Designing cooperatively folded abiotic uni- and multimolecular helix bundles

Soumen De, Bo Chi, Thierry Granier, Ting Qi, Victor Maurizot, Ivan Huc

Designing cooperatively folded abiotic uni- and multimolecular helix bundles

Soumen De,[1] Bo Chi,[1,2] Thierry Granier,[1] Ting Qi,[1,3] Victor Maurizot,[1] Ivan Huc[1*]

[1] CBMN Laboratory, Univ. Bordeaux, CNRS, IPB, Institut Européen de Chimie et Biologie, 2 rue Escarpit, 33600 Pessac, France.

[2] Current address: State key Laboratory of Materials-Oriented Chemical Engineering, Nanjing Tech University, No. 30 south Puzu Road, Pukou District, Nanjing 211816, P. R. China

[3] Current address: School of Chemistry and Chemical Engineering, University of Chinese Academy of Sciences, Beijing 100049, P. R. China

[*] Correspondence to: i.huc@iecb.u-bordeaux.fr.

ABSTRACT

Abiotic foldamers, *i.e.* foldamers that have backbones chemically remote from peptidic and nucleotidic skeletons, may give access to shapes and functions different from those of peptides and nucleotides, just like peptides and nucleotides differ from each. However, design methodologies towards abiotic tertiary and quaternary structures are yet to be developed. Here we report rationally designed interactional patterns to guide the folding and assembly of abiotic helix-bundles. Computations allowed to introduce hydrogen bonding functionalities at defined locations on aromatic amide backbones that eventually promote cooperative folding into helix-turn-helix motifs in organic solvents. The hydrogen bond-directed aggregation of helices not linked by a turn unit produced several thermodynamically and kinetically stable homochiral dimeric and trimeric bundles having structures distinct from the designed helix-turn-helix. Relative helix orientation within the bundles may be changed from parallel to tilted upon subtle solvent variations. Altogether, these results prefigure the richness and uniqueness of abiotic tertiary structure behavior.

INTRODUCTION

In the last twenty years, chemists have shown that numerous non-natural foldamer backbones adopt well-defined conformations analogous to the secondary motifs of biopolymers, such as helices, sheets, and turns[1,2,3,4,5] This new paradigm holds great promise for molecular design. Indeed, biopolymers' efficient and sophisticated functions in terms of recognition, catalysis, information storage and transfer, or energy conversion, rest on their specific shapes obtained through folding of their primary sequence. Folding emerges as the most efficient strategy to organize large and complex arrays of chemical groups in space with atomic precision. In addition, molecules consisting of a linear sequence of building blocks possess the advantage of being inherently modular and amenable to optimization through iterative modifications, *e.g.* mutations, and deletions or additions of monomers or longer segments. In this context, abiotic foldamers that have backbones chemically remote from the natural peptidic and nucleotidic skeletons are particularly interesting because they may give access to shapes, and thus to functions, different from those of peptides and nucleotides, just like the shapes and functions of peptides and nucleotides differ from each other because of their distinct chemical compositions. For example, foldamers having aromatic moieties in their main chain constitute a distinct grouping that has been shown to form unusual structures such as pilars[6], knots[7,8], or helices having a variable diameter[9,10].

However, most of the functions of biopolymers, especially of proteins, emerge at the level of their tertiary structures and would not be achieved by an isolated α-helix or β-sheet. A major challenge thus resides in establishing design principles of artificial tertiary folds comprised of several secondary structural elements. Because of the highly cooperative nature of tertiary folds, protein design itself remains a very challenging endeavour despite great advances[11,12,13,14] including the development of dedicated computational tools[15]. Changing an α-amino acid sequence to optimize interactions *between* secondary motifs may also alter

interactions within them and thus decrease overall conformational stability. Nevertheless, α-helix bundling is a specific pattern that can largely be predicted, allowing the design of self-assembled quaternary bundles[12]**Erreur ! Signet non défini.**[,16,17,18,19] as well as novel tertiary structures[20]. Yet, the exploitation of non-natural monomers in artificial tertiary folds is limited. Several β-amino acids may be tolerated into known peptidic bundles or tertiary structures[21,22,23,24]. Sheets containing non-natural units have been reported as well[25,26]. As key milestones, self-assembled helix bundles of β-peptides[27,28] and of β-ureas[29] have been reported. Stitching helices together within these bundles may in principle yield an artificial tertiary fold but this remains to be implemented. To the best of our knowledge, no genuine tertiary structures are known that do not contain α-amino acids. Approaching examples simply consist in connecting several abiotic helices through short or rigid linkers[30,31,32], or non-covalently threading them onto rod-like templates[33]. Going beyond the world of peptides, tertiary folds based on abiotic backbones remain an unexplored area. Given current standards in the fabrication of long sequences through solid phase synthesis and ligation methods[34,35,36], synthesis itself may not constitute a major stumble block towards artificial tertiary structures. It then appears that, even though interactions *within* abiotic secondary structural motifs are well-understood and often predictable, interactions *between* secondary elements have escaped our prediction capabilities. Furthermore, tertiary folding and helix bundling have until now been limited to water.

In the following, we report our success at rationally designing helix-helix interactional patterns to form tertiary folds within abiotic aromatic amide sequences and the subsequent characterization of different quaternary folds based on the same patterns. The approach rested on rather simple structure-based considerations and computational tools. We exploited the high stability of aromatic amide helices[37,38] that made a hierarchical approach to tertiary structure design possible: helices may not unfold upon interacting with each other and may thus be used

as stable construction modules. Nevertheless, we observed the typical cooperative behavior associated to the simultaneous occurrence of a large number of interactions. We also demonstrate that tertiary folding may occur in apolar solvents and be driven exclusively by hydrogen bonds. These results demonstrate that rational design is a viable approach to artificial tertiary structures and pave the way to engineering more complex and larger molecular shapes with atomic precision. Although function was not considered in the context of this study, abiotic tertiary folded objects prefigure the implementation of sophisticated molecular machineries including controlled motion, recognition, and catalysis in other media than water.

RESULTS AND DISCUSSION

Design of tertiary structures

Oligoamides of 8-amino-2-quinoline-carboxylic acid **Q** (Figure 1a) have been shown to adopt helical conformations stabilized by hydrogen bonds, and electrostatic repulsions, between quinoline nitrogen atoms and adjacent amide NH and CO groups, respectively, as well as by solvophobic forces associated with aromatic stacking.[2]**Erreur ! Signet non défini.**,**Erreur ! Signet non défini.**,[39] Non covalent interactions within these helices thus generate well understood motifs in which, unlike in peptides, amides NH point towards, and amide CO point away from, the helix axis. However, at the start of this work, there was no clue on how to orchestrate interactions between two helices and generate more complex architectures. We set to design a tertiary helix-turn-helix structure *ab initio* by simply examining computer models. We eventually found that a terephthalic acid bis-hydrazide **T** had an appropriate length to constitute a turn unit linking two helical $\mathbf{Q}_n$ segments. **T** inserts itself in the network of hydrogen bonds and prevents steric clashes between short helices while preserving enough proximity to be compatible with helix-helix interactions. Sequences such as $\underline{\mathbf{Q}}_4$-**T**-$\underline{\mathbf{Q}}_4$ in which quinoline rings are deprived of side chains may thus adopt two conformations: a C_2 symmetrical conformer where both $\underline{\mathbf{Q}}_4$ helices have right-handed (*P*) helicity (or left-handed (*M*) helicity) and are on the same side of the turn; and a centro-symmetrical *PM* conformation

in which helices are on opposite sides of the turn (Figure 1e, 1f)[40]. In the absence of helix-helix interactions, the *PP*/*MM* and *PM* conformers are expected to be equiprobable. A notable feature of the *PP*/*MM* conformer of **Q**$_4$-**T**-**Q**$_4$ is the relative proximity (3.5 to 4 Å) between the carbonyl amide of the second quinoline of each **Q**$_4$ segment (counting starts from *N*-terminus), and positions 3 and 4 of the second quinoline ring of the other **Q**$_4$ segment (Figure 1f). Based on this proximity we inferred that a hydrogen bond donor at these positions might reach the carbonyl group and generate attractive interactions between the two helices. The 4-hydroxy-group of unit **X** was thus added and energy minimized models predicted the formation of inter-helix hydrogen bonds stabilizing the start of a helix-turn-helix motif in the *PP*/*MM* conformer (Figure 1b, 1g).

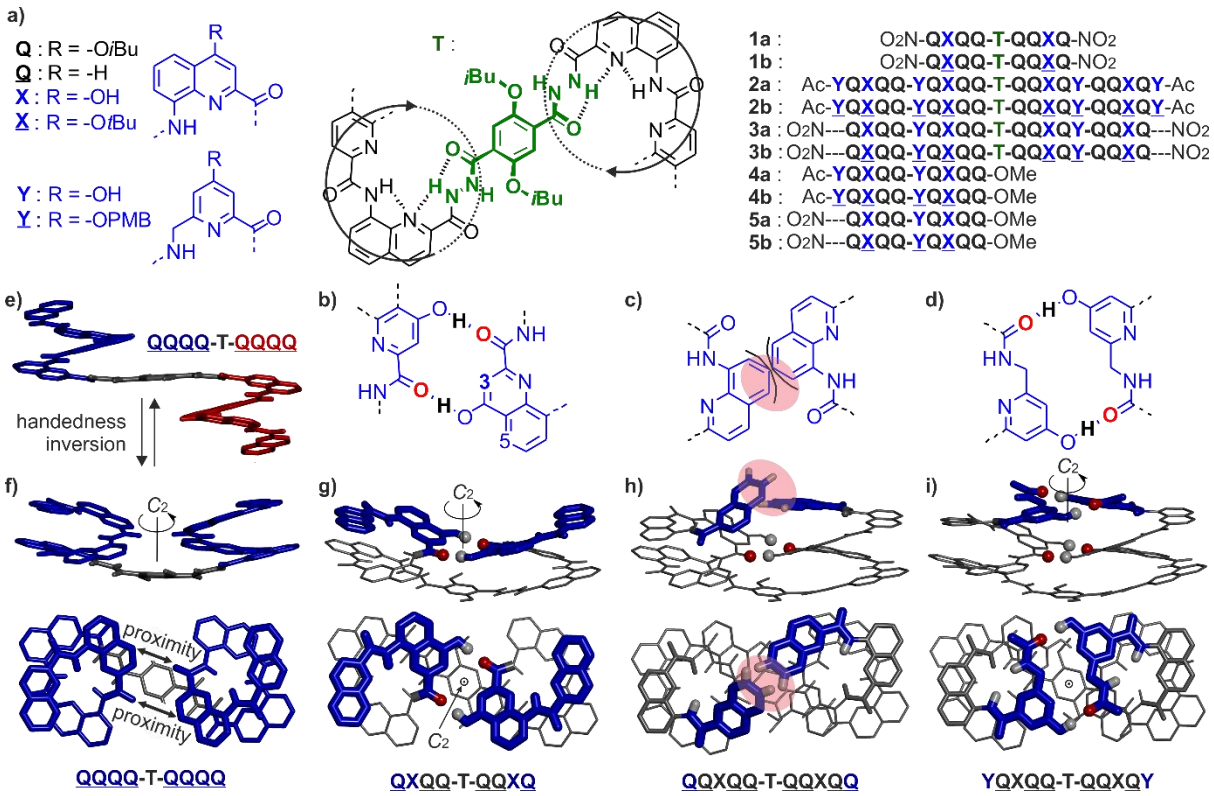

Figure 1. **Foldamer sequence and helix-turn-helix design**. a) Structures of units **Q**, **X**, **Y** and **T** (green part only), and alignment of sequence **1**-**5**. **X** and **Y** are the protected precursors of **X** and **Y**, respectively (PMB: 4-methoxy-benzyl). **Q** lacks any side chain and was used in

modelling studies only. Sequences are labelled "a" when deprotected, "b" when protected. **T** units constitutes an inversion of C→N sequence polarity; sequences containing **T** thus have two N-termini. Two circles schematize the projection of adjacent $\mathbf{Q_n}$ helices (in black) in the plane of **T** (in green). The arrows indicate in plane clockwise rotation when going from C- to N-terminus of a $\mathbf{Q_n}$ helix. The actual helix handedness depends on whether the helix is above or below the plane. For sequences terminated with 8-nitro groups, this group is noted in replacement of the NH group of **Q** or **X** units. b-d) hydrogen bonds or steric hindrance critical in the models shown in g-i), respectively. e-i) Energy minimised models. In e) and f), *P* and *M* helices are shown in blue and red, respectively. In g), h) and i), the units of interest at the top of the structures are shown as thick blue tubes. In the models, the isobutoxy groups of **T** were replaced by methoxy groups. Protons and oxygen atoms involved in hydrogen bonds are shown as white and red balls, respectively. A pink surface indicate a steric clash in c) and h).

This initial prediction was verified by experiments. Sequence **1b** was prepared (for synthetic Schemes see Supplementary Figures 1-4). Its ^{1}H NMR spectrum in $CDCl_3$ at 25 °C shows slightly broadened resonances that split upon cooling, consistent with the coexistence of *PP*/*MM* and *PM* conformers (Figure 2a, 2b)[31]. One species is favored which we assume to be *PM* (see below). On the contrary, the deprotected sequence **1a** shows one set of sharp NMR signals. Diastereotopic motifs of the signals of *i*BuO side chains indicate slow helix handedness inversion on the NMR time scale even at 25 °C (Supplementary Figure 5), implying that only one conformer could be detected. A crystal structure of **1a** was obtained (Figure 3a) which shows a *PP*/*MM* helix-turn-helix essentially superimposable to the computer prediction.

Models of elongated sequences showed that an additional **Q** monomer leads to steric clashes between the helices: CH groups in position 6 and 7 of the most peripheral quinoline rings hinder each other (Figure 1c, 1h). Instead, monomer **Y** (Figure 1a) brings the same

contribution to helix curvature as **X** yet lacks the aryl CH groups responsible for steric hindrance[41]. Modelling predicts that this monomer may allow the formation of two additional inter-helix hydrogen bonds involving the amide carbonyl of the next unit (Figure 1d, 1i).

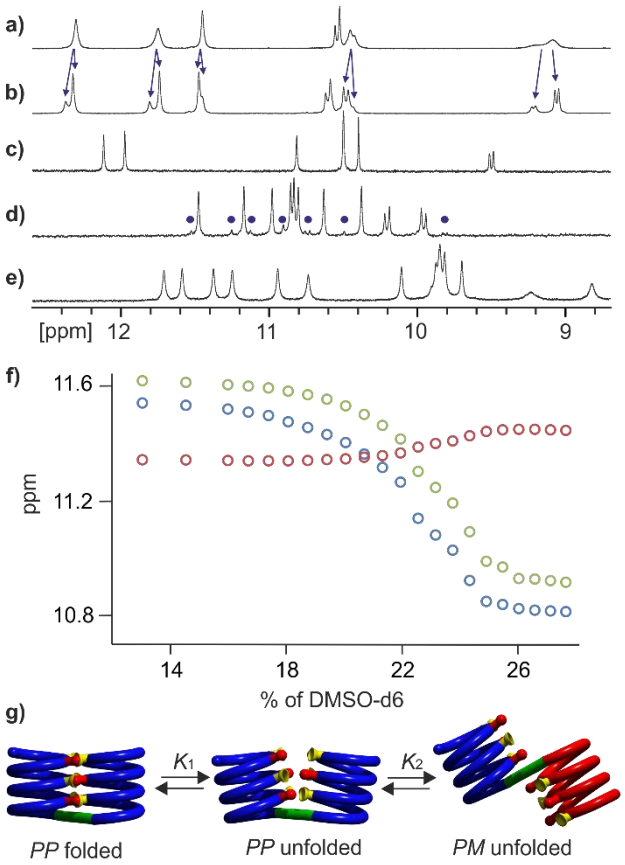

Figure 2. **Solution studies of helix-turn-helix motifs.** Excerpts of ^{1}H NMR spectra in $CDCl_3$ of: a) **1b** at 25 °C; b) **1b** at 0 °C; c) **1a** at 25 °C; d) **2b** at 25° C; e) **2a** at 25 °C. Arrows and circles indicate the presence of a minor and major species assigned to *PP*/*MM* and *PM* conformers, respectively. f) Variations of the chemical shift of selected ^{1}H NMR signals of **2a** upon addition of DMSO-d_6 (see also Supplementary Figure 12). g) Schematic representation of the equilibria involved in polar solvent induced helix-turn-helix unfolding. A similar equilibrium exist with the *MM* conformer.

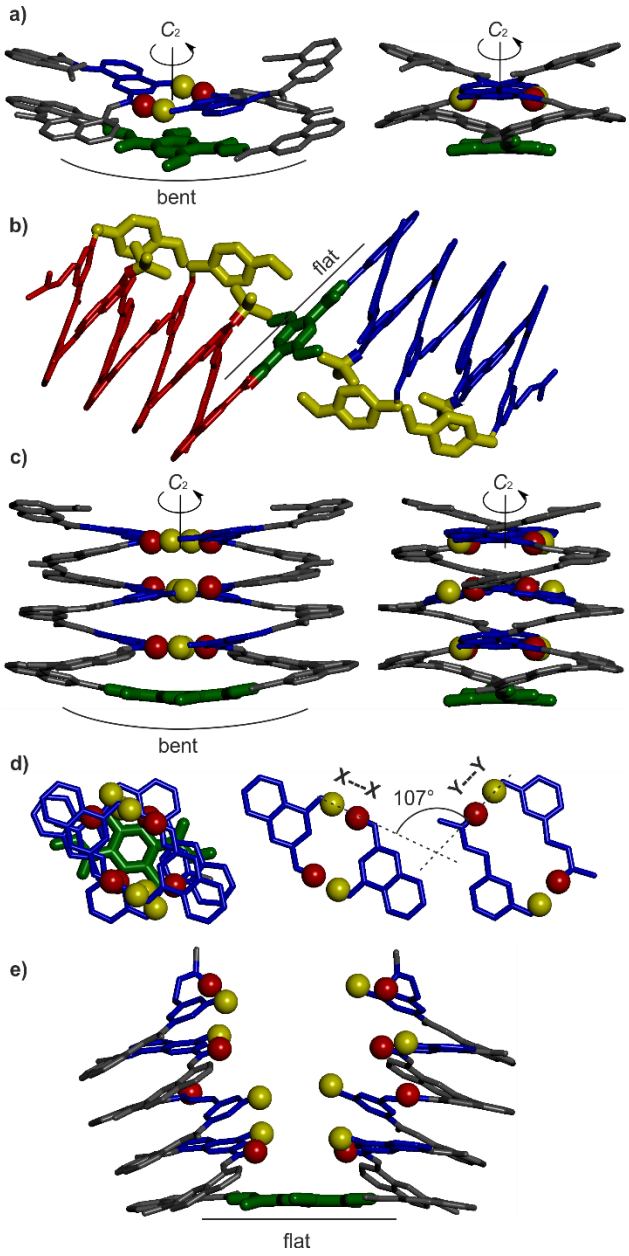

Figure 3. Helix-turn-helix motifs in the solid state. Crystal structures of **1a** (a), **2b** (b), **3a** (c,d), and **2a** (e) represented at the same scale. At left in d) a top view of the stack of **X**, **Y**, and **T** units of the structure of **3a** is shown. The central and right structures are hydrogen bonded pairs of **X** units and of **Y** units, respectively, translated horizontally from the stack, showing a distinct orientation of hydrogen bonds. The hydroxyl protons and carbonyl oxygen atoms of the hydrogen bonding arrays are shown as yellow and red balls respectively. **X** and **Y** units are shown as blue tubes, except in b) where blue tubes are for the *P* helix and red tubes for the *M* helix. In b), the PMB and *t*Bu protecting groups are shown as thick yellow tubes. For clarity,

*i*BuO side chains of **Q** units, included solvent molecules and most hydrogen atoms have been omitted, and only the oxygen atom and the first carbon of each *i*BuO chains of **T** units are shown (in green).

Sequence **2a** was then prepared whose two helices each consists of two contiguous **YQXQQ** segments. $\mathbf{Q_n}$ sequences have been shown to span about five units per two turns[Erreur ! Signet non défini.], suggesting that inter-helix hydrogen-bonds of consecutive pentads may be in register. This assumption is to be related to the behavior of heptad repeat motifs of α-peptidic zippers[19]. Heptads would be in perfect register if the α-helix spanned 3.5 units per turn. The fact that curvature is closer to 3.6 units per turn forces some adjustments, *i.e.* coil-coiling, to maintain helix-helix interactions over long distances[19]. Sequence **2a** contains one turn unit and twenty δ-amino-acids each of which has the same backbone length as a dipeptide. In terms of molecular weight, sequence **2a** is to be compared to a 42mer peptide which is larger than required for small tertiary folds to form[42,43,44]. The ^{1}H NMR of spectrum of protected sequence **2b** again shows two sets of signals reflecting the coexistence of *PP*/*MM* and *PM* species (Figure 2d). In this case, one species is strongly favored that can be assigned to the *PM* conformer by a crystal structure and the NMR spectrum of freshly dissolved crystals (Figure 3b, Supplementary Figure 6). The crystal structure shows that four *t*Bu and four PMB protecting groups would be in contact and generate hindrance if the helices were on the same side of the turn, *i.e.* in the *PP*/*MM* conformer. Upon deprotection into **2a**, only one species is observed (Figure 2e) showing that the two helices interact so as to completely bias the equilibrium. DOSY NMR testified that **2a** is monomeric in solution: **2a** and **2b** have similar diffusion coefficients (Supplementary Figures 7, 8). Assignment of OH protons to downfield shifted resonances between 9 and 10 ppm indicate their involvement in hydrogen bonds (Supplementary Figure 9). As our initial attempts to crystallize **2a** failed, sequence **3a** was

prepared. Its helical segments are both one unit shorter than in **2a** and are terminated by nitro-quinoline rings instead of acetyl-aminomethyl groups, a feature that we have observed to favour crystallization. The crystal structure of **3a** (Figure 3c,d, Supplementary Video 1) validated the predicted helix-turn-helix motif: no steric clash occurred and all six expected inter-helix hydrogen bonds were established (Figure 1b,1d), maintaining the two helices in a parallel orientation with the same handedness. In $CDCl_3$, **3a** shows spectral features similar to those of **2a**: one set of sharp signals whereas its protected precursor **3b** shows two sets (Supplementary Figure 10).

Eventually, diffracting crystals of **2a** were obtained from DMF. From this polar solvent, an open, "unfolded", *PP*/*MM* form crystallized (Figure 3e). This observation triggered the investigation of the solvent dependence of the helix-turn-helix motif. No denaturation of **2** was observed by NMR upon heating in $C_2D_2Cl_4$ up to 70 °C, reflecting high stability and hinting at a small entropy of folding. However, the addition of hydrogen bonding solvents disrupted the tertiary fold. No folding occurs in DMSO-d_6 or pyridine-d_5 as judged by the presence of both *PP*/*MM* and *PM* species (Supplementary Figure 11). Remarkably, DMSO-induced unfolding of **2a** occurs through a sharp transition between 18 and 23% (vol/vol) of the polar solvent (Figure 2f, Supplementary Figure 12) indicative of a cooperative phenomenon. Nothing similar was observed for **2b** (Supplementary Figures 15), confirming that changes observed for **2a** reflect a conformational transition and not a solvent dependence of chemical shift values. A similar transition is observed for **1a** but it is less sharp and requires less DMSO (Supplementary Figures 13, 14). Changes in NMR spectra during the transitions support the existence of two equilibria shown in Figure 2g. The first equilibrium (K_1) between the folded *PP* (or *MM*) helix-turn-helix and its unfolded form is solvent dependant and rapid on the NMR timescale: adding DMSO causes chemical shift variations as the proportions between the two species vary (Figure 2f,g). The second equilibrium between unfolded *PP*/*MM* and *PM* conformations is slow on the

NMR timescale. Unfolding thus also goes along with the emergence of the signals of the *PM* conformer whose protons resonate at frequencies that vary little with the proportion of DMSO (Supplementary Figure 12). The sharp transition between folded and unfolded state is to be related to the overall rigidity of the helix backbone which makes it difficult to not establish all inter-helix hydrogen bonds at the same time. Unfolding does not proceed progressively but through the simultaneous disruption of all hydrogen bonds. In addition, the folded helix-turn-helix is somewhat spring-loaded. X-ray structures of **1a** and **3a** show a certain level of bending of the turn units (Figure 3a,c) whereas it is flat in the unfolded state (Figure 3b,e). This bending had been predicted by modelling (Figure 1g-1i) and reflects that the turn is actually slightly longer than the inter-helix distance in the folded state. Furthermore, helix curvature in the helix-turn-helix is slightly reduced with respect to its conformation in a relaxed state (Supplementary Figure 51), suggesting a certain level of strain that is compensated by hydrogen bonding.

Formation of quaternary structures

We then explored the capacity of inter-helix hydrogen bonds to act intermolecularly. In the absence of a turn unit, additional degrees of freedom are allowed. For instance, **T** imposes hydrogen bonding to occur in register between helices with same handedness and parallel C→N sequence polarity. In the absence of **T**, out of register, anti-parallel arrangements, and association between *P* and *M* helices are allowed. The deca-amide sequence **4a** spans four helix turns and possesses two **X** and two **Y** units. Despite the possibilities mentioned above, its ^{1}H NMR spectrum in $CDCl_3$ shows one set of sharp lines (Figure 4a). Three OH resonances are found at 9.90, 10.90, and 11.36 ppm indicating their involvement in hydrogen bonds, but the fourth is not (δ = 7.70 ppm). A DOSY spectrum of a 1:1 mixture of **4a** and its protected form **4b** shows that **4a** is larger despite having a lower molecular weight, indicating an aggregated state (Supplementary Figures 18, 19). The aggregate does not dissociate upon diluting down to 0.30 mM. Upon adding CD_2Cl_2, a second set of signals emerges indicating the presence of a

distinct structure that becomes major in pure CD_2Cl_2 (Figure 4b,c). In this structure, all OH groups are involved in hydrogen bonds (δ = 10.51, 10.58, 11.07, and at 11.37 ppm). The exchange between the two species is very slow, showing that considerable rearrangements are needed for their interconversion. For instance, evaporating an equilibrated CD_2Cl_2 sample and dissolving the residue in $CDCl_3$ allows monitoring over 30 min (Supplementary Figures 24-30). This hints at an interconversion mechanism involving the simultaneous disruption of all hydrogen bonds within aggregates, possibly *via* a disfavored monomeric state.

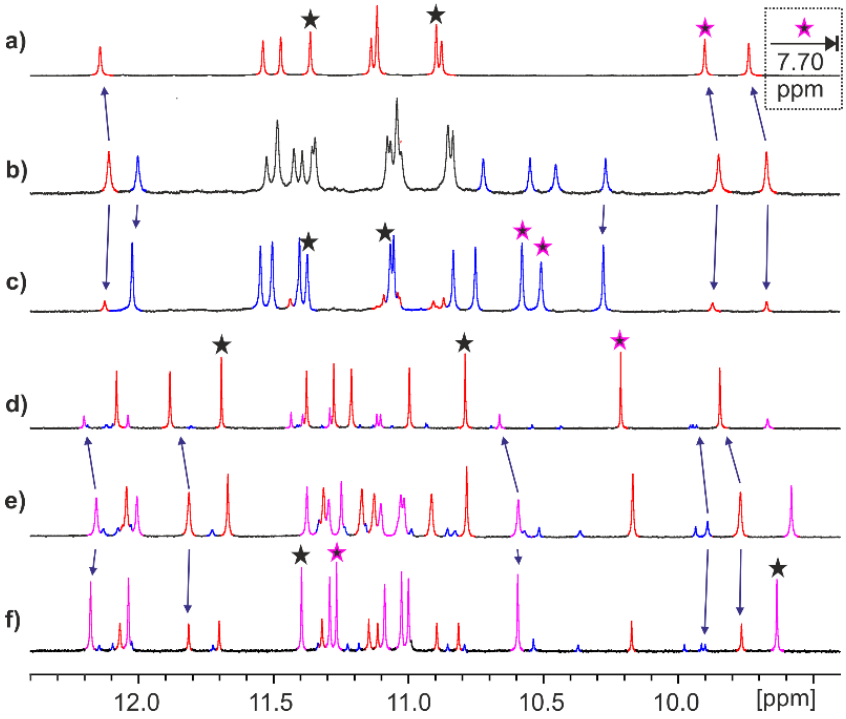

Figure 4. Solution studies of helix bundles. Excerpts of ^{1}H NMR spectra showing NH and OH resonances: 400 MHz of 3.0 mM solutions of: a) **4a** in $CDCl_3$; b) **4a** in $CDCl_3/CD_2Cl_2$ (1:1 vol/vol); c) **4a** in CD_2Cl_2; and 700 MHz spectra of 1.5 mM solutions of d) **5a** in $CDCl_3$; e) **5a** in $CDCl_3/CD_2Cl_2$ (1:1 vol/vol); f) **5a** in CD_2Cl_2; showing OH and NH resonances. Red and magenta lines have been assigned to homochiral dimeric helix bundles with tilted axes; blue lines have been assigned to parallel homochiral trimeric helix bundles. Black lines indicate overlapping signals of the various species. Black stars indicate X-OH signals; magenta star indicates Y-OH. Some arrows denote the corresponding signals in the various solvents.

A crystal structure of **4a** revealed the formation of a trimeric helix bundle comprised of three helices with the same handedness, parallel helix axes, and identical C→N sequence polarity (Figure 5a-f, Supplementary Video 2). A total of twelve hydrogen bonds are established between **X** units and between **Y** units within (**4a**)$_3$, similar to the eight hydrogen bonds of helix-turn-helix **3a**. However, hydrogen bond orientation is different: in both structures O···H-O angles are all close to 180°, but C=O··· H angles are also close to 180° in **3a** whereas they are closer to 120° in (**4a**)$_3$. In (**4a**)$_3$, hydroxyl protons thus point straight to carbonyl lone pairs, an orientation that has been showed to be favoured[45] but that is made impossible in **2a** or **3a** due restrictions imposed by turn units.

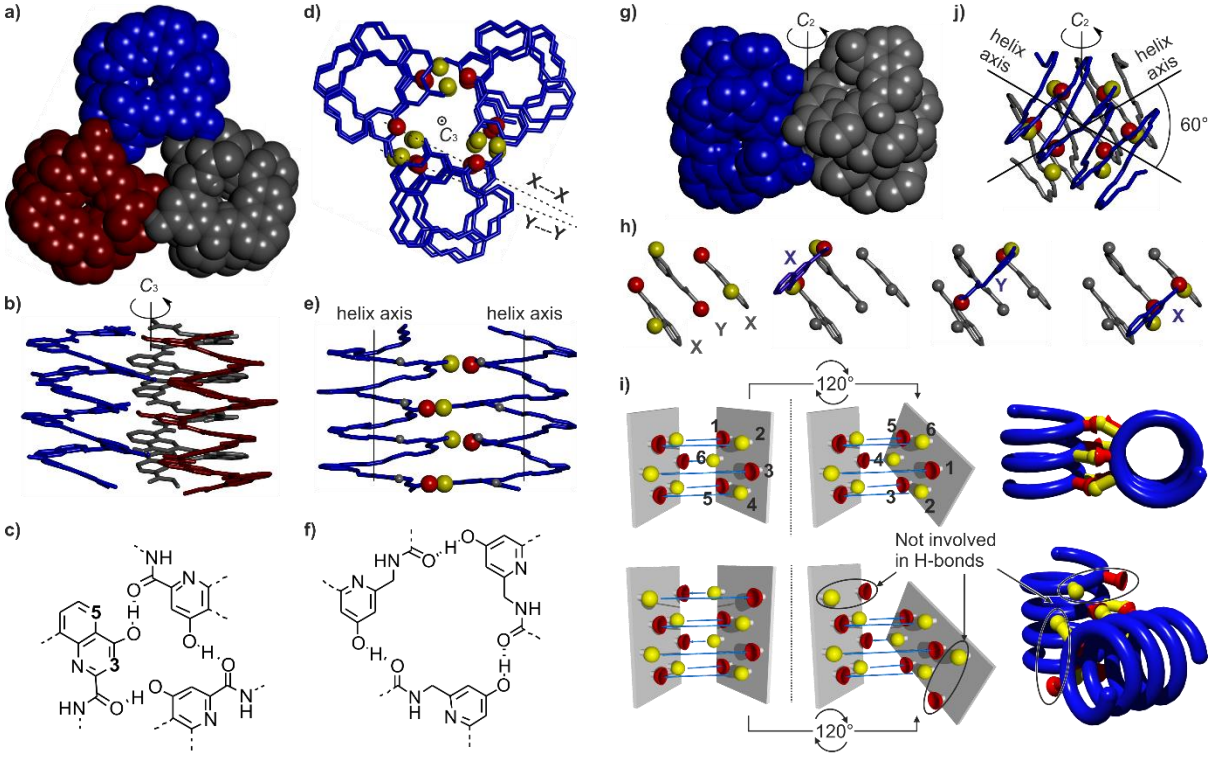

Figure 5. Helix bundles in the solid state. Crystal structures of (**4a**)$_3$ (a-f) and (**5a**)$_2$ (g-j). In all representations, *i*BuO side chains of **Q** units, included solvent molecules and most hydrogen atoms have been omitted for clarity. a,b) Top view (CPK) and side view (tube) of the trimeric helix bundle (**4a**)$_3$. d,e) Top view and side view of (**4a**)$_3$ showing H-bond donors and acceptors

as golden yellow and red spheres, respectively. Only the outer rim of the helices is shown as blue tube. In e), only two helices and one array of inter-helix H-bonds are shown. c,f) hydrogen-bonded trimeric arrangements of **X** and **Y** units. In c), OH groups point toward of the position 5 of the **X** unit to which they belong, in contrast with helix-turn-helix structures in which OH groups point towards position 3 (Figure 1b). Dashed lines in d) show that **X**···**X** and **Y**···**Y** hydrogen bonds are almost parallel, in contrast with the structure of **3a** (Figure 3d). g) Side view of the tilted dimeric bundle (**5a**)$_2$ (CPK representation). h) Front view of (**5a**)$_2$ showing the 60° tilt angle between helix axes. H-bond donors and acceptors are shown as yellow and red spheres, respectively. Only the outer rim of the helices is shown as blue tube. h) Same view as in j, showing exclusively the **X** and **Y** units of one helix in grey (left) and each of the hydrogen bonded **X** and **Y** units of the other helix in blue. i) Schematic representation of how a given array of six hydrogen bond donors (yellow) and acceptors (red) may allow hydrogen bonding both a in parallel arrangement, and after an 120° rotation.

The NMR of freshly dissolved crystals of (**4a**)$_3$ shows that this is the dominant species in CD_2Cl_2 (**4a**)$_3$ (Supplementary Figure 23), consistent with the observation that all OH groups are involved in H-bonds (Figure 4c). Dilution causes the dissociation of (**4a**)$_3$ into the other species which, we have shown above, is also an aggregate, and therefore cannot be anything else than dimeric (Supplementary Figure 32). Ion mobility mass spectrometry (IMS) measurements confirm the coexistence of dimeric and trimeric aggregates (Supplementary Figure 21). The concentration dependence allows to quantify the equilibrium between dimers and trimers: 3x(**4a**)$_2$ <-> 2x(**4a**)$_3$ ($K_{\mathbf{4a}}$) with $K_{\mathbf{4a}}$ = 2.2 10^5 M^{-1} in CD_2Cl_2 and < 0.2 M^{-1} in $CDCl_3$. We also find that other solvents (CCl_4, $C_2D_2Cl_4$, toluene) affect this equilibrium in various ways (Supplementary Figure 33). The reason for this solvent dependence is unclear. Pyridine and DMSO cause the dissociation of all aggregates into monomeric helices

(Supplementary Figure 34). The effect is more progressive and less cooperative than in the helix turn helix bundle.

The solution behaviour of shorter sequence **5a** shed light on the nature of (**4a**)$_2$. Its ^{1}H NMR spectrum in $CDCl_3$ shows a major species and two distinct minor species (Figure 4d). A DOSY spectrum of a mixture of **4b** and **5a** in $CDCl_3$ indicates a larger size for the latter which must therefore be a well-defined aggregate (Supplementary Figures 16, 17). Upon adding CD_2Cl_2, the proportion of one of the two other species increased to become major in pure CD_2Cl_2, whereas the third species remained minor (Figure 4f). In both major species, all OH groups are hydrogen bonded. Their relative proportion does not depend on concentration indicating that they are composed of the same number of molecules. In contrast, the proportion of the minor species increases with increasing concentration showing that it has a higher molecularity (Supplementary Figure 31). Dissolving crystals of **5a** (see below) allowed us to assign the two dominant species to dimers (**5a**)$_2$. IMS also registered the presence of a trimeric aggregate (Supplementary Figure 20). The minor NMR signals were thus assigned to a trimer (**5a**)$_3$ which may be similar to (**4a**)$_3$. The equilibrium constant between the two dimers dimer$_{CDCl3}$ <-> dimer$_{CD2Cl2}$ was calculated to be $K_{\mathbf{5a}}$ = 2.5 in CD_2Cl_2 and 0.24 in $CDCl_3$. Again, we found no obvious explanation to the solvent dependence.

The structure in the solid state of the species that prevails in $CDCl_3$ was elucidated (Figure 5g-j, Supplementary Video 3), its identity being established by NMR spectra of freshly dissolved crystals (Supplementary Figure 22). It shows a dimeric arrangement of two helices with the same handedness. The two helix axes are not parallel but tilted by 120° counter-clockwise (for *P* helices) with respect to the expected parallel head-to-head arrangement (or 60° clockwise with respect to a parallel head-to-tail arrangement, Figure 5j). All hydroxyl groups are involved in hydrogen bonds but these are not established between homomeric pairs of **X** or **Y** units, but directly between **X** and **Y** (Figure 5h). As depicted in Figure 5i (top), the

positions of hydrogen bonding sites in one face of a helix can be rotated by 120° and still lead to a matching array of donors and acceptors. The consequence of the 120° tilt in a helix having not 6 but 8 hydrogen bonding sites like **4a** is that two hydrogen bonding sites will be unoccupied (Figure 5i, bottom). This is consistent with the ^{1}H NMR spectrum of (**4a**)$_2$ in $CDCl_3$ which can reasonably be assigned to a tilted dimer apparently favored in this solvent.

The nature of the (**5a**)$_2$ dimer that prevails in CD_2Cl_2 can be proposed based on the partial assignment of NMR spectra and remarkable NOE correlations within the various complexes of **4a** and **5a** (Supplementary Figures 35-46). Head-to-head arrangements of helices of opposite handedness can be ruled out because they lead to **X**-**X** and **Y**-**Y** mismatches (donors facing donors, acceptors facing acceptors), or **X**-**Y** mismatches (different distances between pairs of donors and acceptors). The parallel head-to-head arrangement of helices of identical handedness, as it is found in the structure of **3a**, are less favored as the formation of (**4a**)$_3$ demonstrates. A similar head-to-tail arrangement of helices of opposite handedness (Supplementary Figure 47) is also ruled out based on observed NOEs. Tilted arrangements of helices of opposite handedness (Supplementary Figure 48) can also be ruled out because they would have no symmetry element and give rise to twice the number of ^{1}H resonances actually observed (*i.e.* no degeneracy). Finally, the only remaining and most probable option is that the (**5a**)$_2$ dimer that prevails in CD_2Cl_2 is also a tilted dimer between helices of identical handedness, but with a tilt direction opposite to that shown in Figure 5j (Supplementary Figures 47-50).

CONCLUSION

Using simple computational considerations, hydrogen bonding has been designed at defined locations on aromatic amide backbones that eventually promote cooperative folding into helix-turn-helix motifs in organic solvents. This approach rested on the inherent stability of aromatic amide helices and may be extended to interacting functionalities other than aromatic hydroxyl

and amide carbonyl groups. The hydrogen bond-directed aggregation of helices not linked by a turn unit produced several thermodynamically and kinetically stable dimeric and trimeric homochiral bundles having structures distinct from the designed helix-turn-helix. Relative helix orientation within the bundles may be changed from parallel to tilted upon subtle solvent variations such as replacing dichloromethane by chloroform. Altogether, these results prefigure the richness and uniqueness of abiotic tertiary structure behavior. Combining parallel and tilted motifs in the same structure using appropriately designed turn units would yield yet another level of complexity in abiotic tertiary folding. The high stability and rigidity of the helix bundles presented here allow one to speculate that they might accommodate the presence of more flexible backbone features. This would in turn enhance folding entropy and result in melting of the structures upon heating in a tunable fashion. Progress along these lines is being made and will be reported in due course.

Supplementary Information is linked to the online version of the paper at http://www.nature.com/nature.

Acknowledgments. This work was supported by the European Research Council under the European Union's Seventh Framework Programme (Grant Agreement No. ERC-2012-AdG-320892), by the European Union under the People program (FP7 PIIF-2009-254156 postdoctoral fellowship to T.Q.) and by the French-Chinese Foundation for Science and its Applications (postdoctoral fellowship to B.C.). The contribution of Dr. Post in optimizing the synthesis of 4-*t*BuO protected quinoline monomers is acknowledged. The authors thank Dr. Kauffmann for assistance with x-ray data collection and structure resolution at IECB's X-ray

diffraction facility and Dr. Ferrer for beamtime and help during data collection on FIP-BM30A at the European Synchrotron Radiation Facility.

Author Contributions. S.D. and B.C. contributed equally to this work. S.D., B.C. and T.Q. synthesized all new compounds, carried out solution studies and grew single crystals. T.G. refined the crystal structures. I.H., and V.M. designed the research and carried out modelling studies. I.H. wrote the manuscript. All authors discussed the results and commented on the manuscript.

Additional Information. The crystallographic data and experimental details of the structural refinement for the X-ray crystal structures of **1a**, **2a**, **2b**, **3a**, **4a** and **5a** have been deposited at the Cambridge Crystallographic Data Centre, under deposition numbers CCDC 1450791, 1451415, 1469843, 1451523, 1470116 and 1451494, respectively. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre (http://www.ccdc.cam.ac.uk/data_request/cif). Reprints and permissions information is available at http://www.nature.com/reprints. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to I.H. (i.huc@iecb.u-bordeaux.fr).

REFERENCES

1 Guichard, G.; Huc, I. Synthetic foldamers. *Chem. Commun.* **47**, 5933–5941 (2011).

2 Zhang, D. W.; Zhao, X.; Hou, J. L.; Li, Z. T. Aromatic amide foldamers: Structures, properties, and functions. *Chem. Rev.* **112,** 5271–5316 (2012).

3 Gellman, S. H. Foldamers: A Manifesto. *Acc. Chem. Res.* **31,** 173–180 (1998).

4 Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. Foldamers as versatile frameworks for the design and evolution of function. *Nat. Chem. Biol.* **3,** 252–262 (2007).

5 Nair, R. V.; Vijayadas, K. N.; Roy, A.; Sanjayan, G. J. Heterogeneous foldamers from aliphatic-aromatic amino acid building blocks: current trends and future prospects. *Eur. J. Org. Chem.* **2014,** 7763–7780 (2014).

6 Scott Lokey, R.; Iverson, B. L. Synthetic molecules that fold into a pleated secondary structure in solution. *Nature* **375,** 303–305 (1995).

7 Brüggemann, J. et al. Spontaneous knotting - From oligoamide threads to trefoil knots. *Angew. Chem. Int. Ed.* **46,** 254–259 (2007).

8 Ponnuswamy, N.; Cougnon, F. B. L.; Clough, J. M.; Pantoş, G. D.; Sanders, J. K. M. Discovery of an organic trefoil knot. *Science* **338,** 783–785 (2012).

9 Hua, Y.; Liu, Y.; Chen, C. H.; Flood, A. H. Hydrophobic collapse of foldamer capsules drives picomolar-level chloride binding in aqueous acetonitrile solutions. *J. Am. Chem. Soc.* **135,** 14401–14412 (2013).

10 Chandramouli, N. et al. Iterative design of a helically folded aromatic oligoamide sequence for the selective encapsulation of fructose. *Nat. Chem.* **7,** 334–341 (2015).

11 Huang, P.-S. et al. De novo design of a four-fold symmetric TIM-barrel protein with atomic-level accuracy. *Nat. Chem. Biol.* **12,** 29–34 (2016).

12 Boyken, S. E. et al. De novo design of protein homo-oligomers with modular hydrogen-bond network-mediated specificity. *Science* **352**, 680–687 (2016).

13 Hilvert, D. Design of protein catalysts. *Annu. Rev. Biochem.* **82,** 447–70 (2013).

14 Khoury, G. A.; Smadbeck, J.; Kieslich, C. A.; Floudas, C. A. Protein folding and de novo protein design for biotechnological applications. *Trends in Biotechnology* **32**, 99–109 (2014).

15 Parmeggiani, F. et al. D. A general computational approach for repeat protein design. *J. Mol. Biol.* **427,** 563–575 (2015).

16 Price, J. L. et al. Design of a three-helix bundle capable of binding heavy metals in a triscysteine environment. *Angew. Chem. Int. Ed.* **49,** 368–371 (2010).

17 Grigoryan, G.; Degrado, W. F. Probing designability via a generalized model of helical bundle geometry. *J. Mol. Biol.* **405,** 1079–1100 (2011).

18 Joh, N. H. et al. De novo design of a transmembrane Zn^{2+}-transporting four-helix bundle. *Science* **346,** 1520–4 (2014).

19 Woolfson, D. N. The Design of Coiled-Coil Structures and Assemblies. *Adv. Protein Chem.* **70**, 79–112 (2005).

20 Gradišar, H. et al. Design of a single-chain polypeptide tetrahedron assembled from coiled-coil segments. *Nat. Chem. Biol.* **9,** 362–366 (2013).

21 Reinert, Z. E.; Lengyel, G. A.; Horne, W. S. Protein-like tertiary folding behavior from heterogeneous backbones. *J. Am. Chem. Soc.* **135,** 12528–12531 (2013).

22 Mayer, C.; Müller, M. M.; Gellman, S. H.; Hilvert, D. Building proficient enzymes with foldamer prostheses. *Angew. Chemie Int. Ed.* **53,** 6978–6981 (2014).

23 Tavenor, N. A.; Reinert, Z. E.; Lengyel, G. A.; Griffith, B. D.; Horne, W. S. Comparison of design strategies for α-helix backbone modification in a protein tertiary fold. *Chem. Commun.* **52,** 2–5 (2016).

24 Horne, W. S.; Price, J. L.; Keck J. L.; Gellman S. H. Helix-Bundle Quaternary Structure from α/β-Peptide Foldamers. *J. Am. Chem. Soc.***129**, 417

8–4180 (2007).

25 Cheng, P. N.; Pham, J. D.; Nowick, J. S. The supramolecular chemistry of β-sheets. *J. Am. Chem. Soc.* **135,** 5477–5492 (2013).

26 Kreutzer, A. G.; Hamza, I. L.; Spencer, R. K.; Nowick, J. S. X-ray Crystallographic Structures of a Trimer, Dodecamer, and Annular Pore Formed by an $A\beta_{17\text{-}36}$ β-Hairpin. *J. Am. Chem. Soc.* **138,** 4634–4642 (2016).

27 Daniels, D. S.; Petersson, E. J.; Qiu, J. X.; Schepartz, A. High-resolution structure of a β-peptide bundle. *J. Am. Chem. Soc.* **129,** 1532–1533 (2007).

28 Petersson, E. J.; Craig, C. J.; Daniels, D. S.; Qiu, J. X.; Schepartz, A. Biophysical characterization of a β-peptide bundle: Comparison to natural proteins. *J. Am. Chem. Soc.* **129,** 5344–5345 (2007).

29 Collie, G. W. et al. Shaping quaternary assemblies of water-soluble non-peptide helical foldamers by sequence manipulation. *Nat. Chem.* **7,** 871–8 (2015).

30 Sharma, G. V. M. et al. Design and synthesis of peptides with hybrid helix-turn-helix (HTH) motif and their conformational study. *J. Org. Chem.* **79,** 8614–8628 (2014).

31 Delsuc, N.; Massip, S.; Léger, J. M.; Kauffmann, B.; Huc, I. Relative helix-helix conformations in branched aromatic oligoamide foldamers. *J. Am. Chem. Soc.* **133,** 3165–3172 (2011).

32 Ichinose, W.; Ito, J.; Yamaguchi, M. Tetrameric αααββ aggregate formation by stereoisomeric bidomain helicene oligomers. *Angew. Chem. Int. Ed.* **52,** 5290–5294 (2013).

33 Gan, Q. et al. Translation of rod-like template sequences into homochiral assemblies of stacked helical oligomers. *Nat. Nanotech.*, 12, 447–452 (2017).

34 Kent, S. B. H. Total synthesis of proteins. *Chem. Soc.Rev.* **38**, 338–351 (2009).

35 Pusterla, I.; Bode, J. W. An oxazetidine amino acid for chemical protein synthesis by rapid, serine-forming ligations. *Nat. Chem.* **7**, 668–672 (2015).

36 Haj-Yahya, M. et al. Synthetic polyubiquitinated α-Synuclein reveals important insights into the roles of the ubiquitin chain in regulating its pathophysiology. *Proc. Natl. Acad. Sci. USA* **110**, 17726–17731 (2013).

37 Jiang, H.; Léger, J. M.; Huc, I. Aromatic δ-peptides. *J. Am. Chem. Soc.* **125,** 3448–3449 (2003).

38 Delsuc, N. et al. Kinetics of helix-handedness inversion: Folding and unfolding in aromatic amide oligomers. *ChemPhysChem* **9,** 1882–1890 (2008).

39 Qi, T. et al. Solvent dependence of helix stability in aromatic oligoamide foldamers. *Chem. Commun.* **48,** 48–51 (2012).

40 Maurizot, V. et al. Design of an inversion center between two helical segments. *J. Am. Chem. Soc.* **126,** 10049–10052 (2004).

41 Sánchez-García, D. et al. Nanosized hybrid oligoamide foldamers: Aromatic templates for the folding of multiple aliphatic units. *J. Am. Chem. Soc.* **131,** 8642–8648 (2009).

42 Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Designing a 20-residue protein. *Nat. Struct. Biol.* **9**, 425–430 (2002).

43 Hodges, A. M.; Schepartz, A. Engineering a Monomeric Miniature Protein. *J. Am. Chem. Soc.* **129**, 11024−11025 (2007).

44 Craven, T. W.; Cho, M. K.; Traaseth, N. J.; Bonneau, R.; Kirshenbaum, K. A Miniature Protein Stabilized by a Cation-π Interaction Network. *J. Am. Chem. Soc.* **138,** 1543–1550 (2016).

45 Lommerse, J. O. S. P. M.; Price, S. L.; Taylor, R. Hydrogen Bonding of Carbonyl, Ether, and Ester Oxygen Atoms with Alkanol Hydroxyl Groups. *J. Comput. Chem.* **18,** 757–774 (1997).

Table of contents summary

Can foldamers remote from peptides and nucleotides also adopt complex conformations composed of several sub-domains? We report that designed arrays of hydrogen bonds between aromatic oligoamide segments generate helix-turn-helix and unexpected dimeric and trimeric helix bundle motifs that show kinetic and thermodynamically stability and cooperative folding in nonpolar solvents.