

Molecular Recognition of β -O-GlcNAc Glycopeptides by a Lectin-Like Receptor: Binding Modulation by the Underlying Ser or Thr Amino Acids

Francisco Corzana,^[a] Alberto Fernández-Tejada,^[a] Jesús H. Busto,^[a] Gururaj Joshi,^[b] Anthony P. Davis,^{*[b]} Jesús Jiménez-Barbero,^[c] Alberto Avenoza,^[a] and Jesús M. Peregrina^{*[a]}

The binding properties of different carbohydrates and glycopeptides containing the β -O-2-deoxy-2-(N-acetyl)-D-glucosaminyl (β -O-GlcNAc) to a synthetically prepared lectin-like receptor have been analyzed. The study combines the use of NMR spectroscopy experiments with extensive MD simulations in explicit water. Notably, the presence of a key hydrogen bond between the receptor and the OMe group of the β -O-GlcNAc-OMe derivative appears to be responsible for the high selectivity observed for this compound. In addition, to study the effect on the binding of the underlying amino acid, we

have prepared different model glycopeptides, which include the non-natural α -methylserine and α -methylthreonine as underlying amino acids. Interestingly, the presence of a methyl group decreases the affinity constant, especially in those cases in which a β -methyl group is present. As a result, the serine-containing glycopeptide exhibited the highest affinity constant of the glycopeptides, and the threonine derivative showed the lowest one. This low selectivity could have its origin in the difficulty to form both specific hydrogen bonds and hydrophobic (CH- π) contacts.

Introduction

β -O-2-Deoxy-2-(N-acetyl)-D-glucosaminyl (β -O-GlcNAc) moiety is a common motif in biological chemistry^[1] (Scheme 1 A). It plays a unique role in protein regulation through linkage to the hydroxy group of serine or threonine. This post-translational^[1c] modification is highly dynamic^[2] and draws comparisons with protein phosphorylation as a biological control mechanism. It has been implicated in gene transcription, nuclear trafficking, protein translation,^[3] signal transduction,^[1a] the regulation of protein–protein interactions,^[1] and the sensing of nutritional levels within the cell.^[4] Furthermore, there is clear evidence that the aberrant O-GlcNAc modification of proteins is correlated with diabetes, tumorigenesis, and even with Alzheimer's disease.^[1,5]

Bearing in mind the singular features of β -O-GlcNAc, it is crucial to know in detail the interaction of this carbohydrate moiety with its biological targets. On this basis, some of us^[6] have recently reported on the synthesis of a simple lectin-like receptor (Scheme 1 B) that is particularly effective for this substrate. In fact, the receptor binds **1** (a model for β -O-GlcNAc) with an affinity constant (K_a) of 630 M^{-1} , very similar to that shown by the lectin wheat germ agglutinin. The receptor is also highly selective for the β -anomer. Indeed, binding to related carbohydrates is significantly weaker (e.g., K_a for **2** and α -anomer **3** are 156 and 24 M^{-1} , respectively) whereas affinities for other monosaccharides are lower still (e.g., glucose 9 M^{-1} ; xylose 5 M^{-1} , galactose 2 M^{-1} and N-acetylgalactosamine 2 M^{-1}). This system presents an opportunity to study the recognition of β -O-GlcNAc by a binding site that, though lectin-like, avoids the size and complexity of a protein scaffold. Here we present molecular dynamics and NMR spectroscopic investigations that

explore the role of aglycon structure in the molecular recognition of β -O-GlcNAc.

Results and Discussion

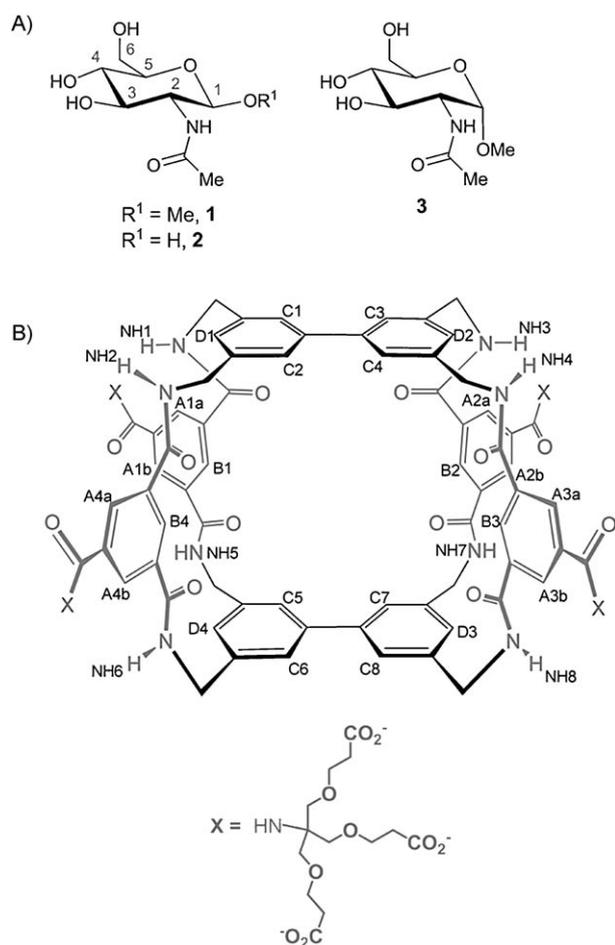
To gain insights into the receptor–ligand interactions that play a pivotal role in the strong selectivity observed for compound **1**, we decided to run 80 ns MD simulations with time-averaged restraints (MD-tar)^[7] on the receptor–**1** complex by using all the experimental distances^[6] as restraints and by following our previously reported protocol.^[8] Figure 1 A shows 15 snapshots taken from these MD-tar simulations. In this particular case, the calculations were carried out without the externally directed, water-solubilizing side chains. The obtained structure was fairly similar to the NOE-based conformation previously reported. Seven conventional intramolecular hydrogen bonds were detected throughout the simulations. The oxygen of the NHAc is involved in a hydrogen bond with NH5, and its NH group in-

[a] Dr. F. Corzana, Dr. A. Fernández-Tejada, Dr. J. H. Busto, Prof. A. Avenoza, Prof. J. M. Peregrina
Departamento de Química, Universidad de La Rioja, UA-CSIC
26006 Logroño (Spain)
Fax: (+34) 941299621
E-mail: jesusmanuel.peregrina@unirioja.es

[b] Dr. G. Joshi, Prof. A. P. Davis
School of Chemistry, University of Bristol
Cantock's Close, Bristol, BS8 1TS (UK)

[c] Prof. J. Jiménez-Barbero
Centro de Investigaciones Biológicas (CSIC)
Ramiro Maeztu 9, 28040 Madrid (Spain)

Supporting information for this article is available on the WWW under <http://dx.doi.org/10.1002/cbic.201000526>.



Scheme 1. A) Carbohydrates studied in this work. B) Structure of the receptor used in this study, showing the labeling system used throughout the text.

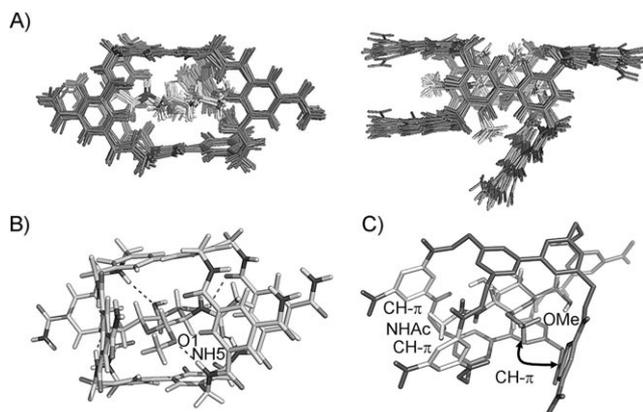


Figure 1. A) Different orientations of the ensemble obtained from the MD-tar simulations for complex receptor:1. The resulting structure reproduces all the NOE-derived distances reported by some of us (ref. [6]). B) Schematic representation of the different hydrogen bonds obtained from the MD-tar simulations in the receptor–1 complex. C) Schematic representation of the hydrophobic interactions between ligand 1 and the receptor. The hydrogen atoms have been removed for clarity.

teracts with the carbonyl group NH1_C=O. On the other hand, O6 of the sugar participates in two hydrogen bonds: O6...NH3 and O6...NH7. Moreover, O4 interacts with NH4 and O3...H with

the carbonyl group NH2_C=O. Finally, an important hydrogen bond was present between the β -glycosidic oxygen (O1) and NH5. These hydrogen bonds are summarized in Figure 1B. On the other hand, hydrophobic contacts were observed between the methyl group of NHAc and the two spacer aromatic rings. The averaged distance between the carbon atom of the methyl group and the centers of the aromatic rings was calculated to be 3.6 Å. Notably, an additional hydrophobic contact was detected between the methoxy group and one spacer aromatic ring. As in the former cases, the distance between the ring and the methyl group was 3.6 Å (Figure 1C). Experimental evidence for this CH- π interaction^[9] is the perturbation of the chemical shift of the OMe group in the bound state with respect to the free sugar. In addition, from a theoretical point of view, when ligand 1 is removed from the complex, the side chain attached to the spacer involved in the CH- π interactions becomes significantly more flexible (see the Supporting Information).

To check the reliability of the above-mentioned MD simulations on these systems, we also ran 25 ns unrestrained MD simulations on the receptor–1 complex in explicit water. Now all the receptor atoms, including the water-solubilizing side chains and the counter ions, were included in the calculations. Strikingly, the distances derived from these simulations were similar to the experimental ones. Moreover, the same interactions between the receptor and compound 1 were detected, which validates the unrestrained MD simulations.

Encouraged by these results, we decided to run unrestrained MD simulations on the complexes formed between the receptor and compounds 2 and 3. As before, the hydrogen bonds observed in the complex with 1 were also found for these substrates. However, in sharp contrast with 1, the population of the hydrogen bond NH5...O1 dropped to only 20% in compound 2 and, moreover, was never populated for 3 (with O1 axially oriented, Figure 2). Considering that the rest of the stabilizing interactions have a similar weight in the three complexes, and assuming that a typical NH...O hydrogen^[10] bond stabilizes around 1.9 kcal mol⁻¹, it could be deduced that the K_a of compound 1 should be around 25 times the affinity constant of derivative 3, which is in excellent agreement with the experimental data commented on above. On the other hand, the low population of this hydrogen bond obtained for the receptor–3 complex indicates that this interaction, with the proper orientation of NH5, plays the key role in the β versus α selectivity of this receptor.

Taking into account the strong selectivity of the receptor for compound 1 and the fact that β -O-GlcNAc appears in most cases linked to a serine or a threonine residue; we decided to use this simple lectin-like receptor to investigate the effect that the underlying amino acid has on the binding properties. On this basis, and given that the incorporation of unnatural residues in the backbone of small glycopeptides can stabilize conformations present in naturally occurring molecules or exhibit some atypical conformations,^[11] we report herein the binding properties of different model glycopeptides derived not only from natural serine and threonine (compounds 4 and 5, respectively) but also derived from the unnatural α -methyl-

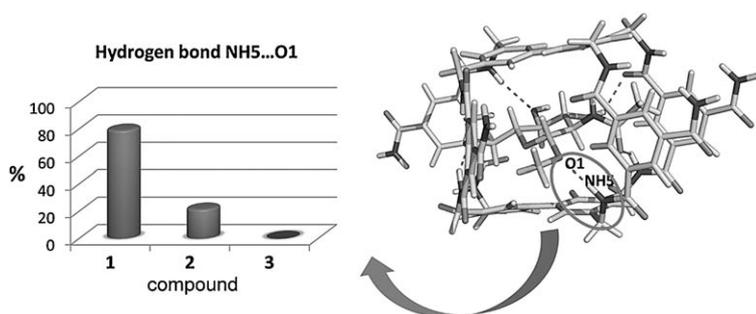
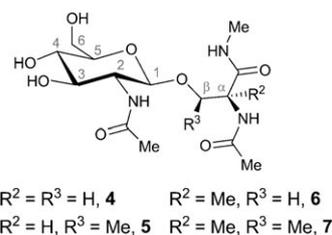


Figure 2. Population of the hydrogen bond NH5...O1 between the receptor and carbohydrates 1, 2, and 3, as obtained from the 25 ns unrestrained MD simulations in explicit water.

serine (MeSer) and α -methylthreonine (MeThr), compounds 6 and 7 (Scheme 2). Thus, the main goal of this study was to investigate in detail the role that the substituents at the C α and/or C β atoms have in the glycopeptide conformations and, consequently, on their binding properties. In these compounds, the amino and carboxylic acid functional groups were transformed into amides to simulate the peptide backbone.

Prior to carrying out the conformational study of the glycopeptides in the bound state, it is important to know the conformation of the glycopeptides in water (free state). In this context, the conformational behavior of compounds 4 and 5 has been previously reported.^[12]

In this case, the main conclusion is that although these compounds have a similar behavior for the peptide backbone, represented by extended conformations, they have different behavior in terms of the ψ_s (C1-O1-C β -C α) dihedral angle. In fact, for derivative 5, this angle showed values around 120°, resulting in an eclipsed conformation for the H β -C β and O1-C1 bonds. This defined conformation can accommodate a water pocket between the peptide and the carbohydrate moieties. Additionally, the lateral chain adopts a *g*(+) conformation ($\chi^1 = 60^\circ$) in both derivatives. On the other hand, the conformational analysis of glycopeptides 6 and 7 was carried out by following our methodology,^[8] which combines NOE-derived distances with MD-tar simulations. The results obtained for these two novel molecules are summarized in Figure 3. Our data show that the peptide backbone of both derivatives 6 and 7 adopted mainly helix-like conformations, which are typical for glycopeptides, which incorporate α,α -disubstituted amino acids. This result is in good agreement with the weak-to-medium NOE observed between the NH protons of the backbone (Supporting Information). As far as the lateral chain (χ^1 torsion angle) is concerned, the simulations suggest that the rotation around χ^1 in the MeThr derivative is restricted with values for this torsion angle close to -60° . In sharp contrast, the lateral chain of the MeSer-containing glycopeptide is rather flexible, showing significant population of each staggered conformation for this torsion angle. Concerning the glycosidic linkage, they also showed a different behavior in terms of the ψ_s dihedral angle. In fact, for 7, this angle showed values around 120°–140°. This result is similar to that obtained for glycopeptide 5, which also possesses a β -methyl group. However, ψ_s was found to be



Scheme 2. Glycopeptides studied in this work, showing the labeling system used throughout the text.

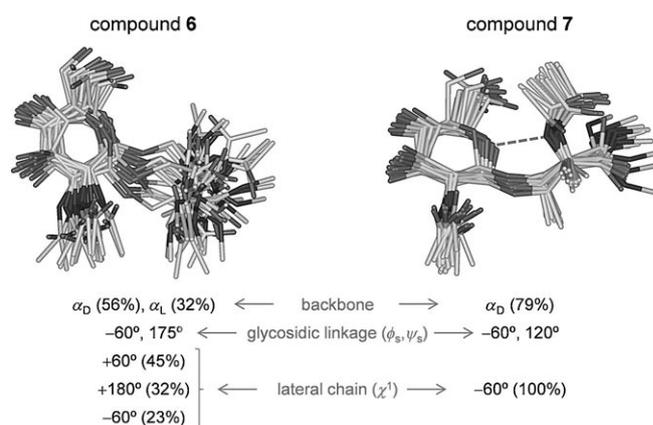


Figure 3. Ensembles obtained from the MD-tar simulations in explicit water for model glycopeptides 6 and 7 together with detailed information about their major conformations found in the free state for the most relevant torsional angles.

much more flexible for the MeSer analogue 6, which shows a major *anti*-type arrangement for the GlcNAc residue and the peptide moiety. Curiously, although the anomeric linkage and the chemical nature of the sugar is completely different, it is noteworthy that an eclipsed ψ_s conformation has been previously reported by us for glycopeptides containing the α -O-GalNAc-Thr unit.^[13] This similar behavior could indicate that the conformational tendencies for ψ_s in O-glycopeptides are mainly determined by the presence (or not) of the β -methyl group in the β -hydroxy- α -amino acid residue (Thr or Ser, respectively), irrespective of the chemical nature of the carbohydrate moiety and of the configuration at the anomeric center.

From the interactions perspective, the different features observed in the glycopeptides could have important implications for molecular recognition processes. To test this hypothesis, we studied the binding of compounds 4–7 to the receptor in D₂O with ¹H NMR spectroscopic titrations, following the methodology previously described.^[6]

As can be observed in Figure 4, when nine equivalents of the glycopeptides 4–7 were added to a solution of the receptor, the signals of this molecule became more complicated, which is consistent with the loss of symmetry upon binding.

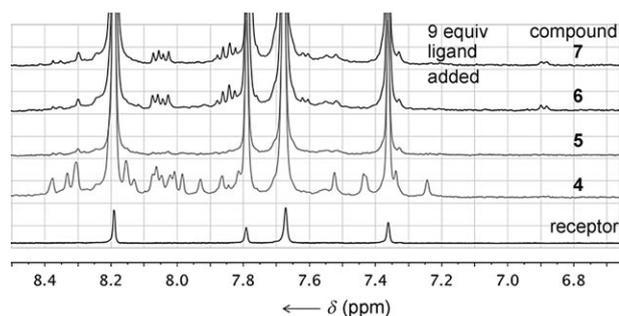


Figure 4. ^1H NMR spectra of the receptor (0.5 mM) in D_2O after addition of nine equivalents of the different glycopeptides. Only aromatic signals of the receptor are shown.

The association constants were measurable through the integral ratios of the bound and unbound receptor signals. The K_a and ΔG values obtained for the different glycopeptides are summarized in Table 1.

Table 1. Association constants (K_a) and ΔG values in aqueous solution for the receptor shown in Scheme 1 B with different substrates.				
	Name	Compound	K_a [M^{-1}]	ΔG [kcal mol^{-1}]
1	$\beta\text{GlcNAcOMe}$	1	630 ^[6]	3.82
2	GlcNAc	2	156 ^[6]	2.99
3	$\alpha\text{GlcNAcOMe}$	3	24 ^[6]	1.88
4	$\beta\text{GlcNAcSer}$	4	91	2.67
5	$\beta\text{GlcNAcThr}$	5	13	1.52
6	$\beta\text{GlcNAcMeSer}$	6	75	2.56
7	$\beta\text{GlcNAcMeThr}$	7	46	2.27

Important conclusions can be drawn from the inspection of these K_a values. Firstly, the presence of the amino acid reduces the affinity constant significantly. Secondly, the presence of α and/or β -methyl groups at the amino acid moiety appears to have negative effects on the binding. Especially relevant is the small K_a value obtained for the Thr-containing glycopeptides. In fact, the two glycopeptides with a β -methyl group (compounds 5 and 7) exhibited the lower K_a values. According to these values, the complex with glycopeptide 4 is, at 298 K, about $1.1 \text{ kcal mol}^{-1}$ more stable than the complex with the Thr derivative 5. Finally, it is important to note that the pattern shown by the complex with ligand 4 (similar to that found for compound 1) differs from that observed for the rest of the glycopeptides, which could indicate that the methyl groups accommodate the sugar moiety in a different way.

To explain these results, we ran extensive MD simulations on the complexes between glycopeptides 4–7 and the receptor. Figure 5 shows 15 snapshots taken from these simulations for each complex.

As shown in Figure 6, the helix-like conformation was particularly populated in the bound state for glycopeptides 5 and 7 (compound with a β -methyl group); this conformer coexists with the G conformation^[14] in derivatives 4 and 6. The helix-

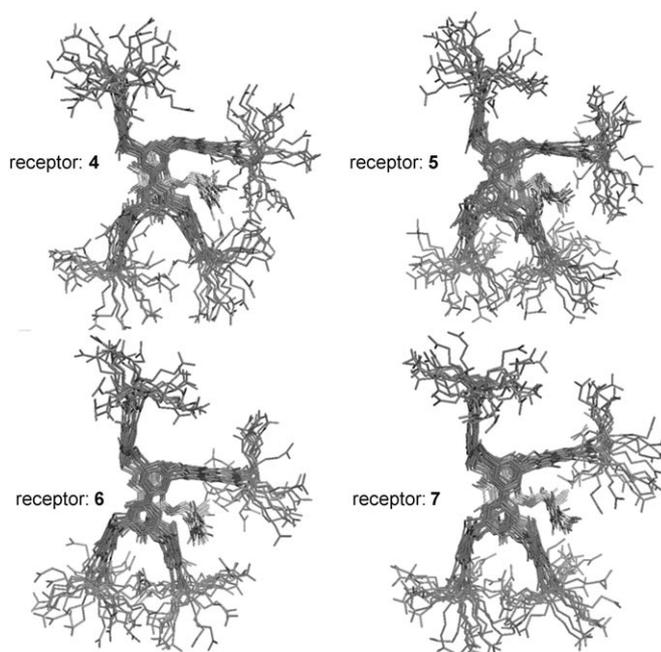


Figure 5. Ensembles obtained from the 25 ns unrestrained MD simulations in explicit water for the complexes between the receptor and glycopeptides 4–7.

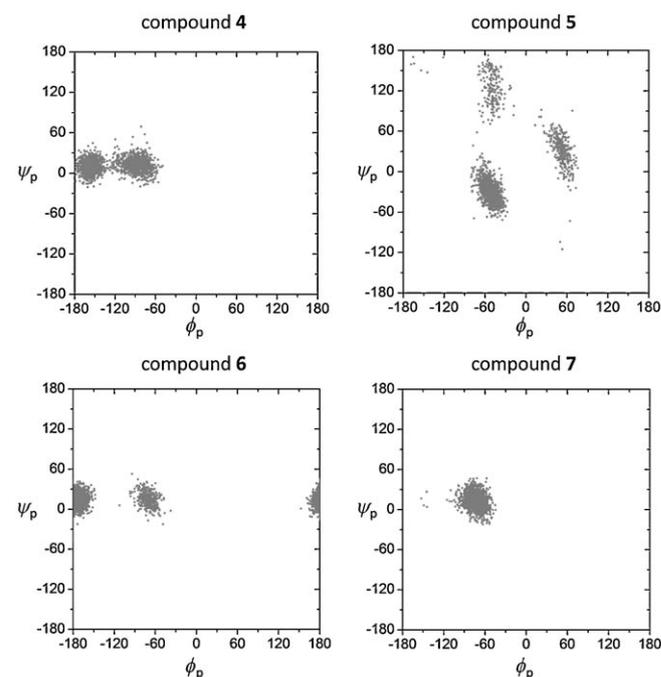


Figure 6. ϕ_p/ψ_p distributions (backbone) obtained from the 25 ns unrestrained MD simulations in explicit water for the bound state of the glycopeptides.

like conformation allows the formation of a weak hydrogen bond between the amino group of the amino acid residues and the carbonyl group $\text{NH1}_C=\text{O}$ of the receptor.

Concerning the glycosidic linkage (Figure 7), glycopeptides 4 and 6 adopted an alternate conformation, which is expected

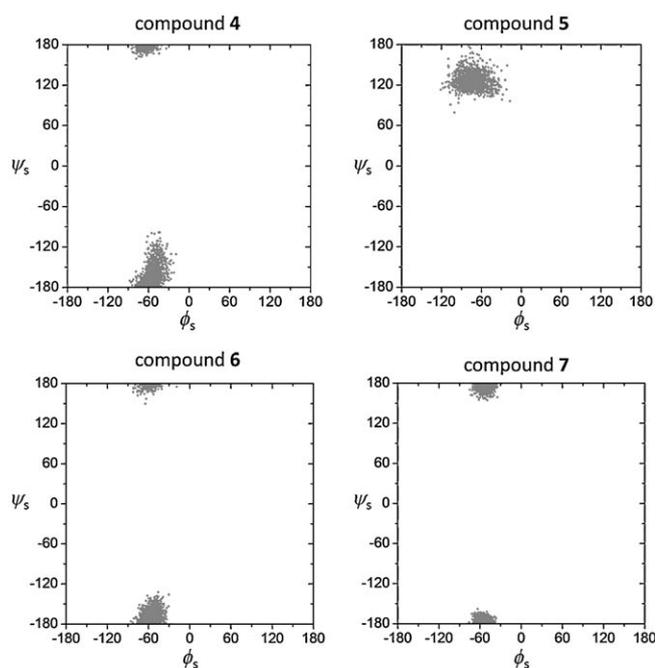


Figure 7. ϕ_s/ψ_s distributions (glycosidic linkage) obtained from the 25 ns unrestrained MD simulations in explicit water for the bound state of the glycopeptides.

for derivatives without a β -methyl group at the backbone.^[12,13] Moreover, glycopeptide **5** showed the typical eclipsed conformation for the glycosidic linkage. Strikingly, and in contrast to the result obtained for the free state, the alternate conformation was also the most populated for derivative **7** in the bound state.

As a next step, the intermolecular hydrogen-bonding analysis for all the complexes was carried out. The most persistent hydrogen bonds between the sugar moiety and the receptor are shown in Figure 8. The population of the hydrogen bond $\text{NH5}\cdots\text{O1}$ considerably differs from that observed for compound **1**; its occupancy was only around 25% in the best case (compound **4**). In particular, this hydrogen bond was very weak ($\leq 5\%$ occupancy) in compounds with a methyl group at $\text{C}\beta$ (derivatives **5** and **7**). This methyl group also inhibits the

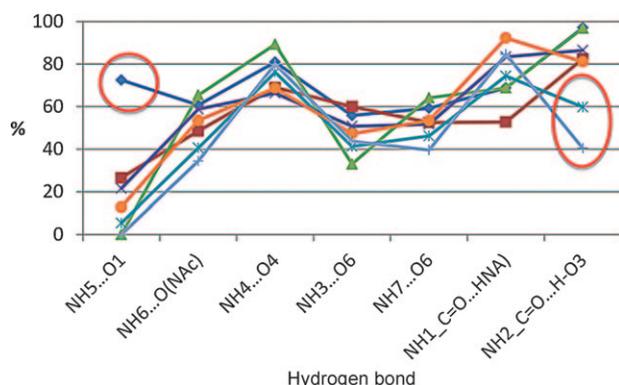


Figure 8. Occupancy of the most significant hydrogen bonds obtained from the 25 ns unrestrained MD simulations for the bound states of compounds **1** (\blacklozenge), **2** (\blacksquare), **3** (\blacktriangle), **4** (\times), **5** ($*$), **6** (\bullet), and **7** ($+$).

formation of a hydrogen bond between the carbonyl group $\text{NH2}_\text{C}=\text{O}$ of the receptor and O3-H of the ligand. These two features could explain the observed very low affinity of the receptor for glycopeptides **5** and **7**.

We guessed that the presence of the methyl groups in the amino acid moiety could also modulate the ligand–receptor interactions and, consequently, the affinity constant. As commented on above, there is one hydrophobic contact ($\text{CH}-\pi$ interaction) that involves the OMe group of compound **1** and the aromatic ring of one of the side chains (Figure 1C). On this basis, Figure 9 shows the distances between the different car-

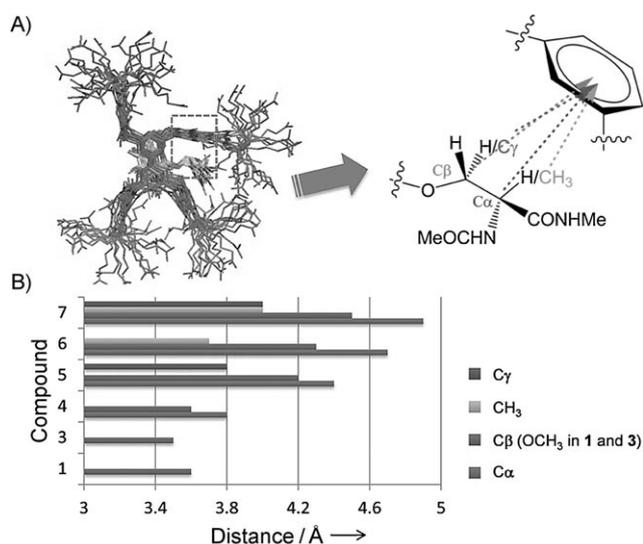


Figure 9. A) Left panel) Superimposition of 10 conformers randomly taken from the MD simulations for the complex with ligand **4**, showing the $\text{CH}-\pi$ interaction of the backbone with the side chain of the receptor. Right panel) Schematic representation of the distances between the different carbon atoms and the center of the aromatic ring. B) Distances $\text{C}\alpha-\text{Ar}$ obtained from the unrestrained MD simulations for the different complexes.

bonds of the backbone and the center of this aromatic ring. For compounds **1** and **3**, the distance between the OMe group and the aromatic ring (Ar) is less than 3.6 \AA . In compound **4**, without methyl groups in the backbone, the distance $\text{C}\beta-\text{Ar}$ is also less than 3.6 \AA . This scenario considerably differs when a methyl group at α and/or β positions is present. The presence of this additional moiety inhibits the $\text{CH}-\pi$ interaction, with distances $\text{C}\beta-\text{Ar}$ and $\text{C}\alpha-\text{Ar}$ greater than 4.2 \AA . Although the additional methyl groups of compounds **5–7** could make contact with the Ar group, these interactions should be rather weak due to the large distance $\text{C}-\text{Ar}$. In fact, only in the case of compound **6** is this value less than 3.8 \AA .

On the other hand, to compare our theoretical data to the experimental K_a values, we calculated the $\Delta\Delta G$ of binding for compounds **4** and **5**. To this end, we used the thermodynamic integration (TI) methodology^[15,16] as implemented in Amber 10 package (see the Supporting Information for computational details). This method has been successfully employed in biophysical studies to predict free energy changes, notably the

effect of amino acid point mutations.^[17] Based on these calculations, the difference in the $\Delta\Delta G_{\text{binding}}$ was calculated to be $0.6 \text{ kcal mol}^{-1}$ in favor of compound **4**, which is in the same direction as the experimental data.

Finally, to explain the high K_a value ($> 1000 \text{ M}^{-1}$) observed for the CKII-based *O*-GlcNAc decapeptide **8**,^[6] we ran 25 ns unrestrained MD simulation in explicit water on the receptor–**8** complex (Figure 10). In this case, the calculations were carried

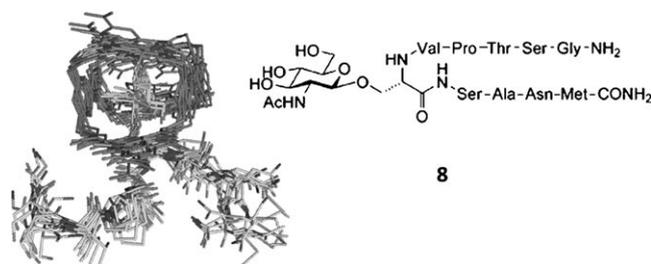


Figure 10. Calculated ensembles obtained from the 25 ns unrestrained MD simulations in explicit water for glycopeptide **8**.

out without the externally directed, water-solubilizing side chains. The MD simulations revealed that the larger backbone not only allow further interactions between the peptide and the receptor, but also increases the fraction of the key hydrogen bond $\text{NH5}\cdots\text{O1}$ up to 60%, which could be at the origin of the large experimentally measured K_a value. Although merely speculative, the high affinity could be also attributed to the electrostatic interactions between the positively charged backbone and the negatively charged side chains of the receptor. The binding free energy for the receptor–**8** complex was estimated with the MM-PBSA method as implemented in Amber 10 (see the Experimental Section for details). This binding energy (without considering the unfavorable entropy contribution) was very high: about $-40 \text{ kcal mol}^{-1}$. This result is qualitatively in agreement with the experimental data.

Conclusions

The recognition of β -*O*-GlcNAc-containing moieties by the synthetic lectin is modulated by a variety of factors, which have been scrutinized by using a combination of NMR spectroscopic methods and modeling protocols. For more complex molecules, additional interactions take place in the process. In this particular case, and for the studied compounds, the presence of a key hydrogen bond between the receptor and the OMe group of compound **1** appears to be responsible for the high selectivity observed for this compound. On the other hand, the presence of the underlying amino acid residue decreases the affinity constant, especially in those cases in which a β -methyl group is present. Thus, the serine-containing glycopeptide exhibits the highest affinity constant of the glycopeptides, and the threonine derivative shows the lowest one. The combination of NMR spectroscopic and MD methods permits us to deduce that the observed selectivity has its origin in the formation of specific hydrogen bonds and hydrophobic (CH– π) con-

tacts, which may only take place for certain orientations of the receptor groups and for the proper stereochemistries and substitutions in the glycopeptide chain.

Experimental Section

Synthesis: The synthesis of the receptor^[6] and compounds **4** and **5**^[12] have been previously reported. The synthetic routes to obtain glycopeptides **6** and **7** are described in the Supporting Information.

Compound 6: $^1\text{H NMR}$ (400 MHz, D_2O): $\delta = 1.42$ (s, 3H), 2.00 (s, 3H), 2.06 (s, 3H), 2.72 (s, 3H), 3.40–3.49 (m, 2H), 3.51–3.59 (m, 1H), 3.67–3.82 (m, 3H), 3.88–3.99 (m, 1H), 4.17 (d, $J = 9.9 \text{ Hz}$, 1H), 4.51 ppm (d, $J = 8.4 \text{ Hz}$, 1H); $^{13}\text{C NMR}$ (100 MHz, D_2O): $\delta = 19.8$, 22.1, 22.2, 26.0, 55.5, 59.4, 60.7, 69.9, 71.7, 73.5, 75.8, 101.5, 173.7, 174.5, 174.7 ppm; elemental analysis calcd (%) for $\text{C}_{15}\text{H}_{27}\text{N}_3\text{O}_8$: C 47.74, H 7.21, N 11.13; found: C 47.81, H 7.22, N 11.10.

Compound 7: $^1\text{H NMR}$ (400 MHz, D_2O): $\delta = 1.00$ (d, $J = 6.4 \text{ Hz}$, 3H), 1.31 (s, 3H), 1.94 (s, 3H), 1.99 (s, 3H), 2.65 (s, 3H), 3.37–3.47 (m, 2H), 3.48–3.55 (m, 1H), 3.61–3.69 (m, 1H), 3.76 (dd, $J = 12.1 \text{ Hz}$, $J = 5.0 \text{ Hz}$, 1H), 3.86–3.99 (m, 2H), 4.51 ppm (d, $J = 8.3 \text{ Hz}$, 1H); $^{13}\text{C NMR}$ (100 MHz, D_2O): $\delta = 13.5$, 14.4, 22.0, 22.3, 26.1, 55.3, 60.7, 62.9, 69.9, 73.4, 75.9, 78.5, 99.4, 173.1, 174.3, 174.7 ppm; elemental analysis calcd (%) for $\text{C}_{16}\text{H}_{29}\text{N}_3\text{O}_8$: C 49.10, H 7.47, N 10.74; found: C 49.21, H 7.44, N 10.70.

NMR spectroscopy experiments: All the NMR spectroscopy experiments were recorded on a Bruker Avance 400 spectrometer at 293 K. ^1H and ^{13}C NMR spectra were recorded in D_2O (chemical shifts are reported in ppm on the δ scale). Magnitude-mode ge-2D COSY spectra were recorded with gradients and by using the cosygpqf pulse program with 90° pulse width. Phase-sensitive ge-2D HSQC spectra were recorded by using a z-filter and selection before t1 by removing the decoupling during acquisition by use of invigpndph pulse program with CNST2 ($J_{\text{HC}} = 145$). 2D NOESY experiments were made by using phase-sensitive ge-2D NOESY for CDCl_3 spectra and phase-sensitive ge-2D NOESY with WATERGATE for $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1) spectra. Selective ge-1D NOESY experiments were carried out by using the 1D-DPFGE NOE pulse sequence. NOEs intensities were normalized with respect to the diagonal peak at zero mixing time.

$^1\text{H NMR}$ spectroscopic titrations: Solutions of glycopeptides **4–7** were made up in D_2O (99.9%). Aliquots were then added to an NMR tube containing the receptor in the same solvent (500 μL , $[\text{receptor}]_{\text{initial}} = 0.50 \text{ mM}$). The sample tube was shaken carefully after each addition and $^1\text{H NMR}$ spectra were recorded at 298 K on a Bruker Avance 500 spectrometer. Affinity constants (K_a) were obtained by integration of signals from both the free and bound receptor. Values were obtained from several spectra and the results averaged. Details are given in the Supporting Information.

Molecular dynamics simulations

Unrestrained MD simulations in explicit water: Simulations were performed by using the AMBER 10 program package^[18] (parm03),^[19] which was implemented with GLYCAM 06 and GAFF parameters^[20,21] to accurately simulate the conformational behavior of the sugar moiety and the receptor, respectively. The starting geometries were generated by superimposing the carbohydrate moiety of the glycopeptides on the structure previously proposed for the receptor–**1** complex.^[6] Counterions (12 Na^+) and a cubic box of pre-equilibrated TIP3P water molecules^[22] were then added by using Xleap module of AMBER. The simulations were run with the

PMEMD module of AMBER with SHAKE algorithm,^[23] by using periodic boundary conditions, a 2 fs time step, a temperature of 300 K, a Langevin-type thermostat^[24] for temperature control, and constant pressure of 1 atm. 9 Å cut-off was applied to the Lennard-Jones interactions, and Ewald sums for the treatment of the electrostatic interactions.^[25] An initial 2500 cycles of minimization (combining steepest descent with conjugate gradient) were run by first restraining the atoms of the complex (Na⁺ ions and water molecules were allowed to move). The whole system was then minimized by using 2000 cycles. This first step was followed by 200 ps of dynamics at constant volume with weak positional restraints on the complex (10 kcal mol⁻¹ Å⁻²). In this step the system was heated from 100 to 300 K. The restraints on the solute were then removed and we ran then 400 ps MD simulations at 300 K and 1 atm to get the appropriate density. Finally, we ran a 25 ns MD simulations by using the conditions commented on above.

MD-tar simulations: They were performed with AMBER 6.0 package^[26] (parm03),^[19] which was implemented with GLYCAM 06 parameters^[20] to accurately simulate the conformational behavior of the sugar moiety. NOE-derived distances were included as time-averaged distance constraints. A $\langle r^{-6} \rangle^{-1/6}$ average was used for the distances. Final trajectories were run by using an exponential decay constant of 8000 ps and a simulation length of 80 ns for the MD-tar simulation with $\epsilon=80$ to simulate the water environment.

Thermodynamic integration method: These simulations were carried out as implemented in AMBER 10^[8] by using the default parameters. In our calculations compound **5** was transformed into derivative **4**. The non-TI parts of the input files were carried out by using the following steps: Firstly, 500 steps of steepest-descent minimization, secondly, 500 ps of density equilibration and finally 200 ps of NTP production MD to collect dV/dλ data. A time step of 2 fs was used together with the Shake algorithm, a 9 Å direct sum cutoff, isotropic pressure scaling, and a Langevin-type thermostat. The λ parameter was set to 0.1, 0.2, ... 0.9. Only for step 2, a soft-core potential were activate (by setting ifsc = 1). This uses a different form of the Lennard-Jones equation, specifically designed for better convergence of TI calculations in the case of appearing or disappearing atoms. Plots showing dV/dλ as function of λ are shown in the Supporting Information.

MM-PBSA method.^[27] The simulation to estimate the binding energy between the receptor and glycopeptide **8** was carried out as implemented in Amber 10 program package,^[18] by using the default keywords in the MMPBSA.py script (http://ambermd.org/tutorials/advanced/tutorial3/py_script/files/mmpbsa_py.tar.gz).

Acknowledgements

We are grateful to the Ministerio de Ciencia e Innovación (CTQ2009-13814/BQU and grant to A.F.-T.) the Universidad de La Rioja (EGI10/65), the Gobierno de La Rioja (Colabora 2010/05), the EPSRC (research grant EP/D060192/1) and the Commonwealth Scholarship Commission (Scholarship to G.J.). We also thank CESGA for computer support.

Keywords: conformational analysis • glycopeptides • molecular dynamics • molecular recognition • NMR spectroscopy

- [1] a) L. Wells, K. Vosseller, G. W. Hart, *Science* **2001**, *291*, 2376–2378; b) J. E. Rexach, P. M. Clark, L. C. Hsieh-Wilson, *Nat. Chem. Biol.* **2008**, *4*, 97–106; c) N. E. Zachara, G. W. Hart, *Chem. Rev.* **2002**, *102*, 431–438; d) R. G.

- Spiro, *Glycobiology* **2002**, *12*, 43R–56R; e) C. Butkinaree, K. Park, G. W. Hart, *Biochim. Biophys. Acta Gen. Subj.* **2010**, *1800*, 96–106; f) R. Hurtado-Guerrero, H. C. Dorfmueller, D. M. F. van Aalten, *Curr. Opin. Struct. Biol.* **2008**, *18*, 551–557.
- [2] G. W. Hart, M. P. Housley, C. Slawson, *Nature* **2007**, *446*, 1017–1022.
- [3] F. I. Comer, G. W. Hart, *J. Biol. Chem.* **2000**, *275*, 29179–29182.
- [4] N. E. Zachara, G. W. Hart, *Biochim. Biophys. Acta Gen. Subj.* **2004**, *1673*, 13–28.
- [5] a) Y. Akimoto, L. K. Kreppel, H. Hirano, G. W. Hart, *Diabetologia* **2000**, *43*, 1239–1247; b) M. Brownlee, *Nature* **2001**, *414*, 813–820; c) K. Vosseller, L. Wells, M. D. Lane, G. W. Hart, *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 5313–5318.
- [6] a) Y. Ferrand, M. P. Crump, A. P. Davis, *Science* **2007**, *318*, 619–622; b) Y. Ferrand, E. Klein, N. P. Barwell, M. P. Crump, J. Jiménez-Barbero, C. Vicent, G. J. Boons, S. Ingale, A. P. Davis, *Angew. Chem.* **2009**, *121*, 1807–1811; *Angew. Chem. Int. Ed.* **2009**, *48*, 1775–1779.
- [7] a) A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *Chem. Phys. Lett.* **1989**, *157*, 289–294; b) D. A. Pearlman, P. A. Kollman, *J. Mol. Biol.* **1991**, *220*, 457–479; c) A. E. Torda, R. M. Scheek, W. F. van Gunsteren, *J. Mol. Biol.* **1990**, *214*, 223–235; d) A. E. Torda, R. M. Brunne, T. Huber, H. Kessler, W. F. van Gunsteren, *J. Biomol. NMR* **1993**, *3*, 55–66; e) D. A. Pearlman, *J. Biomol. NMR* **1994**, *4*, 279–299; f) A. P. Nanzer, W. F. van Gunsteren, A. E. Torda, *J. Biomol. NMR* **1995**, *6*, 313–320.
- [8] a) F. Corzana, J. H. Busto, S. B. Engelsen, J. Jiménez-Barbero, J. L. Asensio, J. M. Peregrina, A. Avenoza, *Chem. Eur. J.* **2006**, *12*, 7864–7871; b) F. Corzana, J. H. Busto, G. Jiménez-Osés, J. L. Asensio, J. Jiménez-Barbero, J. M. Peregrina, A. Avenoza, *J. Am. Chem. Soc.* **2006**, *128*, 14640–14648.
- [9] a) S. Vandenbussche, D. Diaz, M. C. Fernández-Alonso, W. Pan, S. P. Vincent, G. Cuevas, F. J. Cañada, J. Jiménez-Barbero, K. Bartik, *Chem. Eur. J.* **2008**, *14*, 7570–7578; b) G. Terraneo, D. Potenza, A. Canales, J. Jiménez-Barbero, K. K. Baldrige, A. Bernardi, *J. Am. Chem. Soc.* **2007**, *129*, 2890–2900; c) K. Ramírez-Gualito, R. Alonso-Ríos, B. Quiroz-García, A. Rojas-Aguilar, D. Diaz, J. Jiménez-Barbero, G. Cuevas, *J. Am. Chem. Soc.* **2009**, *131*, 18129–18138.
- [10] J. Emsley, *Chem. Soc. Rev.* **1980**, *9*, 91–124.
- [11] a) A. Fernández-Tejada, F. Corzana, J. H. Busto, G. Jiménez-Osés, J. M. Peregrina, A. Avenoza, *Chem. Eur. J.* **2008**, *14*, 7042–7058; b) A. Fernández-Tejada, F. Corzana, J. H. Busto, A. Avenoza, J. M. Peregrina, *Org. Biomol. Chem.* **2009**, *7*, 2885–2893; c) A. Fernández-Tejada, F. Corzana, J. H. Busto, A. Avenoza, J. M. Peregrina, *J. Org. Chem.* **2009**, *74*, 9305–9313.
- [12] a) A. Fernández-Tejada, F. Corzana, J. H. Busto, G. Jiménez-Osés, J. Jiménez-Barbero, A. Avenoza, J. M. Peregrina, *Chem. Eur. J.* **2009**, *15*, 7297–7301; b) C. Aydillo, G. Jiménez-Osés, J. H. Busto, J. M. Peregrina, A. Avenoza, M. M. Zurbano, *Chem. Eur. J.* **2007**, *13*, 4840–4848.
- [13] F. Corzana, J. H. Busto, G. Jiménez-Osés, M. García de Luis, J. L. Asensio, J. Jiménez-Barbero, J. M. Peregrina, A. Avenoza, *J. Am. Chem. Soc.* **2007**, *129*, 9458–9467.
- [14] Letter code employed in S. S. Zimmerman, M. S. Pottle, G. Nemethy, H. A. Scheraga, *Macromolecules* **1977**, *10*, 1–9.
- [15] a) T. Simonson in *Computational Biochemistry and Biophysics* (Eds.: O. Becker, A. D. MacKerell, B. Roux, M. Watanabe) Marcel Dekker, New York, **2001**; b) P. Kollman, *Chem. Rev.* **1993**, *93*, 2395–2417; c) M. L. Lamb, W. L. Jorgensen, *Curr. Opin. Chem. Biol.* **1997**, *1*, 449–457.
- [16] T. Steinbrecher, D. A. Case, A. Labahn, *J. Med. Chem.* **2006**, *49*, 1837–1844.
- [17] T. Simonson, G. Archontis, M. Karplus, *J. Phys. Chem. B* **1997**, *101*, 8349–8362.
- [18] P. A. Kollman et al, AMBER 10, University of California, San Francisco, **2008**.
- [19] Y. Duan, C. Wu, S. Chowdhury, M. C. Lee, G. Xiong, W. Zhang, R. Yang, P. Cieplak, R. Luo, T. Lee, J. Caldwell, J. Wang, P. A. Kollman, *J. Comput. Chem.* **2003**, *24*, 1999–2012.
- [20] 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*, 622–655.
- [21] J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, *J. Comput. Chem.* **2004**, *25*, 1157–1174.
- [22] M. W. Mahoney, W. L. Jorgensen, *J. Chem. Phys.* **2000**, *112*, 8910–8922.
- [23] J.-P. Ryckaert, G. Ciccotti, H. J. C. Berendsen, *J. Comput. Phys.* **1977**, *23*, 327–341.

[24] S. A. Adelman, J. D. Doll, *J. Chem. Phys.* **1976**, *64*, 2375–2388.

[25] T. Darden, H. Lee, L. G. Pedersen, *J. Chem. Phys.* **1995**, *103*, 8577–8593.

[26] P. A. Kollman et al., AMBER 6, University of California, San Francisco, **1999**.

[27] J. Wang, P. Morin, W. Wang, P. A. Kollman, *J. Am. Chem. Soc.* **2001**, *123*, 5221–5230.

Received: September 3, 2010

Published online on December 22, 2010
