# A comparative investigation of the mannose binding interface in DC-SIGN and MRC1 carbohydrate recognition domains with all-atom molecular dynamics simulations

*Sina Geissler and Sophie Sacquin-Mora**

Laboratoire de Biochimie Théorique, Université Paris-Cité, CNRS,

13 rue Pierre et Marie Curie, Paris 75005, France

*Corresponding author e-mail: sacquin@ibpc.fr

**Running title :** Modeling the DC-SIGN and MRC1 mannose binding sites

**Keywords:** Molecular Dynamics, carbohydrate recognition, protein-ligand interaction

A preliminary version of this work was deposited in bioRxiv, doi:

**Abstract**

Protein-carbohydrate interactions play a key role in numerous biological processes, including immune response, and glycan-based ligands that can target specific protein receptors on a cell surface represent promising candidates for therapeutics applications. For example, in retinoblastoma, the DC-SIGN mannose receptor (MR) is overexpressed on the surface of pathogenic cell and represents an interesting target for mannose-based ligands. At the same time, these ligands should not bind the MRC1 receptor, which is expressed by adjacent, healthy, retinal pigment epithelial cells. Therefore, the challenge remains to obtain a detailed picture of the recognition process between carbohydrates and proteins, in order to design effective and selective therapeutic compounds. In this work we used classical, all-atom Molecular Dynamics (MD) simulations to investigate the interaction between several mannose based ligands and the CRDs from DC-SIGN and MRC1. The analysis of the protein-carbohydrate contacts from the resulting trajectory highlights the variability of the mannose binding modes on both CRDs, and shows how the increased affinity of mannose for the MRC1 CRD can be related to a specific mannose binding state that is not accessible in the DC-SIGN CRD.

# 1. Introduction

Carbohydrates, or glycans, represent the most abundant class of organic molecules on Earth,[1] and a central component of life, both as a primary energy source[2] and as key player of cellular functions such as signaling and recognition.[3, 4] Glycan diversity covers a broad range of size, shape and functions,[5, 6] which is mirrored by the diversity of protein-carbohydrate interactions.[7-9] These interactions can be related to an organism's immune response,[10, 11] but also to numerous diseases, such as cancer,[12, 13] and host-pathogen recognition.[14-17] As a consequence, glycans, with their ability to bind specific protein targets, appear as promising candidates for therapeutic applications.[1, 18-21] One example of such biomedical applications is the development of targeted photodynamic therapies (PDTs), that use photosensitizers (PSs) where a porphyrin group is connected to one or several mannose groups.[22, 23] Mannose receptors (MRs), which are overexpressed on the surface of pathogenic cells,[24, 25] represent a potential target for these PSs. The design of adequate PSs however, requires to determine which protein receptor they should preferentially bind. In the case of Retinoblastoma, the DC-SIGN and MRC2 receptors were identified as PSs targets,[26] while PSs should not bind the MRC1 receptor, which is expressed by adjacent, healthy, retinal pigment epithelial cells.[27] All these receptors are transmembrane C-type lectins, a large family of carbohydrate binding, non-enzymatic, non-immunoglobulin proteins, which bind glycans in a $Ca^{2+}$ dependent manner via the carbohydrate recognition domain (CRD).[28-31]

Numerous experimental methods are currently available for the characterization of the structure and affinity of protein-carbohydrate interfaces,[32] including (but not limited to)

X-ray crystallography,[33, 34] NMR spectroscopy,[35] Cryo-EM[36] and isothermal titration calorimetry (ITC).[37] The high flexibility of glycans and the low-affinity of some carbohydrate recognition mechanisms[30] can however represent a challenge for experimental approaches, and computational analyses can help improving our understanding of these interactions.[38-41] In this context, we used classical, all-atom Molecular Dynamics (MD) simulations to investigate the interaction between several mannose based ligands and the CRDs from DC-SIGN and MRC1. The analysis of the protein-carbohydrate contacts from the resulting trajectory highlights the variability of the mannose binding modes on both CRDs, and shows how the increased affinity of mannose for the MRC1 CRD can be related to a specific mannose binding state that is not accessible in the DC-SIGN CRD.

## 2. Material and Methods

### *Starting models and MD simulations*

For the DC-SIGN-CRD, we used the PDB[42] crystal structure 2XR5[43] (DC-SIGN with mannose, 1.42 Å resolution) as the starting conformation. Simulations of the MRC1-CRD were based on the 7JUB crystal structure[44] of the CRD4 of MCR1 (MRC1 with methyl-mannoside, 1.20 Å resolution). For this system, methyl-mannoside was changed to mannose using the CHARMM-GUI[45] web server (https://www.charmm-gui.org/). In addition to single mannose, five mannose-based ligands were also modeled and parameterized on CHARMM-GUI,[46] using NMR-experimental structures provided our collaborators as first input (see Figure 1). The ligands were then superposed to the mannose group in the crystal structures using Pymol[47] (Version 3.1.6.1) to generate a

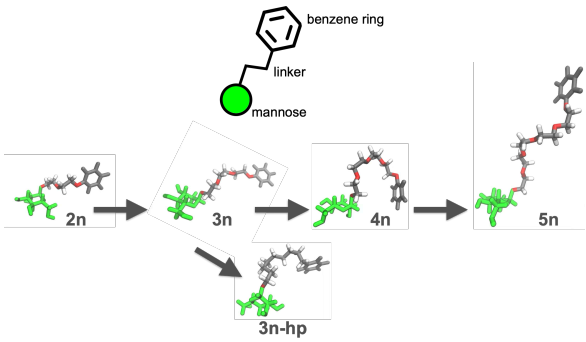


***Figure 1:*** *Graphic summary of the mannose based ligands tested in the simulations. For ligands 2n-5n, the linker between the mannose and benzene ring comprises 2 to 5 (-$CH_2$-$CH_2$-O-) units. For the 3n-hp ligand (lower line) the linker between the mannose and benzene ring comprises 3 (-$CH_2$-$CH_2$-$CH_2$) units. Figures 1, 2b-d, 3d, 5, 6, 7, 8a and 9 were prepared using Visual Molecular Dynamics.*[57]

starting conformation of the ligand bound to the receptor. All starting structures were solvated using the CHARMM-GUI server with TIP3P water.[48] The calcium cations initial positions (three for DC-SIGN and only one for MRC1) were taken from the crystal structures as they are important for the ligand binding. Potassium and chloride ions were added to the solution with a physiological concentration of 0.15 mol.$L^{-1}$. The pH of the system was set to 7 and the histidine residues were unprotonated.

All-atom molecular dynamics simulations were carried out using GROMACS,[49] and the CHARMM36m[50] forcefield, with a time step of 2 fs. First, energy minimization was done in 5000 steps with the steepest decent algorithm and a Verlet cutoff scheme. H-bonds were constrained with the LINCS[51] algorithm, receptor and ligand were position restrained. The equilibration step was done in six simulations with decreasing position restraints, the Verlet integrator, and 125000 steps per simulation. Pressure and temperature coupling was done using the Berendsen[52] algorithm. During the 500 ns of the production step, pressure coupling to 1.0 Bar was done using the Parrinello-Rahman

***Table 1:*** *Summary of the MD simulations performed in the study*

| | Ligand atoms | Protein atoms | $K^+$ and $Cl^-$ ions | Water molecules | Box size [$Å^3$] |
|---|---|---|---|---|---|
| DC | - | 2037 | 23, 27 | 8181 | $64^3$ |
| man-DC | 24 | 2037 | 27, 31 | 9417 | $67^3$ |
| 2n-DC | 48 | 2037 | 38, 42 | 13573 | 75x75x74 |
| 3n-DC | 55 | 2037 | 52, 56 | 18420 | 82x82x83 |
| 3nhp-DC | 61 | 2037 | 56, 60 | 19879 | 85x85x84 |
| 4n-DC | 62 | 2037 | 38, 42 | 13548 | $75^3$ |
| 5n-DC | 69 | 2037 | 47, 51 | 16563 | $80^3$ |
| MRC | - | 1920 | 21, 19 | 6748 | $60^3$ |
| man-MRC | 24 | 1920 | 21, 19 | 6736 | 61x61x59 |
| 2n-MRC | 48 | 1920 | 34, 32 | 11394 | $71^3$ |
| 3n-MRC | 55 | 1920 | 34, 32 | 11380 | $71^3$ |
| 3nhp-MRC | 61 | 1920 | 31, 29 | 10414 | $69^3$ |
| 4n-MRC | 62 | 1920 | 33, 31 | 10971 | $70^3$ |
| 5n-MRC | 69 | 1920 | 33, 31 | 10982 | 71x71x69 |

algorithm,[53] and we used the Nose-Hoover algorithm[54] for temperature coupling to room temperature (295.15 K). For each system we ran three replicas for the 500 ns production steps and frames were saved every 100 ps during production. Table 1 summarizes all the systems that were simulated during this study.

***Analysis of the MD Trajectories***

A sequence alignment of the two receptor pocket sequences was done on the uniprot server,[55] surface electrostatics were calculated on the APBS server (https://server.poissonboltzmann.org/)[56] using the crystal structure PDB files as input. Unbinding of a ligand was defined via the Colvars dashboard in VMD.[57, 58] The collective variable was defined as a distance between the pocket residues and the ligand, the ligand was defined as bound until the last full nanosecond before the collective

variable surpassed the value of 10 Å. MMPBSA[59] binding enthalpies between the ligand and the CRD were calculated using all the simulated frames with the ligand bound across all three replicas of each simulation as input data, for better sampling of the bound positions. RMSD values were calculated with GROMACS,[49] while MDAnalysis[60, 61] was used for performing Principal Component Analysis (PCA) of the trajectories, and for calculating hydrogen bonds. The different ligand binding states were defined based on characteristic distances between the OH-groups of the mannose moiety and specific residues of the CRDs binding pocket. The cutoff distance was 3.2 Å. In order to ensure sufficient sampling, binding enthalpies were only calculated for mannose binding states that lasted for 100 snapshots (i.e. 10 ns) or more.

## Coarse-grain Brownian Dynamics simulations

***Rigidity profile of a protein*** Coarse-grained Brownian Dynamics (BD) simulations were run using the ProPHet (Probing Protein Heterogeneity, available online at https://bioserv.rpbs.univ-paris-diderot.fr/services/ProPHet/) program[62-64]. In this approach, the protein is represented using an elastic network model (ENM). Unlike most common coarse-grained models where each residue is described by a single pseudoatom,[65] ProPHet uses a more detailed representation[66] that involves up to 3 pseudoatoms per residue and enables different amino acids to be distinguished. Pseudoatoms closer than the cutoff parameter $R_c$ = 9 Å are joined by Gaussian springs which all have identical spring constants of $\gamma_{struct}$ = 0.42 N.m$^{-1}$ (0.6 kcal.mol$^{-1}$.Å$^{-2}$). The springs are taken to be relaxed for the initial conformation of the protein. The simulations use an implicit solvent representation via the diffusion and random displacement terms in the equation of motion,[67] and hydrodynamic interactions are included through the diffusion tensor.[68]

Mechanical properties are obtained from 200,000 BD steps at an interval of 10 fs and a temperature of 300 K. The simulations lead to deformations of roughly 1.5 Å root-mean-square deviation with respect to the protein starting conformation (which, by construction, corresponds to the system's equilibrium state). The trajectories are analyzed in terms of the fluctuations of the mean distance between each pseudoatom belonging to a given amino acid and the pseudoatoms belonging to the remaining residues of the protein. The inverse of these fluctuations yields an effective force constant $k_i$ describing the ease of moving a pseudoatom with respect to the overall protein structure.

$$k_i = \frac{3k_B T}{\left\langle \left( d_i - \langle d_i \rangle \right)^2 \right\rangle},$$

where $\langle\rangle$ denotes an average taken over the whole simulation and $d_i = \langle d_{ij} \rangle_{j*}$ is the average distance from particle $i$ to the other particles $j$ in the protein (the sum over $j*$ implies the exclusion of the pseudoatoms belonging to residue $i$). The distance between the $C_\alpha$ pseudoatom of residue $i$ and the $C_\alpha$ pseudoatoms of the adjacent residues $i-1$ and $i+1$ are excluded since the corresponding distances are virtually constant. The force constant for each residue $k$ is the average of the force constants for all its constituent pseudo atoms $i$. We will use the term *rigidity profile* to describe the ordered set of force constants for all the residues of the protein.

## 3. Results and discussion

### *Structural similarity of the DC-SIGN and MRC1 glycan binding sites*

Despite a low sequence identity (around 32%), both CRDs present very similar crystallographic structures (1.3 Å rmsd on the heavy atoms), with a set of five

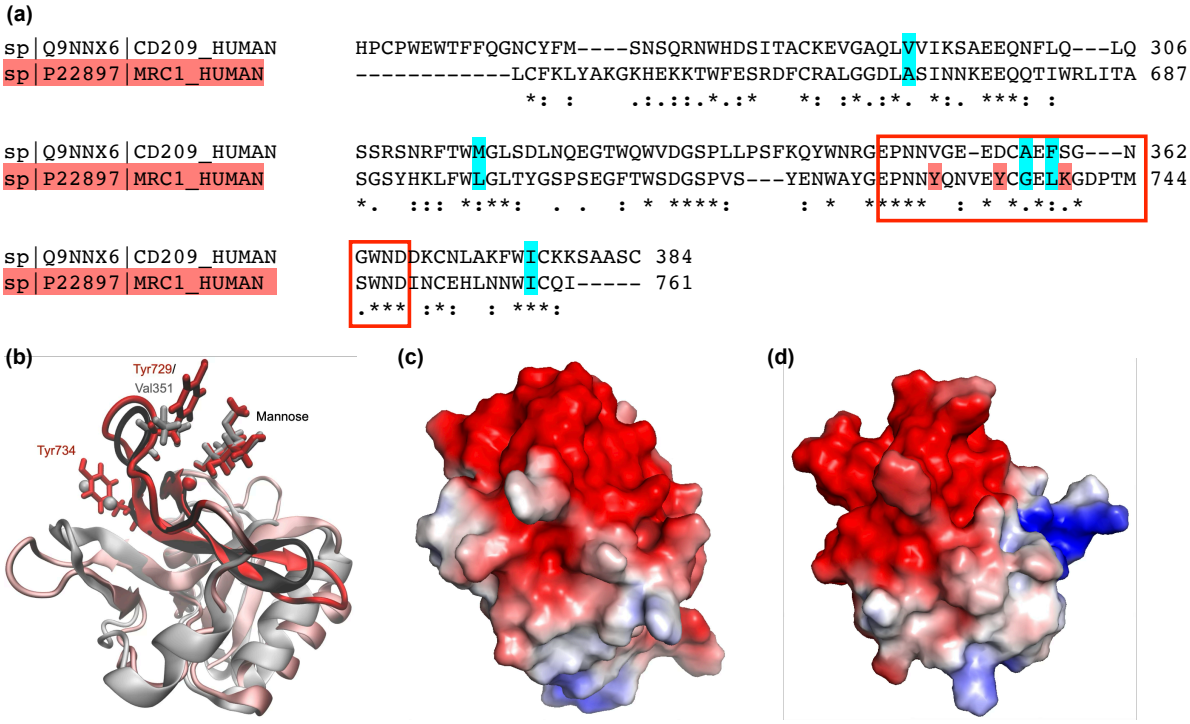


***Figure 2:*** *(a) Sequence alignement for the CRDs of DC-SIGN (upper line) and MRC1 (lower line). The red rectangle highlights the mannose binding pocket, the conserve rigid residues (also shown on Figure 3cd) are highlighted in cyan. (b) Aligned crystallographic structures for the DC-SIGN (in gray) and MRC1 (in red) CDRs. Surface electrostatic potentials for the CRDs, positive potentials are colored in blue and negative potentials in red (c) DC-SIGN, (d) MRC1.*

conserved residues lining the $Ca^{2+}$ cation involved in glycan binding, namely Glu347/725, Asn349/727, Glu354/733, Asn365/747 and Asp366/748 for DC-SIGN and MRC1 respectively (see the sequence alignment in Figure 2a). The most notable differences are the presence of two tyrosine residues in MRC1 (highlighted in red in Figure 2a). Tyr729 lies in the glycan binding pocket (instead of a valine for DC-SI), and Tyr734 lies in the pocket occupied by two additional $Ca^{2+}$ ions in DC-SIGN. In addition, the Lys739-Ser745 loop is slightly longer for MRC1 than its Ser360-Gly363 DC-SIGN counterpart, leading to a deeper groove for the glycan binding pocket (See Figure 2c). The Lys739 residue at the beginning of the loop, instead of Ser360 for DC-

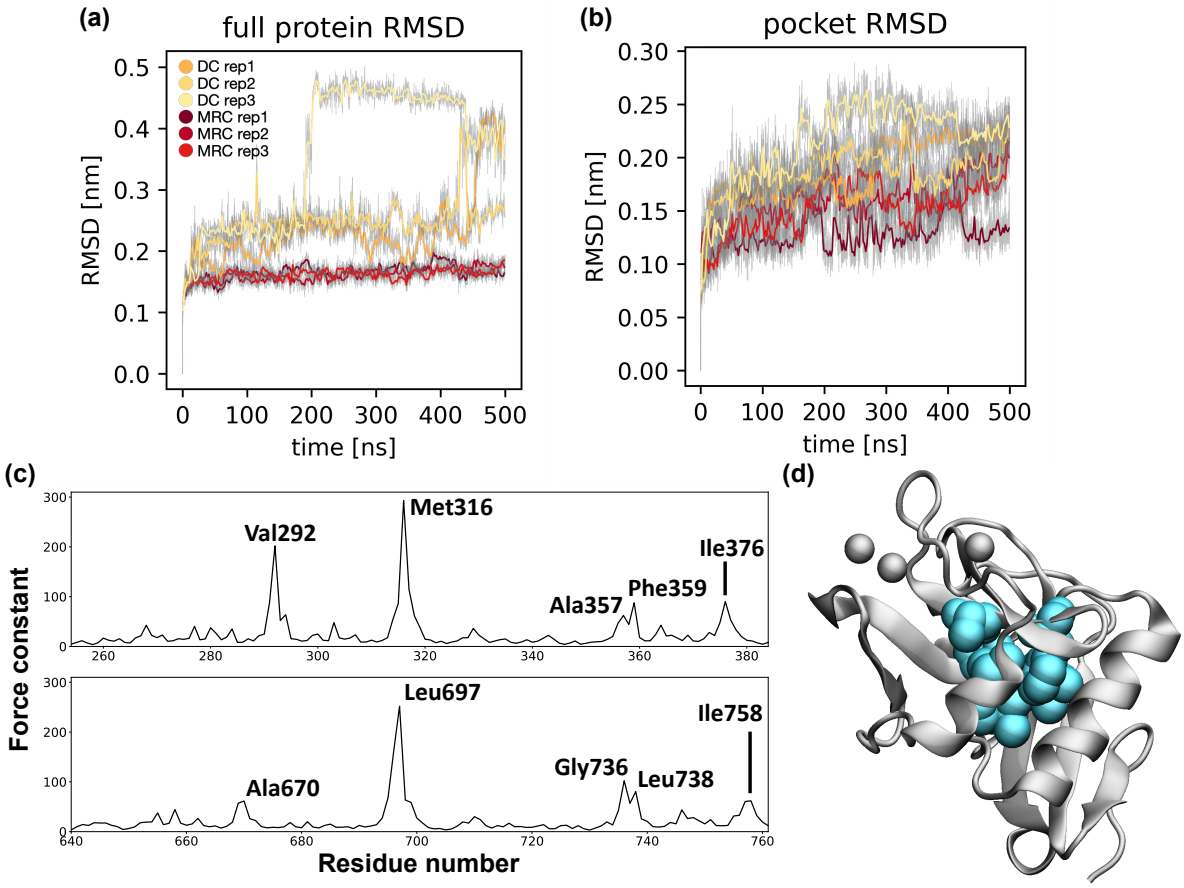


***Figure 3:*** *RMSD as a function of time for the DC-SIGN (yellow lines, three replicas) and MRC1 (red lines, three replicas) CRDs with mannose. (a) Full protein, (b) Mannose binding pocket only. For clarity, all colored curves were smoothed using a 20 frames sliding window. (c) Force constants (in* $mol^{-1} \cdot Å^{-2}$*) for the DC-SIGN (upper line) and MRC1 (lower line) CRDs. (d) Crystallographic structure of the DC-SIGN CRD with the rigid residues annotated in panel c highlighted in cyan.*

SIGN, leads to the presence of a positive patch on the CRD's surface for MRC1 (see Figure 2d). The RMSD evolution along the MD trajectories for the CRDs with mannose shows stable structures, with the RMSD increase along time for the D-SIGN replicas being due to the mobility of the N-terminal tail (see Figure 3a). The RMSD evolution for the glycan binding pocket only suggests that it is less flexible for MRC1 (see Figure 3b). PCA on the trajectories for the CRDs in their apo state show that the first five eigenvectors for DC-SIGN are mostly localised on the N-terminus and on the glycan

binding pocket, while for MRC1, the dynamics are more broadly distributed over the whole protein sequence (see Figure SI1). Finally, for both CRDs, the rigidity profile resulting from the coarse-grain simulations highlight a well-defined, conserved, rigid hydrophobic core (see Figure 3c-d), with the more compact DC-SIGN CRD being slightly more rigid on average.

***Comparing the DC-SIGN and MRC1 binding affinities for the different ligands.***

A striking results from the all-atom MD simulations is the instability of the mannose/ CRD interface for DC-SIGN, with the systematic unbinding of the ligand within the first 100 ns of each trajectory for all the replicas and all the tested ligands (see Figure 4). In the case of MRC1, the ligand/CRD interaction appears to be more stable, especially for the 3n and 4n ligands, which manage to stay bound in the glycan binding pocket for the complete duration of the simulation for at least one replica. For MRC1, the ligands linker section (which connects the mannose moiety to the benzene ring, see Figure 1) is in contact with the polar Tys729 (see Figure SI2), which might contribute to stabilizing the CRD/ligand interface. Hence, for the 3n case, we tested whether a less hydrophobic linker (with no oxygen atoms) could improve the ligand/DC-SIGN

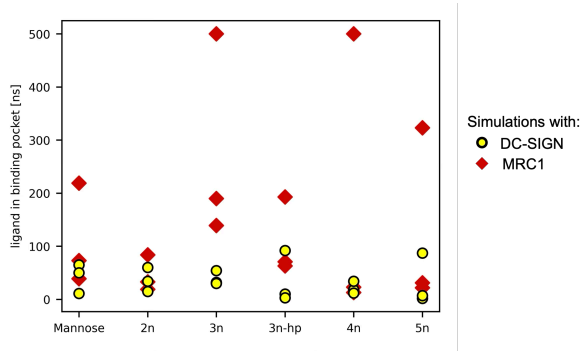


***Figure 4:*** *Ligand residency time in the glycan binding pocket for all possible ligands/CRD combinations (6*2*3 replicas).*

***Table 2:*** *Average Ligand/CRD binding enthalpies (in kcal.mol$^{-1}$)*

| Ligand | DC-SIGN | MRC1 |
|---|---|---|
| Mannose | -25.52 ± 1.19 (1260 frames) | -22.41 ± 0.21 (3313 frames) |
| 2n | -14.03 ± 0.34 (1090 frames) | -23.18 ± 0.38 (1364 frames) |
| 3n | -13.82 ± 0.39 (1160 frames) | -25.21 ± 0.14 (8290 frames) |
| 3n-hp | -16.08 ± 0.46 (643 frames) | -27.22 ± 0.23 (3274 frames) |
| 4n | -12.61 ± 0.50 (643 frames) | -27.02 ± 0.25 (5363 frames) |
| 5n | -17.49 ± 0.49 (964 frames) | -24.98 ± 0.30 (3763 frames) |

interaction while simultaneously disrupting the ligand/MRC1 interaction. The resulting 3n-hp ligand shows no significant increase of its affinity for DC-SIGN, while being slightly more unstable than the 3n ligand when interacting with MRC1. All the resulting ligand/CRD binding affinities are available in Table 2 and the values are of the same order of magnitude as those obtained in earlier simulation works.[69, 70]

***Investigating the mannose binding modes in the glycan binding pockets on DC-SIGN and MRC1***

The mannose binding mode in the CRDs crystallographic structure has already been described in the experimental literature,[43, 44] with the mannose OH3 and OH4 groups forming contacts with the calcium ion from the glycan binding pocket. In addition, the OH2 and OH3 groups interact with Glu354 and Glu733 (highlighted in red on Figure 5), while OH4 interact with Glu 347 and Glu 725 (highlighted in blue on Figure 5) for DC-SIGN and MRC1 respectively.

However, for both receptors, and all ligands, the mannose group leaves its crystallographic initial position at an early stage of the simulation and explores additional binding modes in the CRDs glycan binding pocket that are summarized in

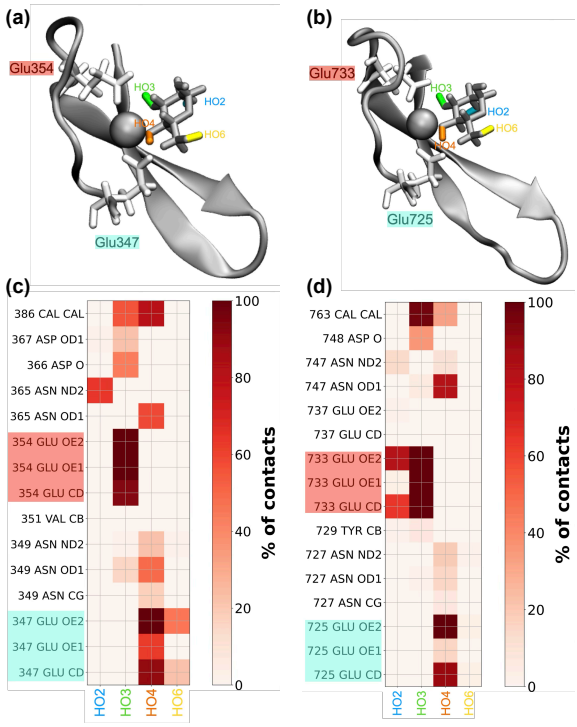


***Figure 5:*** *Mannose binding mode in the crystallographic structures (a) DC-SIGN, (b) MRC1. Corresponding contact maps (c) DC-SIGN, (d) MRC1. The blue/green/orange/ yellow color code to distinguish the mannose OH2/OH3/OH4/OH6 groups is kept in Figure 7.*

***Figure 6:*** *(a) Distribution of the mannose binding states for all the possible ligand/ CRD combinations (6*2 systems, the three replicas for each combination have been concatenated), blue: crystal position, green: A state, orange: B state (DC-SIGN specific), red: C state (MRC1 specific), lilac: other.*
*(b) MMPBSA binding affinities (in kcal.mol$^{-1}$) for the different ligand binding states on the DC-SIGN (circles) and MRC1 (diamonds) CRDs.*

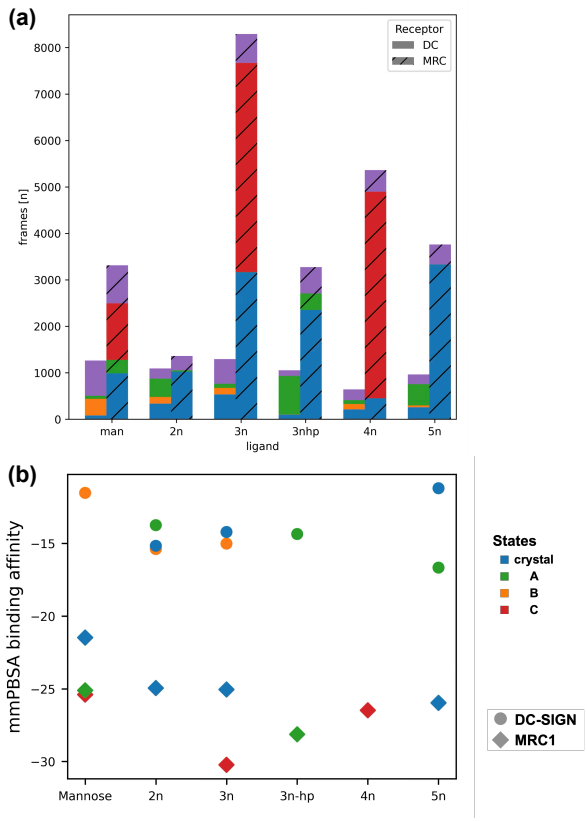

Figure 6. In *state A*, which was observed for both the DC-SIGN and MRC1 CRDs, the mannose moiety rotates so that its OH2 interacts with Glu 347/725, while both OH3 and OH4 form contacts with Glu 354/733 (see Figure 7ac). This binding mode can be related to the minor mannose binding mode described by Cramer,[71] and was also observed by Feinberg et al.[44] for fucose binding to the MRC1 CRD. Two other binding modes were observed for mannose in the CRDs glycan binding pockets. *State B* is specific to the DC-SIGN CRD and sees OH3 form contacts with the Ca2+ ion from the pocket (see Figure 7b). Glu354 is the only CRD residue forming noticeable contacts with the mannose moiety, via OH2 and OH3. This state is much less stable than the others (see the ligand/CRD binding affinities listed in Figure 6). Finally, *state C* was only observed for the mannose binding to the MRC1 CRD. In this case, the mannose

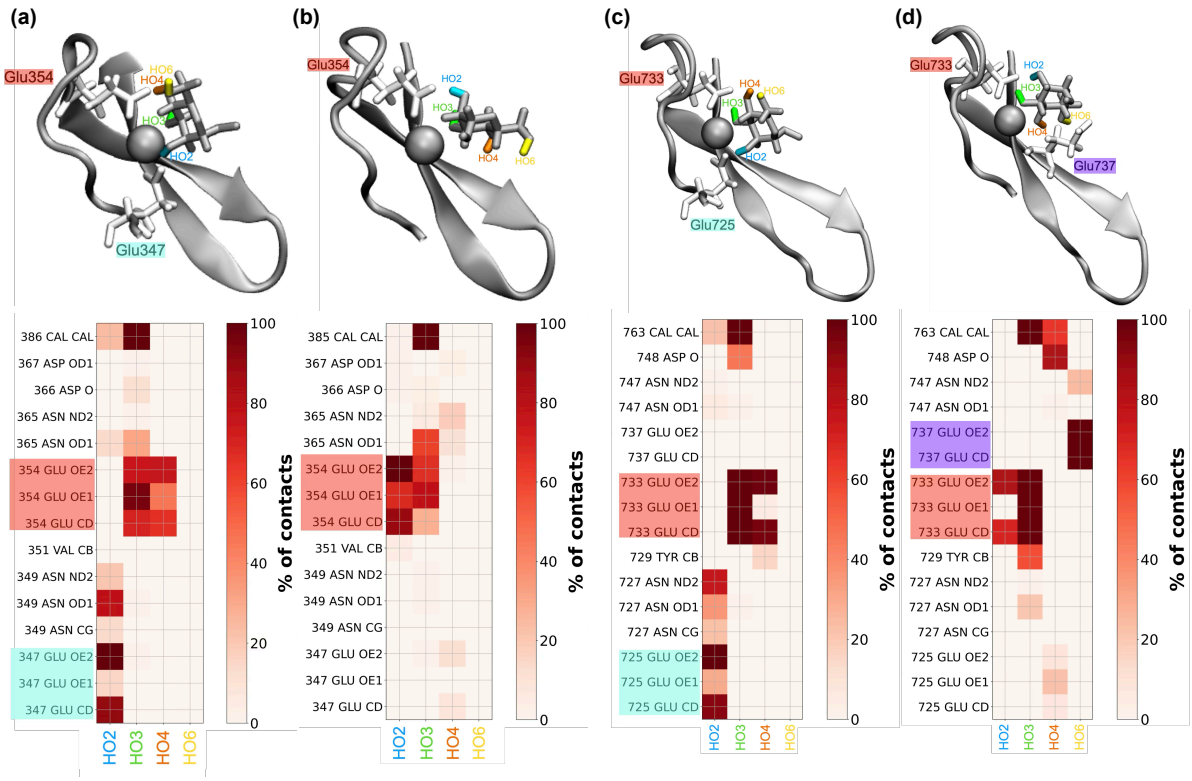


***Figure 7:*** *Mannose binding modes in the CRD glycan binding pockets with the associated contact maps.*

*DC-SIGN: (a) state A, (b) state B. MRC1: (c) state A, (d) state C*

OH3 and OH4 form contacts with the Ca2+ ion, while OH2 interacts with Glu733. Following a shift of the mannose moiety, Glu725 no longer interacts with the ligand. Instead, the OH6 group now forms hydrogen bonds with Glu737 (see Figure 7d). This binding state, which is specific to the MRC1 CRD, is the most stable one (see Figure 6b). In state C, the mannose loss of contacts with Glu347/725 is also compensated by the OH3 group forming contact with Tyr729, which is not possible in the DC-SIGN CRD, where a valine residue occupies the same position. State C was not described in the experimental literature before, but it seems to be responsible for the longer mannose binding times that we observed for the MRC1 CRD compared to the DC-SIGN system.

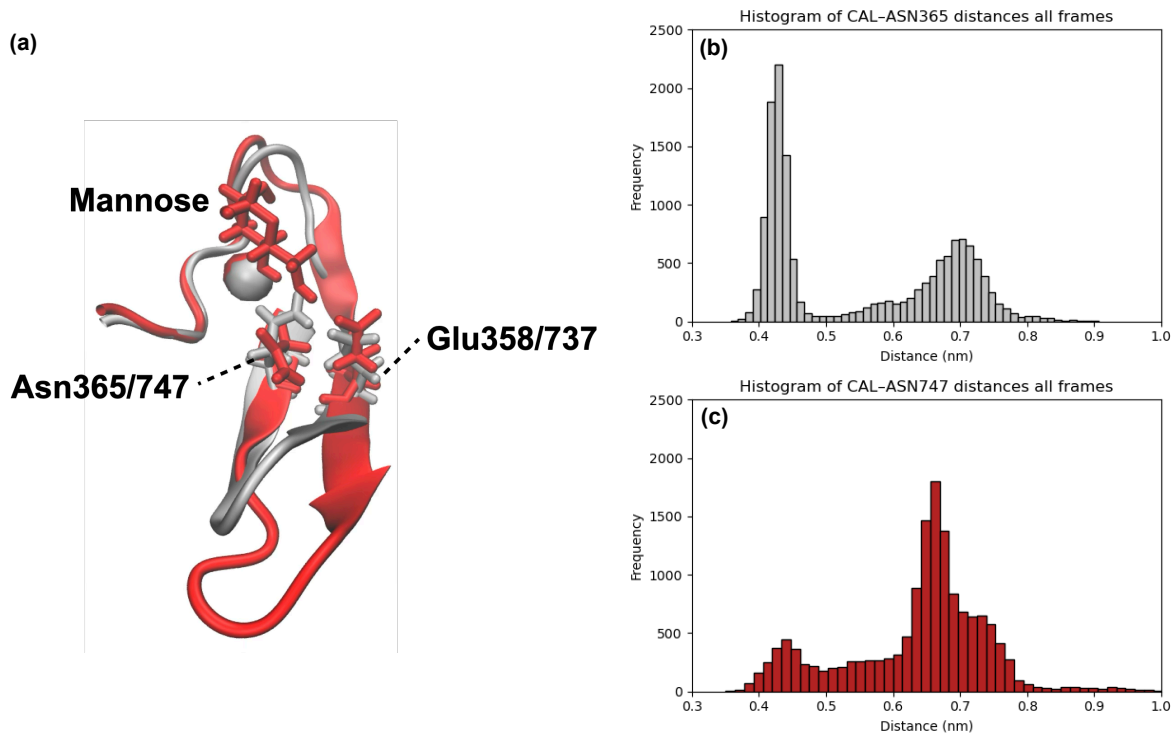


***Figure 8:*** *(a) Aligned structures of the glycan binding pockets for the DC-SIGN (in grey) and MRC1 (in red) CRDs with the mannose positioned in the C binding state. DC-SIGN Asn365 (in grey) is close to the $Ca^{2+}$ ion, while MRC1 Asn747 (in red) has rotated away.*
*(b) Distribution of the $Ca^{2+}$-Nδ2 distance for Asn365 in DC-SIGN for all trajectories (with and without bound mannose).*

Further exploration of the residues lining the glycan binding site show that, for the mannose group to bind in the state C orientation and form contacts with Glu358/737, Asn365/747 must rotate away from the $Ca^{2+}$ ion (see Figure 8a). This rotated conformation (termed *far state*), which is characterized by an increase in the $Ca^{2+}$/Asn-Nδ2 distance, appears to be more frequent in the MRC1 CRD than in DC-SIGN (see the distributions in Figures 8b and 8c). In particular, the distributions of the $Ca^{2+}$/Asn-Nδ2 when considering only structures with mannose bound on the CRDs (Figure S3a) show that Asn365 never occupies the far state in DC-SIGN, meaning that this rotation only occurs in DC-SIGN after mannose unbinding. In the simulation of the CRDs without ligand, Asn365/747 is initially in its close state (which correspond to the crystallographic structure) and stays so for the whole simulation in DC-SIGN, while it rapidly switches to the far state in MRC1 (see Figure S3b). We ran additional MD simulations of the CRDs without ligand (using the same parameters as in Table 1), but starting this time with Asn365/747 in its far state. For both CRDs, Asn365/747 stayed in the far state for the whole trajectory and never switched back to the close state (see Figure S3c).

### *Alternative binding sites on the CRDs surface*

Previous experimental studies using NMR spectroscopy reported secondary binding sites on C-type lectins,[31, 72] and even showed that for DC-SIGN, ligands binding these secondary sites might allosterically enhance glycan recognition by the main glycan binding pocket.[73] During our MD simulations, we also observed short-lived (lasting for 5-10 ns) rebinding events of the mannose-based ligands (but not for simple mannose) after leaving the glycan binding pocket (see Table S1 for a complete summary of the binding residues in the secondary binding sites). While earlier studies on alternative

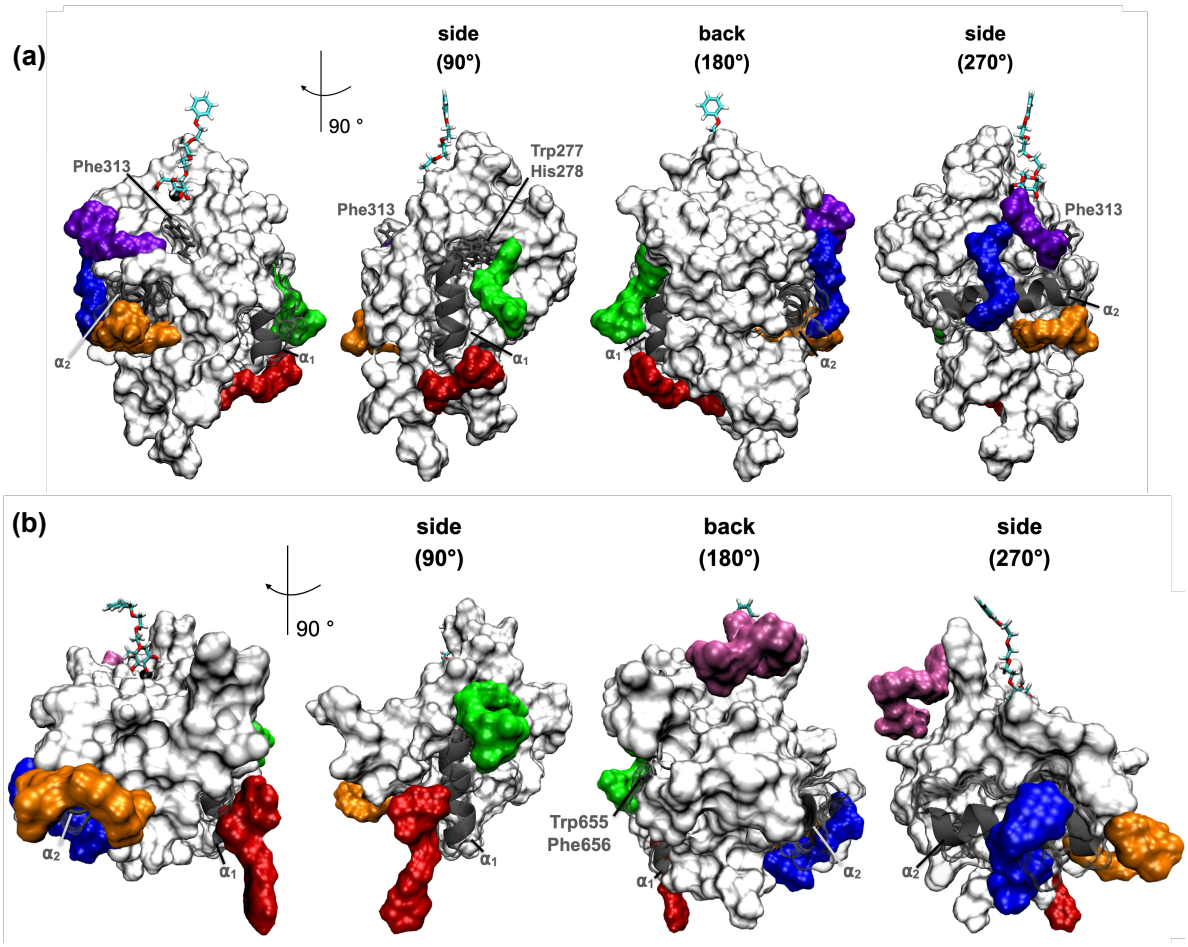


***Figure 9:*** *Alternative binding sites observed for both CRDs. The CRD is shown form foeur different angles in white surface representation with the ligands bound in alternative position in various colors. The 3n ligand bound to the canonical binding pocket is shown in licorice representation. Secondary elements and residues important for ligand binding are*

sites in DC-SIGN are mostly related to aromatic biphenyl substrates like the quinolone scaffold,[74] we could observe a similar phenomenon with only a simple linear linker and a benzene ring. For the simulations with the DC-SIGN CRD, the ligand bound on several occasions next to Phe313 (Figure 9a, in purple), which has been proposed in literature to be part of an extended canonical binding site (CBS) for large enough ligands.[29, 31] Two others secondary binding sites are located near the α2-helix of the lectin fold. One directly covers the helix (in blue on Figure 9a) and concurs with the site

II from ref. 72, while the second one (in orange on Figure 9a) lies below the $\alpha$2-helix. The green site on Figure 9a lies next to the $\alpha$1-helix, with the ligand forming contacts with Trp277 and His278, and can be related to site V from ref. 72. Finally, we also observed frequent interactions between the ligands and the $\alpha$1-helix lower part, on one side, and the protein's tow termini on the other side (in red in Figure 9a), but, as far as we know, these were not previously reported in NMR studies.

Simulations of the MRC1 CRD also led us to observe secondary ligand binding events. Like in DC-SIGN, the ligands bind next the $\alpha$1- (in green and red on Figure 9b) and $\alpha$2-helices (in blue and orange on Figure 9b). The green binding site around Trp655 ans Phe656 can be related to the one observed next to Trp277 and His278 in DC-SIGN, as both residue pairs occupy the same position in the CRDs sequences (see the alignment in Figure S4). Ligand binding behind the $\alpha$2-helix (in orange on Figure 9b) seems to be specific to the MRC1 CRD. Additionally, we could observe ligand binding below the CBS long loop (in mauve on Figure 9b). In contrast to DC-SIGN, there are no calcium atoms located in this pocket (see Figure 2b). The binding site at Phe313 of DC-SIGN is not matched in MRC1, as corresponding Phe695 residue in the MRC1 sequence (see Figure S4) points inwards, towards the core of the protein, thus not forming a convincing site of interaction for the ligands.

## 4. Conclusion

We used all-atom and coarse-grain simulations on the classical level to characterize the carbohydrate recognition domain of two mannose receptors, DC-SIGN and MRC1, and investigate the binding modes of various mannose based ligands on these CRDs. For

both receptors, the simulations highlight the low stability of the ligand in the glycan binding pockets and the reorientation of the mannose group during the MD trajectory, thus leading to three binding states that involve different residues on the CRDs surface. In particular, the binding of mannose in state C, which is specific to the MRC1 CRD, requires the rotation of an asparagine residue (Asn365/747). And this side-chain reorientation does not happen (or happens only after mannose unbinding) in the DC-SIGN CRD. The MD trajectories also show secondary ligand binding events on alternative binding sites on the CRDs surface. Calculations on a single CRD hint at a lower mannose affinity for DC-SIGN than for MRC1. However, both receptors present a more complex structure in the cell that might drastically change the results obtained when working on the CRD alone. Indeed, DC-SIGN assembles as a homotetramer[75] (see Figure 2b in ref. 71) and is thus able to simultaneously bind several mannose groups. Furthermore, recent experimental work by Porkolab et al.[76] highlight an unbinding/rebinding mechanism, where the ligand sequentially binds on the different CRDs, which is likely to enhance its affinity for DC-SIGN. In addition, some of the alternative binding sites that were identified on the surface of the DC-SIGN CRD monomer might no longer be accessible in the homotetramer. On the other hand, the MRC1 CRD is inserted in position4, within a linear chain of C-type lectin-like domains (CTLDs), see Figure 1 in ref. 44. The CRD4 domain is surrounded by CTLD3, which cannot bind mannose, and the C-type CRD5, which is missing the equivalent of Tyr729 and Glu737, thus making the possibility of mannose binding in state C very unlikely.

In this perspective, further work will address the multimeric structures of DC-SIGN and MRC1 by modeling the oligomeric assemblies of several CRDs for both receptors and

investigate how this global description impacts the mannose binding mechanism in both systems.

**Acknowledgments:** This work was supported by the ANR (SIMPA-ANR-23-CE29-0022) and by the "Initiative d'Excellence" program from the French State (Grant "DYNAMO", ANR-11-LABX-0011-01). Simulations were performed using the HPC resources from LBT/HPC and GENCI/TGCC (Grant A018071713)

**Supporting Information Available:** Additional data regarding the PCAs on the CRDs and the ligand alternative binding sites is available as a pdf file. All MD trajectories (without solvent) and their topologies for each simulation are deposited in Zenodo (https://zenodo.org/records/17550324).

## References


[1] Hudak, J. E., and Bertozzi, C. R. (2014) Glycotherapy: new advances inspire a reemergence of glycans in medicine, *Chem Biol 21*, 16-37.
[2] He, X., Agnihotri, G., and Liu Hw, H. W. (2000) Novel enzymatic mechanisms in carbohydrate metabolism, *Chem Rev 100*, 4615-4662.
[3] Varki, A. (2017) Biological roles of glycans, *Glycobiology 27*, 3-49.
[4] Cummings, R. D. (2019) "Stuck on sugars – how carbohydrates regulate cell adhesion, recognition, and signaling", *Glycoconjugate Journal 36*, 241-257.
[5] Sarkar, A., Drouillard, S., Rivet, A., and Perez, S. (2015) Databases of Conformations and NMR Structures of Glycan Determinants, *Glycobiology 25*, 1480-1490.
[6] Marchetti, R., Forgione, R. E., Fabregat, F. N., Di Carluccio, C., Molinaro, A., and Silipo, A. (2021) Solving the structural puzzle of bacterial glycome, *Current Opinion in Structural Biology 68*, 74-83.
[7] Mattox, D. E., and Bailey-Kellogg, C. (2021) Comprehensive analysis of lectin-glycan interactions reveals determinants of lectin specificity, *PLoS Comput Biol 17*, e1009470.
[8] Gheeraert, A., Bailly, T., Ren, Y., Hamraoui, A., Te, J., Vander Meersche, Y., Cretin, G., Leon Foun Lin, R., Gelly, J. C., Perez, S., Guyon, F., and Galochkina, T. (2025) DIONYSUS: a database of protein-carbohydrate interfaces, *Nucleic Acids Res. 53*, D387-D395.
[9] Gheeraert, A., Guyon, F., Perez, S., and Galochkina, T. (2025) Unraveling the diversity of protein-carbohydrate interfaces: Insights from a multi-scale study, *Carbohydr. Res. 550*, 109377.
[10] van Kooyk, Y., and Rabinovich, G. A. (2008) Protein-glycan interactions in the control of innate and adaptive immune responses, *Nat Immunol 9*, 593-601.
[11] Maverakis, E., Kim, K., Shimoda, M., Gershwin, M. E., Patel, F., Wilken, R., Raychaudhuri, S., Ruhaak, L. R., and Lebrilla, C. B. (2015) Glycans in the immune system and The Altered Glycan Theory of Autoimmunity: a critical review, *J Autoimmun 57*, 1-13.
[12] Pinho, S. S., and Reis, C. A. (2015) Glycosylation in cancer: mechanisms and clinical implications, *Nature reviews. Cancer 15*, 540-555.
[13] Nardy, A. F., Freire-de-Lima, L., Freire-de-Lima, C. G., and Morrot, A. (2016) The Sweet Side of Immune Evasion: Role of Glycans in the Mechanisms of Cancer Progression, *Frontiers in oncology 6*, 54.
[14] Brown, A., and Higgins, M. K. (2010) Carbohydrate binding molecules in malaria pathology, *Curr. Opin. Struct. Biol. 20*, 560-566.
[15] Mayer, S., Raulf, M. K., and Lepenies, B. (2017) C-type lectins: their network and roles in pathogen recognition and immunity, *Histochemistry and cell biology 147*, 223-237.
[16] Lin, B., Qing, X., Liao, J., and Zhuo, K. (2020) Role of Protein Glycosylation in Host-Pathogen Interaction, *Cells 9*.
[17] Sztain, T., Ahn, S. H., Bogetti, A. T., Casalino, L., Goldsmith, J. A., Seitz, E., McCool, R. S., Kearns, F. L., Acosta-Reyes, F., Maji, S., Mashayekhi, G., McCammon, J. A., Ourmazd, A., Frank, J., McLellan, J. S., Chong, L. T., and Amaro, R. E. (2021) A glycan gate controls opening of the SARS-CoV-2 spike protein, *Nature chemistry 13*, 963-968.
[18] Ernst, B., and Magnani, J. L. (2009) From carbohydrate leads to glycomimetic drugs, *Nat Rev Drug Discov 8*, 661-677.
[19] Gadekar, A., Bhowmick, S., and Pandit, A. (2020) A Glycotherapeutic Approach to Functionalize Biomaterials-Based Systems, *Advanced Functional Materials 30*, 1910031.
[20] Smith, B. A. H., and Bertozzi, C. R. (2021) The clinical impact of glycobiology: targeting selectins, Siglecs and mammalian glycans, *Nat Rev Drug Discov 20*, 217-243.
[21] Leusmann, S., Menova, P., Shanin, E., Titz, A., and Rademacher, C. (2023) Glycomimetics for the inhibition and modulation of lectins, *Chem. Soc. Rev. 52*, 3663-3740.
[22] Ballut, S., Makky, A., Chauvin, B., Michel, J. P., Kasselouri, A., Maillard, P., and Rosilio, V. (2012) Tumor targeting in photodynamic therapy. From glycoconjugated photosensitizers to glycodendrimeric one. Concept, design and properties, *Organic & Biomolecular Chemistry 10*, 4485-4495.
[23] Lupu, M., Maillard, P., Mispelter, J., Poyer, F., and Thomas, C. D. (2018) A glycoporphyrin story: from chemistry to PDT treatment of cancer mouse models, *Photochem Photobiol Sci 17*, 1599-1611.
[24] East, L., and Isacke, C. M. (2002) The mannose receptor family, *Biochim. Biophys. Acta 1572*, 364-386.
[25] Rojekar, S., Gholap, A. D., Togre, N., Bhoj, P., Haeck, C., Hatvate, N., Singh, N., Vitore, J., Dhoble, S., Kashid, S., and Patravale, V. (2024) Current status of mannose receptor-targeted drug delivery for improved anti-HIV therapy, *Journal of Controlled Release 372*, 494-521.
[26] Gallud, A., Warther, D., Maynadier, M., Sefta, M., Poyer, F., Thomas, C. D., Rouxel, C., Mongin, O., Blanchard-Desce, M., Morère, A., Raehm, L., Maillard, P., Durand, J. O., Garcia, M., and Gary-Bobo, M. (2015) Identification of MRC2 and CD209 receptors as targets for photodynamic therapy of retinoblastoma using mesoporous silica nanoparticles, *Rsc Adv 5*, 75167-75172.
[27] Wilt, S. D., Greaton, C. J., Lutz, D. A., and McLaughlin, B. J. (1999) Mannose receptor is expressed in normal and dystrophic retinal pigment epithelium, *Exp. Eye Res. 69*, 405-411.

[28] Zelensky, A. N., and Gready, J. E. (2005) The C-type lectin-like domain superfamily, *The FEBS journal 272*, 6179-6217.
[29] 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.
[30] Keller, B. G., and Rademacher, C. (2020) Allostery in C-type lectins, *Curr. Opin. Struct. Biol. 62*, 31-38.
[31] Lefebre, J., Falk, T., Ning, Y., and Rademacher, C. (2024) Secondary Sites of the C-type Lectin-Like Fold, *Chemistry 30*, e202400660.
[32] Lisacek, F., Schnider, B., and Imberty, A. (2025) Tools for structural lectinomics: From structures to lectomes, *BBA Adv 7*, 100154.
[33] Agirre, J. (2017) Strategies for carbohydrate model building, refinement and validation, *Acta Crystallogr D Struct Biol 73*, 171-186.
[34] Agirre, J., and van Raaij, M. J. (2018) Carbohydrate structure hits the groove, *Acta crystallographica. Section F, Structural biology communications 74*, 441-442.
[35] Quintana, J. I., Atxabal, U., Unione, L., Arda, A., and Jimenez-Barbero, J. (2023) Exploring multivalent carbohydrate-protein interactions by NMR, *Chem. Soc. Rev. 52*, 1591-1613.
[36] Atanasova, M., Bagdonas, H., and Agirre, J. (2020) Structural glycobiology in the age of electron cryo-microscopy, *Curr. Opin. Struct. Biol. 62*, 70-78.
[37] Murthy, B. N., Sinha, S., Surolia, A., Indi, S. S., and Jayaraman, N. (2008) SPR and ITC determination of the kinetics and the thermodynamics of bivalent versus monovalent sugar ligand-lectin interactions, *Glycoconj J 25*, 313-321.
[38] Casalino, L., Gaieb, Z., Goldsmith, J. A., Hjorth, C. K., Dommer, A. C., Harbison, A. M., Fogarty, C. A., Barros, E. P., Taylor, B. C., McLellan, J. S., Fadda, E., and Amaro, R. E. (2020) Beyond Shielding: The Roles of Glycans in the SARS-CoV-2 Spike Protein, *ACS Central Science 6*, 1722-1734.
[39] Perez, S., Fadda, E., and Makshakova, O. (2021) 1.15 - Computational Modeling in Glycoscience☆, In *Comprehensive Glycoscience (Second Edition)* (Barchi, J. J., Ed.), pp 374-404, Elsevier, Oxford.
[40] Perez, S., and Makshakova, O. (2022) Multifaceted Computational Modeling in Glycoscience, *Chemical Reviews 122*, 15914-15970.
[41] Nieto-Fabregat, F., Lenza, M. P., Marseglia, A., Di Carluccio, C., Molinaro, A., Silipo, A., and Marchetti, R. (2024) Computational toolbox for the analysis of protein-glycan interactions, *Beilstein J Org Chem 20*, 2084-2107.
[42] Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., and Zardecki, C. (2002) The Protein Data Bank, *Acta Crystallogr., sect. D: Biol. Crystallogr. 58*, 899-907.
[43] Thepaut, M., Guzzi, C., Sutkeviciute, I., Sattin, S., Ribeiro-Viana, R., Varga, N., Chabrol, E., Rojo, J., Bernardi, A., Angulo, J., Nieto, P. M., and Fieschi, F. (2013) Structure of a glycomimetic ligand in the carbohydrate recognition domain of C-type lectin DC-SIGN. Structural requirements for selectivity and ligand design, *J. Am. Chem. Soc. 135*, 2518-2529.
[44] Feinberg, H., Jegouzo, S. A. F., Lasanajak, Y., Smith, D. F., Drickamer, K., Weis, W. I., and Taylor, M. E. (2021) Structural analysis of carbohydrate binding by the macrophage mannose receptor CD206, *The Journal of biological chemistry 296*, 100368.
[45] Lee, J., Cheng, X., Swails, J. M., Yeom, M. S., Eastman, P. K., Lemkul, J. A., Wei, S., Buckner, J., Jeong, J. C., Qi, Y., Jo, S., Pande, V. S., Case, D. A., Brooks, C. L., 3rd, MacKerell, A. D., Jr., Klauda, J. B., and Im, W. (2016) CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field, *J Chem Theory Comput 12*, 405-413.
[46] Park, S.-J., Lee, J., Qi, Y., Kern, N. R., Lee, H. S., Jo, S., Joung, I., Joo, K., Lee, J., and Im, W. (2019) CHARMM-GUI Glycan Modeler for modeling and simulation of carbohydrates and glycoconjugates, *Glycobiology 29*, 320-331.
[47] Schrodinger, LLC. (2010) The PyMOL Molecular Graphics System, Version 1.3r1.
[48] Onufriev, A. V., and Izadi, S. (2018) Water models for biomolecular simulations, *WIREs Computational Molecular Science 8*, e1347.
[49] Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., and Lindahl, E. (2015) GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, *SoftwareX 1-2*, 19-25.
[50] Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B. L., Grubmuller, H., and MacKerell, A. D., Jr. (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins, *Nat Methods 14*, 71-73.
[51] Hess, B., Bekker, H., Berendsen, H. J. C., and Fraaije, J. (1997) LINCS: A linear constraint solver for molecular simulations, *J. Comp. Chem. 18*, 1463-1472.
[52] Berendsen, H. J. C., Postma, J. P. M., van Gunsteren, W. F., DiNola, A., and Haak, J. R. (1984) Molecular dynamics with coupling to an external bath, *J. Chem. Phys. 81*, 3684-3690.

[53] Parrinello, M., and Rahman, A. (1981) Polymorphic transitions in single crystals: A new molecular dynamics method, *J. Appl. Phys. 52*, 7182-7190.
[54] Evans, D. J., and Holian, B. L. (1985) The Nose–Hoover thermostat, *The Journal of Chemical Physics 83*, 4069-4074.
[55] Consortium, T. U. (2024) UniProt: the Universal Protein Knowledgebase in 2025, *Nucleic Acids Research 53*, D609-D617.
[56] Jurrus, E., Engel, D., Star, K., Monson, K., Brandi, J., Felberg, L. E., Brookes, D. H., Wilson, L., Chen, J., Liles, K., Chun, M., Li, P., Gohara, D. W., Dolinsky, T., Konecny, R., Koes, D. R., Nielsen, J. E., Head-Gordon, T., Geng, W., Krasny, R., Wei, G. W., Holst, M. J., McCammon, J. A., and Baker, N. A. (2018) Improvements to the APBS biomolecular solvation software suite, *Protein science : a publication of the Protein Society 27*, 112-128.
[57] Humphrey, W., Dalke, A., and Schulten, K. (1996) VMD: visual molecular dynamics, *J. Mol. Graph. 14*, 33-38, 27-38.
[58] Henin, J., Lopes, L. J. S., and Fiorin, G. (2022) Human Learning for Molecular Simulations: The Collective Variables Dashboard in VMD, *J Chem Theory Comput 18*, 1945-1956.
[59] Valdes-Tresanco, M. S., Valdes-Tresanco, M. E., Valiente, P. A., and Moreno, E. (2021) gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS, *J Chem Theory Comput 17*, 6281-6291.
[60] Gowers, J., Linke, Barnoud, Reddy, J. E., Melo, N., Seyler, L., Domański, Dotson, L., Buchoux, and Kenney, M. MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dyna mics Simulations, (Benthall, Ed.), pp 98 - 105.
[61] Michaud-Agrawal, N., Denning, E. J., Woolf, T. B., and Beckstein, O. (2011) MDAnalysis: A toolkit for the analysis of molecular dynamics simulations, *J. Comput. Chem. 32*, 2319-2327.
[62] Sacquin-Mora, S. (2014) Motions and mechanics: investigating conformational transitions in multi-domain proteins with coarse-grain simulations, *Mol. Simul. 40*, 229-236.
[63] Sacquin-Mora, S. (2016) Bridging Enzymatic Structure Function via Mechanics: A Coarse-Grain Approach, *Methods Enzymol. 578*, 227-248.
[64] Sacquin-Mora, S. (2018) Mechanical variations in proteins with large-scale motions highlight the formation of structural locks, *J Struct Biol 203*, 195-204.
[65] Tozzini, V. (2005) Coarse-grained models for proteins, *Curr. Opin. Struct. Biol. 15*, 144-150.
[66] Zacharias, M. (2003) Protein-protein docking with a reduced protein model accounting for side-chain flexibility, *Protein Sci. 12*, 1271-1282.
[67] Ermak, D. L., and McCammon, J. A. (1978) Brownian dynamics with hydrodynamic interactions, *J. Chem. Phys. 69*, 1352-1360.
[68] Pastor, R. W., Venable, R., and Karplus, M. (1988) Brownian dynamics simulation of a lipid chain in a membrane bilayer, *J. Chem. Phys. 89*, 1112-1127.
[69] Mishra, S. K., and Koča, J. (2018) Assessing the Performance of MM/PBSA, MM/GBSA, and QM–MM/GBSA Approaches on Protein/Carbohydrate Complexes: Effect of Implicit Solvent Models, QM Methods, and Entropic Contributions, *The Journal of Physical Chemistry B 122*, 8113-8121.
[70] Zlotnikov, I. D., and Kudryashova, E. V. (2022) Computer simulation of the Receptor–Ligand Interactions of Mannose Receptor CD206 in Comparison with the Lectin Concanavalin A Model, *Biochemistry (Moscow) 87*, 54-69.
[71] Cramer, J. (2021) Medicinal chemistry of the myeloid C-type lectin receptors Mincle, Langerin, and DC-SIGN, *RSC Med Chem 12*, 1985-2000.
[72] Aretz, J., Baukmann, H., Shanina, E., Hanske, J., Wawrzinek, R., Zapol'skii, V. A., Seeberger, P. H., Kaufmann, D. E., and Rademacher, C. (2017) Identification of Multiple Druggable Secondary Sites by Fragment Screening against DC-SIGN, *Angew Chem Int Ed Engl 56*, 7292-7296.
[73] Wawrzinek, R., Wamhoff, E. C., Lefebre, J., Rentzsch, M., Bachem, G., Domeniconi, G., Schulze, J., Fuchsberger, F. F., Zhang, H., Modenutti, C., Schnirch, L., Marti, M. A., Schwardt, O., Brautigam, M., Guberman, M., Hauck, D., Seeberger, P. H., Seitz, O., Titz, A., Ernst, B., and Rademacher, C. (2021) A Remote Secondary Binding Pocket Promotes Heteromultivalent Targeting of DC-SIGN, *J. Am. Chem. Soc. 143*, 18977-18988.
[74] Zhang, H., Daněk, O., Makarov, D., Rádl, S., Kim, D., Ledvinka, J., Vychodilová, K., Hlaváč, J., Lefèbre, J., Denis, M., Rademacher, C., and Ménová, P. (2022) Drug-like Inhibitors of DC-SIGN Based on a Quinolone Scaffold, *ACS Medicinal Chemistry Letters 13*, 935-942.
[75] Tabarani, G., Thepaut, M., Stroebel, D., Ebel, C., Vives, C., Vachette, P., Durand, D., and Fieschi, F. (2009) DC-SIGN neck domain is a pH-sensor controlling oligomerization: SAXS and hydrodynamic studies of extracellular domain, *The Journal of biological chemistry 284*, 21229-21240.
[76] Porkolab, V., Lepsik, M., Ordanini, S., St John, A., Le Roy, A., Thepaut, M., Paci, E., Ebel, C., Bernardi, A., and Fieschi, F. (2023) Powerful Avidity with a Limited Valency for Virus-Attachment Blockers on DC-SIGN: Combining Chelation and Statistical Rebinding with Structural Plasticity of the Receptor, *ACS Cent Sci 9*, 709-718.