J|A|C|S
ARTICLES
Published on Web 03/16/2007

Pyrrolic Tripodal Receptors Effectively Recognizing Monosaccharides. Affinity Assessment through a Generalized Binding Descriptor

Cristina Nativi,[‡,£] Martina Cacciarini,[‡] Oscar Francesconi,[‡] Alberto Vacca,[†] Gloriano Moneti,[#] Andrea Ienco,[¶] and Stefano Roelens*[,§]

Contribution from the Dipartimento di Chimica Organica, Dipartimento di Chimica, Centro Risonanze Magnetiche (CERM), and Centro Interdipartimentale di Spettrometria di Massa (CISM), Università di Firenze, and Istituto di Chimica dei Composti OrganoMetallici (ICCOM) and Istituto di Metodologie Chimiche (IMC), Consiglio Nazionale delle Ricerche (CNR), Polo Scientifico e Tecnologico, I−50019 Sesto Fiorentino, Firenze, Italy

Received December 6, 2006; E-mail: stefano.roelens@unifi.it

Abstract: Pyrrolic and imino (**3**) or amino (**4**) H-bonding ligands were incorporated into a benzene-based tripodal scaffold to develop a new generation of receptors for molecular recognition of carbohydrates. Receptors **3** and **4** effectively bound a set of octylglycosides of biologically relevant monosaccharides, including glucose (Glc), galactose (Gal), mannose (Man), and *N*-acetyl-glucosamine (GlcNAc), showing micromolar affinities in $CDCl_3$ and millimolar affinities in CD_3CN by NMR titrations. Both receptors selectively recognized Glc among the investigated monosaccharides, with **3** generally less effective than **4** but showing selectivities for the all-equatorial β-glycosides of Glc and GlcNAc among the largest reported for H-bonding synthetic receptors. Selectivities in $CDCl_3$ spanned a range of nearly 250-fold for **3** and over 30-fold for **4**. Affinities and selectivities were univocally assessed through the BC_{50} descriptor, for which a generalized treatment is described that extends the scope of the descriptor to include any two-reagent host−guest system featuring any number of binding constants. ITC titrations of βGlc in acetonitrile evidenced, for both receptors, a strong enthalpic contribution to the binding interaction, suggesting multiple H bonding. Selectivity trends toward αGlc and βGlc analogous to those obtained in solution were also observed in the gas phase for **3** and **4** by collision-induced dissociation experiments. From comparison with appropriate reference compounds, a substantial contribution to carbohydrate binding emerged for both the imino/amino and the pyrrolic H-bonding groups but not for the amidic group. This previously undocumented behavior, supported by crystallographic evidence, has been discussed in terms of geometric, functional, and coordinative complementarity between H-bonding groups and glycosidic hydroxyls and opens the way to a new designer strategy of H-bonding receptors for carbohydrates.

Introduction

Molecular recognition of carbohydrates is crucial in many biological processes besides carbohydrate metabolism and transport, including cell-to-cell adhesion, cell infection by pathogens, immune response, and enzyme activity regulation.[1] All such processes rely on selective carbohydrate-to-carbohydrate and carbohydrate-to-protein interactions as a fundamental step. No wonder recognition of carbohydrates has been an actively investigated subject in bioorganic and supramolecular chemistry in the past decade.[2] Because of the complexity of natural processes, carbohydrate recognition has been investigated, focusing mainly on the monosaccharides or short oligosaccharides most frequently encountered as natural epitopes and largely exploiting synthetic receptors to mimic the interaction occurring with natural receptors and to understand the molecular basis of recognition events.[3]

Among the many questions still open, a crucial issue is the comprehension of the structural and functional requirements for

[‡] Dipartimento di Chimica Organica, Università di Firenze.
[†] Dipartimento di Chimica, Università di Firenze.
[£] CERM, Università di Firenze.
[#] CISM, Università di Firenze.
[¶] ICCOM, CNR.
[§] IMC, CNR.

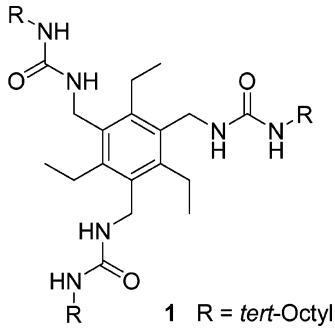
(1) (a) Ernst, B.; Hart, W.; Sinaÿ, P. *Carbohydrates in Chemistry and Biology*; Wiley-VCH: Weinheim, Germany, 2000; Part I, Vol. 2 and Part II, Vol. 4. (b) Lindhorst, T. K. *Essentials of Carbohydrate Chemistry and Biochemistry*; Wiley-VCH: Weinheim, Germany, 2000. (c) *Essentials of Glycobiology*; Varki, A., Cummings, R., Esko, J., Freeze, H., Hart, G., Marth, J., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 1999.

(2) (a) Lutzen, A. In *Highlights in Bioorganic Chemistry*; Schmuck, C., Wennemers, H., Eds.; Wiley-VCH: Weinheim, Germany, 2004; pp 109−119. (b) Striegler, S. *Curr. Org. Chem.* **2003**, *7*, 81−102. (c) Dam, T. K.; Brewer, C. F. *Chem. Rev.* **2002**, *102*, 387−429. (d) Mellet, C. O.; Defaye, J.; García Fernández, J. M. *Chem.−Eur. J.* **2002**, *8*, 1982−1990. (e) Host−Guest Chemistry. Mimetic Approaches to Study Carbohydrate Recognition. In *Topics in Current Chemistry*; Penadés, S., Ed.; Springer-Verlag: Heidelberg, Germany, 2002; Vol. 218. (f) Davis, A. P.; Wareham, R. S. *Angew. Chem., Int. Ed.* **1999**, *38*, 2978−2996. (g) James, T. D.; Sandanayake, K. R. A. S.; Shinkai, S. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 1910−1922. (h) Lee, Y.C.; Lee, R.T. *Acc. Chem. Res.* **1995**, 321−327.

(3) For a recent comprehensive review on the subject, see: Davis, A. P.; James, T. D.; In *Functional Synthetic Receptors*; Schrader, T., Hamilton, A. D., Eds.; Wiley-VCH: Weinheim, Germany, 2005; pp 45−109.

10.1021/ja068754m CCC: $37.00

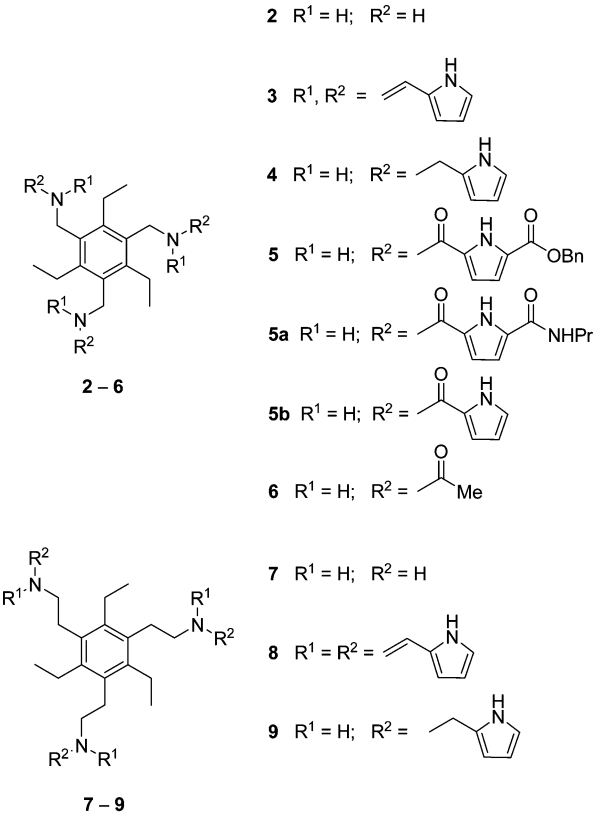
the effective and selective recognition of a specific saccharide. In nature, molecular recognition of carbohydrates relies on weak noncovalent interactions, mainly hydrogen bonding, and the required affinity and selectivity are achieved by multivalency[1,4] through a concerted array of H bonds involving multiple carbohydrate units. In synthetic receptors, such a combination is difficult to achieve by design, but by capitalizing on noncovalent interactions,[5] encouraging results have been obtained in several cases, mostly in organic solvents, even though high selectivity remains the most ambitious goal yet to be achieved.[3] Among the numerous artificial receptors reported to date, benzene-based tripodal structures, both macrocyclic and acyclic, have been successfully explored for the recognition of saccharides.[6] In this context, we have recently reported a new prototypical tripodal receptor for the recognition of monosaccharides, featuring a triethylbenzene scaffold bearing three convergent ureidic H-bonding units (**1**).[7] Binding affinities in the millimolar range and moderate selectivities were measured for **1** in $CDCl_3$ toward a set of representative α- and β-octylglycosides, selected among the most relevant to biological recognition processes. In order to improve on binding ability, we reasoned that more efficient receptors may be obtained by replacing ureidic groups with H-bonding ligands that could exhibit better complementarity with the carbohydrate hydroxy and ether functions but still preserving the flexibility of the tripodal architecture. The excellent results obtained



1 R = *tert*-Octyl

with a cage receptor featuring amino and pyrrole functions[8] showed that these groups can be conveniently employed as alternative H-bonding ligands. Indeed, amino and hydroxy groups have been shown to be complementary H-bonding partners, both geometrically and coordinatively, giving rise to molecular recognition and self-assembly.[9] Likewise, pyrroles, which are well-established H-bonding donors largely employed

Chart 1



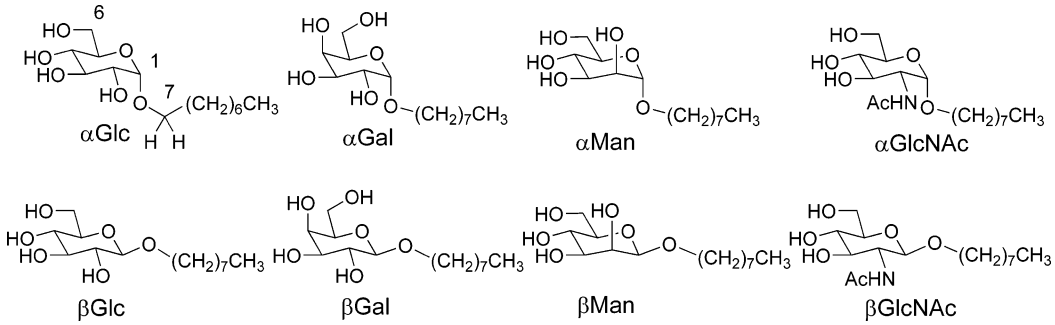
2 R¹ = H; R² = H
3 R¹, R² =
4 R¹ = H; R² =
5 R¹ = H; R² = OBn
5a R¹ = H; R² = NHPr
5b R¹ = H; R² =
6 R¹ = H; R² = Me
2 – 6
7 R¹ = H; R² = H
8 R¹ = R² =
9 R¹ = H; R² =
7 – 9

for anion binding,[10] appear to be, as of yet, essentially unexplored for the recognition of carbohydrates.[11] We thus thought that amino and pyrrolic binding groups could be conveniently assembled on the 1,3,5-triethylbenzene scaffold to afford adaptive receptors of significantly improved recognition properties toward monosaccharides. Following a systematic investigation, we now wish to describe a new type of tripodal receptor, which exhibits some of the best binding affinities toward biologically relevant monosaccharides reported in the literature for neutral synthetic receptors, shows a unique selectivity for the glycosides of β-glucose and *N*-acetyl-β-glucosamine, and opens the way to a new generation of acyclic hosts for recognition of carbohydrates solely based on H-bonding interactions.

A fundamental issue for the assessment of binding properties was the evaluation of affinities on a common scale because the investigated receptor−glycoside systems fit different binding models depending on the glycoside. To address this issue, we also wish to report a generalized treatment we developed to extend the scope of the BC_{50} descriptor, which we previously proposed as a comparative index of binding efficacy,[7] to include systems fitting chemical models of any number of equilibria. In the generalized formulation, the BC_{50} parameter can, thus, be used to assess affinities on the same scale for any two-reagent

(4) (a) Kiessling, L.; Pontrello, L.; Schuster, M. C. In *Carbohydrate-based Drug Discovery*; Wong, C. H., Ed.; Wiley−VCH: Weinheim, Germany, 2003; Vol. 2, pp 575−608. (b) Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, 555−578. (c) Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998**, *37*, 2754−2794.
(5) Although not biomimetic, formation of covalent B−O bonds has also been largely exploited for the recognition of carbohydrates. See ref 3.
(6) For some representative examples from the recent literature, see: (a) Abe, H.; Aoyagi, Y.; Inouye, M. *Org. Lett.* **2005**, *7*, 59−61. (b) Schmuck, C.; Schwegmann, M. *Org. Lett.* **2005**, *7*, 3517−3520. (c) Mazik, M.; Radunz, W.; Boese, R. *J. Org. Chem.* **2004**, *69*, 7448−7462. (d) Welti, R.; Diederich, F. *Helv. Chim. Acta* **2003**, *86*, 494−503. (e) Kim, Y-H.; Hong, J-I. *Angew. Chem., Int. Ed.* **2002**, *41*, 2947−2950. (f) Zhong, Z.; Anslyn, E. V. *J. Am. Chem. Soc.* **2002**, *124*, 9014−9015.
(7) Vacca, A.; Nativi, C.; Cacciarini, M.; Pergoli, R.; Roelens, S. *J. Am. Chem. Soc.* **2004**, *126*, 16456−16465.
(8) Francesconi, O.; Ienco, A.; Moneti, G.; Nativi, C.; Roelens, S. *Angew. Chem., Int. Ed.* **2006**, *45*, 6693−6696.
(9) (a) Hanessian, S.; Simard, M.; Roelens, S. *J. Am. Chem. Soc.* **1995**, *117*, 7630−7645. (b) Ermer, O.; Eling, A. *J. Chem. Soc., Perkin Trans. 2* **1994**, 925−944.
(10) *Anion Receptor Chemistry*; Sessler, J. L., Gale, P. A., Cho, W.-S., Eds.; The Royal Society of Chemistry: Cambridge, U.K., 2006.
(11) Apart from a report on anionic carbohydrate recognition by a cationic receptor through electrostatic interactions in water (ref 6b), we are aware of only one example of pyrrole-containing receptors for sensing monosaccharides. See: Fang, J.-M.; Selvi, S.; Liao, J.-H.; Slanina, Z.; Chen, C.-T.; Chou, P.-T. *J. Am. Chem. Soc.* **2004**, *126*, 3559−3566.

Chart 2



host−guest systems, irrespective of the nature of the binding partners. Making use of the generalized BC_{50} parameter, the affinities of the receptors for the investigated set of glycosides have been univocally assessed and may be compared with those calculated for any other receptor.

Results and Discussion

Design and Synthesis. The design of the new pyrrolic tripodal receptors was based on the same 1,3,5-substituted 2,4,6-triethylbenzene scaffold[12] of the parent ureidic host **1** on account of its marked preference for the alternate substituents pattern, directing the three binding arms toward the same side of the aromatic ring, which was shown to provide the correct geometry for binding monosaccharides.[7] Molecular mechanics calculations indicated that, for replacing the ureidic group with an amino−pyrrole motif, an appropriate assemblage to achieve the correct geometry for binding would require the amine and the pyrrole groups to be spaced by one methylene unit, with the spacer connected to the 2 position of the pyrrole ring. Such an arrangement could be conveniently realized by condensation of the parent triamine **2** with pyrrole-2-carboxaldehyde through the formation of the corresponding Schiff base **3** (Chart 1). Reduction to the amine **4** would provide an amino−pyrrole receptor endowed with the functional, structural, and flexibility features appropriate for H bonding to monosaccharides. Although structurally similar to **4**, the iminic receptor **3** may exhibit significantly different binding properties because the conformational constraint imposed by the iminic double bond, coplanar with the pyrrolic ring due to conjugation, may lock the conformation into a chelating arrangement of the two nitrogen atoms. Since amidic groups are largely employed for H bonding in carbohydrate recognition,[3] the structurally related amido−pyrrole **5** was prepared by standard peptide chemistry from pyrrole-2,5-dicarboxylic acid monobenzylester,[13] with the aim of assessing whether the aminic group would be more effective than the amidic group in binding carbohydrates. The benzylester moiety in the 5 position of pyrrole was necessary to overcome the poor solubility we observed with the corresponding 2,5-diamidic (**5a**) and 5-unsubstituted (**5b**) amidic receptors, which prevented an evaluation of their binding properties. To ascertain the contribution from the pyrrolic groups to the recognition of monosaccharides, the plain triacetamide **6**, to be compared to the amidic receptor **5**, was prepared by acetylation of the triamine **2**; likewise, the plain triamine **2** was used as a reference for the amino−pyrrolic receptor **4**. Eventually, receptors **8** and **9** were prepared from the homologue triamine **7** by condensation with pyrrole-2-carboxaldehyde to elucidate the effect of the spacer length on the binding capabilities of the imino− and amino−pyrrolic receptors and to validate predictions from molecular modeling calculations.

Binding Studies. The binding affinities toward the set of octylglycosides of biologically relevant monosaccharides depicted in Chart 2 were measured by ^{1}H NMR titrations in $CDCl_3$ at $T = 298$ K. Glc, Gal, Man, and GlcNAc were selected among the most frequently encountered monosaccharidic epitopes, present as terminal α- or β-glycosides in more complex oligo- or polysaccharides on cell surfaces and in glycoconjugates.

An in-depth investigation of all of the aspects involved in the correct determination of binding affinities toward glycosides by NMR titrations (reactant self-association, chemical model, choice of signals, etc.) has been described in a previous paper.[7] In the present study, we closely followed the established titration protocol and the optimized analysis of data, which consisted of diluting the receptor with a stock solution of the glycoside and simultaneously fitting of all the available signals of both reagents to the appropriate model of chemical equilibria, respectively. Unfortunately, the hydroxylic proton signals, usually experiencing large shifts due to H bonding, could not be followed in the titration experiments because they all collapsed into a broad average signal, together with the water signal and with the aminic protons for receptor **4**. For the iminic receptor **3**, interestingly, with a lack of aminic protons, the signals involved in fast exchange averaging with the hydroxylic protons were the pyrrolic NH protons. In all cases, however, complexation-induced shifts could be accurately monitored for several CH signals of both binding partners, some of which exhibited remarkably large values (see Supporting Information). The titration of βGlcNAc with **3** is reported in Figure 1 as a typical example, showing the excellent agreement obtained between the experimental and the calculated shifts for all of the detected signals. In general, binding of the receptors to the investigated glycosides fit a model including 1:1 and 2:1 host-to-guest association equilibria, in addition to the dimerization of the receptor. For the latter self-association equilibrium, the corresponding constants were measured independently and used as invariant parameters in the nonlinear least-squares regression analysis.[14] The results obtained for receptors **3** and **4** with the set of glycosides of Chart 2 are reported in Table 1 as cumulative log β values for the formation of the 1:1 and 2:1 receptor-to-

(12) Tripodal receptors based on hexasubstituted benzene rings have become quite popular in the last decade. For a recent review, see: Hennrich, G.; Anslyn, E. V. *Chem.—Eur. J.* **2002**, *8*, 2218−2224.

(13) Schmuck, C.; Schwegmann, M. *J. Am. Chem. Soc.* **2005**, *127*, 3373−3379.

(14) Measured self-association constants, log β_{dim}, in $CDCl_3$ at $T = 298$ K for **1** and the receptors of Chart 1: 1.732 ± 0.009 (**1**); 1.83 ± 0.02 (**2**); 0.92 ± 0.02 (**3**); 1.07 ± 0.01 (**4**); n.d. (**5**); n.d. (**6**).

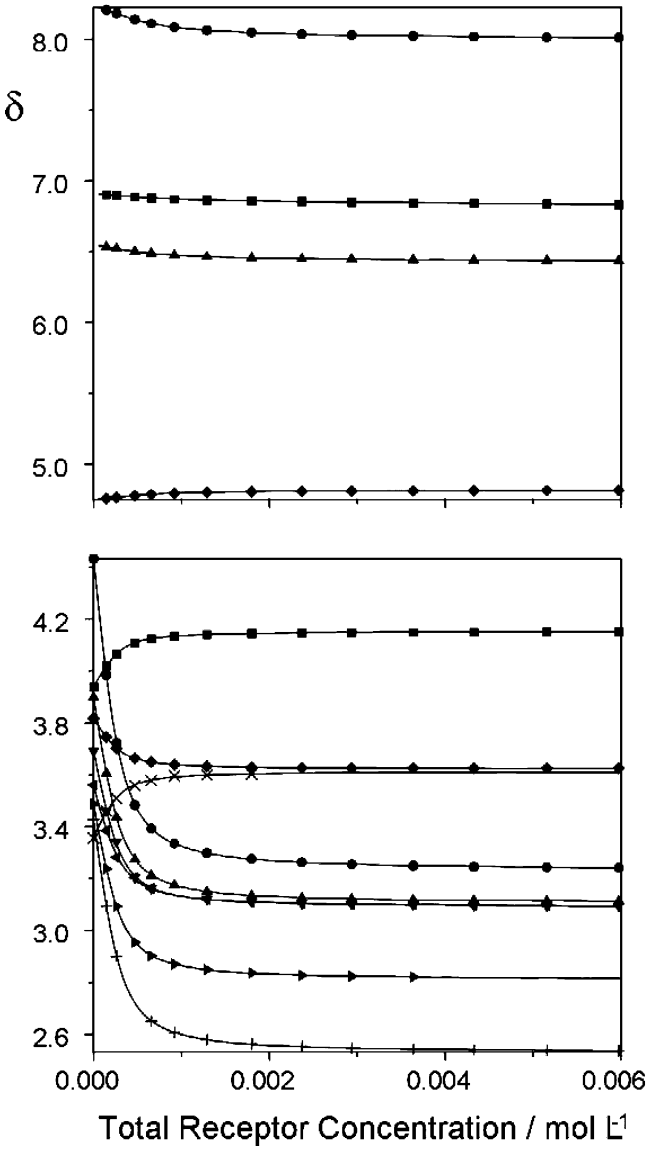


Figure 1. Plot of complexation induced shifts versus the receptor's total concentration in the titration of βGlcNAc (0.27 mM) with **3** in $CDCl_3$ at 400 MHz and $T = 298$ K. Bottom: glycoside CH signals (● CH-**1**; × CH-**2**; ▼ CH-**3**; left arrow CH-**4**; + CH-**5**; ■ CH-**6**; ♦ CH-**6′**; ▲ CH-**7**; right arrow CH-**7′**). Top: receptor CH signals (● CHN-im; ■ CH-pyrr; ▲ CH-pyrr; ♦ CH_2N). For glycoside proton numbering, see Chart 2; for proton assignment, see 2D-NMR spectra in Supporting Information. Symbols are experimental data points; solid lines are best-fit curves obtained through nonlinear regression by simultaneous fit of all data.

glycoside adducts. Corresponding data for **1**[7] are also included for direct comparison. While for αGlcNAc no evidence of the 2:1 complex species was found, in several cases, the concentration of the latter was too low to allow a correct determination of the corresponding constant. In both instances, the system was correctly described with good agreement by the 1:1 association constant. From the inspection of Table 1, particularly large values can be noted for βGlc with receptors **3** and **4**. Since these values were obtained in a noncompetitive solvent, binding constants were also measured in CD_3CN to ascertain whether recognition would still occur in a polar medium. A 1:1 association with βGlc could, indeed, be detected, with log β values of 2.169 ± 0.004 ($K_a = 147 \pm 1.2$ M^{-1}) and 1.721 ± 0.002 ($K_a = 52.5 \pm 0.3$ M^{-1}) for **3** and **4**, respectively. An independent support was obtained by isothermal titration calorimetry (ITC), whose results are reported in Table 2. Besides the very good agreement between the association constants obtained by the ITC and NMR techniques, thermodynamic parameters evidenced a strong enthalpic contribution to the association, which resulted in much smaller binding free-energy values because of the large and adverse entropic contribution. It should be mentioned that receptor **4** was also tested with fully acetylated βGlc, giving no evidence of binding in $CDCl_3$ and clearly showing that the strong binding observed with βGlc is fully depleted in the absence of free hydroxyl groups. Recognition of glycosides by the tripodal pyrrolic receptors can, thus, be ascribed to a strong enthalpic interaction, likely resulting from multiple H bonding.

As the largest constants were observed for βGlc, it was used as a test ligand to evaluate the affinities of the reference compounds **2** and **5**−**9**. The parent triamine **2** showed a 1:1 association equilibrium, with a log β value of 2.616 ± 0.004, together with a significant self-association constant.[14] The trisamidopyrrole **5** showed 1:1 and 2:1 association equilibria with a negligible self-association constant; the corresponding log β values were 2.22 ± 0.01 and 3.59 ± 0.03, respectively. To ascertain any α/β selectivity, binding constants were also measured toward αGlc, giving log β values for the 1:1 and 2:1 adducts of 2.02 ± 0.02 and 3.24 ± 0.06, respectively. In the case of the triacetamido derivative **6**, only a 1:1 association could be revealed for βGlc, with a log β value of 1.21 ± 0.01 and undetectable self-association. Finally, attempts to evaluate the

Table 1. Cumulative Association Constants (log β_n) with Standard Deviations (σ) for 1:1 (Upper Value) and 2:1 (Lower Value) Complexes of Receptors **1**, **3**, and **4** with Octylglycosides[a]

glycoside	**1**[b]	**3**	**4**
αGlc	2.77 ± 0.01	3.573 ± 0.004	3.23 ± 0.01
	4.80 ± 0.02	*c*	5.22 ± 0.04
βGlc[d]	2.67 ± 0.04	5.30 ± 0.05	4.61 ± 0.03
	4.88 ± 0.06	9.04 ± 0.09	7.79 ± 0.06
αGal	2.82 ± 0.01	3.437 ± 0.002	3.10 ± 0.01
	4.82 ± 0.02	*c*	4.75 ± 0.10
βGal	2.92 ± 0.01	3.921 ± 0.004	4.153 ± 0.008
	5.01 ± 0.03	*c*	5.59 ± 0.12
αMan	2.82 ± 0.05	3.583 ± 0.006	4.37 ± 0.01
	4.60 ± 0.11	*c*	5.75 ± 0.08
βMan	2.31 ± 0.18	3.185 ± 0.009	4.43 ± 0.02
	3.98 ± 0.29	*c*	6.45 ± 0.09
αGlcNAc	2.75 ± 0.04	2.937 ± 0.001	4.14 ± 0.04
	4.31 ± 0.13	n.d.[e]	6.05 ± 0.47
βGlcNAc[f]	2.62 ± 0.03	4.49 ± 0.04	4.72 ± 0.04
	4.50 ± 0.06	7.95 ± 0.07	8.25 ± 0.06

[a] Measured by ^{1}H NMR (400 MHz) from titration experiments at $T = 298$ K in $CDCl_3$ on 0.8−1.2 mM stock solutions of glycoside using receptor concentrations up to 20 mM. Binding constants β_{11} (M^{-1}) and β_{21} (M^{-2}) were calculated by simultaneous fit of the shifts of all of the available signals. The receptor's dimerization constant was measured independently under the same conditions and set invariant in the nonlinear regression analysis. For log β_{dim} values, see ref 14. [b] Data from ref 7. [c] Too low to be correctly determined. [d] Determined at 900 MHz. [e] Nondetectable. [f] Stock solution, 0.27 mM.

Table 2. Association Constants K_a (M^{-1}) and Thermodynamic Parameters (kcal mol^{-1}) for 1:1 Association of Receptors **3** and **4** with βGlc[a]

receptor	K_a	$-\Delta G°$	$-\Delta H°$	$-T\Delta S°$
3	129.0 ± 1.6	2.87 ± 0.01	11.3 ± 0.8	8.4
4	87.4 ± 1.7	2.65 ± 0.01	6.0 ± 0.8	3.4

[a] Measured by ITC from titration experiments at $T = 298$ K in CH_3CN in 1.0 mM solutions of receptor-injecting 25−50 mM solutions of glycoside.

affinities of **8** and **9** gave no evidence of binding toward βGlc and a clear-cut answer to the question about the choice of the spacer length, in agreement with predictions based on molecular modeling calculations.

Although a whole set of binding constants was obtained for the recognition of glycosides, the results reported in Table 1 cannot be directly used for assessing the binding ability of the tripodal pyrrolic receptors for several reasons: (a) each constant of the two complexation steps alone does not describe the overall binding ability; (b) comparison of binding constants of different orders is unfeasible; and (c) the investigated systems are not homogeneously characterized by the same model. To assign an univocal value to the affinity of each receptor for each glycoside on a common scale, a general descriptor of binding affinity is necessarily required.

Binding Descriptors. In a previous work,[7] the affinity of a reagent A for a reagent B has been assessed through the BC_{50} parameter, which we proposed as a descriptor of binding affinity. In analogy to the IC_{50} parameter widely employed in biochemistry, the median binding concentration, BC_{50}, was defined as the total concentration of A (T_A, in mol L^{-1}) necessary for binding 50% of B, that is, for which

$$T_B^{\text{bound}} = T_B/2 \tag{1}$$

or, equivalently

$$x_B = T_B^{\text{bound}}/T_B = 0.5 \tag{2}$$

where x_B is the fraction of bound B. Thus, like for IC_{50}, the higher the affinity, the lower the BC_{50} value. BC_{50} becomes essential to assess affinities for systems featuring more than one association equilibrium, in which the binding ability of a species is not directly defined by a single binding constant. In such cases, BC_{50}, which can be calculated from the knowledge of the set of binding constants, has been shown to depend on all of the constants involved and thus takes into account all of the complex species.[7]

BC_{50} is a conditional parameter that depends on the total concentration of the analyte, T_B, at which it is calculated (eq 1), which must be, therefore, specified together with the temperature.[15] A very useful property of BC_{50} is that, when the total concentration of the analyte T_B becomes negligible, the value of BC_{50} becomes constant, that is,

$$\lim_{T_B \to 0} BC_{50} = \text{const} = BC_{50}^0 \tag{3}$$

In contrast to BC_{50}, BC_{50}^0, which we called intrinsic median binding concentration, depends on all of the equilibrium constants involved in the system but not on specific (concentration) conditions.

The BC_{50} descriptor is very useful when comparing binding abilities for different host−guest systems. For example, the affinity of two or more receptors for a ligand or the selectivity of a receptor toward two or more ligands can be directly assessed by comparing the corresponding BC_{50} values, whatever the number of equilibria involved. We have successfully assessed, this way, the affinity and the selectivity of **1** toward the glycosides of Chart 1, all sharing a common three-constant model.[7]

In order to be a general descriptor of binding affinity, BC_{50} must fulfill two basic requirements, namely, (1) it must be a function of variables directly related to binding affinity, and (2) to allow comparison of different systems, it must be calculated under the same binding conditions. For systems fitting the same chemical model, parity of conditions, besides solvent and temperature, is ensured by calculating BC_{50} at the same T_B value. Unfortunately, this is no longer true for systems fitting different models, as T_B appears in the BC_{50} expression with a different coefficient (thus having a different weight) for complexes of different stoichiometry.[16] For a generalized validity of the descriptor, a more convenient quantity to refer to is the fraction of bound A (eq 4), which identifies the extent of saturation of the binding reagent

$$x_A = T_A^{\text{bound}}/T_A \tag{4}$$

BC_{50} is most usefully expressed as a function of this quantity because x_A has two main advantages over T_B or the extent of binding T_A^{bound}/T_B. First, in contrast to T_B, it does not appear in the BC_{50} expression with coefficients dependent on the complex's stoichiometry, as long as complex species multinuclear in B are absent (see Appendix in the Supporting Information). Second, in contrast to T_A^{bound}/T_B, it ranges from 0 to 1 for all systems, irrespective of the model and the stoichiometry of the complexes formed. These properties ensure the univocal definition of parity of conditions, that is, the same extent of saturation of reagent A for systems fitting different models and stoichiometries. Under these conditions, BC_{50} describes, on a common scale, the binding ability of reagents when they have the same tendency to form complexes. Indeed, from the expressions obtained for different models (see Appendix in the Supporting Information), it can be seen that BC_{50} depends solely on x_A and on all binding constants and, therefore, for a specific x_A value, exclusively on affinity, making BC_{50} a binding descriptor of general scope. Thus, comparing the binding abilities of reagents for two host−guest systems fitting different chemical models translates into comparing BC_{50} values calculated at the same fraction of bound reagent x_A, that is, calculating the total concentration of A that binds 50% of B when A is saturated to the same extent for both systems.

Analogous to BC_{50} as a function of T_B, when x_A becomes negligible, the value of BC_{50} becomes constant, that is,

$$\lim_{x_A \to 0} BC_{50} = \text{const} = BC_{50}^0 \tag{5}$$

For a two-reagent system, BC_{50}^0 as a function of T_B and x_A coincide because, when the concentration of B becomes negligible, the fraction of bound A becomes negligible as well. However, BC_{50}^0 as a function of x_A assumes the meaning of the affinity of reagent A for reagent B in the absence of formed complex species, that is, when forming the first complex molecule. Thus, BC_{50}^0 identifies the maximum binding ability of A toward B since, with increasing x_A, A becomes

(15) We name the species (B) toward which the binding affinity is being evaluated the "analyte".

(16) This feature can be appreciated, for example, from the BC_{50} expressions relative to complexes of the type AB, A_2B, and A_3B, respectively. AB: $BC_{50} = \beta_{11}^{-1} + 0.5T_B$. A_2B: $BC_{50} = \beta_{21}^{-1/2} + 1T_B$. A_3B: $BC_{50} = \beta_{31}^{-1/3} + 1.5T_B$.

progressively saturated. When comparing, for example, the affinities of two reagents A and A′ (e.g., two hosts H and H′) for the same ligand (G), while the selectivity ratio $BC_{50}(H)/BC_{50}(H')$ will remain constant in the entire x_H range and equal to the $BC^0_{50}(H)/BC^0_{50}(H')$ ratio, the actual BC_{50} values will increase with increasing x_H from BC^0_{50} to saturation, reflecting the decreasing binding ability of the receptors with the increasing extent of saturation (see Supporting Information). We believe that the BC_{50} and BC^0_{50} descriptors represent the most useful and most general indicators of binding affinity.[17]

It should be emphasized that, although not necessary, BC_{50} and BC^0_{50} can be appropriately employed in cases where a 1:1 association is the only equilibrium involved, allowing for a direct comparison of results with those from more complex systems and, thus, heterogeneous in nature. In this case, it has been demonstrated[7] that BC^0_{50} coincides with the dissociation constant $K_d = 1/K_a$, also known as the affinity constant, where K_a is the association constant of the 1:1 binding equilibrium, making apparent the chemical meaning of the intrinsic median binding concentration, which can be viewed as a "global" affinity constant.[18]

Glycoside Binding Affinity and Selectivity. Using the affinity descriptors, we are now able to quantitatively evaluate the binding ability of the pyrrolic tripodal receptors, which can be most appropriately discussed in terms of intrinsic median binding concentrations. The BC^0_{50} values calculated for **3** and **4** from β_{11}, β_{21}, and β_{dim} of Table 1 are reported in Table 3 together with the corresponding values calculated for **1** from previously reported data[7] for direct comparison. In contrast to **1**, whose affinities are in the millimolar range, BC^0_{50} values for **3** and **4** lie in the micromolar range, demonstrating a dramatically improved binding ability with respect to the ureidic progenitor. Besides the remarkable figure of 4.8 μM shown by **3** for βGlc, which improves on the affinity of **1** by over 400-fold, in general, the observed affinities compete favorably with data reported in the chemical literature, establishing the described tripodal hosts as a new generation of highly effective receptors for monosaccharides.

A peculiar feature emerging from Table 3 is that the aminic receptor **4** is generally more effective than the iminic receptor **3**, whereas the latter is distinctly more selective than the former. Indeed, both are selective for βGlc, but selectivity spans a range of over 30-fold for **4** and nearly 250-fold for **3**. Except for αGlc and αGal, for which lower affinities are observed, all glycosides are strongly bound to **4**; on the contrary, βGlc is preferred by **3** by orders of magnitude with respect to the other monosaccharides.

Table 3. Intrinsic Median Binding Concentration BC^0_{50} (μM) with Standard Deviation (σ) for Complexes of Receptors **1**, **3**, and **4** with Octylglycosides in $CDCl_3$ at $T = 298$ K[a]

	BC^0_{50} (σ)		
glycoside	**1**[b]	**3**	**4**
αGlc	1720(40)	268(2)	570(20)
βGlc	1970(90)	4.8(5)	24(2)
		6780(50)[c]	19000(1000)[c]
		7750(90)[d]	11500 (300)[d]
αGal	1520(30)	368(1)	790(20)
βGal	1190(30)	120(1)	70(1)
αMan	1600(200)	262(4)	43(1)
βMan	6000(200)	660(10)	37(1)
αGlcNAc	2000(200)	1179(3)	72(7)
βGlcNAc	2600(200)	30(2)	18(1)

[a] Calculated from the log β values reported in Table 1. [b] Calculated from log β values from ref 7. [c] Calculated from the log β values measured by NMR in CD_3CN. [d] Calculated from the log β values measured by ITC in CH_3CN.

Table 4. Intrinsic Median Binding Concentration BC^0_{50} (μM) with Standard Deviation (σ) for Complexes of Receptors **2**, **5**, and **6** with Octylglycosides in $CDCl_3$ at $T = 298$ K[a]

receptor	glycoside	BC^0_{50} (σ)
2	βGlc	3690(50)
5	αGlc	8300(300)
5	βGlc	5400(100)
6	βGlc	62000(1000)

[a] Calculated from the log β values reported in text.

To our knowledge, the selectivities shown by **3** for βGlc, not only versus the α-anomer but also versus the other monosaccharides, appear to be among the largest reported for neutral synthetic receptors. The fact that βGlcNAc, which, like βGlc, possesses all equatorial substituents, is bound only 6-fold less effectively than βGlc indicates that the correct complementarity is achieved for equatorial H-bonding groups. An analogous conclusion can be drawn for **4**, for which βGlcNAc is bound even more strongly than βGlc. In contrast, axial hydroxyl groups seem to feature a mismatched binding geometry, affecting **3** distinctly more than **4** and showing that geometric and coordinative requirements are significantly more strict for the former than for the latter. It can be concluded that the pyrrolic tripodal architecture is well suited to preferentially bind to the all-equatorial conformation of glucose and glucosamine, while conformational restrictions imposed by the imine double bonds of **3** significantly improve selectivity with respect to the aminic receptor **4**, as a result of a reduced flexibility.

Comparison with the reference compounds highlights the contributions to binding of the pyrrolic and the imino/amino groups. The BC^0_{50} values calculated for **2**, **5**, and **6** from the measured constants are reported in Table 4. Although the triaminic receptor **2** is indeed capable of binding βGlc with millimolar affinity, the improvement brought by the pyrrolic binding sites is 770-fold for **3** and 150-fold for **4**. It is also worth noting that the affinity of **2** for βGlc is only 2-fold lower than that of the triureidic receptor **1**. It is evident that the contribution of the pyrrolic substituents is substantial, whereas converting the triaminic receptor to the corresponding triureidic

(17) Care must be paid to the presence of complex species multinuclear in the analyte B. From BC_{50} equations (see Appendix in the Supporting Information), it is apparent that, in contrast to all other species, they contribute to the descriptor with a coefficient that depends on the stoichiometry of the species. When comparing systems fitting different models and featuring complex species multinuclear in B of different stoichiometry, the contribution of the latter species to BC_{50} will not be equally scaled unless weighted for the corresponding coefficients. It must be stressed that this represents a matter of concern only when complex species multinuclear in the analyte B of different stoichiometry are present, for example, when comparing binding abilities of two reagents A and A′ toward a common reagent B and species AB_2 and $A'B_3$ are formed. In these cases, it is easily seen that the contributions of these species to BC^0_{50} become indefinite, while those to BC_{50} will not be equally weighted (see Appendix in the Supporting Information). Since, most frequently, multinuclear complex species, if any, appear for only one of the two reagents, typically the titrating agent, an alternative approach in such cases would consist of calculating the BC_{50} value toward the reagent lacking multinuclear complex species.

(18) To expedite the calculation of the descriptors from the values of the binding constants, the computer program "BC_{50} Calculator" has been developed and made available for free upon request. The equations used by the program to calculate BC_{50} values are described in the Appendix in the Supporting Information.

host **1** brings little advantage. Somewhat surprisingly, affinities even lower than that for **1** were detected for the amido−pyrrolic receptor **5**, which also shows essentially no β/α selectivity toward Glc, despite the presence of pyrrolic binding sites. A clue to this unexpected result is provided by the triamidic receptor **6**, for which a drop of an order of magnitude in the affinity for βGlc is observed with respect to **5** and is even more pronounced with respect to **2**, showing that the decreased affinity can be ascribed to the amidic function, whereas pyrrole and amino groups are both effective in binding the glycoside. It thus emerges that aminic and pyrrolic binding groups are far better H-bonding ligands than amidic groups, which are largely employed in synthetic receptors for the recognition of carbohydrates. We believe that the reason for this ability resides in the good donor/acceptor H-bonding complementarity of amine and pyrrole with glycosidic hydroxyls, whereas that of amide groups is poorer than commonly thought.[19]

One comment is due to the results obtained in acetonitrile (Table 3). Although a drop of 3 orders of magnitude in the binding affinity for βGlc was observed with respect to chloroform, it is quite remarkable that an affinity at the millimolar level could still be detected, testifying that the association is strong enough to survive in pure acetonitrile. Calorimetric data confirmed that the interaction of the pyrrolic receptors with βGlc results in a distinct association event, enthalpic in origin, which survives even in a polar solvent. Eventually, it should be underlined that contribution of the spacer length to the binding geometry appears unexpectedly critical; binding was completely depleted by chain elongation of a single methylene group, even though the receptor is nonpreorganized and binding relies on adaptivity to the glycosidic guest.

X-ray Studies. The X-ray structures of receptors **3** and **5b** supported the above conclusions, providing an insight into the origin of the observed binding features. The ORTEP projections of the structure of **3** crystallized from $CHCl_3$/EtOH, depicted in Figure 2, show the expected alternate arrangement of substituents, with the three pyrrolic arms on the same side of the aromatic ring forming a cleft, in the center of which is a captured ethanol molecule. As anticipated, the iminic and the pyrrolic nitrogen atoms lie coplanar in all of the three side chains because of the conjugation of the iminic double bond with the pyrrole ring; rotation about the CH_2−NH single bond brings one of the three arms to converge toward the inside of the cleft and to chelate the alcoholic hydroxyl with the two nitrogen atoms. The H-bonding chelate arrangement is noteworthy not only for the nearly perfect planar geometry of the assembly but also for the matched complementarity of the involved functional groups, with the hydroxyl accepting one H bond from the pyrrole NH and donating one H bond to the imine nitrogen; this way, the chelating donor/acceptor diad of the receptor perfectly matches the dual donor/acceptor nature of the hydroxyl group.

Quite remarkably, a nearly identical structure was observed from crystals obtained from $CHCl_3$/MeOH, indicating that the H-bonded chelate with a hydroxylic species included in the cleft represents a structural preference for the pyrrolic tripodal receptor. Unfortunately, crystals suitable for X-ray structure analysis could not be obtained for any of the adducts with the

(19) These results suggest that the binding of anionic carbohydrates reported for the cationic tripodal receptor featuring amido−pyrrolic binding sites (ref 6b) may be substantially ascribed to electrostatic interactions rather than to H bonding.

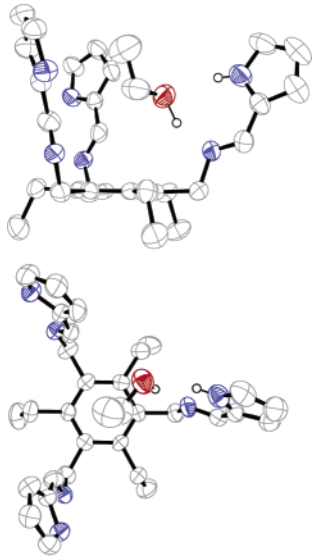


Figure 2. ORTEP projections of the X-ray structure of **3**·EtOH. Left: side view; right: top view. Ellipsoids are at 50% probability. Nitrogen and oxygen atoms are represented as shaded ellipsoids. Hydrogen atoms are omitted for clarity, except for those involved in H bonding. Selected distances and angles: N−H···O, 2.10 Å (161.6°); N(H)···O, 2.93 Å; O−H···N, 1.76 Å (153.6°); O(H)···N, 2.80 Å.

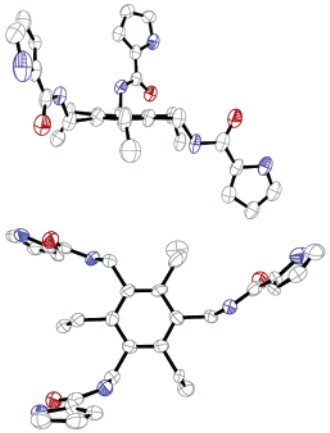


Figure 3. ORTEP projections of the X-ray structure of **5b**. Left: side view; right: top view. Ellipsoids are at 50% probability. Nitrogen and oxygen atoms are represented as shaded ellipsoids. Hydrogen atoms and a water molecule of crystallization are omitted for clarity.

investigated glycosides; however, it is plausible that, in the presence of glycosidic, guests of the appropriate size and possessing appropriately located hydroxyl groups, all three of the pyrrolic side chains may converge to cooperatively engage more than one H bond, giving rise to a reinforced enthalpic interaction and enhanced selectivities. In contrast to **3**, a refined X-ray structure could not be obtained for the aminic receptor **4**. Preliminary diffraction data indicated that disorder in the crystal structure prevented the refinement of data, most likely caused by a much floppier configuration of the receptor and the absence of solvent molecules bound in the cleft, although the alternate substituent pattern appeared to be conserved. A refined structure could instead be obtained for the amidic receptor **5b**, which resulted to be quite informative about the binding properties of the amidic receptor family (Figure 3). The main feature is the unexpected lack of an alternate arrangement of substituents and thus of a binding cleft; only two pyrrolic arms lie on one side of the aromatic ring, the third being located

on the other side together with the three ethyl groups. This evidence demonstrates that the alternate pattern of substituents is not a general feature of hexasubstituted benzenes. Of course, in solution, all conformations are expected to be present, but crystallographic evidence indicates that the alternate distribution may not be the preferred conformation.[20] A further feature is relevant to binding properties; as expected, the amidic group is coplanar with the pyrrole ring due to conjugation, but in contrast to the iminic receptor, the pyrrolic NH lies on the same side of the carbonyl and opposite the amidic NH in a preferred anti configuration. Thus, if the latter would point toward the inside of the virtual binding cleft, the former would point toward the outside of it and would be, therefore, unable to bind. The alternative binding mode, with the pyrrole NH and the carbonyl as chelating elements, does not meet the geometrical requirements for binding, being located one bond further from the appropriate geometry; moreover, the arrangement of the chelating groups would be energetically unfavorable when converging toward the inside of the cleft for binding, and because of the repulsive interactions, it would induce between pairs of carbonyls pointing toward each other. The latter factor may also be responsible for the lack of an alternate conformation in the solid state in order to minimize carbonyl group repulsion. Finally, in the crystal packing, two contiguous molecules are H bonded through a pair of short contacts between the pyrrolic NH of one arm and the amidic CO of the corresponding other molecule's arm in a head-to-tail fashion. Since **5b** was crystallized from the same alcoholic solvent used for **3**, this evidence shows that amidic groups preferentially bind to pyrrolic NH rather than to solvent hydroxyl groups and, most likely, to carbohydrate hydroxyls. Altogether, X-ray studies showed that, at least in the tripodal scaffold, the amidic group performs poorly as a H-bonding ligand for carbohydrates for structural, geometrical, and functional complementarity reasons.

Mass Studies. Independent evidence of the recognition properties of the pyrrolic tripodal receptors, in agreement with binding studies in solution, was obtained in the gas phase from mass experiments. In the positive ion mode ESI−MS spectrum of an equimolar mixture of **3** and βGlc, the [**3**·βGlc+H]$^+$ complex was present as the major peak, after the base peak of the free receptor, together with a peak of smaller intensity for dimeric **3** (Figure 4, bottom). An analogous spectrum was obtained by injecting an equimolar mixture of **3** and αGlc, showing the same set of peaks in comparable intensities, with the [**3**·αGlc+H]$^+$ complex present as the minor of the three peaks. The observed relative intensities suggested the formation of a more stable complex for βGlc than for αGlc. Experiments performed under the same conditions on receptor **4** with αGlc and βGlc gave very similar results, showing the corresponding peaks of the receptor, of its dimer, and of the complex present in comparable intensities for the two anomeric glycosides (Figure 4, top). The common features exhibited by this set of ESI−MS spectra, although just providing a qualitative picture, demonstrated the presence of the complex species as a major component for all of the investigated mixtures. A more quantitative description of the relative affinities of **3** and **4** for αGlc and βGlc could be obtained through collision-induced dissociation (CID) experiments performed on a triple-quadrupole

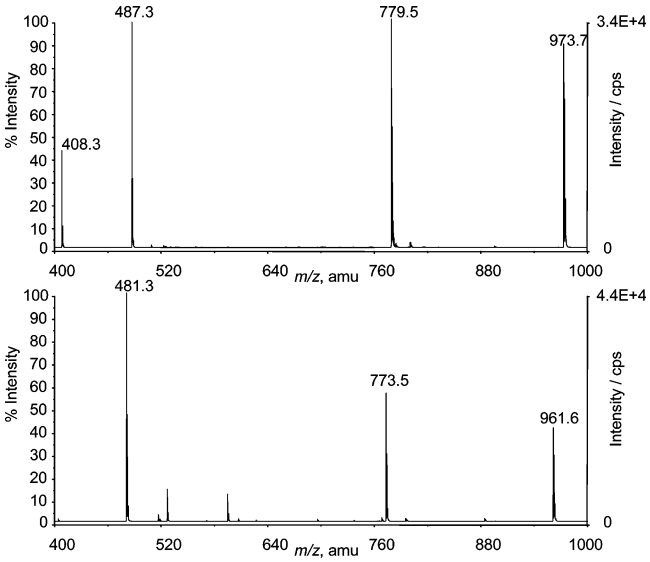


Figure 4. (+)ESI−MS spectra of (bottom) **3** + βGlc, 0.2 mM each; *m/z* 481.3 [**3**+H]$^+$, 773.5 [**3**·βGlc+H]$^+$, 961.6 [**3**·**3**+H]$^+$. (Top) **4** + βGlc, 0.2 mM each; *m/z* 487.3 [**4**+H]$^+$, 779.5, [**4**·βGlc+H]$^+$, 973.7 [**4**·**4**+H]$^+$. Solvent, $CHCl_3/CH_3CN$ 1:1; ESI voltage, 6 kV; sampling cone potential, 56 V.

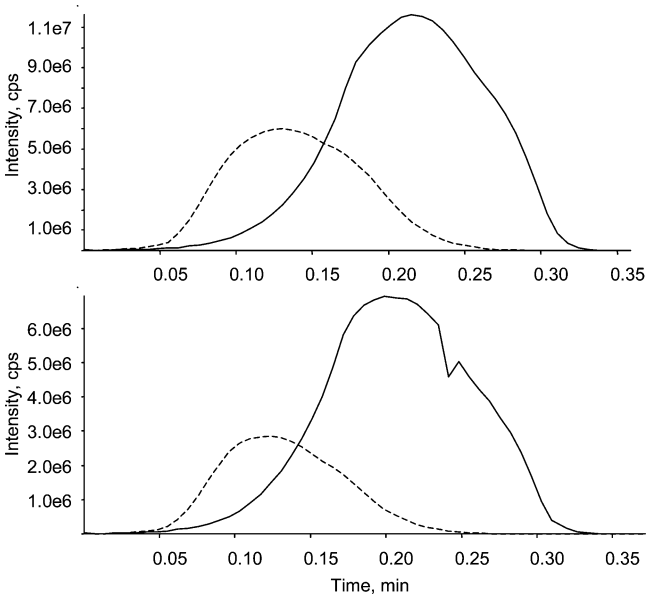


Figure 5. CID MS/MS analysis of the complex detected as the [M+H]$^+$ ion at *m/z* 779.5. Products, 779.5 ([M+H]$^+$, dotted line), 487.3 ([**4**+H]$^+$, solid line). Solvent, $CHCl_3/CH_3CN$ 1:1; ESI voltage, 6 kV; sampling cone potential, 56 V; signal acquisition, 0.36 min; 53 scans over collision energies from − 8.0 to − 34.0 eV in 0.5 eV steps (Q0 = −12 eV); collision-gas pressure, $P = 2.64\ 10^{-5}$ Torr. Top: **4** + βGlc, 0.2 mM each. Bottom: **4** + αGlc, 0.2 mM each.

mass spectrometer. A scan of the intensity of the [**4**·βGlc+H]$^+$ and [**4**+H]$^+$ ions originating from the ion of the complex selected at *m/z* 779 with increasing potential gave the profiles shown in Figure 5 (top), which crossed for a collision energy value of 7.4 eV, corresponding to the energy required to dissociate 50% of the complex under the specific experimental conditions. The corresponding CID profiles originating from the [**4**·αGlc+H]$^+$ ion under identical conditions (Figure 5, bottom) exhibited a crossing point for a collision energy value of 6.4 eV, showing that dissociation of the αGlc complex required a collision energy smaller by 1 eV than that required by the βGlc complex. CID profiles were analogously obtained

(20) This conformation in the solid state does not appear to be exceptional. See, for example, ref 6c.

for complexes of **3** with αGlc and *β*Glc run under identical experimental conditions in order to obtain comparable results. The corresponding profiles, similar in all respects to those depicted in Figure 5, gave crossing points for collision energies of 8.1 and 7.7 eV for *β*Glc and αGlc, respectively, demonstrating a stronger binding to the former glycoside also for **3** (see Supporting Information). Quite gratifyingly, results in the gas phase showed the same trend observed in solution; besides the same *β*/α selectivity order, both glycoside anomers were more strongly bound to **3** than to **4** in solution and in the gas phase.

Conclusion

In the present work, we have shown that very effective receptors for monosaccharides have been obtained by assembling the correctly matching H-bonding groups on a tripodal scaffold appropriately designed to form an adaptive cleft around the carbohydrate moiety. Following previous findings, we demonstrated that pyrrolic donors and iminic/aminic acceptors are far better H-bonding ligands for glycosidic hydroxyls than the most largely employed amidic and ureidic groups, showing geometric, coordinative, and functional complementarity to the donor/acceptor properties of aliphatic hydroxyls. When correctly located and precisely spaced in the side arms of the tripodal scaffold, since spacer length has been shown to be crucial, these H-bonding groups gave rise to a new generation of receptors showing micromolar affinities for monosaccharidic glycosides in chloroform and millimolar affinities in a polar solvent such as acetonitrile due to a strongly enthalpic interaction. Remarkable selectivities were also observed, reaching unprecedented selectivity factors for *β*Glc with the imino−pyrrolic receptor. Affinities and selectivities were univocally and quantitatively assessed on a common scale by generalized binding descriptors, that is, the median binding concentration BC_{50} and the intrinsic median binding concentration BC^0_{50} parameters, which we developed for host−guest systems of any nature and involving any number of binding constants, and which we propose as the most useful and most general indicators of binding affinity. These findings may open the way to a new generation of carbohydrate receptors, featuring hitherto unexplored pyrrolic H-bonding donors, and to a new and general way of assessing binding affinities.

Acknowledgment. The contribution of Dr. P. Gans, University of Leeds, and of Prof. A. Sabatini, Università di Firenze, to the upgrade of the HypNMR program is gratefully acknowledged. High-field NMR spectra were acquired at the Magnetic Resonance Center (CERM) facilities of the Università di Firenze. We thank Mr. M. Lucci for technical assistance. Ente Cassa di Risparmio di Firenze (Italy) is acknowledged for granting a 400 MHz NMR spectrometer.

Supporting Information Available: General methods, materials, and synthetic procedures for compounds **1**−**9**. Data tables, result pages, and plots of experimental and calculated shifts from titrations of glycosides of Chart 2. Plots of ITC titrations and CID profiles from ESI−MS experiments. Crystallographic tables and details. Appendix: generalized treatment of BC_{50} and BC^0_{50} for host−guests systems. Two-dimensional NMR spectra. X-ray crystallographic data in the electronic Crystallographic Information File (CIF) format. This material is available free of charge via the Internet at http://pubs.acs.org.

JA068754M