Electronic Supplementary Material (ESI) for Chemical Communications.

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

Gavin W. Collie, Caterina M. Lombardo, Sung Hyun Yoo, Karolina Pułka-Ziach, Valérie Gabelica, Cameron D. Mackereth, Frédéric Rosu and Gilles Guichard

Supplementary information

Contents:

Supplementary tables

Table S1. Data collection and refinement statistics for X-ray crystal structures of oligourea **H1**$^+$.

	Crystal form 1 Six-helix bundle	***Crystal form 2*** Eight-helix bundle	***Crystal form 3*** Eight-helix bundle
Crystallisation conditions	10 % isopropanol, 200 mM $CaCl_2$, 100 mM sodium acetate (pH 4.6)	20 % PEG2000, 10 mM nickel (II) chloride, 100 mM tris-HCl (pH 8.5)	30 % PEG400, 200 mM sodium citrate, 100 mM Tris-HCl (pH 8.5)
Data Collection			
Space group	*P* 6_3	*P* $2_12_12_1$	*P* 422
a, b, c (Å)	33.89, 33.89, 38.08	54.53, 59.37, 68.40	43.24, 43.24, 43.22
α, β, γ (°)	90.0, 90.0, 120.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	29.35 – 1.15 (1.21 – 1.15)	44.84 – 1.70 (1.81 – 1.70)	43.24 – 1.75 (1.85 – 1.75)
R_{meas} (%) overall	3.2 (82.0)	10.4 (69.9)	2.3 (61.4)
I / σ	23.53 (2.09)	11.86 (2.21)	38.63 (3.51)
Completeness (%)	99.8 (99.2)	98.6 (96.0)	99.6 (100.0)
Reflections (total)	47,565	119,929	35,801
Reflections (unique)	8,989	24,582	4,497
Redundancy	5.29 (4.82)	4.88 (4.72)	7.96 (8.28)
Refinement			
Resolution (Å)	29.35 – 1.15	44.84 – 1.70	43.24 – 1.75
R / R_{free} (%)	16.87 / 23.37	19.34 / 23.71	20.14 / 24.07
Atoms	286	2303	391
Waters	22	330	20
Overall B-factor (Å^2)	15.57	20.60	39.11
RMS Deviations			
Bond lengths (Å)	0.037	0.018	0.011
Bond angles (°)	3.413	1.860	1.777
CCDC code	2085450	2085448	2085449

Supplementary figures

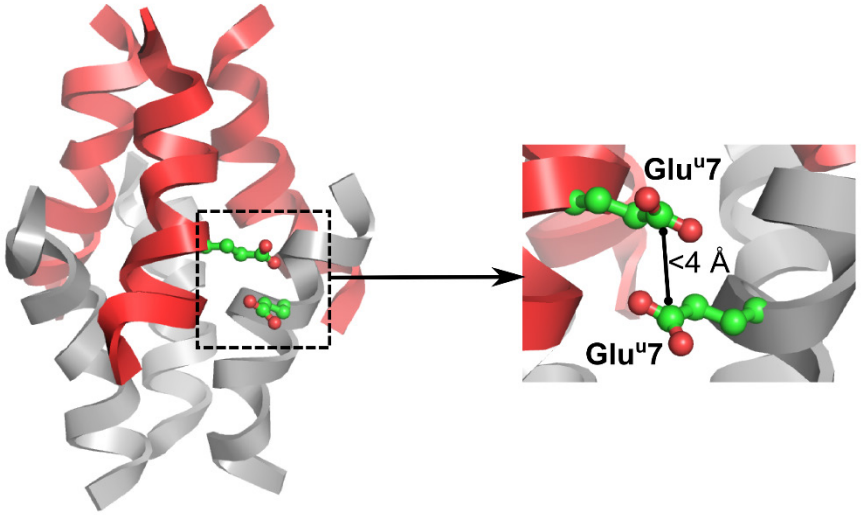


Figure S1. Crystal structure of **H1**[1] six-helix bundle, with close proximity of charged glutamate-type urea residues (Glu^{u}7) highlighted (inset).

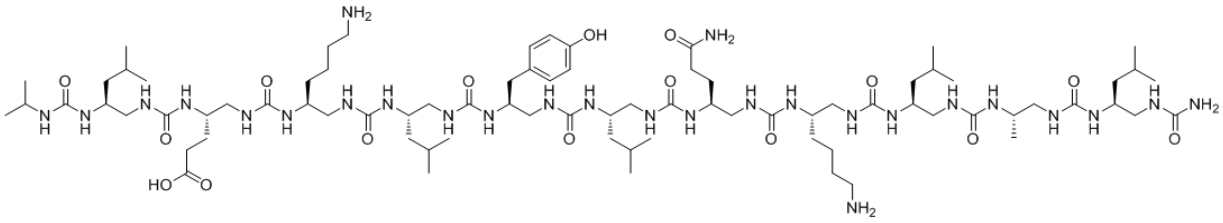


Figure S2. Chemical structure of oligourea **H1**$^+$. Details of the synthesis of this molecule can be found below.

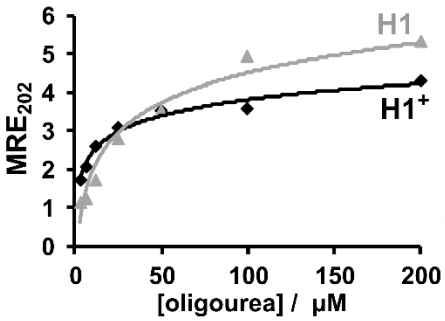


Figure S3. Comparison of variable-concentration circular dichroism analysis of **H1**$^+$ and **H1** (data for **H1** reported previously[1]). Oligourea concentration is plotted on x axis with molar residual ellipticity at 202 nm (MRE_{202}) for **H1**$^+$ and **H1** shown on y axis, with units of deg·cm^2·dmol^{-1}·residue^{-1} (divided by 10^4 for clarity). Data were recorded in pure water for both molecules.

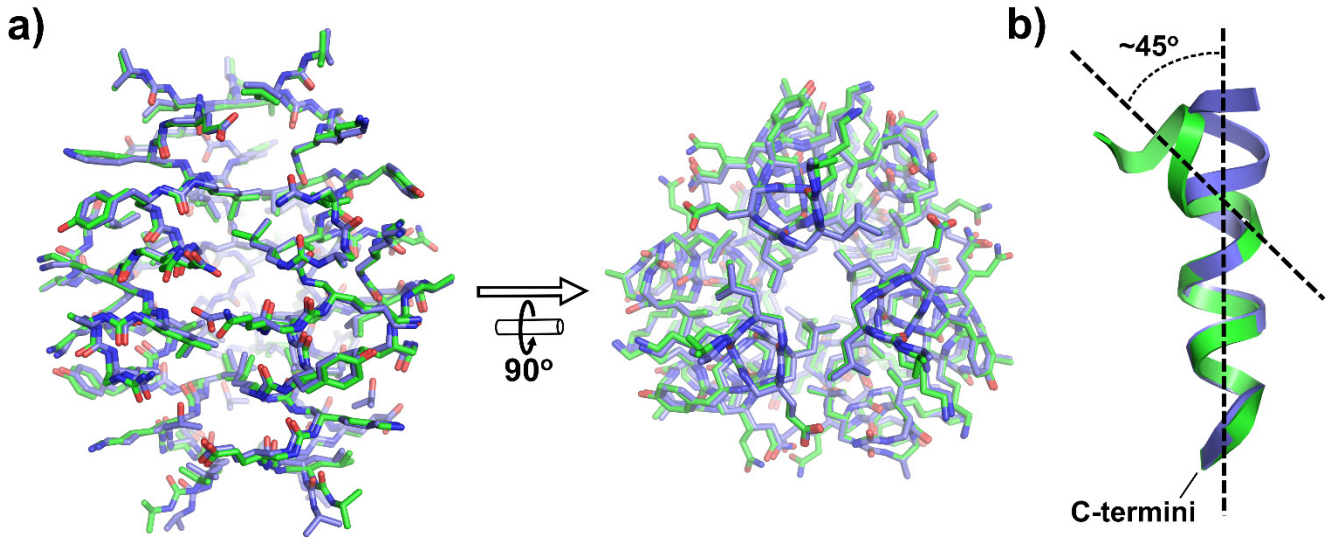


Figure S4. Supporting **H1**$^{+}$ crystallographic figures. a) Structural alignment of six-helix bundles formed by **H1**[1] (blue carbons) and **H1**$^{+}$ (green carbons). RMSD of alignment (Cαs) = 0.13 Å. b) comparison of the helical geometry of single helices from the **H1**$^{+}$ six-helix (blue) and **H1**$^{+}$ eight-helix bundles (green).

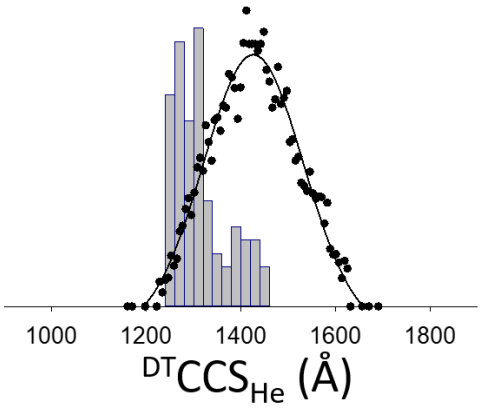


Figure S5. Collision cross sections ($^{DT}CCS_{He}$) of the **H1**$^{+}$ $[8+2Cl^{-}]^{6+}$ species measured by drift tube ion mobility ESI-MS (black circles) with theoretical CCSs calculated from snapshots sampled from molecular dynamics simulations (histograms) using the crystal structure of the **H1**$^{+}$ eight-helix bundle bound by two chloride ions as a starting model.

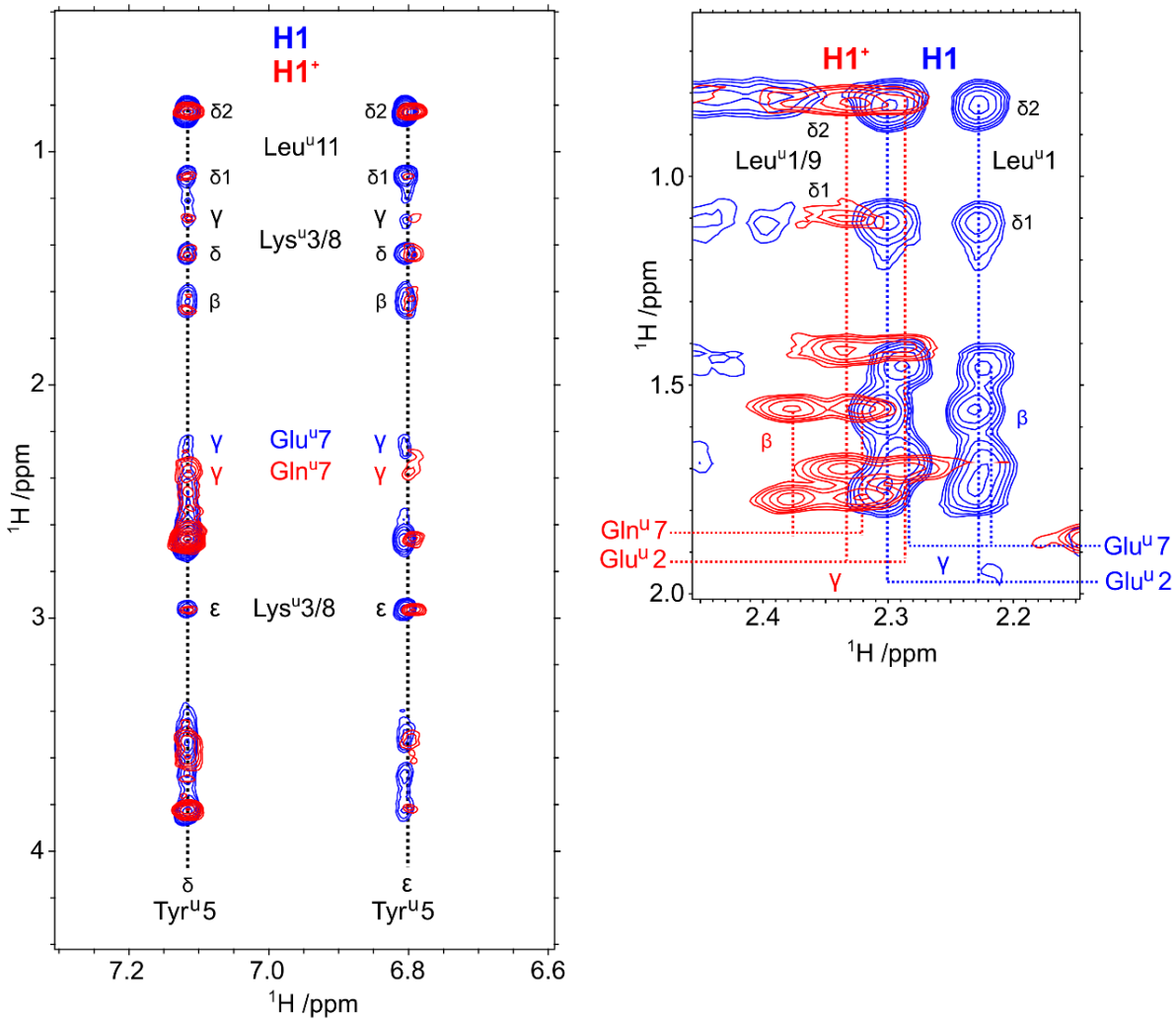


Figure S6. NOESY analysis of **H1**$^+$ at 2 mM in D_2O (red) and comparison to **H1** (blue). The overlay on the left shows NOE crosspeaks that involve the Tyr^{u}5 side chain Hδ and Hε aromatic protons. Inter-residue NOE crosspeaks have been annotated. The overlay on the right shows NOE crosspeaks that involve the Glu^{u}2 and Gln^{u}7 geminal Hγ protons from **H1**$^+$, and the Glu^{u}2 and Gln^{u}7 geminal Hγ protons from **H1**. These protons show NOE crosspeaks to the intra-residue Hβ geminal protons. For **H1** (blue), the Glu^{u}2 protons also show NOE crosspeaks to the Leu^{u}1 methyl group Hδ protons in the neighbouring oligourea chain. For **H1**$^+$ (red), the NOE crosspeaks to the neighbouring oligourea chain involves either Leu^{u}1 in the hexamer form, as well as Leu^{u}9 in the octamer form. NB The Leu^u sidechain δ1 and δ2 protons have not been stereospecifically assigned.

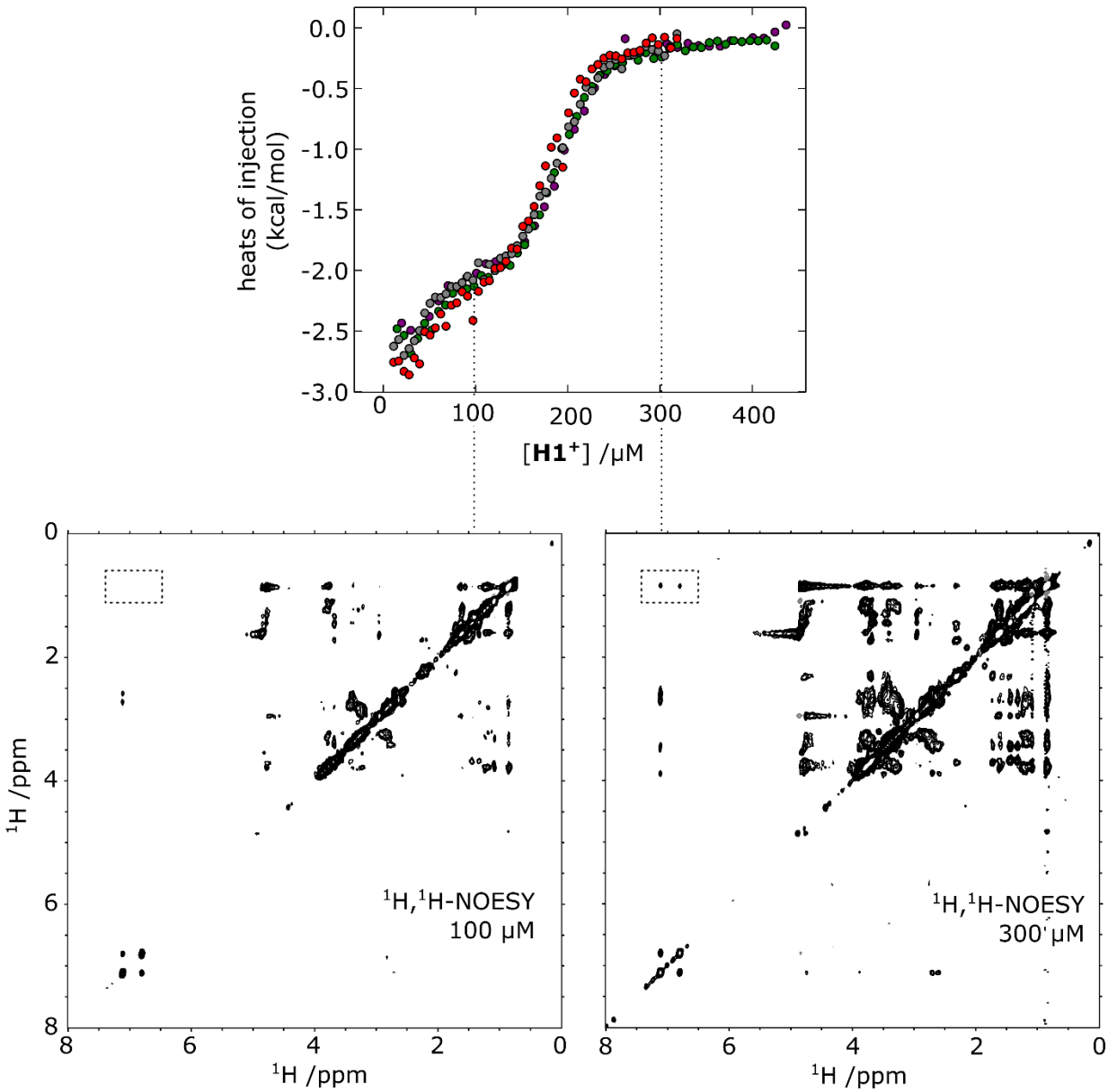


Figure S7. Analysis of **H1**$^{+}$ by isothermal titration calorimetry (ITC) and 2D NOESY NMR. Top: two samples each of 1.5 mM and 2 mM **H1**$^{+}$ in D_2O were diluted into pure D_2O and the heats of injection measured by ITC. An apparent dissociation midpoint at approximately 200 µM **H1**$^{+}$ is followed by a plateau starting at around 300 µM in which there is no further heat of injection. This plateau is in keeping with no further dissociation of the injected **H1**$^{+}$ at concentrations above 300 µM, at which **H1**$^{+}$ has reached an assembly equilibrium. Bottom: ^{1}H,^{1}H-NOESY NMR spectra collected at 100 µM (left) or 300 µM (right) **H1**$^{+}$ at 293 K and 700 MHz. A dotted box indicates a region indicative of assembled **H1**$^{+}$ that corresponds to NOE crosspeaks between Tyr^{U} aromatic ^{1}H and Leu^{U} methyl ^{1}H. These crosspeaks are absent at 100 µM **H1**$^{+}$ and present at 300 µM **H1**$^{+}$. The mixing time was 300 ms and the acquisition used sweepwidths of 7000 Hz in both dimensions and 2,048 × 512 complex points. The contour levels have been adjusted to normalize the concentration difference and number of acquisition scans between the two spectra.

Materials and methods

Chemistry

Oligourea **H1**$^+$ was synthesized using microwave assisted (CEM DiscoveryBio) solid phase methods starting from N_3-protected succinimidyl carbamate building blocks,[2] following the protocol previously described for **H1**.[1] As described for **H1**, the TFA salt of **H1**$^+$ obtained after resin cleavage was purified by semipreparative HPLC (Dionex Ultimate 3000, column: Macherey-Nagel Nucleodur 100-16 C18 ec, 10 x 250, gradient: 35-40% 10 min - 40-55% 15 min acetonitrile 0.1% TFA in water 0.1 % TFA, 4 mL/min). Analytical HPLC characterizations were performed on a Macherey-Nagel column, Nucleodur cc 70/4 100-3 C18 ec, 4.6 x 100, using a gradient of 10-100% 10 min acetonitrile 0.1 % TFA in H_2O 0.1 % TFA, 1 mL/min. The pure product was freeze-dried and TFA was exchanged with HCl by repeated lyophilisations in 0.1 N HCl.

The synthesis of the Glu-,[1] Lys-,[1] Leu-[2] and Ala-[2] type building blocks has been described previously. The Gln-type building block was synthesized starting from the equivalent *N*-Fmoc protected amino acid following previously reported procedures.[2] Fmoc deprotection was performed using 1.5 eq. of piperidine in dry THF.

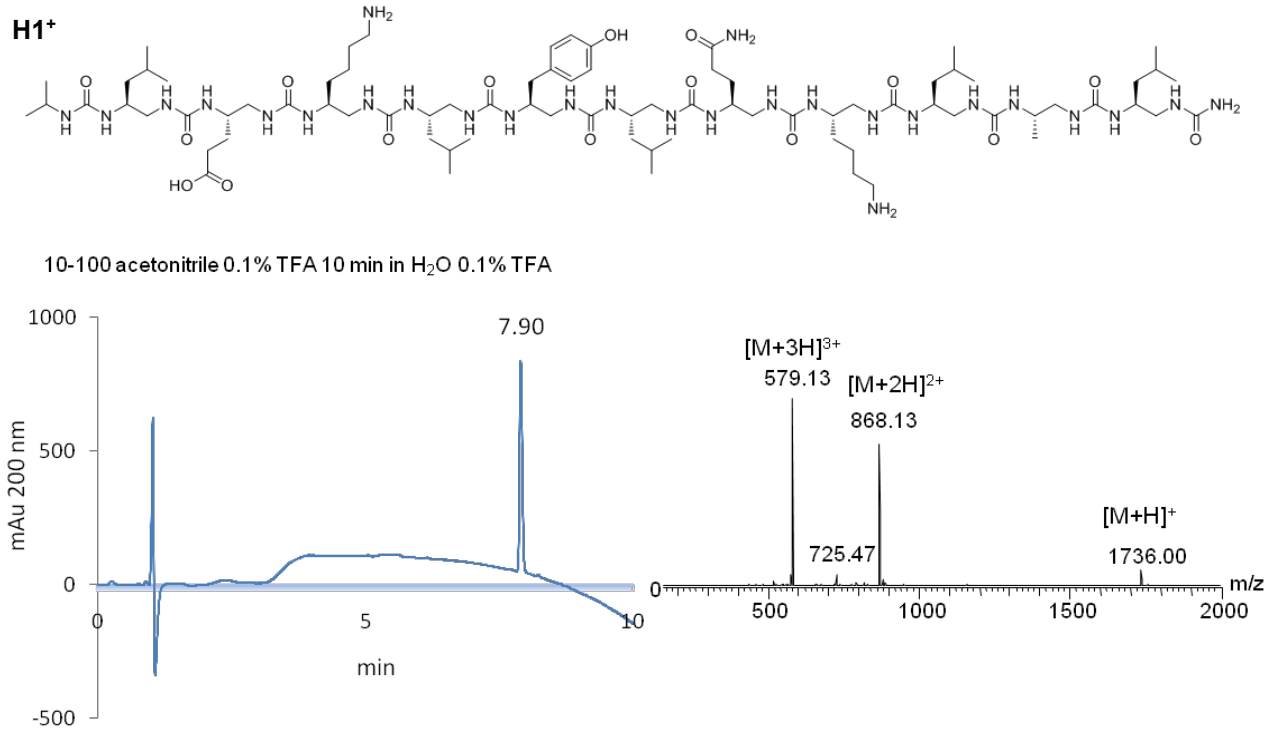
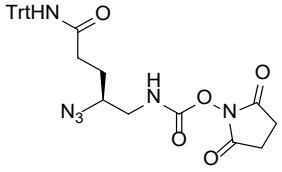


Figure S8. Chemical structure, HPLC profile and ESI-MS spectrum of oligourea **H1**$^+$ (*isopropyl^u-Leu^u-Glu^u-Lys^u-Leu^u-Tyr^u-Leu^u-Gln^u-Lys^u-Leu^u-Ala^u-Leu^u-NH$_2$*). ESI-MS (ESI+) *m/z*: 579.13 $[M+3H]^{3+}$, 868.13 $[M+2H]^{2+}$, 1736.00 $[M+H]^+$; HPLC: R_t= 7.90 min

(S)-2,5-dioxopyrrolidin-1-yl (2-azido-5-oxo-5-(tritylamino)pentyl)carbamate (N_3-Gln^u(Trt)-Osu)



The pure product was obtained as a white solid with an overall yield of 22 % after 6 steps.[2] ^{1}H NMR ($CDCl_3$, 300 MHz) δ: 7.37-7.22 (m, 15H), 6.73 (s, 1H), 5.67 (t, *J* = 6.0 Hz, 1H), 3.68 – 3.50 (m, 1H), 3.25 (dt, *J* = 13.0, 6.5 Hz, 1H), 3.11 (dt, *J* = 14.1, 5.8 Hz, 1H), 2.80 (s, 4H), 2.49 (t, *J* = 6.7 Hz, 2H), 1.92 (q, *J* = 6.5 Hz, 2H); ^{13}C NMR ($CDCl_3$, 75 MHz) δ: 170.83, 169.76, 151.80, 144.46, 128.63, 128.04, 127.15, 70.70, 60.59, 44.36, 32.37, 26.36, 25.44; HRMS (ESI+) *m/z* calcd for $C_{29}H_{29}N_6O_4$ $[M+H]^+$ 541.2201, found 541.2195. (Figs. S9 and S10)

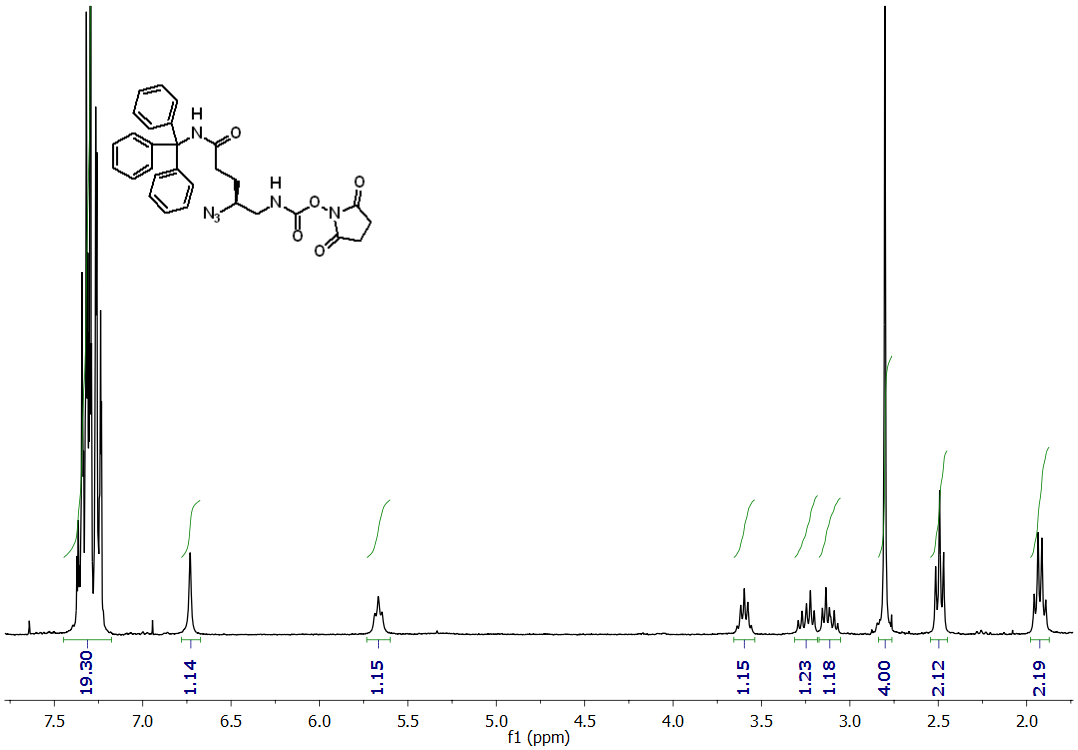


Figure S9. ^{1}H NMR Gln-type building block *(N_3-Gln^u(Trt)-Osu)*

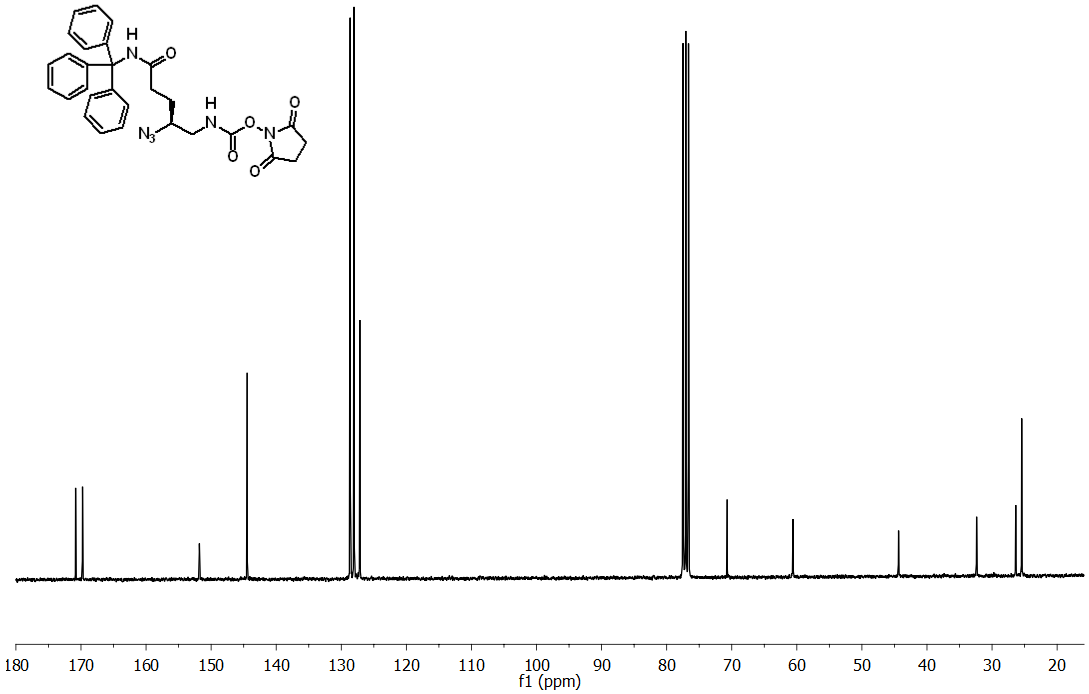


Figure S10. ^{13}C NMR Gln-type building block *(N_3-Gln^u(Trt)-Osu)*

Circular dichroism (CD)

Circular dichroism (CD) experiments were performed on a Jasco J-815 spectrometer. Variable-concentration experiments were performed on oligourea **H1**$^+$ in double-distilled H_2O starting from an oligourea concentration of 200 µM followed by serial two-fold dilutions. Data were recorded at 20 °C between wavelengths of 180 and 250 nm at 0.5 nm intervals at a speed of 50 nm/min with an integration time of 2 seconds. CD-monitored thermal melting experiments were performed in pure water at an oligourea concentration of 200 µM. For these experiments, the sample was heated from 5 °C to 90 °C using a gradient of 1 °C·min^{-1}. The CD signal at 202 nm was monitored for these experiments.

Native mass spectrometry

Experiments were performed on an Agilent 6560 DTIMS-Q-TOF instrument (Agilent Technologies, Santa Clara, CA), with the dual-ESI source operated in the positive ion mode. A syringe pump flow rate of 180 µL/h was used. Capacitance diaphragm gauges are connected to the funnel vacuum chamber and to the drift tube. An in-house modification to the pumping system allows better equilibration of the pressures: a Tri-scroll 800 vacuum pump (Agilent, Santa Clara, CA) is connected to the source region with an Edwards SP16K diaphragm valve connected to the front pumping line, while the original Tri-scroll 800 pump is connected to the Q-TOF region. The helium pressure in the drift tube was 3.89 ± 0.01 Torr, and the pressure in the trapping funnel was 3.69 ± 0.01 Torr. The pressure differential between the drift tube and the trapping funnel ensures only helium is present in the drift tube. The acquisition software version was B.06.00. All spectra were recorded using soft source conditions. The tuning parameters of the instrument (electrospray source, trapping region and post-IMS region (QTOF region)) are optimized as described elsewhere.[3] The source temperature was set at 220 °C and the source fragmentor voltage was set to 350 V. The trapping time was 1500 µs and release time 200 µs. Trap entrance grid delta was set to 8 V.

Step-field experiments (five drift tube voltages for each samples) were performed to determine the collision cross sections (CCS). The arrival time distributions (ATDs) for each charge state of the complexes were fitted with one gaussian peak (or two gaussian peaks if necessary) using OriginPro 2016, to determine the arrival time t_A of the center of the peak. The arrival time t_A is related to ΔV (voltage difference between the entrance and the exit of the drift tube region) by:

$$t_A = \frac{L^2}{K_0}\frac{T_0 p}{p_0 T}\cdot\left(\frac{1}{\Delta V}\right) + t_0$$

t_0 is the time spent outside the drift tube region and before detection. A graph of t_A vs. $1/\Delta V$ provides K_0 from the slope and t_0 as the intercept. The drift tube length is L = 78.1 ± 0.2 cm, the temperature is measured accurately by a thermocouple (T = 297 ± 1 K), and the pressure is measured by a capacitance gauge (p = 3.89 ± 0.01 Torr). The CCS is determined using:

$$CCS = \frac{3ze}{16N_0}\cdot\sqrt{\frac{2\pi}{\mu k_B T}}\cdot\frac{1}{K_0}$$

The relative combined standard uncertainty on the CCS of the peak center is ~2.0%.[4] The reconstruction of the experimental CCS distributions from the arrival time distributions at the lowest voltage is then performed using equation:

$$CCS = a\cdot\frac{z}{\sqrt{\mu}}\times t_A$$

where the factor *a* is determined from the t_A of the peak center at the lowest voltage and the CCS calculated from the regression described above, from the peak centers.

The samples were prepared at 50 µM concentration in pure water or ammonium acetate 100 mM. 2% sulfolane (Sigma-Aldrich, St Quentin Fallavier, France) was added to the 100 mM ammonium acetate solution in order to enable the detection of chlorine adducts.

Calculation of gas-phase structures and collision cross sections (CCS) by molecular dynamics

Molecular dynamics simulations were used to calculate theoretical collision cross sections of the gas-phase **H1**$^{+}$ assemblies. The crystal structures of the **H1**$^{+}$ six-helix and eight-helix bundles were used as starting models. To reach the experimental charge state (5+ or 6+), protons were added to N-terminal and side-chain NHs. The structures were optimised at the PM7 semi-empirical level[5] using Gaussian 16 rev. B.01.[6] Then, Atom-Centered Density Matrix Propagation molecular dynamics (ADMP, 1000 fs, 296 K) at the semi-empirical level (PM7) was performed. The theoretical CCS values were calculated for a structure every 10 fs, using the trajectory model (Mobcal[7], original parameters for helium, N and O parameterized as C, P and K parameterized as Si). Histograms of the calculated CCS values were prepared using Sigmaplot 14.

Crystallography

Crystals of oligourea **H1**$^{+}$ were grown in standard hanging drops incubated at 20 ^{o}C. Hanging drops were prepared by mixing 0.5 µL of a 10 mg/mL solution of oligourea **H1**$^{+}$ (dissolved in double-distilled H_2O) with an equal volume of crystallisation reagent. Standard commercially-available sparse-matrix protein crystallisation screens were used to find suitable crystallisation conditions. Specific crystallisation, data collection, structure solution and refinement details for the three distinct crystal forms of **H1**$^{+}$ are described below with additional details provided in Table S1.

Crystal form 1 ($P6_3$) crystallised from a crystallisation reagent composed of 10 % isopropanol, 200 mM $CaCl_2$ and 100 mM sodium acetate buffer (pH 4.6). For data collection, a single crystal was cryo-protected in a solution composed of the above crystallisation reagent supplemented with 25 % glycerol. Diffraction data were collected on beam line ID23-2 at the European Synchrotron Radiation Facility (ESRF), and processed using XDS.[8] The structure was solved by molecular replacement using a modified version of the previously reported **H1** crystal structure as a search model,[1, 9] using Phaser[10] from the CCP4 suite.[11] Geometric restraints for maximum-likelihood restrained refinement were generated using the PRODRG server.[12] Model building and restrained refinement were performed in Coot[13] and REFMAC5,[14] respectively. The final model was refinement to a resolution of 1.15 Å with R and R_{free} factors of 16.87 % and 23.37 % respectively. Crystal form 2 ($P2_12_12_1$) crystallised from a crystallisation reagent composed of 20 % PEG2000, 10 mM nickel (II) chloride and 100 mM tris-HCl (pH 8.5). For data collection, a single crystal was cryo-protected in a solution composed of 25 % PEG2000, 10 mM nickel (II) chloride and 100 mM tris-HCl (pH 8.5), and flash frozen in liquid nitrogen. Data collection, data processing, model building and refinement were performed as described above. The final model was refined to a resolution of 1.7 Å, with R and R_{free} factors of 19.34 % and 23.71 %, respectively. Crystal form 3 ($P422$) crystallised from a crystallisation reagent composed of 30 % PEG400, 200 mM sodium citrate and 100 mM Tris-HCl (pH 8.5). For data collection, a single crystal was frozen directly in liquid nitrogen. Data collection, data processing, model building and refinement were performed as described above. The final model was refined to a resolution of 1.75 Å, with R and R_{free} factors of 20.14 % and 24.07 %, respectively. Structural analysis was performed using SURFNET,[15] HELANAL[16, 17] and PyMOL. Atomic coordinates and structure factors have been deposited in the Cambridge Crystallographic Data Centre with accession codes listed in Table S1.

NMR spectroscopy

Spectra were recorded on a Bruker Avance 800 or 700 MHz spectrometer equipped with a cryogenic triple-resonance gradient probe or standard temperature probe, respectively. Samples were measured in pure D_2O at a temperature of 293 K. ^{1}H,^{1}H-NOESY spectra used a mixing time of 150-300 ms with excitation sculpting to remove the residual water signal, and using sweep widths of 7000-11000 Hz in both dimensions centred on the water frequency. NMR spectra were collected with the software Topspin and processed with NMRPipe/Draw.[18] Figures were prepared by using Sparky 3 (T. D. Goddard & D. G. Kneller, University of California). Additional details of the experimental NMR conditions are given in the corresponding figure legends.

Isothermal titration calorimetry (ITC)

Dissociation of oligourea **H1**$^+$ assemblies was characterised using a Malvern iTC200 at a temperature of 293 K. Assembled **H1**$^+$ was prepared in pure D_2O at concentrations of 1.5-2 mM and loaded into the syringe. Pure D_2O was placed in the cell and heats of injection were measured with a series of 52 injections of 0.75 µL, with stirring at 500 rpm and sensitivity set to high. The data were processed by using NITPIC[19, 20] and the figure prepared by using GUSSI.[21] Figure S7 includes data from four separate measurements. An apparent concentration at which the assembly occurs could be determined from the data, whereas multiple oligomeric states and likely stepwise disassembly/assembly prevented determination of precise values for the dissociation constant(s).

Supplementary information references

1. 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.
2. C. Douat-Casassus, K. Pulka, P. Claudon and G. Guichard, *Org. Lett.*, 2012, **14**, 3130-3133.
3. V. Gabelica, S. Livet and F. Rosu, *J. Am. Soc. Mass Spectrom.*, 2018, **29**, 2189-2198.
4. V. Calabrese, H. Lavanant, F. Rosu, V. Gabelica and C. Afonso, *J. Am. Soc. Mass Spectrom.*, 2020, **31**, 969-981.
5. J. J. P. Stewart, *J. Mol. Model.*, 2013, **19**, 1-32.
6. Gaussian 16, Revision B.01, M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian, Inc., Wallingford CT, 2016.
7. M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *J. Phys. Chem.*, 1996, **100**, 16082-16086.
8. W. Kabsch, *Acta Crystallogr. D*, 2010, **66**, 125-132.
9. G. W. Collie, K. Pulka-Ziach and G. Guichard, *Chem. Commun.*, 2016, **52**, 1202-1205.
10. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658-674.
11. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. D*, 2011, **67**, 235-242.
12. A. W. Schüttelkopf and D. M. F. van Aalten, *Acta Crystallogr. D*, 2004, **60**, 1355-1363.
13. P. Emsley and K. Cowtan, *Acta Crystallogr. D*, 2004, **60**, 2126-2132.
14. G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr. D*, 2011, **67**, 355-367.
15. R. A. Laskowski, *J. Mol. Graph.*, 1995, **13**, 323-330.
16. M. Bansal, S. Kumar and R. Velavan, *J. Biomol. Struct. Dyn.*, 2000, **17**, 811-819.
17. P. Kumar and M. Bansal, *J. Biomol. Struct. Dyn.*, 2012, **30**, 773-783.
18. F. Delaglio, S. Grzesiek, G. W. Vuister, G. Zhu, J. Pfeifer and A. Bax, *J. Biomol. NMR*, 1995, **6**, 277-293.
19. S. Keller, C. Vargas, H. Zhao, G. Piszczek, C. A. Brautigam and P. Schuck, *Anal. Chem.*, 2012, **84**, 5066-5073.
20. T. H. Scheuermann and C. A. Brautigam, *Methods*, 2015, **76**, 87-98.
21. C. A. Brautigam, *Methods Enzymol.*, 2015, **562**, 109-133.