Structural Basis for Selective Recognition of Endogenous and Microbial Polysaccharides by Macrophage Receptor SIGN-R1

Noella Silva-Martín,1 Sergio G. Bartual,1 Erney Ramirez-Aportela,3 Pablo Chacón,3 Chae Gyu Park,2,4,* and Juan A. Hermoso1,*

1Department of Crystallography and Structural Biology, Institute of Physical Chemistry “Rocasolano”, CSIC, 28006 Madrid, Spain
2Laboratory of Cellular Physiology and Immunology and Chris Browne Center for Immunology and Immune Diseases, The Rockefeller University, 1230 York Avenue, New York, NY 10065, USA
3Department of Biological Chemical Physics, Institute of Physical Chemistry “Rocasolano”, CSIC, 28006 Madrid, Spain
4Laboratory of Immunology, Severance Biomedical Science Institute, Brain Korea 21 PLUS Project for Medical Science, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 120-752, Korea
*Correspondence: chaegyu@yuhs.ac (C.G.P.), xjuan@iqfr.csic.es (J.A.H.)

http://dx.doi.org/10.1016/j.str.2014.09.001

SUMMARY

SIGN-R1 is a principal receptor for microbial polysaccharides uptake and is responsible for C3 fixation via an unusual complement activation pathway on splenic marginal zone macrophages. In these macrophages, SIGN-R1 is also involved in anti-inflammatory activity of intravenous immunoglobulin by direct interaction with sialylated Fcs. The high-resolution crystal structures of SIGN-R1 carbohydrate recognition domain and its complexes with dextran sulfate or sialic acid, and of the sialylated Fc antibody provide insights into SIGN-R1’s selective recognition of α,2,6-sialylated glycoproteins. Unexpectedly, an additional binding site has been found in the SIGN-R1 carbohydrate recognition domain, structurally separate from the calcium-dependent carbohydrate-binding site. This secondary binding site could bind repetitive molecular patterns, as observed in microbial polysaccharides, in a calcium-independent manner. These two binding sites may allow SIGN-R1 to simultaneously bind both immune glycoproteins and microbial polysaccharide components, accommodating SIGN-R1’s ability to relate the recognition of microbes to the activation of the classical complement pathway.

INTRODUCTION

The classical pathway is one of the major complement activation mechanisms and is required for innate protection against pathogenic microorganisms including Streptococcus pneumoniae (Brown et al., 2002). A pivotal step in the complement activation pathway is the assembly of C3 convertase, which digests the complement components to form C3 fragments that bind to microbes and make them recognized by leukocytes. Deposition and activation of C3 convertase over the bacterial surface initiates the complement cascade leading to the elimination of the microbes. However, encapsulated microorganisms like S. pneumoniae are able to resist to the complement pathway by limiting access to the cell surface and reducing the amount of C3 deposition (Abeyta et al., 2003). To overcome this problem, hosts have developed an alternative strategy for specific pathogen detection by pattern recognition receptors (PRRs). PRRs discriminate the molecular patterns expressed by pathogens and thus facilitate differential recognition of pathogens and their products by the immune system (Gordon, 2002).

C-type lectin receptors (CLRs) are one kind of PRRs expressed by a broad spectrum of cells, including dendritic cells (DCs), macrophages, and other antigen-presenting cells (Meyer-Wentrup et al., 2005). Increasing evidence has shown that the roles played by several members of the CLR family are critical for innate immunity. SIGN-R1 or CD209b is one of the eight mouse homologs of human CLR named DC-SIGN or CD209 (Park et al., 2001; Powlesland et al., 2006). Unlike human DC-SIGN/CD209, three of mouse SIGN family molecules are found to be expressed in the immune system: DC-SIGN/CD209a, SIGN-R1/CD209b, and SIGN-R3/CD209d. Whereas human DC-SIGN is expressed in both dendritic cells and macrophages (Geijtenbeek et al., 2000; Granelli-Piperno et al., 2005), the expression of three mouse SIGN homologs is detected in different levels from distinct populations of dendritic cells and macrophages (Cheong et al., 2010; Kang et al., 2003; Nagaoka et al., 2010). Of three homologs, mouse DC-SIGN lacks polysaccharide binding, whereas SIGN-R1, similar to human DC-SIGN, binds a broad variety of microbial organisms and their polysaccharides and SIGN-R3 binds a fewer number of ligands (Caminschi et al., 2006; Galustian et al., 2004; Takahara et al., 2004).

SIGN-R1 is expressed at high levels in macrophages within the splenic marginal zone and lymph node medulla and is also a principal receptor for taking up microbial polysaccharides such as dextran or capsular polysaccharides of S. pneumoniae (CPSs; Geijtenbeek et al., 2002; Kang et al., 2003, 2004). In vivo inactivation of SIGN-R1 inhibits clearance of S. pneumoniae by preventing macrophages from phagocytosis, blocks complement-dependent deposition of CPS on splenic follicles, and therefore makes mice more susceptible to pneumococcal
infection (Geijtenbeek et al., 2002; Kang et al., 2003, 2004, 2006; Lanoue et al., 2004). This defect in innate immunity against S. pneumoniae after inactivation of SIGN-R1 suggests a link between SIGN-R1 and the fixation of C3 opsonins on CPS. This connection was revealed by an unusual complement activation pathway against CPS mediated by SIGN-R1 on splenic marginal zone macrophages (Kang et al., 2006). Binding of CPS to SIGN-R1 initiates an immunoglobulin (Ig)-independent activation of complements for the fixation of C3 on CPS, where the direct interaction between SIGN-R1 and C1q triggers the classical complement pathway. This complement activation strategy in splenic marginal zone macrophages prevents S. pneumoniae from escaping C3 opsonization (Kang et al., 2006). In this mechanism, it is critical that SIGN-R1 molecule can function in both processes, i.e., the recognition and binding of a pathogen and the activation of classical complement pathway, which is likely to require SIGN-R1 to bind both CPS and C1q simultaneously.

In addition, it has been reported that SIGN-R1 clears out the cells via interaction with C1q fixed on apoptotic cells, thus increasing the C3 fixation (i.e., complement activation) on the apoptotic cells, and reducing autoimmunity (Prabagar et al., 2013).

The cause of anti-inflammatory activity of intravenous Ig (IVIG) was contributed to a minor population of the pooled IgG molecules that contains terminal α-2,6-sialic acid linkages on their Fc-linked glycans (Kaneko et al., 2006). It was reported that SIGN-R1 in splenic marginal zone macrophages preferentially binds to the sialylated Fcs and that in vivo inactivation of SIGN-R1 eliminates the anti-inflammatory activity of IVIG and intravenous sialylated Fcs (Anthony et al., 2008). This interaction between SIGN-R1 and sialylated Fcs was shown to accomplish the anti-inflammatory activity via a different Th2 cytokine pathway (Anthony et al., 2011). Therefore, SIGN-R1 is a single member of the mouse SIGN CLR family that functions by binding polysaccharides from both endogenous and microbial origins.

We report herein the high-resolution crystal structures of the carbohydrate recognition domain of SIGN-R1 (CRD_SIGN-R1) as well as their complexes with dextran sulfate (DexS) and sialic acid (N-acetyllneuraminic acid [NeuAc]; Figure 1). In addition, the high-resolution crystal structure of the sialylated Fc antibody portion obtained in this work together with molecular dynamics calculations, provides insights into SIGN-R1’s selective recognition of α-2,6-sialylated glycoproteins relevant to the role of SIGN-R1 in anti-inflammatory activity of IVIG. Structural, computational and biochemical experiments reveal that sialylated glycoproteins, such as C1q and Ig, bind SIGN-R1 through the classical calcium-dependent carbohydrate recognition site, whereas the binding of microbial pathogens can also use the additional calcium-independent substrate-binding site, thereby allowing SIGN-R1 to simultaneously bind both immune glycoproteins and microbial components.

RESULTS AND DISCUSSION

Overall CRD_SIGN-R1 Structure
With the extracellular domain of SIGN-R1 (ECD_SIGN-R1) produced as monomers from Escherichia coli, crystals could not be generated (P. Li, personal communication). In addition, with the ECD_SIGN-R1 produced as His-tagged monomers in E. coli, carbohydrate binding could not be detected even after oligomerization with anti-His antibody (T. Feizi, personal communication). Powlesland and colleagues also reported that the ECD_SIGN-R1 produced from E. coli was unable to bind any carbohydrates as monomers (Powlesland et al., 2006). Therefore, it is likely that the ECD_SIGN-R1 or CRD_SIGN-R1 proteins prepared from E. coli are not properly folded. Previously, however, we had successfully performed the carbohydrate binding analyses of SIGN-R1, where the soluble Fc-fused ECD_SIGN-R1 was expressed and purified directly from the culture supernatant of transfectant Chinese hamster ovary (CHO) cells (Galustian et al., 2004). Therefore, we prepared the soluble CRD_SIGN-R1 (amino acid residues from Cys192 to Gly325) from the culture supernatant of transfectant CHO cells and made its crystals successfully (Silva-Martin et al., 2009).

High-resolution crystal structures for CRD_SIGN-R1 and its complexes with dextran sulfate (SIGN-R1:DexS) and sialic acid (SIGN-R1:NeuAc) have been solved in this study (Table 1). The asymmetric unit is composed of four nearly identical monomers with an average root-mean-square deviation (rmsd)...
value of 0.182 Å over 126 aligned Cα atoms. The CRD_SIGN-R1 structure (Figure 2A) exhibits one large β strand (β2), which divides the framework in two lobes. The upper lobe comprises β strands 2 to 4 and the α3 helix. The lectin characteristic long loop region (LLR) involved in calcium-mediated carbohydrate binding is found between α3 and β3 (Tyr280 to Asp293; Figure 1). The lower lobe includes the N-terminal and the C-terminal extremes, coming close together to form an antiparallel β sheet that is flanked by α1 and α2, one on each side. Eight cysteines are involved in the formation of four disulfide bridges, three of them stabilizing the lower lobe, while the fourth is located between both lobes. The “WMGL” motif present in all C-type lectin proteins reported (Zelensky and Gready, 2003) is also present in SIGN-R1 (residues 253–256). This pattern unites a cluster of hydrophobic residues stabilizing all parts of the structure. The human C-type lectin DC-SIGN (Protein Data Bank [PDB] code 2IT6) is the most closely related structure to SIGN-R1, with 69% of sequence identity and an rmsd of 1.32 Å over 131 aligned C-alpha atoms.

SIGN-R1 presents six sulfate molecules bound to each monomer (Figure 2A); S1 is near the calcium+-mediated carbohydrate binding site, S6 is placed in the lower lobe near the N-terminal and the C-terminal extremes, and four sulfate molecules (S2–S5) are placed at the opposite face of the calcium-mediated carbohydrate binding site (Figure 2A). Interestingly these four sulfates are located inside a positively charged groove in collar pattern disposition, retaining a similar distance among them (8.70 Å, 9.94 Å, and 8.46 Å; Figure 2B).

Calcium-Binding Site in SIGN-R1
In general, altogether, four Ca2+–binding sites occur in different subgroups of the C-type lectin-like domain (CTLD) family (Zelensky and Gready, 2003). In the crystal structure of SIGN-R1 only one, that of the characteristic C-type lectin sugar-binding site, is observed (Figure 2A). This Ca2+ is stabilized by carboxylate or carboxamide groups of Glu285, Asn287, Glu292, and Asp304 and by two water molecules, providing an octahedral coordination for the cation (Figure S1A available online). All these interactions are critical for the stability of the structure. SIGN-R1 presents six sulfate molecules bound to each monomer (Figure 2A); S1 is near the calcium+-mediated carbohydrate binding site, S6 is placed in the lower lobe near the N-terminal and the C-terminal extremes, and four sulfate molecules (S2–S5) are placed at the opposite face of the calcium-mediated carbohydrate binding site (Figure 2A). Interestingly these four sulfates are located inside a positively charged groove in collar pattern disposition, retaining a similar distance among them (8.70 Å, 9.94 Å, and 8.46 Å; Figure 2B).

**Table 1. Data Collection and Refinement Statistics**

<table>
<thead>
<tr>
<th></th>
<th>CRD-SIGN-R1</th>
<th>CRD-SIGN-R1:DexS</th>
<th>CRD-SIGN-R1:NeuAc</th>
<th>FC_Sial</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Space group</td>
<td>C2</td>
<td>C2</td>
<td>C2</td>
<td>P2;2;2;1</td>
</tr>
<tr>
<td>Cell dimensions a, b, c (Å)</td>
<td>146.72, 92.77, 77.06</td>
<td>144.03, 97.96, 73.84</td>
<td>146.20, 94.47, 76.14</td>
<td>49.59, 80.15, 140.21</td>
</tr>
<tr>
<td>a, β, γ (°)</td>
<td>90, 121.66, 90</td>
<td>90, 120.86, 90</td>
<td>90, 121.42, 90</td>
<td>90, 90, 90</td>
</tr>
<tr>
<td>Temperature (K)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Wavelength (Å)</td>
<td>1.07225</td>
<td>0.9334</td>
<td>0.87260</td>
<td>0.999</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>74.53–(1.98–1.87)</td>
<td>61.815–(2.98–2.58)</td>
<td>42.98–(2.95–2.80)</td>
<td>42.17–(2.42–2.30)</td>
</tr>
<tr>
<td>No. unique reflections</td>
<td>52,474 (2,478)</td>
<td>30,140 (2,478)</td>
<td>21,870 (3,185)</td>
<td>25,633 (3,662)</td>
</tr>
<tr>
<td>R_sym ^a</td>
<td>0.06 (0.54)</td>
<td>0.11 (0.28)</td>
<td>0.11 (0.45)</td>
<td>0.09 (0.51)</td>
</tr>
<tr>
<td>Average I/σ(I)</td>
<td>11.80 (1.70)</td>
<td>6.2 (1.8)</td>
<td>9.4 (3.7)</td>
<td>11.7 (3.5)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>93 (72)</td>
<td>99 (99)</td>
<td>100 (100)</td>
<td>99.9 (100)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>5.7 (5.4)</td>
<td>4.4 (2.4)</td>
<td>2.3 (2.3)</td>
<td>6.4 (6.6)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>27.63–1.87</td>
<td>40.00–2.6</td>
<td>40.31–2.80</td>
<td>20.10–2.03</td>
</tr>
<tr>
<td>R/R_free (%)</td>
<td>20.45/24.43</td>
<td>20.05/24.74</td>
<td>18.06/26.21</td>
<td>23.64/28.21</td>
</tr>
<tr>
<td>No. atoms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein</td>
<td>4,292</td>
<td>4,315</td>
<td>4,295</td>
<td>3,316</td>
</tr>
<tr>
<td>Water</td>
<td>407</td>
<td>168</td>
<td>479</td>
<td>82</td>
</tr>
<tr>
<td>Ligand</td>
<td>–</td>
<td>147</td>
<td>63</td>
<td>221</td>
</tr>
<tr>
<td>Metal ions (Ca)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>–</td>
</tr>
<tr>
<td>SO₄</td>
<td>24</td>
<td>12</td>
<td>30</td>
<td>–</td>
</tr>
<tr>
<td>B-factor (Å²)</td>
<td>34.29</td>
<td>30.74</td>
<td>26.55</td>
<td>42.82</td>
</tr>
<tr>
<td>Protein</td>
<td>407</td>
<td>168</td>
<td>479</td>
<td>82</td>
</tr>
<tr>
<td>Ligand</td>
<td>–</td>
<td>63.45</td>
<td>59.69</td>
<td>61.15</td>
</tr>
<tr>
<td>Water</td>
<td>24.88</td>
<td>23.70</td>
<td>24.67</td>
<td>37.63</td>
</tr>
<tr>
<td>Rmsds</td>
<td>0.012</td>
<td>0.011</td>
<td>0.018</td>
<td>0.011</td>
</tr>
<tr>
<td>Bond angles (°)</td>
<td>1.063</td>
<td>1.257</td>
<td>1.141</td>
<td>1.257</td>
</tr>
<tr>
<td>PDB code</td>
<td>3ZHG</td>
<td>4C9F</td>
<td>4CAJ</td>
<td>4CDH</td>
</tr>
</tbody>
</table>

Value for the highest resolution shell is shown in parentheses.

^aR_sym = Σ|I – Iav|/ΣI, where the summation is over symmetry equivalent reflections.
residues are part of two characteristic motifs present in most of the CTLD superfamily sequences: the “EPN” motif (residues 285–287) that belongs to the LLR and the “WND” motif (residues 302–304) from the β4 strand (Figure S1C). Interestingly, the Asn303 in SIGN-R1, which is part of the “WND” motif, is not coordinating the calcium ion (Figure S1A) as occurs in the other known structures of sugar-binding C-type lectins either alone or in complex with a ligand (Figure S1B). Structural analysis also reveals that the LLR region from SIGN-R1 presents a different conformation among all reported CTLD members (Figure 2C). The cavity in which secondary calcium ions (i.e., Ca2 and Ca3) are located in human DC-SIGN is occupied by the LLR loop in SIGN-R1, therefore avoiding interaction with these two calcium atoms (Figure S1D). The L2 loop has been observed in two different conformations (closed and open) in DC-SIGNR, depending on the presence or absence of the secondary calcium ions. In SIGN-R1, the L2 loop is folded in the closed conformation (Figure 2C) without the presence of secondary Ca ions. Two residues, Ile289 from LLR and Ala263 from L2 (neither present in DC-SIGN nor in DC-SIGNR), stack with hydrophobic residues of the cavity masking the secondary calcium-binding cavity (Figure S1D). The presence of a salt bridge interaction between Glu291 and Lys306 further stabilizes the conformation of the LLR loop (Figure S1D). It is worth mentioning that crystallization with an excess of Ca2+ did not result in any modification of the 3D structures of SIGN-R1 either in the native or in the complexes with substrates. Mechanistic implications of these differences in residues and loop conformations in SIGN-R1 and DC-SIGN/DC-SIGNR cannot be overstated. Mutagenesis experiments on residues in DC-SIGN (Val351) and in DC-SIGNR (Ser363), equivalent to Ile289 of SIGN-R1, revealed that they are the main responsible residues for binding specificity of DC-SIGN to fucose and DC-SIGN-R to mannose, respectively (Guo et al., 2004).

Carbohydrate Recognition through Canonical Site
Carbohydrates that bind to SIGN-R1 have been previously explored with glycan-array tests, suggesting the specificity of SIGN-R1 for mannose, fucose, and DexS, among other polysaccharides (Galustian et al., 2004). Besides, SIGN-R1 has been reported to bind sialic acid present on antibodies or glycoproteins such as C1q complement factor (Anthony et al., 2008; Mizuochi et al., 1978).

In the SIGN-R1:DexS complex, three of the four monomers of the asymmetric unit presented electron density for DexS moieties (one tetrasaccharide unit and two disaccharide units...
respectively, see Figure 1B) in the Ca²⁺-mediated primary binding site of the lectin (Figures 2D and 2E; Figure S2). The Ca²⁺-mediated interactions with dextran depend on the presence or absence of a sulfate molecule attached to the glucose sugar. When the sugar contains a sulfate molecule (as occurs in the tetrasaccharide and one of the disaccharides; Figure 2D), the sugar ring interacts with calcium through the O1 and establishes hydrogen bonds between its OH-1 with Asn287 and Asn288, the OH-2 with Glu285, and the OH-3 with Asn303. When the sugar does not possess this sulfate moiety, interaction with Ca²⁺ is made through the OH-4, whereas OH-3 interacts with residues Asn287 and Asn288 (Figure 2E). It is worth noting that due to SIGN-R1’s conformation of LLR, Asn288 plays a key role in carbohydrate binding but not in binding of one of the secondary Ca²⁺ as occurs in human DC-SIGN(R). Protein interactions with the second sugar ring are conserved in all three DexS complexes (Figures 2D and 2E). The tetrasaccharide extends its binding through rings three and four by hydrophobic interactions with Pro251 and Phe312. The fourth ring also establishes hydrogen bonds with Gln213 (Figure 2D).

In the crystal structure of the SIGN-R1:NeuAc complex, all four monomers of the asymmetric unit present the sialic acid attached (Figures 2F and S2). The carboxyl group of the sialic acid replaces the interaction of the two water molecules of native SIGN-R1 in the octahedral coordination of the calcium ion. Sialic acid is strongly stabilized by SIGN-R1 because most of residues involved in calcium binding (Glu285, Asn287, Glu292) also interact with sialic acid, yielding a large hydrogen bond interaction network (Figure 2F). Interestingly, as also observed in SIGN-R1:DexS complexes, the Asn288 contributes to this network interaction in SIGN-R1.

A Secondary Binding Site in SIGN-R1

Analysis of the density maps of SIGN-R1:DexS complex revealed the presence of three well-defined molecules of D-glucose-4-sulfate (Glc-S) at the opposite face of the canonical binding site (Figure 3B). This secondary site is a 30-Å long basic groove that is formed by loops connecting x1 and x2, x3 and l2, and by the lower lobe β sheet. Glc-S molecules occupy the positions of the sulfate molecules S3–S5 found in the native structure (Figure 2). Remarkably, both the sulfate moiety and the glycan part of Glc-S are well stabilized in the SIGN-R1:DexS complex (Figures 3B and S2E).

The repetitive disposition of sulfate/sugar molecules recognized by SIGN-R1 through this secondary binding site pushed us to explore SIGN-R1 binding to other cellular components sharing a similar repetitive structure. Teichoic acids (TAs), one of the main components of the cell wall of all gram-positive bacteria, are formed by repeating units of sugar-phosphate chains (Weidenmaier and Peschel, 2008). These glycopolymers are frequently formed by glycerol or ribitol groups linked by phosphodiester bridges as in TAs from Staphylococcus aureus or Listeria monocytogenes (Weidenmaier and Peschel, 2008). Docking calculations reveal that polyribitol phosphate chain of TAs nicely fits into SIGN-R1’s secondary binding site with phosphate moieties superimposing onto the S3–S5 molecules found in SIGN-R1 (Figure 3C). Another repetitive structure is that of capsular pneumococcal polysaccharide 14 (CPS14; Figure 1). It has been found that SIGN-R1 binds preferentially to this capsular polysaccharide (Kang et al., 2004). Docking calculations also show binding of CPS14 to this region of SIGN-R1 (Figure 3D). Docked sugar rings replace the crystallographic glucose molecules found in the SIGN-R1:DexS complex.

The additional secondary binding site in SIGN-R1 is calcium independent and structurally not related to the canonical
binding site. Whereas the binding of SIGN-R1 to microbial polysaccharides is calcium dependent (Galustian et al., 2004; Nagaoka et al., 2005; Takahara et al., 2004); i.e., EDTA at 5 mM or higher concentration prevents SIGN-R1 from binding to its polysaccharide ligands. This implies that, although the primary and secondary binding sites are spatially separate, polysaccharides might use both sites to bind SIGN-R1. In agreement with this, it has been reported that SIGN-R1 can only bind to dextrans with high molecular weights but other SIGNs can bind to dextrans with high as well as lower molecular weights (Takahara et al., 2004). To test this model, we compared the polysaccharide-binding activities of SIGN-R1 with those of human DC-SIGN (hDC-SIGN), which lacks a calcium-independent secondary binding site. In both cases of FITC-labeled dextran (500 kDa) and PE-labeled CPS14, the CHO cells expressing SIGN-R1 were able to bind these polysaccharides at 0.25–0.5 mM higher concentration of EDTA than the CHO cells expressing hDC-SIGN (Figure 4). This may indicate that SIGN-R1 contains an extra calcium-independent binding site for dextran and CPS14, making
Basis for Selective Recognition by SIGN-R1

**Structure**

Implications in Anti-inflammatory Activity of Intravenous Ig

SIGN-R1 plays an essential role in the anti-inflammatory process induced by IVIG (Anthony et al., 2008). In this process, SIGN-R1 specifically recognizes α,2,6 sialylated antibodies instead of α,2,3 sialylated antibodies. According to our study on the SIGN-R1:NeuAc complex, SIGN-R1 interacts with sialic acid through the carbohydrate moiety. The atomic resolution of an α,2,3 sialylated antibody (PDB code 1MCO; Guddat et al., 1993) shows the carboxylic group of sialic acid occluded and facing the inner part of the antibody, therefore it is not able to interact with SIGN-R1 as found in the SIGN-R1:NeuAc complex.

To shed more light on the interaction between SIGN-R1 and sialylated Fc, we solved the high-resolution structure of a human α,2,6 sialylated antibody Fc portion by X-ray crystallography (Table 1; Figure S3A). The density maps of the α,2,6 Fc structure were of excellent quality and allow us to trace almost the complete glycan chain but without the terminal sialic acid that was not visible in our maps (Figure S3B), indicating a high flexibility of this terminal sugar. The crystal structure of an α,2,6 sialylated antibody has been reported (PDB code 4BYH; Crispin et al., 2013). Despite completely different crystallization conditions, our α,2,6 Fc structure was nearly identical to the reported α,2,6 sialylated antibody in both the backbone (rmsd of 0.563 Å over 294 aligned Ca atoms) and the sugar conformation (Figures S3C and S3D). Interestingly, the α,2,6 Fc structure shows the carboxylic group of arm 6 terminal sialic acid as completely exposed (Figures S3D and S3E), allowing the interaction with SIGN-R1 as observed in the SIGN-R1:NeuAc complex.

Molecular dynamic (MD) simulations further indicate that α,2,3 and α,2,6 sialylations result in two different conformers for the sialic acid. In α,2,3 sialylated antibody both galactose and sialic acid rings establish strong interactions with the protein (Figures S4 and S5; Movie S1), provoking the carboxylate moiety being occluded (accessible surface of 0.3 Å²). In the α,2,6 sialylated antibody, the galactose interacts with the protein while the sialic acid is completely exposed (accessible surface of 5 Å²), releasing some regions of Fc antibody (Figures S5 and S6; Movie S2). This high flexibility is in agreement with what we observe in our crystal structures.

Summing up, whereas the α,2,6 link provokes exposition of carboxylic acid of the sialic acid, the α,2,3 variant does not. Therefore only the α,2,6 link could allow Fc recognition by SIGN-R1. Superimposition of both SIGN-R1 and Fc structural models (Figure 5) indicates that their interaction mediated by sialic acid carboxylate would be the starting anchor for further interaction with other sugars from the carbohydrate moiety linked to Fc. A complementarity exists between the protein surfaces of the SIGN-R1 receptor and the antibody, pointing to a protein-protein interaction upon glycan recognition.

Figure 5. Model of Interaction between SIGN-R1 and α,2,6 Sialylated Antibody

Superimposition of the crystal structure of SIGN-R1:NeuAc (blue ribbon) with the crystal structure of α,2,6 sialylated antibody (molecular surface in brown for chain A and in green for chain B). Canonical binding site of SIGN-R1 is represented with the Ca ion (green sphere) and residues coordinating it (blue capped sticks). The LLR loop is labeled. The secondary binding site of SIGN-R1 is located at the opposite face where sulfate molecules (represented as orange-capped sticks) are found. Oligosaccharides attached to Fc are represented as capped sticks. The terminal α,2,6 sialic acid was not visible in the crystal structure; the coordinates for this terminal sugar have been taken from the energy-minimized model. SIGN-R1:Fc interaction mediated by sialic acid carboxylate would be the starting anchor for further interaction with other sugars from the carbohydrate moiety linked to Fc. A complementarity exists between the protein surfaces of the SIGN-R1 receptor and the antibody, pointing to a protein-protein interaction upon glycan recognition.

Complement Fixation Pathway for Pathogens Initiated by SIGN-R1

SIGN-R1 presents a single calcium-binding site that corresponds to the primary carbohydrate-binding site observed in other CTLDs. The specific conformation of SIGN-R1 for the LLR loop avoids stabilization of Ca2 and Ca3.

Crystal structures of SIGN-R1 alone and in complex with DexS reveal the presence of sulfate or Glc-S moieties, respectively, both of which are attached to SIGN-R1’s calcium-independent secondary binding site in a repetitive manner. This additional binding site is located at the opposite face of the canonical calcium-dependent binding site without apparent structural connection among them. Docking calculations indicate that polyribitol phosphate (backbone of teichoic acids of some microbial pathogens), which follows the same repetitive pattern, can be accommodated in this secondary binding site. The same result is also observed for docking with a model of the capsular...
poly saccharides from pneumococcus serotype 14 (CSP14). Altogether, these results indicate that SIGN-R1 possesses two separate binding sites. The additional noncanonical binding site in SIGN-R1 might recognize distinct carbohydrate patterns from microbial pathogens, such as dextran, teichoic acids, and CPS14.

SIGN-R1 mediates activation of the classical complement pathway through direct interaction with C1q and CPS in spleen (Kang et al., 2006). How exactly these molecules interact to activate the complement pathway is unknown. C1q is a glycoprotein and the structure of its recognition domain (Gaboriaud et al., 2003) and sugar chains has been reported (Mizuochi et al., 1978). Structures of oligosaccharides attached to C1q are represented in Figure 1A (or small variants with only one NeuAc at the end or without the Fuc at the beginning of the chain). In all cases, the sialic acid is exclusively linked at the C-6 position of the galactose residue. Furthermore, it has been reported that SIGN-R1 specifically recognizes IgG molecules that contain terminal α-2,6 sialic acid linkages (Anthony et al., 2008). The crystal structure of the SIGN-R1:NeuAc complex reveals that sialic acid is strongly stabilized by SIGN-R1 at the primary canonical carbohydrate binding site (but not in the secondary site) through carboxylic acid in NeuAc. Considering all these facts, we propose a model that the SIGN-R1 lectin captures C1q, which is bound to C1r and C1s in serum, via a direct interaction between the primary carbohydrate binding site in SIGN-R1 and the polysaccharides ended by an α-2,6 sialic acid in C1q, whereas SIGN-R1 also binds to pathogen patterns through its secondary binding site (Figure 6).

However, seeing that the binding of microbial polysaccharides to SIGN-R1 expressed on cell lines is less but still sensitive to EDTA treatment (Figure 4), this model for simultaneous recognition of two different polysaccharides by SIGN-R1 (Figure 6) may not explain all the binding events mediated by SIGN-R1 on splenic marginal zone macrophages. SIGN-R1 protein in spleen and lymph nodes and expressed in cell lines exists in highly complex forms that have much higher molecular weights than tetrameric forms (Kang et al., 2003). In addition, cell lines expressing SIGN-R1 cannot bind ≤40 kDa dextrans, whereas other CLRs such as SIGN-R3 and Langerin can bind 40 kDa of dextran (Takahara et al., 2004). Based on these observations, we propose that on the surface of splenic marginal zone macrophages, where SIGN-R1 is expressed as large complexes of multiple tetramers, microbial polysaccharides such as dextran and CPS14 bind to SIGN-R1 on both of its calcium-dependent and -independent sites. Meanwhile, Ca-dependent primary binding sites in a significant number of SIGN-R1 molecules in the same complexes might bind to the terminal α-2,6 sialic acid of polysaccharides in C1q (Figure 6). Therefore, the interaction with two different polysaccharides would allow SIGN-R1 complexes to efficiently identify and bring the microbial pathogen to the proximity of C1q, thus promoting the assembly of a C4bC2a, i.e., C3 convertase, and the C3 fixation of microbial polysaccharides. This mechanism could allow the efficient activation and fixation of C3 by S. pneumoniae and CPS upon entry into the bloodstream and spleen. Then, within minutes of intravenous injection, both CPS and S. pneumoniae are fixed with C3 to be sequestered from marginal zones to follicles, and much of the C3 in serum is catabolized (Kang et al., 2006).

Recent evidence suggests that complement and cellular complement receptors also play important roles in the localization and retention of nonbacterial pathogens to follicular dendritic cells, especially HIV-1 (Feinberg et al., 2007). Interestingly, SIGN-R1 binds viral glycoproteins as do human DC-SIGN(R) (Curtis et al., 1992). DC-SIGN binding to HIV can promote viral infection of T cells (Chehimi et al., 2003). It is worth noting that oligosaccharides present on the envelope of various viruses such as HIV also contain terminal α-2,6 sialic acid linkages (Varchetta et al., 2013). Therefore, in the mechanism we proposed, viral glycoproteins could block the interaction of sialylated C1q or Ig with SIGN-R1 through its primary binding site, thus avoiding efficient recognition and response against the pathogen by the immune system.
**Structure**

**Basis for Selective Recognition by SIGN-R1**

**EXPERIMENTAL PROCEDURES**

**Production of SIGN-R1**

A DNA construct encoding the soluble CRD protein of SIGN-R1 (CRD_SIGN-R1) was generated by amplifying the CRD domain from SIGN-R1 cDNA by PCR, cloned into a mammalian expression vector containing signal sequence and FLAG epitope and sequenced, as described previously (Park et al., 2008). The mammalian expression vector encoding this FLAG-tagged CRD_SIGN-R1 (GenBank accession EU697459) was transfected into Chinese hamster ovary (CHO) cells (CHO-S cells, Gibco, Life Technologies) using Lipofectamine 2000 reagent (Life Technologies) following the manufacturer’s instructions. Stable CHO cell lines producing the soluble or mature form of FLAG-tagged CRD_SIGN-R1 protein that consists of an N-terminal, 9 amino acid long FLAG epitope and a 134 amino acid long CRD domain of SIGN-R1 (residues from Cys192 to Gly325) were generated following selection with G418 (1.5 mg/ml) as described previously (Kang et al., 2004). CRD_SIGN-R1 protein was purified from the supernatant of stably transfected CHO cell culture with an affinity column using anti-FLAG M1 agarose gel (Sigma-Aldrich) following the manufacturer’s instructions.

**Binding Experiments**

The cDNA encoding an open reading frame of human DC-SIGN (hDC-SIGN) was cloned by PCR from a human thymic cDNA library (Clontech). CHO cells expressing hDC-SIGN were generated and cloned under G418 (1.5 mg/ml) selection pressure as described above. Stable FITC-labeled dextran (FITC-dextran, 500 kDa) was purchased from Sigma-Aldrich. PE-labeled CPS14 (PE-CPS14) was generated at home as follows. Purified CPS14 (xylophylized pneumococcal polysaccharide powder type 14) was purchased from ATCC, and 2 mg of CPS14 was biotinylated in essence as described previously (Cristagalli and Snapper, 2007; Koesling et al., 2001). After extensive dialysis against Dulbecco’s PBS (GIBCO, Life Technologies), 500 μg of biotinylated CPS14 in DPBS was mixed with 25 μg of PE-conjugated streptavidin (BD Biosciences) by rotator at 4°C overnight. CHO cells expressing SIGN-R1 (Kang et al., 2004) and hDC-SIGN (see above) were t Tinaplemented and plated into 12-well plates. When the plated CHO cells became confluent, each well was replenished with fresh media containing FITC-dextran (10 μg/ml in final concentration) or PE-CPS14 (60 μg/ml in final concentration) with without graded doses of EDTA. Then the plates were further incubated at 5% CO₂, 37°C for 1 hr (FITC-dextran) or 5 hr (PE-CPS14) before being analyzed for polysaccharide uptake with a LSR II flow cytometer (BD Biosciences).

**Production of α,2,6 Sialylated Antibody**

Hypersialylated human Fcs were generated as previously described, with minor modifications (Anthony et al., 2008). Briefly, IVG-derived Fc fragments were treated with 2,3/6 sialidase to remove sialic acid residues. Fc fragments (30 mg/ml) were incubated in 50 mM MOPS, pH 7.2, 20 mM MnCl₂ with 10 μM uridine diphosphate-galactose and 25 mM/10mg substrate of α,β-2,6 Sialyltransferase at 37°C. The glycosylation reactions were confirmed by lectin blotting with Erythrina cristagalli Lectin (terminal galactose), and Sambucus Nigra Lectin (terminal α,2,6 sialic acid).

**Crystalization of SIGN-R1 and SIGN-R1 Complexes**

Crystalization trials were performed at 18°C as previously described (Silva-Martin et al., 2009). Good quality crystals grown by mixing 1 μl of CRD_SIGN-R1 (4 mg/ml in 20 mM Tris/HCl pH 7.5 and 100 mM NaCl) with 1 μl of precipitant solution from commercial JCSG+ crystallization grid consisting of 0.1 M Bis-tris pH 5.5 and 1.6 M (NH₄)₂SO₄ (Quaigen). Drops were equilibrated against 0.5 μl. The best SIGN-R1 native crystal appears after 2 weeks with a maximum dimension of approximately 0.30 × 0.28 × 0.10 mm. Complexes of SIGN-R1 with dextran sulfate polysaccharide were obtained by incubation of a 4 mg/ml protein solution with 5 mM CaCl₂ and 3% dextran sulfate (Hampton research additive number 51). Best complex crystals appear after 3 weeks with maximum dimensions of approximately 0.20 × 0.18 × 0.10 mm. Complexes of the SIGN-R1 with sialic acid were obtained after overnight soaking of SIGN-R1 native crystals with precipitant solution supplemented with 10 mM sialic acid (SIGMA). All crystals were cryoprotected into a mixture of 3M Li₂SO₄ with crystallization conditions in a 1:1 volume proportion. Then, crystals were flash-cooled in liquid nitrogen and maintained at 100 K during data collection.

**Crystallization of the Glycosylated Human Fc**

Crystals were obtained by mixing 1 μl of Fc-Sial (9 mg/ml) with 3 μl of precipitant solution consisting of 0.1 M HEPES pH 7.0 and 0.1 M PEG 6000 under hanging drop vapor diffusion experiment at 18°C. The best Fc-Sial crystals appear after 2 weeks, reaching maximum dimensions of approximately 0.70 × 0.30 × 0.40 mm. Before data collection crystals were cryoprotected in a solution containing mother liquor and 25% glycerol. Then, crystals were flash-cooled in liquid nitrogen and maintained at 100 K during data collection.

**X-Ray Data Collection and Structural Determination**

X-ray data sets of CRD_SIGN-R1 native and ligand complexes were collected using a synchrotron radiation source at the ESRF (Grenoble). The native crystals data set was measured with an ADSC Quantum Q315r detector at beamline ID23-1. SIGN-R1:DexS data sets were measured with an ADSC Quantum 4 detector at beamline ID14-2. Sign-R1:NeuAc data sets were measured with a MAR225 detector at beamline ID23-2. All data sets were processed using MOSFLM (Moras et al., 1992) and scaled using SCALA from the CCP4 package (CCP4, 1994). All SIGN-R1 crystals belong to monoclinic C2 space group containing four independent molecules in the asymmetric unit. Data processing results are summarized in Table 1.

The native structure was determined by molecular replacement with MOLREP program from the CCP4 package (CCP4, 1994) using human DC-SIGN structure (PDB code ISL4) with 69% of identity, as initial model. Molecular replacement rotation and translation correlation coefficients were ranked and yielded four solutions well above the background; the final solution for R and Sdiff were 39.7% and 57.3%, respectively. X-ray data sets for sialylated Fc were measured with a Q315r ADSC CCD detector on ID14-4 beamline (ESRF, Grenoble). Crystals belong to the P2₁2₁2₁ orthorhombic space group containing two molecules in the asymmetric unit. Structural determination was done by molecular replacement using a monomer of the human IgG1 structure (PDB code 1HZ), which shows 99% sequence identity for the Fc portion. Molecular replacement was performed with the MOLREP using a 3.5 Å cutoff limit for resolution. Four single and unambiguous solutions were obtained, with final correlation coefficients of 25% for R and 54% for Sdiff.

In all cases, initial refinement and map calculations were carried out using CNS (Brüger et al., 1998). Bulk solvent correction was applied during the refinement. Water molecules were gradually added to peaks higher than 3σ of Fc-Fc maps with WATER PICK program routine of CNS program. Iterative cycles of refinement and manual model building were done. Model building was performed with Coot (Emsley and Cowtan, 2004) program. Further refinement using maximum likelihood methods was performed with the REFMAC5 from the CCP4 package (CCP4, 1994) and PHENIX (Adams et al., 2010) program. The final geometry of the structures was evaluated using the program PROCHECK (Laskowski et al., 1993). All the final models presented good stereochemistry parameters and R and Rfree values (Table 1).

**Docking Calculations**

Two polysaccharides were docked at the SIGN-R1 secondary binding site: the capsular pneumococcal polysaccharide 14 (CPS14) and the polyribitol phosphate. The protein-ligand docking experiments were performed to test whether it was possible to find a compatible model with the experimental evidence of the repetitive Glucose-Sulfate moieties observed in the crystallographic studies. To this end, we first docked the CPS14, one of the highest affinity capsular polysaccharide reported to bind to SIGN-R1 (Kang et al., 2004). The 3D structure of CPS14 was modeled in the sweet server (http://www.glycosciences.de/modeling/sweet2/doc/index.php), using the unit repetitive tetrasaccharide of the S. pneumoniae serotype 14 [Galα(1,4)-Galβ(1,3)-Galβ(1,4)-Galβ(1,6)-; rNAG as a template. Molecular docking was carried out using GOLD (Genetic Optimization for Ligand Docking) software (Verdonk et al., 2003), using the genetic algorithm (GA). This method allows a partial flexibility of protein and full flexibility of the ligand. For each of the
Molecular Dynamics Simulations

Simulations were based on the crystal structure of the human Fc fragment immunoglobulin reported in this paper. Long simulations were performed to investigate the dynamics of the siatic acid in positions 2–6 and 2–3. We used this crystallographic structure of 1MCO (Guettard et al., 1993) as a reference for the 2–3 conformation. The final glycodies molecules where prepared for the Glycam_06 (Kirschner et al., 2008) force field, whereas the AMBER ff99SB (Hornak et al., 2006) force field was used for the protein structure. Prepared structures were immersed in triclinic boxes of TIP3P water molecules (Jorgensen et al., 1993). Counter ions were also added to maintain electroneutrality. Before and after the addition of ions, two minimizations of 10,000 iterations were performed using a steepest descend and conjugated gradient method.

All minimizations and MD simulations were performed using the GROMACS 4 (Hess et al., 2008). Time steps of 2 fs were used for the integrations of Newton’s equations of motion. The simulations have been carried out using periodic boundary conditions with the NPT (N, total number of atoms; P, pressure; T, temperature) ensemble. Electrostatic interactions were evaluated using the particle-mesh Ewald (PME) method (Darden et al., 1993) with Van der Waals interactions truncated at 14 Å. The systems were then heated and equilibrated in two steps: (1) 300 ps of MD with restricted position of all atoms of the glyco-protein and heating whole system from 50 K to 298 K, and (2) 150 ps of MD with restricted position of the backbone atoms. The equilibrated structures were the starting points for 200 ns MD simulations at constant temperature (298 K) and pressure (1 atm). Bond lengths were constrained using the LINCS algorithm, and water geometries constrained by SETTLE. Analyses were performed using the GROMACS 4 suite of programs. Demonstration of system geometries and interactions were done using CHIMERA.

ACCESSION NUMBERS

The Protein Data Bank accession numbers for the atomic coordinates codes reported in this paper are 3ZH6(SIGN-R1), 4CF9(SIGN-R1:sulfodextran complex), 4CAJ(SIGN-R1:siatic acid complex), and 4CDH (α-2,6 sialylated antibody FC).

SUPPLEMENTAL INFORMATION

Supplemental Information includes six figures and two movies and can be found with this article online at http://dx.doi.org/10.1016/j.str.2014.09.001.
stimulation fully differentiates monocytes to DC-SIGN/CD209(+) dendritic cells for immune T cell areas. Cell 143, 416–429.


