AI85

# Siglecs

Ajit Varki
*University of California, San Diego, California, USA*

**Sialic acid recognizing Ig-superfamily Lectins (Siglecs) are a major subset of the "I-type lectins." The latter are defined as animal proteins other than antibodies that can mediate carbohydrate (glycan) recognition via immunoglobulin(Ig)-like domains. Siglecs share characteristic amino-terminal structural features that are involved in their sialic acid-binding properties, and can be broadly divided into two groups: an evolutionarily conserved subgroup (Siglecs-1, -2, and -4) and a CD33/Siglec 3-related subgroup (Siglecs -3 and -5 to -11). While the precise functions of Siglecs are unknown, they seem to send inhibitory signals to the cells that express them, in response to recognition events on cell surfaces.**

## Historical Background and Definition

Sialic acids (Sias) are a family of nine-carbon acidic sugars that typically occupy a terminal position on glycan chains attached to the cell surface of "higher" animals. The immunoglobulin superfamily (IgSf) is an evolutionarily ancient group of proteins whose appearance predated the emergence of the immunoglobulins themselves. Until the 1990s, it was assumed that IgSf members (other than some antibodies) did not mediate carbohydrate recognition. Independent work on CD22 (eventually Siglec-2, a protein on mature resting B cells) and on sialoadhesin (Sn, eventually Siglec-1, a protein on certain macrophage subsets) revealed that their first Ig V-set-like domains could mediate Sia recognition. Homologous features of this V-set Ig-like domain and the adjacent C2-set domain then led to the discovery that two other previously cloned molecules—CD33 (eventually Siglec-3) and Myelin-associated Glycoprotein (MAG, eventually Siglec-4)—also had Sia-binding properties. Following consultation among all researchers working on these proteins, the common name "Siglec" and a numbering system were agreed upon. Criteria for inclusion of IgSf-related proteins as Siglecs are: (1) the ability to recognize sialylated glycans; and (2) significant sequence similarity within the N-terminal V-set and adjoining C2-set domains. Evaluation of the human and mouse genomes eventually defined 11 human and 8 mouse molecules that fulfill these criteria. Since humans have more Siglecs than mice and cloning of the mouse molecules initially lagged behind, the primary numbering system is based on the human molecules.

## Two Broad Subgroups of Siglecs

While Siglecs -1, -2, and -4 appear to be evolutionarily rather conserved, the CD33/Siglec-3-related subgroup (Human Siglecs -3 and -5 to -11) appear to be rapidly evolving. Some CD33/Siglec-3-related Siglecs appear to have evolved as hybrids of pre-existing genes and/or by gene conversion. For these reasons, sequence comparisons alone do not allow the conclusive designation of the orthologue status of all mouse genes, and additional features such as gene position and exon structure must be taken into account. Until such issues are resolved, some mouse Siglecs have been assigned a temporary alphabetical designation.

## Common Structural Features

All are single-pass Type 1 integral membrane proteins with extra-cellular domains consisting of uniquely similar N-terminal V-set Ig domains, followed by variable numbers of C2-set Ig domains, ranging from 16 in Sn/Siglec-1 to 1 in CD33/Siglec-3. Crystal structures for mouse Siglec-1 and human Siglec-7 indicate that the V-set immunoglobulin-like fold has several unusual features, including an intra-beta sheet disulphide and a splitting of the standard beta strand G into two shorter strands. These features along with certain key amino acid residues appear to be requirements for Sia recognition. In particular, a conserved arginine residue is involved in a salt bridge with the carboxylate of Sia in all instances studied to date.

## Cell-Type Specific Expression

With the exception of MAG/Siglec-4 and Siglec-6, expression appears to be confined to the hematopoeitic

*Encyclopedia of Biological Chemistry, Volume 4.* 

during evolution. Thus, there has been ample time for genetic variation and divergence to occur in the genetic encoding for the receptor proteins. Whatever the process of receptor diversity, the multiple receptor signaling has considerable biological significance. For example, inputs to the dorsal raphe nucleus are integrated into global 5-HT release, via both synaptic and volume transmission, that in turn can modulate multiple and diverse neuronal functions based on receptor subtype. Thus, one transmitter (5-HT) with multiple receptors/multiple pathways can introduce levels of complexity, yet specific physiological responses, that are localized to a given brain region or neural system. These responses can be cell-type specific or even cell-compartment specific. Furthermore, a level of fidelity is introduced based on the receptor's relative affinity for 5-HT. The exact nature of this diversity is at present unclear however the physiological implications are quite apparent, when considering the diverse role of these receptors in many physiological functions and disease states.

## See Also the Following Articles

Adenylyl Cyclases • Neurotransmitter Transporters • Phospholipase C • RNA Editing

## Glossary

**autoreceptor** A receptor on the neuronal cell body or presynaptic terminal can regulate its own cell firing and/or neurotransmitter release and synthesis.

**constitutive activity** Ability of a receptor to activate second-messenger pathways without the binding of an external ligand.

**heteroreceptor** A presynaptic receptor that regulates the release of neurotransmitter other than its own natural ligand.

**promiscuous coupling** The ability of a receptor to couple to more than one signal cascade.

**RNA editing** A process whereby the nucleotide sequence of RNA transcripts is chemically altered.

## Further Reading

Aghajanian, G. K., and Sanders-Bush, E. (2002). Serotonin. In *Psychopharmacology: The Fifth Generation of Progress* (K. L. Davis, D. Charney, J. T. Coyle and C. Nemeroff, eds.) Lippincott, Williams, and Wilkins, Philadelphia.

Barnes, N. M., and Sharp, T. (1999). A review of central 5-HT receptors and their function. *Neuropharmacology* **38**, 1083–1152.

Hoyer, D., Clarke, D. E., Fozard, J. R., Hartig, P. R., Martin, G. R., Mylecharane, E. J., Saxena, P. R., and Humphrey, P. P. A. (1994). VII. International Union of Pharmacology classification of receptors for 5-hydroxytrytamine (serotonin). *Pharmacol. Rev.* **46**, 157–203.

Meltzer, H. Y. (1999). The role of serotonin in antipsychotic drug action. *Neuropsychopharmacology* **21**, 106S–115S.

Raymond, J. R., Mukhin, Y. V., Gelasco, A., Turner, J., Collinsworth, G., Gettys, T. W., Grewal, J. S., and Garnovskaya, M. N. (2001). Multiplicity of mechanisms of serotonin receptor signal transduction. *Pharmacol. Ther.* **92**, 179–212.

## Biography

Elaine Sanders-Bush is a Professor of Pharmacology and Psychiatry at Vanderbilt University School of Medicine, Nashville, Tennessee. She earned a Ph.D. in pharmacology at Vanderbilt in 1967. Her research focuses on serotonin receptors, applying a multidisciplinary approach to define the role of signal transduction molecules and posttranscriptional and posttranslational modifications that alter receptor function. Dr. Sanders-Bush received a Merit Award from the National Institute of Mental Health in recognition of her research accomplishments.

Paul Gresch is a Postdoctoral Fellow in Dr. Sanders-Bush's laboratory. He earned a Ph.D. in cellular and clinical neurobiology from Wayne State University in Detroit, Michigan in 1999.

and immune systems. Within these systems each Siglec is expressed in a cell-type specific fashion, suggesting that each may be involved in discrete functions. However, systematic studies of Siglec expression outside the hematopoeitic system and during development have not yet been done.

## Genomic Organization and Phylogeny

Based on probing for the canonical functional amino acids in the V-set domain of the typical Siglec, there is no evidence for Siglec-like molecules in prokaryotes, fungi or plants, nor in animals of the Protostome lineage, including organisms for which the complete genome is available. In contrast, it is relatively easy to find Siglec-like V-set domains in many vertebrate taxa (Sia recognition by fish and reptile Siglec candidates has not been formally shown as yet). While the relatively conserved Siglecs (-1, -2, and -4) have clear-cut single orthologues that are easy to identify in various species, the remaining "CD33/Siglec-3 related" Siglecs appear to have been evolving rapidly. Most of the latter genes are clustered together in a ~500 kb region on human chromosome 19q13.3–13.4.

## Siglec Recognition of Sialic Acids and Their Linkages

The first two Siglecs discovered (Sn/Siglec-1 and CD22/Siglec-2) had strikingly different binding properties for sialosides—with Sn preferring alpha2–3 linked targets and CD22 being highly specific for alpha2–6 linkages. In the latter case, the binding affinity was in the low micromolar range. MAG/Siglec-4 also has an extended binding site that is even highly specific for the underlying sugar chain. There is also variable preference for certain types of sialic acids, with Sn and MAG not tolerating the common N-glycolyl modification of Sias. However the CD33/Siglec-3-related Siglecs are more promiscuous in their preferences for different types and linkages of Sias. Of course, many of the less common linkages and types of sialic acids have not been studied for Siglec recognition. The Golgi enzymes that are potential regulators of Siglec functions are primarily the sialyltransferases, and to some extent the enzymes which modify sialic acids. Some Siglecs show preferences for certain macromolecular ligands e.g., CD45 for CD22/Siglec-2, the mucins CD43, and Muc-1 for Sn/Siglec-1, and certain brain glycolipids for MAG/Siglec-4.

## Potential Effects of Neu5Gc Loss on Human Siglec Biology

The most common Sias of mammalian cells are N-acetylneuraminic acid (Neu5Ac) and N-glycolylneuraminic acid (Neu5Gc). Humans are an exception, because of a mutation in CMP-sialic acid hydroxylase, which occurred after the time (~5–7 Ma) when we shared a common ancestor with great apes. The resulting loss of Neu5Gc and increase in Neu5Ac in humans could have potentially altered the biology of the Siglecs. For example, human cells have a higher density of Sn/Siglec-1 ligands than great apes, the distribution of Sn-positive macrophages in humans is different, and a much larger fraction of human macrophages is positive. Other emerging evidence suggests that there are further human-specific changes in Siglec biology that may be related to the loss of Neu5Gc.

## Masking and Unmasking of Siglecs Binding Sites on Cell Surfaces

The initial assumption was that Siglecs were involved in intercellular adhesion. However, in most instances, their binding sites appear to be masked by Sias on the same cell surfaces on which they are expressed. Of course, external ligands with very high affinity/avidity may still compete for the endogenous masking ligands. There is also some evidence that unmasking can occur under certain conditions, but it is not known if this is biologically relevant. Overall, the significance of Siglec masking is unclear at this time.

## Signaling Motifs in Cytosolic Tails

The CD33-related Siglecs have conserved tyrosine residues in the cytosolic tails, one of which corresponds to a canonical immunoreceptor tyrosine-based inhibition motif (ITIM). Various *in vitro* manipulations of these receptors indicate that these tyrosines are indeed targets for phosphorylation, and that they can modulate signaling events by recruiting certain tyrosine phosphatases. However, the true *in vivo* biological functions of these signaling motifs remain obscure. Another major unresolved question is: what is the connection between extra-cellular sialic acid recognition and signaling via the cytosolic motifs?

## Known and Putative Functions of the Siglecs

Various lines of evidence indicate that MAG/Siglec-4 is involved in the maintenance of myelin organization

and in the inhibition of neurite outgrowth during regeneration after injury. It is also reasonably clear that CD22/Siglec-2 functions as an inhibitory component of the antigen receptor complex of B Cells, and is thus involved in regulating the humoral immune response. While Sn/Siglec-1 appears to mediate various macrophage adhesion events *in vitro* and *in vivo*, it is as yet unclear what the functions of these interactions are. Little is known about the functions of CD33-related Siglecs. It has been suggested that these molecules are involved in innate immunity. One hypothesis currently being tested is that Siglecs may be sensors for pathogens that have sialylated cell surfaces and/or express extra cellular sialidases.

## See Also the Following Articles

Immunoglobulin (Fc) Receptors • Lectins • Polysialic Acid in Molecular Medicine

## Glossary

**immunoglobulin superfamily (IgSf)** Proteins that have modules homologous to those of antibodies (immunoglobulins). This is an evolutionarily ancient group of proteins whose appearance actually predated the emergence of the immunoglobulins themselves.

**I-type lectins** Proteins (other than antibodies) in which immunoglobulin-like modules mediate binding to glycans (sugar chains).

**sialic acids** These acids are a diverse family of nine-carbon acidic sugars that typically occupy a terminal position on glycan chains attached to the cell surface of "higher" animals of the deuterostome lineage.

**siglecs** A major subset of the I-type lectins. Name is based on their defining properties, as sialic acid recognizing IgSf lectins.

## Further Reading


Angata, T., and Brinkman-Van der Linden, E. (2002). I-type lectins. *Biochim. Biophys. Acta* **1572**, 294.

Angata, T., and Varki, A. (2002). Chemical diversity in the sialic acids and related alpha-keto acids: An evolutionary perspective. *Chem. Rev.* 102, 439–470.

Crocker, P. R., and Varki, A. (2001). Siglecs, sialic acids and innate immunity. *Trends Immunol.* 22, 337–342.

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.

Crocker, P. R., Clark, E. A., Filbin, M., Gordon, S., Jones, Y., Kehrl, J. H., Kelm, S., Le, D. 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, M., and Varki, A. (1998). Siglecs: A family of sialic-acid binding lectins [letter]. *Glycobiology* 8(v2).

Kelm, S., Pelz, A., Schauer, R., Filbin, M. T., Tang, S., De, B. 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.

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

Powell, L. D., Sgroi, D., Sjoberg, E. R., Stamenkovic, I., and Varki, A. (1993). Natural ligands of the B cell adhesion molecule CD22beta carry N-linked oligosaccharides with alpha-2, 6-linked sialic acids that are required for recognition. *J. Biol. Chem.* 268, 7019–7027.


## Biography

Ajit Varki is Professor of Medicine and Cellular and Molecular Medicine, Director of the Glycobiology Research and Training Center, and Coordinator of the project for Explaining the Origin of Humans, at the University of California, San Diego. Dr. Varki's laboratory explores the biology and evolution of sialic acids in health and disease.