

Cite this: *Chem. Sci.*, 2017, **8**, 4056Received 9th December 2016
Accepted 28th March 2017

DOI: 10.1039/c6sc05399h

rsc.li/chemical-science

Introduction

Selective carbohydrate recognition is a long-standing interest of supramolecular chemistry.¹ The problem has clear relevance for biology, where carbohydrate-binding proteins (lectins) mediate a wide range of processes² including fertilization,^{2c-g} neuronal development,^{2d-h} hormonal activities,³ tumour metastasis,⁴ immune surveillance⁵ and inflammatory responses.⁶ At the same time, carbohydrates are exceptionally challenging targets, especially in the natural medium of water.^{1b} Typically they are coated with hydroxyl groups and are therefore highly hydrophilic. They are also “hydromimetic” in their resemblance to clusters of water molecules, and are therefore especially difficult to distinguish from competing solvent. Indeed, the affinities of natural lectins for their substrates are notoriously weak on the general scale of biomolecular interactions.⁷ If the challenge can be met, “synthetic lectins” have various potential applications. For example, they may serve as tools for biological research, as diagnostic and therapeutic agents,^{1c,8} and also as models for probing the basis of natural carbohydrate recognition.⁹

A characteristic feature of carbohydrates is their chirality. Indeed, they are often used to illustrate biological asymmetry, and figure strongly in the history of stereochemistry.¹⁰ Moreover, while natural sugars normally occur as single, specific enantiomers (in most cases *D*), the alternative (*L*)

enantiomers are found in some circumstances.¹¹ Accordingly, the enantioselective recognition of carbohydrates has attracted much interest.^{12,13} Success has been achieved for organic molecular receptors in organic solvents,¹² and also for boron-based systems in water.¹³ However, to date there have been no reports of enantioselective recognition by “synthetic lectins”, *i.e.* receptors operating in water through non-covalent interactions.

In previous work, we have developed a number of synthetic lectins which target the all-equatorial family of carbohydrates, *i.e.* glucose, *N*-acetylglucosamine (GlcNAc) and derivatives.^{1b,14} These substrates are important for various reasons. For example glucose is a major analyte in medicine and biotechnology.¹⁵ Selective glucose receptors have potential as components of glucose monitors, which could be used to aid the management of diabetes and also to follow fermentation and cell growth. Meanwhile GlcNAc, β -linked to serine and threonine, is a dynamic post-translational modification of proteins which is currently under intensive investigation.¹⁶ Selective GlcNAc receptors could be used in stains for modified proteins, and in other tools for β -GlcNAc research.

As illustrated in Fig. 1, our receptors possess a common architecture in which parallel aromatic surfaces (blue) are separated by rigid polar pillars (red). The aromatic surfaces make hydrophobic/CH- π contacts with substrate CH groups, while the pillars form hydrogen bonds to polar equatorial substituents. In line with the cartoon in Fig. 1, we have named these molecules the “temple” family of carbohydrate receptors.^{1b}

Thus far, all variants of the temple design have been achiral, with at least one plane of symmetry. However, as illustrated in Fig. 2a, it is possible to generate asymmetric versions by employing roof and floor units of different symmetries. In

School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK
E-mail: t.j.mooibroek@uva.nl; Anthony.Davis@bristol.ac.uk

† Electronic supplementary information (ESI) available: Experimental details for synthesis, characterisation and binding experiments; NMR spectra and binding analyses. See DOI: [10.1039/c6sc05399h](https://doi.org/10.1039/c6sc05399h)

‡ Present address: Van't Hoff Institute for Molecular Sciences, Universiteit van Amsterdam, Science Park 904, 1098 XH Amsterdam, The Netherlands.



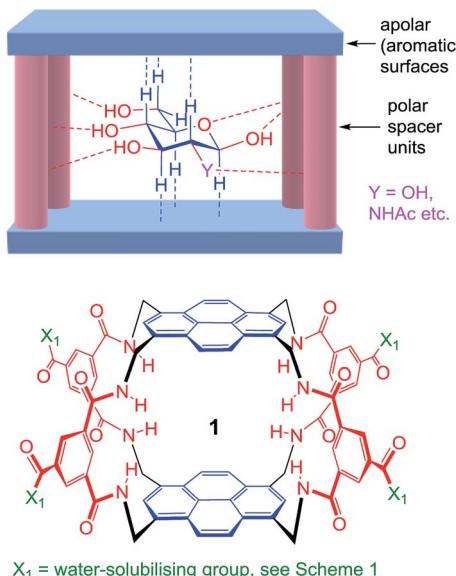


Fig. 1 General approach to synthetic lectins for all-equatorial carbohydrates (β -glucose, β -GlcNAc etc.), with recent achiral example 1 (see ref. 14a).

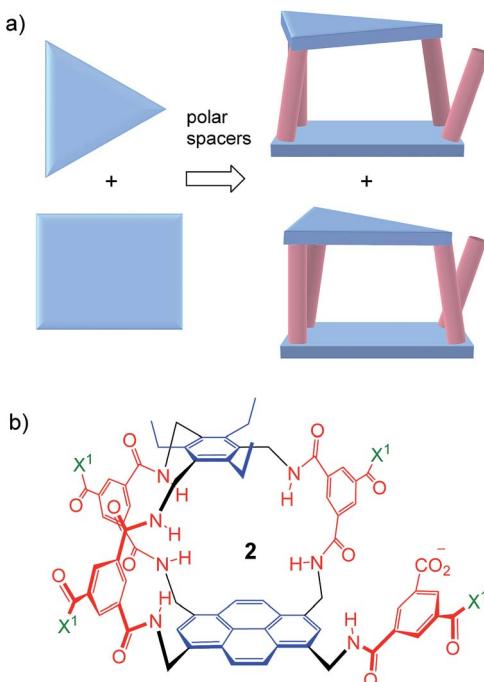


Fig. 2 (a) Chiral receptor architecture derived by combining C_3 and D_2 roof/floor components. (b) Specific design 2 featured in this paper. See Scheme 1 for the structure of water-solubilising group X^1 . The framework of 2 is planar chiral and the enantiomer shown is pS according to standard nomenclature (see Fig. S35, ESI†).

particular, a C_3 -symmetric roof can be combined with a square or rectangular (D_2) floor in just two ways, giving a pair of enantiomeric products. Acting as receptors, these molecules would surround their substrates with chiral cage frameworks which seem likely to favour enantioselectivity. The synthesis

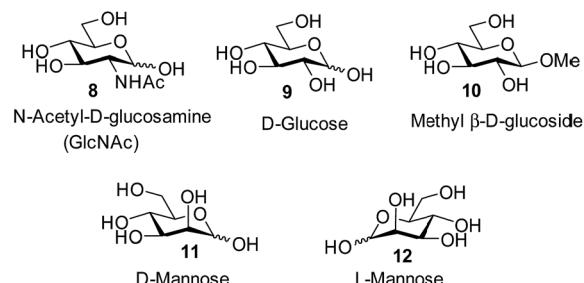
would leave an unreacted spacer unit, but this can be quenched in various ways and could be used to advantage.

Herein we report the realisation of this concept in the form of macrocycle 2 and its enantiomer (Fig. 2b). We show that these prototypes are capable of high levels of enantioselectivity in water, matching the discrimination shown by other systems in unnatural non-aqueous media. In addition, new records are set for binding affinity to simple monosaccharides, suggesting that chirality can improve complementarity in the design of synthetic carbohydrate receptors.

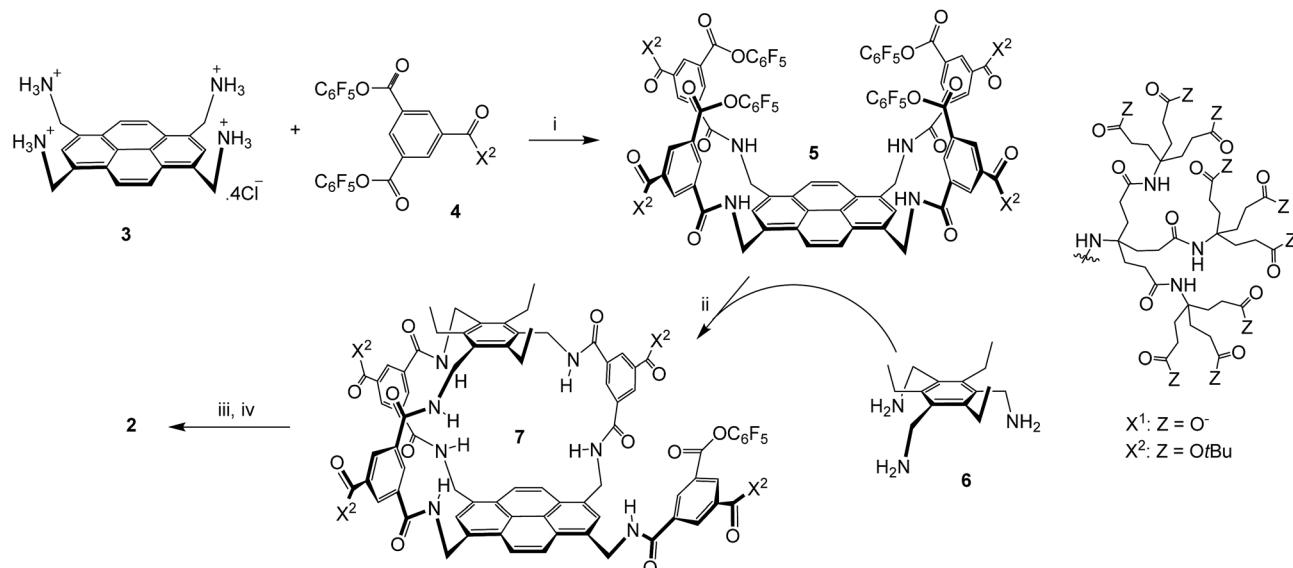
Results and discussion

The sequence employed to synthesise (\pm)-2 is shown in Scheme 1. Intermediate 5 was prepared from protonated tetra-amino pyrene 3 and isophthalate reagent 4, as reported previously for the synthesis of 1.^{14a} Active ester 5 was then combined with triamine 6 under high dilution to give (\pm)-7 in the remarkably good yield of 51%. Triamine 6 was chosen as roof component in the expectation that the ethyl groups would preorganise the amines and improve yields;¹⁷ this scaffold has previously been employed to construct carbohydrate receptors operating in organic solvents.¹⁸ The side-chain *t*-butyl esters were then cleaved with TFA, and the unreacted pentafluorophenyl ester was converted to carboxyl by basic hydrolysis. Finally the pH was adjusted to 7 using acidic ion exchange resin and NaOH, to give (\pm)-2 as a salt-free racemate.

Receptors 2 gave well-resolved ^1H NMR spectra in D_2O with minimal changes on dilution below 0.5 mM, implying that they do not self-aggregate in this concentration range (Fig. S13 and S24†). The spectra were more complex than those of previous synthetic lectins, reflecting the asymmetry of the framework. A full assignment of signals due to framework CH protons could be made using 2D NMR spectroscopy (Fig. S14–S23†).



Attempts to separate the enantiomers of 2 were unsuccessful,¹⁹ but the binding properties of both receptors could be studied through ^1H NMR studies on the racemate. Addition of carbohydrates 8–12 to (\pm)-2 caused substantial movements of receptor signals, many splitting as expected for the formation of diastereomeric complexes (one for each enantiomer of 2). The inward-directed spacer CH signals s6b–d (see Fig. 2) were relatively shielded in free 2 and showed especially large changes during the titrations. The titration with GlcNAc (8) was particularly informative (Fig. 3 and S25†). In this case signals due to spacer protons s6d (see Fig. 3) could be followed throughout the titration for both receptor enantiomers. Both sets of data gave excellent fits to a 1 : 1 binding model, yielding binding



Scheme 1 Synthesis of (\pm) -2. (i) THF/water, **4** (8 equivalents), $\text{EtN}(\text{iPr})_2$, 47%; see ref. 14a. (ii) THF, $\text{EtN}(\text{iPr})_2$, **[5]** and **[6]** ≤ 0.11 mM, 51%. (iii) TFA, DCM. (iv) NaOH , H_2O , then Amberlyst 15 hydrogen form, then NaOH (to $\text{pH} = 7$).

constants K_a of 1280 M^{-1} and 81 M^{-1} for the diastereomeric complexes (Fig. 4).²⁰

These values are significant in two respects. Firstly, the enantioselectivity of $16 : 1$ matches the highest previously measured for monosaccharide binding by synthetic receptors in any medium.^{12i,k} Secondly, the affinity for the more tightly-bound complex is the highest yet observed for this substrate, breaking the record of 520 M^{-1} previously held by the bis-pyrenyl system **13** (the “staggered” regioisomer of **1**).^{14a} Indeed, it appears to be the highest for biomimetic recognition of any underivatised, uncharged²¹ monosaccharide. It also compares well with the binding constant of 410 M^{-1} measured for the lectin Wheat Germ Agglutinin (WGA) for the same

substrate.²² Notably, the high affinity of (one enantiomer of) **2** for **8** is achieved without the benefit of statistical factors available to receptors of higher symmetry. For example, in D_{2h} receptors such as **1** there are four equivalent orientations for a carbohydrate within the binding site, each contributing to the binding affinity.²³ For chiral receptor **2** there is no degeneracy. The additional binding energy from improved complementarity must therefore be sufficient to compensate for this loss.

Although the receptors **2** could not be resolved, the separation of signals due to the diastereomeric complexes presented an opportunity to assign the structure of the stronger binding enantiomer. NMR studies were performed on a mixture of (\pm) -**2** and **8**, chosen so that receptor signals were well-separated and

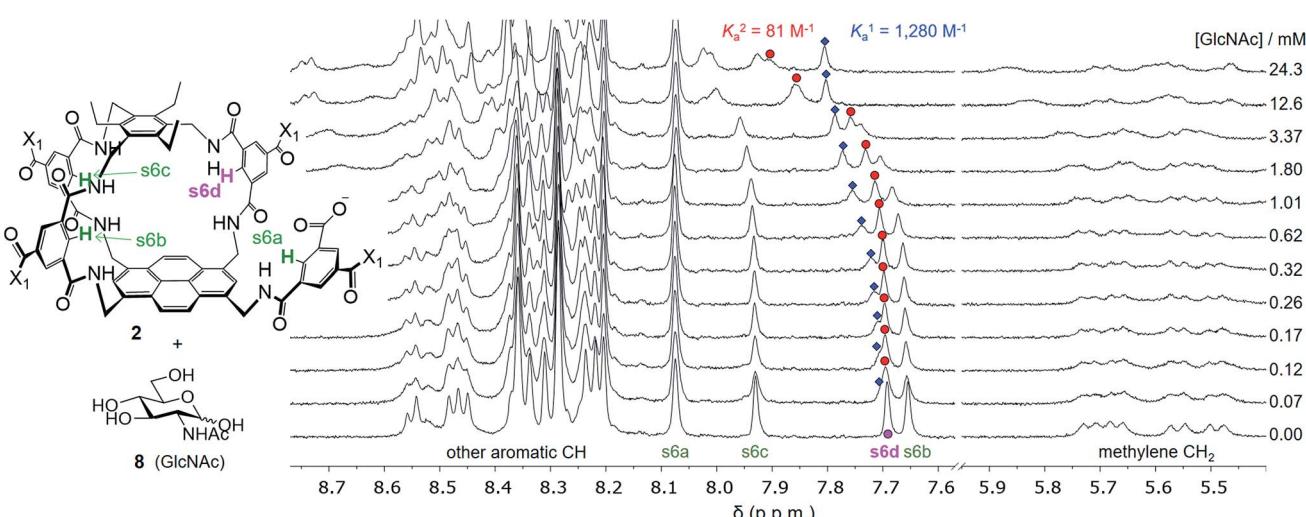


Fig. 3 Selected partial spectra for a ^1H NMR binding study of receptors (\pm) -**2** (0.15 mM each) with d-GlcNAc **8** in D_2O . The labelling system used for **2** and a full NMR assignment are detailed in the ESI.† Signals due to protons **s6a–d** appear in the region 7.6–8.1 ppm and are readily observed during the titration. In particular, the signal due to **s6d** splits into two peaks which can be followed throughout.



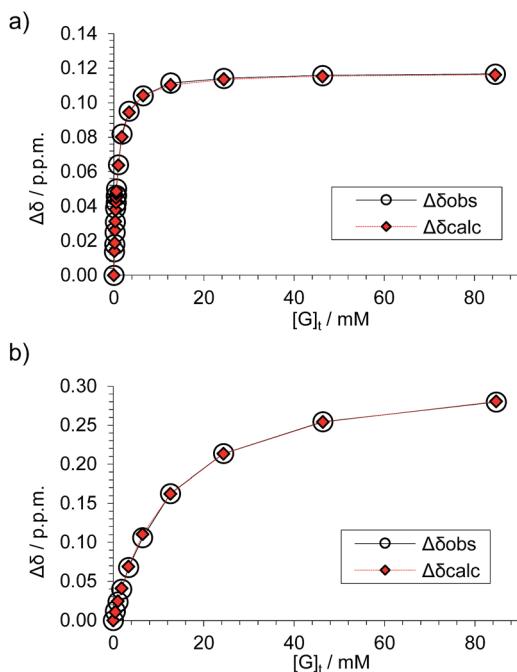
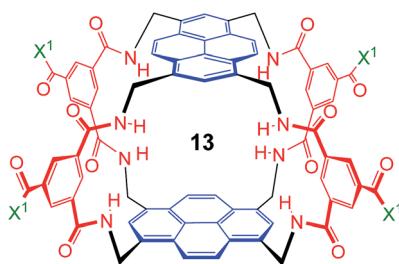


Fig. 4 Data analyses for the ^1H NMR titration of (\pm) -2 with D-GlcNAc 8 (see Fig. 3) assuming a 1 : 1 binding model. The signal due to proton s6d was followed for both diastereomeric complexes. (a) Analysis of peaks marked by blue diamonds in Fig. 3 spectra; $K_a = 1280 \text{ M}^{-1} \pm 2\%$, limiting $\Delta\delta = 0.117 \text{ ppm}$. (b) Analysis of peaks marked by red circles in Fig. 3 spectra, $K_a = 81 \text{ M}^{-1} \pm 5\%$, limiting $\Delta\delta = 0.331 \text{ ppm}$.

that the more strongly-bound complex was present at relatively high concentrations. NOESY and TOCSY spectra allowed a nearly complete assignment of signals due to the more strongly bound enantiomer of 2. Intermolecular NOE signals indicated that the carbohydrate CH_2OH was positioned in the smallest of the three portals of the receptor (between s6b and s6c), with the α -face directed towards the pyrene. Further connections led to the conclusion that stronger-binding enantiomer was pR-2 (*i.e.* the antipode of the structure in Fig. 2b). Details of the assignment and spectra are given in the ESI (see Fig. S35–S55†).



The other all-equatorial carbohydrates **9** and **10** proved more difficult to study, although enantiodiscrimination was apparent for both substrates. In the case of glucose **9**, addition to (\pm) -2 caused receptor ^1H NMR signals to move and split as expected for diastereomeric complex formation, but those due to one diastereomer then broadened (presumably due to slow exchange) and could not be followed (Fig. S27†). A signal from

the second complex could be analysed, yielding a binding constant of 250 M^{-1} (Fig. S28†).²⁰ Again, this breaks previous records for biomimetic recognition of this important substrate; earlier studies have yielded values up to 190 M^{-1} , again for **13**.^{14a} Moreover, while the signals for the second diastereomeric complex could not be analysed, initial movements occurred early during the titration suggesting that the affinity may be still higher.¹⁷ Similar changes were observed when methyl β -D-glucoside **10** was added to **2**. In this case two signals could be followed to the end of the titration, but both appeared to belong to one diastereomeric complex and gave $K_a = 250 \text{ M}^{-1}$ on analysis (Fig. S29 and S30†).

The non-target substrate mannose was studied both as single enantiomer **12** and as a racemate, formed by mixing **11** and **12**. For the single enantiomer, the formation of both diastereomeric complexes could be followed, with $K_a = 7.6$ and 2.3 M^{-1} respectively (Fig. S31 and S32†). For racemic mannose + racemic **2** a single set of signals was observed, as expected for two racemic diastereomeric complexes equilibrating rapidly on the ^1H NMR timescale. Analysis gave an apparent K_a of 4.8 M^{-1} , close to the average of the separately-measured values (Fig. S33 and S34†). The selectivity for all-equatorial substrates **8–10** vs. mannose **11/12** is consistent with previous work on the “temple” family of receptors.^{1b,14}

Conclusions

In conclusion we have demonstrated a straightforward, high-yielding method for introducing chirality into the framework of a synthetic lectin. The approach has yielded the first synthetic receptors capable of enantioselective carbohydrate recognition in water through the application of purely non-covalent interactions. Moreover the level of enantioselectivity matches that achieved by earlier systems in non-aqueous (and therefore unnatural) media. For matched enantiomers, the receptors are also exceptionally powerful. The highest affinity measured, 1280 M^{-1} for GlcNAc, is more than twice that for any earlier receptor with the same substrate, and three times that for the lectin WGA. This success may reflect the principle that a chiral substrate requires a chiral receptor for ideal complementarity, even though low symmetry incurs an entropic penalty.²³

The strategy employed to make **2** also allows generalization to other “roofs”, through use of different triamines in the cyclisation step. All such systems would require resolution, and experience with **2** suggests that the problem may not be trivial. However for receptors with potential applications solutions could surely be found. Further elaboration should also be feasible through nucleophilic attack on the unreacted active ester groups in intermediates such as **7**. For example, this would provide a ready means of attachment to a polymer support, to generate a carbohydrate-selective stationary phase. Alternatively, the binding site could be supplemented by attaching peptidic units, potentially in combinatorial format.²⁴ This approach thus seems promising for the development of improved synthetic lectins for β -glucosyl, β -GlcNAc and related “all-equatorial” carbohydrates.



Acknowledgements

This research was supported by the Engineering and Physical Sciences Research Council through grant number EP/I028501/1, and the Wellcome Trust grant number 099185/Z/12/Z. The Secretaría General de Universidades, Investigación y Tecnología (Junta de Andalucía) is acknowledged for a Talentia Scholarship to P. R. The Netherlands Organisation for Scientific Research (NWO) is thanked for a VIDI grant (723.015.006) to T. J. M. Support is acknowledged from BrisSynBio, a BBSRC/EPSRC Synthetic Biology Research Centre, grant no. BB/L01386X/1, and from the Novo Nordisk STAR programme (Fellowship to M. W.).

Notes and references

- (a) A. P. Davis and R. S. Wareham, *Angew. Chem., Int. Ed.*, 1999, **38**, 2978–2996; (b) A. P. Davis, *Org. Biomol. Chem.*, 2009, **7**, 3629–3638; (c) S. Jin, Y. F. Cheng, S. Reid, M. Y. Li and B. H. Wang, *Med. Res. Rev.*, 2010, **30**, 171–257; (d) M. Mazik, *RSC Adv.*, 2012, **2**, 2630–2642; (e) Y. Nakagawa and Y. Ito, *Trends Glycosci. Glycotechnol.*, 2012, **24**, 1–12; (f) C. E. Miron and A. Petitjean, *ChemBioChem*, 2015, **16**, 365–379.
- 2 See, for example: (a) *The Sugar Code-Fundamentals of Glycoscience*, ed. H.-J. Gabius, Wiley-Blackwell, Weinheim, 2009; (b) *Carbohydrate Recognition: Biological Problems, Methods and Applications*, ed. B. Wang and G.-J. Boons, Wiley, Hoboken, 2011; (c) A. B. Diekman, *Cell. Mol. Life Sci.*, 2003, **60**, 298–308; (d) D. J. Miller, M. B. Macek and B. D. Shur, *Nature*, 1992, **357**, 589–593; (e) W. J. Snell and J. M. White, *Cell*, 1996, **85**, 629–637; (f) R. Shalgi and T. Raz, *Histol. Histopathol.*, 1997, **12**, 813–822; (g) E. Rubinstein, A. Ziyyat, J. P. Wolf, F. Le Naour and C. Boucheix, *Semin. Cell Dev. Biol.*, 2006, **17**, 254–263; (h) H. E. Murrey and L. C. Hsieh-Wilson, *Chem. Rev.*, 2008, **108**, 1708–1731; (i) D. Solis, N. V. Bovin, A. P. Davis, J. Jiménez-Barbero, A. Romero, R. Roy, K. Smetana and H. J. Gabius, *Biochim. Biophys. Acta, Gen. Subj.*, 2015, **1850**, 186–235.
- 3 G. Caltabiano, M. Campillo, A. De Leener, G. Smits, G. Vassart, S. Costagliola and L. Pardo, *Cell. Mol. Life Sci.*, 2008, **65**, 2484–2492.
- 4 K. S. Lau and J. W. Dennis, *Glycobiology*, 2008, **18**, 750–760.
- 5 E. I. Buzas, B. Gyorgy, M. Pasztoi, I. Jelinek, A. Falus and H. J. Gabius, *Autoimmunity*, 2006, **39**, 691–704.
- 6 (a) R. A. Dwek, *Chem. Rev.*, 1996, **96**, 683–720; (b) P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek, *Science*, 2001, **291**, 2370–2376; (c) D. A. Calarese, C. N. Scanlan, M. B. Zwick, S. Deechongkit, Y. Mimura, R. Kunert, P. Zhu, M. R. Wormald, R. L. Stanfield, K. H. Roux, J. W. Kelly, P. M. Rudd, R. A. Dwek, H. Katinger, D. R. Burton and I. A. Wilson, *Science*, 2003, **300**, 2065–2071.
- 7 E. J. Toone, *Curr. Opin. Struct. Biol.*, 1994, **4**, 719–728; M. Ambrosi, N. R. Cameron and B. G. Davis, *Org. Biomol. Chem.*, 2005, **3**, 1593–1608.
- 8 Y. Nakagawa, T. Doi, T. Taketani, K. Takegoshi, Y. Igarashi and Y. Ito, *Chem.-Eur. J.*, 2013, **19**, 10516–10525.
- 9 E. Klein, Y. Ferrand, N. P. Barwell and A. P. Davis, *Angew. Chem., Int. Ed.*, 2008, **47**, 2693–2696; E. Klein, Y. Ferrand, N. P. Barwell and A. P. Davis, *Angew. Chem., Int. Ed.*, 2008, **47**, 2693–2696.
- 10 H. Kunz, *Angew. Chem., Int. Ed.*, 2002, **41**, 4439–4451.
- 11 M. Stefansson and M. Novotny, *J. Am. Chem. Soc.*, 1993, **115**, 11573–11580.
- 12 (a) K. M. Bhattacharai, R. P. Bonar-Law, A. P. Davis and B. A. Murray, *J. Chem. Soc., Chem. Commun.*, 1992, 752–754; (b) R. Liu and W. C. Still, *Tetrahedron Lett.*, 1993, **34**, 2573–2576; (c) R. P. Bonar-Law and J. K. M. Sanders, *J. Am. Chem. Soc.*, 1995, **117**, 259–271; (d) S. Anderson, U. Neidlein, V. Gramlich and F. Diederich, *Angew. Chem., Int. Ed. Engl.*, 1995, **34**, 1596–1600; (e) J. Cuntze and F. Diederich, *Helv. Chim. Acta*, 1997, **80**, 897–911; (f) G. Das and A. D. Hamilton, *Tetrahedron Lett.*, 1997, **38**, 3675–3678; (g) T. Ishi-i, M. A. Mateos-Timoneda, P. Timmerman, M. Crego-Calama, D. N. Reinhoudt and S. Shinkai, *Angew. Chem., Int. Ed.*, 2003, **42**, 2300–2305; (h) C. Li, G. T. Wang, H. P. Yi, X. K. Jiang, Z. T. Li and R. X. Wang, *Org. Lett.*, 2007, **9**, 1797–1800; (i) A. Arda, C. Venturi, C. Nativi, O. Francesconi, G. Gabrielli, F. J. Canada, J. Jiménez-Barbero and S. Roelens, *Chem.-Eur. J.*, 2010, **16**, 414–418; (j) F. Fernandez-Trillo, E. Fernandez-Megia and R. Riguera, *J. Org. Chem.*, 2010, **75**, 3878–3881; (k) O. Perraud, A. Martinez and J. P. Dutasta, *Chem. Commun.*, 2011, **47**, 5861–5863.
- 13 T. D. James, K. Sandanayake and S. Shinkai, *Nature*, 1995, **374**, 345–347; J. Z. Zhao, M. G. Davidson, M. F. Mahon, G. Kociok-Kohn and T. D. James, *J. Am. Chem. Soc.*, 2004, **126**, 16179–16186; J. Z. Zhao and T. D. James, *J. Mater. Chem.*, 2005, **15**, 2896–2901; G. Heinrichs, M. Schellentrager and S. Kubik, *Eur. J. Org. Chem.*, 2006, 4177–4186.
- 14 Recent examples: (a) P. Rios, T. S. Carter, T. J. Mooibroek, M. P. Crump, M. Lisbjerg, M. Pittelkow, N. T. Supekar, G.-J. Boons and A. P. Davis, *Angew. Chem., Int. Ed.*, 2016, **55**, 3387–3392; (b) T. J. Mooibroek, J. M. Casas-Solvas, R. L. Harniman, C. M. Renney, T. S. Carter, M. P. Crump and A. P. Davis, *Nat. Chem.*, 2016, **8**, 69–74; (c) H. Destecroix, C. M. Renney, T. J. Mooibroek, T. S. Carter, P. F. N. Stewart, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2015, **54**, 2057–2061; (d) C. Ke, H. Destecroix, M. P. Crump and A. P. Davis, *Nat. Chem.*, 2012, **4**, 718–723.
- 15 X. L. Sun and T. D. James, *Chem. Rev.*, 2015, **115**, 8001–8037; R. Gifford, *ChemPhysChem*, 2013, **14**, 2032–2044; M. S. Steiner, A. Duerkop and O. S. Wolfbeis, *Chem. Soc. Rev.*, 2011, **40**, 4805–4839.
- 16 E. J. Kim, *ChemBioChem*, 2015, **16**, 1397–1409; S. A. Yuzwa and D. J. Vodadlo, *Chem. Soc. Rev.*, 2014, **43**, 6839–6858; C. Slawson, R. J. Copeland and G. W. Hart, *Trends Biochem. Sci.*, 2010, **35**, 547–555.
- 17 G. Hennrich and E. V. Anslyn, *Chem.-Eur. J.*, 2002, **8**, 2218–2224; K. J. Wallace, R. Hanes, E. Anslyn, J. Morey, K. V. Kilway and J. Siegel, *Synthesis*, 2005, 2080–2083.



18 See, for example ref. 1d and: J. Lippe and M. Mazik, *J. Org. Chem.*, 2015, **80**, 1427–1439; C. Nativi, M. Cacciarini, O. Francesconi, G. Moneti and S. Roelens, *Org. Lett.*, 2007, **9**, 4685–4688; O. Francesconi, A. Ienco, G. Moneti, C. Nativi and S. Roelens, *Angew. Chem., Int. Ed.*, 2006, **45**, 6693–6696.

19 For details, see ESI.†

20 Binding constants for reducing sugars are weighted averages of those for the two anomers; see: N. P. Barwell, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 7673–7676. Given the geometry of the binding site, it is likely that the β -anomers are preferred, and that the affinities for these anomers are therefore higher than those reported.

21 Higher affinities have been observed for charged monosaccharides such as glucosamine, but this is expected for systems where electrostatic forces may be exploited. See ref. 14c and: (a) Y. Jang, R. Natarajan, Y. H. Ko and K. Kim, *Angew. Chem., Int. Ed.*, 2014, **53**, 1003–1007; (b) M. Rauschenberg, S. Bandaru, M. P. Waller and B. J. Ravoo, *Chem.-Eur. J.*, 2014, **20**, 2770–2782; (c) C. Geffert, M. Kuschel and M. Mazik, *J. Org. Chem.*, 2013, **78**, 292–300; (d) C. Schmuck and M. Schwegmann, *Org. Lett.*, 2005, **7**, 3517–3520.

22 See: Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jiménez-Barbero, C. Vicent, G. J. Boons, S. Ingale and A. P. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 1775–1779. WGA is the lectin commonly employed to bind the GlcNAc unit.

23 L. Challinor, E. Klein and A. P. Davis, *Synlett*, 2008, 2137–2141.

24 For previous carbohydrate receptors with aromatic and peptidic components, see for example: J. M. Benito and M. Meldal, *QSAR Comb. Sci.*, 2004, **23**, 117–129; T. Reenberg, N. Nyberg, J. O. Duus, J. L. J. van Dongen and M. Meldal, *Eur. J. Org. Chem.*, 2007, 5003–5009.

