Rhys, G. G., Cross, J. A., Dawson, W. M., Thompson, H. F., Shanmugaratnam, S., Savery, N. J., Dodding, M. P., Höcker, B., & Woolfson, D. N. (2022). De novo designed peptides for cellular delivery and subcellular localisation. *Nature Chemical Biology*, *18*(9), 999-1004. https://doi.org/10.1038/s41589-022-01076-6

De novo designed peptides for cellular delivery and subcellular localisation

Guto G. Rhys[1,†], Jessica A. Cross[2,3,†], William M. Dawson[2,†], Harry F. Thompson[2,3], Sooruban Shanmugaratnam[1], Nigel J. Savery[3,4], Mark P. Dodding[3], Birte Höcker[1,*], and Derek N. Woolfson[2,3,4,*]

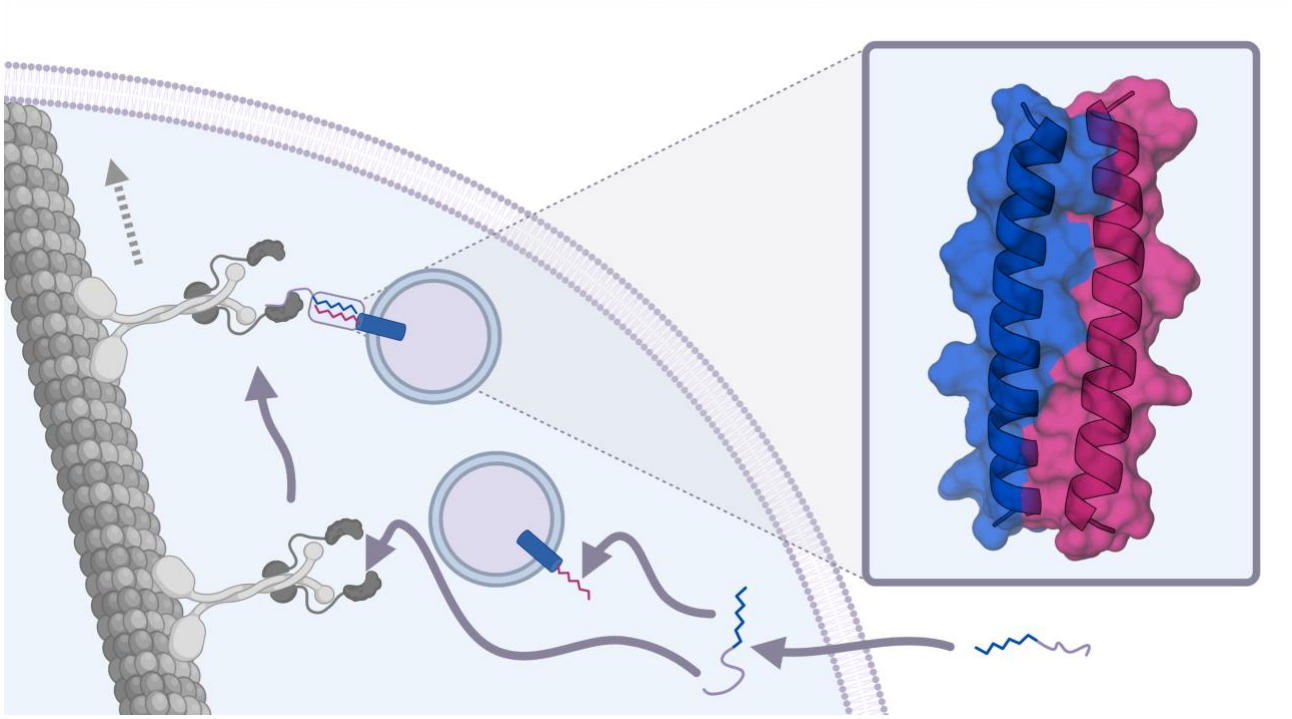
[1]Department of Biochemistry, University of Bayreuth, Universitätsstrasse 30, 95447 Bayreuth, Germany
[2]School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, United Kingdom
[3]School of Biochemistry, University of Bristol, Medical Sciences Building, University Walk, Bristol BS8 1TD, United Kingdom
[4]BrisSynBio, Life Sciences Building, University of Bristol, Bristol, BS8 1TQ, United Kingdom

[†]These authors contributed equally.

*Corresponding authors: birte.hoecker@uni-bayreuth.de and d.n.woolfson@bristol.ac.uk

Abstract

Increasingly, it is possible to design peptide and protein assemblies *de novo* from first principles or computationally. This provides new routes to functional synthetic polypeptides, including designs to target and bind proteins of interest. Much of this work has been developed *in vitro*. Therefore, a challenge is to deliver *de novo* polypeptides efficiently to sites of action within cells. Here, we describe the design, characterization, intracellular delivery, and subcellular localisation of a *de novo* synthetic peptide system. This comprises a dual-function basic peptide, programmed both for cell penetration and target binding, and a complementary acidic peptide that can be fused to proteins of interest and introduced into cells using synthetic DNA. The designs are characterized *in vitro* using biophysical methods and X-ray crystallography. The utility of the system for delivery into mammalian cells and subcellular targeting is demonstrated by marking organelles and actively engaging functional protein complexes.



Introduction

De novo peptide and protein design has advanced to allow the construction of a rapidly increasing variety of stably folded structures and assemblies completely from scratch.[1–3] This ability ushers in new challenges for *de novo* design. These include endowing synthetic polypeptides with functions such as binding and catalysis and, as most designs to date have been made and tested *in vitro*, porting these molecules into living cells, tissues and organisms. Whilst there have been successes with in-cell and *in vivo* applications of *de novo* design,[4–13] considerable challenges remain to make this mainstream activity. Here, we address the challenge of delivering exogenous cargoes to subcellular targets in mammalian cells. We do this by tagging the cargo with a *de novo* designed cell-penetrating peptide that forms a highly specific peptide-peptide interaction with a partner peptide appended to the subcellular target.

Peptide-peptide, peptide-protein and protein-protein interactions (PPIs) are central to most, if not all, biological processes.[14] Therefore, disrupting existing PPIs or generating new ones to monitor or intervene in such processes is a major endeavour in peptide and protein design and engineering.[14–16] Nonetheless, even with bioactive polypeptides that target PPIs in hand, the challenge of delivering these into mammalian cells where they would be most useful remains. Exogenous reagents—usually proteins and larger assemblies—presented to cells can enter actively through endocytosis.[17,18] In this case, however, the reagents end up in the endolysosomal system, which can inhibit release into the cytosol.[19] Alternatively, certain reagents—such as cell-penetrating peptides (CPPs)[20,21]—can enter cells *via* direct or passive translocation across biological membranes. Typically, CPPs are basic or cationic peptides of ≤30 amino acids in length, and the addition or conjugation of such sequences can be used to effect the import of various cargoes into cells.[22–25]

The small size of CPPs has advantages and disadvantages: it facilitates cell uptake, but it leaves little room for internal functionalisation. One strategy to address this is to enhance the cell-permeability of a functional peptide directly by overlaying charged residues at non-target-binding positions.[26] α-Helical coiled coils (CCs) are a class of peptide assembly with potential to be programmed for both cell penetration and protein recognition.[27,28] Effectively, CC peptides are bifaceted: The peptide-peptide interfaces are well defined and can be designed rationally or computationally to deliver various oligomers, topologies and partners.[2,3] This leaves exposed surfaces that can be altered for water or membrane solubility. Moreover, CC peptides are short and synthetically accessible and could fit the CPP paradigm.[29–32]

Here, we describe the design and characterisation of new antiparallel homo- and hetero-dimeric CCs that expand the set of structurally resolved *de novo* CCs.[33–39] We show that the hetero-dimer is effective for both delivery into cells and subcellular targeting. In this system, an arginine-rich basic CC peptide is used as the exogenous reagent, as it incorporates both cell-penetrating and target-binding functionalities. The sequence of its complementary acidic partner can be fused to various subcellular proteins and introduced into cells *via* the transfection of synthetic DNA. We show that this system can tag intracellular proteins, complexes and organelles. Moreover, the two components diversify the labels that can be incorporated for imaging and intervening in biological processes.

Rational *de novo* design of antiparallel coiled-coil dimers

We decided to design a cell-penetrating system using antiparallel rather than parallel CCs. This was because structurally resolved *de novo* designed antiparallel CC dimers are currently poorly represented,[40] which offers an opportunity to develop new design principles

and sequences to address our target. The rational design of CC dimers to tetramers centres on the interfacial ***g-a-d-e*** sites of their ***abcdefg***, heptad repeats (Fig. 1a).[28] This can be guided by the large number of available natural CC sequences and structures. Therefore, we selected a subset of antiparallel two-helix CCs from the CC+ structural database[41] and examined the amino-acid profiles of their repeats (Supplementary Note, Supplementary Figs. 1 and 2). This revealed preferences for isoleucine (Ile, I) and leucine (Leu, L) at the core-forming ***a*** and ***d*** positions, respectively. Unfortunately, this combination also specifies parallel CC dimers.[28] Polar residues at these sites can aid oligomer specification and helix orientation.[28,34,42,43] Comparison of the profiles for antiparallel and parallel dimers indicated that arginine (Arg, R) was enriched at ***a*** sites of the former by almost two-fold. Inspection of antiparallel structures showed that R@***a*** pairs most often (30% of cases) with aspartate (Asp, D) at proximal ***e'*** sites of the partnering helices resulting in inter-side-chain hydrogen bonds or salt bridges (Fig. 1b). Therefore, in addition to I@***a*** and L@***d***, we included a single R@***a*** to interact with a complementary D@***e*** in four-heptad, 30-residue designs, Table 1. The use of oppositely charged residues at core-flanking ***e*** and ***g*** sites is established in CC design.[28] Therefore, to target an antiparallel homo-dimer, apCC-Di, we made the *N*-terminal ***e*** and ***g*** sites glutamate (Glu, E), the most favoured residues at these sites in CC dimers, and the complementary *C*-terminal ***e*** and ***g*** sites Arg and lysine (Lys, K), respectively. The hetero-dimer, apCC-Di-AB, required two peptides: apCC-Di-A had all Glu at ***e*** and ***g***; and apCC-Di-B had a combination of Arg and Lys at these sites (i) to complement the acidic 'A' peptide, and (ii) to promote cell penetration. Several sequences were made by solid-phase peptide synthesis and characterised biophysically (Supplementary Table 1, Supplementary Figs. 3–9, and 20–28), but we focus here on the most successful designs apCC-Di and apCC-Di-AB below (see Supplementary Note for design iterations).

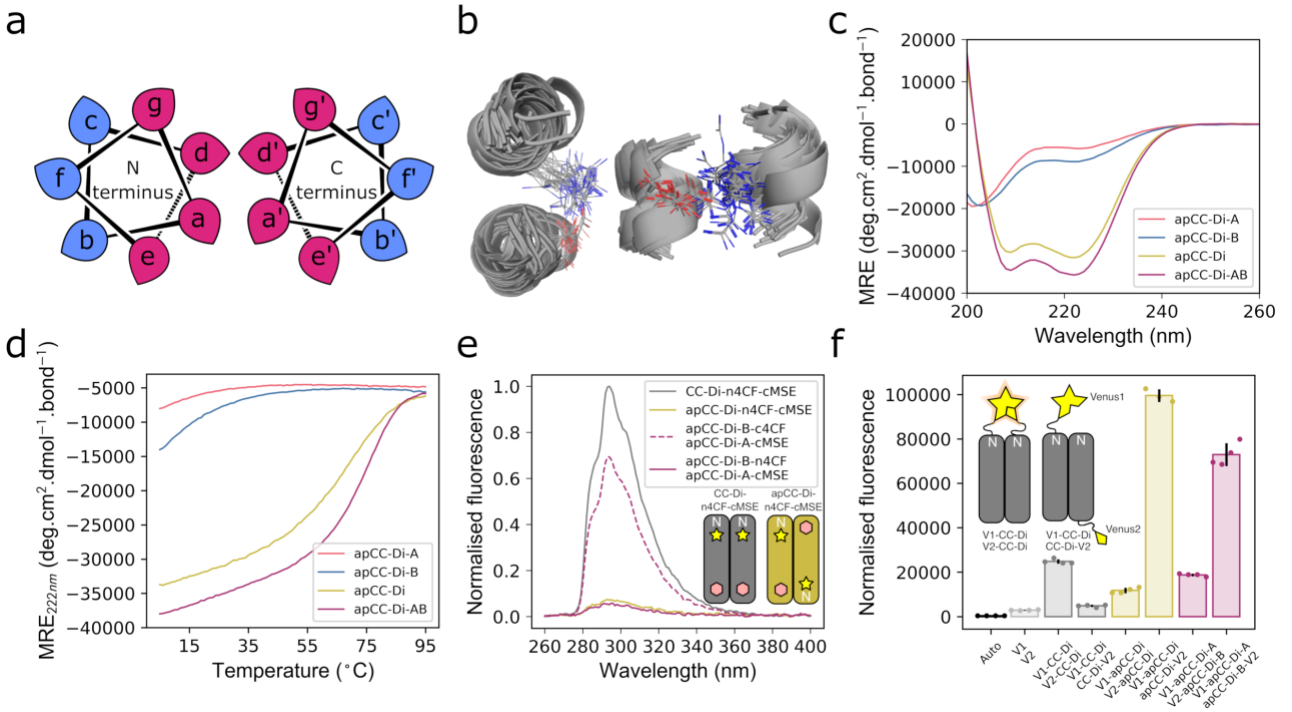


Fig. 1: Design and biophysical characterisation of antiparallel coiled-coil dimers. a, Helical wheels for the seven-residue repeat of CC sequences labelled ***a*** to ***g***. Interface positions (***a***, ***d***, ***e*** & ***g***) are coloured pink and solvent-exposed positions (***b***, ***c*** & ***f***) are coloured blue. **b**, Examples of arginine to aspartic acid interactions observed in antiparallel CC dimers. **c, d,** Circular dichroism spectra at 20 °C (**c**) and thermal denaturation profiles (**d**), for apCC-Di and apCC-Di-AB. Conditions: 100 μM total peptide, phosphate buffered saline (PBS), pH 7.4. **e**, Fluorescence-quenching assay for labelled apCC-Di, apCC-Di-AB, and the control parallel peptide CC-Di. Key: n and c indicate mutations near the *N* and *C* termini, respectively; 4CF, 4-Cyano-L-phenylalanine fluorophore (yellow star); MSE, L-selenomethionine fluorescence quencher (pink hexagon); CC-Di, parallel coiled-coil dimer control. Conditions: 100 μM concentration of each peptide, PBS,

pH 7.4. **f**, Bimolecular fluorescence complementation (BiFC) assay for CC-Di, apCC-Di and apCC-Di-AB. V1 and V2 are *N*- and *C*-terminal fragments of the Venus fluorescent protein, respectively. Peptide names that precede the Venus fragment name denote fusion to the *C* terminus of the CC, *e.g.* apCC-Di-B-V2, and *vice versa*. Fluorescence values are normalised to cell density (OD_{600}) and are presented as mean values +/- 1 SD from n≥3 technical replicate measurements (dots).

Table 1: Peptide sequences and summary of biophysical data. Key: CD, percentage α helix determined from the MRE_{222}; T_M, midpoint of the thermal unfolding transition. AUC, molecular weight measured by analytical ultracentrifugation relative to expected masses for a monomer (*m*) or heterodimer (*d*). XRD, resolution of X-ray crystallography data, and Protein Data Bank deposition id in brackets.

Name	Sequence ***cdefgab cdefgab cdefgab cdefgab***	CD (%)	T_M (°C)	AUC	XRD
apCC-Di	GQLEQELA ALDQQIA ALKQRRA ALKWQIQG	83	63	1.9 x *m*	1.08 Å (7Q1R)
apCC-Di-A	GQLEQELA ALDQEIA AAEQELA ALDWQIQG	15	-	-	-
apCC-Di-B	GQLKQRRA ALKQRIA ALKQRRA ALKWQIQG	23	-	-	-
apCC-Di-AB	apCC-Di-A + apCC-Di-B	93	69	0.9 x *d*	1.68 Å (7Q1T)

Structural characterisation of apCC-Di and apCC-Di-AB

The circular dichroism (CD) spectrum of apCC-Di indicated a fully α–helical structure (Fig. 1c). Furthermore, temperature-dependent CD measurements showed a sigmoidal, reversible, and concentration-dependent unfolding transition as expected for a unique, cooperatively folded species (Fig. 1d). Consistent with this and the design, sedimentation-equilibrium analytical-ultracentrifugation experiments returned dimer molecular weights (Supplementary Figs. 22–28). The apCC-Di-AB design required iteration (Supplementary Note). In isolation, the final peptides, apCC-Di-A and apCC-Di-B, were unfolded (Fig. 1c,d). When mixed 1:1, they gave a fully α–helical species (Fig. 1c), with a reversible thermal unfolding transition (Fig. 1d), and a dimeric molecular weight (Supplementary Fig. 28). We estimated nanomolar affinities for apCC-Di and apCC-Di-AB from the concentration-dependent thermal-denaturation experiments (Supplementary Figs. 29 and 30).[44]

We crystallised and solved X-ray structures for apCC-Di, apCC-Di-AB, and several variants of these (Table 1; Supplementary Table 3). The atomic-resolution structures of apCC-Di and apCC-Di-AB revealed antiparallel helical dimers as designed (Fig. 2a,b; Extended Data Figs. 1 and 2). Moreover, the program SOCKET2[45] only identified knobs-into-holes interactions, the structural signature of CCs, within the designed antiparallel interfaces (Fig. 2c, and Extended Data Figs. 3–5); and the introduced Arg-Asp pairs made salt-bridges in three of four cases (Fig. 2d).

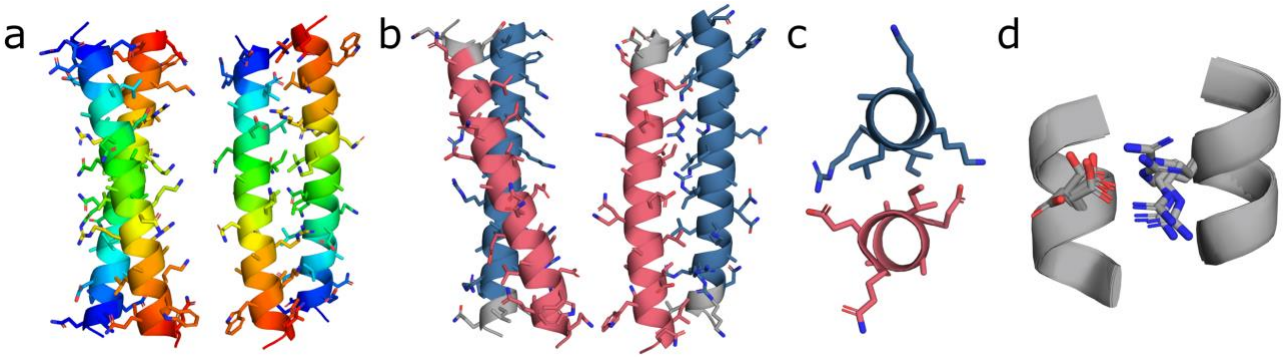


Fig. 2: X-ray crystal structures reveal antiparallel coiled coils as designed.
a, **b**, Orthogonal views of X-ray crystal structures of apCC-Di (**a**) and apCC-Di-AB (**b**). In panel **b** the individual chains are apCC-Di-A (red) and apCC-Di-B (blue). The *N*-terminal turns are coloured grey. **c**, A single heptad from apCC-Di-AB viewed along the superhelical axis showing the definitive knobs-into-holes packing observed in CCs. **d**, Overlay of designed Arg:Asp interactions from the experimental apCC-Di and

apCC-Di-AB structures. Arg-N_η to Asp-O_δ distances of $\leqslant$ 3.5 Å were taken as indicating salt-bridge formation.

Next, we sought to confirm that the folded states were antiparallel in solution. First, guided by the X-ray crystal structures, we introduced the fluorescent 4-cyano-L-phenylalanine (4CF) and a fluorescence quencher L-selenomethionine (MSE)[46] at proximal interhelical ***g*** and ***d*** positions. This was done for apCC-Di and apCC-Di-AB, and with 4CF *N* terminal and MSE *C* terminal and *vice versa* (Supplementary Table 1, Supplementary Figs. 10–16 and Extended data Fig. 6). As controls, we made analogous changes to the parallel CC-Di design[34] and introduced 4CF and MSE at the *C* termini of the apCC-Di-AB design. Only the control peptides fluoresced (Fig. 1e, Extended data Fig. 7). By contrast, fluorescence was completely quenched in the antiparallel designs with juxtaposed 4CF and MSE residues (Fig. 1e).

Second and to test the designs in cells, we reconfigured a bimolecular fluorescence complementation (BiFC) assay to examine peptide assembly and orientation directly in *Escherichia coli* cells.[47] Using a variant of the fluorescent protein Venus split after residue 154 to give two fragments, V1 and V2,[48–50] we fused one *de novo* peptide to the *C* terminus of V1. The other test peptide was fused to either the *N* or *C* terminus of V2 (Extended Data Fig. 8 and Supplementary Figs. 31–34). The hypothesis was that any fragment complementation and fluorescence should report on antiparallel or parallel peptide association, respectively; *i.e.*, antiparallel CCs should increase the fluorescence readout when fused to the *C* terminus of V1 and the *N* terminus of V2. Experimentally, this was done by transforming *E. coli* cells with expression plasmids encoding synthetic genes for the peptide-split protein fusions, and then the interaction was monitored by direct measurement of fluorescence from the cells. The background fluorescence from V1 plus V2 alone was low (Fig. 1f, Extended Data Fig. 8). As a positive control, we tested the homomeric, parallel CC-Di sequence in the system.[34] With CC-Di on the *C* terminus of V1 and the *N* terminus of V2 there was ≈1.4-fold increase in fluorescence above baseline. Whereas, with CC-Di *C* terminal to both fragments the signal increased by ≈7-fold. These results are consistent with the parallel orientation of CC-Di and validate the assay. Turning to the new designs, fusing apCC-Di to the *N* terminus of V2 gave a larger signal and an ≈8-fold increase in fluorescence relative to the *C*-terminal fusion, consistent with apCC-Di being antiparallel. Furthermore, with ap-CC-Di-A fused to the *C* terminus of V1 the better complementary fusion had ap-CC-Di-B at the *N* terminus of V2 with a ≈4-fold increase in fluorescence over the alternative construct. Thus, both apCC-Di and ap-CC-Di-AB are confirmed as antiparallel dimers in solution and in cells.

apCC-Di-B delivers functional cargo to subcellular locations

With a verified antiparallel heterodimer in hand, we turned to the delivery into eukaryotic cells. We resynthesised apCC-Di-A and apCC-Di-B each with an *N*-terminal 5-carboxytetramethylrhodamine (TAMRA) fluorophore (Supplementary Figs. 17–19). The peptides were added separately to cultures of HeLa cells, fixed and imaged after 1 hour. Only TAMRA-apCC-Di-B entered cells (Fig. 3a). Moreover, it did so at 4 ˚C indicating passive uptake directly into the cytosol rather than *via* active endocytosis. To confirm this, we pre-incubated cells with endocytosis inhibitors, MiTMAB or Dynasore, or the electron transport chain inhibitor, NaN_3, for 30 minutes prior to the addition of apCC-Di-B. In all three cases, peptide uptake by the cells was unaffected (Extended data Fig. 9). Thus, apCC-Di-B is a *de novo* cell-penetrating peptide (CPP) consistent with the basic, arginine-rich sequence design. To determine whether apCC-Di-B taken up by mammalian cells can still bind apCC-Di-A, we added TAMRA-apCC-Di-B to HeLa cells transfected with plasmids for transient expression of GFP alone or of a GFP-apCC-Di-A fusion. Cells were lysed, GFP proteins

were immunoprecipitated, and any associated TAMRA fluorescence signal was analysed using a microplate reader. TAMRA-apCC-Di-B signal only co-immunoprecipitated with the GFP-apCC-Di-A fusion (Fig. 3b).

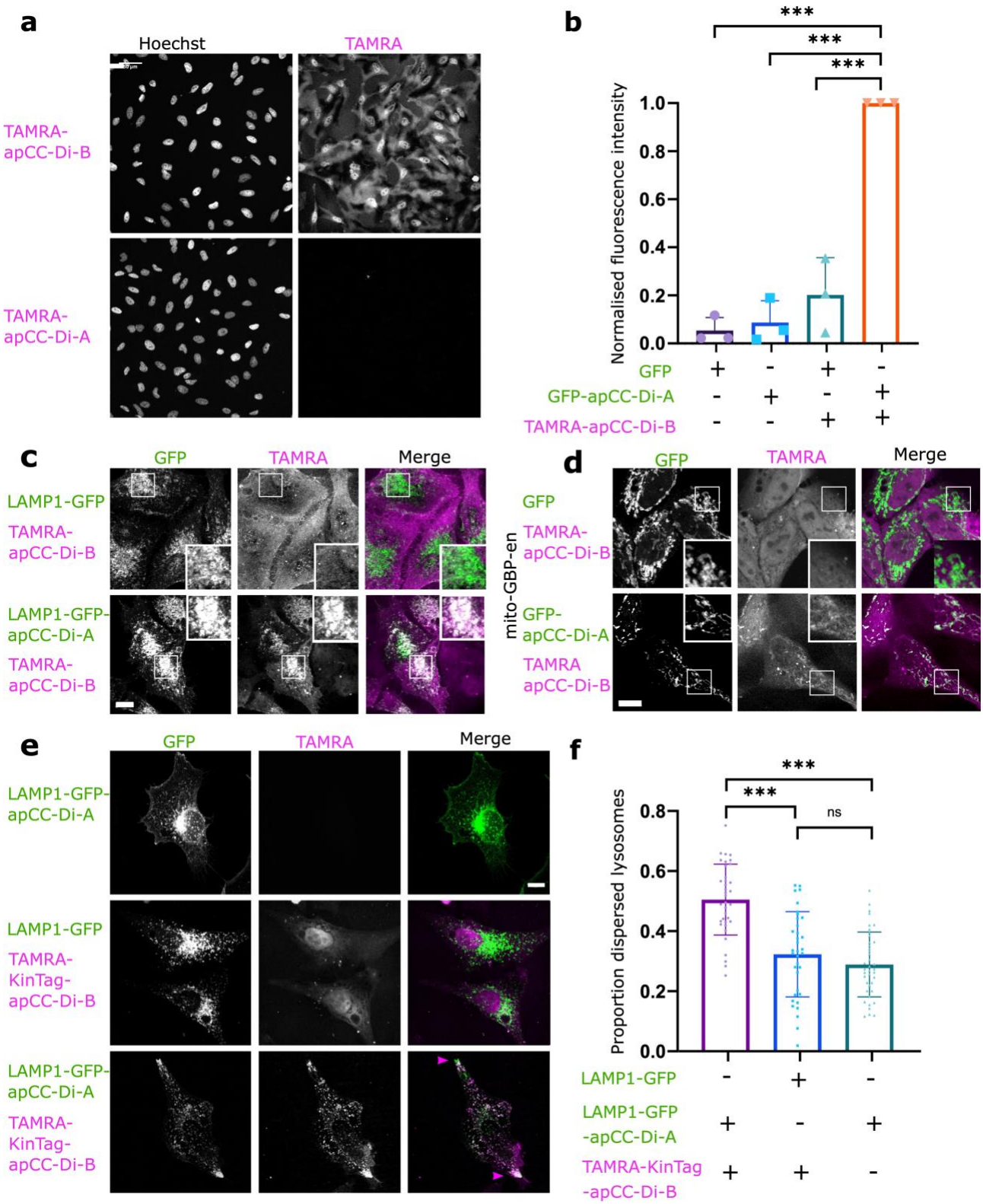


Fig. 3: apCC-Di-B is cell penetrating and binds to apCC-Di-A in mammalian cells. a, HeLa cells treated with 2 μM TAMRA labelled apCC-Di-B (top) or apCC-Di-A (bottom) peptides for 1 hour. apCC-Di-B is cell penetrating shown by TAMRA fluorescence in the cytoplasm and nucleus, apCC-Di-A does not enter cells. Scale bar 50 μm. **b**, Cells transfected with GFP or GFP-apCC-Di-A and treated with TAMRA-apCC-Di-B or vehicle control were subject to GFP-immunoprecipitation and quantification of TAMRA immobilised on beads measured by fluorescence of boiled bead eluant at 555 nm. Fluorescence was normalised to the signal of cells transfected with GFP-apCC-Di-A and treated with TAMRA-apCC-Di-B in three independent experiments. Data

are presented as mean values +/- 1 SD and an unpaired T-test is used for statistical analysis, *** P<0.0001. TAMRA-apCC-DiB binds to GFP-apCC-Di-A in cell lysate and is immobilised on beads. **c**, Cells transfected with LAMP1-GFP (top) or LAMP1-GFP-apCC-Di-A (bottom) and treated with TAMRA-apCC-Di-B. Colocalization of GFP and TAMRA fluorescence on lysosomes indicates apCC-Di-AB dimerization. Scale bar 20 μm. **d**, Cells transfected with a mitochondrially targeted GFP nanobody, mitoGFP-en, and GFP (top) or GFP-apCC-Di-A (bottom) and treated with TAMRA-apCC-Di-B. Colocalization of GFP and TAMRA fluorescence on mitochondria indicates apCC-Di-AB dimerization. Scale bar 20 μm. **e**, Cells transfected with LAMP1-GFPapCC-Di-A (top and bottom) or LAMP1-GFP (middle) and treated with TAMRA-KinTag-apCC-Di-B (middle and bottom) or vehicle control (top). KinTag is delivered to lysosomes by apCC-Di-AB dimerization and promotes transport of lysosomes to the cell periphery, purple arrows indicate peripheral accumulations. Scale bar 10 μm. **f**, Quantification of lysosome dispersal from e by comparison of fluorescence intensity of perinuclear cluster with total LAMP1 fluorescence in a minimum of 35 cells (n = 36, 35, 40 cells from left to right), pooled from three independent experiments. Data are presented as mean values +/- 1 SD and an unpaired T-test is used for statistical analysis, *** P<0.0001, ns P = 0.2421.

Next, we asked if the apCC-Di-AB system could target subcellular organelles. For this, we constructed a fusion with the lysosomal-associated membrane protein 1, LAMP1-mGFP-apCC-Di-A, to present an exposed peptide bait on late endosomes and lysosomes.[51] Transfection of this gene into HeLa cells followed by exogenous treatment with TAMRA-apCC-Di-B showed striking recruitment to LAMP1 compartments (Fig. 3c). In addition, we used a mitochondrially targeted GFP nanobody (pMito-GBPen) to recruit GFP-apCC-Di-A to mitochondria.[52] Treatment of these cells with TAMRA-apCC-Di-B resulted in its localisation to mitochondria (Fig. 3d).

Finally, to determine if apCC-Di-B could deliver functional moieties that drive new PPIs in cells, we extended it with a 'KinTag' sequence (Supplementary Table 1 and 2). Recently, we developed this *de novo* sequence to bind endogenous kinesin-1 light chains and, so, provide a synthetic adapter between kinesin-1 heavy chain motors and potential cargoes.[53] Cells expressing the LAMP1-mGFP-apCC-Di-A fusion were treated with TAMRA-KinTag-apCC-Di-B as an exogenous reagent. TAMRA and GFP fluorescence colocalized to lysosomal puncta indicating formation of the apCC-Di-AB complex (Fig. 3e), implying that the apCC-Di-B component facilitated cell entry and localisation to the intended target. Moreover, compared with controls, the lysosomes dispersed from the perinuclear region of the cell and accumulated at the cell periphery (Fig. 3e,f). Interestingly, this led to altered cell shapes and GFP- and TAMRA-positive projections (Fig. 3e). This is consistent with the KinTag component of the exogenous reagent binding to and strongly activating kinesin, which—mediated by the apCC-Di-AB bridge—transports its lysosome cargo along microtubules to the cell periphery.

Discussion and conclusion

In summary, we have designed and fully characterised homo- and heteromeric antiparallel coiled-coil (CC) dimers. This includes in-cell confirmation of the antiparallel orientation using a variation on the split-Venus bimolecular fluorescence complementation (BiFC) assay, validated with known parallel interacting CCs, which should be useful to others investigating the orientation of PPI domains in cells. Others have reported designs for antiparallel CC dimers.[42,54–62] However, to our knowledge, apCC-DI-AB and apCC-Di are the first heteromeric and the second homomeric *de novo* design to be confirmed structurally to atomistic resolution.[40] In addition, we have shown that the basic, arginine-rich B peptide enters directly into the cytosol of mammalian cells; it is a new cell-penetrating peptide. Once in cells, the B peptide can seek and bind its acidic partner introduced as parts of synthetic genes. This can be exploited using the B peptide as an exogenous reagent for targeting intracellular proteins, complexes and organelles labelled with the complementary A peptide.

Finally, the system can deliver an orthogonally functional peptide, which we demonstrate by recruiting an endogenous motor protein to relocate organelle cargoes.

Others have reported the use of CC peptides, notably natural leucine zippers, in combination with CPPs to import cargoes into cells.[63–65] Our system is different as it combines cell penetration and partner binding within a single CC peptide motif. As we have demonstrated, this allows one half of the system to be functionalised for both cell penetration and partner binding. In turn, this can be used to import dyes for imaging or short bioactive peptides to engage with active subcellular processes. We anticipate that this and related *de novo* designed peptides will add to the armoury of reagents available for identifying and following subcellular assemblies and processes in mammalian cells.

Acknowledgements
G.G.R. is supported by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 88899. J.A.C. is supported by the EPSRC-funded Bristol Centre for Doctoral Training in Chemical Synthesis (EP/G036764/1). H.F.T. is supported by the EPSRC- and BBSRC-funded Centre for Doctoral Training in Synthetic Biology (EP/L016494/1). M.P.D. is a Lister Institute of Preventative Medicine Fellow and work in his lab is supported by BBSRC (BB/S000917/1). We thank the University of Bristol School of Chemistry Mass Spectrometry Facility for access to the EPSRC-funded Bruker Ultraflex MALDI-TOF instrument (EP/K03927X/1), the BBSRC-funded BrisSynBio centre for access to peptide synthesis and a plate reader (BB/L01386X/1), and the Wolfson Bioimaging Facility for their assistance in this work. B.H. acknowledges financial support and allocation of beamtime by HZB and we thank the beamline staff at BESSY for assistance.

Author Contributions
G.G.R., J.A.C. and W.M.D. contributed equally to the project. G.G.R., W.M.D., B.H. and D.N.W. conceived and developed the project idea. G.G.R. conducted the bioinformatic analysis and designed the peptides. W.M.D. and G.G.R. synthesised and purified the peptides and conducted the biophysical analysis. W.M.D. conducted and analysed the fluorescence quenching assays. H.F.T. and N.J.S. conducted and analysed the bimolecular fluorescence complementation (BiFC) assay. S.S. and G.G.R. conducted the X-ray crystallography. J.A.C. and M.P.D. conducted and analysed all mammalian cell experiments. G.G.R, B.H. and D.N.W. wrote the manuscript with contributions from all authors.

Competing Interests
The authors declare no competing interests.

References

1. Huang, P.-S., Boyken, S. E. & Baker, D. The coming of age of de novo protein design. *Nature* **537**, 320–327 (2016).
2. Korendovych, I. V. & DeGrado, W. F. De novo protein design, a retrospective. *Q. Rev. Biophys.* **53**, E3 (2020).
3. Woolfson, D. N. A brief history of de novo protein design: minimal, rational, and computational. *J. Mol. Biol.* **433**, 167160 (2021).
4. Correia, B. E. *et al.* Proof of principle for epitope-focused vaccine design. *Nature* **507**, 201–206 (2014).
5. Langan, R. A. *et al.* De novo design of bioactive protein switches. *Nature* **572**, 205–210 (2019).
6. Foight, G. W. *et al.* Multi-input chemical control of protein dimerization for

programming graded cellular responses. *Nat. Biotechnol.* **37**, 1209–1216 (2019).
7. Silva, D.-A. *et al.* De novo design of potent and selective mimics of IL-2 and IL-15. *Nature* **565**, 186–191 (2019).
8. Smith, A. J., Thomas, F., Shoemark, D., Woolfson, D. N. & Savery, N. J. Guiding biomolecular interactions in cells using de novo protein-protein interfaces. *ACS Synth. Biol.* **8**, 1284–1293 (2019).
9. Fink, T. *et al.* Design of fast proteolysis-based signaling and logic circuits in mammalian cells. *Nat. Chem. Biol.* **15**, 115–122 (2019).
10. Chen, Z. *et al.* De novo design of protein logic gates. *Science* **368**, 78–84 (2020).
11. Cao, L. *et al.* De novo design of picomolar SARS-CoV-2 miniprotein inhibitors. *Science* **370**, 426–431 (2020).
12. Lebar, T., Lainšček, D., Merljak, E., Aupič, J. & Jerala, R. A tunable orthogonal coiled-coil interaction toolbox for engineering mammalian cells. *Nat. Chem. Biol.* **16**, 513–519 (2020).
13. Bryan, C. M. *et al.* Computational design of a synthetic PD-1 agonist. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
14. Titeca, K., Lemmens, I., Tavernier, J. & Eyckerman, S. Discovering cellular protein-protein interactions: Technological strategies and opportunities. *Mass Spectrom. Rev.* **38**, 79–111 (2019).
15. Cheng, S. S., Yang, G. J., Wang, W., Leung, C. H. & Ma, D. L. The design and development of covalent protein-protein interaction inhibitors for cancer treatment. *J. Hematol. Oncol.* **13**, 26 (2020).
16. Kahan, R., Worm, D. J., De Castro, G. V., Ng, S. & Barnard, A. Modulators of protein-protein interactions as antimicrobial agents. *RSC Chem. Biol.* **2**, 387–409 (2021).
17. Mellman, I. Endocytosis and molecular sorting. *Annu. Rev. Cell Dev. Biol.* **12**, 575–625 (1996).
18. Park, J. H. & Oh, N. Endocytosis and exocytosis of nanoparticles in mammalian cells. *Int. J. Nanomedicine* **9**, 51 (2014).
19. Torres-Vanegas, J. D., Cruz, J. C. & Reyes, L. H. Delivery systems for nucleic acids and proteins: Barriers, cell capture pathways and nanocarriers. *Pharmaceutics* **13**, 428 (2021).
20. Guidotti, G., Brambilla, L. & Rossi, D. Cell-penetrating peptides: from basic research to clinics. *Trends Pharmacol. Sci.* **38**, 406–424 (2017).
21. Ruseska, I. & Zimmer, A. Internalization mechanisms of cell-penetrating peptides. *Beilstein J. Nanotechnol.* **11**, 101–123 (2020).
22. Agrawal, P. *et al.* CPPsite 2.0: A repository of experimentally validated cell-penetrating peptides. *Nucleic Acids Res.* **44**, D1098–D1103 (2016).
23. Rhodes, C. A. *et al.* Cell-Permeable Bicyclic Peptidyl Inhibitors against NEMO-IκB Kinase Interaction Directly from a Combinatorial Library. *J. Am. Chem. Soc.* **140**, 12102–12110 (2018).
24. Dougherty, P. G. *et al.* Cyclic Peptidyl Inhibitors against CAL/CFTR Interaction for Treatment of Cystic Fibrosis. *J. Med. Chem.* **63**, 15773–15784 (2020).
25. Buyanova, M. *et al.* Discovery of a Bicyclic Peptidyl Pan-Ras Inhibitor. *J. Med. Chem.* **64**, 13038–13053 (2021).
26. Dietrich, L. *et al.* Cell Permeable Stapled Peptide Inhibitor of Wnt Signaling that Targets β-Catenin Protein-Protein Interactions. *Cell Chem. Biol.* **24**, 958-968.e5 (2017).
27. Lupas, A. N. & Bassler, J. Coiled coils – a model system for the 21st century. *Trends Biochem. Sci.* **42**, 1–11 (2016).
28. Woolfson, D. N. Coiled-Coil Design: Updated and Upgraded. in *Fibrous Proteins: Structures and Mechanisms* (eds. Parry, D. A. D. & Squire, J. M.) 35–61 (Springer,

2017). doi:10.1007/978-3-319-49674-0_2
29. Lamarre, B., Ravi, J. & Ryadnov, M. G. GeT peptides: A single-domain approach to gene delivery. *Chem. Commun.* **47**, 9045–9047 (2011).
30. Nakayama, N., Takaoka, S., Ota, M., Takagaki, K. & Sano, K. I. Effect of the aspect ratio of coiled-coil protein carriers on cellular uptake. *Langmuir* **34**, 14286–14293 (2018).
31. Dhankher, A., Lv, W., Studstill, W. T. & Champion, J. A. Coiled coil exposure and histidine tags drive function of an intracellular protein drug carrier. *J. Control. Release* **339**, 248–258 (2021).
32. Li, J. *et al.* The Coiled-Coil Forming Peptide (KVSALKE)5 Is a Cell Penetrating Peptide that Enhances the Intracellular Delivery of Proteins. *Adv. Healthc. Mater.* 2102118 (2021). doi:10.1002/adhm.202102118
33. Zaccai, N. R. *et al.* A de novo peptide hexamer with a mutable channel. *Nat. Chem. Biol.* **7**, 935–941 (2011).
34. Fletcher, J. M. *et al.* A basis set of de novo coiled-coil peptide oligomers for rational protein design and synthetic biology. *ACS Synth. Biol.* **1**, 240–250 (2012).
35. Thomson, A. R. *et al.* Computational design of water-soluble α-helical barrels. *Science* **346**, 485–488 (2014).
36. Rhys, G. G. *et al.* Maintaining and breaking symmetry in homomeric coiled-coil assemblies. *Nat. Commun.* **9**, 4132 (2018).
37. Rhys, G. G. *et al.* Navigating the structural landscape of de novo α-helical bundles. *J. Am. Chem. Soc.* **141**, 8787–8797 (2019).
38. Edgell, C. L., Savery, N. J. & Woolfson, D. N. Robust de novo-designed homotetrameric coiled coils. *Biochemistry* **59**, 1087–1092 (2020).
39. Dawson, W. M. *et al.* Coiled coils 9-to-5: Rational: de novo design of α-helical barrels with tunable oligomeric states. *Chem. Sci.* **12**, 6923–6928 (2021).
40. Majerle, A. *et al.* A nanobody toolbox targeting dimeric coiled-coil modules for functionalization of designed protein origami structures. *Proc. Natl. Acad. Sci. U. S. A.* **118**, (2021).
41. Testa, O. D., Moutevelis, E. & Woolfson, D. N. CC+: A relational database of coiled-coil structures. *Nucleic Acids Res.* **37**, 1–8 (2008).
42. Oakley, M. G. & Kim, P. S. A buried polar interaction can direct the relative orientation of helices in a coiled coil. *Biochemistry* **37**, 12603–12610 (1998).
43. Boyken, S. *et al.* De novo design of protein homo-oligomers with modular hydrogen bond network-mediated specificty. *Science* **399**, 69–72 (2016).
44. Thomas, F., Boyle, A. L., Burton, A. J. & Woolfson, D. N. A set of de novo designed parallel heterodimeric coiled coils with quantified dissociation constants in the micromolar to sub-nanomolar regime. *J. Am. Chem. Soc.* **135**, 5161–5166 (2013).
45. Kumar, P. & Woolfson, D. N. *Socket2* : a program for locating, visualizing and analyzing coiled-coil interfaces in protein structures. *Bioinformatics* **37**, 4575–4577 (2021).
46. Watson, M. D., Peran, I. & Raleigh, D. P. A non-perturbing probe of coiled coil formation based on electron transfer mediated fluorescence quenching. *Biochemistry* **55**, 3685–3691 (2016).
47. Ghosh, I., Hamilton, A. D. & Regan, L. Antiparallel leucine zipper-directed protein reassembly: application to the green fluorescent protein. *J. Am. Chem. Soc.* **122**, 5658–5659 (2000).
48. Nagai, T. *et al.* A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Nat. Biotechnol.* **20**, 87–90 (2002).
49. Shyu, Y. J., Liu, H., Deng, X. & Hu, C. D. Identification of new fluorescent protein fragments for bimolecular fluorescence complementation analysis under physiological conditions. *Biotechniques* **40**, 61–66 (2006).

50. Nakagawa, C., Inahata, K., Nishimura, S. & Sugimoto, K. Improvement of a venus-based bimolecular fluorescence complementation assay to visualize bfos-bjun interaction in living cells. *Biosci. Biotechnol. Biochem.* **75**, 1399–1401 (2011).
51. Falcón-Pérez, J. M., Nazarian, R., Sabatti, C. & Dell'Angelica, E. C. Distribution and dynamics of Lamp1-containing endocytic organelles in fibroblasts deficient in BLOC-3. *J. Cell Sci.* **118**, 5243–5255 (2005).
52. Küey, C., Larocque, G., Clarke, N. I. & Royle, S. J. Unintended perturbation of protein function using GFP nanobodies in human cells. *J. Cell Sci.* **132**, (2019).
53. Cross, J. A., Chegkazi, M. S., Steiner, R. A., Woolfson, D. N. & Dodding, M. P. Fragment-linking peptide design yields a high-affinity ligand for microtubule-based transport. *Cell Chem. Biol.* **0**, (2021).
54. Monera, O. D., Zhou, N. E., Kay, C. M. & Hodges, R. S. Comparison of antiparallel and parallel two-stranded α-helical coiled-coils. Design, synthesis, and characterization. *J. Biol. Chem.* **268**, 19218–19227 (1993).
55. Monera, O. D., Kay, C. M. & Hodges, R. S. Electrostatic interactions control the parallel and antiparallel orientation of α-helical chains in two-stranded α-helical coiled-coils. *Biochemistry* **33**, 3862–3871 (1994).
56. Campbell, K. M. & Lumb, K. J. Complementation of buried lysine and surface polar residues in a designed heterodimeric coiled coil. *Biochemistry* **41**, 7169–7175 (2002).
57. McClain, D. L., Gurnon, D. G. & Oakley, M. G. Importance of potential interhelical salt-bridges involving interior residues for coiled-coil stability and quaternary structure. *J. Mol. Biol.* **324**, 257–270 (2002).
58. Gurnon, D. G., Whitaker, J. A. & Oakley, M. G. Design and characterization of a homodimeric antiparallel coiled coil. *J. Am. Chem. Soc.* **125**, 7518–7519 (2003).
59. Pagel, K. *et al.* Advanced approaches for the characterization of a de novo designed antiparallel coiled coil peptide. *Org. Biomol. Chem.* **3**, 1189–1194 (2005).
60. Hadley, E. B. & Gellman, S. H. An antiparallel α-helical coiled-coil model system for rapid assessment of side-chain recognition at the hydrophobic interface. *J. Am. Chem. Soc.* **128**, 16444–16445 (2006).
61. Negron, C. & Keating, A. E. A set of computationally designed orthogonal antiparallel homodimers that expands the synthetic coiled-coil toolkit. *J. Am. Chem. Soc.* **136**, 16544–16556 (2014).
62. Nagarkar, R. P., Fichman, G. & Schneider, J. P. Engineering and characterization of a pH-sensitive homodimeric antiparallel coiled coil. *Pept. Sci.* **112**, e24180 (2020).
63. Hakata, Y. *et al.* A novel leucine zipper motif-based hybrid peptide delivers a functional peptide cargo inside cells. *Chem. Commun.* **51**, 413–416 (2015).
64. Bode, S. A. *et al.* Coiled-coil-mediated activation of oligoarginine cell-penetrating peptides. *ChemBioChem* **18**, 185–188 (2017).
65. Hakata, Y., Michiue, H., Ohtsuki, T., Miyazawa, M. & Kitamatsu, M. A leucine zipper-based peptide hybrid delivers functional Nanog protein inside the cell nucleus. *Bioorganic Med. Chem. Lett.* **29**, 878–881 (2019).

Methods

Bioinformatic analysis

The coiled-coil database, CC+ (coiledcoils.chm.bris.ac.uk/ccplus/search/dynamic_interface) was filtered using the following options, Redundancy: <= 50%, α helices: 2, Orientation: antiparallel, Partnering: any, Chains: any, Repeats: any, and Length: >11 residues. The SOCKET text files were downloaded for the filtered list of coiled coils. The knob residue identity at ***a*** and ***d*** positions (and their corresponding hole residue identity and relative heptad positions) were extracted from the SOCKET files into dictionaries for analysis.

11,855 ***a*** positions and 11,885 ***d*** positions, and their corresponding hole residues, were extracted from 1,275 biological units. Scripts were written in Python and are available online as a Jupyter Notebook session (https://github.com/Hoecker-Lab/Designed_apDimers).

Peptide synthesis

Standard Fmoc solid-phase peptide synthesis was performed on a 0.1 mM scale using CEM Liberty Blue automated peptide synthesis apparatus with inline UV monitoring. Activation was achieved with DIC/Cl-HOBt. Fmoc deprotection was performed with 20% v/v morpholine/DMF with addition of 5% formic acid to prevent aspartimide formation. Double couplings were used for β-branched residues and the subsequent amino acid. All peptides were synthesised from *C* to *N* terminus as the *C*-terminal amide on Rink amide resin or ChemMatrix solid support. *N*-terminally acetylated peptides were produced by addition of acetic anhydride (0.25 mL) and pyridine (0.3 mL) in DMF (5 mL) for 30 minutes at room temperature (rt). Fluorescently labelled peptides were produced by addition of TAMRA (0.1 mM, 2 eq.), HATU (0.095 mM, 1.9 eq.) and DIPEA (0.225 mM, 4.5 eq.) in DMF (3 mL) to DMF washed peptide resin (0.05 mM) with agitation for 3 hours. Resin was washed with 20% piperidine in DMF (5 mL) for 2 x 30 minutes to remove any excess dye. All manipulations were carried out under foil to exclude light. Biotinylated peptides were produced by addition of Cl-HOBt (1 ml, 0.5 M) to a solution of biotin (0.25 mM, 5 eq.) and DIC (1 ml, 1 M) in DMF and the mixture added to DMF washed peptide resin with agitation at rt for 2 hours. Peptides were cleaved from the solid support by addition of TFA (9.5 mL), TIPS (0.25 mL) and water (0.25 mL) for 3 hours with shaking at rt. The cleavage solution was reduced to ≈1 mL under a flow of nitrogen. Crude peptide was precipitated upon addition of ice-cold diethyl ether (40 mL) and recovered via centrifugation. The resulting precipitant was dissolved in 1:1 acetonitrile and water (≈15 mL) and lyophilised to yield crude peptide as a solid.

Peptide purification

Peptides were purified by reverse phase HPLC on a Phenomenex Luna C18 stationary phase column (150 x 10 mm, 5 μM particle size, 100 A pore size) using a preparative JASCO HPLC system. A linear gradient of 20–80% acetonitrile and water (with 0.1% TFA) was typically applied over 30 minutes with gradient adjusted for each peptide to optimise separation. For acidic peptides, a linear gradient of 10–60% acetonitrile, and ammonium bicarbonate (25 mM) in water was typically applied over 30 minutes with gradient adjusted for each peptide to optimise separation. Where peptides were insoluble in HPLC solvents, DMSO was added to improve solubility. Chromatograms were monitored at wavelengths of 220 and 280 nm. The identities of the peptides were confirmed using MALDI-TOF mass spectrometry using a Bruker ultrafleXtreme II instrument in reflector mode. Peptides were spotted on a ground-steel target plate using dihydroxybenzoic acid or α-cyano-4-hydroxycinnamic acid (CHCA) as the matrix. Masses quoted are for the monoisotopic mass as the singly protonated species. Masses were measured to 0.1% accuracy. Peptide purities were determined using a JASCO analytical HPLC system, fitted with a reverse-phase Kinetex® C18 analytical column (100 x 4.6 mm, 5 µm particle size, 100 Å pore size). Fractions containing pure peptide were pooled and lyophilised. Peptides were dissolved in buffer, typically PBS or FP assay buffer (25 mM pH 7.4 HEPES buffer with 150 mM NaCl, 5 mM BME) and their concentrations determined by UV–Vis at 280 nm on a ThermoScientific Nanodrop 2000 spectrophotometer using the known extinction coefficients for tyrosine (ε280 = 1280 M^{-1} cm^{-1}) and tryptophan (ε_{280} = 5690 M^{-1} cm^{-1}). The total extinction coefficients for a sequence were calculated by adding the extinction coefficient for each chromophore. For TAMRA coupled peptides concentration was calculated by measurement of UV absorbance at 555 nm (ε_{555}(TAMRA) = 85000 $mol^{-1}cm^{-1}$).

Circular-dichroism spectroscopy

Circular-dichroism (CD) data were collected on a JASCO J-810 or J-815 spectropolarimeter fitted with a Peltier temperature controller (Jasco UK). Peptide samples were dissolved at 100 μM concentration for individual peptides or 50 μM for each peptide in mixed systems, in phosphate buffered saline (PBS; 8.2 mM sodium phosphate, 1.8 mM potassium phosphate, 137 mM sodium chloride, 2.7 mM potassium chloride at pH 7.4). CD spectra were recorded in 1 mm path length quartz cuvettes at 20 °C. The instruments were set with a scan rate of 100 nm min^{-1}, a 1 nm interval, a 1 nm bandwidth and a 1 s response time and scans are an average of 8 scans recorded for the same sample. Thermal denaturation data were acquired at 222 nm between 5 °C and 95 °C, with settings as above and a ramping rate of 40 °C hr^{-1} and are a single recording. Baselines recorded using the same buffer, cuvette and parameters were subtracted from each dataset. The spectra were converted from ellipticities (deg) to mean residue ellipticities (MRE, ($deg.cm^2.dmol^{-1}.res^{-1}$)) by normalizing for concentration of peptide bonds and the cell path length using the equation:

$$MRE(\mathrm{deg.\,cm^2.\,dmol^{-1}.\,res^{-1}}) = \frac{\theta \times 100}{c \times l \times b}$$

Where the variable θ is the measured difference in absorbed circularly polarized light in millidegrees, *c* is the millimolar concentration of the specimen, *l* is the path-length of the cuvette in cm and *b* is the number of amide bonds in the polypeptide, for which the N-terminal acetyl bond was included but not the *C*-terminal amide.

Sedimentation equilibrium analytical ultracentrifugation

AUC sedimentation equilibrium experiments were conducted at 20 °C in a Beckman Optima XL-I or XL-A analytical ultracentrifuge using an An-60 Ti rotor (Beckmann Coulter). Solutions were made up in PBS at 100 μM total peptide concentration. The experiments were run in 2-channel centerpiece. The samples were centrifuged at speeds in the range 44–60 krpm and scans at each recorded speed were duplicated. Data were fitted to single, ideal species models using SEDFIT (v15.2b)/SEDPHAT, comprising a minimum of four speeds. 95% confidence limits were obtained via Monte Carlo analysis of the obtained fits.

Affinity measurements

Affinities were estimated by plotting the melting temperature (T_M) against the total concentration of peptide.[66] Thermal denaturation curves were measured between 5 and 95 °C, at total peptide concentrations between 15 and 100 μM, and the T_M extracted from the second derivative of the spectra. All measurements were performed in PBS (8.2 mM sodium phosphate dibasic, 1.8 mM potassium phosphate monobasic, 137 mM sodium chloride, 2.4 mM potassium chloride, pH 7.4).

X-ray crystallography

Sitting-drop vapour-diffusion experiments were set up in 3-well IntelliPlates (Art Robbins) using a Phoenix pipetting robot (Art Robbins). Drops for the homomers were pipetted in a 3:2 ratio (peptide:reservoir solution) and the heteromers were pipetted in a 1:1 ratio. Crystals were obtained at 20 °C directly from a sparse matrix screen using the commercially available JCSG Core I-IV screens (Qiagen). Hit conditions can be found in the Supplementary Table 4. Crystals were mounted using CryoLoops (Hampton Research), transferred into a cryogenic solution made of the corresponding reservoir solution supplemented with 25% Glycerol, and flash-cooled in liquid nitrogen. Diffraction data were collection at -173 °C in 0.1 ° slices on beamlines BL 14.1 and BL 14.2 at the BESSY II electron storage ring operated by the Helmholtz-Zentrum Berlin.[67] Data were processed using DIALS (2.0.2)[68] or XDSAPP2 (2.0)[69] with XDS (Build 20200417). Resolution cut-off was selected according to $CC_{1/2}$ greater than 0.30, I/σI greater than 0.5, and a completeness

of more than 75% in the highest resolution shell. The data were phased via molecular replacement with Phenix.phaser (2.8.3)[70] using a model poly-alanine α helix without accounting for translational non-crystallographic symmetry. Model building as well as manual and automated refinements were done iteratively using Coot (0.9.6)[71] and either Phenix (1.19.2_4158).refine[72] or Refmac5[73].

Fluorescent quenching assay

Fluorescent-quenching experiments were performed following previously published protocols.[46] In brief, using a Jasco Fluorimeter, peptides (100 μM each) were excited at 240 nm and emission spectra recorded between 260 and 400 nm. All measurements were recorded in 50 mM sodium phosphate, pH 7, and are the average of 5 individual spectra.

Biomolecular fluorescence complementation assay

Genestrings (Eurofins Genomics) encoding individual coiled coils were cloned using standard restriction enzyme-ligation cloning. Both V1 (N-Venus) and V2 (C-Venus) constructs[48,49] were amplified by PCR from the entire protein sequence and then cloned into arabinose inducible expression cassettes in the medium copy plasmid backbones pVRc[8] and pBAD, respectively. Site-directed mutagenesis was carried out following the QuikChange II Protocol (Agilent). Full protein sequences for all constructs are given in the supplementary information (Supplementary Table 2).

To monitor the interaction of the coiled coils in cells, TB28 cells (MG1655 ΔLacIZYA)[74] were transformed with the required plasmid combination as indicated. Colonies were picked in triplicate at least and overnight cultures were grown at 37 °C in M9 minimal media + 0.25% glycerol + 0.5 mM $CaCl_2$ + 2 mM $MgSO_4$ + 2 μg/mL thiamine + 0.2% casamino acids (+ 100 μg/mL ampicillin + 25 μg/mL chloramphenicol where required). The cultures were then used to inoculate 10 mL cultures of the same medium with the addition of 0.2% arabinose, and then were grown at 37 °C until reaching an OD_{600} ~ 0.5. Five mL samples were centrifuged at 5000 rpm for 15 min, the supernatant removed and then resuspended in 250 μL of PBS. Additional 1 mL samples were centrifuged at 13000 rpm for 10 min, the supernatant removed and frozen for western blotting at a later stage. Duplicate 100 μL samples of each culture were placed in a black-bottomed 96-well plate, and the fluorescence read using a CLARIOstar microplate reader (BMG Labtech), using the yellow fluorescent protein (YFP) with the following settings: 497 nm excitation (15 nm bandwidth), 540 nm emission (20nm bandwidth), 517.2 nm dichroic. YFP fluorescence (relative fluorescence units) was normalized by dividing by the OD_{600}.

Western blotting

Pellets were resuspended in appropriate volumes of SDS loading buffer + DTT, boiled at 95 °C for 3 min, then 10 μL loaded, along with 10 μL PageRuler™ Plus Prestained Protein Ladder (Thermo), onto 8–15% Mini-PROTEAN Gels (Bio-rad) and run at 160 V until the loading dye reached the bottom of the gel. Proteins were then transferred onto a Immobilon-P membrane (Merck) in standard transfer buffer (1 x Tris Glycine, 10% methanol) for 2 hours at 230 mA. After this, the membranes were washed in PBS and blocked overnight in 10% milk in PBS with gentle rocking. The blots were then washed in PBS, and then incubated with a His primary antibody (Abcam) (1:5000) in 10% milk in PBS for 2 hours at RT with gentle rocking. After washing, the HRP conjugated secondary antibody was added (Abcam) (1:5000) in 10% milk in PBS and incubated for 1 hour. After washing, 2 mL Amersham ECL detection reagent (GE) was added to each blot and incubated for 1 min, and then placed in a cassette along with standard X-ray film for the desired interval before being placed in a developer.

Cell culture
HeLa cells were maintained in high glucose Dulbecco's Modified Eagle's Medium (Gibco Invitrogen) with 10% (v/v) foetal calf serum (Sigma-Aldrich) and 5% penicillin/streptomycin (PAA) (herein referred to as DMEM) at 37 °C and 5% CO_2. For transfection, cells were seeded in 6-well plates on fibronectin coated 13 mm coverslips at a density of 1 x 10^5 cells per well and incubated at 37 °C and 5% CO_2 for 16 h prior to transfection. Cells were transfected with 0.4 μg DNA using Effectene transfection reagent according to the manufacturer's instructions (Qiagen). After transfection cells were incubated at 37 °C, 5% CO_2 for 16 hours. Peptide treatments were in 1 mL DMEM 37 °C and 5% CO_2 for the time indicated. Cells were fixed by addition of 4% paraformaldehyde in PBS at room temperature for 10 minutes (2 mL per well for 6-well plate) and washed 3 x with PBS. Confocal images were collected using a Leica SP5II system with a 63× objective running Leica LAS X and are presented as maximum intensity projections. Figures were assembled using ImageJ in conjunction with Inkscape (0.92). Image analysis was conducted in ImageJ (1.53c) and data was processed using Graphpad Software Prism 8.0.

GFP-Immunoprecipitation
For GFP immunoprecipitation, HeLa cells were plated at a density of 2 x 10^6 per 10 cm dish 4 hours prior to transfection. Cells were incubated on ice and washed in 2 x 5 mL wash buffer (25 mM HEPES pH 7.4, 150 mM NaCl) followed by 1 mL lysis buffer (25 mM HEPES pH 7.4, 150 mM NaCl, 0.1% NP-40, 0.1% Trition-X 100 and protease inhibitor cocktail) for 10 minutes. Cells were centrifuged at 13000 g for 10 minutes at 4 °C and the supernatant incubated with 30 μL of prewashed (3 x 800 μL wash buffer, 2000 rpm) GFPTrap beads for 120 minutes at 4 °C. 50 μL of supernatant was retained for analysis of cell lysate. Beads were centrifuged at 2000 g for 2 minutes, the supernatant removed and washed (3 x 800 μL wash buffer, 2000 rpm). Samples were analysed for GFP expression by gel electrophoresis and western blotting. Elution from beads was achieved by addition of 1% SDS at 95 °C for 10 minutes. TAMRA fluorescence intensity measurements of eluents were performed at 555 nm on a CLARIOstar (BMG Labtech) microplate reader at room temperature.

Data processing and plotting
Unless otherwise stated, all other data were processed and plotted in Python (3.8.5) using Numpy (1.19.2) and matplotlib (3.1.3).

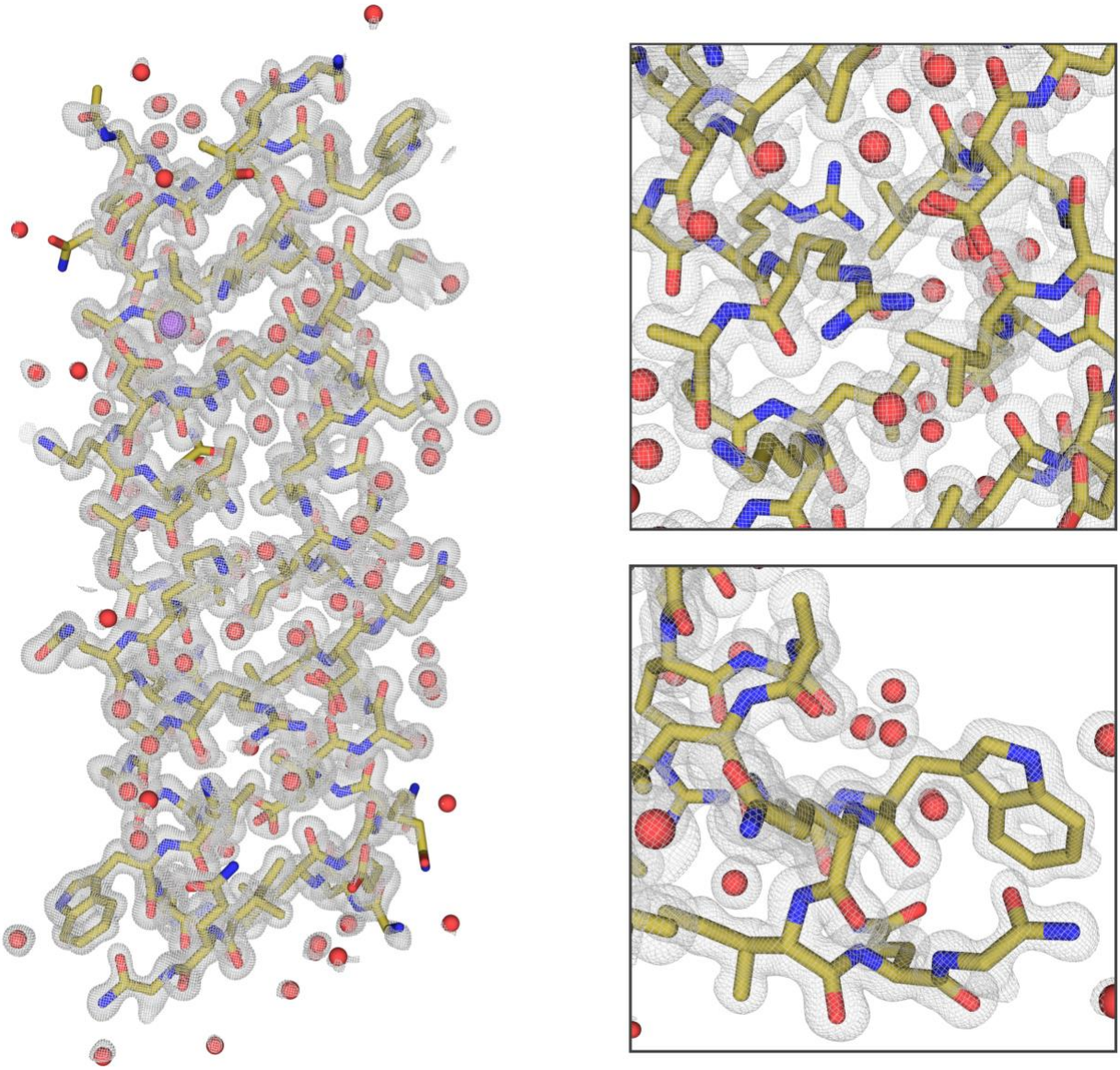
Data availability
The coiled-coil database CC+ is an open and publicly accessible database. The coordinate and structure factor files for Homomer-S, apCC-Di, apCC-Di-AB_var & apCC-Di-AB have been deposited in the Protein Data Bank with accession codes 7Q1Q, 7Q1R, 7Q1S & 7Q1T, respectively. All the raw data used in this publication has been deposited in the Zenodo repository (https://www.doi.org/10.5281/zenodo.6519961).

Code availability
The scripts used for bioinformatic analysis are available from a Zenodo repository (https://www.doi.org/10.5281/zenodo.6518524).

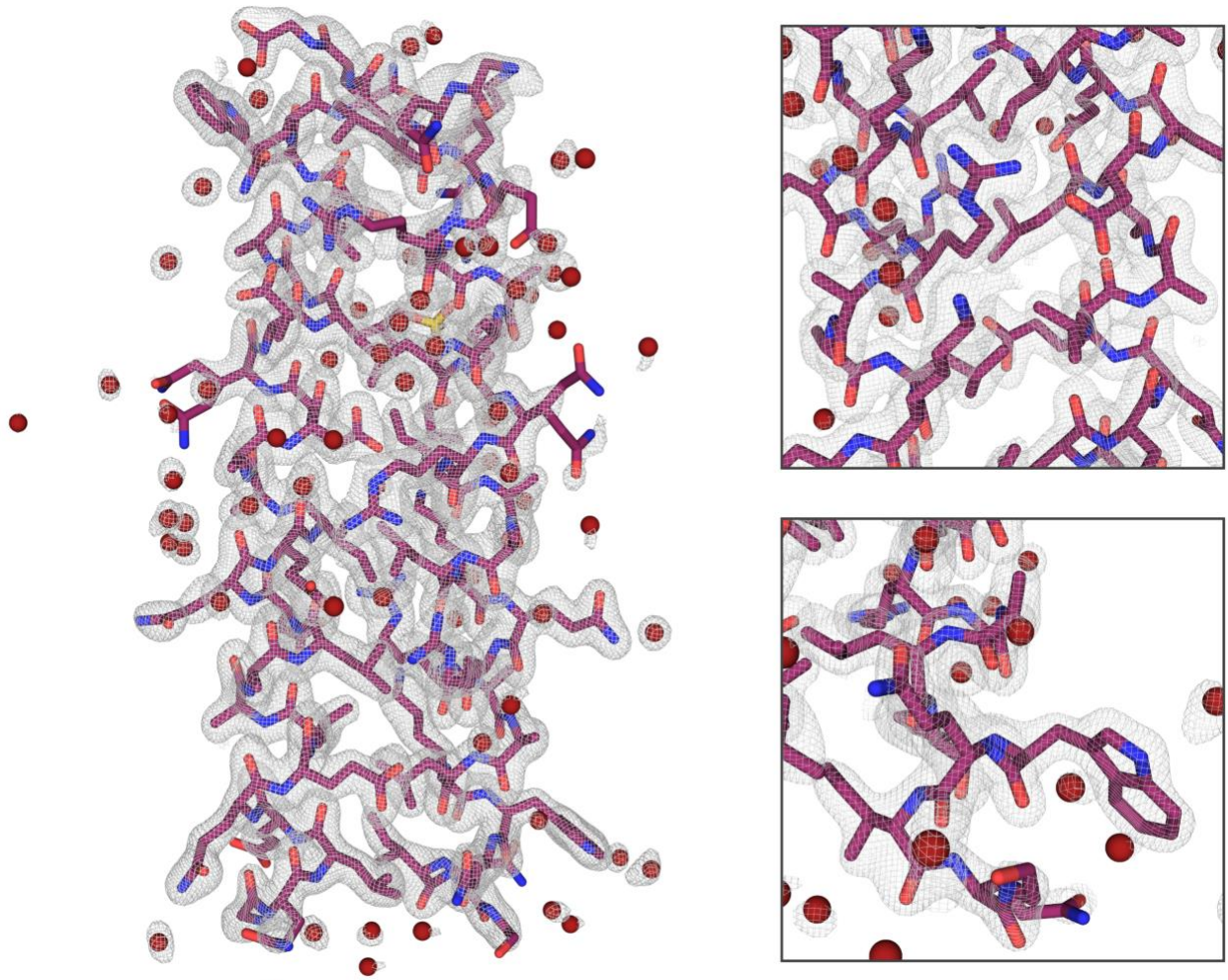
Methods References

66. Marky, L. A. & Breslauer, K. J. Calculating thermodynamic data for transitions of any molecularity from equilibrium melting curves. *Biopolymers* **26**, 1601–1620 (1987).
67. Mueller, U. *et al.* The macromolecular crystallography beamlines at BESSY II of the Helmholtz-Zentrum Berlin: Current status and perspectives. *Eur. Phys. J. Plus* **130**,

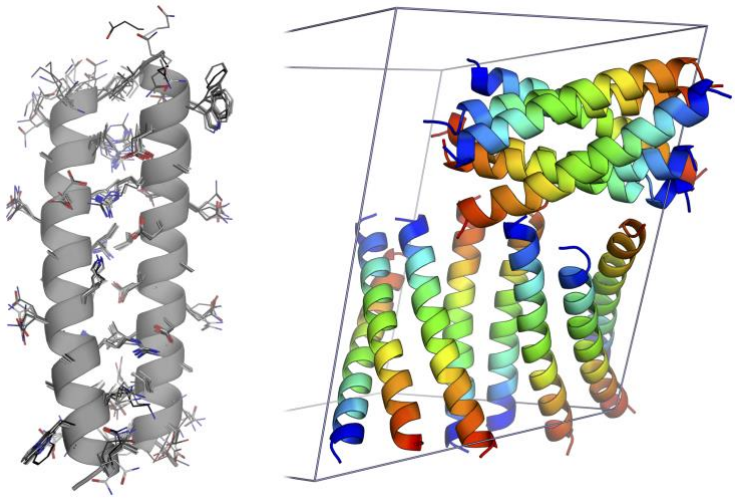
1–10 (2015).
68. Winter, G. *et al. DIALS* : implementation and evaluation of a new integration package. *Acta Crystallogr. Sect. D Struct. Biol.* **74**, 85–97 (2018).
69. Sparta, K. M., Krug, M., Heinemann, U., Mueller, U. & Weiss, M. S. XDSAPP2.0. *J. Appl. Crystallogr.* **49**, 1085–1092 (2016).
70. Mccoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
71. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Cryst* **66**, 486–501 (2010).
72. Afonine, P. V *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Cryst* **68**, 352–367 (2012).
73. Murshudov, G. N. *et al. REFMAC* 5 for the refinement of macromolecular crystal structures. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 355–367 (2011).
74. Bernhardt, T. G. & De Boer, P. A. J. The Escherichia coli amidase AmiC is a periplasmic septal ring component exported via the twin-arginine transport pathway. *Mol. Microbiol.* **48**, 1171–1182 (2003).



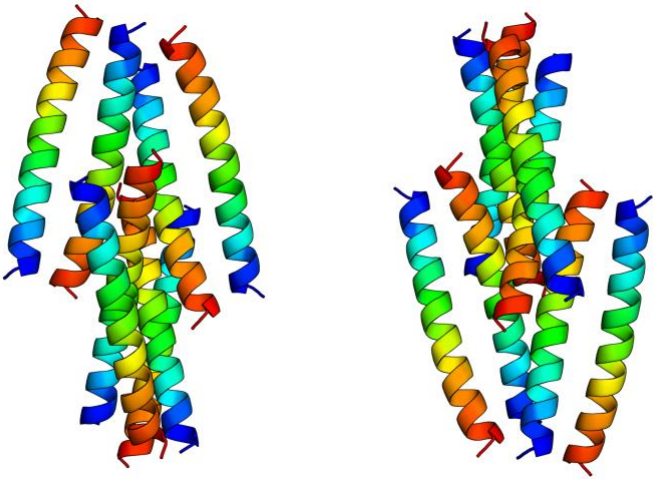
Extended Data Fig. 1 Electron density map of apCC-Di crystal structure. Fo-Fc difference map of the apCC-Di crystal structure represented as a grey wire mesh overlaid on a stick model. The map was contoured at 1 rmsd.



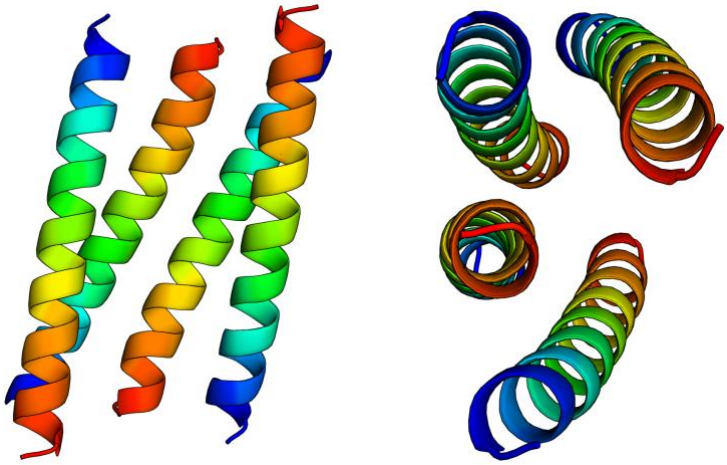
Data Fig. 2 Electron density map of apCC-Di-AB crystal structure. Fo-Fc difference map of the apCC-Di-AB crystal structure represented as a grey wire mesh overlaid on a stick model. The map was contoured at 1 rmsd.



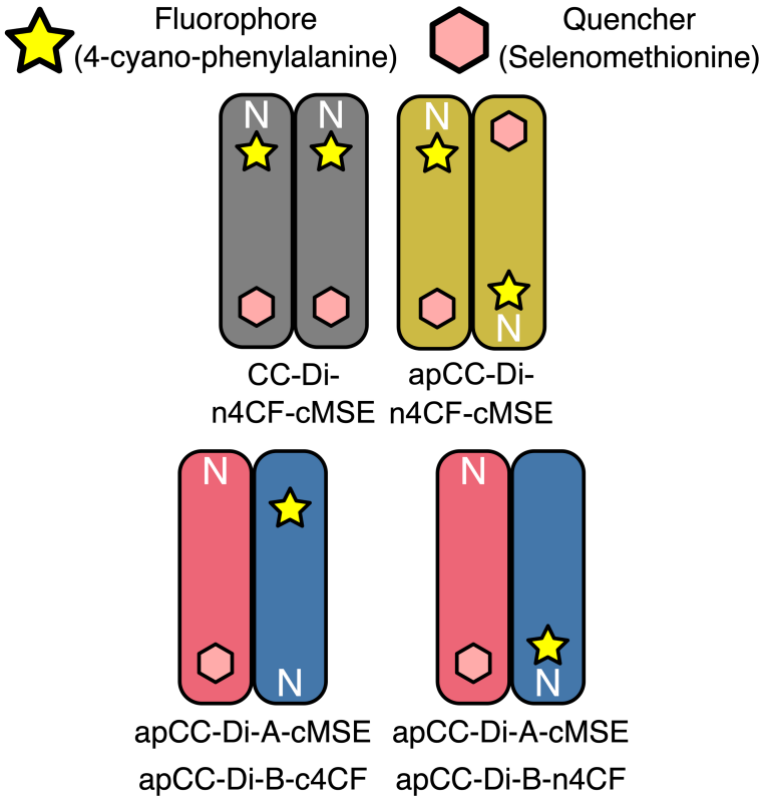
Extended Data Fig. 3 The asymmetric unit of apCC-Di-AB_var is composed of 14 chains. While there are minor differences between the surface-exposed residues of the seven dimer pairs, all the biological units are structurally similar (left). The high number of biological units in the asymmetric unit is a result of crystal packing. In the crystal form, apCC-Di-AB_var forms perpendicular layers of right-handed α-helical fibres (right). Between each dimer pair, the individual peptides self-associate to form KIH parallel interfaces, leading to alternating parallel and antiparallel interfaces.



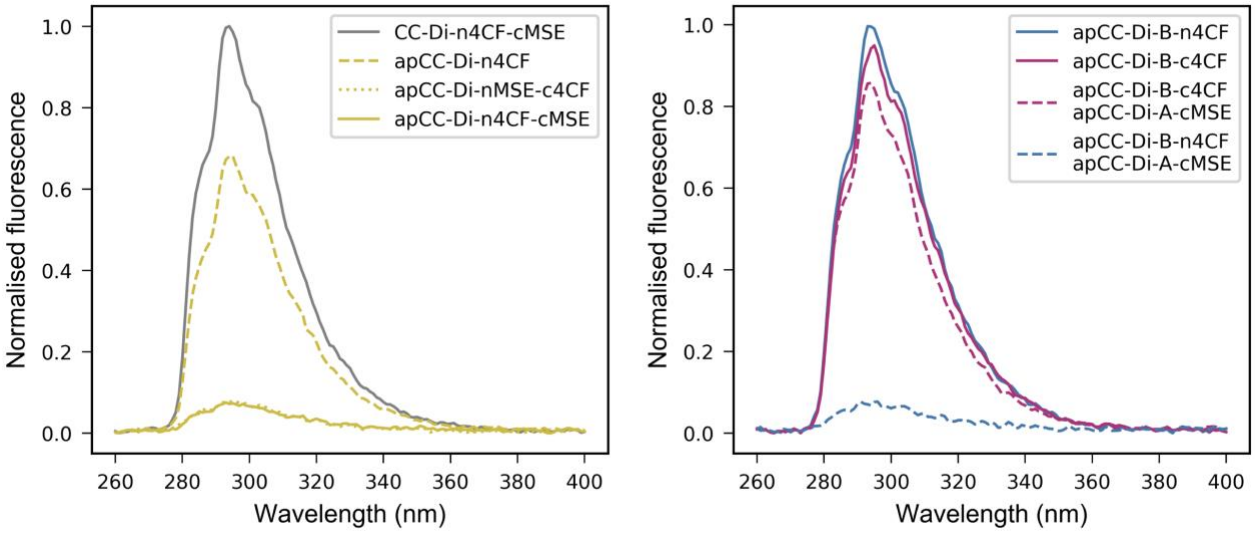
Extended Data Fig. 4 Symmetry mates for the crystal structure of apCC-Di. Four symmetry mates are depicted above that form a repeating pattern in the crystal structure. SOCKET 2 analysis of the structure reveals four antiparallel CC dimers. There are no inter-biological unit CC interactions.



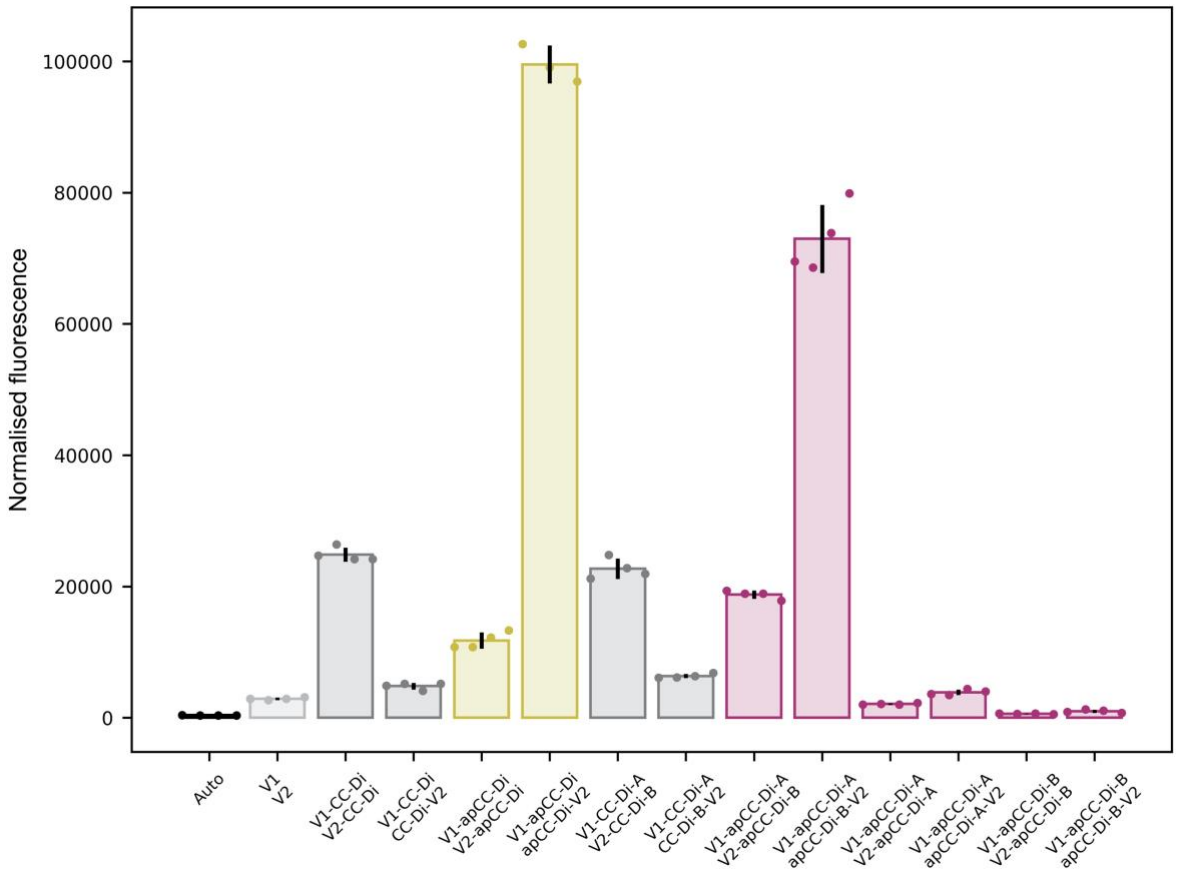
Extended Data Fig. 5 Symmetry mates for the crystal structure of apCC-Di-AB. Two symmetry mates are depicted above that form a repeating pattern in the crystal structure. SOCKET 2 analysis of the structure reveals two antiparallel CC dimers. There are no inter-biological unit CC interactions.



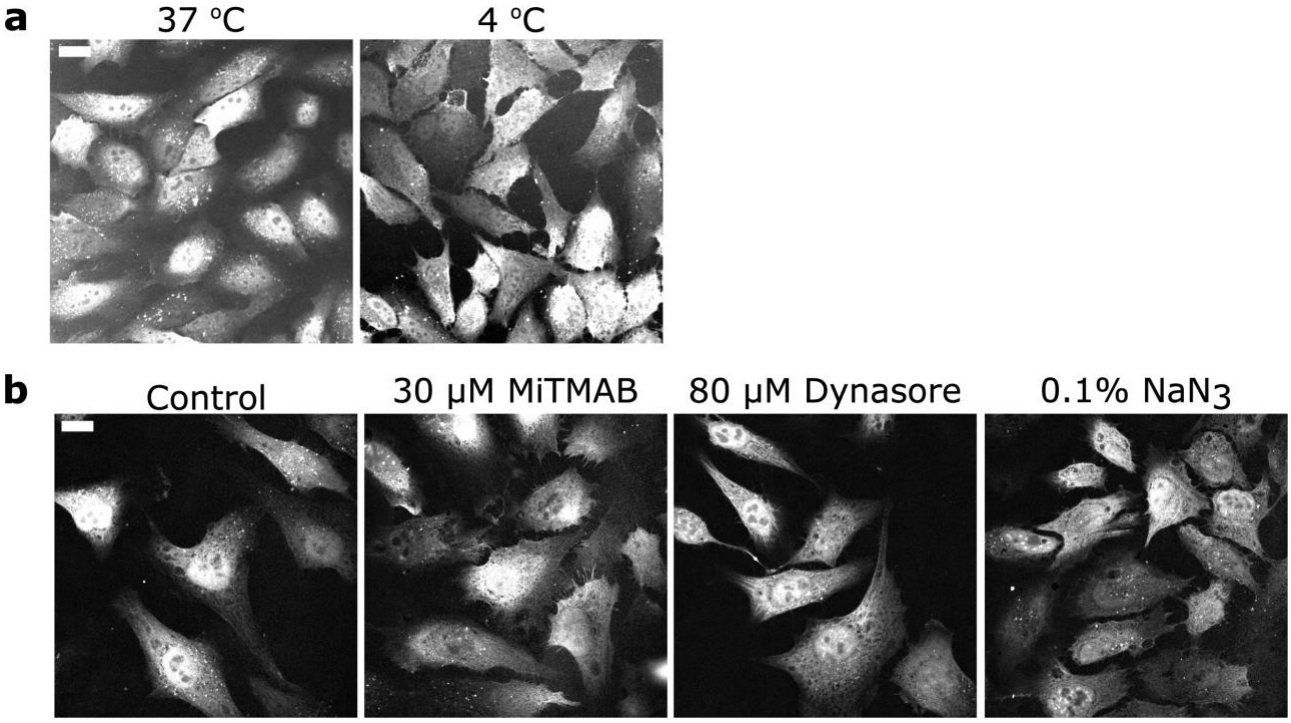
Extended Data Fig. 6 Cartoon representation of mutation positions in designed coiled-coil (CC) dimers for the fluorescence quenching assay. For homomeric CCs, the fluorophore was incorporated near the N termini and a fluorescence quencher near the C termini. For parallel CCs (top left) this should lead to fluorescence, whereas for antiparallel CCs (top right) this should lead to quenching. For antiparallel heteromeric CCs, incorporating the fluorophore and fluorescence quencher near the same termini (bottom left) should lead to fluorescence, whereas incorporating at opposite termini should lead to quenching (bottom right).



Extended Data Fig. 7 Extended plots for the fluorescence-quenching assay for labelled apCC-Di, apCC-Di-AB containing additional controls. (left) Fluorescence quenching assay for homomeric peptides. Controls include the parallel coiled coil CC-Di and swapping the fluorophore and the fluorescence quencher between the termini of apCC-Di. (right) Fluorescence quenching assay for the designed heteromeric peptides. The control peptides include labelling both apCC-Di-A and apCC-Di-B near the C-termini, which does not lead to quenching in an antiparallel orientation. Key: n and c indicate mutations near the N and C termini, respectively; 4CF, 4-Cyano-L-phenylalanine fluorophore; MSE, L-selenomethionine fluorescence quencher. Conditions: 100 μM concentration of each peptide, 50 mM sodium phosphate, pH 7.



Extended Data Fig. 8 Extended bimolecular fluorescence complementation (BiFC) assay containing controls for homomerisation of apCC-Di-A and apCC-Di-B, and controls for heteromerisation of CC-Di-A and CC-Di-B (a parallel heteromeric system). BiFC assay for apCC-Di and apCC-Di-AB. V1 and V2 are N- and C-terminal fragments of the Venus yellow fluorescent protein, where V1 represents the V1 mutant. Peptide names that precede the Venus fragment name denote fusion to the C termini of the coiled coil, for example apCC-Di-B-V2. Values are normalised according to cell density (OD600) and are presented as mean values±1 SD from n ⩾ 3 technical replicate measurements (dots).



Extended Data Fig. 9 Cell penetration of apCC-Di-B occurs at 4 °C and in the presence of endocytosis inhibitors. a, Representative confocal images of HeLa cells treated for 1 h with 2 µM TAMRA-labelled

apCC-Di-B at 37 oC and 4 °C. These representative images are taken from n = 3 biological replicates. **b**, The same for cells at 37 °C pre-incubated with 0.1% DMSO (as a control) and with endocytosis inhibitors 30 μM MiTMAB, 80 μM Dynasore, and 0.1% NaN3 in the DMEM before treatment with 2 μM TAMRA labelled apCC-Di-B peptide in media containing each of the inhibitors for 1 h. apCC-Di-B remains cell penetrating at 4 °C and in the presence of the inhibitors and at 37 °C, as shown by TAMRA fluorescence in the cytoplasm and nucleus. Scale bar 10 μm. These representative images are taken from n = 3 biological replicates.