Diversity in recognition of glycans by F-type lectins and galectins: molecular, structural, and biophysical aspects

Gerardo R. Vasta[1], **Hafiz Ahmed**[2], **Mario A. Bianchet**[3], **José A. Fernández-Robledo**[1], and **L. Mario Amzel**[3]

[1]Department of Microbiology and Immunology, University of Maryland School of Medicine, IMET, Baltimore, Maryland [2]Department of Biochemistry and Molecular Biology, University of Maryland School of Medicine, IMET, Baltimore, Maryland [3]Department of Biophysics & Biophysical Chemistry, Johns Hopkins University School of Medicine, Baltimore, Maryland

Abstract

Although lectins are "hard-wired" in the germline, the presence of tandemly arrayed carbohydrate recognition domains (CRDs), of chimeric structures displaying distinct CRDs, of polymorphic genes resulting in multiple isoforms, and in some cases, of a considerable recognition plasticity of their carbohydrate binding sites, significantly expand the lectin ligand-recognition spectrum and lectin functional diversification. Analysis of structural/functional aspects of galectins and F-lectins —the most recently identified lectin family characterized by a unique CRD sequence motif (a distinctive structural fold) and nominal specificity for L-Fuc—has led to a greater understanding of self/nonself recognition by proteins with tandemly arrayed CRDs. For lectins with a single CRD, however, recognition of self and nonself glycans can only be rationalized in terms of protein oligomerization and ligand clustering and presentation. Spatial and temporal changes in lectin expression, secretion, and local concentrations in extracellular microenvironments, as well as structural diversity and spatial display of their carbohydrate ligands on the host or microbial cell surface, are suggestive of a dynamic interplay of their recognition and effector functions in development and immunity.

Keywords

galectin; F-type lectin; C-type lectin; AAA; MsFBP32; fucolectin

Introduction

Except for a few members of the immunoglobulin superfamily, such as hemolin and other unique mosaic proteins containing immunoglobulin-like (immunoglobulin superfamily) domains,[1–5] no *bona fide* components of the mammalian adaptive immunity, such as immunoglobulins, B and T cells, and their receptors, are present in invertebrates or protochordates, taxa that thus rely solely on innate immunity for defense against microbial infection.[6] Furthermore, it has become now well established that in vertebrates, innate immunity not only carries a substantial burden of the defense functions against infectious diseases but it is also critical for the development of an effective adaptive immune response.

Therefore, great interest has been generated in the structure-function basis of diversity in recognition of its various components, particularly on lectins, Toll, and Toll-like receptors, and other "nonself" recognition receptors.[7–9]

Unlike immunoglobulins, most lectins do not generate diversity in recognition by genetic recombination and they are considered to be "hardwired" in the germline. Therefore, attention has focused on the potential germline-encoded diversity of the lectin repertoires, including tandem gene duplications, formation of chimeric structures by exon shuffling, their allelic variation, and additional mechanisms for expanding the ligand recognition spectrum by alternative splicing and somatic mutation. Further, substantial interest has arisen in the structural basis for the potential plasticity of the carbohydrate binding sites concerning the diversity in ligand topologies recognized by any given lectin.[10–12]

Based on the Medzhitov and Janeway (2002)[13] model for nonself recognition, pattern recognition receptors (PRRs) such as the *Drosophila* Toll and the mammalian Toll-like receptors recognize pathogens via highly conserved and widely distributed microbial surface molecules, such as lipopolysaccharide, flagellin, lipoteichoic acid, or peptidoglycan (a group of molecules known as *pathogen-associated molecular patterns* (PAMPs)), which are essential for the microbe but absent in the host. Given that nonpathogenic microbes also share these surface molecules it has been suggested that these may be more accurately described as *microb-associated molecular patterns* (MAMPs).[14] Endogenous factors such as nuclear or cytosolic components that are released during tissue stress or necrosis under sterile conditions can trigger inflammatory responses have been designated *danger-associated molecular patterns* (DAMPs).[15] More recently, the term *virulence-associated molecular pattern* (VAMP) has been introduced to describe those factors (e.g., microbial toxins, flagellin) that enable the host to discriminate pathogenic microbes from the nonpathogenic ones.[16] Among the best-characterized animal lectins, the mannose-binding lectin (MBL), a member of the C-type lectin family, has been described as the prototypical PRR.[17]

C-type lectins are characterized by their structural fold (C-type lectin domain fold, CTLD), their Ca^{2+} requirement for ligand binding, and, in most family members, the presence of multiple, unrelated structural domains in the polypeptide. They comprise the collectins (MBLs, ficolin, conglutinin, and pulmonary surfactant) (Fig. 1), proteoglycan core proteins, selectins, endocytic receptors, the mannose-macrophage receptor, and DC-SIGN.[18–20] Although some C-type lectins such as selectins and DC-SIGN bind self glycans, others, such as collectins, recognize exposed sugar ligands on microbial surfaces. Collectins are lectins with a collagenous region linked to a CRD that recognizes sugars on microbial surfaces, and upon binding to a serine protease (i.e., MBL-associated serine proteases, MASPs) may activate the complement cascade (Fig. 1).[21] Several lectins homologous to MBLs and ficolins, MASPs, and complement components have been identified in invertebrates and ectothermic vertebrates, suggesting that C-type lectins and the complement system played a pivotal role in innate immunity long before the emergence of adaptive immunity in vertebrates.[22,23]

The oligomerization of the MBL subunits results in binding multivalency and that enables the protein to recognize ligands are displayed 45 Å apart on the microbial surface, thereby increasing the MBL's avidity (Fig. 1). The density of surface ligands and their scaffolding (as glycoproteins or glycolipids) modulates the affinity of the interaction via negative cooperativity.[24] Thus, the binding of MBL to multiple nonreducing terminal carbohydrate ligands on the microbial surface, which are not readily exposed in the mammalian host, leads to agglutination and immobilization of the potential pathogen. Further, the interactions of other MBL domains with additional factors, such as MASP, trigger downstream effector

functions, including complement activation and opsonization or lysis of the agglutinated microbes.[25] Therefore, the MBL binds to nonself glycans and functions as a PRR endowed of both recognition and effector properties.

The F-lectin family: structural and functional diversity of tandemly arrayed CRDs

F-type lectins constitute the most recently identified lectin family, characterized by a unique CRD amino acid sequence motif and structural fold, and a nominal specificity for L-Fuc.[26–29] In this family, the F-type domain can be present either as a single CRD, as in the European eel agglutinin (*Anguilla anguilla* agglutinin; AAA) or as tandemly arranged F-type CRD repeats, in some examples combined with unrelated domains, yielding molecular species of variable sizes and function(s) within a single species.[26–29]

The structure of the AAA/L-Fuc complex provided not only a novel CRD sequence motif that identifies the F-lectin family members, but also a novel fold (F-type fold) for an animal lectin, which consists of a β-jelly roll sandwich composed of three- and five-stranded β-sheets.[28] Binding to αFuc is mediated by hydrogen bonds from a trio of basic side chains that emerge from a shallow pocket and by van der Waals contacts with an unusual disulfide bridge between contiguous cysteines (Cys[82] and Cys[83]). The lectin establishes interactions with the ring O5 and the equatorial 3- and axial 4-OH groups of the α-Fuc using the nitrogen atoms of three conserved residues: N- of His[52] and the guanidinium groups of Arg[79] and Arg[86] (Fig. 2). Van der Waals interactions are established between the disulfide bridge formed between the consecutive cysteine residues and the bonds of ring atoms C1 and C2 of the monosaccharide, to the C6, which docks loosely in a hydrophobic pocket, stacking against the aromatic rings of two residues His[27] and Phe[45]. These residues, together with the residues Leu[23] and Tyr[46], form the binding pocket. Although AAA preferentially binds α-Fuc, it also binds 3-*O*-methyl-D-galactose (also 3-*O*-methyl-D-fucose), a galactose derivative that has topologically equivalent hydroxyls (i.e., axial hydroxyl on C4) and a hydrophobic moiety similar to α-Fuc (Fig. 2). As for most animal lectins, the specificity of AAA for α-Fuc is not absolute, but rather nominal, as carbohydrates (e.g., 3-*O*-methyl-D-galactose and 3-*O*-methyl-D-fucose) that share critical configurational features of α-Fuc (i.e., axial hydroxyl and hydrophobic moiety) also behave as ligands for AAA. Unlike C-type lectins, the role of calcium appears to play in AAA is structural stabilization, rather than participating in direct cation-saccharide interactions.[28]

Recognition of H type 1 and Le[a] oligosaccharides by AAA requires additional interactions of residues in the loops (CDRs 1–5) that surround the binding pocket with subterminal sugars of the tyrisaccharides Fucα1–2 Galβ1–3GlcNAcβ1–3Galβ1–4Glc (blood group H type 1) and Galβ1–3[Fucα1–4]GlcNAcβ1–3Galβ1–4Glc (Le[a]). The equatorial hydroxyls (3-OH and 2-OH) in Gal and the oxygen of the GlcNAc 2-N-Acetyl group in Le[a], or GlcNAc 6-OH and 4-OH groups in H1, are recognized by Glu[26] and His[27] on CDR1. The OH group of Tyr[46] in CDR2 can coordinate the glycosidic bond oxygen between the Gal and GlcNAc. Asp[81] and Arg[79] in CDR4 interact with the 6-OH group of GlcNAc in Le[a], and a water molecule may bridge the Gal 4-OH group with Asp[81] in H1. The lack of binding of AAA for Le[x] is explained by the lack of flexibility of CDR1 that would clash with the 2-N-acetyl, which in Le[x] points toward the Fuc side of the trisaccharide. F-lectins with a shorter CDR1, such as MsaFBP32 (see below), would have a broader specificity for Le oligosaccharides.[29]

One key observation that concerns the potential diversity in carbohydrate recognition by F-type lectins is the presence of multiple isoforms with amino acid substitutions at positions revealed by the structural analysis as key for ligand binding. Variability of critical residues

in the binding pocket and surrounding loops in the multiple isoforms, as expressed in the Japanese eel,[30] suggests that alternative interactions with terminal and subterminal sugar units may expand the range of diverse oligosaccharides recognized by the lectin isoform repertoire.[28] This is an intriguing observation for a protein proposed to mediate the recognition of potential microbial pathogens. Like collectins, the native structure of the European eel (*Anguilla anguilla*) agglutinin is a homotrimer (Fig. 2), which suggests that ligand binding is enhanced through cooperative binding to multivalent glycans. Further, the threefold symmetry of the native AAA, very similar to that observed in collectins,[31] would optimize the spacing and orientation of binding sites for recognition of glycoconjugates displayed on microbial surfaces. The distances between binding sites in the AAA trimer (26 Å) and those in MBL (45 Å) suggest that they bind to differently arrayed surface glycans on the surface of microbes (Fig. 2). Therefore, although F- and C-type lectins may recognize the same monosaccharide (MBL also binds fucose), they may bind to different microorganisms, thereby expanding the immune recognition spectrum in those species that are endowed with both lectin types.

The F-lectin from the striped bass *Morone saxatilis* (MsaFBP32) consists of two tandemly arrayed F-type CRDs. The amino acid sequence of the N-terminal CRD (N-CRD) is more similar that of AAA than the C-terminal CRD (C-CRD). The structure of the complex of MsaFBP32 with L-fucose revealed a subunit that folds as a tail-to-tail arrangement of the two F-type CRDs (total length, 81 Å) in which the CRDs of the carbohydrate recognition sites are at opposite ends of the monomer.[29] The overall structure of the N-CRD is highly similar to that of the C-CRD, and the full range of polar and apolar interactions that mediate fucose recognition by AAA is observed in each CRD of MsaFBP32, although in both N-and C-CRDs of MsFBP32 the CRD1 is shorter than in AAA, making the pocket for the C6 more solvent accessible. Further, there are additional significant differences between the binding sites of the MsFBP32 N- and C-CRD domains. The structure of the N-CRD binding site resembles the AAA site more closely than the C-CRD site, and the C6 pocket in the C-CRD is less open than in N-CRD due to, first, the replacement of Phe[37] by the bulkier Trp[183], and second, by the replacement of the apolar contact of the S–S bridge with the face b of L-fucose observed in AAA and N-CRD by that with a bulkier Phe[220] that tilts the sugar out of the shallow binding pocket. Further, an examination of the solvent-accessible surfaces reveals significant differences in the topology of the N- and C-CRDs, specifically those features corresponding to the extended binding site that surround the primary recognition site where L-fucose binds. These differences strongly suggest that the N-CRD recognizes more complex fucosylated oligosaccharides, and with a relatively higher avidity than the C-CRD. For example, in the N-CRD, a methyl group of a second fucose may dock on top of Phe[37], but in the C-CRD, Trp[183] closes the pocket with its indole ring, thereby imposing steric constraints on the second fucose moiety of Lewis tetrasaccharides.[29]

Like AAA, the physiological state of MsaFBP32 is an approximately cylindrical trimer (81-Å-long and 60-Å-wide) divided into two globular halves: one containing the three N-CRDs and the other containing the three C-CRDs, with all six CRDs containing a bound Ca^{2+} that does not participate in ligand binding. The resulting binding surfaces at the opposite ends of the cylindrical trimer resemble the collectin-like "bouquet" CRD displays, and have the potential to cross-link cell surface or humoral carbohydrate ligands (Fig. 3). The back-to-back orientation of the binding surfaces of the trimeric MsaFBP32 supports the notion that the function of this lectin in circulation is to cross-link fucosylated glycoconjugates displayed on different cells, with an epitope separation of 25Å on the cell surface. Modeling of MsaFBP32 complexed with fucosylated glycans that are widely distributed in prokaryotes and eukaryotes rationalizes the observation that binary tandem CRD F-type lectins can function as opsonins. This would take place by cross-linking nonself carbohydrate ligands and self carbohydrate ligands, such as sugar structures displayed by microbial pathogens and

glycans on the surface of phagocytic cells from the host. For example, MsaFBP32 may cross-link Le^a-containing glycans on the phagocytic cell surface via the N-CRD with glycans on the microbial surface via the C-CRD, such as those containing α-linked L-Fuc, 2-acetoamido L-Fuc, 3-deoxy-L-fucose (colitose) or L-Rha (6-deoxy-L-mannose, present in *E. coli* glycans) as nonreducing terminal residues.[29] Pre-exposure of *E. coli* to a binary tandem F-lectin from sea bass significantly increases their phagocytosis by peritoneal macrophages relative to the unexposed bacteria,[32] confirming that F-lectins with tandemly arrayed CRDs such as MsaFBP32 function as opsonins that mediate innate immune responses against microbial pathogens.

It is noteworthy that the sperm bindin protein from the Japanese oyster (*Crassostrea gigas*) has been recently identified as an F-lectin.[33] Bindins are polymorphic gamete recognition proteins stored in the acrosomal rings that bind sperm to egg during fertilization. Although encoded by a single copy gene, positive selection, recombination, and alternative splicing mechanisms produce highly diversified transcripts (possibly thousands of variants) both in sequence and domain organization within and among individuals among this species. The males however, translate only one or two polymorphic molecular species housing between one and five tandemly arrayed F-lectin domains.[34] The unusual high intraspecific diversity in sequence and domain organization of the oyster bindin F-lectins could represent co-evolution of sperm gamete recognition mechanisms to "catch-up" with the high diversification of egg receptors aimed at avoiding polyspermia.[35] This diversification of the bindin F-lectins in gamete recognition resembles the eel F-lectin isoforms involved in host-pathogen interactions described above.[28]

Although the F-type lectin fold is unique among animal lectins, a structure-based search identified several "unrelated" proteins that showed limited sequence similarity to AAA, yet shared the same jellyroll fold.[36] These include the C-terminal domains (FA58C) of human blood coagulation factor V and VIII, the C-terminal domain of a bacterial sialidase, the NH$_2$-terminal domain of a fungal galactose oxidase, a subunit of the human APC10/DOC1 ubiquitin ligase, the N-terminal domain of the XRCC1 single-strand DNA repair complex, and a yeast allantoicase (PDB 1SG3).[28] In NGOase and CSIase, the hollow conserves two members of the triad of basic residues that hold the axial hydroxyl of the ligand in AAA: the His and one Arg residue. The third partner, the second Arg residue is replaced by a Glu in CSIase and by an Asn residue in a distant position in NGOase. CSIase, the galactose-binding domain of the bacterial sialidase, binds carbohydrate,[37] and it is likely that NGOase also binds its substrate galactose. These residues are absent in human FVa-C2, which is reported to bind phospholipids.[38] Thus, the higher vertebrate F-lectin representatives may have evolved from F-type CRDs that, while conserving the placement of their binding-site among the loops at the opening of the β-barrel, also developed diverse specificities for ligands other than saccharides during evolutionary co-option to fulfill divergent biological roles. The FA58C domain, tandemly arrayed in the C-terminal region of the mammalian coagulation factors V and VIII, is similar to the slime mold discoidin (DS) domain.[37,39] However, unlike the *bona fide* tandemly arrayed F-lectin CRDs, it displays multiple hydrophobic spikes that enable binding to cell surface phospholipids instead of saccharides.[37,38] In contrast, the b1 domain of neuropilin-1 receptor,[39] also a DS homologue, lacks these hydrophobic residues and displays a polar binding cleft. Similarly, the so-called discoidin domain receptors[40] are activated by binding to collagen.[41] The structural analogs of intracellular localization such as APC10/DOC1 and XRCC1 do not appear to bind glycoconjugates either.[42]

Galectins as PRRs: recognition of self and nonself glycans

Galectins are a lectin family of extensive taxonomic distribution and striking evolutionary conservation.[43] They are characterized by a conserved CRD sequence motif and affinity for β-galactosides, with some family members showing preference for *N*-acetyllactosamine (LacNAc; Galβ1,4GlcNAc) and related disaccharides, with dissociation constants in the order of 10^{-5} M.[44,45] Most galectins are nonglycosylated soluble proteins, although a few exceptions have transmembrane domains.[46,47] In spite of their lack of a typical secretion signal peptide, galectins are present not only in the cytosol and the nucleus, but also in the extracellular space.[48] Galectins are secreted by nonclassical mechanisms, possibly by direct translocation across the plasma membrane.[49,50] In the extracellular space, galectins can bind to cell surface glycans, to the extracellular matrix, and to potential parasites and pathogens.[51–58] It has been proposed that their release to the extracellular space under noninfectious stressful conditions is perceived as a DAMP signal that trigger inflammatory responses.[59]

Based on the CRD organization of the polypeptide monomer, galectins have been classified in three types: proto, chimera, and tandem-repeat (TR).[60] Proto-type galectins, such as galectin-1, contain one CRD per subunit and are noncovalently linked homodimers. The chimera-type galectins (galectin 3) have a C-terminal CRD and an N-terminal domain rich in proline and glycine. The dimerization of proto-type galectins is critical for their function in mediating cell–cell or cell–ECM interactions,[61,62] and similar interactions via the N-terminus domain mediate oligomerization (mostly trimers and pentamers) of the chimera galectins.[63] In TR galectins, such as galectin-9, two CRDs are joined by a functional linker peptide. Galectins with four tandemly arrayed CRDs have been described in invertebrates.[64,65] The primary structures and gene organization of mammalian galectins are substantially conserved. Prior to or during early in chordate evolution, duplication of a mono-CRD galectin gene would have led to a bi-CRD galectin gene, in which the N- and C-terminal CRDs subsequently diverged into two different subtypes, defined by exon–intron structure (F4-CRD and F3-CRD). All vertebrate single-CRD galectins belong to either the F3- (e.g., gal-1, -2, -3, -5) or F4- (e.g., gal-7, -10, -13, -14) subtype, whereas TR galectins such as gal-4, -6, -8, -9, and -12 contain both F4 and F3 subtypes.[66] However, phylogenetic analysis of the galectin from the eastern oyster *Crassostrea virginica* (CvGal), that displays four tandemly arrayed CRDs, revealed that these are closely related to the single CRD galectins, suggesting that the CvGal gene is the product of two consecutive gene duplications of a single-CRD galectin gene.[66]

The crystal structure of the dimeric galectin-1 complexed with LacNAc determined at 1.9 Å resolution revealed the galectin structural fold, and allowed the identification of the amino acid residues and the hydroxyl groups of the ligands that participate in protein–carbohydrate interactions (Fig. 4).[67–69] Each subunit is composed of an 11-strand antiparallel β-sandwich and contains one CRD. The carbohydrate binding site is formed by three continuous concave strands (β4–β6) containing all residues involved in direct interactions with LacNac, which include His[44], Asn[46], Arg[48], His[52], Asp[61], Trp[68], Glu[71], and Arg[73] (Ref. 67). Additional interactions involving a water molecule that bridges the nitrogen of the NAc group with His[52], Asp[54], and Arg[73] explains the higher affinity of LacNAc over Lac. The crystal structures of the complexes of *Bufo arenarum* galectin-1 with LacNAc and the TDG (thio-digalactoside) illustrate how this galectin uses similar interactions to recognize the natural ligand (LacNAc) and a synthetic disaccharide (TDG) (Fig. 5A and B).[68] Interestingly, the displacement of the loop containing histidine 52 opens the binding cleft to accommodate the sulfur atom in TDG, illustrating the plasticity of the galectin binding site (Fig. 5C). Unlike galectin-1, galectin-3 has an extended carbohydrate-binding site formed by a cleft open at both ends, in which the LacNAc is positioned in such a way that the reducing

end of the LacNAc (GlcNAc) is open to solvent, but the nonreducing moiety (Gal) is in close proximity to residues in the β3 strand.[70] The extended binding site leads to increased affinity for glycans with multiple lactosamine units, and with their substitution of the nonreducing terminal galactose moiety with ABH blood group oligosaccharides [Fucα1, 2; GalNAcα1,3(Fucα1,2); and Galα1,3(Fucα1,2)]. As observed for C- and F-type lectins, metazoans are also endowed with a complex galectin repertoire constituted by multiple galectin types, subtypes, and isoforms, that suggests functional diversification. Proto- and TR-type galectins comprise several distinct subtypes: galectins-1, -2, -5, -7, -10, -11, -13, -14, and -15 are proto-type. Galectin-3 is the only chimera-type; galectins-4, -6, -8, -9, and -12 are TR-type. As observed in other lectin families, a galectin subtype such as galectin-1 may exhibit multiple isoforms in a single individual.[71] Further, because galectin types and subtypes exhibit notable differences in carbohydrate specificity and bind a broad range of glycans that display the requisite topologies, the galectin repertoire has substantial diversity in recognition properties.[72–77] Since their initial description in the early 1980s, expression of galectins was understood as developmentally regulated, and their functions related to embryogenesis and early development and cancer progression and metastasis (reviewed in Ref. 78). Since the 1990s the roles of galectins in regulation of both innate and adaptive immune responses have been gradually identified and characterized in detail.[79] In general, glycans that display N-acetyllactosamine and polylactosamine chains [(Galβ1,4GlcNAc)n], such as laminin, fibronectin, and mucins are the preferred endogenous ligands for galectins.[72–76] More specifically, endogenous glycans recognized by galectins include β integrin, CD45, GM1, CD44, Tim3, MUC1, podoplanin, CD166, ABH-type oligosaccharides CD43, CD45, CD7, CD71, CD44, TIM3, CTLA4, MUC1, MUC16, and MerTK.[53,80–86] Although it has been proposed that a certain degree of functional redundancy exists among the members of the galectin repertoire, as the subtle aspects of their binding properties and natural ligands are characterized and their biological roles are elucidated in increasing detail, it has become clear that this is not the case.[53]

In recent years it has become clear that galectins can also recognize nonself carbohydrate moieties on the surface of microbial pathogens and parasites, and function as PRRs.[51,54] These ligands on foreign cells can be similar to those displayed on host cells such as ABH or Le blood group oligosaccharides[55] and LacNAc present in viral and bacterial glycans, or structurally different and absent from the host glycome, such as α1–2-mannans in *Candida*[56] and LacdiNAc in *Schistosoma*.[87] While the first scenario can be rationalized as molecular mimicry by the microbial pathogens and parasites, understanding the molecular basis of galectin binding to distinct self and nonself glycans via the same CRD requires additional considerations. In this regard, galectins with tandemly arrayed CRDs such as the TR galectins from vertebrates, and the 4-CRD galectins from invertebrates are intriguing both in their binding properties and functional aspects. Vertebrate TR galectins such as galectin-4, -8, and -9 differ from the proto- and chimera types in that they display two tandemly arrayed CRDs, and like the binary CRD F-type lectins, the N- and C-CRDs of TR galectins are similar but not identical, suggesting that they have distinct recognition properties.[88] The structures of TR galectin-4, -8, and -9 have been partially determined either by X-ray diffraction or NMR analysis of their isolated N- or C-CRDs and revealed differences in their binding specificity and or affinity for oligosaccharides or their scaffolding as glycolipids or glycoproteins.[89–91] The structure of the N-CRD of the mouse galectin-4 revealed binding sites for lactose with different affinities, while galectin-8 binds preferentially to larger glycans such as glycosphingolipids.[89–91] The capacity of TR galectins to cross-link cells with different synthetic glycoconjugates[91,92] strongly suggests significant differences in the binding properties of their N- and C-CRDs. From the functional standpoint, the most striking example is the recognition and killing of *E. coli* O86 that display B-blood group oligosaccharides (BGB+ *E. coli*) by the TR galectin-4 and -8. Mutation of key residues in either CRD revealed that the C-CRD mediates recognition of the BGB+ *E. coli* but does not

affect its viability, while the N-CRD was not affected, suggesting that N-CRD might be endowed with killing activity. This was confirmed by examining the binding and killing activity of the separate CRDs. This observation suggests that in galectin-4 and -8 the N- and C-CRDs not only have different recognition properties, but also they are functionally different in their interactions with the BGB⁺ *E. coli*.[93]

A different scenario is presented by the recognition of glycans on the phagocyte (self) and parasite (nonself) surface by the oyster galectin CvGal.[64] Although the four CRDs of CvGal are not identical, homology modeling suggests that the minor differences in sequence among the CRDs do not result in significant differences in binding properties. Therefore, the cross-linking of similar glycans on the host phagocytic cells and on the surface of the parasite trophozoite by CRDs of similar binding properties can only be rationalized by cross-linking of similar self and nonself glycans by the folding of the 4-CRD polypeptide in a particular geometry, such as a tetrahedral architecture, where the CRDs are oriented in opposite directions. A similar model could be proposed for single-CRD prototype and chimera galectins for which the requisite multivalency and the spatial arrangement and orientation of the CRDs is achieved by oligomerization of single-CRD subunits. However, questions about the molecular and structural basis for recognition and binding to either similar or distinct self or nonself glycans remain open, and several aspects concerning the protein, the carbohydrate ligands, the thermodynamics of lectin–ligand interaction, and the particular microenvironment in which this interaction takes place merit further discussion. With regard to the protein, recognition specificity and affinity are key aspects of the lectin–ligand interaction(s) that modulate downstream effector functions.

Lectin oligomerization, ligand density/scaffolding, and the microenvironmental factors that modulate glycan recognition

In addition to the structural aspects of the CRDs and the landscape of the binding surfaces of the primary and extended binding sites, as described above for binary F-type lectins,[27–29] even when considering monovalent recognition, additional factors contribute to the energetics of the interaction, which are further complicated when addressing multivalent recognition.[92] Although multivalency may be a property of tandemly arrayed CRDs in some lectins, as discussed above for C-, F-lectins and galectins, oligomerization of the monomers results in displays of clustered binding sites that either in their geometry or dimensions can be unique to a particular lectin. For C-type lectins the bouquet-like display of binding sites in the trimer results in multivalency for sugars spaced about 45 Å apart, and increases avidity of the binding.[94] The tissue localization and subcellular compartmentalization of lectins can influence their oligomerization and, in turn, the selective binding of lectins to particular ligands. For galectins, oligomerization, and clustering of CRDs enables binding to and cross-linking of multivalent glycoproteins and glycolipids on the cell surface leading to formation of microdomains and lattices that signal vial specific pathways depending on the glycans and galectins involved.[95–97]

For F-lectins, the identification of the trimer as the physiological state is rationalized by the large accessible surface area buried upon trimer formation, and the Cl⁻ coordinated with positively charged residues (Lys or Arg) in the three monomers.[27–29] In the F-lectin MsaFBP32, the opposite orientation of the binding N- and C-CRD surfaces in the rather rigid trimer strongly suggests that when present in the extracellular space it is able to cross-link cells by recognizing fucosylated glycoconjugates displayed with a separation of 25 Å on the cell surface. Under the low Cl⁻ concentrations in the intracellular environment, the more flexible monomeric form of MsaFBP32 is favored, in which the N- and CCRDs may point in variable directions. When the monomers are secreted to the extracellular space rich in Cl⁻, the trimer is assembled and cross-links cell surface and/or soluble multivalent

glycans.[29] Alternatively, individual F-lectins bound to glycans of the pathogen's membrane may display a modified (2-dimensional) affinity for each other that facilitates its oligomerization.[98] In addition to ionic strength or the presence of particular electrolytes, the redox properties of the intracellular and extra-cellular environments can modulate the activity of lectin–ligand interactions, and the biological outcome. Most galectins are susceptible to inactivation by oxidation of their free Cys residues present in the CRD that require being in the reduced state for binding activity.[93] In the reducing intracellular environment, galectins remain stable, but upon secretion to the mostly oxidative extracellular space, the oxidation of the free cysteines compromises not only the binding activity but also the oligomerization of the galectin subunits.[93] It is noteworthy that binding of galectin-9 to the protein disulfide isomerase at the cell surface increases retention of the enzyme that in turn modulates the redox status at the plasma membrane.[97] In addition to the lectin quaternary structure, as established via structural or hydrodynamic approaches, higher orders of aggregation may occur in environments where the lectin concentration reaches above a certain threshold. Sedimentation velocity and sedimentation equilibrium studies revealed that the galactosyl-binding lectins DCL-I and DCL-II from the protochordate *Didemnum candidum* behave as homotetramers of 14.5 and 15.5 kDa subunits, respectively. However, even at concentrations below 1 mg/mL a positive dependence of the sedimentation coefficient with protein concentration was observed, suggesting that at increasing concentrations the protein associates at orders above the tetramer.[99] For the dimeric galectin-1, further association to form tetramers is temperature-driven.[100] In addition, the ligand density at the cell surface can drive oligomerization of the lectins subunits. In solution, galectin-3 monomers are in equilibrium with higher order oligomers, and when binding to multivalent glycoproteins or cell surface glycans they may recruit additional monomers to form a complex of multivalent interactions via galectin-3 trimers and pentamers.[101] The density of the carbohydrate ligands and their protein or lipid scaffolding context on the cell surface are key factors that determine lectin recognition and affinity. As proposed by Dam and Brewer,[100] in contrast to the well-defined dissociation constants that describe the binding of lectin monomers to monovalent glycans, a wide range of relative dissociation constants is needed to describe lectin binding to multivalent surface glycans, that depends on the density and number of glycans on a surface and increased negative cooperativity. The authors propose that this relative affinity should replace the term avidity.[102] Galectin binding to ligand is enthalpically driven, exhibits enthalpy-entropy compensation, and follows a van't Hoff dependence of the binding constant on temperature, properties that are shared by other lectins. Further, the solvation properties of the CRD, the reorganization or displacement of the water shell, and the establishment of defined water-mediated interactions between the protein and the carbohydrate are factors that affect the thermodynamic aspects of the recognition process and stability of the lectin–ligand interaction.[100,103–105] As discussed above, many functional properties of lectins, such as the triggering of signaling cascades, result from clustered CRDs recognizing multivalent carbohydrate moieties in different scaffolds exposed at the cell surface, followed by the formation of microdomains and lattices. Therefore, a detailed knowledge of the thermodynamic aspects of lectin-mediated cross-linking of complex glycans is critical for understanding of lectin-mediated selective recognition of self and nonself ligands. In this regard, recent studies have shown that high-affinity lectin–mucin interactions are driven by favorable binding entropy of binding associated with a "bind and jump" mechanism. This consists of a dynamic binding process in which the bound lectins jump from carbohydrate to carbohydrate moiety in the multivalent glycans, which, by enhancing entropic effects, facilitates binding and subsequent complex formation.[100] Taken together, the factors discussed above suggest that in the natural context, lectin–ligand binding cannot be interpreted as a simple "lock-and-key" interaction. Instead, multiple factors that concern the protein binding site and lectin oligomerization, the display geometry and density of the recognized carbohydrate moieties in a particular protein or lipid scaffold, and the particular

properties of the environment where the lectin–sugar interaction takes place, may determine not only the binding dynamics and affinity, but also the lectin's preference for any one of the (self or nonself) ligands available.

Conclusions

In contrast with immunoglobulins that generate extensive diversity in recognition by genetic recombination and somatic mutation, a certain degree of somatic variability in the binding properties of lectins is achieved by the presence of multiple isoforms as observed for the eel F-lectins. In F-lectins that display multiple CRDs, differences in their binding properties would endow these proteins with the capacity for binding to a range of ligands that can be recognized on either host or microbial cell surfaces. In these and other lectin families, particularly in C-type family members, and in some F-type lectins, such as the *Drosophila* "furrowed" receptor, the combination of one or more lectin CRDs with effector domains results in proteins with functional multiplicity. Further, during the evolution of the F-lectin family considerable structural and functional divergence has occurred, ranging from extensive diversification in the gamete recognition proteins (bindins) of bivalves, to the binding affinity for phospholipids observed in the FA58C domain, tandemly arrayed in the C-terminal region of the mammalian coagulation factors V and VIII.

By recognizing highly conserved and widely distributed microbial surface glycans (PAMPs or MAMPs), which are essential for the microbe but absent in the host, lectins like the MBL, a member of the C-type lectin family, fit accurately the definition of PRR. Other lectins, such as F-type lectins and the TR galectins, however, can recognize both self and nonself glycans. Because these lectins display tandemly arrayed CRDs of similar but distinct specificity in a single polypeptide monomer, the binding and cross-linking of endogenous and exogenous glycans can be rationalized by the distinct properties of their binding sites. For other lectins, such as the proto- and chimera-type galectins that display a single CRD per monomer, their capacity to recognize both endogenous and exogenous glycans through the same binding site can be explained by taking into consideration multiple factors pertaining to the local lectin concentrations and oligomerization, the geometry of the presentation of the multivalent carbohydrate ligands on the host or microbial cell surface, and the properties of the microenvironment in which interactions take place.

Acknowledgments

References

1. Schmidt O, et al. Role of adhesion in arthropod immune recognition. Annu. Rev. Entomol. 2010; 55:485–504. [PubMed: 19743913]

2. Zhang SM, et al. Diversification of Ig superfamily genes in an invertebrate. Science. 2004; 305:251–254. [PubMed: 15247481]

3. Dishaw LJ, et al. A role for variable region-containing chitin-binding proteins (VCBPs) in host gut-bacteria interactions. Proc. Natl. Acad. Sci. USA. 2011; 108:16747–16752. [PubMed: 21930927]

4. Hernández Prada JA, et al. Ancient evolutionary origin of diversified variable regions demonstrated by crystal structures of an immune-type receptor in amphioxus. Nat. Immunol. 2006; 7:875–882. [PubMed: 16799561]

5. Watson FL, et al. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science. 2005; 309:1874–1878. [PubMed: 16109846]

6. Akira S, Takeda K, Kaisho T. Toll-like receptors: critical proteins linking innate and acquired immunity. Nat. Immunol. 2001; 2:675–680. [PubMed: 11477402]

7. Liu FT, Rabinovich GA. Galectins as modulators of tumour progression. Nat. Rev. Cancer. 2005; 5:29–41. [PubMed: 15630413]

8. Ludwig IS, Geijtenbeek TB, van Kooyk Y. Two way communication between neutrophils and dendritic cells. Curr. Opin. Pharmacol. 2006; 6:408–413. [PubMed: 16750420]

9. Fujita T, Matsushita M, Endo Y. The lectin-complement pathway—its role in innate immunity and evolution. Immunol. Rev. 2004; 198:185–202. [PubMed: 15199963]

10. Khalturin K, et al. Recognition strategies in the innate immune system of ancestral chordates. Mol. Immunol. 2004; 41:1077–1087. [PubMed: 15476919]

11. Iliev DB, et al. Endotoxin recognition: in fish or not in fish? FEBS Lett. 2005; 579:6519–6528. [PubMed: 16297386]

12. Vasta GR, et al. Galectins in teleost fish: zebrafish (Danio rerio) as a model species to address their biological roles in development and innate immunity. Glycoconj. J. 2004; 21:503–521. [PubMed: 15750792]

13. Medzhitov R, Janeway CA Jr. Decoding the patterns of self and nonself by the innate immune system. Science. 2002; 296:298–300. [PubMed: 11951031]

14. Bittel P, Robatzek S. Microbe-associated molecular patterns (MAMPs) probe plant immunity. Curr. Opin. Plant Biol. 2007; 10:335–341. [PubMed: 17652011]

15. Seong S-Y, Matzinger P. Hydrophobicity: an ancient damage-associated molecular pattern that initiates innate immune responses. Nat. Rev. Immunol. 2004; 4:469–478. [PubMed: 15173835]

16. Miao EA, Warren SE. Innate immune detection of bacterial virulence factors via the NLRC4 inflammasome. J. Clin. Immunol. 2010; 30:502–506. [PubMed: 20349122]

17. Garred P, et al. Mannose-binding lectin and its genetic variants. Genes Immun. 2006; 7:85–94. [PubMed: 16395391]

18. Ip WK, et al. Mannose-binding lectin and innate immunity. Immunol. Rev. 2009; 230:9–21. [PubMed: 19594626]

19. Zelensky AN, Gready JE. The C-type lectin-like domain superfamily. FEBS J. 2005; 272:6179–6217. [PubMed: 16336259]

20. Kingeter LM, Lin X. C-typelectinreceptor-induced NF-kappaB activation in innate immune and inflammatory responses. Cell Mol. Immunol. 2012; 9:105–112. [PubMed: 22246129]

21. Weis WI, Taylor ME, Drickamer K. The C-type lectin superfamily in the immune system. Immunol. Rev. 1998; 163:19–34. [PubMed: 9700499]

22. Wallis R. Structural and functional aspects of complement activation by mannose-binding protein. Immunobiology. 2002; 205:433–445. [PubMed: 12396005]

23. Nonaka M. The complement C3 protein family in invertebrates. ISJ. 2011; 8:21–32.

24. Dam TK, Brewer CF. Effects of clustered epitopes in multivalent ligand-receptor interactions. Biochemistry. 2008; 47:8470–8476. [PubMed: 18652478]

25. Matsushita M, et al. Proteolytic activities of two types of mannose-binding lectin-associated serine protease. J. Immunol. 2000; 165:2637–2642. [PubMed: 10946292]

26. Odom EW, Vasta GR. Characterization of a binary tandem domain F-type lectin from striped bass (*Morone saxatilis*). J. Biol. Chem. 2006; 281:1698–1713. [PubMed: 16251191]

27. Vasta GR, Ahmed H, Odom EW. Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. Curr. Opin. Struct. Biol. 2004; 14:617–630. [PubMed: 15465324]

28. Bianchet MA, et al. A novel fucose recognition fold involved in innate immunity. Nat. Struct. Biol. 2002; 9:628–634. [PubMed: 12091873]

29. Bianchet MA, et al. Structure and specificity of a binary tandem domain F-lectin from striped bass (*Morone saxatilis*). J. Mol. Biol. 2010; 401:239–252. [PubMed: 20561530]

30. Honda S, et al. Multiplicity, structures, and endocrine and exocrine natures of eel fucose-binding lectins. J. Biol. Chem. 2000; 275:33151–33157. [PubMed: 10924498]

31. Weis WI, Drickamer K. Trimeric structure of a C-type mannose-binding protein. Structure. 1994; 2:1227–1240. [PubMed: 7704532]

32. Cammarata M, et al. A serum fucolectin isolated and characterized from sea bass *Dicentrarchus labrax*. Biochim. Biophys. Acta. 2001; 1528:196–202. [PubMed: 11687307]

33. Moy GW, et al. Extraordinary intraspecific diversity in oyster sperm bindin. Proc. Natl. Acad. Sci. USA. 2008; 105:1993–1998. [PubMed: 18268333]

34. Moy GW, Vacquier VD. Bindin genes of the Pacific oyster *Crassostrea gigas*. Gene. 2008; 423:215–220. [PubMed: 18675329]

35. Springer SA, et al. Oyster sperm bindin is a combinatorial fucose lectin with remarkable intra-species diversity. Int. J. Dev. Biol. 2008; 52:759–768. [PubMed: 18649288]

36. Holm L, Sander C. Protein structure comparison by alignment of distance matrices. J. Mol. Biol. 1993; 233:123–138. [PubMed: 8377180]

37. Gaskell A, Crennell S, Taylor G. The three domains of a bacterial sialidase: a beta-propeller, an immunoglobulin module and a galactose-binding jelly-roll. Structure. 1995; 3:1197–1205. [PubMed: 8591030]

38. Macedo-Ribeiro S, et al. Crystal structures of the membrane-binding C2 domain of human coagulation factor V. Nature. 1999; 402:434–439. [PubMed: 10586886]

39. Ito N, et al. Novel thioether bond revealed by a 1.7 A crystal structure of galactose oxidase. Nature. 1991; 350:87–90. [PubMed: 2002850]

40. Firbank SJ, et al. Crystal structure of the precursor of galactose oxidase: an unusual self-processing enzyme. Proc. Natl. Acad. Sci. USA. 2001; 98:12932–12937. [PubMed: 11698678]

41. Vogel W, et al. The discoidin domain receptor tyrosine kinases are activated by collagen. Mol. Cell. 1997; 1:13–23. [PubMed: 9659899]

42. Marintchev A, et al. Solution structure of the single-strand break repair protein XRCC1 N-terminal domain. Nat. Struct. Biol. 1999; 6:884–893. [PubMed: 10467102]

43. Cooper DN. Galectinomics: finding themes in complexity. Biochim. Biophys. Acta. 2002; 1572:209–231. [PubMed: 12223271]

44. Vasta GR, et al. C-type lectins and galectins mediate innate and adaptive immune functions: their roles in the complement activation pathway. Dev. Comp. Immunol. 1999; 23:401–420. [PubMed: 10426431]

45. Schwarz FP, et al. Thermodynamics of bovine spleen galectin-1 binding to disaccharides: correlation with structure and its effect on oligomerization at the denaturation temperature. Biochemistry. 1998; 37:5867–5877. [PubMed: 9558320]

46. Lipkowitz MS, et al. Galectin 9 is the sugar-regulated urate transporter/channel UAT. Glycoconj. J. 2004; 19:491–498. [PubMed: 14758072]

47. Gorski JP, et al. New alternatively spliced form of galectin-3, a member of the beta-galactoside-binding animal lectin family, contains a predicted transmembrane-spanning domain and a leucine zipper motif. J. Biol. Chem. 2002; 277:18840–18848. [PubMed: 11886849]

48. Cho M, Cummings RD. Galectin-1, a beta-galactoside-binding lectin in Chinese hamster ovary cells. II: localization and biosynthesis. J. Biol. Chem. 1995; 270:5207–5212. [PubMed: 7890631]

49. Cleves AE, et al. A new pathway for protein export in *Saccharomyces cerevisiae*. J. Cell Biol. 1996; 133:1017–1026. [PubMed: 8655575]

50. Tsay YG, et al. Export of galectin-3 from nuclei of digitonin-permeabilized mouse 3T3 fibroblasts. Exp. Cell Res. 1999; 252:250–261. [PubMed: 10527616]

51. Vasta GR. Roles of galectins in infection. Nat. Rev. Microbiol. 2009; 7:424–438. [PubMed: 19444247]

52. Elola MT, et al. Galectins: matricellular glycan-binding proteins linking cell adhesion, migration, and survival. Cell Mol. Life Sci. 2007; 64:1679–1700. [PubMed: 17497244]

53. Rabinovich GA, Toscano MA. Turning `sweet' on immunity: galectin-glycan interactions in immune tolerance and inflammation. Nat. Rev. Immunol. 2009; 9:338–352. [PubMed: 19365409]

54. Cerliani JP, et al. Expanding the universe of cytokines and pattern recognition receptors: galectins and glycans in innate immunity. J. Clin. Immunol. 2011; 31:10–21. [PubMed: 21184154]

55. Stowell SR, et al. Innate immune lectins kill bacteria expressing blood group antigen. Nat Med. 2010; 16:295–301. [PubMed: 20154696]

56. Kohatsu L, et al. Galectin-3 induces death of *Candida* species expressing specific beta-1,2-linked mannans. J. Immunol. 2006; 177:4718–4726. [PubMed: 16982911]

57. Pelletier I, et al. Specific recognition of *Leishmania major* poly-beta-galactosyl epitopes by galectin-9: possible implication of galectin-9 in interaction between L. major and host cells. J. Biol. Chem. 2003; 278:22223–22230. [PubMed: 12684513]

58. Mercier S, et al. Galectin-1 promotes HIV-1 infectivity in macrophages through stabilization of viral adsorption. Virology. 2008; 371:121–129. [PubMed: 18028978]

59. Sato S, et al. Galectins in innate immunity: dual functions of host soluble beta-galactoside-binding lectins as damage-associated molecular patterns (DAMPs) and as receptors for pathogen-associated molecular patterns (PAMPs). Immunol. Rev. 2009; 230:172–187. [PubMed: 19594636]

60. Hirabayashi J, Kasai K. The family of metazoan metal-independent beta-galactoside-binding lectins: structure, function and molecular evolution. Glycobiology. 1993; 3:297–304. [PubMed: 8400545]

61. Gabius HJ. Animal lectins. Eur. J. Biochem. 1997; 243:543–576. [PubMed: 9057819]

62. Colnot C, et al. Galectins in mouse embryogenesis. Biochem. Soc. Trans. 1996; 24:141–146. [PubMed: 8674632]

63. Rabinovich GA, Rubinstein N, Toscano MA. Role of galectins in inflammatory and immunomodulatory processes. Biochim. Biophys. Acta. 2002; 1572:274–284. [PubMed: 12223275]

64. Tasumi S, Vasta GR. A galectin of unique domain organization from hemocytes of the Eastern oyster (*Crassostrea virginica*) is a receptor for the protistan parasite Perkinsus marinus. J. Immunol. 2007; 179:3086–3098. [PubMed: 17709523]

65. Yoshino TP, et al. Molecular and functional characterization of a tandem-repeat galectin from the freshwater snail *Biomphalaria glabrata*, intermediate host of the human blood fluke Schistosoma mansoni. Gene. 2008; 411:46–58. [PubMed: 18280060]

66. Houzelstein D, et al. Phylogenetic analysis of the vertebrate galectin family. Mol. Biol. Evol. 2004; 21:1177–1187. [PubMed: 14963092]

67. Liao DI, et al. Structure of S-lectin, a developmentally regulated vertebrate beta-galactoside-binding protein. Proc. Natl. Acad. Sci. USA. 1994; 91:1428–1432. [PubMed: 8108426]

68. Bianchet MA, et al. Soluble beta-galactosyl-binding lectin (galectin) from toad ovary: crystallographic studies of two protein-sugar complexes. Proteins. 2000; 40:378–388. [PubMed: 10861929]

69. Lobsanov YD, et al. X-ray crystal structure of the human dimeric S-Lac lectin, L-14-II, in complex with lactose at 2.9-A resolution. J. Biol. Chem. 1993; 268:27034–27038. [PubMed: 8262940]

70. Seetharaman J, et al. X-ray crystal structure of the human galectin-3 carbohydrate recognition domain at 2.1-A resolution. J. Biol. Chem. 1998; 273:13047–13052. [PubMed: 9582341]

71. Ahmed H, et al. Galectin-1 from bovine spleen: biochemical characterization, carbohydrate specificity and tissue-specific isoform profiles. J. Biochem. 1996; 120:1007–1019. [PubMed: 8982870]

72. Sparrow CP, Leffler H, Barondes SH. Multiple soluble beta-galactoside-binding lectins from human lung. J. Biol. Chem. 1987; 262:7383–7390. [PubMed: 2438278]

73. Sato S, Hughes RC. Binding specificity of a baby hamster kidney lectin for H type I and II chains, polylactosamine glycans, and appropriately glycosylated forms of laminin and fibronectin. J. Biol. Chem. 1992; 267:6983–6990. [PubMed: 1313027]

74. Ahmed H, et al. Novel carbohydrate specificity of the 16-kDa galectin from *Caenorhabditis elegans*: binding to blood group precursor oligosaccharides (type 1, type 2, Talpha, and Tbeta) and gangliosides. Glycobiology. 2002; 12:451–461. [PubMed: 12145186]

75. Shoji H, et al. Characterization of the *Xenopus* galectin family. Three structurally different types as in mammals and regulated expression during embryogenesis. J. Biol. Chem. 2003; 278:12285–12293. [PubMed: 12538594]

76. Zhou Q, Cummings RD. The S-type lectin from calf heart tissue binds selectively to the carbohydrate chains of laminin. Arch. Biochem. Biophys. 1990; 281:27–35. [PubMed: 1696449]

77. Fang R, Mantle M, Ceri H. Characterization of quail intestinal mucin as a ligand for endogenous quail lectin. Biochem. J. 1993; 293(Pt 3):867–872. [PubMed: 8352754]

78. Barondes SH, et al. Galectins. Structure and function of a large family of animal lectins. J. Biol. Chem. 1994; 269:20807–20810. [PubMed: 8063692]

79. Di Lella S, et al. When galectins recognize glycans: from biochemistry to physiology and back again. Biochemistry. 2011; 50:7842–7857. [PubMed: 21848324]

80. Guzman-Aranguez A, Mantelli F, Argueso P. Mucin-type O-glycans in tears of normal subjects and patients with non-Sjogren's dry eye. Invest. Ophthalmol. Vis. Sci. 2009; 50:4581–4587. [PubMed: 19407012]

81. Hirabayashi J, et al. Oligosaccharide specificity of galectins: a search by frontal affinity chromatography. Biochim. Biophys. Acta. 2002; 1572:232–254. [PubMed: 12223272]

82. Wu AM, et al. Fine specificity of domain-I of recombinant tandem-repeat-type galectin-4 from rat gastrointestinal tract (G4-N). Biochem. J. 2002; 367:653–664. [PubMed: 12133001]

83. Krzeminski M, et al. Human galectin-3 (Mac-2 antigen): defining molecular switches of affinity to natural glycoproteins, structural and dynamic aspects of glycan binding by flexible ligand docking and putative regulatory sequences in the proximal promoter region. Biochim. Biophys. Acta. 2011; 1810:150–161. [PubMed: 21070836]

84. Zhu C, et al. The Tim-3 ligand galectin-9 negatively regulates T helper type 1 immunity. Nat. Immunol. 2005; 6:1245–1252. [PubMed: 16286920]

85. Functional_Glycomics. http://www.functionalglycomics.org/static/consortium/consortium.shtml

86. Caberoy NB, et al. Galectin-3 is a new MerTK-specific eat-me signal. J. Cell Physiol. 2012; 227:401–407. [PubMed: 21792939]

87. Cummings RD, Nyame AK. *Schistosome* glysoconjugates. Biochim. Biophys. Acta. 1999; 1455:363–374. [PubMed: 10571025]

88. Carlsson S, et al. Affinity of galectin-8 and its carbohydrate recognition domains for ligands in solution and at the cell surface. Glycobiology. 2007; 17:663–676. [PubMed: 17339281]

89. Nagae M, et al. Structural analysis of the recognition mechanism of poly-N-acetyllactosamine by the human galectin-9 N-terminal carbohydrate recognition domain. Glycobiology. 2009; 19:112–117. [PubMed: 18977853]

90. Krejcirikova V, et al. Structure of the mouse galectin-4 N-terminal carbohydrate-recognition domain reveals the mechanism of oligosaccharide recognition. Acta. Crystallogr. D Biol. Crystallogr. 2011; 67:204–211. [PubMed: 21358051]

91. Tomizawa T, et al. Solution Structure of The C-Terminal Gal-Bind Lectin Domain from Human Galectin-4. Structural Genomics and Proteomics Initiative (RSGI), Riken. 2005

92. Ideo H, et al. Galectin-8-N-domain recognition mechanism for sialylated and sulfated glycans. J. Biol. Chem. 2011; 286:11346–11355. [PubMed: 21288902]

93. Stowell SR, et al. Ligand reduces galectin-1 sensitivity to oxidative inactivation by enhancing dimer formation. J. Biol. Chem. 2009; 284:4989–4999. [PubMed: 19103599]

94. Rapoport EM, et al. Galectin-loaded cells as a platform for the profiling of lectin specificity by fluorescent neoglycoconjugates: a case study on galectins-1 and -3 and the impact of assay setting. Glycobiology. 2008; 18:315–324. [PubMed: 18256179]

95. Vokhmyanina OA, et al. Carbohydrate specificity of chicken and human tandem-repeat-type galectins-8 in composition of cells. Biochemistry. 2011; 76:1185–1192. [PubMed: 22098245]

96. Greenspan NS. Dimensions of antigen recognition and levels of immunological specificity. Adv. Cancer Res. 2001; 80:147–187. [PubMed: 11034543]

97. Rabinovich GA, et al. Functions of cell surface galectin-glycoprotein lattices. Curr. Opin. Struct. Biol. 2007; 17:513–520. [PubMed: 17950594]

98. Wu Y, et al. Transforming binding affinities from three dimensions to two with application to cadherin clustering. Nature. 2011; 475:510–513. [PubMed: 21796210]

99. Vasta GR, et al. Galactosyl-binding lectins from the tunicate *Didemnum candidum*. Purification and physicochemical characterization. J. Biol. Chem. 1986; 261:9174–9181. [PubMed: 3722194]

100. Dam TK, Gerken TA, Brewer CF. Thermodynamics of multivalent carbohydrate-lectin cross-linking interactions: importance of entropy in the bind and jump mechanism. Biochemistry. 2009; 48:3822–3827. [PubMed: 19292456]

101. Ahmad N, et al. Galectin-3 precipitates as a pentamer with synthetic multivalent carbohydrates and forms heterogeneous cross-linked complexes. J. Biol. Chem. 2004; 279:10841–10847. [PubMed: 14672941]

102. Brewer CF, Miceli MC, Baum LG. Clusters, bundles, arrays and lattices: novel mechanisms for lectinsaccharide-mediated cellular interactions. Curr. Opin. Struct. Biol. 2002; 12:616–623. [PubMed: 12464313]

103. Dam TK, Brewer CF. Lectins as pattern recognition molecules: the effects of epitope density in innate immunity. Glycobiology. 2010; 20:270–279. [PubMed: 19939826]

104. Di Lella S, et al. Characterization of the galectin-1 carbohydrate recognition domain in terms of solvent occupancy. J. Phys. Chem. B. 2007; 111:7360–7366. [PubMed: 17523619]

105. Echeverria I, Amzel LM. Disaccharide binding to galectin-1: free energy calculations and molecular recognition mechanism. Biophys. J. 2011; 100:2283–2292. [PubMed: 21539798]
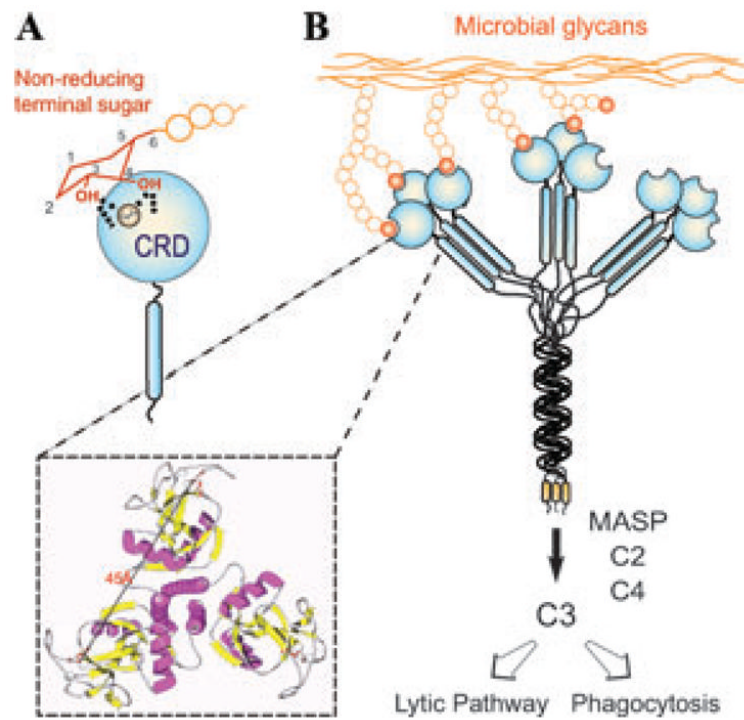
Figure 1.
Recognition and effector activities of the mannose-binding lectin (MBL). (A) The CRD recognizes equatorial hydroxyls on C3 and C4 of nonreducing terminal mannose with participation of the Ca^{2+} atom. (B) The MBL trimer binds to ligands that are displayed about 45 Å apart on the microbial surface, and via association with the MBL-associated serine protease (MASP) may activate the complement cascade, leading to opsonization or lysis of the microbe.
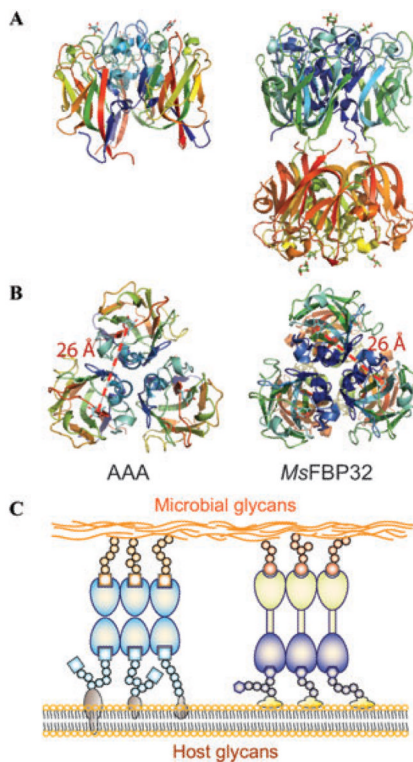
Figure 2.
Structure and binding activities of F-type lectins. (A) Lateral view of the F-type lectins from the European eel(*Anguilla anguilla* agglutinin; AAA, left), a trimer of single-CRD subunits, and the striped bass (*Morone saxatilis* fucose-binding protein; *Ms*FBP32, right) a trimer of binary CRD subunits. (B) Both AAA and *Ms*FBP32-A binds to ligands that are displayed about 25Å apart (measure between fucose's O4) on the microbial or host cell surface. (C) AAA trimers can form oligomers and agglutinates cells; cross-linking of microbes and host phagocytic cells can lead to opsonization. *Ms*FBP32 trimers have N-CRD and C-CRD clusters of distinct specificity. Binary CRD F-lectins have the capacity to crosslink microbes and phagocytes, and facilitate opsonization of bacteria.
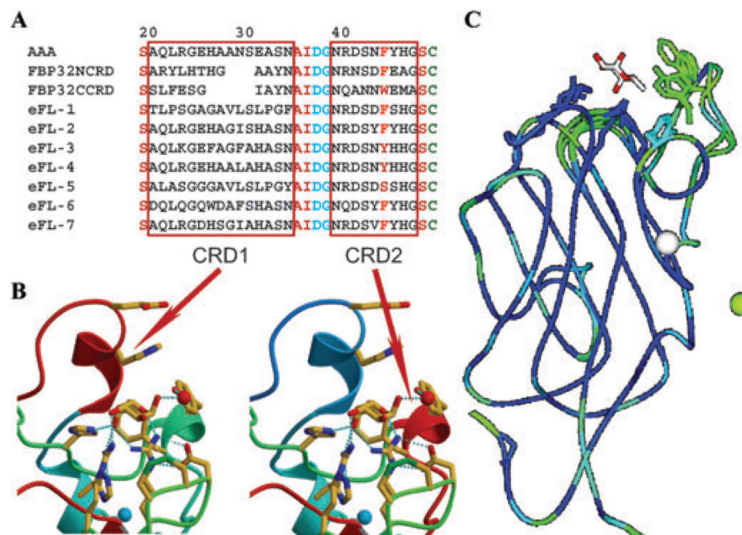
Figure 3.
Sequence variability in the loops surrounding the binding cleft of the eel agglutinin isoforms. (A) Sequence alignment of the AAA with eFL-1-7 isoforms showing variability in the loops CDR1 and CDR2. The amino acid replacements at several positions suggest a broader specificity for some isoforms, particularly eFL-1 and eFL-5. (B) Localization of the replacements and potential interactions are illustrated for CDR1 and CDR2. (C) The regions with highest sequence variability in the loops surrounding the binding cleft are illustrated with lighter blue and green color.
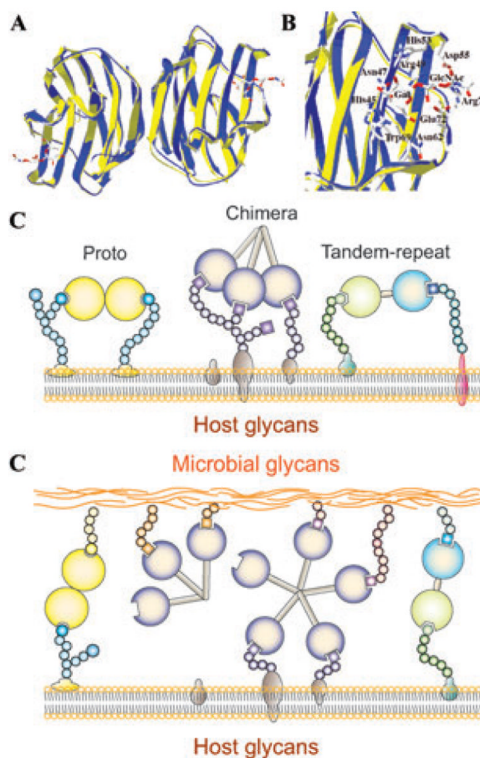
Figure 4.
Structure and binding activities of galectins. (A) Structure of the galectin-1 dimer in complex with LacNAc. (B) Detail of the binding cleft indicating the amino acid residues that interact with the disaccharide. (C) Schematic illustration of *cis*-interactions of proto, chimera, and tandem repeat galectins with host cell surface glycans. (D) Schematic illustration of *trans*-interactions of proto, chimera, and tandem repeat galectins with host cell surface and microbial glycans. Proto and tandem repeat type galectins can cross-link host and microbial glycans. Chimera galectins (galectin-3) can recognize microbial glycans but it is not clear that they can cross-link them to the host cell surface.
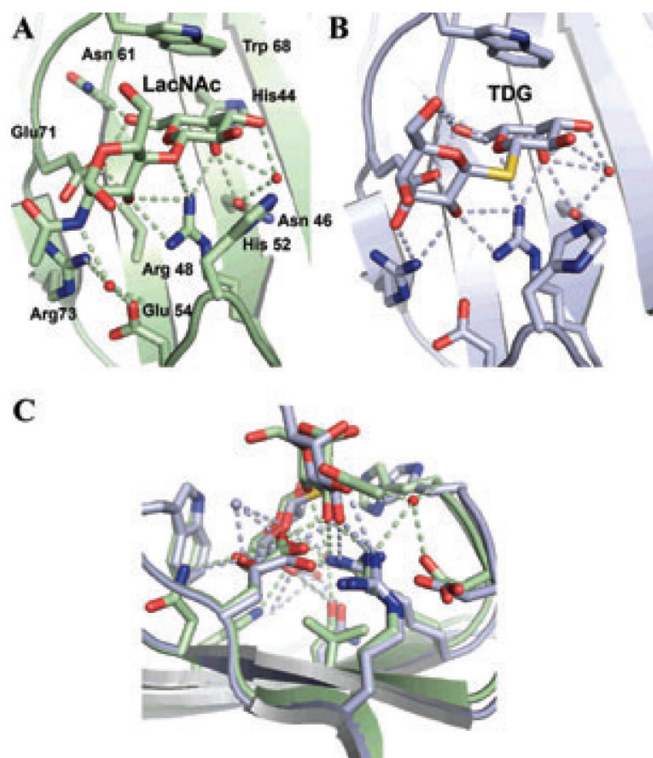

Figure 5.
Recognition of LacNAc and TDG by *Bufo arenarum* galectin-1. (A) Structure of the galectin-1 dimer in complex with LacNAc indicating the amino acid residues that interact with the disaccharides. (B) Structure of the galectin-1 in complex with TDG. (C) Overlap of the binding cleft.