Carbohydrates

DOI: 10.1002/anie.200805497

Synthetic Lectins

*Stefan Kubik**

carbohydrate receptors · carbohydrates · lectins · molecular recognition · supramolecular chemistry

Membranes of plant and animal cells are covered extracellularly with a dense, up to 140-nm-thick carbohydrate layer, the so-called glycocalix. This layer, which consists of oligosaccharides covalently linked to lipids or to proteins associated with the membrane, differs between cell types, thus providing every cell with a characteristically structured outer surface. The absence or presence of certain oligosaccharides on erythrocyte membranes determines, for example, the blood type of humans. Also pathological changes or changes within the life cycle, for example growth and division, have characteristic effects on the carbohydrates on the cell surface.

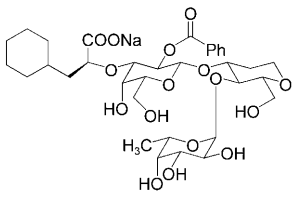
The glycocalix thus contains information about the type and the status of a cell. Consequently, it participates in a number of biological recognition processes all of which involve specific cell–cell interactions.[1] Examples include fertilization, during which receptors on the sperm surface recognize oligosaccharide residues of glycoproteins located on the egg membrane; infection, which requires pathogens such as microbes, bacteria, or viruses to bind to a host cell; and inflammation, which is induced by the interaction of leucocytes with endothelial cells of the blood vessels close to the injured site. Overzealous recruitment of leucocytes can lead to acute or chronic inflammatory diseases such as rheumatoid arthritis, psoriasis, and dermatitis.

The primary step in all of these processes, the actual cell–cell recognition, involves selective interactions between membrane proteins on the surface of one cell and peripheral oligosaccharides in the glycocalix of the second cell. These proteins are termed lectins.[2]

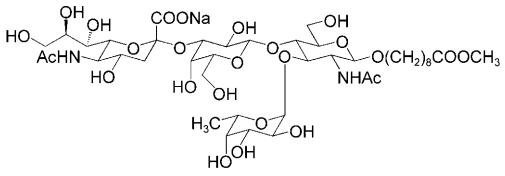
A promising approach to prevent harmful infections or to treat inflammatory diseases is the specific intervention in carbohydrate-mediated cell–cell recognition processes. This concept is in fact more important that the above-mentioned examples indicate because tissue damage in an organ caused by temporary interruption of blood supply, for example during surgery or after a stroke, and metastasis of cancer cells are also initiated by cell–cell adhesion.[1] Efficient therapeutics to prevent or treat these health risks would therefore be of considerable interest for pharmaceutical research.

[*] Prof. S. Kubik
Fachbereich Chemie—Organische Chemie
Technische Universität Kaiserslautern
Erwin-Schrödinger-Strasse, 67663 Kaiserslautern (Germany)
Fax: (+49) 631-205-3921
E-mail: kubik@chemie.uni-kl.de
Homepage: http://www.chemie.uni-kl.de/fachrichtungen/oc/kubik/

There are two complementary approaches to inhibit interactions between lectins and carbohydrates. In the first, the active center of a lectin is blocked with a corresponding antagonist, and in the other, the carbohydrate is protected with a receptor. Compound **1** is an example of the first strategy.[3a] This carbohydrate derivative binds to E-selectin with an IC_{50} value of 1–2 μM, roughly three orders of magnitude less than that of the natural substrate sialyl Lewis^x **2** ($IC_{50} = 1$ mM).[3b] Potentially **1** would thus be able to inhibit an early step of leucocyte recruitment in which E-selectin on endothelial plasma membranes interacts with sialyl Lewis^x on leucocyte membranes.



1



2

The second strategy, the search for synthetic carbohydrate receptors, falls within the field of supramolecular chemistry. Although recent years have seen considerable progress in this area, the development of efficient and selective synthetic receptors for carbohydrates, particularly ones that are effective in competitive solvents, remains a major challenge.[4] Reasons for this are that interactions of a receptor with the OH groups of a carbohydrate-derived substrate do not fundamentally differ from interactions with water molecules, and the structural similarity of many carbohydrates, D-glucose and D-mannose, for example, differ in the configuration of only a single OH group on the ring.

Inspiration for the design of artificial carbohydrate receptors often arose from crystal structures of protein–carbohydrate complexes.[5] These show that substrate binding in the active center of a protein usually involves a well-defined array of hydrogen bonds with amide groups in the

side chains of asparagine residues, carboxylate groups from aspartates, OH groups in serine side chains, and NH groups from lysine, tryptophan, and imidazole residues frequently operating as hydrogen-bond donors and acceptors. Electrostatic interactions can contribute to the complexation of charged substrates, and in some proteins the substrate is also coordinated to a metal center. Interestingly, sandwiching of the substrate between aromatic amino acid side chains is a recurring binding motif in several protein–carbohydrate complexes, suggesting a contribution from C–H···π interactions to substrate binding. As an example, the binding mode with which D-glucose is bound inside the active center of the D-galactose binding protein is depicted in Figure 1.[5b]

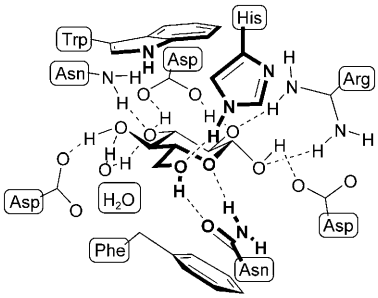
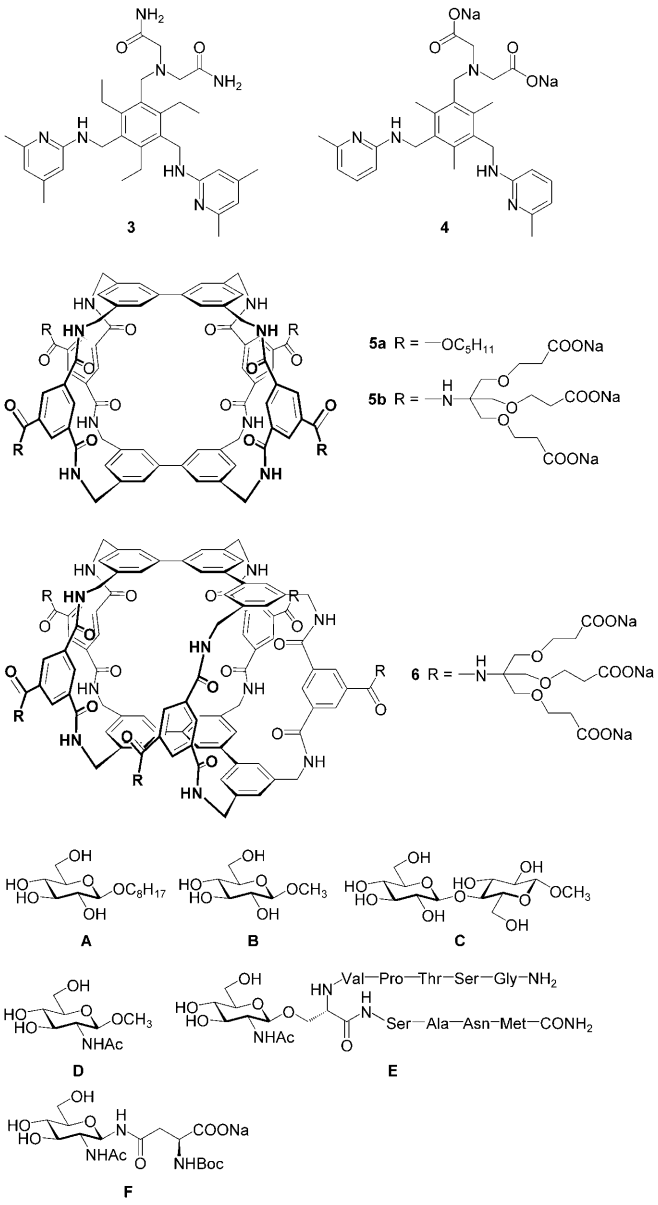


Figure 1. Representation of the interactions of D-glucose with functional groups inside the active center of the D-galactose binding protein.

Recent work from the Waters group demonstrates that C–H···π interactions between a carbohydrate and an arene indeed cooperatively contribute to the overall stability of a complex.[6] In addition, the notion put forward by Lemieux[7]—that the driving force for the binding of carbohydrates in water arises at least in part from solvent reorganization, in particular the release of the water molecules located near the hydrophobic regions of hydrated carbohydrate molecules—recently received support from the experimental findings of the Davis group.[8] Accordingly, solvent reorganization provides an entropic, possibly even a favorable enthalpic contribution to complex formation.

At least two types of interactions detected in protein–carbohydrate complexes are realized in most of the more recently described carbohydrate receptors; most frequently these are hydrogen bonding and a coplanar arrangement of the substrate and aromatic receptor subunits to allow for C–H···π interactions. A number of efficient carbohydrate receptors have been identified, for example, that contain a cyclic arrangement of binding sites along a trisubstituted benzene core or other aromatic scaffolds. An example is receptor **3** which displays high affinity for 1-*O*-n-octyl-β-D-glucopyranoside (**A**) in chloroform (K_a value of the 1:1 complex: 144520 M^{-1}).[9a]

Characterization of the binding properties of such receptors in organic solvents provides important insights into how the type and arrangement of hydrogen-bond donors and acceptors affect affinity and selectivity. In addition, they help in understanding the binding mechanisms relevant to natural systems since the solvent chloroform mimics the environment



inside the active centers of proteins very well. In view of applications, however, carbohydrate receptors must be identified that are effective in aqueous media. One way to achieve this is to make use of stronger types of interactions for substrate binding. Receptor **4**, for example, whose carboxylate groups can form charge-assisted hydrogen bonds to the substrate similar to those of aspartate groups in the active centers of carbohydrate-binding proteins, binds 1-*O*-methyl-β-D-glucopyranoside (**B**) also in water. With an association constant K_a of only 2 M^{-1}, the stability of the 1:1 complex is, however, very low.

An alternative approach was pursued by Davis et al. The carbohydrate receptors they developed consist of *two* aromatic subunits appropriately spaced to allow intercalation of the substrate.[10] Four linkers stabilize the parallel arrangement of these arenes and also provide hydrogen-bond donor and acceptor groups which can converge toward the interior of the cavity to interact with the peripheral OH groups of an

included substrate. The cavity of these receptors therefore very well mimics the situation inside the active centers of carbohydrate-binding proteins.

The first member of this class of receptors, compound **5a**, binds **A** in 5% $CD_3OD/CDCl_3$ with an association constant determined by NMR spectroscopy of $980\,M^{-1}$.[10a] In chloroform, this complex has a remarkable K_a value of $3\times10^5\,M^{-1}$, according to fluorescence spectroscopy. The affinity of **5a** for the corresponding α anomer of **A** or for 1-*O*-n-octyl-β-D-mannopyranoside is significantly lower, most probably because the axial substituents in these monosaccharides prevent efficient inclusion into the receptor cavity.

Binding studies in water were possible with the water-soluble analogue of **5a**, receptor **5b**, and a stability constant of $27\,M^{-1}$ was determined for the complex between **5b** and **B**.[10b] The affinity of **5b** for carbohydrates with axial substituents is significantly lower, in accordance with the binding studies in organic solvents. Thus, these studies clearly demonstrated that a neutral receptor can recognize monosaccharides in the highly competitive solvent water with appreciable affinity even when no charge-assisted hydrogen bonds are used for substrate binding.[11] A likely explanation is that solvent reorganization provides thermodynamically favorable contributions to complex formation.[8] Yet, although carbohydrate affinity of **5b** is considerably higher than that of **4**, it clearly remains below that of natural systems, which typically lies in the millimolar range.[2c]

Receptor **6**, which has a larger cavity than that of **5b**, binds disaccharides in water.[10c] The association constant of the complex between 1-*O*-methyl-β-D-cellobiose (**C**) and **6** amounts to approximately $900\,M^{-1}$, for example. The affinity of **6** for lactose, an epimer of cellobiose with a single axially oriented OH group, is almost two orders of magnitude lower. In terms of affinity and selectivity, the binding properties of **6** thus approach those of natural lectins. It must be considered, however, that the inclusion of cellobiose in the cavity of **6** brings larger hydrophobic surfaces in close contact and establishes more direct receptor–substrate interactions than binding of a monosaccharide to **5b**.

Most recent investigations in the Davis group have now revealed that the affinity of **5b** for certain monosaccharide derivatives is substantially greater than that for glucose.[10d] Specifically, *N*-acetyl-1-*O*-methyl-β-D-glucosamine (**D**) is bound by **5b** with an association constant of $630\,M^{-1}$; this complex is thus more than one order of magnitude more stable than that of **B** or glucose. In addition, binding is highly selective. Of the 21 other monosaccharide derivatives studied, only four bind to **5b** with an association constant greater than $10\,M^{-1}$, and none of these complexes has an association constant exceeding $60\,M^{-1}$. For comparison, with a K_a of $730\,M^{-1}$ the natural lectin wheat germ agglutinin (WGA) displays a similar affinity as **5b** for **D**, and the carbohydrate selectivity of WGA is even lower. The affinity of WGA for *N*-acetyl-1-*O*-methyl-α-D-glucosamine amounts to $480\,M^{-1}$, for example, while the corresponding complex of **5b** has a K_a value of only $24\,M^{-1}$.

These promising results prompted the question as to whether **5b** also exhibits affinity for glycoproteins containing *O*-glycosidically attached *N*-acetylglucosamine residues. This was tested by using peptide **E** as a model substrate, which is based on a sequence from casein kinase II (CKII) after glycosylation of the serine residue in position 6 by *O*-GlcNAc transferase. The binding of **5b** to **E** is indeed remarkably strong; according to an NMR titration the association constant of the complex amounts to $1040\,M^{-1}$. That binding is a result of specific interactions with the monosaccharide residue was demonstrated by comparing affinities of **5b** for **E** and for an analogue of **E** lacking the carbohydrate subunit. Interestingly, only very weak binding of **5b** to asparagine derivative **F** containing an *N*-linked acetylglucosamine moiety was detected; this is also indicative of the high binding selectivity of this receptor.

Structural characterization of the complex between **5b** and **D** using a combination of NMR spectroscopy and molecular modeling revealed that the substrate occupies the cavity between the two arenes, which are aligned parallel to each other (as in the complexes of receptors **5a** and **6**).[10a–c] Numerous hydrogen bonds and C–H···π interactions ensure efficient and well-defined binding. The *N*-acetyl group in the 2-position of the substrate is located between two isophthalamide linkers at one of the smaller portals of the cavity. It is held tightly by several hydrophobic contacts, two hydrogen bonds to the carbonyl oxygen atom, and NH···π interactions. These multiple attractive interactions are presumably the explanation for the higher stability of the *N*-acetyl-1-*O*-methyl-β-D-glucosamine complex of **5b** with respect to the glucose complex.

Compounds **5b** and **6** therefore represent the first carbohydrate receptors for which the term "synthetic lectin" is truly warranted. They demonstrate that efficient and selective recognition of carbohydrates in water can indeed be realized by using synthetic receptors, and they provide information about general principles for future receptor design. Intercalation of the substrate between two aromatic residues, for example, seems to give a significant contribution to binding in water. An exciting question for future studies is now whether such receptors exert a characteristic effect on cell adhesion phenomena.

Published online: December 29, 2008

[1] Y. C. Lee, R. T. Lee, *Acc. Chem. Res.* **1995**, *28*, 321–327; T. K. Lindhorst, *Chem. Unserer Zeit* **2000**, *34*, 38–52; Thematic issue of *Chem. Rev.* on Glycobiology **2002**, *102*, 283–602; M. E. Taylor, K. Drickamer, *Introduction to Glycobiology*, Oxford University Press, Oxford, **2006**; T. K. Lindhorst, *Essentials of Carbohydrate Chemistry and Biochemistry*, Wiley-VCH, Weinheim, **2007**.

[2] a) H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674; b) H. Lis, N. Sharon, *Lectins*, Kluwer, Dordrecht, **2003**; c) T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387–430.

[3] a) P. Sears, C.-H. Wong, *Angew. Chem.* **1999**, *111*, 2446–2471; *Angew. Chem. Int. Ed.* **1999**, *38*, 2300–2324; b) G. Thoma, R. Bänteli, W. Jahnke, J. L. Magnani, J. T. Patton, *Angew. Chem.* **2001**, *113*, 3756–3759; *Angew. Chem. Int. Ed.* **2001**, *40*, 3644–3647.

[4] A. P. Davis, R. S. Wareham, *Angew. Chem.* **1999**, *111*, 3160–3179; *Angew. Chem. Int. Ed.* **1999**, *38*, 2978–2996; A. P. Davis,

T. D. James in *Functional Synthetic Receptors* (Eds.: T. Schrader, A. D. Hamilton), Wiley-VCH, Weinheim, **2005**, pp. 45 – 109.
[5] a) F. A. Quiocho, *Annu. Rev. Biochem.* **1986**, *55*, 287 – 315; b) F. A. Quiocho, *Pure Appl. Chem.* **1989**, *61*, 1293 – 1306; c) W. I. Weis, K. Drickamer, *Annu. Rev. Biochem.* **1996**, *65*, 441 – 473.
[6] Z. R. Laughrey, S. E. Kiehna, A. J. Riemen, M. L. Waters, *J. Am. Chem. Soc.* **2008**, *130*, 14625 – 14633.
[7] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373 – 380.
[8] E. Klein, Y. Ferrand, N. P. Barwell, A. P. Davis, *Angew. Chem.* **2008**, *120*, 2733 – 2736; *Angew. Chem. Int. Ed.* **2008**, *47*, 2693 – 2696.
[9] a) M. Mazik, M. Kuschel, *Eur. J. Org. Chem.* **2008**, 1517 – 1526; b) M. Mazik, H. Cavga, *J. Org. Chem.* **2006**, *71*, 2957 – 2963.
[10] a) A. P. Davis, R. S. Wareham, *Angew. Chem.* **1998**, *110*, 2397 – 2401; *Angew. Chem. Int. Ed.* **1998**, *37*, 2270 – 2273; b) E. Klein, M. P. Crump, A. P. Davis, *Angew. Chem.* **2005**, *117*, 302 – 306; *Angew. Chem. Int. Ed.* **2005**, *44*, 298 – 302; c) Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619 – 622; 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 – 1811; *Angew. Chem. Int. Ed.*, **2009**, *48*, 1775 – 1779.
[11] For other examples of receptors that bind carbohydrates in water, see: R. Yanagihara, Y. Aoyama, *Tetrahedron Lett.* **1994**, *35*, 9725 – 9728; V. Král, O. Rusin, F. P. Schmidtchen, *Org. Lett.* **2001**, *3*, 873 – 876; O. Rusin, K. Lang, V. Král, *Chem. Eur. J.* **2002**, *8*, 655 – 663.

