Molecular Recognition

Allosteric Recognition of Homomeric and Heteromeric Pairs of Monosaccharides by a Foldamer Capsule

Pedro Mateus, Nagula Chandramouli, Cameron D. Mackereth, Brice Kauffmann, Yann Ferrand,* and Ivan Huc*

Abstract: *The recognition of either homomeric or heteromeric pairs of pentoses in an aromatic oligoamide double helical foldamer capsule was evidenced by circular dichroism (CD), NMR spectroscopy, and X-ray crystallography. The cavity of the host was predicted to be large enough to accommodate simultaneously two xylose molecules and to form a 1:2 complex (one container, two saccharides). Solution and solid-state data revealed the selective recognition of the α-$^4C_1$-D-xylopyranose tautomer, which is bound at two identical sites in the foldamer cavity. A step further was achieved by sequestering a heteromeric pair of pentoses, that is, one molecule of α-$^4C_1$-D-xylopyranose and one molecule of β-$^1C_4$-D-arabinopyranose despite the symmetrical nature of the host and despite the similarity of the guests. Subtle induced-fit and allosteric effects are responsible for the outstanding selectivities observed.*

Introduction

The development of selective saccharide receptors is a notoriously difficult endeavor, so much so that few research groups dare challenging it.[1] Saccharides nevertheless con-

[*] Dr. P. Mateus, Dr. N. Chandramouli, Dr. Y. Ferrand, Prof. I. Huc
CBMN (UMR5248), Univ. Bordeaux—CNRS—IPB
Institut Européen de Chimie et Biologie
2 rue Robert Escarpit, 33600 Pessac (France)
E-mail: y.ferrand@iecb.u-bordeaux.fr

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

Dr. C. D. Mackereth
Université de Bordeaux, CNRS, INSERM U1212 (ARNA)
Institut Européen de Chimie et Biologie
2 Rue Robert Escarpit, 33600 Pessac (France)

Prof. I. Huc
Department Pharmazie and Center for Integrated Protein Science
Ludwig-Maximilians-Universität
Butenandtstr. 5–13, 81377 München (Germany)
E-mail: ivan.huc@cup.lmu.de

Supporting information and the ORCID identification number(s) for the author(s) of this article can be found under:
https://doi.org/10.1002/anie.201914929.

stitute a central class of biomolecules, and their chemical synthesis is an important subfield of organic chemistry. Discriminating, sensing, and selectively manipulating saccharides in water and in organic solvents thus remain subjects of broad interest, and also provide genuine opportunities to push forward the boundaries of molecular recognition. At the notable exception of receptors for all equatorial sugars,[2] the ab initio design of selective saccharide receptors has not been achieved. Screening therefore remains a common method. A typical approach has consisted of shaping binding sites from first principles and then screening which sugar binds best. Thus, various families of receptors, such as macrocycles,[2–4] tripods,[5] self-assembled metallo-organic capsules,[6] and helically folded containers,[7,8] have been produced, which often showed good affinity, and some selectivity, including for saccharides other than glucose derivatives.[3,4d,8a] Systematic variations of the receptor structure may then permit improvements of binding selectivity and affinity.

In the case of aromatic amide helical foldamers,[8] conformations are predictable through energy minimization, allowing for the design of the cavity volume and the positioning of binding features. When the helix has a reduced diameter at both extremities, it surrounds its guest and secludes it from the solvent. Guest binding and release then require a local unfolding.[9a] Because of their folding mode, such capsules are relatively rigid in all kinds of solvents and therefore operate as size and shape selectors: it was for example possible to bind selectively a dipentose at the exclusion of dihexoses, which were too large to fit into the cavity.[8b] Furthermore, their modular nature provides quick access to structural variants, using a common synthetic scheme to add, delete, or mutate monomers.

An advance brought by aromatic-foldamer-based saccharide receptors was straightforward access to detailed structural elucidation, including the very first characterization of complexes by single-crystal X-ray diffraction at atomic resolution. Crystal growth was facilitated by the rigid nature of the foldamer helices and by the use of racemic crystallography, through mixing the racemic sugar with the racemic host.[8a,10] Based on this structural information, we showed that it is possible to iteratively design a sugar receptor, that is, to introduce precise modifications so as to enhance selectivity in just a few rounds. Negative design, that is, the preservation of a binding mode to a given guest and the rational introduction of modifications to exclude all other guests was demonstrated.[8a] The rational reversal of guest selectivity was also achieved, using two guests that differ by a single hydroxy group.[9b]

Encouraged by this background, we endeavored to develop a receptor able to bind two different monosaccharides simultaneously. The study of bi- and multimolecular recognition has made it possible to explore new forms of stereoisomerism,[11] to perform chemical reactions in confined spaces,[12] and to construct supramolecular switches and logic gates.[13] Its extension to carbohydrates was initially intended as a curiosity-driven molecular recognition challenge and also as a milestone towards receptor-mediated selective reactions between unprotected saccharides. As shown in the following, our attempt was successful and also proved to be rich with several important discoveries and lessons. First, the shape and selectivity filter of aromatic amide capsules for saccharide binding is shown to be recurrently effective. Selective binding was not an objective of the current study but an essential result: the few guests that have been tested have a prevailing binding mode and are thus in principle amenable to structure-based rational iterative improvements.[8a,9b] In one case, one complex out of 42 possible host–guest combinations selectively forms. Second, heteromeric saccharide recognition was found to prevail by simultaneously binding α-$^4C_1$-D-xylopyranose and β-$^1C_4$-D-arabinopyranose despite the C_2 symmetry of the host, a rather counterintuitive process. Heteromeric guest binding in symmetrical hosts has been implemented before through space filling:[14] when one guest fills more than half the available space, a second, smaller guest is still allowed in the remaining space. However, the mechanism here is different and seems to proceed via a subtle allostery. Third, induced fit and allostery are responsible for the occupation of different binding sites by the same guest depending on whether or not another guest is present. Alternatively, they may prevent the binding of a homomeric pair of a guest but nevertheless can form a heteromeric pair. The intriguing equilibria shown in Figure 1, including the unusual substitution of one of two identical guests by a different molecule, schematize these findings.
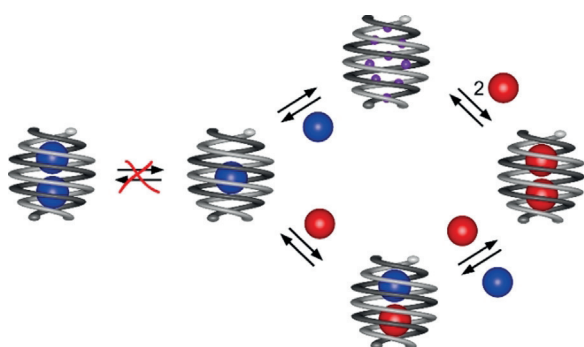
Results and Discussion

Design, Synthesis, and Characterization of a Double Helical Capsule

Taking advantage of the predictability of aromatic oligoamide foldamer structures, we previously designed unimolecular capsule **1** (Figure 2 d), which proved to be efficient at stereoselectively binding small organic acids.[9] In **1**, the quinoline trimers at each extremity close the helix cavity and also prevent its self-assembly into multiple helices (Figure 2 a). Indeed, high helix curvature, as in quinolinecarboxamide oligomers, disfavors the spring-like extension associated with double helix formation.[15] We envisioned that the removal of these trimers would allow the strands to form a stable double helical architecture endowed with a signifi-
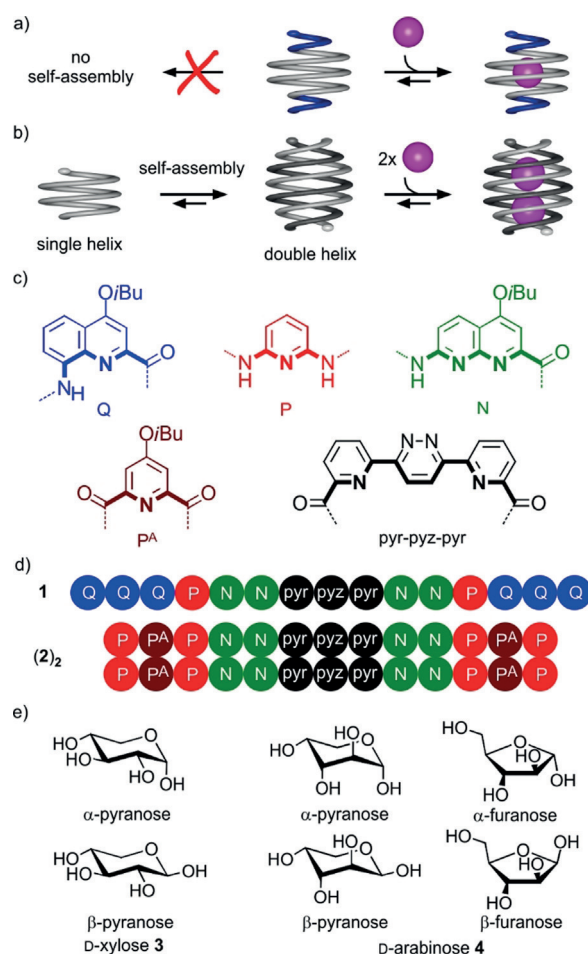

Figure 1. Schematic representation of the reported 1:1 and 1:2 host–guest complexes formed from a double helical host and two monosaccharides (red or blue spheres). Top: in the absence of guest, the host is filled by water molecules or solvent (small purple spheres). The blue guest forms a 1:1 complex and cannot form a 1:2 complex alone, but it does in the presence of the red guest. Note that the blue guest occupies different binding sites in the 1:1 and heteromeric 1:2 complexes.

Figure 2. a) Encapsulation of a guest by a single helical strand possessing a cap (blue) with reduced diameter at both ends preventing dimerization into double helices. b) Single helix–double helix equilibrium (left) and the encapsulation of two guest molecules within the cavity of a duplex (right). c) Letter and color codes of the diamine, diacid, and amino acid monomers, the bonds shown in bold delineate the inner rim of the helix. d) Oligoamide sequences used in this work. In sequence **1**, the two terminal Q units have a terminal 8-nitro group instead of an amino function while in sequence **2** the two terminal P units have a pivaloyl group. e) The guest molecules: most abundant tautomeric forms of D-xylose **3** and D-arabinose **4** in solution.

cantly larger cavity than a single helical analogue (Figure 2b). Oligomer **2** was designed based on this assumption. Its sequence is a modified version of **1** in which the quinoline segments have been replaced by pyridinecarboxamide dimers. The convergent synthesis of **2** involves the coupling of a pivaloyl-PP^A mono-acid with the amine of H_2N-PN$_2$-Boc using PyBOP as the coupling reagent. After Boc cleavage, pivaloyl-P$_3$N$_2$-NH$_2$ was coupled twice to the diacid of pyr-pyz-pyr to provide **2** (see the Supporting Information). The choice of a self-assembled receptor aimed to simplify the synthesis. The different blocks mentioned above can be prepared on multigram scales, allowing us to readily obtain large amounts (> 1 g) of **2**.

The solid-state crystallographic structure of **2** was solved and revealed a 2 nm long duplex in which each strand spans three helical turns (Figure 3). The two strands are helically offset with respect to one another by half a turn and extensively stack on top of each other. No other obvious interstrand interactions were noted. The duplex has three pseudo-C_2 symmetry axes, one along the helix axis and two in orthogonal directions.
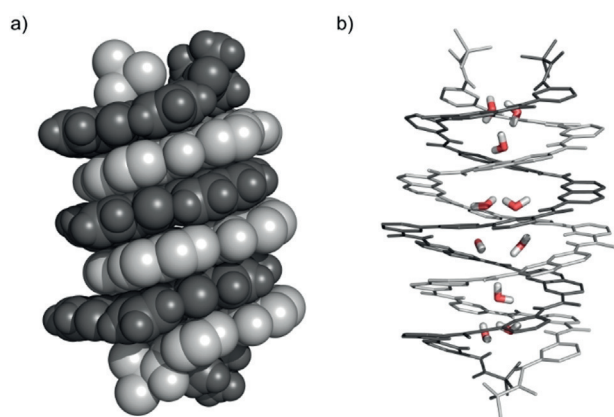

Figure 3. Solid-state structure of (**2**)$_2$: a) Side view shown in CPK representation. b) Side view with 10 encapsulated water molecules. In both representations each strand is colored in a different tone of gray. Side chains have been omitted for clarity.

Evidence of double helix formation was also found in solution. The ^1H NMR spectrum of **2** recorded at 2 mM in CDCl$_3$ (Figure S4) shows slightly broadened peaks. The aromatic amide resonances appear in the 10–8.5 ppm region, at significantly higher field than what is usually found in single helical aryl amide capsules, a hallmark of double helix formation in these systems.[15] Diluting down to 0.05 mM did not allow for the detection of single helix resonances in this solvent. However, the addition of [D$_6$]DMSO (Figure S5), a competitive solvent that disfavors double helix formation, led to the emergence of a second set of sharp signals at lower fields, which were assigned to the single helix. In pure [D$_6$]DMSO, only the single helix is observed. Indeed, the structure of a crystal grown from DMSO was solved and shown to be the single helix (Figure S22). In CDCl$_3$/[D$_6$]DMSO (9:1 v/v), the single helix can be detected as a minor species at 180 µM. Integration of the signals

provided a minimal estimate of the dimerization constant as $K_{dim} = 4 \times 10^5 \, M^{-1}$, a value large enough to consider (**2**)$_2$ to be a single entity at the concentrations used in this work.

Prediction of Polar Guest Binding

A trend has emerged from the host–guest properties of various aromatic amide foldamer capsules studied in independent contexts: tight and selective binding goes along with an occupancy of the host cavity volume by the guest of at least 70%.[8,9] Guests that are smaller than optimal also bind but with a lower affinity. Guests that are too large for the available space do not bind at all. Another aspect to consider is a weak but non negligible ability of the host to adjust its conformation to the volume of the guest through slight changes in helix curvature, the pitch remaining constant and equal to the thickness of one aromatic ring. In the few cases where the structure of the host has been elucidated in the absence of guest, that is, when the host is filled with solvent only, the cavity was found to be slightly smaller than in the presence of a guest. The folded duplex structure (**2**)$_2$ has a polar cavity, as evidenced by the presence of ten crystallographically defined water molecules (Figure 3b), and an inner volume of 280 Å^3 (Figure S23a,e,f). We surmised that this cavity may be large enough to harbor two aldopentose guests. Xylose has a volume of 107 Å^3, and two molecules of xylose would occupy 77% of the cavity volume measured in the absence of guest, and presumably a smaller fraction of the space available in an actual host–guest complex. In contrast, the volume of a hexose (ca. 130 Å^3) is clearly too large to fit twice in the cavity of (**2**)$_2$. We therefore concentrated our efforts on the recognition of pentoses.

Solution and Solid-State Study of D-Xylose Encapsulation

The ability of (**2**)$_2$ to bind pentoses was assessed by titrations in CHCl$_3$/DMSO (9:1 v/v) at 298 K. A circular dichroism (CD) titration of achiral (**2**)$_2$ with D-**3** showed the appearance of a negative induced CD signal (Figure 4b) resulting from helix handedness bias. The changes in ellipticity could be fitted to a 1:2 binding model (Figure 4c), which afforded an overall binding constant $K = 2.19 \times 10^9 \, M^{-2}$ corresponding to K_a values of 69 200 M^{-1} and 31 600 M^{-1} for the binding of a first and second D-xylose guest, respectively. These values reflect a slightly positive cooperativity ($\alpha = 4K_{a2}/K_{a1} = 1.8$), illustrated by the sigmoidal binding isotherm (Figure 4c).[16] A ^1H NMR titration under the same conditions revealed the appearance of a single new set of sharp helix signals upon binding of D-**3** (Figure 4d–g). The number of amide resonances (12) is indicative of a pseudo-twofold symmetry, that is, a lower symmetry than for the double helix in the absence of xylose. In agreement with this result was the fact that the ^1H-^{13}C HSQC spectrum of encapsulated, uniformly ^{13}C-labeled D-**3** (Figure 4h) recorded under the same conditions showed only one set of correlations, suggesting that both bound sugars have the same chemical environment on average. The number of amide signals and the
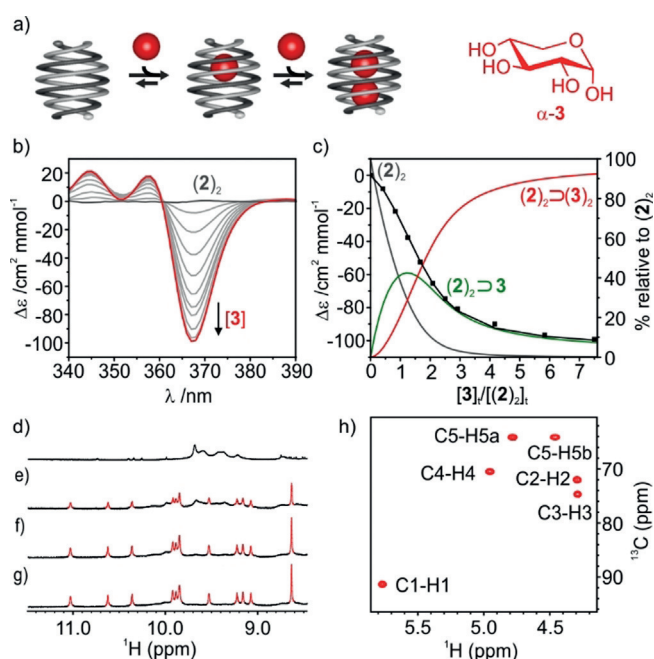
Figure 4. a) Encapsulation of two guests by a double helical capsule. Red balls represent D-xylose **3**. b) Induced CD spectra upon binding of D-**3** by (**2**)₂ in CHCl₃/DMSO (9:1 v/v) at 298 K, [(**2**)₂]ₜₒₜ = 96 μM. The red colored line corresponds to 7.5 equiv of D-**3** added. c) Experimental (■) and calculated values (——) for the ICD binding study of receptor (**2**)₂ vs. D-**3** with the corresponding species distribution diagram. λ = 367 nm. d–g) Excerpts from the 400 MHz ¹H NMR spectra showing the amide resonances of (**2**)₂ at 1 mM (298 K) in CDCl₃/[D₆]DMSO (9:1 v/v) in the presence of d) 0 equiv, e) 1 equiv, f) 2 equiv, and g) 3 equiv of D-**3**. h) Excerpt of the HSQC spectrum showing the ¹H-¹³C correlation signature of two encapsulated, uniformly ¹³C-labeled α-⁴C₁-D-xylopyranose molecules **3**.

simplicity of the HSQC spectrum also reflect a complete selectivity for a single tautomer of D-**3**, associated with the full diastereoselectivity for a given handedness of the double helix. By considering the ¹³C chemical shifts, the dihedral angles between CH and OH groups derived from ³J coupling constants, and COSY and TOCSY two-dimensional NMR experiments, it was possible to determine that the two guests adopt an α-⁴C₁-pyranose puckered conformation.

The spontaneous occurrence of such a level of selectivity is significant. Considering the α/β-anomers of the guest and the *P/M*-helicity of the host, six different 1:2 complexes may form. Solution data not only show that one of them prevails, but also that one binding mode must prevail as well. Despite the many hydrogen bond donors and acceptors on both the guest and host, a particular orientation is preferred.[8]

Further characterization of the complex was performed by 2D and 3D NMR spectroscopy. ¹H-¹H ROESY spectra recorded at 298 K revealed the existence of exchange cross-peaks between protons of the capsule strands, reflecting the dynamic nature of the pseudo-*C₂*-symmetrical complex in solution. The exchange rate between the two populations was measured to be 4.9 ± 0.2 s⁻¹, and is thus on the same timescale as that required by the multidimensional NMR spectra. Decreasing the temperature to 278 K suppressed the exchange (Figure S13). This allowed for extensive ¹H, ¹³C, and

¹⁵N chemical shift assignments of the spectra of (**2**)₂⊃(D-**3**)₂ (Tables S1–S3 and Figures S12 and S13) and the determination of a high-resolution NMR structure of the complex. A final ensemble of 20 structures (Figure 5a) was calculated from distance restraints measured on a sample of ¹³C-labeled D-**3** bound to ¹³C-natural abundance (**2**)₂. The use of ¹³C-edited and -filtered NMR spectra allowed for the collection of 334 unique distance restraints (Figure S12 and Table S4), including 82 intermolecular restraints to accurately position the monosaccharides within the capsule cavity and 190 inter- and intrastrand restraints to position the strands relative to one another, as well as the side chains. The structure also revealed the presence of two bound water molecules located at the extremities of the double helix and the assignment of *P*-helix handedness. In the complex, the two sugar-binding sites are identical and consist of the two inequivalent extremities of the two strands. The dynamic exchange mentioned above can thus be assigned to a sliding motion of the two strands[17] with respect to one another, along with a concomitant repositioning of each sugar within its cavity.

In parallel, the crystal structure of racemic (**2**)₂⊃D/L-**3** was elucidated (Figure 5c). The use of racemic crystallography recently helped delivering the very first structures of receptor–sugar complexes,[8a, 10] but crystal growth is also a consequence of the complex structure being well defined in solution. There is a near-perfect superposition of the ensemble of solution structures with the solid-state crystallographic analysis with a root mean square deviation (rmsd) of 0.054 ± 0.03 Å (Figure S17). The match includes the pseudo-*C₂* symmetry, the presence and position of the two sugars, and the presence of the two water molecules, as well as the *P*-handedness of the double helix containing D-**3**. The solid-state structure also confirmed the α-⁴C₁-pyranose tautomeric form of the sugars. The centrosymmetric *P*-1 space group implies that the crystal lattice also contains two α-¹C₄-L-xylopyranose molecules encapsulated by *M*-(**2**)₂. The segregation of D- and L-**3** in *P*- and *M*-helices, respectively, is in line with the full diastereoselectivity observed in solution (Figure 4d–h). The sugar racemate could also yield a diastereomeric heterocomplex including both D-**3** and L-**3**. However, the latter did not crystallize and was also not observed in solution. For example, when a 1:1 (pseudo-racemic) mixture of ¹³C-labeled D-**3** and ¹³C-natural abundance L-**3** was added to (**2**)₂, HSQC spectra confirmed that only one type of complex formed.

In both the solution and solid-state structures, two water molecules were found to be encapsulated with the carbohydrates, each occupying an extremity of the cavity. These water molecules are held in position through multiple hydrogen bonds with the amide protons of the terminal pyridine monomer of each strand. Additionally, direct water-to-saccharide hydrogen bonding is observed. The structures show an extensive array of eight hydrogen bonds between each sugar hydroxy group and the inner wall of the helix (Table S9 and Figure 5b,d). No intramolecular hydrogen bonds were found between neighboring hydroxy groups of the monosaccharides. Only host–guest intermolecular hydrogen bonds exist. Four of the eight hydrogen bonds involve hydroxy proton donors and the other four involve amide
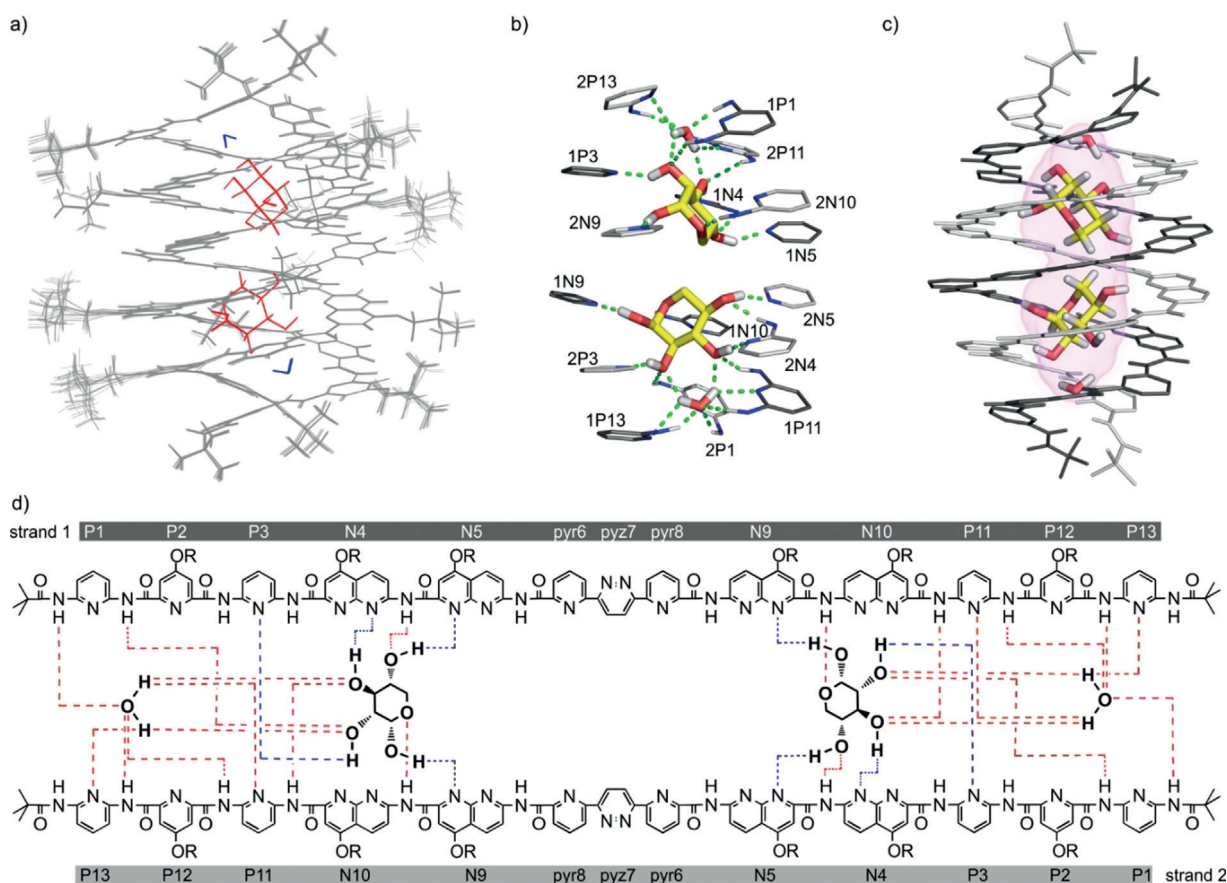
Figure 5. a) Ensemble of 20 overlaid high-resolution NMR-based models of *P*-(**2**)$_2$⊃(α-^{4}C$_1$-D-xylopyranose)$_2$. The sugars are highlighted in red and the two water molecules in blue. b) Enlarged side view of the cavity showing the heterocycles that interact with the guests and the water molecules. The heterocycles are color-coded in light gray or dark gray depending on the strand that they belong to. The 30 hydrogen bonds found in the complex are shown as green dashed lines. Details of these hydrogen bonds can be found in the Supporting Information. c) Solid-state structure of *P*-(**2**)$_2$⊃(α-^{4}C$_1$-D-xylopyranose)$_2$, shown in thin tube representation for the host and thick tube representation for the guests. Each strand is colored in a different shade of gray; the two D-**3** guests are shown in yellow. Non-polar hydrogen atoms, isobutoxy side chains, and cavity-excluded solvent molecules have been removed for clarity. The volume of the cavity (306 Å^3) is shown as a transparent pink isosurface. d) Formula and monomer numbering of each strand of the double helical capsule together with the structures of D-**3** represented as Mills projections. Hydrogen bonds where the sugars act as acceptors or donors are shown as red and blue dashed lines, respectively. R = isobutyl.

proton donors. Each sugar exposes its endocyclic oxygen atom and methylene group to the center of the cavity, forbidding inter-sugar hydrogen bonds. The weak positive binding cooperativity observed is thus mediated by the helix backbone, not by guest–guest interactions, despite the fact that the guests do not break the symmetry of the host. Comparison of the solid-state structures of the empty host and of the (**2**)$_2$⊃(D-**3**)$_2$ complex confirms that a conformational change takes place upon binding: the two strands of (**2**)$_2$ are helically offset with respect to one another by a quarter of a turn in the complex as opposed to half a turn in the empty capsule, revealing some kind of induced fit at one of the rare degrees of structural freedom of the duplex. As a result, the central monomers of each strand are at an angle of about 60° in the complex instead of being in front of each other in the empty capsule. Concomitantly, the inner volume increases from 280 to 306 Å^3 upon guest binding. This allowed us to calculate that 70% of the volume is occupied by the two xylose molecules, as predicted initially.

Solution Studies on D-Arabinose Binding

The binding of D-arabinose **4**, a pentose that differs from D-**3** by only two stereogenic centers, was then evaluated. A titration of (**2**)$_2$ with D-**4** in CHCl$_3$/DMSO (9:1 v/v) at 298 K was monitored by CD spectroscopy. A positive signal centered at 370 nm appeared upon increasing the concentration of D-**4** (Figure 6b). The CD intensity remained significantly weaker than with D-**3**. In addition, the band at 370 nm had an opposite sign but this did not apply to other bands, hinting at variations of the CD spectra not only through the handedness of the helix but possibly also through the relative positioning of the two strands. In the case of D-**4**, the changes in ellipticity were inconsistent with a 1:2 stoichiometry but instead fitted to a 1:1 binding model (Figure 6c), which afforded a K_a value of 21 900 M^{-1}, a lower affinity than for D-**3**. A ^{1}H NMR titration in CDCl$_3$/[D$_6$]DMSO (9:1 v/v) at 298 K was attempted. However, although spectral changes were clearly visible, the
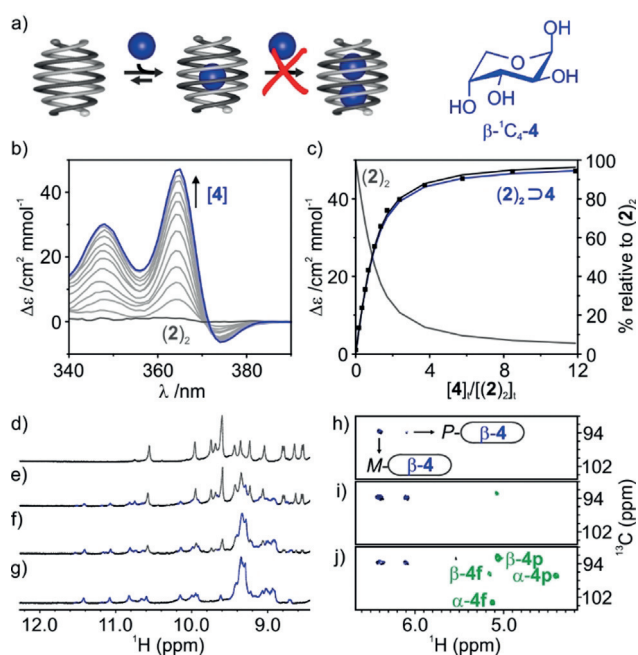
Figure 6. a) Encapsulation of a single guest by a double helical capsule. Blue balls represent D-arabinose **4**. b) Induced CD spectra upon binding of D-**4** by (**2**)$_2$ in CHCl$_3$/DMSO (9:1 v/v) at 298 K, [(**2**)$_2$] = 120 μM. The blue line corresponds to [D-**4**] = 823 μM. c) Experimental (■) and calculated values (——) for the ICD binding study of receptor (**2**)$_2$ vs. D-**4** with the corresponding species distribution diagram. λ = 365 nm. d–g) Excerpts from the 400 MHz ^{1}H NMR spectra showing the amide resonances of capsule (**2**)$_2$ at 1 mM in CDCl$_3$/[D$_6$]DMSO (9:1 v/v) and at 243 K in the presence of d) 0 equiv, e) 0.5 equiv, f) 1.0 equiv, and g) 2.0 equiv of D-**4**. h–j) Excerpts from ^{1}H-^{13}C HSQC spectra showing the ^{1}H correlations of the C1 (anomeric carbon) atom of uniformly ^{13}C-labeled β-^{1}C$_4$-D-arabinopyranose recorded under the following conditions: h) [(**2**)$_2$] = 1.0 mM, [D-**4**] = 0.5 mM; i) [(**2**)$_2$] = 1.0 mM, [D-**4**] = 1.0 mM; j) [(**2**)$_2$] = 1.0 mM, [D-**4**] = 2.0 mM. Encapsulated and free forms of the sugar are represented in blue and green, respectively.

signals were too broad to be interpreted, probably because of guest tumbling and/or the presence of disordered solvent molecules inside the cavity. The same experiment repeated at 243 K (Figure 6d–g) revealed the emergence of reasonably sharp amide peaks. As for the CD titration, the spectrum does not change appreciably after addition of 1 equiv of D-**4**. The same titration was carried out with ^{13}C-labeled D-**4** and monitored by ^{1}H-^{13}C HSQC spectroscopy. Again, addition of more than 1 equiv of guest did not lead to any variation of the complex resonance pattern (highlighted in blue in Figure 6h–j). The HSQC data revealed two distinct resonances for the anomeric C1 carbon atom of the ^{13}C-labeled D-**4** encapsulated in (**2**)$_2$. These cross-peaks have a similar ^{13}C chemical shift, and are thus unlikely to correspond to different anomers of the sugar. Instead, we assigned them to diastereoisomeric complexes of a unique anomer of D-**4** being encapsulated either in *P*- or *M*-(**2**)$_2$. Integration of the two cross-peaks allowed the diastereomeric excess to be calculated to be 30%, which is consistent with the lower CD intensity.[18] Based on the ^{13}C chemical shifts, the dihedral angles between CH and OH groups derived from ^3J coupling constants, and two-

dimensional COSY and TOCSY NMR experiments, it was possible to determine that the guest is in a β-^{1}C$_4$-pyranose puckered conformation. In this conformation, only one guest molecule is allowed in the double helix cavity.

The binding of arabinose thus appears to be less selective than that of xylose. However, it can be inferred from the NMR data that the two complexes observed are well-defined, including the conformation of the sugar. The absence of a 1:2 complex even though arabinose is not larger than xylose is also indicative of tight and selective interactions. Arabinose will not migrate to a location of the host cavity that would allow for a second guest to bind. The size and shape selectivity filters of aromatic amide foldamer cavities evoked above are again at play.

Structural and Thermodynamic Studies of the Formation of a Heterodimeric D-Xylose–D-Arabinose Complex

Next, we sought to evaluate whether a heteromeric pair of pentoses could be encapsulated by (**2**)$_2$. CD monitoring of the addition of D-**4** to (**2**)$_2$ previously equilibrated with excess D-**3** revealed that the initial spectrum, typical of (**2**)$_2$⊃(D-**3**)$_2$, changed to eventually reach a final state suggesting some sort of saturation (Figure 7b,c). The final spectrum was very different from that of (**2**)$_2$⊃(D-**4**), the expected final product if arabinose had simply replaced the two xylose molecules. This result hinted at the possible formation of heteromeric (**2**)$_2$⊃(D-**3**;D-**4**). The changes in ellipticity could indeed be fitted to such a process (Figure 7c) to afford a K_a value of 46 800 M^{-1} for the equilibrium (**2**)$_2$⊃D-**3** + D-**4** ⇌ (**2**)$_2$⊃(D-**3**,D-**4**). In other words, D-**4** has an affinity for (**2**)$_2$D-**3** that is more than twice as large as that for (**2**)$_2$. Conversely, (**2**)$_2$⊃D-**3** has a slightly larger affinity for D-**4** than for D-**3**.

Monitoring the same titration by ^{1}H NMR spectroscopy revealed the emergence of a new set of peaks as D-**4** probably replaces one of the D-**3** guests (Figure 7d–g) to form a heteromeric complex. Consistent with this interpretation, the number of the capsule amide peaks was doubled relative to what was found for the symmetrical (**2**)$_2$⊃(D-**3**)$_2$, implying that the final structure had no symmetry at all. Titrations with ^{13}C-labeled D-**3** and D-**4** monitored by ^{1}H-^{13}C HSQC spectroscopy eventually provided unequivocal evidence for heterocomplex formation (Figure 7h–j). Upon adding D-**4**, two new cross-peaks appeared that correspond to the encapsulated C1 anomeric carbon atoms of both D-**3** and D-**4**. The unambiguous assignment of the sugar resonances was achieved by using ^{13}C-labeled D-**3** in the presence of unlabeled D-**4** and vice versa (Figure S9). We also found that increasing the temperature to 318 K increased the proportion of the heterocomplex to more than 90% (Figure 7k). This led us to study the effect of temperature on the replacement of D-**3** by D-**4**, which can be described by the equilibrium: (**2**)$_2$⊃(D-**3**)$_2$ + D-**4** ⇌ (**2**)$_2$⊃(D-**3**;D-**4**) + D-**3**. The linear van't Hoff plots (Figure S11) showed that the process is enthalpically disfavored and entropy-driven (ΔH = 32 kJ mol^{-1}; ΔS = 0.11 kJ mol^{-1} K^{-1}, i.e., $-T\Delta S$ = −32.8 kJ mol^{-1} at 298 K): below room temperature homocomplex formation is favored while above it heterocomplex formation prevails. The origin
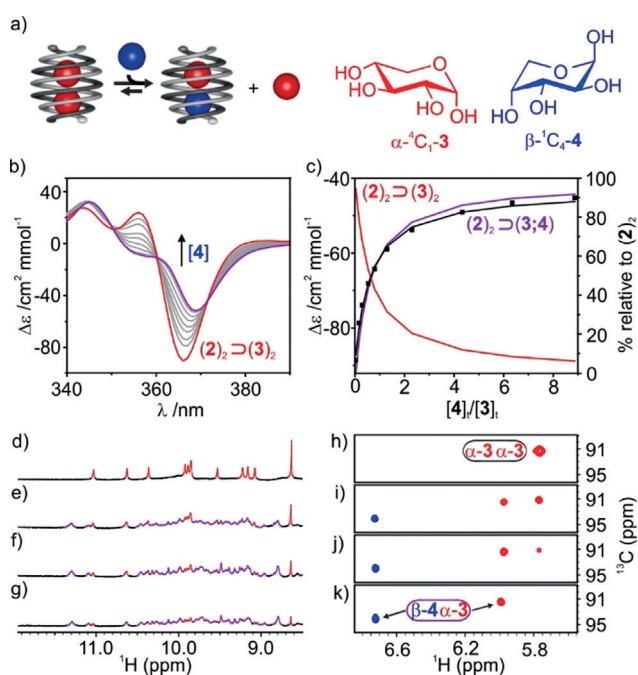
Figure 7. a) Schematic representation of the replacement of a guest by a different one within a double helical capsule to afford a heterocomplex. D-Xylose **3** and D-arabinose **4** are shown in red and blue, respectively. b) Changes in the CD spectra of $(2)_2 \supset (D-3)_2$ upon binding of D-**4** in CHCl$_3$/DMSO (9:1 v/v) at 298 K. $[(2)_2] = 94$ μM; [D-**3**] $= 755$ μM. The purple line corresponds to [D-**4**] $= 4.1$ mM. c) Experimental (■) and calculated values (——) for the ICD binding study of $(2)_2 \supset (D-3)_2$ vs. D-**4** with the corresponding species distribution diagram. $\lambda = 367$ nm. d–g) Excerpts from the 400 MHz ^{1}H NMR spectra showing the amide resonances of $(2)_2 \supset (D-3)_2$ at 1 mM in CDCl$_3$/[D$_6$]DMSO (9:1 v/v) and 298 K in the presence of d) 0 equiv, e) 1 equiv, f) 2 equiv, and g) 3 equiv of D-**4** relative to D-**3**. $(2)_2 \supset (D-3)_2$ and $(2)_2 \supset (D-3,D-4)$ amide resonances are shown in red and purple, respectively. h–k) Excerpts from ^{1}H-^{13}C HSQC spectra recorded in CDCl$_3$/[D$_6$]DMSO (9:1 v/v) at 298 K showing the ^{1}H correlations of the C1 atom of encapsulated, uniformly ^{13}C-labeled α-$^4C_1$-D-xylopyranose and β-$^1C_4$-D-arabinopyranose recorded under the following conditions: h) $[(2)_2] = 1.0$ mM, [**3**] $= 2.0$ mM; i) $[(2)_2] = 1.0$ mM, [**3**] $= 2.0$ mM, [**4**] $= 2.0$ mM; j) $[(2)_2] = 1.0$ mM, [**3**] $= 2.0$ mM, [**4**] $= 6.0$ mM; k) $[(2)_2] = 1.0$ mM, [**3**] $= 2.0$ mM, [**4**] $= 6.0$ mM at 318 K. H1/C1 correlations of D-**3** and D-**4** are shown in red and blue, respectively.

of such a large entropic component for a substitution process is unclear and may be related to the release of encapsulated water molecules (Figure 3). Yet, no such effects have been observed in the other equilibria investigated here or in our earlier studies on foldamer–saccharide recognition.[8]

Detailed structural information could not be gathered using crystallography. However, advanced NMR spectroscopic techniques allowed for the structural determination of the complex. As in the case of $(2)_2 \supset (D-3)_2$, exchange between the two strands of the capsule hampered the resolution of the structure at 298 K. Upon cooling down to 278 K, the exchange signals disappeared, but the amount of heterocomplex present in solution dropped because of the large entropic term mentioned above. Nevertheless, the concentration of $(2)_2 \supset (D-3;D-4)$ remained sufficient for a partial assignment of the resonances (Table S6 and Figures S19 and S20). By using

combinations of natural-abundance and ^{13}C-labeled sugars it was eventually possible to determine an NMR structure of the complex. A final ensemble of 15 structures (Figure 8a) was calculated from a total of 54 distance restraints, including 30 sugar–capsule intermolecular restraints that allowed us to position accurately both pentoses within the capsule cavity, and 24 inter- and intra-sugar restraints to determine the configurations and relative orientations of the guests (Figure S18). The structure revealed that the arabinose adopts a β-$^1C_4$-D-pyranose conformation whereas the xylose remains as an α-$^4C_1$-D-pyranose.

An array of eight hydrogen bonds between the hydroxy groups of each sugar and the inner wall of the helix hold the guests in place (Table S7 and Figure 8b). The xylose position is identical to that of the homo-complex. Both pentoses are oriented with their endocyclic oxygen atom pointing to the same side of the cavity, leaving OH4 of D-**3** relatively close to OH1 and OH2 of D-**4** (3.4 Å). Although it was not possible to confidently obtain distance restraints to accurately position the two water molecules located at the extremities of the complex, these were kept in the structure. The inner volume of the cavity was found to be 325 Å^3, and is slightly larger than that of the homo-complex. It thus appears that the replacement of one D-**3** guest by D-**4** can be mediated by subtle allosteric variations, but that the changes are important enough to prevent a second substitution ($(2)_2 \supset (D-4)_2$ was not observed). The host-mediated allosteric communication between two guests of identical size that differ by two stereogenic centers, xylose and arabinose, without contact between them appears to be unique. Usually, a host meant to bind two different guests would be designed with two different binding sites, for example, to bind an ion pair.[19] Alternatively, symmetrical hosts have been shown to bind to heteromeric pairs of guests when the first guest occupies more than half of the space available, leaving room only for a smaller guest.[14]

At last, we challenged the stereoselectivity of the complexation of both **3** and **4** by $(2)_2$ to draw a parallel with the mutual exclusion of D-**3** and L-**3** mentioned above. The addition of D-**3** to $(2)_2 \supset (D-4)$ first produced $(2)_2 \supset (D-3;D-4)$ and then, with a large excess of D-**3**, led to $(2)_2 \supset (D-3)_2$. In contrast, the addition of L-**3** to $(2)_2 \supset (D-4)$ did not produce any heteromeric complex. Instead, D-**4** is replaced by L-**3** to first produce $(2)_2 \supset (L-3)$ and then $(2)_2 \supset (L-3)_2$ (Figure S10). Note that these competition experiments also involve some changes in helix handedness as $(2)_2 \supset (D-3)_2$ is *P*- whereas $(2)_2 \supset (L-3)_2$ is *M*-helical.

Assessing the selectivity of sugar binding through the screening of a large number of different pentoses or hexoses was not the purpose of the present investigation. Yet, selectivity eventually turned out to be our main finding. The theoretical outcome of mixing D-xylose and D-arabinose with a racemic *P/M*-capsule is that no less than 42 different host–guest complexes may be produced (Figure S21). The observation of a single heteromeric pair of sugars composed of α-$^4C_1$-D-xylopyranose and β-$^1C_4$-D-arabinopyranose in the *P*-helical foldamer cavity is thus outstanding.
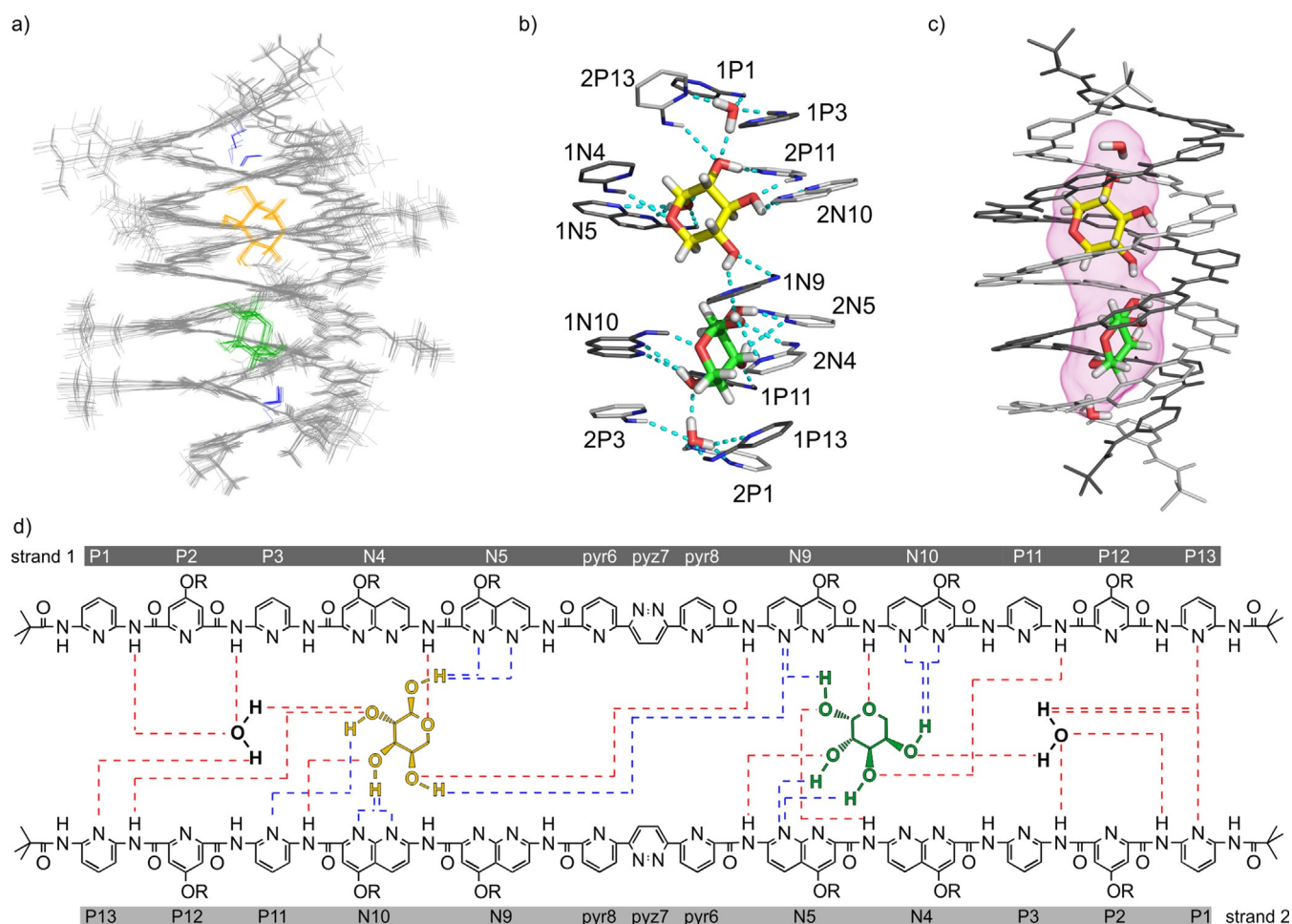
Figure 8. a) Ensemble of 15 overlaid high-resolution NMR-based models of *P*-(2)$_2$⊃(α-^{4}C$_1$-D-xylopyranose;β-^{1}C$_4$-D-arabinopyranose). D-**3** and D-**4** are highlighted in yellow and green, respectively. b) Enlarged side view of the cavity showing the heterocycles that interact with the guests and the water molecules. The heterocycles are color-coded in light gray or dark gray depending on the strand that they belong to. The hydrogen bonds found in the complex are shown as cyan dashed lines. Details of these hydrogen bonds can be found in the Supporting Information. c) Optimized structure of the complex, shown in thin tube representation for the host and in thick tube representation for the guests. Each strand is colored in a different shade of gray. Non-polar hydrogen atoms, isobutoxy side chains, and cavity-excluded solvent molecules have been removed for clarity. The volume of the cavity (325 Å^3) is shown as a transparent pink isosurface. d) Formula and monomer numbering of each strand of the double helical capsule together with the structures of D-**3** and D-**4** shown as Mills projections. Hydrogen bonds where the sugars act as acceptors or donors are shown as red and blue dashed lines, respectively. R = isobutyl.

Conclusion

In conclusion, we have prepared a double helical foldamer container with a large internal cavity by using a strategy combining the folding and the self-assembly of a readily accessible aromatic oligoamide strand. The container stereoselectively encapsulates a single homochiral pair of one xylose tautomer, α-^{4}C$_1$-D-xylopyranose. We then demonstrated the fully selective complexation of a heteromeric pair of pentoses. Together with earlier studies,[15] these results concur to show that aromatic amide helices act as stringent shape and selectivity filters for carbohydrate binding. In reference to the first sentence of the introduction about the challenge of carbohydrate recognition, it appears that general solutions are emerging. An outcome of the formation of well-defined complexes is the possibility to accurately elucidate their structures and unravel the recognition, induced fit, and allosteric mechanism at play. In turn, the obtained host–guest

complex structures may constitute new starting points for structure-based iterative design and, eventually, the further improvement (i.e., exclusion of all sugars but one) or even the reversal of guest selectivity.[8a,9b] For this purpose, advanced predictive computational tools would bring a major advantage, and their development is highly needed. Our results also open up the possibility to precisely design confined spaces that could alter and control the reactivity of native carbohydrates and behave as molecular flasks.

Acknowledgements

Conflict of interest

The authors declare no conflict of interest.

[1] a) O. Francesconi, S. Roelens, *ChemBioChem* **2019**, *20*, 1329; b) S. Tommasone, F. Allabush, Y. K. Tagger, J. Norman, M. Kopf, J. H. R. Tucker, P. M. Mendes, *Chem. Soc. Rev.* **2019**, *48*, 5488.

[2] a) R. A. Tromans, T. S. Carter, L. Chabanne, M. P. Crump, H. Li, J. V. Matlock, M. G. Orchard, A. P. Davis, *Nat. Chem.* **2019**, *11*, 52; b) T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump, A. P. Davis, *Nat. Chem.* **2016**, *8*, 69; c) C. Ke, H. Destecroix, M. P. Crump, A. P. Davis, *Nat. Chem.* **2012**, *4*, 718; d) Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619; e) E. Klein, M. P. Crump, A. P. Davis, *Angew. Chem. Int. Ed.* **2005**, *44*, 298; *Angew. Chem.* **2005**, *117*, 302.

[3] O. Francesconi, M. Martinucci, L. Badii, C. Nativi, S. Roelens, *Chem. Eur. J.* **2018**, *24*, 6828.

[4] a) Y. Jang, R. Natarajan, Y. H. Ko, K. Kim, *Angew. Chem. Int. Ed.* **2014**, *53*, 1003; *Angew. Chem.* **2014**, *126*, 1021; b) A. P. Davis, R. S. Wareham, *Angew. Chem. Int. Ed.* **1998**, *37*, 2270; *Angew. Chem.* **1998**, *110*, 2397; c) S. Anderson, U. Neidlein, V. Gramlich, F. Diederich, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1596; *Angew. Chem.* **1995**, *107*, 1722; d) K. Kobayashi, Y. Asakawa, Y. Kato, Y. Aoyama, *J. Am. Chem. Soc.* **1992**, *114*, 10307; e) O. Perraud, A. Martinez, J.-P. Dutasta, *Chem. Commun.* **2011**, *47*, 5861.

[5] a) C. Geffert, M. Kuschel, M. Mazik, *J. Org. Chem.* **2013**, *78*, 292; b) M. Mazik, H. Cavga, P. G. Jones, *J. Am. Chem. Soc.* **2005**, *127*, 9045; c) H. Abe, Y. Aoyagi, M. Inouye, *Org. Lett.* **2005**, *7*, 59; d) A. Ardá, C. Venturi, C. Nativi, O. Francesconi, G. Gabrielli, F. J. Canada, J. Jimenez-Barbero, S. Roelens, *Chem. Eur. J.* **2010**, *16*, 414; e) C. Schmuck, M. Schwegman, *Org. Lett.* **2005**, *7*, 3517.

[6] M. Yamashina, M. Akita, T. Hasegawa, S. Hayashi, M. Yoshizawa, *Sci. Adv.* **2017**, *3*, e1701126.

[7] a) J.-L. Hou, X.-B. Shao, G.-J. Chen, Y.-X. Zhou, X.-K. Jiang, Z.-T. Li, *J. Am. Chem. Soc.* **2004**, *126*, 12386; b) C. Li, G.-T. Wang, H.-P. Yi, X.-K. Jiang, Z.-T. Li, R.-X. Wang, *Org. Lett.* **2007**, *9*, 1797; c) M. Waki, H. Abe, M. Inouye, *Angew. Chem. Int. Ed.* **2007**, *46*, 3059; *Angew. Chem.* **2007**, *119*, 3119; d) H. Abe, H.

Machiguchi, S. Matsumoto, M. Inouye, *J. Org. Chem.* **2008**, *73*, 4650; e) J. Y. Hwang, H.-G. Jeon, Y. R. Choi, J. Kim, P. Kang, S. Lee, K.-S. Jeong, *Org. Lett.* **2017**, *19*, 5625.

[8] a) N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil, I. Huc, *Nat. Chem.* **2015**, *7*, 334; b) S. Saha, B. Kauffmann, Y. Ferrand, I. Huc, *Angew. Chem. Int. Ed.* **2018**, *57*, 13542; *Angew. Chem.* **2018**, *130*, 13730.

[9] a) Y. Ferrand, I. Huc, *Acc. Chem. Res.* **2018**, *51*, 970; b) G. Lautrette, B. Wicher, B. Kauffmann, Y. Ferrand, I. Huc, *J. Am. Chem. Soc.* **2016**, *138*, 10314.

[10] P. K. Mandal, B. Kauffmann, H. Destecroix, Y. Ferrand, A. P. Davis, I. Huc, *Chem. Commun.* **2016**, *52*, 9355.

[11] a) F. C. Tucci, D. M. Rudkevich, J. Rebek, Jr., *J. Am. Chem. Soc.* **1999**, *121*, 4928; b) A. Shivanyuk, J. Rebek, Jr., *J. Am. Chem. Soc.* **2002**, *124*, 12074; c) A. Scarso, A. Shivanyuk, J. Rebek, Jr., *J. Am. Chem. Soc.* **2003**, *125*, 13981; d) A. Scarso, A. Shivanyuk, O. Hayashida, J. Rebek, Jr., *J. Am. Chem. Soc.* **2003**, *125*, 6239; e) L. C. Palmer, Y.-L. Zhao, K. N. Houk, J. Rebek, Jr., *Chem. Commun.* **2005**, 3667.

[12] a) J. Chen, J. Rebek, Jr., *Org. Lett.* **2002**, *4*, 327; b) M. Yoshizawa, Y. Takeyama, T. Okano, M. Fujita, *J. Am. Chem. Soc.* **2003**, *125*, 3243; c) M. Yoshizawa, M. Tamura, M. Fujita, *Science* **2006**, *312*, 251; d) Y. Nishioka, T. Yamaguchi, M. Yoshizawa, M. Fujita, *J. Am. Chem. Soc.* **2007**, *129*, 7000; e) J.-L. Hou, D. Ajami, J. Rebek, Jr., *J. Am. Chem. Soc.* **2008**, *130*, 7810; f) T. Murase, S. Horiuchi, M. Fujita, *J. Am. Chem. Soc.* **2010**, *132*, 2866.

[13] a) M. Yoshizawa, M. Tamura, M. Fujita, *J. Am. Chem. Soc.* **2004**, *126*, 6846; b) I. Hwang, A. Y. Ziganshina, Y. H. Ko, G. Yun, K. Kim, *Chem. Commun.* **2009**, 416.

[14] a) T. Heinz, D. M. Rudkevich, J. Rebek, Jr., *Nature* **1998**, *394*, 764; b) J. Rebek, Jr., *Angew. Chem. Int. Ed.* **2005**, *44*, 2068; *Angew. Chem.* **2005**, *117*, 2104; c) M. Yamashina, M. M. Sartin, Y. Sei, M. Akita, S. Takeuchi, T. Tahara, M. Yoshizawa, *J. Am. Chem. Soc.* **2015**, *137*, 9266.

[15] E. Berni, B. Kauffmann, C. Bao, J. Lefeuvre, D. M. Bassani, I. Huc, *Chem. Eur. J.* **2007**, *13*, 8463.

[16] a) P. Thordarson, *Chem. Soc. Rev.* **2011**, *40*, 1305; b) C. A. Hunter, H. L. Anderson, *Angew. Chem. Int. Ed.* **2009**, *48*, 7488; *Angew. Chem.* **2009**, *121*, 7624.

[17] a) V. Berl, I. Huc, R. Khoury, M. J. Krische, J.-M. Lehn, *Nature* **2000**, *407*, 720–723; b) Y. Ferrand, Q. Gan, B. Kauffmann, H. Jiang, I. Huc, *Angew. Chem. Int. Ed.* **2011**, *50*, 7572; *Angew. Chem.* **2011**, *123*, 7714.

[18] The diastereomeric excess was calculated here using (([*P*-complex]−[*M*-complex])/([*P*-complex]+[*M*-complex]).

[19] Q. He, G. I. Vargas-Zúñiga, S. H. Kim, S. K. Kim, J. L. Sessler, *Chem. Rev.* **2019**, *119*, 9753.