Carbohydrate Recognition

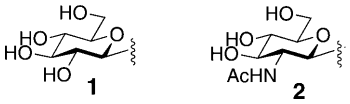
DOI: 10.1002/ange.200903104

A Synthetic Lectin for β-Glucosyl**

*Nicholas P. Barwell, Matthew P. Crump, and Anthony P. Davis**

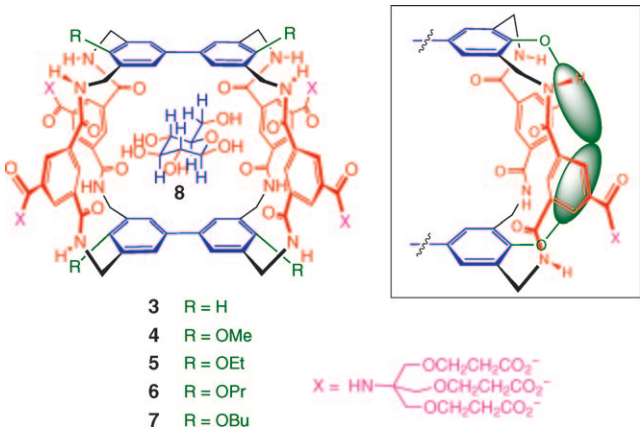
The measurement of blood glucose levels is a regular necessity for millions of diabetics. The standard enzyme-based methodology[1] is reliable and inexpensive, but has certain limitations. In particular, it is difficult to employ in long-term implantable sensors, due to stability problems and/or the need for consumable components.[2] Alternative approaches have therefore been widely studied.[2,3] Some involve carbohydrate binding proteins (lectins),[2,3a,d] while others employ boronic acid units which bind glucose through cyclic boronate formation.[3e–j] The use of "synthetic lectins" (receptors which bind through non-covalent interactions) has been discussed as a third possibility,[3e] but presents technical difficulties. Biomimetic carbohydrate recognition has proved challenging for supramolecular chemists,[4] and the effective binding of glucose in water is still an unsolved problem.

We have been studying a design for synthetic lectins targetting pyranose units with all-equatorial substitution patterns (β-glucosyl **1**, β-GlcNAc **2**, etc.).[5] As illustrated in



Scheme 1, the combination of parallel biphenyl units with isophthalamide spacers in a macrotricyclic framework provides a binding site which complements both apolar and polar moieties on the carbohydrates. Our prototype **3** was found to bind glucose **8** in water with low affinity ($K_a = 9\,M^{-1}$),[5a] but was far more effective with β-GlcNAc **2** (K_a up to $10^3\,M^{-1}$).[5d] While this selectivity for **2** has potential applications, a preference for glucose could be even more valuable. We now report modifications to **3** which raise affinities for glucose, and enhance selectivities for β-glucosyl vs. other carbohydrates (including GlcNAc). The results suggest that, with further development, exploitation in blood glucose monitors is a real possibility.

To modify the binding properties of **3**, we sought changes which might yield useful effects without substantially changing the core structure. Appending substituents to the framework could alter selectivities and enhance binding by adding to the intermolecular interactions (either polar or non-polar).



Scheme 1. Structural formulae for receptors **3**–**7** enclosing glucose **8** (β anomer). Complementary apolar groups are shown in blue, polar units in red. Inset: one end of the structure showing the expected positioning for the alkoxy substituents.

Groups added to the isophthalamide spacers would either occupy the cavity or be directed away from the substrate, while additions to the biphenyl 2/2′ positions would severely distort the conformation. However, substitution in the biphenyl 4 and 4′ positions appeared more promising. Such groups need have little effect on the core conformation, but could fold inward between the isophthalamides to reinforce the cavity walls and make contact with a bound carbohydrate. As a first trial of this strategy we chose to investigate alkoxy 4,4′ substituents, as in **4**–**7** (Scheme 1). For these apolar groups, the inward-directed orientation should be favoured by hydrophobic interactions with the isophthalamide spacers and, in some cases, with each other.

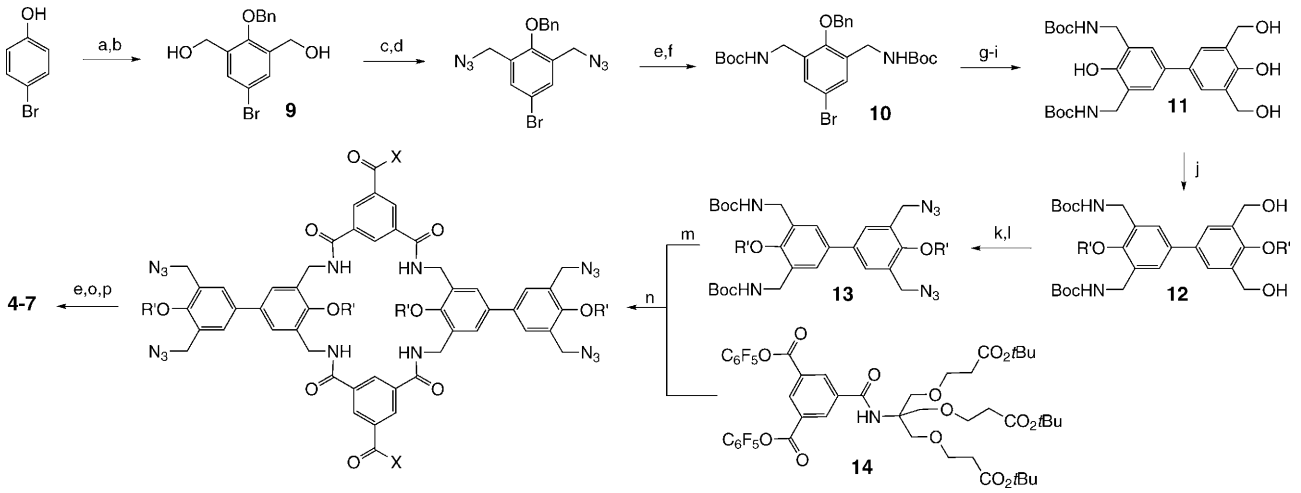
Receptors **4**–**7** were synthesized as shown in Scheme 2. Hydroxymethylation of *p*-bromophenol, followed by phenolic O-benzylation, gave diol **9**. Part of this material was converted to the Boc-protected diamine **10**, then coupled to unchanged **9** using Suzuki–Miyaura methodology. The resulting biphenyl was de-O-benzylated to give biphenol **11**, then re-alkylated to give intermediates **12**. Conversion of the benzylic hydroxy groups to azides gave biphenyls **13** with four masked amino groups, suitable for stepwise cyclisations with spacer unit **14** as described previously.[5] After conversion of the peripheral *tert*-butyl ester groups to carboxylates, the macrotricycles were freely soluble in water and gave well-resolved NMR spectra.

The binding properties of **4**–**7** were studied through ^1H NMR titrations in D_2O.[6] As observed for **3**,[5a,c] addition of most carbohydrates caused movements of signals due to receptor aromatic protons. The motions were consistent with 1:1 binding, and were analysed by non-linear curve-fitting to obtain association constants K_a. In most cases, two signals were analysed and gave closely similar results. *N*-Acetylglu-

[*] N. P. Barwell, Dr. M. P. Crump, Prof. A. P. Davis
School of Chemistry, University of Bristol
Cantock's Close, Bristol, BS8 1TS (UK)
Fax: (+44) 117-929-8611
E-mail: anthony.davis@bristol.ac.uk

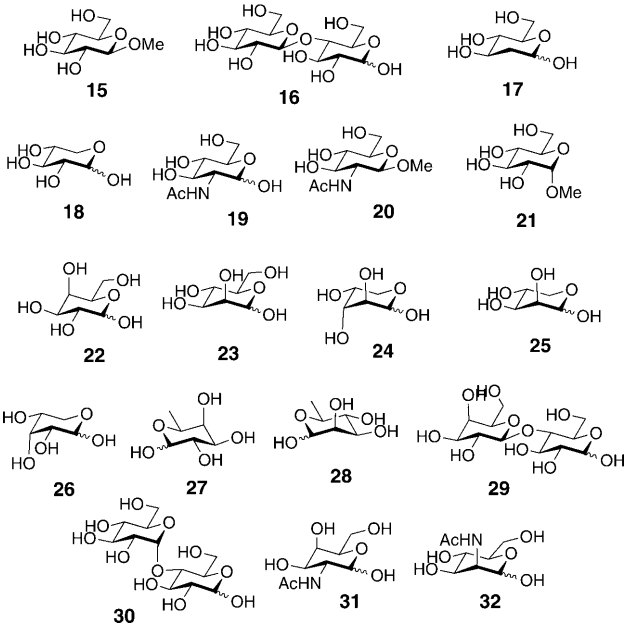
[**] This work was supported by the EPSRC (EP/D060192/1).

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/anie.200903104.



Scheme 2. Synthesis of receptors **4**–**7**: a) KOH, aq. formaldehyde, 48 h, 40 °C; b) K_2CO_3, BnBr, acetone, reflux, 4 h; c) $SOCl_2$, DCM; d) NaN_3, DMF, 60 °C; e) PPh_3, THF, 60 °C, then H_2O, 60 °C; f) Boc_2O, DIPEA, THF; g) bispinacolato diboron, [Pd(dppf)Cl_2] (3 mol%), KOAc (4 equiv), DMF, 80 °C; h) **9**, [Pd(dppf)Cl_2] (5 mol%), Na_2CO_3 aq., DMSO, 80 °C; i) Pd/C, H_2, THF/methanol saturated with NH_3 (1:1); j) R′I, Cs_2CO_3, DMF, RT (R′ = Me, Et) or RBr, Cs_2CO_3, MeCN, 60 °C (R′ = Pr, Bu); k) PPh_3, CBr_4, THF, 0 °C to RT; l) NaN_3, DMF, 60 °C; m) TFA, DCM, 0 °C; n) DIPEA, THF; o) **15**, DIPEA, THF; p) TFA, DCM, 0 °C, then NaOH aq. Bn = benzyl, DCM = dichloromethane, DMF = *N*,*N*′-dimethylformamide, Boc = *tert*-butoxycarbonyl, DIPEA = diisopropylethylamine, dppf = 1,1′-bis(diphenylphosphanyl)ferrocene, DMSO = dimethylsulfoxide, TFA = trifluoroacetic acid.

cosamine (GlcNAc, **19**) and the corresponding β-methyl glycoside (GlcNAcβ-OMe, **20**) were bound with slow exchange on the NMR timescale, allowing determination of



K_a by integration of signals from bound and unbound receptor. The data are summarised in Table 1, along with published results for **3**.[7] In several cases the results were checked using isothermal titration calorimetry (ITC) and/or induced circular dichroism (ICD). In general, good agreement was observed between the different techniques.[6]

Focussing first on receptor **4**, it seems that the methoxy substituents produce a general increase in affinities. The only clear exception is GlcNAc (**19**), for which a small drop in K_a is observed between **3** and **4**. Binding to glucose is enhanced by a factor of ≈4. When the alkoxy group is then lengthened through ethyl to propyl (**4**→**5**→**6**), further increases are observed for the β-glucosyl substrates **8**, **15** and **16**, and for the closely related all-equatorial substrates **17** and **18**. By

Table 1: Association constants K_a for receptors **3**–**7** with carbohydrate substrates in aqueous solution.[a]

Carbohydrate	K_a [M^{-1}]				
	3	**4**	**5**	**6**	**7**
D-glucose (**8**)	9	35[b]	41	60[b]	47
methyl β-D-glucoside (**15**)	27	70[b]	81	130[b]	82
D-cellobiose (**16**)	17	35		71	
2-deoxy-D-glucose (**17**)	7.2	18		29[b]	
D-xylose (**18**)	5	14		17[b]	
N-acetyl-D-glucosamine (**19**)	56	41	10	7	3
GlcNAcβ-OMe (**20**)	630	730[b]		43	
methyl α-D-glucoside (**21**)	7	11	11	15[b]	8
D-galactose (**22**)	2	4	3	3	2
D-mannose (**23**)	≈0	2	≈0	≈0	2
D-arabinose (**24**)	2	7		4	
D-lyxose (**25**)	≈0	≈0		≈0	
D-ribose (**26**)	3	6		6	
L-fucose **27**)	2	3		3	
L-rhamnose (**28**)	≈0	≈0		≈0	
D-lactose (**29**)	≈0	4		8	
D-maltose (**30**)	≈0	≈0		≈0	
N-acetyl-D-galactosamine (**31**)	2	3		3	
N-acetyl-D-mannosamine (**32**)	2	4		4	

[a] Measured by ^{1}H NMR titration in D_2O at 296 K. Results for **3** are repeated from refs. [5a,d]. Values denoted ≈0 were too small for reliable analysis. Errors were estimated at ≤5% for most of the systems which showed fast exchange on the NMR timescale, and ≤10% for systems showing slow exchange. For further details, including individual error estimates, see Supporting Information. [b] Corroborative data obtained using ICD or ITC. For details, see Supporting Information.

contrast, binding to GlcNAc (**19**) and GlcNAcβ-OMe (**20**) is substantially decreased. For most other substrates, affinities remain static or decrease slightly. Finally, the addition of a further methylene group (**6**→**7**) yields a decrease in affinities for most substrates.

Macrotricycle **6** emerges from these studies as a promising receptor for β-glucosyl units, and for glucose itself. When compared to the 4,4′-unsubstituted receptor **3**, it binds glucose more strongly (by a factor of ca. 6) and considerably more selectively. Receptor **6** shows glucose/galactose selectivity of 20:1 and glucose/GlcNAc selectivity of 9:1. The corresponding values for **3** were 4.5:1 and 1:6, respectively. Remarkably, receptor **6** seems to be more glucose-selective than the readily available[8] lectins used for this substrate, i.e. concanavalin A, lens culinaris agglutinin and pisum sativum agglutinin. These proteins bind mannose as well as glucose, and indeed prefer the former.[9] The affinity of **6** for glucose, at $60\,\text{M}^{-1}$, is still lower than those of the natural lectins (540, 230 and $380\,\text{M}^{-1}$, respectively[8]). However, it may be sufficient for the main projected application, i.e. blood glucose monitoring. Physiological glucose levels range from 2 mM (low) through 5 mM (normal) to 10 mM and above (high). When **6** is exposed to these concentrations, the fraction of receptor bound (f) may be calculated as 0.11, 0.23 and 0.38, respectively (assuming the glucose is present in large excess). In principle, the variation of f across this range should not be difficult to measure and translate into a glucose concentration.

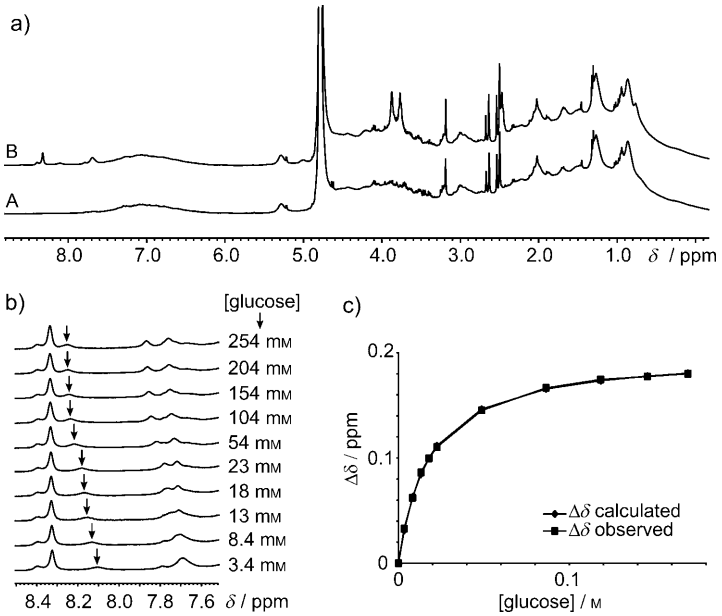


Figure 1. Glucose binding by receptor **6** in blood plasma reconstituted with D_2O. a) ^{1}H NMR of plasma sample without (A) and with (B) added **6** (1 mM). New signals due to **6** are observed in the aromatic region of the spectrum and at $\delta \approx 3.8$ and 2.5 ppm. b) Changes to the aromatic region of the spectrum as the glucose concentration is increased. The peak analysed to give K_a is marked by vertical arrows. c) Observed and calculated binding curves. $K_a = 58\,\text{M}^{-1}$ (cf. $60\,\text{M}^{-1}$ in pure D_2O).

If receptor **6** is to be used for physiological glucose monitoring, it must be able to operate in the presence of the other blood constituents. In a preliminary test, we obtained lyophilised human blood plasma from Sigma–Aldrich and reconstituted it with D_2O. The concentration of glucose was measured at 3.4 mM using a commercial Accucheck Aviva blood glucose meter. NMR spectra were recorded before and after addition of **6** (1 mM) (Figure 1). Further glucose was added incrementally, with monitoring by NMR. Although the signals due to **6** were broadened relative to those in pure D_2O, their movements were essentially unaffected by the plasma contents. Analysis of the data as for a normal binding study gave an excellent fit (Figure 1c) with $K_a = 58\,\text{M}^{-1}$ ($\pm 3\,\%$), almost identical to that in pure D_2O. The results suggest that normal blood does not contain any soluble components that interfere with glucose binding. Of course, NMR is not a viable method for glucose monitoring, but if appropriate transduction methods can be developed there is clear potential for applying this system in practical sensors.

To rationalise the binding properties of **4**–**7**, we first employed NMR and molecular modelling to elucidate, as far as possible, the structures of the complexes. Two systems were studied in detail; methoxy receptor **4** + GlcNAcβ-OMe (**20**), and propoxy receptor **6** + glucose (**8**). In the case of **4**·**20**, we were guided by the previously published NMR structure of **3**·**20**.[5d] Like **3**·**20**, complex **4**·**20** was strongly bound ($K_a = 730\,\text{M}^{-1}$) and showed slow exchange on the NMR timescale. This allowed the assignment of chemical shifts for most of the protons, notable exceptions being the receptor OMe signals. NOESY spectroscopy then yielded a pattern of intermolecular correlations which was almost identical to that for **3**·**20**, implying a very similar structure. On this basis, a model for **4**·**20** was constructed by adding four methoxy groups to the NMR structure of **3**·**20**, then optimising their orientation using Monte Carlo molecular mechanics (MCMM).[10] The calculations suggested a small preference for inward-directed methoxy groups, reflected in the global minimum shown in Figure 2.

In the case of **6**·**8**, the complex was weaker ($K_a = 60\,\text{M}^{-1}$) and in fast exchange, so less information was available. However, NOESY contacts between axial carbohydrate and biphenyl CH groups were clearly discernable, supporting the expected binding geometry. All intermolecular cross-peaks

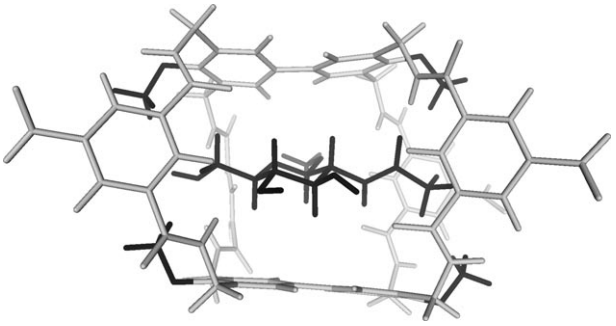


Figure 2. Model of **4**·**20**, based on the NMR structure of **3**·**20** supplemented by MCMM. The carbohydrate guest and (inward-directed) receptor 4,4′-methoxy groups are highlighted in black. The water-solubilising tricarboxylate groups in **4** are omitted for clarity.

correlated with β-glucose, suggesting that only this (the all-equatorial) anomer was bound. Correlations were also observed between propoxy and spacer protons within the receptor, implying that the propoxy groups are inward-directed as in Scheme 1. Modelling showed that receptor propoxy groups could contact each other across the narrow portal of the cavity, presumably a stabilising interaction in aqueous solution.

The structural evidence helps to explain one aspect of the binding results, i.e. the strongly variable glucose/GlcNAc selectivity. As illustrated in Figure 2, GlcNAc units seem to bind in a geometry which places the NHAc in a narrow portal of the cavity. If the receptor alkoxy groups are directed inwards, steric clashes are possible. As the alkoxy groups increase in size it is understandable that binding to GlcNAc is disfavoured. The other effects of the alkoxy groups may be rationalised in several ways. The groups might raise affinities by interfering with the hydration of the empty cavity and/or by making apolar contacts to carbohydrate CH groups (for example the 6-CH_2). Alternatively, they might help to stabilise the precise conformation required for binding. An electronic component is possible, but seems unlikely. In the general case, alkoxy groups can enhance CH–π interactions through raising electron density in aromatic rings.[11] However here the alkoxy group is flanked by CH_2 units and forced out of the plane of the aromatic ring, inhibiting π-donation.[12]

In conclusion, we have shown that the binding properties of our tricyclic monosaccharide receptors can be usefully modified through substituents in the biphenyl-based hydrophobic units. There is further scope for applying this strategy, for example by adding polar functional groups. The alkoxy substituents studied in this work have conferred selectivity for β-glucosyl, and have raised affinities for this unit. Binding to glucose is still quite weak, but potentially strong enough for blood glucose monitoring. In future we hope to realise this important application of synthetic lectins.

Received: June 9, 2009
Revised: July 15, 2009
Published online: September 8, 2009

Keywords: biomimetics · carbohydrates · molecular recognition · receptors · supramolecular chemistry

[1] J. Wang, *Chem. Rev.* **2008**, *108*, 814.
[2] S. M. Borisov, O. S. Wolfbeis, *Chem. Rev.* **2008**, *108*, 423.
[3] Leading references: a) J. S. Marvin, H. W. Hellinga, *J. Am. Chem. Soc.* **1998**, *120*, 7; b) R. Ballerstadt, J. S. Schultz, *Anal. Chem.* **2000**, *72*, 4185; c) K. M. Ye, J. S. Schultz, *Anal. Chem.* **2003**, *75*, 3451; d) S. Chinnayelka, M. J. McShane, *Anal. Chem.* **2005**, *77*, 5501; e) T. D. James, M. D. Phillips, S. Shinkai, *Boronic Acids in Saccharide Recognition*, RSC, Cambridge, **2006**; f) T. D. James, S. Shinkai, *Top. Curr. Chem.* **2002**, *218*, 159; g) L. Y. Wang, Y. D. Li, *Chem. Eur. J.* **2007**, *13*, 4203; h) G. K. Samoei, W. H. Wang, J. O. Escobedo, X. Y. Xu, H. J. Schneider, R. L. Cook, R. M. Strongin, *Angew. Chem.* **2006**, *118*, 5445; *Angew. Chem. Int. Ed.* **2006**, *45*, 5319; i) D. B. Cordes, S. Gamsey, B. Singaram, *Angew. Chem.* **2006**, *118*, 3913; *Angew. Chem. Int. Ed.* **2006**, *45*, 3829; j) X. P. Yang, M. C. Lee, F. Sartain, X. H. Pan, C. R. Lowe, *Chem. Eur. J.* **2006**, *12*, 8491.
[4] Reviews: A. P. Davis, R. S. Wareham, *Angew. Chem.* **1999**, *111*, 3160; *Angew. Chem. Int. Ed.* **1999**, *38*, 2978; A. Lützen in *Highlights in Bioorganic Chemistry: Methods and Applications* (Eds.: C. Schmuck, H. Wennemers), Wiley-VCH, Weinheim, **2004**, p. 109; A. P. Davis, T. D. James in *Functional Synthetic Receptors* (Eds.: T. Schrader, A. D. Hamilton), Wiley-VCH, Weinheim, **2005**, p. 45; M. Mazik, *Chem. Soc. Rev.* **2009**, *38*, 935. Recent contributions: P. B. Palde, P. C. Gareiss, B. L. Miller, *J. Am. Chem. Soc.* **2008**, *130*, 9566; C. He, Z. H. Lin, Z. He, C. Y. Duan, C. H. Xu, Z. M. Wang, C. H. Yan, *Angew. Chem.* **2008**, *120*, 891; *Angew. Chem. Int. Ed.* **2008**, *47*, 877; M. Waki, H. Abe, M. Inouye, *Angew. Chem.* **2007**, *119*, 3119; *Angew. Chem. Int. Ed.* **2007**, *46*, 3059; T. Reenberg, N. Nyberg, J. O. Duus, J. L. J. van Dongen, M. Meldal, *Eur. J. Org. Chem.* **2007**, 5003; C. Nativi, M. Cacciarini, O. Francesconi, A. Vacca, G. Moneti, A. Ienco, S. Roelens, *J. Am. Chem. Soc.* **2007**, *129*, 4377; M. Mazik, M. Kuschel, *Chem. Eur. J.* **2008**, *14*, 2405.
[5] a) E. Klein, M. P. Crump, A. P. Davis, *Angew. Chem.* **2005**, *117*, 302; *Angew. Chem. Int. Ed.* **2005**, *44*, 298; b) Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619; c) E. Klein, Y. Ferrand, N. P. Barwell, A. P. Davis, *Angew. Chem.* **2008**, *120*, 2733; *Angew. Chem. Int. Ed.* **2008**, *47*, 2693; d) Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jiménez-Barbero, C. Vicent, G.-J. Boons, S. Ingale, A. P. Davis, *Angew. Chem.* **2009**, *121*, 1807; *Angew. Chem. Int. Ed.* **2009**, *48*, 1775.
[6] For details see Supporting Information.
[7] Note that the binding constants reported to reducing sugars are averaged values, with contributions from both α and β anomers. It seems likely that the receptors will show anomeric selectivity, probably for β anomers (see the discussion of **6·8** later in the paper). In such cases, affinities for the favoured anomers will be higher than those given in Table 1. For further discussion, see Supporting Information.
[8] *Vector Laboratories Catalogue, 2007/2008*, Vector Laboratories, Burlingame, CA, **2007**, p. 81.
[9] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719.
[10] Performed using MacroModel 9.1 (Maestro 7.5 interface), employing the OPLS 2005 force field and water GB/SA solvation.
[11] M. Nishio, Y. Umezawa, M. Hirota, Y. Takeuchi, *Tetrahedron* **1995**, *51*, 8665.
[12] G. Baddeley, N. H. P. Smith, M. A. Vickars, *J. Chem. Soc.* **1956**, 2455.

