SUPPORTING INFORMATION

Crystal Structure of a Protein – Aromatic Foldamer Composite: Macromolecular Chiral Resolution

Jimi M. Alex,[a] Valentina Corvaglia,[b,c] Xiaobo Hu,[b] Sylvain Engilberge,[a] Ivan Huc,*[,b,c] Peter B. Crowley*[,a]

[a]School of Chemistry, National University of Ireland, University Road, Galway, H91 TK33, Ireland.

[b]University of Bordeaux – CNRS – IPB, CBMN Laboratory (UMR5248), Institut Européen de Chimie et Biologie, Pessac, France.

[c]Department of Pharmacy, Ludwig-Maximilians-Universität, Butenandtstraße 5, München, Germany.

*Correspondence to:

ivan.huc@cup.lmu.de / +49 89 218 077 804

peter.crowley@nuigalway.ie / +353 91 49 24 80

Keywords

Composite material, lysine, molecular glue, protein assembly, quinoline

Materials and Methods

General procedures for synthesis and characterization

Chemicals and reagents were used as commercially supplied without any further purification unless otherwise stated. Low loading Wang resin (0.41 mmol/g) was purchased from Novabiochem. Ghosez reagent (1-chloro-*N*,*N*-2-trimethyl-1-propenylamine) was purchased from Sigma Aldrich. *N,N*-diisopropylethylamine (DIEA) was distilled over calcium hydride. Reactions requiring anhydrous conditions were performed under nitrogen.

Analytical grade organic solvents were used for solid phase synthesis. Anhydrous THF and CH_2Cl_2 for solid phase synthesis were dispensed from an MBRAUN SPS-800 solvent purification system. RP-HPLC quality acetonitrile and Milli-Q water were used for RP-HPLC analyses and purification.

RP-HPLC analyses were performed at 1.0 mL min^{-1} by using a Macherey–Nagel Nucleodur C18 HTec column (4 x 100 mm, 5 µm). The mobile phase was composed of NH_4OAc (12 mM, pH 8.5, solvent C) and CH_3CN (solvent D). Monitoring was performed by UV detection at 214, 254 and 300 nm with a diode array detector. Semi-preparative purifications of oligomers were performed at 5 mL min^{-1} by using a Macherey-Nagel Nucleodur C18 HTeC column (10 x 125 mm, 5 µm). The mobile phase was the same as for the analytic system. Monitoring was performed by UV detection at 300 nm.

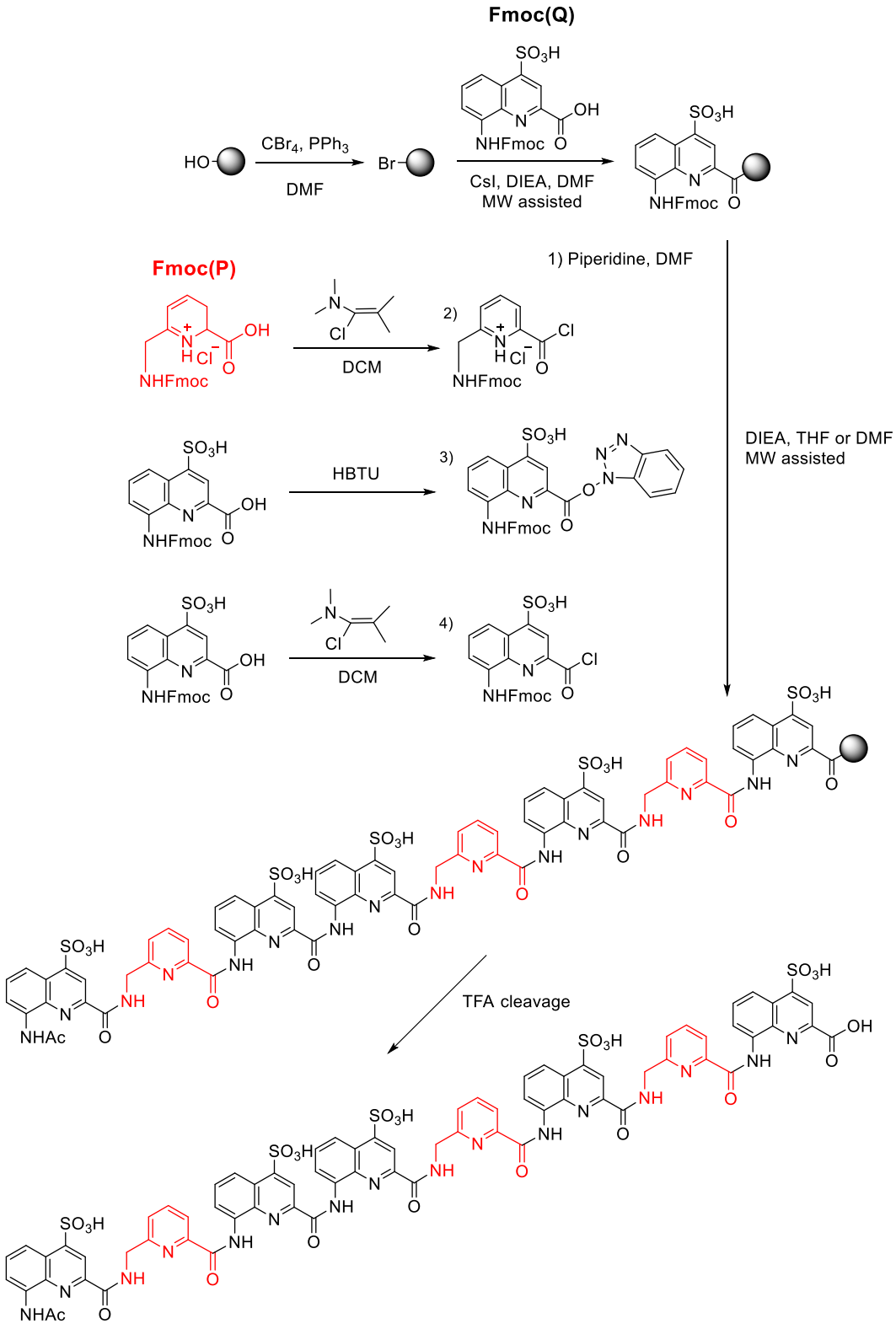
NMR spectra were recorded on Avance III HD 400 MHz NMR spectrometer from Bruker BioSpin equipped with a PABBO BB19F-1H/D broadband probe with Z-GRD. Data processing was performed with Bruker TOPSPIN 2.1 software.

High resolution mass spectrometry (HRMS) was used for the ionization of oligomer **2** in negative ion modes. ESI mass detection was performed on a Thermo Finnigan LTQ FT Ultra Fourier Transform Ion Cyclotron Resonance mass spectrometer. The resolution was 100.000 FWHM at m/z 400 with a scan repetition rate of 1 second, while acquiring a mass range from m/z 200 to 2000. The ESI voltage was 4 kV, the heater capillary temperature 250 °C, the nitrogen sheath gas flow 20 arbitrary units and the sweep gas 5 arbitrary units.

Monomer synthesis

Fmoc (**Q**) and **Fmoc** (**P**) monomers were synthesized as previously described.[1,2]

Oligomer synthesis



Scheme S1. Solid Phase Synthesis of oligomer **2**.

Oligomer **1** was prepared as reported before.[1] Oligomer **2** was prepared (Scheme S1) by using the coupling, Fmoc deprotection and terminal N-acetylation methods recently described,[2] i.e. using acid chloride activation to couple acids on aromatic amines and HBTU activation to couple acids on aliphatic amines. Crude **2** was obtained in 60 % purity. The compound was purified by semi-prep HPLC; Gradient: 0-100% of solvent D in 20 min. Final oligomer was 97 % pure by HPLC at 300 nm and MS confirmed its composition: ESI MS calculated for $C_{73}H_{51}N_{16}O_{25}S_5$ $[M-H]^-$: 1711.1893; Found: 855.5961 $[M-2H]^{2-}$, 427.2924 $[M-4H]^{4-}$, 341.4305 $[M-5H]^{5-}$ (Figure S1-S3).

Sample preparation

Millimolar stock solutions of **1** were prepared in 1 M sodium acetate or water whereas foldamer **2** was prepared in water. ^{15}N-labelled and unlabelled *Saccharomyces cerevisiae* cytochrome *c* (C102T) was prepared as described previously.[3,4]

Co-crystallization trials

Trials were performed with an Oryx 8 Robot (Douglas Instruments) and a sparse matrix screen (JCSG++, Jena Bioscience) at 20° C. Cytochrome *c* (1 mM) and **1** (in sodium acetate) were tested in ratios of 1:1 and 1:4. Crystal growth was achieved with 4 mM of **1** only in condition C7, 10 % PEG 3000, 0.1 M sodium acetate pH 4.5, 0.2 M zinc acetate. These rod-shaped crystals were reproducible by the hanging drop method in 5 - 25 % PEG 3350. Trials with **1** in water yielded thin plates in conditions A9 (20 % PEG 3350, 0.2 M ammonium chloride pH 6.3), G7 (15 % PEG 3350, 0.1 M di-sodium succinate pH 7.0) and G8 (20 % PEG 3350, 0.1 M di-sodium DL-malate pH 7.0). Thin plates were obtained with **2** (in water) in conditions G7 and H6 (17 % PEG 10,000, 0.1 M Bis-TRIS pH 5.5, 0.1 M ammonium acetate).

X-ray data collection

Crystals of ~100 µm dimension were cryo-protected in the reservoir solution supplemented with 20 % glycerol and cryo-cooled in liquid nitrogen. The rod-shaped crystals of **1** at 5 % PEG 3350 diffracted to 2.1 Å resolution at beamline PROXIMA-2A (SOLEIL synchrotron). Datasets were collected using ϕ scans of 0.1° over 360° with an Eiger X 9M detector. The thin plates with **1** or **2** diffracted poorly.

X-ray structure determination

Diffraction data were integrated and scaled using the autoPROC pipeline.[5] The structure was solved by molecular replacement in PHASER, with 5LYC as the search model.[6] JLigand was used to generate the foldamer coordinates and refinement restraints.[7] Iterative cycles of model building in COOT[8] and refinement in BUSTER[9] were performed until no further improvements in the R_{free} or electron density

were obtained. The final structure was validated with MolProbity[10] and deposited in the Protein Data Bank as PDB 6s8y.

Circular dichroism

CD spectra were recorded on a Jasco J-810 spectrometer using quartz cells with a 1 mm path length.[2] Scans were acquired at 20 °C, over the 300-500 nm range, with a 1 s response time and a 100 nm/min scan speed. The CD data (Figure 3) are an average of three scans. Samples contained 90 µM cytochrome *c* and 1 eq. **2** in 25 mM sodium acetate, 25 mM sodium chloride at pH 5.4 and were equilibrated for a minimum of 24 h. Data were baseline-corrected for signal contributions due to the buffer and protein.

NMR spectroscopy

^{1}H-^{15}N HSQC monitored titrations were performed at 30°C on a 600 MHz Varian spectrometer equipped with a HCN coldprobe, as described previously.[3] Samples of 0.1 mM ^{15}N-labelled cytochrome *c* were titrated with 3 - 24 µL aliquots of 8 mM ligand (in water) in 20 mM potassium phosphate, 50 mM sodium chloride at pH 6.0. Spectra were processed in NMRPipe and chemical shift perturbations were determined in CCPN.[11]

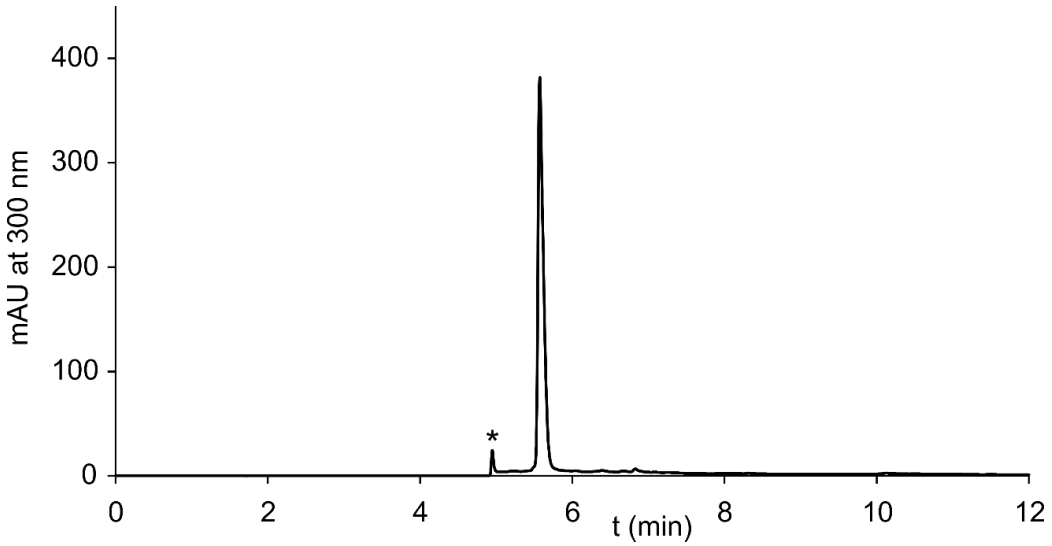


Figure S1. Analytical RP-HPLC profile of oligomer **2**. Elution conditions: Nucleodur C18 HTec column (4 x 100 mm, 5 µm); eluents C and D, where solvent C was NH_4OAc (12 mM, pH 8.5) and solvent D was CH_3CN; gradient program: from 0 % to 100 % of solvent D in 10 min; UV detection at λ = 300 nm. A small impurity from the solvent is denoted with by an asterisk.

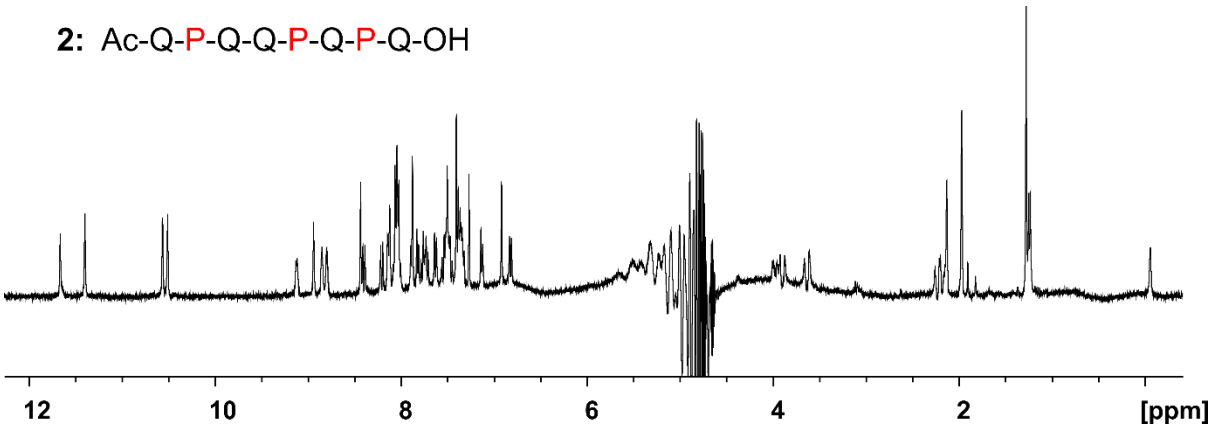


Figure S2. ^{1}H NMR spectrum (400 MHz) at 293 K of oligomer **2** in H_2O/D_2O 9:1 vol/vol, 50 mM NH_4HCO_3.

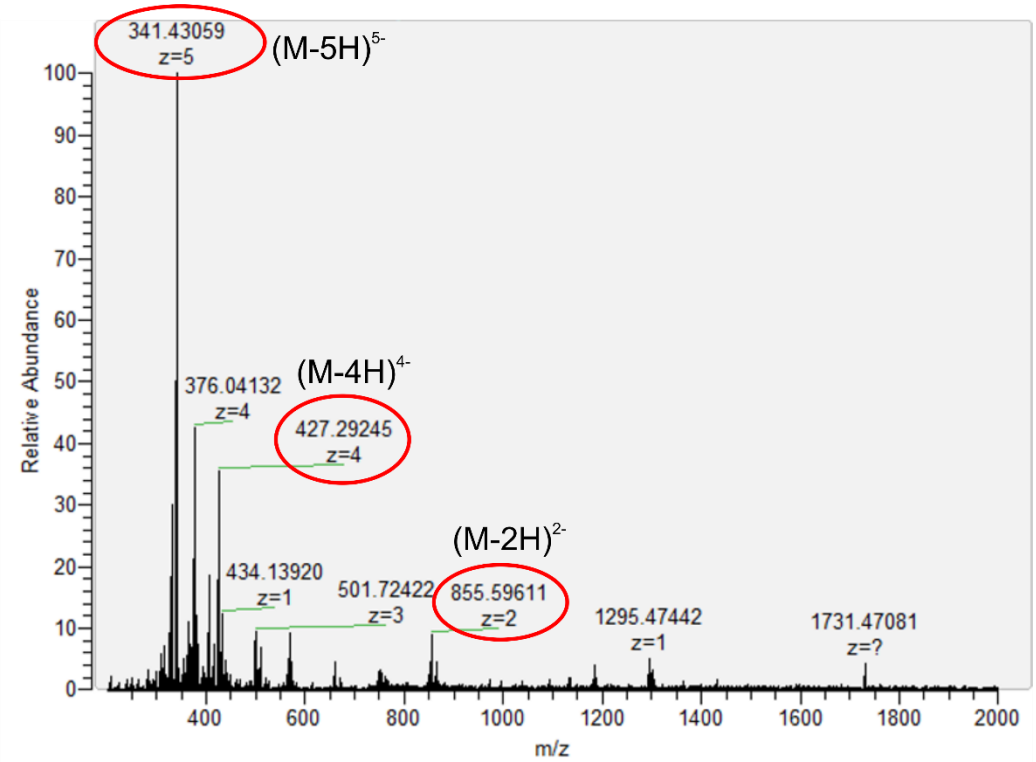


Figure S3. Multicharged species observed by HRMS ESI of oligomer **2**.

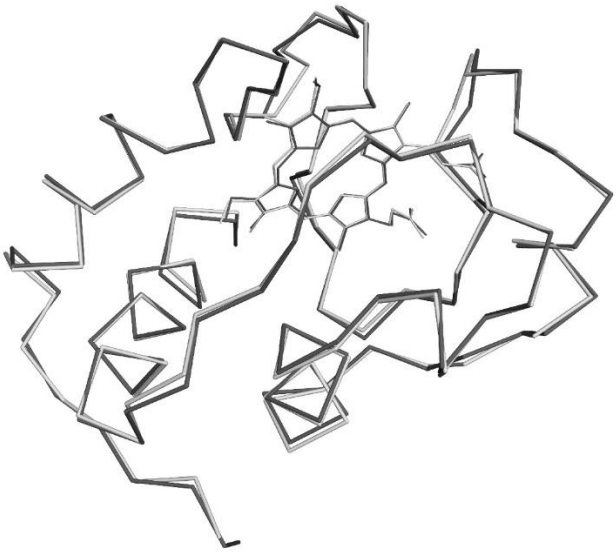


Figure S4. Superposition of the cyt*c* – foldamer **1** complex (light grey, PDB 6s8y) with ligand-free cyt*c* (dark grey, PDB 1ycc) revealed minor structural changes, with an average C^{α} RMSD of 0.4 Å. Only the protein C^{α} traces and haem groups are shown.

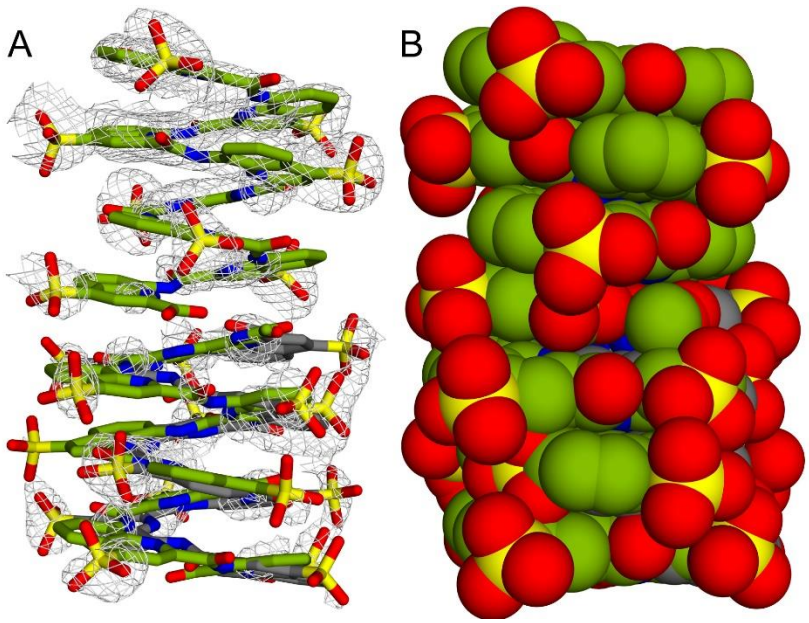


Figure S5. A. Omit map showing the $2F_o$ - F_c electron density (contoured at 1.0 σ) of the foldamers. The grey foldamer corresponds to the head-to-head orientation. **B.** Spacefill representation of the foldamer stack depicting the cylinder with an anionic surface (hydrophobic quinoline rings are buried).

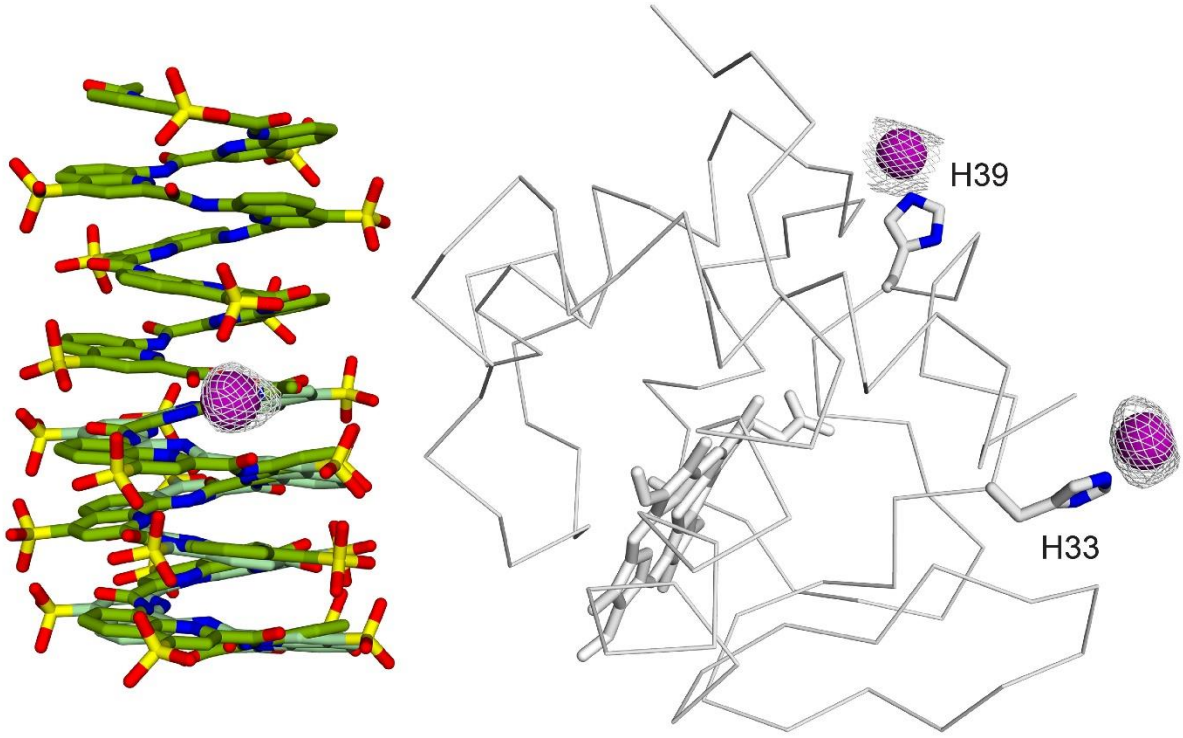


Figure S6. Fourier map (contoured at 3.0 σ) confirmed the presence and position of three zinc ions.

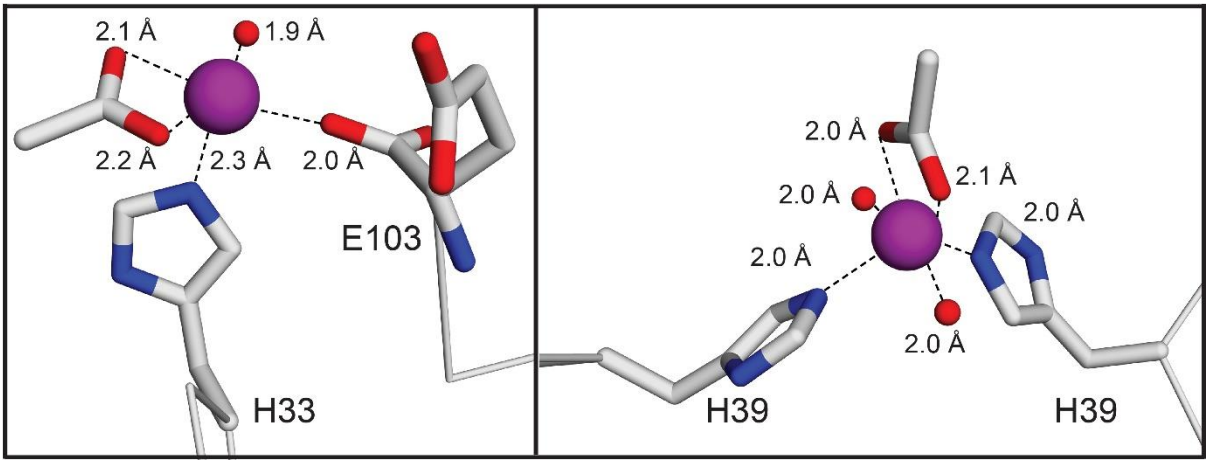


Figure S7. Zinc ions coordinated at His33 or His39 with acetate ions and waters completing the coordination spheres. At His39, the zinc is a bridging ion that contributes to protein assembly in the crystal.

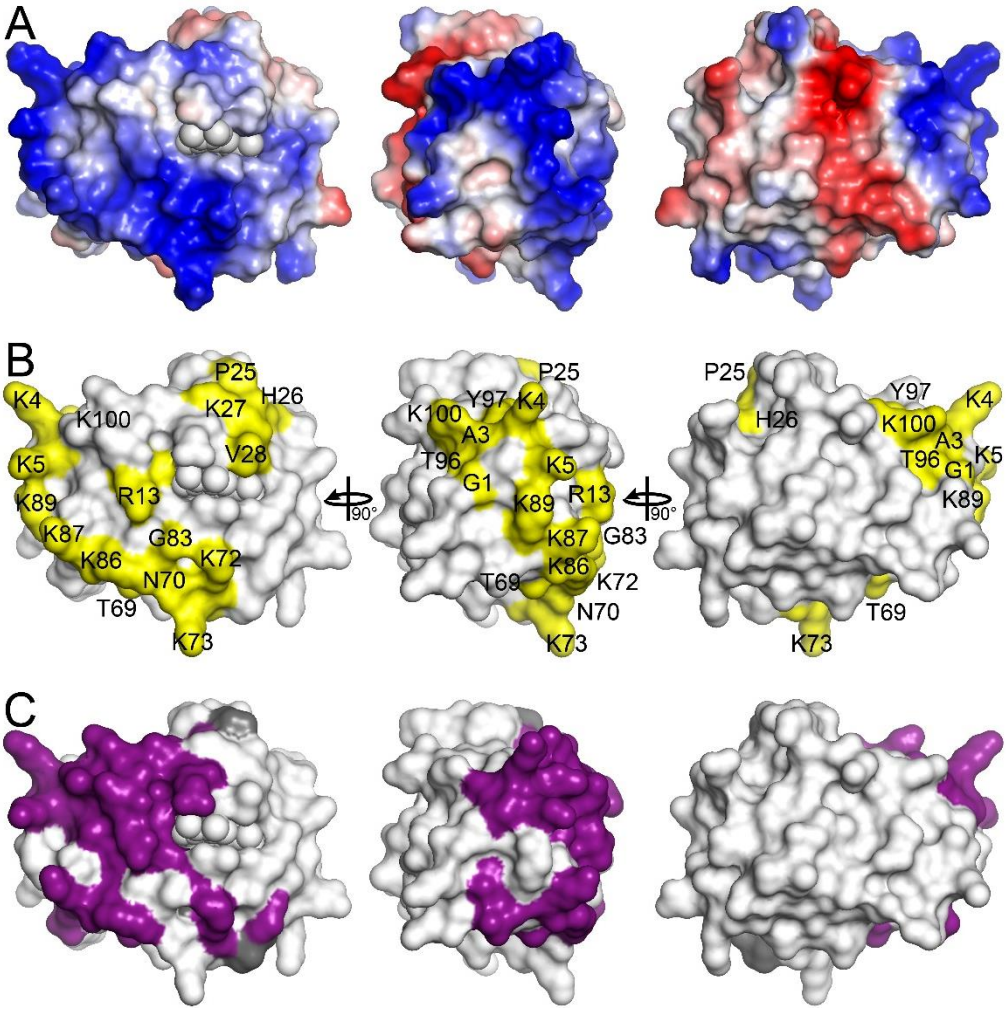


Figure S8. Surface representations of cyt*c* illustrating: **(A)** the electrostatic potential with cationic (blue) and anionic (red) patches, **(B)** the crystallographic binding sites with residues coloured yellow if at least one atom was ≤ 4.5 Å from foldamer **1**, and **(C)** the NMR binding site with residues coloured purple if $\Delta\delta\ {}^{1}H^{N} \geq 0.04$ and/or ${}^{15}N \geq 0.4$, at 0.4 mM foldamer **1**. Prolines are dark grey.

Table S1. X-ray data collection, processing and refinement statistics for cyt*c* – **1**.

Data Collection	
Light source	SOLEIL, PROXIMA-2A
Wavelength (Å)	0.98009
Space group	$P3_221$
Cell constants	$a = b = 39.13$ Å, $c = 187.36$ Å
	$\alpha = \beta = 90^o$, $\gamma = 120^o$
Resolution (Å)	62.45-2.09 (2.12-2.09)
# reflections	199568 (10386)
# unique reflections	10590 (521)
Multiplicity	18.8 (19.9)
I/σ (I)	19.0 (2.3)
Completeness (%)	100 (100)
R_{meas}[b] (%)	8.2 (12.7)
R_{pim}[c] (%)	2.0 (28.3)
$CC_{1/2}$	99.7 (85.6)
Solvent content (%)	50
Refinement	
R_{work}	0.2290
R_{free}	0.2780
rmsd bonds (Å)	0.0101
rmsd angles (°)	1.5249
Asymmetric unit composition	
protein	1
foldamer	2
zinc ion	3
water	26
Ave. B-factor (Å^2)	67.41
Clashscore	2.71
Ramachandran analysis,[d] % residues in	
favoured regions	97.1
allowed regions	100
PDB code	6s8y

[a]Values in parentheses correspond to the highest resolution shell

[b]$R_{meas} = \sum_{hkl} \sqrt{(n/n-1)} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$

[c]$R_{pim} = \sum_{hkl} \sqrt{(1/n-1)} \sum^n_{i=1} |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$

[d]Determined in MolProbity

References

1. X. Hu, S. J. Dawson, P. K. Mandal, X. de Hatten, B. Baptiste, I. Huc, *Chem. Sci.,* **2017**, *1*, 3741.
2. M. Vallade, P. S. Reddy, L. Fischer, I. Huc, *Eur. J. Org. Chem.* 2018, **2018,** 5489.
3. R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power, P. B. Crowley, *Nat. Chem.,* **2012**, *4***,** 527.
4. P. B. Crowley, P. Ganji and H. Ibrahim, *ChemBioChem,* 2008, **9**, 1029.
5. C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack, G. Bricogne, *Acta Crystallogr. D,* 2011, **67**, 293.
6. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *J. Appl. Crystallogr.,* 2007, **40**, 658.
7. E. Potterton, P. Briggs, M. Turkenburg, E. A Dodson, *Acta Crystallogr. D,* 2003, **59***,* 1131.
8. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallogr. D,* 2010, **66**, 486.
9. O. S. Smart, T. O.Womack, C. Flensburg, P. Keller, W. Paciorek, A. Sharff, C. Vonrhein and G. Bricogne, *Acta Crystallogr. D,* 2012, **68**, 368.
10. V. B. Chen, W. B. Arendall, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr. D,* 2010, **66**, 12.
11. W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins,* 2005, **59**, 687.