Characterization of Siglec-5, a Novel Glycoprotein Expressed on Myeloid Cells Related to CD33

By Ann L. Cornish, Sylvie Freeman, Gareth Forbes, Jian Ni, Mei Zhang, Mario Cepeda, Reiner Gentz, Meena Augustus, Kenneth C. Carter, and Paul R. Crocker

We describe the characterization of siglec-5 (sialic acid-binding Ig-like lectin-5), a novel transmembrane member of the immunoglobulin superfamily, highly related to the myeloid antigen, CD33. A full-length cDNA encoding siglec-5 was isolated from a human activated monocyte cDNA library. Sequencing predicted that siglec-5 contains four extracellular immunoglobulin-like domains, the N-terminal two of which are 57% identical to the corresponding region of CD33. The cytoplasmic tail is also related to that of CD33, containing two tyrosine residues embodied in immunoreceptor tyrosine-based inhibitory motif-like motifs. The siglec-5 gene was shown to map to chromosome 19q13.41-43, closely linked to the CD33 gene. When siglec-5 was expressed on COS cells or as a recombinant protein fused to the Fc region of human IgG1, it was able to mediate sialic acid–dependent binding to human erythrocytes and soluble glycoconjugates, suggesting that it may be involved in cell-cell interactions. By using specific antibodies, siglec-5 was found to have an expression pattern distinct from that of CD33, being present at relatively high levels on neutrophils but absent from leukemic cell lines representing early stages of myelomonocytic differentiation. Western blot analysis of neutrophil lysates indicated that siglec-5 exists as a disulfide-linked dimer of approximately 140 kD.

© 1998 by The American Society of Hematology.

APPROXIMATELY 34% of the cell surface glycoproteins identified on leukocytes are members of the immunoglobulin (Ig) superfamily.[1] Proteins of this type mediate a broad range of functions in cellular recognition often involving protein-protein interactions. Recent work in our and other laboratories has shown the existence of a novel subset of structurally related Ig superfamily proteins that mediate protein-carbohydrate interactions, specifically interacting with sialic acids in glycoproteins and glycolipids.[2] The group currently comprises sialoadhesin, CD22, myelin-associated glycoprotein (MAG), Schwann cell myelin protein (SMP), and CD33. Each of these proteins is expressed by specific cell types and involved in distinct functions, such as regulation of B-cell activation (CD22),[3] cell-cell interactions of macrophages (sialoadhesin),[4] and maintenance of myelin-axon interactions (MAG).[5] Each protein has a particular preference for both the nature of sialic acid and its glycosidic linkage to adjacent sugars.[6,7] Recently, these proteins have been designated "siglecs" (sialic acid-binding Ig-like lectins) and, although there has been no proposal to change their existing names, sialoadhesin, CD22, CD33, MAG, and SMP have been classified as siglecs 1, 2, 3, 4a, and 4b, respectively.[8]

Each of the siglecs has an N-terminal V-set domain followed by varying numbers of C2-set domains ranging from 1 in CD33 to 16 in sialoadhesin. Mutagenesis and structural studies have established that the sialic acid binding site of sialoadhesin is on the V-set domain, with key interactions between sialic acid substituents and amino acid side-chains of the G, F, and A β-strands.[9,10] In particular, a conserved arginine on the F-strand of sialoadhesin's V-set domain has been shown by X-ray crystallography to form a salt-bridge with the carboxylate of sialic acid,[10] and site-directed mutagenesis studies with sialoadhesin, CD22, MAG, and CD33 have shown that the corresponding arginine is essential for sialic acid–dependent binding by all siglecs.[9,11-13] Other important interactions shown by X-ray crystallography include hydrophobic contacts between conserved aromatic amino acids on the A and G β-strands of the V-set domain with the N-acetyl and glycerol side groups of N-acetyl neuraminic acid.[10] All siglecs so far characterized also have an unusual arrangement of conserved cysteine residues in the V-set and adjacent C2-set domains.[2] These are predicted to result in the formation of an intrasheet disulfide bond in the V-set domain and an interdomain disulfide bond.[10,14,15]

Apart from MAG and SMP, the siglecs characterized to date are expressed uniquely within the hematopoietic system and have been used extensively as lineage-restricted markers. For example, CD33 is a useful marker of early myeloid progenitor cells, being absent from pluripotential hematopoietic stem cells.[16] Likewise, CD22 is a specific marker of B cells,[17] and sialoadhesin is expressed uniquely on macrophage subsets.[18,19] In the cases of CD33 and CD22, these restricted expression patterns are conserved in myeloid[20] and B lymphoid[21] leukemias, respectively, and their presence can provide useful diagnostic markers as well as targets for immunotherapy, for example using toxin-conjugated or radiolabeled monoclonal antibodies (MoAbs).[22,23]

In this study we describe the characterization of a human cDNA that encodes a new member of the siglec family, designated siglec-5. This protein exhibits a high degree of sequence similarity with CD33 and contains all of the key structural features of siglecs described above. As predicted, siglec-5 is able to mediate sialic acid–dependent binding to cells either when expressed on COS cells or as a recombinant protein immobilized on plastic. By using polyclonal antisera raised to the extracellular region of siglec-5, we show that, among human blood leukocytes and leukemic cell lines, siglec-5 is expressed

in a myeloid-restricted manner that is distinct from that of CD33.

MATERIALS AND METHODS

Materials. Unless indicated otherwise, all reagents and chemicals were purchased from Sigma (Poole, UK or St Louis, MO). Protein A Sepharose and Factor Xa were purchased from Pharmacia (St Albans, UK), *Vibrio cholerae* sialidase was purchased from Calbiochem (La Jolla, CA), and ^{125}I-streptavidin (20-40 mCi/mg) and ^{125}I-sheep anti-mouse IgG (750-3,000 Ci/mmol) were purchased from Amersham (Little Chalfont, UK). Biotinylated polyacrylamide glycoconjugates carrying either NeuAcα2,3Galβ1,4Glc (2,3-PAA) or NeuAcα2,6Galβ1,4Glc (2,6-PAA) were purchased from Syntesome (Munich, Germany). Digoxygenin 11-dUTP, 4′-6-diamidino-2-phenylindole-2 HCl (DAPI) and anti-digoxygenin–tagged rhodamine were purchased from Boehringer Mannheim (Indianapolis, IN). Anti-CD33 MoAb LeuM9 was from Becton Dickinson (Abingdon, UK). Immulon 3 microtiter plates were from Dynatech Laboratories Inc (Chantilly, VA). NUNC tissue culture plastics were purchased from Life Sciences (Paisley, UK). The pIGplus plasmid was purchased from R & D Systems (Abingdon, UK) and pcDNA3 plasmid was from British Biotechnology (Oxford, UK).

Identification and characterization of siglec-5 cDNA. By using the amino acid sequence of CD33, a specific homology search was performed against a database containing more than one million expression sequence tags (ESTs) obtained from over 700 different cDNA libraries.[24-26] A full-length clone in pBluescript II (SK⁻) encoding one of the CD33-like sequences identified in the search was isolated from a human activated monocyte library and designated pHMQCD14. A computer search of nucleotide and protein sequence was performed by using the Blast GeneSearch (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD). Manipulation of sequences and alignments were performed by using Baylor College of Medicine molecular biology software available on the internet (Human Genome Center, Baylor College of Medicine, Houston, TX).

Northern blot analysis. The cDNA insert from pHMQCD14 was labeled with ^{32}P by using the Rediprime DNA labeling system from Amersham Life Sciences according to the manufacturer's instructions. Unincorporated nucleotides were removed from labeled probe by using CHROMA SPIN-100 columns (Clontech, Palo Alto, CA). Two human multiple tissue Northern (MTN) blots containing approximately 2 μg of poly (A)⁺ RNA per lane from various human tissues were purchased from Clontech and hybridized according to the manufacturer's instructions.

Chromosomal mapping by fluorescence in situ hybridization (FISH). The chromosomal localization of the siglec-5 gene was determined by single-copy gene FISH to human male metaphase chromosome spreads. Briefly, a 2-kb HMQCD14 cDNA insert was nick-translated by using digoxygenin 11-dUTP, and FISH was performed as detailed previously.[27] Individual chromosomes were counterstained with DAPI. Color digital images containing both DAPI bands and gene signals detected with anti-digoxygenin–tagged rhodamine were recorded by using a triple-band pass filter set (Chroma Technology Corp, Brattleboro, VT) in combination with a charged coupled-device camera (Photometrics Inc, Tucson, AZ) and variable excitation wave length filters.[28] Images were analyzed using the ISEE software package (Inovision Corp, Durham, NC).

Preparation of recombinant siglec-5. The extracellular region of siglec-5 was amplified by polymerase chain reaction (PCR) with the following forward and reverse primers (5′→3′): ACTCTAGAGTTC-GATCTCCCTTGCAGCAG and ACAGATCTGTTCGATCTCCCTTG-CAGCAG. The PCR product was cloned in-frame into the pIGplus vector, which encodes a Factor Xa cleavage site between the extracellular region and the hinge region of the Fc portion of human IgG1. Plasmids encoding other Fc-proteins used in this study were prepared as described previously.[7,29,30] To generate recombinant proteins, plasmids were transfected into COS-1 cells by diethylaminoethyl-dextran (DEAE) transfection and Fc-proteins were purified from the conditioned supernatants as described.[31] Briefly, supernatant was passed over a protein-A Sepharose column, and the bound protein was eluted with 0.1 mol/L glycine pH 3.0 followed by neutralization with 10% vol/vol 1.0 mol/L Tris pH 8.0. The proteins were dialyzed against 20 mmol/L Tris pH 8.0, and the concentrations were estimated by using the bicinchoninic acid (BCA) assay kit (Pierce, Rockford, IL) with bovine serum albumin (BSA) as a standard. Fc-proteins were shown to be greater than 95% pure by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis.

Generation of antibodies to siglec-5 and enzyme-linked immunosorbent assay (ELISA). Purified recombinant siglec-5 Fc was dialyzed into 20 mmol/L Tris pH 8.0 with 100 mmol/L NaCl and 2 mmol/L CaCl₂, and the protein concentration was adjusted to 1.0 mg/mL and digested overnight at room temperature with Factor Xa at a ratio of 1 μg Factor Xa:50 μg Fc-protein. Undigested material and Fc fragments were removed by passage over a protein-A Sepharose column, and the

```
       Leader peptide   | D1
MLPLLLLPLLWGGSLQEKPVYELQVQKSVTVQEGLCVLVPCSFSYPWRSWYSSPPLYVYW    60

FRDGEIPYYAEVVATNNPDRRVKPETQGRFRLLGDVQKKNCSLSIGDARMEDTGSYFFRV   120
                  | D2
ERGRDVKYSYQQNKLNLEVTALIEKPDIHFLEPLESGRPTRLSCSLPGSCEAGPPLTFSW   180
                                         | D3
TGNALSPLDPETTRSSELTLTPRPEDHGTNLTCQMKRQGAQVTTERTVQLNVSYAPQTIT   240

IFRNGIALEILQNTSYLPVLEGQALRLLCDAPSNPPAHLSWFQGSPALNATPISNTGILE   300
                       | D4
LRRVRSAEKGGFTCRAQHPLGFLQIFLNLSVYSLPQLLGPSCSWEAEGLHCRCSFRAWPA   360

PSLCWRLEEKPLEGNSSQGSFKVNSSSPGPWANSSLILHGGLNSDLKVSCKAWNIYGSQS   420

GSVLLLQGRSNLGTGVVPAALGGAGVMALLCICLCLIFFLIVKARRKQAAGRPEKMDDED   480

PIMGTITSGSRKKPWPDSPGDQASPPGDAPPLEEQKELHYASLSFSEMKSREPKDQEAPS   540

TTEYSEIKTSK                                                   551
```

Fig 1. Predicted amino acid sequence of siglec-5. The leader peptide and transmembrane region are double-underlined. Potential N-linked glycosylation sites are single-underlined. The likely Ig domain boundaries are indicated.

purified extracellular region of siglec-5 was used to immunize a group of three mice, using 10 µg protein per injection. Immune serum was collected 10 days after the third injection. To check for possible cross-reactivity of the immune serum with CD33 and to establish the saturating concentration, siglec-5 Fc and CD33 Fc were immobilized on plastic via antihuman IgG as described[7] and incubated for 1 hour with differing dilutions of preimmune or immune sera. After washing, wells were then incubated with horse radish peroxidase-conjugated goat antimouse IgG and washed and exposed to O-phenylenediamine hydrochloride as substrate, followed by measurement of absorbance at

450 nm. Anti–siglec-5 MoAbs 1A5 and 8H2 were generated from one of the immune mice by standard techniques and shown to react specifically with siglec-5 Fc by ELISA as described above.

Cells. All cell lines were provided by the ICRF Cell Production Service (South Mimms, UK). COS-1 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) with 5% heat-inactivated fetal calf serum (FCS). Other cell lines were cultured in RPMI 1640 medium with 5% or 10% FCS. Human red blood cells (RBCs) were obtained from healthy donors and stored at 4°C in Alsever's solution for up to 2 weeks. Neutrophils, monocytes, and lymphocytes were purified from human
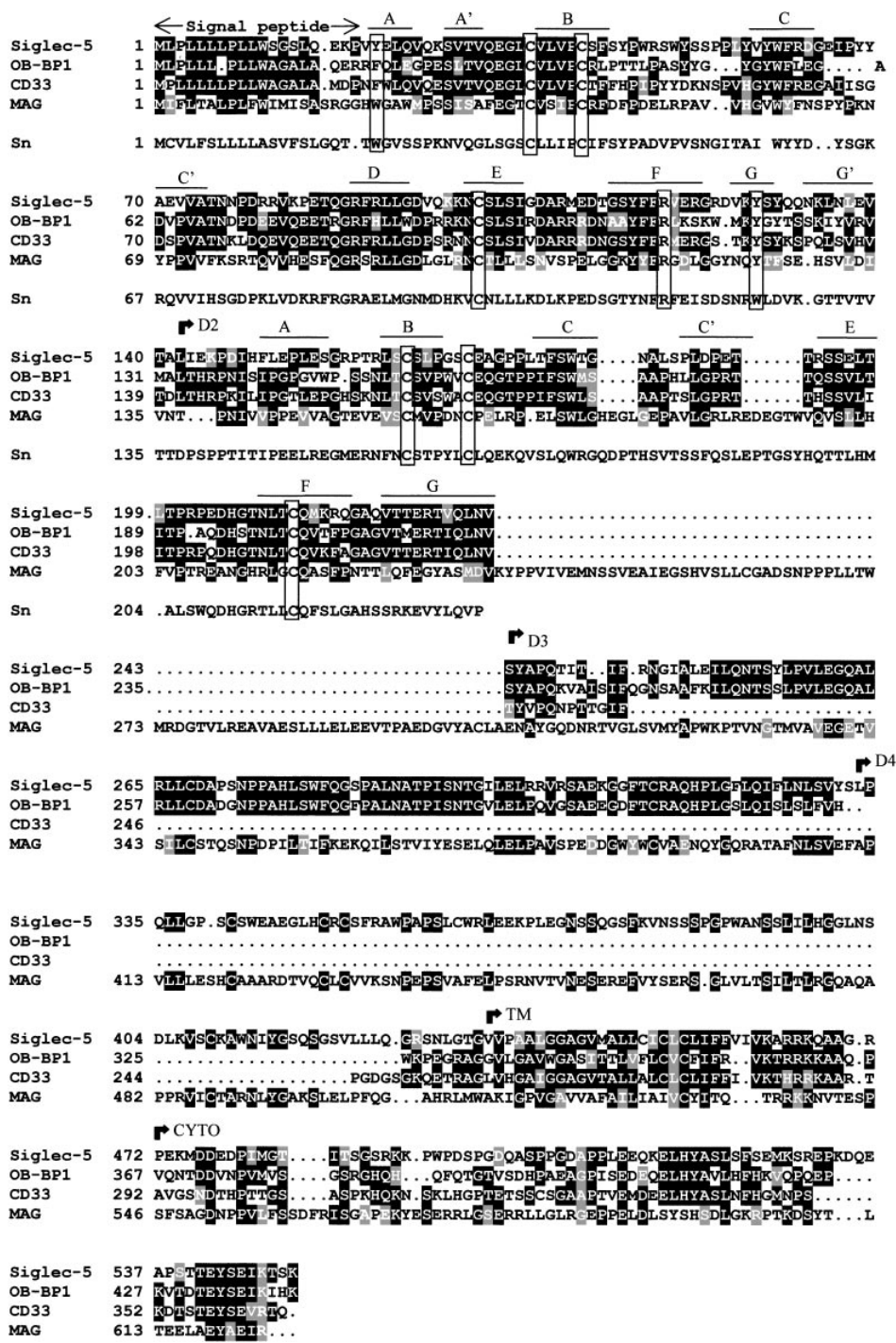

Fig 2. Alignment of siglec-5, OB-BP1, human CD33, human MAG and the first two domains of mouse sialoadhesin. Sialoadhesin is included as a reference for amino acids that are important in sialic acid binding.[10] Alignment was performed with the ClustalW multiple sequence alignment program, version 1.7 and optimized by eye. Residues that are identical in at least half the proteins are boxed in black, similar residues are in grey. Open boxes group the characteristic cysteine residues of siglec proteins and the residues that are important for interaction with sialic acid.[10] The positions of the β-strands in domains 1 and 2 are indicated. The beginning of the transmembrane (TM) and cytoplasmic (CYTO) regions are indicated. Genbank accession numbers of OB-BP1, CD33, MAG, and sialoadhesin are U71382, M23197, M29273, and Z36293, respectively.

blood as follows. Twenty milliliters of whole heparinized blood was mixed 1:5 with 6% Dextran 500 solution made in RPMI 1640 plus 20 mmol/L HEPES and left at room temperature for 20 minutes for RBCs to sediment. The RBC-free supernatant was overlaid onto 5 mL Lymphoprep (Nycomed Pharma AS, Oslo, Norway) and centrifuged at room temperature for 20 minutes. Mononuclear cells at the interface were washed in RPMI plus 20 mmol/L HEPES and added to a 10-cm diameter plastic bacterial petri dish in RPMI plus 10% FCS and incubated for 1 hour at 37°C. Nonadherent cells (lymphocytes) were removed and the adherent cells (monocytes) were detached by incubation in phosphate-buffered saline (PBS) plus 5 mmol/L EDTA for 10 minutes at 4°C. To obtain purified neutrophils, the pellet from the Lymphoprep step was resuspended in 1 mL ice-cold water for 20 seconds to lyse contaminating RBCs. The mixture was diluted to 20 mL with PBS and the neutrophils were washed twice in PBS before use.

Human RBC binding assays. The full-length cDNA insert from pHMQCD14 was excised with *Eco*RI and *Xho* I and subcloned into the corresponding sites of pcDNA3. COS-1 cells were transiently transfected with this plasmid or with the full-length CD33 plasmid in

pCDM8[32] by the DEAE-dextran method as described above, trypsinized 24 hours later, and replated at 2×10^5/well on 6-well tissue culture dishes (Falcon, Becton Dickinson UK Ltd, Oxford, UK) in DMEM containing 1% FCS. COS cells treated identically but without plasmid DNA (sham-transfected) were included in each assay. Binding assays were performed 48 to 72 hours posttransfection. RBCs were washed three times in PBS containing 0.25% BSA (PBA), resuspended at 0.25% vol/vol in DMEM plus 0.2% BSA, and 1 mL of the cell suspension was added to the wells. After 30 minutes at 37°C, nonadherent cells were removed by three gentle washes in PBA, and cellular rosettes were fixed in 0.25% glutaraldehyde. Sialidase pretreatment of COS cells or RBCs was performed with 0.1 U/mL *V cholerae* sialidase in DMEM for 2 to 3 hours at 37°C, followed by three washes in PBA. To quantify the rosetting, the number of COS cells binding more than two RBCs was scored from 20 fields in each well of duplicates using a Carl Zeiss Axioskop (Carl Zeiss Ltd, Welwyn Garden City, UK) fitted with a ×10 objective. Results were expressed as the mean number of rosettes per field.

Solid-phase binding assays with human RBCs were performed as

A

B
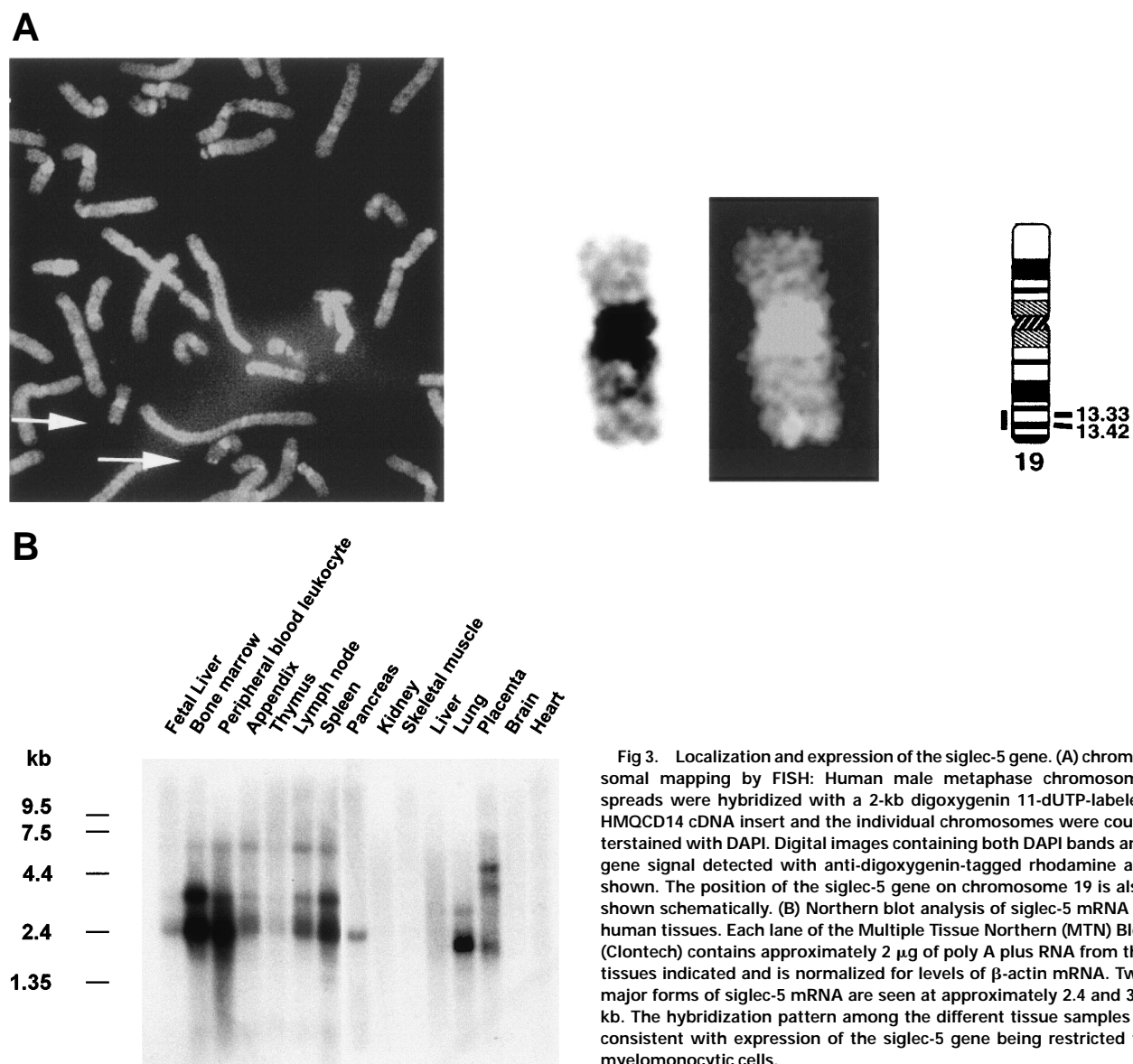
Fig 3.   Localization and expression of the siglec-5 gene. (A) chromosomal mapping by FISH: Human male metaphase chromosome spreads were hybridized with a 2-kb digoxygenin 11-dUTP-labeled HMQCD14 cDNA insert and the individual chromosomes were counterstained with DAPI. Digital images containing both DAPI bands and gene signal detected with anti-digoxygenin-tagged rhodamine are shown. The position of the siglec-5 gene on chromosome 19 is also shown schematically. (B) Northern blot analysis of siglec-5 mRNA in human tissues. Each lane of the Multiple Tissue Northern (MTN) Blot (Clontech) contains approximately 2 μg of poly A plus RNA from the tissues indicated and is normalized for levels of β-actin mRNA. Two major forms of siglec-5 mRNA are seen at approximately 2.4 and 3.4 kb. The hybridization pattern among the different tissue samples is consistent with expression of the siglec-5 gene being restricted to myelomonocytic cells.

described previously.[7] Briefly, RBCs were added to wells of microtiter plates that had been coated with Fc proteins via antihuman Fc IgG. After 30 minutes at room temperature, unbound cells were removed by washing, and bound RBCs were fixed in methanol and quantified by incubation with peroxidase substrate and measurement of absorbance at 450 nm.[7]

Binding assays with polyacrylamide glycoconjugates. Fc proteins at 1 μg/mL were added for 1 hour to wells of Immulon-3 plates previously coated with antihuman Fc IgG as described.[7] Wells were blocked for nonspecific binding with 5% skimmed milk powder for 1 hour at room temperature and incubated with varying dilutions of 2,3-PAA or 2,6-PAA dissolved in PBA for 1 hour at room temperature. Wells were washed in PBA and incubated with ^{125}I-streptavidin diluted in PBA to 0.5 μCi/mL for 1 hour at 4°C. After washing three times in PBA, bound radioactivity was solubilized by addition of 0.1 mol/L NaOH and counted by using a Beckman gamma-counter (Beckman Instruments, Fullerton, CA).

Fluorescence-activated cell sorting (FACS) analysis. All incubations were at 4°C. Cells were washed three times in PBA and incubated at 10^7/mL with anti-CD33 MoAb at 1:50, anti–siglec-5 polyclonal antibody at 1:1,000, or the preimmune serum at 1:1,000 as a negative control. After 1 hour, cells were washed and incubated with phycoerythrin-conjugated goat F(ab)₂-antimouse IgG for 1 hour on ice, washed, fixed in 2% paraformaldehyde, and analyzed on a Becton Dickinson FACS analyzer.

Western blotting. Neutrophils purified from human blood were lysed in 1% Triton-X-100 at a cellular concentration of 5×10^7/mL. Nuclei were removed by centrifugation at 10,000*g* for 15 minutes, and a quarter volume of 4X SDS-PAGE sample buffer, either reducing or nonreducing, was added. Twenty microliters of lysate were separated by SDS-PAGE on a Biorad (Hemel Hempstead, UK) minigel apparatus and transferred to nitrocellulose. The blots were incubated with 5% Marvel skimmed milk in PBS plus 0.1% Tween 20 to block nonspecific binding sites and then incubated with either preimmune serum diluted 1:1,000 or anti–siglec-5 antiserum at 1:2,000 for 1 hour at 4°C. After washing in PBS plus 0.1% Tween 20, the blots were incubated with ^{125}I-sheep antimouse IgG at a concentration of 0.5 μCi/mL for 1 hour, washed, and exposed to autoradiographic film for 24 hours.

RESULTS

Characterization of siglec-5 as a CD33-related cDNA. HMQCD14, a clone derived from a human macrophage cDNA library, was initially characterized as an EST sharing a high degree of sequence similarity with human CD33 cDNA. Examination of its full-length sequence of 2,017 bp showed a single long open reading frame encoding a type-1 transmembrane protein of 551 amino acids belonging to the Ig superfamily. Based on its sequence similarity with other siglecs and its ability to bind sialic acid (see below), this protein has been designated siglec-5 (Fig 1). The extracellular region of siglec-5 contains a hydrophobic signal peptide and four Ig-like domains, which are made up of an N-terminal V-set domain and three C2-set domains. There are eight potential N-linked glycosylation sites. After the transmembrane region, there is a cytoplasmic tail of 89 amino acids.

Homology of siglec-5 to other proteins. Database searches showed that siglec-5 is 99.6% identical to OB-BP2, a recently-deposited human sequence encoding a "leptin-binding protein" for which there is currently no published information. There are two amino acid differences: Glu 309 and Ser 403 in siglec-5 are Lys and Asn, respectively, in OB-BP2. It seems likely that siglec-5 and OB-BP2 are the product of the same gene, in which
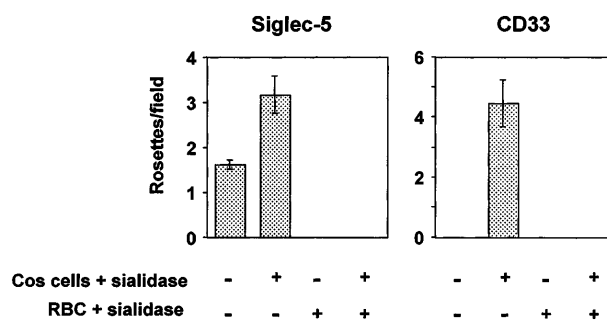

Fig 4. Binding of human RBCs to COS cells transfected with siglec-5 or CD33 cDNAs. Three days after transfection, COS cells were treated with sialidase or left untreated. Human RBCs, untreated or sialidase-treated, were added and allowed to bind for 30 minutes. After washing, cells were fixed and the number of rosettes (binding >2 RBCs) was determined by microscopy from counts of 20 fields in each well. Results show mean ± range of counts from duplicate wells and are representative of three experiments performed.

case the differences could reflect genetic polymorphism or be caused by sequencing errors. Apart from OB-BP2, the most related proteins in the database were OB-BP1 and CD33 (Fig 2). OB-BP1, like OB-BP2, is a recently-deposited human cDNA, which is almost identical to a newly-described CD33-
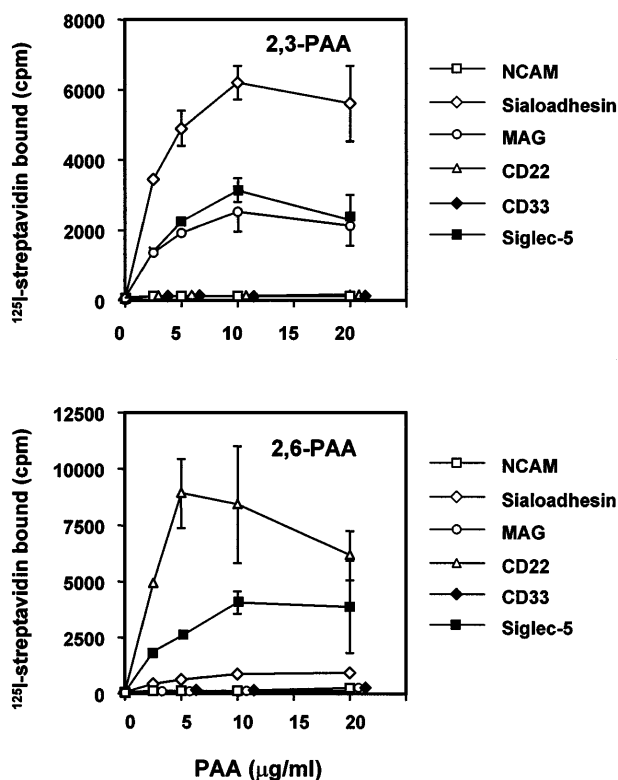

Fig 5. Binding of siglec-5 to polyacrylamide glycoconjugates in comparison with other siglecs. Fc-siglecs at 1 μg/mL were immobilized to plastic via anti-Fc antibody and biotinylated polyacrylamide (PAA) glycoconjugates linked either to 3'sialyllactose (2,3-PAA) or 6'sialyllactose (2,6-PAA) was added at the indicated concentrations. Unbound conjugate was washed off and binding detected with ^{125}I-streptavidin. Data show means ± standard deviations of triplicates and are representative of four experiments performed.
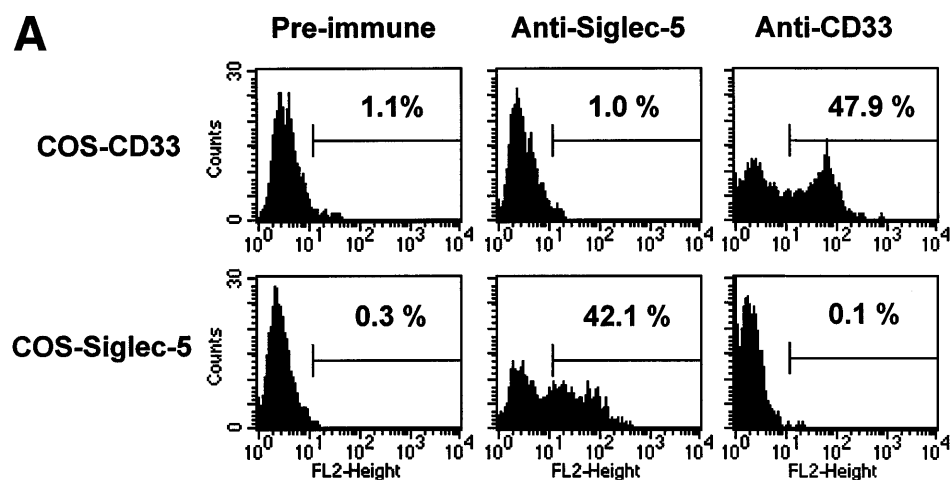
Fig 6. FACS analyses of siglec-5 expression on human hematopoietic cells in comparison to CD33. (A) COS-1 cells transiently transfected with either full-length CD33 cDNA or siglec-5 cDNA and labeled with preimmune serum, mouse anti–siglec-5 antiserum or anti-CD33 MoAb followed by phycoerythrin-conjugated F(ab)$_2$ anti-mouse IgG. Values refer to the percentage of cells within the indicated gate. (B) Human myeloid leukemic cell lines stained with either preimmune serum (open histograms), anti–siglec-5 antiserum or anti-CD33 MoAb (filled histograms). The mean fluorescence intensities observed with either anti–siglec 5 or anti-CD33 antibodies are shown in each box. (C) Human blood leukocytes, stained as described above for (B).

like cDNA, expressed specifically in placenta and designated CD33-L.[33] It contains three extracellular Ig-like domains that are 57% identical to the N-terminal three domains of siglec-5. The two extracellular Ig-like domains of CD33 are 57% identical to the two N-terminal domains of siglec-5 (Fig 2). The membrane proximal Ig-like domain of siglec-5 was found to be most similar to the fifth domain of MAG with 30% sequence identity (Fig 2).

Structural features characteristic of siglecs. Examination of the two N-terminal Ig-like domains of siglec-5 showed the presence of the characteristic structural features of the siglec subgroup of Ig superfamily proteins (Fig 2). There is precise conservation of the unusual pattern of cysteines found in these proteins, in particular the cysteine residues on the B and E strands of domain 1 that give rise to the unusual intrasheet disulfide bond[10,14] and the cysteines on the A-B loop of domain 1 and the B-C loop of domain 2 that are thought to result in an interdomain disulfide bond.[14,15] There is also conservation of key amino acids involved in sialic acid binding, in particular the critical arginine at position 119 and the two conserved aromatic residues at positions 21 and 128 on the A and G strands of the V-set domain. These are both tyrosine in siglec-5 rather than tryptophan found in sialoadhesin (Fig 2). It is also evident from the alignment that OB-BP1 exhibits the same structural features and therefore, is likely to be a functional sialic acid–binding protein.

The transmembrane region and cytoplasmic tail of siglec-5 are very similar to the corresponding regions of OB-BP1 and CD33 with overall identities of 40% and 42%, respectively. Within the cytoplasmic tails of these proteins, there are two highly-conserved regions of seven amino acids centered in both cases around tyrosine residues. The first region (LHYAS/VL) conforms well to the consensus immunoreceptor tyrosine-based inhibitory motif (ITIM) (I/L/VxYxxL/V) defined in several other leukocyte proteins, whereas the second motif (TEYSEI/V)

does not conform, but is similar to ITIM-like motifs described in other leukocyte cell surface proteins.[34]

Chromosomal localization and expression of the siglec-5 gene. The gene encoding siglec-5 was mapped by in situ hybridization on chromosome 19q13.41-q13.43 (Fig 3A) and was closely linked to the CD33 gene at 19q13.3-q13.4.[35] Northern blot analysis (Fig 3B) showed the presence of two major siglec-5 mRNA transcripts of 2.4 kb and 3.4 kb in various tissues with highest levels in hematopoietic organs, notably bone marrow and spleen. High levels were also detected in extracts of peripheral blood leukocytes, with lower levels present in lymph node, lung, appendix, placenta, pancreas, and thymus. Siglec-5 mRNA was undetectable in brain, heart, skeletal muscle, and kidney.

Siglec-5 mediates sialic acid-dependent binding to human RBCs. All siglecs so far characterized are able to mediate sialic acid–dependent binding to human RBCs after transient expression of their full-length cDNAs in COS cells.[7,29,36] However, in the case of CD33, binding was only shown after pretreating the transfected COS cells with sialidase before adding the RBCs.[29] Because CD33 contains only two Ig-like domains, its binding site at the N-terminus may not extend sufficiently from the plasma membrane to avoid *cis*-interactions with sialic acids in the glycocalyx, thus reducing its potential to interact with ligands added in trans. Treatment of the transfected COS cells with sialidase removes the inhibitory sialic acids in the glycocalyx, thereby allowing the protein to bind RBCs. In the present experiments, COS cells transfected with siglec-5 or CD33 cDNAs were either untreated or treated with sialidase before addition of human RBCs. To check that binding was sialic acid dependent, assays were also performed by using sialidase-treated RBCs. COS cells transfected with siglec-5 cDNA bound RBCs at low levels, and this could be increased substantially when the transfected COS cells were pretreated with sialidase (Fig 4). In contrast, binding of RBCs to CD33-
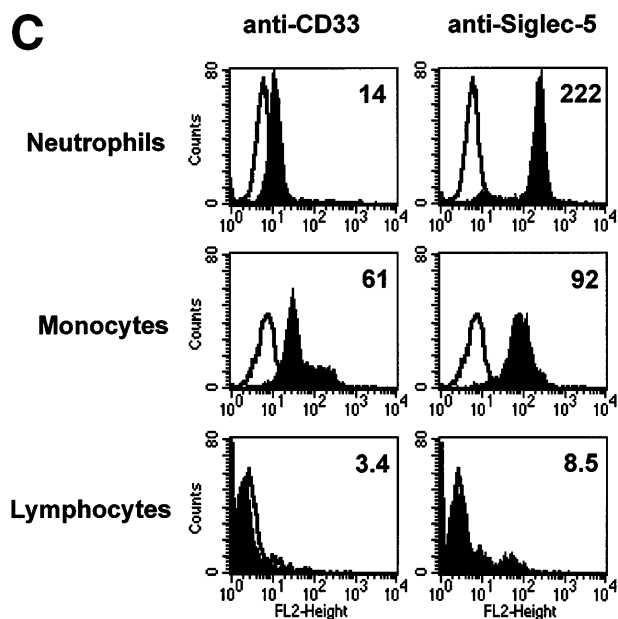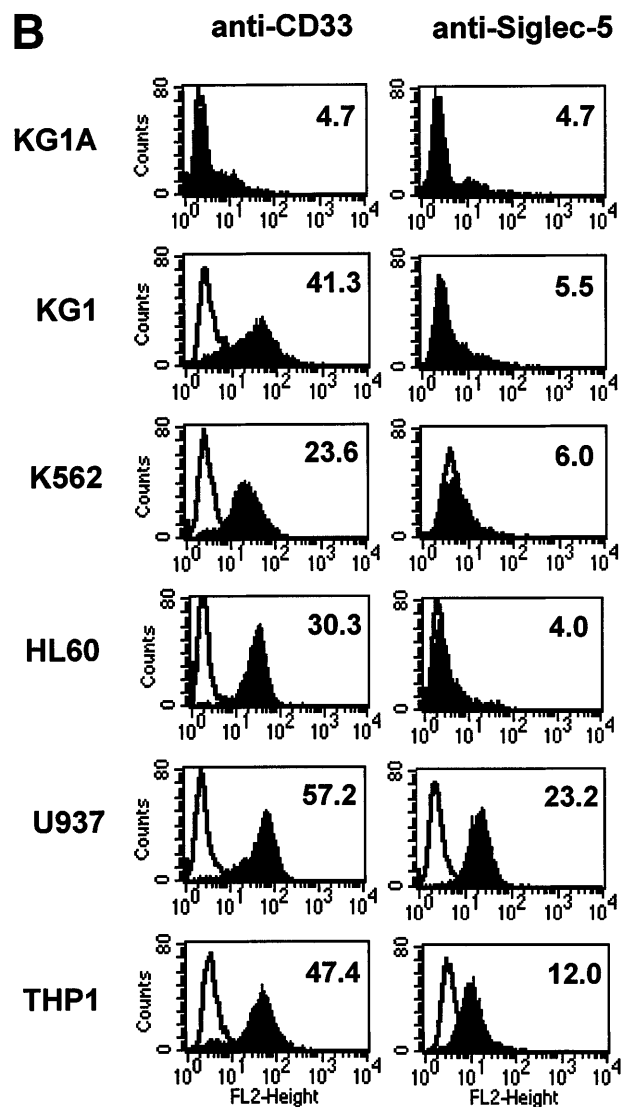
Fig 6 (Cont'd).

transfected COS cells was only observed after sialidase treatment of the COS cells, as shown previously[29] (Fig 4). In all cases, sialidase-pretreatment of the RBCs abrogated binding, indicating that it was sialic acid dependent.

Sialic acid linkage preference of siglec-5. It has been shown previously that siglecs exhibit a preference for the glycosidic linkage of sialic acid to adjacent sugars.[7,29,37,38] Thus, sialoadhesin, CD33, and MAG bind preferentially to oligosaccharides terminating with $\alpha2,3$-linked sialic acid, whereas CD22 binds only to the $\alpha2,6$ linkage. To investigate the sialic acid linkage preference of siglec-5, a fusion protein was generated in which the entire extracellular region was fused to the Fc portion of human IgG1. The recombinant protein was used in solid-phase binding assays with biotinylated polyacrylamide derivatized with either 3'- or 6'-sialyllactose and its binding activity was compared with those of sialoadhesin, CD33, MAG, CD22, and N-CAM as a negative control (Fig 5). Siglec-5 showed a unique specificity, binding equally to both 3'-and 6'-sialyllactose–conjugated polyacrylamide, unlike sialoadhesin and MAG, which bound preferentially to 3' sialyllac-

tose, or CD22, which bound specifically to the 6' sialyllactose glycoconjugate (Fig 5). Unexpectedly, no binding was seen with CD33-Fc. This was not caused by inactivity of the CD33-Fc protein because in solid-phase binding assays with RBCs done in parallel, CD33-Fc and siglec-5–Fc bound RBCs to a similar level (data not shown).

Expression of siglec-5 on human leukemic cell lines and human blood leukocytes. The pattern of mRNA expression observed in Northern blots of human tissues (Fig 3A) is consistent with siglec-5 being a myeloid-restricted protein. Therefore, it was of interest to examine expression of siglec-5 protein in comparison to CD33, which is found predominantly on myeloid cells. CD33 is first detected on mixed progenitors and continuing down the myelomonocytic pathway until being downregulated on granulocytes but retained on monocytes and some tissue macrophages.[39] Expression of siglec-5 on hematopoietic cells was assessed by using a mouse polyclonal antiserum raised to the extracellular region of siglec-5. ELISA experiments with immobilized Fc proteins showed that the anti–siglec-5 antiserum did not cross-react with CD33 (data not

shown). This was confirmed in FACS analysis by using COS cells transiently transfected with either siglec-5 or CD33 full-length cDNAs (Fig 6A).

Leukemic cell lines that are "fixed" at a particular stage of hematopoietic differentiation are useful models to study expression of lineage-restricted antigens. For this purpose we analyzed KG1A (CD34$^+$ CD33$^-$), KG1 (CD34$^-$, CD33$^+$), K562 (CD33$^+$, displays erythroid/megakaryocytic markers), HL60 (CD33$^+$, myelomonocytic), U937 (CD33$^+$, promonocytic), and THP-1 (CD33$^+$, monocytic). FACS analyses showed that the expression pattern of siglec-5 was similar, yet clearly distinct from that of CD33. Neither antigen was expressed on KG1A cells thought to represent early progenitors. In contrast to CD33, siglec-5 expression was not detected on KG1, K562, or HL-60 cells. Both CD33 and siglec-5 were readily detectable on U937 and THP-1 cells (Fig 6B). We next compared expression of CD33 and siglec-5 on human blood leukocytes (Fig 6C). Neither was expressed on peripheral blood lymphocytes (Fig 6C) nor on a range of B and T lymphoblastic cell lines (data not shown). Both were expressed at intermediate levels on monocytes, but only siglec-5 was present at high levels on neutrophils (Fig 6C). This pattern of expression of siglec-5 on human leukocytes was confirmed by using anti–siglec-5 MoAbs 1A5 and 8H2 (data not shown).

Molecular characterization of siglec-5 on neutrophils. To determine the molecular weight of siglec-5 expressed by neutrophils, Western blots of neutrophil lysates were probed with the anti–siglec-5 antiserum at 1:2,000 dilution. No bands were seen with preimmune serum at 1:1,000 dilution (data not shown) but with the immune serum a predominant band was observed at approximately 140 kD under nonreducing conditions and at approximately 70 kD under reducing conditions (Fig 7). This suggests that siglec-5, similar to CD33, exists as a homodimer on the cell surface.

DISCUSSION

The recent availability of extensive EST databases has been of great value in permitting the identification and characterization of new members of preexisiting families of proteins.[24] In the case of the siglec-5 protein presented here, the high degree of sequence similarity with CD33 and conservation of key residues implicated in sialic acid recognition increased the possibility that this protein was a new member of the siglec family of Ig superfamily proteins. In the present study, we were able to show that the full-length cDNA expressed in COS cells behaved similarly to all other previously characterized siglecs in being able to bind human RBCs in a sialic acid–dependent manner. Therefore, siglec-5 fulfills the criteria required for its inclusion in the siglec subgroup of Ig superfamily proteins.[8]

The high degree of sequence similarity within the two N-terminal domains of siglec-5 and CD33 suggested that they might share similar sialic acid–binding properties. However, two potentially important differences were shown in our experiments. First, COS cells that expressed siglec-5 were able to mediate binding to RBCs without sialidase treatment, unlike CD33. Second, siglec-5 showed an unexpected binding specificity, recognizing α2,3- and α2,6-linked sialic acid equally. This is in contrast to all other previously-characterized members of the family that preferentially bind to either α2,3- or α2,6-linked

sialic acids.[7,20,37] The specificity analysis was performed by using polyacrylamide glycoconjugates rather than resialylated human RBCs used extensively in our previous studies. Although the polyacrylamide-based glycoconjugates exhibited good binding to immobilized sialoadhesin, MAG, CD22, and siglec-5, there was no measurable interaction with CD33, despite the fact that the immobilized CD33 was able to bind human RBCs in a sialic acid–dependent manner. This difference between CD33 and siglec-5 could reflect weaker binding affinity of CD33 for sialylated oligosaccharides or could point to a more stringent requirement of CD33 for oligosaccharide presentation on appropriate carrier molecules.

The pattern of expression of siglec-5 on myeloid cell lines and blood leukocytes was also different from that of CD33, being absent from cell lines with an immature myeloid phenotype such as KG1 and K562, but was expressed on cell lines with a more differentiated phenotype such as U937 and THP-1 cells. Among blood leukocytes, a clear difference was also seen with neutrophils that downregulate expression of CD33 but express relatively high levels of siglec-5. This expression pattern indicates that siglec-5 is more likely to be involved in effector functions of myeloid cells rather than in the differentiation process per se. In this respect, it is of interest that CD33 and siglec-5 share two conserved tyrosine residues embedded in ITIM-like motifs that have been found in a growing number of cell surface members of the Ig superfamily.[34,40] In many cases, such ITIM motifs have been shown to mediate negative regulatory signals via recruitment and activation of tyrosine phosphatases like SHP-1, SHP-2, and SHIP.[34,41] Future studies are needed to determine what role, if any, the ITIM-like motifs have in potential signalling functions of CD33 and siglec-5.

An emerging theme for siglecs is that their ability to mediate cell-cell interactions can be influenced by the nature and extent of sialylation on the host cell, the target cell, and also the
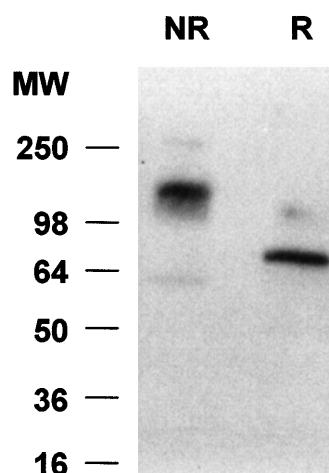
Fig 7. Western blot analyses of neutrophil lysates with anti–siglec-5 antiserum. Samples, either reduced (R) or nonreduced (NR), were separated on a 10% polyacrylamide gel and transferred to nitrocellulose. The blot was incubated with anti–siglec-5 antiserum at 1:2,000 dilution or preimmune serum at 1:1,000 dilution, washed, and incubated with ^{125}I-antirabbit IgG. No signal was observed with preimmune serum (not shown) but a predominant band at approximately 70 kD and approximately 140 kD can be seen under reducing and nonreducing conditions, respectively.

relative length of the siglec. Thus, sialoadhesin with 17 Ig domains is thought to extend the binding site up to 35 nm from the plasma membrane, thereby reducing inhibitory cis-interactions with sialic acids present in the macrophage glycocalyx (Crocker, unpublished observations).[42,43] At the other end of the spectrum, CD33 with only two Ig domains appears unable to mediate trans interactions unless the endogenous sialic acids on the cell surface are removed by sialidase treatment.[29] As shown in the present study, siglec-5, with 4 Ig domains, is able to mediate modest binding to RBCs on COS cells in the absence of sialidase treatment, unlike CD33. This finding, together with the demonstration that siglec-5 is present at high levels on neutrophils, raises the possibility that this receptor could be involved in cellular interactions of neutrophils during acute inflammatory responses.

The recent deposition in the Genbank database of leptin-binding proteins OB-BP1 and OB-BP2 (virtually identical to siglec-5), as well as a recent report of another CD33-like sequence expressed specifically in placenta[33] (virtually identical to OB-BP1), together with our unpublished observations indicate the existence of a novel subgroup of CD33-like genes that are likely to function as sialic acid–binding proteins. Interestingly, siglec-5 maps to chromosome 19q13.4, a region where a large number of other hematopoietically expressed Ig superfamily members have been localized. These include a family of genes encoding killer cell inhibitory receptors expressed on natural killer cells and subsets of T-lymphocytes, immunoglobulin-like transcripts (ILT-1, -2, and -3) expressed on myeloid cells and subsets of lymphoid cells, the gp49 family of receptors expressed on mast cells and natural killer cells, and the gene encoding human myeloid immunoglobulin A Fc receptor (CD89).[44] An emerging theme for these receptors is that they are involved in either activatory or inhibitory signalling functions in hematopoietic cells.[34] A major challenge for the future will be to elucidate the functions of the CD33-like subgroup of sialic acid–binding receptors and to determine the significance of sialic acid recognition.

ACKNOWLEDGMENT

REFERENCES

1. Barclay AN, Brown MH, Law SKA, McKnight AJ, Tomlinson MG, van der Merwe PA: The Leucocyte Antigen Facts Book. San Diego, CA, Academic Press, 1997

2. Crocker PR, Kelm S, Hartnell A, Freeman S, Nath D, Vinson M, Mucklow S: Sialoadhesin and related cellular recognition molecules of the immunoglobulin superfamily. Biochem Soc Trans 24:149, 1996

3. Cyster J, Goodnow C: Tuning antigen receptor signaling by CD22: Integrating cues from antigens and the microenvironment. Immunity 6:509, 1997

4. Crocker PR, Hartnell A, Munday J, Nath D: The potential role of sialoadhesin as a macrophage recognition molecule in health and disease. Glycoconj J 14:601, 1997

5. Fruttiger M, Montag D, Schachner M, Martini R: Crucial role for the myelin-associated glycoprotein in the maintenance of axon-myelin integrity. Eur J Neurosci 7:511, 1995

6. Kelm S, Schauer R, Manuguerra J-C, Gross H-J, Crocker PR: Modifications of cell surface sialic acids modulate cell adhesion mediated by sialoadhesin and CD22. Glycoconj J 11:576, 1994

7. Kelm S, Pelz A, Schauer R, Filbin MT, Tang S, DeBellard M-E, Schnaar RI, Mahoney JA, Hartnell A, Bradfield P, Crocker PR: Sialoadhesin, MAG and CD22 define a new family of sialic acid-dependent adhesion molecules of the immunoglobulin superfamily. Curr Biol 4:965, 1994

8. Crocker PR, Clark EA, Filbin M, Gordon S, Jones Y, Kehrl JH, Kelm S, Douarin NL, Powell L, Roder J, Schnaar RL, Sgroi DC, Stamenkovic I, Schauer R, Schachner M, van den Berg TK, van der Merwe PA, Watt SM, Varki A: Siglecs: A new family of sialic acid-binding lectins. Glycobiology 8:v, 1998

9. Vinson M, van der Merwe PA, Kelm S, May A, Jones EY, Crocker PR: Characterization of the sialic acid-binding site in sialoadhesin by site-directed mutagenesis. J Biol Chem 271:9267, 1996

10. May AP, Robinson RC, Vinson M, Crocker PR, Jones EY: Crystal structure of the N-terminal domain of sialoadhesin in complex with 3' sialyllactose at 1.85A resolution. Molecular Cell 1:719, 1998

11. van der Merwe PA, Crocker PR, Vinson M, Barclay AN, Schauer R, Kelm S: Localization of the putative sialic acid-binding site on the immunoglobulin superfamily cell-surface molecule CD22. J Biol Chem 271:9273, 1996

12. Freeman SD: Studies on the myeloid antigens CD33 and sialoadhesin. PhD thesis, Oxford University, Oxford, UK, 1996

13. Tang S, Shen YJ, DeBellard ME, Mukhopadhyay G, Salzer JL, Crocker PR, Filbin MT: Myelin-associated glycoprotein (MAG) interacts with neurons via a sialic acid binding site at arg118 and a distinct neurite inhibition site. J Cell Biol 138:1355, 1997

14. Williams AF, Davis SJ, He Q, Barclay AN: Structural diversity in domains of the immunoglobulin superfamily. Cold Spring Harbor Symp Quant Biol LIV:637, 1989

15. Pedraza L, Owens GC, Green LAD, Salzer JL: The myelin-associated glycoproteins: Membrane disposition, evidence of a novel disulfide linkage between immunoglobulin-like domains, and posttranslational palmitylation. J Cell Biol 111:2651, 1990

16. Pierelli L, Teofili L, Menichella G, Rumi C, Paoloni A, Iovino S, Puggioni PL, Leone G, Bizzi B: Further investigations on the expression of HLA-DR, CD33 and CD13 surface antigens in purified bone marrow and peripheral blood CD34+ haematopoietic progenitor cells. Br J Haematol 84:24, 1993

17. Dorken B, Moldenhauer G, Pezzutto A, Schwartz R, Feller A, Kiesel S, Nadler L: HD39 (B3), a B lineage-restricted antigen whose cell surface expression is limited to resting and activated human B lymphocytes. J Immunol 136:4470, 1986

18. Crocker PR, Gordon S: Mouse macrophage hemagglutinin (sheep erythrocyte receptor) with specificity for sialylated glycoconjugates characterised by a monoclonal antibody. J Exp Med 169:1333, 1989

19. Steiniger B, Barth P, Harbst B, Hartnell A, Crocker PR: The species-specific structure of microanatomical compartments in the human spleen: Strongly sialoadhesin-positive macrophages occur in the perifollicular zone, but not in the marginal zone. Immunology 92:307, 1997

20. Griffin JD, Linch D, Sabbath K, Larcom P, Schlossman SF: A monoclonal antibody reactive with normal and leukaemic human myeloid progenitor cells. Leukaemia Res 8:521, 1984

21. Mason DY, Stein H, Gerdes J, Pulford KA, Ralfkiaer E, Falini B, Erber WN, Micklem K, Gatter KC: Value of monoclonal anti-CD22 (p135) antibodies for the detection of normal and neoplastic B lymphoid cells. Blood 69:836, 1987

22. McGraw KJ, Rosenblum MG, Cheung L, Scheinberg DA: Characterization of murine and humanized anti-CD33, gelonin immunotoxins reactive against myeloid leukemias. Cancer Immunol Immunother 39:367, 1994

23. Mansfield E, Amlot P, Pastan I, FitzGerald DJ: Recombinant RFB4 immunotoxins exhibit potent cytotoxic activity for CD22-bearing cells and tumors. Blood 90:2020, 1997

24. Adams MD, Kerlavage AR, Fleischmann RD, Fuldner RA, Bult CJ, Lee NH, Kirkness EF, Weinstock KG, Gocayne JD, White O, Sutton G, Blake JA, Brandon RG, Chiu MW, Clayton RA, Cline RT, Cotton MD, Earle-Hughes J, Fine LD, Fitzgerald LM, Fitzhugh WM, Fritchman JL, Goeghagen NSM, Glodek A, Gnehm CL, Hanna MC, Hedblom E, Hinkle PS, Kelley JM, Klimek KM, Kelley JC, Liu LI, Marmaros SM, Merrick JM, Moreno-Palanques RF, McDonald LA, Nguyen DT, Pellegrino SM, Phillips CA, Ryder SE, Scott JL, Saudek DM, Shirley R, Small KV, Spriggs TA, Utterback TR, Weldman JF, Li Y, Barthlow R, Bednarik DP, Cao LA, Cepeda MA, Coleman TA, Collins EJ, Dimke D, Feng P, Ferrie A, Fischer C, Hastings GA, He WW, Hu JS, Huddleston KA, Greene JM, Gruber J, P H, Kim MA, Kozak DL, Kunsch C, Ji HJ, Li HD, Meissner PS, Olson H, Raymond L, Wei YF, Wing J, Xu C, Yu GL, Ruben SM, Dillon PJ, Fannon MR, Rosen CA, Haseltine WA, Fields C, Fraser CM, Venter JC: Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. Nature 377:3, 1995

25. Ni J, Abrahamson M, Zhang M, Fernandez MA, Grubb A, Su J, Yu GL, Li Y, Parmelee D, Xing L, Coleman TA, Gentz S, Thotakura R, Nguyen N, Hesselberg M, Gentz R: Cystatin E is a novel human cysteine proteinase inhibitor with structural resemblance to family 2 cystatins. J Biol Chem 272:10853, 1997

26. Feng GS, Ouyang YB, Hu DP, Shi ZQ, Gentz R, Ni J: Grap is a novel SH3-SH2-SH3 adaptor protein that couples tyrosine kinases to the Ras pathway. J Biol Chem 271:12129, 1996

27. Johnson CV, Singer RH, Lawrence JB: Fluorescent detection of nuclear RNA and DNA: Implications for genome organization. Methods Cell Biol 35:73, 1991

28. Johnson CV, McNeil JA, Carter KC, Lawrence JB: A simple, rapid technique for precise mapping of multiple sequences in two colors using a single optical filter set. Genet Anal Tech Appl 8:75, 1991

29. Freeman SD, Kelm S, Barber EK, Crocker PR: Characterisation of CD33 as a new member of the Sialoadhesin family of cellular interactions molecules. Blood 85:2005, 1995

30. Nath D, van der Merwe PA, Kelm S, Bradfield P, Crocker PR: The amino-terminal immunoglobulin-like domain of sialoadhesin contains the sialic acid binding site. Comparison with CD22. J Biol Chem 270:26184, 1995

31. Simmons DL: Cloning cell surface molecules by transient expression in mammalian cells, in Hartley DA (ed): Cellular Interactions in Development—A Practical Approach. Oxford, UK, IRL Press, 1993, p 93

32. Simmons DL, Seed B: Isolation of a cDNA encoding CD33, a differentiation antigen of myeloid progenitor cells. J Immunol 141:2797, 1988

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

34. Vely F, Vivier E: Conservation of structural features reveals the existence of a large family of inhibitory cell surface receptors and noninhibitory/activatory counterparts. J Immunol 159:2075, 1997

35. Trask BFA, Christensen M, Youngblom J, Bergmann A, Copeland A, de Jong P, Mohrenweiser H, Olsen A, Carrano A, Tynan K: Fluorescence in situ hybridization mapping of human chromosome 19: Cytogenetic band location of 540 cosmids and 70 genes or DNA markers. Genomics 15:133, 1993

36. Tropak MB, Roder JC: Regulation of myelin-associated glycoprotein binding by sialylated cis-ligands. J Neurochem 68:1753, 1997

37. Powell LD, Varki A: The oligosaccharide binding specificities of CD22-beta, a sialic acid-specific lectin of B-cells. J Biol Chem 269:10628, 1994

38. Collins BE, Kiso M, Hasegawa A, Tropak MB, Roder JC, Crocker PR, Schnaar RL: 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, 1997

39. Peiper SC, Leboeuf RD, Hughes CB, Prasthofer EF, Borowitz MJ, Dambuyant CD, Katz DR, Walker WS, Ashmun RA, Look TA: Report on the CD33 cluster workshop: Biochemical and genetic characterisation of gp67, in Knapp W, Dorken B, Gilks WR, Rieber EP, Schmidt RE, Stein H, von dem Borne AEG (eds): Leucocyte Typing IV. White Cell Differentiation Antigens. Oxford, UK, Oxford University Press, 1989, p 814

40. Borges L, Hsu M-L, Fanger N, Kubin M, Cosman D: A family of human lymphoid and myeloid Ig-like receptors, some of which bind to MHC class I molecules. J Immunol 159:5192, 1997

41. Unkeless JC, Jin J: Inhibitory receptors, ITIM sequences and phosphatases. Curr Opin Immunol 9:338, 1997

42. Crocker PR, Kelm S, Dubois C, Martin B, McWilliam AS, Shotton DM, Paulson JC, Gordon S: Purification and properties of sialoadhesin, a sialic acid-binding receptor of murine tissue macrophages. EMBO J 10:1661, 1991

43. Crocker PR, Mucklow S, Bouckson V, McWilliam A, Willis AC, Gordon S, Milon G, Kelm S, Bradfield P: Sialoadhesin, a macrophage-specific adhesion molecule for haemopoietic cells with 17 immunoglobin-like domains. EMBO J 13:4490, 1994

44. Dupont B, Selvakumar A, Steffens U: The killer cell inhibitory receptor genomic region on human chromosome 19q13.4. Tissue Antigens 49:557, 1997