

Self-Assembled Protein–Aromatic Foldamer Complexes with 2:3 and 2:2:1 Stoichiometries

Michal Jewginski,^{†,||} Thierry Granier,^{*,†} Béatrice Langlois d'Estaintot,[†] Lucile Fischer,[†] Cameron D. Mackereth,[§] and Ivan Huc^{*,†} 

[†]CBMN (UMR5248), Univ. Bordeaux, CNRS, IPB, Institut Européen de Chimie et Biologie, 2 rue Robert Escarpit, 33600 Pessac, France

^{||}Department of Organic and Pharmaceutical Technology, Faculty of Chemistry, Wrocław University of Science and Technology, 50-370 Wrocław, Poland

[§]ARNA (U 1212), Univ. Bordeaux, INSERM, Institut Européen de Chimie et Biologie, 2 rue Robert Escarpit, 33600 Pessac, France

Supporting Information

ABSTRACT: The promotion of protein dimerization using the aggregation properties of a protein ligand was explored and shown to produce complexes with unusual stoichiometries. Helical foldamer **2** was synthesized and bound to human carbonic anhydrase (HCA) using a nanomolar active site ligand. Crystal structures show that the hydrophobicity of **2** and interactions of its side chains lead to the formation of an HCA₂-**2**₃ complex in which three helices of **2** are stacked, two of them being linked to an HCA molecule. The middle foldamer in the stack can be replaced by alternate sequences **3** or **5**. Solution studies by CD and NMR confirm left-handedness of the helical foldamers as well as HCA dimerization.

The recognition of protein surface features by smaller molecules is a means to mask and inhibit interactions in which proteins may engage, in particular with other proteins. A possible effect of a surface-binding ligand is thus a disruption of protein assemblies, an outcome that can serve therapeutic¹ or pharmacological² purposes. Conversely, inducing homo- or heteromeric protein assembly can be useful when activity results from the association of several proteins. This is generally achieved by presenting multitopic protein surface ligands that can simultaneously bind two identical or two different proteins,^{3,4} and eventually enhance activity.^{4f,g} Multitopic ligands can also serve in the fabrication of ordered protein arrays.⁵ An alternate, essentially unexplored approach, to induce protein assembly into discrete objects would be to use monotopic ligands that themselves possess self-assembly properties. Along this line, we have recently reported the characterization of a 2+2 complex between two human carbonic anhydrase (HCA) molecules and two molecules of helical aromatic foldamer **1** (Figure 1). This complex is stabilized by protein–foldamer, foldamer–foldamer and protein–protein interactions.⁶ We now have extended this exploratory work and report the intriguing formation of HCA–aromatic foldamer complexes comprised of two proteins and either three identical foldamers (2:3 stoichiometry), or two identical and a third, different, foldamer (2:2:1 stoichiometry).

Several groups have been investigating aromatic foldamers⁷ as rigid helical or linear scaffolds that can be equipped with proteinogenic side chains to recognize protein,^{6,8} nucleic acid,^{9,10} or saccharide¹¹ surfaces. For example, **1** is a stable helical oligoamide of 8-amino-2-quinolinecarboxylic acid.¹² It was identified and characterized while studying interactions between HCA and related foldamer sequences bearing varied side chains protruding from the helix in position 4 of each quinoline unit.⁶ These foldamers possess an appended aryl-sulfonamide HCA nanomolar ligand of the enzyme active site that confines them at the protein surface. Specific interactions with the protein may then be detected by the induction of a handedness bias in the achiral helix backbone that otherwise exists as a racemic mixture of interconverting right-handed (*P*) and left-handed (*M*) enantiomers. In the case of HCA-**1**, circular dichroism (CD) at pH 7.4 in a phosphate buffer shows a positive band in the quinoline absorption region characteristic of *P* helicity (Figure 1c). This observation triggered structural investigations both in the solid state^{6a} (Figure 1b) and in solution^{6b} that eventually showed that HCA-**1** is not a 1:1 but an unprecedented 2:2 complex in which the hydrophobic nature of the aromatic helix cross section promotes foldamer–foldamer interactions that ultimately result in dimerization. Among over 400 HCA structures found in the protein databank (PDB), the packing and intercomplex protein–protein contacts of the (HCA-**1**)₂ structure are unique.

Compound **2** was prepared (see SI) using solid phase synthesis¹³ within a series of variants of **1** to test the scope of the formation of (HCA-**1**)₂. In the sequence of **2**, two monomers of **1** have been swapped. HCA-**2** also shows a positive CD band indicative of *P* helicity at pH 7.4 in phosphate buffer (Figure 1c). However, the HCA-**2** complex failed to crystallize from the Zn²⁺ rich medium in which crystals of (HCA-**1**)₂ grow. Other crystallization conditions were eventually found using commercial sparse-matrix screens and a structure of HCA-**2** was resolved at 1.8 Å resolution and revealed several distinct features (Figure 2, Figure S7).

The structure shows a higher aggregate containing two HCA molecules and not two but three molecules of **2** arranged in a

Received: January 6, 2017

Published: February 7, 2017

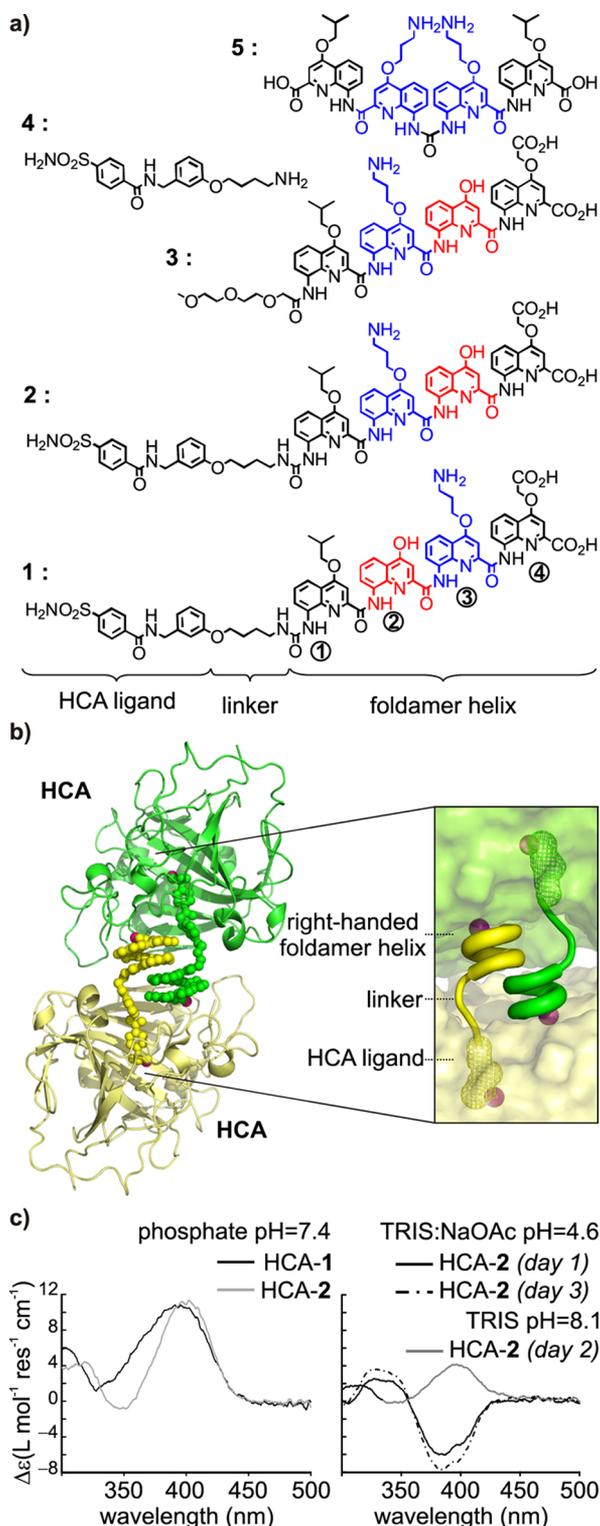


Figure 1. Foldamer–HCA interactions. (a) Formulas of HCA ligand 4 and foldamers 1, 2, 3 and 5. (b) Crystal structure (left) of the (HCA-1)₂ dimeric complex previously reported^{6a} and schematic view (right) of the foldamer–protein interface in that complex. Proteins are shown as ribbons (left) or as solvent-accessible surfaces (right). Foldamers are shown in the same color (yellow or green) as the color of the protein to which their ligand moiety is bound. Zinc ions at the bottom of the ligand binding pocket and at the foldamer–protein interface are shown as purple spheres. (c) CD spectra of HCA-1 and HCA-2 complexes in different buffers.

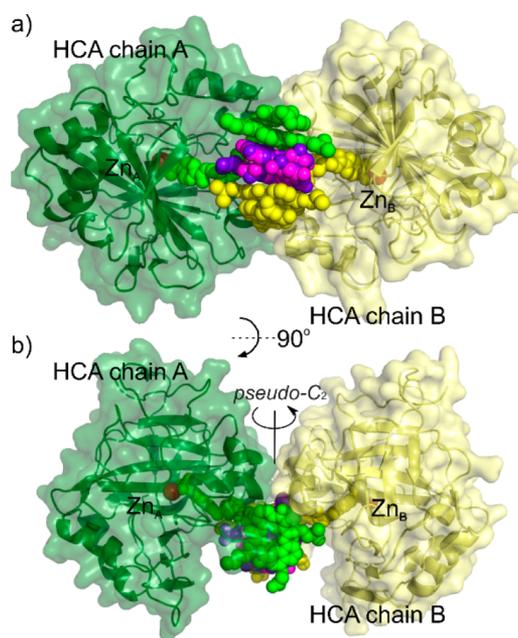


Figure 2. Crystal structure of the HCA₂₋₂₃ complex (PDB# 5L30). (a) Asymmetric unit showing two HCA molecules in green and yellow and three foldamer helices in CPK representation. The green and yellow foldamers have their ligand bound to an HCA active site. Native Zn²⁺ ions are shown as red spheres. A third foldamer (purple and magenta for the two orientations) is sandwiched between the first two. (b) View down the foldamers' helix axis of the complex showed in panel a.

stack, again mediated by contacts between the aromatic helices' cross sections. The two peripheral helices of the stack have their ligand bound to an HCA active site while the ligand of the central helix is not bound and actually not visible in the electron density map due to disordered positions. The complex has an overall noncrystallographic pseudo-C₂ symmetry. Consequently, the central foldamer is in fact present in two degenerate orientations within the stack and was modeled as such. Other remarkable differences exist between the (HCA-1)₂ and HCA₂₋₂₃ structures: (i) the axis of the stack of helices is parallel to the proteins' surfaces in the latter, whereas it is perpendicular in the former. The foldamer side chains thus point directly toward the protein surfaces in HCA₂₋₂₃ and establish remarkable contacts (Figure S7, see also the similar HCA₂₋₂₂₋₃ complex in Figure 3); (ii) these contacts include interactions with N-terminal His3 which (along with Ser2) is visible in the electron density map whereas it is usually not seen in HCA crystal structures; (iii) the relative orientation of the two HCA molecules is completely different; (iv) foldamers are *P*-helical in (HCA-1)₂ and *M*-helical in HCA₂₋₂₃. This latter feature first appeared to contradict the sign of the CD band observed in phosphate buffer. However, it was then realized that the crystallization medium was in fact quite different (see SI). A few controls established that *M*-helicity is favored not by salt or the nature of the buffer but by a lower pH and that it can indeed be observed in solution (Figure 1c). *M* and *P* helicities can be induced alternatively by changing pH (Figure S16). In contrast, the CD band of the (HCA-1)₂ complex is still positive at pH 4.6. The low pH must not only disfavor interactions with the *P* helix as this would result in no CD signal at all but also favor interactions with the *M* helix. In the HCA₂₋₂₂₋₃ complex, a salt bridge is indeed observed between the terminal foldamer

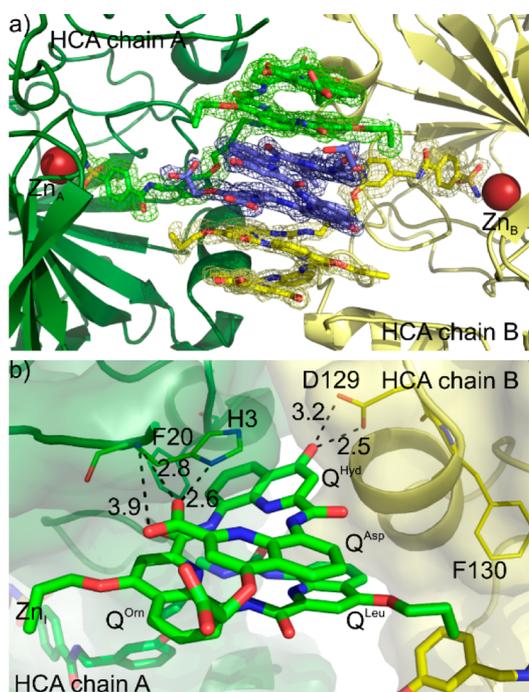


Figure 3. Crystal structure of the HCA₂-2-3 complex at 1.6 Å. (PDB# 5L6K). (a) Foldamer backbones contoured by 2mFo-DFc density maps at 1σ level, only one orientation of the central foldamer is shown. (b) Some relevant contacts of a foldamer helix of compound 2 and neighboring protein surfaces (distances in Å). The imidazole of H3 stacks on a quinoline ring; the imidazole and amide NH groups hydrogen bond to the terminal quinoline carboxylate. F20 (in the back) engages in edge-to-face aromatic contacts with a quinoline.

carboxylate function and His3. Whether this interaction is critical to *M* helicity remains to be demonstrated.

We sought to confirm the assembly features of the HCA-2 complex in solution by using NMR spectroscopy as established for HCA-1 (Figure 4, Figures S13–15).^{6b} At pH = 4.6, ¹⁵N HSQC spectra of [¹⁵N]HCA and [¹⁵N]HCA-4 demonstrated that the protein is stable, well-folded and fully bound by 4 (Figure 4a, note the shift of the Trp208 amide cross-peak located near to the sulfonamide moiety). Intermolecular contacts were then identified by comparing ¹H–¹⁵N HSQC spectra of 2 or 4 bound to [¹⁵N]HCA (Figure 4a,c). Compound 4 lacks a foldamer helix and chemical shift perturbations (CSPs) in the spectrum of HCA-2 as compared to HCA-4 can thus be attributed mainly to foldamer–protein contacts. Quite remarkably, CSPs were principally located at residues involved in protein–foldamer and protein–protein contacts of the HCA₂-2-3 crystal structure (Figure 4e). Size estimate of the HCA-2 complex by measuring ¹H^N T₂ of a 500 μM sample revealed a correlation time (τ_c) of 31.3 ns (Table S2), consistent with a dimeric state (two proteins) but the possible presence of a third foldamer could not be established.^{6b} At 100 μM, dissociation results in a significantly reduced correlation time of 23.2 ns. In the alternate phosphate condition at pH 7.4, ¹H–¹⁵N HSQC spectra again confirm protein folding and complex formation (Figure 4b). In contrast to pH 4.6, the comparison of HCA-4 and HCA-2 spectra at pH 7.4 show limited CSPs suggesting a weaker protein–foldamer interaction. Size estimates by ¹H^N T₂ indicate a mainly monomeric state (Table S2) in which helix handedness induction nevertheless takes place (Figure 1c) as for HCA-1.^{6b}

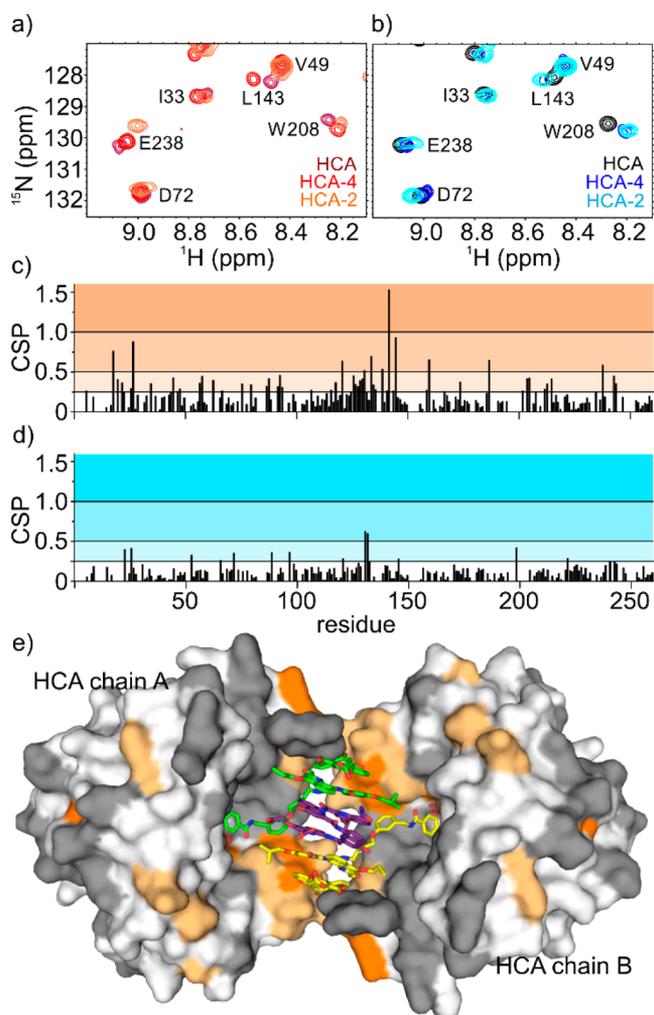


Figure 4. Intermolecular contacts identified by NMR spectroscopy in phosphate buffer. Part of ¹H–¹⁵N HSQC spectra of [¹⁵N]HCA (500 μM), either free (purple or black), or in the presence of 4 (1.5 equiv, red or blue), or of 2 (1.5 equiv, orange or cyan) at pH = 4.6 (a) and pH = 7.5 (b), respectively. CSPs of HCA-2 compared to HCA-4 calculated as a root-mean-square deviation ($\sqrt{((\Delta\delta_H)/0.14)^2 + (\Delta\delta_N)^2}$)^{0.5} at pH = 4.6 (c) and pH = 7.5 (d). (e) Protein surface of the HCA₂-2-3 crystal structure colored as in panel c. Residues for which NMR assignment is unclear are shown in gray.

Because the middle foldamer in the stack of helices does not need to possess an HCA ligand and does not engage in the same interactions with HCA as the peripheral helices, we explored the possibility to replace it by other foldamer sequences. Compound 3 is an analogue of 2 that lacks the HCA ligand. When HCA was set to crystallize in the presence of 2 (1 equiv) and 3 (0.5 equiv) a structure similar to that of HCA₂-2-3 was obtained (Figure 3, Figure S9), yet with a better resolution (1.6 Å). LC-MS analysis after dissolving the crystals ascertained the presence of 3 in the structure, presumably in the middle of the helical stacks (Figure S12). However, given the similarity between 2 and 3, and the absence of their flexible chains in the electron density maps, they cannot be distinguished in the structure. Next, compound 5 was designed which possesses a C₂ symmetrical helical structure and contains twice the two N-terminal residues of 2 and 3. HCA crystals were grown in the presence of 2 (1 equiv) and 5 (0.5 equiv) and a structure at 1.4 Å was obtained (Figure S10, PDB#

SLVS). The structure unambiguously shows the presence of 5 in the middle of the helical stack. Because of its symmetry, its two degenerate positions are identical. The electron density map indicates that a small amount of 1 also occupies the middle position but this could not be modeled accurately (Figure S11). Overall, although the contribution of the middle foldamer to the overall stability of the multimeric assembly is not established, the fact that it can be substituted ascertains the existence of a driving force for that site not to remain unoccupied.

In summary, we have established that linking helical aromatic foldamers to a protein surface leads to the formation of protein complexes with intriguing stoichiometries. Simply swapping two foldamer residues causes subtle changes in foldamer–protein surface interactions, resulting in a major change in assembly behavior at different pH values. The stacking of hydrophobic helical cross sections constitutes a recurrent pattern and presumably a driving force that adds to foldamer–protein and protein–protein interactions. An alternate way to control protein association may be to exploit deliberately the self-assembly properties of aromatic amide helices that show a propensity to form not stacks, but multistranded triple or double helices, including in water.¹⁴

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.7b00184.

Experimental details for synthetic procedure, spectroscopic and crystallographic data (PDF)

Model for 5L3O (PDB)

Model for 5L6K (PDB)

Model for 5LVS (PDB)

■ AUTHOR INFORMATION

Corresponding Authors

*t.granier@cbmn.u-bordeaux.fr

*i.huc@iecb.u-bordeaux.fr

ORCID

Ivan Huc: 0000-0001-7036-9696

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the ESRF and SOLEIL for provision of synchrotron radiation facilities, and the staffs of beamlines ID23-2, ID30A-3 and Proxima-2 for their kind assistance. We also thank the structural biology platform at the Institut Européen de Chimie et Biologie (UMS 3033) for access to NMR spectrometers and technical assistance. Financial support from the IR-RMN-THC Fr3050 CNRS is gratefully acknowledged. M. J. was supported by the Polish Ministry of Science and Higher Education (Mobility Plus Program).

■ REFERENCES

(1) (a) Wells, J. A.; McClendon, C. L. *Nature* **2007**, *450*, 1001. (b) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. *Nat. Chem.* **2013**, *5*, 161. (c) Jayatunga, M. K. P.; Thompson, S.; Hamilton, A. D. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 717. (d) Gopalakrishnan, R.; Frolov, A. I.; Knerr, L.; Drury, W. J., III; Valeur, E. *J. Med. Chem.* **2016**, *59*, 9599.

(2) Rollins, C. T.; Rivera, V. M.; Woolfson, D. N.; Keenan, T.; Hatada, M.; Adams, S. E.; Andrade, L. J.; Yaeger, D.; van Schravendijk, M. R.; Holt, D. A.; Gilman, M.; Clackson, T. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 7096.

(3) (a) Rutkowska, A.; Schultz, C. *Angew. Chem., Int. Ed.* **2012**, *51*, 8166. (b) Fegan, A.; White, B.; Carlson, J. C. T.; Wagner, C. R. *Chem. Rev.* **2010**, *110*, 3315. (c) Luo, Q.; Hou, C.; Bai, Y.; Wang, R.; Liu, J. *Chem. Rev.* **2016**, *116*, 13571. (d) Milroy, L.-G.; Grossmann, T. N.; Hennig, S.; Brunsveld, L.; Ottmann, C. *Chem. Rev.* **2014**, *114*, 4695.

(4) (a) Heitmann, L. M.; Taylor, A. B.; Hart, P. J.; Urbach, A. R. *J. Am. Chem. Soc.* **2006**, *128*, 12574. (b) Barnard, A.; Miles, J. A.; Burslem, G. M.; Barker, A. M.; Wilson, A. J. *Org. Biomol. Chem.* **2015**, *13*, 258. (c) Erhart, D.; Zimmermann, M.; Jacques, O.; Wittwer, M. B.; Ernst, B.; Constable, E.; Zvelebil, M.; Beaufils, F.; Wymann, M. P. *Chem. Biol.* **2013**, *20*, 549. (d) Kopytek, S. J.; Standaert, R. F.; Dyer, J. C. D.; Hu, J. C. *Chem. Biol.* **2000**, *7*, 313. (e) Spencer, D. M.; Wandless, T. J.; Schreiber, S. L.; Crabtree, G. R. *Science* **1993**, *262*, 1019. (f) Bosmans, R. P. G.; Briels, J. M.; Milroy, L.-G.; de Greef, T. F. A.; Merckx, M.; Brunsveld, L. *Angew. Chem., Int. Ed.* **2016**, *55*, 8899. (g) Dang, D. T.; Nguyen, H. D.; Merckx, M.; Brunsveld, L. *Angew. Chem., Int. Ed.* **2013**, *52*, 2915.

(5) (a) McGovern, R. E.; McCarthy, A. A.; Crowley, P. B. *Chem. Commun.* **2014**, *50*, 10412. (b) Li, Q.; So, C. R.; Fegan, A.; Cody, V.; Sarikaya, M.; Vallera, D. A.; Wagner, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 17247. (c) Hou, C.; Li, J.; Zhao, L.; Zhang, W.; Luo, Q.; Dong, Z.; Xu, J.; Liu, J. *Angew. Chem., Int. Ed.* **2013**, *52*, 5590.

(6) (a) Buratto, J.; Colombo, C.; Stupfel, M.; Dawson, S. J.; Dolain, C.; Langlois d'Estaintot, B.; Fischer, L.; Granier, T.; Laguerre, M.; Gallois, B.; Huc, I. *Angew. Chem., Int. Ed.* **2014**, *53*, 883. (b) Jewgiński, M.; Fischer, L.; Colombo, C.; Huc, I.; Mackereth, C. D. *ChemBioChem* **2016**, *17*, 727.

(7) Zhang, D.-W.; Zhao, X.; Hou, J.-L.; Li, Z.-T. *Chem. Rev.* **2012**, *112*, 5271.

(8) (a) Ernst, J. T.; Becerril, J.; Park, H. S.; Yin, H.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2003**, *42*, 535. (b) Barnard, A.; Long, K.; Martin, H. L.; Miles, J. A.; Edwards, T. A.; Tomlinson, D. C.; Macdonald, A.; Wilson, A. J. *Angew. Chem., Int. Ed.* **2015**, *54*, 2960. (c) Azzarito, V.; Miles, J. A.; Fisher, J.; Edwards, T. A.; Warriner, S. L.; Wilson, A. J. *Chem. Sci.* **2015**, *6*, 2434. (d) Saraogi, I.; Hebda, J. A.; Becerril, J.; Estroff, L. A.; Miranker, A. D.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **2010**, *49*, 736. (e) Kumar, S.; Birol, M.; Schlamadinger, D. E.; Wojcik, S. P.; Rhoades, E.; Miranker, A. D. *Nat. Commun.* **2016**, *7*, 11412. (f) Kumar, S.; Birol, M.; Miranker, A. D. *Chem. Commun.* **2016**, *52*, 6391.

(9) (a) Dervan, P. B.; Edelson, B. S. *Curr. Opin. Struct. Biol.* **2003**, *13*, 284. (b) Bando, T.; Sugiyama, H. *Acc. Chem. Res.* **2006**, *39*, 935.

(10) (a) Müller, S.; Laxmi-Reddy, K.; Jena, P. V.; Baptiste, B.; Dong, Z.; Godde, F.; Ha, T.; Rodriguez, R.; Balasubramanian, S.; Huc, I. *ChemBioChem* **2014**, *15*, 2563. (b) Mandal, P. K.; Baptiste, B.; Langlois d'Estaintot, B.; Kauffmann, B.; Huc, I. *ChemBioChem* **2016**, *17*, 1911. (c) Delaurière, L.; Dong, Z.; Laxmi-Reddy, K.; Godde, F.; Toulmé, J.-J.; Huc, I. *Angew. Chem., Int. Ed.* **2012**, *51*, 473.

(11) Choi, S.; Clements, D. J.; Pophristic, V.; Ivanov, I.; Vemparala, S.; Bennett, J. S.; Klein, M. L.; Winkler, J. D.; DeGrado, W. F. *Angew. Chem., Int. Ed.* **2005**, *44*, 6685.

(12) (a) Jiang, H.; Léger, J.-M.; Huc, I. *J. Am. Chem. Soc.* **2003**, *125*, 3448. (b) Qi, T.; Maurizot, V.; Noguchi, H.; Charoenraks, T.; Kauffmann, B.; Takafuji, M.; Ihara, H.; Huc, I. *Chem. Commun.* **2012**, *48*, 6337.

(13) (a) Baptiste, B.; Douat-Casassus, C.; Laxmi-Reddy, K.; Godde, F.; Huc, I. *J. Org. Chem.* **2010**, *75*, 7175. (b) Dawson, S. J.; Hu, X.; Claerhout, S.; Huc, I. *Methods Enzymol.* **2016**, *580*, 279.

(14) (a) Ferrand, Y.; Kendhale, A. M.; Garric, J.; Kauffmann, B.; Huc, I. *Angew. Chem., Int. Ed.* **2010**, *49*, 1778. (b) Shang, J.; Gan, Q.; Dawson, S. J.; Rosu, F.; Jiang, H.; Ferrand, Y.; Huc, I. *Org. Lett.* **2014**, *16*, 4992.