Analysis of Protein−Carbohydrate Interaction at the Lower Size Limit of the Protein Part (15-Mer Peptide) by NMR Spectroscopy, Electrospray Ionization Mass Spectrometry, and Molecu...

ARTICLE *in* BIOCHEMISTRY · AUGUST 2002

Impact Factor: 3.02 · DOI: 10.1021/bi025891x · Source: PubMed

Analysis of Protein−Carbohydrate Interaction at the Lower Size Limit of the Protein Part (15-Mer Peptide) by NMR Spectroscopy, Electrospray Ionization Mass Spectrometry, and Molecular Modeling†

Hans-Christian Siebert,*,‡ Shan-Yun Lü,§ Martin Frank,|| Jasper Kramer,⊥ Rainer Wechselberger,# John Joosten,▽ Sabine André,‡ Kate Rittenhouse-Olson,○ René Roy,× Claus-Wilhelm von der Lieth,|| Robert Kaptein,# Johannes F. G. Vliegenthart,▽ Albert J. R. Heck,⊥ and Hans-Joachim Gabius*,‡

Institut für Physiologische Chemie, Tierärztliche Fakultät, Ludwig-Maximilians-Universität München, Veterinärstrasse 13, 80539 München, Germany, College of Life Sciences, Peking University, Beijing 100871, P.R. China, Zentrale Spektroskopie, Deutsches Krebsforschungszentrum, Im Neuenheimer Feld 280, 69120 Heidelberg, Germany, Department of Biomolecular Mass Spectrometry, Bijvoet Center for Biomolecular Research and Utrecht Institute for Pharmaceutical Sciences, Utrecht University, Sorbonnelaan 16, 3584 CA Utrecht, The Netherlands, Department of NMR Spectroscopy, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands, Department of Bio-Organic Chemistry, Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands, Department of Biotechnical and Clinical Laboratory Sciences, State University of New York, Buffalo, New York 14214, and Centre for Research in Biopharmaceuticals, Department of Chemistry, University of Ottawa, Ottawa, Ontario, Canada K1N 6N5

Received March 28, 2002; Revised Manuscript Received June 1, 2002

ABSTRACT: Structural analysis of minimally sized lectins will offer insights into fundamentals of intermolecular recognition and potential for biomedical applications. We thus moved significantly beyond the natural limit of lectin size to determine the structure of synthetic mini-lectins in solution, their carbohydrate selectivity and the impact of ligand binding on their conformational behavior. Using three disaccharide (Thomsen-Friedenreich antigen; Galβ1,3GalNAcα1,R)-binding pentadecapeptides without internal disulfide bridges as role models, we successfully tested a combined strategy with different techniques of NMR spectroscopy, electrospray ionization mass spectrometry, and molecular modeling. In solution, the peptides invariably displayed flexibility with rather limited restrictions, shown by NMR experiments including nearly complete resonance assignments and molecular dynamics simulations. The occurrence of aromatic/nonpolar amino acids in the sequence did not lead to formation of a hydrophobic core known from microbial chitinase modules. Selectivity of disaccharide binding was independently observed by mass spectrometry and NMR analysis. Specific ligand interaction yielded characteristic NMR signal alterations but failed to reduce conformational flexibility significantly. We have thereby proven effectiveness of our approach to analyze even low-affinity interactions (not restricted to carbohydrates as ligands). It will be useful to evaluate the impact of rational manipulation of lead peptide sequences.

Nucleotides, amino acids, and carbohydrates share vital prerequisites to serve as hardware in bioinformation storage and transfer. Like letters of an alphabet they can form oligo- and polymers making up code words with building blocks which are sufficiently different to minimize risk potential for chemical ambiguities. Remarkably, the potential which carbohydrates offer for isomer synthesis is unsurpassed among the three classes of biomolecules (*1*, *2*). In fact, this high-density coding capacity will keep the size of code units small, for example, in sulfated/sialylated Lewis epitopes as determinant in lymphocyte homing (*3*). In a broader context, carbohydrates are increasingly being unraveled to play a role in biorecognition and initiation of signaling pathways with medical perspectives (functional glycomics) (*4*−*7*). This versatility underscores the relevance to study protein(lectin)−carbohydrate interactions in detail. The aim is to account for enthalpic and entropic terms by the individual contributions of directional hydrogen and coordination bonds and stacking/hydrophobic and ionic interactions, as well as by changes of conformational equilibrium and flexibility, and to explain specificity of lectins in molecular terms (*8*−*16*). An intriguing question in this area concerns elucidation of the minimal structural requirements for ligand binding by polypeptides.

Regarding animal lectins (*13*, *17*), carbohydrate recognition domains of major families such as galectins or C-type and I-type lectins generally consist of about 115−130 amino

† The work benefited from financial support from the EC program Training and Mobility in Research (ERBFMGECT-950032) and the EC program for high-level scientific conferences and the Wilhelm-Sander-Stiftung (Munich, Germany).

* To whom correspondence should be addressed. Phone: +49-89-21802290. Fax: +49-89-21802508. E-mail: hcsiebert@aol.com or gabius@lectins.de or gabius@tiph.vetmed.uni-muenchen.de.

‡ Institut für Physiologische Chemie.

§ College of Life Sciences.

|| Zentrale Spektroskopie.

⊥ Department of Biomolecular Mass Spectrometry.

Department of NMR Spectroscopy.

▽ Department of Bio-Organic Chemistry.

○ Department of Biotechnical and Clinical Laboratory Sciences.

× Centre for Research in Biopharmaceuticals.

10.1021/bi025891x CCC: $22.00
Published on Web 07/04/2002

acids, with further truncation abolishing their activity as demonstrated for two C-type lectins and galectin-3 (*18*, *19*). With 111 amino acids and four disulfide bridges, the sialic-acid-binding lectin with intrinsic RNase activity from bullfrog eggs is another example for a functional lectin of this size class (*20*, *21*). Beyond this limit, lectin occurrence is not frequent. It encompasses the chitin-binding proteins hevein and pseudohevein from the rubber tree (*Hevea brasiliensis*) with 43 amino acids and the two small lectins from the venom of the spider *Selenocosmia huwena* (SHL-I) and the seeds of love-lies-bleeding (*Amaranthus caudatus*; referred to as Ac-AMP2) with 30 and 32 amino acids, respectively. As role models for the group of hevein-domain-containing proteins, the solution structures of the two lectins from the rubber tree were assessed by NMR spectroscopy. The major secondary structural elements are an antiparallel β-sheet formed by two tetrapeptide stretches within the sequence Leu16-Ser26 and a short α-helix from Asp28 to Ser32 together with four disulfide bridges (*22*−*26*). For hevein-ligand interaction, Trp21 and Tyr30 provide favorable stacking and a hydrogen bond (Tyr30 with OH-3 of the nonreducing GlcNAc residue) so that the ligand comes into contact with the aromatic residues (*23*, *24*, *27*). The antimicrobial protein Ac-AMP2 with three disulfide bridges shares the antiparallel β-sheet from Met13−Lys23 to an extent that the root-mean-square deviations of backbone atoms (residues 12−32) to those of hevein amount to only 0.1 nm (*24*, *28*). This structural motif is likewise seen in the mentioned spider lectin (*29*) and the C-terminal part of tachycitin, an antimicrobial protein from the horseshoe crab (*30*). The concept that has emerged attributes origin of this class of plant and invertebrate lectins to convergent evolution (*31*). Although the stability of this folding pattern is not sufficient to be maintained in an aprotic polar solvent in contrast to Ig-domains and receptor domains of galectins (*32*), the combination of an antiparallel β-strand with aromatic amino acids for ligand contact appears to be favorable to build up small lectins. The structural organization of a bacterial chitin-binding domain reinforces this conclusion. Even in the absence of disulfide bonds, a twisted β-sandwich and the core region with aromatic and further hydrophobic side chains aid to make the structure of this domain (45 residues) of the Chitinase A1 of *Bacillus circulans* rigid (*33*). Thus, the natural size limit for a lectin domain has been charted well with these descriptions. To define the minimal structural requirements for carbohydrate binding, it is essential to examine the properties of peptides.

Two approaches have been exploited to gain access to sequences of sugar-binding peptides. Selected sequence stretches from the lectin domain of selectins or the S2 subunit of the pertussis toxin were shown to react with sialic-acid-containing ligands (*34*−*38*). Presence of either Ca^{2+}-ions (selectin-derived peptides) or sialic acid in the ligand was indispensable, carbohydrate binding being abolished in randomized sequences. Dependence of the presence of Ca^{2+}(Mn^{2+})-ions was also observed with active peptides from five plant lectins obtained by affinity chromatography after endoproteinase digestion (*39*, *40*). Thus, either a negative charge or metal ions are crucial in these cases to effect binding. An ionic interaction is also likely to contribute markedly to binding of ganglioside GM_1 to three 15-mer peptides derived from screening of a phage-displayed

P6 : ARVSFWRYSSFAPTY

P10: GSWYAWSPLVPSAQI

P30: HGRFILPWWYAFSPS

FIGURE 1: Sequences of the three TF-binding pentadecapeptides p6, p10 and p30 according to published data (*43*).

pentadecapeptide library (*41*, *42*). Methodologically, such libraries enable the discovery of active sequences without being restricted to lectin fragments. By using the same technique, three apparently high-affinity pentadecapeptides (K_D values of 0.1−1.2 μM in interaction with asialofetuin and of 20−200 μM with disaccharide; for sequences, see Figure 1) to the Thomsen−Friedenreich antigen (TF;[1] Galβ1,3GalNAcα1,R) had recently been described (*43*, *44*). Their synthesis afforded the opportunity to initiate the biophysical study of peptide-carbohydrate interaction without the monitored accessory factors (metal ions, ionic interaction). To gain first insights into the solution structures of carbohydrate-binding peptides and their modes of interaction with the ligand, we devised a combined strategy with mass spectrometry (MS), nuclear magnetic resonance (NMR) spectroscopy, and molecular modeling. This integrated approach is rigorously tested and herein shown to hold promise for further work in this area.

In our study, we first performed analyses of the NMR spectra of the free peptides in solution and resonance assignment. We next addressed the questions whether the frequent occurrence of aromatic residues (3−5 side chains) per pentadecapeptides favors formation of a hydrophobic core or aggregation by NMR spectroscopy, electrospray ionization MS (ESI MS), and molecular dynamics (MD) simulations and whether the peptides (as monomers and/or oligomers?) display specificity in a direct binding assay. Monitoring of the assays is performed independently by NMR spectroscopy and ESI MS. Having ascertained the peptide−sugar interaction by ESI MS and by recording shifts and line broadening of peptide-derived NMR signals, we then searched for structural details of the small peptides responsible for binding the carbohydrate ligand using the NMR-derived information. On the basis of current theoretical models and the experimental input, we also performed molecular modeling for this purpose. In the course of these analyses, we answered the question whether changes in the degree of flexibility of the peptides occurred, which might be brought about by ligand-induced formation of either a hydrophobic core or a secondary structural element. In fact, peptide−aptamer association is characterized by major conformational shifts within the rugged energy landscape and acquisition of secondary structure (*45*). In a more general context, it was asserted that an intrinsic lack of folded structure covering a

[1] Abbreviations: AMBER, assisted model building with energy refinement; CIDNP, chemically induced dynamic nuclear polarization; ESI MS, electrospray ionization mass spectrometry; IgG, immunoglobulin G; MD, molecular dynamics; MLEV-17, composite pulse decoupling sequence; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser and exchange spectroscopy; PNA, peanut agglutinin; Q-TOF, quadrupole time-of-flight; STD, saturation transfer difference; TF, Thomsen−Friedenreich antigen; TLC, thin-layer chromatography; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase incrementation.

sequence stretch of amino acids in the absence of the target structure might provide an intriguing level of control in regulatory proteins (*46*). Drawing on the crystal structure of the TF-specific peanut agglutinin (PNA) in complex with lactose or the TF-antigen (*47*, *48*), we also present results of a comparative computational interaction analysis of the disaccharide with the contiguous sequence stretch of a pentadecapeptide and the lectin's binding site established by four loops. On the basis of our data, we finally discuss strategies how to improve the reactivity of these carbohydrate-binding peptides.

MATERIALS AND METHODS

Materials. Peptides p6, p10 and p30 were chemically synthesized on a solid-phase peptide synthesizer and purified to homogeneity on a HPLC column. Chemical synthesis of the TF-antigen derivative (Galβ1,3GalNAcα1,R, R = $-O-CH_2-CH{=}CH_2$) involved covalently linking the 2-acetamido-2-deoxy-4,6-*O*-(p-methoxybenzylidene)-β-D-galactopyranoside derivative to 2,3,4,6-tetra-*O*-acetyl-α-D-galactosyl bromide in anhydrous nitromethane-benzene (1:1, v/v) in the presence of sodium sulfate and mercuric cyanide, as described in detail elsewhere (*49*). Purity and quality controls of the products in each synthetic step were performed by thin-layer chromatography on silica plates (TLC precoated plastic sheets SIL G/UV, layer = 0.25 mm, Macherey-Nagel, Düren, Germany) and NMR spectroscopy, using a published resonance assignment (*50*). The TF-specific monoclonal antibody (IgG_3) had been raised by immunization with a synthetic T-antigen-bearing neoglycoprotein (*51*−*53*). Supernatants of JAA-F11 hybridoma cells grown in RPMI 1640 medium containing 10% fetal calf serum and supplements (L-glutamine, vitamins, sodium pyruvate and nonessential amino acids) were centrifuged at 20000*g* for 30 min. The resulting pellet was discarded, and the immunoglobulin G (IgG) fraction was purified by ammonium sulfate precipitation at 45% (v/v) saturation and protein A-Sepharose CL-4B chromatography. Following dialysis against water and lyophilization, the antibody fraction was stored at −20 °C and brought into solution before running the experiments by adding D_2O.

Mass Spectrometric Experiments. Aqueous solutions of mixtures of p6, p10, and p30 (20 μM) with the TF-antigen and cellobiose or maltose (100 μM) were used for electrospray (ES) ionization studies. Mass spectrometry measurements were performed on a quadrupole time-of-flight (Q-TOF) instrument (Micromass, Manchester, U.K.) operating in positive ion mode, equipped with a "Z-spray" nanoelectrospray source using in-house pulled and gold-coated needles. Typical conditions were needle voltage, 1300−1600 V; cone voltage, 25 V; quadrupole pressure, 5×10^{-6} mbar; and TOF analyzer pressure, 3×10^{-7} mbar. Mass spectra were averaged over typically 100 scans. The standard mass scanned was 100−4000 Thomson. To deduce binding constants from the relative ion abundance in the ESI MS spectra, the assumption was made that the response factors (e.g., ionizability) for the doubly protonated peptide−peptide or peptide−carbohydrate complexes are similar to those of free peptides and carbohydrates (*54*, *55*). However, this assumption cannot be regarded to be invariably and fully valid, and thus the determined binding constants should be interpreted as estimations.

NMR Spectroscopical Experiments. ^{1}H NMR spectra were recorded in 90% H_2O/10% D_2O with Bruker AMX 500, AMX 600, and Varian Unity 750 MHz spectrometers. Spectra from the two-dimensional experiments were acquired at 33 °C (and 43 °C) with 4 mM solutions of p6, p10, and p30 in the absence or in the presence of an equimolar concentration of the TF-antigen. The total correlation spectroscopy (TOCSY) and nuclear Overhauser and exchange spectroscopy (NOESY) experiments were performed in the phase-sensitive mode using time-proportional phase incrementation (TPPI) method for quadrature detection in F1. Typically, a data matrix of 512×2048 points was chosen to digitize a spectral width of 15 ppm. Eighty scans per increment were used with a relaxation delay of 1 s. Prior to Fourier transformation, zero filling was performed in F1 direction to expand the data to 1024×2048 points. Baseline correction was applied in both dimensions. The TOCSY spectra were recorded applying 10 and 50 ms, respectively, by use of a MLEV-17 isotropic mixing scheme. The NOESY experiments were performed with mixing times of 50, 100, and 200 ms, as described previously in a study of protein−carbohydrate interaction (*56*). Experiments were carried out for the peptides in the absence of ligand and in the presence of the TF-antigen at a molar ratio of peptide:disaccharide set to 1:1. The applied laser light for the laser photo chemically induced dynamic nuclear polarization (CIDNP) experiments was generated by a continuous-wave argon ion laser (Spectra Physics) which operates in the multi-line mode with principal emission wavelengths of 488.0 and 514.5 nm, close to the edge of the 450 nm absorption band of the dye. The laser light was directed to the sample by an optical fiber and chopped by use of a mechanical shutter controlled by the spectrometer. This setup of the equipment prevents harmful heating of the peptide-containing solution. The laser photo CIDNP radical reaction is initiated by the flavin I mononucleotide (N3 of the isoalloxazine ring substituted with CH_2COOH, N10 with CH_3) as laser-reactive dye. The irradiation leads to the generation of peptide-dye radical pairs involving dye-accessible (and therefore surface-exposed) Tyr, Trp, and His residues in a typical experimental setting, as described previously for protein−carbohydrate interactions (*27*, *57*, *58*). The saturation transfer difference (STD) (*59*, *60*) spectrum was obtained by collecting 512 scans for on- ($\Delta = 7.2$ ppm) and off-resonance ($\Delta = 40$ ppm). A number of 16 dummy scans was chosen. The internal subtraction from each other was carried out by phase cycling. Selective pulses of 50 ms duration were used to saturate peptide resonances. Measurements were run on solutions containing 0.5 mM of the peptide (p10) and 5 mM of the TF-antigen. The temperature was kept at 32 °C during the experiment.

Molecular Modeling. Molecular dynamics simulations, docking analyses, and the processing of the acquired data were performed using SGI and IBM workstations and PCs or McIntosh computers as platforms. The conformational behavior of peptides was calculated using the program DISCOVER 2.98 (Accelrys Inc., San Diego, CA), atomic charge assignment of INSIGHT II and the parametrization of the AMBER 1.6 (assisted model building with energy refinement) force field (*61*, *62*). Following a standard equilibration period of 100 ps, molecular dynamics (MD) runs using a distance-dependent dielectric constant $\epsilon = 4r$ were performed at 300 or 400 K including the presence of

explicit water molecules for total sampling phases of 5 or 10 ns in the cases of the peptides and 100 ps for the two lectins hevein and PNA. To prevent random movement of the peptides close to the boundaries of the solvent box with side lengths of 40 Å during the simulation periods, a selected residue (Cα of Leu6 for p30 and Arg2 for p6) was fixed in its original position. An integration step of 1 fs and storage of a conformation after a series of 1000 steps of integration led to 5000 or 10000 individual conformations reflecting the extent of peptide flexibility during the simulation period. To visualize ligand contacts with the peptides, an experimentally derived distance constraint was introduced based on the NMR data: for p30 between Trp8Cβ and the TF-antigen (C1 of the nonreducing sugar unit) and for p6 between Arg2 (Cδ) and the TF-antigen (oxygen atom of the glycosidic linkage). The distance was not allowed to exceed 10 Å. To offer the possibility to view the conformational behavior of the peptides p6 and p30 in the absence and in the presence of the TF-antigen, we have made relevant parts of MD runs publicly accessible under http://www.dkfz.de/spec/publications/ using the Chime program available free-of-charge (http://www.mdlchime.com/chime/users/download.asp, last accessed XXXXX). In addition to giving access to visual inspection of the conformational behavior on the screen we also processed the complete data set to yield the average values (±standard deviations) of interproton distances in order to delineate any ligand-dependent changes of the positioning of amino acid residues quantitatively. As a further indicator of residue flexibility we systematically calculated relative probabilities for residue positioning in volume elements of 1 $Å^3$ (voxels) as a function of the individual residue and its position in the sequence based on the data of the MD runs. In this set of calculations, each structural unit (amino acid or hexopyranose) was transformed into a pseudoatom (geometric mean of heavy atom coordinates). Using independent calibration of each pseudoatom (the most densely populated volume segment was scaled to 100), characteristic isodensity plots for the accessible voxel distribution with an arbitrary scaling set to the half-maximal value were obtained as measure of internal flexibility. Further analysis (distance matrix approach) of the probability of pseudoatom residence in relation to any other sequence part (boundary set to distance <10 Å) derived a measure of spatial correlation of the complete set of building blocks. As positive control for these calculations, the lectin-ligand complexes of hevein and PNA were also subjected to these calculations using the parameters of the NMR solution structure of hevein (*22*−*24*, *27*) and the crystallographic parameters of PNA (*47*, *48*, *63*) to start the MD runs at 300 K in a water box. Any occurrence of spatial vicinity indicative of presence of secondary/tertiary structural elements or contact to the ligand will show up above or beyond the diagonal line. The extent of pseudoatom distances is visualized quantitatively by different degrees of shading. Added to each respective figure part on its right side, the internal calibration of shading relative to distance is shown. To obtain energy values (Coulomb and *van der Waals* terms) for the interaction between two receptor types (peptide p30, PNA) and the TF-antigen for relative comparison, the print molecule−molecule interaction option of Discover was used, as described previously in detail for another example of interaction analysis involving a sugar ligand (*64*).

RESULTS AND DISCUSSION

Probing the size limits of peptide length for operative carbohydrate interaction is encouraged by the natural role models mentioned in the introductory portion of this paper. Especially the rigid and compact core region of the chitin-binding domain of *Bacillus circulans* WL-12 Chitinase A1 (*33*) gives reason to expect a limitation of conformational fluctuations in peptide modules by presence of hydrophobic clusters. Also of interest, a biologically active nonapetide, i.e., the N-terminal part of the HIV-1 Tat protein binding to dipeptidyl peptidase IV, revealed a tendency to adopt a type I β-turn and a left-handed polyproline type II helix (*65*). This example, among others providing evidence for secondary structure propensity overriding unrestrained flexibility (*66*), attests that small peptides−free or in complex with binding partners−are not necessarily devoid of conformational constraints. Because size reductions to 8−15mer peptides will account for acquisition of favorable pharmacokinetic properties in biomedical applications (such as clearance, improved target-to-background ratio and fairly unimpeded accessibility to target organs/sites), such small synthetic products could become potent tools in carbohydrate-mediated drug delivery, imaging, or anti-adhesion/proliferation approaches (*67*, *68*). Moreover, elucidation of the way how a sequence stretch of a peptide interacts with the ligand will definitely enhance our understanding of carbohydrate selectivity. This information can also enable a detailed comparison to the more complex positioning of crucial contact sites from different loops in plant and animal lectins (*13*, *67*, *69*, *70*). As emerges from Figure 1, the three TF-antigen-binding sequences are established by sufficiently different sequences to prompt individual structural analysis. Our experiments started with sequence-specific resonance assignment and searches for secondary structure elements.

Secondary Structure and Resonance Assignment. The NMR-spectroscopical analysis of the three peptides revealed no evidence for long-range connectivities and secondary structure. As exemplarily shown in Figure 2, the amide-CαH region of the 2D-NOESY spectrum of p6 is devoid of such signals. In this spectrum, the cross-peaks involving the N−H group of Arg2 at 8.65 ppm and the sequential NOE to Val3 are highlighted (Figure 2). The cross-peak, assigned to Arg2, and not Arg7 due to the NOE contact, will be important when discussing spectra of this peptide in the presence of the ligand. To be able to attribute any ligand-induced changes to individual side chains, it was necessary to correlate the sequences with the signal pattern. Sequence-specific resonance assignments are presented in the table in the Supporting Information. Signal overlap and the lack of sequential NOEs frequently precluded unambiguous complete assignment in distinct cases. Moreover, especially spectra of p30 showed signal broadening mainly attributable to aromatic residues, e.g., the N−H-protons of the indolyl ring system (Figure 3). Compared to the special property of the spider lectin SHL-I, where partial *trans*-to-*cis* isomerization involving Pro31 accounted for presence of two conformations adopted by the C-terminal section from Leu28 to Trp32 (*29*), this type of addition to the conformational space could definitely be excluded for Pro14 in the C-terminus of p30.

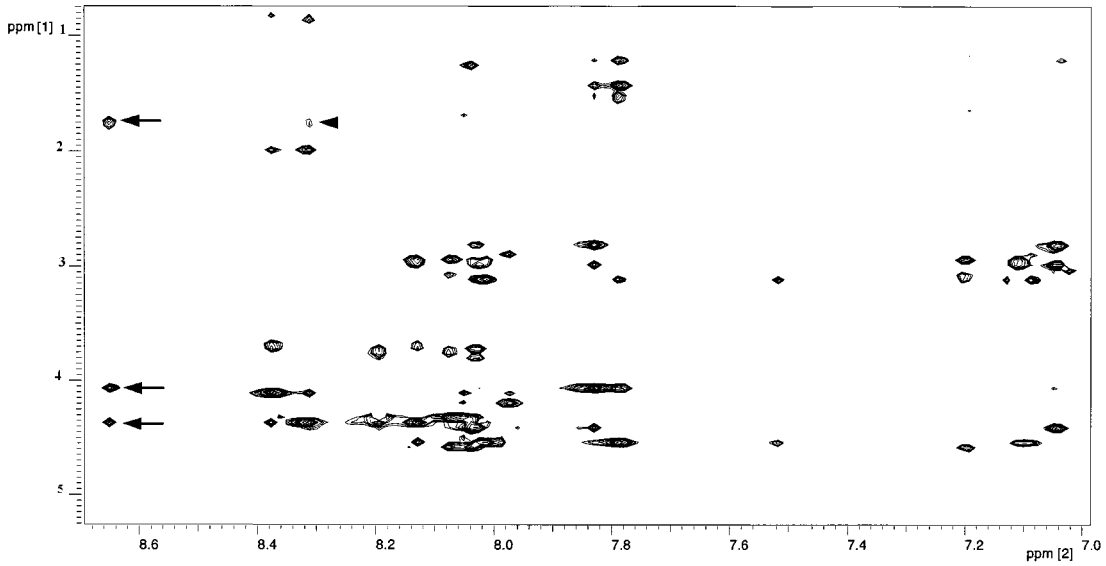

FIGURE 2: Amide-CαH region of a 2D-NOESY spectrum recorded at 750 MHz of p6. The three arrows in its left section identify the cross-peaks originating from the N−H of Arg2 at 8.65 ppm with one interresidual contact to the neighboring Ala1αH at 4.07 ppm and two intraresidual contacts to Arg2βH at 1.77 ppm and to Arg2αH at 4.37 ppm, respectively (see also the table in Supporting Information for resonance assignments). The arrowhead also in the spectrum's left section denotes the NOE contact between Arg2βH and the N−H of the neighboring Val3. For further details, please see information, given in the Materials and Methods.

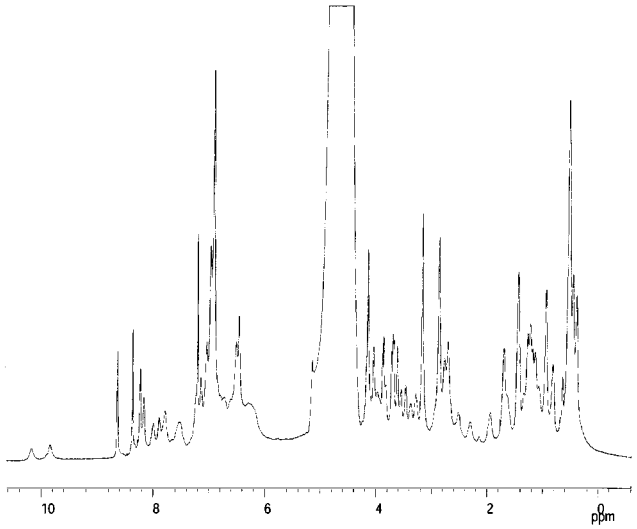

FIGURE 3: 1D-NMR spectrum of p30 recorded at 750 MHz. Measurements were performed in 90% H_2O/10% D_2O at 33 °C with a pentadecapeptide concentration of 4 mM, revealing line broadening especially of the two signals around 10 ppm originating from the N−H protons of the two Trp ring systems (see also the table in Supporting Information for resonance assignments).

In the context of discussing this aspect, it is also noteworthy that carbohydrate-binding activity of the peptides is not dependent on the presence of metal ions (*43*, *44*). They are known to trigger peptide bond isomerization in a mammalian C-type lectin, in concanavalin A or short-chain peptides (*71*−*73*). At this stage of the study it was fair to say that these three peptides are flexible despite the presence of aromatic clusters in their sequences. Using an independent approach, i.e., laser photo CIDNP experiments (*27*, *57*, *58*), we then investigated whether aromatic side chains can be easily contacted by other compounds in solution, as predicted from the previous set of experiments. Indeed, the ring systems sensitive in this technique (Trp, Tyr, His) are fully reactive to the indicator dye, excluding their integration into a hydrophobic core (not shown).

The quantitation of the surface accessibilities of these residues by the laser photo CIDNP technique yielded the following data for selected conformations of p30 as well as of p6 which are compiled in Table 1. These sets of quantitative data indicate that surface accessibility is subject to a certain degree of restriction even for the so-called stretched conformations.

As shown in the example above, the conclusions drawn from the NMR-spectroscopical analysis are supported by theoretical calculations, i.e., MD simulations in the explicit presence of water molecules which will be outlined in detail in the next paragraph.

Molecular Modeling of Ligand-Free Peptides. As noted in the respective section of Materials and Methods, we have made inspection of conformational flexibility calculated by MD simulations for p6 and p30 publicly accessible. When viewing the conformational behavior, it becomes evident that the peptides did not appear to form a rigid hydrophobic core for a substantial fraction of time. To obtain a more subtle measure of flexibility and inherent limits to it not readily registered upon visual inspection, we have processed the complete data set to yield relative probabilities of residue positioning in volume elements of 1 $Å^3$ (voxels), visualized in Figure 4 by plotting isodensity surfaces. The uneven distribution which Figure 4 presents for an exemplary case, i.e., the C-terminal amino acids, is an indication for preferred voxel occupancy. These results argue against a fully random movement of the side chains. Combining the results from the NMR experiments and the computational analysis, we can conclude that we deal with flexible chains in each of the three peptides. Neither formation of a rigid core nor entirely unrestricted fluctuations characterize their conformational behavior. Because aromatic side chains are certain to become accessible to the solvent, the observed line broadening especially for p30 might be explained by intermolecular exchange processes.

To answer the arising question as to whether the peptides form aggregates, we took advantage of the recent develop-

Table 1: Values of the Surface Accessibility (Å^2) of Laser-Photo-CIDNP-Reactive Amino Acids and Phe Calculated for 16 Selected Conformations of p30 (a) and p6 (b)[a]

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
(a)																
His1	170	165	160	136	131	159	142	165	147	145	111	119	165	131	171	136
Phe4	153	183	145	153	154	143	166	176	171	112	99	91	142	138	111	166
Trp8	179	173	181	158	167	164	152	142	207	170	125	142	177	205	180	145
Trp9	154	148	181	143	165	153	162	223	166	95	109	144	136	165	185	177
Tyr10	158	180	137	156	161	136	98	150	193	135	118	164	124	143	123	122
Phe12	162	151	134	163	139	110	120	189	160	140	161	175	146	157	158	121
(b)																
Phe5	182	114	141	110	131	152	149	118	101	166	115	103	115	105	109	150
Trp6	149	158	137	147	167	166	161	147	126	162	131	171	177	175	193	213
Tyr8	128	175	135	159	126	134	143	112	125	133	148	140	162	122	107	180
Phe11	152	125	150	153	155	155	132	171	139	123	109	107	138	156	111	146
Tyr15	198	154	158	166	142	173	130	195	168	177	161	148	162	160	135	125

[a] The first set (1−8) corresponds to stretched conformations, the latter one (9−16) to backfolded conformations, in which the C- and N-termini are close to each other. Dot density, 1; sphere radius, 1.5 Å; for comparison, the values of surface accessibility of the surface-exposed Trp21 ring system in hevein are in the range from 120 to 175 Å^2 (maximal accessibility of Trp residues is about 230 Å^2) (*27*).

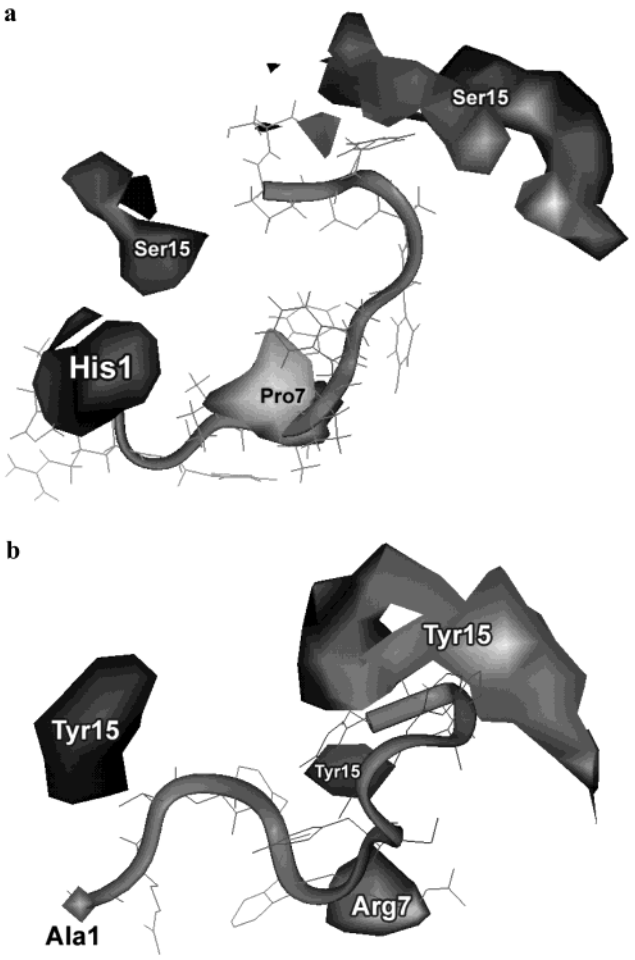

FIGURE 4: Visualization of the relative probabilities for side chain positioning of p30 (a) and p6 (b) in volume elements of 1 Å^3 (voxels) highlighted for exemplary residues by drawing isodensity surfaces. Placing the peptides' amino-terminal residues (His1α and Ala1α) at the coordinate origins and Pro7α (a) or Arg7α (b) on the *x*-axis the isodensity surfaces in the *x*/*y*-plane reveal a nonrandom population of voxels for the C-terminal amino acid residue.

ments of gentle electrospray ionization and also phase transfer from solution to the gas phase. These experimental conditions had even allowed to maintain stability of noncovalent megaDalton protein assemblies (*74*).

Mass Spectrometry. To assay the capacity of the peptides for intermolecular interactions, solutions containing peptide mixtures were prepared. Besides homodimers and oligomers, the peptides might also form heterooligomeric complexes. As shown in Figure 5a for p30 and p6, this is indeed the case. Starting with equimolar concentrations, the presence of homodimers is more pronounced for p30 than for p6, also indicated by the significant line broadening in the corresponding NMR spectrum (Figure 3). Interestingly, intermolecular interactions were restricted to establish homo- and heterodimers under these conditions, with no evidence for further size increase of aggregates beyond dimers. Methodologically, this result shows that peptide−peptide interactions can be reliably picked up by ESI MS even in the absence of matching clusters of basic and acidic residues, as monitored previously for dynorphin fragments with their basic Arg6−Arg7 sequence and the acidic minigastrin with its stretch of Glu residues from position 2 to 5 (*75*). Our data encouraged next examination of the behavior of mixtures of peptide(s) with disaccharides.

Noncovalent binding of the TF-antigen to the peptides could indeed be detected (Figure 5, panels b−d). By running titration in this experimental setting, dissociation constants in the 10 mM range could be calculated based on the ratio of the peak intensities (p6 to p6 + TF-antigen), example given in Figure 5b. An enlarged region of such an ESI MS spectrum of p6 and p30 in the presence of the TF-antigen illustrates the sensitivity of the method and relative differences in the extent of ligand binding for p6 and p30 under these conditions (Figure 5c). Notably, the estimated K_D values are higher than those determined by other authors using fluorescence quench titration (*43*). Apart from the peaks of the peptide−carbohydrate complexes with 1:1 stoichiometry, no evidence for sugar binding to any homo- or heterodimers was detected. To infer the selectivity of binding of the carbohydrate ligand, we analyzed the pattern of complexes formed in the presence of TF-antigen and maltose. As shown in Figure 5d, the diglucoside was also capable to associate with peptides, albeit reaching a comparatively small extent. Similar results were obtained with cellobiose (not shown). In conclusion, the data of Figure 5 documented the capacity of these peptides for dimer formation and for TF-antigen binding with notable carbohydrate

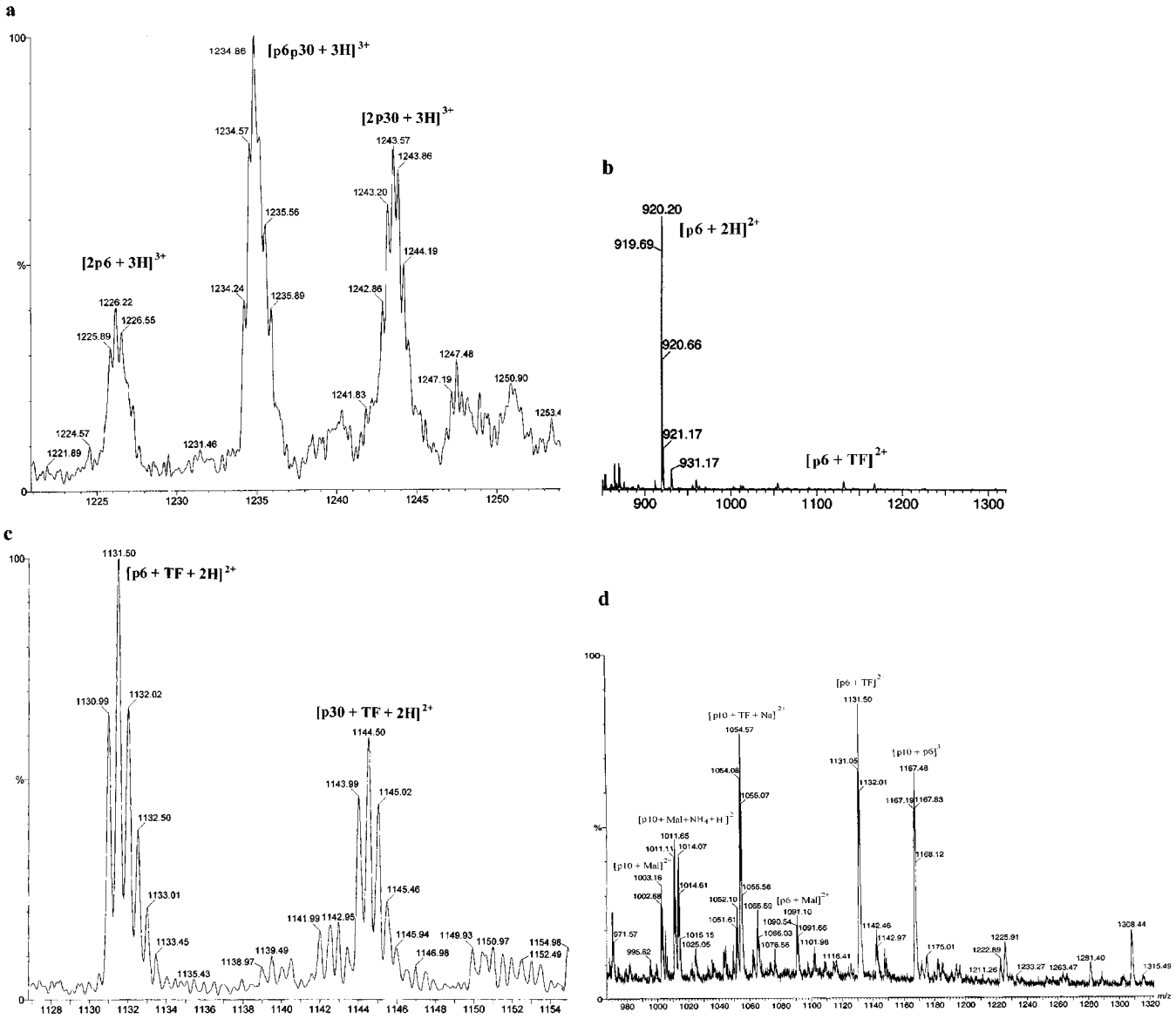

FIGURE 5: Electrospray ionization mass spectra of the pentadecapeptides in the absence or in the presence of carbohydrates (TF-antigen and maltose). The buffer (40 mM ammonium acetate, pH 6.8) contained a mixture of peptides at 20 μM with 100 μM of the carbohydrates. Presence of signals indicative of complex formation involving p6 and p30 (a) and presence of signals for p6 and p6-TF-antigen complexes (b) reveal intermolecular interactions of the reactants. Enlargement of a relevant section of an ESI MS spectrum of the mixture showing signal enhancement for the complexes of p6 and p30 with the TF-antigen illustrates the sensitivity of the method (c). The relative signal intensities for complexes involving either p6 or p10 with either TF-antigen or maltose indicate the level of selectivity of the peptides for carbohydrate binding (d).

selectivity. The presented results underscore the power of ESI MS to gather information on the presence of noncovalent complexes even of rather low affinity. To delineate residues of the peptides responsible for contacting the sugar ligand, we then performed NMR-spectroscopical experiments with solutions containing peptide−carbohydrate mixtures. We also independently confirmed the measured carbohydrate selectivity.

NMR Spectroscopy of Peptide−Carbohydrate Interaction. Signal shifts and/or line broadening are indicative of an association of the carbohydrate with the peptide. We thus scrutinized NMR spectra recorded in the absence and in the presence of sugar for appearance of any alteration(s). In the course of these experimental series, we also confirmed the selectivity determined by ESI MS (see above). Addition of maltose (or cellobiose) to peptide-containing solutions did not affect peak positions and areas, reinforcing the conclusion based on the mass spectrometrical data. Invariably, the NMR spectra of the peptides in the presence of maltose (or cellobiose) were identical to those in the absence of the diglucoside, an example shown for p6 (Figure 6a). Evidently, the phase transfer in ESI MS is not responsible to impair maltose/cellobiose binding to the measured limit, unless one assumes transient association of the disaccharides in solution without any perturbation of peptide-derived proton signals. This negative result further excludes the possibility that the diglucosides with their equatorial 4′-OH-groups could be wedged or even sandwiched between aromatic residues. Such a ligand placement had been seen in crystals for microbial receptors of maltose or cellobiose (*76*, *77*). The combined ESI MS and NMR analyses enabled the conclusion so far that the peptides have selective disaccharide binding properties, extending previous observations (*43*). However, for this conclusion to be fully valid and for shedding light on contact

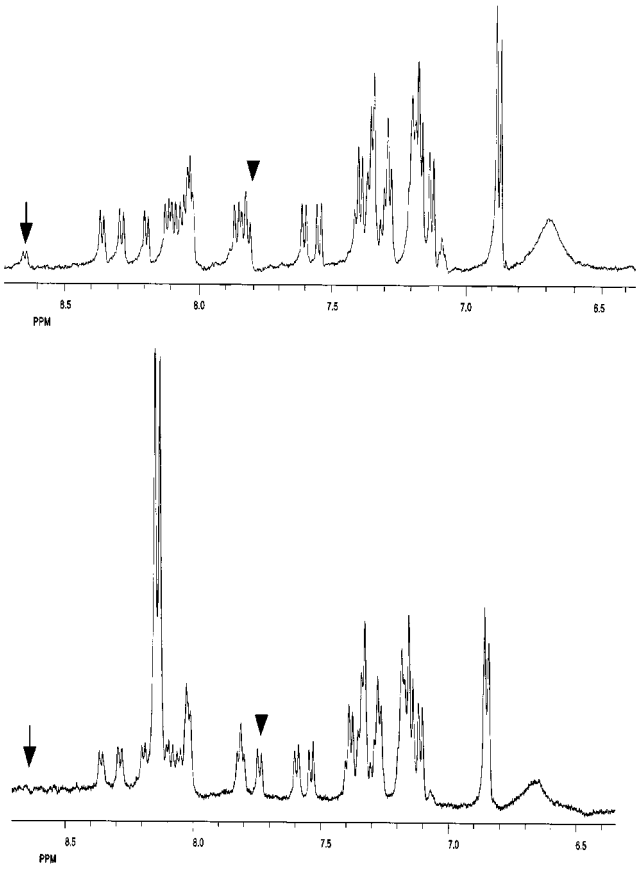

FIGURE 6: Detection of ligand-dependent NMR signal alterations for p6. 1D-NMR spectra of p6 were recorded (for further details, please see information given in Materials and Methods) without carbohydrate addition (a, top) and in the presence of 4 mM TF antigen (b, bottom). Addition of 4 mM maltose to the peptide-containing solution does not affect the signal pattern in the peptide's spectrum. The ligand-induced disappearance of the N−H signal of Arg2 at 8.65 ppm (see also Figure 2) is denoted by an arrow. The upfield shift of the N−H signal of Arg7 is marked by an arrowhead. Please note that the ligand itself introduces a characteristic signal at 8.15 ppm.

sites, the TF-antigen should induce signal alterations in NMR spectroscopy. Consequently, we performed systematic comparative analysis of NMR spectra of the three peptides in the absence and presence of the TF-antigen.

Indeed, we were able to identify signal changes induced by the presence of the ligand. In the spectrum of ligand-free peptide p6 (Figure 6a), the N−H-resonance of Arg2 was clearly visible (see also Figure 2). After addition of the TF-antigen this signal disappeared (Figure 6b). Also, the N−H-signal of Arg7 was subject to an upfield-shift of 0.08 ppm from 7.80 to 7.72 ppm (Figure 6, panels a and b). From the spectra of peptide p10 the loss of the Ser12-N−H resonance in the presence of TF-antigen is readily visible (please see Supporting Information). In addition to this decrease in signal intensity, a shift of the Ala5-N−H signal was detectable. Interestingly, this residue lies within the aromatic cluster of p10 (Figure 1). To obtain information which part of the ligand comes into the vicinity of this section of the peptide, we performed a different set of NMR experiments in this case. Following saturation of peptide resonances in independent STD experiments, we closely examined the signal pattern of the ligand. In line with the rather small but visible alterations of the peptide's spectrum by the ligand (but not

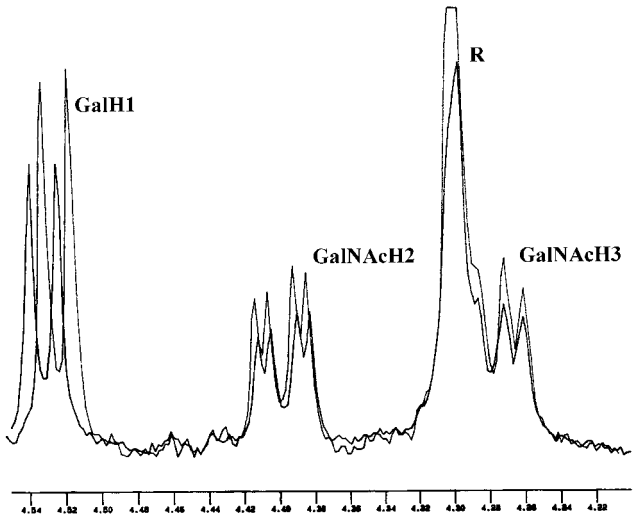

FIGURE 7: Detection of NMR signal shifts of the ligand (TF-antigen) by binding to a receptor protein. The 1D-NMR spectra of the free ligand at the concentration of 4 mM and of the mixture of TF-antigen (4 mM) with TF-specific monoclonal immunoglobulin G (4 μM) were recorded in D_2O at 33 °C. The two spectra were overlayed for direct comparison. Signals of the sugar in complex with the antibody are smaller than those of the sugar in the absence of the receptor. Positions of chemical shifts of GalH1 and GalNAcH2 were altered in the presence of the receptor protein, as denoted by designations in the figure, whereas the signal position of GalNAcH3 was not affected.

by maltose), we were able to pick up contacts between the peptide p10 and the TF-antigen. Irradiation with a frequency characteristic for the signals of aromatic residues and also N−H-protons (7.2 ppm) yielded a weak but significant response of protons from the GalNAcH3 residue (not shown). This detection of a contact to the ligand in this area can be reconciled with laser photo CIDNP data revealing no protection of sensitive rings by the TF-antigen (not shown) by considering the rather low affinity as inferred from ESI MS titrations.

The alanine residue flanked by aromatic amino acids on both sides is also an important source of information for ligand-dependent signal alterations in the case of peptide p30. As shown in the Supporting Information with 2D-NOESY spectra of p30 with and without a ligand, a cross-peak derived from the Ala11-N−H/Ala11αH contact was recorded in the presence of the ligand. This new constraint can arise from altered exchange processes upon dimer dissociation or a ligand-induced conformational change. In the same way, we interpret the significant increase of the cross-peak intensity for the contact Leu6N−H/Ile5αH in the presence of the TF-antigen, documented in the Supporting Information. So far, the analysis had focused on peptide-derived proton signals, monitored directly or exploited in STD experiments. That protons of the TF-antigen likewise have potential to act as sensors in protein-TF-antigen interactions is illustrated in Figure 7 using a monoclonal antibody specific for this disaccharide. When accommodated in the antibody's binding site, the following small but reproducibly detectable changes were picked up: the signal of GalH1 was shifted downfield by 0.01 ppm and that of GalNAcH2 moved upfield by 0.005 ppm (Figure 7). The size of this shift for GalH1 lies in the range of the ppm-alteration measured for the Gal′H1 signal of Gal′β1,2Galβ1,R after the addition of the avian galectin CG-16 (*78*). In contrast to this, e.g., the signal position of

GalNAcH3 was not affected. In addition to these signal shifts, trNOE signals were recorded from the mixture of the TF-antigen and the IgG_3 antibody (not shown). Our experiments with the peptides did not pick up any change in peak position or intensity of the carbohydrate signals. The low affinity of the peptides compared to proteins (antibody and lectins) obviously restricted the possibility to obtain further NMR-derived information. This result corroborated the order of magnitude of affinity, as measured by mass spectrometry. However, the signal alterations of the peptides in the presence of the TF-antigen are helpful to identify interaction sites. This information served as input to the computer-assisted theoretical calculations of the dynamic behavior of peptides in the presence of the TF-antigen.

MD Simulations and Ligand Docking. The first issue to be addressed in this part of the study is whether the presence of the ligand might reduce the inherent flexibility of the peptides in the calculations. As outlined in the introductory portion of this paper, an aptamer can guide a peptide to acquire secondary structure (*45*). In our cases, the NOESY spectra (please see Figure 2 and Supporting Information for examples) of p6, p10, and p30 argued against a remarkable stabilization. As already indicated, these spectra provided salient experimental information for the MD runs concerning the initial positioning of the ligand. Computation of inter-proton distances from selected residues clearly showed that the disaccharide did not exert a rigidifying effect (please see Supporting Information). More graphically, MD runs visualize the relative movement for the TF-antigen and the peptides, respective files for p6 and p30 being publicly accessible (please see Materials and Methods section *Molecular Modeling* for details). Their inspection will underscore that the carbohydrate appears to have no notable effect to induce secondary structure. The MD-derived information was further processed to see whether any nonneighboring amino acid side chains might be engaged in interactions indicative of transient contacts. Within such a correlation plot, any contacts of side chains of amino acids not directly linked are located in the area above or below the diagonal line. While no firm indication for a stable long-range contact was seen in peptides p6 and p30, the positive controls, i.e., running this protocol with complexes of hevein-$(GlcNAc)_2$ and peanut agglutin/TF-antigen found such predicted long-range contacts (please see Supporting Information). These calculated contact sites are in line with published data for the two plant lectins (*22*−*24*, *47*, *48*), a proof of the reliability of the applied computational technique for these proteins. Further processing of the data gleaned from the interaction analysis into an estimation of Coulomb forces and *van der Waals* contributions yielded a value of about 7 kcal/mol for the p30-sugar complex compared to about 30 kcal/mol for the interaction of the disaccharide with the binding site of PNA.

Admittedly, the peptides with their inherent flexibility can at present hardly be suggested to match the technical perfection of the topological arrangement of key residues from four sequentially separated loops held in place by a lectin's backbone scaffold, as seen for PNA (*47*, *48*, *63*). Nonetheless, our combined analysis proves complex formation and ligand selectivity experimentally by ESI MS and NMR spectroscopy. One should on no account expect a peptide without previous cycles of structural optimization to be able to fully mimic the binding architecture of role-model-like lectins. In fact, even the binding mode of a lectin-derived peptide may not accurately reflect the set of interactions which the corresponding sequence encounters in the context of the lectin framework. To give a telling example, the contiguous, nonpolar face of the heptapeptide YYWIGIR−NH_2 (with the essential presence of, e.g., W proven by substitutional analysis), a sialyl-Lewis^x-binding sequence derived from the C-type lectin domain of selectins, hardly enables the conclusion that E- and P-selectin binding of the sialylated Lewis epitopes is almost entirely electrostatic with no evidence for stacking involving a tryptophan (*37*, *79*).

Conclusions and Perspectives. Our data reveal that neither the aromatic sequence cluster nor the interaction with the disaccharide introduced significant constraints to the peptides' conformational flexibility. Therefore, it would at present be premature to deliberate substitutions of distinct amino acids with unnatural analogues in the course of solid-phase synthesis. Enhancements of carbohydrate affinity had been achieved with size increases of one aromatic ring in the case of Ac-AMP2, namely exchange of Phe18 by Trp or β-(1(or 2)-naphthyl)alanine (*80*). The primary aim will be to reduce the apparent conformational freedom and thus the entropic barrier. Toward this end, a salient lesson can be learned from small microbial chitin-binding modules and also the hevein-like domains. Here, cysteine residues often sit in convenient places to form disulfide bonds (*31*, *81*). As in other relevant cases, the chitin-binding protein of *Streptomyces* (CHB1) loses its target-binding capacity after cleavage of its two disulfide bonds (*82*). On the basis of these facts, redesigning by protein engineering (i.e., strategically introducing cysteine residues) had been suggested for the microbial Chitinase A1 mentioned above to further increase its stability with implications for industrial use. In analogy to the structure of a related domain in the endoglucanase Z of *Erwinia chrysanthemi*, the following concept had been proposed: to attempt to exploit the stability-conferring disulfide bonding seen in that domain with its disulfide bridge between the domain extremities at Cys4−Cys61 by introducing Cys residues (*33*). Using the given peptide sequences (Figure 1) as lead compounds, this approach is also worthy to be tested thoroughly in our case. Besides the example of naturally tested chitin-binding domains, the cyclization of an α-bungarotoxin-binding 13-mer peptide by synthetic substitutions with Cys residues (here Met1 and Pro12) convincingly demonstrated how this bridging can increase the target affinity, in this case by 2 orders of magnitude (*83*). If the modified peptide will still access conformation(s) with ligand-binding properties, our combined panel of techniques for probing peptide-disaccharide interactions will be conducive to identify the contact area with sufficient accuracy to provide hints for further engineering. One reasonable option could then be to replace natural amino acids by synthetic derivatives, a notable advantage of working with peptides. Moreover, by having ingeniously mastered the challenge to obtain carbohydrate-binding peptides composed of unnatural enantiomers which resist proteolytic degradation in vivo (*84*), biomedical applications of suitable products could be envisioned in drug delivery, imaging, and anti-adhesion/proliferation approaches. Due to the still small molecular weight per peptide unit, their

clustering on a matrix, e.g., linear, cyclic, or dendrimer scaffolds (*85*−*92*), by covalent tethering might enable to turn multivalency into remarkable target selectivity without reaching very large overall sizes.

In conclusion, we herein report that three TF-antigen-binding pentadecapeptides show target selectivity for the disaccharide, using independent experiments by ESI MS and NMR spectroscopy. Hydrophobic clusters in the sequence were not sufficient to build a compact core. In contrast, conformational flexibility allowed surface presentation of these residues. MD runs complemented the analysis by computationally exploring the conformational space of the backbone and side chains, e.g., visualizing surface accessibilities of the aromatic rings. This factor is likely to contribute to formation of homo- and heterodimers seen in ESI MS spectra. At this level of inherent flexibility, the strength of contact to the ligand, monitored by signal alterations in NMR spectra and formation of noncovalent complexes in ESI MS, was not (yet) sufficient to limit the conformational space of peptides. Whether and to what extent the involvement of an ionic interaction in this type of molecular recognition process can exert a favorable impact, is currently being tested with ganglioside GM_1-binding peptides. On the basis of the discussion in the preceding paragraph and the presented results, rational synthetic tailoring of the three 15-mer peptides (Figure 1) will be the next step to turn the establishment of this integrated approach into new insights into the ways how carbohydrates interact with minimally sized binding sites established by peptides.

ACKNOWLEDGMENT

We sincerely thank Prof. Dr. Janusz Dabrowski from the Max-Planck-Institut für Medizinische Forschung (Heidelberg, Germany) for valuable help and indispensable support of this research project. This study used spectrometers at the European SON NMR Large-Scale Facility in Utrecht (NL). We thank Dr. R. Pipkorn (DKFZ-Heidelberg, Germany) for careful synthesis of the peptides.

SUPPORTING INFORMATION AVAILABLE

Table of ^{1}H-chemical shift assignments for p6, p10, and p30 and figures of 1D-NMR spectra of p10, 2D NOESY spectra of p30, interproton distance computations for p30 and p6, and residue interaction correlation plots of p30, p6, chitin-binding hevein, and peanut agglutinin. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Laine, R. A. (1997) in *Glycosciences: Status and Perspectives* (Gabius, H.-J., and Gabius, S., Eds.) pp 1−14, Chapman & Hall, London, Weinheim.
2. Gabius, H.-J. (2000) *Naturwissenschaften 87*, 108−121.
3. Fukuda, M., Hiraoka, N., and Yeh, J.-C. (1999) *J. Cell Biol. 147*, 467−470.
4. Gabius, H.-J., and Gabius, S., Eds. (1997) *Glycosciences: Status and Perspectives*, Chapman & Hall, London, Weinheim.
5. Varki, A., Cummings, R. D., Esko, J., Freeze, H. H., Hart, G. W., and Marth, J., Eds. (1999) *Essentials of Glycobiology*, Cold Spring Harbor Laboratory Press, New York.
6. Doyle, R. J., Ed. (2000) *Glycomicrobiology*, Kluwer Academic/Plenum Publishers, New York.
7. Ernst, B., Hart, G. W., and Sinay, P., Eds. (2000) *Carbohydrates in Chemistry and Biology*, Wiley-VCH, Weinheim.
8. Quiocho, F. A. (1986) *Annu. Rev. Biochem. 55*, 287−315.
9. Vyas, N. K. (1991) *Curr. Opin. Struct. Biol. 1*, 732−740.
10. Toone, E. J. (1994) *Curr. Opin. Struct. Biol. 4*, 719−728.
11. Gabius, H.-J. (1997) *Cancer Invest. 15*, 454-464.
12. von der Lieth, C.-W., Siebert, H.-C., Kozár, T., Burchert, M., Frank, M., Gilleron, M., Kaltner, H., Kayser, G., Tajkhorshid, E., Bovin, N. V., Vliegenthart, J. F. G., and Gabius, H.-J. (1998) *Acta Anat. 161*, 91−109.
13. Lis, H., and Sharon, N. (1998) *Chem. Rev. 98*, 637−674.
14. Jiménez-Barbero, J., Asensio, J. L., Cañada, F. J., and Poveda, A. (1999) *Curr. Opin. Struct. Biol. 9*, 549−555.
15. Rüdiger, H., Siebert, H.-C., Solís, D., Jiménez-Barbero, J., Romero, A., von der Lieth, C.-W., Díaz-Mauriño, T., and Gabius, H.-J. (2000) *Curr. Med. Chem. 7*, 389−416.
16. Dam, T. K., and Brewer, C. F. (2002) *Chem. Rev. 102*, 387−429.
17. Gabius, H.-J. (1997) *Eur. J. Biochem. 243*, 543−576.
18. Quesenberry, M. S., and Drickamer, K. (1991) *Glycobiology 1*, 615−621.
19. Moriki, T., Kuwabara, I., Liu, F.-T., and Maruyama, I. N. (1999) *Biochem. Biophys. Res. Commun. 265*, 291−296.
20. Nitta, K., Oyama, F., Oyama, R., Sekiguchi, K., Kawauchi, H., Takayanagi, Y., Hakomori, S., and Titani, K. (1993) *Glycobiology 3*, 37−45.
21. Nitta, K., Ozaki, K., Tsukamoto, Y., Hosono, M., Ogawa-Konno, Y., Kawauchi, H., Takayanagi, Y., Tsuiki, S., and Hakomori, S. (1996) *Int. J. Oncol. 9*, 19−23.
22. Andersen, N. H., Cao, B., Rodriguez-Romero, A., and Arreguin, B. (1993) *Biochemistry 32*, 1407−1422.
23. Asensio, J. L., Cañada, F. J., Bruix, M., Rodriguez-Romero, A., and Jiménez-Barbero, J. (1995) *Eur. J. Biochem. 230*, 621−633.
24. Asensio, J. L., Cañada, F. J., Bruix, M., Gonzalez, C., Khiar, N., Rodriguez-Romero, A., and Jiménez-Barbero, J. (1998) *Glycobiology 8*, 569−577.
25. Asensio, J. L., Siebert, H.-C., von der Lieth, C.-W., Laynes, J., Bruix, M., Soedjanaatmadja, U. M. S., Beintema, J. J., Cañada, F. J., Gabius, H.-J., and Jiménez-Barbero, J. (2000) *Proteins 40*, 218−236.
26. Asensio, J. L., Cañada, F. J., Siebert, H.-C., Laynez, J., Poveda, A., Nieto, P. M., Soedjanaatmadja, U. M. S., Gabius, H.-J., and Jiménez-Barbero, J. (2000) *Chem. Biol. 7*, 529−543.
27. Siebert, H.-C., von der Lieth, C.-W., Kaptein, R., Beintema, J. J., Dijkstra, K., van Nuland, N., Soedjanaatmadja, U. M. S., Rice, A., Vliegenthart, J. F. G., Wright, C. S., and Gabius, H.-J. (1997) *Proteins 28*, 268−284.
28. Martins, J. C., Maes, D., Loris, R., Pepermans, H. A. M., Wyns, L., Willem, R., and Verheyden, P. (1996) *J. Mol. Biol. 258*, 322−333.
29. Lü, S., Liang, S., and Gu, X. (1999) *J. Protein Chem. 18*, 609−617.
30. Suetake, T., Tsuda, S., Kawabata, S., Miura, K., Iwanaga, S., Hikichi, K., Nitta, K., and Kawano, K. (2000) *J. Biol. Chem.* 275, 17929−17932.
31. Shen, Z., and Jacobs-Lorena, M. (1999) *J. Mol. Evol. 48*, 341−347.
32. Siebert, H.-C., André, S., Asensio, J. L., Cañada, F. J., Dong, X., Espinosa, J.-F., Frank, M., Gilleron, M., Kaltner, H., Kozár, T., Bovin, N. V., von der Lieth, C.-W., Vliegenthart, J. F. G., Jiménez-Barbero, J., and Gabius, H.-J. (2000) *ChemBioChem 1*, 181−195.
33. Ikegami, T., Okada, T., Hashimoto, M., Seino, S., Watanabe, T., and Shirakawa, M. (2000) *J. Biol. Chem. 275*, 13654−13661.
34. Geng, J. G., Heavner, G. A., and McEver, R. P. (1992) *J. Biol. Chem. 267*, 19846−19853.
35. Heerze, L. D., Chong, P. C., and Armstrong, G. D. (1992) *J. Biol. Chem. 267*, 25810−25815.
36. Heerze, L. D., Smith, R. H., Wang, N., and Armstrong, G. D. (1995) *Glycobiology 5*, 427−433.
37. Briggs, J. B., Larsen, R. A., Harris, R. B., Sekar, K. V., and Macher, B. A. (1996) *Glycobiology 6*, 831−836.
38. Heavner, G. A. (1996) *Drug Discovery Today 1*, 295−304.
39. Yamamoto, K., Konami, Y., and Osawa, T. (1992) *J. Chromatogr. 597*, 221−230.
40. Yamamoto, K., Konami, Y., Osawa, T., and Irimura, T. (1992) *J. Biochem. (Tokyo) 111*, 436−439.
41. Matsubara, T., Ishikawa, D., Taki, T., Okahata, Y., and Sato, T. (1999) *FEBS Lett. 456*, 253−256.
42. Matsubara, T. (2001) *Trends Glycosci. Glycotechnol. 13*, 557−560.
43. Peletskaya, E. N., Glinsky, V. V., Glinsky, G. V., Deutscher, S. L., and Quinn, T. P. (1997) *J. Mol. Biol. 270*, 374−384.

44. Glinsky, V. V., Huflejt, M. E., Glinsky, G. V., Deutscher, S. L., and Quinn, T. P. (2000) *Cancer Res. 60*, 2584−2588.
45. Leulliot, N., and Varani, G. (2001) *Biochemistry 40*, 7947−7956.
46. Wright, P. E., and Dyson, H. J. (1999) *J. Mol. Biol. 293*, 321−331.
47. Banerjee, R., Das, K., Ravishankar, R., Suguna, K., Surolia, A., and Vijayan, M. (1996) *J. Mol. Biol. 259*, 281−296.
48. Ravishankar, R., Ravindran, M., Suguna, K., Surolia, A., and Vijayan, M. (1997) *Curr. Sci. 72*, 855−861.
49. Gabius, H.-J., Schröter, C., Gabius, S., Brinck, U., and Tietze, L.-F. (1990) *J. Histochem. Cytochem. 38*, 1625−1631.
50. Pollex-Kruger, A., Meyer, B., Stuike-Prill, R., Sinnwell, V., Matta, K. L., and Brockhausen, I. (1993) *Glycoconjugate J. 10*, 365−380.
51. Rittenhouse-Diakun, K., Xia, Z., Pickhardt, D., Morey, S., Baek, M.-G., and Roy, R. (1998) *Hybridoma 17*, 165−173.
52. Roy, R., Baek, M.-G., and Rittenhouse-Olson, K. (2001) *J. Am. Chem. Soc. 123*, 1809−1816.
53. Baek, M.-G., Rittenhouse-Olson, K., and Roy, R. (2001) *Chem. Commun.* 257−258.
54. Jorgensen, T. J. D., Roepstorff, P., and Heck, A. J. R. (1998) *Anal. Chem. 70*, 4427−4432.
55. Van Dongen, W. D., and Heck, A. J. R. (2000) *Analyst 125*, 583−589.
56. Gilleron, M., Siebert, H.-C., Kaltner, H., von der Lieth, C.-W., Kozár, T., Halkes, K. M., Korchagina, E. Y., Bovin, N. V., Gabius, H.-J., and Vliegenthart, J. F. G. (1998) *Eur. J. Biochem. 252*, 416−427.
57. Siebert, H.-C., Tajkhorshid, E., von der Lieth, C.-W., Kleineidam, R. G., Kruse, S., Schauer, R., Kaptein, R., Gabius, H.-J., and Vliegenthart, J. F. G. (1996) *J. Mol.* Model. *2*, 446−455.
58. Siebert, H.-C., Adar, R., Arango, R., Burchert, M., Kaltner, H., Kayser, G., Tajkhorshid, E., von der Lieth, C.-W., Kaptein, R., Sharon, N., Vliegenthart, J. F. G., and Gabius, H.-J. (1997) *Eur. J. Biochem. 249*, 27−38.
59. Mayer, M., and Meyer, B. (1999) *Angew. Chem., Int. Ed. Engl. 38*, 1784−1788.
60. Klein, J., Meinecke, R., Mayer, M., and Meyer, B. (1999) *J. Am. Chem. Soc. 121*, 5336−5337.
61. Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, S., Jr., and Weiner, P. (1984) *J. Am. Chem. Soc. 106*, 765−784.
62. Weiner, S. J., Kollman, P. A., Nguyen, D. T., and Case, D. A. (1986) *J. Comput. Chem. 7*, 230−252.
63. Pratap, J. V., Bradbrook, G. M., Reddy, G. B., Surolia, A., Raftery, J., Helliwell, J. R., and Vijayan, M. (2001) *Acta Crystallogr., Sect. D 57*, 1584−1594.
64. Siebert, H.-C., von der Lieth, C.-W., Dong, X., Reuter, G., Schauer, R., Gabius, H.-J., and Vliegenthart, J. F. G. (1996) *Glycobiology 6*, 561−572.
65. Kanyalkar, M., Srivastava, S., and Coutinho, E. (2001) *J. Pept. Sci. 7*, 579−587.
66. Venkatraman, J., Shankaramma, S. C., and Balaram, P. (2001) *Chem. Rev. 101*, 3131−3152.
67. Kaltner, H., and Stierstorfer, B. (1998) *Acta Anat. 161*, 162−179.
68. Reuter, G., and Gabius, H.-J. (1999) *Cell Mol. Life Sci. 55*, 368-422.
69. Gabius, H.-J. (1998) *Pharmaceut. Res. 15*, 23−30.
70. Loris, R., Hamelryck, T., Bouckaert, J., and Wyns, L. (1998) *Biochim. Biophys. Acta 1383*, 9−36.
71. Ng, K. K.-S., and Weis, W. I. (1998) *Biochemistry 37*, 17977−17989.
72. Bouckaert, J., Dewallef, Y., Poortmans, F., Wyns, L., and Loris, R. (2000) *J. Biol. Chem. 275*, 19778−19787.
73. Gaggelli, E., D'Amelio, N., Gaggelli, N., and Valensin, G. (2001) *ChemBioChem 2*, 524−529.
74. van Berkel, W. J. H., van den Heuvel, R. H. H., Versluis, C., and Heck, A. J. R. (2000) *Protein Sci. 9*, 435−439.
75. Woods, A. S., and Huestis, M. A. (2001) *J. Am. Soc. Mass Spectrom. 12*, 88−96.
76. Spurlino, J. C., Lu, G.-Y., and Quiocho, F. A. (1991) *J. Biol. Chem. 266*, 5202−5219.
77. Notenboom, V., Boraston, A. B., Kilburn, D. G., and Rose, D. R. (2001) *Biochemistry 40*, 6248−6256.
78. Siebert, H.-C., Gilleron, M., Kaltner, H., von der Lieth, C.-W., Kozár, T., Bovin, N. V., Korchagina, E. Y., Vliegenthart, J. F. G., and Gabius, H.-J. (1996) *Biochem. Biophys. Res. Commun. 219*, 205−212.
79. Somers, W. S., Tang, J., Shaw, G. D., and Camphausen, R. T. (2000) *Cell 103*, 467−479.
80. Muraki, M., Morii, H., and Harata, K. (2000) *Protein Eng. 13*, 385−389.
81. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. (1991) *Microbiol. Rev. 55*, 303−315.
82. Svergun, D. I., Becirevic, A., Schrempf, H., Koch, M. H. J., and Grüber, G. (2000) *Biochemistry 39*, 10677−10683.
83. Balass, M., Kalef, E., Fuchs, S., and Katchalski-Katzir, E. (2001) *Toxicon 39*, 1045−1051.
84. Kozlov, I. A., Mao, S., Xu, Y., Huang, X., Lee, L., Sears, P. S., Gao, C., Coyle, A. R., Janda, K. D., and Wong, C.-H. (2001) *ChemBioChem 2*, 741−746.
85. Gabius, H.-J. (2001) *Anat. Histol. Embryol. 30*, 3−31.
86. Lee, Y. C., and Lee, R. T., Eds. (1994) *Neoglycoconjugates. Preparation and Applications*, Academic Press, San Diego.
87. Roy, R. (1996) *Trends Glycosci. Glycotechnol. 8*, 79−99.
88. Roy, R. (1996) *Curr. Opin. Struct. Biol. 6*, 692−702.
89. André, S., Cejas Ortega, P. J., Perez, M. A., Roy, R., and Gabius (1999) *Glycobiology* 9, 1253−1261.
90. Yamazaki, N., Kojima, S., Bovin, N. V., André, S., Gabius, S., and Gabius, H.-J. (2000) *Adv. Drug. Deliv. Rev. 43*, 225−244.
91. André, S., Pieters, R. J., Vrasidas, I., Kaltner, H., Kuwabara, I., Liu, F.-T., Liskamp, R. M. J., and Gabius, H.-J. (2001) *ChemBioChem 2*, 822−830.
92. Fulton, D. A., and Stoddart, J. F. (2001) *Bioconjugate Chem. 12*, 655−672.

BI025891X