CHEMISTRY
A EUROPEAN JOURNAL

DOI: 10.1002/chem.200500367

VIP

On the Importance of Carbohydrate–Aromatic Interactions for the Molecular Recognition of Oligosaccharides by Proteins: NMR Studies of the Structure and Binding Affinity of AcAMP2-like Peptides with Non-Natural Naphthyl and Fluoroaromatic Residues

M. Isabel Chávez,[a, b] Cecilia Andreu,[c] Paloma Vidal,[a] Nuria Aboitiz,[a] Felix Freire,[a, d] Patrick Groves,[a] Juan Luis Asensio,[a] Gregorio Asensio,[c] Michiro Muraki,[e] Francisco Javier Cañada,*[a] and Jesús Jiménez-Barbero*[a]

Dedicated to Professor Manuel Martín-Lomas on the occasion of his 64th birthday

Abstract: The specific interaction of a variety of modified hevein domains to chitooligosaccharides has been studied by NMR spectroscopy in order to assess the importance of aromatic–carbohydrate interactions for the molecular recognition of neutral sugars. These mutant AcAMP2-like peptides, which have 4-fluoro-phenylalanine, tryptophan, or 2-naphthylalanine at the key interacting positions, have been prepared by solid-phase synthesis. Their three-dimensional structures, when bound to the chitin-derived trisaccharide, have been deduced by NMR spectroscopy. By using DYANA and restrained molecular dynamics simulations with the AMBER 5.0 force field, the three-dimensional structures of the protein–sugar complexes have been obtained. The thermodynamic analysis of the interactions that occur upon complex formation have also been carried out. Regarding binding affinity, the obtained data have permitted the deduction that the larger the aromatic group, the higher the association constant and the binding enthalpy. In all cases, entropy opposes binding. In contrast, deactivation of the aromatic rings by attaching fluorine atoms decreases the binding affinity, with a concomitant decrease in enthalpy. The role of the chemical nature of the aromatic ring for establishing sugar contacts has been thus evaluated.

Keywords: carbohydrate binding · hevein · molecular dynamics · molecular recognition · NMR spectroscopy · peptides

[a] M. I. Chávez,[+] P. Vidal,[+] Dr. N. Aboitiz, F. Freire, Dr. P. Groves, Dr. J. L. Asensio, Dr. F. J. Cañada, Prof. Dr. J. Jiménez-Barbero
Department of Protein Structure and Function
Centro de Investigaciones Biológicas, CSIC
Ramiro de Maeztu 9, 28040 Madrid (Spain)
Fax: (+34) 91–553–1706
E-mail: jcanada@cib.csic.es
jjbarbero@cib.csic.es

[b] M. I. Chávez[+]
Instituto de Química, UNAM
Ciudad Universitaria, Coyoacán (México DF)

[c] Dr. C. Andreu,[+] G. Asensio
Departamento de Química Orgánica
Facultad de Farmacia
Universidad de Valencia
46100 Burjassot (Spain)

[d] F. Freire
Departamento de Química Orgánica
Univ. de Santiago de Compostela y
Unidad Asociada al CSIC de RMN de Biomoléculas
Santiago de Compostela (Spain)

[e] M. Muraki
Biological Information Research Center
National Institute of Advanced Industrial Science and Technology
Central 6, 1–1–1 Higashi, Tsukuba, Ibaraki 305–8566 (Japan)

[+] These authors contributed equally to this work.

Supporting information for this article is available on the WWW under http://www.chemeurj.org/ or from the author. Schematic view of the NMR signals of the peptides that are significantly affected by the addition of ligand (S1), chemical shifts assignments (S2), sections of NMR spectra for the three polypeptides (S3), list of the intramolecular protein–protein NOEs unambiguously assigned and converted into relevant distance constraints for the peptides (intraresidue, sequential, medium, and long range) (S4), Ramachandran-type plots of the NMR structures of the three polypeptides (S5), titrations and chemical shift perturbation data of the NMR signals of the mutant peptides in the presence of $(GlcNAc)_3$ (S6), schematic view of the two possible arrangements of the trisaccharide and the intermolecular NOE data at the binding site (S7), NMR/AMBER-based structures of the complex for the Phe18Nal peptide (S8), description of the results of the solvated molecular dynamics simulation for Phe18Trp (S9), and temperature dependence of titrations plots of the mutant peptides in the presence of $(GlcNAc)_3$ (S10).

Introduction

Carbohydrate–protein interactions play an important role in a wide variety of biological processes, including immunological and inflammatory responses, organogenesis, metastasis, and diverse infections.[1] In these processes, defense mechanisms against pathogen microorganism invasions such as those produced by bacteria, viruses, and parasites, are included, since on many occasions they essentially depend on the molecular recognition of specific carbohydrates by proteins.[2]

In this context, the determination of the structural and conformational factors, which govern the molecular recognition of these biomolecules, as well as the knowledge of the physicochemical features of these processes, is of paramount importance.[3]

Previous investigations by X-ray crystallography, NMR spectroscopy, molecular modeling, and calorimetric studies have allowed information to be obtained on the structural and thermodynamic characteristics of a variety of protein–carbohydrate complexes.[4] Due to the amphipathic character of oligosaccharides, different kinds of forces may be involved in the recognition process by a given protein. The presence of the oxygen atoms of the hydroxyl groups evidently provides a potential involvement in intermolecular hydrogen bonds with the side chains of polar amino acids within the polypeptide chain. Nevertheless, not only polar forces are involved in carbohydrate recognition. Depending on the stereochemistry of the monomer constituents of the oligosaccharide chain, the presence of a number of rather apolar C−H groups constitute patches that provide van der Waals, CH–π, and hydrophobic interactions.[5]

NMR spectroscopy has been used to determine the conformations of diverse carbohydrates in the binding site of proteins, generally by means of transferred NOE experiments. Conformations are usually determined in combination with molecular dynamics calculations, thus leading to an understanding, to varying degrees, of the factors involved in molecular recognition processes.[6]

Among the processes in which carbohydrates are recognized by proteins, there are some small proteins of plant origin that bind chitin, a polymer of *N*-acetylglucosamine with β(1→4) glycosidic linkages.[7] These protein domains have been related with the defense of plants against insects and fungi,[8] and usually include a structural motif (sequence) of around 30 to 45 amino acids linked with three or four disulfide bonds, denominated the hevein domain.[9]

In the last years, the three-dimensional structures of several hevein domains have been determined, both in the free and in the sugar-associated state, by using X-ray crystallographic analysis and/or NMR spectroscopy. As leading examples, the cases of wheat germ agglutinin[10,11] and its isolated B domain,[12] *Urtica dioica* agglutinin,[13,14] and hevein itself isolated from latex[15–20] have been reported. Additional studies for other domains have also been carried out,[21,22] including one on the smaller antimicrobial peptide AcAMP2.[23,24]

According to these X-ray and NMR studies, the aromatic residues at relative positions 21, 23, and 30 in hevein domains (Phe18, Tyr20, Tyr27 in Ac-AMP2) play an important role in carbohydrate binding by stabilizing the complexes by means of interactions between the aromatic and sugar rings through van der Waals contacts and CH–π stacking. Additionally, Ser19 of hevein (Ser16 in Ac-AMP2) is involved in a hydrogen-bonding interaction with the carbonyl group of the acetamide moiety of one GlcNAc residue, and at the same time the methyl group of this acetamide interacts with the side-chain aromatic rings of residues 23 and 30 (Tyr20 and Tyr27 in Ac-AMP2). As a final key interaction, the hydroxyl group of hevein Tyr30 (AcAMP2 Tyr27) also provides additional hydrogen bonding with OH-3 of the same sugar moiety. The same kind of interactions take place in other proteins with one or more hevein domains on the same polypeptide chain, including pseudohevein, *Urtica Dioica* agglutinin (UDA), pokeweed agglutinin, and wheat germ agglutinin (WGA).[15–22] No detailed analysis of the binding mode to the antifungal AcAMP2 peptide has been reported to date.[23]

The origin of the specificity and stability of protein–carbohydrate complexes has become a topic of major discussion,[5,25–28] and, at this point, it seems relevant to verify the relative roles of carbohydrate–aromatic stacking and hydrogen-bonding interactions for the recognition of sugars. The small size of hevein-like peptides, as well as the availability of their three-dimensional structures, make them a suitable system for the study of the structural features and energetics involved in their sugar interaction processes.[29,30]

On this basis, we are currently engaged in a multidisciplinary project devoted to the modification of the key interacting residues of hevein domains in order to understand the origin of protein–carbohydrate interactions. Thus, synthetic peptides have been prepared by using solid-phase synthesis with the aim of modifying the chemical nature of the key aromatic amino acids involved in sugar binding. One of us has previously reported on the chemical synthesis of mutated AcAMP2 peptides.[31,32] Thus, the aromatic amino acid at relative position 18 (Phe in AcAMP2) was mutated to alanine (non-aromatic residue) and to residues either with larger electron density and aromatic surface such as tryptophan (Trp, Phe18Trp) and naphthylalanine (2-Nal, Phe18Nal), or with the electron-deficient aromatic rings pentafluorophenylalanine and nitrophenylalanine.[5,31,32] Herein, we have studied the three-dimensional structure and thermodynamics of the synthetic tryptophan-, 2-naphtylalanine-, and 4-fluorophenylalanine (Pff)-AcAMP-2 mutants bound to chitooligosaccharides. This last variation has also been combined with the simultaneous mutation of Tyr20 to a second 4-fluorophenylalanine residue (Pff). Thus, it is our aim in this work to deduce the influence of the electronic density of the aromatic rings of the lectin in the energetics of binding to the sugars. The three-dimensional structure of the modified lectins in their free and $(GlcNAc)_3$-bound state has been studied in aqueous solution by NMR spectroscopy and modeling methods, including molecular dynamics. Fur-

thermore, the thermodynamics of the binding processes with the artificial peptides has been characterized both by fluorescence (for Phe18Trp), ^{1}H NMR, and ^{19}F NMR spectroscopy and compared with a chemically synthesized sample of AcAMP-2 that is identical to the natural sequence. Previous qualitative data on the affinity of the Phe18Trp and Phe18Nal towards large polymeric chitin have been reported.[32]

A schematic view of the sequences is given in Figure 1.

Results

Synthesis: The synthesis of Phe18Trp and Phe18Nal have been described elsewhere.[31,32] As with AcAMP2, Phe18Pff/Tyr20Pff is a polypeptide that consists of 30 amino acid residues with three intramolecular disulfide bonds in its structure and was obtained as an amide peptide. In this modified AcAMP2 analogue, Phe18 and Tyr20 of the native AcAMP2 were changed to 4-fluorophenylalanine (Pff).

The amino acids were manually assembled by solid-phase synthesis by using Fluoren-9-ylmethoxycarbonyl (Fmoc) chemistry according to standard protocols, with DCC and 1-hydroxybenzotriazole hydrate (HOBt) as coupling reagents.

The side chains of the six Cys residues of the sequence were protected with the same protecting group (trityl), which were removed by trifluoroacetic acid (TFA) treatment for resin cleavage. After testing several experimental conditions[33] (redox pairs, temperatures, denaturants, pH, and buffers), the fully reduced peptide was subjected to oxidative folding in NH_4Ac buffer (0.1 M, pH 7.8) in high dilution conditions (peptide concentration 3.1×10^{-5} M) in the presence of reduced and oxidized glutathione (GSH/GSSG) as redox reagents for three days at room temperature (molar ratio peptide:GSH:GSSG 1:100:10). After folding, the reaction was stopped by acidification to pH 2.5 with TFA and lyophilized. The crude material was analyzed by RP-HPLC (linear gradient from 10% to 90% of acetonitrile in water containing TFA 0.09%, over 30 min at flow rate of 1 mL min^{-1}) showing the presence of several different isomers (retention times between 10.4 and 13.5 min), but clearly the first peak represented the major product (retention time 10.6 min). Isolation and purification by using semipreparative RP-HPLC afforded the mutant AcAMP2 peptide in 30% yield from the starting material.

The synthesis and full characterization by HPLC and ESMS is detailed in the Experimental Section.

Thermodynamic analysis of chitooligosaccharide binding to AcAMP2, and to the Phe18Trp, Phe18Nal, and Phe18Pff/Tyr20Pff mutants: The equilibrium association constants (K_a) were first obtained by one-dimensional ^{1}H NMR titrations. Thus, the binding constants of $(GlcNAc)_3$ to Phe18Trp, Phe18Nal, and Phe18Pff/Tyr20Pff were obtained (Table 1) by utilising one-dimensional ^{1}H NMR spectra of a series of peptide samples with increasing trisaccharide concentrations.[34] Also, similar experiments were performed for the parent AcAMP2 for testing purposes. For the Phe18Pff/Tyr20Pff peptide, additional spectra were recorded using ^{19}F NMR spectroscopy. For Phe18Trp, additional binding data were obtained by fluorescence titration. In this titration, as observed for hevein, hevein-32,[34] and wheat germ agglutinin,[35] a blue-shift was observed for the tryptophan emission of Phe18Trpwith the peak maximum shifting from

Table 1. Affinity data from the NMR analysis (and fluorescence for Phe18Trp) for the binding of *N*,*N*′,*N*″-triacetylchitotriose to natural AcAMP2 and to the other peptides mutated at positions 18 and 20. Estimated errors amount to ±15%. The thermodynamic parameters estimated from a van't Hoff analysis are also given. An enthalpy–entropy compensation phenomenon is evident.

	K_a [M^{-1}]				Thermodynamic parameters	
	T=298 K	T=303 K	T=308 K	T=313 K	ΔH [kJ mol^{-1}]	ΔS [J K^{-1} mol^{-1}]
AcAMP2	1206	721	537	340	−63.4	−154.1
Phe18Trp	1755	1473	916	614	−54.1	−119.0
Phe18Nal	3527	2314	1638	999	−63.9	−146.6
Phe18Pff/Tyr20Pff	448	348	255	190	−45.1	−101.0

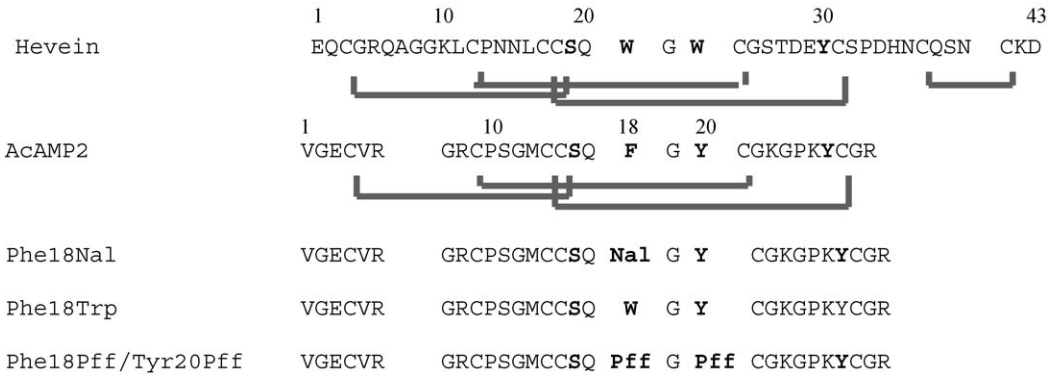

Figure 1. Amino acid sequence alignment of AcAMP2 and hevein (indicating the disulfide bridge pattern) with the three mutant ACAMP2 polypeptides used in this study. The residues involved in the binding site are highlighted.

357 nm in the protein-free form to 352 nm in the protein-bound form (Figure 2). This shift in peak maximum is consistent with a solvent exposed tryptophan of the protein becoming shielded from water molecules upon complexation with the ligand. A 46% increase in fluorescence intensity at 340 nm falls between the 28% increase reported for hevein and 89% increase reported for hevein-32.[34] In the NMR titration, a variety of peptide NMR signals are significantly affected by the addition of ligand (see Supporting Information, S1). The perturbations in chemical shifts allowed the determination of the equilibrium association constants.[36] For AcAMP2, and Phe18Trp, Phe18Nal and Phe18Pff/Tyr20Pff, the backbone NH protons of either Ser16 or other NH resonance signals were followed as a function of the

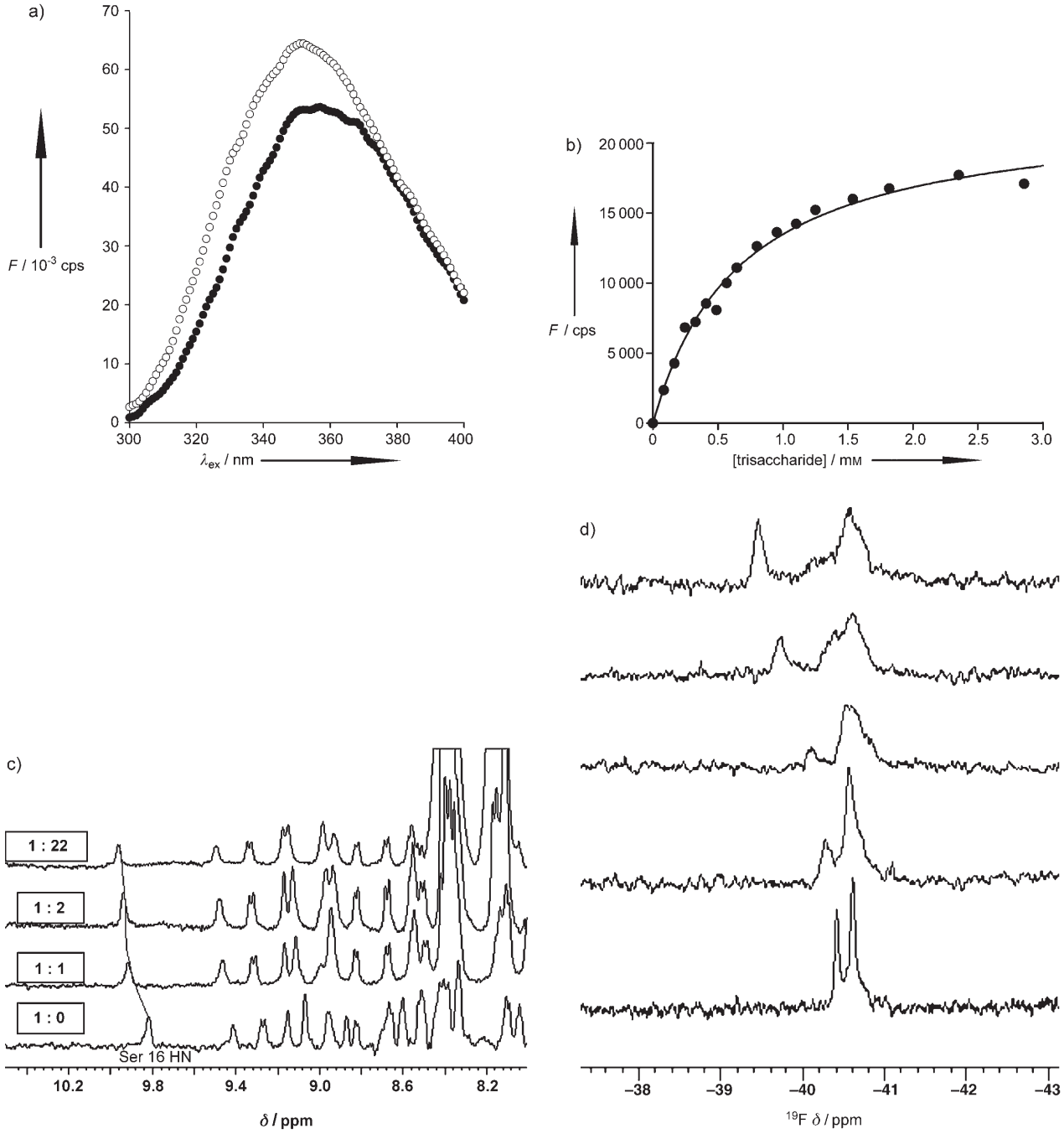

Figure 2. Fluorescence spectroscopy of Phe18Trp. a) Phe18Trp free (filled circles) and trisaccharide-bound (open circles) fluorescence spectra. b) The change in fluorescence of Phe18Trp at 340 nm was monitored as a function of trisaccharide concentration. The data was fitted in SigmaPlot to a first-order binding equation to obtain the binding parameters. c) ^{1}H NMR titrations of Phe18Nal (0.7 mM) with increasing amounts of *N*,*N*′,*N*″-triacetylchitotriose at 298 K and pH 5.6. The variations of Ser16 HN are highlighted. d) ^{19}F NMR titrations of Phe18Pff/Tyr20Pff with *N*,*N*′,*N*″-triacetylchitotriose. The variations of the ^{19}F signals with the ligand to protein molar ratio are shown at 298 K. The presence of a dynamic process is evident from the broadening of the signals. More than one sugar–lectin complex is present in solution.

concentration of added chitotriose to assess the binding constants (Figure 2C, Table 1). For Phe18Pff/Tyr20Pff, the ^{19}F NMR spectra permitted a better evaluation of the titration data.

A van't Hoff plot of NMR-determined K_a values as a function of temperature was used to estimate the equilibrium thermodynamic parameters, $\Delta H°$ and $\Delta S°$ (Figure 3, Table 1).

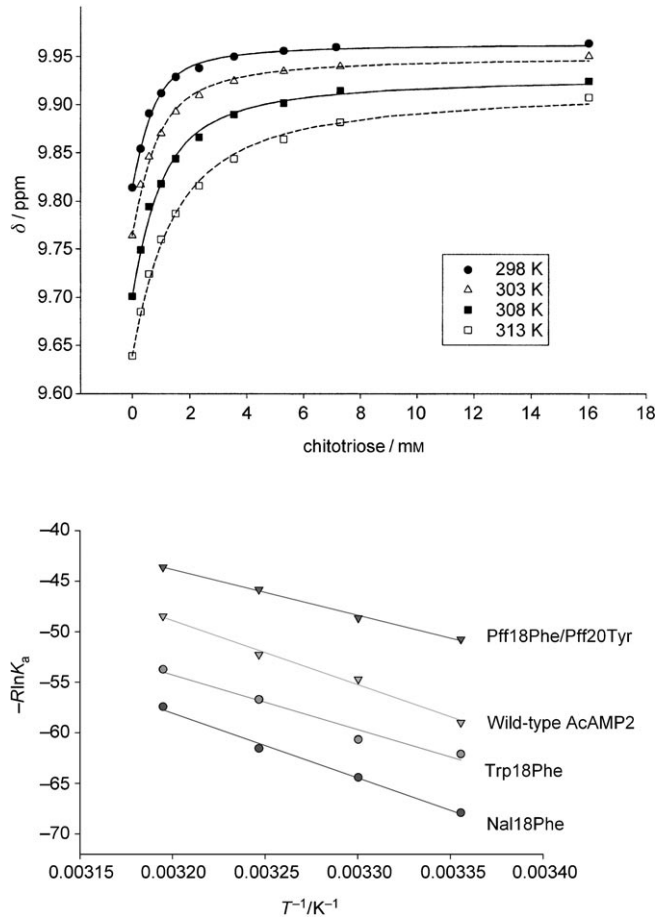

Figure 3. Thermodynamic parameters of $(GlcNAc)_3$ binding to AcAMP2 and its mutant peptides, as determined by NMR spectroscopy. Top: The binding curves derived from the NMR titrations for the association of $(GlcNAc)_3$ to Phe18Nal are shown. Bottom: The van't Hoff approximation for the binding of $(GlcNAc)_3$ to AcAMP2 and the three mutant peptides. In all cases, the binding process is enthalpy driven.

As reported for other hevein domains, the interaction with chitotriose is enthalpy driven and entropy opposes binding.[17–22] It should be noted that although the linear assumption in the van't Hoff plot is only approximated, our previous studies with hevein domains have demonstrated that NMR values differ by $<10\%$ from those obtained by isothermal titration microcalorimetry.[17–22]

^{19}F NMR titration experiments for chitotriose binding to Phe18Pff/Tyr20Pff at 298 K resulted in one of the signals being shifted downfield when the concentration of the sugar increases. In contrast, the other peak is remarkable in being clearly broadened and split at a 1:7 molar ratio, remaining split at larger molar ratios. This experimental observation suggests the presence of different association modes of the trisaccharide to the peptide (see Figure 2 and below). The peptides studied herein only differ in the chemical nature of the aromatic residue at position 18 (and 20 in one case), thus allowing a direct comparison of the influence of this ring (see Discussion).

The three-dimensional NMR structure of the AcAMP2 mutants bound to $(GlcNAc)_3$

The protein: The assignments of the ^{1}H NMR spectra of the free and bound Phe18Trp, Phe18Nal, and Phe18Pff/Tyr20Pff in aqueous solution were facilitated by the reported data for the natural peptide (see Supporting Information for chemical shifts assignments, S2). (The NMR data have been deposited in the BMRB database, with accession numbers 6639 and 6647 for free and bound Phe18Trp, 6657 and 6637 for free and bound Phe18Nal, and 6656 and 6591, for free and bound Phe18Pff/Tyr20Pff.) Indeed, the assignments of the three polypeptides were obtained by using well-established protocols through the assignment of regular two-dimensional TOCSY and NOESY spectra. For the complexes, according to the measured binding constant, K_a (Table 1), and experimental conditions (peptide concentration 0.3–1.0 mM and $(GlcNAc)_3$ at a 8:1 molar ratio), the different polypeptides are more than 60% bound at 298 K. For both the free and bound peptides, different sets of intramolecular protein–protein NOEs were unambiguously assigned and converted into relevant distance constraints (intraresidue, sequential, medium, and long range) as depicted in Table 2

Table 2. Number and type of conformationally significant NOEs for the complexes of the three mutant peptides to $(GlcNAc)_3$.

	Total assigned NOEs	Intra-residue	Sequential	Medium range	Long range
Phe18Nal	348	64	140	52	99
Phe18Trp	331	60	135	59	77
Phe18Pff/Tyr20Pff	248[a]	79	98	24	47

[a] After "distance modify" and CALIBA routines.

(and Supporting Information S3, S4). For Phe18Pff/Tyr20Pff, the low availability of the samples did not allow the use of higher protein concentrations and the number of detected significant NOEs is lower. Thus, starting from 400, 500, and 800 randomized conformations and applying the DYANA protocol,[37] a group of 50 structures with low target-function values were obtained (Table 3). The 25 best DYANA structures of Phe18Trp and Phe18Nal, as well as the 35 best structures of Phe18Pff/Tyr20Pff were subjected to further optimization through a simulated annealing protocol by restricted molecular dynamics (rMD) with the AMBER force field. The average backbone root mean square deviations (rmsd) of the refined AMBER structures was below 0.9 Å,

Table 3. Statistics from DYANA and AMBER restrained MD calculations. The data from the DYANA analysis and AMBER restrained MD calculations are given. The limits of variation interval are in parentheses.

Peptide	Method	Number of structures	Range	Backbone RMSD [Å]	Heavy atom RMSD [Å]	DYANA target function
Phe18Nal	DYANA	25	4–28	0.61 ± 0.15 (0.24–1.05)	1.17 ± 0.17 (0.63–1.68)	0.026 (0.012–0.036)
	AMBER	30	4–28	0.81 ± 0.18 (0.28–1.42)	1.45 ± 0.27 (0.77–2.25)	
Phe18Trp	DYANA	25	4–28	0.73 ± 0.23 (0.25–1.56)	1.53 ± 0.30 (0.76–2.44)	0.43 (0.33–0.50)
	AMBER	26	4–28	0.80 ± 0.23	1.86 ± 0.41	
Phe18Pff/Tyr20Pff	DYANA	35	4–28	0.60 ± 0.13 (0.15–0.90)	1.40 ± 0.19 (0.74±2.02)	0.071 (0.044–0.086)
	AMBER	25	4–28	0.84 ± 0.18 (0.00–1.36)	1.96 ± 0.37 (0.00–2.68)	

while the heavy atoms were below 2.0 Å, considering residues 4–28 in all cases (see Table 3 and Supporting Information S5). Moreover, the obtained structures have very small deviations from ideal geometry and there are no nonbonded contacts (the coordinates of the three three-dimensional structures of the mutant peptides have been deposited in the protein data bank, with accession codes 1ZUV, 1ZWU, and 1ZNT, for Phe18Trp, Phe18Nal, and Phe18Pff/Tyr20Pff, respectively). Schematic representations of the superimposition of the backbone and key side chains of the three mutant peptides, as derived from the NMR data, by using restricted molecular dynamics are depicted in Figure 4. The two segments located between residues Met13 and Lys23 build a two strand antiparallel β-sheet, which is also present in the natural AcAMP2[23] and hevein.[15–20] Indeed, NOEs characteristic for the β-sheet (Cys15HA-Cys21HA, Cys14HN-Lys23HA, and Cys15HN-Gly21HA) were identified for all the mutant polypeptides.

The complex: While the structure of the protein was well defined by the NMR data, for all three modified AcAMP2 peptides there was also a number of intermolecular ligand–protein NOEs that supported the experimental structures for the complexes (see Figure 5, Table 4, and the Supporting Information S7). In spite of the strong overlapping of the chemical shifts of the sugar residues protons, the comparison with previous chitotriose complexes of hevein allowed the unambiguous assignment of some ligand protons[20,21] and their intermolecular cross-peaks. Regarding the saccharide, the solution conformation of $(GlcNAc)_3$ in the free state[38] has been studied. The conformers are in agreement with the occurrence of the exo-anomeric effect and correspond to dihedral angles of $\Phi = 50 \pm 20°$ and $\Psi = 0 \pm 20°$ for the glycosidic linkages.[20,21] These angles also agree with data obtained from standard NMR methods, molecular mechanics, and dynamics calculations for chitooligosaccharides bound to hevein[16–20] and hevamine.[38] The sugar-binding site of diverse hevein domains have been studied by NMR spectroscopy,[15–21] and the contact points correspond to Ser16, Phe18, Tyr20, and Tyr27 in wild-type AcAMP2. In the case of the three modified AcAMP2 peptides under study, the presence of a similar binding arrangement was deduced

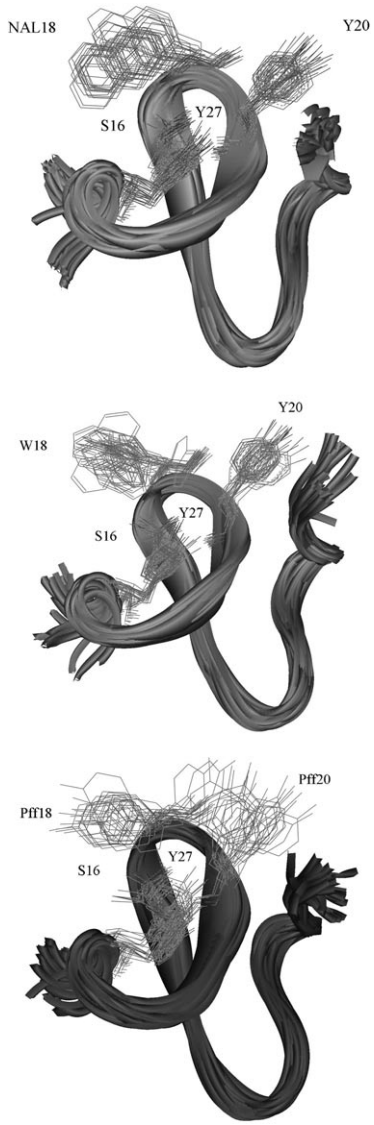

Figure 4. Ribbon representation of the superimposition of the backbone of the NMR-derived structures of the AcAMP2 mutants when bound to $(GlcNAc)_3$. In this representation, the sugar has been omitted for clarity. The structures of Phe18Nal (top), Phe18Trp (middle), and Phe18Pff/Tyr20Pff (bottom) are shown. The location of the side chains of the key amino acids for sugar binding is also highlighted. The N and C termini of the peptides are marked in each panel.

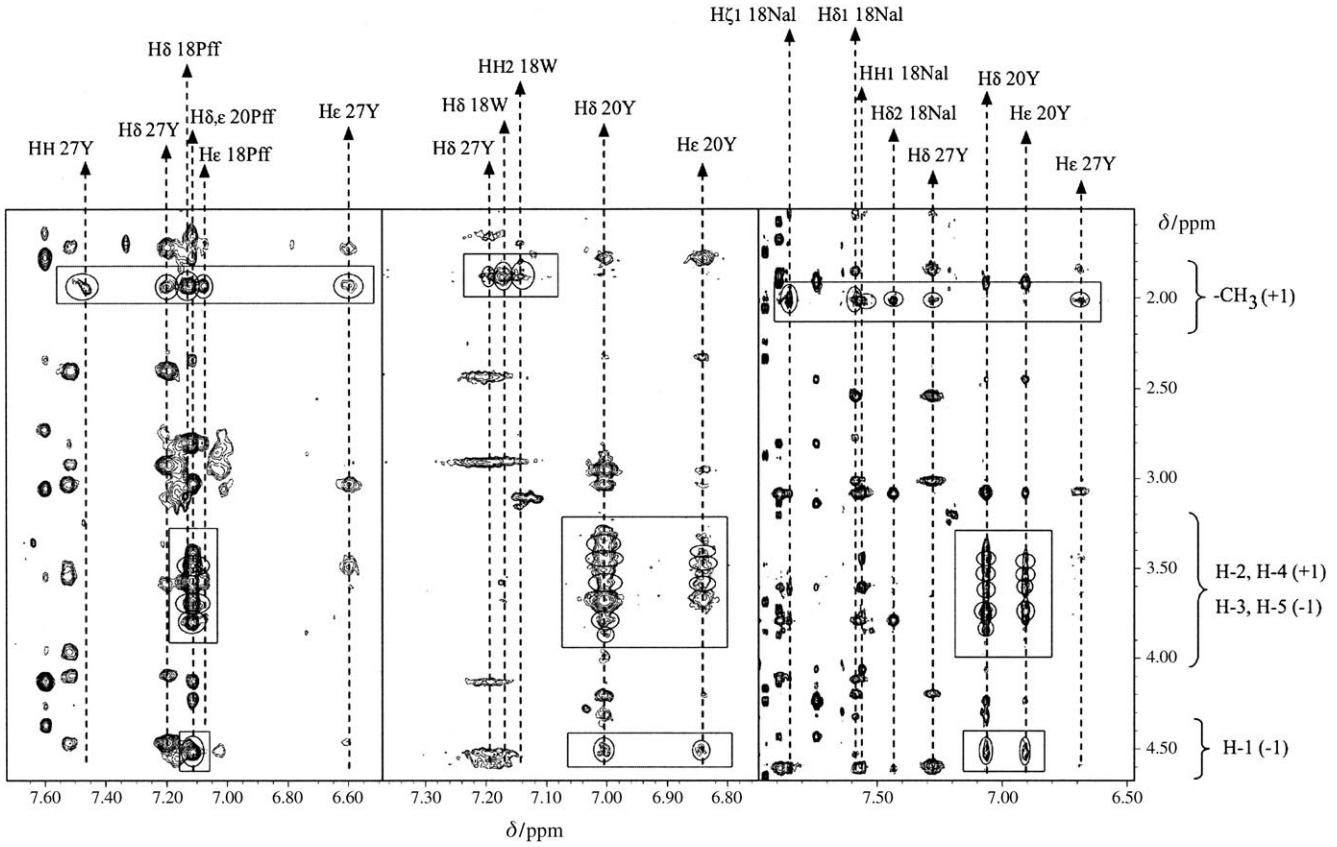

Figure 5. Sections of the 800 MHz NOESY spectra (298 K, 200 ms mixing time), recorded for the complexes of the three polypeptides with $(GlcNAc)_3$, showing the assigned sugar–lectin intermolecular NOEs. Phe18Pff/Tyr20Pff (left), Phe18Trp (middle), and Phe18Nal (right).

from the chemical shift perturbation data of the NMR signals of the mutant peptides (free versus bound, see Supporting Information S1 and S6). In particular, important chemical shift differences between the free and bound species were mainly observed for Ser16HA, as well as for Cys21HA and Cys28HA, located immediately after two of the key aromatic residues.

Moreover, the position of the binding site was corroborated by the existence of key intermolecular interactions involving these particular residues in the NOESY spectra (see below and Figure 5). Previous studies on the interaction of hevein domains with small chitooligosaccharides (up to the trisaccharide in length) have also shown that the lectin binding site is sufficiently extended for interacting with at least two *N*-acetylglucosamine units, placed on the two primary subsites arranged with the side chains of residues 16, 18, 20, and 27. Therefore, for a trisaccharide, it is possible to consider two possible arrangements at the binding site (see Figure 6 and Supporting Information S7). Additional evidence for the existence of two binding modes for chitotriose also came from the ^{19}F NMR titrations of Phe18Pff/Tyr20Pff, as mentioned above.

In fact, the observed intermolecular saccharide–lectin NOESY cross-peaks are fittingly explained when considering the simultaneous existence of two distinct binding modes (Table 4). For example, the intermolecular NOE data for the complex of $(GlcNAc)_3$ to Phe18Pff/Tyr20Pff will be detailed, while those for the other two mutants basically follow the same trends (Figure 6 and Supporting Information S7).

In the NOESY spectra of the bound Phe18Pff/Tyr20Pff, NOE cross-peaks are observed between the acetamide methyl group of a nonreducing sugar unit with the aromatic protons of the residues Pff18 and Tyr27; this places the corresponding sugar unit over the aromatic side chain of residue Pff20 and defines the subsite +1 (Figure 6 top). This binding mode may also explain the observed interactions between the H-2 and the H-4 protons from the nonreducing end sugar[21] with the aromatic protons of Pff20. These cross-peaks are specific to the so-called binding mode 1.

In contrast, the NOE cross-peaks between the anomeric proton of the nonreducing unit of the trisaccharide with the aromatic protons of Pff20 are only possible if this proton belongs to the nonreducing end sugar unit when it is located at subsite (−1). The existence of this NOE indicates the presence of so-called binding mode 2. In this manner, a number of intermolecular NOEs were assigned (Figures 5 and 6, and Supporting Information S7) and related to either of the two possible binding modes. Then, the three-dimensional structures of the two possible complexes were obtained through a simulated annealing protocol from the AMBER force field program, considering the two binding modes previously described. The peptide–peptide (as described above) and the peptide–trisaccharide restrictions, included as distance

Table 4. Intermolecular peptide–carbohydrate NOE interactions for AcAMP2 Phe18Nal, AcAMP2 Phe18Trp, and AcAMP2 Phe18Pff/Tyr20Pff with chitotriose.

$H_{peptide}$	$\delta H_{Phe18Nal}$	$\delta H_{Phe18Trp}$	$\delta H_{Phe18Pff/Tyr20Pff}$	H_{carboh}	δH_{carboh}	Binding mode 1[a] subsites +3, +2, +1	Binding mode 2[a] subsites +2, +1, −1
18 Hδ1	7.57			CH_3	2.01	+1	+1
18 Hδ1		7.17		CH_3	1.96	+1	+1
18 Hδ2	7.43			CH_3	2.01	+1	+1
18 HZ1	7.85			CH_3	2.01	+1	+1
18 HH1	7.55			CH_3	2.01	+1	+1
18 HH2		7.15		CH_3	1.96	+1	+1
18 Hδ1,2			7.12	CH_3	2.03	+1	+1
18 Hε1,2			7.08	CH_3	2.03	+1	+1
20 Hδ1,2	7.06			H1	4.51	–	−1
20 Hδ1,2	7.06			H2	(3.73, 3.82)[b]	+1	+1
20 Hδ1,2	7.06			H3	3.57	–	−1
20 Hδ1,2	7.06			H4	(3.48, 3.62)[b]	+1	+1
20 Hδ1,2	7.06			H5	3.50	–	−1
20 Hδ1,2		7.01		H1	4.48	–	−1
20 Hδ1,2		7.01		H2	3.78	+1	+1
20 Hδ1,2		7.01		H3	3.57	–	−1
20 Hδ1,2		7.01		H4	3.43	+1	–
20 Hδ1,2		7.01		H5	3.48	–	−1
20 Hδ1,2, Hε1,2			7.11	H1	4.52	–	−1
20 Hδ1,2, Hε1,2			7.11	H2	(3.69, 3.80)[b]	+1	+1
20 Hδ1,2, Hε1,2			7.11	H4	(3.48, 3.69)[b]	+1	+1
20 Hε1,2	6.91			H1	4.51	–	−1
20 Hε1,2	6.91			H2	(3.73, 3.82)[b]	+1	+1
20 Hε1,2	6.91			H3	3.57	–	−1
20 Hε1,2	6.91			H4	(3.48, 3.62)[b]	+1	+1
20 Hε1,2	6.91			H5	3.50	–	−1
20 Hε1,2		6.84		H1	4.48	–	−1
20 Hε1,2		6.84		H3	3.57	–	−1
20 Hε1,2		6.84		H4	3.43	+1	–
20 Hε1,2		6.84		H5	3.48	–	−1
27 Hδ1,2	7.28			CH_3	2.01	+1	+1
27 Hδ1,2		7.20		CH_3	1.96	+1	+1
27 Hδ1,2			7.19	CH_3	2.03	+1	+1
27 Hε1,2	6.69			CH_3	2.01	+1	+1
27 Hε1,2			6.60	CH_3	2.03	+1	+1
27 HH			7.46	CH_3	2.03	+1	+1

[a] In binding mode 1, +1 corresponds to the reducing end sugar residue and "–" corresponds to the middle sugar residue. In binding mode 2, when −1 corresponds to the reducing end sugar residue, +1 corresponds to the middle one sugar residue, and "–" corresponds to the sugar residue at the nonreducing end. [b] Chemical shifts for both binding modes, respectively.

constraints, derived from the NOESY spectra were used as input to generate the three-dimensional structures of the complexes.

Besides the intermolecular NOEs, two intermolecular hydrogen bonds were also included between the trisaccharide and the hydroxyl groups of Ser16 and Tyr27. These hydrogen bonds have been described in all the hevein-domain complexes reported to date.[15–20] The NMR/AMBER-based structures of the two complexes for Phe18Pff/Tyr20Pff, which just differ in the position of the trisaccharide in the binding site, are shown in Figure 6 (bottom), while that for Phe18Nal is given in the Supporting Information S8.

Solvated molecular dynamics simulations: In a parallel manner, a series of solvated molecular dynamics simulations starting from the experimental NMR derived structures was performed in order to obtain, by a semi-independent manner,[3,29,39] a complete view of the three-dimensional structures and possible motions of the peptide–chitooligosaccharide complexes, especially to complement Phe18Pff/Tyr20Pff, for which the number of experimental restraints was smaller due to the low availability of the sample. Therefore, to evaluate the influence of the NOE-based restrictions on the conformational properties of the mutant peptides in solution, solvated MD simulations (sMD) were performed for the free (sMD_{free}) and $(GlcNAc)_3$-bound ($sMD_{complex}$) peptides, without NOE restrictions. Again, a description of the results for Phe18Pff/Tyr20Pff will be presented, the results for Phe18Trp and Phe18Nal being very similar (see Supporting Information, S8, S9). Since within the context of the present paper the main aim of the sMD simulations was to support the NMR data for Phe18Pff/Tyr20Pff, a deeper discussion of the results of all the solvated MD simulations will be presented separately. In the case of the complexes, the calculations were performed for the two different binding modes.

In all cases, the solvated MD simulations (sMD) started from the corresponding restricted rMD_{free} and $rMD_{complex}$

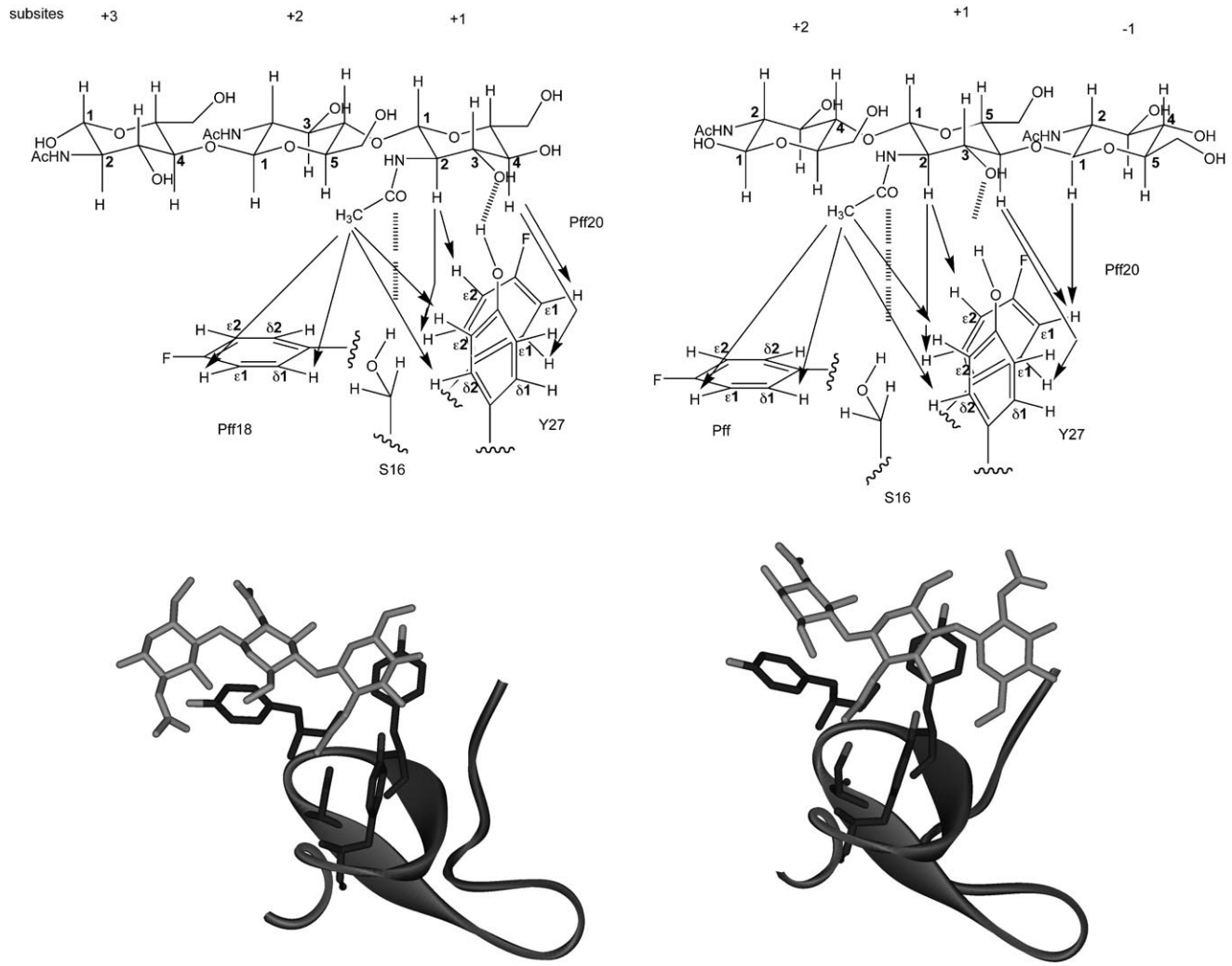

Figure 6. Top: Schematic views of the intermolecular NOEs detected for the complex between $(GlcNAc)_3$ and Phe18Pff/Tyr20Pff. The two possible binding modes are shown. Bottom: Representations of the corresponding three-dimensional structures, as obtained by restrained MD calculations, using the Amber 5.0 force field for bound Phe18Pff/Tyr20Pff. The non-ambiguous intraprotein and intermolecular protein–trisaccharide NOEs were used as restrictions. Two hydrogen bonds between the trisaccharide and the residues Ser16 and Tyr27 were also added.

structures described above. No NMR constraints were applied to retain the starting structure of the proteins or to maintain $(GlcNAc)_3$ in contact with the protein. It is worth noting that the trisaccharide remained at the binding site and orientation during all simulation times in those calculations carried out for the complexes.

From the methodological viewpoint, the sMD were carried out using AMBER 5.0 with explicit water molecules, counterions, and periodic boundary conditions, while Ewald sums were used for the treatment of electrostatic interactions.[40]

A detailed representation of the trajectories of the key torsions that define the orientation of the two aromatic rings of Phe18Pff/Tyr20Pff in the absence and presence of chitotriose are shown in Figure 7.

Discussion

Thermodynamic analysis of peptide binding to $(GlcNAc)_3$: Previous studies have shown that the $\Delta G°$ values for hevein binding to $(GlcNAc)_3$, based on fluorescence ($-22.9\,kJ\,mol^{-1}$) and NMR ($-23.1\,kJ\,mol^{-1}$) data, are very similar to ΔG measured by isothermal titration calorimetry ($-22.6\,kJ\,mol^{-1}$).[20,34] Interestingly, the ΔG of binding of a truncated analogue of hevein, with the same size as the AcAMP2 analogues and dubbed HEV32, for $(GlcNAc)_3$ was shown to be only slightly weaker (1.0 (NMR) and $2.2\,kJ\,mol^{-1}$ (fluorescence))[34] than the corresponding value for hevein (Table 1). These data suggested that the C terminus of hevein does not appear to make a significant contribution to the overall binding affinity of chitooligosaccharide ligands.

The binding data from fluorescence and NMR for Phe18Trp are similar. As mentioned above, a 5 nm blue-

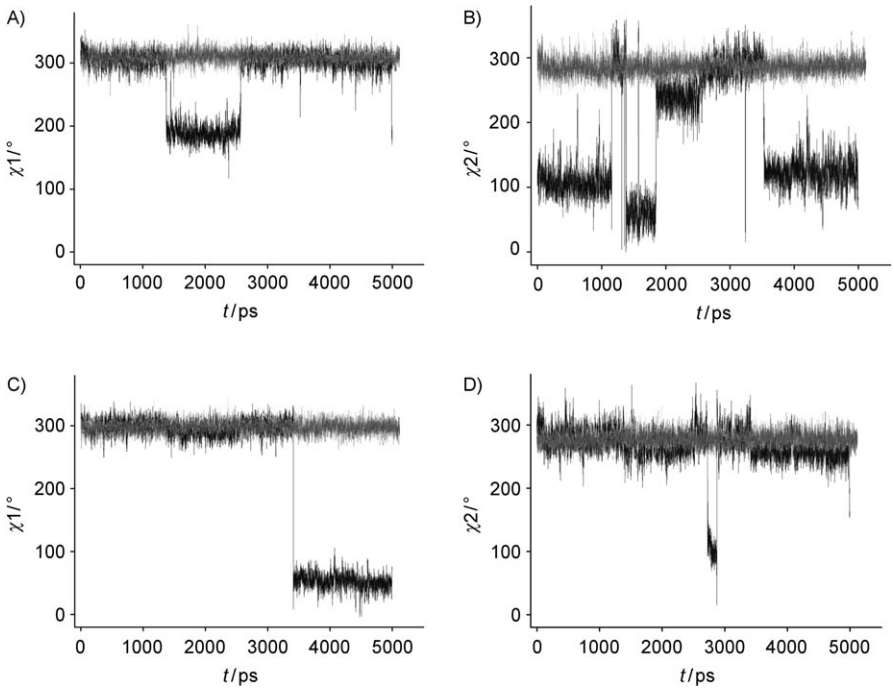

Figure 7. Schematic representation of the trajectory of torsion angles $\chi 1$ (A, C) and $\chi 2$ (B, D), for the side chains at position 18 (A, B) and position 20 (C, D) of Phe18Pff/Tyr20Pff during the solvated MD simulation. A) Free peptide (black), $(GlcNAc)_3$-bound peptide (gray). B) Free peptide (black), $(GlcNAc)_3$-bound peptide (gray). C) Free peptide (black), $(GlcNAc)_3$-bound peptide (gray). D) Free peptide (black),$(GlcNAc)_3$-bound peptide (gray). Both binding modes 1 and 2 (see text) basically gave the same monotonic trajectory. The importance of the presence or absence of the trisaccharide is remarkable.

shift was observed for the Trp emission of Phe18Trp. The observed 46% increase in fluorescence intensity at 340 nm is in between the values reported for hevein and hevein-32.[39] This fluorescence increase is related to changes in fluorescence energy transfer (FRET) between tyrosine and tryptophan residues in the saccharide binding site of the protein. A direct comparison cannot be made as the aromatic residues of Phe18Tyr (2Tyr, 1Trp) are different to those in hevein and hevein-32 (1Tyr, 2Trp). Nonetheless, the increase of 46% in fluorescence unambiguously indicates that changes in the distance and/or relative orientation between Tyr20 and Tyr27 with respect to Trp18 in Phe18Trp occur during the ligand-binding process.

The trisaccharide-induced changes in the tryptophan fluorescence spectrum of Phe18Trp allowed the binding affinity to be measured from a titration experiment. At 25°C, the binding constant, K_a was measured as $1120 \pm 200\,M^{-1}$, on the same order of magnitude than the NMR data for AcAMP2 and the three modified analogues (see Table 1). Despite the fact that the measured K_a values for all the mutants are about one order of magnitude lower than for full-length hevein, the van't Hoff data for the AcAMP2 peptides, analogously to the truncated HEV32, show larger $\Delta H°$ values than that of hevein, but compensated with an increased $\Delta S°$ that highly opposes association (Table 1). Although the origin of this enthalpy–entropy compensation phenomenon remains an open question,[25,26] it has been reported,[41] at this magnitude and sign of ΔS and ΔH, that hydrogen bonding and van der Waals forces should be the most important factors that stabilize the complex. Part of the observed negative entropy of binding could arise from rigidification of the carbohydrate and/or protein lateral chains,[25,26,42] or by reorganization of the water structure. As deduced in earlier studies,[15–21] the maximum loss of conformational entropy by freezing of the $(GlcNAc)_3$ ligand upon binding to lectin domains can reach $17\,kJ\,mol^{-1}$ at 25°C (the conformational entropy was estimated from the conformational distribution map of $(GlcNAc)_3$),[15–20,38] much less than the entropy penalty determined for the trisaccharide binding to the mutants, which is between 30 and $46\,kJ\,mol^{-1}$ (Table 1). Thus, apart of the rigidification of the sugar, an additional loss of entropy is taking place. It seems that these shorter hevein domains must pay an entropic penalty with some reduction of the lateral chains flexibility upon binding, in order to better accommodate the trisaccharide and provide a more favorable binding enthalpy.

The inverse of K_a values (dissociation constants) for $(GlcNAc)_3$ binding are in the millimolar range. However, there is a significant trend in the obtained values that is above experimental error and correlates with the qualitative data of chitin binding obtained previously. The obtained binding constant K_a values for native AcAMP2 and the modified peptides indicated that the larger the aromatic surface of the residue at position 18, the higher the binding affinity. In contrast, the deactivation of the aromatic rings of residues 18 and 20 in Phe18Pff/Tyr20Pff provokes a weaker association constant value in comparison to the other peptides. It has to be considered that this latter peptide (Phe18Pff/Tyr20Pff) involves a double mutation (not only residue 18, but also residue 20). Inspection of the three-dimensional structures of the complexes for the other two peptides (with a Tyr at position 20) indicated that the hydroxyl group of this Tyr residue is not involved in either intra- or intermolecular hydrogen bonding, and thus the measured decrease of affinity for Phe18Pff/Tyr20Pff is not due to the lack of the hydrogen-bond donor character of the hydroxyl group present in Tyr20, but to deactivation of the key aromatic rings.

As mentioned above, the obtained enthalpy and entropy values for all the peptides confirmed that all the molecular recognition processes are enthalpy driven. The entropy of

binding, $\Delta S°$, was found to be negative, as also observed for a variety of chitooligosaccharides interacting with hevein itself, pseudohevein, WGA, and UDA.[11,15–21,43,44]

The binding enthalpies oscillate between −45 and −64 kJ mol^{-1}. In all cases, there is an important loss of entropy that lowers the interaction process, thus increasing the dissociation constant into the millimolar range. The higher enthalpy value is observed for Phe18Nal and the lowest for Phe18Pff/Tyr20Pff. Care should be taken when quantitatively analyzing these thermodynamic data; this observation suggests that the carbohydrate–aromatic interaction (CH/π-type interaction), typical of the hevein domains, is indeed modulated by the chemical nature and the electronic density of the aromatic ring. Although these data should not be overemphasized, some sort of enthalpy–entropy compensation also seems to occur, since the smaller observed binding enthalpy is also accompanied by a lower entropy loss.

Comparative structural analysis of the three mutant polypeptides with AcAMP2 and with hevein: The analysis of the NMR data leads to a three-dimensional perspective of the protein–saccharide complexes for the AcAMP2 mutants. The global folding, as well as the orientation of the key side chains of the AcAMP2 mutants, wild-type AcAMP2, and hevein, when bound to the $(GlcNAc)_3$ are shown in Figures 4, 6, and 8–10. When comparing the free and bound structures, the conformation of the peptide backbone is maintained when the sugar is added and is indeed similar to that deduced for the free peptides, with some slight changes in the orientation of the amino acids involved in the binding site. Indeed, the three-dimensional structures of the complexes of the three mutant peptides were very similar although the orientation of the aromatic ring in Phe18Nal is slightly different to those in the other derivatives.

As mentioned above, the comparison of the refined three-dimensional NMR structures of the mutant peptides with those structures previously reported for wild-type free AcAMP2 by NMR spectroscopy,[23] for hevein by X-ray analysis,[16] and for hevein as determined by NMR spectroscopy in our research group,[17–20] allowed the observation that the molecular topology is very similar. In fact, not only is the polypeptide backbone remarkably similar, but also the orientation of the binding site residues does not reveal significant changes, except for a restriction to motion in the bound state as depicted in Figures 8–10. According to the sMD simulations, the orientations of the key aromatic rings are different in the free and bound states. The orientations of these rings may be defined by the torsions $\chi1$ (Cα−Cβ) and $\chi2$ (Cβ−Cγ). Thus, Figure 7 shows that the aromatic residues at positions 18 and 20 may adopt a variety of values in the free state, while the presence of the sugar strongly influences the available conformational space. In fact, the trisaccharide remarkably restricts the flexibility of these aromatic rings, especially for Pff18 and for Pff20 to a lesser extent. The asymmetry of the aromatic rings in Phe18Trp and Phe18Nal already restrict the motion of these side chains (see Supporting Information S7–S9).

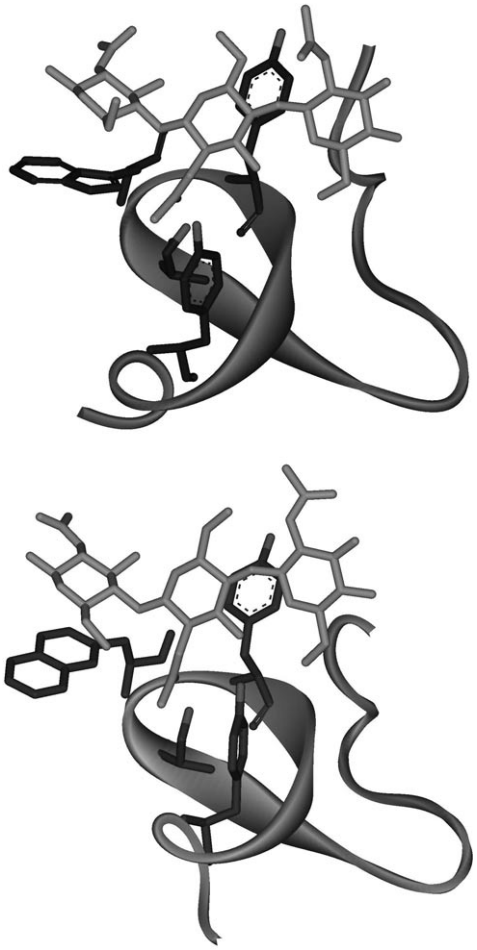

Figure 8. Schematic ribbon representation of the NMR-based structures of Phe18Trp (top) and Phe18Nal (bottom). The key aromatic and Ser16 side chains are depicted. The figure clearly shows the stacking interactions that take place for the molecular recognition of chitotriose by these AcAMP2 mutants. The interaction modes are in total agreement with the detected intermolecular peptide–sugar NOEs. The structures are indeed similar to those depicted for binding mode 2 in Figure 6, for the Phe18Pff/Tyr20Pff double mutant.

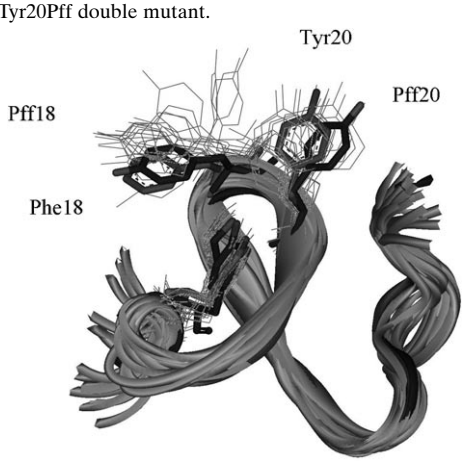

Figure 9. Ribbon representation of the superimposition of the backbone of 18 structures of free Phe18Pff/Tyr20Pff derived from molecular dynamics in explicit solvent (sMDfree). The structure obtained herein by NMR spectroscopy is also shown (gray), as well as that described by Martins et al.[24] for wild-type AcAMP2 (black). The similarity is evident. The N and C termini of the peptide chains are marked for clarity.

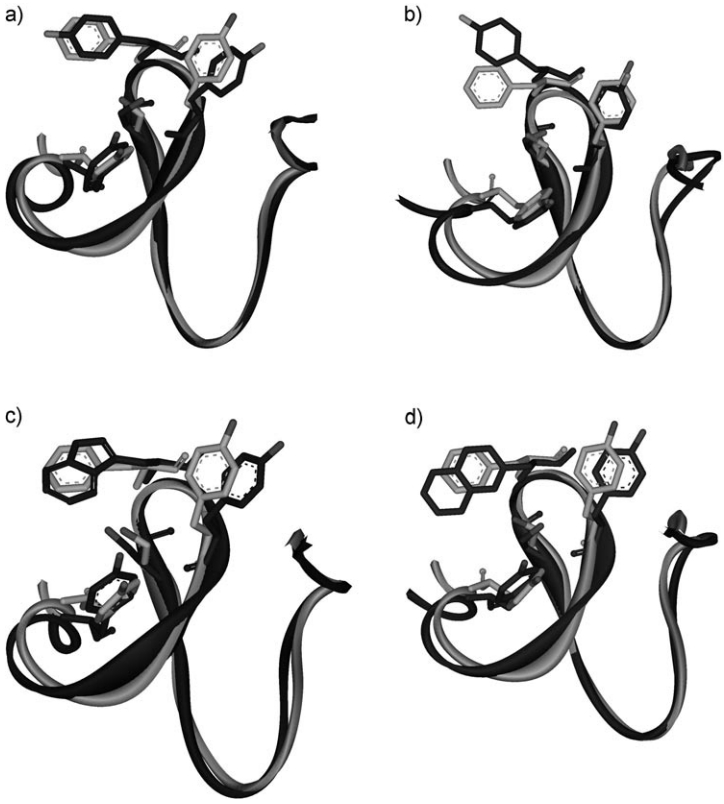

Figure 10. Schematic ribbon representation of the pairwise backbone superimposition of the NMR spectroscopic and solvated MD structures of the three mutant peptides described herein with the NMR structure described by Martins et al.[24] for wild-type AcAMP2. a) Phe18Pff/Tyr20Pff by NMR spectroscopy (black) with wild-type AcAMP2 (gray); b) Phe18Pff/Tyr20Pff by solvated MD (black) with wild-type AcAMP2 (gray); c) Phe18Trp by NMR spectroscopy (black) with wild-type AcAMP2 (gray); d) Phe18Nal by NMR spectroscopy (black) with wild-type AcAMP2 (gray).

The simulations confirm that the trisaccharide provides stabilizing interactions with the polypeptide. A combination of CH/π stacking-type interactions, accompanied by hydrogen bonding and van der Waals contacts, contribute to the maintenance of the complexes (Figures 8 and 9), which, for the double mutant, are expressed in the homogeneous orientation of the aromatic rings of the Pff18 and Pff20 moieties (Figure 9).

Regarding similarities, the superimposition of 18 structures taken from the solvated molecular dynamic trajectory (sMD_{free}, with a 4.9 ns simulation time) of Phe18Pff/Tyr20Pff with the NMR structure deduced above (and also with the NMR-based wild-type AcAMP2 structure) showed a maximum backbone rmsd of only 0.887 Å. Some differences between the sMD_{free} and the NMR structures were observed for the orientation (flexibility) of the aromatic rings of residues 18 and 20, which evidently depend on the presence or absence of the sugar (Figure 9).

When the mutant polypeptides are compared to their parent AcAMP2 structure, the pairwise backbone rmsd are small. In fact, the average NMR structures of Phe18Trp, Phe18Nal, and Phe18Pff/Tyr20Pff have rmsds with AcAMP2 of 0.984, 0.925, and 0.835 Å, respectively. A pairwise superimposition is given in Figure 10.

Concluding remarks

The interaction of AcAMP2 and the three mutants with $(GlcNAc)_3$ has been described in structural terms, making use of a NMR-derived three-dimensional structure and modeling procedures. We have shown that the binding process is enthalpically driven and that both hydrogen bonds and van der Waals forces contribute to the stability of the complexes in aqueous solution. The key role of carbohydrate–aromatic interactions is highlighted. We have demonstrated that the chemical nature of the aromatic residues involved in the protein–carbohydrate interaction strongly influences the binding affinity. Besides, the orientation and mobility of the key aromatic amino acid residues of the lectins implicated in binding depends on the presence of the trisaccharide, experiencing a significant decrease in their conformational freedom upon complexation. Therefore, the nature of the environment of the binding site is significantly important to predict the impact of changes in their constitution for binding affinity. Our data support the physiological relevance and binding ability of small native hevein domains (such as the wild-type AcAMP2) or their mutants. Chitin recognition by a large number of AcAMP2-like molecules would be expected to have a net effect on the dynamic behavior of the polymer, modifying its physical properties. We might speculate that through this mechanism, plant defense proteins, such as AcAMP2, could block the normal development of pathogen organisms, such as fungus, by interfering with the chitin biological function of protection. The role of the C terminus, absent in AcAMP2 with regard to hevein, might be related to other factors, for example, selectivity or as a functional spacer between hevein and other domains of chitinases.

Experimental Section

Source of lectins and ligand: Oligosaccharides were purchased from Toronto Research Chemicals (Canada). The synthesis of the Phe18Trp and Phe18Nal mutants have been described elsewhere.[31]

Synthesis of Phe18Pff/Tyr20Pff—materials and instruments: All Fluoren-9-ylmethoxycarbonyl (Fmoc) amino acids derivatives, including that of 4-fluorophenylalanine (4-Fphe), and the resin TentaGel S-Ram were purchased from Fluka. The side-chain protection scheme was trityl for all the 6 Cys (S-Trt) and Gln (N_δ-Trt); *tert*-butyl for Glu (5-*t*Bu), Ser (O-*t*Bu), and Tyr (O-*t*Bu); *tert*-butoxycarbonyl for Lys (N_ε-Boc), and 2,2,4,6,7-pentamethyldihydrobenzofuransulfonyl for Arg (N_δ-Pbf).

Dicyclohexylcarbodiimide (DCC), 1-hydroxybenzotriazole hydrate (HOBt) and other reagents for peptide synthesis were obtained from Aldrich, and the solvents for the HPLC analysis from Merck.

Reversed-phase high-performance liquid chromatography (RP-HPLC) was performed on a Merck Hitachi Lachrom system. For the analytical work a LiChrospher 100 RP-18 (5 μm) 250–4 column was used, with a linear gradient from 10% to 90% of acetonitrile in water (containing

0.09% TFA) over 30 min, and at flow rate of 1 mL min^{-1}. The wavelength for the detection was fixed at 214 nm.

The semipreparative separation was performed with a LiChrospher 100 RP-18 (10 μm) 250–10 column, using a gradient of 10% to 37% of acetonitrile in water (containing 0.09% TFA) over 20 min at a flow rate of 2.5 mL min^{-1}.

Mass spectrometry was performed using an ion trap mass spectrometer (Esquire 3000 Plus, Bruker) coupled to a liquid chromatograph (Agilent LC 1100 Chemstation). The ionization method was electrospray with positive ion polarity (ESI+). Peptide samples were dissolved in acetonitrile/water (2:3) containing 0.5% formic acid to a concentration of 0.3 mg mL^{-1} and infused into the ion spray using a syringe pump at a flow rate of 4 μL min^{-1} (capillary voltage −3.8 kV, dry gas temperature 200 °C).

Solid-phase synthesis: The AcAMP2 mutant was synthesized as the amide peptide by using a TentaGel S-Ram resin (0.24 mmol g^{-1}). The coupling processes were carried out manually with three equivalents of Fmoc amino acids, three equivalents of HOBt, and three equivalents of DCC in DMF during, at least, 90 min until completion, and checked with Kaiser's test. The Fmoc group in each cycle was removed by washing twice with 25% piperidine in DMF for 5 and 12.5 min, respectively. After synthesis, the resin was washed with methylene chloride and diethyl ether and dried under vacuum.

For resin cleavage, the dried resin was suspended in a fresh mixture of trifluoroacetic acid, water, phenol, and triisopropylsilane (88:5:5:2, 10 mL g^{-1} of resin) for 5–6 h. Then, the cleavage cocktail containing the peptide was filtered into chilled diethyl ether (10 volumes in relation to the cleavage cocktail), and the peptide precipitated was stored at −80 °C overnight. The crude precipitated peptide was separated by centrifugation (3500 rpm × 15 min), washed twice with cold diethyl ether and air dried. The reduced peptide was directly analyzed by HPLC in the conditions described above (retention time 12.6 min) and characterized by ESMS ($[M+2H]^{2+}$ 1604.7, $[M+3H]^{3+}$ 1070.6, $[M+4H]^{4+}$ 804.0; M_r found: 3208.87; M_r calcd: 3208.74).

Disulfide formation and purification: The polythiol peptide was subjected to an oxidative folding reaction under equilibration conditions. The reduced peptide (60 mg) was dissolved in NH_4OAc buffer (0.1 M, pH 7.8, 600 mL) in the presence of reduced (650 mg) and oxidized (128 mg) glutathione (molar ratio of peptide/GSH/GSSG, 1:100:10). After three days of air oxidation at room temperature, the reaction was stopped by acidification of the mixture with TFA to pH 2.5. The total volume of the reaction was lyophilized to 30 mL, and then directly desalted and purified by semi-preparative RP-HPLC under the conditions described above. The isomer mixture of the different folded peptides were eluted under semipreparative conditions in the interval of 18 to 25 min, but only the main peak was collected (retention time 18.6 min). This fraction was dried by lyophilization to yield 18 mg of a product that was analyzed by HPLC in the analytical conditions observing a single homogeneous polypeptide peak (retention time 10.6 min) and characterized by ESMS ($[M+2H]^{2+}$ 1602.2, $[M+3H]^{3+}$ 1069.3, $[M+4H]^{4+}$ 803.3, $[M+5H]^{5+}$ 643.5; M_r found: 3202.33; M_r calcd: 3202.69). The yield of the oxidative folding reaction was 30% over the starting material.

Titration experiments—fluorescence: Fluorescence experiments were performed with 7×10^{-6} M Phe18Trp. The stock chitotrisaccharide (Toronto Research Chemicals) concentration was calculated based on weight. Experiments were performed at 25 °C, as detailed in Aboitiz et al.[34]

Titration experiments—NMR spectroscopy: The binding of the carbohydrate to the lectins was monitored by recording one-dimensional 500 MHz ^{1}H NMR spectra of a series of samples with increasing ligand concentration (ten different concentrations) as previously described.[17] The concentration of the protein during the experiments was kept constant (ca. 0.7 mM for AcAMP2; Phe18Trp; Phe18Nal; 0.3 mM for Phe18Pff/Tyr20Pff). The samples were prepared by dissolving the lyophilized protein in buffer (1.0 mL; 85:15, $^1H_2O/^2H_2O$, 100 mM NaCl, 20 mM NaH_2PO_4, pH 5.6). The one-dimensional NMR spectrum for the sample with the highest ligand/protein ratio was recorded by dissolving the corresponding oligosaccharide (16 mM; 23 mM for Phe18Pff/Tyr20Pff) in the lectin-containing solution (0.5 mL) described above. The titration curve was established by adding small aliquots of the highest ligand/protein ratio sample to the ligand-free protein sample as previously described. Identical experiments were carried out for Phe18Pff/Tyr20Pff using a series of one-dimensional ^{19}F NMR spectra. Thermodynamic equilibrium parameters, ΔS and ΔH, for the lectin–$(GlcNAc)_3$ interaction were determined from van't Hoff plots in which the affinity constants were assessed at 25, 30, 35, and 40 °C.

NMR two-dimensional experiments: The spectra were recorded at 500 and 800 MHz. The samples for the three lectins free (2.5 mM protein) or bound to $(GlcNAc)_3$ complex [1:8 molar ratio, (ca. 0.7 mM for AcAMP2; Phe18Trp; Phe18Nal; 0.3 mM for Phe18Pff/Tyr20Pff)] were prepared in buffer (0.6 mL; 85:15 H_2O/D_2O, 100 mM NaCl, 20 mM NaH_2PO_4, pH 5.6) and degassed by passing argon. TOCSY[45] (50 ms mixing time) was performed using standard sequences at 298 K. Two-dimensional NOESY experiments[46] were performed with mixing times of 200 and 300 ms at 500 and 800 MHz, and 298 K. The NMR data have been deposited in the BMRB database (codes 6639, 6647 (Phe18Trp free and bound, respectively), 6657, 6637 (Phe18Nal free and bound, respectively), 6656 and 6591 (Phe18Pff/Tyr20Pff free and bound, respectively).

Structure calculations: The torsion angle dynamics protocol as implemented in DYANA[37] package was followed. Upper limits for proton–proton distances were obtained from NOESY cross-peak intensities at two mixing times (200 and 300 ms). Cross-peak intensities were analyzed with the XEASY software and initially classified as strong, medium, and weak corresponding to upper limits of 2.6, 3.5, and 5.5 Å, respectively. The 800 MHz cross-peaks that were not present at 500 MHz (except as overlapped) were included as 5.5 Å upper limit constraints. The CALIBA protocol was also used to obtain a better definition of the distances. The lower limit for proton–proton distances was set as the sum of the van der Waals radii of the protons. Constraints involving diastereoisomeric atoms were defined to an intermediate position (pseudoatom) and assigned an additional distance of 2.20 Å. Disulfide linkages were included as distance constraints between S–S ($2.0 < r < 2.1$ Å) and between Cβ–S ($3.0 < r < 3.1$ Å). The calculations were performed on a Linux PC computer using DYANA. A set of distance constraints derived from protein–protein NOEs (Table 2) and between 10 and 15 intermolecular distance restraints relating the lectin to the ligand were used in the final round of calculations. The standard protocol with GLOMSA was used,[34] which allowed the stereospecific assignments of 50% of the prochiral β-methylenes, followed by the application of the "distance modify" and CALIBA modules.

Between 25 to 35 best DYANA structures in terms of target function were subjected to restrained molecular dynamics[47] with the AMBER[48] force field. After an initial restrained energy minimization (REM) with 2000 conjugate gradient iterations, the structures were equilibrated at 600 K for 2 ps and, at this temperature, their conformational behavior for the next 2 ps was simulated by restrained molecular dynamics (RMD). In the next step, the structures were subjected to a cooling regime, in which the temperature was decreased by 100 K every 2 ps until a temperature of 100 K was reached. At this temperature, 4 ps of RMD calculations were carried out. The final structures were energy-minimized (REM) by using 2000 conjugate gradient iterations. The refined structures can be found in the Protein DataBank (PDB ID: 1ZUV, 1ZWU, and 1ZNT, for the Phe18Trp, Phe18Nal, and Phe18Pff/Tyr20Pff, respectively).

Molecular mechanics and dynamics calculations: Φ is defined as H-1″-C-1″-O-C4′/H-1′-C-1′-O-C4 and Ψ as C1″-O-C4′-H4′/C1′-O-C4-H4 for the nonreducing–middle and middle–reducing disaccharide entities, respectively. The MD calculations for the complexes were performed using the AMBER 5.0 package.[49]

For Phe18Nal and Phe18Pff/Tyr20Pff, it was necessary to build up the corresponding library for the non-natural naphtylalanine and Pff residues. As a first step, the corresponding structures were generated and minimized with the MM2 force field. Then, the charges were estimated using Hartree Fock 6-31G calculations with Gaussian 98.[50] Then, the MOLMOL program[51] was used to generate the DYANA and XEASY libraries and the X-LEAP[52] module to build up the library for AMBER.

The starting glycosidic torsion angles were taken from previously described MM3* calculations performed for $(GlcNAc)_2$ and $(GlcNAc)_3$.[20,21] The global minimum for the trisaccharide was used as starting structure, since it has been shown that it is very similar to its hevein-bound conformation (Φ/Ψ 55:5 for both glycosidic linkages).[20,21] For the protein coordinates, the input files were prepared from the NMR-derived structures by using the X-LEAP module of the AMBER package. The obtained initial structures were immersed into a box of 3017 TIP3P water molecules in order to obtain accurate solvation. Cut-offs for nonbonding interactions was set to 11.0 Å. The molecular dynamics simulations were carried out by using the Sander module and were performed using periodic boundary conditions and the particle-mesh Ewald approach to introduce long-range electrostatic effects. The SHAKE algorithm for hydrogen atoms, which used a 2 fs time step, was employed.

All solvated simulations were performed at constant pressure and temperature, using the Berendsen coupling algorithm for the latter. Equilibration of the system was carried out as follows; as a first step, a short minimization with positional restraints on solute atoms was run to remove any potentially bad contacts. The force constant for the positional constraints was 500 kcal mol^{-1} Å. Next a 12.5 ps molecular dynamics calculation was run at 300 K maintaining positional restraints on the ligand in order to equilibrate the water box and ions. A 9 Å cut-off was used for the treatment of the electrostatic interactions. The system was further equilibrated (12.5 ps run at 300 K) by using the mesh Ewald approach for long-range electrostatic effects. Then, the system was subjected to several minimization cycles (each using 1000 steepest descent iterations) gradually reducing positional restraints on the chitooligosaccharides from 500 kcal mol^{-1} Å to 0. Finally, MD trajectories at constant pressure (1 atm) and temperature (300 K) were collected and analyzed using the Carnal module of AMBER. Structures were recorded every 0.5 ps, for a total calculation time around 4.5 ns in all cases. A deeper discussion of the results of the sMD results will be presented separately.

Acknowledgments

Funding from the European Union (SACCSIGNET, HPRN-CT2002–00251), Ministry of Education and Science of Spain (grant BQU2003–3550-C03), and Comunidad de Madrid (grant 08.2/0035.1/2000) is gratefully acknowledged. N.A. also thanks Comunidad de Madrid for financial support. We also thank the CAI of NMR of Universidad Complutense of Madrid and the Parc Cientific of Barcelona for access to the NMR spectrometers. I.C. thanks UNAM of Mexico for a leave of absence. C.A. and G.A. thank Generalitat Valenciana (Grupos 03/120) for funding. M.M. thanks Ministry of Economy, Trade and Industry of Japan for financial support.

[1] H.-J. Gabius, H.-C. Siebert, S. Andre, J. Jimenez-Barbero, H. Rudiger, *ChemBioChem* **2004**, *5*, 740–764.
[2] H. Rudiger, H.-C. Siebert, D. Solis, J. Jimenez-Barbero, A. Romero, C. W. von der Lieth, T. Diaz-Marino, H.-J. Gabius, *Curr. Med. Chem.* **2000**, *7*, 389–416.
[3] a) A. Imberty, S. Perez, *Chem. Rev.* **2000**, *100*, 4567–4588; b) H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637–674.
[4] a) J. Jimenez-Barbero, J. L. Asensio, F. J. Canada, A. Poveda, *Curr. Opin. Struct. Biol.* **1999**, *9*, 549–555; b) H. Kogelberg, D. Solis, J. Jimenez-Barbero, *Curr. Opin. Struct. Biol.* **2003**, *13*, 646–653.
[5] M. Muraki, *Protein Pept. Lett.* **2002**, *9*, 195–209.
[6] T. Weimar, R. J. Woods, in *NMR Spectroscopy of Glycoconjugates* (Eds.: J. Jimenez-Barbero, T. Peters), Wiley-VCH, Weinheim, **2002**, pp. 111–114.
[7] E. J. M. Van Damme, W. J. Peumans, A. Barre, P. Rouge, *Crit. Rev. Plant Sci.* **1998**, *17*, 575–692.
[8] a) W. J. Peumans, E. J. M. van Damme, *Histochem. J.* **1995**, *27*, 253–271; b) F. Garcia-Olmedo, A. Molina, J. M. Alamillo, P. Rodriguez-Palenzuela, *Biopolymers* **1998**, *47*, 479–491.
[9] J. J. Beintema, *FEBS Lett.* **1994**, *350*, 159–163.
[10] a) C. S. Wright, *J. Biol. Chem.* **1992**, *267*, 14345–14352; b) C. S. Wright, G. E. Kellogg, *Protein Sci.* **1996**, *5*, 1466–1476.
[11] G. Bains, R. T. Lee, Y. C. Lee, E. Freire, *Biochemistry* **1992**, *31*, 12624–12628.
[12] J. F. Espinosa, J. L. Asensio, J. L. Garcia, J. Laynez, M. Bruix, C. Wright, H. C. Siebert, H. J. Gabius, F. J. Canada, J. Jimenez-Barbero, *Eur. J. Biochem.* **2000**, *267*, 3965–3978.
[13] F. A. Saul, P. Rovira, G. Boulot, E. J. Damme, W. J. Peumans, P. Truffa-Bachi, G. A. Bentley, *Structure* **2000**, *8*, 593–603.
[14] K. Harata, M. Muraki, *J. Mol. Biol.* **2000**, *297*, 673–681.
[15] A. Rodriguez-Romero, K. G. Ravichandran, M. Soriano-Garcia, *FEBS Lett.* **1991**, *291*, 307–339.
[16] C. A. Reyes-Lopez, A. Hernandez-Santoyo, M. Pedraza-Escalona, G. Mendoza, A. Hernandez-Arana, A. Rodriguez-Romero, *Biochem. Biophys. Res. Commun.* **2004**, *314*, 123–130.
[17] N. H. Andersen, B. Cao, A. Rodriguez-Romero, B. Arreguin, *Biochemistry* **1993**, *32*, 1407–1422.
[18] J. L. Asensio, F. J. Canada, M. Bruix, A. Rodriguez-Romero, J. Jimenez-Barbero, *Eur. J. Biochem.* **1995**, *230*, 621–633.
[19] a) J. L. Asensio, F. J. Canada, M. Bruix, C. Gonzalez, N. Khiar, A. Rodriguez-Romero, J. Jimenez-Barbero, *Glycobiology* **1998**, *8*, 569–577; b) H. C. Siebert, S. Andre, J. L. Asensio, F. J. Canada, X. Dong, J. F. Espinosa, M. Frank, M. Gilleron, H. Kaltner, T. Kozar, N. V. Bovin, C. W. von Der Lieth, J. F. Vliegenthart, J. Jimenez-Barbero, H. J. Gabius, *ChemBioChem* **2000**, *1*, 181–195.
[20] J. L. Asensio, F. J. Canada, H. C. Siebert, J. Laynez, A. Poveda, P. M. Nieto, U. M. Soedjanaamadja, H. J. Gabius, J. Jimenez-Barbero, *Chem. Biol.* **2000**, *7*, 529–543.
[21] J. L. Asensio, H. C. Siebert, C. W. von Der Lieth, J. Laynez, M. Bruix, U. M. Soedjanaamadja, J. J. Beintema, F. J. Canada, H. J. Gabius, J. Jimenez-Barbero, *Proteins* **2000**, *40*, 218–236.
[22] P. Karisola, J. Mikkola, N. Kalkkinen, K. J. Airenne, O. H. Laitinen, S. Repo, O. T. Pentikainen, T. Reunala, K. Turjanmaa, M. S. Johnson, T. Palosuo, M. S. Kulomaa, H. Alenius, *J. Immunol.* **2004**, *172*, 2621–2628.
[23] J. C. Martins, D. Maes, R. Loris, H. A. Pepermans, L. Wyns, R. Willem, P. Verheyden, *J. Mol. Biol.* **1996**, *258*, 322–333.
[24] P. Verheyden, J. Pletinckx, D. Maes, H. A. Pepermans, L. Wyns, R. Willem, J. C. Martins, *FEBS Lett.* **1995**, *370*, 245–249.
[25] R. U. Lemieux, *Acc. Chem. Res.* **1996**, *29*, 373–380.
[26] J. P. Carver, *Pure Appl. Chem.* **1993**, *65*, 763–770.
[27] T. K. Dam, F. Brewer, *Chem. Rev.* **2002**, *102*, 387–429.
[28] M. R. Wormald, A. J. Petrescu, Y. L. Pao, A. Glithero, T. Elliott, R. A. Dwek, *Chem. Rev.* **2002**, *102*, 371–386.
[29] R. Woods, *Glycoconj. J.* **1998**, *15*, 209–216.
[30] a) H.-J. Gabius, *Pharm. Res.* **1998**, *15*, 23–30; b) C. von der Lieth, H. Siebert, T. Kozar, M. Burchert, M. Frank, M. Gilleron, H. Kaltner, G. Kayser, E. Tajkhorshid, N. V. Bovin, J. F. Vliegenthart, H. J. Gabius, *Acta Anat.* **1998**, *161*, 91–109.
[31] M. Muraki, H. Morii, K. Harata, *Protein Pept. Lett.* **1998**, *5*, 193–198.
[32] M. Muraki, M. Morii, K. Harata, *Protein Eng.* **2000**, *13*, 385–389.
[33] a) D. Andreu, E. Nicolas, in *Solid-Phase Synthesis* (Eds.: S. A. Kates, F. Albericio), Marcel Dekker, New York, **2000**, pp. 365–375; b) M. Vila-Perello, A. Sanchez-Vallet, F. Garcia-Olmedo, A. Molina, D. Andreu, *FEBS Lett.* **2003**, *536*, 215–219.
[34] N. Aboitiz, M. Vila-Perello, P. Groves, J. L. Asensio, D. Andreu, F. J. Canada, J. Jimenez-Barbero, *Chembiochem*, **2004**, *5*, 1245–1255.
[35] J. P. Privat, F. Delmotte, G. Mialonier, P. Bouchard, M. Monsigny, *Eur. J. Biochem.* **1974**, *47*, 5–14.
[36] L. Fielding, *Curr. Top. Med. Chem.* **2003**, *3*, 39–53.
[37] a) P. Guntert, W. Braun, K. Wuthrich, *J. Mol. Biol.* **1991**, *217*, 517–530; b) P. Guntert, C. Mumenthaler, K. Wuthrich, *J. Mol. Biol.* **1997**, *273*, 283–298.
[38] A. Germer, C. Mugge, M. G. Peter, A. Rottmann, E. Kleinpeter, *Chem. Eur. J.* **2003**, *9*, 1964–1973.
[39] M. G. Ford, T. Weimar, T. Kohli, R. J. Woods, *Proteins* **2003**, *53*, 229–240.

[40] C. S. A. T. Darden, *Annu. Rev. Biophys. Biomol. Struct.* **1999**, *28*, 155.
[41] K. A. Kronis, J. P. Carver, *Biochemistry* **1985**, *24*, 834–840.
[42] a) J. P. Carver, S. W. Michnik, A. Imberty, D. A. Cumming, *Ciba Found. Symp.* **1989**, *145*, 6–26; b) F. A. Quiocho, *Pure Appl. Chem.* **1989**, *61*, 1293–1306.
[43] E. Garcia-Hernandez, R. A. Zubillaga, A. Rojo-Dominguez, A. Rodriguez-Romero, A. Hernandez-Arana, *Proteins* **1997**, *29*, 467–477.
[44] R. T. Lee, H. J. Gabius, Y. C. Lee, *Glycoconj. J.* **1998**, *15*, 649–655.
[45] A. Bax, D. G. Davis, *J. Magn. Reson.* **1985**, *65*, 355–360.
[46] A. Kumar, R. R. Ernst, K. Wuthrich, *Biochem. Biophys. Res. Commun.* **1980**, *95*, 1–6.
[47] R. M. Scheek, W. F. van Gunsteren, R. Kaptein, *Methods Enzymol.* **1989**, *177*, 204–218.
[48] S. J. Weiner, P. A. Kollman, D. A. Case, U. C. Singh, C. Ghio, G. Alagona, S. Profeta, P. Weiner, *J. Am. Chem. Soc.* **1984**, *106*, 765–784.
[49] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham, S. Debolt, D. Ferguson, G. Seibel, P. Kollman, *Comput. Phys. Commun.* **1995**, *91*, 1–41.
[50] Gaussian 98 (Revision A.11.3), M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, V. G. Zakrzewski, J. A. Montgomery, R. E. Stratmann, J. C. Burant, S. Dapprich, J. M. Millam, A. D. Daniels, K. N. Kudin, M. C. Strain, O. Farkas, J. Tomasi, V. Barone, M. Cossi, R. Cammi, B. Mennucci, C. Pomelli, C. Adamo, S. Clifford, J. Ochterski, G. A. Petersson, P. Y. Ayala, Q. Cui, K. Morokuma, D. K. Malick, A. D. Rabuck, K. Raghavachari, J. B. Foresman, J. Cioslowski, J. V. Ortiz, B. B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R. Gomperts, R. L. Martin, D. J. Fox, T. Keith, M. A. Al-Laham, C. Y. Peng, A. Nanayakkara, C. Gonzalez, M. Challacombe, P. M. W. Gill, B. G. Johnson, W. Chen, M. W. Wong, J. L. Andres, M. Head-Gordon, E. S. Replogle, J. A. Pople, Gaussian, Inc., Pittsburgh, PA, **1998**.
[51] R. Koradi, M. Billeter, K. Wüthrich, *J. Mol. Graphics* **1996**, *14*, 51–55.
[52] C. E. A. F. Schafmeister, W. F. Ross, V. Romanovsky, University of California, San Francisco, **1995**.

Received: April 1, 2005
Revised: August 19, 2005
Published online: October 12, 2005

