Discovery and Classification of Glycan-Recognizing Proteins

Lecture 17

April 27th., 2004

Ajit Varki

OVERALL COURSE OUTLINE

• General Principles
• Structure, Biosynthesis and Turnover
• Protein-Glycan Interactions
• Organismal Diversity in Glycans
• Methods and Applications
Protein-Glycan Interactions

- April 27  Discovery and Classification of Glycan-Recognizing Proteins Varki
  Principles of Glycan Recognition Esko
- April 29  R- and L-type Lectins (Other than Plant lectins)
  Glycoproteins in Quality Control and Secretion Varki
  The "P-type" lectins and the trafficking of lysosomal enzymes Varki
- May 4  The Galectins (formerly "S-type" lectins) Cummings
  The "C-type" lectins and the Selectins Cummings
- May 6  Plant lectins: discovery, characterization and utility Cummings
  Glycans in Parasitic Infections (**rescheduled from May 20) Cummings
- May 11  The "I-type" lectins and the Siglecs (**rescheduled from May 6) Varki
  Glycosaminoglycan-binding proteins Esko
- May 20  Microbial Adhesins, Agglutinins and Toxins (**rescheduled from May 11) Nizet

Glycan-Binding Proteins are Widespread in Nature

M = Micro-organism (Pathogen or Symbiont) or Toxin
P = Bacteriophage
Definition of a Lectin:

“A protein that specifically recognizes and binds to glycans without catalyzing a modification of the glycan.”

Exceptions:
- Anti-carbohydrate antibodies
- Sulfated GAG-binding Proteins
Comparison of the two Major Classes of Eukaryotic Glycan-Binding Proteins

<table>
<thead>
<tr>
<th></th>
<th>Lectins</th>
<th>Sulfated GAG-binding proteins</th>
</tr>
</thead>
<tbody>
<tr>
<td>Shared evolutionary origins</td>
<td>Yes, within each group</td>
<td>No</td>
</tr>
<tr>
<td>Shared structural features</td>
<td>Yes, within each group</td>
<td>No</td>
</tr>
<tr>
<td>Defining arrangement of AA</td>
<td>Typical for each group</td>
<td>Patch of basic amino acid residues</td>
</tr>
<tr>
<td>residues involved in binding</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of glycans recognized</td>
<td>N- &amp; O-glycans, GSLs</td>
<td>Different types of sulfated GAGs</td>
</tr>
<tr>
<td>Location of cognate residues</td>
<td>Typically sequences at the outer ends of glycans (can be internal)</td>
<td>Typically sequences internal to an extended GAG chain</td>
</tr>
<tr>
<td>within glycans</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Specificity for glycans</td>
<td>Stereospecificity often high for specific glycan structures</td>
<td>Often recognize a range of related GAG structures</td>
</tr>
<tr>
<td>Recognized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single-site binding affinity</td>
<td>Often low. High avidity generated by multivalency</td>
<td>Often moderate to high</td>
</tr>
<tr>
<td>Valency of binding sites</td>
<td>Multivalency very common</td>
<td>Often monovalent</td>
</tr>
</tbody>
</table>

Milestones in the Discovery of Plant Lectins

<table>
<thead>
<tr>
<th>Date</th>
<th>Investigators</th>
<th>Discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>1888</td>
<td>H. Stillmark</td>
<td><em>Ricinus communis</em> plant extract has red cell agglutinating properties</td>
</tr>
<tr>
<td>1908</td>
<td>K. Landsteiner &amp; H. Raubitscheck</td>
<td>Different hemagglutinating properties in various plant seeds</td>
</tr>
<tr>
<td>1919</td>
<td>J. Sumner</td>
<td>Crystallization of Concanavalin A</td>
</tr>
<tr>
<td>1936</td>
<td>J. Sumner</td>
<td>Lectins bind sugar - Con A precipitates glycogen</td>
</tr>
<tr>
<td>1940</td>
<td>W. Boyd, R. Reguera &amp; K.O. Renkonen</td>
<td>Lectins specific for some human blood group antigens</td>
</tr>
<tr>
<td>1952</td>
<td>W. Watkins &amp; W. Morgan</td>
<td>Lectins and glycosidases used to prove that blood group antigens are sugars and to deduce structure</td>
</tr>
</tbody>
</table>
### Milestones in the Discovery of Plant Lectins

<table>
<thead>
<tr>
<th>Date</th>
<th>Investigators</th>
<th>Discovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>1954</td>
<td>W. Boyd &amp; E. Shyleigh</td>
<td>The name lectin is proposed to replace “hemagglutinin”</td>
</tr>
<tr>
<td>1960</td>
<td>P.C. Nowell &amp; J.C. Aub</td>
<td>Red kidney bean lectin P. vulgaris mitogenic for resting lymphocytes</td>
</tr>
<tr>
<td>1960’s</td>
<td>M. Burger</td>
<td>Lectins preferentially agglutinate</td>
</tr>
<tr>
<td>1970’s</td>
<td>G. Nicolson</td>
<td>Animal tumor cells</td>
</tr>
<tr>
<td>1980’s</td>
<td>Kornfeld(s) Osawa</td>
<td>Use of immobilized lectins to analyze animal glycans</td>
</tr>
<tr>
<td>1980’s</td>
<td>Kobata</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cummings</td>
<td></td>
</tr>
<tr>
<td>1980’s</td>
<td>D. Kabelitz</td>
<td>Discovery that plant lectins induce apoptosis</td>
</tr>
<tr>
<td>1990’s</td>
<td>D.J. Gee K. Schweizer</td>
<td></td>
</tr>
</tbody>
</table>

### Milestones in the Discovery of Animal Lectins

- 1860-86: S. Wier Mitchell - agglutination of erythrocytes by rattlesnake venom
- 1902: Agglutination of erythrocytes by Horseshoe crab hemolymph
- 1902: Flexner and Noguchi - “Hemagglutination” by various snake venoms
- 1935-1946: Eel agglutinins used to detect blood groups
- 1960s: Ginsburg & colleagues - treated rat blood leukocytes with bacterial glycosidases and injected into circulation - changes in organ homing seen
- 1960s-70s: Ashwell and colleagues - Liver uptake of circulating glycoproteins with terminal [-Gal or GalNAc residues: the “Asialoglycoprotein receptor”
Milestones in the Discovery of Animal Lectins

- **1970-80s:** Many investigators - discovery of various circulating soluble lectins
- **1970-80s:** Many investigators - isolation of various Gal-binding lectins from different cell types by affinity chromatography (most turn out to be galectins)
- **1960s-70s:** Neufeld & colleagues - a glycan-dependent system mediating uptake of lysosomal enzymes by cells.
- **1970-80s:** Kornfeld, Jourdian, Von Figura & others - discovery of Man 6-P receptors, which recognize phosphorylated N-glycans on lysosomal enzymes.
- **1970-80s:** Stahl, Sly and others - Clearance of proteins via recognition of Man and GlcNAc residues - discovery of the macrophage mannose receptor.

---

Milestones in the Discovery of Animal Lectins

- **Late 1980s:** Kurt Drickamer - C-type and S-type lectin motifs recognized in the primary protein sequence. First classification of lectins based on evolutionary homology rather than on functional characteristics
- **1980-90s:** Roden, Seed, others - Hyaluronan binding properties of cartilage link protein, CD44 etc. discovered
- **Late 1980s-1990s:** Rosen, Laskey, Bevilacqua, Paulson, Lowe, McEver, Varki and others - Discovery and characterization of the Selectin family - first prediction of lectin recognition properties on the basis of primary polypeptide sequence
Milestones in the Discovery of Animal Lectins

- 1990s: Baenziger & colleagues - Clearance system recognizing sulfated GalNAc residues found on pituitary glycoprotein hormones
- 1990s: Bergeron, Helenius, Williams - Calnexin and other proteins within the ER-Golgi pathway discovered to be lectins
- 1990s: Varki, Crocker “I-type lectins” recognized and discovered (Sialic acid-recognizing ones designated “Siglecs” in 1998)
- 2001 Drickamer: Genomics-based definition by of “L-type” and “R-type” animal lectins

Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- **“L-type”** lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins
- **“P-type”** lectins (Mannose-6-Phosphate Receptors)
- **“C-type”** lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins” (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin, Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)
Subcellular Trafficking Pathways for Glycoproteins

Calnexin (and Calcireticulin) function during glycoprotein folding in the endoplasmic reticulum

Improperly folded proteins are re-glucosylated by glucosyltransferase which acts as “sensor” for improper folding

3 Glucose Residues

N-GLYCOSYLATION

Other soluble glycoproteins

Lysosomal enzymes

Calnexin

Glc I

Glc II

Glc III

unfolded

folded

Glc I

Glc II

Glc III
Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calciertulin, Calmegin)
- "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- "P-type" lectins (Mannose-6-Phosphate Receptors)
- "C-type" lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins" (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)
Lectins in the ER-Golgi Pathway

Evolutionary relationships of Animal and Plant L-type Lectins

### Current Classification of Lectins

**Families with known protein sequence homologies**
- **Calnexin group** (e.g., Calnexin, Calciheticulin, Calmegin)
- **"L-type" lectins** (e.g., (ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- **"P-type" lectins** (Mannose-6-Phosphate Receptors)
- **"C-type" lectins** (e.g., Selectins, Collectins etc.)
- *Galectins* (formerly "S-type" lectins)
- **"I-type" lectins** (includes Siglec family)
- **"R-type" lectins** (e.g., GaINAc-SO4 receptors, Plant Lectins)
  - "Eglectins" (Frog Egg lectins)
  - Eel Agglutinins (Fucolectins)
  - Hyaluronan-binding proteins
  - Ficolins
  - Pentraxins

**Sequence homologies not known (Examples)**
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)*

### Selective Trafficking of Lysosomal Enzymes to Lysosomes

[Diagram showing the trafficking pathways from translation and N-glycosylation through the rough endoplasmic reticulum, intermediate compartment, Golgi stacks, trans Golgi network, and secretory granule to early and late endosomes, ultimately leading to lysosomes.](image-url)
Current Classification of Lectins

**Families with known protein sequence homologies**

- **Calnexin group** (e.g., Calnexin, Calciereculin, Calmegin)
- **"L-type"** lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- **"P-type"** lectins (Mannose-6-Phosphate Receptors)
- **"C-type"** lectins (e.g., Selectins, Collectins etc.)
- **Galectins** (formerly "S-type" lectins)
- **"I-type"** lectins (includes Siglec family)
- **"R-type"** lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- **"Eglectins"** (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

**Sequence homologies not known (Examples)**

- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphotericin

*Have defined Carbohydrate Recognition Domains (CRDs)*

---

**Evolution and Diversity of C-type Lectins**

- **A - VERTEBRATE C-TYPE LECTINS**
  - I
  - II
  - III
  - IV
  - V
  - VI

- **B - DROSOPHILA PROTEINS CONTAINING CTLDs**
  - A1 (7)
  - A2 (13)
  - A3 (2)
  - A4 (3)
  - B1 (1)

- **C - CTLD EVOLUTION**

---

THE SELECTINS AND THEIR LIGANDS

Most contain sialyl Lewis*  

SELECTIN LIGANDS

Leukocyte

Endothelial Cell

Platelet

L-selectin  P-selectin  E-selectin

Current Classification of Lectins

Families with known protein sequence homologies

- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- "P-type" lectins (Mannose-6-Phosphate Receptors)
- "C-type" lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins" (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)

- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)
Evolution of the Galectins

Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
  - "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins
  - "P-type" lectins (Mannose-6-Phosphate Receptors)
  - "C-type" lectins (e.g., Selectins, Collectins etc.)
  - Galectins (formerly "S-type" lectins)
  - "I-type" lectins (includes Siglec family)
  - "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins" (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyluronan-binding proteins
- Ficolins
- Pentraxins

Sequences homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectins
- Tachylectins
- Annexins
- Amphoterin
  - Have defined Carbohydrate Recognition Domains (CRDs)
Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- **“L-type”** lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- **“P-type”** lectins (Mannose-6-Phosphate Receptors)
- **“C-type”** lectins (e.g., Selectins, Collectins etc.)
- * Galectins (formerly "S-type" lectins)
- **“I-type”** lectins (includes Siglec family)
- **“R-type”** lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- “Eglectins” (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)  Complement Factor H
- TNF, Interleukins & Cytokines  Ameoba lectin  Tachylectins
- Annexins
- Amphoterin  *Have defined Carbohydrate Recognition Domains (CRDs)
Domain Organization of R-Type CRDs in Nature

Distribution of Lectin Families in Nature

Current Classification of Lectins

Families with known protein sequence homologies

- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins
- "P-type" lectins (Mannose-6-Phosphate Receptors)
- "C-type" lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins" (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)

- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)
“Eglectins”

- “Cortical Granule Lectin” (CGL) characterized in Frog Egg Cortical Granules by Hedrick and colleagues
- N-glycosylated protein delivered through ER-Golgi Pathway
- Binds Galactosyl terminal glycans in Frog Egg Jelly (Ca++ dependent)
- Founding member of family present from ascidian to mammalian sources: embryonic epidermal lectin XEEL, serum lectin, blood group B-active membrane glycoprotein; mouse and human intelectins; human lactoferrin receptor; lamprey serum lectin; ascidian plasma lectin
- Possible Biological Roles
  - Blocking of Polyspermy in Eggs (CGL)
  - Assisting Adhesion? - blood group B-active membrane glycoprotein
  - Receptor Function? - lactoferrin receptor
  - Phagocytosis Stimulating? - ascidian plasma lectin, human and mouse “intelectins”

Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins
- "P-type" lectins (Mannose-6-Phosphate Receptors)
- "C-type" lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GaINAc-SO4 receptors, Plant Lectins)
- “Eglectins” (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)
Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- "P-type" lectins (Mannose-6-Phosphate Receptors)
- "C-type" lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins" (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- *Have defined Carbohydrate Recognition Domains (CRDs)

Ficolins

- Soluble mammalian lectins that can recognize the unique sugar patterns on microbes
- Act directly as opsonins that present bound microbes to phagocytes or indirectly through further opsonization with complement proteins
- Found predominantly in body fluids and at the interface with the environment, e.g. the surfaces of the respiratory and mucosal surfaces.
- Lectin domains clustered at the end of triple helical collagen bundles and multiple clusters are displayed with dimensions and flexibility to maximize interactions with unique sugar arrays of microbial surfaces
- Bind to sugar residues that are rich on microbial surfaces, e.g. N-acetyl-D-glucosamine (GlcNAc)
- Structurally similar to Collectins. But, not Calcium-dependent, and no sequence similarities.
Current Classification of Lectins

Families with known protein sequence homologies
- Calnexin group (e.g., Calnexin, Calcireticulin, Calmegin)
- "L-type" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- "P-type" lectins (Mannose-6-Phosphate Receptors)
- "C-type" lectins (e.g., Selectins, Collectins etc.)
- Galectins (formerly "S-type" lectins)
- "I-type" lectins (includes Siglec family)
- "R-type" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- "Eglectins" (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

* Have defined Carbohydrate Recognition Domains (CRDs)

Pentraxins

- Human serum C-reactive protein (CRP) first reported as precipitator of pneumococcal C polysaccharide. High affinity for phosphorylcholine, but also binds galactans and galactose phosphates via separate site
- "Pentraxin" - a term applied to CRP and its homologue, serum amyloid P component (SAP), reflecting unusual quaternary structure in which five identical polypeptide subunits combine to form a ring with a central hole. A related human pentraxin, PTX3, found in blood cells
- Pentraxins also found in serum of many vertebrate species. Division into CRP or SAP is not on basis of primary structural homology, but on binding preferences.
- Large numbers and diversity of pentraxins, with primary sequence homology to vertebrate CRP/SAP, found in invertebrates, including tunicates, and horseshoe crabs.
**Current Classification of Lectins**

Families with known protein sequence homologies
- **Calnexin group** (e.g., Calnexin, Calcireticulin, Calmegin)
  - **“L-type”** lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
  - **“P-type”** lectins (Mannose-6-Phosphate Receptors)
  - **“C-type”** lectins (e.g., Selectins, Collectins etc.)
- **Galectins** (formerly **“S-type”** lectins)
- **“I-type”** lectins (includes Siglec family)
- **“R-type”** lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- **Eglectins** (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

| *Have defined Carbohydrate Recognition Domains (CRDs) |

**Biosynthesis, Trafficking and Regulation of Animal Lectins**

- All membrane-bound and most soluble lectins synthesized in ER-Golgi pathway. Thus, lectins themselves can be glycoproteins.
- Some soluble lectins (galectins and some cytokines) synthesized on free ribosomes and delivered to cell exterior by as yet poorly understood mechanisms for extrusion thru plasma membrane.
- Makes teleological sense, since several lectins can recognize biosynthetic intermediates that occur in the Golgi-ER pathway
- Some lectins (like galectins) are sensitive to redox state and remain active only in reducing environment of cytosol. Upon entering the oxidizing environment of extracellular space, they must immediately bind to ligands, or become inactivated.
- Some membrane-bound lectins internalized upon ligand binding, with delivery to internal acidic compartments (endosomes). Cargo is released, and receptors can recycle back
Typical trafficking pattern for Endocytic Animal Lectins

Possible interactions of an animal lectin with cognate ligands

Negative regulation by cognate sugar chains present on the same molecule or on the same cell surface
**Soluble and Membrane-bound forms of Animal Lectins**

- Animal lectins exist as soluble and membrane-bound molecules
- Membrane-bound lectins likely involved in endocytosis or cell adhesion, and to stay confined to cell type of original synthesis.
- Soluble lectins are capable of diffusing locally and/or entering the blood circulation.

However:

- Lectins that start out as membrane-bound proteins can be proteolytically shed into the extracellular fluid
- Soluble multivalent lectins can become attached to cell surfaces via their carbohydrate binding sites.

---

**Lectins are Generally Multivalent**

- High avidity generated by multivalent binding of low affinity single sites - a common mechanism for optimizing lectin function
- Until 1990s, all known animal lectins naturally multivalent, because of multisubunit structure, or by having multiple carbohydrate binding sites within a single protein.
- First exception to this general rule were Selectins, which have only a single CRD. Same applies to Siglecs. However, these may become functionally multimeric by non-covalent clustering
- Remains to be seen if any biologically significant binding by any lectin can arise from a monovalent interaction.
Nature of the ligands for animal lectins

- Natural ligands for most lectins are complex glycoconjugates carrying clustered arrays of cognate carbohydrate. Cooperation with clustered lectin binding sites generates high avidity binding. This is probably further enhanced by"mass transport effects" (e.g., high local concentration of ligands, on the cell surface).
- In some instances (e.g., selectins), co-operation with other aspects of underlying polypeptide may be necessary for optimal binding.
- Natural ligands of some animal lectins may be present primarily on foreign invaders (e.g., circulating soluble mannan-binding protein may bind and opsonize microorganisms bearing high densities of mannose, e.g., yeasts and other fungi).
- Few instances known where protein-protein interactions can be mediated by protein modules that are evolutionarily related to the carbohydrate recognition domains (e.g., C-type lectins).

Difficulties with Identifying and naming true ligands for Animal Lectins

- Most natural ligands for animal lectins are glycoproteins
- Names of the underlying polypeptide are commonly used to define the nature of a ligand, e.g., "PSGL-1 is the ligand for P-selectin".
- However, unless it is correctly glycosylated and/or otherwise modified, the polypeptide may not itself be a ligand.
- Recombinant lectins used to identify biological ligands are usually multimeric and/or are presented in multivalent clustered arrays in soluble complexes or on solid supports.
- Thus, while many molecules can bind to a recombinant lectin in a glycosylation-dependent manner, only a few may be actually involved in mediating biologically significant interactions.
- The challenge is to tell the difference between what can bind to a recombinant lectin in vitro, and what does actually bind in vivo to the native lectin in a biologically relevant manner.
## Current Classification of Lectins

### Families with known protein sequence homologies
- **Calnexin group** (e.g., Calnexin, Calciertulin, Calmegin)
- "**L-type**" lectins (e.g., ERGIC-53 and VIP-36 in ER-Golgi pathway, Plant Lectins)
- "**P-type**" lectins (Mannose-6-Phosphate Receptors)
- "**C-type**" lectins (e.g., Selectins, Collectins etc.)
- * Galectins (formerly "S-type" lectins)
- "**I-type**" lectins (includes Siglec family)
- "**R-type**" lectins (e.g., GalNAc-SO4 receptors, Plant Lectins)
- “Eglectins” (Frog Egg lectins)
- Eel Agglutinins (Fucolectins)
- Hyaluronan-binding proteins
- Ficolins
- Pentraxins

### Sequence homologies not known (Examples)
- CD11b/CD18 (beta3-integrin, CR3)
- Complement Factor H
- TNF, Interleukins & Cytokines
- Ameoba lectin
- Tachylectins
- Annexins
- Amphoterin

*Have defined Carbohydrate Recognition Domains (CRDs)*