ARTICLES
PUBLISHED ONLINE: 16 MARCH 2015 | DOI: 10.1038/NCHEM.2195
nature chemistry

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

Nagula Chandramouli[1,2], Yann Ferrand[1,2], Guillaume Lautrette[1,2], Brice Kauffmann[3,4,5], Cameron David Mackereth[6,7], Michel Laguerre[1,2], Didier Dubreuil[8,9] and Ivan Huc[1,2]*

The *ab initio* design of synthetic molecular receptors for a specific biomolecular guest remains an elusive objective, particularly for targets such as monosaccharides, which have very close structural analogues. Here we report a powerful approach to produce receptors with very high selectivity for specific monosaccharides and, as a demonstration, we develop a foldamer that selectively encapsulates fructose. The approach uses an iterative design process that exploits the modular structure of folded synthetic oligomer sequences in conjunction with molecular modelling and structural characterization to inform subsequent refinements. 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 crystallize, a sequence that binds to β-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 for the development of selective receptors.

The primary structures of proteins and nucleic acids consist of a modular arrangement in which multiple monomeric units are connected in long sequences. The fine-tuning of structures and folded conformations can be generated via additions, deletions or mutations of monomers, or even the recombination of segments or entire domains. This modular arrangement enables the evolution of polymer sequences without having to modify synthetic or biosynthetic schemes. In recent years, chemists have shown that synthetic molecular strands with backbones chemical remote from those of peptides and nucleotides may also fold into well-defined structures in solution[1–7]. Even in the absence of dedicated enzymatic systems to replicate them, these artificial backbones, termed synthetic foldamers[8], 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 recognize double-stranded DNA in a sequence-selective manner[9]. 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[10,11]. In just a few iterations, an initial sequence with relatively poor guest selectivity evolved into a receptor with fully characterized atomic-scale complementarity for β-D-fructopyranose. This evolutionary approach thus emerges as a novel method to tailor receptors for predefined complex guests. Schematic views of the successive steps of our method are presented in Fig. 1. In practice, the method requires the availability of monomers for which the monomer sequence alone provides 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[12,13].

The molecular recognition of saccharides by both natural and artificial receptors is seriously challenged by the issues of affinity and selectivity. The difficulty in defining high affinities in water or polar solvents is a result of competing hydrogen bonds between the saccharides and solvent molecules[14]. This is less acute for sugars with well-defined hydrophobic patches, such as the multiple axial protons of the 'all-equatorial' glucose derivatives[15,16]. 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[17,18]. Multiple hydrogen bonds to hydroxyl groups also elicit binding that can be very strong in less polar media[19].

Challenges in selectivity arise from the fact that saccharides inherently resemble one another, and most co-exist in several forms at equilibrium. For example, exchange between α and β anomers, or between furanose and pyranose tautomers, further increases the chance of one given saccharide resembling another. Resolving these issues is important not just for proteins and other biopolymers, but also for practical applications in synthetic receptors, as minor saccharide structural differences are often biologically meaningful. Despite the numerous synthetic receptors of saccharides described to date[20–29], progress towards high selectivity has been modest except in relation to all-equatorial glucose derivatives[15,16,20], which have a relatively unique disc-like shape. A few

[1]University of Bordeaux, CBMN (UMR 5248), Institut Européen de Chimie et Biologie, 2 rue Escarpit, Pessac 33600, France. [2]CNRS, CBMN (UMR 5248), France. [3]University of Bordeaux, IECB (UMS 3033), Institut Européen de Chimie et Biologie, 2 rue Escarpit, Pessac 33600, France. [4]CNRS, IECB (UMS 3033), France. [5]INSERM, IECB (US 001), France. [6]University of Bordeaux, ARNA (U 869), Institut Européen de Chimie et Biologie, 2 rue Escarpit, Pessac 33600, France. [7]INSERM, ARNA (U 869), France. [8]University of Nantes, CEISAM (UMR 6230), Faculté des Sciences et des Techniques, 2 rue de la Houssinière, BP 92208, Nantes Cedex 3 44322, France. [9]CNRS, CEISAM (UMR 6230), France. *e-mail: i.huc@iecb.u-bordeaux.fr

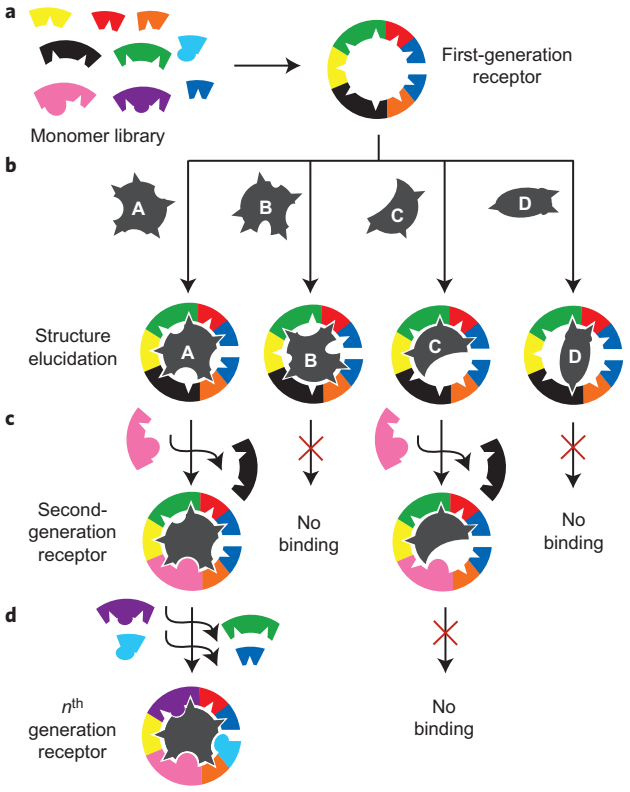

Figure 1 | Schematic representation of structure-based iterative evolution of a foldamer sequence. a, A set of building-block monomers are selected from those available. Based on molecular modelling and first-principle positive design elements, a first-generation sequence is proposed and synthesized 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 second-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 to fill space around the guest to achieve atomic-scale complementarity. **d**, Through iterative repetition of the steps in **b** and **c**, the sequence evolves and quickly converges to a highly selective receptor for the target.

foldamer-based saccharide receptors have been reported in which a sugar is bound in the cavity formed by a helically folded oligomer[28,29]. 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 by screening libraries, but rational improvements are difficult to implement in such cases[30].

In this Article we explore a novel and original structure-based iterative approach that paves the way to a general strategy aiming to discriminate between similar sugars by subtle evolution of a parent heterocyclic foldamer sequence (first-generation receptor, Fig. 1). Our investigation focused on the difficult distinction of fructose from other sugars, in particular the closely resembling mannopyranse structure. The outcome was unmatched selectivity.

Results and discussion

Design principles of the initial host. Following well-established rules (Supplementary Fig. 1)[31,32], a first-generation aromatic oligoamide sequence **1** (Fig. 2b,d) was designed to fold into a helix with 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 has a smaller cavity than capsules formed by metal- or hydrogen-bond-mediated self-assembly of organic building blocks but has the advantage of having no required symmetry[33–35], a desirable property for selectively binding to complex guests. Similar to other folded aromatic oligoamides, local conformational preferences at aryl–amide linkages and intramolecular π–π interactions between aromatic monomers combine to endow the folded conformation with high stability and favourable crystallogenesis[12,13]. The contribution of each monomer to the helix curvature can be predicted with good accuracy based on general force fields and simple energy minimization, 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[36]. 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 (12) and acceptors (20) and their overall even distribution was predicted to be sufficient to engage in the 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 $Å^3$ (the volumes of pentoses and hexoses range from 107 to 129 $Å^3$).

In this first foldamer generation, control over absolute right (*P*) or left (*M*) helix handedness is 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 synthesized, typically on a 100 mg scale, following optimized protocols[37].

Binding properties of first- and second-generation hosts. The intrinsic binding specificities of several monosaccharides to this starting-state host molecule **1** were assessed in DMSO/$CHCl_3$ mixtures 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. 41 for details) as the guest favours one handedness over the other. Note that CD spectroscopy does not enable observation of the preference for α versus β anomers or pyranose versus furanose tautomers of the guest (Fig. 2c). Determination of the selectivity of a host for a tautomer or an anomer of a guest instead relies on NMR spectroscopy (discussed below). Importantly, 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 on decreasing the fraction of DMSO (Supplementary Table 1). In 5:95 DMSO/$CHCl_3$, very tight 1:1 binding of most guests to **1** spanned two orders of magnitude ($K_a \approx 10^3$–10^5 l mol^{-1}). For example, for fructose and ribose, K_a =165,000 l mol^{-1} and 1,900 l mol^{-1}, respectively (Supplementary Figs 3 and 9). These values match those of the best synthetic receptors for monosaccharides in organic solvents[20], and largely exceed those of other helical containers[28,29]. This result validates our proposal 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 the 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

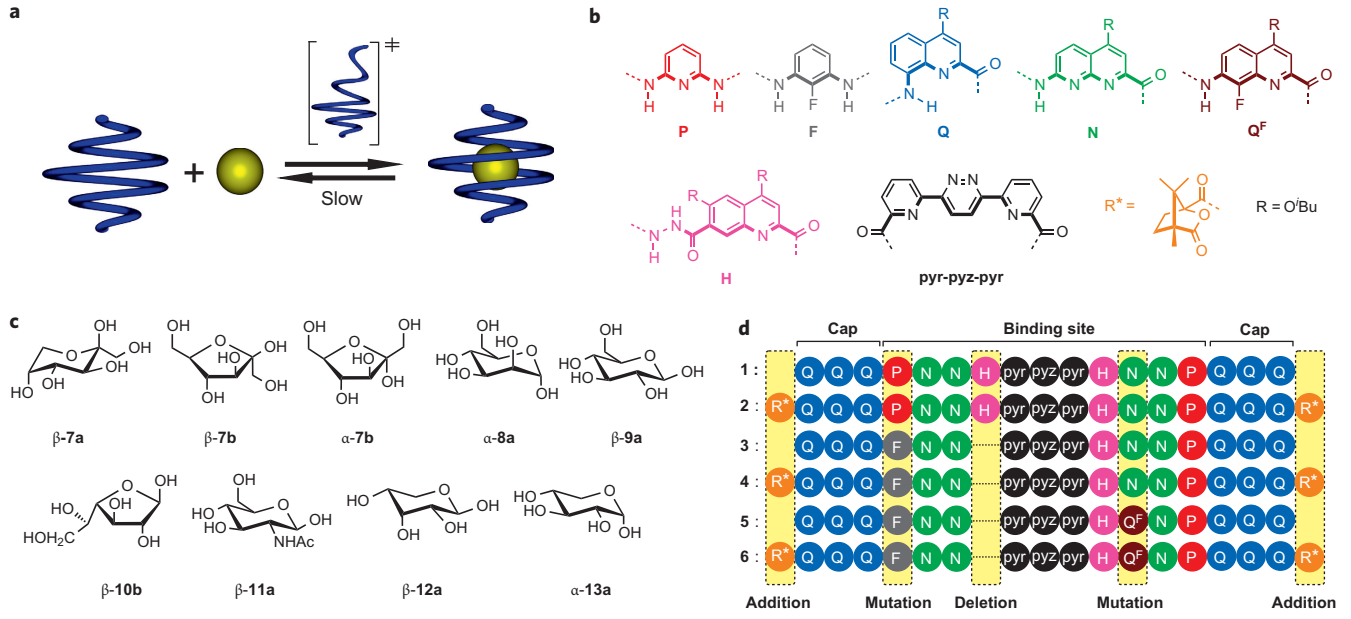

Figure 2 | Host–guest components and assembly. a, Schematic representation of the encapsulation of a guest (yellow sphere) in a molecular helical container (in blue) with a reduced diameter at both ends. **b**, Colour-coded formulae and associated letters of amino acid, diamino and diacid monomers. The inner rim of the helix is marked by thick bonds. **c**, Formulae of some monosaccharide guests. Three dominant forms of D-fructose (**7**) are shown: the β-pyranose (β-**7a**), α-furanose (α-**7b**) and β-furanose (β-**7b**). The α-pyranose tautomer (not shown) is a minor species. Only one form is shown for other guests (D enantiomers). All exist as a mixture of α/β-pyranoses in solution, **10**, **12** and **13** also forming α/β-furanoses. **d**, Sequence alignment of six generations. The two terminal Q units in **1**, **3** and **5** possess a nitro group at position 8.

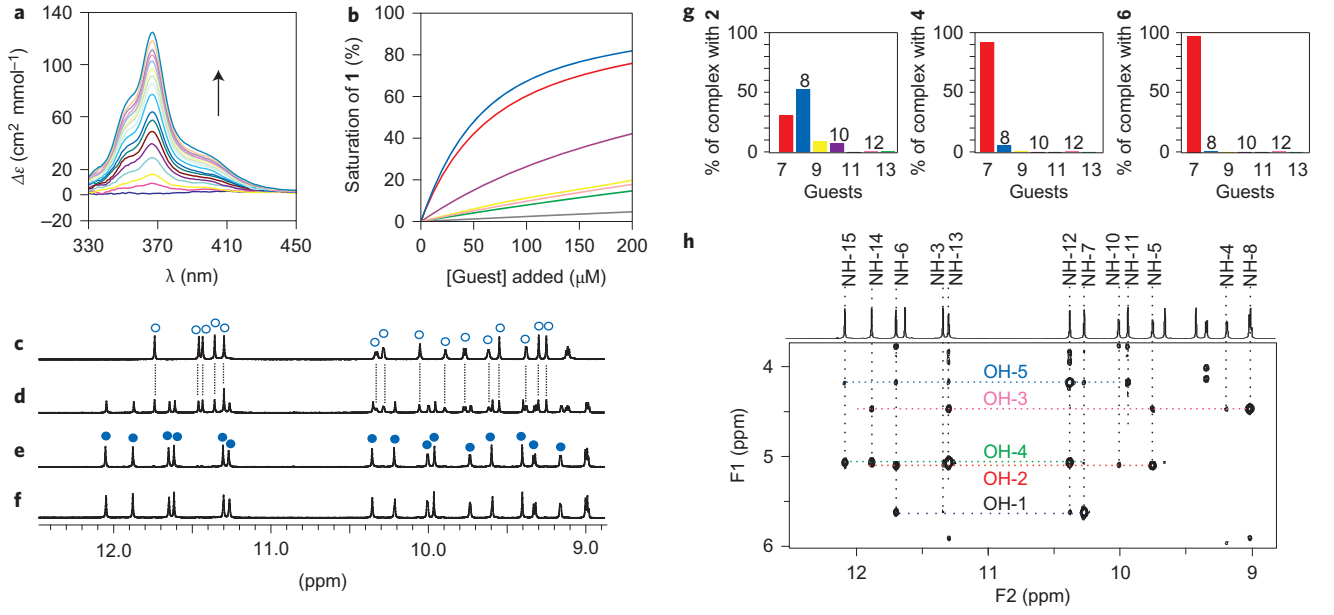

Figure 3 | Spectroscopic assessment of host–guest association. a, Example of variation of the CD spectrum of capsule **1** with the addition of D-**7** in $CDCl_3$:[D_6]-DMSO 80:20 (vol/vol) at 298 K. **b**, Curve fittings of the CD titrations of **1** with D-**7** (red), D-**8** (blue), D-**9** (yellow), D-**10** (purple), D-**11** (grey), D-**12** (pink) and D-**13** (green). **c**–**f**, Part of the 700 MHz ^{1}H NMR spectra of **6** (1 mM) in $CDCl_3$:[D_6]-DMSO 80:20 (vol/vol) at 298 K in the presence of 0 equiv. (**c**), 0.5 equiv. (**d**), 2 equiv. (**e**) of D-fructose **7**, and 3 equiv. each of **7**, **8**, **9**, **10**, **12** and **13** (**f**). Signals of the empty host and of the host–guest complex are marked with open and filled 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 the measured K_a values reported in Table 1 using DCLSim software[47]. The same colour code is used as in **b**. **h**, Expansion of the ^{1}H–^{1}H double-filtered NOESY spectrum recorded with 150 ms mixing time of **6** ⊃ [^{13}C]β-**7a** in $CDCl_3$/[D_6]-DMSO (95/5) at 298 K showing through-space NOE cross peaks (<6 Å) between the capsule and the hydroxyls of the sugar.

the 10^2–10^4 l mol^{-1} range (Table 1). It is 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[38]. Some preferences for either *P* or *M* helix conformation of the host

Table 1 | Host–guest association constants determined in 4:1 $CDCl_3$/[D_6]-DMSO vol/vol at 298 K.

Carbohydrates	K_a (l mol^{-1})					
	1	**2 (+)**	**3**	**4 (+)**	**5**	**6 (+)**
D-fructose (**7**)	13,600(+)	15,200	29,600(+)	33,000	30,500(+)	27,500
D-mannose (**8**)	24,600(–)	28,500	1,200 (+)	1,600	240 (+)	150
D-glucose (**9**)	1,300 (+)	3,900	30	<1	<1	<1
D-galactose (**10**)	3,900 (+)	3,150	<1	<1	<1	<1
D-GlcNAc (**11**)	880 (–)	<10	<1	<1	<1	<1
D-ribose (**12**)	240 (+)	110	190 (–)	40	<20	30
D-xylose (**13**)	1,100 (–)	300	<1	<1	<1	<1

K_a values are determined by CD for **1**, **3** and **5** and by ^{1}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 determined previously[35].

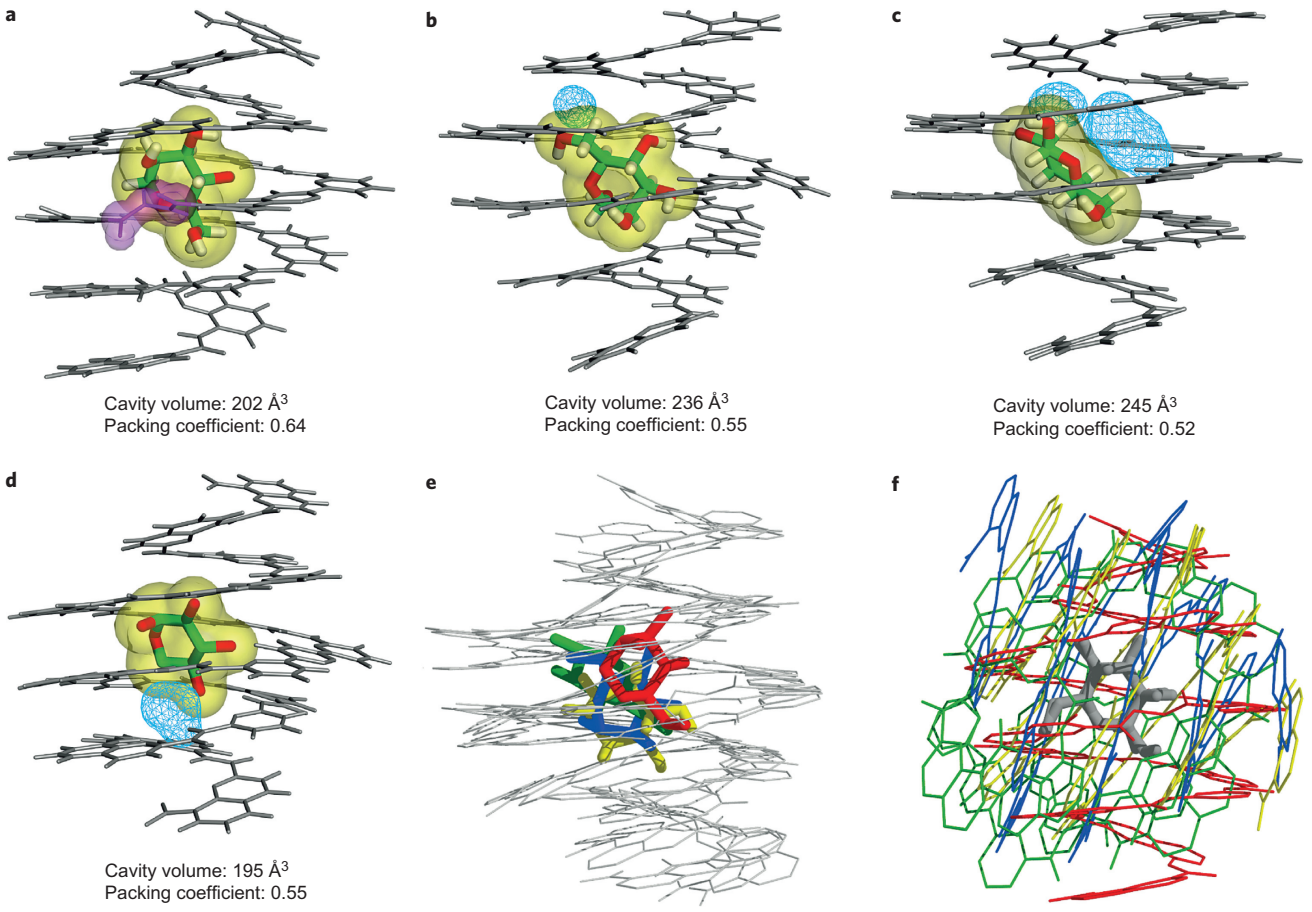

Figure 4 | Crystal structures of complexes of first-generation capsule *P*-1. a–d, Structures with β-$^2C_5$-D-fructopyranose (**a**), α-$^1C_4$-L-mannopyranose (**b**), β-$^4C_1$-D-glucopyranose (**c**) and α-$^1C_4$-L-xylopyranose (**d**). The structures belong to centrosymmetrical space groups and thus the lattices also contain their enantiomers *M*-**1** ⊃ β-$^5C_2$-L-fructopyranose, *M*-**1** ⊃ α-$^4C_1$-D-mannopyranose, *M*-**1** ⊃ β-$^1C_4$-L-glucopyranose and *M*-**1** ⊃ α-$^4C_1$-D-xylopyranose, respectively. Capsule backbones are represented in grey. Guests are shown in a thick tube representation. Guest volume is shown using transparent yellow isosurfaces. Free space within the cavity is shown as a cyan triangle mesh. In **a**, the *gauche* conformation of one of the two hydrazide bonds is highlighted with a transparent purple isosurface. **e**,**f**, Superimposition of foldamer–sugar complexes with the central pyridazine rings (pyz) serving as templates for alignment (**e**) and the six-membered pyranose rings of the carbohydrates serving as templates for alignment (**f**). In **e**, the helix backbone is shown in grey whereas β-D-fructopyranose, α-L-mannopyranose, β-D-glucopyranose and α-L-xylopyranose are shown in red, blue, green and yellow, respectively. In **f**, the overlaid sugars are in grey whereas the backbones of the helices are shown using the same colour code as above. Volumes (Å^3) of the cavity, the guest and the free space are determined using SURFNET v1.4[48]. Only volumes greater than 10 Å^3 are shown. The packing coefficient is defined as the ratio of guest volume to host volume. Isobutoxy side chains and solvent molecules are omitted for clarity.

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 the smaller pentose guests (**12**, **13**), as well as for the D-glucose derivatives (**9**, **11**), which have the particular feature of being all-equatorial, with the exception of the anomeric OH group, which equilibrates between axial (α-**9a**) and equatorial (β-**9a**) positions. To further characterize the binding mechanisms and therefore provide a

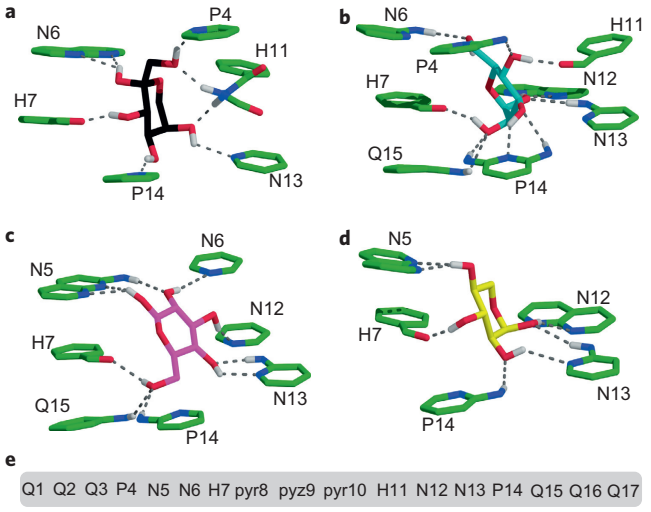

Figure 5 | Sugar-binding modes with first-generation capsule *P*-1. **a**, β-$^{2}C_{5}$-D-fructopyranose. **b**, α-$^{1}C_{4}$-L-mannopyranose. **c**, β-$^{4}C_{1}$-D-glucopyranose. **d**, α-$^{1}C_{4}$-L-xylopyranose. **e**, Numbering of the units of sequence **1** used in this figure. The sugars and those heterocycles that interact with them are shown in a tube representation. Dashed lines indicate hydrogen bonds. Details of these hydrogen bonds (distances, angles) can be found in the Supplementary Table 14.

basis on which to optimize host–guest specificity, it was necessary to obtain structural information about the complex formation, including the anomeric and tautomeric state of the guest (Fig. 2c). In solution, most **1** ⊃ guest complexes yielded complicated NMR spectra assigned to mixtures of α and β anomers of the guest furanose and pyranose forms bound to both *P*-**1** and *M*-**1**, as a result of small d.e. values.

To facilitate analysis by NMR spectroscopy, second-generation sequence **2** was introduced. In this sequence, terminal (1*S*)-(–)-camphanyl groups R* (Fig. 2b) quantitatively induce *P* helicity[36], thus proscribing the formation of diasteromeric complexes in the *M* helix. The resulting NMR spectra are simpler and titrations could be carried out. The K_a values for *P*-**2** measured by ^{1}H NMR are similar to those measured by CD for **1**, except for the guests which show a strong preference for *M*-**1** and thus have a lower affinity for the *P* helix (Table 1). On adding a guest to the empty host, a distinct set of signals emerge for the host–guest complex, 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. On cooling, the signals of the host broaden and split into two sets as tumbling of the guest becomes slow (Supplementary Fig. 42). The furanose versus pyranose forms and α versus β configurations of the guests were assessed using ^{1}H-^{13}C single quantum correlation NMR spectroscopy (^{13}C-HSQC) of ^{13}C-labelled 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. 20). On addition of **2**, some spectra remained complex. However, for **7**, **8** and **10**, the signal patterns radically simplified, showing quantitative selection of a single encapsulated species, assigned to β-D-fructopyranose (β-**7a**), α-D-mannopyranose (α-**8a**) and β-D-galactofuranose (β-**10b**), respectively (Supplementary Figs 20, 22 and 26). Unlike the furanose forms of fructose, the stabilization of β-**7a** is uncommon. Indeed, only four protein structures are reported in the Protein Data Bank that sequester this form[39].

Structures of early host–guest complexes. Structural information near the atomic scale (0.9–1.1 Å) was obtained from the structures of four **1** ⊃ guest complexes in the solid state that could be elucidated by crystallographic analysis (Fig. 4). Because of the large dimensions of the host–guest complexes, these structures contain a fraction of solvent molecules much greater than that usually found in small molecule crystals. The solvent molecules and side chains are generally disordered and their positions could not be determined accurately. However, the position of the helix backbones and the binding modes of the sugars were determined unambiguously. Details about the methods used for structure refinement and the treatment of disordered side chains and solvent molecules can be found in the Supplementary Section 1. These represent the first examples of crystal structures of unsubstituted monosaccharides bound to synthetic receptors, and provide key insights into the roles of shape and hydrogen bonds in monosaccharide selectivity by the host. (Note that a structure with octyl-glucoside has been reported in ref. 26.) To achieve these structures, co-crystallization of the host and guest was facilitated by their high affinity constants and by the use of racemic monosaccharides that, when mixed with *P*/*M*-**1**, produce centrosymmetric lattices[40,41]. In all structures, the D-carbohydrate was found in the helix with 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 $^{2}C_{5}$ β-pyranose and $^{4}C_{1}$ α-pyranose sugar conformations assigned by NMR, respectively. As listed in Supplementary Tables 13 and 14 and shown in Fig. 5, the structures had extensive arrays of up to ten hydrogen bonds between the sugar hydroxyl groups and the inner wall of the helix, a number matching that found in lectin–monosaccharide complexes[11]. Remarkably, the structures show that the intramolecular hydrogen bonds between neighbouring hydroxyl groups that are typically observed within isolated monosaccharides structures[38,42] 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** remained larger than needed to bind monosaccharides.

On average, about two-thirds of the hydrogen bonds involve a hydroxyl proton donor on the sugar, and only one-third involve an amide proton donor on the capsule. However, these proportions vary from 3:1 for fructose to 1:1 for mannose. The 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 ($d_{(O\cdots H)} \leq 2.0$ Å), except with the smaller guest xylose ($d_{(O\cdots H)} = 2.5$ Å). The H units therefore serve the purpose for which they were developed well, suggesting that other efficient hydrogen-bond acceptors such as N oxides would be valuable additions. An interesting case of induced fit is observed in the complex with fructose (Figs 4a and 5a), in which one of the two acyl hydrazide functions is found in a *gauche* conformation, with both NH groups pointing towards the sugar and both carbonyl groups pointing out. This arrangement is attributed to two hydrogen bonds formed by the two acyl hydrazide protons with the O atoms of fructose. The details revealed by these structures validate the initial capsule design and illustrate that accurate predictions of saccharide binding modes have limitations that require 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 of 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

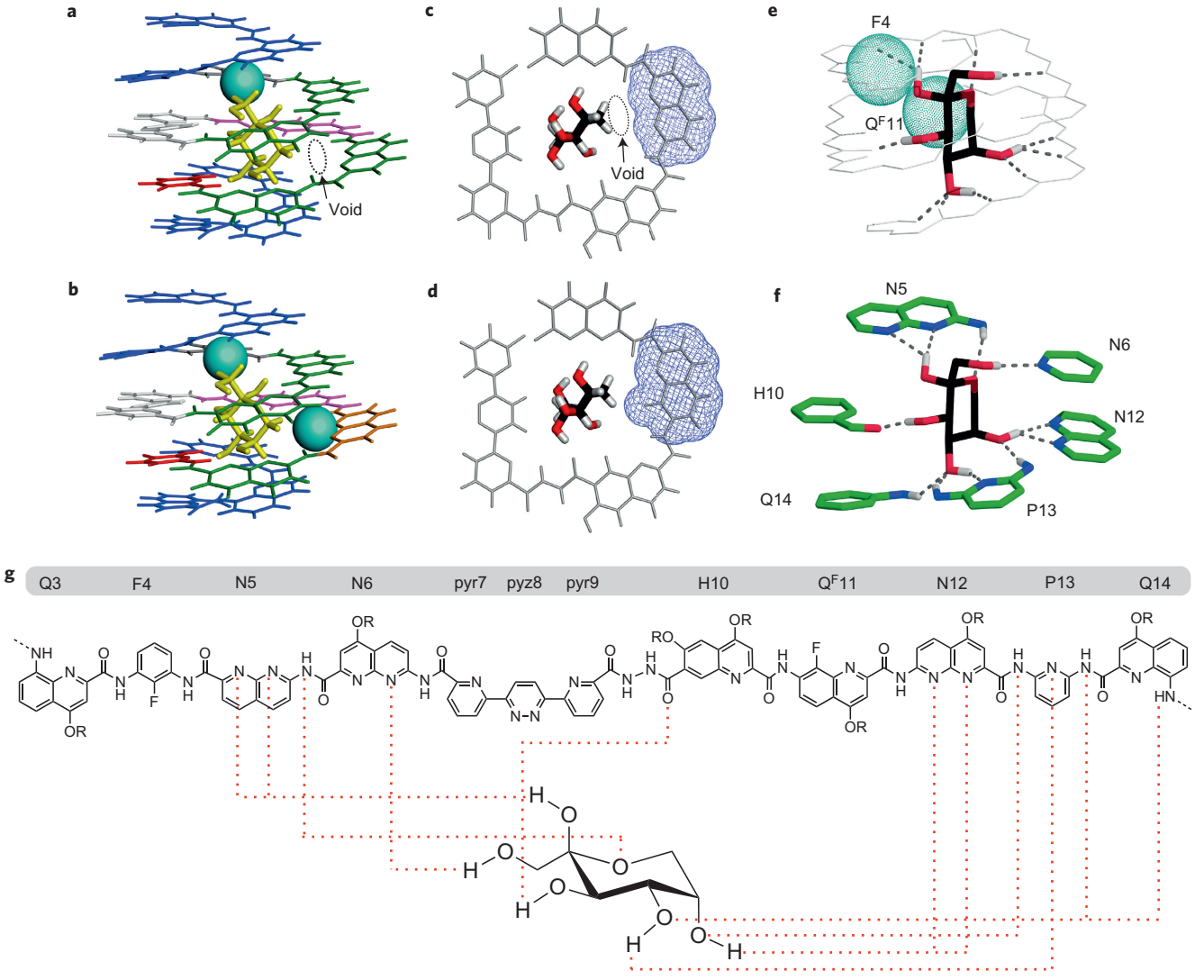

Figure 6 | Iterative receptor design. a,**b**, Crystal structures of capsule *P*-**3** with β-$^{2}C_{5}$-D-fructopyranose (**a**) and capsule *P*-**5** with β-$^{2}C_{5}$-D-fructopyranose (**b**). **c**,**d**, Top views of a slice of capsule *P*-**3** (**c**) and *P*-**5** (**d**) illustrating the volume reduction due to the N→Q^{F} mutation. Volumes of each N and Q^{F} monomer are shown as blue quad mesh. **e**,**f**, Views of the crystal structure of *P*-**5** ⊃ β-D-fructopyranose showing the array of hydrogen bonds. In **e**, only the inner rim of the capsule in light grey representation is shown. The 11 hydrogen bonds are shown as grey dashes. Fluorine atoms are represented as dotted spheres. In **f**, the sugar and those heterocycles that interact with it are shown in a tube representation. Dashed lines indicate hydrogen bonds. Details of these hydrogen bonds (distances, angles) can be found in the Supplementary Table 14. **g**, Formula and monomer numbering of sequence **5** together with the structures of β-D-fructopyranose. The two distal quinoline monomers have been removed for clarity. Note that the monomer numbering differs from that in Fig. 5, which corresponds to sequence **1**. Hydrogen bonds from the capsule wall to the sugar are shown as red dashed lines.

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. 42). 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 (Table 1). The affinity for mannose dropped by one order of magnitude, and the affinities for other guests collapsed. Interestingly, the helix handedness preferences of the **3**-mannose and **3**-ribose complexes were opposite to 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, it was hard to predict the exact consequences of deleting one H unit from the sequence of **1** to produce **3** on binding saccharides other than fructose. 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 β-fructopyranose (Supplementary Fig. 32).

Further structural insights to direct the next iteration of this optimization process were gathered in the solid state. A crystal structure of racemic **3** ⊃ β-**7a** was obtained and confirmed the preservation of most of the interactions seen within **1** ⊃ β-**7a** (Fig. 6). A remaining small void was spotted in the vicinity of the pyranose endocyclic methylene group facing one N unit (Fig. 6a,c). A mutation N→Q^{F} 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, now bound 300 times better than mannose despite the notorious difficulty to distinguish these two sugars (Supplementary Fig. 61)[43]. The selectivity of **6** for fructose is so high that, in the presence of 3 equiv. of **7**, **8**, **9**, **10**, **12** and **13**, the formation of **6** ⊃ β-**7a** remains quantitative, as judged by ^{1}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 β-**7a** was observed (Fig. 3c–e). A crystal structure of **5** ⊃ β-**7a** confirmed that the effect of the last mutation was a complete filling of the cavity space around fructose by the additional F atom, revealing atomic-scale shape complementarity between host and guest (Fig. 6). The occupancy factor of the capsule cavity by β-**7a** reaches 65%, a value much greater than the common 55% reference[44] that reflects tight attractive interactions as well as shape complementarity.

Further characterization of the final complex **6** ⊃ β-**7a** was performed by using NMR spectroscopy to observe the behaviour of the host–guest association in the active form. As a first step, extensive ^{1}H, ^{13}C and ^{15}N chemical shift assignments of **6** ⊃ β-**7a** were determined (Supplementary Tables 2–4 and Figs 47– 51) before calculation of a high-resolution NMR structure of the complex. A final ensemble of 20 structures (Supplementary Fig. 51) were calculated from distance restraints measured on a sample of uniformly ^{13}C-labelled D-fructose bound to natural abundance host molecule. The use of ^{13}C-edited and -filtered NMR spectra allowed for the collection of 311 unique distance restraints (Supplementary Table 5), 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 of the ensemble of solution structures with the solid-state crystallographic analysis (r.m.s.d. of 0.06 ± 0.04 Å).

Conclusion

The approach described above thus provides a general protocol by which to generate new highly selective hosts for given guest molecules, starting with a slightly larger than needed cavity possessing first-principles design features, then iteratively reducing the cavity size while 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 about saccharide recognition at the atomic level 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[32,45], new building blocks may be required to bring hydrophobic moieties or charges into the capsule cavity. Efforts in this direction are currently being made in our laboratory[46]. Expanding our approach to other large and complex guest molecules would enable the fabrication of novel generations of selective receptors, sensors and transporters.

Methods

X-ray crystallography. Methods for the resolution of the crystal structures of host–guest complexes **1** ⊃ β-**7a**, **1** ⊃ α-**8a**, **1** ⊃ β-**9a**, **1** ⊃ α-**13a**, **3** ⊃ β-**7a** and **5** ⊃ β-**7a** are described in detail in the Supplementary Section 1. Solving these structures requires overcoming difficulties that have more in common with macromolecular crystallography than with small-molecule crystallography. These include small crystal size and weak diffraction intensity (hence the use of bright home sources or synchrotron radiation), diffraction at the limit of atomic resolution (0.9–1.1 Å), and a large solvent and side chain content that results from the large size of the host–guest complexes and leads to large volume fractions of disordered areas. Solving and refining the structures allows models to be built with R1 values between 12 and 17%. Nevertheless, the models remain excellent for foldamer main chains, the encapsulated saccharides and the arrays of interactions between them (Supplementary Figs 58 and 59).

Multidimensional NMR spectroscopy. Data required for calculation of the ensemble of solution structures for **6** ⊃ **7a** (β-D-fructopyranose) were collected on a 2 mM sample of natural-abundance foldamer in $CDCl_3/d_6$-DMSO (95:5 vol/vol) and an equimolar amount of $^{13}C_6$-fructose previously equilibrated in d_6-DMSO for 24 h before addition to the capsule. Complete assignment of capsule ^{1}H nuclei (Supplementary Table 2), as well as ^{15}N and ^{13}C directly bound to hydrogen (Supplementary Tables 3 and 4 and Supplementary Figs 47 and 48), were achieved by using a combination of 2D ^{1}H,^{1}H-NOESY (150 and 300 ms mixing times), 2D-^{1}H,^{1}H-COSY, 2D ^{1}H,^{13}C-HSQC and 2D ^{1}H,^{15}N-HSQC. Assignments of fructose ^{1}H and ^{13}C nuclei (Supplementary Tables 2 and 4) were based on 2D ^{1}H,^{13}C-HSQC and 3D ^{1}H,^{13}C-HSQC-NOESY (150 and 300 ms mixing time). All fructose hydroxyl ^{1}H nuclei are observable and assigned based on the 2D ^{1}H,^{1}H-COSY and 3D ^{1}H,^{13}C-HSQC-NOESY (Supplementary Table 2 and Supplementary Fig. 49). Distance restraints were based on 3D ^{1}H,^{13}C-HSQC-NOESY for intramolecular fructose restraints as well as intermolecular restraints between the fructose and the capsule (Supplementary Fig. 50). Additional intermolecular restraints with the fructose hydroxyl ^{1}H atoms as well as intramolecular capsule distance restraints were based on a double-filtered 2D ^{1}H,^{1}H-NOESY (150 and 300 ms mixing times) that results in suppression of ^{1}H signals bound to ^{13}C nuclei in both dimensions of the spectra, leaving only the hydroxyl ^{1}H nuclei remaining from the fructose as well as all capsule ^{1}H nuclei (Supplementary Fig. 49). Calibration of the NOESY crosspeaks was based on the known distances of *ortho* aromatic protons (2.5 Å) for the double-filtered ^{1}H,^{1}H-NOESY and geminal protons (1.8 Å) for the 3D ^{1}H,^{13}C-HSQC-NOESY. The final list of 311 distance restraints includes 199 capsule–capsule, 35 fructose–fructose and 77 intermolecular capsule–fructose restraints (Supplementary Table 5). Calculation of 20 structures was performed with Merck Molecular Force Field (MMFFs) implemented within MacroModel (Maestro v6.5.007) and resulted in an ensemble of well-defined structures with no distance violation greater than 0.35 Å (Supplementary Table 5). The set of distance restraints are dispersed throughout the capsule–fructose **6** ⊃ **7a** complex, with the conformation of the β-D-fructopyranose precisely defined by the intramolecular fructose–fructose restraints, and the position of the fructose within the capsule defined with high resolution due to the intramolecular capsule–fructose restraints (Supplementary Fig. 51). Dihedral orientations of the hydroxyl hydrogens were confirmed by observation of a doublet or singlet in the 1D ^{1}H spectrum due to the larger value of $^{3}J_{HO-HC}$ dipolar coupling for 180° and 0°, as compared to a small value for ~90°, respectively.

Modelling of the hydrogens of carbohydrate hydroxyls. Molecular dynamic simulations were used to position the hydrogens in the hydroxyl groups of the sugars of all complexes **1** ⊃ **7a**, **1** ⊃ **8a**, **1** ⊃ **9a**, **1** ⊃ **13a**, **3** ⊃ **7a** and **5** ⊃ **7a**. Data obtained from the X-ray structures of complexes were imported within MacroModel version 8.6 via Maestro version 6.5 (Schrödinger). The X-ray structures were cleared of the isobutoxy lateral chains and hydrogens were added to the whole complex. The resulting molecules were then typified with an MMFF94 force field, which is well parametrized for the foldamer entities. The extended cutoff option was used throughout. The various complexes were submitted to 5 ns of stochastic molecular dynamics at a temperature of 500 K, with a time step of 1fs. The SHAKE option was not used and all atoms of the foldamer structures and the heavy atoms of the sugar moiety were constrained with a force constant of 100 kJ mol^{-1} Å^{-2}. A total of 200 snapshots were stored, which were subsequently minimized using the Truncated Newton Conjugate Gradient (TNCG). The resulting structures are shown in Supplementary Figs 52–57. In the case of **5** ⊃ **7a**, the position of the hydrogen of carbohydrate hydroxyls obtained by calculation was corroborated by NMR structure elucidation, thus validating this protocol.

Received 2 June 2014; accepted 2 February 2015; published online 16 March 2015

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 β-peptide oligomers. *Nature* **387,** 381–384 (1997).
3. Nelson, J. C., Saven, J. G., Moore, J. S. & Wolynes, P. G. Solvophobically 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. *et al.* A synthetic lectin for *O*-linked β-*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 Publishing, 2006).
19. Davis, A. P. & Wareham, R. S. Carbohydrate recognition through noncovalent interactions: a challenge for biomimetic and supramolecular chemistry. *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–π 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. *Org. Lett.* **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 β-mannose and β-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. Hua, Y., Liu, Y., Chen, C-H. & Flood, A. H. Hydrophobic collapse of foldamer capsules drives picomolar-level chloride binding in aqueous acetonitrile solutions. *J. Am. Chem. Soc.* **135,** 14401–14412 (2013).
33. Ajami, D. & Rebek, J. J. Compressed alkanes in reversible encapsulation complexes. *Nature Chem.* **1,** 87–90 (2009).
34. 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).
35. 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).
36. 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).
37. Qi, T., Deschrijver, T. & Huc, I. Large-scale and chromatography-free synthesis of an octameric quinoline-based aromatic amide helical foldamer. *Nature Protoc.* **8,** 693–708 (2013).
38. Lemieux, R. U. The origin of the specificity in the recognition of oligosaccharides by proteins. *Chem. Soc. Rev.* **18,** 347–374 (1989).
39. Cocinero, E. J. *et al.* Free fructose is conformationally locked. *J. Am. Chem. Soc.* **135,** 2845–2852 (2013).
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. Doores, K. J. *et al.* A nonself sugar mimic of the HIV glycan shield shows enhanced antigenicity. *Proc. Natl Acad. Sci. USA* **107,** 17107–17112 (2010).
44. Mecozzi, S. & Rebek J. Jr The 55% solution: a formula for molecular recognition in the liquid state. *Chem. Eur. J.* **4,** 1016–1022 (1998).
45. Tanatani, A., Mio, M. J. & Moore, J. S. Chain length-dependent affinity of helical foldamers for a rodlike guest. *J. Am. Chem. Soc.* **123,** 1792–1793 (2001).
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).
48. Laskowski, R. A. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. *J. Mol. Graph.* **13,** 323–330 (1995).

Acknowledgements

This work was supported by the Agence Nationale de la Recherche (project no. ANR-09-BLAN-0082-01, post-doctoral fellowship to N.C.), 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, post-doctoral fellowship to G.L.). The authors thank A. Kendhale and C. Blum for preliminary investigations on the preparation of 'H' monomers, the IR-RMN-THC Fr3050 CNRS for high-field NMR time, J-L. Ferrer for beamtime and help during data collection on FIP-BM30A at the European Synchrotron Radiation Facility and P. Nowak (University of Gröningen) for calculations using the DCLSim software.

Author contributions

N.C. and G.L. synthesized all new compounds. N.C., Y.F., G.L. and 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., Y.F. and C.D.M. co-wrote the manuscript. All authors discussed the results and commented on the manuscript. N.C. and Y.F. contributed equally to this work.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at www.nature.com/reprints. Correspondence and requests for materials should be addressed to I.H.

Competing financial interests

The authors declare no competing financial interests.