

# High Resolution Crystal Structures of Siglec-7

INSIGHTS INTO LIGAND SPECIFICITY IN THE SIGLEC FAMILY\*

Received for publication, October 16, 2002, and in revised form, November 15, 2002  
Published, JBC Papers in Press, November 15, 2002, DOI 10.1074/jbc.M210602200

Magnus S. Alphey<sup>‡§</sup>, Helen Attrill<sup>‡§</sup>, Paul R. Crocker<sup>‡¶</sup>, and Daan M. F. van Aalten<sup>§||</sup>

From the Divisions of <sup>‡</sup>Cell Biology and Immunology and <sup>§</sup>Biological Chemistry and Molecular Microbiology, School of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland, United Kingdom

**Sialic acid-binding immunoglobulin-like lectins (Siglecs) recognize sialylated glycoconjugates and play a role in cell-cell recognition. Siglec-7 is expressed on natural killer cells and displays unique ligand binding properties different from other members of the Siglec family. Here we describe the high resolution structures of the N-terminal V-set Ig-like domain of Siglec-7 in two crystal forms, at 1.75 and 1.9 Å. The latter crystal form reveals the full structure of this domain and allows us to speculate on the differential ligand binding properties displayed by members of the Siglec family. A fully ordered N-linked glycan is observed, tethered by tight interactions with symmetry-related protein molecules in the crystal. Comparison of the structure with that of sialoadhesin and a model of Siglec-9 shows that the unique preference of Siglec-7 for  $\alpha(2,8)$ -linked disialic acid is likely to reside in the C-C' loop, which is variable in the Siglec family. In the Siglec-7 structure, the ligand-binding pocket is occupied by a loop of a symmetry-related molecule, mimicking the interactions with sialic acid.**

The Siglecs<sup>1</sup> are a specialized subgroup of the Ig superfamily that share significant sequence similarity and the ability to recognize sialylated glycoconjugates (1). There are at least 11 *bona fide* members in humans and potentially 8 in mice, all of which are type I membrane proteins, containing an N-terminal sialic acid-binding V-set Ig domain and varying numbers of C2-set Ig domains (2). Apart from myelin-associated glycoprotein (MAG, Siglec-4), which is found uniquely in the nervous system, Siglecs are expressed predominantly in the hemopoietic and immune systems and appear to mediate both adhesive and signaling functions (3).

Siglecs can be divided into two subgroups based on sequence similarity in the extracellular and intracellular regions. Sialoadhesin (Siglec-1), CD22 (Siglec-2), and MAG constitute one subgroup, share ~25–30% sequence identity in the extracellu-

lar region, and have divergent cytoplasmic tails. The second subgroup is made up of the CD33-related Siglecs that include CD33 (Siglec-3) and the recently discovered human Siglecs 5–11 (reviewed in Ref. 3). These proteins share 50–80% sequence similarity and have two highly conserved tyrosine-based motifs in their cytoplasmic tails. In humans, the CD33-related Siglec genes are clustered on chromosome 19q13.3–4 and are separate from the genes encoding CD22, MAG (19q13.1), and sialoadhesin (20p). Recent studies using specific monoclonal antibodies have shown that the CD33-related Siglecs are expressed in a partially overlapping manner on all major leukocytes of the innate immune system, suggesting a role for these proteins in regulation of leukocyte function. This is further supported by the finding that human CD33-related Siglecs have two conserved immunoreceptor tyrosine-based inhibitory motif-like sequences, which in all cases studied can interact with the tyrosine phosphatases SHP-1 and SHP-2 following tyrosine phosphorylation (4–8). Recent studies of CD22 (Siglec-2) have also illustrated the importance of sialic acid recognition and specificity in triggering biological functions mediated by Siglecs (9–11).

To date, structural information on the Siglecs is limited to the V-set domain of sialoadhesin (Siglec-1) in complex with 3'-sialyllactose (12). A 1.8-Å crystal structure revealed the basis for recognition of the terminal sialic acid. This mode of recognition is likely to be common to all Siglecs. However, the structure did not give insights into the differential specificity for sialic acid linkages exhibited by the Siglecs. Furthermore, the low sequence similarity shared between sialoadhesin and the CD33-related Siglecs is an obstacle to interpreting mutagenesis data in a structural framework. It would therefore be desirable to obtain the structure of at least one member of the CD33-related subgroup, which could be used as a template for modeling other CD33-related Siglecs as well as for gaining molecular insights into sialic acid linkage recognition in the Siglec family. In this context, Siglec-7 is of particular interest since it is the major Siglec expressed by natural killer cells (13, 14) and possesses a unique preference for  $\alpha(2,8)$ -linked disialic acids over  $\alpha(2,6)$ - and  $\alpha(2,3)$ -linked sialic acids (15). When cross-linked at the cell surface in a redirected killing assay, Siglec-7 was capable of inhibiting the cytotoxicity toward cell targets and is therefore a potential natural killer cell inhibitory receptor (13). Similar to CD22 on B cells, glycan recognition by Siglec-7 is likely to be directly linked to its function in modulating the activation of natural killer cells.

Siglec-7 is closely related to Siglec-9 with 80% overall sequence identity (2, 16). However, Siglec-9 cannot bind  $\alpha(2,8)$ -linked disialic acids and prefers  $\alpha(2,3)$ - and  $\alpha(2,6)$ -linked terminal sialic acids (2, 15, 16). Using protein chimeras, this difference in sialic acid linkage preference was recently shown to reside in a 6-amino acid stretch within the C-C' loop of

\* This work was funded by Biotechnology and Biological Sciences Research Council Grant 94/B14010. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The atomic coordinates and structure factors (code 1O7S and 1O7V) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (<http://www.rcsb.org/>).

¶ Supported by a Wellcome Trust Senior Research Fellowship.

|| Supported by a Wellcome Trust Career Development Research Fellowship. To whom correspondence should be addressed. Fax: 44-1382-345764; E-mail: [dava@davapc1.bioch.dundee.ac.uk](mailto:dava@davapc1.bioch.dundee.ac.uk).

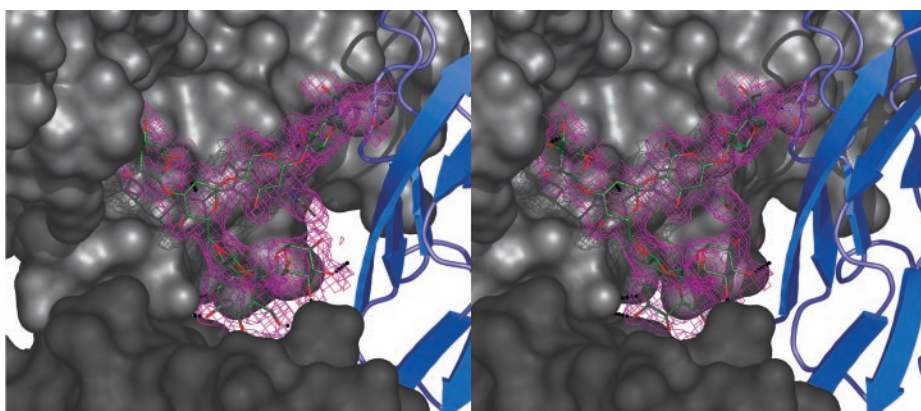
<sup>1</sup> The abbreviations used are: Siglec, sialic acid-binding immunoglobulin-like lectin; CHO, Chinese hamster ovary; MES, 4-morpholineethanesulfonic acid.

TABLE I  
Details of data collection and structure refinement for the two Siglec-7 crystal forms

Values between brackets are for the highest resolution shell. All measured data were included in structure refinement.

Structure	Sig7a	Sig7b
Wave length (Å)	0.953758	0.933
Space group	P <sub>4</sub> <sub>1</sub> 2 <sub>1</sub> 2	P <sub>4</sub> <sub>3</sub> 2 <sub>1</sub> 2
Unit cell (Å)	<i>a</i> = 53.08 <i>b</i> = 53.08 <i>c</i> = 93.77	<i>a</i> = 40.05 <i>b</i> = 40.05 <i>c</i> = 140.77
Resolution (Å)	30 – 1.75 (1.81 – 1.75)	25 – 1.9 (1.97 – 1.90)
Observed reflections	86594	31409
Unique reflections	14144	9206
Redundancy	6.1 (6.2)	3.4
Completeness (%)	100 (100)	94.8 (96.3)
<i>R</i> <sub>merge</sub>	0.065 (0.562)	0.061 (0.312)
<i>I</i> / $\sigma$ <i>I</i>	8.6 (3.4)	11.7 (2.4)
<i>R</i> <sub>free</sub> reflections	544	479
<i>R</i> <sub>cryst</sub>	0.223	0.210
<i>R</i> <sub>free</sub>	0.249	0.255
Number of atoms		
Protein	902	1021
Water	95	86
Glycan	29	83
<B> Protein	28.3	31.7
<B> Water	36.8	39.0
<B> Glycan	63.8	51.6
Root mean square deviation from ideal geometry		
Bond lengths (Å)	0.013	0.007
Bond angles (°)	1.56	1.28
Main chain B (Å <sup>2</sup> )	1.52	1.67

FIG. 1. **Glycan density.** The structure of the *N*-linked glycan attached to Asn-105 is shown. The protein is shown as a ribbon (blue). The molecular surfaces of symmetry-related protein molecules are shown in gray. The glycan is represented by a stick model. Hydrogen bonds are shown as black dashed lines. The final  $2F_o - F_c$ ,  $\phi_{\text{calc}}$  map (contoured at  $1\sigma$ ) is shown in magenta.



Siglecs-7 and -9 (15). In the present study, we have obtained the first crystal structure of the Siglec-7 V-set domain, in two crystal forms refined to 1.75 and 1.9 Å. Comparison of Siglec-7 with a model of Siglec-9 shows that differences in residues in the tip of the C-C' loop may explain the preference of Siglec-7 for  $\alpha(2,8)$ -linked disialic acids. The Siglec-7 structure also reveals a fully ordered *N*-linked glycan and interesting crystallographic packing interactions within the ligand-binding site. The structure will provide a useful template for modeling other CD33-related Siglecs.

#### MATERIALS AND METHODS

**Molecular Cloning, Overexpression, and Purification**—The Siglec-7 V-set domain was PCR-amplified from CD33-like HDPUW68 in the pSPORT vector (Invitrogen) (14) using the forward primer (5'-ACT-TCTAGAGCACCTCCAACCCAGATATG-3') and reverse primer (5'-ACTGGATCCTTATGTCACGTCACAGAGAGCTG-3'). The amplified gene was inserted into the pDEF mammalian expression vector via *Xba*I and *Bam*HI restriction sites introduced by the primers (bold). The resulting plasmid was transformed into Chinese hamster ovary Lec1 (CHO Lec1) cells using FuGENE (Roche Pharmaceuticals). The CHO Lec1 cells stably expressing the Siglec-7 V-set domain were selected in the presence of hygromycin and cultivated in  $\alpha$ -MEM containing 5% fetal calf serum and 1% penicillin/streptomycin mix (Invitrogen). Secreted protein was harvested and filtered through a 0.2- $\mu$ m filter and then passed over an 8-ml anti-Siglec-7 polyclonal antibody column. This

was prepared by coupling affinity-purified sheep anti-Siglec-7 IgG to cyanogen bromide-activated Sepharose CL-4B (Sigma) at 5 mg/ml. Protein was eluted from the antibody affinity column using 8-ml aliquots of 0.1 M glycine buffer, pH 2.5, followed by immediate neutralization with 10% 1.0 M Tris-HCl, pH 8.0. Resulting fractions were buffer-exchanged into 0.025 M Tris, pH 7.5, 0.1 M NaCl. Further purification was achieved using a Superdex 75 16/60 column on an Akta Purifier system (Amersham Biosciences). Purity was assessed by SDS-PAGE and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. The protein was concentrated for crystallization to 5 mg/ml using a Vivaspin 10-kDa cutoff spin concentrator (Vivascience).

**Crystallization and Data Collection**—Crystallization experiments were carried out at 20 °C using hanging drop vapor diffusion. Crystallization conditions were found using a sparse matrix sampling approach with Crystal Screens I and II from Hampton Research. A drop of 0.5  $\mu$ l of protein (5 mg/ml in 25 mM Tris-HCl, pH 8.0) was mixed with 0.5  $\mu$ l of reservoir solution and equilibrated against 250  $\mu$ l of reservoir solution. Small crystals appeared after 2 days in Screen I conditions 9 (0.2 M ammonium acetate, 0.1 M trisodium citrate, pH 5.6, 30% polyethylene glycol 4000) and 28 (0.2 M sodium acetate trihydrate, 0.1 M sodium cacodylate, pH 6.5, 30% polyethylene glycol 8000). The initial crystals (crystal form 1, Sig7a) from condition 9 were cryoprotected using solution 9 from Cryo Screen (Hampton Research) and frozen in a nitrogen stream. Subsequently, a second crystal form (Sig7b) was found in conditions identical to those for Sig7a. These crystals were soaked in a cryoprotectant consisting of 30% polyethylene glycol monomethyl-ether 5000, 0.1 M MES, pH 6.5, 0.2 M ammonium sulfate, and 5%

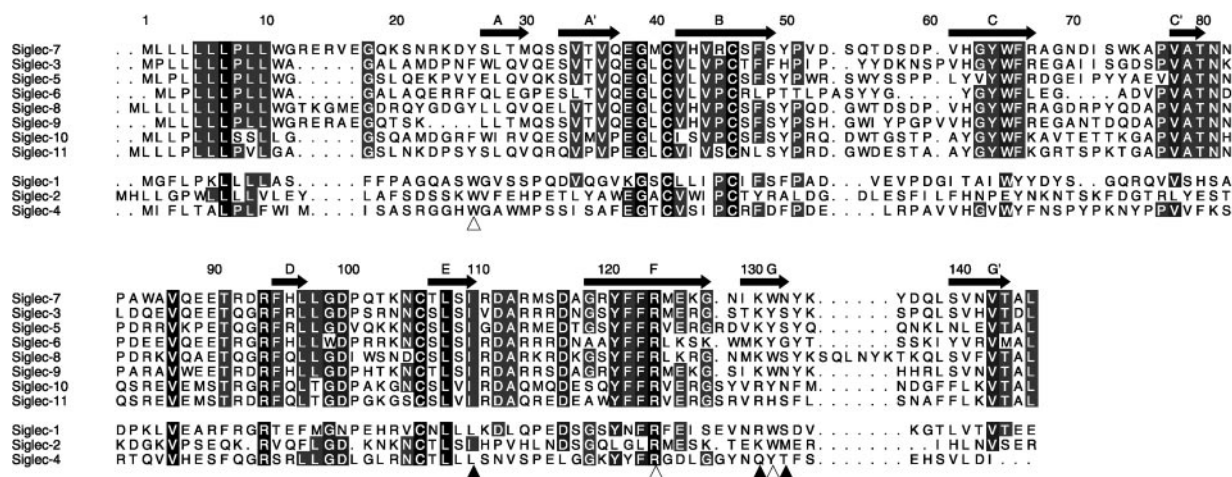
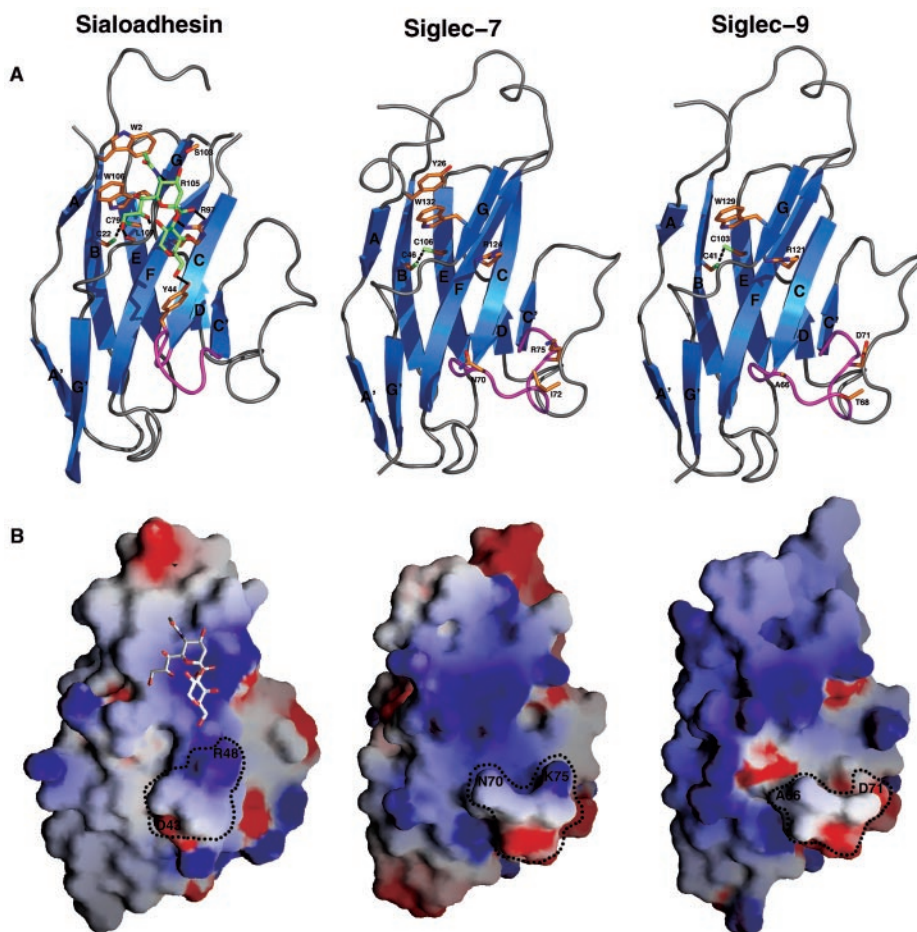


FIG. 2. **Sequence alignment of the human Siglec family.** All currently known human Siglec sequences are shown in a CLUSTAL-W (33) sequence alignment with sequence conservation indicated by shaded boxes. Secondary structure elements are shown as defined by the Siglec-7 structure. Residues surrounding the ligand-binding site are indicated by open triangles (side chain contacts) and closed triangles (backbone contacts).

**FIG. 3. Comparison of sialoadhesin, Siglec-7, and Siglec-9 structures.** The crystal structures of sialoadhesin in complex with sialyllactose (Protein Data Bank entry 1QFO), the crystal structure of Siglec-7 (described here), and a model for Siglec-9 are shown in two representations. In *A*, the backbones are shown as ribbons with the side chains surrounding the ligand-binding site shown as sticks with orange carbons. The C-C' loop is highlighted in magenta. The two pyranose sugars from the sialyllactose ligand are shown as sticks with green carbons. Hydrogen bonds with the protein are shown as black dotted lines. *B*, electrostatic surfaces calculated with GRASP (34) contoured at  $-6$  kiloteslas (red) and  $+6$  kiloteslas (blue). For sialoadhesin, the two pyranose sugars of the ligand are shown in a sticks representation. The surface area corresponding to the C-C' loop is highlighted by dashed circles.



2-methyl-2,4-pentanediol and then frozen in a nitrogen stream.

X-ray data for Sig7a were collected at beamline ID14-EH2 European Synchrotron Radiation Facility (ESRF) on an ADSC Quantum4 CCD at a temperature of 100 K (see Table I). Data for Sig7b were collected at beamline ID29 (ESRF) on an ADSC Q210 CCD at a temperature of 100 K (see Table I). All data were processed and scaled with the HKL suite of programs (17).

**Structure Determination and Refinement**—The Sig7a structure was solved by molecular replacement with AMoRe (18) using the sialoadhesin structure (Protein Data Bank entry: 1QFO (12)) as a search model. A single solution was found with an *R*-factor of 0.505 (correlation

coefficient = 0.337). The resulting model phases were used as input for Arp/wARP (19), which built 82 out of 127 residues. This was followed by refinement with CNS (20) interspersed with model building in O (21). Five percent of the data were set aside for calculation of  $R_{\text{free}}$  (22). After several rounds of refinement, the model was incomplete. No electron density was observed for the B-C and the C-C' loops (residues 53–57 and 68–74, respectively) and the N terminus (residues 18–23). This structure was used as a search model in AMoRe against 8–4-Å diffraction data collected on Sig7b. A single solution was found with an *R*-factor of 0.386 (correlation coefficient = 0.723). This was followed by refinement in CNS and manual model building with O. A well defined

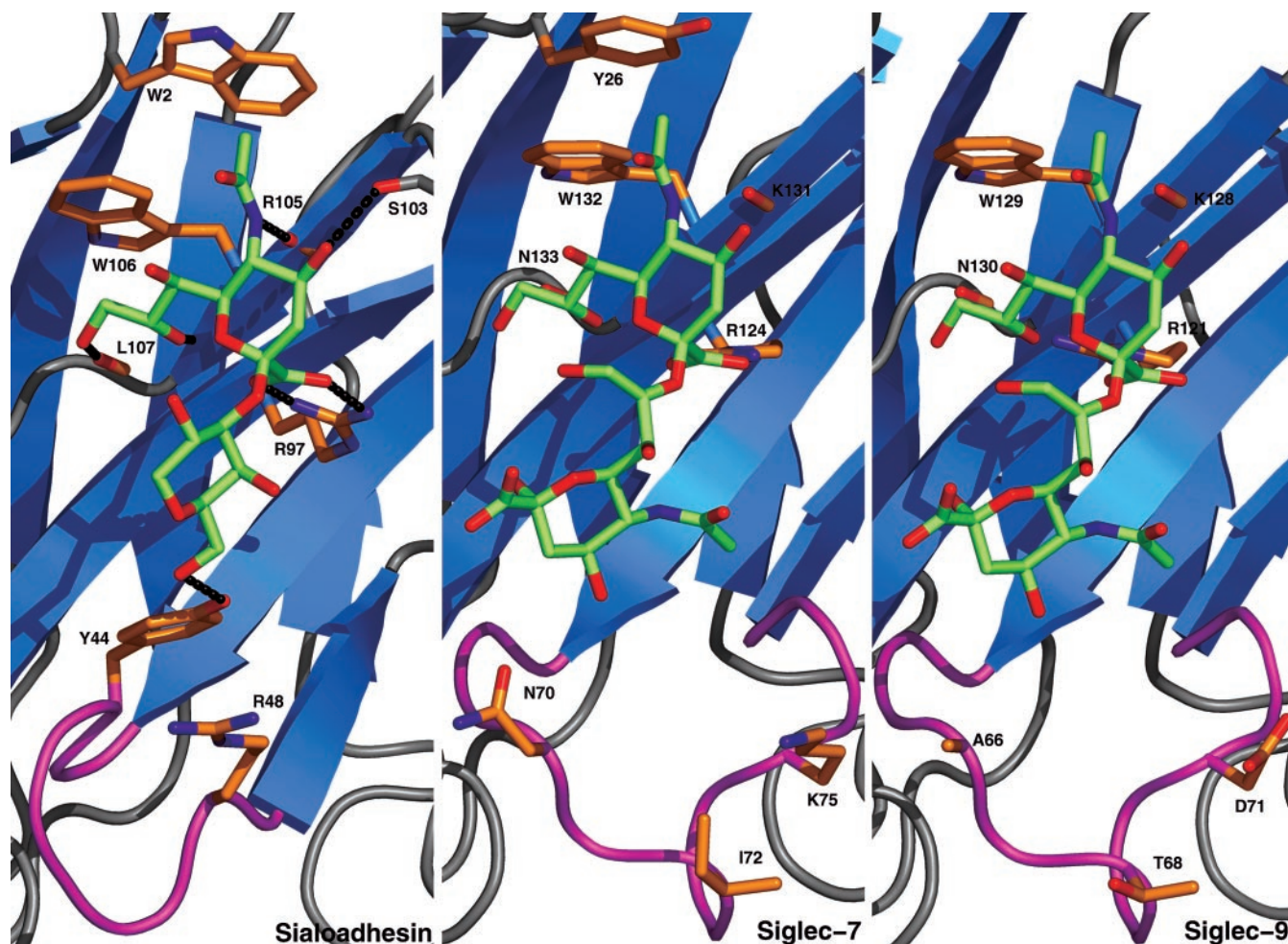


FIG. 4. **Comparison of the ligand-binding sites.** The backbone structures of sialoadhesin, Siglec-7, and Siglec-9 are shown as *ribbons*. Key residues interacting with the terminal pyranose sugars are shown for sialoadhesin (*sticks with orange carbons*) and compared with the equivalent residues in the Siglec-7 structure and the Siglec-9 model. The terminal sialic acid of the sialoadhesin sialyllactose ligand is shown as *sticks with green carbons*. Models for  $\alpha(2,8)$ -linked disialic acid in Siglec-7 and Siglec-9 are shown as *sticks with green carbons*. Hydrogen bonds with the protein are shown as *black dotted lines*.

density for the complete glycan structure attached to Asn-105 was observed (see Fig. 1). In general, the Sig7b electron density maps, although calculated from lower resolution data, were of higher quality than those for Sig7a and allowed building of the flexible loops and extension of the N terminus, resulting in a model with a final  $R$ -factor of 0.210 ( $R_{\text{free}} = 0.255$ , see Table I).

A model of Siglec-9 was produced by homology modeling with WHAT IF (23) using the Sig7b structure as a template. A model of the Siglec-7/9 interaction with disialic acid was produced through superposition of the sialoadhesin-sialyllactose structure (Protein Data Bank entry 1QFO (12)) of the Siglec-7/9 structures followed by superposition of disialic acid (from Protein Data Bank entry 1FV2 (24)) using the coordinates for the first sialic acid.

#### RESULTS AND DISCUSSION

**Overall Structure**—The structure of the Siglec-7 V-set ligand-binding domain was solved by molecular replacement in two crystal forms (Sig7a and Sig7b) and refined to 1.75 Å ( $r = 0.223$ ,  $R_{\text{free}} = 0.249$ ) and 1.9 Å ( $r = 0.210$ ,  $R_{\text{free}} = 0.255$ ) resolution, respectively (Table I). The Sig7a structure is incomplete, with loops B-C, C-C', and the 6 N-terminal residues missing. The Sig7b structure is complete from residue 18 to 144 and includes a fully ordered *N*-linked glycan with a truncated structure as defined by the CHO Lec1 cell line used (discussed below). All analyses described here are based on the Sig7b structure, although all figures shown here include the side chain conformation of Arg-124 (a key ligand-binding residue) as seen in the Sig7a structure. The Sig7b Ramachandran plot

calculated with PROCHECK (25) shows that 98.3% of non-glycine residues are in the most favorable and additionally allowed regions. Two residues, Gln-19 and Asp-93, are in disallowed regions, which are located in disordered regions of the map.

CHO Lec1 mammalian cells cannot synthesize complex oligosaccharides (26) due to a point mutation in the *N*-acetylglucosaminyltransferase I (GnTI), resulting in homogenous glycosylation with the structure  $\text{Man}_5\text{-GlcNAc}_2$  (27)). This is advantageous for protein crystallography studies, which can be hampered by glycan heterogeneity. The crystal structure of Sig7b reveals a complete  $\text{Man}_5\text{-GlcNAc}_2$  glycan *N*-linked to Asn-105 that is well defined in the electron density maps (Fig. 1). This high degree of order can be attributed to crystal packing. There are 3 hydrogen bonds between the glycan and the protein and an additional 16 hydrogen bonds between the glycan and symmetry-related protein molecules (Fig. 1). Thus, a total of 19 hydrogen bonds tether the glycan within the crystal lattice.

**Comparison with Sialoadhesin**—Siglec-7, given its higher sequence identity with other members of the Siglec family as compared with sialoadhesin (Fig. 2), may provide a more suitable template for interpretation of structure-function relationships in the whole Siglec family. Despite the low sequence identity of Siglec-7 with sialoadhesin (Fig. 2), the structures share many common features and superpose with

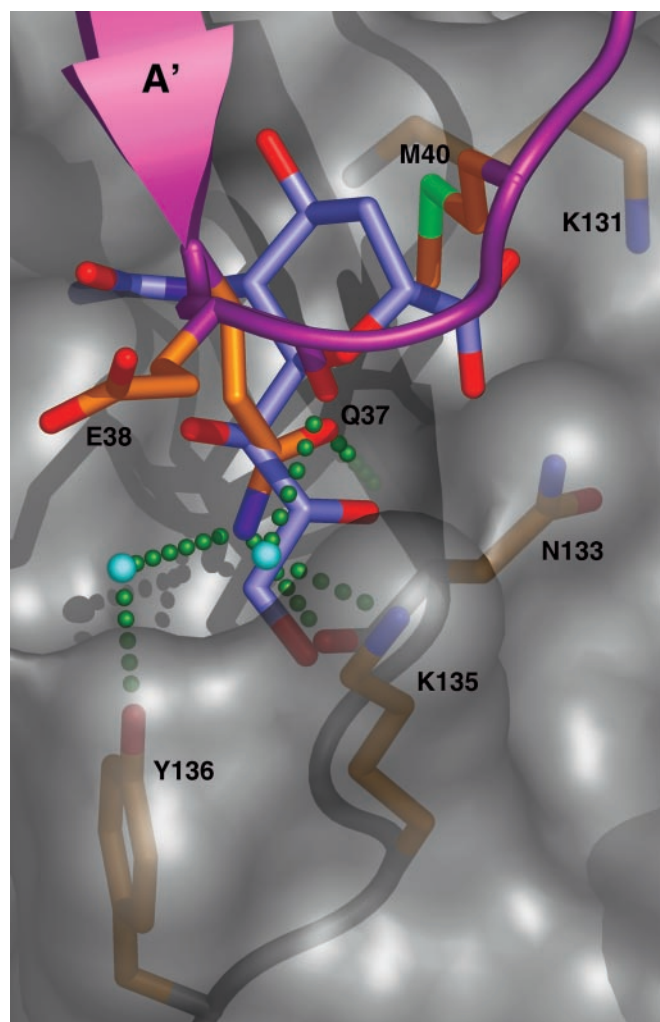


FIG. 5. **A crystal contact in the ligand-binding site.** The molecular surface of the Siglec-7 structure is shown in gray. The A'-B loop of a symmetry-related molecule is shown in magenta. Residues on this loop pointing into the ligand-binding site are shown as sticks with orange carbons. Hydrogen bonds are shown as green dashed lines. The terminal sialic acid as observed in the sialoadhesin-sialyllactose complex is shown as a stick model with blue carbons. Water molecules involved in hydrogen bonding to both the protein and the symmetry-related A'-B loop are shown as cyan spheres.

a root mean square deviation of 1.04 Å using 100 C $\alpha$  atoms (Fig. 3). Siglec-7 shows an Ig-like fold based on a  $\beta$ -sandwich formed by two  $\beta$ -sheets consisting of strands A'GFCC' and ABED, respectively. The G strand is split (forming G and G'), as is the A strand (to give A and A'). As with sialoadhesin, an intra-sheet disulfide is present (between Cys-46 and Cys-106), replacing the inter-sheet disulfide more commonly observed in other Ig-like folds (28). The cysteine on strand F is replaced by Phe-123, which lies 10 Å from Cys-46, farther than the 6–8-Å distance observed in other Ig domains (28). This results in a widening of the cleft between the sheets, exposing residues that may interact with ligand. The largest differences between the sialoadhesin and Siglec-7 structures occur in the flexible loops on the outside of the protein, namely the B-C and C-C' loops, the latter of which is different in length (Figs. 2 and 3). The C-C' loops of both Siglec structures extend away from the main body of the protein (Fig. 3).

**The Ligand-binding Site**—Despite repeated attempts with a range of ligands, either by soaking or by co-crystallization, a crystal of a Siglec-7-ligand complex could not be obtained.

Although the structure of Sig7b was solved without a ligand in the binding site, comparison of the sialoadhesin structure with 3'-sialyllactose-bound (12) provides a model of the interactions that may occur and identifies the residues essential for specificity. The ligand-binding site lies between strands A and G (Figs. 3 and 4). As discussed previously, the partial opening of the  $\beta$ -sandwich, caused by the absence of the inter-sheet disulfide, provides a large flat surface onto which the ligand binds in the sialoadhesin structure (12). Compared with sialoadhesin, this surface has a more basic character in Sig7b, resulting from the additional basic residues Arg-23, Arg-120, and Lys-135 (Fig. 3). The additional positive charge presented by the binding face of Siglec-7 may provide a further docking site for a negatively charged sugar, possibly explaining the Siglec-7 preference for disialylated ligands.

**Structural Basis for Differential Specificity**—Comparison of the sialic acid-binding site of Siglec-7 and sialoadhesin show that many of the residues important in the interaction with sialic acid are conserved (Figs. 2 and 4). An arginine (Arg-124), conserved in all Siglecs, interacts with the carboxyl group on the terminal sialic acid sugar (Fig. 4). Trp-132 provides a hydrophobic interaction with the glycerol moiety. Protein backbone hydrogen bonding with the glycerol and *N*-acetyl groups seen in the sialoadhesin complex are also possible in the Siglec-7-binding site. However, there are some key differences. Sialoadhesin Trp-2 forms a hydrophobic contact with the *N*-acetyl methyl group. This tryptophan is replaced by Tyr-26 in Siglec-7, which could potentially hydrogen-bond with the *N*-acetyl carbonyl. In Siglec-9, this residue is absent. In sialoadhesin, mutation of the equivalent Trp-2 revealed that the hydrophobic interaction is important for *N*-acetyl neuraminic acid interactions (12). This could also explain why sialoadhesin cannot interact with *N*-glycoylneuraminic acid since the additional oxygen atom in the latter form of sialic acid would be expected to result in a steric clash (Fig. 4). Apart from sialoadhesin and MAG, all other siglecs examined so far, including hCD22, hCD33, and Siglec-6 (29), can bind both *N*-acetyl and *N*-glycoylneuraminic acid, suggesting that the hydrophobic contact with the *N*-acetyl neuraminic acid may not be required for sialic acid recognition in all cases. This could also explain why Siglec-9 mediates robust sialic acid binding in the absence of an equivalent aromatic residue (2, 15, 16).

An important aspect of Siglec function is the differential linkage specificity displayed by members of this family. This issue was recently addressed in a domain-swapping experiment between Siglec-7 and -9 (15). Surprisingly, this experiment showed that a single stretch of 6 amino acids in the tip of the C-C' loop could confer Siglec-9-like binding specificity on Siglec-7. The comparison of sialoadhesin with Siglec-7 shows that in the latter, the C-C' loop is longer and extends farther away from the body of the protein, offering the potential for more interactions with additional sugars (Figs. 3 and 4). In particular, the absence of an equivalent of sialoadhesin Tyr-44 in Siglec-7 creates a larger cavity and exposes the basic Lys-75 (Fig. 4). Sequence alignment shows that the C-C' loop is variable in the Siglec family, and it is thus possible that interactions with specific side chains that extend toward the binding site for the second sugar are responsible for binding selectivity. For Siglec-7, these residues are represented by Asn-70, Ile-72, and Lys-75, whereas the equivalent residues in Siglec-9 are Ala-66, Thr-68, and Asp-71. Thus, from a structural point of view, it is possible that the linkage specificity displayed by Siglecs is due to the interaction of side chains in the C-C' loop with subterminal sugars.

**A Symmetry-related Loop Occupies the Binding Pocket**—Although Siglec-7 was crystallized in the absence of any carbo-

hydrate ligands, the binding pocket is not empty. Expansion of the space group symmetry shows that the pocket is occupied by the A'-B loop and C terminus of a symmetry-related protein molecule (Fig. 5), burying a total surface area of 200 Å<sup>2</sup>. Comparison with sialic acid binding in sialoadhesin shows that the amide of Gln-37 mimics the interaction with the sialic acid glycerol side chain (Fig. 5). Met-40 has hydrophobic interactions with the Cβ of Lys-131 and Cα of Trp-132. In addition, there are several water-mediated hydrogen bonds (Fig. 5). Although the interactions with the symmetry-related loop are not extensive, this is one of the few examples of a peptide occupying a carbohydrate-binding pocket (30) and perhaps represents a first step toward a peptide-based Siglec-7 inhibitor. Interesting examples of a such peptide inhibitors have been described recently for family 18 chitinases (31) and concanavalin A (32).

**Conclusions**—The structure of the Siglec-7 sialic acid-binding domain is the first such report on a member of the CD33-related Siglec subset. Our results support previous predictions that sialic acid recognition by Siglecs is based on a common template and provide insights into the molecular basis of sialic acid linkage specificity exhibited by different Siglec family members. The structure presented here could be used as a template for the design of specific inhibitors, which could be used to dissect the precise role of CD33-related Siglecs in the regulation of leukocyte activation. Future studies will be aimed at elucidating the precise structural basis for sialic acid recognition by this family of mammalian lectins.

**Acknowledgments**—We thank the European Synchrotron Radiation Facility, Grenoble, France, for the time at beamline ID14-4. We also thank Rod McEver for providing the CHO Lec1 mutant cells and for helpful discussions on their use in protein expression. We also thank Yvonne Jones for helpful comments.

## REFERENCES

- Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y. Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., van den Berg, T. K., van der Merwe, P. A., Watt, S. M., and Varki, A. (1998) *Glycobiology* **8**, v
- Angata, T., Hingorani, R., Varki, N. M., and Varki, A. (2001) *J. Biol. Chem.* **276**, 45128–45136
- Crocker, P. R. (2002) *Curr. Opin. Struct. Biol.* **12**, 609–615
- Paul, S. P., Taylor, L. S., Stansbury, E. K., and McVicar, D. W. (2000) *Blood*, **96**, 483–490
- Ulyanova, T., Shah, D. D., and Thomas, M. L. (2001) *J. Biol. Chem.* **276**, 14451–14458
- Whitney, G., Wang, S. L., Chang, H., Cheng, K. Y., Lu, P., Zhou, X. D., Yang, W. P., McKinnon, M., and Longphre, M. (2001) *Eur. J. Biochem.* **268**, 6083–6096
- Yu, Z. B., Lai, C. M., Maoui, M., Banville, D., and Shen, S. H. *J. Biol. Chem.* **276**, 23816–23824
- Yu, Z. B., Maoui, M., Wu, L. T., Banville, D., and Shen, S. H. (2001) *Biochem. J.* **353**, 483–492
- Jin, L., McLean, P. A., Neel, B. G., and Wortis, H. H. (2002) *J. Exp. Med.* **195**, 1199–1205
- Kelm, S., Gerlach, J. Brossmer, R., Danzer, C. P., and Nitschke, L. (2002) *J. Exp. Med.* **195**, 1207–1213
- Lanoue, A. Batista, F. D., Stewart, M., and Neuberger, M. S. (2002) *Eur. J. Immunol.* **32**, 348–355
- May, A. P., Robinson, R. C., Vinson, M., Crocker, P. R., and Jones, E. Y. (1998) *Mol. Cell*, **1**, 719–728
- Falco, M., Biassoni, R., Bottino, C., Vitale, M., Sivori, S., Augugliaro, R., Moretta, L., and Moretta, A. (1999) *J. Exp. Med.* **190**, 793–801
- Nicoll, G., Ni, J., Liu, D., Klenerman, P., Munday, J., Dubock, S., Mattei, M. G., and Crocker, P. R. (1999) *J. Biol. Chem.* **274**, 34089–34095
- Yamaji, T., Teranishi, T., Alphey, M. S., Crocker, P. R., and Hashimoto, Y. (2002) *J. Biol. Chem.* **277**, 6324–6332
- Zhang, J. Q., Nicoll, G., Jones, C., and Crocker, P. R. (2000) *J. Biol. Chem.* **275**, 22121–22126
- Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
- Navaza, J. (1994) *Acta Crystallogr. Sect. A* **50**, 157–163
- Perrakis, A., Morris, R., and Lamzin, V. S. (1999) *Nat. Struct. Biol.* **6**, 458–463
- Brunger, A. T., Adams, P. D., Clore, G. M., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **54**, 905–921
- Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
- Brunger, A. T. (1992) *Nature* **355**, 472–474
- Vriend, G. (1990) *J. Mol. Graph.* **8**, 52–56
- Fotinou, C., Emsley, P., Black, I., Ando, H., Ishida, H., Kiso, M., Sinha, K. A., Fairweather, N. F., and Isaacs, N. W. (2001) *J. Biol. Chem.* **276**, 32274–32281
- Laskowski, R. A., McArthur, M. W., Moss, D. S., and Thornton, J. M. (1993) *J. Appl. Cryst.* **26**, 283–291
- Stanley, P., Narasimhan, S., Siminovitch, L., and Schachter, H. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 3323–3327
- Puthalakath, H., Burke, J., and Gleeson, P. A. (1996) *J. Biol. Chem.* **271**, 27818–27822
- Richardson, J. S. (1981) *Adv. Protein Chem.* **34**, 167–339
- Brinkman-Van der Linden, E. C. M., Sjoberg, E. R., Juneja, L. R., Crocker, P. R., Varki, N., and Varki, A. (2000) *J. Biol. Chem.* **275**, 8633–8640
- Johnson, M. A., and Pinto, B. M. (2002) *Aust. J. Chem.* **55**, 13–25
- Houston, D. R., Shiomi, K., Arai, N., Omura, S., Peter, M. G., Turberg, A., Synstad, B., Eijssink, V. G. H., and van Aalten, D. M. F. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99**, 9127–9132
- Jain, D., Kaur, K., Sundaravadivel, B., and Salunke, D. M. (2000) *J. Biol. Chem.* **275**, 16098–16102
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) *Nucleic Acids Res.* **22**, 4673–4680
- Nicholls, A., Sharp, K., and Honig, B. (1991) *Proteins*, **11**, 281–296