

Structure Elucidation of Host–Guest Complexes of Tartaric and Malic Acids by Quasi-Racemic Crystallography**

Guillaume Lautrette, Brice Kauffmann, Yann Ferrand, Christophe Aube, Nagula Chandramouli, Didier Dubreuil, and Ivan Huc*

The propensity of racemic solutions of organic molecules to frequently produce racemic crystals and to only rarely resolve into crystals that contain exclusively one enantiomer (conglomerates) has been known for a long time.^[1] More than a decade of foldamer chemistry has provided strong empirical evidence that this observation also holds true for helical aromatic foldamers. A vast number of racemic crystals that contain both right-handed (*P*) and left-handed (*M*) helices have been reported,^[2,3] whereas conglomerates are extremely rare.^[4] When aromatic-foldamer helices are exclusively one-handed through the introduction of chiral groups that control helix handedness, our experience is that crystals do not grow well and X-ray quality crystals that allow structure resolution are rarely obtained.^[5] In contrast, the corresponding racemates often crystallize readily. Thus, racemic mixtures have purposely been used to assign absolute helix handedness of aromatic foldamers induced by covalently appended chiral residues^[6] or by noncovalently bound chiral guests.^[7] Racemic crystallography has also been used to solve the structure of some proteins that can be produced by chemical synthesis and that were shown to crystallize more readily as a racemic pair than as a single enantiomer.^[8]

As an extension to racemic crystallography, quasi-racemates are sometimes found to co-crystallize, as was originally described by Pasteur for malate and tartrate salts.^[9] A quasi-racemic crystal comprises a pair of molecules, the structures of which are almost, but not exactly, mirror images.^[9,10] Quasi-racemic crystals of small proteins have recently been produced, allowing the resolution of the three-dimensional structure of a slightly modified protein without having to synthesize its enantiomer when the enantiomer of the prototypical protein is available.^[11] In helical aromatic foldamers, quasi-racemic crystals have been encountered as well, with a diastereomeric pair, a *P* helix and an *M* helix with the same chiral terminal residue, co-crystallizing in an almost perfectly centrosymmetric fashion.^[6,12,13]

Herein, we describe the use of quasi-racemic crystallography as an efficient method to simultaneously elucidate the structures of diastereomeric complexes of tartaric and malic acids, the same molecules that allowed the first description of a quasi-racemic crystal by Pasteur 160 years ago,^[9] encapsulated into a helically folded host. We took advantage of the high propensity of enantiomeric *P* and *M* helices to co-crystallize, and demonstrated that co-crystals still form even when the guests introduced in the cavities are not mirror images.

We have previously shown that the aromatic oligoamide sequence **1** (Figure 1) folds into a stable helix with a cavity that binds to tartaric acid (**2**) with high affinity ($K_a = 5300 \text{ L mol}^{-1}$ in $\text{CDCl}_3/[\text{D}_6]\text{DMSO} = 9:1$ (v/v)) and high diastereoselectivity: *L*-tartaric acid is preferentially bound in *M*-**1** and *D*-tartaric acid in *P*-**1**, with a diastereomeric excess $de > 99\%$.^[7a] Upon mixing racemic tartaric acid with **1**, a racemic solution of *P*-**1**⊃*D*-**2** and *M*-**1**⊃*L*-**2** complexes was

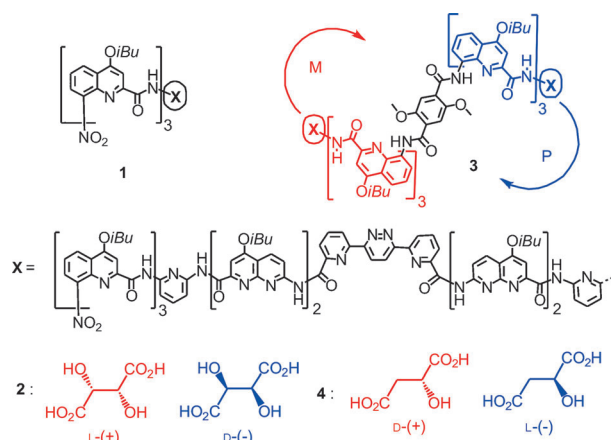


Figure 1. Formula of hosts **1** and **3** together with structures of the isomers of tartaric and malic acids.

[*] G. Lautrette, Dr. Y. Ferrand, Dr. N. Chandramouli, Dr. I. Huc
Université de Bordeaux, CBMN, UMR5248
Institut Européen de Chimie et Biologie
2 rue Escarpit, 33600 Pessac (France)
and
CNRS, CBMN, UMR5248 (France)
E-mail: i.huc@iecb.u-bordeaux.fr

Dr. B. Kauffmann
Université de Bordeaux, UMS3033
Institut Européen de Chimie et Biologie (IECB)
2 rue Escarpit, 33600 Pessac (France)
and
CNRS, IECB; UMS3033 (France)
and
INSERM, IECB, UMS3033 (France)

Dr. C. Aube, Prof. D. Dubreuil
Université de Nantes, CEISAM, UMR6230
Faculté des Sciences et des Techniques
2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3 (France)
and
CNRS, CEISAM, UMR6230 (France)

[**] This work was supported by the Comité Interprofessionnel du Vin de Bordeaux (CIVB) (predoctoral fellowship to G.L.) and by ANR grant (ANR-09-BLAN-0082-01 and predoctoral fellowship to C.A.). We thank Dr. Maurizot for providing terephthalic acid derivatives.

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/anie.201305364>.

formed. This racemate crystallized in the centrosymmetric $P21/n$ monoclinic space group, and its structure was solved, showing an array of hydrogen bonds between the acid and hydroxy functions of **2** and the inner wall of **1**, consistent with data obtained in solution.^[7a] In another experiment, enantiomerically pure tartaric acid L-**2** was mixed with **1** and two diastereomeric complexes P -**1DL-2** and M -**1DL-2** were formed in solution.^[14] Thus, the less favored complex P -**1DL-2** still shows enough stability to form quantitatively at mM concentration, as a kinetic supramolecular product (see below). We called the P -**1DL-2** complex the mismatching complex as opposed to the M -**1DL-2** complex, which has matching host-guest stereochemistry. Intrigued by this phenomenon, we sought for structural information about the mismatching complex to shed light on the better complementarity and higher stability of the matching complex.

The P -**1DL-2** plus M -**1DL-2** mixture may be considered to be a quasi-racemate inasmuch as the two complexes are expected to be mirror images for most part (the large helical host) at the exclusion of the stereochemistry and orientation of the small tartaric acid guest. As such, this mixture represented a candidate to grow quasi-racemic crystals. Unfortunately, the mismatching complex is a transient species. As the P and M helices equilibrate through an unfolding–refolding process, the mismatching complex converts into the more stable matching complex.^[14] Attempts to grow quasi-racemic crystals before the mismatching complex disappeared failed in our hands.

In order to capture the mismatching complex, we considered an original approach to lock helix handedness, making use of an earlier design that allows to introduce a helix reversal within a helical aromatic-foldamer sequence.^[15] Oligomer **3** consists of two segments, each derived from the sequence of host **1** connected at their termini by a terephthalic acid unit (Figure 1, see also Figure 2). As described in detail in the Supporting Information, the preparation of **3** required the development of a new synthetic approach of the host sequences to enable desymmetrization and distinct functionalization at each end. The intramolecular hydrogen-bonding motif at the central terephthalamide is such that the two host segments on each side of the central unit have opposite handedness.^[15] The overall structure of **3** thus possesses two independent binding sites for tartaric acid, one in a P helix, the other in an M helix. On average, **3** possesses an inversion center, as reflected by the multiplicity of ¹H NMR signals (13 amide resonances, Figure 3a and Figure S1). The presence of both P and M segments were expected to facilitate its crystallization. This was indeed the case and the structure of **3** was solved in the $P1$ space group with a single molecule in the unit cell (Figure 4a). It is worth noting that while the two enantiomeric P - and M -helical segments are important for crystallization, the resulting structure does not necessarily have a crystallographic inversion center or plane of symmetry, as was also observed in some racemic protein crystals.^[16] The propensity of the P and M helices to co-crystallize may not be directly related to the availability of centrosymmetric space groups.

All ¹H NMR titrations of **3** were carried out in CDCl₃, using [D₆]DMSO (1%) to dissolve the guest in the stock

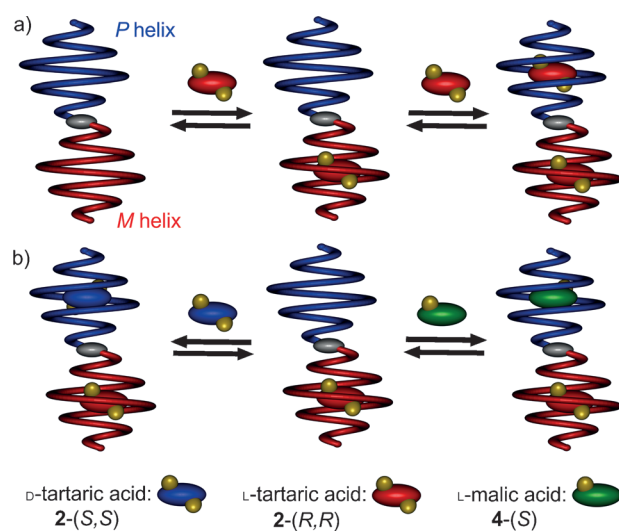


Figure 2. Schematic representation of the encapsulation of a) two molecules of the L-enantiomer of tartaric acid and b) a racemic mixture of D- and L-tartaric acid (left equilibrium) and a mixture of L-tartaric acid and L-malic acid (right equilibrium) in host **3**, which possesses two cavities of opposite handedness. P and M handedness are denoted in blue and red, respectively. The inversion center controlling the stereochemistry of **3** is colored in light grey. The golden balls stand for the hydroxy groups of the guests. Red-in-blue and blue-in-red complexes are mismatching complexes.

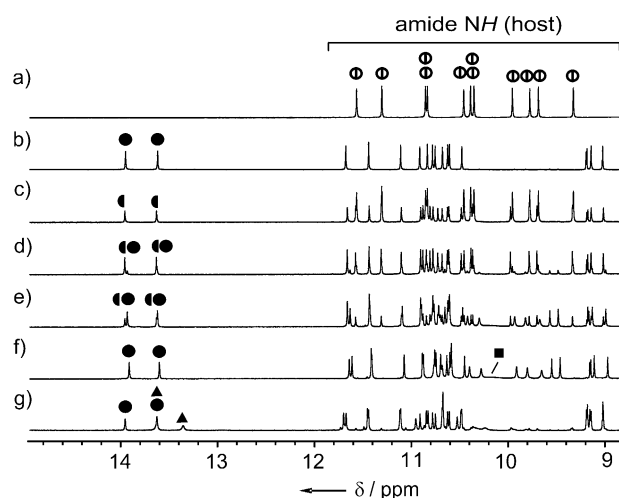


Figure 3. Parts of the ¹H NMR spectra (700 MHz) of **3** (0.5 mM in CDCl₃; [D₆]DMSO = 99:1 (v/v)) at 298 K in the presence of a) 0 equiv of guest; b) 2 equiv of D/L-**2**; c) 0.5 equiv; d) 1 equiv; e) 1.5 equiv; and f) 2 equiv of L-**2**; g) 1 equiv of L-**2** and 3 equiv of L-**4**. For $3 \supset (2-D; 2-L)$ (b) signals of the carboxylic acid protons of **2** in the matching helices are denoted with full black circles. For $3 \supset (2-L)_2$ (c–g) signals of the carboxylic acid protons of L-**2** in the M helix (matching) are marked with half-black circles when the P helix is empty, and with black circles when the P helix also hosts a guest; signals of the carboxylic acid protons of L-**2** in the P helix (mismatching) are marked with a black square. It should be noted that the two mismatching acid resonances are at coalescence at this temperature. For $3 \supset (2-L; 4-L)$, signals of the carboxylic acid protons of L-**2** in the M helix (matching) are marked with black circles and carboxylic acid protons of L-**4** in the P helix (matching) are denoted with black triangles.

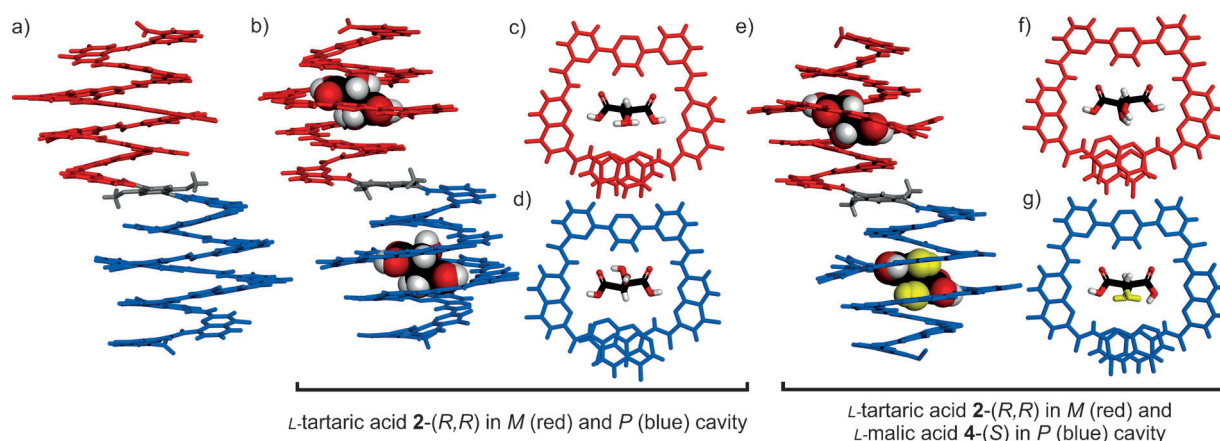


Figure 4. a, b, e) Side view of the solid-state structure in the tube representation of a) **3**, b) $3\supset(2-L)_2$, and e) $3\supset(2-L;4-L)$. c, d) Slice views of $3\supset(2-L)_2$ with c) the left-handed cavity (*M*) showing guest **2-L** with its hydroxy groups pointing toward the naphthyridines, and d) the right-handed cavity (*P*) showing guest **2-L** with its hydroxy groups pointing toward the pyridazine. f, g) Slice views of $3\supset(2-L;4-L)$ with f) the left-handed cavity (*M*) showing guest **2-L**, and g) the right-handed cavity (*P*) showing guest **4-L** with its hydroxy group pointing toward the naphthyridines. The position of malic acid in the *P* helix is disordered, leading to two partially occupied positions of its hydroxy oxygen atom (e, f; yellow). *P* and *M* helices are denoted in blue and red, respectively. The central terephthalic acid unit is colored in grey. Tartaric and malic acid are represented as CPK (b, e) or tube (c, d, f, g) structures. Isobutyl groups and included solvent molecules are not shown for clarity.

solution. As observed previously for **1**,^[7] binding and release of the guest was slow on the NMR time scale at 298 K. A titration of **3** using D/L-**2** led to the emergence of a new set of signals corresponding to a centrosymmetric 1:2 complex ($3\supset(2-D;2-L)$) in which both cavities of **3** are filled with the matching guest (Figure 2 b, left). Evidence for this assignment stemmed from the symmetry of the final complex (Figure S1, 13 amide resonances), which is the same as that of the empty host, from the stoichiometry assigned by integrating the host and guest signals in the complex, and from the chemical shift values of the carboxylic acid protons of the guest at 13.5–14 ppm, which are characteristic of a matching complex.^[7] During the titration, the formation of this 1:2 complex is preceded by a 1:1 matching complex (i.e. $3\supset 2-D$ or $3\supset 2-L$) having one empty cavity and lacking any symmetry (26 amide signals, Figure S2). Interestingly, NMR signals of the guest in both 1:1 and 1:2 complexes show that its own symmetry is broken upon encapsulation, regardless of the symmetry of the overall structure. Distinct signals are observed for each carboxylic acid proton because each cavity of **3** is itself dissymmetric.

A second titration experiment consisted in the addition of a single enantiomer of tartaric acid **2-L** to host **3** (Figure 2 a). The addition of 0.5 equivalents of **2-L** led to the appearance of a new set of signals with twice the number of resonances compared to **3** alone, corresponding to a 1:1 complex in which **2-L** was fully encapsulated in the *M* cavity and not in the *P* cavity. Again, the two sharp signals at 13.96 and 13.63 ppm are those of the hydrogen-bonded acid protons of the guest in a matching configuration and surrounded by a dissymmetric environment (Figure 3 c–e, half-circles). Upon further addition of **2-L**, resonances corresponding to $3\supset 2-L$ were replaced by another set of signals corresponding to $3\supset(2-L)_2$ with both a matching and a mismatching complex. Using ¹³C and ¹⁵N-heteronuclear single-quantum coherence (HSQC) spectroscopy, we identified the signals of the carboxylic acid protons

of the mismatching **2-L** in the *P* helix of **3** (Figure 3, squares) to be broad and to show less upfield-shifted resonances at 10.20 and 10.13 ppm at 273 K (Figure S4), thus indicating weaker hydrogen bonding than in matching complexes.

Unlike $3\supset(2-D;2-L)$, $3\supset(2-L)_2$ is not centrosymmetric. Yet these two 1:2 complexes differ only by the configuration of one of the guests inside one of the two helical cavities of **3**. Thus, $3\supset(2-L)_2$ appeared to be a good candidate for a quasi-racemic crystallographic analysis that would finally allow the elucidation of the structure of a mismatching complex. Indeed, single crystals of $3\supset(2-L)_2$ grew from the slow diffusion of hexane into chloroform solutions and the structure was solved in the *P1* space group. In both the *P* and *M* cavities, **2-L** is positioned perpendicular to the helical axis and adopts a conformation with the two acid and the two hydroxy groups in *trans* and *gauche* position, respectively (Figure 4 b–d). In both mismatching and matching complexes, carboxylic acid groups are doubly hydrogen bonded to the neighboring 7-aminonaphthyridine (*P* helix $d_{OH\cdots N} = 2.67$ Å, $d_{NH\cdots OC} = 3.11$ Å; *M* helix $d_{OH\cdots N} = 2.79$ Å, $d_{NH\cdots OC} = 2.95$ Å) and also hydrogen bonded to the aromatic pyridazine protons (*P* helix $d_{CH\cdots OC} = 3.33$ Å; *M* helix $d_{CH\cdots OC} = 3.11$ Å). Tartaric acid adopts an *anti* arrangement of the O=C–C–OH bonds in the matching configuration, whereas the *syn* conformer is observed in the mismatching configuration. Overall, the guest **2-L** in its mismatching configuration has to shift its position in the cavity, moving away from the pyridazine central unit in order to accommodate its hydroxy groups in between. In contrast, in the matching complex, the CH groups of the guest point toward the pyridazine.

A step further in our study consisted in exploiting the centrosymmetric nature of **3** to serve the crystallographic analysis of the complexes it may form with two different guests. Malic acid (**4**; Figure 1) also binds to **1**, though with lower affinity ($K_a = 70$ L mol⁻¹ in CDCl₃/[D₆]DMSO = 9:1 (v/v)) and poorer diastereoselectivity ($de = 52\%$) than tartaric

acid. A consequence of the poor diastereoselectivity is the coexistence of matching and mismatching complexes, even when using racemic malic acid, which has prevented crystallization and structural analysis until now. L-malic acid (**L-4**) differs from L-tartaric acid (**L-2**) by the lack of one hydroxy group and by the absolute configuration of its only stereogenic center. The two acids are thus quasi-enantiomers and their successful combination in quasi-racemic crystals was demonstrated by Pasteur long ago^[9a] and recently fully characterized.^[9b] We thus endeavored to crystallize both acids in complex with **3**.

The addition of **L-2** (1 equiv) and **L-4** (3 equiv) to **3** resulted in the almost quantitative formation of a single species **3**⊃(**2-L**;**4-L**) (Figure 3g, see also Figure 2b, right). This hetero-trimolecular complex features four low-field resonances that correspond to the carboxylic acid protons of encapsulated **L-2** and **L-4**, both in a matching complex. Upon cooling to 243 K, each of the two signals (at 13.63 and 13.36 ppm) assigned to carboxylic acid protons of **L-4** broadened and then split into two (Figure S7), consistent with two different orientations of **L-4** within the *P* helix of **3**.

Again, quasi-racemic crystallography proved successful to analyze the structure of encapsulated **L-4**, an otherwise elusive complex. X-ray quality crystals of **3**⊃(**2-L**;**4-L**) were obtained and its structure solved in the *P1* space group. Both **2-L** and **4-L** were found to adopt a matching *anti* conformation and to be similarly hydrogen bonded to the 7-aminonaphthyridine (**2-L** in *M* helix $d_{\text{OH}\cdots\text{N}} = 2.80 \text{ \AA}$, $d_{\text{NH}\cdots\text{OC}} = 3.08 \text{ \AA}$; **4-L** in *P* helix $d_{\text{OH}\cdots\text{N}} = 2.78 \text{ \AA}$, $d_{\text{NH}\cdots\text{OC}} = 3.08 \text{ \AA}$) and to the pyridazine protons ($d_{\text{CH}\cdots\text{OC}} = 3.25 \text{ \AA}$ and $d_{\text{CH}\cdots\text{OC}} = 3.26 \text{ \AA}$, respectively). As observed in solution, **4-L** exists in two disordered positions in the solid state, arising from its dissymmetric structure. This disorder relates to the disorder of malic acid in its quasi-racemic crystals with tartaric acid.^[9b]

Our investigation thus adds to very recent reports that quasi-racemic crystallography can be used deliberately to analyze X-ray structures that would be difficult to elucidate otherwise, including those of host-guest complexes.^[11]

Received: June 21, 2013

Published online: September 6, 2013

Keywords: crystallography · foldamers · helical structures · host-guest systems · quasi-racemic

- [1] J. Jacques, A. Collet, S. H. Wilen in *Enantiomers, racemates and resolutions*, 3rd ed., Krieger, Malabar, **1994**.
 [2] For reviews: D.-W. Zhang, X. Zhao, J.-L. Hou, Z.-T. Li, *Chem. Rev.* **2012**, *112*, 5271–5316; I. Huc, *Eur. J. Org. Chem.* **2004**, 17–29.
 [3] a) V. Berl, I. Huc, R. Khoury, M. J. Krische, J.-M. Lehn, *Nature* **2000**, *407*, 720–723; b) H. Jiang, J.-M. Léger, I. Huc, *J. Am. Chem. Soc.* **2003**, *125*, 3448–3449; c) N. Delsuc, S. Massip, J.-M. Léger, B. Kauffmann, I. Huc, *J. Am. Chem. Soc.* **2011**, *133*, 3165–3172; d) B. Gong, H. Zeng, J. Zhu, L. Yuan, Y. Han, S. Cheng, M. Furukawa, R. D. Parra, A. Y. Kovalevsky, J. L. Mills, E. Skrzypczak-Jankun, S. Martinovic, R. D. Smith, C. Zheng, T. Szyperki, X. C. Zeng, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11583–11588; e) Z.-Q. Hu, H.-Y. Hu, C.-F. Chen, *J. Org. Chem.*

- 2006**, *71*, 1131–1138; f) H.-Y. Hu, J.-F. Xiang, Y. Yang, C.-F. Chen, *Org. Lett.* **2008**, *10*, 69–72; g) A. J. Blake, P. A. Cooke, K. J. Doyle, S. Gair, N. S. Simpkins, *Tetrahedron Lett.* **1998**, *39*, 9093–9096.
 [4] To the best of our knowledge, only two examples have been reported: a) E. Ohta, H. Sato, S. Ando, A. Kosaka, T. Fukushima, D. Hashizume, M. Yamasaki, K. Hasegawa, A. Muraoka, H. Ushiyama, K. Yamashita, T. Aida, *Nat. Chem.* **2011**, *3*, 68–73; b) H. Zhao, W. Qiang Ong, F. Zhou, X. Fang, X. Chen, S. F. Y. Li, H. Su, N.-J. Cho, H. Zeng, *Chem. Sci.* **2012**, *3*, 2042–2046.
 [5] a) A. M. Kendhale, L. Poniman, Z. Dong, K. Laxmi-Reddy, B. Kauffmann, Y. Ferrand, I. Huc, *J. Org. Chem.* **2011**, *76*, 195–200; b) H. Masu, M. Sakai, K. Kishikawa, M. Yamamoto, K. Yamaguchi, S. Kohmoto, *J. Org. Chem.* **2005**, *70*, 1423–1431; c) Y. Tanaka, H. Katagiri, Y. Furusho, E. Yashima, *Angew. Chem.* **2005**, *117*, 3935–3938; *Angew. Chem. Int. Ed.* **2005**, *44*, 3867–3870.
 [6] C. Dolain, H. Jiang, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2005**, *127*, 12943–12951.
 [7] a) Y. Ferrand, A. M. Kendhale, B. Kauffmann, A. Grélard, C. Marie, V. Blot, M. Pipelier, D. Dubreuil, I. Huc, *J. Am. Chem. Soc.* **2010**, *132*, 7858–7859; b) Q. Gan, Y. Ferrand, N. Chandramouli, B. Kauffmann, C. Aube, D. Dubreuil, I. Huc, *J. Am. Chem. Soc.* **2012**, *134*, 15656–15659.
 [8] a) L. E. Zawadzke, J. M. Berg, *Proteins Struct. Funct. Genet.* **1993**, *16*, 301–305; b) C. Toniolo, C. Peggion, M. Crisma, F. Formaggio, X. Shui, D. S. Eggleston, *Nat. Struct. Biol.* **1994**, *1*, 908–914; c) B. L. Pentelute, Z. P. Gates, V. Tereshko, J. L. Dashnau, J. M. Vanderkooi, A. Kossiakoff, S. B. H. Kent, *J. Am. Chem. Soc.* **2008**, *130*, 9695–9701.
 [9] a) L. Pasteur, *Ann. Chim. Phys.* **1853**, *28*, 437–483; b) K. A. Wheeler, R. C. Grove, R. E. Davis, W. S. Kassel, *Angew. Chem.* **2008**, *120*, 84–87; *Angew. Chem. Int. Ed.* **2008**, *47*, 78–81.
 [10] a) I. L. Karle, J. Karle, *J. Am. Chem. Soc.* **1966**, *88*, 24–27; b) A. Fredga, *Bull. Soc. Chim. Fr.* **1973**, *1*, 173–182; c) A. M. Lineberry, E. T. Benjamin, R. E. Davis, W. S. Kassel, K. A. Wheeler, *Cryst. Growth Des.* **2008**, *8*, 612–619; d) M. E. Breen, S. L. Tameze, W. G. Dougherty, W. S. Kassel, K. A. Wheeler, *Cryst. Growth Des.* **2008**, *8*, 3863–3870; e) R. C. Grove, S. H. Malehorn, M. E. Breen, K. A. Wheeler, *Chem. Commun.* **2010**, *46*, 7322–7324; f) B. Calmuschi, M. Alesib, U. Englert, *Dalton Trans.* **2004**, 1852–1857.
 [11] a) D. E. Mortenson, K. A. Satyshur, I. A. Guzei, K. T. Forest, S. H. Gellman, *J. Am. Chem. Soc.* **2012**, *134*, 2473–2476; b) K. Mandal, B. L. Pentelute, D. Bang, Z. P. Gates, V. Yu. Torbeev, S. B. H. Kent, *Angew. Chem.* **2012**, *124*, 1510–1515; *Angew. Chem. Int. Ed.* **2012**, *51*, 1481–1486.
 [12] H. Jiang, C. Dolain, J.-M. Léger, H. Gornitzka, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 1034–1035; see also the asymmetric unit of compound **3b** in Ref. [5b].
 [13] These quasi-racemic structures actually showed better statistics when resolved in the centrosymmetric space group *P1̄*, in which the number of parameters to refine is twice as small as in the correct space group *P1*.
 [14] Y. Ferrand, N. Chandramouli, A. M. Kendhale, C. Aube, B. Kauffmann, A. Grélard, M. Laguerre, D. Dubreuil, I. Huc, *J. Am. Chem. Soc.* **2012**, *134*, 11282–11288.
 [15] V. Maurizot, C. Dolain, Y. Leydet, J.-M. Léger, P. Guionneau, I. Huc, *J. Am. Chem. Soc.* **2004**, *126*, 10049–10052.
 [16] a) K. Mandal, B. L. Pentelute, V. Tereshko, V. Thammavongsa, O. Schneewind, A. A. Kossiakoff, S. B. H. Kent, *Protein Sci.* **2009**, *18*, 1146–1154; b) W. R. Patterson, D. H. Anderson, W. F. DeGrado, D. Cascio, D. Eisenberg, *Protein Sci.* **1999**, *8*, 1410–1422; c) L. W. Hung, M. Kohmura, Y. Ariyoshi, S. H. Kim, *J. Mol. Biol.* **1999**, *285*, 311–321.