UNIVERSITY OF BRISTOL

University of Bristol - Explore Bristol Research
### General rights

Platform Synthetic Lectins for Divalent Carbohydrate Recognition in Water

Tom S. Carter, Tiddo J. Mooibroek, Patrick F. N. Stewart, Matthew P. Crump, M. Carmen Galan, and Anthony P. Davis*

Abstract: *Biomimetic carbohydrate receptors ("synthetic lectins") have potential as agents for biological research and medicine. However, although effective strategies are available for "all-equatorial" carbohydrates (glucose, etc.), the recognition of other types of saccharide under natural (aqueous) conditions is less well developed. Herein we report a new approach based on a pyrene platform with polar arches extending from aryl substituents. The receptors are compatible with axially substituted carbohydrates, and also feature two identical binding sites, thus mimicking the multivalency observed for natural lectins. A variant with negative charges forms 1:2 host/guest complexes with aminosugars, with $K_1 > 3000\,M^{-1}$ for axially substituted mannosamine, whereas a positively charged version binds the important α-sialyl unit with $K_1 \approx 1300\,M^{-1}$.*

Carbohydrate recognition is a central biological phenomenon that mediates a range of cellular processes.[1] Carbohydrate-binding molecules are important as research tools for investigating these processes, and potentially as diagnostic and therapeutic agents in medicine.[1–3] Studies in this area most commonly use lectins, the major class of saccharide-binding proteins, but lectins often lack the desired selectivities and tend to show low affinities (generally 10^3–$10^4\,M^{-1}$ for monosaccharides).[4] Moreover, as proteins, their therapeutic potential is limited by issues such as immunogenicity.[3] There is consequently much interest in small-molecule receptors, which could complement lectins and perhaps be developed for new types of application.[1,3,5] However, the design of such molecules has proved difficult, especially for biomimetic systems based on noncovalent bonding.[6] Although a variety of structures have been shown to be active in organic solvents,[7] there are few which can operate in the natural but challenging environment of water.[8]

We have approached this problem by constructing symmetrical cavities with an aromatic "roof" and "floor"
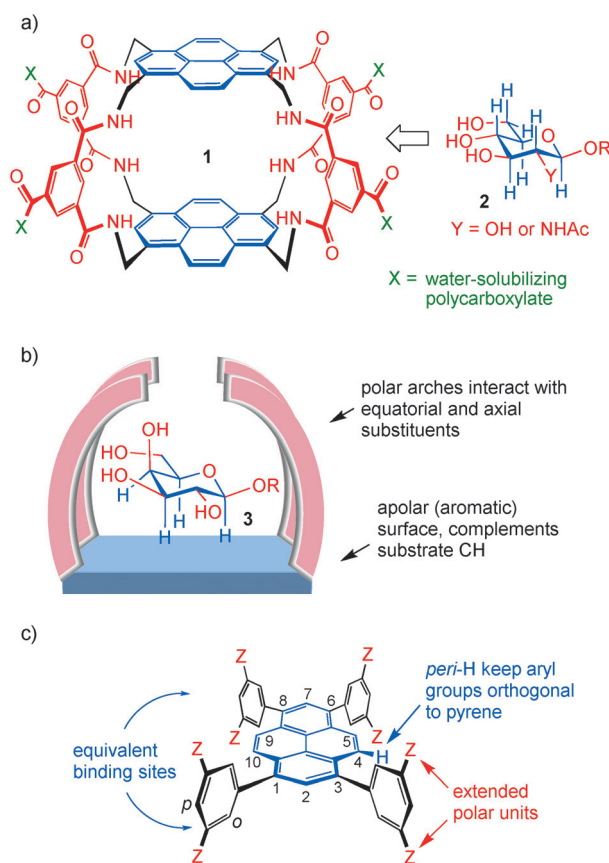
Figure 1. a) A receptor **1** for all-equatorial carbohydrates **2** (see Ref. [9a]). The symmetrical cavity matches the polar (red) and apolar (blue) groups in the substrate. b) A strategy for binding saccharides with axial substituents, illustrated for the β-galactosyl group **3** as the substrate. c) General design of the pyrene-based receptors described herein.

separated by polar spacers (e.g. **1**, Figure 1 a).[9] The designs are complementary to saccharides with all-equatorial substitution patterns (e.g. **2**), and have yielded encouraging results. Selectivities are good, and some affinities are above $10^4\,M^{-1}$, even for uncharged substrates.[9a,b] However, the selectivity for all-equatorial carbohydrates is a constraint on potential applications, as many substrates of interest do not belong to this family. Herein we report an alternative design strategy which rationally targets carbohydrates with axial substituents and which, for the first time, mimics the multivalency exhibited by many lectins.[1]

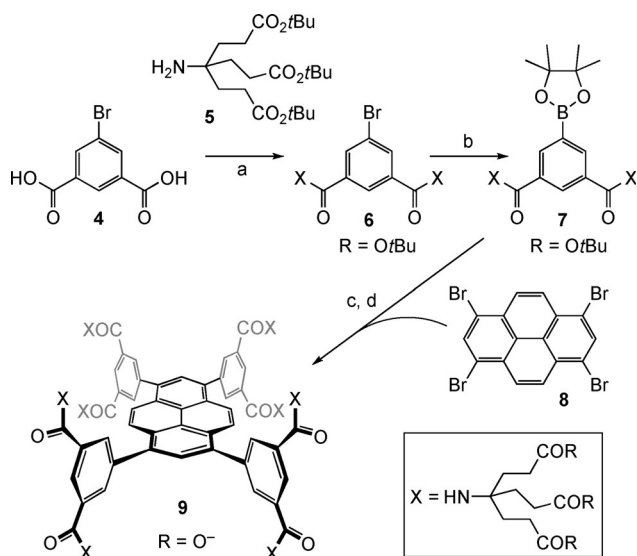While all-equatorial saccharides possess two roughly similar hydrophobic patches, other carbohydrates tend to be

[*] T. S. Carter, Dr. T. J. Mooibroek, P. F. N. Stewart, Prof. M. P. Crump, Dr. M. C. Galan, Prof. A. P. Davis
School of Chemistry, University of Bristol
Cantock's Close, Bristol, BS8 1TS (UK)
E-mail: Anthony.Davis@bristol.ac.uk

facially amphiphilic. The inversion of a stereocenter, as in β-galactosyl **3**, adds to the polarity of one face while decreasing the polarity of the other. A complementary binding site should therefore contain just one extended apolar surface, the remainder being mainly polar.[10] An approach to such structures might involve an aromatic platform with polar substituents that can arch over a bound carbohydrate (e.g. Figure 1 b). When considering options for realizing this architecture, we noted the potential of 1,3,6,8-tetraarylpyrenes (Figure 1 c). These compounds are readily prepared from tetrabromopyrene (**8**) by Suzuki–Miyaura methodology, and are forced to adopt nonplanar conformations owing to interactions between aryl and *peri*-H groups.[11] The *meta* positions on the aryl groups could provide anchor points for the polar "arches". Stereoisomers would be possible in the general case, but could be avoided by using symmetrically substituted aryl groups, as shown in Figure 1 c. Interestingly, this arrangement would generate two equivalent binding sites, thus mimicking the multivalency common in lectins.[1] Groups Z could be used to confer water-solubility as well as to provide polar interactions. For example, polyionic dendrimers, which are highly water solubilizing and capable of hydrogen bonding to polar carbohydrate substituents, could be readily installed.[9c]

As a prototype for this design, we chose the tetracosacarboxylate **9**. Modeling[12] showed that the dendrimers in **9**, though relatively small, possessed sufficient reach to interact with axial groups on a substrate. Protonated mannosamine **10**·H$^+$, with an axial NH$_3^+$ group, was found to be an especially promising substrate (Figure 2). Receptor **9** was prepared in 23% yield over four steps from diacid **4**, amine **5**,[13] and 1,3,6,8-tetrabromopyrene (**8**;[14] Scheme 1).[12] The anionic receptor **9** dissolved freely in water to give well-resolved ^{1}H NMR spectra, which were concentration-inde-

Figure 2. Model of receptor **9** bound to the β anomer of protonated mannosamine **10**·H$^+$. Aromatic portions of the receptor are shown in space-filling mode, side chains in pale green, mannosamine carbon atoms cyan. a) View from the side, with mannosamine in space-filling mode. b) View from above, with mannosamine in stick mode. Hydrogen bonds are shown in yellow.

pendent below 1.2 mM, thus implying a monomeric species. Fluorescence spectra showed an emission maximum at 425 nm (excitation wavelength: 380 nm), detectable down to nanomolar concentrations and with intensity directly proportional to concentration.

Carbohydrate recognition was studied by fluorescence and ^{1}H NMR titrations of **9** with a range of sugars.[12] Solutions were adjusted to pH 7, and the pH value was confirmed to be unchanged after each experiment. Titrations with aminosugars **10**–**12** (protonated at pH 7) yielded clear
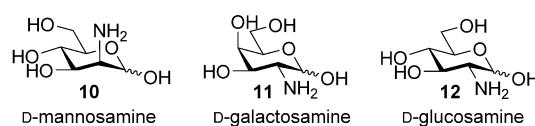
Scheme 1. Synthesis of receptor **9**. a) SOCl$_2$, reflux; then **5**, EtN*i*Pr$_2$, THF, 53%; b) [Pd(dppf)Cl$_2$], bis(pinacolato)diboron, KOAc, dioxane, 80°C, 80%; c) **8**, [Pd(dppf)Cl$_2$], Cs$_2$CO$_3$, dioxane, H$_2$O, 90°C, 53%; d) CF$_3$CO$_2$H, SiHEt$_3$, CH$_2$Cl$_2$; then aqueous NaOH, quantitative. dppf = 1,1′-bis(diphenylphosphino)ferrocene.
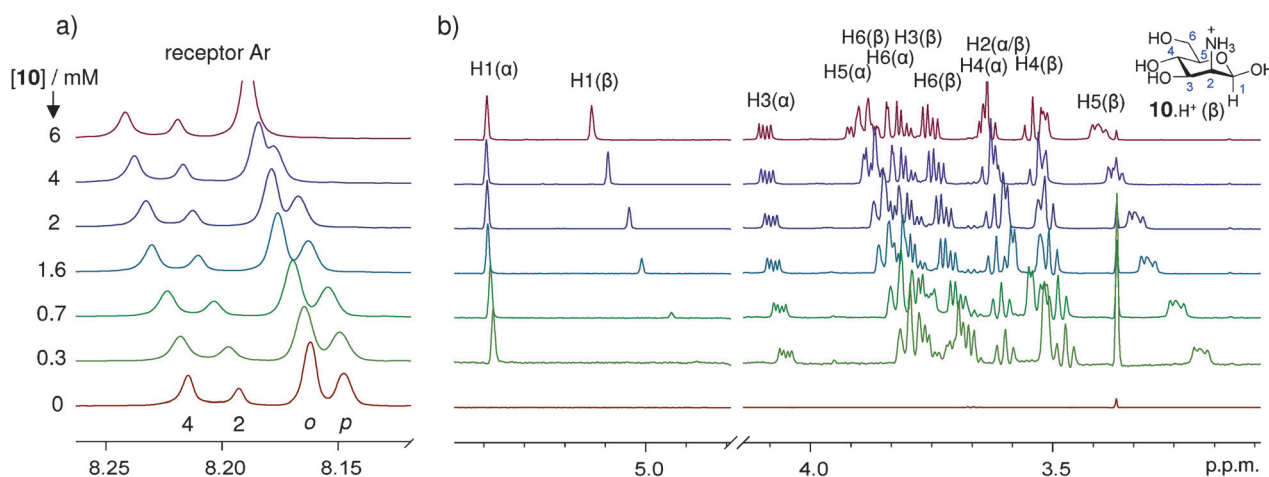
Figure 3. Partial ^{1}H NMR titration spectra of receptor **9** (0.20 mM) with D-mannosamine (**10**) in D$_2$O at pH 7 and 298 K. a) Signals from receptor aromatic hydrogen atoms. For numbering, see Figure 1 c. b) Mannosamine signals, with intensities normalized to compensate for the increase in concentration. The mannosamine H1(β) signal is affected by the water-suppression sequence and is artificially depressed in the early stages of the titration.

evidence of complex formation. For example, in NMR titrations, the addition of mannosamine (**10**) was accompanied by downfield changes in the positions of the receptor aromatic signals (see Figure 3 a). A Job plot[15] based on the receptor 4-H signal displacements gave a maximum at **9**/**10** = 1:2, thus confirming divalency (see Figure S14 in the Supporting Information). Analysis of the receptor signals, assuming the 1:2 binding model, gave stepwise binding constants $K_1 = 3120\,M^{-1}$ and $K_2 = 540\,M^{-1}$ (see Figure S9).[16] As $K_1 = 4\,K_2$ for noncooperative two-site binding,[17] it seems the two associations are almost independent. The results were supported by fluorescence titrations of **9** with **10**, which showed large increases in receptor emission intensity and could be analysed to give almost identical binding constants (Figure 4 and Table 1).

Changes in the positions of carbohydrate signals during the ^{1}H NMR titration were also informative (Figure 3 b). All the signals from **10** moved downfield during the titration implying that, as expected, these protons are shielded in the complex.[18] The spectra provide separate information for
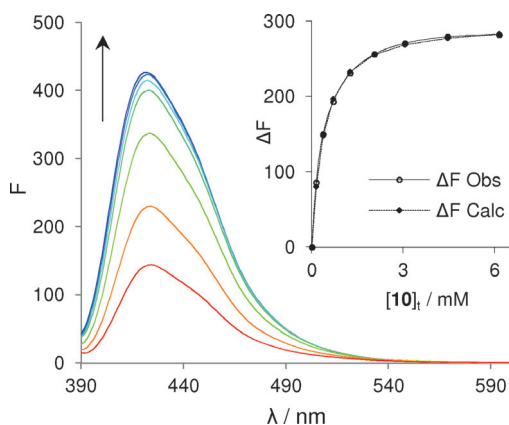
Table 1: Cumulative association constants (K_1, K_2) for 1:1 and 1:2 binding of receptor **9** with aminosugars in aqueous solution, as determined by ^{1}H NMR and fluorescence titrations.[a]

Carbohydrate	K_1 [M^{-1}]	K_2 [M^{-1}]
D-mannosamine (**10**)	3120,[b] 3120[c]	540,[b] 600[c]
D-galactosamine (**11**)	1800,[b] 2000[c]	200,[b] 180[c]
D-glucosamine (**12**)	1040,[b] 1100[c]	410,[b] 360[c]

[a] pH 7, $T = 298$ K. [b] ^{1}H NMR titration. [c] Fluorescence titration.

α and β anomers, present in the ratio 1:1.9.[19] Signals for hydrogen atoms on the α face of the β anomer [for example, H1(β), H5(β)] showed especially large movements, consistent with the modeled structure (Figure 2). From the α anomer, the proton H5(α) signal also moved substantially. NOE data (see Figure S15) confirmed that **9** and the carbohydrate are closely associated. Binding could not be quantified reliably because of the complexity of the system (two substrates, both forming 1:1 and 1:2 complexes). However, the signal movements were consistent with $K_1 \approx 3000\,M^{-1}$ for both anomers.

Titrations with galactosamine (**11**) and glucosamine (**12**) gave similar changes in both ^{1}H NMR and fluorescence spectra. Analysis of the data gave the binding constants shown in Table 1. Affinities were somewhat lower, suggesting that the binding sites of **9** favor axial NH$_3{}^+$ substitution. Uncharged monosaccharides did not appear to bind, but titrations with the disaccharides cellobiose, lactose, and maltose, and also methylamine, gave evidence of weak-complex formation ($K_1 \leq 16\,M^{-1}$).[12] It thus seems that the high affinities for aminosugars result from combining carbohydrate-specific interactions (hydrophobic, CH–π, hydrogen bonding) with electrostatic attraction.

Whereas cationic sugars are relatively uncommon in nature, anionic carbohydrates are widespread and play important roles. Especially significant is the α-linked *N*-acetylneuraminic acid (α-sialyl) unit **13**. This moiety commonly appears as a terminus of oligosaccharides, accessible for binding and therefore an important potential target.[20]

Figure 4. Fluorescence emission spectra and binding curve of receptor **9** (0.33 μM) titrated with D-mannosamine (**10**) in H$_2$O at pH 7 and 298 K. Excitation wavelength: 380 nm.

α-sialyl **13**

14 R = Me

N-acetylneuraminic acid
(β anomer) **15**

Figure 5. Fluorescence emission titration spectra and binding model fit of receptor **16** (0.50 μM) with methyl sialoside (**14**) in H₂O at pH 7 and 298 K. Excitation wavelength: 380 nm.

The substitution pattern in **13** is all-equatorial with an additional axial negatively charged substituent. This compound should be nicely complementary to our platform design (Figure 1 b), provided that the polar arches are positively charged.

To test this concept, we prepared the cationic receptor **16**, possessing 24 guanidinium units (see the Supporting Information). Guanidinium substituents were chosen to ensure

16

X = HN

3 Cl⁻

that the receptor would be fully protonated at pH 7.[21] Spectroscopic studies implied that receptor **16** is monomeric in water below a concentration of 1.2 mM. As expected, the ¹H NMR spectrum of **16** was unaffected by the pH value in the range 6–8, confirming full protonation.

Studies of **16** as a receptor for the α-sialyl unit **13** required a model substrate. The parent saccharide *N*-acetylneuraminic acid is readily available, but exists mainly (>90%) as the β anomer **15**, and is thus unrepresentative of **13**. Simple α-sialosides are not commercially available, and we therefore synthesized the methyl derivative **14** through a variation of a literature procedure.[12] The binding of **14** to **16** was studied by ¹H NMR and fluorescence spectroscopy. NOESY cross-peaks between substrate and receptor aromatic proton signals supported complex formation (see Figure S55). On the other hand, NMR titrations yielded relatively small changes in the signal positions, implying a looser geometry than for **9** + **10**. A Job plot based on a receptor aromatic signal confirmed the expected **16**/**14** = 1:2 binding stoichiometry (see Figure S54). For quantitative analysis, we employed a titration in which receptor **16** was added to substrate **14**. The carbohydrate signals moved upfield as expected, and several could be followed throughout (see Figure S57). Simultaneous analysis of four of these signals was consistent with three successive binding events, with K_1, K_2, and $K_3 = 1310$, 570, and 30 M⁻¹, respectively. Given the high density of positive charge on **16**, it is reasonable to suppose that a third (and possibly fourth) molecule of **14** might bind to the receptor.

Fluorescence titration of **16** with **14** yielded a surprisingly strong effect (Figure 5); receptor emission was reduced almost to zero by addition of the carbohydrate. Analysis of

the changes assuming a 1:2 binding model (see Figure 5) gave stepwise binding constants K_1, $K_2 = 1300$, 790 M⁻¹, consistent with the NMR data. The decrease in emission was found to depend on the chloride counterions; when these were replaced with trifluoroacetate, fluorescence increased on binding. We presume that the addition of **14** to **16** causes a rearrangement of the counterions, thus promoting fluorescence quenching.

Fluorescence titrations were also performed for receptor **16** with several uncharged monosaccharides.[12] Efficient quenching was again observed, although analysis suggested that binding was much weaker than for **14**. Moderate affinities were estimated for methyl β-D-glucoside ($K_1 = 43$ M⁻¹), methyl β-D-galactoside ($K_1 = 46$ M⁻¹), and galactose ($K_1 = 17$ M⁻¹). In these cases, NMR shifts were too small for analysis, but association between the receptor and the carbohydrate was confirmed by NOE enhancements.[22] Sodium acetate was also bound weakly ($K_1 = 17$ M⁻¹). The affinity for glucose was too small to be quantified implying that, as intended, the platform design can reverse the selectivity shown by our earlier synthetic lectins.

In conclusion, we have reported a rational design for synthetic lectins which mimic the multivalency shown by natural lectins and, unlike earlier systems, can accommodate substrates with axial substituents. There is scope for varying the system, for example, by altering/extending the side chains or changing one aryl substituent. The potential for binding the β-sialyl group **13** is especially significant, and will be a focus of future efforts.

Acknowledgements

[1] a) D. Solís, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Smetana, H. J. Gabius, *Biochim. Biophys. Acta Gen. Subj.* **2015**, *1850*, 186; b) B. Wang, G.-J. Boons, *Carbohydrate Recognition: Biological Problems, Methods and Applications*, Wiley, Hoboken, **2011**; c) *The Sugar Code—Fundamentals of Glycoscience* (Ed.: H.-J. Gabius), Wiley-Black-well, Weinheim, **2009**.

[2] J. Balzarini, *Nat. Rev. Microbiol.* **2007**, *5*, 583.

[3] Y. Nakagawa, T. Doi, T. Taketani, K. Takegoshi, Y. Igarashi, Y. Ito, *Chem. Eur. J.* **2013**, *19*, 10516.

[4] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719.

[5] C. E. Miron, A. Petitjean, *ChemBioChem* **2015**, *16*, 365; Y. Nakagawa, Y. Ito, *Trends Glycosci. Glycotechnol.* **2012**, *24*, 1; M. Mazik, *RSC Adv.* **2012**, *2*, 2630; A. P. Davis, *Org. Biomol. Chem.* **2009**, *7*, 3629; A. P. Davis, R. S. Wareham, *Angew. Chem. Int. Ed.* **1999**, *38*, 2978; *Angew. Chem.* **1999**, *111*, 3160.

[6] An alternative strategy is to employ covalent B−O bonds; see, for example, Ref. [1b] and T. D. James, M. D. Phillips, S. Shinkai, *Boronic Acids in Saccharide Recognition*, RSC, Cambridge, **2006**.

[7] For recent examples, see: Y. Ohishi, H. Abe, M. Inouye, *Chem. Eur. J.* **2015**, *21*, 16504; J. Lippe, M. Mazik, *J. Org. Chem.* **2015**, *80*, 1427; N. Chandramouli, Y. Ferrand, G. Lautrette, B. Kauffmann, C. D. Mackereth, M. Laguerre, D. Dubreuil, I. Huc, *Nat. Chem.* **2015**, *7*, 334; A. Schmitt, O. Perraud, E. Payet, B. Chatelet, B. Bousquet, M. Valls, D. Padula, L. Di Bari, J. P. Dutasta, A. Martinez, *Org. Biomol. Chem.* **2014**, *12*, 4211; O. Francesconi, M. Gentili, C. Nativi, A. Arda, F. J. Canada, J. Jiménez-Barbero, S. Roelens, *Chem. Eur. J.* **2014**, *20*, 6081; S. Rieth, M. R. Miner, C. M. Chang, B. Hurlocker, A. B. Braunschweig, *Chem. Sci.* **2013**, *4*, 357.

[8] See, for example: Y. Jang, R. Natarajan, Y. H. Ko, K. Kim, *Angew. Chem. Int. Ed.* **2014**, *53*, 1003; *Angew. Chem.* **2014**, *126*, 1021; M. Rauschenberg, S. Bandaru, M. P. Waller, B. J. Ravoo, *Chem. Eur. J.* **2014**, *20*, 2770; S. M. Bromfield, A. Barnard, P. Posocco, M. Fermeglia, S. Pricl, D. K. Smith, *J. Am. Chem. Soc.* **2013**, *135*, 2911; C. Geffert, M. Kuschel, M. Mazik, *J. Org. Chem.* **2013**, *78*, 292.

[9] For recent examples, see: a) P. Rios, T. S. Carter, T. J. Mooibroek, M. P. Crump, M. Lisbjerg, M. Pittelkow, N. T. Supekar, G.-J. Boons, A. P. Davis, *Angew. Chem. Int. Ed.* **2016**, *55*, 3387; *Angew. Chem.* **2016**, *128*, 3448; b) T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump, A. P. Davis, *Nat. Chem.* **2016**, *8*, 69; c) H. Destecroix, C. M. Renney, T. J. Mooibroek, T. S. Carter, P. F. N. Stewart, M. P. Crump, A. P. Davis, *Angew. Chem. Int. Ed.* **2015**, *54*, 2057; *Angew. Chem.* **2015**, *127*, 2085; d) C. Ke, H. Destecroix, M. P. Crump, A. P. Davis, *Nat. Chem.* **2012**, *4*, 718.

[10] As commonly observed for proteins; see, for example: K. L. Hudson, G. J. Bartlett, R. C. Diehl, J. Agirre, T. Gallagher, L. L. Kiessling, D. N. Woolfson, *J. Am. Chem. Soc.* **2015**, *137*, 15152.

[11] X. Feng, H. Tomiyasu, J. Y. Hu, X. Wei, C. Redshaw, M. R. J. Elsegood, L. Horsburgh, S. J. Teat, T. Yamato, *J. Org. Chem.* **2015**, *80*, 10973.

[12] For further details, see the Supporting Information.

[13] G. R. Newkome, C. D. Weis, *Org. Prep. Proced. Int.* **1996**, *28*, 495–498.

[14] K. Ogino, S. Iwashima, H. Inokuchi, Y. Harada, *Bull. Chem. Soc. Jpn.* **1965**, *38*, 473–477.

[15] C. Y. Huang, *Methods Enzymol.* **1982**, *87*, 509.

[16] Conditions were chosen to minimize ionic strength and maximize electrostatic interactions (see the Supporting Information). As expected, binding constants are lowered by the addition of salts.

[17] P. Thordarson, *Chem. Soc. Rev.* **2011**, *40*, 1305.

[18] The proportion of carbohydrate bound is highest early in the titration.

[19] Binding-constant measurements based on receptor properties (fluorescence, receptor NMR signals) should give average values for the two anomers; for a discussion, see: N. P. Barwell, M. P. Crump, A. P. Davis, *Angew. Chem. Int. Ed.* **2009**, *48*, 7673; *Angew. Chem.* **2009**, *121*, 7809.

[20] T. Angata, A. Varki, *Chem. Rev.* **2002**, *102*, 439; A. Varki, *Trends Mol. Med.* **2008**, *14*, 351.

[21] The corresponding tetracosaamine precursor to **16** was also studied. However, partial protonation of the amino groups resulted in broadened spectra, which could not be interpreted.

[22] As the affinities for **16** + uncharged monosaccharides were quantified by just one technique, they should be seen as tentative.