Selective disaccharide binding by a macrotetracyclic receptor†

Emmanuel Klein, Yann Ferrand, Elizabeth K. Auty and Anthony P. Davis*

Received (in Cambridge, UK) 2nd January 2007, Accepted 23rd February 2007
First published as an Advance Article on the web 13th March 2007
DOI: 10.1039/b618776e

A new carbohydrate receptor possesses a C_3-symmetric polar cavity capable of encapsulating disaccharides; binding to β-maltosyl is preferred, complementing previous systems which have favoured "all-equatorial" substrates.

Carbohydrates play key roles in biology, not only as energy sources and structural building blocks but also as information carriers. Monosaccharides can be oligomerised to give large numbers of structures, far more than for peptides or nucleic acids of similar molecular weight.[1] Nature exploits this resource, using oligosaccharides to label cell surfaces and proteins and thus mediate processes such as cell–cell recognition, protein trafficking and many aspects of the immune response.[2] This importance has fuelled many efforts to mimic biological carbohydrate recognition.[3] However, most work has targeted monosaccharides, instead of the oligosaccharides which are the major focus of biological interest. Of the published biomimetic[4] carbohydrate receptors, there are few which can span or encapsulate oligosaccharides, selecting for and among these larger substrates.[5]

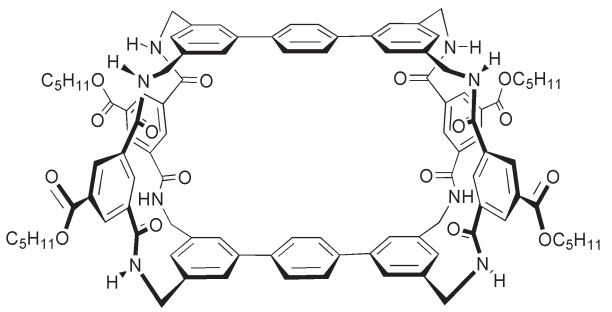
We have previously described the terphenyl-based macrotricycle **1**, a receptor for the "all-equatorial" β-cellobiosyl disaccharide **2a**.[5e] We now report a new *tetra*cyclic system **3**, with a "short, fat" architecture which complements the "long, thin" cavity of **1**. The change in shape is reflected in changes in selectivity, particularly in a preference for the curved structure of β-maltoside **4**.

The preparation of **3** exploited a side-reaction in our previous synthesis of monosaccharide receptor **8**.[6a] As shown in Scheme 1, diamine **5** was treated with diester **6** under high dilution to give a mixture of cyclo-oligomers. After Cbz removal, the major component **7a** was separated and converted into **8**. Further examination of the mixture **7** revealed that higher macrocycles could also be isolated and, in particular, that the trimeric **7b** could be obtained in yields of up to 20%.[7] This hexa-amine might seem unpromising as a cage precursor; at first sight, it may appear that reaction with **6** should yield a complex and intractable mixture of products. However, free rotation about the biphenyl central C–C bond simplifies the situation. Detailed analysis[8] suggests that, in principle, a high proportion of the hexa-amine could be converted to dodeca-amide **3**. In practice, treatment of **7b** with **6** at high dilution gave **3** in 15% yield after purification by HPLC.

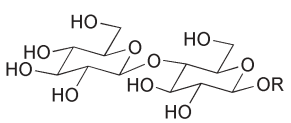
Macrotetracycle **3** was characterised by MALDI MS and NMR. ^{1}H NMR spectra in $CDCl_3$ were broad, but addition of CD_3OD (25%) caused the lines to sharpen. The resulting spectrum showed just 4 absorptions in the aromatic region (δ = 7.2–8.4), consistent with the expected D_{3h} time-average symmetry. Molecular modelling[9] revealed a range of conformations including open structures with roughly cylindrical cavities of length *ca.* 16 Å and internal diameters of *ca.* 7 Å. As illustrated in Fig. 1, these conformations were capable of encapsulating disaccharides such as methyl β-D-maltoside **4b** and methyl β-D-cellobioside **2b**, with formation of at least 6 intermolecular hydrogen bonds.

School of Chemistry, University of Bristol, Cantock's Close, Bristol, UK BS8 1TS. E-mail: Anthony.Davis@bristol.ac.uk; Fax: +44 117 9298611; Tel: +44 117 9546334

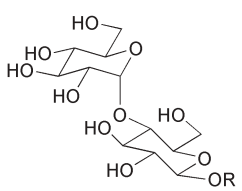
† Electronic supplementary information (ESI) available: Experimental details, including binding curves, and detailed discussion of the cyclisation giving **3**. See DOI: 10.1039/b618776e



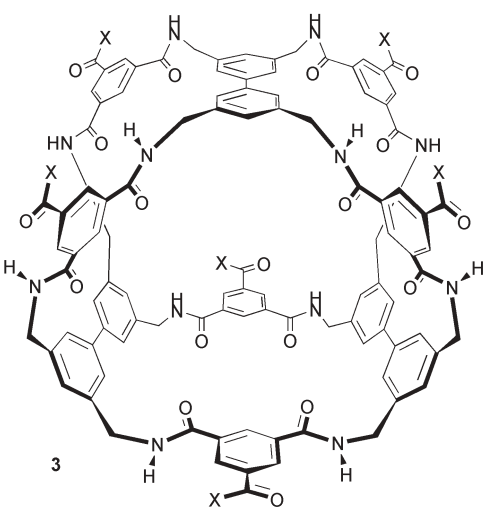
1



2a R = C_8H_{17}
2b R = CH_3

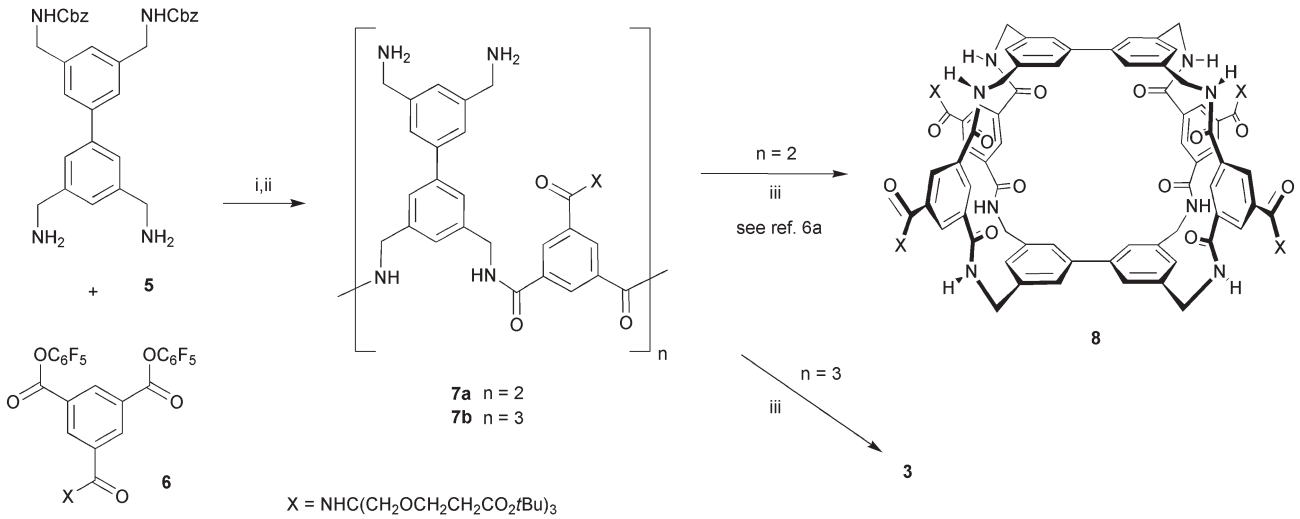


4a R = $C_{12}H_{25}$
4b R = CH_3



3

X = $NHC(CH_2OCH_2CH_2CO_2tBu)_3$



Scheme 1 Synthesis of **3**. *Reagents and conditions*: (i) iPr$_2$NEt, THF, high dilution; (ii) Pd/C, THF, methanolic ammonia, then flash chromatography; (iii) **6**, iPr$_2$NEt, THF, high dilution.

Monosaccharides were too small to fill the cavity and thus made fewer binding contacts.

The recognition properties of **3** were studied using ^{1}H NMR titrations in $CDCl_3$–CD_3OD (75 : 25, as used for the pure receptor), and by induced CD in the corresponding $CHCl_3$–CH_3OH mixture. These solvent systems are more polar than those we have used previously for organic-soluble receptors; for example, terphenyl-based **1** was studied in chloroform–methanol 92 : 8. β-Cellobioside **2a**, β-maltoside **4a**, α-maltoside **9**, β-glucoside **10** and α-glucoside **11** were used as substrates.

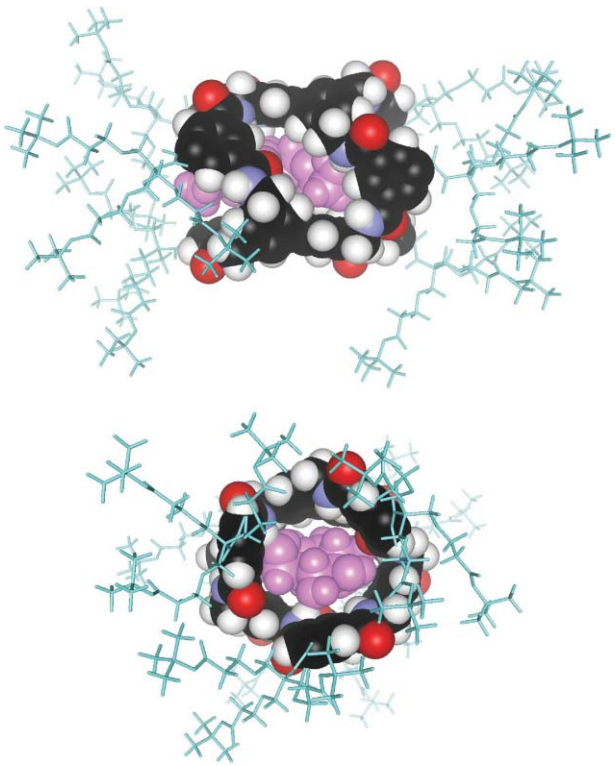


Fig. 1 Model of receptor **3** in an open conformation binding methyl β-D-maltoside **4b**, viewed from two perspectives. Solubilising side-chains are cyan (stick display), while the maltoside is coloured magenta. The conformation and position of the maltoside were optimised using MCMM.[9] Six intermolecular hydrogen bonds are formed between host and guest.

Initial NMR experiments showed that, despite the competitive solvent, receptor **3** is indeed effective for **2a** and **4a**. In both cases the receptor signals broadened, implying complex formation with a rate of exchange between "slow" and "fast" on the NMR timescale. In the case of **4a**, broadening was extreme towards the end of the titration at which point new signals appeared. In the case of **2a** the signals shifted during the titration without the appearance of new absorptions, suggesting that, to a first approximation, the exchange could be considered fast. On this basis, the data for **2a** were analysed according to a 1 : 1 binding model, giving a good fit to theory with K_a = 215 M^{-1}. In contrast to **2a** and **4a**, addition of α-maltoside **9** gave minimal signal motions (< 0.02 ppm) that were linear with concentration, while glucosides **10** and **11** caused no change whatever to the spectrum of **3**.

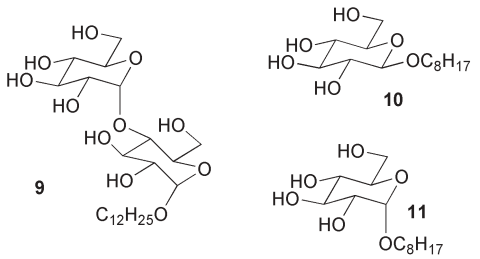
To provide more quantitative data we turned to induced circular dichroism (ICD). This technique is especially useful for achiral UV-absorbing receptors interacting with chiral UV-silent substrates such as carbohydrates;[10] the appearance of a CD signal on mixing is a firm indication of binding. Titrations of receptor **3**

Table 1 Association constants[a] (K_a, M^{-1}) for glycosides forming 1 : 1 complexes with receptor **3** in chloroform–methanol 75 : 25, as determined by CD and ^{1}H NMR titrations

Substrate	K_a (CD)	K_a (^{1}H NMR)[b]
dodecyl β-D-maltoside **4a**	780	n.d.[c]
octyl β-D-cellobioside **2a**	310	215
dodecyl α-D-maltoside **9**	v.s.[d]	v.s.[d]
octyl α-D-glucoside **11**	—[e]	—[e]

[a] From non-linear least-squares curve-fitting to a 1 : 1 binding model, implemented within Excel 2002. For binding curves, see Supplementary Information. [b] $CDCl_3$–CD_3OD. [c] Not determined, due to slow exchange. [d] Very small. Effects are minor and linear with concentration of added glycoside. [e] No effect with either technique.

against glycosides **2a**, **4a**, **9** and **11** gave the results gathered in Table 1. Both **2a** and **4a** gave substantial ICD effects, with excellent fits to a 1 : 1 binding model. The binding constant obtained for **2a** was reasonably close to that derived from NMR, while the value for maltoside **4a** was significantly higher at 780 M^{-1}. The results for α-glycosides **9** and **11** were also consistent with the NMR data; the former gave a CD signal which changed slowly and linearly with concentration, while the latter showed no ICD whatsoever.



Titration of β-glucoside **10** into **3** gave less expected results. In contrast to the ^{1}H NMR data, quite strong ICD effects were observed. "False positive" ICD results are unlikely, and we conclude that binding takes place but does not affect the NMR spectrum. The ICD data were inconsistent with 1 : 1 binding but suggested 1 : 2 (host : guest) stoichiometry. Analysis using a 1 : 1 + 1 : 2 binding model[11] yielded values of ~1 M^{-1} and 1470 M^{-1} for the successive binding constants K_{a1} and K_{a2}; this implies highly cooperative 1 : 2 binding with an overall K_a = 1470 M^{-2}.

The results may be summarised as follows. Firstly macrotetracycle **3** is a powerful carbohydrate receptor, comparing well with terphenyl-based **1**. Although higher K_a values were measured for **1** (up to 7000 M^{-1}), the solvent mixture was less polar (containing only 8% methanol). Secondly, **3** is strongly selective for β-glycosides **4a/2a/10** as against α-glycosides **9/11**. Thirdly, it is selective for *paired* monosaccharide units, forming 1 : 1 complexes with disaccharides **4a/2a** and a 1 : 2 complex with glucoside **10**. Of these, it prefers the disaccharides under the conditions of the titrations.[12] Finally receptor **3** is selective for maltoside **4a** *vs.* cellobioside **2a**, thus reversing the preference of the earlier system **1**.

In conclusion, we report a carbohydrate receptor with a novel tetracyclic architecture, and the rare ability to distinguish between disaccharide stereoisomers. The new receptor is the first to select for maltoside *vs.* cellobioside, complementing the cellobiose-selective **1**. The results confirm that selectivity-tuning is possible in these macropolycyclic polyamide hosts, encouraging the hope that "synthetic lectins" with a range of specificities may ultimately be possible.

Financial support from the EU and EPSRC is gratefully acknowledged. Mass spectra were provided by the EPSRC National MS Service Centre at the University of Swansea. We thank Prof. M. J. Hynes for access to the EQNMR binding analysis programme, and Drs P. Gale and G. Bates for assistance with its use.

Notes and references

1 R. A. Laine, *Glycobiology*, 1994, **4**, 759.
2 Leading references: H. J. Gabius, H. C. Siebert, S. André, J. Jiménez-Barbero and H. Rüdiger, *ChemBioChem*, 2004, **5**, 740; R. A. Dwek and T. D. Butters, *Chem. Rev.*, 2002, **102**, 283; C. R. Bertozzi and L. L. Kiessling, *Science*, 2001, **291**, 2357; C.-H. Wong, *Carbohydrate-based drug discovery*, Wiley-VCH, Weinheim, 2003.
3 Reviews: (*a*) A. P. Davis and T. D. James, in *Functional Synthetic Receptors*, ed. T. Schrader and A. D. Hamilton, Wiley-VCH, Weinheim, 2005, p. 45; (*b*) A. P. Davis and R. S. Wareham, *Angew. Chem., Int. Ed.*, 1999, **38**, 2978. Recent examples: (*c*) M. Mazik and H. Cavga, *J. Org. Chem.*, 2006, **71**, 2957; (*d*) J. D. Lee, N. T. Greene, G. T. Rushton, K. D. Shimizu and J. I. Hong, *Org. Lett.*, 2005, **7**, 963; (*e*) H. Abe, N. Masuda, M. Waki and M. Inouye, *J. Am. Chem. Soc.*, 2005, **127**, 16189; (*f*) M. Mazik, H. Cavga and P. G. Jones, *J. Am. Chem. Soc.*, 2005, **127**, 9045; (*g*) H. Abe, Y. Aoyagi and M. Inouye, *Org. Lett.*, 2005, **7**, 59; (*h*) C. Schmuck and M. Schwegmann, *Org. Lett.*, 2005, **7**, 3517; (*i*) H. P. Yi, X. B. Shao, J. L. Hou, C. Li, X. K. Jiang and Z. T. Li, *New J. Chem.*, 2005, **29**, 1213; (*j*) A. Vacca, C. Nativi, M. Cacciarini, R. Pergoli and S. Roelens, *J. Am. Chem. Soc.*, 2004, **126**, 16456; (*k*) O. Francesconi, A. Ienco, G. Moneti, C. Nativi and S. Roelens, *Angew. Chem., Int. Ed.*, 2006, **45**, 6693.
4 We use the term "biomimetic" for receptors which, as a minimum criterion, operate through non-covalent interactions. Effective but non-biomimetic carbohydrate recognition is possible through reversible covalent B–O bond formation. See ref. 3*a* and: S. Striegler, *Curr. Org. Chem.*, 2003, **7**, 81; T. D. James and S. Shinkai, *Top. Curr. Chem.*, 2002, **218**, 159; T. D. James, K. R. A. S. Sandanayake and S. Shinkai, *Angew. Chem., Int. Ed. Engl.*, 1996, **35**, 1911.
5 (*a*) U. Neidlein and F. Diederich, *Chem. Commun.*, 1996, 1493; (*b*) A. S. Droz and F. Diederich, *J. Chem. Soc., Perkin Trans. 1*, 2000, 4224; (*c*) A. S. Droz, U. Neidlein, S. Anderson, P. Seiler and F. Diederich, *Helv. Chim. Acta*, 2001, **84**, 2243; (*d*) R. D. Hubbard, S. R. Horner and B. L. Miller, *J. Am. Chem. Soc.*, 2001, **123**, 5810; (*e*) G. Lecollinet, A. P. Dominey, T. Velasco and A. P. Davis, *Angew. Chem., Int. Ed.*, 2002, **41**, 4093; (*f*) O. Rusin, K. Lang and V. Král, *Chem.–Eur. J.*, 2002, **8**, 655; (*g*) M. Dukh, D. Saman, K. Lang, V. Pouzar, I. Cerny, P. Drasar and V. Král, *Org. Biomol. Chem.*, 2003, **1**, 3458; (*h*) M. Mazik and A. König, *J. Org. Chem.*, 2006, **71**, 7854. For other systems with notable oligosaccharide binding properties, see ref. 3*c* and: (*i*) J. Otsuki, K. Kobayashi, H. Toi and Y. Aoyama, *Tetrahedron Lett.*, 1993, **34**, 1945; (*j*) J. L. Hou, X. B. Shao, G. J. Chen, Y. X. Zhou, X. K. Jiang and Z. T. Li, *J. Am. Chem. Soc.*, 2004, **126**, 12386; (*k*) Y. Q. Chen, X. Z. Wang, X. B. Shao, J. L. Hou, X. Z. Chen, X. K. Jiang and Z. T. Li, *Tetrahedron*, 2004, **60**, 10253; (*l*) J. Billing, H. Grundberg and U. J. Nilsson, *Supramol. Chem.*, 2002, **14**, 367.
6 (*a*) E. Klein, M. P. Crump and A. P. Davis, *Angew. Chem., Int. Ed.*, 2005, **44**, 298. For related systems see (*b*) A. P. Davis and R. S. Wareham, *Angew. Chem., Int. Ed.*, 1998, **37**, 2270; (*c*) T. J. Ryan, G. Lecollinet, T. Velasco and A. P. Davis, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 4863. Like **1**, these receptors prefer all-equatorial substrates.
7 For details of the synthesis and binding experiments, see Supplementary Information.
8 For further discussion see Supplementary Information.
9 Monte Carlo Molecular Mechanics (MCMM) searches were performed using MacroModel 9.0 (Maestro 7.0 interface), employing the MMFFs force field and GB/SA chloroform solvation.
10 Y. Kikuchi, K. Kobayashi and Y. Aoyama, *J. Am. Chem. Soc.*, 1992, **114**, 1351.
11 The programme EQNMR was employed; see M. J. Hynes, *J. Chem. Soc., Dalton Trans.*, 1993, 311.
12 Comparisons are difficult as the stoichiometries, and therefore the units of the binding constants, are different. However, an estimate can be obtained by comparing the concentrations at which guests occupied 50% of receptor binding sites during the titrations. For **4a** and **10** the figures are 1.4 mM and 25 mM respectively. By this measure, the receptor favours disaccharide **4a** by a factor of 18.

