Exquisite Binding Mechanism of Human Macrophage Galactose-Type Lectin: Insights from the Plasticity of Its Carbohydrate Recognition Domain

Ana Diniz

Helena Coelho

Jorge S. Dias]

Abstract: Human macrophage galactose-type lectin (MGL), expressed on macrophages and dendritic cells (DCs), modulates distinct immune cell responses by recognizing GalNAc-containing structures present on pathogens, self-glycoproteins and tumor cells. Herein, we used NMR spectroscopy and molecular dynamics (MD) simulations to investigate the structural preferences of MGL against different GalNAc-containing structures derived from blood group A antigen, Forssman antigen and the GM2 glycolipid. NMR analyses of the MGL carbohydrate recognition domain (MGL-CRD, C181-H316) in the absence and presence of α-methyl GalNAc, a simple monosaccharide, shows that the MGL-CRD is highly dynamic and its structure is strongly altered upon ligand binding. This plasticity of the MGL-CRD structure explains the ability of MGL to accommodate different GalNAc-containing molecules. However, key differences are observed in the recognition process reliant whether the GalNAc is part of the blood group A antigen, the Forssman antigen or GM2-derived structures. These results are in accordance with molecular dynamics simulations that suggest the existence of a distinct MGL binding mechanism depending on the context of GalNAc moiety presentation. These results afford new perspectives for the rational design of GalNAc modifications able to fine tune MGL immune responses in distinct biological contexts, especially in malignancy.

Introduction

The human macrophage galactose-type lectin (MGL) is an immunological C-type lectin expressed on the surface of macrophages and dendritic cells (DCs) that recognizes glycans with exposed *N*-acetylgalactosamine (GalNAc) residues in a Ca^{2+}-dependent manner.[1,2] MGL is a type II transmembrane protein that contains a cytoplasmic domain, a transmembrane domain, a neck region and a carbohydrate recognition domain (CRD). [3] Physical and chemical methods have demonstrated that the extracellular domain of MGL, constituted by the neck region and the CRD domain, is a trimeric structure, mainly stabilized by the helical coiled-coil neck region and where the CRDs act as independent domains.[4] The MGL-CRD is composed by the characteristic Gln-Pro-Asp (QPD) motif which is responsible for the selective recognition of GalNAc and galactose (Gal) residues at the CRD.[5] Indeed, the availability and orientation of the equatorial/axial 3-OH and 4-OH groups of Gal/GalNAc residues are crucial for Ca^{2+} coordination at the CRD binding site.[5] Nevertheless, additional contacts established by the 2-acetamido group (NHAc) render the MGL binding preference for GalNAc unit over Gal.[6] The molecular recognition of more complex GalNAc-containing structures by MGL has been extensively investigated by glycan microarrays, surface plasmon resonance and cellular binding assays showing high binding specificity to *O*-glycan Tn-structures (αGalNAc-Ser/Thr); Helminth glycans (fucosylated and non-fucosylated LacdiNAc structures); glycosphingolipid structures (GM2 and GD2), Forssman antigen and blood group A antigen.[1,7] In this context, by sensing Tn-structures present on self-glycoproteins (CD43/CD45) and on tumor-associated MUC1 glycoprotein, MGL induces immunosuppressive responses in malignancy, characterized by increased IL-10, production and induction of regulatory T cells and the effector T cell apoptosis.[8–12] Furthermore, the MGL ligands GM2/GD2 and Forssman antigen are known tumor-associated antigens present in several tumors[1,7,13] which may correlate with the role of MGL in malignancy. Additionally, in a rat model, the MGL binding to the blood group A antigen correlates with increased apoptosis resistance and the ability of cancer cells to escape to immune surveillance, which may also be correlated with the

---

[a] A. Diniz, Dr. H. Coelho, Dr. J. S. Dias, Prof. Dr. E. J. Cabrita, Dr. F. Marcelo
UCIBIO, REQUIMTE, Departamento de Química,
Faculdade de Ciências e Tecnologia,
Universidade NOVA de Lisboa,
2829-516 Caparica, (Portugal)
E-mail: filipa.marcelo@fct.unl.pt
[b] Dr. H. Coelho, Prof. Dr. J. Jiménez-Barbero
CIC bioGUNE,
Bizkaia Technology Park, Building 801A,
48170 Derio (Spain)
[c] Dr. H. Coelho, Prof. Dr. J. Jiménez-Barbero
Departament of Organic Chemistry II,
Faculty of Science & Technology,
University of the Basque Country,
Leioa, 48940 Bizkaia, Spain
[d] Dr. S. J. van Vliet
Amsterdam UMC, Vrije Universiteit Amsterdam,
Department of Molecular Cell Biology and Immunology, Cancer Center Amsterdam,
Amsterdam Infection and Immunity Institute,
Amsterdam, the Netherlands
[e] Prof. Dr. F. Corzana
Departamento de Química, Universidad de La Rioja,
Centro de Investigación en Síntesis Química,
E-26006 Logroño, Spain
[f] Prof. Dr. J. Jiménez-Barbero
Ikerbasque, Basque Foundation for Science,
Maria Diaz de Haro 13,
48009 Bilbao, Spain

[+] These authors contributed equally to this work.

Supporting information for this article is given via a link at the end of the document.

immunosuppressive properties of MGL.[14] Therefore, targeting of MGL could provide a novel therapeutic approach to alleviate MGL-mediated immune suppression and to elicit anti-tumor immunity. To efficiently target MGL it is highly important to explore and characterize the molecular determinants of MGL recognition. In this context, we and others have earlier reported, from a ligand perspective, the structural elements of the molecular recognition process of Gal/GalNAc and Tn/sTn glycopeptides by MGL using saturation transfer-difference NMR (STD-NMR) in tandem with molecular modelling protocols.[6,15] However, the lack of human MGL-CRD X-ray crystal structure has precluded a structure-guided design of molecules targeting MGL. Here, by NMR spectroscopy in tandem with molecular dynamics (MD) simulations, we have unveiled that the exquisite structural preferences of MGL critically depend on the context of the GalNAc ligand. We have achieved the ^{1}H,^{15}N backbone resonance assignments of the MGL-CRD in the free and GalNAc complexed states and analyzed its interactions with distinct biologically relevant GalNAc-containing moieties (Figure 1), including the blood group A antigen (**1**), Forssman antigen (**2**) and GM2/asialo GM2 glycolipid structures (**3**)/(**4**) using ^{1}H,^{15}N HSQC titrations and STD-NMR studies. Analysis of the data has permitted us to define the exquisite differential ligand recognition patterns employed by the MGL-CRD. The results reported herein represent a significant improvement in the understanding of factors that determine the MGL fine specificity and the mode of action of MGL towards distinct GalNAc-containing ligands.

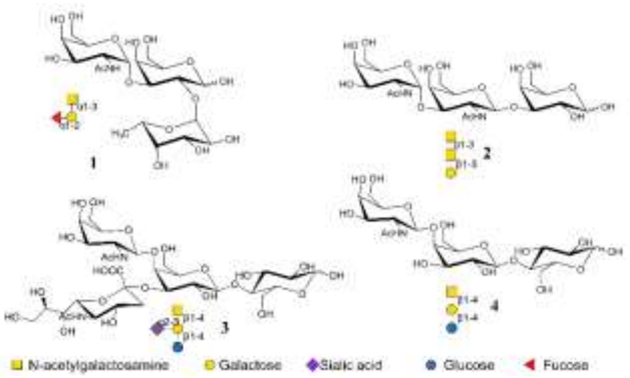


Figure 1. Structure of the GalNAc-containing fragments derived from blood group A antigen (**1**), Forssman antigen (**2**), GM2 (**3**) and asialo GM2 (**4**).

Results and Discussion

MGL-CRD apo vs MGL-CRD/α-methyl GalNAc complex

NMR methods [16] were employed to study the structure and GalNAc-binding mechanism of the human MGL-CRD in solution. The recombinant MGL-CRD fragment (137 residues, C181-H316, Figure S1) was prepared with uniform ^{13}C,^{15}N-labelling in presence of 20 mM of Ca^{2+} (see experimental section). Purification of MGL-CRD using a lactose-agarose column allowed the discrimination of the functional folded fraction of MGL-CRD and the non-functional and incorrectly folded protein. The ^{1}H,^{15}N HSQC spectra of the purified ^{15}N-MGL-CRD reveals a well folded protein (Figure 2A – black spectrum). Additionally, ^{1}H,^{15}N HSQC spectra of ^{15}N-MGL-CRD at fixed concentration (200µM) were acquired with increasing amounts of α-methyl GalNAc (Figure S2). A gradual disappearance of two representative cross-peaks of MGL-CRD (G279 and Q297) corresponding to the unbound state and the concomitant appearance of new cross-peaks corresponding to residues of MGL-CRD bound to α-methyl GalNAc can be observed in Figure S2-B/C. The relative intensities of the two sets of ^{1}H,^{15}N HSQC cross peaks reflect the population of the free and bound lectin states along the titration (Figure S2-B/C). The slow exchange rate observed during the ^{1}H,^{15}N HSQC titrations of MGL-CRD with α-methyl GalNAc is in accordance with a low dissociation rate (K_{off}), between the free and bound forms, and compatible with the dissociation constant (K_D) value of ca. 23 µM, previously determined for the MGL-FL/GalNAc complex, by STD-NMR competition binding experiments.[6,15]

Resonance assignments of the N, C, CA, CB and H nuclei for both MGL-CRD in apo and GalNAc-bound state (MGL-CRD/ α-methyl GalNAc 1:10 molar ratio) using standard triple resonance approach was carried out (experimental section, Figure S3-S4 and table S1 in supporting information). In the presence of α-methyl GalNAc, a total of 123 out of 130 amide resonances were unambiguously assigned (MGL-CRD construct has 137 amino acid residues including the initial methionine and 6 proline residues, Figure S1). In the α-methyl GalNAc-bound state 11 amide resonances were assigned that were totally absent in the apo state form of MGL-CRD. These resonances are annotated in the ^{1}H,^{15}N-HSQC with a * (Figure 2A – red spectrum) and correspond to the following amino acids residues: D270, Q272, H274, G275, L276, D288, N292, D293, D294, V295, R298. Chemical shift perturbation (CSP) was also analyzed taking in to account the variation of chemical shift in the ^{1}H and ^{15}N dimensions from the MGL-CRD apo state and MGL-CRD bound state in presence of 10 equivalents of α-methyl GalNAc (see supporting information for details).[17] The histogram of the associated combined chemical shift perturbation (Δδcomb) are shown in Figure 2B. Two cut-off lines at 0.15 and 0.08 ppm were considered to discriminate the amide resonances that strongly and moderately shift in presence of α-methyl GalNAc. In particular, the amide cross peaks of Y236, G279, H286 residues, as well as, the indole NH of W271, experienced high values of Δδcomb upon addition of α-methyl GalNAc (Δδcomb > 0.15 ppm). Additionally, the amide resonances of A235, K264, Q267, D269, W271, G277, G278, E280, C296, Q297 residues are moderately perturbed in presence of α-methyl GalNAc (0.15 ppm > Δδcomb > 0.08 ppm). All the amino acids that experience CSP upon α-methyl GalNAc addition are annotated in the ^{1}H,^{15}N-HSQC without * (Figure 2A – red spectrum).

Additionally, the variations of the ^{1}H,^{15}N-HSQC cross peak intensities, in absence and presence of α-methyl GalNAc, were also scrutinized (Figure 2C). Interestingly, several resonances showed a significant increase in their intensities in the presence of α-methyl GalNAc. As expected, most of them correspond to those amino acids whose resonances become observable only upon addition of α-methyl GalNAc (residues labelled in Figure 2C). However, several other residues also showed an increase in the intensity of their cross-peak resonances. These results clear highlight that α-methyl GalNAc binding induces strong effects on the MGL-CRD dynamics. In addition, this alteration is general and extended to regions located far away from the QPD-motif (gray bar displayed in Figure 2B and 2C).

A 3D homology model of MGL-CRD, generated from the X-Ray structure of asialoglycoprotein (ASGP, pdb 1DV8)[18], which shares 73% identity with MGL, has been previously built in apo and α-methyl GalNAc-bound state.[6] The secondary structural elements considering this homology model are shown in Figure 2 on top of the histograms. With the backbone assignment in hand the secondary structural elements of the MGL bound to α-methyl GalNAc were determined through the analysis of the chemical shift index (CSI) using ^{13}Cα and ^{13}Cβ resonances (see supporting for details,[19]) (Figure S5) and further compared with the secondary structure predicted by the 3D model.

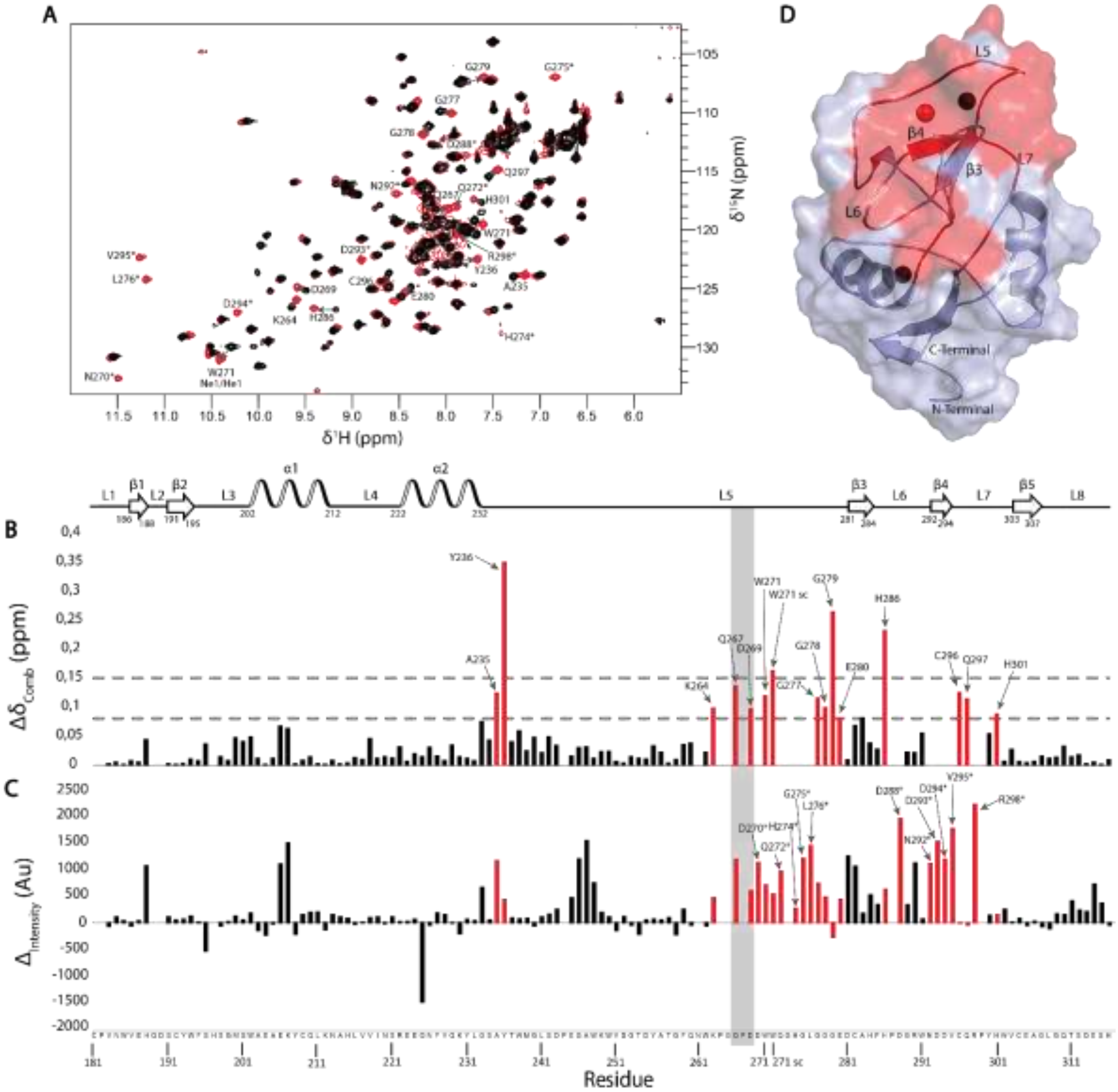


Figure 2. NMR analysis of MGL-CRD in apo and α-methyl GalNAc-bound states. **A.** Overlay of the ^{1}H,^{15}N HSQC NMR spectra of MGL-CRD in the apo (black) and α-methyl GalNAc bound states (red). The residues labelled with * correspond to those residues that appear and without * those that strongly and moderately shift upon addition of GalNAc. **B.** Histogram of the ΔδComb for each amino acid of MGL-CRD/ α-methyl GalNAc complex in presence of 10 equivalents of GalNAc. Two cut-off lines at 0.08 ppm and 0.15 ppm, displayed as gray dash line, are considered to discriminate the residues that strongly and moderately shift in presence of α-methyl GalNAc. **C.** Histogram representing the difference in intensity (ΔIntensity) of MGL-CRD residues in presence and absence of α-methyl GalNAc. For the residues which resonances only become detectable in presence of α-methyl GalNAc an intensity of zero was considered for MGL-CRD in the apo state On top of the histograms it is displayed the secondary structural elements from the 3D MGL-CRD homology model.[6] α-Helices and β-sheets are represented by waves and arrows, respectively. The red bars in histograms indicate the amino acids affected in the presence of α-methyl GalNAc and the gray box highlights the QPD-motif. **D.** Structural mapping of the residues perturbed upon α-methyl GalNAc binding highlighted in 3D MGL-CRD homology model. The three Ca^{2+} present on the MGL-CRD structure are shown as spheres and the Ca^{2+} ion coordinated to the QPD-motif is shown in red. The amino acids affected in the presence of α-methyl GalNAc (residues that appear and strongly/moderately shift) are displayed in red color in the model.

The CSI-prediction of MGL-CRD bound to α-methyl GalNAc is, in general, in good agreement with the 3D homology model, concerning to the location of the α-helical and β-sheets secondary structural elements, with exception of the region from A214 to V217 and from T237 to M239 to (Figure S5). In addition, the lack of prediction to long stretches of the sequence is consistent with a highly flexible structure for the MGL-CRD domain with the presence of several and long loop regions (especially between the residues 233 and 281 of loop L5, Figure S5). Therefore, we used the 3D homology model to map the amino acid residues whose amide resonances appeared or shift (strongly and moderately) upon α-methyl GalNAc addition (Figure 2D). Chemical shift mapping shows that many amino acid signals are perturbed indicating that the MGL-CRD structure is strongly altered upon GalNAc binding. Both the amino acids Q267 and D269 (cross-peaks moderately perturbed) from the crucial QPD-motif are affected in presence of α-methyl GalNAc. These two residues are involved in the complexation of Ca^{2+} (Figure 2D – Ca^{2+} ion shown in red) that, in turn, binds to the OH-3 and OH-4 of GalNAc in MGL-CRD binding site. Nevertheless, several other signals are additionally perturbed in presence of GalNAc, with several of them located far away from the QPD-motif (Figure 2D). Most of the amino acid residues affected in the presence of GalNAc belong to the loop L5 (strongly/moderately shift: A235, Y236, K264, Q267, D269, W271, G277, G278, G279, E280 and appeared: D270, Q272, H274, G275, L276). The perturbation of these 15 amino acids (Figure 2B and 2C) pinpoints that the loop L5 is highly dynamic and flexible, adopting several conformations in the MGL-CRD apo-state. Upon GalNAc complexation, the dynamics of this loop is strongly altered, and many amide resonances that are totally absent in the MGL-CRD of the apo form are now detected in the $^{1}H,^{15}N$ HSQC. Furthermore, the amide resonances of the amino acids N292, D293 and D294, located at the β4 region, also appear in presence of GalNAc (Figure 2C). Additional perturbations in a few amino acids at the loops L6 and L7 are also observed (appeared: D288, V295 and R298; shift: H286, C296 and Q297). All these facts together unequivocally point out that GalNAc binding induces strong conformational and dynamics changes on the MGL-CRD structure.

A dual binding mode has previously been proposed for α-methyl GalNAc based on the concerted application of saturation-transfer difference NMR (STD-NMR) experiments, MD and CORCEMA-ST analysis.[6] The coexistence of at least two binding modes for the GalNAc moiety was proposed since the STD experimental intensities could not be explained using a single binding mode for α-methyl GalNAc to MGL. In both proposed binding modes, the GalNAc moiety establishes key interactions (H-bonds and CH/π) with several amino acids at the protein surface. MD analysis with 20 ns of simulation time revealed H-bonds between GalNAc and K264, Q267, D269, E280, N292 and D294 of the protein, as well as key CH/π contacts[20] through the aromatic amino acids Y236 and W271.[6] Indeed, all the amide resonances of these amino acids in the $^{1}H,^{15}N$-HSQC are perturbed in more or less extension (appeared or shift) when passing from the free to the α-methyl GalNAc bound state. Indeed, the now observed chemical shift perturbations (CSP) are not sufficient to ascertain the existence of single of dual binding mode in solution for α-methyl GalNAc.

Remarkably, the high flexibility of MGL-CRD deduced by the analysis of $^{1}H,^{15}N$ HSQC in the absence and presence of α-methyl GalNAc could explain the ability of MGL to recognize a large repertoire of GalNAc-containing ligands. In this context, we decided to explore the fine structural features of MGL-CRD binding towards GalNAc-containing ligands, **1-4** (Figure 1).

MGL-CRD complexed with GalNAc-containing ligands 1-4

Different GalNAc-containing fragments derived from the blood group A antigen (**1**), the Forssman antigen (**2**), the GM2 (**3**) and asialo GM2 (**4**) were selected based on their differences in GalNAc moiety presentation (Figure 1). Specifically, the GalNAc moiety of blood group A antigen (**1**) and Forssman antigen (**2**) are α-substituted by Gal or GalNAc units, respectively. In contrast, the GM2 (**3**) and asialoGM2 (**4**) are substituted by a β-Gal moiety. Furthermore, (**2**) and (**4**) are linear structures, while (**1**) and (**3**) show a branched fucose and a branched sialic acid unit, respectively. From the functional perspective, all the selected GalNAc-containing fragments **1-4** are intrinsically associated with MGL tumor biology. [1,7,13,14]

*Analysis of the $^{1}H,^{15}N$ HSQC NMR spectra of MGL-CRD in presence of **1-4***

MGL-CRD was titrated with increasing concentrations of compounds **1-4** (see experimental section). Figure S6-S9 displays the $^{1}H,^{15}N$ HSQC spectra of MGL-CRD and the corresponding Δδcomb in absence and presence of **1-4** at MGL-CRD/ligand with a 1:10 molar ratio. As described above for GalNAc, ligands **1-4** bind to MGL-CRD in slow exchange regime. The new cross-peaks that appeared in the presence of **1-4** correspond to the amide resonances of those amino acids that also become detectable in presence of α-methyl GalNAc. Furthermore, **1-4** induce large CSP on the key amino acids Y236, G279 and H286 (Δδcomb > 0.15) (Figures S6B-S9B). These results indicate the existence of a common interaction mode around the GalNAc moiety for all ligands.

Nevertheless, to investigate if there is any specific difference in the interaction of **1-4** with MGL, the $^{1}H,^{15}N$-HSQC of MGL-CRD/**1-4** complexes were compared to that obtained for the MGL-CRD/ α-methyl GalNAc, using the same NMR experimental conditions (Figure S10-S13). Figure 3 displays the histogram of Δδcomb based on the CSP in the ^{1}H and ^{15}N dimensions for the MGL-CRD/**1-4** complexes using the $^{1}H,^{15}N$ HSQC of the MGL/ α-methyl GalNAc complex, as a reference, for the calculation. The CSP calculated in this way allows to account for the differences between the bound states of the different complexes. To discriminate the residues that display higher perturbation in presence of **1-4** with respect to the complex MGL-CRD/α-methyl GalNAc only CSPs of $^{1}H,^{15}N$ cross-peaks higher than 0.05 ppm were considered. For comparison, Figure S14 in Supporting information additionally contains the individual CSP analysis of **1-4** with respect to the MGL-CRD at apo state.

Interestingly, no major CSP was detected between the complex of MGL-CRD/**1** with that of MGL-CRD/GalNAc (Figure 3A). In contrast, important changes in the observed CSP were evidenced between the MGL-CRD/ α-methyl GalNAc complex and the MGL-CRD/**2-4** complexes (Figure 3B-D). For MGL-CRD/**2**, subtle differences in CSP were deduced for residues H274, L276, G279 and lateral chain of W271 (0.11 ppm > Δδcomb > 0.05 ppm) in comparison with MGL-CRD/GalNAc (Figure 3B). These residues correspond to those amino acids that only become observable in the presence of α-methyl GalNAc and are located at loop L5 of MGL-CRD model at the right side of the QPD-motif in the amino acid sequence of MGL. As for MGL-CRD/**2**, both MGL-CRD/**3** and MGL-CRD/**4** complexes present a large CSP of H274 with respect to that for MGL-CRD/GalNAc (Δδcomb > 0.11 ppm, Figure 3C-D).

In addition, these last two complexes display a large CSP of K264 and Q267 (Δδcomb > 0.07 ppm, Figure 3C-D), undetectable for the MGL-CRD/**1-2** complexes. K264 is located at loop L5 of MGL-CRD model at the left side of the QPD-motif in the amino acid sequence of MGL (opposite site than H274), while Q267 belongs to the QPD-motif. Interestingly, the careful analysis of Figures 3C-D allowed detecting subtle changes in a few other amino acids for the MGL-CRD/**3** complex. Accordingly, in presence of **3**, a considerable CSP was observed for V295 (Δδcomb = 0.10 ppm) and a more modest variation of CSP was detected for D269, W291 and Q297 (0.06 ppm > Δδcomb > 0.05 ppm). On the other hand, for compound **4**, CSP differences with respect to GalNAc arose for H286 that was not perceived in the presence of **3**.

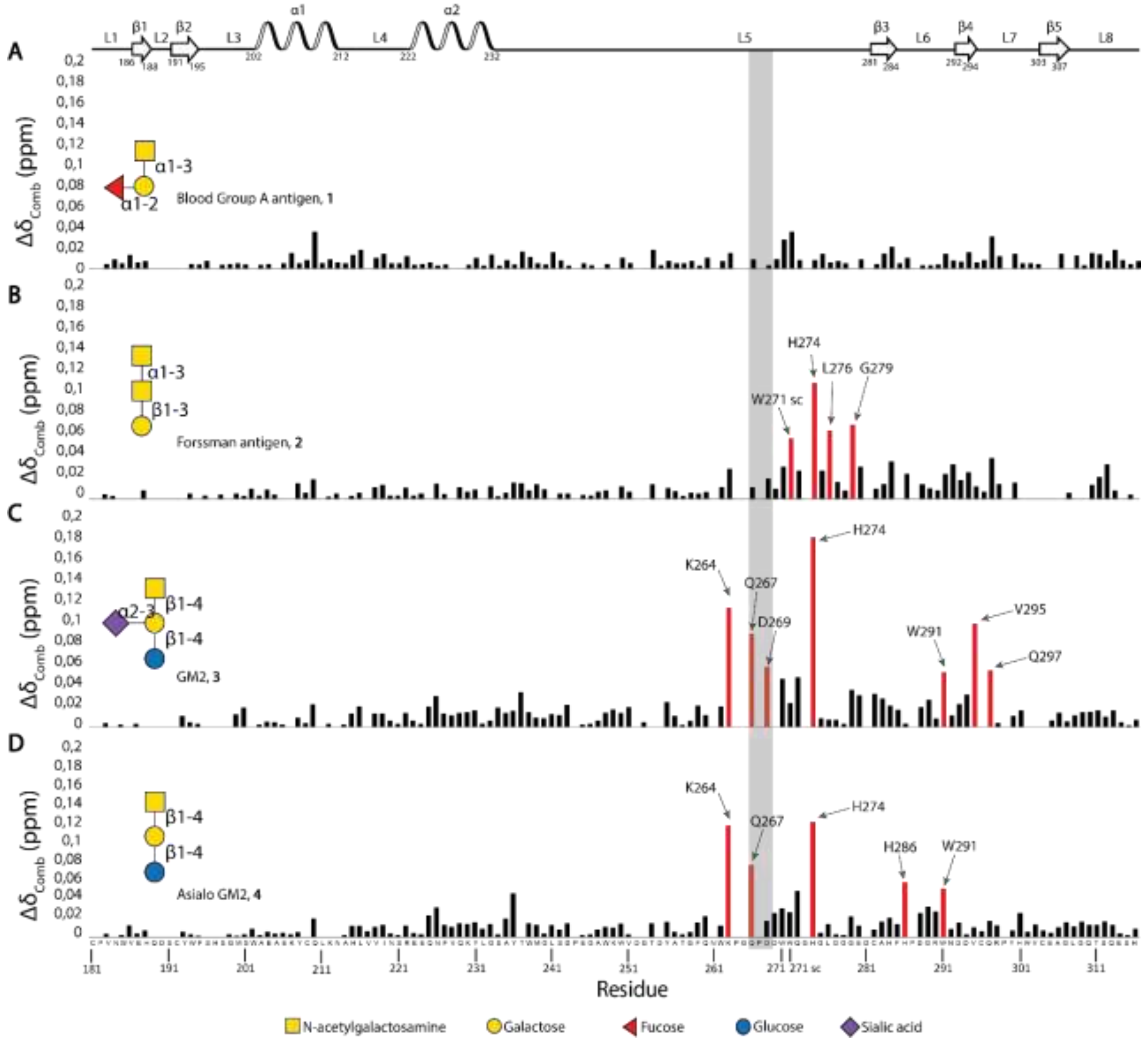


Figure 3. Histogram of the associated ^{1}H,^{15}N combined chemical shift (ΔδComb) of the distinct bound forms using MGL-CRD/ α-methyl GalNAc complex as a reference: **A.** MGL-CRD/**1**. **B.** MGL-CRD/**2**. **C.** MGL-CRD/**3**. **D.** MGL-CRD/**4**. The main amino acids with major differences in CSP (> 0.05 ppm) with respect to the MGL-CRD/ α-methyl GalNAc complex are annotated in the histograms and displayed as red bars. The secondary structural elements from the MGL-CRD derived homology model are shown on the top of the plot. α-Helices and β-sheets are represented by waves and arrows, respectively. The region of the QPD-motif is displayed in a gray box.

Figures 4 displays a closer view of the cross-peaks of K264, H274, H286 and V295 residues for MGL-CRD complexed with GalNAc and overlapped with those for the lectin complexed with **1-4**. The CSP and the pattern observed for K264 (Figure 4A) is totally different for the β- (**3**/**4**) and α-substituted molecules (**1**/**2**). Remarkably, in presence of **3** and **4**, two cross-peaks with similar intensity were assigned to K264, indicating that this amino acid can hold two alternative conformations with the same population in the complexed form. This behavior is not observed in the presence of ligands **1** and **2**. These results strongly suggest that K264 is a key amino acid to engage ligands **3** and **4**. In the case of H274 (Figure 4B), and with exception of **1,** all the ligands induced significant differences in the CSP of H274 relatively to that measured for the MGL-CRD/α-methyl GalNAc complex**.** Furthermore, the extension and the directionality in ^{1}H dimension of Δδcomb is different depending on the presentation of the GalNAc unit, in the Forssman antigen (**2**) or in the GM2-derived structures (**3**, **4**). Particularly, **2** induced a Δδcomb of 0.11 to high field, while **3** and **4** prompted a Δδcomb to low field of 0.18 and 0.12, respectively. These results indicate differences in the chemical environment around H274 depending on the GalNAc moiety presentation. Figures 4C-D illustrate the sensitivity of

H286 and V295 residues to the absence/presence of the sialic acid moiety. Clear differences in the CSP arise when MGL is bound to either **3** or **4** versus GalNAc. Hence, the fine differences in the mechanism of MGL binding against compounds **1-4** can be inferred by chemical shift perturbation analysis.

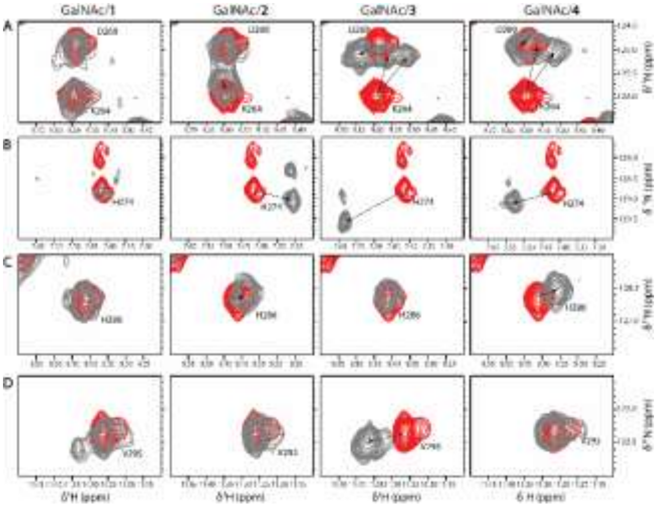


Figure 4. Overlay of regions of the ^{1}H,^{15}N HSQC spectra of MGL-CRD/ α-methyl GalNAc complex (red) and MGL-CRD/ligands **1-4** (gray) showing the cross-peaks of key amino acids that show differential perturbation depending of the ligand context. K264 (**A**); H274 (**B**); H286 (**C**); V295 (**D**).

*STD-NMR derived epitope mapping of ligands **1-4***

The differences in the molecular recognition of MGL towards the GalNAc-containing ligands **1-4** were further investigated by STD-NMR experiments (Figure S15-S18) using the MGL-FL construct (Q61-H316), as previously described.[6] Figure 5 shows the STD-based epitope map derived for the lectin/**1-4** complexes. The STD-based epitope mapping of α-methyl GalNAc, previously reported by us,[6] is also shown in Figure 5 for comparison purposes. As expected, the highest values of STD intensity (> 75%) are presented for most of the protons of GalNAc unit in all the MGL-FL/**1-4** complexes. However, the adjacent sugar units in **2-4**, namely the central unit and the reducing end, display additional STD responses (> 35%). In contrast, for **1**, only the GalNAc unit shows significant STD response (> 35%). This result indicates that only the GalNAc moiety of **1** establishes interactions with MGL, which correlates very well with the fact that no major differences in the CSP of the MGL-CRD amino acids were detected between MGL-CRD/**1** and MGL-CRD/GalNAc complexes (Figure 3A). On the other hand, for **2-4**, other residues beyond the GalNAc moiety establish contacts with the MGL surface. Thus, when the STD-based epitope map of the central and reducing end units of the linear Forssman antigen (**2**, GalNAc α-substituted ligand) was compared to that obtained for asialo GM2 (**4**, GalNAc β-substituted ligand) ligands, it was observed that the asialo GM2 (**4**) shows stronger STD enhancements (> 75%) than Forssman antigen **2** (35-54%). This difference points out that the binding mode of these two ligands to MGL should be somehow different. These results also correlate with the observed CSP differences measured for **2** and **4**, with respect to those for α-methyl GalNAc. Furthermore, differences in the STD-based epitope mapping stand up when the two GM2 derived structures (**3** and **4**) are compared. In particular, the STD observed for the central unit of these two ligands is moderately different. In the case of **4**, the central sugar residue displays higher STD enhancements (> 75%) than those presented by the central sugar unit of **3** (35-75%). Hence, these changes might indicate that the presence of the sialic acid moiety at the Gal central unit of GM2 structure modulates the binding mode of **3** against MGL. The sialic acid in **3** shows STD intensities around 35-55%, in line with those observed for protons at the central and reducing end of **3**.

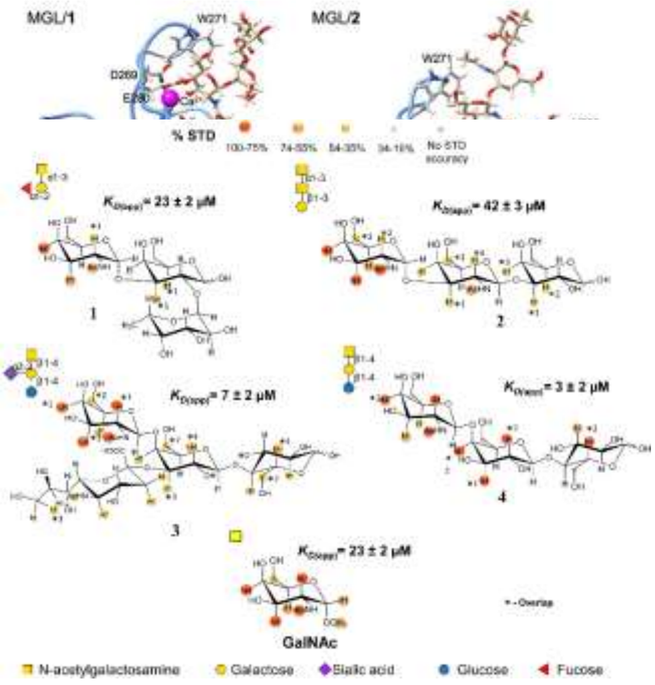


Figure 5. STD-based epitope mapping for **1-4** in presence of MGL-FL. The STD-derived epitope mapping of α-methyl GalNAc is included for comparison.[6] The apparent K_D ($K_{D(app)}$) values determined by STD-competition binding experiments are also displayed on top of the structures.

In general, the STD-NMR results point out subtle differences in the recognition process of the GalNAc-containing ligands **1-4** by MGL. To understand if those differences have impact on the binding strength, STD-based competition experiments were also carried out in order to estimate the corresponding apparent dissociation constants ($K_{D(app)}$) of the complexes, using α-methyl galactose (Gal) as the reference ligand with known affinity (900 µM), and employing the protocol previously described.[6] The intensity changes in the STD spectrum of α-methyl Gal in the presence of MGL were monitored after addition of **1-4** (see experimental section). The estimated $K_{D(app)}$ values are also displayed in Figure 5. The inspection of the $K_{D(app)}$ values shows that ligands **1-4** have an affinity to MGL with the same order of range than α-methyl GalNAc, and where the ligands **3** and **4** exhibit a slightly lower $K_{D(app)}$ values, indicating an improved affinity towards the lectin.

*A 3D view of the interaction of **1-4** to MGL*

To get a 3D view of the interaction of the ligands **1-4** to MGL molecular dynamics (MD) simulations on complexes MGL-CRD/**1-4** were carried out, using the 3D homology model previously generated for the MGL-CRD as the lectin structure (Figure 1D).[6] As previously described, the 3D model was built using the coordinates of the CRD of the mammalian asialoglycoprotein receptor (PDB code 1DV8) as a template.[18] The two binding modes A and B, with alternative relative orientations of GalNAc ring towards MGL, and previously proposed for the MGL/ α-methyl GalNAc complex were used to create the starting MGL/**1-4** model complexes.[6] In contrast to the MGL/ α-methyl GalNAc complex,[6] only binding mode A revealed to be stable for the MGL/**1-4** complexes (Figure S19-S22). Indeed, when mode B was used as starting geometry, important steric clashes are observed between the central/reducing end units of **1-4** and Loop 5 of MGL-CRD that destabilize the key interaction between OH-3/OH-4 and the Ca^{2+} ion. These results strongly suggest that, in contrast to α-methyl GalNAc, ligands **1-4** binds MGL with the GalNAc moiety binding mode A presentation, indicating that this binding geometry should be selected in the context of more complex and biological relevant GalNAc-containing ligands.

In the case of the complexes MGL-CRD/**1-4** and for binding mode A, the interactions between OH-3/OH-4 and he Ca^{2+} ion are maintained through the entire simulation time (Figure S23-S26). The φ/ψ distributions of the glycosidic linkages of ligands **1-4** were also monitored (Figure S27-S30) showing that the branched ligands **1** and **3** present a reduced flexibility around the φ/ψ glycosidic dihedral angles, while **2** shows a certain degree of flexibility around the glycosidic linkages.

According to the MD simulations (Figure 6), the GalNAc unit of ligands **1**-**4** is recognized by the protein through a similar pattern of interactions than α-methyl GalNAc (Figure SX-SX). Thus, the α-face of the sugar (H4, H5, H6a/b) and the methyl group of NHAc are engaged in CH/π interactions with the aromatic ring of W271 and the side chain of Y236, respectively (Figure S31-S38). In addition, significant hydrogen bonds are also formed in all the complexes involving OH-3 and OH-4 of GalNAc ring (Figure S39-S42). Additionally, for **1**, **2** and **4,** the hydrogen bond between OH-6 and D269 is populated around 14, 28 and 19%, respectively (a similar hydrogen bond is formed between the OH-6 of α-methyl GalNAc and D269, but only populated around 10%, Figure SX). In contrast, in compound **3**, OH-6 is simultaneously engaged in a hydrogen bond with D269 populated around 27% and with OH-9 of sialic acid populated around 35%, which might explain the difference in the chemical shift of D269 detected in the case of **3** (Figure 4A).

Figure 6. Representative frames obtained from 100 ns MD simulations for the different complexes. The protein is shown as blue ribbons and the ligands as sticks. The carbon atoms of the ligands are shown in brown. The carbon atoms of the protein that are interacting with the antigen are shown in gray. Ca^{2+} is shown as a magenta sphere. The key amino acids residues involved in the engagement of the ligands 1-4 are identified.

For compound **1** a low populated hydrogen bond involving the carbonyl group of the NHAc of GalNAc and the aromatic ring of H286 is detected (population *ca.* 14% of the total trajectory time). This result is similar to that observed for GalNAc (300 ns trajectory time, Figure SX-SX). Of note, no interactions between the two additional sugar units and the lectin are observed, which is in line with the STD-derived epitope mapping shown in Figure 5. The simulations propose a similar scenario for **2** although, in this case, the hydrogen bond between the GalNAc and H286 is almost negligible. As for antigen **1**, no significant interactions connecting the central and reducing end units of **2** and MGL are observed. Therefore, the binding mode proposed by the simulations corroborates the low STD signals recorded for these two extra residues in the Forssman antigen.

For ligand **3**, the hydrogen bond between GalNAc and H286 is also similar populated as in the case of GalNAc (12% of the trajectory). However, this residue can be also engaged in CH/π interactions with the reducing end of **3** (Figure S43, distance between the center of the involved rings of 5.5 ± 0.8 Å). Furthermore, there are alternative hydrogen bonds between the sidechain of K264 and hydroxyl groups OH-2 and OH-3 of the reducing Glc residue, with populations around 41% and 10%, respectively (Figure S44). This outcome can explain the differences detected in Δδcomb for K264 residue in presence of **3** (Figure 3C). The alternative hydrogen bonds with OH-2 and OH-3 of the Glc unit fits well with the double conformation detected for K264 in the ^{1}H,^{15}N HSQC measured in the presence of **3** (Figure 4A). K264 also interacts with OH-2 (population above 20%) and OH-3 (population around 15%) of the reducing end of **4** (Figure S45), and as for **3** this result is in agreement with the NMR data. Accordingly, to the 3D MGL-CRD model K264 and W291 are connected by hydrogen bonds therefore the CSP observed for the residue W291 in the presence of **3** and **4** (Figure 3C and 3D) with respect to the MGL-CRD/ α-methyl GalNAc complex can be explain by alterations in the chemical environment of K264 induced by the ligands **3** and **4**. Also Q267 xxxx

Noteworthy, in the case of ligand **4**, H286 is now interacting with GalNAc through a relevant hydrogen bond (with a population ca. 50%, Figure S42) and through a CH/π interaction with the reducing end of the ligand (Figure S46, distance between the center of the involved rings of 5.4 ± 0.7 Å). Hence, the variation of the frequency of hydrogen bond population involving H286 might explain the difference in Δδcomb of **4** for GalNAc (Figures 3D and Figure 4C). Particularly, the importance of H286 in the recognition of GM2-related compounds has been previously demonstrated,[7] which is fully compatible with the data presented herein. The present data reinforce the conclusion that MGL-CRD displays an extended binding pocket that is far from the QPD motif and where H286 is critical for the optimal recognition of the extended ligands. Furthermore, the engagement of the reducing end of GM2 derived molecules (**3** and **4**) by MGL, unraveled by the MD, is in total agreement with the increase of STD intensities at this region for **3** and **4** (Figure 5).

Finally, in the case ligand **3** additional interactions involving the sialic moiety and the guanidinum group of R298 (Figure S41) can justify the CSP observed by the residue Q297 in Figure 3.

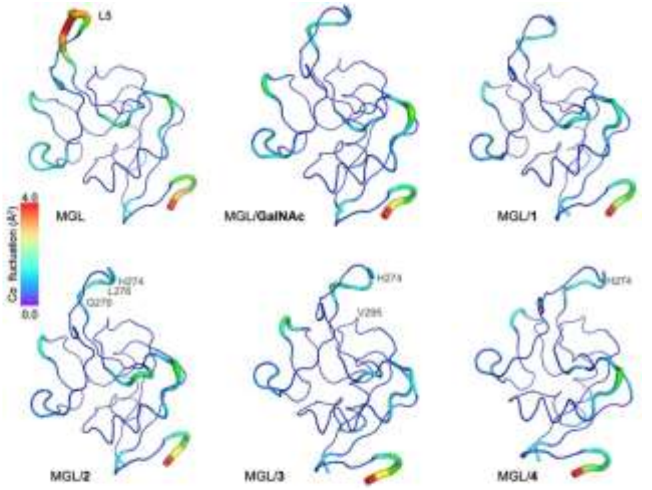
Analysis of the atomic fluctuations was also accomplished (Figure 7) and shows that, in general, MGL becomes more rigid upon ligand binding (α-methyl GalNAc and ligands **1-4**).

Figure 7. Atomic fluctuation (Cα) analysis of MGL and the different complexes derived from 100 ns MD simulations. The data correspond to the average protein structure throughout the simulations.

This result matches with the NMR collected data and stresses the conclusion that MGL structure is strongly altered upon ligand binding. In particular, for all the complexes, a strong decrease in the atomic fluctuations of the amino acids from 4.0 Å to a value less than 1.4 Å of loop L5 was clearly observed. Noteworthy, accordingly to this analysis the residues from H274 to G279 region located in the loop L5 still maintains atomic fluctuations even upon binding. Finally, fine differences in the conformation of the loop L5 can also be deduced depending of the tumor-associated GalNAc-containing ligand. These subtle differences might correlate with the detected alterations in the NMR analysis for some residues in this loop. This is the case of H274 for **2**, **3**, and **4**, or L276/G279 in the case of **2**. Lastly, the experimental difference displayed by V295 in **3** can be correlated with its close proximity to H274.

Conclusions

By decoding the structural features of the glycan-lectin interactions that occur between tumor-associated glycans and the immune system could lead to the design of novel and complementary anti-tumor immunotherapies. Specifically, MGL triggering by GalNAc-containing structures, commonly (over)expressed in tumors, is known to mediate immune suppression. [9–11,21] In this perspective, our studies provide compelling evidence that GalNAc binding induces significant structural and dynamics changes on the MGL-CRD structure. This result pinpoints the high plasticity of the MGL-CRD and might explain the ability of MGL to recognize different tumor-associated GalNAc-containing structures. Remarkably, our data also shows that a different degree of interactions, between distinct GalNAc-containing ligands and MGL takes place, reliant whether the GalNAc comes from blood group A antigen, Forssman antigen or GM2-derived structures. Therefore, the structure and dynamics of MGL-CRD are critically dependent of the precise GalNAc-containing structure. In fact, the signaling transmission of C-type immune lectins is usually accompanied by conformational changes on the CRD structure induced by the ligand binding.[22,23]



On this basis, it is tempting to speculate that by modulating the structure/dynamics of MGL-CRD it is possible to modify the signaling outcome in immune cells. From this point of view, our studies afford unique structural insights for the rational design of GalNAc-based structures to potential modulate the suppressive MGL anti-tumor immune responses.

Experimental Section

Recombinant expression and purification of MGL-FL and MGL-CRD

Recombinant expression and purification of MGL-FL

The expression and purification were based on the previously reported work with slightly differences noted below.[24] The gene for MGL-FL (Q61-H316, Figure S1) was synthesized by NZYTech using a codon optimization strategy for *E. coli* expression. This region was further sub cloned into the expression vector pET21a by NZYTech. The MGL was overexpressed in *E. coli* BL21(DE3) cells grown in LB medium and induced with 1mM IPTG overnight at 298 K. Cells were harvested, resuspended in 50mM TRIS buffer pH 8.0, 150 mM NaCl and lysed by sonication and centrifuged. MGL is expressed in insoluble inclusion bodies. The pellet was washed twice with 50mM TRIS buffer pH 8.0, 150 mM NaCl, 0.5% Triton X-100 and 10 mM EDTA, and once with water. The inclusion bodies were solubilized with 2 M of NH_4OH and dialyzed against 10 mM Tris buffer pH 7.5 containing 150 mM NaCl and $CaCl_2$ 20mM during 16h. Soluble recombinant MGL was purified using affinity chromatography performed using an α-Lactose agarose column (Sigma-Aldrich L7634) and the protein was eluted in buffer containing 150 mM of lactose. To remove the lactose, a desalting step (HiTrap^{TM} Desalting, GE Healthcare) followed by a dialysis against 10 mM Tris-HCl pH 7.5, 150 mM NaCl and 20 mM $CaCl_2$ during 16h was carried out. The purity was confirmed by SDS-PAGE. The protein yield was approximately 15 mg of pure and functional MGL-FL per liter of culture.

Recombinant expression and purification of MGL-CRD

The gene for MGL-CRD was synthesized by NZYTech using a codon optimization strategy of *E. coli* expression. The region corresponding to residues Q181–H316 (named MGL-CRD along the text, Figure S1) was sub cloned into the expression vector pET21a by NZYTech. The ^{15}N and ^{13}C labelled MGL-CRD was overexpressed in *E. coli* BL21(DE3) cells grown in M9 minimal medium with $^{15}NH_4Cl$ and ^{13}C-glucose as exclusive nitrogen and carbon sources, respectively, and induced with 1 mM IPTG at 298 K during 16h. The purification procedure was similar to that one used in MGL-FL. The purity of the MGL-CRD was confirmed by SDS-PAGE. The protein yield was approximately 2.5 mg of pure and functional MGL-CRD per liter of culture.

NMR Experiments

All NMR experiments were performed in a 600-MHz Bruker Avance III spectrometer equipped with 5mm inverse detection triple-resonance z-gradient cryogenic probe head.

Carbohydrates

Carbohydrates were purchased from Elicityl or Carbosynth. The ^1H NMR resonances of the GalNAc-containing fragments **1-4** were completely assigned through standard 2D-TOCSY (30 and 60 ms mixing time), 2D-ROESY/2D-NOESY (500 ms mixing time) and 2D-^1H,^{13}C-HSQC experiments. Typical concentrations were 3 mM for the heteronuclear experiments in D_2O Tris(D_{11})-DCl buffer solutions with pD 7.5. The assignment was accomplished at 298K (Table S2-S5). The resonance of 2,2,3,3-tetradeutero-3-trimethylsilylpropionic acid (TSP) was used as a chemical shift reference in the ^1H NMR experiments (δ TSP = 0 ppm).

Sequential backbone assignment

Sequence-specific backbone assignment of MGL-CRD was performed using data from the following experiments: 2D ^{1}H,^{15}N-HSQC, ^{1}H,^{13}C-HSQC, 3D HNCO, HN(CA)CO, HNCA, HN(CO)CA, HNCACB, CBCACONH, ^{15}N-edited NOESY-HSQC (mixing time, 80 ms), ^{15}N-edited TOCSY-HSQC (mixing time, 45 ms) spectra. All these experiments were acquired at 293 K using ^{13}C,^{15}N MGL-CRD at 350 µM in 10 mM TRIS buffer pH 7.5, 75 mM NaCl and 20 mM $CaCl_2$ in H_2O:D_2O 90:10, uncorrected pH meter reading 7.5, in absence and presence of GalNAc (MGL-CRD/GalNAc of 1:10 molar ratio). All data was processed in Bruker TopSpin 3.5 and analyzed with the assistance of computer aided resonance assignment (CARA) software.[25]

^{1}H,^{15}N HSQC carbohydrate titrations

^{1}H,^{15}N HSQC titration experiments were carried out by adding increasing amounts of the ligands GalNAc and GalNAc-containing fragments 1-4 to the MGL-CRD at constant concentration. The concentration of MGL-CRD was set to 100-200µM in 10 mM TRIS buffer, 75 mM NaCl and 20 mM $CaCl_2$ in H_2O:D_2O 90:10 at uncorrected pH meter reading 7.5. All the ligands were added to achieved protein/ligand (P/L) of 1:0.1; 1:0.25; 1:0.5; 1:1; 1;2; 1:5 and 1:10 molar ratios. The ^{1}H,^{15}N HSQC experiments were acquired at 293K with 2048x128 points and 32 scans in a spectral window of 9615.4 Hz (center at 2801 Hz) x 1946 Hz (center at 7175 Hz), in ^{1}H and ^{15}N sweep's width, respectively. The data was processed in Bruker TopSpin 3.5, CcpNMR Analysis Version 2.4[26] and CARA software.[25]

STD-NMR experiments

For the STD-NMR experiments it was used the MGL-FL construct. For this purpose the buffer was exchanged by means of several cycles of dilution-ultrafiltration (Vivaspin Turbo 15, 10.000 dalton molecular weight cut-off, Sartorius Stedim Biotech, Germany) to 10 mM perdeuterated Tris (D11)-DCl (Sigma-Aldrich) pD 7.5 in deuterated water with 20 mM $CaCl_2$, 75mM NaCl and and 0.09% sodium azide. STD-NMR experiments were performed with a 1:15 MGL-FL/ligand molar ratio at 310K. The STD-NMR spectra were acquired with 2048 transients with 16k data. An excitation sculpting module with gradients was employed to suppress the water proton signals. Selective saturation of the protein resonances (on resonance spectrum) was performed by irradiating at −0.5 ppm using a series of Eburp2.1000-shaped 90° pulses (50 ms, 1 ms delay between pulses) for a total saturation time of 2.0 s. For the reference spectrum (off resonance), the samples were irradiated at 100 ppm. A blank STD experiment only with protein, without any ligand, was recorded and the subtracted to the STD spectrum of the ligands allowing to eliminate the signal background from the protein. To determine the epitope mapping of each ligand shown in Figure 5 the STD intensities of each proton were normalized with respect to that with the highest response. The protons that could not be analyzed with accuracy were displayed by the blue sphere in the STD-based epitope mapping (Figure 5). Proton resonances that appear at the same chemical shift region of the spectrum are identified with * in Figure 5.

Dissociation constants estimated by STD-NMR

The dissociation constants (K_D) of GalNAc-containing fragments 1-4 were estimated by STD competition experiments, using methyl α-Galactose as reference ligand with known affinity (900 µM) and by employing the protocol previously described.[6] The STD intensity changes of the H4 and/or H2 proton of Gal in the STD spectrum were monitored after addition of the ligands 1-4.

Molecular Dynamics (MD) Simulations

MD simulations were performed using Amber16[26] implemented with ff14SB[28] and GLYCAM 06j-1 force fields.[29] The 3D models of the MGL complexed with ligands 1-4 were generated according to Marcelo et al.[6] In particular for ligand 3, the conformer with the *trans* arrangement between the carboxylate group of sialic acid moiety and C-3 of Gal central unit [ψ (C1sialic-C2sialic-O-C3Gal)=180º] was used as starting structure. This structure was selected since it represents the conformer with minimum energy in solution. All complexes were further immersed in a box with a 10 Å buffer of TIP3P[30] water molecules, Ca^{2+}[31] and neutralized by adding explicit counter ions (Na^{+}). A two-stage geometry optimization approach was used: a) first minimization involving exclusively solvent molecules and ions, and then b) second an unrestrained minimization of the entire system. The systems were then gently heated by incrementing the temperature from 0 to 300 K under a constant pressure of 1 atm and periodic boundary conditions. Harmonic restraints of 10 kcal/mol were applied to the solute, and the Andersen temperature coupling scheme [30] was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages. Long-range electrostatic effects were modelled using the particle-Mesh-Ewald method.[32] An 8-Å cut-off was applied to Lennard-Jones interactions. Each system was equilibrated for 2 ns with a 2 fs time step at a constant volume and temperature of 300 K. Production trajectories, under the same simulation conditions, were then run for 100 ns.

Acknowledgements

The authors acknowledge the Portuguese Foundation for Science and Technology (FCT-Portugal) for the projects IF/00780/2015 and PTDC/BIA-MIB/31028/2017 and for the UCIBIO project funding (UID/Multi/04378/2019). The authors also thank FCT-Portugal for the PhD grant attributed to AD (PD/BD/142847/2018) and for the Norma transitória DL 57/2016 Program Contract to JSD. FC thanks Ministério de Ciencia, y Universidades (No. RTI2018-099592-B-C21). JJB thanks the European Research Council for financial support (ERC-2017-AdG, Project number 788143-RECGLYCANMR) and Agencia Estatal Investigación (AEI, Spain) for funding (grants CTQ2015-64597-C2-1-P and Severo Ochoa Excellence Accreditation SEV-2016-0644). The NMR spectrometers are part of the National NMR Facility supported by FCT (ROTEIRO/0031/2013 - PINFRA/22161/2016, co-financed by FEDER through COMPETE 2020, POCI and PORL and FCT through PIDDAC). Finally, the authors also acknowledge Aldino Viegas for the fruitful discussions and assistance in the preparations of the figures.

Keywords: human macrophage galactose-type lectin • tumor carbohydrate associated antigens • molecular recognition • NMR spectroscopy • molecular modeling

[1] S. J. van Vliet, E. van Liempt, E. Saeland, C. A. Aarnoudse, B. Appelmelk, T. Irimura, T. B. H. Geijtenbeek, O. Blixt, R. Alvarez, I. van Die, and Y. van Kooyk, *Int. Immunol.*, **2005,** *17*, 661–669.

[2] S. J. van Vliet, E. van Liempt, T. B. H. Geijtenbeek, and Y. van Kooyk, *Immunobiology*, **2006,** *211*, 577–585.

[3] N. Higashi, K. Fujioka, K. Denda-Nagai, S. I. Hashimoto, S. Nagai, T. Sato, Y. Fujita, A. Morikawa, M. Tsuiji, M. Miyata-Takeuchi, Y. Sano, N. Suzuki, K. Yamamoto, K. Matsushima, and T. Irimura, *J. Biol. Chem.*, **2002**, *277*, 20686–20693.

[4] S. A. Jégouzo, A. Quintero-Martínez, X. Ouyang, Á. Dos Santos, M. E. Taylor, and K. Drickamer, *Glycobiology,* **2013,** 23, 7, 853–864.

[5] A. R. Kolatkar, W. I. Weis, and B. Chem, *Biochemistry*, **1996**, *271*, 6679–6685.

[6] F. Marcelo, F. Garcia-Martin, T. Matsushita, J. Sardinha, H. Coelho, A. Oude-Vrielink, C. Koller, S. André, E. J. Cabrita, H. J. Gabius, S. I.

Nishimura, J. Jiménez-Barbero, and F. J. Cañada, *Chem. - A Eur. J.,* **2014,** *20*, 16147–16155.

[7] F. Marcelo, N. Supekar, F. Corzana, J. C. van der Horst, I. M. Vuist, D. Live, G.-J. P. H. Boons, D. F. Smith, and S. J. van Vliet, *J. Biol. Chem.,* **2019,** *294*, 1300–1311.

[8] E. Saeland, S. J. Van Vliet, M. Bäckström, V. C. M. Van Den Berg, T. B. H. Geijtenbeek, G. A. Meijer, and Y. Van Kooyk, *Cancer Immunol. Immunother.*,**2007**, *56*, 1225–1236.

[9] S. J. van Vliet, S. Bay, I. M. Vuist, H. Kalay, J. J. García-Vallejo, C. Leclerc, and Y. van Kooyk, *J. Leukoc. Biol.,* **2013**, 94, 2, 315–323.

[10] S. J. van Vliet, S. I. Gringhuis, T. B. H. Geijtenbeek, and Y. van Kooyk, *Nat. Immunol.*, **2006,** *7*, 1200–1208.

[11] K. Lenos, J. A. C. M. Goos, I. M. Vuist, S. H. den Uil, P. M. D. Diemen, E. J. T. Belt, H. B. A. C. Stockmann, H. Bril, M. de Wit, B. Carvalho, S. Giblett, C. A. Pritchard, G. A. Meijer, Y. van Kooyk, R. J. A. Fijneman, and S. J. van Vliet, *Oncotarget,* **2015**, *6*, 26278–26290.

[12] D. Li, G. Romain, A.-L. Flamar, D. Duluc, M. Dullaers, X.-H. Li, S. Zurawski, N. Bosquet, A. K. Palucka, R. Le Grand, A. O'Garra, G. Zurawski, J. Banchereau, and S. Oh, *J. Exp. Med.,* **2012**, *209*, 109–121.

[13] S. Hakomori, S.-M. Wang, and W. W. Young*, Proc. Natl. Acad. Sci.*, **1977**, *74*, 3023–3027.

[14] S. Marionneau, B. Le Moullac-Vaidye, and J. Le Pendu, *Glycobiology*, **2002**, *12*, 851–856.

[15] N. Mortezai, H. N. Behnken, A. K. Kurze, P. Ludewig, F. Buck, B. Meyer, and C. Wagener, *Glycobiology*, **2013**, *23*, pp. 844–852.

[16] A. Ardá and J. Jiménez-Barbero, *Chem. Commun.,* ***2018,*** *54*, 4761–4769.

[17] F. H. Schumann, H. Riepl, T. Maurer, W. Gronwald, K. P. Neidig, and H. R. Kalbitzer, *J. Biomol. NMR*, **2007**, *39,* 275–289.

[18] M. Meier, M. D. Bider, V. N. Malashkevich, M. Spiess, and P. Burkhard, *J. Mol. Biol.,* **2000**, *300*, 857–865.

[19] D. S. Wishart, B. D. Sykes, and F. M. Richards, *Biochemistry*, **1992**, *31*, 1647–1651.

[20] J. L. Asensio, A. Ardá, F. J. Cañada, and J. Jiménez-Barbero, *Acc. Chem. Res.,* **2013**, *46*, 946–954.

[21] N. M. Sahasrabudhe, J. C. van der Horst, V. Spaans, G. Kenter, C. de Kroon, T. Bosse, S. J. van Vliet, and E. S. Jordanova, *Front. Oncol.,* **2019**, *9*, 1–8.

[22] A. A. Watson, A. A. Lebedev, B. A. Hall, A. E. Fenton-May, A. A. Vagin, W. Dejnirattisai, J. Felce, J. Mongkolsapaya, A. S. Palma, Y. Liu, T. Feizi, G. R. Screaton, G. N. Murshudov, and C. A. O'Callaghan*, J. Biol. Chem.,* **2011**, *286*, 24208–24218.

[23] F. Probert, S. B. M. Whittaker, M. Crispin, D. A. Mitchell, and A. M. Dixon, *J. Biol. Chem.,* **2013**, *288,* 22745–22757.

[24] N. Suzuki, K. Yamamoto, S. Toyoshima, T. Osawa, and T. Irimura, *J. Immunol.,***1996,** *156,* 128–135.

[25] R. Keller, The computer aided resonance assignment tutorial. **2004**, pp. 1-81.

[26] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, and E. D. Laue, *Proteins Struct. Funct. Genet.,* **2005**, *59*, 687–696.

[27] D. A. Case, R. M. Betz, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, H. T. Nguyen, I. Omelyan, A. Onufriev, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, and P. A. Kollman, AMBER 2016, University of California, San Francisco. **2016**.

[28] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, and C. Simmerling, *Proteins Struct. Funct. Genet.,* **2006***, 65*, 712–725.

[29] S. Genheden and U. Ryde, *J. Comput. Chem.,* **2010**¸ *32*, 187–195.

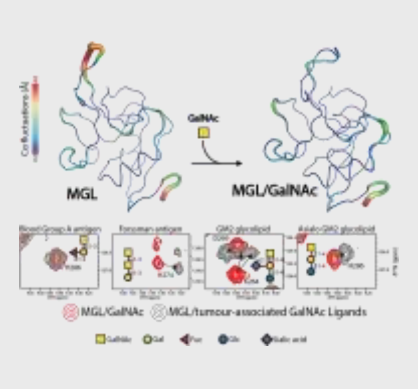
[30] T. A. Andrea, W. C. Swope, and H. C. Andersen, *J. Chem. Phys.*, **1983,** *79*, 4576–4584.

[31] G. M. Bradbrook, T. Gleichmann, S. J. Harrop, J. Habash, J. Raftery, J. Kalb, J. Yariv, I. H. Hillier, and J. R. Helliwell, *J. Chem. Soc. - Faraday Trans.,* **1998**, *94*, 1603–1611.

[32] T. Darden, D. York, and L. Pedersen, *J. Chem. Phys.,***1993**, *98*, 10089, 1993.

FULL PAPER

Flexibility orders binding: The carbohydrate recognition domain of human macrophage galactose-type lectin (MGL) is highly flexible and its structure and dynamics is strongly altered upon GalNAc binding. This intrinsic plasticity dictates the ability of MGL to recognize different tumor-associated GalNAc-containing ligands and key differences on MGL binding are observed depending on the GalNAc presentation.



Ana Diniz^{+}*, Helena Coelho*$^{+}$*, Jorge S. Dias, Sandra J. van Vliet, Jesús Jiménez-Barbero, Francisco Corzana, Eurico J. Cabrita and Filipa Marcelo**

Page No. – Page No.

The Plasticity of Carbohydrate Recognition Domain Dictates the Exquisite Mechanism of Binding of Human Macrophage Galactose-Type Lectin