Host–Guest Induced Peptide Folding with Sequence-Specific Structural Chirality

David E. Clarke, † Guanglu Wu, † Ce Wu, and Oren A. Scherman*

ABSTRACT: Controlling the spatial and temporal behavior of peptide segments is essential in the fabrication of functional peptide-based materials and nanostructures. To achieve a desired structure, complex sequence design is often required, coupled with the inclusion of unnatural amino acids or synthetic modifications. Herein, we investigate the structural properties of 1:1 inclusion complexes between specific oligopeptides and cucurbit[8]uril (CB[8]), inducing the formation of turns, and by alteration of the peptide sequence, tunable structural chirality. We also explore extended peptide sequence binding with CB[8], demonstrating a simple approach to construct a peptide hairpin.

Control over the spatial and temporal behavior of peptide segments is of great importance for precisely engineering hierarchical protein structures,1,2 inhibiting peptide and protein aggregation,3,4 and developing stimuli-responsive materials.5 Peptide architectures can be altered significantly by covalent modifications or through noncovalent molecular recognition.6,7 These approaches typically require precise sequence design, along with the inclusion of unnatural amino acids to achieve the desired structural function.

Cucurbit[n]urils (CB[n]), n = 5–8, 10) are a family of synthetic macrocycles that can bind specific amino acids with high affinity and selectivity.8–10 Specifically, the large cavity of CB[8] is versatile in accommodating peptide side chains and can adopt various configurations. In addition to 1:1 complexation, CB[8] can encapsulate one residue with an auxiliary guest to form a 1:1 heteroternary motif11 or bind two identical residues to form a 2:1 homoternary complex.12,13 Namely, the motif containing two Phe-Gly-Gly has been applied to generate protein dimers or supramolecular polymers via binding with CB[8].14,15

CB[8] binding motifs can be classified through their binding stoichiometry coupled with their corresponding enthalpy perturbations.16 Urbach and co-workers have reported 1:1 binding between CB[8] and specific tripeptide sequences such as Tyr-Leu-Ala (YLA) and Met-Leu-Ala (MLA), which exhibit abnormally large enthalpy changes similar to those found for ternary complexes.17,18 Through investigating their size-based diffusion properties, we elucidated that rather than forming arbitrary n:n aggregates, this unique complex contained exactly one CB[8] and one peptide.19 This configuration, two adjacent side chains of the peptide simultaneously reside inside the cavity, resembling a ternary motif but with the exception that the encapsulated moieties are connected via a covalent amide bond.20–22 This binding motif exhibits a host-mediated adjacent side-chain pairing (HASP), which we have termed a “HASP knot”, representing a pseudostatic complexation tecton.

The wider applications of HASP knots are relatively unexplored, with previous studies focusing on the thermodynamic aspects of complexation.24–26 The main findings can be summarized as follows: (1) an adjacent residue pair from any combination of Tyr (Y), Leu (L), Lys (K), Phe (F), Met (M), and Arg (R) is inclined to form a HASP knot;24,25 (2) the binding affinity is more susceptible to the position of the pair (K_N-terminus > 10K_C-terminus) than its sequence order (K_{FL} ≈ K_{KL}); (3) although only two residues are necessary, an additional third residue or further extension in the sequence can stabilize the pair-inclusion configuration, leading to quantitative formation of HASP knots with a substantial suppression on other competitive pathways.26 Formation of a HASP knot requires the peptide backbone to adopt a constrained conformation, whereby the amino acid side chains rotate around the amide bond so they are located on the same side plane of the peptide. This distortion angle, referred to as the angle of rotation between the two side chains needs to approach 0° (Figure S1). In a survey of approximately 200 dipeptide crystal structures found in the CCDC database, <20% were smaller than 90°, and only 10% were within 10°, with these predominantly existing in aromatic–aromatic species such as FF.27,28 Therefore, it is conformationally unfavorable for a dipeptide to adopt such an acute distortion angle.

Herein, we investigated the hypothesis whether HASP knots could be a feasible approach to precisely manipulate peptide conformation. We explore their structural properties focusing on dipeptide isomers (constitutional- and stereo-), leading us to the discovery of both their sequence-specific structural
chirality and the amplification of their circular dichroism (CD) signals. In addition, through the design and site-specific CB[8] labeling of extended peptide sequences, we demonstrate the formation of a peptide hairpin via Förster resonance energy transfer (FRET).

Through utilizing CD spectroscopy, we identified that the formation of HASP knots can induce a significant conformational change. As shown in Figure 1b, YLA (all peptides have a C-terminal amide, unless stated otherwise) and YAL exhibit similar CD signatures with a negative minimum at 195 nm and a smaller positive in reflection at 225 nm, suggesting typical random coil conformations. When complexed with CB[8] at a 1:1 ratio, the CD spectra of YLA@CB[8] (orange trace) displayed a new positive maxima at around 200 nm and a large increase in signal intensity and the in reflection at 225 nm is no longer pronounced. Conversely, YAL@CB[8] which does not form a HASP knot,24 has a signature similar to its uncomplexed species resembling a random coil. These significant transformations of YLA@CB[8]’s CD spectra suggest the generation of a new secondary structure arising from the formation of a HASP knot. The same spectral behavior was also observed for YMA, YKA, YRA, and FLA, which all readily form HASP knots with CB[8] (Figure 1b). This phenomenon is also found in dipeptide sequences such as YL@CB[8] (Figure 1c, blue trace), although slight spectral differences exist when compared to YLA@CB[8]. Therefore, to avoid unnecessary complexity introduced by Ala, which is positioned outside the CB[8] cavity, we opted to concentrate on dipeptide sequences for subsequent chirality studies.

In line with previous studies on tripeptides,24,25 isothermal titration calorimetry and NMR experiments revealed that Y residues coupled with an aliphatic amino acid (YL/LY, YM/MY, YK/KY) or an additional Y residue (YY) were all found to favor the formation of HASP knots (Figures S2–S9). The CD spectra of these dipeptides display similar CD signatures, with minima around 190 nm and positive inflections at 200 and 225 nm (Figure 1c). When complexed with CB[8] at a 1:1

Figure 1. (a) Proposed conformational change of the YL dipeptide when forming a 1:1 HASP knot with CB[8]. CD spectra of tripeptides (b) and dipeptides (c) and their (1:1) CB[8] complexes.

Figure 2. (a) “Clockwise” turns are generated through Leu rotation in YL@CB[8] (top) or via d-Leu in YLd@CB[8] (bottom). (b) “Counterclockwise” turns are formed through Leu rotation in LY@CB[8] (top) or via D-Tyr in YdL@CB[8] (bottom). (c) CD spectra of Tyr and Leu containing dipeptides and their CB[8] complexes, highlighting CW (top) or CCW (bottom) turn formation.
ratio, the aromatic–aliphatic dipeptide (YM/YK@CB[8]) CD spectra both undergo similar transformations to YL@CB[8]. Interestingly, the aliphatic–aromatic dipeptides (LY/MY/KY@CB[8]) displayed different transformations to their aromatic–aliphatic counterparts, exhibiting an inverse negative minima around 195–200 nm (Figure 1c). A chiral inversion induced through swapping the sequence order of adjacent residues is highly unique and offers an approach to manipulate peptide folding in a controllable manner.

CD signals of peptide sequences in the near UV are typically generated by $\pi \rightarrow \pi^*$ (190–200 nm) and $n \rightarrow \pi$ (215–230 nm) transitions of the amide bond. Therefore, the intense peaks observed around 195–200 nm in the HASP knots originate from an electronic transition from NH to a carbonyl bond ($\pi \rightarrow \pi^*$) and resemble spectra of $\beta$-turns. The acute angle required for the dipeptide to fold into the CB[8] cavity, it suggests that this signal arises from an induced $\pi \rightarrow \pi^*$ transition present in the complex’s constrained conformation (Figure 1a). This is further supported by the CD spectra of YL at different ratios of CB[8], where both the evolution and saturation of the positive maxima can be visualized (Figure S10).

CD spectra of the aromatic–aromatic dipeptides (FF and YY) already display a small increase in signal intensity with S10. This is further supported by the hyperchromism displayed for these complexes, where the induced turn permits a $\pi \rightarrow \pi^*$ transition which can be visualized in their CD spectra (Figures 2c and S14b). These isomeric complexes suggest YY has a $\pi \rightarrow \pi^*$ transition generated in opposite directions and this can be visualized in the 200 nm region corresponding to a complex containing three CB[8]s (Figure S21). Following HASP knot formation with CB[8], NOE correlations could be recognized between the two adjacent side chains, which suggests rotation around the amide bond and the formation of a turn-like structure (Figure S21).

To apply our findings, we investigated whether HASP knots could be applied to extended peptide sequences as an elementary operation to manipulate their spatial structure. We rationalized the sequence YLAGFLAGGAL, which has three binding sites: FL (midchain), YL (N-terminus), and LY (C-terminus) (Figure S23). These three HAPs were coupled with neighboring Ala residues, which have been reported to stabilize pair inclusion complexes. In the presence of excess CB[8] (1:3.6 peptide:CB[8]), the NMR spectrum of YLAGFLAGGAL is almost identical to the superposition of the spectra of YLA@CB[8], AFLA@CB[8], and ALY@CB[8] (Figure S23). Diffusion ordered spectroscopy (DOSY) NMR measurements also displayed a diffusion value corresponding to a complex containing three CB[8]s (Figure S24 and Table S2), confirming that HASP knots can be generated independently at multiple sites and are not restricted to the peptide’s termini.

Heterochiral sequences of YL containing both L- and D-amino acids were also investigated. The side chains of YL and YLd already reside on the same side of the amide bond but are in opposite directions with respect to each other (Figure 2a and b). This was verified by the mirrored CD signatures obtained for YLd and YL (Figures 2c and S14a). These unnatural heterodipeptides inherently have an acute angle between side chains as they reside on the same side of the amide bond. With this in mind, HASP knot formation with CB[8] resulted in a similar spatial conformation being retained through complexation (Figure 2a and b). However, since YL and YLd are chiral isomers, the CB[8] induced turns are generated in opposite directions and this can be visualized in their CD spectra (Figures 2c and S14b). These isomeric complexes suggest YYL@CB[8]/YLd@CB[8] (CW) and LY@CB[8]/YL@CB[8] (CCW) adopt similar structures. Therefore, in both the YL@CB[8] and LY@CB[8] complexes, we can suggest the L side chain rotates toward the Y residue during complexation to form turns that are opposite in direction (Figure 2a and b).

We validated our model by performing NOESY measurements on the YLd and YL dipeptides. When uncomplexed, YLd has no detectable correlation between side chains, confirming they reside on opposite sides of the amide bond (Figure S20). Following HASP knot formation with CB[8], NOE correlations could be recognized between the two adjacent side chains, which suggests rotation around the amide bond and the formation of a turn-like structure (Figure S21). In contrast, different behavior is observed for YL, where NOE correlations between side chains are present both before and after complexation (Figures S16 and S17). This provides further evidence for our hypothesis, where the side chains of YLd remain on the same side of the amide bond and following HASP knot formation, adopt a turn conformation. These CD and NMR studies demonstrate that simple changes in the ordering of a peptide sequence can be utilized to control a localized chiral environment and the corresponding direction of the induced turn.

To apply our findings, we investigated whether HASP knots could be applied to extended peptide sequences as an elementary operation to manipulate their spatial structure. We rationally designed the sequence YLAGFLAGGALY, which has three binding sites: FL (midchain), YL (N-terminus), and LY (C-terminus) (Figure S23). These three HAPs were coupled with neighboring Ala residues, which have been reported to stabilize pair inclusion complexes. In the presence of excess CB[8] (1:3.6 peptide:CB[8]), the NMR spectrum of YLAGFLAGGALY is almost identical to the superposition of the spectra of YLA@CB[8], AFLA@CB[8], and ALY@CB[8] (Figure S23). Diffusion ordered spectroscopy (DOSY) NMR measurements also displayed a diffusion value corresponding to a complex containing three CB[8]s (Figure S24 and Table S2), confirming that HASP knots can be generated independently at multiple sites and are not restricted to the peptide’s termini.

To demonstrate that a hairpin structure can be induced by the formation of a HASP knot, we designed a sequence WGGAYLAGG-Dansyl (h-peptide) containing a FRET pair of Trp (W) and a Dansyl fluorophore positioned at opposite termini, coupled with a HASP (AYLA-) midchain binding site. NMR of the h-peptide at different ratios of CB[8] displayed broad peaks and indistinguishable binding modes until an excess of CB[8] was reached (Figures S26). At a ratio
of 1:6 (h-peptide:CB[8]), 1H and DOSY NMR showed that a HASP knot formed at the -AYLA- site accompanying 1:1 CB[8] binding of the W and Dansyl residues (Figures S27, S28 and Table S2). Considering that both Dansyl and W are each independently encapsulated by CB[8] macrocycles, only long-range energy transfer through space (e.g., FRET) is permitted between these two chromophores.

Steady-state emission spectra were recorded for the h-peptide sequence before and after the addition of CB[8] (Figure 3b). Without CB[8], two emission bands were observed at 360 and 560 nm (λex 225 nm) corresponding to the photoluminescent emissions of W and Dansyl, respectively. In the presence of excess CB[8] (1:6 h-peptide:CB[8]), the emission band of W diminishes by 70% and is coupled with a 2.3 fold increase in Dansyl emission (Figure 3b). This observed FRET indicates the close proximity of the N- and C- termini and confirms the formation of a hairpin structure induced by a HASP knot.

Through investigating the structural chirality of dipeptides and their CB[8]-mediated adjacent side-chain pairing, i.e., HASP knot, we discovered the formation of an induced turn-like structure. Interestingly, by altering the ordering of the peptide sequence or the inclusion of unnatural amino acids, the direction of the turn can be manipulated. This sequence-specific structural chirality could be incredibly powerful in tuning localized chiral environments for asymmetric catalysis. We also explored how HASP motifs can be applied to extended peptide sequences as a simple operation to influence their wider structure. By combining missequence binding and the ability to induce a turn, we utilized a peptide sequence containing a FRET pair to confirm that CB[8] binding can induce the formation of a peptide hairpin. This work provides a simple approach to manipulate peptide segments, where mutations containing HASP motifs can be introduced into intrinsically disordered regions of proteins to provide structural function. Moreover, there is potential for this approach to be explored in the design and fabrication of chiral supramolecular complexes and other engineered nanostructures.

ASSOCIATED CONTENT

* Supporting Information
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c00342.

Materials and methods; synthesis and characterization of peptide@CB[8] complexes; isothermal titration thermograms and thermodynamic data of CB[8]-mediated complexes; circular dichroism spectra of peptide@CB[8] complexes; 1H NMR titrations of peptides into CB[8]; NOESY 2D NMR of YL and Y before and after addition of CB[8]; DOSY 2D NMR of YLAGGALY@CB[8] and h-peptide@CB[8] complexes; UV/vis of the h-peptide@CB[8] complex (PDF)

AUTHOR INFORMATION

Corresponding Author
Oren A. Scherman – Melville Laboratory for Polymer Synthesis, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom; orcid.org/0000-0001-8032-7166; Email: oas23@cam.ac.uk

Authors
David E. Clarke – Melville Laboratory for Polymer Synthesis, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom
Guanglu Wu – Melville Laboratory for Polymer Synthesis, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom
Ce Wu – Melville Laboratory for Polymer Synthesis, Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, United Kingdom

Complete contact information is available at: https://pubs.acs.org/10.1021/jacs.1c00342

Author Contributions
D.E.C. and G.W. contributed equally to this work.

Notes
The authors declare no competing financial interest.
ACKNOWLEDGMENTS

This work was supported by an EPSRC Programme Grant (No. CH, EP/L027151/1), the Leverhulme Trust (project: "Natural material innovation for sustainable living") and the Marie Curie FP7 ITN (SASSYPOL, 607602). The authors thank Dr Peter Grice and Mr Duncan Howe for the helpful discussions surrounding NMR analysis.

REFERENCES

(1) Sasaki, E.; Böhrringer, D.; van de Waterbeemd, M.; Leibundgut, M.; Zschoche, R.; Heck, A. J. R.; Ban, N.; Hilvert, D. Nat. Commun. 2017, 8, 14663.
(2) Pyles, H.; Zhang, S.; de Yoreo, J. J.; Baker, D. Nature 2019, 571, 251–256.
(3) Miranda, E.; Nordgren, I. K.; Male, A. L.; Lawrence, C. E.; Hoakwie, F.; Cuda, F.; Court, W.; Fox, K. R.; Townsend, P. A.; Packham, G. K.; Eccles, S. A.; Tavassoli, A. J. Am. Chem. Soc. 2013, 135, 10418–10425.
(4) Satav, T.; Korevaar, P.; de Greef, T. F. A.; Huskens, J.; Jonkheijm, P. Chem. - Eur. J. 2016, 22, 12675–12679.
(5) Kumar, M.; Ing, N. L.; Narang, V.; Wijerathne, N. K.; Hooiveld, A. I.; Ulijn, R. V. Nat. Chem. 2018, 10, 596–703.
(6) Clarke, D. E.; Parmenter, C. D. J.; Scherman, O. A. Angew. Chem., Int. Ed. 2018, 57, 7709–7713.
(7) Clarke, D. E.; Olesińska, M.; Mönch, T.; Schoenaers, B.; Stesmans, A.; Scherman, O. A. Chem. Commun. 2019, 55, 7354–7357.
(8) Mendive-Tapia, L.; Preciado, S.; García, J.; Ramón, R.; Kielland, N.; Albericio, F.; Lavilla, R. Nat. Commun. 2015, 6, 7160.
(9) Pappas, C. G.; Shafi, R.; Sasselli, I. R.; Siccardi, H.; Wang, T.; Narang, V.; Abzalimov, R.; Wijerathne, N.; Ulijn, R. V. Nat. Nanotechnol. 2016, 11, 960–967.
(10) Yang, D.; Li, W.; Qi, J.; Luo, S.-W.; Wu, Y.-D. J. Am. Chem. Soc. 2003, 125, 13018–13019.
(11) Maynard, S. J.; Almeida, A. M.; Yoshimi, Y.; Goldman, S. H. J. Am. Chem. Soc. 2014, 136, 16683–16688.
(12) Hosseinzaehdeh, P.; Bhardwaj, D.; Mulligan, V. K.; Shortridge, M. D.; Craven, T. W.; Pardo-Avila, F.; Rettie, S. A.; Kim, D. E.; Silva, D.-A.; Ibrahim, Y. M.; Webb, I. K.; Cort, J. R.; Adkins, J. N.; Varani, G.; Baker, D. Science 2017, 358, 1461–1466.
(13) Lee, J. W.; Samal, S.; Selvapalam, N.; Kim, H.-J.; Kim, K. Acc. Chem. Res. 2003, 36, 621–630.
(14) Laguna, J.; Mukhopadhyay, P.; Chakrabarti, S.; Isaacs, L. Angew. Chem., Int. Ed. 2005, 44, 4844–4870.
(15) Barrow, S. J.; Kasera, S.; Rowland, M. J.; del Barrio, J.; Scherman, O. A. Chem. Rev. 2015, 115, 12320–12406.
(16) Urbach, A. R.; Ramalingam, V. Isr. J. Chem. 2011, 51, 664–678.
(17) Kim, H.-J.; Heo, J.; Jeon, W. S.; Lee, E.; Kim, J.; Sakamoto, S.; Yamaguchi, K.; Kim, K. Angew. Chem., Int. Ed. 2001, 40, 1526–1529.
(18) Bush, M. E.; Bouley, N. D.; Urbach, A. R. J. Am. Chem. Soc. 2005, 127, 14511–14517.
(19) Heitmann, L. M.; Taylor, A. B.; Hart, P. J.; Urbach, A. R. J. Am. Chem. Soc. 2006, 128, 12574–12581.
(20) Cavaottta, E.; Jonkheijm, P.; Huskens, J. Chem. - Eur. J. 2017, 23, 4046–4050.
(21) Nguyen, H.; Dang, D.; van Dongen, J. L. J.; Brunsveld, L. Angew. Chem., Int. Ed. 2010, 49, 895–898.
(22) Tan, X.; Yang, L.; Liu, Y.; Huang, Z.; Yang, H.; Wang, Z.; Zhang, X. Polym. Chem. 2013, 4, 5378–5381.
(23) Wu, G.; Olesińska, M.; Wu, Y.; Matak-Vinkovic, D.; Scherman, O. A. J. Am. Chem. Soc. 2017, 139, 3202–3208.
(24) Smith, L. C.; Leach, D. G.; Blaylock, B. E.; Ali, O. A.; Urbach, A. R. J. Am. Chem. Soc. 2015, 137, 3663–3669.
(25) Hirani, Z.; Taylor, H. F.; Babcock, E. F.; Bockus, A. T.; Varnado, C. D.; Bielawski, C. W.; Urbach, A. R. J. Am. Chem. Soc. 2018, 140, 22623–22629.
(26) Wu, G.; Clarke, D. E.; Wu, C.; Scherman, O. A. Org. Biomol. Chem. 2019, 17, 3514–3520.
(27) Görbitz, C. H. Chem. - Eur. J. 2001, 7, 5153–5159.
(28) Mason, T. O.; Chirgadze, D. Y.; Levin, A.; Adler-Abramovich, L.; Gazit, E.; Knowles, T. P. J.; Buell, A. K. ACS Nano 2014, 8, 1243–1253.
(29) Greenfield, N. J. Anal. Biochem. 1996, 235, 1–10.
(30) Woody, R. Theory of Circular Dichroism of Proteins; Springer, Boston, MA, 1996, pp 25–67.
(31) Bush, C. A.; Sarkar, S. K.; Kopple, K. D. Biochemistry 1978, 17, 4951–4954.
(32) Migliore, M.; Bonvicini, A.; Tognetti, V.; Guillaume, L.; Baaden, M.; Oulyadi, H.; Joubert, L.; Ségala-Milazzo, I. Phys. Chem. Chem. Phys. 2020, 22, 1611–1623.
(33) Handelman, A.; Kuritz, N.; Natan, A.; Rosenman, G. Langmuir 2016, 32, 2847–2862.
(34) Wu, G.; Szabó, I.; Rosta, E.; Scherman, O. A. Chem. Commun. 2019, 55, 13227–13230.