A Biomimetic Synthetic Receptor Selectively Recognising Fucose in Water

CHEMISTRY
A European Journal

A Journal of
ChemPubSoc Europe

Accepted Article

Title: A biomimetic synthetic receptor selectively recognizing fucose in water.

Authors: Oscar Francesconi, Marco Martinucci, Lorenzo Badii, Cristina Nativi, and Stefano Roelens

To be cited as: *Chem. Eur. J.* 10.1002/chem.201800390

Link to VoR: http://dx.doi.org/10.1002/chem.201800390

Supported by
ACES

WILEY-VCH

10.1002/chem.201800390

FULL PAPER

A biomimetic synthetic receptor selectively recognizing fucose in water.

Oscar Francesconi,* [a] Marco Martinucci, Lorenzo Badii, Cristina Nativi, and Stefano Roelens

Abstract: Carbohydrate recognition in water by biomimetic receptors is an attractive but very challenging goal. Despite the advances achieved on glucose recognition, little or no success has been obtained in the recognition of other saccharidic epitopes of paramount importance in biological processes. Here we report the unprecedented recognition of fucose in water by an artificial receptor that shows affinities closely comparable to those of several lectins. The receptor has been build-up by assembling a hydrogen bonding element (carbazole), a hydrophobic aromatic moiety (anthracene), and a water-solubilizing function (phosphonate) into a macrocyclic structure providing the appropriate binding geometry. The described receptor binds fucose with sub-millimolar affinity in water at physiological pH, showing enthalpic binding that can be ascribed to H-bonding to saccharidic hydroxyls and to CH-π interactions between the sugar backbone and the aromatic moieties. Experimental NOE contacts coupled to conformational search calculations return a picture of a binding site in which fucose assumes a staggered orientation reminiscent of that shown by fucose when bound to the bacterial lectin RSL.

Introduction

Carbohydrates, assembled on the surface of eukaryotic cells as glycoconjugates of lipids and proteins, the so called glycocalix, are essential mediators of a wide range of both physiological and pathological processes.[1] All these processes rely on the recognition of specific saccharidic units, exerted by several classes of proteins, such as lectins.[2] The presence of epimers and the polar hydrophilic nature of saccharides make their recognition a non trivial task for lectins, which often bind to monosaccharides with modest selectivity and weak affinities. [3,4] Although widely used as tools in glycobiology and diagnostics, development of lectins as therapeutics is hampered by their proteic nature.[5-7] Biomimetic synthetic receptors for carbohydrates, often called “synthetic lectins”,[8-10] developed for mimicking the function of lectins in specific recognition events,[11-14] are small-molecules exploiting the noncovalent interactions used by their natural counterparts to bind saccharides.[15,16] Although over the last decades a significant effort has been dedicated to the design of biomimetic receptors for carbohydrates,[17-20] at present only a few examples of structures effective in the competitive acqueous environment have been reported.[21] While good affinities toward charged sugars were reported,[22,23] the recognition of neutral carbohydrates is generally more challenging, so that effective recognition was achieved only toward “all-equatorial” carbohydrates, like glucose and related sugars.[24-28] Because many biologically relevant carbohydrates present α glycosidic linkages and/or axial hydroxyls in their structures, it would be of paramount importance to expand the realm of biomimetic receptors to different classes of saccharides, such as galactose, fucose, mannose, and N-acetylgalactosamine.[2] Among these, L-fucose (6-deoxy-L-galactose) is peculiar for two reasons: a) is a rare example of naturally occurring saccharide of the L-series, and b) it lacks the hydroxyl in the 6 position. In mammals, L-fucose, generally linked as α anomer, is a common component of many N- and O-glycans, crucial in transfusion reactions, selectin mediated leukocyte-endothelial adhesion, host-microbe interactions, and numerous ontogenic events.[29] Indeed, fucose is present in most of Lewis antigens, such as the blood group antigens (A, B and H) and the sialyl-Lex.[30] Moreover, recognition of fucosylated glycans mediated by bacterial lectins is a critical step of the infection mechanism by several pathogens of high health risk, like *Pseudomonas aeruginosa*.[31] Alterations in the expression of fucosylated oligosaccharides have also been observed in several pathological processes, including cancer and atherosclerosis.[32,33] For example, abnormal fucosylation in cancer cells induces the formation of fucosyl-GM1, a ganglioside tumoral antigen associated with small-cell lung carcinoma.[34]

Herein we report design, synthesis, binding properties, and structural studies of the first member of a new generation of biomimetic synthetic receptors that shows remarkable affinity for fucose in water. To the best of our knowledge, recognition of fucose in water by biomimetic receptors is unprecedented in the chemical literature.

Results and Discussion

Design and synthesis. The design of the receptor was based on three constituting elements: i) a hydrogen (H)-bonding unit, ii) an aromatic extended π unit, and iii) a hydrosolubilizing group, which were assembled into a macrocyclic architecture providing the correct binding geometry (Figure 1a). Based on our previous work on diaminopyrrolic structures,[35-38] 1,8-diaminocarbazole was selected as the appropriate H-bonding unit, because the analogous diaminopyrrolic chelating arrangement of Figure 1b was shown to be optimal for H-bonding to the saccharidic hydroxyls. Replacement of the diaminopyrrolic with a diaminocarbazolic unit was believed to be a significant

10.1002/chem.201800390

FULL PAPER

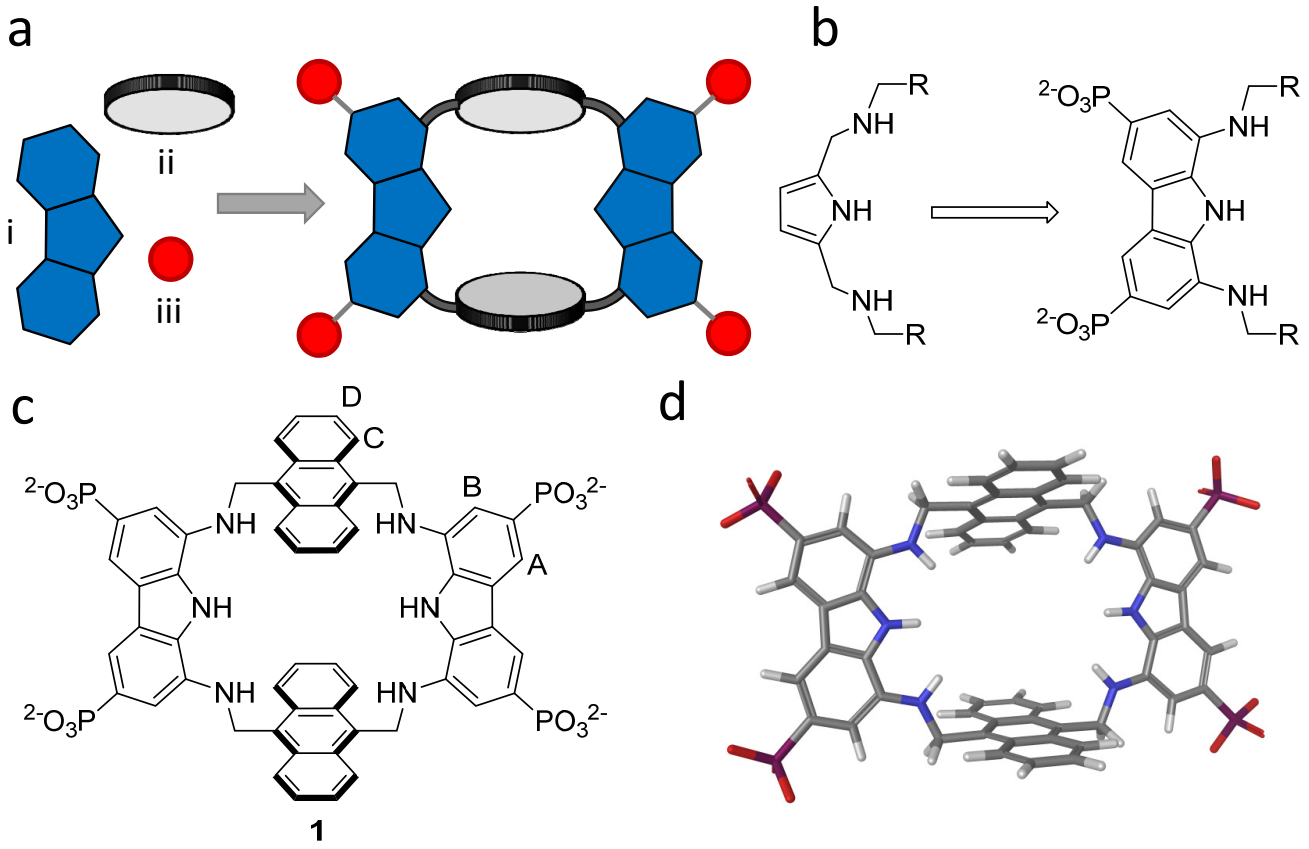


Figure 1. Design of receptor 1. a) constituting elements of the receptor design: i) H-bonding unit; ii) aromatic extended π unit; iii) hydrosolubilizing group. b) Replacement of the H-bonding unit preserving the chelating arrangement. c) Molecular structure of receptor **1** with proton labelling. d) Minimum energy conformation of **1**, as predicted by molecular mechanics conformational search, showing the convergent disposition of H-bonding groups in a shape-persistent aromatic cavity endowed with outward-facing solubilizing groups.

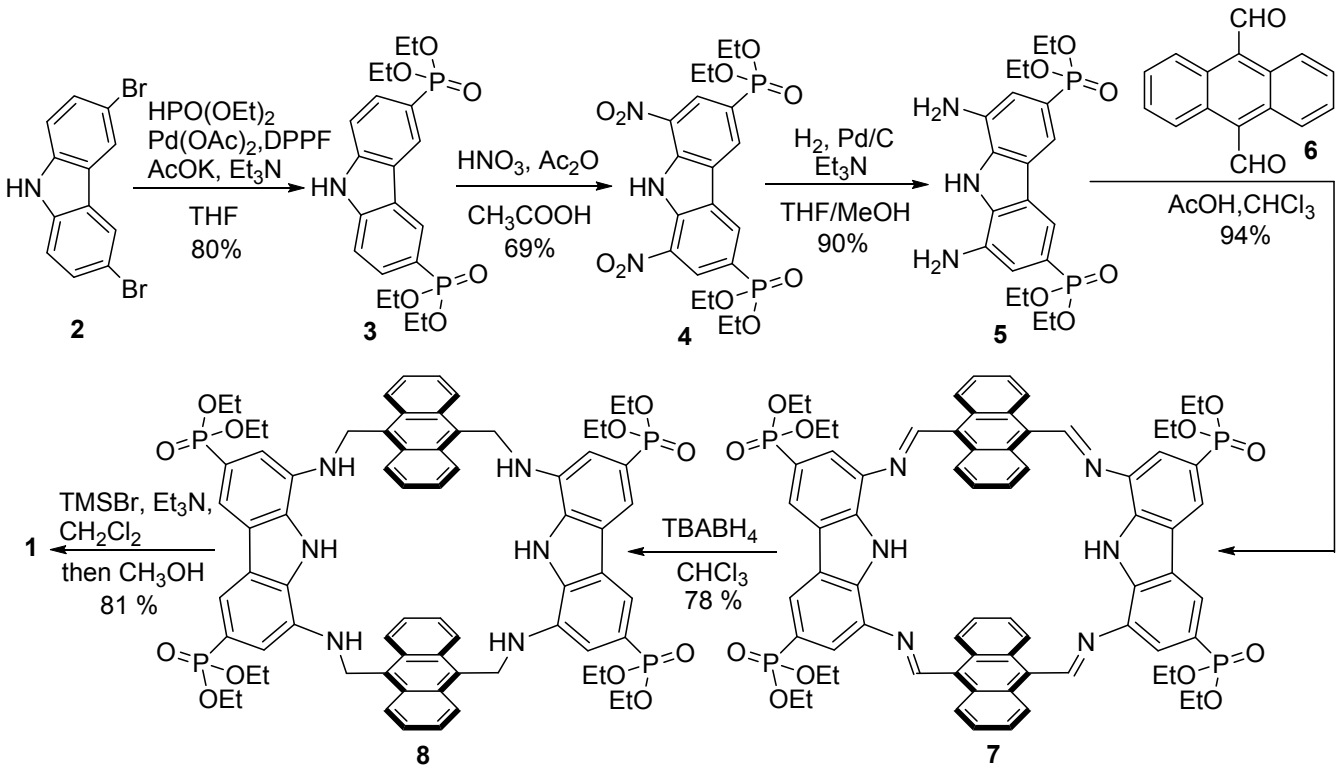
improvement, imparting rigidity to the building block while preserving the same H-bonding groups arrangement. Additionally, the decreased basicity of the aromatic amino groups should be beneficial to the binding ability, which is inhibited by protonation on nitrogen.[39] The carbazole unit is decorated with two phosphonate groups to ensure appropriate solubility in water, while an anthracene moiety provides the extended π unit required to establish CH-π interactions with the saccharide backbone and to create a shape persistent hydrophobic cavity. Eventually, the constituting elements are assembled into a macrocyclic structure of the appropriate size to host a monosaccharidic guest, in which all binding elements are convergent. The result of this design is the structure **1** shown in Figure 1c, whose 3D model is shown in Figure 1d, as predicted by a molecular mechanics conformational search.

Receptor **1** was easily prepared on a multigram scale in six steps with 30 % overall yield, from inexpensive, commercially available 3,6-dibromocarbazole **2**, as outlined in Scheme 1. The diphosphonate **3** was obtained from **2** by a Pd-catalyzed aromatic variant of the Arbuzov reaction, and subsequently nitrated in 1,8 to afford the dinitrocarbazole **4**. The latter was hydrogenated to the diaminocarbazole **5** and condensed with 9,10-diformylanthracene **6** to afford the macrocyclic tetraimine **7** in high yield without resorting to high dilution techniques. Borohydride reduction of **7** gave the corresponding tetraamine, which was hydrolized to **1** by reacting the phosphonate tetraester **8** with trimethyl bromosilane.

Binding studies. The affinities of receptor **1** toward saccharides were determined by NMR and ITC techniques. In a preliminary screening by ^{1}H-NMR, the binding ability of **1** was tested towards a set of pentoses, hexoses and 9-carbon monosaccharides, including glucose (Glc), rhamnose (Rha), fucose (Fuc), xylose (Xyl) and sialic acid (Neu5Ac) (Chart 1), monitoring the shifts of the proton signals of the sugar upon addition of an equimolar amount of **1**. While in some cases little (Rha, Xyl) or no (Neu5Ac) variations were observed, a marked upfield shift was appreciated for Glc, larger for the β anomer than for the α anomer, as well as for Fuc, for which a larger shift was observed for the α anomer, with concomitant broadening of

10.1002/chem.201800390

FULL PAPER



Scheme 1. Schematic pathway for the synthesis of receptor **1**.

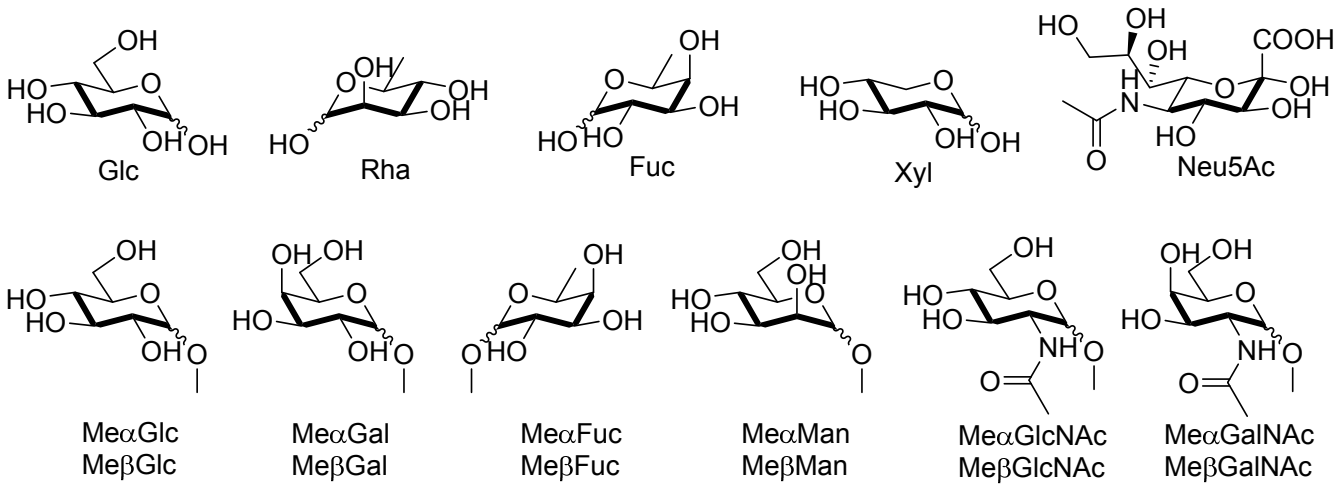


Chart 1. Monosaccharides used in recognition experiments.

signals, suggesting slow exchange likely due to strong binding (Figures S12, S13). A more detailed and quantitative investigation was thus carried out on Glc and Fuc, together with related monosaccharides, including galactose (Gal), mannose (Man), N-acetylglucosamine (GlcNAc) and N-acetylgalactosamine (GalNAc), all of which were used as α and β methyl glycosides to ascertain any anomeric selectivity (Chart 1).

A detailed analysis of the receptor behaviour in solution was preliminarily required for a quantitative investigation. A titration showed that **1** is freely soluble in water over a range of pH from 3.5 to 12 (Figure S14). Precipitation is observed at pH < 3.5, due

10.1002/chem.201800390

FULL PAPER

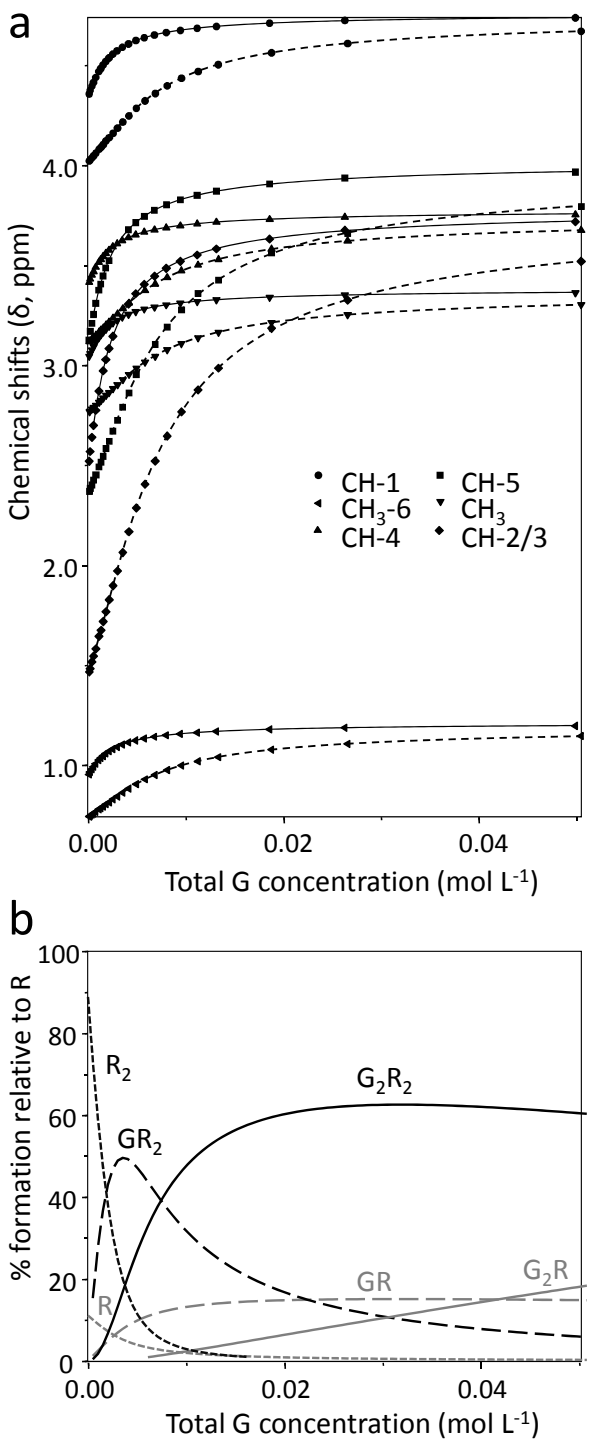


Figure 2. ^{1}H-NMR spectroscopic titrations (500 MHz, D_2O, pD 11, 298 K) of receptor **1** (R) with MeαFuc (G). a) Plot of the complexation-induced shifts of the glycoside proton signals along the titrations. Symbols are experimental data points; lines (titration 1, [R] = 1.03 10^{-3} mol L^{-1}, solid lines; titration 2, [R] = 4.02 10^{-3} mol L^{-1}, dashed lines) are best fit curves obtained through nonlinear regression by simultaneous fit of all available signals from the two independent titrations at different reactant concentrations. b) Distribution diagram of the species along the titration in percentage relative to R with [R] = 1.03 10^{-3} mol L^{-1}.

to high degree of protonation of phosphonate groups. At pH > 11 receptor **1** is fully deprotonated and shows sharp signals in ^{1}H-NMR spectra. At pH < 11, phosphonate protonation gives rise to several partially protonated microspecies; correspondingly, ^{1}H-NMR spectra show several set of overlapping broad signals (Figure S15). Because this feature complicates the investigation of binding properties by ^{1}H-NMR titrations at physiological pH (7.4), affinities were determined at pH 11, using all the available signals from the receptor and the saccharide, subject to verification that under these conditions affinities would not be affected. To this end, binding constants for MeαFuc were measured both, at pH 11 by NMR (Figure S17), and at pH 7.4 by ITC and by ^{1}H-NMR titrations following the saccharide signals. Allowing for the different accuracy obtained when using only the shifts of the sugar, the comparable affinities observed (see below in Table 1) demonstrated that binding properties were hardly affected by pH variations and that, therefore, the values measured at pH 11 reliably describe the binding ability of **1** also at physiological pH.

Finding the model. Apart from the above pH dependence, occurrence of receptor self-association was clearly evident from chemical shift changes with concentration; therefore, multiple association equilibria must be expected for the interaction of **1** with saccharides. Dilution experiments of receptor **1** fitted a self-association model including three clusters, in which the dimer was by far the dominant species (see Supporting Information, pp. S24-S26). The fit gave a dimerization constant log β_{dim} = 3.95±0.11, which was set invariant in the nonlinear regression analysis of the receptor-glycoside binding data. The strong self-association, most likely due to π-stacking of the aromatic moieties, generates a species distribution in solution in which the dimeric form of the receptor is prevalent, giving rise to complexes featuring two binding cavities that can be filled by 1 or 2 sugar molecules, and consequently to stoichiometries higher than 1:1. Cumulative association constants were thus measured by ^{1}H NMR titrations in D_2O (pD 11) at 298 K according to a previously described protocol,[40] which relies on simultaneous fit of the complexation induced shifts of all the available signals from both the receptor and the glycoside to the appropriate association model by nonlinear regression analysis. Where necessary, to avoid high correlation of association constants in multiple binding equilibria, data from two independent titrations at different reactant concentrations were simultaneously fitted. The association model most closely describing the system was assessed by careful analysis of the error of the fit, scrutinizing models including an increasing number of species of higher stoichiometry. The results of such an analysis for binding of MeαFuc to **1** are reported in Supporting Information (pp. S27-S46). It can be easily appreciated that the simplest model giving a satisfactory fit includes, in addition to the receptor dimer and the 1:1 adduct, also the G_2R, the GR_2, and the G_2R_2 species (G = glycoside, R = receptor) (Figure 2a), whereas any simpler model appears clearly inadequate. Furthermore, from the species distribution plot (Figure 2b and p. S43) it can be noted that the dominant complexes are the GR_2 species below 8 mM, and the G_2R_2

10.1002/chem.201800390

Table 1. Cumulative formation constants (log β_n)[a] and intrinsic median binding concentration (BC_{50}^0, mM)[b] for receptor to glycoside (R:G) complexes of **1** with methyl glycosides, measured at 298 K from NMR data in D_2O at pD11[c] and from ITC data in H_2O at pH 7.4.[d]

	R:G	NMR		ITC	
Glycoside		log β	BC_{50}^0	log β	BC_{50}^0
MeαGlc	1:1	2.35±0.03	3.12±0.63		7.40±3.17
	2:1	6.45±0.02		6.03±0.01	
	2:2	7.88±0.05			
MeβGlc	2:1	6.91±0.01	1.30±0.24	6.92±0.04	0.87±0.23
	2:2	8.82±0.03		9.53±0.13	
	1:2	4.57±0.05		5.51±0.07	
MeαGal	1:1	2.88±0.04	1.19±0.20		1.01±0.31
	2:1	6.81±0.02		6.91±0.03	
	2:2	8.92±0.05		9.26±0.04	
	1:2	4.94±0.07			
MeβGal	1:1	2.69±0.03	7.90±1.35		14.9±6.6
	2:1	5.58±0.01		5.71±0.02	
	1:2	3.99±0.04			
MeαFuc[e]	1:1	3.28±0.02	0.36±0.09		0.52±0.14
	2:1	7.34±0.01		7.26±0.02	
	2:2	9.83±0.03		9.88±0.02	
MeαFuc	1:1	2.98±0.03	0.65±0.11		
	2:1	7.15±0.01			
	2:2	9.50±0.02			
	1:2	4.41±0.12			
MeβFuc	2:1	5.67±0.01	20.6±5.0	5.85±0.02	10.9±4.8
MeαMan[f]					
MeβMan	2:1	5.58±0.02	25.2±6.3		
MeαGlcNAc	2:1	5.40±0.06	37.7±10.6		
MeβGlcNAc	2:1	5.68±0.01	20.1±4.9		
MeαGalNAc[f]					
MeβGalNAc[f]					

[a] Formation constants were obtained by nonlinear least-square regression analysis of NMR and ITC data. [b] Calculated from the log β values using the "BC50 Calculator" program.[41] [c] The receptor dimerization constant at pD 11 (log β_{dim} = 3.95±0.11) was measured independently and set invariant in the nonlinear regression analysis of NMR data measured at pD 11. [d] The dimerization constant at pH 7.4 (log β_{dim} = 3.84±0.20) was optimized in the ITC titration of receptor **1** with MeαFuc and set invariant in the non-linear regression analysis of NMR and ITC data measured at pH 7.4. [e] from NMR data in D_2O at pD 7.4. [f] no chemical shift variations were detected.

10.1002/chem.201800390

species above 8 mM, as expected from the prevalence of the dimeric receptor, accounting for 50 to 60% of bound species at least up to 50 mM, whereas the 1:1 adduct never exceeds 15% of bound species. In this context it is worth noting that the naive approach of considering a model including the 1:1 adduct only (pp. S44-S46), which may appear good at first sight when looking at the shifts of the sugar signals exclusively, is indeed incorrect and misleading, as can be appreciated from the plot of the receptor chemical shifts and from the systematic, abnormally high error of the sugar signals.

The results obtained following the above approach for the whole set of sugars are reported in Table 1 as cumulative formation constants. Because multiple binding constants were measured in all cases, affinities were assessed through the BC_{50}^{0} parameter, a generalized affinity descriptor univocally defining the overall binding ability of a receptor in chemical systems involving multiple complex species.[40-45] The BC_{50}^{0} (intrinsic median binding concentration) descriptor, is defined as the total concentration of receptor necessary for binding 50% of the ligand when the fraction of bound receptor is zero, that is, when forming the first complex molecule, and is calculated from binding constants taking into account all the species contributing to ligand recognition. For 1:1 association equilibria, BC_{50}^{0} coincides with the dissociation constant K_d. The BC_{50}^{0} values calculated from the measured binding constants are also reported in Table 1.

At a first glance, from the results reported in Table 1, four classes of recognition levels can be identified: a) top affinity: MeαFuc is bound with an outstanding affinity in the micromolar range. Not only such an affinity in water is remarkable, but also, to the best of our knowledge, biomimetic receptors for fucose are unprecedented in the chemical literature. Likewise, the α/β selectivity shows an outstanding value larger than 30-fold, which places **1** within the most in-teresting receptors to date. b) high affinity: Glc and Gal, in both anomeric forms, display affinities in the low millimolar range (1-8 mM). While a 1.3 mM affinity for the β anomer of Glc is in line with the highest values reported in the literature, the 3.1 mM affinity for the α anomer, hardly recognized by "synthetic lectins" reported as potential tools for glucose monitoring, is noteworthy.[26] Even more noteworthy is recognition of galactosides: specifically, receptor **1** appears to be one of the most effective examples of biomimetic receptors recognizing α-galactosides in water,[46] showing a remarkable 1.2 mM affinity for MeαGal. c) low affinity: Apart from the above discussed MeβFuc, the α and β anomers of GlcNAc and MeβMan show affinity values between 20 and 40 mM, which places **1** within the class of interesting receptors, even though not exceptional. d) no affinity: **1** appears to be essentially incapable of binding to both anomers of GalNAc and to MeαMan, suggesting that requirements for binding are particularly strict toward substituents in the 2-position of the monosaccharide.

The peculiar feature emerging from inspection of Table 1 is that the receptor cavity, besides accepting the "all-equatorial" glucoside, appears to be willing to preferentially accomodate 1/4 axial and 1,4-diaxial hydroxyls, which is an unexpected novelty in the biomimetic receptors panorama. Furthermore, considering that Fuc is the enantiomer of 6-deoxy-Gal, receptor **1** appears to be suited for the selective recognition of this specific monosaccharidic structure.

It is worth noting that the observed affinities are closely comparable to those reported for several fucose/galactose binding natural lectins, such as, for example, tunicate C-type lectin (TC14) from *Polyandrocarpa misakiensis*,[47] winged bean basic lectin (WBAI) from *Psophocarpus tetragonolobus*,[48] and human galectin-1,[49] which places **1** in the range of receptors useful for biological applications.

Investigation by spectrophotometric techniques was attempted exploiting the carbazole and anthracene UV/fluorescence chromophores. UV/Vis studies carried out on the constituting elements of receptor **1** showed a hyperchromic effect when diaminocarbazole **5** and anthracene are assembled into the macrocyclic structure (Figure S21). Unfortunately, addition of MeαFuc to **1** induces only a modest change in absorbance, which cannot be used to accurately determine binding constants (Figure S22). Disappointingly, even fluorescence spectroscopy could not be employed because, although diaminocarbazole **5** and anthracene are good fluorophores (Figures S23 and S24), fluorescence is strongly quenched when they are assembled into the macrocyclic structure and is not turned on upon addition of the fucoside (Figure S25).

The binding affinities obtained by NMR spectroscopy were further confirmed by ITC, as an independent technique, in H_2O at physiological pH 7.4, for the glucose, galactose and fucose glycosides (see Supporting Information, pp. S96-S138). To remove ambiguities in the definition of the binding model, 3 to 5 independent titrations run at different reactant concentrations were combined into a simultaneous fit of all data. The receptor dimerization constant at pH 7.4 (log β_{dim} = 3.84±0.20), optimized in the data analysis of the ITC titrations of receptor **1** with MeαFuc (Figure S26), turned out to be closely similar to that found by NMR at pD 11 and was set invariant in the non linear regression analysis of receptor-glycoside binding data performed at pH 7.4, both by ITC and NMR technique. Cumulative association constants, together with affinity values, are reported in Table 1 for a direct comparison with the NMR data. The generally good agreement between the ITC and the NMR results, observed for the most relevant cases, confirms the above discussed affinities. The agreement is, however, not excellent in all cases. The discrepancy is clearly apparent in the models detected from the two techniques, as up to four binding constants were measured by NMR, whereas no more than two (three in only one case) could be obtained by ITC. Such a discrepancy can be ascribed to the higher definition of the equilibrium systems that can be obtained by NMR with respect to ITC, because the larger number of data points (many shift data for several signals of both reactants) and the higher sensitivity of the shift modulation to the presence of different complex species, allow for a much finer deconvolution of the binding isotherm by the former technique.[50] As a result, the number of species (and therefore of binding constants) that can be appreciated by ITC is limited by the intrinsic resolution of the technique.

10.1002/chem.201800390

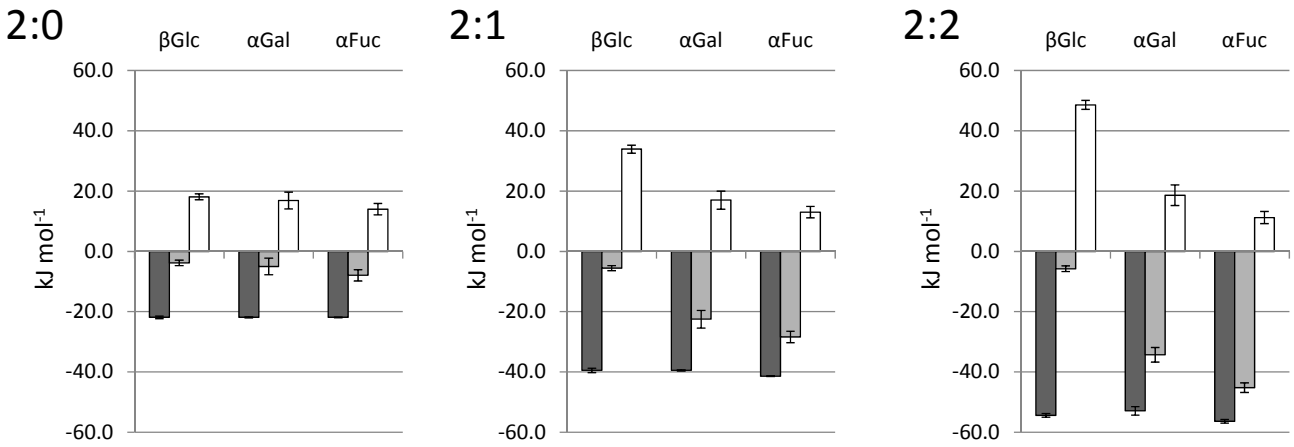


Figure 3. Thermodynamic parameters. Bar plot (with standard deviation) of thermodynamic parameters (dark grey, ΔG^0; light grey, ΔH^0; and white, $-T\Delta S^0$) for the formation of receptor to glycoside (R:G) complexes of MeβGlc, MeαGal and MeαFuc in H_2O, pH 7.4 at 298 K.

Thermodynamic parameters obtained from ITC data for the formation of the receptor dimer (R_2) and of the major complexes (R_2G and R_2G_2) between **1** and three glycosides (MeβGlc, MeαGal and MeαFuc) are reported in Table S2 and presented as a bar plot in Figure 3 for comparison. Quite interestingly, while the dimer formation is invariably entropically driven, as can be expected for the hydrophobic stacking of aromatic units, a shift from entropic to enthalpic driving force is appreciated for complexes of Gal and Fuc with respect to those of Glc. This evidence is consistent with a stronger contribution of H-bonding for the former complexes with respect to the latter, suggesting that desolvation gives a major contribution to the inclusion of glucosides into the hydrophobic cavity, whereas H-bonding appears to be the driving force for complexation of galactosides and even more of fucosides.

Structural studies. To verify the latter hypothesis and to gain a deeper insight into the binding mode characterizing the complex between **1** and MeαFuc, a description of the binding mode was attempted by combining NMR techniques with molecular modeling calculations. Experimental information was obtained by NOESY experiments performed at pD 11, from which unambiguous intermolecular NOE contacts were identified (Figure S29), and schematically represented in Figure 4 (see also Table S3). The strongest contacts were observed between the H-C anthracene protons and the H-1, the H-4, and the CH_3O protons of the fucoside (for receptor labelling, see Figure 1); a strong contact was also observed between the H-D proton of the anthracene and the H-6 of fucose. The evidence that intermolecular cross-peaks were observed between the anthracene/methylene protons of the receptor and the protons of both the α and the β face of the fucoside strongly supports that the sugar is located inside the receptor cavity.

Based on experimental NOE data, a well-tested molecular mechanics computational protocol[51,52] (see Supporting Information) has been applied to attempt a three-dimensional description of the complex between **1** and MeαFuc. Because NOE contacts provided information on the proximity of sugar protons to the receptor cavity, based on the assumption that dimerization of the receptor is most likely caused by π-π stacking of the anthracene moieties in water,[53] the binding site has been modeled using the cavity of the monomeric receptor as the active structure and considering the dimeric receptor as constituted by two independent binding sites. Under this assumption, the conformational search of the molecular mechanics protocol returned only one family of conformers in very good agreement with the NMR data within 6 kJ mol^{-1} from

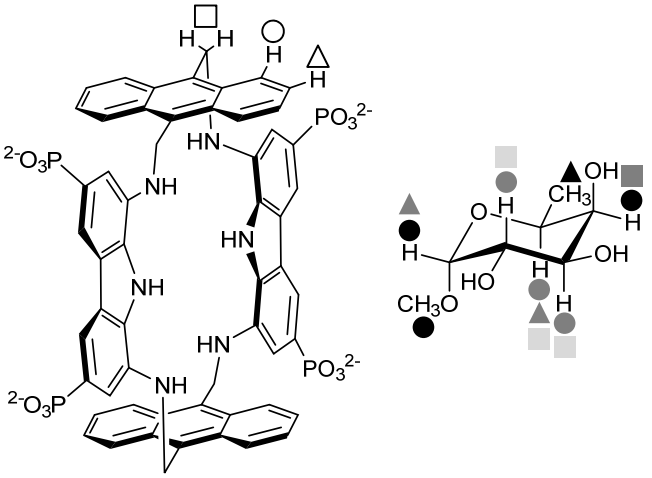


Figure 4. Schematic representation of intermolecular NOE cross-peaks found in the NOESY spectrum of an equimolar mixture of MeαFuc and **1**. NOE contacts of the sugar protons with the C and D protons of the anthracene moiety (circles and triangles, respectively) and with the methylene protons (squares) are depicted with relative intensities shown as dark grey (strong), medium grey (medium), and light grey (weak).

10.1002/chem.201800390

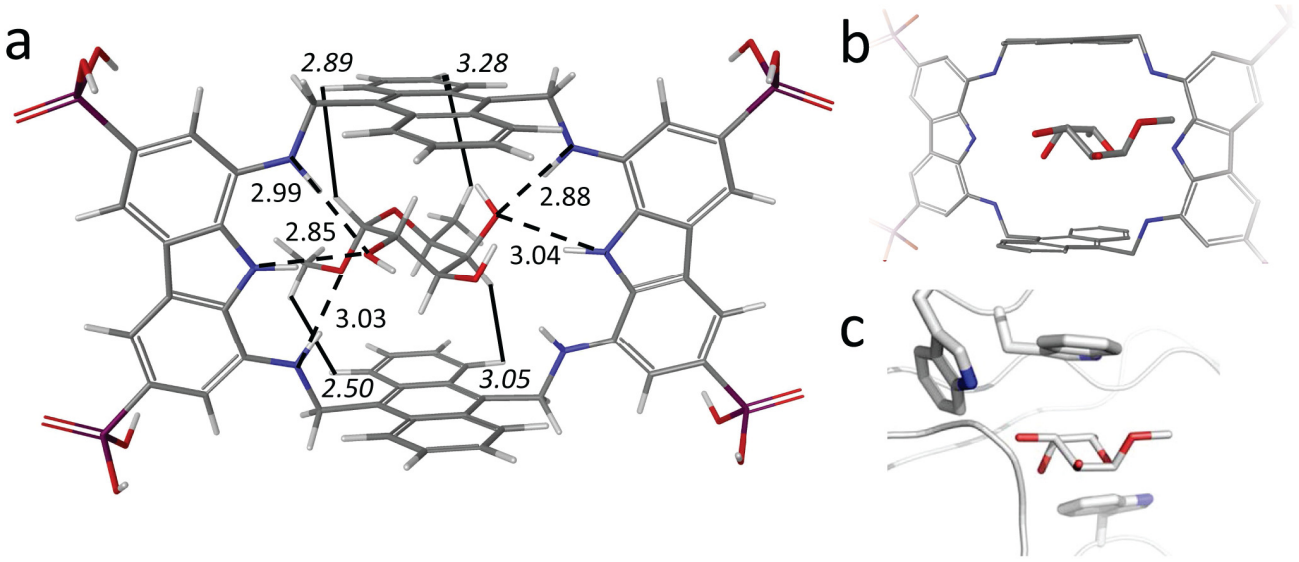


Figure 5. Structural studies of the complex between **1** and MeαFuc. a) Global minimum structure of the **1**•MeαFuc complex. The strongest intermolecular NOEs found between **1** and MeαFuc are indicated as solid lines together with the corresponding distances (Å) calculated for the lowest energy conformer. Hydrogen-bonding interactions found in the calculated structure are indicated in dashed lines together with the corresponding oxygen/nitrogen distances (Å). b) Structure of the receptor binding cleft showing the orientation of MeαFuc in the minimum energy conformation (hydrogens were omitted for clarity). c) A simplified representation of the binding site of the fucose-binding lectin (RSL) from the bacterium *Ralstonia solanacearum* in the complex with MeαFuc. (Adapted by permission from ref. [54]. Copyright (2016) American Chemical Society).

the global minimum. The global minimum structure, depicted in Figure 5a and in Figure S32, shows the MeαFuc bound inside the receptor cavity in a tilted geometry, orienting the methyl group at the 6-position toward one of the CH-D protons of the anthracene, and the CH_3O, the H-1, and the H-4 protons toward three CH-C protons, in agreement with the proximities inferred from the strongest NOE contacts. Intermolecular proton distances ranging from 3.3 to 4.2 Å and from 4.4 to 5.1 Å, compatible with the observed medium and weak NOE contacts, respectively, were also found in the global minimum structure (Figure 5a).

From the above model, all O•••N interatomic distances shorter than the sum of the van der Waals radii could be calculated, selecting those distances that comply with H-bonding criteria (Figure 5a). Several H-bonding interactions have thus been detected, showing that both carbazolic NH and three of the four aminic NH are indeed involved in H-bonding to the fucoside OH-2, OH-4, and the α-anomeric oxygen. CH-π interactions from the axial protons facing the anthracene rings at relatively short distances (H-2, 2.52 Å; H-3, 2.49 Å; H-4, 2.78 Å) provide additional contribution to binding. This picture sheds light on the origin of the observed 1,4-diaxial preference shown by the receptor and the crucial role of the 2-position in rejecting substituents other than the equatorial hydroxyl, such as Man, GlcNAc and GalNAc. It also explains the strong enthalpic contribution to binding, due to the extensive of H-bonding network established with the fucoside, which, synergically acting with hydrophobic desolvation, compete favourably with the acqueous solvent.

Very interestingly, a close similarity could be found between the orientation of MeαFuc inside the binding cavity of **1** and inside the binding site of the fucose-binding lectin (RSL) from bacterium *Ralstonia solanacearum*, indicating an unforeseen correlation between this artificial receptor and the natural lectins (Figure 5b,c).[54] As a matter of fact, in the binding pocket of RSL, the MeαFuc features a titled orientation with respect to the indole ring of one Trp, closely similar to that observed in the cavity of **1** with respect to the anthracene moieties. Moreover, in the RSL complex, the NH of a second Trp is H-bonded to one of the hydroxyl groups of MeαFuc, in a similar fashion to carbazolic NH in receptor **1**.

Conclusions

In the present work, we have shown that a water soluble synthetic receptor mimicking the saccharide recognition function of a lectin can be rationally designed and easily prepared. The binding properties of receptor **1**, as investigated by NMR and ITC techniques, showed that effective recognition of non “all equatorial” monosaccharides can be achieved in water at physiological pH with affinities comparable to several lectins. Unprecedented for a synthetic receptor, the α anomer of fucose is recognized with an outstanding affinity of 360 μM and an α/β selectivity of over 30-fold, whereas α-galactose shows an unforeseen affinity of 1 mM for **1**. Calorimetric measurements showed that recognition is enthalpically driven, which is unusual in water, due to extensive H-bonding, while structural studies

indicated a binding mode reminiscent of that of fucose into the binding pocket of RSL, a fucose bacterial lectin. In the current panorama of synthetic receptors for carbohydrates, the carbazolic structure described here represents a new tool for glycobiology, alternative to natural lectins, opening the way to potential therapeutic applications.

Acknowledgements

This work was financially supported by the Ministero dell'Istruzione dell'Università e della Ricerca (MIUR) through PRIN2010-2011 (grant no. 2010JMAZML) and PRIN2015 (grant no. 2015RNWJAM). Ente Cassa di Risparmio di Firenze (Italy) is acknowledged for granting an ITC nano calorimeter and a high-filed NMR. UV/Vis and fluorescence were acquired at the Magnetic Resonance Center (CERM) of the University of Florence. HRMS were acquired at Centro di Servizi di Spettrometria di Massa (CISM) of the University of Florence. We are grateful to the Erasmus Plus exchange program for a grant to Ms Dafni Chandrinou, University of Crete, who participated in the synthetic work and to Professor Andrea Bencini, Department of Chemistry, University of Florence, for the potentiometric measurements.

Keywords: synthetic receptors • molecular recognition • carbohydrates • macrocycles • selective binding

[1] B. Ernst, W. Hart, P. Sinay, *Carbohydrates in Chemistry and Biology*, Wiley-VCH, Weinheim, Germany, **2000**.
[2] H.-J. Gabius, *The Sugar Code: Fundamentals of Glycosciences*, Wiley-VCH, Weinheim, Germany, **2009**.
[3] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*,719-728.
[4] M. Ambrosi, N. R. Cameron, B. G. Davis, *Org. Biomol. Chem.* **2005**, *3*, 1593-1608.
[5] B. Wang, G.-J. Boons, *Carbohydrate Recognition: Biological Problems, Methods and Applications*, Wiley, Hoboken, New Jersey, **2011**.
[6] J. Balzarini, *Nat. Rev. Microbiol.* **2007**, *5*, 583-597.
[7] Y. Nakagawa, T. Doi, T. Taketani, K. Takegoshi, Y. Igarashi, Y. Ito, *Chem. Eur. J.* **2013**, *19*, 10516-10525.
[8] Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619-622.
[9] S. Kubik, *Angew. Chem. Int. Ed.* **2009**, *48*, 1722-1725.
[10] A. P. Davis, *Org. Biomol. Chem.* **2009**, *7*, 3629-3638.
[11] Y. Nakagawa, Y. Ito in *Advances in Carbohydrate Chemistry and Biochemistry*, *Vol. 68* (Ed.: D. Horton) Academic Press, Oxford, UK, **2012**, pp. 1-58.
[12] C. Nativi, O. Francesconi, G. Gabrielli, I. De Simone, B. Turchetti, T. Mello, L. Di Cesare Mannelli, C. Ghelardini, P. Buzzini, S. Roelens, *Chem. Eur. J.* **2012**, *18*, 5064-5072.
[13] O. Francesconi, C. Nativi, G. Gabrielli, I. De Simone, S. Noppen, J. Balzarini, S. Liekens, S. Roelens, *Chem. Eur. J.* **2015**, *21*, 10089-10093.
[14] S.-H. Park, Y.P. Choi, J. Park, A. Share, O. Francesconi, C. Nativi, W. Namkung, J. L.Sessler, S. Roelens, I. Shin, *Chem. Sci.* **2015**, *6*, 7284-7292.
[15] A. P. Davis, R. S. Wareham, *Angew. Chem. Int. Ed.* **1999**, *38*, 2978-2996.
[16] S. Jin, Y. Cheng, S. Reid, M. Li, B. Wang, *Med. Res. Rev.* **2010**, *30*, 171-257.
[17] M. Mazik, *RSC Adv.* **2012**, *2*, 2630-2642.
[18] J. Arnaud, A. Audfray, A. Imberty, *Chem. Soc. Rev.* **2013**, *42*, 4798-4813.
[19] A. Draganov, D. Wang, J. Holmes, K. Ji, B. Wang, S. K. Shaw, B. D. Smith in *Synthetic Receptors for Biomolecules: Design Principles and Applications* (Ed.: B. D. Smith), The Royal Society of Chemistry, Cambridge, UK, **2015**, pp. 177-203.
[20] M. Gentili, C. Nativi, O. Francesconi, S. Roelens in *Carbohydrate Chemistry, Vol. 41* (Eds.: A. P. Rauter, T. K. Lindhorst, Y. Queneau), The Royal Society of Chemistry, Cambridge, UK, **2016**, pp. 149-186.
[21] C. E. Miron, A. Petitjean, *ChemBioChem* **2015**, *16*, 365-379.
[22] Y. Jang, R. Natarajan, Y. H. Ko, K. Kim, *Angew. Chem. Int. Ed.* **2014**, *53*, 1003-1007.
[23] T. S. Carter, T. J. Mooibroek, P. F. N. Stewart, M. P. Crump, C. Galan, A. P. Davis, *Angew. Chem. Int. Ed.* **2016**, *55*, 9311-9315.
[24] Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jiménez-Barbero, C. Vicent, G.-J. Boons, S. Ingale, A. P. Davis, *Angew. Chem. Int. Ed.* **2009**, *48*, 1775-1779.
[25] N. P. Barwell, M. P. Crump, A. P. Davis, *Angew. Chem. Int. Ed.* **2009**, *48*, 7673-7676.
[26] C. Ke, H. Destecroix, M. P. Crump, A. P. Davis, *Nat. Chem.* **2012**, *8*, 718-723.
[27] P. Rios, T. S. Carter, T. J. Mooibroek, M. P. Crump, M. Lisbjerg, M. Pittlekow, N. T. Supekar, G.-J. Boons, A. P. Davis, *Angew. Chem. Int. Ed.* **2016**, *55*, 3387-3392.
[28] P. Rios, T. J. Mooibroek, T. S. Carter, C. Williams, M. R. Wilson, M. P. Crump, A. P. Davis, *Chem. Sci.* **2017**, *8*, 4056-4061.
[29] D. J. Becker, J. B. Lowe, *Glycobiology* **2003**, *13*, 41R-53R.
[30] B. Ma, J. L. Simala-Grant, D. E. Taylor, *Glycobiology* **2006**, *16*, 158R-184R.
[31] C. Chemani, A. Imberty, S. de Bentzmann, M. Pierre, M. Wimmerová, B. P. Guevry, K. Faure, *Infect. Immun.* **2009**, *77*, 2065-2075.
[32] S. Pinho, C. A. Reis, *Nat. Rev. Cancer* **2015**, *15*, 540-555.
[33] A. Breiman, M. D. L. Robles, S. de Carné Trécesson, K. Echasserieau, K. Bernardeau, K. Drickamer, A. Imberty, S. Barillé-Nion, F. Altare, J. Le Pendu, *Oncotarget* **2016**, *7*, 14064-14082.
[34] S. Zhang, C. Cordon-Cardo, H. S. Zhang, V. E. Reuter, S. Adluri, W. B. Hamilton, K. O. Lloyd, P. O. Livingston, *Int. J. Cancer* **1997**, *73*, 42-49.
[35] O. Francesconi, A. Ienco, G. Moneti, C. Nativi, S. Roelens, *Angew. Chem. Int. Ed.* **2006**, *43*, 6693-6696.
[36] M. Cacciarini, C. Nativi, M. Norcini, S. Staderini, O. Francesconi, S. Roelens, *Org. Biomol. Chem.* **2011**, *9*, 1085-1091.
[37] O. Francesconi, M. Gentili, S. Roelens, *J. Org. Chem.* **2012**, *77*, 7548-7554.
[38] O. Francesconi, M. Gentili, C. Nativi, A. Ardá, F. J. Cañada, J. Jiménez-Barbero, S. Roelens, *Chem. Eur. J.* **2014**, *20*, 6081-6091.
[39] Unpublished results from this laboratory.
[40] A. Vacca, C. Nativi, M. Cacciarini, R. Pergoli, S. Roelens, *J. Am. Chem. Soc.* **2004**, *126*, 16456-16465.
[41] A. Vacca, O. Francesconi, S. Roelens, *Chem. Rec.* **2012**, *12*, 544-566.
[42] C. Nativi, M. Cacciarini, O. Francesconi, A. Vacca, G. Moneti, A. Ienco, S. Roelens, *J. Am. Chem. Soc.* **2007**, *129*, 4377-4385.
[43] S. Roelens, A. Vacca, C. Venturi, *Chem. Eur. J.* **2009**, *15*, 2635 –2644.
[44] S. Roelens, A. Vacca, O. Francesconi, C. Venturi, *Chem. Eur. J.* **2009**, *15*, 8296-8302.
[45] C. Nativi, O. Francesconi, G. Gabrielli, A. Vacca, S. Roelens, *Chem. Eur. J.* **2011**, *17*, 4814-4820.
[46] We are aware of only one report describing effective recognition of α-Galactosides; see: M. Rauschenberg, S. Bandaru, M. P. Waller, B. J. Ravoo, *Chem. Eur. J.* **2014**, *20*, 2770-2782.
[47] S. F. Poget, G. B. Legge, M. R. Proctor, P. J. G. Butler, M. Bycroft, R. L. J. Williams, *Mol. Biol.* **1999**, *290*, 867-879.
[48] C. P. Swaminathan, D. Gupta, V. Sharma, A. Surolia, *Biochemistry* **1997**, *36*, 13428-13434.
[49] C. Meynier, M. Feracci, M. Espeli, F. Chaspoul, P. Gallice, C. Schiff, F. Guerlesquin, P. Roche, *Biophys. J.* **2009**, *97*, 3168–3177.

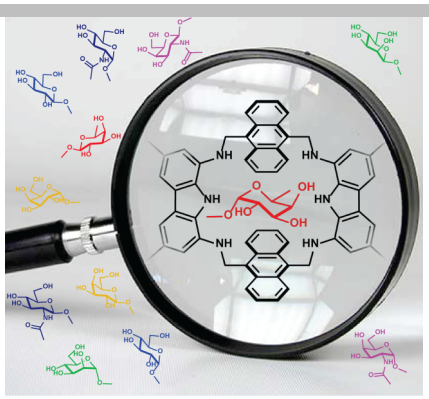
10.1002/chem.201800390

[50] O. Francesconi, C. Nativi, G. Gabrielli, M. Gentili, M. Palchetti, B. Bonora, S. Roelens, *Chem. Eur. J.* **2013**, *19*, 11742-11752.
[51] A. Ardá, C. Venturi, C. Nativi, O. Francesconi, F. J. Cañada, J. Jiménez-Barbero, S. Roelens, *Eur. J. Org. Chem.* **2010**, 64-71.
[52] A. Ardá, F. J. Cañada, C. Nativi, O. Francesconi, G. Gabrielli, A. Ienco, J. Jiménez-Barbero, S. Roelens, *Chem. Eur. J.* **2011**, *17*, 4821-4829.
[53] The assumption is supported by the negligible self-association observed for the corresponding receptor in which the anthracene moiety has been replaced by a benzene moiety (unpublished results from this laboratory).
[54] P. M. Antonik, A. N. Volkov, U. N. Broder, D. Lo Re, N. A. J. van Nuland, P. B. Crowley *Biochemistry* **2016**, *55*, 1195-1203.

10.1002/chem.201800390

FULL PAPER

Focus on fucose. A synthetic receptor build-up by assembling a hydrogen bonding element (carbazole), a hydrophobic aromatic moiety (anthracene), and a water-solubilizing function (phosphonate) into a macrocyclic structure providing the appropriate binding geometry selectively binds fucose with sub-millimolar affinity in water at physiological pH.



Oscar Francesconi, Marco Martinucci, Lorenzo Badii, Cristina Nativi, and Stefano Roelens*

Page No. – Page No.

A biomimetic synthetic receptor selectively recognizing fucose in water.