Key regulators of galectin–glycan interactions

Nourine A. Kamili, Emory University
Connie M. Arthur, Emory University
Christian Gerner-Smidt, Emory University
Eden Tafesse, Emory University
Anna Blenda, Emory University
Marcelo Dias-Baruffi, University of Sao Paulo
Sean R. Stowell, Emory University

Journal Title: Proteomics
Volume: Volume 16, Number 24
Publisher: Wiley: 12 months | 2016-12-01, Pages 3111-3125
Type of Work: Article | Post-print: After Peer Review
Publisher DOI: 10.1002/pmic.201600116
Permanent URL: https://pid.emory.edu/ark:/25593/s8k6x

Final published version: http://dx.doi.org/10.1002/pmic.201600116

Copyright information:
© 2016 WILEY-VCH Verlag GmbH & Co.

Accessed November 22, 2019 12:28 AM EST
Key regulators of galectin–glycan interactions

Nourine A. Kamili¹, Connie M. Arthur¹, Christian Gerner-Smidt¹, Eden Tafesse¹, Anna Blenda¹,2, Marcelo Dias-Baruffi³, and Sean R. Stowell¹,3

¹Center for Transfusion Medicine and Cellular Therapies, Department of Laboratory Medicine and Pathology, Emory University School of Medicine, Atlanta, GA, USA
²Department of Biology, Erskine College, Due West, SC, USA
³Department of Clinical Analyses, Toxicology and Food Sciences, Faculty of Pharmaceutical Sciences of Ribeirao Preto, University of Sao Paulo, Ribeirao Preto, Brazil

Abstract

Protein–ligand interactions serve as fundamental regulators of numerous biological processes. Among protein–ligand pairs, glycan binding proteins (GBPs) and the glycans they recognize represent unique and highly complex interactions implicated in a broad range of regulatory activities. With few exceptions, cell surface receptors and secreted proteins are heavily glycosylated. As these glycans often represent highly regulatable post-translational modifications, alterations in glycosylation can fundamentally impact GBP recognition. Among GBPs, galectins in particular appear to engage a diverse set of glycan determinants to impact a broad range of biological processes. In this review, we will explore factors that impact galectin activity, including the effect of glycan modification on galectin–glycan interactions.

Keywords

Frontal Affinity Chromatography; Galectin; Glycan; Glycan Binding Protein; Glycoproteomics; Surface Plasmon Resonance

1 Introduction

Protein–ligand interactions play a critical role in virtually every biological process, from direct recognition of microbes to the regulation of fundamental metabolic pathways.

While many examples of protein–ligand interactions exist, glycan binding proteins (GBPs) and their respective carbohydrate (glycan) ligands have recently emerged as one of the most fundamental, yet unique examples of regulatory networks capable of modulating biological processes [1]. GBPs possess the ability to recognize a broad range of distinct carbohydrate determinants that often completely envelope a cell [2]. As more than half of the molecular weight of many glycoproteins is often glycan in origin and nearly 80% of all human proteins...
posses glycan modifications, glycans provide a very unique and highly regulatable substrate for GBPs to interact with cells [2]. Specific glycan modifications are not directly encoded by the genetic signature responsible for a given glycoprotein, but instead are governed by an entire array of different glycosyltransferase enzymes [2]. This allows cells to rapidly change the types of glycan modifications on a given glycoprotein by simply altering the repertoire, location and activity of the various glycosyltransferases [2]. In this way, glycan modifications provide a plastic substrate that can directly govern cellular sensitivity to GBPs, providing an additional and complex regulatory feature of GBP–glycan interactions that can occur independent of protein receptor expression [1–6].

Among GBP–glycan interactions that appear to influence a wide variety of complex processes, galectins perhaps regulate the broadest range of cellular functions [4–6]. This in part likely reflects the expression of various galectin family members in the vast majority of tissues and the ability of galectins to recognize β-galactosides, some of the most common terminating carbohydrate determinants on mammalian glycans [7]. Despite the ability of galectins to engage β-galactoside-containing glycans, modifications of β-galactosides can specifically impact glycan recognition by individual galectin family members, with distinct consequences on cellular sensitivity to galectin-mediated activities [4–6]. In this review, we will provide an overview of the factors that regulate galectin activity, including the unique relationship between galectin–glycan interactions and galectin function. In doing so, we will not attempt to provide an exhaustive review of the impact of glycan modifications on galectin–glycan interactions. Instead, we hope to provide several key examples that illustrate how specific glycan modifications can differentially impact the ability of galectins to engage various glycan ligands, with unique consequences on cellular sensitivity to galectin activity.

2 Discovery of galectins

In an effort to understand how biology may decode complex cell surface carbohydrate structures into meaningful biological outcomes, several investigators sought to determine whether mammals, like plants, possess GBPs. The first example of a vertebrate GBP was provided by Ashwell and Morell, who isolated a hepatic GBP, the Ashwell–Morell or asialoglycoprotein receptor, previously shown to clear asialoglycoproteins [8,9]. While early studies suggested a role for the Ashwell–Morell receptor in the uptake of desialylated glycoproteins, it was not until nearly 40 years later that the functional consequence of this receptor in platelet clearance became apparent [10]. However, despite the uncertainty of its function for many years, the pioneering work of Ashwell and Morell suggested that like plants, mammals actually possess functional GBPs.

In the wake of Ashwell and Morell's discovery, many researchers sought to determine whether additional vertebrate GBPs exist capable of engaging terminal galactose-containing glycans. In 1975, a soluble β-galactoside binding carbohydrate binding protein termed “electrolectin” was purified from the electric organ of the electric eel [11]. This carbohydrate binding protein is now known to be the first described member of the galectin family. Subsequent studies led by the groups of Barondes and Kornfeld demonstrated that similar β-galactoside binding proteins exist in chicken and bovine tissue, respectively [12, 13]. These results suggested that this protein might be widely expressed in different species.
Indeed, galectins are present in all metazoans, making these GBPs one of the most evolutionarily ancient GBP families expressed in mammals [14]. Fifteen members of the galectin family have been described, with eleven of these found in humans (Fig. 1) [15].

3 Key regulators of galectin activity

While the discovery of galectins resulted from their ability to recognize affinity supports containing β-galactoside-containing ligands, early attempts to purify galectins also isolated another key variable that regulates the carbohydrate binding activity of several galectin family members. Although several groups simultaneously sought to isolate β-galactoside-binding proteins, only Teichberg and colleagues originally used reducing conditions in all isolation buffers [11]. As a result, other attempts initially failed to successfully isolate similar galectin family members, even when using protocols that were otherwise essentially identical [16]. Subsequent studies demonstrated that isolated electrolectin required reducing conditions for sustained activity [17, 18], strongly suggesting that the reducing requirements for galectin activity reflected an intrinsic galectin requirement as opposed to simply facilitating the purification process.

Due to this unique dependence of early galectin family members on reduced thiols for carbohydrate recognition, in 1988, these lactose-binding GBPs were termed “S-type Lectins” as a result of their putative requirement of reduced cysteine residues for glycan engagement [19]. The dependence of these early galectins for reduced thiols appeared to be analogous to C-type GBPs, which require calcium for glycan recognition and suggested a general theme for key additional regulators that impact mammalian GBP function [19]. However, in 1991 Hirabayashi and Kasai substituted serine for cysteine residues and demonstrated that these mutants exhibited virtually no change in the ability of galectin-1 (Gal-1) to bind to asialofetuin-agarose [20]. These results were corroborated by subsequent studies and demonstrated that the cysteine residues of Gal-1 are not necessary for glycan-binding [20–23]. Instead, these studies demonstrated that the reduced thiols of cysteine residues are important for maintaining protein stability and that oxidation at such residues eliminates Gal-1 carbohydrate binding activity [20–23].

While reducing agents can protect Gal-1 from oxidative inactivation, as the extracellular environment is oxidative by nature, the unique sensitivity of Gal-1 to alterations in redox potential likely evolved to specifically regulate Gal-1 function [24, 25]. For example, as ligand engagement partially protects Gal-1 from oxidation, failure to bind ligand following release from a cell may not only facilitate oxidative inactivation, but also provides a unique form of spatial and temporal regulation of Gal-1 activity [24,25]. This may be especially important in the modulation of galectin-mediated leukocyte turnover during tissue injury and inflammation. In this setting, Gal-1 released immediately following tissue injury may quickly saturate available glycan ligands, causing the residual Gal-1 to undergo oxidative inactivation [26]. As Gal-1 can impede chemotaxis and target neutrophils for turnover, this unique form of regulation would be predicted to allow neutrophils to accumulate unrestrained, thereby facilitating pathogen neutralization and the removal of necrotic tissue [27]. However, as neutrophils encroach on and damage viable cells surrounding an area of tissue injury, reduced and therefore active Gal-1 may engage these neutrophils, inhibiting

Proteomics. Author manuscript; available in PMC 2018 March 02.
their chemotaxis and inducing their turnover [28]. As galectins possess the unique ability to induce neutrophil turnover independent of apoptosis, a process called prepareasis, this may allow neutrophils to maintain membrane integrity in the face of inflammation until successfully phagocytosed [29]. Similar redox-based regulatory networks may govern Gal-1 activity in the setting of fetal-maternal tolerance [30, 31]. As a result, while the requirement for reducing conditions became one of the earliest biochemical descriptors of galectins, the unique sensitivity of Gal-1 and other galectins to redox potential suggests a fundamental feature that regulates the activity of several members of the galectin family [32].

While much remains unknown regarding the redox regulation of galectins, several key studies provide biochemical insight into the mechanism whereby oxidation eliminates Gal-1 activity. As Gal-1 possesses six Cys residues, oxidation can result in intramolecular and intermolecular disulfide bond formation, depending on Gal-1 concentration [22]. Intramolecular disulfide bond formation results in profound conformation changes in redox sensitive galectins [33, 34]. In Gal-1 in particular, this appears to be due to disulfide bond formation between Cys2 and Cys130, cysteine residues that normally reside just outside the dimer interface, in addition to other disulfide pairs, including Cys16-Cys88 and Cys42-Cys60 [23, 32, 35]. Disulfide bond formation-induced alterations in the tertiary structure of Gal-1 prevent further dimerization and effectively eliminate carbohydrate recognition [36]. As monomeric mutants display enhanced sensitivity to oxidative inactivation, monomeric intermediates likely facilitate disulfide bond formation through enhanced flexibility of individual Cys residues [23, 27]. Similarly, the ability of glycan ligand to reduce Gal-1 sensitivity to oxidation may reflect an ability to modulate Gal-1 dimerization, as engagement by glycan ligand enhances Gal-1 dimerization, while carbohydrate recognition fails to similarly protect monomeric versions of Gal-1 from oxidative inactivation [17, 27, 37] (Fig. 2). Recent studies suggest that nitrosylation of Cys57 can protect Gal-2 from oxidative inactivation, providing another intriguing example of redox regulation of galectin activity [38, 39].

While the unique sensitivity of Gal-1 to oxidative inactivation can complicate the study of Gal-1 in biological systems [29, 40], the ability of redox potential to regulate carbohydrate independent activities of Gal-1 represents a relatively unexplored, yet critical factor responsible for Gal-1 function [23, 27, 36, 41, 42]. This is especially important when considering that while oxidation appears to impact the carbohydrate binding capacity of Gal-1, subsequent studies suggested that oxidized Gal-1 actually possesses biological activity that occurs through glycan-independent processes [23, 36, 41, 43–46]. Indeed, oxidized Gal-1 appears to regulate neurite outgrowth and may be important in overall neuron regeneration following injury [23, 36, 41, 43–47]. Whether oxidation of other galectins likewise converts these proteins into alternative products with distinct biologically activity remains unknown. These results suggest that galectins in some ways may belong to the morpheeins, a class of proteins that possess drastically different biological activities depending on the conformation state of the protein [48].

While some of the first galectins described appeared to possess a unique requirement for reducing conditions to maintain glycan recognition, subsequent studies demonstrated that other members of the galectin family, in particular galectin-3 (Gal-3), do not display the
same sensitivity to changes in redox environment [49]. As these additional family members retain the ability to bind β-galactoside-containing glycans, despite failing to display a requirement for reduced thiols, “S-type lectin” ineffectively captured all members of this GBP family. Given their common preference for β-galactoside-containing glycans, the carbohydrate ligand preference became their defining feature [50]. Therefore, members of this protein family were re-named galectins due to their common affinity for β-galactosides [50]. Isolation of different galectins not only revealed distinct sensitivities for thiol reduction, but also demonstrated unique tertiary and quaternary configurations. As the first galectin described, galectin-1 became an example of the prototypical galectin subtype, which exists largely as homodimers composed of two identical carbohydrate recognition domains (CRDs). In contrast, galectin-3 (Gal-3), the only member of the chimeric subgroup, exists as a pentamer through self-association mediated by an N-terminal collagenous domain not capable of carbo-hydrate recognition. Finally, several galectin family members belong to the tandem repeat subfamily composed of two distinct CRDs tethered by a linker peptide (Fig. 1) [51]. It should be noted that while galectins are defined by their ability to recognize β-galactoside-containing glycans, many, if not all of these proteins possess important regulatory roles through intracellular interactions that often occur independent of carbohydrate recognition [52]. Although these activities will not be reviewed here, they represent an active area of study that continues to illustrate the diverse range of biological activities attributable to this protein family.

While not all galectins display the same sensitivity to oxidative inactivation, evolution appears to have selected alternative regulatory pathways capable of governing the activity of galectin family members insensitive to redox potential. For example, although galectin-3 (Gal-3) does not undergo oxidative inactivation, proteolytic cleavage of the N terminal collagenous domain required for oligomerization prevents the Gal-3 CRD from existing in a multivalent configuration [53]. As Gal-3 oligomerization is required for most of its carbohydrate-dependent functions, cleavage of the N terminal domain of Gal-3 effectively eliminates many of its biological activities [54, 55]. Similar cleavage of the linker peptide responsible for tethering the distinct domains of tandem-repeat galectins likewise eliminates many of their biological functions [56]. While detailed studies regarding the consequences of prototypical galectin oxidation remain to be tested, oxidative inactivation appears to be most apparent with Gal-1, Gal-2 and Gal-7. As prototypical galectin signaling likely also requires functional bivalency [57], oxidative inactivation may not only prevent carbohydrate recognition, but also inhibit prototypical galectin dimerization. In this way, regulatory features appear to prevent carbohydrate recognition or severely compromise the quaternary organization required for effective galectin function (Fig. 2).

**4 General features of galectin–glycan recognition**

Regardless of the overall quaternary and unique domain organization of galectin subfamily members, the shared affinity of each galectin CRD for β-galactoside-containing glycans ultimately appears to reflect a collection of key conserved amino acids within each CRD [50, 58, 59]. Early studies demonstrated that site directed mutagenesis of several conserved amino acids could eliminate carbohydrate recognition [20]. Crystallographic studies corroborated these findings and demonstrated that distinct hydrogen bonds are responsible...
for galectin-glycan interaction. For example, in Gal-1 His45 and Arg49 interact with the 4-OH of galactose, Arg49, Asn62 and Glu72 interact with the 6-OH of galactose and Arg49 and Glu72 interact with the 3-OH of N-Acetylg glucosamine [60–66] (Fig. 3). Three additional conserved residues (Glu55, Arg74, Asn47) help stabilize this binding, while a conserved tryptophan residue likely stabilizes the hydrophobic face of galactose through van der waals interactions [60–66]. Although the crystal structures of the entire galectin family remain to be solved, key features of these conserved residues likely form similar LacNAc interactions across all known galectins [60–66].

While galectins within and across species clearly share common domains and a low-level binding affinity for β-galactoside-containing glycans such as lactose and LacNAc, the unique and overlapping preference of individual galectin family members for various presentations of these simple sugars likely underlies the distinct and complementary activities of each family member. Discrete residues that reside in close approximation to amino acids required for LacNAc recognition likely dictate the impact of β-galactoside modifications on glycan recognition [67]. This allows variation in β-galactoside modifications to differentially regulate cellular recognition by individual galectins [2, 68]. In this way, relatively simple glycan modifications largely driven by epigenetic alterations in glycosyltransferase expression, localization and activity can directly impact the sensitivity of cells to the signaling effects of different galectins [1, 2, 4–6].

While glycan variation can impact galectin engagement and therefore signaling function, the specific impact of glycan modification on galectin–glycan interactions remained elusive for many years. This was due in large part to the technical difficulty of examining GBP–glycan interactions. Unlike protein-protein interactions, where site-directed mutagenesis can provide an effective tool to examine the critical residues required for ligand engagement, as complex glycan synthesis represents the coordinated effort of many different enzymes, carbohydrates cannot be readily “cloned” in the same fashion as their protein counterparts [69]. As a result, efforts to directly introduce very specific alterations in glycan structure to determine the key requirements for galectin binding often requires complex chemoenzymatic processes [70]. Thus, early studies seeking to examine the specificity of galectins for glycan ligands often utilized relatively simple carbohydrate structures to ascertain general galectin–glycan preferences [11–13]. However, the increased availability of more complex glycan structures, provided through highly diverse and well-characterized glycan libraries, has completely changed our understanding of key features of galectin–glycan interactions [71–73]. As a result of data generated using these glycan libraries, in addition to unique glycan binding preference observed using different assay platforms, distinct patterns of galectin interactions with glycan ligands have emerged, continuing to provide important insight into galectin function.

5 Polylactosamine: A common ligand for galectins

Some of the earliest studies seeking to identify physiological ligands for galectins utilized metabolically labeled glycans released from cells to assess which glycans are preferentially engaged by Gal-1 [74, 75]. By employing a variety of glycosidases to elucidate the general composition and motifs of bound glycans, these studies elegantly demonstrated that Gal-1
exhibits a significant preference for polylactosamine (polyLacNAc) containing glycans [74,75]. Subsequent studies using solid-phase assay systems with defined glycan libraries, such as glycan microarrays, produced similar results and suggested that additional galectin family members, including Gal-2, Gal-3, Gal-7, Gal-8 and Gal-9 likewise prefer polyLacNAc glycans [67, 71, 76–81]. However, other studies failed to corroborate some of these findings. For example, key results obtained using frontal affinity chromatography (FAC), surface plasmon resonance (SPR) and isothermal calorimetry (ITC) demonstrated that while some galectins, such as Gal-3, Gal-7, Gal-8 and Gal-9, exhibit enhanced binding to polyLacNAc [71, 80, 81], neither Gal-1 nor Gal-2 display a similar polyLacNAc preference [71, 76, 79–82].

While individual studies examining galectin-polyLacNAc interactions appear to be conflicting, the binding outcomes in each assay system may actually reflect unique modes of polyLacNAc recognition displayed by individual galectins (Fig. 4). For example, in solution-based assays, where the terminal galactose of LacNAc and polyLacNAc is presented equivalently, Gal-1 and Gal-2 likewise exhibit similar recognition [71, 80]. These results suggest that Gal-1 and Gal-2 specifically engage only the terminal LacNAc motif irrespective of polyLacNAc length. Consistent with this, removal of the terminal β-galactose of polyLacNAc completely prevents Gal-1 and Gal-2 binding regardless of the assay system [76, 77, 80]. However, despite only recognizing the terminal LacNAc motif, Gal-1 and Gal-2 exhibit a clear preference for polyLacNAc following glycan immobilization [76]. As Gal-1 and Gal-2 exist as relatively rigid dimers with CRDs facing opposite directions [60–63], the preference of Gal-1 and Gal-2 for polyLacNAc in solid-phase systems might reflect preferential binding and optimal crosslinking of the terminal LacNAc motifs of polyLacNAc, a process that only becomes apparent following glycan immobilization (Fig. 4). In contrast, as Gal-3 does not exist as a rigid dimer, but instead oligomerizes through a flexible linker [83], Gal-3 appears to have evolved a different mechanism for polyLacNAc binding. Unlike Gal-1 and Gal-2, Gal-3 recognition of polyLacNAc remains unaffected following removal of the terminal β-galactose [76, 80], strongly suggesting that Gal-3 intrinsically recognizes polyLacNAc through internal LacNAc engagement. Consistent with this, Gal-3 displays a distinct preference for polyLacNAc glycans in either solution-based or solid-phase assays [71, 76, 80]. The distinct modes of polyLacNAc recognition exhibited by Gal-1, Gal-2 and Gal-3 extend to other members of the galectin family, including Gal-7 and Gal-8 [67, 80], suggesting that unique interactions with polyLacNAc glycans may likewise impact the relative preference for polyLacNAc by additional galetins in various assay systems.

Unique modes of polyLacNAc engagement may not only be relevant when interpreting galectin–glycan interactions in disparate assays, but might also allow distinct polyLacNAc modifications to differentially impact galectin–glycan interactions. The most classic and well-studied polyLacNAc modification capable of differentially impacting galectin binding is sialylation, which typically occurs as a α2-3 or α2-6 terminal β-galactoside modification. While the addition of α2-6 sialic acid often inhibits LacNAc recognition by all galectins [71, 80, 82], this same modification on polyLacNAc fails to impact Gal-3 recognition, likely due to the ability of Gal-3 to primarily recognize internal LacNAc motifs within polyLacNAc [76]. Given the preferential engagement of the terminal LacNAc motif by Gal-1 and Gal-2,
α2-6 sialylation completely prevents polyLacNAc recognition by these two galectins [29, 71, 77, 79, 80, 84] (Fig. 5). As α2-3 sialylation inhibits Gal-2 LacNAc recognition, but fails to affect Gal-1 binding, α2-3 sialylation of polyLacNAc also inhibits Gal-2 binding, while failing to significantly impact engagement by Gal-1 [76]. In this way, α2-3 versus α2-6 sialylation can differentially impact the binding of Gal-1, Gal-2 and Gal-3 toward polyLacNAc glycans. However, in contrast to Gal-1, Gal-2 and Gal-3, sialylation and even sulfation actually appears to positively influence glycan recognition by other members of the galectin family. For example, sulfation of LacNAc can substantially enhance binding by Gal-4, and both sulfation and sialylation can similarly augment glycan recognition by Gal-8 [66, 67, 73, 85]. As Gal-4 and Gal-8 represent tandem repeat galectins, occasionally the distinct domains of these galectins can exhibit preferential engagement of different forms of LacNAc modifications, as shown for sulfated or sialylated glycans by Gal-8 [66, 67, 73, 86].

In contrast to terminal modifications of polyLacNAc, such as sialylation, internal modification of polyLacNAc appears to produce the opposite effect on galectin recognition. Given the preference of Gal-3 for internal LacNAc motifs within polyLacNAc, addition of α1-3 fucose to internal LacNAc motifs, which uniformly inhibits galectin recognition of LacNAc, prevents Gal-3 polyLacNAc recognition. In contrast, this same modification fails to prevent polyLacNAc engagement by Gal-1 and Gal-2 [67, 76]. Thus, distinct modes of polyLacNAc binding by individual members of the galectin family allow simple glycan modifications to differentially impact glycan engagement (Fig. 6).

The impact of glycan modification is not limited to biochemical analysis of galectins with purified glycans, but also results in the unique regulation of cellular sensitivity to galectin-induced signaling. Alterations in cell surface sialylation specifically regulate the ability of Gal-1 to recognize T cells, which in turn influences the response of distinct T cell populations to Gal-1. Seminal studies demonstrated that following differentiation, CD4 T helper (Th) cells possess unique glycan signatures. CD4 Th2 T cells cap glycans with α2-6 sialic acid, while CD4 Th1 and Th17 T cells instead modify cell surface glycans with α2-3 sialic acid [87]. As Gal-1 can recognize α2-3 sialylated glycans, the ability of Gal-1 to readily engage Th1 and Th17 cells renders these CD4 T cell populations sensitive to Gal-1-induced signaling. In contrast, Gal-1 fails to similarly bind or alter the activity of α2-6 sialylated Th2 cells [87]. Unlike the differential sensitivity displayed by CD4 T cells to Gal-1, distinct CD4 T cell populations remain equally sensitive to Gal-3, strongly suggesting that unlike Gal-1, alterations in T cell surface sialylation fail to impact Gal-3 binding or signaling [87]. Unique modes of polyLacNAc recognition displayed by Gal-1 and Gal-3 may therefore differentially impact the sensitivity of distinct CD4 T cell populations to individual galectin family members [87].

The unique influence of sialylation on galectin binding and signaling does not appear to be limited to CD4 T cells. Gal-1, Gal-2, Gal-3 and Gal-8 also display preferential binding to polyLacNAc glycans on myeloid cells, while retaining differential sensitivity to cell surface sialylation [76]. Similar to CD4 T cells, α2-6 sialylation also inhibits binding by Gal-1 to myeloid cells, while α2-3 sialylation fails to significantly impact Gal-1 engagement. Addition of sialic acid with either linkage inhibits binding by Gal-2, while neither modification appears to readily influence Gal-3 recognition [76]. Similar examples of the differential impact of sialic acid modifications can be observed for Gal-1, Gal-2, Gal-3 and
Gal-8 binding to Chinese hamster ovary (CHO) cells [67, 76, 78, 88]. The distinct impact of sialylation appears to reflect terminal versus internal LacNAc recognition of polyLacNAc-containing glycans, as removal of the terminal β-galactose severely inhibits Gal-1 and Gal-2 recognition, while failing to similarly impact Gal-3 [76]. These binding preferences do not appear to reflect non-specific engagement of the cell surface glycolcalyx, as alterations in galectin binding directly correlate with changes in cellular outcomes, including the preferential internalization of cell-bound galectin [67, 76, 78]. It should be noted that while polyLacNAc may serve as a common galectin ligand, some cell surface glycoprotein substrates might not possess polyLacNAc structures. Thus, on the surface of some cells, α2-6 sialylation can impact cellular engagement by Gal-3, where this modification can protect neoplastic cells from Gal-3-induced apoptosis [89, 90]. As a result, differences in glycan modification and the types of LacNAc-containing glycans available to galectins influence the sensitivity of a given cell to different members of the galectin family.

6 Impact of core glycan presentation on galectin–glycan interactions

While polyLacNAc represents a common ligand for many members of the galectin family, these structures typically do not exist as isolated glycans, but instead represent specific modifications on core glycan structures [2]. Common core structures on glycoproteins can be elongated to possess polyLacNAc glycans, which in turn can be further altered to possess additional modifications. Most glycoprotein modifications reside on core N-glycans, which are initiated in the ER as an attachment of a presynthesized core structure to asparagine or O-glycans, which are synthesized de novo onto serine/threonine residues in the Golgi apparatus [2]. While galectins can occasionally recognize key features of core glycans, the terminal modifications of these glycans most commonly represent the dominant glycan motif that supports galectin–glycan interactions [67, 71, 91]. As similar modifications can terminate N- or O-glycans, many studies have evaluated whether core N- or O-glycans bear key gly-can ligands responsible for conveying galectin-induced signaling in a given cell. This has been largely accomplished by pharmacological or genetic engineering that directly impact the production of particular core glycans, followed by determining the impact on galectin recognition and consequent alterations in cellular behavior. Using this general approach, T cells deficient in particular core O-glycan structures often display a key requirement for complex O-glycans for galectin-induced signaling [92, 93]. This signaling can occur through engagement of a variety of different glycoproteins and can result from binding to either core 1 or core 2 O-glycans, depending on the cell type and particular glycoprotein receptor involved [93, 94]. Pharmacological inhibition of core glycans likewise suggested that O-glycans serve as the dominant carriers of galectin ligands on T cells [95], although in some T cell populations, including Jurkat T cells, which are deficient in complex O-glycans due to a mutation in Cosmc, N-glycans can also mediate T cell responses [21, 96].

In contrast to a variety of T cell lines, N-glycans expressing polyLacNAc ligands on the surface of myeloid cells mediate galectin recognition and subsequent signaling [97]. Similar N-glycan engagement by Gal-3 on the surface of T cells may negatively regulate T cell signaling through engagement of polyLacNAc glycans mediated by the extension of tetra-antennary N-glycans generated through the activity of MgatV [98]. Loss of MgatV prevents...
Gal-3 from removing key constituents from the immunological synapse, resulting in a heightened T cell response that increases the probability of autoimmunity. In other cells, such as CHO cells, Gal-1 and Gal-3 likewise recognize complex N-glycans [88]. These studies illustrate that many different core glycans express galectin ligands, allowing for a variety of cell surface glycoproteins to become potential ligands for different members of the galectin family.

As Gal-1 and Gal-2 in particular prefer extended carbohydrate structures that terminate in LacNAc largely irrespective of the type of core structure, elongated N- or O-glycans may provide the same degree of LacNAc flexibility needed to optimally bind LacNAc, as observed for poly-LacNAc glycans [88, 91]. Similar to polyLacNAc, the preference for terminal LacNAc bearing N-glycans can only be observed for some galectins in surface bound assays, as solution-based approaches fail to demonstrate a similar preference [79]. However, in addition to presenting polyLacNAc or simple LacNAc glycans in a favorable manner following glycan immobilization, core glycan structures can also present branched glycan ligands, each of which terminate in LacNAc [2]. Presentation of multiple terminating Lac-NAc motifs on branched structures enriches galectin ligand density, increasing the probability of effective LacNAc engagement and therefore enhancing the overall strength of galectin binding [99]. In this scenario, although the binding strength between individual CRDs and LacNAc may remain unaltered [100], the combined binding of multiple CRDs increases the overall avidity of this interaction [101]. Thus, while the microscopic $K_a$ between galectin CRDs and Lac-NAc may remain intrinsically unchanged, the lattice structure generated between homodimeric or oligomeric galectins and branched glycan structures not only stabilizes these interactions, but also mediates the complex networks of cell surface glycoproteins needed to propagate distinct signaling events [99, 102]. This likely allows galectins to induce alterations in the half-life, localization and potential lattice formation of glycan bearing receptors [99, 103–106]. Galectin interactions with branched and other cell surface glycans likely mediate many of the immunoregulatory and other activities of galectins.

Although branched structures appear to serve as key ligands for galectins in a variety of settings, attempts to determine the effect of branching on binding affinity have at times produced conflicting results. This may in part reflect similar differences in results obtained following examination of galectin binding toward polyLacNAc, where the assay system can provide different results regarding the impact of LacNAc presentation on galectin recognition, yet, in so doing, offer important information regarding the potential mode of galectin-induced lattice formation. For example, screening of galectins by FAC indicated a higher affinity for most human galectins with increased branching [71]. This increased affinity is especially pronounced for Gal-7 and Gal-9 when comparing interactions with biennary to tetraennary structures. A recent microarray study confirmed this preference for branched structures by Gal-1 and Gal-3 [76]. In contrast to these solid-state assays, ITC studies on Gal-1 and Gal-3 show that neither has an increased affinity for branched structures over non-branched structures [81]. As the impact of simultaneous tethering of branched structures by galectins in the setting of FAC or microarrays may result in higher avidity interactions not realized in assays where glycan and galectin are free in solution, these differences may reflect preferences largely dictated by alterations in avidity primarily
appreciated when branched glycans are bound to a solid-state surface. Importantly, as surface bound glycans reflect the type of orientation that galectins evolved to recognize when signaling mammalian cells, this orientation likely dictated an evolutionary selection for specific lattice orienting interactions that facilitate galectin binding and signaling [98, 99, 103–106]. Thus, while results obtained using different assay systems may occasionally appear conflicting, in many ways these studies provide important insight into the modes and mechanisms whereby galectins preferentially engage different glycan ligands with implications in the overall function of galectins in the regulation of cellular processes.

7 Blood groups, glycolipids and other galectin ligands

In addition to sialylation, LacNAc or terminal β-galactoside-containing glycans can be modified in a variety of ways that can likewise differentially impact recognition by members of the galectin family. A classic example of this can be found in early studies suggesting that Gal-1 can bind the ganglio-side GM1, where engagement of this glycolipid actually appears to induce cellular signaling [107]. FAC and microar-ray studies demonstrate that other galectin family members can recognize a variety of glycolipids, with Gal-8 in particular displaying high affinity for GM3 and GD1a [67, 71, 108]. While these later studies further illustrate key differences in the binding profile exhibited by individual galectins, the biological consequences of these later interactions remain less well understood. In addition to glycolipids, galectins similarly display differential recognition of other modifications to the core LacNAc structure, including the T antigen (Galβ1-3GalNAc), LacDiNAc (GalNAcβ1-4GlcNAc) and type 1 (Galβ1-3GlcNAc) versus type II (Galβ1-4GlcNAc) LacNAc glycan structures [80, 109, 110]. Several studies suggest that interactions with T antigen in particular may be critical in the ability of some galectins to regulate tumor metastasis, which has resulted in the development of inhibitors designed to block galectin-mediated cancer progression [111, 112].

Perhaps the most well-known β-galactose modifications are ABO(H) blood group antigens. The addition of α1-2 fucose to LacNAc generates the type II H antigen and resides on the surface of red blood cells (RBCs) of blood group O individuals. This H antigen can be additionally modified by attachment of a α 1–3 N-acetylgalactosamine (GalNAc) or α 1–3 galactose to the terminal galactose that results in the formation of blood group A or blood group B, respectively [113]. Unlike alloimmunization that occurs following transfusion or transplantation [114, 115], the formation of anti-A and anti-B antibodies occurs spontaneously early in life and represents the first recognized immunological barrier to tissue transfer between individuals [113]. Early studies suggested a preference of some galectins for ABO(H) blood group antigens and recent glycan microarray studies largely corroborated these results [71, 116], with Gal-2, Gal-3, Gal-4 and Gal-8 displaying high binding to ABO(H) blood groups [67, 76]. Conversely, blood group A and B modifications of LacNAc actually inhibit Gal-1 interactions [76] (Fig. 6). Preference for blood group antigens does not appear to be an artifact of the glycan microarrays, as Gal-1, Gal-2, Gal-3 and Gal-8 displayed a similar preference of blood group A and B bearing red blood cells. However, as ABO(H) antigens are polymorphic in nature, these carbohydrate antigens do not likely represent functional mammalian receptors for galectin family members.

Proteomics. Author manuscript; available in PMC 2018 March 02.
Although several studies suggest a potential impact of polymorphic ABO(H) antigen expression in the human population [117–119], the ability of several galectin family members to recognize these glycans may actually reflect an evolutionary ancient role of this protein family. As microbes exist that decorate themselves in blood group antigens and blood group positive individuals cannot make anti-blood group antibodies, galectins appear to fill an important gap in adaptive immunity by selectively targeting microbes that utilize blood group molecular mimicry [15, 120, 121]. Interestingly, while galectins recognize the same antigen on RBCs, galectins failed to directly induce loss of membrane integrity following RBC engagement [120]. This differential outcome observed following galectin binding of the same glycan ligand not only represents a fundamental aspect of host protection against molecular mimicry, but also demonstrates that GBP engagement of the exact same glycan can result in a completely different outcome depending on the cell. As a result, unlike many different protein–ligand interactions where discrete receptor–ligand pairs induce predictable signaling pathways, the impact of galectin–glycan interactions can be directly influenced by the type of protein or lipid expressing a given glycan ligand on the cell surface.

In addition to providing innate immunity against molecular mimicry [122], many key studies demonstrated that galectins recognize a diverse range of pathogens with varied outcomes [123–126]. For example, Gal-3 is antimicrobial toward *C. albicans*, can engage *Neisseria meningitidis*, *Helicobacter pylori* and *Pseudomonas aeruginosa* and can opsonize *T. cruzi* [124, 127–130]. Gal-1 can also bind various viral pathogens, where it inhibits Influenza or Nipah viral entry or actually enhances HIV infection [131–133]. While several of these interactions likely reflect engagement of self-like structures, it clearly remains possible that galectins recognize many microbes through the binding of unrelated glycan antigens [134]. As galectins are present in all metazoans and represent one of the oldest animal GBP families described, some of their earliest biological activity likely reflected recognition and removal of potential pathogens [135]. Understanding these binding preferences may not only provide important insight into the biochemical nature of galectin–glycan interactions, but will also likely facilitate new appreciation of the breadth and consequences of galectin engagement of different microbes.

**8 Conclusions**

Galectins are unique proteins with the ability to recognize and decode a diverse array of glycan motifs. As the presentation, modification and overall context of glycans can influence galectin binding, these interactions represent highly regulatable modifications with significant implications on a variety of cellular processes [1, 4–6, 136]. Recent studies using fluorescence polarization have sought to take advantage of the distinct binding properties of individual galectins by designing small molecule inhibitors capable of docking adducts into unique sites that surround the CRD of each galectin [137, 138]. These inhibitors will not only provide a unique opportunity to further understand the function of individual members of the galectin family, but also enable specific targeting of individual or multiple galectins involved in a wide variety of diseases. As these tools and additional techniques become available, including ever-expanding glycan libraries, the impact of glycan modification on the sensitivity of various cells to galectin-mediated effects will become increasingly clear.
As galectins regulate a broad range of biological activities, including cells involved in the adaptive immune responses germane to autoimmunity, infectious disease, cancer, transplantation and even transfusion [4, 5, 139–142], understanding the impact of glycan modifications on the binding affinity and overall biological activities of different galectin family members will provide important insight into their function with broad implications in a variety of biomedical disciplines.

References


Abbreviations

CRD carbohydrate recognition domain

FAC frontal affinity chromatography

Gal galectin

GBP glycan binding proteins
Figure 1.
Galectin family members. Galectins can be divided into three subfamilies, based on the organization of the carbohydrate recognition domain: prototypical (galectin-1, galectin-2, galectin-7 and galectin-10), chimeric (galectin-3) and tandem repeat (galectin-4, galectin-8, galectin-9 and galectin 12). Each of these subfamilies also exhibit unique quaternary structures, with prototypical and at least some tandem repeat galectins forming dimers and chimeric galectin-3 forming higher order oligomers.
Figure 2.
Regulation of galectin activity. Galectin-1 (Gal-1) can undergo oxidative inactivation through the formation of intramolecular or intermolecular disulfide bond formation, depending on Gal-1 concentration. In contrast, galectin-3 (Gal-3), which is resistant to oxidative inactivation, can be inactivated by proteolytic cleavage of the N-terminal domain required for Gal-3 oligomerization. Similar cleavage of the intervening linker between the carbohydrate recognition domains of tandem-repeat galectins, as shown for galectin-8 (Gal-8), can also eliminate their biological activity.
Figure 3.
Structure of galectin-1 carbohydrate recognition domain. Prototypical galectins exist as homodimers with carbohydrate recognition domains that face in opposite directions. Structural studies elucidated the carbohydrate binding pocket of the galectin family, highlighting conserved amino acids responsible for the $\beta$-galactoside specificity of galectins. Lactosamine modification can differential impact core carbohydrate recognition domain interactions with $\beta$-galactoside ligands. Here, galectin-1 is shown associated with its inhibitor TDG. The crystal structure data was obtained from open source NCI database (PDBID 3OYW) and reconstructed using SwissPBD Viewer. Key residues involved in glycan recognition for galectin-1, galectin-3 and galectin-7 are highlighted in red.
Quaternary structure can influence galectin–glycan interactions. While the majority of human galectins exhibit a preference for polylactosamine (polyLacNAc) glycans in solid-phase assays, this predilection is not observed for some galectins in solution-based approaches. This may in part reflect different modes of polyLacNAc recognition coupled with unique aspects of the quaternary structure of individual galectin family members. Galectin-1 (Gal-1) and other prototypical galectins primarily exist as rigid dimers with carbohydrate recognition domains facing opposite directions. Several prototypical galectins also appear to preferentially engage the terminal lactosamine of polyLacNAc. As a result, co-engagement of lactosamine-containing ligands is likely facilitated by the conformational flexibility afforded by extended polyLacNAc structures. In contrast, in solution-based assays, crosslinking interactions appreciated on solid supports are no longer apparent. Unlike Gal-1, galectin-3 (Gal-3) and other tandem repeat galectins oligomerize through flexible linkers and therefore do not exhibit the same rigid carbohydrate recognition domain orientation. These galectins appear to have instead evolved the ability to primarily interact with polyLacNAc through internal lactosamine motifs, allowing the same polyLacNAc preference to be observed in solution and solid-phase-based assays. These differences in polyLacNAc engagement allow simple polyLacNAc modifications to differentially impact
galectin–glycan interactions with significant consequences on cellular sensitivity to galectin-induced signaling.
Individual galectin family members differentially recognize polylactosamine (polyLacNAc). While each galectin family member displays a common ability to interact with polyLacNAc, the mode of interaction between different family members can fundamentally differ. For example, galectin-1 (Gal-1) preferentially engages the terminal lactosamine (LacNAc) unit of polyLacNAc glycans. Consequently, terminal modifications, such as α2,6-sialylation, are more likely to impact recognition by Gal-1. In contrast, galectin-3 (Gal-3) preferentially engages internal LacNAc units within polyLacNAc and therefore may not be impacted by terminal modifications, but may instead be influenced by internal polyLacNAc modifications.
Figure 6.
Lactosamine modification differentially impacts galectin binding. While galectin proteins are defined by their common affinity for β-galactoside, each galectin varies in its recognition of modified galactose residues. For galectin-1 (Gal-1), galectin-2 (Gal-2) and galectin-3 (Gal-3), the affinity for modified lactosamine (LacNAc) motifs compared to LacNAc alone is shown on the left panel. Similarly, binding preferences toward different polylactosamine (polyLacNAc) modifications is compared to polyLacNAc in the right panel. The black Ys indicate positive binding between each respective galectin and the glycan structure, while the blue Ys indicate increased binding compared to baseline binding (LacNAc for the left panel and polyLacNAc for the right panel). Gal-1, Gal-2 and Gal-3 all display higher binding toward polyLacNAc than LacNAc in solid-phase assays. The red Ns indicate no binding toward the respective glycan.