Tables of Representative Carbohydrate-Binding Proteins

To date numerous CBPs have been identified and demonstrated or implicated in mediating various cellular events through protein-carbohydrate interactions. The Consortium will target representatives of three major families of CBPs 1) the C-type lectin family (including the selectins, Table 1), the Siglec family (including CD22, Table 2) and the galectin family (Table 3). In addition, within the last five years T cell antigen receptors (TCR) have been documented to recognize carbohydrate antigens when presented by non-classical (CD1) and classical major histocompatibility complex (MHC) antigen presenting molecules. For the purposes of this project, the sub-family of TCR and the corresponding presenting molecules represent a fourth family (Table 4) that will be investigated. In addition to these four families, several other CBPs have been identified which are not homologous members of these families, but that still fall within the scope of the program. Please see the Consortium CBP DB for a comprehensive listing of targets within the scope of the Consortium.

<table>
<thead>
<tr>
<th>CBP</th>
<th>Group</th>
<th>Species</th>
<th>Cell types</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>CD23a</td>
<td>2</td>
<td>Human/Murine</td>
<td>B cells</td>
<td>• Binds CD11/CD18 via carbohydrate&lt;br&gt;• Binds IgE via protein determinant</td>
</tr>
<tr>
<td>DC-SIGNb</td>
<td>2</td>
<td>Human</td>
<td>Dendritic cell subset</td>
<td>• Binds T cells to promote primary immune response&lt;br&gt;• Mannose binding motif&lt;br&gt;• Binds ICAM3 as ligand, also HIV gp120</td>
</tr>
<tr>
<td>DC-SIGNRc</td>
<td>2</td>
<td>Human</td>
<td>Dendritic cells</td>
<td>• May be a co-receptor with DC-SIGN</td>
</tr>
<tr>
<td>Dectin-1d</td>
<td>2</td>
<td>Murine</td>
<td>Dendritic cells in Spleen, thymus and Skin (langerhan)</td>
<td>• Dendritic cells in spleen, thymus and skin (Langerhans)</td>
</tr>
<tr>
<td>Dectin-2e</td>
<td>2</td>
<td>Murine</td>
<td>Same as Dectin-1</td>
<td>• Similar expression pattern to dectin-1</td>
</tr>
<tr>
<td>HECLf</td>
<td>2</td>
<td>Human</td>
<td>Macrophage Erythroid precursors</td>
<td>• Putative cytoplasmic localization</td>
</tr>
<tr>
<td>Langering</td>
<td>2</td>
<td>Human</td>
<td>Langerhan cells (epidermal DC)</td>
<td>• Inducer of Birbeck granules</td>
</tr>
<tr>
<td>Layilinh</td>
<td>2</td>
<td>Murine</td>
<td>Numerous</td>
<td>• Located in membrane ruffles&lt;br&gt;• Binds talin</td>
</tr>
<tr>
<td>MCLi</td>
<td>2</td>
<td>Murine</td>
<td>Macrophage</td>
<td>• Macrophage restricted</td>
</tr>
<tr>
<td>Minclej</td>
<td>2</td>
<td>Murine</td>
<td>Macrophage</td>
<td>• Inducible C-type lectin (TNF, IL-6, INF-γ)</td>
</tr>
<tr>
<td>MMGLk</td>
<td>2</td>
<td>Murine</td>
<td>Macrophages</td>
<td>• Binds Gal/GalNac&lt;br&gt;• Removes asialo-glycoproteins from serum&lt;br&gt;• Implicated in DTH</td>
</tr>
<tr>
<td>E-selectinl</td>
<td>4</td>
<td>Murine/Human</td>
<td>Endothelial cells</td>
<td>• Binds SLeX containing carbohydrates&lt;br&gt;• Mediates leukocyte trafficking to sites of inflammation</td>
</tr>
<tr>
<td>P-selectinm</td>
<td>4</td>
<td>Murine/Human</td>
<td>Platelets Endothelial</td>
<td>• Binds SLeX containing O-linked sugars on PSGL-1&lt;br&gt;• Mediates leukocyte trafficking to sites of inflammation&lt;br&gt;• Mediates platelet binding to neutrophils</td>
</tr>
<tr>
<td>L-selectin</td>
<td>4</td>
<td>Murine/Human</td>
<td>Neutrophils Lymphocytes</td>
<td>• Binds sulfated-SLeX containing carbohydrates on O-linked oligosaccharides&lt;br&gt;• Mediates neutrophil trafficking to sites of inflammation&lt;br&gt;• Mediates lymphocyte recirculation</td>
</tr>
<tr>
<td>DEC-205o</td>
<td>6</td>
<td>Murine/Human</td>
<td>Dendritic Thymic endothelial</td>
<td>• Endocytic receptor involved in antigen processing&lt;br&gt;• Existing knock out mouse</td>
</tr>
<tr>
<td>Endo180p</td>
<td>6</td>
<td>Human/Murine</td>
<td>Macrophages Fibroblasts Endothelial Chondrocytes</td>
<td>• Endocytic receptor&lt;br&gt;• Binds GlcNAc</td>
</tr>
<tr>
<td>Mannose/(GalNAc-SO₄) receptorq</td>
<td>6</td>
<td>Murine</td>
<td>Macrophages Hepatic endothelial</td>
<td>• Role in innate immunity, possible uptake of antigen for presentation by CD1 and MHC class II&lt;br&gt;• Mannose specific as monomer in macrophages&lt;br&gt;• GalNAc-SO₄ specific as dimer in endothelial cells</td>
</tr>
<tr>
<td>Phospholipase A(2) receptorr</td>
<td>6</td>
<td>Human/Murine</td>
<td>Various</td>
<td>• Binds to phospholipase A2 through peptide determinant</td>
</tr>
</tbody>
</table>

References: a 1-3; b 4-7; c 5; d 8; e 9; f 10; g 11; h 12; i 13; j 14; k 15-19; l 20; m 21-23; n 24-26; o 27-29; p 30; q 31-36; r 37-40
<table>
<thead>
<tr>
<th>Siglec (# Ig Domains)</th>
<th>Species</th>
<th>Cell types</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Sialoadhesin/Siglec-1\(^a\) (17) | Human Murine | Macrophages | • Subsets of resident and inflammatory macrophages  
• Does not require unmasking to mediate binding  
• Does not have tyrosine-based motifs in cyt. Tail  
• Prefers 2,3 linked sialic acid over 2,6 linkage |
| CD22/Siglec-2\(^b\) (7) | Human Murine | B cells | • Regulator of B cell activation  
• Recruits SHP-1 to cytoplasmic ITIM motifs  
• Potential role in B cell homing to bone marrow  
• Binding site masked on resting B cells  
• Strong preference for 2,6-linked sialic acids |
| CD33/Siglec-3\(^c\) (2) | Human Murine | Myeloid progenitors Monocytes | • Commonly used as a marker of early myeloid cells  
• Co-cross-linking with Fc receptor inhibits activation  
• Two cytoplasmic ITIM-like motifs  
• Recruits SHP-1 and SHP-2 to ITIM motifs |
| MAG/Siglec-4\(^d\) (4) | Human Murine | Oligodendrocytes Schwann cells | • Role in maintenance of myelin, inhibitor of axonal growth  
• Two forms with and without ITIM motif  
• Single cytoplasmic tyrosine-based motif  
• Prefers 2,3 linked sialic acid over 2,6 linkage |
| Siglec-5\(^e\) (4) | Human | Monocytes Neutrophils | • Presence of two ITIM-like motifs  
• Binds 2,3- 2,6- and 2,8-linked sialic acids |
| Siglec-6\(^f\) (3) | Human | Trophoblast B cells | • Identified as low affinity leptin receptor  
• Presence of two ITIM-like motifs  
• Restricted specificity for Sialyl Tn structure |
| Siglec-7\(^g\) (3) | Human | Natural killer cells Monocytes | • Identified as NK cell inhibitory receptor  
• Alternatively spliced form lacking D2 identified  
• Presence of two ITIM-like motifs  
• Binds equally to 2,3 and 2,6-linked sialic acids |
| Siglec-8\(^h\) (3) | Human | Eosinophils | • Alt. Spliced forms have two ITIM-like motifs  
• Prefers 2,3 linked sialic acid over 2,6 linkage |
| Siglec-9\(^i\) (3) | Human | Monocytes Neutrophils | • Presence of two ITIM-like motifs  
• Binds equally to 2,3 and 2,6-linked sialic acids |
| Siglec-10\(^j\) (5) | Human | NK-like cells B cells | • Presence of two ITIM-like motifs  
• Binds equally to 2,3 and 2,6-linked sialic acids |
| Siglec-11 | Human | Under study | • Under investigation |

References a 41, 42; b 43–46; c 47, 48; d 47, 48, 9, 49, 50; e 51, 52; f 52; g 53–55; h 56; i 57, 58; j 59
<table>
<thead>
<tr>
<th>Galectin / Species</th>
<th>Structure</th>
<th>Organs/Cell types</th>
<th>Comments</th>
</tr>
</thead>
</table>
| Galectin-1<sup>a</sup>  
• Human  
• Mouse  
• Rat | ![Structure](image1) | Abundant in most organs (muscle, heart, lung, liver, lymph node, thymus, colon) | • Probably most widely distributed galectin  
• Binds terminal -galactosides  
• Occurs as a 14.9 kDa non-covalent dimer  
• Expressed in epithelial endothelial, fibroblasts, and smooth muscle cells |
| Galectin-2<sup>b</sup>  
• Human  
• Rat | ![Structure](image2) | Stomach epithelial cells | • Thought to be expressed at a minor level  
• Occurs as a 14.5 kDa non-covalent dimer |
| Galectin-3<sup>c</sup>  
• Human  
• Mouse  
• Rat | ![Structure](image3) | Colon, macrophages, many other epithelial and fibroblastic cells | • N-terminus has a collagen-like domain  
• Also known as the Mac-2 suface antigen  
• Occurs as a 27.5 kDa monomer |
| Galectin-4<sup>d</sup>  
• Human  
• Mouse  
• Rat | ![Structure](image4) | Gastrointestinal | • Occurs as a divalent monomeric protein with tandem CRDs that may have different sugar binding specificity |
| Galectin-5<sup>e</sup>  
• Rat | ![Structure](image5) | Erythrocytes | • The single CRD has 85% homology to the CRD of galectin-9 |
| Galectin-6<sup>f</sup>  
• Human  
• Mouse | ![Structure](image6) | Gastrointestinal | • 85% identical to galectin-6  
• Galectin-4 and galectin-6 genes are closely linked |
| Galectin-7<sup>g</sup>  
• Human  
• Mouse | ![Structure](image7) | Skin | • Used as a marker of stratified epithelium |
| Galectin-8<sup>h</sup>  
• Human  
• Mouse  
• Rat | ![Structure](image8) | Liver, kidney, cardiac muscle, lung, and brain | • Occurs as a divalent monomeric protein with tandem CRDs (joined by a link peptide) that may have different sugar binding specificity  
• Its link peptide is totally unique from Galectin4/6 |
| Galectin-9<sup>i</sup>  
• Human  
• Mouse  
• Rat | ![Structure](image9) | Thymus, kidney, Hodgkin’s lymphoma | • Galectin-9 and its (most probably) allelic variant ecalectin, are potent eosinophil chemoattractants  
• Occurs as a divalent monomeric protein with tandem CRDs (joined by a link peptide) that may have different sugar binding specificity |
| Galectin-10<sup>l</sup> (Charcot-Leyden Crystal Protein)  
• Human  
• Mouse | ![Structure](image10) | Eosinophils, basophils | • Most abundant protein in eosinophils, formerly called Charcot-Leyden Crystal Protein  
• Previously thought to have lysosphospholipase activity, but recently shown to be distinct  
• Unlike galectin-1, galectin-10 binds mannose |
| Galectin-11<sup>k</sup> (Grifin)  
• Rat | ![Structure](image11) | Lens | • May represent a new lens crystallin, (galectin-related inter-fiber protein)  
• Exhibits lack of affinity for simple β-galactosides  
• Like galectin-1, galectin-11 can dimerize |

References: a 60-62, b 63, c 64, d 65, e 66, f 67, g 68, h 69, i 70, j 71, k 72
### Table 4. T cells Reactive with Antigens Containing Carbohydrates

<table>
<thead>
<tr>
<th>Antigen</th>
<th>T cell</th>
<th>Antigen presenting molec.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Peptides</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glycopeptides</td>
<td>αβ T cells</td>
<td>Classical class I or Classical class II</td>
</tr>
<tr>
<td>Glycopeptides</td>
<td>γδ T cells</td>
<td>None—direct recognition</td>
</tr>
<tr>
<td>Synthetic glycosphingolipids</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Microbial glycolipids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Glucose monomycolate</td>
<td>αβ T cells</td>
<td>CD1a</td>
</tr>
<tr>
<td>Lipoarabinomannan</td>
<td>αβ T cells</td>
<td>CD1b</td>
</tr>
<tr>
<td>Phosphatidyl inositol</td>
<td>αβ T cells</td>
<td>CD1c</td>
</tr>
<tr>
<td>Hexosyl-1-phosphoisoprenoids</td>
<td>αβ T cells</td>
<td></td>
</tr>
<tr>
<td>Mannosyl-phosphodolichols</td>
<td>αβ T cells</td>
<td></td>
</tr>
<tr>
<td>Undefined microbial</td>
<td>γδ T cells</td>
<td>CD1c</td>
</tr>
<tr>
<td><strong>Autologous or synthetic glycolipids</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undefined autologous</td>
<td>αβ T cells</td>
<td>CD1a</td>
</tr>
<tr>
<td>Brain glycosphingolipids</td>
<td>αβ T cells</td>
<td>CD1b</td>
</tr>
<tr>
<td>including gangliosides and sulfatide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undefined autologous</td>
<td>αβ T cells</td>
<td>CD1c</td>
</tr>
<tr>
<td>α-galactosylceramide</td>
<td>NK T cells</td>
<td>CD1d</td>
</tr>
<tr>
<td>Undefined autologous</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### References:


