Electronic Supplementary Material (ESI) for Physical Chemistry Chemical Physics.
This journal is © the Owner Societies 2020

SUPPORTING INFORMATION FOR:

Impact of Oxetane Incorporation on the Structure and Stability of Alpha-helical Peptides

Eleanor S. Jayawant[a], Jonathan D. Beadle[a], Ina Wilkening[a], Piotr Raubo[b], Michael Shipman[a], Rebecca Notman[a] and Ann M. Dixon[a*]

[a] Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK.

[b] Medicinal Chemistry, Research and Early Development, Oncology R&D, AstraZeneca, Cambridge, UK.

*To whom correspondence should be addressed: Dr Ann Dixon, Department of Chemistry, University of Warwick, Coventry, CV4 7AL, UK, Telephone: +44 2476 150037; FAX: +44 2476 524112; email: ann.dixon@warwick.ac.uk

Molecular dynamics simulations

To obtain insights into the effect of the oxetane ring on the structure of helical peptides, we carried out molecular dynamics (MD) computer simulations of the 18-residue parent peptide Ac-[KAAAA]$_3$-KGY (**1**); modified peptides **2a** and **2b** which contain an oxetane-modified alanine residue, A_{ox}, at positions 8 and 3 respectively; and modified peptide **2c** which contains an oxetane-modified glycine residue, G_{ox}, at position 3.

1.1 Model Building

Starting configurations of the peptides were generated using the Avogadro program version 1.2.0.[1] The peptides were built assuming an ideal α-helical conformation (ϕ and ψ backbone dihedral angles were set to -60° and -40° respectively) and uncharged ends. The ends were then replaced with N-terminal acetyl and C-terminal amide caps. For peptides **2a** and **2b-c**, the C=O oxetane substitution was made at positions 8 or 3 respectively. This was followed by a brief steepest descents energy minimisation in Avogadro using the Universal Force Field,[2] which generated the starting structure of each helical peptide for subsequent MD simulation in Gromacs 5.1.4.[3] Input topology files for Gromacs were generated using the Gromacs pdb2gmx tool.

1.2 Force Field Parameters

MD simulations were carried out using the CHARMM27 force field for proteins[4]. Methanol parameters were obtained from the CHARMM General Force Field[5]. Additional parameters for the oxetane ring were incorporated into the CHARMM27 force field, details of which are provided in our previous work.[6]

1.3 Simulation Parameters

Prior to solvation, each structure was relaxed by performing energy minimisation. The structures were solvated with a pre-equilibrated box of 4931–4935 MeOH molecules and 4 chloride ions were added to balance the charge.

The initial structures were relaxed by performing 50000 steps of steepest descent energy minimisation in vacuum. The final structures were then solvated with MeOH and the system

was subjected to a further 50000 steps of steepest descents energy minimization. The systems were equilibrated in two stages, with position restraints of 1000 kJ mol^{-1} applied to peptide atoms. This equilibration consisted of 50000 steps of simulation at 310 K in the NVT ensemble, followed by 50000 steps of simulation at 310 K and 1 bar in the NPT ensemble to equilibrate the temperature and density of the system respectively.

Production MD simulations were performed for 40 ns and position restraints were released except for the Cα atom of Lys1, which was used as an immobile reference for the pulling simulations. For each helical structure, the Cα atom of Tyr18 was pulled along the *z*-axis, using a spring constant of 25 kJ mol^{-1} nm^{-2} and a pull speed of 0.00025 nm ps^{-1}. These parameters were selected following a series of tests in which spring constants of 100 to 10 kJ mol^{-1} nm^{-2} and pull speeds of 0.02 to 0.00005 nm ps^{-1} were assessed using peptide **1**.

Following initial observations that pull speeds of 0.02 to 0.001 nm ps^{-1} resulted in rapid helix unwinding even in short tests, further repeats showed that with slow pull speeds, use of a higher spring constant results in less error between repeats but more overall noise in the force extension curves (Figure S5a & b). Thus a spring constant of 25 kJ mol^{-1} nm^{-2} and a pull speed of 0.00025 nm ps^{-1} were selected, as these parameters allowed for slow, gentle unwinding over the course of a longer simulation, while producing force extension curves with minimal noise. Slower pulling rates of 0.0001 and 0.00005 nm ps^{-1} resulted in the production of similar trajectories and force extension curves (Figure S5c), thus the faster pull speed was selected to reduce computational expense while maintaining integrity of the data. Preliminary tests showed that pulling from the N-terminus rather than the C-terminus with these conditions did not alter the key unwinding events of the trajectory.

In all MD simulations, all bonds were constrained using the LINCS algorithm[7] and a simulation timestep of 2 fs was used. Periodic boundary conditions were applied in all directions. Lennard-Jones interactions were cutoff at 1.0 nm. Electrostatic interactions were handled using the particle mesh Ewald approach with a real-space cutoff of 1.0 nm. In initial temperature equilibration simulations, the temperature was controlled using velocity rescaling with a time constant of 0.1 ps.[8] For the density equilibrations, temperature was maintained using the Berendsen thermostat and the pressure was isotropically maintained at 1 bar using the Berendsen barostat with a time constant of 2.0 ps and compressibility of 4.5×10^{-5} bar^{-1}.[9]

In the steered MD production run, temperature was controlled using the Nosé-Hoover thermostat.[10]

1.4 Linear Pentapeptide

Simulations of linear oxetane-modified pentapeptide $LAG_{ox}AY$-OMe were prepared based on methods previously described by us[11], with some minor changes. In brief, starting coordinates of the peptides were generated using the Avogadro program version 1.2.0.[1] The peptide was built assuming a linear conformation (ϕ, ψ and ω backbone dihedral angles were set to 180°) and uncharged ends. The C-terminus of the peptide was replaced with a methyl ester end group, and the C=O oxetane substitution was made. This was followed by a steepest descents energy minimisation in Avogadro using the Universal Force Field,[2] which generated the starting structure of the peptide for subsequent MD simulation in Gromacs 5.1.4.[3] MD simulations were carried out using the CHARMM27 forcefield for proteins[4] and DMSO[12] with modified parameters for the oxetane ring as previously described.[6] To improve the quality of our model, selected NOE distances from NMR experiments described previously[11] were incorporated as distance restraints in the MD simulations.

The initial structure was relaxed by performing 50000 steps of MD simulation in the NVT ensemble in vacuum at 300 K. The final structure was then solvated with a pre-equilibrated box of 424 DMSO molecules and the system was subjected to 50000 steps of steepest descents energy minimization. This was followed by 100000 steps of simulation at 300 K in the NVT ensemble and 50000 steps of simulation at 300 K and 1 bar in the NPT ensemble to equilibrate the temperature and density of the system respectively. The peptide was then simulated for 100 ns at 500 K in the NVT ensemble. Cluster analysis using the algorithm proposed by Daura et al.[13] was performed to group peptide conformations according to their structural similarity. The central structure of the top five most populated clusters (which accounted for > 99.9% of the total population) was then used as the starting configuration for five independent 100 ns simulations at 300 K and 1 bar in the NPT ensemble. Accordingly, $LAG_{ox}AY$-OMe was simulated for a total simulation time of 500 ns.

All bonds were constrained using the LINCS algorithm[7] and a simulation timestep of 2 fs was used. Periodic boundary conditions were applied in all directions. Lennard-Jones interactions were cutoff at 1.0 nm. Electrostatic interactions were handled using the particle mesh Ewald approach with a real-space cutoff of 1.0 nm. The temperature was controlled using velocity

rescaling with a stochastic thermostat with a time constant of 0.1 ps[8] and the pressure was isotropically maintained at 1 bar using the Parrinello-Rahman barostat with time constant 2.0 ps and compressibility 4.5 × 10^{-5} bar^{-1}.[10,14] Analysis of ϕ/ψ angles was carried out for each residue for the entire length of the trajectory, and dihedral angles were plotted using Matplotlib[15].

1.5 Simulation Results

1.5.1 *Work Done*

By integrating the area under the force extension curves, the total work done to unwind the helices was calculated in kJ mol^{-1}. A cut-off point of 20 ns, equivalent to 5 nm of extension, was selected for the calculation, as this approximately corresponds to loss of helicity. For each peptide, the average work done across 10 repeats was calculated and plotted using OriginPro 2019.[16] Leave-one-out cross-validation was performed, by calculating the mean work done value while systematically excluding each repeat and plotting using Origin 2016. The cross-validation shows that 10 repeats are sufficient to reach convergence of data (Figure S6). Statistical analysis of work done was carried out using one-way analysis of variance to assess statistically significant differences between the means for the four helices, and using independent sample t-tests to compare the means for peptides **2b** and **2c**. All statistical analysis was performed using IBM SPSS Statistics 24.

1.5.2 Hydrogen bonding

Analysis of the MD simulations was carried out on first 20 ns of the trajectory. Beyond this point, helicity is lost but the peptide continues to experience pulling. Hydrogen bonds were identified using geometric criteria whereby a hydrogen bond is said to exist if the donor---acceptor angle ≤ 30° and the donor-hydrogen---acceptor distance ≤ 0.35 nm. The data were plotted using Matplotlib[15]. Number of hydrogen bonds vs. time for each residue across 10 repeats for peptides **1**, **2a**, **2b** and **2c** are shown in Figure S8.

1.5.3 Structural Analysis

In order to assess the effect of the oxetane on the structure, representative snapshots were prepared after 5 ns of simulation. This point in the trajectory was selected as the peptides have had time to relax from position-restrained ideal helices, but have not yet been affected

by the pulling: in all repeats, 90–100% of initial hydrogen bonds are maintained at 5 ns, suggesting the helix is not yet unwinding.

1.5.4 Dihedral Analysis

Analysis of ϕ/ψ angles was carried out for the period of the trajectory between 2 and 5 ns. The initial 2 ns was omitted to allow the peptides to relax from their position-restrained ideal helix starting configurations, and, as previously discussed, at 5 ns the structures of the peptides are not yet affected by the pulling (as confirmed by the maintenance of initial hydrogen bonds). Dihedral angles were plotted for all 10 repeats using Matplotlib[15].

1.5.5 Comparison to Linear Pentapeptide

Analysis of ϕ/ψ angles was carried out for each residue of $LAG_{ox}AY$-OMe for the entire length of the trajectory, and dihedral angles were plotted using Matplotlib[15].

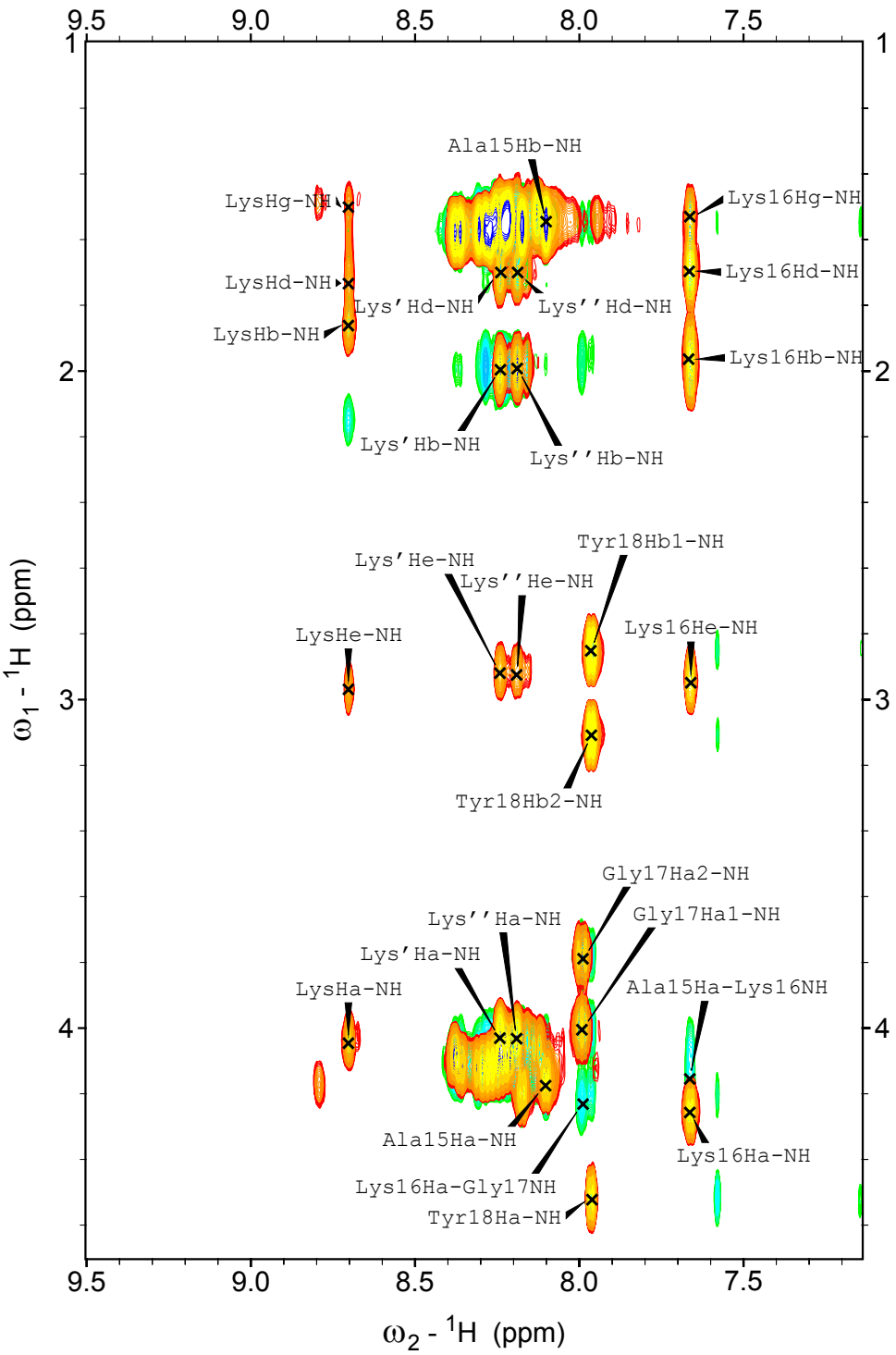

Figure S1. Fingerprint region of the ^{1}H-^{1}H TOCSY spectrum (orange contours, 140 ms mixing time) for peptide **1** solubilised to a final concentration of 2 mM in 80% MeOD-d_4 + 20% H_2O. The TOCSY spectrum is overlayed onto the ^{1}H-^{1}H NOESY spectrum (green contours, 150 ms mixing time). Non-sequential Lys assignments are denoted Lys, Lys', and Lys''.

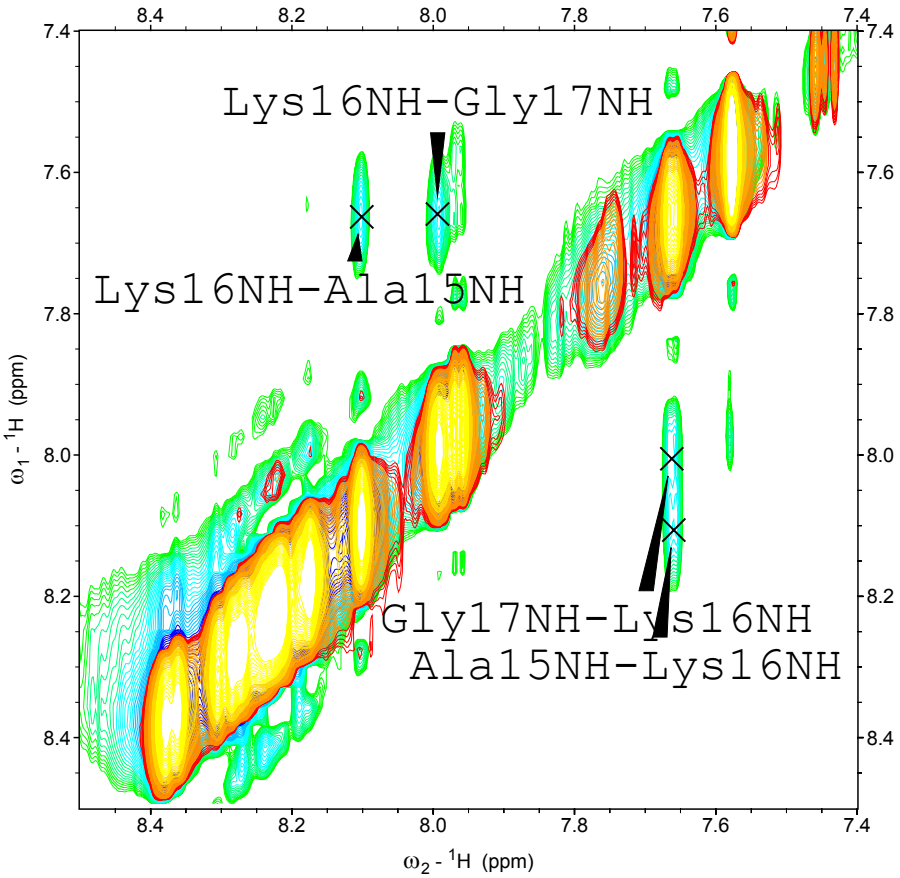

Figure S2. NH-NH NOE correlations supporting the sequential assignment of residues 15-17 in peptide **1**.

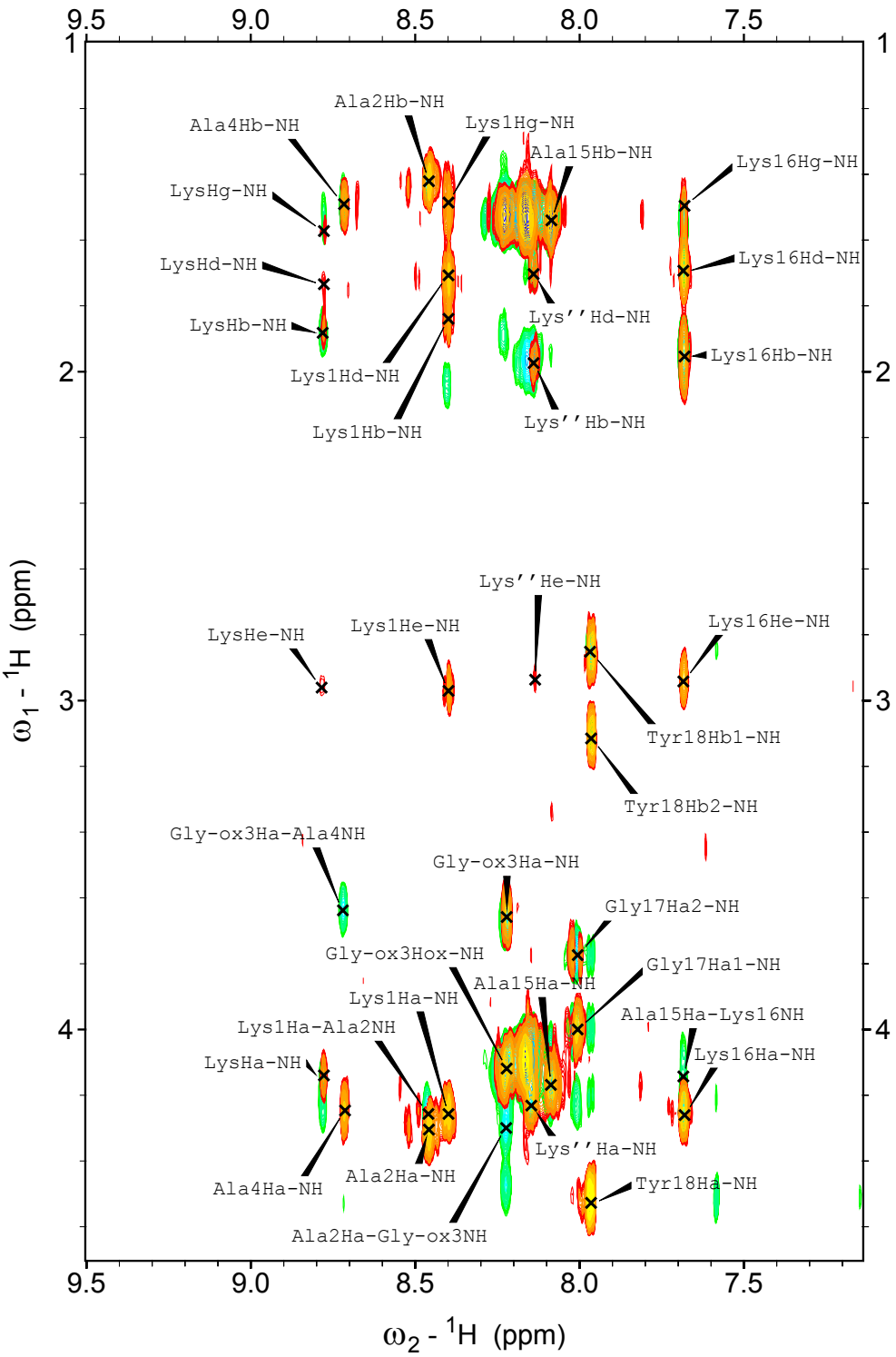

Figure S3. Fingerprint region of the ^{1}H-^{1}H TOCSY spectrum (orange contours, 140 ms mixing time) for peptide **2c** solubilised to a final concentration of 2 mM in 80% MeOD-d_4 + 20% H_2O. The TOCSY spectrum is overlayed onto the ^{1}H-^{1}H NOESY spectrum (green contours, 150 ms mixing time). Non-sequential Lys assignments are denoted Lys and Lys''.

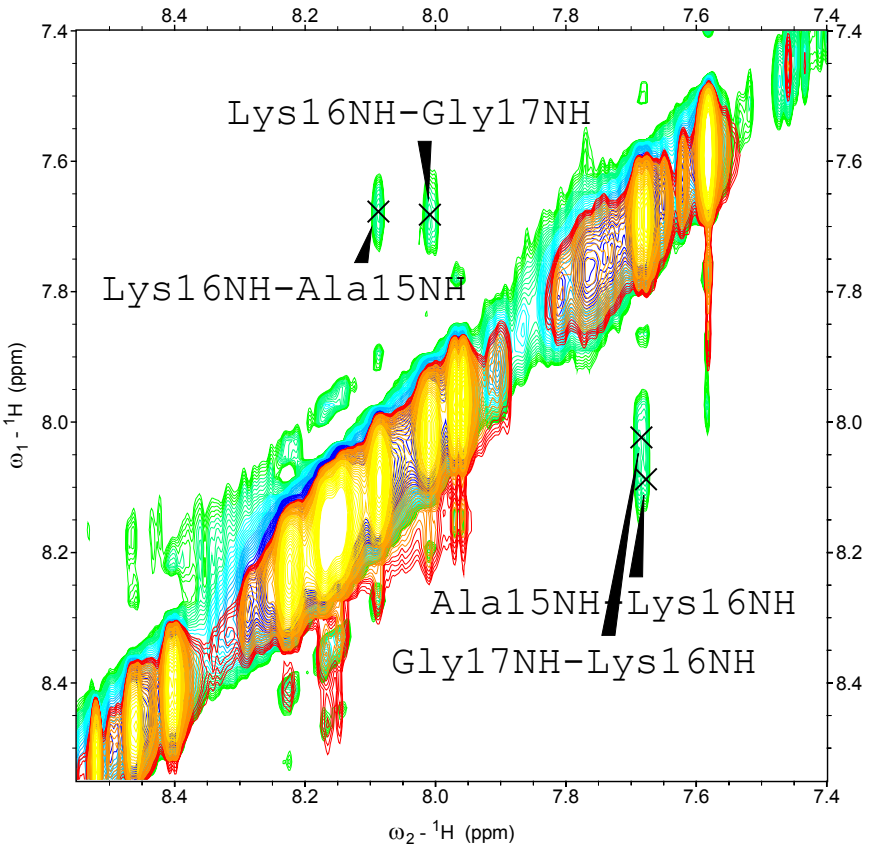

Figure S4. NH-NH NOE correlations supporting the sequential assignment of residues 15-17 in peptide **2c**.

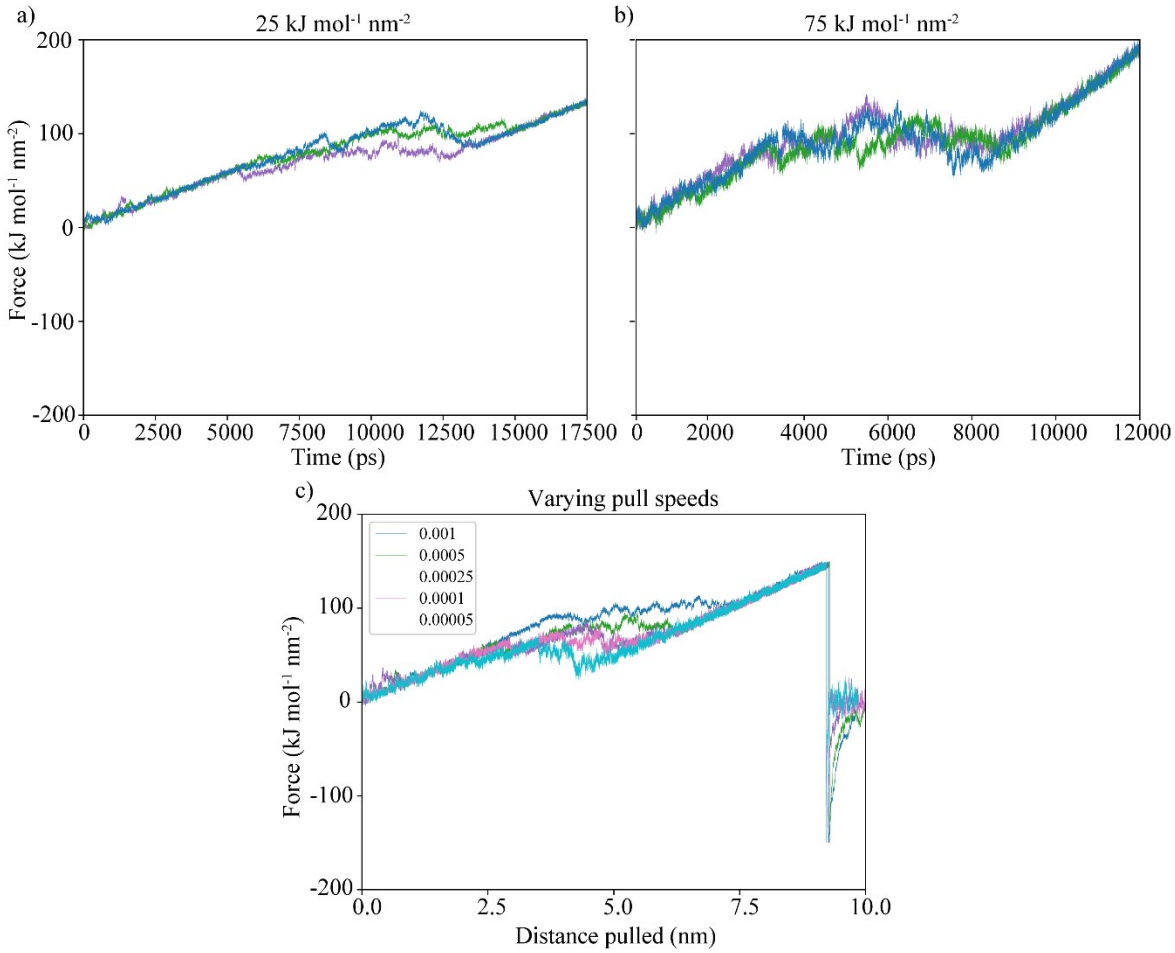

Figure S5. Representative force extension curves showing three repeats with different starting co-ordinates for different spring constants (a) 25 kJ mol^{-1} nm^{-2} and (b) 75 kJ mol^{-1} nm^{-2}, while maintaining the pull speed at 0.0005 nm ps^{-1}, and (c) a single repeat for five different pull speeds (0.001 to 0.00005 nm ps^{-1}) with a constant spring constant of 25 kJ mol^{-1} nm^{-2}. Based on these tests, a spring constant of 25 kJ mol^{-1} nm^{-2} and a pull speed of 0.00025 nm ps^{-1} were selected.

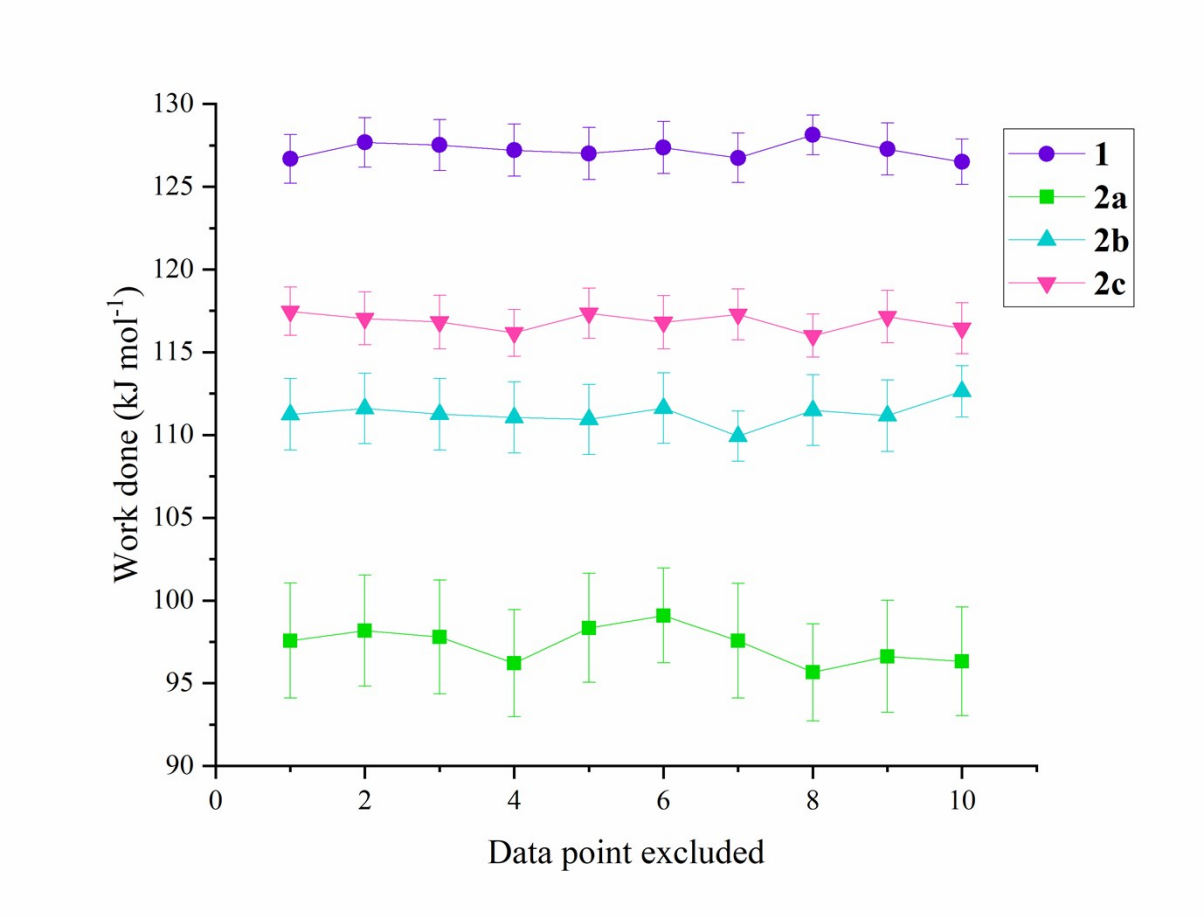

Figure S6. Leave-one-out cross-validation of simulation data for peptides **1**, **2a**, **2b** and **2c**. Each data point for a repeat was systematically left out, to assess how the mean changes. Error bars represent standard error between the 9 data points.

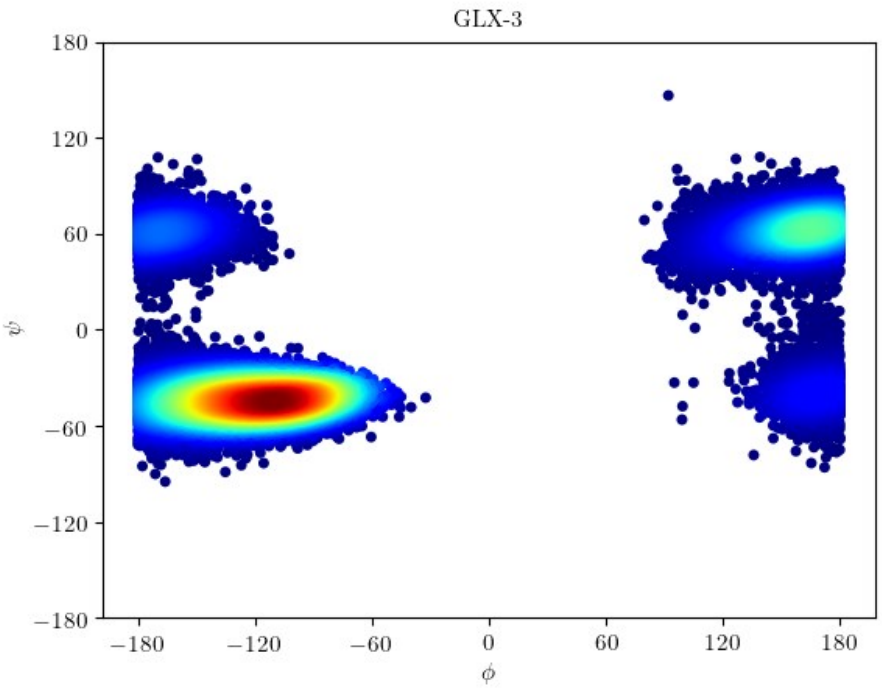

Figure S7. Ramachandran plot for G_{ox} residue in a short linear peptide, $LAG_{ox}AY$-OMe.

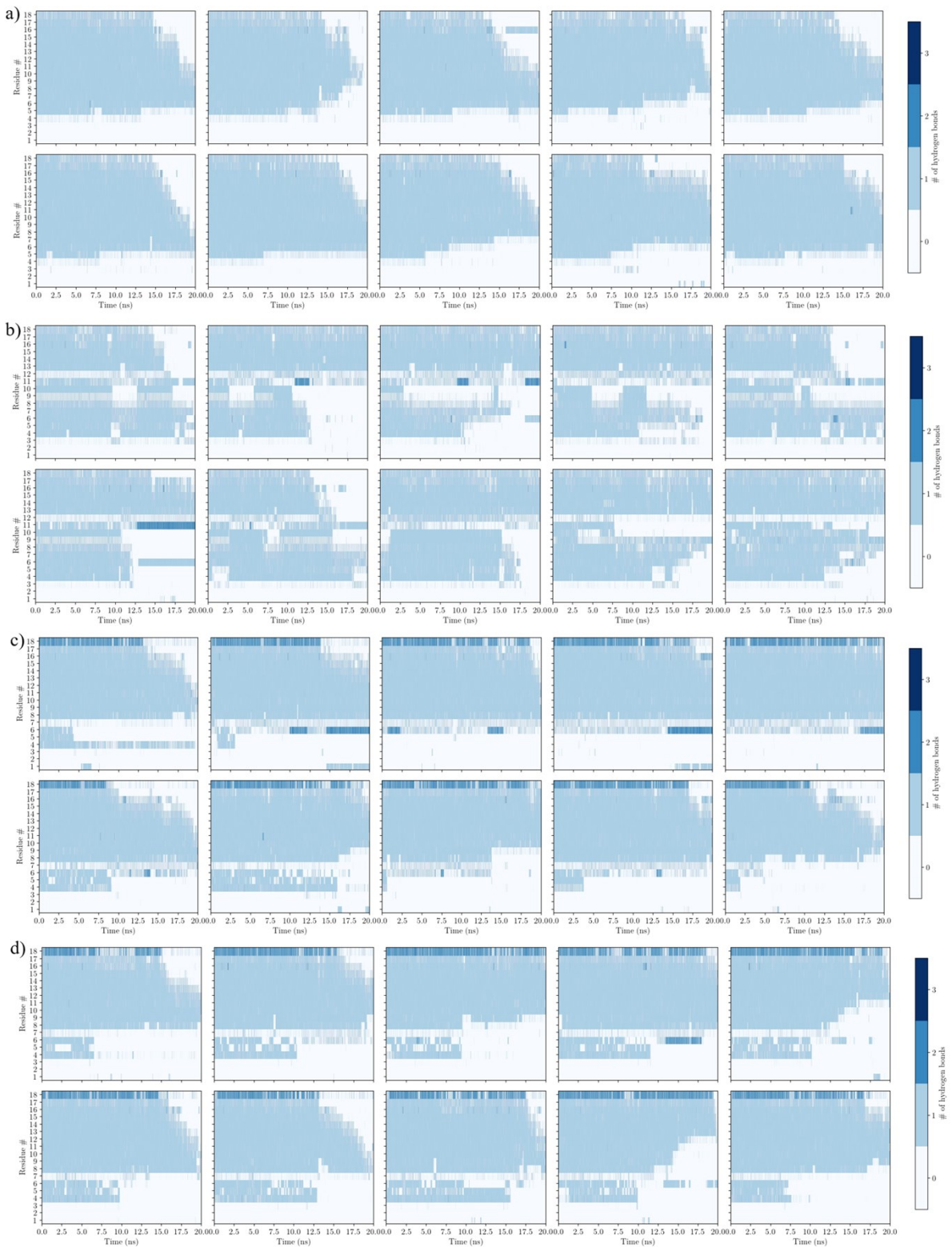

Figure S8. Number of hydrogen bonds per residue over the course of the 20 ns trajectory for all 10 repeats for (a) parent peptide **1**, (b) peptide **2a**, (c) peptide **2b**, and (d) peptide **2c**.

References

1 M. D. Hanwell, D. E. Curtis, L. D. C, T. Vandermeersch, E. Zurek and G. R. Hutchinson, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, *J. Cheminform.*, 2012, **4**, 17.

2 A. K. Rappé, C. J. Casewit, K. S. Colwell, W. A. Goddard and W. M. Skiff, UFF, a Full Periodic Table Force Field for Molecular Mechanics and Molecular Dynamics Simulations, *J. Am. Chem. Soc.*, 1992, **114**, 10024–10035.

3 M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX*, 2015, **1**–**2**, 19–25.

4 A. D. Mackerell Jr., N. Banavali and N. Foloppe, Development and Current Status of the CHARMM Force Field for Nucleic Acids, *Biopolym. Orig. Res. Biomol.*, 2001, **56**, 257–265.

5 K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell Jr., CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J. Comput. Chem.*, 2010, **31**, 671–690.

6 N. H. Powell, G. J. Clarkson, R. Notman, P. Raubo, N. G. Martin and M. Shipman, Synthesis and structure of oxetane containing tripeptide motifs, *Chem. Commun.*, 2014, **50**, 8797.

7 B. Hess, C. Kutzner, D. Van Der Spoel and E. Lindahl, GRGMACS 4: Algorithms for highly efficient, load-balanced, and scalable molecular simulation, *J. Chem. Theory Comput.*, 2008, **4**, 435–447.

8 G. Bussi, D. Donadio and M. Parrinello, Canonical sampling through velocity rescaling, *J. Chem. Phys.*, 2007, **126**, 014101.

9 H. J. C. Berendsen, J. P. M. Postma, W. F. Van Gunsteren, A. Dinola and J. R. Haak, Molecular dynamics with coupling to an external bath, *J. Chem. Phys.*, 1984, **81**,

3684–3690.

10 S. Nosé and M. L. Klein, Constant pressure molecular dynamics for molecular systems, *Mol. Phys.*, 1983, **50**, 1055–1076.

11 S. Roesner, G. J. Saunders, I. Wilkening, E. Jayawant, J. V. Geden, P. Kerby, A. M. Dixon, R. Notman and M. Shipman, Macrocyclisation of small peptides enabled by oxetane incorporation, *Chem. Sci.*, 2019, **10**, 2465–2472.

12 M. L. Strader and S. E. Feller, A flexible all-atom model of dimethyl sulfoxide for molecular dynamics simulations, *J. Phys. Chem. A*, 2002, **106**, 1074–1080.

13 X. Daura, K. Gademann, B. Jaun, D. Seebach, W. F. van Gunsteren and A. E. Mark, Peptide Folding: When Simulation Meets Experiment, *Angew. Chemie Int. Ed.*, 1999, **38**, 236–240.

14 M. Parrinello and A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

15 J. D. Hunter, Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.*, 2007, **9**, 99–104.

16 OriginPro 2019b, OriginLab Corportation, Northampton, MA, USA, 2019.