Crystal structures capture multiple stoichiometric states of an aqueous self-assembling oligourea foldamer

Gavin W. Collie,[a*] Caterina M. Lombardo,[b†] Sung Hyun Yoo,[b†] Karolina Pułka-Ziach,[c] Valérie Gabelica,[d] Cameron D. Mackereth,[d] Frédéric Rosu[e] and Gilles Guichard[b*]

a. Discovery Sciences, R&D, AstraZeneca, Cambridge, UK.
b. Univ. Bordeaux, CNRS, Bordeaux INP, CBMN, UMR 5248, Institut Européen de Chimie et Biologie, 2 rue Robert Escarpit, 33607, Pessac, France.
c. University of Warsaw, Faculty of Chemistry, Pasteura 1, 02-093, Warsaw, Poland.
d. Univ. Bordeaux, CNRS, INSERM, ARNA, UMR 5320, U1212, IECB, F-33600 Bordeaux, France.
e. Univ. Bordeaux, CNRS, INSERM, IECB, UMS 3033, F-33600 Pessac, France.
† These authors contributed equally to this work

* Email: gavin.collie@astrazeneca.com , g.guichard@iecb.u-bordeaux.fr

Electronic Supplementary Information (ESI) available: full experimental details and additional figures. See DOI: 10.1039/x0xx00000x

Abstract : We report here an oligourea foldamer able to self-assemble in aqueous conditions into helix bundles of multiple stoichiometries. Importantly, we report crystal structures of several of these stoichiometries, providing a series of high-resolutions snap-shots of the structural polymorphism of this foldamer and uncovering a novel self-assembly.

There is currently considerable interest in the development of oligomers and polymers able to self-assemble into precise, well-ordered three-dimensional structures in aqueous conditions. While there has been significant progress with respect to self-assembling systems based on natural backbones – such as DNA[1] and α-peptides/proteins[2, 3] – much remains to be achieved in the design of aqueous self-assembling systems based on fully non-natural folded backbones[4] (i.e. biotic or abiotic foldamers[5]). A number of foldamers have been shown to self-assemble in non-aqueous conditions (i.e. in organic solvent), including intertwined multiple helices,[6-9] helix bundles[10] and multi-stranded sheets formed from abiotic oligomers.[11] However, few reports have described the aqueous self-assembly of foldamers with atomic precision. Among abiotic foldamers, the association of water-soluble aromatic helical foldamers into dimeric helices[12] and the crystal structure of nanometric cage formed by the self-assembly of a short amphiphilic arylamide oligomer[13] are particularly notable. Atomically precise nano-sized structures formed from biotic foldamer backbones are also rare and are limited to fibril-like assemblies of macrocyclic β-sheet mimics,[14] tetrameric helix bundles formed by α/β hybrid peptides,[15, 16] and an eight-helix bundle formed from amphiphilic helical β-peptides.[17, 18]

We have previously reported the strong propensity of amphiphilic aliphatic oligoureas (a class of peptidomimetic foldamer) bearing proteinaceous side-chains to self-assemble into protein-like

architectures in a controllable manner in aqueous conditions.[19-21] In particular, we have described a six-helix bundle formed from the self-assembly of a zwitterionic oligourea foldamer, referred to as '**H1**'.[21] This foldamer helix bundle, however, displays a molten-globule-like character (with a modest T_m of 46 °C) and we were interested in improving the thermal stability of the **H1** helix bundle to enable further functional development. Here, we report the structure-guided design and in-depth characterisation of an analogue of the zwitterionic **H1** foldamer in which the formal charge was increased to +1. While this modification did not improve bundle thermal stability as hoped, a series of crystal structures allowed us to capture atomic-scale snap-shots of different stoichiometric states of this new foldamer, one of which represents a novel aqueous foldamer self-assembly with promising features for potential functional development.

Inspection of the previously reported crystal structure of the **H1**[21] six-helix bundle revealed a number of possible regions of the foldamer that could be amenable to engineering with the aim of improving the thermal stability of the assembly. In particular, the distribution and orientation of the charged side-chains appeared sub-optimal. Indeed, despite **H1** bearing four chargeable side-chains – two lysine-type (Lys^u)[22] and two glutamate-type (Glu^u) – no salt-bridges were evident in the crystal structure of the **H1** six-helix bundle. In addition, potential charge clashes (of the Glu^u7 residues) were also evident from this crystal structure (Fig. S1), suggesting that the assembly is unlikely to be favoured at neutral or higher pH (i.e. at a pH above the pKa of the side-chain carboxylates). In order to address this issue, we designed and synthesised an analogue of **H1** with residue Glu^u7 mutated to a glutamine-type residue (Gln^u7), resulting in a foldamer with a formal charge of +1, consequently termed **H1**$^+$ (Figs. 1a and S2). **H1**$^+$ was synthesised on solid support following previously reported procedures.[21, 23]

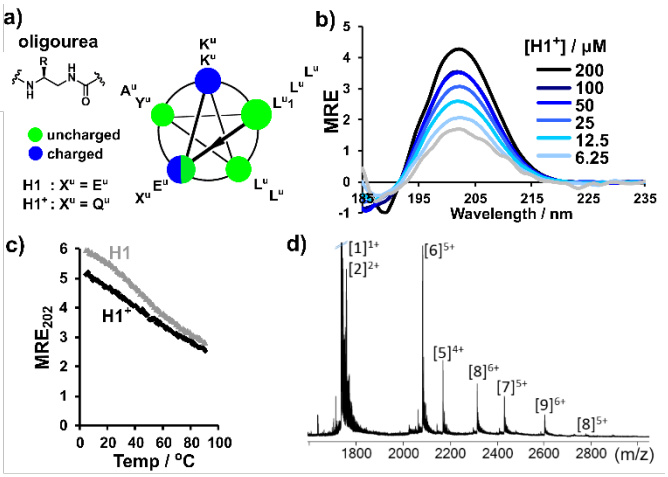

Fig. 1 a) Details of oligoureas **H1** (previously reported[21]) and **H1**$^+$ (reported here). b) Variable-concentration circular dichroism (CD) analysis of **H1**$^+$. MRE: molar residual ellipticity (deg·cm^2·dmol^{-1}·residue^{-1}). MRE values divided by 10^4 for clarity. c) CD-monitored thermal melting profiles of **H1** (grey) and **H1**$^+$ (black) at 200 µM. CD data shown for **H1** reported previously.[21] d) Analysis of **H1**$^+$ at 50 µM in pure water by native mass spectrometry.

Preliminary assessment of the solution behaviour of **H1**$^+$ by circular dichroism (CD) confirmed this molecule to fold into a helical conformation in aqueous conditions as shown by a positive peak at 202 nm, with variable-concentration CD analysis showing further that this foldamer displayed a tendency to self-assemble in aqueous solution comparable to that of **H1** (Figs. 1b and S3). We then assessed the thermal stability of **H1**$^+$ by CD-monitored thermal melting experiments. Disappointingly however, the thermal melting profile of **H1**$^+$ did not indicate an increase in T_m (relative to **H1**) (Fig. 1c). We next analysed **H1**$^+$ by native electrospray ionisation mass spectrometry (ESI-MS), with spectra revealing a variety of stoichiometries ranging from 5- to 9-mers (Fig. 1d), with a hexameric species appearing to be dominant, again, in line with equivalent analysis of **H1**.[21]

We next turned to X-ray crystallography to provide atomic-scale details of the assemblies indicated by the above biophysical data. Fortunately, we were able to grow crystals of **H1**$^+$ suitable for X-ray diffraction studies using standard (aqueous) protein crystallisation methods, with crystals obtained from multiple crystallisation conditions (Table S1). Diffraction data collected for a crystal grown from a crystallisation solution composed of 10 % isopropanol, 200 mM $CaCl_2$ and 100 mM sodium acetate (pH 4.6) yielded a 1.15 Å crystal structure (belonging to space group $P6_3$), revealing a six-helix bundle arrangement (Fig. 2a) highly similar to that formed by **H1** in similar conditions[21] (Fig. S4). We also collected high-quality diffraction data for **H1**$^+$ crystals grown from alternative crystallisation conditions (25 % PEG2000, 10 mM nickel (III) chloride and 100 mM tris-HCl (pH 8.5)) which appeared to belong to an alternative space group ($P2_12_12_1$) (Table S1). Subsequent structure determination of this second crystal form yielded a surprisingly different assembly, with **H1**$^+$ forming an eight-helix bundle under these conditions (Fig. 2b). Despite clear differences in bundle stoichiometry, size and geometry, the six-helix and eight-helix bundles formed by **H1**$^+$ are constructed in a remarkably similar fashion – indeed, both helix bundles are constructed from the same type of helix dimer (Fig. 2). This dimer involves the anti-parallel arrangement of two helices held together by a combination of conserved hydrophobic Leu^u-Leu^u interactions and a mirrored hydrogen bond between the amine of the Lys^u8 side-chain and the penultimate backbone urea carbonyl. The **H1**$^+$ six-helix and eight-helix bundles are then constructed from the side-to-side packing of three or four of these dimers, respectively.

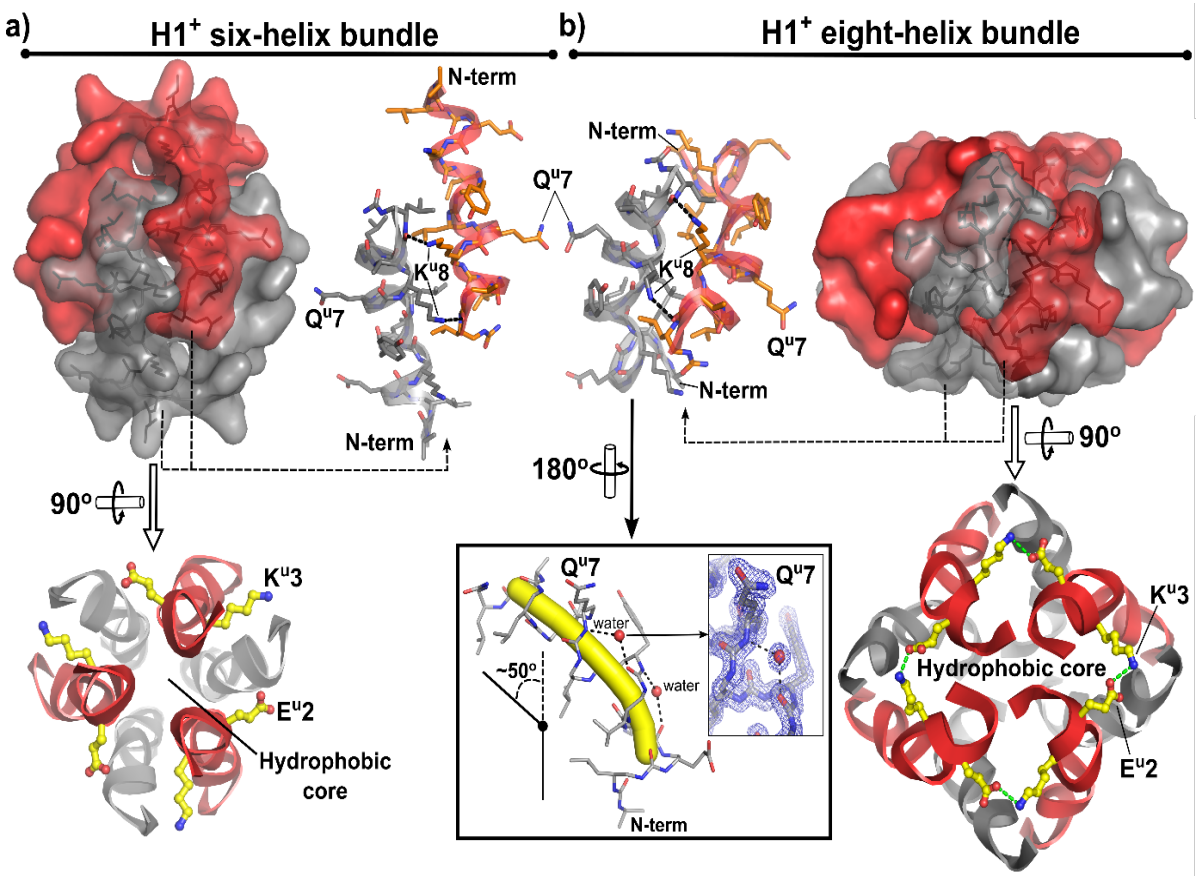

Fig. 2 H1⁺ crystal structures revealing six-helix (a) and eight-helix (b) bundle arrangements. Helix dimer repeat unit of helix bundles is shown as stick and ribbon representation to the side of each bundle. In ribbon cartoon view of helix bundles, residues Glu^{u}2 (E^{u}2) and Lys^{u}3 (K^{u}3) are shown as yellow sticks, with electrostatic contacts shown as green dashes. Inset: single **H1⁺** helix from eight-helix crystal structure highlighting the role of water molecules (red spheres) in the bent helix geometry, with electron density shown to the right (1.7 Å $2F_o$-F_c map at 1.5 σ level).

In addition to precise hydrophobic packing interactions, electrostatic contacts also play a major role in the folding of the **H1⁺** eight-helix bundle (Fig. 2b, bottom right panel), contrasting starkly with the six-helix bundle formed from **H1⁺** (and **H1**[21]), within which no electrostatic contacts are evident (Fig. 2a, bottom left panel). The salt-bridges of the **H1⁺** eight-helix bundle (eight in total) are all inter-helical and are formed between the Glu^{u}2 and Lys^{u}3 side-chains of neighbouring helices. Interestingly, in order for these salt-bridges to form, the helices need to 'bend' significantly in a directional manner. Indeed, all helices of the **H1⁺** eight-helix bundle do not conform to the expected canonical oligourea helix geometry, but instead are all 'bent' by around 50° (Fig. 2b, inset). This helical bending results in the insertion of ordered water molecules into the subsequently enlarged helix clefts, and, importantly, the helical hydrogen bonding is not broken, with the water molecules bridging the carbonyl oxygen of residue *i* to the backbone NHs of residues *i* + 2 and *i* + 3.[24]

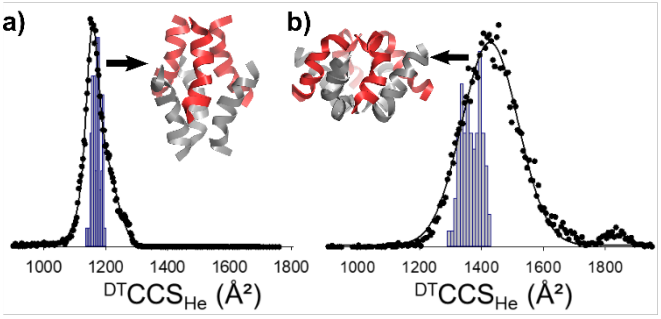

Fig. 3. Collision cross sections ($^{DT}CCS_{He}$) of the **H1**$^{+}$ $[6]^{5+}$ (a) and $[8]^{6+}$ (b) species measured by drift tube ion mobility ESI-MS (black circles) with theoretical CCSs calculated from snapshots sampled from molecular dynamics simulations (histograms) using the crystal structures (shown as inset cartoons) as starting

We were then curious as to whether the **H1**$^{+}$ crystal structures of the six- and eight-helix bundles correlated with the hexameric and octameric species observed by ESI-MS (Fig. 1d). We therefore used drift-tube ion mobility spectrometry to measure experimental collisional cross sections (CCS) in helium for the $[6]^{5+}$ and the $[8]^{6+}$ species observed by native ESI-MS, and compared these experimental CCS distributions with theoretical values. Ensembles of gas-phase structures were generated using molecular dynamics at the PM7 semi-empirical level[25] using the X-ray derived coordinates for the six-helix and eight-helix bundles described above as starting models. The theoretical CCS values in helium were calculated with the trajectory method.[26] Gratifyingly, the theoretical and experimental CCS distributions for both the **H1**$^{+}$ six-helix and eight-helix bundles correlate very well (Fig. 3), strongly implying the **H1**$^{+}$ self-assemblies observed crystallographically to match those transferred from the solution to the gas-phase by native ESI-MS.

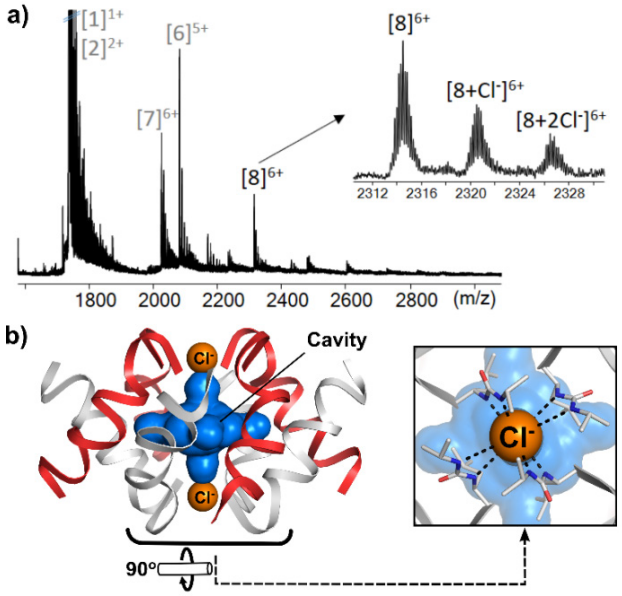

Fig. 4. a) Analysis of **H1**$^{+}$ by ESI-MS in the presence of 2 % sulfolane (plus 100 mM ammonium acetate) revealing the presence of octameric species ($[8]^{6+}$) bound by one and two chloride ions (inset). b) Crystal structure of **H1**$^{+}$ eight-helix bundle bound by two chloride ions (orange spheres). The internal cavity (volume ~480 Å^3) is shown as a blue surface.

In addition to the native ESI-MS experiments performed above, **H1**$^+$ was further analysed by ESI-MS in the presence of 2 % sulfolane (a polar non-volatile additive that typically "supercharges" proteins[27]) in an attempt to generate higher charge states and probe the resistance of the assemblies to Coulomb repulsion. While the addition of 2 % sulfolane did not exert a supercharging effect on **H1**$^+$, we noted new peaks in the spectra corresponding to an **H1**$^+$ octamer species ([8]$^{6+}$) with 1 and 2 chloride ion adducts (Fig. 4a). Interestingly, we were able to obtain a third crystal form of **H1**$^+$,[28] revealing an eight-helix bundle isomorphous with that described above, but differing in the surprising coordination of two chloride ions within the bundle – each ion octahedrally bound by eight free N-terminal NHs of four helices (Fig. 4b). These chlorides effectively 'close' the internal void space of the bundle, creating an enclosed cavity with a volume of 479.7 Å^3,[29] and offering a theoretical means of controlling solvent access to the cavity. CCS values were then determined experimentally for the [8+2Cl$^-$]$^{6+}$ species (Fig. S5) and were seen to closely match those of the crystal structure of the chloride-bound (and chloride-free) **H1**$^+$ eight-helix bundles.

Finally, we used high-field NMR in order to understand these data and the general self-assembling properties of **H1**$^+$ further. A comparison of the crystal structures predicts that the **H1**$^+$ eight-helix bundle would present mainly similar NOE patterns as the **H1** six-helix bundle, and indeed similarity to **H1** is observed in the NOESY analysis of **H1**$^+$ (Figs. 5 and S6). One notable difference in the **H1**$^+$ eight-helix bundle structure concerns Tyr^u: this sidechain forms a close intermolecular contact with Leu^u methyls in the six-helix bundles formed by **H1** and **H1**$^+$, but in the **H1**$^+$ eight-helix bundle the contact is only close for four of the eight Tyr^u residues. Nevertheless, the Tyr^u-Leu^u cross-peak intensities for **H1**$^+$ are close to those observed for **H1** (instead of 50 %), suggesting a significant population of the **H1**$^+$ six-helix bundle coexists with the **H1**$^+$ eight-helix bundle. We therefore find that **H1**$^+$ forms a stable assembled equilibrium in a mix of oligomeric states, consistent with the populations observed by mass spectrometry.

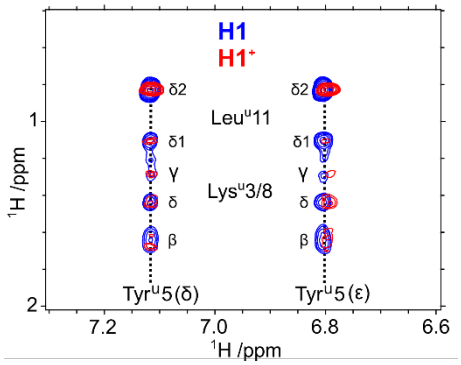

Fig. 5. NOESY analysis of **H1**$^+$ (2 mM in D_2O[30]) and comparison to **H1**. NB δ1 and δ2 have not been stereospecifically assigned. See Fig. S6 for further data and experimental details.

We have reported here an oligourea foldamer (**H1**$^+$) able to self-assemble in aqueous conditions into helix bundles of multiple stoichiometries. Importantly, we have been able to determine

crystal structures of several of these stoichiometries, providing a series of high-resolution snap-shots of the structural polymorphism of this foldamer. Furthermore, the novel eight-helix bundle formed by **H1**$^{+}$ adds to the currently limited number of reports describing aqueous foldamer quaternary structure, and further possesses several unexpected yet promising features inducive for future functional development, such as substrate binding within an enclosed cavity,[31] and perhaps even eventual catalysis.[32, 33]

We thank the SOLEIL and ESRF synchrotrons for providing access to beamlines PX1, ID29 and ID23-2. A post-doctoral fellowship to S.H.Y. from IdEx Bordeaux (ANR-10-IDEX-03-02), a program of the French government managed by the Agence Nationale de la Recherche and Marie Curie FP7-PEOPLE-2010-IEF-273224 and FP7-PEOPLE-2012-IEF-330825 postdoctoral fellowships to K.P.-Z. and C.M.L. are gratefully acknowledged. This work has benefited from the IECB Biophysical and Structural Chemistry platform (BPCS), CNRS UMS3033, Inserm US001, Univ. Bordeaux.

Notes and references

1. a) Y. Zhang, J. Tu, D. Wang, H. Zhu, S. K. Maity, X. Qu, B. Bogaert, H. Pei and H. Zhang, *Adv. Mater.*, 2018, **30**, e1703658; b) S. Nummelin, J. Kommeri, M. A. Kostiainen and V. Linko, *Adv. Mater.*, 2018, **30**, e1703721.
2. W. M. Dawson, G. G. Rhys and D. N. Woolfson, *Curr. Opin. Chem. Biol.*, 2019, **52**, 102-111.
3. P. Makam and E. Gazit, *Chem. Soc. Rev.*, 2018, **47**, 3406-3420.
4. W. S. Horne and T. N. Grossmann, *Nat. Chem.*, 2020, **12**, 331-337.
5. G. Guichard and I. Huc, *Chem. Commun.*, 2011, **47**, 5933-5941.
6. V. Berl, I. Huc, R. G. Khoury, M. J. Krische and J. M. Lehn, *Nature*, 2000, **407**, 720-723.
7. Y. Ferrand, A. M. Kendhale, J. Garric, B. Kauffmann and I. Huc, *Angew. Chem. Int. Ed.*, 2010, **49**, 1778-1781.
8. D. Haldar and C. Schmuck, *Chem. Soc. Rev.*, 2009, **38**, 363-371.
9. Y. Tanaka, H. Katagiri, Y. Furusho and E. Yashima, *Angew. Chem. Int. Ed.*, 2005, **44**, 3867-3870.
10. S. De, B. Chi, T. Granier, T. Qi, V. Maurizot and I. Huc, *Nat. Chem.*, 2018, **10**, 51-57.
11. J. Atcher, P. Mateus, B. Kauffmann, F. Rosu, V. Maurizot and I. Huc, *Angew. Chem. Int. Ed.*, 2021, **60**, 2574–2577.
12. H. Goto, H. Katagiri, Y. Furusho and E. Yashima, *J. Am. Chem. Soc.*, 2006, **128**, 7176-7178.
13. V. Pavone, S. Q. Zhang, A. Merlino, A. Lombardi, Y. Wu and W. F. DeGrado, *Nat. Commun.*, 2014, **5**, 3581.
14. R. K. Spencer, H. Li and J. S. Nowick, *J. Am. Chem. Soc.*, 2014, **136**, 5595-5598.
15. W. S. Horne, J. L. Price, J. L. Keck and S. H. Gellman, *J. Am. Chem. Soc.*, 2007, **129**, 4178-4180.
16. W. S. Horne, J. L. Price and S. H. Gellman, *Proc. Natl. Acad. Sci. U. S. A.*, 2008, **105**, 9151-9156.
17. C. J. Craig, J. L. Goodman and A. Schepartz, *Chembiochem*, 2011, **12**, 1035-1038.
18. J. L. Goodman, E. J. Petersson, D. S. Daniels, J. X. Qiu and A. Schepartz, *J. Am. Chem. Soc.*, 2007, **129**, 14746-14751.
19. S. H. Yoo, G. W. Collie, L. Mauran and G. Guichard, *Chempluschem*, 2020, **85**, 2243-2250.
20. C. M. Lombardo, G. W. Collie, K. Pulka-Ziach, F. Rosu, V. Gabelica, C. D. Mackereth and G. Guichard, *J. Am. Chem. Soc.*, 2016, **138**, 10522-10530.
21. G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth and G. Guichard, *Nat. Chem.*, 2015, **7**, 871-878.

22. Ureido residues are denoted Xaa^{u} by analogy to the three letter code of the corresponding Xaa α-amino acid residues.
23. C. Douat-Casassus, K. Pulka, P. Claudon and G. Guichard, *Org. Lett.*, 2012, **14**, 3130-3133.
24. A similar phenomenon has been observed previously for an oligourea in non-aqueous conditions: J. Fremaux, C. Dolain, B. Kauffmann, J. Clayden and G. Guichard, *Chem. Commun*., 2013, **49**, 7415-7417.
25. J. J. P. Stewart, *J. Mol. Model.*, 2013, **19**, 1-32.
26. M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *J. Phys. Chem.*, 1996, **100**, 16082-16086.
27. a) H. Metwally, R. G. McAllister, V. Popa and L. Konermann, *Anal. Chem*., 2016, **88**, 5345-5354; b) R. R. Ogorzalek Loo, R. Lakshmanan and J. A. Loo, *J. Am. Soc. Mass. Spectrom*., 2014, **25**, 1675-1693.
28. Crystallisation conditions: 30 % PEG400, 200 mM sodium citrate and 100 mM Tris-HCl (pH 8.5).
29. Cavity volume determined using SURFNET: R. A. Laskowski, *J. Mol. Graph*., 1995, **13**, 323-330.
30. Dissociation experiments using ITC and NMR show the oligourea is fully assembled at 300 µM **H1**$^{+}$ in D_2O (Fig. S7)
31. G. W. Collie, R. Bailly, K. Pulka-Ziach, C. M. Lombardo, L. Mauran, N. Taib-Maamar, J. Dessolin, C. D. Mackereth and G. Guichard, *J. Am. Chem. Soc.*, 2017, **139**, 6128-6137.
32. Z. C. Girvin and S. H. Gellman, *J. Am. Chem. Soc.*, 2020, **142**, 17211-17223.
33. P. S. P. Wang, J. B. Nguyen and A. Schepartz, *J. Am. Chem. Soc.*, 2014, **136**, 6810-6813.