Glycobiology vol. 10 no. 4 pp. 431–438, 2000

Siglec-7: a sialic acid-binding lectin of the immunoglobulin superfamily

Takashi Angata and Ajit Varki[1]

Glycobiology Research and Training Center, Department of Medicine and Cancer Center, University of California San Diego, La Jolla, CA 92093, USA

Received on October 6, 1999; revised on November 5, 1999; accepted on November 8, 1999

The Siglecs are a recently discovered family of sialic acid-binding lectins of the immunoglobulin (Ig) superfamily. We report a molecule showing homology to the six first reported Siglecs, with the closest relationship to Siglec-3(CD33), Siglec-5, and Siglec-6(OBBP-1). The extracellular portion has two Ig-like domains, with the amino-terminal V-set Ig domain including amino acid residues known to be involved in sialic acid recognition by other Siglecs. The cytoplasmic domain has putative sites of tyrosine phosphorylation shared with some Siglecs, including an Immunoreceptor Tyrosine-based Inhibitory Motif (ITIM). Expression of the full-length cDNA induces sialic acid-dependent binding to human erythrocytes. A recombinant chimeric form containing the extracellular Ig domains selectively recognizes the sequence Neu5Acα2–6Galβ1–4Glc, and binding requires the side chain of sialic acid. Mutation of an arginine residue predicted to be critical for sialic acid binding abolishes both interactions. Taken together, our findings justify designation of the molecule as Siglec-7. Analysis of bacterial artificial chromosome (BAC) clones spanning the known human genomic location of Siglec-3 indicates that the Siglec-7 gene is also located on chromosome 19q13.3–13.4. Human tissues show strong expression of Siglec-7 mRNA in spleen, peripheral blood leukocytes, and liver. The combination of an extracellular sialic acid binding site and an intracellular ITIM motif suggests that this molecule is involved in trans-membrane regulatory signaling reactions.

Key words: receptors/Siglecs/sialic acids/lectins/Ig superfamily

Introduction

Sialic acids (Sias) are typically attached to the outer end of cell surface and secreted glycoconjugates of higher animals (Schauer, 1982; Varki, 1992; Kelm and Schauer, 1997). They can repel intermolecular and cell–cell interactions by virtue of their negative charge, as well as provide binding targets for viruses, bacteria, parasites, and toxins (Varki, 1997; Karlsson, 1998). Their marked structural complexity with regard to internal modifications and linkages (Schauer, 1982; Varki, 1992; Ye *et al.*, 1994; Kelm and Schauer, 1997) was thought to be related mostly to roles in host–pathogen interactions (Gagneux and Varki, 1999). Indeed, a decade ago, the only example of a vertebrate Sia binding protein was factor H of the alternative complement pathway (Fearon, 1978; Pangburn and Muller-Eberhard, 1978). It is now clear that Sias can be recognized by many specific lectins expressed by the same organism that synthesizes them (Varki, 1997). Among these is a family of lectins called the Siglecs (Sialic acid-binding Ig-superfamily lectins) (Crocker *et al.*, 1998). The initially recognized members were Siglec-1(sialoadhesin, Sn, on macrophages), Siglec-2(CD22, on B cells), Siglec-3(CD33, on myelomonocytic cells), and Siglec-4a/4b (myelin-associated glycoprotein/ Schwann cell myelin protein, found on oligodendrocytes and Schwann cells in the nervous system) (Crocker *et al.*, 1998). Siglec-5 was subsequently discovered on neutrophils and monocytes (Cornish *et al.*, 1998) and Siglec-6/OBBP-1 on B cells and placental trophoblasts (Patel *et al.*, 1999). Each Siglec preferentially recognizes aspects of its sialylated ligands, such as the side chain, the N-acyl group, the linkage from the 2-position, and sometimes the underlying sugars (Powell and Varki, 1994; Sjoberg *et al.*, 1994; Powell and Varki, 1995; Powell *et al.*, 1995; Kelm *et al.*, 1996; Shi *et al.*, 1996; Collins *et al.*, 1997a,b; May *et al.*, 1998; Schnaar *et al.*, 1998; Crocker *et al.*, 1999; Sawada *et al.*, 1999). Siglec-6 also has a protein ligand, i.e., leptin (Patel *et al.*, 1999).

Assignment of specific functions to the Sia binding phenotype of the Siglecs is complicated, because these binding sites are often masked by endogenous ligands (Braesch-Andersen and Stamenkovic, 1994; Freeman *et al.*, 1995; Hanasaki *et al.*, 1995b; Sgroi *et al.*, 1995; Sgroi *et al.*, 1996; Collins *et al.*, 1997b; Tropak and Roder, 1997), and can be unmasked by sialidase treatment or cellular activation (Razi and Varki, 1998). Recently, we reported that human blood leukocytes have Sia binding sites that can be unmasked by sialidase treatment (Razi and Varki, 1999). Such unmasking occurred not only on B cells, monocytes and neutrophils (known to carry Siglecs) but also on natural killer cells and on a minority of mature T cells. The masking of such sites on unactivated cells may explain why many of these lectins have not been previously discovered. Here we report the discovery and characterization of a seventh member of the Siglec family. While this work was being prepared for submission, an apparent splice-variant of the same gene (p75/AIRM1) was independently reported by another group studying natural killer cells and also shown to bind erythrocytes in Sia-dependent manner (Falco *et al.*, 1999). We are also aware that yet another group has independently cloned the same gene product (P.Crocker, personal communication).

[1]To whom correspondence should be addressed

© 2000 Oxford University Press

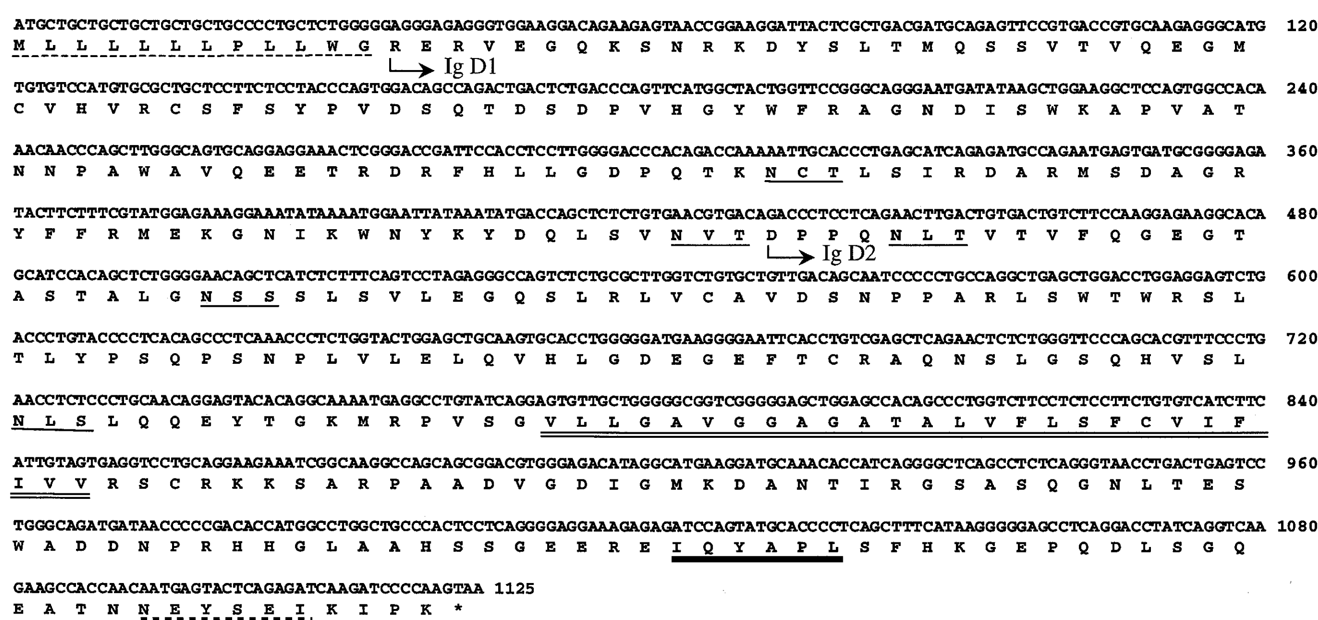

Fig. 1. Nucleotide and derived amino acid sequences of Siglec-7. The dotted line and the double line indicate the putative signal sequence and transmembrane sequence respectively. Potential sites of N-glycosylation are indicated with thin lines. The bold line indicates the potential ITIM motif, and the bold dotted line the motif similar to the SAP-binding site of SLAM. The amino-terminal ends of the two Ig domains are indicated.

Results

Identification of a cDNA clone homologous to Siglecs

The first 120 amino acid sequences of mouse Siglec-1 (GenBank accession #Z36293) and human Siglecs 2–6 (X59350, 4502654, M29273, 4502658, and 4502656, respectively) were used to carry out homology searches of the dbEST division of GenBank database. A human gall bladder cDNA (EST50650; GenBank accession #AA344713, The Institute for Genomic Research) was identified as a Siglec candidate, based on >50% homology to Siglecs 3, 5, and 6, and conservation of several amino acid residues common to previously known Siglecs (Cornish *et al.*, 1998; Patel *et al.*, 1999). This clone was obtained via ATCC, and sequenced.

Cloning of a full-length cDNA encoding the putative Siglec-7

Since the clone lacked a coding sequence for the first half of an Ig-like domain, 5′-RACE and RT-PCR were performed, using human PBMC total mRNA as a template (PBMC were chosen based on preliminary Northern analysis using the partial clone). The full-length clone contained an open reading frame encoding 374 amino acids, starting with a typical signal sequence and two Ig-like domains, followed by a probable transmembrane domain and cytosolic tail (Figure 1). The cytosolic tail of the full-length protein contains two tyrosine residues, of which the first is contained in an Immunoreceptor Tyrosine-based Inhibitory Motif (ITIM; S/I/L/VxYxxL/V). The second is in a motif [NEYSEI] similar, but not identical to the binding site on SLAM (Signaling Lymphocyte Activating Molecule) for SLAM-associated protein (SAP) (Sayos *et al.*, 1998). Nucleotide identity with other human Siglecs (first 750 nucleotides) are: Siglec-1, 45.9%; Siglec-2, 48.0%; Siglec-3, 63.6%; Siglec-4, 45.4%; Siglec-5, 64.0%; Siglec-6, 63.0%, and many of the residues typical of Siglecs are conserved (Figure 2a). Siglec-7 formed a closely related cluster with Siglecs 3, 5, and 6 on an unrooted phylogram (Figure 2b).

Sia-dependent erythrocyte-binding is abolished by conservative substitution of a single Arg residue

COS-7 cells were transfected with full-length Siglec-7 cDNA and interactions with human erythrocytes analyzed by a rosetting assay (Cornish *et al.*, 1998). Transfected COS cells bound human erythrocytes (Figure 3a), and sialidase pretreatment of the erythrocytes completely abrogated the interaction (Figure 3b), indicating that Siglec-7 recognizes sialylated glycans on the erythrocyte surface. Binding occurred even when the COS cells were not pretreated with sialidase to remove potential *cis*-ligands that can engage some Siglecs. A specific Arg residue in Siglecs 1–4 is known to be required for Sia recognition, and is conserved among all Siglecs, including Siglec-7 (Figure 2a). Introduction of conservative point-mutation (Arg124 to Lys) completely eliminated the binding of erythrocytes to transfected COS cells (Figure 3c,d).

Recombinant Siglec-7 recognizes a specific sialyl linkage and requires the Sia side chain

A fusion protein of extracellular domains of Siglec-7 and human IgG Fc tail was produced and used in an ELISA assay. Among the ligands used, Neu5Acα2–6Galβ1–4Glc showed the best binding (Figure 4). There was much weaker binding to probes carrying α2–3-linked Sias, including Neu5Acα2–3Galβ1–4Glc which differs only in the linkage of the Sia. Thus, Siglec-7 prefers α2–6-linked Sias. However, Neu5Acα2–6GalNAc also showed less binding, suggesting that the underlying glycan structure affects recognition. Sialidase treatment of the Neu5Acα2–6Galβ1–4Glc probe

a

```
Siglec7 MLLLLLLPLLWGRERVEGQKSNRKDYSLTMQSSVTVQEGMCVHVRCSFSYPVDSQTDSDPVHGYWFRAGNDISWKAPVATNNPAWAVQEETRDRFHLLGDPQTKNCTLSIRDARMSDAGR 120
Siglec3 MPLLLLLPLLWA-----GALAMDPNFWLQVQESVTVQEGLCVLVPCTFFHPIPYYDKNSPVHGYWFREGAIISGDSPVATNKLDQEVQEETQGRFRLLGDPSRNNCSLSIVDARRRDNGS 115
Siglec5 MLPLLLLPLLWG-----GSLQEKPVYELQVQKSVTVQEGLCVLVPCSFSYPWRSWYSSPPLYVYWFRDGEIPYYAEVVATNNPDRRVKPETQGRFRLLGDVQKKNCSLSIGDARMEDTGS 115
Siglec6 MLP-LLLPLLWA-----GALAQERRFQLEGPESLTVQEGLCVLVPCRLPTTLPASYYG---YGYWFLEG----ADVPVATNDPDEEVQEETRGRFHLLWDPRRKNCSLSIRDARRRDNAA 107
        *   *******      *       : *    *:*****:** * * :             : ***  *        ****     *: **: **:** *   :**:*** ***  *

Siglec7 YFFRMEKGN-IKWNYKYDQLSVNVTDPPQNLTVTVFQGEGTASTALGNSSSLSVLEGQSLRLVCAVDSNPPARLSWTWRSLTLYPSQPSNPLVLELQVHLGDEGEFTCRAQNSLGSQHVS 239
Siglec3 YFFRMERGS-TKYSYKSPQLSVHVTDLTHRPKILIPGTLEPGHSKNLTCSVSWACEQGTPPIFSWLSAAPTSLGPRTTHSSVLIITPRPQDHGTNLTCQVKFAGAGVTTERTIQLNVTYV 234
Siglec5 YFFRVERGRDVKYSYQQNKLNLEVTALIEKPDIHFLEPLESGRPTRLSCSLPGSCEAGPPLTFSWTGNALSPLDPETTRSSELTLTPRPEDHGTNLTCQMKRQGAQVTTERTVQLNVSYA 235
Siglec6 YFFRLKSKW-MKYGYTSSKLSVRVMALTHRPNISIPGTLESGHPSNLTCSVPWVCEQGTPPIFSWMSAAPTSLGPRTTQSSVLTITPRPQDHSTNLTCQVTFPGAGVTMERTIQLNVSYA 226
        ****::     *: *   :* : *        :                *     *                    * :*  *  :   :    :*  ::   *      :

Siglec7 LNLSLQQEYTG------------------------------------------------------------------------------------------------------------- 250
Siglec3 PQNPTTGIFPG-------------------------------------D----------------------------------------------------------------------- 246
Siglec5 PQTITIFRNGIALEILQN---TSYLPVLEGQALRLLCDAPSNPPAHLSWFQGSPALNATPISNTGILELRRVRSAEEGGFTCRAQHPLGFLQIFLNLSVYSLPQLLGPSCSWEAEGLHCR 352
Siglec6 PQKVAISIFQGNSAAFKILQNTSSLPVLEGQALRLLCDADGNPPAHLSWFQGFPALNATPISNTGVLELPQVGSAEEGDFTCRAQHPLGSLQISL------------------------- 321
         :

Siglec7 ----------------------------------------------------------------------------KMRPVSGVLLGAVGGAGATALVFLSFCVIFIVVRSCRKKSARPA 294
Siglec3 -------------------------------------------------------------------------GSGKQETRAGLVHGAIGGAGVTALLALCLCLIFFI-----VKTHRRK 288
Siglec5 CSFRARPAPSLCWRLEEKPLEGNSSQGSFKVNSSSAGPWANSSLILHGGLSSDLKVSCKAWNIYGSQSGSVLLLQGRSNLGTGVVPAALGGAGVMALLCICLCLIFFLI----VKARRKQ 468
Siglec6 ----------------------------------------------------------------------SLFVHWKPEGRAGGVLGAVWGASITTLVFLCVCFIFR------VKTRRKK 365
                                                                                    :    :* :  *: **   :*: :  *  **        *: *

Siglec7 ADVGDIGMKDANTIRGSASQGNLTESWAD--DNPRHHGLAAHSSGEEREIQYAPLSFHKGEPQDLSGQ-EATNNEYSEIKIPK 374
Siglec3 AARTAVGSNDTHPTTGSASPKHQKNSKLHGPTETSSCSGAAPTVEMDEELHYASLNFHGMNPS------KDTSTEYSEVRTQ- 364
Siglec5 AAGRPEKMDDEDPIMGTITSGSRKKPWPDSPGDQASPPGDAPPLEEQKELHYASLSFSEMKSREPKDQEAPSTTEYSEIKTSK 551
Siglec6 AAQPVQNTDDVNPVMVSGSRGHQHQFQTG--IVSDHPAEAGPISEDEQELHYAVLHFHKVQPQEP----KVTDTEYSEIKIHK 442
        *        *      : :     :                     : *::** * *  :           :  ****::
```

b

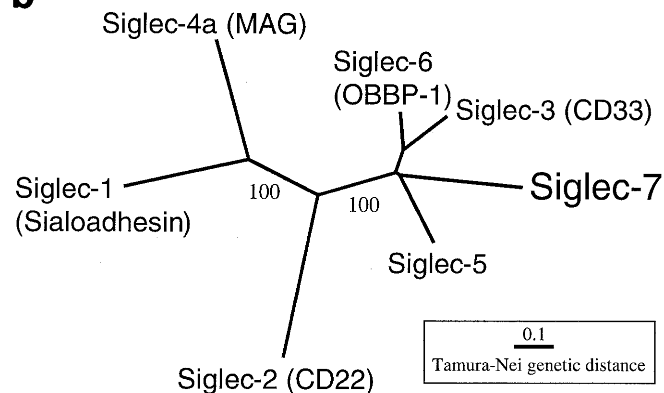

Fig. 2. Sequence relationships of Siglec-7 with other Siglecs. (**a**) Amino acid sequence alignment of the Siglec-7 with the most closely related Siglecs (3, 5, and 6) performed using Clustal W, with minimal manual adjustment. The * indicate identical residues, **:** indicates conservative differences. The conserved Arg residue that is mutated in this study is indicated with a box. (**b**) Phylogenetic analysis of the Siglec family using PAUP* 4.0 with results presented as an unrooted phylogram. Extent of divergence is indicated by the length of the lines between any two genes. Numbers near internodes indicate the Boot-strap values for 500 resamplings. Siglec-7 forms a clade with 3, 5, and 6 that appears to be distinct from the other three Siglec family members. A maximum likelihood-based reconstruction using Quartet Puzzling (Strimmer and von Haeseler, 1996) gave similar results.

abolishes the binding of Siglec-7 (Figure 5). Mild periodate treatment, which truncates the glycerol-like side chain of Sias and abolishes recognition by Siglecs 1–4 (Sjoberg *et al.*, 1994; Shi *et al.*, 1996; Collins *et al.*, 1997b) also reduced recognition by Siglec-7 (Figure 5), indicating that the Sia side chain is important for binding.

Tissue distribution of message expression

Northern blot analysis of normal human tissues (Figure 6) indicates high expression in PBMC, spleen, and liver, and less prominent expression in lung and placenta. While the major mRNA species is estimated to be about 1.7 kb, there seem to be a shorter isoform (~1.5 kb), notably in spleen. This result suggests that expression of isoforms may differ amongst tissues. The PBMC and splenic expression may be explained by the finding that NK cells carry this gene product (Falco *et al.*, 1999).

The Siglec-7 gene is on chromosome 19

The genes for human Siglecs 2–6 are known to be on chromosome 19, with Siglecs 3, 5, and 6 localized around 19q13.3–13.4 (Mucklow *et al.*, 1995; Takei *et al.*, 1997; Cornish *et al.*, 1998). From this fact and the high homology of Siglec-7 with the 3 Siglecs, we hypothesized a “Siglec gene cluster” around 19q13.3–13.4. Several BAC clones contiguously covering the chromosomal region containing the Siglec-3/CD33 gene were obtained, and screened for a clone containing Siglec-7 gene using PCR. Sequencing of PCR products revealed that a BAC clone which carries the Siglec-3/CD33 gene also contains the Siglec-7 gene.

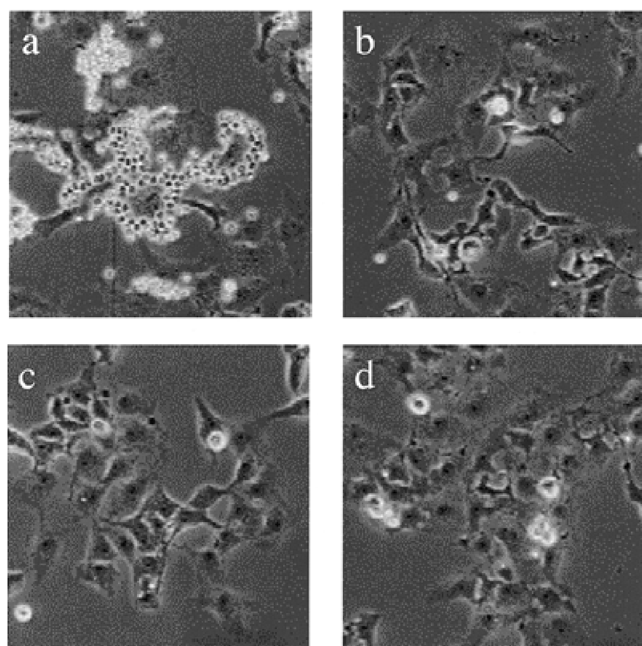

Fig. 3. Siglec-7 binds human erythrocytes in a Sia-dependent manner. COS-7 cells were transfected with the full-length cDNA for Siglec-7 (wild-type or R124K mutant) in pcDNA3.1 or sham-transfected with empty vector. After 48 h, binding of human erythrocytes was assessed (*Materials and methods*). Examples shown are: (**a**) wild-type and untreated erythrocytes; (**b**) wild-type and sialidase-treated erythrocytes; (**c**) R124K mutant and untreated erythrocytes; (**d**) R124K mutant and sialidase-treated erythrocytes.

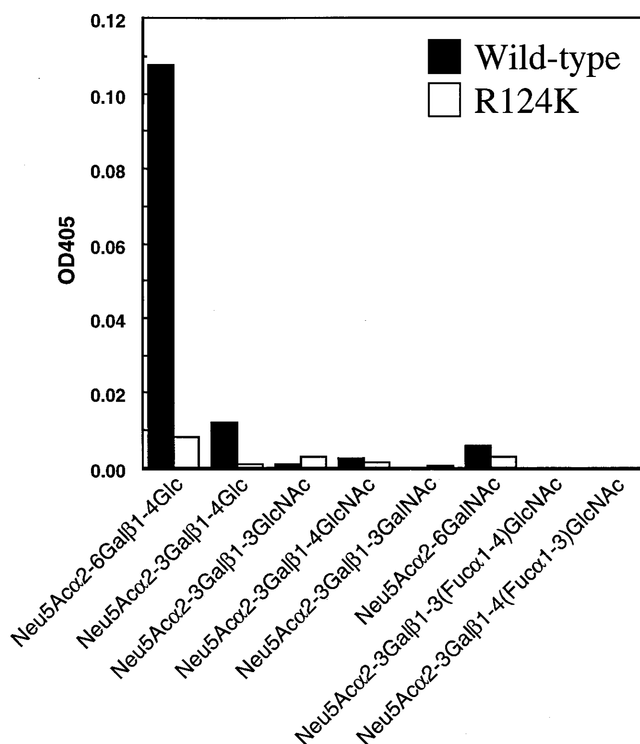

Fig. 4. Binding of sialylated ligands to the extracellular domain of Siglec-7. Siglec-7-Fc chimera was immobilized via protein A on a microtiter plate and binding of biotinylated polyacrylamide arrays conjugated with various sialyloligosaccharides was determined (*Materials and methods*). Data shown are mean values of triplicates.

Discussion

We have reported here a seventh member of the Siglec family, which shares the following features: a type 1 transmembrane topology with two or more Ig-like extracellular domains; an amino-terminal V-set Ig-like domain with characteristic sequence motifs and conserved cysteine residues; and specific recognition of Sia-containing glycans. Closer examination of the sequences indicates an ancestral relationship of the new gene with those of Siglecs 3, 5, and 6. In keeping with this, the four genes appear to be clustered around chromosome 19q13.3–13.4.

The Sia-binding properties of Siglec-7 are interesting. Among the potential targets studied, the preferred recognition motif is Neu5Acα2–6Galβ1–4Glc, which is very similar to the naturally occurring Siaα2–6Galβ1–4GlcNAc sequences of N-glycans in vertebrate systems. This motif is also recognized by Siglec-5 (Cornish *et al.*, 1998), and is the preferred ligand for Siglec-2 (Powell *et al.*, 1993; Powell and Varki, 1994). This may also explain why it was not necessary to treat Siglec-7 transfected COS cells with sialidase to unmask the binding site—COS cells have low endogenous levels of this structure. As with most other Siglecs, the actual binding affinity is likely to be low, and detection of specificity is achieved by the presentation of the cognate structures in multivalent arrays. While erythrocytes give robust Sia-dependent binding, this also represents an artificial situation. In the *in vivo* situation, erythrocytes are constantly bathed in plasma proteins that carry at least ~1 mM concentration of glycosidically linked Siaα2–6Galβ1–4GlcNAc sequences (Hanasaki *et al.*, 1995a). Thus, the location and nature of the natural sialylated ligands of Siglec-7 remain to be determined.

Site-directed mutagenesis studies of Siglec-1 and Siglec-2 identified many residues in the amino-terminal V-set domain that were suggested to be involved in recognition of sialylated ligands (Van der Merwe *et al.*, 1996; Vinson *et al.*, 1996). However, the subsequently obtained crystal structure of Siglec-1 in complex with sialyllactose indicated that most of these effects had been due to a general disruption of protein folding (May *et al.*, 1998). However, one Arg residue (R97) that was originally mutated and led to loss of binding (Crocker *et al.*, 1999) was in fact found to form a critical salt bridge with the carboxylate of the bound Sia (May *et al.*, 1998). Mutation of the corresponding Arg residue in Siglec-2 (Van der Merwe *et al.*, 1996), Siglec-3 (Taylor *et al.*, 1999), and Siglec-4a (Tang *et al.*, 1997) also resulted in loss of binding. This Arg residue is conserved among all the remaining Siglecs reported to date (see Figure 2a). We therefore conservatively mutated this residue in Siglec-7 to a Lys residue, giving a complete loss of binding to erythrocytes, and to semi-synthetic ligands. Thus, one can predict that this Arg residue is critical to the binding of Sias by all Siglecs. This also allows us to suggest an *in vivo* approach to explore the function of the sialic binding properties of Siglecs. Partial or wholesale deletion of a Siglec

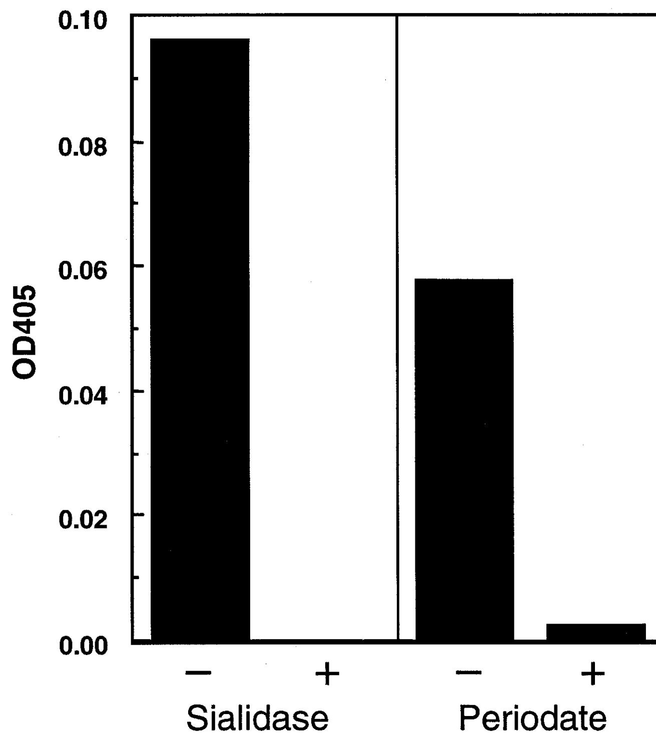

Fig. 5. Mild periodate and sialidase treatments abolish binding of sialylated ligands to the extracellular domain of Siglec-7. Siglec-7-Fc chimera to biotinylated polyacrylamide arrays carrying Siaα2–6Galβ1–4Glc was assayed as in Figure 4, except that aliquots of the probe were pretreated with mild periodate or sialidase (*Materials and methods*). Data shown are mean values of triplicates.

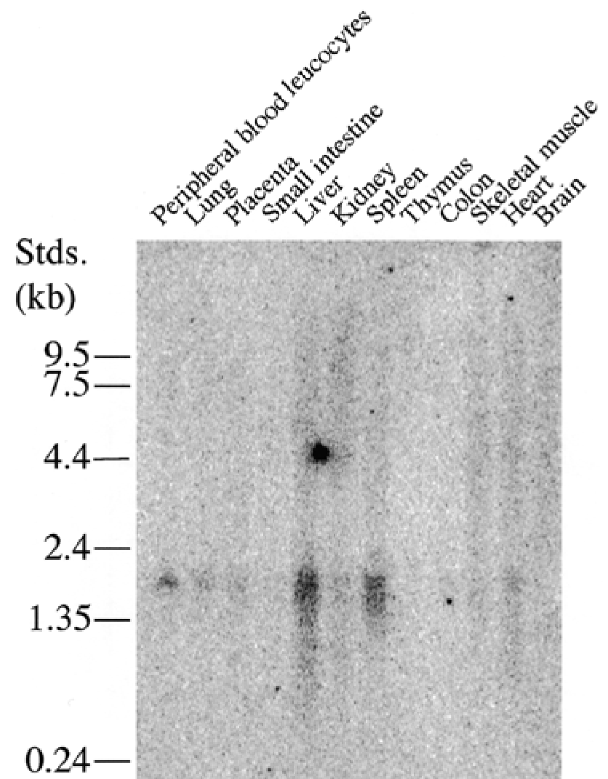

Fig. 6. Expression of messenger RNA for Siglec-7 in different human tissues. Northern blot analysis was carried out as described under *Materials and methods*.

molecule in the intact mouse is informative, but the results can be difficult to interpret, because more than just the sialic-acid binding property is eliminated. We suggest instead a "knock-in" strategy, where the only change made would be the mutation of the critical Arg residue of a given Siglec into a Lys. This would leave the Siglec with all of its domains intact, but lacking only its Sia binding site. Such an experiment should be more informative with regard to the function of Sia recognition.

Regarding the biological significance of Sia recognition, the best studied example is Siglec-2. Genetic elimination in the intact mouse gives a phenotype of B cell dysregulation (O'Keefe *et al.*, 1996; Otipoby *et al.*, 1996; Sato *et al.*, 1996; Cornall *et al.*, 1998). Genetic disruption of the ST6Gal I sialyltransferase that generates the CD22 ligand (Siaα2–6Galβ1–4GlcNAc) also gives suppression of B cell function, which is however more severe than that obtained by disruption of CD22 itself (Hennet *et al.*, 1998). This may now be explained by the fact that there are other Siglecs (3, 5, and 7) that can recognize the ST6Gal I product. The functions of Siglec-1 remain unclear, although it may be involved in macrophage interactions with developing myeloid cells (Crocker *et al.*, 1990) and/or cellular trafficking (Shi *et al.*, 1996). Regarding Siglec-4a, genetic disruption gives alterations in myelin sheath stability (Filbin, 1995), which correlates with the Wallerian degeneration seen upon genetic elimination of its cognate ganglioside ligands (Sheikh *et al.*, 1999).

The biological functions of the CD33-related Siglecs (3, 5, 6, and 7) remain to be elucidated. There is a conserved consensus sequence surrounding the intracellular tyrosine residues of all four of these molecules [E(I/L)xYAxL–(12–18 residues)–(T/N)EYSE(I/V)(K/R)] suggesting that they may associate with common intracellular signal transducers in their respective cell types. Indeed, as with Siglec-2 and 4a, the cytoplasmic tail of Siglec-3 is tyrosine-phosphorylated upon engagement (Taylor *et al.*, 1999). Once the tyrosines are phosphorylated, the first one is part of an ITIM (Immunoreceptor Tyrosine-based Inhibitory Motif; consensus: L/I/V/SxYxxL/V), which is found in many members of the Ig superfamily (including Siglec-2/CD22), and forms a potential docking site for the SHP-1 tyrosine phosphatase. Of note, the second tyrosine-containing motif [(T/N)EYSE(I/V)] is similar to the sequence (TxYxxI/V) that has been reported in SLAM (Sayos *et al.*, 1998), an immunoregulatory molecule of the Ig superfamily. This motif is the docking site in SLAM for an SH2-containing molecule called SAP (SLAM-associated protein), which blocks recruitment of the tyrosine phosphatase SHP-2 to its docking site in the SLAM cytoplasmic region. We have postulated that a similar interplay occurs between the recruitment of phosphatases to the ITIM motif in Siglecs and the presence of SAP or SAP-like inhibitors to the SAP binding motif (Patel *et al.*, 1999).

Given the combination of an extracellular Sia binding site and an intracellular ITIM motif, it is reasonable to suggest that Siglec-7 is involved in transmembrane regulatory signaling of the types mentioned above. In this regard, a paper has just

appeared reporting what appears to be the identical gene (the major cDNA reported has an additional extracellular C2-set Ig domain) (Falco *et al.*, 1999). These authors isolated this cDNA based on a monoclonal antibody directed against human natural killer cells, and showed that the protein negatively regulates the natural killer cell activity upon cross-linking by antibody. They also showed that the protein, when expressed on COS-7 cell, binds erythrocytes in Sia-dependent manner. While the relationship between Sia-binding at extracellular domain and signal transduction via intracellular domain of Siglec-7 is yet the be established, our recent finding that Sia binding sites on peripheral blood natural killer cells are masked by endogenous Sias (Razi and Varki, 1999) is of note in this context. Thus, if Siglec-7 indeed serves as an inhibitory receptor on natural killer cells, the Sias that regulate the signaling could be either on the same cell surface as the Siglec, or on another cell type. It also remains to be seen if Siglec-7 is expressed and functionally important on other cell types outside the hematopoietic system.

Materials and methods

5′-Rapid amplification of cDNA end (5′-RACE)

Total RNA from human peripheral blood mononuclear cells (PBMC) was obtained using an RNeasy Midi Kit (Qiagen). First strand cDNA was synthesized by reverse transcription (Superscript II, Life Technologies) of 0.5 µg of PBMC total RNA with a gene-specific primer SP1 (5′-GCCTTCTCCTTG-GAAGACAG-3′), and poly(dA) was added with dATP and terminal deoxynucleotidyl transferase. The cDNA was purified by QiaExII (Qiagen) and subjected to PCR. Gene-specific fragments were amplified by two rounds of nested PCR. In the first round, PS-NXB-dT (5′-GCTAGCTCGAGGATCCT$_{18}$-3′) and SP2 (5′-ACAGTCAAGTTCTGAGGAGG-3′) were used as primers. The second round PCR used the first round PCR product as a template, and PS-NXB (5′-TTTGCTAGCTCGA-GGATCCT$_4$-3′) and SP3 (5′-TCACGTTCACAGAGAGCT-GG-3′) as primers. The PCR product from the second round was directly sequenced.

Cloning and mutation of a full-length cDNA for Siglec-7

RT-PCR was performed using primers based on the DNA sequence of the EST clone and 5′ RACE product. First strand cDNA was synthesized by reverse transcription of 1 µg of PBMC total RNA with random DNA hexamers as primer. The reaction product was subjected to two rounds of PCR using *Pwo* DNA polymerase (Roche), using SLX-5′UTR (5′-CTCG-GATCCCTGGCACCTCTAACCC-3′) and SLX-3′UTR (5′-GGTCTAGAACCCTCAAACAAGCCC-3′) as primers. The PCR products were digested with *Bam*HI and *Xba*I, ligated to *Bam*HI–*Xba*I sites of pBluescript II KS(-) (Stratagene) and sequenced. A point mutation converting an Arg residue (R124) to Lys was introduced using QuickChange Site-Directed Mutagenesis Kit (Stratagene), following the manufacturer's protocol.

Northern blot analysis

A human 12-lane Multiple Tissue Nothern Blot (Clontech) was probed with the *Nco*I–*Xho*I fragment of the EST clone labeled with the Strip-EZ DNA kit (Ambion) and [α-^{32}P]dATP (NEN). Hybridization signals were visualized using a Phosphor Imager.

Phylogenetic analysis of Siglecs

DNA sequences of human Siglecs 1–7 encoding the first two Ig-like domains (750 nt; the human Siglec-1 sequence was kindly provided by P.Crocker, Dundee, UK) were aligned by Clustal W at the European Bioinformatics Institute web site, and also subjected to phylogenetic analysis using PAUP* 4.0 (Sinauer Associates). The phylogenetic tree was constructed using the neighbor joining method (Saitou and Nei, 1987). The distance matrix was based on Tamura-Nei genetic distances (Tamura and Nei, 1993).

Expression of full-length Siglec-7 on COS-7 cells and erythrocyte rosetting

The full-length coding sequence of Siglec-7 (wild-type and R124K mutant) was amplified by PCR using the primers SLX 5′Chi (5′-CCTGTCGACGCCACCATGCTGCTGCTGCTG-CTGCTGCCC-3′) and SLX 3′UTR. The PCR fragment was treated with *Sal*I and subcloned into the *Xho*I–*Eco*RV sites of pcDNA3.1(-) (Invitrogen), and sequenced. Constructs were transfected using LipofectAMINE reagent (Life Technologies) into COS-7 cells. After 48 h, cells were washed twice with PBS, treated with or without 10 mU *Arthrobacter ureafaciens* sialidase at 37°C for 1 h, and washed three times with rosetting assay solution (DMEM with 0.25% bovine serum albumin). Human erythrocytes (0.25% v/v; also pretreated with or without sialidase) in rosetting assay solution were layered on the COS-7 cells and incubated at 37°C for 30 min. Unbound erythrocytes were gently washed away, and rosettes observed under a microscope.

Production of recombinant chimeric proteins (Siglec-7-Fc)

A DNA fragment of Siglec-7 (wild-type and R124K) encoding the first two Ig-like domains was amplified by PCR using SLX 5′Chi and SLX 3′Chi (5′-CCTCATTTTGCCTGTGTACT-CCTG-3′) as primers. The fragment was cloned into the expression vector EK-Fc-pEdDC (prepared in this laboratory by Hui-ling Han), giving rise to a fusion protein of Siglec-7 extracellular domains and a human IgG Fc tail with a FLAG epitope tag (DYKDDDDK) in between. The constructs were transfected using LipofectAMINE into COS-7 cells or CHO-TAg cells, culture supernatants collected, and the chimeric proteins purified on Protein A–Sepharose (Amersham Pharmacia Biotech).

Binding specificity of Siglec-7

Binding to sialylated oligosaccharides on polyacrylamide arrays (Glycotech) was performed as described (Patel *et al.*, 1999). Briefly, microtiter plate wells (Nunc, catalog #269620) were coated with protein A (0.5 µg/well) in 50 mM sodium carbonate-bicarbonate buffer, pH 9.5, at 4°C overnight, washed three times with ELISA buffer (20 mM HEPES, 1% bovine serum albumin, 125 mM NaCl, 1 mM EDTA, pH 7.45), blocked with ELISA buffer (RT, 1 h), and sequentially incubated at RT with the following (each incubation followed by 3 washes with ELISA buffer): Siglec-7-Fc (0.5 µg/well, human gamma globulin as negative control), 2 h; probes (1 µg/well), 2 h; streptavidin-conjugated alkaline phosphatase (1/1000 diluted from stock solution, Life Technologies), 1 h. After a

final wash, *p*-Nitrophenyl Phosphate Liquid Substrate (Sigma) was added, incubated at RT for varying times (product formation was linear up to 18 h), and absorbance measured at 405 nm.

Mild periodate treatment of probes

The polyacrylamide probe carrying Neu5Acα2–6Galβ1–4Glc was subjected to mild periodate treatment, to truncate the glycerol-like side chain of Sia (Van Lenten and Ashwell, 1971). Typically, 10 µg of the probe was incubated in 100 µl of 2 mM $NaIO_4$ in PBS for 30 min on ice in the dark. After the incubation, 100 µl of 20 mM $NaBH_4$ in PBS was added to the mixture and further incubated at RT for 1 h in the dark, to reduce aldehydes generated by periodate treatment. The mixture was diluted with 800 µl ELISA buffer and directly used in the assay.

Chromosomal localization of Siglec-7

From the physical map of human chromosome 19 (Human Genome Center, Laurence Livermore National Laboratory), BAC clones contiguously covering the region of chromosome 19 containing the Siglec-3/CD33 gene were identified and obtained through Research Genetics. The BACs were prepared by a modified alkaline lysis method (BACPAC Resources, Rosewell Park Cancer Institute), and subjected to PCR using SP1 and SP-5′ (5′-AAACTCGGGACCGATTCCAC-3′) as primers. PCR products were purified and directly sequenced.

Acknowledgments

We thank Hiromu Takematsu and Els Brinkman-Van der Linden for general advice, and Pascal Gagneux for help with the PAUP analysis. This work was supported by USPHS grant RO1-GM32373. T.A. was partially supported by the Naito Foundation (Tokyo, Japan). Data deposition: sequence data reported here is deposited in GenBank (AF193441).

Abbreviations

Neu5Ac, N-acetyl-neuraminic acid; Sia, sialic acid, type unspecified; Sn, sialoadhesin; MAG, myelin-associated glycoprotein; PBMCs, peripheral blood mononuclear cells; 5′-RACE, 5′-rapid amplification of cDNA end; RT, room temperature; BAC, bacterial artificial chromosome; ITIM, Immunoreceptor Tyrosine-based Inhibitory Motif; SLAM, Signaling Lymphocyte Activating Molecule; and SAP, SLAM-associated protein.

Note added in proof

After the acceptance of this manuscript, a paper describing a 3-domain form of Siglec-7 appeared: Nicoll,G., Ni,J., Liu,D., Klenerman,P., Munday,J., Dubock,S., Mattei,M.G. and Crocker,P.R. (1999) Identification and characterization of a novel siglec, siglec-7, expressed by human natural killer cells and monocytes. *J. Biol. Chem.*, **274**, 31441–31447.

References

Braesch-Andersen,S. and Stamenkovic,I. (1994) Sialylation of the B lymphocyte molecule CD22 by α2,6-sialyltransferase is implicated in the regulation of CD22-mediated adhesion. *J. Biol. Chem.*, **269**, 11783–11786.

Collins,B.E., Kiso,M., Hasegawa,A., Tropak,M.B., Roder,J.C., Crocker,P.R. and Schnaar,R.L. (1997a) Binding specificities of the sialoadhesin family of I-type lectins—sialic acid linkage and substructure requirements for binding of myelin-associated glycoprotein, Schwann cell myelin protein and sialoadhesin. *J. Biol. Chem.*, **272**, 16889–16895.

Collins,B.E., Yang,L.J.S., Mukhopadhyay,G., Filbin,M.T., Kiso,M., Hasegawa,A. and Schnaar,R.L. (1997b) Sialic acid specificity of myelin-associated glycoprotein binding. *J. Biol. Chem.*, **272**, 1248–1255.

Cornall,R.J., Cyster,J.G., Hibbs,M.L., Dunn,A.R., Otipoby,K.L., Clark,E.A. and Goodnow,C.C. (1998) Polygenic autoimmune traits: Lyn, CD22 and SHP-1 are limiting elements of a biochemical pathway regulating BCR signaling and selection. *Immunity*, **8**, 497–508.

Cornish,A.L., Freeman,S., Forbes,G., Ni,J., Zhang,M., Cepeda,M., Gentz,R., Augustus,M., Carter,K.C. and Crocker,P.R. (1998) Characterization of siglec-**5**, a novel glycoprotein expressed on myeloid cells related to CD33. *Blood*, **92**, 2123–2132.

Crocker,P.R., Werb,Z., Gordon,S. and Bainton,D.F. (1990) Ultrastructural localization of a macrophage-restricted sialic acid binding hemagglutinin, SER, in macrophage-hematopoietic cell clusters. *Blood*, **76**, 1131–1138.

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) Siglecs: a family of sialic-acid binding lectins [letter]. *Glycobiology*, **8**, v.

Crocker,P.R., Vinson,M., Kelm,S. and Drickamer,K. (1999) Molecular analysis of sialoside binding to sialoadhesin by NMR and site-directed mutagenesis. *Biochem. J.*, **341**, 355–361.

Falco,M., Biassoni,R., Bottino,C., Vitale,M., Sivori,S., Augugliaro,R., Moretta,L. and Moretta,A. (1999) Identification and molecular cloning of p75/AIRM1, a novel member of the sialoadhesin family that functions as an inhibitory receptor in human natural killer cells. *J. Exp. Med.*, **190**, 793–801.

Fearon,D.T. (1978) Regulation by membrane sialic acid of beta1H-dependent decay-dissociation of amplification C3 convertase of the alternative complement pathway. *Proc. Natl. Acad. Sci. USA*, **75**, 1971–1975.

Filbin,M.T. (1995) Myelin-associated glycoprotein: A role in myelination and in the inhibition of axonal regeneration. *Curr. Opin. Neurobiol.*, **5**, 588–595.

Freeman,S.D., Kelm,S., Barber,E.K. and Crocker,P.R. (1995) Characterization of CD33 as a new member of the sialoadhesin family of cellular interaction molecules. *Blood*, **85**, 2005–2012.

Gagneux,P. and Varki,A. (1999) Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology*, **9**, 747–755.

Hanasaki,K., Powell,L.D. and Varki,A. (1995a) Binding of human plasma sialoglycoproteins by the B cell-specific lectin CD22. Selective recognition of immunoglobulin M and haptoglobin. *J. Biol. Chem.*, **270**, 7543–7550.

Hanasaki,K., Varki,A. and Powell,L.D. (1995b) CD22-mediated cell adhesion to cytokine-activated human endothelial cells. Positive and negative regulation by α2–6-sialylation of cellular glycoproteins. *J. Biol. Chem.*, **270**, 7533–7542.

Hennet,T., Chui,D., Paulson,J.C. and Marth,J.D. (1998) Immune regulation by the ST6Gal sialyltransferase. *Proc. Natl. Acad. Sci. USA*, **95**, 4504–4509.

Karlsson,K.A. (1998) Meaning and therapeutic potential of microbial recognition of host glycoconjugates. *Mol. Microbiol.*, **29**, 1–11.

Kelm,S., Schauer,R. and Crocker,P.R. (1996) The sialoadhesins—a family of sialic acid-dependent cellular recognition molecules within the immunoglobulin superfamily. *Glycoconjugate J.*, **13**, 913–926.

Kelm,S. and Schauer,R. (1997) Sialic acids in molecular and cellular interactions. *Int. Rev. Cytol.*,**175**, 137–240.

May,A.P., Robinson,R.C., Vinson,M., Crocker,P.R. and Jones,E.Y. (1998) Crystal structure of the N-terminal domain of sialoadhesin in complex with 3′ sialyllactose at 1.85 A resolution. *Mol. Cell*, **1**, 719–728.

Mucklow,S., Hartnell,A., Mattei,M.-G., Gordon,S. and Crocker,P.R. (1995) Sialoadhesin (*Sn*) maps to mouse chromosome 2 and human chromosome 20 and is not linked to the other members of the sialoadhesin family, CD22, MAG and CD33. *Genomics*, **28**, 344–346.

O'Keefe,T.L., Williams,G.T., Davies,S.L. and Neuberger,M.S. (1996) Hyperresponsive B cells in CD22-Deficient mice. *Science*, **274**, 798–801.

Otipoby,K.L. Andersson,K.B., Draves,K.E., Klaus,S.J., Farr,A.G., Kerner,J.D., Perlmutter,R.M., Law,C.L. and Clark,E.A. (1996) CD22 regulates thymus-independent responses and the lifespan of B cells. *Nature*, **384**, 634–637.

Pangburn,M.K. and Muller-Eberhard,H.J. (1978) Complement C3 convertase: cell surface restriction of beta1H control and generation of restriction on neuraminidase-treated cells. *Proc. Natl. Acad. Sci. USA*, **75**, 2416–2420.

Patel,N., Brinkman-Van der Linden,E.C.M., Altmann,S.W., Gish,K., Balasubramanian,S., Timans,J.C., Peterson,D., Bell,M.P., Bazan,J.F., Varki,A. and Kastelein,R.A. (1999) OB-BP1/Siglec-6—a leptin- and sialic acid-binding protein of the immunoglobulin superfamily. *J. Biol. Chem.*, **274**, 22729–22738.

Powell,L.D. and Varki,A. (1994) The oligosaccharide binding specificities of CD22β, a sialic acid-specific lectin of B cells. *J. Biol. Chem.*, **269**, 10628–10636.

Powell,L.D., Sgroi,D., Sjoberg,E.R., Stamenkovic,I. and Varki,A. (1993) Natural ligands of the B cell adhesion molecule CD22β carry *N*-linked oligosaccharides with α-2,6-linked sialic acids that are required for recognition. *J. Biol. Chem.*, **268**, 7019–7027.

Powell,L.D., Jain,R.K., Matta,K.L., Sabesan,S. and Varki,A. (1995) Characterization of sialyloligosaccharide binding by recombinant soluble and native cell-associated CD22. Evidence for a minimal structural recognition motif and the potential importance of multisite binding. *J. Biol. Chem.*, **270**, 7523–7532.

Powell,L.D. and Varki,A. (1995) I-Type lectins. *J. Biol. Chem.*, **270**, 14243–14246.

Razi,N. and Varki,A. (1998) Masking and unmasking of the sialic acid-binding lectin activity of CD22 (Siglec-2) on B lymphocytes. *Proc. Natl. Acad. Sci. USA*, **95**, 7469–7474.

Razi,N. and Varki,A. (1999) Cryptic sialic acid binding lectins in blood leukocytes can be unmasked by sialidase treatment of cellular activation. *Glycobiology*, **9**, 1225–1234.

Saitou,N. and Nei,M. (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.

Sato,S., Miller,A.S., Inaoki,M., Bock,C.B., Jansen,P.J., Tang,M.L.K. and Tedder,T.F. (1996) CD22 is both a positive and negative regulator of B lymphocyte antigen receptor signal transduction: altered signaling in CD22-deficient mice. *Immunity*, **5**, 551–562.

Sawada,N., Ishida,H., Collins,B.E., Schnaar,R.L. and Kiso,M. (1999) Ganglioside GD1α analogues as high-affinity ligands for myelin-associated glycoprotein (MAG). *Carbohydr. Res.*, **316**, 1–5.

Sayos,J., Wu,C., Morra,M., Wang,N., Zhang,X., Allen,D., van Schaik,S., Notarangelo,L., Geha,R., Roncarolo,M.G., Oettgen,H., de Vries,J.E., Aversa,G. and Terhorst,C. (1998) The X-linked lymphoproliferative-disease gene product SAP regulates signals induced through the co-receptor SLAM. *Nature*, **395**, 462–469.

Schauer,R. (1982) *Sialic Acids: Chemistry, Metabolism and Function, Cell Biology Monographs, Volume 10.* Springer-Verlag, New York.

Schnaar,R.L., Collins,B.E., Wright,L.P., Kiso,M., Tropak,M.B., Roder,J.C. and Crocker,P.R. (1998) Myelin-associated glycoprotein binding to gangliosides—structural specificity and functional implications. *Ann. N.Y. Acad. Sci.*, **845**, 92–105.

Sgroi,D., Koretzky,G.A. and Stamenkovic,I. (1995) Regulation of CD45 engagement by the B-cell receptor CD22. *Proc. Natl. Acad. Sci. USA*, **92**, 4026–4030.

Sgroi,D., Nocks,A. and Stamenkovic,I. (1996) A single *N*-linked glycosylation site is implicated in the regulation of ligand recognition by the I-type lectins CD22 and CR33. *J. Biol. Chem.*, **271**, 18803–18809.

Sheikh,K.A., Sun,J., Liu,Y.J., Kawai,H., Crawford,T.O., Proia,R.L., Griffin,J.W. and Schnaar,R.L. (1999) Mice lacking complex gangliosides develop Wallerian degeneration and myelination defects. *Proc. Natl. Acad. Sci. USA*, **96**, 7532–7537.

Shi,W.X., Chammas,R., Varki,N.M., Powell,L. and Varki,A. (1996) Sialic acid 9-*O*-acetylation on murine erythroleukemia cells affects complement activation, binding to I-type lectins and tissue homing. *J. Biol. Chem.*, **271**, 31526–31532.

Sjoberg,E.R., Powell,L.D., Klein,A. and Varki,A. (1994) Natural ligands of the B cell adhesion molecule CD22β can be masked by 9-*O*-acetylation of sialic acids. *J. Cell Biol.*, **126**, 549–562.

Strimmer,K. and von Haeseler, A (1996) Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Mol. Biol. Evol.*, **13**, 964–969.

Takei,Y., Sasaki,S., Fujiwara,T., Takahashi,E., Muto,T. and Nakamura,Y. (1997) Molecular cloning of a novel gene similar to myeloid antigen CD33 and its specific expression in placenta. *Cytogenet. Cell Genet.*, **78**, 295–300.

Tamura,K. and Nei,M., (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.*, **10**, 512–526.

Tang,S., Shen,Y.J., DeBellard,M.E., Mukhopadhyay,G., Salzer,J.L., Crocker,P.R. and Filbin,M.T. (1997) Myelin-associated glycoprotein interacts with neurons via a sialic acid binding site at ARG118 and a distinct neurite inhibition site. *J. Cell Biol.*, **138**, 1355–1366.

Taylor,V.C., Buckley,C.D., Douglas,M., Cody,A.J., Simmons,D.L. and Freeman,S.D. (1999) The myeloid-specific sialic acid-binding receptor, CD**33**, associates with the protein-tyrosine phosphatases, SHP-1 and SHP-2. *J. Biol. Chem.*, **274**, 11505–11512.

Tropak,M.B. and Roder,J.C. (1997) Regulation of myelin-associated glycoprotein binding by sialylated *cis*-ligands. *J. Neurochem.*, **68**, 1753–1763.

Van der Merwe,P.A., Crocker,P.R., Vinson,M., Barclay,A.N., Schauer,R. and Kelm,S. (1996) Localization of the putative sialic acid-binding site on the immunoglobulin superfamily cell-surface molecule CD22. *J. Biol. Chem.*, **271**, 9273–9280.

Van Lenten,L. and Ashwell,G. (1971) Studies on the chemical and enzymatic modification of glycoproteins. A general method for the tritiation of sialic acid-containing glycoproteins. *J. Biol. Chem.*, **246**, 1889–1894.

Varki,A. (1992) Diversity in the sialic acids. *Glycobiology*, **2**, 25–40.

Varki,A. (1997) Sialic acids as ligands in recognition phenomena. *FASEB J.*, **11**, 248–255.

Vinson,M., Van der Merwe,P.A., Kelm,S., May,A., Jones,E.Y. and Crocker,P.R. (1996) Characterization of the sialic acid-binding site in sialoadhesin by site-directed mutagenesis. *J. Biol. Chem.*, **271**, 9267–9272.

Ye,J., Kitajima,K., Inoue,Y., Inoue,S. and Troy,F.A.,II. (1994) Identification of polysialic acids in glycoconjugates. *Methods Enzymol.*, **230**, 460–484.