



**HAL**  
open science

## Iterative design of a helically folded aromatic oligoamide sequence for the selective encapsulation of fructose

Nagula Chandramouli, Yann Ferrand, Guillaume Lautrette, Brice Kauffmann, Cameron David Mackereth, Michel Laguerre, Didier Dubreuil, Ivan Huc

### ► To cite this version:

Nagula Chandramouli, Yann Ferrand, Guillaume Lautrette, Brice Kauffmann, Cameron David Mackereth, et al.. Iterative design of a helically folded aromatic oligoamide sequence for the selective encapsulation of fructose. *Nature Chemistry*, 2015, 7 (4), pp.334-341. 10.1038/nchem.2195 . hal-01509470

**HAL Id: hal-01509470**

**<https://hal.science/hal-01509470v1>**

Submitted on 19 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 design of a helically folded aromatic oligoamide sequence for the selective encapsulation of fructose**

Nagula Chandramouli,<sup>1,2,†</sup> Yann Ferrand,<sup>1,2,†</sup> Guillaume Lautrette,<sup>1,2</sup> Brice Kauffmann,<sup>3,4,5</sup>  
Cameron D. Mackereth,<sup>6,7</sup> Michel Laguerre,<sup>1,2</sup> Didier Dubreuil,<sup>8,9</sup> Ivan Huc<sup>1,2\*</sup>

<sup>1</sup> Univ. Bordeaux, CBMN (UMR 5248), Institut Européen de Chimie et Biologie, 2 rue Escarpit 33600 Pessac, France.

<sup>2</sup> CNRS, CBMN (UMR 5248), France.

<sup>3</sup> Univ. Bordeaux, IECB (UMS 3033/US 001), Institut Européen de Chimie et Biologie, 2 rue Escarpit 33600 Pessac, France.

<sup>4</sup> CNRS, IECB (UMS 3033), France.

<sup>5</sup> INSERM, IECB (US 001), France.

<sup>6</sup> Univ. Bordeaux, ARNA (U 869), Institut Européen de Chimie et Biologie, 2 rue Escarpit 33600 Pessac, France.

<sup>7</sup> INSERM, ARNA (U869), France.

<sup>8</sup> Univ. Nantes, CEISAM (UMR6230), Faculté des Sciences et des Techniques, 2 rue de la Houssinière, BP 92208, 44322 Nantes Cedex 3, France.

<sup>9</sup> CNRS, CEISAM (UMR6230), France.

\* To whom correspondence should be addressed. E-mail: i.huc@iecb.u-bordeaux.fr.

† These authors contributed equally to this work

## ABSTRACT

The *ab initio* design of synthetic molecular receptors for a given biomolecular guest remains an elusive objective. We describe a powerful and novel approach to produce receptors of exquisite selectivity in the challenging context of monosaccharide recognition by using an iterative evolution process that exploits the modular structure of folded synthetic oligomer sequences coupled with molecular modelling and structural characterisation. Starting from a first principles design taking size, shape and hydrogen-bonding ability into account, and using the high predictability of aromatic oligoamide foldamer conformations and their propensity to crystallise, a sequence that binds to  $\beta$ -D-fructopyranose in organic solvents with atomic scale complementarity was obtained in just a few iterative modifications. This scheme, which mimics the adaptable construction of biopolymers from a limited number of monomer units, provides a general protocol by which to create selective receptors towards an eventual wide range of monosaccharides or other biological and synthetic compounds.

## INTRODUCTION

The primary structure of proteins and nucleic acids - multiple monomeric units connected in long sequential arrangements - reflects exquisite modularity. Fine tuning of structures and folded conformations are generated *via* additions, deletions or mutations of monomers, or the recombination of segments and entire domains, to allow for the evolution of polymer sequences without having to modify synthetic or biosynthetic schemes. In recent years, chemists have shown that synthetic molecular strands having backbones remote from those of peptides and nucleotides may also fold into well-defined structures in solution<sup>1-7</sup>. Even in the absence of dedicated enzymatic systems to replicate them, these artificial backbones, termed synthetic foldamers<sup>8</sup>, are expected to be amenable to some sort of bioinspired evolutionary process. The approach would exploit the modularity of their primary sequences and their relatively easy and repetitive synthesis to accelerate the emergence of functions. For example, pyrrole-imidazole oligomers have been created to recognise double stranded DNA in a sequence selective manner<sup>9</sup>. Here, we show that the structure-directed iterative optimization of an aromatic oligoamide enables rapid production of helically-folded containers that bind monosaccharide guests, arguably some of the most challenging targets in the field of molecular recognition<sup>10,11</sup>. In just a few iterations, an initial sequence with relatively poor guest selectivity was evolved into a receptor that has fully characterised atomic-scale complementarity for  $\beta$ -D-fructopyranose. This evolutionary approach thus emerges as a novel method to tailor receptors for predefined complex guests. A schematic view of the successive steps of our method is shown in **Fig. 1**. In practice, the method requires the availability of monomers for which the monomer sequence alone provides a good predictability of overall receptor shape and binding features. In addition, the host-guest complexes must be amenable to accurate structural elucidation, including for difficult guests. These features are all combined in aromatic oligoamide foldamers<sup>12,13</sup>.

The molecular recognition of saccharides by both natural and artificial receptors is seriously challenged by issues of affinity and selectivity. The difficulty in defining high affinities in water or polar solvents is due to competing hydrogen bonds between saccharides and solvent molecules<sup>14</sup>. This is less acute for sugars with well-defined hydrophobic patches, such as the multiple axial protons of the “all-equatorial” glucose derivatives<sup>15,16</sup>. A possible way to improve binding in polar media is to introduce boronic acid moieties on the receptor to reversibly form boronate esters with two vicinal hydroxyl groups of the saccharides<sup>17,18</sup>. Multiple hydrogen bonds to hydroxyl groups also elicit binding that can be very strong in less polar media<sup>19</sup>.

Challenges in selectivity arise from the fact that saccharides inherently resemble each other and that most of them co-exist under several forms at equilibrium. For example, exchange between  $\alpha$  vs  $\beta$  anomers, or furanose vs pyranose tautomers, further increases the chance for one given saccharide to resemble another. Resolving these issues is important not just for proteins and other biopolymers, but also for practical uses in synthetic receptors, as minor saccharide structural differences are often biologically meaningful. Despite the numerous synthetic receptors of saccharides described to date<sup>20,21,22,23,24,25,26,27,28,29</sup>, progress towards high selectivity has been modest except for all-equatorial glucose derivatives<sup>15,16,20</sup> which possess a relatively unique disc-like shape. A few foldamer-based saccharide receptors have been reported in which a sugar is bound in the cavity formed by a helically folded oligomer<sup>28,29</sup>. However, these did not perform better than other types of receptors and were not easily amenable to optimization due to the lack of accurate information about host-guest complex structures. Effective synthetic receptors may also be identified upon screening libraries but rational improvements are difficult to implement in this case<sup>30</sup>. In contrast, the structure-based iterative evolution presented below (**Fig. 1**) represents a new and powerful methodological

approach in this field. The selectivity for  $\beta$ -fructopyranose that we attained, in particular with respect to closely resembling sugars such as mannopyranose, is unmatched.

In this work, we explore a novel and original structure-based iterative approach paving the way to a general strategy aiming to discriminate between resembling sugars by subsequent subtle evolution of a parent heterocyclic foldamer sequence (1st generation receptor Fig. 1). To illustrate our purpose, investigation was focused on the difficult distinction of fructose against other sugar homologues, in particular closed mannose structure

## RESULTS AND DISCUSSION

### Design principles of the initial guest

Following well established rules (**Supplementary Fig. S1**)<sup>31</sup>, a first generation aromatic oligoamide sequence **1** (**Fig. 2b-c**) was designed to fold into a helix having a reduced diameter at both ends, thereby creating a cavity that can completely surround and bind guest molecules in solution (**Fig. 2a**). This type of molecular container possesses a smaller cavity than capsules formed by metal or hydrogen-bond mediated self-assembly of organic building blocks but have the advantage of having no required symmetry<sup>32,33,34</sup>, a desirable property to selectively bind to complex guests. Similar to other folded aromatic oligoamides, local conformational preferences at aryl-amide linkages and intramolecular  $\pi$ - $\pi$  interactions between aromatic monomers combine to endow the folded conformation with high stability and favourable crystallogenes<sup>12,13</sup>. The contribution of each monomer to the helix curvature can be predicted with good accuracy based on general force fields and simple energy minimisation, thereby allowing initial modelling of the size and shape of the central cavity, as well as the position of groups that line the interior and determine molecular recognition properties. In such constructs, guest capture and release occur *via* dynamic conformational changes of the backbone<sup>35</sup>. Sequence **1** was designed so that most amide protons and endocyclic nitrogen atoms, as well

as two carbonyl groups belonging to novel H monomers, are located on the helix inner rim. The number of accessible hydrogen bond donors (x12) and acceptors (x20) and their overall even distribution was predicted to be sufficient to engage in multiple interactions generally required to bind to monosaccharides, but their detailed spatial arrangement was unbiased and not designed to target any particular guest. The outcome is a large chiral cavity with a predicted volume of 216 Å<sup>3</sup> whilst the volume of pentoses and hexoses ranges from 107 to 129 Å<sup>3</sup>. In this first foldamer generation, control over absolute right (*P*) or left (*M*) helix handedness was not implemented. The sequence is symmetrical, and comprises two identical hemi-capsules that result in a mismatch with the targeted guests which lack any symmetry element. Sequence **1** and the subsequent generations were synthesised, typically on a 100 mg scale, following optimised protocols<sup>36</sup>.

### **Binding properties of 1<sup>st</sup> and 2<sup>nd</sup> generation hosts**

The intrinsic binding specificity of several monosaccharides to this starting state host molecule **1** was assessed in DMSO/CHCl<sub>3</sub> mixtures by using circular dichroism (CD) spectroscopy (**Fig. 3a,b**) which gives access to affinity constants ( $K_a$ ) and the extent of differential binding by *P* and *M* helices (diastereomeric excess, *d.e.*, see **Supplementary Fig. S41** for details) as the guest favours one handedness over the other. Note that CD spectroscopy does not however enable observation of the preference for  $\alpha$  vs  $\beta$  anomers, or pyranose vs furanose tautomers of the guest (**Fig. 2d**). Conformational selectivity of the guests instead relied on NMR spectroscopy, to be discussed below. Note that DMSO is capable of dissolving the guest molecules but also competes for hydrogen-bonds with the host. Consistent with this effect,  $K_a$  values were found to increase upon decreasing the fraction of DMSO (**Supplementary Table S1**). In 5:95 DMSO/CHCl<sub>3</sub> very tight 1:1 binding of most guests to **1** spanned two orders of magnitude ( $K_a \sim 10^3$ - $10^5$  L mol<sup>-1</sup>). For example,  $K_a = 165\,000$  L mol<sup>-1</sup> and 1900 L mol<sup>-1</sup> for fructose and ribose, respectively (**Supplementary Figs. S3, S9**). These

values matched those of the best synthetic receptors for monosaccharides in organic solvents<sup>20</sup> and largely exceeded those of other helical containers<sup>28,29</sup>. This result validated our design that the multiple hydrogen bond donors and acceptors provided by the capsule cavity should result in efficient monosaccharide binding. For the purpose of our investigation, the proportion of DMSO was adjusted to 20% to bring  $K_a$  values down to a range where they can be accurately measured not only by CD but also by NMR spectroscopy. In this solvent,  $K_a$  values were found to be in the  $10^2$ - $10^4$  L mol<sup>-1</sup> range (**Fig. 2e**). It was significant that even in this first iteration clear selectivity for certain monosaccharides was achieved despite the high level of similarity between the different sugar molecules that make monosaccharide discrimination particularly challenging even for their natural receptors<sup>37</sup>. Some preferences for either *P* or *M* helix conformation of the host molecule were also observed. From the CD spectroscopy data, binding was shown to be strongest for D-mannose, although with only a weak preference for the *M* helix (*d.e.* = 14%), and also strong for D-fructose with a marked preference for the *P* helix (*d.e.* = 72%). Binding was weaker for smaller pentose guests (**12**, **13**), and for D-glucose derivatives (**9**, **11**), which have the particular feature to be all-equatorial, at the exclusion of the anomeric OH group which equilibrate between axial ( $\alpha$ -**9a**) and equatorial ( $\beta$ -**9a**) positions. In order to further characterise the binding mechanisms and therefore provide a basis by which to optimise host-guest specificity, it was necessary to obtain structural information of the complex formation including the anomeric and tautomeric state of the guest (**Fig. 2d**). In solution, most **1**⊃guest complexes yielded complicated NMR spectra assigned to mixtures of  $\alpha$  and  $\beta$  anomers of the guest furanose and pyranose forms bound to both *P*-**1** and *M*-**1** as a result of small *d.e.* values.

In order to facilitate analysis by NMR spectroscopy, a second generation sequence **2** was introduced in which terminal (*1S*)-(-)-camphanyl groups R\* (**Fig. 2b**) quantitatively induce *P* helicity,<sup>35</sup> thus proscribing the formation of diastomeric complexes in the *M* helix.

The resulting NMR spectra were simpler and titrations could be carried out.  $K_a$  values for *P-2* measured by  $^1\text{H}$  NMR were similar to those measured by CD for **1** except for guests which had shown a strong preference for *M-1* and thus had a lower affinity for the *P* helix (**Fig. 2e**). Upon adding a guest to the empty host, a distinct set of signals of the host-guest complex emerged indicating slow guest capture and release on the NMR timescale. The multiplicity of these signals was the same as for the empty host showing that the host-guest complex retains an average  $C_2$  symmetry as the guest tumbles rapidly in the cavity at 25°C. Upon cooling, signals of the host broadened and split into two sets when tumbling of the guest becomes slow (**Fig. S42**). The furanose vs pyranose forms and  $\alpha$  vs  $\beta$  configurations of the guests were assessed using  $^1\text{H}$ - $^{13}\text{C}$  single quantum correlation NMR spectroscopy ( $^{13}\text{C}$ -HSQC) of  $^{13}\text{C}$ -labeled monosaccharides in the context of unlabelled host molecules. The acquired spectra revealed the expected coexistence of several forms or anomers of each free guest (**Supplementary Fig. S20**). Upon addition of **2**, some spectra remained complex. But in the case of **7**, **8** and **10**, the signal patterns radically simplified showing quantitative selection of a single encapsulated species assigned to  $\beta$ -D-fructopyranose ( $\beta$ -**7a**),  $\alpha$ -D-mannopyranose ( $\alpha$ -**8a**) and  $\beta$ -D-galactofuranose ( $\beta$ -**10b**), respectively (**Supplementary Figs. S20, S22, S26**). Unlike the furanose forms of fructose, the stabilisation of  $\beta$ -**7a** is uncommon; only four protein structures are reported in the PDB that sequester this form<sup>38</sup>.

### High resolution structures of early host-guest complexes

High resolution information (0.9-1.1 Å) was obtained *via* the structures of four **1**⊃guest complexes in the solid state that could be elucidated by crystallographic analysis (**Fig. 4**). Because of the large size of the host-guest complexes, these structures possess a fraction of solvent molecules much above that usually found in small molecule crystals. Solvent molecules and side chains are generally disordered. However, the position of the helix backbones and the binding modes of the sugars were determined with great accuracy. Details about the methods

used for structure refinement and the treatment of disordered side chains and solvent molecules can be found in the supplementary material. These represent the first examples of crystal structures of unsubstituted monosaccharides bound to synthetic receptors, which provide key insights into the role of shape and hydrogen bonds towards monosaccharide selectivity by the host<sup>39</sup>. To achieve these structures, co-crystallisation of the host and guest was facilitated by the high affinity constants, and by the use of racemic monosaccharides which, when mixed with *P/M-1*, produce centrosymmetric lattices<sup>40,41</sup>. In all structures, the D-carbohydrate was found in the helix having the handedness favoured by the guest in CD titrations. The structures of *P-1*⊃β-D-**7a** and *M-1*⊃α-D-**8a** confirmed the <sup>2</sup>C<sub>5</sub> β-pyranose and <sup>4</sup>C<sub>1</sub> α-pyranose sugar conformations assigned by NMR, respectively. As listed in Supplementary Tables S13 and S14 and shown in Figure 5, the structures revealed extensive arrays of up to 10 hydrogen bonds between the sugar hydroxyl groups and the inner wall of the helix, a number matching that found in lectin-monosaccharide complexes<sup>11</sup>. The arrays of interaction between the guests and their host, including the positioning of the hydroxyl protons, could be determined with a much greater accuracy than with any protein-saccharide complex described previously. Remarkably, the structures reveal that intramolecular hydrogen bonds between neighbouring hydroxyl groups that are typically observed within isolated monosaccharides structures<sup>38,42</sup> do not form. Instead, host-guest intermolecular hydrogen bonds prevail. The orientation of the sugars with respect to the host was found to vary greatly (**Fig. 4e-f**). The volume of the cavity was also found to vary, showing some induced fit of the host by the guest, but in all cases the cavity of **1** remains larger than needed to bind monosaccharides.

On average about two thirds of hydrogen bonds involve a hydroxyl proton donor on the sugar and only one third involve an amide proton donor on the capsule. Yet these proportions vary from 3:1 for fructose to 1:1 for mannose. Hydrogen bond acceptors on the capsule inner rim are naphthyridine and pyridine endocyclic nitrogen atoms, and one or both inner carbonyl

group of H units which are engaged in short hydrogen bonds ( $d(\text{O}\cdots\text{H}) \leq 2.0 \text{ \AA}$ ), except with the smaller guest xylose ( $d(\text{O}\cdots\text{H}) = 2.5 \text{ \AA}$ ). Thus, H units serve well the purpose for which they were developed, suggesting that other efficient hydrogen bond acceptors such as *N*-oxides would represent valuable additions. An interesting case of induced fit is observed in the complex with fructose (**Figs. 4a, 5a**) in which one of the two acyl-hydrazide functions is found in *gauche* conformation with both NH groups pointing towards the sugar whilst both carbonyl groups point out. This arrangement is attributed to two hydrogen bonds formed by the two acyl hydrazide protons with oxygen atoms of fructose. The details revealed by these structures validate the initial capsule design and illustrate that accurate prediction of saccharide binding modes have limitations which necessitate experimental atomic resolution data.

### Structure-based iterative evolution

Based on the available accurate structural information, we submitted oligomeric sequence **2** to iterative evolution to improve its binding properties and produce a more selective host. The strategy consisted in filling the host molecule inner space around a given guest so as to sterically exclude all other guests, while at the same time preserving key interactions with the targeted substrate (**Fig. 1c,d**). Fructose (**7**) was selected as a target having both high affinity and a strong diastereoselectivity for *P*-**1**. The structure of **1**⊃β-**7a** showed that most hydrogen bonds are established with one half of the helix and that the other half could be drastically reduced without altering these interactions. As mentioned above, helix symmetry breaking was also established in solution by the signal multiplicity of the low temperature NMR spectrum of **2**⊃β-**8a** (**Supplementary Fig. S42**). We thus proceeded to the simultaneous deletion of one H unit and to a *P*→*F* mutation in the same half-sequence. *F* units feature a fluorine atom that points towards the host cavity, thus filling a small space without engaging in strong attractive or repulsive interactions. Third foldamer generation **3** was then produced, which showed a slightly enhanced affinity for fructose and sharply reduced affinities for other sugars (**Fig. 2e**).

Affinity for mannose dropped by one order of magnitude; affinities for other guests collapsed. Interestingly, helix handedness preferences of the **3**-mannose and **3**-ribose complexes were reverted from those of the complexes with **1**. Due to the induced-fit capabilities of the foldamers, including helix inversion when handedness is not controlled by a terminal camphanyl group, and due to the structural variability of the sugars, the exact consequences of deleting one **H** unit from the sequence of **1** to produce **3** on binding saccharides other than fructose were hard to predict. NMR titrations with the *P* analogous sequence **4** confirmed the results obtained with **3**. As a slight drawback, sequence **4** has a poorer selectivity than **2** for the different forms of **7**, one of the furanose anomers being bound to some extent along with  $\beta$ -fructopyranose (**Supplementary Fig. S32**).

Further structural insights to direct the next iteration of this optimisation process were gathered in the solid state. A crystal structure of racemic **3** $\supset$  $\beta$ -**7a** was obtained and confirmed the preservation of most of the interactions seen within **1** $\supset$  $\beta$ -**7a** (**Fig. 6**). A remaining small void was spotted in the vicinity of the pyranose endocyclic methylene group facing one N unit (**Fig. 6a, 6c**). A mutation N $\rightarrow$ Q<sup>F</sup> was then implemented to further increase host-guest shape complementarity and racemic sequence **5** was produced along with its *P* analogue **6**. Titrations revealed a further enhanced selectivity for fructose which is now bound 300 times better than mannose despite the notorious difficulty to distinguish these two sugars (**Supplementary Fig. S61**)<sup>44</sup>. The selectivity of **6** for fructose is so high that, when placed in presence of 3 equiv. of **7**, **8**, **9**, **10**, **12**, and **13**, the formation of **6** $\supset$  $\beta$ -**7a** remains quantitative as judged by <sup>1</sup>H NMR (**Fig. 3f**) and by the calculated proportions of the different complexes (**Fig. 3g**). In contrast with the behaviour of **4**, no complex of **6** with forms of **7** other than  $\beta$ -**7a** was observed (**Fig. 3c-e**). A crystal structure of **5** $\supset$  $\beta$ -**7a** confirmed that the effect of the last mutation was a complete filling of the cavity space around fructose by the additional fluorine atom, revealing atomic scale shape complementarity between host and guest (**Fig. 6**). The occupancy factor of

the capsule cavity by  $\beta$ -**7a** reaches 65%, a value much above the common 55% reference<sup>45</sup> that reflects tight attractive interactions as well as shape complementarity.

Further characterisation of the final complex **6** $\supset$  $\beta$ -**7a** was performed by using NMR spectroscopy in order to observe the behaviour of the host-guest association in the active form. As a first step, extensive <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift assignments of **6** $\supset$  $\beta$ -**7a** were determined (**Supplementary Tables S2-S4** and **Supplementary Figs. 47-51**) prior to calculation of a high-resolution NMR structure of the complex. A final ensemble of 20 structures (**Supplementary Fig. S51**) were calculated from distance restraints measured on a sample of uniformly <sup>13</sup>C-labelled D-fructose bound to natural abundance host molecule. The use of <sup>13</sup>C-edited and -filtered NMR spectra allowed for the collection of 311 unique distance restraints (**Supplementary Table S5**), including 77 intermolecular restraints to accurately position the monosaccharide within the capsule cavity. Direct observation of hydroxyl protons by NMR spectroscopy confirmed the hydrogen bonding network predicted from the crystal structure, and overall there is a near perfect superposition between the ensemble of solution structures with the solid state crystallographic analysis (rmsd of  $0.06 \pm 0.04$  Å).

## CONCLUSION

The selection process described above thus provides a general protocol by which to generate new highly selective hosts for given guest molecules, starting from a slightly larger-than-needed cavity possessing first principles design features, and iteratively reducing cavity size whilst increasing shape complementarity with the guest. In this strategy, exploiting the modularity and crystal growth ability of aromatic amide foldamer sequences was found to greatly accelerate the emergence of function. Using fructose as a test-case, extensive and unprecedented structural information at the atomic level about saccharide recognition was eventually provided, shedding light on the binding mechanism of these intractable natural substrates. With their polar inner rim, the aromatic amino-acid monomers used in this study

are well suited to binding polar guests in organic solvents. For other families of guests, or to achieve binding in polar or even protic media, new building blocks may be required in order to bring hydrophobic moieties or charges into the capsule cavity. Efforts in this direction are currently being made in our laboratory<sup>46</sup>. Expanding our approach to other large and complex guest molecules would enable the fabrication of novel generations of selective receptors, sensors and transporters.

## REFERENCES

1. Lokey, R. S., Iverson, B. L. Synthetic molecules that fold into a pleated secondary structure in solution. *Nature* **375**, 303–305 (1995).
2. Appella, D. H. *et al.* Residue-based control of helix shape in  $\beta$ -peptide oligomers. *Nature* **387**, 381–384 (1997).
3. Nelson, J. C., Saven, J. G., Moore, J. S., Wolynes P. G. Solvophobic driven folding of nonbiological oligomers. *Science* **277**, 1793–1796 (1997).
4. Hamuro, Y., Geib, S. J., Hamilton, A. D., Oligoanthranilamides. Non-peptide subunits that show formation of specific secondary structure. *J. Am. Chem. Soc.* **118**, 7529–7541 (1996).
5. Cuccia, L. A., Lehn, J.-M., Homo, J.-C., Schmutz, M. Encoded helical self-organization and self-assembly into helical fibers of an oligoheterocyclic pyridine–pyridazine molecular strand. *Angew. Chem. Int. Ed.* **39**, 233–237 (2000).
6. Berl, V., Huc, I., Khoury, R., Krische, M., Lehn J.-M. Interconversion of single and double helices formed from synthetic molecular strands. *Nature* **407**, 720–723 (2000).
7. Zhu, J. A. *et al.* New class of folding oligomers: crescent oligoamides. *J. Am. Chem. Soc.* **122**, 4219–4220 (2000).
8. Guichard, G., Huc, I. Synthetic foldamers. *Chem. Commun.* **47**, 5933–5941 (2011).
9. White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan P. B. Recognition of the four Watson–Crick base pairs in the DNA minor groove by synthetic ligands. *Nature* **391**, 468–471 (1998).
10. Davis, A. P. Sticking to sugars. *Nature* **464**, 169–170 (2010).
11. Kubik, S. Synthetic lectins. *Angew. Chem. Int. Ed.* **48**, 1722–1725 (2009).
12. Huc, I. Aromatic oligoamide foldamers. *Eur. J. Org. Chem.* **1**, 17–29 (2004).
13. Zhang, D.-W., Zhao, X., Hou, J.-L., Li Z.-T. Aromatic amide foldamers: structures, properties, and functions. *Chem. Rev.* **112**, 5271–5316 (2012).

14. Klein, E., Ferrand, Y., Barwell, N. P., Davis A. P. Solvent effects in carbohydrate binding by synthetic receptors: implications for the role of water in natural carbohydrate recognition. *Angew. Chem. Int. Ed.* **47**, 2693–2696 (2008).
15. Ferrand, Y., Crump, M. P., Davis, A. P. A Synthetic lectin analog for biomimetic disaccharide recognition. *Science* **318**, 619–622 (2007).
16. Ferrand, Y., Klein, E., Barwell, N. P., Crump, M. P., Jiménez-Barbero, J., Vicent, C., Boons, G. J., Ingale, S., Davis A. P. A synthetic lectin for *O*-linked  $\beta$ -*N*-acetylglucosamine *Angew. Chem. Int. Ed.* **48**, 1775–1779 (2009).
17. James, T. D., Sandanayake, K. R. A. S., Shinkai, S. Chiral discrimination of monosaccharides using a fluorescent molecular sensor. *Nature* **374**, 345–347 (1994).
18. James, T. D., Phillips, M. D., Shinkai, S. Boronic acids in saccharide recognition. *RSC Cambridge* (2006).
19. Davis, A. P., Wareham, R. S. *Angew. Chem. Int. Ed.* **38**, 2978–2996 (1999).
20. Walker, D. B., Joshi, G., Davis A. P. Progress in biomimetic carbohydrate recognition. *Cell. Mol. Life Sci.* **66**, 3177–3191 (2009).
21. Kobayashi, K., Asakawa, Y., Kato Y., Aoyama Y. Complexation of hydrophobic sugars and nucleosides in water with tetrasulfonate derivatives of resorcinol cyclic tetramer having a polyhydroxy aromatic cavity: importance of guest-host CH- $\pi$  interaction. *J. Am. Chem. Soc.* **114**, 10307–10313 (1992).
22. Das, G., Hamilton, A. D. Molecular recognition of carbohydrates: strong binding of alkyl glycosides by phosphonate derivatives *J. Am. Chem. Soc.* **116**, 11139–11140 (1994).
23. Anderson, S., Neidlein, U., Gramlich, V., Diederich, F. A New Family of Chiral Binaphthyl-Derived Cyclophane Receptors: Complexation of Pyranosides *Angew. Chem. Int. Ed.* **34**, 1596–1600 (1995).
24. Davis, A. P., Wareham, R. S. A tricyclic polyamide receptor for carbohydrates in organic media. *Angew. Chem. Int. Ed.* **37**, 2270–2273 (1998).
25. Bitta, J., Kubik, S. Cyclic hexapeptides with free carboxylate groups as new receptors for monosaccharides *Organic Letters* **3**, 2637-2640 (2001).
26. Mazik, M., Cavga, H., Jones, P. G. Molecular recognition of carbohydrates with artificial receptors: mimicking the binding motifs found in the crystal structures of protein-carbohydrate complexes. *J. Am. Chem. Soc.* **127**, 9045–9052 (2005).
27. Ardá, A. *et al.* A chiral pyrrolic tripodal receptor enantioselectively recognizes  $\beta$ -mannose and  $\beta$ -mannosides. *Chem. Eur. J.* **16**, 414–418 (2010).
28. Inouye, M., Waki, M., Abe, H. Saccharide-dependent induction of chiral helicity in achiral synthetic hydrogen-bonding oligomers. *J. Am. Chem. Soc.* **126**, 2022–2027 (2004).

29. Hou, J.-L. *et al.* Hydrogen bonded oligohydrazide foldamers and their recognition for saccharides. *J. Am. Chem. Soc.* **126**, 12386–12394 (2004).
30. Pal, A., Bérubé, M., Hall, D. G. Design, synthesis, and screening of a library of peptidyl bis-boroxoles as low molecular weight receptors for complex oligosaccharides in water: identification of a receptor for the tumour marker TF-antigen. *Angew. Chem. Int. Ed.* **49**, 1492–1495 (2010).
31. Bao, C. *et al.* Converting sequences of aromatic amino acid monomers into functional three-dimensional structures: second-generation helical capsules. *Angew. Chem. Int. Ed.* **47**, 4153–4156 (2008).
32. Ajami, D., Rebek J. J. Compressed alkanes in reversible encapsulation complexes. *Nature Chem.* **1**, 87–90 (2009).
33. Sawada, T., Yoshizawa, M., Sato, S., Fujita M. Minimal nucleotide duplex formation in water through enclathration in self-assembled hosts. *Nature Chem.* **1**, 53–56 (2009).
34. Mugridge, J. S., Rudi van Eldik, A. Z., Bergman, R. G., Raymond, K. N. Solvent and pressure effects on the motions of encapsulated guests: tuning the flexibility of a supramolecular host. *J. Am. Chem. Soc.* **135**, 4299–4306 (2013).
35. Ferrand, Y. *et al.* Long-range effects on the capture and release of a chiral guest by a helical molecular capsule. *J. Am. Chem. Soc.* **134**, 11282–11288 (2012).
36. Qi, T., Deschrijver, T., Huc, I. Large-scale and chromatography-free synthesis of an octameric quinoline-based aromatic amide helical foldamer. *Nat. Protoc.* **8**, 693–708 (2013).
37. Lemieux, R. U. The origin of the specificity in the recognition of oligosaccharides by proteins. *Chem. Soc. Rev.* **18**, 347–374 (1989).
38. Cocinero, E. J. *et al.* Free fructose is conformationally locked. *J. Am. Chem. Soc.* **135**, 2845–2852 (2013).
39. A structure with octyl-glucoside has been reported: see ref. 26.
40. Pentelute, B. L. *et al.* X-ray structure of snow flea antifreeze protein determined by racemic crystallization of synthetic protein enantiomers. *J. Am. Chem. Soc.* **130**, 9695–9701 (2008).
41. Lautrette, G. *et al.* Structure elucidation of host-guest complexes of tartaric and malic acid by quasi-racemic crystallography. *Angew. Chem. Int. Ed.* **52**, 11517–11520 (2013).
42. Çarcabal, P. *et al.* Hydrogen bonding and cooperativity in isolated and hydrated sugars: mannose, galactose, glucose, and lactose. *J. Am. Chem. Soc.* **127**, 11414–11425 (2005).
43. Laskowski, R.A. *J. Mol. Graph.* **13**, 323–330 (1995).
44. Doores, K. J. *et al.* A nonself sugar mimic of the HIV glycan shield shows enhanced antigenicity. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 17107–17112 (2010).
45. Mecozzi, S., Rebek Jr., J. The 55% solution: a formula for molecular recognition in the liquid state. *Chem. Eur J.* **4**, 1016-1022 (1998).

46. Singleton, M. L., et al. Synthesis of 1,8-diazaanthracenes as building blocks for internally functionalized aromatic oligoamide foldamers. *J. Org. Chem.* **79**, 2115–2122 (2014).
47. Corbett, P. T., Otto, S., Sanders, J. K. M. Correlation between host-guest binding and host amplification in simulated dynamic combinatorial libraries. *Chem. Eur. J.* **10**, 3139–3143 (2004).

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

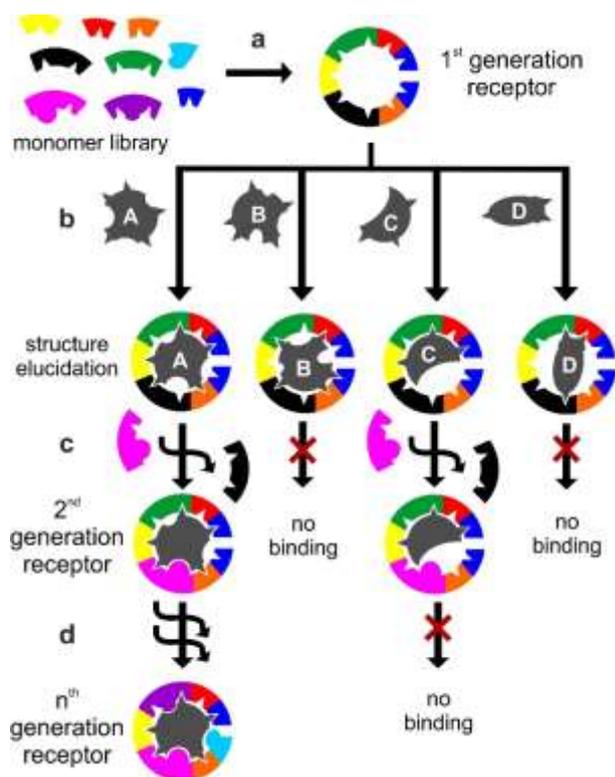
**Acknowledgments** This work was supported by the Agence Nationale de la Recherche (Project No. ANR-09-BLAN-0082-01 to NC), by the Conseil Interprofessionnel du Vin de Bordeaux (predoctoral fellowship to GL), and by the European Research Council under the European Union's Seventh Framework Programme (Grant Agreement No. ERC-2012-AdG-320892). The authors thank Dr. Amol Kendhale, and Ms. Catherine Blum for preliminary investigations on the preparation of "H" monomers, the TGIR-RMN-THC Fr3050 CNRS for high-field NMR time, Jean-Luc Ferrer for beamtime and help during data collection on FIP-BM30A at the European Synchrotron Radiation Facility, and Piotr Nowak (Univ. Gröningen) for calculations using the DCLSim software.

**Author Contributions** N.C. and G.L. synthesized all new compounds. N.C., Y.F., G.L., C.D.M carried out solution studies. B.K. collected X-ray data and solved the crystal structures. M.L. performed modelling studies. I.H., D.D. and Y.F. designed the study. I.H. and Y.F. wrote the manuscript. All authors discussed the results and commented on the manuscript.

**Additional Information** The crystallographic data and experimental details of the structural refinement for the X-ray crystal structures reported in this paper, **1 $\beta$ -7a**; **1 $\alpha$ -8a**; **1 $\beta$ -9a**; **1 $\alpha$ -13a**; **3 $\beta$ -7a** and **5 $\beta$ -7a**, have been deposited at the Cambridge Crystallographic Data Centre, under deposition numbers CCDC 999559, CCDC 999560, CCDC 999617, CCDC 999616, CCDC 999615 and CCDC 999614, respectively. These data can be obtained free of charge from the Cambridge Crystallographic Data Centre

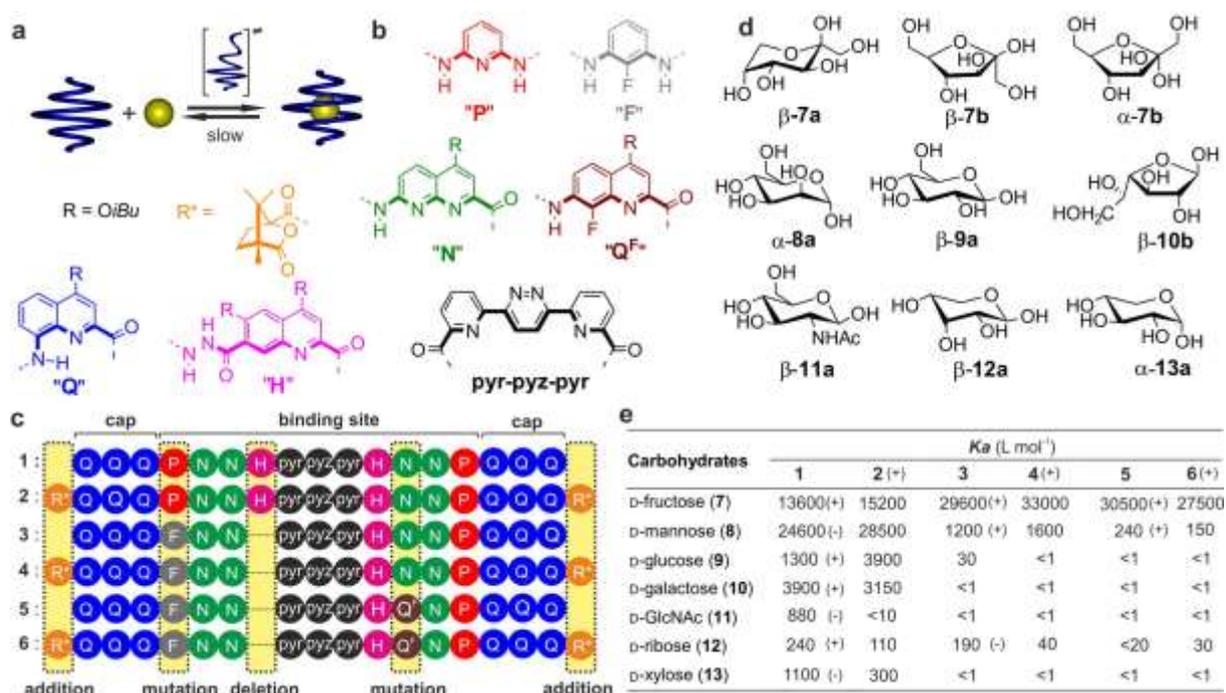
([http://www.ccdc.cam.ac.uk/data\\_request/cif](http://www.ccdc.cam.ac.uk/data_request/cif)). Reprints and permissions information is available at <http://www.nature.com/reprints>. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to I.H. ([i.huc@iecb.u-bordeaux.fr](mailto:i.huc@iecb.u-bordeaux.fr)).

## FIGURE LEGENDS



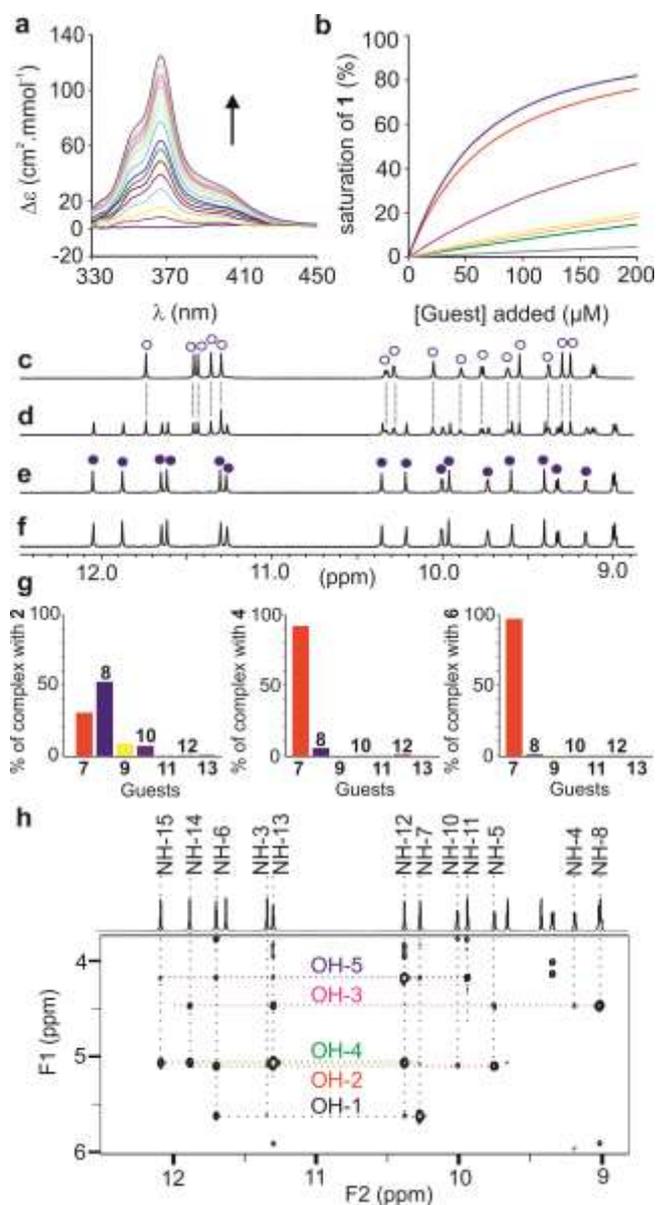
**Figure 1. Schematic representation of structure-based iterative evolution of a foldamer sequence.** **a**, A set of building block monomers are selected from those available. Based on molecular modelling and first principle positive design elements, a 1<sup>st</sup> generation sequence is proposed and synthesized so as to create a cavity slightly larger than needed for, and able to engage in multiple favourable interactions with, a family of guest molecules. Accurate prediction of these interactions and of selectivity is not achieved at this stage. **b**, Host-guest interactions are assessed and the structure of host-guest complexes are determined for a series of guests (A, B, C, D...). **c**, Based on the results of step **b**, a target guest is selected, and a 2<sup>nd</sup> generation sequence is proposed using negative design elements that are meant to exclude other

guests while preserving interactions with the target. Typically, an appropriate mutation, deletion or addition is implemented in the sequence in order to fill space around the guest so as to achieve atomic scale complementarity. **d**, Through iterative repetition of steps **b** and **c**, the sequence evolves and quickly converges to a highly selective receptor for the target.



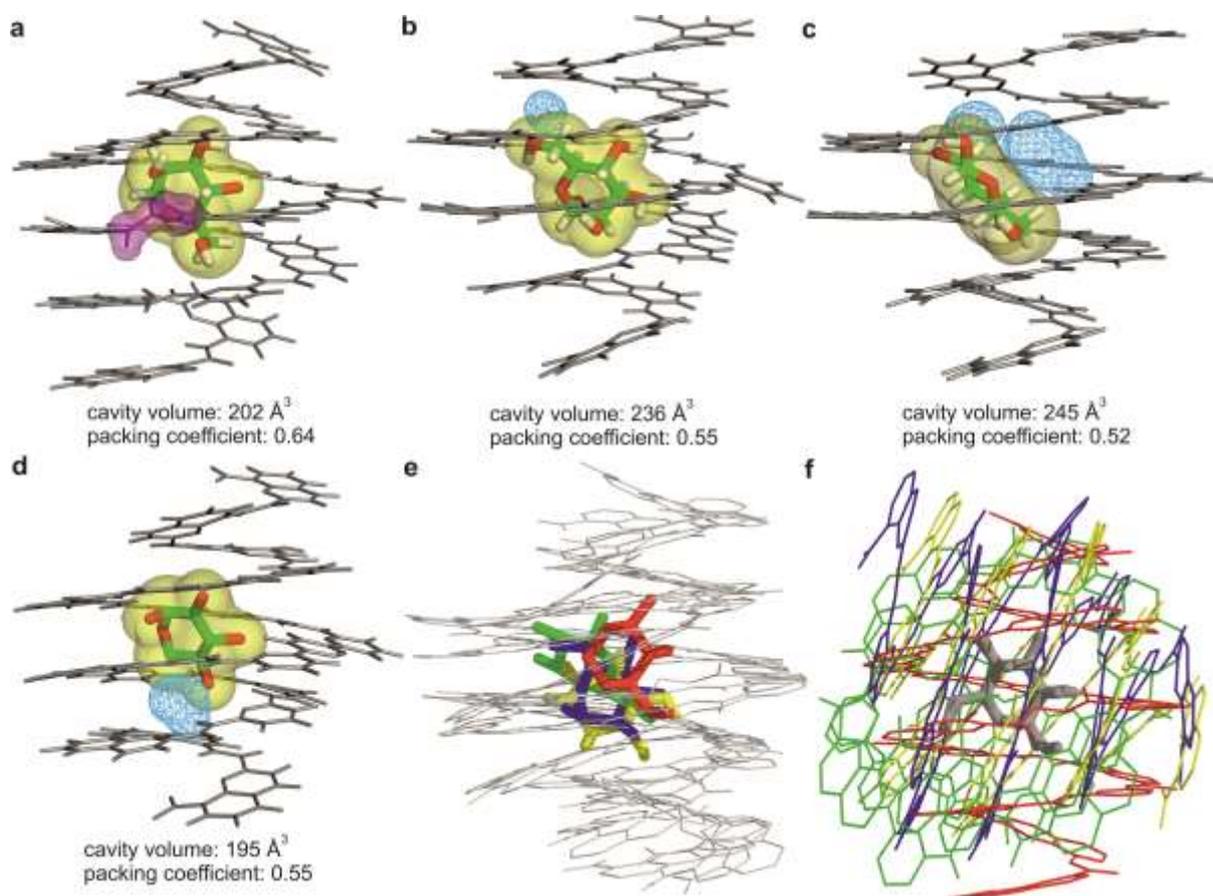
**Figure 2. Host-guest components and assembly.** **a**, Schematic representation of the encapsulation of a guest (yellow sphere) in a molecular helical container (in blue) with a reduced diameter at both ends. **b**, Color-coded formulas and associated letters of amino acid, diamino and diacid monomers. The inner rim of the helix is marked by thick bonds. **c**, Sequence alignment of six generations. The two terminal Q units in **1**, **3** and **5** possess a nitro group in position 8. **d**, Formulas of some monosaccharide guests. Three dominant forms of D-fructose (**7**) are shown: the  $\beta$ -pyranose ( $\beta$ -7a),  $\alpha$ -furanose ( $\alpha$ -7b) and  $\beta$ -furanose ( $\beta$ -7b). The  $\alpha$ -pyranose tautomer (not shown) is a minor species. Only one form is shown for other guests (D enantiomers). All exist as a mixture of  $\alpha/\beta$ -pyranoses in solution, **10**, **12** and **13** also forming  $\alpha/\beta$ -furanoses. **e**, Host-guest association constants determined in 4:1 CDCl<sub>3</sub>/[D<sub>6</sub>]-DMSO vol/vol at 298 K.  $K_a$  values are determined by CD for **1**, **3**, and **5**, and by <sup>1</sup>H NMR for **2**, **4**, and

6. The (+) or (–) signs indicate whether a *P* or *M* helicity is induced by the guest, respectively. Absolute assignment of helix handedness has been previously determined<sup>35</sup>.



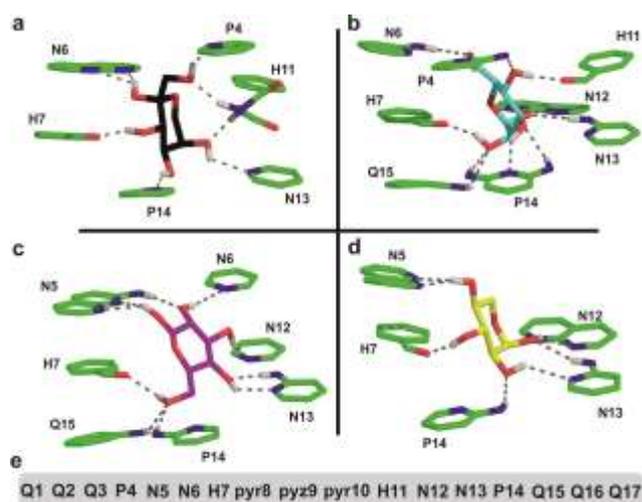
**Figure 3. Spectroscopic assessment of host-guest association.** **a**, Example of variation of the CD spectrum of capsule **1** upon the addition of **D-7** in  $\text{CDCl}_3$ : $[\text{D}_6]$ -DMSO 80:20 (vol/vol) at 298 K. **b**, Curve fittings of the CD titrations of **1** with **D-7** (red), **D-8** (blue), **D-9** (yellow), **D-10** (purple), **D-11** (grey), **D-12** (pink) and **D-13** (green). **c-f**, Part of the 700 MHz  $^1\text{H}$  NMR spectra of **6** (1mM) in  $\text{CDCl}_3$ : $[\text{D}_6]$ -DMSO 80:20 (vol/vol) at 298 K in the presence of: **c**, 0 equiv.; **d**, 0.5 equiv.; **e**, 2 equiv. of **D-fructose 7**; **f**, 3 equiv. each of **7**, **8**, **9**, **10**, **12** and **13**. Signals of the

empty host and of the host–guest complex are marked with empty blue, and blue circles, respectively. **g**, Bar diagrams showing the proportions of host guest complexes in a mixture of either **2**, **4** or **6** (1 mM) and 3 equiv. each of **7**, **8**, **9**, **10**, **11**, **12** and **13**. Proportions were calculated based on measured  $K_a$  values reported in **Fig. 2e** using the DCLSim software<sup>47</sup>. The same color code is used as in **b**. **h**, Expansion of the  $^1\text{H}$ - $^1\text{H}$  double-filtered NOESY spectrum recorded with 150 ms mixing time of **6** $\supset$ [ $^{13}\text{C}$ ]**7a** in  $\text{CDCl}_3/[\text{D}_6]\text{-DMSO}$  (95/5) at 298 K showing through-space NOE cross peaks ( $< 6 \text{ \AA}$ ) between the capsule and the hydroxyls of the sugar.



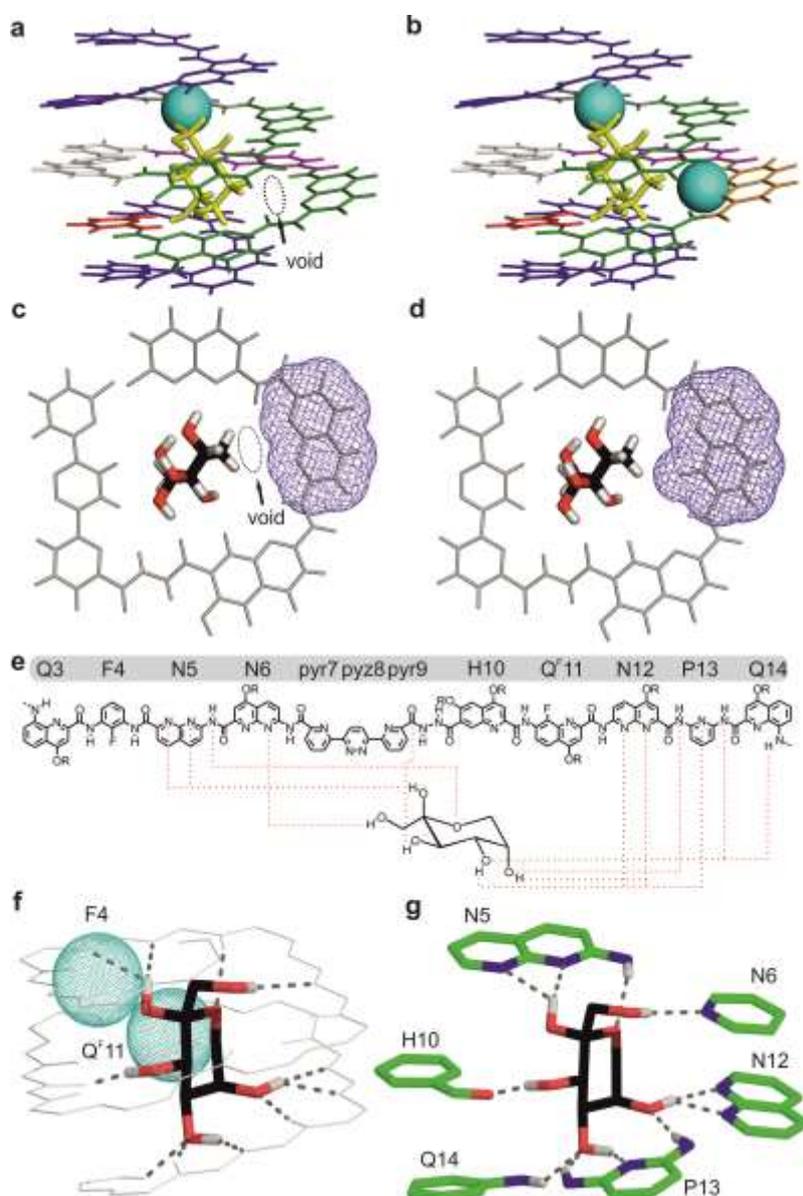
**Figure 4.** Crystal structures of complexes of first generation capsule *P-1* with: **a**,  $\beta\text{-}^2\text{C}_5\text{-D-fructopyranose}$ ; **b**,  $\alpha\text{-}^1\text{C}_4\text{-L-mannopyranose}$ ; **c**,  $\beta\text{-}^4\text{C}_1\text{-D-glucopyranose}$ ; **d**,  $\alpha\text{-}^1\text{C}_4\text{-L-xylopyranose}$ . The structures belong to centrosymmetrical space groups and thus the lattices also contain their enantiomers *M-1* $\supset\beta\text{-}^5\text{C}_2\text{-L-fructopyranose}$ , *M-1* $\supset\alpha\text{-}^4\text{C}_1\text{-D-mannopyranose}$ ,

$M-1 \supset \beta\text{-}^1\text{C}_4\text{-L-glucopyranose}$  and  $M-1 \supset \alpha\text{-}^4\text{C}_1\text{-D-xylopyranose}$ , respectively. Details about the sugar binding modes can be found in **Fig. 5**. The backbones of the capsules are represented in grey (tube) whereas the guest is shown in thick tube representation. The volume of the guest is shown using transparent yellow isosurfaces whereas the free space within the cavity is shown in cyan triangle mesh. In **a**, the gauche conformation of one of the two hydrazide bonds is highlighted with a transparent purple isosurface. Superimposition of foldamer-sugar complexes with: **e**, the central pyridazine rings (pyz) serving as templates for alignment and **f**, the six membered pyranose rings of the carbohydrates serving as templates for alignment. In **e**, the helix backbone is shown as light grey tube whereas  $\beta\text{-D-fructopyranose}$ ,  $\alpha\text{-L-mannopyranose}$ ,  $\beta\text{-D-glucopyranose}$  and  $\alpha\text{-L-xylopyranose}$  are shown in red, blue, green and yellow color, respectively. In **f**, the overlaid sugars are in grey tube whereas the backbones of the helices are shown using the same color code as above. Volumes ( $\text{\AA}^3$ ) of the cavity, the guest and the free space are determined using SURFNET v1.4<sup>43</sup>. Only the volumes higher than  $10 \text{\AA}^3$  are shown. The packing coefficient is defined here as the ratio of the guest volume to the host volume. Isobutoxy side chains and solvent molecules were omitted for clarity.



**Figure 5. Sugar binding modes with first generation capsule *P-1*:** **a**,  $\beta\text{-}^2\text{C}_5\text{-D-fructopyranose}$ ; **b**,  $\alpha\text{-}^1\text{C}_4\text{-L-mannopyranose}$ ; **c**,  $\beta\text{-}^4\text{C}_1\text{-D-glucopyranose}$ ; **d**,  $\alpha\text{-}^1\text{C}_4\text{-L-xylopyranose}$ . **e**, numbering of the units of sequence **1** used in this Figure. The sugars and those

heterocycles that interact with them are shown in tube representation. Dashed lines indicate hydrogen bonds. Details of these hydrogen bonds (distances, angles) can be found in the Supplementary material.



**Figure 6. Iterative evolution.** Crystal structures of: **a**, capsule *P-3* with  $\beta$ -<sup>2</sup>C<sub>5</sub>-D-fructopyranose and **b**, capsule *P-5* with  $\beta$ -<sup>2</sup>C<sub>5</sub>-D-fructopyranose. Top view of a slice of **c**, capsule *P-3* and **d**, *P-5* illustrating the volume reduction due to the N $\rightarrow$ Q<sup>F</sup> mutation. Volumes of each N and Q<sup>F</sup> monomer are shown as blue quad mesh. **e**, Formula and monomer numbering of sequence **5** together with the structures of  $\beta$ -D-fructopyranose. The two distal quinoline

monomers have been removed for clarity. Note that monomer numbering differ from that of **Fig. 5** which correspond to sequence **1**. Hydrogen bonds from the capsule wall to the sugar are shown as red dashed lines. **f** and **g**, views of the crystal structure of *P-5*β-D-fructopyranose showing the array of hydrogen bonds. In **f**, only the inner rim of the capsule in white tube representation is shown. The eleven hydrogen bonds are shown as yellow dashes. Fluorine atoms represented as dotted spheres. In **g**, the sugar and those heterocycles that interact with it are shown in tube representation. Dashed lines indicate hydrogen bonds. Details of these hydrogen bonds (distances, angles) can be found in the Supplementary material