Lectins

Nathan Sharon, *Weizmann Institute of Science, Rehovot, Israel*

Based in large part on the previous version of this Encyclopedia of Life Sciences (ELS) article, Lectins by Nathan Sharon and Halina Lis.

Lectins, a class of sugar-specific and cell-agglutinating proteins of nonimmune origin that are devoid of enzymatic activity, are ubiquitous in nature. Plant lectins are invaluable tools for the study of carbohydrates, in solution and on cells, and are also employed for purging of bone marrow for transplantation into 'bubble children'. Bacterial cell surface lectins mediate the attachment of the organisms to host cell surfaces in the initiation of infection; their blocking by suitable sugars can serve as a basis of antiadhesion therapy of bacterial diseases. Those of animals control the biosynthesis of glycoproteins, play key roles in cell interactions in the immune system and serve as innate immunity agents against microbial pathogens. They also monitor the migration of leucocytes in blood vessels and contribute to proliferation and metastasis of tumour cells.

Introduction

Lectins (from Latin, *legere*, to select or choose) are proteins that bind mono- and oligosaccharides specifically and reversibly similarly to anticarbohydrate antibodies, but are not products of an immune response. However, they differ from antibodies in several important aspects. Thus, numerous lectins are present in plants, microorganisms and viruses, which are not capable of an immune response. Another marked difference between the two classes of protein is that antibodies are structurally similar, whereas lectins are structurally diverse. In general, lectins are oligomeric proteins composed of subunits, one or more of which carries a sugar-binding site. They vary, however, in size, amino acid composition, metal requirement, domain organization, subunit number and assembly, as well as in three-dimensional structure (**Figure 1**). In their structural diversity, lectins are akin to enzymes, although they are devoid of catalytic activity. In spite of their structural variation, lectins can be grouped in families of homologous proteins.

Lectins typically contain two or more carbohydrate-combining sites per molecule, that is, they are divalent or polyvalent, although exceptions also occur. Therefore, binding of a lectin to sugars on the surface of cells, for example, erythrocytes may cause cross-linking of the cells and their subsequent precipitation, a phenomenon referred to as cell agglutination. The erythrocyte agglutinating, or haemagglutinating, activity of lectins is a major attribute of these proteins and serves routinely for their detection and characterization. It was actually by this activity that lectins were first detected in extracts of plant seeds at the turn of the nineteenth century. For a long time such haemagglutinating proteins were known as 'phytohaemagglutinins', because they were found almost exclusively in plants. A turning point in lectin research came in 1936 with the work of James B Sumner on jackbean lectin, concanavalin A, still the best-characterized protein of this class. He reported that concanavalin A also precipitates polysaccharides and glycoproteins and that both the haemagglutinating and precipitating activities are inhibited by mannose and glucose (the sugars are of the D-configuration except for fucose which is L). With much foresight, Sumner suggested that these activities might be the consequence of a reaction of the lectin with carbohydrates on the erythrocyte surface. In fact, testing the inhibition of haemagglutination or polysaccharide precipitation by a panel of sugars is still the simplest way to establish the specificity of a lectin. **See also**: Membrane Proteins; Sumner, James Batcheller

Another turning point was the discovery, in the late 1940s, by William C Boyd (who coined the term lectin) and independently by Karl O Renkonen, that certain lectins exhibit blood type A, B or O specificity. Soon thereafter, such lectins played a crucial role in the identification, by Walter JT Morgan and Winifred M Watkins, of the chemical nature of the blood-type ABO determinants: α-*N*-acetylgalactosamine for the A type, α-galactose for the B type and α-fucose for the O type. However, lectin research started to gain momentum only in the 1960s, thanks to two major developments. The first was the finding by Peter C Nowell that some lectins are mitogenic, that is that they stimulate lymphocytes to undergo mitosis. This discovery had a revolutionary impact on immunology in that it
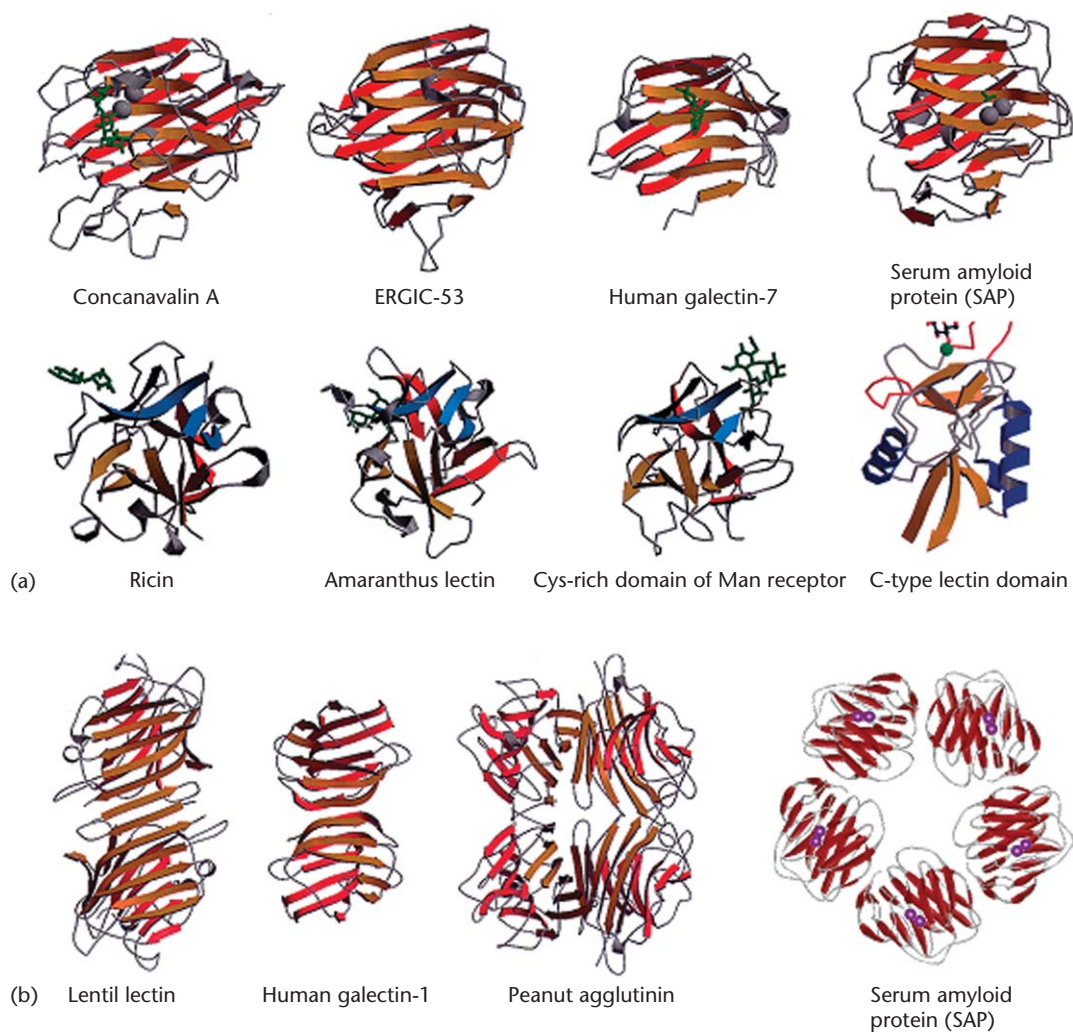
Figure 1 Structures of different lectins represented as ribbon diagrams. (a) Upper row shows monomers of plant (concanavalin A) and animal lectins (ERGIC-53, human galectin-7 and serum amyloid protein) that share the jellyroll or lectin fold. First three (from left) in the lower row all exhibit the β-trefoil fold; the first two of these are from plants, the third (and the fourth) from animals. (b) Variations in quaternary lectin structures. In all figures, the grey spheres represent metal ions; bound carbohydrate is shown in ball and stick representation. Reprinted with permission from Loris R (2002) Principles of structures of animal and plant lectins. *Biochimica et Biophysica Acta* **1572**: 198–208. Elsevier, Oxford.

shattered the view, held until then, that lymphocytes are dead-end cells that could neither divide nor differentiate further. The second stimulus was the accumulation of evidence that lectins are powerful tools for the structural and functional investigation of complex carbohydrates, especially glycoproteins, for the study of their biosynthesis and for the examination of changes that occur on cell surfaces during physiological and pathological processes, from cell differentiation to cancer. **See also**: Glycoproteins; Lymphocytes

Lectins are now known to be present in most organisms, ranging from viruses and bacteria to animals, although appreciation of their wide occurrence in higher animals came only with the discovery of the selectins (see later discussion) in the late 1980s. They are readily obtainable in purified form, mostly by affinity chromatography on immobilized sugars, a method pioneered by Irwin J Goldstein in 1965, and recently also by recombinant deoxyribonucleic acid (DNA) techniques. Numerous lectins have been isolated and characterized (**Table 1**), and the high-resolution three-dimensional structures of close to 200 have been established, almost all also in complexes with specific carbohydrates (e.g. see **Figure 1**). As well as varying in their primary, secondary, tertiary and quaternary structures, lectins differ also in the constitution of their binding sites. Nonetheless, many of them belong to distinct protein families with similar sequences and structural features. These families are often along taxonomic lineages (e.g. legume and cereal lectins), although sometimes common features are found across phylogenetic barriers as well (as in the case of structurally similar legume lectins and animal galectins). Indeed, sequence similarities with known lectins

Table 1 Some well-studied lectins[a]

Specificity	Source (common abbreviated name in parentheses)		
	Plants[b]	Animals	Microorganisms
Mannose[c]	Jackbean (conA)	Human serum (MBP)	*E. coli* type 1 fimbriae
	Lentil (LCL)		
	Pea (PSL)		
	Snowdrop (GNL)		
N-Acetylglucosamine	Wheat germ (WGA)	Avian liver	*E. coli* type 17F fimbriae
Galactose/*N*-acetylgalactosamine	Castor bean (RCA)		
	Coral tree seed (ECorL)	Rabbit liver (HBP, RHL)	Amoeba
		Garden snail	Slime mould
	Lima bean (LBA)	Vertebrates (galectins)	*Actinomyces naeslundii type II*
	Peanut (PNA)[d]		
	Soybean (SBA)		
Fucose	Asparagus pea	Eel	
		Rat liver	
Sialic acid	Elderberry (SNL)	Horseshoe crab	*Influenza virus*
		Lobster	*Polyoma virus*

[a]This table does not include lectins that bind only oligosaccharides. Examples are phytohaemagglutinin (PHA, specific for Gal(β1–4)GlcNAc(β1–2)[Gal(β1–4)GlcNAc(β1–6)]Man); potato and tomato lectins (specific for [GlcNAc(β1–4)$_{2–4}$]; *E. coli* type P lectin (specific for Gal(α1–4)Gal); cyanovirin-N (specific for oligomannosides), selectins (specific for sLe^x and sLe^a, see **Figure 7**), dectin-1 (specific for β-glucans) and calnexin and calreticulin (specific for GlcMan$_9$GlcNAc$_2$).
[b]All the plant lectins listed, and many others, are in wide use as carbohydrate-specific reagents (see section on Applications).
[c]Most lectins in this group also bind glucose, often with similar affinity to that of mannose; exceptions among those listed in the table are the snowdrop lectin and type 1 fimbriae that do not bund glucose.
[d]Does not bind *N*-acetylgalactosamine.

provide a novel guideline for the detection and identification of new ones. **See also**: Affinity Chromatography

At present, lectins are the focus of intense attention because of the realization that they act as recognition determinants in diverse biological processes. These include control of intracellular protein glycosylation and glycoprotein traffic, clearance of glycoproteins from the circulatory system, adhesion of infectious agents to host cells, recruitment of leucocytes to inflammatory sites, as well as cell interactions in the immune system, in malignancy and metastasis. The basis of this recognition is a molecular fit between pairs of complementary structures, in this case the combining site of a lectin and a carbohydrate (often referred to as the lectin ligand), on the surfaces of interacting cells or between those on a cell and on a molecule in solution. It is analogous to that between an enzyme and its substrate (or inhibitor) and between an antibody and antigen, and can be blocked by appropriate sugars not only *in vitro* but also *in vivo*. This inhibition is the basis of current attempts to develop novel therapies for microbial infections, inflammation and other diseases. **See also**: Antigen–Antibody Complexes; Membrane Proteins

Carbohydrate Specificity

As mentioned, the carbohydrate specificity of lectins is customarily established examining the ability of different monosaccharides, oligosaccharides or glycopeptides to inhibit either haemagglutination or polysaccharide (or glycoprotein) precipitation by the lectin. The most effective inhibitor defines the specificity of the lectin tested. Alternatively, the specificity can be examined with the aid of carbohydrate microarrays (or glycoarrays) that consist of collections of tens to hundreds of different carbohydrates immobilized on microtiter plates or other kinds of solid support; here, the specificity is defined by the carbohydrate to which the lectin binds best.

Of the numerous monosaccharides found in nature, almost all lectins recognize just a few, primarily mannose, glucose, galactose, *N*-acetylglucosamine, *N*-acetylgalactosamine, fucose and *N*-acetylneuraminic acid; in addition, they combine specifically with a large number of the oligosaccharides composed of these monosaccharides. A striking feature of the above mentioned monosaccharides is that they are all regular constituents of animal glycoconjugates and are also present on the surfaces of eukaryotic cells, including erythrocytes.

Based on their specificity, lectins are classified in the following five groups: mannose, galactose/*N*-acetylgalactosamine, *N*-acetylglucosamine, fucose and *N*-acetylneuraminic (sialic) acid (**Table 1**). Many lectins specific for mannose also react with glucose (two monosaccharides that differ in the configuration of the hydroxyl at their carbon-2) but none of these reacts with galactose (which differs from glucose in the configuration of the hydroxyl at carbon-4), nor do those

specific for galactose bind mannose. Similarly, members of the *N*-acetylglucosamine specificity group do not combine with *N*-acetylgalactosamine (or vice versa). However, most lectins that bind galactose interact also with *N*-acetylgalactosamine, in some cases preferentially. For this reason they are classified in a single specificity group, Gal/GalNAc, even though a few (e.g. peanut lectin, PNA) do not bind *N*-acetylgalactosamine at all. Lectins of the same specificity group may differ markedly in their affinity for various derivatives of the corresponding monosaccharide, especially oligosaccharides, while certain lectins interact with oligosaccharides only (see **Table 1**). Binding of monosaccharides to lectins is usually weak, with association constants in the range 10^3 to $5 \times 10^4 \, \text{mol L}^{-1}$ whereas the affinity of oligosaccharides is often much stronger (up to 1000 times).

Plant Lectins

Plants continue to be the richest source of lectins, with several hundred well studied to date. Many are from the plant seeds, but they have been obtained also from virtually all kinds of vegetative tissues, such as bark, leaves, stems, fruits and roots. When isolated from the same family and tissue, they are structurally similar, but may have different specificities (e.g. concanavalin A and the coral tree lectin, both from leguminous seeds, are specific for mannose and galactose, respectively).

Legumes

Lectins of leguminous seeds comprise the largest family of proteins of this class, with approximately 200 members. They all consist of two or four nearly identical subunits of 25–30 kDa, each with a single carbohydrate-combining site and one Ca^{2+} and one Mn^{2+} per subunit that are essential for carbohydrate binding. They have similar amino acid sequences, which include a number of conserved residues, among them most of those that participate in the binding of the ligand and of the metal ions. Concanavalin A occupies a special position since, as a result of unusual steps during biosynthesis, its N-terminal half is homologous to the C-terminal half of the other legume lectins, and its C-terminal half to their N-terminal half, a phenomenon known as 'circular homology' (see later discussion).

The three-dimensional structures of the subunits of all legume lectins examined are almost identical, and are also similar to those of certain animal lectins, although their primary structures are different. However, their quaternary structures vary (**Figure 1**). The subunits are dome-shaped, with the carbohydrate-combining site located at the top of the dome, in proximity to the metal ions. The invariant binding site residues (an aspartic acid, an asparagine and an aromatic amino acid) occupy identical spatial positions in all these lectins. Discrimination between closely related sugars, for example glucose/mannose and galactose, is achieved by the different positioning of the respective monosaccharides in the combining sites (**Figure 2a**), due to
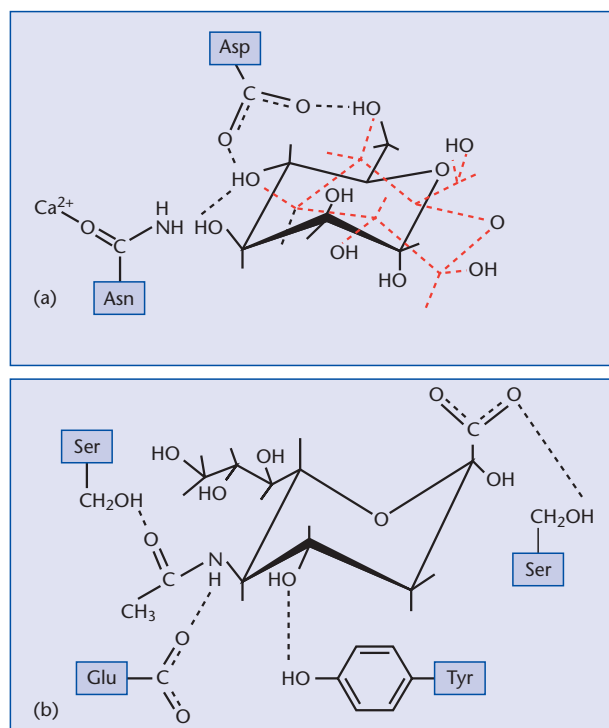

Figure 2 (a) Binding site of legume lectins. Key hydrogen bonds with amino acid side chains (dotted lines) holding galactose (red lines) in the combining site of a galactose-specific lectin (e.g. from the coral tree) and glucose (solid line) in a glucose-/mannose-specific lectin (e.g. concanavalin A or pea lectin). (b) *N*-acetylneuraminic acid in the combining site of wheat germ agglutinin.

differences in the structure of the variable amino acids that line the binding pocket or surround it. In some cases, the combining sites of animal lectins (e.g. ERGIC-53) are similar to those of the legume lectins.

The legume seed lectins are coded by single genes and are synthesized as proproteins or pre-proproteins, which after processing are converted into the mature lectins. This includes cotranslational cleavage of the signal peptide and posttranslational proteolytic removal of a short C-terminal peptide. Concanavalin A, one of the few legume lectins that are not glycoproteins (i.e. do not contain covalently linked carbohydrate chains in their final, mature form), is exceptional in that it is formed from a glycosylated precursor that is processed by a series of rare steps that end up with the rearrangement of the N- and C-terminal halves of the molecule (**Figure 3**).

Cereals

Another family of plant lectins is that of the cereals, a prominent example of which is wheat germ agglutinin (WGA). The lectins of this family are all specific for *N*-acetylglucosamine and *N*-acetylneuraminic acid and consist too of two identical subunits, but they differ markedly from those of the legumes. For instance, they are exceptionally rich in cysteine, which in legume lectins is generally absent, and they are devoid of metals. In WGA, the only
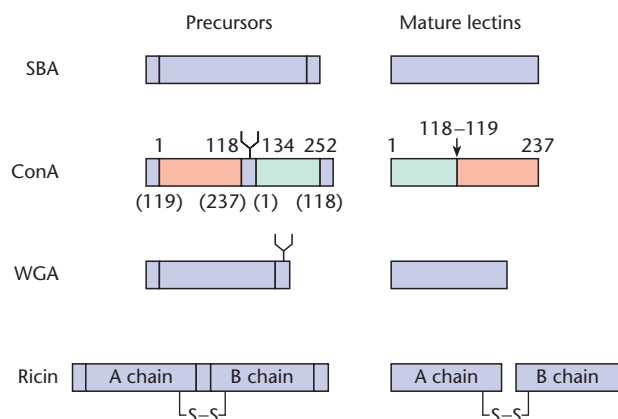
Figure 3 Schematic representation of the precursors of soybean agglutinin (pre-SBA), concanavalin A (pre-ConA), wheat germ agglutinin (pre-WGA) and ricin (pre-ricin) and of their processing to the mature lectins. Arrows denote positions of cleavages, and numbers in parentheses in pre-ConA give the corresponding positions in mature concanavalin A. Processing of pro-SBA and pro-WGA (like that of the precursors of most legume and cereal lectins) involves sequential cleavage of the signal peptide and of the C-terminal peptide (which in the cereal lectins is glycosylated). Pre-ConA contains in addition to the earlier peptides a glycosylated spacer. After removal of the signal peptide, the precursor is deglycosylated; an endopeptidase then cleaves the spacer and residues 118 and 119 are ligated enzymatically, with the concomitant removal of the C-terminal peptide, resulting in the rearrangement of the primary sequence of the precursor. In pro-ricin, the polypeptides of the two subunits (A and B) are separated by a spacer peptide, but already linked by an S–S bridge, and processing involves the cleavage of the signal peptide, the C-terminal and the spacer peptides (N-linked oligosaccharide).

Figure 4 N-Acetylneuraminic acid in the combining site of influenza virus lectin.

cereal lectin characterized in molecular detail, each subunit is made up of four homologous and similarly folded sub-domains, with four identically positioned disulfide bridges. There are four combining sites per lectin dimer, each located at the interface of the subunits, and formed by amino acids from both subunits. The latter include several tyrosines that interact with the bound sugar both hydrophobically and via hydrogen bonds, as well as two serines and a glutamic acid, the side chains of which form bonds of the latter type with the ligand (**Figure 2b**).

The cereal lectins are coded by several closely related genes. Although not glycoproteins, they too are formed from glycosylated precursors that are processed by the co-translational removal of the signal peptide and posttranslational cleavage of a glycosylated C-terminal peptide.

Other plant lectins

Beans of the castor tree (*Ricinus communis*) contain two closely related lectins, ricin and *R. communis* agglutinin, RCA; the former is one of the deadliest poisons known and, according to some estimates, a single molecule is sufficient to kill a cell. Ricin is a heterodimeric glycoprotein with a molecular mass of 60 kDa, composed of two S–S-linked chains, A and B. The cytotoxic activity resides in the A chain, which acts by enzymatically inactivating the ribonucleic acid (RNA) 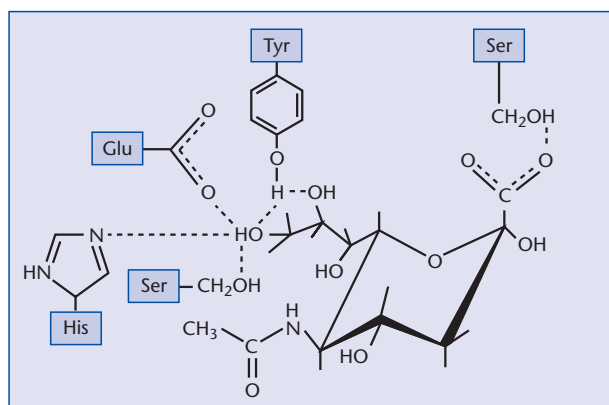involved in protein synthesis. The B chain is made up of two topologically similar globular domains, each binding galactose (and lactose) in a shallow cleft. As with the WGA subunit, this chain is stabilized by several disulfide-linked cysteines. RCA is a dimer of two subunits, each of which is similar to the ricin molecule, but it is not toxic.

Ricin and RCA are encoded by a multigene family of about eight members. The A and B chains are derived from a single polypeptide precursor through the excision of a linker sequence between the C-terminus of the A chain and the N-terminus of the B chain, but remain linked by a disulfide bridge.

Microbial Lectins

Several viruses, for example, those of influenza, polyoma, Sendai and Newcastle disease are classified as sialic acid-specific. They carry on their surfaces lectins (or haemagglutinins) the preference of which to particular sialyl-oligosaccharides determines their host tropism. Thus, strains of influenza virus A that are specific for N-acetylneuraminic acid linked α2–6 to galactose are infectious to humans, whereas those specific for N-acetylneuraminic acid linked α2–3 to galactose are infectious to chicken but not to humans. The influenza virus lectin is a trimer that consists of subunits composed of two polypeptides, HA1 and HA2, covalently linked by a disulfide bond. Each subunit contains a globular domain, made up of HA1 only, on top of an elongated stem, projecting 135 Å from the viral membrane. The globular domain contains the combining site of the lectin, which is in the form of a pocket composed of amino acids that are largely conserved in the numerous strains of the virus. Binding of N-acetylneuraminic acid is by hydrogen bonds to the hydroxyls of two serines and one tyrosine, the carboxylate of a glutamic acid and a ring nitrogen of histidine, as well as by hydrophobic interactions with the indole ring of the tryptophan (**Figure 4**). **See also**: Influenza Viruses

Many bacterial species express surface lectins, usually in the form of fimbriae (or pili), a few hundreds of which protrude from their cell surface. These filamentous, multisubunit organelles, a several nanometres in diameter and 100–200 nm in length, consist typically of helically arranged subunits (pilins) of several different types. Only one of the subunits, usually a minor component of the fimbriae, possesses a carbohydrate-binding site, for example, for mannose (in type 1 fimbriae of *Escherichia coli* and related organisms) or galabiose, Gal(α1–4)Gal (in P fimbriae, also of *E. coli*) and *N*-acetylglucosamine (in *E. coli* F17). The specificity of the bacterial surface lectins may determine the animal tropism of the bacteria. A case in point is *E. coli* K99, specific for *N*-glycolylneuraminic acid. This strain can cause diarrhoea in piglets and calves but not in humans that lack *N*-glycolylneuraminic and that have instead the closely related *N*-acetylneuraminic acid. **See also**: Bacterial Pili and Fimbriae

Besides the surface lectins discussed earlier, a small number of soluble bacterial lectins have been described, of which the recently isolated cyanovirin-N deserves special mention. It is an oligomannose-specific cyanobacterial lectin that potently inactivates all strains of human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) and has a high potential for use as a systemic compound to control viral load and to prevent primary viral infection.

Animal Lectins

Animal lectins started to attract attention only around 1990 and their study is currently growing at an explosive rate, faster than any other area of lectin research. It is greatly facilitated by (a) the application of recombinant DNA techniques, which ease the search for new lectins and make animals deficient in particular lectins accessible for investigation of lectin function; (b) the vast improvement in methods of structural analysis of the carbohydrates of glycoconjugates and (c) the availability of glycoarrays with hundreds of different saccharides for the detection of lectins and identification of their specificity.

Well over 200 animal lectins have been identified, mostly from vertebrates, and the functions of a considerable number of these has been clarified. Their carbohydrate-binding activity often resides in a limited polypeptide segment, designated the carbohydrate-recognition domain (CRD). Several CRD types have been discerned, each of which shares a pattern of invariant and highly conserved amino acid residues at a characteristic spacing. The CRDs are often highly evolutionarily conserved.

The animal lectins are classified in a number of families and superfamilies as described in the next sections (**Table 2**).

Galectins

The galectins are a family of soluble, animal lectins specific for β-galactosides, such as lactose and

Table 2 Functions of animal lectins

Function	Lectins involved
Glycoprotein folding, sorting and secretion, targeting, degradation and clearance	Calnexin, calreticulin, ERGIC-53, RHL VIP-36, Man-6-P receptors
Cell–cell interactions, signalling and transport; immune regulation; malignancy and metastasis	Galectins, selectins, siglecs
Recognition and effector mediators in innate immunity	MBL, SP-A and SP-D, dectin-1, DC-SIGN

N-acetyllactosamine. They are found both inside the cytoplasm and nucleus of cells and occasionally also on the cell surface and outside the cell. The expression of galectins in the organism is developmentally regulated, that is their synthesis in a given tissue occurs only during particular developmental or physiological stages.

The galectins occur as monomers and homodimers of subunits with molecular masses of approximately 14 kDa, as well as larger polypeptides (30–35 kDa), each of which contains one or two copies of a carbohydrate recognition domain, known as S-CRD. Their three-dimensional structure is similar to that of the legume lectins, despite the absence of significant sequence homology, the lack of metal ions and a different location and structure of the combining site (**Figure 1**). Of particular importance in the binding of the carbohydrate to galectins are hydrogen bonds between the C4 hydroxyl group of the ligand and the side chains of three amino acids (histidine, asparagine and arginine), invariant among all galectins sequenced, as well as hydrophobic interactions with a likewise conserved tryptophan (**Figure 5a**).

Calnexins

Calnexin and calreticulin are related proteins that participate in an endoplasmic reticulum (ER) chaperone system. This system ensures the proper folding and quality control of newly synthesized glycoproteins. The specificity calnexin and calreticulin for glycoproteins is conferred by a combining site that recognizes an early oligosaccharide processing intermediate $Glc_1Man_9GlcNAc_2$ covalently attached to the folding immature glycoprotein. The above lectins, and related ones, associate with most, if not all, glycoproteins that pass through the ER. **See also**: Glycoproteins

C-type lectins

The C-type lectins form a superfamily of animal lectins that require calcium ions for their activity. A characteristic feature of all C-type lectins is that they are composite, multidomain molecules, in which the CRD is attached to a variable number of polypeptide domains of different kinds. Most lectins of this superfamily contain a hydrophobic
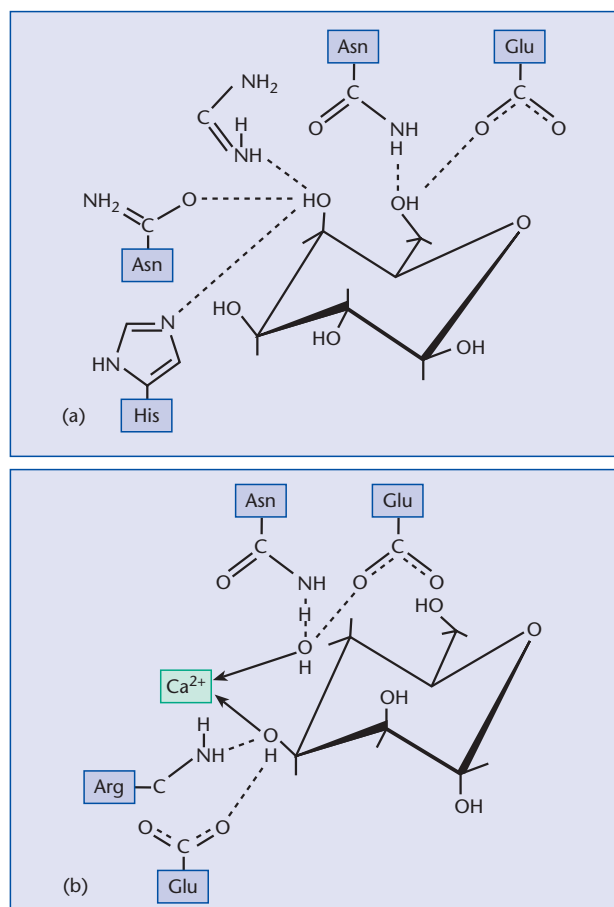
Figure 5 Combining sites of animal lectins. (a) Bovine galectin with bound galactose. (b) Mannose-binding protein with bound mannose. Hydrogen bonds are shown as broken lines and coordination bonds as dotted lines.
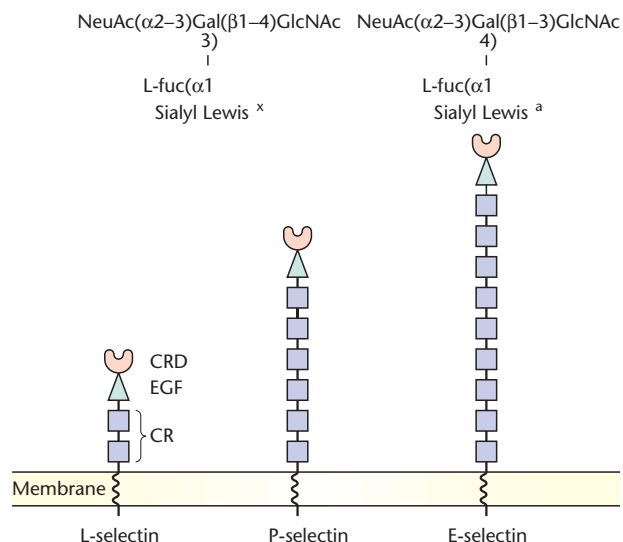

Figure 6 Schematic representation of the multidomain structure of the selectins. CR, complement regulatory repeats; EFG, epidermal growth factor-like repeats and CRD, carbohydrate-recognition domain. Structures on top are of the selectin ligands.

stretch by which they are anchored to the membrane. C-type lectins are grouped into three main families – selectins, collectins and endocytic lectins.

Selectins

This family consists of three members, E-, P- and L-selectin (**Figure 6**). They are so named because they mediate selective contacts between leucocytes and endothelial cells in blood vessels, and thus monitor the migration of the leucocytes to sites of inflammation. All selectins recognize the sialic acid containing tetrasaccharide sLe^x and to a lesser extent its positional isomer, sLe^a (**Figure 6**). For P- and L-selectins, the biological ligand also requires the presence of a sulfate group.

Collectins

The collectins, a group of five soluble proteins found in the serum of mammals and birds, received their name because they contain a collagen-like domain. An important member of this group is the mannose (or mannan)-binding protein MBP (or MBL), the structural unit of which is a trimer

of 32 kDa subunits based on a triple helix formed by its collagenous portion (**Figure 7**). MBP circulates in serum of rodents and man as a hexamer of such trimers. As shown by X-ray crystallography, it binds mannose via Ca^{2+} that serves as the nucleus of the combining site and interacts with the 3- and 4-OH of the ligand (**Figure 5b**). Four of the five additional bonds that coordinate the metal ion are provided by the side chains of two glutamic acids and two asparagines that also are H-bonded to the same (3 and 4) mannose hydroxyls. In all C-type lectins specific for mannose, two of the four aforementioned amino acids are in the sequence Glu-X-Asn, whereas in galactose-specific C-type lectins, the amide and carboxyl groups are switched and the corresponding sequence is Gln-X-Asp. **See also**: Macromolecular Structure Determination by X-ray Crystallography

Endocytic lectins

A prominent representative of the endocytic lectins is the rabbit hepatic asialoglycoprotein receptor (RHL), known also as hepatic-binding protein (HBP), the first mammalian lectin to be described, in 1974. It is found on hepatocytes of different mammals and is specific for galactose and *N*-acetylgalactosamine, whereas its avian homologue is specific for *N*-acetylglucosamine. Another endocytic lectin is the mannose-specific lectin (receptor) of macrophages. In contrast to the other members of this family that are type II transmembrane proteins (i.e. with the N-terminal in the cytoplasm and the C-terminal outside the cell), it is a type I protein (i.e. the N-terminal is outside). The mannose-specific lectin of the macrophages is also exceptional in that it contains eight CRDs in its extracellular domain.
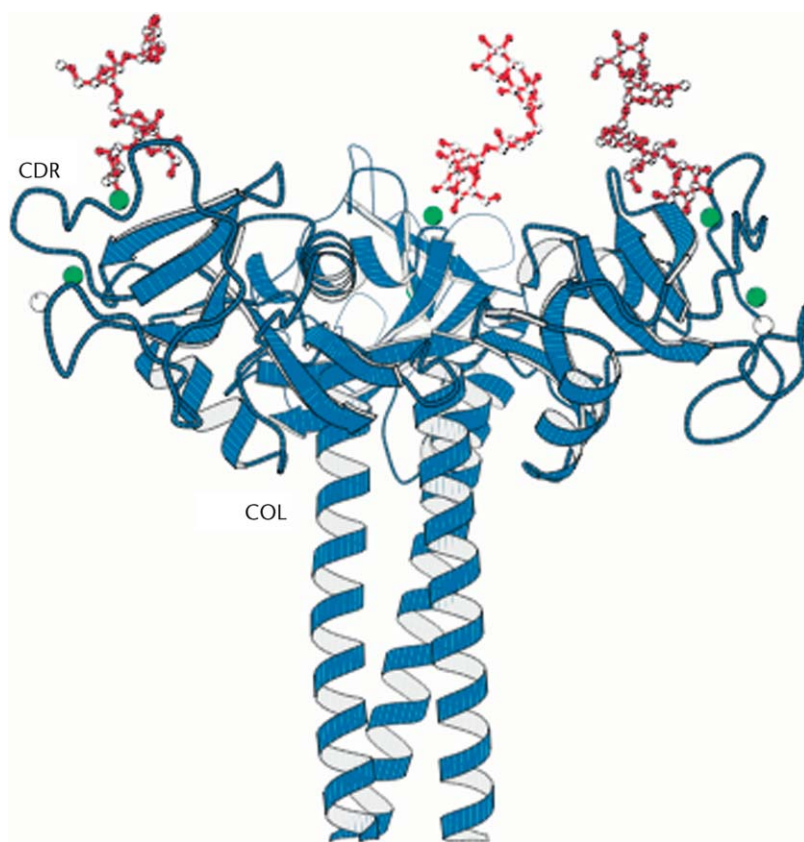
Figure 7 Bouquet-like structure of MBP with mannose oligosaccharides bound to its three carbohydrate recognition. The green and white circles next to the sites denote differently attached calcium ions. COL, collagenous region and CRD, carbohydrate-recognition domain. Reprinted with permission from Weis WI and Drickamer K (1994) Trimeric structure of a C-type mannose-binding protein. *Structure* **2**: 1227–1240.

P-type lectins

A carbohydrate-recognition domain, different from S- and C-CRD, and referred to as P-type, is present in two closely related lectins that bind mannose-6-phosphate (Man-6-P), a sugar found almost exclusively on lysosomal enzymes; they are known as Man-6-P receptors. One of these lectins is of high molecular mass (approximately 300 kDa), and does not require cations for activity; the other is of low mass (approximately 45 kDa) and is Ca^{2+} dependent. Both are type I transmembrane glycoproteins found on membranes of intracellular organelles and on the plasma membrane. The extracellular domain of the high molecular mass receptor contains 15 contiguous, homologous repeating units of which two possess high-affinity binding sites for Man-6-P, whereas that of the low molecular mass receptor contains a single Man-6-P-binding domain similar in size and in sequence to the repeating units of the high molecular mass receptor.

Lectins as Recognition and Adhesion Molecules

As mentioned, the main role of lectins is in cell recognition and adhesion. Accordingly, legume lectins may function in the establishment of symbiosis between nitrogen-fixing bacteria, mainly rhizobia, and leguminous plants, where they serve for the attachment of the rhizobia to the plant roots. Since some plant lectins are toxic to phytopathogenic fungi and insects, it has been suggested that they may serve in defence of plants against these pests. The evidence for these functions is controversial, however.

The viral and bacterial surface lectins mediate the adhesion of the organisms to host cells, a prerequisite for infection to occur (**Figure 8**). This was first demonstrated in the 1940s for the influenza virus haemagglutinin. Thus, removal of *N*-acetylneuraminic acid (for which the viral lectin is specific) from the cell membranes abolishes viral binding and prevents infection, whereas reattachment of the monosaccharide restores the ability of the cells to bind the virus and to be infected. **See also**: Immune Response: Evasion and Subversion by Pathogens

Compelling evidence for the role of lectins in bacterial infection derives from experiments in which blocking the lectins by suitable sugars provided protection against infection in animals. As early as 1979, it was shown that injection of mannose and methyl α-mannoside into the urinary tract of mice or rats prevented infection of the animals by type 1 fimbriated bacteria; glucose (or methyl α-glucoside), which is not an inhibitor of type 1 fimbriae, had
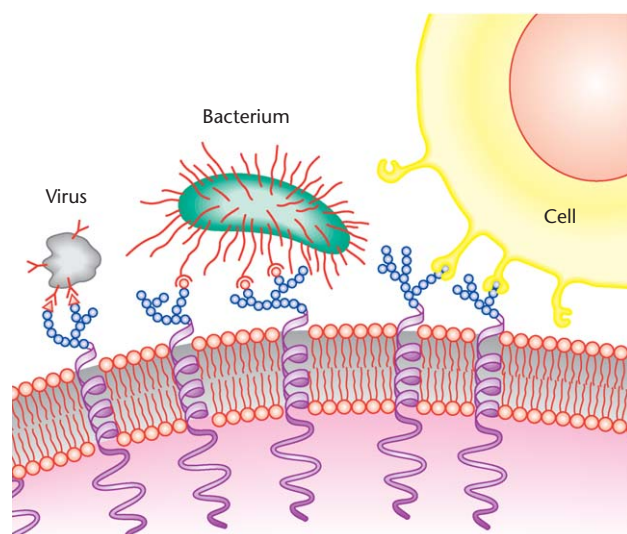
Figure 8 Surface lectins mediate cell adhesion by binding to corresponding carbohydrates on apposing cells. Reproduced from Sharon and Lis (1993) Carbohydrates in cell recognition. *Scientific American* **268**(1): 82–89.

no effect. These results were subsequently confirmed and extended to other experimental systems such as prevention by the tetrasaccharide lacto-*N*-neotetraose of lung infection in rabbits by *Streptococcus pneumoniae* and by NeuAcα(2–3)Galβ(1–4)Glc (sialyllactose) of stomach infection in monkeys by *Helicobacter pylori*. They serve as an impetus for attempts to develop antiadhesive drugs for the treatment of microbial infections in humans.

Galectins are thought to be essential for the normal development and differentiation of all multicellular organisms, acting as cell-adhesion molecules. Galectin–glycoprotein interactions *in vitro* and *in vivo* are required for critical physiological and cellular functions. In some cell types, this formation of galectin–glycan lattices or scaffolds is needed for organizing plasma membrane domains, such as lipid rafts, or for targeted delivery of glycoproteins to the apical or basolateral surface of epithelial cells. Galectin–glycan lattice formation is also involved in regulating the signalling threshold of some cell-surface glycoproteins, including T-cell receptors and growth factor receptors. Such lattices can determine receptor residency time by inhibiting endocytosis of glycoprotein receptors from the cell surface, thus modulating the magnitude or duration of signalling from the cell surface. The elevated level of a particular galectin present on the surface of metastatic murine and human cancer cells may be responsible for the adhesion of these cells to target organs, a step necessary for metastasis to occur. **See also**: Adhesive Specificity and the Evolution of Multicellularity

The siglecs perform a variety of functions, primarily in the immune system. Thus, siglec-1 (sialoadhesin) may act in innate immunity as a macrophage pattern-recognition molecule for sialylated pathogens, such as *Neisseria meningitidis* and *Trypanosoma cruzi*, that coat themselves with sialic acids in an attempt to evade other forms of immune recognition from the infected host. Siglec-2 (CD22) is a well-established regulator of B-cell activation. The CD33-related siglecs act as signalling proteins, and appear in both inhibitory and activating forms. Siglec-4 (myelin-associated glycoprotein, MAG) plays a role in the neural system, such as in axon–glia and/or glia–glia interactions.

The selectins serve as another paradigm for the role of sugar–lectin interactions in biological recognition. They mediate the primary adhesion of circulating leucocytes to endothelial cells of blood vessels, a prerequisite for the exit of the former cells from the circulation and their migration into tissues. L-selectin is found on all leucocytes and is involved in the recirculation of lymphocytes, directing them specifically to peripheral lymph nodes or other secondary lymph organs. E- and P-selectins are expressed on endothelial cells only when these cells are activated by inflammatory mediators (e.g. histamine, interleukin 2 or tumour necrosis factor) released from tissue cells in response to infection or ischaemia, for example. They then trap the leucocytes via their surface sugars and thus enable the leucocytes to leave the blood vessels on their way to the infected or otherwise damaged tissues, where they act to eliminate bacteria and other intruders. As expected, individuals unable to synthesize the selectin ligands sLe^x and sLe^a or their sulfo derivatives suffer from recurrent bacterial infections. The mechanism that helps the leucocytes to breach the endothelial barrier as a step in the fulfilment of their infection-fighting duties is also responsible for their undesirable accumulation in tissues where they do not belong, for instance joints, thereby causing tissue damage, swelling and pain. Another example is reperfusion injury, in which the leucocytes destroy tissues damaged by temporary lack of oxygen, as happens during a heart infarct or cerebral stroke. Prevention of adverse inflammatory reactions by inhibition of leucocyte–endothelium interactions, another application of antiadhesion therapy, has become a major aim of several pharmacological efforts. As shown in animal models, oligosaccharides recognized by the selectins protect against experimentally induced lung injury and tissue damage caused by myocardial ischaemia and reperfusion. In addition to their involvement in inflammation, selectins may play a role in the spread of cancer cells from the primary tumour and formation of metastases throughout the body. **See also**: Histamine Biosynthesis and Function; Interleukins; Stroke; Tumour Necrosis Factors

A particularly well-characterized participant in the innate immune system of humans and other animals is MBL (or MBP). It too functions as a pathogen-recognition molecule, opsonizing infectious microorganisms and initiating the complement cascade. MBL-deficient patients or even those that are partially deficient suffer from repeated infections. Low levels of MBL are often associated with a wide range of infections as well, and also with noninfectious diseases including systemic lupus erythematosus, rheumatoid arthritis, cystic fibrosis and common variable immunodeficiency. MBL-replacement therapy may be beneficial in affected MBL-deficient patients, and

recombinant human MBL (rhMBL) is in development as a therapeutic approach for this purpose.

The human lung-associated collectins, SP-A, SP-D, DC-SIGN and the mannose macrophage-binding receptor (MMR), contribute significantly to innate immunity against pulmonary infections by pathogens such as *Klebsiella pneumoniae* and *S. pneumoniae*.

Calnexin, calreticulin and other ER-associated lectins (among them ERGIC-53) function in the ER to assure correct folding of the glycoproteins formed in these subcellular organelles. The Man-6-P receptors target lysosomal enzymes to their subcellular compartment. A defect in the synthesis of the Man-6-P marker results in I-cell disease (also called mucolipidosis II or MLII), an extremely rare and lethal inherited lysosomal storage disease. It is characterized by a lack in the lysosomes of the afflicted individuals of all those enzymes that normally carry the marker and result in the secretion of the enzymes in the urine and in the intracellular accumulation of undigested glycoconjugates.

Applications

Native lectins, many of which are commercially available, are used predominantly for applications that are based on precipitation and agglutination reactions or for mitogenic stimulation of lymphocytes (see later discussion); lectin derivatives (also commercially available) are required for a variety of other purposes. Thus, lectins derivatized with fluorescent dyes, gold particles or enzymes are employed as histochemical and cytochemical reagents for detection of glycoconjugates in tissue sections, on cells and subcellular organelles, and in investigations of intracellular pathways of protein glycosylation. Lectin histochemistry is also employed in the investigation of cancer-specific carbohydrate markers. Lectins have been used to demonstrate that membrane receptors for hormones, growth factors, neurotransmitters and toxins are glycoconjugates. Immobilized lectins, such as those that are covalently bound to Sepharose, are indispensable for the purification and isolation by affinity chromatography of glycoproteins, glycopeptides and oligosaccharides. The recently developed lectin arrays, that consist of tens of different carbohydrate-binding proteins attached to microtiter plates or cover glasses, greatly facilitate the analysis of soluble glycoproteins and cell surface glycans. **See also**: Agglutination Techniques for Detecting Antigen–Antibody Reactions

The cytotoxic properties exhibited by certain lectins (e.g. ricin) make them ideal reagents for selection of lectin-resistant clones of animal cell mutants with altered surface carbohydrates. Such lectin-resistant mutants provide an excellent tool for investigations of the genetics, biosynthesis and function of glycoconjugates, and also serve in the biotechnology industry for the production of therapeutics glycoproteins with tailor-made glycans.

A small number of lectins are in clinical use. Soybean agglutinin is utilized for the purging of human bone marrow for transplantation by selective agglutination of the mature T cells present in the marrow. These cells are responsible for the lethal graft-versus-host reaction that occurs when histo-incompatible (mismatched) marrow is used for transplantation. Donor marrow thus depleted of T cells is employed routinely for transplantations into children born with severe combined immune deficiency ('bubble children', since they are highly susceptible to microbial infections and have to be kept in a plastic bubble) with close to 70% success. The same method has been in experimental use during the 1980s and early 1990s for bone marrow transplantation of end-stage leukaemic patients, but has been replaced by other techniques for T-cell depletion, such as by monoclonal antibodies. **See also**: Bone Marrow; Monoclonal Antibodies; Transplantation

Another clinical application of lectins is in blood typing. Thus, the lectins from *Lotus tetragonolobus* and *Ulex europaeus*, both specific for fucose, serve for the identification of blood-type O cells and of secretors of blood group substances. The lectin from *Dolichos biflorus* is used to distinguish between A_1 and A_2 subgroups and that from *Vicia graminea* to identify blood-type N erythrocytes. In addition, peanut agglutinin, specific for Gal(β1–3)GalNAc, is employed in the detection of 'polyagglutination' (or 'polyagglutinability'), a condition accompanying certain bacterial and viral infections in which human erythrocytes become agglutinable by antibodies normally present in the sera of nearly all adults. If not diagnosed in time, it may lead to serious complications and death.

Several lectins (e.g. concanavalin A and the lectin from the red kidney bean, known as phytohaemagglutinin or PHA), are potent mitogens that activate lymphocytes and induce them to divide. Such lectins are polyclonal activators, in that they activate lymphocytes irrespective of their antigenic specificity. Before the advent of monoclonal antibodies to cell surface antigens, mitogenic lectins were the major tool for studies of the mechanism of the transfer of signals from the outer surface of the cell to its interior. Mitogenic stimulation by lectins provides a facile means to assess the immunocompetence of patients suffering from a diversity of diseases, including acquired immunodeficiency syndrome (AIDS), and to monitor the effects of various immunosuppressive and immunotherapeutic manipulations. It is also employed for the preparation of chromosome maps for different purposes, such as karyotyping, sex determination and detection of chromosome defects, since the chromosomes are easily visualized in the stimulated cells.

Further Reading

Blixt O, Head S, Mondala T *et al*. (2004) Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proceedings of the National Academy of Sciences of the USA* **101**: 17033–17038.

Dommett RM, Klein N and Turner MW (2006) Mannose-binding lectin in innate immunity: past, present and future. *Tissue Antigens* **68**: 193–209.

ENCYCLOPEDIA OF LIFE SCIENCES © 2009, John Wiley & Sons, Ltd. www.els.net

Kilpatrick DC (2002) Animal lectins: a historical introduction and overview. *Biochimica et Biophysica Acta* **1572**: 187–197.

Rosenfeld R, Bangio H, Gerwig GJ *et al.* (2007) A lectin array-based methodology for the analysis of protein glycosylation. *Journal of Biochemical and Biophysical Methods* **70**: 415–426.

Sharon N (2006) Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochimica et Biophysica Acta* **1760**: 527–537.

Sharon N and Lis H (2003) *Lectins*, 2nd edn, 454pp. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Sharon N and Lis H (2004) History of lectins: from hemagglutinins to biological recognition molecules. *Glycobiology* **14**: 53R–63R.

Taylor ME and Drickamer K (2007) Paradigms for glycan-binding receptors in cell adhesion. *Current Opinion in Cell Biology* **19**: 572–577.

Van Damme EJM, Peumans WJ, Pusztai A and Bardocz S (1998) *Handbook of Plant Lectins: Properties and Biomedical Applications*, 452pp. Chichester: Wiley.

Van Kooyk Y and Rabinovich GA (2008) Protein–glycan interactions in the control of innate and adaptive immune responses. *Nature Immunology* **9**: 593–601.

Wellens A, Garofalo C, Nguyen H *et al.* (2008) Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the fimH–oligomannose-3 complex. *PLoS* **3**: e2040.