JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Molecular Recognition of Complex-Type Biantennary *N*-Glycans by Protein Receptors: a Three-Dimensional View on Epitope Selection by NMR

Ana Ardá,[†] Pilar Blasco,[†] Daniel Varón Silva,[§] Volker Schubert,[§] Sabine André,[∥] Marta Bruix,[‡] F. Javier Cañada,[†] Hans-Joachim Gabius,[∥] Carlo Unverzagt,*[,§] and Jesús Jiménez-Barbero*[,†]

[†]Chemical and Physical Biology, Centro de Investigaciones Biológicas, CSIC, Ramiro de Maeztu 9, 28040 Madrid, Spain

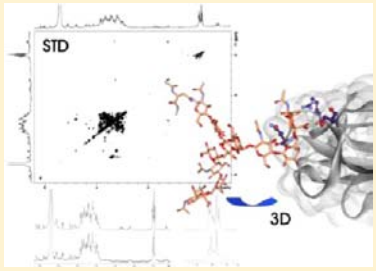
[‡]Instituto de Química-Física Rocasolano, CSIC, Serrano 119, 28006 Madrid, Spain

[§]Bioorganische Chemie, Gebäude NWI, Universität Bayreuth, 95440 Bayreuth, Germany

[∥]Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University Munich, Veterinärstr. 13, 80539 Munich, Germany

Supporting Information

ABSTRACT: The current surge in defining glycobiomarkers by applying lectins rekindles interest in definition of the sugar-binding sites of lectins at high resolution. Natural complex-type *N*-glycans can present more than one potential binding motif, posing the question of the actual mode of interaction when interpreting, for example, lectin array data. By strategically combining *N*-glycan preparation with saturation-transfer difference NMR and modeling, we illustrate that epitope recognition depends on the structural context of both the sugar and the lectin (here, wheat germ agglutinin and a single hevein domain) and cannot always be predicted from simplified model systems studied in the solid state. We also monitor branch-end substitutions by this strategy and describe a three-dimensional structure that accounts for the accommodation of the α2,6-sialylated terminus of a biantennary *N*-glycan by viscumin. In addition, we provide a structural explanation for the role of terminal α2,6-sialylation in precluding the interaction of natural *N*-glycans with lectin from *Maackia amurensis*. The approach described is thus capable of pinpointing lectin-binding motifs in natural *N*-glycans and providing detailed structural explanations for lectin selectivity.



INTRODUCTION

Molecular recognition is at the heart of essential biological events. The recognition of saccharides presented as part of natural scaffolds (protein, sphingolipids) by protein receptors (lectins) mediates key processes underlying many aspects of life as reflected by the term "sugar code".[1] Matching the wide spectrum of glycan structures (glycome complexity[2]), lectins are found in all branches of the phylogenetic tree and have proven to be versatile tools for glycan detection and structural characterization.[3] Indeed, there is nowadays particular interest in the identification of glycoprotein glycan structures as a result of altered glycoprotein expression occurring in diseases, hereby defining functional glycobiomarkers of diagnostic value.[4] Recent technological advances have led to the development of lectin and glycan microarrays, which offer a comparative view of the lectin−glycan interactions and provide a high-throughput approach to screen a wide range of glycan profiles and to determine glycan binding specificity of the lectin.[5]

With the growing realization that the presentation of a glycan epitope can affect its recognition by lectins,[3b] it is timely to study these interactions with natural glycans of the proper size and then rigorously define the contact sites, especially in cases with more than one possible interaction mode.

NMR and X-ray crystallography provide key information at atomic resolution on biomolecules, including complexes. However, in the protein−carbohydrate interaction field, these techniques have mainly focused on acquiring information of complexes formed by lectins with small/medium-sized fragments of natural glycan chains. Often, it is an open question how affinity/specificity is regulated on the level of complex glycans.

Herein, we have studied the recognition in solution of two naturally occurring complex-type *N*-glycans by different lectins commonly used for glycan detection and epitope identification, in search of biomarkers. This has been made possible due to the remarkable progress in both chemical synthesis of *N*-glycans[6] and powerful NMR-based techniques.[7] The study has yielded unexpected results for hevein and a multihevein-domain in terms of epitope recognition, revealing that the characteristics of glycan and protein may preclude extrapolations, for example, from work with small ligands.

N-Glycans are common modifications of membrane/secreted proteins conferring specific properties to the associated protein

Received: October 30, 2012
Published: January 29, 2013

dx.doi.org/10.1021/ja3104928 | *J. Am. Chem. Soc.* 2013, 135, 2667−2675

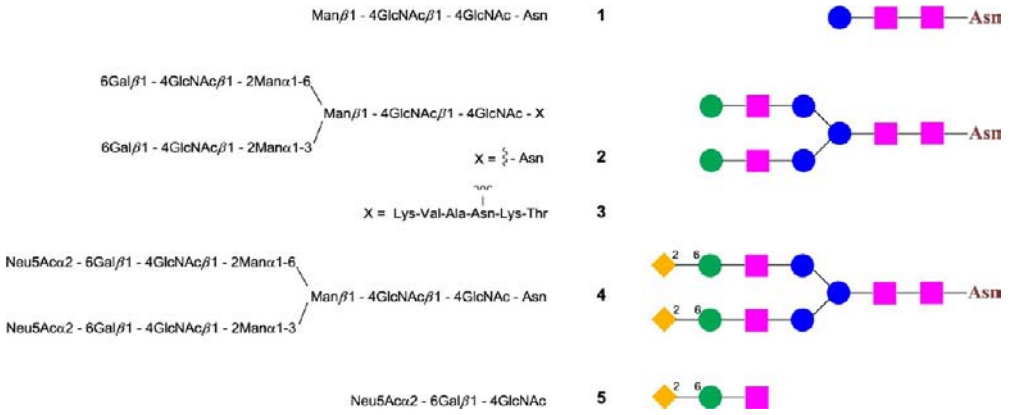


Figure 1. Glycan structures used in this study.

that will ultimately determine its fate. The structures of *N*-glycans are widely diverse but share common structural motifs. The core pentasaccharide is composed of a chitobiose unit and a trimannoside section introducing the branch point. Both antennae can be extended by *N*-acetyllactosamine (LacNAc) units terminated by α2,6- or α2,3-linked sialic acid. As abundant species among *N*-glycans, we have studied the biantennary complex-type nonasaccharide−Asn (with LacNAc termini) (**2**) and its α2,6-sialylated undecasaccharide−Asn (**4**), which were isolated from egg yolk (Figure 1). These structures represent common motifs of *N*-glycans on glycoproteins, with α2,6-sialylation acting as the salient signal for masking/docking.[8] From the receptor side, different lectins have been selected, which according to published data, should be able to bind different epitopes of natural *N*-glycans. In particular, two agglutinins described with binding affinity to the *N*,*N*′-diacetylchitobiose ($(GlcNAc)_2$) part of the core have been chosen, namely, wheat germ agglutinin (WGA)[9] and a WGA-derived single domain similar to hevein.[10] Hevein, one of the smallest plant lectins with 43 amino acids, was the first lectin whose structure in complex with its sugar ligand has been solved in solution.[11] Its reactivity is represented by the so-called hevein domain, present in WGA. This lectin indeed harbors eight hevein domains and displays a dual specificity for both GlcNAc and Neu5Ac,[9] which are monosaccharides found at opposing ends of natural *N*-glycans. Whether both sugars are reactive in an *N*-glycan or whether the lectin exhibits selectivity is not known, prompting us to perform this study.

The approach used can likewise be used to obtain structural view on target−site interaction. To illustrate this application, the *Viscum album* agglutinin (viscumin, VAA) was selected. It binds LacNAc termini on *N*-glycans and also in the case of α2,6-sialylated *N*-glycans.[13] As a negative control, a mixture of lectins, for whom α2,6-sialylgalactose terminus blocks recognition, was chosen. In this context, the two agglutinins from the seeds of *Maackia amurensis* (the leukoagglutinin MAL and the hemeagglutinin MAH) have been selected.[9,12]

We show that the mode of epitope presentation is crucial for glycan recognition and that the selectivity observed in assays with free sugars can change after subtle alterations of the *N*-glycan structure. The same can hold true for lectins differing in quaternary structure.

METHODS

NMR. The ^{1}H NMR resonances of the ligands were completely assigned through standard TOCSY (60 ms mixing time), NOESY (300 and 500 ms mixing time), and HSQC experiments on 800 MHz (cryo) and 600 MHz spectrometers. Typical concentrations were 1 mM for the homonuclear experiments and 2 mM for the heteronuclear experiments.

The samples for saturation-transfer difference (STD) experiments as adapted to the individual lectins[14] were prepared in phosphate-buffered saline (pH = 5.7 for MAA and WGA, and pH = 7.3 for viscumin) using ligand/lectin ratios varying from 1:20 to 1:100. The applied temperatures varied between 283 and 303 K. Molar ratio and temperature were optimized in each case. Representative experiments with significant STD responses are presented in the figures.

Data from one-dimensional (1D) STD experiments were acquired at 600 and 500 MHz using a Gaussian pulse (49 ms) cascade separated by 1 ms delays. STD-TOCSY experiments were acquired at 500 MHz, using 256 increments and a isotropic mixing time of 60 ms. In all cases, the on-resonance frequency was set at the aromatic or aliphatic regions (6.9 or 0.5 ppm, respectively), and a series of experiments were recorded, where the saturation time was varied between 200 ms and 2 s. The off-resonance frequency was always set at 100 ppm. Under these conditions, the free ligand in solution showed residual STD intensities in the 1D STD spectra that were taken into account when analyzing the STD spectra.

Molecular Modeling. The templates for the modeling procedures were built based on X-ray structures of WGA complexed with small oligosaccharides (pdb codes: 2UVO and 2CWG), 1DBN for the *M. amurensis* isolectin and 1PUU for viscumin. The structure of the *N*-glycan was superimposed in the binding site working with the most populated conformation found for the free state (according to a standard NOE/molecular modeling approach). The complex structure was then submitted to a short molecular dynamics (MD) run, followed by energy minimization with a low gradient convergence threshold (0.02) in 5000 steps. In all cases, the OPL2005 force field[15] was employed, as integrated in the Schroedinger MAESTRO suite of programms.[16]

Lectins. WGA and MAA were purchased from Sigma−Aldrich, VAA was obtained and controlled for purity as described,[17] and WGA-B domain was also obtained as previously described.[18]

Ligands. Compound **5** was purchased from Sigma−Aldrich. Compounds **1**,[19] **2**, and **4** were obtained as previously described.[20] Compound **3** was prepared from its corresponding sialylated undecasaccharide−glycopeptide, obtained from egg yolk in analogy to published procedures.[21] This undecasaccharide−glycopeptide (5.7 mg, 2 μmol) was dissolved in 400 μL of phosphate buffer (75 mM; pH 6), and a solution of 0.2 mg (0.6 U) of neuraminidase from *Clostridium perfringens* (Sigma−Aldrich, EC: 3.2.1.18) in 50 μL of phosphate buffer (75 mm, pH 6) was added. The mixture was incubated at ambient temperature for 3 days (TLC: 2-propanol/1 M ammonium acetate 1.5:1) and cleared by centrifugation, and the supernatant was purified by gel filtration (Superdex 30 (1.6 cm × 60 cm); flow rate, 1 mL/min; eluent, 0.1 *m* NH_4HCO_3; detection, 214 nm). The fractions eluting at 77 min were collected, lyophilized, and desalted by gel filtration

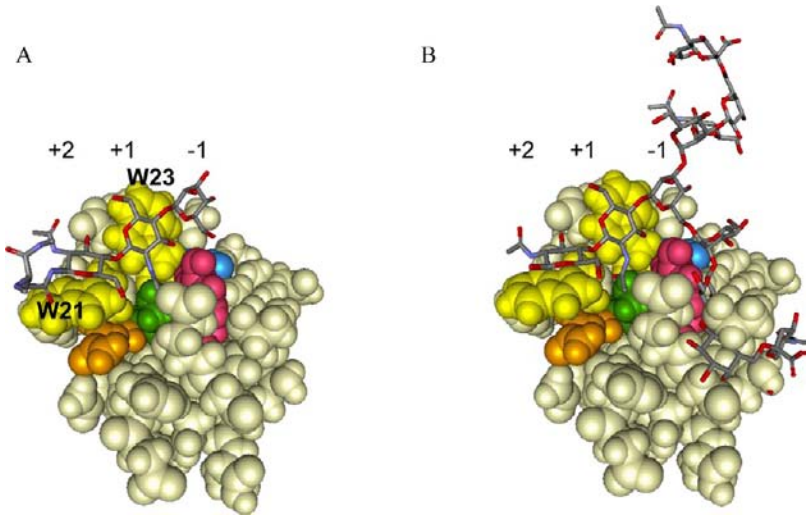


Figure 2. (A) Complex of hevein with the trisaccharide core (**1**) as deduced from NMR.[19] (B) Attempted docking of the extended *N*-glycan structure **4** obtained by superimposition of the sialylated *N*-glycan in the trisaccharide core template (see panel A). The extension at position O6 of the central Man moiety in the core causes steric clashes with the protein. The *gg* orientation is shown at Man C5−C6. The *gt* rotamer is sterically also forbidden. Subsites are indicated on top, and the key amino acids, i.e., Trp21 and Trp23, are annotated.

(Sephadex G25 (2.5 cm × 70 cm); flow rate, 1 mL/min; eluent, 5% ethanol in water; detection 214 nm). The fractions eluting at 201 min were lyophilized and yielded 4.05 mg of nonasaccharide−glycopeptide **3** (98.2%)

RESULTS

Epitope Selection of Monomeric versus Multimeric Lectins (Hevein Domain and WGA) between Two Potential Contact Sites. WGA constitutes the classical model for chitooligosaccharide binding. Additionally, WGA is often included among the sialic acid binding lectins. Therefore, two potential epitopes are presented in sialylated glycopeptides for this lectin. In fact, the recognition features of chitooligosaccharides by hevein domains have been well established through NMR.[10,11] More recently, we have demonstrated that the trisaccharide *N*-glycan core (**1**) is also recognized by hevein, with GlcNAc1 (linked to Asn) on subsite +2 (on top of W21), GlcNAc2 on subsite +1 (on top of W23), and *β*Man at subsite −1, leaving the Asn residue outside of the binding site, without establishing interactions with the lectin.[19] This structural model was used as a template for analyzing the recognition of complex-type *N*-glycans. However, when we titrated a 0.3 mM solution of a single hevein domain (the B domain of WGA[18]) with increasing amounts of the sialylated *N*-glycan−Asn **4** (Figure S1, Supporting Information), no changes were observed for the ^{1}H NMR resonances of the lectin. Similarly, attempts to dock the undecasaccharide−Asn **4** using the binding mode of the core trisaccharide−Asn (**1**) (Figure 2) were unsuccessful. In fact, branching at position 6 of the central *β*Man residue causes a steric clash of the sugar extension with different protein residues, independent of the presence of either *gg* or *gt* rotamers at the C5−C6 linkage. In contrast, WGA is known to bind *N*-glycans.[9]

WGA is a dimeric lectin where each subunit consists of an assembly of four hevein domains.[22] Therefore, WGA is composed of eight hevein domains. It has been described to be specific for terminal Neu5Ac and GlcNAc moieties by inhibition studies,[9] crystallography,[22] and NMR spectroscopy.[23] The structure of complexes, also with a small sialoglycopeptide,[22c] reveals that the primary binding site is constituted by three conserved aromatic residues from one hevein domain but further polar residues (S114 and E115) from the neighboring hevein domain contribute to the stabilization of the complex, providing structural evidence for an intraprotein cooperation of lectin domains.

Strikingly, and in contrast with the observations for the single hevein domain, the STD experiments performed with the sialylated *N*-glycan−Asn **4** in the presence of the multidomain lectin showed STD signals, evidencing the existence of interaction. Analysis of the STD spectra (Figure 3) clearly showed that the terminal Neu5Ac residue is not the key point in the recognition. Only a very weak effect is observed for the methyl group of the 5-acetamide, while the ring hydrogens do not show up in the STD spectrum (Figures 3 and S3, Supporting Information). In fact, the residues in closer contact with the protein are at the opposite end of the molecule, namely, GlcNAc1 and Asn. Thus, the binding mode is different from that occurring between a single hevein domain (Figure 2A) and the trisaccharide core (**1**), in which the Asn moiety was far from the lectin (positive NOEs in the NOESY spectrum of the complex).[19] For hevein itself, GlcNAc1 and GlcNAc2 are located at subsites +2 and +1, respectively. However, for WGA, the docking/minimization procedure based on the experimental observations showed that the binding mode is shifted by one subsite: the Asn residue now occupies subsite +2 (on top of W21); GlcNAc1 is on subsite +1 (on top of W23), and GlcNAc2 sits on subsite −1. In this situation, the *β*Man residue is shifted one subsite away from the center of binding to the protein, and therefore, the typical branching at O6 does not cause steric clashes with the protein (Figure 4). Obviously, the following question arises: why does the single hevein domain not use this binding mode for recognizing the large *N*-glycan **4**? Detailed comparisons point to the presence of the additional polar residues in the neighboring hevein domain in WGA; that is, S114 and E115 appear as key factors stabilizing the complex. They form hydrogen bonds with the *N*-glycan (Figure 4), which are absent in the single hevein domain. Probably, this binding mode (GlcNAc1−Asn on subsites +1 and +2, respectively) is energetically not favorable in the case of just

dx.doi.org/10.1021/ja3104928 | J. Am. Chem. Soc. 2013, 135, 2667−2675

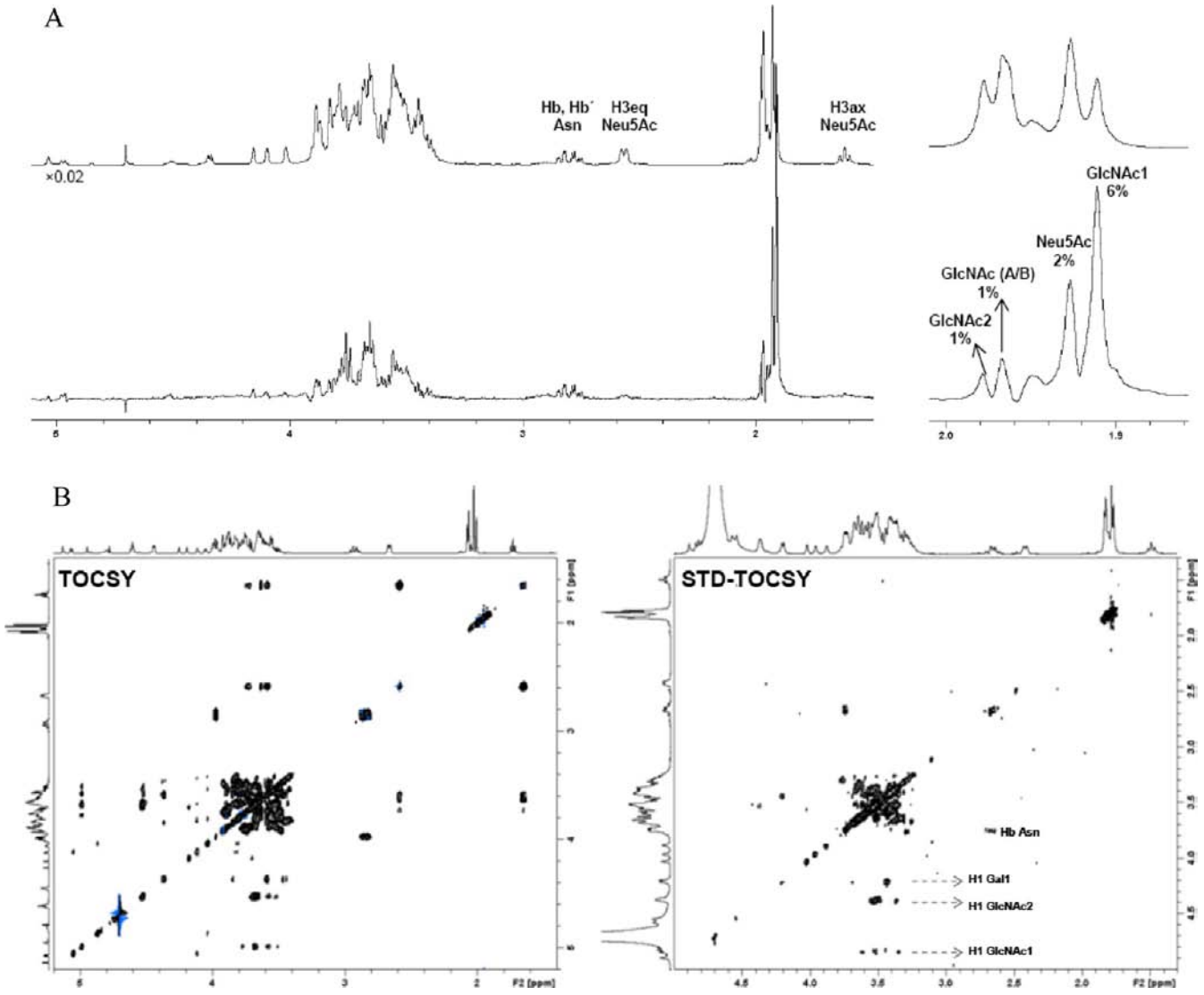


Figure 3. (A) STD (below) and off-resonance spectrum (on top) of the sample containing WGA and the sialylated *N*-glycan **4**. The region of the methyl protons of the Ac groups is enlarged on the right. (B) left, TOCSY of ligand **4**; right, STD-TOCSY of the mixture of WGA and sialylated *N*-glycan **4**. The sialyl residue does not contribute to the binding event. In contrast, the Asn residue provided clear STD signals indicating its direct involvement in the recognition process.

a single hevein domain,[19] where these new interactions cannot contribute to the binding enthalpy.

Analogous STD experiments were carried out with the non-sialylated *N*-glycan–Asn **2** (Figure 5), which were a control to ascertain the binding mode. The results were very similar to those obtained with **4**: Asn and GlcNAc1 display the strongest STD intensities, with lower intensities for the branches.

As already mentioned, WGA is often included among the sialic acid binding lectins. However, our experimental STD data have shown that, although **4** is indeed recognized by WGA, the terminal sialic acid of the undecasaccharide does not establish primary interactions with the lectin. Instead, the interactions occur via the GlcNAc1 and Asn moieties at the stem region of the glycan.

To prove reactivity to a sialylated but GlcNAc-free compound, one additional STD experiment with 6′-sialyl-LacNAc (**5**) was acquired. The STD spectrum (Figure S2, Supporting Information) clearly showed that only protons from Neu5Ac show STD intensity, especially those from the acetyl group, in agreement with the X-ray crystallographic data.[22] The other STD signals correspond to H5, H6, H7, and H4 of Neu5Ac. Thus, when presented in the context of an *N*-glycan, the sialic acid is therefore not a primary contact site, as it is for the branch-end trisaccharide **5**. Explicitly, in the **4**–WGA complex, only the acetyl group of Neu5Ac showed weak STD, and protons H5, H6, and H7 did not show STD signals (Figure S3, Supporting Information).

This result underscores the inherent value of the given approach to team up binding-site identification by STD NMR spectroscopy with *N*-glycan preparation. When more than one possible contact site is presented, the given strategy can determine the preferred epitope in the context of the *N*-glycan.

Since the recognition of **2** and **4** by WGA mainly involves GlcNAc1 and the Asn residue, it remained to be clarified whether there is an influence of a peptidic moiety, as occurring in glycoproteins. Thus, an STD experiment was carried out under the same experimental conditions for a mixture of glycopeptide **3** and WGA. The spectrum showed STD intensities similar to those obtained for **2** (Figure S4, Supporting Information). In particular, the protons of the

dx.doi.org/10.1021/ja3104928 | J. Am. Chem. Soc. 2013, 135, 2667–2675

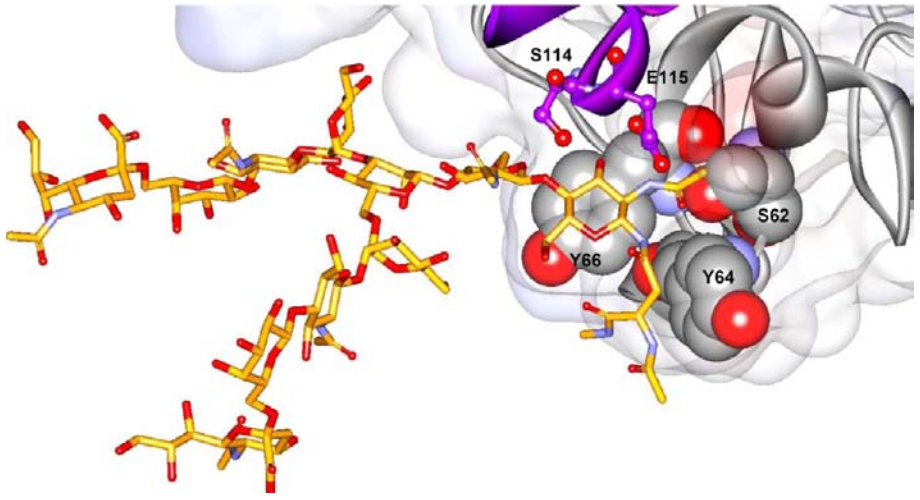


Figure 4. Deduced complex of the sialylated *N*-glycan (**4**, in orange) in the primary binding site of WGA, as deduced from the docking, minimization, and MD process. Amino acids Y66, Y64, and S62 (equivalents to W21, W23, and S19 in hevein) from one hevein domain are represented in gray CPK, while S114 and E115 from the neighboring hevein domain are represented in balls and sticks, in purple. The sialyl residues are exposed to the solvent and not directly involved in WGA binding.

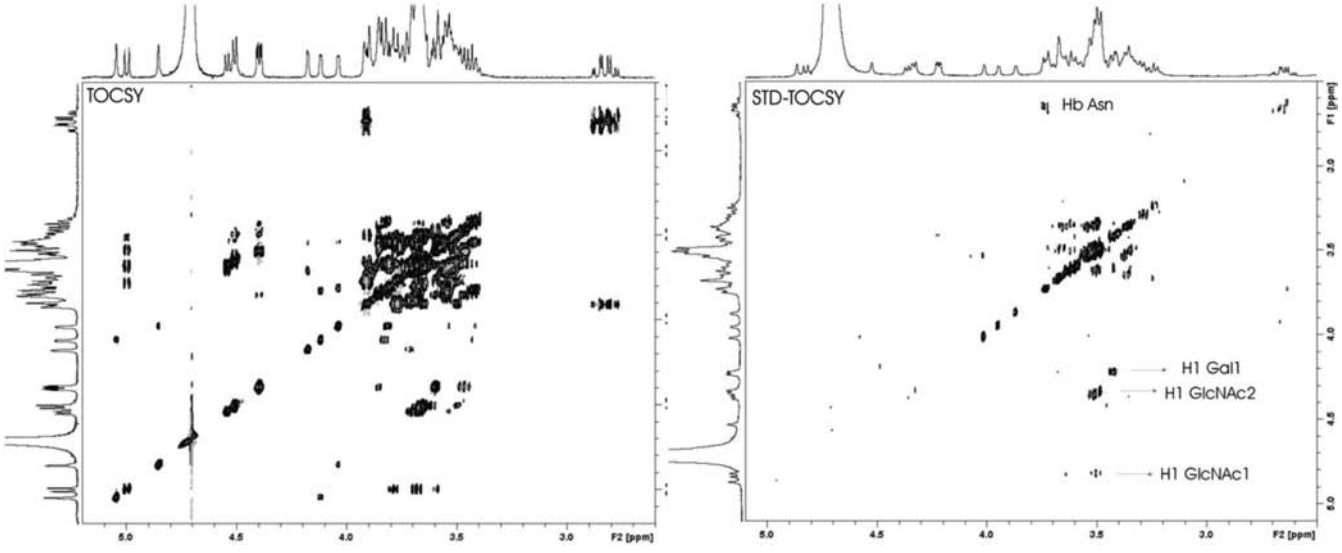


Figure 5. (left) TOCSY spectrum of non-sialylated *N*-glycan (**2**). (right) STD-TOCSY of WGA/non-sialylated *N*-glycan, **2**. As for **4**, the Asn and GlcNAc1 residues provided clear STD signals indicating their direct involvement in the recognition process.

additional peptide residues did not show up in the spectrum, indicating exclusive recognition of **3** by WGA through the GlcNAc−Asn region. We next proceeded to a case study of actual branch-end recognition.

Recognition of the Terminal Nonreducing End: Viscumin. This lectin is a toxic AB-type protein, which in addition to the A chain with rRNA *N*-glycosidase activity, contains a B chain with lectin activity to galactosides. The B chain harbors two carbohydrate-binding sites, characterized by W38 and Y249. In solution, the lectin activity of VAA can be attributed primarily to the Tyr site.[27] The recognition mode of lactose by VAA has been characterized by X-ray crystallography, chemical mapping, and modeling.[28] It has been reported that VAA strongly binds Neu5Acα2−6Galβ1−4GlcNAc on gangliosides and glycoproteins, while gangliosides with terminal galactoses are poorly recognized.[29] α2,6-Sialylation can improve inhibitory capacity of lactose.[13b] Nevertheless, a structural rationalization to those facts has not yet been given.

Figure 6 shows the STD spectra, along with their corresponding reference spectra, for VAA in the presence of the sialylated *N*-glycan **4** (on top) and of the non-sialylated analogue **2** (below). Both compounds are bound by VAA, but the STD intensities were significantly stronger for **4** than for **2**. Since strictly the same experimental conditions were kept, the results suggest respective grading of interaction for VAA. Furthermore, the distinct protons of the Neu5Ac residue (H3ax, H3eq, and Ac) clearly revealed that this residue is in close contact with VAA. This fact was confirmed by the STD-TOCSY spectrum (Figure S5, Supporting Information). Thus, in this case, the terminal Neu5Ac moiety is clearly involved in the recognition process, and its absence weakens but does not prevent glycan−ligand interaction.

Three-dimensional (3D) structures of the corresponding complexes were generated by using molecular modeling techniques, taking the available experimental X-ray coordinates for the VAA−lactose complex as a starting geometry. In the

dx.doi.org/10.1021/ja3104928 | *J. Am. Chem. Soc.* 2013, 135, 2667−2675

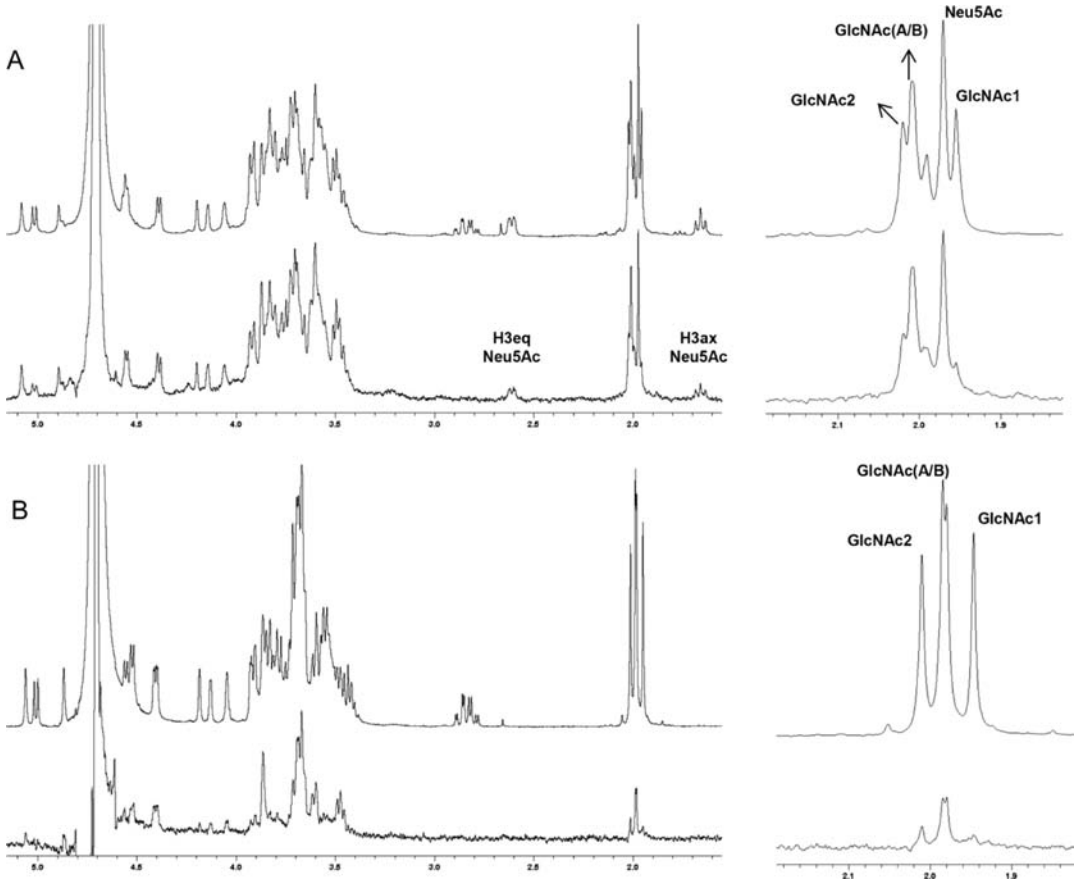


Figure 6. STD spectra of the mixtures of VAA with (A) **4** and (B) **2**. The methyl groups regions are enlarged at the right-hand side, and the resonances are annotated for each case. The STD spectra are shown below, while the corresponding off-resonance spectra are on top. In all cases, VAA binds both molecules (see the Discussion section).

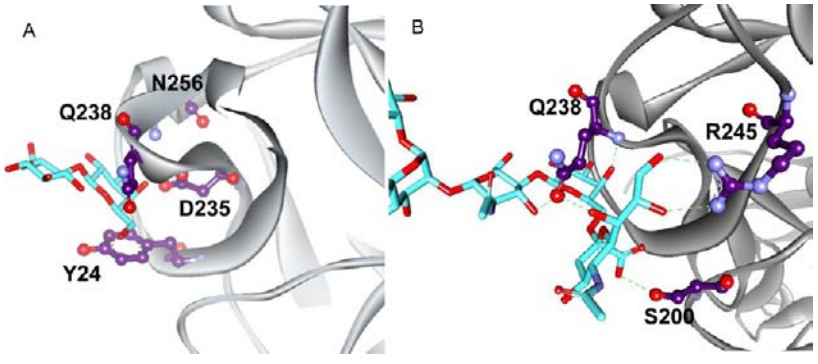


Figure 7. (A) X-ray structure (pdb code 1PUU) of VAA complexed with lactose. The amino acids implicated in interactions with the ligand are highlighted. (B) Docking pose of the sialylated *N*-glycan **4** on the Tyr site in subdomain 2γ. The amino acids providing new further interactions (hydrogen bonds depicted in green) with the Neu5Ac residues are highlighted.

resulting 3D structure, the galactose-binding site is formed by amino acids D335, N256, and Q238 that establish hydrogen bonding with the Gal residue and by Y249, which establishes stacking interaction with the α-face of the Gal moiety, explicitly with H3, H4, and H5 (Figure 7A). Inspection of this experimental geometry detects space for a sugar extension at Gal O6. This location is surrounded by polar amino acids that may efficiently interact with sialic acid residues; they are S200, R245, and Q238. Docking of **4** at this locus, using the lactose-binding model as a template, provided a stable pose, which was further optimized through an energy minimization and MD protocol. Of note, further interactions can then take place between Neu5Ac and the above-mentioned polar residues (Figure 7B): R245 establishes a bidentate hydrogen bond with the glycerol chain of the sialic acid residue, S200 is hydrogen bonded to the carboxylic acid, and Q238 establishes a cooperative hydrogen bond with OH7 of Neu5Ac and OH3 of the branching GlcNAc. Evidently, these interactions appear to underlie the preferential affinity reported above. Terminal LacNAc units are recognized, but further sialylation to Neu5Acα2−6LacNAc enhances binding.

Precluding Recognition by Terminal α2,6-Sialylation: *M. amurensis* Lectins. The mixture of the two lectin species from the seeds of *M. amurensis*, MAL and MAH, was

dx.doi.org/10.1021/ja3104928 | *J. Am. Chem. Soc.* 2013, 135, 2667−2675

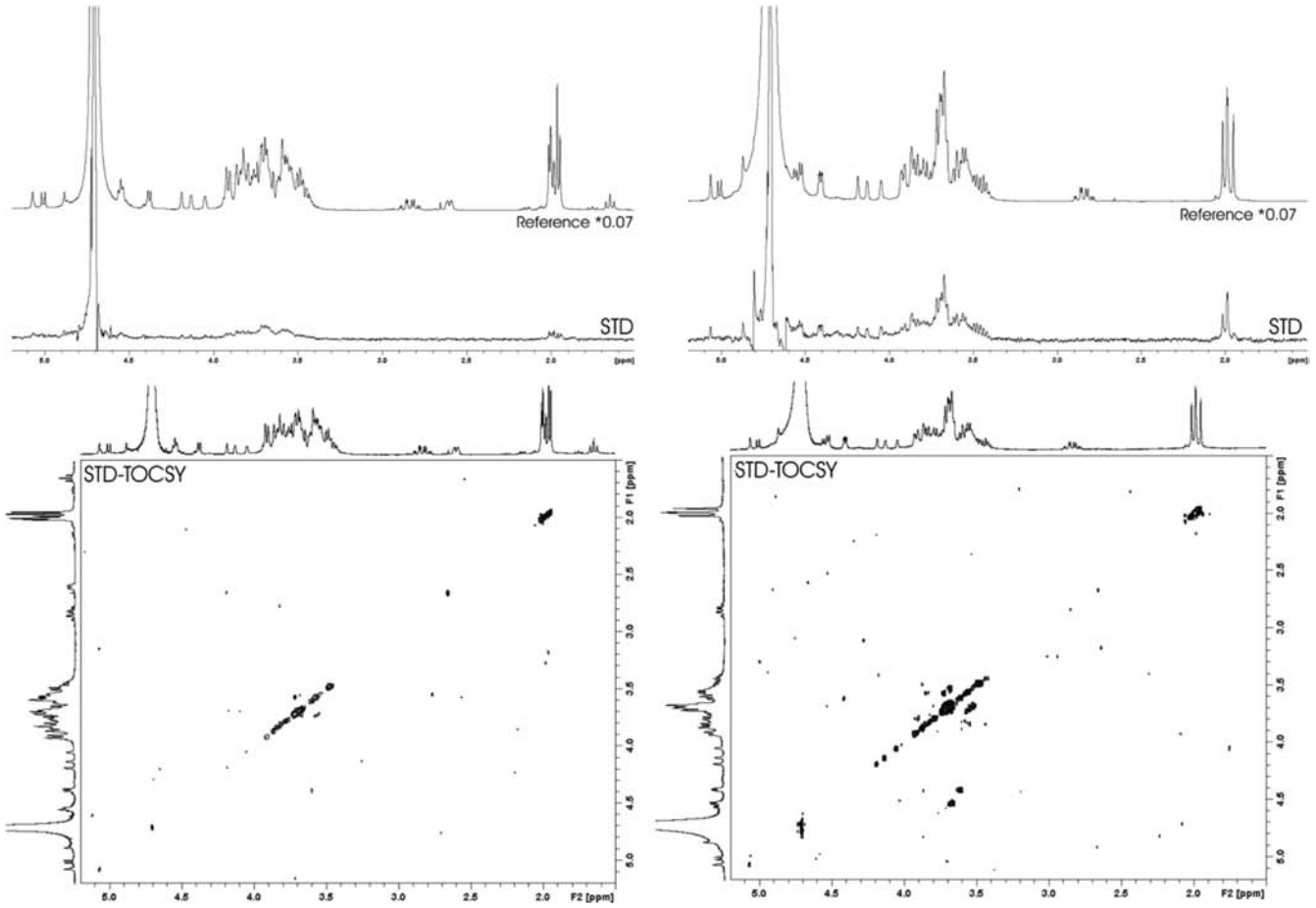


Figure 8. (left) STD and STD-TOCSY spectra corresponding to MAA−sialylated *N*-glycan (**4**). No STD is detected. (right) Spectra corresponding to MAA−non-sialylated *N*-glycan (**2**). The terminal *N*-acetyllactosamine residues provided clear STD signals indicating their direct involvement in the recognition process.

employed.[12] Both isolectins are known to be able to bind to α2,3-sialylgalactose, although with different preferences: the leukoagglutinin (MAL) is known to preferentially react with Neu5Acα2,3Galβ1,4βGlcNAc moieties present in *N*-glycans,[24] while the hemagglutinin (MAH) has higher affinity for the disialylated tetrasaccharide Neu5Acα2,3Galβ1,3[Neu5Ac-(2,6)]αGalNAc present in *O*-glycans.[4e,12,25] The binding sites of both isolectins are very similar, with the only differences being the amino acids Y221 and E222 of MAL, which are replaced by Ala in MAH. The X-ray crystal structure of MAL complexed with 3′-sialyllactose has been published.[26]

1D STD and two-dimensional STD-TOCSY experiments were acquired for **2** and **4**, using samples with a protein/ligand 1:20 (50 μM/1 mM) ratio. The corresponding spectra are shown in Figure 8. The comparison of the spectra on the left (sialylated *N*-glycan-Asn **4**) and on the right (non-sialylated **2**) enabled the assessment that the MAL−MAH mixture indeed does not recognize the Neu5Acα(2,6)Gal-containing molecule (**4**), providing a negative control for specificity, while it does efficiently bind to non-sialylated **2**. The documented lack of reactivity in lectin-affinity chromatography or in glycan arrays serves as further control.[4e,12] From the analysis of this STD spectrum, it was deduced that the residues interacting with the lectin are the nonreducing end terminal galactose and GlcNAc residues (LacNAc unit), while in this case, the GlcNAc1 and Asn stem at the other end is not in contact with the protein. This observation is in agreement with the previously reported specificity data and with the published crystallographic structure (pdb code 1DBN). Indeed, the binding site of MAL is unusually narrow for a lectin. Both the terminal galactose and GlcNAc residues are sandwiched between Y136 and Y221, leaving Gal O6 deep inside the binding site and establishing hydrogen bonding with D137, clearly precluding the extension through Gal O6, as it is the case of 2,6-sialylated *N*-glycan **4**. The absence of sialic acid allows the terminal LacNAc unit of *N*-glycan **2** to perfectly fit inside the binding pocket, following the typical feature for lectin binding[3b] by establishing a combination of CH−π stacking interactions with Y136 and Y221 and hydrogen bonds with Y131, D87, D137, and E224.[26] Further interactions are established with GlcNAc2 and ManB through Y136 and Y221 (Figure 9). Terminal elongation of **2** with Neu5Ac but with an α2−3 linkage was then explored by modeling procedures (Figure S6, Supporting Information), yielding a complex in which Neu5Ac not only fits in the narrow lectin binding site but further contributes to the binding by hydrogen bonding to amino acids K107, S104, and Y136, as also observed in the X-ray structure for sialylα2,3-lactose.[26] Thus, in this particular example, the presence of the α2,6-linked Neu5Ac moiety impairs the *N*-glycan ligand binding to the lectin. Interestingly, in the absence of Neu5Ac, recognition of the underlying LacNAc unit occurs.

dx.doi.org/10.1021/ja3104928 | *J. Am. Chem. Soc.* 2013, 135, 2667−2675

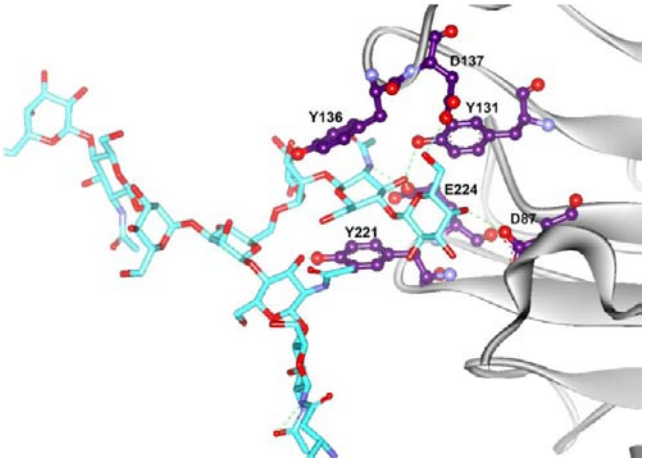


Figure 9. Modeled structure (using a docking and MD protocol) for the complex formed between MAL and **2**, using the deposited X-ray structure (pdb code 1DBN) as the template structure. The *N*-acetyllactosamine at the 1→6 branch has been docked at the binding site.

DISCUSSION

The preparation of *N*-glycans facilitates definition of the docking sites of lectins on complex glycans. Shown herein, the nature of the target epitope strictly depends on the nature of the tested lectin, which can select distinct determinants of the complex-type saccharide chain with exquisite specificity, depending on the relative presentation of the different residues and the architecture of the binding site.

We have underscored the importance of the coexistence of binding sites from different domains for achieving glycan binding by comparing a single hevein domain and WGA. The monomeric unit does not recognize the large glycan, with the branching at the βMan residue being the key structural element for precluding the interaction. In contrast, the multidomain lectin WGA binds the terminal reducing end of either **2** or **4**, namely, the GlcNAc−Asn epitope, owing to the possibility of achieving additional intersite stabilizing interactions from one of the neighboring hevein domains. In this context, it is worth emphasizing that WGA does not bind to **2** or **4** through the $(GlcNAc)_2$ moiety, as it would be expected. The chemical nature of the branched *N*-glycans excludes the possibility of efficient interactions of the chitobiose stem with the two consecutive aromatic residues typical for hevein domains, and thus, the recognized epitope is shifted by one position to the GlcNAc−Asn fragment. Despite being considered as a sialic acid binding lectin, our data have clearly shown that the sialyl residue of the *N*-glycan−Asn **4** is not a key player in the binding process. Thus, in naturally occurring *N*-glycopeptides, although presented at the spatially less readily accessible nonreducing end, binding occurs via the *N*-glycan−Asn stem region, a result of conspicuous importance for the interpretation of data using WGA in array platforms or in cyto- or histochemical glycophenotyping. Given the assumed significance of sialylation, for example, in tumor suppressor activities in response to environmental effectors or in immunity,[30] caution should thus be exercized regarding WGA as a tool to detect alterations in this parameter. As demonstrated, on the level of small glycans, WGA can accommodate sialic acid residues. Indeed, using α2,6-sialyllactose, we clearly demonstrate that this small trisaccharide is recognized by WGA, but the situation is different on the level of *N*-glycans. Considering the importance of terminal sialylation, we have extended our study by analyzing the respective recognition structurally.

Work with VAA has yielded insights into how α2,6-sialylation can be accommodated by this lectin, providing enhanced binding in natural *N*-glycans when compared to the non-sialylated LacNAc-terminated analogue. Our methodology has permitted identification of the direct involvement of the sialic acid, as well as a network of interactions.

Conversely, the results for the MAL−MAH mixture have permitted definition of the role of the α2,6-linked sialic acid residue in blocking binding of the corresponding *N*-glycan **4**. The binding to the LacNAc terminus of the non-sialylated glycan **2** is detectable, providing the control reference for positive binding.

CONCLUSION

In conclusion, the wide variety found both within the glycome and lectin architectures is a challenge for analyzing carbohydrate−lectin interactions. Advances in the macroscopic analysis of oligosaccharide binding using array-based and other technologies[31] can strategically be complemented by the given approach, to rigorously define the contact sites also of complex glycans. Thus, the combination of *N*-glycan preparation, STD NMR spectroscopy, and modeling protocols gives new insights and proves its potency to detect and characterize epitope selection and interaction. They are of particular relevance for lectins used for biomedical applications.

ASSOCIATED CONTENT

Supporting Information

^{1}H NMR titration of a solution of B domain of WGA with increasing amounts of sialyated *N*-glycan; edited HQSC spectrum of **4**; STD spectra of **3** and WGA, **4**−WGA, and **5** with WGA; TOCSY of **4** and STD-TOCSY of **4**−VAA; and model for the recognition of a undecasaccharide with terminal Neu5Ac in α2−3 linkage to Gal by MAL. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

jjbarbero@cib.csic.es

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support from the Ministry of Economy and Competitiveness of Spain (Grants CTQ2011-22514, CTQ2009-08536, and CTQ2012-30820), from Comunidad de madrid (MHit project), and from the EU (Projects GlycoHit and GlycoPharm, and COST action CM1102) is acknowledged.

REFERENCES

(1) (a) Haltiwanger, R. S.; Lowe, R. B. *Annu. Rev. Biochem.* **2004**, *73*, 491−537. (b) *The Sugar Code: Fundamentals of Glycosciences*; Gabius, H.-J., Ed.; Wiley-VCH: Weinheim, Germany, 2009.

(2) Cummings, R. D. *Mol. BioSyst.* **2009**, *5*, 1087−1104.

(3) (a) Kiessling, L. L.; Splain, R. A. *Annu. Rev. Biochem.* **2010**, *79*, 619−653. (b) Gabius, H.-J.; André, S.; Jiménez-Barbero, J.; Romero, A.; Solís, D. *Trends Biochem. Sci.* **2011**, *36*, 298−313.

(4) (a) Arnold, J. N.; Saldova, R.; Hamid, U. M.; Rudd, P. M. *Proteomics* **2008**, *8*, 3284−3293. (b) Drake, P. M.; Cho, W.; Li, B.; Prakobphol, A.; Johansen, E.; Anderson, N. L.; Regnier, F. E.; Gibson,

B. W.; Fisher, S. J. *Clin. Chem.* **2010**, *56*, 223−236. (c) Reis, C. A.; Osorio, H.; Silva, L.; Gomes, C.; David, L. *J. Clin. Pathol.* **2010**, *63*, 322−329. (d) Gabius, H.-J. *Biochem. Soc. Trans.* **2011**, *39*, 557−565. (e) Amano, M.; Eridsson, H.; Manning, J. C.; Detjen, K. M.; André, S.; Mishimura, S.-I.; Lehtiö, J.; Gabius, H.-J. *FEBS J.* **2012**, *279*, 4062−4080.

(5) (a) Etxebarria, J.; Calvo, J.; Martin-Lomas, M.; Reichardt, N. C. *ACS Chem. Biol.* **2012**, *7*, 1729−1737. (b) Song, X.; Yu, H.; Chen, X.; Lasanajak, Y.; Tappert, M. M.; Air, G. M.; Tiwari, V. K.; Cao, H.; Chokhawala, H. A.; Zheng, H.; Cummings, R. D.; Smith, D. F. *J. Biol. Chem.* **2011**, *286*, 31610−31622. (c) Song, X. Z.; Xia, B. Y.; Stowell, S. R.; Lasanajak, Y.; Smith, D. F.; Cummings, R. D. *Chem. Biol.* **2009**, *16*, 36−47. (d) Blixt, O.; Head, S.; Mondala, T.; Scanlan, C.; Huflejt, M. E.; Alvarez, R.; Bryan, M. C.; Fazio, F.; Calarese, D.; Stevens, J.; Razi, N.; Stevens, D. J.; Skehel, J. J.; van Die, I.; Burton, D. R.; Wilson, I. A.; Cummings, R.; Bovin, N.; Wong, C.-H.; Paulson, J. C. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17033−17038. (e) Padler-Karavani, V.; et al. *J. Biol. Chem.* **2012**, *287*, 22593−22608.

(6) (a) André, S.; Kožár, T.; Schuberth, R.; Unverzagt, C.; Kojima, S.; Gabius, H.-J. *Biochemistry* **2007**, *46*, 6984−6995. (b) André, S.; Kožár, T.; Kojima, S.; Unverzagt, C.; Gabius, H.-J. *Biol. Chem.* **2009**, *390*, 557−565.

(7) Roldós, V.; Cañada, F. J.; Jiménez-Barbero, J. *ChemBioChem* **2011**, *12*, 990−1005.

(8) (a) Zuber, C.; Roth, J. *N*-Glycosylation. In *The Sugar Code: Fundamentals of Glycosciences*; Gabius, H.-J., Ed.; Wiley-VCH: Weinheim, Germany, 2009; pp 87−110. (b) Schauer, R. *Curr. Opin. Struct. Biol.* **2009**, *19*, 507−514.

(9) Goldstein, I. J.; Poretz, R. D. In *The Lectins: Properties, Functions and Applications in Biology and Medicine*; Liener, I. E., Sharon, N., Goldstein, I. J., Eds.; Academic Press: San Diego, CA, 1986; pp 33−247.

(10) (a) Jiménez-Barbero, J.; Cañada, F. J.; Asensio, J. L.; Aboitiz, N.; Vidal, P.; Canales, A.; Groves, P.; Gabius, H.-J.; Siebert, H.-C.; Derek, H. *Adv. Carbohydr. Chem. Biochem.* **2006**, *60*, 303−354.

(11) (a) Asensio, J. L.; Cañada, F. J.; Bruix, M.; Rodriguez-Romero, A.; Jiménez-Barbero, J. *Eur. J. Biochem.* **1995**, *230*, 621−633. (b) Asensio, J. L.; Cañada, F. J.; Siebert, H.-C.; Laynez, J.; Poveda, A.; Nieto, P. M.; Soedjanaatmadja, U. M.; Beintema, J. J.; Gabius, H.-J.; Jiménez-Barbero, J. *Chem. Biol.* **2000**, *7*, 529−543.

(12) (a) Kawaguchi, T.; Matsumoto, I.; Osawa, T. *J. Biol. Chem.* **1974**, *249*, 2786−2792. (b) Knibbs, R. N.; Goldstein, I. J.; Ratcliffe, R. M.; Shibuya, N. *J. Biol. Chem.* **1991**, *266*, 83−88. (c) Geisler, C.; Jarvis, D. L. *Glycobiology* **2011**, *21*, 988−993.

(13) (a) André, S.; Unverzagt, C.; Kojima, S.; Dong, X.; Fink, C.; Kayser, K.; Gabius, H.-J. *Bioconjugate Chem.* **1997**, *8*, 845−855. (b) Galanina, O. E.; Kaltner, H.; Khraltsova, L. S.; Bovin, N. V.; Gabius, H.-J. *J. Mol. Recognit.* **1997**, *10*, 139−147. (c) Unverzagt, C.; André, S.; Seifert, J.; Kojima, S.; Fink, C.; Srikrishna, G.; Freeze, H.; Kayser, K.; Gabius, H.-J. *J. Med. Chem.* **2002**, *45*, 478−491.

(14) (a) Mayer, M.; Meyer, B. *Angew. Chem., Int. Ed.* **1999**, *38*, 1784−1788. (b) Jiménez-Barbero, J.; Dragoni, E.; Venturi, C.; Nannucci, F.; Ardá, A.; Fontanella, M.; André, S.; Cañada, F. J.; Gabius, H.-J.; Nativi, C. *Chem.—Eur. J.* **2009**, *15*, 10423−10431. (c) Martín-Santamaría, S.; André, S.; Buzamet, E.; Caraballo, R.; Fernández-Cureses, G.; Morando, M.; Ribeiro, J. P.; Ramírez-Gualito, K.; de Pascual-Teresa, B.; Cañada, F. J.; Menéndez, M.; Ramström, O.; Jiménez-Barbero, J.; Solís, D.; Gabius, H.-J. *Org. Biomol. Chem.* **2011**, *9*, 5445−5455.

(15) Kaminski, G. A.; Friesner, R. A.; Tirado-Rives, J.; Jorgensen, W. J. *J. Phys. Chem. B* **2001**, *105*, 6474.

(16) *Maestro, A Powerful, All-Purpose Molecular Modeling Environment*, version 8.5; Schroedinger, LLC: New York, 2008.

(17) Gabius, H.-J.; Darro, F.; Remmelink, M.; André, S.; Kopitz, J.; Danguy, A.; Gabius, S.; Salmon, I.; Kiss, R. *Cancer Invest.* **2001**, *19*, 114−126.

(18) Espinosa, J. F.; Asensio, J. L.; García, J. L.; Laynez, J.; Bruix, M.; Wright, C.; Siebert, H.-C.; Gabius, H.-J.; Cañada, F. J.; Jiménez-Barbero, J. *Eur. J. Biochem.* **2000**, *267*, 3965−3978.

(19) Hernández-Gay, J. J.; Ardá, A.; Eller, S.; Mezzato, S.; Leeflang, B. R.; Unverzagt, C.; Cañada, F. J.; Jiménez-Barbero, J. *Chem.—Eur. J.* **2010**, *16*, 10715−10726.

(20) Kajihara, Y.; Suzuki, Y.; Yamamoto, N.; Sasaki, K.; Sakakibara, T.; Juneja, L. R. *Chem.—Eur. J.* **2004**, *10*, 971−985.

(21) Seko, A.; Koketsu, M.; Nishizono, M.; Enoki, Y.; Ibrahim, H. R.; Juneja, L. R.; Kim, M.; Yamamoto, T. *Biochim. Biophys. Acta, Gen. Subj.* **1997**, *1335*, 23−32.

(22) (a) Wright, C. S. *J. Mol. Biol.* **1984**, *178*, 91−104. (b) Wright, C. S. *J. Mol. Biol.* **1990**, *215*, 635−651. (c) Wright, C. S. *J. Biol. Chem.* **1992**, *267*, 14345−14352. (d) Wright, C. S.; Jaeger, J. *J. Mol. Biol.* **1993**, *232*, 620−638.

(23) (a) Kronis, K. A.; Carver, J. P. *Biochemistry* **1982**, *21*, 3050−3057. (b) Kronis, K. A.; Carver, J. P. *Biochemistry* **1985**, *24*, 826−833.

(24) Wang, W. C.; Cummings, R. D. *J. Biol. Chem.* **1988**, *263*, 4576−4585.

(25) Konami, Y.; Yamamoto, K.; Osawa, T.; Irimura, T. *FEBS Lett.* **1994**, *342*, 334−338.

(26) Imberty, A.; Gautier, C.; Lescar, J.; Perez, S.; Wyns, L.; Loris, R. *J. Biol. Chem.* **2000**, *275*, 17541−17548.

(27) Jiménez, M.; André, S.; Siebert, H.-C.; Gabius, H.-J.; Solís, D. *Glycobiology* **2006**, *16*, 926−937.

(28) (a) Mikeska, R.; Wacker, R.; Arni, R.; Singh, T. P.; Mikhailov, A.; Gabdoulkhakov, A.; Voelter, W.; Betzel, C. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2005**, *61*, 17−25. (b) Jiménez, M.; André, S.; Barillari, C.; Romero, A.; Rognan, D.; Gabius, H.-J.; Solís, D. *FEBS Lett.* **2008**, *582*, 2309−2312.

(29) Müthing, J.; Meisen, I.; Bulau, P.; Langer, M.; Witthohn, K.; Lentzen, H.; Neumann, U.; Peter-Katalinic, J. *Biochemistry* **2004**, *43*, 2996−3007.

(30) (a) André, S.; et al. *FEBS J.* **2007**, *274*, 3233−3256. (b) Patsos, G.; André, S.; Roeckel, N.; Gromes, R.; Gebert, J.; Kopitz, J.; Gabius, H.-J . *Glycobiology* **2009**, *19*, 726−734. (c) van de Wouwer, M.; André, S.; Gabius, H.-J.; Villalobo, A. *Nitric Oxide* **2011**, *24*, 91−101. (d) Varki, A.; Gagneux, P. *Ann. N. Y. Acad. Sci.* **2012**, *1253*, 16−36.

(31) (a) Solís, D.; Romero, A.; Menéndez, M.; Jiménez-Barbero, J. Protein−carbohydrate interactions: basic concepts and methods for analysis. In *The Sugar Code: Fundamentals of Glycosciences*; Gabius, H.-J., Ed.; Wiley-VCH: Weinheim, Germany, 2009; pp 233−245. (b) Tateno, H.; Kuno, A.; Hirabayashi, J. How to determine specificity: from lectin profiling to glycan mapping and arrays. In *The Sugar Code: Fundamentals of glycosciences*; Gabius, H.-J., Ed.; Wiley-VCH: Weinheim, Germany, 2009; pp 247−259.