

Identification, characterization and leucocyte expression of Siglec-10, a novel human sialic acid-binding receptor

James MUNDAY*, Sheena KERR*, Jian NI†, Ann L. CORNISH*, Jiquan Q. ZHANG*, Gavin NICOLL*, Helen FLOYD*, Marie-Genevieve MATTEI‡, Paul MOORE†, Ding LIU† and Paul R. CROCKER*¹

*The Wellcome Trust Biocentre, School of Life Sciences, University of Dundee, Dundee DD1 5EH, Scotland, U.K., †Human Genome Sciences Inc., 9410 Key West Avenue, Rockville, MD 20850-3338, U.S.A. and ‡INSERM U491, Unité de Génétique Médicale et Développement, Faculté de Médecine, Boulevard Jean Moulin, 13385 Marseille, Cedex 5, France

Here we characterize Siglec-10 as a new member of the Siglec family of sialic acid-binding Ig-like lectins. A full-length cDNA was isolated from a human spleen library and the corresponding gene identified. Siglec-10 is predicted to contain five extracellular Ig-like domains and a cytoplasmic tail containing three putative tyrosine-based signalling motifs. Siglec-10 exhibited a high degree of sequence similarity to CD33-related Siglecs and mapped to the same region, on chromosome 19q13.3. The expressed protein was able to mediate sialic acid-dependent binding to human

erythrocytes and soluble sialoglycoconjugates. Using specific antibodies, Siglec-10 was detected on subsets of human leucocytes including eosinophils, monocytes and a minor population of natural killer-like cells. The molecular properties and expression pattern suggest that Siglec-10 may function as an inhibitory receptor within the innate immune system.

Key words: immunoglobulin superfamily, inhibitory receptor, innate immunity, lectin.

INTRODUCTION

Siglecs (sialic acid-binding Ig-like lectins) are members of the Ig superfamily expressed at the cell surface [1]. Structurally, they have a characteristic N-terminal V-set Ig-like domain containing the sialic acid-binding site [2] followed by varying numbers of C2-set domains. Most Siglecs have one or more immune receptor tyrosine-based inhibitory motifs (ITIMs) in their cytoplasmic tails implicated in negative regulatory signalling functions [3–9]. The first Siglecs to be defined were sialoadhesin/Siglec-1, CD22/Siglec-2, CD33/Siglec-3 and myelin-associated glycoprotein (MAG)/Siglec-4 expressed by macrophages, B-cells, myeloid cells and myelin-forming cells respectively [10–14]. Recent studies have revealed the existence of a new subset of human Siglecs that are highly related to CD33/Siglec-3, namely Siglecs 5, 6, 7, 8 and 9 [6,15–22]. Each protein exhibits distinct sialic acid-binding properties and is expressed in a characteristic manner. With the exception of Siglec-6, which is found mostly in the placenta [16,23] but also on B-cells [16], the CD33-related group of Siglecs is largely expressed on discrete subsets of haemopoietic cells. Interestingly, the highest levels of expression of these proteins are found on effector cells of the innate immune system, including monocytes (Siglecs 3, 5, 7 and 9), neutrophils (Siglecs 5 and 9), eosinophils (Siglec-8) and natural killer cells (NK cells; Siglec-7). This observation, together with the common theme of sialic acid recognition and the presence of conserved, putative tyrosine-based inhibitory signalling motifs [4–6,9], has led to the suggestion that several of the CD33-related Siglecs might be involved in regulating activation of leucocytes via sialic acid recognition [21,22].

In this paper we describe the properties of a new member of the Siglec family, Siglec-10. Amongst peripheral blood leucocytes Siglec-10 was found to be expressed on eosinophils, monocytes and B-cells. Higher levels of this novel Siglec were found on a

minor subset of NK-like cells that express the CD16 low-affinity Fc receptor but lack the CD56 antigen. The molecular properties and expression pattern suggest that Siglec-10 may have evolved to function as an inhibitory receptor of the innate immune system.

EXPERIMENTAL

Materials

Unless specified otherwise, all reagents and chemicals were purchased from Sigma. ¹²⁵I-Streptavidin (20–40 mCi/mg) and Protein A-Sepharose were obtained from Amersham Pharmacia Biotech. *Vibrio cholerae* sialidase was purchased from Calbiochem. Biotinylated polyacrylamide (PAA) glycoconjugates carrying either NeuAc α 2,3Gal β 1,4Glc (2,3-PAA) or NeuAc α 2,6Gal β 1,4Glc (2,6-PAA) were obtained from Syntosome (Munich, Germany). These conjugates have a molecular mass of \approx 30 kDa and contain 20 mol% of saccharide and 5 mol% of biotin. Phycoerythrin-conjugated monoclonal antibodies (mAbs) against the following human CD antigens were purchased from Serotec (Kidlington, Oxford, U.K.): CD3, CD4, CD8, CD16, CD19 and CD56. FITC-conjugated F(ab)₂ anti-mouse IgG was from Dako (Cambridge, U.K.).

Identification and characterization of Siglec-10 cDNA

Using the amino acid sequence of CD33, a specific homology search was performed in the Human Genome Sciences database, containing more than 1 million expressed-sequence tags (ESTs) obtained from over 700 different cDNA libraries. A total of 21 clones encoding a potential novel Siglec were identified in cDNA libraries prepared from the following human sources: bone marrow, unstimulated B-cells, eosinophils, primary dendritic cells, spleen and chronic lymphocytic leukaemia

Abbreviations used: Siglec, sialic acid-binding Ig-like lectin; CHO, Chinese hamster ovary; ITIM, immune receptor tyrosine-based inhibitory motif; PAA, biotinylated polyacrylamide; Lac-PAA, Gal β 1,4Glc coupled to PAA; mAb, monoclonal antibody; MAG, myelin-associated glycoprotein; Siglec-10-Fc, five extracellular domains of Siglec-10 coupled to the Fc part of human IgG1; Siglec-10-CHO, CHO cells stably expressing Siglec-10; 2,3-PAA, NeuAc α 2,3Gal β 1,4Glc coupled to PAA; 2,6-PAA, NeuAc α 2,6Gal β 1,4Glc coupled to PAA; NK cells, natural killer cells; RBC, red blood cell.

¹ To whom correspondence should be addressed (e-mail p.r.crocker@dundee.ac.uk).

Table 1 Sequence comparisons of Siglec-10 with CD33-related Siglecs and MAG

Amino acid sequences from the indicated regions of Siglec-10 were aligned with the indicated regions of the different Siglecs using the SIM sequence-alignment program. The percentage identities are shown. LP, leader peptide; D1–5, domains 1–5; NA, not applicable.

Siglec-10	Identity (%)						
	CD33/Siglec-3	Siglec-5	Siglec-6	Siglec-7	Siglec-8	Siglec-9	MAG/ Siglec-4
LP + D1 + D2	49 (LP + D1 + D2)	48 (LP + D1 + D2)	40 (LP + D1 + D2)	43 (LP + D1 + D2)	48 (LP + D1 + D2)	45 (LP + D1 + D2)	23 (LP + D1 + D2)
D3	NA	42 (D3)	42 (D3)	49 (D3)	45 (D3)	45 (D3)	31 (D3)
D4	NA	50 (D3)	56 (D3)	67 (D3)	67 (D3)	60 (D3)	35 (D4)
D5	NA	71 (D4)	NA	NA	NA	NA	30 (D5)

cells. Inserts from four clones were sequenced: pHEOMH10 (eosinophil), pHEOOV77 (eosinophil), pHDPIB36 (dendritic cell) and pHDPCLO5 (dendritic cell). Since none of these was found to encode a correctly spliced full-length form of this novel Siglec, a human spleen cDNA library in λ ZAPII [24] was screened with a *KpnI*–*SacI* fragment from HDPCLO5. A single positive phage clone, p2.2, was identified and the phage insert subcloned into pBluescript and sequenced. Since clone p2.2 contained a correctly spliced full-length open reading frame, the insert was subcloned into the mammalian expression vector pcDNA3 (Invitrogen, Groningen, The Netherlands) and used as a template in subsequent molecular characterization. On the basis of its sialic acid-binding activity, this novel Siglec is herein referred to as Siglec-10. To identify proteins related to Siglec-10, a computer search of the GenBank nucleotide and protein sequence databases was carried out using the Blast GeneSearch (National Center for Biotechnology Information, National Institutes of Health, Bethesda, MD, U.S.A.).

Sequence comparisons and phylogenetic analysis

Manipulations of sequences and alignments were performed using ClustalW sequence-alignment software available on the worldwide web, either at the Human Genome Center, Baylor College of Medicine, Houston, TX, U.S.A. (<http://dot.imgen.bcm.tmc.edu:9331/multi-align/multi-align.html>) or at the European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambs., U.K. (<http://www.ebi.ac.uk/clustalw/>). Protein alignments of individual domains shown in Table 1 were performed using the SIM alignment program [25]. For phylogenetic analysis, the protein sequences corresponding to either the leader peptide and first two Ig-like domains or the transmembrane and cytoplasmic regions of CD33-related Siglecs were aligned using ClustalW and then analysed with the PHYLIP 3.6 package [26], available at <http://evolution.genetics.washington.edu/phyliip.html>. Bootstrapping was performed with SEQBOOT and phylogenies estimated using the parsimony method with PROTPARS. A consensus unrooted phylogeny tree was obtained using CONSENSE and plotted with DRAWTREE.

Chromosomal localization

Metaphase spreads prepared from phytohaemagglutinin-stimulated human lymphocytes were hybridized with a biotinylated 3 kb insert from pHEOMH10 as described in [27]. Metaphase spreads from 50 cells were analysed.

Northern-blot analysis

Two human Multiple Tissue Northern (MTN) Blots containing approx. 2 μ g of poly(A)⁺ RNA per lane from various human

tissues were purchased from Clontech (Palo Alto, CA, U.S.A.) and hybridized with a ³²P-labelled insert from pHEOMH10 as described previously [15].

Cells

The following cell lines were provided by the ICRF Cell Production Service: COS-1, Balb/c 3T3 A31, Chinese hamster ovary K1 (CHO), KG1b, HL-60, U937, THP-1 and Daudi. The NK-like cell line, YT, was obtained from Dr Gillian Griffiths (Sir William Dunn School of Pathology, University of Oxford, Oxford, U.K.). Cells were cultured as described previously [17]. Human red blood cells (RBCs) were obtained from healthy donors and stored at 4 °C in Alsever's solution for up to 2 weeks. Human blood leucocytes were obtained from whole blood by dextran sedimentation followed by lysis of contaminating RBCs. Mononuclear fractions for flow cytometry were obtained by density-gradient centrifugation using Ficoll-Paque (Amersham Pharmacia Biotech).

Production of Fc proteins

Recombinant chimaeras containing the entire extracellular region (five domains) of Siglec-10 fused to the Fc region of human IgG1 (Siglec-10-Fc) were prepared by PCR amplification using p2.2 as the template with the following forward and reverse primers: 5'-ACAAGCTTTGCGCCTCCTATGCGG-AGATG-3' and 5'-CCCCTCGAGCTTATCTGGCAGCTGC-AGGAT-3'. The PCR product was cloned in-frame into the pIGplus vector, which encodes a Factor Xa cleavage site between the extracellular region and the Fc hinge region of human IgG1. Siglec-10-Fc was produced in transiently transfected COS cells and purified with Protein A-Sepharose as described in [13].

Generation of mouse polyclonal and monoclonal antibodies to Siglec-10

Balb/c 3T3 A31 cells were transfected by electroporation with Siglec-10 cDNA in the pcDNA3 vector as described in [17]. G418-resistant clones expressing Siglec-10 were identified by their ability to bind human RBCs. To generate polyclonal antibodies, Balb/c mice were immunized twice intraperitoneally, with an interval of 14 days, with 10⁷ live 3T3 cells expressing Siglec-10. After a final boost, the immune serum was collected, IgG-purified by Protein G-Sepharose affinity chromatography and passed over a Siglec-10-Fc column prepared by coupling 1.0 mg of purified Siglec-10-Fc to CNBr-activated Sepharose CL-4B. Bound IgG was eluted with 0.1 M glycine buffer, pH 2.5, and neutralized with 0.1 vol. of 1.0 M Tris, pH 8.0. To generate a mAb, Balb/c mice were immunized with Siglec-10-Fc and hybridomas generated by fusing immune spleen cells with the

SP2 myeloma following standard methods [28]. A positive well reacting specifically with Siglec-10-Fc was identified by ELISAs as described in [21]. The hybridoma was cloned three times by limiting dilution and the mAb designated 5G6 (IgG1). 5G6 was used as a tissue-culture supernatant in all experiments.

Human RBC-binding assays to Siglec-10 expressed on CHO cells and COS cells

CHO cells stably expressing Siglec-10 (Siglec-10-CHO) were generated by transfecting CHO cells with Siglec-10 cDNA in the pcDNA3 vector. G418-resistant clones expressing Siglec-10 were identified by their ability to bind the anti-Siglec-10 mAb 5G6. COS cells were transiently transfected with Siglec-10 cDNA by electroporation. RBC-binding assays with Siglec-10-transfected COS cells and Siglec-10-CHO cells were carried out as described in [13].

Binding assays with polyacrylamide glycoconjugates

COS-1 cells were transiently transfected by electroporation with cDNAs encoding Siglec-10 or CR1 (CD35) as a negative control and binding assays carried out 48–72 h later, as described previously [17]. Briefly, sialidase-treated and control cells were incubated with saturating concentrations (20 µg/ml) of biotinylated 2,3-PAA or 2,6-PAA for 1 h at room temperature, washed and incubated with ¹²⁵I-streptavidin at 0.5 µCi/ml for 1 h at 4 °C. Bound radioactivity was counted using a Beckman γ-counter.

FACS analysis

Single and double labelling of cells for flow cytometry were performed following standard protocols [28]. Following staining, cells were fixed in 2% formaldehyde and analysed on a Becton-Dickinson FACS analyser.

Immunoprecipitation

Wild-type CHO cells, Siglec-10-CHO and Daudi cells at 2×10^7 /ml were surface-biotinylated using Sulpho-NHS-biotin (Pierce) according to the manufacturer's suggestions and lysates prepared in 1% Triton X-100. Immunoprecipitations were carried out following standard procedures [28] and precipitated material was run on 4–12% gradient SDS/PAGE gels followed by Western blotting on to nitrocellulose. The blots were blocked and biotinylated proteins revealed using streptavidin-peroxidase followed by addition of the chemiluminescent ECL[®] reagent (Amersham Pharmacia Biotech).

RESULTS AND DISCUSSION

Identification and properties of Siglec-10

In a search for novel human Siglec cDNAs, 21 unique ESTs were identified in eight different cDNA libraries that encoded the same novel Siglec-like protein. The libraries were as follows (with frequencies of ESTs in each library shown in parentheses): eosinophils (10), primary dendritic cells (6), bone marrow (2), unstimulated B-cells (1), tonsils (1) and spleen (1). Inserts from four of the 21 clones were selected for sequencing but none was found to contain an intact open reading frame due to variable alternative splicing (results not shown). A full-length clone encoding this novel Siglec was successfully isolated from a human spleen cDNA library. Based on sequence similarity with other Siglecs (Figure 1) and its ability to bind sialic acid (see below) this protein has been designated Siglec-10. The extracellular region of Siglec-10 contains a hydrophobic signal peptide

and five Ig-like domains that are made up of an N-terminal V-set domain and four C2-set domains. There are five potential N-linked glycosylation sites. Following the transmembrane region, there is a cytoplasmic tail of 126 amino acids.

Siglec-10 contains all of the characteristic features of the Siglec subgroup of Ig superfamily proteins (Figure 1). These include the critical arginine at position 120 that has been shown in the ligand-bound crystal structure of sialoadhesin to interact with the carboxy group of sialic acid [2]. In addition, Siglec-10 contains two conserved aromatic residues, Phe²⁶ and Tyr¹³⁴, on the A and G strands of the V-set domain that in the case of sialoadhesin were shown to make important hydrophobic interactions with the N-acetyl and glycerol side chains of sialic acid [2]. Siglec-10 also exhibits the unusual pattern of cysteines in domains 1 and 2 that are thought to form intra-sheet and inter-domain disulphide bonds in Siglecs [2]. Within the cytoplasmic tail there are three potential tyrosine-based motifs. The most membrane proximal of these, Y⁵⁹⁷INV, is found within an extended region of the cytoplasmic tail that is not present in other CD33-related Siglecs (Figure 1). This motif fits the consensus sequence Y(L/V/I)N(V/P) for tyrosine-phosphorylated receptors that bind to the SH2 domain of Grb2 [29–31]. Grb2 is an adaptor molecule that is able to activate the Ras signalling pathway via interactions with the Ras-specific guanine nucleotide-exchange factor, Sos [32]. Although this motif has not been observed in CD33-related Siglecs previously, a similar motif (Y⁸²⁸ENV) is present in CD22/Siglec-2 where it has been shown to be capable of interacting with Grb2 following tyrosine phosphorylation of the CD22 cytoplasmic tail [33]. The interaction of CD22 with Grb2 was found recently to be important for the formation of a quaternary complex of CD22, SHIP, Grb2 and Shc implicated in the signalling functions of CD22 [34].

The other two tyrosine-based motifs in the cytoplasmic tail of Siglec-10 are more typical of the CD33-related Siglec subset (Figure 1). Although Siglec-8 was reported previously to differ significantly from other CD33-related Siglecs in the cytoplasmic tail [19], we have recently isolated alternatively spliced Siglec-8 cDNAs that encode a cytoplasmic tail that is very similar to those in the other CD33-related Siglecs (Figure 1). In Siglec-10, the motif LHY⁶⁶⁷ATL fits the ITIM consensus, (L/I/V/S)XYXX(L/V), and is similar to the corresponding motif in CD33/Siglec-3, LHY³⁴⁰ASL, which has been shown to be dominant in tyrosine phosphorylation and recruitment of SHP-1 [4,5,9] and SHP-2 [4,9]. Similar to other CD33-related Siglecs, the membrane distal motif, ADY⁶⁹¹AEV, does not fit the ITIM consensus but is well-conserved amongst the Siglecs (Figure 1). It has been suggested that the membrane distal tyrosine-based motif in some of the CD33-related Siglecs shares similarity with the binding site, TIYXX(V/I), on signalling lymphocyte-activation molecule (SLAM) for SLAM-associated protein (SAP) [16,18,22]. It is noteworthy that the membrane distal motif in Siglec-10, like that of Siglec-7 (Figure 1), conforms poorly to this consensus binding site. As yet there is no evidence that the membrane distal motif in any of the CD33-related Siglecs can bind SAP, but the corresponding phosphopeptide for CD33 has been shown to be capable of recruiting SHP-2 [4,9]. The presence of several potential tyrosine-based motifs in Siglec-10 suggests that this protein may be involved in signalling functions, but more studies are needed to explore this possibility.

Similarity with other CD33-related Siglecs and phylogenetic analysis

Database searches showed that Siglec-10 shares the highest sequence similarity with the CD33-related Siglec subset, namely

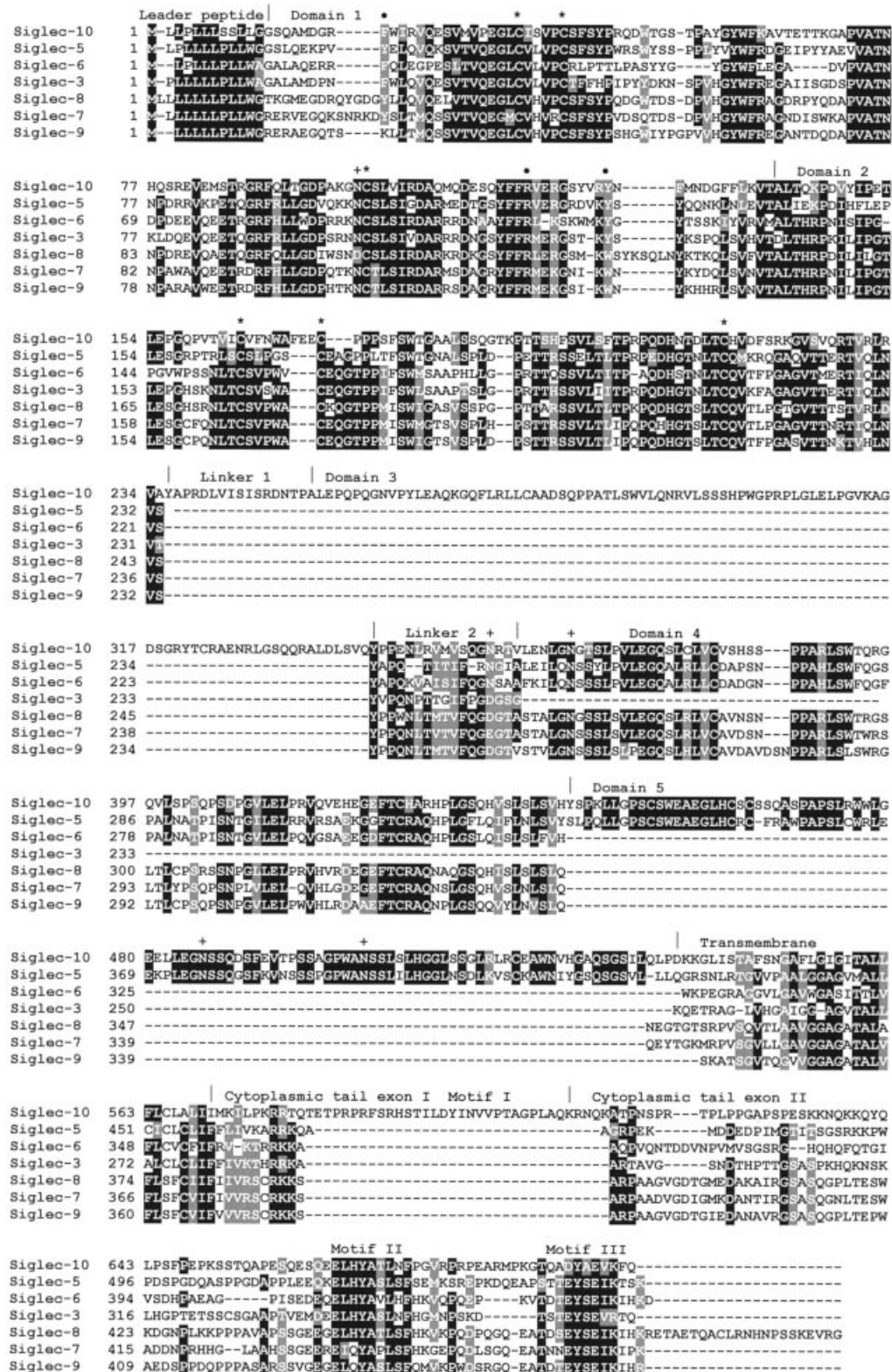


Figure 1 Predicted protein sequence of Siglec-10 and alignment with closely related Siglecs

Alignment was performed with the ClustalW multiple-sequence-alignment program and optimized by eye. Residues that are identical or similar in more than 70% of the aligned sequences are boxed, either in black (identical) or grey (similar). *, Positions of the cysteine residues characteristic of Siglecs; ●, residues important for sialic acid binding [2]; +, potential N-linked glycosylation sites in Siglec-10. Vertical lines indicate positions of intron-exon boundaries, as deduced from the sequence of the gene encoding Siglec-10. Positions of the domains, linkers, transmembrane region, cytoplasmic tail (enclosed by two exons) and the tyrosine-based motifs are indicated. GenBank[®] accession numbers are as follows. CD33/Siglec-3, M23197; Siglec-5, AF170484; Siglec-6, NM-001245; Siglec-7, AF170485; Siglec-8, AF195092; Siglec-9, AF247180; Siglec-10, AF310233; clone containing the Siglec-10 gene, CTD-2616J11. The extended cytoplasmic tail shown here for Siglec-8 is derived from a recently isolated cDNA clone (accession number AF310234).

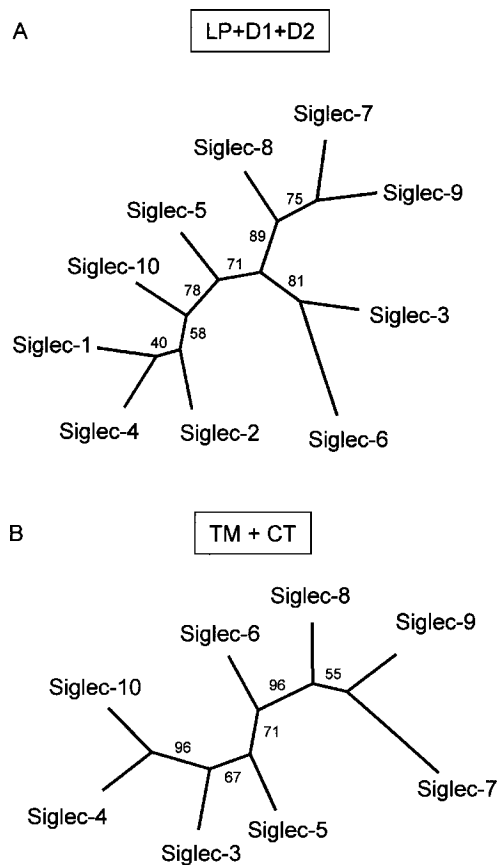


Figure 2 Phylogenetic analysis of CD33-related Siglecs

Amino acid sequences corresponding to either the leader peptide+domain 1+domain 2 (LP + D1 + D2; **A**) or the transmembrane and cytoplasmic tail regions (TM + CT; **B**) for each of the indicated Siglecs were aligned using the ClustalW multiple-sequence-alignment program and analysed for phylogenetic relationship using the PHYLIP 3.6 package. Unrooted phylogenies were constructed using the protein-sequence parsimony method. The values at the internodes indicate the number of times the adjacent forks occurred in 100 trees created by bootstrapping [46].

Siglecs 3, 5, 6, 7, 8 and 9. Within the first two Ig-like domains, Siglec-10 was found to be 40–48% identical to these proteins (Figure 1 and Table 1). So far, Siglec-10 is the least similar of the CD33-related Siglecs, which have been shown previously to share between 50 and 80% identity within the first two Ig-like domains [15–19,21]. The database searches also identified a human chromosome 19 clone (CTD-2616J11) from the completed human genome sequence that contains the entire Siglec-10 gene. This allowed us to map accurately the intron–exon boundaries within the cDNA sequence (Figure 1). This revealed that, besides an additional Ig-like domain (domain 3) not represented in any of the other CD33-related Siglecs (Figure 1), Siglec-10 has an extra linker region at the end of domain 2 encoded by a separate exon (Figure 1). This linker is similar in length (16 amino acids) and sequence (63% identity at the nucleotide level) to the linker at the end of domain 3 which is represented at the end of domain 2 in all other CD33-related Siglecs (Figure 1). Since domains 3 and 4 of Siglec-10 share 46% amino acid sequence identity, it seems likely that this region of the molecule arose by tandem duplication of two exons encoding a linker and an associated Ig-like domain (see Figure 3). Alignment of the Siglec-10 amino acid sequence with other CD33-related Siglecs

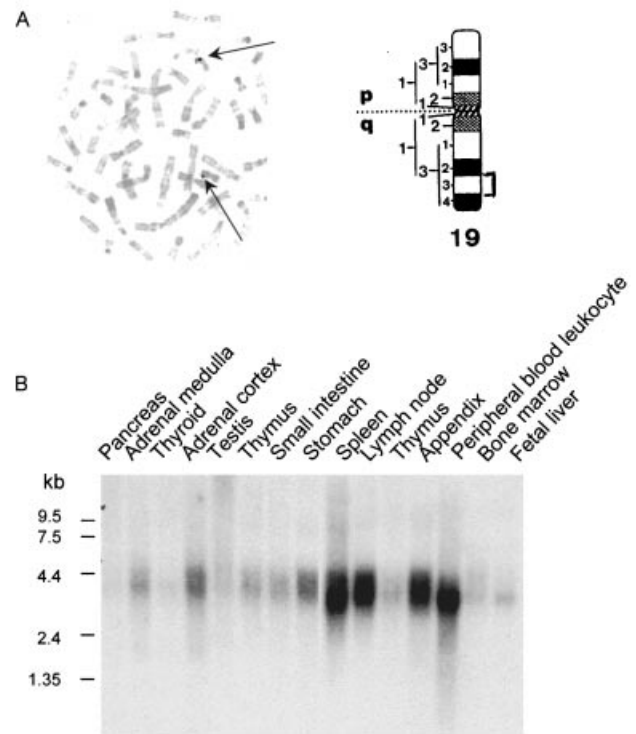


Figure 3 Localization and expression of the Siglec-10 gene

(**A**) Human lymphocyte metaphase spreads were hybridized with a 3 kb biotinylated insert from pHEOMH10 followed by fluorescein-avidin and the chromosomes counterstained with propidium iodide. The digital image is reversed to illustrate the hybridization signals (arrows) on the long arm of chromosome 19. The position of the Siglec-10 gene on chromosome 19q13.3 is also shown schematically. (**B**) Northern-blot analysis of Siglec-10 mRNA in human tissues. Each lane of the Multiple Tissue Northern (MTN) Blot (Clontech) contained approx. 2 μ g of poly(A)⁺ RNA from the tissue indicated and was normalized for levels of β -actin mRNA. A major form of Siglec-10 mRNA is seen at around 3 kb in most tissues with particularly high expression in spleen, lymph node, appendix and peripheral blood leucocytes.

showed that domain 3 of Siglecs 5, 6, 7, 8 and 9 is highly related to domain 4 of Siglec-10 (50–67% identity, Table 1). Domain 3 of Siglecs 5, 6, 7, 8 and 9 also aligned well with domain 3 of Siglec-10 (44–49% identity, Table 1). This further supports the notion that domains 3 and 4 of Siglec-10 arose through a duplication process. The alignment also showed that domain 5 of Siglec-10 was highly related to domain 4 of Siglec-5 (71% identity, Table 1). Interestingly, Siglec-10 also shared low, but significant similarity with MAG/Siglec-4 in all five Ig-like domains (Table 1).

To investigate further the possible evolutionary relationships between Siglec-10 and the other CD33-related Siglecs, unrooted phylogenetic trees were created using both the N-terminal regions (leader peptide, domain 1 + domain 2; Figure 2A) and C-terminal regions (transmembrane region + cytoplasmic tail; Figure 2B). All currently known human Siglecs were included in the N-terminal analysis but sialoadhesin/Siglec-1 and CD22/Siglec-2 were omitted from the C-terminal phylogenies since their cytoplasmic tails share no significant sequence similarity with the others. The N-terminal phylogenetic tree positioned Siglec-10 at an intermediate position between Siglecs 1, 2 and 4 and the other CD33-related Siglecs, which group together, as shown previously [22]. A similar overall pattern was observed with the C-terminal phylogenetic tree (Figure 2B).

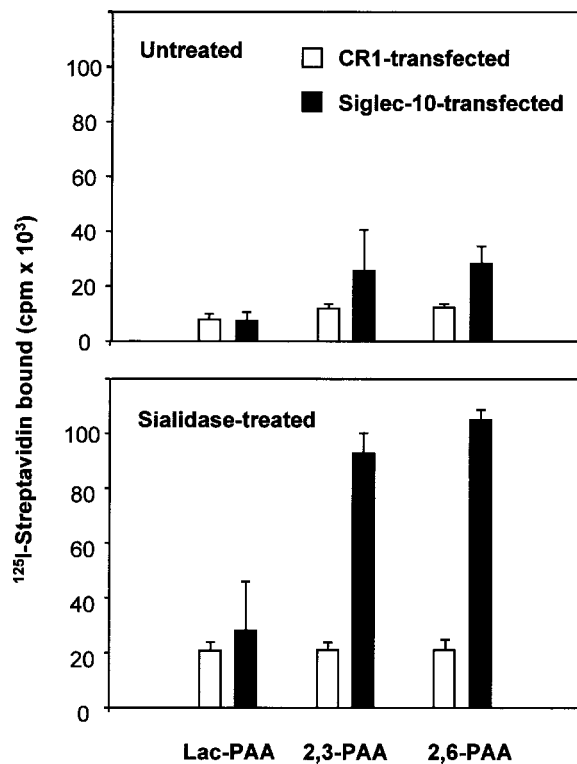


Figure 4 Binding of Siglec-10 expressed on COS cells to polyacrylamide conjugates

CR1 was included as a negative control to measure non-specific binding. After transient transfection (3 days), COS cells expressing the indicated proteins were incubated with biotinylated polyacrylamide (PAA) glycoconjugates linked to 3' sialyl-lactose (2,3-PAA), 6' sialyl-lactose (2,6-PAA) or lactose (Lac-PAA) at 20 $\mu\text{g}/\text{ml}$ or with buffer alone. Unbound conjugate was washed off and binding detected with ^{125}I -streptavidin. Data show means \pm S.D. from quadruplicates and are representative of three experiments.

Chromosomal localization and expression of the Siglec-10 gene

The Siglec-10 gene was mapped by *in situ* hybridization to the long arm of chromosome 19, in the 19q13.3 band (Figure 3A), closely linked to the other CD33-related Siglecs [15–17,19,22,35]. Northern-blot analysis (Figure 3B) revealed the presence of a major Siglec-10 mRNA transcript of ≈ 3.0 kb, with highest levels in spleen, lymph node, blood leucocytes and appendix. These tissues are rich in cells of haemopoietic origin, which suggests that Siglec-10 is expressed predominantly by these cells, similar to most other CD33-related Siglecs. Readily detectable mRNA could also be seen in several other tissues, except for pancreas, thyroid and testis, in which the mRNA was at low or undetectable levels (Figure 3B).

Siglec-10 mediates sialic acid-dependent binding to human RBCs and to glycoconjugate

To investigate the potential sialic acid-binding properties of Siglec-10, we initially performed binding assays in which native and sialidase-treated human RBCs were added to transiently transfected COS cells. Very little binding could be detected with untreated COS cells expressing Siglec-10. In contrast, following sialidase treatment, high levels of RBC binding were seen (results not shown). Similar results were obtained using CHO cells stably transfected with Siglec-10 (results not shown). Much of the

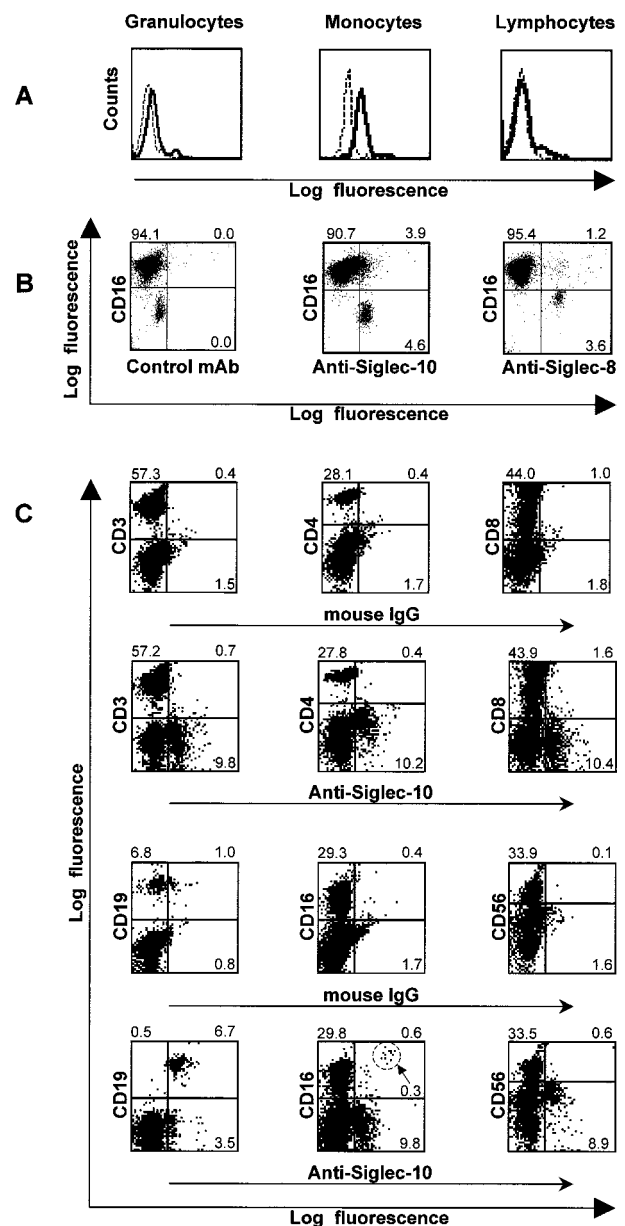


Figure 5 Expression of Siglec-10 on human peripheral blood leucocyte subsets

(A) FACS histograms showing expression of Siglec-10 on granulocytes, monocytes and lymphocytes, gated in each case according to their characteristic side- and forward-scatter properties. Thick lines show staining with 5G6 anti-Siglec-10 mAb and dashed lines show staining in the presence of an irrelevant isotype-matched mAb used as a negative control. (B) Double labelling of granulocytes with anti-CD16 (neutrophils) and anti-Siglec-10 mAbs compared with anti-Siglec-8 mAb. Anti-Siglec-10 shows clear labelling of the CD16-negative eosinophils and some of the neutrophils are also stained weakly. Values represent percentages of the total granulocytes analysed. (C) Double labelling of the lymphocyte population with antibodies to CD19 (B-cells), CD3 (pan T-cell), CD4 and CD8 (T-cell subsets), CD16 and CD56 (NK cells) and affinity-purified anti-Siglec-10 polyclonal antibody. Siglec-10 is expressed by most CD19⁺ B-cells and a small subset of CD16⁺ cells that do not express the CD56 NK-cell marker (circled). Values represent percentages of the total lymphocytes analysed.

increase in binding seen after treating Siglec-expressing cells with sialidase is thought to be due to unmasking the sialic acid-binding site on Siglecs from *cis*-interactions with sialic acids at the cell surface [14,15,17,19,21,22,36,37]. All binding of RBCs to

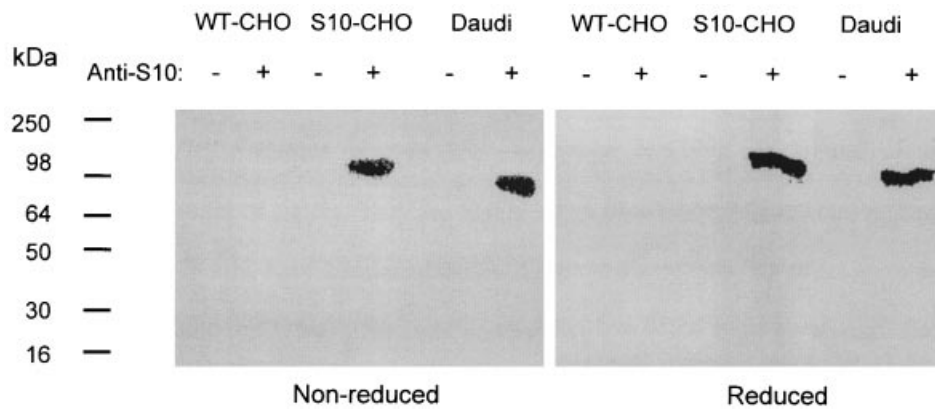


Figure 6 Molecular characterization of Siglec-10

Wild-type CHO cells (WT-CHO), stably transfected CHO cells expressing Siglec-10 (S10-CHO) or Daudi cells were surface-biotinylated, lysed and immunoprecipitations performed without or with the mouse anti-Siglec-10 polyclonal antibody (anti-S10). Precipitates were run under either non-reducing or reducing conditions on 4–12% gradient SDS/polyacrylamide gels, transferred to nitrocellulose and probed with horseradish peroxidase-conjugated streptavidin followed by ECLTM reagent. Siglec-10 migrates as a single monomeric species at ≈ 120 kDa from Siglec-10-CHO cells and at ≈ 100 kDa from Daudi cells.

sialidase-treated cells expressing Siglec-10 was abolished when the RBCs were pretreated with sialidase, demonstrating that the binding was sialic acid-dependent (results not shown).

To determine the sialic acid linkage preference of Siglec-10, binding assays were carried out with synthetic polyacrylamide conjugates, carrying either 3' or 6' sialyl-lactose or lactose. In these experiments, COS cells were transiently transfected with Siglec-10 or CR1 as a negative control. FACS analysis showed that 20–30% of the cells expressed each molecule three days after the transfection (results not shown). Transfected cells were either untreated or treated with sialidase immediately before the binding assay to remove potentially inhibitory sialic acids in the COS cell glycocalyx (Figure 4). Very little binding of PAA conjugates could be observed with untreated cells, but after sialidase treatment, Siglec-10-transfected COS cells bound strongly to glycoconjugates carrying sialic acid in either the $\alpha 2,3$ or $\alpha 2,6$ linkage (Figure 4). No specific binding was observed with lactose-PAA used as a negative control.

Expression of Siglec-10 on human peripheral blood leucocytes

An affinity-purified mouse polyclonal antibody was prepared to the extracellular region of Siglec-10 and shown by FACS analysis not to cross-react with Siglecs 3, 5, 7 and 8 expressed on CHO cells (results not shown). A mAb, 5G6, was also raised against Siglec-10 and shown not to cross-react with Siglecs 3, 5, 7, 8 and 9 expressed on CHO cells (results not shown). A detailed analysis of the expression of Siglec-10 was carried out by FACS analysis, using human peripheral blood leucocytes. Indistinguishable results were obtained using either the polyclonal or monoclonal antibodies. First, expression on granulocyte, monocyte and lymphocyte subsets was compared (Figure 5A). This revealed low levels of Siglec-10 on monocytes and on subsets of granulocytes and lymphocytes. Since Siglec-10 transcripts had been identified frequently in an eosinophil cDNA library, we asked whether Siglec-10 was expressed normally on eosinophils within the granulocyte population. Using anti-CD16 mAb to distinguish eosinophils from neutrophils, the CD16-negative eosinophils were found to stain specifically with 5G6 (Figure 5B). To date, Siglec-8 is the only other Siglec that has been shown to be expressed on eosinophils, although in this case its expression

is highly restricted [19,20]. It is evident that Siglec-8 is expressed at higher levels on eosinophils than Siglec-10 (Figure 5B), but further studies are needed to determine the potential roles of these Siglecs in eosinophil biology and explore the possibility that there is functional redundancy due to the co-expression of these two related Siglecs.

To characterize the lymphocyte-reactive cells in more detail, double labelling was carried out in which staining for Siglec-10 was combined with staining for CD3 (pan T-cell), CD4 and CD8 (T-cell subsets), CD19 (pan B-cell), CD16 and CD56 (NK cells; Figure 5C). The dominant population expressing Siglec-10 was made up of CD19⁺ B-cells that were labelled weakly. Higher levels of Siglec-10 were seen on a very small subset ($\approx 0.5\%$) of cells in the lymphocyte gate that expressed high levels of CD16 but no detectable CD56 (Figure 5C). A similar pattern of Siglec-10 staining was observed in eight independent experiments with different donors, using either the polyclonal or monoclonal antibody. In a previous study on Siglec-9 [21], the same population of CD16⁺CD56⁻ cells was shown to be labelled at a level similar to that shown here. Although these cells constitute only a very minor subset of the total blood leucocytes, it is striking that they express two different Siglecs at relatively high levels. Currently the nature and potential functions of these cells are unclear, but they may correspond to a minor population of CD16⁺CD56⁻ cells that were identified previously in human blood and shown to have the morphological features of NK cells and exhibit low levels of natural cytotoxicity [38,39]. CD16⁺CD56⁻ NK-like cells have been shown to be enriched in fetal cord blood compared with adult blood and it has been speculated that they may represent an earlier stage of NK-cell differentiation than typical CD16⁺CD56⁺ NK cells [40]. Further studies are needed to examine the expression of Siglecs on cord-blood leucocytes and explore the potential physiological significance of Siglec expression on this interesting population of CD16⁺CD56⁻ cells.

Finally, FACS staining of various human cell lines was also performed. Consistent with the staining pattern with blood leucocytes, weak positive labelling was observed with the U937 pro-monocytic cell line and the Daudi B-cell line. No staining was seen with HL-60 (myelomonocytic), THP-1 (monocytic) or YT (NK-like) cells (results not shown).

Molecular characterization of Siglec-10

To investigate the molecular mass of Siglec-10 expressed at the cell surface, Siglec-10-CHO and Daudi cells were surface-biotinylated and immunoprecipitations performed using the affinity-purified anti-Siglec-10 antibody. Specific bands were obtained for both Siglec-10-CHO cells and Daudi cells but not for wild-type CHO cells used as a control (Figure 6). In both cases, a single band of ≈ 100 –120 kDa was observed under both reducing and non-reducing conditions, demonstrating that Siglec-10 exists as a monomer in the plasma membrane (Figure 6). Since the observed molecular mass was considerably greater than that predicted from the amino acid sequence of the mature protein (74 560 Da), it is likely that most or all of the five potential N-linked glycosylation sites are occupied by carbohydrate. Furthermore, the slightly slower migration observed with Siglec-10 precipitated from Siglec-10-CHO cells compared with Daudi cells is likely to reflect differential glycosylation.

Conclusions and perspectives

The results presented here with Siglec-10 extend our previous work characterizing novel human Siglecs related to CD33/Siglec-3. Similar to most other CD33-related Siglecs, Siglec-10 is found on cells of the innate immune system, namely eosinophils, monocytes and a poorly characterized subset of NK-like cells that express high levels of CD16 without the CD56 marker. It is possible that CD33-related Siglecs evolved to lower activation thresholds of cells within the innate immune system, analogous to the well-established role of CD22/Siglec-2 in modulating B-cell activation mediated by the B-cell receptor (reviewed in [41]). Sialic acid-dependent ligation of the CD33-related Siglecs could, in principle, provide a mechanism that prevents reactivity towards host cells while allowing effective killing of non-sialylated pathogens.

Based on searches of genomic databases, it has been proposed recently that Siglecs co-evolved with sialic acid-biosynthesis pathways [22]. Genomic analysis of the protostome-lineage species *Caenorhabditis elegans* and *Drosophila melanogaster* failed to uncover obvious Siglec homologues or sequences encoding known enzymes of the sialic acid-biosynthesis pathway [22]. Sialic acid biosynthesis is thought to have evolved in deuterostome-lineage animals such as starfish [42,43], but whether Siglecs are present in these species is currently unknown. However, MAG/Siglec-4 has been identified immunochemically in a wide variety of higher non-mammalian species including fish, frogs, birds, snakes and lizards [44,45]. Given the apparent relatedness of Siglec-10 to MAG, it will be interesting to determine if a Siglec-10 orthologue is expressed in these species and, if so, whether more recently evolved Siglecs such as Siglecs 7, 8 and 9 are also present. Future genome-sequencing projects of both vertebrate and invertebrate species should be instrumental in providing answers to these questions.

We are grateful to Maggie Chambers for help with immunization, Stuart Dubock for help with the polyacrylamide-conjugate-binding assays and to Lars Nitschke for discussions. We thank Ajit Varki, Pascale Gagneux and Takashi Angata for helpful advice on phylogenetic analyses. This work was supported by the Wellcome Trust.

REFERENCES

- Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J. et al. (1998) Siglecs: a family of sialic-acid binding lectins. *Glycobiology* **8**, v
- 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 Å resolution. *Mol. Cell* **1**, 719–728
- Nitschke, L., Carsetti, R., Ocker, B., Kohler, G. and Lamers, M. C. (1997) CD22 is a negative regulator of B-cell receptor signalling. *Curr. Biol.* **7**, 133–143
- 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, CD33, associates with the protein-tyrosine phosphatases, SHP-1 and SHP-2. *J. Biol. Chem.* **274**, 11505–11512
- Ulyanova, T., Blasioli, J., Woodford-Thomas, T. A. and Thomas, M. L. (1999) The sialoadhesin CD33 is a myeloid-specific inhibitory receptor. *Eur. J. Immunol.* **29**, 3440–3449
- Falco, M., Biondi, 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**, 793–802
- Vitale, C., Romagnani, C., Falco, M., Ponte, M., Vitale, M., Moretta, A., Bacigalupo, A., Moretta, L. and Mingari, M. C. (1999) Engagement of p75/AIRM1 or CD33 inhibits the proliferation of normal or leukemic myeloid cells. *Proc. Natl. Acad. Sci. U.S.A.* **96**, 15091–15096
- Ferlazzo, G., Spaggiari, G. M., Semino, C., Melioli, G. and Moretta, L. (2000) Engagement of CD33 surface molecules prevents the generation of dendritic cells from both monocytes and CD34+ myeloid precursors. *Eur. J. Immunol.* **30**, 827–833
- Paul, S. P., Taylor, L. S., Stansbury, E. K. and McVicar, D. W. (2000) Myeloid specific human CD33 is an inhibitory receptor with differential ITIM function in recruiting the phosphatases SHP-1 and SHP-2. *Blood* **96**, 483–490
- Crocker, P. R., Mucklow, S., Bouckson, V., McWilliam, A., Willis, A. C., Gordon, S., Milon, G., Kelm, S. and Bradfield, P. (1994) Sialoadhesin, a macrophage sialic acid binding receptor for haemopoietic cells with 17 immunoglobulin-like domains. *EMBO J.* **13**, 4490–4503
- Powell, L. D., Sgroi, D., Sjöberg, E. R., Stamenkovic, I. and Varki, A. (1993) Natural ligands of the B cell adhesion molecule CD22 beta carry N-linked oligosaccharides with alpha-2,6-linked sialic acids that are required for recognition. *J. Biol. Chem.* **268**, 7019–7027
- Sgroi, D., Varki, A., Braesch-Andersen, S. and Stamenkovic, I. (1993) CD22, a B cell-specific immunoglobulin superfamily member, is a sialic acid-binding lectin. *J. Biol. Chem.* **268**, 7011–7018
- Kelm, S., Pelz, A., Schauer, R., Filbin, M. T., Tang, S., de Bellard, M. E., Schnaar, R. L., Mahoney, J. A., Hartnell, A., Bradfield, P. and Crocker, P. R. (1994) Sialoadhesin, myelin-associated glycoprotein and CD22 define a new family of sialic acid-dependent adhesion molecules of the immunoglobulin superfamily. *Curr. Biol.* **4**, 965–972
- 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
- 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
- Patel, N., Linden, E. C., 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
- 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**, 34089–34095
- Angata, T. and Varki, A. (2000) Siglec-7: a sialic acid-binding lectin of the immunoglobulin superfamily. *Glycobiology* **10**, 431–438
- Floyd, H., Ni, J., Cornish, A. L., Zeng, Z., Liu, D., Carter, K. C., Steel, J. and Crocker, P. R. (2000) Siglec-8. A novel eosinophil-specific member of the immunoglobulin superfamily. *J. Biol. Chem.* **275**, 861–866
- Kikly, K. K., Bochner, B. S., Freeman, S. D., Tan, K. B., Gallagher, K. T., D'Alessio, K. J., Holmes, S. D., Abrahamson, J. A., Erickson-Miller, C. L., Murdock, P. R. et al. (2000) Identification of SAF-2, a novel siglec expressed on eosinophils, mast cells, and basophils. *J. Allergy Clin. Immunol.* **105**, 1093–1100
- Zhang, J. Q., Nicoll, G., Jones, C. and Crocker, P. R. (2000) Siglec-9, a novel sialic acid binding member of the immunoglobulin superfamily expressed broadly on human blood leukocytes. *J. Biol. Chem.* **275**, 22121–22126
- Angata, T. and Varki, A. (2000) Cloning, characterization, and phylogenetic analysis of siglec-9, a new member of the CD33-related group of siglecs. Evidence for co-evolution with sialic acid synthesis pathways. *J. Biol. Chem.* **275**, 22127–22135
- 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

- 24 Hartnell, A., Steel, J., Turley, H., Jones, M., Jackson, D. G. and Crocker, P. R. (2001) Characterization of human sialoadhesin, a sialic acid binding receptor expressed by resident and inflammatory macrophage populations. *Blood* **97**, 288–296
- 25 Huang, X. and Miller, W. (1991) A time-efficient, linear-space local similarity algorithm. *Adv. Appl. Math.* **12**, 337–357
- 26 Felsenstein, J. (1989) PHYLIP - Phylogeny Inference Package (version 3.2). *Cladistics* **5**, 164–166
- 27 Ollendorff, V., Mattei, M., Fournier, E., Adelaide, J., Lopez, M., Rosnet, O. and Birnbaum, D. (1998) A third human CBL gene is on chromosome 19. *Int. J. Oncol.* **13**, 1159–1161
- 28 Harlow, E. and Lane, D. (1988) Monoclonal antibodies. In *Antibodies: A Laboratory Manual*, pp. 139–243. Cold Spring Harbor Press, Cold Spring Harbor
- 29 Skolnik, E. Y., Lee, C. H., Batzer, A., Vicentini, L. M., Zhou, M., Daly, R., Myers, Jr., M. J., Backer, J. M., Ullrich, A., White, M. F. et al. (1993) The SH2/SH3 domain-containing protein GRB2 interacts with tyrosine-phosphorylated IRS1 and Shc: implications for insulin control of ras signalling. *EMBO J.* **12**, 1929–1936
- 30 Ogura, K., Tsuchiya, S., Terasawa, H., Yuzawa, S., Hatanaka, H., Mandiyan, V., Schlessinger, J. and Inagaki, F. (1999) Solution structure of the SH2 domain of Grb2 complexed with the Shc-derived phosphotyrosine-containing peptide. *J. Mol. Biol.* **289**, 439–445
- 31 Rahuel, J., Gay, B., Erdmann, D., Strauss, A., Garcia-Echeverria, C., Furet, P., Caravatti, G., Fretz, H., Schoepfer, J. and Grutter, M. G. (1996) Structural basis for specificity of Grb2-SH2 revealed by a novel ligand binding mode. *Nat. Struct. Biol.* **3**, 586–589
- 32 Buday, L. and Downward, J. (1993) Epidermal growth factor regulates p21ras through the formation of a complex of receptor, Grb2 adapter protein, and Sos nucleotide exchange factor. *Cell* **73**, 611–620
- 33 Yohannan, J., Wienands, J., Coggeshall, K. M. and Justement, L. B. (1999) Analysis of tyrosine phosphorylation-dependent interactions between stimulatory effector proteins and the B cell co-receptor CD22. *J. Biol. Chem.* **274**, 18769–18776
- 34 Poe, J. C., Fujimoto, M., Jansen, P. J., Miller, A. S. and Tedder, T. F. (2000) CD22 forms a quaternary complex with SHIP, Grb2, and Shc. A pathway for regulation of B lymphocyte antigen receptor-induced calcium flux. *J. Biol. Chem.* **275**, 17420–17427
- 35 Peiper, S. C., Ashmun, R. A. and Look, A. T. (1988) Molecular cloning, expression, and chromosomal localization of a human gene encoding the CD33 myeloid differentiation antigen. *Blood* **72**, 314–321
- 36 Braesch-Andersen, S. and Stamenkovic, I. (1994) Sialylation of the B lymphocyte molecule CD22 by alpha 2,6-sialyltransferase is implicated in the regulation of CD22-mediated adhesion. *J. Biol. Chem.* **269**, 11783–11786
- 37 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. U.S.A.* **95**, 7469–7474
- 38 Lanier, L. L., Corliss, B. C., Wu, J., Leong, C. and Phillips, J. H. (1998) Immunoreceptor DAP12 bearing a tyrosine-based activation motif is involved in activating NK cells. *Nature (London)* **391**, 703–707
- 39 McKenzie, R. S., Simms, P. E., Helfrich, B. A., Fisher, R. I. and Ellis, T. M. (1992) Identification of a novel CD56-lymphokine-activated killer cell precursor in cancer patients receiving recombinant interleukin 2. *Cancer Res.* **52**, 6318–6322
- 40 Gaddy, J. and Broxmeyer, H. E. (1997) Cord blood CD16+56- cells with low lytic activity are possible precursors of mature natural killer cells. *Cell. Immunol.* **180**, 132–142
- 41 Cyster, J. G. and Goodnow, C. C. (1997) Tuning antigen receptor signaling by CD22: integrating cues from antigens and the microenvironment. *Immunity* **6**, 509–517
- 42 Schauer, R. (2000) Biochemistry of sialic acid diversity. In *Carbohydrates in Chemistry and Biology Part II: Biology of Saccharides*, vol. 3 (Ernst, B., Hart, G. W. and Sinary, P., eds.), pp. 227–243. Wiley-VCH, Weinheim
- 43 Kelm, S. and Schauer, R. (1997) Sialic acids in molecular and cellular interactions. *Int. Rev. Cytol.* **175**, 137–240
- 44 O'Shannessy, D. J., Willison, H. J., Inuzuka, T., Dobersen, M. J. and Quarles, R. H. (1985) The species distribution of nervous system antigens that react with anti-myelin-associated glycoprotein antibodies. *J. Neuroimmunol.* **9**, 255–268
- 45 Tropak, M. B., Jansz, G. F., Abramow-Newerly, W. and Roder, J. C. (1995) Conservation of functionally important epitopes on myelin associated glycoprotein (MAG). *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* **112**, 345–354
- 46 Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**, 783–791

Received 20 October 2000/2 January 2001; accepted 8 February 2001