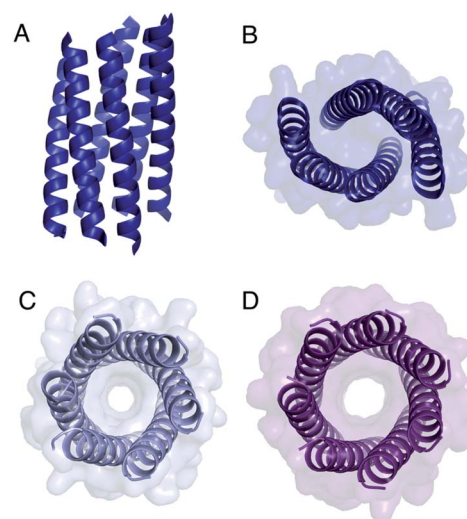

Fig. 4 Structures of Gly@g variants. (A and B) Orthogonal views of the collapsed hexameric-form of CC-Type2-$(G_gL_aI_d)_4$ (PDB 7A1T). (C and D) The hexameric (C) and heptameric (D) form of CC-Type2-$(G_gI_aI_d)_4$ (PDB 7BAT & 7BAW).‡
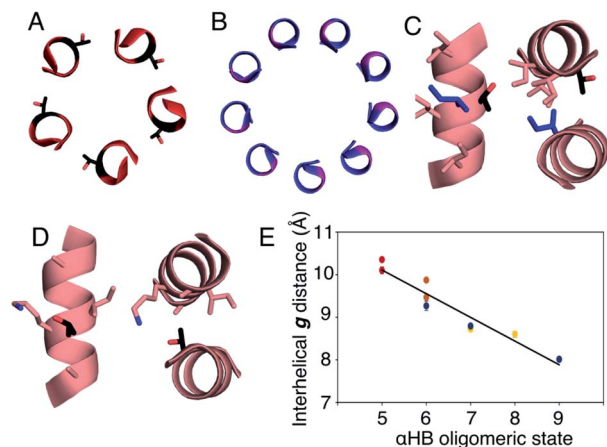
Fig. 5 Analysis of pentameric to nonameric αHBs. (A and B) Cross-sections through CC-Type2-$(T_gI_aI_d)_4$ (red) and CC-Type2-$(G_gL_aI_d)_4$ (blue) showing the geometry of residues at g (black and pink, respectively). (C and D) Knobs-into-holes interaction showing how these residues (black) contribute to a d'_{-1}-g'_{-1}-a'-d' hole when Ile is the knob residue (blue) (C); and when the Thr at the g position is a knob residue (D). (E) Interhelical distances in the αHBs with Thr (red), Ser (orange), Ala (yellow) and Gly (blue) at g. Errors are the standard deviation of measurements from the central heptads of each structure.

states are the dominant species, and that the higher oligomers observed by X-ray crystallography are meta-stable. This is consistent with smaller oligomers being entropically favoured. In addition, the crystallization conditions for the octamer and nonamer contained isopropanol. This increases the hydrophobicity of the bulk solvent, which potentially supports larger, hydrophobic pores that would otherwise be energetically unfavourable.[29] Nevertheless, these are legitimate states to consider as they are clearly accessible on the CC free-energy landscape.

Summarizing these data, Type-2 CC peptides with sequence repeats LppIApZ and Z = Thr, Ser, or Ala, form pentameric, hexameric, and heptameric αHBs, respectively. Adding Gly to the series accesses a nonameric open barrel, but only in the crystal state. Similarly, an IppIApA peptide forms an octamer in the crystal state.[18] These X-ray crystal structures enabled us to examine the structural transitions in detail.

In all of the structures, the a and d residues contribute both to the lumens and to interactions between neighbouring helices. SOCKET[30,31] analysis revealed that residues at a form knobs that fit into holes made by d'_{-1}-g'_{-1}-a'-d' of a neighbouring helix.[32] Thus, these knobs are complemented by interactions formed by the residues at g'_{-1} varied herein. The Cα-to-Cβ bond vectors of side chains at g point directly towards the adjacent helix, Fig. 5A–D. In classical CCs, this is called perpendicular packing, and it restricts how close the helices can approach.[3,32,33] Thus, mutations at g might be expected to influence the quaternary structure. The changes made in the

Gly→Ala→Ser→Thr series progressively add a single heavy atom to that side chain: Gly (0 heavy atoms)→Ala (1)→Ser (2)→Thr (3), Fig. 1C. It is gratifying, but still surprising, that this leads to unitary changes in oligomer state, at least for the Ala, Ser, and Thr variants. This is manifest in adjacent helix–helix distances through the series, Fig. 5E. The average distance increases from 8.0 ± 0.1 Å in the Gly@g nonamer to 10.4 ± 0.1 Å in the Thr@g pentamer. Thus, through the series, neighbouring helices are pushed apart effectively expelling helices from the assembly.

The judicious placement of Gly may prove useful in designing αHBs to unlock previously unseen architectures. However, using Gly presents challenges that must be met to allow its full exploitation. The first challenge is incorporating multistate design into αHBs, *i.e.* considering multiple conformations and/or assemblies that may become accessible in the CC free-energy landscape.[29] This approach is being applied to other systems.[34–36] It is tractable to model many possible Type-2 αHBs to direct computational design.[16] However, this becomes difficult with increasing off-target states, *e.g.* collapsed and anti-parallel structures. The second challenge is to stabilize the larger, but clearly accessible, oligomer states in solution. One possibility would be to introduce networks of polar residues to reduce the penalty of all-hydrophobic channels.[37] The hexameric and heptameric assemblies of CC-Type2-$(G_gI_aI_d)_4$ imply that, whilst Gly@g is necessary to access larger states, it is not the only factor that dictates oligomer state of the helical assembly. *De novo* αHBs have proved useful in functional protein design.[11–13] Reliably accessing scaffolds with significantly larger pores systematically and with minimal changes in primary sequence, would expand the scope for this and future applications.

Conclusions

We have combined rational design, computational modelling, and structural biology for a series of α-helical barrels (αHBs) with mutations from Gly→Ala→Ser→Thr at all g sites of a coiled-coil (CC) sequence repeat. Minimal and stepwise changes in size of the residue at these sites, combined with Leu/Ile at a and Ile at d, control the oligomeric state of the assembly. This expands the range of αHBs that can be designed systematically from pentamer (with Thr at g) to a nonamer (with Gly at g). Inspection of X-ray crystal structures rationalizes the role of side-chain bulk at g in dictating inter-helical packing distance, angles, and, thus, oligomeric state. CC-Type2-$(G_gL_aI_d)_4$ is the first example of a stand-alone nonameric CC. However, it appears that high oligomeric states (such as 8 and 9) are on the edge of what is possible for such 'Type-2' CC sequences as they are not favoured in solution.[16,18,38–40] Nonetheless, the X-ray crystal structures show that they are accessible. The rarity of such assemblies in nature[6,17,27,28] and their potential as scaffolds for functional *de novo* design makes these large αHBs tantalizing targets for design. For example, *de novo* designed and engineered αHBs have already proven useful in constructing peptide nanotubes,[8,9,12] peptide-based switches,[10,29] a rudimentary catalyst,[11] membrane-spanning ion-channels[41] and small-molecule receptors.[13]

‡ X-ray crystal structures for CC-Type2-$(T_gL_aI_d)_4$-W19BrPhe, CC-Type2-$(T_gI_aI_d)_4$-W19BrPhe, CC-Type2-$(G_gL_aI_d)_4$, CC-Type2-$(G_gL_aI_d)_4$-W19BrPhe and CC-Type2-$(G_gI_aI_d)_4$ are available from the Protein Data Bank. Accession codes: 7A1T, 7BAS, 7BAT, 7BAU, 7BAV, 7BAW & 7BIM.

Author contributions

WMD and DNW designed the study. WMD performed the biophysical characterisation of all peptides. WMD and FJOM crystallised the peptides. FJOM and GGR collected X-ray crystal data. FJOM, GGR, KLS and RLB determined the peptide crystal structures of the peptides. WMD and DNW wrote the manuscript. All authors read and contributed to the preparation of the manuscript.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

Notes and references

1 A. N. Lupas and M. Gruber, *Adv. Protein Chem.*, 2005, **70**, 37–38.

2 A. N. Lupas and J. Bassler, *Trends Biochem. Sci.*, 2017, **42**, 130–140.

3 P. B. Harbury, T. Zhang, P. S. Kim and T. Alber, *Science*, 1993, **262**, 1401–1407.

4 J. M. Fletcher, A. L. Boyle, M. Bruning, G. J. Bartlett, T. L. Vincent, N. R. Zaccai, C. T. Armstrong, E. H. C. Bromley, P. J. Booth, R. L. Brady, A. R. Thomson and D. N. Woolfson, *ACS Synth. Biol.*, 2012, **1**, 240–250.

5 D. N. Woolfson, G. J. Bartlett, A. J. Burton, J. W. Heal, A. Niitsu, A. R. Thomson and C. W. Wood, *Curr. Opin. Struct. Biol.*, 2015, **33**, 16–26.

6 A. Niitsu, J. W. Heal, K. Fauland, A. R. Thomson and D. N. Woolfson, *Philos. Trans. R. Soc., B*, 2017, 372.

7 J. Hume, J. Sun, R. Jacquet, P. D. Renfrew, J. A. Martin, R. Bonneau, M. L. Gilchrist and J. K. Montclare, *Biomacromolecules*, 2014, **15**, 3503–3510.

8 C. Xu, R. Liu, A. K. Mehta, R. C. Guerrero-Ferreira, E. R. Wright, S. Dunin-Horkawicz, K. Morris, L. C. Serpell, X. Zuo, J. S. Wall and V. P. Conticello, *J. Am. Chem. Soc.*, 2013, **135**, 15565–15578.

9 N. L. Ing, R. K. Spencer, S. H. Luong, H. D. Nguyen and A. I. Hochbaum, *ACS Nano*, 2018, **12**, 2652–2661.

10 R. Lizatović, O. Aurelius, O. Stenström, T. Drakenberg, M. Akke, D. T. Logan and I. André, *Structure*, 2016, **24**, 946–955.

11 A. J. Burton, A. R. Thomson, W. M. Dawson, R. L. Brady and D. N. Woolfson, *Nat. Chem.*, 2016, **8**, 837–844.

12 N. C. Burgess, T. H. Sharp, F. Thomas, C. W. Wood, A. R. Thomson, N. R. Zaccai, R. L. Brady, L. C. Serpell and D. N. Woolfson, *J. Am. Chem. Soc.*, 2015, **137**, 10554–10562.

13 F. Thomas, W. M. Dawson, E. J. M. Lang, A. J. Burton, G. J. Bartlett, G. G. Rhys, A. J. Mulholland and D. N. Woolfson, *ACS Synth. Biol.*, 2018, **7**, 1808–1816.

14 J. Liu, Q. Zheng, Y. Deng, C.-S. Cheng, N. R. Kallenbach and M. Lu, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 15457–15462.

15 N. R. Zaccai, B. Chi, A. R. Thomson, A. L. Boyle, G. J. Bartlett, M. Bruning, N. Linden, R. B. Sessions, P. J. Booth, R. L. Brady and D. N. Woolfson, *Nat. Chem. Biol.*, 2011, **7**, 935–941.

16 A. R. Thomson, C. W. Wood, A. J. Burton, G. J. Bartlett, R. B. Sessions, R. L. Brady and D. N. Woolfson, *Science*, 2014, **346**, 485–488.

17 J. Walshaw and D. N. Woolfson, *J. Struct. Biol.*, 2003, **144**, 349–361.

18 G. G. Rhys, C. W. Wood, E. J. M. Lang, A. J. Mulholland, R. L. Brady, A. R. Thomson and D. N. Woolfson, *Nat. Commun.*, 2018, **9**, 4132.

19 P. Teare, C. F. Smith, S. J. Adams, S. Anbu, B. Ciani, L. J. C. Jeuken and A. F. A. Peacock, *Dalton Trans.*, 2018, **47**, 10784–10790.

20 A. L. Boyle, M. Rabe, N. S. A. Crone, G. G. Rhys, N. Soler, P. Voskamp, N. S. Pannu and A. Kros, *Chem. Sci.*, 2019, **10**, 7456–7465.

21 C. W. Wood, J. W. Heal, A. R. Thomson, G. J. Bartlett, A. Á. Ibarra, R. L. Brady, R. B. Sessions and D. N. Woolfson, *Bioinformatics*, 2017, **33**, 3043–3050.

22 S. McIntosh-Smith, T. Wilson, A. Á. Ibarra, J. Crisp and R. B. Sessions, *Comput. J.*, 2011, **55**, 192–205.

23 S. McIntosh-Smith, J. Price, R. B. Sessions and A. A. Ibarra, *Int. J. High Perform. Comput. Appl.*, 2015, **29**, 119–134.

24 A. Chakrabartty, J. A. Schellman and R. L. Baldwin, *Nature*, 1991, **351**, 586–588.

25 A. Chakrabartty, T. Kortemme and R. L. Baldwin, *Protein Sci.*, 1994, **3**, 843–852.

26 C. Nick Pace and J. Martin Scholtz, *Biophys. J.*, 1998, **75**, 422–427.

27 P. Abrusci, M. Vergara-Irigaray, S. Johnson, M. D. Beeby, D. R. Hendrixson, P. Roversi, M. E. Friede, J. E. Deane, G. J. Jensen, C. M. Tang and S. M. Lea, *Nat. Struct. Mol. Biol.*, 2013, **20**, 99–104.

28 M. Podobnik, P. Savory, N. Rojko, M. Kisovec, N. Wood, R. Hambley, J. Pugh, E. J. Wallace, L. McNeill, M. Bruce, I. Liko, T. M. Allison, S. Mehmood, N. Yilmaz, T. Kobayashi, R. J. C. Gilbert, C. V. Robinson, L. Jayasinghe and G. Anderluh, *Nat. Commun.*, 2016, **7**, 11598.

29 W. M. Dawson, E. J. M. Lang, G. G. Rhys, K. L. Shelley, C. Williams, R. L. Brady, M. P. Crump, A. J. Mulholland and D. N. Woolfson, *Nat. Commun.*, 2021, **12**, 1530.

30 J. Walshaw and D. N. Woolfson, *J. Mol. Biol.*, 2001, **307**, 1427–1450.

31 J. W. Heal, G. J. Bartlett, C. W. Wood, A. R. Thomson and D. N. Woolfson, *Bioinformatics*, 2018, **34**, 3316–3323.

32 D. N. Woolfson, *Subcell. Biochem.*, 2017, **82**, 35–61.

33 F. Wang, O. Gnewou, C. Modlin, L. C. Beltran, C. Xu, Z. Su, P. Juneja, G. Grigoryan, E. H. Egelman and V. P. Conticello, *Nat. Commun.*, 2021, **12**, 407.

34 A. Leaver-Fay, R. Jacak, P. B. Stranges and B. Kuhlman, *PLoS One*, 2011, **6**, e20937.

35 K. Y. Wei, D. Moschidi, M. J. Bick, S. Nerli, A. C. McShan, L. P. Carter, P. S. Huang, D. A. Fletcher, N. G. Sgourakis, S. E. Boyken and D. Baker, *Proc. Natl. Acad. Sci. U. S. A.*, 2020, **117**, 7208–7215.

36 A. D. St-Jacques, M. E. C. Eyahpaise and R. A. Chica, *ACS Catal.*, 2019, **9**, 5480–5485.

37 S. E. Boyken, Z. Chen, B. Groves, R. A. Langan, G. Oberdorfer, A. Ford, J. M. Gilmore, C. Xu, F. DiMaio, J. H. Pereira, B. Sankaran, G. Seelig, P. H. Zwart and D. Baker, *Science*, 2016, **352**, 680–687.

38 J. Walshaw and D. N. Woolfson, *Protein Sci.*, 2001, **10**, 668–673.

39 E. H. Egelman, C. Xu, F. DiMaio, E. Magnotti, C. Modlin, X. Yu, E. Wright, D. Baker and V. P. Conticello, *Structure*, 2015, **23**, 280–289.

40 C. R. Calladine, A. Sharff and B. Luisi, *J. Mol. Biol.*, 2001, **305**, 603–618.

41 K. R. Mahendran, A. Niitsu, L. Kong, A. R. Thomson, R. B. Sessions, D. N. Woolfson and H. Bayley, *Nat. Chem.*, 2017, **9**, 411–419.