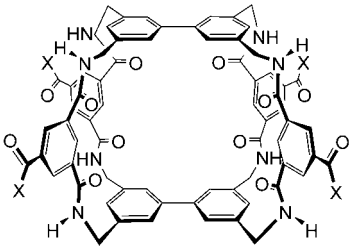
Molecular Recognition

Carbohydrate Recognition in Water by a Tricyclic Polyamide Receptor**

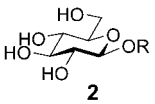
*Emmanuel Klein, Matthew P. Crump, and Anthony P. Davis**

Carbohydrate recognition in aqueous solution remains an important challenge in supramolecular chemistry.[1] Saccharides are complex, irregular, and multifunctional, while their hydroxylated exteriors blend well with water. Even nature has difficulty; binding constants of proteins with monosaccharides peak at approximately $10^7\,M^{-1}$, a remarkably low value for biological molecular recognition.[1a,2] Synthetic receptors that exploit covalent B–O bonds have been relatively successful,[1b,3] but there are still few reports of binding in water through noncovalent interactions.[1a,4] Such systems are of special interest as they shed light on natural carbohydrate recognition, a process of fundamental biological significance.[5]

A few years ago, we reported the octa-amide **1a** as a receptor for octyl glycosides in organic media.[6] The tricyclic



1a: X = OC_5H_{11}
1b: X = O^-
1c: X = NH(OBn)(OBn)(OBn)
1d: X = NH(CH$_2$CO$_2^-$)$_3$



2

core **1** was specifically targeted at β-glucosyl derivatives **2**. It was supposed that the axial hydrogens in **2** would make C–H–π contacts with the biphenyl surfaces, while the equatorial substituents would form hydrogen bonds to the isophthalamide spacers. Accordingly, **1a** showed high affinities for **2**

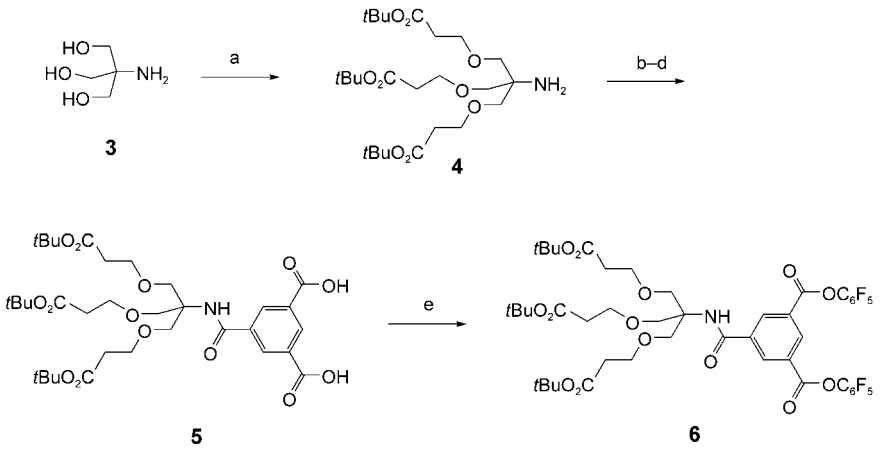
[*] Dr. E. Klein, Dr. M. P. Crump, Prof. A. P. Davis
School of Chemistry
University of Bristol
Cantock's Close
Bristol BS8 1TS (UK)
Fax: (+44) 117-929-8611
E-mail: anthony.davis@bristol.ac.uk

[**] Financial support was provided by the European Commission (network contract HPRN-CT-2002-00190). Assistance from Dr. C. Dempsey (Department of Biochemistry, University of Bristol) and the EPSRC National Mass Spectrometry Service Centre, Swansea, is gratefully acknowledged.

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

  DOI: 10.1002/anie.200461409

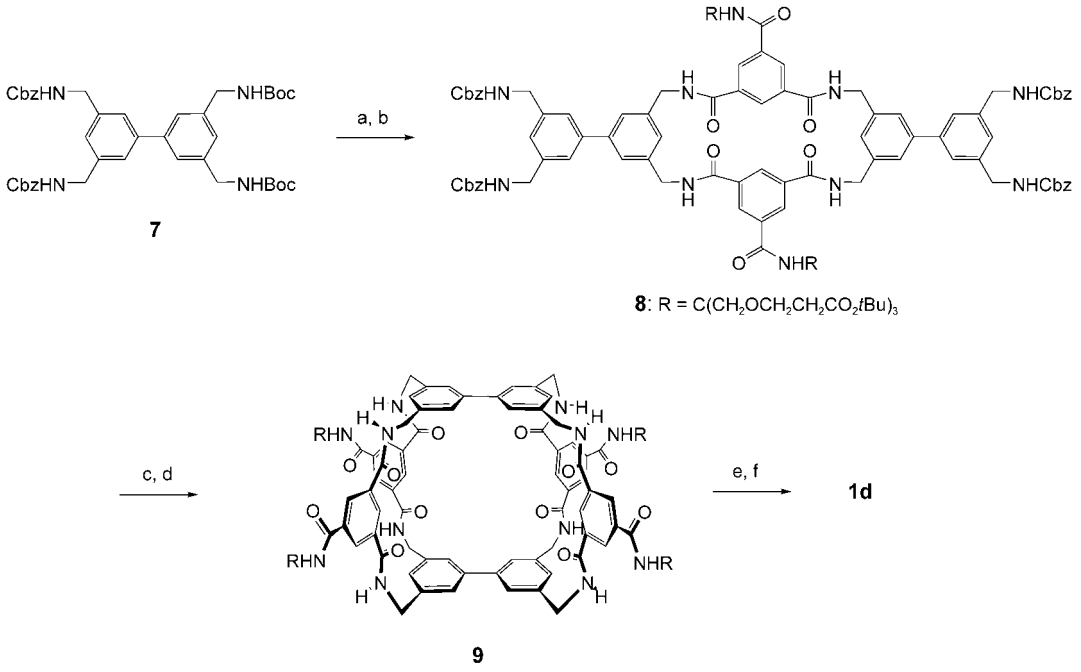
(R = octyl) and was less effective for the α anomer or octyl β-D-galactoside. Our prime objective, however, was recognition in water. In this more-challenging medium, hydrogen bonding would provide less (if any) impetus for binding, but, in compensation, the C–H–π interactions should be reinforced by the hydrophobic effect. Tetracarboxylate **1b** could not be employed owing to aggregation in water. Attempts to use a tris(hydroxymethyl)methyl solubilizing group also foundered, although precursor **1c** proved useful for phase-transfer experiments.[7] We now describe the synthesis of dodecacarboxylate **1d** and report that this polyanion provides, at last, an opportunity to study carbohydrate recognition by **1** in homogeneous aqueous solution.



Scheme 1. Synthesis of **6**: a) *tert*-butyl acrylate, NaOH, DMSO, H_2O; b) 1,3,5-benzenetricarbonyl trichloride (excess), *i*Pr_2NEt, THF; c) aq. NaOH (2 M); d) aq. HCl (3.5 M); e) C_6F_5OH, DCC, THF. DMSO = dimethylsulfoxide, DCC = *N*,*N*′-dicyclohexylcarbodiimide.

Receptor **1d** was prepared as shown in Schemes 1 and 2. The active ester **6** was synthesized from tris(hydroxymethyl)aminomethane **3** through the aminotriester **4**[8] and diacid **5** (Scheme 1). The biphenyl unit **7** was prepared as previously reported.[7] Acid deprotection of **7** followed by cyclization with **6** under high dilution gave macrocycle **8** (Scheme 2). Hydrogenolysis of **8** followed by a second [2+2] macrolactamisation gave the protected macrocycle **9**. The sodium salt of dodecacarboxylate **1d** was obtained through deprotection of **9** with trifluoroacetic acid followed by treatment with sodium hydroxide. ^{1}H NMR spectra of solutions in D_2O/H_2O consisted of sharp signals (Figure 1) and did not vary with concentration in the range 1–4 mM. Receptor **1d** was therefore assumed to be monomeric in aqueous solution at approximately 1 mM.

Preliminary NMR titration studies on the dodeca-*tert*-butyl ester **9** confirmed that, as expected, this macrocycle binds carbohydrates in the organic solvent system $CDCl_3$/CD_3OD (92:8). The K_a value for octyl β-D-glucopyranoside was 600 M^{-1}, which is similar to those obtained previously for **1a**[6] and **1c**.[7] Variations in the chemical shifts were also similar to those for the earlier systems. More significantly, the addition of glucose **10** to **1d** in D_2O/H_2O yielded qualitatively similar NMR effects. In particular, the signal for the inward-directed spacer protons, B (for labeling, see Figure 1), moved downfield, whereas that for the annular amide protons, NH^a, moved upfield and split into two (Figure 2). The signal for the protons C of the biphenyl group also split and moved



Scheme 2. Synthesis of **1d**: a) TFA, CH_2Cl_2; b) **6**, *i*Pr_2NEt, THF, high dilution; c) H_2, 10% Pd/C, THF/methanolic ammonia (1:1); d) **6**, *i*Pr_2NEt, THF, high dilution; e) TFA, CH_2Cl_2; f) aq. NaOH (1 M). Boc = *tert*-butoxycarbonyl, Cbz = benzyloxycarbonyl, TFA = trifluoroacetic acid.

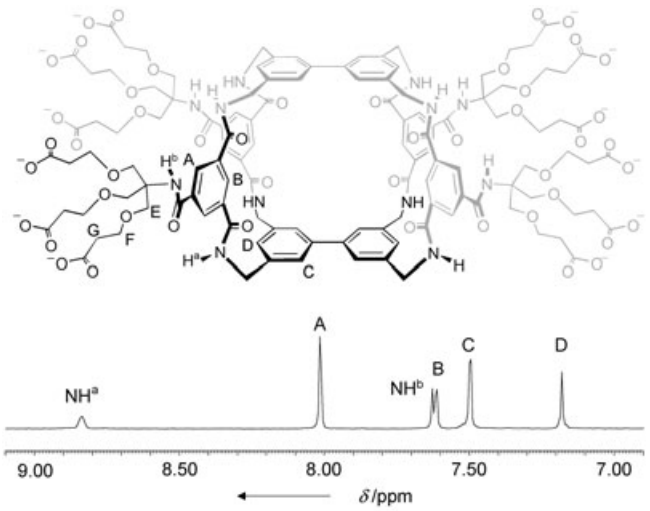


Figure 1. Partial ^{1}H NMR spectrum (600 MHz, 93:7 H_2O/D_2O) of receptor **1d**.

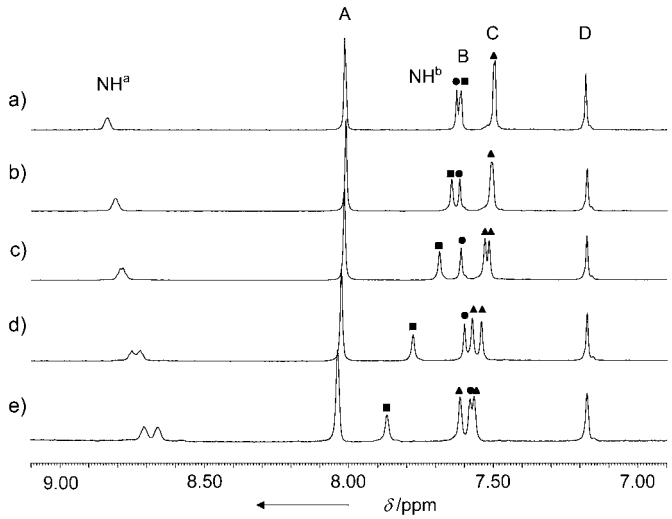


Figure 2. ^{1}H NMR spectra (600 MHz, 93:7 H_2O/D_2O) of receptor **1d** after addition of increasing amounts of D-glucose **10**: a) 0, b) 20, c) 50, d) 200, and e) 1000 equivalents. ● = NH^b, ■ = protons B, ▲ = protons C.

downfield. The splitting of the eight equivalent NH^a and C protons into groups of 4 is expected for a D_{2h} receptor that interacts with an asymmetrical substrate.[6] As illustrated in Figure 3a the motions could be fitted, to high accuracy, to a 1:1 binding model. The resulting binding constants varied according to the history of the glucose solution. For freshly dissolved glucose with an anomeric ratio α/β of 72:28 a K_a value of 4.6 M^{-1} was obtained in D_2O. For glucose that had been equilibrated overnight (α/β = 40:60), the binding constant was 9.2 M^{-1}. The figures imply substantial selectivity for the β anomer; indeed, they are consistent with $K_a \approx 14\,M^{-1}$ for β-glucopyranose and $< 1\,M^{-1}$ for the α anomer.[9]

Binding to other carbohydrates (**11**–**24**) was followed by NMR spectroscopy to yield the association constants shown in Table 1, page 301. Errors may be high for the smaller K_a

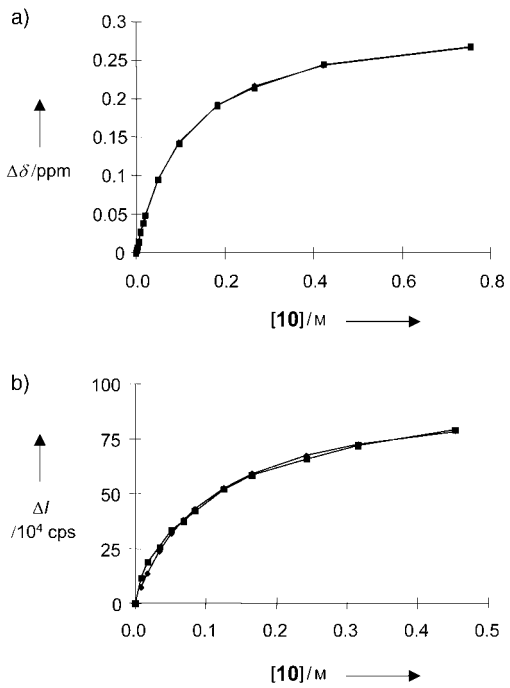


Figure 3. Theoretical (◆) and experimental (■) binding curves for a) NMR titration of **1d** + **10** (proton B) and b) fluorescence emission titration of **1d** + **10** (cps = counts per second; λ_{ex} = 266 nm, λ_{em} = 475 nm).

values (especially when $K_a \approx 2\,M^{-1}$), but selectivities are clearly significant. The receptor shows a notable preference for β-glucosyl units in accordance with the original design and with previous work on **1a** and **1c**. Thus the highest K_a values are to methyl β-D-glucoside **15** and cellobiose **22**. The selectivity for **22** relative to lactose **23** and maltose **24** is especially striking and mirrors the results in organic media for an "extended", terphenyl-based analogue of **1a**.[10] Other units with all-equatorial substitution (2-deoxy-D-glucose **17** and D-xylose **18**) are also bound well by the receptor. Remarkably, the removal of hydroxyl groups does not enhance binding—**17** is no more strongly bound than glucose, while rhamnose **20** and fucose **21** are poor substrates.

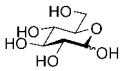
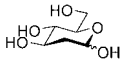
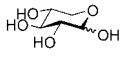
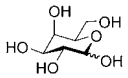
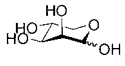
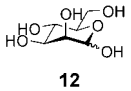
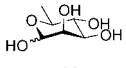
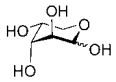
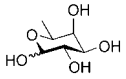
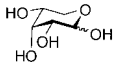
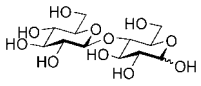
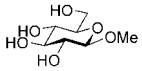
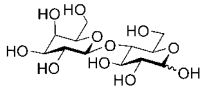
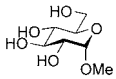
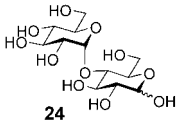
The formation of the complexes was also detected by means of fluorescence spectroscopy. On addition of substrates to **1d**, the fluorescence emission was enhanced and shifted slightly in wavelength (see Supporting Information). Analysis by curve fitting (see Figure 3b) yielded the binding constants shown in Table 2 which are in good agreement with the values obtained through ^{1}H NMR titrations.

Table 2: Association constants (K_a) between receptor **1d** and carbohydrates as measured in H_2O by fluorescence titration.[a]

Carbohydrate	K_a [M^{-1}]	Carbohydrate	K_a [M^{-1}]
D-glucose (**10**; α/β 40:60)	9.5	methyl α-D-glucoside (**16**)	5.7
methyl β-D-glucoside (**15**)	32	D-xylose (**18**)	3.8

[a] T = 298 K. For binding curves and further details, see Supporting Information.

Table 1: Association constants (K_a) between receptor **1d** and carbohydrates as measured in D_2O by ^1H NMR titration.[a]

Carbohydrate		K_a [M^{-1}]	Carbohydrate		K_a [M^{-1}]
 **10** (α/β 72:28)[b]	D-glucose	4.6	 **17**	2-deoxy-D-glucose[c]	7.2
10 (α/β 40:60)[c]	D-glucose	9.2	 **18**	D-xylose[c]	4.6
 **11**	D-galactose[c]	2.1	 **19**	D-lyxose[c]	v.s.[d]
 **12**	D-mannose[c]	v.s.[d]	 **20**	L-rhamnose[c]	v.s.[d]
 **13**	D-arabinose[c]	2.2	 **21**	L-fucose[c]	2.1
 **14**	D-ribose[c]	3.1	 **22**	D-cellobiose[c]	16.6
 **15**	methyl β-D-glucoside	27.3	 **23**	D-lactose[c]	v.s.[d]
 **16**	methyl α-D-glucoside	6.9	 **24**	D-maltose[c]	v.s.[d]

[a] $T = 296$ K. Data were analyzed by using a nonlinear least-squares curve-fitting program implemented within MS Excel 2000. Quoted figures were derived from the peak for proton B (see Figure 1); peaks for protons C and (sometimes) A could also be followed and gave closely consistent results. For binding curves and further details, see Supporting Information. [b] Freshly dissolved in D_2O. Anomeric ratio was monitored by ^1H NMR spectroscopy and was found to remain roughly constant throughout the titration. [c] Solution in D_2O allowed to equilibrate before use. For further information on anomeric ratios, see Supporting Information. [d] Very small. Minor signal movements, almost linear with substrate concentration.

Information on the geometry of the **1d**·glucose complex was provided by NOESY experiments. A portion of the spectrum is shown in Figure 4. Cross-peaks are observed between glucose protons (β anomer only) and receptor protons B, C, and NH^a, whereas connections from substrate to externally directed protons A, NH^b, and D are weak or

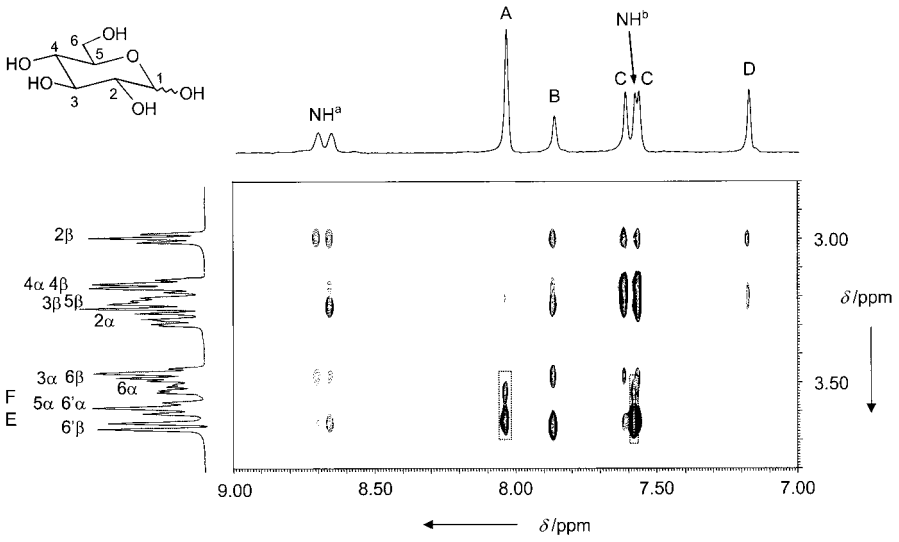


Figure 4. Partial NOESY spectrum of receptor **1d** and 1000 equivalents of D-glucose (**10**) in H_2O/D_2O 93:7 (600 MHz). The horizontal axis covers the region 7.0–9.0 ppm, which includes aromatic and NH protons of the receptor (for atom labeling, refer to Figure 1). The vertical axis covers 2.8–3.8 ppm, which includes protons 2–6 of glucose. The region plotted vertically also includes protons F and E of the receptor which are invisible in the 1D spectrum owing to the disparity in concentrations. Intramolecular cross-peaks that involve these signals are framed by dotted lines (•••••). All other cross-peaks represent intermolecular contacts. For further details see Supporting Information.

nonexistent. These data provide strong evidence that the carbohydrate is bound in the interior of the macrotricyclic structure. Both protons B and C appear to show cross-peaks with all of the carbohydrate protons, but the cross-peaks between protons B and proton H^6 of glucose, and those between proton C and the axial protons H^3, H^4, and H^5, are especially strong. Our proposed binding geometry, with "axial→biphenyl" and "equatorial→isophthaloyl" orientations, thus receives some support.

In conclusion we have shown that, in the form of **1d**, the tricyclic cage **1** can bind carbohydrates in water. Affinities are low, but selectivities are significant. The designed preference for β-glucosyl which was previously demonstrated in organic solvents is retained in the saccharides' natural environment. The formation of complexes has been characterized by ^{1}H NMR spectroscopy and confirmed by a second, independent technique, namely fluorescence spectroscopy. Tricycle **1** is thus established as a carbohydrate receptor with consistent behavior across a wide range of media. This property should allow its use in studies which complement those on natural receptors and further elucidate the principles underlying biological carbohydrate recognition.

Received: July 23, 2004

Keywords: carbohydrates · hydrophobic effect · molecular recognition · receptors · supramolecular chemistry

[1] a) A. P. Davis, R. S. Wareham, *Angew. Chem.* **1999**, *111*, 3160; *Angew. Chem. Int. Ed.* **1999**, *38*, 2978; b) T. D. James, K. R. A. S. Sandanayake, S. Shinkai, *Angew. Chem.* **1996**, *108*, 2038; *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1911; c) S. Striegler, *Curr. Org. Chem.* **2003**, *7*, 81.

[2] F. A. Quiocho, *Pure Appl. Chem.* **1989**, *61*, 1293.

[3] For recent examples, see: W. Q. Yang, J. Yan, H. Fang, B. H. Wang, *Chem. Commun.* **2003**, 792; C. J. Ward, P. Patel, T. D. James, *Org. Lett.* **2002**, *4*, 477; Z. L. Zhong, E. V. Anslyn, *J. Am. Chem. Soc.* **2002**, *124*, 9014; J. N. Camara, J. T. Suri, F. E. Cappuccio, R. A. Wessling, B. Singaram, *Tetrahedron Lett.* **2002**, *43*, 1139; H. S. Cao, D. I. Diaz, N. DiCesare, J. R. Lakowicz, M. D. Heagy, *Org. Lett.* **2002**, *4*, 1503; W. Yang, H. He, D. G. Drueckhammer, *Angew. Chem.* **2001**, *113*, 1764; *Angew. Chem. Int. Ed.* **2001**, *40*, 1714.

[4] R. D. Hubbard, S. R. Horner, B. L. Miller, *J. Am. Chem. Soc.* **2001**, *123*, 5810; V. Král, O. Rusin, F. P. Schmidtchen, *Org. Lett.* **2001**, *3*, 873; N. Sugimoto, D. Miyoshi, J. Zou, *Chem. Commun.* **2000**, 2295. For leading references to studies in nonaqueous media, see ref [1a] and: T. Ishi-i, M. A. Mateos-Timoneda, P. Timmerman, M. Crego-Calama, D. N. Reinhoudt, S. Shinkai, *Angew. Chem.* **2003**, *115*, 2402; *Angew. Chem. Int. Ed.* **2003**, *42*, 2300; M. Segura, B. Bricoli, A. Casnati, E. M. Munoz, F. Sansone, R. Ungaro, C. Vicent, *J. Org. Chem.* **2003**, *68*, 6296; K. Wada, T. Mizutani, S. Kitagawa, *J. Org. Chem.* **2003**, *68*, 5123; R. Welti, F. Diederich, *Helv. Chim. Acta* **2003**, *86*, 494; M. Mazik, W. Radunz, W. Sicking, *Org. Lett.* **2002**, *4*, 4579; Y. H. Kim, J. I. Hong, *Angew. Chem.* **2002**, *114*, 3071; *Angew. Chem. Int. Ed.* **2002**, *41*, 2947; K. Ladomenou, R. P. Bonar-Law, *Chem. Commun.* **2002**, 2108; J. H. Liao, C. T. Chen, H. C. Chou, C. C. Cheng, P. T. Chou, J. M. Fang, Z. Slanina, T. J. Chow, *Org. Lett.* **2002**, *4*, 3107; J. Bitta, S. Kubik, *Org. Lett.* **2001**, *3*, 2637; S. Tamaru, M. Yamamoto, S. Shinkai, A. B. Khasanov, T. W. Bell, *Chem. Eur. J.* **2001**, *7*, 5270.

[5] R. A. Dwek, T. D. Butters, *Chem. Rev.* **2002**, *102*, 283 (and succeeding articles); C. R. Bertozzi, L. L. Kiessling, *Science* **2001**, *291*, 2357; T. Feizi, B. Mulloy, *Curr. Opin. Struct. Biol.* **2001**, *11*, 585 (and succeeding articles).

[6] A. P. Davis, R. S. Wareham, *Angew. Chem.* **1998**, *110*, 2397; *Angew. Chem. Int. Ed.* **1998**, *37*, 2270.

[7] T. J. Ryan, G. Lecollinet, T. Velasco, A. P. Davis, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 4863.

[8] C. M. Cardona, R. E. Gawley, *J. Org. Chem.* **2002**, *67*, 1411.

[9] This estimate is subject to large errors, and is therefore tentative.

[10] G. Lecollinet, A. P. Dominey, T. Velasco, A. P. Davis, *Angew. Chem.* **2002**, *114*, 4267; *Angew. Chem. Int. Ed.* **2002**, *41*, 4093.

