COMMUNICATION

Supramolecular dimerisation of middle-chain Phe pentapeptides *via* CB[8] host–guest homoternary complex formation†

Silvia Sonzini, Seán T. J. Ryan and Oren A. Scherman*

Pentapeptides containing a Phe residue in the middle of the sequence exhibit ternary complex formation in the presence of cucurbit[8]uril, thus opening new perspectives on supramolecular peptide dimerisation studies.

Supramolecular host–guest chemistry in peptide sequences is a clever approach to reversibly control not only peptide conformations, but also their aggregation into dimers and higher order structures. Peptide dimers can be useful as scaffolds to bind DNA, as well as proteins, to regulate their activity.[1,2] Additionally, dimers may become important building blocks to further study hierarchical aggregation pathways of greater complexity. A simple dimerisation technique exploiting host–guest supramolecular interactions, which does not require synthetic modification but only a Phe residue in the sequence, is herein explored.

Cucurbit[*n*]uril, a family of synthetic, macrocyclic host molecules, has already been shown to selectively bind specific amino acid side chains.[3–5] Such interactions have been exploited for sensing of protonated and aromatic residues.[6,7] One of the larger members of this family, cucurbit[8]uril (CB[8]), can simultaneously accommodate two guest molecules inside its cavity.[8–10] This has allowed the dimerisation of large (bio)molecules, which have suitable binding motifs incorporated at their termini. A previous study on the strength of CB[8] binding with aromatic amino acids as 1 : 2 homoternary complexes was carried out by Urbach *et al.*,[11] highlighting the importance of having the binding residue at the N-terminal position. Furthermore, complexes between CB[8] and aromatic residues have also been reported in bioconjugate systems such as BSA and PEG through a heteroternary (1 : 1 : 1) complex.[12]

None of the reported studies, however, suggest the possibility of significant ternary complex formation when the guest residues exist in the middle of the peptide sequence. In fact, the contrary was even proposed.[11] We report three examples of pentapeptide

Melville Laboratory for Polymer Synthesis, Department of Chemistry, University of Cambridge, Lensfield Road, CB2 1EW, Cambridge, UK. E-mail: oas23@cam.ac.uk
† Electronic supplementary information (ESI) available. See DOI: 10.1039/c3cc45420g
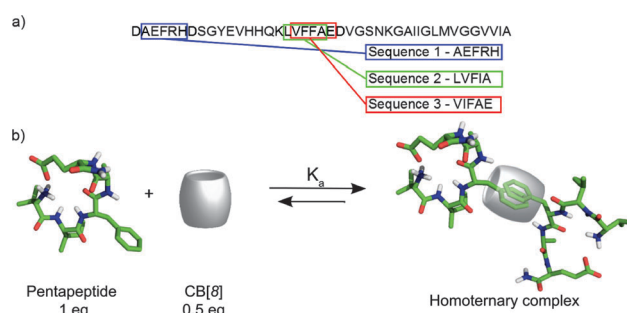

Fig. 1 (a) Aβ 1–42 sequence with highlighted chosen fragments. (b) Schematic illustrating the homoternary complex formation *via* host–guest interaction of pentapeptides and CB[8].

sequences with a general structure $H_3N^+–X_1X_2F_3X_4X_5–CONH_2$, which are capable of exhibiting strong 2 : 1 binding, through the middle-chain Phe residue, with CB[8] (Fig. 1b). The three chosen sequences, reported in Fig. 1a, are short fragments adapted from the wild type human Amyloid Beta 1–40/42 (WT Aβ) sequences.

The interaction of the Phe residues with CB[8] was investigated in an effort to better understand the oligomerisation pathways of Aβ. Sequence **1** (AEFRH) was directly taken from residues 2–6 of the WT Aβ. Sequences **2** (LVFIA) and **3** (VIFAE) were derived from residues 17–21 and 18–22, respectively, of the WT peptide. In the latter two sequences, however, both residues 19 and 20 are Phe, therefore, one amino acid per fragment has been exchanged for Ile, which maintains the hydrophobic nature of the pentapeptides without inserting a second aromatic residue.[13] This avoided potential CB[8] interactions with more than one residue per chain without dramatically changing the surrounding environment, thus allowing us to better comprehend mid-chain Phe–CB[8] interactions.

The pentapeptides were synthesised by solid phase peptide synthesis (SPPS) using standard Fmoc protocols and characterised by HPLC and ESI-MS to ensure their purity (see ESI†). Ternary complex formation with CB[8] was studied using three different methods, all of which support strong 2 : 1 homoternary complexation. The sequences were first analysed by fluorescence titration to evaluate any difference in the presence or absence of CB[8]. When Phe is irradiated, its emission maximum is 284 nm.[14] This is readily

observed for a solution of **1** in the absence of CB[8], which is a sequence suitably hydrophilic and well solvated in buffer. On the other hand, fibres formed from Phe-rich sequences display a red-shifted emission at about 303 nm, which is attributed to π–π stacking.[15,16] A change in the emission band at 303 nm in the presence of CB[8] is observed for all of the three sequences (see Fig. S9–S11, ESI†) and is indicative of Phe dimerisation in the cavity and variations in the aggregation state, as expected.

Sequence **1** initially shows an intrinsic major peak at 284 nm and a shoulder at 303 nm; upon increasing CB[8] concentration a dramatic enhancement of the peak at 303 nm is observed. **2**, the most hydrophobic of the three sequences, exhibits only one peak at 303 nm, which is quenched upon addition of CB[8]. **3** alone presents two intrinsic peaks at 284 and 303 nm, which both increase upon CB[8] addition, however, the peak at 284 nm transitions from being the minor peak to the major one.

We evaluated the change in emission at 284 and 303 nm of **1–3** upon increasing CB[8] concentration (Fig. 2). Notably, analysing the ratio between fluorescence upon CB[8] addition (F_i) and the intrinsic fluorescence (F_0) all of the sequences show a change in their emission behaviour at a 0.5 ratio, supporting the theory of a 2:1 binding complex. In addition, **1** presents a second binding event at higher CB[8] concentrations, which might be related to interactions with the C-terminal His residue. Further studies, in full agreement with the CB[8] data, were also conducted evaluating any changes in the emission spectra upon addition of CB[7] (see Fig. S1, ESI†).

^{1}H-NMR spectra of host–guest complexes usually exhibit shifted and broadened peaks, therefore ^{1}H-NMR was also used to investigate 2:1 complex formation between the pentapeptide sequences and CB[8]. The concentration of the peptides was held constant (either at 0.5 mM or 0.25 mM in the case of **2**) and the spectra were recorded at various concentrations of CB[8]. The titration spectra of **3** are reported in Fig. 3, while the spectra of **1** and **2** are available in the ESI† (see Fig. S12 and S13).

^{1}H-NMR spectra of **3** (Fig. 3) showed a clear upfield shift of Phe aromatic peaks from 7.26 and 7.20 ppm to 6.68 and 6.40 ppm, respectively. Additionally, the H_α and CH_2 protons of Phe exhibit an upfield shift as well. The aromatic peaks are completely shifted upfield at a ratio of 2:1 peptide:CB[8]. These shifts in the Phe
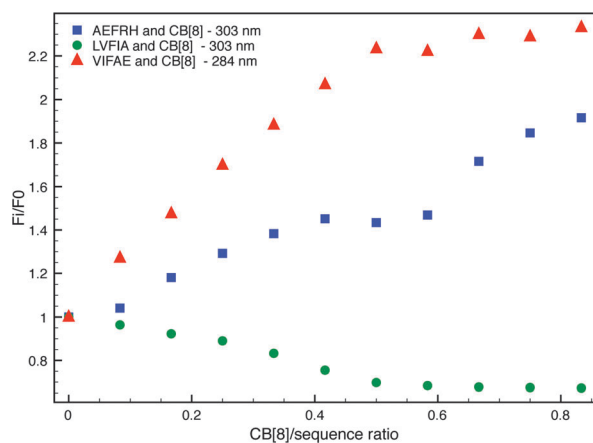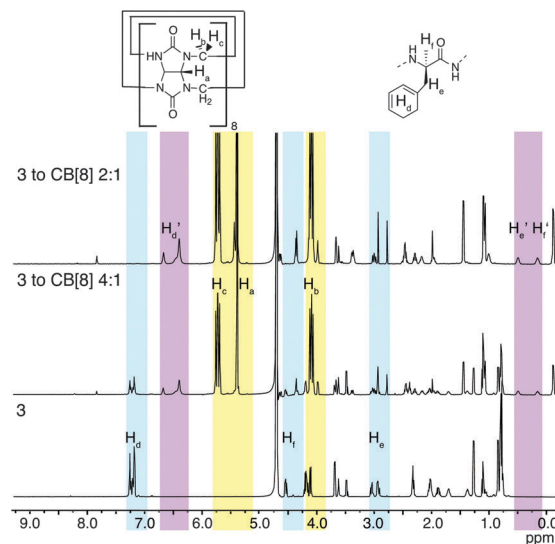

Fig. 3 ^{1}H-NMR in D_2O of sequence **3** (0.5 mM) with an increasing amount of CB[8].

aromatic signals have been reported previously as evidence of homoternary complex formation between Phe and CB[8].[11] Sequence **1** exhibited similar behaviour; in fact, at a 2:1 ratio of pentapeptide to CB[8] the only peak in the region of 7.3 ppm is at 7.36 ppm, which is one of the His side chain peaks, together with the signal at 8.64 ppm (see Fig. S12, ESI†). **2**, on the other hand, exhibits a dramatic broadening of the Phe side chain peaks and the shifted peak is barely visible at a ratio of 2:1 of **2** to CB[8] (see Fig. S13, ESI†); this different behaviour is attributed to a fast exchange rate of the ternary complex.

A final confirmation of 2:1 homoternary complexation between the pentapeptide sequences and CB[8] was obtained by isothermal titration calorimetry (ITC). The ITC data clearly indicate that all three pentapeptide sequences exhibit 2:1 binding stoichiometry with CB[8]. The binding isotherm obtained for **3** is shown in Fig. 4. The 2:1 binding stoichiometry was first checked by fitting the data to a one site binding model; the data points were, then, fitted with a sequential binding site model using the Origin-Microcal software as reported previously.[7] The binding isotherms of **1** and **2** are reported in the ESI† (see Fig. S15 and S16). K_{a1} and K_{a2} are about 10^5 and 10^4 M^{-1}, respectively, for both **1** and **2**. **3** shows higher values of *ca.* 10^7 and 10^5 M^{-1}, respectively. The binding parameters obtained for the homoternary complexes are all reported in Table 1.

It should be noted that the three sequences exhibit very different characteristics: sequence **1** (AEFRH) has a negatively charged (Glu) and a positively charged (Arg) residue flanking the Phe, sequence **2** (LVFIA) is completely hydrophobic and presents a challenging steric environment around Phe, and sequence **3** (VIFAE) possesses a negatively charged (Glu) residue at its C-terminus.

The difference in binding abilities reflects the different environments of the Phe–CB[8] complexes in the three sequences. Furthermore, the data obtained imply that the pentapeptide sequences present different self-assembly in solution, which thus cause a different strength of the interaction with CB[8]. **1** shows preferential binding between Phe and CB[8], but fluorescence
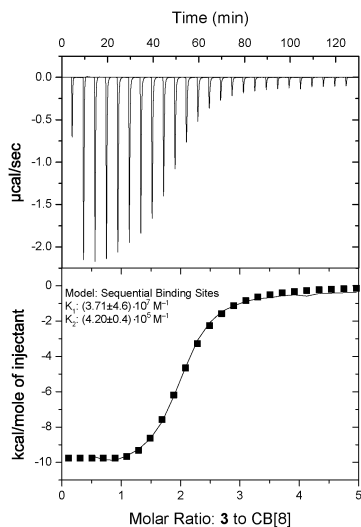

Fig. 2 Relative emission at 284 nm or 303 nm *vs.* the CB[8]/peptide ratio (peptide sequences 40 μM, CB[8] 0–35 μM).

Fig. 4 ITC titration of sequence **3** into a CB[8] solution.

Table 1 ITC data for CB[8] ternary complexes of **1**–**3**

	K_a (M^{-2})	K_{a1}/K_{a2}	ΔH (kcal mol^{-1})	$T\Delta S$ (kcal mol^{-1})
1	$(3.74 \pm 0.7) \times 10^{10}$	15.3	-15.3 ± 0.2	-0.88 ± 0.3
2	$(7.68 \pm 2.3) \times 10^{9}$	2.04	-9.02 ± 0.6	-4.49 ± 0.9
3	$(1.58 \pm 0.3) \times 10^{13}$	88.3	-20.8 ± 0.2	-2.76 ± 0.3

and ^{1}H-NMR data suggest an additional secondary interaction with His. **2** is highly prone to fibre formation and likely undergoes a conformational adjustment prior to binding CB[8]. The breaking of the strong π–π interactions that are already present within the peptide bundle causes quenching of the fluorescence at 303 nm. Moreover, the significantly higher binding constants measured for **3** are very likely related to the circular conformation that it adopts in water and its exposed Phe residue. The cationic N-terminus and the anionic Glu side chain at the C-terminus are held in close proximity pushing the Phe side chain far from the backbone and, therefore, allowing better interaction with CB[8]. This hypothesis is also supported by computational energy minimisation of **3** (see ESI†).

In conclusion, we have reported three pentapeptide sequences, differing in charge, hydrophobicity and steric hindrance, all of which form strong homoternary complexes with CB[8]. A binding stoichiometry of 2:1 Phe–CB[8] was confirmed by three independent methods. Nevertheless, the sequences present different binding features strictly related to their amino acid composition and conformation in solution, confirming that guest environment plays a significant role in CB[8] binding. This study suggests that CB[8] can bind Phe (and likely most if not all of the aromatic residues) not only located at the N-terminus, but also in many places along a peptide sequence. Homoternary complexation can be very useful for further Phe-rich peptide dimerisation studies, such as Aβ that is currently under investigation in our laboratory and other pathologically related peptides.

References

1 B. K. Kim, H. Kang, K. O. Doh, S. H. Lee, J. W. Park, S. J. Lee and T. J. Lee, *Bioorg. Med. Chem. Lett.*, 2012, **22**, 5415–5418.
2 S. Aggarwal, P. Singh, O. Topaloglu, J. T. Isaacs and S. R. Denmeade, *Cancer Res.*, 2006, **66**, 9171–9177.
3 P. Rajgariah and A. Urbach, *J. Inclusion Phenom. Macrocyclic. Chem.*, 2008, **62**, 251–254.
4 D. T. Dang, H. D. Nguyen, M. Merkx and L. Brunsveld, *Angew. Chem., Int. Ed.*, 2013, **52**, 2915–2919.
5 H. D. Nguyen, D. T. Dang, J. L. J. van Dongen and L. Brunsveld, *Angew. Chem., Int. Ed.*, 2010, **49**, 895–898.
6 S. W. Heo, T. S. Choi, K. M. Park, Y. H. Ko, S. B. Kim, K. Kim and H. I. Kim, *Anal. Chem.*, 2011, **83**, 7916–7923.
7 J. M. Chinai, A. B. Taylor, L. M. Ryno, N. D. Hargreaves, C. A. Morris, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2011, **133**, 8810–8813.
8 J. W. Lee, S. Samal, N. Selvapalam, H.-J. Kim and K. Kim, *Acc. Chem. Res.*, 2003, **36**, 621–630.
9 L. Isaacs, *Chem. Commun.*, 2009, 619–629.
10 E. Masson, X. Ling, R. Joseph, L. Kyeremeh-Mensah and X. Lu, *RSC Adv.*, 2012, **2**, 1213–1247.
11 L. M. Heitmann, A. B. Taylor, P. J. Hart and A. R. Urbach, *J. Am. Chem. Soc.*, 2006, **128**, 12574–12581.
12 F. Biedermann, U. Rauwald, J. M. Zayed and O. A. Scherman, *Chem. Sci.*, 2011, **2**, 279–286.
13 O. D. Monera, T. J. Sereda, N. E. Zhou, C. M. Kay and R. S. Hodges, *J. Pept. Sci.*, 1995, **1**, 319–329.
14 H. Du, R. C. A. Fuh, J. Li, L. A. Corkan and J. S. Lindsey, *Photochem. Photobiol.*, 1998, **68**, 141–142.
15 M. J. Krysmann, V. Castelletto, A. Kelarakis, I. W. Hamley, R. A. Hule and D. J. Pochan, *Biochemistry*, 2008, **47**, 4597–4605.
16 K. E. Marshall, K. L. Morris, D. Charlton, N. O'Reilly, L. Lewis, H. Walden and L. C. Serpell, *Biochemistry*, 2011, **50**, 2061–2071.