ChemComm

RSC Publishing

COMMUNICATION

View Article Online
View Journal | View Issue

An accessible bicyclic architecture for synthetic lectins†

Joshua D. Howgego, Craig P. Butts, Matthew P. Crump and Anthony P. Davis*

Cite this: *Chem. Commun.*, 2013, **49**, 3110

Received 13th February 2013,
Accepted 4th March 2013

DOI: 10.1039/c3cc41190g

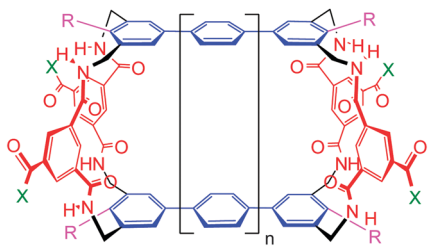
www.rsc.org/chemcomm

Bicyclic carbohydrate receptors are easier to synthesise than tri- or tetra-cyclic relatives, and are better adapted to bind monosaccharide residues with bulky appendages. Disaccharides containing β-glucosyl units are preferred substrates.

Biomimetic carbohydrate recognition remains a significant problem for supramolecular chemistry.[1] There is general interest in binding polar molecules in water, and saccharides are especially challenging due to their hydromimetic[2] character. Indeed, the affinities shown by carbohydrate-binding proteins (lectins) are notably weak, at least to monosaccharides.[3] Even so, carbohydrate recognition is an important natural phenomenon, playing roles in many biological processes.[4] Synthetic carbohydrate receptors can throw light on the principles which underlie lectin binding, could potentially complement these proteins as research tools for glycobiology, and may eventually find applications in medicine.[1c,5]

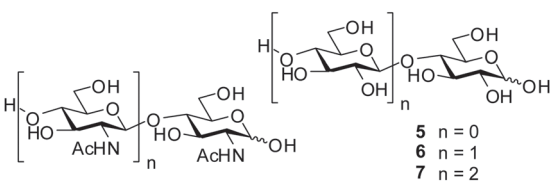
Among the more effective "synthetic lectins" are the macropolycyclic oligoamides developed in our group.[6] Targeted at carbohydrates with all-equatorial substitution patterns, these structures employ parallel biphenyl or terphenyl units to complement the axial carbohydrate CH groups, and polar spacers (mostly isophthalamides) to match polar substituents. The biphenyl-based systems (*e.g.* **1**[6a,c]) are effective for monosaccharides such as *N*-acetylglucosamine (GlcNAc) **3** and glucose **5**, while the terphenyl-based structures (*e.g.* **2**[6f]) are complementary to disaccharides such as *N*,*N*′-diacetylchitobiose **4** and cellobise **6**.

A feature common to these receptors has been the ability to surround the mono- or disaccharide target. This has conferred interesting size-selectivity; for example, **1** binds **3** but not **4**, while **2** shows a strong preference for **6** *vs.* **5** or **7**. However, the cost has been limited versatility. For applications such as glycoprofiling, it may be useful to bind just part of an oligosaccharide structure, for example a specific terminal unit. For this, a more open structure is required, so that unbound portions of the substrate can protrude from the cavity without steric clashes. Herein we describe a new bicyclic architecture for synthetic lectins, possessing larger portals than earlier tri- or tetracyclic systems. The system retains the ability to bind carbohydrates in water despite increased flexibility, and is particularly effective for sterically demanding substrates. Moreover, it is easier to synthesise than more highly-connected relatives, and thus has greater promise for variation and optimisation.



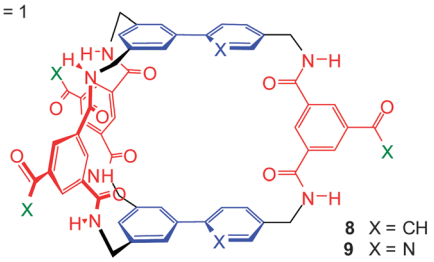
1 n = 0, R = H
2 n = 1, R = OMe

X = HN, $O(CH_2)_2CO_2^-$, $O(CH_2)_2CO_2^-$, $O(CH_2)_2CO_2^-$



3 n = 0
4 n = 1

5 n = 0
6 n = 1
7 n = 2



8 X = CH
9 X = N

The new family of receptors is exemplified by prototype **8**. Molecular modelling[7] of **8** predicts generally similar properties to **1**, *i.e.* an open binding site in water appropriately sized for a monosaccharide and apparently compatible with all-equatorial

School of Chemistry, University of Bristol, Cantock's Close, Bristol BS1 1TS, UK. E-mail: Anthony.Davis@bristol.ac.uk; Fax: +44 (0)117 9298611; Tel: +44 (0)117 9546334

† Electronic supplementary information (ESI) available: Experimental and modelling details, spectra for compound characterisation and binding studies, and fitting curves for binding analyses. See DOI: 10.1039/c3cc41190g

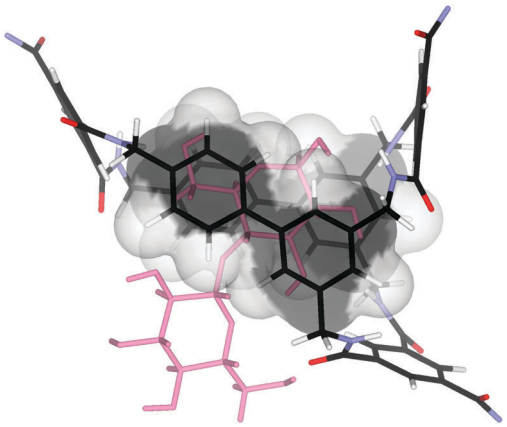
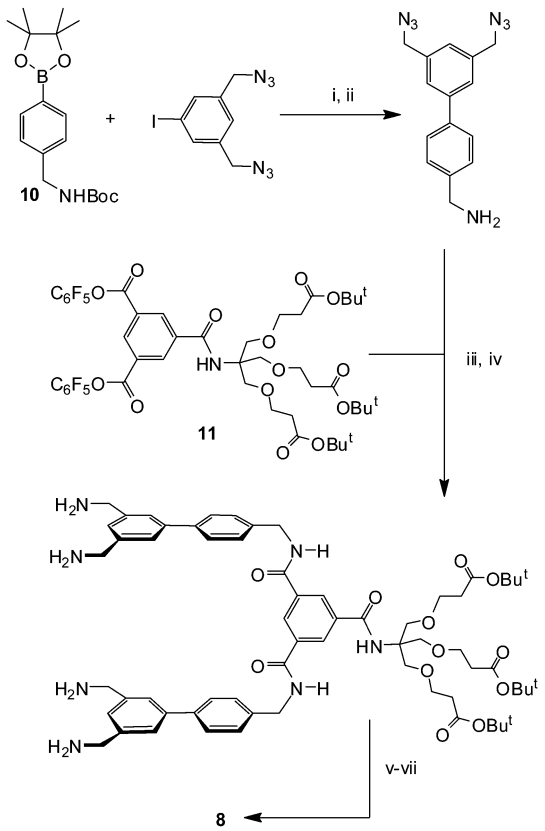


Fig. 1 Model of **8** binding β-D-cellobiose **6** (magenta), viewed from above the plane of the biphenyl units. Side chains are omitted, and the biphenyls are shown with transparent CPK surfaces.

substrates. However clear differences emerge with disaccharide substrates. As illustrated in Fig. 1, receptor **8** can bind a glucosyl residue in cellobiose **6** while leaving ample space for the unbound monosaccharide unit. In corresponding structures of **1** + **6**, the unbound glucosyl is hindered and twisted from its preferred orientation.[7] Similar disparities are observed for **8** and **1** binding maltose **13**.



Scheme 1 (i) $Pd(PPh_3)_3$, K_3PO_3, 80 °C, DMF, 4 h, 57%; (ii) TFA, CH_2Cl_2, 0 °C, 3 h, 76%; (iii) **11**, DIPEA, THF, 3 h, 90%; (iv) PPh_3, THF-H_2O, 60 °C, 12 h, 86%; (v) **11**, DIPEA, THF (high dilution), RT, 5 days, 19%; (vi) TFA, CH_2Cl_2, 0 °C, 3 h; (vii) NaOH (aq), ~100%.

Receptor **8** was prepared as shown in Scheme 1. Compared to the synthesis of tricyclic **1**, the sequence is considerably more straightforward. In particular, the biphenyls are initially joined by a simple linear coupling which proceeds in 90% yield. The corresponding step in the synthesis of **1** is a high-dilution macrocyclisation (~50% yield after preparative HPLC). ^{1}H NMR spectra of **8** (≤0.5 mM) in D_2O were well-resolved and invariant with concentration, suggesting that the system is monomeric under these conditions.

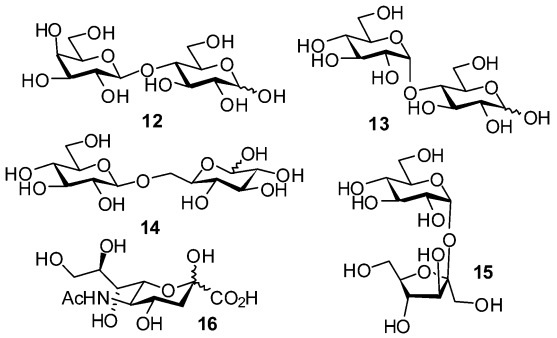
Binding of **8** to carbohydrates in D_2O was studied using ^{1}H NMR titrations.[8] In many cases, addition of the substrates caused significant changes to the spectrum of the receptor, including signal movements of up to 0.25 ppm and loss of some equivalences.[6a] Where feasible, the movements were analyzed according to a 1 : 1 binding model, giving mostly good fits with internal consistency. The resulting binding constants K_a are listed in Table 1, along with the corresponding data for **1**. For the hexose monosaccharides such as glucose and GlcNAc, the change from tri- to bicyclic architecture results in a lowering of affinities. This is perhaps unsurprising given the reduction in preorganisation (for example, the increased rotational freedom within the biphenyl units[9]). Interestingly, pentoses such as xylose are bound slightly more strongly, perhaps reflecting a decrease in the average size of the cavity. However, more significant is the increased affinity for disaccharide substrates. Bicyclic receptor **8** shows moderate-good affinities for cellobiose **6**, lactose **12**, maltose **13** and gentiobiose **14**. By contrast the tricyclic system **1** fails to bind lactose and maltose, and is three times less effective for cellobiose. The obvious interpretation is that, unlike **1**, receptor **8** can bind β-glucosyl residues in a wide variety of structural environments. Although the cavity cannot accept more than one monosaccharide residue (Fig. 1), interactions to parts of a second residue can presumably enhance binding. Although more forgiving than **1**, receptor **8** still discriminates between disaccharides, failing to bind sucrose **15** (which lacks β-glucosyl). Lastly the small but

Table 1 Association constants (K_a) for receptors **8** and **1** with carbohydrates[a]

	K_a [M^{-1}]	
Carbohydrate	**8**[a]	**1**[b]
D-Glucose **5**	4	9
Me β-D-glucoside	5	27
D-Mannose	~0	~0
D-Galactose	~0	2
D-Xylose	8	5
D-Lyxose	~0	~0
2-Deoxy-D-glucose	6	7
D-Ribose	7	3
N-Acetyl-D-glucosamine **3** (GlcNAc)	19	56
D-Cellobiose **6**	52	17
D-Lactose **12**	37	~0
D-Maltose **13**	14	~0
D-Gentiobiose **14**	22	n.m.
D-Sucrose **15**	~0	n.m.
N-Acetylneuraminic acid (Neu5Ac) **16**	7	~0

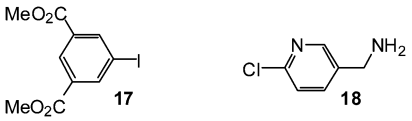
[a] Determined by ^{1}H NMR titration. Values denoted ~0 were too small for analysis. Errors were estimated at ≤±10% for most cases where $K_a \geq 10\ M^{-1}$. [b] These values have been previously reported (see ref. 6c). n.m. = not measured.

measurable affinity for Neu5Ac **16**, another bulky substrate, provides further evidence for receptor flexibility.



Following our normal practice we attempted to find alternative methods which could be used to support the data in Table 1. Unfortunately UV-vis/fluorescence titrations or isothermal titration calorimetry (ITC) could not be applied. However, addition of cellobiose and maltose to **8** caused measurable induced circular dichroism (ICD), which could be analysed to give K_a values consistent with those from NMR. In the case of Neu5Ac **16** we also employed a less usual method, that of ^{13}C NMR titration. It is well established that changes in ^{13}C chemical shifts can be used to detect binding, and there is precedent for the determination of binding constants in biochemical systems using signals from ^{13}C-enriched substrates.[10] On the other hand, as far as we can determine, the method has not been applied to unenriched synthetic supramolecular systems. In the present case we observed the ^{13}C NMR signals of **8** (2 mM in D_2O) on addition of NeuAc (6 aliquots) up to 0.22 M. Movements of up to 0.3 ppm were observed for several aromatic carbons. These were readily followed and analysed to yield K_a = 5.5 M^{-1}, close to the value achieved using ^{1}H NMR. The titration took several days, as it was performed with a proton-sensitive probe that was not ideal for the task.[11] However, the use of a ^{13}C-observe cryoprobe should reduce this time to a few hours. As such probes become more widely available, ^{13}C NMR titration could become a routine method for quantitative binding studies in supramolecular chemistry.

Finally, as mentioned earlier, an advantage of the bicyclic architecture is the ease of synthesis and thus the potential for variation. To illustrate the possibilities we prepared a second receptor **9** in which the *p*-phenylene unit was replaced by a 2,5-linked pyridylene. An interesting feature of this structure was the ability of the pyridyl nitrogen to reach inwards towards the substrate, potentially assisting binding through polar interactions. Receptor **9** was synthesized in ~7 steps from commercially available starting materials **17** and **18**.[7] Binding of **9** to carbohydrates was studied by ^{1}H NMR titrations as for **8**. Disappointingly the pyridyl nitrogen produced a uniformly negative effect on binding, resulting in a general lowering of affinities. Binding constants to cellobiose and maltose were 14 and 10 M^{-1} respectively, while the affinity for glucose was too weak to be measured.[7] Binding to galactose and mannose, with axial OH groups, was also negligible. The lower affinities may result from weaker CH–π interactions due to the electron-deficient nature of the pyridyl ring,[4c,6f] or reduced hydrophobic interactions[12] due to hydration of the pyridyl nitrogen. Unfortunately it seems that polar interactions to the pyridyl nitrogen are not significant, at least for the substrates tested.



In conclusion, we have shown that bicyclic designs are viable for biphenyl-based synthetic lectins, and can provide complementary binding properties. The systems discussed in this paper retain the ability to bind carbohydrates in water and (in the case of **8**) show improved performance with bulky substrates. The new receptors are more accessible than their predecessors, and therefore more tuneable. Their activity towards disaccharides suggests potential for this architecture in the binding of oligosaccharides, the key long-term objective for work on synthetic lectins.

This work was supported by the Engineering and Physical Sciences Research Council (grant number EP/D060192/1 and DTA studentship to JH). JH is grateful to J.-M. Casas-Solvas and C. Ke for useful discussions.

Notes and references

1 For reviews, see: (*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, *Chem. Soc. Rev.*, 2009, **38**, 935–956; (*e*) Y. Nakagawa and Y. Ito, *Trends Glycosci. Glycotechnol.*, 2012, **24**, 1–12. For recent examples, see: (*f*) C. Nativi, O. Francesconi, G. Gabrielli, I. De Simone, B. Turchetti, T. Mello, L. D. C. Mannelli, C. Ghelardini, P. Buzzini and S. Roelens, *Chem.–Eur. J.*, 2012, **18**, 5064–5072; (*g*) G. Fukuhara and Y. Inoue, *J. Am. Chem. Soc.*, 2011, **133**, 768–770; (*h*) A. Pal, M. Berube and D. G. Hall, *Angew. Chem., Int. Ed.*, 2010, **49**, 1492–1495; (*i*) M. Rauschenberg, S. Bomke, U. Karst and B. J. Ravoo, *Angew. Chem., Int. Ed.*, 2010, **49**, 7340–7435.
2 C. Ke, H. Destecroix, M. P. Crump and A. P. Davis, *Nat. Chem.*, 2012, **4**, 718–723.
3 E. J. Toone, *Curr. Opin. Struct. Biol.*, 1994, **4**, 719–728.
4 (*a*) *The Sugar Code – Fundamentals of Glycoscience*, ed. H.-J. Gabius, Wiley-Blackwell, Weinheim, 2009; (*b*) B. Wang and G.-J. Boons, *Carbohydrate Recognition: Biological Problems, Methods and Applications*, Wiley, Hoboken, 2011; (*c*) H. J. Gabius, S. Andre, J. Jiménez-Barbero, A. Romero and D. Solis, *Trends Biochem. Sci.*, 2011, **36**, 298–313.
5 J. Balzarini, *Nat. Rev. Microbiol.*, 2007, **5**, 583–597.
6 (*a*) E. Klein, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2005, **44**, 298–302; (*b*) Y. Ferrand, M. P. Crump and A. P. Davis, *Science*, 2007, **318**, 619–622; (*c*) 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, **49**, 1775–1779; (*d*) N. P. Barwell, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2009, **48**, 7673–7676; (*e*) N. P. Barwell and A. P. Davis, *J. Org. Chem.*, 2011, **76**, 6548–6557; (*f*) B. Sookcharoenpinyo, E. Klein, Y. Ferrand, D. B. Walker, P. R. Brotherhood, C. Ke, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2012, **51**, 4586–4590.
7 For further details see ESI†.
8 P. Thordarson, *Chem. Soc. Rev.*, 2011, **40**, 1305–1323.
9 In models of unbound **8**, the internal dihedral angles of the biphenyl groups are ~50–55° (see Fig. S39, ESI†). On binding to cellobiose these are reduced by ~15°, representing an energetic cost of ~4 kJ mol^{-1}.
10 K. J. Neurohr, N. M. Young, I. C. P. Smith and H. H. Mantsch, *Biochemistry*, 1981, **20**, 3499–3504; E. Berman, J. H. Brown, H. Lis and N. Sharon, *Eur. J. Biochem.*, 1985, **152**, 447–451; M. A. Kenyon and J. A. Hamilton, *J. Lipid Res.*, 1994, **35**, 458–467.
11 Varian VNMRS 600 with H{C,N} coldprobe. For ^{13}C measurements this instrument requires ×25 longer acquisition times than current generation direct-observe ^{13}C{H} cryogenically cooled probes.
12 E. Klein, Y. Ferrand, N. P. Barwell and A. P. Davis, *Angew. Chem., Int. Ed.*, 2008, **47**, 2693–2696.

