Biochemical and Biophysical Research Communications **278,** 775–781 (2000)
doi:10.1006/bbrc.2000.3866, available online at http://www.idealibrary.com on IDEAL®

Molecular Characterization of a Siglec8 Variant Containing Cytoplasmic Tyrosine-Based Motifs, and Mapping of the Siglec8 Gene

George Foussias,*,† George M. Yousef,*,† and Eleftherios P. Diamandis*,†,[1]

*Department of Pathology and Laboratory Medicine, Mount Sinai Hospital, Toronto, Ontario, Canada; and
†Department of Laboratory Medicine and Pathobiology, University of Toronto, Toronto, Ontario, Canada

Received October 23, 2000

Through efforts to investigate the CD33-like subgroup of sialic acid binding immunoglobulin-like lectins (Siglecs), which are believed to be located on chromosome 19q13.4, we have identified the precise genomic region containing the Siglec8 gene. It is located on chromosome 19q13.4, approximately 330 kb downstream of the Siglec9 gene. Further, we have identified a novel Siglec8 variant, named Siglec8-Long (Siglec8-L), which differs in its last two exons from the previously published mRNA sequence of Siglec8 (GenBank Accession No. AF195092). Both Siglec8 and Siglec8-L are comprised of seven exons, of which the first five are identical, followed by marked differences in exon usage and mRNA splicing. The 499 amino acid protein encoded by the Siglec8-L open reading frame has a molecular weight of 54 kDa. Like the other members of the CD33-like subgroup of Siglecs, except for the previously published Siglec8, Siglec8-L also contains the two tyrosine-based motifs that have been found to recruit both SH2 domain-containing tyrosine and inositol phosphatases.

Key Words: **siglec gene family; CD33-like subgroup; immunoglobulin superfamily; sialoadhesins; immunoreceptor tyrosine kinase inhibition motif (ITIM); alternative splicing; gene mapping.**

Abbreviations used: Siglec, sialic acid binding immunoglobulin-like lectin; Siglec8-L, Siglec8-Long; RT-PCR, reverse transcription-polymerase chain reaction; IgSF, immunoglobulin superfamily; EST, expressed sequence tag; ITIM, immunoreceptor tyrosine kinase inhibitory motif; RACE, rapid amplification of cDNA ends; SLAM, signaling lymphocyte activation molecule; SAP, SLAM-associated protein; SH2, src homology 2; SHP, SH2 domain-containing protein tyrosine phosphatase; SHIP, SH2 domain-containing inositol phosphatase.

[1] To whom correspondence should be addressed at Mount Sinai Hospital, Department of Pathology and Laboratory Medicine, 600 University Avenue, Toronto, Ontario M5G 1X5, Canada. Fax: 416-586-8628. E-mail: ediamandis@mtsinai.on.ca.

Sialic acid-binding immunoglobulin-like lectins (Siglecs) are a family of recently discovered type 1 transmembrane proteins belonging to the immunoglobulin superfamily (IgSF) (1). In addition to their type 1 transmembrane topology, these proteins are characterized by the presence of one N-terminal V-set Ig-like domain, a variable number of downstream C2-set domains, and the ability to bind sialic acid in glycoproteins and glycolipids (2). So far there have been nine Siglec family members described in humans, each with its own unique expression pattern: Siglec1 (sialoadhesin) expressed on macrophages (3); Siglec2 (CD22) on B lymphocytes (4); Siglec3 (CD33) on myeloid progenitor cells and monocytes (5); Siglec4a (myelin-associated glycoprotein (MAG)) on oligodendrocytes and Schwann cells (6); Siglec5 on neutrophils (7); Siglec6 on B lymphocytes (8); Siglec7 (p75/AIRM1) on natural killer cells (9–11); Siglec8 on eosinophils (12); and Siglec9 on neutrophils, monocytes, and various lymphocytes (13–15).

The members of the Siglec family are highly homologous in their extracellular, Ig-like, domains, particularly for the CD33-like subgroup, which includes Siglec3, 5, 6, 7, 8, and 9 (13, 15). This subgroup of genes has been localized to human chromosome 19q13.3-13.4, and it has been suggested that it may be the result of relatively recent gene duplication and exon shuffling (14, 15). In addition to the highly homologous extracellular domains, all the CD33-like Siglecs except Siglec8 show conservation of two cytoplasmic tyrosine-based motifs (12). The first of these contains the consensus sequence for the immunoreceptor tyrosine kinase inhibitory motif (ITIM), (ILV)xYxx(LV) (x being any amino acid) (16, 17). This motif has been shown to be the binding site for the SH2 (src homology 2) domains of the SH2 domain-containing protein tyrosine phosphatases SHP-1 and SHP-2 (18, 19), as well as the SH2 domain-containing inositol phosphatases SHIP1 and SHIP2 (20). The second motif displays homology to a

0006-291X/00 $35.00

tyrosine-based motif, TxYxx(IV), identified in the signaling lymphocyte activation molecule (SLAM) and was found responsible for its association with the SLAM-associated protein (SAP), which in turn blocks the binding of SHP-2 to phosphorylated SLAM (21, 22).

As mentioned previously, of the CD33-like subgroup of Siglecs, only Siglec8 does not possess the characteristic two tyrosine-based motifs (12). In the present study, we describe the identification and characterization of a novel variant of Siglec8, termed Siglec8-Long (Siglec8-L), which possesses these two tyrosine-based motifs, consistent with findings for the other members of the CD33-like subgroup of Siglecs. Further, we have determined the genomic organization of the Siglec8 gene, finding that it is comprised of seven exons, and is localized approximately 330 kb downstream of Siglec9 on chromosome 19q13.4. We have further identified the exact nucleotide at which the Siglec8 and Siglec8-L mRNAs deviate, producing proteins of 431 and 499 amino acids, respectively. The difference arises in exon six, where the Siglec8-L mRNA has a splice site after 97 nucleotides followed by a unique exon seven, while the Siglec8 mRNA continues right through exon 6 and then into a further downstream exon seven.

MATERIALS AND METHODS

Identification of the genomic area containing Siglec8. Genomic DNA sequences derived from BAC clones covering chromosome 19q13.4 were identified and obtained from the Lawrence Livermore National Laboratory (LLNL) Human Genome Center. These sequences were compared to the mRNA sequence for Siglec8 (GenBank Accession No. AF195092), which has been reported to be linked to this area (12), using the BLASTN nucleotide alignment tool (23). In addition, genomic regions found to match Siglec8 were also analyzed by the Grail exon prediction program (24), in order to determine the existence of any new Siglecs, as well as possible additional exons for Siglec8. Prediction results were compared to the human EST database by the BLAST alignment tool (23). Further, the genomic region containing Siglec8 was localized to a specific region of chromosome 19q13.4 through the aid of the WebCutter restriction analysis program and comparison of the fragments to the previously published *Eco*R1 map for chromosome 19q13.4 (25).

Molecular characterization of Siglec8-L. Based on the results of exon prediction and the known sequence of the Siglec8 mRNA, we designed PCR primers to determine the sequence of the Siglec8-L mRNA, as well as to confirm the published Siglec8 mRNA sequence, both through RT-PCR. The primers used were: S8-Forward (common), ACAAGTGACACTGGCAGCAG; S8-L-Reverse, AGCTGAGGGTTGCATAATGG; S8-L-Reverse2, TACTGCATAGCATGGGGCTC; S8-Reverse, AGAAGAGCAGGGGAAACCAC. The fetal liver cDNA used was prepared as described elsewhere (13). The PCR conditions were as follows: 2.5 units of HotStarTaq polymerase (Qiagen, Valencia, CA), 1X PCR buffer with 1.5 mM $MgCl_2$ (Qiagen), 1 μL cDNA, 200 μM dNTPs (deoxynucleoside triphosphates), and 250 ng of each primer, using the Mastercycler gradient thermocycler (Eppendorf Scientific Inc., Westbury, NY). The temperature profile was: denaturation at 95°C for 15 min followed by 94°C for 30 s, annealing at 60°C for 30 s, and extension at 72°C for 1 min for a total of 35 cycles, followed by a final extension at 72°C for 10 min. The PCR product was subjected to electrophoresis on a 2% agarose gel containing ethidium bromide. The product bands were then extracted from the gel, and the purified DNA was directly sequenced using an automated sequencer.

We also utilized Marathon-Ready fetal liver cDNA (Clontech, Palo Alto, CA) to perform nested 3′-RACE in order to verify the 3′ end of the Siglec8 mRNA. The procedure was carried out according to the manufacturer's instructions, with some minor modifications. Briefly, the first round of the 3′ RACE reaction utilized the forward gene-specific primer (GSP1) which is identical to the above mentioned Siglec8-Forward (common) and the provided adapter primer AP1. The nested 3′ RACE reaction was carried out using GSP2 (CCTTCCTGTCCTTCTGCATC), and the provided AP2. The touchdown PCR method was utilized as recommended by the manufacturer, with a slight temperature profile modification. The annealing temperatures used were: 70°C for 4 min for 5 cycles, then 68°C for 4 min for 5 cycles, followed by 66°C for 4 min for 25 cycles. The denaturation temperature was set to 94°C for 5 s for every cycle, with an extension temperature of 72°C for 1 min After all cycles were completed a final extension at 72°C for 10 min was performed. The reaction was carried out using 2.5 units HotStarTaq (Qiagen, Valencia, CA), 1× PCR buffer with 1.5 mM $MgCl_2$ (Qiagen), 200 μM dNTPs, primer concentrations of 200 ng (GSP1 and 2) and 1 μM (AP1 and 2), 1 μL Marathon-Ready cDNA, or 5 μL of the first round 3′ RACE product (50 μL total volume). All 3′ RACE reactions were performed using the Perkin Elmer GeneAmp PCR System 9600 thermocycler (Perkin Elmer, Norwalk, CT). The products of both rounds of the RACE reaction were subjected to electrophoresis on a 2% agarose gel containing ethidium bromide, with any bands evident being extracted and the DNA directly sequenced with an automated sequencer.

Based on RT-PCR and 3′ RACE results, as well as the initial alignment of the Siglec8 mRNA species to the genomic sequence covering chromosome 19q13.4, we proceeded to map the exons of the Siglec8 gene. This was achieved through the use of the BLASTN nucleotide alignment tool, which enabled us to localize the mRNA sequences to specific regions of genomic DNA.

Following final characterization of Siglec8-L, the primary structure of the protein encoded by the Siglec8-L mRNA was determined. This was compared to the published protein sequence for Siglec8, as well as to all the members of the CD33-like subgroup of Siglecs using the CLUSTALX multiple alignment tool (26).

RESULTS

Identification of the Genomic Area Containing Siglec8

The CD33-like subgroup of Siglecs has been mapped, through various means, to chromosome 19q13.4. The Siglec9 gene has recently been characterized and localized to this area (13), immediately following the end of the kallikrein gene family. Further, the mRNA species for the remainder of this subgroup have also been characterized, and mapped to this region primarily through fluorescence *in situ* hybridization (FISH), as well as somatic cell hybridization (7–10, 12). During our examination of this area, we identified a clone, CTD-3073N11, from the CalTech Human BAC library D, that contained the Siglec8 gene. Upon exon prediction analysis of this genomic region, we identified two putative exons at the 3′ terminus of the Siglec8 gene which differed from the previously published mRNA sequence. One of these exons was much shorter in length and the other was not present at all. We proceeded to search the human EST database, both with the published mRNA sequence, as well as with the two putative exons, and were unable to find any significant matches.

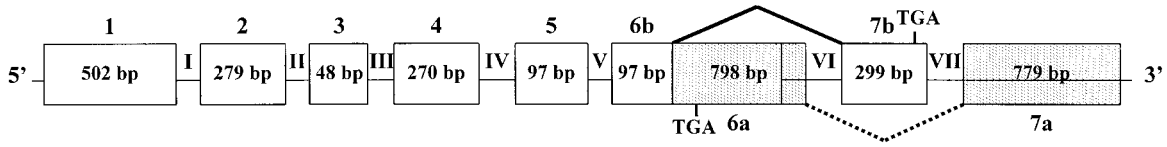


FIG. 1. Genomic organization of the Siglec8 gene. Based on our own experimental results, as well as the previously published sequence for the Siglec8 mRNA (GenBank Accession No. AF195092), we have established that both Siglec8 and our Siglec8-L are composed of seven exons (Arabic numerals). The two mRNA species are identical until exon six, where the Siglec8 mRNA contains the whole of exon six (6a and 6b) and continues to exon 7a, indicated by the broken line, while the Siglec8-L mRNA contains only exons 6b and 7b, shown by the solid line. The location of the stop codon (TGA) is shown for both mRNA species, and differs due to a change in the open reading frame. Splice sites are conserved (-mGT. . .AGm-, where m is any base) in Siglec8-L between exons 6b and 7b, but not between exons 6a and 7a in the Siglec8 sequence reported by Floyd *et al.* (12).

Based on the sequence information from the clone on which Siglec8 is localized, we used the WebCutter restriction analysis tool to determine the size of the *Eco*R1 fragments produced. By comparing these results with the available *Eco*R1 restriction map for chromosome 19q13.4, we determined that the Siglec8 gene is located in the more centromeric region of 19q13.4, and is approximately 330 kb downstream from the Siglec9 gene.

Molecular Characterization of Siglec8-L

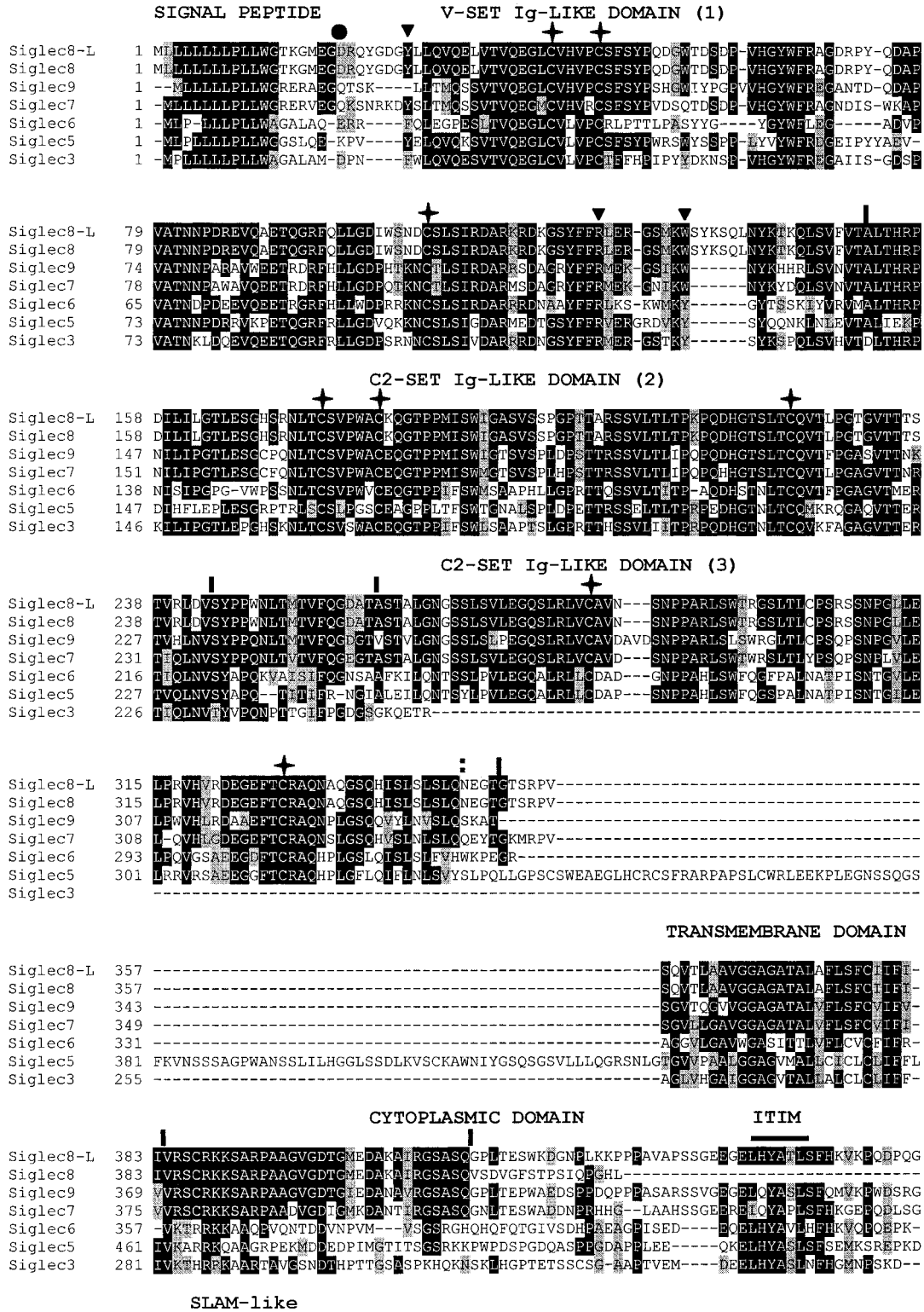
Using RT-PCR and primers derived from sequence shared by both Siglec8 and Siglec8-L, as well as from the 3′ termini of Siglec8 and the putative Siglec8-L, we were able to confirm the existence of both putative exons mentioned above. Through automated sequencing and subsequent alignment of the mRNA sequence with the genomic sequence for our BAC clone, we were able to determine the exact genomic organization for these last two exons. However, using the PCR primers specific for the published sequence of Siglec8 we obtained a very faint band, after agarose gel electrophoresis, which we were unable to sequence. In conjunction with 3′ RACE and the Marathon-Ready fetal liver cDNA we identified additional sequence at the 3′ terminus of Siglec8-L, including both the termination codon and a portion of the 3′ untranslated region. The nested 3′ RACE reaction produced a major product with an approximate size of 300 bp, as well as a much fainter high molecular weight band. Attempts to sequence the latter band were unsuccessful.

Based on the results obtained through the initial alignment, as well as from the characterization of the last two exons of Siglec8-L, we were able to characterize the genomic organization of the Siglec8 gene. As is shown in Fig. 1, both Siglec8 and Siglec8-L are comprised of seven exons. The first five exons are identical in both Siglec8 and Siglec8-L, with lengths of 502, 279, 48, 270, and 97 bp. The first exon of 502 bp contains a 5′ untranslated region of 48 nucleotides, with the possibility that there is even more upstream sequence. For Siglec8-L, exon six is 97 nucleotides long, while exon seven contains at least 299 bp, of which 252 bp code for amino acid residues, and 44 bp being part of the 3′ untranslated region. Exons six and seven of Siglec8, on the other hand, are 895 bp and 779 bp long, respectively. Examining the splice donor and acceptor sites for each of these exons, we noticed that the first five exons, as well as exons six and seven of Siglec8-L, all possessed sequences closely related to the consensus splice sites (-m**GT**AAGT. . .C**AG**m-, where m is any base) (27). However, in the case of Siglec8, these splice sites were not present. Examination of the open reading frame of Siglec8-L revealed a 499 amino acid residue protein with a molecular weight of 54.04 kDa, excluding any posttranslational modifications which may be present. The sequence of Siglec8-L is identical to that of Siglec8 until residue 415, after which Siglec8-L contains a sequence homologous to the C-terminal sequences of the CD33-like subgroup of Siglecs, including the two tyrosine-based motifs (Fig. 2).

DISCUSSION

Through efforts to investigate the CD33-like subgroup of Siglecs, we have identified the area of chromosome 19q13.4 which contains the Siglec8 gene, located approximately 330 kb downstream of the recently published Siglec9 gene (13). Examination of this area revealed the existence of an alternative form of Siglec8, named Siglec8-L. The protein product of this mRNA species differs markedly from the previously published Siglec8 (GenBank Accession No. AF195092) at its C-terminus (12). Siglec8-L is a 499 residue protein with a molecular weight of 54 kDa. It is encoded by seven exons, and unlike Siglec8, it shows a high degree of homology at its C-terminus to the other members of the CD33-like subgroup of Siglecs. Consistent with its inclusion in this subfamily of Siglecs, Siglec8-L also possesses the two characteristic tyrosine-based motifs.

The Siglec8 mRNA published by Floyd *et al.* (2000) contains an abbreviated C-terminus, lacking the characteristic tyrosine-based motifs reported in other members of the CD33-like subgroup of Siglecs (12). Further, based on our characterization of its genomic structure, the Siglec8 mRNA species reported by Floyd *et al.*


SIGNAL PEPTIDE
V-SET Ig-LIKE DOMAIN (1)
C2-SET Ig-LIKE DOMAIN (2)
C2-SET Ig-LIKE DOMAIN (3)
TRANSMEMBRANE DOMAIN
CYTOPLASMIC DOMAIN
ITIM
SLAM-like
Siglec8-L
Siglec8
Siglec9
Siglec7
Siglec6
Siglec5
Siglec3

contains approximately 1.5 kb of untranslated sequence at its 3′ end. By comparison, and in keeping with the hypothesis that this subgroup of Siglecs arose through gene duplication relatively recently in vertebrate evolution (14, 15), none of the other Siglecs that belong to this subgroup have such an extensive 3′ untranslated region, or any untranslated exons at the 3′ end. Furthermore, the intron–exon splice sites for the last two exons of Siglec8 (based on our genomic sequence and the mRNA sequence of Floyd *et al.*) are not consistent with the characteristic splice donor and acceptor sites (27), unlike the first five exons and the two exons identified by us (Fig. 1).

The identification of Siglec8 by Floyd *et al.* (2000) was accomplished through the use of an EST that showed homology to CD33 (12). We have not as yet been able to identify this clone, and believe that the discrepancy between our sequence and theirs may be due to the inclusion of genomic sequence in the EST. Wolfsberg and Landsman have reported that they have indeed found several instances in which ESTs contained intron sequences or unusual splice sites (28). If this is the case with Siglec8, the EST used by Floyd *et al.* (12) may represent a partially spliced or incorrectly spliced sequence, which results in the inclusion of two pieces of genomic DNA (not present in any mRNA species we were able to identify) as well as the absence of an entire exon. During their identification of Siglec8, the authors report an unsuccessful attempt of Northern blot analysis for Siglec8 using a probe derived from the coding and 3′ untranslated sequences of their mRNA species. Based on the above noted problems with ESTs, the lack of specific bands in the Northern blot may be due to such intronic sequence contamination. This contamination likely resulted in a change in the open reading frame of the Siglec8 mRNA which causes premature termination and loss of the two tyrosine-based motifs. Despite efforts to identify the mRNA species reported by Floyd *et al.* (12), we have been unsuccessful in obtaining any PCR product that could be sequenced. However, we can still not rule out the possibility that this mRNA species does in fact represent an alternatively spliced form that lacks the two tyrosine motifs and is present in tissues not tested by us.

During their characterization of Siglec8, the authors demonstrated eosinophil-specific expression for this member of the CD33-like subgroup of Siglecs (12). This was accomplished by generating monoclonal antibodies to the extracellular domains of Siglec8. Given that Siglec8 and Siglec8-L are identical in their extracellular domains, the expression of Siglec8-L is also expected to be restricted to eosinophils.

For a few of the other members of the Siglec family, alternative splice forms have been described. For CD22, two isoforms have been identified in humans, with either four or six C2-set Ig-like domains (4, 29, 30). In the mouse there have been three isoforms of MAG identified (31, 32). One of the isoforms lacks an untranslated exon in the 5′ end of the mRNA, which has no effect on the size of the resultant polypeptide. The two other forms differ in the presence of exon 12, which is 45 nucleotides in length and introduces a termination codon when included in the mRNA. In humans, however, there has been no report of any MAG isoforms. Further, two isoforms for CD33 have been reported in the mouse, which differ by an 83 nucleotide in-frame insertion in the cytoplasmic domain (33). Human CD33 has also been reported to exist as two different size transcripts, which is believed to be through the use of alternative polyadenylation signals, with no change in the size of the polypeptide (34). Therefore, although there is evidence of alternative splicing in members of the Siglec family, it appears to occur primarily in nonhuman members, with the exception of the alternative number of Ig-like domains for human CD22. However, even these cases do not compare to the drastic differences seen in the splicing patterns of Siglec8 and Siglec8-L.

The Siglec family of transmembrane proteins, and in particular the CD33-like subgroup, is a very recently described member of the IgSF (1, 14, 15). There has been only limited progress into the investigation of the biological function of these proteins, with the exception of Siglec7, and more recently CD33. Siglec7, identified initially by Falco *et al.* (9), was found to be tyrosine-phosphorylated in its ITIM motif, resulting in recruitment of SHP1 with a consequent inhibition of natural killer cell cytotoxicity. More recently, it has been reported that engagement of Siglec7 and CD33, through

FIG. 2. Protein sequence alignment for the CD33-like subgroup of Siglecs. The sequence of Siglec8-L was aligned to those of Siglec8, as well as the remaining members of the CD33-like subgroup of Siglecs, using the ClustalX multiple alignment tool (26). The solid vertical lines indicate the positions of the exon boundaries. In all but one case, shown by the broken vertical line, the exon boundaries match those found for Siglec9 (13). The conserved cysteine residues responsible for the intra- and interdomain disulfide bonds (2, 37, 38) are indicated by the star (★), while the triangles (▼) denote the aromatic residues believed to be important for sialic acid binding, based on findings for Siglec1 (sialoadhesin) (39). The signal peptide cleavage site for Siglec8-L, indicated by the solid circle (●), was predicted using the SignalP program (40). The Ig-like domain assignments, as well as those for the transmembrane and cytoplasmic domains, are based on previous reports (12) and the one domain–one exon rule (41). The positions of the two tyrosine-based motifs, ITIM and SLAM-like, are indicated. The GenBank Accession Nos. are as follows: Siglec8-L, AF287892; Siglec8, AF195092; Siglec7,: AF170485; Siglec6, NM_001245; Siglec5, NM_003830; Siglec3 (CD33), M23197.

the use of monoclonal antibodies, inhibits the proliferation of both normal and leukemic myeloid cells *in vitro* (35). These effects are believed to be the result of phosphorylation of the two tyrosine-based motifs present in the cytoplasmic domain of both CD33 and Siglec7. In addition, Taylor *et al.* (36) have found that CD33 recruits the protein-tyrosine phosphatases SHP1 and SHP2, both *in vitro* and *in vivo*, and is the result of tyrosine phosphorylation in the ITIM motif. Mutation of the tyrosine in this ITIM motif of CD33 resulted in increased red blood cell binding by CD33-expressing COS cells. These findings suggest that, in addition to the recruitment of SHP1 and 2 inhibiting the activating signaling pathways that lead to cell proliferation and survival, this recruitment may also modulate the receptor's ligand-binding activity (36). It is quite likely, given the high degree of homology within the CD33-like subgroup of Siglecs, that the remainder of this group, including Siglec8-L, play a similar inhibitory role in their respective cell types.

REFERENCES

1. Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le Douarin, N., Powell, L., Roder, J., Schnaar, R. L., Sgroi, D. C., Stamenkovic, K., Schauer, R., Schachner, M., van den Berg, T. K., van der Merwe, P. A., Watt, S. M., and Varki, A. (1998) Siglecs: A family of sialic-acid binding lectins [letter]. *Glycobiology* **8,** v.
2. Crocker, P. R., Kelm, S., Hartnell, A., Freeman, S., Nath, D., Vinson, M., and Mucklow, S. (1996) Sialoadhesin and related cellular recognition molecules of the immunoglobulin superfamily. *Biochem. Soc. Trans.* **24,** 150–156.
3. 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.
4. Stamenkovic, I., and Seed, B. (1990) The B-cell antigen CD22 mediates monocyte and erythrocyte adhesion. *Nature* **345,** 74–77.
5. 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.
6. Kelm, S., Schauer, R., Manuguerra, J. C., Gross, H. J., and Crocker, P. R. (1994) Modifications of cell surface sialic acids modulate cell adhesion mediated by sialoadhesin and CD22. *Glycoconj. J.* **11,** 576–585.
7. 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.
8. Patel, N., Brinkman-Van der 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.
9. Falco, M., Biassoni, R., Bottino, C., Vitale, M., Sivori, S., Augugliaro, R., Moretta, L., and Moretta, A. (1999) Identification and molecular cloning of p75/AIRM1, a novel member of the sialoadhesin family that functions as an inhibitory receptor in human natural killer cells. *J. Exp. Med.* **190,** 793–802.
10. 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.
11. Angata, T., and Varki, A. (2000) Siglec-7: A sialic acid-binding lectin of the immunoglobulin superfamily. *Glycobiology* **10,** 431–438.
12. 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.
13. Foussias, G., Yousef, G. M., and Diamandis, E. P. (2000) Identification and molecular characterization of a novel member of the Siglec family (SIGLEC9). *Genomics* **67,** 171–178.
14. 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.
15. 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.
16. Burshtyn, D. N., Yang, W., Yi, T., and Long, E. O. (1997) A novel phosphotyrosine motif with a critical amino acid at position -2 for the SH2 domain-mediated activation of the tyrosine phosphatase SHP- 1. *J. Biol. Chem.* **272,** 13066–13072.
17. Vivier, E., and Daeron, M. (1997) Immunoreceptor tyrosine-based inhibition motifs. *Immunol. Today* **18,** 286–291.
18. Borges, L., Hsu, M. L., Fanger, N., Kubin, M., and Cosman, D. (1997) A family of human lymphoid and myeloid Ig-like receptors, some of which bind to MHC class I molecules. *J. Immunol.* **159,** 5192–5196.
19. Le Drean, E., Vely, F., Olcese, L., Cambiaggi, A., Guia, S., Krystal, G., Gervois, N., Moretta, A., Jotereau, F., and Vivier, E. (1998) Inhibition of antigen-induced T cell response and antibody-induced NK cell cytotoxicity by NKG2A: Association of NKG2A with SHP-1 and SHP-2 protein-tyrosine phosphatases [published erratum appears in *Eur. J. Immunol.* (1998) Mar. **28**(3):1122]. *Eur. J. Immunol.* **28,** 264–276.
20. Muraille, E., Bruhns, P., Pesesse, X., Daeron, M., and Erneux, C. (2000) The SH2 domain containing inositol 5-phosphatase SHIP2 associates to the immunoreceptor tyrosine-based inhibition motif of Fc gammaRIIB in B cells under negative signaling. *Immunol. Lett.* **72,** 7–15.
21. Coffey, A. J., Brooksbank, R. A., Brandau, O., Oohashi, T., Howell, G. R., Bye, J. M., Cahn, A. P., Durham, J., Heath, P., Wray, P., Pavitt, R., Wilkinson, J., Leversha, M., Huckle, E., Shaw-Smith, C. J., Dunham, A., Rhodes, S., Schuster, V., Porta, G., Yin, L., Serafini, P., Sylla, B., Zollo, M., Franco, B., Bentley, D. R., *et al.* (1998) Host response to EBV infection in X-linked lymphoproliferative disease results from mutations in an SH2-domain encoding gene [see comments]. *Nat. Genet.* **20,** 129–135.
22. Sayos, J., Wu, C., Morra, M., Wang, N., Zhang, X., Allen, D., van Schaik, S., Notarangelo, L., Geha, R., Roncarolo, M. G., Oettgen, H., De Vries, J. E., Aversa, G., and Terhorst, C. (1998) The X-linked lymphoproliferative-disease gene product SAP regulates signals induced through the co-receptor SLAM [see comments]. *Nature* **395,** 462–469.
23. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402.
24. Murakami, K., and Takagi, T. (1998) Gene recognition by combination of several gene-finding programs. *Bioinformatics* **14,** 665–675.

25. Ashworth, L. K., Batzer, M. A., Brandriff, B., Branscomb, E., de Jong, P., Garcia, E., Garnes, J. A., Gordon, L. A., Lamerdin, J. E., Lennon, G., *et al.* (1995) An integrated metric physical map of human chromosome 19. *Nat. Genet.* **11,** 422–427.

26. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G., and Gibson, T. J. (1998) Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23,** 403–405.

27. Iida, Y. (1990) Quantification analysis of 5′-splice signal sequences in mRNA precursors. Mutations in 5′-splice signal sequence of human beta-globin gene and beta-thalassemia. *J. Theor. Biol.* **145,** 523–533.

28. Wolfsberg, T. G., and Landsman, D. (1997) A comparison of expressed sequence tags (ESTs) to human genomic sequences. *Nucleic Acids Res.* **25,** 1626–1632.

29. Wilson, G. L., Fox, C. H., Fauci, A. S., and Kehrl, J. H. (1991) cDNA cloning of the B cell membrane protein CD22: A mediator of B-B cell interactions. *J. Exp. Med.* **173,** 137–146.

30. Wilson, G. L., Najfeld, V., Kozlow, E., Menniger, J., Ward, D., and Kehrl, J. H. (1993) Genomic structure and chromosomal mapping of the human CD22 gene. *J. Immunol.* **150,** 5013–5024.

31. Fujita, N., Sato, S., Kurihara, T., Kuwano, R., Sakimura, K., Inuzuka, T., Takahashi, Y., and Miyatake, T. (1989) cDNA cloning of mouse myelin-associated glycoprotein: A novel alternative splicing pattern. *Biochem. Biophys. Res. Commun.* **165,** 1162–1169.

32. Fujita, N., Sato, S., Kurihara, T., Inuzuka, T., Takahashi, Y., and Miyatake, T. (1988) Developmentally regulated alternative splicing of brain myelin-associated glycoprotein mRNA is lacking in the quaking mouse. *FEBS Lett.* **232,** 323–327.

33. Tchilian, E. Z., Beverley, P. C., Young, B. D., and Watt, S. M. (1994) Molecular cloning of two isoforms of the murine homolog of the myeloid CD33 antigen. *Blood* **83,** 3188–3198.

34. Simmons, D., and Seed, B. (1988) Isolation of a cDNA encoding CD33, a differentiation antigen of myeloid progenitor cells. *J. Immunol.* **141,** 2797–2800.

35. 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. USA* **96,** 15091–15096.

36. 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.

37. Pedraza, L., Owens, G. C., Green, L. A., and Salzer, J. L. (1990) The myelin-associated glycoproteins: Membrane disposition, evidence of a novel disulfide linkage between immunoglobulin-like domains, and posttranslational palmitylation. *J. Cell Biol.* **111,** 2651–2661.

38. Williams, A. F., Davis, S. J., He, Q., and Barclay, A. N. (1989) Structural diversity in domains of the immunoglobulin superfamily. *Cold Spring Harb. Symp. Quant. Biol.* **54,** 637–647.

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

40. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int. J. Neural Syst.* **8,** 581–599.

41. Williams, A. F., and Barclay, A. N. (1988) The immunoglobulin superfamily—domains for cell surface recognition. *Annu. Rev. Immunol.* **6,** 381–405.