Structure Elucidation of Helical Aromatic Foldamer–Protein Complexes with Large Contact Surface Areas

Post Sai Reddy, Béatrice Langlois d'Estaintot, Thierry Granier, Cameron Mackereth, Lucile Fischer, Ivan Huc

Structure Elucidation of Helical Aromatic Foldamer–Protein Complexes with Large Contact Surface Areas

Post Sai Reddy,[a] B8atrice Langlois d'Estaintot,[a] Thierry Granier,[a] Cameron D. Mackereth,[b] Lucile Fischer,[a] and Ivan Huc*[a, c]

Abstract: The development of large synthetic ligands could be useful to target the sizeable surface areas involved in protein–protein interactions. Herein, we present long helical aromatic oligoamide foldamers bearing proteinogenic side chains that cover up to 450 a^2 of the human carbonic anhydrase II (HCA) surface. The foldamers are composed of aminoquinolinecarboxylic acids bearing proteinogenic side chains and of more flexible aminomethyl-pyridinecarboxylic acids that enhance helix handedness dynamics. Crystal structures of HCA-foldamer complexes were obtained with a 9- and a 14-mer both showing extensive protein–foldamer hydrophobic contacts. In addition, foldamer–foldamer interactions seem to be prevalent in the crystal packing, leading to the peculiar formation of an HCA superhelix wound around a rod of stacked foldamers. Solution studies confirm the positioning of the foldamer at the protein surface as well as a dimerization of the complexes.

Aromatic foldamers[1] emerge as a new class of folded oligomers that may be decorated with proteinogenic side chains to interact with proteins[2,3] and nucleic acids,[4,5] and eventually serve as inhibitors of nucleic acid–protein and protein–protein interactions. Amphipathic structures have also been shown to interact with, or to insert themselves in, membranes.[6,7] Some possess antibiotic activity.[6] Both linear[2,5,6] and helical[3,4,7] fol- damers have been developed and varied targets have been identified, including *h*DM2 and B-cell lymphoma-2 (Bcl-2) regu- lator proteins,[2a,d,e] protein precursors of amyloids,[2c,3a–e] G-quadruplex DNA[4] and some DNA-binding enzymes.[3f] Advantages of aromatic foldamers include their ease of synthesis, for example through solid-phase methodologies,[8] and the predictability and stability of their folded conformations in both protic and aprotic solvents.[9] Because relatively large and well-defined folded objects can be produced using aromatic amide backbones,[10] it may be envisaged to cover large surface areas of proteins and nucleic acids. For example, we recently reported protein binding using a 9.2 kDa foldamer mimicking a 16 base–pair DNA duplex.[3f] Nevertheless, designing objects that can recognize large surface areas of proteins is difficult: which side chains are to be selected and where should they be located? Some of the published work concerned mimetics of *a*-helices,[2] B-DNA,[3f] or natural products.[5] Other approaches use screening through directed evolution methods.[11] It remains that no general approach exists for the ab initio design of large ligands for a protein surface. Structural information about aromatic foldamer–protein interactions would constitute a firm stepping-stone for further design, but it can hardly be obtained without having reasonable binding affinity in the first place.

To overcome this sort of deadlock, we endeavored to investigate foldamer–protein interfaces by confining foldamers at the surface of a protein.[12,13] For example, helical oligoamides based on 8-aminoquinoline carboxylic acid Q were maintained in close proximity to the surface of human carbonic anhydrase (HCA) as a model system by means of a nanomolar HCA ligand (Figure 1 a). Interactions were first detected through induction of foldamer helix handedness in response to contacts with chiral elements at the protein surface, and subjected to structural investigations both in the solid state and in solution.[13]

Crystal structures proved their usefulness in that they revealed multiple features that could not have been designed in the first place, including unusual foldamer–protein complex stoichiometries,[13c] and now constitute starting points for iterative improvements. They also relate to crystal structures of complexes between proteins and other medium-sized aromatic ligands such as calixarenes,[14] suramine,[15] 1,3,6,8-pyrenetetra- sulfonic acid,[16] or molecular tweezers.[17]

However, these earlier studies only concerned short (tetrameric) foldamer segments and that could not cover very large protein surface areas. To extend this approach to longer sequences, we recently demonstrated that incorporating more flexible P units (Figure 1 b) into Q_n sequences enhance helix dynamics and allow for protein-mediated handedness induction in helical foldamers such as nonamer 3 and tetradecamer 5.[18]

[a] Dr. P. S. Reddy, Dr. B. Langlois d'Estaintot, Dr. T. Granier, Dr. L. Fischer, Prof. I. Huc
CBMN (UMR5248), Univ. Bordeaux–CNRS–INP
Institut Europ8en de Chimie et Biologie
2 rue Escarpit, 33600 Pessac (France)

[b] Dr. C. D. Mackereth
ARNA (U1212), Univ. Bordeaux–INSERM–CNRS
Institut Europ8en de Chimie et Biologie
2 rue Escarpit, 33600 Pessac (France)

[c] Prof. I. Huc
Department Pharmazie and Center for Integrated Protein Science
Ludwig-Maximilians-Universit-t
Butenandtstr. 5–13, 81377 Menchen (Germany)
E-mail: ivan.huc@cup.lmu.de

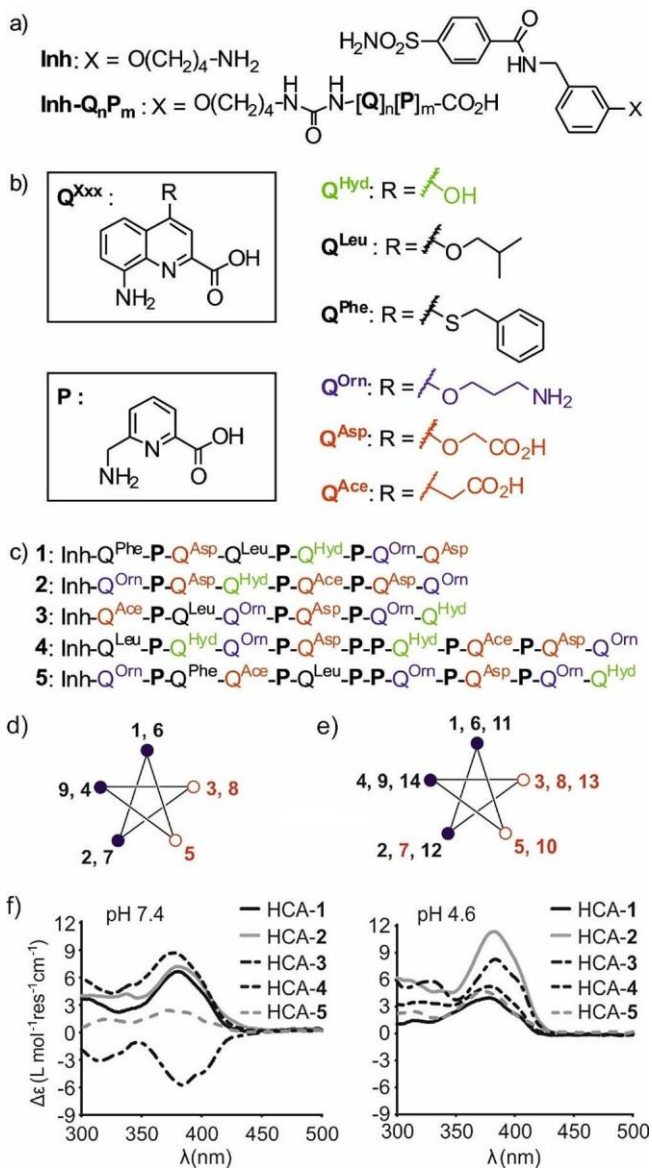

Figure 1. a) Formula of functionalized HCA ligands. b) Formula of amino acid units color coded according to their side chain: hydrophobic (black), polar neutral (green), cationic (blue), anionic (red). c) Foldamer sequences 1–5. Five-pointed star representations of amphipathic hybrid Q/P foldamers 1–3 (d) and 4–5 (e). Monomers are counted from the C-terminus. Numbers in red indicate the location of P units. f) CD spectra of HCA–foldamer complexes after an 8 day equilibration at two different pH conditions in phosphate buffer at 25 8C.

In the following, we introduce additional sequences 1, 2 and 4 and report the structure elucidation of complexes HCA-2 and HCA-4. The structures again reveal an ensemble of hard-to-predict features, including extended shape complementarity between the cylinder-like helices and a shallow groove at the protein surface as well as multiple hydrophobic contacts with the face of the foldamer that was initially not intended to interact with the protein. The extensiveness of foldamer–protein and foldamer–foldamer contacts is also obvious in the crystal packing. These results thus pave the way to the ab initio design of large foldamer-based ligands of protein surfaces.

Sequences 1, 2 and 4 were designed following the same principles as for 3 and 5: 1) each has a benzenesulfonamide HCA ligand at its N-terminus; 2) side chains of Q units were selected with no other prejudice than to generate some foldamer surface diversity, that is, with hydrophobic, polar neutral and charged groups; 3) P units aim at enhancing helix dynamics and were positioned on one face of the foldamer helix to allow for interactions between the other face and the protein (Figure 1 d,e); and 4) sequences contain no stereogenic center and initially fold as a racemic mixture of right- and left-handed helices, but this equilibrium may be biased by foldamer–protein interactions. Oligomer solid-phase synthesis[8] and characterization are reported in the Supporting Information. Crude products typically have 75–80 % purity. After reversed-phase HPLC purification, yields from initial Wang resin loadings range from 37 to 66 %.

We used CD in the absorption region of quinoline chromophores at 360 nm to detect helix handedness induction in presence of HCA (Figure 1 f and Figure S4, Supporting Information). At equilibrium, all five foldamers showed a CD response[19] with slight variations depending on sequence. As previously observed with shorter sequences,[13c] one foldamer (3) showed an inversion of CD sign, and thus of preferred helix handedness, suggesting the involvement of charged residues in the interaction: depending on pH, foldamer–protein interactions vary which may result in favoring the *P* or the *M* helix handedness. Given these encouraging results, crystallization of HCA–foldamer complexes was attempted for all compounds but 5, which gives the weakest CD. The structures of HCA-2 and HCA-4 could be solved and refined at 2.7 and 2.9 a resolution, respectively (Figure 2 and Figure S5, Supporting Information).

Despite the different lengths and side chain composition of 2 and 4, their complexes with HCA share multiple features: the HCA ligands are well located in the HCA active site with the sulfonamide coordinated to Zn^{II}; canonical helical conformations are retained, even when two consecutive flexible P units are present as in the case of 4; the helices are right-handed, in agreement with their positive CD bands;[20] the helices lie down on the protein with the helix axis parallel to the protein surface and cover a sizeable area, fulfilling our main objective. Specifically, the foldamer–protein contact area (i.e. the interface per component, protein or foldamer) at the exclusion of the HCA ligand was measured with PDBePISA[21] to be 308 and 448 a^2 in HCA-2 and HCA-4, respectively. In both complexes, the foldamer helix is located in a wide and shallow groove of the protein exposing its P units to HCA and its side chains to the solvent. Side chains thus do not contribute to direct interactions with the protein to which the foldamer ligand is bound, consistent with the lack of effect of pH on helix handedness induction. This arrangement appears to be driven by shape complementarity—the groove surface is smooth and so is the face of the helix where P units are located—and by hydrophobic effects. Indeed, the groove is lined with hydrophobic residues: Phe20, Pro21, Ile22, Val134, Pro201 and Leu203 towards which the helices exclusively expose aryl CH groups of P and Q units. One may infer from these structures that the

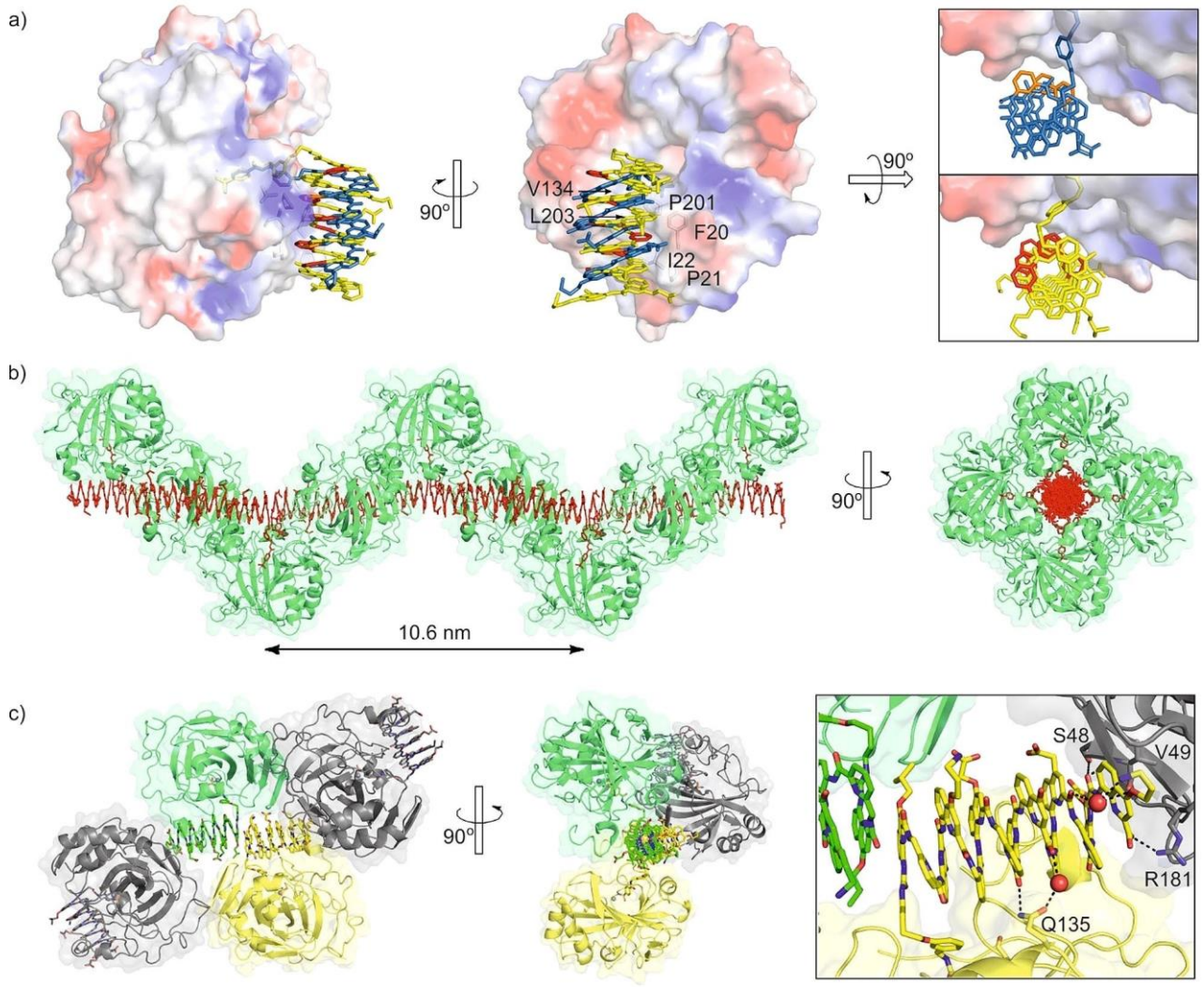

Figure 2. Structures in the solid state of HCA-4 (PDB# 6Q9T) and HCA-2_2 (PDB# 6HZX). a) Side view (left), top view (middle) and front view (right) of the two complexes. In the side and top views, the structures are overlaid. HCA is shown as an isosurface color labelled according to electrostatic potential (blue: cationic, red: anionic, white: neutral). Hydrophobic surface residues are indicated. The foldamers are shown in tube representation in yellow (Q units) and red (P units) for 4, and in blue (Q units) and orange (P units) for 2. The top and front views show the horizontal and angular shifts of the foldamer positions. b) Solid-state packing of HCA-2_2 showing the helical arrangement of HCA molecules around a columnar stack of foldamers along a 4-fold screw axis. Every other foldamer in the stack is not bound to an HCA molecule. c) Solid-state packing of HCA-4 showing stacked dimers of foldamers surrounded by four HCA molecules. In the top view, one protein structure has been removed for clarity. At side, a zoom of the foldamer–foldamer contact and of the surrounding protein residues are shown. In b and c), foldamers are shown as space-filling models and the proteins as a ribbon representation within a transparent iso-surface. In c each foldamer and the protein bound to it have the same color.

HCA-3 complex is different and involves some quinoline side chains responsible for its pH dependence.

The relative positions of 2 and 4 in their HCA complexes also show some differences. They are shifted horizontally with respect to the protein surface and axially, that is, around the helix axis (Figure 2 a). This positioning appears to be influenced by some directional interactions, for example, a hydrogen bond between the primary amide of Gln135 and two main chain carbonyl groups of 4, either directly, or through a water bridge (Figure S7b, Supporting Information). It may also be slightly influenced by crystal packing (see below). Overall, the large protein–foldamer contacts, the presentation of the foldamer aromatic edges—which can be functionalized—to the protein, and the simple stoichiometry (one foldamer per protein), make HCA-2 and HCA-4 much better starting points for structure-based design than earlier 2:2 and 2:3 complexes.[12,13]

Packing in HCA-2 and HCA-4 crystals differ much from each other and also from over 400 reported HCA structures. Nevertheless, as with shorter foldamers,[13] foldamer-foldamer interactions appear to be a strong driving force in both cases, extending the concept of "molecular glue" proposed for protein–calixarene complexes.[14b] The structure of HCA-2 is actually a structure of HCA-2_2 in which a second foldamer is included with its HCA ligand not bound to a protein.[22] This second foldamer has few contacts with surrounding proteins (Figure S6, Supporting Information) but it inserts itself in continuous foldamer columns in alternation with the foldamer bound to HCA (Figure 2 b). Extensive head-to-head and tail-to-tail contacts

thus occur alternatively between the aromatic helix cross-sections. The $P4_3$ symmetry of the crystal then results in a unique left-handed helical arrangement of HCA molecules around foldamer columns through a 4-fold screw axis. The HCA-4 complexes also involve stacks of foldamers but these are limited to dimers which are all surrounded by four HCA molecules in the $P2_12_12$ lattice (Figure 2 c). Some foldamer side chains and one foldamer cross-section are involved in contacts with HCA molecules other than the one to which the ligand is bound (Figure 2 c, right, Figures S6 and S7, Supporting Information). The differences between the packing of HCA-2 and HCA-4 together with their similar foldamer–protein contacts suggest that the foldamers indeed influence packing but packing itself does not cause major differences in the foldamer–protein interactions.

We then sought for information about the structures of HCA-2 and HCA-4 in solution by using surface plasmon resonance (SPR) and NMR spectroscopy taking HCA-Inh as a reference, as established for HCA-short foldamer complexes[13b,c] (Figure 3 and Figures S8–S10, Supporting Information). We opted for an investigation at physiological pH to allow for comparison with earlier studies, rather than at the lower pH of the crystallization drops. The $^1H,^{15}N$ HSQC spectra of [^{15}N]HCA- Inh demonstrated that the protein is stable, well-folded and fully bound by the HCA ligand (Figure 3 a,b and Figure S8, Sup- porting Information). Intermolecular contacts were then identi- fied by comparing the $^1H,^{15}N$ HSQC spectra of [^{15}N]HCA-2 (Fig- ure S9) or [^{15}N]HCA-4 (Figure S10) with that of [^{15}N]HCA-Inh (Figure 3 a, b). Compound Inh lacks a foldamer helix and chem- ical shift perturbations (CSPs) observed in the spectra of [^{15}N]HCA-2 or [^{15}N]HCA-4 can thus be attributed mainly to fol- damer–protein contacts. We distinguished weak and strong CSPs, and measured HSQC signal broadening. Quite remarka- bly, signal broadening beyond detection and strong CSPs were principally located at residues involved in protein–foldamer and protein–protein contacts of the HCA-2_2 and HCA-4 crystal structures (Figure 3 e,f), suggesting a positioning of the foldam- ers in solution comparable to that in the solid state. The broadening beyond detection of some signals, a phenomenon known for calixarenes[14e] but not previously observed with shorter foldamers, was attributed to some dynamic phenom- ena, perhaps related with the mobility of the foldamer in the protein groove, as suggested by the slightly different positioning observed in the two crystal structures.

Measurements of the $^1H^N$ T2 delays allowed for an estimate of the correlation times (t_c) of the complexes in solution and thus to assess their size and thereby their aggregation state. For HCA-Inh, a correlation time of 19.5 ns was measured indicating a mainly monomeric state in solution. In contrast, HCA-2 and HCA-4 at a 200 *m*M concentration had t_c values of 36.2 and 35.1 ns, respectively (Table S2, Supporting Information), consistent with a dimeric state (two proteins and two foldamers).[13b,c] This aggregation might contribute to the CSPs observed at residues involved in some protein–protein contacts in the crystal lattice.

We also assessed the strength of interactions using SPR. Titration data could all be fitted to a 1:1 binding model. A K_d value of $4.2 \pm 10^{\pm 9}$ M$^{\pm 1}$ was found for HCA-2, which is very simi-

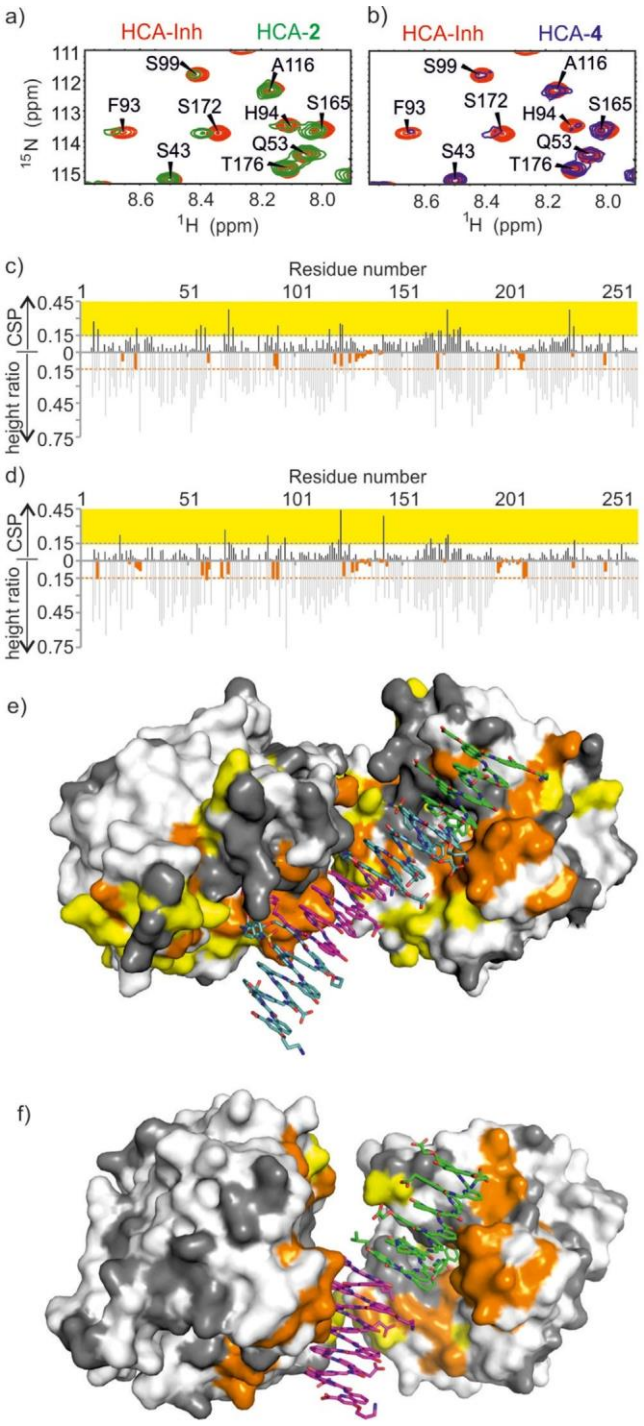

Figure 3. NMR chemical shift variations of [^{15}N]HCA (200 *m*M) in complex with Inh, 2 or 4 (1.3 equiv) in Tris buffer (10 mM, pH 8.0). Part of the superimposed $^1H,^{15}N$ HSQC spectra of: (a) HCA-2 and HCA-Inh; (b) HCA-4 and HCA-Inh. (c,d) CSP that is, chemical shift perturbations (Dd_{NH}) calculated as a root-mean-square deviation $(((DdH)/0.14)^2 + (DdN)^2)^{0.5}$ and height ratio calculated as a ratio of peak intensities. (c) HCA-2 compared to HCA-Inh; (d) HCA-4 compared to HCA-Inh. Residues marked in orange exhibit significant line-broadening in their (HCA-2 or HCA-4) HSQC signal with height ratio < 0.15. e) Protein surface of the HCA$_2$-2$_4$ crystal structure colored as in panel c. Residues for which NMR assignment is unclear are shown in gray. f) Protein surface of the HCA$_2$-4$_2$ crystal structure colored as in panel d. Residues with signal overlap or ambiguous NMR assignment are shown in gray.

lar to the K_d of HCA-Inh (5×10^{-9} M).[13a] However, association and dissociation were both about three times slower for HCA-2 ($k_{on} = 4.9 \times 10^{5}$ M^{-1}s^{-1}, $k_{off} = 2.0 \times 10^{-3}$ s^{-1}) than for HCA-Inh ($k_{on} = 1.5 \times 10^{6}$ M^{-1}s^{-1}, $k_{off} = 7.7 \times 10^{-3}$ s^{-1}), illustrating the involvement of the foldamer in the interactions. HCA-4 was found to be slightly less stable (K_d = 30x10^{-9} M^{-1}), as a consequence of a slightly slower association (k_{on} = 0.6x10^{5} M^{-1} s^{-1}) while dissociation remained as slow as for HCA-2 (k_{off} = 1.9x10^{-3} s^{-1}). Interpretation of these values must take into account that helix handedness inversion takes place only partially in the course of the SPR titration, meaning that the values average the binding of the *P* helix and of the less favored *M* helix.

In summary, we showed that the tethering approach has allowed for the identification of structurally defined foldamer–HCA complexes with large contact surface areas. Good foldamer–protein shape complementarity and hydrophobic contacts seem to be prevailing parameters within these complexes. Structure elucidation provides an accurate description of the protein–foldamer contact and a starting point to further design the foldamer–protein interaction by the introduction of tailored foldamer side chains. The ultimate objective is to observe tight and selective binding in the absence of a tether. Efforts in this direction are currently in progress in our laboratories and will be reported in due course.

Acknowledgements

This work was supported by the European Union (H2020-MSCA-IF-2016-751019-PROFOLIG, postdoctoral fellowship to P.S.R.). It benefited from the facilities and expertise of the Biophysical and Structural Chemistry platform at IECB, CNRS UMS3033, INSERM US001, Bordeaux University, France. We thank Ms. L. Minder for assistance with SPR measurements and Dr. M. Savko for data collection at beamline PROXIMA-2 (SOLEIL, proposal n820170745).

Conflict of interest

The authors declare no conflict of interest.

Keywords: aromatic oligoamides · foldamers · protein surface recognition · structure elucidation · X-ray crystallography

[1] a) D.-W. Zhang, X. Zhao, J.-L. Hou, Z.-T. Li, *Chem. Rev.* 2012, *112*, 5271–5316; b) I. Huc, *Eur. J. Org. Chem.* 2004, 17–29.

[2] Rodlike arylamides: a) J. T. Ernst, J. Becerril, H. S. Park, H. Yin, A. D. Hamilton, *Angew. Chem. Int. Ed.* 2003, *42*, 535–539; *Angew. Chem.* 2003, *115*, 553–557; b) I. Saraogi, J. A. Hebda, J. Becerril, L. A. Estroff, A. D. Miranker, A. D. Hamilton, *Angew. Chem. Int. Ed.* 2010, *49*, 736–739; *Angew. Chem.* 2010, *122*, 748–751; c) S. Kumar, A. D. Hamilton, *J. Am. Chem. Soc.* 2017, *139*, 5744–5755; d) S. Kumar, A. Henning-Knechtel, M. Magzoub, A. D. Hamilton, *J. Am. Chem. Soc.* 2018, *140*, 6562–6574; e) A. Barnard, K. Long, H. L. Martin, J. A. Miles, T. A. Edwards, D. C. Tomlinson, A. Macdonald, A. J. Wilson, *Angew. Chem. Int. Ed.* 2015, *54*, 2960–2965; *Angew. Chem.* 2015, *127*, 3003–3008; f) V. Azzarito, J. A. Miles, J. Fisher, T. A. Edwards, S. L. Warriner, A. J. Wilson, *Chem. Sci.* 2015, *6*, 2434–2443.

[3] Helical arylamides: a) S. Kumar, M. Birol, D. E. Schlamadinger, S. P. Wojcik, E. Rhoades, A. D. Miranker, *Nat. Commun.* 2016, *7*, 11412; b) S. Kumar, M. Birol, A. D. Miranker, *Chem. Commun.* 2016, *52*, 6391–6394; c) S. Kumar, M. A. Brown, A. Nath, A. D. Miranker, *Chem. Biol.* 2014, *21*, 775–781; d) S. Kumar, A. D. Miranker, *Chem. Commun.* 2013, *49*, 4749–4751; e) S. Kumar, A. Henning-Knechtel, I. Chehade, M. Magzoub, A. D. Hamilton, *J. Am. Chem. Soc.* 2017, *139*, 17098–17108; f) K. Ziach, C. Chollet, V. Parissi, P. Prabhakaran, M. Marchivie, V. Corvaglia, P. P. Bose, K. Laxmi-Reddy, F. Godde, J.-M. Schmitter, S. Chaignepain, P. Pourquier, I. Huc, *Nat. Chem.* 2018, *10*, 511–518.

[4] a) S. Meller, K. Laxmi-Reddy, P. V. Jena, B. Baptiste, Z. Dong, F. Godde, T. Ha, R. Rodriguez, S. Balasubramanian, I. Huc, *ChemBioChem* 2014, *15*, 2563–2570; b) P. K. Mandal, B. Baptiste, B. Langlois d'Estaintot, B. Kauffmann, I. Huc, *ChemBioChem* 2016, *17*, 1911–1914; c) L. DelauriHre, Z. Dong, K. Laxmi-Reddy, F. Godde, J.-J. Toulm8, I. Huc, *Angew. Chem. Int. Ed.* 2012, *51*, 473–477; *Angew. Chem.* 2012, *124*, 488–492.

[5] a) P. B. Dervan, B. S. Edelson, *Curr. Opin. Struct. Biol.* 2003, *13*, 284–299; b) T. Bando, H. Sugiyama, *Acc. Chem. Res.* 2006, *39*, 935 –944.

[6] a) G. N. Tew, R. W. Scott, M. L. Klein, W. F. DeGrado, *Acc. Chem. Res.* 2010, *43*, 30–39; b) R. W. Scott, W. F. DeGrado, G. N. Tew, *Curr. Opin. Biotechnol.* 2008, *19*, 620–627.

[7] a) C. Lang, W. Li, Z. Dong, X. Zhang, F. Yang, B. Yang, X. Deng, C. Zhang, J. Xu, J. Liu, *Angew. Chem. Int. Ed.* 2016, *55*, 9723–9727; *Angew. Chem.* 2016, *128*, 9875–9879; b) C. Lang, X. Deng, F. Yang, B. Yang, W. Wang, S. Qi, X. Zhang, C. Zhang, Z. Dong, J. Liu, *Angew. Chem. Int. Ed.* 2017, *56*, 12668–12671; *Angew. Chem.* 2017, *129*, 12842–12845; c) Y. Huo, H. Zeng, *Acc. Chem. Res.* 2016, *49*, 922–930; d) P. Xin, P. Zhu, P. Su, J.-L. Hou, Z.-T. Li, *J. Am. Chem. Soc.* 2014, *136*, 13078–13081; e) X. Wei, G. Zhang, Y. Shen, Y. Zhong, R. Liu, N. Yang, F. Y. Al-mkhaizim, M. A. Kline, L. He, M. Li, Z.-L. Lu, Z. Shao, B. Gong, *J. Am. Chem. Soc.* 2016, *138*, 2749–2754.

[8] a) N. S. Murphy, P. Prabhakaran, V. Azzarito, J. P. Plante, M. J. Hardie, C. A. Kilner, S. L. Warriner, A. J. Wilson, *Chem. Eur. J.* 2013, *19*, 5546–5550; b) X. Hu, S. J. Dawson, Y. Nagaoka, A. Tanatani, I. Huc, *J. Org. Chem.* 2016, *81*, 1137 – 1150.

[9] T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara, I. Huc, *Chem. Commun.* 2012, *48*, 6337–6339.

[10] a) N. Delsuc, J.-M. L8ger, S. Massip, I. Huc, *Angew. Chem. Int. Ed.* 2007, *46*, 214–217; *Angew. Chem.* 2007, *119*, 218–221; b) S. De, B. Chi, T. Granier, T. Qi, V. Maurizot, I. Huc, *Nat. Chem.* 2018, *10*, 51–57.

[11] a) Y. X. Wu, Y. J. Kwon, *Methods* 2016, *106*, 21– 28 ; b) D. Lipovsek, A. Pleckthun, *J. Immunol. Methods* 2004, *290*, 51– 67.

[12] a) M. Vallade, M. Jewginski, L. Fischer, J. Buratto, K. Bathany, J.-M. Schmitter, M. Stupfel, F. Godde, C. D. Mackereth, I. Huc, *Bioconjugate Chem.* 2019, *30*, 54– 62; b) A. Jain, S. G. Huang, G. M. Whitesides, *J. Am. Chem. Soc.* 1994, *116*, 5057 –5062; c) A. Jain, G. M. Whitesides, R. S. Alexander, D. W. Christianson, *J. Med. Chem.* 1994, *37*, 2100–2105; d) T. Andersson, M. Lundquist, G. T. Dolphin, K. Enander, B.-H. Jonsson, J. W. Nilsson, L. Baltzer, *Chem. Biol.* 2005, *12*, 1245 –1252 ; e) L. T. Tegler, K. Fromell, B.-H. Jonsson, J. Viljanen, C. Winander, J. Carlsson, L. Baltzer, *ChemBioChem* 2011, *12*, 559–566; f) K. Enander, G. T. Dolphin, L. Baltzer, *J. Am. Chem. Soc.* 2004, *126*, 4464–4465; g) A. L. Banerjee, M. Swanson, B. C. Roy, X. Jia, M. K. Haldar, S. Mallik, D. K. Srivastava, *J. Am. Chem. Soc.* 2004, *126*, 10875–10883; h) M. Cigler, T. G. Meller, D. Horn-Ghetko, M.-K. von Wrisberg, M. Fottner, R. S. Goody, K. Lang, *Angew. Chem. Int. Ed.* 2017, *56*, 15737–15741; *Angew. Chem.* 2017, *129*, 15943–15947; i) D. A. Keedy, Z. B. Hill, J. T. Biel, E. Kang, T. J. Rettenmaier, J. Brandao-Neto, N. M. Pearce, F. von Delft, J. A. Wells, J. S. Fraser, *eLife* 2018, *7*, e36307; j) J. Gao, J. A. Wells, *Chem. Biol. Drug Des.* 2012, *79*, 209–215; k) B. Yang, S. Tang, C. Ma, S.-T. Li, G.-C. Shao, B. Dang, L. Wang, *Nat. Commun.* 2017, *8*, 2240.

[13] a) J. Buratto, C. Colombo, M. Stupfel, S. J. Dawson, C. Dolain, B. Langlois d'Estaintot, L. Fischer, T. Granier, M. Laguerre, B. Gallois, I. Huc, *Angew. Chem. Int. Ed.* 2014, *53*, 883–887; *Angew. Chem.* 2014, *126*, 902–906; b) M. Jewginski, L. Fischer, C. Colombo, I. Huc, C. D. Mackereth, *ChemBioChem* 2016, *17*, 727–736; c) M. Jewginski, T. Granier, B. Langlois d'Estaintot, L. Fischer, C. D. Mackereth, I. Huc, *J. Am. Chem. Soc.* 2017, *139*, 2928–2931.

[14] a) R. E. McGovern, H. Fernandes, A. R. Khan, N. P. Power, P. B. Crowley, *Nat. Chem.* 2012, *4*, 527–533; b) R. E. McGovern, A. A. McCarthy, P. B. Crowley, *Chem. Commun.* 2014, *50*, 10412– 10415; c) R. E. McGovern,

B. D. Snarr, J. A. Lyons, J. McFarlane, A. L. Whiting, I. Paci, F. Hof, P. B. Crowley, *Chem. Sci.* 2015, *6*, 442 – 449; d) M. L. Rennie, G. C. Fox, J. P8rez, P. B. Crowley, *Angew. Chem. Int. Ed.* 2018, *57*, 13764 – 13769; *Angew. Chem.* 2018, *130*, 13960 – 13965; e) M. L. Rennie, A. M. Doolan, C. L. Raston, P. B. Crowley, *Angew. Chem. Int. Ed.* 2017, *56*, 5517 – 5521; *Angew Chem..* 2017, *129*, 5609 – 56013.

[15] a) L. M. T. R. Lima, C. F. Becker, G. M. Giesel, A. F. Marques, M. T. Cargnelutti, M. O. Neto, R. Q. Monteiro, H. Verli, I. Polikarpov, *Biochim. Biophys. Acta Proteins Proteomics* 2009, *1794*, 873 – 881; b) H. P. Morgan, I. W. McNae, M. W. Nowicki, W. Zhong, P. A. M. Michels, D. S. Auld, L. A. Fothergill-Gilmore, M. D. Walkinshaw, *J. Biol. Chem.* 2011, *286*, 31232 – 31240; c) E. A. Arag¼o, D. S. Vieira, L. Chioato, T. L. Ferreira, M. R. Lourenzoni, S. R. Silva, R. J. Ward, *Arch. Biochem. Biophys.* 2012, *519*, 17 – 22; d) C. Ren, K. Morohashi, A. N. Plotnikov, J. Jakoncic, S. G. Smith, J. Li, L. Zeng, Y. Rodriguez, V. Stojanoff, M. Walsh, M.-M. Zhou, *Chem. Biol.* 2015, *22*, 161 – 168.

[16] H. P. Morgan, I. W. McNae, K.-Y. Hsin, P. A. M. Michels, L. A. Fothergill-Gilmorea, M. D. Walkinshaw, *Acta Crystallogr. Sect. F* 2010, *66*, 215 – 218.

[17] D. Bier, S. Mittal, K. Bravo-Rodriguez, A. Sowislok, X. Guillory, J. Briels, C. Heid, M. Bartel, B. Wettig, L. Brunsveld, E. Sanchez-Garcia, T. Schrader, C. Ottmann, *J. Am. Chem. Soc.* 2017, *139*, 16256 – 16263.

[18] M. Vallade, P. S. Reddy, L. Fischer, I. Huc, *Eur. J. Org. Chem.* 2018, 5489 – 5498.

[19] This is significant in that these five sequences represent all those we tested, and not a selection out of large number of sequences.

[20] A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* 2011, *76*, 195 – 200.

[21] E. Krissinel, K. Henrick, *J. Mol. Biol.* 2007, *372*, 774 – 797.

[22] The asymmetric unit in fact contains two essentially superimposable HCA-2_2 complexes.