

## Glyco-Forum section

### Natural ligands for CD33-related Siglecs?

#### Ajit Varki

Departments of Medicine and Cellular and Molecular Medicine, and the Glycobiology Research and Training Center, University of California, San Diego, La Jolla, CA 92093-0687, USA

Finding natural ligands involved in protein–protein interactions is relatively easy, and false-positives are rare. In contrast, the natural ligands for glycan-binding proteins (GBPs) can be quite elusive, and artifacts are common. For example, it took almost 10 years from the detection of P-selectin's glycan binding properties (Moore et al. 1991) to reach the definitive conclusion that a certain specific posttranslationally modified form of the polypeptide PSGL-1 was the functionally relevant natural ligand for this GBP (Epperson et al. 2000). There are many reasons for such difficulties, including the fact that glycan binding is often of relatively low affinity and relies on multivalency to achieve adequate avidity, as well as the complexities of local re-binding effects such as the recently described “jump and bind” mechanisms for lectin–mucin interactions (Dam and Brewer 2008). Meanwhile, investigators continue to incorrectly state that “PSGL-1 is the ligand for P-selectin,” without realizing that while the PSGL-1 polypeptide backbone is expressed in many cell types (Laszik et al. 1996), it is only the correctly sialylated, fucosylated, and tyrosine-sulfated form found in certain cells that can function as a P-selectin ligand. In this regard, an excellent prior suggestion is to define the ligand as the glycan(s) bound by the GBP, the protein or lipid to which the glycan(s) are bound as the “carrier,” and the composite of the glycan and the carrier as the “counter–receptor” (Crocker and Feizi 1996) (see Notes).

The Siglecs (Sialic acid-binding Ig super-family lectins) are a recently discovered family of GBPs (Crocker and Varki 2001; Varki and Angata 2006; Crocker et al. 2007). They are divided into two groups: the evolutionary conserved ones (Siglec-1, -2, -4, and -15) and the rapidly evolving CD33-related Siglecs (CD33rSiglecs, Siglec-3, 5–14, and -16 in primates, and Siglec-3 plus E-H in rodents). Within the first group there is evidence for specific sialylated ligands (for examples, see Sgroi et al. 1993; Schnaar et al. 1998; Nitschke et al. 1999; Van den Berg et al. 2001; Kimura et al. 2007). In contrast, it has been unclear what the natural ligands (or functions) of CD33-related Siglecs are. Based on the rapid evolution of their sialic acid-recognizing amino-terminal V-set domains, the inhibitory “ITIM” signaling motifs in their cytosolic tails, and the rapid evolving “sialome” of most vertebrate species, it was postulated that CD33rSiglecs function to recognize “self” in the

form of the host sialic acids (Crocker and Varki 2001; Varki and Angata 2006), thereby dampening unwanted responses by innate immune cells, wherein CD33rSiglecs are prominently expressed. A corollary hypothesis was that pathogens expressing sialic acids are “taking advantage” of the inhibitory properties of CD33rSiglecs in these host cells (Crocker and Varki 2001; Varki and Angata 2006). However, these hypotheses were without experimental proof, and most “functional” studies involved artificial interactions of recombinant soluble CD33rSiglecs with a variety of synthetic or semi-synthetic sialylated ligands. It is only recently that supporting evidence for these hypotheses emerged via elucidation of specific natural ligands for these GBPs – a story complicated by the discovery of additional activatory CD33rSiglecs (-13, -14, -16, and -H) that do not have ITIMs, but can instead recruit the ITAM-containing adapter protein DAP-12 (Angata et al. 2006; Blasius et al. 2006; Cao et al. 2008, and our unpublished observations).

The first evidence for natural ligands of CD33rSiglecs came from a study showing that alpha<sub>1</sub>-acid glycoprotein (AGP) could elicit an intracellular calcium response in neutrophils (Gunnarsson et al. 2007). Evidence that this signaling effect was mediated by Siglec-5 included the use of anti-Siglec-5 antibodies, affinity chromatography of Siglec-5 on sialylated AGP, and abrogation of interactions by treating the AGP with sialidase or with mild periodate (which selectively truncates the side chain of sialic acids). As AGP is a heavily sialylated N-linked glycoprotein found at high concentrations in blood plasma and also shows sialylation changes in response to inflammation (Brinkman-Van der Linden et al. 1996), it was reasonable to suggest that this interaction is a natural one, in which sialic acids on a “self” glycoconjugate send a signal to innate immune cells (Gunnarsson et al. 2007). Of course, the response in question was activatory, rather than inhibitory. Thus, it possibly also involved Siglec-14, a subsequently discovered “paired receptor” of Siglec-5, which cross-reacts with most known antibodies against Siglec-5, due to concerted evolution involving exons encoding the first two Ig-like domains (Angata et al. 2006; Yamanaka et al. 2009). Alternatively, this could be an example wherein an “inhibitory” receptor can function in activation (Barrow and Trowsdale 2006). Meanwhile, indirect evidence indicated that local upregulation of Siglec-F ligands during an allergic response could potentially engage Siglec-F on mouse eosinophils and mediate negative feedback inhibition (Zhang et al. 2007), a finding more in tune with the “self-recognition” hypothesis.

A recent study in *Science* reported another highly suggestive example, showing that the interaction of Siglec-10/G with CD24 (“heat stable antigen”) protects the host against a lethal response to pathological cell death by discriminating a “danger-associated molecular pattern” (DAMP) and repressing damage-induced immune responses during a noninfectious

To whom correspondence should be addressed: Tel: +1-858-534-2214; Fax: +1-858-534-5611; e-mail: avarki@ucsd.edu

inflammatory reaction (Chen et al. 2009). In this fascinating and important work, the authors discovered that interactions between Siglec-10 and CD24 mediated recognition of “self,” and thus limited damage of host cells by innate immune cells during a noninfectious inflammatory process, in which endogenous DAMPs such as HMGB1 were involved in a complex with CD24 and Siglec-10. While this paper clearly shows the functional significance of the Siglec-10:CD24 interaction and mentions that CD24 is >90% by weight carbohydrate, neither it nor the accompanying analysis (Bianchi and Manfredi 2009) actually discusses sialic acid-dependent interactions. However, the authors did find that an “inactivated” Siglec-10 was non-functional, involving a mutation that eliminates an arginine residue known to be critical for sialic acid binding (Varki and Angata 2006; Crocker et al. 2007). Moreover, CD24 is a GPI-anchored polypeptide with only ~30 amino acids, which carry a large number of sialylated *N*- and *O*-linked glycans (Kay et al. 1991). Given the results, it is very likely that the interaction studied by these authors (Chen et al. 2009) is between the sialic acid-binding site of Siglec-10 and the heavily sialylated glycans of CD24. Definitive proof requires further work of the kind done in the earlier study with AGP and Siglec-5/14. Meanwhile, it is reasonable to speculate that certain sialylated glycoforms of heavily glycosylated proteins such as AGP and CD24 might be ligands/counter-receptors for other Siglecs.

Since the authors of the *Science* paper (Chen et al. 2009) used a recombinant soluble form of CD24 expressed in CHO cells (Bai et al. 2000), it also remains to be seen what are the actual sialylation patterns of the native ligand recognized by Siglec-10. Another potential complexity arises because the authors did not pre-treat their commercially purchased Siglec-Fc chimeras with sialidase to remove *cis*-inhibition by sialylated *N*-glycans on the chimeras themselves (Sgroi et al. 1996). If they had done so, some of the other CD33rSiglecs might have also bound CD24 in a sialic acid-dependent manner. Additionally, Cheng et al. assume that “protein–polysaccharide interactions largely depend on cations,” and therefore use EDTA to disrupt interactions. As divalent cations are not known to be involved in Siglec binding, the effect is more likely due to the high density of carboxylates on EDTA, as previously found for selectins (Koenig et al. 1998), and/or only involves the CD24/HMGB1 interaction. Overall, while the jury is out on such specifics, it is reasonable to say that certain sialylated forms of CD24 are likely to carry functionally relevant ligands for one or more CD33rSiglecs *in vivo*. However, given the tissue-specific variations in glycosylation of CD24 (Ohl et al. 2003), it is very likely incorrect to simply say that “CD24 is a ligand for Siglec-10” (Bianchi and Manfredi 2009; Chen et al. 2009). Indeed, some glycoforms of CD24 are even preferred ligands/counter-receptors for P-selectin and others for the neural cell adhesion molecule L1 (see Notes), depending on their state of glycosylation (Sammar et al. 1997).

Taken together, all these data support the hypothesis that a primary function of the inhibitory ITIM-bearing CD33rSiglecs is to recognize self in the form of the host sialome, and thus dampen unwanted innate immune reactivity during homeostasis and/or in settings of inflammation. It seems likely that there will be heavily glycosylated molecules other than CD24 and AGP that are involved in this type of self-recognition. Given the importance of avidity and “jump and bind” effects (Dam and Brewer 2008), other mucin-type and/or heavily *N*-glycosylated

glycoproteins are candidate ligands. Meanwhile, results have recently emerged to support the secondary hypothesis that bacteria expressing sialic acids on their capsular polysaccharides use them to engage host CD33rSiglecs, by mimicking natural ligands. For example, the dense array of sialic acids presented on group B *Streptococcus* engages Siglec-9 on human neutrophils and dampens their reactivity to this human commensal/pathogen (Carlin et al. 2009). In this case, the “natural” ligand is generated by the pathogen and appears to take advantage of a native self-recognition system to blunt the immune response.

As mentioned earlier, the simple logic of the sialic acid “self-recognition” hypothesis has been complicated by the discovery of additional CD33rSiglecs that should deliver an activatory, rather than an inhibitory signal. Perhaps the emergence of these activatory CD33rSiglecs was an evolutionary response to sialylated bacteria “taking advantage” of the inhibitory CD33rSiglecs? Consistent with this idea is the fact that at least two of them (–14 and –16) are paired in binding properties with typical inhibitory counterparts (Siglec-5 and –11, respectively) (Angata et al. 2006, Cao et al. 2008). In this regard, cross-reactions of monoclonal antibodies within the 5/14 and 11/16 pairs have likely been confounding many studies to date.

Regardless of all these speculations, the road forward must include identification of the true natural ligands of CD33rSiglecs, which are likely to be certain specifically sialylated “counter–receptor” forms of very heavily glycosylated molecules such as AGP and CD24, whose function would be modulated not only by expression of the polypeptide backbone, but also by the glycosyltransferase repertoire of the cells expressing them. And these ligands must function in a milieu in which other lectins can recognize them, as well as many other potentially competing low-density and/or lower affinity cognate glycans. Furthermore, it is likely that false ligands will be detected during *in vitro* binding studies, interactions that actually never occur *in vivo*. Scientists who study protein–protein interactions have it easier!

## Notes

Another option is to use a superscript to indicate that a particular glycoform of a glycoprotein is a ligand for a particular GBP, e.g., the glycoform of PSGL-1 specifically recognized by P-selectin would be PSGL-1<sup>PSL</sup>. In this approach, the glycoform of CD24 that is recognized by Siglec-10 would be CD24<sup>S10L</sup>, differentiating it from the glycoform of CD24 that is a ligand for L1 in the brain, which would be designated CD24<sup>L1L</sup>.

## Funding

National Institute of Health (P01-HL057345).

## Acknowledgements

The author thanks Shiv Pillai of Harvard Medical School and Paul Crocker of the University of Dundee for insightful comments.

## Abbreviations

DAMP, danger-associated molecular pattern; GBPs, glycan-binding proteins.

## References

- Angata T, Hayakawa T, Yamanaka M, Varki A, Nakamura M. 2006. Discovery of Siglec-14, a novel sialic acid receptor undergoing concerted evolution with Siglec-5 in primates. *FASEB J*. 20:1964–1973.
- Bai XF, Liu JQ, Liu X, Guo Y, Cox K, Wen J, Zheng P, Liu Y. 2000. The heat-stable antigen determines pathogenicity of self-reactive T cells in experimental autoimmune encephalomyelitis. *J Clin Invest*. 105:1227–1232.
- Barrow AD, Trowsdale J. 2006. You say ITAM and I say ITIM, let's call the whole thing off: The ambiguity of immunoreceptor signalling. *Eur J Immunol*. 36:1646–1653.
- Bianchi ME, Manfredi AA. 2009. Immunology. Dangers in and out. *Science*. 323:1683–1684.
- Blasius AL, Cella M, Maldonado J, Takai T, Colonna M. 2006. Siglec-H is an IPC-specific receptor that modulates type I IFN secretion through DAP12. *Blood*. 107:2474–2476.
- Brinkman-Van der Linden ECM, Van OECR, Van DW. 1996. Glycosylation of alpha1-acid glycoprotein in septic shock: Changes in degree of branching and in expression of sialyl Lewisx groups. *Glycoconjugate J*. 13:27–31.
- Cao H, Lakner U, de Bono B, Traherne JA, Trowsdale J, Barrow AD. 2008. SIGLEC16 encodes a DAP12-associated receptor expressed in macrophages that evolved from its inhibitory counterpart SIGLEC11 and has functional and non-functional alleles in humans. *Eur J Immunol* 38:2303–2315.
- Carlin AF, Uchiyama S, Chang YC, Lewis AL, Nizet V, Varki A. 2009. Molecular mimicry of host sialylated glycans allows a bacterial pathogen to engage neutrophil Siglec-9 and dampen the innate immune response. *Blood*. 113:3333–3336.
- Chen GY, Tang J, Zheng P, Liu Y. 2009. CD24 and Siglec-10 selectively repress tissue damage-induced immune responses. *Science*. 323:1722–1725.
- Crocker PR, Feizi T. 1996. Carbohydrate recognition systems: Functional triads in cell–cell interactions. *Curr Opin Struct Biol*. 6:679–691.
- Crocker PR, Paulson JC, Varki A. 2007. Siglecs and their roles in the immune system. *Nat Rev Immunol*. 7:255–266.
- Crocker PR, Varki A. 2001. Siglecs, sialic acids and innate immunity. *Trends Immunol* 22:337–342.
- Dam TK, Brewer CF. 2008. Effects of clustered epitopes in multivalent ligand–receptor interactions. *Biochemistry*. 47:8470–8476.
- Epperson TR, Patel KD, McEver RP, Cummings RD. 2000. Noncovalent association of P-selectin glycoprotein ligand-1 and minimal determinants for binding to P-selectin. *J Biol Chem*. 275:7839–7853.
- Gunnarsson P, Levander L, Pahlsson P, Grenegard M. 2007. The acute-phase protein alpha 1-acid glycoprotein (AGP) induces rises in cytosolic Ca<sup>2+</sup> in neutrophil granulocytes via sialic acid binding immunoglobulin-like lectins (siglecs). *FASEB J*. 21:4059–4069.
- Kay R, Rosten PM, Humphries RK. 1991. CD24, a signal transducer modulating B cell activation responses, is a very short peptide with a glycosyl phosphatidylinositol membrane anchor. *J Immunol*. 147:1412–1416.
- Kimura N, Ohmori K, Miyazaki K, Izawa M, Matsuzaki Y, Yasuda Y, Takematsu H, Kozutsumi Y, Moriyama A, Kannagi R. 2007. Human B-lymphocytes express alpha2-6-sialylated 6-sulfo-N-acetylactosamine serving as a preferred ligand for CD22/Siglec-2. *J Biol Chem*. 282:32200–32207.
- Koenig A, Norgard-Sumnicht KE, Linhardt R, Varki A. 1998. Differential interactions of heparin and heparan sulfate glycosaminoglycans with the selectins – Implications for the use of unfractionated and low molecular weight heparins as therapeutic agents. *J Clin Invest*. 101:877–889.
- Laszik Z, Jansen PJ, Cummings RD, Tedder TF, McEver RP, Moore KL. 1996. P-Selectin glycoprotein ligand-1 is broadly expressed in cells of myeloid, lymphoid, and dendritic lineage and in some nonhematopoietic cells. *Blood*. 88:3010–3021.
- Moore KL, Varki A, McEver RP. 1991. GMP-140 binds to a glycoprotein receptor on human neutrophils: Evidence for a lectin-like interaction. *J Cell Biol*. 112:491–499.
- Nitschke L, Floyd H, Ferguson DJ, Crocker PR. 1999. Identification of CD22 ligands on bone marrow sinusoidal endothelium implicated in CD22-dependent homing of recirculating B cells. *J Exp Med*. 189:1513–1518.
- Ohl C, Albach C, Altevogt P, Schmitz B. 2003. N-Glycosylation patterns of HSA/CD24 from different cell lines and brain homogenates: A comparison. *Biochimie*. 85:565–573.
- Sammar M, Aigner S, Altevogt P. 1997. Heat-stable antigen (mouse CD24) in the brain: Dual but distinct interaction with P-selectin and L1. *Biochim Biophys Acta Protein Struct Mol Enzymol*. 1337:287–294.
- Schnaar RL, Collins BE, Wright LP, Kiso M, Tropak MB, Roder JC, Crocker PR. 1998. Myelin-associated glycoprotein binding to gangliosides – Structural specificity and functional implications. *Ann NY Acad Sci*. 845:92–105.
- SgROI D, Nocks A, Stamenkovic I. 1996. A single N-linked glycosylation site is implicated in the regulation of ligand recognition by the I-type lectins CD22 and CD33. *J Biol Chem*. 271:18803–18809.
- SgROI D, Varki A, Braesch-Andersen S, Stamenkovic I. 1993. CD22, a B cell-specific immunoglobulin superfamily member, is a sialic acid-binding lectin. *J Biol Chem*. 268:7011–7018.
- Van Den Berg TK, Nath D, Ziltener HJ, Vestweber D, Fukuda M, Van Die I, Crocker PR. 2001. Cutting edge: CD43 functions as a T cell counterreceptor for the macrophage adhesion receptor sialoadhesin (Siglec-1). *J Immunol*. 166:3637–3640.
- Varki A, Angata T. 2006. Siglecs – the major subfamily of I-type lectins. *Glycobiology*. 16:1R–27R.
- Yamanaka M, Kato Y, Angata T, Narimatsu H. 2009. Deletion polymorphism of SIGLEC14 and its functional implications. *Glycobiology* doi:10.1093/glycob/cwp052.
- Zhang M, Angata T, Cho JY, Miller M, Broide DH, Varki A. 2007. Defining the in vivo function of Siglec-F, a CD33-related Siglec expressed on mouse eosinophils. *Blood*. 109:4280–4287.