

Insights into the Dynamics and Molecular Recognition Features of Glycopeptides by Protein Receptors: The 3D Solution Structure of Hevein Bound to the Trisaccharide Core of *N*-Glycoproteins

José Juan Hernández-Gay,^[a] Ana Ardá,^[a, c] Steffen Eller,^[b] Stefano Mezzato,^[b]
Bas R. Leeftang,^[c] Carlo Unverzagt,^[b] F. Javier Cañada,^{*[a]} and
Jesús Jiménez-Barbero^{*[a]}

Dedicated to Professor Horst Kessler on the occasion of his 70th birthday

Abstract: Protein-carbohydrate interactions are at the heart of a variety of essential molecular recognition events. Hevein, a model lectin related to the superantigen family, recognizes the trisaccharide core of *N*-glycoproteins (**1**). A combined approach of NMR spectroscopy and molecular modeling has permitted us to demonstrate that an Asn-linked Man(GlcNAc)₂ (**2**) is bound with even higher affinity than

(GlcNAc)₃. The molecular recognition process entails conformational selection of only one of the possibilities existing for chitooligosaccharides. The deduced 3D structure of the hevein/**2** complex permits the extension of poly-

peptide chains from the Asn moiety of **2**, as well as glycosylation at Man O-3 and Man O-6 of the terminal sugar. Given the ubiquity of the Man-(GlcNAc)₂ core in all mammalian *N*-glycoproteins, the basic recognition mode presented herein might be extended to a variety of systems with biomedical importance.

Keywords: allergies • carbohydrate binding • glycoproteins • hevein • NMR spectroscopy

Introduction

The study of molecular recognition processes from the chemical perspective is of paramount importance to understand and modulate key processes in nature. Chemical synthesis methods employed in combination with other spectro-

scopic, biophysical, and theoretical protocols is a powerful tool to obtain detailed structural information of molecules with atomic resolution. The features that mediate the interactions between biomolecules related to a variety of biological and biomedical problems can be understood through this tool box. In this context, and within glycosciences,^[1–6] lectins are carbohydrate-specific binding proteins that mediate the transfer of biological information from the sugar code and have been widely used as tools in different areas of (bio)-chemical investigations.^[1–6] Some of these lectins from plant origin, dubbed hevein domains, bind reversibly to chitin, a key structural component of the fungi cell wall and invertebrate exoskeletons.^[7–12] Many of the proteins containing hevein domains have been associated with antimicrobial and plant defense functions.^[1–6,9] Additionally, hevein itself (43 amino acids; Scheme 1), found in *Hevea brasiliensis*

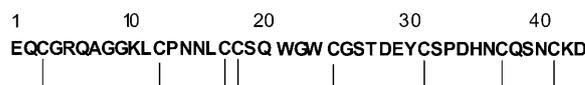
[a] Dr. J. J. Hernández-Gay,⁺ Dr. A. Ardá,⁺ Prof. Dr. F. J. Cañada, Prof. Dr. J. Jiménez-Barbero
Chemical and Physical Biology
Centro de Investigaciones Biológicas, CSIC
Ramiro de Maeztu 9, 28040 Madrid (Spain)
Fax: (+34)91-536-0432
E-mail: jjbarbero@cib.csic.es
jcanada@cib.csic.es

[b] S. Eller, S. Mezzato, Prof. Dr. C. Unverzagt
Bioorganische Chemie, Gebäude NW1
Universität Bayreuth, 95440 Bayreuth (Germany)

[c] Dr. A. Ardá,⁺ Dr. B. R. Leeftang
Bijvoet Center, Faculty of Sciences
Utrecht University, Padualaan 8
NL-3584 CH Utrecht (The Netherlands)

[⁺] These two authors have equally contributed to this work

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/chem.201000939>.



Scheme 1. Hevein's sequence. Disulfide bridges: 3–18, 12–24, 17–31, and 37–41 are indicated.

latex,^[9,13–15] and some other very close homologues present in several fruits have been identified as major allergens responsible for allergic syndromes to latex and fruits.^[7,10–12]

X-ray crystallography,^[11,12] NMR spectroscopy,^[13–16] and microcalorimetry^[15,17,18] studies have provided basic structural information about hevein and its binding features to different glycoligands. Indeed, conformational and dynamic features, as well as thermodynamic data, have been deduced.^[11–18] From the chemical perspective, there are several key amino acid residues involved in the recognition of chitooligosaccharides. The aromatic residues placed in the relative positions Trp 21, Trp 23, and Tyr 30 stabilize the complexes by CH– π interactions^[11–23] and van der Waals contacts. Additionally, the hydroxyl groups of the conserved residues Ser and Tyr (19 and 30 in hevein) are involved in hydrogen bonding with the acetamide carbonyl group and the hydroxyl group at position 3 of a key GlcNAc residue (Figure 1).

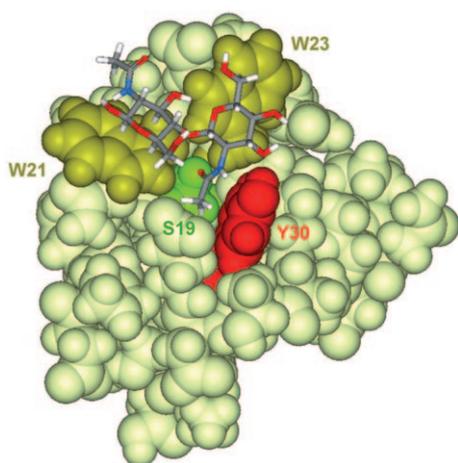


Figure 1. Structure of the complex between hevein and *N,N'*-diacetyl chitobiose (amino acids in the binding site are highlighted).

One of the key features of the molecular recognition process of hevein by chitooligosaccharides is that it displays interesting dynamic features. Indeed, when passing from chitobiose to chitotriose, it has been demonstrated that hevein domains recognize the chitin trimer in two different manners (Figure 2).^[15] There is a first binding mode in which the terminal nonreducing GlcNAc residue is placed at the so-called subsite +1, interacting with Trp 23 (CH– π stacking), with Ser 19 (hydrogen bond), and with Tyr 30 (CH₃– π stacking and hydrogen bond). The intermediate residue is located at subsite +2 interacting with Trp 21 (CH– π stacking), whereas the reducing GlcNAc residue provides very few contacts with the lectin, at subsite +3. In the second orientation, the reducing residue is placed at subsite +2, interacting with Trp 21, and the middle residue is placed at subsite +1, interacting with Trp 23 (CH– π stacking), Ser 19 (hydrogen bond), and with Tyr 30 (CH₃– π stacking and hydrogen

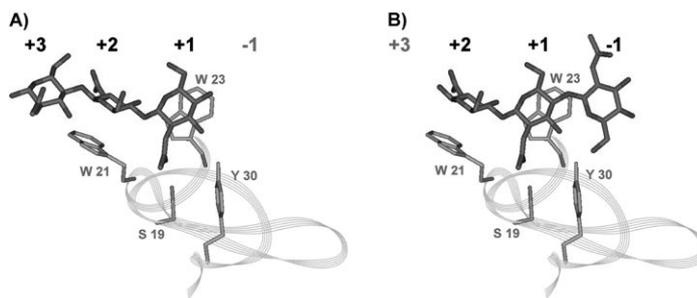
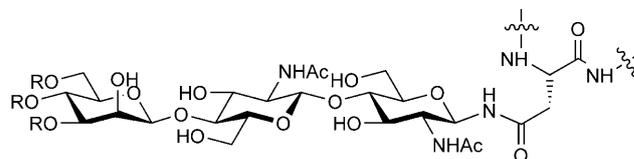


Figure 2. Representation of the two possible binding modes of (GlcNAc)₃ to hevein. A) Orientation +3, +2, +1. B) Orientation +2, +1, –1 (see main text).

bond), whereas the terminal nonreducing end displays some contacts with the lectin at the so-called subsite –1.

For longer chitin chains, the process is multivalent and, depending on the length of the chitooligosaccharide chain, different hevein domains may simultaneously bind to the same chitin chain.^[15]

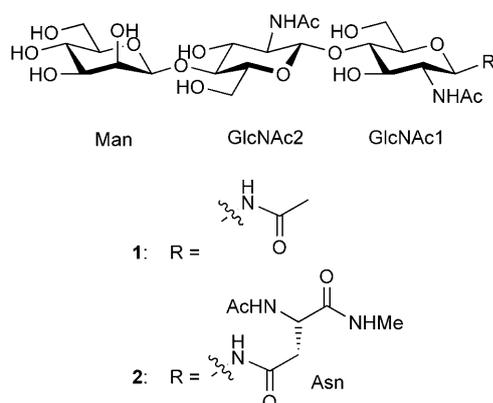
Interestingly, and in relation with its allergenic character, a conformational immunoglobulin binding epitope that includes the aromatic residues has been described in hevein^[11,12,24,25] and hevein recognition by neutrophils is inhibited by chitooligosaccharides and glycoproteins containing *N*-glycosidically linked glycans.^[26] Furthermore it has been described that the *Urtica dioica* lectin (UDA), a covalent hevein-homologue dimer that equally binds chitooligosaccharides,^[26,27] behaves as a superantigen to T cells, inducing exclusive proliferation of V β 8.3 lymphocytes.^[28] In this case, it seems that a key interaction for the biological response takes place through the *N*-glycan chains of mammalian glycoproteins.^[28] It is well known that the *N*-glycan chains in glycoproteins have a common core pentasaccharide. This core contains the *N,N'*-diacetyl chitobiose unit (minimal element recognized by hevein) that connects the glycan chain to the protein through an *N*-glycosidic bond to an asparagine side chain (Scheme 2).



Scheme 2. General structure of the trisaccharide core region of an *N*-glycan chain. In mammalian glycoproteins, the terminal mannose displays additional substitutions.

To further understand a possible recognition of the sugar moiety of mammalian *N*-glycoproteins by hevein domains, and thus its relationship to the described super-antigen behavior of UDA, we have initiated the study of the molecular recognition features of the interaction process between hevein and the trisaccharide core of the *N*-glycoproteins at-

tached to Asn (Scheme 3) by using a combination of organic synthesis, NMR spectroscopy, and molecular modeling procedures. Thus, herein we describe the 3D solution structure



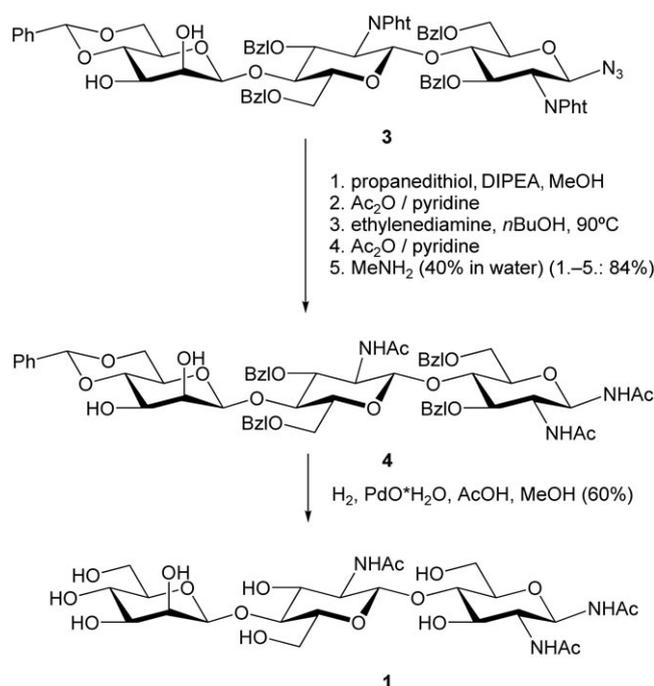
Scheme 3. The structure of the target synthetic models of the trisaccharide *N*-glycan core: Compound **1** is a trisaccharide *N*-acetylated at the reducing end and compound **2** a trisaccharide *N*-linked asparagine at the reducing end.

at atomic resolution of the corresponding complex, as well as key atomic and thermodynamic parameters for the interaction. We aimed at understanding how the presence of the mannose residue at the nonreducing end of chitobiose affects the binding to hevein, as well as the effect of the amino acid moiety at the reducing end. It should be interesting to know how the presence of the additional mannose residue affects (by reducing or enhancing) the dynamic features of the binding process, and the possibility of the formed complex to further interact with other receptors. Thus, extension of the *N*-glycan chain at the mannose residue with natural branches and of the Asn moiety with a polypeptide chain could allow for an interaction of the *N*-glycan-containing glycoprotein with additional receptors.

Results and Discussion

Synthesis of trisaccharide **1 and the glycosylamino acid **2**:** By starting from synthetic core trisaccharide **3**,^[29] compound **4** was obtained after a sequence of five reactions^[30] performed as a one-pot conversion (Scheme 4). After selective reduction of the anomeric azido group of trisaccharide **3** by using propanedithiol, the intermediate glycosylamine was acetylated. Removal of the phthalimido and *O*-acetyl groups followed by complete acetylation and selective *O*-deacetylation gave the benzylated trisaccharide **4** in 84% isolated yield over five steps. The benzyl groups were cleaved by catalytic hydrogenation furnishing the anomerically *N*-acetylated core trisaccharide **1** in 60% yield after purification by gel filtration.

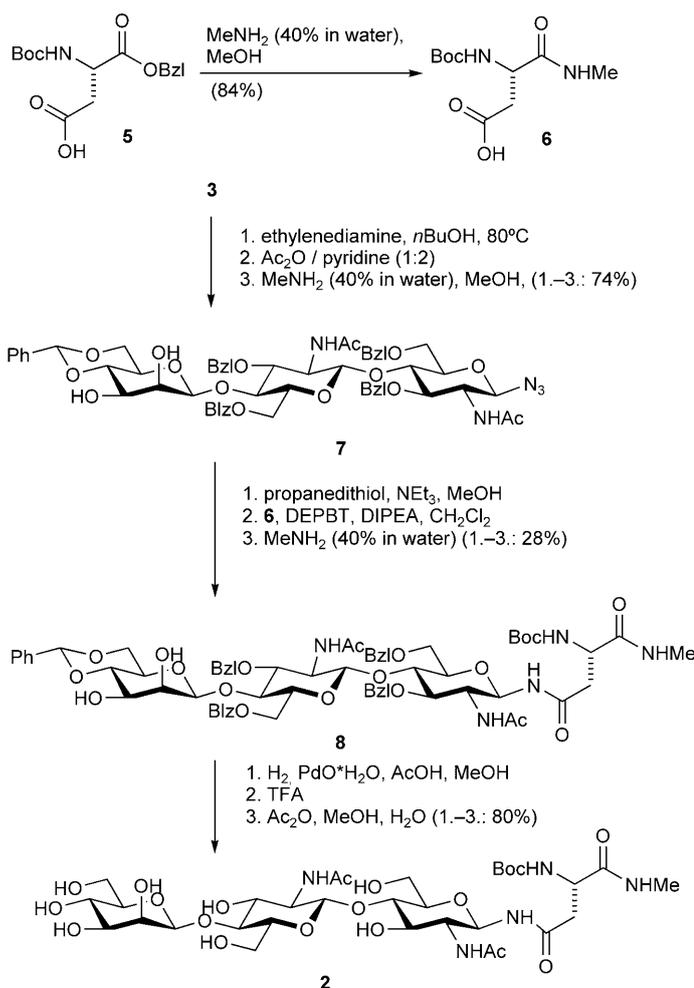
As a step towards the investigation of potential interactions of hevein with glycopeptides containing the core trisaccharide, a second model compound **2** was designed and synthesized (Scheme 5). The rationale for compound **2** was the



Scheme 4. Synthesis of compound **1**. Bzl = benzyl; DIPEA = *N,N*-diisopropylethylamine.

natural linkage of the core trisaccharide to asparagine combined with two minimal amide bonds of the peptide backbone mimicking the linkages of native *N*-glycopeptides. Additionally, by using the *N*-acetylated glycosylasparagine *N*-methylamide instead of the corresponding free asparagine, electrostatic interactions with the lectin were precluded.

In the synthesis of **2**, the introduction of the desired *N*-acetylated asparagine *N*-methyl amide proved unexpectedly demanding. It was initially planned to couple commercially available amino acid building blocks to the glycosylamine generated from the *N*-acetylated trisaccharide **7**. Compound **7** was obtained from **3** in a three-step one-pot dephthaloylation sequence. After reduction of the azido group of **7** with propanedithiol (*Z*)-Asp-OBzl was coupled to the glycosylamine by using benzotriazol-1-yl-oxytripyrrolidinophosphonium hexafluorophosphate (PyBOP) activation (data not shown). Reaction with aqueous methylamine cleaved *O*-acetylated side products and converted the benzyl ester to the methyl amide. Removal of the residual protecting groups by catalytic hydrogenation showed that the main product was *N*-methylated by in situ generated formaldehyde as shown by NMR spectroscopy and mass spectrometry. To avoid *N*-methylation, the Boc-protected aspartic acid benzyl ester **5** was employed instead for the reaction sequence (data not shown). However, after coupling to the glycosylamine followed by *N*-methyl amide formation, catalytic hydrogenation, cleavage of the Boc group, and subsequent *N*-acetylation two products of the desired mass were obtained, which could not be separated. Thus, the *N*-methyl amide was introduced first by reacting Boc aspartic acid benzyl ester **5** with methylamine, giving compound **6**. Amino acid **6** was activat-



Scheme 5. Synthesis of compound **2**. Boc = *tert*-butoxycarbonyl; DEPBT = 3-(diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3*H*)-one; TFA = trifluoroacetic acid.

ed with DEPBT to avoid *O*-acylation of the trisaccharide **8**. Because of residual propanedithiol in the reaction mixture an excess of activated **6** was added. Evaporation of the solvent gave some *O*-acylation, which was removed by reaction with methylamine. After purification by flash chromatography, product analysis by NMR spectroscopy and HPLC-MS showed an anomeric mixture of glycosylamides (α/β 1:2), which could only be separated by RP-HPLC and led to a reduced overall yield (28%). The aromatic protecting groups of the β -linked trisaccharide-asparagine **8** were removed by catalytic hydrogenation. The Boc group was cleaved by TFA followed by selective *N*-acetylation, leading to the desired core trisaccharide-asparagine conjugate **2**.

Conformations of compounds **1** and **2** in the free state

Molecular modeling: MD calculations: A conformational study of both compounds **1** and **2** was carried out by using molecular dynamics (MD) calculations, which were performed by using the MM3* force field,^[31] as implemented in MAESTRO software.^[32] The total time of the simulations

was 6 ns and the GB/SA (generalized Born solvent-accessible surface area) solvent model was used.^[33]

The glycosidic torsions were defined as Φ (H1-C1-O-C4) and Ψ (C1-O-C4-H4) for Man β 1 \rightarrow 4GlcNAc and GlcNAc β 1 \rightarrow 4GlcNAc bonds, and Φ (H1-C1-*N*-C γ) and Ψ (C1-*N*-C γ -C β) for the GlcNAc β 1 \rightarrow Asn glycosylamide bond.

These simulations (Figure 3A) indicated that both Man-GlcNAc and GlcNAc-GlcNAc linkages present a typical *syn-Φ/syn-Ψ* conformation, with Φ and Ψ values of around 60 and 0°, respectively, thus in agreement with the *exo*-anomeric effect.^[34] In contrast, for the GlcNAc-Asn bond, the Φ torsion did not show well-defined minima, but displayed fluctuations between 60 and –60°, whereas Ψ torsion remained in the 180° anti-region as expected for an amide-type bond. The behavior of this glycosyl amide linkage is in accordance with the torsion angles most frequently found (Figure 3B) for the GlcNAc-Asn linkage present in *N*-glycoproteins deposited in the PDB.^[35]

Conformational study by NMR spectroscopy: The obtained simulation results were compared to the experimental data obtained for both compounds **1** and **2** by NMR spectroscopy to assess the modeling conclusions. Thus, in a first step, the ¹H NMR spectroscopic resonances were assigned by using standard COSY, TOCSY, and ROESY/NOESY experiments (see Tables S1 and S2 in the Supporting Information).

The proton scalar coupling constant analysis indicated that all the pyranose rings adopt a typical ⁴C₁ chair geometry. The observed *J*_{5,6} intermediate values indicate that the hydroxymethyl moieties are in concordance with a conformational equilibrium between *gg:gt* conformers in Glc and Man rings, as described for these residues.^[36]

In a second step, ROESY and NOESY experiments were performed to obtain the relevant 3D information.^[37,38] In these spectra, the unique observed inter-residual contacts were those corresponding to cross-peaks between H1 from one ring and H4 from the contiguous one. These peaks are exclusive for a *syn-exo*-anomeric conformation, which permits us to confirm that this is the adopted geometry by both glycosidic bonds. The presence of an anti-type disposition between the anomeric proton of the reducing end GlcNAc1 and the Asn-NH δ proton was confirmed by the existence of an NOE between the Asn-NH δ and H2 from GlcNAc1, thus in accordance to the calculated most-stable conformation of compound **2** (Φ angle of around 0° in the right panel of Figure 3A).

The interaction with hevein: Thermodynamic analysis of the binding:

The interaction of **1** and **2** with hevein was studied by NMR spectroscopy by using a well-established methodology.^[13–16] The association was studied by 1D ¹H NMR titrations by following the procedure described in the Materials and Methods Section, which is based on the analysis of 1D spectra recorded for series of samples containing a constant concentration of polypeptide with increasing ligand concentrations.

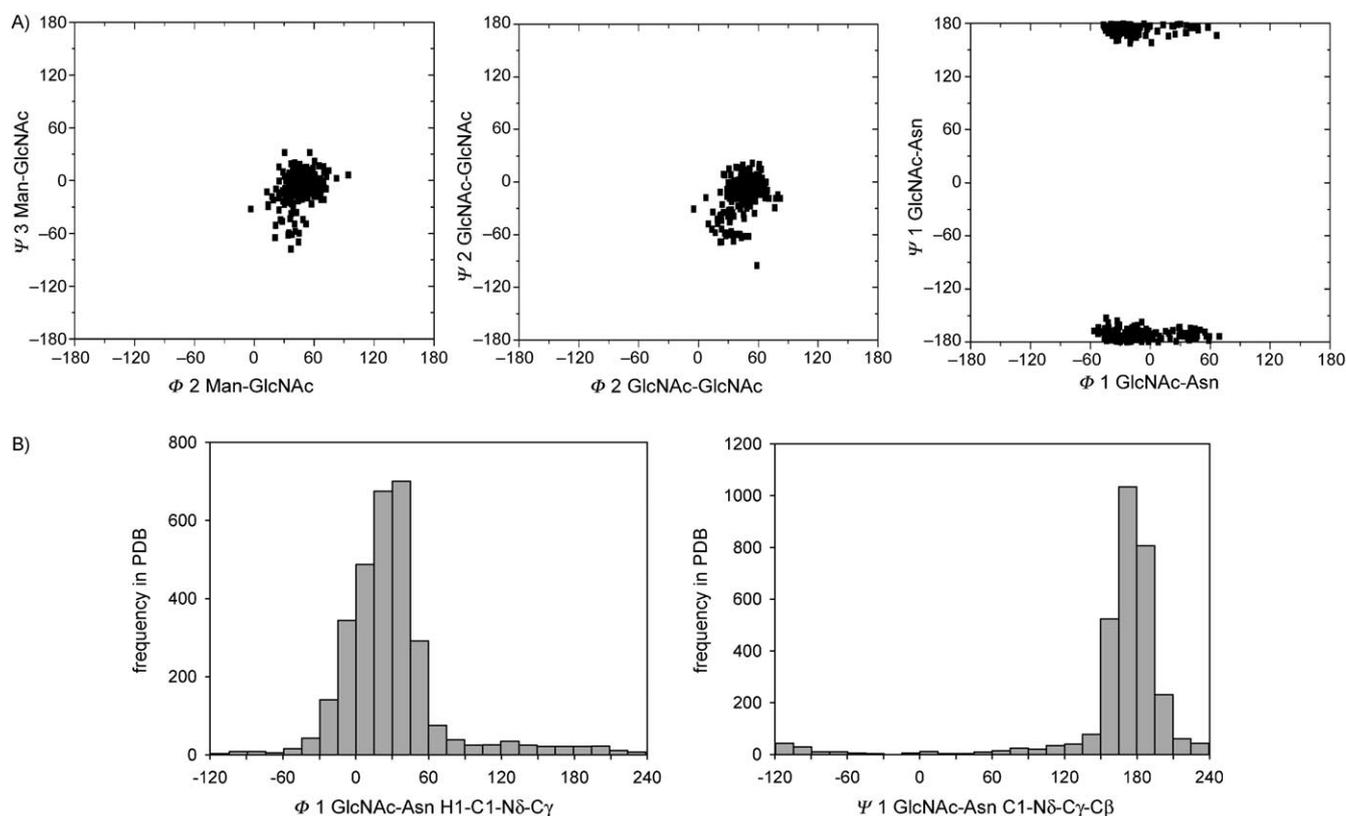


Figure 3. A) Φ versus Ψ representations for the different glycosidic torsions of **2**. From left to right, Man-GlcNAc, GlcNAc-GlcNAc, and GlcNAc-Asn linkages. B). Frequency distribution of GlcNAc-Asn bond torsion angles Φ and Ψ found in the PDB.^[35]

The observed perturbations in the chemical shifts of the peptide upon sugar addition (see Figure S1 in the Supporting Information) clearly proved the formation of specific complexes between hevein and **1** or **2**. The continuous variation of the chemical shifts when increasing quantities of ligand were added indicated that these changes can be used to determine the association constant of the equilibrium between the free and bound species. Since the observed proton chemical shift changes of the protein are proportional to the molar fraction of the present complex in solution, the association constants (K_a) were then determined by non-linear least-square fitting of the observed chemical shifts perturbations (Figure 4) versus different ligand/receptor molar ratios. Several amidic protein protons changed their chemical shifts after ligand addition in a noteworthy way. For this case, the signal belonging to the Trp 21 side-chain NH was monitored as a function of the added ligand concentration to determine the association constant values.

The titrations were performed at four different temperatures, 298, 303, 308, and 313 K, providing, in the case of the trisaccharide amino acid **1**, K_a values of around 12000, 10000, 8500, and 7000 M^{-1} , respectively (Table 1). These values were used to obtain the thermodynamic parameters through a van't Hoff approximation (Figure 4) by the plot of $\ln(K_a)$ versus $1/T$.

It should be recognized that the use of van't Hoff plots should be considered with caution, since there are several

approximations regarding the lack of heat capacity dependence with temperature, which have not been demonstrated for these systems, although NMR spectroscopy and microcalorimetry data have been shown to be fairly similar for chitobiose and chitotriose binding to hevein and related domains.^[13–18] In any case, the observed negative values of ΔH° and ΔS° are in agreement with an enthalpy-driven process. The K_a data analysis and the obtained thermodynamic parameters for the **1**/hevein and **2**/hevein complexes (Table 1) permitted us to conclude that the hevein affinity for these new ligands is much higher than for chitobiose and even higher than for chitotriose. This fact indicates that the presence of the mannose residue is indeed stabilizing. In contrast, the presence of the Asn moiety is not beneficial for the interaction, since the affinity for **2** is lower than for **1**. Nevertheless, the origin of this extra stabilization of **2** and especially **1** versus the $(\text{GlcNAc})_3$ seems to be entropic, since the interaction enthalpy is better for $(\text{GlcNAc})_3$. In fact, the change of the nonreducing terminal GlcNAc residue in chitotriose to Man precludes for **1** and **2** the existence of one of the two existing binding modes of chitotriose to hevein,^[13–16] (Figure 2), with the corresponding loss of enthalpic stabilization. Additionally the improvement of the affinity for compounds **1** and **2** relative to the $(\text{GlcNAc})_3$ trisaccharide, existing as a mixture of anomers α and β , could partially arise from the fixing of the β configuration at the reducing end in both compounds **1** and **2**.

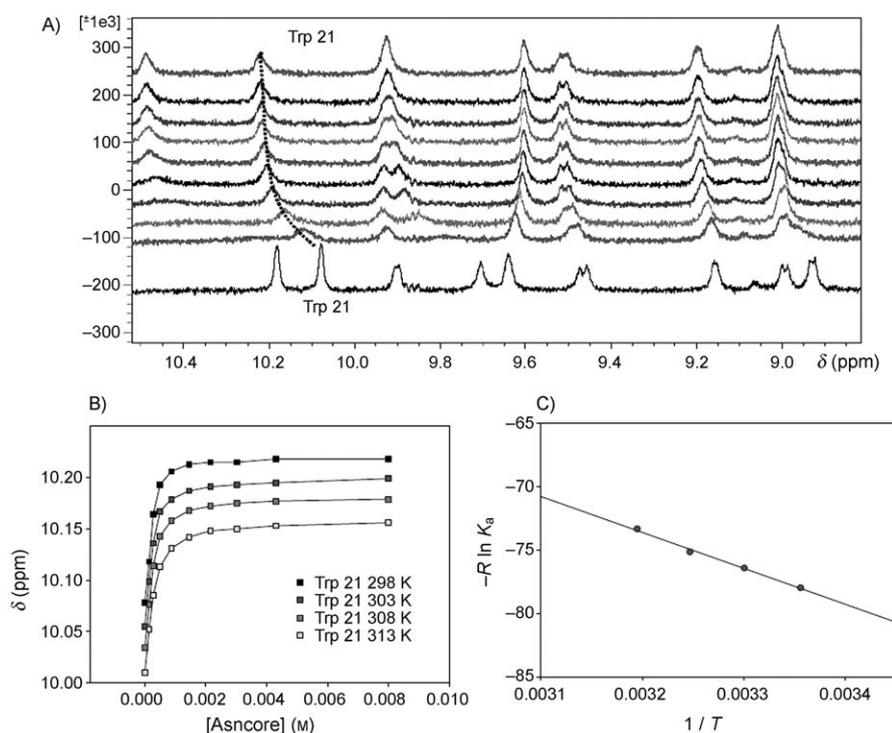


Figure 4. Titration experiments: A) Amide region of the ^1H NMR spectra at 298 K of hevein at increasing concentrations of **2**. The chemical shift perturbation of the Trp 21 side-chain NH signal is highlighted. B) Titration curves at different temperatures for the Trp 21 NH signal: chemical shift variation versus concentrations of **2**. C) van't Hoff representation for the hevein/**2** interaction.

Table 1. Binding affinities and thermodynamic parameters for the interaction of hevein with **1** and **2**, compared with those previously determined for chitooligosaccharides (GlcNAc)₂, methyl-β (GlcNAc)₂, and (GlcNAc)₃.^[14]

	K_a [M^{-1}]				Thermodynamic parameters	
	298 K	303 K	308 K	313 K	ΔH° [kJ mol^{-1}]	ΔS° [J mol^{-1}]
<i>N,N'</i> -diacetyl chitobiose	620	460	381	337	-31.3	-52.5
<i>N,N'</i> -diacetyl methyl β-chitobiose	1225	1069	882	647	-32.1	-53.5
<i>N,N,N'</i> -triacetyl chitotriose	11500	8700	6900	5700	-36.4	-45.1
glycosylamino acid 2	11863	9842	8437	6792	-28.3	-17.0
trisaccharide 1	14898	-	10111	8245	-30.4	-22.2

When the estimated ΔH° and ΔS° are compared with those previously described for the other ligands, it is evident that the data show an enthalpy/entropy compensation phenomenon, typical in association processes between proteins and carbohydrates.^[39] Better binding enthalpy is often accompanied by a larger entropic loss. The enthalpy driven processes are probably due to van der Waals interactions and hydrogen bonds in the ligand-receptor complex, as has been demonstrated in the case of hevein domains. The corresponding analysis was corroborated by determining the 3D structure of the complex in solution by NMR spectroscopic methods and modeling protocols.

NMR spectroscopic structure: The analysis of the NOESY experiments (100 and 500 ms mixing times) of the complex between hevein and glycosylamino acid **2** was performed

(500 and 900 MHz, 1:8 molar ratio). From the technical viewpoint, and within the framework of the FP6 project EU-ROCarbDB, new tools were introduced to the CCPN data-model^[40] and its derived software tools for the knowledge of carbohydrates down to the atomic level. In this study, the feasibility of the CCPN Analysis package^[40] was demonstrated for the computer-assisted book-keeping of assignments of these complex (glycopeptide and protein) molecular systems. Regarding the cross-peaks corresponding to intraprotein protons, 446 cross-peaks (Figure 5) could be unequivocally assigned, which were basically identical to those previously found for hevein when complexed to a variety of chitooligosaccharides.^[13–16] Thus, it could be safely assessed that the topology, folding and 3D shape of the lectin is preserved upon binding to glycosylamino acid **2**. Also, the perturbation of chemical shifts in the protein protons was particularly focused on the region between amino acids 19 and 31, thus exclusively corresponding to the chitooligosaccharide binding region (see Figure S1 in the Supporting Information).

The use of a very high field (900 MHz) permitted the detection of the presence of many intermolecular NOE peaks between the lectin and the ligand (Table 2). In fact, the intermolecular cross-peaks were observed between the sugar ring protons and the corresponding acetamide methyl groups of the GlcNAc residues with different protons of the hevein amino acids Trp 21, Trp 23, Tyr 30, and Ser 19. No cross-peaks between the Asn side chain of the glycosylamino acid and protons of the protein could be detected. These experimental data indeed assessed that the recognition of the glycosylamino acid **2** involves the same binding site as for chitotriose. More importantly, some of the intermolecular NOEs permitted the unequivocal location of the glycosylamino acid in the binding site of the lectin, in a +2, +1, -1 disposition, as shown in Figure 2B; cross-peaks involving H3 and H4 protons of the GlcNAc residues, and z3, h2, and d1 protons of Trp 23 and h2 of Trp 21 required some interpretation due to certain overlapping.

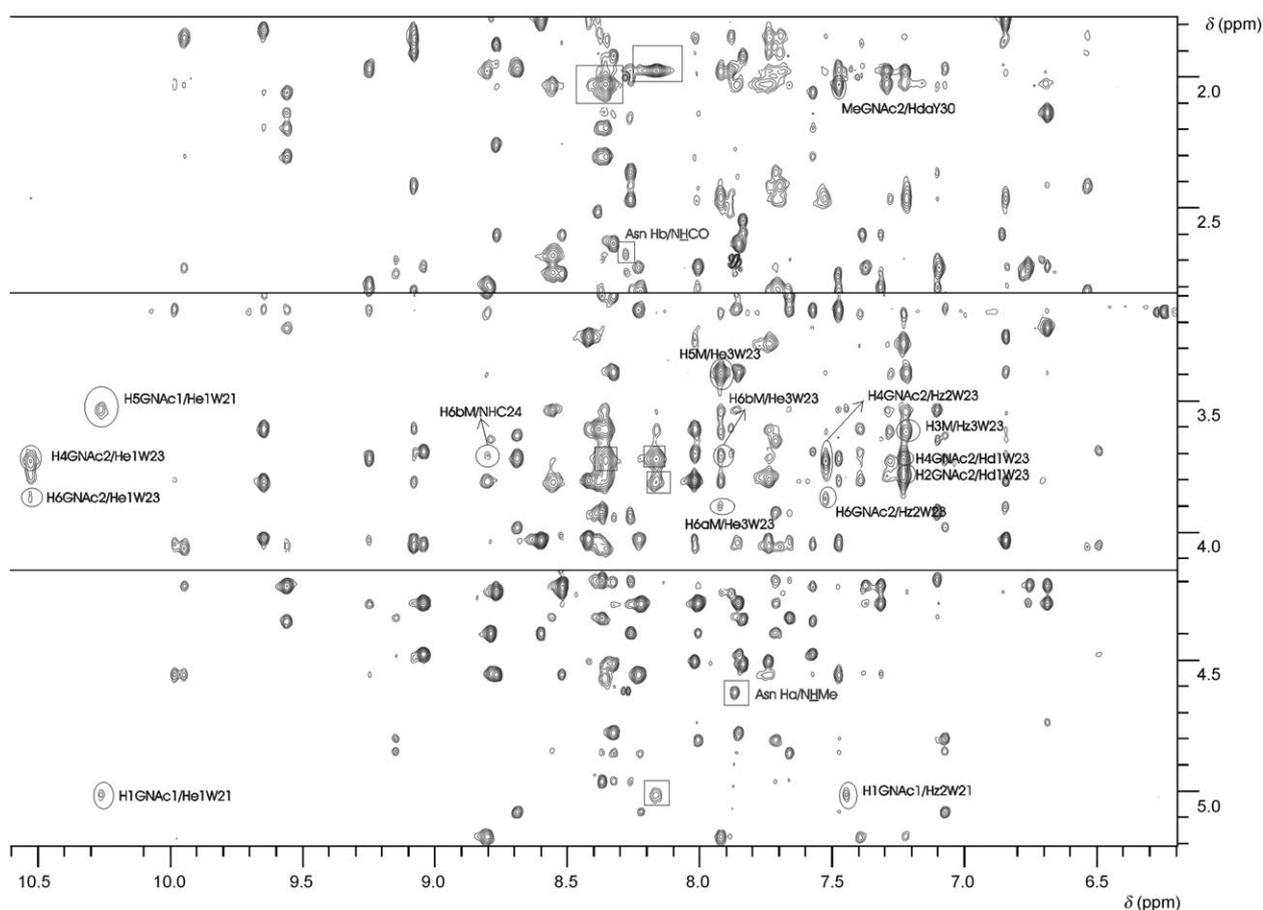


Figure 5. Amide region of the NOESY spectrum of hevein/2 (1:8) acquired at 900 MHz with 500 ms mixing time. Inside circles: Some selected intermolecular NOE cross-peaks are listed in Table 2. Inside squares: Intramolecular NOEs of **2**, two of which are positive and correspond to the Asn residue. M = Man, GNAc2 = GlcNAc2, and GNAc1 = GlcNAc1 (Scheme 3).

Table 2. Intermolecular NOEs between compound **2** and hevein (900 MHz NOESY, 100 and 500 ms mixing), supporting the +2, +1, -1 sites occupancy, and the corresponding distance in that binding mode.

2 proton	Hevein proton	Distance in final model [Å]
mannose	H3 z3 W23	3.8
	H5 z3 W23	3.1
	e3 W23	2.4
	H6a e3 W23	3.6
	H6b NHC24	3.0
	e3 W23	2.9
GlcNAc2	H2 d1 W23	1.9
	H4 e1 W23	3.5
	d W23	3.8
	z2 W23	3.8
	H6 z2 W23	3.1
	e1 W23	3.6
Me	da Y30	3.1
	NHS19	3.9
GlcNAc1	H1 z2 W21	3.4
	e W21	3.3
	H5 e1 W21	2.6

Thus, all assigned NOE interactions shown in Table 2 indeed indicate the recognition site +2, +1, -1 as the main one. Besides, other intermolecular NOEs (Me GlcNAc1/

HdY30 or Me GlcNAc2Hz3W21) would indicate minor contributions of a different disposition of the ligand in the binding site in which GlcNAc1 would be located at site +1 or GlcNAc2 on site -1.

From the molecular recognition point of view and the ligand mobility in the complex, it has to be emphasized that the intraresidual NOESY cross-peaks for the GlcNAc and Man residues displayed the same phase as the diagonal peaks, which indicated that this part of the molecule moves with a global motion rotational correlation time similar to that of the protein with relatively slow mobility. On the contrary, the NOESY cross-peaks for the amide protons of the Asn moiety showed the opposite phase to the diagonal, which indicated a faster effective correlation time, as a possible consequence of weaker interactions with the protein. This fact shows that the major interaction of the ligand and the lectin occurs in the oligosaccharide part with the Asn moiety being more exposed to the solvent.

The nature of the cross-peaks permitted the confirmation that the binding of the glycosylamino acid **2** to hevein occurs mainly in the binding mode with a +2, +1, -1 disposition (see Figure 6 and below in the molecular modeling section).

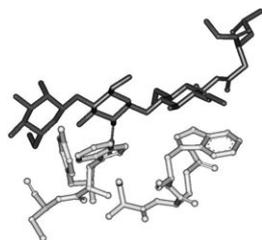


Figure 6. Binding site of hevein (amino acids Ser 19, Gln 20, Trp 21, Trp 23, Cys 24, Tyr 30) with compound **2** docked on site +2, +1, -1.

Molecular modeling: Finally, molecular-modeling calculations were also performed to independently confirm the NMR spectroscopic results, which point to the existence of one unique binding mode, and to also substantiate the possibility of adopting a modeling protocol to study this type of protein/carbohydrate complexes. First, docking calculations were performed with AutoDock.^[41] Additionally, the modeling approach should also permit the

assessment of a further extension of the polypeptide at the Asn moiety (within a glycoprotein) and of additional glycosylations at the terminal Man residue.

The two major clusters of structures obtained from the calculations for both ligands **1** and **2** satisfactorily fitted the binding orientation described in Figure 6 with different orientations for the Asn moiety.

The major complexes of **1** and **2** with hevein completely backed up the existence of only one binding orientation (+2, +1, -1). The glycosylamino acid **2** is located so that the reducing GlcNAc residue is placed at subsite +2, the middle GlcNAc at subsite +1, and the nonreducing Man ring at -1, leaving the Asn residue very exposed to the solvent, at subsite +3. The change of the chemical nature of the sugar and the variation of the stereochemistry at position C2 at the nonreducing end moiety (from GlcNAc to Man) precludes the possibility of establishing proper interactions between the terminal Man and subsite +1. Thus, of the two possibilities existing for the recognition of chitooligosaccharides, only one possibility of sugar-lectin interaction remains possible in the case of the core trisaccharide of *N*-glycoproteins.

Molecular dynamics: Finally, to assess the conformational stability of the proposed complex, the 3D models obtained from the docking calculations, in agreement with the NMR spectroscopic data, were used as input structures for molecular dynamics simulations,^[42] by using explicit solvent with the AMBER 9 program,^[43] with no experimental NMR spectroscopic restraints. The simulation time was beyond four ns. It was observed that the complex remained completely stable during the complete simulation, showing minor motion of the lateral side chains of the essential amino acids for the molecular recognition; Trp 21, Trp 23, Ser 19, and Tyr 30. Regarding the key intermolecular contacts between the ligand and the protein (i.e., the hydrogen bond between Ser 19 OH and the carbonyl group of the intermediate GlcNAc, the hydrogen bond between Tyr 30 OH and OH3 of the intermediate GlcNAc, and the stacking interaction between the acetamide methyl group and the center of Tyr 30), the corresponding distances were also monitored, showing a very good stability with time, within

the ranges expected for stabilizing interactions. In contrast, fluctuations with time were observed for the Asn moiety, as also deduced from the NMR spectroscopic data, by monitoring the distance between its terminal groups and the NH of Trp 23, also ruling out any kind of important interaction between the Asn fragment and the lectin binding site.

In addition, the glycosidic torsions for the ligand remained fairly stable during the simulation, close to those deduced in the free-state. All these data support the idea of preorganization of both hevein and its ligands for the molecular recognition process. There are no important changes in the global shape of the lectin or of ligands **1** and **2** between the free and bound states, only some restriction to motion due to the intermolecular protein-ligand contacts.

Finally, the obtained 3D structure of the complex was employed to predict the possibility of extensions at the Asn and Man moieties of the glycosylamino acid **2**. Inspection of Figure 7 permits the assessment that, in the complex with

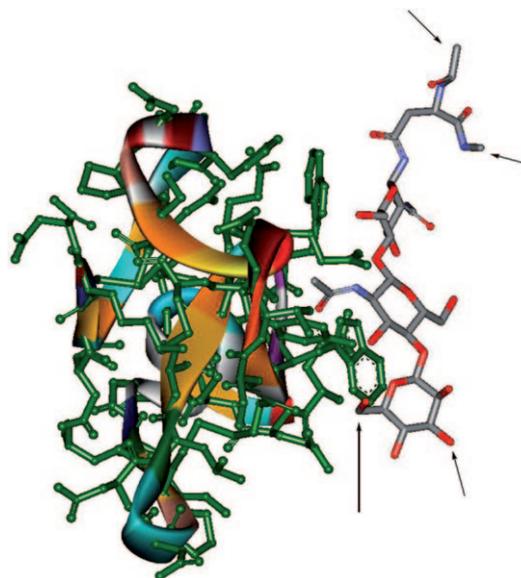


Figure 7. The 3D structure of the complex of hevein/**2**, showing the positions for both peptide elongation from the asparagine residue and the mannose branching at positions 3 and 6 of the terminal Man residue that would exist in *N*-glycoproteins.

hevein, both the Asn and Man residues adopt orientations that allow the extension of the polypeptide chain from the Asn moiety with no steric hindrance, as well as branching glycosylation at O-3 and O-6 of the terminal mannose. Thus, in principle, hevein could also recognize *N*-linked glycoproteins by using the same binding mode outlined above.

Conclusion

NMR spectroscopy and modeling calculations have permitted us to demonstrate that hevein, a model lectin related to the superantigen family, recognizes the trisaccharide core of

N-glycoproteins (**1**) and also an Asn-linked glycosylamino acid **2** with even higher affinity than (GlcNAc)₃. The binding mode is similar to those described for regular chitooligosaccharides. However, for the core trisaccharide, the mode of binding selects only one of the possibilities existing for chitooligosaccharides. The deduced 3D structure of the complex permits the extension of polypeptide chains from the Asn moiety, as well as branching glycosylation at O-3 and O-6 of the terminal mannose. Given the ubiquity of the Man(GlcNAc)₂ core in all mammalian *N*-glycoproteins, the basic recognition mode presented herein might be extended to a variety of systems with increased complexity and with biomedical importance. In any case, and due to the very important issue that latex allergies bring in to public health, the fact that the carbohydrate binding site of hevein could accept *N*-glycoproteins and, at the same time, is part of the immunoglobulin binding epitope,^[24–26] makes it interesting to explore the potential beneficial use of chitin or its oligosaccharide derivatives,^[26] mimics, and analogues in protective formulations against latex allergies (or detrimental behavior acting as local concentrators of the hevein antigen). It should be pointed that chitin itself is already present in diverse commercial skincare formulations.

Experimental Section

General methods: Solvents were dried according to standard methods. Optical rotations were measured on a Perkin–Elmer 241 polarimeter at 589 nm. NMR spectra were recorded on a Bruker Avance 360 instrument. Coupling constants are reported in Hz. ESI-TOF mass spectra were recorded on a Micromass LCT instrument coupled to an Agilent 1100 HPLC. Flash chromatography was performed on silica gel 60 (230–400 mesh, Merck Darmstadt). The reactions were monitored by TLC on coated aluminum plates (silica gel 60 GF₂₅₄, Merck Darmstadt). Spots were detected by UV light or by charring with a 1:1 mixture of 2N H₂SO₄ and 0.2% resorcinol monomethyl ether in ethanol.

Synthesis

***O*-(4,6-*O*-Benzylidene-β-D-mannopyranosyl)(1→4)-*O*-(2-acetamido-3,6-di-*O*-benzyl-2-deoxy-β-D-glucopyranosyl)(1→4)-2-acetamido-*N*¹-acetyl-3,6-di-*O*-benzyl-2-deoxy-β-D-glucopyranosylamine (**4**):** Trisaccharide **3** (301 mg, 0.16 mmol) was dissolved in freshly distilled methanol (12.5 mL). DIPEA (420 μL) and 1,3-propanedithiol (1.25 mL, 12.4 mmol) were added and the solution was stirred at room temperature. After 2 h (TLC: hexane/acetone 1.5:1), the solvent was removed under reduced pressure and the residue was dried under high vacuum. The remaining mixture was treated with pyridine/acetic anhydride (30 mL, 2:1) at room temperature. After complete acetylation (TLC: CH₂Cl₂/methanol 15:1), the solvent was removed in vacuo and the residue was co-distilled with toluene (3×) and subsequently dried under high vacuum. The residue was dissolved in mixture of *n*-butanol (28 mL) and ethylenediamine (7 mL, 0.10 mol) and the solution was stirred at 90 °C. After 17 h (TLC: CH₂Cl₂/methanol 15:1), the volatiles were removed under reduced pressure and the residue was co-distilled with toluene (3×) and dried under high vacuum. The remaining solid was treated with pyridine/acetic anhydride (30 mL, 2:1) for 1 h (TLC: CH₂Cl₂/methanol 15:1). After complete acetylation, the solvent was removed in vacuo and the residue was co-distilled with toluene (3×) and dried under high vacuum. The residue was dissolved in aqueous methylamine (20 mL, 40%) and stirred for 1 h at room temperature. After disappearance of the starting material (TLC: CH₂Cl₂/methanol 15:1), the volatiles were removed in vacuo and the remainder was co-distilled with toluene (3×) and dried under high vacuum. The crude product was purified by flash chromatography (CH₂Cl₂/methanol 20:1). Yield: 221 mg (84.2%); *R*_f (glycosylamine) = 0.26 (hexane/acetone 1.5:1), *R*_f (acetamide) = 0.56 (CH₂Cl₂/methanol 15:1), *R*_f (diamine) = 0.25 (CH₂Cl₂/methanol 15:1), *R*_f (peracetate) = 0.31 (CH₂Cl₂/methanol 15:1), *R*_f (**4**) = 0.27 (CH₂Cl₂/methanol 15:1); [α]_D²⁵ = -29.4 (*c* = 0.5 in CH₂Cl₂); ESI-MS (100% acetonitrile): *m/z*: calcd for C₅₉H₆₉N₅O₁₆: 1075.5; found: 1098.3 [M+Na]⁺.

***O*-(4,6-*O*-Benzylidene-β-D-mannopyranosyl)(1→4)-*O*-(2-acetamido-2-deoxy-β-D-glucopyranosyl)(1→4)-2-acetamido-*N*¹-acetyl-2-deoxy-β-D-glucopyranosylamine (**1**):** Trisaccharide **4** (125 mg, 116.2 μmol) was dissolved in a mixture of freshly distilled methanol (16 mL) and acetic acid (4 mL). Palladium(II) oxide hydrate (120 mg) was added and the suspension was stirred under a hydrogen atmosphere for 3 d. After disappearance of **4** (TLC: isopropanol/1 M ammonium acetate 4:1), the solution was diluted with methanol and the catalyst was removed by filtration. After evaporation of the solvents, the crude product was purified by gel-filtration chromatography (Pharmacia Hi Load Superdex 30 prep grade (600×16 mm), 0.1 M ammonium hydrogen carbonate in water, flow rate = 0.75 mL min⁻¹). Yield: 43.8 mg (60.0%), *R*_f = 0.33 (isopropanol/1 M ammonium acetate 4:1); [α]_D²⁵ = -17.9 (*c* = 1.43 in water); ESI-MS (water): *m/z*: calcd for C₂₄H₄₁N₅O₁₆: 627.3; found: 650.8 [M+Na]⁺; ¹H NMR (360 MHz, D₂O + [D₆]DMSO as an internal standard): δ = 4.84 (d, *J*_{1,2} = 9.5 Hz, 1H; H-1¹), 4.57 (d, *J*_{1,2} < 1 Hz, 1H; H-1²), 4.41 (d, *J*_{1,2} = 7.2 Hz, 1H; H-1³), 3.86 (dd, *J*_{1,2} < 1, *J*_{2,3} = 3.0 Hz, 1H; H-2³), 3.76–3.33 (m, 16H; H-6a³, H-6a², H-2¹, H-6a¹, H-2², H-6b², H-3², H-3¹, H-4², H-6b³, H-6b¹, H-3³, H-4¹, H-5², H-4³, H-5¹), 3.22 (m, 1H; H-5³), 1.86 (s, 3H; NAc), 1.80 ppm (s, 6H; NAc); ¹³C NMR (90 MHz, D₂O + [D₆]DMSO as an internal standard): δ = 176.4, 176.2, 176.0 (C=O NAc), 102.9 (C-1³), 101.6 (C-1²), 80.3 (C-4¹), 80.2 (C-4²), 79.9 (C-1¹), 77.9 (C-5³), 77.8 (C-5¹), 76.2 (C-5²), 74.3 (C-3³), 74.3 (C-3¹), 73.5 (C-3²), 72.1 (C-2³), 68.2 (C-4³), 62.5 (C-6³), 61.6 (C-6²), 61.4 (C-6¹), 56.6 (C-2²), 55.3 (C-2¹), 23.7, 23.6, 23.5 ppm (NAc).

***O*-(4,6-*O*-Benzylidene-β-D-mannopyranosyl)(1→4)-*O*-(2-acetamido-3,6-di-*O*-benzyl-2-deoxy-β-D-glucopyranosyl)(1→4)-2-acetamido-3,6-di-*O*-benzyl-2-deoxy-β-D-glucopyranosylamide (**7**):** Trisaccharide **3** (200 mg, 0.16 mmol) was dissolved in *n*-butanol (26 mL). Ethylenediamine (6.5 mL, 97.2 mmol) was added and the solution was stirred at 80 °C. After 20 h (TLC: CH₂Cl₂/methanol 15:1), the volatiles were removed under reduced pressure. The residue was co-distilled with toluene (3×) and dried under high vacuum. Subsequently, the solid was treated with pyridine/acetic anhydride (20 mL, 2:1) for 3 h at room temperature. After complete acetylation (TLC: CH₂Cl₂/methanol 15:1), the reagents were removed in vacuo followed by co-distillation with toluene (3×). The dried residue was dissolved in methanol (10 mL) and aqueous methylamine (13 mL, 40%) was added. After disappearance of the starting material (13 h, TLC: CH₂Cl₂/methanol 15:1), the mixture was concentrated in vacuo and the residue co-distilled with toluene (3×) and dried under high vacuum. The crude product was purified by flash chromatography (CH₂Cl₂/methanol 20:1). Yield: 127 mg (74.1%); *R*_f (amine) = 0.40 (CH₂Cl₂/methanol 15:1), *R*_f (peracetate) = 0.46 (CH₂Cl₂/methanol 15:1), *R*_f (**7**) = 0.18 (CH₂Cl₂/methanol 15:1); [α]_D²⁵ = -45.4 (*c* = 0.5 in CH₂Cl₂); ESI-MS (50% acetonitrile): *m/z*: calcd for C₅₇H₆₅N₅O₁₅: 1059.5; found: 1060.5 [M+H]⁺, 1082.6 [M+Na]⁺, 1098.5 [M+K]⁺; ¹H NMR (360 MHz, [D₆]DMSO): δ = 8.03 (d, *J*_{NH,2} = 8.9 Hz, 2H; NH), 7.49–7.13 (m, 25H; Ar), 5.51 (s, 1H; Ph-CH), 5.00–4.89 (m, 3H; CH₂O, CH₂O, OH-3³), 4.86 (d, *J*_{OH,2} = 4.3 Hz, 1H; OH-2³), 4.66–4.41 (m, 8H; CH₂O, H-1², CH₂O, H-1¹, CH₂O, H-1³, CH₂O, CH₂O), 4.38 (d, *J*_{gem} = 12.5 Hz, 1H; CH₂O), 3.96–3.42 (m, 16H; H-6a³, H-4¹, H-4², H-6a¹, H-2¹, H-2³, H-6b¹, H-4³, H-6a², H-3², H-2², H-5¹, H-3¹, H-6b², H-6b³, H-3³), 3.27–3.20 (m, 1H; H-5²), 3.07–2.99 (m, 1H; H-5³), 1.81 (s, 3H; OAc), 1.79 ppm (s, 3H; OAc); ¹³C NMR (90 MHz, [D₆]DMSO): δ = 169.3, 169.2 (C=O NAc), 139.3, 139.2, 138.5, 138.3, 137.9, 128.7, 128.2, 127.7, 127.2, 127.0, 126.3 (C Ar), 101.0 (Ph-CH), 100.6 (C-1³), 100.0 (C-1²), 88.0 (C-1¹), 80.5 (C-3¹), 79.9 (C-3²), 78.3 (C-4³), 77.0 (C-4²), 76.3 (C-5¹), 74.9 (C-4¹), 74.4 (C-5²), 73.6 (CH₂O), 73.4 (CH₂O), 72.2 (CH₂O), 71.9 (CH₂O), 70.9 (C-2³), 70.0 (C-3³), 68.5 (C-6²), 68.1 (C-6¹), 67.9 (C-6³), 66.8 (C-5³), 55.3 (C-2²), 53.6 (C-2¹), 22.9, 22.8 ppm (NAc).

***N*²-*tert*-Butyloxycarbonyl-L-aspartic acid-1-methylamide (**6**):** Boc aspartic acid benzyl ester **5** (200 mg, 0.69 mmol) was dissolved in freshly distilled methanol (4 mL). Aqueous methylamine (4 mL, 40%) was added and

the solution was stirred at room temperature. After 2 h, the solvents were removed by lyophilization and the residue was purified by ion-exchange chromatography (DOWEX 50WX8-100, column: 20×500 mm, eluent: water). Yield: 141.7 mg (84.1%); $R_f=0.35$ (CH₂Cl₂/methanol 10:1 with 0.1% acetic acid); $[\alpha]_D^{25} = +6.3$ ($c=0.9$ in water); IR (KBr): $\tilde{\nu}=1732, 1695, 1643$ cm⁻¹ (C=O); ESI-MS (100% water): m/z : calcd for C₁₀H₁₈N₂O₅: 246.1; found: 493.2 [2M+H]⁺, 515.2 [2M+Na]⁺, 531.2 [2M+K]⁺; ¹H NMR (360 MHz, [D₆]DMSO): $\delta=12.22$ (s, 1H; CO₂H), 7.72 (d, $J_{\text{NH,Me}}=4.6$ Hz, 1H; NHMe), 6.99 (d, $J_{\text{NH},\alpha}=8.2$ Hz, 1H; NHBoc), 4.24–4.16 (m, 1H; α CH-Asn), 2.62–2.56 (m, 1H; β CHa-Asn), 2.55 (d, $J_{\text{NH,Me}}=4.6$ Hz, 1H; Me), (dd, $J_{\alpha,\beta}=8.6$, $J_{\text{gem}}=16.1$ Hz, 1H; β CHb-Asn), 1.36 ppm (s, 9H; *t*Bu); ¹³C NMR (90 MHz, [D₆]DMSO): $\delta=171.9, 171.3, 155.1$ (C=O), 78.2 (qC *t*Bu), 51.0 (C- α Asn), 36.5 (C- β Asn), 28.2 (*t*Bu), 25.8 ppm (Me).

N⁴-[O-(4,6-O-Benzylidene- β -D-mannopyranosyl)(1→4)-O-(2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl)(1→4)-(2-acetamido-3,6-di-O-benzyl-2-deoxy- β -D-glucopyranosyl)]-N²-tert-butylloxycarbonyl-L-asparagine methylamide (8): Trisaccharide **7** (25 mg, 23.6 μ mol) was dissolved in freshly distilled methanol (2 mL) under an argon atmosphere. Triethylamine (131 μ L, 0.94 mmol) and 1,3-propanedithiol (473 μ L, 4.7 mmol) were added and the solution was stirred at room temperature. After 4 h (TLC: CH₂Cl₂/methanol 15:1), the mixture containing the glycosylamine was concentrated under reduced pressure and dried under high vacuum. Boc aspartic acid methylamide (**6**) (116 mg, 0.47 mmol) and DEPBT (141 mg, 0.47 mmol) were dissolved in freshly distilled CH₂Cl₂ (1 mL) and DIPEA (202 μ L, 1.2 mmol) was added. After shaking for 10 min, this solution was added to the freshly prepared glycosylamine under an argon atmosphere and the reaction mixture was stirred for 40 min. After complete reaction (TLC: CH₂Cl₂/methanol 10:1) the solvent was removed under reduced pressure and the residue was dried under high vacuum. The remaining mixture was dissolved in methanol (2.5 mL) and treated with aqueous methylamine (2.5 mL, 40%). After disappearance of the faster migrating spots (TLC: CH₂Cl₂/methanol 10:1), the mixture was lyophilized. The crude product was purified by flash chromatography (CH₂Cl₂/methanol 20:1), followed by HPLC (Agilent C8 XBD 15×4.65 mm, gradient: 50–65% acetonitrile/water (0.1% formic acid), flow rate=1 mL min⁻¹). Yield: 8.2 mg (27.5%); R_f (glycosylamine)=0.30 (CH₂Cl₂/methanol 10:1), R_f (**8**)=0.36 (CH₂Cl₂/methanol 10:1); $[\alpha]_D^{25} = -17.7$ ($c=0.3$ in methanol); ESI-MS (50% acetonitrile): m/z : calcd for C₆₇H₈₃N₅O₁₉: 1261.6; found: 1262.5 [M+H]⁺, 1284.6 [M+Na]⁺; ¹H NMR (360 MHz, [D₆]DMSO): $\delta=8.29$ (d, $J_{\text{NH,1}}=8.7$ Hz, 1H; NH), 8.01 (d, $J_{\text{NH,2}}=8.4$ Hz, 1H; NHAc), 7.85 (d, $J_{\text{NH,2}}=8.9$ Hz, 1H; NHAc'), 7.63 (d, $J_{\text{NH,Me}}=4.21$ Hz, 1H; NHMe), 7.44–7.13 (m, 15H; Ar), 6.60 (d, $J_{\text{NH},\alpha}=8.4$ Hz, 1H; NHBoc), 5.50 (s, 1H; Ph-CH), 4.99–4.88 (m, 4H; CH₂O, H-1', CH₂O, OH-3'), 4.81 (d, $J_{\text{OH,2}}=4.5$ Hz, 1H; OH-2'), 4.68–4.36 (m, 8H; H-1³, H-1³, CH₂O, CH₂O, CH₂O, CH₂O, CH₂O, CH₂O), 4.23–4.14 (m, 1H; α CH-Asn), 3.95–3.85 (m, 2H; H-6a³, H-4¹), 3.83–3.42 (m, 13H; H-4², H-2¹, H-6a¹, H-4³, H-2³, H-3², H-6b¹, H-3¹, H-2², H-6a², H-6b³, H-6b², H-3³), 3.42–3.17 (m, 2H; H-5¹, H-5²), 3.08–2.99 (m, 1H; H-5³), 2.54 (d, $J_{\text{NH,Me}}=4.2$ Hz, 3H; Me), 2.46–2.28 (m, 2H; β CHa,b-Asn), 1.82 (s, 3H; NHAc), 1.76 (s, 3H; NHAc), 1.35 ppm (s, 9H; *t*Bu). ¹³C NMR (90 MHz, [D₆]DMSO): $\delta=171.6, 169.9, 169.4, 165.8, 165.1$ (C=O NHAc, Asn, Boc), 139.3, 138.5, 138.4, 137.9, 128.7, 128.2, 127.9, 127.8, 127.6, 127.3, 127.2, 127.1, 127.0, 126.9, 126.3 (C Ar), 100.8 (Ph-CH), 100.4 (C-1³), 99.6 (C-1²), 81.4 (C-3¹), 79.6 (C-3²), 78.2 (qC *t*Bu), 78.1 (C-1^{1 β} , $^1J_{\text{C-1,H-1}}=156.6$ Hz), 78.1 (C-4³), 76.8 (C-4²), 75.9 (C-5¹), 74.4 (C-4¹), 74.2 (C-5²), 73.5 (CH₂O), 73.3 (CH₂O), 72.2 (CH₂O), 71.8 (CH₂O), 71.0 (C-2³), 70.0 (C-3³), 68.6 (C-6²), 67.9 (C-6¹), 66.8 (C-6³), 66.8 (C-5³), 55.4 (C-2²), 53.3 (C-2¹), 50.9 (C- α Asn), 37.4 (C- β Asn), 28.1 (CH₃ *t*Bu), 25.7 (Me), 22.9, 22.8 ppm (NHAc).

N⁴-[O- β -D-Mannopyranosyl(1→4)-O-(2-acetamido-2-deoxy- β -D-glucopyranosyl)(1→4)-(2-acetamido-2-deoxy- β -D-glucopyranosyl)]-N²-acetyl-L-asparagine methylamide (2): Palladium(II) oxide hydrate (23 mg) was suspended in freshly distilled methanol (0.5 mL) containing acetic acid (50 μ L) and stirred under a hydrogen atmosphere. After 3.5 h, a solution of glycosylated amino acid **8** (10.6 mg, 8.4 μ mol) in freshly distilled methanol (1.2 mL) containing acetic acid (120 μ L) was added and the suspension was stirred for 2 d under a hydrogen atmosphere. After disappearance of **8** (TLC: isopropanol/1 M ammonium acetate 4:1), the catalyst was removed by centrifugation and washed with methanol and water (3×).

The filtrate was lyophilized and used in the next step without further purification. $R_f=0.60$ (isopropanol/1 M ammonium acetate 4:1); ESI-MS (100% water): m/z : calcd: 813.4; found: 814.3 [M+H]⁺, 836.3 [M+Na]⁺, 1649.8 [2M+Na]⁺. The lyophilized solid was dissolved in trifluoroacetic acid (0.3 mL) and kept for 5 min at room temperature. The trifluoroacetic acid was removed under reduced pressure followed by high vacuum. The remainder was dissolved in aqueous acetic acid (10%) and lyophilized. R_f (amine)=0.40 (isopropanol/1 M ammonium acetate 2:1); ESI-MS (100% water): m/z : calcd: 713.3; found: 736.4 [M+Na]⁺.

The amine was dissolved in a mixture of freshly distilled methanol (2 mL) and water (0.3 mL). Acetic anhydride (0.2 mL) was added and the reaction mixture was stirred for 90 min. After complete conversion of the starting material (TLC: isopropanol/1 M ammonium acetate 2:1), the mixture was concentrated and the residue was dried under high vacuum. The crude product was purified by gel-filtration chromatography (Bio-Gel P-4 fine (750×15 mm), eluent: water, flow rate=1.5 mL min⁻¹). Yield: 5.1 mg (80.3%); R_f (**2**)=0.56 (isopropanol/1 M ammonium acetate 2:1); $[\alpha]_D^{25} = -11.2$ ($c=0.5$ in water); ESI-MS (100% water): m/z : calcd for C₂₉H₄₉N₅O₁₈: 755.3; found: 778.6 [M+Na]⁺; ¹H NMR (360 MHz, D₂O): $\delta=4.89$ (d, $J_{1,2}=9.7$ Hz, 1H; H-1¹), 4.61 (d, $J_{1,2}<1$ Hz, 1H; H-1³), 4.50 (dd, $J_{\alpha,\beta}=6.6$ Hz, 1H; α CH-Asn), 4.45 (d, $J_{1,2}=7.7$ Hz, 1H; H-1²), 3.90 (dd, $J_{1,2}<1$, $J_{2,3}=2.7$ Hz, 1H; H-2³), 3.80–3.71 (m, 2H; H-6a³, H-6a²), 3.70–3.54 (m, 8H; H-2¹, H-6a¹, H-2², H-3², H-6b², H-3¹, H-4², H-6b³), 3.54–3.37 (m, 6H; H-3³, H-4¹, H-6b¹, H-5², H-4³, H-5¹), 3.30–3.22 (m, 1H; H-5³), 2.67–2.51 (m, 5H; β CHa,b-Asn, Me), 1.90 (s, 3H; NHAc), 1.87 (s, 3H; NHAc), 1.85 ppm (s, 3H; NAc); ¹³C NMR (90 MHz, D₂O): m/z : $\delta=174.3, 174.1, 173.7, 172.3, 172.2$ (C=O NHAc, Asn), 100.8 (C-1³), 99.7 (C-1²), 78.2 (C-4²), 78.2 (C-4¹), 77.8 (C-1¹), 76.0 (C-5³), 75.7 (C-5¹), 74.2 (C-5²), 72.4 (C-3¹), 72.3 (C-3³), 71.5 (C-3²), 70.1 (C-2³), 66.2 (C-4³), 60.5 (C-6³), 59.7 (C-6²), 59.5 (C-6¹), 54.6 (C-2²), 53.2 (C-2¹), 49.9 (C- α Asn), 36.5 (C- β Asn), 25.5 (Me), 21.7, 21.4 ppm (NHAc).

Conformational analysis

NMR spectroscopy: The corresponding spectra for structure determination of the complex were recorded at 800 MHz in a Bruker Avance spectrometer. The samples for free and bound Hev32S19D (0.5 mM) were prepared in a buffer (90% H₂O/10% D₂O, 100 mM NaCl, 20 mM NaH₂PO₄, pH 5.6). TOCSY^[44] (50 and 70 ms of mixing time) experiments were performed by using standard sequences at 298 K, by using the Watergate module for water suppression. NOESY^[45] experiments were acquired with 200 and 300 ms of mixing times at 298 K, by using the Watergate module for water suppression.

Titration experiments: Titration experiments were performed by recording a series of 1D ¹H NMR spectra, in a Bruker Avance 500 MHz spectrometer, for different mixtures of hevein with the glycosylamino acid **2**, by following the procedure previously described.^[13–15] Firstly, the ¹H NMR spectra of two samples were recorded: for one 0.5 mL aliquot of a 5 mL solution of hevein, as a zero-point of the titration, and for a 0.5 mL aliquot of a 5 mL solution of a mixture of hevein (0.28 mM) and the corresponding ligand (8 mM), as final points of the titration, corresponding to highest ligand/peptide ratio (ca. 29:1). To build up the titration curve, small aliquots of the highest ligand/peptide ratio sample were added to the ligand-free peptide sample in a systematic way, as previously described.^[13–15] For each titration point, with different concentrations of carbohydrate, but maintaining a constant concentration of hevein, the ¹H NMR spectra were acquired at four different temperatures (298, 303, 308, and 313 K). These data allowed the qualitative estimation of the thermodynamic parameters (ΔS^0 and ΔH^0) of the interaction of hevein with both oligosaccharides, by using van't Hoff plots. It should be recognized that the use of van't Hoff plots should be considered with caution, since there are several approximations regarding the lack of heat capacity dependence with temperature that have not been demonstrated for these systems.

Structure determination: The 3D structure of free and ligand-bound hevein has already been determined.^[13–15] In any case, a complete NOESY cross-peak assignment was again performed to deduce the intermolecular ligand–hevein cross-peaks. In a first step, the spin systems of all amino acids and sugar residues that constitute the protein and the gly-

cosylamino acid **2** were assigned through the CcpNmr Analysis program.^[40] Subsequently, the cross-peak volumes were determined by manual peak integration. The CYANA program (version 2.1)^[46] was used to calculate the structure, by following the standard protocol through seven iterative cycles, by starting with 100 randomized conformers. The 20 best conformers, with the lowest final CYANA target function values, were retained for analysis and used as starting geometries for the next cycle. The obtained folding was basically identical to that previously reported.

In any case, the best CYANA structure was minimized in a box of explicit water molecules, by using the conjugated gradient method, with the AMBER 9 program.^[43] The free peptide was immersed in a TIP3P water box (3000 molecules), with a thickness of 12 Å. The restrained energy minimization process was carried out as follows: Initially, to eliminate the bad contacts between the water molecules and the polypeptide, a 500 steps minimization was performed only to the water molecules, keeping the position of the peptide atoms fixed, and by using a force constant of 100 kcal mol⁻¹ and constant volume. A subsequent minimization was then carried out for which the peptide was relaxed with the NOE-based experimental restrictions and the water molecules were kept fixed. Finally, the restrained energy minimization was performed by taking into account both the solvent and the peptide, by using 3000 steps with the force field of Cornell et al.^[47]

Molecular modeling

The glycosylamino acid **2:** The low-energy conformers of the glycosylamino acid **2** were calculated by using the MM3* force field, in MAESTRO. The ϕ/ψ torsion angles were defined as Man H-1'-Man C-1'-O-GlcNAc C-4' and GlcNAc H-4'-GlcNAc C-4'-O-Man C-1' and GlcNAc H-1'-GlcNAc C-1'-O-GlcNAc C-4' and GlcNAc H-4'-GlcNAc C-4'-O-GlcNAc C-1', respectively, for the terminal nonreducing and the intermediate glycosidic linkages, and GlcNAc H-1-GlcNAc C-1-N-CH₂ and GlcNAc C-1-N-CH₂-CH for the glycosylamino acid linkage. Potential-energy maps were generated in a systematic manner as previously described.^[34] In all the molecular mechanics and dynamics calculations, the GB/SA solvation model for water was used.

Also, molecular dynamics simulations were performed by using the MM3* force field. A temperature of simulation of 300 K was employed with a time step of 1.5 fs and an equilibration time of 100 ps. The total simulation time was 5 ns.

For the C5-C6 torsion of the GlcNAc and Man moieties, only the *gt* geometry (ω , defined as C4-C5-C6-O6, ca. 180°) was considered.

Autodock: The most stable conformers of the glycosylamino acid, in agreement with the *exo*-anomeric effect for both glycosidic linkages (as also observed by NMR spectroscopy) was docked into the deduced NMR spectroscopic structure of hevein for the complex. AutoDock 4.0 simulations with the multiple Lamarckian Genetic Algorithm were performed. Only local searches were performed centered around the known GlcNAc binding site.^[11,12] Grids of probe atom interaction energies and electrostatic potential were generated by the AutoGrid program present in AutoDock 4.0.^[41] A grid spacing of 0.275 Å was used. 100 docking runs were performed by using a population of 200 individuals and an energy evaluation number of 3×10^6 .

Molecular dynamics: Finally, 4 ns MD were run of the complex deduced by Audock employing the protocol described above for the NMR spectroscopic structure determination but without the experimental restraints. The complex structure was found to be completely stable during the MD run, as described in the text.

Acknowledgements

We thank Prof. Dr. Hans Kamerling for his interest and for helpful discussions. We thank the EU for financial support (MRTN CT2006-035546) and the MICINN (Spain) for grants CTQ2006-10874-CO2-01 and CTQ2009-8536. We also thank the EU for RIDS Contract number 011952 (for EUROCarbDB: <http://www.eurocarbdb.org>) and EU-NMR (<http://www.eu-nmr.eu>) for the EC FP6 Research Infrastructure action—

Integrated Infrastructure Initiative—Contract # RII3-026145. C.U. thanks the Deutsche Forschungsgemeinschaft and Fonds der Deutschen Chemischen Industrie for financial support.

- [1] C. Bewley, *Protein–Carbohydrate Interactions and Infectious Diseases*, RSC, London, **2005**.
- [2] H.-J. Gabius, *The Sugar Code: Fundamentals of Glycosciences*, Wiley-VCH, Weinheim, **2009**.
- [3] H.-J. Gabius, H.-C. Siebert, S. André, J. Jiménez-Barbero, H. Rüdiger, *ChemBioChem* **2004**, *5*, 740.
- [4] H. L. N. Sharon, in *Glycosciences: Status and Perspectives* (Eds.: H.-J. Gabius), Chapman & Hall, New York, **1997**, p. 133.
- [5] O. Seitz, *ChemBioChem* **2000**, *1*, 214.
- [6] M. R. Wormald, A. J. Petrescu, Y. L. Pao, A. Gilthero, T. Elliott, R. A. Dwek, *Chem. Rev.* **2002**, *102*, 371.
- [7] H. Alenius, N. Kalkkinen, T. Reunala, K. Turjanmaa, T. Palosuo, *J. Immunol.* **1996**, *156*, 1618.
- [8] B. L. Archer, *Biochem. J.* **1960**, *75*, 236.
- [9] J. J. Beintema, *FEBS Lett.* **1994**, *350*, 159.
- [10] A. Diaz-Perales, C. Collada, C. Blanco, R. Sanchez-Monge, T. Carrillo, C. Aragoncillo, G. Salcedo, *J. Allergy Clin. Immunol.* **1998**, *102*, 127.
- [11] C. A. Reyes-López, A. Hernández-Santoyo, M. Pedraza-Escalona, G. Mendoza, A. Hernández-Arana, A. Rodríguez-Romero, *Biochem. Biophys. Res. Commun.* **2004**, *314*, 123.
- [12] A. Rodríguez-Romero, K. G. Ravichandran, M. Soriano-García, *FEBS Lett.* **1991**, *291*, 307.
- [13] J. L. Asensio, F. J. Cañada, M. Bruix, C. Gonzalez, N. Khiar, A. Rodríguez-Romero, J. Jimenez-Barbero, *Glycobiology* **1998**, *8*, 569.
- [14] J. L. Asensio, F. J. Cañada, M. Bruix, A. Rodríguez-Romero, J. Jimenez-Barbero, *Eur. J. Biochem.* **1995**, *230*, 621.
- [15] J. L. Asensio, F. J. Cañada, H.-C. Siebert, J. Laynez, A. Poveda, P. M. Nieto, U. M. Soedjanaamadja, H.-J. Gabius, J. Jiménez-Barbero, *Chem. Biol.* **2000**, *7*, 529.
- [16] N. H. Andersen, B. Cao, A. Rodríguez-Romero, B. Arreguin, *Biochemistry* **1993**, *32*, 1407.
- [17] E. García-Hernández, R. A. Zubillaga, A. Rojo-Domínguez, A. Rodríguez-Romero, A. Hernández-Arana, *Proteins* **1997**, *29*, 467.
- [18] A. Hernández-Arana, A. Rojo-Domínguez, M. Soriano-García, A. Rodríguez-Romero, *Eur. J. Biochem.* **1995**, *228*, 649.
- [19] M. C. Fernandez-Alonso, F. J. Cañada, J. Jimenez-Barbero, G. Cuevas, *J. Am. Chem. Soc.* **2005**, *127*, 7379.
- [20] M. Muraki, *Protein Pept. Lett.* **2002**, *9*.
- [21] M. Muraki, H. Morii, K. Harata, *Protein Eng.* **2000**, *13*, 385.
- [22] G. Terraneo, D. Potenza, A. Canales, J. Jimenez-Barbero, K. K. Baldrige, A. Bernardi, *J. Am. Chem. Soc.* **2007**, *129*, 2890.
- [23] S. Vandenbussche, D. Díaz, M. C. Fernández-Alonso, W. Pan, S. P. Vincent, G. Cuevas, F. J. Cañada, J. Jiménez-Barbero, K. Bartik, *Chem. Eur. J.* **2008**, *14*, 7570.
- [24] M. Pedraza-Escalona, B. Becerril-Luján, C. Agundis, L. Domínguez-Ramírez, A. Pereyra, L. Riaño-Umbarila, A. Rodríguez-Romero, *Mol. Immunol.* **2009**, *46*, 668.
- [25] E. Rojas, P. Llinas, A. Rodríguez-Romero, C. Hernández, M. Linares, E. Zenteno, R. Lascurain, *Glycoconjugate J.* **2001**, *18*, 339.
- [26] R. T. Lee, H.-J. Gabius, Y. C. Lee, *Glycoconjugate J.* **1998**, *15*, 649.
- [27] K. Harata, M. Muraki, *J. Mol. Biol.* **2000**, *297*, 673.
- [28] F. A. Saul, P. Rovira, G. Boulot, E. J. M. Van Damme, W. J. Peumans, P. Truffa-Bachi, G. A. Bentley, *Structure* **2000**, *8*, 593.
- [29] C. Unverzagt, *Chem. Eur. J.* **2003**, *9*, 1369.
- [30] C. Unverzagt, *Carbohydr. Res.* **1997**, *305*, 423.
- [31] N. L. Allinger, Y. H. Yuh, J. H. Lii, *J. Am. Chem. Soc.* **1989**, *111*, 8551.
- [32] Maestro, A powerful, all-purpose molecular-modeling environment, Schrodinger LLC, **2005**.
- [33] W. C. Still, A. Tempczyk, R. C. Hawley, T. Hendrickson, *J. Am. Chem. Soc.* **1990**, *112*, 6127.

- [34] J. L. Asensio, F. J. Cañada, A. Garcia-Herrero, M. T. Murillo, A. Fernandez-Mayoralas, B. A. Johns, J. Kozak, Z. Zhu, C. R. Johnson, J. Jimenez-Barbero, *J. Am. Chem. Soc.* **1999**, *121*, 11318.
- [35] Calculated with Glytorsion (<http://www.glycosciences.de/tools>); T. Lütteke, M. Frank, C.-W. von der Lieth, *Nucleic Acids Res.* **2005**, *33*, D242.
- [36] K. Bock, J. O. Duus, *J. Carbohydr. Chem.* **1994**, *13*, 513.
- [37] A. Imberty, S. Perez, *Chem. Rev.* **2000**, *100*, 4567.
- [38] J. Jiménez-Barbero, J. L. Asensio, F. J. Cañada, A. Poveda, *Curr. Opin. Struct. Biol.* **1999**, *9*, 549.
- [39] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719.
- [40] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins Struct. Funct. Bioinf.* **2005**, *59*, 687.
- [41] G. M. Morris, D. S. Goodsell, R. S. Halliday, R. Huey, W. E. Hart, R. K. Belew, A. J. Olson, *J. Comput. Chem.* **1998**, *19*, 1639.
- [42] K. N. Kirschner, A. B. Yongye, S. M. Tschampel, J. González-Outeirino, C. R. Daniels, B. L. Foley, R. J. Woods, *J. Comput. Chem.* **2008**, *29*, 622.
- [43] AMBER 9, D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, I. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross, P. A. Kollman, University of California, San Francisco, **2006**.
- [44] A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *65*, 355.
- [45] A. Kumar, R. R. Ernst, K. Wüthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1.
- [46] P. Güntert, *Methods Mol. Biol.* **2004**, *278*, 353.
- [47] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179.

Received: April 13, 2010
Published online: July 22, 2010