



HAL
open science

Iterative Evolution of an Abiotic Foldamer Sequence for the Recognition of Guest Molecules with Atomic Precision

Guillaume Lautrette, Barbara Wicher, Brice Kauffmann, Yann Ferrand, Ivan Huc

► **To cite this version:**

Guillaume Lautrette, Barbara Wicher, Brice Kauffmann, Yann Ferrand, Ivan Huc. Iterative Evolution of an Abiotic Foldamer Sequence for the Recognition of Guest Molecules with Atomic Precision. *Journal of the American Chemical Society*, 2016, 138 (32), pp.10314-10322. 10.1021/jacs.6b05913 . hal-01509476

HAL Id: hal-01509476

<https://hal.science/hal-01509476>

Submitted on 18 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Iterative evolution of an abiotic foldamer sequence for the recognition of guest molecules with atomic precision

Guillaume Lautrette,^{†,¶} Barbara Wicher,^{†,¶} Brice Kauffmann,^{†,§,‡} Yann Ferrand,^{*,†,¶} and Ivan Huc^{*,†,¶}

[†] Univ. Bordeaux, CBMN, UMR 5248, Institut Européen de Chimie Biologie, 2 rue Escarpit 33607 Pessac, France.

[¶] CNRS, CBMN, UMR 5248, Pessac, France.

[‡] Univ. Bordeaux, IECB, UMS3033, 2 rue Escarpit 33607 Pessac, France.

[§] CNRS, UMS3033, 2 rue Escarpit 33607 Pessac, France.

[‡] INSERM, US001, IECB, 2 rue Escarpit 33607 Pessac, France.

KEYWORDS: *foldamer, helical capsule, molecular recognition, malic acid, tartaric acid, x-ray crystallography, NMR, circular dichroism.*

ABSTRACT: A synthetic helical aromatic oligoamide foldamer receptor with high affinity and selectivity for tartaric acid was subjected to a structure-based evolution of its sequence *via* mutations, additions and deletions of monomers to produce a new receptor having high affinity and selectivity for malic acid, a guest that differs from tartaric acid by a single oxygen atom. Seven iteratively modified sequences were synthesized. Detailed structural investigations of host-guest complexes were carried out systematically to guide the design of the next generation. A first outcome was a reversal of selectivity of the receptors, with a starting preference for tartaric acid over malic acid of over 10^2 and an ending sequence showing a preference for malic acid over tartaric acid of over 10^2 . Another outcome was a very strong enhancement of the affinity for malic acid, despite the fact it has fewer recognition features for binding through polar interactions such as hydrogen bonds. Such level of discrimination between very much resembling guests exemplifies the amenability of foldamers to outstanding achievements in molecular recognition. Altogether, our results demonstrate the viability of a rational receptor design approach that exploits the modularity of foldamer sequences and, in the case of aromatic amide foldamers, their amenability to structural elucidation, their relative ease of synthesis and the predictability of their structure.

INTRODUCTION

Folding of polymeric molecular strings into well-defined structures is the process nature has selected to produce molecular shapes from which sophisticated functions derive. For example, folding allows to precisely control the positioning of functional groups in space and achieve efficient and selective molecular recognition between biopolymers as well as between biopolymers and smaller molecules. Inspired by this, chemists have developed non-natural backbones that are also able to fold into well-defined objects in solution, giving rise to a class of molecular architectures named “foldamers”.¹ Depending on whether they are chemically related to, or distant from, biopolymers, foldamers may be considered to be “biotic” or “abiotic”, respectively. Because folding is a reversible process,² foldamers are dynamic objects. They bring novel, bioinspired solutions to the design of synthetic receptors³ that contrast with previous approaches that relied on the formation of more rigid structures such as macropolycycles to generate arrays of functional groups converging towards a binding cavity.⁴ In recent years, a number of helically folded abiotic oligomers have been shown to bind guests within a central cavity.⁵ When the helix diameter is reduced at both extremities, the guest may be completely surrounded by the helical host and isolated from the solvent,⁶ allowing guest capture and release only through dynamic structural changes of the helix.⁷ Foldamers have also been decorated with proteinogenic side chains to achieve molecular recognition at their surface, including protein⁸ and nucleic acid recognition.⁹

In this context, aromatic foldamers in general and aromatic amide foldamers in particular^{3a,10} constitute a promising family as their structures, and thus the positioning of recognition groups in space, are predictable; their synthesis is relatively easy; and they are often amenable to full structural elucidation *via* x-ray crystallography.

Common to biopolymers and foldamers is their inherent modularity: they consist of a primary sequence of monomers whose order and nature encode the information necessary to generate the tridimensional folded structure. Sequence, structure and properties may be fine-tuned through trivial operations such as additions, deletions and mutation of monomers, or swapping of segments. Since foldamer chemical synthesis generally consists of repetitive couplings between monomers or blocks, one may envisage to quickly produce sequence variants in order to improve properties in a sort of evolutionary process reminiscent of biopolymer evolution, albeit in the absence of any replication machinery.

In a proof-of-concept study, we recently showed that an evolutionary approach could be applied to abiotic foldamers. Based on detailed structural analyses of host-guest complex structures, an abiotic receptor showing unprecedented binding to β -D-fructopyranose was produced in just a few iterations.¹¹ Starting with a good, but poorly selective, first principle design fructose receptor, successive generations were conceived not to improve interactions with the target guest but to preserve these interac-

tions while filling space around the guest so as to disfavor binding of other guests. This negative design strategy led to exquisite selectivity; even the structurally related α -D-mannopyranose was discriminated from β -D-fructopyranose by a factor higher than 10^2 whilst most other monosaccharides were discriminated by factors over 10^4 .

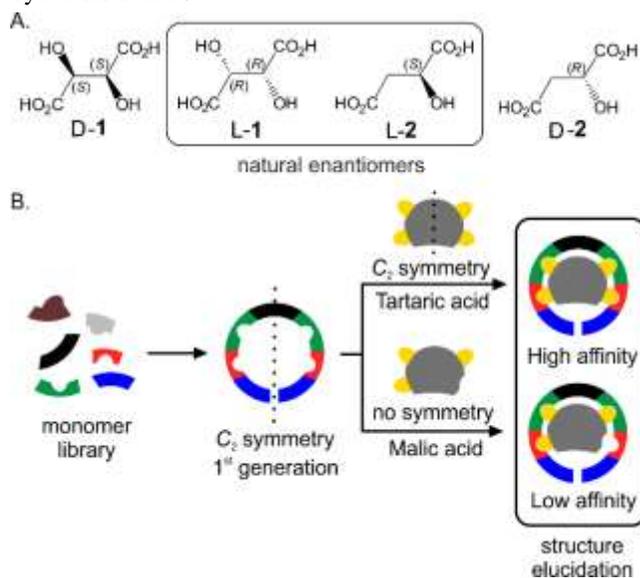


Figure 1. (A) Formula and stereochemical configurations of D/L tartaric acid **1** and D/L malic acid **2**. (B) Schematic representation of the preparation of the first generation of C_2 symmetrical receptor for the selective recognition of **1**.¹²

To further validate and extend the scope of structure-based iterative foldamer evolution, we challenged ourselves to achieve a guest selectivity reversal. Specifically, we started from a helically folded aromatic amide sequence previously shown to bind tartaric acid **1** (Figure 1A) with high affinity, complete diastereoselectivity and good selectivity towards others organic acids such as malic acid **2**,¹² and engaged an iterative evolution to convert that sequence into a receptor that binds malic acid much better than tartaric acid. Both acids occur in wine, tartaric acid being far more abundant, and exist in nature as single enantiomers having opposite stereochemistry (L-**1** (*R,R*) and L-**2** (*S*), Figure 1A). They differ *only* by a single oxygen atom, resulting in a loss of symmetry in the structure of **2**. Their discrimination by molecular recognition thus requires atomic precision and represent a real challenge in receptor design. Several reports on artificial receptors for tartrate and malate exist in the literature,¹³ most often showing a preference for tartrate. Yet very few receptors are able to target the protonated form of these guests.¹⁴ To our knowledge, the selective recognition of malic acid in the presence of tartaric acid has never been reported. Indeed, with its additional hydroxy group tartaric acid may engage in more hydrogen bonds with a host, leading to a stronger binding enthalpy. However, malic acid is (very slightly) smaller and one can envision a cavity in which malic acid fits whereas tartaric acid would cause some strain. In addition, their opposite stereochemistry certainly provides an opportunity for their discrimination by a diastereoselective receptor.

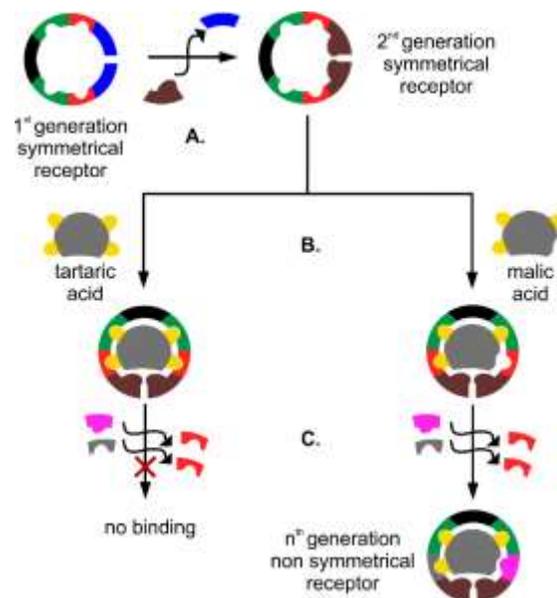


Figure 2. Schematic representation of structure-based iterative evolution of a foldamer sequence for the selective recognition of malic acid. (A) Fine tuning of the C_2 symmetrical sequence through mutations or deletions. (B) Structure elucidation (x-ray, NMR, CD) of the complexes obtained with **1** and **2**. Based on this knowledge, step (C) consists in modifying the sequence so as to sterically exclude tartaric acid through a desymmetrization and the projection of atoms in the cavity creating hindrance against tartaric acid. Iterations of steps (B) and (C) allows the sequence to evolve quickly to a highly selective receptor for L-**2**.

As shown in the following, structure directed iterative evolution proceeded in two phases to successfully yield a receptor with reverted selectivity. Starting with a two order of magnitude preference in favor of tartaric acid, the outcome was a two order magnitude preference in favor of malic acid while the strength of binding was preserved. The first phase focused on guest similarities and aimed at improving binding of the α -hydroxy acid moiety found once in malic acid and twice in tartaric acid. For this purpose, sequences that fold into C_2 symmetrical helical hosts were used and changes (mutations, deletions) were implemented at two symmetrical positions in the sequence (Figure 1B). The outcome was an improvement of tartaric acid binding while preserving a strong selectivity in its favor. The second phase focused on guest differences. The sequence optimized for the binding of an α -hydroxy acid moiety was conserved in one half of the host whilst the other half was evolved so as to best fit the methylene group of malic acid. Sequences that code for a helical host lacking any symmetry element were thus synthesized. Using structural elucidation (*i.e.* CD, NMR and x-ray) after each variation of the sequence to guide the design of the following iteration, appropriate modifications (mutation, deletion or addition) were implemented in only one half of the sequence in order to fill voids around **2** and eventually achieve atomic-scale complementarity (Figure 2). At last, benefit was taken from the opposite stereochemistry of naturally occurring L-**1** and L-**2**, and helix handedness control was implemented through the addition of chiral units at each extremity of the sequence.

Our results further establish the high amenability of helically folded aromatic oligoamide receptors to both structural elucidation and rational iterative modifications for the quick emergence of unprecedented molecular recognition properties.

RESULTS AND DISCUSSION

The starting point of our investigation was receptor **3** (Figure 3D) which tightly and selectively binds to tartaric acid **1** to form a complex in which the guest is completely surrounded by the helically folded host.¹² In sequence **3**, each monomer may code for structural and recognition features, as well as solubility properties. Specifically, monomers' size and contribution to curvature are arranged in the sequence to form a helical capsule with a wide diameter in the center and a narrow diameter at each extremity (Figure 3A,C). Recognition features are defined by the inner rim of the aromatic oligoamide strand which includes hydrogen bond donors and acceptors such as amide protons and naphthyridine endocyclic nitrogen atoms, respectively. Solubility is provided by side chains anchored on the outer rim of the strand; here isobutoxy groups provide solubility in a wide range of chlorinated, aromatic and polar organic solvents. Folding relies on well-established principles. It is directed by locally favored conformations of the aryl-amide bonds and reinforced by intramolecular π - π interactions between aromatic monomers.¹⁰ In these folded architectures, guest entrance and release occur via dynamic conformational changes of the aromatic oligoamide backbone.⁷ To match with guest **1**, capsule **3** was designed to be C_2 symmetrical. Peripheral quinoline trimers **Q**₃ serve as caps closing the cavity and also prevent the formation of multi-helical assemblies.¹⁵ The binding site is composed of monomers coding for a large diameter of the helix (**PN**₂ segments) and possesses multiple hydrogen bond donors (x 10) and acceptors (x 12) to interact with polar guests. In particular, naphthyridines were chosen for their ability to interact strongly with carboxylic acids (Figure 3B).¹⁶ In the final step of synthesis, two identical hemicapsules **Q**₃**PN**₂-NH₂ were coupled to a central diacid monomer **pyr-pyz-pyr** which codes for a large diameter. C_2 symmetrical capsule **3** was found to bind **1** with high affinity even in presence of DMSO ($K_a = 5300 \text{ L mol}^{-1}$, Table 1) and high diastereoselectivity ($de > 99\%$) in organic solvents.¹² Selectivity towards others organic acids was also found to be high. For example K_a for malic acid **2** was two orders of magnitude lower than for **1** (Table 1).

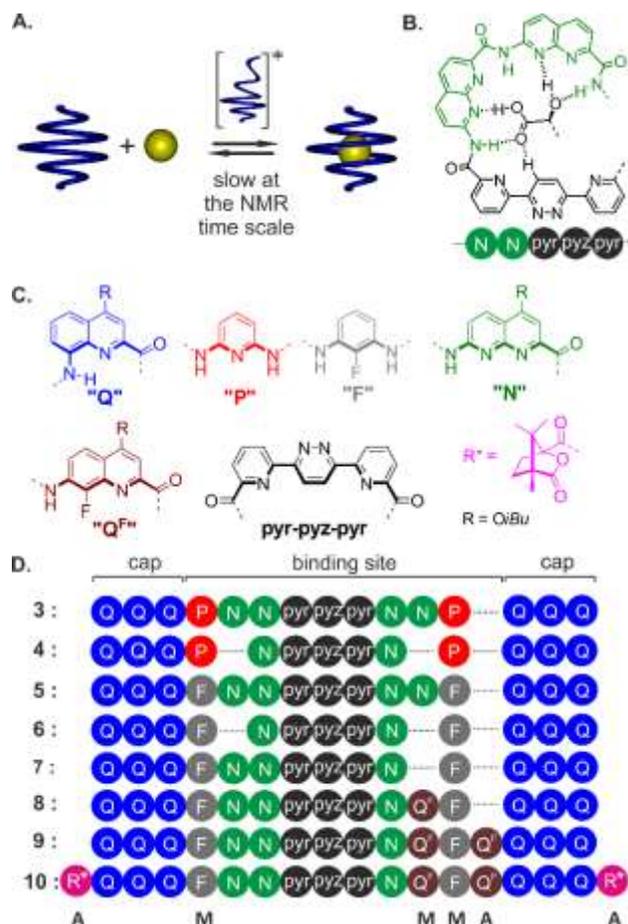


Figure 3. (A) Schematic representation of the encapsulation of a guest in a helical molecular container with a reduced diameter at both ends. (B) Hydrogen bonding pattern of a **N**₂-**pyr-pyz-pyr** aromatic oligoamide segment with the α -hydroxy acid moiety of a guest molecule. (C) Color-coded formulas and associated letters of amino acid, diamino and diacid monomers used in the receptor sequences. (D) Sequences of the 8 generations of receptors. Apart for receptor **10** both terminal **Q** units possess a nitro group in position 8. Mutations and additions of monomers are noted as **A** and **M** respectively.

Iterative improvements for tartaric acid binding. The first phase of the iterative evolution focused on improving binding of the α -hydroxy-acid moiety shared by tartaric and malic acids. During that phase modifications were implemented twice at symmetrical positions of the sequence, and the main indicator was tartaric acid binding. Using the atomic scale characterization of **3**⊃**1**,¹² we thus submitted sequence **3** to modifications in order to improve its affinity and selectivity towards guest **1**. Earlier studies¹⁷ and our own experience¹¹ showed that reducing space around the guest can have positive effects on binding. Thus, we first considered to reduce the volume of the cavity while preserving hydrogen bonds to tartaric acid which concern the central **N**₂-**pyr-pyz-pyr**-**N**₂ segment. A first modification was the removal of two naphthyridines. The central segment then becomes **PN-pyr-pyz-pyr-NP**, which has similar features to those of the original. The synthesis and characterization of the corresponding sequence **4** as well as those of all other capsules presented in this manuscript are reported in detail in the Supporting Information. Solution syntheses are all based on optimized procedures that can be scaled up.¹⁸ The double naphthyridine deletion caused a 15% reduction of the cavity volume

(Table 2) associated with a twofold decrease of the affinity for **1** ($K_a = 2600 \text{ L mol}^{-1}$). Binding of the smaller malic acid was unaltered. Crystals of **4**⊃**1** were obtained from the slow diffusion of hexane in a chloroform solution of the complex. The crystallographic structure allowed us to validate the encapsulation of **1** and the highly constricted nature of the complex (Figure S13). The volume fraction of the host cavity occupied by the guest was calculated using SURFNET¹⁹ and reaches 83%, to be compared with 70% for **3**⊃**1** and to the common benchmark of 55%¹⁷ (Table 2).

Considering this first result, we stepped back and a less drastic structural change than a deletion was envisaged. As previously for fructose,¹¹ we introduced fluorinated aromatic amino acid **F** in replacement of monomer **P** to very slightly fill the interior of the cavity with fluorine atoms without interfering with key interactions with the guest. The double mutation **P**→**F** yielded C_2 symmetrical receptor **5**, **Q**₃**FN**₂-pyr-pyz-pyr-**N**₂**FQ**₃. As for other host-guest complexes, the affinity of **5** for **1** and **2** was assessed by ¹H NMR titrations using a competitive solvent mixture (10% [D₆]-DMSO in CDCl₃) so as to lower the binding constants down to a level at which ¹H NMR titrations are accurate. Upon increasing the concentration of guest in a solution of **5**, new sets of signals emerge corresponding to host-guest complexes in slow exchange with the free capsule on the NMR time scale (Figure S4). In the case of tartaric acid D/L-**1** a single set of sharp signals corresponding to *P*-**5**⊃**1**/*M*-**5**⊃**1** was observed reflecting a completely diastereoselective encapsulation. In contrast, the addition of D/L malic acid **2** led to the emergence of two set of resonances (Figure 4c) corresponding to enantiomeric complexes *P*-**5**⊃**2**/*M*-**5**⊃**2** with matching chirality (major) and enantiomeric complexes *P*-**5**⊃**2**/*M*-**5**⊃**2** with mismatching chirality (minor).^{12,20} The proportions between the two species allow to calculate a diastereomeric excess of host-guest recognition of 80%. In matching and mismatching complexes, the number of amide resonances (six per complex) indicates that the helix backbone remains symmetrical on average despite the dissymmetrical nature of the guest, reflecting that **2** tumbles fast in the cavity, even though its capture and release are slow on the NMR time scale. The broadness of some resonances (green crosses, Figure 4C) likely reflects this tumbling.

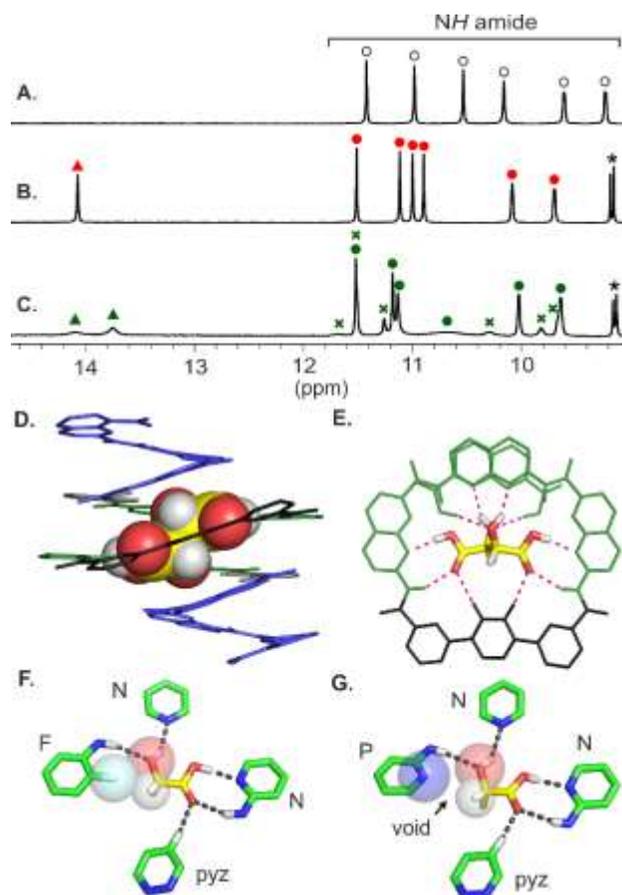


Figure 4. Excerpts of the 300 MHz ¹H NMR spectra of capsule **5** (1 mM) in 99:1 CDCl₃/[D₆]-DMSO (vol/vol) at 298 K: (A) without guest, in the presence of (B) 1.1 equiv. of D/L-**1** and (C) 1.1 equiv. of D/L-**2**. Signals of the empty host **5** are marked with empty circles. Signals of matching *P*-**5**⊃**1**/*M*-**5**⊃**1** and *P*-**5**⊃**2**/*M*-**5**⊃**2** complexes are marked with red and green circles, respectively. Green crosses indicate signals of the mismatching complexes *P*-**5**⊃**2**/*M*-**5**⊃**2**. Triangles denote carboxylic acid resonances from bound guests whereas stars denote aromatic resonances from the hosts. (D) Crystal structures of the matching complex *P*-**5**⊃**1** in tube (host) and CPK (guest) representations. (E) Top view of a slice of the same complex, both host and guest are shown in tube representation. Hydrogen bonds are shown as dashed lines. (F) Top view of the α -hydroxyacid moiety of **1** and those heterocycles of the **FN**₂-pyr-pyz- segment that interact with it (tube representation). (G) Similar view but from extracted from the previously published crystal structure of **3**⊃**1**.¹² In (F,G) the array of 5 hydrogen bonds is shown as grey dashes whereas a fluorine atom, a nitrogen atom and the α -hydroxyl group are shown in transparent CPK representation. Isobutoxy side chains and solvent molecules were omitted for clarity.

The double **P**→**F** mutation resulted in a strong enhancement of the affinities for both guests, with a 5 and 8 fold increase for **1** ($K_a = 29000 \text{ L mol}^{-1}$) and **2** ($K_a = 550 \text{ L mol}^{-1}$), respectively (Table 1). The x-ray structure of **5**⊃**1** was elucidated (Figure 4D,E) providing atomic scale information to be compared with that of **3**⊃**1** and **4**⊃**1**. As for **3**⊃**1**, tartaric acid lies flat in capsule **5** forming an array of hydrogen bonds between its acid functions and central naphthyridines on one hand and its hydroxy functions and peripheral naphthyridines on the other hand (Figure 4E). This binding mode corresponds to the complex with

matching stereochemistry *P-5*→*D-1*/*M-5*→*L-1*. The fluorine atoms of **F** monomers do not appear to be involved in direct interactions with the guest. Unexpectedly, the inner volume of **5**→**1** was found to be slightly larger (146 Å³) than that of **3**→**1** (142 Å³) despite the introduction of two fluorine atoms in the cavity. This reflects the plasticity of the aromatic amide backbone which undergoes subtle conformational changes upon binding a guest. Nevertheless, the absence of cavity volume reduction and of direct interactions between the fluorobenzene and the guest make it difficult to explain why binding with **5** is enhanced. This enhancement was anyway a positive result and exploited in the second phase of the directed evolution.

Worth noting is the last iteration of this phase. When the double naphthyridine deletion and the double **P**→**F** mutation are performed simultaneously (sequence **6**) binding to tartaric acid is suppressed and binding to malic acid is reduced by 50 fold. This can be explained by the critical involvement of pyridine motifs (in **P** or **N** monomers) in hydrogen bonds with hydroxyl groups of the guests in reduced sequence **4**. Replacing the **P** units of **4** by **F** monomers in **6** causes a loss of interactions and probably also generates some steric hindrance with both guests. In principle, sequence **6** provides a first example of guest preference reversal, but not in a satisfactory manner. We thus endeavored to desymmetrize sequence **5** to elicit selectivity for malic acid.

Table 1. Host-guest association constants (K_a) determined by ¹H NMR in 9:1 CDCl₃/[D₆]-DMSO (vol/vol) at 298 K.

Guest	K_a (L mol ⁻¹) ^a (<i>d.e.</i>) ^b							
	3 ^c	4	5	6	7	8	9	<i>P-10</i>
D/L- 1	5300 (>99)	2600 (>99)	29000 ^d (>99)	<1	<1	2600 (25)	2100 (>99)	240 ^j
D/L- 2	70 (82) ^e	60 ^e	550 (80)	<10	200 ^f	5000 ^g (88)	23000 ^h (71)	40000 ⁱ

^aCalculated from integrating ¹H NMR signals of free host and host-guest under slow exchange on NMR timescale. ^bCalculated diastereomeric excess (*d.e.* in %). ^cFrom reference 12. ^dFrom reference 21. ^eFast exchange on NMR timescale. ^fSignal broadening prevented *d.e.* calculation. ^gBinding constant was found to be 5030 L mol⁻¹ using CD. ^hBinding constant was found to be 20200 L mol⁻¹ using CD. For (f), (g) and (i) natural L-**2** enantiomer was used. ^jNatural L-**1** enantiomer was used.

Helix symmetry breaking amounted to prepare capsules lacking any symmetry, composed for one half of a **Q₃FN₂** segment optimized to bind an α -hydroxyacid moiety, and for the other half of a segment optimized for the methylene group of **2**. The method to join two distinct helical cones to a central **pyr-pyz-pyr** unit has been presented in reference 11. It is not straightforward in the sense that we could not find efficient ways to desymmetrize **pyr-pyz-pyr** under the form of a simple mono-acid mono-ester. Instead, desymmetrization was achieved by producing **N-pyr-pyz-pyr**. This segment was elongated to **Q₃FN₂-pyr-pyz-pyr** which served as a common precursor to sequences **7** to **10**, as described in the Supporting Information.

Starting from capsule **5**, the objective was to preserve key interactions of one half of the capsule with the α -hydroxy acid moiety of malic acid, while filling the void left around the methylene group of malic acid in the other half of the cavity. Too large a capsule would result in a lack of selectivity, whereas too small a capsule would result in a drop of affinity. A first iteration consisted in testing a naphthyridine deletion, as in sequence **6**, but in one half of the sequence. The behavior of **6** had clearly shown that combining a naphthyridine deletion and the presence of an **F** unit was detrimental to binding an α -hydroxy acid. Dissymmetrical receptor **Q₃FN₂-pyr-pyz-pyr-NFQ₃** (sequence **7**) was thus generated. As for **6**, binding to tartaric acid is negligible, while binding to malic acid stands at 200 L mol⁻¹ (Figure S6), a value above the affinity of our starting sequence **3** for malic acid. At this stage, we could have considered our initial objective to be achieved, having used negative design to exclude tartaric acid and to completely revert guest selectivity, while preserving interactions with malic acid. It remained that a K_a of 200 L mol⁻¹ for malic acid is a modest value, which is

probably responsible for our inability to grow crystals and obtain a structure of **7**→**2** and further improve this design.

Table 2. Calculated complex cavity volume.

Capsule	Volume ^a (Å ³)	P. C. ^b (%)
3 ^c	136	-
3 → 1	142	70.4
4 → 1	120	83.3
5 → 1	146	68.5
8 → 1	145	69.0
8 → 2	140	67.8
9 → 1	160	62.5
9 → 2	156	60.9

^a volume of capsule cavities determined using SURFNET v1.4 (see SI). ^b Packing coefficients defined here as the ratio of the guest volume to the host cavity volume. Volumes of guests **1** and **2** were found to be 100 and 95 Å³, respectively. ^c Predicted capsule structure obtained by molecular modeling (MMFFs force field) using Maestro v6.5.

Further iterations (sequences **8** to **10**) were motivated by curiosity, by opportunism (the availability of **Q^F** monomer), and by our empirical experience that removing space through deletions is quite drastic whereas bringing in multiple fluorine atoms sometimes leads to unexpected (and difficult to explain) enhancement of binding (sequence **5** being an example). The

positive outcome of the iterations reported below can thus be considered to be in part serendipitous. We stepped back to sequence **5** and considered that another way to disfavor interactions with an α -hydroxy acid and enhance complementarity with a methylene group was not to delete **N** (which amounts to place **F** at that position of the sequence as in **7**), but instead to implement an $N \rightarrow Q^F$ mutation, Q^F having one ring identical to that of **F**. Capsule **8**, Q_3FN_2 -pyr-pyz-pyr- NQ^FFQ_3 , was then prepared and its selectivity towards guests **1** and **2** was assessed by 1H NMR. Binding properties towards both **1** and **2** were surprising. Affinity for **1** did not collapse ($K_a = 2600 \text{ L mol}^{-1}$), showing that the fluorine atom of Q^F fails to achieve what a deletion does. In addition, the NMR spectrum shows two sets of signals (Figure 5B), revealing the presence of two diastereomeric complexes, and thus a major drop in diastereoselectivity. The two hydrogen bonded tartaric acid protons are inequivalent and show distinct resonances, reflecting the dissymmetrical environment of the capsule and the fact that the guest does not tumble rapidly in the cavity. The chemical shift values of bound acid resonances (acid protons are more deshielded in matching complexes)^{7,20} suggest that the major species (63 %) is still the matching complex $P-8\supset D-1/M-8\supset L-1$. Nevertheless, the mismatching complex $P-8\supset L-1/M-8\supset D-1$ represents a substantial fraction (37 %). Crystals of **8** \supset **1** were obtained and allowed to solve simultaneously the structure of the matching (Figure 5D,G) and of the mismatching (Figure 5E,H) host-guest-complexes. The structure revealed a disorder of the tartaric acid hydroxyl groups, meaning that matching and mismatching complexes had co-crystallized. Such disorder is never observed in highly diastereoselective complexes. The disorder could eventually be resolved and a 43/57 ratio of matching and mismatching complexes in the crystal was modelled. The structure of the mismatching complex $P-8\supset L-1/M-8\supset D-1$ shows how the hydroxyl groups of tartaric acid escape hindrance from the Q^F fluorine atom by pointing towards the pyr-pyz-pyr segment.

Binding of capsule **8** with malic acid **L-2** revealed an unanticipated ten-fold enhancement of affinity ($K_a = 5000 \text{ L mol}^{-1}$) with respect to **5** (Figure 5C). This value was confirmed by a CD titration of **8** by the natural malic acid enantiomer **L-2** which results in an intense induced positive CD band revealing induction of *P* handedness²² (Figure S1). Thus, capsule **8** actually binds malic acid with an affinity twice as large as for tartaric acid. NMR also reveals the presence of diastereomeric complexes in different proportions. Two sharp acid resonances were observed for each diastereoisomer (Figure 5C). This indicates that, for each diastereoisomer, malic acid adopts a single well defined orientation within the dissymmetrical cavity, or that it tumbles rapidly between two orientations.²³ However, the latter is less likely as it would presumably result in broader NMR signals (see Figure 4C) and because there is evidence that tartaric acid tumbles slowly (see above). Crystals of **8** \supset **2** were also obtained. The structure revealed the matching configuration of **2** with its hydroxyl group pointing toward the naphthyridine escaping the hindrance of the fluorine atom of the Q^F monomer (Figure 5F,I). The exact origin of the enhanced binding of **2** by **8** is unclear, but it comforted us in pursuing into further changes of similar nature.

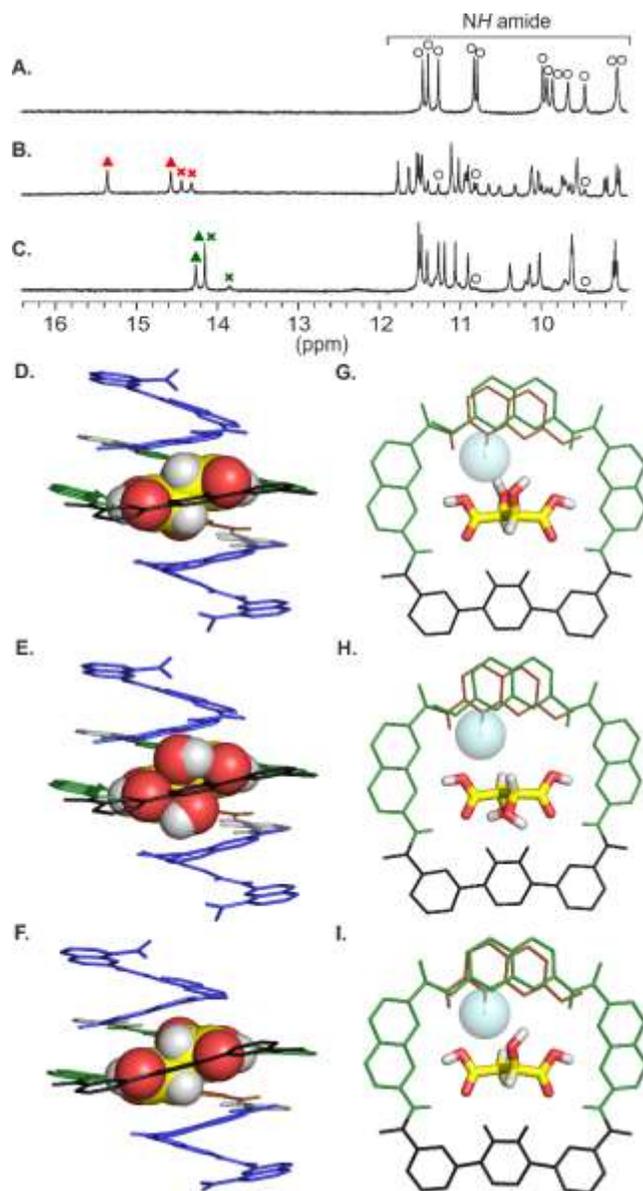


Figure 5. Excerpts of the 400 MHz 1H NMR spectra of capsule **8** (1 mM) in 90:10 $CDCl_3/[D_6]-DMSO$ (vol/vol) at 298 K: (A) without guest, in the presence of (B) 3 equiv. of **D/L-1** and (C) 6 equiv. of **D/L-2**. Signals of the empty host **8** are marked with empty circles. Acid resonances of the guests in the matching $P-8\supset D-1/M-8\supset L-1$ and mismatching $P-8\supset L-1/M-8\supset D-1$ complexes are marked with red triangles and red crosses, respectively. Similarly, matching $P-8\supset L-2/M-8\supset D-2$ and mismatching $P-8\supset D-2/M-8\supset L-2$ complexes are marked with green triangles and green crosses, respectively. Crystal structures of: (D) the tartaric acid matching complex $P-8\supset D-1$, (E) the tartaric acid mismatching complex $P-8\supset L-1$ and (F) the malic acid matching complex $P-8\supset L-2$. The three structures are shown in tube (host) and CPK (guest) representations. Top views of a slice of the same complexes: (G) $P-8\supset D-1$, (H) $P-8\supset L-1$ and (I) $P-8\supset L-2$. For the three structures both host and guest are shown in tube representation. The fluorine atom of the Q^F unit is shown as a transparent light blue sphere. Isobutoxy side chains and solvent molecules were omitted for clarity.

The structure of the mismatching complex $P-8\supset L-1/M-8\supset D-1$ points to the necessity of disfavoring not only matching but also mismatching association with tartaric acid. Based on this structure and on those of the other complexes, we proposed the

next iteration to consist in the addition of a peripheral Q^F monomer projecting a third exocyclic fluorine atom in the vicinity of expected location of the methylene group of malic acid. This addition led to sequence **9**, Q_3FN_2 -pyr-pyz-pyr- $NQ^FFQ^FQ_3$. Again, binding properties were somewhat surprising. Disappointingly, binding to **1** was not significantly altered ($K_a = 2100 \text{ L mol}^{-1}$), but the diastereoselectivity partly lost in **8** was completely recovered (Figure S9), showing that the addition of Q^F is effective at disfavoring mismatching complexes. Indeed, the crystal structure of **9****1** (Figure S17) confirmed that the unique complex observed in solution is having matching chirality (P -**9** \supset **D**-**1**/ M -**9** \supset **D**-**1**). Binding to **2** by sequence **9** was again enhanced with respect to **8**, this time by a factor of four, and eventually became ten times larger than binding to **1**. CD titrations confirm the value measured by NMR (Figure 6E, 6F). With respect to original sequence **3**, four fluorine atoms have been added and binding for malic acid has increased by a factor over 300. As for **8**, diastereomeric complexes can be observed on ^1H NMR spectra, indicating the coexistence of the matching P -**9** \supset **L**-**2**/ M -**9** \supset **D**-**2** and mismatching P -**9** \supset **D**-**2**/ M -**9** \supset **L**-**2** species in solution (*d.e.* = 71%, Figure 6B). The difference in proportion is reflected by an intense induced CD observed upon adding **L-2** to **9**. The positive sign of the CD band allows to assign preferred binding of **L-2** in P helices.²²

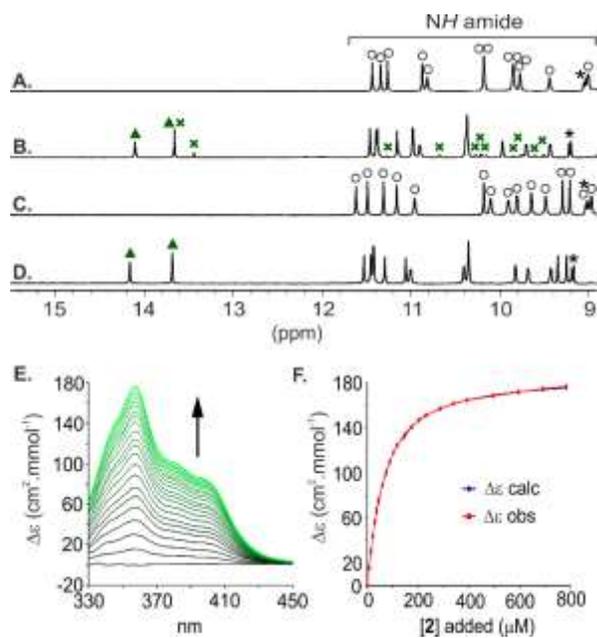


Figure 6. Excerpts of the 300 MHz ^1H NMR spectra in 90/10 $\text{CDCl}_3/[\text{D}_6]\text{-DMSO}$ (vol/vol) at 298 K of: (A) capsule **9** (1 mM) without guest (empty circles) and in the presence of (B) 3 equiv. of **D/L-2**; (C) capsule **10** (1 mM) without guest (empty circles) and in the presence of (D) 2 equiv. of **L-2**. Acid resonances of the guest in the matching P -**9** \supset **L**-**2**/ M -**9** \supset **D**-**2** and P -**10** \supset **L**-**2** complexes are marked with green triangles. Mismatching complexes P -**9** \supset **D**-**2**/ M -**9** \supset **L**-**2** are denoted with green crosses. (E) Variation of the CD spectrum of capsule **9** upon the addition of **L-2** in 90:10 $\text{CDCl}_3/[\text{D}_6]\text{-DMSO}$ (vol/vol) at 298 K. (F) Curve fitting of the CD titration of **9** with **L-2** monitored at 358 nm. $K_a = 20200 \text{ L mol}^{-1}$. $\Delta\epsilon_{\text{max}} = 186.6 \text{ cm}^2 \text{ mmol}^{-1}$.

A crystal structure of P -**9** \supset **L**-**2**/ M -**9** \supset **D**-**2** (Figure 7) confirmed the handedness assignment based on CD as well as the positioning of malic acid in the receptor cavity, with its α -hydroxy acid surrounded by the Q_3FN_2 segment of the host and its $\text{CH}_2\text{CO}_2\text{H}$ moiety surrounded by the $NQ^FFQ^FQ_3$ segment,

with three fluorine atoms positioned around the methylene group. In this complex, the occupancy factor of the cavity by the guest was 60.9% (Table 2), the lowest of all in the whole series. In these crystals, the asymmetric unit contained two independent molecules of host **9**. One as the matching complex P -**9** \supset **L**-**2**/ M -**9** \supset **D**-**2**, mentioned above, and the other incidentally containing complex P -**9** \supset **D**-**2**/ M -**9** \supset **L**-**2** with mismatching chirality of the host and guest. The latter is shown in Figure 7F. It reveals that malic acid manages to binding helices of mismatching chirality by a completely different mechanism than tartaric acid which is found with its hydroxyl groups pointing toward the pyridazine unit in mismatching complexes (Figure 5E,H, see also reference 20). Instead malic acid undergoes a rotation about its CH-CH_2 bond which allows one of its carboxylic acid function to hydrogen bond to a naphthyridine opposite to **pyr-pyz-pyr** unit in the helix. This binding mode is probably responsible for the lower *d.e.* values observed for malic acid binding in these receptors in general.¹²

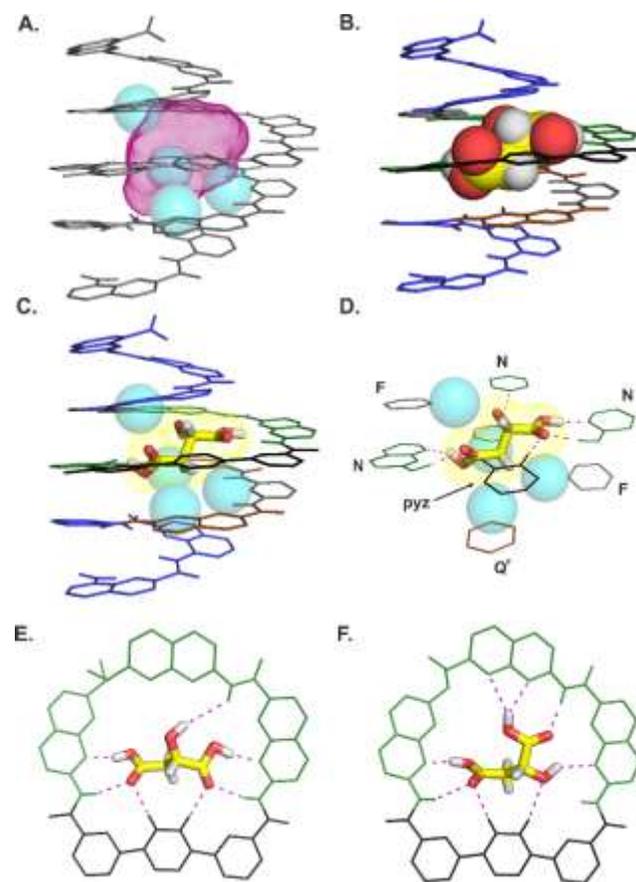


Figure 7. Side views of the crystal structure of P -**9** \supset **L**-**2**: (A) the capsule appears in grey tube representation, the guest is not shown whereas the volume of the cavity (156 \AA^3) is shown as a transparent purple isosurface; (B) the host and the guest are shown in tube and CPK representation, respectively. (C) The backbone of the capsule and the guest are represented in tube; (D) **D**-malic acid **2** and those heterocycles that interact with it are shown in a tube representation. Views of a slice of: (E) the matching P -**9** \supset **L**-**2** complex and (F) the mismatching P -**9** \supset **D**-**2** host-guest complex. For each complex the array of hydrogen bonds is shown as magenta dashes. In (A), (C) and (D) fluorine atoms appear as transparent cyan CPK spheres. In (C) and (D), the volume of the guest (95 \AA^3) is shown as a transparent yellow isosurface. Monomers are color coded as in Figure 3.

Finally, a last iteration was implemented to take full advantage of the diastereoselectivity of the hosts and of the fact that naturally occurring L-1 and L-2 have opposite stereochemistry. Through the addition of two terminal (1S)-(-)-camphanyl groups (**R***) to sequence **9**, the resulting sequence **10** was made to be exclusively right-handed.²² Upon titration of *P*-**10** with malic acid L-2 which possess matching stereochemistry, a K_a of 40000 L mol⁻¹ was measured. This value is twice as high as that of **9** for D/L-2, as can be expected since all complexes are productive when mixing *P*-**10** and L-2 whereas only half of possible host-guest combinations are matching complexes when mixing **9** and D/L-2. It can be noted that the titration of *P*-**10** with L-2 gives rise to the emergence of only one set of peaks corresponding to the matching complex. This provides evidence, a posteriori, that the two species observed in the titrations of **8** and **9** with D/L-2 do correspond to diastereomeric complexes with matching or mismatching chirality, and not to different orientations of malic acid within the same complex.

When titrating *P*-**10** with tartaric acid L-1, K_a is only 240 L mol⁻¹ because this combination corresponds to a host and a guest having mismatching chirality. The overall outcome is that the selectivity of **10** for L-2 with respect to L-1 amounts to a factor over 160. That is maybe not as high as the selectivity of **7** but it is achieved with a 200 fold higher affinity for malic acid. When compared to starting sequence **3**, the selectivity for malic acid of sequence **10** has changed by a ratio over 12000, and the affinity for malic acid has been increased by almost 600 fold.

CONCLUSION

In conclusion, we have been able to modify and fine-tune host-guest binding properties of helically folded aromatic oligoamide receptors through seven iterative modifications of their sequences followed by structural investigations of host-guest complexes. One outcome was to meet the initial objective of reversal of selectivity of the receptors, with a starting preference for tartaric acid over malic acid of over 10² and an ending sequence showing a preference for malic acid over tartaric acid of over 10². Another outcome was a very strong enhancement of the affinity for malic acid, despite the fact it has fewer recognition features for binding through polar interactions such as hydrogen bonds. Such level of discrimination between very much resembling guests exemplifies the amenability of foldamers to outstanding achievements in molecular recognition. Modifications (deletions, mutations or additions) from one iteration to the next were guided by the rational consideration of structural information. In many cases, the effect of modifications on guest binding properties was as expected, indicating a high level of predictability of these systems. In some cases, unexpected and often unexplained (yet favorable) effects occurred, such as the progressive buildup of binding affinity for malic acid as fluorine atoms are introduced so as to surround its methylene group. Altogether, our results demonstrate the viability of a rational receptor design approach that exploits the modularity of foldamer sequences and, in the case of aromatic amide foldamers, their amenability to structural elucidation, their relative ease of synthesis – only one of us carried out all syntheses reported here – and the predictability of their structure. A desirable extension of this approach would include advanced computational tools to rate possible iterations and formulate an even more reliable choice of the iteration most likely to bring about enhanced properties.

ASSOCIATED CONTENT

Supporting Information

Experimental details for synthetic procedures, spectroscopic data and crystallographic information files for **4**→**1**, **5**→**1**, **8**→**1**, **8**→**2**, **9**→**1**, and **9**→**2** (CIF). This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Authors

*y.ferrand@iecb.u-bordeaux.fr, i.huc@iecb.u-bordeaux.fr

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

This work was supported by the Conseil Interprofessionnel du Vin de Bordeaux (predoctoral fellowship to G.L.) and by the European Research Council under the European Union's Seventh Framework Programme (grant agreement no. ERC-2012-AdG-320892, postdoctoral fellowship to G.L. and B.W.). We thank Dr. Jean Luc Ferrer for beamtime and help during data collection on FIP BM30A at the ESRF.

REFERENCES

- (1) (a) Guichard, G.; Huc, I. *Chem. Commun.* **2011**, 47, 5933–5941. (b) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, 101, 3893–4011. (c) Gellman S. H. *Acc. Chem. Res.* **1998**, 31, 173–180.
- (2) *note*: when it does not encounter kinetic traps that are common for large folded structures.
- (3) (a) Zhang, D.-W.; Zhao, X.; Hou, J.-L.; Li, Z.-T. *Chem. Rev.* **2012**, 112, 5271–5316. (b) Saraogi, I.; Hamilton A. D. *Chem. Soc. Rev.* **2009**, 38, 1726–1743. (c) Juwarker, H.; Suk, J.-m.; Jeong K.-S. *Chem. Soc. Rev.* **2009**, 38, 3316–3325.
- (4) (a) Warmuth, R.; Yoon, J. *Acc. Chem. Res.* **2001**, 34, 95–105. (b) Cram, D. J.; Karbach, S.; Kim, Y. H.; Baczynskyj, L.; Kalleyman, G. W. *J. Am. Chem. Soc.* **1985**, 107, 2575–2576. (c) Garel, L.; Dutasta, J.-P.; Collet, A. *Angew. Chem. Int. Ed.* **1993**, 32, 1169–1171. (d) Timmerman, P.; Verboom, W.; van Veggel, F. C. J. M.; van Hoorn, W. P.; Reinhoudt, D. N. *Angew. Chem. Int. Ed.* **1994**, 33, 1292–1295. (e) Hunter, C. A.; Sarson, L. D. *Angew. Chem. Int. Ed.* **1994**, 33, 2313–2316.
- (5) (a) Hou, J.-L.; Shao, X.-B.; Chen, G.-J.; Zhou, Y.-X.; Jiang, X.-K.; Li, Z.-T. *J. Am. Chem. Soc.* **2004**, 126, 12386–12394. (b) Li, C.; Wang, G.-T.; Yi, H.-P.; Jiang X.-K.; Li, Z.-T.; Wang, R.-X. *Org. Lett.* **2007**, 9, 1797–1800. (c) You, L.-Y.; Chen, S.-G.; Zhao, X.; Liu, Y.; Lan, W.-X.; Zhang, Y.; Lu, H.-J.; Cao, C.-Y.; Li Z.-T. *Angew. Chem., Int. Ed.* **2012**, 51, 1657–1661. (d) Waki, M.; Abe, H.; Inouye, M. *Angew. Chem., Int. Ed.* **2007**, 46, 3059–3061. (e) Abe, H.; Machiguchi, H.; Matsumoto, S.; Inouye, M. *J. Org. Chem.* **2008**, 73, 4650–4661. (f) Chang, K.-J.; Kang, B.-N.; Lee, M.-H.; Jeong, K.-S. *J. Am. Chem. Soc.* **2005**, 127, 12214–12215. (g) Hua, Y.; Flood, A. H. *J. Am. Chem. Soc.* **2010**, 132, 12838–12840. (h) Hua, Y.; Liu, Y.; Chen, C.-H.; Flood, A. H. *J. Am. Chem. Soc.* **2013**, 135, 14401–14412. (i) Wang, Y.; Bie, F.; Jiang, H. *Org. Lett.* **2010**, 12, 3630–3633. (j) Nishinaga, T.; Tanatani, A.; Oh, K.; Moore, J. S. *J. Am. Chem. Soc.* **2002**, 124, 5934–5935. (k) Tanatani, A.; Hughes, T. S.; Moore, J. S. *Angew. Chem. Int. Ed.* **2002**, 41, 325–328. (l) Gan, Q.; Ferrand, Y.; Bao, C.; Kauffmann, B.; Grélard, A.; Jiang, H.; Huc, I. *Science* **2011**, 331, 1172–1175. (m) Ferrand, Y.; Gan, Q.; Kauffmann, B.; Jiang, H.; Huc, I. *Angew. Chem. Int. Ed.* **2011**, 50, 7572–7575. (n) Gan, Q.; Ferrand, Y.; Chandramouli, N.; Kauffmann, B.; Aube, C.; Dubreuil, D.; Huc, I. *J. Am. Chem. Soc.* **2012**, 134, 15656–15659.
- (6) (a) Garric, J.; Léger, J.-M.; Huc, I. *Angew. Chem. Int. Ed.* **2005**, 44, 1954–1958. (b) Bao, C.; Kauffmann, B.; Gan, Q.; Srinivas, K.; Jiang, H.; Huc, I. *Angew. Chem. Int. Ed.* **2008**, 47, 4153–4156.

- (7) Ferrand, Y.; Chandramouli, N.; Kendhale, A. M.; Aube, C.; Kauffmann, B.; Grélard, A.; Laguerre, M.; Dubreuil, D.; Huc, I. *J. Am. Chem. Soc.* **2012**, *134*, 11282–11288. Abramyan, A. M.; Liu Z.; Pophristic, V. *Phys. Chem. Chem. Phys.* **2014**, *16*, 20406–20410.
- (8) (a) Yin, H.; Hamilton, A. D. *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163. (b) Azzarito, V.; Long, K.; Murphy, N. S.; Wilson, A. J. *Nature Chem.* **2013**, *5*, 161–173. (c) Boersma, M. D.; Haase, H. S.; Petersson-Kaufman, K. J.; Lee, E. F.; Clarke, O. B.; Colman, P. M.; Smith, B. J.; Horne, S. W.; Fairlie, W. D.; Gellman, S. H. *J. Am. Chem. Soc.* **2012**, *134*, 315–323. (d) Michel, J.; Harker, E. A.; Tirado-Rives, J.; Jorgensen, W. L.; Schepartz, A. *J. Am. Chem. Soc.* **2009**, *131*, 6356–6357. (e) Park, H. S.; Lin, Q.; Hamilton, A. D. *J. Am. Chem. Soc.* **1999**, *121*, 8–13.
- (9) (a) Jena, P. V.; Shirude, P. S.; Okumus, B.; Laxmi-Reddy, K.; Godde, F.; Huc, I.; Balasubramanian, S.; Ha, T. *J. Am. Chem. Soc.* **2009**, *131*, 12522–12523. (b) Shirude, P. S.; Gillies, E. R.; Ladame, S.; Godde, F.; Shin-ya, K.; Huc, I.; Balasubramanian, S. *J. Am. Chem. Soc.* **2007**, *129*, 11890–11891.
- (10) Huc, I. *Eur. J. Org. Chem.* **2004**, 17–29.
- (11) Chandramouli, N.; Ferrand, Y.; Lautrette, G.; Kauffmann, B.; Mackereth, C. D.; Laguerre, M.; Dubreuil, D.; Huc, I. *Nature Chem.* **2015**, *7*, 334–341.
- (12) Ferrand, Y.; Kendhale, A. M.; Kauffmann, B.; Grélard, A.; Marie, C.; Blot, V.; Pipelier, M.; Dubreuil, D.; Huc, I. *J. Am. Chem. Soc.* **2010**, *132*, 7858–7859.
- (13) (a) Lavigne, J. J.; Anslyn E. V. *Angew. Chem. Int. Ed.* **1999**, *38*, 3666–3669. (b) Wiskur, S. L.; Floriano, P. N.; Anslyn, E. V.; McDevitt, J. T. *Angew. Chem. Int. Ed.* **2003**, *42*, 2070–2072. (c) Piątek, A. M.; Bomble, Y. J.; Wiskur, S. L.; Anslyn, E. V. *J. Am. Chem. Soc.* **2004**, *126*, 6072–6077. (d) Alfonso, I.; Rebolledo, F.; Gotor, V. *Chem. Eur. J.* **2000**, *6*, 3331–3338. (e) Alfonso, I.; Dietrich, B.; Rebolledo, F.; Gotor, V.; Lehn J.-M. *Helv. Chim. Acta* **2001**, *84*, 280–295. (f) González-Álvarez, A.; Alfonso, I.; Díaz, P.; García-España, E.; Gotor V. *Chem. Commun.* **2006**, 1227–1229. (g) Faggi, E.; Porcar, R.; Bolte, M.; Luis, S. V.; García-Verdugo, E.; Alfonso, I. *J. Org. Chem.* **2014**, *79*, 9141–9149. (h) Gamsey, S.; Miller, A.; Olmstead, M. M.; Beavers, C. M.; Hirayama, L. C.; Pradhan, S.; Wessling, R. A. Singaram, B. *J. Am. Chem. Soc.* **2007**, *129*, 1278–1286.
- (14) (a) Ghosh, K.; Adhikari S. *Tetrahedron Lett.* **2006**, *47*, 3577–3581. (b) Bencini, A.; Coluccini, C.; Garau, A.; Giorgi, C.; Lippolis, V.; Messori, L.; Pasini, D.; Puccioni, S. *Chem. Commun.* **2012**, *48*, 10428–10430.
- (15) Berni, E.; Garric, J.; Lamit, C.; Kauffmann, B.; Léger, J.-M.; Huc, I. *Chem. Commun.* **2008**, 1968–1970.
- (16) Garcia-Tellado, F.; Geib, S. J.; Goswami, S.; Hamilton, A. D. *J. Am. Chem. Soc.* **1991**, *113*, 9265–9269.
- (17) Mecozzi, S.; Rebek, J. Jr *Chem. Eur. J.* **1998**, *4*, 1016–1022.
- (18) Qi, T.; Deschrijver, T.; Huc, I. *Nat. Protoc.* **2013**, *8*, 693–708.
- (19) Laskowski, R. A. *J. Mol. Graph.* **1995**, *13*, 323–330.
- (20) Lautrette, G.; Kauffmann, B.; Ferrand, Y.; Aube, C.; Chandramouli, N.; Dubreuil, D.; Huc, I. *Angew. Chem. Int. Ed.* **2013**, *52*, 11517–11520.
- (21) Chandramouli, N.; El-Behairy, M. F.; Lautrette, G.; Ferrand, Y.; Huc, I. *Org. Biomol. Chem.* **2016**, *14*, 2466–2472.
- (22) Kendhale, A. M.; Poniman, L.; Dong, Z.; Laxmi-Reddy, K.; Kauffmann, B.; Ferrand, Y.; Huc, I. *J. Org. Chem.* **2011**, *76*, 195–200.
- (23) Another interpretation of the two sets of peaks could be that they correspond to two distinct orientations of malic acid in slow exchange within only one diastereomeric complex. Yet it is difficult to conceive that multiple orientations would not result in a loss of diastereoselectivity. In addition, this interpretation is experimentally ruled out in the case of sequence **9** (see below) for which absolute helix handedness control (sequence **10**), *i.e.* making it impossible to form one diastereoisomer, results in the disappearance of the minor species on NMR spectra.

Graphic entry for the Table of Contents (TOC)

