

## Supplementary information

## Supplementary information

### ***Experimental Details***

Peptide solutions were prepared prior to mass spectrometric analysis by dissolving the lyophilised peptide into 70:29:1 methanol : water : formic acid mixtures and subsequently filtering the freshly prepared sample using 20 $\mu$ m filters.

### ***ECD and CID***

Protein samples were directly infused via an external nanospray ionization source (Advion Biosciences, Ithaca, NY) into a 7T Thermo Finnigan LTQ FT mass spectrometer (Thermo Fisher Scientific, Bremen, Germany). Ion isolation was performed in the front-end linear ion trap. The isolation width was 50 Th. Automatic gain control (AGC) was used to accumulate precursor cations in the ion trap (target  $1 \times 10^6$ , maximum fill time 2 s) before transporting them into the ICR cell with a trapping voltage of 1 V.

Electrons for ECD were produced by an indirectly heated barium-tungsten cylindrical dispenser cathode (5.1 mm diameter, 154 mm from the cell, 1 mm off axis) (Heat-Wave Labs, Watsonville, CA). The current across the electrode was  $\sim 1.1$  A. Raw MS data were analyzed by use of Xcalibur 2.05 software (Thermo Fisher Scientific), where the Xtract program was used for calculating monoisotopic masses (44% fit factor, 25% remainder). ProSight PTM (<https://prosightptm.scs.uiuc.edu>) was used to search for *c*, *z* and *b*, *y* protein fragment ions. The mass accuracy for the search was set at 10 ppm. The lists of masses from ECD MS/MS spectra were searched both for “standard” *c'*, *z•* ions, and for “hydrogen transfer” *c•*, *z'* fragments. Manual inspection of MS/MS spectra confirmed *c•*, *z'* assignments.

### ***Ion Mobility Mass Spectrometry***

For IMMS experiments, ions were generated by nano-electrospray ionisation using glass capillaries pulled in-house.

Linear, low-field IMMS was performed on a Micromass QToF-1 (equipped with an off-axis z-spray source) modified in-house by the incorporation of a 5.1cm-long drift tube before the first quadrupole analyser<sup>1</sup>. Experiments were carried out at a known temperature (*ca.* 298K) and pressure of helium buffer gas (*ca.* 3.3 torr). A linearly decreasing field is applied along the longitudinal axis of the drift tube imparting a constant force on the ions. When the external force is sufficiently low in comparison to the buffer gas density, the drift velocity in the direction of the field (**v**) is directly proportional to the force applied (*qE*):

$$\mathbf{v} = K\mathbf{E} \quad \underline{\text{Equation 1}}$$

The mobility, *K*, (estimated at a range of different field strengths, ranging between 2 and 12 V/cm) is found under such conditions to be inversely proportional to the collision cross section  $\Omega$ , an indicator of the size of the analyte ion as evidenced by the scattering of neutral buffer gas molecules.

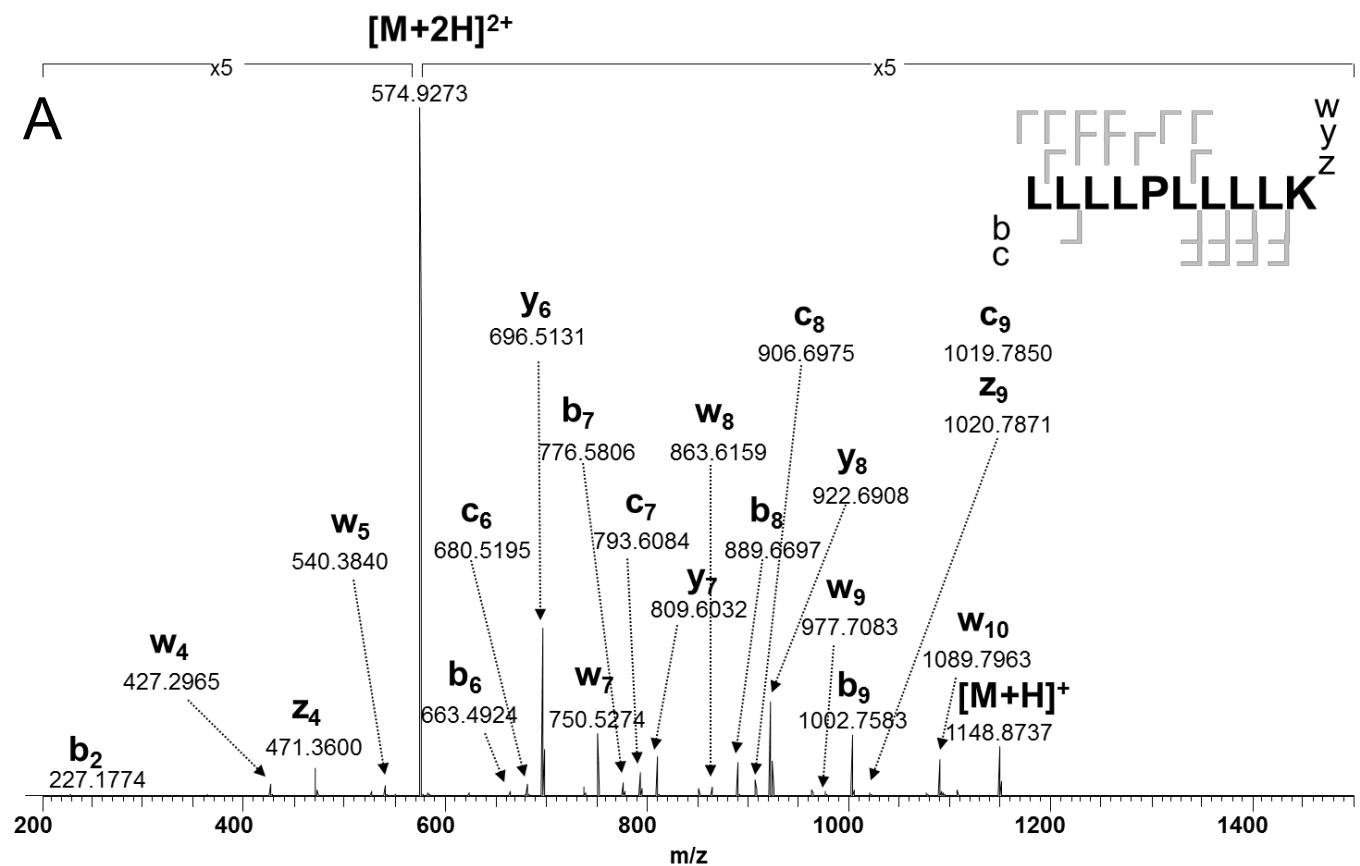
## Molecular Dynamics

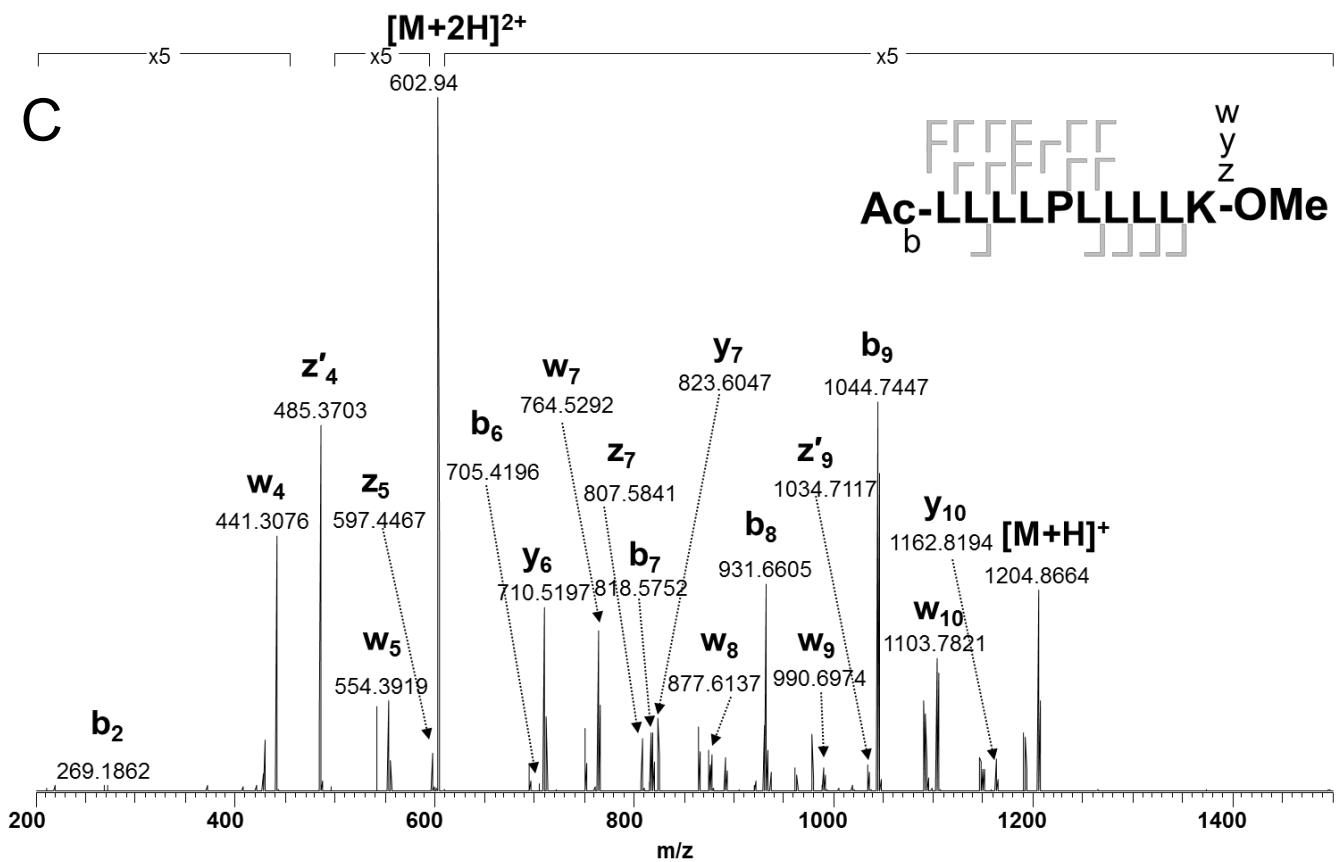
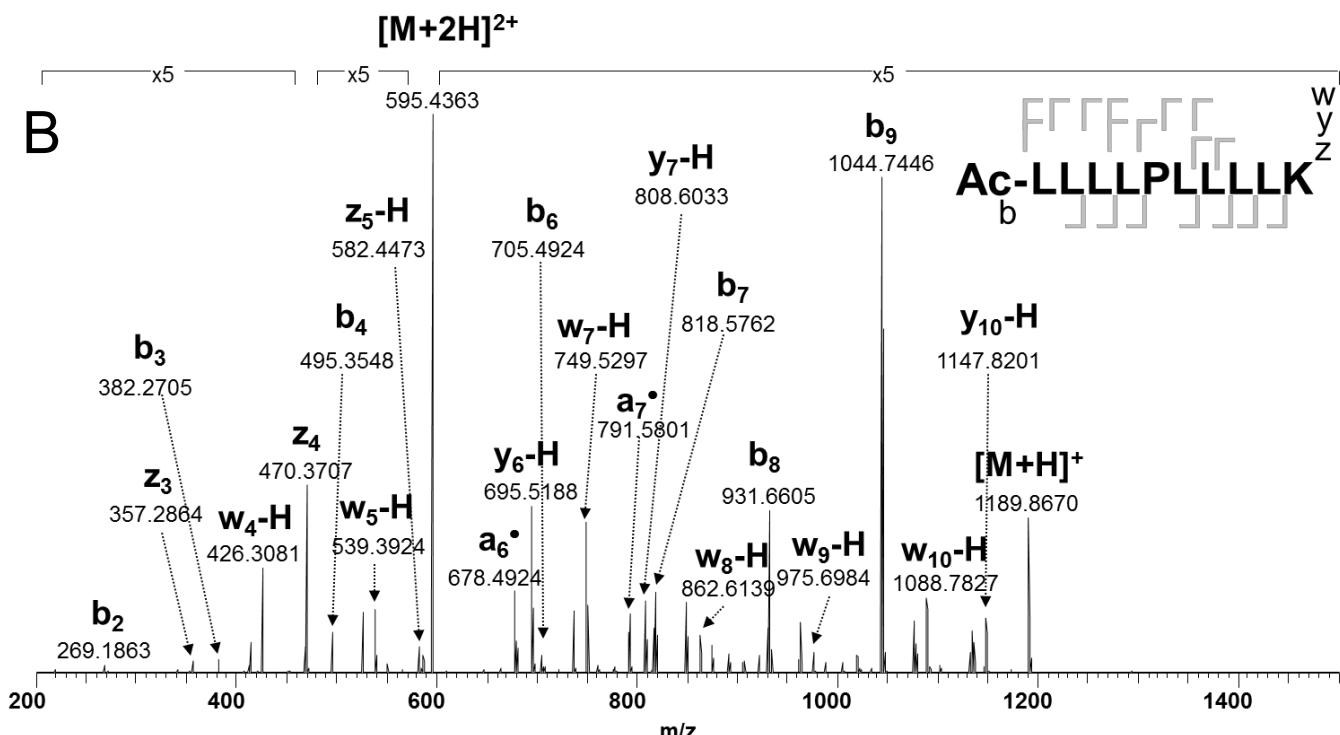
Classical atomistic molecular mechanics simulations were performed with the AMBER ff99SB empirical force field parameters<sup>2-5</sup> with additional point-charges for residues not present in the existing libraries being obtained using the multiconformational RESP procedure<sup>6-8</sup> with an optimisation at a B3LYP/6-31G(d,p)//HF/6-31G level of theory, using Gaussian software<sup>9</sup>. All simulations were performed *in vacuo*. First, to obtain a sample of potential energy minima a simulated annealing procedure was followed which involved heating the system to 800K for 50ps of simulation time, followed by a linear cooling schedule for 30ps and subsequent energy minimisation. The process was repeated 500 times, yielding a set of candidate geometries for each construct simulated, ranked by (simulated) potential energy.

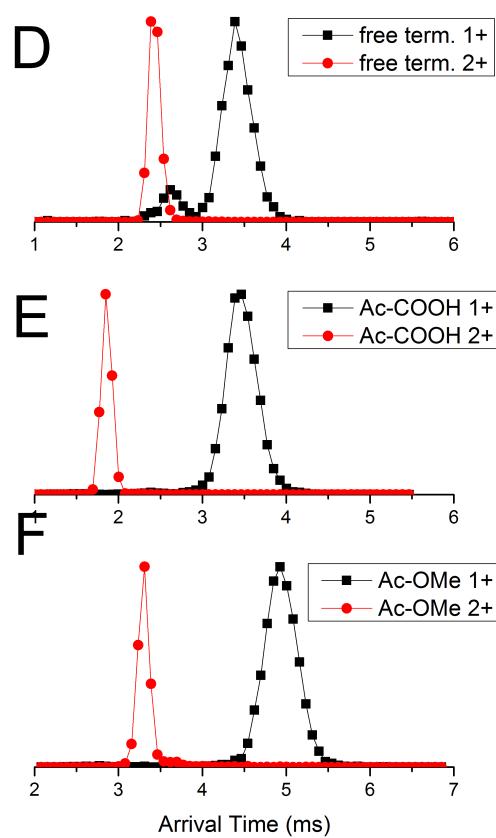
The lowest potential energy minima sampled by simulated annealing were used as initial geometries for molecular mechanics simulations. All  $[M+2H]^{2+}$  L<sub>4</sub>PL<sub>4</sub>K constructs, the  $[M+H]^+$  ion with free termini and the  $[M+H]^+$  Ac-NH<sub>2</sub>-capped L<sub>9</sub>K (protonated at K10) ion were all subject to further simulations. For each of the initial structures (three for all L<sub>4</sub>PL<sub>4</sub>K systems and two for the L<sub>9</sub>K 1+ system), the peptide was heated over 0.7ns of simulation time (with a time-step of 1fs) to 350K in fourteen 25K discrete steps, following which it was relaxed at 350K for 0.1ns before MD was run for a total time of 10ns, giving an aggregate time of 30ns for each L<sub>4</sub>PL<sub>4</sub>K construct. Coordinates were saved every picosecond of simulation time, which were used for further analysis. Secondary structure content for all MD trajectories were calculated using the DSSP model<sup>10</sup>.

Algorithms exist to calculate the orientationally-averaged collision cross section numerically for objects of arbitrarily complex geometries<sup>11, 12</sup>; these approaches allow structural assignments of detected ions on the basis of agreement of their drift speeds with those predicted for model structures, provided that the set of candidate geometries are distinguishable experimentally. The collision cross sections of all optimised geometries (generated by simulated annealing) were calculated using the trajectory algorithm<sup>11</sup>.

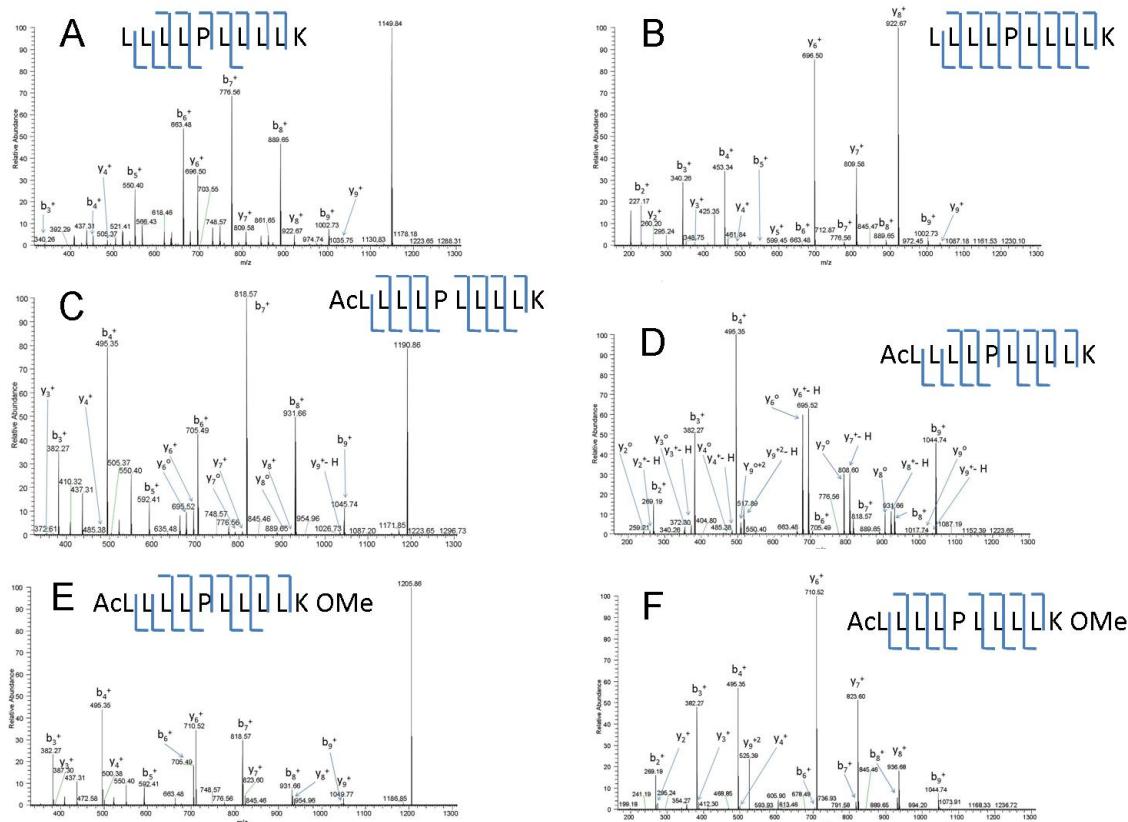
## Additional figures





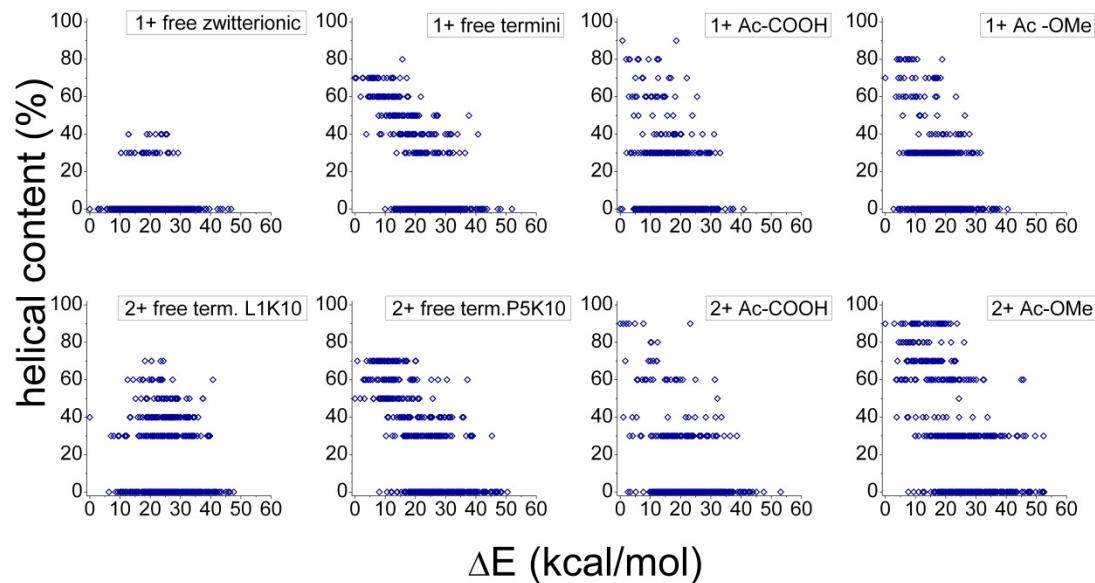


**Figure S1** ECD spectra resulting from  $[M+2H]^{2+}$  parent ion of the three  $L_4PL_4K$  constructs studied: A) free termini, B) acetyl-capped N-terminus, free acid C-terminus, C) acetyl-capped N-terminus, methyl ester capped C-terminus. D-F) Typical arrival time distributions observed at a field of 4 V/cm at a pressure of 3.0 torr D) free termini, E) acetyl-capped N-terminus, free acid C-terminus, F) acetyl-capped N-terminus, methyl ester capped C-terminus.

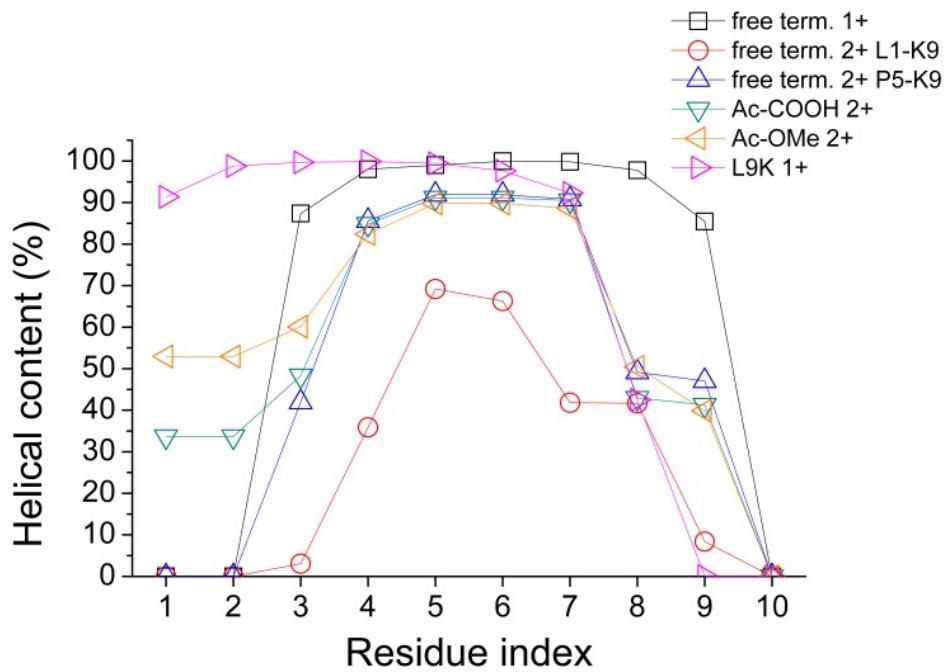


**Figure S2** CID spectra resulting from  $[M+H]^+$  (A, C, E)  $[M+2H]^{2+}$  (B, D, F) parent ion of the three  $L_4PL_4K$  constructs studied. A-B) free termini, C-D) acetyl-capped N-terminus, free acid C-terminus, E-F) acetyl-capped N-terminus, methyl ester capped C-terminus.

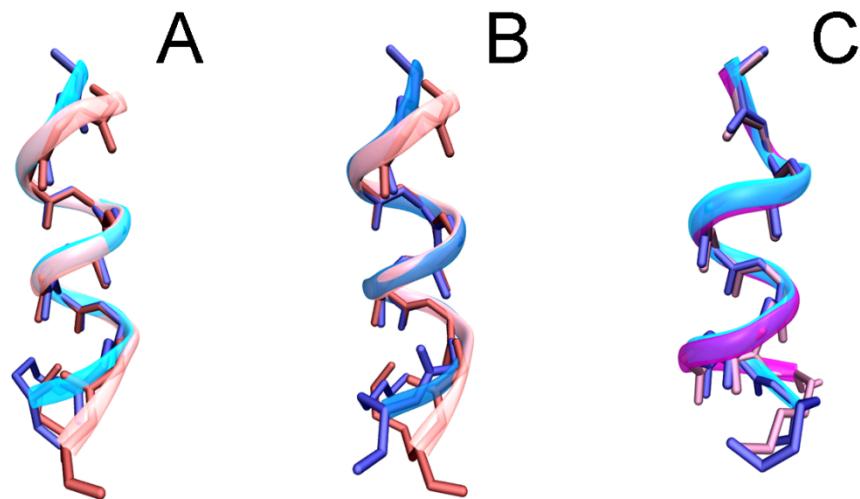
Low molecular weight b ions are significantly less abundant in the peptides with acetylated N-termini, an effect that can be attributed to the significantly lowered proton affinity of the L1 amide over the free amine. This effect is more prominent in  $[M+2H]^{2+}$  ion CID spectra (rhs). The N-terminus acetylated / free acid C-terminus  $[M+2H]^{2+}$  ion (panel D) is characterised by abundant water loss, indicating proton transfer to the C-terminal carboxyl group. This transition is disfavoured when the free N-terminal amine captures the second proton; moreover water loss is not observed in the methyl ester capped C-term. peptides, hinting to the formation of an unstable germinal diol intermediate prior to water loss. y5 is absent of very weak, a trait often seen in CID.



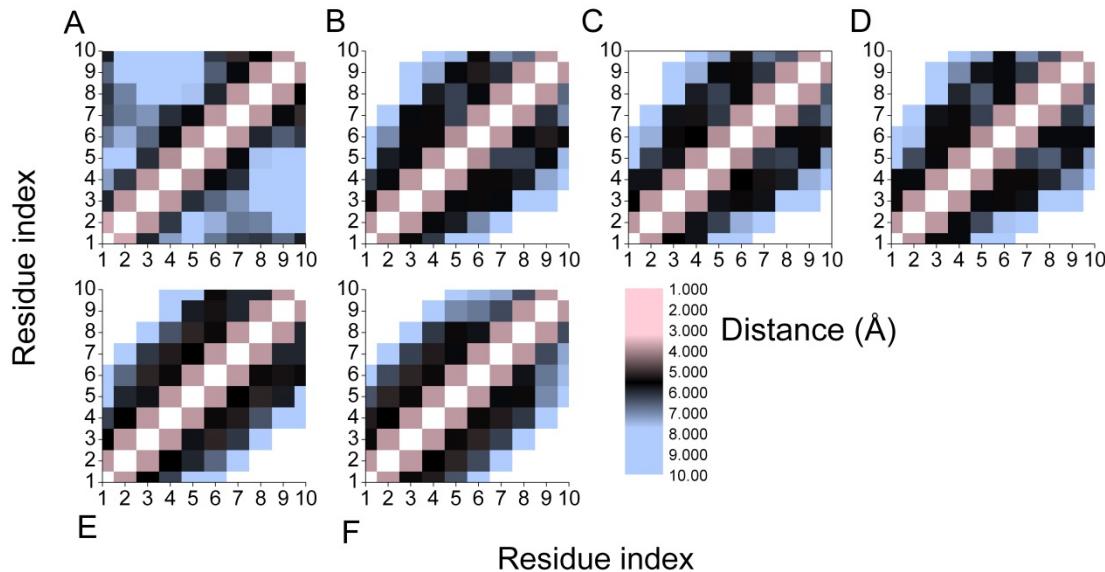
**Figure S3** Scatter plots of % helical content against  $\Delta E$ , the potential energy difference (in kcal/mol) between each sampled minimum and the lowest energy minimum for each construct sampled 500 iterations of the simulated annealing algorithm. All constructs contain a positively charged K10; Ac-COOH and Ac-OMe constructs contain an extra proton on the P5 carbonyl; for the free-termini construct the more probable state protonated at the N-terminus was also simulated. Low-lying energy minima with high helical content ( $3_{10}$ ,  $\alpha$  or  $\pi$ , calculated with the DSSP model<sup>10</sup>) were readily sampled for constructs which contained a neutral N-terminal amine. In the case of the  $[M+H]^{1+}$  ion with zwitterionic termini and  $[M+2H]^{2+}$  ion protonated at the N-terminus and K10 helical structures were very sparse among the ones of lowest energy.



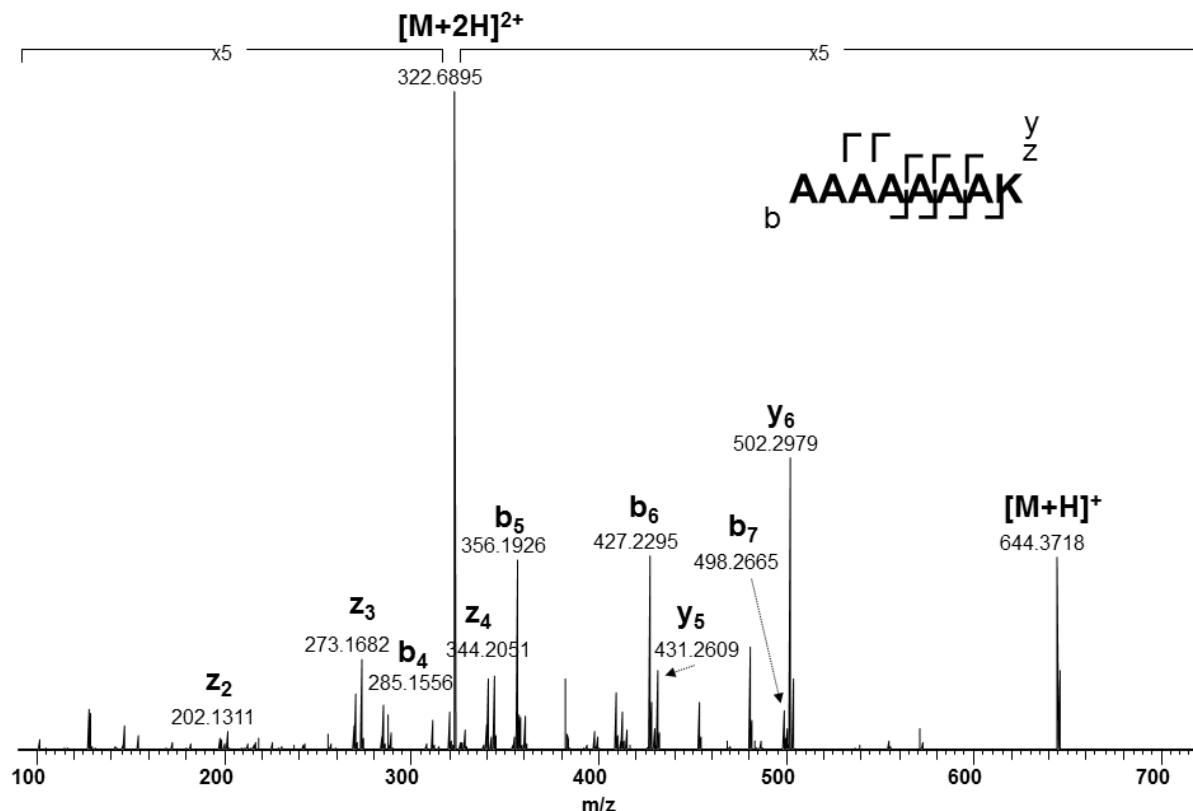
**Figure S4.** Helical content per residue (either  $\text{3}_{10}$ - $\text{\alpha}$ - or  $\pi$ -helix, calculated with the DSSP model) for all peptides simulated by MD. P5 is incorporated in helical stretches, showing that secondary structural preferences in this leucine-based peptide can override the “helix breaking” character of proline. Helicity is less favoured in the construct protonated at the N-terminus.



**Figure S5** Backbone heavy-atom alignments between lowest-energy structures sampled by simulated annealing. Backbone heavy atoms and the K10 side-chain are shown (terminal atoms have been omitted along with all other side-chains as well as protons for clarity) A) Alignment between residues 2-8 for L<sub>9</sub>K 1+ (red) and free-termini 1+ L<sub>4</sub>PL<sub>4</sub>K (cyan). B) Alignment between residues 2-8 for L<sub>9</sub>K 1+ (red) and free-termini 2+ L<sub>4</sub>PL<sub>4</sub>K protonated at P5 and K10 (blue). C) Alignment between residues 1-8 for 1+ L<sub>4</sub>PL<sub>4</sub>K (cyan) and 2+ L<sub>4</sub>PL<sub>4</sub>K protonated at P5 and K10 (purple). Clearly, all three constructs share a common structural core, with greatest differences emerging near the termini. The two L<sub>4</sub>PL<sub>4</sub>K constructs differ from L<sub>9</sub>K in the backbone configuration of L1, which in the former forms a hydrogen bond with L4, the latter forming a hydrogen bond with L5 instead. Due to the length of the peptide, a helix can form with an internal proline residue, which is stabilised by the interaction between the K10 ammonium and the L7 and L8 carbonyls. Addition of a proton at the P5 carbonyl does not prevent the lowest energy geometry sampled to also form this helical structure.



**Figure S6**  $\alpha$ -carbon distance matrices for six constructs investigated by molecular dynamics simulations: A) free-termini  $[M+2H]^{2+}$  ion protonated at the N-terminus and at K10; B-C-D) free termini, Ac-COOH and Ac-OMe  $[M+2H]^{2+}$  ions protonated at the P5 carbonyl and at K10; E) free-termini  $[M+H]^{1+}$  construct protonated at K10; F) acetyl-amine-capped  $[M+H]^{1+}$   $L_9K$  ion protonated at K10. The distance range shown is between 1 and 10 Å. Intermediate distances are shown in black in order to increase the contrast over that range. Patches of contacts at  $i-i+4$  and  $i-i+3$  separations, which are characteristic of peptide helices stabilised by H-bonding, are exhibited by all constructs with the exception of the free-termini  $[M+2H]^{2+}$  species protonated at the N-terminus. For the latter, the N-terminal leucine forms stable contacts with L7, L8 and K10 which are driven by the need of the N-terminal ammonium to be solvated by the peptide. Notably, the behaviour of all constructs containing a charge at P5 and K10 behave in a nearly identical manner, indicating that for these relatively small peptide ions the gas-phase conformation is strongly affected by the location of the charge, the effect of the capping groups per se being less pronounced (beyond their ability to change the localisation of the extra proton(s)).



**Figure S7** ECD spectrum of  $\text{NH}_2\text{-A}_7\text{K-COOH}$   $[\text{M}+2\text{H}]^{2+}$  ion. The  $[\text{M}+\text{H}]^+$  ion of this octopeptide is expected to form a stable gas-phase helix, owing to the positioning of lysine at the N-terminus. Notably this peptide does give rise to *b* fragment ions (*b*4-7) by ECD, an observation that is consistent with the behaviour of  $\text{L}_4\text{PL}_4\text{K}$ . On the other hand,  $\text{KL}_4\text{PL}_4$  and  $\text{RL}_4\text{PL}_4$  do not form *b* ions<sup>13</sup>, which is consistent with their lower gas-phase helical propensity.

## References

1. B. J. McCullough, J. Kalapothakis, H. Eastwood, P. Kemper, D. MacMillan, K. Taylor, J. Dorin and P. E. Barran, *Analytical Chemistry*, 2008, **80**, 6336-6344.
2. D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *Journal of Computational Chemistry*, 2005, **26**, 1668-1688.
3. D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel and P. Kollman, *Computer Physics Communications*, 1995, **91**, 1-41.
4. P. K. Weiner and P. A. Kollman, *Journal of Computational Chemistry*, 1981, **2**, 287-303.
5. V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins-Structure Function and Bioinformatics*, 2006, **65**, 712-725.
6. J. M. Wang, P. Cieplak and P. A. Kollman, *Journal of Computational Chemistry*, 2000, **21**, 1049-1074.
7. W. D. Cornell, P. Cieplak, C. I. Bayly and P. A. Kollman, *Journal of the American Chemical Society*, 1993, **115**, 9620-9631.
8. C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *Journal of Physical Chemistry*, 1993, **97**, 10269-10280.
9. M. J. T. Frisch, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; and Pople, J. A.; Gaussian, Inc., Wallingford C, 2004.
10. W. Kabsch and C. Sander, *Biopolymers*, 1983, **22**, 2577-2637.
11. M. F. Mesleh, J. M. Hunter, A. A. Shvartsburg, G. C. Schatz and M. F. Jarrold, *Journal of Physical Chemistry*, 1996, **100**, 16082-16086.
12. A. A. Shvartsburg and M. F. Jarrold, *Chemical Physics Letters*, 1996, **261**, 86-91.
13. H. J. Cooper, *Journal of the American Society for Mass Spectrometry*, 2005, **16**, 1932-1940.