Cracking the carbohydrate code for selectin recognition

Selectins are central in the inflammatory response; the discovery that they bind to carbohydrate ligands has galvanized carbohydrate chemists to search for inhibitors of the process. Recent progress in identifying and analyzing physiological selectin counter-receptors suggests new approaches to the design of ligands that bind to specific selectins.

Selectin family of adhesion molecules has three members, E-, P- and L-selectin (for a recent review see [1]). These share a common domain structure, which includes an amino-terminal calcium-dependent (C-type) lectin domain, the sequence of which is highly conserved between family members, an epidermal growth factor (EGF)-like domain, and a series of short consensus repeat sequences that are homologous to complement regulatory proteins. Expression of E- and P-selectin on the surface of endothelial cells is induced in response to inflammatory cytokines and microbial-derived toxins. E- and P-selectin interact with carbohydrate-based ligands on leukocytes. In contrast, L-selectin is constitutively expressed on all classes of leukocytes in the blood and its ligands, also carbohydrate-based, are found on endothelial cells. All three receptors have been shown in animal models to be important in acute and chronic inflammatory leukocyte recruitment; in different disease states, the selectin or selectins that are most important for recruitment vary.

The discovery of a C-type lectin domain in the three selectins initiated the race to uncover complementary carbohydrate ligands. What emerged from these efforts was the discovery of a common motif recognized by all three selectins, the sialylated and fucosylated tetrasaccharide.
sialyl Lewis x (sLe^x), NeuAcα2,3Galβ1,4(Fucoα1,3)GlcNAc, compound 1, Fig. 2) and a related structure, sialyl Lewis a (sLe^a, NeuAcα2,3Galβ1,3(Fucoα1,4)GlcNAc) [3,4]. These two oligosaccharides differ in their linkages of galactose and fucose to N-acetylglucosamine, but share a similar spatial orientation of galactose, fucose and sialic acid. The functional groups that interact directly with the selectin binding pocket are thought to be on the sialic acid, fucose and galactose moieties, whereas N-acetylglucosamine serves primarily as a scaffold. Selectin binding to sLe^x under equilibrium conditions is fairly weak; the K_d for sLe^x binding to E-selectin has been reported to be in the millimolar range [5,6]. Nonetheless, sLe^x is an effective anti-inflammatory agent in several acute disease models [7-10] and is currently in clinical trials. The components of sLe^x that contribute to selectin binding have been probed extensively using synthetic sLe^x analogs. It is remarkable that derivatives in which the sialic-acid residue was replaced with a sulfate ester (3'-sulfo Le^x) were equally active in binding all three selectins, implying that any anionic substituent at the 3-position of the galactose residue might impart selectin-binding activity [11-14]. This discovery was embraced by carbohydrate chemists interested in inhibitor design, as it offered a welcome reprieve from the difficulties of sialic acid chemistry. The hydroxyl groups of fucose were found to be critical for binding to all three selectins [14,15], whereas a variety of substitutions appear to be permitted at the 2-position of N-acetylglucosamine [16]. These observations indicate that the recognition domain on sLe^x comprises the carbohydrate group of sialic acid and the hydroxyl groups of fucose and galactose (Fig. 2), which are all found on one face of the molecule in the solution structure of sLe^x as determined by NMR [17,18].

Central to any drug design effort is an accurate structural model of the receptor–ligand interaction. Unfortunately, a high-resolution crystal structure of a selectin bound to sLe^x is not yet available. The three-dimensional structure of E-selectin without bound ligand has been reported [19], however, and several models for selectin–sLe^x binding based on the crystal structure of E-selectin and the bound conformation of sLe^x determined by NMR. In this model, sLe^x binds in a shallow pocket delineated by Lys99 (orange) and the bound calcium ion (cyan). A critical residue, Arg97, interacts with the carboxylate of sialic acid, and the 2- and 3-hydroxyl groups of fucose are coordinated to calcium.

Fig. 2. Sialyl Lewis x (sLe^x), a recognition motif common to the selectins. The highlighted functional groups are essential for selectin binding and form a recognition domain along one face of the tetrasaccharide. NeuAc = N-acetylneuraminic acid (sialic acid), Gal = galactose, Fuc = fucose, GlcNAc = N-acetylglucosamine. Key positions referred to in the text are numbered.

Fig. 3. A proposed model for sLe^x bound in the lectin domain of E-selectin. This model is based on the crystal structure of E-selectin and the bound conformation of sLe^x determined by NMR. In this model, sLe^x binds in a shallow pocket delineated by Lys99 (orange) and the bound calcium ion (cyan). A critical residue, Arg97, interacts with the carboxylate of sialic acid, and the 2- and 3-hydroxyl groups of fucose are coordinated to calcium.

Kiesling and coworkers [27] examined the effects of galactose modifications on selectin binding to the Lewis a...
Selectin recognition of glycoproteins

Table 1. Glycomimetics of sLex and their relative binding potencies with E-selectin.

<table>
<thead>
<tr>
<th>Compound</th>
<th>Potency</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0.7</td>
<td>[28]</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
<td>[28]</td>
</tr>
<tr>
<td>4</td>
<td>1.5</td>
<td>[29]</td>
</tr>
<tr>
<td>5</td>
<td>0.05</td>
<td>[30]</td>
</tr>
<tr>
<td>6</td>
<td>0.04</td>
<td>[31]</td>
</tr>
<tr>
<td>7</td>
<td>[32]</td>
<td></td>
</tr>
</tbody>
</table>

Compounds with values greater than 1 bind E-selectin with greater affinity than sLex, whereas compounds with values less than 1 bind E-selectin with weaker affinity.

trisaccharide. It is interesting that sulfation of the 6 OH of galactose in 3'-sulfo Le' caused a dramatic reduction in E selectin binding activity. In the second model for E-selectin binding, the 6-OH of galactose is involved in a critical hydrogen bond with Tyr94 and is positioned near Glu80, which might explain the weakened binding of the sulfated derivative. Future binding studies with mutant selectins and complementary synthetic sLe' analogs may allow further refinement of ligand-binding models. Still, the ultimate proof of any proposed model awaits the structure of a selectin-ligand co-crystal.

**Glycomimetics as selectin antagonists**

Sialylated and fucosylated oligosaccharides such as sLe' are difficult and expensive to synthesize. Thus, armed with a collection of data from structure-activity relationships and binding models, several synthetic chemists have designed sLe' mimetics endowed with the minimal requirements for selectin binding. Some examples are shown in Table 1 with their E-selectin binding potencies compared to that of sLe' (compound 1). Compound 2 contains fucose and galactose residues joined by a short linker; a simple acidic substituent replaces sialic acid [28]. Compound 3 is further simplified with replacement of galactose by an amino-acid diol [28]. An even smaller derivative, compound 4, is composed of a monosaccharide mimic of fucose (mannose) tethered to a carboxylic acid by a rigid hydrophobic spacer [29]. It is remarkable that mimetics 2–4 are comparable to sLe' in their binding to E-selectin. Compounds 5 and 6, in which fucose and sialic-acid residues are tethered by a rigid or flexible linker, respectively, show diminished E-selectin binding activity compared to sLe' [30,31]. The most effective glycomimetic E-selectin inhibitor described to date is compound 7, which was designed based on the results of a computer-aided search using sLe' as a basis structure [32]. But as yet, no truly potent (by which I mean, having an IC<sub>50</sub> in the nanomolar or low micromolar range) selectin inhibitor has been derived from sLe' pharmacophores.

**Physiological selectin counter-receptors provide new leads for drug design**

The pursuit of selectin antagonists has centered on the common sLe' binding activity of the three selectins in vitro. It is now apparent, however, that selectin binding in vivo is a more complicated matter, involving distinct binding specificities in different tissue sites. The picture emerging from biological studies is that each selectin recognizes a discrete set of glycoprotein molecules on opposing cells, and that these physiological counter-receptors present the proper carbohydrate-based epitopes for optimal selectin binding in vivo. Several selectin counter-receptors have now been cloned and biochemically characterized, exposing a number of different mechanisms by which nature has achieved selectin specificity (Fig. 4).

Three physiological counter-receptors for L-selectin have been identified, GlyCAM-1 [33], CD34 [34] and
MAdCAM-1 [35,36]. These glycoproteins share two structural features: dense clusters of serine and threonine residues that are modified with O-linked oligosaccharide chains, and oligosaccharide structures that bind to L-selectin. Although the GlyCAM-1, CD34 and MAdCAM-1 polypeptides are expressed in different tissues, the proper glycoforms for L-selectin binding are restricted to endothelial cells at sites of leukocyte recruitment.

The oligosaccharides on GlyCAM-1 have been analyzed in detail, culminating in the discovery of two novel structures, 6'-sulfo sLe\(^{\alpha}\) (NeuAcα2,3(SO\(_3\)-6)Galβ1,4-(Fucα1,3)GlcNAc) and 6-sulfo sLe\(^{\alpha}\) (NeuAcα2,3-Galβ1,4(Fucα1,3)(SO\(_3\)-6)GlcNAc) [37-39], which cap the majority of the oligosaccharide chains. Based on the observation that sulfation of GlyCAM-1 is required for L-selectin binding [40], it was proposed that sulfation of the 6-hydroxyl group of galactose or N-acetylglucosamine increases the affinity of sLe\(^{\alpha}\) for L-selectin. Consequently, a new strategy for the design of L-selectin antagonists might involve the installation of sulfate esters on the appropriate positions of a core structure related to sLe\(^{\alpha}\) [41]. Indeed, the disaccharide lactose 6',6-disulfate, which is sulfated at positions analogous to those found on the GlyCAM-1 oligosaccharides, is more potent than sLe\(^{\alpha}\) as an L-selectin inhibitor [42]. Other structural features of GlyCAM-1, such as the organization of oligosaccharides along the peptide backbone, may also be relevant to inhibitor design.

The only physiological counter-receptor for P-selectin that has been identified at the molecular level is the leukocyte-associated glycoprotein PSGL-1 [43,44]. Like the L-selectin counter-receptors, PSGL-1 possesses regions of densely clustered oligosaccharides bound to serine and threonine residues, and the oligosaccharide chains are capped with sLe\(^{\alpha}\)-like structures [45]. PSGL-1 has an additional feature, however, that is not shared by the L-selectin counter-receptors — an amino-terminal domain with a consensus sequence for tyrosine sulfation. Recent data suggest that sulfated tyrosine residues within this domain and sLe\(^{\alpha}\)-like epitopes on a separate region of the glycoprotein are both essential for functional P-selectin binding activity [46,47]. This startling discovery suggests a model for P-selectin binding that includes protein recognition in addition to carbohydrate recognition, dramatically diverging from traditionally held views. The discovery of a tyrosine-sulfated peptide as a structural motif for P-selectin recognition offers new approaches for the design of experimentally and therapeutically useful molecules. Novel inhibitors can be designed to include both tyrosine sulfates and sLe\(^{\alpha}\) moieties. In addition, readily available aryl sulfatases, which cleave tyrosine sulfate esters, may have therapeutic value in preventing P-selectin adhesion to PSGL-1.

The bifunctional nature of P-selectin/PSGL-1 recognition invites speculation about the mechanics of the interaction. Sako et al. [47] proposed a model in which the lectin domain of P-selectin binds specifically to sLe\(^{\alpha}\)-like epitopes, while a secondary binding region, perhaps within the EGF domain, engages in non-specific electrostatic interactions with the sulfated region of PSGL-1. From this perspective, P-selectin binding to PSGL-1 may be mechanistically similar to DNA recognition by a subset of DNA-binding proteins known to engage in both non-specific electrostatic interactions with the phosphodiester backbone and in specific hydrogen bonds within the major groove [48]. The sequence-non-specific electrostatic interactions greatly increase the association rate. Likewise, non-specific electrostatic interactions...
between P-selectin and PSGL-1 may increase the association rate and enable tethering of leukocytes to endothelial cells under conditions of shear stress.

A physiological counter receptor for E-selectin, ESL-1, has also been identified [49]. This glycoprotein is unique among the selectin counter-receptors in that the oligosaccharide substituents appear to be primarily asparagine-linked (N-linked). Furthermore, ESL-1 has only five potential sites of N-linked glycosylation, in stark contrast to GlyCAM-1, CD34 and PSGL-1, which have mucin-like structures with many potential serine/threonine-linked (O-linked) glycosylation sites. The multivalent interaction of a mucin-like glycoprotein with several selectin molecules may contribute strongly to high-avidity L- and P-selectin binding and may be a requirement for adhesion under conditions where the cell is under shear stress, as it is in a blood vessel due to blood flow. Clearly ESL-1 is not endowed with the same degree of multivalency, and high-avidity E-selectin binding must involve different mechanisms.

The report of unusual tetra-antennary N-linked oligosaccharides with high affinity for E-selectin [50] is worthy of note. These oligosaccharides bear a difucosylated SLε- like epitope (d-6Le') on one branch, and various sialylated structures on the other three branches. The $K_a$ values of the monomeric oligosaccharides with E-selectin were estimated to be $<1 \mu M$, in contrast to $1 \mu M$ for SLε. The structural origin of the unusually high binding affinity has not yet been determined, but may be a consequence of favorable contacts with carbohydrate residues beyond the SLε capping group. Alternatively, the environment of the SLε tetrasaccharide within the branched structure may impose conformational restrictions that lower the entropic cost of E-selectin complexation [51]. If the ESL-1 oligosaccharides bind E-selectin with affinities similar to those of the tetra-antennary oligosaccharides, a high degree of multivalency may not be required to achieve cell adhesion under the conditions of blood flow. These oligosaccharide structures, which bind to E-selectin with high avidity, may provide new insights into the design of E-selectin antagonists.

**Future perspectives**

In summary, SLε is a good starting point for the design of selectin antagonists, but further progress will require a better understanding of the structural requirements for selectin binding and specificity. Although much has been learned in the six years since the discovery of the selectin family, the molecules responsible for selectin-mediated adhesion in vivo are only now being characterized at the molecular level. The physiological selectin counter-receptors bear sialylated and fucosylated oligosaccharides related to SLε, a shared feature required for selectin binding. But it is the differences among the selectin counter-receptors that are most intriguing and that offer chemists new avenues for therapeutic intervention: that is, sialylated, O-linked oligosaccharides on GlyCAM-1 (L-selectin), tyrosine sulfate on PSGL-1 (P-selectin), and sparse N-linked oligosaccharides on ESL-1 (E-selectin). The organization of key epitopes on the peptide backbone may also be important for selectin binding in vivo. How these structural elements contribute to selectin binding at the molecular level is an open field of investigation. A goal for chemists is to combine these unique structural features in the design of potent antagonists with specificity for a single selectin.

**Acknowledgements** I thank Tim Kogan of Texas Biotechnology Corp. for kindly providing a figure of the proposed model for E-selectin binding to SLε. I also thank Tim Kogan, Gray Shaw and Brian Seed for sharing unpublished results, and Steve Rosen and Pati Giblin for helpful comments on the manuscript. Support from the Camille and Henry Dreyfus Foundation is gratefully acknowledged.

**References**


a-t-mannopyranosyloxybiphenyl substituted carboxylic acids. /.
the vascular sialomucin CD34. Science 262, 436-438.
The synthesis of novel mimics of the sialyl Lewis x determinant.
derived from a pharmacophore search are selectin inhibitors with
synthesis of small molecule, non-oligosaccharide selectin inhibitors:
single amino acid residue can determine the ligand specificity of
selectin is a novel mucin-like molecule. Cell69, 927-938.
(19931. L-selectin-mediated lymphocyte rolling on MadCAM-I.
the amino terminus of PSGL-1 is critical for P-selectin binding. Cell
75, 1179-l 186.
ESL-1 is a variant of a receptor for fibroblast growth factor. Nature
373, 615-620.
and characterization of natural protein-associated carbohydrate
act as signals? Pure Appl. Chem. 65, 763-770.

Carolyn R Bertozzi, Department of Chemistry, University of California, Berkeley, CA 94720, USA.