DOI: 10.1002/chem.200801777

The Nature and Sequence of the Amino Acid Aglycone Strongly Modulates the Conformation and Dynamics Effects of Tn Antigen's Clusters

Francisco Corzana,*[a] Jesús H. Busto,[a] Marisa García de Luis,[a] Jesús Jiménez-Barbero,[b] Alberto Avenoza,[a] and Jesús M. Peregrina*[a]

Abstract: Synthetic oligosaccharide vaccines based on carbohydrate epitopes are currently being evaluated as potential immunotherapeutics in the treatment of cancer. In an effort to study the role that the amino acid moiety (L-serine and/or L-threonine residues) plays on the global shape of the resulting glycopeptides and on the dynamics of the carbohydrate moiety, diverse glycopeptides based on the Tn antigen have been synthesized and studied in aqueous solution by combining NMR spectroscopic experiments and molecular dynamics simulations. Our results demonstrate that although the effect of the clustering of Tn on the peptide backbone is not remarkable, it substantially modifies the dynamics, and thus, the presentation features of the carbohydrate moiety. In fact, the selected sequence has a crucial influence on both the orientation and flexibility of the sugar region. Thus, although a serine–threonine pair shows a well-defined spatial disposition of the Tn epitopes, its analogue sequence threonine–serine allows a certain degree of mobility that could favor the interaction with a diversity of receptors without a major energy penalty. These features can be explained by attending to the different conformational behavior of the glycosidic linkage of threonine-containing glycopeptides when compared with those of the serine analogues. On this basis, and taking into account that these carbohydrates interact with components of the immune system, these findings could have implications for further design of new cancer vaccines.

Keywords: antigens · conformation analysis · glycopeptides · molecular dynamics · NMR spectroscopy

Introduction

Mucin MUC1 is an extensively *O*-glycosylated protein that is largely composed of an expressed variable number of tandem repeats of the sequence PDTRPAPGSTAPPAHGVTSA.[1] Expression of this mucin MUC1 is drastically increased in tumor cells. This over-expression is accompanied by down-regulation of glucosaminyltransferase activity. As a result of this incomplete glycosylation, the normally hidden core Tn (α-*O*-GalNAc-Ser/Thr), sialyl-Tn (STn), and Thomsen–Friedenreich (TF) carbohydrate-based antigens are now accessible[2] and may be associated with cancer progression and metastasis.[3,4] Therefore, mucins and their derivatives are attracting real interest as potential targets for immunotherapy in the development of vaccines for the treatment of cancer.[5]

Two main approaches have been taken for carbohydrate-based vaccines designs: the monomer approach, using isolated carbohydrate-based antigens, and the cluster approach, achieved by glycosylation of contiguous L-serine (S or Ser) and/or L-threonine (T or Thr) residues. Thus, dimeric (or even trimeric) clusters of carbohydrate–antigens have been attached to a Ser or Thr peptide scaffold.[6]

From a structure-based design perspective, and to discern how mucins interact with their biological targets, it is essential to study in detail their conformational preferences and the factors that can modify these preferences. At the present moment, it is well known that α-O-glycosylation forces the underlying peptide into an extended conformation,[7] which minimizes the steric crowding of the proximal carbohydrate structures and could favor the binding between the sugar and the components of the immune system.[8] This finding has been explained by forming specific hydrogen bonds be-

[a] Dr. F. Corzana, Dr. J. H. Busto, M. García de Luis, Prof. A. Avenoza, Dr. J. M. Peregrina
Departamento de Química, Universidad de La Rioja
Grupo de Síntesis Química de La Rioja
UA-CSIC. Madre de Dios 51, 2006 Logroño (Spain)
Fax: (+34) 941-299-621
E-mail: jesusmanuel.peregrina@unirioja.es

[b] Prof. J. Jiménez-Barbero
Centro de Investigaciones Biológicas (CSIC)
Ramiro Maeztu 9, 2840 Madrid (Spain)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/chem.200801777.

tween the peptide moiety and the first α-D-*N*-acetylgalactosamine (GalNAc) unit, which "locks" the orientation of the sugar with respect to the peptide backbone.[7,9] In this context, we have recently reported that the existence of bridging-water molecules between the sugar and the peptide moieties could also explain the stabilization of the extended conformation of the backbone.[10a] Interestingly, the conformational analyses of LI-cadherin glycopeptides reported by Kuhn and Kunz not only confirm the effect of the carbohydrate moiety on the peptide backbone but also show that this effect is dependent on the carbohydrate type.[11]

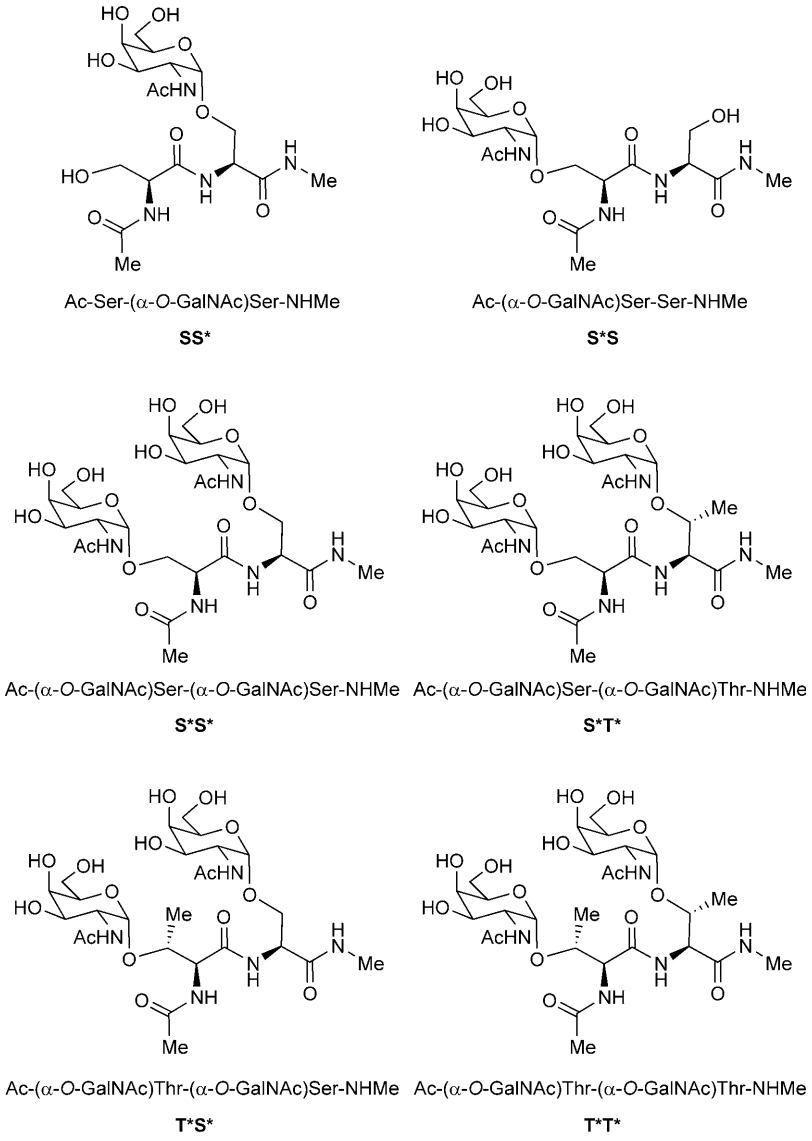
On the other hand, although an important number of studies have been reported on the preferred conformations of structurally complex and relevant mucin-related glycopeptides,[7,11] few conformational studies have been carried out on Tn glycopeptides containing "clustered" Ser/Thr sequences. An excellent example was reported by Kunz and co-workers.[12] The study revealed that the glycosylation of the STAPPA and PAPGSTAPPA peptides with αGalNAc results in a stiffening effect at the site of glycosylation. The authors conclude that this feature could have important implications on the binding affinity between the antibody and the MUC1-core-related glycopeptides.

From a biological point of view, some important features have been observed. First, the amino acids flanking the carbohydrate antigen have been shown to modulate antibody recognition.[13] Second, recent studies have experimentally shown that the aglyconic part of the Tn structure (Ser or Thr residues) could also play a key role for antibody recognition. Thus, the STT backbone was found to be the most permissive scaffold for induction of anti-Tn antibodies.[14] Moreover, in recent work[5m] concerning the study of synthetic MUC1-based glycopeptides used as antigen-presenting cells, only the processed glycopeptides that carried GalNAc α-*O*-linked to the G*ST*A motif of the peptide were capable of activating hybridoma T-cells. Glycopeptides with glycosylation at the V*TS*A motif were inactive.

In this context, we have recently reported that the glycosidic linkages of GalNAc-Ser and GalNAc-Thr adopt completely different 3D orientations.[15] These distinct spatial arrangements of the carbohydrates could have important inferences on the global shape and presentation of the epitope, which could be of paramount importance for a proper interaction between the sugars and the components of the immune system.

Importantly, although great efforts have been made to understand the influence of the carbohydrate moiety on the peptide backbone, not enough information has been reported concerning the dynamics and conformational preferences of the carbohydrate, particularly, with regard to the conformational behavior of the glycosidic linkage.

Taking into account the above comments and to gain detail on the influence that the underlying amino acid (Ser or Thr) has on the global structure as well as on the dynamics of the carbohydrate moiety, we have synthesized and carried out the conformational study in an aqueous solution of the model glycopeptides as shown in Scheme 1. The confor-



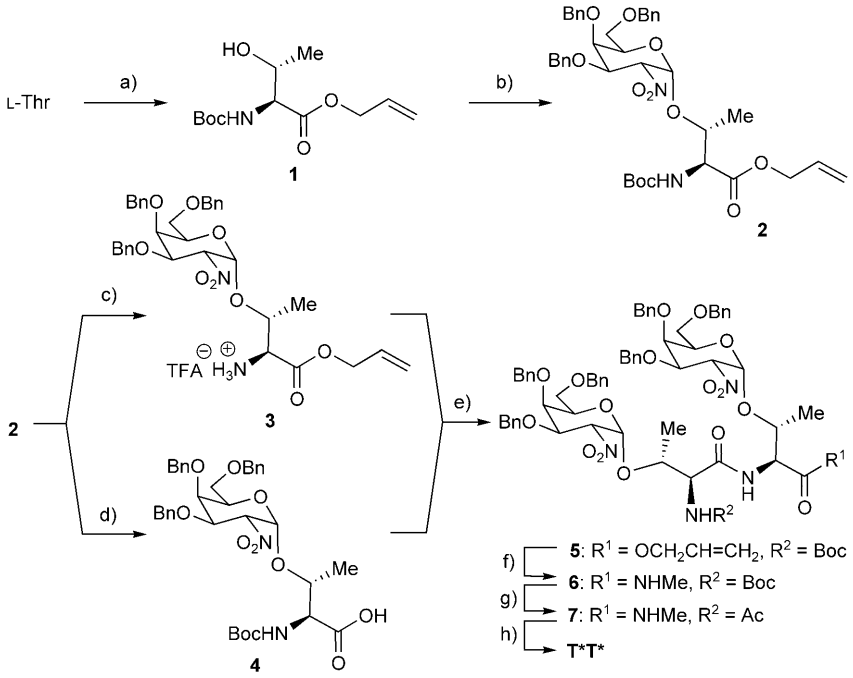
Scheme 1. Glycopeptides studied in this work.

mational analysis of all these compounds has been carried out by using NMR spectroscopy. In particular, NOE and coupling constants have been interpreted with the assistance of molecular dynamics (MD) simulations, following a protocol previously described by our group.[10]

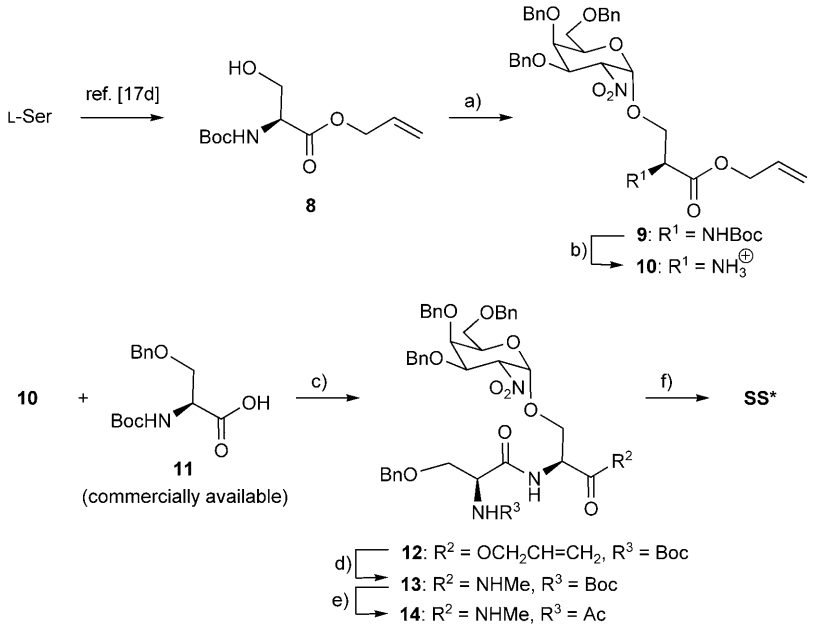
Results and Discussion

Synthesis: As key examples, Scheme 2 and Scheme 3 show the synthesis of **T*T*** and **SS*** glycopeptides. The rest of the molecules were synthesized as described in the Supporting Information and follow a similar methodology.

The synthesis starts with L-Thr, whose carboxylic acid group was protected as an allyl ester, a selectively removable carboxy-protecting function,[16] and the amino group as a *tert*-butyloxycarbonyl (Boc) carbamate, giving compound **1**. The treatment of the protected Thr derivative **1** with 3,4,6-tri-*O*-benzyl-2-nitrogalactal, following the procedure described by Schmidt and co-workers,[17a–c] gave the linked 2-nitroglycoside **2** in good yield. Then, selective deprotection of the amino or acid groups in this compound gave, almost quantitatively, derivative **3** or **4**, respectively. Further treatment of a mixture of **3** and **4** with *N*,*N*,*N*′,*N*′-tetramethyl-*O*-(benzotriazol-1-yl)uronium tetrafluoroborate (TBTU) as a coupling agent and diisopropylethylamine (DIEA) as a base led to the doubly glycosylated compound **5**. Selective removal of the allyl group was carried out with Pd^0-catalyzed allyl transfer to morpholine. Under these conditions, the deprotection proceeded without affecting other labile groups.[18] The subsequent transformation of the acid group into methyl amide gave derivative **6**. Deprotection of the amino group and its further acetylation with acetic anhydride (Ac_2O) and pyridine (Py) led to **7** in a



Scheme 2. Synthesis of **T*T***. a) i) Allylic alcohol, *p*TsOH, toluene, reflux, 24 h. ii) $(Boc)_2O$, Na_2CO_3, THF/H_2O (5:1), 25 °C, 12 h, 71 % overall yield; b) 3,4,6-Tri-*O*-benzyl-2-nitrogalactal, *t*BuOK, THF, 25 °C, 12 h, 68 %; c) TFA/CH_2Cl_2 (2:5), 0 to 25 °C, 3 h, quantitative; d) $[Pd(PPh_3)_4]$, morpholine, THF, 25 °C, 3 h, 87 %; e) TBTU, DIEA, CH_3CN, 25 °C, 10 h, 67 %; f) i) $[Pd(PPh_3)_4]$, morpholine, THF, 25 °C, 3 h. ii) $MeNH_2$·HCl, DIEA, CH_3CN, 25 °C, 10 h, 64 % overall yield; g) i) TFA/CH_2Cl_2 (1:2), 0 to 25 °C, 3 h. ii) Ac_2O, Py, 25 °C, 3 h, 80 % overall yield; h) i) H_2/Ni Raney (T4), EtOH, 25 °C, 7 h. ii) Ac_2O, Py, 25 °C, 3 h. iii) H_2/Pd-C, MeOH/ethyl acetate (2:1), 25 °C, 12 h, 33 % overall yield.



Scheme 3. Synthesis of **SS***. a) 3,4,6-Tri-*O*-benzyl-2-nitrogalactal, *t*BuOK, THF, 25 °C, 12 h, 74 %; b) TFA/CH_2Cl_2 (2:5), 0 to 25 °C, 3 h, quantitative; c) TBTU, DIEA, CH_3CN, 25 °C, 10 h, 69 %; d) i) $[Pd(PPh_3)_4]$, morpholine, THF, 25 °C, 3 h. ii) $MeNH_2$·HCl, DIEA, CH_3CN, 25 °C, 10 h, 65 % overall yield; e) i) TFA/CH_2Cl_2 (1:2), 0 to 25 °C. ii) Ac_2O, Py, 25 °C, 3 h, 83 % overall yield; f) i) H_2/Ni Raney (T4), EtOH, 25 °C, 7 h. ii) Ac_2O, Py, 25 °C, 3 h. iii) H_2/Pd-C, 25 °C, MeOH/ethyl acetate (2:1), 12 h, 46 % overall yield.

good yield. The reduction of the nitro groups was carried out by hydrogenation of **7** with nickel Raney (T4). Finally, acetylation of the amino groups and deprotection of the hydroxy groups of the sugar moieties gave the desired model glycopeptide **T*T*** with a moderate yield (Scheme 2).

Glycopeptide **SS*** was obtained as follows. First, 2-nitroglycoside **9** was synthesized starting from the Ser derivative **8**[17d] and following the methodology described above for its analogue, **6**. Then, selective deprotection of the amino group afforded compound **10**, which was treated with the commercially available Ser derivative **11** in the presence of TBTU and DIEA to give **12**. Finally, this compound was transformed into glycopeptide **SS*** by using the strategy described above for the synthesis of **T*T*** (Scheme 3).

NMR spectroscopic studies: In a first step, full assignment of ^{1}H NMR spectra of all of the compounds was carried out by using standard COSY, HSQC, and NOESY experiments. This assignment is displayed in the Supporting Information. The torsional angles and the numbering used in this work for the key compounds are shown in Figure 1.

Chemical shift deviation: Generally, in proteins and large peptides, Hα chemical-shift deviations (CSD, $\Delta\delta H\alpha = \delta H\alpha_{observed} - \delta H\alpha_{random\,coil}$) exhibit an average value of −0.39 ppm when a residue is placed in a helical conformation. In contrast, a mean shift of +0.37 ppm is observed when the residue is found in an extended conformation.[19] As can be seen in Figure 2, the obtained CSD values[20] of

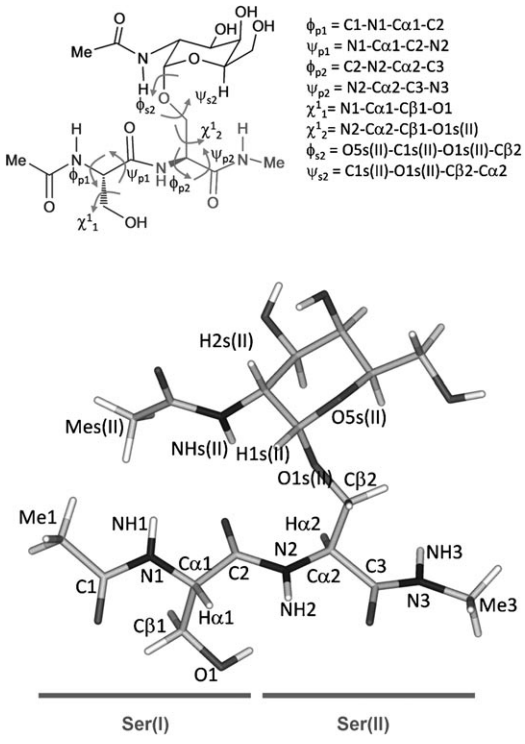


Figure 1. Molecular structure of compound **SS*** with definitions of the torsional angles and the numbering of the atoms. The same definitions were used for the other model glycopeptides.

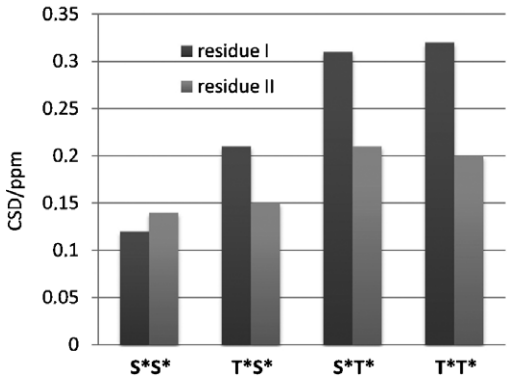


Figure 2. Chemical-shift deviations (CSD) of Hα of different glycopeptides. The CSD for Ser and Thr residues was calculated taking $\delta H\alpha_{random\,coil}$ as 4.47 and 4.35, respectively (see reference [20]).

the Hα of the glycosylated residues showed a significant downfield shift. Therefore, although our model glycopeptides are too short to adopt a well-defined secondary structure, this result is in agreement with the increase in the population of extended peptide structures upon glycosylation.[10]

2D NOESY cross-peaks: 2D NOESY experiments were then carried out in H_2O/D_2O (9:1) (20 °C, pH 5.2) for the model glycopeptides (see the Supporting Information and Figure 3 and Figure 4). In all cases, strong consecutive Hα,NH (i,i+1) connectivities were found (Figure 3). This finding, together with the weak or absence of consecutive NH,NH (i,i+1) NOE values (Figure 4) indicated the predominance of extended backbone conformations.[21] On the other hand, when a Thr residue was present, a small–medium NOE value between the NH hydrogen of the corresponding GalNAc (NHs) and the NH of the attached Thr was observed, indicating the presence of an eclipsed conformation[15] for the ψ_s dihedral angle of the GalNAc-Thr moiety (Figure 4). Furthermore, no significant NOE values were detected between the methyl or the NHs proton of the *N*-acetyl group of the GalNAc and the NH (i+1) proton of the peptide backbone. This suggests the absence of relevant hydrogen bonds between the sugar and the peptide moieties.[10a]

As a next step, distances involving NH protons were semi-quantitatively determined by integrating the corresponding cross-peak volumes in the 2D NOESY spectrum (Table 1).

^3J coupling constants: In addition, ^3J coupling constants were measured from the splitting of the resonance signals in the 1D spectrum (Table 2). The relatively high values of the peptide backbone $^3J_{NH,H\alpha}$ coupling constants for all the residues also suggest the existence of a preferred extended structure for all the glycopeptides.[22] The $^3J_{NHs,H2s}$ values in the GalNAc residues were all close to 9 Hz, thus suggesting that the torsion angle between the H2 and NH protons to be ≈180°.[23] Accordingly, the orientation of the *N*-acetyl group relative to the sugar framework seems essentially

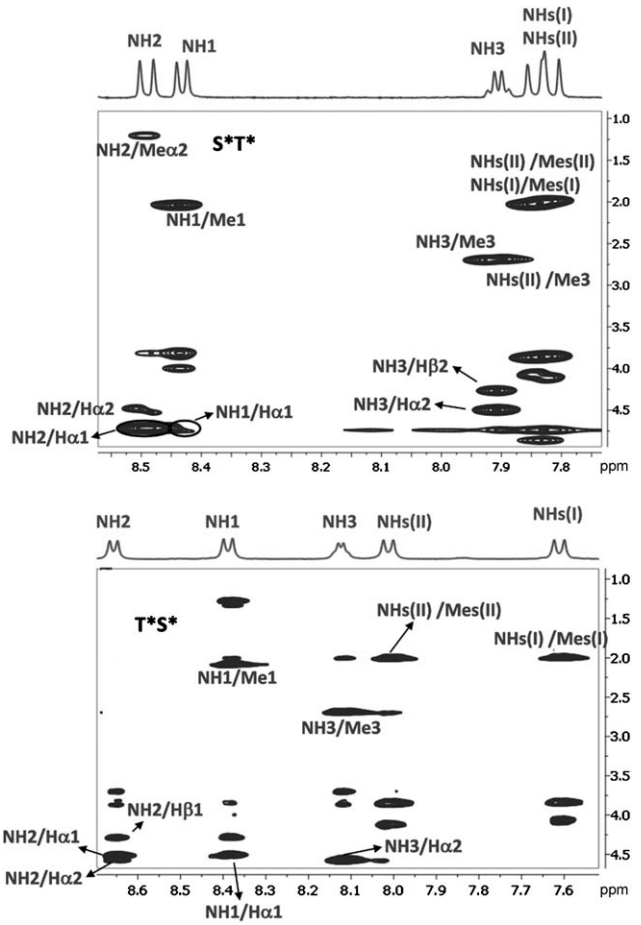


Figure 3. Sections of the 800 ms NOESY spectra (400 MHz) of compounds **S*T*** and **T*S*** in H_2O/D_2O (9:1) at 20 °C showing amide–aliphatic cross-peaks. The NOE contacts are represented as positive cross-peaks.

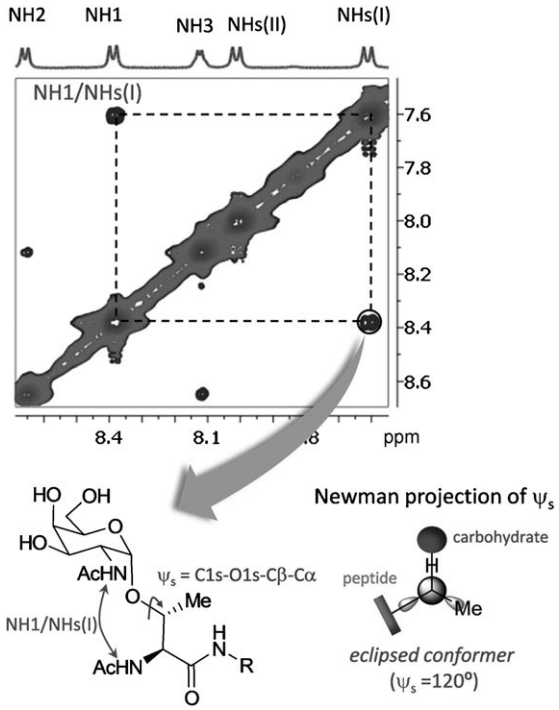


Figure 4. Section of the 800 ms NOESY spectrum (400 MHz) of compounds **T*S*** in H_2O/D_2O (9/1) at 20 °C showing the amide–amide cross-peaks. The NOE contact between NH1-NHs(I) is representative of an eclipsed conformation of the glycosidic linkage ($\psi_s = 120°$). Diagonal peaks are negative and the NOE contacts are represented as positive cross-peaks.

fixed. Additionally, $^3J_{H\alpha,H\beta}$ values were also determined to provide insights into the χ^1 angles for the Thr and Ser side chains. Strikingly, the coupling constants for the Thr residues had remarkably small values, falling between 2.4 and 2.7 Hz. These low values preclude any significant torsional averaging and strongly indicate that the rotation around χ^1 is rather restricted.[24]

MD simulations: To get an experimentally derived ensemble, MD-tar (MD with time-averaged restraints)[25] simulations were carried out by the inclusion of the experimental distances and ^3J coupling constants (Table 1 and Table 2) as time-averaged restraints. The results obtained from the MD-tar simulations (see the Supporting Information) showed a close agreement, in numerical terms, between the calculated and the experimental proton–proton distances and ^3J coupling constants.

The calculated φ_p,ψ_p distribution maps for the peptide backbone of both the singly and doubly Tn-glycosylated derivatives were similar to the typical values expected for extended conformations, such as polyproline II and the β-sheet conformations. Indeed, only a small number of conformers showed φ_p,ψ_p dihedral values corresponding to helix-like conformations (Figure 5 and the Supporting Information). This result is in good agreement with the observed pattern of NOE values described above.

Figure 6 shows the side-chain distribution (χ^1 dihedral angle) obtained from the MD-tar simulations. Important conclusions can be drawn from these data. First, the side chain of the non-glycosylated residues was more flexible than those of their glycosylated counterparts, showing a significant population of the three possible staggered conformers [*anti*, with χ^1 around 180°; *gauche* (+) or *g*(+), with χ^1 around 60°; and *gauche* (−) or *g*(−), with χ^1 around −60°]. Second, the rotation around χ^1 for the glycosylated Thr residues was significantly restricted, with a clear preference for a *g*(+) conformation, except for the derivative **T*S***, for which both the *anti* and *g*(+) conformers seem to be similarly populated. These findings are all consistent with the small $^3J_{H\alpha,H\beta}$ coupling constant values for Thr residues.

On the other hand, no significant hydrogen bonding was detected over the course of MD-tar simulations between the *N*-acetyl group of the GalNAc residue and the carbonyl group of the underlying amino acid in any of the studied glycopeptides. Indeed, if any, this inter-residual hydrogen bonding was present in less than 1% of the total trajectory for all glycosylated Ser residues. In contrast, in the case of Thr residues, this hydrogen bonding was present about 10% of time. This result is in accordance with the experimental

Table 1. Calculated and experimental (in brackets) distances obtained for the glycopeptides.[a]

Distance	**SS***	**S*S**	**S*S***	**S*T***	**T*S***	**T*T***[b]
NH1–NH2	>3.0	>3.0	>3.0	>3.0	>3.0	>3.0
	(3.0)	(3.1)	(3.3)	(3.4)	(3.5)	(3.0)
NH2–NH3	2.9	3.1	3.1	2.8	3.3	>3.0
	(3.1)	(3.1)	(3.2)	(3.4)	(3.2)	(3.1)
NHs(I)–NH1	–	–	–	–	2.8	2.9
	(–)	(3.6)	(3.2)	(–)	(2.9)	(2.7)
NHs(II)–NH2	–	–	–	2.8	–	2.8
	(3.4)	(–)	(3.5)	(3.9)	(–)	(2.7)
Hα1–NH1	2.7	2.9	2.5	2.7	2.5	2.7
	(2.8)	(2.9)	(2.7)	(2.7)	(2.7)	(2.8)
Hα1–NH2	2.3	2.3	[c]	2.2	2.3	2.3
	(2.5)	(2.4)	(2.6)	(2.5)	(2.5)	(2.5)
Hα2–NH2	3.2	2.8	[c]	2.9	3.1	2.8
	(2.9)	(2.9)	(2.6)	(2.8)	(2.8)	(2.7)
Hα2–NH3	2.3	2.2	2.2	2.2	2.1	2.2
	(2.5)	(2.4)	(2.4)	(2.4)	(2.3)	(2.4)
NH1–Hβ1[d]	[c]	[c]	2.8/2.8[d]	2.6/2.7	2.9	2.9
	(2.7/2.7)	(2.7/3.0)	(2.7/2.8)	(2.6/2.9)	(2.9)	(3.0)
NH2–Hβ1[d]	[c]	[c]	[c]	[c]	2.9	2.6
	(3.0/2.8)	(3.2/2.5)	(3.0/2.5)	(3.1/2.5)	(2.7)	(2.4)
NH2–Hβ2[d]	[c]	[c]	[c]	[c]	3.6/3.4	3.2
	(2.8/3.0)	(2.8/2.7)	(2.7/3.1)	(3.4)	(3.2/3.2)	(3.4)
NH3–Hβ2[d]	[c]	[c]	[c]	2.4	3.5/2.9	2.3
	(3.0/2.5)	(3.0/2.7)	(3.1/2.6)	(2.4)	(3.2/2.7)	(2.4)

[a] The distances are given in Å. The maximum estimated experimental error is 10%. [b] In the case of the **T*T*** sequence, NOE build-up curves were carried out to give H1s(I)–Hβ1 (2.1 Å) and H1s(II)–Hβ2 (2.2 Å) distances. [c] Not determined owing to an overcrowded NMR spectrum. [d] The first value is referred to Hβ*proR* and the second one to Hβ*proS*.

Table 2. Calculated and experimental (in brackets) ^3J coupling constants obtained for the glycopeptides.[a]

^3J	**SS***	**S*S**	**S*S***	**S*T***	**T*S***	**T*T***
Hα1, Hβ1[b]	5.8[c]	5.1[c]	5.3[c]	6.0/4.5	2.5	2.7
	(4.8/4.9)	(5.5/4.8)	(5.6/5.1)	(6.4/4.3)	(2.8)	(2.8)
Hα2, Hβ2[b]	5.1[c]	5.4[c]	4.5[c]	2.7	5.7[c]	2.4
	(4.7/5.1)	(5.2/5.7)	(4.9/4.3)	(3.1)	(5.7/4.8)	(3.3)
NH1,Hα1	6.6	7.5	6.6	6.9	8.3	8.5
	(6.8)	(7.4)	(6.9)	(6.7)	(7.0)	(7.4)
NH2,Hα2	7.2	7.7	7.2	9.1	7.4	9.1
	(7.3)	(7.4)	(5.6)	(7.0)	(7.2)	(6.7)
NHs(I),H2s(I)	–	9.4	9.2	9.7	9.6	9.5
	(–)	(9.4)	(9.4)	(9.4)	(9.4)	(9.5)
NHs(II),H2s(II)	9.8	–	9.2	9.7	9.2	8.8
	(9.4)	(–)	(9.4)	(9.4)	(9.4)	(9.4)

[a] ^3J coupling constants are given in Hz. The estimated experimental error is 10%. [b] The first value is referred to Hβ*proR* and the second one to Hβ*proS*, which were extracted from *J* resolved experiments (see the Supporting Information). [c] Experimentally observed as a pseudotriplet.

findings reported by Gururaja and co-workers,[26] who previously stated that, in mucin derivatives, this hydrogen bond was only present when a Thr residue was incorporated in the glycopeptide.

As we have recently reported,[15] one of the most important differences between the glycosidic linkages of Ser and Thr concerns the ψ_s dihedral angle value (C1s-O1s-Cβ-Cα). In general, when a Thr is present, this angle adopts a value close to 120° (showing an eclipsed conformation around O1s-Cβ as shown in Figure 4). In contrast, for Ser derivatives a value around 180° is preferred. The different behavior of both glycosidic linkages can be seen in Figure 7. As a consequence of the different ψ_s values in Ser and Thr derivatives, the carbohydrate moiety adopts a complete different orientation. In fact, although the GalNAc moiety in Thr is almost perpendicular to the peptide backbone, in Ser residues, the sugar moiety adopts a parallel disposition. It is evident that these differences must also have implications in the global shape of the molecules, and also in the flexibility of the systems. Figure 8 shows the superimposition of 20 conformers randomly taken from the MD-tar simulations. As can be seen, both **S*T*** and **T*T*** can properly accommodate the two GalNAc residues, giving a rather rigid structure. However, in the case of **T*S*** and **S*S***, a larger degree of flexibility was observed. Additionally, the global shape of **T*S*** differs completely from that derived for the rest of the model glycopeptides. In fact, the sugar moieties are located at different regions of the plane defined by the peptide. On the other hand, the **S*T*** glycopeptides exhibits one of the lowest root mean square deviation (RMSD) values (2.2 Å), whereas the **T*S*** derivative is rather flexible with an RMSD close to 2.6 Å.

To assess the flexibility of this type of molecule, the following experimental approach was considered. Selective NOE values were obtained upon inversion of the Hβ proton of the Thr residue in **T*S*** and **S*T*** compounds at different temperatures (Figure 9). Interestingly, it was observed that at 288 K, compound **T*S*** (RMSD = 2.6 Å) presented a positive NOE with Hα. In contrast, **S*T*** (RMSD = 2.2 Å) exhibited an apparent negative NOE. This observation is consistent with a shorter local correlation time for the **T*S***, which is in agreement with the larger flexibility of this molecule, which was previously deduced from the analysis of the key cross-peaks.

With the aim of verifying whether the features commented above are reproduced in larger systems, we compared our results with those previously reported by Kunz and coworkers[12] on the Tn-glycohexapeptide **S*T***APPA. The careful inspection of both the $^3J_{NH,H\alpha}$ coupling constants (8.26 and 10.1 Hz for Ser and Thr residues, respectively) as

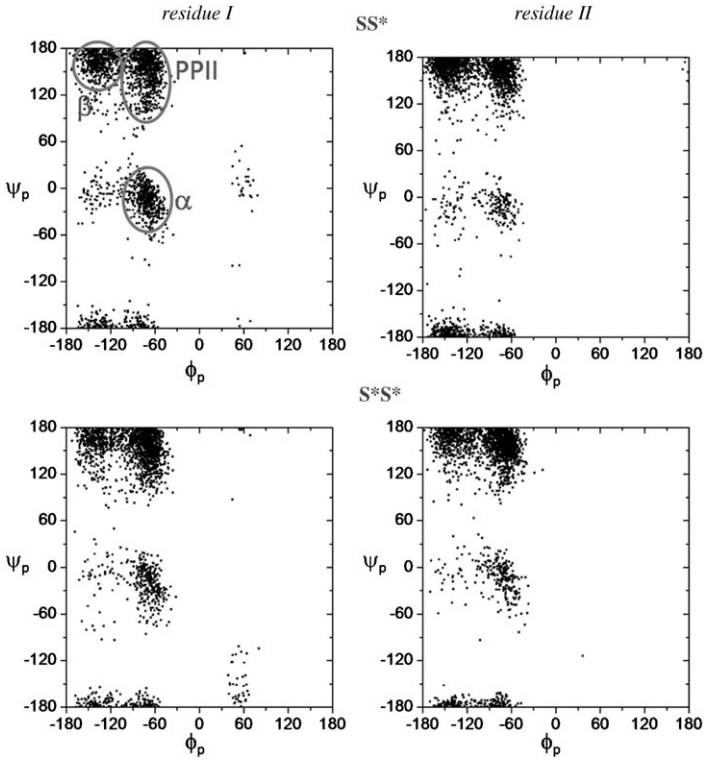


Figure 5. Distributions obtained from the MD-tar simulations for the peptide backbone (φ_p/ψ_p) of **SS*** and **S*S*** glycopeptides.

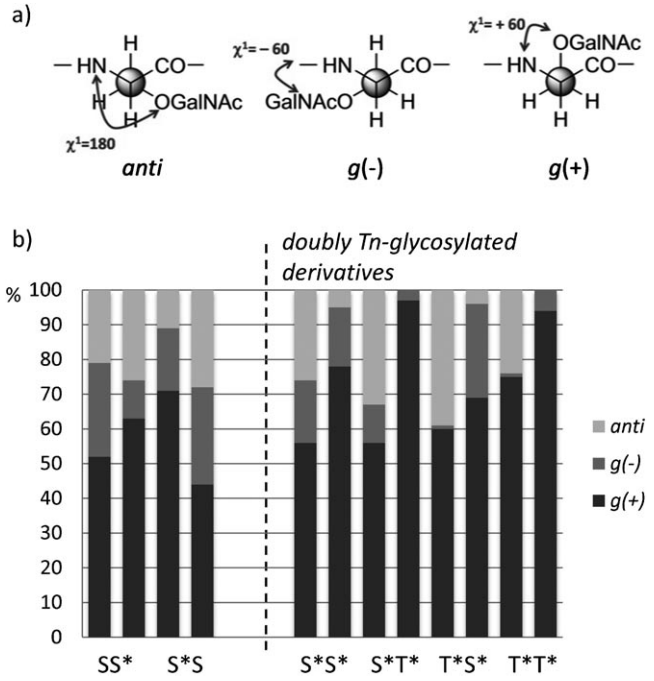


Figure 6. a) More stable conformers of the side chain (χ^1) for a glycosylated Ser residue. b) Distributions for the side chain (χ^1) obtained from the MD-tar simulations for the glycopeptides.

well as some key NOE contacts (Thr^2 NH-Ser^1 Hα and Ala^3 NH-Thr^2 Hα) indicate that the peptide adopts an extended conformation in this region of the peptide chain. Moreover, the presence of a NOE contact between Thr^2 NH and the NH proton of the GalNAc attached to this Thr corroborates the eclipsed conformation of the ψ_s torsion angle. As expected, this key NOE contact was not observed in the Ser residue. In addition, all these experimental data and the $^3J_{H\alpha,H\beta}$ coupling constants were used as restraints in the MD-tar simulations carried out on this glycohexapeptide. As can be seen in Figure 10, the dynamics and the global shape of the carbohydrate moieties are quite similar to those described in this work for the **S*T*** species.

On the other hand, the eclipsed conformation of the ψ_s glycosidic linkage of the αGalNAc-Thr residue can be experimentally observed in recent detailed studies on glycopeptides, such as the glycosylated MUC1 eicosapeptide, studied by Kunz's group,[7a] and the triglycosylated sequence STTAV reported by Danishefsky and co-workers.[7b] Furthermore, the rigidity exhibited for **S*T*** and **T*T*** glycopeptides is in accordance with the conformationally highly stable structure predicted for the STTAV molecule.

In summary, the results obtained in this work demonstrate that the peptide sequence used for clustering of Tn antigens plays a significant role, not only on the flexibility of the systems, but also, and more interestingly, on their different global shape. This result is, to some extend, in good accordance with the results previously obtained by Lo-Man and co-workers.[14] In this study, the STT backbone was found to be the most satisfactory of all those tested for the induction of anti-Tn antibodies.

Conclusion

Our results demonstrate that, although the effect of clustering of Tn on the peptide backbone is not remarkable, the clustering strongly modifies the dynamics of the carbohy-

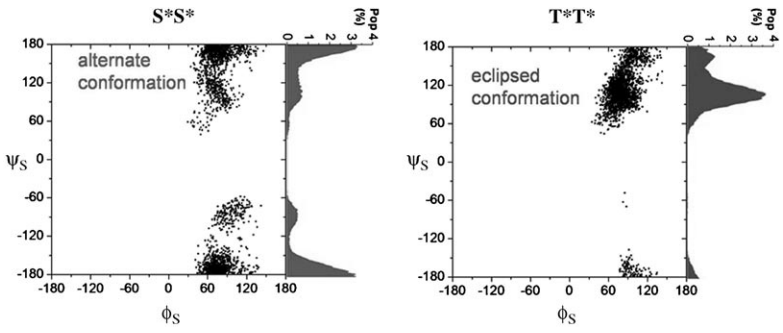


Figure 7. φ_s/ψ_s distributions obtained from the MD-tar simulations for the glycosidic linkage of residue I of **S*S*** and **T*T*** glycopeptides, together with the histograms of ψ_s dihedral angle.

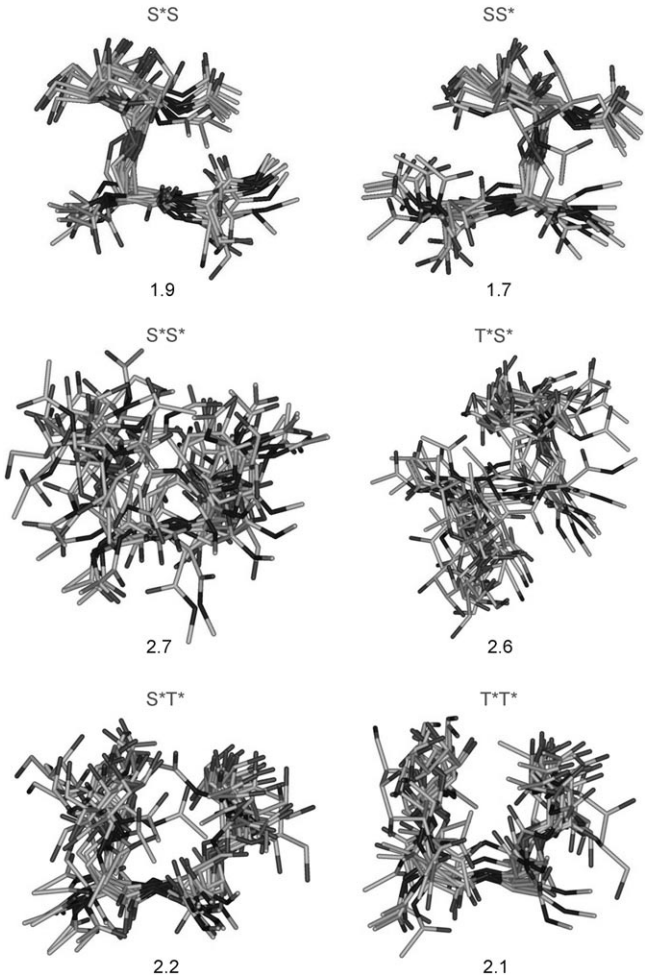


Figure 8. Calculated ensembles obtained from the MD-tar simulations for the glycopeptides. The numbers indicate the RMSDs (Å) for heavy-atom superimposition with respect to the average structure of the trajectory.

drate moiety and thus, its presentation ability. In fact, the selected sequence has a crucial influence on both the orientation and flexibility of the carbohydrate moiety. Consequently, although the ST sequence somehow fixes the spatial disposition of the Tn epitopes, the alternative analogue sequence, TS, allows for a significant degree of mobility for the carbohydrate moiety. This feature can be explained by considering the different conformational behavior of the glycosidic linkage of the Thr residue when compared with that of the Ser one. On this basis, and taking into account that these carbohydrates presumably interact with components of the immune system, these findings could be implications for designing new cancer vaccines.

Experimental Section

General procedures: Solvents were purified according to standard procedures. Analytical TLC was performed by using Polychrom SI F254 plates. Column chromatography was performed by using silica gel 60 (230–400 mesh). ^{1}H and ^{13}C NMR spectra were recorded on Bruker ARX 300 and

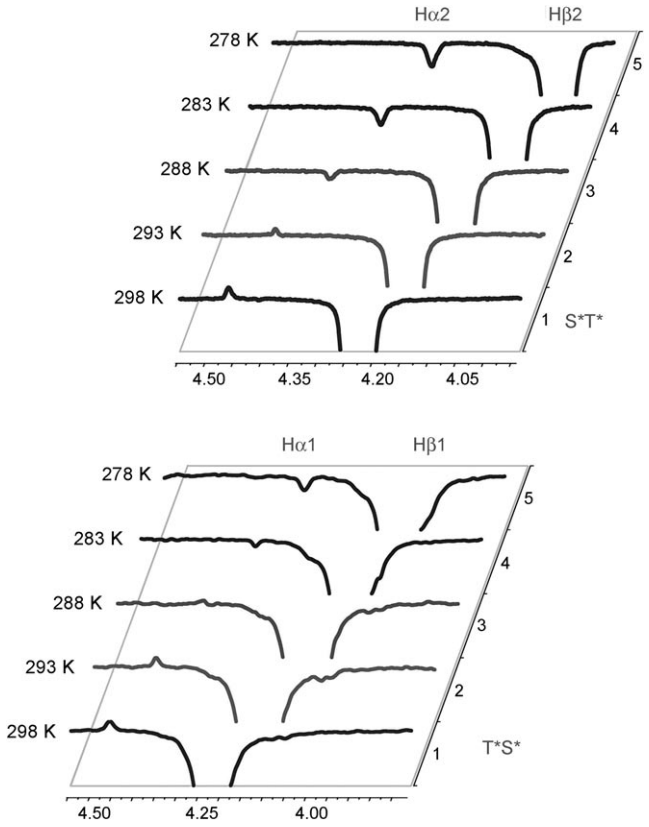


Figure 9. Selective 1D NOESY spectra of compounds **T*S*** and **S*T*** at different temperatures.

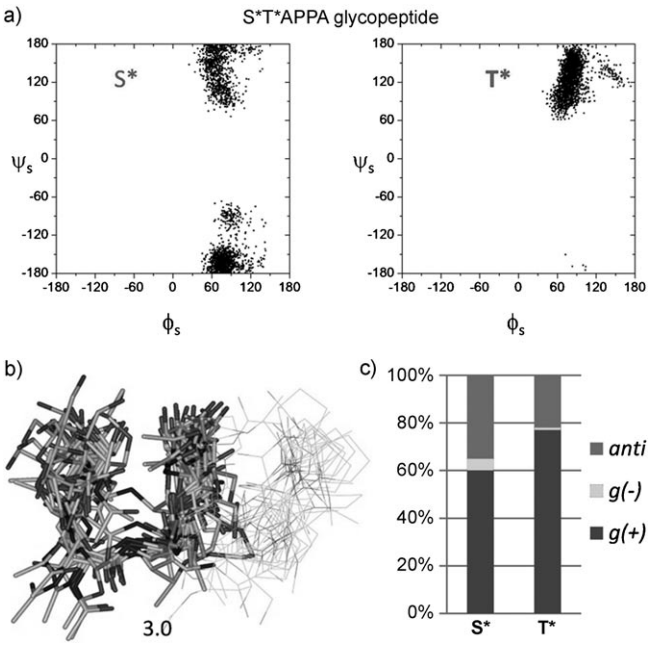


Figure 10. a) φ_s/ψ_s distribution obtained from the 80 ns MD-tar simulations for **S*T***APPA glycopeptide, showing the eclipsed conformation of the glycosidic linkage in the Thr residue. b) Calculated ensembles obtained from the MD-tar simulations for **S*T***APPA glycopeptide. The RMSDs (Å) was calculated for heavy-atom superimposition of the αGalNAc-Ser and αGalNAc-Thr residues with respect to the average structure of the trajectory. c) Distributions for the side chain (χ^1) obtained from the MD-tar simulations for the glycopeptide.

Bruker Avance 400 spectrometers. ^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$, CD_3OD, and D_2O with TMS as the external standard by using a coaxial microtube (chemical shifts are reported in ppm on the δ scale, coupling constants in Hz). Melting points were determined on a Büchi B-545 melting-point apparatus and are uncorrected. Optical rotations were measured on a Perkin-Elmer 341 polarimeter. Microanalyses were carried out on a CE Instruments EA-1110 analyser and are in good agreement with the calculated values.

Synthesis of compound 1: A mixture of allylic alcohol (20.0 mL, 0.296 mol), L-Thr (4.00 g, 33.6 mmol), *p*-TsOH (7.65 g, 40.4 mmol), and toluene (120 mL) was refluxed for 24 h by using a Dean-Stark trap and then concentrated. The crude salt was suspended in THF/H_2O (5:1, 120 mL) and $Na_2CO_3{\cdot}10H_2O$ (28.6 g, 0.10 mol) was added. The resulting suspension was cooled to 0 °C and a solution of Boc_2O (7.33 g, 33.6 mmol) in THF (50 mL) was added. The mixture was allowed to warm to 25 °C and stirred for 12 h. The solvent was then removed and the mixture partitioned between brine (40 mL) and ethyl acetate (100 mL). The organic layer was washed with 0.1 M HCl (2×20 mL) and 5% $NaHCO_3$ (2×20 mL), dried, filtered, and evaporated to give a residue, which was purified by silica gel column chromatography (hexane/ethyl acetate, 3:2) to give **1** (6.22 g, 71%) as a colorless oil. $[\alpha]_D^{25}=-13.2$ ($c=1.06$, $CHCl_3$); ^{1}H NMR (300 MHz, $CDCl_3$): $\delta=1.20$ (d, 3H, $J=6.3$ Hz), 1.40 (s, 9H), 2.89 (br s, 1H), 4.16–4.34 (m, 2H), 4.61 (d, 2H, $J=5.6$ Hz), 5.15–5.36 (m, 2H), 5.46 (d, 1H, $J=9.0$ Hz), 5.75–5.97 ppm (m, 1H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta=19.8$, 28.2, 58.8, 65.9, 67.9, 79.9, 118.6, 131.4, 156.1, 171.2 ppm; elemental analysis calcd for $C_{12}H_{21}NO_5$: C 55.58, H 8.16, N 5.40; found: C 55.17, H 8.11, N 5.39.

Synthesis of compound 2: 3,4,6-Tri-*O*-benzyl-2-nitrogalactal (1.10 g, 2.65 mmol), which was prepared as described in the literature,[17b] and Thr derivative **1** (529 mg, 2.04 mmol) were dissolved in THF (40 mL) under an argon atmosphere, and molecular sieves were then added. The reaction mixture was stirred at 25 °C for 15 min, and then a 1 M potassium *tert*-butoxide solution in THF (0.2 mL, 0.20 mmol) was added. The reaction was stirred for 12 h, the molecular sieves were then filtered off, and all solvents were removed by evaporation. The residue was purified by a silica gel column chromatography (hexane/ethyl acetate, 4:1) to give **2** (1.00 g, 68%) as a colorless oil. $[\alpha]_D^{25}=+71.5$ ($c=1.20$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta=1.23$ (d, 3H, $J=6.3$ Hz), 1.40 (s, 9H), 3.41–3.52 (m, 2H), 3.91–3.99 (m, 2H), 4.19–4.44 (m, 6H), 4.58 (d, 2H, $J=5.7$ Hz), 4.61–4.69 (m, 2H), 4.75 (d, 1H, $J=11.2$ Hz), 4.85 (dd, 1H, $J_1=10.7$ Hz, $J_2=4.2$ Hz), 4.99 (d, 1H, $J=9.8$ Hz), 5.18 (d, 1H, $J=10.4$ Hz), 5.24–5.32 (m, 2H), 5.81–5.92 (m, 1H), 7.10–7.31 ppm (m, 15H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta=18.2$, 28.3, 58.2, 66.5, 68.2, 69.9, 73.0, 73.6, 75.0, 75.0, 77.5, 80.1, 84.2, 97.5, 118.9, 127.7, 127.8, 127.9, 128.0, 128.1, 128.3, 128.4, 128.5, 131.7, 137.2, 137.6, 137.8, 156.1, 169.6 ppm; elemental analysis calcd for $C_{39}H_{48}N_2O_{11}$: C 64.99, H 6.71, N 3.89; found: C 64.86, H 6.68, N 3.86.

Synthesis of compound 3: Trifluoroacetic acid (TFA; 2 mL) was added to a solution of compound **2** (0.50 g, 0.69 mmol) in CH_2Cl_2 (5 mL) at 0 °C. The reaction was maintained at 0 °C for 30 min, at 25 °C for 2.5 h, and concentrated. The crude compound **3** was quantitatively obtained and used without further purification. ^{1}H NMR (400 MHz, CD_3OD): $\delta=$ 1.41–1.49 (m, 3H), 3.51–3.59 (m, 2H), 4.11–4.21 (m, 3H), 4.39–4.57 (m, 4H), 4.54–4.62 (m, 1H), 4.64–4.85 (m, 5H), 4.86–4.95 (m, 1H), 5.25–5.33 (m, 1H), 5.35–5.46 (m, 1H), 5.46–5.53 (m, 1H), 5.93–6.09 (m, 1H), 7.15–7.38 ppm (m, 15H); ^{13}C NMR (100 MHz, CD_3OD): $\delta=18.3$, 58.7, 68.7, 70.1, 72.0, 73.8, 74.6, 74.8, 76.0, 76.3, 76.5, 85.5, 99.2, 120.2, 128.9, 129.0, 129.1, 129.2, 129.3, 129.4, 129.5, 129.5, 132.7, 139.1, 139.3, 139.5, 167.7 ppm.

Synthesis of compound 4: A solution of derivative **2** (500 mg, 0.69 mmol) in THF (15 mL) was stirred under an argon atmosphere at 25 °C, [Pd(PPh_3)$_4$] (8 mg, 7.1×10^{-3} mmol) and morpholine (0.4 mL, 4.87 mmol) were subsequently added. After stirring for 3 h, the solvent was evaporated and the residue was taken up in 50 mL of CH_2Cl_2. The resulting solution was extracted with 1 N HCl (3×20 mL), dried, and concentrated in vacuo. The desired acid **4** was obtained as a pallid yellow oil (408 mg, 87%), which was then used directly used for the subsequent transformations. ^{1}H NMR (400 MHz, $CDCl_3$): $\delta=1.16$–1.25 (m, 3H), 1.38 (s, 9H), 3.40–3.52 (m, 2H), 3.88–4.00 (m, 2H), 4.22–4.46 (m, 6H), 4.57–4.66 (m, 2H), 4.73 (d, 1H, $J=10.7$ Hz), 4.80–4.88 (m, 1H), 5.08 (d, 1H, $J=$ 8.5 Hz), 5.30–5.39 (m, 1H), 7.12–7.30 ppm (m, 15H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta=18.0$, 28.3, 57.7, 68.2, 69.9, 72.9, 73.0, 73.5, 75.0, 75.0, 77.1, 80.2, 84.2, 97.3, 127.7, 127.8, 127.9, 128.0, 128.1, 128.3, 128.4, 128.6, 137.3, 137.6, 137.9, 156.0, 173.7 ppm.

Synthesis of compound 5: A solution of acid **4** (681 mg, 1.0 mmol) in acetonitrile (30 mL) was treated with DIEA (0.7 mL, 4.2 mmol), compound **3** (735 mg, 1.0 mmol), and (benzotriazol-1-yl)-1,1,3,3-tetramethyluronium TBTU (0.42 g, 1.3 mmol). The reaction mixture was stirred at 25 °C for 10 h and then partitioned between brine (20 mL) and ethyl acetate (50 mL). The organic layer was washed with 0.1 N HCl (2×20 mL) and 5% $NaHCO_3$ (2×15 mL). Finally, the organic layer was dried, filtered, and evaporated to give a residue that was purified by a silica gel column chromatography (hexane/ethyl acetate, 4:1) to give **5** as a colorless oil (860 mg, 67%). $[\alpha]_D^{25}=+108.0$ ($c=1.30$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta=1.19$ (d, 3H, $J=5.8$ Hz), 1.29 (d, 3H, $J=6.0$ Hz), 1.44 (s, 9H), 3.45–3.55 (m, 4H), 3.94–4.01 (m, 2H), 4.10–4.25 (m, 4H), 4.32–4.51 (m, 7H), 4.53–4.58 (m, 2H), 4.61–4.69 (m, 2H), 4.70–4.81 (m, 4H), 4.83–4.97 (m, 4H), 5.10 (dd, 1H, $J_1=10.6$ Hz, $J_2=3.2$ Hz), 5.23–5.39 (m, 2H), 5.41 (d, 1H, $J=3.7$ Hz), 5.55 (d, 1H, $J=6.2$ Hz), 5.59 (d, 1H, $J=3.0$ Hz), 5.90–6.03 (m, 1H), 6.87 (d, 1H, $J=8.7$ Hz), 7.13–7.40 ppm (m, 30H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta=16.2$, 18.0, 28.3, 56.6, 57.5, 66.7, 68.3, 68.5, 70.0, 70.1, 72.7, 72.8, 73.4, 73.5, 73.5, 73.6, 74.8, 75.0, 75.0, 75.7, 77.9, 80.0, 84.3, 85.3, 97.4, 97.6, 119.1, 127.6, 127.7, 127.8, 127.9, 128.0, 128.1, 128.2, 128.3, 128.3, 128.4, 131.7, 137.5, 137.7, 137.8, 137.9, 138.0, 138.0, 155.1, 168.7, 169.0 ppm; elemental analysis calcd for $C_{70}H_{82}N_4O_{19}$: C 65.51, H 6.44, N 4.37; found: C 65.41, H 6.42, N 4.39.

Synthesis of compound 6: The cleavage of the allyl ester of **5** was achieved by using the same procedure described for the preparation of derivative **4** with the following amounts: compound **5** (821 mg, 0.64 mmol), [Pd(PPh_3)$_4$] (7 mg, 6.4×10^{-3} mmol), and morpholine (0.37 mL, 4.48 mmol) to give the desired acid (663 mg, 83%) as a colorless oil, following the workup described for compound **4**. A solution of this acid (665 mg, 0.53 mmol) in acetonitrile (25 mL) was treated with DIEA (0.35 mL, 2.1 mmol), methylamine hydrochloride (71 mg, 1.06 mmol), and (benzotriazol-1-yl)-1,1,3,3-tetramethyluronium TBTU (204 mg, 0.64 mmol). The reaction mixture was stirred at 25 °C for 10 h, and then partitioned between brine (15 mL) and ethyl acetate (50 mL). The organic layer was washed with 0.1 N HCl (2×15 mL) and 5% $NaHCO_3$ (2× 15 mL). Finally, the organic layer was dried, filtered, and evaporated to give a residue that was purified by a silica gel column chromatography (hexane/ethyl acetate, 1:1) to give **6** as a colorless oil (514 mg, 77%). $[\alpha]_D^{25}=+$ 105.4 ($c=1.10$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta=1.15$ (d, 6H, $J=5.9$ Hz), 1.39 (s, 9H), 2.72 (d, 3H, $J=4.1$ Hz), 3.41–3.56 (m, 4H), 3.92–4.06 (m, 5H), 4.20–4.48 (m, 10H), 4.59–4.78 (m, 7H), 4.91–5.01 (m, 2H), 5.28 (d, 1H, $J=5.7$ Hz), 5.38–5.44 (m, 1H), 5.48 (d, 1H, $J=2.0$ Hz), 6.02–6.11 (m, 1H), 7.06–7.33 ppm (m, 31H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta=16.9$, 17.6, 26.4, 28.3, 56.1, 58.5, 68.1, 68.3, 70.1, 70.3, 72.8, 73.0, 73.0, 73.4, 73.5, 73.5, 74.2, 75.0, 75.0, 75.1, 75.2, 80.7, 85.0, 96.5, 97.3, 127.7, 127.8, 127.8, 128.0, 128.0, 128.1, 128.3, 128.4, 128.5, 137.3, 137.7, 137.9, 156.1, 168.2, 169.1 ppm; elemental analysis calcd for $C_{68}H_{81}N_5O_{18}$: C 65.01, H 6.50, N 5.57; found: C 64.83, H 6.47, N 5.61.

Synthesis of compound T*T*: TFA (3 mL) was added to a solution of compound **6** (0.50 g, 0.40 mmol) in CH_2Cl_2 (6 mL) at 0 °C. The reaction was maintained at 0 °C for 30 min, at 25 °C for 2.5 h and concentrated. The crude product was dissolved in pyridine (5 mL) and acetic anhydride (2 mL) was added. The resulting mixture was stirred for 3 h at 25 °C. The solvent was evaporated and the crude product purified by silica gel column chromatography (hexane/ethyl acetate, 1:4) to give **7**, as a colorless oil (382 mg, 80%). Then, Platinized Raney-nickel (T4) catalyst was freshly prepared as described in the literature.[27] The catalyst obtained by using 2 g of Raney nickel/aluminum alloy was suspended in ethanol (10 mL) and pre-hydrogenated for 10 min before the addition of compound **7** (150 mg, 0.12 mmol) in ethanol (5 mL). The reaction mixture was shaken under H_2 (1 atm) for 7 h at 25 °C. The catalyst was filtered off and the solvent evaporated. The residue was dissolved in pyridine/acetic anhydride (2:1, 6 mL) and stirred at 25 °C for 3 h. After removing

the volatile products, a solution of the crude in MeOH/ethyl acetate (2:1, 6 mL) was treated with 10% palladium–carbon (15 mg) as a catalyst. The reaction mixture was shaken under H_2 (1 atm) for 12 h at 25°C. Removal of the catalyst and a further purification of the residue first by a silica gel column chromatography (aq. NH_3/EtOH/butanol/$CHCl_3$, 8:5:4:2) and then with a C_{18} reverse-phase sep-pak cartridge gave **T*T*** as a colorless oil (27 mg, 33%). $[\alpha]_D^{25} = +96.0$ ($c = 0.82$, H_2O); ^{1}H NMR (400 MHz, D_2O): $\delta = 1.25$ (d, 3H, $J = 6.2$ Hz), 1.36 (d, 3H, $J = 6.2$ Hz), 2.03 (s, 3H), 2.10 (s, 3H), 2.14 (s, 3H), 2.72 (s, 3H), 3.72–3.80 (m, 4H), 3.85–3.95 (m, 2H), 3.97–4.01 (m, 2H), 4.01–4.08 (m, 2H), 4.08–4.15 (m, 2H), 4.24–4.32 (m, 1H), 4.34–4.42 (m, 1H), 4.55 (brs, 1H), 4.68 (brs, 1H), 4.89 (d, 1H, $J = 3.3$ Hz), 4.95 (d, 1H, $J = 3.3$ Hz) ppm; ^{1}H NMR (400 MHz, H_2O/D_2O): $\delta = 7.51$ (d, 1H, $J = 9.6$ Hz, NHs(I)), 7.91–8.00 (m, 2H, NHs(II) + NH3), 8.39 (d, 1H, $J = 8.5$ Hz, NH1), 8.62 ppm (d, 1H, $J = 9.1$ Hz, NH3); ^{13}C NMR (100 MHz, D_2O): $\delta = 18.1$, 18.3, 21.7, 22.1, 22.3, 25.8, 49.8, 49.8, 57.2, 57.7, 61.3, 67.6, 68.0, 68.5, 68.6, 71.3, 76.2, 76.8, 98.8, 99.3, 171.1, 172.0, 173.9, 174.3, 174.5 ppm; elemental analysis calcd for $C_{27}H_{47}N_5O_{15}$: C 47.57, H 6.95, N 10.27; found: C 48.98, H 6.91, N 10.23.

Synthesis of compound 9: 3,4,6-Tri-*O*-benzyl-2-nitrogalactal (1.10 g, 2.65 mmol), which was prepared as described in the literature,[17b] and Ser derivative **8**[17d] (500 mg, 2.04 mmol) were dissolved in THF (35 mL) under an argon atmosphere, and molecular sieve was then added. The reaction mixture was stirred at 25°C for 15 min and a 1 M potassium *tert*-butoxide solution in THF (0.2 mL, 0.20 mmol) was then added. The reaction was stirred for 12 h, the molecular sieve was then filtered off and all solvents were removed by evaporation. The residue was purified by a silica gel column chromatography (hexane/ethyl acetate, 4:1) to give **9** (1.07 g, 74%) as a colorless oil. $[\alpha]_D^{25} = +62.4$ ($c = 1.30$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 1.45$ (s, 9H), 3.48–3.63 (m, 2H), 3.87–3.96 (m, 2H), 3.96–4.05 (m, 2H), 4.34–4.54 (m, 5H), 4.59–4.67 (m, 2H), 4.69–4.76 (m, 2H), 4.82 (d, 1H, $J = 11.2$ Hz), 4.95 (dd, 1H, $J_1 = 10.6$ Hz, $J_2 = 4.2$ Hz), 5.23–5.38 (m, 3H), 5.47 (d, 1H, $J = 8.2$ Hz), 5.83–6.00 (m, 1H), 7.17–7.39 ppm (m, 15H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 28.3$, 53.9, 66.6, 68.1, 70.0, 72.9, 73.1, 73.6, 75.0, 75.2, 80.3, 84.1, 97.1, 119.2, 127.9, 128.1, 128.2, 128.4, 128.6, 131.6, 137.3, 137.7, 137.9, 155.4, 169.4 ppm; elemental analysis calcd for $C_{38}H_{46}N_2O_{11}$: C 64.58, H 6.56, N 3.96; found: C 64.42, H 6.58, N 3.95.

Synthesis of compound 10: TFA (2 mL) was added to a solution of compound **9** (0.50 g, 0.71 mmol) in CH_2Cl_2 (5 mL) at 0°C. The reaction was maintained at 0°C for 30 min, at 25°C for 2.5 h and concentrated. The crude compound **10** was quantitatively obtained and used without further purification. ^{1}H NMR (400 MHz, CD_3OD): $\delta = 3.44$–3.56 (m, 2H), 3.83–3.92 (m, 1H), 3.95–4.09 (m, 3H), 4.25–4.45 (m, 4H), 4.49–4.73 (m, 6H), 4.83–4.91 (m, 1H), 5.16–5.34 (m, 3H), 5.80–5.95 (m, 1H), 7.06–7.29 ppm (m, 15H); ^{13}C NMR (100 MHz, CD_3OD): $\delta = 54.1$, 67.6, 68.6, 69.9, 71.8, 73.8, 74.6, 74.7, 75.9, 76.3, 85.4, 98.4, 120.2, 128.9, 129.0, 129.1, 129.2, 129.3, 129.4, 129.5, 132.5, 139.0, 139.1, 139.4, 167.7 ppm.

Synthesis of compound 12: A solution of the commercially available acid **11** (384 mg, 1.3 mmol) in acetonitrile (30 mL) was treated with DIEA (0.7 mL, 4.2 mmol), compound **10** (721 mg, 1.0 mmol), and (benzotriazol-1-yl)-1,1,3,3-tetramethyluronium TBTU (0.42 g, 1.3 mmol). The reaction mixture was stirred at 25°C for 10 h and then partitioned between brine (20 mL) and ethyl acetate (50 mL). The organic layer was washed with 0.1 N HCl (2×20 mL) and 5% $NaHCO_3$ (2×15 mL). Finally, the organic layer was dried, filtered, and evaporated to give a residue that was purified by a silica gel column chromatography (hexane/ethyl acetate, 4:1) to give **12** as a colorless oil (860 mg, 69%). $[\alpha]_D^{25} = +63.0$ ($c = 1.40$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 1.44$ (s, 9H), 3.46–3.55 (m, 2H), 3.59 (dd, 1H, $J_1 = 9.2$ Hz, $J_2 = 6.6$ Hz), 3.82–3.89 (m, 1H), 3.89–4.00 (m, 4H), 4.30 (dd, 1H, $J_1 = 10.6$ Hz, $J_2 = 1.9$ Hz), 4.34–4.45 (m, 3H), 4.49 (d, 1H, $J = 11.8$ Hz), 4.53–4.71 (m, 6H), 4.71–4.82 (m, 2H), 4.91 (dd, 1H, $J_1 = 10.6$ Hz, $J_2 = 4.2$ Hz), 5.24–5.37 (m, 3H), 5.44–5.55 (m, 1H), 5.83–5.98 (m, 1H), 7.16–7.37 ppm (m, 21H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 28.3$, 52.7, 53.9, 66.6, 68.2, 69.3, 69.9, 70.2, 72.9, 73.3, 73.5, 74.9, 75.1, 80.1, 84.0, 97.0, 119.3, 127.6, 127.8, 127.8, 127.9, 128.0, 128.1, 128.3, 128.4, 128.5, 128.5, 131.4, 137.2, 137.5, 137.6, 137.8, 155.5, 168.6, 170.3 ppm; elemental analysis calcd for $C_{48}H_{57}N_3O_{13}$: C 65.22, H 6.50, N 4.75; found: C 65.02, H 6.51, N 4.73.

Synthesis of compound 13: The cleavage of the allyl ester of **12** was achieved by using the same procedure described for the preparation of derivative **4** with the following amounts: compound **12** (566 mg, 0.64 mmol), $[Pd(PPh_3)_4]$ (7 mg, 6.4×10^{-3} mmol), and morpholine (0.37 mL, 4.48 mmol) to give the desired acid (469 mg, 87%) as a colorless oil, following the workup described for compound **4**. A solution of this acid (447 mg, 0.53 mmol) in acetonitrile (25 mL) was treated with DIEA (0.35 mL, 2.1 mmol), methylamine hydrochloride (71 mg, 1.06 mmol), and (benzotriazol-1-yl)-1,1,3,3-tetramethyluronium (TBTU) (204 mg, 0.64 mmol). The reaction mixture was stirred at 25°C for 10 h, and then partitioned between brine (15 mL) and ethyl acetate (50 mL). The organic layer was washed with 0.1 N HCl (2×15 mL) and 5% $NaHCO_3$ (2×15 mL). Finally, the organic layer was dried, filtered, and evaporated to give a residue that was purified by a silica gel column chromatography (hexane/ethyl acetate, 1:1) to give compound **13** as a colorless oil (359 mg, 65%). $[\alpha]_D^{25} = +74.6$ ($c = 1.00$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 1.46$ (s, 9H), 2.63 (d, 3H, $J = 4.7$ Hz), 3.51–3.61 (m, 2H), 3.68 (dd, 1H, $J_1 = 9.1$ Hz, $J_2 = 6.0$ Hz), 3.74 (dd, 1H, $J_1 = 11.0$ Hz, $J_2 = 4.8$ Hz), 3.84 (dd, 1H, $J_1 = 9.5$ Hz, $J_2 = 5.1$ Hz), 3.95 (t, 1H, $J = 6.6$ Hz), 3.99–4.04 (m, 1H), 4.08–4.17 (m, 1H), 4.26–4.33 (m, 1H), 4.35 (dd, 1H, $J_1 = 10.7$ Hz, $J_2 = 2.8$ Hz), 4.41 (d, 1H, $J = 11.6$ Hz), 4.44–4.52 (m, 2H), 4.52–4.70 (m, 5H), 4.81 (d, 1H, $J = 11.2$ Hz), 4.97 (dd, 1H, $J_1 = 10.6$ Hz, $J_2 = 4.1$ Hz), 5.34 (d, 1H, $J = 4.1$ Hz), 5.40 (d, 1H, $J = 4.1$ Hz), 6.54–6.66 (m, 1H), 6.71 (d, 1H, $J = 8.2$ Hz), 7.18–7.39 ppm (m, 20H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 26.4$, 28.3, 52.9, 55.2, 67.4, 68.2, 69.8, 70.1, 72.8, 73.6, 73.7, 75.1, 80.9, 84.6, 96.7, 127.8, 127.9, 128.0, 128.1, 128.1, 128.4, 128.5, 128.7, 137.2, 137.6, 137.9, 156.1, 169.0, 170.1 ppm; elemental analysis calcd for $C_{46}H_{56}N_4O_{12}$: C 64.47, H 6.59, N 6.54; found: C 64.32, H 6.58, N 6.57.

Synthesis of compound 14: TFA (3 mL) was added to a solution of compound **13** (0.50 g, 0.58 mmol) in CH_2Cl_2 (6 mL) at 0°C. The reaction was maintained at 0°C for 30 min and then warmed to 25°C for 2.5 h and concentrated. The crude was dissolved in pyridine (5 mL) and acetic anhydride (2 mL) was added. The resulting mixture was stirred for 3 h at 25°C. The solvent was evaporated and the crude purified by a silica gel column chromatography (CH_2Cl_2/MeOH, 15:1) to give **14**, as a colorless oil (386 mg, 83%). $[\alpha]_D^{25} = +54.3$ ($c = 1.00$, $CHCl_3$); ^{1}H NMR (400 MHz, $CDCl_3$): $\delta = 1.99$ (s, 3H), 2.53 (d, 3H, $J = 4.4$ Hz), 3.44–3.52 (m, 2H), 3.58–3.68 (m, 1H), 3.68 (dd, 1H, $J_1 = 11.0$ Hz, $J_2 = 4.3$ Hz), 3.80–3.91 (m, 2H), 3.93–3.98 (m, 1H), 4.06 (dd, 1H, $J_1 = 11.0$ Hz, $J_2 = 3.2$ Hz), 4.27–4.64 (m, 10H), 4.74 (d, 1H, $J = 11.2$ Hz), 4.90 (dd, 1H, $J_1 = 10.6$ Hz, $J_2 = 3.8$ Hz), 5.25 (d, 1H, $J = 3.6$ Hz), 6.33 (d, 1H, $J = 6.1$ Hz), 6.46–6.53 (m, 1H), 6.60 (d, 1H, $J = 8.3$ Hz), 7.16–7.33 ppm (m, 20H); ^{13}C NMR (100 MHz, $CDCl_3$): $\delta = 23.2$, 26.3, 53.0, 53.6, 67.1, 68.1, 69.3, 70.1, 72.7, 72.7, 73.6, 73.7, 75.0, 75.0, 84.6, 96.5, 127.8, 127.9, 127.9, 128.0, 128.1, 128.1, 128.2, 128.3, 128.5, 128.5, 128.6, 137.0, 137.1, 137.5, 137.8, 168.9, 169.7, 171.1 ppm; elemental analysis calcd for $C_{43}H_{50}N_4O_{11}$: C 64.65, H 6.31, N 7.01; found: C 64.25, H 6.28, N 6.98.

Synthesis of compound SS*: Platinized Raney-nickel (T4) catalyst was freshly prepared as described in the literature.[27] The catalyst obtained by using 2 g of Raney nickel/aluminum alloy was suspended in ethanol (10 mL) and pre-hydrogenated for 10 min before the addition of compound **14** (200 mg, 0.25 mmol) in ethanol (5 mL). The reaction mixture was shaken under H_2 (1 atm) for 7 h at 25°C. The catalyst was filtered off and the solvent evaporated. The residue was dissolved in pyridine/acetic anhydride (2:1, 6 mL) and stirred at 25°C for 3 h. After removing the volatile products, a solution of the crude products in MeOH/ethyl acetate (2:1, 6 mL) was treated with 10% palladium–carbon (15 mg) as a catalyst. The reaction mixture was shaken under H_2 (1 atm) for 12 h at 25°C. Removal of the catalyst and a further purification of the residue first by a silica gel column chromatography (aqueous NH_3/EtOH/butanol/$CHCl_3$, 8:5:4:2) and then with a C_{18} reverse-phase sep-pak cartridge gave **SS*** as a colorless oil (52 mg, 46%). $[\alpha]_D^{25} = +51.6$ ($c = 0.76$, H_2O/MeOH, 1:1); ^{1}H NMR (400 MHz, D_2O): $\delta = 2.05$ (s, 3H), 2.08 (s, 3H), 2.77 (s, 3H), 3.73–3.78 (m, 2H), 3.81–3.93 (m, 5H), 3.96 (dd, 1H, $J_1 = 10.8$ Hz, $J_2 = 4.6$ Hz), 3.98–4.01 (m, 1H), 4.17 (dd, 1H, $J_1 = 11.0$ Hz, $J_2 = 3.5$ Hz), 4.50 (t, 1H, $J = 5.7$ Hz), 4.60–4.65 (m, 1H), 4.91 (d, 1H, $J = 3.5$ Hz); ^{1}H NMR (400 MHz, H_2O/D_2O): $\delta = 7.98$–8.04 (m, 2H, NHs(II) + NH3), 8.35 (d, 1H, $J = 6.6$ Hz, NH1), 8.55 (d, 1H, $J = 7.2$ Hz, NH2),

8.62 ppm (d, 1H, $J=9.1$ Hz, NH3); ^{13}C NMR (100 MHz, D_2O): $\delta=24.1$, 24.4, 28.3, 52.1, 56.2, 57.8, 63.4, 63.6, 69.2, 69.9, 70.8, 73.7, 100.1, 173.9, 174.7, 176.9, 176.9 ppm; elemental analysis calcd for $C_{17}H_{30}N_4O_{10}$: C 45.33, H 6.71, N 12.44; found: C 45.29, H 6.68, N 12.42.

NMR spectroscopic experiments: All the NMR spectroscopic experiments were recorded on a Bruker Avance 400 spectrometer at 293 K. ^{1}H and ^{13}C NMR spectra were recorded in $CDCl_3$ and CD_3OD with TMS as the internal standard and 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 degree pulse width. Phase-sensitive ge-2D ^{1}H-^{13}C edited-HSQC spectra[28] were recorded by using z-filter and selection before t1 removing the decoupling during acquisition by use of invigpndph pulse program with CNST2 (JHC) = 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 H_2O/D_2O (9:1) spectra. Selective ge-1D NOESY experiments were carried out by using the 1D-DPFGE NOE pulse sequence.

MD-tar simulations: MD-tar simulations were performed with AMBER 6.0[29] (parm94),[30] which was implemented with GLYCAM 04 parameters[31] to accurately simulate the conformational behavior of the sugar moiety. NOE-derived distances were included as time-averaged distance constraints and scalar coupling constants ^3J as time-averaged coupling constraints.

For each restraint (distance or coupling constant), six main constants (r1 to r4, rk2 and rk3) have to be defined. These parameters define the 'shape' of the restraining potential (Figure 11): R is the experimental distance or ^3J couplings. The values of r1 to r4 have to be specified in Å or Hz, depending on the restraint type. rk2 and rk3 are defined in kcal mol^{-1} Å^{-2} for distance restraints and in kcal mol^{-1} rad^{-2} for dihedral restraints. The experimental key proton-pair distances and ^3J couplings were implemented as structural restraints with a 10% margin by using a flat well potential as described above.

$$\begin{cases} R < \mathrm{r1} & \rightarrow \quad E = \mathrm{rk2}\,(\mathrm{r1}-\mathrm{r2})^2 + 2\,\mathrm{rk2}\,(\mathrm{r1}-\mathrm{r2})(\mathrm{r1}-R) \\ \mathrm{r1} < R < \mathrm{r2} & \rightarrow \quad E = \mathrm{rk2}\,(R-\mathrm{r2})^2 \\ \mathrm{r2} < R < \mathrm{r3} & \rightarrow \quad E = 0 \\ \mathrm{r3} < R < \mathrm{r4} & \rightarrow \quad E = \mathrm{rk3}\,(R-\mathrm{r3})^2 \\ R > \mathrm{r4} & \rightarrow \quad E = \mathrm{rk3}\,(\mathrm{r4}-\mathrm{r3})^2 + 2\,\mathrm{rk3}\,(\mathrm{r3}-\mathrm{r4})(\mathrm{r4}-R) \end{cases}$$

Figure 11. The violation energy is a well with a square bottom between r2 and r3, with parabolic sides out to a defined distance or ^3J (r1 and r4 for lower and upper bounds, respectively).

The input-file used to start a MD-tar simulation consists of two parts. The first part is an unrestrained energy minimization of 1000 steps of steepest-descents minimization followed by 1000 steps of conjugate-gradient minimization. The second part contains a 80 ns molecular dynamics simulation at 300 K in which the distance restraints are time-averaged according to R^{-6} by using a exponential memory function with τ equal to 8 ns. Similarly, scalar coupling restraints are time averaged according to J by using the same memory function. To assure a smooth run of the calculation, both types of restraints are introduced slowly over the first 1 ns of the calculation.

All the simulations were performed by using a dielectric constant of 80 to simulate the water environment. Additionally, a 20 ns MD-tar simulation in explicit water (TIP3P water molecules[32]) on glycopeptide **S*S** was carried out and obtained a similar result (see the Supporting Information).

$^3J_{H\alpha,H\beta}$ and $^3J_{NH,H\alpha}$ coupling constants were estimated by using the Karplus equations given in reference [33a, b], respectively.

Acknowledgements

We thank the Ministerio de Educación y Ciencia and FEDER (project CTQ2006–05825/BQU and Ramón y Cajal contract to F. C.) and the Universidad de La Rioja (grant to M.G.L.). We also thank CESGA for computer support. We appreciate the comments of the referees, which have resulted in an improvement in the manuscript.

[1] a) F.-G. Hanisch, S. Müller, *Glycobiology* **2000**, *10*, 439–449; b) S. J. Gendler, *J. Mammary Gland Biol. Neoplasia* **2001**, *6*, 339–353.
[2] S. H. Itzkowitz, M. Yuan, C. K. Montgomery, T. Kjeldsen, H. K. Takahashi, W. L. Bigbee, Y. S. Kim, *Cancer Res.* **1989**, *49*, 197–204.
[3] G. F. Springer, *J. Mol. Med.* **1997**, *75*, 594–602.
[4] a) K. Terasawa, H. Furumoto, M. Kamada, T. Aono, *Cancer Res.* **1996**, *56*, 2229–2232; b) J. Taylor-Papadimitriou, J. Burchell, D. W. Miles, M. Dalziel, *Biochim. Biophys. Acta* **1999**, *1455*, 301–313.
[5] a) J. N. Blattman, P. D. Greenberg, *Science* **2004**, *305*, 200–205; b) P. Moingeon, *Vaccine* **2001**, *19*, 1305–1326; c) S. J. Danishefsky, J. R. Allen, *Angew. Chem.* **2000**, *112*, 882–912; *Angew. Chem. Int. Ed.* **2000**, *39*, 836–863; d) S. Jarcz, J. Chen, L. V. Kuznetsova, I. Ojima, *Bioorg. Med. Chem.* **2005**, *13*, 5043–5054; e) S. Dziadek, A. Hobel, E. Schmitt, H. Kunz, *Angew. Chem.* **2005**, *117*, 7803–7808; *Angew. Chem. Int. Ed.* **2005**, *44*, 7630–7635; f) T. Buskas, S. Ingale, G.-J. Boons, *Glycobiology* **2006**, *16*, 113R–136R; g) D. H. Dube, C. R. Bertozzi, *Nat. Rev.* **2005**, *4*, 477–488; h) M. A. Hollingsworth, B. J. Swanson, *Nat. Rev. Cancer* **2004**, *4*, 45–60; i) C. Brocke, H. Kunz, *Bioorg. Med. Chem.* **2002**, *10*, 3085–3112; j) Y. J. Kim, A. Varki, *Glycoconjugate J.* **1997**, *14*, 569–576; k) M. A. Tarp, H. Clausen, *Biochim. Biophys. Acta* **2008**, *1780*, 546–563; l) B. Acres, J.-M. Limacher, *Expert Rev. Vaccines* **2005**, *4*, 493–502; m) F.-G. Hanisch, *Biochem. Soc. Trans.* **2005**, *33*, 705–708; n) F.-G. Hanisch, T. Ninkovic, *Curr. Protein Pept. Sci.* **2006**, *7*, 307–315; o) T. Becker, S. Dziadek, S. Wittrock, H. Kunz, *Curr. Cancer Drug Targets* **2006**, *6*, 491–517.
[6] G. Ragupathi, L. Howard, S. Cappello, R. R. Koganty, D. X. Qiu, B. M. Longenecker, M. A. Reddish, K. O. Lloyd, P. O. Livingston, *Cancer Immunol. Immunother.* **1999**, *48*, 1–8.
[7] a) S. Dziadek, C. Griesinger, H. Kunz, U. M. Reinscheid, *Chem. Eur. J.* **2006**, *12*, 4981–4993; b) D. M. Coltart, A. K. Royyuru, L. J. Williams, P. W. Glunz, D. Sames, S. D. Kuduk, J. B. Schwarz, X. T. Chen, S. J. Danishefsky, D. H. Live, *J. Am. Chem. Soc.* **2002**, *124*, 9833–9844; c) J. Schuman, D. X. Qiu, R. R. Koganty, B. M. Longenecker, A. P. Campbell, *Glycoconjugate J.* **2000**, *17*, 835–848; d) Y. Mimura, Y. Inoue, N. J. Maeji, R. Chujo, *Int. J. Pept. Protein Res.* **1989**, *34*, 363–368.
[8] D. I. R. Spencer, M. R. Price, S. J. B. Tendler, C. I. DeMatteis, T. Stadie, F.-G. Hanisch, *Cancer Lett.* **1996**, *100*, 11–15.
[9] G. I. Csonka, G. A. Schubert, A. Perczel, C. P. Sosa, I. G. Csizmadia, *Chem. Eur. J.* **2002**, *8*, 4718–4733.
[10] a) 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; b) F. Corzana, J. H. Busto, S. B. Engelsen, J. L. Asensio, J. Jiménez-Barbero, J. M. Peregrina, A. Avenoza, *Chem. Eur. J.* **2006**, *12*, 7864–7871.
[11] A. Kuhn, H. Kunz, *Angew. Chem.* **2007**, *119*, 458–462; *Angew. Chem. Int. Ed.* **2007**, *46*, 454–458.
[12] P. Braun, G. M. Davies, M. R. Price, P. M. Williams, S. J. B. Tendler, H. Kunz, *Bioorg. Med. Chem.* **1998**, *6*, 1531–1545.
[13] E. Osinaga, S. Bay, D. Tello, A. Babino, O. Pritsch, K. Assemat, D. Cantacuzene, H. Nakada, P. Alzari, *FEBS Lett.* **2000**, *469*, 24–28.
[14] R. Lo-Man, S. Vichier-Guerre, R. Perraut, E. Dériaud, V. Huteau, L. BenMohamed, O. M. Diop, P. O. Livingston, S. Bay, C. Leclerc, *Cancer Res.* **2004**, *64*, 4987–4994.
[15] 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.
[16] H. Waldmann, H. Kunz, *Liebigs Ann. Chem.* **1983**, 1712–1715.

[17] a) G. A. Winterfeld, Y. Ito, T. Ogawa, R. R. Schmidt, *Eur. J. Org. Chem.* **1999**, 1167–1171; b) G. A. Winterfeld, R. R. Schmidt, *Angew. Chem.* **2001**, *113*, 2718–2721; *Angew. Chem. Int. Ed.* **2001**, *40*, 2654–2657; c) R. R. Schmidt, Y. D. Vankar, *Acc. Chem. Res.* **2008**, *41*, 1059–1073; d) P. Baumhof, R. Mazitschek, A. Giannis, *Angew. Chem.* **2001**, *113*, 3784–3786; *Angew. Chem. Int. Ed.* **2001**, *40*, 3672–3674.
[18] a) H. Kunz, H. Waldmann, *Angew. Chem.* **1984**, *96*, 49–50; *Angew. Chem. Int. Ed. Engl.* **1984**, *23*, 71–72; b) H. Kunz, C. Unverzagt, *Angew. Chem.* **1984**, *96*, 426–427; *Angew. Chem. Int. Ed. Engl.* **1984**, *23*, 436–437.
[19] D. S. Wishart, B. D. Sykes, F. M. Richards, *J. Mol. Biol.* **1991**, *222*, 311–333.
[20] D. S. Wishart, C. G. Bigam, A. Holm, R. S. Hodges, B. D. Sykes, *J. Biol. NMR* **1995**, *5*, 67–81.
[21] H. J. Dyson, P. E. Wright, *Annu. Rev. Biophys. Biophys. Chem.* **1991**, *20*, 519–538.
[22] A. Pardi, M. Billeter, K. Wüthrich, *J. Mol. Biol.* **1984**, *180*, 741–751.
[23] K. Wuthrich, *NMR of Proteins and Nucleic Acids*, Wiley, New York, **1986**.
[24] V. F. Bystrov, *Prog. Nucl. Magn. Reson. Spectrosc.* **1976**, *10*, 41–81.
[25] D. A. Pearlman, *J. Biomol. NMR* **1994**, *4*, 1–16.
[26] G. A. Naganagowda, T. L. Gururaja, J. Satyanarayana, M. J. Levine, *J. Pept. Res.* **1999**, *54*, 290–310.
[27] S. Nishimura, *Bull. Chem. Soc. Jpn.* **1959**, *32*, 61–64.
[28] T. Parella, F. Sanchez-Ferrando, A. Virgili, *J. Magn. Reson.* **1997**, *126*, 274–277.
[29] a) D. A. Pearlman, D. A. Case, J. W. Caldwell, W. R. Ross, T. E. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel, P. A. Kollman, *Comput. Phys. Commun.* **1995**, *91*, 1–41; b) P. A. Kollman, AMBER 6, University of California, San Francisco **1999**.
[30] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, P. A. Kollman, *J. Am. Chem. Soc.* **1995**, *117*, 5179–5197.
[31] R. J. Woods, R. A. Dwek, C. J. Edge, B. Fraser-Reid, *J. Phys. Chem. B.* **1995**, *99*, 3832–3846.
[32] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, M. L. Klein, *J. Chem. Phys.* **1983**, *79*, 926–935.
[33] a) A. Marco, M. Llinas, K. Wuthrich, *Biopolymers* **1978**, *17*, 617–636; b) G. W. Vuister, A. Bax, *J. Am. Chem. Soc.* **1993**, *115*, 7772–7777.

Received: August 28, 2008
Revised: December 10, 2008
Published online: February 19, 2009

