

2.7 Synthetic Approaches to Study Multivalent Carbohydrate–Lectin Interactions

Valentin Wittmann

2.7.1 Introduction

The specific recognition of carbohydrate structures in biological systems (Box 5) by carbohydrate-binding proteins (lectins) is the basis of numerous intra- and intercellular events ranging from the control of protein folding to cell–cell communication during development, inflammation, and cancer metastasis [1]. Investigation of carbohydrate–lectin interactions can be approached from two directions. One is characterization of the protein part by molecular biology and structure determination (X-ray crystallography, NMR spectroscopy) [2]. In the other approach, which relies on synthetic organic chemistry, the specificity and affinity of modified or artificial lectin ligands and their effect on lectin function is studied [3, 4]. High-affinity lectin ligands are, furthermore, of considerable medicinal interest in the diagnosis and inhibition of carbohydrate-mediated processes such as inflammation or microbial adhesion [5]. The generation of high-affinity lectin ligands, however, is not trivial because most saccharide ligands bind to their protein receptors only weakly with dissociation constants typically in the milli- to micromolar range. Because many lectins have several binding sites or occur in oligomeric or clustered form on cell membranes, the creation of multivalent carbohydrate derivatives is a promising means of producing high-affinity ligands [3, 6–8].

2.7.2 Mechanistic Aspects of Multivalent Interactions

Multivalent interactions are characterized by the simultaneous binding of several ligands on one biological entity (surface, macromolecule) with several receptors on another entity (Figure 2.7.1) [8]. This type of interaction has unique collective properties that are qualitatively different from those of the corresponding monovalent systems. Not only it is possible to regulate the strength of an interaction by the number of receptor–ligand contacts, multivalent interactions also have different kinetic properties. As Whitesides et al. have demonstrated, it is possible to increase

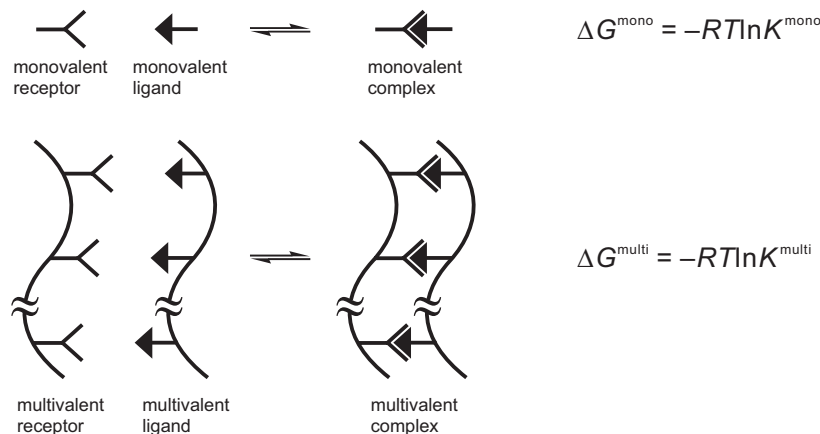


Fig. 2.7.1. Comparison of monovalent and multivalent interactions.

the rate of dissociation (k_{off}) of a multivalent complex by addition of a monovalent ligand [9]. A complete mechanistic description of multivalent binding is, however, difficult because of the complexity of such systems. The formation of a multivalent interaction involves many possible intermediates. Beside intramolecular binding of a multivalent receptor and a multivalent ligand, intermolecular binding may occur, leading to cross-linking and precipitation. The next section therefore focuses on the basic thermodynamics of the simplest multivalent system, the interaction of a bivalent ligand with a bivalent receptor, according to an analysis by Whitesides et al. [8].

The parameter ΔG^{multi} is made up of enthalpic (ΔH^{multi}) and entropic (ΔS^{multi}) components (Eq. 1) which have to be considered separately.

$$\Delta G^{\text{multi}} = \Delta H^{\text{multi}} - T\Delta S^{\text{multi}} \quad (1)$$

The enthalpy of binding (ΔH^{multi}) is, to a first approximation, the sum of the enthalpies of the individual monovalent interactions, i.e. for a bivalent system $\Delta H^{\text{bi}} = 2\Delta H^{\text{mono}}$ (Figure 2.7.2A, Case 1). This, however, only applies if the bivalent complex is unstrained and the binding events do not interfere with each other. If the binding of the first ligand interferes with binding of the second, the enthalpy of binding is less favorable (less negative) and $\Delta H^{\text{bi}} > 2\Delta H^{\text{mono}}$ (Figure 2.7.2A, Case 2). Such binding is enthalpically diminished and might occur if the bivalent complex is strained or if the first binding event exerts a negative allosteric effect on the second. Enthalpically enhanced binding is observed if the second binding event is more favorable than the first, because of a positive allosteric effect or because of favorable secondary interactions between the tether and the receptor and $\Delta H^{\text{bi}} < 2\Delta H^{\text{mono}}$ (Figure 2.7.2A, Case 3). A possible example of enthalpically enhanced binding is the interaction of the pentameric cholera toxin with five GM₁ molecules on cell surfaces [10].

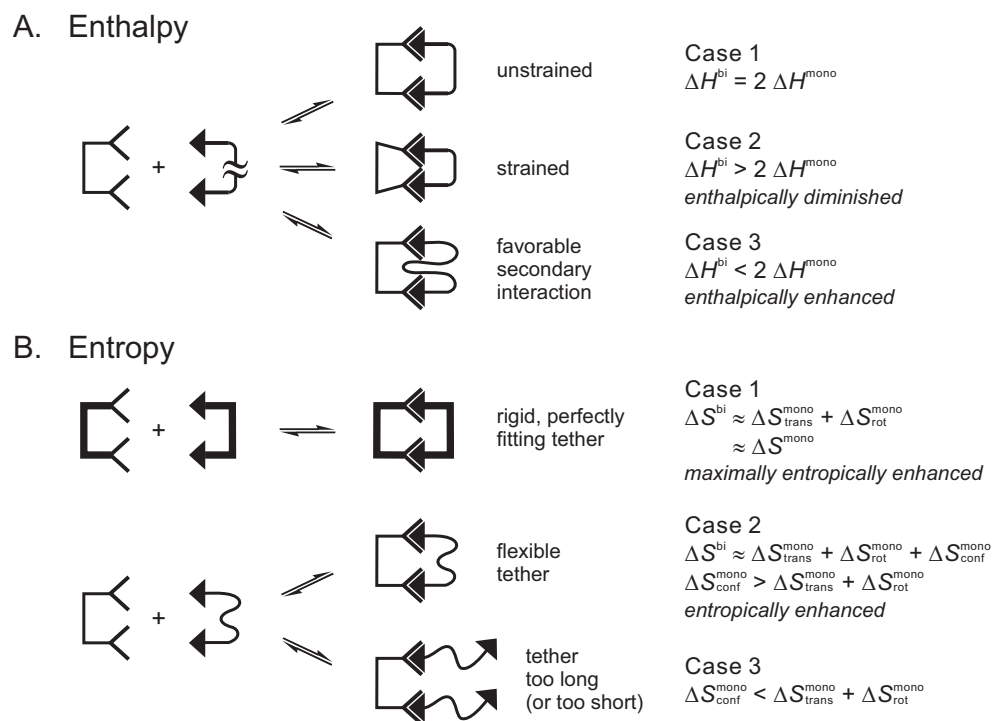


Fig. 2.7.2. Enthalpy and entropy of different binding modes of bivalent interactions.

The entropy of binding (ΔS^{multi}) of a multivalent interaction can be divided into contributions from changes in translational (ΔS_{trans}^{multi}), rotational (ΔS_{rot}^{multi}), conformational (ΔS_{conf}^{multi}), and hydrational ($\Delta S_{H_2O}^{multi}$) entropy. The latter is assumed to be similar in each situation and is therefore ignored in this discussion. Also the weak logarithmic dependence of translational and rotational entropy on the mass and size of different molecules is ignored. For a bivalent interaction several cases can again be distinguished. If the two ligands and the two receptors are connected by rigid, perfectly fitting spacers, $\Delta S_{conf}^{multi} = 0$ and the interaction occurs with an entropy equivalent to a single monovalent interaction (Figure 2.7.2B, Case 1). This case of maximum entropic enhancement is, in general, unrealistic, because all tethers are somewhat flexible and ΔS_{conf}^{multi} is almost always unfavorable (less than zero). If this conformational cost is less than the total translational and rotational cost ($\Delta S_{conf}^{multi} > \Delta S_{trans}^{multi} + \Delta S_{rot}^{multi}$), the bivalent association is still entropically enhanced and favored over an intermolecular interaction (Figure 2.7.2B, Case 2). If $\Delta S_{conf}^{multi} < \Delta S_{trans}^{multi} + \Delta S_{rot}^{multi}$, bivalent binding is entropically diminished and a (1+2) association is favored (Figure 2.7.2B, Case 3).

According to this discussion, ΔG^{multi} for a (theoretical) bivalent system with rigid perfectly fitting spacers is given by Eqs (2) and (3).

$$\Delta G^{\text{bi}} = 2\Delta H^{\text{mono}} - T\Delta S^{\text{mono}} \quad (2)$$

$$\Delta G^{\text{bi}} = \Delta G^{\text{mono}} + \Delta H^{\text{mono}} \quad (3)$$

Because ΔS^{mono} is usually less than zero, in such a system ΔG^{bi} is even more favorable than $2\Delta G^{\text{mono}}$ and, therefore, $K^{\text{bi}} > (K^{\text{mono}})^2$. For systems with greater flexibility calculation of the binding enhancement is more complex. A suitable theoretical model has been published recently [11].

The discussion of a bivalent interaction in the general sense is applicable to N -valent interactions and illustrates that the design of the tethers connecting individual binding sites within a multivalent ligand is critical to obtaining high-affinity ligands. Usually, however, the optimization of multivalent ligands is performed by trial and error. In these cases the mechanisms by which binding occurs are less critical. To describe the binding enhancement in such uncharacterized multivalent systems compared with the corresponding monovalent system, Whitesides et al. proposed the empirical parameter β , which is the ratio of the association constants K^{multi} and K^{mono} as defined by Eqs (4) and (5) [8].

$$\beta = K^{\text{multi}}/K^{\text{mono}} \quad (4)$$

$$\Delta G^{\text{multi}} = \Delta G^{\text{mono}} - RT \ln \beta \quad (5)$$

Systems with high values of β are useful, irrespective of their mechanism of action. In fact, many multivalent systems with large values of β do not reach the binding enhancement which would be possible in the case of maximum entropic enhancement and enthalpic additivity.

Many examples of multivalent ligands employing numerous scaffolds have been described; these differ in size, carbohydrate content, and flexibility [3, 6–8]. Glycopolymers, for example, are able to cover large areas of cell surfaces and bridge several membrane-located lectins (“statistical” multivalency). Low-valent glycoclusters (miniclusters), on the other hand, bind preferentially to several binding sites of a single (oligomeric) lectin proximate in space and may be tailored to lectins with known 3D structure (“directed” multivalency).

2.7.3

Low-valent Glycoclusters for “Directed Multivalency”

Low-valent glycoclusters have been important for defining the structural features required for high-affinity binding to multivalent receptors. As discussed above, rigid miniclusters are particularly affine ligands – if the carbohydrates are properly oriented, enabling unstrained multidentate binding. Rigid miniclusters are, in principle, moreover, able to differentiate between various multivalent lectins with the same carbohydrate specificity but varying orientation of their binding sites. If, however, the 3D structure of the targeted lectin is unknown, large numbers of potential ligands have to be synthesized and screened to identify the required presentation of the sugar residues.

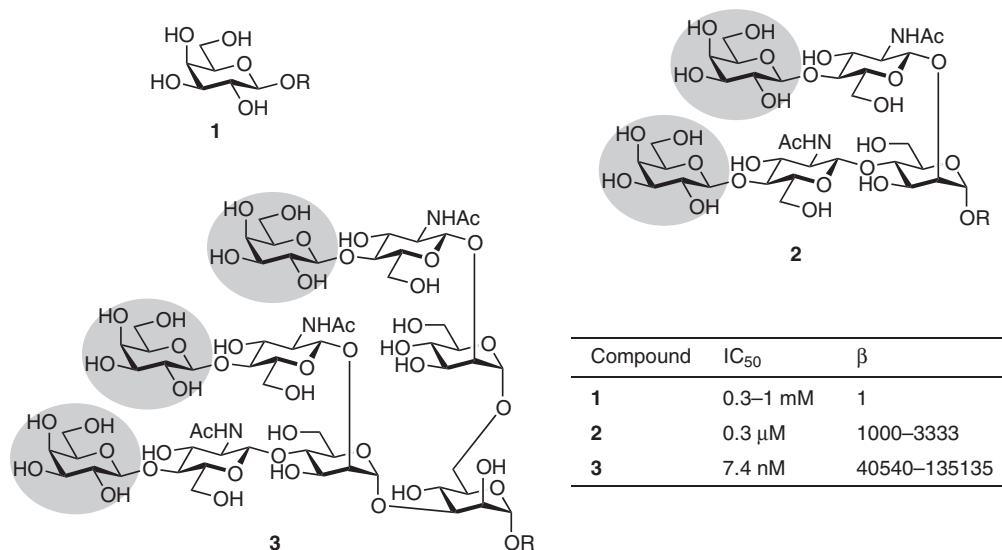
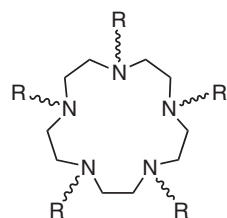


Fig. 2.7.3. Cluster glycosides synthesized by Lee et al. and their binding to ASGPR on isolated rabbit hepatocytes [7]. Carbohydrate residues involved in multivalent binding are highlighted in gray. The residual carbohydrates are assumed to function as a scaffold.

In pioneering investigations, Lee and coworkers have synthesized many glyco-clusters employing different scaffolds during the development of high-affinity ligands for the asialoglycoprotein receptor (ASGPR) on intact hepatocytes and in the soluble form [4, 7, 12]. A remarkable series of compounds with increasing valency is shown in Figure 2.7.3. The IC₅₀ values for inhibition of a radiolabeled ligand for ASGPR binding to hepatocytes, which under the conditions used approach dissociation constants, were 0.3–1 mM for β-galactosides 1, 0.3 μM for divalent oligosaccharide 2, and 7.4 nM for trivalent oligosaccharide 3. The phenomenon that the binding affinity of an oligovalent glycocluster increases geometrically with a linear increase in the number of sugar residues was termed the “glycoside cluster effect” by Lee et al.

Two research groups used the known X-ray structures of the heat-labile enterotoxin and the shiga-like toxin from *E. coli* to design inhibitors for these members of the AB₅ family of bacterial toxins. The most potent ligand of the heat-labile enterotoxin 7 prepared by Fan et al. had an IC₅₀ value of 560 nM, corresponding to an enhancement of β = 10⁵ compared with monomeric galactose (Figure 2.7.4) [13]. Bundle and coworkers designed the decavalent glycocluster 8, named Starfish, with an IC₅₀ value of 0.4 nM being comparable with the estimated affinity of the native ganglioside–pentamer interaction (Figure 2.7.5) [14]. Interestingly, the crystal structure of the toxin–inhibitor complex showed that each Starfish molecule was complexed by two toxin pentamers; this was, however, not planned.



Compound	n	IC ₅₀ (μM)	β
galactose		58000	1
4	1	242	240
5	2	16	3600
6	3	6	10000
7	4	0.56	104000

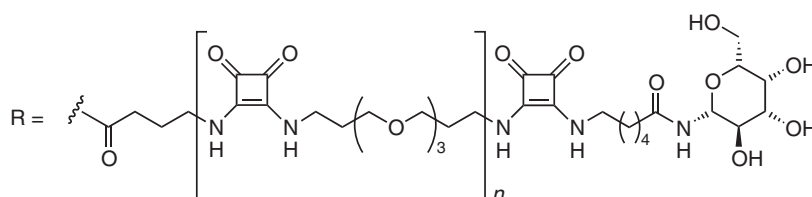


Fig. 2.7.4. Pentameric inhibitors of the heat-labile enterotoxin from *E. Coli* [13].

2.7.4

Spatial Screening of Lectin Ligands

To accelerate the process of finding the required presentation of carbohydrates when no structural information on the receptor is available, we developed a screening procedure for multivalent lectin ligands comprising four steps [15–17]:

1. split-mix synthesis (Box 11) of a library of scaffold molecules containing side-chain amino groups in varying amounts and spatial orientation;
2. attachment of several copies of a carbohydrate ligand to the amino groups;

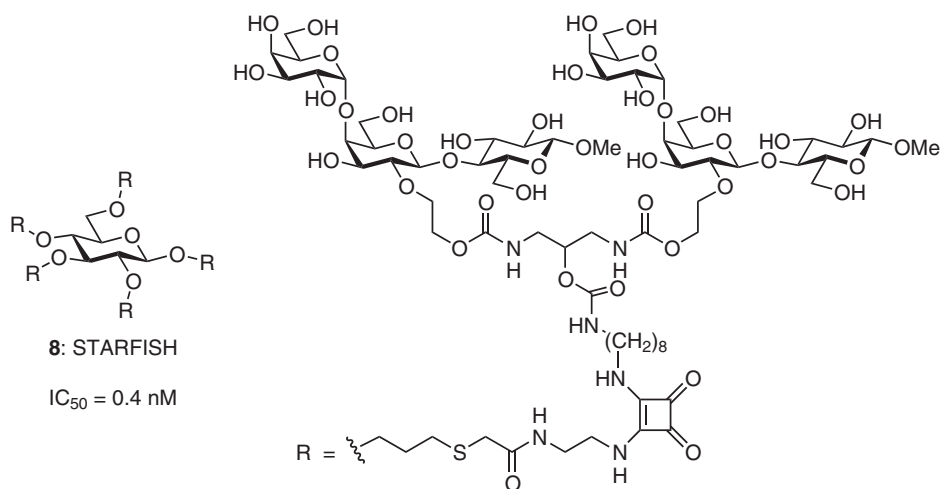


Fig. 2.7.5. Potent decameric inhibitor of the shiga-like toxin from *E. Coli* [14].

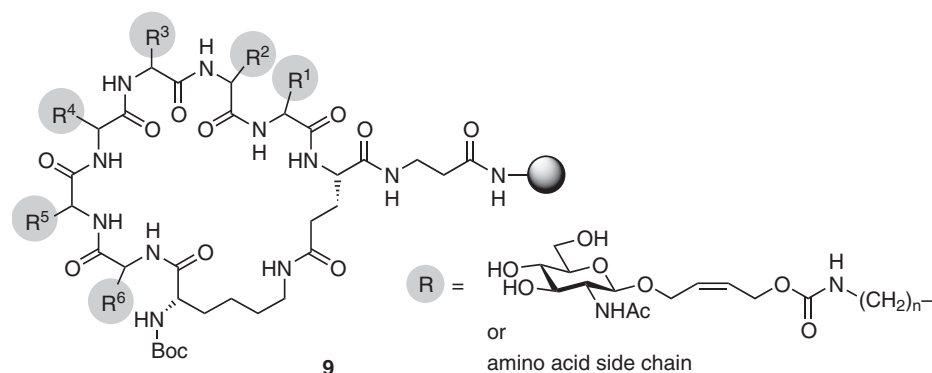


Fig. 2.7.6. Structure of multivalent neoglycopeptides used for spatial screening of lectin ligands [15, 17].

3. on-bead screening of the library for lectin-binding properties; and
4. identification of potent ligands by single-bead analysis.

2.7.4.1

Design and Synthesis of a Library of Cyclic Neoglycopeptides

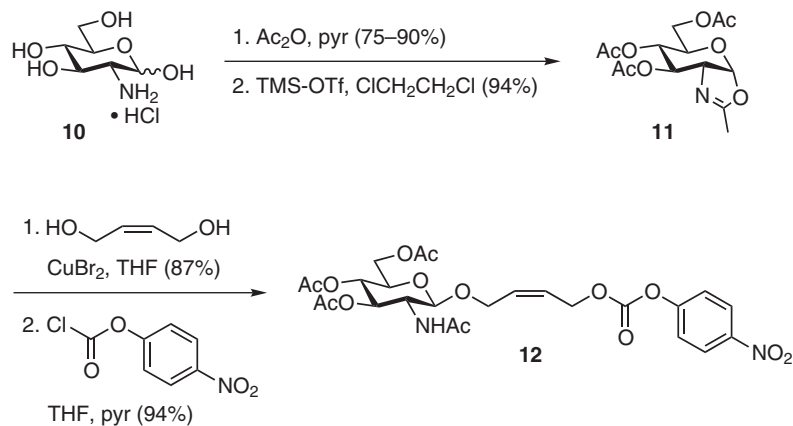
As scaffolds for the multivalent presentation of carbohydrate ligands we chose cyclic peptides of general type **9** (Figure 2.7.6) [15]. At the combinatorially varied positions indicated by gray circles, D- and L-amino acids without side-chain functionality and D- and L-diamino acids such as lysine, diaminobutyric acid, or diaminopropionic acid are incorporated. The latter represent the points of attachment of the carbohydrates. This library design enables generation of spatial diversity in two dimensions. Positional diversity generates different carbohydrate patterns displayed on the scaffolds. Varying the stereochemistry of the amino acids increases spatial diversity by generating different backbone folds [18].

For attachment of the carbohydrates a new urethane-type linker based on the Alloc protecting group has been developed (Scheme 2.7.1) [15, 19] (for other linkers used in solid-phase synthesis see Chapter 6.1). In contrast with glycosylation reactions employing solid phase-bound peptides [20], the formation of an urethane bond proceeds in virtually quantitative yield. Scheme 2.7.2 shows the convergent solid-phase peptide synthesis of the 19,440 compounds containing library **15** of cyclic neoglycopeptides [17]. N-acetylglucosamine (GlcNAc) residues were attached to side-chain amino groups by employing active carbonate **12**. The carbohydrate content of the library members ranges from 0 (2.6% of all compounds) through 1 (14.5%), 2 (30.3%), 3 (30.9%), 4 (16.6%), 5 (4.5%) and 6 (0.5%).

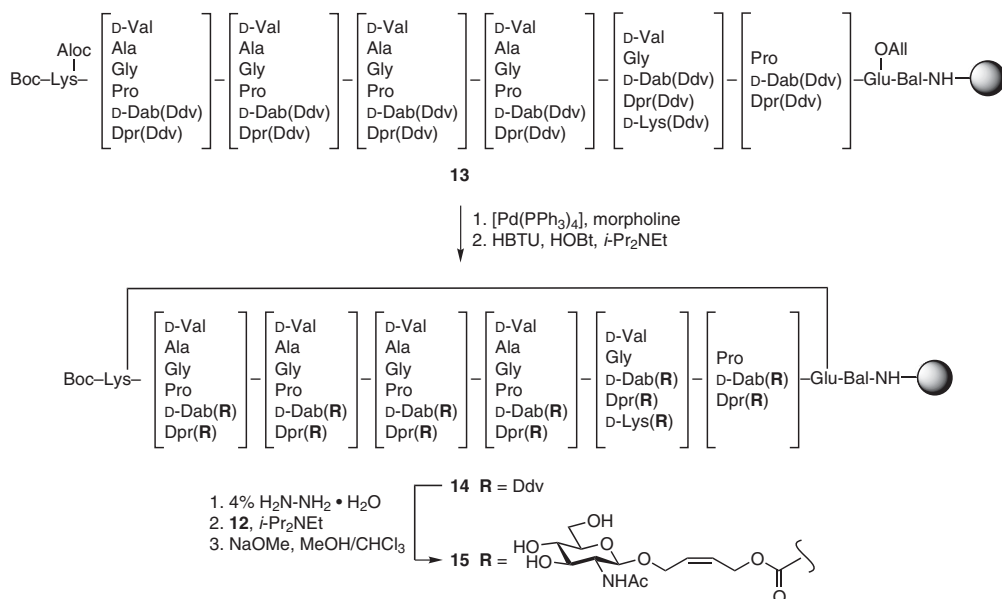
2.7.4.2

On-bead Screening and Ligand Identification

Library **15** was screened for binding properties to wheat germ agglutinin (WGA) [17]. WGA is a 36 kDa lectin composed of two glycine- and cysteine-rich subunits.



Scheme 2.7.1. Synthesis of the activated carbonate **12** for high-yield modification of side-chain amino groups of diamino acids [15, 19].



Scheme 2.7.2. Convergent solid-phase peptide synthesis of the library **15** of cyclic neoglycopeptides on amino-functionalized TentaGel without employing a linker [17]. (All = allyl, Aloc = allyloxycarbonyl, Bal = β -alanine, D-Dab = D-2,4-diaminobutyric acid, Ddv =

1-(4,4-dimethyl-2,6-dioxocyclohexylidene) isovaleryl, Dpr = L-2,3-diamino propionic acid, HBTU = O-benzotriazol-1-yl-N,N,N',N'-tetramethyluronium hexafluorophosphate, HOBT = 1-hydroxybenzotriazole.)

- 16: Boc-Lys-D-Dab(R)-D-Val-D-Dab(R)-D-Dab(R)-D-Val-D-Dab(R)-Glu-Bal-NH₂
- 17: Boc-Lys-Ala-Ala-Dpr(R)-D-Dab(R)-D-Dab(R)-Dpr(R)-Glu-Bal-NH₂
- 18: Boc-Lys-D-Dab(R)-Pro-Dpr(R)-D-Dab(R)-D-Lys(R)-Dpr(R)-Glu-Bal-NH₂
- 19: Boc-Lys-Dpr(R)-D-Dab(R)-D-Dab(R)-D-Dab(R)-D-Lys(R)-Dpr(R)-Glu-Bal-NH₂
- 20: Boc-Lys-D-Dab(R)-Dpr(R)-D-Dab(R)-D-Dab(R)-D-Dab(R)-D-Dab(R)-Glu-Bal-NH₂
- 21: Boc-Lys-D-Dab(R)-D-Dab(R)-D-Dab(R)-D-Dab(R)-D-Dab(R)-Dpr(R)-Glu-Bal-NH₂
- 22: Boc-Lys-D-Dab(R)-D-Dab(R)-D-Dab(R)-D-Dab(R)-D-Lys(R)-D-Dab(R)-Glu-Bal-NH₂

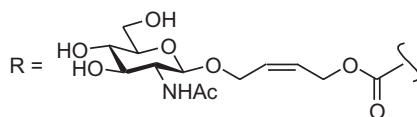


Fig. 2.7.7. Identified WGA ligands from spatial screening [17].

Each subunit contains four carbohydrate binding sites for *N*-acetylglucosamine and oligomers thereof, thus WGA is a promising candidate for multivalent interaction. Briefly, the resin-bound neoglycopeptides were incubated with biotinylated WGA followed by addition of an anti-biotin alkaline-phosphatase conjugate. Beads with bound lectin were detected by means of an alkaline phosphatase-catalyzed color reaction. When the assay was conducted in the presence of a competing monovalent ligand (GlcNAc), a small part (approx. 0.1%) of the beads stained very darkly. These beads were manually selected under a microscope and treated with [Pd(PPh₃)₄]/morpholine to remove the carbohydrates. After cleavage of the *N*-terminal Boc protecting group, “hit” structures were identified by automated single-bead Edman degradation.

Figure 2.7.7 shows the WGA ligands 16–22 identified from the screening process. If the binding assay responds to multivalency, it is expected to find glyco-clusters with a large amount of GlcNAc residues. For compounds 16–22 with four to six sugars each, this is clearly the case. Interestingly, not all hexavalent glycopeptides contained in library 15 (0.5%) led to stained beads. Furthermore, beside the hexavalent compounds, one pentavalent and two tetravalent compounds which caused similar staining were identified.

To quantify their binding to WGA in solution, glycopeptides 16, 18, and 22 were re-synthesized as single compounds. IC₅₀ values for inhibition of the binding of porcine stomach mucin to peroxidase-labeled WGA (Table 2.7.1) were determined by an enzyme-linked lectin assay (ELLA) as described by Zanini and Roy [21]. The IC₅₀ values shown in Table 2.7.1 confirm that the binding behavior of 16, 18, and

Tab. 2.7.1. Inhibition of binding of porcine stomach mucin to peroxidase-labeled WGA by neoglycopeptides **16**, **18**, and **22**.

<i>Compound</i>	<i>IC₅₀ (mM)</i>	<i>β</i>
GlcNAc	83	1
16	0.381	218
18	0.134	619
22	0.146	568

22 in solution is similar to that on the solid phase, so the result of the on-bead screening was not a surface effect, as has been observed by Kahne et al. with an immobilized oligosaccharide library [22].

The fact that **18** and **22** have similar IC_{50} values illustrates that the binding affinity is not only determined by the number of GlcNAc residues but also by the ligand architecture. The β values in Table 2.7.1 are the highest ever reported for oligovalent GlcNAc clusters of that size binding to WGA. With the presented convergent synthetic strategy it is possible to attach any desired carbohydrate ligand to a once-prepared cyclopeptide library, enabling rapid screening of different lectins.

2.7.5

Conclusion

Multivalency provides a means of enhancing the affinity of weak carbohydrate–lectin interactions. High-affinity lectin ligands can function as inhibitors or effectors of carbohydrate-mediated biological processes such as the inflammatory cascade or microbial adhesion to host cells. If the three-dimensional structure of the targeted lectin is known, multivalent carbohydrate ligands may be obtained by rational drug design. Spatial screening of multivalent lectin ligands on the other hand enables rapid identification of ligands with enhanced binding affinity without knowledge of the lectin structure.

References

- 1 R. A. DWEK, *Chem. Rev.* **1996**, 96, 683–720; A. VARKI, R. CUMMINGS, J. ESKO, H. FREEZE, G. HART, J. MARTH, eds., *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, **1999**; R. S. HALTIWANGER, *Curr. Opin. Struct. Biol.* **2002**, 12, 593–598.
- 2 J. M. RINI, *Annu. Rev. Biophys. Biomol. Struct.* **1995**, 24, 551–577; K. DRICKAMER, *Curr. Opin. Struct. Biol.* **1999**, 9, 585–590; R. LORIS, *Biochim. Biophys. Acta* **2002**, 1572, 198–208.
- 3 B. T. HOUSEMAN, M. MRKSICH, *Top. Curr. Chem.* **2002**, 218, 1–44.
- 4 Y. C. LEE, R. T. LEE, *Acc. Chem. Res.* **1995**, 28, 321–327.
- 5 J. ALPER, *Science* **2001**, 291, 2338–2343; H. RÜDIGER, H.-C. SIEBERT, D. SOLIS, J. JIMENEZ-BARBERO, A.

- ROMERO, C.-W. VON DER LIETH, T. DIAZ-MAURINO, H.-J. GABIUS, *Curr. Med. Chem.* **2000**, *7*, 389–416.
- 6 J. J. LUNDQUIST, E. J. TOONE, *Chem. Rev.* **2002**, *102*, 555–578; T. K. LINDHORST, *Top. Curr. Chem.* **2002**, *218*, 201–235; L. L. KIESSLING, T. YOUNG, K. H. MORTELL, *Glycoscience: Chemistry and Chemical Biology*, Vol. II, B. FRASER-REID, K. TATSUTA, J. THIEM, eds., Springer, Heidelberg, **2001**, pp. 1817–1861; R. ROY, *Top. Curr. Chem.* **1997**, *187*, 241–274.
- 7 R. T. LEE, Y. C. LEE, *Glycoconjugate J.* **2001**, *17*, 543–551.
- 8 M. MAMMEN, S.-K. CHOI, G. M. WHITESIDES, *Angew. Chem.* **1998**, *110*, 2908–2953.
- 9 J. RAO, J. LAHIRI, L. ISAACS, R. M. WEIS, G. M. WHITESIDES, *Science* **1998**, *280*, 708–711; J. RAO, J. LAHIRI, R. M. WEIS, G. M. WHITESIDES, *J. Am. Chem. Soc.* **2000**, *122*, 2698–2710.
- 10 A. SCHON, E. FREIRE, *Biochemistry* **1989**, *28*, 5019–5024.
- 11 J. M. GARGANO, T. NGO, J. Y. KIM, D. W. K. ACHESON, W. J. LEES, *J. Am. Chem. Soc.* **2001**, *123*, 12909–12910.
- 12 Y. C. LEE, R. T. LEE, eds., *Neoglycoconjugates. Preparation and Applications*, Academic Press, San Diego, **1994**.
- 13 E. FAN, Z. ZHANG, W. E. MINKE, Z. HOU, C. L. M. J. VERLINDE, W. G. J. HOL, *J. Am. Chem. Soc.* **2000**, *122*, 2663–2664.
- 14 P. I. KITOV, J. M. SADOWSKA, G. MULVERY, G. D. ARMSTRONG, H. LING, N. S. PANNU, R. J. READ, D. R. BUNDLE, *Nature* **2000**, *403*, 669–672.
- 15 V. WITTMANN, S. SEEBERGER, *Angew. Chem.* **2000**, *112*, 4508–4512.
- 16 V. WITTMANN, S. SEEBERGER, H. SCHÄGGER, *Tetrahedron Lett.* **2003**, *44*, in press.
- 17 V. WITTMANN, S. SEEBERGER, *Angew. Chem.* **2003**, *115*, in press.
- 18 R. HAUBNER, D. FINSINGER, H. KESSLER, *Angew. Chem.* **1997**, *109*, 1440–1456.
- 19 V. WITTMANN, D. LENNARTZ, *Eur. J. Org. Chem.* **2002**, 1363–1367.
- 20 A. SCHLEYER, M. MELDAL, M. RENIL, H. PAULSEN, K. BOCK, *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1976–1978.
- 21 D. ZANINI, R. ROY, *Bioconjugate Chem.* **1997**, *8*, 187–192.
- 22 R. LIANG, L. YAN, J. LOEBACH, M. GE, Y. UOZUMI, K. SEKANINA, N. HORAN, J. GILDERSLEEVE, C. THOMPSON, A. SMITH, K. BISWAS, W. C. STILL, D. KAHNE, *Science* **1996**, *274*, 1520–1522.