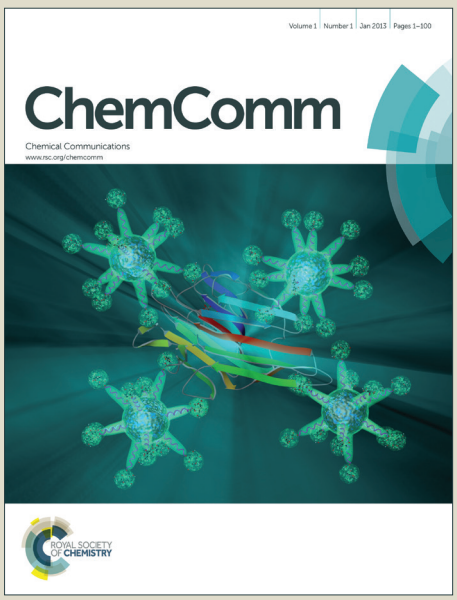
# ChemComm

Accepted Manuscript





ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/chemcomm

Journal Name

RSCPublishing

COMMUNICATION

# Structural Elucidation of Foldamers with No Long Range Conformational Order

Cite this: DOI: 10.1039/xoxxoooooox

Mayumi Kudo,[a,b,c] Victor Maurizot,[b,c] Hyuma Masu,[d] Aya Tanatani*[,a] and Ivan Huc*[,b,c]

Received ooth January 2012,
Accepted ooth January 2012

DOI: 10.1039/xoxxoooooox

www.rsc.org/

**The synthesis and structural investigations of aromatic-aliphatic oligoamide foldamers reveal a zig-zag tape conformation with local conformational variability that precludes long range order.**

Partly folded structures are thought to be common but, due to the coexistence of several conformers, they generally escape structural investigations and little is known about the possible structural patterns that they may display. As a rare example of its kind, this study presents an accurate structure elucidation of a partly-folded foldamer and sheds light into how organization and disorganization may coexist at various levels within a given foldamer sequence.

Synthetic foldamers – artificial folded molecular architectures produced by step-wise synthesis – attract widespread attention.[1] Investigations to discover foldamer backbones and new folding patterns have been directed by design when reasonable predictions about the folded structures could be made from first principles.[2] However, the field has also been curiosity driven, particularly so in the case of heteromeric hybrid sequences.[3,4] It is thus not uncommon that investigations start without preconceived ideas of whether folding will occur at all. A single conformation amenable to straightforward structural assignment may emerge. Yet, more often than not, partial folding occurs and the coexistence of multiple conformation impedes their structural characterization. Foldamers may possess local but not long range conformational order. Conversely, foldamers may be organized overall (e.g. helical) but show multiple local structural fluctuations as the *cis-trans* tertiary amide isomerism in peptoids.[4]

Foldamers with only partially defined folding are not necessarily less important than stable structures, and may have as much potential for applications as is well illustrated by peptoids.[4] However, ill-folded oligomers are often ill characterized. Indications of folding can be obtained by the chain-length dependence of some spectroscopic properties.[5] Yet, this may not allow to discriminate unique from multiple coexisting conformations. Circular dichroism (CD) bands which sum up the signatures of all species, and NMR resonances, which represent pondered averages of the signals of species under rapid exchange, may not be deconvoluted.[6,7] Even x-ray structures may be misleading when they correspond to a snapshot of potentially diverse conformations in solution. Few spectroscopic techniques allow the direct observation of coexisting conformers.[8]

This work developed while exploring the folding of hybrid sequences comprised of aromatic and aliphatic monomers. In such sequences, backbone chemical diversity and folding patterns rapidly expand upon combining different monomers.[3,9-11] Aromatic oligoamides derived from 8-amino-2-quinolinecarboxylic acid **Q** (Chart 1) adopt exceptionally stable helical conformations.[12] When aliphatic units are incorporated into these helices, quinoline monomers dictate their folding behaviour to the aliphatic monomers[13] unless the proportion of the latter is high, in which case new folding patterns have been observed such as herringbone helices.[11] These findings hold upon combining **Q** and α-amino acids, despite the absence of feature that would make their folding be compatible *a priori*. Sequences comprised of leucine (**L**) in an (**LQ$_2$**) repeat motif adopt stable helical canonical conformations directed by the propensity of crescent-shaped **Q$_2$** dimers to stack on top of each other.[14] To challenge the folding-directing behaviour of **Q** monomers, we endeavoured to increase the proportion of **L** units and prepared (**LQ**)$_n$ oligomers with n = 1-8 (Fig. 1).[†] The synthesis proceeded as before[14] using an iterative segment doubling approach based on optimized procedures.[15] For the purpose of racemic crystallographic investigations, syntheses were carried in both the L and D series.[16]

The conformational behaviour of the oligomers in solution was first analysed by CD (Fig. 2). While dipeptide **1** showed no CD signal in the 230-400 nm region, longer oligomers **2**-**4** showed CD bands of moderate intensity suggesting the existence of some defined structures. Maximal Δε values for **2**-**4** were 23.7 (253.4 nm), 50.2 (257.2 nm), and 78.3 (258.2 nm) cm$^2$ mmol$^{-1}$, respectively. They thus increase with oligomer length and show a slight red shift

of the band. However, Δε values normalized by the number of quinolines vary little, indicating a lack of cooperativity, in contrast with $(\mathbf{LQ_2})_n$ oligomers.[14]

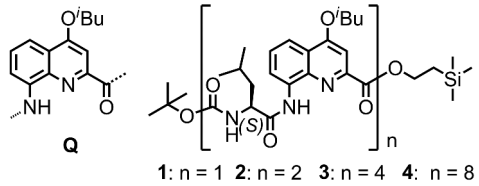



**Fig. 1**. Structures of (LQ)n oligomers

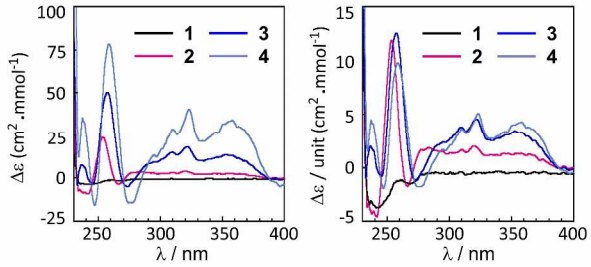



**Fig. 2**. Circular dichroism spectra of **1-4** in $CHCl_3$ at 25°C.

A crystal structure of **1** provided no hint about what conformations of longer oligomers may be, except for the expected coplanarity of the quinoline ring and of the amide or ester moieties that it bears.[14] Multiple attempts to grow crystals of **2-4** (single L-enantiomers) that would be suitable for x-ray crystallographic analysis proved unsuccessful. Our experience is that racemic or quasi-racemic crystals of aromatic oligoamides or $(\mathbf{LQ_2})_n$ oligomers grow much more readily,[14,16] as is the case for other peptides and small proteins.[17] We thus prepared rac-$(\mathbf{LQ})_2$ and rac-$(\mathbf{LQ})_4$ by mixing the corresponding sequences synthesized in the L and D series. This effort proved to be rewarding as the structure of octapeptide rac-$(\mathbf{LQ})_4$ in the solid state could be solved in the *P*21/c centrosymmetrical space group (Fig. 3). This structure stands as an unusual case in the vast body of literature on foldamers in that it shows no regular pattern despite the relative length of the sequence. Examples of crystallo-graphic data of at least partially unfolded foldamers are uncommon.[13a,18] The various views of $(\mathbf{LQ})_4$ shown in Fig. 3 illustrate that quinoline rings and leucine side chains adopt various orientations with respect to each other. Nevertheless, the very fact that crystals formed suggests that the conformational space available to $(\mathbf{LQ})_4$ re-mains relatively limited – too flexible strands are not good candidates for crystal growth.

Unlike all other oligomers containing **Q** units known so far,[11-14] the structure of $(\mathbf{LQ})_4$ shows no intramolecular aromatic stacking between quinoline rings. The only apparent organization of $(\mathbf{LQ})_4$ consists of a central stretch arranged in a zig-zag tape held by intramolecular hydrogen bonds (Table 1), as can be viewed in Fig. 3c, 3d and Fig. 4. Two quinolines and two leucines are involved in this segment. At each quinoline ring, two amide NHs form bifurcated hydrogen bonds with a single amide carbonyl oxygen atom. This pattern has been commonly encountered in systems comprised of 2,6-pyridinedicarboxamide units.[19] The shortest hydrogen-bonded rings amount to peptidic γ-turns centered at each leucine. The largest hydrogen-bonded rings amount to much less common peptidic pseudo-α-turns,[20,#] an unprecedented motif in hybrid aliphatic-aromatic oligomers. These hydrogen bonds are not formed at the C- and N-terminal **LQ** dipeptides. Several factors may contribute to this. At the C-terminus an ester function precludes the formation of the pseudo-α-turn and might also cause an electrostatic repulsion that hampers the formation of the γ-turn. At the N-terminus, the required carbonyl group belongs to a carbamate, a weaker hydrogen bond acceptor than amides. As another factor, one may invoke packing in the crystal structure, the N-terminal and C-terminal quinoline rings being involved in face-to-face stacks in the crystal lattice (see supporting information).

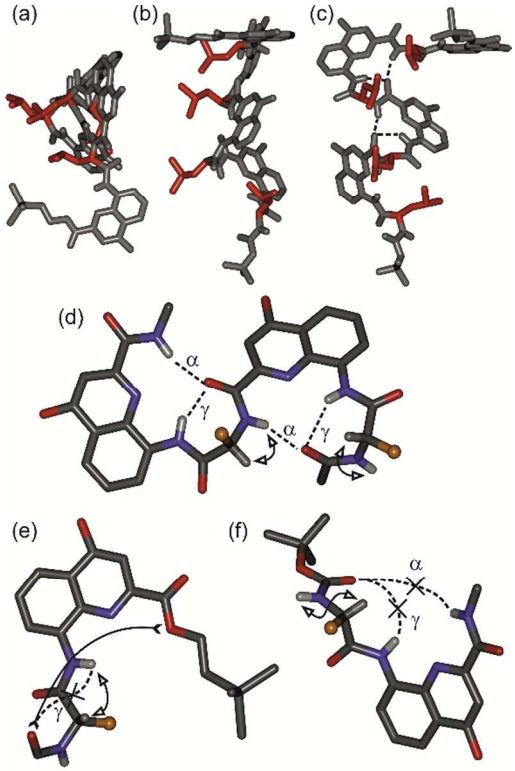



**Fig. 3**. Crystal structure of *rac*-$(\mathbf{LQ})_4$. Only the L-enantiomer is shown. (a-c) Top view, front view and side view. Leucine side-chains are shown in red. Hydrogen atoms other than NH, isobutyl from **Q** monomers and included solvent molecules have been removed for clarity. (d-f) Side views of the two central, of the C-terminal, and of the N-terminal **LQ** units, respectively. Intramolecular hydrogen bonds defining γ-turn and pseudo-α-turn motifs are shown as dashed lines and barred with a cross when not established. Double headed arrows delineate different NH-$C_\alpha$H dihedral angles. In (e), a doubled headed inverted arrow indicates a potential electrostatic repulsion. Hydrogen atoms have been removed except NH and $C_\alpha$H. Isobutyl groups have been removed from **Q** monomers and replaced by a golden ball in **L** monomers for clarity.

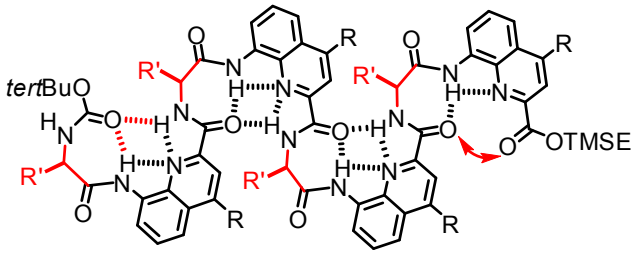



**Fig. 4**.Schematic view of the zig-zag tape conformation of **3**. Parts in red indicate where structural variability may occur: different conformations of the γ-turns, weak hydrogen bonds with the N-terminal Boc-carbonyl group, electrostatic repulsion (double headed arrow) with the C-terminal ester function.

One may infer from the structure of **3** that hydrogen bonded α- and γ-turns would endow longer sequence with a certain degree of organization, the terminal Q units only not being involved in the zig-

zag shaped tape. While, NMR data support this hypothesis (see below), the crystal structure of **3** also reveals an additional degree of freedom that precludes long range order. Indeed, conformations about the γ-turns appear to be variable, with the $C^{\alpha}$-H and N-H bonds almost eclipsed in one case, and close to be in a *trans* conformation in the other case (Fig. 3d). As a result, the corresponding leucine side chains are found on the same side of the tape, whereas a conserved arrangement would result in an alternation of the side chains in positions above and below the tape. Such variability is clearly expressed in the $\psi$ and $\phi$ values shown in Table 2.

Table 1. Angle and distance parameters of intramolecular hydrogen bonds in the crystal structure of 3.

| | γ-turns | | α-turns | |
|---|---|---|---|---|
| $d_{N\text{-}O}$ (Å) | 2.8 | 3.0 | 3.0 | 2.9 |
| NHO (°) | 156 | 136 | 160 | 146 |

Table 2. $\phi$ and $\psi$ angle values at each leucine residue in the crystal structure of 3.

| | L1[a] | L2[a] | L3[a] | L4[a] |
|---|---|---|---|---|
| $\phi$ (°) | 148 | 73 | -60 | 166 |
| $\psi$ (°) | -72 | -89 | 71 | -74 |

[a]Leucine units are numbered starting from the N terminus.

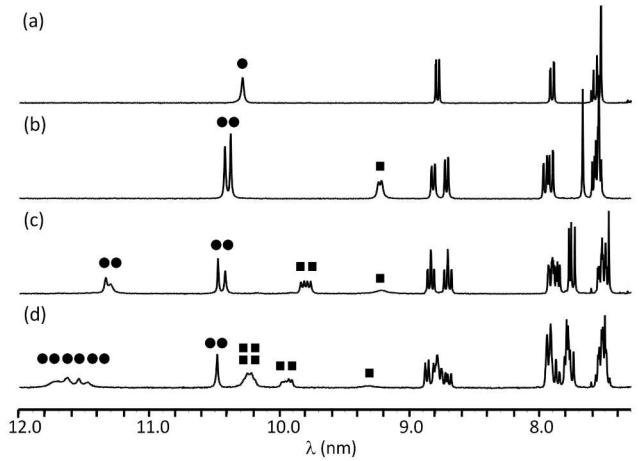



**Fig. 5**. Part of the 300 MHz $^1$H NMR spectra of **1**-**4** at 25°C in $CDCl_3$. (a) **1**; (b) **2**; (c) **3**; (d) **4**. Circles and squares indicate aromatic and aliphatic NH resonances, respectively.

The $^1$H NMR spectra of **1**-**4** are presented in Fig. 5. Resonances assigned to aromatic protons show no significant dependence on oligomer length. In agreement with the structure of **3**, this hints at the absence of intramolecular aromatic stacking in solution. Resonances assigned to aromatic amide protons are found near 10.5 ppm for **1** and **2** which are too short to form pseudo-α-turns as those observed in the central part of the structure of **3**. In contrast, in the spectra of **3** and **4**, two aromatic amide resonances are found at 10.5 ppm, and all the others near 11.5 ppm. This 1 ppm difference is consistent with their participation to hydrogen bonded pseudo-α-turns. Chemical shift values of aliphatic amide protons are also in agreement with the solid state structure. The spectrum of **2** shows one resonance near 9.3 ppm, whereas the spectra of **3** and **4** show one resonance near 9.3 ppm and all the others above 9.8 ppm, consistent with their involvement in hydrogen bonded γ-turns. The two hydrogen bonded aliphatic aromatic resonances of **3** have similar $^3J_{(NH\ C\alpha H)}$ coupling constants (7.5 and 7.6 Hz) in contrast with the different dihedral angles about the $C^{\alpha}$H-NH bond in the solid state. This discrepancy may reflect different γ-turn conformations under rapid exchange on the NMR time scale in solution. Overall, the consistency of NMR data with the solid state structure suggest that longer oligomers also behave like **3**.

## Conclusions

In conclusion, structural investigations and in particular a rare crystal structure allowed us to give a detailed description of the conformational behaviour of $(\mathbf{LQ})_n$ oligomers. The overall rigidity of the backbone is locally enhanced by hydrogen-bonded α- and γ-turns in sequences where n>2, resulting in a zig-zag tape structure that much differs from the helix previously observed in $(\mathbf{LQ_2})_n$ oligomers in which dominant **Q** units dictate their folding behaviour to leucine monomers. Nevertheless, the few degrees of freedom that are left in the tape structure, including wiggling of the terminal units and different conformations of the γ-turns lead to an only partially folded state that does not have long range order. The regular arrangement of LQ blocks, that nevertheless possess an internal degree of freedom at the γ-turns constitute a new motif that may serve as a model for some of the many ill-folded oligomers whose structure has not been elucidated.

This work was supported by the International Training Pro-gram of JSPS (Predoctoral Fellowship to MK). We thank Dr. Dr. Ito (Institute of Biomaterials and Bioengineering, Tokyo Medical and Dental University) for assistance with MS measurements.

## Notes and references

[a] Department of Chemistry, Faculty of Science, Ochanomizu University, 2-1-1 Otsuka, Bunkyo-ku, Tokyo 112-8610, Japan
[b] Univ. Bordeaux, CBMN (UMR 5248), Institut Européen de Chimie Biologie, 2 rue Escarpit 33600 Pessac, France
[c] CNRS, CBMN (UMR 5248), 2 rue Escarpit 33600 Pessac, France
[d] Center for Analytical Instrumentation, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

† Throughout the manuscript sequences abbreviated $(\mathbf{LQ})_n$ possess a Boc group at the N terminus and a 2-trimethylsilyl-ester at the C terminus.

‡ Crystal data: ($C_{90}H_{122}N_{12}O_{15}Si$), $M$ = 1640.08, $T$ = 120 K, Monoclinic, space group $P2_1/n$, $a$ = 16.9924(7), $b$ = 27.8784(11), $c$ = 26.0023(10) Å, β = 97.315(3)°, $V$ = 12217.6(8) Å$^3$, $D_c$ =0.892 g cm$^{-3}$, $Z$ = 4, 56324 reflections measured, 20816 independent reflections ($R_{int}$ =0.0904), Refinement on $F^2$ against all reflections. The weighted $R$-factor w$R$ and goodness of fit GOF are based on $F^2$, GOF = 0.845, $R[F^2 > 2\sigma(F^2)]$ = 0.1210, w$R_2$ = 0.3059. CCDC # 1003833.

# 8-amino-2-quinolinecarboxylic acids are δ-amino acids and thus are equivalent to a dipeptide.

Electronic Supplementary Information (ESI) available: synthetic schemes, synthetic procedures, characterization of new compounds, crystallographic procedures, variable temperature UV and CD spectra. See DOI: 10.1039/c000000x/

1 G. Guichard and I. Huc, *Chem. Commun.*, 2011, **47**, 5933.
2 (a) G. S. Hanan, J.-M. Lehn, N. Krystsaka and J. Fisher, *J. Chem. Soc., Chem. Commun.*, 1995, 765; (b) J. C. Nelson, J. G. Saven, J. S. Moore and P. G. Wolynes, *Science*, 1997, **277**, 1793; (c) J. Zhu, R. D. Parra, H. Q. Zeng, E. Skrzypczak-Jankun, X. C. Zeng and B. Gong, *J. Am. Chem. Soc.*, 2000, **122**, 4219; (d) V. Berl, I. Huc, R. G.

Khoury and J.-M. Lehn, *Chem. Eur. J*., 2001, **7**, 2798; (e) D. H. Appella, L. A. Chris-tianson, I. L. Karle, D. R. Powell and S. H. Gellman, *J. Am. Chem. Soc.*, 1996, **118**, 13071; (f) I. M. Mándity, E. Wéber, T. A. Martinek, G. Olajos, G. K. Tóth, E. Vass and F. Fülöp, *Angew. Chem., Int. Ed.*, 2009, **48**, 2171.

3 For reviews, see: (a) A. Roy, P. Prabhakaran, P. K. Baruah and G. J. Sanjayan, *Chem. Commun*., 2011, **47**, 11593; (b) W. S. Horne and S. H. Gellman, *Acc. Chem. Res*., 2008, **48**, 1399; (c) T. A. Martinek and F. Fülöp, *Chem. Soc. Rev.*, 2011, **41**, 687.

4 (a) J. A. Crapster, I. A. Guzei and H. E. Blackwell, *Angew. Chem., Int. Ed.*, 2013, **52**, 5079; (b) C. Caumes, O. Roy, S. Faure and C. Taillefumier, *J. Am. Chem. Soc.*, 2012, **134**, 9553; (c) J. S. Laursen, J. Engel-Andreasen, P. Fristrup, P. Harris and C. O. Olsen, *J. Am. Chem. Soc.* 2013, **135**, 2835; (d) J. R. Stringer, J. A. Crapster, I. A. Guzei and H. E. Blackwell, *J. Am. Chem. Soc.*, 2011, **133**, 15559; (e) C. W. Wu, K. Kirshenbaum, T. J. Sanborn, J. A. Patch, K. Huang, K. A. Dill, R. N. Zuckermann and A. E. Barron, *J. Am. Chem. Soc.*, 2003, **125**, 13525; (f) N. H. Shah, G. L. Butterfoss, K. Nguyen, B. Yoo, R. Bonneau, D. L. Rabenstein and K. Kirshenbaum, *J. Am. Chem. Soc.*, 2008, **130**, 16622.

5 M. T. Stone, J. M. Heemstra and J. S. Moore, *Acc. Chem. Res.*, 2006, **39**, 11.

6 A. Glättli, X. Daura, D. Seebach and W. F. van Gunsteren, *J. Am. Chem. Soc.*, 2002, **124**, 12972.

7 (a) J. P. Snyder, A. S. Lakdawala and M. J. Kelso, *J. Am. Chem. Soc.*, 2003, **125**, 632; (b) C. R. Landis, L. L. Luck and J. M. Wright, *J. Magn. Reson. Series B*, 1995, **109**, 44.

8 (a) E. E. Baquero, W. H. James, S. H. Choi, S. H. Gellman and T. S. Zwier, *J. Am. Chem. Soc.*, 2008, **130**, 4795; (b) R. Kusaka, D. Zhang, P. S. Walsh, J. R. Gord, B. F. Fisher, S. H. Gellman and T. S. Zwier, *J. Phys. Chem. A*, 2013, **117**, 10847.

9 (a) P. G. Vasudev, S. Chatterjee, N. Shamala and P. Balaram, *Chem. Rev.*, 2011, **111**, 657; (b) A. Hayen, M. A. Schmitt, F. N. Ngassa, K. A. Thomasson and S. H. Gellman, *Angew. Chem., Int. Ed.*, 2004, **43**, 505; (c) G. V. Sharma, K. R. Reddy, P. R. Krishna, A. R. Sankar, K. Narsimulu, S. K. Kumar, P. Jayaprakash, B. Jagannadh and A. C. Kunwar, *J. Am. Chem. Soc.*, 2003, **125**, 13670; (d) A. Hetényi, G. K. Tóth, C. Somlai, E. Vass, T. A. Martinek and F. Fülöp, *Chem.–Eur. J.* 2009, **15**, 10736; (e) P. Claudon, A. Violette, K. Lamour, M. Decossas, S. Fournel, B. Heurtault, J. Godet, Y. Mély, B. Jamart-Grégoire, M.-C. Averlant-Petit, J.-P. Briand, G. Duportail, H. Monteil and G. Guichard, *Angew. Chem., Int. Ed.*, 2010, **49**, 333; (f) L. Guo, A. M. Almeida, W. Zhang, A. G. Reidenbach, S. H. Choi, I. A. Guzei and S. H. Gellman, *J. Am. Chem. Soc.*, 2010, **132**, 7868; (g) N. Pendem, Y. Reddy Nelli, C. Douat, L. Fischer, M. Laguerre, R. Ennifar, B. Kauf-fmann and G. Guichard, *Angew. Chem., Int. Ed.*, 2013, **52**, 4147.

10 (a) P. Prabhakaran, S. S. Kale, V. G. Puranik, P. R. Rajamohanan, O. Chetina, J. A. Howard, H. J. Hofmann and G. J. Sanjayan, *J. Am. Chem. Soc.*, 2008, **130**, 17743; (b) D. Srinivas, R. Gonnade, S. Ravindranathan and G. J. Sanjayan, *J. Org. Chem.*, 2007, **72**, 7022; (c) A. Roy, P. Prabhakaran, P. K. Baruah and G. J. Sanjayan, *Chem. Commun.*, 2011, **47**, 11593; (d) V. V. E. Ramesh, K. N. Vijayadas, S. Dhokale, R. G. Gonnade, P. R. Rajamohanan and G. J. Sanjayan, *Org. Biomol. Chem.*, 2013, **11**, 7072.

11 N. Delsuc, F. Godde, B. Kauffmann, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2007, **129**, 11348.

12 (a) H. Jiang, J.-M. Léger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448; (b) C. Dolain, A. Grélard, M. Laguerre, H. Jiang, V. Maurizot and I. Huc, *Chem. Eur. J.*, 2005, **11**, 6135. (c) N. Delsuc, T. Kawanami, J. Lefeuvre, A. Shundo, H. Ihara, M. Takafuji and I. Huc, *ChemPhysChem*, 2008, **9**, 1882; (d) T. Qi, V. Maurizot, H. Noguchi, T. Charoenraks, B. Kauffmann, M. Takafuji, H. Ihara and I. Huc, *Chem. Commun.*, 2012, **48**, 6337.

13 (a) D. Sánchez-García, B. Kauffmann, T. Kawanami, H. Ihara, M. Takafuji, M.-H. Delville and I. Huc, *J. Am. Chem. Soc.*, 2009, **131**, 8642; (b) B. Baptiste, C. Douat-Casassus, K. Laxmi-Reddy, F. Godde and I. Huc, *J. Org. Chem.*, 2010, **75**, 7175; (c) N. Delsuc, L. Poniman, J.-M. Léger and I. Huc, *Tetrahedron*, 2012, **68**, 4464.

14 M. Kudo, V. Maurizot, B. Kauffmann, A. Tanatani and I. Huc, *J. Am. Chem. Soc.*, 2013, **135**, 9628.

15 T. Qi, T. Deschrijver, I. Huc and *Nat. Protoc.*, 2013, **8**, 693.

16 G.Lautrette, B. Kauffmann, Y. Ferrand, C. Aube, N. Chandramouli, D. Dubreuil and I. Huc, *Angew. Chem., Int. Ed.*, 2013, **52**, 11517.

17 (a) Z. Hayouka, D. E. Mortenson, D. F. Kreitler, B. Weisblum, K. T. Forest and S. H. Gellman, *J. Am. Chem. Soc.*, 2013, **135**, 15738; (b) M. Lee, J. Shim, P. Kang, I. A. Guzei and S. H. Choi, *Angew. Chem., Int. Ed.*, 2013, **52**, 12564.

18 (a) J. Fremaux, C. Dolain, B. Kauffmann, J. Clayden and G. Guichard, *Chem.Commun.*, 2013, **49**, 7415; (b) R. V. Nair, S. Kheria, S. Rayavarapu, A. S. Kotmale, B. Jagadeesh, R. G. Gonnade, V. G. Puranik, P. R. Rajamohanan and G. J. Sanjayan, *J. Am. Chem. Soc.*, 2013, **135**, 11477.

19 (a) J. J. Gassensmith, J. M. Baumes, J. Eberhard and B. D. Smith, *Chem. Commun.*, 2009, 2517; (b) J. J. Gassensmith, J. M. Baumes and B. D. Smith, Chem. Commun., 2009, 6329; (c) Y. Ferrand, Q. Gan, B. Kauffmann, H. Jiang and I. Huc, *Angew. Chem., Int. Ed.*, 2011, **50**, 7572; (d) E. A. Kataev, G. V. Kolesnikov, R. Arnold, H. V. Lavrov and V. N. Khrustalev, *Chem. Eur. J.*, 2013, **19**, 3710.

20 V. Pavone, G. Gaeta, A. Lombardi, F. Nastri, O. Maglio, C. Isernia and M. Saviano, *Biopolymers*, 1996, **38**, 705.

