Soft Matter

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/softmatter

Soft Matter

RSCPublishing

COMMUNICATION

The influence of the localised charge of C- and N-termini on peptide self-assembly

Cite this: DOI: 10.1039/x0xx00000x

C. Bortolini,[a] N. C. Jones,[b] S. V. Hoffmann,[b] F. Besenbacher [a] and M. Dong*[a]

Received 00th January 2015,
Accepted 00th January 2015

DOI: 10.1039/x0xx00000x

www.rsc.org/

The charge of a peptide influences final assembled structures. It is important to consider not only global charge, but also local, such as that found on the terminal residues. This work investigates the change of peptide self-assembly through the selection of different amino acid sequences and by varying the local charge of the residues on the C- and N- termini.

Self-assembled structures are characterised by a higher degree of order than the starting units; their assembly is usually spontaneous and reversible.[1] During the self-assembly process, isolated building blocks (e.g. nanoparticles, molecules) are capable of organising into well-ordered structures[2] through local interactions, without the need for external forces.[3, 4] Self-assembly of biological compounds is mainly driven and influenced by factors such as temperature, concentration, pH,[5] charge and shaking speed[6] during incubation. As previously shown in a study looking at the position of individual units (hydrophobic residues) within a sequence[7], it is not only the global charge of the structure which needs to be considered, but also local charge on the residues plays an important role.

One factor affecting peptide self-assembly that is not yet well understood is the importance of the position of the carboxyl-terminus (C-terminus) within the sequence. The C-terminus is known to be involved in amyloid formation[8, 9], but its role has never been extensively studied. Therefore in this study the C-terminus has been swapped with the N-terminus and the effect of this modification has been assessed. The choice of sequence is crucial, so four different types of amino acids were chosen based on the polarity and charge of their side chains (also referred to as the R-group) - it is customary to distinguish between non-polar (or hydrophobic), polar un-charged (neutral side-group), polar charged (R-group is charged either positive or negative at physiological pH) and aromatic residues.

The amino acids used in this study are:

- Alanine (A). This small non-polar organic compound is characterised by the presence of a methyl group ($-CH_3$) as the side chain. Being hydrophobic, it is expected to actively contribute to the formation of ordered structures, possibly peptide fibrils.[10-13]
- Aspargine (N). A polar un-charged molecule which has an amide group in its side chain and is hydrophilic.
- Aspartic Acid (D). Containing a carboxyl group which has low pKa in its side chain, this residue is acidic. The polarity and negative charge of the R-group mean this amino acid can be easily deprotonated at physiological pH.
- Histidine (H). This residue is characterised by an imidazole group as the side chain making it an aromatic amino acid. The imidazole group, having pKa = 6, gives histidine a net positive charge at physiological pH. Below pH 6, the aromatic ring is likely to be protonated, however, it maintains its aromaticity at all pH values (i.e. the imidazole ring retains the capability of cyclic delocalisation and resonance).

Combining these residues, three different sequences were designed, AAAAANNNNN (A5N5), DDDDDNNNNN (D5N5) and DDDDDHHHHH (D5H5). Another 3 samples were synthesised with the same sequences, but with the C-terminus swapped, NNNNNAAAAA (N5A5), NNNNNDDDDD (N5D5) and HHHHHDDDDD (H5D5). Charge charts and the isoelectric points are the same for each sequence and the corresponding sample with the C-terminus swapped, as the residue composition is the same for both.

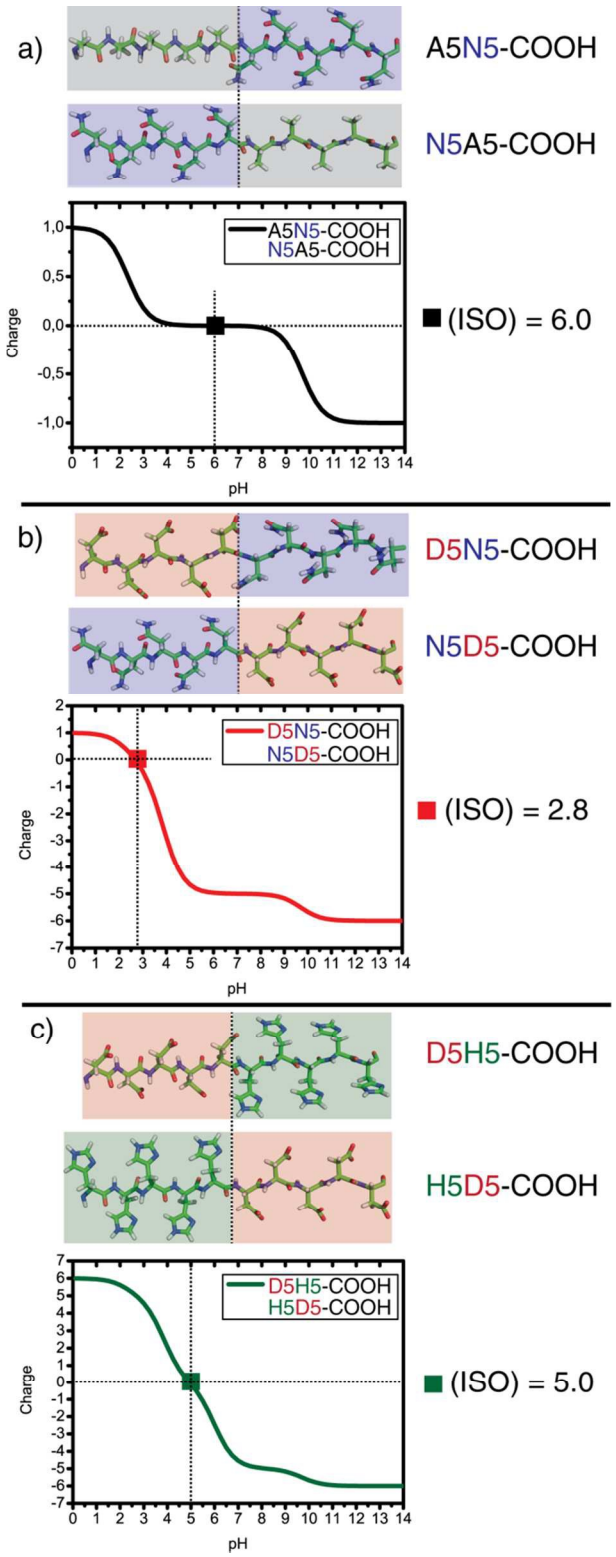

Fig. 1 Peptides charge charts and molecular models. Isoelectric points are labelled with a square on the charge curve. a) A5N5 and N5A5 peptides are made of hydrophobic alanines and hydrophilic aspargines. b) D5N5 and N5D5 peptides are made of acidic aspartic acids and hydrophilic aspargines. c) D5H5 and H5D5 peptides are made of acidic aspartic acids and basic histidines.

In general, in the case of short peptides, supramolecular assemblies are mainly formed by the hydrophobic residues interacting with each other. We expect our short sequences containing a significant amount of hydrophobic amino acids (i.e. A5N5 and N5A5) both to form fibrillar structures, similar to amyloids, which are characterised by a significant amount of β-sheet in their secondary structure.

Figure 1 shows molecular models and charge charts associated with the peptide sequences. Isoelectric points are indicated with a square, on the charge curve. *Figure S1* compares the charge curves of all sequences.

A5N5 and N5A5 are made of 5 alanines and 5 aspargines that make the peptide amphiphilic. For D5N5 and N5D5 alanine is replaced with aspartic acid. The global charge is now negative as the charge rapidly decreases even at low pH values. D5H5 and H5D5 possess only charged amino acids as aspargines are replaced by histidines. This sequence is very sensitive to pH changes, with very high charge values, +6, at low pH and -6 at high pH values.

Incubation of the monomers at room temperature (RT) in a suitable buffer (Millipore water in our case) needs to occur at an appropriate pH value. Based on previous experience,[7, 10] a pH value of 2.5 was chosen, which is below the isoelectric point of all sequences, i.e. when the charge is positive.

After 2 days of incubation, the solutions containing the self-assembled peptides were analysed using SR-CD and tapping mode AFM.[14-16] AFM images and line profiles are shown in Figure S3 in the *Electronic Supplementary Information (ESI)*.

The nature of the secondary structure of these peptide nanostructures is also of interest, and also to what extent this folding is affected by swapping the carboxyl and amino-termini. The qualitative and quantitative content of secondary structure of a self-assembled peptide structure can be assessed by a technique called circular dichroism (CD) spectroscopy.[17] The synchrotron radiation CD facility at ASTRID2, Aarhus University, was used for this investigation.[18, 19]

Figure 2 shows SRCD spectra for the sequences under investigation. *Figure 2a* presents the A5N5 and N5A5 spectra: interestingly, even as the tertiary structures appeared to be significantly different in the AFM analysis, the SRCD spectra are highly similar. In *Figure 2b*, it is evident that the situation is the same: D5N5 and N5D5 exhibited different tertiary structures, but the secondary structures appear unchanged. However the sequences having only charged residues, i.e. D5H5 and H5D5, display very different secondary structures as seen in *Figure 2c*.

A quantitative analysis of these CD spectra can be carried out by computing the specific amount of secondary structures within each SRCD spectrum. *Table S1* in *ESI* presents the computed content of peptides predicted using the web based server Dichroweb[20, 21].

If A5N5 and N5A5 are compared, it is clear that the secondary structure content is almost identical, with no more than 4% difference in the amount of random coil (51% vs. 53%), α-helices (4% vs. 0%) or -strands (27% vs. 28%).

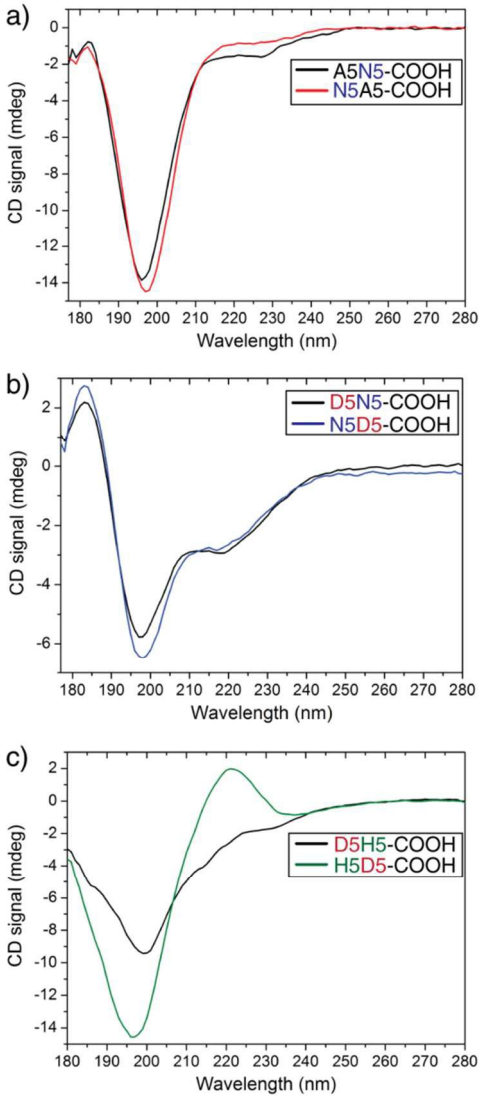

Fig. 2 SRCD measurements of the six sequences previously analysed by using AFM. Peptide solutions utilised for the measurements are similar to the AFM experiments. a) SRCD spectra of A5N5 (black line) and N5A5 (red line); b) Spectra of D5N5 (black line) and N5D5 (blue line); c) Spectra of D5H5 (black line) and H5D5 (green line).

A similar situation is found when comparing the calculated percentage of secondary structures for D5N5 and N5D5. In this case, the content showed an even higher degree of similarity: random coil (43% vs. 44%), α-helices (8% vs. 7%) and β-strands (34% vs. 33%). Comparing the results for D5H5 and H5D5 a different situation is found. D5H5 is much richer in α-helices than H5D5 (25% vs. 5%), while the latter is more abundant in random coil (53% vs. 39%) and β-strands (26% vs. 20%). These results suggest that, apart from the sequences containing amino acids which are charged and/or aromatic, the secondary structure does not change when the C-terminus is displaced, even though the tertiary structure can be highly affected.

Figure 3 offers a graphical view of the data presented in *Table S1* in *ESI*. It is clear that all sequences produced a high amount of random coil (above 0.38). The contribution to the CD signal of the turns is very similar having values ranging from 0.15 to 0.18. Except for D5H5, the content of β-sheets predominates if compared to α-helices, being on average 5 times higher.

The lower panel of *Figure 3* presents AFM images of the fibrillar-like nanostructures self-assembled in solution and deposited onto a Mica surface to facilitate the analysis.

The first sequence, A5N5, formed tiny peptide aggregates (see line profile in *ESI*) 0.8 ± 0.1 nm tall. Conversely, the same sequence with swapped C-terminus, N5A5, gave rise to twisted amyloid-like fibrils 4.0 ± 0.2 nm tall. By replacing alanine with aspartic acid (D5N5) non-twisted fibrils 3.0 ± 0.2 nm tall were observed. Its swapped counterpart (N5D5) exhibited a porous film structure (see *Figure S3* in *ESI*). Even after further incubation for 1 day at 35°C the structure is unaltered. By replacing aspargine with histidine, D5H5, only particles and small peptide fragments 1.4 ± 0.1 nm tall are visible in the AFM analysis. As for the N5D5 sequence, incubation at higher temperature did not change the outcome.

For the last sequence, H5D5, a uniform array of peptide patches was observed (see *Figure S3* in *ESI*).

It is worth observing that within the 4 alanine-free sequences all residues are hydrophilic and either neutral or charged. Not having an amphiphilic sequence suggests that peptide-peptide interactions may lack of an easily predictable driving force; to ease the self-assembly we increased the temperature to 35°C, but thermal energy did not help the peptide assembly. Nevertheless, SR-CD data exhibit the presence of a considerable amount of ordered secondary structures (i.e. alpha-helices, beta-sheet and turns) indicating a considerable level of organisation for the bulk peptide in solution (i.e. there are well-defined patterns of hydrogen bonds).

The different sequences investigated gave rise to very different peptide assemblies. It is evident that changing the position of the nearest neighbour's residues on the C- and N- termini leads to a considerably different tertiary structure, regardless of the amino acids composing the sequence. The tertiary structure is therefore highly influenced not only by the composition of the sequence, but also by the local charge provided by the residues in the proximity of the C- and N-termini.

Journal Name RSCPublishing

COMMUNICATION

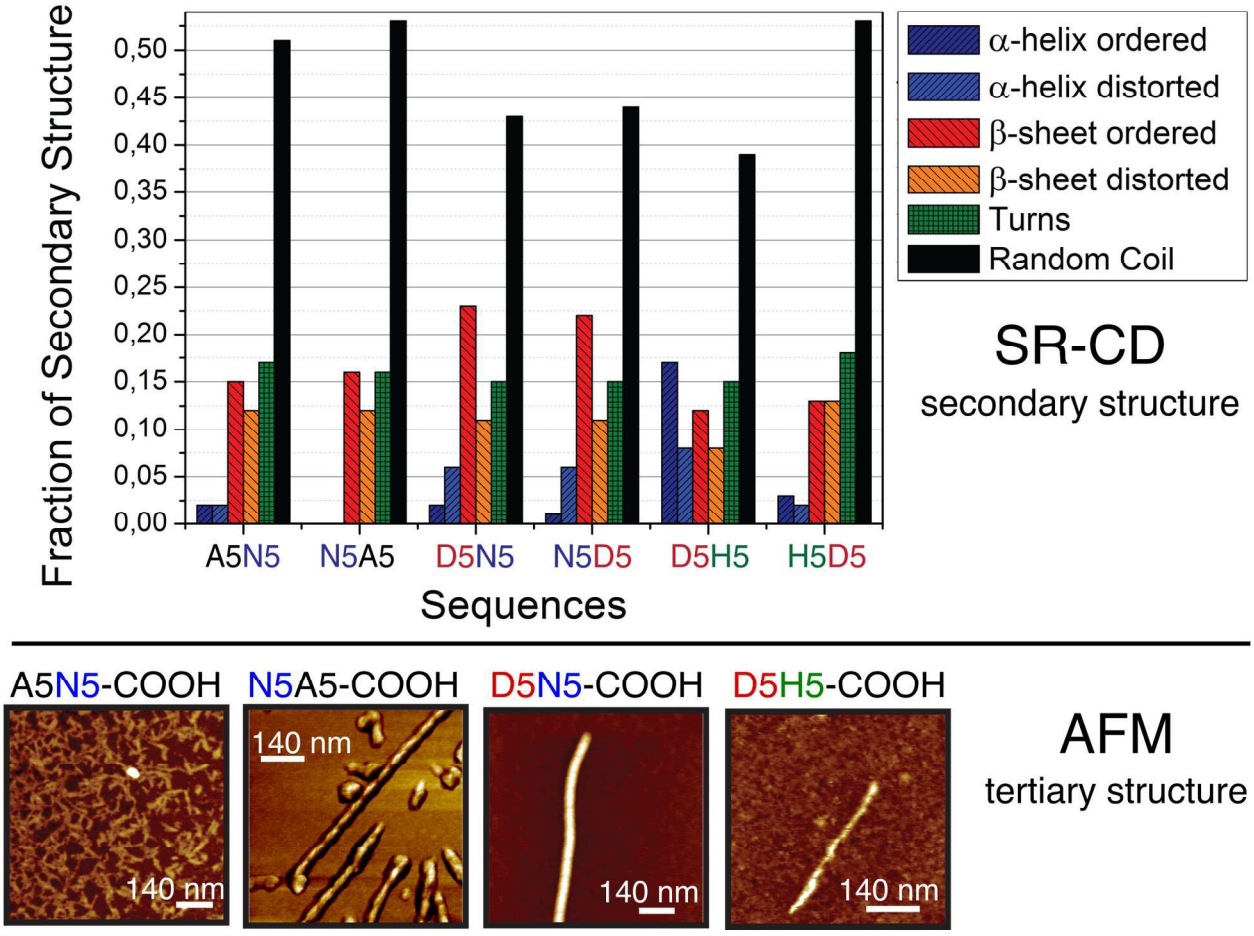

Fig. 3 (Higher panel) Histograms indicating the fractions of secondary structure measured by SR-CD. Cool colours indicate the α-helices, warm hues the β-sheets, green the turns and black the random coil. (Lower panel) AFM images of the peptides incubated for 2 days at pH 2.5. The sequence is listed above each AFM image.

Conclusion

In this work, the effect of changing the nearest neighbour's residues on the C- and N- termini on the secondary and tertiary structures of *ad hoc* designed synthetic peptides has been investigated. Our results can be combined with the previous observations made when investigating the role of hydrophobic residues[7]. Peptide sequences have been carefully chosen based on their different properties given by the specific side chains. AFM has been employed to characterise the tertiary structure, whilst SR-CD has proven to be helpful in identifying and quantifying the secondary structures. The findings presented here indicate that while some peptides fold into different tertiary structures when the nearest neighbour's residues on the C- and N- termini are changed, the secondary structure is apparently unaltered. An exception to this was found for the sequence containing charged and aromatic residues, where both secondary and tertiary structures were affected. This investigation suggests that information regarding the global charge of a peptide is not sufficient to fully understand peptide self-assembly, as local charge can have a significant effect and therefore cannot be neglected.

The results of this investigation may be helpful for scientists interested in peptide self-assembly. These new insights provide a new knowledge of peptide self-assembly and how it is affected by modifying the nearest neighbour residues of the C- and N- termini.

Acknowledgements

This work was supported by grants from the Danish National Research Foundation and the Danish Research Agency through support for the iNANO Center and the Danish Council for Strategic Research to iDEA project. M.D. acknowledges a STENO Grant for the Danish Research Council and the VKR Young Investigator Program in Denmark.

Notes and references

[a] Interdisciplinary Nanoscience Center (iNANO), Aarhus University, Gustav Wieds 14, Building 1590, Aarhus C, Denmark; E-mail: dong@inano.au.dk.

[b] ISA, Department of Physics and Astronomy, Aarhus University, 8000 Aarhus C, Denmark.

Electronic Supplementary Information (ESI) available: ESI includes the experimental details, a charge chart combining the sequences and their isoelectric points, SR-CD and Absorbance spectra of all peptides analysed in this work and a table containing data corresponding to Figure 4 in the main text. In addition, ESI include the AFM images of all peptide sequences and their line profiles. See DOI: 10.1039/c000000x/

1. G. M. Whitesides, J. P. Mathias and C. T. Seto, *Science*, 1991, **254**, 1312-1319.
2. L. Liu, Y. Li, D. Xia, C. Bortolini, S. Zhang, Y. Yang, J. S. Pedersen, C. Wang, F. Besenbacher and M. Dong, *Nanoscale*, 2015.
3. R. V. Ulijn and A. M. Smith, *Chem Soc Rev*, 2008, **37**, 664-675.
4. I. Choi, I. S. Park, J. H. Ryu and M. Lee, *Chem Commun*, 2012, **48**, 8481-8483.
5. Y. Arii and K. Hatori, *Biochem Bioph Res Co*, 2008, **371**, 772-776.
6. W. Qiang, K. Kelley and R. Tycko, *J Am Chem Soc*, 2013, **135**, 6860-6871.
7. C. Bortolini, L. Liu, T. M. A. Gronewold, C. Wang, F. Besenbacher and M. Dong, *Soft Matter*, 2014, **10**, 5656-5661.
8. J. T. Jarrett, E. P. Berger and P. T. Lansbury, *Biochemistry-Us*, 1993, **32**, 4693-4697.
9. L. Liu, L. Niu, M. Xu, Q. Han, H. Duan, M. Dong, F. Besenbacher, C. Wang and Y. Yang, *Acs Nano*, 2014, **8**, 9503-9510.
10. C. Bortolini, L. Liu, Z. Li, K. Thomsen, C. Wang, F. Besenbacher and M. Dong, *Advanced Materials Interfaces*, 2014, **1**, n/a-n/a.
11. T. P. Knowles, A. W. Fitzpatrick, S. Meehan, H. R. Mott, M. Vendruscolo, C. M. Dobson and M. E. Welland, *Science*, 2007, **318**, 1900-1903.
12. M. B. Hovgaard, M. D. Dong, D. E. Otzen and F. Besenbacher, *Biophys J*, 2007, **93**, 2162-2169.
13. D. E. Otzen, C. B. Andersen, C. Rischel, M. B. Hovgaard, M. Dong, F. Besenbacher and J. S. Pedersen, *Biophys J*, 2007, 220a-221a.
14. M. D. Dong, M. B. Hovgaard, W. Mamdouh, S. L. Xu, D. E. Otzen and F. Besenbacher, *Nanotechnology*, 2008, **19**.
15. D. J. Muller and A. Engel, *Biophys J*, 1997, **73**, 1633-1644.
16. D. J. Muller, D. Fotiadis, S. Scheuring, S. A. Muller and A. Engel, *Biophys J*, 1999, **76**, 1101-1111.
17. B. M. Bulheller, A. Rodger and J. D. Hirst, *Phys Chem Chem Phys*, 2007, **9**, 2020-2035.
18. A. J. Miles, S. V. Hoffmann, Y. Tao, R. W. Janes and B. A. Wallace, *Spectrosc-Int J*, 2007, **21**, 245-255.
19. A. J. Miles, R. W. Janes, A. Brown, D. T. Clarke, J. C. Sutherland, Y. Tao, B. A. Wallace and S. V. Hoffmann, *J Synchrotron Radiat*, 2008, **15**, 420-422.
20. L. Whitmore and B. A. Wallace, *Nucleic Acids Res*, 2004, **32**, W668-W673.
21. L. Whitmore and B. A. Wallace, *Biopolymers*, 2008, **89**, 392-400.

Even though the exchange of N- and C-termini has a significant impact on peptide self-assembly and peptide structure, the secondary structure do not seem to be influenced.