DOI: 10.1002/ejoc.200700295

Mimicking the Binding Motifs Found in the Crystal Structures of Protein–Carbohydrate Complexes: An Aromatic Analogue of Serine or Threonine Side Chain Hydroxyl/Main Chain Amide

Monika Mazik*[a] and Alexander König[a]

Keywords: Molecular Recognition / Receptors / Carbohydrates / Hydrogen Bonds / Supramolecular Chemistry

An aromatic analogue of the side chain hydroxyl/main chain amide of Ser or Thr was used for the construction of artificial receptors for *N*-acetylneuraminic acid (Neu5Ac), the most common occurring sialic acid. The acyclic receptors, incorporating only neutral recognition sites, are able to bind Neu5Ac with an overall binding constant β_2 of 10^5 M^{-2} in a competitive solvent, such as water-containing [D_6]DMSO. Both receptors show remarkable selectivity for Neu5Ac over D-glucuronic acid (GlcA).

Introduction

The binding motifs observed in the crystal structures of protein–carbohydrate complexes[1] provide much of the inspiration for the development of artificial carbohydrate receptors.[2] The protein–carbohydrate interactions involve hydrogen bonding, van der Waals forces, interactions of sugar CHs with aromatic residues of the protein, and metal coordination. Quiocho et al. pointed out that "hydrogen bonds are the main factors in conferring specificity and affinity to protein–carbohydrate interactions".[1a] The hydrogen bonds have both neutral and ionic character, and are both direct and water-mediated. Ion pairing and ionic hydrogen bonding are frequently observed in the complexation of proteins with charged sugars, such as *N*-acetylneuraminic acid (**3**), which is the most commonly occurring sialic acid. It should be noted that the design of both selective and effective carbohydrate receptors operating through noncovalent interactions in competitive media still represents a significant challenge.

Our interest in the area of sugar recognition with artificial receptors concentrates on receptors that possess a relatively simple, acyclic structure and that are expected to complex uncharged carbohydrates through neutral and charge-reinforced hydrogen bonds in combination with the CH–π interactions between the sugar CHs and the aromatic rings of the receptor.[3] Recently, we showed that binding motifs (including cooperative and bidentate hydrogen bonds) observed in the complexation of uncharged sugars with lectins can be successfully mimicked with acyclic pyridine- and pyrimidine-based receptors.[3f] In this study we focused on mimicking the binding motifs shown in Figure 1a, b.

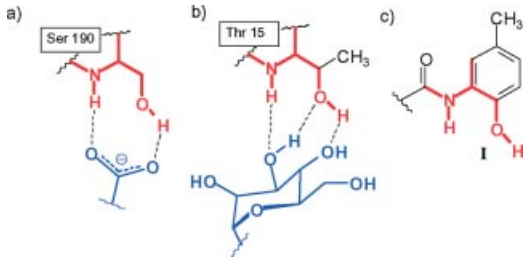


Figure 1. (a) Examples of charge-reinforced hydrogen bonds in the complex between 2-α-*O*-methyl *N*-acetylneuraminic acid and rhesus rotavirus hemagglutinin.[1e,4a] (b) Neutral hydrogen bonds in the complex between Manα6(Manα3)Man and concanavaline A.[1c,4b] (c) Aromatic analogue of the side chain hydroxyl/main chain amide groups of Ser or Thr.

The analyses of the crystal structures of sialic acid binding lectins with bound *N*-acetylneuraminic acid (Neu5Ac) or Neu5Ac-containing ligands showed that, with the exception of the polyoma virus, the carboxylate moiety of Neu5Ac interacts with the main-chain amide groups, polar side chains (especially serine, see Figure 1a), and ordered water molecules rather than fully charged side chains[1c,1d] (in contrast, formation of ion pairs with the Neu5Ac carboxylate appears to be a common feature of neuraminidases[1d]).

The binding motif shown in Figure 1a was found, for example, in the crystal structure of 2-α-*O*-methyl-*N*-acetylneuraminic acid with rhesus rotavirus hemagglutinin.[1e,4a] The noncovalent interactions shown in Figure 1a consist of two charge-reinforced hydrogen bonds: one from the side

[a] Institut für Organische Chemie der Technischen Universität Braunschweig,
Hagenring 30, 38106 Braunschweig

Supporting information for this article is available on the WWW under http://www.eurjoc.org or from the author.

chain hydroxyl group of Ser190 to the carboxylate oxygen, and one from the main chain amide functionality of Ser190 to the sugar carboxylate. Neutral hydrogen bonds formed between sugar hydroxyl groups and the side chain hydroxyl moiety of Ser or Thr, as well as the main chain amide functionality of the same amino acid, could also be observed in the crystal structures of protein–carbohydrate complexes. In the crystal structure of concanavaline A with Manα6-(Manα3)Man (see Figure 1b), for example, the α1-3-linked mannose contributes to the binding of the trisaccharide by hydrogen bonds between the 3- and 4-OH groups and the hydroxyl moiety of Thr15, as well as between the 3-OH and the main chain NH group of the same amino acid.[1c,4b]

The binding motifs shown in Figure 1a, b inspired the design of an aromatic analogue of the side chain hydroxyl/main chain amide groups of Ser or Thr, as shown in Figure 1c (recognition unit **I**). Artificial receptors consisting of more than one recognition unit **I** were expected to complex Neu5Ac (**3**) through multiple interactions, incorporating charge-reinforced hydrogen bonds with the Neu5Ac carboxylate and neutral hydrogen bonds with the sugar hydroxy groups (2-OH, 4-OH, and the OHs of the glyceryl side chain). Furthermore, interactions with the acetamido moiety and the CH units of Neu5Ac were expected to provide further stabilization of the receptor–sugar complexes. In the complexes of sialic acid binding lectins, the glyceryl side chain of Neu5Ac is often involved in hydrogen bonds; the acetamido moiety of Neu5Ac is frequently a significant recognition determinant.

Neu5Ac and Neu5Ac-containing ligands play a key role in a variety of biological processes,[1e,4a,5] including different cellular recognition processes. Specificity for Neu5Ac-containing ligands is expressed by the hemagglutinins of numerous viruses, notably of influenza (including H5N1 influenza A viruses[5]), as well as several others, such as Sendai, Newcastle disease, and polyoma viruses.[1e] Furthermore, Neu5Ac is known to be overexpressed on the cell surface of tumor cells. Thus, the design of artificial receptors for the recognition of Neu5Ac may serve as a basis for the development of new therapeutics or sensors.[6]

Results and Discussion

Molecular modeling calculations indicated that a 3,3′,5,5′-tetrasubstituted diphenylmethane or a 3,3′,5,5′-tetrasubstituted biphenyl scaffold (see formula **1** and **2**, respectively) provides a cavity of the correct shape and size for Neu5Ac encapsulation. According to the molecular modeling studies, receptors **1** and **2**, including four recognition units **I**, should be able to form both 1:1 and 1:2 receptor/sugar complexes with Neu5Ac (see Figure 2).

Receptors **1** and **2** were prepared by the reaction of 2-amino-4-methylphenol (**15**) with 3,3′,5,5′-diphenylmethanetetracarbonyl (**14**) or 3,3′,5,5′-biphenyltetracarbonyl tetrachloride (**22**) (see Schemes 1 and 2, and the Supporting Information).

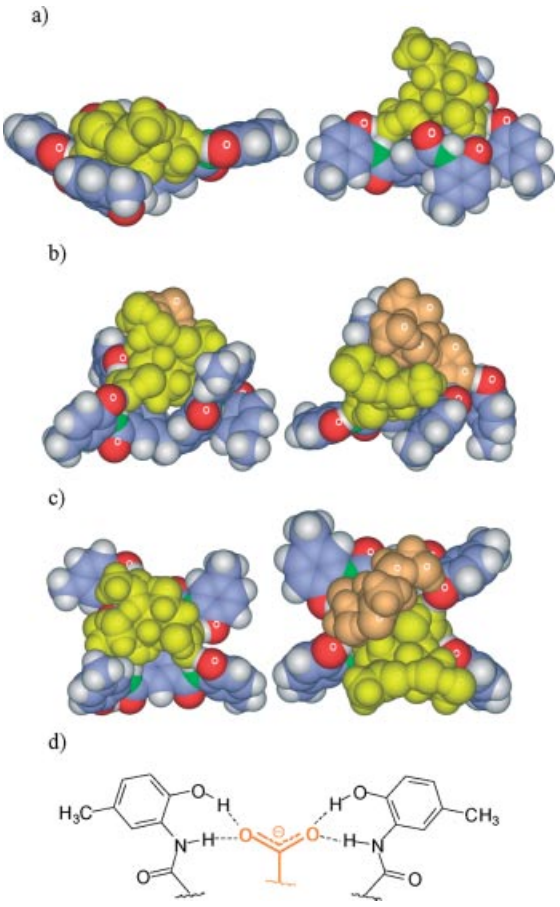
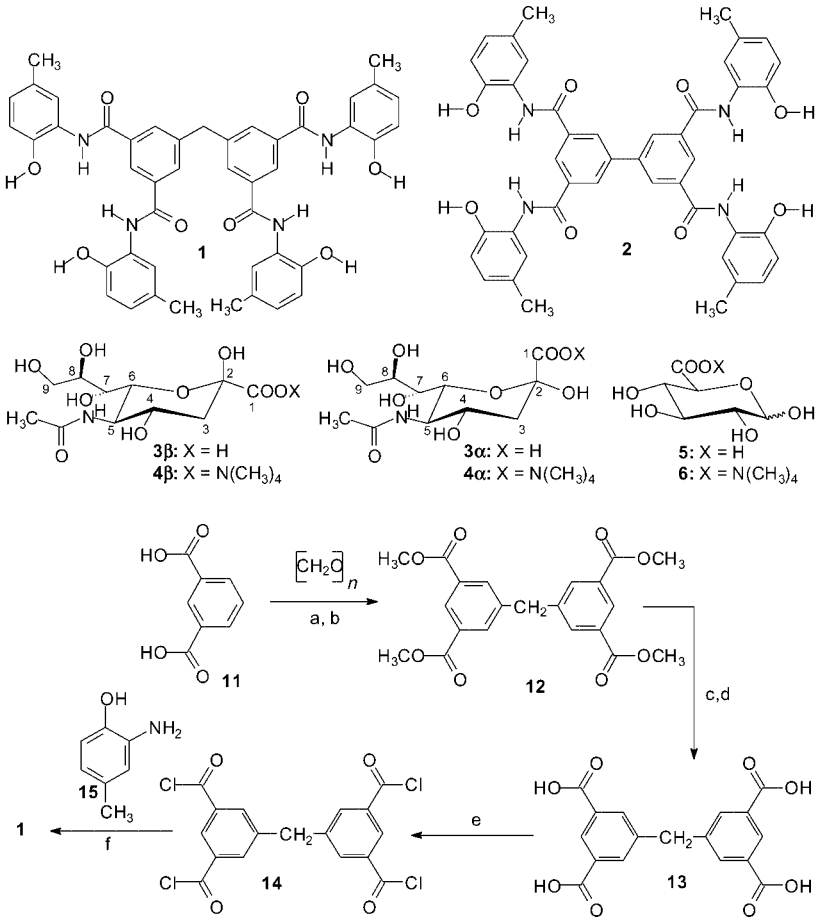


Figure 2. Energy-minimized structure of the (a) 1:1 complex between receptor **1** and Neu5Ac **3β** (two different representations), (b) 1:2 receptor/sugar complex between **1** and **3β** (two different representation), (c) 1:1 (left) and 1:2 (right) receptor/sugar complex between **2** and **3β**. (d) Examples of charge-reinforced hydrogen bonds indicated by molecular modeling (MacroModel V.8.5, OPLS-AA force field, MCMM, 50000 steps). Color code: receptor C, blue; receptor N, green; receptor O, red; the sugar molecule is highlighted in yellow or orange.

The interactions of receptors **1** and **2**[7] with Neu5Ac were investigated by ^{1}H NMR spectroscopy in water-containing [D$_6$]DMSO ([receptor]/[H$_2$O] ≈ 1:100; the solubility behavior of **1** and **2** prevents the binding studies in both pure water and nonpolar solvents, such as chloroform). The binding properties of **1** and **2** towards Neu5Ac were compared with those towards glucuronic acid (GlcA, **5**), an anionic sugar actively participating in biotransformations and detoxification processes of a number of endogenous compounds, through hepatic glucuronidation.[8] The titration experiments were carried out by adding increasing amounts of the tetramethylammonium salt of Neu5Ac (for α- and β-anomer, see formulas **4α** and **4β**) or tetramethylammonium D-glucuronate (**6**) to a solution of receptor **1** or **2**.[9] The ^{1}H NMR spectroscopic binding titration data were analyzed by using the Hostest 5.6 program.[10] Additionally, comparative binding studies were carried out with neutral sugars, such as methyl β-D-glucopyranoside (**7**), methyl β-D-galactopyranoside (**8**), D-maltose (**9**), and D-lactose (**10**).



Scheme 1. a) 20% Oleum; b) CH_3OH, HCl; c) 10% NaOH; d) 50% H_2SO_4; e) $SOCl_2$, THF; f) **15** (9 equiv.), THF.

Scheme 2. a) $NaNO_2$, 20% HCl, KI; b) $PdCl_2$(dppf), KOAc/DMF, 80 °C, 2 h; c) **17** (2 equiv.), $PdCl_2$(dppf), CsF, 80 °C, 12 h; d) 10% NaOH; e) 50% H_2SO_4; f) $SOCl_2$, THF; g) **15** (10 equiv.), THF.

During the titration of **1**[11] with **4** (used as an anomeric mixture), the signals due to the NH and OH protons of **1** shifted downfield with strong broadening of the peaks and they were almost unobservable after the addition of about 1 equiv. of **4** (see Figure 3c), which indicates the important contribution of the OH and NH groups of **1** to the formation of the complex in the competitive solvent. The complexation between **1** and **4** was further evidenced by chemical shift changes of the CH units of **1** (in the range of 0.03–0.07 ppm; Figure 3a, b), which were monitored as a function of sugar concentration.[12] The best fit of the titration data was obtained with the mixed 1:1 and 1:2 receptor/sugar binding model (for comparison of the plots obtained for the different binding models, see Figures S1a and S1b,

Supporting Information); the binding constants were found to be K_{a1} = 140 M^{-1} and K_{a2} = 1450 M^{-1} (β_2 = 2.03×10^5 M^{-2}). Typical titration curves are shown in Figure 3d, e.

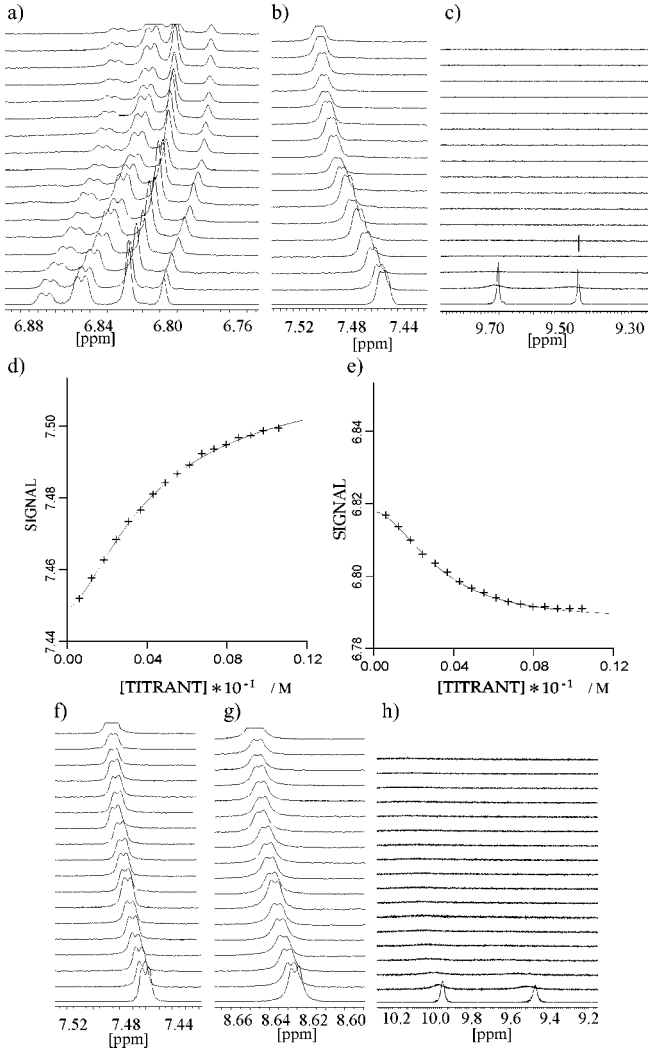


Figure 3. (a–c) Partial ^{1}H NMR spectra (400 MHz; water-containing [D$_6$]DMSO) of receptor **1** after the addition of (from bottom to top) 0.0–19 equiv. of **4** ([**1**] = 0.51 mM). Shown are chemical shifts of the phenyl CH (a, b), OH and NH (c) resonances of **1**. (d,e) Plots of the chemical shifts of the phenyl CHs of **1** as a function of added **4**. (f–h) Partial ^{1}H NMR spectra of receptor **2** after the addition of 0.0–21 equiv. of **4** ([**2**] = 0.50 mM). Shown are chemical shifts of the phenyl CH (f), biphenyl CH (g), OH and NH (h) resonances of **2**.

Similar to **1**, the NH and OH signals of receptor **2** broaden during the titration with **4**, and were almost unobservable after the addition of about 1 equiv. of **4** (see Figure 3h). Furthermore, the ^{1}H NMR spectra showed small changes in the chemical shifts of the CH protons of **2** (in the range of 0.02–0.05 ppm; Figure 3f, g). The spectroscopic changes indicated stronger 1:1 binding followed by weaker association of the second sugar molecule. The best fit of the titration data was obtained with the mixed 1:1 and 1:2 receptor-sugar binding model; the binding constants were found to be K_{a1} = 930 M^{-1} and K_{a2} = 120 M^{-1} (β_2 = 1.11×10^5 M^{-2}).[11d]

The comparison of the overall binding constants β_2 indicates that diphenylmethane-based receptor **1** exhibits about a twofold higher affinity for sugar **4** than biphenyl-based receptor **2** in the competitive solvent (Figure 3).

The ^{1}H NMR binding studies between **1** or **2** and sugar **6** indicated that the interactions with GlcA are less favorable than those with Neu5Ac. The chemical shifts of the OH, NH, and CH signals (see Figure 4) were affected weakly (some CH signals almost did not move, $\Delta\delta$ < 0.01 ppm) during the titrations with **6**. The NH and OH signals of **1** or **2** broaden during the titrations, but in contrast to the titrations with **4**, were still observable, even after the addition of 20 equiv. of **6**. Among the CH signals, the signal of the CH proton in the neighboring position to the OH group of **1** or **2** shows the largest shift ($\Delta\delta_{max}$ = 0.018 ppm in the case of **2**; in the case of **1** the changes were smaller). However, the binding constants could not be accurately determined on the basis of the small chemical shift changes; the Hostest program indicated $K_a \approx$ 100 M^{-1} for **2**·**6** (the titration curve is shown in Figure S1c). Thus, the ^{1}H NMR spectroscopic binding studies suggested high selectivity for Neu5Ac over GlcA (of a factor of about 1000 in the case of **2**), which indicates an important contribution of the glyceryl side chain and the acetamido moiety of Neu5Ac to the formation of the complex, in line with the observation in natural complexes.

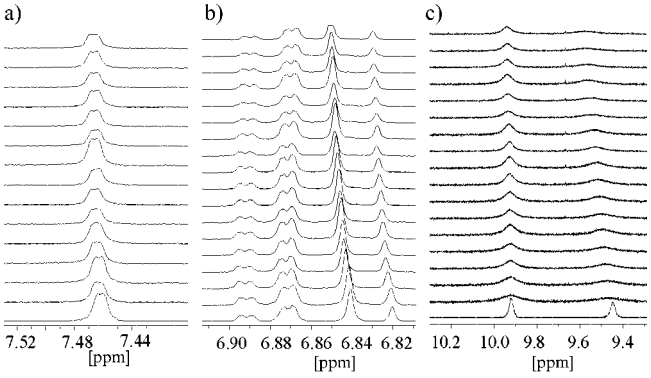


Figure 4. (a) Partial ^{1}H NMR spectra (400 MHz; water-containing [D$_6$]DMSO) of receptor **1** after the addition of 0.0–19 equiv. of **6** ([**1**] = 0.45 mM). Shown are the phenyl CH signals of **1**. (b,c) Partial ^{1}H NMR spectra of receptor **2** after the addition of 0.0–19 equiv. of **6** ([**2**] = 0.59 mM). Shown are the phenyl CHs (b), NH and OH signals (c) of **2**.

The comparative binding studies with neutral sugars **7**–**10** showed that the interactions between the neutral binding partners are unable to effectively compete with the interactions with the solvent molecules. The complexation-induced shifts observed after the addition of 20 equiv. of **7**–**10** into the solution of **1** or **2** in water-containing [D$_6$]DMSO were very small (largest changes ≈ 0.015 ppm for NH/OH signals, ≈ 0.005 for CH signals of **1** or **2**). It should be noted that the spectral changes observed after the addition of disaccharides **9** and **10** were more substantial than those observed in the presence of monosaccharides **7** and **8**. These

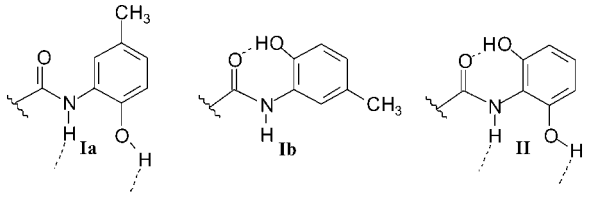
results indicate that neutral sugars need receptors containing both neutral and ionic recognition sites for effective complexation in competitive media (see also ref.[3h]).

Conclusions

The analysis of the binding motifs in the protein–carbohydrate complexes has inspired the design of an aromatic analogue of the side chain hydroxyl/main chain amide groups of Ser or Thr. This recognition unit (Figure 1c) was used for the construction of artificial receptors for the complexation of Neu5Ac, the most naturally abundant sialic acid. Receptors **1** and **2** are able to form 1:2 receptor/sugar complexes with Neu5Ac with the overall binding constant β_2 of 10^5 M^{-2} in water-containing [D_6]DMSO ([receptor]:[water] ≈ 1:100). The affinity of **1** is about twofold higher than that of **2**. Binding of Neu5Ac is achieved through a combination of neutral and charge-reinforced hydrogen bonds as well as CH–π interactions.

Receptors **1** and **2** are able to distinguish between two anionic sugars, Neu5Ac and GlcA, with high selectivity for Neu5Ac over GlcA (of a factor of about 1000 in the case of **2**). Thus, acyclic receptors **1** and **2**, which contain only neutral recognition sites and operate through noncovalent interactions, exhibit remarkable binding selectivity in the competitive solvent.

Further improvement of the binding affinity is expected with receptors constructed on the basis of recognition unit **II** as the intermolecular interactions formed through such receptors should compete more effectively with the intramolecular hydrogen bonds (see **Ia**, **Ib**, and **II**).



The synthesis of new receptors of this type and complexation studies with different sugar molecules, including 2-α-*O*-methyl Neu5Ac, which serves as a model for Neu5Ac bound to a neighboring unit in glycoproteins,[6b,13] are in progress.

The acyclic scaffold of the receptors provides simplicity in the synthetic plan for many modifications to the structure of the receptor, which provides a basis for systematic studies directed towards recognition motifs of carbohydrates.

Supporting information (see footnote on the first page of this article): Syntheses of **1** and **2**; additional titration curves.

Acknowledgments

Financial Support by the Deutsche Forschungsgemeinschaft is gratefully acknowledged.

[1] a) F. A. Quiocho, *Pure Appl. Chem.* **1989**, *61*, 1293–1306; b) R. U. Lemieux, *Chem. Soc. Rev.* **1989**, *18*, 347–374; c) H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674; d) W. I. Weiss, K. Drickamer, *Annu. Rev. Biochem.* **1996**, *65*, 441–473; e) H. Lis, N. Sharon, *Lectins*, Kluwer Academic Publishers, Dordrecht, The Netherlands, **2003**.

[2] For reviews on carbohydrate recognition with artificial receptors, see: a) A. P. Davis, R. S. Wareham, *Angew. Chem.* **1999**, *111*, 3161–3179; *Angew. Chem. Int. Ed.* **1999**, *38*, 2979–2996; b) A. P. Davis, T. D. James in *Functional Synthetic Receptors* (T. Schrader, A. D. Hamilton, Eds.), Wiley-VCH, Weinheim, Germany, **2005**, p. 45–109; c) For a review on boronic acid based receptors, see: T. D. James, S. Shinkai, *Top. Curr. Chem.* **2002**, *218*, 159–200.

[3] a) M. Mazik, H. Bandmann, W. Sicking, *Angew. Chem.* **2000**, *112*, 562–565; *Angew. Chem. Int. Ed.* **2000**, *39*, 551–554; b) M. Mazik, W. Sicking, *Chem. Eur. J.* **2001**, *7*, 664–670; c) M. Mazik, W. Radunz, W. Sicking, *Org. Lett.* **2002**, *4*, 4579–4582; d) M. Mazik, W. Radunz, R. Boese, *J. Org. Chem.* **2004**, *69*, 7448–7462; e) M. Mazik, W. Sicking, *Tetrahedron Lett.* **2004**, *45*, 3117–3121; f) M. Mazik, H. Cavga, P. G. Jones, *J. Am. Chem. Soc.* **2005**, *127*, 9045–9052; g) M. Mazik, A. König, *J. Org. Chem.* **2006**, *71*, 7854–7857; h) M. Mazik, H. Cavga, *J. Org. Chem.* **2006**, *71*, 2957–2963; i) M. Mazik, M. Kuschel, W. Sicking, *Org. Lett.* **2006**, *8*, 855–858; j) M. Mazik, H. Cavga, *J. Org. Chem.* **2007**, *72*, 831–838.

[4] a) P. R. Dormitzer, Z. Y. Sun, G. Wagner, S. C. Harrison, *EMBO J.* **2002**, *21*, 885–897; b) J. H. Naismith, R. A. Field, *J. Biol. Chem.* **1996**, *271*, 972–976.

[5] K. Shinya, M. Ebina, S. Yamada, M. Ono, N. Kasai, Y. Kawaoka, *Nature* **2006**, *440*, 435–436.

[6] For examples of boronic acid based receptors for Neu5Ac (using covalent interactions for sugar binding), see: a) H. Otsuka, E. Uchimura, H. Koshino, T. Okano, K. Kataoka, *J. Am. Chem. Soc.* **2003**, *125*, 3493–3502; b) K. Djanashvili, L. Frullano, J. A. Peters, *Chem. Eur. J.* **2005**, *11*, 4010–4018; c) M. Yamamoto, M. Takeuchi, S. Shinkai, *Tetrahedron* **1998**, *54*, 3125–3140; d) M. He, R. J. Johnson, J. O. Escobedo, P. A. Beck, K. K. Kim, N. N. S. t. Luce, C. J. Davis, P. T. Lewis, F. R. Fronczek, B. J. Melancon, A. A. Mrse, W. D. Treleaven, R. M. Strongin, *J. Am. Chem. Soc.* **2002**, *124*, 5000–5009; e) T. Zhang, E. V. Anslyn, *Org. Lett.* **2006**, *8*, 1649–1652.

[7] For examples of other carbohydrate receptors incorporating phenolic hydroxy groups, see: a) H. Abe, Y. Aoyagi, M. Inouye, *Org. Lett.* **2005**, *7*, 59–61; b) S. Anderson, U. Neidlein, V. Gramlich, F. Diederich, *Angew. Chem.* **1995**, *107*, 1722–1726; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1596–1600; c) A. Bähr, A. S. Droz, M. Püntener, U. Neidlein, S. Anderson, P. Seiler, F. Diederich, *Helv. Chim. Acta* **1998**, *81*, 1931–1963; d) T. Mizutani, T. Kurahashi, T. Murakami, N. Matsumi, H. Ogoshi, *J. Am. Chem. Soc.* **1997**, *119*, 8991–9001; e) Y. Aoyama, Y. Tanaka, S. Sugahara, *J. Am. Chem. Soc.* **1989**, *111*, 5397–5404.

[8] M. Segura, V. Alćazar, P. Prados, J. de Mendoza, *Tetrahedron* **1997**, *53*, 13119–13128, and references therein.

[9] Tetralkylammonium ions are commonly used as countercations in the binding studies of anions. For a recent discussion of the solvent and countercation effects, see: J. L. Sessler, D. E. Gross, W.-S. Cho, V. M. Lynch, F. P. Schmidtchen, G. W. Bates, M. E. Light, P. A. Gale, *J. Am. Chem. Soc.* **2006**, *128*, 12281–12288.

[10] C. S. Wilcox, N. M. Glagovich, *Program HOSTEST 5.6*; University of Pittsburgh, Pittsburgh, PA, **1994**.

[11] a) Dilution experiments show that receptors **1** and **2** do not self-aggregate in the used concentration range; b) For each system, at least four titrations were carried out; for each titration 15–21 samples were prepared; c) Errors in K_a are <15%; d) The binding constants were determined at 25 °C. K_{a1} corresponds to the 1:1 association constant. K_{a2} corresponds to the 1:2 receptor/sugar association constant.

[12] As noted by Schneider and Yatsimirsky, under properly chosen conditions reliable association constants can be obtained even if the maximally observed shift changes are <0.03 ppm, see: H.-J. Schneider, A. Yatsimirsky, *Principles and Methods in Supramolecular Chemistry*, John Wiley & Sons, Chichester, **2000**.
[13] The naturally occurring oligosaccharides contain α-linked *N*-acetylneuraminic acid, see: a) H. Osborn, T. Khan, *Oligosaccharides: Their Synthesis and Biological Roles*, Oxford University Press, New York, **2000**; b) T. K. Lindhorst, *Essentials of Carbohydrate Chemistry and Biochemistry*, Wiley-VCH, Weinheim, Germany, **2000**; c) R. Schauer, *Angew. Chem.* **1973**, *85*, 128–140; *Angew. Chem. Int. Ed. Engl.* **1973**, *12*, 127–138.
Received: April 3, 2007
Published Online: May 29, 2007

