Organic & Biomolecular Chemistry

This article is part of the

OBC 10th anniversary

themed issue

All articles in this issue will be gathered together online at

www.rsc.org/OBC10

Organic & Biomolecular Chemistry

Cite this: *Org. Biomol. Chem.*, 2012, **10**, 5916

www.rsc.org/obc

PAPER

α-*N*-Linked glycopeptides: conformational analysis and bioactivity as lectin ligands†‡§

Filipa Marcelo,[a,b] Francisco Javier Cañada,[a] Sabine André,[c] Cinzia Colombo,[d] Fabio Doro,[d] Hans-Joachim Gabius,[c] Anna Bernardi[d] and Jesús Jiménez-Barbero*[a]

Received 19th December 2011, Accepted 29th February 2012
DOI: 10.1039/c2ob07135e

Natural *N*-glycosylation involves a β-anomeric linkage connecting the sugar to one asparagine residue of the protein. We herein report NMR- and modelling-based data on glycomimetics containing α-glycosidic linkages. The bioactivity of α-Gal-containing glycopeptides has been documented by revealing binding to two plant lectins, *i.e.* a potent β-trefoil toxin (*Viscum album* agglutinin) and β-sandwich lectin (*Erythrina corallodendron* agglutinin), by NMR protocols. Docking provided insights into the 3D structures of the resulting complexes. These results provide the basis to introduce α-substituted neoglycopeptides to the toolbox of scaffold for the design of potent lectin inhibitors.

Introduction

Based on the fundamental importance of cellular glycans for coding biological information, analysis of their interaction with specific receptors (lectins) is of continued interest.[1] Viewed from the perspective of the design of potent lectin inhibitors, analogues of glycopeptides offer the potential for affinity increases by tailoring the peptide portion as well the sugar moiety.[2] Physiologically, the criteria of rigidity, planarity and linearity make the common β-*N*-glycosidic linkage of GlcNAc to asparagine the common acceptor for *N*-glycosylation in the three kingdoms of life.[3] In addition, β-linkage to Glc is also known, as are conjugation to L-rhamnose and GalNAc.[4] Starting with the analysis of nephritogenoside, a 21-amino-acid peptide with a trisaccharide (Glcα1,6Glcβ1,6Glcα1,N),[5] attention has been drawn to the possibility for α-linked *N*-glycosylation, also computationally.[6] In fact, it is possible to find some glycoprotein structures in the PDB (3RY6, 1U65, 3IYW) where the *N*-glycan is formally attached in α-configuration, however the low crystallographic resolution associated to the unusual geometry of the sugar attached to Asn makes this configuration more than suspicious. Since a β-linked galactoside has been shown to maintain lectin-binding properties,[2d] the pertinent question on properties of respective α-*N*-linked compounds has prompted this study. Based on initial synthetic work[7] it was possible to take the next step to address the issue of analysing binding properties of α-*N*-linked glycopeptides carrying a galactose moiety.

In this report, NMR-based experiments,[8] in combination with molecular modelling and docking protocols, have been applied to investigate the 3D solution structure of two model α-*N*-linked galactosyl-peptides. To reveal affinity for lectins the interaction of glycopeptides **1** and **2** (Fig. 1) was tested with the plant toxin *Viscum album agglutinin* (VAA), also known as a model for lectin drug design,[9] and the *Erythrina cristagalli* agglutinin (from coral tree, ECA).[10] Of note, the two lectins have different folding (β-trefoil and β-sandwich), and their contact sites for the

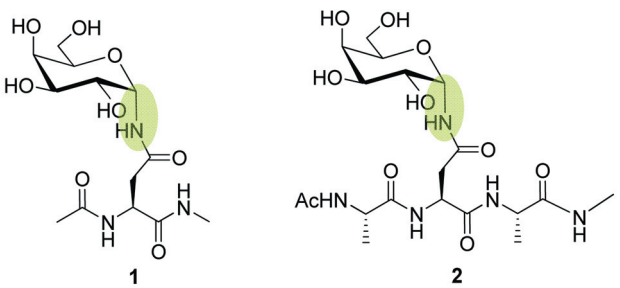


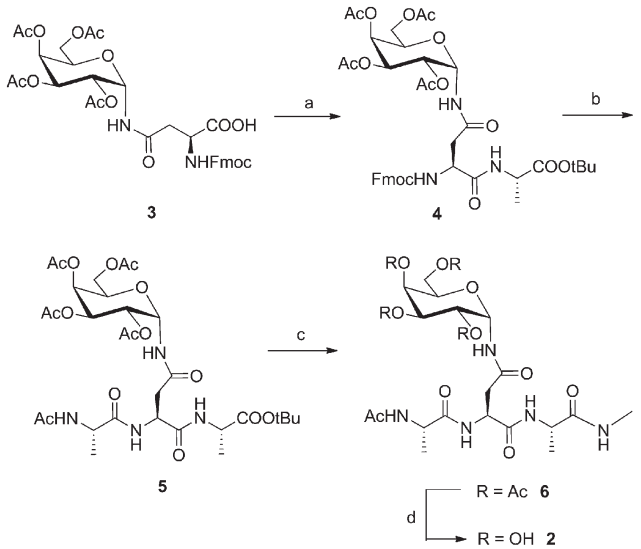
Fig. 1 α-*N*-Glycopeptides **1** and **2**.

[a]*Chemical and Physical Biology, Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Ramiro de Maeztu 9, 28040 Madrid, Spain. E-mail: jjbarbero@cib.csic.es; Fax: +34915360432; Tel: +34918373112 ext:4370*
[b]*REQUIMTE, CQFB, Departamento de Química, Faculdade de Ciências e Tecnologia, UNL, 2829-516 Caparica, Portugal*
[c]*Institute of Physiological Chemistry, Faculty of Veterinary Medicine, Ludwig-Maximilians-University Munich, Veterinärstr. 13, 8053 Munich, Germany*
[d]*Dipartimento di Chimica Organica e Industriale, Universita' degli Studi di Milano, via Venezian 21, 20133 Milano, Italy*

†This article is part of the *Organic & Biomolecular Chemistry* 10th Anniversary issue.

‡Electronic supplementary information (ESI) available: NMR spectra, STD competition binding experiments and Docking models of the glycopeptide **1** and **2** in presence of VAA and ECA. See DOI: 10.1039/c2ob07135e

§Dedicated to Prof. M. A. Miranda on the occasion of his 60th birthday.



Scheme 1 Synthesis of the model glycopeptide Ac-Ala-Asn-(α-*N*-Gal)-Ala-NHMe$_2$. Reagents and conditions: (a) 1. H-Ala-O*t*Bu·HCl, PyBROP, iPr$_2$NEt, CH_2Cl_2, 0 °C to room temperature, 83%; (b) 1. octanethiol, cat. DBU, THF; 2. Ac-Ala-OH, HATU, iPr$_2$NEt, DMF, 65% over two steps; (c) 1. TFA, CH_2Cl_2; 2. NH$_2$Me·HCl, PyBROP, iPr$_2$NEt, DMF, 54%; (d) cat. K_2CO_3, MeOH, 85%.

sugar are structurally well-defined, recently also for VAA at physiological pH.[11] Experimentally, STD-NMR[12] and trNOESY[13] protocols were used. Moreover, both proteins have proven to be robust and reliable model systems for interaction studies under the given conditions[14] so that the combination of NMR-derived data with docking to visualize 3D structures for the glycopeptides–lectin complexes is feasible.

Results and discussion

Synthesis

The synthesis of Ac-Asn-(α-*N*-Gal)-NHMe **1** has previously been reported.[7] The model glycopeptide Ac-Ala-Asn-(α-*N*-Gal)-Ala-NHMe **2** was synthesized in solution starting from N_α-fluorenylmethoxycarbonyl-N_γ-tetra-*O*-acetyl-α-D-galactopyranosyl-L-asparagine **3** (Scheme 1)[7] by coupling with alanine *t*-butylester hydrochloride, using PyBROP (bromotripyrrolidinophosphonium hexafluorophosphate) as the coupling agent. The reaction produced **4** in 83% yield (Scheme 1), which was subjected to Fmoc removal in solution, using octanethiol (10 equiv.) and a catalytic amount of DBU (0.5 equiv.). The crude product, washed several times with cold diethylether and pentane 1 : 1, was used without further purification for the coupling reaction with *N*-acetylalanine with HATU (*O*-(7-azabenzotriazol-1-yl)-*N*,*N*,*N*′,*N*′-tetramethyluronium hexafluorophosphate) to afford **5** in 65% yield over two steps. Removal of the *tert*-butyl ester from **5** with TFA in dichloromethane and subsequent reaction with methylamine and PyBROP afforded **6** (54% yield over two steps). Removal of the acetyl groups from the carbohydrate moiety was performed with catalytic amounts of K_2CO_3 in MeOH under carefully controlled pH conditions (pH 8–9) to give **2** in 85% yield.

Conformational analysis of the free α-*N*-linked glycopeptides

The solution conformation of the two α-*N*-linked glycopeptides **1** and **2** was first investigated by NMR spectroscopy and molecular mechanics calculations. Coupling constants and NOE data were carefully analysed and enabled to determine the conformational features of the peptide backbone in water solution (Fig. 2 and Fig. S11 and S12 in ESI‡). The large experimental $^3J_{NH,H\alpha}$ coupling constants values along the peptide chain strongly suggested the presence of an extended conformation for the peptide backbone in solution. Nevertheless, the $J_{H\alpha,H\beta1}$/$J_{H\alpha,H\beta2}$ values (5.1/6.5 Hz (for **1**) and 8.1/7.3 Hz (for **2**) of Asn amino acid revealed that compound **1** somehow presents more flexibility around χ1 (Hα-Cα-Cβ-Hβ) than glycopeptide **2**. The absence of non-vicinal medium-range NOE contacts supported the notion that glycopeptide **2** adopts, as its main conformation, an extended form of the peptide backbone when free in solution (Fig. 2 and 3). Inter-residual NOEs contacts between the side chain NH of Asn amino acid (NH1 on the NMR spectra) and H3 and H5 of the galactose residue were also detected (Fig. 2 and Fig. S13 in ESI‡), indicating that the galactose ring adopts the usual $^4C_1$ chair conformation and that the NH bond is buried below the ring. The inspection of these NOE cross-peaks disclosed a certain degree of flexibility around the glycosidic linkage. Noteworthy, in the case of glycopeptide **2**, the NOE between NH of Asn (NH1) and H5 appears to be stronger when compared to that observed between NH of Asn (NH1) and H3 proton (Fig. 2), evidence for a preferred orientation of this linkage. In addition, the high value for the glycosidic coupling constant ($^3J_{H1,NH}$ = 8.3 Hz for **1**, and 8.2 Hz for **2**) denotes the existence of a major *anti*-type orientation between the NH of Asn (NH1) and H1 of the galactose. An unconstrained conformational analysis of both α-*N*-glycopeptides was carried out using MC/EM[15] calculations with AMBER* force field and a GB/SA continuum water model,[16] as implemented in Macromodel.[17] These calculations came up with mostly extended conformations of the peptide chains, despite the known tendency of the force field to overestimate folded conformations such as γ-turns for small peptides.[18] The same results were also obtained from dynamic simulations of **1** and **2** performed both with implicit (Macromodel, MC/SD,[19] GB/SA water model) and explicit water (AMBER 9,[20] TIP3P water, periodic boundary conditions). This structural feature, which is agreement with the NMR data, may be favoured by the formation of H-bond interaction between the sugar and the peptide chain. Fig. 3 shows representative low-energy conformations calculated for **2** and highlights the intramolecular H-bonds predicted. To the best of our knowledge, no data are available on the preferred conformation of the Ala-Asn-Ala (ANA) tripeptide, whereas many AXA tripeptides have been found to exist mostly in extended conformation. A study published in 2004, for example, showed that AXA tripeptides (X being valine, tryptophan, histidine, and serine) predominantly adopt an extended β-strand conformation, while AXA tripeptides, for which X is lysine and proline, prefer a polyproline II-like (PPII) structure.[21]

In contrast, the Ala-Phe-Ala sequence was found to fold as an inverse γ-turn in water.[22] Therefore, it is not unambiguously clear whether the extended conformation observed for **1** and **2** is a direct consequence of peptide glycosylation. One additional

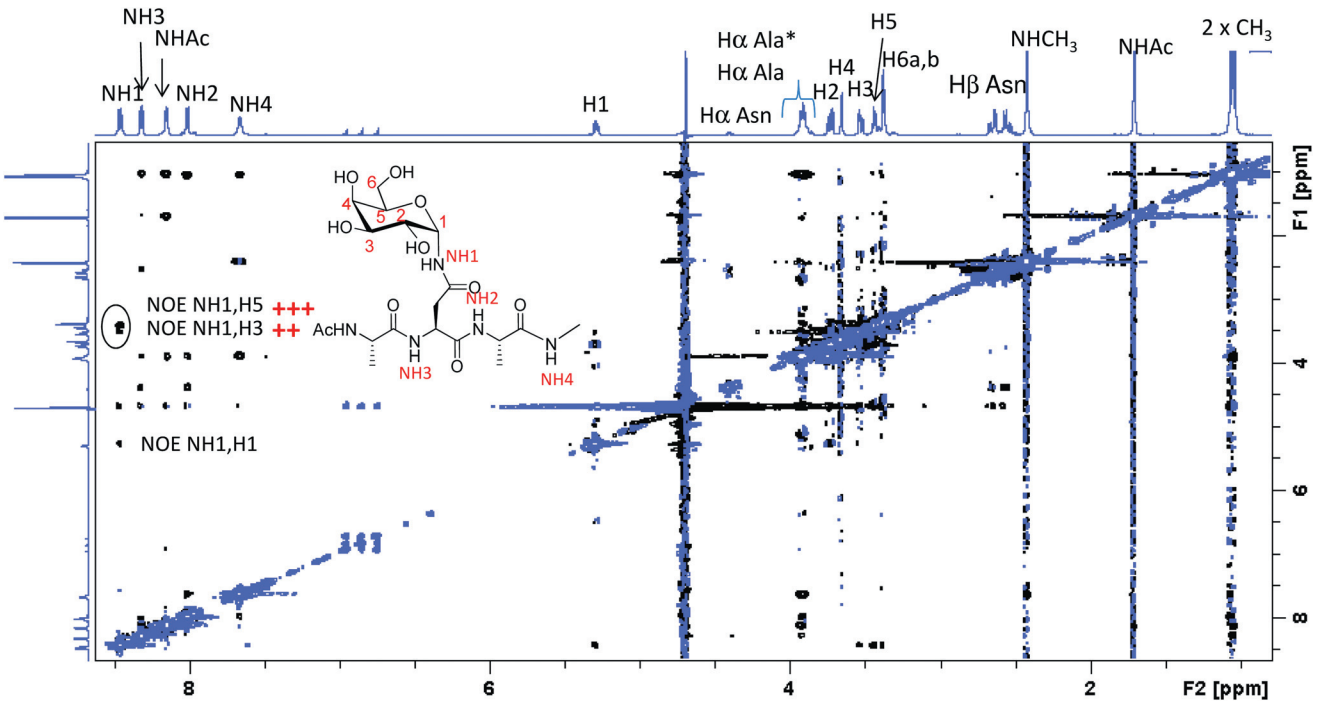


Fig. 2 2D-NOESY spectrum obtained for glycopeptide **2** in H_2O–D_2O 90 : 10 recorded at 500 MHz with 600 ms of mixing time and at 278 K.

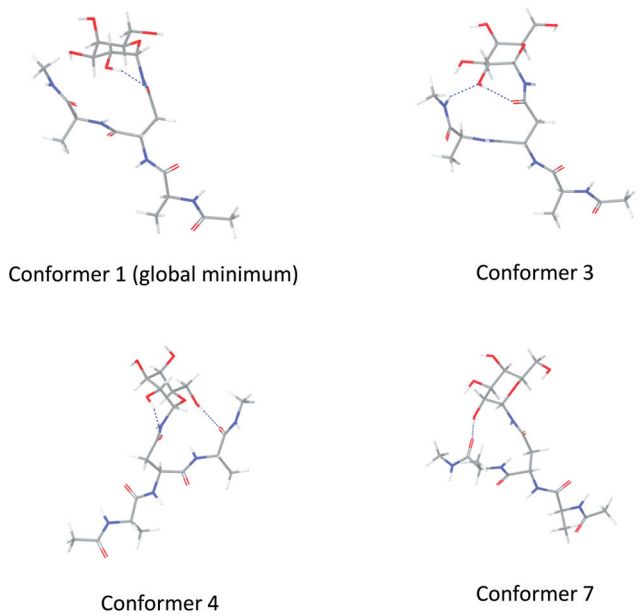


Fig. 3 Representative low-energy conformations (within 2 kcal mol^{-1} from the global minimum) from the conformational search of **2**, illustrating the extended conformation of the peptide and the H-bond interactions predicted to occur between the sugar and the peptide chain.

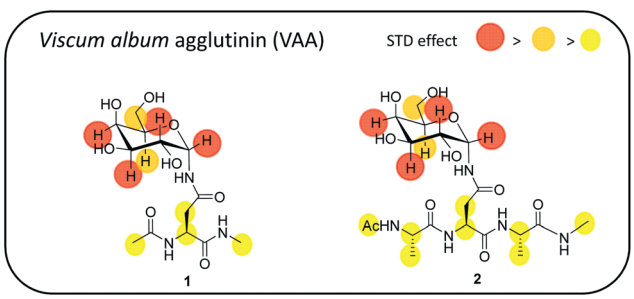


Fig. 4 Epitope mapping obtained for glycopeptides **1** and **2** with VAA.

experimental feature well-reproduced in the calculations is the orientation around the glycosydic linkage, which agrees well with the observed NOE and NH–H1 coupling constants. In contrast, the conformational flexibility of the χ1 bond is not represented by the calculations, which predict exclusively the existence of *anti* (N-Cα-Cβ-Cγ) conformations. Statistical analysis of the occurrence of conformers for side chains from X-ray structures of proteins shows that the frequency for the *anti*-conformation of Asn χ1 is similar as the one for the *gauche*(−) conformation.[23] On the contrary, rotamer libraries generated by computational simulations of short peptides in explicit water show a large prevalence of the *anti*-conformation of the χ1 bond,[24] a characteristic also observed in our conformational searches.

The bound conformation of the α-*N*-linked glycopeptides to galactose-specific proteins

Having performed structural analysis on the free compounds, their ligand properties were studied for the two test lectins. Of note, coming from different families, the vicinity of the primary docking site for galactose is different for VAA and ECA. First, STD-NMR data were collected for the individual glycopeptides to establish their respective binding epitopes to the two receptors.

Clear STD signals were detected for both glycopeptides in the presence of both lectins (Fig. 4 and 5 and Fig. S16 in ESI‡). To

prove specificity, STD-based competition experiments were carried out with both glycopeptides and the proteins, in the presence of lactose. Thus, the interference of lactose on the recognition of glycopeptides by both lectins was demonstrated by assessing the decrease of STD intensity of glycopetide's proton signals after addition of lactose. STD competition binding experiments clearly indicated that the neoglycopeptides (**1** and **2**) and lactose (chosen as "natural" ligand) indeed compete for the same binding site (Fig. S17–S19 in ESI‡). The estimated binding affinities for **1** and **2** were 0.93 and 0.89 mM, respectively. As can be seen in Fig. 4 and 5, the most intense signals corresponded, in all cases, to the sugar moiety, highlighting the predominant role of the carbohydrate moiety as central binding epitope to the lectins. The detailed epitope mapping of **1** and **2** was then achieved by normalizing the observed STD signal intensities with respect to that of the highest STD response. Reflecting structural differences between the lectins, some subtle disparities in the molecular recognition features of both α-*N*-glycopeptides by VAA and ECA were detected. In particular, for VAA, Gal H4 was the proton receiving more percentage of saturation, followed by Gal H2 and Gal H3. In contrast, for ECA, Gal H1 and H2 received the highest extent of saturation, followed by Gal H3. The differences on the binding epitopes of both α-*N*-glycopeptides **1** and **2** match perfectly with the respective epitope mapping obtained for the corresponding natural ligand, in the presence of VAA and ECA, respectively. These results demonstrate that the sugar moiety of these non-natural glycopeptides is properly recognized by both lectins, but with slightly different contacts.

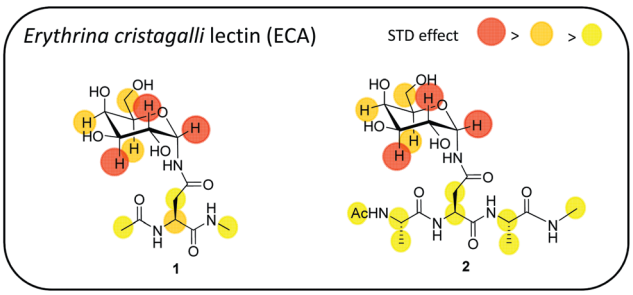


Fig. 5 Epitope mapping obtained for glycopeptides **1** and **2** with ECA.

Furthermore, for both lectins, the transfer of magnetization was rather uniform to all Hα protons of glycopeptide **2**. This observation suggests that the peptide backbone adopts an extended conformation in the bound state. Regarding interplay of the peptide chain of compound **2** with the two proteins, slight differences in the binding epitope were also recorded. The relative STD responses for Asn Hα, Hβ1 and Hβ2, as well as for Ala Hα were significantly higher in the presence of ECA than with VAA (Fig. 4 and 5 and Fig. S16 in the ESI‡). Moving beyond the STD experiments, the TR-NOESY approach was tested. Due to the relatively small size of ECA (dimer of 28 763 Da per subunit),[10c] the corresponding experiments were not completely satisfactory and will thus not be discussed. In contrast, for VAA as $(AB)_2$ tetramer of 119.2 kDa in solution,[9c] strong negative NOE were observed, indicating efficient recognition (Fig. 6 and S20 in the ESI‡). The comparison of the NOE-cross peak pattern of glycopeptides **1** and **2** in the free and VAA-bound states indicated differences in the NOE intensities of the proton pairs that define the conformation of the glycosidic linkage (Fig. 2 and 6; Fig. S13 and S20 in the ESI‡). In the bound state, the Gal NH–H3 NOE is now stronger than the Gal NH–H5 NOE (NH1–H3 and NH1–H5 on the NMR spectra, respectively), the opposite pattern to that observed when free in

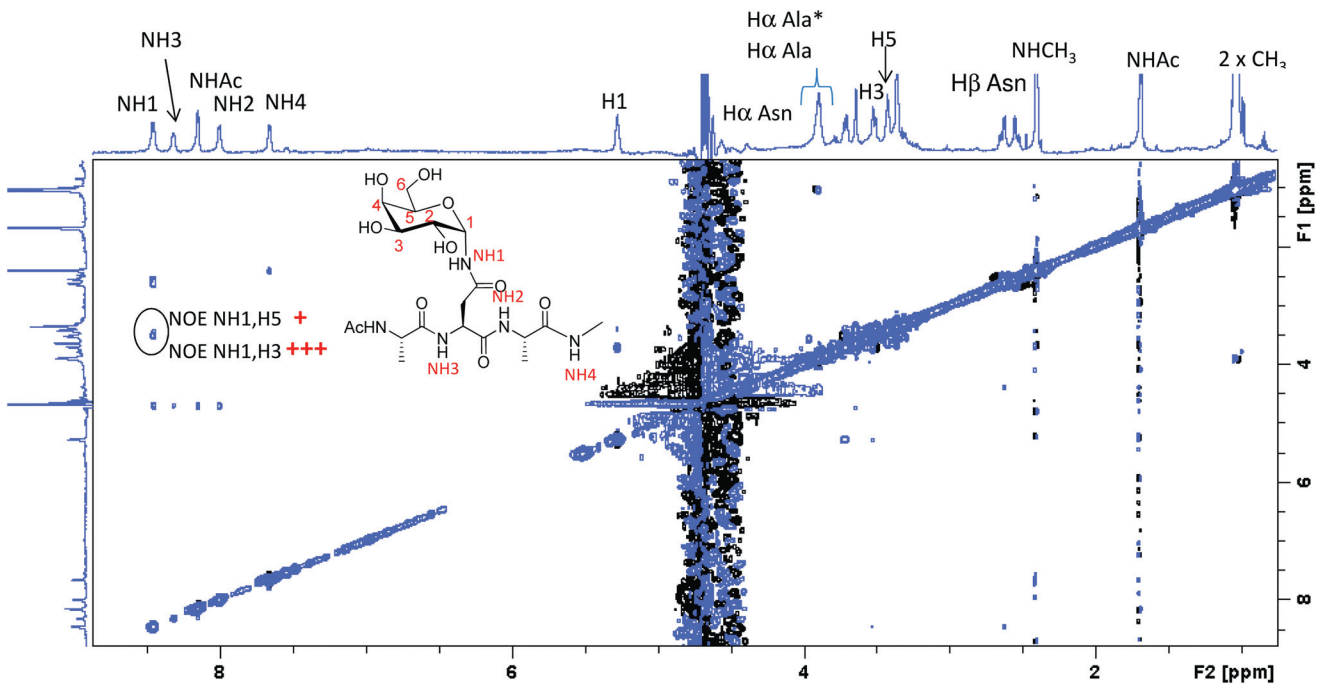


Fig. 6 trNOESY spectrum recorded at 600 MHz with 200 ms of mixing time of glycopeptide **2** in the presence of VAA (NaPi 20 mM, NaCl, 100 mM, pH = 7.4 at 278 K), with a glycopeptide–VAA molar ratio of 7 : 1. All cross-peaks present the same sign (negative) as the diagonal.

solution. Thus, a conformational selection process around the glycosidic linkage of **1** and **2** is taking place upon binding to VAA. Fittingly, the NOE cross-peak patterns also provided salient information on different rotational mobility around the different regions of the molecules. The intra-residual NOESY cross-peaks for the Gal and Asn residues were clearly negative, while those for the Ala residues were close to zero. This observation indicates that the major contacts with lectin are provided by the sugar and the Asn moiety, while the rest of the peptide chain only provides transient interactions with the receptor. Indeed, a much faster effective rotational correlation occurs for the peptide backbone than for the sugar and Asn residues. This result is in perfect accord with the STD-NMR data.

Molecular modelling

In order to generate a 3D model of the complex formed between the two neoglycopeptides (**1** and **2**) and the lectins docking was performed. The docking calculations were performed using AutoDock4.2.[25] The starting coordinates for the glycopeptides were those obtained after energy minimization of the NMR-based structure generated in the Schrödinger suite of programs. Then, the 3D structures of **1** and **2** were docked into the binding site of VAA (pdb code: 1PUM) and ECA (pdb code: 1GZC). The obtained Autodock-based poses were carefully compared to the available STD and trNOESY data, enabling to select those which satisfactorily correlated with the NMR experimental data. In solution, lectin functionality is attributed to the Tyr249 site at subdomain 2γ.[26] As example, Fig. 7 shows the 3D model of the glycopeptide **2**–VAA complex, highlighting the key intermolecular interactions. Indeed, they are basically identical to those observed in the presence of lactose in X-ray crystallographic analysis. There are hydrogen bonds involving Asp235, Asn256, Gln238, with C–H···π interaction[27] between the non-polar face of galactose and Tyr249. Answering the question on relevance of the peptide portion, no further H-bonds were observed between the peptide backbone and the protein surface.

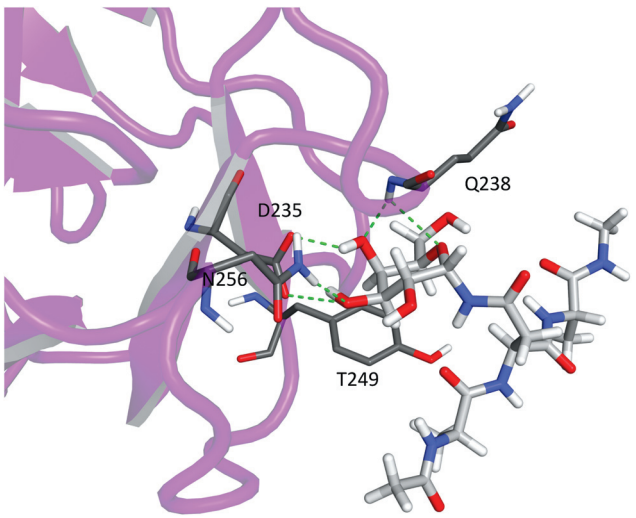


Fig. 7 3D model of the interaction of glycopeptide **2** and VAA protein deduced by combination of NMR and modelling. The H-bonds interactions are depicted as green dashed lines.

Obviously, this peptide is therefore less conducive to make contacts than anomeric extensions such as an isoxazole extended by furan or a 2-benzothiazolyl moiety.[28] This 3D structure fits well with the STD and NOE data. Of note, it shows an extended conformation for the peptide backbone. Following the same methodology, different models were deduced for the other three complexes (Fig. S21–S23 in the ESI‡), which show similar interaction features (H-bonds and CH···π interaction) between the glycopeptide and the lectins.

Conclusions

The structural properties of two α-*N*-linked glycopeptides (with different extents of sequence length of the peptide portion) were defined. The bioactivity of the sugar unit was disclosed, and the strategic combination of NMR and modelling protocols enabled the deduction, in a non-ambiguous manner, of the key factors involved in the recognition of these glycopeptides by selected lectins. α-*N*-Linked neoglycopeptides warrant further studies taking these models as scaffolds for the design of compounds to target medically relevant lectins with extended sites that could potentially interact with the peptide portion in similar fashion as indicated for human galectins.[1c,2d,29]

Experimental section

Synthesis

General: Dichloromethane, methanol and *N*,*N*-diisopropylethylamine (DIPEA) were dried with calcium hydride, *N*,*N*-dimethylformamide (DMF) was dried with activated molecular sieves (3 Å). Chemicals, including coupling agents (PyBROP and HATU), were purchased from Sigma Aldrich. Amino acids (H-Ala-O*t*Bu·HCl; Ac-Ala-OH) were purchased from Bachem. Reactions requiring anhydrous conditions were performed under nitrogen. ^{1}H and ^{13}C NMR spectra were recorded at 400 MHz with a Bruker AVANCE-400 instrument. Chemical shifts (δ) are expressed in ppm relative to internal Me_4Si as standard. Signals are abbreviated as follows: s, singlet; br. s, broad singlet; d, doublet; t, triplet; q, quartet; m, multiplet. Mass spectra were recorded with a Bruker ion-trap Esquire 3000 apparatus (ESI ionization) or a FT-ICR APEX II mass spectrometer and Xmass 4.7 Magnet software (Bruker Daltonics). Thin-layer chromatography (TLC) was carried out with precoated Merck F254 silica gel plates. Flash chromatography was carried out with Macherey–Nagel silica gel 60 (230–400 mesh) or with Biotage® SNAP KP-C18-HS cartridges for reversed-phase chromatography. HPLC-MS analyses were performed with an Agilent 1100 instrument with a quaternary pump, diode array detector, autosampler, thermostatted column holder coupled to a Bruker ion-trap Esquire 3000 mass spectrometer equipped with ESI ionization.

***N*α-Fluorenylmethoxycarbonyl-*N*γ-(2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranosyl)-L-asparagyl-L-alanine *tert*-butyl ester (4).** Compound **3**[7] (100 mg, 0.146 mmol, 1 equiv.), H-Ala-O*t*Bu hydrochloride (58 mg, 0.321 mmol, 2.2 equiv.) and PyBROP (150 mg, 0.321 mmol, 2.2 equiv.) were dissolved in dry CH_2Cl_2 (1.5 mL) under nitrogen at 0 °C. DIPEA (132 μL, 0.760 mmol,

5.2 equiv.) was added and the reaction mixture was stirred at 0 °C for 1 h and then at room temperature for 3 h (TLC, 8 : 2 chloroform–methanol and 4 : 6 hexane–EtOAc). The solvent was evaporated, the residue was dissolved in EtOAc and the organic phase was washed with 1 M HCl and saturated $NaHCO_3$ and then dried with sodium sulfate. The solvent was evaporated and the crude was purified by flash chromatography (4 : 6 hexane–EtOAc) to afford **4** (99 mg) in 83% yield. $[\alpha]_D^{20}$ = +65.3 (c = 0.5, MeOH). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 7.76 (d, J = 7.4 Hz, 2 H, 4-H-, 5-H-Fmoc), 7.60 (d, J = 7.3 Hz, 2 H, 1-H-, 8-H-Fmoc), 7.40 (t, J = 7.4 Hz, 2 H, 3-H-, 6-H-Fmoc), 7.36–7.30 (m, 2 H, Gal-NH-Asn, Asn-NH-Ala), 7.31 (t, J = 7.4 Hz, 2 H, 2-H-, 7-H-Fmoc), 6.38 (br. s, 1 H, NH-Fmoc), 5.89 (dd, $J_{1,NH}$ = 7.6, $J_{1,2}$ = 5.2 Hz, 1 H, 1-H), 5.43–5.40 (m, 1 H, 4-H), 5.38 (dd, J_{2-3} = 10.9, $J_{2,1}$ = 5.2 Hz, 1 H, 2-H), 5.34–5.28 (m, 1 H, 3-H), 4.54 (br. s, 1 H, α-H-Asn), 4.46–4.34 (m, 3 H, CH_2-Fmoc, α-H-Ala), 4.26–4.20 (m, 2 H, 9-H-Fmoc, 5-H), 4.16–4.00 (m, 2 H, 6-H), 2.94–2.85 (m, 1 H, β-CH_2-Asn), 2.75–2.64 (m, 1 H, β-CH_2-Asn), 2.14 (s, 3 H, CH_3CO), 2.02 (s, 3 H, CH_3CO), 2.00 (s, 6 H, CH_3CO), 1.45 (s, 3 H, O*t*Bu), 1.37 (d, J = 7.2 Hz, 3H, CH_3-Ala) ppm. ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 171.6–169.7 (CO), 143.9, 143.8 (C_{quat}Fmoc), 141.5 (C_{quat}F-moc), 128.0 (C-2-, C-7-Fmoc), 127.3 (C-3-, C-6-Fmoc), 125.3 (C-1-, C-8-Fmoc), 120.2 (C-4-, C-5-Fmoc), 82.3 (C_{quat}-O*t*Bu), 75.1 (C-1), 67.7 (CH_2-Fmoc), 67.7, 67.4 (C-2, C-3, C-4), 66.2 (C-5), 61.6 (C-6), 51.6 (α-C-Asn), 49.3 (α-C-Ala), 47.3 (C-9-Fmoc), 38.5 (β-CH_2-Asn), 28.2 (CH_3O*t*Bu), 20.8 (4 × OAc), 18.2 (CH_3-Ala) ppm. MS (ESI): m/z = 834.3 $[M + Na]^+$. FT-ICR MS (ESI): calcd for $[C_{40}H_{49}N_3O_{15}Na]^+$ 834.30559; found 834.30448.

***N*α-(L-*N*-Acetylalanyl)-*N*γ-(2,3,4,6-tetra-*O*-acetyl-α-D-galacto-pyranosyl)-L-asparagyl-L-alanine *tert*-butyl ester (5).** Compound **4** (99 mg, 0.122 mmol, 1 equiv.) was dissolved in dry THF (1.2 mL) under nitrogen. Octanethiol (211 µL, 1.219 mmol, 10 equiv.) and DBU (10 µL, 0.070 mmol, 0.5 equiv.) were added sequentially and the reaction mixture was stirred for 1 h (TLC, 9 : 1 chloroform–methanol and 2 : 8 hexane–EtOAc). The solvent was evaporated. The residue was washed thoroughly with a mixture of cold diethyl ether and pentane (1 : 1). The crude solid was then dissolved in dry DMF (1.5 mL) with Ac-Ala-OH (40 mg, 0.305 mmol, 2.5 equiv.) and HATU (93 mg, 0.244 mmol, 2 equiv.) under nitrogen at 0 °C. DIPEA (75 µL, 0.427 mmol, 3.5 equiv.) was added and the reaction mixture was stirred at 0 °C for 1 h and then at room temperature for 4 h (TLC, 9 : 1 chloroform–methanol). The solvent was evaporated, the residue was dissolved in EtOAc and the organic phase was washed with 1 M HCl and saturated $NaHCO_3$ and then dried with sodium sulfate. The solvent was evaporated and the crude was purified by flash chromatography (95 : 5 chloroform–methanol) to afford **5** (55 mg) in 65% yield. $[\alpha]_D^{20}$ = +57.7 (c = 1.05, CH_2Cl_2). ^{1}H NMR (400 MHz, $CDCl_3$, 25 °C): δ = 8.28 (d, J = 8.6 Hz, 1 H, Gal-NH-Asn), 8.08 (d, J = 8.6 Hz, 1 H, Gal-NH-Ala), 7.51 (d, J = 7.4 Hz, 1 H, NHAsn), 7.47 (d, J = 7.4 Hz, 1 H, NHAc), 6.54 (d, J = 7.4 Hz, 1 H, NHAla), 5.93 (dd, $J_{1,NH}$ = 8.6, $J_{1,2}$ = 5.3 Hz, 1 H, 1-H), 5.47–5.31 (m, 3 H, 4-H, 2-H, 3-H), 4.78–4.70 (m, 1 H, α-H-Asn), 4.48–4.39 (m, 1 H, α-H-AlaO*t*Bu), 4.38–4.32 (m, 2 H, α-H-AcAla, 5-H), 4.14 (dd, $J_{6-6'}$ = 11.0, J_{6-5} = 7.3 Hz, 1 H, 6-H), 4.03 (dd, $J_{6-6'}$ = 11.0, $J_{6'-5}$ = 6.1 Hz, 1 H, 6′-H), 2.81 (dd, J_{gem} = 15.4, $J_{\alpha,\beta}$ = 4.1 Hz, 1 H, β-HAsn), 2.73 (dd, J_{gem} = 15.4, $J_{\alpha,\beta}$ = 6.2 Hz, 1 H, β-H-Asn), 2.14 (s, 3 H, CH_3CO), 2.06 (s, 3 H, CH_3CO), 2.04 (s, 6 H, CH_3CO, $NHCH_3$), 2.00 (s, 3 H, CH_3CO), 1.45 (s, 3 H, O*t*Bu), 1.40 (d, J = 7.1 Hz, 3H, CH_3-Ala), 1.36 (d, J = 7.2 Hz, 3H, CH_3-Ala) ppm. ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 172.9–169.9 (CO), 82.2 (C_{quat}-O*t*Bu), 75.0 (C-1), 67.9 (C-2), 67.7 (C-3, C-4), 66.6 (C-5), 61.6 (C-6), 50.0 (α-C-Asn), 50.0 (α-C-Ala), 49.4 (α-C-Ala) 38.7 (β-CH_2-Asn), 28.2 (CH_3O*t*Bu), 23.1 ($NHCH_3$), 20.9–20.8 (4 × OAc), 18.1 (CH_3-Ala) ppm. MS (ESI): m/z = 725.2 $[M + Na]^+$. FT-ICR MS (ESI): calcd for $[C_{30}H_{46}N_4O_{15}Na]^+$ 725.28519; found 725.28404.

***N*α-(L-*N*-Acetylalanyl)-*N*γ-(2,3,4,6-tetra-*O*-acetyl-α-D-galacto-pyranosyl)-L-asparagyl-L-alanine *N*-methylamide (6).** Compound **5** (20 mg, 0.028 mmol, 1 equiv.) was dissolved in dry CH_2Cl_2 (1.5 mL) under nitrogen. TFA (212 µL, 0.280 mmol, 100 equiv.) was added and the reaction mixture was stirred for 2 h (TLC, 9 : 1 chloroform–methanol). The solvent was co-evaporated with toluene and with CH_2Cl_2. The crude product was dissolved in dry DMF (500 µL) together with PyBROP (29 mg, 0.062 mmol, 2.2 equiv.) and methylamine hydrochloride (8 mg, 0.120 mmol, 4.2 equiv.) under nitrogen at 0 °C. DIPEA (75 µL, 0.427 mmol, 3.5 equiv.) was added and the reaction mixture was stirred at 0 °C for 1 h and then at room temperature for 2 h (TLC, 9 : 1 chloroform–methanol). The solvent was evaporated and the crude was purified by automated chromatography on a reversed-phase C-18 column (CH_3CN–H_2O 5 to 30% tr = 5 min) to afford 10 mg of **6** (54% over two steps). $[\alpha]_D^{20}$ = +69.5 (c = 0.8, MeOH). ^{1}H NMR (400 MHz, CD_3OD, 25 °C): δ = 5.92 (d, $J_{1,2}$ = 5.5 Hz, 1 H, 1-H), 5.52 (dd, J_{3-2} = 11.1, J_{3-4} = 3.5 Hz, 1 H, 3-H), 5.44 (m, 1 H, 4-H), 5.24 (dd, J_{3-2} = 11.1, J_{2-1} = 5.5 Hz, 1 H, 2-H), 4.70 (dd, $J_{\alpha,\beta}$ = 7.2 Hz, $J_{\alpha,\beta}$ = 5.3 Hz, 1H, α-H-Asn), 4.32–4.27 (m, 1 H, α-H-AlaO*t*Bu), 4.24–4.11 (m, 3 H, α-H-AcAla, 5-H, 6-H), 4.05–3.99 (m, 1 H, 6-H), 2.93 (dd, J_{gem} = 15.5, $J_{\alpha,\beta}$ = 7.2 Hz, 1 H, β-HAsn), 2.73 (dd, J_{gem} = 15.5, $J_{\alpha,\beta}$ = 5.3 Hz, 1 H, β-H-Asn), 2.15 (s, 3 H, CH_3CO), 2.05 (s, 3 H, CH_3CO), 2.02 (s, 6 H, CH_3CO, $NHCH_3$), 1.99 (s, 3 H, CH_3CO), 1.38 (d, J = 7.2 Hz, 3H, CH_3-Ala), 1.36 (d, J = 7.2 Hz, 3H, CH_3-Ala) ppm. ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 175.4–171.6 (CO), 75.6 (C-1), 69.9 (C-2), 69.1 (C-3, C-4), 67.9 (C-5), 62.7 (C-6), 51.5 (α-C-Asn), 51.4 (α-C-Ala), 50.9 (α-C-Ala) 38.1 (β-CH_2-Asn), 26.6 ($CONHCH_3$), 22.5 ($NHCH_3$), 20.7–20.6 (4 × OAc), 17.9–17.6 (CH_3-Ala) ppm. MS (ESI): m/z = 682.3 $[M + Na]^+$.

***N*α-(L-*N*-Acetylalanyl)-*N*γ-(α-D-galactopyranosyl)-L-asparagyl-L-alanine *N*-methylamide (2).** Compound **6** (10 mg, 0.015 mmol, 1 equiv.) was dissolved in dry methanol (250 µL) and a catalytic amount of K_2CO_3 (0.1 equiv., pH 8–9) was added. The reaction mixture was stirred for 1 h (TLC, 6 : 4 chloroform–methanol and 8 : 2 chloroform–methanol). IRA H^+ 120 was added to neutral pH. The mixture was filtered and washed with methanol. The solvent was evaporated and the compound was purified by preparative HPLC (C-18 reverse phase, 100 : 0 to 60 : 40 H_2O–CH_3CN in 14 min; tr = 5 min) to afford **2** (6 mg) in 85% yield. $[\alpha]_D^{20}$ = +46.3 (c = 0.1, MeOH). ^{1}H NMR (400 MHz, CD_3OD, 25 °C): δ = 5.58 (d, $J_{1,2}$ = 5.6 Hz, 1 H, 1-H), 4.67 (t, J = 6.0 Hz, 1 H, α-H-Asn), 4.30–4.21 (m, 2 H,

α-H-Ala), 3.98 (dd, $J_{2,3}$ = 10.2 Hz, $J_{2,1}$ = 5.6 Hz, 1 H, 2-H), 3.87 (d, J = 3.4 Hz, 1 H, 4-H), 3.74 (dd, J_{3-2} = 10.2 Hz, J_{3-4} = 3.4 Hz, 1 H, 3-H), 3.70–3.60 (m, 3 H, 5-H, 6-H), 2.85–2.78 (m, 2 H, β-CH_2-Asn), 2.72 (s, 1 H, $NHCH_3$), 1.98 (s, 3 H, NHAc), 1.36 (d, J = 7.3 Hz, 3H, CH_3Ala), 1.33 (d, J = 7.2 Hz, 3H CH_3Ala) ppm. ^{13}C NMR (100 MHz, $CDCl_3$, 25 °C): δ = 175.5–172.6 (CO), 78.6 (C-1), 73.8 (C-5), 71.4 (C-3), 70.7 (C-4), 68.4 (C-2), 62.8 (C-6), 51.7 (α-C-Asn), 51.3 (α-C-Ala), 50.9 (α-C-Ala), 38.5 (β-CH_2-Asn), 26.6 ($NHCH_3$), 22.5 ($CONHCH_3$), 17.9, 17.6 (CH_3Ala) ppm. MS (ESI): m/z = 492.2 $[M + H]^+$. FT-ICR MS (ESI): calcd for $[C_{19}H_{33}N_5O_{10}Na]^+$ 514.21196; found 514.21128.

NMR spectroscopy

For the conformational analysis of the two α-*N*-linked glycopeptides in solution, the experiments were recorded in H_2O–D_2O 90 : 10 on a Bruker Avance 500 MHz spectrometer at 278 K. 2D-NOESY experiments were carried out with mixing times of 300 and 600 ms. The concentration of glycopeptides for the NMR experiments was set to 10 mM. STD NMR experiments were recorded at 298 K on a Bruker Avance 500 MHz spectrometer. VAA and ECA were dissolved in D_2O buffer (20 mM NaPi, 100 mM NaCl, pH = 7.4) and the final concentration measured by UV spectroscopy. For binding studies, a stock solution of the glycopeptides was prepared, thus the glycopeptides were suspended in a buffer solution to a final concentration of 40 mM. STD experiments were performed for a molar ratio of 100/1 (glycopeptides **1** or **2**–protein). The final concentration of the protein in the NMR tube was 40 μM. A series of Gaussian-shaped pulses of 49 ms each were applied, separated by 1 ms delay, with a total saturation time for the protein envelope of 2 s and a maximum B_1 field strength of 50 Hz. An off-resonance frequency of δ = 100 ppm (where no proteins signals are present) and on-resonance frequency of δ = 7.2 ppm and −1 ppm (protein aromatic signals region) were employed. No significant differences on the epitope mapping were observed between the two on-resonance frequencies. A total number of 1024 scans were acquired and the spectra were multiplied by an exponential line broadening function of 1 Hz prior to Fourier transformation. All experiments were recorded with a 15 ms spin lock pulse, which minimizes the protein background resonances. trNOESY experiments (mixing time 200 ms) were recorded on a Bruker Avance 600 MHz spectrometer equipped with a triple channel cryoprobe at 278 K. These experiments were accomplished for a molar ratio of 7 : 1 (glycopeptides **1** or **2**–VAA) using a 60 μM of lectin as final concentration in the NMR tube. For this particular experiment, VAA and ECA were dissolved in H_2O buffer solution (20 mM NaPi, 100 mM NaCl, pH = 7.4). No purging spin-lock period to remove the NMR signals of the macromolecule background was employed. To properly analyse the sign change of the NOE cross-peaks, a NOESY spectrum of each glycopeptide at 600 MHz was also recorded (see S14–S15 in ESI‡).

Conformation analysis and dynamic simulations

MC/EM calculations[15] were performed using MacroModel 9.5.[17] The AMBER* force field with the Senderowitz–Still parameters[30] has been used. Water solvation was simulated using GB/SA continuum solvent model.[16] Extended non-bonded cut off distances (a van der Waals cut off of 8.0 Å and an electrostatic cut-off of 20.0 Å) were used. The MC/EM procedure was carried out applying 6000 and 10 000 steps for glycopeptides 1 and 2, respectively. Backbone (Φ–Ψ) and side chain (χ^1–χ^2) dihedral angles were all varied during the simulation, along with the pseudo-Φ anomeric torsion. Only conformers among 5.00 kcal mol^{-1} from the global minimum were analyzed and clustered based on their backbone and side chain conformation. The MC/SD dynamic simulations[19] were run using the AMBER* all-atom force field and van der Waals and electrostatic cut-offs of 25 Å, together with a hydrogen bond cut-off of 15 Å. The same degrees of freedom of the MC/EM searches were used. All simulations were performed at 300 K, with a dynamic time-step of 1.5 fs and a frictional coefficient of 0.1 ps^{-1}. Runs of 5 ns for **1** and 10 ns for **2** were performed, starting from conformations selected from the MC/EM outputs. The acceptance ratios for glycopeptides **1** and **2** were 4.5 and 4.0, respectively. MD simulations in explicit water with periodic boundary condition were performed using AMBER 9[20] with ff99SB[31] and glycam04[32] force fields. TIP3P water model and a truncated octahedron box with 12.0 Å buffer were used. An integration step of 1 fs and a cutoff of 10.0 Å were applied. Glycopeptides **1** and **2** were simulated at 300 K (using the Langevin thermostat) and 1 atm for 20 ns and 50 ns, respectively, after an equilibration time of 450 ps.

Docking calculations

The 3D geometries of each glycopeptide–lectin complex were deduced by using AutoDock 4.2.[25] The 3D coordinates of glycopeptides (**1** and **2**) were those obtained after the NMR-based analysis of the compound in the free state. Starting from the available coordinates for VAA and ECA in the Protein Data Bank (pdb codes 1PUM and 1GZC, respectively), molecular models of both proteins were constructed, using the protein preparation wizard of Macromodel, as integrated in the Schrödinger package. The pdb structures for both proteins were modified by removing all the ions and waters of crystallization, with exception of W1120 and W1239, in the case of VAA, and W2262, in the case of ECA. Hydrogens were then added and the structures were minimized with the OPLS-2005 force field,[33] using the convergence method Truncated Newton Conjugate Gradients, in the presence of the water molecules. The existing ligands in the binding site were fixed to their crystallographic positions. Finally, the natural ligand was removed from the binding pocket and the structure of the proteins saved as a new pdb file.

Different torsional degrees of freedom for the glycopeptides **1** and **2** were considered as a flexible ligand to facilitate the docking process. The lectins were always considered as rigid receptors. However, once the torsion angles around the peptide chain (for compound **2**) were optimized, only eight active torsions were considered. Autodock default charges were used for the glycopeptides and the protein receptors. In brief, grid maps (with a grid spacing of 0.375 Å) were constructed using 54 × 54 × 54 points for the box dimensions. A total of 100 Lamarckian genetic algorithm runs were performed using 2 × 10^7

evaluations. The analysis of the different binding modes was performing after clustering the results using a rmsd of 2.0 Å. Finally, the Autodock-based geometries that satisfactorily correlated with the NMR experimental data were further minimized using Macromodel and analysed.

Acknowledgements

Filipa Marcelo thanks FCT-Portugal for a post-doctoral research grant (SFRH/BPD/65462/2009). Financial support also came from Universita' degli Studi di Milano (PUR2009) and Ministero dell'Istruzione Universita' e Ricerca (PRIN 2008, Prot. 2008J4YNJY). The group at Madrid is grateful to MICINN for funding through grant CTQ2009-08536. The groups of AB and JJB are part of the CM1102 COST Action; the groups of JJB and HJG are part of the EC GlycoHIT project (contract no. 260600).

Notes and references

1 (*a*) D. P. Gamblin, E. M. Scanlan and B. G. Davis, *Chem. Rev.*, 2009, **109**, 131; (*b*) D. Solís, A. Romero, M. Menéndez and J. Jiménez-Barbero, in *The Sugar Code. Fundamentals of Glycosciences*, ed. H.-J. Gabius, Wiley-VCH, Weinheim, 2009, pp. 233–245; (*c*) H.-J. Gabius, S. André, J. Jiménez-Barbero, A. Romero and D. Solís, *Trends Biochem. Sci.*, 2011, **36**, 298.

2 (*a*) C. M. Taylor, *Tetrahedron*, 1998, **38**, 11317; (*b*) O. Seitz, *ChemBioChem*, 2000, **1**, 214; (*c*) J. van Ameijde, H. B. Albada and R. M. J. Liskamp, *J. Chem. Soc., Perkin Trans. 1*, 2002, 1042 and references therein (*d*) S. André, C. E. J. Maljaars, K. M. Halkes, H.-J. Gabius and J. P. Kamerling, *Bioorg. Med. Chem. Lett.*, 2007, **17**, 793–798.

3 (*a*) T. Lakshmanan, D. Sriram, K. Priya and D. Loganathan, *Biochem. Biophys. Res. Commun.*, 2003, **312**, 405; (*b*) S. R. Hanson, E. K. Culyba, T.-L. Hsu, C.-H. Wong, J. W. Kelly and E. T. Powers, *Proc. Natl. Acad. Sci. U. S. A.*, 2009, **106**, 3131; (*c*) I. B. H. Wilson, K. Paschinger and D. Rendič, in *The Sugar Code. Fundamentals of Glycosciences*, ed. H.-J. Gabius, Wiley-VCH, Weinheim, 2009, pp. 136–154; (*d*) C. Zuber and J. Roth, in *The Sugar Code. Fundamentals of Glycosciences*, ed. H.-J. Gabius, Wiley-VCH, Weinheim, 2009, pp. 87–110; (*e*) A. Larkin and B. Imperiali, *Biochemistry*, 2011, **50**, 4411.

4 R. G. Spiro, *Glycobiology*, 2002, **12**, 43R–56R.

5 S. Shibata, T. Takeda and Y. Natori, *J. Biol. Chem.*, 1988, **263**, 12843.

6 C. J. Bosques, S. M. Tschampel, R. J. Woods and B. Imperiali, *J. Am. Chem. Soc.*, 2004, **26**, 8421.

7 C. Colombo and A. Bernardi, *Eur. J. Org. Chem.*, 2011, 3911–3919.

8 (*a*) H. Kogelberg, D. Solís and J. Jiménez-Barbero, *Curr. Opin. Struct. Biol.*, 2003, **13**, 646; (*b*) V. Roldós, F. J. Cañada and J. Jiménez-Barbero, *ChemBioChem*, 2011, **12**, 990.

9 (*a*) H.-J. Gabius, H. Walzel, S. S. Joshi, J. Kruip, S. Kojima, V. Gerke, H. Kratzin and H.-J. Gabius, *Anticancer Res.*, 1992, **12**, 669; (*b*) H.-J. Gabius, F. Darro, M. Remmelink, S. André, J. Kopitz, A. Danguy, S. Gabius, I. Salmon and R. Kiss, *Cancer Invest.*, 2001, **19**, 114; (*c*) M. Jiménez, J. L. Sáiz, S. André, H.-J. Gabius and D. Solís, *Glycobiology*, 2005, **15**, 1386; (*d*) L. He, S. André, V. M. Garamus, H.-C. Siebert, C. Chi, B. Niemeyer and H.-J. Gabius, *Glycoconjugate J.*, 2009, **26**, 111.

10 (*a*) B. Shaanan, H. Lis and N. Sharon, *Science*, 1991, **254**, 862; (*b*) H.-C. Siebert, R. Adar, R. Arango, M. Burchert, H. Kaltner, G. Kayser, E. Tajkhorshid, C.-W. von der Lieth, R. Kaptein, N. Sharon, J. F. G. Vliegenthart and H.-J. Gabius, *Eur. J. Biochem.*, 1997, **249**, 27; (*c*) S. Elgavish and B. Shaanan, *J. Mol. Biol.*, 1998, **277**, 917; (*d*) K. Turton, R. Natesh, N. Thiyagaraja, J. A. Chaddock and K. R. Acharya, *Glycobiology*, 2004, **14**, 923.

11 M. Jiménez, S. André, C. Barrilari, A. Romero, D. Rognan and H.-J. Gabius, *FEBS Lett.*, 2008, **582**, 2309.

12 (*a*) M. Mayer and B. Meyer, *Angew. Chem., Int. Ed.*, 1999, **38**, 1784; (*b*) M. Mayer and B. Meyer, *J. Am. Chem. Soc.*, 2001, **123**, 6108; (*c*) B. Meyer and T. Peters, *Angew. Chem., Int. Ed.*, 2003, **42**, 864.

13 (*a*) F. Ni, *Prog. Nucl. Magn. Reson. Spectrosc.*, 1994, **26**, 517; (*b*) *NMR Spectroscopy of Glycoconjugates*, ed. J. Jiménez-Barbero and T. Peters, Wiley-VCH, Weinheim, 2002.

14 (*a*) J. Jiménez-Barbero, E. Dragoni, C. Venturi, F. Nannucci, A. Ardá, M. Fontanella, S. André, F. J. Cañada, H.-J. Gabius and C. Nativi, *Chem.–Eur. J.*, 2009, **15**, 10423; (*b*) M. Kolympadi, M. Fontanella, C. Venturi, S. André, H.-J. Gabius, J. Jiménez-Barbero and P. Vogel, *Chem.–Eur. J.*, 2009, **15**, 2861; (*c*) J. P. Ribeiro, S. André, F. J. Cañada, H.-J. Gabius, A. P. Butera, R. J. Alves and J. Jiménez-Barbero, *ChemMedChem*, 2010, **5**, 415; (*d*) F. Corzana, J. H. Busto, F. Marcelo, M. García de Luis, J. L. Asensio, S. Martín-Santamaría, J. Jiménez-Barbero, A. Avenoza and J. M. Peregrina, *Chem.–Eur. J.*, 2011, **17**, 3105.

15 G. Chang, W. C. Guida and W. C. Still, *J. Am. Chem. Soc.*, 1989, **111**, 4379.

16 W. C. Still, A. Tempzyk, R. Hawley and T. Hendrickson, *J. Am. Chem. Soc.*, 1990, **112**, 6127.

17 MacroModel, version 9.5, Schrödinger, LLC, New York, NY, 2007; F. Mohamadi, N. G. J. Richard, W. C. Guida, R. Liskamp, M. Lipton, F. Caufield, G. Chang, T. Hendrickson and W. C. Still, *J. Comput. Chem.*, 1990, **11**, 440–467.

18 S. Gnanakaran and A. E. Garcia, *J. Phys. Chem. B*, 2003, **107**, 12555.

19 F. Guarnieri and W. C. Still, *J. Comput. Chem.*, 1994, **15**, 1302.

20 D. A. Case, T. A. Darden, T. E. Cheatham III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, K. M. Merz, D. A. Pearlman, M. Crowley, R. C. Walker, W. Zhang, B. Wang, S. Hayik, A. Roitberg, G. Seabra, K. F. Wong, F. Paesani, X. Wu, S. Brozell, V. Tsui, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, P. Beroza, D. H. Mathews, C. Schafmeister, W. S. Ross and P. A. Kollman, *AMBER 9*, University of California, San Francisco, 2006.

21 F. Eker, K. Griebenow, X. Cao, L. A. Nafie and R. Schweitzer-Stenner, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 10054.

22 A. Motta, M. Reches, L. Pappalardo, G. Andreotti and E. Gazit, *Biochemistry*, 2005, **44**, 14170.

23 R. L. Dunbrack and F. E. Cohen, *Protein Sci.*, 1997, **7**, 1661.

24 M. W. van der Kamp, R. D. Schaeffer, A. L. Jonsson, A. D. Scouras, A. M. Simms, R. D. Toofanny, N. C. Benson, P. C. Anderson, E. D. Merkley, S. Rysavy, D. Bromley, D. A. C. Beck and V. Daggett, *Structure*, 2010, **18**, 423.

25 G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell and A. J. Olson, *J. Comput. Chem.*, 2009, **16**, 2785.

26 (*a*) M. Jiménez, S. André, H.-C. Siebert, H.-J. Gabius and D. Solís, *Glycobiology*, 2006, **16**, 926; (*b*) S. Martín-Santamaría, S. André, E. Buzamet, R. Caraballo, G. Fernández-Cureses, M. Morando, J. P. Ribeiro, K. Ramírez-Gualito, B. de Pascual-Teresa, F. J. Cañada, M. Menéndez, O. Ramström, J. Jiménez-Barbero, D. Solís and H.-J. Gabius, *Org. Biomol. Chem.*, 2011, **9**, 5445.

27 (*a*) G. Terraneo, D. Potenza, A. Canales, J. Jiménez-Barbero, K. K. Baldridge and A. Bernardi, *J. Am. Chem. Soc.*, 2007, **129**, 2890–2900; (*b*) S. Vandenbussche, F.-A. Díaz, W. Pan, S. P. Vincent, G. Cuevas, F. J. Cañada, J. Jiménez-Barbero and K. Bartik, *Chem.–Eur. J.*, 2008, **14**, 7570–7578.

28 (*a*) S. André, D. Giguère, T. K. Dam, C. F. Brewer, H.-J. Gabius and R. Roy, *New J. Chem.*, 2010, **34**, 2229; (*b*) D. Giguère, S. André, M.-A. Bonin, M.-A. Bellefleur, A. Provencal, P. Cloutier, B. Pucci, R. Roy and H.-J. Gabius, *Bioorg. Med. Chem.*, 2011, **19**, 3280–3287.

29 (*a*) R. Schwartz-Albiez, in *The Sugar Code. Fundamentals of Glycosciences*, ed. H.-J. Gabius, Wiley-VCH, Weinheim, pp. 447–467; (*b*) H. Kaltner and H.-J. Gabius, *Histol. Histopathol.*, 2012, **27**, 397–416.

30 H. Senderowitz and W. C. Still, *J. Org. Chem.*, 1997, **62**, 1427.

31 T. E. Cheatham III, P. Cieplak and P. A. Kollman, *J. Biomol. Struct. Dyn.*, 1999, **16**, 845.

32 R. J. Woods, R. A. Dwek, C. J. Edge and B. Fraser-Reid, *J. Phys. Chem.*, 1995, **99**, 3832.

33 W. L. Jorgensen, D. S. Maxwell and J. Tirado-Rives, *J. Am. Chem. Soc.*, 1996, **118**, 11225.