Identification of a secondary binding site in human macrophage galactose-type lectin by microarray studies: Implications for the molecular recognition of its ligands

Received for publication, July 19, 2018, and in revised form, November 26, 2018 Published, Papers in Press, November 30, 2018, DOI 10.1074/jbc.RA118.004957

Filipa Marcelo[‡1], **Nitin Supekar**[§], **Francisco Corzana**[¶2], **Joost C. van der Horst**[‖], **Ilona M. Vuist**[‖3], **David Live**[§], **Geert-Jan P. H. Boons**[§4], **David F. Smith**[**], and **Sandra J. van Vliet**[‖5]

From the [‡]*Departamento de Química, Faculdade de Ciências e Tecnologia, UCIBIO, REQUIMTE, 2829-516 Caparica, Portugal, the* [§]*Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30602, the* [¶]*Departamento de Química, Universidad de La Rioja, Centro de Investigación en Síntesis Química, E-26006 Logroño, Spain, the* [‖]*Department of Molecular Cell Biology and Immunology, Cancer Center Amsterdam, Amsterdam Infection and Immunity Institute, Amsterdam UMC, Vrije Universiteit Amsterdam, 1081 HZ Amsterdam, The Netherlands, and the* [**]*Department of Biochemistry, Emory Comprehensive Glycomics Center, Emory University School of Medicine, Atlanta, Georgia 30322*

Edited by Gerald W. Hart

The human macrophage galactose-type lectin (MGL) is a C-type lectin characterized by a unique specificity for terminal GalNAc residues present in the tumor-associated Tn antigen (α-GalNAc-Ser/Thr) and its sialylated form, the sialyl-Tn antigen. However, human MGL has multiple splice variants, and whether these variants have distinct ligand-binding properties is unknown. Here, using glycan microarrays, we compared the binding properties of the short MGL 6C (MGL^{short}) and the long MGL 6B (MGL^{long}) splice variants, as well as of a histidine-to-threonine mutant (MGL^{short} H259T). Although the MGL^{short} and MGL^{long} variants displayed similar binding properties on the glycan array, the MGL^{short} H259T mutant failed to interact with the sialyl-Tn epitope. As the MGL^{short} H259T variant could still bind a single GalNAc monosaccharide on this array, we next investigated its binding characteristics to Tn-containing glycopeptides derived from the MGL ligands mucin 1 (MUC1), MUC2, and CD45. Strikingly, in the glycopeptide microarray, the MGL^{short} H259T variant lost high-affinity binding toward Tn-containing glycopeptides, especially at low probing concentrations. Moreover, MGL^{short} H259T was unable to recognize cancer-associated Tn epitopes on tumor cell lines. Molecular dynamics simulations indicated that in WT MGL^{short}, His^{259} mediates H bonds directly or engages the Tn-glycopeptide backbone through water molecules. These bonds were lost in MGL^{short} H259T, thus explaining its lower binding affinity. Together, our results suggest that MGL not only connects to the Tn carbohydrate epitope, but also engages the underlying peptide via a secondary binding pocket within the MGL carbohydrate recognition domain containing the His^{259} residue.

The C-type lectin family is a large family of soluble and transmembrane proteins that share the specific recognition of carbohydrate structures, which in most cases is Ca^{2+}-dependent (1). Carbohydrate binding is facilitated via a common carbohydrate recognition domain (CRD).[6] The primary binding site within this CRD is composed of a conserved three-amino acid motif, which largely determines the class of glycans the lectin is able to engage (*e.g.* the Gln-Pro-Asp (QPD) motif predicts binding to galactose or GalNAc) (2). In addition, secondary binding sites fine-tune the carbohydrate recognition, thereby creating enormous diversity in C-type lectin specificity and function.

The human QPD-containing macrophage galactose-type lectin (MGL; CLEC10A or CD301) is exclusively expressed by macrophages and dendritic cells within the immune system (3). Triggering of MGL on these human dendritic cells conveys immune inhibitory signals, leading to production of the anti-inflammatory cytokines and the ability to block unwanted inflammatory responses (4–6). Humans harbor only one *MGL* gene, which is subject to extensive splicing, generating different long and short MGL isoforms. Alternative splicing mainly occurs around exon 6, encoding the terminal stalk region (7). The most abundantly expressed MGL subtype in dendritic cells is the short

This work was supported by a bridging grant under National Institutes of Health Grant U54GM62116 (to the Consortium for Functional Glycomics) and P41GM103390 (to the Research Resource for Integrated Glycotechnology). The authors declare that they have no conflicts of interest with the contents of this article. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This article contains Table S1 and Figs. S1–S3.

The MGL array data for this paper are available from the Glycan array data Consortium for Functional Glycomics.

[1] Supported by the FCT-Portugal IF project (IF/00780/2015) and the UCIBIO funding UID/Multi/04378/2013, co-financed by the FEDER (POCI-01-0145-FEDER-007728).

[2] Supported by Ministerio de Economía y Competitividad Projects CTQ2015-67727-R and UNLR13-4E-1931.

[3] Present address: Swammerdam Institute for Life Sciences, University of Amsterdam, 1098 XH Amsterdam, The Netherlands.

[4] Present address: Dept. of Pharmaceutical Sciences and Chemistry, Utrecht University, 3584 CG Utrecht, The Netherlands.

[5] To whom correspondence should be addressed: Dept. of Molecular Cell Biology and Immunology, Amsterdam UMC, Location VU University Medical Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands. Tel.: 31-204440355; E-mail: s.vanvliet@vumc.nl.

[6] The abbreviations used are: CRD, carbohydrate recognition domain; MGL, macrophage galactose-type lectin; LDN, LacdiNAc; LacNAc, *N*-acetyl-D-lactosamine; PAA, polyacrylamide; MD, molecular dynamics; 3D, three-dimensional; B4GALNT3, β1,4-*N*-acetylgalactosaminyltransferase III; GM2, GalNAcβ1–4(Neu5Acα2–3)Galβ1–4Glc; GM3, NeuAcα2,3Galβ1,4Glc-ceramide; GD2, GalNAcβ1–4(Neu5Acα2–8Neu5Acα2–3)Galβ1–4Glc; Gb4Cer, GalNAcβ1–3Gal.

MGL 6C splice variant (Fig. 1*A*, from here on denoted as MGL^{short}) (7). Another relatively abundant variant is the MGL 6B variant, which compared with the MGL^{short} contains an additional 27 amino acids in its stalk region (Fig. 1*A*, MGL^{long}).

The carbohydrate specificity of the MGL^{short} variant has been widely studied through surface plasmon resonance, glycan microarrays, and cellular binding assays, revealing its exclusive specificity for terminal GalNAc moieties, such as the Tn antigen (αGalNAc-Ser/Thr), the LacdiNAc structure (LDN; GalNAcβ1–4GlcNAc), and the GalNAc-Tyr moiety (8–12). In QPD-containing C-type lectins, the equatorial/axial configuration of the 3-OH and 4-OH groups in the GalNAc are crucial for glycan binding and chelation of the Ca^{2+} ion (13). Indeed, from glycan microarray analysis, GalNAc structures extended at the 3- or 4-OH position fail to bind MGL (8). It is noteworthy that MGL tolerates substitutions at the 6-OH position and therefore recognizes the sialylated Tn antigen (Neu5Acα2–6GalNAcα-Ser/Thr) (10, 12).

Mice contain two functional copies of the *MGL* gene, termed *MGL1* (CD301a) and *MGL2* (CD301b) (14). Despite their high homology and the shared QPD motif, MGL1 and MGL2 differ substantially in their carbohydrate specificities. MGL1 mainly interacts with Lewis X (Galβ1–4(Fucα1–3)GlcNAc) and Lewis A (Galβ1–3(Fucα1–4)GlcNAc) (15), whereas MGL2 recognizes terminal α-/β-GalNAc ((sialyl)Tn antigen, LDN), as well as the terminal galactose in the core 1/T antigen and core 2 *O*-glycan structures (15). These disparities in ligand recognition have been further elucidated by Sakakura *et al.* (16), showing that whereas the galactose moiety of Lewis X is engaged by the QPD motif in the MGL1 CRD, the fucose residue is bound by a secondary binding site composed of Ala^{89} and Thr^{111} (corresponding to Ala^{256} and Thr^{278} in the full-length MGL1; Fig. 1*A*). These amino acids are replaced by arginine and serine in MGL2, respectively. Although no crystal structure of human MGL is available, glycan accommodation by MGL^{long} has been studied by saturation transfer-difference NMR measurements and molecular modeling, revealing fine-structural insights into the interaction of Gal/GalNAc and MUC1-Tn-glycopeptides with MGL (17). These NMR binding studies have verified that galactose is a much weaker ligand than GalNAc. The NHAc group of GalNAc participates in additional H-bond and CH-π interactions, which are absent in the case of galactose. Moreover, an increased number of H bonds were formed when MGL engaged a Tn-glycopeptide; however, this was not reflected by an improved affinity, probably due to the subtle enthalpy–entropy equilibrium of any ligand–protein recognition event.

Our knowledge on MGL has substantially increased over the last decade; nevertheless, the carbohydrate fine specificities and ligand-binding properties of the different human MGL splice variants have never been directly compared and have been assumed to be identical. Here, we aimed to elucidate the molecular recognition features of human MGL^{short} and MGL^{long} isoforms by evaluating their carbohydrate specificities using glycan and Tn-peptide microarrays. In addition, we mutated the His^{259} in the CRD of human MGL, equivalent to Thr^{111} in MGL1, to Thr in the MGL^{short} variant (H259T) to elucidate whether this amino acid is also crucial for the exclusive GalNAc-binding properties of human MGL.

Results

Carbohydrate recognition profiles of MGL variants

The MGL 6B (MGL^{long}) and 6C (MGL^{short}) splice variants vary in a region of 27 amino acids present in the terminal stalk domain of MGL^{long} only (Fig. 1*A*). We constructed recombinant MGL-Fc proteins, consisting of the extracellular domains of MGL^{long} and MGL^{short} fused to the human IgG1-Fc domain for detection. Previously, Ala^{89} and Thr^{111} (Ala^{256} and Thr^{278} in the full-length MGL1) were shown to be crucial for the Lewis X preference of murine MGL1 (16). Based on these findings and our previous work using the MGL^{short} splice variant (8), we mutated human MGL^{short} at the corresponding amino acids (H259T and K237A) to elucidate whether these amino acids also contribute to the fine specificity of human MGL (Fig. 1*A*). In our initial screens, the MGL^{short} K237A/H259T double mutant showed identical binding characteristics as the MGL^{short} single H259T mutant (data not shown); therefore, we omitted the K237A/H259T mutant from further analyses. All MGL-Fc proteins displayed equal binding to three independent anti-MGL antibodies (Fig. S1*B*), confirming the correct folding and indicating that the H259T mutation did not induce significant structural changes.

We next screened the glycan array developed by the Consortium for Functional Glycomics (http://www.functionalglycomics.org),[7] containing 610 glycans of natural and synthetic origin. As reported previously (8), MGL^{short} recognized the *O*-glycans αGalNAc, sialyl-Tn, and core 6 (GlcNAcβ1–6GalNAcα); the helminth glycans LDN and its fucosylated derivate (LDNF; GalNAcβ1–4(Fucα1–3)GlcNAc); and the glycosphingolipid structures GM2 (GalNAcβ1–4(Neu5Acα2–3)Galβ1–4Glc) and GD2 (GalNAcβ1–4(Neu5Acα2–8Neu5Acα2–3)Galβ1–4Glc) (Fig. 1*B*; full results in Table S1). Although differences were subtle, MGL^{long} displayed significantly higher binding to core 6, LDN, and LDNF than MGL^{short} and significantly lower binding to GM2 and GD2, as deduced by glycan microarrays. In contrast, the MGL^{short} H259T failed to interact with sialyl-Tn, LDNF, GM2, or GD2 on the array, whereas αGalNAc, the core 6 structure, and LDN were still bound (Fig. 1*B*). None of the MGL variants interacted with the core 1/T antigen, LacNAc, Lewis X, lactosylceramide, or GM3 structures on this array.

Next, we performed ELISA-based binding assays using polyacrylamide (PAA)-coupled glycoconjugates. All MGL variants, including the mutant, strongly bound the PAA-coupled GalNAc (Fig. 1*C*). Recognition of PAA-coupled sialyl-Tn and LDN by MGL^{short} and MGL^{long} was comparable, whereas MGL^{short} H259T showed diminished binding to both. Again, no binding could be detected to galactose, core 1/T antigen, or Lewis X. In contrast to the glycan microarray, no differential binding between MGL^{short} and MGL^{long} to the LDN moiety could be observed. The discrepancy with glycan microarrays might be explained by the multivalent nature of the PAA-glycoconjugate (20–25 glycans/molecule) *versus* the monovalent presentation on the array.

[7] Please note that the JBC is not responsible for the long-term archiving and maintenance of this site or any other third party hosted site.

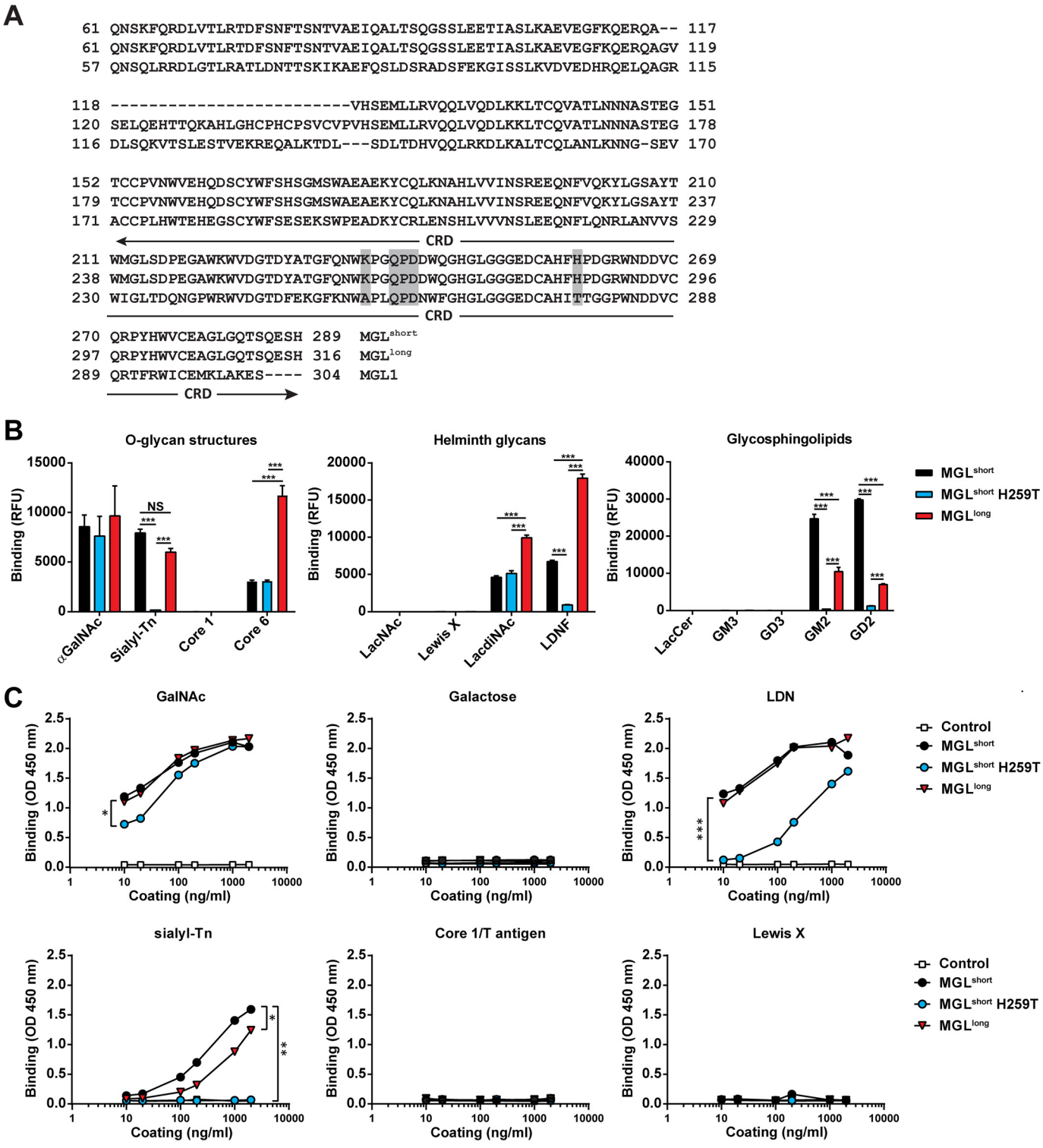


Figure 1. Differential GalNAc binding of MGL variants. *A*, amino acid sequence of the human MGL splice variants and the mouse MGL1. The stalk region and CRD are depicted, showing the 27-amino acid insertion in the stalk region of the MGL^{long} variant. *Boxed* in *gray* are the primary binding site (QPD) of MGL and the localization of the K237A and H259T mutations. *B*, binding of MGL variants determined by glycan array. Depicted are the binding to selected *O*-glycan core structures (*left*), Helminth-associated glycans (*middle*), and glycosphingolipids (*right*). Full glycan array results can be found in Table S1. Significance was calculated using two-way analysis of variance and Bonferroni's multiple-comparison test. Data represent mean ± S.E. *LacCer*, lactosylceramide. *C*, binding of the MGL variants to PAA-coated glycoconjugates was determined using an ELISA-based assay. Human IgG1 was used as a negative control. One representative experiment of three is shown. Significance was calculated by nonlinear regression followed by one-way analysis of variance and Bonferroni's multiple-comparison test. *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.005$.

Differential glycan binding of the MGL^{short} H259T mutant

Overall, the carbohydrate recognition profiles of MGL^{short} and MGL^{long} appeared to be quite similar; therefore, we focused our further investigations on the differential recognition of MGL^{short} and the MGL^{short} H259T mutant. Although the Thr^{111} amino acid (corresponding to His^{259} in human

Table 1
Glycan binding motifs of MGL^{short} and MGL^{short} H259T

The CFG glycan array data of MGL^{short} (*A*) and MGL^{short} H259T (*B*) (200 and 2 μg/ml) were analyzed using the data mining GlycoPattern software (https://glycopattern.emory.edu)[7] (18), and the identified glycan-binding motifs are listed below. Associated with each motif are the number of "binders" and "non-binders," which were determined using a statistical analysis (*z*-score) of the glycan array data based on the relative fluorescence units. A binder was defined as an MGL-binding structure that contains the motif. A non-binder contains the glycan motif but failed to interact with MGL. *Yellow squares*, GalNAc; *blue squares*, GlcNAc; *yellow circles*, galactose.

(A) MGL^{short}

Glycan structure	binders	non-binders
Galβ1-3GlcNAcβ1-6GalNAc (*Extended Core 6*)	6	2
GalNAcβ1-4GlcNAc (*LacdiNAc*)	9	7
GalNAcβ1-3Gal (*Globoside Gb4Cer*)	5	10

(B) MGL^{short} H259T

Glycan structure	binders	non-binders
Galβ1-3GlcNAcβ1-6GalNAc (*Extended Core 6*)	5	3
Galβ1-4GlcNAcβ1-6GalNAc (*Extended Core 6*)	6	26
GlcNAcβ1-6GalNAc (*Core 6*)	13	35

Table 2
High-affinity ligands for MGL^{short} and MGL^{short} H259T

Listed are the top five binding glycans (at 2 μg/ml) on the CFG glycan array for MGL^{short} (*A*) and MGL^{short} H259T (*B*). RFU, relative fluorescence units. *Yellow squares*, GalNAc; *blue squares*, GlcNAc; *yellow circles*, galactose; *red triangles*, fucose; *purple diamonds*, sialic acid.

(A) MGL^{short}

Glycan structure	RFU	stdev
Galβ1-4GlcNAcβ1-6GalNAcα (*Extended Core 6*)	54398	2348
(6S)GalNAcβ1-4GlcNAc (*Sulfated LacdiNAc*)	48863	3880
GalNAcα1-4(Fucα1-2)Galβ1-4GlcNAcβ (*Synthetic*)	43633	5609
GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-6GalNAc (*Extended Core 6*)	43374	4868
GalNAcβ	40039	2069

(B) MGL^{short} H259T

Glycan structure	RFU	stdev
GalNAcα	19927	3316
Galα1-3Galβ1-3GlcNAcβ1-6GalNAcα (*Extended Core 6*)	18669	546
Galα1-3Galβ1-4GlcNAcβ1-6GalNAcα (*Extended Core 6*)	18637	945
Neu5Acα2-3Galβ1-4GlcNAcβ1-6GalNAcα (*Extended Core 6*)	18005	1024
GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-6GalNAc (*Extended Core 6*)	16750	841

MGL) has been implicated in the exclusive Lewis X specificity of mouse MGL1 (16), neither the MGL^{short} H259T nor the K237A/H259T double mutant recognized Lewis X on the glycan array or in the ELISA-based assay, indicating that these two amino acids are not sufficient to convey Lewis X recognition. Nevertheless, the His^{259} seems essential for binding extended GalNAc structures, such as the sialyl-Tn and glycosphingolipids, and might therefore be a critical component of a secondary binding site in the MGL CRD.

To further extract carbohydrate determinants associated with MGL binding, we analyzed our array results using the publicly available Glycopattern web-based resource (https://glycopattern.emory.edu)[7] (18). We first employed the GlycanMotifMiner to discover glycan binding motifs for MGL^{short} and MGL^{short} H259T. A binding motif was defined as a structural glycan element that contained the motif and was classified as an MGL-binding glycan. For both MGL variants, we identified the extended core 6 *O*-glycan Galβ1–3GlcNAcβ1–6GalNAc as the main glycan-binding motif (Table 1). In addition, MGL^{short} recognized the LDN motif and the globoside glycolipid GalNAcβ1–3Gal (Gb4Cer). The MGL^{short} H259T only seemed to favor core 6–derived structures (Table 1). Although not identified as a glycan-binding motif, some additional glycans emerged as high-affinity ligands for MGL on the glycan microarray. These included the single α- and β-GalNAc and 6-sulfated LDN (Table 2).

To evaluate the differential glycan binding of MGL^{short} and MGL^{short} H259T, we used the side-by-side comparison tool in the Glycopattern software. Here, we identified 10 glycan structures that displayed high binding to MGL^{short}, but failed to interact with the MGL^{short} H259T (Table 3). Five of these contain the GalNAcα1–3(Fucα1–2)Galβ blood group A antigen. Furthermore, we observed diminished binding to the synthetic GalNAcβ1–3GalNAcα epitope, the ganglioside GM2, sialylated LDN, a synthetic ganglioside containing polysialic acid, and the Forssman antigen (GalNAcα1–3GalNAcβ) (Table 3).

Tn-peptide specificity of MGL variants

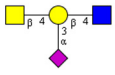
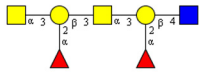
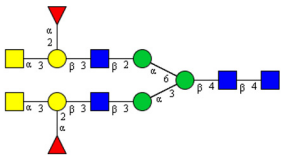
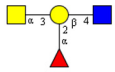
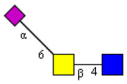
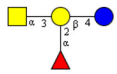
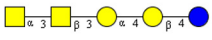
The lectin and antibody-mediated recognition of GalNAc moieties in Tn-containing mucin structures seems to be fine-tuned by the peptide backbone (19). Moreover, the underlying Ser of Thr residues in Tn antigen play a key role in the recognition process (20, 21). In the case of MGL, additional contacts with the peptide backbone upon binding of MUC1 Tn-glycopeptides have been described previously (17, 22). To address whether the His^{259} plays a role in the fine specificity of MGL toward Tn-containing glycopeptides, we employed a mucin glycopeptide array that included the known MGL ligands MUC1, MUC2, and CD45, which we interrogated using the MGL^{short}- and MGL^{short} H259T-Fc proteins (Fig. 2) (19). We chose to probe these MGL ligands based on their prior identification as major MGL ligands (6, 23, 24).

At low concentrations, binding of MGL^{short} was avid for CD45 and MUC2 glycopeptides, whereas the well-known MGL ligand MUC1 (17, 22, 23) displayed weaker binding to MGL^{short} (Fig. 2*A*). Curiously, MGL^{short} preferred the diglycosylated T*T* or T**X*T* MUC2 peptides (where the asterisk denotes the presence of a Tn antigen on Ser or Thr) compared with the mono- and triglycosylated peptides, whereas the diglycopeptide S*T* MUC1 displayed much weaker binding to MGL^{short} (Fig. 2*A*). Additionally, MGL^{short} may favor a Thr-based Tn antigen compared with the Ser-linked Tn antigen on this array (compare peptides 29/30 with 31/32).

Table 3

Differential glycan binding of MGL^{short} H259T

The CFG glycan array data of the MGL^{short} and MGL^{short} H259T (200 and 2 μg/ml) were analyzed using the data mining GlycoPattern software (https://glycopattern.emory.edu)[7] (18). Listed are glycans that failed to bind MGL^{short} H259T and showed binding of >10,000 relative fluorescence units for MGL^{short}. *Yellow squares*, GalNAc; *blue squares*, GlcNAc; *yellow circles*, galactose; *red triangles*, fucose; *purple diamonds*, sialic acid; *blue circles*, glucose; *green circles*, mannose.

Glycan structure	
GalNAcβ1-4(Neu5Acα2-3)Galβ1-4GlcNAcβ (*ganglioside GM2*)	
GalNAcβ1-3GalNAcα (*synthetic*)	
GalNAcα1-3(Fucα1-2)Galβ1-3GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ (*A antigen*)	
GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-2Manα1-6(GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1-2Manα1-3)Manβ1-4GlcNAcβ1-4GlcNAcβ (*A antigen*)	
GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ (*A antigen*)	
Neu5Acα2-6GalNAcβ1-4GlcNAcβ (*sialylated LDN*)	
GalNAcα1-3GalNAcβ (*synthetic*)	
GalNAcα1-3(Fucα1-2)Galβ1-4Glcβ (*A antigen*)	
GalNAcα1-3GalNAcβ1-3Galα1-4Galβ1-4Glcβ (*Globoside, Forssman antigen*)	
GalNAcα1-3(Fucα1-2)Galβ (*A antigen*)	

Interestingly, in the glycopeptide array, MGL^{short} H259T lost high-affinity binding toward MUC1 structures, even at high probing concentrations, while displaying a residual affinity toward MUC2- and CD45-derived Tn-glycopeptides (Fig. 2*A*), indicating that replacement of His^{259} by a threonine residue has an impact on the Tn-glycopeptide recognition by MGL. Indeed, the loss in affinity of the MGL^{short} H259T mutant toward Tn-glycopeptides could imply that the His^{259} residue interacts with the surrounding peptide backbone of Tn-glycopeptides.

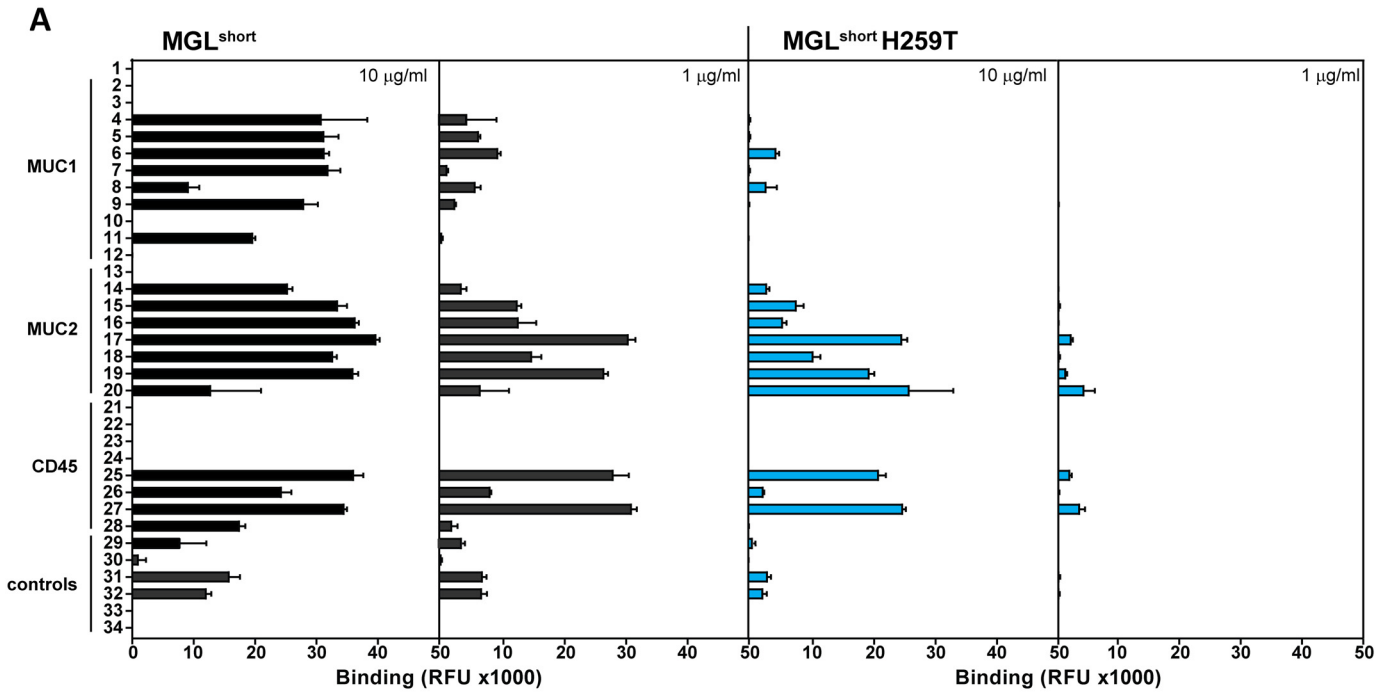
MGL^{short} H259T has a diminished capacity to recognize cellular ligands

To confirm that the MGL^{short} H259T mutant also has a diminished capacity to interact with cellular MGL ligands, we performed flow cytometry–based MGL-Fc binding assays to colorectal carcinoma and Jurkat cell lines. The colorectal carcinoma cell lines HT29 and SW1398 were previously shown to contain large numbers of MGL-binding ligands (25). Whereas MGL^{short} displayed high binding to the colorectal carcinoma cell lines, the MGL^{short} H259T failed to recognize the HT29 cells and displayed a 10-fold decreased binding to the SW1398 cells (Fig. 3*A*). The Jurkat cell line has a deficiency in the COSMC chaperone leading to an absence of T-synthase activity and thus a complete lack of *O*-glycan elongation beyond the Tn stage (26). Again, we observed a reduced binding of MGL^{short} H259T to Jurkat cells (Fig. 3*B*), verifying that also on cells, MGL^{short} H259T has a reduced ability to engage Tn antigen. Together, our data indicate that although the H259T MGL mutant can still recognize Tn antigen, it does so with a strongly reduced affinity.

Molecular dynamics simulations of MGL and the H259T mutant

The involvement of His^{259} in the MGL-mediated recognition of Tn-glycopeptides seems to indicate that this residue is part of what we dubbed the secondary binding site in the MGL CRD. To check whether this secondary binding site establishes additional interactions with the surrounding peptide backbone of Tn-glycopeptides, we conducted molecular dynamics (MD) simulations on both MGL/Tn-glycopeptide and MGL H259T/Tn-glycopeptide complexes. The Tn-glycopeptides MUC2B Ac-PTT*TPLK-NH_2 (peptide 15 in Fig. 2*B*) and MUC1 H_2N-TRPAPGST*APPA-NH_2 (peptide 6 in Fig. 2*B*) were selected based on the array results. A 3D model of the MGL CRD construct and the corresponding H259T mutant was built employing the homology model generated for MGL (residues Cys^{181}–Leu^{308}), as previously reported by us (17). Intensive 1-μs MD simulation was initially accomplished on both the WT MGL^{short} (MGL-His^{259}) and the H259T mutant (MGL-Thr^{259}) in explicit water and Ca^{2+} ions to evaluate the stability of both models. A structural ensemble of both proteins is depicted in Fig. S2. The calculated root mean square deviation shows that these molecules are equally stable and that the replacement of His^{259} by Thr does not induce major conformational perturbations (Fig. S2). Complexes of MGL-His^{259} and MGL-Thr^{259} with MUC2B and MUC1 glycopeptides were further generated as described by Marcelo *et al.* (17). Our previous MD and CORCEMA-ST analysis allowed us to conclude that the GalNAc monosaccharide interacts with MGL, employing two distinct binding modes (A and B), with alternative relative orientations of the galactosyl ring toward MGL (17). From a recognition point of view, histidine residues are commonly involved in H bonds and water-mediated interactions (27), and in GalNAc-binding mode A, the NHAc fragment of GalNAc is close to the His^{259} to potentially establish a stable H bond. Thus, MD simulations of the generated complexes in binding mode A were further submitted to evaluate their stability and the existence of intermolecular interactions involving either His^{259} (WT MGL) or Thr^{259} (mutant) and the surrounding peptide structure in GalNAc-containing peptides. In the case of MUC2B, transient H bonds were detected with different peptide residues (Fig. 4, *A–D*). Despite the fact that these H bonds are not highly populated, the sum of these H bonds in the surrounding peptide sequence can help to stabilize the complex and thereby significantly increase the binding affinity of MGL toward Tn-glycopeptides. No H bonds were detected between Thr^{259} and peptide backbone in the MD simulation analysis of mutant MGL-Thr259/MUC2B complex. The replacement of His^{259} by Thr thus may abolish these additional interactions. Furthermore, as expected, a stable H bond was detected between the CO of the NHAc group of GalNAc and the polar hydrogen Hϵ of His^{259}.

Interestingly, stable H bonds were not present between His^{259} and the peptide backbone of the MUC1-derived Tn-gly-



B

Chart ID	Peptide	Sequence [a]	Chart ID	Peptide	Sequence [a]
1	PBS		18	MUC2E	Ac-PT*TT*PLK-NH_2
2	MUC1 (A)	H_2N -AHGVTSAPDTRPAPGSTAPPA-NH_2	19	MUC2F	Ac-PTT*T*PLK-NH_2
3	MUC1 peptide 2R	H_2N-APGSTAPP-NH_2	20	MUC2G	Ac-PT*T*T*PLK-NH_2
4	MUC1 peptide 2	H_2N-APGS*T*APP-NH_2	21	CD45 exon 5 (A)	H_2N-GVSSVQTPHLPTHAD-NH_2
5	MUC1 peptide 3-1	H_2N-TRPAPGS*TAPPA-NH_2	22	CD45 exon 5 (B)	H_2N-ADSQTPSAGTDTQTF-NH_2
6	MUC1 peptide 3-2	H_2N-TRPAPGST*APPA-NH_2	23	CD45 exon 5 (C)	H_2N-QTFSGSAANA-NH_2
7	MUC1 peptide 4-1	H_2N-APGS*T*APPA-NH_2	24	CD45 exon 5 (D)	H_2N-NAKLNPTPGSNAISDV-NH_2
8	MUC1 peptide 4-2	H_2N-AHGVTS*APDT-NH_2	25	CD45 peptide 1	H_2N-GVS*S*VQT*PHL-NH_2
9	MUC1 peptide 4-3	H_2N-AHGVT*SAPDT-NH_2	26	CD45 peptide 2	H_2N-PHLPT*HADS-NH_2
10	MUC1 peptide 5R	H_2N-TSAPDTRPAP-NH_2	27	CD45 peptide 3	H_2N-SAGT*DT*QT*FS*GSA-NH_2
11	MUC1 peptide 5	H_2N-TSAPDT*RDAP-NH_2	28	CD45 peptide 4	H_2N-LNPT*PGS*NAI-NH_2
12	MUC1 (B)	H_2N–APDTRPAPGSTAPPAHGVTSAPDTRP-NH_2	29	Ser-GalNAc 1	H_2N-Ser(α-D-GalNAc)-NH_2
13	MUC2R	Ac-PTTTPLK-NH_2	30	Ser-GalNAc 2	H_2N-Ser(α-D-GalNAc)-OH
14	MUC2R	Ac-PT*TTPLK-NH_2	31	Thr-GalNAc 1	H_2N-Thr(α-D-GalNAc)-NH_2
15	MUC2B	Ac-PTT*TPLK-NH_2	32	Thr-GalNAc 2	H_2N-Thr(α-D-GalNAc)-OH
16	MUC2C	Ac-PTTT*PLK-NH_2	33	PBS	
17	MUC2D	Ac-PT*T*TPLK-NH_2	34	Biotin	

** denotes the presence of a Tn antigen on Ser or Thr*

Figure 2. Binding profile of MGL variants to the Tn-glycopeptide array. *A*, glycopeptides and corresponding parental peptides were printed at 100 μM and probed using the MGL^{short}-Fc (*left panels*) and MGL^{short} H259T-Fc (*right panels*) at the indicated concentrations. Relevant protein carriers are indicated in the figure. Data represent mean ± S.D. (*error bars*). *RFU*, relative fluorescence units. *B*, list of peptide backbones and Tn attachment sites of glycopeptides present on the Tn-glycopeptide array. *, presence of a Tn antigen on Ser or Thr.

copeptide; however, a water pocket could be observed between the carbonyl CO of Thr linked to GalNAc and the polar hydrogen of the aromatic ring of His^{259} (Fig. 5). In contrast, no water oxygen density was detected involving the Thr^{259} amino acid for the complex generated using the MGL^{short} H259T mutant model. Overall, the MD results suggest that His^{259} plays a critical role in the MGL recognition of Tn-glycopeptides by establishing intermolecular interactions with the peptide backbone.

Discussion

In this paper, we further defined the fine specificity of the human C-type lectin MGL through microarray analysis and MD simulations. Although MGL binding to the CFG glycan array has been analyzed before (8, 28) (see also the CFG glycan array data website: http://www.functionalglycomics.org),[7] the carbohydrate recognition profiles of the different human MGL splice variants and the H259T mutant have never been directly compared. In our glycan array analyses, all identified MGL ligands contained a nonreducing terminal GalNAc with the equatorial/axial 3-OH and 4-OH groups exposed, allowing the interaction with Ca^{2+} at the primary binding site. None of the 3- or 4-extended GalNAc structures were recognized by MGL in the glycan microarray, confirming its unique preference for terminal GalNAc residues (8–10, 12, 29). It is notewor-

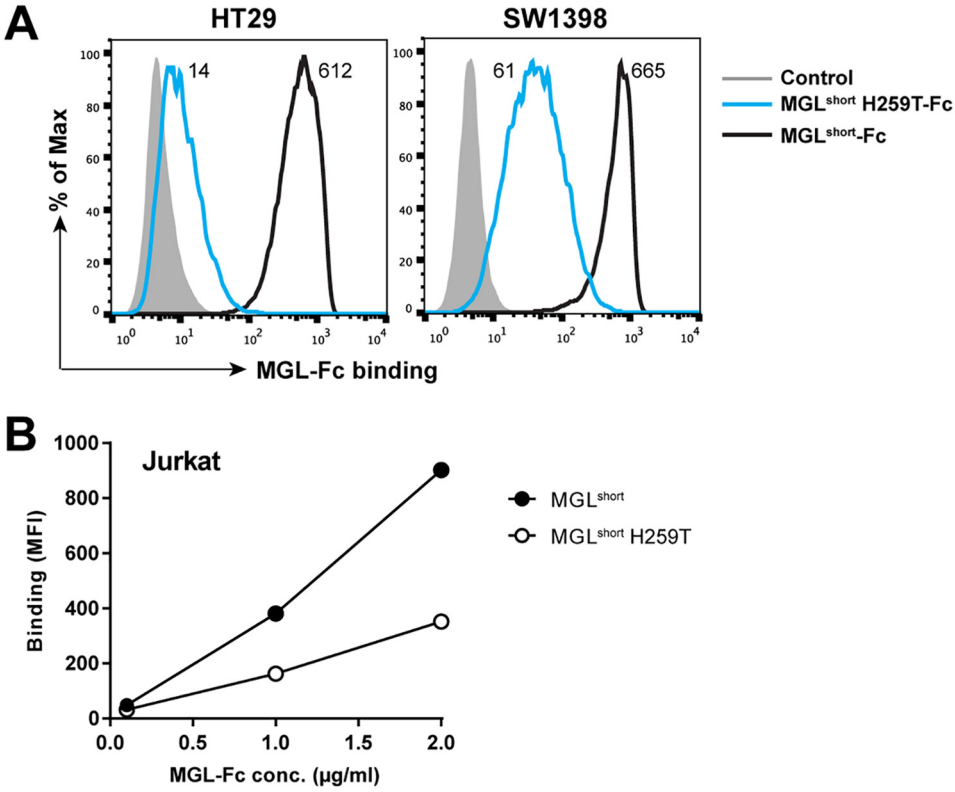


Figure 3. MGL^{short} H259T displays poor recognition of GalNAc-containing glycan epitopes on tumor cells. Colorectal cancer cell lines HT29 and SW1398 (*A*) and the COSMC-deficient Jurkat cell line (*B*) were incubated with 5 μg/ml or the indicated concentrations of the MGL-Fc variants for 30 min at 4 °C. Binding was detected using an FITC-labeled goat anti-human Fc secondary antibody and analyzed by flow cytometry. One representative experiment of three is shown.

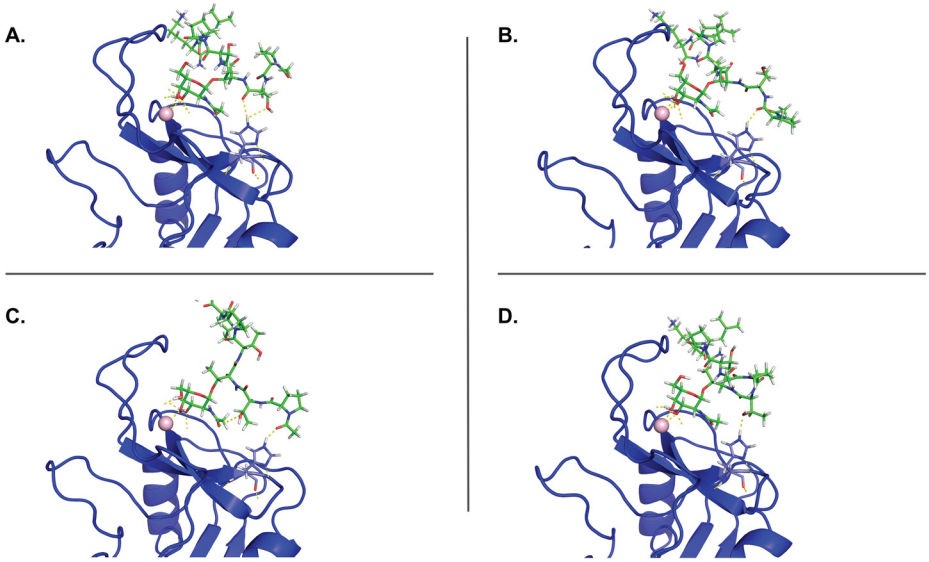


Figure 4. Selected frames of the MGL-His259/MUC2B complex extracted from 200-ns MD simulations. *A–D*, transient H bonds are established between the His^{259} and the peptide backbone (shown as *yellow dashed lines*). No H bonds were detected in the MGL-Thr259/MUC2B complex.

thy that extension at C-6 was allowed. Elongation or sulfation of the 3-OH or 4-OH abrogated MGL binding, thus explaining why some structures that do contain an identified MGL-binding motif were still categorized as nonbinders by the Glycopattern software (Table 1).

Strikingly, the identified MGL-binding elements are highly restricted in their expression pattern, suggesting that MGL may have some unique and distinctive functions within the human body. The LDN motif is mainly found on intestinal helminths (8) and in humans on glycodelin, a human glycoprotein with potent immunosuppressive and contraceptive activities (30). Sulfated LDN is a marker for pituitary glycoprotein hormones, such as lutropin (31). The enzyme that synthesizes LDN, β1,4-*N*-acetylgalactosaminyltransferase III (B4GALNT3) is overexpressed in colon cancer (32). Noticeably, both high MGL binding and overexpression of B4GALNT3 are associated with poor survival of colon cancer patients (25, 32). The preferred core 6 ligand has only been detected on human intestinal mucins and

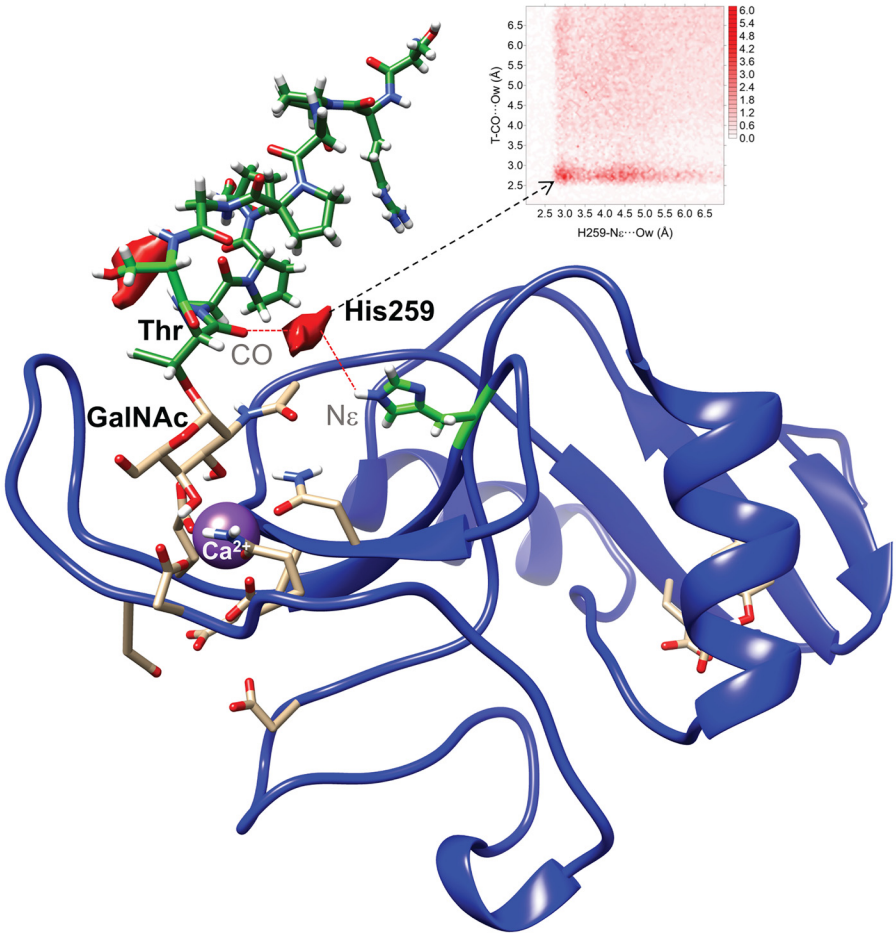


Figure 5. Water-mediated binding between Tn-glycopeptide MUC1 (H_2N-TRPAPGST*APPA-NH_2) and His^{259} of MGL^{short}. Shown is the 2D-rdf function derived from 200-ns MD simulations between carbonyl (*CO*) of the glycosylated Thr and the polar nitrogen Nϵ of the aromatic ring of His^{259} (*top*), together with the water density (shown in *red*) calculated for the MGL-His^{259}/MUC1–derived Tn-glycopeptide complex. No water density was detected in the case of MGL-Thr^{259}/MUC1-derived Tn-glycopeptide complex.

in the human ovary (33) (http://www.functionalglycomics.org),[7] whereas the Gb4Cer glycolipid is most abundant on human embryonic stem cells and up-regulated on colon cancer (34). Together, these studies indicate that MGL may recognize cancer cells through their glycolipid framework and/or LDN expression and may be involved in hormone action or turnover. Both potential roles have so far not been studied in relation to MGL biology, although the mouse MGL1 was recently shown to clear WT and hyposialylated von Willebrand factor (35), indicating that nonimmune-related functions might exist for human MGL as well.

Intriguingly, we could detect some subtle nuances in glycan specificity between the MGL^{short} and MGL^{long} splice variants, even though their CRDs are completely identical. The additional 27 amino acids in MGL^{long} stalk region could potentially affect multimerization of the receptor. The neck region of MGL^{short} has been shown to form trimers, generating a cluster of binding sites for glycans abundant on pathogens and tumor cells (28). So far, there are no indications that multimerization might be different for MGL^{long}. The MGL CRDs have some flexibility to engage differentially spaced glycans on the counteracting interface (28). It is therefore tempting to speculate that the 27-amino acid region alters the orientation and spacing of the individual MGL CRDs, thereby modulating the range of ligands that are able to interact. This phenomenon has already been demonstrated for DC-SIGNR, where the variable number of neck repeats in DC-SIGNR affects the orientation of the individual CRDs and subsequently the relative binding affinities for surface glycoproteins (36). Structural studies of MGL^{short} and MGL^{long} trimers are required to formally prove the effect of the extended neck region in MGL^{long}, as well as a more quantitative analysis of binding affinities of each interacting glycan with surface plasmon resonance.

Engagement of Tn antigens by MGL on human dendritic cells is thought to lead to immune tolerance via the MGL-induced secretion of the anti-inflammatory cytokine IL-10 and the ability to instruct the differentiation of suppressive $CD4^+$ T cells (4, 5). Moreover, through direct engagement of CD45 on activated T cells, MGL is able to inhibit T cell proliferation and cytokine release, while at the same time promoting T cell apoptosis (6). MGL also specifically recognizes the cancer-associated MUC1 and MUC2 (12, 23, 24), suggesting that MGL might promote immune evasion by Tn antigen–positive tumors as well (37). Therefore, we constructed a Tn-glycopeptide array harboring these immune-relevant MGL ligands. On our Tn-glycopeptide array, WT MGL^{short} seemed to prefer MUC2- and

CD45-derived glycopeptides over the MUC1 glycopeptides. However, the MUC1 glycopeptides in this particular array represent only a subset of possible glycosylation patterns of the MUC1 repeats; thus, we cannot exclude the possibility that other MUC1 glycopeptides may be dominant for MGL binding. MGL^{short} also appeared to favor diglycosylated peptides, as well as GalNAc-Thr over GalNAc-Ser. In general, lectin affinity increases with the increased valency of glycan epitopes in a molecule (38). From the four CD45 glycopeptides tested, MGL^{short} preferred the tri- and tetraglycosylated peptides over the mono- and diglycosylated peptides. The preference for Tn clustering by MGL might be caused by the use of dimeric MGL-Fc constructs, which contain two MGL molecules coupled to one Fc-tail. However, further experiments are required to determine the minimal distance required in a multi-Tn-peptide to accommodate two MGL proteins. Differential recognition of GalNAc-Ser and GalNAc-Thr has been observed before and might be explained by the relatively fixed position of the GalNAc-Thr compared with the more flexible GalNAc-Ser (19, 39–43).

We did, however, find major differences in the glycan-binding profiles of the MGL^{short} and its H259T mutant, whereby the MGL^{short} H259T was still able to recognize a single GalNAc monosaccharide (Fig. 1*C*) and core 6 structures (Table 2). Nevertheless, the H259T mutant displayed a diminished capacity to bind elongated structures, such as the LDN epitope and sialyl-Tn (Fig. 1), as well as the blood group A determinant, the ganglioside GM2, sialylated LDN, and the Forsmann antigen (Table 3). Blood group antigens are highly expressed on erythrocytes and in a variety of other tissues, including the vascular endothelium, mucus secretions, and epithelial surfaces. Intriguingly, in a rat model for colon cancer, expression of blood group A antigen increased apoptosis resistance and facilitated immune escape (44), a feature that may be linked to the immunosuppressive properties of MGL (3, 5). Forssman antigen has been demonstrated in stomach and colon cancer (45) and sialylated LDN in prostate cancer (46), again pointing to an influential role for MGL in tumor biology (12, 23, 25).

Also on the Tn-glycopeptide array, a reduced binding capacity of the MGL^{short} H259T mutant could be observed. The local sequence context around the GalNAc-modified Ser or Thr residue thus emerges as an important factor for MGL binding specificity, which is further corroborated by the loss of affinity of MGL^{short} H259T to particular glycan structures as well as Tn-containing glycopeptides and Tn-expressing cells. This loss of affinity directly implies that MGL recognizes not only the GalNAc monosaccharide, but also the underlying glycan or protein backbone, through a secondary binding site, of which the His^{259} amino acid is a key element. Through MD simulations, we could confirm the MGL-mediated recognition of the peptide backbone in MUC1 (17, 22) as well as in MUC2 through the His^{259} residue. Strikingly, the type of interactions, either through H bonds or water bridges, depended on the peptide sequence of the glycopeptide. We assume that also in the recognition of GalNAc-Tyr (11), MGL is able to engage the underlying tyrosine residue. Actually, the MGL binding mode shows clear parallels to the first structurally characterized Tn-glycopeptide–specific 237mAb antibody (21). Similar to MGL, the primary binding site of 237mAb is anchored by the Tn antigen, whereas the peptide backbone of podoplanin provides the additional specificity, thus conveying a dual recognition mode toward both Tn antigen and the podoplanin peptide backbone (21). The additional interactions involving the His^{259} residue, conveying fine specificity toward Tn-containing proteins or peptides, clearly illustrate the importance of exploring molecular recognition events for the optimal design of MGL-targeting structures for anti-tumor vaccines in light of the proposed role of MGL in impacting immunological responses through its highly exclusive recognition of tumor-associated glycan structures (47).

Experimental procedures

Generation and production of MGL-Fc constructs

The $MGL\text{-}Fc^{short}$ variant (corresponding to the 6C splice variant (7)) was constructed as described (8). cDNA encoding the extracellular domains of the MGL^{long} isomer (splice variant 6B (7)) was bought from Baseclear and cloned into the SigpIgG1-Fc vector. The MGL^{short} H259T mutant was generated by site-directed mutagenesis using the QuikChange II site-directed mutagenesis kit (Stratagene) and confirmed by sequencing. All MGL-Fc proteins were produced in CHO-K1 cells and purified from supernatants using HiTrap Protein A HP columns (GE Healthcare). MGL-Fc concentrations were determined by ELISA using human IgG1 as a standard, and purity was verified by Western blot analysis (Fig. S1*A*).

Antibody- and lectin-binding ELISAs

Anti-MGL antibodies 18E4 and 1G6.6 were generated as described before (3). Anti-MGL clone 125A10 was obtained from Dendritics. Isotype-matched control antibodies were purchased from eBioscience. To test the reactivity of these anti-MGL antibodies toward the different MGL-Fc proteins, NUNC MaxiSorp plates were coated overnight at room temperature with 5 μg/ml of the purified MGL-Fc proteins in 0.2 M $NaHCO_3$ (pH 9.2). After washing with TSM (20 mM Tris-HCl, pH 7.4, 150 mM NaCl, 2 mM $MgCl_2$, 1 mM $CaCl_2$), plates were blocked for 30 min at 37 °C with TSM containing 1% BSA. Anti-MGL antibodies and isotype controls were added to the wells at 5 μg/ml in TSM plus 0.5% BSA and incubated for 1 h at room temperature. Plates were washed with TSM plus 0.05% Tween 20, and bound antibodies were detected using peroxidase-labeled goat anti-mouse IgG Fc-specific secondary antibodies (0.3 μg/ml in TSM plus 0.05% Tween 20). TMB (100 μg/ml TMB and 0.003% H_2O_2 in 0.1 M NaAc, pH 4) was used as a substrate to visualize the presence of bound peroxidase, and optical densities were measured at 450 nm.

To investigate MGL-Fc binding properties, NUNC MaxiSorp plates were coated overnight at room temperature with the indicated concentrations of polyacrylamide-coupled glycoconjugates (Lectinity) in 0.2 M $NaHCO_3$ (pH 9.2). After washing with TSM, plates were blocked with 1% BSA plus TSM. MGL-Fc proteins (0.5 μg/ml) were added and incubated for 2 h at room temperature. Plates were washed with TSM plus 0.05% Tween 20, and bound MGL-Fc was detected using a peroxidase-labeled goat anti-human IgG-Fc (Jackson Immunoresearch Laboratories; 0.3 μg/ml in TSM plus 0.05% Tween 20). The reaction was developed using TMB (100 μg/ml TMB and

0.003% H_2O_2 in 0.1 M NaAc, pH 4), and optical densities were measured at 450 nm.

Glycan microarray analysis

The mammalian glycan array (version 5.1, containing 610 glycan structures) was developed by the Consortium of Functional Glycomics (www.functionalglycomics.org) from a library of natural and synthetic mammalian glycans with amino-linkers and printed in replicates of six onto *N*-hydroxysuccinimide–activated glass slides, forming covalent amide linkages. 70 μl of the MGL-Fc proteins (2 and 200 μg/ml in TSM containing 1% BSA and 0.05% Tween 20) was applied to the array and incubated at room temperature in a humidified chamber for 1 h. After extensive washing in TSM, bound MGL-Fc was detected using an Alexa 488–labeled anti-human IgG secondary antibody (1 μg/ml in TSM containing 1% BSA and 0.05% Tween 20). Fluorescence intensities were detected using a PerkinElmer Life Sciences ScanArray 5000 scanner, followed by image analyses using IMAGENE image analysis software (for representative images, see Fig. S3). The glycan array data (2 and 200 μg/ml) was analyzed using the Data mining feature in the GlycoPattern (https://glycopattern.emory.edu)[7] software (18) to identify glycan binding motifs and glycans that differentially bound the MGL variants. Glycan structures were drawn using the Glycoworkbench software (48). All glycan microarray analyses were carried out in accordance with the MIRAGE (minimum information required for a glycomics experiment) guidelines for reporting glycan microarray-based data (49).

Tn-glycopeptide array

Glycopeptides were synthesized by solid-phase peptide synthesis as carboxamides at the C terminus following published procedures (50). Those incorporating lysine were acetylated at the N terminus with the lysine side chain providing the amino group for the reaction with the *N*-hydroxysuccinimide–functionalized slide. The others had a free N terminus through which they were immobilized on the slide (19, 51). The glycopeptides were adjusted to the same concentration (100 μM) in printing buffer (300 mM sodium phosphate, pH 8.5) and spotted in replicates of 6. The Tn-glycopeptide was probed using MGL-Fc at a concentration of 10 and 1 μg/ml. Bound MGL-Fc was detected using an Alexa 488–labeled anti-human IgG secondary antibody at 5 μg/ml. Other assay conditions, buffers, and methods for imaging and processing of the data were identical to the screening of the glycan microarray and as described previously (Fig. S3) (19).

MGL-Fc staining of tumor cells

The human colon cancer cell lines SW1398 and HT29 were maintained in Dulbecco's modified Eagle's medium (Thermo Fischer Scientific) supplemented with 10% fetal calf serum and antibiotics. The Jurkat cell line was cultured in RPMI (Thermo Fisher Scientific) supplemented with 10% fetal calf serum and antibiotics. Colon cancer and Jurkat cell lines were incubated for 30 min at 4 °C with 5 μg/ml or the indicated concentrations of the different MGL-Fc variants. After washing with Hanks' buffered saline solution containing 0.5% BSA, the bound MGL-Fc was counterstained for 30 min at 4 °C using an FITC-labeled goat anti-human IgG Fc (Jackson Immunoresearch). Cells were analyzed by flow cytometry on a CyAn flow cytometer (Beckman Coulter). Dead cells were excluded by 7-aminoactinomycin D discrimination (Molecular Probes).

MD simulations

MD simulations were performed using Amber16 implemented with the force field ff14SB (52) and GLYCAM 06j-1 (53) to properly reproduce the conformational behavior of the glycopeptides. The 3D models of the MGL^{short} and the corresponding mutant H259T were generated according to Marcelo *et al.* (17). 1-μs MD simulations were performed to compare the stability of both models for MGL^{short} and the MGL^{short} H259T. The MUC2 glycopeptide Ac-PTT*TPLK-NH_2 (peptide 15; Fig. 2*B*) and the MUC1 glycopeptide H2N-TRPAPGST*APPA-NH_2 (peptide 6; Fig. 2*B*) were selected to generate the corresponding MGL^{short}/Tn-glycopeptide and MGL^{short} H259T/Tn-glycopeptide complexes.

The 3D models of MGL^{short} and corresponding H259T mutant, as well as the MGL/Tn-glycopeptide complexes, were immersed in a box with a 10-Å buffer of TIP3P (54) water molecules plus Ca^{2+} and neutralized by adding explicit counterions (Na^+). A two-stage geometry optimization approach was used. The first stage minimizes only the positions of solvent molecules and ions, and the second stage is an unrestrained minimization of all of the atoms in the simulation cell. 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 30 kcal/mol were applied to the solute, and the Andersen temperature coupling scheme (54) was used to control and equalize the temperature. The time step was kept at 1 fs during the heating stages. Water molecules are treated with the SHAKE algorithm, such that the angle between the hydrogen atoms is kept fixed. Long-range electrostatic effects are modeled using the particle mesh–Ewald method (55). 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 an additional 1 μs in the case of MGL^{short} and MGL^{short} H259T and 200 ns in the case of Tn-glycopeptide–containing complexes.

Statistics

Significant differences were evaluated using the GraphPad Prism software. A p value < 0.05 was considered to be significant.

Author contributions—F. M., F. C., D. L., and S. J. v. V. conceptualization; F. M., F. C., D. L., and S. J. v. V. formal analysis; F. M., N. S., F. C., J. C. v. d. H., I. M. V., D. L., D. F. S., and S. J. v. V. investigation; F. M., N. S., F. C., J. C. v. d. H., I. M. V., D. L., G.-J. P. B., D. F. S., and S. J. v. V. methodology; F. M., F. C., D. L., and S. J. v. V. writing-original draft; N. S., J. C. v. d. H., I. M. V., G.-J. P. B., and D. F. S. writing-review and editing; D. L. and S. J. v. V. funding acquisition; G.-J. P. B. supervision.

References

1. Drickamer, K., and Taylor, M. E. (2015) Recent insights into structures and functions of C-type lectins in the immune system. *Curr. Opin. Struct. Biol.* **34,** 26–34 CrossRef Medline

2. Zelensky, A. N., and Gready, J. E. (2005) The C-type lectin-like domain superfamily. *FEBS J.* **272,** 6179–6217 CrossRef Medline
3. van Vliet, S. J., van Liempt, E., Geijtenbeek, T. B., and van Kooyk, Y. (2006) Differential regulation of C-type lectin expression on tolerogenic dendritic cell subsets. *Immunobiology* **211,** 577–585 CrossRef Medline
4. van Vliet, S. J., Vuist, I. M., Lenos, K., Tefsen, B., Kalay, H., García-Vallejo, J. J., and van Kooyk, Y. (2013) Human T cell activation results in extracellular signal-regulated kinase (ERK)-calcineurin-dependent exposure of Tn antigen on the cell surface and binding of the macrophage galactose-type lectin (MGL). *J. Biol. Chem.* **288,** 27519–27532 CrossRef Medline
5. Li, D., Romain, G., Flamar, A. L., Duluc, D., Dullaers, M., Li, X. H., Zurawski, S., Bosquet, N., Palucka, A. K., Le Grand, R., O'Garra, A., Zurawski, G., Banchereau, J., and Oh, S. (2012) Targeting self- and foreign antigens to dendritic cells via DC-ASGPR generates IL-10-producing suppressive $CD4^+$ T cells. *J. Exp. Med.* **209,** 109–121 CrossRef Medline
6. van Vliet, S. J., Gringhuis, S. I., Geijtenbeek, T. B., and van Kooyk, Y. (2006) Regulation of effector T cells by antigen-presenting cells via interaction of the C-type lectin MGL with CD45. *Nat. Immunol.* **7,** 1200–1208 CrossRef Medline
7. Higashi, N., Fujioka, K., Denda-Nagai, K., Hashimoto, S., Nagai, S., Sato, T., Fujita, Y., Morikawa, A., Tsuiji, M., Miyata-Takeuchi, M., Sano, Y., Suzuki, N., Yamamoto, K., Matsushima, K., and Irimura, T. (2002) The macrophage C-type lectin specific for galactose/*N*-acetylgalactosamine is an endocytic receptor expressed on monocyte-derived immature dendritic cells. *J. Biol. Chem.* **277,** 20686–20693 CrossRef Medline
8. van Vliet, S. J., van Liempt, E., Saeland, E., Aarnoudse, C. A., Appelmelk, B., Irimura, T., Geijtenbeek, T. B., Blixt, O., Alvarez, R., van Die, I., and van Kooyk, Y. (2005) Carbohydrate profiling reveals a distinctive role for the C-type lectin MGL in the recognition of helminth parasites and tumor antigens by dendritic cells. *Int. Immunol.* **17,** 661–669 CrossRef Medline
9. Suzuki, N., Yamamoto, K., Toyoshima, S., Osawa, T., and Irimura, T. (1996) Molecular cloning and expression of cDNA encoding human macrophage C-type lectin: its unique carbohydrate binding specificity for Tn antigen. *J. Immunol.* **156,** 128–135 Medline
10. Mortezai, N., Behnken, H. N., Kurze, A. K., Ludewig, P., Buck, F., Meyer, B., and Wagener, C. (2013) Tumor-associated Neu5Ac-Tn and Neu5Gc-Tn antigens bind to C-type lectin CLEC10A (CD301, MGL). *Glycobiology* **23,** 844–852 CrossRef Medline
11. Gibadullin, R., Farnsworth, D. W., Barchi, J. J., Jr., and Gildersleeve, J. C. (2017) GalNAc-tyrosine is a ligand of plant lectins, antibodies, and human and murine macrophage galactose-type lectins. *ACS Chem. Biol.* **12,** 2172–2182 CrossRef Medline
12. Beatson, R., Maurstad, G., Picco, G., Arulappu, A., Coleman, J., Wandell, H. H., Clausen, H., Mandel, U., Taylor-Papadimitriou, J., Sletmoen, M., and Burchell, J. M. (2015) The breast cancer-associated glycoforms of MUC1, MUC1-Tn and sialyl-Tn, are expressed in COSMC wild-type cells and bind the C-type lectin MGL. *PLoS One* **10,** e0125994 CrossRef Medline
13. Kolatkar, A. R., and Weis, W. I. (1996) Structural basis of galactose recognition by C-type animal lectins. *J. Biol. Chem.* **271,** 6679–6685 CrossRef Medline
14. Tsuiji, M., Fujimori, M., Ohashi, Y., Higashi, N., Onami, T. M., Hedrick, S. M., and Irimura, T. (2002) Molecular cloning and characterization of a novel mouse macrophage C-type lectin, mMGL2, which has a distinct carbohydrate specificity from mMGL1. *J. Biol. Chem.* **277,** 28892–28901 CrossRef Medline
15. Singh, S. K., Streng-Ouwehand, I., Litjens, M., Weelij, D. R., García-Vallejo, J. J., van Vliet, S. J., Saeland, E., and van Kooyk, Y. (2009) Characterization of murine MGL1 and MGL2 C-type lectins: distinct glycan specificities and tumor binding properties. *Mol. Immunol.* **46,** 1240–1249 CrossRef Medline
16. Sakakura, M., Oo-Puthinan, S., Moriyama, C., Kimura, T., Moriya, J., Irimura, T., and Shimada, I. (2008) Carbohydrate binding mechanism of the macrophage galactose-type C-type lectin 1 revealed by saturation transfer experiments. *J. Biol. Chem.* **283,** 33665–33673 CrossRef Medline
17. Marcelo, F., Garcia-Martin, F., Matsushita, T., Sardinha, J., Coelho, H., Oude-Vrielink, A., Koller, C., André, S., Cabrita, E. J., Gabius, H. J., Nishimura, S., Jiménez-Barbero, J., and Cañada, F. J. (2014) Delineating binding modes of Gal/GalNAc and structural elements of the molecular recognition of tumor-associated mucin glycopeptides by the human macrophage galactose-type lectin. *Chemistry* **20,** 16147–16155 CrossRef Medline
18. Agravat, S. B., Saltz, J. H., Cummings, R. D., and Smith, D. F. (2014) GlycoPattern: a web platform for glycan array mining. *Bioinformatics* **30,** 3417–3418 CrossRef Medline
19. Borgert, A., Heimburg-Molinaro, J., Song, X., Lasanajak, Y., Ju, T., Liu, M., Thompson, P., Ragupathi, G., Barany, G., Smith, D. F., Cummings, R. D., and Live, D. (2012) Deciphering structural elements of mucin glycoprotein recognition. *ACS Chem. Biol.* **7,** 1031–1039 CrossRef Medline
20. Madariaga, D., Martínez-Sáez, N., Somovilla, V. J., Coelho, H., Valero-González, J., Castro-López, J., Asensio, J. L., Jiménez-Barbero, J., Busto, J. H., Avenoza, A., Marcelo, F., Hurtado-Guerrero, R., Corzana, F., and Peregrina, J. M. (2015) Detection of tumor-associated glycopeptides by lectins: the peptide context modulates carbohydrate recognition. *ACS Chem. Biol.* **10,** 747–756 CrossRef Medline
21. Brooks, C. L., Schietinger, A., Borisova, S. N., Kufer, P., Okon, M., Hirama, T., Mackenzie, C. R., Wang, L. X., Schreiber, H., and Evans, S. V. (2010) Antibody recognition of a unique tumor-specific glycopeptide antigen. *Proc. Natl. Acad. Sci. U.S.A.* **107,** 10056–10061 CrossRef Medline
22. Artigas, G., Monteiro, J. T., Hinou, H., Nishimura, S. I., Lepenies, B., and Garcia-Martin, F. (2017) Glycopeptides as targets for dendritic cells: exploring MUC1 glycopeptides binding profile toward macrophage galactose-type lectin (MGL) orthologs. *J. Med. Chem.* **60,** 9012–9021 CrossRef Medline
23. Saeland, E., van Vliet, S. J., Bäckström, M., van den Berg, V. C., Geijtenbeek, T. B., Meijer, G. A., and van Kooyk, Y. (2007) The C-type lectin MGL expressed by dendritic cells detects glycan changes on MUC1 in colon carcinoma. *Cancer Immunol. Immunother.* **56,** 1225–1236 CrossRef Medline
24. Napoletano, C., Rughetti, A., Agervig Tarp, M. P., Coleman, J., Bennett, E. P., Picco, G., Sale, P., Denda-Nagai, K., Irimura, T., Mandel, U., Clausen, H., Frati, L., Taylor-Papadimitriou, J., Burchell, J., and Nuti, M. (2007) Tumor-associated Tn-MUC1 glycoform is internalized through the macrophage galactose-type C-type lectin and delivered to the HLA class I and II compartments in dendritic cells. *Cancer Res.* **67,** 8358–8367 CrossRef Medline
25. Lenos, K., Goos, J. A., Vuist, I. M., den Uil, S. H., Delis-van Diemen, P. M., Belt, E. J., Stockmann, H. B., Bril, H., de Wit, M., Carvalho, B., Giblett, S., Pritchard, C. A., Meijer, G. A., van Kooyk, Y., Fijneman, R. J., and van Vliet, S. J. (2015) MGL ligand expression is correlated to BRAF mutation and associated with poor survival of stage III colon cancer patients. *Oncotarget* **6,** 26278–26290 Medline
26. Ju, T., and Cummings, R. D. (2002) A unique molecular chaperone Cosmc required for activity of the mammalian core 1 β3-galactosyltransferase. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 16613–16618 CrossRef Medline
27. Liao, J. H., Chien, C. T., Wu, H. Y., Huang, K. F., Wang, I., Ho, M. R., Tu, I. F., Lee, I. M., Li, W., Shih, Y. L., Wu, C. Y., Lukyanov, P. A., Hsu, S. T., and Wu, S. H. (2016) A multivalent marine lectin from *Crenomytilus grayanus* possesses anti-cancer activity through recognizing globotriose Gb3. *J. Am. Chem. Soc.* **138,** 4787–4795 CrossRef Medline
28. Jégouzo, S. A., Quintero-Martínez, A., Ouyang, X., dos Santos, Á., Taylor, M. E., and Drickamer, K. (2013) Organization of the extracellular portion of the macrophage galactose receptor: a trimeric cluster of simple binding sites for *N*-acetylgalactosamine. *Glycobiology* **23,** 853–864 CrossRef Medline
29. Iida, S., Yamamoto, K., and Irimura, T. (1999) Interaction of human macrophage C-type lectin with *O*-linked *N*-acetylgalactosamine residues on mucin glycopeptides. *J. Biol. Chem.* **274,** 10697–10705 CrossRef Medline
30. Dell, A., Morris, H. R., Easton, R. L., Panico, M., Patankar, M., Oehniger, S., Koistinen, R., Koistinen, H., Seppala, M., and Clark, G. F. (1995) Structural analysis of the oligosaccharides derived from glycodelin, a human glycoprotein with potent immunosuppressive and contraceptive activities. *J. Biol. Chem.* **270,** 24116–24126 CrossRef Medline
31. Fiete, D., Srivastava, V., Hindsgaul, O., and Baenziger, J. U. (1991) A hepatic reticuloendothelial cell receptor specific for SO4–4GalNAc

β1,4GlcNAcβ1,2Manα that mediates rapid clearance of lutropin. *Cell* **67,** 1103–1110 CrossRef Medline

32. Che, M. I., Huang, J., Hung, J. S., Lin, Y. C., Huang, M. J., Lai, H. S., Hsu, W. M., Liang, J. T., and Huang, M. C. (2014) β1,4-*N*-acetylgalactosaminyltransferase III modulates cancer stemness through EGFR signaling pathway in colon cancer cells. *Oncotarget* **5,** 3673–3684 Medline
33. Brockhausen, I., Schachter, H., and Stanley, P. (2009) *O*-GalNAc glycans. In *Essentials of Glycobiology*, 2nd Ed. (Varki, A., Cummings, R. D., Esko, J. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E., eds) Chapter 9, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY Medline
34. Joo, E. J., Weyers, A., Li, G., Gasimli, L., Li, L., Choi, W. J., Lee, K. B., and Linhardt, R. J. (2014) Carbohydrate-containing molecules as potential biomarkers in colon cancer. *OMICS* **18,** 231–241 CrossRef Medline
35. Ward, S. E., O'Sullivan, J. M., Drakeford, C., Aguila, S., Jondle, C. N., Sharma, J., Fallon, P. G., Brophy, T. M., Preston, R. J. S., Smyth, P., Sheils, O., Chion, A., and O'Donnell, J. S. (2018) A novel role for the macrophage galactose-type lectin receptor in mediating von Willebrand factor clearance. *Blood* **131,** 911–916 CrossRef Medline
36. Feinberg, H., Tso, C. K., Taylor, M. E., Drickamer, K., and Weis, W. I. (2009) Segmented helical structure of the neck region of the glycan-binding receptor DC-SIGNR. *J. Mol. Biol.* **394,** 613–620 CrossRef Medline
37. Rodríguez, E., Schetters, S. T. T., and van Kooyk, Y. (2018) The tumour glyco-code as a novel immune checkpoint for immunotherapy. *Nat. Rev. Immunol.* **18,** 204–211 CrossRef Medline
38. Dam, T. K., and Brewer, C. F. (2010) Lectins as pattern recognition molecules: the effects of epitope density in innate immunity. *Glycobiology* **20,** 270–279 CrossRef Medline
39. Madariaga, D., Martínez-Sáez, N., Somovilla, V. J., García-García, L., Berbis, M. Á., Valero-Gónzalez, J., Martín-Santamaría, S., Hurtado-Guerrero, R., Asensio, J. L., Jiménez-Barbero, J., Avenoza, A., Busto, J. H., Corzana, F., and Peregrina, J. M. (2014) Serine *versus* threonine glycosylation with α-*O*-GalNAc: unexpected selectivity in their molecular recognition with lectins. *Chemistry* **20,** 12616–12627 CrossRef Medline
40. Corzana, F., Busto, J. H., Jiménez-Osés, G., Asensio, J. L., Jiménez-Barbero, J., Peregrina, J. M., and Avenoza, A. (2006) New insights into α-GalNAc-Ser motif: influence of hydrogen bonding versus solvent interactions on the preferred conformation. *J. Am. Chem. Soc.* **128,** 14640–14648 CrossRef Medline
41. Corzana, F., Busto, J. H., Jiménez-Osés, G., García de Luis, M., Asensio, J. L., Jiménez-Barbero, J., Peregrina, J. M., and Avenoza, A. (2007) Serine *versus* threonine glycosylation: the methyl group causes a drastic alteration on the carbohydrate orientation and on the surrounding water shell. *J. Am. Chem. Soc.* **129,** 9458–9467 CrossRef Medline
42. Martínez-Sáez, N., Peregrina, J. M., and Corzana, F. (2017) Principles of mucin structure: implications for the rational design of cancer vaccines derived from MUC1-glycopeptides. *Chem. Soc. Rev.* **46,** 7154–7175 CrossRef Medline
43. Bermejo, I. A., Usabiaga, I., Compañón, I., Castro-López, J., Insausti, A., Fernández, J. A., Avenoza, A., Busto, J. H., Jiménez-Barbero, J., Asensio, J. L., Peregrina, J. M., Jiménez-Osés, G., Hurtado-Guerrero, R., Cocinero, E. J., and Corzana, F. (2018) Water sculpts the distinctive shapes and dynamics of the Tn antigens: implications for their molecular recognition. *J. Am. Chem. Soc.* **140,** 9952–9960 CrossRef Medline
44. Marionneau, S., Le Moullac-Vaidye, B., and Le Pendu, J. (2002) Expression of histo-blood group A antigen increases resistance to apoptosis and facilitates escape from immune control of rat colon carcinoma cells. *Glycobiology* **12,** 851–856 CrossRef Medline
45. Hakomori, S., Wang, S. M., and Young, W. W., Jr. (1977) Isoantigenic expression of Forssman glycolipid in human gastric and colonic mucosa: its possible identity with "A-like antigen" in human cancer. *Proc. Natl. Acad. Sci. U.S.A.* **74,** 3023–3027 CrossRef Medline
46. Hirano, K., Matsuda, A., Shirai, T., and Furukawa, K. (2014) Expression of LacdiNAc groups on *N*-glycans among human tumors is complex. *Biomed. Res. Int.* **2014,** 981627 Medline
47. Freire, T., Zhang, X., Dériaud, E., Ganneau, C., Vichier-Guerre, S., Azria, E., Launay, O., Lo-Man, R., Bay, S., and Leclerc, C. (2010) Glycosidic Tn-based vaccines targeting dermal dendritic cells favor germinal center B-cell development and potent antibody response in the absence of adjuvant. *Blood* **116,** 3526–3536 CrossRef Medline
48. Ceroni, A., Maass, K., Geyer, H., Geyer, R., Dell, A., and Haslam, S. M. (2008) GlycoWorkbench: a tool for the computer-assisted annotation of mass spectra of glycans. *J. Proteome Res.* **7,** 1650–1659 CrossRef Medline
49. York, W. S., Agravat, S., Aoki-Kinoshita, K. F., McBride, R., Campbell, M. P., Costello, C. E., Dell, A., Feizi, T., Haslam, S. M., Karlsson, N., Khoo, K. H., Kolarich, D., Liu, Y., Novotny, M., Packer, N. H., *et al.* (2014) MIRAGE: the minimum information required for a glycomics experiment. *Glycobiology* **24,** 402–406 CrossRef Medline
50. Movahedin, M., Brooks, T. M., Supekar, N. T., Gokanapudi, N., Boons, G. J., and Brooks, C. L. (2017) Glycosylation of MUC1 influences the binding of a therapeutic antibody by altering the conformational equilibrium of the antigen. *Glycobiology* **27,** 677–687 Medline
51. Liu, M., Barany, G., and Live, D. (2005) Parallel solid-phase synthesis of mucin-like glycopeptides. *Carbohydr. Res.* **340,** 2111–2122 CrossRef Medline
52. Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins* **65,** 712–725 CrossRef Medline
53. Kirschner, K. N., Yongye, A. B., Tschampel, S. M., González-Outeiriño, J., Daniels, C. R., Foley, B. L., and Woods, R. J. (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. *J. Comput. Chem.* **29,** 622–655 CrossRef Medline
54. Andrea, T. A., Swope, W. C., and Andersen, H. C. (1983) The role of long ranged forces in determining the structure and properties of liquid water. *J. Chem. Phys.* **79,** 4576–4584 CrossRef
55. Darden, T., York, D., and Pedersen, L. (1993) Particle mesh Ewald: an *N*.log(*N*) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089–10092 CrossRef

Identification of a secondary binding site in human macrophage galactose-type lectin by microarray studies: Implications for the molecular recognition of its ligands

Filipa Marcelo, Nitin Supekar, Francisco Corzana, Joost C. van der Horst, Ilona M. Vuist, David Live, Geert-Jan P. H. Boons, David F. Smith and Sandra J. van Vliet

J. Biol. Chem. 2019, 294:1300-1311.
doi: 10.1074/jbc.RA118.004957 originally published online November 30, 2018